

ИЗДАТЕЛЬСТВО
«МИР»

**THE MATHEMATICAL APPROACH
TO BIOLOGY AND MEDICINE**

NORMAN T. J. BAILEY

Professor of Biomathematics

Cornell University Graduate School of Medical Sciences

and

Member of the Sloan-Kettering Institute for Cancer Research

**JOHN WILEY AND SONS
LONDON-NEW YORK-SYDNEY**

1967

Н. БЕЙЛИ

МАТЕМАТИКА В БИОЛОГИИ И МЕДИЦИНЕ

Перевод с английского

Е. Г. КОВАЛЕНКО

Предисловие

Э. Л. НАШЕЛЬБАУМА

ИЗДАТЕЛЬСТВО «МИР» Москва 1970

Нашим читателям хорошо знакомо имя Н. Бейли, книга которого «Статистические методы в биологии», вышедшая в русском переводе двумя изданиями (1962 и 1964 гг.), пользуется заслуженным успехом. Новая книга Бейли состоит из двух частей. Первая часть посвящена методологии применения различных разделов математики в экспериментальных исследованиях, а также вопросу о преподавании математики биологам и об организации исследований по математической биологии. Во второй части проводится математический анализ конкретных биологических и медицинских проблем.

Предназначена для биологов и врачей различных специальностей, а также для математиков и инженеров, занимающихся моделированием сложных систем и применением методов исследования операций.

Редакция биологической литературы

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Имя Н. Бейли, руководителя математического отдела Всемирной организации здравоохранения, уже известно советскому читателю по его книге «Статистические методы в биологии», опубликованной в русском переводе¹. И хотя этому автору принадлежит большое число других работ и, в частности, книг по теории эпидемий и теории генетического сцепления, исследования по применению математических методов в биологии наиболее характерны для его научных интересов. Проф. Бейли в первую очередь статистик, и именно глазами статистика он смотрит на биологический материал. Хорошо известно, что сильная вариабельность биологических экспериментальных данных и вероятностный характер многих биологических явлений предоставляют большие возможности для статистической постановки биологических исследований. Этим и объясняется широта научных интересов проф. Бейли, нашедшая отражение в его новой книге, предлагаемой вниманию советского читателя.

Цель этой книги состоит в том, чтобы обрисовать место и роль математических методов в современной биологии и медицине и содействовать установлению плодотворного сотрудничества между биологами, физиологами и медиками, с одной стороны, и математиками — с другой. Книга состоит из двух частей. В первой, общей части речь идет о методологии использования математики в биологии и медицине. Здесь автор говорит о тех выгодах, которые сулит количественное описание явлений, о роли различных вероятностных методов, о месте математического моделирования в общем процессе научного познания, а также о том, как повлияло на научные исследования появление быстродействующих вычислительных машин. В этой же части затрагиваются некоторые интересные и важные вопросы организационного порядка, касающиеся сотрудничества математиков и биологов, уровня математической подготовки, необходимой для биологов, и преподавания соответствующих предметов.

Материал первой части книги служит базой для второй части, более конкретной, но и более фрагментарной. Эта часть состоит

¹ Н. Бейли, Статистические методы в биологии, ИЛ, М., 1962 и 1964 гг.

из семи, по сути дела самостоятельных, глав, в которых приводятся конкретные примеры применения математического подхода к биологическим исследованиям. Выбор тем для этих глав диктовался личными пристрастиями автора и еще раз свидетельствует о широте его интересов. Здесь рассматриваются методы классификации в применении к задачам биологической систематики и медицинской диагностики, модели генетического сцепления, распространения эпидемий и роста численности популяций. Последняя глава посвящена использованию методов исследования операций в организационных вопросах, связанных с медицинским обслуживанием. Эта тема в книгах по математической биологии и медицине обычно не затрагивается, а между тем она представляет большой интерес для читателя.

В общем и целом автор дает довольно широкую картину возможностей математики в биологии и медицине. Но при этом он остается в плену своих чисто статистических представлений. Именно поэтому в книге не нашлось места для математических моделей тех биологических и физиологических явлений, в которых вероятностные аспекты играют подчиненную роль и которые связаны с аппаратом теории управления или эвристического программирования. В первой части книги увлечение автора общеметодологическими рассуждениями иногда уводит его слишком далеко: его рассуждения носят подчас слишком общий характер с точки зрения конкретной задачи книги и в то же время недостаточно полновесны, если относить их к серьезной философии науки и математики. В то же время во второй части автор излагает материал, пожалуй, слишком конкретно, если учесть, что все главы этой части носят обзорный характер. Ведь специалист все равно предпочтет ознакомиться с деталями того или иного исследования по оригинальным статьям, а неспециалисту приведенные в книге тонкости вряд ли будут понятны. Тем не менее книга представляет несомненный интерес. Она охватывает разнообразные области знания, в которых математический подход становится существенно необходимым элементом биологического исследования и многие из которых недостаточно широко известны. Книга не требует от читателя солидной математической подготовки, и многие ее главы читаются довольно легко. Можно надеяться, что она будет способствовать дальнейшему улучшению взаимопонимания между математиками и биологами.

Гл. 1—6 отредактированы Э. Л. Наппельбаумом, гл. 7 — проф. Е. С. Смирновым, гл. 8 — д-ром биол. наук Г. А. Викторовым, гл. 9, 11 и 12 — канд. мед. наук Л. Б. Хейфецом.

Э. Наппельбаум

ИЗ ПРЕДИСЛОВИЯ АВТОРА

Важная роль математики в такой точной науке, как физика, общепризнана, однако ценность и целесообразность применения математических методов в «менее строгих» науках — биологии и медицине — нередко ставится под сомнение. В большинстве случаев биологический и медицинский материал крайне изменчив, подвержен влиянию многочисленных сложных факторов, взаимодействующих между собой, и для его описания требуется огромное количество разнообразных данных; на этом основании обычно считают точный математический анализ в этой области нецелесообразным или невозможным.

В ответ на эти возражения нужно сказать следующее. В столь сложных и запутанных ситуациях истинный научный прогресс будет проходить медленно и, так сказать, на ощупь, если для описания научных фактов будут использоваться преимущественно интуитивные понятия, выраженные чисто словесно. Чтобы по-настоящему проникнуть в исследуемые процессы или явления и управлять ими, необходимо найти соответствующий математический аппарат, который мог бы обеспечить более точный и логически строгий метод анализа. Такой аппарат в настоящее время существует. Биологическую изменчивость можно достаточно адекватно описать соответствующим распределением вероятностей. Правильно выбрав план эксперимента, можно выявить влияние каждого из многочисленных факторов в отдельности. Альтернативные гипотезы можно сравнивать на количественной основе, используя соответствующие критерии статистической значимости. На электронных вычислительных машинах можно очень оперативно обработать большое количество данных, а также быстро и точно выполнить все необходимые вычисления.

Хотя математика уже давно использовалась при исследовании отдельных вопросов, относящихся к различным разделам биологии и медицины, лишь в настоящее время стал возможным интегрированный математический подход ко всей этой области знания. Все разделы прикладной математики и математической статистики, а также вычислительные методы, связанные с биологией и медициной, принято рассматривать под общим названием *мате-*

математическая биология. Общая цель математической биологии — сделать для биологии и медицины то, что математическая физика сделала для физики.

В этой книге я стремился изложить математический подход к биологии и медицине таким образом, чтобы как можно яснее донести до читателя его значение и смысл. В первой части книги рассматриваются общие принципы и основные положения математической биологии. Здесь проводится относительно неспециальное обсуждение областей применения математики и статистики, научных методов, методов исследования операций, электронных вычислительных машин, а также организации биоматематических исследований. Во второй части книги рассматривается ряд конкретных приложений математических методов к таким областям биологии и медицины, как таксономия, экология, теория эпидемий, генетика, медицинская диагностика и организация медицинской службы. Эти главы содержат, естественно, очень много математического материала, однако биолог или врач, не имеющие достаточной математической подготовки, смогут проследить, во всяком случае, за общим направлением работ в знакомой им области. Так, специалист в области таксономии вполне сможет понять принципы численного подхода, изложенного в гл. 7, хотя, возможно, он не всегда с одобрением отнесется к описанным методам; аналогичным образом врач, интересующийся возможностями аналитических и численных методов в диагностике, вероятно, без труда поймет большую часть гл. 11.

Такое деление книги на две части было предложено издателями, и я очень благодарен им за это предложение, так как оно сильно облегчило мою задачу; кроме того, я полагаю, что в таком виде материал, безусловно, будет более доступным для читателя.

Надеюсь, что эта книга будет интересна и полезна биологам и врачам, стремящимся ближе познакомиться с математическим подходом к их предмету. Всю первую часть и отдельные главы второй части, вероятно, полезно прочитать студентам биологических и медицинских специальностей. Я думаю, что в книге найдут нужный материал и те многочисленные читатели, которых интересует, какое значение имеет более широкое применение математики в научных исследованиях для общества в целом.

Норман Т. Дж. Бейли

Часть I

ОБЩИЕ ПРИНЦИПЫ И МЕТОДЫ

Глава 1

ПОТРЕБНОСТЬ В МАТЕМАТИКЕ

1.1. СЧЕТ И ИЗМЕРЕНИЕ

Кратковременное посещение обезьянника в богатом зоопарке всегда доставляет удовольствие и часто оказывается поучительным. Впечатление, вероятно, будет особенно сильным, если вы придете туда после нескольких часов нескончаемых споров в каком-либо учреждении или прозаседав на так называемой научной конференции. Аналогия, кроме чисто внешнего сходства, состоит в том, что и там и здесь выражаются самые разнообразные чувства, но сообщается ничтожно малое количество информации. Подходящей аналогией менее серьезной человеческой деятельности может служить оживленная болтовня поугаев в клетке. Но наблюдение за поведением приматов особенно полезно. Здесь есть все: непрерывная смена равнодушия, любопытства, желания, гнева, хитрости, подозрительности и т. д., некоторая понятливость и способность решать элементарные повседневные задачи. Единственным важным видом деятельности, отсутствие которого сразу бросается в глаза, является речь в нашем понимании.

Очевидно, что развитие языка — необходимый этап в процессе целенаправленного сознательного покорения окружающего мира. Речь, несомненно, служит более или менее существенным фактором, обеспечивающим быстрый обмен большим количеством информации между различными представителями какой-либо группы живых существ. С возникновением письменности появляется возможность регистрировать информацию, запасы которой можно непрерывно корректировать и пополнять. Однако недостаточно лишь передавать и хранить информацию. Необходимо принимать решения о том, какая информация необходима, существенна, интересна, имеет смысл и т. д. Дар речи позволил человечеству совершить огромный скачок вперед не столько в области простейшей обработки данных, сколько в способности управлять самой мыслью. Нет необходимости обсуждать здесь

вопрос о способности к мышлению у различных представителей животного мира, в частности у человекообразных обезьян. Суть дела в том, что язык позволяет сформулировать понятия, которыми можно манипулировать независимо от определяемых ими объектов. Чтобы употребление языка было успешным и, с одной стороны, помогало воздействовать на внешний мир, а с другой — обогащало духовную жизнь, необходимы определенные правила. Таким образом, язык и его грамматика развиваются, по-видимому, естественным образом. В соответствующих курсах грамматика введена в систему, однако важно отдавать себе отчет в том, что ученые лишь фиксируют фактически употребляемые правила языка. Сознательное и целенаправленное изменение языка происходит на значительно более позднем этапе развития.

Итак, язык не только служит средством выражения мыслей об окружающем мире, но по существу является важным элементом, без которого вряд ли может существовать сама мысль. По мере развития цивилизации возникли различные языки, которые стали более тонким средством выражения мыслей и общения между людьми. Наиболее поразительные успехи достигнуты в развитии, передаче и создании эмоциональных состояний, что иллюстрируется великолепными образцами ораторского искусства, поэзии и художественной литературы. Как ни странно, но сами эти успехи основаны на одном из крупных недостатков обычной речи, а именно на ее неточности. Хотя философы изучают логику еще со времен Древней Греции, ее влияние на большинство человеческих дел пока довольно незначительно. Действительно, великие философы, по-видимому, занимались в основном построением широких картин мира и не очень заботились о разработке более точного языка.

Развитие в нашем столетии лингвистического анализа вселяет большие надежды на то, что в конце концов мы глубже пойдем функции языка и научимся более тонко и более точно употреблять его. Однако достигнутые к настоящему времени практические результаты еще весьма скудны. Слишком часто теоретики-лингвисты дают основание для упреков в том, что они учатся выражать все более и более точно все меньшее и меньшее число понятий. Основное беспокойство вызывает разрыв между теорией и практикой. В большинстве случаев тот, кто старался облагораживать язык, не был достаточно связан с его повседневным употреблением в процессе той или иной конкретной практической деятельности, например при лабораторных научных исследованиях. В результате, несмотря на обилие философской литературы, посвященной научному методу, большинство ученых овладевают мастерством подобно ремесленникам — путем длительной практики, обучаясь у более опытных людей. Ускорить этот про-

цесс могло бы более сознательное освоение научных принципов и целенаправленное применение их в конкретной практической деятельности (аналогично тому, что в настоящее время происходит в промышленности, где старые медленные способы обучения сменяются новыми, более быстрыми способами, основанными на специально разработанных методах).

В области научных исследований сейчас создается критическое положение. Новые научные исследования ведутся во все возрастающем темпе, и специалистам любой области становится все труднее оставаться на уровне новых достижений в собственном предмете, не говоря уже о соответствующих смежных дисциплинах. В такой ситуации большую роль играет умение ставить эксперименты таким образом, чтобы они действительно приводили к важным выводам, сэкономили время, рабочую силу и материалы, легко интерпретировались и давали ясные результаты. В так называемых точных науках за последние три столетия достигнуты значительные успехи, и сейчас существуют все необходимые предпосылки для еще более быстрого накопления знаний в этих областях. В таких «менее строгих» науках, как биология, медицина, физиология и социология, прогресс идет медленнее и успехи менее явны. Одна из очевидных причин этого отставания состоит в том, что здесь объект исследования значительно более изменчив и сложен, чем, скажем, в физике или химии. Поэтому в биологических и общественных науках гораздо труднее использовать основные методы, которые так успешно применяются в точных науках. Каковы же главные причины непрерывных успехов точных наук?

Возьмем, к примеру, самую древнюю из точных наук — астрономию. Фантастические достижения астрономии в Вавилоне и Египте общеизвестны. Вавилоняне и египтяне не только проводили продолжительные точные наблюдения неба, но и умели распознавать существенные изменения в расположении светил, на основе которых им удалось разработать довольно точный календарь. Этот календарь предсказывал сезонные изменения и, несомненно, был крайне полезен при планировании сельскохозяйственных работ. Интересно проследить, каким образом даже на этом начальном этапе результаты точной науки использовались в практической деятельности, которую в наши дни можно было бы назвать управлением сельским хозяйством. Эти две стороны человеческой деятельности, несомненно, были взаимосвязаны и взаимодействовали друг с другом. Чтобы достигнуть таких результатов, необходимо было не просто смотреть на небо и наблюдать за множеством более или менее ярких светил, а делать нечто значительно большее. Важно было наблюдать за расположением видимых светил и измерять их положение. Кроме того, вы-

полнение таких измерений для целей предсказания означает существование весьма точных арифметических методов. Такие методы были разработаны математиками Вавилона и Египта отнюдь не случайно. Вероятнее всего, математика, астрономия и их сельскохозяйственные приложения развивались одновременно и оказывали взаимное влияние друг на друга. Впоследствии, с упадком этих цивилизаций динамическое равновесие теории и практики было утрачено, хотя некоторые отрывочные конкретные знания сохранились. В Европе точные науки не достигали столь высокого уровня вплоть до XV в., но начиная с этого времени развитие их благодаря исключительно благоприятной интеллектуальной атмосфере эпохи Возрождения шло быстро и непрерывно.

Важный вывод, который необходимо сделать из всего сказанного, состоит в том, что даже в древнем мире нельзя было получить и использовать точные знания без основных понятий счета и измерения, а также арифметики, необходимой для выполнения этих операций. Тем не менее могли существовать и действительно существовали такие цивилизации, где эти знания находились в зачаточном состоянии. Мы уже указывали, что неточность обычной речи, по-видимому, обусловила успехи в тех областях культуры, которые связаны с передачей и созданием эмоциональных состояний. Но даже в этом плане относительно точные утверждения могут быть высказаны только с помощью понятий, которые выражаются математически хотя бы в неявном виде. При любом детальном анализе картины или архитектурного сооружения невозможно избежать обращения к таким понятиям, как форма, стиль, пропорция, пространство, перспектива, вес, равновесие, цвет и т. д. Все это влечет за собой определенные численные измерения и даже использование элементарной стереометрии. Аналогичным образом при любом детальном обсуждении музыкального или литературного произведения неизбежно употребление некоторых математических понятий. Можно возразить, что все по-настоящему важные стороны искусства и литературы не поддаются численной оценке. Это справедливо, если речь идет о субъективных впечатлениях. Но как только начинается обсуждение вопросов, означающих нечто большее, чем просто выражение личных симпатий и антипатий, приходится переходить на язык, содержащий заметную долю количественных элементов.

Вопрос о том, как далеко можно развить эту идею, выходит за рамки нашего изложения. Но мы вновь повторяем, что большинство утверждений, предназначенных для передачи точной информации, можно сформулировать математически. Хотя многие люди, если спросить их мнение усомнятся в этом, большинство из них на самом деле поступает так, как если бы они считали, что

эта мысль верна. Человек может критиковать здание, найдя его безобразным. Но как только он скажет, что именно ему не нравится, выяснится, что дело сводится к неудовлетворительным пропорциям или измерениям или даже к тому, что это здание нельзя будет достаточно эффективно использовать по назначению. Другой человек может предложить некоторый образ действий, представляющий собой выбор меньшего из двух зол. Это автоматически предполагает существование некоторого грубого способа измерения и сравнения количества «зла», связанного с каждым образом действия. Вот еще один пример, более близкий к основному предмету настоящей книги. В одном из медицинских журналов автор статьи жалуется на неуместность применения математической статистики к некоторым проблемам. Затем он начинает опровергать статистиков, приводя в качестве доказательства собственные факты и цифры, рассматривая средние значения, сравнивая проценты и т. д.

Разумеется, совершенно справедливо, что во многих случаях очень трудно получить действительно удовлетворительное количественное определение. Меньше всего затруднений в этом плане возникает в таких точных науках, как физика и химия, больше всего — в области искусства и этики. Биология и медицина находятся где-то посередине, и основная цель этой книги состоит в том, чтобы показать, чего можно здесь добиться. В следующем разделе более детально рассматриваются проблемы построения математических моделей в этой конкретной области знания.

1.2. МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Мы уже указывали, что обычный язык не только позволяет нам упорядочивать наши мысли и идеи настолько, чтобы ими можно было манипулировать, т. е. записывать их и передавать, но в значительной мере определяет и существование самих мыслей. Это взаимодействие мысли и языка и фактически их совместное развитие можно сравнить с соответствующей взаимозависимостью мысли и математики, когда мы переходим к использованию последней для более точных описаний. Если рассматриваются такие сложные вопросы, как теория относительности или квантовая теория, то справедливость этого утверждения самоочевидна. Некоторые основные идеи этих теорий можно проиллюстрировать, рассмотрев, скажем, известный парадокс часов и переход элементарных частиц с одной орбиты на другую. И даже здесь потребуется определенный объем элементарной математики. Однако описание наиболее важных вопросов, составляющих существо этих теорий, возможно лишь с помощью очень сложного математического аппарата; без такого описания дальней-

шая творческая мысль невозможна. Большинство биологических и медицинских исследований не требует столь сложного математического анализа. Однако достаточно рассмотреть лишь применение статистики при планировании и анализе экспериментов, теории случайных процессов — при изучении роста биологических популяций или теории информации — при обсуждении некоторых вопросов биосинтеза белка, чтобы понять, что в биологической науке существует много важных проблем, где математические методы действительно обеспечивают такое проникновение в суть дела, которого невозможно достигнуть чисто описательным путем. Некоторые из этих приложений будут рассмотрены далее более подробно. Сейчас наша цель состоит в том, чтобы познакомиться поближе лишь с первым этапом построения математических моделей биологических явлений.

Бурное развитие прикладной математики в XVII в. сразу же привело к впечатляющим успехам в изучении реального мира. Рассмотрим, например, область ньютоновой механики и оптики. Наиболее смелые люди отваживались даже в то время использовать математический подход при изучении форм живых организмов и их развития. Так, Галилей рассматривал связь между прочностью материалов и относительными пропорциями соответствующих конструкций не только в применении к постройке судов или сооружению дворцов, но и в связи со строением растений и животных. Например, он пытался установить максимальную высоту дерева, при превышении которой небольшое отклонение ствола приводит к постоянному изгибу, изучить влияние увеличения размеров тела животного на размеры и прочность скелета и мышц, а также выявить различные физические изменения, происходящие в организме, когда такое животное, как кит, погружается в воду и вес его тела уравновешивается весом вытесненной им воды. В XVII в. Уильям Гарвей подходил к объяснению таких физиологических явлений, как кровообращение, с позиций механики, хотя более точные описания — через гидродинамическую теорию — появились только в XIX в. При изучении многих вопросов такого рода могут найти применение методы физики и химии, а также некоторые разделы прикладной математики. Однако в этой книге нас будут интересовать прежде всего те аспекты биологии и медицины, которые лежат вне области, непосредственно связанной с точными науками. В то же время нельзя забывать, что физические и химические понятия лежат в основе большинства (если не всех) сравнительно абстрактных исследований, скажем таких, как изучение поведения животных или эволюционного развития.

Начнем с того, что большинство биологических работ носит описательный характер. При описании растений и животных рас-

смаатриваются их форма, размер, цвет, поведение, распространение, сходство с другими организмами или отличие от них и т. д. Это, в сущности, стадия естественной истории. В простейшем виде такое описание носит главным образом словесный характер и насыщено различными подробностями. Однако по мере того, как наблюдения становились более точными и обширными, появилась возможность применения математического языка для описания или отображения огромного многообразия форм жизни. Высшая ступень в развитии такого рода описательно-математической естественной истории была отмечена появлением в 1917 г. книги д'Арси Томпсона «Рост и форма», которая в настоящее время считается классической. Значительная ее часть посвящена применению обычной статики, динамики, физики и химии для получения более детального и более сложного описания живых организмов, чем это возможно с помощью словесных средств. Значительное внимание уделяется также правильному выбору математических понятий и теорий для объяснения сложности и запутанности наблюдаемых форм.

Рассмотрим, например, замечательные гексагональные образования в архитектуре ячеек медовых сотов, детальное описание которых восходит еще к Паппусу из Александрии. На первый взгляд эту конфигурацию можно рассматривать как результат сжатия двойного слоя одинаковых цилиндров с круговым поперечным сечением; при таком сжатии поперечные сечения должны принять вид шестиугольников. Однако истинная форма получающихся многогранников требует более тщательного изучения. Можно показать, как это сделал Кеплер еще в XVII в., что пространственная симметрия медовых сотов должна привести к расположению плоскостей и углов, наблюдаемому в правильном ромбо-додекаэдре. Впоследствии оказалось, что такое тело имеет определенные оптимальные свойства, а именно площадь его поверхности при определенных условиях минимальна. На этом основании некоторые ученые предположили, что пчела осмысленно выбирает эту определенную форму медовых сотов, чтобы экономить воск. Однако в дальнейшем геометрическая форма сотов была уточнена, и это предположение пришлось изменить. Так, наблюдаемый ромбо-додекаэдр содержит ряд гексагональных призм с ромбоидальными пирамидами в качестве оснований. Допустим теперь, что основанием служит плоскость. Можно показать, что такое тело будет всего на 2% менее эффективным. Вряд ли можно предположить, что пчела производит настолько точные вычисления. Кроме того, при более тщательном исследовании оказалось, что форма медовых сотов отклоняется от идеальной математической фигуры: толщина их стенок неодинакова, ячейки никогда не располагаются строго по горизонтали, край

не совсем прямые, на гранях имеются неровные подтеки воска и т. д. Более вероятно, что наблюдаемая геометрическая форма возникает в результате совместного действия силы тяжести, взаимного давления и поверхностного натяжения. Основная ценность таких исследований заключается в том, что они иллюстрируют огромную силу математического описания, или, как сейчас говорят, математической модели. Отыскание в данном конкретном случае правильного многогранника, столь удачно описывающего объект, встречающийся в природе, дает нам не только определенное эстетическое удовлетворение. Такого рода модель служит полезной основой для дальнейших рассуждений и исследований. Следует отметить также, что такой подход обладает большой эффективностью и гибкостью, несмотря на то (и скорее даже благодаря тому) что математическая модель является лишь аппроксимацией действительности. В самом деле, если бы модель слишком точно имитировала реальную действительность, математические выражения оказались бы чрезмерно сложными и их обработка была бы связана с непреодолимыми трудностями. Далее мы остановимся на этом противоречии более подробно.

Еще один интересный вопрос, дающий широкие возможности для применения математики,— листорасположение у деревьев и других растений; этот вопрос был также довольно детально рассмотрен д'Арсии Томпсоном. Не только листья на стеблях многих растений, но и отдельные цветочки в соцветии подсолнечника, чешуйки в еловой шишке и т. д. образуют замечательно правильные спирали, возбуждающие интерес многих биологов и математиков. Так, у шишек норвежской ели имеется пять рядов чешуек, круто поднимающихся в одном направлении, и три ряда, идущих более полого в противоположном направлении. У обычной лиственницы число таких рядов равно соответственно пяти и восьми. Цветорасположение у гигантского подсолнечника имеет аналогичную схему, и числа рядов равны соответственно 34 и 55, 55 и 89, 89 и 144 и т. д. Хотя имеют место и исключения, пары чисел, встречающиеся наиболее часто, являются двумя последовательными членами арифметического ряда Фибоначчи (итальянский математик, известный еще как Леонардо из Пизы, 1170—1250 гг.):

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144 ...

Ряд начинается с единицы, и каждый его член в точности равен сумме двух предыдущих. Одно из наиболее важных математических свойств этого ряда состоит в том, что последовательность дробей, образованных путем деления каждого члена на последующий, т. е. ряд

$$1, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \frac{8}{13} \dots,$$

стремится к так называемому золотому сечению, точное значение которого равно $\frac{1}{2}(\sqrt{5} - 1) = 0,618$. . . Если отрезок прямой разделен на две части золотым сечением, т. е. так, что отношение большей части ко всему отрезку равно $0,618$. . . , то отношение меньшей части к большей также является золотым сечением. Эта пропорция и ее свойства были хорошо известны древним грекам, и ей часто придавали важную роль в эстетических дискуссиях об искусстве и архитектуре. Разве не удивительно, что такие же числа встречаются и в живой природе?

Рассмотрим рост одиночного стебля растения (считая его идеальным круговым цилиндром), на поверхности которого по спирали на одинаковом расстоянии друг от друга появляются новые листья: A, B, C, D ... Можно ожидать, что листья расположатся при этом в несколько вертикальных рядов, равноудаленных друг от друга по поверхности цилиндра. Однако в общем случае такого расположения не возникает. При внимательном рассмотрении стебля глаз легко выделит на нем некоторую другую спираль, например A, C, E ... Действительно, на основе геометрических соображений можно показать, что на стебле имеются две дополняющие друг друга системы спиралей: одна — правосторонняя, а другая — левосторонняя, и обе они отличаются от исходной генетической спирали A, B, C ... Кроме того, если O — лист, к которому можно прийти из точки A за m шагов, двигаясь по одной спирали, и за n шагов, двигаясь по соответствующей дополнительной спирали, то можно показать, что m и n представляют собой последовательные члены ряда Фибоначчи и что путем надлежащего выбора спирали можно выделить все предыдущие пары членов ряда. Хотя эти соотношения приближенно справедливы только для почти цилиндрических стеблей или ветвей, при соответствующей деформации цилиндра топологически идентичное построение можно выполнить и для таких структур, как еловая шишка или соцветие подсолнечника. По-видимому, различные на глаз дополнительные спирали вместе с соответствующими числами Фибоначчи являются неизбежным следствием регулярного роста листьев на основной генетической спирали. В известном смысле появление определенных арифметических соотношений является математическим совпадением. Однако единство математики состоит в том, что результаты и теоремы, вытекающие из данной группы допущений, применимы во всех тех случаях, когда эти допущения можно считать справедливыми.

Еще больше места (около 100 страниц) д'Арси Томпсон уделяет часто встречающейся в природе логарифмической спирали, которая связана с определенными видами роста. Наиболее простой вид имеет спираль, которую вычерчивает точка P при вращении

радиуса-вектора OP с постоянной угловой скоростью относительно начала координат O при одновременном движении точки P из точки O вдоль линии OP с постоянной скоростью. Это так называемая архимедова спираль; примерно такую форму принимает канат, аккуратно свернутый в бухту на палубе судна и имеющий витки одинаковой толщины. В полярных координатах уравнение спирали имеет вид $r = a\theta$. Однако у очень многих моллюсков, например у *Nautilus pompilius*, *Turritella duplicata*, *Ammonites* и т. д., последовательные витки не одинаковы, а все более и более утолщаются. Во многих случаях приближенные значения толщины последовательных витков образуют геометрическую прогрессию. Спираль, обладающая этим свойством, называется логарифмической; ее можно построить, заставив точку P двигаться из начала координат O не с постоянной скоростью, как в описанной выше архимедовой спирали, а с возрастающей скоростью, пропорциональной расстоянию от точки O . Построенная спираль может быть описана уравнением в полярных координатах $r = a^{\theta}$, или $\log r = \theta \log a$. Эта спираль обладает рядом интересных свойств, в том числе следующими. Во-первых, кривая пересекает все радиусы под одним и тем же углом. Отсюда одно из ее названий — равноугольная. Во-вторых, как уже указывалось, расстояния между последовательными витками образуют геометрическую прогрессию. В-третьих, последовательность длин радиусов, образующих одинаковые углы друг с другом, также составляет геометрическую прогрессию. Кроме того, наконец, образующиеся в процессе расширения секторы, отсекаемые такими радиусами, подобны друг другу.

Хотя саму раковину моллюска нельзя назвать живой, она образуется растущим организмом. Один из простейших способов наращивания нового вещества автоматической приводит к образованию некоторой фигуры, очень близкой к логарифмической спирали. Раковина моллюска *Nautilus* представляет собой, конечно, трехмерную структуру, однако любая последовательность соответствующих точек ее поверхности (например, точек, определяющих контуры раковины в процессе ее роста) лежит более или менее в одной плоскости. Рост раковин таких моллюсков, как *Turritella duplicata*, происходит в трех измерениях. Тем не менее в этом случае пригодно аналогичное математическое описание, хотя на этот раз мы должны оперировать с более широким понятием логарифмической спирали, навернутой на конус.

Во многих раковинах обнаруживается поразительно близкое совпадение между результатами измерений и теоретическими значениями, ожидаемыми для точной логарифмической спирали. Еще более удивительно, что во многих случаях аналогичная кривая довольно точно описывает образование крышечки, закры-

вающей вход в раковину. Эта крышечка растет путем постепенного нарастания с одного края, одновременно передвигаясь вдоль раковины и поворачиваясь вокруг своей оси таким образом, что в конце концов образуется плотная пробка. Ясно, что дать строгое геометрическое описание такого великолепного приспособления довольно трудно.

Из предыдущих примеров видно, что математические построения сравнительно простого типа описывают некоторые биологические явления с удивительной степенью правдоподобия. Конечно, если попытаться описать надлежащими геометрическими кривыми форму тела лошади, то можно быстро прийти к выводу, что математика не подходит для описания объектов такого рода. Именно из-за огромной сложности мира живых организмов многие биологи сомневаются в возможности широкого применения математических методов. Но хотя простые математические методы находят в биологии лишь ограниченное применение, это не означает, что сложные явления жизни недоступны математическому анализу. Напротив, как уже указывалось в предыдущем разделе, мы сталкиваемся с математическими понятиями всякий раз, когда пытаемся внести любое уточнение при обсуждении идей и передаче информации.

Поэтому на самом деле вопрос состоит в том, какого рода математическое исследование подходит к той или иной конкретной задаче, какие ограничения имеют место в каждом случае и каким образом эти методы можно развивать дальше.

Не следует, разумеется, забывать, что любая математическая модель — это идеализированное, абстрактное построение, которое в лучшем случае лишь частично соответствует действительности. Так, мы видели, что с определенной степенью точности поперечное сечение пчелиных сотов можно считать более или менее шестиугольным, однако при достаточно близком рассмотрении наблюдаются заметные отклонения от идеальной геометрической формы. Аналогично логарифмическая спираль, выбранная надлежащим образом, приближенно отражает форму раковины моллюска *Nautilus*, однако при более тщательных наблюдениях и измерениях легко обнаруживается, что она заметно отклоняется от теоретической кривой. И в лучшем случае удастся описать лишь основную общую форму раковины; при попытке математического описания ее тонкой структуры, не говоря уже об описании находящегося в раковине живого организма, возникают значительно более сложные проблемы.

Чем сложнее рассматриваемое явление, тем труднее построить его достаточно точную количественную модель; частично это обусловлено самой природой некоторых процессов, затрудняющей их измерение, а частично тем, что слишком сложные математи-

ческие модели чрезмерно громоздки и не имеют практической ценности. При описании целого организма нам часто приходится ограничиваться такими общими показателями, как вес и размер, а также общим описанием анатомии и физиологии, которое удается подкрепить лишь отдельными количественными характеристиками некоторых частей тела или некоторых функций. При изучении групп животных степень абстракции еще больше. В этих случаях могут рассматриваться число индивидуумов в популяции, их возраст и пол, возможно, некоторые фенотипические признаки (соответствующие весьма небольшому числу генетических локусов) или некоторые легко поддающиеся наблюдению схемы поведения и т. п.

Для того чтобы решить, какие факторы можно считать несущественными в том или ином случае, необходимо хорошо разобраться в исследуемой задаче. Довольно большое число деталей приходится опускать даже при чисто словесном биологическом описании. Еще больше деталей опускают при построении математических моделей биологических явлений, но за счет этого все, что включается в модель, может быть описано очень точно. В то же время следует учесть, что сложные экологические системы могут быть связаны с таким огромным количеством биологически существенной информации, которую человек не в состоянии охватить полностью, пользуясь лишь обычными словесными описаниями. В этом случае ценную помощь могут оказать статистические методы компактного представления и обработки информации. Кроме того, больших успехов можно добиться с помощью существующих в настоящее время методов обработки данных на больших электронных вычислительных машинах. Более полно эти вопросы обсуждаются в последующих главах.

Нередко мы сталкиваемся с противоречием между реалистичностью модели и возможностью оперировать ею. Очень точное описание всех известных фактов, скажем, о некотором развивающемся виде не позволит выполнить какие-либо полезные вычисления или осуществить какое-либо прогнозирование. В то же время чрезмерно упрощенная математическая модель позволит легко оценить численные значения определенных коэффициентов, но они будут крайне неточны, поскольку были опущены некоторые существенные факты. Это извечная дилемма, возникающая при любой научной работе. Она отнюдь не специфична для биологии и медицины, хотя принимает здесь особенно острую форму. Более подробно мы коснемся этого важного вопроса в гл. 3, где рассматривается процесс научного исследования в целом.

1.3. ДИНАМИЧЕСКИЕ МОДЕЛИ

Все живые существа рождаются, растут, а затем стареют, претерпевают непрерывные изменения и превращения и в конце концов умирают; иными словами, все они всегда вовлечены в какие-то динамические процессы развития во времени. В мире неживой природы также непрерывно протекают различные динамические процессы, и некоторые философы, как, например, Гераклит Эфесский в Древней Греции, положили понятие непрерывного изменения и движения в основу своего мировоззрения. Многие задачи чисто физического характера — движение планет солнечной системы, траектория снаряда, поток тепловой энергии или электрический ток, периодическое движение приливной волны и другие задачи динамики — могут быть решены с достаточной степенью точности в математической форме. Изменение во времени обычно приводит непосредственно к выводу дифференциальных уравнений. В настоящее время исследован широкий круг дифференциальных уравнений, особенно в области физики, и прикладная математика получила в высшей степени полезный аппарат, известный под общим названием «уравнений математической физики».

Живые существа с их саморегуляцией, способностью к приспособлению, целенаправленной активностью и сложными схемами поведения труднее втиснуть в рамки общих математических законов. В разд. 1.2, где говорилось о математическом описании статических структур, были детально рассмотрены пчелиные соты, листорасположение у растений и форма раковины у моллюсков. Даже здесь мы не могли не коснуться процесса роста и развития, в результате которого появилась рассматриваемая структура, и естественно, что этот процесс до некоторой степени определяет выбор соответствующего математического метода. Однако в настоящем разделе нам необходимо исследовать более конкретно те ситуации, в которых динамическое изменение и развитие обнаруживаются в явной форме с самого начала. По-видимому, наиболее простыми процессами такого рода являются развитие индивидуума и рост популяции. Впервые эти вопросы широко рассмотрел Кетле в 1835 г. в своей знаменитой книге «*Essai de Physique Sociale*».

Очевидно, что вес и еще один-два простых показателя лишь довольно грубо описывают развитие отдельного индивидуума. Тем не менее общепризнано (и совершенно правильно), что такие показатели, если уделяется должное внимание и другим факторам, весьма полезны в качестве ориентира. Предполагается, что если детеныш к определенному возрасту не набрал вес, близкий к среднему значению для данного вида, то, по-видимому, что-то неладно с ним самим или с его питанием. Однако это справедливо не во

всех случаях. Кривые роста многих здоровых индивидуумов значительно отличаются от кривых роста для большинства индивидуумов того же вида. Единственное, что можно сказать по этому поводу, это то, что такое расхождение нужно рассматривать как указание на необходимость дальнейшего исследования. Страховые компании обычно подозрительно относятся к значительным отклонениям в весе (особенно в сторону увеличения) за относительно короткие промежутки времени. Хотя в отдельных случаях в основе таких отклонений может и не быть никакой патологии, данные страховых таблиц, составленных с учетом возраста, веса и смертности, достаточно хорошо обоснованы и весьма серьезно принимаются во внимание любой страховой компанией.

Средний здоровый ребенок прибавляет в весе в первые несколько месяцев довольно устойчиво. Со временем темпы несколько снижаются, но общий вес продолжает расти. Наблюдения, проведенные в течение нескольких лет, показывают (во всяком случае, для некоторых групп населения), что у мальчиков прибавление в весе становится минимальным, когда они достигают примерно семилетнего возраста, затем вес увеличивается быстрее, и скорость прибавления в весе достигает максимума примерно к четырнадцати годам, после чего снова снижается. Для девочек график несколько сложнее и представляет собой двухвершинную кривую с максимумами примерно в десять и шестнадцать лет. Аналогичную форму имеют и кривые роста, хотя в деталях они отличаются от кривых веса; рост значительно замедляется перед основным максимумом, и этот максимум — единственный как у мальчиков, так и у девочек.

Кривые значений веса и роста и их приращений могут быть точно описаны математически. Иногда в литературе сообщается о том, что для подбора многочленов высокого порядка, возможно точнее описывающих полученную экспериментальную кривую, были выполнены громоздкие вычисления. По нашему мнению, вряд ли стоит это делать. Кривая, построенная от руки и проходящая через все точки кривой, дает практически всю требуемую информацию. В частности, графики приращений веса или роста совпадают не только с повседневными наблюдениями, показывающими, что в некоторых интервалах времени вес почти не меняется, а в других быстро растет, особенно в период полового созревания, но и хорошо согласуются с результатами более детальных физиологических и биохимических исследований. Таким образом, измерение роста или веса дает некоторую количественную информацию о жизнедеятельности растущего организма и весьма элементарно характеризует динамику процесса развития. Трудность заключается в том, что такие приближенные методы не дают ничего, кроме простого описания. Вряд ли можно назвать

такое математическое описание моделью и на его основе получить какие-либо обобщающие выводы.

Обратимся теперь к росту популяции в целом. Под популяцией мы обычно подразумеваем совокупность определенного *числа* индивидуумов (возможно, различающихся по возрасту и полу), а не совокупность результатов измерений какого-либо физического признака. Очевидно, что число организмов в популяции представляет собой лишь один аспект в бесконечном многообразии биологического материала. Тем не менее эта величина служит важным ключом к пониманию поведения всей группы в целом и предоставляет широкие возможности для математического исследования. Во многих биологических работах исследуются такие проблемы, как развитие и эволюция видов, конкуренция между видами, влияние факторов окружающей среды, распространение эпидемий и т. д. Ни одно из этих исследований не может быть сколь-нибудь точным, если оно не начинается с построения некоторой достаточно приемлемой математической модели рассматриваемой популяции. Здесь, как и всегда, приходится идти на компромисс между двумя крайностями: очень простую модель легко построить математически, но она слишком нереалистична, чтобы ей можно было доверять, а с очень сложной моделью, которая значительно ближе к реальной действительности, очень трудно оперировать и ее очень трудно объяснить.

Одна из простейших моделей роста популяции принадлежит Т. Мальтусу, который в конце XVIII в. заметил, что популяции имеют тенденцию увеличиваться в геометрической прогрессии. Мальтуса беспокоило то, что, по его мнению, средства существования могут возрастать только в арифметической прогрессии и что рано или поздно их станет недостаточно. Здесь не место обсуждать экономические и социальные следствия из теории Мальтуса. В природе численность большинства живых существ действительно способна увеличиваться в геометрической прогрессии, однако рост популяций в достаточной мере сдерживают такие факторы, как борьба за существование, болезни, естественная гибель и уничтожение хищниками. Обычно если популяция начинает развиваться в среде с достаточным количеством пищи и относительно небольшим количеством хищников, то сначала ее численность растет очень быстро. С течением времени запасы пищи истощаются, перенаселенность приводит к условиям, менее благоприятным для выживания, плодовитость снижается и смертность увеличивается. При определенных условиях достигается равновесное состояние и численность популяции становится более или менее постоянной. Очевидно, что очень важно знать точное соотношение между численностью популяции в различные моменты времени и скоростями размножения и гибели.

Математическую форму этой типичной S-образной кривой роста популяции впервые получил Ферхюльст, современник Кетле. Он использовал следующий подход. Во-первых, удобно рассматривать численность популяции n как непрерывную переменную, что вполне допустимо, если n довольно велико. Во-вторых, рассматривается непрерывное время t , а не дискретные поколения. Допустим, что средняя скорость роста популяции при благоприятных условиях составляет m на одного индивидуума, так что за время dt численность популяции увеличивается на $mndt$. Это означает, что $dn = mndt$. Поэтому изменение численности популяции описывается дифференциальным уравнением

$$\frac{dn}{dt} = mn, \quad (1.1)$$

решение которого имеет вид

$$n = ae^{mt}, \quad (1.2)$$

где a — число индивидуумов в начальный момент времени $t = 0$. Экспоненциальный рост непрерывной популяции в непрерывном времени, описываемый формулой (1.2), эквивалентен геометрической прогрессии для дискретной численности популяции в предположении дискретной смены поколений.

Идея Ферхюльста состояла в наложении на экспоненциальный рост, выраженный формулой (1.2), некоторого фактора, характеризующего замедление и увеличивающегося с ростом популяции. Простейшее из возможных допущений состоит в том, что степень замедления роста для одного индивидуума пропорциональна размеру популяции, т. е. что результирующая скорость роста равна не m , а $m - rn$, где r — коэффициент замедления роста. В этом случае дифференциальное уравнение (1.1) принимает вид

$$\frac{dn}{dt} = mn - rn^2, \quad (1.3)$$

а его решение выражается формулой

$$n = \frac{m}{r + (m/a - r)e^{-mt}} \quad (1.4)$$

(здесь, как и ранее, a есть численность популяции в начальный момент $t = 0$). Уравнение (1.4) описывает S-образную кривую (так называемая *логистическая кривая* Ферхюльста), наклон которой вначале монотонно возрастает, как у экспоненты, а затем постепенно уменьшается до нуля; при больших значениях t кривая сливается с горизонтальной прямой $n = m/r$ (m/r есть равновесное значение, к которому стремится размер популяции).

На практике скорость роста m определяется скоростями размножения и гибели. Если скорость размножения β такова,

что за время dt появляется βndt новых индивидуумов, а скорость гибели μ такова, что за то же время погибает μndt индивидуумов, то $m = \beta - \mu$. Значение m может быть положительным, отрицательным или нулевым, и в зависимости от этого популяция будет расти, уменьшаться или оставаться неизменной. Аналогичным образом можно включить в рассмотрение и многие другие факторы, например миграцию.

Логистическая кривая широко используется при описании роста популяций животных и людей. Например, в демографии предпринимались попытки подобрать логистические кривые для данных о численности населения различных стран, в частности США, в прошлом и в настоящее время с целью использовать продолжение полученной кривой для прогноза численности населения при достижении равновесного состояния. Хотя такие методы имеют положительное значение, поскольку благодаря применению математики обсуждение этих проблем становится более конкретным, к оценке результатов, получаемых с их помощью, необходимо подходить с большой осторожностью. Окружающая среда и социальные условия могут измениться, и, следовательно, изменятся скорости размножения и гибели и коэффициент замедления роста. Модель может быть приближенно справедливой для того периода, по которому имеются данные, однако это не гарантирует ее справедливости в будущем. Экстраполяция всегда сопряжена с неопределенностью, однако в данном случае некоторые трудности можно преодолеть путем более детального анализа популяции, рассматривая скорости размножения и гибели для различных возрастных групп с учетом пола.

Достаточно полное описание первых работ, посвященных различным формам кривых роста, полученных как для отдельных представителей вида, так и для целых популяций, можно найти в гл. 3 книги д'Арси Томпсона «Рост и форма».

Важной вехой в динамическом подходе к биологическим явлениям было издание в 1925 г. книги Лотки «Элементы физической биологии» [40], которая впоследствии переиздавалась под названием «Элементы математической биологии». Эта книга состоит из четырех разделов: «Общие принципы», «Кинетика», «Статика» и «Динамика». Значительная часть книги посвящена вопросам, которые относятся главным образом к физике и химии, однако практически в каждой главе иллюстрируется важная роль изменений, движения и развития в биологии. Даже раздел, озаглавленный «Статика», посвящен в основном состояниям равновесия, достигаемым в различных непрерывных процессах, независимо от того, являются ли они физико-химическими (как, например, круговорот азота и фосфора в природе) или преимущественно биологическими (например, конкуренция между видами). Книга

Лотки, примечательная тем, что в ней рассматривается необычайно широкий круг явлений, оказала значительное влияние на развитие математической биологии.

Приведенные выше уравнения (1.1) и (1.3) напоминают уравнения движения, получаемые при математическом описании динамических систем. Это впечатление еще более усиливается после прочтения нескольких глав книги Лотки, посвященных фундаментальным уравнениям кинетики развивающихся систем. Выше, рассматривая (на очень элементарном уровне) рост популяций, мы сделали несколько упрощающих допущений, в частности о том, что рост представляет собой непрерывный процесс — как в отношении размера популяции, так и во времени. Кроме того, подразумевалось, что биологическая изменчивость или случайные статистические колебания либо отсутствуют, либо настолько малы, что не имеют существенного значения. Если популяция достаточно велика, то это допущение часто оказывается близким к истине. Кроме того, даже в том случае, если размер популяции испытывает заметные колебания, можно все же применять эти уравнения, полагая, что они относятся к *средним* значениям. Однако необходимо иметь в виду следующее важное обстоятельство. Для некоторых простых явлений, как, например, размножение, гибель и миграция, можно спокойно пренебречь присущей им заметной изменчивостью и выводить уравнения движения для средних значений, как если бы эти средние значения были фактически наблюдаемыми величинами, не подверженными воздействию статистических колебаний. В то же время при исследовании, например, конкуренции между видами, развития эпидемий и вообще любых процессов, в которых участвуют взаимодействующие группы, средние значения, получающиеся из уравнений, выведенных при допущении об отсутствии статистических колебаний, обычно отличаются от истинных средних значений. Поэтому мы откладываем обсуждение таких ситуаций до следующей главы, где рассматривается роль теории вероятностей и математической статистики при построении математических моделей.

Необходимо упомянуть и о том, что если принять во внимание статистические колебания, то даже рассмотренная выше логистическая модель, аналогичная модели простой эпидемии без случаев выздоровления, должна будет претерпеть значительные изменения. Следовательно, *детерминистский* подход необходимо применять с большой осторожностью. Многие результаты, приводимые в старых книгах, например в книге Лотки, справедливы лишь для тех случаев, когда на протяжении всего процесса численность популяции остается достаточно большой.

Практическое значение математических моделей, рассмотренных в данном разделе, состоит в том, что они дают предваритель-

ное количественное представление об изучаемых процессах. Используемые в них параметры (например, скорость размножения) имеют определенный биологический смысл, и это позволяет проверить соответствие модели тому реальному процессу, который, как предполагается, она описывает. На основании полученных данных можно вычислить соответствующие значения параметров и использовать их как основу для дальнейшего исследования.

1.4. ОБЛАСТИ ПРИМЕНЕНИЯ МАТЕМАТИЧЕСКИХ МЕТОДОВ

Несколько лет назад, когда автор этой книги работал консультантом по вопросам математической статистики в небольшой медицинской научно-исследовательской группе, разговоры о возможности проложить математическую тропинку через густые дебри экологических факторов часто заканчивались довольно скептическим покачиванием головой и утверждением, что «медицина — это все-таки искусство». Отчасти это, конечно, верно в том смысле, что интуиция и воображение для врача действительно необходимы. В то же время большинство больных и потенциальных больных, несомненно, надеются на непрерывное развитие и расширение научных аспектов медицины. А наука означает применение математики.

Существенно важен вопрос о том, в каких областях применимы математические методы. В разд. 1.1 мы уже отмечали, что потребность в математическом описании появляется при любой попытке вести обсуждение в точных понятиях и что это касается даже таких сложных областей, как искусство и этика. В настоящем разделе мы несколько конкретнее рассмотрим области применения математики в биологии и медицине.

Хорошо известно, что один из подходов к описанию картины природы — это построение иерархии уровней организации, изучаемых различными науками; по уровню абстракции, свойственному каждой из них, эти науки можно расположить в такой последовательности: физика, химия, биохимия, физиология, психология, социология. Мы начинаем с основных материальных элементов реального мира, т. е. с субатомного уровня, и заканчиваем необычайно разносторонними проявлениями духовной жизни человеческого общества. В этой последовательности уровней организация и сложность непрерывно повышаются. На каждом уровне действуют свои собственные законы, и поэтому их можно изучать до некоторой степени независимо друг от друга. Однако любой из них неразрывно связан с закономерностями, действующими на более низких уровнях. Так, законы физики и химии отчасти распространяются и на психологию, хотя понятия и законы последней выходят за пределы физических и хими-

ческих законов. Проблемы, касающиеся организации и деятельности больниц, следует отнести к более высокому уровню абстракции, чем, скажем, физиологию и патологию человека. Но хотя в определенной степени логическое содержание этого более высокого уровня независимо от более низкого, вопросы физиологии и патологии неизбежно должны учитываться при решении любой проблемы, касающейся организации больничных служб. Мы не собираемся углубляться здесь в эти философские рассуждения или обсуждать отдельные их детали, а хотим лишь подчеркнуть, что описанная последовательность уровней приближенно соответствует порядку возрастания трудностей при использовании научных методов и проведении математических исследований.

Как мы уже отмечали, прикладная математика добилась крупных и бесспорных успехов в области физики и химии, однако в данной книге мы не будем касаться этих вопросов. В разд. 1.2 было показано, что математические описания, связанные с биологическими формами, охватывают широкий круг вопросов и могут быть проведены достаточно точно. В разд. 1.3 мы познакомились с динамическими моделями развития и коснулись проблем, связанных со случайными колебаниями численности популяций. Изложение этих вопросов требовало достаточной степени абстракции, однако именно использование упрощающих допущений позволило нам получить некоторое представление о законах, регулирующих рост популяций. Было отмечено, что при рассмотрении такого рода проблем неизбежно приходится сталкиваться с фактором статистической изменчивости, подробное обсуждение которого переносится в гл. 2.

При переходе на более высокие уровни абстракции мы сталкиваемся не только с более сложными вопросами, но и с возрастающей степенью изменчивости, по большей части непредсказуемой. Например, полная картина конкуренции между несколькими видами, обитающими в определенной среде, включает огромное множество факторов. В области научных экологических описаний, выполненных главным образом в словесной форме, достигнуты значительные успехи, однако разработка математических моделей находится здесь еще на самом элементарном уровне. Другим примером может служить область медицинской диагностики. Для постановки диагноза врач совместно с другими специалистами часто бывает вынужден учитывать самые разнообразные факты, опираясь отчасти на свой личный опыт, а отчасти на материалы, приводимые в многочисленных медицинских руководствах и журналах. Общее количество информации увеличивается со все возрастающей интенсивностью, и есть такие болезни, о которых уже столько написано, что один человек не в состоянии в точности изучить, оценить, объяснить и использовать всю имеющуюся информацию

при постановке диагноза в каждом конкретном случае. Разумеется, хороший диагност, используя свой большой опыт и интуицию, может отобрать необходимую часть важных данных и дать достаточно точное заключение. Однако, как это ни парадоксально звучит, по мере накопления знаний положение ухудшается.

Именно в такого рода ситуациях, когда разум одного человека не способен справиться со сложностями стоящих перед ним задач и описать их решение даже в общей словесной форме, специалисты в области так называемых неточных наук (включая, разумеется, биологию и медицину) часто утверждают, что математический анализ несовершенен, неуместен, приводит к ошибочным заключениям или невозможен, и поэтому его лучше избегать. Это возражение содержит рациональное зерно в том смысле, что современная математика, возможно, еще недостаточно совершенна; однако пройдет время, и мы увидим, что справедливо как раз обратное. В тех случаях, когда задача содержит большое число существенных взаимозависимых факторов, каждый из которых в значительной мере подвержен естественной изменчивости, *только* с помощью правильно выбранного статистического метода можно точно описать, объяснить и углубленно исследовать всю совокупность взаимосвязанных результатов измерений. Если число факторов или важных результатов настолько велико, что человеческий разум не в состоянии их обработать даже при введении некоторых статистических упрощений, то обработка данных может быть произведена на электронной вычислительной машине. Использование статистических методов и вычислительной техники рассматривается в гл. 2 и 5 соответственно.

Основная причина недоверия к математическим и вычислительным методам, по-видимому, состоит в следующем. Математическая модель некоторого биологического явления будет приемлемой для биолога только в том случае, если выраженная в словесной форме информация об этом явлении, которой он располагает, достаточно полна для того, чтобы можно было судить об адекватности модели. Ясно, что получение такой информации представляет собой первый и наиболее важный этап биологического исследования и что на этом этапе математика играет второстепенную роль. Естественно, возникает мысль, что по мере того, как вопросы становятся более трудными и сложными, математика приобретает все меньшее и меньшее значение. Однако не всегда учитывается то обстоятельство, что, достигнув достаточной степени сложности, математика развивается далее по своим собственным законам и дает биологу понятия и образ мышления, которых у него раньше не было. Будем надеяться, что эта книга хотя бы в некоторой степени проиллюстрирует справедливость этого утверждения.

До сих пор мы имели в виду главным образом те биологические и медицинские исследования, которые требуют более высокого уровня абстракции, чем физика и химия, но тесно связаны с этими последними. Далее мы перейдем к проблемам, связанным с поведением животных и психологией человека, т. е. к использованию прикладных наук для достижения некоторых более общих целей. Эту область довольно расплывчато называют *исследованием операций*, и более детально она рассматривается в гл. 4. Пока мы лишь отметим, что речь будет идти о применении научных методов при решении административных и организационных задач, особенно тех, которые непосредственно или косвенно связаны с биологией и медициной. Лесоводство, животноводство, общие вопросы сельскохозяйственного производства, проектирование больниц и организация медицинского обслуживания — таковы лишь немногие вопросы, относящиеся к этой категории.

Разумеется, не все задачи административного управления можно решить на научной основе, используя методы исследования операций. Однако применение этих методов там, где оно возможно (а они применимы ко многим задачам такого рода), имеет большие преимущества, так как позволяет расширить область точных исследований и сократить область неопределенных словесных рассуждений. Благодаря этому интуиция и здравый смысл человека могут быть направлены на решение тех вопросов, где невозможно применение шаблонных методов. Еще более сложные вопросы, к которым приравниваются какие-либо этические соображения. Но иногда математический анализ может помочь даже и в этих случаях. Например, в медицине часто возникают сложные проблемы, связанные с применением лекарственных препаратов, которые еще находятся на стадии испытания. Морально врач обязан предложить своему больному наилучший из существующих препаратов, но фактически он не может сделать выбор, пока испытание не будет закончено. В этих случаях применение правильно спланированных *последовательностных* статистических испытаний позволяет сократить время, требуемое для получения окончательных результатов. Этические проблемы при этом не снимаются, однако такой математический подход несколько облегчает их решение. О последовательностных методах более подробно говорится в разд. 2.3.

Основное положение настоящего раздела состоит в том, что математические методы применимы к самому широкому кругу вопросов — от физики элементарных частиц до моральных проблем. Удобно (хотя вовсе не обязательно) рассматривать некую иерархию уровней. По мере перехода на более абстрактные уровни математические методы оказываются менее разработанными и применять их становится все труднее. Тем не менее при пра-

вильном применении математический подход не отличается существенно от подхода, основанного просто на здравом смысле. Математические методы просто более точны и в них используются более четкие формулировки и более широкий набор понятий, но в конечном счете они должны быть совместимы с обычными словесными рассуждениями, хотя, вероятно, и идут дальше их.

Глава 2

РОЛЬ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

2.1. БИОЛОГИЧЕСКАЯ ИЗМЕНЧИВОСТЬ И ВЕРОЯТНОСТЬ

«Если для вашего эксперимента нужна статистика, вам следовало бы провести его получше», — сказал Резерфорд. Хотя эти слова, возможно, и были справедливы для многих разделов физики в прошлом, к современной биологии и медицине это, безусловно, не относится. Даже в так называемых точных науках наличие «ошибок» измерений признавалось уже давно, однако большинство исследователей исходило из того, что путем усреднения результатов нескольких повторных измерений такие ошибки можно свести к пренебрежимо малой величине. В большинстве разделов физики и химии действительно можно отбрасывать не представляющие интереса отклонения и, формулируя те или иные гипотезы и прогнозы, исходить из того, что результаты любых измерений можно считать более или менее точными. Этим не отрицается тот факт, что в некоторых случаях, например при описании радиоактивного распада и броуновского движения, а также в квантовой теории, случайный характер самих физических явлений требует применения специальных методов исследования, основанных на теории вероятностей.

В биологии и медицине изменчивость выражена гораздо сильнее и имеет большее научное значение. В разд. 1.3 мы уже уделили некоторое внимание кривым веса человека. При повторных измерениях веса одного и того же человека, проводимых в одно и то же время, можно легко обнаружить небольшие колебания результатов, однако они не представляют особого интереса. Если же повторные измерения проводить через короткие промежутки времени, то можно обнаружить колебания вследствие прибавления или потери веса за счет приема пищи, дыхания, выделений и т. д. Все эти аспекты, безусловно, имеют важное значение с биологической точки зрения, однако по сравнению с более значительными изменениями, происходящими, скажем, за неделю или за месяц и связанными с общим процессом роста, такие кратковременные изменения могут считаться несущественными. Пойдя дальше и сравнив значения соответствующих численных показателей у различных индивидуумов, мы немедленно обнаружим изменчивость внутри популяции. Известно, что отдельные представи-

тели любого данного вида могут значительно отличаться друг от друга по весу или размерам тела, и обычно идея описания популяции средними показателями не встречает серьезных возражений. Вес и рост — настолько знакомые для большинства из нас показатели, что усредненные кривые роста или таблицы среднего веса для людей определенного возраста, пола и роста принимают за стандарты, позволяющие судить о степени отклонения от нормы в каждом конкретном случае.

Однако даже у таких простых показателей, как рост и вес, наблюдаются иногда очень большие колебания вследствие обычной естественной изменчивости. Автору известно об одном исследовании веса младенцев в течение первых десяти дней пребывания в родильном доме, которое проводилось для сравнения результатов кормления грудью и результатов искусственного кормления с учетом таких факторов, как вес ребенка при рождении, его пол, возраст матери и т. п. Кривая среднего веса для нескольких сотен нормальных детей, получавших искусственное питание, в течение всего периода исследования непрерывно поднималась вверх. Средний вес младенцев, вскармливаемых грудью, в первые два дня резко падал, как и ожидалось, а затем начинал быстро расти и уже через несколько дней совпадал с весом детей-искусственников. Можно было бы сказать, что это служит наглядной демонстрацией способности организма преодолевать первоначальную нехватку пищи и достигать устойчивой скорости роста. Однако примечательно то, что, хотя на основе кривых для средних значений можно попытаться сделать какие-то общие выводы, данные, записанные для отдельных младенцев, оказываются совершенно хаотичными: одни дети непрерывно прибавляли в весе, другие непрерывно теряли, а у остальных вес то возрастал, то снижался, т. е. наблюдались резкие колебания. При этом никакой очевидной связи между этими различными случаями и различными исследовавшимися факторами обнаружить не удалось. Упорядоченность и регулярность легко обнаруживаются лишь в средних значениях, взятых по большому числу индивидуумов. Поэтому при использовании общей кривой среднего веса в качестве стандарта для суждения о развитии отдельного новорожденного необходимо проявлять большую осторожность. Исключительно важно учитывать возможные отклонения, чтобы основная математическая модель определяла не только средний вес, которого следует ожидать при данном возрасте ребенка и при данном режиме питания, но и позволяла измерить имеющееся отклонение от нормы.

Применение статистических методов для получения выводов на основе данных об изменчивости рассматривается в разд. 2.2. Здесь же нас интересует лишь необходимость *измерения* измен-

чивости. В разд. 1.4 уже упоминалось о распространенном мнении, согласно которому многие биологические явления настолько сложны и разнообразны, что математический анализ в этой области неизбежно приведет к чрезмерным упрощениям, дающим ошибочные результаты. Однако истина состоит совсем в другом: если не предпринимать серьезных попыток разработать надлежащие математические методы, то это только уменьшит возможность точного описания биологических процессов.

Как хорошо известно, одним из самых плодотворных способов описания характера изменчивости является применение соответствующего закона распределения, который определяет вероятность того, что результат измерения какого-либо параметра индивидуума, выбранного случайным образом, будет иметь любое заданное значение или лежать в определенном интервале значений. Такие непрерывные параметры, как рост, вес и т. п., нередко удовлетворительно описываются кривой нормального, или гауссова, распределения (несмотря на то, что теоретически эта кривая лежит в интервале от $-\infty$ до $+\infty$)

$$y = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \quad (-\infty < x < \infty), \quad (2.1)$$

где μ — математическое ожидание, а σ — среднее квадратическое отклонение. Если такая кривая применяется для описания распределения людей по росту, то вероятность того, что рост данного человека находится в интервале от x_1 до x_2 , равна

$$\int_{x_1}^{x_2} y \, dx.$$

Нормальное распределение является одним из простейших с точки зрения математики. Кроме того, существует ряд теоретических оснований, позволяющих предполагать, что многие распределения, встречающиеся на практике, должны быть близки к нормальному, и это предположение действительно часто подтверждается. Этих соображений вполне достаточно для того, чтобы нормальное распределение заняло важное положение в теории вероятностей и математической статистике.

Для описания дискретных величин в тех случаях, когда имеется ограниченное число альтернативных наблюдений (например, таких, как число детей-альбиносов в семье данного состава), может оказаться пригодным биномиальное распределение. Если имеется n индивидуумов и вероятность того, что какой-либо из них обладает определенным признаком, равна p (независимо от других индивидуумов), то вероятность наблюдения r индивидуумов

с данным признаком имеет биномиальное распределение и равна

$$\frac{n!}{r!(n-r)!} p^r (1-p)^{n-r} \quad (0 \leq r \leq n). \quad (2.2)$$

Распределение числа радиоактивных частиц, испускаемых за данный промежуток времени некоторой большой массой радиоактивного вещества, числа дорожно-транспортных происшествий, происходящих за данный промежуток времени при определенных условиях, или числа лейкоцитов, наблюдаемых в одном квадрате гемоцитометра, лучше всего описывается законом Пуассона, согласно которому вероятность наблюдения r событий равна

$$e^{-m} m^r / r! \quad (0 \leq r < \infty), \quad (2.3)$$

где m — среднее значение случайной величины.

Мы привели всего три наиболее распространенных и наиболее простых распределения из числа встречающихся на практике, однако с их помощью можно охватить поразительно большое множество случаев естественной изменчивости в биологии и медицине, не обращаясь к более сложным описаниям. Некоторое представление о содержании и возможностях теории распределений можно почерпнуть из книг по теории вероятностей (см., например, книгу Феллера [22]) или математической статистике (см., например, книгу Кендалла и Стюарта [38]).

Применение распределений вероятностей — отнюдь не новый способ описания биологической изменчивости. Кетле, работавший вначале в области астрономии и метеорологии, был, по-видимому, первым, кто применил нормальное распределение для описания биологического материала (он ввел его при изучении распределения людей по росту, о чем уже говорилось выше). Позже Фрэнсис Гальтон [27] широко применял кривую нормального распределения при статистическом исследовании наследственности, и она сыграла фундаментальную роль в глубокой работе Карла Пирсона по вопросам биометрии, написанной в конце прошлого века. С тех пор различные типы распределений начали применять в самых разнообразных областях биологии — в молекулярной биологии, таксономии, экологии, генетике, психологии и т. д.

Как с исторической, так и с логической точки зрения распределения вероятностей представляют собой просто более совершенные варианты математических моделей, рассмотренных в гл. 1. Они позволяют свести огромное многообразие наблюдений к одному закону, который можно охарактеризовать очень небольшим числом параметров: двумя в случае нормального распределения, одним-единственным в случае пуассоновского распределения и т. д. Это дает возможность более точно описать изменяющиеся явления и облегчает их понимание. По существу это то, что Р. Фишер

называл «редукцией данных». Численную информацию можно точно записывать, хранить, передавать и обсуждать. Затем эти описания можно преобразовать к такому виду, который и принято рассматривать как собственно математическую модель, т. е. аналог реальной действительности, наделенный такой структурой, которая позволяет применять обычные методы научного исследования. Это означает, что с помощью модели выводятся следствия и прогнозы, справедливость ее проверяется по соответствующим наблюдениям и в случае необходимости в модель вносятся изменения. Проверка моделей связана со статистическими методами, которые будут рассматриваться в следующем разделе. Более детальное изложение всего процесса научного исследования мы отложим до гл. 3.

Разумеется, математические модели (даже вероятностные) часто не удовлетворяют биологов, которые считают их чрезмерно упрощенными. Для специалиста в области экологии современные вероятностные модели конкуренции между видами (см. гл. 8) вполне могут показаться слишком примитивными. Однако все дело в том, что такой подход позволяет более уверенно охватить все многообразие и сложность природы. При использовании современных математических и статистических методов и вычислительной техники метод построения математических моделей может быть развит до такой степени, что появится возможность сделать для биологии то, что математическая физика сделала для физики.

2.2. ПРИМЕНЕНИЕ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

В предыдущем разделе мы показали, что самые разнообразные типы изменчивости, встречающиеся в природе, часто удается описать с помощью одного надлежащим образом выбранного распределения вероятностей. Оно, разумеется, не позволяет определить конкретное значение какого-либо данного параметра или результата наблюдения, но в то же время точно описывает ожидаемую общую картину. Говоря об изменчивости, случайных колебаниях и распределениях, мы употребляем слова «вероятностный» и «статистический» почти в одинаковом смысле, отдавая некоторое предпочтение первому термину. На данном этапе удобно провести определенное различие между ними, хотя в мнениях разных авторов на этот счет нет полного единства.

Распределения вероятностей и связанные с ними теоретические построения представляют собой по существу математические конструкции, хотя они и выводятся из повседневных представлений о случае и шансах и должны приводить к результатам, которые можно истолковать в практическом смысле. Для построения правильной математической модели некоторого явления, характе-

ризующегося сильными случайными колебаниями, вводятся вероятностные идеи, и для раскрытия смысла модели используется теория вероятностей. Чтобы модель была удовлетворительной хотя бы в некоторых отношениях, она должна обеспечить возможность практической проверки получаемых с ее помощью результатов. Это означает, что математические результаты должны в определенной мере соответствовать данным, которые были или могли быть получены. Задачей математической статистики и является изучение соответствия между теоретическими моделями и реальной действительностью и проверка их адекватности. Коротко говоря, в то время как теория вероятностей относится к области дедуктивной логики, математическая статистика охватывает помимо всего прочего и область индуктивной логики. Роль этих различных дисциплин и их значение для всего процесса научного исследования подробно рассматриваются в гл. 3.

В разд. 2.1 мы уже указывали, каким образом с помощью понятий теории вероятностей может быть произведена редукция данных до объема, поддающегося обработке. Для этого сосредотачивают внимание лишь на наиболее существенных аспектах реального явления, а остальными факторами не пренебрегают, но и не дифференцируют их, учитывая их влияние в совокупности (независимо от того, какие распределения вероятностей используются). Успех такого упрощающего допущения зависит от того, позволит ли построенная статистическая модель получить полезные результаты, несмотря на то что в нее не входят в явном виде факторы, которые в определенном смысле несомненно имеют важное значение. Поэтому крайне важно еще на начальном этапе проверить, достаточно ли хорошо согласуется принятая модель с полученными данными. Вследствие существования естественной изменчивости — основной причины, по которой приходится применять теорию вероятностей, — на практике не могут быть в точности воспроизведены прогнозы или результаты, вытекающие из модели. Но ее можно считать вполне удовлетворительной, если соответствие окажется достаточно хорошим при достаточно большой выборке данных. Совершенно ясно, что при оценке модели существенно важен выбор критериев, на основании которых решается вопрос о том, является ли соответствие достаточно хорошим и число наблюдений достаточно большим.

Для статистической проверки вероятностной модели важнейшую роль играет понятие статистического *критерия значимости*. Конкретный способ выполнения проверок в некоторой степени зависит от того, какая из нескольких моделей статистического вывода выбирается. Не следует удивляться тому, что хорошо обоснованные разные модели нередко могут приводить к одинаковым практическим методам и результатам. Этот вопрос подни-

мается снова и подробно рассматривается в разд. 3.3. Здесь мы остановимся лишь на наиболее широко используемой так называемой «частотной» интерпретации вероятности. В соответствии с этим подходом необходимо проводить четкое различие между «вероятностью» применительно к результатам экспериментов, которые были или могли быть проведены, и «вероятностью», характеризующей степень нашей уверенности в некоторой гипотезе. По возможности мы употребляем слово «вероятность» в первом значении, что согласуется с определением вероятностных моделей, описывающих реальные процессы, которое давалось в разд. 2.1. Таким образом, «вероятность» характеризуется частотой, которую, во всяком случае в принципе, можно наблюдать.

Построенная первоначально математическая модель представляет собой так называемую *нулевую гипотезу*. Нулевая гипотеза может быть довольно простой, например допущение о том, что распределение числа детей в семье по полу является биномиальным с $p = 0,5$ (т. е. отношение полов равно единице). Чтобы проверить справедливость этой гипотезы для данной семьи, нужно изучить фактические данные; при этом может оказаться, что, скажем, в семье пять детей и все они девочки. Оценим теперь вероятность фактического события, которое произошло, и вероятность любого другого столь же вероятного или более редкого события, которое может произойти. Несколько необычная форма данной методики вызывается тем, что какой-либо конкретный экспериментальный результат может иметь очень малую вероятность, особенно если число различных наблюдаемых событий очень велико. (Строго говоря, для непрерывных случайных величин вероятность появления любого заданного значения равна нулю.) Поэтому нас больше интересует, принадлежит ли наблюдаемый результат к классу необычных событий, резко отличающихся от наиболее вероятного результата. В приведенном выше примере наиболее вероятным является наличие в семье, насчитывающей пятеро детей, двух или трех девочек; менее вероятно появление одной или четырех девочек; наименее вероятно отсутствие девочек или наличие пяти девочек. Наблюденное значение, равное пяти, принадлежит к последнему классу «0 или 5», вероятность P которого равна $1/16$, или 6,25%.

Если вероятность P наблюдаемого события велика, например не менее 30%, то это означает, что оно является весьма распространенным. Если же значение P достаточно мало, например менее 5%, то наблюдаемое событие принадлежит к классу довольно редких. В этом случае можно вообще отвергнуть нулевую гипотезу, вместо того чтобы придерживаться ее и считать, что произошло маловероятное событие. Выводы, к которым мы приходим при таком подходе, в значительной мере зависят от того, где про-

водится граница между приемлемыми и неприемлемыми значениями P , т. е. от *уровня значимости*. Разумеется, необходимо отдавать себе отчет в том, что применение такого критерия для проверки значимости еще не гарантирует полностью отсутствия ошибки. Даже если нулевая гипотеза справедлива, то в доле случаев, равной P , при использовании данного критерия она будет отвергнута. Если значение P очень велико, то довольно часто истинная гипотеза будет отвергаться, что приведет к ложным выводам. Если же значение P очень мало, то мы лишимся возможности отвергнуть ложную гипотезу и развить новые идеи. Широко используется 5%-ный уровень значимости, который, как показывает опыт, в общем случае вполне пригоден. В некоторых особых случаях могут ставиться и другие требования. Ясно, что более осторожный исследователь (скажем, тот, кто занимается испытанием сильнодействующих лекарственных препаратов) будет считать желательным более низкий уровень значимости — возможно, 1%-ный или даже меньше.

Возвращаясь к приведенному выше примеру, можно видеть, что наблюдаемое значение $P = 6,25\%$ как раз и не является статистически значимым при 5%-ном уровне. Принимая такой уровень значимости, мы допускаем, что первоначальная гипотеза ($p = 0,5$) все еще приемлема и вероятность того, что следующий ребенок в семье будет мальчик, по-прежнему равна $1/2$. Однако если в семье появилось шесть девочек, то значение P становится равным примерно 3%, и следует отвергнуть нулевую гипотезу, согласно которой соотношение полов равно единице.

Проверка значимости рассматривается здесь лишь для того, чтобы на простом конкретном примере проиллюстрировать характер связанных с ней рассуждений. Более точное изложение теоретических вопросов, связанных с проверкой значимости, и описание разнообразных критериев, используемых для проверки в различных случаях, читатель найдет в любом учебнике по математической статистике. Основная мысль, которую мы хотели подчеркнуть, заключается в том, что существуют численные методы, позволяющие определить соответствие конкретной математической модели эмпирическим данным. Если модель не согласуется с эмпирическими данными, то необходимо либо переделать ее каким-то образом (возможно, изменяя значения входящих в нее параметров), либо, в более серьезном случае, полностью отказаться от нее и перейти к некоторому другому описанию изучаемого явления.

Если рассматривается семья, в которой все шестеро детей — девочки, то, возможно, придется отклонить исходную гипотезу, согласно которой $p = 0,5$, хотя, по-видимому, потребуется сохранить подразумеваемое допущение, что при любом значении p

распределение будет биномиальным. Какое же значение (или значения) p приемлемо в данном случае? Этот вопрос приводит нас к проблеме нахождения *оценок* — другой важной области математической статистики. Существует несколько способов решения этой задачи. Один из них состоит в нахождении некоторой области значений параметра (ее называют *доверительным интервалом*), в пределах которой с заданной вероятностью находится истинное значение. При частотной интерпретации вероятности это правило получает следующую формулировку: если при многократном повторении эксперимента вычисляется 95%-ный доверительный интервал, то вычисленный интервал будет включать истинное значение в 19 случаях из 20. (Заметим, что при частотной интерпретации вероятности только интервал может иметь некоторое распределение. Истинным значением параметра является некоторая *фиксированная*, хотя и неизвестная постоянная.) Можно показать, что для семьи, где все шестеро детей девочки, 95%-ным доверительным интервалом является интервал от 0,61 до 1,00 (пределы его вычислены приближенно). Этот интервал не включает значения $p = 0,5$ первоначальной нулевой гипотезы, соответствующего данным, которые статистически значимы для этого значения при 5%-ном уровне.

Разумеется, в этом примере выборка очень мала и соответственно велик доверительный интервал. Однако при достаточно больших выборках можно ожидать, что интервалы будут значительно меньше, и в других случаях они, по-видимому, будут более приемлемы. Как известно, нередко средние значения имеют распределения, близкие к нормальному, даже если отдельные значения параметров и не распределены по нормальному закону. Это означает, что в качестве показателя достоверности среднего значения можно использовать среднее квадратическое отклонение нормального распределения. Обычно среднее квадратическое отклонение среднего значения или другой величины (например, оценки неизвестного параметра, полученной методом максимального правдоподобия), имеющей приближенно нормальное распределение, предпочтительнее называть *средней квадратической ошибкой*. Так, если расчет покажет, что при употреблении данного лекарственного препарата при определенном заболевании выздоравливает 25% больных, а средняя квадратическая ошибка σ составляет 2%, то можно ожидать, что при повторных экспериментальных проверках выборок такого же объема в 95% случаев истинный процент выздоравливающих будет лежать в интервале, равном этому среднему значению $\pm 2\sigma$, т. е. в интервале $25 \pm 4\%$. Проще говоря, мы будем на 95% уверены в том, что истинный процент выздоравливающих находится в интервале между 21 и 29%.

Со всеми этими рассуждениями, которые довольно хорошо известны (во всяком случае, на элементарном уровне) большинству научных работников в области биологии и медицины, связан следующий важный вопрос: насколько полезны статистические методы на практике и насколько близки они к тем проблемам, которые в действительности старается решить ученый-практик? Более подробно о логике научных исследований говорится в гл. 3. На данном этапе мы только заметим, что большинство научных исследований (если не все) начинается с принятия гипотезы, сформулированной на основе существующих знаний или предположений, а затем эта гипотеза проверяется с помощью соответствующих экспериментальных данных. Если согласие оказывается удовлетворительным, то гипотеза сохраняется и развивается дальше, в противном случае она отвергается или изменяется. Поворотным моментом в развитии и применении такого статистического подхода, по-видимому, явилась вышедшая в 1925 г. первым изданием книга Р. Фишера «Статистический метод исследований» [23], которая с тех пор выдержала много изданий. Критерий Фишера, по сути дела, переводит обычную практику интуитивных оценок результатов исследований на количественную основу. Для очень широкого круга приложений такое уточнение понятий оказалось крайне полезным. Если в некоторых частных случаях статистическая проверка гипотез и встречает определенные затруднения, то, во всяком случае, она дает хорошую основу для строгого вывода более адекватных методов.

2.3. ПЛАНИРОВАНИЕ ЭКСПЕРИМЕНТОВ

Статистический подход, рассмотренный в предыдущем разделе, можно использовать в любом случае, когда требуется выяснить, отличается ли реально или, во всяком случае, статистически значимо некоторая полученная в эксперименте величина — среднее значение, коэффициент корреляции и т. п. — от ее теоретического значения. Нас, скажем, может интересовать, достаточно ли отличается от 0,5 наблюдаемая частота рекомбинаций между двумя генетическими локусами в потомстве от возвратного скрещивания для того, чтобы можно было заключить, что данные локусы лежат в одной и той же хромосоме. Или, например, требуется выяснить, значимо ли статистически наблюдаемое различие между числом образующихся мужских и женских гамет. Далее, может возникнуть необходимость в определении частоты случаев выздоровления после какого-то заболевания — либо при испытании того или иного препарата, либо при сравнении эффективности двух препаратов. Замечательная особенность такого статистического анализа состоит в том, что все виды неизбежной естествен-

ной изменчивости, составляющей как бы «фон», на котором проявляется изменчивость, связанная с исследуемым фактором, учитываются в комплексе путем использования соответствующего распределения вероятностей. Если фоновая изменчивость очень велика, то для получения окончательных результатов может потребоваться очень большое число наблюдений, а когда она сравнительно мала, результат будет получен значительно быстрее.

Если какой-либо эффект вызывается очень большим числом различных факторов, то вполне возможно, что фоновая изменчивость будет весьма велика. В таких случаях целесообразно попытаться выделить некоторые из этих факторов, даже если их невозможно полностью контролировать или исключить. Часто оказывается возможным разбить общую изменчивость на отдельные компоненты, из которых один соответствует исследуемому фактору, несколько других — другим воздействиям, допускающим возможность раздельной оценки, и последний — остальным воздействиям, раздельная оценка которых невозможна. Поскольку влияние последней группы факторов, безусловно, будет слабее, чем влияние исследуемого фактора, то это обеспечивает более точную статистическую проверку. Искусство располагать наблюдения в определенном порядке или проводить специально спланированные проверки с целью полного использования возможностей этих методов и составляет содержание предмета «планирование эксперимента». Детальное изложение существующих методов планирования эксперимента можно найти в литературе (см., например, [14]). В настоящем разделе мы осветим лишь некоторые основные преимущества сознательного и продуманного планирования эксперимента.

Допустим, например, что требуется сравнить болеутоляющее действие двух различных лекарственных препаратов А и В. Пусть подобрано 16 больных и принято решение разделить их случайным образом (во избежание какой-либо сознательно или непроизвольно вносимой систематической ошибки) на две группы, по 8 больных в каждой. Одна группа получает препарат А, а другая — препарат В. Затем измеряют время, в течение которого каждый из больных испытывает облегчение, и сравнивают средние значения по обеим группам. Если среднее время для препарата А значительно превышает среднее время для препарата В, то можно сделать вывод, что первый препарат более эффективен. (В данном случае несущественно, какой статистический критерий используется. Поскольку рассматривается небольшое число объектов, это может быть один из критериев Стьюдента.) Известно, что больные по-разному реагируют на один и тот же лекарственный препарат, поэтому продолжительность периода облегчения обычно сильно варьирует, что значительно понижает точность сравне-

ния этих двух препаратов. Однако в данном эксперименте различия между больными не представляют для нас особого интереса, и этот источник погрешности можно исключить следующим образом. Вместо того чтобы делить больных на две группы, проверяют на каждом из них оба препарата, назначая их последовательно через достаточно большие промежутки времени (чтобы избежать взаимодействия) и в случайном порядке (или, возможно, в одном порядке для одной половины больных и в другом порядке для другой). Теперь для каждого больного определяют *относительное* преимущество препарата А перед препаратом В, для чего вычисляют суммарную продолжительность периода облегчения для каждого из них и находят разность этих двух величин. Таким образом получают 16 чисел, характеризующих относительное преимущество одного препарата перед другим, что позволяет проверить, значимо ли отличается от нуля их среднее значение. Положительная разность $t_A - t_B$ свидетельствует о статистически значимом преимуществе препарата А, отрицательная — об обратном соотношении. Рассматривая показатели относительного преимущества, мы исключаем влияние реакции отдельных больных и в общем случае добиваемся более эффективного сравнения этих двух лекарств.

Такая простая проверка методом попарного сравнения представляет собой простейший план эксперимента, имеющий целью извлечь максимальное количество информации из данного числа наблюдений. Заметим, что этот план имеет и свои дополнительные особенности, так как требует особого внимания к ряду практических вопросов, например к тому, чтобы препараты назначались в случайном порядке (во избежание нежелательной систематической ошибки) и через достаточно большие промежутки времени (для исключения эффектов взаимодействия); однако здесь мы не можем детально рассматривать эти вопросы.

Мы показали, каким образом при проверке методом попарного сравнения можно контролировать или исключать из рассмотрения какой-либо один важный и явный источник изменчивости. В более общем случае могут быть спланированы *факторные* эксперименты, с помощью которых можно определить вклад каждого из нескольких факторов в общую изменчивость. Некоторые из этих факторов могут представлять особый интерес, тогда как другие имеют второстепенное значение. Идея и практическое применение этого нового подхода, принадлежащего главным образом Р. Фишеру, получили широкое распространение после появления его книги «Планирование экспериментов», вышедшей первым изданием в 1935 г. Большинство фундаментальных работ в области планирования эксперимента было посвящено сельскохозяйственным приложениям. Допустим, требуется сравнить среднюю урожай-

ность нескольких сортов пшеницы при применении различных удобрений в различной концентрации, учитывая при этом колебания в плодородии почвы на достаточно больших участках земли, которые можно разбить на делянки подходящих размеров. Для начала можно попытаться составить план эксперимента, в котором будут рассматриваться все возможные комбинации значений, или уровней, различных факторов. Так, если имеются четыре сорта пшеницы и три различных вида удобрений, применяемых в трех различных концентрациях, то общее число комбинаций условий будет равно 36. Таким образом, исходное число делянок в одном блоке факторного эксперимента будет равно 36 — по одной делянке на каждую комбинацию условий. Вследствие возможного колебания в плодородии почвы от одного блока к другому может оказаться целесообразным иметь не менее двух полных блоков.

Применение факторного плана вместо классической схемы, согласно которой каждый раз изменяется только один фактор, имеет ряд серьезных и даже несколько неожиданных преимуществ. Прежде всего в этом случае становится значительно более полной картина влияния каждого фактора, поскольку оно изучается в самых различных условиях (вследствие одновременного изменения других факторов). Во-вторых, большое число комбинаций факторов, используемых в эксперименте, облегчает предсказание результатов, которые могут быть достигнуты при определенной комбинации условий. В-третьих, если эффекты, вызываемые каждым фактором, включаемым в эксперимент, статистически независимы, то о каждом факторе можно получить не меньше информации, чем если бы в процессе эксперимента изменялся только один этот фактор, а остальные оставались постоянными. В-четвертых, если (как это часто бывает) различные факторы не являются независимыми, а вызывают эффекты, которые в большей или меньшей степени коррелированы, то в этом случае только факторный эксперимент может дать информацию о характере этих взаимодействий. При наличии нескольких взаимосвязанных существенных факторов обойтись без постановки факторного эксперимента невозможно. Для ряда часто встречающихся специальных задач разработано большое число стандартных планов такого типа.

Согласно некоторым из этих простейших планов, эксперимент проводят на нескольких блоках и внутри каждого из них на отдельных делянках проверяют влияние всех уровней какого-то одного фактора. При правильном планировании получают рандомизированный блочный план. В сельскохозяйственных задачах блоками могут служить участки земли на различных полях, а уровнями одного фактора — ступенчатая последовательность концентраций удобрений или просто различные сорта пшеницы.

В лабораторном эксперименте, в котором, скажем, проверяется влияние различных рационов питания на крыс, рационы питания будут испытываемыми условиями, а крысы — отдельными экспериментальными единицами (соответствующими деланкам в сельскохозяйственном эксперименте). Если бы мы могли использовать в эксперименте животных отдельных пометов, подвергая каждому воздействию по одному животному из каждого помета, то каждый помет можно было бы рассматривать как отдельный блок.

В рассмотренной выше простой проверке методом попарного сравнения также можно было бы применить рандомизированный блочный план; тогда каждого больного можно было бы рассматривать как отдельный блок, а лекарственные препараты — как условия эксперимента.

Хотя иногда бывает трудно перенести планы экспериментов, разработанные для одной области, особенно для сельского хозяйства, в совершенно другую область, лежащая в их основе логическая схема часто оказывается весьма сходной. Поэтому целесообразно тщательно обдумать возможность того, чтобы при надлежащей интерпретации элементов какого-либо определенного плана эксперимента можно было обеспечить его успешное применение в задачах совершенно иного характера. Это иллюстрирует большие возможности математических методов планирования эксперимента. В основе планирования должна, разумеется, лежать некоторая исходная математическая модель. Опишем самую простую из них, которая в том или ином варианте используется наиболее широко. Пусть требуется исследовать влияние только двух факторов A и B . Допустим, что наблюдаемое на некоторой экспериментальной единице влияние i -го уровня фактора A и j -го уровня фактора B можно записать в виде

$$y_{ij} = m + a_i + b_j + z_{ij}, \quad (2.4)$$

где y_{ij} — наблюдаемая величина, m — общее среднее, a_i и b_j — относительные вклады этих двух факторов при заданных уровнях каждого из них, а z_{ij} — случайное изменение, налагаемое на основную линейную аддитивную схему. Кроме того, часто принимается, что все величины имеют одно и то же нормальное распределение и независимы друг от друга. Эти ограничения весьма серьезные, однако часто приняты в качестве первого приближения вполне оправданно. Так, если влияние этих факторов мало, то заметную величину будут иметь только линейные члены и возможными членами второго порядка можно пренебречь. При независимости факторов формула (2.4) вполне удовлетворительна. Но если они взаимодействуют друг с другом, то следует включить в нее дополнительные члены c_{ij} , учитывающие это взаимодействие. Можно, однако, выполнить проверку значимости на ос-

нове формулы (2.4), чтобы убедиться, нужны ли члены, характеризующие взаимодействие. Кроме того, если случайные величины z_{ij} не распределены по нормальному закону, то можно использовать какую-либо функцию эмпирических результатов (например, квадратные корни или логарифмы), для которой сохраняется нормальный закон распределения.

На основе элементарной формулы (2.4) легко построить модели, учитывающие множество факторов, блоков, взаимодействий и других усложнений, вызываемых практической необходимостью в каждом данном эксперименте. Дело в том, что в очень многих случаях необходимые вычисления относительно просты и выполняются непосредственно. Обычно приходится производить повторяющиеся вычисления сумм и сумм квадратов данных, выбранных соответствующим образом. Результаты представляют в виде таблицы дисперсионного анализа, с помощью которой можно установить значимость всех различных факторов, влияющих на результаты эксперимента.

Одним из современных вариантов планирования экспериментов, который следует рассмотреть особо, является последовательностная схема эксперимента. В эксперименте стандартного типа необходимо заранее решить, сколько наблюдений нужно набрать. Если после анализа обнаружится, что число наблюдений слишком мало, то нужно попытаться продолжить эксперимент, однако может оказаться, что на данном этапе сделать это трудно или невозможно. Если же выяснится, что получено значительно больше наблюдений, чем необходимо для достижения требуемой точности, то будут потеряны время и деньги. В медицинских задачах это имеет особенно существенное значение. Ни один врач не заинтересован в том, чтобы эксперимент длился дольше, чем это строго необходимо, так как его цель — дать своим больным наилучший из существующих препаратов, как только он пройдет клинические испытания. Таким образом, в медицине выбор и планирование эксперимента теснейшим образом связаны с этическими соображениями. Последовательностная схема предусматривает проведение эксперимента отдельными сериями. Оценка результатов производится на каждом этапе, с тем чтобы немедленно можно было решить, применять препарат А, препарат В или же продолжать эксперимент, поскольку окончательного вывода сделать еще нельзя. При такой схеме эксперимента длительность его будет минимальна и он закончится значительно раньше, чем в любом другом случае. Кроме того, в медицине часто бывает очень трудно или даже вообще невозможно провести обычную экспериментальную проверку, так как после нескольких неудачных исходов, которые могут закончиться смертью больного, начинаются острые споры о том, следует ли продолжать экспери-

мент вообще. Последовательная схема означает, что заранее можно тщательно и спокойно рассмотреть различные линии поведения, обуславливаемые различными исходами эксперимента. При этом значительно легче выбрать наилучшие решения непосредственно в ходе эксперимента и совместить требования этики со статистической эффективностью. Более детально методы последовательного анализа в медицине рассмотрены в книге Эрмитажа [1].

Недостаток места не позволяет продолжать здесь изложение теории планирования эксперимента, тем более что этому предмету посвящена уже огромная литература¹. Наша основная цель состояла в том, чтобы показать, каким образом с помощью простой математической модели процесса, на который одновременно воздействуют несколько факторов (носящих, возможно, в значительной мере вероятностный характер), можно с достаточной точностью выяснить степень влияния каждого из этих факторов в отдельности. Это позволяет применять для проверки значимости влияния крайне изменчивых и одновременно действующих факторов простые статистические критерии, описанные в разд. 2.2. А именно такой подход и необходим для исследования всего огромного многообразия явлений, встречающихся в биологии и медицине.

2.4. СЛУЧАЙНЫЕ ПРОЦЕССЫ

В разд. 1.3 уже подчеркивалось значение динамических процессов изменения и развития для любых биологических исследований. Мы видели, что некоторые важные свойства растущих популяций можно выделить из очень сложной совокупности реальных свойств и описать с помощью относительно простых дифференциальных уравнений. В приведенных примерах эти уравнения легко разрешались и давали экспоненциальные и логистические кривые. Мы уже указывали, что, хотя для построения математических моделей требуется вводить значительные упрощения, этот метод позволяет глубже проникнуть в суть биологических процессов и создает основу для дальнейших исследований. Серьезный недостаток детерминистских моделей, рассмотренных в разд. 1.3, заключается в том, что в них игнорируется естественная биологическая изменчивость (в отличие от изменений, возникающих за длительный период) или, во всяком случае, предполагается, что она пренебрежимо мала. Когда число объектов достаточно велико, детерминистские модели могут оказаться вполне удовлетворительными (в пределах налагаемых ограни-

¹ См., например, Х и к с Ч., Основные принципы планирования эксперимента, изд-во «Мир», 1967.— *Прим. перев.*

чений). Однако при рассмотрении многих важных ситуаций, например эпидемий или взаимодействия между видами, необходимо учесть, что число индивидуумов в определенные моменты времени может стать весьма небольшим. И в этих случаях статистические колебания обязательно будут оказывать значительное влияние на протекание всего процесса.

Для построения демографических прогнозов структуры населения детерминистский анализ, основанный на скоростях роста численности в различных возрастных группах мужчин и женщин, может оказаться достаточно точным. Возможность предсказывать численность населения на несколько десятилетий вперед с точностью до нескольких сотен единиц в большинстве случаев вполне удовлетворит экономистов и социологов. Допустим теперь, что мы пытаемся провести количественное исследование десяти пар (самцов и самок) какой-нибудь редкой птицы. Совершенно ясно, что случайность играет важнейшую роль в отношении таких факторов, как плодовитость, размер выводка, наличие пищи, климатические условия, конкуренция, болезни, несчастные случаи и гибель. При любом описании возможных изменений, происходящих в такой небольшой популяции, необходимо учитывать, что процесс может пойти в различных направлениях. Это не означает, что точное математическое описание таких ситуаций невозможно. Отсюда лишь следует, что для описания общей картины возможных событий и их вероятностей необходима стохастическая модель¹. Иными словами, нужно исходить из того, что в каждый данный момент времени существует не единственное значение размера популяции (как предполагалось в детерминистской модели), а некоторое распределение вероятностей этих значений. Таким образом, в отличие от рассмотренного ранее детерминистского процесса в данном случае мы имеем *вероятностный (стохастический, или случайный) процесс*.

Аналогичные соображения применимы и к описанию «эпидемии» в небольшой группе, например в семье, где имеется четверо детей. В такой маленькой группе не представляет большого интереса даже среднее число случаев через определенный промежуток времени. Но если бы мы смогли вычислить вероятность сроков заболевания каждого ребенка, учитывая особенности

¹ Здесь необходимо сделать одно уточнение. Вероятностное описание процессов или событий действительно позволяет разобраться в более тонких деталях и изучать их более индивидуализированно, чем при рассмотрении средних. Однако содержательность всякого утверждения типа «вероятность события A равна p » в такой же степени зависит от массовости события A , что и утверждение «математическое ожидание величины x равно m ». Всякая вероятностная модель предполагает принципиальную возможность многократного повторения независимого эксперимента и не может относиться к уникальным событиям. — *Прим. ред.*

участия в эпидемическом процессе восприимчивых индивидуумов, носителей инфекции и тех, кто был изолирован или выздоровел, а также принимая во внимание существование латентного периода и различную продолжительность заразного периода, то мы были бы на пути к такому описанию всего процесса, которое бы имело определенный биологический смысл. Такая модель не только послужила бы средством исследования эпидемии, но и создала бы прочную основу для вероятностного предсказания возможности появления новых случаев. Более того, теории распространения эпидемий в популяции удалось бы более надежно обосновать отдельными клиническими и эпидемиологическими наблюдениями. Гл. 9 специально посвящена математической теории эпидемий, а в гл. 8 проводится более подробное обсуждение некоторых простых задач, связанных с ростом и развитием популяции. В данном разделе мы лишь в общих чертах покажем, с чем связана попытка описать весь процесс с помощью распределений вероятностей.

Рассмотрим кратко стохастический аналог упрощенного процесса роста популяции с учетом только размножения. При детерминистском подходе мы считали бы, что существует определенная скорость размножения β , такая, что численность популяции n за время dt увеличивается на $dn = \beta n dt$. Это приводит к экспоненциальному закону

$$n = ae^{\beta t} \quad (2.5)$$

(где a — численность популяции при $t = 0$), уже встречавшемуся нам в формуле (1.2). При включении вероятностных элементов в эту ситуацию возможны различные варианты. Простейшим является допущение о том, что вероятность появления одного потомка у данного индивидуума в интервале времени dt равна βdt . Таким образом, вероятность появления одного нового индивидуума в целой популяции за время dt равна $\beta n dt$. Теперь мы можем оперировать непосредственно с фактическим дискретным размером популяции и обозначим через $p_n(t)$ вероятность того, что в момент t в популяции имеется ровно n индивидуумов. Размер популяции в момент t можно связать с размером популяции в момент $t + dt$ с помощью довольно элементарных рассуждений. Так, если число индивидуумов в момент $t + dt$ равно n , то это означает, что либо в момент t было n индивидуумов и за время dt это число не изменилось, либо в момент t их было $(n - 1)$ и за время dt появился еще один. Комбинируя эти вероятности, можно записать соотношение

$$p_n(t + dt) = p_{n-1}(t) \beta (n - 1) dt + p_n(t) (1 - \beta n dt),$$

откуда путем перестановки членов и деления на dt получаем дифференциальное уравнение

$$\frac{dp_n}{dt} = \beta(n-1)p_{n-1} - \beta np_n. \quad (2.6)$$

Это уравнение справедливо при $n > a$. Легко показать, что соответствующее уравнение для $n = a$ несколько проще и имеет вид

$$\frac{dp_a}{dt} = -\beta ap_a, \quad (2.7)$$

так как в том случае, когда процесс начинается при значении $n = a$, отсутствует член, содержащий p_{a-1} .

Системы дифференциально-разностных уравнений, аналогичных уравнению (2.7), которые можно рассматривать как *уравнения движения* для случайного процесса, иногда довольно трудно разрешить в общем виде. Некоторые из методов решения таких систем будут проиллюстрированы более детально в гл. 8 и 9. Здесь же мы рассмотрим относительно элементарный подход. Наиболее простой способ состоит в непосредственном интегрировании уравнения (2.7) с учетом того обстоятельства, что $p_a(0) = 1$; это дает $p_a(t) = e^{-a\beta t}$. Затем подставляем полученный результат в уравнение для $n = a + 1$, интегрируем, используя начальное условие $p_{a+1}(0) = 0$, и находим, что $p_{a+1}(t) = ae^{-(a+1)\beta t}(e^{\beta t} - 1)$. В свою очередь этот результат подставляется в последующее уравнение, и весь процесс повторяется. После вычисления нескольких последовательных членов легко записать результат в общем виде:

$$p_n(t) = \binom{n-1}{a-1} e^{-n\beta t} (e^{\beta t} - 1)^{n-a}, \quad n \geq a. \quad (2.8)$$

Трудность состоит в том, что при таком способе алгебраические выкладки становятся все сложнее и сложнее. Лучше поступать следующим образом.

Поскольку для интегрирования уравнения (2.6) необходимо умножить каждый член на интегрирующий множитель $e^{n\beta t}$, удобно ввести обозначение $q_n = e^{n\beta t} p_n$; тогда, заменив уравнение (2.6) уравнением

$$\frac{d}{dt}(e^{\beta n t} p_n) = \beta(n-1)e^{\beta t}(e^{\beta(n-1)t} p_{n-1}),$$

можно записать его в виде

$$\frac{dq_n}{dt} = \beta(n-1)e^{\beta t} q_{n-1}.$$

В целях дальнейшего упрощения заменим переменную t переменной T , где

$$dT = \beta e^{\beta t} dt. \tag{2.9}$$

Теперь имеем

$$\left. \begin{aligned} \frac{dq_n}{dT} &= (n-1) q_{n-1}, \quad n > a, \\ \frac{dq_a}{dT} &= 0. \end{aligned} \right\} \tag{2.10}$$

Интегрируя уравнение (2.9) и выбирая постоянную интегрирования из условия $T = 0$ при $t = 0$, получаем формулу для преобразования переменных в явном виде:

$$T = e^{\beta t} - 1. \tag{2.11}$$

Теперь очень легко проинтегрировать последовательно систему уравнений (2.10), используя начальные условия: $q_a(0) = 1$ и $q_n(0) = 0$ при $n > 0$. Сразу же получаем

$$q_a = 1, \quad q_{a+1} = aT, \quad q_{a+2} = a(a+1)T^2/2! \text{ и т. д.}$$

Очевидно, что

$$q_n = \frac{a(a-1)\dots(n-1)}{(n-a)!} T^{n-a} = \binom{n-1}{a-1} T^{n-a}, \quad n > a. \tag{2.12}$$

Подставляя соотношение (2.11) в уравнение (2.12) и учитывая, что $p_n = e^{-n\beta t} q_n$, непосредственно получаем общее решение (2.8).

Выражение (2.8) определяет фактическое распределение вероятностей для любого данного момента времени, заменяющее при вероятностном описании то единственное значение $n = ae^{\beta t}$, которое рассматривалось в детерминистской модели. Тем, кто знаком с биномиальным распределением, известно, что выражение (2.8) является частным случаем этого распределения. С помощью стандартных выражений можно найти математическое ожидание и дисперсию:

$$m(t) = ae^{\beta t}, \tag{2.13}$$

$$\sigma^2(t) = ae^{\beta} (e^{\beta t} - 1). \tag{2.14}$$

Эти величины можно получить, хотя и с большим трудом, непосредственно из формулы (2.8). Кроме того, существуют более общие методы вычисления таких статистик, как математическое ожидание и дисперсия, без нахождения всего распределения. Эти методы иллюстрируются в гл. 8 и 9. Такой подход может оказаться крайне полезным в том случае, когда все распределение вычислить трудно или невозможно.

Легко заметить, что математическое ожидание (2.13) совпадает с детерминистским средним (2.5). Таким образом, при большом числе индивидуумов детерминистское описание будет удовлетворительно заменять любую стохастическую модель, в которой основное внимание уделяется нахождению средних значений¹. С другой стороны, когда число индивидуумов мало, например когда начальный размер популяции a составляет всего лишь несколько единиц, изменчивость, характеризующая дисперсии (2.14), может быть весьма значительной. Вычисление математических ожиданий часто используют в качестве первого шага при описании случайных процессов; это, очевидно, удобно, когда детерминистское среднее, которое вычислить легче, имеет такое же значение. Однако, как уже указывалось ранее, это бывает далеко не всегда, и на этот случай полагаться не следует.

Интерпретация выражения (2.8) требует известной осторожности. При детальном рассмотрении какой-либо определенной популяции, для которой подходит принятая модель, в каждый момент времени мы, разумеется, будем наблюдать только одно численное значение. График роста популяции обнаружит значительные колебания. Возникает вопрос, каким образом эти колебания связаны с распределением вероятностей. Смысл выражения (2.8) состоит в том, что если имеется некоторое большое число популяций и в начальный момент $t = 0$ численность каждой из них равна a , то доля этих популяций, имеющих в момент t численность n , теоретически равна $p_n(t)$ с математическим ожиданием $m(t)$ и дисперсией $\sigma^2(t)$. Кривая роста любой отдельно взятой популяции может значительно отклоняться от кривой математического ожидания, хотя последняя совместно с дисперсией может все же служить некоторым показателем изменчивости, характерной для данного процесса.

Мы рассмотрели некоторые математические детали стохастической модели простейшего процесса роста популяции (с учетом только размножения) для того, чтобы показать характер подхода, используемого для описания процессов со значительными элементами случайности. В более общем случае, когда учитывается гибель особей или какие-либо другие факторы, вызывающие уменьшение численности, появляется по крайней мере один новый аспект, а именно возможность вымирания популяции.

¹ См. примечание на стр. 48. Отметим лишь дополнительно, что совпадение результатов, полученных с помощью детерминистской модели, с математическим ожиданием распределения, полученного из вероятностной модели, определяется линейностью этих двух моделей. В противном случае такого совпадения может и не быть. — *Прим. ред.*

Так, простейшее детерминистское описание популяции, учитывающее как размножение, так и гибель, приводит к уравнению (1.1), где m — средняя скорость роста. Если скорость размножения превышает скорость гибели, то m — положительная величина, и это детерминистское уравнение показывает, что численность популяции экспоненциально растет. Но при стохастическом описании всегда существует некоторая вероятность вымирания популяции в результате случайной гибели достаточно большого числа особей. (Разумеется, при отрицательном m размер популяции, согласно детерминистскому описанию, постепенно уменьшается до нуля.) Вероятность вымирания является важной характеристикой процесса, которая не учитывается в детерминистской модели.

Как будет показано в гл. 9, где рассматривается теория эпидемий, бывают и другие ситуации, в которых стохастический подход приводит к совершенно иным результатам, нежели детерминистский. Например, детерминистская модель абсолютно непригодна для описания повторяющихся вспышек эпидемий, так как она предсказывает затухающие колебания, которые на практике не наблюдаются; в то же время стохастическая модель предсказывает как раз то, что наблюдается в действительности, а именно повторяющуюся последовательность незатухающих вспышек различной интенсивности.

Из этого краткого изложения очевидно, что теория случайных процессов играет важную роль при изучении биологических явлений. Ясно также, что для построения даже очень простых стохастических моделей может потребоваться непредвиденно большой объем математических выкладок. При попытке построить более реалистические модели положение еще более усложняется. Например, рассматривая выше процесс роста популяции с учетом только размножения, мы приняли очень удобное допущение о том, что вероятность появления потомка у некоторого отдельно взятого индивидуума за время dt равна βdt . Можно легко показать, что это означает, будто интервалы времени от момента рождения до момента размножения имеют экспоненциальное распределение $f(t) = \beta e^{-\beta t}$, $t \geq 0$. На самом деле такое распределение описывает процессы, обладающие большей изменчивостью, чем это обычно встречается в природе. Например, для бактерий лучшей приближение к действительности обеспечивает одновершинная кривая, почти совпадающая с кривой нормального распределения. Поэтому нужно быть готовым к значительному усложнению стохастических моделей, если мы хотим, чтобы они достаточно точно описывали реальные явления.

Здесь возникает двойное затруднение. С одной стороны, сильно упрощенная стохастическая модель весьма полезна в том

отношении, что она сразу дает некоторое представление о том, что нового (по сравнению с детерминистским описанием) дает вероятностная модель. С этим мы сталкиваемся, например, в теории эпидемий при рассмотрении больших популяций. Но, с другой стороны, статистическую проверку адекватности модели и оценку тех ее параметров, которые имеют биологический смысл, стоит производить лишь для достаточно реалистичных моделей. В случае эпидемий это можно сделать, рассматривая отдельные небольшие группы индивидуумов (семьи). Однако обычно более реальные и более сложные модели приводят к громоздким математическим выкладкам. Эта проблема будет разрешаться частично за счет новых достижений в математике, особенно в области теории аппроксимаций. В этом плане трудно переоценить значение современных быстродействующих вычислительных машин. С одной стороны, они позволяют получить численные решения очень сложных систем уравнений, описывающих какой-либо реальный случайный процесс. С другой стороны, они открывают широкие возможности для исследования самих математических моделей. В этом случае вычислительная машина на основе математической модели строит большое число искусственных реализаций процесса, а затем с некоторой приемлемой точностью вычисляет значения математического ожидания, дисперсии и т. п. (разд. 2.5).

В этом кратком изложении роли случайных процессов в биологии мы ограничились рассмотрением таких процессов, в которых время можно считать непрерывным. Однако часто бывает более удобно или целесообразно оперировать с дискретным временем, как, например, при рассмотрении популяций, в которых последовательные поколения не перекрываются. Кроме того, при аппроксимации процессов с дискретными переменными диффузионными процессами как размер популяции, так и время оказываются непрерывными¹. В настоящее время существуют большие возможности выбора самых разнообразных моделей, и для любого исследователя, занимающегося проблемами биологии или медицины, весьма желательно хотя бы поверхностно познакомиться с различными вариантами. Относительно простое (и нестрогое) изложение элементов теории случайных процессов в рамках прикладной математики можно найти в другой книге автора [9]. Более строго эта теория изложена в книгах Бартлетта [11], Баруча-Рида [13] и Парцена [50], а в книге Гарленда [31] приводятся некоторые конкретные приложения теории к биологии и медицине.

¹ Впрочем, это необязательно, так как известны математические процессы с непрерывным временем, но дискретным пространством состояний. — *Прим. ред.*

2.5. ФИЗИЧЕСКОЕ МОДЕЛИРОВАНИЕ И МЕТОД МОНТЕ-КАРЛО

Моделирование, о котором упоминалось в предыдущем разделе, представляет собой по существу статистический метод исследования выборочных распределений. Когда некоторый случайный процесс удается полностью описать с помощью математической модели, то при вычислении математических ожиданий, дисперсий и вероятностей автоматически полностью учитывается вся область возможных изменений. При исследовании методом физического моделирования получают достаточно большое число возможных реализаций рассматриваемого процесса, а затем на основе этой ограниченной выборки моделированных наблюдений вычисляются искомые величины. Разумеется, эти величины не будут получены с абсолютной точностью, характерной для чисто аналитических методов, а, как и все выборочные статистики, будут иметь определенную погрешность, характеризующую соответствующей средней квадратической ошибкой. Вследствие того что в результаты, получаемые при физическом моделировании, вносится некоторая погрешность за счет случайного характера выборок, оно дает в известной степени смазанное представление о процессе. Кроме того, при таком моделировании приходится использовать фиксированные значения параметров (таких, как размер популяции, скорости размножения и гибели, число наблюдаемых поколений и т. д.). Поэтому, для того чтобы достаточно полно охватить различные комбинации значений параметров, при физическом моделировании может потребоваться очень большое число реализаций процесса. Но даже и в этом случае может оказаться, что наиболее существенные соотношения между различными параметрами модели выявить не удастся. В то же время чисто математическое исследование, если оно выполнимо, позволяет более плодотворно и всеобъемлюще раскрыть основную структуру и смысл всего процесса. По этим причинам часто говорят, что применение физического моделирования означает признание математической несостоятельности. Отчасти это, разумеется, справедливо, но в ответ можно сказать, что во многих случаях полные и изящные математические решения удается получить лишь для чрезвычайно упрощенных моделей. Поэтому для правильного использования физического моделирования нужно знать, когда оно действительно необходимо для дополнения чисто теоретических исследований. На практике может оказаться, что результаты какого-то модельного опыта подскажут, в каком конкретном направлении следует проводить новые теоретические исследования, и таким образом будет обеспечен синтез этих двух различных подходов.

Как подчеркивается в гл. 8, упрощенные стохастические модели роста популяции, построенные с учетом только процессов

размножения и гибели, крайне полезны в том отношении, что они позволяют сразу увидеть, со свойствами какого рода мы столкнемся в той или иной экологической задаче, отказываясь от негибких детерминистских моделей и вводя ряд вероятностных элементов. Однако практическая применимость таких упрощающих допущений, как, например, экспоненциальное распределение величин, характеризующих продолжительность жизни или отсутствие латентного периода при развитии инфекционного заболевания, по-видимому, весьма ограничена. Принимая эти допущения, мы можем получить лишь некоторое общее представление о процессе, но при детальном исследовании какой-либо реальной ситуации они неприменимы, в частности потому, что не позволяют оценить биологически важные параметры.

Для оценки этих параметров необходимо, чтобы модель была в достаточной степени реалистичной. Возможно, придется допустить, что значения продолжительности жизни имеют приближенно нормальное или, скажем, гамма-распределение; что каждая инфекция складывается из латентного и клинического периодов, имеющих некоторое соответствующее совместное распределение; что популяции имеют определенную структуру с четко выраженными возрастными и половыми группами, причем каждая такая группа имеет свои особые характеристики; что имеет место некоторая пространственная дифференциация популяции, которую необходимо принимать во внимание, и т. д. Как только предпринимается хотя бы малейшая попытка ввести эти факторы в исходную математическую модель, обычно исключается возможность получить математические выражения в явном виде или, во всяком случае, довести выкладки до конца. И все же исключительно важно уметь построить такую более сложную математическую модель, поскольку только она может служить основой для точного рационального рассмотрения проблемы. В некоторых случаях удается получить удобные математические аппроксимации, однако обычно единственным способом добиться успеха является моделирование на вычислительной машине.

Физическое моделирование применяется при исследовании самых разнообразных биологических и медицинских явлений: взаимодействия между видами, эволюционного поведения генетических систем, распространения и повторных вспышек эпидемий, функционирования нейронных сетей и других физиологических систем и т. п. Другая плодотворная область применения моделирования — исследование операций, о котором уже кратко упоминалось в разд. 1.4. Более подробно исследование операций рассматривается в гл. 4, где нас будут интересовать в основном административные и организационные задачи. Если такие задачи и поддаются реалистической математической формулировке, то она

обычно оказывается довольно сложной, так как должна учитывать множество самых разнообразных случайных событий. Здесь редко удается получить аналитически разрешимые математические уравнения, и реальный прогресс может быть достигнут только с помощью моделирования. Другие аспекты физического моделирования в задачах исследования операций рассматриваются в разд. 4.3. В гл. 12 излагаются некоторые конкретные примеры применения методов исследования операций в медицине, в частности вопросы, связанные с проектированием и деятельностью больниц и организацией амбулаторного приема больных.

Термин «метод Монте-Карло» часто используется как синоним «стохастического моделирования» (на вычислительных машинах), хотя первоначально он отражал логически совсем иную идею. Выше мы говорили о физическом моделировании, при котором реализации стохастической модели по своей физической природе были точно такими же, как и явления в моделируемом объекте. В отличие от такого физического моделирования метод Монте-Карло (происхождение этого термина довольно очевидно) был разработан в годы второй мировой войны для решения вероятностными методами очень сложных математических уравнений, получающихся главным образом в чисто детерминистских задачах. Метод Монте-Карло состоит в отыскании такой вероятностной задачи, решение которой приводит к тем же уравнениям, что и исследуемая детерминистская задача, после чего решение этой вероятностной задачи находится с помощью выборочных экспериментов, проведение которых по существу и есть моделирование. Простейшим примером может служить вычисление площади, ограниченной некоторой плоской замкнутой кривой. Построим вокруг этой кривой квадрат площадью A и с помощью обычного статистического метода произведем случайный выбор точек, лежащих внутри квадрата. Если доля точек, лежащих внутри данной кривой, равна p , то pA служит оценкой площади, ограниченной этой кривой. Эту идею можно легко обобщить на случай многомерных задач, для которых обычные числовые методы могут оказаться очень трудоемкими и для получения ответа потребуется огромное число точек. Кроме того, точность этих методов пропорциональна лишь $n^{-1/d}$, где d — размерность задачи. В отличие от числового подхода метод Монте-Карло позволяет получить некоторые результаты при любых значениях n и дает оценки с точностью, пропорциональной $n^{-1/2}$, которая при больших n очень высока.

На основе моделирования возник ряд новых важных методов. Прекрасным пособием для знакомства с этим предметом может служить книга Точера [62]. Более конкретные приложения метода Монте-Карло изложены в книге Хэммерсли и Хендскома [32].

Глава 3

ПРОЦЕСС НАУЧНОГО ИССЛЕДОВАНИЯ

3.1. НАУЧНЫЙ МЕТОД

«Единство всей науки,— писал Карл Пирсон в своей «Грамматике науки» [52], — заключается лишь в ее методе, а не в ее материале». Хотя нас интересует главным образом область биологии и медицины, стоит все же кратко рассмотреть общие принципы научного исследования. После этого можно продолжить более подробное изучение различных проблем, возникающих в интересующей нас области в связи с построением и применением математических моделей.

Вообще говоря, научный метод представляет собой непрерывный процесс проверки, изменения и развития идей и теорий в соответствии с имеющимися фактическими данными. В известной степени научный метод — это просто развитие обычного рационального подхода, основанного на здравом смысле. В то же время многие глубокие научные исследования действительно настолько сложны, специфичны и изложены таким непонятным языком, что не сразу можно обнаружить их связь с повседневной жизнью. Тем не менее требование, предъявляемое к любой научной теории, — чтобы в конце концов она приводила к строгим утверждениям, — означает возможность объяснить общий ее смысл (хотя бы приближенно) обычным понятным языком.

Рассматривая более детально применение научного метода в любой ситуации, можно выделить ряд четко различимых и взаимосвязанных этапов. Первый этап — это этап наблюдений, который можно назвать «естественноисторическим». На этом этапе происходит просто накопление огромной массы разнородного материала, характер которого преимущественно зависит от случайных интересов одного или нескольких исследователей; часть его основана на точных измерениях, а другая часть представляет собой лишь отрывочные описательные данные. Затем предпринимается попытка систематизировать имеющиеся факты и, возможно, получить некоторое систематическое описание всей совокупности данных. Типичным примером этого первого этапа служит работа биолога, связанная со сбором материала по тем или иным видам животных: описываются общий вид, биология, поведение отдельных особей; особое внимание может быть уделено способам добывания пищи, а также брачному поведению, размножению и заботе

о потомстве; отмечается влияние конкретных особенностей окружающей среды — климата, географических условий и т. д.

Однако рано или поздно возникают некоторые специальные вопросы и проблемы. Допустим, например, что у какого-либо вида животных наблюдаются две различные окраски тела: светлая и темная. В этом случае можно задать вопрос, не имеем ли мы здесь дело с двумя различными подвидами, определенно связанными с двумя различными типами среды, или же это пример сбалансированного полиморфизма по какому-либо признаку в пределах одного вида. На данном этапе формулируется некоторая удобная гипотеза, которая если и не полностью, то в значительной мере объясняет эмпирические факты и, что особенно важно, может быть в дальнейшем опровергнута новыми данными. Принципиальное значение имеет возможность проверить справедливость гипотезы. Такая гипотеза, которую невозможно опровергнуть (хотя бы в принципе), по существу бессмысленна. Разумеется, она может привести к некоторым более плодотворным теоретическим построениям, но значительно чаще ее результатом будет создание целой абстрактной теории, проверка которой невозможна. Наилучшие гипотезы — это простые гипотезы, которые легко подтвердить экспериментально, если они верны, и легко опровергнуть с помощью надлежащим образом подобранных решающих экспериментов или наблюдений, если они неверны. В приведенном выше примере можно предположить, что две наблюдаемые окраски обусловлены заменой одного гена в одном локусе и, возможно, связаны с наличием доминантного и рецессивного аллелей. Если изучаемых животных легко добыть и они достаточно быстро размножаются в неволе (как, например, мыши или полевки), то для проверки этой гипотезы можно произвести требуемое число скрещиваний. Может оказаться, что светлая форма является рецессивной, а темная — доминантной и что наличие этих двух форм может поддерживаться неопределенно долгое время за счет расщепления при возвратных скрещиваниях. Потребуется определенное время, прежде чем можно будет с достаточной уверенностью считать, что найдено удовлетворительное объяснение и что наблюдаемые фенотипы не связаны ни с какими аномалиями.

Если бы здесь действовал какой-либо иной механизм, то это выявилось бы довольно рано, так как фактические соотношения при расщеплении отличались бы от тех, которых следовало ожидать на основе простой гипотезы об одном локусе, т. е. от соотношения 1 : 1 при возвратном скрещивании и соотношения 3 : 1 при скрещивании между гибридами. Когда число скрещиваемых животных достаточно велико, можно определить, реализуются ли эти соотношения хотя бы приближенно или же наблюдаются значительные отклонения от них.

Однако очень часто встречаются менее определенные ситуации. Число животных может быть не очень велико, и в выборках могут иметь место значительные случайные колебания. В этих случаях могут возникнуть соотношения, которые мало отличаются от теоретически ожидаемых, но которым можно дать иное объяснение. При небольшой численности потомства мы не сможем, например, легко провести различие между соотношением 3 : 1, получаемым при моногибридных скрещиваниях, и соотношением 9 : 7, наблюдающимся при наличии двух независимо расщепляющихся комплементарных генов. В данном случае крайне важно выбрать правильный метод статистического вывода, определяющий направление дальнейших действий. В общем случае простая первоначальная гипотеза будет отклонена, если наблюдения покажут достаточно большое расхождение с ней. Обычно именно на это и направлены научные исследования. Однако из-за существования значительной естественной изменчивости достоверный результат получить невозможно. Заметные расхождения между теорией и наблюдениями могут возникать чисто случайным образом. В точных науках этот аспект обычно не имеет большого значения. Но в биологии часто возникает необходимость проверки статистической значимости полученных результатов (см. разд. 2.2). Если наблюдаемые расхождения достаточно велики, то первоначальная, т. е. нулевая, гипотеза отвергается в пользу некоторой альтернативной гипотезы.

Это подводит нас к следующему этапу процесса научного исследования. Если принятие некоторой гипотезы приводит к значительному расхождению между теорией и практикой, то эта гипотеза не может больше сохраняться в своем первоначальном виде. В тех случаях, когда гипотеза имеет прочную основу, например если она выводится из какой-то существующей вполне установленной теории, исследователь будет скорее склонен вносить в эту гипотезу некоторые изменения, а не отвергать ее полностью. Допустим, например, что другие исследователи на основе многочисленных данных установили, что различная окраска упомянутых выше животных обусловлена просто заменой доминантного гена на рецессивный. Мы не можем полностью отвергнуть эту гипотезу на основе отклонений, наблюдаемых при новых экспериментах. Разумеется, возможно, что мы имеем дело с каким-то другим генетическим фактором. Это можно проверить путем скрещивания наших животных с теми, на которых работали другие исследователи. Если окажется, что исследовался один и тот же фактор, то можно предположить, что полученные нами соотношения отклоняются от обычных вследствие различий в жизнеспособности различных фенотипов и генотипов. Эти идеи мы должны включить в нашу гипотезу о типе наследования и использовать

их в новых экспериментах, спланированных для проверки справедливости этой видоизмененной гипотезы.

Теперь мы прошли полный цикл, хотя новая ситуация является более сложной и интересной, чем первоначальная. Выполнены эксперименты и проведено сравнение с результатами, полученными другими исследователями. Появились дополнительные усложнения, которые в свою очередь также необходимо было усложнить. При этом могут возникать и другие побочные проблемы, например небольшие различия между соотношениями расщепления при мужском и женском гаметогенезе; поколения с отдельными случаями «непроявления» наследуемого признака; изменение соотношений под действием условий среды, например изменений температуры, и т. д. Постепенно накапливается полный объем знаний, касающихся не только тех признаков, с которых началось исследование, но и других признаков, возможно генетически сцепленных с первыми, влияния окружающей среды и т. д.

В то время как изолированную гипотезу можно полностью отвергнуть, если она противоречит экспериментальным данным, гипотеза, выведенная из вполне установившейся теории, имеет прочную основу, даже если она не дает удовлетворительного объяснения новых наблюдений. В этом случае необходимо привести некоторые небольшие изменения, которые позволили бы согласовать новые данные с существующей теорией. При этом основная масса знаний остается нетронутой, хотя при необходимости она подвергается пересмотру, изменению или развитию. Теория не только дает удобный и доступный способ описания широкого круга взаимосвязанных явлений, но и подсказывает новые гипотезы, которые могут быть проверены и дают новые знания. С практической точки зрения весь этот процесс позволяет непрерывно расширять наши знания и проверять их.

Такова в общих чертах сущность научного метода. Его философия является в основном прагматической и эмпирической. Гипотезы и теории ценны лишь до тех пор, пока они действуют. Наши знания не являются абсолютной истиной — они справедливы лишь до тех пор, пока они не противоречат известным фактам. Обычно развитие науки происходит на основе пересмотра и изменения существующих идей. При этом иногда отвергается вся теория как непригодная. Однако даже и в этом случае создаваемая ею интеллектуальная основа может сыграть полезную роль в том смысле, что она способствует выполнению точных исследований и стимулирует проведение новых экспериментов. Нередко в конечном счете приходится возвращаться к, казалось бы, устаревшим теориям, и их некоторые частично справедливые положения включаются в новые более полные теории. Например, на смену примитивным корпускулярной и волновой теориям света в конечном

счете пришла более общая волновая механика. Работы в области биометрии, появившиеся в конце XIX в., затмили простые корпускулярные идеи Грегора Менделя. Однако впоследствии Р. Фишер в результате синтеза этих двух направлений создал прочную основу для построения генетической теории, позволяющей объяснить как количественную, так и качественную изменчивость и послужившей толчком для широких количественных исследований механизма эволюционного развития.

«Единство всей науки заключается лишь в ее методе, а не в ее материале», — повторим высказывание Карла Пирсона. А от себя добавим — «и не в мотивации». Направление научных исследований, безусловно, в большой мере зависит от круга интересов отдельных ученых и их любознательности, но не менее важное значение имеют разнообразные общественные факторы. Наличие денег и научной аппаратуры, атмосфера, способствующая проведению научных исследований, потребности общества — все это в значительной мере определяет, какими проблемами нужно заниматься и какими — нет. Все эти вопросы выходят за рамки обсуждения научного метода как такового. Мы, несомненно, можем рассматривать научный метод как главное и наиболее мощное средство рационального познания. Однако он служит лишь средством для достижения цели. А цели выбираются не на рациональной основе. «Разум является или должен являться рабом страсти», — сказал где-то Бертран Рассел, подразумеваемая под страстью любое желание, а также воображение, интуицию и моральные соображения, определяющие выбор целей. Разум и рациональное поведение служат лишь наилучшим средством достижения этих целей.

Для четкого понимания сути научного метода важно проводить различие между этими эмоциональными мотивациями и рациональной формой самого научного процесса. Этот момент особенно важно иметь в виду в следующем разделе, где мы перейдем к обсуждению значения математических моделей.

Выше мы подчеркнули роль эмоциональных факторов как мотиваторов, побуждающих нас к проведению научных исследований. Такого же рода факторы играют крайне важную роль еще в одном отношении. Когда мы переходим к деталям какого-либо конкретного научного исследования, то обнаруживаем, что успех при выборе проблем и построении гипотез в значительной мере зависит от интуиции и воображения. Детальная проверка выбранной модели на внутреннюю логичность, построение прогнозов и т. п. в значительной мере относятся к дедуктивной логике, а проверка гипотез и оценка параметров связывают нас с индуктивной логикой; обе эти формы логических построений по существу являются рациональными. Однако лауреат Нобелевской премии в области науки

должен не только уметь строго рассуждать и эффективно использовать лабораторное оборудование, но и обладать гениальной способностью выбирать правильные проблемы и формулировать их таким образом, чтобы получить плодотворные результаты.

Сформулировать какие-либо общие правила, позволяющие развивать интуицию и воображение, невозможно, так как они по определению не являются рациональными или, во всяком случае, лежат вне пределов рационального. Однако при более внимательном изучении обстоятельств, в которых проявляются интуиция и воображение, мы сможем лучше узнать, какие аспекты нашей работы выиграют от эффективного применения таких рациональных средств, как научная методология, математические модели, быстродействующие вычислительные машины и т. д., и какие аспекты определяются совершенно иными соображениями.

3.2. ЗНАЧЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ

В гл. 1 мы высказали убеждение в преимуществе метода математического моделирования как средства, позволяющего достигнуть значительно большей ясности и точности, чем чисто словесные методы. Вначале мы остановились на относительно статических ситуациях, а затем рассмотрели динамические процессы, изменяющиеся во времени. В гл. 2 были введены вероятностные аспекты, и это привело нас к обсуждению случайных процессов, а также вопросов математической статистики, связанных с проверкой адекватности математических моделей на основе фактических данных. Как было указано в предыдущем разделе, такая проверка гипотез составляет существенную часть научного метода. Любая разумная научная гипотеза имеет хотя бы некоторые количественные аспекты, однако в данной книге нас больше интересуют такие математические описания, которые достаточно детальные, чтобы заслуживать названия математической модели. Но такая модель имеет в точности такой же логический статус, что и любая гипотеза с гораздо меньшим количественным содержанием. Это означает, что математическая модель дает частичное описание определенных аспектов реальной действительности и ее справедливость целиком зависит от точности этого изображения.

Учитывая крайнюю сложность большинства биологических систем, нетрудно понять, что в биологии простые и легко поддающиеся исследованию математические модели представляют собой чрезвычайно грубые приближения. И, что еще хуже, математики могут с энтузиазмом приняться за глубокую теоретическую разработку моделей, неадекватность которых известна заранее, только потому, что это не составляет большого труда (во всяком случае, для них). Может также случиться, что математическому

анализу будет подвергнута вполне правдоподобная модель, но она не удовлетворит требованиям научного метода, поскольку нельзя найти способа проверить полученные результаты. Значительная часть математической биологии не защищена от такого рода возражений, которые, если придавать им слишком большое значение, могут причинить большой вред. Биологи и врачи совершенно справедливо с подозрением относятся к любой работе, которая представляет собой лишь клубок математических абстракций и не вписывается в непрерывный процесс развития науки.

По нашему мнению, математическая модель, если ее правильно понимать и правильно применять, имеет точно такой же логический статус, как и любая другая научная гипотеза, и поэтому при обращении с ней и проверке ее справедливости необходимо исходить из тех же критериев. Именно потому, что модель формулирует задачу, так сказать, в «очищенном» виде, значительно легче почувствовать трудности этой задачи. С этими трудностями можно частично справиться путем отыскания лучших способов исследования моделей, достаточно сложных для близкого соответствия с реальными процессами (примером может служить моделирование сложных случайных процессов на вычислительной машине). Однако это еще не все. В конце раздела 2.4, посвященного случайным процессам, кратко говорилось о том, что простые модели, отражающие лишь немногие свойства реального процесса, важны тем, что они дают общее представление о процессе, тогда как для достижения статистического соответствия с фактическими данными необходимы более сложные модели. Это противоречие между требованиями, которым должны удовлетворять модели, дающие общее представление о процессе, и модели, дающие реальное его изображение, имеет существенно важное значение в математической биологии, и необходимо в полной мере отдавать себе отчет в связанных с ним трудностях.

Так, на элементарном уровне детерминистский экспоненциальный закон может оказаться вполне приемлемой моделью роста популяции при решении таких демографических задач, как грубое предсказание роста численности населения, т. е. в тех случаях, когда рассматриваются большие числа. Более точные результаты можно получить, разбивая популяцию на группы по возрасту и полу, для каждой из которых характерны свои собственные скорости роста или уменьшения численности. Допустим теперь, что необходимо исследовать более мелкие группы, в которых наблюдаются очень сильные статистические колебания. Как показано в разд. 2.4, в этих случаях возникают значительные сложности даже при изучении простейшего случайного процесса размножения; что же говорить о трудностях, которые вносят такие факторы, как гибель, миграция, пространственная деформация и т. д.

Более того, в большинстве аналитически разрешимых моделей принимается ряд необоснованных допущений, например допущение об экспоненциальном распределении продолжительностей жизни. Выбор более реалистических распределений значительно затрудняет формулировку и описание рассматриваемых процессов. Тем не менее существенно, что даже чрезмерно упрощенные модели находят применение и имеют известную научную ценность.

Допустим, что мы построили вначале детерминистскую модель. Тогда перед нами встает вопрос: к каким последствиям приведет включение в нее вероятностных элементов, т. е. рассмотрение соответствующих распределений? Окажет ли это существенное влияние на результаты? Какие новые общие свойства будет иметь стохастическая модель по сравнению с детерминистской моделью? И так далее. На все эти вопросы можно дать общий ответ, исследовав некоторые простые случайные процессы и вычислив такие величины, как математические ожидания, дисперсии, вероятности вымирания популяции, распределения размеров популяции и т. п. Все это должно дать более глубокое представление об исследуемом процессе. Возможно, мы придем к заключению, что одни его аспекты необходимо рассматривать, а другими можно пренебречь. Например, при изучении мутаций у бактерий необходимо выбрать вероятностный подход при рассмотрении небольших групп мутантов, так как здесь наблюдаются большие статистические колебания, но основная популяция, состоящая из клеток дикого типа, во многих случаях достаточно велика, и в качестве первого приближения вполне приемлем детерминистский подход (см. разд. 8.3). Простейшее исследование повторяющихся эпидемий (см. разд. 9.4) вероятностными методами показывает, что такого рода математическое описание позволяет в общих чертах объяснить важное свойство таких эпидемий — периодическое возникновение вспышек примерно одинаковой интенсивности, тогда как детерминистская модель дает ряд затухающих колебаний, что не согласуется с наблюдаемыми явлениями. При желании разработать более детальные, реалистические модели мутаций у бактерий или повторяющихся эпидемий эта информация, полученная с помощью предварительных упрощенных моделей, будет иметь очень большую ценность. В конечном счете успех всего направления научных исследований определяется возможностью моделей, построенных для объяснения и предсказания реальных наблюдений.

Одно из больших преимуществ правильно построенной математической модели состоит в том, что она дает довольно точное описание структуры исследуемого процесса. С одной стороны, это позволяет осуществлять ее практическую проверку с помощью соответствующих физических, химических или биологических экспериментов. С другой стороны, математический анализ расши-

ряет наши теоретические знания. Если основные уравнения можно решить аналитическим путем, то, подставляя в них различные значения рассматриваемых параметров, мы автоматически получаем решение всех возможных вариантов задачи. Кроме того, сама по себе аналитическая форма решений может углубить наши знания, позволяя выразить наблюдаемые биологические закономерности в виде математических теорем. Примерами такого рода служат пороговая теорема в теории эпидемий и условия достижения равновесия в популяции, в которой благоприятные и неблагоприятные мутации генетически сцеплены.

По возможности нужно применять чисто математические методы исследования модели, так как это позволяет наиболее полно использовать мощные аналитические возможности. К сожалению, во многих случаях получить решение основных уравнений аналитическими методами не удастся и необходимо обращаться к численным решениям. Численный анализ, который рассматривается далее в разд. 5.2, полон ловушек, подстерегающих неосторожного исследователя. Однако при соблюдении достаточной осторожности численные решения часто дают значительный объем полезной информации о свойствах модели. По мере усложнения моделей и приближения их к реальным процессам уменьшается возможность получения лаконичных изящных решений в явном виде и все более возрастает необходимость обращаться к тем или иным формам численных решений. Поэтому в настоящее время исключительно важное значение приобретают быстродействующие вычислительные машины.

В некоторых случаях возникают более серьезные трудности. Может оказаться, что полученные дифференциальные уравнения движения для некоторого сложного биологического процесса (это могут быть дифференциальные уравнения в частных производных высокого порядка) не только неразрешимы аналитически, но и не поддаются решению существующими методами численного анализа. В этом случае наиболее целесообразно применение физического моделирования (разд. 2.5). Как и типичные методы численного анализа, физическое моделирование обладает тем недостатком, что оно не позволяет получить аналитические выражения для рассматриваемого процесса. Однако этот недостаток компенсируется такими преимуществами, как простота и гибкость метода и возможность избежать сложных численных расчетов, полагаясь на статистические свойства достаточно большого числа повторных вычислений.

Последний вопрос, заслуживающий упоминания в связи с использованием физического моделирования, состоит в следующем. Хотя в принципе такое моделирование дает не больше информации, чем чисто математические расчеты, на практике оно

обеспечивает значительное приближение к условиям реального эксперимента. Моделирование может приобрести особенно важное значение при изучении сложных биологических систем и уже широко применяется при исследовании операций (эта область довольно детально рассматривается в гл. 4). Многие задачи исследования операций возникают из необходимости внедрить некоторую близкую к оптимальной схему действия, однако часто оказывается, что характер задачи не позволяет провести эксперименты в реальных условиях. В этих случаях возможно экспериментирование на достаточно сложных моделированных системах, обеспечивающих высокую степень приближения к реальности.

3.3. МОДЕЛИ СТАТИСТИЧЕСКОГО ВЫВОДА

При обсуждении в разд. 2.2 одного из наиболее распространенных методов проверки статистической значимости упоминались различные модели статистического вывода. Хотя на практике использование разных теоретических подходов чаще всего приводит к выводам, мало отличающимся друг от друга, при общем описании процесса научного исследования следует упомянуть о многообразии логических доказательств. Нас интересуют здесь не столько детали дедуктивных рассуждений (словесных или математических), в результате которых на основе четко определенных теоретических допущений получаются логически правильные выводы, сколько более неопределенный тип индуктивного, или недостоверного (эвристического), вывода, при котором делается попытка получить обоснованные обобщения реальных явлений.

В настоящее время большинству авторов научных статей известны простейшие статистические рассуждения, особенно благодаря тому, что редакторы журналов часто просят авторов «вставить статистику», которая, по их мнению, придает статьям необходимую респектабельность. Конечно, пытаться включать статистический материал в уже завершённую экспериментальную программу не имеет смысла. Чтобы достигнуть наилучших результатов, нужно планировать экспериментальную работу таким образом, чтобы в ней с самого начала была предусмотрена соответствующая статистическая обработка данных. Разумеется, множество глубоких биологических и медицинских исследований было успешно выполнено без особого внимания к статистическим тонкостям. Но во многих случаях планирование эксперимента, предусматривающее достаточное использование статистики, значительно повышает эффективность работы и обеспечивает получение большего объема информации о большем числе факторов при меньшем числе наблюдений. В противном случае эксперимент может оказаться неэффек-

тивным и неэкономичным и даже привести к неверным выводам. В этих случаях новые гипотезы, построенные на таких необоснованных выводах, не смогут выдержать проверку временем. Если проводится непрерывная последовательность экспериментов, то рано или поздно свойственный научному методу механизм самокоррекции позволит исправить большинство допущенных ошибок.

Отсутствием статистического подхода можно в какой-то мере объяснить периодическое появление «модных» препаратов или методов лечения. Очень часто врачи ухватываются за те или иные новые препараты или методы лечения и начинают широко применять их только на основании кажущихся благоприятных результатов, полученных на небольших выборках данных и обусловленных чисто случайными колебаниями. По мере того как у медицинского персонала накапливается опыт применения этих препаратов или методов в больших масштабах, выясняется, что возлагавшиеся на них надежды не оправдываются. Однако для такой проверки требуется очень много времени и она весьма ненадежна и неэкономична; в большинстве случаев этого можно избежать путем правильно спланированных испытаний на самом начальном этапе.

В настоящее время специалисты в области биоматематики настоятельно рекомендуют применять статистические методы при проверке гипотез, оценке параметров, планировании экспериментов и обследований, принятии решений или изучении работы сложных систем. Все эти виды деятельности постоянно связаны с использованием эвристических выводов. Целесообразно рассмотреть различные методы статистического вывода, существующие в настоящее время, иначе исследователь может упустить из виду или вообще игнорировать некоторые из них просто потому, что неправильно понимает их или считает, что они противоречат другим известным ему методам.

Детальное математическое изложение этих вопросов можно найти в соответствующих руководствах, например в книге Кендалла и Стюарта [38]. В данном разделе мы остановимся лишь на основных чертах различных методов. Рассмотрим, например, простую статистическую проверку, описанную в разд. 2.2. Суть ее состоит в том, что мы строим какую-либо подходящую нулевую гипотезу, а затем смотрим, не соответствуют ли имеющиеся данные появлению события, которое следовало бы считать довольно редким, если предположить, что нулевая гипотеза верна. Если это так, то исходная нулевая гипотеза отвергается (при выбранном уровне значимости) и рассматривается некоторая видоизмененная гипотеза, которая в свою очередь подвергается проверке. Фигурально выражаясь, мы

ставим перед собой ряд деревянных кукол¹, чтобы сбивать их по очереди, если этого потребуют обстоятельства.

Многим, вероятно, покажется, что в тех случаях, когда сложные суждения можно сжато выразить простыми терминами, большинство ученых работает именно таким или, во всяком случае, весьма аналогичным методом. Но существует и другой метод проверки гипотез, хорошо известный тем, кто занимается приемочным контролем в промышленности. Допустим, что имеется простая альтернатива для нулевой гипотезы; скажем, последняя утверждает, что некоторый параметр θ (например, математическое ожидание некоторой случайной величины, распределенной по нормальному закону) равен нулю, а альтернативная гипотеза утверждает, что он принимает некоторое другое значение θ_1 . В данном случае с помощью критерия Фишера для проверки значимости (о котором уже говорилось) можно установить некоторый уровень значимости, например p_1 , означающий вероятность того, что верная нулевая гипотеза будет отклонена; однако об альтернативной гипотезе ничего конкретного сказать нельзя. Если она точно известна, то можно построить критерий, который задает уровень значимости p_1 и в то же время позволяет минимизировать вероятность p_2 того, что будет отвергнута верная альтернативная гипотеза (т. е. принята ложная нулевая гипотеза). Величины p_1 и p_2 обычно называют ошибками первого и второго рода соответственно. Этот подход по существу и составляет основу *теории статистических испытаний Неймана — Пирсона*. Если вероятностям p_1 и p_2 можно задать приемлемые значения, то эта теория позволяет получить весьма удовлетворительное объяснение статистических критериев и дает удобный способ нахождения новых критериев, когда они потребуются.

В промышленности при выборочной проверке некоторого продукта обычно можно в денежном выражении оценить убытки, причиняемые при допущении ошибок первого и второго рода. Руководствуясь мотивом прибыли, можно выбрать оптимальные значения p_1 и p_2 . В том случае, когда имеется несколько альтернативных гипотез или даже непрерывный спектр альтернатив, ситуация значительно усложняется, хотя основные идеи останутся неизменными. В чисто научном плане эта задача состоит в том, чтобы решить, существует ли какая-либо основа для выбора определенной альтернативной гипотезы. Многие считают, что научное исследование — это по существу бесконечно продолжающийся процесс, и если окажется, что какая-либо гипотеза

¹ Автор имеет в виду старинную народную игру Aunt Sally (тетушка Салли), которая состоит в том, что игрок с определенного расстояния должен выбить трубку изо рта деревянной куклы.— *Прим. перев.*

не подходит, то нужно рассмотреть возможные альтернативы в свете самых последних данных, а не предрешать возможность существования какой-либо определенной гипотезы. Эта точка зрения естественным путем приводит к методу Фишера.

В целом для ученых представляет большой интерес проверка конкретных нулевых гипотез, а теория Неймана — Пирсона кажется им громоздким и необоснованным усложнением. Однако, поскольку теория Неймана — Пирсона определяет свои задачи более точно, она в большей степени допускает возможность математических обобщений и выводов. По этой причине многие статистики утверждают, что ученые должны изменить свое мнение и воспользоваться преимуществами этого более сложного подхода. В некоторых частных случаях теория Неймана — Пирсона может применяться совершенно естественным образом. Однако в общем случае трудно предвидеть, к каким практическим последствиям приведут ошибки второго рода, даже если предполагается, что конкретные альтернативные гипотезы могут быть сформулированы достаточно четко. Довольно трудно также решить, какими должны быть наилучшие вероятности для ошибок первого рода, т. е. обычные уровни значимости, хотя с практической точки зрения объяснить их значительно легче.

Еще один подход к решению всех этих задач дает так называемая *теория статистических решений*. Совершенно ясно, что проверки значимости связаны с решениями о принятии или отклонении нулевой гипотезы, и очень многие статистические задачи проверки значимости и получения оценок можно сформулировать как задачи принятия решений. Обращение к тому или иному методу в значительной мере зависит от личных склонностей отдельных математиков. Однако существует много практических ситуаций, особенно в задачах исследований операций, когда сам характер задачи предполагает необходимость принятия решений. Так, врач может принять решение посетить одного больного, а не другого; он может выбрать один предварительный диагноз вместо другого и, следовательно, в дальнейшем будет вынужден действовать соответствующим образом; он может сам лечить больного или же направить его на консультацию в больницу и т. д. Если мы начинаем искать точный ответ на вопрос о том, как найти правильный, или оптимальный, образ действий, возникают самые разнообразные осложнения. Нужно суметь учесть все возможные решения, а также вероятности различных последствий, к которым может привести каждое из них. Кроме того, потребуется некоторая оценка значения этих последствий для врача, больного и общества. Если всю эту информацию удастся получить, то в принципе можно рассчитать оптимальный образ действий. На практике, когда необходимо быстро принимать решения на основе весьма неполных

данных, обычно приходится полагаться на субъективные суждения в надежде, что они позволят получить довольно хорошие, если не наилучшие, результаты.

Может показаться, что подход к анализу сложных реальных проблем с точки зрения теории статистических решений является безнадежно абстрактным и утопическим. Однако имеется ряд причин считать, что это не так. Естественно, что в книгах по теории решений для иллюстрации приводятся лишь чрезмерно упрощенные варианты реальных задач. Конечно, в тех случаях, когда то или иное действие нужно произвести немедленно, такие методы применяются пока редко, однако вполне возможно, что с дальнейшим развитием быстродействующих вычислительных устройств они найдут более широкое применение. Еще важнее, что теория статистических решений дает основу для количественного подхода к очень сложным административным и организационным задачам, которые до самого последнего времени считались недоступными для математических методов. Таким образом, мы вправе сказать, что эта теория позволяет получить представление о том, что происходит на уровне административного управления, а также помогает более четко представить, что здесь в действительности требуется. Следующий этап, на котором будут предприниматься попытки производить выбор между различными моделями принятия решений, еще только начинает развиваться. В настоящее время такие модели имеют главным образом теоретическое значение, однако в будущем они, по-видимому, приобретут значительно большую практическую ценность.

С проверкой гипотез связан еще один фундаментальный вопрос: можно ли считать, что гипотезы характеризуются некоторым распределением вероятностей? При частотной интерпретации те или иные значения вероятности можно приписать лишь результатам таких экспериментов, которые, во всяком случае в принципе, можно повторить, причем эти вероятности определяются через наблюдаемые частоты. Гипотезы же бывают либо справедливы, либо ложны и обычно не имеют распределений, хотя они с большей или меньшей вероятностью соответствуют имеющимся данным. Критерии для проверки значимости, предложенные Фишером или Нейманом и Пирсоном, не связаны с допущениями о существовании априорных вероятностей гипотез. Однако в частных случаях эти априорные вероятности могут существовать. Если это так, то, используя правило Бейеса, можно определить апостериорные вероятности самих гипотез.

Допустим, например, что для дальнейшего экспериментирования мы выбрали какую-то особь с доминантным признаком, полученную при скрещивании $Aa \times Aa$. Априорная вероятность того, что эта особь обладает генотипом AA , равна $1/3$, а вероятность

генотипа Aa равна $\frac{2}{3}$; третья возможность, aa , для нее исключается. Если мы теперь проведем аналитическое скрещивание с особью, несущей два рецессивных аллеля, aa , то появление хотя бы одной особи с рецессивным признаком покажет, что проверяемая особь является гетерозиготой Aa . Допустим, что вместо этого мы получили три особи с доминантным признаком. Вероятность этого события равна единице (т. е. это событие достоверно), если проверяется особь AA , и $\frac{1}{8}$ при генотипе Aa . Таким образом, апостериорные вероятности равны $\frac{1}{3} \times 1$ и $\frac{2}{3} \times \frac{1}{8}$, или $\frac{1}{3}$ и $\frac{1}{12}$ соответственно. Нормируя эти относительные частоты таким образом, чтобы их сумма была равна единице, находим, что вероятности генотипов AA и Aa равны соответственно $\frac{4}{5}$ и $\frac{1}{5}$. С полным основанием можно сказать, что с вероятностью $\frac{4}{5}$ проверяемая особь гомозиготна по доминантному гену. Этот вывод подтверждается реально только в том случае, если существуют априорные вероятности. Если при использовании критерия Фишера достигается уровень значимости, равный, скажем, 1%, то часто говорят, что шансы на то, что нулевая гипотеза справедлива, составляют 1 : 100. Хотя с логической точки зрения это утверждение не вполне строго, оно вряд ли покажется кому-нибудь недостаточно ясным, разве только теоретикам.

Невозможность оценить шансы за и против гипотезы при частотной интерпретации вероятности часто рассматривается как недостаток этой теории по сравнению со здравым смыслом. С другим подходом мы встречаемся в так называемой *теории инверсных (субъективных) вероятностей*, превосходное изложение которой можно найти в книге Джеффри [37]. Эта теория развита на основе допущений, предполагающих использование распределения вероятностей для измерения степени уверенности или доверия. Так, в том случае, если о некотором параметре θ ничего не известно, за исключением того, что он может принимать любое из трех значений a , b и c , этим значениям приписываются одинаковые априорные вероятности, равные $\frac{1}{3}$. Если первоначально отдается предпочтение значению $\theta = c$, то этому событию приписывается большее значение вероятности. Хотя можно спорить о том, правильную или неправильную форму имеет то или иное априорное распределение, преимущество этого метода состоит в том, что он позволяет учесть уже имеющиеся данные, тогда как при частотной интерпретации вероятности сделать это трудно. Метод субъективных вероятностей имеет много привлекательных свойств; некоторые его приложения более подробно излагаются в гл. 11, где рассматривается ряд задач медицинской диагностики. Типичная

задача постановки диагноза легко и совершенно естественно формулируется через субъективные вероятности. Составление окончательного перечня различных возможных диагнозов, каждому из которых ставится в соответствие определенная субъективная вероятность, по-видимому, ближе к клинической практике, чем методы, основанные на частотной интерпретации вероятности. Разумеется, как уже отмечалось ранее, при постановке диагноза более эффективным был бы подход, основанный на теории решений, однако использовать его на практике пока еще трудно, поскольку обычно отсутствуют очень многие существенные элементы, необходимые для применения этого метода.

Мы кратко остановились на методе субъективных вероятностей в связи с вопросом о проверке гипотез. Однако сюда легко включить и получение оценок. Если рассматривается некоторый непрерывный параметр θ , то его апостериорное распределение показывает, какая степень уверенности ставится в соответствие всей области возможных значений этого параметра. При достаточном количестве данных это распределение будет сконцентрировано около некоторого предпочтительного значения, которое можно использовать как оценку. Концентрация вероятности вокруг этого значения измеряет соответствующую степень уверенности и во многих случаях выражается известной нам средней квадратической ошибкой.

Многие исследователи выдвигают принципиальные возражения против использования априорных вероятностей в тех случаях, когда они вводятся для измерения степени уверенности, а не как фактические частоты, как в рассмотренном выше примере из области генетики. Чтобы обойти это возражение, Фишер ввел так называемое *фидуциальное распределение*. В этом методе никакие допущения относительно априорного распределения неизвестного параметра θ не принимаются, однако данные используются таким образом, что для параметра θ можно составить математические выражения, *эквивалентные* вероятностному распределению θ . При частотной интерпретации вероятности такого распределения не существует, и в этом состоит суть дела. Однако при интерпретации вероятности как степени уверенности рекомендуется использовать это фидуциальное распределение как показатель того, какая степень уверенности ставится в соответствие различным возможным значениям θ . По форме фидуциальное распределение выглядит в точности так же, как и апостериорное распределение, хотя по своей логической основе они заметно отличаются друг от друга. Во многих простых задачах математической статистики, например при оценке математического ожидания нормального распределения, применение фидуциального распределения, апостериорного распределения или же частотной интерпретации вероятности дает

практически одинаковые результаты, однако в более сложных задачах появляются различия. Еще важнее существование некоторых аномалий и очевидных логических несоответствий при использовании фидуциальных распределений. По этим причинам многие математики очень критически относятся ко всему этому подходу в целом. Если бы эти трудности удалось преодолеть, то фидуциальные распределения стали бы более популярны среди теоретиков и их охотнее применяли бы ученые, ведущие практическую исследовательскую работу. Последним, естественно, необходим такой статистический метод, который позволял бы ставить в соответствие неизвестным гипотезам или различным значениям неизвестного параметра определенные вероятности, однако явно небезопасно использовать метод, имеющий логические изъяны.

Чем кончатся дебаты о различных интерпретациях вероятности, сказать трудно. Но пока философы и математики обязаны искать ответ на ряд насущных вопросов. Чем следует руководствоваться, решая, какому из нескольких хорошо разрекламированных продуктов стоит отдать предпочтение? Будет ли исследование более эффективным, если основным принципом получения информации явится теория статистических решений, а не критерий статистической значимости Фишера и не фидуциальное распределение? Будут ли результаты, полученные в сотрудничестве со специалистом в области биометрии, предпочитающим использовать субъективные распределения, отвергнуты научным журналом, редакторы которого являются сторонниками ортодоксальной частотной интерпретации вероятности? Ученый-практик может избежать бесконечных дебатов и бесплодного обсуждения мелких вопросов логического характера, заняв ту же самую прагматическую позицию, которой он руководствуется в своих делах и решениях, касающихся более знакомых ему технических деталей в его собственной области. В своей практической работе он привык к широкому использованию моделей, сформулированных главным образом в словесной форме или же преимущественно математических (в данной книге мы отстаиваем необходимость более широкого применения именно этих моделей). Различные теории и методы статистического вывода можно рассматривать точно в таком же плане, хотя и на более абстрактном уровне. Это просто разные модели статистического вывода, разные способы получения обоснованных выводов при определенных исходных данных и допущениях. Очень часто получаемые выводы оказываются совершенно одинаковыми, какую бы модель мы ни выбирали, и этому вряд ли стоит удивляться. Однако иногда различные модели, представляющие собой по существу различным образом построенные изображения некоторого сложного процесса, приводят к неодинаковым выводам. Если эти выводы дополняют друг друга, то никаких

проблем не возникает, и налицо определенный выигрыш. Если же получаемые выводы несовместимы, то необходимы дальнейшие исследования, и при попытке разрешить возникшие противоречия можно получить новые данные.

Цель этого раздела состоит не в детальном описании различных методов статистического вывода, а скорее в том, чтобы показать многообразие существующих методов и подчеркнуть, что в определенной ситуации тот или иной из них может оказаться наиболее подходящим. Выбор метода может зависеть от целого ряда факторов, в том числе от формулировки исходной задачи, характера вопросов, наличия соответствующего математического аппарата, возможности использовать электронную вычислительную машину и, наконец, от личных склонностей самого исследователя.

3.4. МНОГООБРАЗИЕ МАТЕМАТИЧЕСКИХ МЕТОДОВ

Как мы видели, наука сильна своим операционным и прагматическим подходом. Все, что «работает», годится, по крайней мере в настоящее время. Противоположный по своему характеру подход, присущий науке (его можно назвать эстетическим), — это поиск логического единства, стройности и простоты. Именно этот поиск служит постоянным стимулом новых идей как в области теоретических исследований, так и в области эксперимента. Присущи ли вселенной внутренняя простота и единство или же эти допущения можно рассматривать лишь как удобные первые приближения для пытливых умов — все это составляет хорошо известный предмет философских споров. Каков бы ни был ответ, несомненно одно — что важнейшими составными частями научного процесса в целом является поиск общих законов и объяснений, устранение очевидных противоречий, построение моделей, дающих полное описание различных частных явлений, разработка теорий, имеющих универсальное применение и справедливых в общем случае, и т. п. Существует определенное диалектическое единство мысли и действия, которое можно сознательно и целеустремленно использовать.

Часто обнаруживается, что исследования, ведущиеся на стыке нескольких научных дисциплин, оказывают особенно плодотворное воздействие на широкий круг проблем. Именно такой областью является генетика, которая связана не только с наследственностью, но и с физиологией, биохимией, экологией, эмбриологией, психологией, патологией, судебной медициной и т. д., не говоря уже о множестве математических методов, которые были разработаны и используются в генетических исследованиях. Таким предметом является и сама прикладная математика. Попытки решить частные задачи, возникающие в каких-то конкретных

обстоятельствах, могут привести к разработке общих математических идей, которые не только позволяют решить первоначальные задачи, но и оказываются применимыми в совершенно иных областях. Так называемые уравнения математической физики и связанные с ними специальные функции находят применение в самых различных биологических и социологических задачах: гипергеометрические функции и многочлены Якоби встречаются в теории распространения эпидемических заболеваний; функции Бесселя позволяют исследовать процессы массового обслуживания в медицинских учреждениях; гамма-функции используются для определения наиболее вероятного диагноза.

Значение математики для науки заключается в том, что она дает исключительно точный язык и систему понятий, позволяющие исследовать самые разнообразные вопросы. Мы начинаем с простого счета и измерения, а затем постепенно вводим все более сложные идеи, соответствующие требованиям стоящих перед нами задач. В предыдущих двух главах мы рассмотрели некоторые возможные приложения математики в биологии и медицине. Теперь уже должно быть ясно, что, хотя в некотором смысле математика является единым предметом, она охватывает очень широкую область идей, понятий, приемов и методов. Эти возможности математики уже широко используются в точных науках, особенно в физике, но значение ее для биологии и медицины еще только начинают оценивать по достоинству.

Математический подход не только облегчает точное количественное описание определенной задачи путем построения той или иной подходящей модели, но и дает (или может дать) средство к решению этой задачи. В разд. 3.2 мы уже рассмотрели в общем плане значение математических моделей. Трудности, возникающие в различных частных случаях, более полно будут рассмотрены во второй части книги. В данный момент следует подчеркнуть, что центральной проблемой моделирования является построение самой модели. Как правило, достаточно четкая формулировка задачи всегда дает возможность получить решение в той или иной форме. Математические уравнения могут быть неразрешимы аналитическим путем, однако ценные с научной точки зрения результаты можно получить и с помощью соответствующих аппроксимаций, численных расчетов для большого числа частных случаев, моделирования и т. д. Если же задача сформулирована неудовлетворительно или принята модель недостаточно реалистична, то при любом количестве абсолютно точных математических выкладок будет получен ошибочный результат. Основной проблемой прикладной математики является выбор первоначальной математической модели, и ни в одной области знания это не чувствуется так остро, как в биологии и медицине.

Итак, допустим, что построена некая приемлемая и имеющая смысл математическая модель, позволяющая начать детальные исследования. Для того чтобы их выполнить, необходимо знакомство с методами, пригодными для этой цели. Вероятно, лучше всего поручить эту работу профессиональному математику. Целесообразно вводить его в курс дела еще в процессе построения модели, так как тогда он полнее уяснит себе содержание задачи и ему легче будет интерпретировать получаемые результаты. Однако нередко биологи и врачи, имеющие математическую подготовку, предпочитают хотя бы часть работы выполнять самостоятельно, обращаясь за консультацией к специалистам в области биометрии и математической биологии. Тем, для кого математика не является основной специальностью, полезно еще раз напомнить, что существует большое число разнообразных и дополняющих друг друга методов решения математических задач. Какой бы безнадежной с математической точки зрения или недоступной ни казалась модель, почти всегда имеется какой-либо способ ее исследования.

В качестве примера, иллюстрирующего диапазон методов, имеющихся в нашем распоряжении, кратко остановимся на одной задаче теории эпидемий, более детально описанной в разд. 9.2.

Рассмотрим изолированную группу равномерно общающихся друг с другом восприимчивых индивидуумов. В эту группу попадает единственный зараженный индивидуум, в результате чего заболевание начинает распространяться. В простейшем случае допускается, что случаи выздоровления не имеют места (или наступают через очень длительный промежуток времени и ими можно пренебречь, поскольку рассматриваются начальные стадии болезни). Не представляет труда построить простую стохастическую модель, до некоторой степени аналогичную простой модели процесса размножения (разд. 2.4). Внутреннюю динамику процесса описывает система дифференциально-разностных уравнений, подобная системе (2.6). Точные аналитические решения этой системы возможны, но они очень сложны. Поскольку речь идет о существующих методах решения, одним из них является последовательное решение дифференциальных уравнений, начиная с простейшего. К сожалению, при этом для отыскания решения в общем виде требуются очень громоздкие и длительные математические выкладки. Более разумный путь — использование преобразований Лапласа искомых вероятностей, так как оперировать с ними весьма удобно; правда, при обратном переходе к вероятностям возникают определенные трудности, но имеются некоторые математические приемы, позволяющие облегчить эту работу. Другой возможный подход — применение какой-либо производящей функции (такого многочлена от переменной x , что коэффициент при x^r равен $p_r(t)$, т. е. вероятности того, что в момент t

будет наблюдаться r случаев заболевания). Это позволяет привести систему дифференциально-разностных уравнений к одному дифференциальному уравнению в частных производных параболического типа второго порядка, которое может быть точно решено, хотя и с некоторыми трудностями. С помощью этого единственного уравнения можно, вероятно, определить многие свойства исследуемого процесса, однако это еще дело будущего.

Кроме упоминавшихся здесь чисто аналитических исследований (которые подходят для простой эпидемии без случаев выздоровления, но неприемлемы для большинства более реалистических моделей), можно также рассмотреть численные решения различных уравнений. Этот подход приобретает тем большее значение, чем сложнее модель. Так, систему дифференциальных уравнений можно решить, используя стандартные методы численного интегрирования, например метод Рунге — Кутты. Кроме того, численные решения дифференциальных уравнений в частных производных для производящей функции можно получить с помощью релаксационного метода или итерационного метода Кранка — Никольсона и т. д. Все эти методы требуют применения электронных вычислительных машин, но даже и это далеко не всегда обеспечивает требуемую точность. Еще одна возможность, для практического использования которой также необходима вычислительная техника, — это моделирование, описание которого давалось в разд. 2.5. Моделирование имеет свои недостатки, однако это очень гибкий метод, и в будущем он может получить более широкое распространение. Вопросы, связанные с применением вычислительной техники при моделировании, более подробно рассматриваются в разд. 5.5.

Из всего сказанного видно, сколь велико число различных способов получения математического решения любой данной задачи. Что именно можно сделать в каком-либо конкретном случае, зависит от математических знаний и навыков исследователя и от наличия вычислительной техники. Такое многообразие методов служит гарантией того, что в будущем настойчивые усилия будут вознаграждены хотя бы частичным успехом.

Еще одно следствие широты математической теории состоит в том, что не только существует большое число способов решения данной задачи, но и сама задача может быть сформулирована различными способами, с использованием различных понятий, что в высшей степени полезно. В разд. 3.3 мы уже рассматривали различные методы получения выводов на основе имеющихся данных. В частности, мы отмечали важность применения в отдельных случаях теории статистических решений, которая создает вполне реалистическую основу для расширения исследований сложных административных и организационных проблем (не

говоря уже о более частных приложениях к статистической теории). Таким образом, новый подход, основанный на новых понятиях, обеспечивает значительное продвижение вперед.

Возможность переноса идей из одной области приложений в другую всегда крайне привлекательна. Если это делается умело, то часто дает новые плодотворные результаты. В то же время необходимо проявлять осторожность и не «вгонять» исследуемый материал в неподходящие для него рамки. Как известно, статистический подход к планированию и анализу экспериментов находит особенно широкое применение в агрономической науке, где он существенно необходим и в высшей степени эффективен. Многие биологические эксперименты, проводимые как в полевых, так и в лабораторных условиях, по своему общему характеру не очень сильно отличаются от агрономических. В то же время, например, при клинических испытаниях лекарственных препаратов возникает ряд дополнительных трудностей, которые не всегда легко разрешить в рамках методов, применяемых в агрономии. Некоторые из этих трудностей связаны с множеством побочных эффектов и возможными отдаленными последствиями, проявляющимися при назначении того или иного препарата; другие связаны с моральными и этическими проблемами, возникающими при проведении экспериментов на человеке, и необходимостью найти и применить возможно быстрее наилучший метод лечения в каждом конкретном случае (см. разд. 2.3).

В последние годы много говорят о применении теории информации. Ни один инженер, физиолог или психолог не сомневается в исключительной важности этого предмета. При решении вопросов, связанных с передачей информации в обычном смысле слова, значительное преимущество проведения точного анализа с помощью моделей, в которых используется математическое понятие количества информации, очевидно. Споры ведутся о том, в какой степени понятия теории информации можно применять в совершенно иных областях исследования. Остановимся, например, на некоторых вопросах, изложенных в трудах симпозиума по применению теории информации в биологии [66]. Так, при изучении кодирующей роли ДНК в синтезе белка может быть принят криптографический подход. Такие задачи связаны главным образом с хранением и передачей информации (надлежащим образом определенной и измеренной). Оказалось, что нередко при переходе с одного уровня организации на другой объем информации меняется на целый порядок. К объяснению этого явления можно подойти с самых разных позиций. Можно спросить себя, означает ли это, что в таких условиях передача информации неизбежно происходит при наличии шума и потому точная передача обеспечивается лишь при большой избыточности; или что структуры,

которые определяют конфигурации, необходимые для основных биологических функций, поставляют, так сказать, в качестве побочного продукта и ряд других, «бессмысленных» конфигураций; или, наконец, что дополнительная информация каким-то образом связана с обработкой основного сообщения, а не с самим сообщением и т. д. Конечно, ни один из этих подходов не заменяет точного биологического исследования на молекулярном уровне. Однако применение теории информации позволяет комбинировать большое число количественных результатов и использовать их для теоретических исследований, которые будут уже не носить преимущественно словесный характер, а основываться на некоторой системе законов и принципов.

С точки зрения потери информации можно также изучать такие проблемы, как заболевания дегенеративного характера, лучевое повреждение, старение и смерть. Как известно, организмы могут жить и размножаться, несмотря на некоторые дефекты в той генетической информации, которую они в себе несут. В результате увеличения числа таких ошибок под влиянием облучения организмы, которые ранее были близки к летальному пределу, но все же обладали некоторой жизнеспособностью, могут вообще оказаться нежизнеспособными. Это положение можно сформулировать в понятиях теории информации, предположив, что под влиянием облучения повышается частота мутаций, или, что равнозначно, уровень шума в канале связи. Вследствие этого пропускная способность канала уменьшается в такой степени, что у некоторых организмов передача достаточно точных сообщений оказывается невозможной.

Информационный подход в сочетании с принципом отрицательной обратной связи позволил разработать большое число интересных моделей биологических систем управления. Такие модели особенно полезны в приложении к физиологии, где они позволяют выяснить многие вопросы, касающиеся механизма гомеостаза. На этом фундаменте построена новая наука — кибернетика, охватывающая любые процессы управления в самых разнообразных системах — технических, биологических и социальных [2, 64].

Вряд ли можно утверждать, что те приложения теории информации, о которых говорилось выше, позволили получить много таких результатов в биологии и медицине, которые не были бы уже получены другими способами. И все же понятие информации, без сомнения, ценно тем, что оно служит полезным количественным инструментом для теоретических построений и описаний, а также унифицирует идеи и задачи, возникающие в самых различных областях знания. Как минимум это должно стимулировать научные исследования, а в идеале может привести к новым результатам, получение которых маловероятно при других формах анализа.

Все математические модели, о которых здесь упоминалось или которые рассматривались в этой книге, имеют явно численную основу. Эти модели выведены на основе простых идей счета и измерения, в большинстве случаев характеризуются не очень большой степенью абстракции и содержат описания или формулы, которые довольно легко представить в наглядном виде. Последнее свойство абсолютно несущественно. Обычно оно помогает на начальном этапе развития науки, но при переходе к более сложным исследованиям может даже служить помехой (во многих разделах современной физики дело обстоит именно так, но сомнительно, чтобы это относилось и к биологии). Счет и измерение находятся на первой ступени абстракции, которая дает начальный импульс развитию математики. Однако значительная часть математики находится на более высоких уровнях абстракции. Первым этапом после чисто метрических исследований является изучение соотношения между объектами, т. е. изучение их в реляционном аспекте. Большинство разделов геометрии носит именно такой характер. На более сложных уровнях находятся теория групп, теория множеств и топология.

Рапешевский [54, 55] и его школа сделали очень много для применения реляционных методов в биологии. Математический аппарат, применяемый в этой области, довольно тонок и сложен, и детальное его обсуждение выходит за рамки настоящей книги. Мы только отметим, что с его помощью удалось сформулировать ряд общих представлений и теорем, имеющих большое биологическое значение. Хотя формулировка этих теорем носит преимущественно словесный характер и потому может создаться впечатление, что они были получены на основе неточных, чисто описательных рассуждений, большая их ценность состоит в том, что они являются результатом строго математического вывода. Так, основанная на чисто топологических соображениях теорема Розена гласит, что в любом организме, каким бы простым или сложным он ни был, имеются некоторые компоненты, регенерация которых невозможна. Например, при наличии ядра цитоплазма отдельной клетки может регенерировать, однако разрушенные ядра не способны к регенерации. Розен доказал также другую теорему: если некоторый компонент организма, не способный к регенерации, получает входные воздействия непосредственно из внешней среды, то утрата этого компонента приведет к гибели всего организма. Применение топологического подхода при изучении головного мозга [67] дало возможность получить довольно общую абстрактную модель различных функций мозга, позволяющую по-новому взглянуть на важные реляционные аспекты поведения.

Подлинное значение этих более абстрактных математических исследований еще не получило достойной оценки. Простые модели легко описать, объяснить и проверить. Но чем больше степень абстракции, тем труднее понять, в какой степени связаны между собой модель и реальное явление и не выхолощена ли эта модель настолько, что перенос любых выводов, полученных с ее помощью, на реальную действительность будет малооправданным. Если через некоторое время окажется, что математические модели реляционного типа дают полезные результаты (а это весьма возможно), то, значит, затраченные усилия не пропали даром.

Глава 4

МЕТОДЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

4.1. СОДЕРЖАНИЕ ПРЕДМЕТА

Хотя научный подход в целом и математические методы в частности применимы почти в любой области, связанной с систематизацией и изучением фактов, все же удобно провести некоторые различия. Например, существует традиционное деление наук на чистые и прикладные. Провести четкую границу между ними невозможно, но в широком смысле принято считать, что чистые науки занимаются преимущественно изучением явлений природы, а прикладные — применением полученных знаний на практике для покорения природы и использования ее ресурсов для нужд человека. В обоих случаях научное содержание (т. е. физика, химия, биология явлений) играет фундаментальную роль и должно быть предметом точного исследования.

Однако, для того чтобы полезные результаты прикладных исследований стали доступны обществу в целом, необходимо решить ряд проблем экономического и организационного характера. Финансирование промышленных предприятий, распределение и сбыт продукции, реклама, служба информации, общественное планирование и планирование в масштабе фирмы и т. д. — таковы лишь некоторые из аспектов, которые необходимо принимать во внимание. Многие из этих видов деятельности в большей мере связаны с организационным управлением ресурсами (как людскими, так и материальными), чем с природой самих ресурсов. Короче говоря, здесь основную роль играют проблемы, с которыми сталкиваются исполнительные и административные органы. Всю эту область можно сделать предметом специального научного исследования и объединить под названием «исследование операций»¹. Определение исследования операций как совокупности научных методов изучения организационных и адми-

¹ Отметим, что такая терминология никоим образом не является общепринятой. Чаще «исследованием операций» называют определенную совокупность задач управления, допускающих математическое моделирование, которая рассматривается как часть теории организационных решений.—
Прим. ред.

нистративных проблем, безусловно, очень ценно, так как подчеркивает важное отличие всей этой области знания как от чистых, так и от прикладных наук. Отдельным разделам этой области уже давно уделяется серьезное внимание. Сюда можно отнести исследования по таким традиционным вопросам, как изучение рабочих операций (тейлоризм) и производства, организация управления, научные методы административного управления, инженерная психология и т. д. Совсем недавно появились такие новые дисциплины, как эргономика и кибернетика. В нашу задачу не входит определение и систематизация всех этих различных предметов. Некоторые из них имеют четкие границы, другие перекрываются друг с другом; для одних характерна определенная область приложения, для других — используемые методы. Основная задача данной главы состоит в том, чтобы привлечь внимание к большой и быстро развивающейся группе научных методов, связанных больше с формами организации, чем с конкретными объектами, подлежащими организации. Естественно, что оба эти аспекта — форма и содержание — взаимосвязаны, и для успешного применения рассматриваемых методов необходимо достаточно хорошо знать и то и другое. Методы исследования операций играют все возрастающую роль в практических мероприятиях, связанных с использованием теоретических знаний в области биологии, и их значение при поисках оптимальной организационной структуры медицинского обслуживания уже общепризнано.

Часто в литературе используется термин *системотехника*, который мы пока не вводили. Некоторые авторы употребляют его как синоним исследования операций, а другие проводят между этими терминами четкое различие. Например, можно сказать, что специалисты в области исследования операций в большей мере связаны с процедурами, которые уже существуют, тогда как специалисты в области системотехники больше занимаются комплексом «человек — машина» в целом, особенно в связи с проектированием и разработкой новых систем. Разные авторы придают этим терминам разный смысл. Во многих случаях важно отличать исследования, связанные с организацией и функционированием целой системы процессов, от детального исследования работы какой-либо отдельной части этой системы. Так, изучение порядка приема амбулаторных больных можно рассматривать как исследование операций, а термин «системотехника», или *анализ систем*, целесообразно употреблять при рассмотрении всей системы медицинского обслуживания в каком-то большом районе или в целой стране. В этой книге не имеет особого смысла проводить такое различие, и мы будем считать, что любой научный подход к изучению функционирования той или иной системы относится к исследованию операций.

Административные и организационные проблемы независимо от их конкретного содержания имеют между собой очень много общего. Рассмотрим абонентов, ожидающих телефонного вызова; клиентов, стоящих в очередях в банке или в почтовом отделении; самолеты, выстроившиеся над аэродромом в эшелон в ожидании разрешения на посадку; автомобили, скопившиеся на дороге в часы пик; больных, подлежащих госпитализации в порядке очереди, или амбулаторных больных, ожидающих в приемной вызова на консультацию, которая им была назначена. Во всех этих ситуациях имеется нечто общее, а именно существует определенная конъюнктура спроса (потребности в некотором виде обслуживания) и определенная конъюнктура предложения (обслуживания), и их взаимодействие приводит к образованию очередей, размеры которых колеблются случайным образом. В связи с этим математическая *теория массового обслуживания* играет важную роль при использовании метода исследования операций для изучения любых проблем, связанных со спросом и предложением.

Далее, во многих промышленных процессах приходится решать по существу одинаковые задачи оптимизации, связанные с выпуском той или иной продукции. Например, при крекинге нефти могут получаться различные количества бензина, парафинов, мазута и т. д. Допустим, что нам известны соотношения между выходом различных продуктов, а также их рыночные цены и стоимость сырой нефти. В этом случае важно найти такое оптимальное соотношение между выходами каждого продукта, при котором фирма получит максимальную общую прибыль. Аналогичную теоретическую форму имеет задача отыскания такой схемы распределения товаров между данным числом предприятий и складов, при которой затраты на перевозки будут минимальными. Такого рода задачи на оптимизацию обычно можно решить с помощью какого-либо *метода математического программирования*, например с помощью *линейного программирования*.

Существует целый ряд математических методов, играющих важную роль в исследовании операций, и, более того, многие из них получили значительное развитие именно вследствие того, что они находят применение в этой области. Более подробно некоторые из этих теоретических методов рассматриваются в разд. 4.3.

Методы исследования операций и системотехники подробно рассмотрены в превосходной работе Флегля, Хаггинса и Роя [25], а более конкретное описание различных математических методов и их приложений можно найти в книге Сасени, Яспана и Фридмана [57].

4.2. ВОЗНИКНОВЕНИЕ И РАЗВИТИЕ МЕТОДОВ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Возникновение методов исследования операций, несомненно, восходит к работе Тейлора, в которой он употребляет термин «научные методы организационного управления» (scientific management). Эта работа была опубликована в США еще до первой мировой войны. В период между двумя мировыми войнами были достигнуты определенные успехи в ряде областей (разработка методов учета, изучение рабочих операций, статистический контроль качества и т. д.), связанных с совершенствованием количественных методов принятия решений при организационном управлении. В начале второй мировой войны интерес к этим методам, которые вскоре стали объединять под названием «исследование операций», резко повысился. Сразу стало ясно, какое кардинальное значение имеет теория, позволяющая максимально быстро и эффективно принимать оптимальные решения по самым разнообразным военным проблемам (боевое использование радиолокационных станций и зенитной артиллерии, выбор состава морского конвоя, планирование налетов бомбардировочной авиации, разработка противолодочного оружия и т. д.). В каждой из этих областей основное внимание уделялось, естественно, использованию самых последних научных и технических достижений. Но оставалась важнейшая проблема — как наилучшим образом использовать существующие образцы аппаратуры, оборудования или боевой техники. Была организована группа под руководством П. Блэккетта (она получила известность под названием «группа Блэккетта»), состоявшая из специалистов различного профиля: в нее входили несколько математиков, физиков и биологов, а также астрофизик, топограф и офицер сухопутных войск. Такое многообразие специальностей служило гарантией того, что при разработке той или иной проблемы внимание не будет сосредоточено на каком-либо одном научном или техническом ее аспекте, и обеспечивало надлежащее рассмотрение и оценку противоречивых требований и условий. Группа Блэккетта исследовала многие из упомянутых выше задач, и часто о ней говорят как о прототипе современной группы исследования операций.

В качестве иллюстрации применения методов исследования операций к проблемам военного времени рассмотрим проблему, которой автор занимался сам. В течение долгого времени было принято определять возможное местоположение радиопередатчика путем засечки направлений на источник сигналов, производимой из нескольких удобных точек. Если бы направления определялись без ошибок, то соответствующие прямые проходили бы через фактическое место расположения радиопередатчика и требовалось

бы не более двух отсчетов. К сожалению, направления, особенно в случае высокочастотных сигналов, измеряются со значительной неопределенностью. Поэтому радиопередатчик трудно запеленговать достаточно точно, даже если произвести большое число отсчетов. Разумеется, повышение эффективности и улучшение рабочих характеристик аппаратуры пеленгования до определенной степени помогают преодолеть некоторые из этих трудностей. Однако от таких воздействий, как сильные непрогнозируемые ионосферные возмущения, избавиться с помощью технических средств невозможно; в лучшем случае их можно лишь учесть, обрабатывая данные соответствующими статистическими методами.

Обычно направления наносят на карту с прямоугольной системой координат, на которой они имеют вид прямых линий и поэтому их легко представить с помощью нитей, надлежащим образом закрепленных на обоих концах. Затем изучается область пересечения этих нитей на карте и на глаз оценивается наиболее вероятное место расположения радиопередатчика; одновременно прикидывается точность этой оценки. Как ни парадоксально, но такие визуальные методы обладают той особенностью, что при небольшом числе измерений, произведенных под большими углами, они могут оказаться поразительно точными, однако при увеличении числа отсчетов работать становится значительно сложнее и картина получается более запутанной, хотя объем полезной информации соответственно возрастает.

Путем построения математической модели этого процесса удалось найти вычислительный метод, с помощью которого можно было рассчитать наиболее вероятное место расположения радиопередатчика, а также определить прямоугольник, внутри которого он находится с некоторой заданной вероятностью (скажем, 90%). Специально подготовленные карты и математические таблицы позволяли довольно быстро выполнять необходимые вычисления (даже при большом числе отсчетов) с помощью обычной настольной счетной машины. В настоящее время такие вычисления можно выполнять еще быстрее и точнее на электронных вычислительных машинах.

Вычисленные направления на передатчики, место расположения которых было известно, сравнивали с визуальными отсчетами, чтобы оценить относительную точность этих двух методов. По мере совершенствования механизированных методов все очевиднее становилась целесообразность их применения во многих случаях. В то же время встречались важные ситуации, в которых полученные данные оказывались неопределенными или требовалось принять решение немедленно. В этих случаях приходилось полагаться на субъективные суждения, однако теперь эти суждения можно было корректировать и уточнять с помощью вычисли-

тельных методов. Таким образом, имело место плодотворное взаимодействие между интуицией человека и математическими методами. Очень похожая ситуация возникает при использовании вычислительной машины для медицинской диагностики (гл. 11).

Параллельно развитие методов исследования операций в годы войны происходило также в США. Оно охватывало такие вопросы, как разработка противолодочного оружия и оценка эффективности действия военно-воздушных сил. С этого времени группы исследования операций начали организовывать во всех военных министерствах США¹.

После войны область применения методов исследования операций значительно расширилась. Эти методы применяют теперь в самых разнообразных промышленных и общественных организациях и предприятиях: на железных дорогах, в больницах и универсальных магазинах, в редакциях газет, в организациях, ведающих снабжением углем и электроэнергией, проектированием и строительством зданий и т. д.

Очевидно, что методы исследования операций должны найти широкое применение и при решении ряда биологических и медицинских проблем. В таких областях, как лесоводство, сельское хозяйство и рыболовство, большое место занимает организационная и административная деятельность. Для достижения оптимальных результатов в организации научного руководства лесным хозяйством, сельскохозяйственным производством или инкубаторными станциями математические методы, обычно используемые в теоретических и прикладных науках, недостаточны. Проектирование больниц, организация их деятельности и вообще эффективное оказание медицинской помощи также связаны с решением проблем, которые по своему характеру не являются чисто медицинскими. В применении методов исследования операций при решении проблем, связанных с медицинским обслуживанием, достигнуты значительные успехи, и в гл. 12 детально рассмотрены некоторые работы, выполненные в этой области.

4.3. НЕКОТОРЫЕ МЕТОДЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ

Суть метода исследования операций заключается в применении научного метода для решения организационных задач: однако часто термин «исследование операций» относят к группе математических методов, оказавшихся особо ценными при решении большинства таких задач. Например, в связи с широким спросом на определенные виды товаров или услуг и необходимостью удо-

¹ В США три военных министерства: армии, авиации и военно-морских сил.— *Прим. перев.*

влетворения этого спроса возникают задачи массового обслуживания, определения очередности, составления расписаний, управления ресурсами и т. д. Приступая к решению какой-либо конкретной задачи, прежде всего необходимо построить математическую модель соответствующей операции и исследовать ее свойства. Затем требуется найти метод определения оптимальной (по отношению к соответствующим критериям) схемы протекания реального процесса. О теории массового обслуживания и линейном программировании уже кратко говорилось в разд. 4.1. Далее мы несколько подробнее остановимся на вопросах, представляющих особый интерес и играющих важнейшую роль в исследовании операций.

Теория массового обслуживания

Любая ситуация массового обслуживания связана с наличием непрерывной, хотя и меняющейся во времени, потребности в определенных видах товаров или услуг и организацией обслуживания для удовлетворения этих потребностей. В качестве простейшего примера возьмем покупателя, подошедшего к прилавку, чтобы купить какие-то товары. Если в этот момент продавец свободен, то покупатель будет обслужен немедленно. Но если продавец уже занят обслуживанием покупателя, пришедшего раньше, то вновь прибывший вынужден ждать, и обычно для этого он должен встать в очередь. Этот простой случай может иметь большое число вариантов, обобщений и усложнений. В настоящее время уже существует обширная литература по теории массового обслуживания. Библиографию и подробное изложение математического аспекта предмета читатель найдет в книгах Кокса и Смита [15], Саати [56] или в гл. 11 другой книги автора [9]. В действительности описанная выше предельно простая ситуация, отнесенная нами к области интересов массового обслуживания, представляет собой лишь частный случай более общих процессов преобразования входных величин в выходные. К таким процессам относятся, например, регуляция спроса и предложения в системе медицинского обслуживания или регуляция обмена веществ в организме. Однако вернемся к элементарной модели массового обслуживания и отметим ее характерные свойства.

Во-первых, рассмотрим *входной процесс*, представляющий собой поток клиентов, нуждающихся в обслуживании. Часто бывает удобно и целесообразно предположить, что они прибывают случайным образом, хотя и с некоторой постоянной средней интенсивностью. Это означает, что длительности интервалов между последовательными моментами прибытия клиентов имеют экспоненциальное распределение и что число клиентов, прибываю-

щих в течение любого фиксированного промежутка времени, имеет пуассоновское распределение¹. Можно представить себе также систему предварительной записи, когда клиенты прибывают в фиксированные моменты времени. Кроме того, существует целая область промежуточных возможностей.

Во-вторых, необходимо задать определенную *дисциплину очереди*. Простейшим допущением является принцип «первым пришел — первым обслужен». В этом случае клиенты ожидают в очереди в порядке прибытия, и в любой данный момент времени обслуживается только тот, кто стоит первым. Более сложные виды дисциплины очереди могут быть заданы с помощью определенной системы приоритетов или возникнуть при поступлении срочных «заказов», например таких, как неотложные случаи, требующие немедленного оказания медицинской помощи.

В-третьих, существует *механизм обслуживания*, определяющий выходной поток или окончание элементарного процесса обслуживания. Здесь прежде всего нужно знать число обслуживаемых устройств и распределение длительности обслуживания в каждом из них. Естественно, что простейшей является система с одним обслуживаемым устройством.

Очевидно, что если бы клиенты прибывали через постоянные промежутки времени и время обслуживания было постоянным, то при длительности обслуживания, не превышающей длительности интервала между прибытием последовательных клиентов, очередь не образовалась бы; в противном случае очередь будет увеличиваться бесконечно с постоянной скоростью. В таком простом детерминированном случае никаких сложных проблем не возникает, но, как только вводятся те или иные вероятностные элементы, появляется ряд специфических особенностей.

Для закрепления этих понятий рассмотрим одноканальную систему массового обслуживания (систему с одним обслуживаемым устройством), в которой клиенты прибывают случайным образом; пусть средний интервал между моментами прибытия равен λ и длительность обслуживания характеризуется некоторым распределением с математическим ожиданием μ и дисперсией σ^2 . Здравый смысл подсказывает, что при $\mu > \lambda$ потребность в обслуживании превышает пропускную способность системы, длина очереди будет непрерывно возрастать и создастся крайне неблагоприятная ситуация. Если же $\mu < \lambda$, то средняя потребность в обслуживании будет меньше средней возможной пропускной

¹ Это обычные, но необязательные допущения теории массового обслуживания. Для того чтобы они выполнялись, нужно еще предположить, что вероятность появления очередного клиента никак не зависит от того, что происходило до его появления. — *Прим. ред.*

способности системы; поэтому ситуация будет довольно благоприятной, однако вследствие статистических колебаний временами будут возникать очереди. Ясно, что точное соотношение между длиной очереди и параметрами λ и μ требует исследования. Частный случай $\lambda = \mu$, когда средняя пропускная способность и средняя потребность в обслуживании равны, представляет особый интерес и обладает совершенно неожиданными свойствами.

С помощью довольно элементарных рассуждений можно показать, что средняя длина очереди \bar{q} описывается формулой

$$\bar{q} = \frac{\rho(2-\rho) + \sigma^2/\lambda^2}{2(1-\rho)}, \quad (4.1)$$

где $\rho = \mu/\lambda$ — так называемая *загрузка системы*. Если длительность обслуживания постоянна, то $\sigma^2 = 0$ и формула (4.1) принимает вид

$$\bar{q} = \frac{\rho(2-\rho)}{2(1-\rho)}. \quad (4.2)$$

Если же длительность обслуживания имеет экспоненциальное распределение, то дисперсия σ^2 будет равна $\mu^2 = \rho^2\lambda^2$. В этом случае формула (4.1) принимает вид

$$\bar{q} = \frac{\rho}{1-\rho}. \quad (4.3)$$

Можно также вычислить среднее время \bar{w} , в течение которого клиенту приходится ждать, когда его начнут обслуживать. Можно показать, что общее выражение для \bar{w} имеет вид

$$\bar{w} = \frac{\rho^2 + \sigma^2/\lambda^2}{2(1-\rho)/\lambda}. \quad (4.4)$$

При постоянной длительности обслуживания

$$\bar{w} = \frac{\rho\mu}{2(1-\rho)}, \quad (4.5)$$

а в случае экспоненциального распределения длительности обслуживания

$$\bar{w} = \frac{\rho\mu}{1-\rho}. \quad (4.6)$$

Сравнение формул (4.5) и (4.6) показывает, что при переходе от постоянной длительности обслуживания к экспоненциальному распределению этой величины средняя длительность ожидания увеличивается вдвое.

Разумеется, формулы (4.1) — (4.6) справедливы только при $\rho < 1$. Если ρ приближается к единице, то средняя длина очереди и средняя длительность ожидания стремятся к бесконечности. Таким образом, когда спрос и предложение в точности сбалан-

сированы, возникает крайне нежелательная ситуация. Хотя в настоящее время имеется огромная литература по математической теории массового обслуживания, одним из простейших и наиболее ценных практических результатов является требование, чтобы загрузка системы была меньше единицы (к этому результату вполне можно было бы прийти просто на основании здравого смысла). Главная ценность теории массового обслуживания состоит в том, что она позволяет точно вычислить, сколько продлится среднее ожидание при любой данной загрузке системы. Для наиболее эффективного использования времени обслуживающего устройства загрузка должна быть как можно ближе к единице, тогда как для обеспечения минимального времени ожидания она должна быть как можно меньше. Следовательно, если имеется возможность рассчитать точные последствия различных методов эксплуатации системы, то это позволяет спланировать оптимальное соотношение между этими противоречивыми требованиями.

Аналогичные соображения в принципе применимы и к более сложным системам, например к многоканальным системам, к системам, в которых дисциплина очереди предполагает применение специальных правил установления приоритета, и системам с входными потоками произвольного или сложного характера. Некоторые конкретные приложения теории массового обслуживания к медицинскому обслуживанию рассматриваются в гл. 12.

Математическое программирование

В большинстве организационных задач, решаемых методами исследования операций, существенную роль играют экономические аспекты. Эти аспекты могут быть выражены более или менее неявно, как в только что рассмотренном случае, где речь шла об экономии времени клиентов и об экономичном использовании рабочего времени обслуживающего устройства. Если бы потерю времени клиентами и обслуживающим устройством можно было выразить какими-то денежными показателями, то удалось бы провести общий анализ, показывающий чистые издержки для общества. Однако в общем случае отыскание наилучшего соотношения между противоречивыми требованиями, которые трудно выразить количественно в единой шкале, основывается на интуитивном суждении.

Но во многих задачах экономический аспект выражен в явном виде уже в исходной формулировке. Рассмотрим типичную производственную задачу, характерную и для многих других областей, где основной проблемой является оптимальное распределение ресурсов. Допустим, что главной проблемой, стоящей перед некоторой фирмой, является получение максимальной прибыли от продажи

ряда товаров при различных ограничениях. Пусть имеется n различных видов товаров, и пусть выпуск i -го товара за некоторый определенный промежуток времени составляет x_i единиц. Если прибыль на единицу i -го товара равна c_i , то выражение для общей прибыли z имеет вид

$$z = \sum_{i=1}^n c_i x_i. \quad (4.7)$$

Если какое-либо значение c_i отрицательно, то соответствующий товар продается с убытком.

Вообще говоря, нам необходимо найти значения x_i , максимизирующие z при данных значениях c_i и при определенных условиях, которые вытекают из ограничений, налагаемых производственным процессом. Прежде всего, разумеется, значения x_i не могут быть отрицательными. Таким образом, одна группа ограничений имеет вид

$$x_i \geq 0, \quad 1 \leq i \leq n. \quad (4.8)$$

Допустим далее, что общий объем имеющихся ресурсов постоянен и составляет b человеко-часов и что для производства одной единицы i -го товара необходимо a_i человеко-часов. Следовательно, общее количество ресурсов, необходимое для производства количеств товара x_1, x_2, \dots, x_n , равно

$$\sum_{i=1}^n a_i x_i$$

и не может превышать b , т. е.

$$\sum_{i=1}^n a_i x_i \leq b. \quad (4.9)$$

Аналогичные соображения применимы также к числу часов машинного времени определенного станка, запасам различных материалов, ресурсам, имеющимся для распределения и продажи, и т. д. Если имеется всего m линейных ограничений, то их можно выразить в виде следующего обобщения формулы (4.9):

$$\sum_{i=1}^n a_{ij} x_i \leq b_j, \quad 1 \leq j \leq m. \quad (4.10)$$

Таким образом, математическая задача состоит в том, чтобы максимизировать функцию z , заданную формулой (4.7), при ограничениях, налагаемых выражениями (4.8) и (4.10). В этом довольно простом варианте задачи, часто встречающейся на практике, как функция z , так и ограничивающие условия являются линейными функциями x_i . Такая модель называется *моделью линейного программирования*. Для решения задач такого рода существует множество стандартных методов, но если число пере-

менных или ограничений велико, то на практике может потребоваться значительный объем алгебраических выкладок. К счастью, разработаны программы для электронных вычислительных машин, позволяющие получить решение таких задач.

Мы не будем рассматривать здесь в деталях теорию построения моделей линейного программирования, хотя приводимая далее геометрическая интерпретация может оказаться очень полезной для наглядного представления некоторых важных особенностей этой задачи. (Исчерпывающее изложение теории линейного программирования читатель может найти в любом руководстве по данному предмету, например в книге Гарвина [28].) Допустим вначале, что имеются только два различных товара в количествах x_1 и x_2 , т. е. $n = 2$. Все возможные количества товаров можно изобразить на плоскости точками с координатами (x_1, x_2) . Так как необходимо, чтобы $x_1 \geq 0$ и $x_2 \geq 0$, сразу же ограничиваемся рассмотрением положительного квадранта плоскости. Рассмотрим далее j -е линейное ограничение $a_{1j}x_1 + a_{2j}x_2 \leq b_j$. Все точки, в которых удовлетворяется это неравенство, лежат в полуплоскости, расположенной по одну сторону прямой $a_{1j}x_1 + a_{2j}x_2 = b_j$ (там же, где и начало координат, если $b_j > 0$). Легко показать, что в общем случае все решения, допускаемые при этих ограничениях, лежат внутри многоугольника, образованного $m + 2$ прямыми:

$$\left. \begin{aligned} x_1 = 0, \quad x_2 = 0; \\ a_{1j}x_1 + a_{2j}x_2 = b_j, \quad 1 \leq j \leq m. \end{aligned} \right\} \quad (4.11)$$

Наконец, функция z постоянна на любой линии $c_1x_1 + c_2x_2 = d$, где она и принимает значение $z = d$. После недолгого размышления обнаруживаем, что наибольшее значение z , допускаемое при наложенных ограничениях, находится на прямой, наиболее далеко отстоящей от начала координат и лишь касающейся многоугольника, заданного формулой (4.11). Точное положение этой прямой (если решение существует) при надлежащем значении d , обеспечивающем максимальное значение z , теперь можно определить путем применения соответствующих вычислительных методов.

Нетрудно обобщить изложенные выше идеи на случай n различных товаров ($n > 2$), а также допустить неравенства вида $\sum a_{ij}x_i \geq b_j$. Основной задачей специалиста в области исследования операций является определение возможности хотя бы приближенно сформулировать конкретную реальную задачу в виде задачи линейного программирования, с тем чтобы можно было применять стандартные вычислительные методы.

Мы умышленно выбрали пример, когда функция z и соответствующие ограничения являются линейными. Однако в экономике часто действует закон уменьшения прибыли, согласно кото-

рому увеличение производства какого-либо товара в какой-то момент приводит к уменьшению прибыли на единицу продукции. Кроме того, может случиться, что прибыли, получаемые при сбыте различных товаров, не независимы, как молчаливо здесь предполагалось, а определенным образом связаны друг с другом. Для того чтобы учесть влияние таких факторов, необходимо выбирать функцию z более общего вида. Если z можно выразить как квадратическую функцию от x_i , на которую налагаются определенные ограничения, являющиеся по-прежнему линейными, то в этом случае говорят, что получена *модель квадратического программирования*.

В более общем случае условие линейности ограничений можно ослабить и рассматривать *модель математического программирования* в следующем виде:

$$\left. \begin{aligned} z &= f(x_1, \dots, x_n); \\ x_i &\geq 0, & 1 \leq i \leq n; \\ g_j(x_1, \dots, x_n) &= 0, & 1 \leq j \leq m. \end{aligned} \right\} \quad (4.12)$$

Очевидно, что в этом случае как теоретическая формулировка задачи, так и характер вычислений могут стать весьма сложными. По возможности желательно иметь дело с моделью линейного программирования, если только она не окажется слишком далекой от реальной действительности.

Еще более сложная задача, имеющая большое значение для практики, состоит в том, чтобы найти оптимальную стратегию для определенного промежутка времени. Допустим, известно, что условия изменяются от месяца к месяцу. Казалось бы, что в этом случае можно решать задачу линейного программирования для каждого месяца в отдельности. Но на самом деле нет никакой уверенности в том, что такой метод вообще даст наилучший результат, так как способ производства в один из месяцев может повлиять в ту или другую сторону на условия работы в следующем месяце. Для решения такого рода задач можно использовать метод *динамического программирования*, основные принципы которого разработаны главным образом в работах Р. Беллмана.

Таким образом, существует ряд мощных математических методов, позволяющих решать самые разнообразные задачи на отыскание оптимальной стратегии. Принятие окончательных решений производится на основе анализа, выполняемого совместно с руководителями или исполнителями работы. Однако самое важное состоит в том, что для исследования значительной части проблем, которые прежде решались лишь на основе индивидуального опыта и интуиции, теперь имеются значительно более точные методы, и это позволяет человеку сосредоточиться на тех вопросах, где его интуиция и опыт действительно незаменимы.

Моделирование и операционные игры

Общее понятие *моделирования* уже рассматривалось в разд. 2.5 в связи с исследованием сложных случайных процессов. Поскольку при исследовании операций реальные ситуации не должны слишком сильно упрощаться (если мы хотим, чтобы эти исследования приносили пользу), соответствующие модели человеческой деятельности почти наверняка будут включать некоторые процессы, получить описание которых крайне трудно. С одной стороны, нам необходима конкретная информация относительно общего характера этих процессов, а с другой — мы должны уметь определять возможные последствия изменения некоторых важных параметров, поддающихся регулированию. В обоих случаях единственно возможным способом достижения успеха является моделирование. Таким образом, оно становится одним из наиболее важных математических методов исследования операций, и поэтому целесообразно несколько подробнее остановиться на его роли.

Допустим, к примеру, что мы хотим изучить работу международного аэропорта, обратив особое внимание на интенсивности потоков пассажиров и их багажа. По существу здесь перед нами сложный многофазовый процесс массового обслуживания, в котором вследствие образования очередей в различных пунктах возможно возникновение узких мест. Пассажиры прибывают в аэропорт главным образом в соответствии с расписанием, определяющим время прилета и отлета самолетов, но это время подвержено значительным статистическим колебаниям. Перед посадкой пассажиры должны зарегистрироваться у соответствующей стойки, оформить багаж, пройти таможенный досмотр и дожидаться объявления о своем рейсе. Самолет может прилететь точно по расписанию, но может и опоздать вследствие задержки в полете или потому, что из-за занятости взлетно-посадочной полосы аэродрома ему не дают разрешения на посадку. Прибывающие пассажиры проходят весь этот процесс в обратном порядке. Построить адекватную стохастическую модель такой ситуации не слишком сложно; для этого требуется, во-первых, детально знать те состояния и этапы, которые надо в нее включить (короче говоря, знать общую схему процесса), и, во-вторых, знать распределения входных потоков и длительность обслуживания в различных пунктах системы. При тщательном обследовании всей системы можно получить эти данные с достаточной степенью точности.

Но одно дело — построить такую модель, и совсем другое — попытаться провести ее теоретический анализ. В общем случае эта задача вряд ли выполнима. Однако путем моделирования на вычислительной машине вполне возможно изучить поведение всей системы в течение относительно длительного промежутка

времени. При этом мы не только получим очень полезное количественное описание системы в процессе ее функционирования, но и сможем предсказать возможные последствия любых предлагаемых переделок и усовершенствований. Очень часто бывает трудно внести конкретные изменения в процедуру или порядок эксплуатации; может потребоваться очень много времени, прежде чем удастся определить, дадут ли они значительно лучшие или худшие результаты по сравнению с предыдущей программой, и в случае неудачи эти изменения могут привести к большим снижениям прибыли и потере доверия у клиентов и персонала. Моделирование предлагаемого нового режима работы позволяет теоретически и совершенно безболезненно решить многие спорные вопросы. Только в том случае, если широкие исследования такого рода покажут, что новый режим может привести к заметному улучшению, следует приступить к его проверке на практике. Конечно, при этом могут возникнуть какие-либо непредвиденные трудности, но, во всяком случае, они будут значительно меньше и не приведут к столь серьезным последствиям, как в том случае, когда моделирование не проводится. Кроме того, совершенно очевидно, что за одну человеческую жизнь можно подвергнуть непосредственной проверке лишь очень небольшое число новых мероприятий. С помощью же быстродействующей электронной вычислительной машины можно получить на модели за несколько часов или даже минут вариант работы реальной системы в течение нескольких лет. Поэтому целесообразно сначала исследовать этим методом широкий круг различных реорганизаций, с тем чтобы потом представить один-два наилучших варианта для окончательной практической проверки.

На значительно более абстрактном уровне находится применение методов физического моделирования в так называемых *операционных играх*, имитирующих ситуации, в которых требуется многократное принятие решений в условиях непрерывного возникновения различных сложных проблем. Только что рассмотренный здесь пример работы аэропорта весьма похож на ситуации, наблюдаемые в промышленном производстве, где различного рода материалы должны обрабатываться стандартным образом как можно быстрее и экономичнее. Если задана некоторая конкретная схема организации и соответствующие ей параметры, то можно рассчитать количественные характеристики результатов работы и, таким образом, найти различие между разными вариантами организации. После того как найдена система, близкая к оптимальной, основной проблемой является поддержание ее в рабочем состоянии. В отличие от такого рода приложений сосредоточим теперь внимание на задачах, в которых первостепенное значение имеет принятие решений. Одна из наиболее

полно изученных задач такого рода — выбор оптимальной тактики и стратегии в различной боевой обстановке. Эта работа выполнялась в США группой тактических военных игр Управления исследования операций.

Очевидно, что проведение любых боевых операций в весьма значительной степени зависит от ряда случайных факторов, и часто бывает крайне трудно выбрать стратегию, которая была бы оптимальной. Некоторые аспекты задачи о работе аэропорта присутствуют здесь в сильно усложненном виде. Так, в большинстве случаев оказывается возможным придерживаться только одной из большого числа разнообразных мыслимых стратегий. Практическая проверка различных вариантов не только сложна и требует много времени, но попросту невозможна. Поэтому в течение долгого времени командиры выбирали наилучшую стратегию и тактику, прикидывая в голове грубые «мысленные модели» или проводя воображаемые боевые операции на карте. Ввиду существования большого числа случайных факторов даже наилучшая стратегия может привести к неудачному исходу, и поэтому проблема получения статистически правильных выводов на основе данных о практических результатах, полученных другими командирами в другой обстановке, невероятно сложна. Однако самое важное заключается в том, что принятие решений на основе здравого смысла представляет собой процесс, который по существу есть не что иное, как элементарная форма моделирования. Следовательно, целесообразно попытаться поставить весь этот процесс на значительно более количественную основу.

Был испытан, в частности, следующий метод. На бумаге расписывали искусственную боевую обстановку и разбивали участников игры на две партии, каждая из которых самостоятельно этап за этапом принимала решения на основе представляемой ей информации. Таким образом, на каждом конкретном этапе именно люди принимали решения о начале атаки, отступлении, открытии огня, вызове танкового подкрепления и т. д., хотя последствия этих решений определялись искусственно путем вычисления вероятностей. Конечно, весьма желательно использовать для этой цели вычислительную машину, чтобы с ее помощью надлежащим образом рассчитывать результаты принятия решений, регистрировать развитие боя, а также выполнять необходимый анализ и подведение итогов. Существует много трудностей теоретического порядка и возражений против искусственных игр такого рода. Во-первых, неизбежная сложность любого боя обычно означает, что для завершения одной игры может потребоваться несколько недель и, следовательно, большое число повторений сопряжено с недопустимо большими затратами средств и времени. Поэтому трудно установить достаточно точно преимуще-

ства и недостатки разных стратегий, хотя участники игры могут получить большую практическую пользу, поскольку они вынуждены в большей мере мыслить количественными категориями и имеют возможность обдумывать последствия своих решений в относительно спокойной обстановке. Во-вторых, поскольку в процессе игры человек обучается и его суждения претерпевают определенные изменения, общие условия игры не могут оставаться неизменными. Это усложняет проблему выбора одной из нескольких различных стратегий или различных систем оружия при прочих равных условиях.

Более перспективная методика — построение математической модели всего процесса боя, включая принятие решений. Это не умаляет важности принятия решений, а, наоборот, позволяет включить этот процесс в модель в качестве ее важной составной части и, таким образом, лучше изучить его. Вообще говоря, любые решения, которые могут приниматься в процессе боя, должны кодироваться таким образом, чтобы их можно было принимать повторно (с введением соответствующих поправок на статистические колебания) столько раз, сколько это необходимо. Затем весь процесс боя можно ввести в вычислительную машину, благодаря чему время, требуемое для принятия решения, уменьшается на несколько порядков. В настоящее время для сравнения статистически приемлемым способом эффективности различных стратегий и систем оружия может быть применено моделирование методом Монте-Карло. Такого рода *тактическая военная игра* была разработана Управлением исследования операций США под кодовым названием «Кармонет». Детальное ее описание читатель может найти в гл. 24 книги Флегля, Хаггинса и Роя [25].

Возможность применения операционных игр как одного из методов исследования операций в биологии и медицине еще не изучена в полной мере. Одной из областей, в которой процесс принятия решений имеет особо важное значение, является организация здравоохранения. На основании общих соображений, опирающихся на клинические, биологические, экономические и социологические данные, принимаются определенные меры по совершенствованию организации здравоохранения, однако точно предсказать последствия какой-либо конкретной меры часто бывает трудно и даже невозможно. Некоторый прогресс в исследовании структуры принятия решений был достигнут корпорацией разработки систем в Лос-Анджелесе, где рассматривалась модель города под названием «Эпивиль». Эта превосходная работа более подробно обсуждается в разд. 12.5.

Теория игр

Широкое применение в исследовании операций может найти и так называемая *теория игр*. Существует важное различие между собственно теорией игр и операционными играми, о которых говорилось выше. Термин «операционная игра» по существу обозначает моделирование некоторого очень сложного и изменчивого процесса принятия решений с помощью игры, проводимой по определенным образом закодированным правилам; в ней могут принимать участие и люди, но лучше проводить ее с воображаемыми участниками, решения которых определяются путем соответствующих вычислений. Теория же игр связана главным образом с методами и принципами выбора наилучшей стратегии в определенной ситуации конкурентной борьбы. Практических примеров конкурентного поведения в торговле и промышленности можно найти сколько угодно. Эти примеры большей частью оказываются слишком сложными для непосредственного анализа, хотя уже достигнуты кое-какие успехи в анализе ситуаций, в которых конкурирующие компании добиваются заключения контрактов, предлагая различные условия поставки. Теоретические исследования обычно посвящаются исключительно простым и искусственно подобранным ситуациям, однако именно по этим причинам они позволяют выявить некоторые наиболее важные аспекты реальных проблем. Чтобы получить некоторое представление о теории игр, рассмотрим следующий простой пример.

Допустим, что два лица А и В участвуют в определенной игре, конкретные детали которой нас не интересуют. Игроки стремятся к тому, чтобы выиграть каждую партию при ограничениях, налагаемых правилами игры. Целью, или конечным состоянием, каждой игры является выигрыш для А, выигрыш для В или ничья. В заключение каждой игры один из игроков платит другому определенную сумму денег, размер которой зависит от того, как закончилась игра. Ничья означает нулевые платежи, но сумма, которую получает игрок А в случае выигрыша, может зависеть от характера его победы. Как уже говорилось ранее, в теории игр нас интересуют не детали отдельных шагов или ходов какой-либо конкретной игры, а общий план действий, или *стратегия*, которую нужно выбрать. Допустим, что игроку А известны три различные стратегии: A_1 , A_2 и A_3 , одну из которых он может выбрать, в то время как у игрока В имеются два возможных плана действий: V_1 и V_2 . Допустим теперь, что нам известно, какую сумму в конечном счете один игрок выплачивает другому, если выбрана какая-либо данная пара стратегий: например, игрок В выигрывает 20 очков у игрока А, если последний выби-

рает стратегию A_1 , а игрок В выбирает стратегию B_1 . Результаты для шести комбинаций стратегий можно представить в виде матрицы, в которой показаны суммы, выплачиваемые игроком В игроку А при выборе каждой пары стратегий. *Платежная матрица* для данного примера имеет следующий вид:

		Игрок А		
		A_1	A_2	A_3
Игрок В	B_1	-20	-10	10
	B_2	20	30	20

Если нас интересует игрок В, то сразу видно, что стратегия B_2 для него плоха, поскольку он всегда проигрывает и должен платить игроку А. Если же игрок В выбирает стратегию B_1 , то он выигрывает, когда игрок А выбирает стратегию A_1 или стратегию A_2 , но проигрывает 10 очков, когда игрок А поступает достаточно разумно и выбирает стратегию A_3 . Наилучшей стратегией для игрока В всегда является B_1 . В то же время игроку А лучше всего придерживаться стратегии A_2 , если игрок В выбирает стратегию B_2 . Однако по изложенным выше причинам маловероятно, чтобы игрок В поступил таким образом. Если он выбирает стратегию B_1 , то игрок А получает максимальный выигрыш, действуя по плану A_3 . Таким образом, наилучшими стратегиями для А и В оказываются такие, в результате выбора которых игрок В проигрывает игроку А 10 очков. Эта величина называется *ценой игры*. Если, как в данном случае, игроку лучше придерживаться одного и того же плана в каждой партии, то говорят, что он использует *чистую стратегию*.

Оптимальные стратегии можно найти с помощью полезного правила — *принципа минимакса*, согласно которому целесообразно выбирать стратегию, минимизирующую максимально возможный проигрыш. Так, для игрока А наибольшими проигрышами являются +20, +10 и -10 при стратегиях A_1 , A_2 и A_3 соответственно. Наименьшим проигрышем здесь является -10. Следовательно, минимаксной стратегией является A_3 . Применение аналогичных рассуждений к игроку В показывает, что для него максимальными проигрышами являются +10 и +30 при стратегиях B_1 и B_2 соответственно. Таким образом, минимаксной стратегией для игрока В является B_1 .

Итак, получен тот же результат, что и ранее. Если стратегии игрока А соответствуют столбцам платежной матрицы, то для выбора оптимальной стратегии нужно найти элемент, являющийся

наибольшим в своем ряду и наименьшим в своем столбце. Если такой элемент существует, то говорят, что это *седловая точка* и соответствующие стратегии являются наилучшими для двух рассматриваемых игроков.

Однако часто бывает так, что седловая точка отсутствует, т. е. не существует какой-либо пары стратегий, каждая из которых является оптимальной для данного игрока при всех обстоятельствах. В этом случае наилучший метод состоит в принятии игроками *смешанной стратегии*, которая означает, что каждый игрок случайным образом производит выбор определенного плана из числа всех возможных, причем вероятность выбора зависит от конкретного набора чисел, входящих в платежную матрицу. Полное решение таких задач может быть довольно сложным и осуществлено несколькими способами. Один из них состоит в превращении всей проблемы в задачу линейного программирования и использовании соответствующего метода решения этой преобразованной задачи.

Хотя изложенный выше подход к играм, при котором нас интересовал главным образом выбор одной из альтернативных стратегий без рассмотрения всех деталей игры, совершенно оправдан, можно дать и иное определение конкурентной ситуации, которое будет учитывать эти детали. Так, при игре в покер, бридж или в шахматы нас интересует, какой картой лучше всего ходить или какой фигурой лучше всего играть на каком-либо этапе игры. Можно показать, что при определенных общих условиях для игр, формулируемых в такой *эктенсивной форме*, каждый игрок имеет свою оптимальную чистую стратегию. К сожалению, число возможных вариантов обычно оказывается слишком большим, и потому этот подход не может иметь практического значения, хотя с теоретической точки зрения он заслуживает исследования.

Мы лишь поверхностно коснулись здесь вопросов, связанных с теорией игр. Эта теория может быть обобщена для ряда более сложных случаев: игры с участием нескольких лиц, образующих или не образующих коалиции; игры, в которых проигрыши являются не дискретными функциями, как в данном случае, а непрерывными; игры, в которых проигрыш одного из игроков не означает автоматически выигрыша для другого (это имеет место, например, в случае экономической конкуренции между фирмами), и т. д. Классической работой по теории игр является книга фон Неймана и Моргенштерна, впервые опубликованная в 1944 г. С тех пор появилась обширная литература по этому вопросу.

Насколько успешным оказалось применение теории игр на практике, пока сказать трудно. Сообщалось о применении теории игр в военном деле, однако по понятным причинам результаты опубликованы не были. В опубликованных работах по применению

теории игр в экономике, промышленности и административной деятельности чаще рассматриваются различные варианты и возможные выводы, но не содержится конкретных данных, свидетельствующих об их практической ценности. Однако в целом общепризнано, что теория игр значительно расширила наши возможности ориентироваться в целом ряде очень сложных задач. Такие, понятия, как участники игры, чистые и смешанные стратегии, платежные матрицы и т. д., создают новую систему представлений о сложных ситуациях, связанных с конфликтами и конкуренцией. Поэтому теорию игр нужно рассматривать как один из основных методов исследования операций, используемых при построении количественных моделей очень сложных процессов принятия решений.

Мы не стремились исчерпывающим образом излагать здесь все важнейшие методы, применяемые в исследовании операций, а лишь коснулись некоторых из них, для того чтобы дать почувствовать разницу между математическими моделями, возникающими в этой дисциплине, и моделями «чистых» и прикладных наук, рассмотренными в предыдущих главах. Изложение таких предметов, как теория графов, теория динамических систем, теория информации и принцип обратной связи, сетевые модели и т. д., читатель может найти в других руководствах.

4.4. КОМПЛЕКСНЫЕ ИССЛЕДОВАНИЯ НА СТЫКЕ РАЗЛИЧНЫХ НАУК

Наука вообще и прикладная математика в частности всегда носили в значительной степени эклектический и эмпирический характер. Любой научный аппарат годится до тех пор, пока он работает, и если различные аппараты, использовавшиеся в разные эпохи, в известной мере были несовместимы друг с другом, то какое-то время это не имело существенного значения, хотя попытки достигнуть в конце концов их объединения всегда были желательны и приводили к новым успехам. Например, в оптике в зависимости от задачи, стоящей перед исследователем, оказывались пригодными и корпускулярные, и волновые модели распространения света, хотя более строгое объяснение всех оптических явлений дала в конце концов квантовая механика, носящая более универсальный характер. Вообще говоря, при проведении исследований в любой области науки, скажем в биологии, часто возникает потребность в использовании знаний, полученных совсем в других областях — в физике, химии, математической статистике, электронике, технике и т. д. Некоторые разделы науки пересекаются с особенно большим числом других научных дисциплин. Например, генетика затрагивает ботанику, зоологию, цитологию, физиологию, гема-

тологию, молекулярную биологию, микробиологию, антропологию, судебную медицину, математическую статистику, вычислительную технику и т. д. Как правило, новые направления, возникающие на стыке нескольких вполне сложившихся научных дисциплин, не только развиваются весьма бурно, но и оказывают большое влияние на смежные дисциплины. Именно в таком положении находится математическая биология. Вначале это была довольно узкая область исследований, доступная главным образом тем, кто имел глубокие познания в математике и находил удовольствие в применении их к какой-либо частной биологической проблеме. Но постепенно количественные методы начали проникать во все более обширные области биологии. Это не только внесло определенный вклад в отдельные направления научных исследований, но и позволило установить новые связи и соотношения между ними благодаря использованию одних и тех же методов.

Таким образом, для научных исследований уже давно характерно объединение представителей различных дисциплин, но обычно это бывали научные работники не слишком далеких друг от друга специальностей. С возникновением метода исследования операций эта тенденция к объединению значительно усилилась и приобрела качественно новый характер, так как при операционном подходе для решения какой-либо одной конкретной проблемы требуется значительно более широкий диапазон специальностей и такое объединение становится уже неизбежным. Легко показать, каким образом это происходит, рассмотрев, например, историю возникновения исследования операций в годы второй мировой войны (разд. 4.2).

Военные задачи относятся к типичным задачам исследования операций, поскольку они с самого начала связаны с преодолением трудностей, возникающих при попытке организации ряда совершенно разнородных элементов и управления ими. Эти задачи должны решать прежде всего военные — представители инженерной службы, интенданты, связисты и штабные работники различного ранга. С военными вопросами тесно связаны и различные гражданские лица — ученые и инженеры, а также представители соответствующих отраслей промышленности, занимающиеся разработкой и производством оружия, судов, самолетов, танков и т. д. Кроме того, на высшем уровне вопросами управления и планирования занимаются различные правительственные организации. Поиск наилучшего способа проведения любого сражения или операции связан не только с глубоким изучением общей стратегии и тактики при планировании операций, но и с необходимостью принимать во внимание технические возможности и недостатки боевой техники. Поэтому, естественно, может потребоваться консультация специалистов в области взрывчатых веществ,

баллистики, самолетного оборудования, радиопеленгаторных устройств, экономики, психологии и т. д. Короче говоря, практически все системы «человек — машина» имеют множество аспектов, которые можно отнести к широкому кругу дисциплин научного, технического и организационного характера.

Таким образом, для того чтобы изучение какой-либо проблемы методом исследования операций оказалось успешным, в нем должны принимать участие многие специалисты, каждый из которых подготовлен в определенной области знаний, охватывающей только часть изучаемого вопроса (см. также разд. 6.3). Так, в группу, с которой автор был связан лично и которая занималась вопросами проектирования и деятельности больниц, входили архитектор, социолог, врач, медицинская сестра, историк, отставной офицер сухопутных войск и специалист в области математической статистики. Ценность такой смешанной группы состоит прежде всего в том, что это позволяет в одинаковой мере охватить все основные аспекты центральной проблемы, хотя обычно нет необходимости иметь специалистов по каждому отдельному вопросу. Кроме того, этот подход ценен тем, что именно с такого рода трудностями сталкиваются должностные лица и руководители при попытке провести организационную деятельность в области, в которой сами они не являются специалистами. Группа исследования операций смешанного состава решает поставленную задачу в таких условиях, которые значительно облегчают уяснение и разрешение основных трудностей. Если члены такой смешанной группы научатся работать совместно, общаться друг с другом, черпать друг у друга специальные термины, заимствовать образ мышления и сотрудничать при отыскании практических решений, то им скорее и легче удастся передать свои идеи и предложения тем руководителям, которые в конечном счете должны воплотить в жизнь результаты проведенных исследований. Таким образом, создание смешанной группы является составной частью исследования операций и облегчает не только изучение и анализ основной задачи, но и реализацию ее решения.

Глава 5

РОЛЬ ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКИ

5.1. ОБЩИЕ ПРИНЦИПЫ

За исключением лишь небольшого числа одаренных людей, человечество с незапамятных времен смотрит на арифметические вычисления как на тяжелую нудную работу, от которой необходимо избавиться любой ценой и любыми доступными средствами. Долгую историю имеют конторские счеты (по существу это примитивное цифровое вычислительное устройство с выполнением операций вручную), которые, по-видимому, были изобретены независимо в различных частях мира и кое-где применяются до сих пор. После того как в конце XVI в. были изобретены логарифмы, незаменимым механическим инструментом стала логарифмическая линейка. Первая логарифмическая линейка появилась в 20-х годах XVII столетия. Она дала возможность выполнять более сложные вычисления, хотя аналоговый характер прибора (числа изображаются расстояниями) значительно ограничивает точность при многих расчетах. Тем не менее логарифмической линейкой по-прежнему широко пользуются математики, ученые и инженеры во всем мире. Существует также и ряд специализированных счетных устройств: таблицы готовых расчетов в бухгалтерии, автоматические весы в бакалейно-гастрономических магазинах или счетчики на автозаправочных станциях. Во всех этих устройствах используются заранее подготовленные таблицы или шкалы, позволяющие очень быстро производить операции умножения при различных денежных расчетах.

Такие технические приспособления распространены в настоящее время очень широко и служат простейшим способом избавления от большого объема простых, но утомительных арифметических вычислений. На несколько более высоком уровне сложности находятся конторские арифмометры и кассовые аппараты, которые легко выполняют большое число последовательных операций сложения и вычитания и печатают результаты выполняемых операций, а также различные промежуточные и общие суммы. На основе этих относительно простых аппаратов разработаны разнообразные более совершенные счетные машины. Человек, поль-

зующийся таким прибором, может и не иметь полного представления о том, как он устроен, но обычно ему нетрудно понять, что это просто удобное механическое приспособление для более эффективного выполнения той работы, которую он и сам мог бы сделать вручную с помощью карандаша и бумаги, если бы у него было достаточно времени и он мог работать без усталости и никогда не ошибаться.

Аналогичные соображения применимы также и к более сложным настольным счетным машинам. Это механические цифровые устройства, предназначенные для выполнения основных арифметических действий (сложения, вычитания, умножения и деления) и имеющие ряд вспомогательных устройств для накопления результатов и выполнения операций сокращения. Более дешевые счетные машины обычно управляются вручную и их применяют, когда объем вычислений относительно невелик, а также для целей обучения. Для очень продолжительных вычислений, когда требуется большая скорость и гибкость, необходимы более крупные и более дорогие счетные машины. Обычно они приводятся в действие электромотором, хотя по принципу действия по-прежнему являются механическими. Настольные счетные машины такого общего типа используются уже очень много лет, и до самого последнего времени с их помощью выполняли практически все сколько-нибудь сложные научные расчеты. Огромные преимущества настольных счетных машин по сравнению с методами вычислений вручную общеизвестны. Числа вводятся в машину быстро и точно путем поворота ручек или нажима кнопок; все арифметические действия выполняются внутри самой машины; при тщательном планировании последовательности вычислений можно избежать записи большого числа промежуточных результатов. Разумеется, при вычислениях могут возникать ошибки, хотя и не так часто, как при работе вручную, поэтому обычно в последовательности вычислений включается ряд операций по проверке результата.

Появление настольных счетных машин, естественно, расширило область возможных вычислений, однако вскоре наступило время, когда был достигнут естественный предел быстродействия этих машин. Легко убедиться в том, что при выполнении большой последовательности вычислений, например при суммировании большого числа произведений, ограничивающим фактором оказывается не столько скорость работы самой машины, сколько время, необходимое для введения в нее чисел, чтения и переписывания результатов, а также принятия решений о последующих действиях. Поэтому в итоге время, требуемое для решения сложной задачи, уменьшается очень незначительно, даже если машина выполняет основные арифметические действия практически мгновенно. Чем сложнее работа, тем больше будет сказываться уста-

лость оператора и тем вероятнее возникновение ошибок. Кроме того, если вычисления определенного типа должны повторяться многократно, то правильная последовательность действий при одном вычислении не дает гарантии, что в следующий раз не будет допущена ошибка. Другая трудность состоит в том, что обычно для сокращения времени и повышения точности при продолжительных вычислениях целесообразно поручать основную работу опытному оператору, хотя это и означает, что составитель задачи должен затратить много труда на подготовку работы и необходимые объяснения оператору. И если этот же математический метод будет использоваться снова для других данных, то может случиться, что все эти объяснения придется давать вновь уже другому оператору.

Хотя настольные счетные машины представляют собой большой шаг вперед по сравнению с другими методами вычислений, их недостатки довольно очевидны, и поэтому появились серьезные основания для разработки машины совершенно другого типа. Соответствующая электронная схема, в которой используются лампы или транзисторы и электрические цепи, безусловно, позволяет гораздо быстрее выполнять основные арифметические действия, чем чисто механические устройства. Однако преимуществами этих средств нельзя воспользоваться, если вмешательство человека не будет сведено к минимуму. Это означает, что необходимо не только избавиться от медленной записи, чтения и передачи чисел человеком, но нужно каким-то образом перестроить весь план работы, с тем чтобы оператор не должен был принимать никаких решений в процессе вычислений. Именно в решении этих проблем, а не только в увеличении скорости внутренних операций состоит настоящая революция, вызванная современными автоматическими электронными вычислительными машинами.

Вычислительные машины начали разрабатывать в конце второй мировой войны. Первыми машинами, имевшими основные конструктивные особенности современных вычислительных устройств, были вычислительная машина EDSAC, начавшая работать в 1949 г. в Кембриджском университете, и вычислительная машина SEAC, изготовленная Национальным бюро стандартов США в 1950 г. В этих первых машинах использовались электронные лампы, которые в настоящее время заменены транзисторами, позволившими уменьшить габариты вычислительных машин и обладающими значительно более высокой надежностью. К настоящему времени достигнуты большие успехи в области микроминиатюризации аппаратуры. Все это позволило создать настольные электронные вычислительные машины и снизить продолжительность основных операций до наносекунд, что означает тысячи миллионов операций в секунду. В настоящее время выпущен ряд превосходных руководств, в которых читатель может найти детальное описание суще-

ствующих машин и конкретные рекомендации по работе с ними. Мы обсудим здесь лишь основные принципы в той мере, в какой они связаны с темой данной книги.

Рассмотрим важнейшие свойства современной электронной вычислительной машины. Прежде всего важно уяснить, что она по-прежнему выполняет те же основные арифметические действия, что и настольная счетная машина, и поэтому по существу делает то же, что в принципе может сделать человек, работающий только с бумагой и карандашом. Различие состоит, с одной стороны, в огромном увеличении технической эффективности, а с другой — в осуществлении логического контроля за последовательностью операций. Поэтому понять то, что в конечном счете делает вычислительная машина, ничуть не сложнее, чем понять обычный способ решения арифметических задач. Считать ли, что электронная вычислительная машина обнаруживает признаки интеллекта или что она способна выполнять работу, близкую к функциям мозга, в значительной мере зависит от того, как мы определяем эти понятия. Самое главное состоит в том, что вычислительная машина, как бы она ни была сложна и совершенна, по существу представляет собой лишь одно из очень сложных устройств, предназначенных для определенных целей, и как таковую ее следует рассматривать во многих отношениях в таком же плане, что и любую другую сложную научную аппаратуру, например электронный микроскоп или линейный ускоритель частиц.

В обычную настольную счетную машину вводится пара чисел для выполнения некоторого арифметического действия, например сложения, и эти числа могут храниться в ней после выполнения операции. В машине могут также храниться еще одно-два числа, например число, полученное в результате сложения или умножения нескольких чисел, или накопленная сумма произведений. Однако общая емкость запоминающего устройства редко превышает пять-шесть чисел, причем даже и в этом случае число разрядов чисел крайне ограничено. В электронной вычислительной машине имеется соответствующее электромагнитное устройство, в котором может храниться несколько тысяч чисел (с весьма большим числом разрядов) в виде, позволяющем производить быструю выборку; на магнитном диске или магнитной ленте может храниться несколько миллионов чисел в форме, позволяющей производить относительно медленную выборку (медленную по электронным стандартам). Это дает возможность исключить вмешательство человека при обработке промежуточных результатов, а также осуществлять автоматическую обработку огромного множества численных данных. Поскольку данные записаны на перфокартах, бумажной ленте, магнитных дисках или магнитной ленте, они могут использоваться многократно без приложения умственных

или зрительных усилий со стороны человека; они просто поступают на соответствующее внешнее считывающее устройство вычислительной машины.

Но самое главное состоит в том, что удалось разработать способы планирования всей последовательности вычислений. Для этого в вычислительную машину вводится *программа* команд, которая хранится там вместе с соответствующими данными. Эти команды записаны в соответствующем численном коде и связаны прежде всего с выполнением основных арифметических действий над парами чисел, хранящихся в определенных частях запоминающего устройства. Если программа и необходимые данные введены в машину, то все остальные операции выполняются самой машиной со скоростью, зависящей от электромагнитной схемы. Окончательные результаты либо записываются на перфокартах или бумажной ленте, либо подаются непосредственно на телетайп или другое воспроизводящее устройство. Программа для любой данной последовательности вычислений должна быть продумана очень тщательно. Зато когда программа составлена и как следует проверена, ее можно многократно использовать без дальнейшей проверки. Ясно, что это дает значительную экономию времени и труда. Надлежащим образом проверенную программу могут использовать тысячи раз многие люди, работающие на различных вычислительных машинах. Не менее важно и то, что определенная программа может содержать команду повторить выполнение какого-либо конкретного задания (например, решение определенной системы уравнений) несколько сотен тысяч раз, причем каждый раз с использованием различных данных. Тогда часть программы должна быть посвящена решению одной системы таких уравнений. Если эта часть программы составлена правильно, то в пределах данной программы к ней можно обращаться сколь угодно часто с полной уверенностью, что каждый раз решение уравнений будет производиться правильно. Этим устраняется один из важнейших недостатков настольных счетных машин, о котором уже говорилось выше, а именно то, что правильная работа настольной счетной машины в одном случае не гарантирует отсутствие ошибки при повторных вычислениях.

Работая на настольной счетной машине, оператор на всех этапах наблюдает за производимыми вычислениями. Если происходит что-либо необычное или непредвиденное, он сразу же может принять необходимые меры и таким путем избежать серьезных ошибок. Против применения электронных вычислительных машин часто возражают на том основании, что здесь существует возможность очень большого числа незамеченных ошибок на промежуточных этапах вычислений и ввиду этого окончательные результаты следует считать в высшей степени сомнительными. К счастью, эту труд-

ность можно в значительной мере исключить путем хорошего программирования. Как мы видели, в обычной настольной счетной машине также имеются очень большие возможности появления ошибок при повторных вычислениях. При этом только крупные ошибки выявляются сразу, а для того, чтобы предупредить накопление небольших ошибок, требуется проводить тщательные проверки на всех этапах. Эти проверки составляют важную часть всей последовательности вычислений. Но ведь такие же проверки составляют неотъемлемую часть программы электронной вычислительной машины. Вследствие своего большого быстродействия электронная вычислительная машина позволяет осуществить значительно больше арифметических проверок, чем настольная счетная машина. Следовательно, при хорошо составленной программе электронная вычислительная машина обеспечивает вовсе не меньшую, а значительно большую надежность результатов.

Выше шла речь об ошибках, появляющихся из-за недостаточного контроля за операциями, выполняемыми вычислительной машиной; но возможны также ошибки, обусловленные повреждением электронной схемы. Для выявления таких ошибок можно, например, предусмотреть автоматическую остановку вычислительной машины при появлении ошибок, обнаруживаемых так называемой проверкой на четность. В большинстве машин числа хранятся в двоичной записи как последовательность нулей и единиц. Можно приписать дополнительную цифру, равную нулю или единице, в зависимости от того, четной или нечетной является сумма единиц в каждом хранящемся числе. В определенные моменты времени, например при считывании числа из запоминающего устройства, приписанную цифру можно сверить с самим числом. Такую схему проверки нельзя, конечно, считать абсолютно надежной, так как могут возникать компенсационные ошибки; но она все же очень полезна тем, что привлекает внимание к ошибкам, вызываемым неполадками в электронной схеме.

Таковы некоторые основные идеи, связанные с применением электронных вычислительных машин. Широкое распространение вычислительной техники радикально повлияло на наши вычислительные возможности, на масштабы обработки данных и в итоге — на общее направление научно-исследовательских работ и выбор методов их осуществления. К этому аспекту мы еще вернемся в разд. 5.5 и 5.6, а пока более внимательно рассмотрим некоторые основные задачи, для решения которых целесообразно применение вычислительных машин, и методы, позволяющие исследователям использовать преимущества этой новой техники.

Прекрасным пособием для знакомства с вычислительными методами может служить недавно выпущенная книга Холлингдейла и Тутилла [35].

5.2. НАУЧНЫЕ ВЫЧИСЛЕНИЯ И ЧИСЛЕННЫЙ АНАЛИЗ

Как уже было сказано, одно из важнейших достоинств электронной вычислительной машины — ее способность выполнять с большой скоростью огромное множество арифметических действий. Поэтому естественно, что вычисления, выполнявшиеся ранее на настольных счетных машинах, в настоящее время все чаще производят на электронных вычислительных машинах. Кроме того, теперь стало возможным численное решение многих задач, которые ранее считались слишком сложными или громоздкими. Мы уже говорили о том, что если программы вычислительных машин надлежащим образом испытаны и предусматривают выполнение необходимых проверок, то повышается не только скорость вычислений (на несколько порядков), но и их точность, а также значительно увеличивается охват материала и гибкость программы. Тем не менее необходимо соблюдать известную осторожность, чтобы некоторые трудности, характерные для настольных счетных машин, не проявились здесь в еще более резкой форме.

Решение огромного множества научных задач, требующее специальных вычислений, связано с областью математики, обычно называемой *численным анализом*. Сюда относятся методы интерполяции, сглаживания (подбора) кривых, численное дифференцирование и численное интегрирование, суммирование рядов, решение алгебраических и трансцендентных уравнений, обыкновенных дифференциальных уравнений и дифференциальных уравнений в частных производных, вычисление коэффициентов корреляции и коэффициентов множественной регрессии, статистическая обработка экспериментальных результатов и т. д. При решении всех этих задач весьма возможно появление ошибок, если вычисления производят лица, не имеющие хотя бы минимальной математической подготовки. Вопрос о том, какой именно уровень подготовки нужно установить, служит предметом бурных споров. В прошлом очень многие ученые, обладавшие умеренными математическими познаниями, умели выполнять все виды трудных числовых расчетов для собственных нужд, не получая слишком неблагоприятных отзывов о своей работе со стороны профессиональных математиков. Однако с появлением больших дорогостоящих электронных вычислительных машин и возникновением крупных вычислительных лабораторий занятый здесь персонал нередко принимает большое личное участие в численных расчетах и проявляет к ним профессиональный интерес. При этом затушевывается тот факт, что при выполнении сложных численных расчетов независимо от того, используется настольная счетная машина или же автоматическое вычислительное устройство, существует реальная опасность появления ошибок.

Пусть, например, принято допущение (вероятно, ошибочное), что для решения некоторой практической научной задачи необходимо просуммировать бесконечный ряд

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$$

Исследователь нематематик, которому неизвестны строгие условия сходимости рядов такого типа, естественно, решит, что нужно последовательно суммировать члены ряда до тех пор, пока не окажется, что добавление новых слагаемых мало меняет сумму. Вначале ряд сходится довольно медленно. Если взять первые сто членов, то сумма их составит около 5,19, и при этом последний член будет равен всего 0,01. Добавление еще 10 членов, каждый из которых не превышает 0,01, увеличит сумму почти на 0,1. При выполнении этих вычислений на настольной счетной машине (довольно трудоемкая работа) наблюдательный исследователь легко придет к выводу, что здесь что-то не так. Обратившись за консультацией к математику, он сразу узнает о том общеизвестном факте, что в действительности этот ряд расходится и в зависимости от числа членов сумма его может оказаться сколь угодно большой. Это, вероятно, заставит его сделать вывод, что первоначально принятое допущение в чем-то ошибочно и его необходимо тщательно исследовать и изменить. Посмотрим теперь, что произойдет, если для суммирования такого ряда будет запрограммирована вычислительная машина. Обычно в учебниках рекомендуется прекращать вычисления, если последнее слагаемое меньше некоторого произвольно выбранного малого числа. Но это правило может оказаться неудовлетворительным даже для абсолютно сходящихся рядов (если ряд сходится очень медленно), не говоря уже о расходящихся. Однако вычислительная машина и в этом случае все же выдаст конечное число и, следовательно, допустит грубую ошибку, поскольку конечного результата здесь вообще не существует. Обычно ряды, подлежащие суммированию, сходятся довольно быстро, и никаких трудностей не возникает. Например, никаких затруднений не вызывает ряд $1, -\frac{1}{2}, +\frac{1}{3}, -\frac{1}{4}, \dots$. Но иногда ошибка такого рода может быть допущена как раз тогда, когда решается какая-то важная задача.

В такие по существу математические ловушки исследователь может попасть из-за отсутствия необходимых математических знаний или неумения применять их на практике. Можно привести множество других, более сложных примеров. Например, стандартные формулы для численного интегрирования определенной функции могут оказаться непригодными, если в области интегрирования функция имеет особенность. Часто эту трудность можно обойти с помощью каких-либо преобразований, сделанных перед нача-

лом вычислений. Аналогичные проблемы часто возникают и при решении дифференциальных уравнений.

Другой возможный источник ошибок заключается в том, что в машине числа приходится представлять в виде десятичных дробей с точностью, зависящей от емкости ее запоминающего устройства. Это приводит к ошибке округления, достигающей иногда половины единицы последнего десятичного разряда. Округление любых отдельно взятых чисел, особенно чисел, представляющих собой исходные значения или данные, может не иметь существенного значения, однако округления в процессе последующих расчетов могут привести к серьезной погрешности. Многим известно об опасностях, связанных с вычитанием одного числа из очень близкого к нему другого числа. Так, если числа 100,02 и 100,03 заданы с точностью до пяти значащих цифр, то разность между ними 0,01 может содержать ошибку до единицы второго десятичного разряда и вообще не иметь точных значащих цифр. При планировании вычислений необходимо проявлять большую осторожность, чтобы избежать потери значащих цифр на промежуточных этапах, ибо при длительных вычислениях ошибка, накапливаемая в результате округления, может иметь катастрофические последствия. При выполнении вычислений на настольной счетной машине внимательный и опытный оператор сумеет заметить основные источники неполадок в случае их появления и либо примет необходимые меры, либо обратится за консультацией к составителю задачи. Однако операторов требуемой квалификации часто не хватает, и, кроме того, при длительной работе человек может устать и тогда внимание его ослабевает. При автоматических вычислениях, как уже было сказано, невозможно исследовать промежуточные этапы, и поэтому в программу необходимо вводить соответствующие операции, предупреждающие появление ошибок. Во многих случаях составить программу, надежно защищенную от ошибок, бывает довольно трудно и это требует много времени; но если такая программа уже составлена, то ее можно использовать сколь угодно часто, не опасаясь возникновения ошибки.

Необходимость отбрасывать большое число десятичных разрядов иногда может приводить к результатам, которые менее очевидны на практике, но могут иметь серьезные последствия. Возьмем, например, многочлен

$$(x-1)(x-2)(x-3) \dots (x-20) \equiv x^{20} - 210x^{19} + \dots$$

Его нулями являются целые числа 1, 2, 3, ..., 20. Поэтому при любом математическом методе нахождения нулей многочлена, записанного в правой части равенства, должны быть получены именно эти числа (во всяком случае, если вычисления могут быть выполнены с абсолютной точностью). Но допустим, что коэффи-

циент при втором члене определен с небольшой погрешностью и вместо 210 мы имеем $210 + 2^{-23}$. Если мы теперь начнем искать нули этого видоизмененного многочлена, снова предполагая, что вычисления производятся с абсолютной точностью, то естественно ожидать, что новые нули не будут в точности равны первоначальным, но что различия будут весьма невелики. Однако полной неожиданностью для нас явится тот факт, что первоначальные целочисленные нули 16 и 17 в действительности заменяются парой сопряженных комплексных чисел, приближенно равных $16,7 \pm 2,8i$, где мнимая часть довольно значительна. Если иметь в виду, что в большинстве вычислительных машин все коэффициенты вычисляются с небольшими погрешностями, величина которых зависит от того, как был получен многочлен, то понятно, что в отдельных случаях небольшие ошибки могут увеличиться до огромных размеров. При этом следует учесть, что для упрощения мы пренебрегли ошибками округления, появляющимися в процессе вычитания чисел.

Неустойчивость может быть также связана с применяемым методом вычислений, а не с характером самой задачи, как в описанном только что случае. Поэтому к выбору метода вычислений необходимо подходить с большой осторожностью. Часто для получения достаточно точных результатов необходимо вычислять нули многочлена в определенном порядке. Известный метод Рунге — Кутты, используемый для интегрирования обыкновенных дифференциальных уравнений некоторых типов, также может оказаться ненадежным, когда решение содержит показательные функции с отрицательными показателями степени или четные тригонометрические функции. Возможны также некоторые другие виды неустойчивости, и, только обладая достаточными знаниями в области численного анализа, можно определить, удастся ли избежать неустойчивости при решении определенной задачи с помощью надлежащим образом выбранного метода.

Ясно, что в идеале любые вычисления, выполняемые как на настольной счетной машине, так и на электронной вычислительной машине, должны были бы подвергаться тщательному исследованию, выполняемому квалифицированным специалистом в области численного анализа, но это, к сожалению, осуществимо далеко не всегда. Решение проблемы, по-видимому, заключается в том, чтобы все научные работники, как самостоятельно выполняющие свои вычисления, так и составляющие программы для вычислительных машин, приобрели хотя бы такие познания в математике, чтобы отчетливо понимать, в каких случаях им нужно обратиться за советом к специалисту.

Вполне вероятно, что в реальной научной работе такие действительно сложные проблемы возникают не столь часто, как это

склонны считать математики, хорошо знакомые с такими печальными примерами. На основе статистики можно показать, что, как правило, накапливаемая ошибка округления будет приводить к недопустимой погрешности довольно редко. Кроме того, появление чрезвычайно неточных результатов вследствие, например, возникновения различных видов неустойчивости часто легко обнаруживается по их явному несоответствию физическим или биологическим особенностям исходной задачи. Часто практическая интуиция, основанная на хорошем знании исследуемого реального явления, позволяет предвидеть характер возможных математических затруднений. Не следует недооценивать значения общей научной интуиции. Именно ей мы обязаны, по-видимому, тем, что ученые допускают грубые числовые ошибки не так уж часто, как этого можно было бы ожидать. Но если ученый имеет возможность работать в тесном сотрудничестве с профессиональным математиком, для которого исследуемая проблема также представляет определенный интерес, то это, конечно, в высшей степени полезно.

Хорошим введением в элементный численный анализ, где специально рассматривается использование вычислительных машин в научных исследованиях, может служить книга Хемминга [33].

5.3. ОБРАБОТКА ДАННЫХ

В разд. 5.1 мы уже указывали, что электронная вычислительная машина способна не только выполнять с огромной скоростью арифметические действия, но и хранить огромное количество данных. Это привело к развитию новой широкой области математики, связанной с *обработкой данных*. Эта область в значительной степени отличается от численного анализа, хотя обе они, безусловно, взаимосвязаны и дополняют друг друга.

В большинстве научных дисциплин, а также в области права, торговли, промышленности, государственного управления и т. д. непрерывно накапливается огромное количество все более сложных данных, зафиксированных в виде записей. В сфере административного управления записи используются для установления различных правил и предписаний, в сфере науки — для формулировки и проверки научных гипотез. Как показано в гл. 2, в некоторых случаях удобно обобщать большое количество численных данных путем применения соответствующего статистического метода. Если вычислены надлежащие статистические показатели, то они могут использоваться и при последующих численных расчетах, и при этом не возникает необходимости снова детально изучать исходный материал. Однако нередко исходные данные необходимо сохранять почти полностью в их первоначальном виде — как для проверки различных методов исследования, так

и потому, что некоторые из них могут иметь особо важное значение.

В простейшем случае регистрация данных связана с хранением исходных документов — семантической информации или результатов количественных измерений. Если документов очень много, то их необходимо расклассифицировать и хранить по определенной системе; при этом нужна специальная картотека, позволяющая быстро найти любой требуемый материал. Огромное множество историй болезни в больницах и в поликлиниках хранится именно в такой форме. В этом случае можно сравнительно быстро найти истории болезни отдельных больных, но когда требуется получить сводные данные в виде таблиц, показывающих частоту появления тех или иных событий или связь между различными аспектами материала, обычно возникают почти непреодолимые трудности, если только не принимаются специальные меры, облегчающие такой анализ. При небольшом числе факторов удобно проводить текущие проверки соответствующих вопросов по мере накопления записей. Например, больных, поступающих в больницу на стационарное лечение, можно регистрировать в специальной книге, где записывается возраст, пол, диагноз, длительность пребывания в больнице, фамилия лечащего врача и т. д., а затем каждую неделю или каждый месяц составлять специальные таблицы, отражающие их состояние вплоть до выписки. Выполнение такого рода исследований обычными канцелярскими методами требует больших затрат труда и времени, нередко бывает неточным, значительно ограничивает объем материала, который можно обработать, и позволяет применять лишь ограниченное число методов анализа.

Более совершенный метод — запись соответствующих данных на специальные карточки, на которых пробиваются определенным образом расположенные отверстия, кодирующие классификацию данных. Эти отверстия (перфорации) могут пробиваться у краев карты, как в системе Коуп-Чета и системе Макби, так что сортировку карт можно выполнять вручную на конторском столе с помощью спицы или какого-либо другого инструмента. Например, отверстие может означать отсутствие определенного фактора, а сквозная щель или прорезь — его присутствие. Вставляя спицу в пачку карточек через отверстие, занимающее определенную позицию, и приподнимая спицу, можно отсортировать все карточки с прорезями, соответствующие группе больных, у которых наблюдается данный фактор. При надлежащем выборе комбинаций доверстий можно получить численную или алфавитную классификацию.

Еще удобнее использовать всю площадь карточки. Так, типичная перфокарта вычислительной машины IBM имеет 80 столб-

цов по 12 перфорационных позиций в каждом. Это позволяет записать значительно большее количество информации, и такие карты можно сортировать и подсчитывать автоматически с помощью электрического датчика, сообщающего о наличии отверстия в данной перфорационной позиции. В настоящее время во многих больницах помимо основных историй болезни ведутся записи на перфокартах, используемые для административных и научных целей. Такой метод обработки материала все же еще довольно громоздок и требует много времени, хотя это значительный шаг вперед по сравнению с сортировкой карточек вручную на столе. Он значительно облегчает составление обычных административных отчетов, если только они не требуются слишком часто. Обычно нетрудно найти годовые или квартальные цифры, но данные за сутки в большинстве случаев получить невозможно. Составить представление об общем положении дел можно довольно быстро, но детальные исследования, сопряженные с составлением больших многокоординатных таблиц и, главное, с проведением сложных статистических вычислений, могут оказаться невыполнимыми.

В настоящее время электронные вычислительные машины позволяют в огромной степени ускорить обработку таких записей. Для удобства данные переносят сначала из первоначальных записей на перфокарты. Затем машина считывает информацию, записанную на перфокартах, и хранит ее на магнитных дисках или магнитной ленте в очень компактной и удобной форме. Вся последующая обработка и анализ данных выполняются машиной с очень большой скоростью. По мере надобности данные, хранящиеся в машине, считывают с магнитного диска или магнитной ленты, но и эта операция производится очень быстро. Кроме того, в результате однократного просмотра хранимого материала можно составить большое число сводных таблиц или выполнить другие виды анализа. Сравним эти возможности с трудоемкой многократной сортировкой оригинальных карточек, производимой при описанных выше механических методах классификации. Если требуется обработать лишь несколько сотен записей, то это можно сделать на конторском столе, используя оригинальные документы или их варианты на карточках с краевой перфорацией. Однако если имеется несколько тысяч записей, то почти всегда стоит приложить дополнительные усилия, чтобы передать материал на хранение в электромагнитное запоминающее устройство и составить соответствующие программы для вычислительной машины, позволяющие обрабатывать данные, записанные в таком виде. Это значительно повышает скорость, точность и гибкость обработки. Вопрос повышения скорости вряд ли нуждается в специальном комментарии. Точность повышается благодаря тому, что вычисления производятся почти без вмешательства человека и имеется возможность предусмо-

треть большое число перекрестных проверок и испытаний непротиворечивости. Увеличение гибкости означает, что значительно расширяется диапазон возможных видов анализа и что операции сортировки и обработки данных могут играть роль предварительного этапа научных вычислений, о которых шла речь в предыдущем разделе. Примерам взаимосвязи между обработкой данных и научными вычислениями будет уделено особое внимание в гл. 11 и 12.

Наряду с обработкой данных, компактно представленных в обычном письменном или печатном виде, существует проблема обработки большого количества сырого материала такого типа, как, например, электрокардиограммы, электроэнцефалограммы и т. п. Существуют самые различные способы анализа таких данных, и выбор какого-либо одного из них зависит от специфики конкретной задачи, стоящей перед исследователем. Но смысл любого такого способа заключается по существу в увеличении отношения сигнала к шуму, с тем чтобы более отчетливо выявить основные характеристики формы сигнала. Так, усреднение реакции дает возможность получить более четкую картину формы повторяющегося сигнала. Этот метод очень ценен, если известно, что определенные формы сигнала имеют диагностическое значение. Кроме того, можно попытаться выделить основные компоненты колебаний с помощью частотного анализа. В тех случаях когда основной биологический процесс имеет характер случайной функции, может оказаться предпочтительным корреляционный подход. В настоящее время все эти методы можно применять в реальном масштабе времени, когда наблюдаемые величины в виде электрических напряжений вводятся непосредственно в соответствующую вычислительную машину. Для некоторых целей вполне удобна аналоговая техника, которая часто применяется на первом этапе исследований. Однако при использовании ряда более сложных математических методов необходимо преобразовывать данные в цифровую форму.

Дальнейшее развитие методов обработки данных на ЭВМ в реальном масштабе времени привело к созданию систем автоматического обследования больных. Сейчас разрабатываются системы, позволяющие непрерывно регистрировать различные показатели состояния больного (температура, частота пульса, кровяное давление, частота дыхания и т. д.). Прежде всего, конечно, нужно найти удовлетворительные способы, которые позволяли бы непрерывно измерять и регистрировать эти физиологические параметры, надлежащим образом представлять их визуально, получать сигналы о наступлении тревожного или опасного состояния и т. п. без применения сложных вычислительных методов. Но в тех случаях, когда, например, необходимо анализировать форму сигналов на электрокардиограммах и электроэнцефалограммах, вычисли-

тельные методы могут оказаться крайне полезными. Предполагается, что в общем случае системы автоматического обследования больных позволят очень быстро определить резкое ухудшение состояния у тяжелых больных. При возникновении тревожных или опасных ситуаций медицинские сестры и врачи должны принимать соответствующие меры. Однако вполне возможно предусмотреть автоматическое выполнение некоторых терапевтических приемов, особенно в тех случаях, где важна быстрота.

Еще один тип обработки данных возникает в рамках задачи распознавания образов, например при автоматической классификации хромосом. Для оценки фотографических данных используются методы электронного сканирования, позволяющие перейти от графической формы к цифровому коду и путем вычислений произвести соответствующий анализ.

Широкое обсуждение современных работ в области получения и обработки данных читатель может найти в опубликованных трудах Рочестерских конференций [19—21].

5.4. ПРОГРАММИРОВАНИЕ

После того как установлено, к какому общему типу относится задача, которую требуется решить, и выбран подходящий метод анализа (идет ли речь об обработке данных, численном анализе или же некоторой комбинации того и другого), вычислительная машина должна получить инструкцию, в какой последовательности должны выполняться вычисления. Как уже указывалось в разд. 5.1, это делается с помощью надлежащим образом составленной программы, т. е. последовательности команд. Таким образом, программирование составляет существенный элемент исследования, связанного с использованием электронных вычислительных машин. Обычно составлением новых программ или переделкой существующих занимаются специально подготовленные программисты (так же как при широком использовании настольных счетных машин целесообразно иметь квалифицированных операторов). В обоих случаях важно, однако, чтобы составитель задачи имел определенные сведения о типе машины, которая будет производить вычисления, так как это может оказать решающее влияние на выбор метода анализа. Этому вопросу мы коснемся в разд. 5.5. Здесь же для тех, кто, возможно, не знаком с этим вопросом, будут рассмотрены основные аспекты программирования.

Прежде всего заметим, что, хотя в большинстве электронных вычислительных машин числа выражаются в двоичной системе, т. е. в виде определенной последовательности нулей и единиц, для целей программирования исходные данные или числа, входя-

щие в формулы, можно записывать в обычной знакомой нам десятичной системе. Когда эти числа с перфокарт или магнитной ленты вводятся в машину, их запись автоматически преобразуется в двоичную. В различных ячейках, или регистрах, запоминающего устройства вычислительной машины могут храниться либо числа, либо команды на выполнение определенных арифметических действий. Команды выглядят точно так же, как числа (машина должна получить указание, каким образом отличать команды от чисел), и также записаны в двоичной системе. Для удобства выбирается специальный код с тем, чтобы машина могла легко получить указание, какое арифметическое действие необходимо выполнять. В арифметическом блоке вычислительной машины обычно имеется накапливающий регистр (точно так же, как и в настольной счетной машине). Поэтому для вычисления суммы двух чисел, уже хранящихся в вычислительной машине, одно из них переносят в накапливающий регистр, затем прибавляют к нему второе число и, наконец, переносят сумму обратно в такую ячейку запоминающего устройства, чтобы это не изменило хранящейся в нем полезной информации.

Рассмотрим, например, *машинный код* ЭВМ «Эллиотт-803». В этой машине все отдельные ячейки запоминающего устройства пронумерованы и каждая ячейка может содержать одно число или две команды. Допустим, что два числа x и y хранятся в определенной численной форме в ячейках 251 и 252. Требуется найти сумму $x + y$ и поместить ее в ячейку 253, сохранив x и y в их первоначальном месте. Допустим, что команда для записи числа в накапливающий регистр кодируется числом 30, за которым следует номер, или адрес, ячейки, где хранится это число. Таким образом, первая команда, согласно которой число x передается в накапливающий регистр, имеет вид 30 251. Теперь к x нужно прибавить y . Команда на сложение кодируется числом 04, и, следовательно, вторая команда имеет вид 04 252. Наконец, необходимо перенести сумму $x + y$, находящуюся в накапливающем регистре, в ячейку 253. Операция такого рода выполняется по команде 20; следовательно, третья команда кодируется числом 20 253. Итак, окончательная последовательность команд записывается на бланке программы следующим образом:

30 251 04 252

20 253

Операции заканчиваются тем, что число x , как и вначале, оказывается в ячейке 251, число y — в ячейке 252, а в ячейке 253 появляется новая величина $x + y$. Разумеется, это очень простой пример, однако он ясно иллюстрирует принцип, на котором основана передача команд вычислительной машине.

Хотя имеется всего четыре основных арифметических действия — сложение, вычитание, умножение и деление, — удобно задавать более широкий круг комбинированных операций. Так, может потребоваться перенести некоторое число в накапливающий регистр и освободить ячейку, из которой оно взято; прибавить некоторое число к другому числу, которое уже находится в накапливающем регистре, и перенести это второе число в ячейку, из которой было взято первое. В коде ЭВМ «Эллиотт» насчитывается 30—40 различных комбинаций такого рода, и он очень прост, гибок и легко запоминается (во всяком случае, если пользоваться им постоянно).

Так как все вычисления построены на элементах такого рода, ясно, что с помощью машинного кода можно составить программу для выполнения любой последовательности арифметических действий. Все операции, которые можно выполнять на настольной счетной машине, в том числе запись чисел, передачу их или воспроизведение, имеют соответствующие эквиваленты в машинном коде. Разумеется, при большой последовательности вычислений составление программы для вычислительной машины — весьма утомительное занятие, поглощающее много времени; однако хорошо составленная программа позволяет эффективно использовать возможности вычислительной машины по быстродействию. Программисты с большим опытом работы составляют хорошие программы для ЭВМ довольно быстро. Но для тех лиц, которым приходится пользоваться вычислительной техникой и у которых нет времени или, возможно, способностей приобрести эти навыки, необходимы некоторые более простые средства общения с машиной. Для этой цели разработаны различные виды автоматического программирования, или *автокоды*, в которых команды записываются в обычной словесной или алгебраической форме, а затем машина сама переводит их в эквивалентные варианты машинного кода.

Так, рассмотренный здесь пример машинного кода для ЭВМ «Эллиотт» в варианте автокода записывается просто через x , y и третье число, которое можно обозначить через z , в следующем виде:

$$Z = X + Y$$

(в этом автокоде используются только заглавные буквы). Здесь X , Y и Z обозначают три ячейки, а выражение $Z = X + Y$ представляет собой не уравнение, а команду «заменить содержимое ячейки Z суммой чисел из ячеек X и Y ». Если эту сумму нужно хранить в ячейке, первоначально занятой числом X , то команда записывается в виде $X = X + Y$. Это выражение более наглядно показывает, что мы имеем дело с командами, алгебраическими по

своей форме, хотя это и не настоящая алгебра в истинно математическом смысле. Однако между символическим кодированием и настоящей алгеброй существует настолько близкое соответствие, что, умея обращаться с алгебраическими выражениями, можно легко и быстро перейти к символическому кодированию. Именно поэтому автокоды так популярны.

Таким путем можно записать все основные арифметические действия, но с одним ограничением: одновременно можно выполнять действия только над *двумя* числами. Если требуется сложить три числа A , B и C и поместить результат в ячейку D , то это запишется примерно в таком виде:

$$D = A + B,$$

$$D = D + C,$$

т. е. сумма $A + B$ вначале помещается в ячейку D , а затем к этому числу прибавляется C . Автокод ЭВМ «Эллиотт» предусматривает также возможность выполнения операций над простейшими алгебраическими и тригонометрическими функциями (логарифмическими, показательными, квадратными корнями, синусами, косинусами и т. д.), а также позволяет легко считывать данные и печатать выходные данные.

В качестве простого примера полной программы, записанной через значения X и Y с выдачей на печать результата $\sqrt{X^2 + Y^2}$, можно выбрать следующий:

```

ПЕРЕМЕННЫЕ XYZ
ФУНКЦИЯ КВАДРАТНЫЙ КОРЕНЬ
ЧИСЛО ЦИКЛОВ 1
1) СЧИТАТЬ X
СЧИТАТЬ Y
X = X * X
Y = Y * Y
Z = X + Y
Z = КВАДРАТНЫЙ КОРЕНЬ Z
ПЕЧАТАТЬ Z
ОСТАНОВ
ПЕРЕХОД К КОМАНДЕ 1

```

Мы не предполагаем рассматривать здесь этот вопрос более детально. Однако заметим, что операция умножения обозначается звездочкой, а в начале программы мы «объявляем» названия переменных X , Y и Z и название функции, которую мы собираемся вычислять: КВАДРАТНЫЙ КОРЕНЬ. Такие программы можно

легко составить для выполнения довольно сложных вычислений. Как уже подчеркивалось, при этом необходимо предусмотреть обеспечение точности работы и применять специальные приемы для проверки вычислений.

Одна из самых важных особенностей автоматического программирования состоит в том, что оно позволяет очень просто и компактно записывать команды для выполнения продолжительного цикла повторных операций. Допустим, например, что на магнитной ленте записаны N чисел X_1, X_2, \dots, X_N и требуется найти их сумму S . В автокоде ЭВМ «Эллиотт» эти операции выражаются следующей последовательностью команд (которая может составлять небольшую часть значительно большей программы):

```

 $S = 0;$ 
ИЗМЕНЯТЬ  $I$  ОТ 1 ДО  $N$  С ШАГОМ 1
СЧИТАТЬ  $X$ 
 $S = S + X$ 
ВЕРНУТЬСЯ К  $I$ 

```

Часть программы между ИЗМЕНЯТЬ I и ВЕРНУТЬСЯ К I выполняется многократно для $I = 1, 2, 3, \dots, N$. Каждый раз, когда подается команда СЧИТАТЬ X , с магнитной ленты считывается новое число, а в следующей строке оно прибавляется к числу, уже имеющемуся в ячейке S . Чтобы вычисления начинались с X_1 , перед переходом к *циклу программы* устанавливаются $S = 0$ (в противном случае X_1 прибавлялось бы к числу, которое осталось в ячейке S после выполнения некоторых предыдущих вычислений). Когда данный цикл программы завершен, вычислительная машина переходит к выполнению следующей части программы. Существует несколько различных способов обозначения таких циклов. Возможен также переход назад или вперед в определенные пункты программы при подаче команды перехода.

Существуют самые различные типы автокодов, но обычно каждый из них применим лишь для машин определенного типа или близких по типу машин. Поэтому были созданы более универсальные языки программирования, которые могут применяться в ЭВМ различного типа. Записанная на таком языке программа, разработанная для одной машины, пригодна для управления вычислениями на любой другой машине, которая может перевести универсальный язык в собственный машинный код. В Европе широко используется язык программирования АЛГОЛ (акроним, образованный из начальных букв выражения ALGOrithmic Language — алгоритмический язык). Во-первых, это более мощный и изящный язык для записи порядка вычислений, чем любой автокод. Во-вторых, в настоящее время электронные вычислительные

машины самых различных типов имеют трансляторы (*компилирующие устройства*) для перевода с языка АЛГОЛ, хотя нередко машины несколько отличаются друг от друга по своим характеристикам. Для многих машин наиболее общая форма языка АЛГОЛ (часто обозначаемая АЛГОЛ-60) неприменима, а входные и выходные устройства машин различного типа обычно отличаются друг от друга. Тем не менее благодаря применению языка АЛГОЛ достигается большая универсальность и взаимозаменяемость.

Мы не будем подробно рассматривать здесь язык АЛГОЛ, а ограничимся лишь двумя простыми примерами. Полное изложение этого вопроса можно найти в книгах Дийкстра [18], Мак-Кракена [42] и Наура [46]. Одна из важных особенностей языка АЛГОЛ состоит в том, что алгебраические формулы почти любой сложности могут быть включены в программу в более или менее первоначальном виде, без расчленения их на простые этапы, как в автокоде ЭВМ «Эллиотт». Так, элементы рассмотренной выше программы в автокоде для вычисления $\sqrt{X^2 + Y^2}$ на языке АЛГОЛ для ЭВМ «Эллиотт-803» имеют вид ¹

```
read x, y;
print sqrt (x * x + y * y);
```

а пример программы вычисления суммы, содержащей цикл, принимает вид ²

```
s := 0;
for i := 1 step 1 until n do
begin
read x; s := s + x
end
```

Заметим, что теперь вместо знака = используется :=, что отдельные утверждения отделяются друг от друга точкой с запятой и что слова **begin** и **end** рассматриваются как скобки с тем, чтобы слово **do** относилось к обоим последующим утверждениям.

Другим важным машинным языком, широко используемым в США, а также в Европе и других частях мира, является язык ФОРТРАН (акроним, составленный из начальных букв слов FORMula TRANslator — преобразователь формул), используемый в вычислительных машинах ИВМ. Между сторонниками языков ФОРТРАН и АЛГОЛ идет большая полемика по поводу относительных достоинств каждого из них. В нашу задачу не вхо-

¹ read — считать; print — печатать. — *Прим. ред.*

² for i := 1 step 1 until n do — вычислять $i=1$ с шагом 1 до n ; begin — начало; end — конец. — *Прим. ред.*

дит обсуждение этого вопроса, хотя ради справедливости следует отметить, что оба языка очень удобны для научных вычислений и применяются с большим успехом. Обработка данных представляет собой более сложную задачу; в этом случае выбор языка зависит от объема данных, подлежащих обработке, и от требуемого объема численного анализа. Для решения торгово-промышленных задач очень удобен язык КОБОЛ (COBOL — акроним, составленный из начальных букв выражения COmmon Business Oriented Language — общий язык для коммерческих расчетов). Он составлен главным образом из слов и выражений английского языка, но содержит также некоторые элементарные арифметические операции.

Научиться *кодировать* вычисления в автокоде или на каком-либо другом символическом языке, например АЛГОЛ или ФОРТРАН, относительно несложно. Чтобы стать квалифицированным программистом, необходимы достаточная практика и опыт, однако научиться составлять простые программы можно за несколько часов. Более трудной задачей является составление эффективного алгоритма вычислений, или *программирование* в узком смысле слова. Некоторые вычисления, абсолютно правильные с математической точки зрения, на практике могут оказаться неудовлетворительными вследствие ограничений, свойственных программирующему языку (и ограничений, присущих используемой вычислительной машине). Часто для облегчения выбора методики вычислений крайне важно знать порядок кодирования последовательности арифметических действий. Многие научные работники предпочитают по возможности использовать для кодирования своих задач специально обученный персонал, однако ответственность за общую структуру вычислений все же лежит на них. Поэтому особенно большую роль играет программист, который может разработать общую схему и составить последовательность вычислений, а также обеспечить наиболее эффективное использование средств соответствующего машинного языка.

5.5. ВЫЧИСЛИТЕЛЬНЫЕ МАШИНЫ И НАУЧНЫЕ ИССЛЕДОВАНИЯ

Вряд ли стоит говорить о том, что появление современных вычислительных машин оказало огромное влияние на развитие исследований во всех областях науки, хотя к настоящему времени возможности ЭВМ используются далеко не полностью. В предыдущих разделах были рассмотрены общие принципы вычислительной техники, круг задач, которые могут решаться с помощью ЭВМ, а также методы программирования, посредством которых можно передавать в машину информацию и команды. Появ-

ление вычислительных машин совершило настоящую революцию в области научных исследований. Вычисления могут теперь выполняться значительно быстрее и точнее, чем раньше; вычислительная машина может хранить, обрабатывать и анализировать огромное количество данных, и поэтому перед учеными открываются такие области исследования, которые ранее были просто недоступны. Более того, возможность воспользоваться вычислительной техникой существенно влияет на выбор методов научного исследования и даже на выбор проблем, которые целесообразно исследовать. Причина этого заключается в том, что наличие вычислительных машин не только экономит время научного работника, но и оказывает на него психологическое воздействие.

Обсудим сначала более узкий вопрос о выборе методов решения какой-либо конкретной задачи. Допустим, что изучается развитие эпидемии (более подробно этот вопрос обсуждается в гл. 9). Рассмотрим обычную модель простой стохастической эпидемии для непрерывного времени со случайной передачей инфекции от носителя инфекции к восприимчивым индивидуумам, но без случаев выздоровления (математическое решение этой задачи мы уже затрагивали в разд. 3.4, а вычислительные аспекты отложили до настоящего раздела). Такой процесс можно описать системой дифференциально-разностных уравнений относительно вероятностей. Эти уравнения могут быть точно решены с помощью преобразований Лапласа, хотя формулы для отдельных вероятностей очень сложны и их трудно найти математическим путем. По различным техническим причинам довольно трудно найти численные значения вероятностей и соответствующие математические ожидания и дисперсии путем решения этих уравнений в случае не очень больших групп восприимчивых индивидуумов. Поэтому встает вопрос, не лучше ли, учитывая *существование электронных вычислительных машин*, выбрать другой математический метод.

Если мы собираемся вычислять вероятности, математические ожидания и т. д., то это можно сделать непосредственно с помощью системы дифференциально-разностных уравнений, не прибегая к преобразованиям Лапласа. Решение системы обыкновенных дифференциальных уравнений можно найти стандартными численными методами Рунге — Кутта. Если эпидемия начинается в группе из 50 восприимчивых индивидуумов при одном случае инфекции, занесенной извне, то требуется решить 51 уравнение. В принципе достаточно хорошая вычислительная машина может справиться с этой задачей, однако для обеспечения точности численного анализа необходима большая осторожность, особенно потому, что методы Рунге — Кутта могут оказаться ненадежными, когда решения имеют вид сумм членов с отрицательными показателями степени (что имеет место в данном случае). Возмож-

но, лучшим окажется какой-либо другой метод численного интегрирования, однако, насколько известно автору, весь этот подход в целом еще не нашел практического применения.

Существует другой метод вычислений, основанный на том, что производящая функция распределения вероятностей удовлетворяет некоторому дифференциальному уравнению в частных производных второго порядка параболического типа. Такие уравнения могут быть решены методом Кранка — Никольсона, который позволяет получить численные значения производящей функции для большого числа точек в области интегрирования. Но для нахождения отдельных значений вероятностей или математических ожиданий и т. д. необходимо применять численное дифференцирование, которое пользуется дурной славой из-за своей ненадежности, и, кроме того, для выполнения таких вычислений требуется довольно много времени. Для данного стохастического процесса удается получить достаточно точные значения математических ожиданий, однако необходимо иметь в виду, что мы рассматриваем простую эпидемию без случаев выздоровления. Для более реалистических моделей и этот и все другие упомянутые нами методы при современном состоянии численного анализа и вычислительной техники, по-видимому, малоэффективны.

Рассмотренные численные методы вполне приемлемы при очень малых группах восприимчивых индивидуумов, однако в этих случаях решение можно получить и алгебраическими методами. Что же можно сделать в случае больших групп? Одна из возможностей — использование физического моделирования, рассмотренного в разд. 2.5. Существуют различные способы физического моделирования. Например, можно непосредственно промоделировать процесс для непрерывного времени. Кроме того, можно использовать дискретный аналог цепочно-биномиального типа, а затем промоделировать этот дискретный процесс. Сделать это очень просто. Соответствующую программу для вычислительной машины записать нетрудно, вычисления с требуемой точностью можно относительно быстро выполнить на небольшой вычислительной машине, например IBM-1620 или «Эллиотт-803», и этот метод легко распространить на более общие случаи. Например, оказалось, что он дает возможность без большого труда получить весьма ценную информацию о развитии эпидемий в определенной географической области (см. разд. 9.7).

При решении эпидемиологических задач такого рода физическое моделирование не только во многих отношениях значительно эффективнее других методов, но особенно ценно тем, что оно приложимо и к более сложным и более реалистическим моделям, чем те, с которых мы начинали. Перед научными исследованиями открываются новые перспективы. Возможность изучения

эпидемий в пространственно дифференцированных популяциях, имеющих неоднородную структуру, в настоящее время представляется значительно менее отдаленной, чем при попытках чисто математического анализа. Вычислительная техника существенно повлияла на стандартные математические методы и подсказала, в каком направлении следует искать более плодотворные методы исследования. Более того, она стимулирует и развитие чисто математических исследований, так что эти различные направления взаимно усиливают друг друга.

Примеры, свидетельствующие о влиянии вычислительной техники на научные исследования, можно найти почти в любой области науки. В биологии и медицине применение вычислительных машин должно иметь особенно далеко идущие последствия. И действительно, уже достигнуты поистине замечательные успехи в самых различных областях — от молекулярной биологии до организации управления научными ресурсами в крупном масштабе.

Наряду с конкретным влиянием вычислительной техники на тонкую структуру научных исследований существует более широкий вопрос, связанный с состоянием общественного мнения. Тот, кто действительно понимает, что вычислительная машина представляет собой просто новое средство, в огромной степени облегчающее стоящую перед человечеством задачу изучения и покорения природы, смотрит на новые возможности с оптимизмом и энтузиазмом. Такой человек будет не только приветствовать развитие новых вычислительных машин, но и захочет точно знать, какое влияние они окажут на его работу и каким образом полностью использовать их возможности. Кроме того, он остро заинтересован в том, чтобы как можно большее число людей смогло освоить и использовать новые средства. Но часто — вследствие ли полного невежества или неправильного понимания новых идей и методов — наблюдается поразительное противодействие их распространению. Широкая публика боится развития новых сил, которые она ошибочно представляет себе как нечто мистическое и непостижимое, профсоюзы видят в них новую угрозу жизненному уровню рабочих, а специалисты в области вычислительной техники не желают уступать свое привилегированное положение верховных жрецов, владеющих тайнами, скрытыми от непосвященных. Весь дух настоящей книги направлен против таких тенденций, где бы они ни возникали. По мнению автора, для общества исключительно важно, чтобы как можно больше людей поняло, что могут делать вычислительные машины, и имело возможность пользоваться ими, когда в этом возникнет необходимость.

Эти соображения имеют непосредственное отношение к таким вопросам, как, скажем, вопрос об относительной роли больших и малых вычислительных машин. Большие машины обладают

большим быстродействием, более универсальны, могут обрабатывать большее количество данных и использовать большее число языков программирования, более эффективно выполняют всю последовательность вычислений и т. д. Если такое вычислительное устройство используется эффективно, то эти выгоды оказываются весьма значительными. Но вместе с тем контакт между человеком и машиной значительно ослабляется. Квалифицированный математик, который знает возможности вычислительных машин, после составления программы спокойно передаст технические детали оператору. Однако исследователь, привыкший работать в лаборатории, не будет чувствовать большой уверенности в результатах, полученных с помощью аппаратуры, находящейся где-то в другом месте и управляемой людьми, работу которых он не совсем понимает. Такой специалист будет совершенно иначе относиться к устройствам, которые он может видеть, ходить возле них и обращаться с ними точно так же, как с любым другим сложным оборудованием своей лаборатории. По этим причинам очень важно, чтобы получили широкое распространение небольшие вычислительные машины, работу которых можно тесно увязать с деятельностью их потребителей. По мере того как все большее число людей будет знакомиться с вычислительной техникой, этот аспект станет, вероятно, менее важным, однако в настоящее время он имеет весьма существенное значение. Особенно это относится к биологии и медицине, где необходимо всячески поощрять распространение математических методов. Если врачи и биологи ознакомятся с принципом работы небольших вычислительных машин, они значительно быстрее поймут их значение для своей научной работы. Разумеется, эксплуатацией таких машин должны заниматься специалисты по вычислительной технике, но можно организовать работу так, чтобы они более или менее непосредственно участвовали в научных исследованиях.

Помимо весьма значительного психологического эффекта, к которому приводит установление непосредственного контакта научных работников с вычислительной машиной, небольшие вычислительные устройства имеют и ряд других преимуществ. Так, если программа требует неоднократных усовершенствований, во многих случаях удобнее отлаживать ее на небольших вычислительных машинах, находящихся в самой лаборатории. Прежде всего в сложных программах могут быть ошибки, которые необходимо исправить. Кроме того, даже если программа верна, может оказаться желательным получить результаты немедленно, чтобы их можно было сразу же обсудить, модифицировать программу и получить новые результаты. Если предстоит длинная последовательность обсуждений и модификаций программы, то почти всегда удобнее иметь практически немедленный доступ к небольшой вычисли-

тельной машине, чем использовать большую машину, установленную в другом месте. Исключение составляют лишь те случаи, когда имеется свободный доступ к пульту дистанционного управления большой вычислительной машиной, так как это равносильно наличию вычислительной машины в собственном отделе или лаборатории.

Между тем исследования, требующие многократного изменения программы, обычно сильно задерживаются длительным ожиданием возможности отладить ее на большой центральной вычислительной машине. Одна из основных причин такой задержки заключается в нерентабельности работы большой вычислительной машины, когда эта машина не загружена на полную мощность, т. е. когда ее относительная загрузка далека от единицы. Из теории массового обслуживания (см. разд. 4.3) хорошо известно, что средняя длительность ожидания стремится к бесконечности, если загрузка превышает единицу, и все еще очень высока, если загрузка лишь немногим меньше единицы. Например, в простой системе массового обслуживания со случайным прибытием клиентов и экспоненциальным распределением длительности обслуживания среднее время ожидания, согласно формуле (4.6), равно $\rho\mu/(1 - \rho)$, где ρ — загрузка, а μ — среднее время обслуживания. Если, например, $\rho = 0,95$, то это означает, что среднее время ожидания в 19 раз превышает среднее время обслуживания. Если речь идет о вычислительном устройстве, то в эффективное время обслуживания входят не только те несколько секунд, которые требуются для самих вычислений на машине, но и время, затрачиваемое на выполнение канцелярской и административной работы и на передачу программы и данных из лаборатории на вычислительное устройство. Если не принять специальных мер для его сокращения, суммарное время обслуживания может достигать нескольких часов. Таким образом, при загрузке, близкой к единице, среднее время ожидания может измеряться не часами, а сутками. Поэтому ясно, что увеличение коэффициента использования машинного времени имеет важнейшее, а в некоторых случаях решающее значение.

По всем этим причинам многие достаточно крупные научно-исследовательские группы предпочитают иметь собственные небольшие вычислительные машины, которые могут работать при загрузке, меньшей единицы, и обеспечивать очень быстрый доступ и быструю отладку программы. Кроме того, индивидуальная вычислительная машина обладает значительно большей гибкостью и более пригодна для решения практических задач, особенно тех, которые решаются методами исследования операций. В то же время, конечно, для решения более сложных задач и получения более точных результатов важно иметь доступ к большой быстрой вычислительной машине.

Глава 6

ОРГАНИЗАЦИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ

6.1. ВВЕДЕНИЕ

«Философы лишь различным образом объясняли мир, но дело заключается в том, чтобы изменить его», — гласит одиннадцатый тезис Карла Маркса о Фейербахе¹. Мы уже подчеркивали, особенно в гл. 3, необходимость проверки теоретических построений с помощью эмпирических наблюдений или эксперимента. Однако Маркс идет дальше. Его высказывание означает, что теоретические построения в конечном счете оправданы лишь постольку, поскольку из них следуют выводы, которые могут быть динамически включены в непрерывно изменяющуюся и развивающуюся жизнь общества. Это поднимает ряд важнейших вопросов, связанных с планированием научных исследований и руководством ими.

Нередко указывают, что необычайный взлет творческой мысли в Древней Греции начался совершенно самопроизвольно, без какой-либо сознательной к нему подготовки, и что до сих пор он служит для нас источником вдохновения и многих практически ценных идей. Это, конечно, неоспоримо. Но часто отсюда делают вывод, что творческую и исследовательскую мысль не следует даже и пытаться направлять в какое-либо определенное русло. Ясно, однако, что никакие сколько-нибудь обоснованные заключения на этот счет нельзя делать без учета особенностей данной исторической эпохи.

До древних греков никто столь много не размышлял над самыми различными сторонами деятельности человека. Естественно поэтому, что многие идеи, возникшие в те времена, а затем разработавшиеся с достаточным упорством и энтузиазмом, сохранили свою ценность до наших дней. Однако нельзя забывать, что, хотя в конечном счете некоторые теории древних греков оказались более или менее правильными (как, например, примитивные атомистические идеи Демокрита), многие другие оказались ошибочными. Знаний, накопленных к тому времени, часто не хватало для того, чтобы решить, какие теории верны, а какие ошибочны.

¹ К. М а р к с, Тезисы о Фейербахе, Сочинения, изд. 2-е, т. 3, стр. 4.—
Прим. ред.

Случилось так, что мы задним числом удостоили признания тех мыслителей, идеи которых по чистой случайности оказались правильными. А следовало бы воздать должное всем тем, в чьих трудах мы сегодня находим вдохновение, изобретательность и воображение независимо от справедливости высказанных в них идей. В современном высокообразованном обществе, способном критически оценивать результаты научной деятельности, дело обстоит совершенно иначе. В настоящее время отсутствие определенной системы планирования и организации было бы просто весьма расточительным и, быть может, даже опасным, хотя конкретная форма, которую должна иметь такая система, служит предметом жарких споров. Разумеется, чисто теоретические работы совершенно необходимы, но не менее важны и прикладные исследования. Соотношение этих двух аспектов науки имеет решающее значение, и именно оно является центральной проблемой, рассматриваемой в этой главе.

Далее нужно иметь в виду, что большинство научных исследований обходится очень дорого. Даже затраты на чисто математические исследования не столь уж малы, хотя основными статьями расхода являются лишь заработная плата, помещения для работы и литература. Если же исследования проводятся в лабораториях и требуют специального оборудования, то здесь необходимые затраты столь велики, что главную роль должны играть экономические соображения. Именно в таком положении находятся ядерная физика и космические исследования. Но даже математическая биология может потребовать довольно больших расходов — например, для проведения широких статистических обследований или практических испытаний в рамках исследования операций, особенно если используются большие вычислительные машины. По сравнению с другими областями исследования на математическую биологию в настоящее время выделяется относительно мало ассигнований, но для того чтобы выполнить большую часть тех работ, которые уже сейчас можно предвидеть, потребуются довольно крупные средства. В настоящее время значительная часть научных исследований прямо или косвенно финансируется правительственными организациями или промышленными корпорациями, для которых основную роль играют практические соображения. Правда, как правило, внимание уделяется не только прикладным исследованиям и исследованию операций, но и чисто теоретическим работам, однако вполне естественно, что получить ассигнования на эти последние тем легче, чем теснее они связаны с какими-либо конкретными приложениями.

Очевидно, что решения о том, на какие темы следует направлять отпускаемые средства, должны приниматься руководителями высокого уровня, что неизбежно в значительной мере ограничи-

вает свободу проведения независимых научных исследований. Даже в тех научно-исследовательских институтах, которые считаются самостоятельными, имеются известные ограничения, связанные с основным профилем работы института, но в еще большей степени их свободу ограничивают правительственные организации и фонды, к которым они обращаются за субсидиями. Обычно гораздо большей свободой в проведении научных исследований пользуются университеты, но и им иногда приходится убеждать поддерживающие органы в том, что та или иная работа заслуживает поддержки.

Высказанные выше замечания относятся, разумеется, к любой области научных исследований. Но в этой книге нас интересует прежде всего широкий круг вопросов математической биологии, особенно те вопросы, которые были рассмотрены в предыдущих пяти главах. Поэтому мы рассмотрим более подробно различные виды научной деятельности в этой области.

Так, если взять типичный факультет университета, связанный с математической биологией, биометрией или биостатистикой (впрочем, совокупность математических методов, применяемых в биологии и медицине, может иметь и другое название), то всю его деятельность можно разбить на три взаимосвязанные категории: преподавание, проведение консультаций и научные исследования (разд. 6.2). Далее, хотя некоторые исследования в области математической биологии можно проводить относительно изолированно, обычно, как мы уже неоднократно подчеркивали, необходимо поддерживать контакт с реальными экспериментами и наблюдениями. Поэтому специалисты в области математической биологии должны быть тесно связаны с соответствующими биологическими и медицинскими научными учреждениями, которые часто нуждаются в консультациях по математике и статистике. Но для разработки комплексного подхода требуется более тщательно спланированное сотрудничество (разд. 6.3). Кроме того, когда речь идет об организационных проблемах, важнейшее значение приобретает исследование операций. Оно связано с выполнением совместных научных исследований более сложного типа, так как обычно в работе должно участвовать много специалистов с самой различной подготовкой.

Возможно, а иногда и необходимо, чтобы целые научные учреждения сосредоточивали все свое внимание преимущественно на одном виде деятельности; это может быть преподавание, проведение консультаций, теоретические исследования, прикладные исследования, исследование операций или что-либо еще. Но поскольку все эти виды деятельности взаимосвязаны и успехи в одной области могут оказать существенное влияние на работу в другой, необходимо принять все меры к обеспечению плодотворного взаи-

модействия различных направлений. В небольших учреждениях часто приходится ограничиваться лишь одним-двумя видами научной деятельности, но в более крупных организациях между всеми этими направлениями может поддерживаться некоторое динамическое равновесие. Это значительно облегчит и ускорит восприятие новых идей, их развитие и передачу тем, кто занимается преподаванием и проведением консультаций.

6.2. ПРЕПОДАВАНИЕ, ПРОВЕДЕНИЕ КОНСУЛЬТАЦИЙ И НАУЧНЫЕ ИССЛЕДОВАНИЯ

Как уже указывалось в предыдущем разделе, для того чтобы математические методы приобрели сколько-нибудь существенное значение для биологии и медицины, необходимо показать, что они не только имеют научную ценность, но и дают полезные результаты для общества в целом. Автор считает, что оба эти критерия выполняются. Следовательно, нужно добиться, чтобы с результатами исследований в области математической биологии могли познакомиться как научные работники, так и преподаватели. Первого можно достигнуть с помощью консультаций, понимаемых в весьма широком смысле — от непосредственного участия специалистов по математике, статистике или вычислительной технике в биологических или медицинских исследованиях и до собственно консультаций, когда математика лишь помогает исследователю овладеть тем или иным шаблонным методом расчета. Вторая цель связана с преподавательской деятельностью, и с нее мы и начнем наше обсуждение. (Более подробно вопросы, связанные с преподаванием математической биологии, изложены в трудах Каллоуэйской конференции 1962 г. [41].)

Преподавание

Здесь будет идти речь о том, как знакомить с применением математических методов в биологии и медицине тех, кто желает побольше узнать об этом предмете. Обучение этим методам отличается от преподавания таких конкретных дисциплин, как отдельные разделы прикладной математики или вычислительные методы, которые читаются на самых различных уровнях в школах и университетах. Для того чтобы биолог или врач сумели лучше понять математические методы и научились применять их на практике, преподаватель должен учитывать подготовку слушателей и на занятиях уделять особое внимание биологическим и медицинским задачам.

В точных науках трудности не столь велики, так как большой объем прикладной математики, изучаемой в школе (во всяком

случае, в Англии), автоматически влечет за собой знакомство со многими вопросами математической физики (движение планет вокруг Солнца, основы баллистики, теплопроводность, электромагнетизм и т. д.). Таким образом, те, кто впоследствии специализируется в области экспериментальной, теоретической или математической физики, знакомятся с основными идеями и принципами математического подхода еще на школьной скамье.

Те же, кто решает изучать биологию и медицину, иногда делают это отчасти потому, что их меньше интересует математика или они менее способны к ней. Разумеется, это усугубляет трудности исследовательской работы в этих областях, поскольку, как мы уже подчеркивали, биологический материал часто бывает настолько сложен и изменчив, что для удовлетворительного его описания нужны сложные математические модели, а для обработки и объяснения результатов могут потребоваться довольно сложные математические и вычислительные методы. Поэтому желательно, чтобы тот, кто изучает в школе биологию и получает начальную медицинскую подготовку, намереваясь в дальнейшем специализироваться в этих областях, получал также достаточный объем знаний и по математической биологии.

Вопрос о том, какой объем подготовки считать достаточным, нельзя решать догматически. Возможно, что лица со слабыми математическими способностями не продвинутся дальше курса элементарной алгебры, но, во всяком случае, это даст им возможность разобраться в смысле математических моделей, понять роль теории вероятностей в изучении биологических явлений и принцип работы вычислительных машин. Можно предположить, что другие пойдут дальше и, например, научатся решать элементарные дифференциальные уравнения, применять некоторые простые методы численного анализа и составлять программы для вычислительной машины.

Крайне досадно, что многих студентов, способных изучить основы математической статистики и вычислительной техники, не побуждают к этому или не создают им необходимых для этого условий до тех пор, пока они, поступив в аспирантуру, не начнут заниматься научно-исследовательской работой. Может случиться, что аспиранту в срочном порядке потребуется усвоить довольно большой объем знаний в области математической биологии при работе над темой своей диссертации, и попытка быстро овладеть незнакомым методом исследования и соответствующими специальными методиками окажется для него непосильной. Всего этого можно избежать, если в школьные программы будет введено изучение основных элементов математической биологии, а в программах высших учебных заведений предусмотрено более глубокое ее изучение. Необходимо читать отдельные курсы для тех, кто

получает преимущественно биологическую подготовку, и для тех, кто занимается по программе медицинского учебного заведения. Здесь имеется много общего, однако есть и некоторые специфические требования. Например, врачу необходимо больше знать о методах исследования операций применительно к работе лечебных учреждений и об использовании понятий математической статистики для решения проблем этического характера при испытаниях новых методов лечения, тогда как для биолога большее значение имеет классическая прикладная математика и анализ сложных математических моделей (как детерминистских, так и стохастических).

В будущем преподавание математической биологии в школах и высших учебных заведениях обеспечит широкое распространение ее основных принципов и методов. Однако в ближайшее время необходимо создать различные специальные курсы. Так, студентам и молодым научным работникам нередко читают лекции по статистическим методам в надежде, что и те и другие быстро научатся более эффективно анализировать свои экспериментальные данные. Это, разумеется, весьма ценная и необходимая мера, но обычно бывает очень трудно добиться, чтобы у слушателей за короткое время выработался надлежащий строй мыслей, без которого нельзя обеспечить правильную интерпретацию полученных результатов. Точно так же и студентам-медикам и практических врачей можно познакомить с основными понятиями и результатами применения методов исследования операций и обработки данных. Однако, поскольку они незнакомы с количественным подходом как таковым, им очень трудно будет усвоить эти новые идеи, и они не сумеют использовать их именно в тех случаях, когда это особенно необходимо. В критические моменты вновь приобретенные знания обычно забываются и применяются лишь устоявшиеся методы. Отсюда следует, что методы математической биологии необходимо внедрять медленно и постепенно, в течение долгого времени.

До сих пор мы рассматривали преподавание математической биологии биологам и врачам или тем, кто готовится к этой профессии. И это, по-видимому, самая неотложная задача. Сейчас в области математической биологии работают главным образом люди, имеющие основной специальностью математику, но проявляющие особый интерес к биологии и медицине. Вопрос о том, насколько удовлетворительно такое положение вещей, служит предметом больших споров, и позже мы еще раз вернемся к этому. Хотя вовсе не исключено, что биологи и врачи смогут стать достаточно компетентными в математической биологии, чтобы самостоятельно выполнять всю необходимую работу, сочетая, таким образом, в одном лице несколько специальностей, все же, по-

мнению автора, гораздо лучше, чтобы проблемами математической биологии занимались специально подготовленные люди.

В настоящее время специалистами по математической биологии становятся в основном лица, знающие математическую статистику. Хотя сейчас статистику начинают наконец вводить в школьные программы, нередко за ее изучение берутся лишь в аспирантуре. При этом наблюдается значительный уклон в сторону математической статистики. Равновесие восстанавливается до некоторой степени за счет чтения специальных курсов (и введения экзаменов) по математической биологии, однако ее рассматривают просто как иллюстрацию практического применения математической теории. Лишь немногие студенты решаются выбрать своей специальностью математическую биологию как таковую. Это, по-видимому, в значительной мере обусловлено тем, что они знакомятся с этим предметом сравнительно поздно.

Кроме того, многие профессиональные математики склонны считать, будто чистая математика — это математика высшего сорта и лишь второразрядные математики занимаются математической статистикой. Прикладная статистика котируется еще ниже и совсем низко — исследование операций. Этот интеллектуальный снобизм вреден тем, что отпугивает многих первоклассных математиков от проблем, для решения которых их знания крайне необходимы и с которыми связано множество интересных и увлекательных задач.

В некоторых университетах преподавание математической статистики ведется на высоком уровне и соответствующие курсы содержат большой объем чистой и прикладной математики. С течением времени значительная часть материала, который прежде изучался только в аспирантуре, будет даваться студентам. Это благоприятная тенденция. Точно так же как преподавание математической биологии, многие элементы математической статистики можно ввести в школьную программу.

Если теория вероятностей и математическая статистика будут изучаться в школе в большем объеме, то легче будет включить многие вопросы математической биологии в курсы прикладной математики. Вместо того чтобы концентрировать свое внимание главным образом на физических приложениях, учащиеся, обладающие достаточными способностями, могли бы изучать основы математической генетики или некоторые простые стохастические процессы в популяциях.

Наши конкретные рекомендации заключаются в следующем: математическую биологию необходимо преподавать на всех стадиях обучения, начиная со школы, чтобы уровень знаний повышался постепенно и непрерывно. Занятия необходимо строить, применяясь к потребностям слушателей в зависимости от того,

интересуются ли они в основном математикой или биологией. До тех пор пока этот принцип не будет проведен в жизнь, специфические идеи и понятия математической биологии не станут частью интеллектуального багажа студентов и лишь немногие из них смогут впоследствии достаточно легко овладеть этими важными методами.

Проведение консультаций

По мере того как биологические и медицинские исследования становятся все более сложными и идеи, о которых шла речь в предыдущих главах этой книги, получают все большее распространение, растет потребность в квалифицированном совете и помощи. Мы уже указывали, что сейчас (и вряд ли это положение изменится в ближайшем будущем) биологи и врачи стремятся к тому, чтобы работы в области математической биологии выполняли специалисты, вместо того чтобы браться за них самим. Надо надеяться, что со временем эти две группы специалистов станут ближе друг к другу — биологи будут лучше знать математическую биологию, а специалисты в области математической биологии будут лучше знать биологию. Однако определенное разделение труда здесь, как и везде, представляется более удобным и экономически целесообразным.

Обычно встречается следующая знакомая всем ситуация: исследователь-биолог, работающий в поле или в лаборатории, ставит довольно сложный эксперимент, связанный с большим числом различных факторов, и нуждается в консультации при составлении общей схемы эксперимента, а также при математической обработке результатов. В разд. 2.5 мы уже рассматривали причины, обуславливающие важность планирования эксперимента. Теперь уже многие понимают, что без применения надлежащих статистических методов при постановке сложных и дорогостоящих экспериментов большая часть работы может оказаться напрасной. В связи с этим встает ряд вопросов: какого рода консультации необходимы биологу, какие знания должен иметь специалист в области математической статистики, какая степень сотрудничества необходима и т. д.

Если биолог знаком с принципами планирования эксперимента и может самостоятельно решить, на каком плане следует остановиться, то ему потребуются помощь только в математической обработке результатов. В этом случае математик выполняет лишь незначительные вспомогательные функции. Однако довольно часто бывает, что выбрать наилучший план эксперимента не так-то легко, так как для этого необходимо не только понимать суть исследуемой биологической проблемы, но и иметь

солидные математические знания. В этом случае роль математика становится значительно более важной и он должен оказывать весьма существенную помощь экспериментатору. А иногда математик вынужден принимать еще более непосредственное участие в исследовательской работе. Например, может случиться, что гипотезы, справедливость которых биолог желает установить, невозможно проверить в одном эксперименте; что необходимо разработать новый план эксперимента; что необходимо изменить гипотезы или поставить несколько другие вопросы, выбрать другие математические модели и т. п. В этих случаях очень важно, чтобы биолог и математик тесно сотрудничали друг с другом и хорошо понимали, что именно пытаются делать каждый из них. Это по существу совместное научное исследование, в которое каждый специалист вносит свой собственный вклад. При публикации результатов такого исследования статья должна представляться как совместная работа обоих авторов, тогда как в первых двух случаях обычно достаточно выразить благодарность за помощь в работе.

Изложенные выше соображения, касавшиеся специфических вопросов планирования статистических экспериментов, в равной мере относятся к любому виду работы в области математической биологии. Чтобы специалисты в этой области были способны к тесному сотрудничеству с биологами и врачами, они должны иметь соответствующую подготовку. О некоторых ее аспектах — преподавательской деятельности — уже говорилось выше. Важен также вопрос о том, как должен быть подготовлен специалист в области математической биологии, работающий в качестве консультанта, и чему в свою очередь он должен учить других. Необходимо также рассмотреть положение, которое занимают научные исследования в области математической биологии. Более подробно на этом вопросе мы остановимся в следующем подразделе.

Вообще говоря, желательно, чтобы каждый специалист в той или иной степени принимал участие во всех этих трех видах деятельности.

Хотя некоторая часть консультаций носит чисто вспомогательный характер, поразительно часто в результате беседы между биологами (или врачами) и специалистами в области математической биологии вскрываются непредвиденные затруднения, для разрешения которых необходимы специальные знания. В этих случаях научное сотрудничество в значительной степени представляет собой, так сказать, взаимное обучение. Специалист в области математической биологии не только приобретает ценные биологические или медицинские знания, но и по существу проводит неофициальные учебные занятия по своему предмету.

Если в таких консультациях постоянно нуждаются несколько человек, то целесообразно официально организовать чтение лекций по соответствующим предметам. Это облегчило бы работу по оказанию консультационной помощи, и в то же время проводимые консультации могли бы подсказать форму чтения лекций. И еще одно замечание. Само по себе обучение может оказаться скучным и обескровленным, если оно не будет связано с практикой. Поэтому крайне важно, чтобы все преподаватели независимо от того, работают они в школе или в аспирантуре, принимали определенное участие в консультационной работе. Взаимосвязь консультационной работы и преподавания показывает, что эти два вида деятельности всегда необходимо рассматривать совместно; точное соотношение между ними зависит от обстоятельств.

Научно-исследовательская работа

Преподавание и проведение консультаций связаны в основном с передачей специальных знаний, тогда как научные исследования представляют собой средство получения этих знаний. Одни люди занимаются научно-исследовательской работой, потому что любят ее, другим приходится делать это в силу необходимости. Преподавание математической биологии и проведение консультаций — это особая специальность, требующая глубоких знаний и большого опыта. Но иногда вопросы, задаваемые студентами, или проблемы, возникающие у научных работников, обратившихся за консультацией, могут потребовать постановки специальных научных исследований. Таким образом, преподаватели и консультанты оказываются втянутыми в научные исследования, даже если это и не входит в круг их основных интересов. И наоборот, те, кто занимается в основном научно-исследовательской работой, обычно заинтересованы в том, чтобы передавать другим полученные ими знания и готовить новое поколение научных работников. Таким образом, проведение консультаций и преподавательская деятельность являются почти неизбежными спутниками научно-исследовательской работы. С другой стороны, тесная связь с научно-исследовательской работой оживляет и обогащает обучение и увеличивает его практическую ценность.

Как и в других науках, исследования по математической биологии вносят в предмет дух поиска независимо от того, посвящены ли они углублению уже имеющихся знаний или же изучению пока еще неизведанных областей. Научными исследованиями могут заниматься как отдельные работники, так и группы исследователей; те, кто почти все свое время посвящает научно-исследовательской работе, и те, кто занимается в основном обучением или консультациями; сотрудники университетов, раз-

личных организаций или научных учреждений и т. д. К сожалению, надо признать, что, несмотря на широкие возможности для научно-исследовательской работы по математической биологии, достигнутые к настоящему времени успехи в этой области прискорбно малы.

В предыдущих пяти главах было дано общее представление о математической биологии, а во второй части книги довольно глубоко рассматривается ряд специальных вопросов. Свободный поиск составляет самую суть научной работы, несмотря на то что на него часто накладываются некоторые ограничения общего характера — в отношении как выбора области исследования (разд. 6.1), так и формальных логических требований научного метода (гл. 3).

Научные исследования в области математической биологии удобно разбить на четыре основных направления, соответствующие названиям предыдущих глав, за исключением третьей: 1) преимущественно *математическое* направление (поскольку в математической биологии нужно уметь применять и развивать любые понятия и методы чистой и прикладной математики); 2) *теория вероятностей* и *математическая статистика*, имеющие важное значение вследствие изменчивости всех без исключения биологических явлений; 3) *исследование операций* (поскольку знания в области математической биологии нередко можно применять для отыскания оптимальных схем организации) и 4) *вычислительная техника*, без которой в настоящее время невозможно представить себе ни одно из трех названных здесь направлений.

Любое отдельно взятое исследование может охватывать все эти четыре направления, и для каждого из них может потребоваться специалист, глубоко знающий предмет. Поэтому во многих научных исследованиях принимает участие несколько человек. Иногда один из них является ведущим исследователем, отвечающим за тему в целом, тогда как другие лишь дают квалифицированные консультации по тем или иным узким вопросам. В других случаях имеет место подлинное сотрудничество нескольких научных работников, каждый из которых вносит важный вклад в совместную работу. (Такое положение настолько распространено в математической биологии, что этот случай особо рассматривается в следующем разделе.) Таким образом, для создания условий, в которых научно-исследовательская работа будет плодотворной, необходимо, чтобы в ней были представлены в достаточной степени все четыре перечисленных выше направления. В противном случае исследование может не дать нужных результатов или будет направлено по неверному пути, поскольку один из его существенных элементов не будет проработан достаточно глубоко.

Научно-исследовательская работа, помимо того, что она важна сама по себе, создает условия для обучения новых научных работников. Конечно, основные знания по математической биологии приобретаются в учебных заведениях. Однако посвящение в рыцари науки может состояться только через непосредственное участие в настоящем научном исследовании. Поэтому подготовка молодых научных работников должна быть организована так, чтобы они были тесно связаны с различными лабораториями, в которых ведутся научные исследования в области математической биологии, сами участвовали в некоторых из них и таким образом могли выбрать направление, в котором они хотели бы работать в будущем.

6.3. СОВМЕСТНЫЕ НАУЧНЫЕ ИССЛЕДОВАНИЯ

Хотя в математической биологии широко используются результаты и методы, заимствованные из чистой математики, сама она по существу представляет собой прикладную научную дисциплину.

Как уже указывалось в разд. 6.2 и в других местах книги, преподавание классической прикладной математики строится таким образом, что она включает в себя широкий круг физических приложений. Поэтому специалисты в области прикладной математики нередко сами могут выполнять важные исследования в этих областях, не прибегая к непосредственному сотрудничеству с другими учеными. Кроме того, физики сами чаще всего достаточно хорошо знают математику и поэтому крайне редко нуждаются в математических консультациях. В биологии и медицине дело обстоит совсем иначе. Большинство математиков в процессе своей подготовки не получают глубоких биологических или медицинских знаний, а большинство биологов и врачей не имеют специальной математической подготовки. Поэтому сотрудничество между этими специалистами является важной особенностью почти всех научных исследований в области математической биологии.

Существуют ситуации, когда требуется весьма незначительное сотрудничество. Так, биолог, имеющий некоторую математическую подготовку, сможет довольно точно вывести дифференциальное уравнение в частных производных, описывающее сложный физиологический процесс, однако он не сможет найти его решение. Эту задачу можно передать непосредственно математику с простой просьбой «получить ответ». Такой порядок может оказаться удовлетворительным, если не возникнут какие-либо затруднения. В этом случае работа математика носит преимущественно вспомогательный характер, и настоящего сотрудничества здесь не требуется.

Однако вполне возможно, что для решения уравнений нужны некоторые дополнительные условия или допущения, либо их трудно решить именно в той форме, в какой они представлены. В этом случае математик может ввести дополнительные ограничения или произвести некоторые изменения, позволяющие решить эти уравнения. Но может оказаться, что произведенные им изменения не соответствуют духу первоначальной биологической задачи, и в результате будет затрачено много сил на сложные, но бесполезные математические расчеты в поисках точного решения ошибочной задачи. Для того чтобы математик узнал, что именно в конечном счете допустимо с точки зрения биологии, он должен проявить интерес к самой биологической задаче и познакомиться с ней во всех деталях. Тесное сотрудничество между математиком и биологом должно начинаться по возможности на самом начальном этапе научно-исследовательской работы и продолжаться до ее завершения. Биолог должен быть готов скорректировать или изменить свои концепции и гипотезы в соответствии с возможностями математических и вычислительных методов, а математику не придется двигаться в ложном направлении.

Об аналогичном положении мимоходом упоминалось в разд. 1.2, где приводился конкретный пример проведения консультаций в области математической статистики. Нетрудно привести и другие, самые разнообразные примеры из области математической биологии. По-видимому, целесообразно выработать специальные меры, обеспечивающие надлежащее сотрудничество между математиками и учеными, работающими в других областях.

В значительной мере этому способствует развитие математической биологии как самостоятельной научной дисциплины, требующей проведения консультаций и научных исследований на стыке между математическими и биологическими науками. Специалисты различных направлений объединяются для совместной работы и учатся друг у друга (см., например, разд. 3.4 и 4.4). Часто такое сотрудничество оказывается очень полезным и в других важных областях, возникших на стыке нескольких различных дисциплин. Примером может служить блестящее развитие молекулярной генетики.

Многие университеты (если не большинство) считают своей основной задачей обучение студентов, хотя в последнее время возрастает внимание к подготовке аспирантов и специалистов, уже имеющих ученые степени. С точки зрения общества подготовка нового поколения преподавателей и специалистов имеет, возможно, наиболее важное значение. Однако неизменно важно также и расширение научных исследований. Научная работа должна быть неотделима от других видов деятельности, а не рас-

смагиваться как некая добровольная деятельность. Когда учебные программы становятся крайне напряженными, первыми приносятся в жертву научные интересы преподавателей. Это тормозит научные исследования, ухудшает качество преподавания и может иметь неблагоприятные последствия.

Конечно, многие понимают, что для проведения и стимулирования научных исследований необходимо принимать специальные меры, и поэтому в университетах обычно существуют отделы, занимающиеся в основном научно-исследовательской работой. К сожалению, при существующей нехватке финансовых средств, помещений и оборудования вновь возникающим научным дисциплинам значительно сложнее пробить себе дорогу. Математическая биология находится именно в таком положении. Хотя работы в этой области ведутся уже довольно давно, однако лишь совсем недавно ее начали рассматривать как вполне установившуюся область исследований. Для успешного развития математической биологии необходимо, чтобы этим вопросом занималось значительно больше научных учреждений.

Рассмотрим более детально значение конкретных мероприятий в области научных исследований. Мы уже подчеркивали, насколько важно, чтобы сотрудничество между математиками и биологами (или врачами) начиналось с самого первого этапа работы. Однако довольно часто дело не ограничивается дружеским сотрудничеством двух коллег; может потребоваться совместная работа нескольких специалистов и применение специального оборудования. Допустим, например, что предпринимается широкое исследование экологии двух конкурирующих видов животных в пределах обширного четко ограниченного места обитания. Прежде всего может потребоваться не один, а несколько биологов — зоолог, физиолог, эколог, генетик, биохимик и т. д., так как круг вопросов, подлежащих рассмотрению, зависит от обстоятельств. Состав группы частично зависит от имеющихся специалистов, а частично — от объема лабораторных и полевых исследований. Кроме того, может потребоваться статистическое обследование места обитания, построение достаточно точных стохастических моделей роста популяции (с учетом таких факторов, как размножение, гибель, миграция, взаимодействия хищник — жертва и т. д.), а также применение вычислительной техники. Возможно, достаточно будет группы, состоящей из двух-трех биологов, математика и специалиста по вычислительной технике.

До сих пор мы рассматривали главным образом совместное выполнение биологических или медицинских научных исследований. Другой областью, в которой неизбежна совместная научная работа, является исследование операций. Общие методы

и принципы этой научной дисциплины подробно рассматривались в гл. 4. Мы видели, что часто исследованием операций занимается большой круг специалистов с совершенно различной подготовкой. Любое учреждение обычно занимается рядом различных видов деятельности, и, чтобы разобраться в них, может потребоваться несколько специалистов. В дальнейшем для составления программы исследований необходима помощь одного-двух человек, имеющих специальные знания и опыт в применении методов исследования операций. Статистические обследования, обработка данных, численный анализ, вычислительная техника — все это может потребоваться при выполнении таких исследований. Более того, для проверки правильности предварительных результатов нужно провести испытания в реальных условиях, после чего может выясниться, что было бы полезно осуществить частичную реорганизацию. Естественно, при этом необходимо проявлять максимальную благожелательность, с тем чтобы сотрудники учреждения, работа которого проверяется, охотно шли навстречу всем этим мероприятиям.

Перейдем теперь к исследованиям, более детально рассматриваемым в разд. 12.3 (см. также разд. 4.4). В поисках наиболее разумной организации работы больницы нужно учитывать интересы больных, врачей, медицинских сестер и другого обслуживающего персонала, а также сотрудников, работающих в рентгеновском кабинете, патологоанатомическом отделении, регистратуре и т. д. Кроме того, могут потребоваться специальные знания в области архитектуры и техники. Это немедленно выдвигает важнейшую задачу о составе группы специалистов по исследованию операций.

Конечно, нереально пытаться организовывать группы, в которые входили бы специалисты по всем требуемым областям знания. Такая попытка приведет к тому, что группа окажется слишком большой и успешное руководство ею будет невозможным. Неизбежен определенный компромисс, когда основные направления поручаются отдельным специалистам, каждый из которых имеет также второстепенные обязанности. Так, в состав старшего персонала группы Наффилдского медицинского центра, разрабатывающей проект больницы, входят архитектор с инженерной подготовкой, врач, медицинская сестра, социолог, историк, офицер сухопутных войск в отставке, имеющий опыт исследования операций в вооруженных силах, и статистик, работавший в медицинском учебном заведении. Такой состав группы близок к оптимальному, хотя в зависимости от условий он может варьировать.

Члены группы должны прежде всего изучить и обсудить существующее положение дел, чтобы иметь достаточные осно-

вания принять общую программу научных исследований. По мере появления новых фактов и выдвижения новых предложений необходимо непрерывное общение, способствующее выработке достаточно одинаковых взглядов.

Легко видеть, что при исследовании операций возникают еще более сложные проблемы координации, чем при обычной совместной научно-исследовательской работе. Здесь не только требуется участие большего числа представителей самых разнообразных специальностей, но и существует значительно больше возможных источников критики после опубликования результатов. Так, новые предложения о проекте больничной палаты или целой больницы должны быть приемлемы для медицинского и обслуживающего персонала, хорошо продуманы с точки зрения архитектуры и оправданны с научной точки зрения, не говоря уже об отношении к проекту основного потребителя — больного.

Работы по исследованию операций в таких областях, как, например, промышленность, торговля, военное дело и т. д., в силу необходимости выполняются преимущественно на месте применения и не всегда связаны с научными исследованиями, проводимыми в университетах. Но в целом желательно, чтобы некоторая взаимосвязь все же существовала, поскольку это полезно для обеих сторон. Так, преподаватели и научные работники тех университетов, где есть специальные факультеты исследования операций, могут участвовать в решении производственных задач методами исследования операций — в качестве консультантов или временных членов исследовательских групп. И наоборот, специалисты, работающие в промышленности, могут участвовать в проведении семинаров или читать специальные курсы. То же относится и к использованию методов исследования операций в сельском хозяйстве, рыбном хозяйстве, лесоводстве или здравоохранении.

Часть II

ЧАСТНЫЕ ПРИЛОЖЕНИЯ

Глава 7

ЧИСЛОВАЯ ТАКСОНОМИЯ

7.1. ВВЕДЕНИЕ

Большинство биологических исследований связано с открытием новых организмов или описанием новых признаков, анализом различных биохимических и физиологических процессов, изучением поведения отдельных особей, описанием процессов роста и изменения популяции и т. п. Накоплено огромное количество материала — в форме музейных экспонатов или полевых и лабораторных наблюдений. Однако успехи в области использования и классификации этой информации, достигнутые за последние два столетия, относительно невелики.

Одна из основных проблем биологии состоит в том, чтобы, дав название организму, поместить его в соответствующую систематическую группу. Хотя биномиальная номенклатура Линнея, согласно которой каждый организм имеет двойное название — родовое и видовое, — применяется повсеместно, ее нельзя считать достаточной. Она не позволяет обойтись, с одной стороны, без более крупных категорий — семейств, порядков, классов и т. д., а с другой — без разграничения по более тонким различиям, например расовым и подвидовым. К сожалению, ни одна из этих категорий не имеет четко определенной количественной основы. Хотя они, несомненно, имеют известную ценность в качестве условного шифра и позволяют создать определенную систему обозначений, истинно научное значение их сомнительно. Теория таких методов классификации, или *таксономия*, остается пока чисто интуитивной и является скорее искусством, чем наукой.

В гл. 1 мы говорили о том, что использование математических моделей при анализе биологических явлений позволяет сделать этот анализ более точным. Такие модели могут быть либо чисто статическими, либо динамическими; к первому типу относятся, например, модели пчелиных сотов или модели листорасположе-

ния у некоторых деревьев или других растений; ко второму — модели процессов роста популяций или распространения эпидемий. Эти вопросы довольно подробно освещаются в последующих главах. В настоящей главе мы рассмотрим возможность создания математической основы для всей области таксономии. (Более подробно с этой областью исследований можно познакомиться по книге Сокэла и Снита [59].)

Прежде чем переходить к изложению основных вопросов, дадим вначале определение некоторых фундаментальных терминов, прежде всего таких, как «систематика», «классификация» и «таксономия». В эти термины вкладывают иногда различный смысл, но наиболее распространено следующее их употребление (рекомендованное, например, Симпсоном [58]). Под *систематикой* подразумевается научное исследование многообразия организмов, включающее не только общее описание этих организмов и их расположение по определенным группам, но и объяснение такого расположения. Сюда входит, например, изучение эволюционных механизмов, приводящих к наблюдаемому многообразию. Согласно этому определению, систематика является очень широким понятием. Под *классификацией* (в специальном значении этого термина) понимают расположение организмов по группам на основе определенных критериев, например таких, как морфологическое сходство или как общность происхождения. *Таксономия* определяется как наука о классификации, включающая теорию, правила, методы и приложения. Именно таксономия в таком смысле слова составляет основное содержание настоящей главы. Особенно нас интересуют возможности применения математических методов в этой области.

В разных смыслах применяют также и слово «близость» (*relationship*). Оно может обозначать близость, обусловленную общим происхождением, или же просто указывать на то или иное сходство по каким-то признакам. Во втором случае часто говорят о «*фенетической близости*», о которой судят на основании фенотипов, а не на основании предполагаемого филогенеза.

Наконец, *таксон* — это любая таксономическая группа, получаемая в результате применения определенного метода классификации. Организмы могут быть сгруппированы в один таксон на основании: 1) значительной фенетической близости; 2) наличия признаков, заставляющих предполагать общее происхождение, или 3) действительного родства. Однако обычно информация типа 3 выводится из информации типа 1 или 2, а информация типа 2 может вытекать из информации типа 3. Такого рода порочный круг, т. е. рассуждения, при которых различные критерии не являются полностью независимыми, встречается, к сожалению, слишком часто.

Допустим, что таксон T определен на основе некоторого комплекса признаков C . Все виды, обладающие комплексом признаков C , включаются в таксон T , так что все представители таксона T имеют комплекс признаков C . Тогда каждый вновь рассматриваемый вид будет либо обладать комплексом C , либо не обладать им, хотя в последнем случае он может быть довольно близок к этому комплексу. Фактически сходство видов, лежащих вне комплекса C , с видами таксона T может быть даже большим, чем сходство некоторых видов этого таксона между собой. Из этого следует, что классификация, основанная на некоей ограниченной совокупности признаков, страдает известным формализмом и не дает возможности объективно оценивать положение новых видов.

Создается впечатление, что таксономия пытается принять на себя сразу слишком много различных функций — классификацию, именование, указание сходства и общего происхождения или родства. Можно показать, что практически и теоретически невозможно достигнуть этих целей одновременно. Дело усложняется еще и тем, что в таксономии существуют две четко различимые школы. Сторонники эволюционистской школы считают, что при создании таксонов необходимо прежде всего учитывать филогенетическое происхождение, тогда как старая, эмпирическая школа принимает во внимание только данные, полученные при сравнительном изучении реальных растений или животных. Представители эмпирической школы считают, что таксономия должна стремиться к созданию наилучшей возможной классификации, а исследование филогенеза или механизма эволюции составляет предмет таких специальных дисциплин, как палеонтология или генетика. В то же время представители новой систематики совершенно правильно указывают на то, что многие естественные группы организмов, несомненно, сложились в результате определенных филогенетических процессов.

Эту дискуссию можно продолжать до бесконечности, приводя для иллюстрации множество конкретных биологических примеров. Однако это не входит в нашу задачу. По мнению многих исследователей, особенно сторонников той или иной формы числовой, или нумерической, таксономии, эта путаница возникает вследствие того, что трудно или даже невозможно дать удовлетворительное количественное описание рассматриваемых объектов. Кроме того, часто существует неправильное представление о различии между результатами наблюдений и гипотезой, построенной для их объяснения. В этом плане может оказаться полезным вспомнить общие идеи научного метода, рассмотренного в гл. 3.

Научный процесс в целом — это непрерывная цепь поисков, производимых методом проб и ошибок; он состоит в том, что

некоторым исходным фактическим данным дают предварительное объяснение, выдвигая те или иные гипотезы, которые в дальнейшем отвергают или видоизменяют в зависимости от того, насколько правдоподобными они оказываются при сопоставлении с новыми фактическими данными. Какова роль таксономии в этой схеме? Каковы исходные данные? Каковы гипотезы и объяснения? В чем состоит проверка справедливости гипотез? До тех пор пока не будет получен ответ на эти вопросы, данному предмету вряд ли удастся дать достаточно строгое определение и тем самым обеспечить плодотворное его развитие.

С точки зрения научного метода очевидно, что упомянутые выше сторонники той и другой школы просто делают упор на различных аспектах одного и того же общего процесса. С одной стороны, имеется множество биологических данных, основанных на детальном наблюдении реальных организмов. С другой стороны, для объяснения этих фактов существуют различные гипотезы и теории. Эволюция, видообразование, филогенез и т. д. порождают научные гипотезы, которые строятся специально для объяснения имеющихся данных. Но чтобы объяснить эти данные, их необходимо надлежащим образом упорядочить, систематизировать или редуцировать в статистическом смысле (см. разд. 2.1). Если сюда относится построение некоторой системы таксонов, то сами по себе они не имеют самостоятельной ценности, а важны лишь постольку, поскольку с их помощью легче обеспечить общий прогресс научного исследования. При таком подходе можно будет избежать порочного круга, к которому мы приходим, например, когда пытаемся определять таксоны на основе филогенетических соображений, так как эти последние сами являются гипотезами, для проверки которых необходимо обращаться к таксономическим данным. Таким образом, необходимое условие состоит в том, чтобы наши теоретические построения или гипотезы можно было проверить независимо друг от друга.

Допустим, например, что особи определенного вида кузнечиков, собранные на обширной территории, были на основании их фенетических признаков сгруппированы в несколько таксономических групп. Если бы затем оказалось, что эти группы имеют какие-то неожиданные дополнительные признаки, то можно было бы считать, что эти группы действительно имеют значение, а не служат просто удобными ярлыками. Это могло бы иметь место, если бы вдруг обнаружилось, что разные группы имеют совершенно различную плодовитость или ведут себя по-разному в эволюционном отношении при соответствующем воздействии отбора.

Аналогично допустим, что с помощью некоторого таксономического метода создана новая классификация стрептококков,

отличная от обычной. Если бы оказалось, что этим новым группам соответствуют различия в чувствительности стрептококков к определенным антибиотикам и, таким образом, с помощью новой классификации можно довольно точно определить, какой лекарственный препарат лучше всего назначать в каждом конкретном случае стрептококковой инфекции, то такую классификацию, несомненно, можно было бы считать научно обоснованной.

7.2. ОСНОВНЫЕ ПРИНЦИПЫ КЛАССИФИКАЦИИ

Многие фундаментальные идеи классификации объектов любого рода восходят к понятиям аристотелевой логики. На эти понятия опирались, в частности, авторы первых систем классификации — Чезальпино и даже Линней, — пытавшиеся построить таксономические группы, «сущность» или «истинный характер» которых можно определить. В современной логике свойства, которыми обладают представители такой группы, выводятся из основного определения группы. К сожалению, строго дедуктивные принципы чистой логики нелегко применить к эмпирическому индуктивному миру науки, хотя такие специфические с философской точки зрения термины, как «вид» и «род», были перенесены в таксономию из логики и прижились там.

Постепенно обнаружилось, что ни для какой таксономической группы невозможно найти жесткое определение, выраженное через некоторый единственный диагностический признак или комплекс признаков. Ученые-практики (в отличие от представителей теоретической логики) ощущают, что существуют определенные естественные таксоны, каждый из которых объединяет организмы, обладающие большим числом общих признаков. В некоторых случаях те или иные организмы приходится без всяких сомнений относить к определенному таксону, несмотря на то что у них отсутствует один-два признака, считающихся для него обязательными. Так, у некоторых видов рыб отсутствуют эритроциты, хотя последние обычно считаются признаком, обязательным для всех позвоночных.

Следует различать так называемые *монотетические* и *политетические* группы, т. е. группы, имеющие либо одинаковые, либо различные наборы признаков (эта терминология была введена Снитом). Монотетическую группу получают путем надлежащей последовательности логических делений (классическая аристотелева система, или древо Порфирия), так что обладание определенным комплексом признаков оказывается необходимым и достаточным условием принадлежности экземпляра к данной группе. Построение монотетических групп при анализе фенетических данных всегда сопряжено с большими трудностями, так как орга-

низм, отклоняющийся от нормы по одному-единственному признаку, входящему в определяемый комплекс, будет автоматически исключен из группы. Более того, если этот признак используется для составления первого подразделения ряда, то организм может легко попасть в группу, очень далеко отстоящую от той, с которой он в действительности имеет наибольшее сходство. При такой монотетической классификации не удастся получать группы, соответствующие естественным таксонам.

Политетические группы определяются на основании большого числа общих признаков, ни один из которых не является необходимым или достаточным условием принадлежности к данной группе. Этот принцип значительно лучше согласуется с интуитивными представлениями о естественных таксонах, хотя в этом случае группы могут в некоторой мере перекрываться друг с другом, а построение ключей и иерархий более затруднительно, чем при монотетических классификациях.

Для иллюстрации основного различия между монотетическими и политетическими группами рассмотрим табл. 1, где наличие какого-либо признака обозначается знаком +. Легко видеть, что первые четыре индивидуума составляют политетическую группу, в которой каждый индивидуум обладает только тремя из четырех признаков *A*, *B*, *C* и *D*, но ни один из них не обладает признаками *E*, *F*, *G* и *H*. Другие четыре индивидуума образуют монотетическую группу. Все они обладают признаками *E* и *F* и могут быть подразделены на две самостоятельные подгруппы. В первую подгруппу входят индивидуумы 5 и 6, имеющие признак *G*, а во вторую — индивидуумы 7 и 8 с признаком *H*.

Таблица 1

Таблица, иллюстрирующая различие между монотетическими и политетическими группами (см. текст)

		Признаки							
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>G</i>	<i>H</i>
Индивидуумы	1	+	+	+					
	2	+	+		+				
	3	+		+	+				
	4		+	+	+				
	5					+	+	+	
	6					+	+	+	
	7					+	+		+
	8					+	+		+

Разумеется, на практике число рассматриваемых признаков может быть значительно больше, чем в данном примере, и достигать 50—100. В этих случаях определение таксономических групп соответственно усложняется, особенно если имеются данные по большому числу индивидуумов (последние могут представлять собой отдельные организмы, но обычно это более общие биологические единицы, например виды). Чтобы эта работа не оказалась непосильной, необходимо применение специальных вычислительных методов, которые рассматриваются в последующих разделах.

Какие признаки следует выбирать для исследования в каждом конкретном случае, решить трудно. Характер их может быть самым различным (они могут быть морфологическими, физиологическими, гистологическими, серологическими, цитологическими, поведенческими, экологическими и т. д.) и требовать применения таких специальных методов, как хроматография, электрофорез, инфракрасная спектроскопия и т. п. Вообще говоря, надо думать, что наиболее целесообразно образовывать таксономические группы на основе очень широкого круга признаков. Однако вполне возможно, что таксоны, составленные для некоторой определенной цели, будут отличаться от таксонов, составленных с другой целью. Правильное решение этих вопросов можно получить только на эмпирической основе. Любая система классификации имеет объективную ценность лишь в том случае, если конкретное описание, которое она может обеспечить на основании имеющихся данных, позволяет осуществлять прогнозирование.

7.3. ПРИНЦИПЫ ЧИСЛОВОЙ ТАКСОНОМИИ

Попытки использования некоторых количественных методов в таксономии восходят еще к концу XIX в., когда возникла наука, называемая биометрией. Так, в 1898 г. Гейнке использовал определенный показатель фенетического расстояния, чтобы установить различие между отдельными расами сельдей. Позже при решении ряда задач физической антропологии широко применялся введенный Карлом Пирсоном коэффициент расового сходства, который довольно точно характеризовал степень таксономического подобия. Такого же рода коэффициентом является обобщенный показатель расстояния Махалаанобиса. Однако все эти методы были разработаны в основном для того, чтобы облегчить распределение индивидуумов по уже существующим таксонам.

Нас же интересует вопрос о том, каким образом формировать сами таксоны, пользуясь теми эмпирическими данными, которые непосредственно касаются биологических единиц, подлежащих классификации. Как уже указывалось, эти единицы могут быть

и отдельными организмами (как в случае внутривидовых классификаций), но для сравнения на более высоких уровнях в ботанике и зоологии обычно за единицу принимается целый вид. Хотя по поводу определения некоторых видов имеются известные разногласия, как правило, считают, что виды установлены с достаточной степенью надежности. Основную единицу, используемую в каком-либо определенном таксономическом исследовании, обычно называют *операбельной таксономической единицей*, или О. Т. Е. (operational taxonomic unit — OTU). Изложенные ниже различные методы вычисления сходства относятся к любым операбельным таксономическим единицам, будь то отдельные организмы или какой-нибудь абстрактный класс единиц.

В предыдущем разделе было уже введено важное понятие политетической классификации, основанной на наличных признаках. В числовой таксономии предполагается, что все признаки можно представить в количественной форме. Простейший случай классификации — дихотомическая классификация, производимая в соответствии с наличием или отсутствием того или иного признака; в этом случае наличие может обозначаться, скажем, через 1, а отсутствие — через 0. Признаки могут также образовывать ступенчатый ряд дискретных классов, обозначаемых соответствующими числами, или же представлять собой непрерывные переменные (например, размер или вес). В некоторых случаях, например при вычислении обобщенных расстояний или коэффициентов корреляции, можно оперировать непосредственно с непрерывными переменными. Но для вычисления многих коэффициентов ассоциации необходимо обращаться к дихотомии. Более сложную задачу составляет описание качественных признаков, распадающихся не на две группы, а на большее число групп, расположение которых не характеризуется определенной упорядоченностью. В этом случае может оказаться удобным разделить каждый признак на две или большее число отдельных дихотомических частей. Большое значение имеет проблема правильного кодирования используемых признаков, однако этот вопрос относится не только к работам данного типа. В большей или меньшей степени эта проблема возникает при любой форме количественных исследований. Здесь мы ее рассматривать не будем. Более полное изложение специальных проблем кодирования, встречающихся в числовой таксономии, читатель найдет в разд. 5.3 книги Сокаэла и Снита [59].

Основные этапы, на которые распадается процедура образования таксонов с помощью математических методов, кратко состоят в следующем (детали будут изложены в последующих разделах):

1. Прежде всего выбирают t операбельных таксономических единиц, подлежащих классификации, и по каждой из них исследуют и кодируют достаточно большое число n соответствующих

признаков. Обычно рекомендуется, чтобы n было порядка 50—100 или больше. Если n будет слишком мало, то окончательная конструкция таксонов может оказаться чрезмерно чувствительной к произвольному выбору признаков и будет слишком мала статистическая точность коэффициентов сходства. Исходные данные можно записать в виде матрицы $n \times t$. Очевидно, что в общем случае эта матрица будет содержать весьма значительный объем числового материала.

2. После этого приступают к сравнению всех О. Т. Е. друг с другом, чтобы определить степени сходства на основе соответствующей совокупности n численно закодированных признаков. Одним из простейших показателей сходства некоторой пары О. Т. Е. служит доля признаков, которые у них совпадают (т. е. имеют одно и то же кодированное значение). Однако во многих случаях более точные результаты дает применение иного, более сложного метода (см. разд. 7.4). Веса всех признаков обычно полагают одинаковыми, так как невозможно найти какие-либо разумные доводы в пользу той или иной схемы предпочтений, не впадая в порочный круг. Полученное множество измеренных *коэффициентов сходства* можно представить в виде матрицы $t \times t$. Эта матрица, конечно, симметрична, и поэтому необходимо вычислить и записать только половину ее элементов, лежащих под главной диагональю или над ней. (Элементы матрицы, лежащие на главной диагонали, описывают лишь сходство операльной таксономической единицы с самой собой и не дают никакой специальной информации.)

3. Теперь на основе матрицы вычисленных коэффициентов сходства можно расположить все О. Т. Е. по *группам*. Часто таких групп бывает несколько. Внутри любой группы таксономические единицы обнаруживают значительно большее взаимное сходство, чем единицы, принадлежащие к различным группам. Во многих случаях, объединяя по несколько групп в группы более высокого ранга, можно получить иерархическую структуру. Полную схему такого таксономического «древа» представляют затем в виде *дендрограммы*. Существует целый ряд способов определения групп, и некоторые из них детально рассматриваются в разд. 7.5.

Полученные таким способом группы представляют собой по существу вычисленные таксоны. Сокэл и Снит называют их *фенонами*, поскольку они основаны на чисто фенетических признаках. «Таксон» является более общим термином, обозначающим точно определенную таксономическую группу.

4. Последний этап этой процедуры — проверка правильности классификации путем экспериментальной оценки возможности прогнозирования. Если вновь определенные феноны имеют какую-

либо объективную ценность в том смысле, как это было указано в разд. 7.1, то любой новый индивидуум можно будет поместить в соответствующую группу. Другими словами, для любого нового индивидуума можно будет найти такую группу, что его свойства или признаки будут достаточно точно совпадать со свойствами или признаками, характерными для данной группы.

Все эти операции, за исключением, разумеется, первоначального выбора признаков, могут быть выполнены более или менее автоматически. Другими словами, для этой цели может быть использован весь арсенал мощных математических и вычислительных методов. Таким образом, мы получаем метод таксономического анализа, имеющий четкую логическую основу. Если с его помощью будут получены результаты, имеющие практическую ценность, то тем самым применение данного метода уже будет оправдано. Если же окажется, что полученный результат не достигает цели и не выдерживает критики, то все же можно будет более отчетливо, чем при преимущественно словесном описании, увидеть, в чем здесь ошибка и какие дальнейшие исследования необходимы.

В последующих трех разделах более подробно рассматриваются вопросы, кратко затронутые в пп. 2—4.

7.4. ОЦЕНКА СХОДСТВА

Итак, допустим, что у нас есть исходные данные, записанные в матричной форме для t операбельных таксономических единиц, характеризуемых n признаками; рассмотрим теперь этапы, которые необходимо пройти на второй стадии процесса, описанного в предыдущем разделе. Любой элемент матрицы $n \times t$, например элемент x_{ij} , представляет собой закодированное обозначение i -го признака j -й таксономической единицы. При простой дихотомии элемент x_{ij} может представлять собой знак $+$ или $-$, означающий наличие или отсутствие признака; в этом случае для облегчения числового анализа лучше применять обозначения 1 или 0. Если имеется несколько дискретных классов ступенчатой последовательности признаков, то элемент x_{ij} может быть целым числом, показывающим соответствующий ранг. Кроме того, признак можно выражать в процентах или же пользоваться непосредственными результатами измерений в случае непрерывной переменной.

Существует три основных способа количественного выражения сходства — с помощью коэффициентов ассоциации, коэффициентов корреляции и показателей расстояния. Рассмотрим эти методы по порядку. Подробное их изложение можно найти в гл. 6 книги Сокэла и Снита [59].

Коэффициенты ассоциации

В ботанике, зоологии, экологии, психологии, антропометрии и т. д. используется множество различных коэффициентов ассоциации. Некоторые из них совершенно независимы, а некоторые связаны друг с другом. Однако все они основаны на сравнении двух множеств признаков для какой-либо определенной пары операбельных таксономических единиц (или каких-либо других объектов, используемых в соответствующей области исследования) и выражаются через число совпадений и расхождений в наблюдаемых признаках. Рассмотрим, например, j -ю и k -ю таксономические единицы. Пусть число признаков, имеющихя у обоих О. Т. Е., равно n_{11} , а число «отсутствующих» признаков равно n_{00} ; общее число признаков, имеющихя у первой и отсутствующих у второй таксономической единицы, обозначим через n_{10} , а обратную комбинацию — через n_{01} . (В данном случае нет необходимости употреблять символы j и k для точного различения j -й и k -й О. Т. Е.) Эти числа представлены в табл. 2, где указаны также суммы по рядам и столбцам. Для этих сумм используются принятые в математической статистике обозначения, т. е. $n_{00} + n_{01} = n_{0\cdot}$ и т. д.

Таблица 2

Сравнение признаков для двух таксономических единиц

		k -я О. Т. Е.		
		0	1	Сумма
j -я О. Т. Е.	0	n_{00}	n_{01}	$n_{0\cdot}$
	1	n_{10}	n_{11}	$n_{1\cdot}$
Сумма		$n_{\cdot 0}$	$n_{\cdot 1}$	n

Принято также записывать число совпадающих и несовпадающих пар признаков, обозначаемых соответственно m и u , где

$$m = n_{00} + n_{11}, \quad u = n_{01} + n_{10}, \quad m + u = n. \quad (7.1)$$

Таким образом, существуют m признаков, по которым в обоих О. Т. Е. имеется совпадение, т. е. у обоих 0 или 1, и u признаков, по которым имеется расхождение.

Простейший коэффициент ассоциации, о котором уже мимоходом упоминалось в предыдущем разделе, представляет собой просто относительное число совпадающих признаков

$$S_1 = \frac{m}{n}. \quad (7.2)$$

Это отношение имеет следующее свойство: $S_1 \rightarrow 1$ при $m \rightarrow n$, т. е. при $u \rightarrow 0$; $S_1 \rightarrow 0$ при $m \rightarrow 0$. Таким образом, полное совпадение выражается единицей, а полное несовпадение — нулем.

Существует много различных коэффициентов, и выбор того или иного зависит до некоторой степени от практической целесообразности. Некоторые коэффициенты по существу эквивалентны, другие значительно отличаются друг от друга, в частности в отношении совпадений по отсутствию признаков. Так, коэффициент S_1 основан как на n_{00} , так и на n_{11} . Допустим, однако, что сравниваемые таксономические единицы имеют довольно отдаленную связь, например одна из них представляет собой некоторый вид млекопитающих, а другая — некоторый вид червей. Целесообразно ли в этом случае приписывать им некоторую степень сходства на основе того, что и у тех и у других *отсутствуют* крылья? (И наоборот, имеет ли смысл приписывать сходство по *наличию* крыльев летучим мышам, птицам и насекомым?) Аналогично в бактериологии принято применять определенные стандартные тесты к самым разнообразным группам бактерий. Это может привести к появлению большой группы отрицательных признаков, учет которых искусственно приведет к увеличению степени сходства. В таких случаях исключение совпадений по отсутствию признаков, несомненно, будет разумной мерой предосторожности. В каждой работе нужно принимать решение, исходя из конкретных условий, в зависимости от конкретных биологических свойств исследуемого материала.

При исключении совпадений по отсутствию признаков предпочтительнее использовать коэффициент

$$S_2 = \frac{n_{11}}{n_{11} + u}. \quad (7.3)$$

Как и ранее, при полном совпадении $u \rightarrow 0$ и $S_2 \rightarrow 1$, а при полном несовпадении $n_{11} \rightarrow 0$ и $S_2 \rightarrow 0$.

Для сравнения коэффициентов ассоциации необходимо, чтобы все они измерялись в единой шкале. Удобнее всего использовать шкалу, заключенную в интервале между 0 и 1. И S_1 и S_2 являются коэффициентами именно такого типа.

Используются и другие коэффициенты, например коэффициент

$$S_3 = \frac{m}{n + u}, \quad (7.4)$$

введенный Роджерсом и Танимото; здесь несовпадающие пары имеют в знаменателе вдвое больший вес, чем совпадающие. Этот коэффициент также изменяется в интервале от 0 до 1, однако сомнительно, чтобы он давал какой-либо выигрыш.

Приведем также известный коэффициент ассоциации Юла

$$S_4 = \frac{n_{11}n_{00} - n_{01}n_{10}}{n_{11}n_{00} + n_{01}n_{10}}. \quad (7.5)$$

Он принимает значение $+1$ при $n_{01} = 0$ или $n_{10} = 0$ и -1 при $n_{11} = 0$ или $n_{00} = 0$; его равенство нулю указывает на статистическую независимость признаков. В некоторых случаях этот коэффициент может оказаться очень полезным, и его можно применять, используя преобразование $S_5 = 1/2 (S_4 + 1)$, с тем чтобы он изменялся в интервале от 0 до 1. И в этом случае нет уверенности в целесообразности дополнительного усложнения вычислений.

Продолжать обсуждение различных коэффициентов ассоциаций не имеет смысла (в книге Сокэла и Снита приводится не менее 16 таких коэффициентов). При современном состоянии вопроса целесообразнее всего пользоваться самыми простыми из них. В случае дихотомических признаков, по-видимому, в большинстве случаев вполне пригодны коэффициенты S_1 или S_2 . Однако со временем, возможно, удастся исследовать математически относительную эффективность различных коэффициентов, и может оказаться, что более сложные коэффициенты позволяют точнее оценивать степень сходства или различия таксономических единиц.

При рассмотрении признаков, имеющих более двух состояний, способ описания, естественно, зависит от того, появляются ли эти состояния в определенной последовательности или такой последовательности не наблюдается. В первом случае, пожалуй, удобнее всего применять описываемые ниже коэффициенты корреляции или показатели расстояния. Во втором необходимо либо использовать обобщенную форму рассмотренных выше коэффициентов, либо попытаться разбить признаки на две или даже несколько отдельных групп. Последнее, по-видимому, более предпочтительно, так как применение в числовой таксономии двух существующих обобщенных показателей (обобщенного коэффициента Роджерса и Танимото и таксономического отношения Смирнова) встречает ряд возражений¹.

¹ Автор имеет в виду критику «таксономического отношения», предложенного Е. С. Смирновым, со стороны Сокэла и Снита (1963). Исчерпывающий ответ на эту критику дан в книге Е. С. Смирнова «Таксономический анализ» (1969).— *Прим. ред.*

Коэффициенты корреляции

Коэффициенты корреляции уже давно используются для измерения степени связи между двумя признаками. Обычно их применяют к непрерывным переменным (или, во всяком случае, к дискретным переменным, характеризующимся некоторой упорядоченностью). Лучшее из известных коэффициентов корреляции смешанных моментов, который легко использовать для измерения сходства между двумя О. Т. Е. следующим образом. Допустим, что все элементы основной матрицы $n \times t$ являются непрерывными переменными. Тогда для j -й таксономической единицы мы имеем n измерений, обозначаемых x_{ij} , $i = 1, \dots, n$. Пусть среднее из этих величин равно \bar{x}_j . Аналогично для k -й таксономической единицы имеются измерения x_{ik} , $i = 1, \dots, n$ со средним \bar{x}_k . Наблюдаемый коэффициент корреляции смешанных моментов между этими двумя группами измерений равен

$$r_{jk} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\left[\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 \right]^{1/2}}. \quad (7.6)$$

Этот коэффициент изменяется от -1 в случае полной отрицательной корреляции до $+1$ в случае полной положительной корреляции. Определить точный его смысл часто бывает трудно, однако ясно, что это определенный численный показатель ассоциации. Если допустить, что рассматриваемые случайные величины характеризуются двумерным нормальным распределением, то можно применить статистические критерии значимости и установить, например, является ли значимым отличие данного коэффициента от нуля или различие двух коэффициентов друг от друга. К сожалению, для распределений, отличных от двумерного нормального, такая проверка пока невозможна.

Если отдельные переменные получены в результате измерений самого различного характера, то в этом случае нельзя предположить даже приближенно, что некоторая совокупность случайных величин x_{ij} (с варьированием по индексу i) имеет какое-либо закономерное распределение, не говоря уже о нормальном. Некоторого улучшения можно добиться путем нормировки элементов по рядам. Это означает, что все величины в данном ряду измеряются относительно их среднего значения и делятся на их среднее квадратическое отклонение. Полученные таким образом новые величины, скажем z_{ij} , имеют нулевое среднее значение и единичное среднее квадратическое отклонение. Следовательно, в этом случае все n величин в любом столбце имеют нулевое математическое

ожидание и единичное среднее квадратическое отклонение. Хотя, вообще говоря, нормировка желательна, так как при этом материал становится более однородным, во многих случаях, когда речь идет о фенетических соотношениях, она не дает ничего существенно нового.

Показатели расстояния

Еще один способ установления степени сходства двух таксономических единиц состоит в использовании некоторого показателя «расстояния» между ними. Поясним это понятие, рассмотрев для простоты случай, когда имеется всего три признака (т. е. $n = 3$). Пусть измерения показали, что эти признаки характеризуются значениями x_{1j} , x_{2j} и x_{3j} для j -й О. Т. Е. и x_{1k} , x_{2k} и x_{3k} для k -й О. Т. Е. Тогда эти две таксономические единицы можно однозначно изобразить в трехмерном пространстве точками (x_{1j}, x_{2j}, x_{3j}) и (x_{1k}, x_{2k}, x_{3k}) , расстояние между которыми Δ_{jk} определяется выражением

$$\Delta_{jk}^2 = (x_{1j} - x_{1k})^2 + (x_{2j} - x_{2k})^2 + (x_{3j} - x_{3k})^2. \quad (7.7)$$

Если бы признаков было бы только два, то каждую таксономическую единицу можно было бы представить точкой на двумерном графике. В общем случае число признаков n очень велико, и для изображения таксона необходимо n -мерное пространство. Обобщенное выражение для расстояния между j -й и k -й таксономическими единицами в этом пространстве имеет вид

$$\Delta_{jk}^2 = \sum_{i=1}^n (x_{ij} - x_{ik})^2. \quad (7.8)$$

Если значения x_{ij} и x_{ik} для всех i (т. е. для каждого признака) в точности совпадают, то расстояние Δ_{jk} между точками равно нулю. Любое различие приводит к возрастанию этой величины, и потому интуитивно представляется вполне разумным рассматривать расстояние как показатель различия.

Определенные проблемы возникают при нормировке, поскольку относительные веса различных признаков очень сильно зависят от используемой шкалы. Один из возможных способов состоит в кодировании каждой случайной величины таким образом, чтобы ее значения лежали в интервале от 0 до 1. Это всегда можно выполнить путем соответствующих математических преобразований. Если, например, x изменяется от 0 до ∞ , то $y = x/(a + x)$ (где a — некоторая произвольная положительная постоянная) изменяется от 0 до 1. Но если мы изменим шкалу таким образом, что наибольшее наблюдаемое значение не будет превосходить единицы, то при рассмотрении новых таксономических единиц с более высокими

максимальными значениями x всегда будет существовать опасность того, что значительную часть данных потребуется перекодировать заново. Другой метод состоит в нормировке данных по рядам матрицы, как это было описано при рассмотрении коэффициентов корреляции. В этом случае все признаки имеют одно и то же математическое ожидание и одно и то же среднее квадратическое отклонение, равные соответственно нулю и единице, что придает им значительно большую взаимную эквивалентность. При добавлении новой О. Т. Е. возникает еще одна проблема теоретического характера, поскольку первоначально определенные средние значения и средние квадратические отклонения всех рядов при этом изменятся и все элементы матрицы придется нормировать заново. Однако введение только одной-двух новых таксономических единиц должно потребовать лишь ничтожных изменений и на практике ими можно будет пренебречь. Так, на основе первоначальной нормировки необходимо пересчитывать данные только по новым таксономическим единицам.

Если мы решаем оперировать с нормированными измерениями z_{ij} , то введенную ранее формулу (7.8) для расстояния нужно заметить формулой

$$\Delta_{jk}^2 = \sum_{i=1}^n (z_{ij} - z_{ik})^2. \tag{7.9}$$

Величина Δ_{jk} зависит не только от самих результатов измерений z_{ij} , но и от числа сравниваемых признаков n . Лишь в одном случае, когда сравнения между всеми парами таксономических единиц основаны на одном и том же значении n , это не имеет существенного значения. Для облегчения сравнения в тех случаях, когда используются различные значения n , лучше вычислять средний квадрат расстояния

$$d_{jk}^2 = \frac{\Delta_{jk}^2}{n}. \tag{7.10}$$

С точки зрения числовой таксономии показатели таксономического расстояния (7.9) и (7.10) являются простейшими и, по-видимому, наиболее удобными. В частных случаях используются некоторые другие показатели расстояния или связанные с ними коэффициенты (например, коэффициент расового сходства Карла Пирсона), однако вряд ли есть необходимость рассматривать их здесь. В частности, для диагностики широко используется *обобщенный показатель расстояния Махаланобиса* D^2 . Однако для рассматриваемых здесь проблем числовой таксономии этот коэффициент не подходит, так как мы не учитываем колебаний данного признака внутри одной таксономической единицы.

Если в каком-либо частном случае все признаки имеют одинаковый характер, т. е. все они непрерывны или дискретны, то можно выбрать метод измерения сходства, который будет приемлем в общем случае. Но если одни признаки дискретны, а другие непрерывны, то возникают известные трудности (аналогичная проблема встречается при использовании математических методов для диагностики; см. гл. 11). В этом случае можно либо применять для всех признаков коэффициенты корреляции (или показатели расстояния), рассматривая непрерывные случайные величины обычным способом, а дискретные — описывая числовыми значениями 0 и 1, либо произвольным образом разбить непрерывные случайные величины на дихотомические классы и для каждого признака использовать коэффициент ассоциации. С теоретической точки зрения ни один из этих методов нельзя считать удовлетворительным, однако на современном начальном этапе развития числовой таксономии они могут иметь известную практическую ценность. По мере того как точность числовой таксономии будет повышаться, потребуются, конечно, более точные методы анализа.

Для тех, кто знаком с выполнением статистических исследований, ясно, что любой процесс получения оценок связан с такими понятиями, как значимость и точность. Можно, например, поинтересоваться, значимо ли различие двух коэффициентов сходства. До сих пор в этом отношении сделано очень мало. Нелегко разработать математическую модель, которая бы учитывала совместное влияние таксономических единиц и признаков на результаты измерений и удовлетворительно описывала бы изменения, ожидаемые при повторном выборочном измерении. Кроме того, нас интересует не сравнение каких-либо двух коэффициентов сходства, а оценка всей $t \times t$ -матрицы коэффициентов (см. разд. 7.5). Эта задача теории выборочного метода весьма сложна, и она может приобрести исключительно важное значение при более широком использовании методов числовой таксономии. Мы описали здесь применение нормировки применительно к некоторым распределениям, хотя это выполнялось в довольно эвристической манере, а не на основе какой-либо точно сформулированной модели повторного выбора. Поэтому на данном этапе описанные здесь методы необходимо рассматривать как полезное средство, точность которого неизвестна.

7.5. ПОСТРОЕНИЕ ФЕНОНОВ

Итак, у нас уже имеется матрица $t \times t$, составленная из вычисленных коэффициентов сходства, и мы можем перейти к третьему этапу процесса, описанного в разд. 7.3. В дальнейшем мы будем считать, что все коэффициенты нормированы и заключены между

0 и 1. Как уже указывалось в предыдущем разделе, таким образом обычно нормируют коэффициенты ассоциации, связанные с дискретными случайными величинами; коэффициенты корреляции лежат между -1 и $+1$, а показатели расстояния могут лежать совсем в другом интервале. Если для всех случайных величин используются коэффициенты одного типа, то никаких затруднений не возникает, поскольку их можно сравнивать друг с другом, какая бы шкала ни была выбрана. Если же в одном и том же исследовании используются коэффициенты разного типа, то их можно сделать сравнимыми, произведя некоторые простые преобразования. Это очевидный, хотя и важный пункт.

С помощью матрицы сходства мы хотим разбить операбельные таксономические единицы на такие группы, чтобы единицы одной группы обладали близким сходством, а между единицами, принадлежащими к разным группам, сходство было значительно меньшим. Проиллюстрируем эту операцию на искусственном примере с небольшим числом таксономических единиц. Допустим, что у нас есть 10 О. Т. Е., для которых нижняя половина 10×10 -матрицы коэффициентов имеет вид, показанный в табл. 3, где для простоты все коэффициенты округлены до 0,1.

Таблица 3

Типичная 10×10 -матрица коэффициентов сходства для 10 О. Т. Е.

Номер О. Т. Е.	1	1,0									
	2	0,9	1,0								
	3	0,6	0,5	1,0							
	4	0,5	0,6	0,5	1,0						
	5	0,9	0,8	0,6	0,5	1,0					
	6	0,6	0,6	0,8	0,5	0,6	1,0				
	7	0,6	0,6	0,8	0,5	0,5	0,8	1,0			
	8	0,5	0,5	0,5	0,9	0,5	0,5	0,5	1,0		
	9	0,8	0,9	0,6	0,5	0,9	0,6	0,6	0,6	1,0	
	10	0,6	0,6	0,7	0,6	0,6	0,8	0,7	0,5	0,6	1,0
		1	2	3	4	5	6	7	8	9	10
		Номер О. Т. Е.									

Минимальное значение коэффициента сходства в табл. 3 равно 0,5. Произвольно примем 0,5 и 0,6 за относительно слабое, а 0,7, 0,8 и 0,9 (особенно последние два значения) за относительно сильное сходство. Изучение табл. 3 показывает, что О. Т. Е. № 1 имеет близкое сходство с О. Т. Е. № 2, 5 и 9, причем послед-

ние три имеют очень близкое сходство друг с другом; связь любой из этих четырех О. Т. Е. с любой из остальных шести значительно слабее. Таким образом, О. Т. Е. № 1, 2, 5 и 9 выглядят как особая группа. Кроме того, следует выделить О. Т. Е. № 4 и 8. Они сильно связаны друг с другом, но не связаны с какими-либо другими О. Т. Е. Наконец, оставшиеся О. Т. Е. № 3, 6, 7 и 10 также можно собрать в группу, хотя в этом случае два «внутренних» коэффициента сходства равны всего 0,7. Разобьем эти 10 О. Т. Е. на три названные группы. Тогда матрица примет вид, показанный в табл. 4.

Таблица 4

**Перегруппированный вариант матрицы
коэффициентов сходства,
изображенной в табл. 3**

Номер О. Т. Е.	1	1,0								
	2	0,9 1,0								
	5	0,9 0,8 1,0								
	9	0,8 0,9 0,9 1,0								
	3	0,6 0,5 0,6 0,6				1,0				
	6	0,6 0,6 0,6 0,6				0,8 1,0				
	7	0,6 0,6 0,5 0,6				0,8 0,8 1,0				
	10	0,6 0,6 0,6 0,6				0,7 0,8 0,7 1,0				
	4	0,5 0,6 0,5 0,5				0,5 0,5 0,5 0,6				1,0
	8	0,5 0,5 0,5 0,6				0,5 0,5 0,5 0,5				0,9 1,0
	1	2	5	9	3	6	7	10	4	8

Номер О. Т. Е.

Структура исходной матрицы теперь довольно ясна. Правда, нам точно не известна статистическая значимость этих трех групп, которые можно рассматривать как три фенона, определенных на основе первоначальных данных о признаках при 70%-ном уровне значимости, но, во всяком случае, у нас теперь есть основа для числовой таксономической классификации.

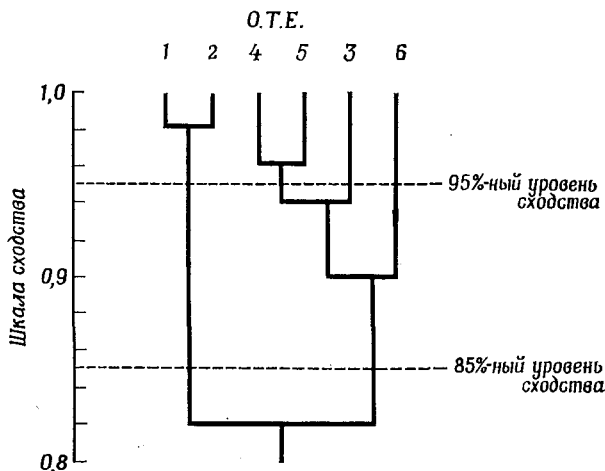
К сожалению, описанный выше процесс распределения операционных таксономических единиц по группам, основанный на визуальном изучении матрицы коэффициентов сходства, крайне субъективен и его трудно выполнить, когда число таксономических единиц больше 10. Сокэл и Снит [59] описывают метод, который по существу аналогичен данному и в котором интервал значений коэффициентов сходства указывается с помощью квадратов, имеющих различную густоту штриховки. В этом случае

быть помещена только такая новая таксономическая единица, у которой коэффициент сходства с любым членом группы достигает требуемого уровня. Естественно, что это довольно строгое условие приводит к образованию довольно компактных, четко определенных групп.

Более приемлемым следует считать критерий «средней связи» («average-linkage criterion»), предложенный Сокэлом и Миченером. Как следует из названия, этот критерий допуска новой таксономической единицы в группу основан на ее среднем сходстве со всеми представителями этой группы. Существует несколько способов формирования групп в этом случае. Один из простейших (и наиболее легкий для программирования при использовании ЭВМ) — так называемый взвешенный парно-групповой метод («weighted pair-group» method). Согласно этому методу, первую группу образуют из двух наиболее близких таксономических единиц. Затем матрицу коэффициентов преобразуют путем замены первой пары таксономических единиц соответствующей группой. После этого вычисляют новые коэффициенты сходства, показывающие связь между данной группой и любой другой таксономической единицей. Так, если О. Т. Е. № 1 и 2 образуют первую группу (обозначим ее, скажем, № 1'), то коэффициент сходства для О. Т. Е. № 3 и группы № 1' равен среднему значению коэффициентов сходства для О. Т. Е. № 3 и 1 и О. Т. Е. № 3 и 2. Изучение новой матрицы с целью нахождения максимального коэффициента сходства показывает, какая таксономическая единица должна объединяться с группой № 1' или какие две таксономические единицы должны образовать следующую новую группу. Этот процесс можно выполнять просто и однозначно шаг за шагом, образуя на каждом этапе вычислений только одну новую группу или включая одну новую таксономическую единицу в существующую группу. Разумеется, существует ряд технических деталей, связанных, в частности, с тем, используются ли коэффициенты ассоциации, коэффициенты корреляции или же показатели расстояния. В нашу задачу не входит обсуждение всех этих вопросов, поскольку нас интересуют лишь основные принципы числовой таксономии. Более полно этот вопрос рассматривается в упомянувшейся ранее книге Сокэла и Снита [59], особенно в приложении А.3, где дается несколько конкретных примеров.

Результаты таких вычислений удобно представлять схематически в виде дендрограммы, показывающей, при каком уровне сходства образуются различные группы. Допустим, что применение взвешенного парно-группового метода к шести таксономическим единицам дало результат, изображенный на фиг. 1, где порядок таксономических единиц выбран лишь для графического удобства. Из этой схемы можно видеть, что наибольшим сходством

обладают О. Т. Е. № 1 и 2 (коэффициент сходства 0,98). После объединения этих двух таксономических единиц и вычисления новой матрицы коэффициентов оказывается, что наибольшим сходством обладают О. Т. Е. № 4 и 5 (коэффициент сходства 0,96). Таким образом, эти таксономические единицы образуют вторую



Ф и г. 1.

группу. На следующем этапе находим, что О. Т. Е. № 3 объединяется с группой, состоящей из О. Т. Е. № 4 и 5, при уровне 0,94, а затем при уровне 0,90 добавляется О. Т. Е. № 6. После этого группа, состоящая из О. Т. Е. № 1 и 2, объединяется с группой, образованной О. Т. Е. № 3—6, при коэффициенте сходства 0,82. Таким образом, эта диаграмма изображает построенное в результате вычислений таксономическое древо, в котором различные группы определяют вычисленные таксоны или феноны.

Ясно, что рассмотренная выше методика автоматически выявляет иерархическую структуру, которую можно довольно гибко описать через номенклатуру фенонов. Допустим, например, что мы оцениваем дендрограмму при 95%-ном уровне сходства. Посмотрим, где этот уровень пересекает древо. На фиг. 1 он отделяет феноны, образованные О. Т. Е. № 1, 2 и 4 и 5, а О. Т. Е. № 3 и 6 остаются разделенными. Однако при 85%-ном уровне можно различить только два фенона. Один из них составлен, как и ранее, из О. Т. Е. № 1 и 2, а другой состоит из остальных четырех О. Т. Е. и имеет определенную внутреннюю структуру. Очевидно, что можно образовать группы, или феноны, более

высокого «ранга», чем те, которые получены при первой группировке основных таксономических единиц. По своему общему виду таксономическое древо, полученное вычислительными методами, и древо, полученное преимущественно на основании биологических соображений, несомненно, близки. Вопрос о надежности таких построений имеет исключительно большое значение, и в следующем разделе мы обсудим с этой точки зрения некоторые результаты, полученные к настоящему времени с помощью методов числовой таксономии.

7.6. ПРИЛОЖЕНИЯ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В настоящее время ведутся большие споры о ценности и потенциальных возможностях числовой таксономии для биологической классификации, и, безусловно, еще слишком рано делать какие-либо окончательные выводы. Тем не менее уже сейчас следует дать числовому подходу некоторую предварительную оценку, особенно если учесть, что он создает для таксономии значительно более точную научную основу, чем та, которая существовала до настоящего времени.

Прежде всего нужно сказать, что методы, описанные в последних трех разделах, уже были применены для классификации самых разнообразных организмов. Например, в зоологии такие авторы, как Миченер, Рольф и Сокэл (более детальные ссылки на эти работы приводятся в книге Сокэла и Снита [59], а также в других работах, упоминаемых далее), исследовали 97 видов пчел комплекса *Hoplitis*, объединяющего четыре рода семейства Megachilidae. Вычисленные коэффициенты сходства и полученные дендрограммы хорошо согласуются с системой, предлагаемой ортодоксальной таксономией. Кроме того, были получены новые данные о тонкой структуре иерархии таксонов и о положении в системе некоторых отклоняющихся видов. Выполненное Рольфом числовое исследование 48 видов комаров рода *Aedes* также привело к системе, весьма похожей на ту, которая явилась результатом других исследований этой группы. Аналогичным образом Литгл, используя методы числовой таксономии, проанализировал на основе 36 признаков 103 вида губок рода *Cliona* и пришел к системе, хорошо согласующейся с современными таксономическими идеями.

В ботанике применение числовых методов также дало хорошее совпадение с существующей системой, особенно в работе Морисима и Ока по исследованию различных видов риса (*Oryza*) и в работе Сориа и Хейзера по исследованию рода *Solanum*.

Значительную работу выполнили Снит и многие другие исследователи по применению числовых методов при классификации

бактерий. Некоторые ранее полученные таксономические группировки подтвердились, а в некоторых случаях запутанная классификация была перестроена и уточнена.

Кроме того, в области клинической гематологии Хейхо, Квалино и Долл [34] использовали числовые таксономические методы при изучении случаев острого лейкоза. Они получили четыре четко определенные группы (с некоторым перекрытием), которые более или менее соответствуют традиционной классификации, но имеют более определенную количественную основу.

Нет необходимости продолжать этот перечень примеров успешного применения числовых методов к разнообразным биологическим организмам и явлениям. Важно то, что методы числовой таксономии приводят к результатам, в значительной степени согласующимся с принятой классификацией, и действительно позволяют выявить естественные таксоны, которые биологи и медики пытаются построить эмпирически или на основе интуиции. Таким образом, это доказывает, что математические методы в высшей степени полезны для решения реальных задач биологической систематики. Каковы же тогда их преимущества по сравнению с существующими методами?

Прежде всего математические количественные методы числовой таксономии обеспечивают более точную основу классификации. С их помощью легче обнаружить возникающие трудности и проблемы, и при исследовании этих проблем остается меньше места для чисто субъективных оценок. Далее, числовая таксономия дает возможность строить классификацию на основе значительно большего числа признаков. Представители традиционной таксономии основывают свои идеи, конечно, на общем впечатлении о сходстве организмов или других объектов, однако число рассматриваемых ими признаков обычно весьма ограничено. Методы числовой таксономии позволяют принимать во внимание значительно большее число признаков: обычно рекомендуемое число не менее 50, а еще лучше — до 100. Кроме того, с помощью электронных вычислительных машин удастся довольно легко выполнять обширные и трудоемкие вычисления. Например, при первоначальном исследовании пчел, о котором Миченер и Сокэл сообщили в 1957 г., выполнение расчетов на настольной счетной машине потребовало около шести недель. В настоящее время быстродействующая вычислительная машина может легко справиться с этой работой за один час.

Если будет принят общий подход, изложенный нами в разд. 7.1, а именно построение таксономической классификации на основе наблюдаемых фенетических признаков как основы для проверки различных гипотез, связанных с филогенезом, эволюцией и т. д., то мы будем иметь определенное основание утверждать, что с по-

мощью числовой таксономии можно достигнуть того, чего мог бы достигнуть каждый специалист в области таксономии при наличии у него достаточных возможностей и времени. Таким образом, этот метод становится просто еще одним инструментом научного исследования. Возможно, что существующие числовые методы еще слишком приближенны и нуждаются в изменении и уточнении, для чего необходимы дальнейшие научные исследования и разработки, но, во всяком случае, математический фундамент для такой работы уже заложен.

Позволит ли применение числовой таксономии пойти дальше тех целей, которые ставят перед собой представители традиционной таксономии, покажет будущее. В разд. 7.1 было подчеркнуто, что любая система классификации имеет реальную научную ценность (помимо удобства номенклатуры и каталогизации) лишь в том случае, если она дает возможность делать прогнозы, которые могут быть проверены. Если бы числовая таксономия позволила создать недостижимую другими способами классификацию, ценность которой в смысле прогнозирования подтвердилась бы на практике (хотя бы в некоторых случаях), то это послужило бы чрезвычайно убедительным аргументом в ее пользу.

Глава 8

ЭКОЛОГИЯ И РОСТ ПОПУЛЯЦИЙ

8.1. ВВЕДЕНИЕ

Одна из основных проблем экологии — это исследование роста и изменения популяции под углом зрения как ее внутренних свойств, так и влияния окружающей среды. Некоторые простые динамические модели роста популяции были приведены в разд. 1.3, а в разд. 2.4 обсуждался элементарный подход к более сложным процессам стохастического типа. В настоящей главе мы более детально рассмотрим ряд экологических задач и остановимся на методах построения и применения стохастических моделей. Важно всеми силами развивать этот подход, поскольку маловероятно, что без него удастся достигнуть надлежащей реалистичности, разве лишь в некоторых весьма ограниченных областях.

В отличие от чисто статических задач классификации, рассмотренных в предыдущей главе, здесь мы сосредоточим наше внимание на динамическом описании популяций, учитывающем влияние самых разнообразных факторов: размножения, гибели, миграций, мутаций, конкуренции и т. д. Надо сразу сказать, что математический аппарат, которым нам придется пользоваться, довольно сложен, и для дальнейшего изучения теории случайных процессов читателю следует обратиться к некоторым книгам, рекомендованным в разд. 2.4.

Изложенный здесь материал преследует двойную цель. Во-первых, крайне важно знать, как строятся вероятностные модели различных биологических процессов, и представлять себе смысл принимаемых допущений. Даже если биолог, имеющий математическую подготовку, не пожелает самостоятельно проанализировать все следствия, вытекающие из той или иной математической модели, ему все же будет значительно легче обсуждать этот вопрос с профессиональным математиком, если он с самого начала сможет разобраться, приемлема ли данная модель с биологической точки зрения хотя бы на первый случай, и будет полностью осведомлен обо всех основных допущениях, принятых в этой модели. После этого программистам и специалистам в области численного анализа останется лишь составить логическую последовательность операций.

Во-вторых, для тех, кто пожелает самостоятельно заняться применением теории случайных процессов, здесь дается введение в существующие методы анализа. Естественно, что для более полного понимания этих методов необходимо более детальное изучение соответствующих руководств (см., например, [9]).

Основное внимание уделяется здесь не техническим деталям, а общему направлению и методике в надежде, что изложенный материал поможет читателю определить направление дальнейших исследований и разработок. Для дополнительного изучения вопросов, связанных с построением стохастических моделей популяции, полезно обратиться к книге Бартлетта [12].

8.2. ПРОЦЕССЫ РАЗМНОЖЕНИЯ И ГИБЕЛИ

Остановимся вначале на простом процессе размножения, о котором говорилось в разд. 2.4. Мы предположили там, что в момент $t = 0$ рассматриваемая популяция содержит ровно a особей; вероятность того, что в момент t в ней было n особей, равна $p_n(t)$, а вероятность размножения какой-либо особи в интервале Δt равна $\beta \Delta t$. Таким образом, вероятность того, что в популяции, содержащей n особей, в интервале Δt появится новая особь, равна $\beta n \Delta t$. Как уже указывалось ранее, одно из важных следствий принятия такого допущения состоит в том, что для любой данной особи интервал времени между ее рождением и размножением будет иметь экспоненциальное распределение с плотностью $f(t) = \beta e^{-\beta t}$, $t \geq 0$. Если окажется, что такое допущение является чрезмерно упрощенным и принять его нельзя, то потребуется значительно более сложная модель (см., например, разд. 8.4).

Хотя самым простым подходом кажется составление системы дифференциально-разностных уравнений для исходных вероятностей $p_n(t)$ и их непосредственное решение, такой метод приводит к недопустимым усложнениям при попытке применить его в случае сколь угодно сложных моделей. В связи с этим во многих случаях предпочтительно пользоваться более общим подходом. Как обычно, будем рассматривать вместо самих вероятностей производящие функции распределения вероятностей

$$P(x, t) = \sum_{n=0}^{\infty} p_n(t) x^n \quad (8.1)$$

или производящие функции моментов

$$M(\theta, t) \equiv P(e^\theta, t). \quad (8.2)$$

В разд. 2.4 мы на основании элементарных вероятностных соображений выражали вероятность $p_n(t + \Delta t)$ того, что в момент $t + \Delta t$ в популяции будет ровно n особей, через соответствующие

вероятности $p_n(t)$, $p_{n-1}(t)$ и т. д. для момента t . Теперь попытаемся выразить $P(x, t + \Delta t)$ через $P(x, t)$, что обычно приводит к дифференциальным уравнениям в частных производных относительно P . Легче всего это выполнить одним методом, более подробно описанным в другой книге автора [19], разд. 7.4). Этот метод позволяет, в частности, получить следующий результат.

Допустим, что случайная величина $X(t)$ описывает размер популяции в момент t , а переходы в интервале Δt обозначаются через $\Delta X(t)$. Пусть имеет место конечное число возможных переходов с вероятностями

$$P\{\Delta X(t) = j | X(t)\} = f_j(X) \Delta t, \quad j \neq 0, \quad (8.3)$$

где j — ненулевое целое число, а f_j — некоторая подходящая неотрицательная функция случайной величины X . Таким образом, вероятность отсутствия переходов в интервале Δt равна

$$P\{\Delta X(t) = 0 | X(t)\} = 1 - \sum_{j \neq 0} f_j(X) \Delta t. \quad (8.4)$$

Как видно из (8.3), мы, кроме того, предполагаем, что вероятности перехода в интервале Δt зависят только от размера популяции в момент t и что информация о более ранних периодах времени не существенна. Процессы, протекание которых в будущем, при известном состоянии системы в настоящее время, не зависит от прошлого, называются *марковскими*. Указанное ограничение не столь серьезно, как могло бы показаться на первый взгляд. Во всяком случае, оно значительно упрощает математические выкладки и, кроме того, позволяет сделать ряд обобщений, обеспечивающих большую реалистичность моделей, о которых идет речь (разд. 8.4).

Теперь можно показать, что определенный выше процесс можно охарактеризовать одним уравнением относительно производящей функции:

$$\frac{\partial P(x, t)}{\partial t} = \sum_{j \neq 0} (x^j - 1) f_j \left(x \frac{\partial}{\partial x} \right) P(x, t), \quad (8.5)$$

где $f_j(\partial/\partial x)$ — оператор, применяемый к $P(x, t)$. Соответствующее выражение для производящей функции моментов имеет вид

$$\frac{\partial M(\theta, t)}{\partial t} = \sum_{j \neq 0} (e^{\theta j} - 1) f_j \left(\frac{\partial}{\partial \theta} \right) M(\theta, t). \quad (8.6)$$

Часто j принимает только значения ± 1 , и функция f_j имеет очень простую форму. В этом случае дифференциальное уравнение в частных производных (8.5) или (8.6) записывается без всякого труда, как только заданы вероятности перехода.

Простой процесс размножения

В случае простого процесса размножения возможен переход только одного вида, а именно $j = 1$. Функция $f_1(X)$ имеет вид βX , так как вероятность рождения новой особи в интервале Δt при наличии ровно n особей равна $\beta n \Delta t$. Подстановка этих выражений в формулу (8.5) дает

$$\frac{\partial P}{\partial t} = \beta x (x - 1) \frac{\partial P}{\partial x} \quad (8.7)$$

при начальном условии

$$P(x, 0) = x^\alpha. \quad (8.8)$$

Линейное дифференциальное уравнение с частными производными (8.7) легко решить стандартными методами. Так, соответствующие характеристические уравнения имеют вид

$$\frac{dt}{1} = \frac{dx}{\beta x(1-x)} = \frac{dP}{0}. \quad (8.9)$$

Два независимых интеграла этой системы имеют вид

$$P = \text{const}, \quad e^{\beta t} (1 - x^{-1}) = \text{const}. \quad (8.10)$$

Таким образом, общее решение можно записать в виде

$$P = \Psi [e^{\beta t} (1 - x^{-1})], \quad (8.11)$$

где Ψ — произвольная функция.

Отыскание вида этой произвольной функции обычно не вызывает особых затруднений, так как он полностью определяется начальными условиями. В данном случае имеем

$$x^\alpha = P(x, 0) = \Psi(1 - x^{-1}),$$

откуда следует, что

$$\Psi(u) = (1 - u)^{-\alpha}. \quad (8.12)$$

Следовательно, искомое решение имеет вид

$$P(x, t) = [1 - e^{\beta t} (1 - x^{-1})]^{-\alpha}, \quad (8.13)$$

а соответствующая производящая функция моментов

$$M(\theta, t) = [1 - e^{\beta t} (1 - e^{-\theta})]^{-\alpha}. \quad (8.14)$$

Эту функцию можно найти из формулы (8.13), подставляя в нее $x = e^\theta$ или непосредственно путем решения дифференциального уравнения в частных производных, соответствующего уравнению (8.7) и выводимого из (8.6).

Вычислив коэффициенты при x^n в формуле (8.13), находим значение $p_n(t)$:

$$p_n(t) = \binom{n-1}{a-1} e^{-n\beta t} (e^{\beta t} - 1)^{n-a}, \quad n \geq a. \quad (8.15)$$

Эта полученная ранее формула (2.8). Математическое ожидание и дисперсию легче всего найти обычным путем по формуле (8.14):

$$m(t) = ae^{\beta t}, \quad (8.16)$$

$$\sigma^2(t) = ae^{\beta t} (e^{\beta t} - 1). \quad (8.17)$$

Полученные выражения совпадают с формулами (2.10) и (2.11).

В этом очень простом примере удалось найти вероятности $p_n(t)$ в явном виде. Для более сложных процессов это может оказаться невозможным или, во всяком случае, потребует крайне сложных выкладок. Проще исследовать свойства этого процесса с помощью моментов или семинвариантов, так как во многих случаях эти величины можно получить, не прибегая к решению основного дифференциального уравнения в частных производных. Так, уравнение для производящей функции моментов, выводимое непосредственно из выражения (8.6), имеет вид

$$\frac{\partial M}{\partial t} = \beta (e^\theta - 1) \frac{\partial M}{\partial \theta}, \quad (8.18)$$

а уравнение для производящей функции семинвариантов $K = \log M -$

$$\frac{\partial K}{\partial t} = \beta (e^\theta - 1) \frac{\partial K}{\partial \theta} \quad (8.19)$$

при начальном условии

$$K(\theta, t) = a\theta. \quad (8.20)$$

Вместо того чтобы решать уравнение (8.19), можно просто приравнять коэффициенты при одинаковых степенях θ в обеих частях равенства, используя обычное разложение

$$K \equiv k_1\theta + k_2\theta^2/2! + \dots \quad (8.21)$$

В данном случае это дает систему легко разрешимых обыкновенных дифференциальных уравнений. Первые два из них имеют вид

$$\frac{dk_1}{dt} = \beta k_1 \quad (8.22)$$

и

$$\frac{dk_2}{dt} = 2\beta k_2 + \beta k_1. \quad (8.23)$$

Интегрирование уравнения (8.22) сразу же дает результат (8.16), а подстановка полученного значения k_1 в формулу (8.23) — урав-

нение, содержащее только k_2 . Интегрируя его, легко получаем выражение (8.17). Этот метод значительно облегчает выкладки и в высшей степени удобен при анализе процессов, в которых вероятности перехода являются линейными функциями размера популяции. К сожалению, при рассмотрении нелинейных функций, встречающихся, например, во многих моделях эпидемий, он неприемлем, поскольку в этом случае первое уравнение содержит не только k_1 , но и k_2 , второе уравнение содержит k_3 и т. д. Таким образом, мы не можем решить первое уравнение, чтобы начать процесс последовательного решения системы уравнений.

Простой процесс размножения и гибели

Рассмотрим теперь более сложный процесс — процесс размножения и гибели, в котором возможны переходы двух видов. Как и ранее, полагаем, что вероятность появления одного потомка у одной особи в интервале Δt равна $\beta \Delta t$, поэтому для всей популяции вероятность увеличения ее численности на единицу равна $\beta n \Delta t$. Кроме того, допустим также, что вероятность гибели одной особи в интервале Δt равна $\mu \Delta t$, а для всей популяции вероятность гибели одной особи составляет $\mu n \Delta t$. Теперь переменная j принимает одно из двух значений -1 и $+1$ при $f_{+1}(X) = \beta X$ и $f_{-1}(X) = \mu X$. Следовательно, основное дифференциальное уравнение в частных производных для производящей функции моментов имеет вид

$$\frac{\partial M}{\partial t} = [\beta (e^\theta - 1) + \mu (e^{-\theta} - 1)] \frac{\partial M}{\partial \theta} \quad (8.24)$$

при начальном условии

$$M(\theta, 0) = e^{a\theta}. \quad (8.25)$$

Это также стандартное линейное уравнение, и его можно решить непосредственно. Решение имеет вид

$$M(\theta, t) = \left(\frac{\mu v - 1}{\beta v - 1} \right)^a, \quad (8.26)$$

где

$$v \equiv \frac{(e^\theta - 1) e^{(\beta - \mu)t}}{\beta e^\theta - \mu}. \quad (8.27)$$

Из формулы (8.26) вытекают многие важные свойства процесса.

Выражения для вероятностей $p_n(t)$ в этом случае также можно найти в явном виде, хотя они и будут несколько сложнее. Положив в формуле (8.26) $e^\theta = x$, получаем производящую функцию

вероятностей, правую часть которой можно разложить в ряд по степеням x . Искомый коэффициент при x^n имеет вид

$$\left. \begin{aligned}
 p_n(t) &= \sum_{j=0}^{\min(a, n)} \binom{a}{j} \binom{a+n-j-1}{a-1} g^{a-j} h^{n-j} (1-g-h)^j, \\
 p_0(t) &= g^a, \\
 \text{где} \\
 g &= \frac{\mu [e^{(\beta-\mu)t} - 1]}{\beta e^{(\beta-\mu)t} - \mu}, \quad h = \frac{\beta g}{\mu}.
 \end{aligned} \right\} \quad (8.28)$$

Таким образом, даже в случае простейшего стохастического процесса размножения и гибели общее выражение для $p_n(t)$ оказывается довольно сложным. Это говорит о том, что для более сложных процессов выразить $p_n(t)$ в явном виде скорее всего не удастся.

Математическое ожидание и дисперсию можно легко найти с помощью формулы (8.26), вычислив коэффициенты при θ и θ^2 в выражении для $K = \ln M$. Кроме того, можно вообще избежать решения дифференциального уравнения в частных производных (8.24), перейдя к записи $M = e^K$. Получаем

$$\frac{\partial K}{\partial t} = [\beta (e^\theta - 1) + \mu (e^{-\theta} - 1)] \frac{\partial K}{\partial \theta}. \quad (8.29)$$

Как и в случае простого процесса размножения, приравниваем коэффициенты при θ и θ^2 в обеих частях уравнения и получаем обыкновенные дифференциальные уравнения

$$\left. \begin{aligned}
 \frac{dk_1}{dt} &= (\beta - \mu) k_1, \\
 \frac{dk_2}{dt} &= 2(\beta - \mu) k_2 + (\beta + \mu) k_1,
 \end{aligned} \right\} \quad (8.30)$$

легко решаемые одно за другим. Решение имеет вид

$$\left. \begin{aligned}
 m(t) &= ae^{(\beta-\mu)t}, \\
 \sigma^2(t) &= \frac{a(\beta+\mu)}{\beta-\mu} e^{(\beta-\mu)t} [e^{(\beta-\mu)t} - 1].
 \end{aligned} \right\} \quad (8.31)$$

Сразу же заметим, что, как и в случае простого процесса размножения, математическое ожидание численности совпадает со значением численности, получаемым в детерминистской модели, хотя выражение для дисперсии свидетельствует о том, что имеет место значительная изменчивость. Чтобы получить дополнительную информацию о характере изменчивости, рассмотрим частный

случай, когда размножение и гибель уравновешивают друг друга, т. е. когда $\mu = \beta$. Математическое ожидание и дисперсию находим с помощью формул (8.31), просто полагая, что $\mu \rightarrow \beta$, и используя во втором выражении правило Лопиталья для раскрытия неопределенности вида $\frac{0}{0}$. Получаем

$$\left. \begin{aligned} m(t) &= a, \\ \sigma^2(t) &= 2a\lambda t. \end{aligned} \right\} \quad (8.32)$$

Первое выражение представляет собой ожидаемый результат, а именно средний размер популяции сохраняет свое начальное значение. Второе выражение показывает, что дисперсия размера популяции возрастает пропорционально длительности интервала времени, в течение которого протекает процесс. Ясно, что должны иметь место распределения с большой асимметрией, так как число особей в популяции не может быть отрицательным.

Вероятность вымирания популяции

Еще одной важной особенностью таких процессов является *вымирание*. Как уже указывалось в разд. 2.4, детерминистская модель в тех случаях, когда скорость размножения превышает скорость гибели, предсказывает устойчивое экспоненциальное увеличение размера популяции. Однако в стохастической модели учитывается, что всегда существует определенная вероятность такого большого числа случаев гибели, при котором популяция полностью вымирает. Таким образом, вероятность вымирания служит важной характеристикой одного из аспектов стохастической модели. Обозначим через $p_0(t)$ вероятность того, что в момент времени t не останется ни одной живой особи. Из уравнений (8.28) можно найти выражения для этой вероятности в явном виде:

$$p_0(t) = \left\{ \frac{\mu [e^{(\beta-\mu)t} - 1]}{\beta e^{(\beta-\mu)t} - \mu} \right\}^a, \quad \beta \neq \mu. \quad (8.33)$$

В частном случае, когда $\mu = \beta$, это выражение принимает вид

$$p_0(t) = \left(\frac{\beta t}{\beta t + 1} \right)^a, \quad \beta = \mu. \quad (8.34)$$

Вероятность того, что вымирание популяции произойдет *рано* или *поздно*, можно найти, полагая $t \rightarrow \infty$. В пределе при $t \rightarrow \infty$ выражения (8.33) и (8.34) для случаев $\beta < \mu$, $\beta = \mu$ и $\beta > \mu$ можно записать следующим образом:

$$\left. \begin{aligned} \lim_{t \rightarrow \infty} p_0(t) &= 1, & \beta &\leq \mu; \\ \lim_{t \rightarrow \infty} p_0(t) &= (\mu/\beta)^a, & \beta &> \mu. \end{aligned} \right\} \quad (8.35)$$

Следовательно, если скорость размножения не превышает скорости гибели, вымирание рано или поздно *обязательно* произойдет. Если же скорость размножения выше скорости гибели, то вероятность вымирания популяции составляет $(\mu/\beta)^a$.

Интересно, что даже в том случае, когда $\beta = \mu$ и математическое ожидание численности имеет постоянную величину, вероятность полного вымирания все же равна единице. На самом деле происходит следующее: несколько популяций увеличиваются до очень больших размеров, тогда как большинство популяций вымирает, и в результате сохраняется некоторое постоянное среднее. Этот результат, полученный при исследовании самого простого процесса размножения и гибели, ясно показывает, что нельзя сосредоточивать внимание только на математических ожиданиях, даже если они совпадают с соответствующими детерминистскими средними. Новые свойства, приобретаемые при введении в модель вероятностных элементов, должны всегда исследоваться в высшей степени тщательно.

В модель простого процесса размножения и гибели можно ввести ряд более общих предположений, приближающих ее к реальным биологическим явлениям. Например, можно принять, что скорости размножения и гибели не постоянны, а являются функциями времени (возможно, периодическими). В этом случае нужно заменить β и μ на $\beta(t)$ и $\mu(t)$. Из рассмотрения описанного выше метода, используемого для получения основных дифференциальных уравнений в частных производных для производящей функции вероятностей или производящей функции моментов, следует, что общий вид этих уравнений останется без изменений. Поэтому уравнения (8.24) и (8.29) можно использовать в том виде, как они записаны, имея в виду, что теперь β и μ являются функциями времени t . К счастью, эти уравнения по-прежнему разрешимы в замкнутом виде, и из них можно получить выражения для отдельных вероятностей, вероятности вымирания популяции, математического ожидания и дисперсии размера популяции и т. д. Мы не будем останавливаться здесь на обсуждении этого вопроса; более детально он изложен в разд. 9.5 другой книги автора [9]. Главное состоит в том, что изменение модели с целью охватить случаи, когда скорости размножения и гибели зависят от времени, вполне возможно и осуществимо. А именно это и требуется при решении многих экологических задач, поскольку скорости размножения и гибели в популяции частично регулируются факторами окружающей среды, которые определенным образом изменяются с течением времени.

Влияние миграции

Еще одна особенность, которая должна быть учтена в модели, чтобы она была более реалистичной, заключается в том, что очень многие биологические популяции в той или иной степени подвержены миграции. Что касается эмиграции, то она влияет на размер популяции так же, как и гибель особей. Таким образом, ее можно учесть путем соответствующей коррекции скорости гибели, поскольку разумно предположить, что вероятность эмиграции одной особи в интервале Δt , как и вероятность гибели одной особи, пропорциональна размеру популяции.

Но в случае иммиграции дело обстоит несколько иначе, поскольку простейшим вероятностным допущением является предположение о том, что новые особи появляются в популяции случайным образом с частотой ν . Таким образом, вероятность того, что в интервале Δt в популяции появится новая особь за счет иммиграции, равна $\nu \Delta t$. Следовательно, суммарная вероятность появления новой особи в интервале Δt за счет размножения и иммиграции равна $\beta n \Delta t + \nu \Delta t$. В принятых нами обозначениях j по-прежнему принимает только значения $+1$ и -1 , однако теперь $f_1 = \beta X + \nu$, тогда как $f_{-1} = \mu X$, как и ранее. Сразу же находим, что дифференциальное уравнение в частных производных для производящей функции моментов имеет вид

$$\frac{\partial M}{\partial t} = [\beta(e^\theta - 1) + \nu(e^{-\theta} - 1)] \frac{\partial M}{\partial \theta} + \nu(e^\theta - 1) M \quad (8.36)$$

при начальном условии

$$M(\theta, 0) = e^{a\theta}. \quad (8.37)$$

Уравнение (8.36) отличается от уравнения (8.24) лишь наличием в правой части дополнительного члена, который описывает пуассоновский процесс. Решение этого линейного дифференциального уравнения также можно получить без большого труда стандартным методом, и оно имеет вид

$$M(\theta, t) = \left. \begin{aligned} & \frac{(\beta - \mu)^{\nu/\beta} [\mu(T-1) - (\mu T - \beta)e^\theta]^a}{[(\beta T - \mu) - \beta(T-1)e^\theta]^{a + \nu/\beta}}, \\ & T = e^{(\beta - \mu)t}. \end{aligned} \right\} \quad (8.38)$$

Рассматривая ряд частных случаев, можно получить некоторое представление о характере влияния иммиграции на простой процесс размножения и гибели. Прежде всего рассмотрим случай $a = 0$. Первоначально размер такой популяции равен нулю, но рано или поздно с появлением иммигрировавшей особи популяция начнет расти. По той же причине не произойдет и полного

вымирания популяции, хотя время от времени ее численность может падать до нуля. Подставим в формулу (8.38) $a = 0$ и $e^{\theta} = x$ и образуем производящую функцию вероятностей

$$P(x, t) = \left(\frac{\beta - \mu}{\beta T - \mu} \right)^{v/\beta} \left[1 - \frac{\beta(T-1)}{\beta T - \mu} x \right]^{-v/\beta}. \quad (8.39)$$

Таким образом, размер популяции имеет отрицательное биномиальное распределение с математическим ожиданием

$$m(t) = \frac{v}{\beta - \mu} [e^{(\beta - \mu)t} - 1]. \quad (8.40)$$

При $\beta > \mu$ средний размер популяции с увеличением t экспоненциально возрастает с интенсивностью $\beta - \mu$, причем масштаб зависит от интенсивности иммиграции v . При $\beta < \mu$ предельный размер популяции при больших t равен $v/(\mu - \beta)$. Если же размножение и гибель в точности уравновешены, то $m(t) = vt$ для всех t и математическое ожидание прямо пропорционально времени. Объединяя эти результаты, можно записать

$$\left. \begin{aligned} m(t) &\sim \frac{v}{\beta - \mu} e^{(\beta - \mu)t}, & \beta > \mu, \\ m(t) &= vt, & \beta = \mu, \\ m(t) &\sim \frac{v}{\mu - \beta}, & \beta < \mu. \end{aligned} \right\} \quad (8.41)$$

В частном случае, когда $\beta < \mu$, не только математическое ожидание приближается к некоторому фиксированному значению, но и все распределение имеет предельную форму. При $t \rightarrow \infty$ имеем $e^{(\beta - \mu)t} \rightarrow 0$ и $T \rightarrow 0$. Следовательно, формула (8.39) принимает вид

$$P(x, \infty) = \left(\frac{\mu - \beta x}{\mu - \beta} \right)^{-v/\beta}, \quad \beta < \mu. \quad (8.42)$$

Нас может интересовать также и тот случай, когда размножение отсутствует, т. е. $\beta = 0$. Такая модель применима, в частности, к системе, состоящей из частиц, находящихся в некотором небольшом объеме. Чандрасекхар использовал ее для изучения коллоидных частиц в суспензии, а Ротшильд — для исследования движения сперматозоидов. Предполагается, что иммиграция из окружающей среды представляет собой случайный пуассоновский процесс с параметром v , а эмиграция или гибель в исследуемой области пропорциональна числу находящихся в ней частиц. Из формулы (8.39) нетрудно вывести соответствующее распределение вероятностей, полагая $\beta \rightarrow 0$ (хотя в некоторых отношениях

было бы несколько проще сделать это с самого начала). Предельное распределение имеет вид

$$\lim_{\beta \rightarrow 0} P(x, t) = \exp[(v/\mu)(1 - e^{-\mu t})(x - 1)]. \quad (8.43)$$

Это пуассоновское распределение с параметром $(v/\mu)(1 - e^{-\mu t})$, который при больших t стремится к v/μ .

Распределение другого типа можно получить, рассматривая случай, когда значение v очень мало, т. е. когда $v \rightarrow 0$, а β и μ остаются постоянными. Это означает, что общий эффект иммиграции пренебрежимо мал, но вполне достаточен для возобновления роста популяции, если произойдет вымирание. Примем в формуле (8.39) $v/\beta = \gamma$ и перепишем ее теперь в виде

$$P(x, t) = A^\gamma (1 - Bx)^{-\gamma}, \quad (8.44)$$

где

$$A = \frac{\beta - \mu}{\beta T - \mu}, \quad B = \frac{\beta(T-1)}{\beta T - \mu}. \quad (8.45)$$

Полагая в формуле (8.44) $\gamma \rightarrow 0$, получаем вырожденное распределение, когда все распределение вероятности сосредоточено в нулевом значении. Но во многих биологических системах нулевой класс вообще не наблюдается или неотличим от других несущественных данных. Число ночных бабочек данного вида, попавших в световую ловушку, может иметь смысл лишь в том случае, если будет поймана хотя бы одна; если в ловушке не будет ни одной бабочки, то это может лишь свидетельствовать о том, что по каким-то особым причинам ни одна из них не пролетала вблизи. Поэтому нас больше интересует распределение, в котором нулевая численность невозможна. Исключая первый член формулы (8.44) и деля на γA^γ , получаем

$$Bx + \frac{\gamma+1}{2!} (Bx)^2 + \frac{(\gamma+1)(\gamma+2)}{3!} (Bx)^3 + \dots$$

Это выражение равно $-\log(1 - Bx)$. Нормируя, получаем логарифмическое распределение, для которого производящая функция вероятностей имеет вид

$$P(x, t) = \frac{\log(1 - Bx)}{\log(1 - B)}. \quad (8.46)$$

Таким образом, мы видим, что на практике при изучении стохастических явлений появляется множество различных распределений: пуассоновское, логарифмическое, отрицательное биномиальное и т. д. К сожалению, каждое из этих распределений может быть получено при самых различных допущениях, поэтому появление какого-либо конкретного распределения мало что говорит

о механизме, его порождающем. Как обычно, математика и статистика не заменяют здесь настоящих научных исследований, а служат лишь в высшей степени полезным аппаратом, помогающим интерпретировать наблюдаемые явления.

8.3. МНОГОМЕРНЫЕ ПРОЦЕССЫ

Лишь в простейших случаях допустимо считать, что популяция состоит из особей только одного типа, обычно же изучают две или большее число групп одновременно. Например, можно рассматривать популяцию, состоящую из особей двух полов, или колонию бактерий, содержащую как особей дикого типа, так и мутантов. Кроме того, можно рассматривать одну популяцию, но изучать одновременно как общее число живых особей, так и общее число когда-либо живших особей, т. е. суммарную численность популяции. Примером может служить различие между подсчетом жизнеспособных бактерий методом выращивания свежих культур из проб и подсчетом общего количества бактерий методом измерения оптической плотности.

Не представляет труда обобщить результаты, полученные в начале предыдущего раздела, на случай двух или большего числа различных типов особей. Допустим, что имеются только два типа особей, общее число которых в момент времени t выражается случайными величинами $X(t)$ и $Y(t)$. Тогда вместо формулы (8.3) мы получим следующее выражение для совместной вероятности одновременных переходов в интервале Δt :

$$P[\Delta X(t) = j, \Delta Y(t) = k | X(t), Y(t)] = f_{jk}(X, Y) \Delta t, \quad (8.47)$$

j и k не равны нулю одновременно.

Использовавшийся ранее метод на этот раз приводит к следующим довольно очевидным обобщениям формул (8.5) и (8.6):

$$\frac{\partial P(x, y, t)}{\partial t} = \sum_{\substack{j \text{ и } k \text{ не} \\ \text{равны нулю} \\ \text{одновременно}}} (x^j y^k - 1) f_{jk} \left(x \frac{\partial}{\partial x}, y \frac{\partial}{\partial y} \right) P(x, y, t) \quad (8.48)$$

и

$$\frac{\partial M(\theta, \varphi, t)}{\partial t} = \sum_{\substack{j \text{ и } k \text{ не} \\ \text{равны нулю} \\ \text{одновременно}}} (e^{j\theta + k\varphi} - 1) f_{jk} \left(\frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi} \right) M(\theta, \varphi, t). \quad (8.49)$$

Более детальное изложение этой теории и более полное рассмотрение приложений дается в гл. 10 другой книги автора [9]. В данном разделе мы весьма кратко рассмотрим несколько примеров, чтобы охарактеризовать область применения моделей такого типа.

Популяции, состоящие из особей двух полов

В качестве непосредственного обобщения простого процесса размножения и гибели предположим, что в популяции имеются особи двух полов — женского и мужского, — число которых описывают случайные величины $X(t)$ и $Y(t)$. Скорость размножения (разумеется, только для самок) равна β , а скорость гибели самок и самцов равна соответственно μ и μ' . Допустим далее, что отношение числа самок к числу самцов при рождении равно p/q , где $q = 1 - p$. Таким образом, вероятность того, что в интервале Δt от данной самки родится особь женского или мужского пола, равна соответственно $\beta p \Delta t$ и $\beta q \Delta t$.

Следовательно, вероятность того, что вся популяция в интервале Δt произведет новую особь женского или мужского пола, равна соответственно $\beta p X \Delta t$ и $\beta q X \Delta t$. Соответствующие вероятности гибели одной особи равны $\mu X \Delta t$ и $\mu' Y \Delta t$. В данном случае функция f принимает только четыре значения, отличных от нуля:

$$f_{1;0} = \beta p X, \quad f_{-1;0} = \mu X, \quad f_{0;1} = \beta q X, \quad f_{0;-1} = \mu' Y, \quad (8.50)$$

и снова можно сразу же записать соответствующие дифференциальные уравнения в частных производных. Уравнение для производящей функции моментов находится из уравнения (8.49) и имеет вид

$$\begin{aligned} \frac{\partial M}{\partial t} = & [(e^\theta - 1) \beta p + (e^{-\theta} - 1) \mu + (e^\varphi - 1) \beta q] \frac{\partial M}{\partial \theta} + \\ & + (e^{-\varphi} - 1) \mu' \frac{\partial M}{\partial \varphi} \end{aligned} \quad (8.51)$$

при начальном условии

$$M(\theta, \varphi, 0) = e^{a\theta + b\varphi}, \quad (8.52)$$

если в начальный момент $t = 0$ число самок равно a , а число самцов b .

Хотя найти решение уравнения (8.51) в явном виде довольно трудно, можно быстро получить информацию о математических ожиданиях и дисперсиях тем же способом, что и ранее, полагая $M = e^K$, находя соответствующее уравнение для производящей функции семиинвариантов и приравнивая затем коэффициенты при θ , φ , θ^2 , $\theta\varphi$, φ^2 и т. д. Это дает систему обыкновенных дифференциальных уравнений, в которых первые пять уравнений содержат семиинварианты двумерных распределений до второго

порядка включительно:

$$\left. \begin{aligned} \frac{dk_{10}}{dt} &= (\beta p - \mu) k_{10}, & \frac{dk_{01}}{dt} &= \beta q k_{10} - \mu' k_{01}, \\ \frac{dk_{20}}{dt} &= (\beta p + \mu) k_{10} + 2(\beta p - \mu) k_{20}, \\ \frac{dk_{11}}{dt} &= (\beta p - \mu - \mu') k_{11} + \beta q k_{20}, \\ \frac{dk_{02}}{dt} &= \beta q k_{10} + \mu' k_{01} + 2\beta q k_{11} - 2\mu' k_{02}. \end{aligned} \right\} (8.53)$$

Заметим, что первые два уравнения для математического ожидания тождественны соответствующим уравнениям для детерминированного случая. Все пять уравнений разрешимы, и их можно решить последовательно в определенном порядке; математические ожидания имеют вид

$$k_{10} = ae^{(\beta p - \mu)t}, \quad (8.54)$$

$$k_{01} = be^{-\mu't} + \frac{\beta qa}{\beta p - \mu + \mu'} [e^{(\beta p - \mu)t} - e^{-\mu't}]. \quad (8.55)$$

Выражения для дисперсий и коэффициентов корреляции получить совсем нетрудно, но они очень громоздки.

Мы не предполагаем углубляться здесь в этот вопрос. Более детально он рассматривается в разд. 4.2 книги Баруча-Рида [13], где приведена также библиография. Наша основная цель состоит в том, чтобы показать, каким образом можно построить стохастическую модель для двух случайных величин и каким образом приступить к ее анализу. В принципе данный метод можно распространить на любое число случайных величин, однако ясно, что при этом математическое исследование будет быстро усложняться.

Суммарная численность популяции

Как уже упоминалось ранее, нас иногда интересует не только число особей некоторого сообщества, фактически живущих в какой-либо момент времени, но и общее число когда-либо живших особей. Решить эту задачу значительно легче, чем это может показаться с первого взгляда. Рассмотрим, к примеру, простой процесс размножения и гибели, описанный в разд. 8.2, со скоростями размножения и гибели, равными соответственно β и μ . Для описания числа особей, живущих в популяции в момент времени t , используем случайную величину $X(t)$, а для описания суммарной численности популяции — случайную величину $Y(t)$. Тогда случайная величина $Y(t)$ будет повторять положительные скачки случайной величины $X(t)$, но не будет повторять ее отрицательные скачки, так как она изменяется только при появлении

в популяции новых членов. Таким образом, единственными ненулевыми значениями функции f являются

$$f_{1;1} = \beta X, \quad f_{-1;0} = \mu X. \quad (8.56)$$

Легко видеть, что основное дифференциальное уравнение в частных производных для производящей функции моментов имеет вид

$$\frac{\partial M}{\partial t} = [\beta (e^{\theta+\varphi} - 1) + \mu (e^{-\theta} - 1)] \frac{\partial M}{\partial \theta} \quad (8.57)$$

при начальном условии

$$M(\theta, \varphi, 0) = e^{a\theta + a\varphi} \quad (8.58)$$

(так как $X(0) = Y(0) = a$).

Уравнение (8.57) по-прежнему справедливо, даже если β и μ являются функциями t . Поэтому вполне применим обычный метод, основанный на выводе уравнения для производящей функции семиинвариантов и обыкновенных дифференциальных уравнений для семиинвариантов, причем для определенных функций $\beta(t)$ и $\mu(t)$ этот метод может оказаться очень полезным. Однако получить решение в явном виде для общего случая пока невозможно.

Когда β и μ — постоянные, разрешить обыкновенные дифференциальные уравнения довольно легко. Что касается математического ожидания и дисперсии числа живых особей популяции, то должны вновь получиться значения, заданные формулами (8.33) и (8.34), так как мы не предполагали изменения законов развития этой группы особей. Однако математическое ожидание и дисперсия для популяции выражаются другими формулами:

$$k_{01} = \frac{a}{\mu - \beta} [\mu - \beta e^{(\beta - \mu)t}] \quad (8.59)$$

и

$$\begin{aligned} k_{02} = & \frac{a\beta(\mu + \beta)}{(\mu - \beta)^2} [1 - e^{(\beta - \mu)t}] - \frac{4a\beta^2\mu t e^{(\beta - \mu)t}}{(\mu - \beta)^2} + \\ & + \frac{a\beta^2(\mu + \beta)}{(\mu - \beta)^3} [1 - e^{2(\beta - \mu)t}]. \end{aligned} \quad (8.60)$$

Если $\beta < \mu$, то предельные значения при $t \rightarrow \infty$ имеют вид

$$k_{01} \rightarrow \frac{a\mu}{\mu - \beta}, \quad k_{02} \rightarrow \frac{a\mu\beta(\mu + \beta)}{(\mu - \beta)^3}, \quad (8.61)$$

так что среднее число особей, когда-либо существовавших в популяции (при $\beta < \mu$ определенно произойдет вымирание), равно $a\mu/(\mu - \beta)$; сюда входят a особей, имевшихся в начальный момент $t = 0$.

Далее можно найти в явном виде выражение для производящей функции вероятностей. Полагая в формулах (8.57) и (8.58) $e^{\theta} = x$ и $e^{\varphi} = y$, находим, что соответствующее уравнение для $P(x, y, t)$ имеет вид

$$\frac{\partial P}{\partial t} = [\beta(x^2y - x) + \mu(1 - x)] \frac{\partial P}{\partial x} \quad (8.62)$$

при начальном условии

$$P(x, y, 0) = x^a y^a. \quad (8.63)$$

(Уравнение (8.62) можно было также записать непосредственно, используя общую формулу (8.48).) Решая это уравнение обычными методами, получаем

$$P(x, y, t) = y^a \left\{ \frac{\xi(\eta - x) + \eta(x - \xi)e^{-\beta y(\eta - \xi)t}}{(\eta - x) + (x - \xi)e^{-\beta y(\eta - \xi)t}} \right\}^a, \quad (8.64)$$

где ξ и η — функции переменной y , определяемой корнями следующего квадратного уравнения:

$$\beta y x^2 - (\beta + \mu)x + \mu = 0, \quad 0 < \xi < 1 < \eta. \quad (8.65)$$

(Это квадратное уравнение образуется автоматически при решении одного из характеристических уравнений.)

Выражение (8.64) достаточно компактно, хотя и сложно, и из него выводятся некоторые интересные свойства. Например, можно получить в явном виде асимптотическое распределение суммарной численности популяции, для чего нужно просто вычислить $P(1, y, \infty)$. Имеем

$$P(1, y, \infty) = (y\xi)^a, \quad (8.66)$$

где

$$\xi = \frac{(\beta + \mu) - [(\beta + \mu)^2 - 4\beta\mu y]^{1/2}}{2\beta y}. \quad (8.67)$$

При $a = 1$ функцию (8.66) легко разложить в ряд по степеням y . При $a > 1$ удобно использовать разложение в ряд Лагранжа (см. другую книгу автора [9], разд. 10.3). Выполнив все преобразования, находим, что асимптотическое распределение имеет вид

$$q_r = \frac{a(2r - a - 1)! \beta^{r-a} \mu^r}{(r - a)! r! (\beta + \mu)^{2r-a}}, \quad r \geq a, \quad (8.68)$$

где q_r — вероятность того, что общее число особей, когда-либо живших в популяции, равно r .

Мутации у бактерий

В предыдущих примерах мы рассматривали ситуации, в которых требовалось учитывать существование двух групп и для каждой из них строить стохастическую модель. В некоторых случаях построение таких моделей сопряжено с чрезвычайно большими трудностями, однако можно принять удобное упрощающее допущение (когда оно обосновано) о детерминистском поведении хотя бы одной группы. Типичным примером служит рост колонии бактерий, в которой протекает случайный мутационный процесс. В такой колонии существуют два типа особей — особи дикого типа и мутанты. Численность первых обычно очень велика, и поэтому вполне допустимо рассматривать ее детерминистскими методами и считать, что она равна общему числу бактерий в популяции. Мутанты появляются очень редко, и их число обычно невелико. Поэтому для них стохастический анализ существенно необходим. Таким путем двумерная стохастическая модель превращается в одномерную, что значительно упрощает исследование.

Допустим, что скорости размножения особей дикого типа и мутантов одинаковы и что гибель отсутствует. Если число бактерий в начальный момент $t = 0$ равно a , то общее число особей в момент t равно $ae^{\beta t}$. Поскольку обычно число мутантов мало по сравнению с общим числом бактерий, то $ae^{\beta t}$ можно рассматривать приближенно как число особей дикого типа в момент времени t . Допустим, как и прежде, что вероятность деления одного мутанта в интервале Δt равна $\beta \Delta t$. Если число мутантов в момент t равно n , то вероятность появления нового мутанта за счет деления особей этой группы равна $\beta n \Delta t$. Однако число мутантов может увеличиваться также за счет мутаций в популяции дикого типа. Допустим, что частота мутаций очень мала и равна p ; это означает, что при делении особи дикого типа существует вероятность p того, что вместо двух особей дикого типа образуется одна особь дикого типа и одна мутантная. Среднее число новых особей, появившихся в группе особей дикого типа в интервале Δt , равно $\beta ae^{\beta t} \Delta t$, а вероятность появления одного мутанта — $p \beta ae^{\beta t} \Delta t$. Таким образом, суммарная вероятность перехода для одного нового мутанта равна $\beta n \Delta t + p \beta ae^{\beta t} \Delta t$.

Это единственный тип перехода, для которого $j = +1$ и $f_1 = \beta X + p \beta ae^{\beta t}$. Следовательно, дифференциальное уравнение в частных производных для производящей функции моментов имеет вид

$$\frac{\partial M}{\partial t} = \beta \gamma ae^{\beta t} (e^{\theta} - 1) M + \beta (e^{\theta} - 1) \frac{\partial M}{\partial \theta} \quad (8.69)$$

при начальном условии

$$M(\theta, 0) = 1, \quad (8.70)$$

так как в начальный момент $t = 0$ численность популяции мутантов равна нулю.

Уравнение для семиинвариантов

$$\frac{\partial K}{\partial t} = \beta p a e^{\beta t} (e^\theta - 1) + \beta (e^\theta - 1) \frac{\partial K}{\partial \theta} \quad (8.71)$$

при

$$K(\theta, 0) = 0, \quad (8.72)$$

как и предполагалось, решить несколько легче. Решение находится обычным методом (хотя необходима особая тщательность при вычислении интегралов; см. другую книгу автора [9], разд. 10.4) и имеет вид

$$K(\theta, t) = p a e^{\beta t} (e^{-\theta} - 1) \log [1 + e^\theta (e^{-\beta t} - 1)]. \quad (8.73)$$

Не представляет труда найти первые два семиинварианта:

$$m(t) = \beta p a t e^{\beta t}, \quad (8.74)$$

$$\sigma^2(t) = 2 p a e^{\beta t} (e^{\beta t} - 1) - \beta p a t e^{\beta t}. \quad (8.75)$$

Если положить $N = a e^{\beta t}$, где N — общее число бактерий во всей колонии, то для больших t имеем

$$\left. \begin{aligned} m &= p N \log(N/a), \\ \sigma^2 &\approx 2 p N^2/a. \end{aligned} \right\} \quad (8.76)$$

Хотя этот результат представляет общий интерес, на практике можно пойти еще дальше, так как уравнение (8.73) определяет производящую функцию вероятностей, которую удобно записать в виде

$$P(x, t) = \exp \left\{ p a e^{\beta t} \left(\frac{1-x}{x} \right) \log [1 + (e^{-\beta t} - 1) x] \right\}. \quad (8.77)$$

Это выражение можно разложить в ряд и получить отдельные вероятности, хотя и в довольно сложной форме. Легко найти вероятность того, что новые мутанты появляться не будут:

$$p_0(t) = \exp[-p a (e^{\beta t} - 1)]. \quad (8.78)$$

Асимптотическую форму уравнения (8.77) при $t \rightarrow \infty$ можно записать как

$$P_N(x) = \exp \left[\frac{p N (1-x)}{x} \log(1-x) \right]. \quad (8.79)$$

Это несколько необычное распределение вероятностей, у которого отсутствуют конечные моменты. Ли и Кулсон показали, что, нормируя распределение (8.79), можно получить приближенно нормальное распределение с нулевым математическим ожиданием и единичным средним квадратическим отклонением, имеющим вид

$$w = \frac{11,6}{(r/M) - \log M + 4,5} - 2,02, \quad M \geq 4, \quad (8.80)$$

где $M = pN$, а r — наблюдаемое число мутантов в колонии. Таким образом, мы имеем практический метод, позволяющий оценить частоту мутаций p на основании изучения параллельных культур.

Многофазовый процесс размножения

Во всех стохастических процессах, рассмотренных выше, мы предполагали, что переходы различного рода происходят случайным образом. Хотя для таких явлений, как мутационный процесс, это предположение довольно близко к истине, на процессы размножения и гибели оно накладывает довольно сильное ограничение, поскольку это означает, что промежутки времени от рождения до размножения или гибели особей распределены по экспоненциальному закону. Для общего исследования эффекта включения вероятностных элементов в процесс, который в противном случае должен был бы изучаться детерминистскими методами, такие допущения приемлемы. Однако для более точного описания биологического механизма роста популяции необходимо нечто более близкое к реальной действительности. Например, непосредственное наблюдение бактерий *Bacterium aerogenes* показало, что значения продолжительности жизни лучше описываются распределением χ^2 с числом степеней свободы 25—100.

Большое преимущество предположения о случайном характере переходов состоит в том, что это позволяет нам оперировать с марковскими процессами, для которых легко получить основные дифференциальные уравнения в частных производных, и во многих случаях эти уравнения решаются сравнительно просто. Как только мы отходим от экспоненциального распределения значений продолжительности жизни (или какого-либо другого соответствующего интервала), процессы становятся немарковскими, и при этом математический анализ значительно усложняется. Способ обойти эту трудность был предложен Кендаллом. (Сводка некоторых результатов, полученных Кендаллом, приводится в другой книге автора [9], разд. 10.5.) Рассмотрим процесс размножения, в котором каждая особь на протяжении всей своей жизни последовательно проходит через k фаз, начиная с первой фазы и кончая

k -й. Применительно к бактериям мы предполагаем, что, когда бактерия покидает последнюю фазу, она заменяется двумя новыми особями, находящимися в первой фазе.

Для удобства по-прежнему будем считать, что переходы из одной фазы в другую происходят случайным образом, но с частотой $k\beta$, так что вероятность перехода в интервале Δt какой-либо особи в следующую фазу равна $k\beta\Delta t$. Это означает, что длительность пребывания в какой-либо фазе имеет экспоненциальное распределение

$$f(u) = k\beta e^{-k\beta u}, \quad 0 \leq u < \infty. \quad (8.81)$$

По существу это распределение χ^2 с двумя степенями свободы (если не учитывать масштабный множитель), поэтому длительность прохождения всех k фаз пропорциональна случайной величине, распределенной по закону χ^2 с $2k$ степенями свободы.

Математическое ожидание случайной величины, заданной выражением (8.81), равно $(k\beta)^{-1}$, а средняя длительность прохождения k фаз равна β^{-1} и не зависит от k . Изменяя значение k , можно изменять форму распределения, сохраняя среднее значение неизменным. Конечно, разбиение всей продолжительности жизни на k фаз — это просто искусственный математический прием (обычно они не имеют какого-либо биологического смысла), предназначенный для получения более приемлемого общего распределения. Но поскольку предположение о случайном характере переходов при этом сохраняется, мы по-прежнему будем иметь марковский процесс, если только все фазы рассматриваются совместно, т. е. если мы рассматриваем соответствующий k -мерный процесс.

Пусть число особей, находящихся в j -й фазе, является случайной величиной $X_j(t)$. Вероятность того, что в интервале Δt произойдет переход, в результате которого одна из этих особей скачком перейдет в $(j + 1)$ -ю фазу, равна $k\beta X_j \Delta t$. Если переход происходит, то при $j < k$ величина X_{j+1} увеличивается на единицу, а X_j уменьшается на единицу. Если $j = k$, то при переходе X_k теряет одну единицу, а X_1 приобретает две единицы. Очевидное k -мерное обобщение уравнения (8.49) имеет вид

$$\frac{\partial M}{\partial t} = k\beta \sum_{j=1}^{k-1} (e^{-\theta_j + \theta_{j+1}} - 1) \frac{\partial M}{\partial \theta_j} + k\beta (e^{-\theta_k + 2\theta_1}) \frac{\partial M}{\partial \theta_k} \quad (8.82)$$

при начальном условии

$$M(\theta_1, \theta_2, \dots, \theta_k; 0) = e^{\theta_1}, \quad (8.83)$$

если в начальный момент $t = 0$ в первой фазе находилась одна особь.

Кое-что может дать решение соответствующего линейного уравнения для производящей функции вероятностей, но получить решение в явном виде для отдельных вероятностей при $k > 1$ пока не удалось. Однако, воспользовавшись обычным методом приравнивания коэффициентов при θ_j , из уравнения (8.82) можно получить систему обыкновенных дифференциальных уравнений для моментов:

$$\left. \begin{aligned} (k\beta)^{-1} \frac{dm_1}{dt} &= -m_1 + 2m_k, \\ (k\beta)^{-1} \frac{dm_j}{dt} &= -m_j + m_{j-1}, \quad 2 \leq j \leq k, \end{aligned} \right\} \quad (8.84)$$

где m_j — среднее число особей, находящихся в j -й фазе в момент времени t . Очевидно, что эти уравнения идентичны соответствующим уравнениям для детерминистского случая. Они разрешимы, хотя и с некоторыми трудностями. Обозначив математическое ожидание для всей популяции как

$$m(t) = \sum_{j=1}^k m_j,$$

в конечном счете получаем

$$m(t) = e^{-k\beta t} \sum_{n=0}^{\infty} \sum_{r=0}^{k-1} \frac{2^n (k\beta t)^{nk+r}}{(nk+r)!}. \quad (8.85)$$

Можно показать, что для больших t

$$m(t) \approx \frac{\exp[(2^{1/k} - 1)k\beta t]}{2k(1 - 2^{-1/k})}. \quad (8.86)$$

На первый взгляд кажется, что приведенные здесь результаты (а также другие результаты такого рода, которые здесь не рассматривались) не имеют большого значения для биологии. Однако такое мнение было бы ошибочным. Хотя рассмотрение многофазового процесса размножения действительно приводит к очень сложному распределению размера популяции, не имеющему предельной формы для больших t , часто случается, что для процессов более общего характера, содержащих многофазовые компоненты, существуют стабильные предельные распределения, найти которые значительно легче. (Например, это довольно часто наблюдается в теории массового обслуживания.) Таким образом, в целом метод введения искусственных фаз очень важен и в перспективе обещает сыграть большую роль в построении более реалистичных моделей, чем многие из тех, которые применялись до сих пор. Можно надеяться, что такие модели не будут

слишком сложными с математической точки зрения; простота их структуры означает, что они подходят для исследований путем моделирования на вычислительных машинах и что соответствующие программы, по-видимому, не будут очень сложными.

8.4. ВЗАИМОДЕЙСТВИЕ МЕЖДУ ВИДАМИ

Еще одной важной экологической проблемой является взаимодействие между видами, например конкуренция за общие источники пищи или взаимоотношения хищник — жертва. На практике такие ситуации могут быть очень сложными, и пока что сделаны лишь первые, но многообещающие шаги на пути математического исследования этих сложных взаимоотношений. Классический детерминистский анализ многих проблем взаимодействия между видами провели еще в 20-х и 30-х годах Лотка и Вольтерра. Но сейчас, когда широкое признание получил стохастический подход, необходимо предпринять новые усилия и провести анализ вероятностными методами. К сожалению, эта задача значительно сложнее, чем изучение простых процессов размножения и гибели, рассмотренных ранее. Обычно начинают с построения какой-либо детерминистской модели, а затем пытаются определить, где в эту модель наиболее целесообразно включить вероятностные элементы, чтобы сделать ее более реалистичной. Рассмотрим кратко лишь одну-две частные задачи. Более полно эти вопросы рассматриваются в гл. 4 и 5 книги Бартлетта [12]; полезно также обратиться к гл. 13 другой книги автора [9].

Конкуренция между двумя видами

В разд. 1.3 мы уже рассматривали простую детерминистскую модель популяции, в которой наблюдается некоторое торможение роста, хотя в целом ее численность увеличивается. Если врожденная скорость роста популяции равна m , то ее размер будет возрастать экспоненциально и в момент t будет равен $n = ae^{mt}$. Но если ввести некоторый коэффициент замедления r , то результирующая скорость роста станет равной не m , а $m - rn$, и основное дифференциальное уравнение примет вид

$$\frac{dn}{dt} = mn - rn^2 \quad (8.87)$$

(это приводившееся ранее уравнение (1.3)). Решение этого уравнения имеет вид

$$n = \frac{m}{r + (m/a - r)e^{-mt}}, \quad (8.88)$$

что соответствует логистической кривой.

Уравнение (8.37) можно легко обобщить к виду

$$\frac{dn}{dt} = [\beta - f(n)] n, \quad (8.89)$$

где β обозначает врожденную скорость размножения, а некоторая функция $f(n)$ — скорость гибели. Считается, что эта функция зависит от размера популяции и учитывает «давление» окружающей среды, обуславливаемое нехваткой пищи, скученностью и т. д.

Пусть теперь имеются две популяции n_1 и n_2 со скоростями размножения β_1 и β_2 и скоростями гибели f_1 и f_2 соответственно. Тогда можно выписать два дифференциальных уравнения, аналогичных уравнению (8.89):

$$\left. \begin{aligned} \frac{dn_1}{dt} &= (\beta_1 - f_1) n_1, \\ \frac{dn_2}{dt} &= (\beta_2 - f_2) n_2, \end{aligned} \right\} \quad (8.90)$$

где величины f_1 и f_2 удобно рассматривать как линейные функции размеров обеих популяций, т. е.

$$f_i = \mu_{i1} n_1 + \mu_{i2} n_2, \quad i = 1, 2. \quad (8.91)$$

Уравнения (8.90) имеют стационарные решения, получаемые при $dn_1/dt = 0$ и $dn_2/dt = 0$. Решив систему уравнений

$$\left. \begin{aligned} \mu_{11} n_1 + \mu_{12} n_2 &= \beta_1, \\ \mu_{21} n_1 + \mu_{22} n_2 &= \beta_2, \end{aligned} \right\} \quad (8.92)$$

получаем точку равновесия

$$(A, B) \equiv \left(\frac{\beta_1 \mu_{22} - \beta_2 \mu_{12}}{\mu_{11} \mu_{22} - \mu_{12} \mu_{21}}, \frac{\beta_2 \mu_{11} - \beta_1 \mu_{21}}{\mu_{11} \mu_{22} - \mu_{12} \mu_{21}} \right) \quad (8.93)$$

при условии, что $\mu_{11} \mu_{22} \neq \mu_{12} \mu_{21}$. Можно исследовать небольшие отклонения от равновесного состояния, записывая

$$n_1 = A(1 + \xi), \quad n_2 = B(1 + \eta), \quad (8.94)$$

где ξ и η — малые величины. Подстановка этих выражений в систему (8.90) дает дифференциальные уравнения

$$\left. \begin{aligned} \left(\frac{d}{dt} + A\mu_{11} \right) \xi + B\mu_{12}\eta &= 0, \\ A\mu_{21}\xi + \left(\frac{d}{dt} + B\mu_{22} \right) \eta &= 0. \end{aligned} \right\} \quad (8.95)$$

Характер равновесия зависит от вида корней квадратного уравнения

$$x^2 + (A\mu_{11} + B\mu_{22})x + AB(\mu_{11}\mu_{22} - \mu_{12}\mu_{21}) = 0. \quad (8.96)$$

Можно показать, что равновесие устойчиво при $\mu_{11}\mu_{22} > \mu_{12}\mu_{21}$.

Если же $\mu_{11}\mu_{22} < \mu_{12}\mu_{21}$, то равновесие будет неустойчивым и в конце концов точка (n_1, n_2) переместится в зависимости от начальных условий на ту или на другую ось. Таким образом, в этой детерминистской модели вопрос о том, какая популяция выживет, зависит прежде всего от того, в каком направлении начинает развиваться процесс. Если теперь подумать об изменениях, которые могут возникнуть при стохастической формулировке этой задачи, то следует ожидать, что случайные изменения в начале процесса окажут существенное влияние на конечный результат. Более того, вполне возможно, что при одинаковых начальных условиях выживет любая из двух популяций. В действительности жизненные циклы конкурирующих видов обычно имеют ряд осложняющих особенностей, и у нас нет возможности подробно рассматривать здесь различные конкретные случаи. Превосходное описание экспериментальных исследований, проведенных Парком на двух видах мучных хрущаков, можно найти в гл. 5 книги Бартлетта [12].

Заметим только, что, используя уравнение (8.49), нетрудно записать дифференциальное уравнение в частных производных для производящей функции моментов полной стохастической модели всего процесса, рассмотренного выше. Но, к сожалению, решения этих уравнений в явном виде пока не получены. Более того, даже вывод обычным способом обыкновенных дифференциальных уравнений для моментов или семиинвариантов не очень облегчает задачу, так как первые два уравнения содержат пять неизвестных и начать процесс последовательного решения уравнений невозможно.

Модель системы хищник — жертва (или паразит — хозяин)

Другая ситуация, заслуживающая краткого рассмотрения, возникает в том случае, когда один из видов представляет собой частичный источник пищи для другого вида. Допустим, что первый вид — это жертва (или хозяин), а другой — хищник (или паразит). При классическом подходе, как и ранее, принимается допущение о том, что жертва имеет некоторую врожденную скорость размножения β_1 , а скорость ее гибели пропорциональна размеру популяции хищника и равна, например, $\mu_1 n_2$. Для хищника врожденная скорость гибели равна μ_2 , а скорость размножения пропорциональна размеру популяции жертвы, т. е. равна $\beta_2 n_1$. Таким образом, получаем основные уравнения движения:

$$\left. \begin{aligned} \frac{dn_1}{dt} &= \beta_1 n_1 - \mu_1 n_1 n_2, \\ \frac{dn_2}{dt} &= \beta_2 n_1 n_2 - \mu_2 n_2. \end{aligned} \right\} \quad (8.97)$$

Ясно, что точкой равновесия является точка $(\mu_2/\beta_2, \beta_1/\mu_1)$, и можно показать, что эта точка нейтральна в том смысле, что при движении в нее из других точек, лежащих в плоскости (n_1, n_2) , отсутствует затухание.

Разделив первое из уравнений (8.97) на второе, получим

$$\frac{dn_1}{dn_2} = \frac{(\beta_1 - \mu_1 n_2) n_1}{(\beta_2 n_1 - \mu_2) n_2},$$

что после интегрирования дает

$$f(n_1, n_2) \equiv -\mu_2 \log n_1 + \beta_2 n_1 - \beta_1 \log n_2 + \mu_1 n_2 = \text{const.} \quad (8.98)$$

Решение этого уравнения выражается семейством замкнутых кривых $f(n_1, n_2) = \text{const}$ при различных значениях постоянной.

Для возможных кривых, проходящих вблизи точки равновесия, можно положить

$$n_1 = \frac{\mu_2}{\beta_2} (1 + \xi), \quad n_2 = \frac{\beta_1}{\mu_1} (1 + \eta), \quad (8.99)$$

где ξ и η — малые величины. Подставляя эти выражения в формулу (8.98) и беря только члены, содержащие ξ и η лишь в первой и второй степенях, получаем множество эллипсов

$$\mu_2 \xi^2 + \beta_1 \eta^2 = \text{const.} \quad (8.100)$$

Такая форма кривой сохраняется только вблизи точки равновесия. Более крупные кривые будут деформироваться и терять форму эллипса.

На условия нейтрального равновесия указывает тот факт, что изменения размера популяции должны происходить в соответствии с движением точки (n_1, n_2) вдоль некоторого определенного контура семейства $f(n_1, n_2) = \text{const}$, и тщательное рассмотрение дифференциалов показывает, что кривые обтекаются в направлении против движения часовой стрелки.

Чтобы более детально выяснить поведение модели вблизи точки равновесия, подставим соотношения (8.99) в уравнения (8.97). Для малых ξ и η получаем

$$\frac{d\xi}{dt} = -\beta_1 \eta, \quad \frac{d\eta}{dt} = \mu_2 \xi. \quad (8.101)$$

Следовательно,

$$\frac{d^2 \xi}{dt^2} = -\beta_1 \mu_2 \xi, \quad (8.102)$$

т. е. движение является периодическим с периодом

$$T = 2\pi (\beta_1 \mu_2)^{-1/2}. \quad (8.103)$$

Рассмотрим теперь изменения, которые должно вызвать включение в модель вероятностных элементов. Как и в случае конкуренции, рассмотренном выше, здесь также можно разработать полную стохастическую модель, но, как и ранее, для нее нельзя получить решения в явном виде. Поэтому мы будем искать более слабые допущения. Допустим, например, что точка равновесия находится на некотором расстоянии от каждой из осей. Тогда для контуров, на которых значения n_1 и n_2 остаются достаточно большими, вполне удовлетворительной будет детерминистская модель. Но если в некоторой точке какого-то соответствующего контура какая-либо переменная не очень велика, то существенное значение могут приобрести случайные колебания. Если они смогут сыграть свою роль, то точка (n_1, n_2) переместится на одну из осей, что означает вымирание соответствующего вида. Таким образом, стохастическая модель оказывается неустойчивой, так как стохастический «дрейф» рано или поздно приводит к вымиранию одного из видов. В такого рода модели хищник в конечном счете, безусловно, вымирает; это может произойти либо чисто случайно, либо вследствие того, что вначале элиминируется популяция его жертвы.

Наше изложение велось в довольно общих количественных выражениях. Его можно сделать несколько более конкретным для некоторых частных случаев. Так, допустим, что популяция жертвы весьма велика. Вероятность ее вымирания в соответствующей стохастической модели мала, и для рассмотрения этого вида вполне подойдет детерминистское описание. Однако если численность хищников невелика, то случайные колебания будут существенно влиять на нее и это влияние будет наибольшим, когда точка, соответствующая популяции хищников, будет находиться на наименьшем расстоянии от оси n_1 . Можно показать (см. разд. 4.5 книги Бартлетта [12] и разд. 13.3 книги Бейли [9]), что в окрестности этой точки численность популяции хищника изменяется примерно так же, как в процессе размножения и гибели, неоднородном во времени, со скоростью размножения $\mu_2 + \gamma \sin(2\pi t/T)$ и скоростью гибели μ_2 (постоянная γ зависит от конкретного контура, по которому следует процесс). Поэтому в данном случае можно использовать теорию неоднородных процессов размножения и гибели, в частности для того, чтобы получить грубую оценку вероятности вымирания на данном этапе цикла.

Бартлетт приводит интересный пример использования этого метода для анализа данных Гаузе, полученных при выращивании *Paramecium aurelia* на дрожжевых клетках. Описанная модель применима здесь, по-видимому, лишь в самом общем приближенном виде, однако на основании этих данных получаются численные значения $\beta_1 \sim 1$, $\mu_2 \sim 0,45$ (время измеряется в сутках);

точка равновесия имеет координаты $(1,5 \cdot 10^7; 100)$. Отсюда следует, что $\mu_1 \sim 0,01$, $\beta_2 \sim 3 \cdot 10^{-8}$. Таким образом, длительность цикла составляет около 9,4 суток. Дальнейшие вычисления показывают, что вероятность вымирания составляет около 0,003, т. е. очень мала для любого конкретного цикла.

8.5. ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Из всего изложенного выше должно быть ясно, что существует множество математических методов, приемлемых для построения и анализа моделей процессов, происходящих в популяциях. В большинстве случаев это пока довольно элементарные и чрезмерно упрощенные модели, однако они указывают направление более реалистичных исследований в будущем.

Хотя вначале целесообразно изучать соответствующую детерминистскую модель, необходимо соблюдать большую осторожность и не слишком полагаться на ее результаты, если числа не столь велики, чтобы можно было пренебречь случайными колебаниями. Вообще говоря, не следует бояться трудностей, возникающих при построении стохастических моделей, и мы видели, что часто без больших усилий можно получить такие показатели, как математическое ожидание, дисперсия, вероятность вымирания и т. д. Обычно эти процессы можно описывать более или менее эвристически, методами прикладной математики, не обращая слишком большого внимания на математические строгости. Однако в этом таится определенная опасность. Вполне возможно, что построенная модель будет обладать неожиданными свойствами и потребует тщательного математического исследования.

Рассмотрим только один простой пример. Допустим, что имеет место детерминированный процесс размножения, в котором скорость размножения *в расчете на одну особь* пропорциональна размеру популяции. Это означает, что приращение популяции в интервале Δt за счет размножения одной особи можно считать равным $\beta n \Delta t$, где n — размер популяции. Следовательно, общее приращение популяции равно $\beta n^2 \Delta t$, и уравнение движения имеет вид

$$\frac{dn}{dt} = \beta n^2. \quad (8.104)$$

При начальном условии $n(0) = a$ решение этого уравнения имеет вид

$$n(t) = \frac{a}{1 - \beta a t}. \quad (8.105)$$

Ясно, что

$$\lim_{t \rightarrow 1/(\beta a)} n(t) = \infty, \quad (8.106)$$

т. е. за конечный промежуток времени численность популяции стремится к бесконечности. Такой *расходящийся* или *взрывоподобный* рост на практике происходить не может, хотя иногда эта модель может служить полезной аппроксимацией очень быстрого роста популяций в течение довольно коротких промежутков времени.

Соответствующие условия существуют и в случае стохастических моделей. В вероятностном аналоге описанной выше ситуации взрывоподобный рост происходит в том случае, когда $\sum_n p_n(t) < 1$

для некоторых t ; иногда может потребоваться исследовать стохастическую модель, чтобы убедиться в том, что она не обладает нежелательным (обычно) свойством взрывоподобности. Однако, как правило, физической и биологической интуиции достаточно для обнаружения моделей с неудобными математическими свойствами.

Часто оказывается возможной замена дискретных случайных величин, описывающих популяцию, непрерывными и использование соответствующей методики предельного перехода к диффузионному процессу. Эти приемы находят весьма широкое применение, особенно при изучении флуктуаций частот генов в генетике, однако в этих случаях преобладают вырожденные формы уравнений диффузии, и для получения точных результатов приходится применять специальные методы математического анализа.

Исследования методом физического моделирования сопряжены с меньшими затруднениями, и многие основные идеи, о которых говорилось выше, могут быть использованы для построения довольно реалистичных моделей. В книге Бартлетта [12] рассматриваются некоторые его собственные работы в этом направлении, а также работы Лесли и Гоуэра. По мере распространения электронных вычислительных машин метод физического моделирования будет, по-видимому, приобретать все более важное значение в популяционных и экологических исследованиях.

Глава 9

ТЕОРИЯ ЭПИДЕМИЙ

9.1. ВВЕДЕНИЕ

Повсеместное распространение эпидемий чумы и голод, несомненно, были самыми главными причинами несчастий и страданий человечества. Общее число людей, погибших от эпидемий за многие столетия, измеряется астрономическими цифрами, и очень часто в отдельных странах погибала значительная часть всего населения. Приведем только один пример. В XIV в. в Европе «черная смерть» погубила около 25% всего населения, насчитывавшего в то время примерно 100 млн. человек. Хотя современные развитые страны избавились от бедствий такого масштаба, в Африке и на Востоке еще часто наблюдаются массовые эпидемии. Они представляют серьезную опасность не только для тех стран, где они возникают: благодаря развитию средств сообщения в наше время существует постоянный риск передачи эпидемии и в те районы, где естественный иммунитет слабее, хотя служба здравоохранения поставлена значительно лучше.

Окончательная победа над эпидемическими заболеваниями всегда была одной из первостепенных задач медицинской науки. В настоящее время появился ряд лекарственных препаратов, излечивающих многие болезни, однако наибольший эффект дают все же предупредительные меры. Например, во многих районах малярия исчезла просто благодаря осушению топей и болот. Таким образом, методы профилактики — надежное удаление сточных вод, обеспечение чистоты источников водоснабжения, организация карантина в случае подозрения на инфекцию, применение прививок и т. д. — имеют важнейшее значение. Однако для успешной борьбы с эпидемиями этого еще недостаточно; здесь необходимы объединенные усилия специалистов самых различных направлений. Существуют медицинские проблемы, связанные с диагностикой и прогнозированием, а также выбором определенных терапевтических мер в каждом отдельном случае. С этими вопросами тесно связано фармакологическое изучение новых лекарственных препаратов и попытки более глубокого проникновения в биологический механизм болезни в целом. Существуют также и чисто эпидемиологические проблемы, касающиеся распространения болезни по стране в целом. Для того чтобы органы

общественного здравоохранения могли принять наиболее эффективные меры в борьбе с эпидемией, необходимо уметь количественно оценивать сравнительные достоинства различных методов — введения карантина, вакцинации, выявления контактов, закрытия школ и общественных бань и т. д.

К строгому изучению всех аспектов этой проблемы можно приступить лишь на основе правильно сформулированных математических моделей независимо от того, идет ли речь о клиническом прогнозе, испытании различных методов лечения, глубоких биологических исследованиях или же мероприятиях, проводимых органами общественного здравоохранения. Начало этим исследованиям было положено в XVII в. количественной работой Гронта и Петти, составивших «билли о смертности в Лондоне»; однако только в начале XX в. появились первые модели, при разработке которых вводились определенные предположения о механизме распространения эпидемии, основанные на результатах бактериологических исследований, проведенных в конце XIX в. С тех пор достигнуты значительные успехи, особенно в области построения стохастических моделей, тесно связанных с результатами клинических и биологических исследований, однако создание настоящей математической эпидемиологии, по существу, еще только начинается. Более полное рассмотрение этого вопроса и ссылки на литературу даются в другой книге автора [6]. В настоящей главе мы лишь приведем некоторые наиболее важные математические результаты и укажем возможные направления дальнейшей работы.

Прежде чем приступать к математическому описанию различных явлений, связанных с развитием эпидемии, целесообразно остановиться на некоторых основных медицинских и биологических понятиях. Этот вопрос можно обсуждать на самых различных уровнях — от деталей поведения отдельных бактериальных клеток и вирусных частиц до макроскопических явлений распространения эпидемий в человеческом обществе.

Для нас основной интерес представляет последний уровень, однако мы всегда будем стараться выбирать такие модели, в которые при необходимости можно включить более тонкие детали. Представим себе, что индивидуум, *восприимчивый* к определенной инфекции, получает ее прямым или непрямым путем от другого зараженного индивидуума, или *источника инфекции*. Сначала в течение некоторого *латентного периода* развитие болезни у вновь зараженного индивидуума не имеет никаких внешних проявлений. Затем следует *заразный период*, в течение которого зараженный индивидуум может передавать заразное начало другим индивидуумам, восприимчивым к данной инфекции. Наконец, у зараженного индивидуума обнаруживаются различные

симптомы болезни, и его изолируют от общения с другими членами общества до тех пор, пока он не выздоровеет (или, возможно, умрет). В случае выздоровления он может стать невосприимчивым к этой болезни в дальнейшем, во всяком случае в течение некоторого времени. Изоляция по существу обрывает заразный период. Промежуток времени от заражения до появления симптомов называется *инкубационным периодом*. Кроме того, необходимо принимать во внимание промежуток времени от момента появления симптомов у одного индивидуума до момента появления симптомов у другого индивидуума, непосредственно заразившегося от первого (будем называть этот промежуток времени *серийным интервалом*). В большинстве случаев это как раз тот эпидемиологический показатель, определить который особенно просто, однако для того, чтобы модель ближе соответствовала биологическому механизму передачи инфекции, по-видимому, необходим более детальный анализ с рассмотрением латентного и заразного периодов.

Обычно для простоты предполагают, что все индивидуумы сообщества непрерывно и равномерно перемешиваются. Это условие приемлемо в качестве первого приближения, однако ясно, что оно представляет собой чрезмерное упрощение. В действительности же большинство реальных популяций имеют крайне сложную структуру, обусловленную социальным расслоением, разнообразием географических условий и сложными временными и пространственными схемами перемещения. Пока что для рассмотрения всех этих аспектов сделано очень мало.

В математической теории эпидемий еще резче, чем при исследовании популяций, рассмотренном в предыдущей главе, обнажается серьезное противоречие между требованием реалистичности модели и возможностью ее анализа (об этом уже говорилось в первой части книги). Даже простейшие стохастические модели распространения эпидемий оказываются крайне сложными с математической точки зрения. Поэтому вопрос о том, какую модель выбрать для исследования и какие методы анализа использовать, имеет решающее значение.

Существует очень удобное деление теории эпидемий на две части: изучение больших групп и изучение малых групп. Теория больших групп занимается общим исследованием характера возникновения эпидемий в целом сообществе или в больших популяциях и рассматривает довольно общие модели распространения эпидемий в весьма упрощенной форме. Основное значение этих исследований состоит в том, что они связаны с работой органов общественного здравоохранения. Теорию малых групп можно разработать более детально. Она позволяет не только составлять общие прогнозы возможного развития той или иной эпидемии

среди группы школьников или в семье, но и получать информацию по вопросам, имеющим более конкретное клиническое или биологическое значение (например, данные о продолжительности латентного и заразного периодов). Разумеется, эти два подхода дополняют друг друга, однако до некоторой степени они могут использоваться независимо друг от друга.

9.2. ПРОСТЫЕ ЭПИДЕМИИ

Рассмотрим вначале эпидемию простейшего вида, т. е. случай, когда заболевание распространяется среди группы восприимчивых индивидуумов, но удаления их из популяции за счет гибели, выздоровления или изоляции не происходит. Такая аппроксимация может оказаться приемлемой для начальных стадий некоторых заболеваний верхних дыхательных путей, так как при таких заболеваниях может пройти большой промежуток времени, прежде чем источник инфекции будет удален из популяции. Допустим, что имеется n индивидуумов, восприимчивых к данному заболеванию, и что в момент времени $t = 0$ в группу попадает один источник инфекции. Естественно начинать исследование с детерминистской модели, хотя можно предположить, что она не очень подойдет для эпидемий, начинающихся при небольшом числе источников инфекции, так как статистические колебания в группе зараженных индивидуумов будут ощутимы даже в том случае, если значение n достаточно велико.

Детерминистская модель

Рассмотрим однородно перемешанную группу, состоящую из $n + 1$ индивидуумов. Пусть в момент t в этой группе имеется x восприимчивых индивидуумов и y источников инфекции, т. е. $x + y = n + 1$. Разумно предположить, что среднее число новых случаев заболевания, появляющихся в интервале Δt , будет пропорционально как числу источников инфекции, так и числу восприимчивых индивидуумов. Если частота контактов между членами этой группы равна β , то среднее число новых случаев заболевания, появляющихся в интервале Δt , будет равно $\beta xy \Delta t$, т. е.

$$\Delta x = -\beta xy \Delta t.$$

Удобно изменить временную шкалу, введя переменную $\tau = \beta t$; тогда уравнение движения для данного процесса принимает вид

$$\frac{dx}{d\tau} = -x(n - x + 1) \quad (9.1)$$

при начальном условии

$$x = n, \quad \tau = 0. \quad (9.2)$$

По существу, мы предполагаем здесь, что зараженный индивидuum становится заразным для остальных восприимчивых индивидуумов сразу после того, как он сам заразится, т. е. что латентный период равен нулю.

Решение уравнения (9.1) при условии (9.2) имеет вид

$$x = \frac{n(n+1)}{n + e^{(n+1)\tau}}. \quad (9.3)$$

На практике во время эпидемии регистрируется обычно число новых случаев, появляющихся за сутки или за неделю. Поэтому более удобно рассматривать динамику нарастания числа новых случаев, описываемую так называемой *эпидемической кривой*. Соответствующее ей уравнение имеет вид

$$-\frac{dx}{d\tau} = x(n-x+1) = \frac{\beta n(n+1)^2 e^{(n+1)\tau}}{\{n + e^{(n+1)\tau}\}^2}. \quad (9.4)$$

Это симметричная одновершинная кривая с максимумом в точке

$$\tau = \frac{\ln n}{n+1}.$$

Таким образом, получено характерное свойство эпидемий: число новых случаев сначала быстро возрастает, в какой-то момент достигает максимума, а затем уменьшается до нуля. Такая форма эпидемической кривой является чисто математическим следствием принятого допущения о том, что среднее число новых случаев пропорционально как числу восприимчивых индивидуумов, так и числу источников инфекции. Нет необходимости объяснять форму кривой изменением вирулентности, хотя это и может иметь место.

Стохастическая модель

Рассмотрим теперь свойства аналогичной стохастической модели и обозначим через $X(t)$ случайную величину, характеризующую число восприимчивых индивидуумов в момент t . Естественно допустить, что вероятность появления одного нового случая пропорциональна числу восприимчивых индивидуумов и числу источников инфекции. Таким образом, фактическую вероятность перехода для интервала Δt можно записать как $\beta X(n-X+1)\Delta t$. Снова изменим масштаб времени, введя переменную $\tau = \beta t$, и используем метод, применявшийся в разд. 8.2. Единственным значением функции f , отличным от нуля, является значение $f_{-1} = X(n-X+1)$. Отсюда непосредственно следует, что дифференциальное уравнение в частных производных для производящей функции вероятностей имеет вид

$$\frac{\partial P}{\partial \tau} = (1-x) \left(n \frac{\partial P}{\partial x} - x \frac{\partial^2 P}{\partial x^2} \right) \quad (9.5)$$

при начальном условии

$$P(x, 0) = x^n. \quad (9.6)$$

Как и в случае детерминистской модели, здесь снова неявно предполагалось, что латентный период равен нулю.

Уравнение (9.5) представляет собой стандартное дифференциальное уравнение в частных производных второго порядка, решение которого можно получить в виде разложения по собственным функциям. В том виде, как оно записано, это уравнение имеет кратные собственные значения, и в целом данный метод довольно сложен. При нецелочисленных $n = N$ решение уравнения (9.5) содержит ряд гипергеометрических функций. Затем полагаем $N \rightarrow n$; этот предельный переход влечет за собой ряд усложнений. Во многих отношениях легче перейти непосредственно к дифференциально-разностным уравнениям для отдельных вероятностей $p_r(\tau)$. Эти уравнения, которые можно вывести из уравнения (9.5) или получить обычным способом, имеют вид

$$\begin{aligned} \frac{dp_r}{d\tau} &= (r+1)(n-r)p_{r+1} - r(n-r+1)p_r, \quad 0 \leq r \leq n-1, \\ \frac{dp_n}{d\tau} &= -np_n \end{aligned} \quad (9.7)$$

при начальном условии

$$p_n(0) = 1. \quad (9.8)$$

Последовательное решение уравнений (9.7) в принципе возможно, но на практике оно неосуществимо. Однако, используя преобразования Лапласа, легко находим

$$q_r = \frac{n!(n-r)!}{r!} \prod_{j=1}^{n-r+1} \{s+j(n-j+1)\}^{-1}, \quad 0 \leq r \leq n, \quad (9.9)$$

где

$$q_r = \int_0^{\infty} e^{-s\tau} p_r(\tau) d\tau. \quad (9.10)$$

Путем обратного преобразования правой части выражения (9.9) можно непосредственно найти значения $p_r(\tau)$, хотя получаемые результаты очень громоздки. Это обусловлено появлением кратных корней при разложении на элементарные дроби.

Например, если n — четное число и $r > \frac{1}{2}n$, то методами элементарной алгебры находим

$$p_r(\tau) = \sum_{k=1}^{n-r+1} c_{rk} e^{-k(n-k+1)\tau}, \quad (9.11)$$

где

$$c_{rk} = \frac{(-1)^{k-1} (n-2k+1) n! (n-r)! (r-k-1)!}{r! (k-1)! (n-k)! (n-r-k+1)!}. \quad (9.12)$$

При $r \leq \frac{1}{2}n$ в разложении на элементарные дроби появляются множители, соответствующие кратным корням, и выражение для $p_r(\tau)$ оказывается значительно сложнее. Основная цель вывода результатов (9.11) и (9.12) состоит в том, чтобы показать, насколько сложны выражения для точных значений вероятностей даже в случае простейшей стохастической модели эпидемии. Столь же сложны и выражения для математического ожидания. Так как

$$m(\tau) = \sum_{r=1}^n r p_r(\tau),$$

то, подставляя сюда значение $p_r(\tau)$ из формулы (9.11), после очень сложных алгебраических преобразований получаем

$$m(\tau) = \sum_{r=1}^n \left[\frac{n!}{(n-r)! (r-1)!} \left\{ (n-2r+1)^2 \tau + 2 - \right. \right. \\ \left. \left. - (n-2r+1) \sum_{w=r}^{n-r} w^{-1} \right\} e^{-r(n-r+1)\tau} \right], \quad (9.13)$$

где r стремится к $\frac{1}{2}n$ при четном n и к $\frac{1}{2}(n+1)$ при нечетном n .

В последнем случае член, соответствующий $r = \frac{1}{2}(n+1)$, необходимо уменьшить вдвое. С помощью уравнения (9.13) эпидемическую кривую $-dm/d\tau$ можно оценить численно. Результаты для таких значений n , как 10, 20 и 30, показывают, что стохастическая эпидемическая кривая достигает своего максимума почти в то же время, что и соответствующая кривая, полученная на основании детерминистской модели, но она асимметрична и поднимается и спускается более полого.

Ясно, что, хотя получение точных результатов чисто алгебраическими методами может представлять интерес с точки зрения математики, на практике это задача весьма неблагоприятная. В разд. 3.4 и 5.5 мы уже упоминали о других математических и вычислительных методах, пригодных для решения этой задачи, и во избежание повторения отсылаем читателя к этим разделам. Можно попытаться найти численное решение дифференциального уравнения в частных производных (9.5) или системы обыкновенных дифференциальных уравнений (9.7), однако при этом способе трудно получить требуемую точность. Методы моделирования, вероятно, гораздо более перспективны, и их легко применять

даже в случае процессов более общего типа, чем тот, который рассматривается нами сейчас. Нельзя, однако, забывать, что такой подход в значительной мере лишает нас возможности реального проникновения в математическую структуру задачи. Поэтому результаты моделирования желательно использовать для проверки новых математических результатов, особенно полученных приближенными методами.

Наконец, заметим, что в простой стохастической модели эпидемии имеется один параметр, определить который несколько легче. Это *длительность эпидемии*, т. е. время, требуемое для того, чтобы вся группа восприимчивых индивидуумов оказалась зараженной. Если имеется j источников инфекции и $n - j + 1$ восприимчивых индивидуумов, то вероятность появления в интервале $\Delta\tau$ нового источника инфекции равна $j(n - j + 1)\Delta\tau$, а длительность промежутка времени τ_j до момента появления этого нового источника инфекции имеет экспоненциальное распределение

$$f(\tau_j) = j(n - j + 1) \exp\{-j(n - j + 1)\tau_j\}. \quad (9.14)$$

Длительность эпидемии T равна сумме величин τ_j :

$$T = \sum_{j=1}^n \tau_j. \quad (9.15)$$

Далее, r -й семиинвариант распределения (9.14) равен $(r - 1)! j^{-r} (n - j + 1)^{-r}$. Все значения τ_j независимы и, следовательно, r -й семиинвариант k_r распределения случайной величины T имеет вид

$$k_r = (r - 1)! \sum_{j=1}^n j^{-r} (n - j + 1)^{-r}. \quad (9.16)$$

Для малых n семиинварианты можно вычислить непосредственно по формуле (9.16). Для больших значений n имеются асимптотические аппроксимации

$$\left. \begin{aligned} k_1 &= \frac{2(\gamma + \log n)}{n+1}, \\ k_r &= \frac{2(r-1)}{n^r} \zeta(r), \end{aligned} \right\} \quad (9.17)$$

где γ — постоянная Эйлера, а $\zeta(r)$ — функция Римана. Легко показать, что при $n \rightarrow \infty$ показатели асимметрии и эксцесса γ_1 и γ_2 приближаются соответственно к 0,8 и 1,2, что указывает на заметное отклонение от нормальности. Кроме того, даже для средних групп коэффициент изменчивости довольно велик: при $n = 20$ он равен 0,27. Отсюда следует, что характер течения эпидемии подвержен резким колебаниям, обусловленным чисто

случайными причинами, и в тех характерных случаях, когда эпидемия распространяется очень медленно или слишком быстро, не следует поспешно относить это за счет каких-то особых причин — скажем, необычно низкой вирулентности возбудителя или особо высокой контагиозности инфекции.

9.3. ОБЩИЙ СЛУЧАЙ ЭПИДЕМИИ

С практической точки зрения больший интерес представляет обобщение рассмотренной здесь модели простой эпидемии на случай, когда зараженные индивидуумы могут удаляться из коллектива. Как и ранее, рассмотрим вначале соответствующую детерминистскую модель.

Детерминистская модель

Допустим, что имеется группа из n однородно перемешивающихся индивидуумов, в которой в момент t имеется x восприимчивых индивидуумов, y источников инфекции и z удаленных (т. е. изолированных, умерших или выздоровевших и ставших невосприимчивыми к инфекции) индивидуумов. Таким образом, $x + y + z = n$. Вновь допустим, что частота контактов равна β , так что среднее число новых случаев в интервале Δt равно $\beta xy \Delta t$. Теперь следует учесть также частоту случаев удаления γ , т. е. что в интервале Δt зараженную группу покидают $\gamma y \Delta t$ индивидуумов.

Уравнения движения для этого процесса имеют вид

$$\left. \begin{aligned} \frac{dx}{dt} &= -\beta xy, \\ \frac{dy}{dt} &= \beta xy - \gamma y, \\ \frac{dz}{dt} &= \gamma y \end{aligned} \right\} \quad (9.18)$$

при начальном условии $(x, y, z) = (x_0, y_0, 0)$ в момент $t = 0$. Если исходное число источников инфекции, т. е. y_0 , очень мало, то можно считать, что $x_0 \approx n$. Из второго уравнения системы (9.18) непосредственно следует, что эпидемия не может начаться, если не выполняется условие $x_0 > \gamma/\beta$. Назовем величину $\rho = \gamma/\beta$ относительной частотой удаления зараженных индивидуумов; тогда пороговым значением этой величины будет значение $\rho = x_0 \approx n$. При плотности восприимчивых индивидуумов, лежащей ниже этого значения, первоначальные случаи заболевания исчезнут раньше, чем инфекция передастся другим индивидуумам. Но если эта плотность выше пороговой, то эпидемия возникнет даже в том случае, если первоначальное число источников инфек-

дии очень незначительно. Естественно, интересно узнать, насколько большой вспышки заболевания можно ожидать.

Хотя для системы уравнений (9.18) можно найти точное решение, в данном случае можно воспользоваться приближенными методами. Неизвестное y можно исключить, разделив первое уравнение на третье. В результате имеем

$$dx/dz = -x/\rho,$$

что после интегрирования дает

$$x = x_0 e^{-z/\rho}. \quad (9.19)$$

Подставляя в третье уравнение системы (9.18) значение $y = n - x - z$ и взяв x из формулы (9.19), имеем

$$\frac{dz}{dt} = \gamma(n - x - z) = \gamma(n - z - x_0 e^{-z/\rho}). \quad (9.20)$$

Разложение правой части уравнения (9.20) в ряд до члена, содержащего z^2 , дает уравнение, разрешимое обычными методами. Нет необходимости выписывать точное его решение. Достаточно лишь указать, что функция dz/dt описывается колоколообразной симметричной кривой. Ее удобно рассматривать как эпидемическую кривую для данной модели, так как изолируют обычно только тех зараженных индивидуумов, у которых проявляются практически различные симптомы. Так, статистика общественного здравоохранения регистрирует каждый день или каждую неделю число именно таких индивидуумов, а не число новых источников инфекции.

Если нас интересует полный размер эпидемии z_∞ , то его можно определить непосредственно из уравнения (9.20). При $t \rightarrow \infty$ имеем $dz/dt \rightarrow 0$. Если разложить правую часть уравнения (9.20) в ряд до члена, содержащего z^2 , и произвести перегруппировку членов, то, подставляя $x_0 \approx n$, сразу же находим

$$z_\infty \approx 2\rho \left(1 - \frac{\rho}{n}\right). \quad (9.21)$$

Как указывалось выше, при $\rho > n$ эпидемия не возникает. При $\rho < n$ запишем

$$n = \rho + v. \quad (9.22)$$

Подставляя это соотношение в формулу (9.21), для малых v получаем

$$z_\infty \approx 2v, \quad (9.23)$$

т. е. начальная плотность восприимчивых индивидуумов $\rho + v$ в конечном счете уменьшается до $\rho - v$, т. е. до величины, отстоящей ниже порогового значения настолько, насколько эта плотность первоначально была выше порогового значения. Эта известная теорема Кермака и Мак-Кендрика соответствует тому обще-

известному факту, что эпидемия легче возникает в том случае, если вследствие скученности плотность восприимчивых индивидуумов высока, а вследствие отсутствия необходимых знаний и неудовлетворительной изоляции зараженных индивидуумов частота их удаления из коллектива относительно низка. Если благодаря хорошим социальным условиям обеспечивается низкая плотность восприимчивых индивидуумов, а благодаря хорошей организации здравоохранения и службы профилактики частота удаления зараженных индивидуумов из коллектива относительно велика, то эпидемия примет ограниченные размеры.

Естественно, что при рассмотрении соответствующей стохастической модели результаты будут более расплывчатыми, и это может иметь серьезное практическое значение, когда эпидемии начинаются при небольшом числе источников инфекции, даже если популяция велика.

Стохастическая модель

Стохастический вариант даже простой эпидемии достаточно сложен. Не удивительно, что в общем случае для анализа стохастической модели эпидемии требуется еще более сложный математический аппарат. По-настоящему удовлетворительное описание основных характеристик такого процесса еще не достигнуто, но ряд отдельных полезных результатов уже получен.

Рассмотрим вначале исходную модель и вывод основных уравнений движения. В данном случае имеются две существенно различные случайные величины. Пусть, как и ранее, $X(t)$ обозначает число восприимчивых индивидуумов в момент времени t , а $Y(t)$ — число источников инфекции. Таким образом, мы имеем дело с двумерным процессом, аналогичным тому, который был рассмотрен в разд. 8.3. Здесь возможны переходы двух видов. Снова примем частоту контактов равной β ; тогда вероятность появления в интервале Δt нового источника инфекции будет равна $\beta XY \Delta t$. Если частота удаления из коллектива зараженных индивидуумов равна γ , то вероятность того, что в интервале Δt будет удален один индивидуум, составит $\gamma Y \Delta t$. В данном случае возможны два значения функции f , отличные от нуля; в обозначениях, принятых в разд. 8.2 и 8.3, они имеют вид $f_{-1,1} = \beta XY$ и $f_{0,-1} = \gamma Y$. Если изменить временной масштаб, перейдя к $\tau = \beta t$, и обозначить через $\gamma/\beta = \rho$ относительную частоту удаления, то, используя уравнение (8.48), получим следующее дифференциальное уравнение в частных производных для производящей функции вероятностей:

$$\frac{\partial P}{\partial \tau} = (y^2 - xy) \frac{\partial^2 P}{\partial x \partial y} + \rho (1 - y) \frac{\partial P}{\partial y} \quad (9.24)$$

при начальном условии

$$P(x, y, 0) = x^n y^a \quad (9.25)$$

(в предположении, что процесс начинается при наличии n восприимчивых индивидуумов и a источников инфекции).

До сих пор непосредственно решить уравнение (9.24) в простом замкнутом виде еще не удалось. Попытки использовать обыкновенные дифференциальные уравнения для моментов или семиинвариантов, выведенные обычным способом, также не увенчались успехом по тем же причинам, что и в случае модели конкуренции между двумя видами, рассмотренной в разд. 8.4. (Такая же трудность возникает даже в случае простой стохастической эпидемии.) Однако не исключено, что уравнение (9.24) можно будет использовать как основу для дальнейших исследований.

Если вероятность того, что в момент τ имеется j восприимчивых индивидуумов и k источников инфекции, равна $p_{jk}(\tau)$, то подстановка производящей функции вероятностей

$$P(x, y, \tau) = \sum_{jk} p_{jk}(\tau) x^j y^k,$$

в уравнение (9.24) дает систему дифференциальных уравнений

$$\left. \begin{aligned} \frac{dp_{jk}}{d\tau} &= (j+1)(k-1)p_{j+1, k-1} - k(j+\rho)p_{jk} + \rho(k+1)p_{j, k+1}, \\ \frac{dp_{na}}{d\tau} &= -a(n+\rho)p_{na}, \end{aligned} \right\} \quad (9.26)$$

где $0 \leq j+k \leq n+a$, $0 \leq j \leq n$, $0 \leq k \leq n+a$, и

$$p_{na}(0) = 1. \quad (9.27)$$

В принципе эти уравнения можно решить непосредственно с помощью преобразований Лапласа. Однако получающиеся алгебраические выражения столь громоздки, что практически этот метод совершенно непригоден.

Некоторого успеха можно добиться в предельном случае при $t \rightarrow \infty$, когда $dp_{jk}/d\tau \rightarrow 0$. Здесь можно получить довольно простую треугольную систему линейных уравнений, решение которой дает вероятность P_w того, что дополнительно к первоначальным случаям эпидемия охватит еще w индивидуумов. Для получения конкретных результатов необходимо провести численные расчеты; были рассчитаны распределения общего числа зараженных индивидуумов для $n = 10, 20$ и 40 при $a = 1$ и различных значениях ρ . Как и ожидалось, при $\rho \geq n$ все распределения имеют J-образную форму с максимальным значением в точке $w = 0$. Если же $\rho \ll n$, то распределения имеют U-образную форму, т. е. возможна очень малая или очень большая вспышка, тогда как промежуточные состояния наблюдаются редко.

Таким образом, хотя при столь малых значениях n (не более 40) резкие переходы отсутствуют, имеются две различные схемы распространения эпидемии.

При больших n справедлива теорема о стохастическом пороговом значении, принадлежащая Уиттлу. Не входя во все детали анализа, проведенного Уиттлом, с помощью следующих приближенных рассуждений легко показать, чего именно можно ожидать в этом случае. Если n достаточно велико, то ρ (во всяком случае, в начальный период) численность группы источников инфекции изменяется примерно по тому же закону, которому подчиняется процесс размножения и гибели со скоростями размножения и гибели, равными соответственно βn и γ . Теперь используем формулу (8.35), выражающую вероятность вымирания популяции, заменив λ на βn и μ на γ . Из нее следует, что вероятность прекращения эпидемического процесса равна 1 при $\rho \geq n$ и $(\rho/n)^n$ при $\rho < n$. В первом случае исходная группа источников инфекции, безусловно, элиминирует и можно ожидать, что общее число заболеваний будет мало. Во втором случае с вероятностью $(\rho/n)^n$ можно ожидать малой вспышки и с вероятностью $1 - (\rho/n)^n$ — большой вспышки эпидемии.

Стохастические модели с такими общими свойствами весьма полезны, хотя и до известного предела. Несмотря на присущие им ограничения, эти модели, соответствующим образом обобщенные и измененные, смогут, по-видимому, сыграть важную роль при исследовании широкого круга эпидемических явлений, наблюдаемых в больших популяциях. Однако очевидно, что для изучения более тонких деталей эти модели не подойдут. Так, в рассмотренной выше стохастической модели предполагалось, что не только латентный период равен нулю, но и длительность заразного периода имеет экспоненциальное распределение; для большинства болезней ни одно из этих допущений не справедливо. Для более реалистичного описания биологических и клинических деталей можно было бы построить модели для многофазовых процессов аналогично тому, что было сделано в конце разд. 8.3. Затем для различных интервалов можно выбрать распределения χ^2 , сохраняя при этом марковский характер всего процесса. В определенных случаях оказываются применимыми модели, рассмотренные в разд. 9.5 и 9.6.

9.4. ПОВТОРЯЮЩИЕСЯ ЭПИДЕМИИ

Хорошо известно, что в ряде популяций многие заболевания повторяются более или менее регулярно и в определенном смысле они являются как эндемическими, так и эпидемическими заболеваниями. Например, в некоторых местностях эпидемии кори

циклически повторяются с интервалом в один-два года. Поэтому значительный интерес представляет математическое исследование таких моделей, которые предусматривали бы возможность периодического повторения эпидемий. Как обычно, рассмотрим вначале детерминистскую модель.

Детерминистская модель

Простейшая модификация, которую необходимо ввести в детерминистскую модель, рассмотренную в начале разд. 9.3, состоит в допущении, что число восприимчивых индивидуумов непрерывно пополняется, скажем, со скоростью μ . Это означает, что после какой-либо одной вспышки эпидемии, в результате которой плотность восприимчивых индивидуумов упадет ниже критического значения, наступает период относительного затишья, длящийся до тех пор, пока снова не будет достигнуто критическое значение и не возникнет новая вспышка. В интервале Δt группа восприимчивых индивидуумов, с одной стороны, уменьшается на $\beta xy \Delta t$ за счет заражения части из них, а с другой — увеличивается на $\mu \Delta t$. Допустим далее, что этот приток новых восприимчивых индивидуумов уравновешивается гибелью индивидуумов, удаленных из популяции, и, таким образом, общий объем популяции остается постоянным.

Уравнения, соответствующие системе (9.18), по крайней мере первые два из них (фактически нам и нужны только эти два уравнения), имеют вид

$$\left. \begin{aligned} \frac{dx}{dt} &= -\beta xy + \mu, \\ \frac{dy}{dt} &= \beta xy - \gamma y. \end{aligned} \right\} \quad (9.28)$$

Совершенно очевидно, что в данном случае возможно равновесное состояние (x_0, y_0) . Так, положив в уравнениях (9.28) $dx/dt = 0$ и $dy/dt = 0$, {получаем

$$x_0 = \gamma/\beta, \quad y_0 = \mu/\gamma. \quad (9.29)$$

Небольшие отклонения от равновесного значения можно исследовать обычным способом, записывая

$$x = x_0(1 + u), \quad y = y_0(1 + v), \quad (9.30)$$

где u и v — малые величины.

Подставив выражения (9.30) в уравнения (9.28), получим

$$\left. \begin{aligned} \sigma \frac{du}{dt} &= -(u + v + uv), \\ \tau \frac{dv}{dt} &= u(1 - v), \end{aligned} \right\} \quad (9.31)$$

где $\sigma = \gamma/(\beta\mu)$, $\tau = 1/\gamma$.

Так как u и v малы, произведением uv можно пренебречь; тогда, исключив u из двух уравнений (9.31), получим одно дифференциальное уравнение второго порядка

$$\frac{d^2v}{dt^2} + \frac{1}{\sigma} \frac{dv}{dt} + \frac{v}{\sigma\tau} = 0. \quad (9.32)$$

Уравнение (9.32) имеет следующее решение:

$$\left. \begin{aligned} v &= v_0 e^{-t/2\sigma} \cos \xi t, \\ \xi^2 &= \frac{1}{\sigma\tau} - \frac{1}{4\sigma^2}, \end{aligned} \right\} \quad (9.33)$$

где

а результат для u имеет вид

$$\left. \begin{aligned} u &= v_0 (\tau/\sigma)^{1/2} e^{-t/2\sigma} \cos (\xi t + \psi), \\ \cos \psi &= -\frac{1}{2} (\tau/\sigma)^{1/2}, \quad 0 \leq \psi \leq \pi. \end{aligned} \right\} \quad (9.34)$$

где

Интереснейшая особенность решений (9.33) и (9.34) заключается в том, что они содержат последовательности *затухающих* гармонических колебаний. Исследуя заболеваемость корью в Лондоне, Сопер принял τ равным двум неделям, что примерно совпадает с длительностью инкубационного периода; согласно данным Сопера, период $2\pi/\xi$ составляет около 74 недель, а значение σ — примерно 68. Отсюда следует, что коэффициент затухания между максимумами, равный $\exp(-\pi/\sigma\xi^{1/2}) \approx 0,6$, довольно велик. Следовательно, через несколько циклов колебательный процесс прекратится и будет достигнуто установившееся эндемическое состояние. К сожалению, это совсем *не то*, что встречается на практике; интенсивность повторяющихся эпидемий, несомненно, подвержена статистическим колебаниям, и тем не менее они не затухают в течение многих лет. Исследуем теперь свойства аналогичной стохастической модели и посмотрим, не приведет ли она к более приемлемому описанию наблюдаемых фактов.

Стохастическая модель

Проще всего использовать модель, описанную в предыдущем разделе, и ввести в нее случайный источник новых восприимчивых индивидуумов; тогда вероятность увеличения этой группы в интервале $\Delta(t)$ на одну единицу составит $\mu\Delta t$. Если $X(t)$ и $Y(t)$ — случайные величины, описывающие, как и ранее, соответственно число восприимчивых и зараженных индивидуумов, то имеем три f -функции: $f_{-1,1} = \beta XY$, $f_{0,-1} = \gamma Y$ и $f_{1,0} = \mu$. Таким образом, в этом случае дифференциальное уравнение

в частных производных для производящей функции вероятностей имеет вид

$$\frac{\partial P}{\partial t} = \beta (y^2 - xy) \frac{\partial^2 P}{\partial x \partial y} + \gamma (1 - y) \frac{\partial P}{\partial y} + \mu (x - 1) P \quad (9.35)$$

при начальном условии

$$P(x, y, 0) = x^n y^a. \quad (9.36)$$

Решить уравнение (9.35), а также соответствующую систему обыкновенных дифференциальных уравнений для отдельных вероятностей пока не удалось. Однако на основе некоторых приближенных рассуждений можно показать, к каким общим выводам приводит включение в модель вероятностных элементов. В разд. 9.3 мы уже использовали процесс размножения и гибели для приближенного описания поведения группы зараженных индивидуумов, предполагая, что при достаточно больших значениях n скорости размножения и гибели равны соответственно βn и γ . Допустим сначала, что новые восприимчивые индивидуумы появляются не случайно, а детерминированно с постоянной скоростью μ . Тогда скорость размножения для группы зараженных индивидуумов будет равна $\beta (n + \mu t)$. Таким образом, это соответствует неоднородному процессу размножения и гибели, о котором упоминалось в разд. 8.2, но который детально нами не рассматривался (более подробно этот вопрос изложен в разд. 9.3 другой книги автора [9]).

При $\mu = 0$ имеем полученное ранее пороговое значение. Но если $\mu \neq 0$, то резкого перехода при $\rho = n$ наблюдаться не будет и вероятность прекращения эпидемии будет изменяться непрерывно. Можно показать, что при значениях параметров, приближенно соответствующих тем данным о заболеваемости корью, которые были рассмотрены выше в детерминистской модели, вероятность прекращения эпидемии будет близка к 1. В общих чертах произойдет следующее: после сильной вспышки эпидемии в небольшом коллективе как число восприимчивых индивидуумов, так и число источников инфекции будет мало. В этом случае значение ρ/n для соответствующей группы индивидуумов будет велико, и соответственно будет очень велика вероятность элиминации источников инфекции. Однако популяция восприимчивых индивидуумов всегда непрерывно увеличивается (например, в школах — за счет поступления новых учеников, в целом сообществе — за счет размножения). Поэтому в конце концов наступит момент, когда случайное появление нового источника инфекции может вызвать новую стохастическую эпидемию. После этого весь процесс повторится сначала. Ясно, что при этих условиях будет иметь место последовательность повто-

ряющихся вспышек без затухания и устойчивое равновесие невозможно. Таким образом, стохастическая модель позволяет объяснить наблюдаемые явления, тогда как детерминистская привела к неверным выводам.

Допущение о том, что время от времени в коллектив поступают новые источники инфекции, разумеется, вполне реалистично. Новые источники инфекции могут прибывать из других районов; циркулирующие среди населения непатогенные микроорганизмы вследствие мутации могут становиться вирулентными; заболеваемость в некоторых случаях может возобновиться в результате заражения от некоторых насекомых или животных, являющихся резервуаром инфекции. Математически это требует определенного изменения модели. Можно допустить, например, что вероятность поступления нового источника инфекции в интервале Δt равна $\varepsilon \Delta t$. Построить полную стохастическую модель для этого случая пока еще невозможно, однако можно использовать принятый ранее эвристический подход и получить приближенное распределение длительностей интервала времени между последовательными вспышками эпидемии. Очевидно, что такие модели заслуживают дальнейшего исследования и разработки, однако мы не имеем возможности более детально обсуждать здесь этот вопрос.

9.5. ЦЕПОЧЕЧНО-БИНОМИАЛЬНЫЕ МОДЕЛИ

Хотя рассмотренные выше модели предназначены главным образом для описания эпидемий в больших популяциях, их, безусловно, можно использовать и в случае малочисленных групп, просто приняв размер популяции достаточно малым. Затруднение состоит в том, что при изучении более тонких деталей в такой небольшой группе, как, скажем, семья, допущения о нулевом латентном периоде и экспоненциальном распределении длительностей заразного периода оказываются чрезмерно упрощенными. Поэтому в таких случаях целесообразно остановиться на одной-двух моделях другого типа.

В случае кори, например, обычно допускается, что имеет место очень короткий заразный период, характеризуемый высокой contagiousностью, и инкубационный период почти постоянной длительности. Для теоретических целей сожмем заразный период в точку. Тогда после появления одного первоначального случая заболевания (или нескольких одновременно) дальнейшие случаи будут появляться дискретными группами, отделенными друг от друга инкубационным периодом. Рассмотрим более детально соответствующую модель для дискретного времени (в отличие от изучавшихся ранее моделей для непрерывного времени).

Допустим, что u_t — число восприимчивых индивидуумов в группе непосредственно перед моментом t и что v_t — число индивидуумов, заразившихся в этот момент. Удобно также определить *вероятность достаточного контакта* p ($= 1 - q$), т. е. вероятность того, что любые два члена группы в какой-либо момент времени будут находиться в контакте, достаточном для появления нового источника инфекции, если один из индивидуумов восприимчив к инфекции, а другой является источником инфекции.

Тогда вероятность того, что какой-либо данный восприимчивый индивидуум избежит заражения в момент t , равна q^{v_t} , а вероятность того, что он заразится, — $(1 - q^{v_t})$. Таким образом, условная вероятность появления v_{t+1} новых случаев в момент $t + 1$ имеет биномиальное распределение

$$P \{v_{t+1} | u_t, v_t\} = \left. \begin{aligned} & \binom{u_t}{v_{t+1}} (1 - q^{v_t})^{v_{t+1}} q^{v_t u_{t+1}}, \\ & u_t = u_{t+1} + v_{t+1}. \end{aligned} \right\} \quad (9.37)$$

Итак, этот процесс определяется последовательностью биномиальных распределений, откуда и происходит его название — *цепочно-биномиальный* процесс.

Распределение (9.37) первоначально использовали Рид и Фрост. Более простой вариант, принадлежащий Гринвуду, основан на допущении, что вероятность заражения не меняется при изменении фактического числа источников инфекции, циркулирующих в группе, при условии, что имеется хотя бы один из них. В этом случае вероятность заражения данного индивидуума в какой-либо момент времени равна p , и выражение, соответствующее (9.37), принимает вид

$$P \{v_{t+1} | u_t, v_t\} = \binom{u_t}{v_{t+1}} p^{v_{t+1}} q^{u_{t+1}}. \quad (9.38)$$

С помощью формулы (9.37) или (9.38) не представляет труда записать вероятности наблюдения любой последовательности случаев. К сожалению, общая теория эпидемических процессов для дискретного времени, определяемых таким способом, еще не разработана, но поскольку можно записать вероятность появления любого наблюдаемого значения, то можно хотя бы оценить неизвестные параметры и, если имеется достаточное число степеней свободы, выполнить проверку с помощью критериев значимости. Рассмотрим элементарный пример, используя некоторые данные, полученные во время эпидемии кори в г. Провиденс за 1924—1934 гг.

Любую цепь случаев, начинающуюся в момент $t = 0$, можно описать последовательностью (v_0, v_1, \dots, v_k) , где v_k — послед-

ний отличный от нуля член последовательности значений v_i . В семье, где имеется трое детей и первоначально заболевает один из них, единственно возможными последовательностями являются следующие: (1), (1,1), (1,1,1), (1,2). Мы будем условно записывать их в виде (1), (1²), (1³), (12). Математические ожидания для n таких семей и результаты наблюдений приведены в табл. 6. (В группах такого размера отсутствует различие между вариантом Гринвуда и вариантом Рида и Фроста.)

Таблица 6

Биномиальная цепь для семей с тремя детьми

Тип цепи	Математическое ожидание числа семей	Наблюдаемое число семей	Данные о заболеваемости корью в г. Провиденс	Значения, получаемые путем подбора кривой распределения
(1)	nq^2	a	34	14,9
(1 ²)	$2npq^2$	b	25	23,5
(1 ³)	$2np^2q$	c	36	87,7
(12)	np^2	d	239	207,9
Всего	n	n	334	334,0

Допустим, что n семей распределены между четырьмя типами цепей так, как показано в третьем столбце табл. 6. Метод максимального правдоподобия дает следующую оценку для p и ее дисперсию по большой выборке:

$$\left. \begin{aligned} \hat{p} &= \frac{b + 2c + 2d}{2a + 3b + 3c + 2d}, \\ \sigma^2(\hat{p}) &= \frac{pq}{2n(1 + pq)}. \end{aligned} \right\} \quad (9.39)$$

Подставляя в выражения (9.39) данные, полученные для эпидемии кори в г. Провиденс, получаем следующую оценку: $\hat{p} = 0,789 \pm 0,015$. Приведенные в последнем столбце табл. 6 значения, полученные путем подбора кривой распределения, явно неудовлетворительны. Действительно, величина χ^2 для двух степеней свободы при $P = 0,001$ равна 59,8. Нетрудно выполнить аналогичный анализ и для более крупных семей, однако мы не будем останавливаться здесь на этом вопросе. Для семей, в которых имеется четверо детей, также наблюдается плохое соответствие с данными, полученными в г. Провиденс. Возможно, это обусловлено неоднородностью данных, которые собирались в течение нескольких лет.

Существует, однако, один вид неоднородности, учесть который относительно легко. Предположим, что для различных семей характерны различные значения p и при этом вероятность достаточного контакта внутри любой данной семьи остается постоянной. Это вполне могло иметь место для данных, полученных в г. Провиденс, особенно если они относились к группам, живущим в различных географических и социальных условиях. Для выполнения математического анализа очень удобно допущение о том, что вероятность p следует бета-распределению, т. е.

$$dF = \frac{1}{B(x, y)} p^{x-1} q^{y-1} dp, \tag{9.40}$$

где $x > 0$, $y > 0$, а $B(x, y)$ — бета-функция. Путем соответствующего выбора x и y можно получить ряд распределений, имеющих самую различную форму.

Применим теперь это допущение к математическим ожиданиям, записанным во втором столбце табл. 6, которые должны быть усреднены по распределению (9.40). По чисто техническим статистическим соображениям проводить проверку на значимость и искать оценки, пользуясь непосредственно величинами x и y , очень неудобно. Лучше вместо x и y использовать величины

$$u = \frac{x}{x+y}, \quad z = \frac{y}{x+y}, \tag{9.41}$$

из которых $u (= 1 - v)$ представляет собой среднее значение вероятности p . Результаты, полученные при таком способе анализа, приведены в табл. 7.

Таблица 7

**Биномиальная цепь для семей с тремя детьми
(в различных семьях значение p различно)**

Тип цепи	Математическое ожидание числа семей	Наблюдаемое число семей	Данные о заболеваемости корью в г. Провиденс	Значения, получаемые путем подбора кривой распределения
(1)	$\frac{nv(v+z)}{1+z}$	a	34	34,9
(1 ²)	$\frac{2nvw(v+z)}{(1+z)(1+2z)}$	b	25	22,7
(1 ³)	$\frac{2nvw(u+z)}{(1+z)(1+2z)}$	c	36	37,6
(1 ²)	$\frac{nu(u+z)}{1+z}$	d	239	238,8
Всего	n	n	334	334

Естественно, что получить оценки для u и z сложнее, чем для одного параметра, и в этом случае лучше применить метод максимального правдоподобия, который дает следующие максимально правдоподобные оценки для u и z :

$$\left. \begin{aligned} S_u &= \frac{b+c+d}{u} - \frac{a+b+c}{v} + \frac{c+d}{u+z} - \frac{a+b}{v+z}, \\ S_z &= -\frac{n}{1+z} - \frac{2(b+c)}{1+2z} + \frac{c+d}{u+z} + \frac{a+b}{v+z}. \end{aligned} \right\} \quad (9.42)$$

При этом *наблюдаемые* информационные функции имеют вид

$$\left. \begin{aligned} I_{uu} &= \frac{b+c+d}{u^2} + \frac{a+b+c}{v^2} + \frac{c+d}{(u+z)^2} + \frac{a+b}{(v+z)^2}, \\ I_{uz} &= \frac{c+d}{(u+z)^2} - \frac{(a+b)}{(v+z)^2}, \\ I_{zz} &= -\frac{n}{(1+z)^2} - \frac{4(b+c)}{(1+2z)^2} + \frac{c+d}{(u+z)^2} + \frac{a+b}{(v+z)^2}. \end{aligned} \right\} \quad (9.43)$$

Вычислить эти функции с помощью настольной счетной машины довольно легко, хотя для повторных вычислений необходимы соответствующие начальные значения. Оценки для u и z можно получить методом максимального правдоподобия только на первой стадии эпидемии, т. е. объединяя цепи второго и третьего типов. Полученные оценки имеют вид

$$\hat{u} = \frac{b+c+2d}{2n}, \quad \hat{z} = \frac{4ad-(b+c)^2}{2n(b+c)}. \quad (9.44)$$

Применяя этот метод в целом к данным, полученным в г. Провиденс, находим следующие результаты: $\hat{u} = 0,81 \pm 0,02$, $\hat{z} = 0,74 \pm 0,18$. Цифры, приведенные в последнем столбце табл. 7, говорят о том, что на этот раз соответствие оказывается довольно хорошим: величина χ^2 для одной степени свободы равна 0,32. Интересно заметить, что соответствующие значения x и y равны 1,09 и 0,26, что определяет J-образную форму кривой бета-распределения с бесконечной ординатой при $p = 1$. Полученные результаты являются весьма обнадеживающими. На основе данных о заболеваемости корью в г. Провиденс аналогичные оценки были получены для семей, имеющих четырех детей, и соответствие вновь оказалось вполне удовлетворительным. Конечно, весьма желательно повторение таких исследований на другом материале.

Цепочно-биномиальный метод можно несколько видоизменить и обобщить. Например, колебания длительности инкубационного периода могут помешать точному определению типа цепи, хотя общее число случаев будет зарегистрировано безошибочно. В этой ситуации лучше строить оценки на основе наблюдаемого распре-

деления общего числа случаев. Включая новые параметры в соответствующим образом построенную модель, можно также принять во внимание и ряд других факторов (например, ошибочную диагностику, обусловленную незаметной иммунизацией или потерей постинфекционного иммунитета у ранее болевшего лица). Увеличение числа параметров усложняет анализ в случае малых семей, так как число степеней свободы оказывается недостаточным для получения оценок. Тем не менее стоит всемерно развивать этот подход, так как многие факторы, влияние которых может быть изучено таким способом, имеют существенное значение с биологической и клинической точек зрения. Однако для такого анализа необходимы тщательно собранные данные об эпидемиях в семьях достаточно большого размера.

9.6. ЗАРАЗНЫЙ И ЛАТЕНТНЫЙ ПЕРИОДЫ

Мы только что указали на необходимость более широкого исследования моделей, включающих параметры, имеющие биологическое значение. Как далеко мы сможем продвинуться в этом направлении путем видоизменения цепочечно-биномиального метода, сказать трудно. Другой метод, который стоит обсудить в связи с этой проблемой, — это непосредственное рассмотрение заразного и латентного периодов в предположении, что длительности каждого из них характеризуются определенным распределением. Рассмотрим сначала семьи, в которых имеется только двое детей. После заболевания одного из них (первичный случай) второй может заболеть, а может и не заболеть. Следовательно, часть имеющихся у нас данных будет охватывать семьи только с одним или двумя случаями заболевания. Из данных, касающихся двух случаев, можно получить дополнительную ценную информацию о длительности интервала между этими случаями.

В табл. 8 приведены типичные данные о заболеваемости корью в семьях, где имеется двое детей, полученные в 1946—1952 гг. д-ром Симпсоном в г. Сайренсестере. Всего исследовалось 264 семьи, и в 45 семьях наблюдался только один случай заболевания. В остальных 219 семьях наблюдались два случая, и это дало возможность вычислить длительность интервала между ними. Из таблицы видно, что распределение имеет несколько неожиданную форму. При значении, равном 11 суткам, наблюдается заметный пик с большим рассеянием в обе стороны, т. е. сериальный интервал (и, возможно, инкубационный период) испытывает заметные статистические колебания. Но, кроме того, наблюдается также некоторая концентрация данных вблизи начальной точки распределения с небольшим пиком при значении, равном одним суткам. Этот пик обусловлен скорее всего одновременным появ-

Таблица 8

Интервалы между случаями кори в семьях с двумя детьми

Интервал, сутки	Наблюдаемое число семей	Число двойных первичных случаев	Число одиночных первичных случаев	Значения, получаемые путем подбора кривой распределения
0	5	5		4,67
1	13	13		8,58
2	5	5		6,73
3	4	4		4,53
4	3	2	1	2,78
5	2		2	2,25
6	4		4	3,97
7	11		11	8,85
8	5		5	16,83
9	25		25	24,72
10	37		37	29,44
11	38		38	29,28
12	26		26	25,44
13	12		12	19,99
14	15		15	14,28
15	6		6	9,02
16	3		3	4,82
17	1		1	2,09
18	3		3	0,71
19	.			0,18
20	.			0,04
21	1		1	0,00
Всего	219	29	190	219,00

лением двух первичных случаев в результате контакта с одним и тем же внешним источником. Основная же часть распределения отражает передачу инфекции от одного члена семьи к другому.

Какие изменения необходимо внести в рассмотренные ранее модели, чтобы объяснить данные такого рода? Простейшим обобщением строгой цепочечно-биномиальной модели явилось бы допущение о том, что длительность латентного периода варьирует, тогда как заразный период по-прежнему очень короток. Легко видеть, что в этих условиях сериальный интервал будет следовать распределению длительностей латентного периода. Допустим, что

дисперсия этого распределения равна ν . Тогда интервалы между последовательными случаями заболевания, вызванными двойными первичными случаями, соответствуют абсолютной разности между двумя независимо распределенными латентными периодами. Таким образом, абсолютная величина начального момента равна 2ν . Если в табл. 8 выделить два распределения (принимая какое-либо произвольное решение о перекрывающихся участках, которое не имеет существенного значения), то абсолютная величина начального момента для двойных первичных случаев окажется приближенно равной $\frac{1}{2}\nu$. Таким образом, предлагаемое обобщение является несостоятельным.

Более общим было бы допущение о наличии более продолжительного заразного периода некоторой постоянной длительности (его также можно сделать переменным, однако даже постоянный, но более продолжительный интервал уже сам по себе вносит некоторые дополнительные вероятностные элементы). Допустим, например, что латентный период ξ характеризуется приближенно нормальным распределением с математическим ожиданием μ и дисперсией σ^2 , т. е. что функция $f(\xi)$ имеет вид

$$f(\xi) = (2\pi\sigma^2)^{-1/2} \exp\{-\frac{1}{2}(\xi - \mu)^2 / (2\sigma^2)\}, \quad -\infty < \xi < \infty, \quad (9.45)$$

и что заразный период имеет постоянную длительность α . Предполагается, что вероятность заражения восприимчивого индивидуума в любом интервале Δt заразного периода вплоть до момента заболевания или момента окончания заразного периода равна $\lambda\Delta t$. Это означает, что механизм передачи инфекции представляет собой случайный пуассоновский процесс, происходящий в интервале длительностью α .

Из сказанного выше следует, что распределение абсолютной длительности интервала τ между двойными первичными случаями имеет вид

$$f_-(\omega) = \frac{1}{\sigma\tau^{1/2}} e^{-\omega^2/4\sigma^2}, \quad 0 \leq \omega < \infty. \quad (9.46)$$

Рассмотрим теперь семьи с одним первичным случаем заболевания и последующим вторичным. Пусть τ — промежуток времени между началом заразного периода и моментом возникновения вторичного случая. Так как длительность интервала τ должна иметь экспоненциальное распределение, усеченное в точке $\tau = \alpha$, то соответствующая плотность распределения имеет вид

$$f(\tau) = \lambda e^{-\lambda\tau} (1 - e^{-\lambda\alpha})^{-1}, \quad 0 \leq \tau \leq \alpha. \quad (9.47)$$

Если длительность сериального интервала равна ξ , то очевидно, что $\xi = \xi + \tau$. С помощью выражений (9.45) и (9.47) можно

записать совместное распределение случайных величин ξ и τ . Заменяя ξ на $\zeta - \tau$ и исключая τ путем интегрирования, получаем безусловное распределение случайной величины ζ :

$$f(\zeta) = \frac{\lambda \exp \left\{ -\lambda \left(\zeta - \mu - \frac{1}{2} \lambda \sigma^2 \right) \right\}}{1 - \exp(-\lambda \alpha)} \int_{u'}^u \frac{\exp \left(-\frac{1}{2} t^2 \right) dt}{(2\pi)^{1/2}}, \quad (9.48)$$

где

$$u = \sigma^{-1} \{ \zeta - (\mu + \lambda \sigma^2) \}, \quad u' = u - \alpha \sigma^{-1}. \quad (9.49)$$

Таким образом, распределение (9.46) описывает длительность сериальных интервалов при двойных первичных заболеваниях, а распределение (9.48) — при одном первичном случае. Кроме того, существует еще и третий источник данных — семьи только с одним случаем заболевания. Пусть число таких семей равно C , а число семей как с первичными, так и с вторичными случаями заболевания равно B . Так как вероятность того, что второй восприимчивый член семьи не заболеет в течение заразного периода, равна $e^{-\lambda \alpha}$, вероятность появления наблюдаемых чисел имеет биномиальное распределение с плотностью

$$\binom{B+C}{B} e^{-C\lambda \alpha} (1 - e^{-\lambda \alpha})^B. \quad (9.50)$$

Теперь можно записать вероятность наблюдения B семей при данном значении $B + C$, а также вероятность появления интервалов распределений, представленных в табл. 8.

Используя выражения (9.46), (9.48) и (9.50), можно приступить к отысканию максимально правдоподобных оценок всех четырех параметров λ , α , μ и σ обычным способом, т. е. применяя метод максимального правдоподобия и вычисляя информационные функции. Приближенные исходные значения, необходимые для того, чтобы начать вычисления итерационным методом, можно получить, приравнявая эмпирическое среднее и эмпирическую дисперсию распределения длительности интервала при одном первичном случае, эмпирический второй начальный момент распределения — длительности интервала при двух первичных случаях и наблюдаемое число семей только с одним случаем заболевания — теоретическим значениям соответствующих параметров. Все эти формулы здесь не приводятся. Вычисление максимально правдоподобных оценок на настольной счетной машине требует много времени и очень утомительно; возможно, по этой причине большинство эпидемиологов в прошлом не принимало этот метод всерьез. Однако теперь, при наличии электронных вычислительных машин, эту работу можно спокойно возложить на математиков и программистов. Биологов и медиков интересует лишь окончательный

результат, т. е. фактические оценки с соответствующими средними квадратическими ошибками. Для рассматриваемых здесь данных эти оценки имеют следующие значения:

$$\left. \begin{aligned} \hat{\lambda} &= 0,25 \pm 0,03, \\ \hat{\alpha} &= 6,6 \pm 0,8 \text{ суток}, \\ \hat{\mu} &= 8,6 \pm 0,3 \text{ суток}, \\ \hat{\sigma} &= 1,8 \pm 0,1 \text{ суток}. \end{aligned} \right\} (9.51)$$

Отсюда, в частности, можно найти, что средняя длительность инкубационного периода, равная $\hat{\alpha} + \hat{\mu}$ (эти оценки имеют отрицательную корреляцию), составляет $15,2 \pm 0,5$ суток. Кроме того, можно вычислить критерий согласия и определить, насколько близки наблюдаемые числа к теоретическим значениям. Некоторые группы данных вследствие их малого размера необходимо объединить, однако для интервалов длительностью 1 и 2 недели существует большая вероятность появления систематической ошибки. Поэтому данные для 7 и 14 суток были объединены с данными, относящимися к предыдущему и последующему дням. Это дает удовлетворительное значение χ^2 , равное 12,9 при 8 степенях свободы.

Оценки (9.51) особенно интересны тем, что они дают для кори довольно большой заразный период длительностью почти в неделю, а средняя длительность латентного периода превышает неделю. Таким образом, сумма этих двух величин очень близка к действительной длительности инкубационного периода. Важно иметь в виду, что принятое нами упрощенное представление о почти постоянном латентном периоде и очень коротком заразном периоде совершенно неприемлемо, если сериальные интервалы варьируют так сильно, как в наблюдениях, представленных в табл. 8. Необходимо, конечно, разработать какую-то более совершенную модель. Рассмотренная здесь модель обладает тем достоинством, что довольно хорошо согласуется с некоторыми результатами наблюдений.

Аналогичный анализ можно выполнить и для более крупных семей, хотя с увеличением размера группы задача быстро усложняется. Для групп из трех человек также получены максимально правдоподобные оценки и информационные функции, и весьма обнадеживает то, что при использовании данных о заболеваемости корью, полученных Симпсоном, найденные оценки основных параметров не отличаются значительно от оценок, определяемых по формулам (9.51).

Разумеется, исследования аналогичного характера необходимо проводить в гораздо большем масштабе и для самых различных

заболеваний. Хотя они сопряжены со значительными математическими и вычислительными трудностями, у медиков и биологов нет реальных оснований для серьезного беспокойства по этому поводу. Рассмотренная здесь модель довольно проста, и входящие в нее параметры имеют прямой практический смысл. Поэтому такой количественный анализ мог бы играть значительно более важную роль в эпидемиологических исследованиях, чем он играет сейчас.

9.7. ПРОСТРАНСТВЕННЫЕ МОДЕЛИ

До последнего времени географические факторы, оказывающие существенно важное влияние на распространение заболеваний, исследовались сравнительно мало. Справедливость предположения об однородном перемещивании населения в небольшом городе или деревне уже давно ставилась под сомнение, хотя вполне допустимо в качестве первого приближения принять, что перемещения источников инфекции носят случайный характер и во многом напоминают движение частиц в коллоидном растворе. Тем не менее необходимо, конечно, иметь некоторое представление о том, к какому эффекту может привести наличие большого числа восприимчивых индивидуумов в пунктах, удаленных на довольно большие расстояния от любого данного источника инфекции.

В детерминистской модели, принадлежащей Д. Кендаллу, предполагается существование бесконечного двумерного континуума популяции, в которой на единицу площади приходится σ индивидуумов. Рассмотрим область dS , окружающую точку P , и допустим, что числа восприимчивых, зараженных и удаленных из коллектива индивидуумов равны соответственно $\sigma x dS$, $\sigma y dS$ и $\sigma z dS$. Величины x , y и z могут быть функциями времени и положения, однако их сумма должна равняться единице. Основные уравнения движения, аналогичные системе (9.18), имеют вид

$$\left. \begin{aligned} \frac{\partial x}{\partial t} &= -\beta \sigma x \tilde{y}, \\ \frac{\partial y}{\partial t} &= \beta \sigma x \tilde{y} - \gamma y, \\ \frac{\partial z}{\partial t} &= \gamma y, \end{aligned} \right\} \quad (9.52)$$

где \tilde{y} — пространственно взвешенное среднее значение

$$\tilde{y}(P, t) = \iint \lambda(PQ) y(Q, t) dS. \quad (9.53)$$

Пусть β и γ — постоянные, dS — элемент площади, окружающий точку Q , и $\lambda(PQ)$ — неотрицательный весовой коэффициент.

Допустим, что начальная концентрация заболеваний равномерно распределена в некоторой небольшой области, окружающей первоначальный очаг. Заметим также, что в произведение $\beta\sigma xy$ в явном виде введен множитель σ , с тем чтобы скорость распространения инфекции оставалась независимой от плотности популяции. Если бы y оставалось постоянным на плоскости, то интеграл (9.53) наверняка сходился бы. В этом случае удобно было бы потребовать, чтобы

$$\iint \lambda(PQ) dS = 1. \quad (9.54)$$

Описанная модель позволяет довольно далеко продвинуть математические исследования. Можно показать (с одной-двумя оговорками), что пандемия охватит всю плоскость в том и только в том случае, если плотность популяции превышает пороговое значение $\rho = \gamma/\beta$. Если пандемия возникла, то ее интенсивность ξ определяется единственным положительным корнем уравнения

$$\xi = 1 - e^{-\sigma\xi/\rho}. \quad (9.55)$$

Смысл этого выражения состоит в том, что доля индивидуумов, заболевающих в конце концов в любой области, как бы далеко она ни отстояла от первоначального эпидемического очага, будет не меньше ξ . Очевидно, что эта теорема Кендалла о пороге пандемии аналогична пороговой теореме Кермака и Мак-Кендрика, в которой пространственный фактор не учитывался.

Можно также построить модель для следующего частного случая. Пусть x и y — пространственные плотности восприимчивых и зараженных индивидуумов соответственно. Если считать инфекцию локальной и изотропной, то нетрудно показать, что уравнения, соответствующие первым двум уравнениям системы (9.18), можно записать в виде

$$\left. \begin{aligned} \frac{\partial x}{\partial t} &= -\beta x (y + \alpha \nabla^2 y), \\ \frac{\partial y}{\partial t} &= \beta x (y + \alpha \nabla^2 y) - \gamma y, \end{aligned} \right\} \quad (9.56)$$

где $x \equiv x(\xi, \eta, t)$, $y \equiv y(\xi, \eta, t)$ [ξ и η — пространственные координаты] и

$$\nabla^2 \equiv \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2}.$$

Для начального периода, когда x можно приближенно считать постоянной величиной, второе уравнение системы (9.56) примет вид

$$\frac{\partial y}{\partial t} = Ay + B\nabla^2 y, \quad (9.57)$$

где $A = \beta x - \gamma$, $B = \alpha \beta x$. Это стандартное уравнение диффузии, решение которого имеет вид

$$y = \frac{C}{2Bt} \exp \left(At - \frac{\xi^2 + \eta^2}{4Bt} \right), \quad (9.58)$$

где постоянная C зависит от начальных условий.

Общее число зараженных индивидуумов, находящихся вне круга радиусом R , равно

$$y_R = \int_{\xi^2 + \eta^2 \geq R^2} y d\xi d\eta = 2\pi C \exp \left(At - \frac{R^2}{4Bt} \right). \quad (9.59)$$

Следовательно,

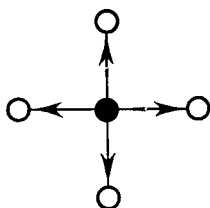
$$R = 2(AB)^{1/2} t \left\{ 1 - \frac{\log(y_R/2\pi C)}{At} \right\}^{1/2},$$

и если $t \rightarrow \infty$, то $R \rightarrow 2(AB)^{1/2} t$. Радиус R , соответствующий какому-либо выбранному значению y_R , растет со скоростью R/t . Эту величину можно рассматривать как *скорость распространения эпидемии*, и ее предельное значение для больших t равно $2(AB)^{1/2}$. В одном из случаев эпидемии кори в Глазго в течение почти полугода скорость распространения составляла около 135 м в неделю.

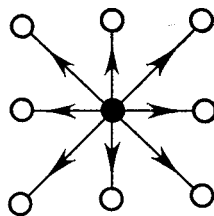
Уравнения (9.56) легко видоизменить так, чтобы была учтена миграция восприимчивых и зараженных индивидуумов, а также появление новых восприимчивых индивидуумов. Как и в случае повторяющихся эпидемий, рассмотренных в разд. 9.4, здесь возможно равновесное решение, однако небольшие колебания затухают столь же быстро или даже быстрее, чем в непространственной модели. Таким образом, ясно, что в данном случае детерминистский подход имеет определенные ограничения. В принципе следовало бы, конечно, предпочесть стохастические модели, но обычно анализ их сопряжен с огромными трудностями, во всяком случае если он проводится чисто математическим путем.

Было выполнено несколько работ по моделированию этих процессов. Так, Барглетт [12] использовал ЭВМ для изучения нескольких последовательных искусственных эпидемий. Пространственный фактор был учтен введением сетки ячеек 6×6 . Внутри каждой ячейки использовались типичные непространственные модели для непрерывного или дискретного времени и допускалась случайная миграция зараженных индивидуумов между ячейками, имеющими общую границу. Была получена информация о критическом объеме популяции, ниже которого происходит затухание эпидемического процесса. Основные параметры модели были получены на основе фактических эпидемиологических и демографических данных.

Недавно автор этой книги предпринял ряд аналогичных исследований, в которых была сделана попытка построить пространственное обобщение стохастических моделей для простого и общего случаев, рассмотренных в разд. 9.2 и 9.3. Допустим, что имеется квадратная решетка, каждый узел которой занят одним восприимчивым индивидуумом. В центре квадрата помещается источник инфекции и рассматривается такой процесс цепочно-биномиального типа для дискретного времени, в котором опасности заражения подвергаются только индивидуумы, непосредственно примыкающие к какому-либо источнику инфекции. Это могут быть либо только четыре ближайших соседа (схема 1), либо также индивидуумы, расположенные по диагонали (схема 2); во втором



С х е м а 1.



С х е м а 2.

случае всего будет восемь индивидуумов, лежащих на сторонах квадрата, центр которого занимает источник инфекции. Очевидно, что выбор схемы произволен, однако в нашей работе использовалось последнее расположение.

Сначала была рассмотрена простая эпидемия без случаев выздоровления. Для удобства использовалась решетка ограниченного размера, и информация о состоянии каждого индивидуума (т. е. восприимчив ли он к инфекции или является ее источником) хранилась в вычислительной машине. В процессе моделирования проводилась текущая запись изменений состояния всех индивидуумов и подсчитывалось общее число новых случаев заболевания во всех квадратах с первоначальным источником инфекции в центре. В памяти машины фиксировались также текущие значения суммы и суммы квадратов числа случаев. Это позволило довольно легко вычислить средние значения и средние квадратические ошибки. Детали этого исследования будут опубликованы в отдельной статье, а здесь мы отметим лишь одну-две частные особенности этой работы. Например, ясно, что при очень высокой вероятности достаточного контакта будет иметь место почти детерминированное распространение эпидемии, при котором на каждом новом этапе развития эпидемии будет добавляться новый квадрат с источниками инфекции. При меньших вероятностях

будет иметь место действительно стохастическое распространение эпидемии. Так как каждый источник инфекции может заразить только восемь своих ближайших соседей, а не всю популяцию, то можно ожидать, что эпидемическая кривая для всей решетки будет возрастать не столь резко, как при однородном перемешивании всей популяции. Этот прогноз действительно оправдывается, и число новых случаев увеличивается с течением времени более или менее *линейно* до тех пор, пока не начнут сказываться краевые эффекты (поскольку решетка имеет ограниченную протяженность).

Таблица 9

**Пространственная стохастическая модель
простой эпидемии, построенная
на решетке 21×21**

Номер этапа эпидемии	Среднее общее число новых случаев	Номер этапа эпидемии	Среднее общее число новых случаев
1	4,4	9	63,5
2	11,8	10	72,2
3	18,6	11	40,1
4	26,7	12	18,1
5	33,1	13	5,1
6	41,2	14	0,6
7	48,5	15	0,2
8	56,0		

В табл. 9 приведены результаты, полученные для решетки 21×21 при наличии одного исходного источника инфекции и вероятности достаточного контакта, равной 0,6. Можно видеть, что между первым и десятым этапами эпидемии среднее число новых случаев каждый раз увеличивается примерно на 7,5. После этого начинает преобладать краевой эффект, и эпидемическая кривая резко падает вниз.

Можно также определить среднее число новых случаев для любой данной точки решетки и найти таким образом эпидемическую кривую для этой точки. Удобно проводить усреднение по всем точкам, лежащим на границе квадрата, в центре которого находится источник инфекции, хотя симметрия в этом случае не будет полной. Сравнение результатов для квадратов различного размера дает картину эпидемической волны, движущейся от первоначального источника инфекции. Здесь мы имеем после-

довательность распределений, моды которых увеличиваются в линейной прогрессии, а дисперсия непрерывно возрастает.

Было также выполнено более детальное исследование эпидемии общего типа с удалением зараженных индивидуумов. Безусловно, все это очень упрощенные модели. Однако важно понять, что они могут быть значительно усовершенствованы. Чтобы учесть мобильность популяции, надо допустить, что восприимчивые индивидуумы заражаются и от тех источников инфекции, которые не являются их ближайшими соседями. Возможно, здесь придется использовать какой-то весовой коэффициент, зависящий от расстояния. Видоизменения, которые нужно будет ввести при этом в программу вычислительной машины, сравнительно невелики. На следующем этапе, возможно, удастся описать таким способом реальные или типичные популяции с самой разнообразной структурой. Это откроет возможность оценивать эпидемиологическое состояние реальных популяций с точки зрения опасности возникновения эпидемий различного типа.

Глава 10

ГЕНЕТИЧЕСКОЕ СЦЕПЛЕНИЕ И ХРОМОСОМНЫЕ КАРТЫ

10.1. ВВЕДЕНИЕ

Многие явления, связанные с биологической наследственностью, привлекали внимание человека с незапамятных времен, однако реальная основа для научного исследования этих явлений возникла лишь во второй половине XIX в., когда появилась работа Грегора Менделя, проведенная на садовом горохе. Выводы Менделя впервые были опубликованы в 1866 г.; по существу он предложил атомистическую модель наследственности, согласно которой единицы, обуславливающие наследственные признаки, передаются без изменения от поколения к поколению. Впоследствии на основе совершенно элементарных комбинаторных соображений были выведены закон расщепления и закон независимого распределения генов. Поскольку все это, во всяком случае на начальных этапах, не требовало никакой сложной математики, интересно поразмыслить, почему подобные идеи не возникли у какого-нибудь смелого мыслителя много веков назад. Возможно, это объясняется тем, что всякая новая теоретическая концепция получает прочную основу лишь в том случае, если она наглядно подтверждается экспериментально наблюдаемыми фактами. В разд. 3.2 мы уже обращали внимание на опасность широких теоретических построений, которые невозможно надлежащим образом проверить. Разумеется, теперь мы знаем, что идеи Менделя были в основе своей верны. Однако они не смогли привлечь к себе должного внимания при его жизни, так как среди известных в то время биологических механизмов не было ни одного, которому бы они соответствовали. Лишь после того как в 80-е и 90-е годы прошлого столетия были достигнуты значительные успехи в области цитологии, настало время для вторичного открытия законов Менделя. Это произошло в 1900 г., и с тех пор наука о наследственности развивается бурно и непрерывно.

Современная генетика, несомненно, является областью биологии с наиболее сильно развитой математической основой. Генетические исследования имеют широкий диапазон — от изучения макроскопических явлений на уровне популяции, в частности естественного отбора и эволюции, до изучения субмикроскопического мира на уровне молекул дезоксирибонуклеиновой кисло-

ты (ДНК). Математические модели можно построить для всех уровней исследования, и между моделями, используемыми на различных уровнях, уже существует довольно большая связь, хотя еще и не столь полная, как хотелось бы. В предыдущей главе, посвященной теории эпидемий, мы видели, что в этой области наметились лишь первые шаги в установлении связи между тонкой структурой и явлениями, протекающими в популяции. В генетике дело обстоит значительно лучше. Из обширной области математической генетики мы кратко обсудим здесь интерпретацию результатов экспериментальных исследований генетического сцепления и попытки объяснения макроскопических явлений с помощью моделей, описывающих микроскопическую (хотя и не молекулярную) структуру хромосом. Все эти вопросы более подробно рассматриваются в другой книге автора [7], в которой читатель может найти интересующие его детали.

Прежде всего нам, по-видимому, следует рассмотреть некоторые основные понятия генетики. Они изложены здесь в упрощенной форме, с тем чтобы ясно показать, в каком направлении нужно делать первые шаги для построения соответствующих математических моделей. На более высоком уровне возможно создание моделей, которые гораздо ближе к реальным биологическим явлениям.

Основная концепция генетики состоит в том, что наследственность по существу обусловлена очень малыми материальными частицами, называемыми *генами*, которые передаются от одного поколения к другому более или менее неизменными и обуславливают реально наблюдаемые наследственные признаки. Обычно гены встречаются парами, и у любого данного вида имеется большое число характерных для него пар генов. Допустим, что гены определенной пары встречаются в двух различных формах, или *аллелях*, A и a . Индивидуумы, несущие пару AA или пару aa , называются *гомозиготными* по рассматриваемым генам, а индивидуумы, несущие пару Aa , — *гетерозиготными*. Если рассматривать только одну пару аллелей, то возможны три разные генетические конституции, или три разных *генотипа*: AA , Aa и aa . Реальное же проявление, или *фенотип*, зависит от характера действия генов. Если ген A — *доминантный*, а ген a — *рецессивный*, то возможны лишь два альтернативных признака: признак a , обусловленный парой aa , и признак A , обусловленный парами AA или Aa . В данном случае генотипы AA и Aa фенотипически неразличимы, но их удастся различить с помощью экспериментов по скрещиванию. Иногда явление доминантности и рецессивности отсутствует, как в случае группы крови MN у человека. В данном случае имеются три генотипа: MM , MN и NN , каждый из которых соответствует определенному, отличному от других фенотипу.

Во всех обычных (соматических) клетках взрослого организма присутствуют оба гена данной пары, тогда как половые клетки, или *гаметы*, содержат только по одному случайным образом отобранному представителю каждой пары. Если ген встречается в виде двух различных аллелей A и a , то возможны два типа яйцеклеток и сперматозоидов — с аллелем A и с аллелем a . Если же взрослый индивидуум является гомозиготным, он может производить гаметы только одного типа. Так, при генотипе AA возможны только гаметы A , а при генотипе aa — только гаметы a . Однако гетерозиготный индивидуум с генотипом Aa производит и гаметы A , и гаметы a , причем примерно в одинаковых количествах. Пары аллелей восстанавливаются при оплодотворении в результате соединения отдельных генов, находящихся в яйцеклетке и сперматозоиде.

После этих предварительных замечаний можно изложить первый закон Менделя, или, как его теперь называют, *закон расщепления*. Наблюдаемые признаки контролируются парами генов, которые разделяются, или *расщепляются*, при образовании гамет и воссоединяются в следующем поколении. Таким образом, при расщеплении происходит перетасовка генов, имевшихся у родительских особей, и перераспределение их у потомков. Сами признаки также в известной мере расщепляются, что частично зависит от доминантности или рецессивности генов, а частично — от реальных генотипов родительских особей. Рассмотрим *скрещивание между двумя гибридами с генотипами Aa* . Каждый из них производит гаметы A и a в одинаковых количествах. Вероятность того, что оплодотворенная яйцеклетка получит ген A по женской линии, равна $1/2$, и вероятность того, что она получит ген A по мужской линии, также равна $1/2$. Таким образом, вероятность того, что потомство будет относиться к генотипу AA , равна $1/4$. Аналогичным образом вероятность появления генотипа aa также равна $1/4$. Далее, ген A может быть получен по женской, а ген a — по мужской линии, и наоборот. Таким образом, вероятность появления генотипа Aa равна $1/2$. Итак, ожидаемое соотношение генотипов AA , Aa и aa составит $1 : 2 : 1$. Если ген A доминирует над геном a , то генотипы AA и Aa неразличимы и имеют фенотип A . Поэтому можно ожидать, что фенотипы A и a будут встречаться в соотношении $3 : 1$. Существует несколько других типов скрещивания, в частности *возвратное скрещивание $Aa \times aa$* , при котором соотношение фенотипов A и a будет равно $1 : 1$. В каждом случае соотношение фенотипов при расщеплении можно найти простым комбинаторным методом, как это показано выше для скрещивания между гибридами.

До сих пор мы рассматривали элементарный случай только одной пары генов. Однако гораздо больший интерес представляет

изучение поведения одновременно двух пар генов или большего их числа. Второй закон Менделя, или *закон независимого распределения*, гласит, что при одновременном расщеплении двух или большего числа пар генов каждая пара расщепляется независимо от других. Оказалось, что из этого закона имеются некоторые очень важные исключения, но в тех случаях, когда он действует, можно объяснить совместное расщепление следующим образом. Рассмотрим две пары генов A, a и B, b . Чего можно ожидать при дигибридном скрещивании $AaBb \times AaBb$? Очевидно, что каждая родительская особь производит гаметы четырех типов — AB, Ab, aB и ab — в равных количествах. Следовательно, любая гамета от каждой родительской особи может участвовать в 16 комбинациях, однако не все эти комбинации действительно различны. Исследование их показывает, что только 9 комбинаций из 16 отличаются друг от друга. Более того, если гены A и B доминантны относительно генов a и b соответственно, то возможны только четыре класса фенотипов — AB, Ab, aB и ab . Легко найти, что соответствующее соотношение при расщеплении составляет $9 : 3 : 3 : 1$.

Исключения из второго закона Менделя возникают вследствие того, что гены расположены в линейном порядке в микроскопических структурах, называемых *хромосомами*. Обычно хромосомы располагаются в клетках организма гомологичными парами, причем общее число пар для данного вида постоянно. Каждая пара генов занимает определенное место, или *локус*, в определенной хромосоме. Различные хромосомы распределяются независимо друг от друга; гены, лежащие в разных хромосомах, также распределяются независимо друг от друга. Однако те пары генов, локусы которых лежат в одних и тех же хромосомах, связаны друг с другом, причем эта связь не абсолютная, а зависит от расстояния между локусами. Это явление называется *сцеплением генов*.

Рассмотрим образование индивидуума, гетерозиготного по двум парам аллелей $AaBb$, причем соответствующие два локуса сцеплены. Здесь имеются только две гомологичные хромосомы, по одной от каждой из родительских особей. Если одна хромосома содержит гены A и B , а другая — гены a и b , то генотип можно обозначить как AB/ab , где косая черта разделяет две хромосомы. При этом говорят, что скрещивание находится в *фазе сцепления*. Возможен другой случай, когда гены A и B лежат в различных гомологичных хромосомах, т. е. случай Ab/aB . Такое скрещивание находится в *фазе отталкивания*. Причина того, что обычно сцепление не является полным, состоит в следующем.

При образовании гамет происходит обмен генетическим материалом между сближившимися (конъюгировавшими) гомологичными хромосомами, в которых гомологичные локусы расположены

в одинаковой последовательности. Точнее, происходит разрыв хромосом, после чего, обменявшись участками, хромосомы вновь воссоединяются. Этот процесс обмена называют также *кроссинговером*. Так, если одна из родительских особей имеет генотип AB/ab , то гаметы AB и ab образуются при четном числе точек обмена между двумя локусами, а гаметы Ab и aB — при нечетном числе таких точек. В последнем случае возникают типы хромосом, или *рекомбинанты*, которых нет у родительских особей. Если доля рекомбинантных хромосом, или *частота рекомбинаций*, равна y , то очевидно, что четыре типа гамет AB , Ab , aB и ab должны появиться в следующем соотношении: $1/2(1 - y) : 1/2y : 1/2y : 1/2(1 - y)$. Если одна из родительских особей несет сцепленные доминантные аллели AB/ab , а другая — сцепленные рецессивные аллели ab/ab (y последней все гаметы будут ab независимо от того, имел ли место кроссинговер), то потомство будет распределяться по классам фенотипов AB , Ab , aB и ab точно в таком же соотношении. Теперь можно оценить частоту рекомбинаций непосредственно по соотношению типов рекомбинантов у потомства. Заметим, что если y гетерозиготной родительской особи эти факторы находятся в фазе отталкивания, то частота рекомбинаций будет измеряться по частоте родительских (а не рекомбинантных) фенотипов в потомстве.

Очевидно, что если $y = 1/2$, то наблюдается независимое расщепление, или отсутствие сцепления (если не считать того, что значения, близкие к $1/2$, возможны также для локусов, лежащих в одной хромосоме, но на большом расстоянии друг от друга). Частота обменов между двумя локусами служит показателем расстояния между ними: она мала, если локусы расположены близко друг к другу, и велика, если локусы удалены друг от друга. Удобно определять *расстояние на хромосомной карте* между любыми двумя локусами как среднее число точек обмена, находящихся на данном участке хромосомы. Это расстояние подчиняется закону аддитивности, даже если распределение точек обмена вдоль хромосомы не является случайным. Отклонение от чисто случайного распределения называется *интерференцией*. Так как расстояние на хромосомной карте равно среднему числу точек обмена, а частота рекомбинаций — частоте появления нечетного числа точек обмена, то ясно, что расстояния на хромосомных картах в отличие от частоты рекомбинаций нельзя наблюдать непосредственно. В отсутствие интерференции между расстоянием на хромосомной карте и частотой рекомбинаций существует простое соотношение, однако при многих типах интерференции это соотношение оказывается более тонким, что требует введения специальных *метрик*, объясняющих как явление интерференции, так и рекомбинацию.

В одном очень важном аспекте приведенное здесь описание кроссинговера чрезмерно упрощено. Рассматривая обмен генетическим материалом только между двумя гомологичными хромосомами, мы подразумеваем *модель двух нитей*. На самом же деле каждая хромосома продольно расщепляется на две хроматиды, которые остаются связанными лишь в одном специфическом локусе при помощи структуры, называемой *центромерой*. Таким образом, нам следовало бы рассматривать *модель четырех нитей*, в которой кроссинговер проходит между хроматидами различных хромосом гомологичной пары. Можно показать, что при элементарном рассмотрении, без учета интерференции, модель двух нитей вполне удовлетворительна; если же рассматривается интерференция, то следует обратиться к модели четырех нитей.

Хотя мы допускаем, что определенные выше расстояния на хромосомных картах в известной мере отражают реальные физические расстояния между локусами, однако это соответствие нельзя считать абсолютным, поскольку хромосомы могут иметь различную способность к кроссинговеру. Тем не менее построение хромосомных карт служит одним из очень важных способов топографического описания генетического материала; кроме того, при этом становится возможным математический анализ и формулировка гипотез, которые можно проверить экспериментально. Экспериментальная проверка сопряжена со многими трудностями, так как в большинстве случаев гены и локусы невозможно наблюдать непосредственно. Проверка гипотез по большей части основана на выявлении их внутренней логичности и возможности использовать их для точного предсказания, хотя в некоторых случаях возможны приближенные сравнения между длиной генетической карты целых хромосом и наблюдаемыми физическими расстояниями, как, например, в случае плодовой мушки *Drosophila melanogaster*.

10.2. ПРОСТЫЕ ЭКСПЕРИМЕНТЫ СО СЦЕПЛЕНИЕМ

Рассмотрим теперь более подробно некоторые проблемы, связанные с оценкой частоты рекомбинаций. Последнюю можно вывести довольно непосредственно на основании результатов нескольких различных скрещиваний, важная особенность которых состоит в том, что на образование гамет хотя бы у одной из родительских особей влияет кроссинговер. В предыдущем разделе мы видели, что нужную информацию может дать простое дигибридное возвратное скрещивание. Однако на практике возникает ряд трудностей. Наиболее распространенная из них — нарушение простых соотношений $1 : 1$ и $3 : 1$, обусловленное разным выживанием различных гамет или различных оплодотворенных яйце

клеток (так называемая *дифференциальная жизнеспособность*). Нас прежде всего интересует частота рекомбинаций, однако мы не сможем точно определить эту частоту, если модель, выбранная для описания всего процесса, не учитывает должным образом случайно возникающих осложнений. Поэтому необходим тщательный статистический анализ, позволяющий определить, значимы ли такие эффекты. Если эти эффекты значимы, то их необходимо учитывать. Более полно материал данного раздела изложен в гл. 3 и 4 другой книги автора [7].

В общем случае возвратное скрещивание дает гораздо больше информации, чем скрещивание между гибридами, если оба фактора обнаруживают доминантность, однако ситуация становится прямо противоположной, если доминирование отсутствует и гетерозиготы можно отличить от обеих гомозигот. Поэтому для получения хороших результатов необходимо тщательное планирование эксперимента. В дальнейшем изложении мы примем, что гены *A* и *B* доминируют над своими аллелями, и для иллюстрации будем использовать главным образом анализ экспериментов по возвратному скрещиванию.

В табл. 10 приведены значения математического ожидания числа потомков, относящихся к различным классам фенотипов, при дигибридном возвратном скрещивании, находящемся в фазе сцепления и в фазе отталкивания.

Таблица 10

Дигибридное возвратное скрещивание со сцеплением

Классы фенотипов	AB	Ab	aB	ab	Сумма
Математическое ожидание для $AB/ab \times ab/ab$	$\frac{1}{2} n (1-y)$	$\frac{1}{2} ny$	$\frac{1}{2} ny$	$\frac{1}{2} n (1-y)$	n
Математическое ожидание для $AB/aB \times ab/ab$	$\frac{1}{2} ny$	$\frac{1}{2} n (1-y)$	$\frac{1}{2} n (1-y)$	$\frac{1}{2} ny$	n
Наблюдаемое число	a	b	c	d	n

Очень важно представить эти результаты в таком виде, чтобы можно было простейшим путем найти максимально правдоподобные оценки рассматриваемых параметров. Из данного примера

легко видеть, что при отсутствии отклонений в жизнеспособности соответствующая оценка частоты рекомбинаций y имеет вид

$$\left. \begin{aligned} \hat{y} &= \frac{b+c}{n} \quad \text{для фазы сцепления} \\ \text{или} \\ \hat{y} &= \frac{a+d}{n} \quad \text{для фазы отталкивания.} \end{aligned} \right\} \quad (10.1)$$

Оценка дисперсии в обоих случаях имеет вид

$$\sigma^2(\hat{y}) = \frac{\hat{y}(1-\hat{y})}{n}. \quad (10.2)$$

Разумеется, можно начать с того, чтобы проверить, имеется ли статистически значимое отклонение от простейшей гипотезы, предполагающей отсутствие влияния рекомбинаций или различий в жизнеспособности. При $y = 1/2$ можно ожидать, что все четыре класса фенотипов, рассматриваемые в табл. 10, будут встречаться одинаково часто. Легко показать, что критерий χ^2 с тремя степенями свободы можно разбить на три независимые компоненты: χ_A^2 — для проверки соотношения при расщеплении аллелей пары Aa ; χ_B^2 — для проверки соотношения при расщеплении аллелей пары Bb ; χ_L^2 — для проверки на сцепление:

$$\left. \begin{aligned} \chi_A^2 &= \frac{(a+b-c-d)^2}{n}, \\ \chi_B^2 &= \frac{(a-b+c-d)^2}{n}, \\ \chi_L^2 &= \frac{(a-b-c+d)^2}{n}. \end{aligned} \right\} \quad (10.3)$$

Каким именно критерием для проверки значимости следует пользоваться, зависит от того, какой информацией мы располагаем вначале. Если известно, что лишь одно из соотношений отклоняется от ожидаемого, то можно использовать критерий χ_L^2 для проверки существования сцепления, но если известно, что отклоняются оба соотношения, то нужно использовать описанный ниже другой критерий.

Допустим далее, что имеются три пары локусов и соответствующие пары генов $A, a; B, b$ и C, c . Сцепление между генами A и B (с частотой рекомбинаций y_1) можно обнаружить с помощью возвратных скрещиваний, затрагивающих только эти два фактора; сцепление между генами B и C (с частотой рекомбинаций y_2) можно обнаружить с помощью другой группы экспериментов, в которых затрагиваются только эти факторы. Что можно сказать

о порядке расположения этих трех локусов? Ничего, если у нас нет данных о наличии сцепления между генами A и C . Таким образом, нам нужна третья группа скрещиваний. Положим, что соответствующая частота рекомбинаций равна y_3 .

Допустим теперь, что истинный порядок расположения локусов — ABC . Легко показать, что между генами A и C произойдет рекомбинация, если она произойдет между генами A и B , но не между генами B и C , и наоборот. Если допустить, что рекомбинация в соответствующих двух участках происходит независимо, то y_3 находится по формуле Трой:

$$y_3 = y_1(1 - y_2) + (1 - y_1)y_2 = y_1 + y_2 - 2y_1y_2. \quad (10.4)$$

Очевидно, что это правило для частот множественных рекомбинаций не является аддитивным (чего следовало бы ожидать на основе материала, рассмотренного в конце предыдущего раздела). Если данных не очень много, а статистическая изменчивость велика, то нелегко определить, будет ли соотношение (10.4) более вероятно, чем два других, соответствующих случаям ACB и BAC . Если эти три локуса располагаются близко друг к другу и все три частоты рекомбинаций малы, то справедливы следующие приближенные равенства:

$$\left. \begin{aligned} y_3 &\approx y_1 + y_2, \text{ порядок } ABC; \\ y_2 &\approx y_1 + y_3, \text{ порядок } BAC; \\ y_1 &\approx y_2 + y_3, \text{ порядок } ACB. \end{aligned} \right\} \quad (10.5)$$

Однако лучше определять порядок расположения локусов на потомстве, полученном от тригибридного возвратного скрещивания типа $ABC/abc \times abc/abc$. Это тригибридное скрещивание в фазе сцепления. Не представляет труда вычислить математические ожидания числа потомков с различными фенотипами. Для порядка ABC эти математические ожидания приведены в табл. 11.

Таблица 11

Тригибридное возвратное скрещивание со сцеплением

Классы фенотипов	ABC, abc	aBC Abc	ABc, aBc	aBc, AbC	Всего
Математическое ожидание для $ABC/abc \times abc/abc$	$n(1-y_1)(1-y_2)$	$ny_1(1-y_2)$	$n(1-y_1)y_2$	ny_1y_2	n
Наблюдаемое число	a	b	c	d	n

Заметим, что мы объединили здесь комплементарные классы, т. е. классы типа ABC и abc , для которых математические ожидания одинаковы. Аналогичным способом можно получить соответствующие результаты и для трех других типов скрещивания. Максимально правдоподобные оценки для y_1 и y_2 , а также оценки их дисперсий в этом случае имеют вид

$$\left. \begin{aligned} \hat{y}_1 &= \frac{b+d}{n}, & \sigma^2(\hat{y}_1) &= \frac{\hat{y}_1(1-\hat{y}_1)}{n}; \\ \hat{y}_2 &= \frac{c+d}{n}, & \sigma^2(\hat{y}_2) &= \frac{\hat{y}_2(1-\hat{y}_2)}{n}. \end{aligned} \right\} \quad (10.6)$$

Если наше предположение о том, что ABC — истинный порядок следования локусов, верно, то число потомков с фенотипом (aBc , AbC) будет наименьшим; оно имеет математическое ожидание ny_1y_2 и отражает рекомбинацию в обоих участках AB и BC одновременно. Обычно при тригибридном возвратном скрещивании нетрудно установить таким способом наименее многочисленный класс фенотипов и определить вероятный порядок расположения локусов. Так, если окажется, что наименьшее число потомков относится к классу (ABc , abc), то искомым порядком локусов должен быть порядок ACB .

Существуют, однако, и некоторые другие преимущества применения полигибридного возвратного скрещивания по сравнению с попытками комбинировать результаты, полученные в нескольких отдельных экспериментах с дигибридным скрещиванием. Эти преимущества станут ясны, если учесть осложнения, связанные, в частности, с дифференциальной жизнеспособностью, которую мы сейчас и рассмотрим для случая простого дигибридного возвратного скрещивания; более сложные эксперименты с тригибридным скрещиванием мы обсудим в разд. 10.3.

Дифференциальная жизнеспособность

Очень часто различные генотипы или фенотипы характеризуются различной жизнеспособностью, и во многих случаях важно учитывать этот фактор. Допустим, что относительная жизнеспособность фенотипа A по сравнению с фенотипом a равна u и что соответствующая величина для фенотипа B по сравнению с фенотипом b равна v . Если величины u и v независимы друг от друга и действуют мультипликативно, то математические ожидания числа потомков с различными фенотипами при дигибридном возвратном скрещивании в фазе сцепления, представленные в табл. 10, изменятся и примут значения, приведенные в табл. 12,

где делитель D определяется формулой

$$D = uv(1-y) + uy + vy + (1-y). \quad (10.7)$$

Эта таблица имеет три степени свободы, что позволяет оценить три неизвестных; легко найти, что соответствующие максимально правдоподобные оценки выражаются следующим образом:

$$\left. \begin{aligned} \hat{y} &= \frac{(bc)^{1/2}}{(ad)^{1/2} + (bc)^{1/2}}, & \sigma^2(\hat{y}) &= \frac{\hat{y}^2(1-\hat{y})^2}{h}; \\ \hat{u} &= \left(\frac{ab}{cd}\right)^{1/2}, & \sigma^2(\hat{u}) &= \frac{\hat{u}^2}{h}; \\ \hat{v} &= \left(\frac{ac}{bd}\right)^{1/2}, & \sigma^2(\hat{v}) &= \frac{\hat{v}^2}{h}, \end{aligned} \right\} \quad (10.8)$$

где

$$\frac{4}{h} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}.$$

Если имеются данные только по скрещиванию в фазе сцепления (или только в фазе отталкивания), то ни одной степени свободы для проверки модели у нас не остается; но если мы располагаем

Таблица 12

Дигибридное возвратное скрещивание в фазе сцепления при наличии дифференциальной жизнеспособности

Классы фенотипов	AB	Ab	aB	ab	Сум- ма
Математическое ожидание для $AB/ab \times ab/ab$	$\frac{nuv(1-y)}{D}$	$\frac{nuy}{D}$	$\frac{nvu}{D}$	$\frac{n(1-y)}{D}$	n
Наблюдаемое число	a	b	c	d	n

данными по обоим фазам, то можно произвести проверку с помощью критерия χ^2 .

Нетрудно провести проверку на сцепление, допустив, что оба моногибридных расщепления нарушены. Так, если мы принимаем нулевую гипотезу $y = \frac{1}{2}$, то из табл. 12 видно, что приведенные в ней математические ожидания соответствуют тем, которые получаются из таблицы сопряженности признаков 2×2

с независимыми рядами и столбцами. Эта таблица имеет вид

$$\begin{array}{c|c} a & b \\ \hline c & d \end{array} \quad (10.9)$$

В этом случае можно применять обычный критерий χ^2 или точный критерий значимости.

Однако вполне возможно, что величины u и v не будут независимы друг от друга, и поэтому лучше принять допущение

Таблица 13

Сбалансированное возвратное скрещивание при наличии дифференциальной жизнеспособности

Классы фенотипов	AB, ab	Ab, aB	Сумма
Математическое ожидание для $Ab/ab \times ab/ab$	$\frac{n_1 \alpha (1-y)}{\alpha (1-y) + y}$	$\frac{n_1 y}{\alpha (1-y) + y}$	n_1
Наблюдаемое число	a	b	n_1
Математическое ожидание для $Ab/aB \times ab/ab$	$\frac{n_2 \alpha y}{\alpha y + (1-y)}$	$\frac{n_2 (1-y)}{\alpha y + (1-y)}$	n_2
Наблюдаемое число	c	d	n_2

о том, что относительная жизнеспособность двух классов родительских особей AB и ab по сравнению с жизнеспособностью классов рекомбинантов Ab и aB равна, скажем, α . В этом случае важно иметь данные по скрещиванию как в фазе сцепления, так и в фазе отгалкивания (сбалансированное скрещивание). Допустим, что получены величины, приведенные в табл. 13.

На этот раз нужно получить оценки лишь для двух параметров, и имеются как раз две степени свободы — по одной для данных по каждой фазе. Максимально правдоподобные оценки имеют вид

$$\left. \begin{aligned} \hat{y} &= \frac{(bc)^{1/2}}{(ad)^{1/2} + (bc)^{1/2}}, & \sigma^2(\hat{y}) &= \frac{\hat{y}^2 (1-\hat{y})^2}{h}; \\ \hat{\alpha} &= \left(\frac{ac}{bd}\right)^{1/2}, & \sigma^2(\hat{\alpha}) &= \frac{\hat{\alpha}^2}{h}, \end{aligned} \right\} \quad (10.10)$$

где

$$\frac{4}{h} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}.$$

Формально результаты (10.8) и (10.10) очень похожи, хотя они имеют различную интерпретацию.

В качестве примера рассмотрим приведенные в табл. 14 данные о сцеплении для рецессивного признака *fidget* (*fi*) и доминантного признака *Danforth's short tail* (*Sd*) у домового мыши. Заметим, что во всех случаях знак + обозначает нормальный аллель.

Таблица 14

Данные о сцеплении между генами *fi* и *Sd* у домового мыши

Классы фенотипов	+ Sd	++	fi Sd	fi +	Сумма
$\frac{+ Sd}{fi +} \times \frac{fi +}{fi +}$ (фаза сцепления)	411	31	25	412	279
$\frac{+ +}{fi Sd} \times \frac{fi +}{fi +}$ (фаза оттаивания)	32	141	73	25	271

Этот материал довольно типичен для реальных данных, получаемых при таких нарушениях соотношений расщепления, которые обусловлены скорее всего дифференциальной жизнеспособностью и сильным сцеплением между этими двумя локусами. С помощью формулы (10.10) легко найти, что оценка для частоты рекомбинаций между локусами *fi* и *Sd* составляет $(20,5 \pm 1,7)\%$.

Аналогичный анализ можно провести для многих других видов скрещивания, например для скрещивания между гибридами или комбинации его с возвратным скрещиванием, однако обычно все же предпочтительнее рассматривать полигибридное возвратное скрещивание с участием двух или большего числа факторов. Кроме того, можно принимать во внимание и другие эффекты, например неполную пенетрантность, которая наблюдается, когда у некоторой доли особей ожидаемые наследственные признаки не проявляются. Когда мы говорим, что ген *A* доминантен, это означает, что при фенотипе *Aa* проявится доминантный признак. Однако при неполной пенетрантности этот признак может не проявиться, и особь с фенотипом *Aa* будет ошибочно принята за особь с генотипом *aa*. Такого рода усложнения в принципе всегда можно учесть, вводя в модель новые параметры. Однако при слишком большом числе параметров может случиться, что их не удастся оценить из-за нехватки числа степеней свободы. Для выявления и оценки влияния различных факторов важнейшее значение имеет тщательное планирование эксперимента.

10.3. СБАЛАНСИРОВАННОЕ ТРИГИБРИДНОЕ ВОЗВРАТНОЕ СКРЕЩИВАНИЕ

В предыдущем разделе мы убедились в значительных преимуществах проведения анализа на сцепление двух факторов методом сбалансированного возвратного скрещивания, поскольку здесь требуются менее строгие допущения о характере влияния дифференциальной жизнеспособности. Если рассматривается порядок расположения нескольких локусов, то в экспериментах с тригибридным возвратным скрещиванием легче определить истинный порядок их расположения, чем при изучении трех отдельных дигибридных возвратных скрещиваний (хотя мы рассматривали лишь тот случай, когда влияние дифференциальной жизнеспособности отсутствует). Существуют также и некоторые другие аргументы в пользу тригибридных (или даже полигибридных) скрещиваний. Так, хотя частота рекомбинаций и является довольно постоянной величиной, у разных пород животных и в разных условиях среды она может слегка варьировать. Поэтому лучше исследовать как можно больше параметров в пределах одного эксперимента. Кроме того, возникают большие сложности, когда точки обмена не распределены вдоль хромосомы случайным образом, т. е. имеет место интерференция. В этом случае соотношение (10.4) уже не справедливо и труднее описать частоты различных типов рекомбинаций; поэтому здесь необходимо проводить эксперименты с использованием нескольких сцепленных факторов. Многие доводы в пользу такого рода экспериментов очень близки к доводам в пользу факторных экспериментов при агрономических полевых испытаниях или других, по смыслу аналогичных исследованиях.

Полный математический анализ даже тригибридного возвратного скрещивания сопряжен с рядом трудностей (см. другую книгу автора [7], разд. 4.5), хотя интерпретация получаемых результатов с точки зрения генетики вполне возможна. Рассмотрим оценку параметров для одной модели по методу Р. Фишера. Как и ранее, допустим, что имеются три пары локусов A, a , B, b и C, c , и предположим, что истинным порядком локусов является ABC . При образовании гамет рекомбинации могут вообще не происходить, происходить только на первом участке AB , только на втором участке BC или на обоих участках одновременно. Пусть вероятности каждого из этих четырех возможных случаев равны соответственно p , q , r и s , где $p + q + r + s = 1$. Мы не принимаем здесь допущения о том, что рекомбинации на обоих участках происходят независимо друг от друга. Ясно, что частоты рекомбинаций между генами A и B и между генами B и C равны

соответственно

$$y_1 = q + s, \quad y_2 = r + s, \quad (10.11)$$

и мы не предполагаем, что в общем случае $s = y_1 y_2$.

Если мы хотим получить полностью сбалансированные данные, нам необходимы данные по каждому из четырех тройных гетерозигот: ABC/abc , aBC/Abc , ABc/abC и aBc/AbC . Пусть число потомков, полученных при этих четырех типах скрещивания, равно соответственно n_1 , n_2 , n_3 и n_4 . Математические ожидания в случае тригибридного возвратного скрещивания в фазе сцепления при отсутствии дифференциальной жизнеспособности приводились в табл. 11. Допустим, что средние значения относительной жизнеспособности для четырех пар фенотипов, встречающихся в потомстве, равны соответственно u , v , w и t , где $u + v + w + t = 4$. Тогда математические ожидания и наблюдаемые значения для всей рассматриваемой совокупности сбалансированных возвратных скрещиваний можно записать в виде, показанном в табл. 15, где приняты следующие обозначения:

$$\left. \begin{aligned} D_1 &= pu + qv + rw + st, \\ D_2 &= pv + qu + rt + sw, \\ D_3 &= pw + qt + ru + sv, \\ D_4 &= pt + qw + rv + su. \end{aligned} \right\} \quad (10.12)$$

Таблица 15

Сбалансированное тригибридное возвратное скрещивание при наличии дифференциальной жизнеспособности

	Отсутст- вие ре- комбина- ций	Рекомби- нации только на первом участке	Рекомби- нации только на втором участке	Рекомби- нации на обоих участках	Сум- ма
$\frac{ABC}{abc}$ Математическое ожидание	$n_1 pu/D_1$	$n_1 qv/D_1$	$n_1 rw/D_1$	$n_1 st/D_1$	n_1
$\frac{abc}{abc}$ Наблюдаемое число	a_1	a_2	a_3	a_4	n_1
$\frac{aBC}{Abc}$ Математическое ожидание	$n_2 pv/D_2$	$n_2 qu/D_2$	$n_2 rt/D_2$	$n_2 sw/D_2$	n_2
$\frac{Abc}{Abc}$ Наблюдаемое число	b_1	b_2	b_3	b_4	n_2
$\frac{ABc}{abC}$ Математическое ожидание	$n_3 pw/D_3$	$n_3 qt/D_3$	$n_3 ru/D_3$	$n_3 sv/D_3$	n_3
$\frac{abC}{abC}$ Наблюдаемое число	c_1	c_2	c_3	c_4	n_3
$\frac{aBc}{AbC}$ Математическое ожидание	$n_4 pt/D_4$	$n_4 qw/D_4$	$n_4 rv/D_4$	$n_4 su/D_4$	n_4
$\frac{AbC}{AbC}$ Наблюдаемое число	d_1	d_2	d_3	d_4	n_4
Наблюдаемая сумма	P	Q	R	S	N

Математические ожидания, приведенные в этой таблице, обнаруживают очевидную комбинаторную симметрию, и можно надеяться, что с помощью какого-либо изящного математического приема можно будет получить в явном виде максимально правдоподобные оценки и средние квадратические ошибки. Так, если для удобства принять обозначения

$$\left. \begin{aligned} U &= a_1 + b_2 + c_3 + d_4, \\ V &= a_2 + b_1 + c_4 + d_3, \\ W &= a_3 + b_4 + c_1 + d_2, \\ T &= a_4 + b_3 + c_2 + d_1, \end{aligned} \right\} \quad (10.13)$$

то функцию правдоподобия для рассматриваемой совокупности наблюдаемых величин можно записать в следующем виде:

$$e^L \sim p^P q^Q r^R s^S u^U v^V w^W t^T D_1^{-n_1} D_2^{-n_2} D_3^{-n_3} D_4^{-n_4}. \quad (10.14)$$

Это соотношение необходимо максимизировать при условиях $p + q + r + s = 1$ и $u + v + w + t = 4$. Поэтому начнем с вычисления таких основных характеристик, как

$$\frac{\partial L}{\partial p} = \frac{P}{p} - \frac{n_1 u}{D_1} - \frac{n_2 v}{D_2} - \frac{n_3 w}{D_3} - \frac{n_4 t}{D_4} \quad (10.15)$$

и т. д.

К сожалению, чисто математический анализ сопряжен здесь с очень сложными алгебраическими выкладками, хотя можно получить ряд интересных результатов для некоторых частных случаев. Допустим, например, что необходимо начать с проверки существования дифференциальной жизнеспособности. Принимаем нулевую гипотезу $u = v = w = 1$ и приступаем к вычислению максимально правдоподобных оценок для этих параметров, используя максимально правдоподобные оценки частот рекомбинаций в отсутствие дифференциальной жизнеспособности. Легко видеть, что эти оценки имеют вид

$$\hat{p} = P/N, \quad \hat{q} = Q/N, \quad \hat{r} = R/N, \quad \hat{s} = S/N. \quad (10.16)$$

Далее находим дифференциальные характеристики

$$\left. \begin{aligned} \frac{\partial L}{\partial u} &= U - (n_1 P + n_2 Q + n_3 R + n_4 S)/N, \\ \frac{\partial L}{\partial v} &= V - (n_1 Q + n_2 P + n_3 S + n_4 R)/N, \\ \frac{\partial L}{\partial w} &= W - (n_1 R + n_2 S + n_3 P + n_4 Q)/N, \\ \frac{\partial L}{\partial t} &= T - (n_1 S + n_2 R + n_3 Q + n_4 P)/N, \end{aligned} \right\} \quad (10.17)$$

которые практически являются оценками для u , v , w и t , если исключить условие, что сумма этих последних должна быть равна 4. Очень удобно использовать преобразование Оуэна

$$\mathbf{u} = \mathbf{Mz}, \quad 4\mathbf{z} = \mathbf{Mu}, \quad (10.18)$$

где

$$\mathbf{u} = \begin{bmatrix} u \\ v \\ w \\ t \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} \quad (10.19)$$

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \quad (10.20)$$

чтобы оперировать с линейно независимыми преобразованными величинами z_2 , z_3 и z_4 . Поскольку величины z_i являются линейными функциями параметров u , v , w и t , можно образовать соответствующие коэффициенты S_i :

$$\left. \begin{aligned} S_2 &= (U - V + W - T) - \\ &\quad - (P - Q + R - S) (n_1 - n_2 + n_3 - n_4)/N, \\ S_3 &= (U + V - W - T) - \\ &\quad - (P + Q - R - S) (n_1 + n_2 - n_3 - n_4)/N, \\ S_4 &= (U - V - W + T) - \\ &\quad - (P - Q - R + S) (n_1 - n_2 - n_3 + n_4)/N. \end{aligned} \right\} \quad (10.21)$$

Если положить

$$\mathbf{S} = \begin{bmatrix} S_2 \\ S_3 \\ S_4 \end{bmatrix}$$

то для проверки нулевой гипотезы можно использовать критерий χ^2 с тремя степенями свободы:

$$\chi_3^2 = \mathbf{S}'\mathbf{V}^{-1}\mathbf{S}, \quad (10.22)$$

где \mathbf{V} — корреляционная матрица коэффициентов S_i . Элементы этой матрицы легче всего вычислить непосредственно, а не с по-

мощью информационных функций. Корреляционная матрица имеет вид

$$V = \frac{16}{N^3} \times \begin{bmatrix} (n_1 + n_3)(n_2 + n_4)(P + R)(Q + S) & (n_1 n_4 - n_2 n_3)(PS - QR) \\ (n_1 n_4 - n_2 n_3)(PS - QR) & (n_1 + n_2)(n_3 + n_4)(P + Q)(R + S) \\ (n_1 n_2 - n_3 n_4)(PQ - RS) & (n_1 n_3 - n_2 n_4)(PR - QS) \\ & (n_1 n_2 - n_3 n_4)(PQ - RS) \\ & (n_1 n_3 - n_2 n_4)(PR - QS) \\ & (n_1 + n_4)(n_2 + n_3)(P + S)(Q + R) \end{bmatrix} \quad (10.23)$$

В табл. 16 приведены данные Парсонса о рецессивных генах *scute* (*sc*), *cross-veinless* (*cv*) и *vermilion* (*v*) плодовой мушки *Drosophila*.

Рассматривая эту таблицу как таблицу сопряженности признаков 4×4 , получаем $\chi^2 = 54,88$ с девятью степенями свободы, т. е. высокосignificantный результат. При допущении о существовании сцепления возможное влияние дифференциальной жизнеспособности можно обнаружить описанным выше методом. Так, диф-

Таблица 16

Сбалансированное тригибридное возвратное скрещивание плодовых мушек *Drosophila*

	Рекомбинации отсутствуют	Рекомбинации только на первом участке	Рекомбинации только на втором участке	Рекомбинации на обоих участках	Сумма
$\frac{+++}{sc\ cv\ v}$	1669	243	480	12	2404
$\frac{sc++}{+cv\ v}$	1646	286	490	11	2433
$\frac{++v}{sc\ cv\ +}$	1350	278	427	15	2070
$\frac{sc+v}{+cv\ +}$	1606	407	485	27	2525
Сумма	6271	1214	1882	65	9432

ференциальные характеристики, заданные выражениями (10.17), равны

$$\begin{aligned}\frac{\partial L}{\partial u} &= +67,0777, & \frac{\partial L}{\partial v} &= -56,1229, \\ \frac{\partial L}{\partial w} &= +50,2908, & \frac{\partial L}{\partial t} &= -61,2456,\end{aligned}$$

а коэффициенты, заданные выражениями (10.21), равны

$$S_2 = +234,7370, \quad S_3 = +21,9096, \quad S_4 = +11,6642.$$

Тогда корреляционная матрица V (10.23) принимает следующий вид:

$$V = \begin{bmatrix} 4410,614 & -37,003 & 88,868 \\ -37,003 & 6176,294 & -260,878 \\ 88,868 & -260,878 & 8302,053 \end{bmatrix}.$$

Теперь по формуле (10.22) можно вычислить χ^2 с тремя степенями свободы; он равен 12,79. Хотя это и значимый результат, он, к сожалению, составляет лишь небольшую часть суммарного χ^2 , равного 54,88. Остающаяся часть χ^2_6 , равная 42,09, также высокозначима. Парсонс предположил, что эта неоднородность, возможно, обусловлена некоторым «эффектом положения», при котором частота рекомбинации изменяется с изменением генотипа родительской гетерозиготной особи.

Для нахождения максимально правдоподобных оценок всех параметров Фишер и Оуэн использовали несколько иную модель, в которой предполагается, что при повторных выборках только общая сумма N сохраняется постоянной, а число потомков с разными фенотипами, n_1 , n_2 , n_3 и n_4 , может изменяться. Это требует иной постановки эксперимента и сопряжено с появлением дополнительных параметров. В конечном счете этот метод дает систему из девяти уравнений, которые можно решить итерационным методом и таким образом получить оценки всех параметров. Применительно к данным, приведенным в табл. 16, окончательно получаем

$$\hat{p} = 0,666, \quad \hat{q} = 0,128, \quad \hat{r} = 0,200, \quad \hat{s} = 0,0068.$$

Из соотношений (10.11) находим, что частота рекомбинаций для $sc-cv$ составляет 13,45%, а для $cv-v$ — 20,67%. Очевидно, что частота двойных рекомбинаций, составляющая около 0,7%, значительно меньше, чем можно было бы ожидать. Это свидетельствует о значительной положительной интерференции.

До сих пор при таком подходе средние квадратические ошибки еще не вычисляли, и в этом его большой недостаток для практики.

С генетической точки зрения (в отличие от математической), по-видимому, лучше всего возвратиться к первой модели и находить максимально правдоподобные оценки путем непосредственного их вычисления. Основные коэффициенты вычисляются по формуле (10.15), а информационные функции легко найти обычным путем, хотя алгебраические выкладки будут довольно громоздкими. Но для преодоления этих технических трудностей, связанных с оценкой параметров, можно составить соответствующую программу для вычислительной машины. Это позволит биологу сосредоточиться на численных результатах и их интерпретации, и ему не придется пугаться сложных алгебраических преобразований, описанных в учебниках по общей математике.

10.4. ПРОСТЫЕ ХРОСОМНЫЕ КАРТЫ

Изучив некоторые проблемы, связанные с оценкой общей частоты рекомбинаций или частоты некоторых видов рекомбинаций (например, рекомбинаций на одном участке хромосомы при отсутствии рекомбинаций на соседнем участке), обратимся теперь к интерпретации этих результатов, исходя из линейного расположения генов на хромосомной карте (см. разд. 10.1, а также [7] гл. 8 и 9). Прежде всего допустим, что рекомбинации на непрерывающихся участках хромосомы происходят независимо, т. е. что интерференция отсутствует. Рассмотрим, как и в разд. 10.2, три локуса, истинным порядком следования которых является порядок ABC , и допустим, что частоты рекомбинаций между локусами A и B , B и C , A и C равны соответственно y_1 , y_2 и y_3 . Мы уже применяли формулу Трой (10.4)

$$y_3 = y_1 + y_2 - 2y_1y_2, \quad (10.24)$$

однако для удобства приводим ее снова.

Чисто математический вывод аддитивной хромосомной карты можно выполнить следующим образом. Формулу (10.24) можно переписать в виде

$$1 - 2y_3 = (1 - 2y_1)(1 - 2y_2). \quad (10.25)$$

Поскольку это выражение представляет собой произведение, то, логарифмируя, получаем сумму. Определим функцию $x(y)$ следующим образом:

$$x(y) = -\frac{1}{2} \ln(1 - 2y). \quad (10.26)$$

После обратного преобразования получаем

$$y(x) = \frac{1}{2}(1 - e^{-2x}). \quad (10.27)$$

Постоянный множитель $-1/2$ в правой части формулы (10.26) выбран для удобства, чтобы сделать x и y одинаковыми при малой частоте рекомбинаций. Переходя к функции x , которую можно рассматривать как расстояние на хромосомной карте, формулу (10.24) можно заменить простой суммой:

$$x_3 = x_1 + x_2. \quad (10.28)$$

Обобщение на случай любого числа смежных участков очевидно и осуществляется элементарно.

Обосновать преобразования (10.26) и (10.27) можно также и иначе, рассматривая основные биологические данные. При нашем упрощающем допущении, согласно которому точки обмена распределены вдоль хромосомы случайным образом (во всяком случае, если используется некоторая подходящая шкала измерений), вероятность того, что между локусами, удаленными друг от друга на расстояние x , находится ровно r таких точек, определяется законом Пуассона:

$$p_r(x) = \frac{e^{-x} x^r}{r!}, \quad r = 0, 1, 2, \dots \quad (10.29)$$

Частота рекомбинаций между этими двумя локусами равна вероятности появления *нечетного* числа точек обмена. Выражение

$$y = \sum_{r=0}^{\infty} p_{2r+1}(x) = \frac{1}{2}(1 - e^{-2x}),$$

приводящее к формуле (10.27), называется формулой Холдейна. Обычно расстояние между локусами измеряется в *морганидах* (по имени Т. Моргана) или в *сантиморганидах*, обозначаемых *cM*. Расстояние между двумя локусами считается равным одной морганиде, если среднее число точек обмена между ними равно единице.

Хотя описанные здесь результаты, полученные при допущении об отсутствии интерференции, удобны в качестве первого приближения, на практике обычно необходима более реалистичная модель. При анализе достаточно обширных данных обычно обнаруживается, что рекомбинации на смежных участках не являются независимыми; более того, неслучайность в кроссинговере хромосом можно наблюдать и цитологически. Связь между рекомбинациями на смежных участках проще всего измерять частотой двойного кроссинговера. Пусть частота рекомбинаций одновременно на участках *AB* и *BC* равна y_{12} . Если рекомбинации на участке *AB* не зависят от рекомбинаций на участке *BC*, то $y_{12} = y_1 y_2$. При наличии положительной интерференции, означающей, что появление одной точки обмена препятствует образованию

другой точки обмена в непосредственной близости от первой, частота двойного кроссинговера уменьшается, т. е. $y_{12} < y_1 y_2$.

Используя выражение (10.11) и учтя, что $y_{12} \equiv s$, находим путем простых вероятностных рассуждений следующее соотношение:

$$y_3 = q + r = y_1 + y_2 - 2y_{12} \quad (10.30)$$

(определение величин p , q , r и s давалось в разд. 10.3). Введя обозначение

$$C = \frac{y_{12}}{y_1 y_2} \quad (10.31)$$

(величина C носит название *коинциденции*), соотношение (10.30) можно записать в виде

$$y_3 = y_1 + y_2 - 2C y_1 y_2. \quad (10.32)$$

Это выражение является обобщением формулы Трой.

Для того чтобы продвинуться дальше, примем соответствующие допущения о коинциденции C . Известно, например, что у плодовой мушки *Drosophila* коинциденция изменяется от одной хромосомы к другой и зависит от длины и положения рассматриваемых участков. Удобно несколько конкретизировать наши рассуждения следующим образом. Положим в формуле (10.32) $y_1 = y(x)$, $y_2 = y(dx) = dx$ и $y_3 = y(x + dx)$; другими словами, будем рассматривать некоторый участок длиной x и примыкающий к нему очень короткий участок длиной dx . После подстановки этих значений в формулу (10.32) получаем

$$y(x + dx) = y(x) + dx - 2C_m(y) y(x) dx, \quad (10.33)$$

где $C_m(y)$ — *краевая (marginal) коинциденция* для участка с частотой рекомбинации y и примыкающего к нему короткого участка. В пределе соотношение (10.33) принимает вид

$$\frac{dy}{dx} = 1 - 2yC_m(y). \quad (10.34)$$

Если функция $C_m(y)$ задана, то из формулы (10.34) путем интегрирования можно получить соответствующую отображающую функцию. Согласно Козамби, вид функции $C_m(y)$ устанавливается на основании следующих рассуждений. При достаточно длинных участках интерференция должна исчезнуть, т. е. $C_m \rightarrow 1$ при $y \rightarrow 1/2$, а для очень коротких смежных участков вполне возможно, что $C_m \rightarrow 0$ при $y \rightarrow 0$. Простейшей функцией с такими свойствами является $C_m(y) = 2y$. В этом частном случае дифференциальное уравнение (10.34) принимает вид

$$\frac{dy}{dx} = 1 - 4y^2. \quad (10.35)$$

После интегрирования получаем

$$\left. \begin{aligned} y &= \frac{1}{2} \operatorname{th} 2x, \\ x &= \frac{1}{4} \ln \frac{1+2y}{1-2y}. \end{aligned} \right\} \quad (10.36)$$

Из формулы (10.36) легко выводится правило сложения частот рекомбинаций:

$$y_3 = \frac{1}{2} \operatorname{th} 2x_3 = \frac{1}{2} \operatorname{th} (2x_1 + 2x_2) = \frac{\frac{1}{2} \operatorname{th} 2x_1 + \frac{1}{2} \operatorname{th} 2x_2}{1 + \operatorname{th} 2x_1 \operatorname{th} 2x_2} = \frac{y_1 + y_2}{1 + 4y_1 y_2}. \quad (10.37)$$

Отметим, что это выражение поразительно напоминает правило сложения скоростей в специальной теории относительности.

Отображающая функция Козамби действительно правильно описывает законы рекомбинации и учитывает явления интерференции у целого ряда организмов, в том числе у плодовой мушки *Drosophila*, душистого горошка, риса и домовой мыши. Рассмотрим применение этой функции на примере данных, полученных М. Уоллес в опытах со сбалансированным тригибридным возвратным скрещиванием для трех локусов *a^t fi Sd*, расположенных на пятой хромосоме домовой мыши (табл. 17). Предполагалось, что в данных, приведенных в табл. 17, влияние дифференциальной жизнен-

Таблица 17

Анализ на сцепление трех локусов *a^t fi Sd* у домовой мыши методом тригибридного возвратного скрещивания

	Отсутствие рекомбинаций	Рекомбинации только на первом участке	Рекомбинации только на втором участке	Рекомбинации на обоих участках	Сумма
Наблюдаемое число	290	127	110	22	549
Математическое ожидание	$n(1 - y_1 - y_2 + y_{12})$	$n(y_1 - y_{12})$	$n(y_2 - y_{12})$	ny_{12}	n

способности в основном исключено. В этой таблице приведены также математические ожидания, причем необходимо иметь в виду, что y_{12} можно выразить через y_1 и y_2 , если принимается правило Козамби. Действительно, используя соотношения (10.30) и (10.37), все математические ожидания можно выразить как функции двух параметров y_1 и y_2 . Методом максимального правдоподобия нахо-

дятся следующие оценки: $\hat{y}_1 = (26,8 \pm 1,9)\%$ для a^t-fi и $\hat{y}_2 = (23,8 \pm 1,8)\%$ для $fi-Sd$; соответствующие расстояния на хромосомной карте равны 29,9 и 25,8 *сМ*. Так как мы имеем три степени свободы, одна степень свободы используется для проверки согласия. Такая проверка дала вполне удовлетворительное значение критерия χ^2 , равное 2,79.

Отображающая функция Козамби вполне пригодна для некоторого ограниченного класса ситуаций, однако она носит слишком эмпирический характер и не может быть полезной в общем случае. Например, она предполагает наличие определенной фиксированной величины интерференции и может быть справедлива только для определенных участков хромосомы. Кроме того, она не учитывает роли расстояния относительно центромеры. И наконец, эта функция вообще не учитывает множественных рекомбинаций на длинных участках хромосомы. Очевидно, что необходимо найти какой-то другой, более общий подход, для чего нужно рассмотреть весь процесс обмена с самого начала. Об этом и пойдет речь в следующем разделе.

10.5. МЕТРИКИ ИНТЕРФЕРЕНЦИИ

В конце разд. 10.1 указывалось, что при рассмотрении характера рекомбинаций при наличии интерференции целесообразнее использовать модель четырех нитей, согласно которой каждая гомологичная хромосома состоит из двух одинаковых хроматид, связанных лишь в одном определенном локусе при помощи центромеры. В кроссинговере участвуют все четыре хроматиды; считается, что этот процесс начинается из центромеры и распространяется к свободным концам хроматид. Полностью описать весь этот случайный процесс чрезвычайно трудно, поскольку нужно учесть существование четырех отдельных хроматид и нескольких различных типов интерференции. В первом приближении большинства этих трудностей можно избежать, ограничившись рассмотрением процесса, происходящего вдоль одной хроматиды, или нити. Мы сосредоточим внимание на точках, в которых происходит обмен генетического материала, и попытаемся найти такое распределение вероятностей для величины интервалов между последовательными точками обмена, которое удовлетворительно описывало бы различные возможные эффекты интерференции. Будем считать, что нити, расположенные по разные стороны от центромеры, независимы, хотя и имеются данные о том, что иногда происходит интерференция и через центромеру.

Итак, допустим, что рассматриваемая нить представляет собой полубесконечный континуум точек, которые можно нанести на полубесконечную прямую с началом в центромере ([7], гл. 11). Таким образом, любой точке нити соответствует определенная

точка на прямой, расстояние которой t от начала отсчета измеряется в некоторой удобной метрике. Эту метрику мы и будем искать. Предположим, что точки обмена образуются вдоль нити последовательно. Начнем с центромеры. Пусть все интервалы между последовательными точками обмена при измерении в метрике представляют собой независимые случайные величины с одинаковыми распределениями

$$dF = f(u) du. \quad (10.38)$$

Соответствующие функции распределения имеют вид

$$F(u) = \int_0^u f(v) dv. \quad (10.39)$$

Нетрудно видеть, что такой подход приводит к известному процессу восстановления. Поэтому с помощью теории восстановления можно непосредственно получить целый ряд результатов. К сожалению, нас мало интересуют общие теоремы, так как мы пытаемся определить конкретные результаты для модели, описывающей наблюдаемый характер рекомбинаций. Кроме того, чтобы модель была более достоверной, в нее необходимо внести некоторые изменения, что значительно затруднит применение стандартной теории восстановления. Поэтому часто целесообразно начинать исследование с самого начала. Однако, прибегнув к аналогии с теорией восстановления, некоторую предварительную информацию можно все же получить путем следующих рассуждений.

Пусть вероятность появления n точек обмена (исключая центромеру) в интервале $(0, t)$ равна $p_n(t)$ с производящей функцией вероятностей

$$P(z, t) = \sum_{n=0}^{\infty} p_n(t) z^n. \quad (10.40)$$

На данном этапе удобнее воспользоваться преобразованием Лапласа, которое для произвольной функции $\varphi(t)$ определяется следующим образом:

$$\varphi^*(s) = \int_0^{\infty} e^{-st} \varphi(t) dt, \quad \operatorname{Re}(s) > 0. \quad (10.41)$$

В теории восстановления существует стандартный результат, согласно которому преобразование $P^*(z, s)$ функции $P(z, t)$ можно выразить через преобразование $f^*(s)$ функции $f(t)$ следующим образом:

$$P^*(z, s) = \frac{1 - f^*(s)}{s \{1 - z f^*(s)\}}. \quad (10.42)$$

В данном случае расстояние на хромосомной карте $x(t)$ равно среднему числу точек обмена в интервале $(0, t)$. Следовательно, преобразование $x^*(s)$ функции $x(t)$ равно $[\partial P^*/\partial z]_{z=1}$, т. е.

$$x^*(s) = \frac{f^*(s)}{s\{1 - f^*(s)\}}. \quad (10.43)$$

Частота рекомбинаций в интервале $(0, t)$ равна

$$y(t) = \sum_{n=0}^{\infty} p_{2n+1}(t) = \frac{1}{2} \{1 - P(-1, t)\},$$

а преобразованием функции $y^*(s)$ является выражение

$$\frac{1}{2} \{s^{-1} - P^*(-1, s)\},$$

т. е.

$$y^*(s) = \frac{f^*(s)}{s\{1 + f^*(s)\}}. \quad (10.44)$$

В частном случае, когда интерференция отсутствует, точки обмена распределены случайным образом и можно принять, что $f(u) = e^{-u}$; преобразование этой функции имеет вид $f^*(s) = (s + 1)^{-1}$. Подставляя это соотношение в формулу (10.43), получаем $x^*(s) = s^{-2}$, т. е. $x = t$. Таким образом, в данном случае эта специальная меррика совпадает с расстоянием на хромосомной карте. Затем, подставляя значение $f^*(s)$ в формулу (10.44), получаем $y^*(s) = s^{-1}(s + 2)^{-1}$, откуда путем обратного преобразования находим $y(t) = \frac{1}{2}(1 - e^{-2t})$ или $y(x) = \frac{1}{2}(1 - e^{-2x})$, как в полученной ранее формуле (10.27).

Разумеется, больший интерес представляют функции $f(u)$ более общего вида, особенно функция, предложенная Оуэном (см. другую книгу автора [7], разд. 12.2):

$$f(u) = 4ue^{-2u}. \quad (10.45)$$

Эта функция по существу является нормированным распределением χ^2 с четырьмя степенями свободы, у которого математическое ожидание равно единице, т. е. случайная величина u распределена как $\frac{1}{4}\chi_4^2$. Преобразование функции (10.45) имеет вид

$$f^*(s) = 4(s + 2)^{-2}. \quad (10.46)$$

Подставляя значения $f^*(s)$ в формулы (10.43) и (10.44), после обратного преобразования получаем

$$x(t) = -\frac{1}{4} + t + \frac{1}{4}e^{-4t} \quad (10.47)$$

и

$$y(t) = \frac{1}{2} - \frac{1}{2} e^{-2t} (\cos 2t + \sin 2t). \quad (10.48)$$

При $t \rightarrow \infty$ имеем $y(t) \rightarrow \frac{1}{2}$, но интересно, что при конечном t возникает ряд затухающих колебаний. Легко убедиться в том, что первое стационарное значение появляется в точке $t = \frac{1}{2} \pi$, когда частота рекомбинаций составляет около 52,2% и расстояние на хромосомной карте равно 133 *cM*. Особый интерес представляет возможность появления в данной модели частоты рекомбинаций, превышающей 50%. Такие значения иногда наблюдаются, хотя и не всегда легко определить, обусловлены ли они интерференцией рассматриваемого типа или же некоторыми другими эффектами.

Рассмотренный здесь вывод можно распространить и на более общий случай, когда интервал (t_1, t_2) не начинается в центромере; правда, описание при этом значительно усложняется. В этом случае при использовании модели Оуэна возможно появление несколько более высокой частоты рекомбинаций — до 53,3%.

Для простоты изложения мы рассматривали теоретически бесконечные хромосомные нити. Поэтому необходимо внести определенную коррекцию с учетом того, что нити имеют конечную длину. Выполнить это можно различными способами. Пусть, например, нить имеет длину T . Тогда можно использовать рассмотренную выше модель бесконечной нити и просто считать, что значения имеют лишь события, происходящие в интервале $(0, T)$, а события в интервале (T, ∞) несущественны. Другой способ основан на следующих соображениях. Имеются некоторые основания полагать, что должна существовать по крайней мере одна точка обмена на пару хромосом; поэтому есть смысл рассматривать для данной нити вероятностную ситуацию, обуславливаемую этим событием. Однако наличие хотя бы одной точки обмена на пару хромосом (содержащую четыре хроматиды) — не то же самое, что наличие хотя бы одной точки обмена в любой данной хроматиде, поэтому при таком подходе нужно учитывать совместную конфигурацию всех четырех хроматид, что весьма сложно. Другой метод, предложенный Оуэном, состоит в том, что рассматривается распределение событий в интервале $(0, T)$ при условии, что в интервале (T, ∞) точки обмена отсутствуют. С математической точки зрения такой прием, вообще говоря, приемлем, но вряд ли он будет привлекателен для биолога, не имеющего глубокой математической подготовки. Интересно, что в такой модели частота рекомбинаций может достигать 60%.

Оценка параметров

С точки зрения метрической теории как расстояния на хромосомной карте, так и частоты рекомбинаций являются производными величинами, а основными параметрами являются именно метрические расстояния. Поэтому возникает вопрос, каким образом правильно оценить их. По существу, следуя обычным статистическим принципам, нужно вначале выразить математические ожидания для всех наблюдаемых типов рекомбинаций через соответствующие метрические длины. После этого в случае удачи можно получить необходимые оценки стандартным методом максимального правдоподобия (см. гл. 13 другой книги автора [7]).

До сих пор не предпринималось попыток оценить параметры в общем случае при каком-либо данном распределении интервалов $f(u)$, хотя теория оценок для частного случая метрики Оуэна разработана. Данная метрика содержит интервалы с распределениями, которые по существу являются распределениями χ^2 с четырьмя степенями свободы, т. е. каждый из них можно рассматривать как свертку двух независимых случайных интервалов. Таким образом, необходимое вероятностное описание можно получить сравнительно просто с помощью распределения Пуассона. Мы опускаем детали и приводим лишь конечные результаты, необходимые для получения оценок.

Рассмотрим начинающуюся с центromеры хромосомную нить, состоящую из n участков и завершающуюся свободным концом. Эта нить содержит $n - 1$ внутренних локусов, т. е. $n - 1$ идентифицируемых пар аллелей. Пусть метрические длины этих участков равны t_i ($i = 1, 2, \dots, n$); тогда общая длина нити определяется выражением

$$T = \sum_{i=1}^n t_i. \quad (10.49)$$

Если проводится анализ на сцепление ($n - 1$) факторов методом возвратного скрещивания, то необходимо знать вероятности появления всех возможных типов рекомбинаций между экспериментально различимыми локусами. Можно показать, что вероятность появления какой-либо определенной схемы множественных рекомбинаций можно записать в виде

$$CG_2G_3 \dots G_{n-2}G_{n-1}D/\text{ch } 2T, \quad (10.50)$$

где

$$C = [\text{ch } 2t_1 \text{ sh } 2t_1]. \quad (10.51)$$

$$D = \begin{bmatrix} \text{ch } 2t_n \\ \text{sh } 2t_n \end{bmatrix} \quad (10.52)$$

и

$$\left. \begin{aligned}
 G_i &\equiv R_i = \begin{bmatrix} \gamma(t_i) & \delta(t_i) \\ \beta(t_i) & \gamma(t_i) \end{bmatrix} \\
 & \text{(если рекомбинации происходят на } i\text{-м участке),} \\
 G_i &\equiv N_i = \begin{bmatrix} \alpha(t_i) & \beta(t_i) \\ \delta(t_i) & \alpha(t_i) \end{bmatrix} \\
 & \text{(если рекомбинации на } i\text{-м участке не происходят).}
 \end{aligned} \right\} (10.53)$$

Функции $\alpha(x)$, $\beta(x)$, $\gamma(x)$ и $\delta(x)$ имеют вид

$$\left. \begin{aligned}
 \alpha(x) &= \frac{1}{2} (\operatorname{ch} x + \cos x), & \beta(x) &= \frac{1}{2} (\operatorname{sh} x + \sin x), \\
 \gamma(x) &= \frac{1}{2} (\operatorname{ch} x - \cos x), & \delta(x) &= \frac{1}{2} (\operatorname{sh} x - \sin x).
 \end{aligned} \right\} (10.54)$$

Вследствие того что в выражении (10.50) крайний множитель слева является вектор-строкой, а крайний множитель справа — вектор-столбцом, это выражение действительно скалярно.

В качестве краткой иллюстрации рассмотрим приведенные в табл. 18 данные Парсонса, полученные им при анализе на сцепление четырех факторов методом возвратного скрещивания домашних мышей. Парсонс исследовал локусы генов *polydactyly* (*py*),

Таблица 18

Оценка метрических расстояний при четырехгибридном возвратном скрещивании

Номера участков	Математическое ожидание числа рекомбинаций	Наблюдаемое число рекомбинаций
(Отсутствие рекомбинаций)	$nCN_2N_3N_4D/\operatorname{ch} 2T$	1129
2	$nCR_2N_3N_4D/\operatorname{ch} 2T$	499
3	$nCN_2R_3N_4D/\operatorname{ch} 2T$	89
4	$nCN_2N_3R_4D/\operatorname{ch} 2T$	678
2-3	$nCR_2R_3N_4D/\operatorname{ch} 2T$	8
2-4	$nCR_2N_3R_4D/\operatorname{ch} 2T$	148
3-4	$nCN_2R_3R_4D/\operatorname{ch} 2T$	15
2-3-4	$nCR_2R_3R_4D/\operatorname{ch} 2T$	7
Всего	n	2573

leaden (*ln*), *splotch* (*Sp*) и *fuzzy* (*fz*). Предполагалось, что при составлении табл. 18 в основном было исключено влияние дифференциальной жизнеспособности; несколько неполная пенетрантность гена *polydactyly* игнорировалась. Если все локусы расположены по одну и ту же сторону от центромеры, то математические ожидания будут именно такими, которые приводятся в таблице. Надо сказать, что вычисление соответствующих коэффициентов и информационных функций — весьма утомительное занятие. Оценки пяти параметров, вычисленные методом максимального правдоподобия, имеют следующие значения:

$$\hat{t}_1 = 0,54 \pm 3,73,$$

$$\hat{t}_2 = 0,27 \pm 0,06,$$

$$\hat{t}_3 = 0,052 \pm 0,005,$$

$$\hat{t}_4 = 0,30 \pm 0,06,$$

$$\hat{t}_5 = 0,24 \pm 0,32.$$

Оценки для первого и последнего участков фактически не имеют смысла, так как данных недостаточно для того, чтобы получить точные результаты. Тем не менее можно найти способ представления этого материала. Ниже приведена хромосомная карта для части 13-й группы сцепления у домашней мыши, содержащая четыре гена-маркера, на которой помимо расстояний на хромосомной карте указаны также частоты рекомбинаций и метрические расстояния; все величины даются в процентах.

	Метрическое расстояние,%	Расстояние на хромосомной карте,%	Частота реком- бинаций,%
○ Центромера	(54)	(32)	(27)
py	27	26	25
ln	5	5	5
Sp	30	36	33
fz	(24)	(39)	(36)
Свободный конец			

У относительно близко расположенных локусов разница между этими тремя измеренными параметрами не очень заметна, однако

при увеличении длины участков она становится более заметной. Модель в этом случае оказывается не вполне удовлетворительной, так как вследствие однородности критерия χ^2 имеем $\chi_1^2 = 3,96$, и если пятый и восьмой классы в табл. 18 объединить, то для оценки пяти параметров остается шесть степеней свободы. Теоретически при анализе на сцепление четырех факторов методом возвратного скрещивания нужно было бы анализировать полные данные, позволяющие в явном виде оценивать влияние дифференциальной жизнеспособности (а не надеяться, что с помощью комбинированных данных удастся исключить этот эффект).

Наконец, заметим, что для таких моделей очень трудно выполнять вычисления с помощью настольных счетных машин, но с этой задачей легко справиться, составив надлежащую программу для электронной вычислительной машины. Тогда биологу не придется заниматься техническими вопросами, связанными с получением результатов, и это позволит ему сосредоточить внимание на оценке параметров и их интерпретации.

Глава 11

МАТЕМАТИЧЕСКИЕ МЕТОДЫ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

11.1. ВВЕДЕНИЕ

Вряд ли кто станет отрицать, что диагностика играет в медицине важнейшую роль и что постановка диагноза требует от врача большого мастерства, знаний и интуиции. Точность диагноза и быстрота, с которой его можно поставить, зависят, разумеется, от очень многих факторов: от состояния больного, от имеющихся данных о симптомах и признаках заболевания и результатах лабораторных анализов, от общего объема медицинской информации о наблюдении таких симптомов при самых различных заболеваниях и, наконец, от квалификации самого врача. Своевременно поставленный точный диагноз часто облегчает выбор метода лечения и значительно повышает вероятность выздоровления больного. Исходя из всех этих соображений, вполне естественно попытаться определить условия, при которых диагноз может быть поставлен максимально быстро и точно. В течение многих веков врачи с переменным успехом предпринимали попытки решить эту задачу. Однако в последние годы благодаря применению современных методов лечения и диагностики, основанных на новейших достижениях науки и техники, возможности получения успешных результатов значительно возросли. Поэтому важно найти точные методы описания, исследования, оценки и контроля процесса постановки диагноза. Как уже неоднократно указывалось в первой части книги и особенно в разд. 1.4, наилучший путь к точности и логике рассуждений при решении любой задачи — это математический подход. В принципе этот подход можно выбирать независимо от того, насколько труден и сложен рассматриваемый вопрос. Если мы имеем дело с большим числом взаимозависимых факторов, обнаруживающих значительную естественную изменчивость, то для достаточно эффективного описания сложной схемы их влияния существует *лишь один* способ — использование соответствующего статистического метода. Если число факторов или число категорий данных очень велико, то желательно, или даже необходимо, использовать электронную вычислительную машину, чтобы искомые результаты можно было получить за достаточно короткое время. Такой подход ни в коей мере не умаляет значения интуиции и воображения. Напротив, он открывает еще

большой простор для проявления этих качеств, освобождая врача от необходимости заниматься такими проблемами, которые можно сформулировать в численной и логической форме и, следовательно, решать математическими методами и с помощью вычислительной техники.]

Итак, что же можно сделать для того, чтобы применить эти идеи к медицинской диагностике? Как известно, среди математиков, специалистов в области вычислительной техники и врачей уже имеется ряд энтузиастов, работающих над применением математики и вычислительной техники в этой области. Естественно, что симпатии автора на стороне этих энтузиастов. В то же время важно иметь четкое представление о том, что теоретически достижимо, что можно было бы сделать в определенных реальных условиях и какие последствия окажет это в целом на взаимоотношения врача и больного.

Даже если бы практическое использование вычислительных машин для диагностики показалось бы кому-нибудь нежелательным, это все равно не умалило бы важности математического анализа рассматриваемых процессов, поскольку такой анализ должен значительно расширить и углубить наши знания. Имеющиеся в настоящее время данные свидетельствуют о том, что вычислительные машины, несомненно, могут играть важную роль при постановке диагноза. Главное состоит в том, чтобы определить возможности вычислительной техники, и в дальнейшем мы рассмотрим этот вопрос более подробно. В данный момент укажем лишь некоторые важные направления: постановка дифференциального диагноза в соответствующих условиях; оценка точности диагнозов, которые ставят врачи, с целью повышения общего уровня диагностики; создание учебных пособий для студентов, а также сбор, обобщение и обработка клинических данных для квалифицированного использования их врачами при постановке диагноза.

В настоящей главе будут кратко изложены вопросы построения математических моделей, описывающих некоторые аспекты процесса постановки диагноза, и применения их для оказания помощи врачам при оценке хотя бы некоторых сложных и крайне изменчивых медицинских данных. Разработка методов диагностики с помощью вычислительных машин находится пока еще на самой начальной стадии, однако исследователями, работающими в ряде стран, уже получены весьма обнадеживающие результаты, и дальнейшие изыскания в этой области следует считать весьма перспективными.

Разумеется, концентрация внимания на постановке дифференциального диагноза является во многих отношениях чрезмерно упрощенным или, во всяком случае, ограниченным подходом

к проблеме в целом. Мы будем предполагать, что все альтернативные диагнозы, из которых нужно выбрать один, четко и однозначно определены. Однако на практике дело обстоит совсем не так (во всяком случае, иногда). Мнения специалистов о наилучших способах классификации болезней (в отличие от распределения больных по заранее определенным классам) нередко расходятся, и новые данные могут потребовать пересмотра существующих схем. С этой проблемой связаны, естественно, вопросы медицинской таксономии, и, возможно, потребуется изучить на широкой основе применение методов числовой таксономии, рассмотренных в гл. 7 в общем биологическом плане.

Кроме того, успех лечения в каждом конкретном случае во многом зависит от предварительного диагноза. Этот диагноз может быть пересмотрен, если метод лечения, который считался наилучшим, оказывается неэффективным или если больной реагирует на него неожиданным образом. Фактически реакцию на лечение можно рассматривать как проверку правильности предварительного диагноза, и она служит дополнительным источником информации. Разумеется, этот способ широко применяется в клинической практике. Однако главное здесь в том, что нам может потребоваться математическое описание всего процесса — классификации болезней, постановки дифференциального диагноза и анализа результатов лечения, прежде чем при таком подходе мы сможем добиться сколько-нибудь значительных успехов.

В литературе имеется довольно много статей по этому вопросу, однако по-настоящему авторитетного руководства еще не написано. Заслуживает внимания очень интересный отчет о конференции, состоявшейся в Мичиганском университете в 1964 г. [36], в котором дается общий обзор широкого круга проблем, связанных с медицинской диагностикой. Отдельные статьи на эту тему имеются в трудах Рочестерских конференций [19—21]. Мы не собираемся давать здесь полный перечень работ в этой области, но, помимо названных книг, сошлемся в дальнейшем еще на несколько важных и полезных статей.

11.2. ДИФФЕРЕНЦИАЛЬНЫЙ ДИАГНОЗ В ЭЛЕМЕНТАРНОМ СЛУЧАЕ ВЫБОРА МЕЖДУ ДВУМЯ ЗАБОЛЕВАНИЯМИ

Рассмотрим прежде всего простейшую задачу — постановку дифференциального диагноза в том случае, когда требуется произвести выбор между двумя заболеваниями. Допустим, что имеется стандартный список симптомов, признаков, данных лабораторного обследования, данных анамнеза и т. д., наличие которых обычно проверяют у обследуемых больных, если у них подозревается одно из этих двух заболеваний. Рано или поздно в каждом

случае ставится окончательный диагноз. Такого рода списки содержат определенный объем данных, относящихся к каждой из этих болезней. При поступлении нового больного лечащий врач может просмотреть имеющиеся списки симптомов и, исходя из них, поставить диагноз, который кажется ему наиболее вероятным. Такой просмотр перечня симптомов обычно производится преимущественно по памяти, хотя за подтверждением первых впечатлений, основанных на интуиции и опыте, врач может обращаться и к количественным сводным данным. Итак, что же можно (или нужно) было бы сделать, чтобы посредством извлечения и использования максимального количества статистической информации, содержащейся в таких списках, найти точную математическую оценку ситуации?

Более детально этот вопрос обсуждается в следующем разделе, где рассматривается более широкая проблема дифференциальной диагностики одного из нескольких заболеваний. Здесь же мы остановимся на одном простом, однако весьма успешном численном методе дифференциальной диагностики в случае двух заболеваний, основанном на некоторых идеях статистического дискриминантного анализа. Согласно этому методу, изучаемые объекты распределяются между двумя группами на основе некоторого показателя, равного сумме нескольких надлежащим образом взвешенных компонент (очков), каждая из которых измеряет какую-то конкретную характеристику. Если сумма очков, вычисленная для данного объекта, превышает некоторую пороговую величину, то этот объект попадает, скажем, в первую группу, а в противном случае — во вторую. Разумеется, этот процесс небезупречен, и возможно появление ошибок. Однако для весовых коэффициентов и порога можно найти такие оптимальные значения, чтобы возможность ошибочного диагноза была сведена к минимуму.

Крукс, Мёррей и Уэйн [16] применяли этот метод при диагностике тиреотоксикоза. Хотя клиническая картина этого заболевания не очень сложна, нередко встречаются случаи, когда либо диагноз сомнителен, либо мнения врачей расходятся, либо один и тот же врач со временем изменяет свое мнение. Поэтому было решено разработать простой метод численной оценки, с тем чтобы получить возможность более точной постановки диагноза, особенно в сомнительных случаях. В табл. 19 приведен перечень признаков и симптомов, которым уделялось особое внимание, а также число очков для каждого признака или симптома при его наличии или отсутствии.

Первоначально число очков для каждого отдельного признака или симптома получали не на основе сложных статистических исследований, а исходя из более или менее интуитивной и эмпирической оценки имеющихся данных. Те характеристики, которые

Таблица 19

**Число очков для признаков и симптомов у больных
с подозрением на тиреотоксикоз**

	Наличие	Отсутствие
Одышка при физической нагрузке	+1	0
Сердцебиение	+2	0
Утомляемость	+2	0
Чувствительность к низкой температуре (предпочитают жару)	0	-5
Чувствительность к высокой температуре (предпочитают холод)	+5	0
Нечувствительность к колебаниям температуры	0	0
Потливость	+3	0
Нервозность	+2	0
Повышенный аппетит	+3	0
Пониженный аппетит	0	-3
Увеличение веса	0	-3
Уменьшение веса	+3	0
Увеличение щитовидной железы по данным пальпации	+3	-3
Шумы над щитовидной железой	+2	-2
Экзофтальм	+2	0
Ретракция века	+2	0
Отставание века	+1	0
Гиперкинезы	+4	-2
Тремор пальцев рук	+1	0
Руки		
горячие	+2	-2
влажные	+1	-1
Мерцание предсердий	+4	0
Ритмичный пульс		
свыше 90	0	-3
от 80 до 90	0	0
ниже 80	+3	0

считались более важными для диагностики, получали большее число очков. Классификация производилась таким образом, что «наличие» признака всегда отмечалось тем или иным положительным числом, а «отсутствие» — отрицательным (нули означают,

что в данном случае ни то ни другое не имеет диагностического смысла). Затем в таблице были произведены некоторые изменения, например были уменьшены очень большие значения, с тем чтобы ослабить влияние колебаний результатов, получаемых различными врачами.

После этого были выбраны две типичные группы индивидуумов и для каждой был произведен подсчет суммарного числа очков. В одной группе было 83 больных, несомненно страдающих тиреотоксикозом, а во второй — 99 человек, у которых не было тиреотоксикоза. Одни представители второй группы были совершенно здоровы, а у других были различные заболевания, в том числе простой зоб. Для каждого индивидуума в соответствии с принятыми правилами вычислялась сумма очков. Естественно, что наблюдались значительные колебания этой суммы, однако важно то, что весь ряд чисел, полученных для группы больных тиреотоксикозом, оказался выше соответствующего ряда для группы индивидуумов, не страдающих этим заболеванием. После этого были внесены соответствующие изменения, с тем чтобы добиться наибольшего различия между этими двумя группами и исключить чрезмерное влияние отмечавшихся ранее колебаний в числе очков, приписываемых различными наблюдателями. В табл. 19 приведены окончательные значения весовых коэффициентов для отдельных признаков и симптомов.

При использовании этих измененных весовых коэффициентов общее число очков для достоверных случаев тиреотоксикоза колеблется в пределах от +21 до +42, а для группы индивидуумов, не страдающих тиреотоксикозом, в пределах от -16 до +10. Таким образом, оказалось, что все больные, у которых число очков составляет не менее +21, страдают тиреотоксикозом, а у индивидуумов с числом очков не более +10 этого заболевания нет. Следовательно, деление больных на две группы, основанное на сумме очков, полностью оправдало себя. Хотя в данном случае всех индивидуумов можно было совершенно четко отнести к той или иной группе, при значительно большем объеме выборок следует ожидать появления некоторых неопределенных случаев. Лакмусовой бумажкой, позволяющей проверить надежность эмпирических методов такого рода, служит их пригодность или непригодность в применении к новым данным. Проявляя достаточную избирательность при обработке определенной совокупности данных, всегда можно разработать правила прогнозирования, обладающие поразительной точностью. Справедливость предложенных критериев была соответствующим образом проверена на совершенно другой группе больных.

Казалось бы, что нужно применить эти методы подсчета числа очков к новым больным, часть которых страдает тиреотоксикозом,

а другая часть — нет, и убедиться в том, что правила, специально разработанные для первоначальных групп, по-прежнему приемлемы. Тщательные исследования методами математической статистики требовали бы именно такого подхода. Однако Крукс, Мёррей и Уэйн пошли по более трудному пути. Они отобрали 118 случаев, в каждом из которых врачи испытывали затруднения при постановке диагноза. Тех больных, относительно которых у врачей никаких сомнений не возникало, исследователи не учитывали, а сконцентрировали все внимание на сомнительных случаях, в которых окончательный диагноз был установлен только после продолжительных наблюдений и после учета результатов лечения. В 51 случае врачи в конце концов пришли к выводу, что имеют дело с тиреотоксикозом. В 88% этих случаев число очков составляло не менее +20, а в остальных 12% случаев оно лежало в области сомнительных значений между +11 и +19. Из остальных 67 больных, у которых тиреотоксикоз был врачами исключен, 88% получили число очков +10 и менее, около 10,5% — число очков, лежащее в области сомнительных значений от +11 до +19, и лишь один случай (1,5%) ошибочно попал в область, характерную для тиреотоксикоза. С аналогичной точностью (около 85%) диагноз был установлен при оценке состояния 171 больного и в четырех других медицинских учреждениях.

Особенно важно, что такой простой математический метод позволил точно отделить больных, страдающих тиреотоксикозом, от других больных именно в тех случаях, когда клиническая оценка заболевания была сопряжена с трудностями. Очевидно, что в тех случаях, когда врач может быстро и легко принять правильное решение сам, применять вычислительные методы ни к чему. Но если этот метод может помочь специалисту прийти к определенному выводу в тех случаях, когда сам он затрудняется в постановке диагноза, то это может иметь большое практическое значение.

Описанный здесь численный метод можно сделать несколько более точным, проведя полный статистический анализ с использованием дискриминантной функции. Насколько известно автору, такая работа выполнялась на материале тиреотоксикоза, однако к моменту написания книги результаты еще не были опубликованы. Следует ожидать, что они окажутся примерно теми же, но точность их будет выше. Поскольку метод дискриминантных функций трудно поддается обобщению, мы не будем останавливаться здесь на его математических деталях, а в следующем разделе покажем, чего можно достигнуть, не обращаясь к таким высоким материям.

11.3. ДИФФЕРЕНЦИАЛЬНЫЙ ДИАГНОЗ В ОБЩЕМ СЛУЧАЕ ВЫБОРА МЕЖДУ НЕСКОЛЬКИМИ ЗАБОЛЕВАНИЯМИ

Часто врач должен произвести выбор между несколькими возможными диагнозами. В этом случае разработать удовлетворительную эмпирическую систему подсчета очков, аналогичную описанной в предыдущем разделе, крайне трудно. Кроме того, отсутствие четко сформулированной исходной математической модели сильно осложняет интерпретацию и оценку результатов. Разумеется, можно было бы попытаться использовать метод дискриминантных функций, о котором говорилось в предыдущем разделе, но при этом в свою очередь возникает ряд новых проблем. Прежде всего при классическом подходе к делению больных на две группы предполагается, что корреляционные матрицы отмеченных симптомов в каждой группе одинаковы и что распределения являются многомерными нормальными. Первое допущение не имеет существенного значения, однако позволяет упростить математические выкладки. Вполне можно рассмотреть более общий случай, когда корреляционные матрицы неодинаковы, но он почти наверняка потребует большого объема вычислений на ЭВМ. Допущение о многомерном нормальном распределении также вряд ли приемлемо в общем случае. Еще более серьезные затруднения возникают при обобщении метода дискриминантных функций на случай нескольких групп (если их нельзя расположить в определенном однозначном линейном порядке и если не выполняются допущения о многомерном нормальном распределении и равенстве корреляционных матриц). Можно было бы использовать обобщенный показатель расстояния Махаланобиса, однако и в этом случае возникает много затруднений в применении теории и интерпретации получаемых результатов.

Однако существует целый ряд статистических методов, которые стоило бы испробовать. Масуяма [44] перечисляет 11 методов, но из них только три-четыре существенно отличаются друг от друга. По мнению автора, один-два метода обладают значительными преимуществами. Одним из них является анализ вероятностей появления отдельных симптомов при определенных заболеваниях, позволяющий вычислить вероятность того, что данный больной страдает определенным заболеванием, относящимся к рассматриваемой группе (см., например, [39]). По существу это так называемый байесовский подход (по названию одного правила теории вероятностей, носящего имя Томаса Бейеса). Поскольку основные идеи выражаются через простые вероятности, не представляет труда описать исходную математическую модель и понять смысл практических результатов. Рассмотрим этот метод более детально.

Допустим, что рассматривается ограниченная группа d различных заболеваний (D_1, D_2, \dots, D_d) и что каждый больной страдает только одним из них (т. е. из этого упрощенного описания исключается более сложный случай сочетания нескольких заболеваний). Допустим также, что имеется список s признаков, симптомов или результатов лабораторных анализов (на основе которых обычно получают необходимую информацию), обозначаемых S_1, S_2, \dots, S_s . Часто бывает удобно рассматривать всю совокупность признаков, симптомов и т. д.; для этой цели используется символ S без индекса. Пока мы будем считать, что все симптомы дискретны, т. е. каждый из них относится к одному из двух или нескольких конкретных классов (например, «наличие» и «отсутствие» признака в случае простой дихотомии).

Допустим далее, что за прошлые годы накоплен обширный материал по всем этим заболеваниям, и поэтому для любого данного заболевания D_j можно задать вероятность $P(S | D_j)$ наблюдения определенного симптомокомплекса S . Если этот материал достаточно представлен, то можно также определить вероятность $P(D_j)$ того, что больной, выбранный случайным образом, страдает заболеванием D_j ; эта вероятность, определяемая еще до того, как будет принято во внимание клиническое состояние данного больного, является *априорной вероятностью* наличия заболевания D_j . Хотя врач, производящий обследование больного, может и не пользоваться этими статистическими категориями при обдумывании возможного диагноза, он скорее предположит какое-нибудь распространенное заболевание, чем заболевание, встречающееся крайне редко.

Далее врач сообщает информацию о симптомах, наблюдаемых у данного больного. Вполне возможно, что предварительный диагноз будет поставлен довольно быстро, но при этом врач нередко будет иметь в виду и другие возможности. В конце концов наступает момент, когда накоплено достаточное (во всяком случае, на данный момент) количество существенной информации. Эти данные характеризуют определенный симптомокомплекс S . Для того чтобы немедленно начать лечение и, возможно, назначить специальные дополнительные проверки, необходимо поставить предварительный диагноз. Выражаясь математическим языком, врачу нужно знать вероятность каждого заболевания *при данном* симптомокомплексе, т. е. $P(D_j | S)$. Это по существу просто численный способ описания того, что происходит в каждодневной клинической практике. Если заболевание D_1 диагностируется с большой уверенностью, то вероятность $P(D_1 | S)$ будет относительно велика, а все остальные вероятности будут значительно меньше (возможно, в 10 или 100 раз). Если же два заболевания D_1 и D_2 кажутся почти одинаково возможными, а все остальные

исключаются, то каждая из вероятностей $P(D_1 | S)$ и $P(D_2 | S)$ будет равна примерно 0,5, а все остальные вероятности будут близки к нулю и т. д. Обычно уверенность врача в правильности поставленного им диагноза покоится на его общем опыте и субъективном мнении. Но последние должны всегда основываться в конечном счете на накопленных ранее знаниях или клинических данных (наблюдавшихся им лично или другими врачами и опубликованных в медицинских журналах и руководствах), а также на данных, полученных в результате обследования больного.

Известно, что *апостериорная вероятность* $P(D_j | S)$ пропорциональна произведению априорной вероятности $P(D_j)$ на функцию правдоподобия $P(S | D_j)$, т. е.

$$P(D_j | S) \sim P(D_j) P(S | D_j). \quad (11.1)$$

Так как сумма всех апостериорных вероятностей для различных заболеваний должна быть равна единице, более точно выражение (11.1) можно записать в виде

$$P(D_j | S) = \frac{P(D_j) P(S | D_j)}{\sum_{j=1}^d P(D_j) P(S | D_j)}, \quad j = 1, 2, \dots, d. \quad (11.2)$$

По существу это формула Бейеса, позволяющая вычислить вероятность справедливости некоторой гипотезы на основании ее априорной вероятности и некоторых эмпирических данных. Этот метод в строгом смысле оправдан, только если альтернативные гипотезы (в данном случае заболевания) имеют априорные вероятности. Однако всегда имеется возможность выбрать модель статистического вывода (см. разд. 3.3), в которой априорные вероятности существуют и задаются соответствующими правилами даже при полном отсутствии информации.

В принципе апостериорную вероятность любого заболевания вычислить совсем нетрудно, так как на основе имеющегося материала можно вычислить как $P(D_j)$, так и $P(S | D_j)$. Однако к величине $P(S | D_j)$ необходимо относиться с известной осторожностью, так как даже при наличии обширного материала он может содержать мало данных о частоте появления определенного симптомокомплекса S . Так, если имеется s различных симптомов и все они являются дихотомическими, то число различных симптомокомплексов будет равно 2^s . При $s = 20$ это число столь велико ($> 10^6$), что применение данного способа оказывается невозможным. Поэтому обычно необходимы некоторые упрощающие предположения. Чаще всего принимается допущение о том, что отдельные симптомы имеют независимые распределения. Это допущение не настолько серьезно, как могло бы показаться. Так, ван Верком

и Бродмен [65] пришли к заключению, что в их материале отсутствует сколько-нибудь значительная корреляция между различными симптомами. Неопубликованные работы, выполненные автором совместно с Дж. Андерсоном, подтверждают эту точку зрения, но, конечно, необходимо исследовать этот вопрос более тщательно.

Если считать, что s симптомов статистически независимы, то можно записать соотношение

$$P(S | D_j) = \prod_{i=1}^s P(S_i | D_j), \quad j = 1, 2, \dots, d \quad (11.3)$$

и на основе имеющегося материала с достаточно высокой точностью вычислить вероятности $P(S_i | D_j)$.

Хотя для простоты мы ввели предположение о дискретности симптомов, это ограничение можно легко ослабить и распространить наш метод на непрерывные величины. Конечно, для описания непрерывных переменных можно было бы задать некоторую «нормальную» область значений этих величин, а все значения, лежащие за пределами этой области, считаются «аномальными». Тогда мы имели бы простую дихотомию типа «наличие» и «отсутствие» признака. Однако при этом теряется большой объем статистической информации, и по возможности нужно использовать непрерывное распределение. Допустим, что k -й симптом является непрерывным с плотностью распределения $f_{kj}(x)$ для j -го заболевания. Это означает, что вероятность наблюдения величины, лежащей в интервале $(x, x + dx)$, равна $f_{kj}(x) dx$. При подстановке этого выражения в выражение (11.3) в формуле для $P(S | D_j)$ появляется сомножитель dx . Этот сомножитель появляется также в числителе и в каждом члене знаменателя правой части выражения (11.2), а поэтому он сокращается и не вызывает никаких осложнений. Таким образом, переход к непрерывным случайным величинам связан лишь с заменой дискретной вероятности соответствующей непрерывной плотностью распределения, вычисленной для рассматриваемой величины.

Конечно, приведенное выше вероятностное описание процесса постановки диагноза крайне упрощенно и в некоторых отношениях, возможно, даже несколько наивно. Тем не менее, как мы вскоре увидим, в ряде частных случаев оно оказывается плодотворным. Рассмотрим теперь одно из наиболее серьезных возражений против этого метода. Дело в том, что врач, сталкиваясь с реальным больным, часто может поставить диагноз на основе самых незначительных симптомов, заметных лишь опытному наблюдателю, не вдаваясь в анализ механизма развития этих симптомов. В то же время при вычислении вероятностей, которое возможно лишь на основе тщательно зарегистрированных данных, теряется мно-

гое из того, что могут дать непосредственные наблюдения. По поводу этого возражения необходимо сделать два замечания. Во-первых, при определенных условиях с помощью математической оценки можно получить такие же (если не лучшие) результаты, к каким врач приходит сам, и, следовательно, при такой оценке учитывается практически все, что имеет диагностическое значение. Во-вторых, если «вычисленные» диагнозы (т. е. диагнозы, поставленные на основании вычисления вероятностей) оказываются менее точными, то это означает, что либо имеется какой-то недочет в самом математическом методе, либо объем обрабатываемой информации значительно меньше, чем при постановке диагноза врачом. В обоих случаях выявление и устранение недочетов даст возможность получить новые знания о процессе постановки диагноза и, вполне возможно, разработать значительно лучший вычислительный метод.

Одним из самых первых и наиболее интересных приложений описанного здесь метода была работа Уорнера, Торонто, Визи и Стефенсона [63], посвященная врожденным порокам сердца. Авторы пользовались перечнем 33 заболеваний и 50 различных дискретных симптомов. Таким образом, исходная матрица «симптомы — заболевания» содержала 1650 элементов. Ясно, что вычисление данных для каждого нового больного с помощью уравнений (11.2) и (11.3) требует большого объема хотя и элементарных, но утомительных и поглощающих много времени арифметических вычислений. Поэтому для этой цели крайне желательно и даже необходимо применять электронную вычислительную машину. В связи с этим возникает важный вопрос использования в будущем вычислительных машин для решения этих задач. Поскольку число поддающихся наблюдению симптомов, и особенно объем лабораторных тестов, непрерывно увеличивается, вскоре может наступить такой момент, когда даже опытный врач не сможет воспользоваться всей той информацией, которую он будет иметь в своем распоряжении. Таким образом, мы сталкиваемся с парадоксальной ситуацией — потенциально у нас есть возможность получить значительно более точные результаты, но на самом деле использовать ее мы не можем и даже оказываемся в худшем положении именно из-за наличия очень большого количества данных (аналогичная ситуация была описана в разд. 4.2). Применение вычислительной техники для хранения и обработки исходных данных, а также для вычисления вероятностей приобретает в этой ситуации кардинальнейшее значение.

По данным Уорнера и др. [63], использование вероятностного метода оказалось весьма успешным. Правильность предположительных диагнозов, поставленных, с одной стороны, врачами-кардиологами, а с другой — с помощью вычислительной машины,

проверялась затем по данным катетеризации сердца или хирургического вмешательства. При анализе 36 случаев оказалось, что вычисленный наиболее вероятный диагноз совпадает с истинным диагнозом не реже, чем заключения, сделанные на основе тех же клинических данных тремя опытными врачами-кардиологами. Эти поразительные результаты убедительно подтвердили возможности математики и вычислительной техники при правильном выборе области их приложения.

К сожалению, Уорнер и его коллеги не подвергли результаты сравнения клинических и вычисленных диагнозов тщательному статистическому анализу. Позже Густафсон и др. (см. [20]) выполнили вполне аналогичную работу на группе из 38 больных с врожденными пороками сердца. В этой работе использовался значительно измененный вариант первоначальной программы для вычислительной машины. Авторы приводят сравнительные цифры, характеризующие правильность диагнозов, поставленных вычислительной машиной и разными врачами на основании перечня симптомов, наблюдавшихся у каждого больного. Шести врачам — двум педиатрам-кардиологам, двум терапевтам-кардиологам, педиатру и врачу-хирургу (специалисту по сердечно-сосудистой хирургии) — предложили составить список всех возможных диагнозов в порядке убывания вероятностей, а программа для вычислительной машины была составлена таким образом, чтобы на печать выдавались четыре наиболее вероятных диагноза (в том же порядке). Полученные результаты, приведенные в табл. 20, доста-

Таблица 20

Сравнение правильности диагнозов, поставленных вычислительной машиной и разными врачами

	Относительное число случаев, %				Среднее число очков	
	правильный диагноз стоит в списке первым	правильный диагноз стоит в списке вторым	правильный диагноз упомянут в числе возможных	правильный диагноз вообще не упомянут		
Вычислительная машина	63	16	18	3	2,39	
Педиатры-кардиологи	1	56	20	8	2,16	
	2	74	5	8	2,39	
	3	36	32	4	28	1,75
Другие врачи	4	52	4	18	26	1,96
	5	36	12	28	24	1,57
	6	48	0	4	48	1,48

точно наглядны. Так, диагноз, выбранный вычислительной машиной в качестве наиболее вероятного, был правильным в 63% случаев, диагноз, выбранный вторым, оказался правильным в 16% случаев, и в 18% случаев правильный диагноз упоминался в числе возможных. Только в 3% случаев правильный диагноз вообще не был упомянут. Каждому врачу и машине ставили по три очка за правильный диагноз, выбранный первым, по два — за правильный диагноз, выбранный вторым, и по одному — за упоминание правильного диагноза.

Хотя можно считать, что в данном случае применение вычислительной машины оказалось довольно успешным, авторы подчеркивают, что опытный педиатр-кардиолог, работающий на своем собственном материале, поставил бы правильный диагноз в 90% случаев. Но при этом следует учесть, что он располагал бы такими материалами, как электрокардиограммы и рентгеновские снимки, которые не были включены в перечень симптомов, так что ни вычислительная машина, ни шесть врачей не могли ими воспользоваться. Авторы отмечают также, что, хотя специалисты, принимавшие участие в эксперименте, и смогли обработать всю полученную ими информацию, менее опытные врачи не справились бы с этой задачей. На основании этих результатов делается вывод, что даже при современном состоянии вопроса математический метод позволяет извлечь из количественных записей столько же (если не больше) информации, сколько ее извлекает самый опытный врач. В то же время в некоторых случаях врачи имеют то преимущество, что они могут использовать материал, который пока еще трудно ввести в вычислительную машину. Очевидно, необходимо приложить все усилия к тому, чтобы найти способы введения в машину этих дополнительных данных.

Очень близкие результаты (главным образом в области кардиологии) были получены несколькими другими исследователями в различных странах (см., например, работу К. Такахаша и др. [60] и статью А. А. Вишневого, И. И. Артоболевского и М. Л. Быковского [68]); все они использовали упоминавшийся нами ранее байесовский подход или какой-либо аналогичный статистический метод. Мы не собираемся обсуждать здесь достоинства и недостатки различных методов, хотя в начале этого раздела были высказаны некоторые замечания по этому поводу и в разд. 11.5 будет рассмотрен другой подход.

Определенные успехи достигнуты также в области идентификации бактерий, выделяемых из исследуемого материала, где используются по существу те же принципы (см. [51]). Так, из существующих данных можно установить частоту обнаружения определенных стандартных признаков у самых разнообразных бактерий. При появлении новой неопознанной культуры производится

ряд исследований и осуществляется предварительная идентификация. Если предварительные данные оказываются недостаточно надежными, производятся дополнительные опыты до тех пор, пока не будет получено приемлемое заключение. В современной микробиологической практике этот процесс обычно основывается на субъективном суждении и личном опыте, и опытный микробиолог может довольно быстро получить правильные результаты. Однако существуют более сложные ситуации, когда результаты лабораторных тестов сильно варьируют. В этих случаях может оказаться целесообразным ставить вероятностный диагноз на основе какой-то ориентировочной группы тестов. Если удовлетворительная дифференциация не достигается, то можно подобрать более действенную группу тестов для дальнейшего исследования материала. Эти идеи столь же важны и в области диагностики, где обычно желательно свести к минимуму объем исследований, требуемых для принятия решения (особенно если некоторые тесты болезненны или даже опасны).

11.4. ИСПОЛЬЗОВАНИЕ НЕБОЛЬШИХ ВЫБОРОК

Во многих случаях альтернативные симптомокомплексы весьма специфичны, и хороший врач может поставить диагноз без большого труда и, более того, довольно быстро. В этих случаях такие сложные средства, как вычислительные машины и вероятностные модели диагностического процесса, практически не нужны, разве что для выполнения научных исследований и обучения. Другими словами, мозг врача, на основе интуиции производящий с достаточной точностью (и не требуя больших затрат) необходимые вычисления, оказывается здесь самой лучшей машиной. Однако существуют и более трудные ситуации. В предыдущем разделе мы уже указывали на сложности, связанные с непрерывным увеличением числа наблюдаемых симптомов и с появлением все более разнообразных методов медицинского обследования. В конце концов само обилие данных вынудит врачей обратиться к автоматическим методам их обработки и интерпретации.

Другая трудность заключается в том, что некоторые заболевания встречались или описывались в прошлом не очень часто, и при появлении нового случая диагноз приходится ставить, исходя из чисто субъективных соображений, поскольку имеющиеся данные крайне недостоверны. Так, изучая материалы по мегалобластической анемии (любезно предоставленные проф. Лесли Уитсом из Радклиффской больницы в Оксфорде), автор и его коллега Дж. Андерсон обнаружили 92 документально описанных случая, на которых можно основываться при постановке диагноза у новых

больных. Всего существует 12 основных форм этого заболевания, однако две из них не были представлены вообще, а четыре другие — всего одним случаем каждая. И только данные по остальным шести формам можно было считать пригодными для постановки диагноза математическими методами. Самая малочисленная из групп, соответствующих этим формам, содержала всего два случая, а самая многочисленная насчитывала пять случаев. Конечно, на практике врач при постановке диагноза новому больному черпает данные о других случаях из журналов и книг. Но если он самостоятельно разрабатывает новые тесты или имеет основания выбрать несколько другую систему классификации, то ему приходится полагаться только на собственный опыт, который неизбежно будет ограниченным. Поэтому крайне важно научиться извлекать максимальное количество информации из малочисленных данных, тем более что они могут иметь очень большое значение с научной точки зрения. Но если имеющиеся выборки данных малы, а статистическая изменчивость значительна, то обычно получение правильных выводов сопряжено с известными трудностями. Именно здесь вероятностный метод может оказаться крайне полезным (см. [10]).

В предыдущем разделе мы предполагали, что данные за прошлое время весьма обширны, и поэтому вполне можно допустить, что вероятности $P(D_j)$ и $P(S | D_j)$ известны с высокой точностью. В этом случае можно использовать формулу (11.1). Но если имеющийся материал не очень велик, а возможно, совсем скуден, то необходимо действовать более осторожно. Прежде всего обозначим его символом R . Теперь $P(D_j)$ есть априорная вероятность того, что случайно выбранный больной страдает заболеванием D_j , определенная *до того*, как будет принята во внимание какая-либо конкретная информация, содержащаяся в исходном материале. Нужно найти вероятность $P(D_j | R)$ того, что некоторый случайным образом выбранный больной страдает заболеванием D_j , *при условии*, что информация, содержащаяся в исходном материале, известна, но симптомокомплекс еще не определен. После этого требуется вычислить вероятность $P(S | D_j, R)$ появления симптомокомплекса S *при условии*, что имеет место данное заболевание и известна информация, содержащаяся в исходных данных. Теперь для определения апостериорной вероятности $P(D_j | S, R)$ того, что больной страдает заболеванием D_j , можно использовать формулу Байеса

$$P(D_j | S, R) \sim P(D_j | R) P(S | D_j, R). \quad (11.4)$$

Это выражение является обобщением соответствующей формулы (11.1). Более точно, мы должны получить формулу

$$P(D_j | S, R) = \frac{P(D_j | R) P(S | D_j, R)}{\sum_{j=1}^d P(D_j | R) P(S | D_j, R)}, \quad j = 1, 2, \dots, d, \quad (11.5)$$

соответствующую выражению (11.2). Наконец, если, как и ранее, все s симптомов взаимно независимы, формулу (11.3) можно заметить соотношением

$$P(S | D_j, R) = \prod_{i=1}^s P(S_i | D_j, R), \quad j = 1, 2, \dots, d. \quad (11.6)$$

Теперь необходимо определить количество информации в материале R . Допустим, что имеется представительный набор случаев с n_j примерами заболевания D_j и что n_{ij} случаев заболевания D_j изучались в отношении симптома S_i . Допустим, что симптомы дискретны и поэтому S_i можно подразделить на l_i классов. Пусть имеется a_{ijk} случаев заболевания D_j с симптомом S_i , относящимся к k -му классу ($k = 1, 2, \dots, l_i$). В случае простой дихотомии $l_i = 2$.

Если нам совершенно неизвестна относительная частота различных заболеваний в той конкретной ситуации, с которой мы встречаемся (например, в определенной больнице или клинике), то естественно принять, что все априорные вероятности $P(D_j)$ равны другу другу, т. е.

$$P(D_j) = 1/d, \quad j = 1, 2, \dots, d. \quad (11.7)$$

Разумеется, при выборе различных значений $P(D_j)$ могла бы учитываться и любая дополнительная информация (помимо материала R). Теперь, используя обобщение правила Лапласа на случай мультиномиальной классификации (см., например, [37], разд. 3.23), находим выражение для вероятности $P(D_j | R)$:

$$P(D_j | R) = \frac{n_j + 1}{\sum_{j=1}^d (n_j + 1)}, \quad j = 1, 2, \dots, d. \quad (11.8)$$

Если окажется, что исходный материал не содержит надежной информации о частоте заболеваний, то необходимо допустить, что все $n_j = 0$. В этом случае выражение (11.8) переходит в (11.7).

Применяя далее это правило, можно получить выражение для вероятности $P(S_i | D_j, R)$. Вероятность того, что симптом S_i у больного с заболеванием D_j относится к k -му классу, равна

$$\frac{a_{ijk} + 1}{\sum_{k=1}^{l_i} (a_{ijk} + 1)} = \frac{a_{ijk} + 1}{n_{ij} + l_i}.$$

Используя формулу (11.6), находим функцию правдоподобия для всего симптомокомплекса:

$$P(S | D_j, R) = \prod_{i=1}^s \frac{a_{ijk} + 1}{n_{ij} + l_i}, \quad j = 1, 2, \dots, d, \quad (11.9)$$

где каждое $k \equiv k_{ij}$.

Теперь, подставляя выражения (11.8) и (11.9) в формулу (11.5), получаем искомое апостериорное распределение различных заболеваний. В принципе эти вычисления довольно просты, хотя они могут быть весьма громоздкими при рассмотрении большого числа симптомов, особенно если перейти от дискретных симптомов к непрерывным.

В разд. 11.3 мы показали, как обращаться с непрерывными случайными величинами в том простом случае, когда исходных данных много; аналогичные идеи можно использовать и здесь. Вначале необходимо определить количество информации в материале R , касающейся какой-либо непрерывной случайной величины. Допустим, что i -я случайная величина непрерывна и распределена по нормальному закону (возможно, после соответствующего преобразования) с математическим ожиданием μ_{ij} и средним квадратическим отклонением σ_{ij} для j -го заболевания. Имеющиеся данные (полученные, скажем, на основании n_{ij} наблюдений) дают эмпирические оценки \bar{x}_{ij} и s_{ij} . Функция правдоподобия $P(S_i | D_j, R)$ в данном случае заменяется оценкой плотности $f(x_i | D_j, R)$, где x_i — наблюдаемая интенсивность i -го симптома у данного нового больного. Эта плотность определяется выражением

$$f(x_i | D_j, R) = \left\{ \frac{n_{ij}}{(n_{ij}^2 - 1)\pi} \right\}^{1/2} \frac{\Gamma\left(\frac{1}{2} n_{ij}\right)}{s_{ij} \Gamma\left(\frac{1}{2} n_{ij} - 1\right)} \times \\ \times \left\{ 1 + \frac{n_{ij}(x_i - \bar{x}_{ij})^2}{(n_{ij}^2 - 1)s_{ij}^2} \right\}^{-1/2 n_{ij}}. \quad (11.10)$$

Вывод этой формулы на основании строгих вероятностных рассуждений можно найти в литературе [29, 37]. Однако форму выражения (11.10) можно было бы легко предсказать, исходя из следующих общих рассуждений. Случайная величина $x_i - \bar{x}_{ij}$ имеет нормальное распределение в повторных выборках с нулевым средним и дисперсией $\sigma^2(1 + 1/n_{ij})$, а s_{ij}^2 — независимая оценка дисперсии σ_{ij}^2 , основанная на материале с $n_{ij} - 1$ степенями свободы. Следовательно, случайная величина

$$t_{n_{ij}-1} = \frac{x_i - \bar{x}_{ij}}{s_{ij}(1 + 1/n_{ij})} \quad (11.11)$$

имеет t -распределение Стьюдента с $n_{ij} - 1$ степенями свободы и, следовательно, ее плотность распределения описывается формулой (11.10). Однако этот вывод необходимо строго проверить, учитывая все входящие в формулу различные условные вероятности.

Если наблюдается какая-то непрерывно распределенная случайная величина, например для k -го симптома, то вероятность того, что наблюдаемое измерение будет лежать в интервале $(x, x + dx)$, можно записать в виде $f_{kj}(x) dx$, где $f_{kj}(x) \equiv f(x_k | D_j, R)$. При подстановке этого выражения в формулу (11.6), как и ранее, в выражении для вероятности $P(S | D_j)$ появляется сомножитель dx . Но если мы переходим к апостериорной вероятности, заданной формулой (11.5), то этот сомножитель появляется как в числителе, так и в каждом члене знаменателя правой части формулы и поэтому сокращается.

Запрограммировать этот математический метод для расчетов на электронной вычислительной машине нетрудно, но, чтобы его можно было с успехом применять в реальных условиях, необходимо обратить особое внимание на ряд моментов. Бейли и Андерсон (неопубликованные данные) проверили этот метод на упоминавшихся ранее данных о мегалобластической анемии, и для очень небольшой группы из девяти больных в семи случаях он себя полностью оправдал. Полученные результаты можно считать обнадеживающими, хотя, разумеется, для того, чтобы сделать окончательный вывод, необходимо провести проверку на значительно большей выборке. Основные трудности вызываются тем, что, когда исходных данных не очень много или даже совсем мало, нередко их можно вообще упустить из виду. Если отсутствует зарегистрированная информация о каком-либо определенном симптоме, связанном со всеми заболеваниями рассматриваемой группы, то никаких проблем не возникает, поскольку в таком случае этот симптом можно просто игнорировать. Но если для одних заболеваний данные о симптоме имеются, а для других таких данных нет, то в этом случае, естественно, необходимо использовать любые возможности дифференциальной диагностики, а не игнорировать данный симптом полностью.

Если рассматриваются дискретные случайные величины, то с этой трудностью справиться легко. Так, допустим, что в исходных данных отсутствует информация об i -м симптоме в связи с j -м заболеванием. Это означает, что a_{ijh} и n_{ij} в соответствующем сомножителе правой части формулы (11.9) равны нулю и остается априорная вероятность, равная $1/l_i$.

Случай непрерывных вероятностей более сложен, так как при отсутствии информации трудно сформулировать априорные утверждения о случайных величинах с бесконечным диапазоном изме-

нений, и неясно, каким образом можно выполнить соответствующие подстановки в формулу (11.5). Эта задача еще не исследована полностью. Один из практических методов состоит в превращении непрерывного симптома, о котором имеется лишь частичная информация, в дискретную форму, что приводит к случаю, рассмотренному в предыдущем абзаце. На самом деле при $n_{ij} < 2$ формулу (11.10) все равно применять нельзя. Вместо этого можно воспользоваться излагаемой в конце следующего раздела идеей об объединении имеющейся информации по всем симптомам в одну переменную [T_j в формуле (11.16)]. Это позволяет избежать трудностей, о которых говорилось при выводе формулы (11.5), хотя справедливость данного метода вызывает некоторые сомнения.

При рассмотрении непрерывных случайных величин возникает еще одно затруднение, состоящее в том, что формулу (11.10) нельзя использовать даже при $n_{ij} \geq 2$, если $s_{ij} = 0$, что имело бы место, если бы все результаты наблюдений были одинаковыми в пределах достижимой точности. Необходимо либо провести более точный анализ, принимая во внимание, что в конечном счете наблюдения имеют дискретный характер, либо применить какой-нибудь особый прием, например использовать взвешенные значения s_{ij} , основанные на данных, полученных по другим заболеваниям (как это предложил Дж. Андерсон), либо действовать так, как если бы данные по всем заболеваниям, для которых $s_{ij} = 0$, отсутствовали (последнее сопряжено, однако, с потерей ценной информации).

До сих пор в большинстве случаев мы предполагали, что при любом заболевании отдельные симптомы статистически независимы. Как указывалось в разд. 11.3, иногда такое упрощающее допущение приближенно выполняется, однако принимать независимость симптомов за общее правило, конечно, нельзя. В принципе необходимо выяснить, к каким последствиям приводит наличие корреляции между различными симптомами. Однако на практике возникают две серьезные трудности. Первая из них состоит в том, что, хотя в упоминавшейся ранее статье Гейссера [29] приводятся результаты, полезные для выполнения соответствующего анализа непрерывных случайных величин, имеющих многомерное нормальное распределение, анализ дискретных случайных величин более сложен. Возможно, здесь подойдет недавно разработанная теория взаимодействий в таблицах сопряженности (см., например, [17, 30]), однако пригодность ее для медицинской диагностики еще не ясна. Вторая трудность более существенна. Она состоит в том, что если имеющиеся материалы недостаточно полны, то они дают очень мало информации о распределениях отдельных компонент многомерной случайной величины, не говоря уже о характере корреляции между ними. В настоящее время мы не располагаем никакими количественными данными о характере или сте-

пени ожидаемых ассоциаций между различными симптомами. Детальное исследование этого вопроса имело бы очень большое значение, так как это позволило бы определить, в каких случаях можно с уверенностью принять допущение о независимости симптомов и какие дополнительные *минимальные* допущения можно принять, когда есть основания предполагать наличие некоторой корреляции. При этом, конечно, нужно было бы стремиться к тому, чтобы число параметров, подлежащих оценке, было сведено к минимуму. Некоторые дальнейшие замечания по этому вопросу приводятся в разд. 11.6.

11.5. КРИТЕРИИ ЗНАЧИМОСТИ

Как при рассмотрении в разд. 11.2 методики распознавания заболевания, так и в случае байесовского подхода, описанного в разд. 11.3 и 11.4, мы предполагали, что каждый больной страдает только одним из нескольких взаимно исключающих заболеваний. Однако нередко у больного бывает не одно, а несколько заболеваний, и это обстоятельство должно быть учтено в любом математическом методе медицинской диагностики. Кроме того, возможно, что у данного больного вообще нет ни одного заболевания из числа рассматриваемых. Но если уж байесовский аппарат пущен в ход, то больного неизбежно отнесут к той группе, вероятность которой является относительно наибольшей, даже если в абсолютном выражении она совсем невелика. В определенной степени эту трудность можно преодолеть, сделав одну из категорий достаточно расплывчатой, например обозначив ее как «прочие». Однако вполне возможен случай, когда заболевание ряда больных просто не попадает в ту ограниченную группу, на которой мы остановили свое внимание. Иногда больной с неопределенным симптомокомплексом первоначально будет совершенно неправильно отнесен к какой-то категории. В этом случае строгий байесовский метод все же припишет ему одно из заболеваний, относящихся к этой категории.

Другой подход состоит в том, чтобы проверить, сходен ли симптомокомплекс, обнаруживаемый у данного больного, с симптомокомплексами, характерными для определенного заболевания. Если сходство отсутствует, а говоря точнее, если различие оказывается статистически значимым, то можно допустить, что у больного этого конкретного заболевания практически нет. Если же окажется, что симптомы, обнаруживаемые у больного, значительно отличаются от симптомов, характерных для всех заболеваний группы, то можно предположить, что у него имеется какое-то совсем иное заболевание, которого раньше не предполагали. Если данные ограничены, то вполне возможно, что установить статистическую

значимость различий не удастся. В этом случае поставить правильный диагноз без дополнительных данных невозможно.

Рассмотрим вначале случай, когда имеются обширные данные за прошлое время и все симптомы независимы и распределены по нормальному закону. Допустим, что, как и ранее, случайная величина x_i , характеризующая i -й симптом, имеет математическое ожидание μ_{ij} и среднее квадратическое отклонение σ_{ij} для j -го заболевания. Тогда при нулевой гипотезе, согласно которой данный больной имеет j -е заболевание, можно проверить значимость отклонения случайной величины x_i от μ_{ij} обычным способом, рассматривая нормально распределенную случайную величину $(x_i - \mu_{ij})/\sigma_{ij}$, имеющую нулевое математическое ожидание и единичную дисперсию. В частности, случайная величина $(x_i - \mu_{ij})^2/\sigma_{ij}^2$ имеет распределение χ^2 с одной степенью свободы. Следовательно, соответствующий критерий для проверки значимости всех s симптомов, рассматриваемых совместно, имеет распределение χ^2 с s степенями свободы:

$$\chi_s^2 = \sum_{i=1}^s \frac{(x_i - \mu_{ij})^2}{\sigma_{ij}^2}. \quad (11.12)$$

Малые значения χ^2 свидетельствуют о близком сходстве между симптомами, наблюдаемыми у больного, и известным перечнем симптомов, характерных для данного заболевания, а при значительных отклонениях появляются большие значения χ^2 . Фактическая значимость проверяется обычным способом.

Допустим далее, что рассматриваются дискретные случайные величины, обнаруживающие простую дихотомию, например «наличие» или «отсутствие» признака. Обозначим вероятность наличия i -го симптома при j -м заболевании через p_{ij} , а его отсутствие — через $q_{ij} = 1 - p_{ij}$. Введем случайную величину X_{ij} , принимающую значение 1 при наличии симптома и значение 0 в его отсутствие. Таким образом, X_{ij} — биномиально распределенная случайная величина, основанная на единичной выборке. Пусть x_i — значение i -го симптома, наблюдаемого у больного. Если бы величина x_i была получена на основе довольно большой выборки объемом n , то для измерения отклонений от математического ожидания можно было бы построить приближенный критерий χ^2 :

$$\chi_1^2 = \frac{(x_i - np_{ij})^2}{np_{ij}q_{ij}}. \quad (11.13)$$

При $n = 1$ маловероятно, что эта аппроксимация окажется удовлетворительной. Однако рассмотрим статистику

$$w_{ij} = (x_i - p_{ij})^2, \quad (11.14)$$

во всяком случае, измеряющую отклонение x_i от математического ожидания. Математическое ожидание случайной величины w_{ij} равно $p_{ij}q_{ij}$, а ее дисперсия составляет $p_{ij}q_{ij}(1-4p_{ij}q_{ij})$. Поэтому мы могли бы использовать статистику z_{ij} , полученную из предыдущей с помощью соответствующей нормировки (т. е. статистику с нулевым математическим ожиданием и единичным средним квадратическим отклонением):

$$z_{ij} = \frac{(x_i - p_{ij})^2 - p_{ij}q_{ij}}{\{p_{ij}q_{ij}(1 - 4p_{ij}q_{ij})\}^{1/2}}. \quad (11.15)$$

Сумма s таких статистик по всем симптомам, т. е.

$$T_j = \prod_{i=1}^s z_{ij}, \quad (11.16)$$

согласно центральной предельной теореме, имеет приближенно нормальное распределение с нулевым математическим ожиданием и дисперсией s . При этом необходимо построить односторонний критерий значимости для положительных отклонений T_j от нуля.

Полученный критерий можно использовать и для непрерывных случайных величин, если заменить выражение (11.12) соответствующим аналогом для непрерывного случая. Каждый член правой части выражения (11.12) распределен как χ^2 с одной степенью свободы, т. е. имеет единичное математическое ожидание и дисперсию, равную 2. Соответствующая нормированная случайная величина z имеет вид

$$z_{ij} = 2^{-1/2} \left\{ \left(\frac{x_i - \mu_{ij}}{\sigma_{ij}} \right)^2 - 1 \right\}. \quad (11.17)$$

Таким образом, в общем случае можно использовать для проверки статистику T_j , заданную формулой (11.16), где члены z_{ij} будут вычисляться по формуле (11.15) для дискретных случайных величин и по формуле (11.17) для непрерывных случайных величин.

На практике должны быть учтены и некоторые дальнейшие осложнения. Прежде всего необходимо рассмотреть дискретные случайные величины, распадающиеся более чем на два класса. Основная проблема здесь состоит по существу в том, чтобы измерить возможную значимость одного наблюдения, полученного из совокупности, имеющей мультиномиальное распределение. Для этого можно приписать случайным величинам X_{ij} значения, определяемые вероятностями самого мультиномиального распределения. Другими словами, если вероятность того, что i -й симптом при j -м заболевании относится к k -му классу, равна p_{ijk} , то случайная величина X_{ij} принимает значение p_{ijk} с вероятностью

p_{ijh} . В случае дихотомии этот способ не отличается от предыдущего, так как мы лишь производим линейное преобразование случайной величины i , как и ранее, завершаем выкладки выводом выражения (11.15). При наличии более чем двух классов нормированная случайная величина z_{ij} , соответствующая формуле (11.15), вычисляется просто, хотя и оказывается несколько громоздкой. По существу, нам необходим показатель

$$z_{ij} = \frac{(x_i - \mu_1'(ij))^2 - \mu_2(ij)}{[\mu_4(ij) - \mu_2^2(ij)]^{1/2}}, \quad (11.18)$$

где $\mu_1'(ij)$ — математическое ожидание, $\mu_2(ij)$ — дисперсия, а $\mu_4(ij)$ — четвертый центральный момент случайной величины x_i при наличии j -го заболевания. Выражение (11.18) можно записать в более развернутом виде, однако вряд ли это целесообразно. На практике все необходимые величины будут, по-видимому, специально вычисляться на электронной вычислительной машине.

Еще одно важное затруднение, как указывалось в разд. 11.4, возникает вследствие относительной неполноты исходных данных. Если рассматриваемые случайные величины непрерывны, то необходимо заменить в формуле (11.11) случайную величину, имеющую t -распределение Стьюдента, соответствующей нормированной случайной величиной. Для дискретных случайных величин при дихотомии необходимо рассмотреть статистику w_{ij} , выраженную формулой (11.14), заменив p_{ij} на a_{ij2}/n_{ij} (в обозначениях, принятых в разд. 11.4), и получить выражение, соответствующее формуле (11.15). Аналогичные видоизменения возможны и для дискретных случайных величин, распадающихся более чем на два класса. Это сопряжено со значительным усложнением алгебраических выражений, и к моменту написания книги необходимые для этой модели расчеты еще не были закончены.

Кроме того, возникает проблема взаимосвязанных симптомов, и при установлении значимости также необходимо преодолевать затруднения, аналогичные тем, о которых шла речь в разд. 11.4.

Существуют также затруднения, вызываемые отсутствием некоторых данных, о чем уже говорилось в предыдущем разделе. Очевидно, что сумма T_j в некотором смысле объединяет всю имеющуюся информацию относительно s рассматриваемых симптомов. Если s велико, то случайная величина $U_j = T_j s^{-1/2}$ имеет приближенно нормальное распределение с нулевым математическим ожиданием и единичным средним квадратическим отклонением, не зависящими от s . Таким образом, мы можем сравнивать, скажем, U_j и U_k , даже если дифференциальные диагнозы j -го и k -го заболеваний должны быть основаны на различных совокупностях

симптомов, т. е. на различных значениях s (например, вследствие того, что некоторые вероятности p_{ij} или их оценки отсутствуют).

Далее, можно рассматривать функцию правдоподобия каждого наблюдаемого значения U_j как замену функции правдоподобия $P(S | D_j)$ или $P(S | D_j, R)$. Поэтому мы имеем возможность использовать в формулах (11.2) или (11.5) правило Бейеса даже в тех случаях, когда данные по некоторым симптомам отсутствуют или недостаточно полны (при условии, конечно, что байесовский подход удовлетворителен в остальных отношениях). Логическая обоснованность этого предложения не вполне безупречна, однако оно заслуживает дальнейшего рассмотрения.

Рассмотренные в данном разделе методы проверки значимости позволяют подойти к исследованию проблемы постановки множественного диагноза. Что станет с предложенными критериями, если у больного два или больше заболеваний? На первый взгляд этот вопрос кажется в значительной степени эмпирическим. Если, например, у больного имеется j -е заболевание, то вероятность появления i -го (дискретного) симптома равна p_{ij} , а если у него есть k -е заболевание, то соответствующая вероятность равна p_{ik} . Какова в этом случае вероятность появления i -го симптома, если у больного имеются оба эти заболевания? И так далее для всего симптомокомплекса. Этот вопрос очень важен, и для его решения требуется провести специальное эмпирическое исследование.

11.6. ЗАКЛЮЧЕНИЕ

В этой главе мы лишь поверхностно ознакомились с возможностями математических и вычислительных методов в медицинской диагностике. Результаты применения этих методов для диагностики тиреотоксикоза (разд. 11.2) ясно показывают, что в определенных случаях даже самый простой математический подход, бесспорно, полезен. Более сложные методы, описанные в разд. 11.3 и 11.4, оказались весьма перспективными уже на самом начальном этапе научных исследований и разработок. Надо думать, что дальнейшая работа в этом направлении приведет к значительному расширению возможностей быстрой и точной оценки диагностического материала.

Для того чтобы такими методами можно было пользоваться повседневно, необходимо найти удовлетворительное решение ряда проблем, которых мы уже касались. Это проблемы, возникающие из-за относительной неполноты и малочисленности исходных данных, возможной взаимосвязи различных симптомов, наличия нескольких заболеваний у одного лица и необходимости во многих случаях начинать лечение до постановки окончательного диагноза.

Возьмем, к примеру, проблему взаимосвязанных симптомов. Удовлетворительной модели, которая учитывала бы корреляцию, или взаимосвязь дискретных и непрерывных переменных, у нас еще нет. И даже если бы такая модель была разработана, то неясно, как можно было бы оценить все ее параметры при наличии лишь ограниченных исходных данных. Может показаться, что вообще невозможно располагать таким объемом статистической информации, который был бы достаточен для принятия практических решений. Однако, как известно, существует множество прекрасных врачей-диагностов, и это показывает, что количество «внутренней» информации у врача несколько больше, чем можно было бы ожидать. Ответ, по-видимому, состоит в следующем. Ставя диагноз, врач на основании своих знаний и опыта всегда подсознательно принимает ряд допущений, которые на математическом языке равносильны допущениям о наличии определенной скрытой структуры взаимосвязи симптомов. Таким образом, необходимы специальные исследования, направленные на изучение этой структуры в явном виде. Предположим, окажется, что определенные симптомы обнаруживают довольно постоянную взаимосвязь (возможно, обусловленную какими-то физиологическими причинами) независимо от характера заболевания. Тогда соответствующие параметры можно было бы оценить на основе обширных исходных данных, а затем использовать их, когда имеющиеся данные скудны.

Разумеется, важно исследовать реальные проблемы, а не чисто математические абстракции. В связи с этим встает общий вопрос о том, какие математические модели процесса постановки диагноза можно считать действительно приемлемыми. В этой главе мы рассмотрели ряд точных, хотя и чрезмерно упрощенных математических моделей того, что делают (или пытаются делать) врачи при постановке диагноза. Для построения более реалистичных моделей необходимо гораздо глубже изучить процессы постановки диагноза в клинической практике. Как мы видели, вычисления, основанные на самых простых идеях, нередко позволяют получить такие же (если не лучшие) результаты, как и те, которых достигает хороший врач на основании того же перечня симптомов. Но если врач непосредственно общается с больным, то он может поставить диагноз значительно точнее, так как в этом случае он получает дополнительную информацию, которой не имеет вычислительная машина. Однако применение современных методов автоматической записи и анализа клинических данных (например, физиологических параметров) или разрабатываемых сейчас машин, способных к распознаванию образов (применяемых, например, при подсчете и классификации хромосом), позволит, вероятно, быстрее и полнее обрабатывать большой объем информации,

чем это может сделать врач. Если это осуществится, то обработку диагностического материала можно будет производить на машине с такой скоростью и эффективностью, которые сейчас даже невозможно себе представить.

Самая насущная задача заключается сейчас в проведении комплексных исследований всех аспектов процесса постановки диагноза в целом, с тем чтобы математические методы и методы составления программ для ЭВМ были тесно увязаны с конкретными клиническими задачами. Только в том случае, если результаты этих научных исследований смогут выдержать длительную проверку практикой, можно надеяться, что задачи, описанные в данной главе, будут выполнены.

Глава 12

ИССЛЕДОВАНИЕ ОПЕРАЦИЙ В МЕДИЦИНЕ

12.1. ВВЕДЕНИЕ

Медицина — это наука, целиком направленная на оказание помощи людям. Главные персонажи здесь — врач и больной; весь смысл работы врача заключается в том, чтобы облегчить страдания больного. Хотя медицинские познания и способности врача — это важнейший фактор, определяющий результаты лечения, они тесно связаны с широким кругом других видов человеческой деятельности — с рядом теоретических и прикладных наук, техникой, экономикой и социологией, а также с решением сложных юридических, моральных и этических проблем. Теоретически возможности новых достижений в медицине неограниченны, однако на практике обычно ощущается нехватка врачей и медицинских сестер, недостаток лекарств, помещений, финансов и т. д. В связи с этим возникает множество неотложных проблем, решение которых позволило бы использовать имеющиеся ограниченные ресурсы с максимальной эффективностью. Эти проблемы относятся к области исследования операций, и в настоящее время важность этого предмета для медицины в целом получает все большее признание.

Общие цели исследования операций и объекты, изучаемые этими методами, уже рассматривались достаточно подробно в гл. 4. В настоящей главе мы обсудим применение этих идей к некоторым вопросам, связанным с организацией и планированием медицинского обслуживания. В этой области применение самых простых методов очень часто позволяет добиться значительных успехов. Таким образом, сложные математические методы (подобные тем, о которых говорилось в разд. 4.3) требуются далеко не всегда, но в определенных ситуациях они могут быть очень полезны.

Получить точное и удовлетворительное решение многих важных задач организационного характера совсем нелегко, хотя формулировка их крайне проста. Например, насколько большой должна быть больница? Где она должна располагаться? Как она должна быть спроектирована и каким образом ею нужно руководить? Каким образом больница будет связана с врачами общей практики и лечебными учреждениями, находящимися в ведении местных властей? Органам здравоохранения приходится непрерывно при-

нимать решения по вопросам такого рода при минимуме количественных данных, и получаемые результаты нередко оказываются неудовлетворительными. Одной из задач исследования операций является построение реалистичных количественных моделей всех этих сложных процессов с тем, чтобы административные решения принимались на основе более точных данных, чем в отсутствие таких моделей. Как мы увидим далее, нередко это означает получение некоторого способа прогнозирования (обычно в вероятностном или статистическом смысле) возможного результата, к которому приведет выбор одного из альтернативных решений. Конечно, ни в коем случае это не исключает суждений, основанных просто на здравом смысле. Но заблаговременное знание конкретных последствий различных линий поведения помогает действовать более точно.

Конечная цель состоит в том, чтобы увеличить объем и повысить качество медицинского обслуживания и добиться того, чтобы каждому больному уделялось столько внимания, сколько требуется. Исследование операций направлено на то, чтобы административные органы, пользуясь более точным научным подходом в тех случаях, где это возможно, могли быстрее достигнуть этой цели. Посмотрим теперь, как это делается в некоторых конкретных областях.

12.2. ПОТРЕБНОСТЬ В МЕДИЦИНСКОМ ОБСЛУЖИВАНИИ

Одна из самых важных проблем обеспечения населения медицинской помощью заключается в том, чтобы с достаточной точностью оценить требуемое количество персонала и помещений. Следует различать такие основные виды обслуживания, как общая практика, амбулаторное лечение, лечение в стационаре, оказание помощи на дому, а также другие виды медицинской помощи. Как правило, необходима также классификация по медицинским специальностям, особенно для госпитализируемых больных. Не следует забывать и о таких специализированных учреждениях, как родильные дома, дома для престарелых и психиатрические клиники. В каждой из этих областей существуют свои собственные сложные проблемы. Многие области в известной степени перекрываются, и полный анализ структуры «всей системы» пока еще не предпринимался. Естественно, что детали структуры и организации медицинского обслуживания в разных странах различны, но общие цели и задачи примерно одни и те же.

Рассмотрим в качестве примера задачу вычисления коечного фонда, необходимого для обслуживания определенной группы населения. Первое исследование фактической потребности в медицинском обслуживании было выполнено в 1955 г. Наффилдским

медицинским центром; целью этой работы было отыскание соотношения между требуемым числом коек и той действительной численностью обслуживаемого населения, среди которого могут оказаться больные, подлежащие госпитализации. При исследовании проводилась классификация по специальностям и были выбраны те районы, где условия жизни считаются достаточно хорошими.

До выполнения этой работы необходимое число коек оценивалось (и, к сожалению, еще слишком часто оценивается) чисто предположительно, разумеется, после тщательного рассмотрения всех существенных факторов. Обычно планирующие органы устанавливают такие цифры, как 200, или 450, или 700 коек для тяжело больных на 100 000 человек, фактически совершенно произвольно. Получаемые оценки колеблются в очень широких пределах — в отдельных случаях они различаются на целый порядок. Так, оценки необходимого числа коек для больных с патологией органов грудной полости колеблются от 10 до 120 на 1 млн. человек. Хотя определенные цифры, которые иногда считаются общепризнанными, устанавливаются после внимательного исследования, надлежащая рациональная основа для их получения, к сожалению, отсутствует. Другими словами, отсутствует приемлемая математическая модель, и нет способа, с помощью которого можно было бы получить истинные статистические оценки и их средние квадратические ошибки. Следовательно, трудно определить, каким образом можно использовать такие численные данные при решении количественных задач, например при планировании строительства больниц в новом городе, где ряд существенных факторов может оказаться совершенно иным, чем в районах, охваченных предыдущим обследованием.

Первым этапом указанной выше работы явилось исследование фактически обслуживаемых групп населения. Хотя органы здравоохранения (во всяком случае, в Великобритании) работают в пределах четко определенных географических районов, где численность населения и имеющиеся лечебные учреждения известны, фактическая картина возможностей госпитализации определяется значительно более сложными факторами. Так, больные, нуждающиеся в госпитализации, могут по тем или иным причинам стремиться попасть в определенное лечебное учреждение. Это частично зависит от их личного желания, а частично от мнения врача общей практики, к которому обращается больной. Обычно больные стараются пойти в ближайшую больницу, даже если она находится в другом административном районе. Исследования показывают, что данная больница (или группа близко расположенных больниц) принимает около 100% больных, нуждающихся в госпитализации и живущих в непосредственной близости от этого лечебного

учреждения. По мере удаления от данной больницы (или группы больниц) этот процент постепенно уменьшается и практически становится равным нулю в окрестности соседней группы больниц. Посредине между двумя большими группами больниц процент случаев госпитализации составляет около 50. Таким образом, этот показатель непрерывно изменяется с расстоянием независимо от того, где проходят границы административных районов.

Действительную численность населения, обслуживаемого какой-то определенной группой больниц (будем называть этот показатель *эффективной* численностью), можно вычислить следующим образом. Берется список больных, поступивших в эти лечебные учреждения, скажем за целый год, а затем на основании адресов больные распределяются по географическим районам, откуда они прибыли, например из городских и из сельских районов. Вообще говоря, это позволяет раскрыть только часть потребностей какого-либо определенного района, поскольку некоторые больные могли обратиться в другие больницы. Поэтому такой анализ повторяется для всех тех групп больниц, которые находятся в окрестности данного района и могли привлечь значительное число больных, в противном случае поступивших бы в больницы, выбранные для исследования. Таким образом, для i -го административного района можно выделить две группы больных: a_i — больные, поступившие в основную исследуемую на данной территории группу больниц, и b_i — остальные больные. Если данные переписи показывают, что в определенном районе проживает N_i человек, то эффективную численность населения данного района, обслуживаемого группой расположенных на его территории больниц, можно оценить как $a_i N_i / (a_i + b_i)$. Суммирование по всем территориям, расположенным как внутри, так и за пределами данного административного района, дает оценку общей эффективной численности обслуживаемого населения:

$$M = \sum_i \frac{a_i N_i}{a_i + b_i} \quad (12.1)$$

Нетрудно найти среднюю квадратическую ошибку для M , которая в известной степени зависит от принятых допущений относительно характера выборочной процедуры. Если a_i и b_i можно рассматривать как случайные величины, распределенные по закону Пуассона (что во многих случаях вполне допустимо), то легко видеть, что дисперсия случайной величины M по большой выборке имеет вид

$$\sigma^2(M) = \sum_i \frac{a_i b_i N_i^2}{(a_i + b_i)^3} \quad (12.2)$$

(в предположении, что значение N_i известно более или менее точно).

Применение этих формул к материалу, собранному в районе Норфолка и Норвича за 1951 г., показало, что эффективная численность составляла $484\,600 \pm 800$ человек. Эта оценка была получена на основе анализа 41 000 адресов больных, прибывших примерно из 60 городских и сельских районов. Фактическая численность всего населения этой административной единицы составляла 546 600 человек, что почти на 13% превышает эффективную численность. В некоторых случаях расхождение может быть еще большим, и в значительной мере оно зависит от местных условий.

Получив оценку численности населения, потенциально нуждающегося в госпитализации, рассмотрим теперь определение фактической потребности в стационарном лечении по разным медицинским специальностям для основной группы исследуемых больниц. Сделать это совсем несложно, так как потребность в медицинском обслуживании можно оценить на основании числа больных, поступивших в стационар за год (или числа выписавшихся и умерших) с учетом изменений, которые произошли в списках очередности, если таковые ведутся. При условии, что эти списки ведутся правильно и регулярно уточняются, их увеличение указывает на возникновение дополнительных потребностей в медицинском обслуживании, и число больных, поступивших в стационар, должно быть увеличено, а в противном случае его нужно уменьшить.

Если среднесуточная потребность в лечении по данной специальности составляет d больных и средняя длительность пребывания их в стационаре равна s суток, то следует ожидать, что потребное число койко-дней составит

$$c = ds. \quad (12.3)$$

Хотя это выражение приближенно справедливо, необходимо сделать одну важную оговорку. При значительной потребности в стационарном лечении в отделении, имеющем ограниченную пропускную способность, вначале возможно образование очереди. Как уже указывалось в разд. 4.3, в этом случае важно, чтобы загрузка системы была меньше единицы, иначе время ожидания будет очень велико (теоретически оно будет равно бесконечности). На практике это означает, что пропускная способность больницы должна несколько превосходить так называемое *критическое* значение, вычисляемое по формуле (12.3). Требуемое дополнительное число коек может быть совсем небольшим (см. [4]); покажем это на примере самой простой модели.

Пусть в отделении имеется k коек, а длительность пребывания в больнице испытывает значительные колебания и является слу-

чайной величиной, распределенной по экспоненциальному закону

$$f(v) = \frac{e^{-v/s}}{s} dv, \quad 0 \leq v < \infty \quad (12.4)$$

(средняя длительность пребывания в стационаре, как и ранее, составляет s суток). Допустим также, что суточная потребность в числе коек имеет пуассоновское распределение с математическим ожиданием d . Тогда загрузка системы будет равна

$$\rho = \frac{sd}{k}. \quad (12.5)$$

Эта величина характеризует отношение требуемой пропускной способности к фактической (если не учитывать дополнительных практических трудностей административного характера). Если при $\rho < 1$ процесс достигает равновесного состояния, то можно показать, что среднее время w , в течение которого больной должен ожидать госпитализации, имеет вид

$$w = \frac{s}{(1-\rho)k} \left[1 + \frac{(1-\rho) \{1 - P(k, sd)\}}{p(k, sd)} \right]^{-1}, \quad (12.6)$$

где $p(x, m)$ — члены, а $P(c, m)$ — суммы членов ряда Пуассона (для вычисления которых имеются очень удобные таблицы, составленные Молина [45]).

Рассмотрим, например, отделение больницы в случае, когда $d = 1$ и среднее время пребывания в стационаре $s = 20$ суток. Допустим, что в отделении имеется 25 коек, но, учитывая административные соображения, при наличии списка очередности можно держать занятыми только 22 койки. Таким образом, $k = 22$ и $\rho = 0,91$. С помощью формулы (12.6) можно легко найти, что $w = 5,7$ суток. Во многих случаях можно считать, что это довольно короткое время ожидания (за исключением неотложных случаев, которые следует рассматривать отдельно). Обычно необходимо добавить одну-две койки сверх критического значения.

Для оценки критического числа c на основе фактических данных нужно знать среднюю квадратическую ошибку. Если d оценивается по числу n вновь поступающих больных (часть которых будет помещена в стационар, а часть останется в списке очередности) за определенный промежуток времени и если s вычисляется по n' наблюдениям, то дисперсия случайной величины c , взятая по большой выборке, будет иметь вид

$$\sigma^2(c) = c^2 \left(\frac{1}{n} + \frac{1}{n'} \right). \quad (12.7)$$

Наконец, оценка критического числа коек в пересчете на одного человека имеет вид

$$R = c/M, \quad (12.8)$$

где

$$\sigma^2(R) = R^2 \left\{ \frac{\sigma^2(c)}{c^2} + \frac{\sigma^2(M)}{M^2} \right\} \quad (12.9)$$

находится путем совместного использования формул (12.2) и (12.7). Итак, получена статистически обоснованная оценка величины R и соответствующая средняя квадратическая ошибка. Это означает, что дальнейшее обсуждение этого вопроса можно поставить на значительно более прочную количественную основу. Например, проектируя строительство больниц в новом городе, можно на основании значений R для разных специальностей и предполагаемой численности населения в городе, N_0 , оценить соответствующее критическое число как RN_0 .

Разумеется, при таком подходе должна существовать достаточная уверенность в том, что в новом городе условия жизни населения и условия работы больницы будут близки к тем, для которых было получено первоначальное значение R . Обсуждение различных осложнений этой проблемы увело бы нас слишком далеко в сторону. Достаточно сказать, что мы должны быть готовы произвести некоторые изменения в приведенных выше формулах. Например, можно предположить, что в новых условиях средняя потребность в госпитализации будет примерно такой же, как и в остальных местах, но медицинский персонал сочтет нужным изменить среднюю длительность пребывания в больнице s . В этом случае в формулу (12.3) или (12.8) нужно будет подставить новое значение s вместо того значения, которое вытекало из обследованного материала.

Одно из обычных возражений против такого подхода состоит в том, что в лучшем случае мы измеряем лишь текущий спрос на госпитализацию вместо реальной потребности. Это, конечно, верно, но до тех пор пока не появится возможность лучше оценивать реальные потребности, руководители будут планировать только удовлетворение постоянно растущего спроса. В то же время некоторые исследователи (см., например, [26]) отмечают явление обратного порядка — «клинически необоснованную госпитализацию», приводящую к увеличению наблюдаемого спроса. Обычно это относится к тем случаям госпитализации, которые обусловлены влиянием социальных и экономических факторов. Если обеспечить соответствующее лечение в домашних условиях, то, естественно, потребность в стационарном лечении уменьшится.

Описанные выше основные соотношения между спросом и предложением остаются справедливыми даже в том случае, если будут учтены указанные нами осложнения. Однако, желая составить прогноз для какой-либо конкретной больницы и группы населения, необходимо решить, какими именно могут быть значения d и s в этом случае. Зная эти значения, можно вычислить соответ-

ствующую критическую величину s и, следовательно, предсказать, насколько удовлетворительным будет то или другое запланированное число коек.

В некоторых случаях основную проблему планирования можно разрешить значительно проще. Допустим, например, что принимается перестройка, переоборудование или реорганизация существующей больницы и поэтому требуется заново оценить фактический спрос с целью определения реальной потребности в помещениях. Если при некоторых незначительных изменениях и уточнениях можно принять существующую потребность за заданную, то не имеет смысла производить обширные исследования фактической потребности в стационарном лечении для определения эффективной численности обслуживаемого населения. Нужно только знать, сколько коек необходимо для обслуживания определенного числа жителей. Такая задача решалась в 1960 г. в связи с переоборудованием Оксфордской объединенной группы больниц. Было известно, что население, обслуживаемое этим лечебным учреждением, рассредоточено на территории обширного географического района (чтобы провести исследование фактической потребности в стационарном лечении, потребовалось бы обследовать накопленный материал более чем по 400 лечебным учреждениям). К тому же не было ясности, с какими другими районами можно было бы обоснованно сравнивать Оксфордскую группу больниц. Кроме того, спрос по некоторым медицинским специальностям был особенно велик и, помимо всего прочего, возникало много споров вокруг клинически обоснованной длительности пребывания в больнице.

Рассмотрим, например, приведенные в табл. 21 данные о потребности в госпитализации по отделению общей хирургии за 5 лет, предшествовавших этой реорганизации.

Таблица 21

**Потребность в госпитализации по отделению общей хирургии
в Оксфордской объединенной группе больниц**

Год	Число выписавшихся	Число ожидающих на 31 декабря	Изменения в списке ожидающих	Общая потребность	Среднее время пребывания в больнице, сутки
1955	5176	1176	+193	5369	9,2
1956	5256	887	-289	4967	9,2
1957	5260	1163	+276	5536	8,9
1958	5459	1413	+250	5709	9,8
1959	5583	1650	+237	5820	9,8

Суммарная фактическая потребность в 1959 г. составила 5820 человек — совершенно независимо от эффективной численности обслуживаемой группы населения. При средней длительности пребывания в больнице, равной 9,8 суток, и среднем времени простоя койки, равном 2,2 суток (это среднее время, в течение которого койка остается свободной после выписки больного), критическое значение необходимого числа коек будет равно

$$c = \frac{5820}{365} \cdot (9,8 + 2,2) = 191.$$

Так как в среднем имелось всего 183 койки, то ясно, что в рассматриваемое время загрузка несколько превышала единицу. Поэтому очередь должна была быть непрерывно расти, даже если бы не происходило заметного увеличения общего спроса.

Очевидно, что во всех этих вычислениях важнейшую роль играет средняя длительность пребывания в больнице. Разумеется, фактическое время пребывания в больнице каждого отдельного больного может колебаться в широких пределах, однако средняя продолжительность зависит частично от состояния больного, частично от системы лечения, а частично от имеющихся возможностей. Поэтому среднее время пребывания в больнице можно несколько изменять или уточнять, смотря по обстоятельствам. Так, хирурги или терапевты могут считать, что фактическое среднее время пребывания в больнице недостаточно, но обусловлено нехваткой коек. И наоборот, при наличии соответствующих условий для домашнего лечения можно было бы исходить из несколько меньшего времени пребывания в больнице. В Оксфордской группе больниц хирурги полагали, что среднее время пребывания в больнице, равное 10 суткам, слишком мало. Большинство предпочитало увеличить его до 15 суток, хотя некоторые считали, что достаточно 12 суток. Какой бы срок ни был выбран, не представляет труда вычислить соответствующее критическое значение числа коек. Так, при длительности пребывания в больнице 12 суток и времени простоя койки, равном двум суткам, для потребности 1959 г. критическое значение c составляло

$$\frac{5820}{365} \cdot (12 + 2) = 223 \text{ (койки)},$$

т. е. значительно больше числа коек, фактически имевшихся в наличии в то время.

Разумеется, такие вычисления не дают точных окончательных оценок потребностей. Однако при некоторых конкретных допущениях они позволяют определить последствия, автоматически вытекающие из этих оценок. Это значительно облегчает планирование, так как возможность заранее рассчитать результаты от внедрения той или иной схемы дает численную основу для дальней-

ших рассуждений и описаний. Рассмотрим вновь данные, приведенные в табл. 21. Допустим, что потребность будет ежегодно увеличиваться примерно на 2%. Каким образом найти какое-либо удовлетворительное решение? Одна из многих возможностей состоит в следующем: произвести вычисления на основании той средней длительности пребывания в больнице, которая необходима по мнению большинства хирургов, т. е. 15 суток (при этом критическое число коек будет равно 272), а затем считать этот показатель равным 12 суткам. Это даст возможность обслужить за год

$$\frac{272 \cdot 365}{(12 + 2)} = 7100 \text{ больных}$$

и практически избавиться от очереди. После этого можно было бы установить среднее время пребывания в больнице равным 15 суткам, но в случае продолжающегося увеличения потребности в госпитализации постепенно несколько уменьшать его, чтобы предупредить повторное образование большой очереди.

Хотя основные идеи такого рода расчетов изложены здесь весьма элементарно (более детальное изложение и библиографию можно найти в литературе [8]), они все же показывают, каким образом простая математическая модель может быть использована для реалистического решения многих важных общих вопросов, связанных с обслуживанием стационарных больных.

12.3. ПРОЕКТИРОВАНИЕ БОЛЬНИЦ

Предыдущий раздел был посвящен главным образом определению общей потребности в средствах для стационарного лечения, выраженной через число коек по различным медицинским специальностям. Допустим теперь, что эта задача решена применительно к новой больнице, строительство которой только намечается. Каким образом спроектировать здание так, чтобы деятельность всей сложной системы больничных служб была максимально эффективной? Во многом это по существу архитектурная задача. В принципе заказчик указывает, какие именно цели он ставит перед собой, а архитектор предлагает один или несколько возможных проектов, в которых надлежащим образом должны быть учтены эстетические, экономические, технические и функциональные требования. На практике больничный комитет может изучить в чертежах проекты зданий, представленные различными архитекторами, и выбрать тот из них, который, по их мнению, является наилучшим. Никто не сомневается в том, что обычно заказчик знает, что именно ему нужно. Трудность состоит в том, чтобы перевести общую цель эффективного оказания медицинской

помощи в такие категории, на которых можно было бы основывать выбор конкретных деталей архитектурных проектов.

По-видимому, еще далеко то время, когда можно будет математически рассчитать полный проект здания, наиболее эффективно удовлетворяющего определенным требованиям. Однако рассчитать некоторые последствия принятия любого данного проекта и, следовательно, провести количественное сравнение разных проектов можно уже сейчас. Таким образом, в некоторых вопросах, где раньше можно было высказывать только догадки, мы теперь можем пользоваться численными критериями. Такой подход позволит заказчику более четко формулировать основные требования к проекту больницы, а архитекторам — оценивать, насколько будут приемлемы их предварительные планы. Трудно предвидеть, каким путем пойдет развитие в дальнейшем, однако ясно, что пока достигнут лишь первый этап в разработке научных методов проектирования зданий, базирующийся на значительной более количественной основе.

Рассмотрим вначале ряд проблем, связанных с проектированием отделения больницы, поскольку их можно рассматривать независимо от общего проекта больницы в целом. Большая часть основной работы первоначально была выполнена Наффилдским медицинским центром [48]; с вопросами статистики при проектировании больниц можно познакомиться в одной из работ автора [5].

Одно из самых важных требований, предъявляемых к отделению, — его относительная компактность. Если оно слишком длинно и различные вспомогательные помещения расположены на одном его конце, то медицинским сестрам и санитаркам придется совершать много ненужных переходов. Относительные достоинства различных проектов нельзя обнаружить непосредственно путем простого наблюдения, так как маршруты движения медицинских сестер от койки к койке, между палатой и сестринским постом, изолятором и т. д. также значительно влияют на общую длину проходимого ими расстояния. Один из способов измерения этой деятельности, которую никогда по-настоящему не проверяли, состоит в определении средней типичной схемы движения между всеми основными элементами отделения; для этого необходимо провести обследование в целой группе больниц, выбранной надлежащим образом. На основании такой типичной схемы для отделения с определенной специализацией можно будет установить среднее число переходов, скажем, между сестринским постом и палатами за сутки или за смену. Возможно, потребуется также описать более точно схему движения персонала внутри одной палаты между различными койками. Какой бы метод получения стандартной схемы ни был выбран, ее можно будет использовать

для различных проектов отделения и получить показатель его относительной компактности. Разумеется, этому показателю нельзя будет придавать решающее значение, однако он явится одним из численных параметров, которые должны приниматься во внимание. При прочих равных условиях необходимо выбирать наиболее компактный проект, обеспечивающий минимальную дальность передвижения за время дежурства.

Существуют, однако, пределы целесообразной компактности. В частности, недопустима скученность, которая способствует передаче инфекции; важно также наличие свободного пространства вокруг коек, так как иначе медицинский персонал не будет иметь возможности выполнять свои обязанности; все эти соображения необходимо учитывать при планировании расстояния между койками. Количественный показатель связи между скученностью и передачей инфекции еще не разработан, но учет потребностей медицинского персонала — это вполне выполнимая задача. В упоминавшемся здесь исследовании, проведенном Наффилдским центром, различные операции, выполняемые медицинскими сестрами, были засняты на киноплёнку. Полы в палате были разбиты на квадраты со стороной 30 см, а над каждой койкой были установлены часы. После этого не представило труда построить кривую плотности распределения доли времени, в течение которого персоналу требовалось различное пространство вокруг коек. Так, при заправке постели 86% времени санитарка находится не далее 60 см от койки, а 98% времени — не далее 75 см. При более сложных процедурах, например при внутривенных вливаниях или при усаживании больного в кресло-каталку, требуется больше места, но и в этих случаях персонал не выходит за пределы 1,2 м. В некоторых частных случаях требуется дополнительное пространство. Так, для установки кислородной палатки над постелью больного нужны две медицинские сестры и две санитарки. Около 28% времени вся эта операция производится на расстоянии от койки, превышающем 1,2 м, но так как вся она занимает меньше минуты, ясно, что детальный статистический анализ здесь не нужен. Практически расстояние между койками 1,2 м вполне достаточно и удобно. Лишь изредка такое пространство может оказаться для медицинской сестры тесноватым. Если принять ширину койки равной 90 см, то ясно, что достаточное расстояние между серединой коек составляет 2,1 м.

В проекте отделения следует учесть также необходимость в таких помещениях, как туалеты, ваннные комнаты и холлы. Потребность в этих помещениях в значительной степени зависит от числа ходячих больных и больных, находящихся не на строгом постельном режиме, а это в свою очередь зависит от специализации отделения, а также от характера лечебного учреждения. Оче-

видно, что эта величина значительно варьирует ото дня ко дню, однако некоторая численная информация о ее статистическом характере была получена следующим образом. В течение определенного промежутка времени врач обследовал больных в отделениях различного профиля и не только регистрировал число ходячих больных и больных со строгим и нестрогим постельным режимом, но и оценивал, какой была бы численность этих трех категорий больных, если бы: 1) была принята традиционная система ведения больных и 2) была принята система, предусматривающая раннюю отмену постельного режима. Естественно, что многие из этих оценок весьма субъективны, однако численные данные для больных хирургического отделения, приведенные в табл. 22, весьма наглядны.

Таблица 22

Относительное число ходячих больных хирургического отделения при различных системах ведения больного

Система ведения больного	Относительное число больных, %		
	со строгим постельным режимом	с нестрогим постельным режимом	ходячих
Традиционная	67	17	16
Фактически наблюдавшаяся	50	20	30
Ранняя отмена постельного режима	16	35	49

Разница между процентом больных со строгим постельным режимом при традиционной системе и процентом, ожидаемым при ранней отмене постельного режима, очень значительна, и учет этого обстоятельства может вызвать весьма серьезные изменения в проекте отделения. Для того чтобы использовать эти данные при проектировании, нужно было бы вычислить предполагаемое число больных, которые будут находиться в различных помещениях данного отделения. Чтобы сделать это правильно, потребуются уточненные материалы обследований, а также информация о том, какой принцип ведения больных будет принят. Конечно, могут возникнуть различные неясности и трудности, но главное — это то, что при наличии такого количественного справочного материала архитекторам и врачам будет значительно легче найти приемлемое решение.

Еще одним аспектом, который следует учитывать при разработке отделения, является число коек, которое необходимо уста-

новить в изоляторах, в отличие от числа коек в общих палатах. Необходимость в нескольких однокоечных палатах обусловлена целым рядом причин. Например, у больного может быть инфекционное заболевание или он обладает повышенной восприимчивостью к инфекции; бывают также очень беспокойные больные или больные с повышенной чувствительностью к внешним раздражителям. Потребность в изоляторах также сильно варьирует, однако очень важно определить их оптимальное число для каждого отделения. Выявление важнейших особенностей этой задачи требует более детального математического анализа, который, оказывается, возможен. Этот анализ позволяет проиллюстрировать нередко возникающее затруднение общего характера, а именно столкновение между двумя противоречивыми требованиями. Очевидно, что если бы все койки были установлены в отдельных палатах, то потребность в таких палатах всегда удовлетворялась бы полностью (за исключением того редкого случая, когда все отделение переполнено больными, остро нуждающимися в изоляции). Однако при такой проектировке изоляторы значительную часть времени не будут использоваться по своему прямому назначению. Это может оказаться неэкономичным и привести к определенным неудобствам для медицинского персонала. В то же время если в проекте будет предусмотрено лишь очень небольшое число отдельных палат, то они почти всегда будут использоваться по назначению, однако значительная часть потребности останется неудовлетворенной. Ясно, что необходимо найти какое-то компромиссное решение.

Для этого можно провести в каких-либо других больницах обследование отделения такого же профиля, как и то, для которого производится расчет, и, исходя из полученных данных, найти вероятность того, что в какой-либо из дней данному больному потребуется отдельная палата. На этом основании для отделения определенного размера можно будет найти среднее число больных, для которых потребуется изолятор. Однако вследствие значительных статистических колебаний, наблюдающихся в небольших отделениях, это среднее число не обязательно будет оптимальным. Для проведения более точного расчета введем следующие два понятия: 1) *коэффициент обеспеченности*, т. е. та доля фактической потребности в однокоечных палатах, которая может быть удовлетворена, и 2) *коэффициент использования*, т. е. доля времени, в течение которого эти палаты используются по своему прямому назначению. Например, в одном из обследований вероятность того, что какого-либо больного в какой-то день потребуется поместить в изолятор, оказалась равной примерно 0,3. Следовательно, для отделения на 16 человек потребуется $0,3 \cdot 16 = 4,8$ койки, т. е. около 5 коек. Что же даст такое число коек на практике?

Довольно простая математическая модель состоит в следующем. Пусть вероятность того, что данному больному в один из дней потребуется отдельная палата, равна p . Тогда для отделения на n человек вероятность того, что r больным потребуется отдельная палата, будет иметь биномиальное распределение

$$p_r = \binom{n}{r} p^r (1-p)^{n-r}, \quad (12.10)$$

а среднее число таких палат составит np . Представим себе, что мы обеспечены произвольным числом a отдельных палат. Среднее число больных, находящихся в таких палатах и действительно нуждающихся в изоляции, равно

$$S = \sum_{r=0}^a r p_r + \sum_{r=a+1}^n a p_r. \quad (12.11)$$

Первый член выражения (12.11) соответствует числу больных в те дни, когда спрос может удовлетворяться полностью, а второй член относится к случаям, когда спрос превышает предложение. Если коэффициент обеспеченности равен E_p , а коэффициент использования равен E_u , то легко видеть, что

$$E_p = S/np, \quad E_u = S/a. \quad (12.12)$$

Простые вычисления, основанные на формулах (12.11) и (12.12), дают результаты, показанные в табл. 23.

Таблица 23

**Коэффициенты обеспеченности
и коэффициенты использования для
отделения на 16 человек ($p=0,3$)**

Число изоляторов	Коэффициент обеспеченности	Коэффициент использования
2	41	99
3	60	96
4	76	91
5	87	83
6	94	75

Если в отделении имеется пять однокоечных палат, то спрос будет удовлетворен на 87%, а койки будут использоваться по своему прямому назначению в 83% случаев. Это, по существу, оптимальное равновесие, хотя коэффициент обеспеченности можно, например, увеличить до 0,94 за счет уменьшения коэффициента использования до 0,75. Если изолированных палат будет меньше,

то равновесие сдвинется в другую сторону. В общем если число изоляторов примерно равно среднему спросу (4,8 койки в приведенном здесь примере), то в больших отделениях достигается довольно оптимальное соотношение между этими противоречивыми требованиями. Однако в небольших отделениях трудно поддерживать высокие значения обоих коэффициентов одновременно. В этих случаях целесообразнее воспользоваться вычисленными данными, аналогичными приведенным в табл. 23. На основании результатов, представленных в таком виде, математик, правда, не сможет сообщить медицинскому персоналу необходимые для них данные или посоветовать архитектору, какой план следует выбрать, но все же каждый из них получит более точное представление о том, что скрывается за различными вариантами проекта, и это позволит проводить совместное обсуждение проекта на более рациональной основе.

Разумеется, существует много других аспектов организации отделения больницы, которые можно подвергнуть количественному анализу с целью более эффективной разработки новых проектов. То же относится и к большинству других помещений больницы: операционных, рентгеновских кабинетов, амбулаторного отделения и т. д. Помимо вопросов, связанных с улучшением функциональной эффективности различных отделений больницы, существует также вопрос их взаимодействия. Рассмотрим, например, потоки людей и различных материальных объектов, циркулирующие внутри больницы. Вообще говоря, участки с большой интенсивностью взаимных потоков должны располагаться поблизости, а участки, меньше связанные друг с другом, находиться на большем расстоянии один от другого. Отступления от этого принципа могут вызываться необходимостью оказания неотложной помощи или удобством административного управления. Различают два основных понятия: *общую интенсивность потока* между любыми двумя участками, которую желательно свести к минимуму, чтобы сократить расстояния перехода для больных и обслуживающего персонала, а также уменьшить количество шума и грязи, и *плотность потока*, изменяющуюся с течением времени и в зависимости от места. Ее также по возможности необходимо уменьшать хотя бы в такой степени, чтобы исключить образование местных перегрузок и узких мест.

Основные проблемы обычно связаны с приемом амбулаторных больных, приемом посетителей и работой пассажирских лифтов. Часто бывает достаточно рассмотреть предложенные проекты больницы с учетом этих основных источников потока. Конкретные численные оценки интенсивности и плотности потоков можно определить с помощью сложных и дорогостоящих обследований существующих лечебных учреждений, однако полученные данные

могут не подойти к новому проекту. По-видимому, более целесообразно построить *априорные* предсказания. Предполагаемое общее число амбулаторных больных, которых придется обслуживать по той или иной специальности, можно получить с помощью исследования, аналогичного описанному в разд. 12.2 для стационарных больных. Конкретная схема потока амбулаторных больных зависит частично от проекта лечебного учреждения, а частично от принятого порядка обслуживания, поэтому необходимо учитывать предполагаемую *организацию* обслуживания амбулаторных больных.

Поток посетителей часто определяется соотношением «два человека на одного больного». Влияние этого потока на общую интенсивность потока зависит от расписания приема посетителей и расположения лестниц и лифтов. Правильная организация приема посетителей имеет важнейшее значение для преодоления трудностей, связанных с движением потоков.

Перевозка людей в лифтах представляет собой более сложную проблему, и на практике часто можно наблюдать, что подъемные устройства обладают недостаточной пропускной способностью. Расчет потребности в лифтах часто строится на коммерческих эмпирических методах, не имеющих рациональной основы, и маловероятно, чтобы эти методы были пригодны для больниц. Работа лифтов связана с некоторыми довольно сложными процессами массового обслуживания. Вопрос о целесообразности детального математического исследования работы лифтов еще требует дополнительного обсуждения. Возможно, было бы полезно заблаговременно рассчитать, к каким последствиям приведет применение подъемников различного типа.

12.4. СИСТЕМЫ ПРИЕМА АМБУЛАТОРНЫХ БОЛЬНЫХ

Хотя проблема обеспечения эффективной деятельности больниц частично связана с общим проектом здания, схемой расположения его частей и числом коек, исключительно важную роль играет правильное использование имеющихся средств. Как было указано в конце предыдущего раздела, решение сложных проблем, связанных с движением потоков внутри помещений, в такой же степени зависит от планирования времени обследования больных и приема посетителей, как и от расположения палат и амбулаторного отделения. Общую последовательность операций, происходящих на некотором ограниченном участке, часто можно изучать относительно изолированно. Хорошим примером служит работа амбулаторного отделения, требующая решения многих задач, связанных с административным управлением, не говоря уже о специализированном обслуживании больных, которое осуществляют врачи и медицинские сестры.

В течение ряда лет очень многие амбулаторные больные постоянно жаловались на то, что им приходится очень долго ожидать приема. Обследование, выполненное в 1952 г. [48] и охватившее 34 различные клиники, показало, что среднее время ожидания от момента регистрации до момента вызова на прием составляет 56 мин. В одной из лондонских учебных хирургических клиник больные ожидали в среднем 1 час 40 мин, 17% из них ожидали более 2,5 час, а около 2% — даже 3,5 час. Более позднее обследование, выполненное в 1964 г. [49], показало, что общее среднее время ожидания составляет 25 мин; это свидетельствует о некотором улучшении обслуживания, хотя, строго говоря, эти две выборки данных несравнимы друг с другом. При этом втором обследовании оказалось, что 34% больных вынуждены ожидать более получаса, а 11% — более часа. Между обследованными больницами наблюдалось значительное различие; в некоторых из них положение было лучше, а в других — значительно хуже.

Как выяснилось, некоторые амбулаторные отделения имеют сравнительно гибкие системы приема больных с довольно коротким временем ожидания, другие же работают малоэффективно; так, всем больным назначается одно и то же время приема, и поэтому длительность ожидания велика. По этому вопросу еще ведутся большие споры. Одни специалисты полагают, что можно значительно сократить время ожидания, другие утверждают, что длительное ожидание неизбежно. Возможно, что та или иная клиника достигла удовлетворительной схемы обслуживания амбулаторных больных почти случайно, однако, как только такая схема устанавливается, она сохраняется и в дальнейшем. Трудности возникают там, где условия неудовлетворительны. Часто считают, что ситуация в целом настолько сложна и внутренние факторы настолько изменчивы, что заметных улучшений достигнуть не удастся, хотя общепризнано, что длительное ожидание по ряду причин нежелательно.

По мнению автора, создание надлежащим образом разработанной системы приема амбулаторных больных совершенно необходимо, и эта точка зрения подтверждается здравым смыслом, наблюдениями и теорией. Здравый смысл подсказывает, что время ожидания можно значительно уменьшить, назначая больных на прием через достаточно большие промежутки времени. Это может, конечно, привести к тому, что врач будет ожидать больного. В этом случае возникает проблема сбалансирования противоречивых требований, как в описанной выше задаче определения числа однокоечных палат. Согласно наблюдениям, некоторые клиники действительно добились успехов в этом отношении, и многие торговые предприятия также избавили своих клиентов от длительного ожидания. Согласно теории, можно проанализиро-

вать эту ситуацию методами теории массового обслуживания, используя модель, обладающую основными чертами всей системы; эта модель показывает, что удовлетворительные оптимальные схемы приема амбулаторных больных вполне достижимы.

Рассмотрим роль таких теоретических исследований. Чтобы иметь надлежащую рациональную основу для обсуждения, нам необходима математическая модель, описывающая прохождение больных через амбулаторное отделение. На первый случай эта модель может быть довольно простой, но она должна обладать большинством важнейших свойств реальной ситуации (если не всеми). Эта модель должна обязательно учитывать сильно варьирующее время обследования отдельных больных и дать возможность сравнивать различные системы приема амбулаторных больных, особенно в отношении времени ожидания для больного и, возможно, времени простоя врача.

Если рассматривается простейшая схема, согласно которой врач обслуживает непрерывный поток посетителей, то ясно, что мы имеем определенную форму процесса массового обслуживания (см. разд. 4.3), который поддается анализу стандартными математическими методами. Однако здесь есть одно важное отличие от других процессов такого рода. В разд. 12.2 мы использовали определенную модель процесса массового обслуживания, чтобы показать, каким образом больница может удовлетворить потребности в госпитализации. Для практических целей нам требовалось лишь вычислить такие величины, как среднее время ожидания госпитализации, при допущении, что установилось некоторое равновесное состояние процесса. Такие средние значения, взятые по бесконечным интервалам времени, во многих случаях вычислить довольно легко. Однако время, в течение которого производится прием больных, обычно имеет ограниченную длительность, а для конечных интервалов времени применение математической теории массового обслуживания крайне усложняется даже в случае простейших моделей. Так как в общем случае процесс приема амбулаторных больных даже не является марковским, то его легче исследовать путем моделирования или с помощью методов Монте-Карло (см., например, разд. 2.5 и 4.3).

Рассмотрим прием, на который записано 25 больных, и допустим, что продолжительность обследования отдельных больных крайне изменчива и имеет по существу распределение χ^2 с пятью степенями свободы (если не учитывать изменения масштаба), которое очень похоже на многие наблюдаемые распределения. Примем *среднее* время обследования равным 5 мин (считая временем обследования весь период, в течение которого больной занимает время врача или пользуется его вниманием). Сюда входит сбор анамнеза, обследование, дача рекомендаций, раздевание

и одевание больного, а в учебных клиниках также время, затрачиваемое на объяснение случая студентам. Таким образом, общая длительность приема составляет около 2 час 5 мин, если врачу самому не приходится какое-то время ожидать больных. Допустим, что больные прибывают на прием один за другим в соответствии с некоторым заранее установленным правилом, или системой назначения. Можно изучить различные системы назначения и сравнить их достоинства.

С помощью соответствующего статистического метода выбора случайных величин, имеющих распределение χ^2 , можно составить искусственную последовательность, состоящую из 25 типичных отрезков времени, необходимых для обследования, которые можно рассматривать как моделированную группу больных, пришедших на амбулаторный прием в определенную клинику. Для того чтобы проверить, к чему приводят различные допущения о системе назначения больных, автор [3] рассмотрел 50 таких воображаемых клиник. Допустим, например, что врач начинает работу, когда уже ожидает приема первоначальная группа, состоящая из пяти больных, а остальным 20 больным время приема назначено с интервалом 5 мин. С помощью простых, хотя и довольно утомительных вычислений можно показать, сколько времени должен в этом случае ожидать вызова каждый больной и, наоборот, сколько времени врач должен ожидать следующего больного после окончания обследования предыдущего. В этом первоначальном исследовании использовались обычные методы вычислений вручную с помощью карандаша и бумаги, однако при современных более обширных исследованиях должна применяться вычислительная техника.

Было изучено влияние самых различных факторов, в частности размера первоначальной группы больных и длительности интервала между прибытием больных на прием. По-видимому, интервал между прибытием больных лучше всего принять равным средней длительности обследования, но оказалось, что решающее значение имеет выбор размера первоначальной группы (табл. 24).

Из табл. 24 хорошо видно столкновение противоречивых требований. Среднее время ожидания для больного можно уменьшить только за счет общего времени простоя врача, и наоборот. Разумеется, на практике врач часто бывает надежно застрахован от возможности ожидания за счет того, что он начинает прием при наличии очень большой группы больных. Действительно, как показывает табл. 24, при первоначальной группе, насчитывающей 6 больных, общее время простоя врача в среднем составляет всего 0,5 мин, а среднее время ожидания для больного — 24 мин, так что общее время ожидания для всех больных равно $24 \cdot 25 = 600$ мин. Значительно более приемлемое равновесие достигает-

Таблица 24

Среднее время ожидания для больного и среднее время простоя врача при разном размере начальной группы больных (среднее время обследования и интервал между прибытием больных 5 мин)

Число больных в первоначальной группе	Среднее время ожидания для больного, мин	Среднее время простоя врача, мин
1	7 (18) *	9
2	9 (20)	6
3	12 (23)	3
4	16 (27)	2
5	20 (31)	1
6	24 (35)	0,5

* Один больной из восьми ожидает больше времени, указанного в скобках.

ся, по-видимому, при наличии в первоначальной группе только двух больных. В этом случае время ожидания для больного составляет 9 мин, а время простоя врача — 6 мин. Но, по-видимому, в большинстве случаев время простоя будет на самом деле даже меньше, так как некоторые больные приходят раньше назначенного времени; кроме того, врачу приходится заниматься и неизбежной канцелярской работой — читать бумаги, производить записи или говорить по телефону.

Изучался также вопрос о том, как влияет на время ожидания форма кривой распределения времени обследования, однако общие результаты оказались мало чувствительными к изменениям числа степеней свободы в пределах от 4 до 7.

Следует заметить, что при предложенной здесь схеме обслуживания загрузка системы фактически равна единице. Если бы рассматривался бесконечный процесс массового обслуживания, то это вызвало бы определенные затруднения, однако при относительно коротких процессах, имеющих конечную длительность, критические пороговые явления не наблюдаются.

Описанные выше моделированные клиники можно также проанализировать другими способами, чтобы получить, например, некоторое представление о распределениях различных случайных величин. Так, в табл. 24 приведены данные, относящиеся к области больших отклонений распределения времени ожидания для больных. Кроме того, может оказаться целесообразным рассмотреть распределение общего числа больных, ожидающих в какой-

либо данный момент времени. Можно, например, показать, что для рекомендованной системы с первоначальной группой из двух больных 22% времени вообще не будет ожидающих больных, а 86% времени будут ожидать приема не более трех больных. Очевидно, что эти данные должны были бы приниматься во внимание архитектором при проектировании залов ожидания. Необходимо также учитывать возможность прибытия больных позже или раньше назначенного срока, а также наличие сопровождающих лиц.

Хотя мы рассмотрели свойства довольно простой модели, общая картина уже ясна. Время приема больных должно назначаться таким образом, чтобы установить приемлемое равновесие с временем простоя врача. Сделать это нетрудно: определив на основе достаточно большой выборки больных среднее время обследования, нужно приравнять к нему интервал между назначенным временем приема следующих друг за другом больных (возможно, округляя вычисленные интервалы до 3 мин).

Конечно, работа многих амбулаторных отделений клиник (если не большинства) протекает гораздо сложнее, чем в описываемой модели. Так, прием могут производить несколько человек, например один врач-консультант и один или несколько регистраторов. Если провести различие между консультациями, проводимыми отдельными врачами, невозможно, то нужно рассматривать всю группу врачей как единое целое и сделать так, чтобы общее время обследования больного примерно равнялось интервалу между прибытием больных. Кроме того, следует учесть, что некоторых больных приходится направлять в рентгеновский кабинет, после чего они вновь должны прийти к врачу. При вычислении среднего времени обследования одного больного необходимо принимать во внимание время, которое он проводит во всех кабинетах, но учитывать каждого больного один раз. Другие трудности связаны с тем, что некоторые больные прибывают раньше или позже назначенного срока или вообще не являются на прием. Как правило, практические последствия, к которым приводит такая неаккуратность, вычислить довольно легко (во всяком случае, с определенной степенью приближения), и нужно должным образом учесть их. Если возникают серьезные сомнения в возможных результатах влияния тех или иных факторов, то не представляет труда провести соответствующим образом спланированное исследование методом моделирования. Однако самые главные трудности носят не теоретический, а практический характер: обычно бывает очень трудно убедить административных деятелей (как врачей, так и неспециалистов), что достаточно эффективные системы приема амбулаторных больных в действительности совсем нетрудно разработать и провести в жизнь.

12.5. ИЗУЧЕНИЕ СЛОЖНЫХ СИСТЕМ И УПРАВЛЕНИЕ ИМИ

Врачи, медицинские сестры, руководители лечебных учреждений и научные работники повсеместно и неустанно собирают медицинскую документацию в надежде, что когда-нибудь эти данные можно будет использовать для научных целей. Чаще всего это преимущественно клинические данные, связанные с анамнезом, постановкой диагноза, лечением и прогнозом, касающиеся отдельных больных. Такие сводные материалы, позволяющие, например, определить среднюю частоту определенного заболевания и частоту появления различных симптомов или количественно оценить результаты различных методов лечения, представляют ценный вклад в общий фонд медицинских знаний. Они помогают врачу в выборе соответствующих методов лечения в каждом конкретном случае, а также могут служить основой для дальнейших научных исследований.

О некоторых возможностях современных методов обработки такого рода данных уже упоминалось в разд. 5.3. Эти данные часто содержат большое количество информации, которая была бы исключительно полезной, если бы ее удалось извлечь и проанализировать. К сожалению, они нередко бывают непредставительными или неполными. В этих случаях статистические трудности, связанные с введением поправок на систематические ошибки и выявлением роли различных взаимодействующих факторов, могут быть столь велики, что справиться с ними не удастся (см., например, разд. 2.3). Если медицинская документация специально не приспособлена к требованиям конкретно сформулированных научно-исследовательских задач, то она может оказаться почти бесполезной. Идея накопления обширных стандартных записей в надежде, что в конечном счете они помогут ретроспективным исследованиям, на практике почти всегда разочаровывает, хотя на их основе и можно получить полезные рекомендации и указания. За немногими исключениями, лишь такой принцип ведения документации, который специально нацелен на будущие исследования, может быть по-настоящему удовлетворительным, поскольку только в этом случае можно провести правильно спланированные эксперименты и обследования.

Тем не менее повседневная медицинская документация имеет практическое значение в повседневной работе медицинского учреждения. На индивидуальном уровне — в каждом отдельном случае — она необходима для регистрации симптомов, проверки выполнения назначений и наблюдения за течением болезни; на более общем уровне она используется для планирования всей работы по обслуживанию целых групп больных. Применение надлежа-

щей аппаратуры позволяет значительно ускорить обработку данных на обоих этих уровнях.

Рассмотрим, например, обычную современную больницу. В таком учреждении административные проблемы, связанные с выполнением всех операций по обслуживанию амбулаторных и стационарных больных врачами, сестрами, техническим персоналом и т. д., крайне сложны. Это приводит к тому, что административными и канцелярскими обязанностями все больше приходится заниматься и тем работникам, чьи знания и способности должны быть полностью направлены на выполнение совсем других дел. Количество данных, касающихся больного, крайне велико. Одни поступают из приемного отделения, другие накапливаются у лечащего врача, некоторые могут уже ранее храниться в центральной картотеке, а некоторые поступают из таких специальных отделений, как рентгеновский кабинет, прозектура и т. д. Есть угроза, что эта обширная разветвленная система потоков данных будет становиться все более сложной. Что же можно сделать, во-первых, для повышения эффективности управления такой системой в повседневной практике, а во-вторых, для изучения поведения таких систем с целью рекомендовать те или иные улучшения и найти новые схемы управления этой деятельностью?

В отношении повседневных методов управления стоит изучить опыт применения вычислительной машины для обработки медицинской документации в детской больнице на 260 мест в г. Акрон. Одна из основных причин, побудивших ускорить обработку и передачу документации, состояла в том, что стало ясно, какую значительную часть своего времени медицинский персонал тратит на оформление различных бумаг (графиков, диаграмм, записей результатов клинического обследования, лабораторных анализов, врачебных назначений, касающихся медикаментов и диеты), на составление распорядка работы клиники, расписаний работы операционных и т. д. В некоторых случаях высококвалифицированные медицинские сестры тратили до 40% своего времени на выполнение по существу канцелярской работы. Существовавшая система не позволяла им уделять достаточное внимание больным и находиться непосредственно у их постели столько времени, сколько требуется.

К счастью, выписка счетов и вся бухгалтерская работа в этой больнице уже была ранее автоматизирована и производилась с помощью вычислительной машины. Эту машину связали с 14 периферийными устройствами ввода и вывода данных, расположенными в различных сестринских постах и специальных отделениях больницы. В результате всю информацию, обычно накапливаемую в разных местах в виде записей, отыскание и обработка которых поглощала массу времени, можно было хранить в цент-

ральной вычислительной машине и в случае надобности получать практически мгновенно. Это не только резко сократило время, необходимое для записи и поиска информации, но и позволило установить более эффективный контроль (см. гл. 5), что крайне важно при обращении с лекарствами и при обработке рецептов. Кроме того, с помощью специально разработанной программы на центральной вычислительной машине могут решаться такие сложные и трудоемкие задачи, как распределение ежедневных обязанностей персонала или составление графика пользования операционными. Экономия только на одном распределении обязанностей персонала оказалась весьма значительной: ранее этой работой занимались две старшие медицинские сестры по 8 час в сутки каждая, а вычислительная машина выполнила ее за 7 мин.

Нет необходимости перечислять все те разнообразные задачи, которые решает вычислительная машина в этой больнице. Главное состоит в том, что автоматизация такого рода работы вполне возможна, и вместо удаления человека из сферы медицинского обслуживания, о чем твердят некоторые скептики, достигается как раз обратный результат, поскольку врачи и медицинские сестры освобождаются от чисто канцелярской и технической работы.

Приведенный пример показывает, какой выигрыш может получить такая сложная система, как современная больница, при введении автоматических методов обработки данных. Однако одна эта мера недостаточна для улучшения деятельности существующих систем. Необходимо исследовать свойства и возможности новых форм организации. По ряду причин во многих случаях практически невозможно экспериментально проверить большое число вариантов (более подробно этот вопрос разбирается в разд. 4.3). Поэтому необходим метод, который позволял бы предварительно осуществлять теоретическую проверку новых идей, чтобы лишь наиболее перспективные из них подвергались практическим испытаниям. Методы моделирования, рассмотренные в разд. 2.5 и 4.3, наилучшим образом подходят для решения задач такого рода.

Исследовательская группа фирмы «Системс девелопмент» и Управления по делам ветеранов в Лос-Анджелесе разработала модель больницы, приняв за основу больницу в г. Уодсуэрт, в состав которой входят терапевтические и хирургические отделения (см. журнал SDC Magazine, январь 1964 г.). Было решено, что в основу модели будет положено описание работы одного отделения, и для исследования были выбраны шесть основных аспектов этой работы: прием и размещение больных; перемещение больных внутри отделения и их выписка; проведение лабораторных анализов (например, биохимических и гематологических); обеспечение быта больных; лечебная деятельность и, наконец, медицин-

ская статистика. На основании данных, извлеченных из соответствующих обследований фактического материала, в модель были включены многие характеристики реальной больницы. Методом моделирования исследовались последствия применения полностью интегрированной системы обработки информации того же общего типа, как и использовавшаяся в детской больнице г. Акрон, хотя, возможно, и более обширной. Это позволило весьма детально исследовать все вопросы, связанные с обработкой информации, и с помощью квазиэкспериментов определить возможные последствия реализации предлагаемых новых методов, не осуществляя их на практике. Такой метод предоставляет почти неограниченные возможности для проведения научных исследований планирования в области организации больничных служб.

Другой аспект моделирования, уже рассмотренный в разд. 4.3, — это применение операционных игр. Фирма «Системс девелопмент» и Управление общественного здравоохранения штата Калифорния использовали этот метод для решения учебных и научно-исследовательских задач в области здравоохранения (см. журнал SDC Magazine, январь 1964 г.), разработав модель гипотетического города под названием Эпивилль; в эту сложную модель были включены в явном виде все обычные действия и функции персонала, работающего в области здравоохранения. В исследовании участвовали реальные люди, которые должны были принимать определенные решения, сталкиваясь с рядом проблем, возникающих при моделированных вспышках эпидемии. Качество решений, принимавшихся каждым конкретным лицом, оценивалось впоследствии при совместных обсуждениях полученных результатов. При этом особое внимание уделялось следующим вопросам: была ли получена достаточная информация об эпидемических вспышках, были ли приняты правильные решения, удалось ли достичь успеха в борьбе с эпидемией и т. д.

Ценность таких операционных игр с точки зрения обучения несомненна. Большинство участников сообщали, что они испытывали такое же беспокойство и состояние напряженности и такое же чувство ответственности, как и в реальной обстановке. Следовательно, персонал, занятый в сфере общественного здравоохранения, может приобрести в процессе таких игр большой и ценный опыт работы в сложной обстановке, возникающей на практике изредка.

Помимо целей обучения такие модели позволяют сравнивать относительные достоинства различных типов организации, что применимо ко всей области медицинского обслуживания и здравоохранения. При изучении общей системы обеспечения медицинской помощью методом исследования операций моделирование

будет, вероятно, играть особенно важную роль, так как чисто математический анализ окажется здесь скорее всего невозможным.

Широкое взаимодействие человека, вычислительной техники и математики во всей области медицины по существу еще только начинается, однако оно может вызвать настоящую революцию в медицинской науке. Такое взаимодействие нужно понимать в том смысле, что все чисто канцелярские, организационные и технические функции должны быть переданы вычислительным машинам, которые будут осуществлять их значительно более эффективно, чем человек; в результате у врачей и медицинских сестер освободится время для выполнения их прямых обязанностей, требующих интуиции, воображения, здравого смысла и чувства сострадания, которых ни математика, ни вычислительная техника заметить не могут.

ЛИТЕРАТУРА

При составлении библиографии автор не стремился дать подробный список работ по всем вопросам, рассмотренным в данной книге. Здесь приведен лишь перечень литературы для дополнительного чтения. В тексте читатель может найти ссылки на существующие общие руководства, где соответствующие вопросы излагаются более полно и приводится основной библиографический материал. Если такие руководства отсутствуют, то даются ссылки на некоторые важнейшие научные статьи, с которыми читателю полезно познакомиться.

1. Armitage P., Sequential Medical Trials, Blackwell, Oxford, 1960.
2. Ashby R. W., Design for a Brain, Wiley, New York, 1962 (У. Р. Эшби, Конструкция мозга, ИЛ, М., 1962).
3. Bailey N. T. J., A study of queues and appointment systems in hospital outpatient departments, with special reference to waiting-times, J. Roy. Statist. Soc., B, 14, 185 (1952).
4. Bailey N. T. J., Queueing for medical care, Applied Statistics, 3, 137 (1954).
5. Bailey N. T. J., Statistics in hospital planning and design, Applied Statistics, 5, 146 (1956).
6. Bailey N. T. J., The Mathematical Theory of Epidemics, Griffin, London, 1957.
7. Bailey N. T. J., The Mathematical Theory of Genetic Linkage, Clarendon, Oxford, 1961.
8. Bailey N. T. J., Calculating the scale of inpatient accommodation. In Towards a Measure of Medical Care, Oxford University Press, London, 1962.
9. Bailey N. T. J., The Elements of Stochastic Processes, Wiley, London, 1964.
10. Bailey N. T. J., Probability methods of diagnosis based on small samples. В книге: Mathematics and Computer Science in Biology and Medicine, Proceedings of an M.R.C. Conference at Oxford, H.M.S.O., London, 1965.
11. Bartlett M. S., An Introduction to Stochastic Processes, Cambridge University Press, Cambridge, 1955 (Барлетт М. С., Введение в теорию случайных процессов, ИЛ, М., 1958).
12. Bartlett M. S., Stochastic Population Models in Ecology and Epidemiology, Methuen, London, 1960.
13. Bhattacharjee A. T., Elements of the Theory of Markov Processes and their Applications, McGraw-Hill, London, 1960. (Баруча-Рид А. Т.,

элементы теории марковских процессов и их приложения, М., «Наука», 1969).

14. Cochran W. G., Cox G. M., *Experimental Designs* (2nd ed.), Wiley, London, 1957.
15. Cox D. R., Smith W. L., *Queues*, Methuen, London, 1961 (Рокс Д. Р., Смит В. Л., Теория очередей, изд-во «Мир», М., 1966).
16. Crooks J., Murray I. D. C., Wayne E. J., *Statistical methods applied to the clinical diagnosis of thyrotoxicosis*, *Quart. J. Med.*, 28, 211 (1959).
17. Darroch J. N., *Interactions in multifactor contingency tables*, *J. Roy. Statist. Soc.*, B, 24, 251 (1962).
18. Dijkstra E. W., *A Primer of ALGOL 60 Programming*, Academic Press, London, 1962.
19. Enslein K. (ed.), *Data Acquisition and Processing in Biology and Medicine*, Pergamon, Oxford, 1962.
20. Enslein K. (ed.), *Data Acquisition and Processing in Biology and Medicine*, vol. 2, Pergamon, Oxford, 1964.
21. Enslein K. (ed.), *Data Acquisition and Processing in Biology and Medicine*, vol. 3, Pergamon, Oxford, 1964.
22. Feller W., *An Introduction to Probability Theory and its Applications*, vol. 1, Wiley, New York, 1957 (Феллер В., Введение в теорию вероятностей и ее приложения, 2-е изд., изд-во «Мир», М., т. 1, 1964; т. 2, 1967).
23. Fisher R. A., *Statistical Methods for Research Workers*, Oliver and Boyd, London, 1925 (Фишер Р. А., Статистический метод исследований, 1936, Стеклографическое издание).
24. Fisher R. A., *The Design of Experiments*, Oliver and Boyd, London, 1935.
25. Flagle C. D., Huggins W. H., Roy R. H., *Operations Research and Systems Engineering*, Johns Hopkins, Baltimore, 1960.
26. Forsyth G., Logan R. F. L., *The Demand for Medical Care*, Oxford University Press, London, 1960.
27. Galton F., *Natural Inheritance*, Macmillan, London, 1889.
28. Garvin W. W., *Introduction to Linear Programming*, McGraw-Hill, London, 1960.
29. Geisser S., *Posterior odds for multivariate normal classifications*, *J. Roy. Statist. Soc.*, B, 26, 69 (1964).
30. Goodman L. A., *On methods for comparing contingency tables*, *J. Roy. Statist. Soc.*, A, 126, 94 (1963).
31. Gurland J. (ed.), *Stochastic Models in Medicine and Biology*, University of Wisconsin, Madison, 1964.
32. Hammersley J. M., Handcomb D. C., *Monte Carlo Methods*, Methuen, London, 1964.
33. Hamming R. W., *Numerical Methods for Scientists and Engineers*, McGraw-Hill, London, 1962.
34. Hayhoe F. G. J., Quagliano D., Doll R., *The Cytology and Cytochemistry of Acute Leukaemias: A Study of 140 cases* (M.R.C. Special Report No. 304), H.M.S.O., London, 1964.
35. Hollingdale S. H., Tootill G. C., *Electronic Computers*, Penguin Books, Harmondsworth, London, 1965.

36. J a s q u e z J. A. (ed.), The Diagnostic Process, University of Michigan, 1964.
37. J e f f r e y s H., Theory of Probability, 3rd ed., Clarendon, Oxford, 1961.
38. K e n d a l l M. G., S t u a r t A., The Advanced Theory of Statistics, Griffin, London, 1958 (К е н д а л л М., С т ю а р т А., Теория распределений, «Наука», М., 1966).
39. L e d l e y R. S., L u s t e d L. B., Mathematical diagnosis and modern decision making, In Mathematical Problems in the Biological Sciences, Proceedings of a symposium in applied mathematics, Vol. 14, p. 117, American Math. Soc., 1961.
40. L o t k a A. J., Elements of Physical Biology, Baltimore, Williams and Wilkins, 1925; впоследствии эта книга была переиздана под названием Elements of Mathematical Biology, Dover, 1956.
41. L u c a s H. L. (ed.), The Cullowhee Conference on Training in Biomathematics, Raleigh, North Carolina, Typing Service, 1962.
42. M e s s r a c k e n D. D., A Guide to ALGOL Programming, London, Wiley, 1962 (М а к - К р а к е н Дж., Программирование на АЛГОЛе, изд-во «Мир», 1964).
43. M a l t h u s T., An Essay on the Principle of Population, London, 1798 (М а л ь т у с Т. П., Опыт о законе народонаселения, т. 1—2, Спб., 1868).
44. M a s u y a m a M., Various approaches to automatic diagnosis by discrete variables, Rep. Statist. Appl. Res. Un. Jap. Sci. Engors, 10, 151 (1963).
45. M o l i n a E. C., Poisson's Exponential Binomial Limit. New. York, Van Nostrand, 1942.
46. N a u r P. (ed.), Revised report on the algorithmic language ALGOL 60, The Computer Journal, 5, 349 (1963).
47. V o n N e u m a n n J., M o r g e n s t e r n O., Theory of Games and Economic Behaviour, Princeton, 1947.
48. Nuffield Provincial Hospitals Trust, Studies in the Function and Design of Hospitals, London, Oxford University Press, 1955.
49. Nuffield Provincial Hospitals Trust, Waiting in Outpatient Departments, Oxford, Oxford University Press, 1965.
50. P a r z e n E., Stochastic Processes, San Fransisco, Holden-Day, 1960.
51. P a y n e L. C., Toward medical automation, Wld. Med. Electronics, 2, 7 (1963).
52. P e a r s o n K., The Grammar of Science, Everyman edition, 1938 (П и р с о н К., Грамматика науки, Спб., 1914).
53. Q u e t e l e t L. A. J., Sur l'homme et le développement de ses facultés, un essai de physique sociale, Paris, Bachelier, 1835.
54. R a s h e v s k y N., Mathematical Biophysics, 3rd ed., 2 vols., New York, Dover, 1960.
55. R a s h e v s k y N., Mathematical Principles in Biology, Springfield, Illinois, Thomas, 1961.
56. S a a t y T. L., Elements of Queuing Theory, London, McGraw-Hill, 1961 (С а а т и Т. Л., Элементы теории массового обслуживания и ее приложения, изд-во «Советское радио», М., 1965).
57. S a s i e n i M., Y a s p a n A., F r i e d m a n L., Operations Research, London, Wiley, 1959.
58. S i m p s o n G. G., Principles of Animal Taxonomy, New York, Columbia University Press, 1961.

59. S o k a l R. R., S n e a t h P. H. A., Principles of Numerical Taxonomy, London, Freeman, 1963.
60. T a k a h a s h i K., D o h i Y., S a t o T., H a s e g a w a T., O u c h i A., T a k a h a s h i T., F u k a m a c h i M., Logic of medical diagnosis and a matrix-type electric diagnostic machine, 5th International Conference on Medical Electronics at Liège, 1963.
61. T h o m p s o n D'A. W., On Growth and Form, 2 vols., 2nd ed., Cambridge University Press, 1942.
62. T o c h e r K. D., The Art of Simulation, London, English University Press, 1963.
63. W a r n e r H. R., T o r o n t o A. F., V e a s e y L. G., S t e p h e n s o n R., A mathematical approach to medical diagnosis, J. Amer. Med. Ass., 177 (1961).
64. W i e n e r N., Cybernetics, London, Wiley, 1948 (В и н е р Н., Кибернетика, изд-во «Советское радио», М., 1968).
65. V a n W o e r k o w A. J., B r o d m a n K., Statistics for a diagnostic model, Biometrics, 17, 299 (1961).
66. Y o s k e y H. P. (ed.), Symposium on Information Theory in Biology, London, Pergamon, 1958 («Теория информации в биологии», ИЛ, М., 1960).
67. Z e e m a n E. C., Topology of the brain. Mathematics and Computer Science in Biology and Medicine, Proceedings of an M.R.C. Conference at Oxford, p. 277, London, H.M.S.O., 1965.
68. В и ш н е в с к и й А. А., А р т о б о л е в с к и й И. И., Б ы к о в с к и й М. Л., Кибернетические методы в медицине, Scientific World, 8, 13 (1964).

О Г Л А В Л Е Н И Е

Предисловие к русскому изданию	5
Из предисловия автора	7

ЧАСТЬ I. ОБЩИЕ ПРИНЦИПЫ И МЕТОДЫ

Глава 1. ПОТРЕБНОСТЬ В МАТЕМАТИКЕ	9
1.1. Счет и измерение	9
1.2. Математические модели	13
1.3. Динамические модели	21
1.4. Области применения математических методов	27
Глава 2. РОЛЬ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ	32
2.1. Биологическая изменчивость и вероятность	32
2.2. Применение математической статистики	36
2.3. Планирование экспериментов	41
2.4. Случайные процессы	47
2.5. Физическое моделирование и метод Монте-Карло	55
Глава 3. ПРОЦЕСС НАУЧНОГО ИССЛЕДОВАНИЯ	58
3.1. Научный метод	58
3.2. Значение математических моделей	63
3.3. Модели статистического вывода	67
3.4. Многообразие математических методов	75
Глава 4. МЕТОДЫ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ	83
4.1. Содержание предмета	83
4.2. Возникновение и развитие методов исследования операций	86
4.3. Некоторые методы исследования операций	88

Теория массового обслуживания	89
Математическое программирование	92
Моделирование и операционные игры	96
Теория игр	100
4.4. Комплексные исследования на стыке различных наук	103
Глава 5. роль вычислительной техники	106
5.1. Общие принципы	106
5.2. Научные вычисления и численный анализ	112
5.3. Обработка данных	116
5.4. Программирование	120
5.5. Вычислительные машины и научные исследования	126
Глава 6. организация научных исследований	132
6.1. Введение	132
6.2. Преподавание, проведение консультаций и научные исследования	135
Преподавание	135
Проведение консультаций	139
Научно-исследовательская работа	141
6.3. Совместные научные исследования	143
 ЧАСТЬ II. ЧАСТНЫЕ ПРИЛОЖЕНИЯ	
Глава 7. числовая таксономия	148
7.1. Введение	148
7.2. Основные принципы классификации	152
7.3. Принципы числовой таксономии	154
7.4. Оценка сходства	157
Коэффициенты ассоциации	158
Коэффициенты корреляции	161
Показатели расстояния	162
7.5. Построение фенонов	164
7.6. Приложения и обсуждение результатов	170
Глава 8. экология и рост популяций	173
8.1. Введение	173
8.2. Процессы размножения и гибели	174
Простой процесс размножения	176

Простой процесс размножения и гибели	178
Вероятность вымирания популяции	180
Влияние миграции	182
8.3. Многомерные процессы	185
Популяции, состоящие из особей двух полов	186
Суммарная численность популяции	187
Мутации у бактерий	190
Многофазовый процесс размножения	192
8.4. Взаимодействие между видами	195
Конкуренция между двумя видами	195
Модель системы хищник — жертва (или паразит — хозяин)	197
8.5. Обсуждение результатов	200
Глава 9. ТЕОРИЯ ЭПИДЕМИЙ	202
9.1. Введение	202
9.2. Простые эпидемии	205
Детерминистская модель	205
Стохастическая модель	206
9.3. Общий случай эпидемии	210
Детерминистская модель	210
Стохастическая модель	212
9.4. Повторяющиеся эпидемии	214
Детерминистская модель	215
Стохастическая модель	216
9.5. Цепочно-биномиальные модели	218
9.6. Заразный и латентный периоды	223
9.7. Пространственные модели	228
Глава 10. ГЕНЕТИЧЕСКОЕ СЦЕПЛЕНИЕ И ХРОСОМНЫЕ КАРТЫ	234
10.1. Введение	234
10.2. Простые эксперименты со сцеплением	239
Дифференциальная жизнеспособность	243
10.3. Сбалансированное тригибридное возвратное скрещивание	247
10.4. Простые хромосомные карты	253
10.5. Метрики интерференции	257
Оценка параметров	261
Глава 11. МАТЕМАТИЧЕСКИЕ МЕТОДЫ МЕДИЦИНСКОЙ ДИАГНОСТИКИ	265
11.1. Введение	265
11.2. Дифференциальный диагноз в элементарном случае выбора между двумя заболеваниями	267

11.3. Дифференциальный диагноз в общем случае выбора между несколькими заболеваниями	272
11.4. Использование небольших выборок	279
11.5. Критерии значимости	285
11.6. Заключение	289
Глава 12. ИССЛЕДОВАНИЕ ОПЕРАЦИЙ В МЕДИЦИНЕ	292
12.1. Введение	292
12.2. Потребность в медицинском обслуживании	293
12.3. Проектирование больниц	301
12.4. Системы приема амбулаторных больных	308
12.5. Изучение сложных систем и управление ими	314
Л и т е р а т у р а	319

Н. Бейли

**МАТЕМАТИКА
В МЕДИЦИНЕ И БИОЛОГИИ**

Редактор *А. Левина*

Художник *В. Карпов*

Художественный редактор *Ю. Максимов*

Технический редактор *Л. Кондюкова*

Корректор *И. Максимова*

Сдано в производство 23/I-70 г.

Подписано к печати 5/V-70 г.

Бум. № 2 60×90¹/₁₆=10,75 бум. л.

Печ. л. 20,5. Уч.-изд. л. 19,77

Изд. № 4/5105. Цена 1 р. 54 к. Зак. 52

ИЗДАТЕЛЬСТВО «МИР»

Москва, 1-й Рижский пер., 2

Московская типография № 16

Главполиграфпрома Комитета по печати
при Совете Министров СССР.

Москва, Трехпрудный пер., 9