

НОВОЕ В ЗАРУБЕЖНОЙ ДЕМОГРАФИИ

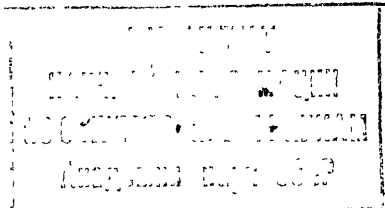
# СТАТИСТИЧЕСКИЙ АНАЛИЗ В ДЕМОГРАФИИ

Сборник статей под редакцией  
А. Г. ВОЛКОВА



МОСКВА «СТАТИСТИКА» 1980

ББК 60.7  
С78



НИИ ЦСУ СССР

ОТДЕЛ ДЕМОГРАФИИ

509  
80

## НОВОЕ В ЗАРУБЕЖНОЙ ДЕМОГРАФИИ

Редакционная коллегия:

В. А. Белова, В. А. Бирюков, А. Я. Боярский,  
А. Г. Волков, Л. Е. Дарский



*Вышли из печати сборники:*

1. РОЖДАЕМОСТЬ И ЕЕ ФАКТОРЫ. 1968 г.
2. МЕТОДЫ ДЕМОГРАФИЧЕСКИХ ИССЛЕДОВАНИЙ. 1969 г.
3. НАСЕЛЕНИЕ И ЭКОНОМИКА. 1970 г.
4. ТЕОРЕТИЧЕСКИЕ ПРОБЛЕМЫ ДЕМОГРАФИИ. 1970 г.
5. ИЗУЧЕНИЕ МНЕНИЙ О ВЕЛИЧИНЕ СЕМЬИ. 1971 г.
6. ДЕМОГРАФИЯ ПОКОЛЕНИЙ. 1972 г.
7. ДЕМОГРАФИЧЕСКИЕ ПРОГНОЗЫ. 1973 г.
8. БРАК И СЕМЬЯ. ДЕМОГРАФИЧЕСКИЙ АСПЕКТ. 1975 г.
9. УРБАНИЗАЦИЯ И РАССЕЛЕНИЕ. 1975 г.
10. ДЕМОГРАФИЧЕСКИЕ МОДЕЛИ. 1977 г.
11. ИЗУЧЕНИЕ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ. 1977 г.
12. РАЗВОД. ДЕМОГРАФИЧЕСКИЙ АСПЕКТ. 1979 г.
13. БРАЧНОСТЬ, РОЖДАЕМОСТЬ, СЕМЬЯ ЗА ТРИ ВЕКА. 1979 г.

*ГОТОВИТСЯ К ПЕЧАТИ  
СЛЕДУЮЩИЙ ВЫПУСК СЕРИИ:*

ИМИТАЦИОННЫЕ МОДЕЛИ В ДЕМОГРАФИИ

С  $\frac{10805^*-081}{008(01)-80}$  64-80 0703000000

\* Второй индекс 10803.

## ПРЕДИСЛОВИЕ

Демографические процессы издавна были и продолжают оставаться одним из главных объектов приложения статистических методов. Массовый характер демографических явлений, кажущаяся устойчивость закономерностей, сравнительная легкость наблюдения демографических событий — все это создавало благодатную почву для развития статистики, не говоря уже о практических нуждах изучения населения. Без статистических методов описание и анализ этих процессов кажутся невыполнимыми. Может быть, именно поэтому демография долгое время отожествлялась со статистикой населения и позже других областей знания, применяющих статистические методы, обособилась в самостоятельную область научного знания.

Действительно, демографические процессы представляют гораздо более, чем другие социально-экономические явления, подходящими для статистического исследования. Эти процессы — массовые явления, т. е. всегда слагаются из множества случаев рождения, одновременного существования или смерти людей. Это процессы достаточно однородные в том смысле, что речь всегда идет о переходе из одного демографического состояния в другое человеческих индивидов. Хотя каждый человек и его судьба единственны и неповторимы, всегда можно выделить достаточно однородную и достаточно многочисленную группу людей, сходных по значениям каких-либо признаков, и считать различия между ними в отношении других признаков несущественными. Демографические события, далее, можно при определенных допущениях рассматривать как случайные величины, и, следовательно, к ним применимы в принципе положения теории вероятностей.

Демографические процессы обладают тем свойством, что значения тех или иных признаков у отдельных объек-

тов могут изменяться, варьировать. Поэтому те или иные закономерности в них выступают только в среднем, при достаточно большом числе наблюдений. Для демографических процессов нельзя построить детерминистскую модель, которая по какому-либо одному значению измеряемой величины позволяла бы получить почти точное значение некоторой другой величины. Однако возможно построение вероятностных или стохастических моделей, позволяющих вычислить вероятность того, что некоторое значение искомой величины будет находиться в определенном интервале. Именно с помощью статистических методов мы отыскиваем в изменении демографических событий некоторую тенденцию, проявляющуюся при совместной обработке большого числа индивидуальных значений. Эта тенденция, освобожденная от влияния случайных колебаний, более или менее точно выражает закон, которому следует в своем изменении данное явление.

Современный этап развития демографии характеризуется все более широким применением математико-статистических методов к исследованию демографических явлений. Тем более важно предостеречь от формального их применения безотносительно к условиям места и времени, к содержанию поставленной исследовательской задачи. Каждый прием анализа, каждый метод основан на тех или иных предположениях и предпосылках. При решении конкретной задачи необходимо оценить, соблюдаются ли в данном случае необходимые для применения этого метода предпосылки, а если нет, то к каким результатам это может привести. Всегда следует помнить, что метод — лишь инструмент исследования, и поэтому особенно важен выбор такого инструмента, который соответствует поставленной задаче. Какой бы изощренной ни была методика исследования, она сама по себе не гарантирует точности результата и не страхует от ошибочных выводов. Особая опасность подстерегает исследователя в том случае, когда он обрабатывает — и часто довольно сложными способами — имеющийся фактический материал, а затем начинает «интерпретировать» полученные результаты. Такая интерпретация может быть совершенно произвольной, хотя скрупулезность обработки и создает иллюзию строгой обоснованности выводов. Статистический анализ сам по себе не показывает ничего. Та или иная статистическая обработка приносит плоды только в том случае, когда ее результаты подтверждают или опровергают априорно выдвинутую содержательную ги-

потезу. Никакой метод, пусть даже самый совершенный, не может заменить содержательного теоретического анализа, опирающегося на логические представления об изучаемом демографическом явлении. Поэтому никакое продвижение вперед в познании закономерностей демографических процессов и их социальной обусловленности невозможно без соответствующего развития теоретической демографии.

Все эти соображения по мере возможности учитывались при составлении сборника. Представляя в нем наиболее интересные применения современных приемов математико-статистического анализа, составитель стремился включить работы, наиболее тщательно продуманные в методическом отношении, такие, в основу исследования которых положены детально разработанные гипотезы, выбор метода обоснован с точки зрения особенностей решаемой задачи, а полученные результаты тщательно проверяются.

В современной зарубежной демографии статистические методы применяются в основном в трех направлениях: для обработки и статистического описания демографических процессов в случаях недостаточной или недостоверной информации; для анализа демографических закономерностей и социально-демографических связей и, наконец, для обобщения более или менее разнородных демографических характеристик и получения некоторых сводных показателей воспроизводства населения. Имея в виду практическое применение зарубежного опыта, мы представили в сборнике в основном два последних направления исследований.

Естественно, что приоритет отдавался статьям, наиболее интересным по содержанию, знакомящим с новыми результатами или раскрывающим новые аспекты исследования. Внимательный читатель заметит, что большинство статей сборника посвящено изучению динамики закономерностей и факторов рождаемости. Это не удивительно, так как рождаемость остается до сих пор одной из наиболее важных проблем демографии. К тому же современный аппарат статистического исследования применяется в демографическом анализе преимущественно в экономически развитых странах, где рождаемость низка и ее изучение особенно актуально.

Для современного этапа демографического анализа характерно изучение воздействия на демографические процессы одновременно ряда социально-демографических факторов и соизмерение силы их влияния. Для этого успешно

применяются различные варианты многомерной группировки, дающие возможность рассматривать воздействие на изучаемое явление как нескольких факторов вместе, так и каждого из них в отдельности при исключении влияния остальных.

Открывающая сборник статья Рафикула Хьюда Чаудхури посвящена изучению с помощью одного из методов многомерной группировки зависимости между благосостоянием и рождаемостью — проблемы, занимающей демографов уже на протяжении нескольких десятилетий. В отличие от своих предшественников автор оперирует не абсолютными, а относительными показателями уровня дохода, а в качестве характеристик рождаемости привлекает не только традиционный показатель числа рожденных детей, но и календарь рождений и мнения женщин о желаемом числе детей. Исследование ведется применительно к небольшой специально подобранной однородной совокупности жительниц Торонто (Канада) и заключается в последовательной проверке логически четко обоснованных гипотез о воздействии условий жизни на рождаемость. По существу, мы имеем дело с хорошо спланированным статистическим экспериментом. Следует отметить тщательность методики исследования, в частности подбора индикаторов для измерения тех или иных факторов, хотя некоторые из этих индикаторов и могут показаться искусственными (например, планы супругов относительно образования детей в качестве характеристики норм потребления). Учитывая слабую чувствительность метода многомерной группировки к взаимодействиям, автор специально исследует с помощью дисперсионного анализа эффект взаимодействия между испытываемыми переменными.

Исследование подтвердило положительную связь относительного дохода и рождаемости, однако она оказалась сильнее выраженной в группах более высокого социального статуса. Автор полагает, что предложенная им модель подходит лишь для социальных групп, имеющих мотивы ограничивать деторождение. Такое ограничение не умаляет ценности модели. Однако, несмотря на четкость исходных предпосылок модели, механизм связи благосостояния и демографического поведения остается все же недостаточно выясненным, что еще раз подчеркивает необходимость его теоретического осмысления.

Следующая статья сборника дает хороший пример применения множественной регрессии для изучения факторов,

определяющих территориальные демографические различия. Объектом исследования Джеймса Уида послужил известный факт нарастания частоты разводов по штатам США с Востока на Запад, который многие исследователи объясняли «духом Запада», особой атмосферой переселенческой вольницы. Наиболее существенным фактором различий в коэффициентах разводимости по штатам оказался не сам эффект переселенческой среды, а более раннее вступление в брак, ведущее к повышенной частоте разводов. Специальное изучение возможного влияния различий в законодательстве о разводах по штатам показало, что они не влияют заметно на показатели разводимости. Следует отметить не формальный, а содержательный подбор переменных для уравнения множественной регрессии. Для большей надежности выводов автор характеризует эффект переселенческой среды не одной, а тремя переменными, показывающими в общем одинаковый результат, и повторяет анализ для двух периодов, разделенных десятилетним промежутком. Методика этого исследования может оказаться полезной для аналогичного анализа по областям нашей страны на материалах переписи населения 1979 г.

В последние полтора-два десятилетия развивается так называемый спектральный анализ временных рядов. В отличие от обычного анализа временных рядов, когда составляющие ряда рассматриваются как функция времени, при спектральном анализе исследуется частота колебаний значений ряда вокруг некоторого неизменного или меняющегося среднего уровня. Анализ так называемого взаимного спектра двух рядов дает возможность измерить запаздывание колебаний одного ряда по сравнению с другим, усиление их при переходе на другой ряд, степень согласованности колебаний обоих рядов. Тем самым оценивается степень связанности изменений изучаемых явлений во времени. Работа Рональда Демоса Ли представляет собой один из немногих примеров применения спектрального анализа к изучению демографических явлений. Автор анализирует ряды чисел рождений и браков для нескольких популяций за длительный период (около 150 лет) с целью проверить гипотезы о влиянии брачности на рождаемость в доиндустриальной Европе. Он показывает, что более частые колебания в рядах чисел рождений вызывались колебаниями брачной рождаемости, а не колебаниями брачности, причем обнаруженная тесная связь колебаний брачности и брачной рождаемости предполагает существование намерен-

ного регулирования деторождения в браке, тогда как до сих пор предполагалось, что в доиндустриальной Европе сознательное ограничение деторождения отсутствовало. Анализ спектров дал автору возможность показать, кроме того, что сезонность рождаемости, по крайней мере для США, независима от сезонности браков, и что влияние экономических циклов на рождаемость через брачность не столь существенно, как это обычно предполагается. Интересные выводы получены и в отношении теории циклов поколений. Согласно «классической» демографии ряды чисел рождений изменяются циклично с периодом, соответствующим времени, необходимому для того, чтобы родившиеся дети сами вступили в брак и стали родителями. Спектральный подход позволил автору скорректировать и расширить это представление. Хотя применение спектрального анализа ограничено из-за отсутствия достаточно длинных рядов демографических данных, он представляет собой полезный инструмент в историко-демографических исследованиях и по мере накопления данных сулит большие перспективы.

Наряду с анализом влияния тех или иных частных закономерностей исследователей всегда привлекала задача отыскания характеристик, обобщающих совместное влияние комплекса условий воспроизводства населения в целом или отдельных демографических процессов. В последние годы, особенно в социалистических странах, стала весьма актуальной также задача демографического районирования, т. е. выделения территорий, сходных с точки зрения действия тех или иных факторов воспроизводства населения, что важно как для народнохозяйственного планирования, так и для разработки эффективной демографической политики. В методическом плане обе эти задачи сводятся к проблеме более компактного представления разнообразной информации о демографических процессах, которая часто решается с помощью таксономической классификации.

Обычная таксономия, в том числе и применяемая при демографической классификации, состоит в упорядочении объектов в пространстве. Тадеуш Грабиньский и Казимеж Заёнец применяют многомерную классификацию во временном аспекте. С помощью методов таксономии они пытаются провести периодизацию стадий развития демографических процессов в Польше на протяжении послевоенных лет. Эта работа представляет собой одну из немногих попыток применения таксономических методов к анализу многомерных временных рядов демографических явлений. Несомненным



достоинством работы служит то, что в ней весьма подробно представлены ход исследования и содержательные предпосылки, положенные в основу анализа, а также оценка надежности полученных результатов. Несмотря на коррелированность принятых для анализа переменных и сильную автокорреляцию в рядах, авторам удалось выделить в послевоенном демографическом развитии три достаточно выраженных периода.

В демографических исследованиях стала уже традиционной задача подбора подходящей кривой для выражения закономерностей изменения с возрастом тех или иных демографических явлений. Особенно многочисленными были попытки построения модели кривой возрастных коэффициентов рождаемости, начиная с исследования выдающегося советского статистика Б. С. Ястремского, за которым последовали работы Питера Мазура, Иво Лаха и др. В сборнике эта область приложения статистических методов представлена двумя работами.

В статье С. Митры и А. Романюка рассматриваются возможности применения для выравнивания ряда возрастных коэффициентов рождаемости кривой Пирсона первого типа, позволяющей восстановить все значения возрастных коэффициентов рождаемости в определенном интервале на основе значения модальной ординаты. Авторы предлагают метод упрощенной оценки параметров кривой и рассматривают возможности применения ее для оценки будущих показателей рождаемости на основе экстраполяции изменения среднего и модального возрастов деторождения.

Статья Дональда Р. Мак-Нейла, Т. Джеймса Трассела и Джона С. Тернера посвящена применению для этой цели так называемых сплайнов. Сплайн—это своеобразное чертежное лекало с изменяемой поверхностью. Он представляет собой длинную тонкую рейку, закрепленную на месте, с подвешенными к ней грузилами, придающими рейке некоторый изгиб. Изменяя точки подвески грузил и положение рейки, чертежник добивается такого изгиба рейки, при котором она проходила бы через заданные точки. Математическим сплайном называется приближенное представление деформированной таким образом рейки кусками параболы третьей или более высоких степеней (обычно разными между каждой парой смежных точек). Приведенный авторами пример выравнивания с помощью сплайна ряда возрастных коэффициентов рождаемости для Италии пока-

зывает преимущества метода, который может быть испробован и на других аналогичных кривых.

Нужно заметить, что ни тот, ни другой метод не претендует на аналитическое выражение закономерностей изменения рождаемости с возрастом: оба они — лишь технический прием обработки данных. Как особенности формы этой кривой, так и лежащие в ее основе закономерности изменения рождаемости остаются задачей самостоятельного теоретического исследования.

Приведенные в сборнике примеры применения статистических методов в демографических исследованиях не исчерпывают, конечно, ни всего разнообразия конкретных приемов исследования, ни тем более всех аспектов их применения. Тем не менее они дают некоторое представление о современной «технологии» демографического анализа и должны принести практическую пользу.

*А. Г. ВОЛКОВ*

*Рафикул Хьюда Чаудхури*

## ОТНОСИТЕЛЬНЫЙ ДОХОД И РОЖДАЕМОСТЬ

Rafiqul Huda Chaudhury. Relative Income and Fertility. *Demography*, vol. 14, № 2, May 1977, p. 179—195.

### ВВЕДЕНИЕ

Проведенные до сих пор исследования [7], [3], [8], [10], [2], [4], [6], в общем, показали, что существует положительная, хотя иногда и слабая, связь между относительным доходом и рождаемостью. Однако эти исследования отличаются рядом очевидных недостатков:

а) ни в одном из них нет прямой оценки желаемого уровня жизни, с которым супружеские пары могли бы сравнивать фактический уровень своей жизни;

б) чаще всего в них пользуются накопленными показателями рождаемости, а не характеристиками ее современного уровня;

в) при проверке гипотез о влиянии относительного дохода в них обычно уделяется мало внимания согласованности относительного дохода и переменных рождаемости.

В настоящей статье сделана попытка исправить эти недостатки.

Относительный доход и рождаемость определяются здесь разными способами. Мы ввели понятия относительного дохода в прошлом и относительного дохода в настоящее время и воспользовались несколькими показателями рождаемости, а именно интервалом между вступлением в брак и первым рождением, интервалами между последующими

рождениями и показателями рождаемости по продолжительности брака. (В нашем исследовании для проверки этой гипотезы впервые применяются в качестве характеристик рождаемости понятия «календарь рождений» и «желаемое число детей».)

В модели влияния относительного дохода на рождаемость нас интересуют следующие аспекты:

1. Изучение связи между относительным доходом и (а) нормами\* рождаемости, (б) нормами потребления, (в) репродуктивным поведением и (г) потребительским поведением.

2. Результаты проявления этих связей в целом. Для доказательства верности концептуальной основы модели требуется, чтобы (а) при изменении относительного дохода оставались неизменными нормы рождаемости; б) при изменении относительного дохода оставались неизменными нормы потребления; в) группы населения с низким относительным доходом отличались более низкой фактической рождаемостью, чем группы с высоким относительным доходом; г) при изменении относительного дохода не изменялось потребительское поведение.

3. Исследование других связей, которые могли бы показать, что поведение групп с низким относительным доходом отличается от поведения групп с высоким относительным доходом, и которые, по логике вещей, могли бы ослаблять влияние низкого уровня относительного дохода на поддержание стандартов потребления данной группы. Примером такой связи, поддающейся проверке в этом отношении, может служить связь между уровнем относительного дохода и тем, работает ли жена вне дома\*\*.

Для проверки модели влияния относительного дохода наиболее существенны пункты 1 и 2. Пункт 3 должен дать дополнительное подтверждение справедливости модели. Как отмечалось ранее, проводившиеся до сих пор исследования ограничивались пунктом 1(в), а единственный общий результат проявления связи, 2(в), и не мог быть никаким иным.

---

\* Норма понимается здесь не в статистическом, а в аксиологическом смысле, т. е. не как обычный «нормальный» уровень рождаемости, а как социально обусловленное и принятое в данной социальной группе представление о числе детей в семье (см. с. 29), в соответствии с которым супругам свойственно определенное поведение в этой сфере. Таким же образом понимаются и нормы потребления и потребительское поведение. — *Примеч. ред.*

\*\* Буквально «входит ли жена в состав рабочей силы» (wife's labor force participation). — *Примеч. ред.*

## ИСХОДНЫЕ ДАННЫЕ И МЕТОД ИССЛЕДОВАНИЯ

### Источник данных

Данные для этого исследования взяты из канадского исследования увеличения семьи [Canadian Family Growth Study], проведенного в метрополитенском Торонто\* в 1967—1968 гг. [1]. Канадское исследование увеличения семьи было основано на случайной выборке женщин в возрасте до 46 лет, состоящих в первом браке и проживающих вместе с мужьями. Данные собирались путем обхода домов и опроса по заранее составленной схеме. В процессе обследования были проведены 1632 полных опроса\*\* соответствующих респондентов.

### Исследуемое население

Наше исследование ограничивалось подвыборкой из данных канадского исследования увеличения семьи, включавшей коренных горожанок, способных иметь детей и успешно регулирующих деторождение. Опрашиваемыми, «успешно регулирующими деторождение», считались те, кто на вопрос «сколько раз вы беременели, несмотря на применение контрацептивов?» ответили: «Ни разу». С целью выяснения способности к деторождению (fecundity) всем женщинам в канадском исследовании увеличения семьи ставился вопрос: «Есть ли у вас причины полагать, что вы и ваш муж не сможете иметь еще одного ребенка, если бы захотели этого?». Ответившие «Нет» считались способными к деторождению. Аналогичным путем выяснялась способность к деторождению и в американском исследовании увеличения семьи [12]. «Коренными горожанками» считались женщины, родившиеся в метрополитенском ареале\*\*\* или в большом городе — высокоурбанизированном поселении с населением свыше 500 000 и в большом городе с населением 100 000—499 999. Коротко говоря, подвыборка «коренных горожанок, способных иметь детей и успешно

---

\* В центральной части Торонтской городской агломерации. —

*Примеч. ред.*

\*\* В которых были получены ответы на все вопросы. —

*Примеч. ред.*

\*\*\* В центральной части городской агломерации. — *Примеч. ред.*

регулирующих деторождение» включала женщин, родившихся в больших городах или в высокоурбанизированных поселениях Канады, сообщивших, что их способность к деторождению не снизилась и что они ни разу не беременели во время применения контрацептивов. Основу для анализа в настоящем исследовании составили в общей сложности 304 таких женщины. (В действительности, однако, в анализ включены не 304 женщины, а меньше, поскольку от некоторых из них не были получены ответы на один или несколько вопросов, ставившихся в этом исследовании.) Эта группа охватывает 23% всех способных к деторождению женщин и 19% всех женщин выборки.

Есть особые причины для того, чтобы ограничить исследование лишь «коренными горожанками, способными иметь детей и успешно регулирующими деторождение». Это соответствует той предпосылке принятой гипотезы относительного дохода, согласно которой пары намеренно контролируют деторождение и, таким образом, располагают средствами для выполнения своих решений в отношении деторождения. Устранение влияния фактора способности к деторождению необходимо для любого исследования дифференциальной рождаемости, цель которого — анализ небиологических факторов. Выборка была ограничена только коренными горожанками для того, чтобы исследовать более однородную совокупность, освободившись от влияния на вкусы и стиль жизни таких факторов, как место жительства и место рождения.

### Показатели относительного дохода

В этом исследовании мы пользуемся двумя показателями относительного дохода. *Относительный доход в настоящее время* вычисляется как соотношение между фактическим доходом мужа в 1967 г. и его ожидаемым (predicted) доходом, вычисленным на основе его теперешнего занятия, возраста, уровня образования и места рождения. *Относительный доход в прошлом* определяется соотношением дохода мужа на момент вступления в брак и его ожидаемого дохода, вычисленного на основе его возраста при вступлении в брак, занятия на момент вступления в брак, места рождения и уровня его образования на момент обследования.

Применение в качестве независимой переменной уровня образования на момент обследования вместо уровня обра-

зования на момент вступления в брак не вызвало смещения нашей оценки ожидаемого дохода при вступлении в брак в сторону увеличения. (Уровень образования в настоящий момент оказался наименее значимым фактором дохода при вступлении в брак. Бета-коэффициент для переменной «уровень образования» оказался равным 0,07, а для переменных «возраст при вступлении в брак», «место рождения» и «занятие на момент вступления в брак» — соответственно 0,17; 0,08 и 0,28.)

Ожидаемый доход был получен с помощью метода многомерной группировки (multiple classification analysis), который дает коэффициенты в виде отклонений от общей средней, стандартизованных с целью устранить эффект других независимых переменных.

Уровень относительного дохода классифицируется как (а) высокий, (б) средний и (в) низкий следующим образом. Ожидаемый (predicted) доход мужа кодируется по группам дохода так же, как и фактический доход. Если номер кода для фактического дохода равен номеру кода для ожидаемого дохода, то такой относительный доход попадает в категорию среднего дохода. Если номер кода для фактического дохода выше номера кода для ожидаемого дохода, то относительный доход считается «высоким», если же ниже, то относительный доход считается «низким».

### Методы анализа

Для анализа данных применяется метод многомерной группировки (ММГ). Этот метод дает среднее значение зависимой переменной для каждого подкласса независимой переменной и отклонение от общей средней, стандартизованное по эффектам всех других рассматриваемых переменных и одновременно по их взаимосвязям. ММГ основывается на аддитивной модели, т. е. модель предполагает, что каждая объясняющая переменная (explanatory variable) влияет на зависимую переменную безотносительно к значениям других независимых переменных (predictors). Нечувствительность ММГ к взаимодействиям побудила нас, прежде чем мы применили этот метод, исследовать эффекты взаимодействий. Интересуясь главным образом влиянием относительного дохода, мы проверили, была ли аддитивной взаимосвязь относительного дохода с испытываемыми переменными при всех уровнях других независимых переменных.

Для исследования эффекта взаимодействия между объясняющими переменными применялся дисперсионный анализ. Избранная нами модель предназначена для случая неравных частот в клетках корреляционной матрицы [13].

## РЕЗУЛЬТАТЫ

Результаты представлены в следующем порядке. Во-первых, рассматриваются результаты исследования связи между относительным доходом и репродуктивным поведением, нормами рождаемости и потребительским поведением. Во-вторых, обсуждается результат анализа этих связей в целом. В-третьих, представлены соображения, касающиеся возможности предвидеть некоторые типы поведения на основе модели относительного дохода.

### Анализ связи относительного дохода с репродуктивным поведением

Исследование связи относительного дохода в настоящее время с репродуктивным поведением проводилось для следующих показателей рождаемости:

- 1) накопленной рождаемости, измеряемой числом когда-либо рожденных детей;
- 2) показателей накопленной рождаемости по продолжительности брака — 0—4 года, 5—9 лет, 10—14 лет, 15 лет и более;
- 3) календаря рождений по интервалам между первым и вторым и между вторым и третьим рождениями.

В некоторых случаях для большей конкретизации гипотезы проводились дополнительные исследования.

Связь относительного дохода *в прошлом* с репродуктивным поведением рассматривалась на примере только одного показателя, как связь относительного дохода в прошлом с интервалом между моментом вступления в брак и первым рождением.

### Связь относительного дохода в настоящее время с общим числом когда-либо рожденных детей

Среднее число когда-либо рожденных детей в группах с высоким, средним и низким относительным доходом оказалось равным соответственно 2,12; 1,85 и 1,89.



Эти результаты согласуются с гипотезой, что среднее число детей наиболее велико в группах с высоким относительным доходом, но различия между группами относительного дохода статистически незначимы.

Однако различие в числе детей между крайними группами (т. е. между группой с высоким и группой с низким относительным доходом) оказывается статистически значимым на уровне значимости 0,10.

В приведенных результатах, основанных на измерении парной взаимосвязи (zero-order relationship), не устранен эффект других переменных, которые могут быть связаны с рождаемостью и, возможно, с относительным доходом. Для определения чистой взаимосвязи между относительным доходом и числом когда-либо рожденных детей сделана поправка на эффект следующих переменных: доход мужа, работа жены вне дома в данное время, работа жены вне дома в течение жизни, уровень образования жены, религия жены, занятие мужа, уровень образования мужа, продолжительность брака, возраст жены на момент вступления в брак, практика планирования семьи в данное время.

Кроме того, прежде чем применить ММГ для исследования влияния относительного дохода на рождаемость в чистом виде, внося поправку на эффект других независимых переменных (predictor variables), мы исследовали эффект взаимодействия относительного дохода с другими независимыми переменными, учитываемыми в модели. Оказалось, что из всех независимых переменных, учитываемых в модели, относительный доход значимо взаимодействует с уровнем образования и занятием мужа, т. е. влияние относительного дохода на рождаемость существенно изменяется в зависимости от уровня образования и занятия мужа. Это взаимодействие ясно видно на рис. 1 и 2, где кривые для групп с высоким, средним и низким уровнем образования (рис. 1) и для занятых физическим и умственным трудом (рис. 2) пересекаются.

В табл. 1 представлены парные нестандартизованные и стандартизованные (с исключением влияния других переменных) взаимосвязи между относительным доходом и рождаемостью для каждого уровня образования мужа и для групп, занятых физическим и умственным трудом. Покажем кратко, как интерпретировать табл. 1 и последующие таблицы, полученные с помощью ММГ. Числа в строках «нестандартизованные» и «стандартизованные» даны в виде отклонений от общей средней. Например, — 0,15 вверху

первого столбца табл. 1 означает, что лица, принадлежащие к группе с высоким относительным доходом, с низким уровнем образования имеют детей на 0,15 меньше среднего их числа для группы.

Из табл. 1 можно видеть, что до поправки на эффект других переменных у супружеских пар из группы с низким уровнем образования и у работников физического труда между относительным доходом и рождаемостью существует

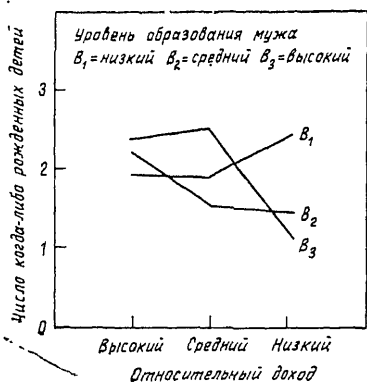


Рис. 1. Характер взаимосвязи между относительным доходом и числом когда-либо рожденных детей для каждого уровня образования мужа

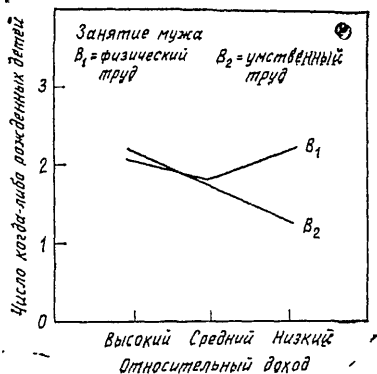


Рис. 2. Характер взаимосвязи между относительным доходом и числом когда-либо рожденных детей для каждого рода занятий мужа

обратная связь. Однако, когда влияние этих прочих факторов устраняется, эта связь почти исчезает. С другой стороны, у супружеских пар в группах со средним и высоким уровнем образования, так же как и у работников умственного труда, связь относительного дохода и рождаемости оказывается положительной как по стандартизованным, так и по нестандартизованным отклонениям.

Таким образом, приведенные результаты заставляют предположить, что связь между относительным доходом и рождаемостью имеет ограниченный характер, поскольку она подтверждается лишь для средней и высшей социально-экономических групп.

Перейдем теперь к исследованию независимого влияния относительного дохода на рождаемость после устранения эффекта тех независимых переменных, которые также свя-

**Нестандартизованные и стандартизованные отклонения  
в среднем числе рожденных детей  
по относительному доходу в настоящее время  
для разных уровней образования и занятия мужа**

Уровень образования и занятие мужа	Относительный доход в настоящее время		
	высокий	средний	низкий
Уровень образования: низкий (8—10 лет):			
<i>n</i>	44	26	40
нестандартизован- ные*	-0,15	-0,22	0,31
стандартизован- ные <sup>1*</sup>	-0,04	0,09	-0,01
средний (11—13 лет):			
<i>n</i>	56	29	21
нестандартизованные	0,31	-0,33	-0,37
стандартизованные	0,13	-0,12	-0,19
высокий (колледж и выше):			
<i>n</i>	22	8	15
нестандартизованные	0,40	0,50	-0,86
стандартизованные	0,12	0,31	-0,34
Занятые:			
физическим трудом:			
<i>n</i>	57	45	51
нестандартизованные	0,01	-0,23	0,18
стандартизованные	-0,04	0,02	0,02
умственным трудом:			
<i>n</i>	66	22	27
нестандартизованные	0,22	-0,003	-0,65
стандартизованные	0,27	0,12	-0,24

\* Нестандартизованные отклонения от общей средней.

\*\* Стандартизованные отклонения от общей средней; исключено влияние следующих переменных: (1) работа жены вне дома; (2) занятость жены в прошлом; (3) продолжительность брака; (4) уровень образования жены; (5) религия; (6) практика планирования семьи в данное время; (7) доход мужа; (8) возраст при вступлении в брак.

Примечание. Общая средняя для низкого уровня образования равна 2,10, для среднего уровня 1,85, для высокого уровня 1,33, для занятых физическим трудом 2,05, для занятых умственным трудом 1,91.

заны с рождаемостью и относительным доходом, но взаимно действуют с относительным доходом незначимо. В табл. 2 даны стандартизованные и нестандартизованные отклонения в среднем числе детей по относительному доходу в настоящее время для всей исследуемой подвыборки.

Нестандартизованные и стандартизованные отклонения  
в среднем числе рожденных детей  
по относительному доходу в настоящее время

Относительный доход в настоящее время	Число случаев	Среднее число рожденных детей	Отклонения от общей средней		
			нестандартизованные	стандартизованные 1*	стандартизованные 2**
Высокий	123	2,13	0,14	0,05	0,05
Средний	67	1,85	-0,14	0,03	0,03
Низкий	78	1,89	-0,10	-0,11	-0,10

\* Стандартизованы по всем переменным, кроме указанных под номерами (1) и (2) в списке \*\*к табл. 1.

\*\* Стандартизованы по всем переменным, перечисленным в списке \*\* к табл. 1.

Примечание. Независимый эффект относительного дохода (полученный после повторного прогона программы с исключением рассматриваемой независимой переменной из анализа и учитывающий уменьшение общей объясненной суммы квадратов) составляет 0,70; общая дисперсия, объясняемая вариацией всех независимых переменных, равна 46,46. Общая средняя равна 1,99.

В табл. 2 показаны два ряда стандартизованных данных. Ряд 1 не включает работу жены вне дома и ее занятость в прошлом в число независимых переменных, влияние которых устранялось, чтобы найти чистое влияние относительного дохода на рождаемость, а ряд 2 включает в число устраняемых и эти два фактора. Работа жены вне дома и ее занятость в прошлом оставлены исходя из предположения, что для семей с доходом ниже нормального поступление на работу вне дома — весьма существенная возможность как повысить доход семьи, так и сдержать увеличение семьи и что попытки семей действовать таким образом улавливаются этими двумя переменными. Поэтому для того, чтобы измерить влияние на рождаемость относительного дохода мужа, нам нужно было оставить работу жены вне дома и ее занятость в прошлом как независимые переменные и допустить, чтобы влияние этих двух факторов проявлялось как реакция на относительный доход. Однако изучение табл. 2 показывает, что влияние относительного дохода на рождаемость остается почти неизменным, независимо от того, входят или не входят работа жены вне дома и ее занятость в прошлом в число независимых переменных. Этот результат

не был неожиданным, поскольку корреляция таких переменных, как работа жены вне дома и ее занятость в прошлом, с относительным доходом мала. (Коэффициенты корреляции ( $r$ ) между работой жены вне дома и относительным доходом и между занятостью жены в прошлом и относительным доходом оказались равными соответственно 0,08 и 0,07.)

Сравнение среднего числа рожденных детей в табл. 2 свидетельствует о небольших различиях в рождаемости между супружескими парами со средним и низким относительным доходом.

Однако группа с высоким относительным доходом дает отклонение от общей средней в том направлении, что и предполагалось. Если устранить влияние других контрольных переменных, то положительная связь между относительным доходом и числом рожденных детей остается заметной, хотя и слабой, причем разница в числе детей между крайними группами по доходу составляет 0,15.

Поскольку модель относительного дохода предполагает сознательное планирование некоторыми супружескими парами как деторождения, так и потребления, то можно было бы ожидать, что гипотеза о влиянии относительного дохода более четко подтвердится среди тех, кто, как правило, планирует свою жизнь. В обследовании был один вопрос, который позволил нам принять это во внимание: «Некоторые люди строят определенные планы на будущее и планируют свою жизнь во многих отношениях на несколько лет вперед. А как поступаете вы?». Стандартизованные и нестандартизованные отклонения от среднего числа рожденных детей по относительному доходу для тех, кто, как правило, планирует свою жизнь на несколько лет вперед, приведены далее (среднее число рожденных детей в группах с высоким, средним и низким относительным доходом составляет соответственно 2,02; 1,29 и 1,20).

Отклонения	Относительный доход		
	высокий ( $n=48$ )	средний ( $n=27$ )	низкий ( $n=20$ )
Нестандартизованные	0,38	-0,34	-0,44
Стандартизованные	0,09	-0,04	-0,17

Здесь мы видим, что связь между относительным доходом и рождаемостью выражена более явно, чем в табл. 2, как по стандартизованным, так и по нестандартизованным данным. В то время как разность в числе детей между группами с вы-

соким и низким относительным доходом в нестандартизованных рядах всей подвыборки (см. табл. 2) составляет 0,24, соответствующая разность для супружеских пар, ответивших, что они планируют свою жизнь, составляет 0,82. Хотя эта разность и сокращается после стандартизации, но среди тех, кто обычно планирует свою жизнь, различия в относительном доходе все же объясняют более значительную часть вариации рождаемости: 1,20% по сравнению с 0,7% во всей подвыборке.

### **Связь относительного дохода в настоящее время с рождаемостью, по продолжительности брака**

Исследование связи между относительным доходом и рождаемостью в группах с разной продолжительностью брака преследовало цель достигнуть большей согласованности во времени показателей относительного дохода и накопленной рождаемости. В табл. 3 раскрывается характер связи между относительным доходом и рождаемостью для каждой группы по продолжительности брака. Можно видеть, что если устраняется эффект других независимых переменных, то между относительным доходом и рождаемостью в каждой группе существует положительная связь. Эта связь более выражена среди тех, кто состоит в браке 0—4 года и 5—9 лет, хотя иногда число наблюдений невелико.

Далее, табл. 3 показывает, что часть вариации рождаемости, объясняемая только вариацией относительного дохода, исключая другие переменные, наиболее велика в группе состоящих в браке менее 10 лет (см. примечание к табл. 3. — *Примеч. ред.*).

В соответствии с нашей гипотезой мы ожидали, что связь относительного дохода с рождаемостью наиболее сильно проявится в группе женщин, состоящих в браке 0—4 года, поскольку именно для этой группы период, за который взят относительный доход, более всего соответствует периоду, за который взята накопленная рождаемость. Однако связь оказалась наиболее тесной в группе состоящих в браке 5—9 лет (см. табл. 3). Возможно, что одна из причин этого — высокая доля добрачных зачатий в группе с низким относительным доходом. Из общего числа 30 добрачных зачатий 13 пришлось на супружеские пары, принадлежащие к группе с низким относительным доходом. Это равносильно не-

Нестандартизованные и стандартизованные отклонения  
в среднем числе рожденных детей  
по относительному доходу и продолжительности брака

Продолжительность брака	Относительный доход		
	высокий	средний	низкий
0—4 года:			
<i>n</i>	15	18	18
нестандартизованные <sup>1</sup>	0,30	-0,01	-0,23
стандартизованные <sup>2</sup>	0,12	-0,007	-0,10
5—9 лет:			
<i>n</i>	29	20	24
нестандартизованные	-0,06	0,19	-0,08
стандартизованные	0,12	0,12	-0,24
10—14 лет:			
<i>n</i>	33	12	18
нестандартизованные	-0,15	-0,06	0,32
стандартизованные	0,03	0,16	-0,17
15 лет и более:			
<i>n</i>	46	17	18
нестандартизованные	-0,03	-0,04	0,13
стандартизованные	-0,02	0,19	-0,12

\* Нестандартизованные отклонения от общей средней.

\*\* Стандартизованные отклонения от общей средней; исключено влияние всех переменных кроме (3) из сноски \*<sup>1</sup> к табл. 1.

Примечание. Независимый эффект относительного дохода (полученный после повторного прогона программы с исключением рассматриваемой независимой переменной из анализа и учитывающий уменьшение общей объясненной суммы квадратов) составляет 1,20 для 0—4 лет пребывания в браке; 1,60 — для 5—9 лет пребывания в браке; 0,40 — для 10—14 лет пребывания в браке и 0,60 — для 15 и более лет пребывания в браке. Общая дисперсия, объясняемая вариацией всех независимых переменных, равна 52,16 (0—4 года брака), 23,95 (5—9 лет), 11,87 (10—14 лет) и 9,54 (15 лет и более). Средние числа рожденных детей для выделенных групп по продолжительности брака равны соответственно 0,62; 1,75; 2,39 и 2,75.

пропорциональному увеличению периода риска иметь ребенка, а возможно, и повышению рождаемости в группе с низким относительным доходом. Стоило исключить супружеские пары, имевшие добрачные зачатия, и ограничить анализ парами, зачавшими только после вступления в брак, как обнаружилась весьма тесная связь относительного дохода и рождаемости в группе с продолжительностью брака 0—4 года. Для всех супружеских пар с продолжительностью брака 0—4 года разность в числе детей между группами с вы-

соким и низким доходом составляла по стандартизованным данным 0,22; когда мы ограничили анализ парами, зачавшими только после вступления в брак, соответствующая разность достигла величины 0,40.

Итак, хотя число входящих в эту группу и невелико (всего 21 случай), обнаружение тесной связи между относительным доходом и рождаемостью в группе с более короткой продолжительностью брака для зачавших после вступления в брак согласуется с нашей гипотезой.

### **Связь относительного дохода с календарем рождений**

В этом разделе мы рассматриваем связь относительного дохода с интервалом между вступлением в брак и первым рождением, а также с календарем рождений следующих очередностей для женщин, родивших трех детей. Наша цель — уточнить зависимую переменную и достичь большей согласованности во времени между мерой относительного дохода и мерой рождаемости. Для этого мы оценили относительный доход на момент вступления в брак (относительный доход в прошлом) и соотнесем его с интервалом между вступлением в брак и первым рождением. Интервалы между первым и вторым рождениями и между вторым и третьим рождениями мы исследуем в связи с относительным доходом в настоящее время.

В соответствии с гипотезой о влиянии относительного дохода мы ожидаем, что относительный доход будет находиться в обратном отношении к интервалу между вступлением в брак и первым рождением, а также к интервалам между рождениями следующих очередностей. Эта гипотеза строится на предположении о том, что если супружеские пары из группы с низким относительным доходом действительно сокращают свою рождаемость под влиянием переживаемых ими сейчас материальных затруднений, то им также понадобится и сравнительно много времени для того, чтобы произвести на свет данное число детей. Более растянутый период деторождения у супружеских пар с низким относительным доходом дает женщине сравнительно больше возможностей для работы вне дома и тем самым способствует сохранению стандартов потребления данной группы. Остерия [11] определил для данной выборки, что интервал между рождениями положительно связан с накоплением имущества. Кроме того, более длительные интервалы между



рождениями в группе с низким относительным доходом менее обременительны для бюджета семьи.

Сначала исследовалась связь между относительным доходом в прошлом и интервалом между вступлением в брак и первым рождением. (Супружеские пары с детьми, зачатыми до брака были исключены из анализа по двум причинам. Во-первых, для добрачных рождений вычисление календаря рождений было бы произвольным, поскольку нет единого способа определить точку отсчета для первого интервала. Во-вторых, случаев добрачного рождения немного (всего 30) и исходная связь между относительным доходом в прошлом и интервалом между вступлением в брак и первым рождением остается почти неизменной независимо от того, включены добрачные рождения в анализ или нет.)

В отношении связи между относительным доходом в прошлом и интервалом между вступлением в брак и первым рождением следовало ожидать более четкого подтверждения гипотезы относительного дохода, поскольку обе величины — и относительный доход и мера рождаемости (интервал) — хорошо согласуются во времени. Далее приведены нестандартизованные и стандартизованные отклонения от общей средней (3,20 года), отражающие связь между относительным доходом в прошлом и интервалом с момента вступления в брак до первого рождения.

Отклонения	Относительный доход		
	высокий ( $n=73$ )	средний ( $n=71$ )	низкий ( $n=75$ )
Нестандартизованные	-0,43	-0,11	0,53
Стандартизованные	-0,29	0,07	0,22

Согласно стандартизованным данным, группа с низким доходом ждала появления первых детей по крайней мере на шесть месяцев (0,51 года) дольше, чем группа с высоким доходом. Относительный доход в прошлом объясняет 3% всей вариации длительности интервала между вступлением в брак и первым рождением, исключая вариацию, объясняемую другими независимыми переменными. Эти результаты согласуются с гипотезой относительного дохода.

Затем мы рассмотрели связь относительного дохода с интервалом между первым и вторым и между вторым и третьим рождениями.\* Интервалы между последующими рождениями не рассматривались, поскольку более трех детей имели лишь немногие супружеские пары. Предполагалось, что

Нестандартизованные и стандартизованные отклонения  
в среднем числе лет, прошедших между первым и вторым  
и между вторым и третьим рождениями  
по относительному доходу в настоящее время

Интервалы между рождениями	Относительный доход		
	высокий	средний	низкий
Между первым и вто- рым рождениями:			
<i>n</i>	94	39	48
нестандартизованные*	-0,09	0,02	-0,20
стандартизованные**	0,03	0,003	-0,06
Между вторым и треть- им рождениями:			
<i>n</i>	40	18	22
нестандартизованные	0,07	0,16	-0,26
стандартизованные	-0,64	-0,12	1,26

\* Нестандартизованные отклонения от общей средней.

\*\* Стандартизованные отклонения от общей средней; исключено влия-  
ние переменных, перечисленных в списке \*\* к табл. 1.

Примечание. Независимый эффект относительного дохода (полученный после повторного прогона программы с исключением рассматриваемой независимой переменной из анализа и учитывающей уменьшение общей объясненной суммы квадратов) составляет 0,15 для интервала между первым и вторым рождениями и 4,02 — для интервала между вторым и третьим рождениями. Общая дисперсия, объясняемая вариацией всех независимых переменных, для интервалов между первым и вторым и между вторым и третьим рождениями равна соответственно 14,46 и 21,52. Общая средняя для интервалов между первым и вторым и между вторым и третьим рождениями равна соответственно 3,20 и 3,45.

гипотеза относительного дохода более четко подтвердится тогда, когда будет рассматриваться интервал между вторым и третьим рождениями, а не интервал между первым и вторым рождениями, так как решение иметь третьего ребенка принимается, вероятно, более осторожно, чем решение иметь второго. В США и Канаде трех детей имеют далеко не все супружеские пары и интервал между рождением второго и третьего ребенка больше, чем между рождением первого и второго. Следовательно, решение иметь более двух детей может быть связано с гораздо более тщательной оценкой таких социальных и экономических факторов, как доход, уровень относительного дохода и т. д. Поэтому мы ожидали, что различия в длительности интервала, обусловленные различиями в относительном доходе, возрастают с каждым рождением.

**Нестандартизованные и стандартизованные отклонения  
в среднем числе лет, прошедших между вторым и третьим  
рождениями, для женщин, имеющих трех и более детей,  
и для женщин, имеющих только трех детей**

	Относительный доход		
	высокий	средний	низкий
Женщины, имеющие трех и более детей:			
<i>n</i>	40	18	22
нестандартизованные*	0,07	0,16	-0,26
стандартизованные <sup>1</sup>	-0,64	-0,12	1,26
Женщины, имеющие только трех детей:			
<i>n</i>	29	9	10
нестандартизованные	-0,12	0,19	0,18
стандартизованные	-0,65	0,08	1,81

\* Нестандартизованные отклонения от общей средней.

<sup>1</sup> Стандартизованные отклонения от общей средней; исключено влияние всех переменных, перечисленных в сноске \*<sup>1</sup> к табл. 1.

**Примечание.** Независимый эффект относительного дохода (полученный после повторного прогона программы с исключением рассматриваемой независимой переменной из анализа и учитывающий уменьшение общей объясненной суммы квадратов) составляет 4,02 для женщин с тремя и более детьми и 7,88 — для женщин только с тремя детьми. Общая дисперсия, объясняемая вариацией всех независимых переменных, для женщин с тремя и более детьми и для женщин только с тремя детьми равна соответственно 21,52 и 21,99. Общая средняя для женщин с тремя и более детьми и для женщин только с тремя детьми равна соответственно 3,45 и 3,91.

Табл. 4 показывает тесную обратную связь между относительным доходом и интервалом между вторым и третьим рождениями, даже когда вносится поправка на эффект других независимых переменных, причем связь, гораздо более тесную, чем для интервала между первым и вторым рождениями. Как показывают стандартизованные данные, промежуток времени между первым и вторым рождениями в группах с высоким, средним и низким относительным доходом составляет соответственно 3,23, 3,20 и 3,14 года. В противоположность этому тот же промежуток между вторым и третьим рождениями для групп с высоким, средним и низким относительным доходом составляет соответственно 2,81, 3,33 и 4,71 года. Относительный доход объясняет 0,15% всей вариации длины интервала между первым и вторым рождениями, исключая вариацию, объясняемую другими незави-

## Процент желающих, чтобы их дети продолжили свое образование после средней школы, по относительному доходу\*

Продолжат ли дети свое образование после средней школы	Относительный доход		
	высокий	средний	низкий
Да:			
<i>n</i>	110	58	68
процент	88	86	87
Нет:			
<i>n</i>	3	3	2
процент	2	4	2
Не решено:			
<i>n</i>	7	3	5
процент	5	4	6
Прочие:			
<i>n</i>	5	3	3
процент	4	4	4
Всего:			
<i>n</i>	125	67	78
процент	100	100	100

\* Сумма процентов может оказаться не равной 100, так как десятичные знаки опущены.

симыми переменными. Соответствующая величина для интервала между вторым и третьим рожденьями составляет 4,0% (см. табл. 4 и примечание к ней).

Для того чтобы достичь еще большей согласованности во времени между интервалом от второго до третьего рожденья и мерой относительного дохода в настоящее время, мы рассматривали эти соотношения для женщин, имеющих только трех детей. При таком ограничении обнаруживается отчетливая обратная связь между относительным доходом и интервалом между вторым и третьим рожденьями (см. табл. 5). Согласно стандартизованному данным, женщинам из группы с низким относительным доходом, имеющим трех детей, для рождения третьего ребенка потребовалось на 2,46 года больше, чем женщинам из группы с высоким относительным доходом. Относительный доход для женщин, имеющих только трех детей, объясняет 8,0% всей вариации длины интервала от второго до третьего рожденья, исключая вариацию, объясняемую другими независимыми переменными (см. табл. 5 и примечание к ней).

Эти результаты, учитывающие интервалы между рожденьями — наиболее убедительное подтверждение гипотезы

относительного дохода. Подтверждение гипотезы, полученное нами, отчасти может объясняться более тщательным согласованием во времени мер относительного дохода и рождаемости. Кроме того, календарь рождений, особенно интервал между вторым и третьим рождением, может быть более важной переменной при проверке гипотезы относительного дохода, чем число когда-либо рожденных детей.

### Связь между относительным доходом и желаемым числом детей

Модель относительного дохода предполагает, что в то время, как фактическая рождаемость варьирует в зависимости от уровня относительного дохода, отражая реакцию на нормы потребления, нормы рождаемости от уровня относительного дохода не зависят. За индикатор нормы рождаемости принято число желаемых детей, которое женщина назвала во время опроса. (Составляет ли указанная при опросе величина семьи норму в социологическом смысле, т. е. отражают ли выраженные предпочтения те установки и ценности, которые одобряются данной социальной группой, — проблема, широко обсуждаемая в литературе. Здесь она решена быть не может. Однако одно из последних исследований подтверждает мысль о том, что социальная группа оказывает давление на те супружеские пары, которые отклоняются от групповых стандартов в деторождении [9]). Анализируя связь между относительным доходом и желаемым числом детей, мы получили следующие нестандартизованные и стандартизованные отклонения от общей средней (среднее число желаемых детей в группах с высоким, средним и низким доходом составляло соответственно 2,99; 3,01 и 3,01):

Отклонения	Относительный доход		
	высокий ( $n=121$ )	низкий ( $n=66$ )	средний ( $n=77$ )
Нестандартизованные	-0,01	0,01	0,009
Стандартизованные	0,04	-0,01	-0,05

Как можно видеть по нестандартизованным данным, между группами по относительному доходу различий в желаемом числе детей, по существу, нет; после стандартизации возникает слабая прямая связь, но различия незначимы. Таким образом, результат согласуется с предполагавшимся характером связи между относительным доходом и желаемой рождаемостью.

## Связь относительного дохода с нормами потребления и потребительским поведением

Модель относительного дохода предполагает, что ни нормы потребления, ни потребительское поведение не различаются между группами по относительному доходу. Торонтское исследование было бедно вопросами, необходимыми для определения норм потребления и потребительского поведения. Несмотря на это, было решено предпринять некоторые попытки рассмотреть эти нормы и поведение при проверке предположений, сделанных применительно к концептуальной основе модели относительного дохода. Индикатором норм потребления служили планы супружеской пары относительно образования ее детей, а индикатором потребительского поведения — количество собственных автомобилей.

В табл. 6 приведены нестандартизованные показатели связи между относительным доходом и планами супружеской пары относительно образования ее детей. Здесь можно видеть, что между группами по относительному доходу, по существу, нет различий в доле тех, кто хотел бы, чтобы их дети продолжали образование после средней школы.

Очень незначительными оказались и различия между группами по относительному доходу и по признаку владения автомобилями. Далее приводятся нестандартизованные и стандартизованные отклонения от общей средней, показывающие связь между относительным доходом и числом собственных автомобилей (среднее число собственных автомобилей в группах с высоким, средним и низким относительным доходом равно соответственно 1,28, 1,07 и 1,00).

Отклонения	Относительный доход		
	высокий ( $n=125$ )	средний ( $n=67$ )	низкий ( $n=78$ )
Стандартизованные	0,13	-0,07	-0,15
Нестандартизованные	0,08	-0,06	-0,07

Слабая положительная связь, обнаруженная по нестандартизованным данным, становится еще слабее после исключения влияния других переменных. Эти результаты подтверждаются данными табл. 7, которые показывают, что при каждом уровне относительного дохода коэффициент вариации для числа детей больше, чем для числа автомобилей. Иными словами, показатели групп по относительному доходу меньше отличаются от общей средней для числа автомобилей, чем для числа детей.

Средние, стандартные отклонения и коэффициенты вариации\*  
в числе рожденных детей и числе собственных автомобилей  
по относительному доходу

	Относительный доход		
	высокий	средний	низкий
Число рожденных детей:			
средняя	2,12	1,85	1,89
стандартное отклонение	1,22	1,31	1,39
коэффициент вариации	0,57	0,70	0,73
Число собственных автомобилей:			
средняя	1,27	0,52	0,40
стандартное отклонение	1,07	0,53	0,49
коэффициент вариации	1,00	0,52	0,47

\* Коэффициент вариации есть отношение стандартного отклонения к средней.

## РЕЗУЛЬТАТЫ АНАЛИЗА

Для оценки адекватности модели относительного дохода мы исследовали связь между относительным доходом и а) репродуктивным поведением, б) нормами рождаемости, в) нормами потребления и г) потребительским поведением. Во введении к статье было указано, что адекватность модели зависит от определенных результатов ее испытания, а именно от того, что:

а) репродуктивное поведение должно меняться в прямой зависимости от относительного дохода;

б) нормы рождаемости не должны обнаруживать связи с относительным доходом;

в) нормы потребления не должны обнаруживать связи с относительным доходом;

г) потребительское поведение не должно обнаруживать связи с относительным доходом.

Результаты всех четырех испытаний в общем подтверждают правильность модели относительного дохода.

Итак, мы подошли к вопросу, будет ли относительный доход обладать прогностической силой\*, кроме указанной,

\* Т. е. можно ли по значениям относительного дохода как независимой переменной определять значения каких-либо зависимых переменных. — *Примеч. ред.*

причем такой, чтобы она не противоречила четким гипотезам, положенным в основу модели. Иными словами, связан ли относительный доход с чем-нибудь, кроме репродуктивного поведения, норм рождаемости, норм потребления и потребительского поведения,—с чем-нибудь, что можно было бы логически включить в модель? К этому вопросу мы обратимся в следующем разделе.

### ДОПОЛНИТЕЛЬНОЕ ИССЛЕДОВАНИЕ ДЛЯ ДОКАЗАТЕЛЬСТВА ПРАВИЛЬНОСТИ МОДЕЛИ ОТНОСИТЕЛЬНОГО ДОХОДА

Модель относительного дохода предполагает, что группы с низким относительным доходом сдержанны в своем репродуктивном поведении (т. е. что они ограничивают число детей и/или увеличивают интервалы между рожденьями) для того, чтобы их потребление соответствовало нормам потребления группы. Возникает вопрос о том, объясняет ли низкий относительный доход также и поведение, которое способствует ограничению деторождения или увеличивает материальные возможности достичь норм потребления группы. Пример поведения, способствующего *ограничению* деторождения для того, чтобы достичь норм потребления, — это практика более ранней и более эффективной контрацепции. Пример поведения, способствующего *увеличению* материальных возможностей для того, чтобы достичь норм потребления, — работа жены вне дома.

В табл. 8 показана в каждой из трех групп по относительному доходу доля супружеских пар, практикующих контрацепцию, по числу детей. Эти данные подтверждают предположение, что в группе с низким относительным доходом бездетные пары прилагают больше усилий с помощью контрацептивов не допустить или отложить рождение первого ребенка, чем супружеские пары с другим уровнем относительного дохода. Различия в доле применяющих контрацептивы между группами по относительному доходу в пределах каждой группы по числу детей статистически незначимы. Однако среди бездетных пар уровень применения контрацепции между группами по относительному доходу различается значимо. Фактически в группе с низким относительным доходом применяли контрацептивы на момент обследования все бездетные пары, тогда как в группе с высоким относительным доходом — только 54,5% пар, и это различие оказалось статистически значимым ( $\chi^2 = 8,54$  при



Процент пользующихся контрацептивами\* в настоящее время по относительному доходу и числу детей среди способных к деторождению коренных горожанок, успешно регулирующих деторождение

Относительный доход	Женщины с числом детей				Все женщины
	0	1	2	3 и более	
Высокий: <i>n</i>	11	19	54	41	125
процент пользующихся контрацептивами	54,5	89,4	83,3	87,8	83,2
Средний: <i>n</i>	11	17	21	18	67
процент пользующихся контрацептивами	81,8	70,5	90,4	88,8	83,5
Низкий: <i>n</i>	15	15	26	22	78
процент пользующихся контрацептивами	100,0	86,6	92,3	90,9	92,3

\* Для выявления пользующихся контрацептивами обследуемым был задан вопрос: «Применяете ли вы какой-нибудь из этих методов (предлагался список методов контрацепции) для предотвращения зачатия?».

Примечание.  $\chi^2=12,7$ , число степеней свободы равно 8, незначимо на уровне значимости 0,10. (Нулевая гипотеза состоит в том, что для каждой из групп по числу детей поведение, т. е. процент применяющих контрацептивы, не различается между группами по относительному доходу.)

2 степенях свободы, уровень значимости 0,02). Таким образом, по-видимому, для супружеских пар из группы с низким относительным доходом отсрочка рождения первого ребенка представляет собой одно из средств достичь норм потребления группы.

Далее был изучен вопрос, ограничивают ли деторождение супружеские пары с низким доходом для достижения определенных норм потребления посредством того, что жены работают вне дома чаще, чем в группах с другим уровнем относительного дохода. Есть доказательства того, что это, по-видимому, именно так. Табл. 9 показывает, что за исключением бездетных пар в группе с низким относительным доходом, по всей вероятности, женщин вне дома работает больше, чем в группах с высоким относительным доходом. В общем, доля работающих женщин в группе с низким относительным доходом более чем на 50% выше, чем в груп-

Процент женщин, работающих вне дома,  
по относительному доходу и числу детей среди способных  
к деторождению коренных горожанок,  
успешно регулирующих деторождение

Относительный доход	Женщины с числом детей				Все женщины
	0	1	2	3 и более	
Высокий:					
<i>n</i>	11	19	54	41	125
процент работающих	100,0	10,5	29,6	19,5	29,6
Средний:					
<i>n</i>	11	17	21	18	67
процент работающих	72,7	17,6	14,2	16,6	25,3
Низкий:					
<i>n</i>	15	15	26	22	78
процент работающих	86,6	13,3	34,6	50,0	44,8

Примечание.  $\chi^2=14,5$ , число степеней свободы равно 8, значимо на уровне 0,10, но очень близко к уровню 0,05. (Нулевая гипотеза состоит в том, что для каждой из групп по числу детей поведение, т. е. работа женщины вне дома, не различается между группами по относительному доходу.)

пе с высоким относительным доходом. Различия в степени занятости женщин между группами по относительному доходу оказались статистически значимыми также в пределах групп по числу детей.

Приведенные результаты, не предусмотренные конкретно в модели относительного дохода, но согласующиеся с ней, указывают на дополнительные прогностические возможности модели. Поэтому можно предположить, что другие типы поведения, которые либо способствуют ограничению деторождения (например, откладывание брака), либо усиливают его ограничение, чтобы достичь определенного уровня потребления, будут чаще обнаруживаться в группах с низким относительным доходом, чем в группах с высоким относительным доходом.

## ОБСУЖДЕНИЕ И ВЫВОДЫ

Помимо демографического объяснения рождаемости ее исследователей часто интересует проверка каких-либо конкретных гипотез. Модель относительного дохода

помогает сконцентрировать внимание на группе взаимосвязанных гипотез и ведет к конкретному указанию ожидаемых результатов.

Прогностические возможности переменной относительного дохода возрастают, когда мы переходим от всей изучаемой подвыборки к тем супружеским парам, которые, как правило, планируют свою дальнейшую жизнь, состоят в браке непродолжительное время, дети которых зачаты только после вступления в брак, а также тогда, когда мы рассматриваем относительный доход в настоящее время вместе с интервалом между вторым и третьим рождением. Мы видели, что в нашей подвыборке прогностические возможности относительного дохода не одинаковы для всех социальных слоев, т. е. положительная связь между относительным доходом и накопленной рождаемостью не сохраняется в группах с низким уровнем образования и в низших профессиональных группах, однако она более определенно выражена в группах с более высоким уровнем образования и более высоким профессиональным статусом.

Эти результаты предполагают, что, проверяя наши гипотезы, мы систематически продвигались к определению прогностических возможностей модели, не выходя из круга экономических понятий. Коротко говоря, результат нашего исследования есть *экономическая* модель, обладающая, вероятно, прогностическими возможностями, но с ограниченной областью применения. Она подходит для ситуации, когда существует мотивация для регулирования рождаемости и доступны соответствующие средства регулирования, но ограничена социальной средой, в которой супружеские пары имеют возможности максимизировать блага (utilities). Иными словами, если группа или группы сознают, что у них нет средств и/или возможностей для максимизации потребления, то можно ожидать, что у них будут отсутствовать и мотивы не иметь детей. На то, что у групп низкого социального статуса (по уровню образования и профессиональному уровню) малые возможности потребления, указывают следующие данные: в группе с низким уровнем образования 79% опрошенных хотели, чтобы их дети получили образование выше среднего, тогда как в группе с высоким уровнем образования этого хотели 100% опрошенных. Эти данные заставляют предполагать, что группы с низким уровнем образования, по крайней мере, с одной точки зрения реалистически оценивают свои возможности.

Следуя в своих рассуждениях дальше, мы полагаем, что

модель относительного дохода неприменима для общества или подгруппы, где большие семьи считаются обеспечивающими максимизацию экономических или неэкономических благ. Модель полезна, по-видимому, для оценки репродуктивного поведения только избранных групп населения. Наши результаты, таким образом, предполагают, что высокие прогностические возможности модели зависят не от одной лишь эффективной практики контрацепции; необходимо также, чтобы все социальные слои могли максимизировать блага.

## ЛИТЕРАТУРА

1. Balakrishnan T. R., Kantner J. F. and Allingham J. D. Fertility and Family Planning in a Canadian Metropolis. Montreal, 1975.
2. Bernhardt E. M. Fertility and Economic Status. Some Recent Findings on Differentials in Sweden. *Population Studies*, 1972, 26, 175—184.
3. Banks James A. Prosperity and Parenthood. London, 1954.
4. Chaudhury Rafiqul H. A Multivariate Analysis of Relative Income's Relationship to Fertility. Unpublished thesis. Baltimore, School of Hygiene and Public Health, The Johns Hopkins University, 1972.
5. Easterlin Richard A. Toward a Socio-Economic Theory of Fertility. A Survey of Recent Research on Economic Factors in American Fertility. In: Samuel J. Behrman, L. Corsa Jr. and R. Freedman (eds.). Fertility and Family Planning. A World View. Ann Arbor. 1969, p. 127—156.
6. Easterlin Richard A. Relative Economic Status and the American Fertility Swing. Unpublished working paper. Philadelphia. University of Pennsylvania, 1972.
7. Edin Karl A. and Hutchinson E. P. Studies of Differential Fertility in Sweden. London, 1935.
8. Freedman D. The Relation of Economic Status to Fertility. *American Economic Review*, 1963, 53, 416—426.
9. Griffith Jeanne E. Contemporary U. S. Family Size Norms. Unpublished dissertation. Baltimore, The Johns Hopkins University, 1973.
10. Kunz P. R. The Relation of Income and Fertility. *Journal of Marriage and the Family*, 1965, 27, 509—513.
11. Osteria Trinidad. Child-Spacing and Economic Position of the Family in Toront. Unpublished doctoral dissertation. Baltimore, The Johns Hopkins University, 1972.
12. Whelpton Pascal K. and Patterson J. E. Fertility and Family Planning in the United States. Princeton, 1966.
13. Winer B. J. Statistical Principles in Experimental Design. New York, 1962.

*Джеймс А. Уид*

## **ВЛИЯНИЕ ВОЗРАСТА ПРИ ВСТУПЛЕНИИ В БРАК НА РАЗЛИЧИЕ КОЭФФИЦИЕНТОВ РАЗВОДИМОСТИ ПО ШТАТАМ**

James A. Wee d. Age at Marriage as a Factor in State Divorce Rate Differentials. *Demography*, vol. 11, № 3, August 1974, p. 361—375.

Одна из наиболее устойчивых и поразительных закономерностей, обнаруженных в проведенных до сих пор исследованиях географии разводов в США, заключается в том, что уровень коэффициентов разводимости резко возрастает с востока на запад. В 1960 г. коэффициенты разводимости (на 1000 супружеских пар) по районам США колебались от низкого уровня (3,8) на Северо-Востоке до высокого (15,0) на Западе. Коэффициенты 8,9 и 12,4 преобладали соответственно в Северном Центральном и в Южном районах (см. [30, табл. 3-B], [26]). Различия такого же порядка прослеживались почти в течение века [6], и, с тех пор как Уилкоккс [32] выступил в печати с критикой доклада о разводах, сделанного в 1889 г. тогдашним руководителем Комиссии США по труду Райтом [33], сравнительное влияние на разводы законодательства и социальных факторов стало предметом широкого исследования и обсуждения. Уилкоккс [32, с. 61] пришел к заключению, что различия в разводимости в «трех зонах (Восточной, Центральной и Западной) не объясняются различием в законах» и что налагаемые законом «ограничения на развод оказывают лишь незначительное влияние на коэффициенты разводимости».

Не так давно Киркпатрик [14, с. 595] указал, что он не обнаружил зависимости между значениями коэффициентов

разводимости и числом необходимых для развода оснований по различным штатам. С другой стороны, Мазур [18, с. 283] утверждает, что «расхождения в американских законах о разводе настолько значительны, что они почти целиком определяют различия между коэффициентами разводимости по штатам страны». Однако ни точку зрения Мазура, ни точку зрения Уилкокса нельзя принять безоговорочно. Часть колебаний в значениях коэффициентов разводимости по штатам, особенно в крайних случаях, вероятно, можно приписать влиянию расхождений в законах о разводе в соответствующих штатах. Но все же значительная часть приведенных в литературе доказательств свидетельствует в пользу предположения, что различия в коэффициентах разводимости зависят также от социальных факторов, факторов социальной среды (ecological factors) и от демографических факторов (см. [9], [12], [8], [2]).

Многие авторы предполагают, что, кроме факторов, связанных с законодательством (таких, как место проживания и основания для развода), на коэффициенты разводимости, вычисленные для штатов в целом, влияет структура населения по вероисповеданию, цвету кожи, а также по месту постоянного проживания, происхождению, месту рождения, образованию и уровню дохода. Необычайно резкое изменение уровня коэффициентов разводимости по штатам с востока на запад часто приписывали воздействию сложного фактора социальной среды, который можно назвать «эффектом переселенческой среды» (frontierness)\*. Он охватывает влияние подвижности населения на его социальную целостность. В сравнении с ним демографические факторы, в частности возрастное распределение населения и возраст при вступлении в брак, очень редко упоминались или анализировались при исследовании различий в коэффициентах разводимости в зависимости от социальной среды. Это представляется довольно странным ввиду того, что данные отдельных обследований неизменно показывали, что в США ранние браки, а особенно браки в возрасте до 20 лет (teenage marriages), в силу ряда причин подвергаются относительно большому риску развода [16]. Цель нашей работы — заново проанализировать влияние эффекта переселенческой среды на различие коэффициентов разводимости по территории с учетом упомянутых демографических переменных.

---

\* Имеется в виду влияние обстановки, которая создается на осваиваемых переселенцами территориях. — *Примеч. пер.*

В этой статье мы покажем с помощью линейных статистических методов, что после стандартизации коэффициентов разводимости для отдельных штатов по возрасту наилучшим статистическим предиктором\* уровня разводимости в 1960 и 1970 гг. становится доля вступивших в брак в молодых возрастах. В круг переменных, включенных нами в регрессионный анализ, входят показатели как связанных, так и не связанных с законодательством факторов, которые, как считалось в предшествующих эмпирических и теоретических исследованиях, могут оказать существенное влияние на уровень разводимости. Полученные результаты дают возможность показать, что в предшествующих исследованиях разводимости в зависимости от социальной среды, по крайней мере на уровне штата в целом, влияние социального климата или среды на прекращение брака переоценивалось, вероятно, из-за пренебрежения факторами, относящимися к обстоятельствам самого заключения брака.

## РАЗЛИЧИЯ В РАЗВОДИМОСТИ

### Социальные факторы

Цвет кожи, вероисповедание, место проживания (город, сельская местность) и место рождения — четыре социальных фактора, которые часто принимают в качестве переменных, влияющих на изменение уровня разводимости по территории [6], [17], [20]. Как Уилкоккс в 1897 г. [32], так и Якобсон в 1959 г. [12] выбрали эти четыре переменные для специального изучения, и существуют некоторые данные, доказывающие, что для США в целом эти переменные связаны со значительными различиями в коэффициентах разводимости [8, с. 246]. Есть свидетельство того, что коэффициенты разводимости, вычисленные для совокупности семей, уменьшаются с увеличением дохода [27]. Хотя с повышением уровня образования неустойчивость брака, по-видимому, снижается, такая зависимость заметно меняется, если устранить влияние возраста при вступлении в брак и других факторов [3], [4]. Различия в уровне образования будут рассмотрены более подробно в следующем разделе.

В нескольких исследованиях была установлена корреляционная связь между данными о разводах и показателя-

---

\* См. примеч. ред. на с. 49.

ми социальных, демографических переменных и переменных социальной среды. Кэннон и Джинглс [5] исследовали связи между коэффициентами разводимости и семью социальными факторами при обследовании 34 наиболее урбанизированных графств\* штата Небраска. Они обнаружили статистически значимую парную корреляцию (zero-order correlations) между коэффициентами разводимости в графстве и показателями религиозной принадлежности, возрастного состава (процентом лиц в возрасте от 18 до 39 лет), уровня образования, подвижности населения, доли иностранцев, уровня дохода и урбанизации. Вариация коэффициентов разводимости между графствами на 83% определялась последними тремя переменными.

Броэл-Плэтерис ([2], см. также [25, приложение А]) попытался по данным переписи населения США 1950 г. всесторонне исследовать зависимость разводимости от 19 переменных, не связанных с законодательством, в разрезе штатов и типов поселений. В качестве показателя уровня разводимости для каждого штата и типа поселения он принял «процент разведенных» по данным переписи населения США 1950 г. и обнаружил, что из девятнадцати отобранных переменных в населении крупных городов не коррелировали (на уровне значимости 0,05) с процентом разведенных только две, в то время как в сельском фермерском и сельском нефермерском населении не коррелировали с процентом разведенных десять переменных. Конечно, доля разведенных в этих категориях населения в тот или иной момент времени зависит не только от числа разводов, но и от частоты повторных браков, продолжительности раздельного проживания и направления миграционных потоков. Следовательно, результаты, полученные Броэлом-Плэтерисом, следует рассматривать с соответствующими оговорками.

Фенелон [9] решил подвергнуть эмпирической проверке понятие эффекта переселенческой среды, соотнося по штатам коэффициенты переселения для графств и коэффициенты разводимости в 1960 г. Взяв данные переписи населения США 1960 г. о проценте городского населения, проценте небелых, доходе на душу населения и проценте лиц, имеющих родителей-иностранцев, а также данные о проценте католиков из Католического альманаха за 1961 г. [10], он обнаружил, что только два последних показателя статисти-

---

\* Графство (county) — административно-территориальная единица в США, меньшая, чем штат. — *Примеч. ред.*



чески значимо коррелировали с коэффициентами разводимости по штатам (на 1000 супружеских пар) и что ни один из показателей не привел к существенному ослаблению частной корреляции\* между коэффициентами переселения для графств и коэффициентами разводимости. Фенелон [9, с. 323] исключил из анализа штаты Алабама, Арканзас, Айдахо и Невада для того, чтобы «избежать сравнения с теми штатами, где простота получения развода по закону оказывает очевидное влияние на уровень коэффициентов разводимости».

### Эффект переселенческой среды

Пожалуй, чаще всего для объяснения изменений в коэффициентах разводимости с востока на запад предлагается такой фактор, как эффект переселенческой среды (frontierness). Иногда его называют «атмосферой переселенческой среды» или «духом Запада». В посвященной этому литературе понятие *переселенческая среда* подразумевает относительно недавно заселенный район, в котором «нет установившихся традиций и норм» и где «социальный контроль слабее, чем в давно заселенных районах, в которых люди живут вместе уже на протяжении трех-четырёх поколений» [15, с. 494]. Для характеристики переселенческих традиций Запада, способствующих более высоким коэффициентам разводимости, предлагались такие черты, как большая степень личной свободы [13, с. 118], либеральные отношения [19, с. 635], менее охотное подчинение общим нормам [13, с. 582], отсутствие корней [16, с. 592].

Ключевым элементом понятия эффекта переселенческой среды служит миграция, в частности историческое движение населения на Запад. «Переселенцы, со свойственным им духом предприимчивости, создавали новую обстановку и новые условия жизни, побуждавшие их разрушать старые обычаи и тем самым порывать с прошлым» [6, с. 132]. Благодаря открывшимся возможностям в этих приграничных областях образовалась «разнородная совокупность эмоционально независимых и предприимчивых людей... Пионерский дух способствовал индивидуалистическому характеру

---

\* Т. е. исключение влияния каждого из этих показателей путем вычисления коэффициентов частной корреляции не изменило существенно тесноты связи между показателями переселения и разводимости. — *Примеч. ред.*

поведения этих людей, которым никакие препятствия не мешали проявлять их своенравие» [17, с. 118]. Миграция приводит к тому, что «люди (на Западе) зачастую разрываю́т все узы родства, расставшись с родственниками при передвижении на Запад» [15, с. 494]. Наконец, миграция может привести к созданию атмосферы изоляции, обезличивания и к ослаблению социального контроля вследствие «снижения институционального и традиционного воздействия семьи, влияния друзей, а также меньшей боязни клейма развода... Популярный (у завоевателей Запада) образ человека часто связывается с такими чертами характера, как независимость, уверенность в себе и предприимчивость» [20, с. 229, 231].

Было высказано предположение, что миграция представляет собой фактор, определяющий социальную целостность, которая, в свою очередь, непосредственно определяет степень, в какой данная община способна диктовать нормы поведения. «Новые, недавно сложившиеся общины не обнаруживают той целостности, которая существует в более старых и более устойчивых общинах,... поскольку навязывание каких-либо норм поведения в этих районах сравнительно слабее» [19, с. 326], см. также [2]). Следовательно, чем больше доля новоприбывших в данном районе, тем меньше степень социальной целостности или устойчивости и тем меньше социальная значимость развода. Результаты эмпирического анализа Фенелона, по существу, подтверждают эту теорию, поскольку коэффициент парной корреляции между коэффициентами разводимости по штатам в 1960 г. и уровнем миграции, равный 0,72, не был сколько-нибудь заметно ослаблен ни одной из выбранных им контрольных переменных.

### **Возраст при вступлении в брак**

Уилкоккс писал: «Закономерно, по-видимому, что в общинах, где наиболее распространены ранние браки, бывает меньше всего разводов» [32, с. 66]. Даже если это и было справедливо для США до 1900 г., — а Уилкоккс пользовался очень ограниченными данными, — то нет никаких оснований предполагать, что эта зависимость существует и сегодня. Не так давно Риль в своей статье [23] обратил внимание на относительно высокий медианный возраст женщин при вступлении в первый брак в некоторых штатах Северо-Восточного района и низкий — в большинстве Цент-

ральных Южных и Западных штатов. (Сравнение проводилось по 29 штатам, предоставившим данные о возрасте при вступлении в брак в 1959 г. Национальному управлению статистики естественного движения населения.) Пенг и Хансон обнаружили, что «на Западе самые высокие в стране коэффициенты брачности в возрасте до 20 лет... Ранние браки имеют большую вероятность закончиться разводом в силу таких факторов, как неустойчивость личности, финансовые трудности, короткий срок знакомства, а также беременность» [20, с. 232—233].

Зависимость между возрастом при вступлении в брак и разводимостью изучалась в ряде исследований [16, с. 427—429]. Данные были получены из всеамериканского обследования экономических возможностей (Survey of Economic Opportunity), проведенного Бюро переписи для Управления экономического развития [27]. В табл. 1 приведено процентное распределение лиц, указавших при обследовании, что они разведены, по полу, возрасту при вступлении в первый брак и по продолжительности брака к 1967 г. Как для мужчин, так и для женщин процент разведенных среди вступивших в брак в самом раннем возрасте почти вдвое больше, чем в двух других возрастных группах, и эта закономерность сохраняется (с одним исключением у мужчин) для всех групп по числу лет с момента вступления в первый брак. Из других данных обследования 1967 г. видно, что у женщин, вступивших в брак до 18 лет (возраст учащихся в средней школе), вероятность развода намного больше, независимо от продолжительности брака. Бумпасс и Суит [3] обнаружили, что обратная зависимость между возрастом при вступлении в брак и вероятностью развода сохраняется и продолжает быть тесной также после того, как устранено влияние других переменных, включая образование жены, ее религиозную принадлежность в детские годы, семейное положение в возрасте 14 лет, место проживания в детские годы, брачное состояние во время первой беременности и первого рождения, а также состоял ли первый муж в браке ранее и сколько времени прошло с момента первого замужества.

Следует отметить, что для 32 штатов, предоставивших в 1960 г. данные о возрасте женщины при вступлении в брак в Национальный центр санитарной статистики, коэффициент корреляции между медианным возрастом при вступлении в брак и коэффициентами разводимости в 1960 г. (на 1000 замужних женщин старше 14 лет) составил 0,61 и был

Таблица 1

Процент обследованных (все цвета кожи) в 1967 г.,  
указавших, что они разведены,  
по полу, возрасту при вступлении в первый брак  
и числу лет с момента вступления в первый брак

Возраст при вступлении в первый брак	Число лет с момента вступления в первый брак			
	все годы	до 10 лет	от 10 до 19 лет	20 лет и более
Мужчины				
От 14 до 69 лет	14,9	7,8	14,9	18,5
От 14 до 21 года	20,2	9,4	19,0	27,6
От 22 до 27 лет	11,7	5,5	12,6	13,8
От 28 до 69 лет	10,9	8,5	11,7	11,3
Женщины				
От 14 до 69 лет	16,7	8,9	16,3	20,1
От 14 до 19 лет	22,4	11,4	23,4	27,3
От 20 до 24 лет	11,5	6,0	9,8	14,5
От 25 до 69 лет	11,1	6,4	9,3	13,1

Источник. [27].

значимым на уровне 0,001. Можно предположить, что на этот показатель корреляции повлияли и различия между штатами в определенном законом минимальном возрасте вступления в брак. В любом случае желательно расширить анализ, чтобы охватить все штаты страны. С этой целью мы воспользуемся в качестве показателя возраста при вступлении в брак процентом когда-либо состоявших в браке для конкретных когорт женщин, выделенных по возрасту, по данным переписей населения США 1960 и 1970 гг.

#### ДАННЫЕ ЗА 1960 г.

Сложную проблему представляет собой разработка измерителя, который характеризовал бы степень дозволенности развода законом согласно законодательству того или иного штата. Очевидно, единственная попытка создать такой измеритель была предпринята Броэлом-Плэтерисом [1] в его докторской диссертации. Прежде всего он провел опрос преподавателей, читающих курс семейного

законодательства, и членов Секции семейного законодательства Американской ассоциации адвокатов. Цель этого опроса — получить представление о мнении экспертов, которым, по всей вероятности, был известен подход к закону о разводе в судах разных штатов. Их попросили указать штаты («проголосовать» за них), где законы о разводе были наиболее мягкими, а также те штаты, где оформить развод наиболее трудно, и штаты, где требования к расторгающим брак средние. Опрос был произведен весной 1958 г. На основании 64 возвращенных анкет Броэл-Плэтерис разработал «шкалу дозволенности развода», теоретически ранжированную от 0 до 100. Шкала была разработана путем присвоения каждому штату ранга на основании полученных им «голосов» опрошенных. Штаты с наиболее мягкими законами получили наивысшие ранги. В недавно вышедшей книге Рейнстайна [25, приложение А] дается конспективное изложение диссертации Броэла-Плэтериса, включая таблицу, указывающую ранги, присвоенные закону каждого штата.

Насколько надежна и обоснована шкала законов Броэла-Плэтериса, по существу, неизвестно, и поэтому анализ данных 1960 г. произведен нами в два этапа. На первом этапе вычислялась регрессия коэффициентов разводимости по показателям всех выбранных факторов, не связанных с законодательством; на втором этапе — регрессия коэффициентов разводимости как по этим факторам, так и по шкале законов Броэла-Плэтериса. Этот подход позволяет, во-первых, оценить относительное значение различных не связанных с законодательством факторов, во-вторых, установить, в какой степени включение в анализ шкалы законов повлияло на результаты, полученные на первом этапе анализа.

Такой двухэтапный подход требует особого внимания к выбору перечня штатов для регрессионного анализа. Поскольку задача первого этапа заключается в анализе воздействия на коэффициенты разводимости факторов, не связанных с законодательством, целесообразно исключить два штата — Неваду и Алабаму. По данным Плэтериса [21, с. 512], коэффициенты разводимости 1960 г., вычисленные для постоянного населения и для мигрантов, составили в Неваде соответственно 3,4 и 26,2 (на 1000 населения), а в Алабаме — 2,7 и 33,1. В обоих штатах коэффициент разводимости для мигрантов был настолько велик, что нельзя было ожидать, чтобы в средних коэффициентах для шта-

та нашли правильное отражение социальные характеристики постоянного населения. Значительно меньше была разница между соответствующими коэффициентами для постоянного населения и мигрантов в штатах Айдахо, Флорида и Арканзас, поэтому ни один из них не был исключен из анализа. Таким образом, на первом этапе в исследование вошли 48 штатов и округ Колумбия.

Для того чтобы достичь сопоставимости двух этапов путем включения в анализ на обоих этапах одних и тех же штатов, необходимо было присвоить ранги законам в штатах Аляска, Гавайи и Южная Каролина. Броэл-Плэтернс присвоил законам о разводе в Южной Каролине ранг 1, так как по существовавшим до 1949 г. законам этого штата брак мог быть расторгнут только путем его аннулирования, а собственно развод законом был запрещен [12]. Однако он указывает, что для 1958 г. ранг, присвоенный Южной Каролине, должен был быть больше, чем ранги штатов Нью-Йорк или Массачусетс, но явно меньше, чем ранг штата Нью-Джерси. Исходя из этого Южной Каролине был присвоен ранг 11. В 1950 г. ни Аляска, ни Гавайи не были включены в число штатов, имеющих законодательство о разводе, и поэтому не получили ранги на шкале законов. Для того чтобы включить эти штаты в состав данных 1960 г., мы присвоили Аляске ранг 70, а Гавайям — 50. Присвоенный Аляске ранг находится в верхней части группы штатов, допускающих развод, посередине между рангами штатов Нью-Мексико и Оклахома. В 1960 г. это были единственные три штата, где достаточным основанием для развода считалась «несовместимость». Ранг, присвоенный Гавайям, более произволен и отражает наше представление о том, что законы о разводе в этом штате «близки к средним», несмотря на то что необходимый для развода срок постоянного проживания в этом штате составляет 2 года. Для того чтобы проверить, как влияют эти выбранные нами значения, регрессионный анализ по данным 1960 г., приводимый далее, был также проведен без учета Аляски и Гавайев. Сколько-нибудь существенного изменения результатов не произошло. Итак, все приведенные далее результаты, полученные для 1960 г., были основаны на данных по 49 судебным ведомствам, относящихся к 48 штатам и округу Колумбия; исключены были только штаты Невада и Алабама.

В предыдущем разделе мы выделили некоторые факторы, которые, как предполагается, могли оказать влияние на коэффициенты разводимости в штатах. Показатели фак-

торов, не связанных с законодательством, для их анализа в данном исследовании 1960 г. были взяты, за одним исключением, из различных таблиц первого тома данных переписи населения США 1960 г. [24].

Ввиду особенно высокой вероятности развода для женщин, вступивших в брак в возрасте до 20 лет, мы сочли целесообразным выбрать в качестве показателя уровня ранних браков в каждом штате процент когда-либо состоявших в браке среди женщин от 15 до 19 лет. Были также взяты данные о проценте когда-либо состоявших в браке среди женщин в возрасте 20—24 лет и среди женщин от 15 до 24 лет.

Мы выбрали три переменные, которые могли бы послужить в качестве самостоятельных показателей различных явлений, связанных с понятием переселенческой среды. Это: (1) процент лиц, родившихся за пределами штата, т. е. процент жителей штата в 1960 г., проживавших не в том штате, где они родились; (2) процент переселившихся в штат, т. е. процент лиц в возрасте 5 лет и старше в 1960 г., которые в 1955 г. проживали в другом штате; (3) процент переселившихся в графство, т. е. процент жителей штата в возрасте 5 лет и старше в 1960 г., которые в 1955 г. проживали в другом графстве. Третья переменная — это тот же показатель, который применил Фенелон [9], и она аналогична показателю, взятому Броэлом-Плэтерисом [1]. Переселившиеся в графство включают как переселившихся в данный штат, так и лиц, прибывших из другого графства в том же штате. Большое число приезжих в определенной степени влияет на социальную целостность в общине, независимо от происхождения новых членов общины и от отдаленности районов, откуда они прибыли. В той же мере и процент переселившихся в графство может служить более или менее подходящим показателем уровня миграции. Однако традиционное понятие переселенческой среды, которым пользуются Карсон [6], Лихтенбергер [17] и Лэндис [15], должно лучше отображаться такими показателями, как «процент родившихся за пределами штата» и «процент переселившихся в штат». Соответственно результаты будут приведены для всех трех показателей.

В анализ включены шесть контрольных переменных. Эти переменные для каждого штата следующие: (1) процент негритянского населения; (2) процент первых двух поколений иммигрантов (foreign stock); (3) процент живущих в городах с населением 50 000 и больше; (4) медианное число

лет обучения для лиц в возрасте 25 лет и старше; (5) медианный доход семьи; (6) процент католиков. Шестой социальный показатель был взят из Католического альманаха [10].

Для того чтобы снизить влияние различий между штатами в возрастном распределении населения, для каждого

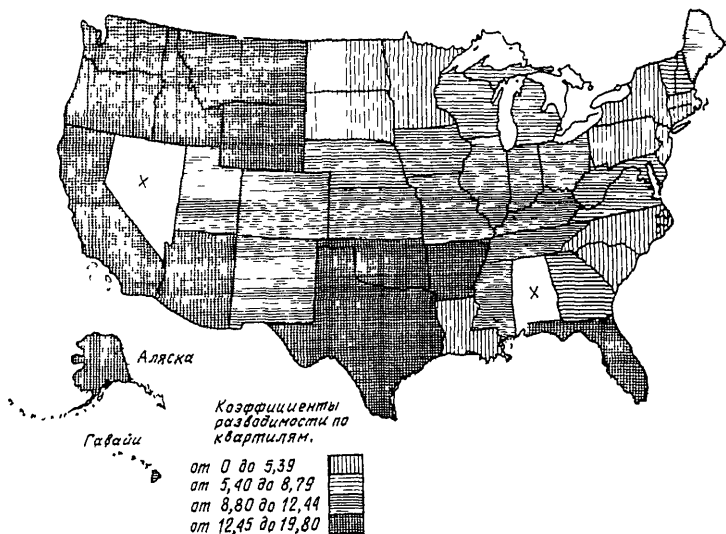


Рис. 1. Точечная карта, иллюстрирующая стандартизованные коэффициенты разводимости 1960 г. (на 1000 замужних женщин) в кватрилях.

Примечание. Невада и Алабама (x) были исключены из анализа из-за очень высоких коэффициентов разводимости для мигрантов в обоих штатах.

Источники: [26], [29, табл. 3-4], [30, табл. 3-В]

штата путем косвенной стандартизации был вычислен коэффициент разводимости (на 1000 замужних женщин), стандартизованный по возрасту. За стандарт были приняты коэффициенты разводимости 1960 г. (на одну замужнюю женщину) по возрасту на момент развода, полученные путем объединения данных для четырех штатов: Гавайи, Айова, Теннесси и Висконсин. Данные о возрасте на момент развода, предоставляемые этими штатами Национальному центру санитарной статистики США [30, табл. 3-4], достаточно полны: они охватывают 97,4 % случаев развода в этих штатах. Распределение замужних женщин по возраст-



ту для каждого штата было взято из данных переписи населения США 1960 г. Для того чтобы уменьшить влияние случайных колебаний, число разводов при вычислении стандартизованного коэффициента разводимости для каждого штата было получено как средняя арифметическая из общего числа разводов и случаев аннулирования брака в 1959, 1960 и 1961 гг. [30, табл. 3-В].

Стандартизованные коэффициенты разводимости для 1960 г. представлены на рис. 1. Схема составлена таким образом, что штаты, попавшие в первый квартиль распределения, имеют наименьшую плотность точек, а в четвертый квартиль — наибольшую плотность. Изменение уровня разводимости с севера на юг и с востока на запад остается ясно выраженным даже после стандартизации по возрасту и брачному состоянию.

### РЕЗУЛЬТАТЫ АНАЛИЗА ДАННЫХ ЗА 1960 г.

В этом исследовании для того, чтобы соизмерить относительное влияние различных факторов на коэффициенты разводимости для штатов, мы воспользовались стандартизованными частными коэффициентами регрессии, т. е. бета-коэффициентами\*.

Регрессия вычислялась методом наименьших квадратов (нестандартизованные коэффициенты регрессии можно вычислить по данным, приведенным в таблицах).

Результаты, полученные на первом этапе анализа данных за 1960 г., приведены в табл. 2. В первых трех столбцах чисел показаны стандартизованные коэффициенты регрессии (бета-коэффициенты) отдельно для каждого из трех уравнений множественной регрессии. Различаются эти уравнения лишь тем, какая взята переменная миграции.

---

\* Бета-коэффициенты (beta-weights) — это коэффициенты в уравнении множественной регрессии, выраженные в долях среднего квадратичного отклонения для каждого фактора или независимой переменной. Бета-коэффициенты показывают, на сколько средних квадратичных отклонений должна измениться вариация каждого фактора для того, чтобы значение зависимой переменной в уравнении регрессии (в данном случае коэффициента разводимости) изменилось на 1 среднее квадратичное отклонение. В этом смысле значение зависимой переменной как бы «предсказывается» (predicted) по значениям включенных в уравнение множественной регрессии независимых переменных, которые поэтому в англо-американской статистической литературе называются предикторами (predictors). — *Примеч. ред.*

Итоговая таблица множественного регрессионного анализа.  
Первый этап: 1960 г.

Переменные	Бета-коэффициенты			Средние	Средние квадратичные отклонения
	модель 1	модель 2	модель 3		
Процент женщин 15—19 лет, когда-либо состоявших в браке	0,810*	0,734*	0,642*	16,54	4,13
Процент женщин 20—24 лет, когда-либо состоявших в браке	0,246	0,268	0,184	73,12	6,16
Процент переселившихся в графство	-0,010	—	—	19,30	7,05
Процент переселившихся в штат	—	0,053	—	11,31	6,92
Процент родившихся за пределами штата	—	—	0,290*	29,41	14,41
Процент первых двух поколений иммигрантов	0,446*	0,416	0,346	16,87	11,39
Процент католиков	-0,254	-0,254	-0,247	20,68	14,63
Процент негритянского населения	-0,263	-0,274	-0,354*	9,40	11,98
Процент городского населения	0,307*	0,334*	0,319*	27,94	18,46
Медианное число лет обучения	0,170	0,149	0,022	10,63	1,09
Медианный доход семьи	-0,187	-0,212	-0,277	5 361	985
Коэффициенты разводимости (1960 г.)	—	—	—	9,16	4,05
Коэффициент детерминации	0,755	0,756	0,779	—	—

\* Статистически значимо на уровне 0,05 ( $N=49$ ).

Источники. [26], [29, табл. 4-3], [30, табл. 3-B], [10].

Так, в первой модели в качестве единственной переменной миграции взят процент переселившихся в графство и бета-коэффициенты помещены в первом столбце. Во второй модели взят процент переселившихся в штат, в третьей — процент родившихся за пределами штата. В четвертом и пятом столбцах табл. 2 показаны соответственно средние величины и средние квадратичные отклонения самих переменных. В табл. 3 приведена матрица коэффициентов парной корреляции. Коэффициенты для 1960 г. расположены в этой таблице ниже диагонали.

Из трех моделей в табл. 2 мы видим, что наибольшие значения бета-коэффициентов характерны для процента женщин в возрасте 15—19 лет, когда-либо состоявших в браке, что свидетельствует об относительном значении этой переменной как предиктора коэффициента разводимости для штата, независимо от того, какая переменная миграции была взята. Действительно, переменные миграции имеют наименьшие бета-коэффициенты как в первой, так и во второй модели; лишь переменная «процент родившихся за пределами штата» имеет небольшой (хотя статистически значимый на уровне 0,05) коэффициент регрессии. Стоит также отметить, что в табл. 2 значения бета-коэффициентов для процента женщин в возрасте 20—24 лет, когда-либо состоявших в браке, неизменно остаются намного меньшими, чем для женщин 15—19 лет, когда-либо состоявших в браке. Это согласуется с отмеченным ранее явлением, что для женщин, вступивших в брак до 20 лет вероятность развода почти вдвое выше, чем для женщин, возраст которых при вступлении в брак 20—24 года (см. табл. 1). Бета-коэффициенты (для шести контрольных переменных) довольно невыразительны, за исключением, пожалуй, лишь переменной «процент первых двух поколений иммигрантов», у которой оказался неожиданно большой положительный бета-коэффициент. Это могло быть в значительной степени результатом тесной корреляции (0,83) между этой переменной и другой независимой переменной — «процент католиков», у которой бета-коэффициент отрицателен.

Для того чтобы выявить возможный эффект коллинеарности между двумя переменными — долей женщин 15—19 лет и долей женщин 20—24 лет, когда-либо состоявших в браке, — показатели для этих двух возрастных подгрупп были заменены переменной «процент женщин в возрасте 15—24 лет, когда-либо состоявших в браке» и расчеты по всем трем моделям в табл. 2 были повторены с учетом этой новой переменной. Полученные коэффициенты (здесь они не приводятся) не изменили сколько-нибудь существенно результаты, представленные в табл. 2: у переменной «доля женщин в возрасте 15—24 лет, когда-либо состоявших в браке» были самые большие бета-коэффициенты во всех трех моделях, в то время как у переменных «процент переселившихся в графство» и «процент переселившихся в штат» бета-коэффициенты остались очень малыми.

В табл. 4 показаны результаты второго этапа исследования, на котором в дополнение к независимым переменным

Матрица коэффициентов парной корреляции  
(данные 1960 г расположены под диагональю, данные 1970 г — над диагональю)

Переменные	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Коэффициент разводимости	—	0,57	0,55	0,60	0,53	0,47	0,56	-0,31	-0,45	-0,05	-0,05	0,13	-0,19
2. Процент женщин 15—19 лет, когда либо состоявших в браке	0,71	—	0,74	0,94	0,36	0,31	0,09	-0,78	-0,73	0,24	-0,29	-0,56	-0,47
3. Процент женщин 20—24 лет, когда либо состоявших в браке	0,70	0,78	—	0,88	0,56	0,39	0,21	-0,56	-0,57	-0,32	-0,71	-0,24	-0,42
4. Процент женщин 15—24 лет, когда либо состоявших в браке	0,72	0,91	0,89	—	0,53	0,45	0,22	-0,66	-0,67	-0,01	-0,13	-0,41	-0,35
5. Процент переселившихся в графство	0,63	0,58	0,63	0,74	—	0,86	0,75	-0,10	-0,21	-0,32	-0,27	0,37	0,15
6. Процент переселившихся в штат	0,53	0,48	0,45	0,63	0,92	—	0,80	-0,01	-0,09	-0,14	-0,19	0,33	0,28
7. Процент родившихся за пределами штата	0,60	0,37	0,41	0,54	0,81	0,83	—	0,11	-0,01	-0,10	0,05	0,55	0,35
8. Процент первых двух поколений иммигрантов	-0,34	-0,79	-0,54	-0,61	-0,17	-0,07	0,02	—	0,83	-0,39	0,26	0,58	0,67
9. Процент католиков	-0,44	-0,66	-0,56	-0,56	-0,24	-0,11	-0,09	0,79	—	-0,31	0,26	0,15	0,56
10. Процент негритянского населения	-0,13	0,30	-0,20	0,05	-0,16	-0,08	-0,11	-0,51	-0,33	—	0,44	-0,40	-0,20
11. Процент городского населения	-0,11	-0,26	-0,59	-0,31	-0,18	-0,07	0,11	0,30	0,34	0,28	—	0,23	0,35
12. Медианное число лет обучения	0,29	-0,13	0,12	0,14	0,46	0,47	0,65	0,44	0,22	-0,49	0,21	—	0,63
13. Медианный доход семьи	-0,06	-0,41	-0,23	-0,13	0,21	0,30	0,45	0,61	0,51	-0,42	0,43	0,71	—
14. Ранг на шкале дозволенности развода по закону	0,76	0,72	0,78	0,75	0,62	0,50	0,48	-0,16	-0,53	-0,11	-0,36	0,19	-0,24

Примечание  $N=49$  случаям в 1960 г,  $N=45$  случаям в 1970 г

Источники Для 1960 г [26], [29, табл 4-3], [30, табл 3-В], [10] для 1970 г [28], [31, табл. 7], [22, табл С], [11] Шкала дозволенности развода взята у Брозла-Плэтериса [1, табл 12]

Итоговая таблица множественного регрессионного анализа.  
Второй этап: 1960 г.

Переменные	Бета-коэффициенты		
	модель 1	модель 2	модель 3
Процент женщин 15—19 лет, когда-либо состоявших в браке	0,704*	0,665*	0,575*
Процент женщин 20—24 лет, когда-либо состоявших в браке	0,182	0,184	0,149
Процент переселившихся в графство	—0,041	—	—
Процент переселившихся в штат	—	0,004	—
Процент родившихся за пределами штата	—	—	0,244
Процент первых двух поколений иммигрантов	0,412	0,394	0,329
Процент негритянского населения	—0,217	—0,216	—0,222
Процент городского населения	—0,208	—0,214	—0,304*
Медианное число лет обучения	0,281	0,294	0,305*
Медианный доход семьи	0,111	0,098	—0,002
Шкала дозволенности развода по закону	—0,100	—0,114	0,209
Коэффициент детерминации	0,243	0,236	0,167
	0,771	0,770	0,786

\* Статистически значимо на уровне 0,05 ( $N=49$ ).

Источники: [26], [29, табл. 3-4], [30, табл. 3-B], [10], [1, табл. 12, с. 149].

первого этапа в анализ была включена шкала законов Брозла-Плэтериса. Стандартизованные коэффициенты регрессии приведены для тех же трех моделей, что и в табл. 2, но дополненных шкалой законов. В целом второй этап показал очень небольшое изменение результатов по сравнению с первым. Почти все бета-коэффициенты в табл. 4 несколько меньше по абсолютной величине, чем соответствующие коэффициенты в табл. 2, и почти все знаки остались прежними; таким образом, относительные величины бета-коэффициентов почти не изменились. Так, у переменной «процент женщин 15—19 лет, когда-либо состоявших в браке» во всех моделях сохраняются наибольшие бета-коэффициенты, а у переменных «процент переселившихся в графство» и «процент переселившихся в штат» — наименьшие бета-коэффициенты в соответствующих моделях. В табл. 4 пере-

менная «процент родившихся за пределами штата» имеет умеренный бета-коэффициент, но он статистически незначим. Может показаться странным, что значения бета-коэффициентов для ранга на шкале дозволенности развода относительно малы во всех трех моделях, несмотря на то что у коэффициентов разводимости более тесная парная корреляция с рангом на шкале законов, чем с любой другой независимой переменной (см. табл. 3). Более того, включение ранга на шкале законов увеличило коэффициент детерминации ( $R^2$ ) в каждой модели очень незначительно, не более чем на 2%. Из всего этого следует, что, хотя шкала Броэла-Плэтериса и измеряет связанный с законодательством фактор, результаты первого этапа не были серьезно смещены из-за того, что в них не было устранено влияние степени дозволенности развода по закону. Это заключение согласуется с теми выводами, к которым пришел ранее Уилкоккс.

#### ПОВТОРНОЕ ИССЛЕДОВАНИЕ ПО ДАННЫМ 1970 г.

Пытаясь удостовериться в том, что результаты анализа по данным 1960 г. были получены не под влиянием случайных факторов, действовавших лишь в этом переписном году, мы повторили исследование, пользуясь данными 1970 г. Однако в анализ были внесены некоторые изменения. Во-первых, прежней шкалой законов Броэла-Плэтериса нельзя было пользоваться при данных 1970 г., так как во многих штатах и законы о разводах, и процедура бракоразводных процессов за десятилетие с 1960 по 1970 г. значительно изменились. А применение новой шкалы законов, которую предложил Картер [7], потребовало бы проведения нового исследования. Поэтому в повторном исследовании 1970 г. не будет устранено влияние фактора, связанного с законодательством.

Вторая проблема, связанная с повторным исследованием 1970 г., — выбор штатов для анализа. Как и прежде, мы исключили Неваду, где крайне высоки коэффициенты разводимости у мигрантов. Благодаря изменениям в толковании закона о разводе в Алабаме после 1960 г. общее число разводов у мигрантов в этом штате, по всей вероятности, снизилось [21]. Соответственно Алабама включена в исследование 1970 г. И наконец, последнее ограничение, налагаемое на перечень штатов, выбранных для анализа, вытекает из того, что последним годом, за который можно было

получить из публикаций Национального центра санитарной статистики США окончательные данные о числе разводов для всех штатов, был 1969 г. Было желательно включить в анализ штаты, для которых можно легко получить среднюю арифметическую из общих чисел разводов за 1969, 1970 и 1971 гг.; вследствие этого Аризона, Индиана, Луизиана, Массачусетс и Вермонт вышли из исследования 1970 г. Итак, повторное исследование 1970 г. основывалось на данных для 44 штатов и округа Колумбия; исключены были Невада и пять штатов, по которым не было данных.

Коэффициенты разводимости для замужних женщин, стандартизованные по возрасту на момент развода, были снова вычислены косвенным методом. В качестве стандарта в исследовании 1970 г. были взяты коэффициенты разводимости 1969 г. (на 1000 замужних женщин) по возрасту на момент развода, вычисленные по данным для штатов, входящих в зону регистрации разводов (Divorce Registration Area)\* [22, табл. С]. Числа разводов, принятые при вычислении стандартизованного коэффициента для каждого штата, были получены как средние из общего числа разводов и случаев аннулирования брака в данном штате за 1969, 1970 и 1971 гг. [31, табл. 7]. Распределение замужних женщин по возрасту для каждого штата было взято из таблиц первого тома переписи населения США 1970 г. [28], так же как и значения других переменных, принятых для исследования в 1960 г. Процент католиков был получен из Католического альманаха за 1970 г. [11].

В табл. 5 показаны результаты регрессионного анализа по данным 1970 г. для каждой из трех описанных ранее моделей (см. табл. 2). В целом бета-коэффициенты, полученные на основе данных 1970 г., очень близки по величине к коэффициентам, полученным по данным 1960 г., совпадают и знаки соответствующих коэффициентов. Переменная «процент женщин 15—19 лет, когда-либо состоявших в браке» имеет для 1970 г. во всех моделях несколько больший бета-коэффициент, чем для 1960 г. Из показателей миграции изменилась лишь переменная «процент родившихся за пределами штата», причем ее бета-коэффициент стал даже несколько меньше. Таким образом, в повторном исследо-

---

\* Зона регистрации разводов в США включает штаты, которые представляют статистические сведения о разводах. В 1972—1974 гг. она охватывала 29 штатов. — *Примеч. ред.*

## Итоговая таблица множественного регрессионного анализа: 1970 г.

Переменные	Бета-коэффициенты			Средние	Средние квадратичные отклонения
	модель 1	модель 2	модель 3		
Процент женщин 15—19 лет, когда-либо состоявших в браке	0,944*	0,797*	0,780*	12,28	3,90
Процент женщин 20—24 лет, когда-либо состоявших в браке	0,243	0,219	0,049	66,02	7,15
Процент переселившихся в графство	-0,097	—	—	18,70	5,53
Процент переселившихся в штат	—	0,092	—	10,82	5,09
Процент родившихся за пределами штата	—	—	0,332*	28,81	12,46
Процент первых двух поколений иммигрантов	0,497*	0,459*	0,404*	13,79	9,29
Процент католиков	-0,243	-0,248	-0,241	19,38	13,89
Процент негритянского населения	-0,012	-0,038	-0,150	10,25	12,95
Процент городского населения	0,264	0,314	0,270	27,92	17,48
Медианное число лет обучения	0,750*	0,619*	0,426*	11,83	0,68
Медианный доход семьи	-0,397*	-0,427*	-0,451*	9 131	1 493
Коэффициенты разводимости (1970 г.)	—	—	—	14,82	5,00
Коэффициент детерминации	0,734	0,734	0,776	—	—

\* Статистически значимо на уровне 0,05 ( $N=45$ ).

Источники. [28], [31, табл. 7], [22, табл. С], [11].

вании для 1970 г. сохранились все соотношения, обнаруженные при анализе для 1960 г.

Однако между соответствующими бета-коэффициентами в табл. 2 и 5 наблюдается одно заметное расхождение. В 1970 г. коэффициенты для переменной «медианное число лет обучения» оказались значительно больше соответствующих коэффициентов для 1960 г., особенно в моделях 1 и 2, и значение бета-коэффициента для уровня образования



приблизилось к его значению для процента женщин 15—19 лет, когда-либо состоявших в браке. Почему это произошло, можно объяснить, приняв во внимание два обстоятельства. Во-первых, в 1970 г. распределение штатов по медианному числу лет обучения было отчетливо бимодально: моды пришлись примерно на 10,6 и 12,1 года. Южные штаты (Арканзас, Джорджия, Кентукки, Луизиана, Миссисипи, Северная Каролина, Южная Каролина, Теннесси и Западная Вирджиния) сгруппировались вокруг меньшей моды, между 9,9 и 10,8, в то время как у неюжных штатов значение соответствующего показателя было 12,0 и более. Миссури, Род-Айленд, Вирджиния и Техас — «промежуточные» штаты с показателем между 11,5 и 11,8. Для сравнения: в 1960 г. лишь у нескольких штатов медианное число лет обучения превышало 12,0, а распределение всех штатов даже не напоминало бимодальное, как в 1970 г. Таким образом, для переменной уровня образования в 1970 г. характерно гораздо большее различие между южными и неюжными штатами, чем это было в 1960 г. Во-вторых, мы заметили, что когда была вычислена регрессия коэффициентов разводимости по проценту женщин 15—19 лет, когда-либо состоявших в браке, то почти для всех южных штатов остаточная дисперсия оказалась отрицательной. В 1960 г. таких случаев было не так много. Следовательно, в 1970 г. переменная уровня образования объясняет большую часть вариации коэффициента разводимости и поэтому имеет и большие бета-коэффициенты, чем в 1960 г. Более того, с аналитической точки зрения бета-коэффициент для переменной уровня образования в 1970 г. и должен был быть положительным, потому что ряд штатов с низким уровнем образования после того, как было исключено влияние переменной «процент женщин 15—19 лет, когда-либо состоявших в браке», имеет и низкие коэффициенты разводимости.

Положительный бета-коэффициент для переменной уровня образования — не неожиданность. Кэннон и Джинглс [5] обнаружили существование весьма значимой положительной корреляции между уровнем образования и коэффициентом разводимости для графств штата Небраска. Броэл-Плэтерис [2] получил тесную корреляцию между процентом разведенных (в 1950 г.) и медианным числом лет обучения на уровне штатов. Может показаться, что эти результаты противоречат наблюдениям Бумпасса и Суита [3], [4], которые установили наличие обратной зависимости между неустойчивостью брака и уровнем образования же-

ны по данным Национального обследования рождаемости 1970 г. Однако, когда эти авторы устранили влияние других переменных, в частности возраста при вступлении в брак, то для белых женщин дифференцирующее влияние уровня образования стало сказываться значительно слабее. Так, различия в коэффициентах неустойчивости брака для групп женщин с числом лет обучения 9—11, 12, 13—15, 16 и более лет почти исчезли. Что касается негритянок, то Бумпасс и Суит [4] обнаружили, что в 1970 г. для женщин моложе 45 лет показатели неустойчивости брака были низкими для женщин с уровнем образования 12, 16 и более лет; довольно высокими они были для учившихся 9—11 лет, а для учившихся 13—15 лет были еще выше. Возможно, таким образом, что переменная «медианное число лет обучения» имеет в настоящем исследовании положительный бета-коэффициент по той причине, что южные и неюжные штаты имеют разное распределение замужних женщин по уровню образования и по цвету кожи. В любом случае следует заметить, что Бумпасс и Суит считали неустойчивостью брака те случаи, когда первый брак закончился раздельным жительство супругов (separation) или разводом. Броэл-Плэтерис [2] обнаружил, что между процентом разведенных и медианным числом лет обучения существует положительная корреляционная связь, но между процентом живущих раздельно и числом лет обучения эта связь оказалась отрицательной. Так как в настоящем исследовании рассматриваются только коэффициенты разводимости, трудно сказать, как обнаруженные Бумпассом и Суитом различия (относящиеся к случаю, когда развод и раздельное проживание были объединены) отразились при нашем анализе регрессии на бета-коэффициентах переменной уровня образования. Наконец, отметим, что Бумпасс и Суит нашли, что при устранении влияния таких переменных, как образование, возраст при вступлении в брак и других, неустойчивость брака в регионе «крайний Юг» оказывается ниже, чем в регионе «остальные штаты США».

Таким образом, в настоящем исследовании переменная «медианное число лет обучения» дифференцирует южные и неюжные штаты, по-видимому, лучше любой другой контрольной переменной, и вполне возможно, что сама она заменяет какую-то необнаруженную независимую переменную, которая более непосредственно влияет на различия в коэффициентах разводимости между южными и неюжными штатами.

## ЗАКЛЮЧЕНИЕ

Результаты регрессионного анализа, подытоженные в табл. 2 и 5, указывают, что из всех рассмотренных переменных процент женщин в возрасте 15—19 лет, когда-либо состоявших в браке, оказался самым важным предиктором коэффициента разводимости в штатах как в 1960, так и в 1970 г. Тот факт, что процент когда-либо состоявших в браке для младшей возрастной группы (15—19 лет) лучше «предсказывает» коэффициент разводимости, чем этот процент для старшей группы (20—24 года), подтверждает, по-видимому, выводы других исследований о подверженности ранних браков относительно большому риску развода.

Мы также обнаружили, что и в 1960, и в 1970 г. среди трех принятых в исследовании переменных миграции лишь у одной — «процент родившихся за пределами штата» — стандартизованный коэффициент частной регрессии оказался статистически значимым, хотя он и был более чем вдвое меньше коэффициента для переменной «процент женщин 15—19 лет, когда-либо состоявших в браке». Переменные «процент переселившихся в графство» и «процент переселившихся в штат» имели наименьшие бета-коэффициенты в 1960 г. и почти самые малые в 1970 г. Таким образом, корреляционная связь между миграцией и коэффициентами разводимости ослабевает, если устранить влияние других соответствующих переменных. Это указывает на необходимость пересмотреть традиционную гипотезу о влиянии переселенческой среды на частоту разводов.

Итак, что же можно сказать в свете полученных результатов о причинных зависимостях между этими тремя переменными: разводимостью, миграцией и возрастом при вступлении в брак? Сомнительно, по-видимому, что долю ранних браков можно с достаточным основанием считать переменной, «объясняющей» зависимость между переселенческой средой и разводимостью, поскольку причинно она предшествует и тому и другому явлению. А для того, чтобы доказать в данном контексте, что сама эта зависимость обманчива, понадобилось бы подтвердить, что высокий уровень ранних браков вызывает высокие коэффициенты миграции. Однако, если бы даже все браки и разводы в 1960 г. имели результатом какую-либо подвижность, то среди переживших место жительства в 1959 и 1960 гг. на долю недавно разведенных или недавно вступивших в брак пар, образовавших

новые домохозяйства, приходилось бы не более 10%. Кроме того, мало вероятно, что ранние браки могут создавать, как это обычно предполагается, так называемую атмосферу переселенческой среды.

Гораздо более правдоподобно предположение, что уровень ранних браков представляет собой как бы промежуточное звено между эффектом переселенческой среды и разводимостью. Так, атмосфера переселенческой среды может сделать более вероятным понижение возраста при вступлении в брак или увеличение доли вступающих в брак в раннем возрасте. Это, в свою очередь, может привести к росту коэффициентов разводимости в силу объективно установленного более высокого риска развода для ранних браков. Если эта причинная модель верна, то следует пересмотреть объяснение того, каким образом миграция одновременно служит как причиной, так и показателем переселенческой среды. Если раньше подчеркивалось воздействие миграции на социальную целостность и на общественную цену развода во вновь осваиваемых переселенцами районах, то теперь более уместно, по-видимому, подчеркнуть воздействие миграции на брачность.

Теоретически миграция может повлиять на уровень ранних браков несколькими путями. При относительно более формальном подходе можно выделить воздействие новых благоприятных условий в районах типа вновь осваиваемых переселенцами территорий, дифференциацию миграции по полу и возрасту, необычное соотношение численностей полов, наконец, «брачный рынок», благоприятствующий ранним бракам женщин. Подходя более абстрактно, надо отметить, что большой наплыв мигрантов в тот или иной район может нарушить социальную целостность этого района, что, в свою очередь, может ослабить действенность норм, направленных против ранних браков. Фактически различные доводы, связывающие миграцию с общественной ценой развода, в равной степени применимы и для объяснения связи между миграцией и общественной ценой ранних браков.

Точно таким образом можно сместить акцент и по отношению к зависимости между процентом первых двух поколений иммигрантов или процентом католиков и коэффициентами разводимости. Хотя в населенных, где доля первых двух поколений иммигрантов и католиков высока, могут существовать строгие нормы, противодействующие разводам, важнее, вероятно, то, что в таких обществах существу-

ют еще более строгие нормы, поощряющие поздний возраст вступления в брак (см., например, [24]). Коэффициенты корреляции в табл. 3, видимо, подтверждают это предположение. Так, например, частный коэффициент корреляции для 1960 г. между процентом католиков и коэффициентами разводимости при устранении влияния процента женщин 15—19 лет, когда-либо состоявших в браке, составляет 0,05. Он указывает на значительное ослабление исходного коэффициента парной корреляции, равного 0,44. Этот результат находит подтверждение в выводах Бумпасса и Суита [3] о том, что исходные различия в неустойчивости брака между католиками и протестантами значительно ослабляются при устранении влияния других факторов, включая возраст при вступлении в брак. Среди переменных, изученных Бумпассом и Суитом, возраст при вступлении в брак был признан преобладающим по силе воздействия на устойчивость брака. Не было бы ничего удивительного, если бы этот фактор оказывал сильное влияние на социальную среду.

Итак, результаты нашего регрессионного анализа указывают на то, что в предшествующих исследованиях разводимости в США в зависимости от социальной среды, во всяком случае за последние десятилетия, воздействие социального климата или социальной устойчивости на прекращение брака было или преувеличено, или неправильно истолковано. Становится очевидным, что в будущих исследованиях наиболее плодотворный путь состоит в изучении факторов, которые могут влиять на разводимость через возраст при вступлении в брак.

## ЛИТЕРАТУРА

1. Broel-Plateris Alexander. Marriage Disruption and Divorce Law. Unpublished Ph. D. dissertation. Chicago, University of Chicago, 1961.
2. Broel-Plateris Alexander. Associations between Marriage Disruption, Permissiveness of Divorce Laws and Selected Social Variables. In: Ernest W. Burgess and Donald J.ogue (eds.). Contributions to Urban Sociology. Chicago, University of Chicago Press, 1964.
3. Bumpass Larry L. and Sweet James A. Differentials in Marital Instability, 1970. *American Sociological Review*, 1972, 37, 754—766.
4. Bumpass Larry L. and Sweet James A. Differentials in Marital Instability of the Black Population, 1970. Unpublished paper presented at the annual meeting of the Population Association of America. New Orleans, 1973.

5. Cannon Kenneth L. and Gingles Ruby. Social Factors Related to Divorce Rates for Urban Counties in Nebraska. *Rural Sociology*, 1956, 21, 34—40.
6. Carson William E. *The Marriage Revolt*. New York, Hearst's International Library Co., 1915.
7. Carter Hugh. Review of Marriage Stability, Divorce and the Law by Max Rheinstein. *Contemporary Sociology*, 1973, 2, 367—368.
8. Carter Hugh and Glick Paul. *Marriage and Divorce. A Social and Economic Study*. Cambridge, Harvard University Press, 1970.
9. Fencelon Bill. State Variations in United States Divorce Rates. *Journal of Marriage and the Family*, 1971, 33, 321—327.
10. Foy Felician A. (ed.). 1961 Catholic Almanac. Paterson, N. J., St. Anthony's Guild, 1961.
11. Foy Felician A. (ed.). 1970 Catholic Almanac. Paterson, N. J., St. Anthony's Guild, 1970.
12. Jacobson Paul H. *American Marriage and Divorce*. New York, Rinehart, 1959.
13. Kephart William M. *The Family, Society and the Individual*. (2nd ed.). Boston, Houghton — Mifflin, 1966.
14. Kirkpatrick Clifford. *The Family as Process and Institution*. (2nd ed.). New York, Ronald Press Co., 1963.
15. Landis Paul H. *Making the Most of Marriage*. New York, Appleton Century — Crofts, Inc., 1955.
16. Leslie Gerald R. *The Family in Social Context*. New York, Oxford University Press, 1967.
17. Lichtenberger J. P. *Divorce*. New York, McGraw-Hill Book Co., 1931.
18. Mazur D. Peter. Correlates of Divorce in the U.S.S.R. *Demography*, 1969, 6, 279—286.
19. Nimkoff Meyer F. *Marriage and the Family*. Cambridge, Riverside Press, 1947.
20. Pang Henry and Hanson Sue Marie. Highest Divorce Rates in Western United States. *Sociology and Social Research*, 1968, 52, 228—236.
21. Plateris Alexander A. The Statistics of Divorce. — In: Jeffrey K. Hadden and Marie L. Borgatta (eds.). *Marriage and the Family*. Ithaca, Ill., F. E. Peacock Publishers, Inc., 1969.
22. Plateris Alexander A. Divorces. Analysis of Changes: United States, 1969. *Vital and Health Statistics*, Series 21, № 22. U. S. National Center for Health Statistics. Washington, D. C., Government Printing Office, 1973.
23. Rele J. R. Trends and Differentials in the American Age at Marriage. *Milbank Memorial Fund Quarterly*, 43 (Part 1), 1965, 219—234.
24. Rosenwaille Ira. Factors Associated with Religious and Civil Marriages. *Demography*, 1972, 9, 129—141.
25. Rheinstein Max. *Marriage Stability, Divorce and the Law*. Chicago, University of Chicago Press, 1972.
26. U. S. Bureau of the Census. U. S. Census of Population, 1960. Vol. 1, Characteristics of the Population. Washington, D. C., Government Printing Office, 1964.

27. U. S. Bureau of the Census. Social and Economic Variations in Marriage, Divorce and Remarriage, 1967. *Current Population Reports*, Series P-20, № 223. Washington, D. C., Government Printing Office, 1971.
28. U. S. Bureau of the Census. U. S. Census of Population, 1970. Vol. 1, Characteristics of the Population. Washington, D. C., Government Printing Office, 1973.
29. U. S. National Center for Health Statistics. Vital Statistics of the United States, 1960. Vol. 3, Marriage and Divorce. Washington, D. C., Government Printing Office, 1964.
30. U. S. National Center for Health Statistics. Vital Statistics of the United States, 1961. Vol. 3, Marriage and Divorce. Washington, D. C., Government Printing Office, 1965.
31. U. S. National Center for Health Statistics. Births, Marriages, Divorces and Deaths for December 1971. *Monthly Vital Statistics Report*. Vol. 20, № 12. Washington, D. C., Government Printing Office, 1972.
32. Willcox Walter F. The Divorce Problem. A Study in Statistics. (2nd ed.). New York, Columbia University Press, 1897.
33. Wright C. D. A Report on Marriage and Divorce in the United States, 1867 to 1886. Washington, D. C., Government Printing Office, 1889.

*Рональд Демос Ли*

**ЕСТЕСТВЕННАЯ РОЖДАЕМОСТЬ,  
ДЕМОГРАФИЧЕСКИЕ ЦИКЛЫ  
И СПЕКТРАЛЬНЫЙ АНАЛИЗ  
РОЖДЕНИЙ И БРАКОВ**

Ronald Demos Lee. Natural Fertility, Population Cycles and the Spectral Analysis of Births and Marriages. *Journal of the American Statistical Association*, vol. 70, № 350, June 1975, p. 295—304.

**1. ВВЕДЕНИЕ**

Существует огромное количество демографических данных в виде продолжительных временных рядов, которые охватывают периоды, простирающиеся от средних веков до настоящего времени. Это собрание было недавно пополнено рядами данных о крещениях, погребениях и браках, извлеченными из приходских метрических книг в разных странах, но главным образом для доиндустриальной Европы. Однако методы, применяемые для анализа этих данных, со времени тщательных исследований Дж. Э. Юла на рубеже двух столетий [37] почти не получили развития. Выводы делаются часто совершенно произвольно, а когда применяются такие методы, как множественная корреляция, мало внимания уделяется особенностям исходных данных. Правильный анализ требует тщательного рассмотрения демографических структур, порождающих эти ряды, а также применения подходящих статистических методов.

В последние несколько лет возрастает интерес к возможности применения к этим данным методов спектрального анализа (см. [7], [11], [24], [25], [29]). В этой статье

© Journal of the American Statistical Association



показывается, как можно анализировать поведение рядов чисел рождений и браков и проверять содержательные гипотезы, сочетая ясные стохастические модели демографических порождающих механизмов, представленных как совместные авторегрессионные процессы, с теорией и практическими приемами спектрального анализа. Мы исследуем, в частности, гипотезу естественной рождаемости и разработаем приемы анализа демографических циклов. Рассмотрим вкратце эти проблемы.

Часто утверждают, что в доиндустриальной Европе брачная рождаемость сохранялась относительно постоянной на «естественном» уровне и супружеские пары не контролировали деторождение. В соответствии с этим утверждением изменения в числах рождений от года к году происходили вследствие изменений в возрасте вступления в брак и в частоте браков, т. е. вследствие изменений в брачности (см., например, [35, с. 488]). Другие исследователи оспаривают эти утверждения [2], [5], [26]. Пользуясь временными рядами чисел рождений и браков, мы проверим четыре гипотезы, касающиеся этих противоположных суждений. Гипотеза *A*, или сильная гипотеза естественной рождаемости, утверждает, что брачная рождаемость постоянна, а вся вариация в рядах рождений и браков возникает из-за колебаний брачности. Гипотеза *B* утверждает противоположное, а именно что брачность постоянна, а вся вариация обусловлена брачной рождаемостью. Гипотеза *C*, слабая гипотеза естественной рождаемости, утверждает, что и брачность и брачная рождаемость с течением времени изменяются, однако изменяются независимо, предположительно под влиянием разных возмущающих сил. В то время как брачность подвергается сознательному контролю, брачная рождаемость может быть, например, подвержена влиянию не зависящего от воли людей уменьшения физиологической плодовитости (*fecundity*). Наконец, в соответствии с гипотезой *D* вариации в брачности и брачной рождаемости тесно коррелируют, как предполагается, под влиянием одних и тех же социальных и экономических изменений. Это должно означать, что супруги регулируют деторождение, хотя возможны и другие объяснения. Пользуясь стохастической моделью возобновления населения (*population renewal*), мы выводим спектральное выражение каждой из этих гипотез и сравниваем их с оценками взаимных спектров (*cross-spectra*) фактических рядов чисел рождений и браков.

Спектральный анализ этой модели дает также возможность истолковать 30-летние циклы в рядах чисел рождений и браков в доиндустриальный период. В демографической литературе были предложены две интерпретации этих циклов: одна основывается на представлении о последовательности поколений, другая — на автокорреляции, придаваемой вариации в числах рождений сохраняющимися колебаниями брачности. Мы проведем более полный анализ как этих, так и других компонент указанных процессов.

## 2. ОСНОВНАЯ МОДЕЛЬ

Рассмотрим ряды чисел рождений ( $B_t$ ) и браков ( $M_t$ ), хотя и подверженные значительной стохастической вариации, но не сильно изменяющие свои средние значения за длительные периоды времени — порядка одного или двух столетий (классический пример таких рядов — данные по деревне Крюле, см. [10, с. 60, 61]). Если все рождения происходят в браке<sup>1</sup>, то два этих ряда будут связаны следующими соотношениями:

$$B_t = \sum F_{x,t} M_{t-x}; \quad (2.1)$$

$$M_t = \sum W_{a,t} B_{t-a}, \quad (2.2)$$

где  $F_{x,t}$  — коэффициент рождаемости для браков продолжительностью  $x$  лет с поправкой на прекращение браков в результате смерти или развода<sup>2</sup>;  $W_{a,t}$  — чистая функция брачности женщин (female net nuptiality function), описывающая возрастное распределение браков женщин всех очередностей брака и очищенная от влияния смертности,

<sup>1</sup> Если добрачные зачатия ведут к бракам, то значения  $F_x$  будут положительными для некоторых  $x$ , меньших 0. По Крюле наблюдалось очень небольшое число добрачных зачатий и внебрачных рождений было меньше 1%. В Швеции в XIX в. незаконнорожденные составляли 5—6% рождений, а на некоторых других территориях их доля была намного выше.

<sup>2</sup> Такая функция иногда применялась для того, чтобы скорректировать ряды чисел рождений, устранив влияние колебаний в параллельных рядах чисел браков (см. [2], [5], [6], [37]). Более подробно она рассматривается в [28]. Эта функция будет, очевидно, зависеть от возрастного распределения женской брачной когорты, к которой она относится, и от коэффициентов смертности и разводимости, а также от брачной рождаемости. См. приложение А.

по отнесенная ко всей когорте родившихся, включая мужчин. Уравнения (2.1) и (2.2) можно преобразовать в уравнение возобновления, выразив их только через рождения. После перестановки членов находим:

$$B_t = \sum B_{t-a} (\sum F_{x,t} W_{a-x, t-x}) = \sum B_{t-a} \Phi_{a,t}. \quad (2.3)$$

Очевидно, что выражение в скобках должно представлять приближенно чистую функцию рождаемости<sup>3</sup> (net maternity function)  $\Phi_{a,t}$ . Обозначим через  $B$ ,  $M$ ,  $F_x$ ,  $W_a$  и  $\Phi_a$  средние значения соответствующих величин за длительный промежуток времени, а через  $F$  и  $W$  — суммы соответственно  $F_x$  и  $W_a$  по  $a$ <sup>4</sup>. Тогда  $F$  будет средним числом рождений на один брак, а  $W$  — средним числом браков женщин, необходимых для одного рождения<sup>5</sup>. Из (2.1) и (2.2) и предположения о стационарности средних следует, что  $B = FM$  и  $M = WB$ . Вместе эти выражения предполагают  $FW = 1$ , что отражает длительную стационарность демографического развития и согласно уравнению (2.3) может считаться эквивалентным тому, что чистый коэффициент воспроизводства населения равен единице<sup>6</sup>. В произвольно взятом году чистая брачная рождаемость и чистая брачность будут отличаться от  $F$  и  $W$ ; эти отклонения во времени можно измерить посредством  $F_t$  (сумма  $F_{x,t}$  по всем  $x$ ) и посредством  $W_t$  (сумма  $W_{a,t}$  по всем  $a$ ). Мы будем интер-

<sup>3</sup> Вычисления показывают, что это довольно хорошее приближение, хотя старшим возрастам был придан слишком большой вес, поскольку одинаковая рождаемость приписана вступившим в брак как в 20, так и в 50 лет.

<sup>4</sup> Хотелось бы интерпретировать эти величины как ожидаемые значения, но это вызвало бы ряд затруднений. Предположение об устойчивости ковариации, принимаемое в следующих разделах, требует, чтобы чистый коэффициент воспроизводства (net reproduction rate), которому соответствует  $\sum \Phi_a$ , был меньше единицы. Но в этом случае собственный коэффициент прироста (intrinsic rate of increase) отрицателен, а ожидаемые значения числа рождений и браков будут равны 0. Этого затруднения можно избежать естественным в данном случае путем, изменив модель так, чтобы равновесие достигалось при положительном значении собственного коэффициента прироста, как это предложено в [24]. Такой прием, не меняя результаты, усложнил бы изложение, за исключением случаев, когда исследуются колебания с очень большим периодом.

<sup>5</sup> В XVIII в. в Западной Европе  $F$  было приблизительно 4 или 5, а  $W$  — примерно от 0,20 до 0,25.  $W$  отражает соотношение численностей полов при рождении, а также различия в вероятностях дожития и вступления в брак мужчин и женщин.

<sup>6</sup> Чистый коэффициент воспроизводства определяется как сумма значений чистой функции рождаемости  $\Phi_a$  по всем возрастам.

претировать эти величины как вариацию в брачной рождаемости и брачности, хотя они, несомненно, отражают также изменения в смертности и, в случае  $F$ , изменения в возрасте вступления в брак, что рассмотрено в приложении А.

Согласно (2.3) мы можем рассматривать  $B_t$  как функцию  $B_{t-a}$ ,  $W_{a-x, t-x}$  и  $F_{x, t}$  для каждого  $a$  и  $x$  при данном  $t$ . Для того чтобы получить для  $B_t$  в (2.3) приближение Тейлора первого порядка, берем сначала для каждого  $a$  и  $x$  производные по  $B_{t-a}$ ,  $F_{x, t}$  и  $W_{a-x, t-x}$ . Эти производные, взятые в точке среднего значения, равны:

$$\partial B_t / \partial B_{t-a} = \sum_x F_x W_{a-x} \approx \Phi_a; \quad (2.4)$$

$$\partial B_t / \partial F_{x, t} = \sum_a W_a B = BW = M; \quad (2.5)$$

$$\partial B_t / \partial W_{a-x, t-x} = BF_x. \quad (2.6)$$

Тогда приближение Тейлора будет:

$$\begin{aligned} B_t \approx & B + \sum_a \Phi_a (B_{t-a} - B) + M \sum_x (F_{x, t} - F_x) + \\ & + B \sum_x F_x \sum_a (W_{a-x, t-x} - W_{a-x}). \end{aligned} \quad (2.7)$$

Заметим, что  $\sum_x (F_{x, t} - F_x) = F_t - F$

и что  $\sum_a (W_{a-x, t-x} - W_{a-x}) = W_{t-x} - W$ .

Уравнение (2.7) можно упростить, выразив его через относительные отклонения от средних значений. Положим,  $b_t = (B_t - B)/B$ ;  $f_x = F_x/F$ ;  $\theta_t = (W_t - W)/W$ ;  $\varepsilon_t = (F_t - F)/F$ . Тогда имеем<sup>7</sup>:

$$b_t = \sum_a c_a b_{t-a} + \sum_x f_x \theta_{t-x} + \varepsilon_t. \quad (2.8)$$

Это выражение разлагает относительную вариацию в числах рождений на три приближенно аддитивные компоненты: прошлую вариацию в числах рождений (т. е. возрастную структуру населения в данное время), прошлую вариацию

<sup>7</sup> Эффект исключения членов второго порядка зависит от амплитуды отклонений и от их распределения по частоте (спектру), причем большие амплитуды и низкие частоты ведут к большим ошибкам. Эффект исключения членов более высоких порядков в детерминистском уравнении возобновления населения был исследован Коулом [3], который нашел, что в большинстве случаев этот эффект незначителен.

цию в брачности (т. е. структуру по продолжительности брака в данное время) и вариацию в брачной рождаемости в данное время.

Эта стохастическая модель процесса возобновления населения отличается от других, например от приведенных в [20], [32], тем, что в ней разграничена вариация, порождаемая брачностью ( $\theta_t$ ), оказывающей систематическое влияние, и брачной рождаемостью ( $\varepsilon_t$ ), оказывающей преходящее влияние. Значение этого для циклического характера рядов рождений будет исследовано в разделе 7.

Отметим, что по соображениям симметрии аналогичное разложение справедливо для относительной вариации в числах браков. Полагая  $c_0 = 1$  и  $c_j = -\Phi_j$ , можно выразить два эти одновременно протекающих процесса следующим образом:

$$\sum c_j b_{t-j} = \sum f_x 0_{t-x} + \varepsilon_t; \quad (2.9)$$

$$\sum c_j m_{t-j} = \sum \omega_a \varepsilon_{t-a} + \theta_t, \quad (2.10)$$

где  $m_t = (M_t - M)/M$  и  $\omega_a = W_a/W$ .<sup>8</sup>

На основании (2.9) и (2.10) мы можем определить источники дисперсии и ковариации в рядах рождений и браков согласно гипотезам, предложенным в разделе 1. В соответствии с гипотезой *A* брачная рождаемость постоянна ( $\sigma_\varepsilon^2 \approx 0$ ), так что вся вариация в числах рождений и браков порождается колебаниями брачности ( $\sigma_\theta^2 > 0$ ). Это «сильная гипотеза естественной рождаемости». Согласно гипотезе *B* брачность постоянна ( $\sigma_\theta^2 \approx 0$ ), а вся вариация порождается брачной рождаемостью ( $\sigma_\varepsilon^2 > 0$ ). Гипотеза *C* утверждает, что брачность и брачная рождаемость изменяются независимо под влиянием совершенно иных факторов ( $\sigma_\theta^2 > 0$ ;  $\sigma_\varepsilon^2 > 0$ ;  $\sigma_{\theta, \varepsilon} \approx 0$ ). Наиболее общая гипотеза *D* приписывает часть вариации взаимоисключающим факторам и в то же время допускает, что другая компонента вариации обязана своим происхождением факторам, воздействующим как на брачность, так и на брачную рождаемость, возможно, с запаздыванием (лагом).

Эту гипотезу можно представить как

$$\theta_t = \theta'_t + \eta_t; \quad (2.11)$$

$$\varepsilon_t = \varepsilon'_t + \sum k_j \eta_{t-j}, \quad (2.12)$$

<sup>8</sup> Следуя предпосылкам и методу, разработанному в [22], этим представлением процесса возобновления можно воспользоваться для прогнозирования одновременно ряда рождений и ряда браков.

где  $\theta'_t$ ,  $\varepsilon'_s$  и  $\eta_r$  для всех  $r, s, t$  не коррелированы. Например, если числа браков и *зачатий* реагируют на общие факторы примерно одинаково, то изменения в рождаемости будут запаздывать по сравнению с изменениями в числах браков приблизительно на 9 месяцев. В этом случае  $k_0 = 0,25$ ,  $k_1 = 0,75$  и  $k_j = 0$ , причем  $\sigma_{\varepsilon}^2 > 0$ ;  $\sigma_0^2 > 0$ ;  $\sigma_{\eta}^2 > 0$ .

До сих пор мы предполагали, что ряды были приближенно стационарны в среднем за длительные периоды. Однако, в более общем случае, средние значения  $F_x$  и  $W_a$  будут приводить к среднему чистому коэффициенту воспроизводства, отличающемуся от единицы, и тогда ряды  $B_t$  и  $M_t$  будут изменяться по экспоненциальному закону. Весь анализ проводится для этого более общего случая<sup>9</sup> с помощью соответствующих преобразований функций и переменных.

<sup>9</sup> Ле Бра показал [20], что с каждым вектором стохастических стационарных возрастных чистых коэффициентов рождаемости, подобных нашему  $\Phi_{a, t}$ , может быть связан единственный коэффициент прироста. Коэн [4] показал это при более строгих ограничениях. Вообще этот коэффициент не будет собственным коэффициентом естественного прироста, соответствующим вектору ожидаемых значений  $\Phi_a$ , но для процессов, представляющих практический интерес, он будет близок к этому коэффициенту и будет равен ему при отсутствии сериальной корреляции (см. [32]). Наиболее неблагоприятный случай возникает, когда чистая рождаемость изменяется циклично с периодом, равным среднему возрасту деторождения; но в этом случае, даже если амплитуда цикла равна 0,5 при чистом коэффициенте воспроизводства, равном 1,0, коэффициент роста будет отклоняться от нуля только на 0,0025 (см. [20, с. 551] или [3, с. 186]). Таким образом, коэффициент роста стохастической популяции приблизительно равен собственному коэффициенту прироста,  $r$ , соответствующему  $\Phi_a$  (где  $r$  определяется выражением  $1 = \exp(-ra) \Phi_a$ ). Отсюда следует, что преобразованная стохастическая функция чистой рождаемости  $\Phi_{a, t}^* = \exp(-ra) \Phi_{a, t}$  будет иметь средний собственный коэффициент, равный 1, и, следовательно, порожденный ею ряд будет приближенно стационарным в среднем за длительный период. Из этого вытекает, что приближенно стационарным в среднем будет и процесс, порожденный  $F_{x, t}^* = \exp(-rx) F_{x, t}$  и  $W_{x, t}^* = \exp(-ra) W_{a, t}$ , в чем можно убедиться, подставляя соответствующие величины в уравнение (2.3). Порожденные ряды чисел рождений и браков будут связаны с первоначальными рядами следующим образом:

$$B_t^* = \exp(-rt) B_t \text{ и } M_t^* = \exp(-rt) M_t.$$

Следовательно, весь анализ, проведенный в этой статье, можно применить к функциям и рядам, преобразованным, как только что указывалось. Мы будем предполагать, что такие преобразования сделаны и опустим обозначение \*.

### 3. ТЕОРЕТИЧЕСКИЕ СПЕКТРАЛЬНЫЕ СОТНОШЕНИЯ

Исследуем теперь стохастическую модель, касающуюся рождений и браков, с точки зрения поведения этих двух рядов. Для этого спектральный анализ будет особенно полезен, так как он разлагает их дисперсию и ковариацию по частоте и дает возможность рассматривать отдельно эффекты колебаний с малым и с большим периодом. (Подход, учитывающий временной аспект изменений, см. в [2] или [23].) Кратко охарактеризуем применяемый метод. (Более подробное его изложение см. в [16], [13] или [27].)

Спектр процесса  $x_t$ , обозначаемый  $g_x(\lambda)$ , дает распределение дисперсии  $x_t$  по частоте  $\lambda$ . На рис. 1—6 частота выражена числом циклов на одно столетие. Частота в 20 циклов на столетие соответствует периоду или продолжительности цикла  $100/20$ , или 5 лет. Если  $x$  — ряд некоррелированных случайных величин, то его спектр будет постоянным на всех частотах. Если  $x$  имеет тенденцию изменяться циклично с частотой  $\lambda_0$ , то  $g_x(\lambda)$  покажет пик на частоте  $\lambda_0$ .

Соотношение между двумя рядами описывается взаимным спектром и выражается через «усиление» (gain), «угол фазы» (phase angle) и «согласованность» (coherence). Усиление,  $G(\lambda)$ , указывает, в какой степени колебания в одном ряде, скажем  $x$ , усиливаются или ослабляются, когда они переходят на другой ряд,  $y$ , на некоторой конкретной частоте. Оно аналогично коэффициенту регрессии  $y$  по  $x$  на данной частоте.

Угол фазы,  $\Phi(y)$ , обозначает на частоте  $\lambda$  степень, в которой колебания в ряде  $y$  отстают или опережают колебания в ряде  $x$ . Мы измеряем его в долях цикла; так, угол фазы 0,1 на частоте в 25 циклов на столетие указывает на то, что на данной частоте  $y$  отстает от  $x$  на одну десятую цикла, или на 0,4 года.

Согласованность,  $C(\lambda)$ , аналогична корреляции между двумя рядами на частоте  $\lambda$ ; квадрат согласованности,  $C^2(\lambda)$ , указывает, какую часть дисперсии  $y$  можно на частоте  $\lambda$  объяснить вариацией  $x$ .

Все эти функции можно вычислить либо по теоретической модели, либо по фактическим значениям временных рядов. Сравнение результатов двух способов вычисления дает возможность проверить адекватность модели.

В этом разделе и в приложении *B* мы получаем спектры  $b$  (колебания чисел рождений) и  $m$  (колебания чисел браков), а также значения квадрата согласованности, квадрата угла и соответствующего им сдвига фазы. Это сделано в соответствии с каждой из четырех гипотез относительно дисперсии и ковариации брачной рождаемости ( $\varepsilon$ ) и брачности ( $\theta$ ). Дисперсии  $\varepsilon$  и  $\theta$  представлены их спектрами,  $g_\varepsilon(\lambda)$  и  $g_\theta(\lambda)$ . Ковариация этих процессов представлена спектром их общей компоненты,  $g_\eta(\lambda)$ , и эффектом структуры запаздывания (лага), связывающего  $\varepsilon$  и  $\eta$ ; он представлен коэффициентами  $k_j$  и их передаточной функцией  $K(\lambda)$  (см. (2.12)).

Остальная часть данной статьи посвящена четырем конкретным примерам значений  $g_\varepsilon(\lambda)$ ,  $g_\theta(\lambda)$ ,  $g_\eta(\lambda)$  и  $K(\lambda)$ , соответствующим четырем гипотезам, предложенным в разделах 1 и 2. Численные значения, положенные в основу каждого примера, частично получены из результатов эмпирического анализа, проводимого в следующих разделах.

*Гипотеза А* утверждает, что брачная рождаемость постоянна; это сильная гипотеза естественной рождаемости. Таким образом,

$$g_\varepsilon(\lambda) = 0; g_\eta(\lambda) = 0 \text{ и } g_\theta(\lambda) = 1.$$

*Гипотеза В* утверждает, что брачность постоянна. Таким образом,

$$g_\theta(\lambda) = 0; g_\eta(\lambda) = 0 \text{ и } g_\varepsilon(\lambda) = 1.$$

*Гипотеза С* утверждает, что брачность и брачная рождаемость изменяются независимо; это слабая гипотеза естественной рождаемости. Таким образом,

$$g_\varepsilon(\lambda) = g_\theta(\lambda) = 1 \text{ и } g_\eta(\lambda) = 0.$$

*Гипотеза D* утверждает, что брачная рождаемость и брачность коррелированы с соответствующим лагом. Таким образом,

$$g_\varepsilon(\lambda) = g_\theta(\lambda) = 1; g_\eta(\lambda) = 0,9; k_0 = 0,67 \text{ и } k_1 = 0,33.$$

Эти конкретные характеристики выбраны для гипотезы *D*, поскольку они хорошо согласуются с данными соответствующих рядов для Швеции. Характеристики  $k_0$  и  $k_1$  предполагают, что  $\varepsilon_t$  отстает от  $\theta_t$  на четыре месяца или что колебания чисел брачных зачатий опережают колебания брачности на пять месяцев. Относительные значения  $g_\eta(\lambda)$ ,  $g_\varepsilon(\lambda)$  и  $g_\theta(\lambda)$  предполагают, что для колебаний с большим периодом коэффициент корреляции  $\varepsilon$  и  $\theta$  равен 0,9; для колебаний с меньшим периодом применение дискретных



длинн времени, равных году, ослабляет корреляцию, что отражается в уменьшении значений  $|K(\lambda)|^2$  на высоких частотах<sup>10</sup>.

Следующая таблица, полученная из анализа, проведенного в приложении В, представляет в сокращенной записи спектральные и взаимно-спектральные выражения согласно каждой из этих гипотез.

Спектры и взаимные спектры  $b$  и  $m$  согласно четырем гипотезам относительно брачности и брачной рождаемости\*

Спектр	А брачная рождаемость постоянна	В брачность постоянна	С брачная рождаемость и брачность независимы	Д брачная рождаемость и брачность коррелируют
$g_b(\lambda)$	$a_\lambda^2 f_\lambda^2$	$a_\lambda^2$	$a_\lambda^2 (f_\lambda^2 + 1)$	$a_\lambda^2 [f_\lambda^2 + 1 + \text{RE}(f_\lambda k_\lambda)]$
$g_m(\lambda)$	$a_\lambda^2$	$a_\lambda^2 w_\lambda^2$	$a_\lambda^2 (w_\lambda^2 + 1)$	$a_\lambda^2 [w_\lambda^2 + 1 + \text{RE}(w_\lambda \bar{k}_\lambda)]$
$g_{bm}(\lambda)$	$a_\lambda^2 f_\lambda$	$a_\lambda^2 w_\lambda$	$a_\lambda^2 (\bar{f}_\lambda + w_\lambda)$	$a_\lambda^2 (\bar{f}_\lambda + w_\lambda + 0,5 \bar{k}_\lambda + 0,5 k_\lambda w_\lambda \bar{f}_\lambda)$

\* Характеристику гипотез А, В, С, Д см. в тексте. В таблице применены следующие обозначения:

$a_\lambda^2 = 1/|C(\lambda)|^2$ ;  $f_\lambda = F(\lambda)$ ;  $f_\lambda^2 = |F(\lambda)|^2$ ;  $w_\lambda = W(\lambda)$ ;  $w_\lambda^2 = |W(\lambda)|^2$ ;  $k_\lambda = K(\lambda)$ ; черточка над символом указывает на сопряженность;  $\text{RE}(\ )$  означает действительную часть аргумента.

Демографические функции  $c$ ,  $F$  и  $W$  известны для многих популяций, и мы воспользовались ими для получения численных значений теоретических спектральных функций. Эти функции, а также эмпирические взаимно-спектральные оценки будут рассмотрены далее.

<sup>10</sup> Для того чтобы получить результаты взаимного спектра, нужно лишь принять на каждой частоте некоторые соотношения  $g_e(\lambda)$ ,  $g_\theta(\lambda)$  и  $g_\eta(\lambda)$ . Эти соотношения будут отчасти зависеть от величины изучаемых популяций, поскольку для малых популяций тот факт, что коэффициенты брачности намного ниже коэффициентов рождаемости, может сделать соотношения более неустойчивыми.

Для того чтобы получить спектры  $b$  и  $m$ , необходимо также указать, какую форму имеют  $g_e(\lambda)$ ,  $g_\theta(\lambda)$  и  $g_\eta(\lambda)$ . В качестве первого приближения можно предположить постоянные спектры (белый шум), хотя хорошо известно, что такие показатели, как урожай, реальная заработная плата и безработица, каждый из которых влияет на коэффициенты рождаемости и брачности, автокоррелированы и квазициклически.

#### 4. ВЛИЯНИЕ ЧИСЕЛ БРАКОВ НА ЧИСЛА РОЖДЕНИИ

Прежде чем воспользоваться полной моделью, рассмотренной в разделе 3, полезно применить спектральный анализ для решения более простой задачи, а именно для оценки непосредственного влияния изменений в рядах браков, переданного функцией  $f$  на параллельные ряды рождений. Значение этой взаимосвязи между рядами можно исследовать, вычисляя квадрат усиления  $f$ , т. е.  $|F(\lambda)|^2$ . Он укажет, в какой степени эта функция (или «фильтр») на каждой частоте усиливает или ослабляет дисперсию в ряде браков, когда она передается на ряд рождений (см. [16], [13], [27]). Таким образом,  $|F(\lambda)|^2 g_m(\lambda)$  — это часть дисперсии в ряде рождений, обусловленная на частоте  $\lambda$  дисперсией в ряде браков.

Начнем с рассмотрения эффекта колебаний в числах браков с очень коротким периодом. Например, можно ли объяснить сезонные колебания чисел рождений сезонными колебаниями чисел браков? В некоторых эмпирических исследованиях (см. [15], [34]) было показано, после сравнения сезонности первых рождений и рождений более высоких очередностей, что такое объяснение неправомерно.

Функция квадрата усиления позволяет подойти к решению иначе. Обозначим через  $F_x^1$  функцию, выражающую чистую\* помесечную частоту *первых* рождений по продолжительности брака, и положим  $f_x^1 = F_x^1 / \Sigma F_j^1$ . Для вычисления функции квадрата усиления мы воспользовались данными о значениях  $f^1$  для современного населения США и обнаружили, что при сезонной частоте, равной одному циклу в год, она была равна 0,007. Таким образом, на колебания чисел первых рождений передается менее одного процента сезонной дисперсии браков и, очевидно, передавалось бы еще меньше, если рассматривать рождения всех очередностей. Ясно поэтому, что сезонные колебания в числах браков, по-видимому, не могут оказать сильного влияния на числа рождений.

Как же обстоит дело с колебаниями, имеющими больший период? На рис. 1 показаны три функции квадрата усиления для частот от 0 до 50 циклов на столетие. Одна из них относится только к первым рождениям и основана на современных данных для Франции (см. [28, с. 230]),

\* В том же смысле, в каком раньше шла речь о «чистой» рождаемости вообще, т. е. с поправкой на смертность. — *Примеч. ред.*

вторая — к рожденьям всех очередностей и также основана на современных данных для Франции (см. [28, с. 209]), третья — к рожденьям всех очередностей, она основана на данных XVIII столетия для французского прихода Крюле (вычислена по таблицам в [10]). Сравнение двух функций по современным данным для Франции показывает,

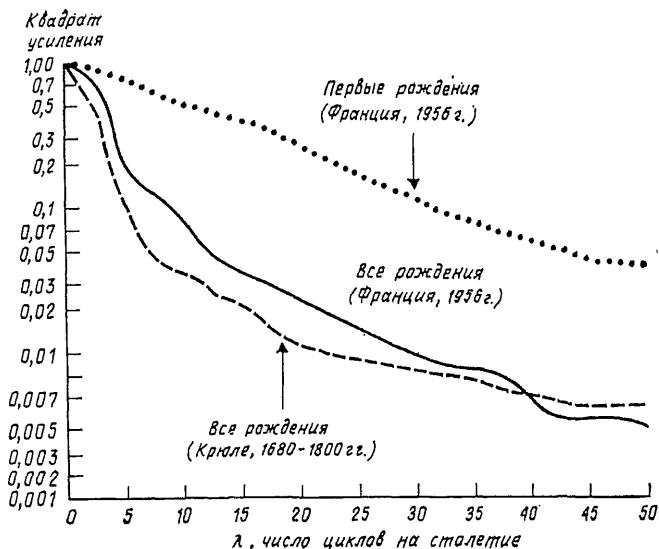


Рис. 1. Квадрат усиления для чистых функций брачной рождаемости.

Источники:  $F_x^1$  для Франции [28, с. 200],  $F_x$  для Франции [28, с. 230],  $F_x$  для Крюле вычислено по таблицам в [10]

что на большей части спектра первые рожденья примерно в 12 раз более чувствительны к колебаниям в числах браков, чем рожденья всех очередностей. Сравнение функций по современным данным и по данным для Франции XVIII столетия показывает поразительно малое различие между ними, хотя среднее число рожденьев на один брак в Крюле было приблизительно вдвое больше, чем в современной Франции, а ожидаемая продолжительность жизни более чем в два раза меньше<sup>11</sup>. Таким образом, рис. 1, по-видимому, дает некоторые основания для обобщения.

<sup>11</sup> Напомним, однако, что и тот и другой показатели были преобразованы так, чтобы вместе они давали чистый коэффициент воспроизводства, равный единице.

Наиболее важная особенность рисунка заключается в резком уменьшении квадрата усиления на более высоких частотах, указывающем на то, что за исключением низких частот на числа рождений передается очень небольшая часть дисперсии чисел браков. Функции для рождений всех очередностей показывают, что при колебаниях продолжительностью в 20 лет на рождения передается около  $1/10$  дисперсии; для 15 лет эта доля снижается до  $1/20$ ; для 10 лет — ниже чем до  $1/30$ ; для 5 лет — до  $1/80$  и для 2 лет она снижается до  $1/200$ . Рассмотрев характер изменения  $|F(\lambda)|^2$ , можно заключить, что влияние на числа рождений колебаний чисел браков с малым периодом почти полностью прекратится, а колебаний с большим периодом — станет существенно слабее. Следовательно, изменения брачности *в принципе* не могут объяснять большую часть быстрых изменений в ряде чисел рождений. Это предполагает, что значительные быстрые изменения, которые фактически обнаруживает ряд рождений, следует объяснять изменением брачной рождаемости.

## 5. ПОЛНАЯ МОДЕЛЬ И ГИПОТЕЗА ЕСТЕСТВЕННОЙ РОЖДАЕМОСТИ

Предшествующее обсуждение функции квадрата усиления поставило под большое сомнение сильную гипотезу естественной рождаемости еще до изучения каких-либо фактических рядов. В данном разделе мы рассмотрим оценки функций взаимного спектра для реальных исторических временных рядов в сравнении с теоретическими функциями, отвечающими гипотезам *A*, *C*, *D*. (Гипотеза *B* опущена ввиду того, что она мало разработана в литературе и ее введение усложнило бы диаграммы. Она будет рассмотрена в разделе, посвященном циклам.) Эмпирические оценки взаимных спектров были получены для следующих населений: пять французских приходов, включая Крюле (1680—1800 гг.); Швеция (1749—1880 гг.); Англия и Уэльс (1780—1870 гг.); два английских прихода (1545—1838 гг.); Акерсхюс, Норвегия (1735—1865 гг.); США (1920—1970 гг.). Поскольку результаты в каждом случае весьма схожи, ограничимся анализом оценок для Швеции<sup>12</sup>.

<sup>12</sup> Данные для Швеции взяты из [31, с. 38—44]. Эмпирические расчеты были проведены по программе, составленной с применением окна Парзена профессором Филипом Хоури (Phillip Howrey),

Начнем с квадрата усиления (squared gain), показанного на рис. 2. Согласно гипотезе *A* или *C* теоретический квадрат усиления очень быстро понижается при увеличении

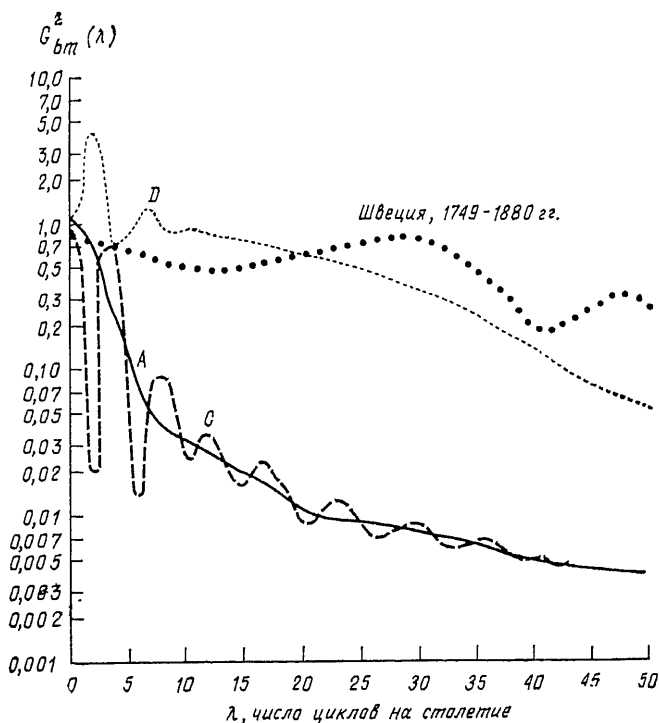


Рис. 2. Теоретические и оцененные значения квадрата усиления вариации чисел рождений по сравнению с вариацией чисел браков.

Примечание. Гипотезы *A*, *C*, *D* интерпретируются следующим образом. Гипотеза *A*: брачная рождаемость постоянна. Гипотеза *C*: брачная рождаемость и брачность варьируют независимо. Гипотеза *D*: брачная рождаемость и брачность коррелируют; лаг составляет 4 месяца.

Источник. Для теоретического квадрата усиления см. текст. Данные по Швеции взяты из [31, с. 38—44]

которую он великодушно предоставил в наше распоряжение. Данные имелись за 132 года, и оценки основаны на 20 лагах. Анализ динамики этих данных во времени см. в [33]. Источники данных для других рядов: Крюле [10, с. 242—244]; Акерсхюс [8, с. 169—173; 180—184]; Англия [25]; английские приходы — ряды для Колитона и Хартленда великодушно предоставлены профессором Э. А. Ригли (E. A. Wrigley).

частоты, что отражает ослабевающее влияние колебаний чисел браков на рождения, как это было показано на рис. 1. Однако согласно гипотезе  $D$  квадрат усиления остается относительно высоким; это происходит не в результате влияния числа браков на число рождений, а, скорее, вследствие влияния общей компоненты вариации в брачности и брачной рождаемости. Как видно, для Швеции оценка квадрата усиления для чисел рождений против чисел браков сравнительно хорошо соответствует гипотезе  $D$  и совершенно не согласуется с гипотезами  $A$  и  $C$ .

Рис. 3 показывает угол фазы (phase angle) — то запаздывание (лаг) чисел рождений по сравнению с числами браков, которое дает максимальную корреляцию между ними на каждой частоте. Согласно гипотезе  $A$  сдвиг фазы получается исключительно за счет передаточной функции  $F(\lambda)$ , и в большей части спектра он мало отличается от сдвига фазы согласно гипотезе  $C$ . Предполагаемый временной лаг равен одному году для колебаний с наименьшим периодом и постепенно увеличивается до шести лет для колебаний с периодом в 30 лет и больше. В отличие от него лаги, предполагаемые гипотезой  $D$  и по фактическим данным для Швеции, гораздо короче. Действительно, короткий лаг, вычисленный по данным для Швеции, можно объяснить лишь взаимодействием, которое подразумевает гипотеза  $D$ : брачная рождаемость и брачность сильно коррелированы и реагируют на одни и те же или сходные факторы. Реакция брачной рождаемости отстает от реакции брачности примерно на четыре месяца. Это означает, что зачатия должны опережать браки приблизительно на пять месяцев<sup>13</sup>.

На рис. 4 представлен квадрат согласованности (coherence squared) в рядах чисел рождений и чисел браков, показывающий долю дисперсии каждого ряда, которую на каждой частоте можно объяснить их взаимодействием. Согласно гипотезе  $A$  вся вариация в ряде чисел рождений прямо или косвенно определяется вариацией в числах браков, так что квадрат согласованности на всех частотах равен 1. Согласно гипотезе  $C$  брачная рождаемость составляет источник вариации в ряде чисел рождений, не зависящий от брачности, и поскольку на более высоких частотах влияние числа браков на число рождений становится все

<sup>13</sup> Доверительные интервалы для оценок  $\Phi(\lambda)$  зависят от  $\gamma^2(\lambda)$ . Когда оцененное значение  $\gamma^2(\lambda)$  равно 0,5 (см. рис. 4), 95%-ный доверительный интервал для  $\Phi(\lambda)$  равен  $\pm 0,5$  (см. [16, с. 381]).

слабее, этот независимый источник вариации начинает преобладать, поэтому квадрат согласованности быстро снижается. Отметим также отдельный пик, наблюдаемый приблизительно на периоде в 30 лет\*, отражающий влияние цикла поколения (который будет рассмотрен в разделе 7) как на ряд рождений, так и на ряд браков.

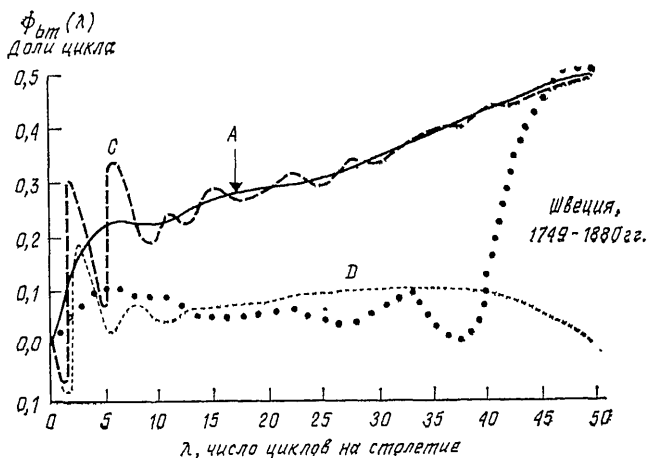


Рис. 3. Теоретические и оцененные значения сдвига фазы чисел рождений по сравнению с числами браков.

Примечание. См. примечание к рис. 2.

Источник. Для теоретического сдвига фазы см. текст. Данные по Швеции взяты из [31, с. 38—44]

Согласно гипотезе *D* квадрат согласованности снижается приблизительно линейно почти по всему спектру<sup>14</sup>. Поскольку кривая для гипотезы *C* показывает только согласованность чисто демографического происхождения, разница между ней и кривой для гипотезы *D* выражает дополнительную согласованность, возникающую вследствие корреляции брачности и брачной рождаемости. Из рис. 4 видно, что для частот выше пяти циклов на столетие почти вся согла-

\* Т. е. при частоте 3 цикла на столетие. — *Примеч. ред.*

<sup>14</sup> Это снижение происходит в результате того, что с возрастанием частоты передаточная функция  $K(\lambda)$  от ряда браков передает на ряд рождений все меньшие доли их общей компоненты дисперсии,  $g_\eta(\lambda)$ . Можно полагать, что в модели с непрерывным временем этого бы не происходило.

сованность по гипотезе  $D$  возникает из-за этой корреляции под влиянием одинаковых возмущений; демографическое влияние числа браков на число рождений не вызывает практически никакой согласованности.

Отметим, что согласно каждой из трех гипотез при частоте 0 согласованность равна 1; при колебаниях с очень большим периодом числа браков и числа рождений будут изменяться совместно, поскольку основным фактором из-

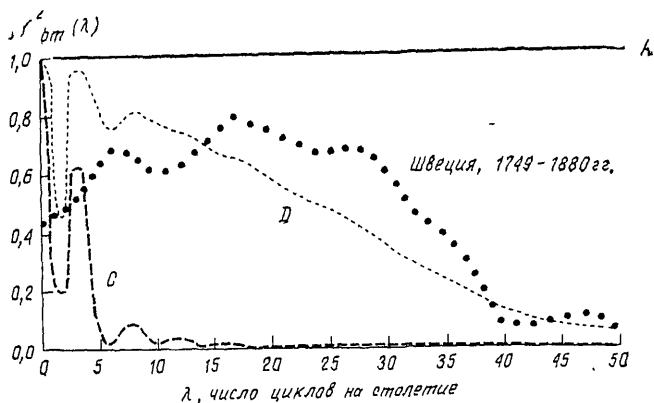


Рис. 4. Теоретические и оцененные значения квадрата согласованности вариации чисел рождений и браков.

Примечание. См. примечание к рис. 2.

Источник. Для теоретического квадрата согласованности см. текст. Данные по Швеции взяты из [31, с. 38—44]

менения будет исходная численность населения. Снова становится ясно, что только гипотеза  $D$  дает достаточно хорошее приближение к оценке исследуемой зависимости для Швеции<sup>15</sup>.

Следовательно, мы можем окончательно отвергнуть как сильную, так и слабую гипотезы естественной рождаемости; остается гипотеза о том, что брачная рождаемость и брачность сами по себе очень тесно коррелируют<sup>16</sup>. Отражает ли эта связь намеренное регулирование деторождения в браке,

<sup>15</sup> Оценка  $r^2$  ( $\lambda$ ), большая, чем 0,25, значимо отличается от 0 при 95%-ном уровне значимости (см. [16, с. 379]).

<sup>16</sup> Этот результат наиболее справедлив для больших популяций. В популяциях величиной в приход с одной или двумя тысячами человек большую роль играют чисто случайные колебания рождений и браков.



нельзя сказать с уверенностью. Однако в других исследованиях было показано при анализе взаимных спектров по многим переменным, что частная согласованность чисел рождений с числами браков намного сильнее, чем с числами смертей [25]; это заставляет сильно сомневаться в справедливости объяснения колебаний чисел рождений «сокращением физиологической плодовитости».

## 6. НЕКОТОРЫЕ ПРИЛОЖЕНИЯ МОДЕЛИ

Существует большое количество работ о влиянии экономических циклов на рождаемость и брачность (см. например, [18], [30]). Как правило, в этих исследованиях после исключения кратковременных колебаний с помощью скользящих средних или выравнивания обнаруживается тесная корреляция брачности и экономического индекса, а также умеренная частная корреляция рождаемости с брачностью и этим индексом. В своей хорошо известной работе Керк, основываясь на таких результатах, делает вывод о том, что «экономические условия определяют почти половину годовой дисперсии рождаемости около ее тренда — причем половина этого влияния передается через брачность, а другая половина сказывается непосредственно на рождаемости» [18, с. 253—254].

Из разделов 4 и 5 должно быть ясно, что такой вывод произволен, а включение показателей брачности в уравнение регрессии неправомерно. Сама операция по исключению тренда устраняет почти всю дисперсию рождаемости, которая не «непосредственна», т. е. не обусловлена брачной рождаемостью. Это хорошо показывает рис. 5, который изображает  $\gamma_{be}^2(\lambda)$  — квадрат согласованности вариации чисел рождений и брачной рождаемости согласно гипотезе С. Период экономического цикла длится от 3 до 9 лет; в этом диапазоне  $\gamma^2$  выше 0,95 и брачность, в сущности, не оказывает влияния на число рождений. Отсюда можно сделать вывод о том, что результаты этих исследований должны быть интерпретированы заново и что при дальнейших исследованиях в этой области нужно более тщательно подбирать переменные для анализа.

С другой стороны, сильная согласованность, показанная на рис. 5, указывает на то, что невыравненный ряд чисел рождений можно эффективно применять для изучения вариации брачной рождаемости, не прибегая к более точным показателям, исключаяющим влияние возрастной струк-

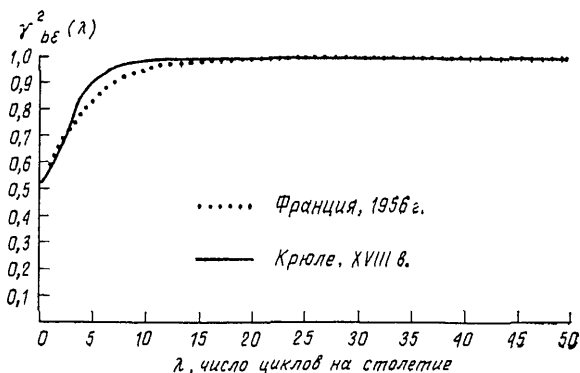


Рис. 5. Теоретические значения квадрата согласованности вариации чисел рождений и брачной рождаемости.

Примечание. Согласно гипотезе С брачная рождаемость и брачность варьируют независимо.

$$1 - \gamma_{бг}^2(\lambda) = \gamma_{б\theta}^2(\lambda).$$

Источник. Метод вычисления см. в тексте. Исходные данные взяты из [28, с. 230] и [10]

туры и брачности. Например, для колебаний с периодом, меньшим 15 или 20 лет, взаимный спектр ряда цен на зерно и ряда крещений по приходской метрической книге мог бы дать представление о влиянии цен на брачную рождаемость. Согласованность при годичном цикле охватывает от 85 до 90% спектра. Аналогичные соображения справедливы для чисел браков и брачности.

## 7. ТЕОРИЯ ЦИКЛОВ ПОКОЛЕНИИ

Специалисты по исторической демографии часто отмечают тот факт, что многие ряды чисел крещений и браков подвержены, по-видимому, длительным волнообразным колебаниям с периодом от 25 до 35 лет (см. [1], [9], [12], [21], [36]). Очевидно, что эти волны порождены не колебаниями, обусловленными циклическим движением ведущих рядов, как в случае более коротких циклов. Предположение о наведенном характере этих колебаний отвергается отчасти из-за того, что основные экономические ряды не обнаруживают соответствующей периодичности (см., например,

[14]), а также потому, что существуют другие объяснения, опирающиеся на внутреннюю структуру процесса возобновления населения.

Здравый смысл подсказывает, что в последовательности поколений более многочисленные когорты по достижении зрелости сами произведут на свет более многочисленные когорты. Этот вывод получает формальное выражение в математической демографии и представляет собой классическое объяснение циклов поколений. Другой аргумент приписывает существование циклов в первую очередь институту брака, придающему процессу рождаемости инерцию (автокорреляцию первого порядка).

Классический анализ основывается на детерминистском уравнении возобновления населения  $B_t = \sum B_{t-a} \Phi_a$ , или в принятых в этой работе обозначениях  $\sum c_i b_{t-i} = 0$ . При стабильных темпах роста  $B_t = e^{rt} B_0$  и  $b_t = 0$  для всех  $t$ . Однако, когда возрастная структура нарушается вследствие какого-либо резкого воздействия извне, то ряды, продолжая стремиться к стабильности, начинают колебаться, причем периоды колебания можно определить на основании комплексных корней уравнения  $\sum c_i y^i = 0$ . Преобладающий период колебаний приблизительно равен среднему возрасту деторождения, который в зависимости от конкретного населения составляет от 25 до 33 лет (см. [1], [3], [17], [19]). Такое объяснение циклов предполагает, что после первоначального толчка коэффициенты рождаемости и смертности остаются относительно постоянными и затем колебания очень быстро затухают. В действительности наблюдаемым циклам, как правило, не предшествует такой отдельный исключительный толчок, а циклы слишком устойчивы, чтобы согласоваться с теорией. Классическое объяснение приходится признать недостаточным.

Стохастическая версия нашего анализа разрешает эти затруднения. Пусть  $\Phi_{a,t}$  будет элементом стационарного стохастического вектора чистых\* возрастных коэффициентов рождаемости, так что уравнение возобновления населения принимает вид  $B_t = \sum \Phi_{a,t} B_{t-a}$  (см. (2.3)). После преобразований, рассмотренных в разделе 2, приближение Тейлора первого порядка можно выразить следующим образом<sup>17</sup>:

\* См. примеч. ред. на с. 74.

<sup>17</sup> Более подробно эта проблема рассматривается в работах [24] и [22].

$$\sum c_i b_{t-i} = \delta_t, \quad (7.1)$$

где  $\delta_t$  — отклонение чистого коэффициента воспроизводства от единицы, или  $\Sigma \Phi_{a,t} - 1 = \delta_t$ . Здесь не проводится различие между вариацией  $\delta$ , обусловленной брачностью или брачной рождаемостью.

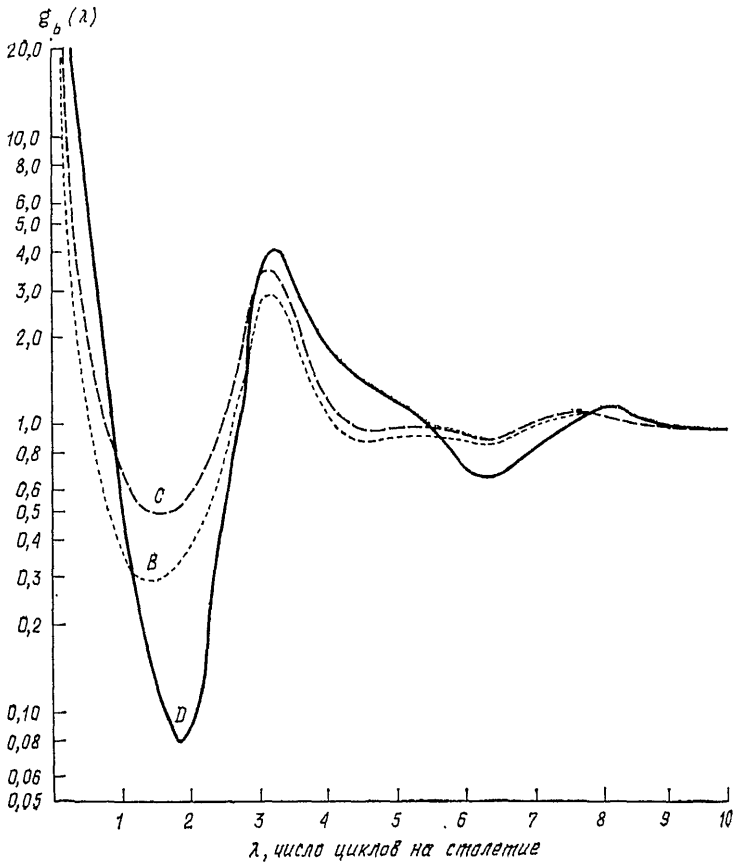


Рис. 6. Теоретический спектр чисел рождений.

Примечание. Гипотезы *B*, *C* и *D* интерпретируются следующим образом. Гипотеза *B*:  $\delta$  есть белый шум. Гипотеза *C*: брачная рождаемость и брачность варьируют независимо. Гипотеза *D*: брачная рождаемость и брачность коррелируют.

Источник. Метод вычисления см. в тексте. Исходные данные вычислены по таблицам в [10]

Уравнение (7.1) предполагает следующее спектральное соотношение:

$$g_b(\lambda) = \{1/|C(\lambda)|^2\} g_\delta(\lambda). \quad (7.2)$$

Функция квадрата усиления,  $1/|C(\lambda)|^2$ , показывает, как возрастная структура населения изменяет характер вариации коэффициентов естественного движения при передаче ее на числа рождений. Вид функции квадрата усиления показан как гипотеза  $B$  на рис. 6. Наблюдается выраженный пик на периоде, соответствующем среднему возрасту деторождения, показывающий, что некоррелированная случайная вариация (белый шум) будет преобразовываться в квазициклические волны поколений<sup>18</sup>.

Такое обобщение результатов классического анализа исключает нереалистическое предположение о катастрофическом толчке и последующем постоянстве демографических коэффициентов. Оно показывает, как из продолжительных рядов незначительных нарушений образуется энергия, позволяющая сохраниться циклическому характеру колебаний. Однако можно заметить, что это обобщение все еще недостаточно.

Иное толкование циклов было предложено Карлссоном [2]. Он предположил, что чистая функция рождаемости по продолжительности брака,  $f_x$ , когда в нее вводится элемент, представляющий умеренно автокоррелированные нарушения, придает ряду чисел рождений автокорреляцию первого порядка, достаточную для того, чтобы породить средний период взаимных колебаний примерно в 30 лет. Согласно этой точке зрения «тридцатилетний «период» не имеет ничего общего с промежутком между двумя поколениями, например, между рождением и вступлением в брак» [2, с. 417].

Если бы это утверждение было правильным, то мы могли бы ожидать, что автокорреляционные функции фактических рядов чисел рождений и браков будут монотонно снижаться с некоторым запаздыванием (лагом). Однако оценки значений этих функций показывают пики приблизительно в 30 лет [21]. Таким образом, теория Карлссона

---

<sup>18</sup> Анализ стохастической модели мог бы быть теснее связан с классическим детерминистским анализом путем получения автокорреляционной функции с теми же комплексными корнями, что и в детерминистском анализе, но с иными весами (см. [16] или [24]).

неверна. Причина ясна: эта теория игнорирует обратную связь в системе. Правильно подчеркивая, что числа рождений зависят от чисел браков, она отвергает то, что в равной степени верно, а именно что и числа браков зависят от чисел рождений. Не учитывать это — все равно, что, основываясь при анализе на формуле (2.1), исключить формулу (2.2).

Тем не менее Карлссон прав в том, что институт брака, добавляя устойчивость случайной вариации, увеличивает автокорреляцию в рядах чисел рождений. Это усиливает колебания в числах рождений с большим периодом, включая цикл поколений, и ослабляет колебания с меньшим периодом<sup>19</sup>. Таким образом определение (7.1) не охватывает всю предшествующую демографическую информацию относительно механизма, порождающего вариацию чисел рождений. Если учесть эти устойчивые влияния вариации в брачности на чистую функцию рождаемости, то мы получим

$$\delta_t = \varepsilon_t + \sum_{i=x}^{\infty} \theta_{t-i}. \quad (7.3)$$

Справедливость этого выражения можно подтвердить подстановкой из (7.3) в (7.1) и сравнением результата с (2.5).

Теоретический спектр чисел рождений, выраженный через спектры брачности  $[g_0(\lambda)]$  и брачной рождаемости  $[g_e(\lambda)]$ , становится теперь более сложным, чем (7.2). Он фактически уже был указан в таблице [см. с. 73] согласно гипотезам *A*, *B*, *C* и *D*. Сейчас гипотеза *A* нас не интересует. Гипотеза *B* дает тот же результат, что и стохастическая версия классической теории, а соответствующий спектр нанесен на рис. 6 точечной линией. Гипотеза *C* сочетает классический эффект и эффект Карлссона, она изображена на рис. 6 пунктирной линией. Эта гипотеза слегка подчеркивает цикл поколений, но эффект этот выражен не резко. Гипотеза *D* добавляет эмпирически оправданное предположение о том, что вариации брачной рождаемости и брачности тесно коррелируют; соответствующий спектр нанесен сплошной линией. Рельефность цикла значительно

---

<sup>19</sup> Очевидно, колебания в смертности дают тот же устойчивый эффект, однако мы не располагаем возможностью их здесь анализировать.

возрастает, особенно в отношении колебаний с относительно большим периодом<sup>20</sup>.

До сих пор мы рассматривали только циклы рождаемости, однако анализ такого рода вполне применим, разумеется, и к циклам брачности; относящиеся к нему уравнения также приведены в таблице. Спектры рядов браков во всех случаях имеют несколько более острые пики, чем спектры рядов рождений, показывая более выраженную циклическую тенденцию. Согласно гипотезе  $C$  квадрат согласованности для периода в 30 лет равен 0,63 и числа рождений отстают на шесть лет; согласно гипотезе  $D$  согласованность равна 0,96 с лагом, составляющим только четыре года (см. рис. 3 и 4)

## 8. ВЫВОДЫ

Как было показано, влияние вариации в рядах чисел браков на изменение параллельных рядов чисел рождений гораздо слабее, чем обычно считалось; так сезонные колебания в числах рождений — даже первых рождений — в сущности независимы от сезонности браков. Даже колебания в рядах чисел браков, имеющие относительно большую периодичность — до 15 или 20 лет, сказываются на рождениях очень незначительно. Неудивительно поэтому, что гипотеза естественной рождаемости применительно к временным рядам должна быть отвергнута. Быстрые колебания в рядах рождений в доиндустриальный период вызывались колебаниями брачной рождаемости, а не колебаниями брачности. Фактически показатели вариации для брачности и брачной рождаемости с малым периодом колебаний были приближенно равны. Кроме того, колебания брачности и брачной рождаемости были тесно коррелированы, что делает правдоподобным существование намеренного регулирования супругами деторождения в браке.

Проведенный анализ дает возможность сделать еще два заключения: (1) быстрые колебания в рядах рождений и браков представляют собой прекрасную модель быстрых колебаний в рядах брачной рождаемости и брачности и (2) значение «косвенного» влияния циклов деловой активности

---

<sup>20</sup> Эти уточненные определения также несколько изменяют длину цикла поколений. Дополнительное предположение, что члены, выражающие внешние возмущающие воздействия ( $\varepsilon$  и  $\theta$ ), обладают автокорреляцией первого порядка с  $\rho = 0,5$ , значительно увеличивает пик по сравнению со спектром на высокой частоте, но не может заметно изменить длину цикла.

на рождаемость — через брачность — было сильно преувеличено; фактически это влияние несущественно.

Нами также разработана методика анализа циклов поколений. Стохастическая версия классической теории, предполагающая, что возмущающие воздействия представляют собой белый шум, объясняет в значительной мере циклический характер изменения рядов. Пик спектра на частоте, соответствующей длине поколения, усилен автокорреляцией, придаваемой изменению рождаемости существованием самого института брака, но ее влияние не очень значительно. Если учесть также высокую корреляцию колебаний брачности и колебаний брачной рождаемости, то циклический характер изменения рождаемости еще более усилится.

#### ПРИЛОЖЕНИЕ А. ВЛИЯНИЕ НА ЧИСЛА РОЖДЕНИЙ ВОЗРАСТА ВСТУПЛЕНИЯ В БРАК И СМЕРТНОСТИ

Модель, разработанная в разделе 2, выделяет влияние на числа рождений изменений в числах браков, и в то же время она явно не учитывает возраста вступления в брак — переменной, роль которой часто подчеркивалась в литературе, посвященной доиндустриальной Европе. Это требует пояснения. Мы предложим две альтернативные модели, учитывающие изменение возраста вступления в брак, и покажем, что в обоих случаях значительное влияние на числа рождений в действительности оказывает лишь вызванное им изменение в числах браков, а не само изменение среднего возраста вступления в брак.

Обозначим через  $n(a)$  коэффициент брачности для никогда не состоявших в браке женщин возраста  $a$ <sup>21</sup>. Первая модель предполагает, что брачность изменяется во всех возрастах пропорционально множителю  $k$  так, что новый ряд возрастных показателей брачности равен:  $n^*(a) = kn(a)$ . Если  $k$  больше 1, то брачность увеличивается и число браков (при прочих равных условиях) будет также увеличиваться. Сначала это не вызовет изменений в среднем

---

<sup>21</sup>. Этот коэффициент брачности связан с чистым распределением брачности, введенным в разделе 2, следующим соотношением:

$$w_a = n_a p_a \prod_{i=0}^{a-1} (1 - n_i).$$



возрасте вступления в брак; женщины всех возрастов будут лишь более часто вступать в брак. В течение длительного периода средний возраст вступления в брак снизится из-за относительно большего истощения «запаса» незамужних женщин в старших возрастах. Однако это потребует продолжительного времени и не будет иметь силы для колебаний в брачности, имеющих малый период. Таким образом, в этом случае первый непосредственный эффект увеличения брачности скажется на числе браков, а не на среднем возрасте вступления в брак; последний изменится лишь во вторую очередь, причем этот эффект будет иметь значение только в течение длительного периода.

Вторая модель, по-видимому, менее правдоподобная, приводит к быстрым колебаниям в возрасте при вступлении в брак и небольшой вариации в числах браков. Предположим, что ряд возрастных показателей брачности сдвигается к более ранним возрастам на  $d$  лет, но в остальном остается неизменным, так что новый ряд их можно записать в виде  $n^i(a) = n(a - d)$ . Тогда средний возраст вступления в брак понизится почти точно на  $d$  лет. Это временно увеличит число браков, так как те же коэффициенты будут относиться к несколько большим числам незамужних (полностью этот эффект скажется спустя длительное время), и постоянно будет незначительно увеличивать числа браков за счет небольшого снижения смертности. Эффект этих изменений проявляется через числа браков. Скажется также чистое влияние возрастного состава, поскольку женщины, раньше вступившие в брак, имеют больше времени для деторождения. Произведя вычисления, находим, что снижение среднего возраста вступления в брак для когорты на один год (т. е.  $d = 1$ ) привело бы, через длительное время, к увеличению числа детей на 0,15 в расчете на один брак, или на 4%.

Для короткого периода эффект этих изменений будет, конечно, значительно меньше.

Влияние изменения в ожидаемой продолжительности жизни при рождении,  $e_0$ , также несущественно. Изменение  $e_0$  на один год вызывает изменение числа детей на 0,025 в расчете на один брак, или на 0,6%, причем также через длительное время. Основное влияние изменений смертности сказывается на числах вступающих в брак через изменения в чистых показателях брачности.

В противоположность этим довольно слабым влияниям изменения среднего возраста вступления в брак и изменения

смертности любое изменение чисел вступающих в брак за длительный период передается на числа рождений полностью (с эластичностью, равной единице). Следовательно, правильно вводить в модель изменения в числах вступающих в брак.

## ПРИЛОЖЕНИЕ В. ВЫВОД ТЕОРЕТИЧЕСКИХ ФУНКЦИИ ВЗАИМНОГО СПЕКТРА

Любой стационарный процесс  $x$ , не являющийся строго периодическим, может быть представлен в виде спектра  $z_x(\lambda)$ , обладающего таким свойством, что  $E [z_x(\lambda) \overline{z_y(\lambda)}] = g_{y,x}(\lambda)$ , где  $g_{y,x}(\lambda)$  — взаимный спектр  $y$  и  $x$  (см. [16] или [13]).

Обозначим через  $z_\theta(\lambda)$ ,  $z_\epsilon(\lambda)$ ,  $z_b(\lambda)$  и  $z_m(\lambda)$  соответствующие спектральные представления процессов  $\theta$ ,  $\epsilon$ ,  $b$  и  $m$ . Уравнения (2.9) и (2.10) выражают отношения между этими переменными во временном аспекте через функции  $f$ ,  $\omega$ ,  $c$ . Подобным же образом можно выразить отношения между спектрами этих переменных в частотном аспекте, выразив их с помощью соответствующих функций, которые обозначим  $F(\lambda)$ ,  $W(\lambda)$  и  $C(\lambda)$ . Например, передаточные функции  $f_x$  вычисляются как  $F(\lambda) = \sum_x f_x \exp(-ix\lambda)$ , где  $i = \sqrt{-1}$  (см. [16] или [13]). Передаточная функция, умноженная на собственное сопряжение, равна ее «квадрату усиления»: так  $F(\lambda) \overline{F(\lambda)} = |F(\lambda)|^2$ , т. е. квадрату усиления  $F(\lambda)$ .

Из (2.9) и (2.10) имеем:

$$C(\lambda) z_m(\lambda) = F(\lambda) z_\theta(\lambda) + z_\epsilon(\lambda); \quad (B.1)$$

$$C(\lambda) z_m(\lambda) = W(\lambda) z_\epsilon(\lambda) + z_\theta(\lambda). \quad (B.2)$$

Спектры  $b$  и  $m$  получают путем умножения каждой из частей уравнений (B.1) и (B.2) на их соответствующие сопряжения и вычисления математических ожиданий. Подобным же образом взаимный спектр  $b$  и  $m$  можно найти, умножая обе части уравнения (B.2) на сопряжение соответствующей части уравнения (B.1).

Когда колебания брачности и брачной рождаемости при всех лагах не коррелированы, члены  $z_\theta(\lambda)$  и  $z_\epsilon(\lambda)$  и их сопряжения равны нулю. Вообще, однако, они не будут равны нулю и их можно получить из соответствующих спектральных выражений; так, из (2.11) и (2.12) имеем:

$$z_\epsilon(\lambda) = z_{\epsilon'}(\lambda) + K(\lambda) z_\eta(\lambda); \quad (B.3)$$

$$z_0(\lambda) = z_0'(\lambda) + z_\eta(\lambda). \quad (B.4)$$

Взаимный спектр  $\varepsilon$  и  $\theta$  будет тогда:

$$g_{\varepsilon\theta}(\lambda) = \overline{K(\lambda)} g_\eta(\lambda). \quad (B.5)$$

Теперь члены, содержащие  $z_0(\lambda)$   $\overline{z_\varepsilon(\lambda)}$  или их сопряжения, должны быть введены в спектры  $b$  и  $m$  и в их взаимный спектр. Эти спектры можно выразить следующим образом:

$$g_b(\lambda) = (1/|C(\lambda)|^2) \{ |F(\lambda)|^2 g_0(\lambda) + g_\varepsilon(\lambda) + 2g_\eta(\lambda) RE [F(\lambda) K(\lambda)] \}; \quad (B.6)$$

$$g_m(\lambda) = (1/|C(\lambda)|^2) \{ |W(\lambda)|^2 g_\varepsilon(\lambda) + g_0(\lambda) + 2g_\eta(\lambda) RE [W(\lambda) K(\lambda)] \}; \quad (B.7)$$

$$g_{bm}(\lambda) = (1/|C(\lambda)|^2) \{ W(\lambda) g_\varepsilon(\lambda) + \overline{F(\lambda)} g_\theta(\lambda) + \overline{K(\lambda)} g_\eta(\lambda) + W(\lambda) \overline{F(\lambda)} K(\lambda) g_\eta(\lambda) \}, \quad (B.8)$$

где функция  $RE( )$  есть действительная часть комплексного аргумента.

Из основных и спектральной и взаимноспектральной функций, представленных в таблице (см. с. 73), можно получить значения усиления, обозначаемого через  $G(\lambda)$ , сдвига фазы, обозначаемого через  $\Phi(\lambda)$ , и квадрата согласованности, обозначаемого через  $\gamma^2(\lambda)$ :

$$G_{bm}(\lambda) = g_{bm}(\lambda)/g_m(\lambda); \quad (B.9)$$

$$\Phi_{bm}(\lambda) = \text{Arctan} \{ IM [g_{bm}(\lambda)] / RE [g_{bm}(\lambda)] \}; \quad (B.10)$$

$$\gamma_{bm}^2(\lambda) = g_{bm}(\lambda)^2 / [g_b(\lambda) g_m(\lambda)]. \quad (B.11)$$

Когда брачная рождаемость постоянна (согласно гипотезе А), они принимают простую форму. Усиление есть просто усиление  $f$ , т. е.  $|F(\lambda)|$ ; сдвиг фазы равен просто сдвигу  $f$ , т. е.  $\text{Arctan} \{ IM [F(\lambda)] / RE [F(\lambda)] \}$ ; квадрат согласованности равен 1,0 на всех частотах. Сходные результаты можно получить согласно гипотезе В. Вообще, однако, эти выражения гораздо сложнее.

## ЛИТЕРАТУРА

1. Bernardelli Harro. Population Waves. *Journal of the Burma Research Society*, 1941, april, 31, 1—18.
2. Carlsson Gosta. Nineteenth Century Fertility Oscillations *Population Studies*, 1970, november, 24, 413—422.

3. Coale Ansley. The Growth and Structure of Human Populations. Princeton, Princeton University Press, 1972.
4. Cohen Joel. Ergodicity of Populations with Markovian Vital Rates. Unpublished manuscript, Department of Biology, Harvard University, 1973.
5. Connor L. R. Fertility of Marriages and Population Growth. *Journal of the Royal Statistical Society*, 1926, may, 89, 553—566.
6. Deprez P. The Demographic Development of Flanders in the Eighteenth Century. In: D. V. Glass and D. E. C. Eversley (eds.). *Population in History*, Chicago, Aldine Publishing, 1965, 608—630.
7. Deprez P., Hum D. and Spencer B. Spectral Analysis and the Study of Seasonal Fluctuations in Historical Demography. Unpublished manuscript, Department of Economics, University of Manitoba, 1974.
8. Drake Michael. *Population in Industrialization*. London, Cambridge University Press, 1969.
9. Eversley D. E. C. Population, Economy and Society. In: D. V. Glass and D. E. C. Eversley (eds.), *Population in History*. Chicago, Aldine Publishing, 1965, 23—69.
10. Gautier Etienne and Henry Louis. *La Population de Crulai*. Paris, Presses Universitaires de France, 1958.
11. Gluckman Perry, Herlihy David and Pori Mary. A Spectral Analysis of Deaths in Florence. Unpublished manuscript, Center for Advanced Study of the Behavioral Sciences, Stanford University, 1973.
12. Goubert Pierre. Recent Theories and Research in French Population Between 1500 and 1700. In: D. V. Glass and D. E. C. Eversley (eds.). *Population in History*. Chicago, Aldine Publishing, 1965, 457—473.
13. Granger C. W. J. and Hatanaka M. Spectral Analysis of Economic Time Series. Princeton, Princeton University Press, 1964. Русский перевод: Гренджер К., Хатанакка М. Спектральный анализ временных рядов в экономике. М., Статистика, 1972.
14. Granger C. W. J. and Hughes A. O. A New Look at Some Old Data. The Beveridge Wheat Price Series. *Journal of the Royal Statistical Society*, 1971, Series A, 134, part 3, 413—428.
15. Henripin J. *La Population Canadienne au Début du XVIII-e Siècle*. Paris, I. N. E. D., 1954.
16. Jenkins Gwilym and Watts Donald. *Spectral Analysis and Its Applications*. San Francisco, Holden — Day Inc., 1968. Русский перевод: Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. М., Мир, 1971, вып. 1; 1972, вып. 2.
17. Keyfitz Nathan. *Introduction to the Mathematics of Population*. Reading, Massachusetts, Addison — Wesley Publishing Co., 1968.
18. Kirk Dudley. The Influence of Business Cycles on Marriage and Birth Rates. In: Universities — National Bureau for Economic Research. *Demographic and Economic Change in Developed Countries*. Princeton, Princeton University Press, 1960, 241—259.

19. Le Bras Hervé. Retour d'une Population a L'état Stabl Après une «Catastrophe». *Population*, 1969, september — october, 24, 861—896.
20. Le Bras Hervé. Eléments pour une Théorie des Populations Instables. *Population*, 1971, may — june, 26, 525—572.
21. Lee Ronald. Econometric Studies of Topics in Demographic History. Unpublished Ph. D. dissertation, Department of Economics, Harvard University, 1970.
22. Lee Ronald. Forecasting Births in Post — Transition Populations. Stochastic Renewal with Serially Correlated Fertility. *Journal of the American Statistical Association*, 1974, september, 69, № 347, 607—617.
23. Lee Ronald. Estimating Series of Vital Rates and Age Structures from Baptisms and Burials. A New Technique, with Applications to Preindustrial England. *Population Studies*, 1974, november, 28, 495—512.
24. Lee Ronald. The Formal Dynamics of Controlled Populations and the Echo, the Boom and the Bust. *Demography*, 1974, november, 11, 563—585.
25. Lee Ronald. Models of Preindustrial Population Dynamics, with Applications to England. In: Charles Tilly ed. Historical Studies of Changing Fertility. Princeton, Princeton University Press, 1975.
26. Ohlin Goran. The Positive and the Preventive Check. Unpublished Ph. D. dissertation, Department of Economics, Harvard University, 1955.
27. Parzen Emanuel. Stochastic Processes. San Francisco, Holden — Day Inc., 1962.
28. Pressat Roland. Demographic Analysis. Chicago and New York, Aldine — Atherton, 1972.
29. Schofield Roger. Historical Demography. Statistical Problems. Paper presented at the Conference of the International Union for the Scientific Study of Population. Liege, 1973.
30. Silver M. Births, Marriages and Business Cycles in the U. S. *Journal of Political Economy*, 1965, june, 73, 237—255.
31. Swedish Central Bureau of Statistics. Historical Statistics of Sweden, vol., 1, Population. Stockholm, Statens Reproduktion — Sanstalt, 1955.
32. Sykes Z. M. Some Stochastic Versions of the Matrix Model for Population Dynamics. *Journal of the American Statistical Association*, 1969, march, 64, 111—131.
33. Thomas Dorothy S. Social and Economic Aspects of Swedish Population Movements 1750—1933. New York, The Macmillan Co., 1941.
34. U. S. National Center for Health Statistics. Seasonal Variation of Births in the United States 1933—1963. *Vital and Health Statistics*, 1966, may, Series, 21, № 9.
35. Van de Walle Etienne. Marriage and Marital Fertility. *Daedalus*, 1968, spring, 486—501.
36. Wrigley E. A. Population and History. New York, McGraw—Hill Book Co., 1969.
37. Yule G. Udney. Changes in the Marriage and Birth Rates in England and Wales During the Past Half Century. *Journal of the Royal Statistical Society*, 1906, march, 69, 18—132.

*Тадеуш Грабиньский, Казимеж Заёц*

## ТАКСОНОМИЧЕСКИЕ МЕТОДЫ ОПРЕДЕЛЕНИЯ ФАЗ РАЗВИТИЯ ДЕМОГРАФИЧЕСКИХ ПРОЦЕССОВ

Tadeusz G r a b i ń s k i, Kazimierz Z a j ą c. Taksonomiczne metody określania faz rozwojowych procesów demograficznych. *Studia demograficzne*, № 43, 1976, s. 15—39.

### 1. ВВОДНЫЕ ЗАМЕЧАНИЯ

Цель работы — представить новые численные методы исследования многомерных временных рядов. На основе этих методов была сделана попытка проверить гипотезу о том, что развитие населения за длительный период представляет собой неоднородный процесс, поскольку оно обусловлено несколькими действующими в разных направлениях главными причинами, которые обозначают поворотные моменты в тенденции развития демографических явлений. Нам представляется, что до сих пор этой важной проблеме отводилось в исследованиях слишком мало места. Следует заметить также, что в имеющихся статистических разработках понятие однородности связано лишь с разрезом на данный момент времени. Временной же аспект однородности учитывается, в общем, косвенным образом, путем установления лишь временных рамок анализа.

Построению любых моделей, относящихся к временным рядам, должна предшествовать оценка однородности протекания исследуемых явлений во времени. Между тем до сих пор необходимые данные получали методом сравнительно-описательного анализа, который часто приводил к субъ-

ективным выводам. Как правило, разработки такого характера не позволяли однозначно выделить существенные фазы развития в рассматриваемой области.

Таким образом, многие ошибки в оценках тенденций развития были обусловлены недооценкой этого типа анализа.

Витольд Куля пишет: «Временные ряды, полученные историками, чаще всего представляют картину беспорядочных колебаний из года в год, с виду лишенных какой-либо тенденции развития. В науке нередки случаи, когда исследователь, построив динамический ряд, не замечает проявляющихся в нем тенденций. Несомненно, если имеешь дело с количественным материалом, то установление наличия или отсутствия тенденции развития не может быть предоставлено лишь наблюдательности историка. С помощью соответствующих операций она должна быть обнаружена и количественно оценена» [10].

Приведенная выдержка из рассуждений В. Кули позволяет оценить, насколько фундаментальную методическую проблему, даже в гуманитарных науках, представляет собой статистическая группировка, которой при изучении хронологических рядов до сих пор пренебрегали.

Статистико-исторический анализ преследует цель найти неизменные элементы в изменяющемся мире экономики, утверждает С. Кузнец [11 — цитируется по работе В. Кули]. Таким образом, отсутствие правильного понимания проблемы однородности временных рядов может привести к тому, что в условиях большой изменчивости анализируемых — чисто описательным методом — явлений их важный компонент может остаться незамеченным. Причина может заключаться в том, что на протяжении длительного времени этот компонент также изменяется. Особенно это касается демографических процессов.

Задача данной работы представить статистическую методику, позволяющую, по возможности однозначно, установить фазы развития демографических процессов в Польше за тридцать лет. Речь идет о разделении исследуемого периода на несколько не установленных заранее разомкнутых, исчерпывающих и возможно более однородных подпериодов, охватывающих наблюдения, максимально сходные с точки зрения рассматриваемых признаков. В результате это приведет к выделению фаз развития, существенно различающихся между собой.

При этом принципиально важно понятие периодизации, которое дает ответ на следующие вопросы (см. [8]):

1) протекало ли развитие выделенных демографических явлений в Польше за тридцатилетие равномерно и гармонично или в нем можно выделить определенные фазы?

2) если последнее возможно, то каким образом установить границы отдельных фаз развития?

3) как иерархизировать полученные границы по степени их важности?

Хотя работа имеет прежде всего методологический характер, она не лишена аспектов, касающихся существа проблемы, поскольку наши рассуждения относятся к демографическим процессам, протекающим в конкретных условиях места и времени. Следует подчеркнуть, что в прежних исследованиях в этой области объектами классификации были пространственные единицы, а иногда, при двойственных методах, и признаки, характеризующие эти единицы. Это автоматически влекло за собой статичность таксономического анализа и отсутствие возможности учесть динамический аспект изменений, происходящих в структуре исследуемых явлений.

В данной же работе представлена иная точка зрения, а именно предметом классификации служат периоды времени, в которые проводится статистическое наблюдение. Таким образом, открываются новые методические возможности в области изучения временных рядов, а также значительно облегчается выявление существующих закономерностей развития, которые не принимались во внимание в ходе проводившихся до сих пор исследований. Следствием такого подхода становится многовариантность суждений, что позволяет расширить интерпретацию результатов. При этом оказывается возможным более четко охарактеризовать связи, возникающие между демографическими и социально-экономическими факторами.

## 2. МЕТОДЫ ИССЛЕДОВАНИЯ

Периодизация, выполняемая с помощью статистических методов, представляет собой сложный процесс, в котором каждый из этапов одинаково важен с точки зрения правильности получаемых конечных результатов. В самом общем виде в процессе периодизации можно выделить три этапа исследования, тесно связанные между собой.

На первом предварительном этапе периодизации определяются предмет изучения по существу, временной диапазон и пространственные границы. Эти проблемы детерминиро-



ваны, с одной стороны, поставленной целью исследования и предварительно сформулированными рабочими гипотезами, а с другой — доступностью достоверных и сопоставимых эмпирических данных.

Второй, чрезвычайно важный этап периодизации связан с установлением перечня признаков, характеризующих развитие изучаемого демографического субъекта в рассматриваемый период. При исследовании социально-экономических явлений мы относительно редко имеем дело с ситуациями, в которых существующая теория дает исчерпывающие сведения относительно диагностичности тех или иных признаков. Не всегда известно также, в достаточной ли степени характеризует установленный набор признаков в целом рассматриваемые явления. В связи с этим содержательные предпосылки (*przesłanki merytoryczne*) должны быть дополнены сведениями, основанными на интуиции исследователя, ограничениями, обусловленными возможностями получения эмпирического материала, либо, наконец, чисто формальными критериями диагностичности признаков.

Что касается последнего, то в специальной литературе можно встретить, в частности, следующие требования, предъявляемые к признакам, обладающим большой диагностичностью [3]:

отражение наиболее существенных свойств анализируемых явлений;

независимость от внешних воздействий среды;

наименьшая корреляция с остальными диагностическими признаками;

наибольшая корреляция с признаками, непосредственно не учтенными в исследовании;

высокая степень изменчивости во времени.

Следует подчеркнуть, что в традиционных таксономических проблемах, где предметом рассмотрения служат пространственные единицы, от диагностических признаков требуется стабильность во времени, а условие большой изменчивости признаков выдвигается в отношении пространственного аспекта. При периодизации эти критерии должны быть изменены, поскольку признак, принимающий в отдельные моменты времени одинаковые значения, не вносит в рассуждения ничего нового.

Как показывает практика таксономических исследований, число принятых для анализа переменных не должно быть слишком большим, поскольку по мере его роста увеличивается опасность включения недиагностических призна-

ков, которые могут вызвать искажение окончательных результатов исследования.

С учетом сформулированных ранее замечаний нам представляется, что при выборе признаков для проведения периодизации хорошие результаты можно получить, применяя формально-математические методы в сочетании с содержательными сведениями. Сущность такого подхода заключается в том, что сначала устанавливается широкий набор признаков, по возможности всестороннее характеризующих изучаемый демографический процесс. При этом следует руководствоваться прежде всего знанием существа проблемы, а также реальными возможностями сбора соответствующих числовых материалов. Далее предварительно установленный набор признаков сводится к более узкому набору в соответствии с формальными критериями диагностичности, причем особенное внимание уделяется требованиям, предъявляемым применяемой методикой периодизации.

Пространство признаков предлагается сокращать, применяя общеизвестные таксономические методы в их двойственном виде<sup>1</sup>. Правда, проблема выбора диагностических признаков часто решается при помощи факторного анализа, но, как показывают результаты последних исследований, проведенных вроцлавской школой, и краковский опыт, таксономические методы дают, в общем, не худшие результаты, чем факторный анализ, причем они гораздо менее трудоемки. Небезразличен и тот факт, что применение факторного анализа приводит в итоге к замене исходных значений признаков новыми, связанными с ними, признаками, четкая содержательная интерпретация которых не всегда возможна.

Среди многих разработанных до сих пор таксономических алгоритмов в данной работе выбраны две методики, представляющие различные группы методов с точки зрения применяемых вычислительных приемов. Это сделано с целью взаимного контроля правильности принимаемых для периодизации совокупностей признаков, полученных при помощи разных подходов.

Перед тем как приступить к краткой характеристике рассматриваемых далее методов, следует напомнить, что в двойственных таксономических алгоритмах предметом классификации служат статистические признаки, рассмат-

<sup>1</sup> Впервые двойственные таксономические методы были применены с этой целью Ю. Стечковским в работе [17].

риваемые как отдельные точки в многомерном пространстве объектов. В процессе периодизации следует провести дополнительные модификации, учитывающие, что роль объектов играют не пространственные единицы, а единицы времени, в которых выполнены статистические наблюдения.

Первая из применяемых методик известна в литературе под названием вроцлавской таксономии [4] и представляет собой группу методов, в которых отдельным парам признаков приписывается заранее установленная мера подобия, а далее наиболее подобные один другому признаки объединяются в группы, вторая относится к территориальным методам и была разработана группой научных работников Высшей экономической школы во Вроцлаве [2]. В территориальных методах пространство объектов делится на разомкнутые подпространства, а принадлежащие к ним признаки рассматриваются как отдельные классы.

Отправным моментом во вроцлавской таксономии служит расчет показателей подобия между всеми возможными парами классифицируемых признаков. Далее на основе некоторой заданной матрицы подобия признаков строится вроцлавский дендрит (*dendryt wrocławski*) следующим образом:

указывается пара наиболее подобных один другому признаков;

находится следующий признак, наиболее подобный уже выбранным признакам;

такой порядок действий продолжается до тех пор, пока не будут исчерпаны все элементы классифицируемого множества признаков.

В результате получается совместный граф, определяющий упорядочение объектов и оказывающийся кратчайшим из всех возможных дендритных упорядочений. Отсекая последовательно самые длинные ветви дендрита, мы получаем желаемое разбиение множества признаков на произвольное число разомкнутых и исчерпывающих подмножеств, содержащих признаки, максимально подобные один другому в рамках данного подмножества и вместе с тем максимально отличающиеся от признаков, входящих в другие подмножества. Е. Перкаль [13] предложил иной способ разбиения дендрита, основанный на понятии силы разбиения.

Алгоритм второго метода, представленный в упомянутой работе [2], также очень прост и не требует больших затрат труда. Порядок действий можно изложить следующим образом:

регистрируется принадлежность отдельных признаков к гиперсферам с произвольно установленным радиусом, а также координатами центра, определенными путем наблюдения значений данного признака в последовательные годы;

гиперсфера с наибольшим числом признаков рассматривается как первая более однородная группа, и относящиеся к ней признаки исключаются из дальнейшего рассмотрения;

в отношении остальных признаков повторяется описанная процедура вплоть до исчерпания всех элементов разлагаемого таким способом множества.

Располагая распределением исходного множества признаков по группам подобных один другому признаков, можно выбрать те признаки, которые представляют отдельные группы. Представителем группы, насчитывающей более двух элементов (см. [14]), считается признак, наиболее подобный всем остальным в данной группе. Кроме того, такими представителями будут и все признаки, которые образуют группы, состоящие из одного элемента. Напротив, из групп, содержащих два элемента, выбирается тот признак, который наименее подобен ранее установленным признакам-представителям. Полученное свернутое множество признаков служит непосредственной основой для проведения процесса периодизации.

Последний, третий по счету этап периодизации заключается в установлении отдельных фаз в развитии изучаемого демографического процесса на основе полученного на предыдущем этапе набора диагностических признаков. Понятие фазы развития, принимаемое в соответствующих теоретических дисциплинах интуитивно, для нужд данной разработки должно быть определено ясно и четко. Требуется, чтобы определение фазы развития было дано в таком виде, который допускал бы его практическое применение в количественных методах. Именно такие методы составляют предмет наших дальнейших рассуждений.

Под фазой развития мы понимаем некоторый подпериод не установленной заранее продолжительности, заключенный в рамках анализируемого периода времени и характеризующийся тем, что все признаки, учитываемые при периодизации, будут приобретать более близкие одно к другому значения для данной фазы развития, чем в отношении соседних подпериодов.

Приведенное определение, хотя оно и не слишком точно, позволяет в общем очертить круг вычислительных методов, которые можно с успехом применить для проведения периоди-

зации. Нам представляется, что специфике проблемы в на и большей степени отвечают методы численной таксономии занимающейся теоретическими принципами, а также процедурами классификации и упорядочения объектов.

Прибегнув к геометрической интерпретации таксономических задач, проблему периодизации можно сформулировать следующим образом: исследуемые периоды времени, описанные с помощью выбранного перечня статистических признаков, рассматриваются как множество точек многомерного пространства. Множество этих точек следует подразделить на несколько разомкнутых подгрупп, максимально удаленных одна от другой в пространстве признаков и одновременно содержащих точки (периоды), максимально близкие одна к другой в смысле априорно установленной многопризнаковой меры отдаленности.

Стоит обратить внимание на то, что применение разработанных до сих пор таксономических алгоритмов небезопасно, поскольку они, как правило, не предусматривают условие линейного упорядочения классифицируемых объектов, которое, несомненно, должно соблюдаться при периодизации. К тому же применение известных методик либо требует серьезных их модификаций, либо влечет за собой необходимость проводить произвольные изменения получаемой классификации периодов.

Кроме таксономических методов периодизацию можно проводить с помощью других методов. Из наиболее распространенных здесь можно упомянуть дискриминантный анализ (в понимании Р. А. Фишера), некоторые разновидности факторного анализа и, наконец, алгоритмы быстро развивающегося в последние годы раздела кибернетики, называемого теорией распознавания образов и декодирования информации.

Выделение фаз развития с помощью упомянутых подходов сводится в принципе к адаптации их к условиям данной задачи и заключается в соединении этих методов с новой для них областью эмпирического применения. Хотя такое решение вполне возможно, лучшие результаты должны быть получены путем разработки методов, приспособленных исключительно для проведения периодизации.

Один из таких методов, хорошо учитывающий специфику проблемы периодизации, представлен в данной работе. В этом методе переплетаются элементы, характерные для таксономических процедур, с элементами теории статистических выводов, разработанными многомерной статистикой.

Построение алгоритма в предлагаемом методе отвечает условию линейного упорядочения классифицируемых периодов времени.

Пусть исследуемый демографический процесс будет характеризоваться  $m$ -элементным множеством признаков, каждый из которых приобретает заданные значения в течение  $n$  периодов времени. Наблюдаемые значения признаков можно представить в виде следующей матрицы исходных данных:

$$\mathbf{X} = \begin{bmatrix} x_{11} & \dots & x_{1j} & \dots & x_{1m} \\ \vdots & & \vdots & & \vdots \\ x_{l1} & \dots & x_{lj} & \dots & x_{lm} \\ \vdots & & \vdots & & \vdots \\ x_{n1} & \dots & x_{nj} & \dots & x_{nm} \end{bmatrix}. \quad (1)$$

С точки зрения многомерной статистики числа, находящиеся в последовательных строках матрицы  $\mathbf{X}$ , представляют собой отдельные реализации  $m$ -мерных случайных переменных:

$$\Xi = [\xi_1, \dots, \xi_j, \dots, \xi_m] \quad (t = 1, 2, \dots, n), \quad (2)$$

где  $\xi_j$  означает  $j$ -ю одномерную случайную переменную.

Упорядоченная во времени последовательность  $m$ -мерных случайных переменных,  $\Xi_t$ , образует  $n$ -элементное пространство  $\Xi^T$ :

$$\Xi^T = [\Xi_1, \dots, \Xi_t, \dots, \Xi_n]. \quad (3)$$

Матрица случайных переменных

$$\Xi^T = \begin{bmatrix} \xi_{11} & \dots & \xi_{1j} & \dots & \xi_{1m} \\ \vdots & & \vdots & & \vdots \\ \xi_{l1} & \dots & \xi_{lj} & \dots & \xi_{lm} \\ \vdots & & \vdots & & \vdots \\ \xi_{n1} & \dots & \xi_{nj} & \dots & \xi_{nm} \end{bmatrix}. \quad (4)$$

будет определяться статистической моделью демографического объекта (процесса). С формальной точки зрения эта модель — дискретное конечное и линейно упорядоченное множество  $T$  точек  $t$ , на котором задана многомерная случайная функция  $\Xi_t$ .

Модель демографического объекта считается статистически однородной, если соблюдается зависимость

$$|F_{t'}(\mathbf{X}) - F_{t''}(\mathbf{X})| = 0 \quad (5)$$

для всех  $t', t'' \in T$  и всех векторов реализации

$$x_t = (x_{t1}, \dots, x_{tj}, \dots, x_{tm}). \quad (6)$$

Символ  $F_t(\mathbf{X})$  означает многомерную функцию распределения случайного вектора  $\Xi_t$ , определяемую формулой

$$F_t(\mathbf{X}) = P(\xi_{t1} \leq x_{t1}, \dots, \xi_{tj} \leq x_{tj}, \dots, \xi_{tm} \leq x_{tm}), \quad (7)$$

где  $x_{tj}$  — конкретное значение реализации одномерной случайной переменной  $\xi_j$  в период  $t$ .

Если пространство  $\Xi^T$  статистически однородно, то каждое его подмножество или класс подмножеств также однородны. В таком случае выделение фаз развития демографического объекта не отвечает поставленной цели.

Однородность пространства  $\Xi^T$  в смысле соотношения (5) означает, что все многомерные распределения случайных переменных  $\Xi_t$  имеют идентичный вид. Если это условие не выполняется, то модель демографического объекта определяется как статистически неоднородная. В этой ситуации по крайней мере для одного вектора реализации  $\mathbf{X}$  должна существовать такая пара  $t', t'' \in T$ , что

$$|F_{t'}(\mathbf{X}) - F_{t''}(\mathbf{X})| > 0. \quad (8)$$

Констатация факта неоднородности пространства  $\Xi^T$  служит предварительным условием, которое определяет возможность провести разделение этого пространства на более однородные части и получить таким образом искомые границы фаз развития.

Приведенные определения имеют очень общий характер и непосредственно не пригодны для практического применения. С этой целью следует уточнить аналитический вид распределения отдельных случайных векторов  $\Xi_t$ . Много предпосылок, как теоретических, так и практических, говорит за то, что случайные векторы  $\Xi_t$  можно считать подчиняющимися многомерному нормальному распределению. С учетом этой предпосылки строят нормальные модели демографических объектов.

В нормальной модели однородного демографического объекта для всех векторов реализации  $\mathbf{X}$  и всех пар  $t', t'' \in T$  выполняется следующее соотношение:

$$|f_{t'}(\mathbf{X}, \mu_{t'}, \Sigma_{t'}) - f_{t''}(\mathbf{X}, \mu_{t''}, \Sigma_{t''})| = 0. \quad (9)$$

Символ  $f_t(\mathbf{X}, \mu_t, \Sigma_t)$  означает функцию плотности многомерного нормального распределения случайной переменной  $\Xi_t$ , определяемую формулой

$$f_t(\mathbf{X}, \mu_t, \Sigma_t) = 2\pi^{-m/2} |\Sigma_t|^{-1/2} \times \exp \left\{ -\frac{1}{2} (\mathbf{X} - \mu_t)' \Sigma_t^{-1} (\mathbf{X} - \mu_t) \right\}. \quad (10)$$

Параметрами распределения служат вектор математических ожиданий

$$\mu_t = [\mu_{t1}, \dots, \mu_{tj}, \dots, \mu_{tm}], \quad (11)$$

где  $\mu_{tj}$  — математическое ожидание одномерной случайной переменной  $\xi_j$  в период  $t$ :

$$\mu_{tj} = E(\xi_{tj}), \quad (12)$$

и матрица дисперсий и ковариаций

$$\Sigma_t = \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1j} & \dots & \sigma_{1m} \\ \vdots & & \vdots & & \vdots \\ \sigma_{i1} & \dots & \sigma_{ij} & \dots & \sigma_{im} \\ \vdots & & \vdots & & \vdots \\ \sigma_{m1} & \dots & \sigma_{mj} & \dots & \sigma_{mm} \end{bmatrix}, \quad (13)$$

причем

$$\sigma_{ij} = E[(\xi_i - \mu_i)' (\xi_j - \mu_j)] \quad (14)$$

означает для  $i = j$  дисперсию одномерной случайной переменной  $\xi_i$  в период  $t$ , а для  $i \neq j$  ковариацию  $\xi_i$  и  $\xi_j$  в этом периоде.

Поскольку параметры  $\mu_t$  и  $\Sigma_t$  однозначно определяют нормальное распределение, условие (9) равнозначно двум рядам:

$$\mu_{t'} = \mu_{t''}; \quad (15)$$

$$\Sigma_{t'} = \Sigma_{t''} \quad (16)$$

для всех  $t', t'' \in T$ . Таким образом, проблема изучения однородности пространства  $\Xi^T$  сводится с математической стороны к проверке ряда нулевых гипотез об эквивалентности двух многомерных распределений<sup>2</sup>.

<sup>2</sup> Правила статистического вывода, касающиеся такого рода гипотез, можно найти в монографиях по многомерной статистике [1], [12].



В нормальной модели неоднородного демографического объекта возникает неравенство

$$|f_{t'}(\mathbf{X}, \mu_{t'}, \Sigma_{t'}) - f_{t''}(\mathbf{X}, \mu_{t''}, \Sigma_{t''})| > 0, \quad (17)$$

по крайней мере, для одной пары  $t', t'' \in T$  и одного вектора реализации  $\mathbf{X}$ . Это условие можно задать также требованием, чтобы возникало, по крайней мере, одно неравенство типа

$$\mu_{t'} \neq \mu_{t''}. \quad (18)$$

Анализ нормальных демографических объектов хотя и возможен теоретически, но очень сложен, если иметь в виду большую трудоемкость вычислений. В связи с этим будет рассмотрен частный случай, в котором исходят из того, что исследуемые демографические процессы можно охарактеризовать с помощью сферически нормальной модели<sup>3</sup>. Это значит, что все случайные векторы  $\Xi_t$ , образующие пространство  $\Xi^T$ , имеют многомерные сферически нормальные распределения  $N(\mathbf{X}, \mu_t, \sigma^2 \mathbf{I})$  с функцией плотности, определяемой формулой

$$f_t(\mathbf{X}, \mu_t, \sigma^2 \mathbf{I}) = 2\pi^{-m/2} |\sigma^2 \mathbf{I}|^{-1/2} \times \\ \times \exp\left\{-\frac{1}{2}(\mathbf{X} - \mu_t)' \sigma^2 \mathbf{I} (\mathbf{X} - \mu_t)\right\}, \quad (19)$$

где, кроме ранее поясненных символов,  $\sigma^2$  означает дисперсию, одинаковую для всех составляющих случайного вектора  $\Xi_t$ , а  $\mathbf{I}$  — единичная матрица размерности  $m \times m$ .

Запись условий однородности или неоднородности сферически нормальной модели демографического объекта — *mutatis mutandis* — аналогична условиям, выраженным соотношениями (9) и (17). Изучение этих свойств сводится при таком положении исключительно к проверке того, справедливы ли соотношения (15) или (18). Однако предполагается, что соотношение (16) всегда выполняется.

Исходя из сформулированных дефиниций предлагается выделять фазы развития в демографических процессах согласно следующему алгоритму. Исходная матрица информации  $\mathbf{X}$  делится  $n - 1$  раз на две части, и каждый раз получают две многомерные случайные выборки, содержащие соответственно  $n_1$  и  $n_2$  наблюдений, причем  $n_1 + n_2 = n$ . Далее в отношении полученных выборок проверяется ряд

<sup>3</sup> В настоящее время ведутся дальнейшие исследования в направлении устранения этого ограничения.

нулевых гипотез об эквивалентности представленных этими выборками распределений.

Принимая, что рассматриваются сферически нормальные модели демографических объектов, проверяемый ряд нулевых гипотез можно выразить в виде

$$H_0 : \mu_{t'} = \mu_{t''} \quad (t', t'' \in T). \quad (20)$$

Гипотеза  $H_0$  проверяется с помощью обобщенной статистики Хотеллинга, выраженной формулой

$$T^2 = \frac{n_1 \cdot n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2), \quad (21)$$

которая при условии справедливости гипотезы  $H_0$  имеет распределение  $\chi^2$  с  $m$  степенями свободы (см. [15]).

В формуле (21)  $\bar{X}_1$  и  $\bar{X}_2$  означают векторы средних арифметических сравниваемых случайных выборок, а  $S^{-1}$  — это матрица, обратная матрице дисперсий и ковариаций, вычисленной на основе всех выборок.

Если окажется, что вычисленное значение критерия  $T^2$  превысит критическое значение, взятое из таблиц распределения  $\chi^2$  при принятом уровне значимости  $\alpha$  и  $m$  степенях свободы, то гипотезу  $H_0$  следует отклонить. В противном случае оснований для отклонения проверяемой нулевой гипотезы нет; однако это вовсе не означает, что она должна быть принята.

В случае если последняя ситуация возникнет для всех возможных разбиений матрицы  $X$ , то можно утверждать, что пространство  $E^T$  статистически однородно и в развитии рассматриваемого демографического процесса нет отдельных фаз, поскольку весь анализируемый период представляет собой единую однородную фазу развития.

Если же хотя бы в одном случае гипотезу  $H_0$  следует отклонить, то пространство  $E^T$  рассматривается как статистически неоднородное, что, в свою очередь, позволяет разделить его, по крайней мере, на две более однородные части. Разбиение производится на основе максимального значения статистики  $T^2$ , поскольку она служит также обобщенной мерой отдаленности двух многомерных случайных выборок.

Далее аналогичным способом проверяется однородность каждой полученной при разбиении подматрицы, и в случае, если она оказывается неоднородной, выделяется еще одна фаза развития. Эта процедура продолжается до тех пор, по-

ка все подматрицы не будут представлять статистически однородные выборки.

Дихотомический характер разбиения матрицы  $X$  в определенных ситуациях может привести к выделению фаз развития, которые после рассмотрения их вместе по окончании процесса разбиения не будут значимо отличаться одна от другой в смысле критерия (18). Поэтому на заключительном этапе процедуры периодизации в отношении случайных выборок, предварительно выделенных и смежных во времени, вновь проверяется ряд гипотез (20) об идентичности распределений совокупностей, из которых взяты эти выборки.

Если в каком-либо случае нет оснований отклонить<sup>4</sup> гипотезу  $H_0$ , то соответствующие выборки, а вместе с ними и предварительно установленные фазы развития объединяются и разделяющая их граница считается фиктивной (несуществующей). Такой порядок действий применяется до тех пор, пока все соседние подпериоды не будут значимо отличаться один от другого.

Полученные в конечном счете значения эмпирических статистик  $T^2$ , будучи обобщенными мерами отдаленности выделенных фаз развития в  $m$ -мерном пространстве признаков, вместе с тем служат измерителями надежности разделяющих их границ.

### 3. ПОПЫТКА ПЕРИОДИЗАЦИИ ДЕМОГРАФИЧЕСКИХ ПРОЦЕССОВ В ПОЛЬШЕ ЗА ТРИДЦАТИЛЕТНИЙ ПЕРИОД

Представленные в предыдущем разделе методы были применены в исследовании развития демографических процессов, происходящих в Польше после второй мировой войны. Временной диапазон анализа охватывает 1946—1973 гг. Исследование проведено в трех вариантах: в целом по всей стране и отдельно для города и села.

Отправным пунктом нашего исследования было установление широкого набора переменных, которые по возможности всесторонне характеризуют важнейшие демографические явления. Принимая во внимание относительно большую длительность анализируемого периода и необходимость получения идентичных групп признаков для всех выделен-

---

<sup>4</sup> В целях проверки гипотезы  $H_0$  на этом этапе применяются те же статистики и критерии, что и ранее.

ных вариантов, в состав статистических материалов, которые можно получить из общедоступных источников, мы включаем только 12 показателей:

число браков на 1000 населения ( $x_1$ );

число рожденных живыми на 1000 населения ( $x_2$ );

число смертей на 1000 населения ( $x_3$ );

естественный прирост ( $x_4$ );

число младенческих смертей на 1000 рожденных живыми ( $x_5$ );

число разводов на 1000 населения ( $x_6$ );

число разводов на 1000 заключенных браков ( $x_7$ );

доля рожденных живыми в общем числе рождений ( $x_8$ );

доля рожденных в браке в общем числе рожденных живыми ( $x_9$ );

валовый коэффициент воспроизводства населения ( $x_{10}$ );

чистый коэффициент воспроизводства населения ( $x_{11}$ );

коэффициент демографической активности ( $x_{12}$ ).

Числовые данные почерпнуты из Демографических ежегодников ГСУ 1945—1968 гг. и 1974 г. и Статистического ежегодника ГСУ 1974 г. При этом добавим, что недостающие данные для признаков  $x_8$  и  $x_9$  за 1946—1947 гг., для признаков  $x_6$  и  $x_7$  за 1946—1948 гг. и для признаков  $x_{10}$  и  $x_{11}$  за 1946—1949 гг. были рассчитаны путем экстраполяции тенденций соответствующих переменных. Аналитический вид функций, примененных с этой целью, каждый раз выбирался из нескольких разных двухпараметрических элементарных функций на основе анализа методом наименьших квадратов стохастических свойств полученных моделей тенденций развития. Поскольку объем расчетов был не слишком велик (5% всей совокупности), все связанные с ними выкладки опущены ввиду ограниченных рамок разработки.

Собранные и дополненные упомянутым способом эмпирические материалы для Польши в целом и с разбивкой на город и село содержатся в табл. 1, 2, 3. Расчетные данные отмечены звездочкой. Приведенная ранее нумерация переменных сохранена во всех таблицах.

Затем на основе представленного числового материала сделан выбор оптимального состава признаков, непосредственно принимаемых для периодизации. В обоих применяемых с этой целью методиках — вроцлавской таксономии и вроцлавском территориальном методе — требуется определить матрицы отдаленности признаков в многомерном пространстве периодов.

Значения избранных демографических признаков в Польше  
за период 1946—1973 гг.

Таблица 1

Год	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$
1946	11,3	26,2	10,2	16,0	119,8	3,9*	37,8*	98,0*	92,4*	2,031*	1,714*	2,574
1947	13,0	28,7	10,9	17,8	114,7	4,0*	39,0*	98,2*	92,5*	1,991*	1,686*	2,636
1948	13,3	29,4	11,2	18,2	112,1	4,1*	40,1*	98,2	92,3	1,950*	1,657*	2,637
1949	11,2	29,7	11,5	18,2	108,4	4,6	40,6	98,1	92,2	1,910*	1,628*	2,585
1950	10,8	30,7	11,6	19,1	111,2	4,4	41,2	98,2	92,4	1,790	1,491	2,643
1951	10,7	31,0	12,4	18,6	117,6	4,4	41,2	98,3	92,4	1,806	1,504	2,509
1952	10,4	30,2	11,1	19,1	96,4	4,9	46,7	98,4	92,7	1,765	1,537	2,717
1953	10,0	29,7	10,2	19,5	88,4	4,9	49,0	98,6	92,6	1,751	1,526	2,923
1954	9,8	29,1	10,3	18,8	83,3	4,6	47,2	98,7	93,0	1,732	1,510	2,815
1955	9,5	29,1	9,6	19,5	82,2	4,9	51,4	98,6	93,7	1,742	1,519	3,034
1956	9,4	28,1	9,0	19,1	70,9	5,0	53,1	98,7	94,4	1,695	1,478	3,124
1957	9,1	27,6	9,5	18,1	77,2	5,5	61,1	98,7	94,7	1,687	1,472	2,907
1958	9,2	26,3	8,4	17,9	72,1	5,5	59,6	98,7	94,8	1,621	1,462	3,130
1959	9,5	24,7	8,6	16,1	71,4	5,3	55,6	98,7	95,0	1,544	1,393	2,864
1960	8,2	22,6	7,6	15,0	54,8	5,0	60,7	98,8	95,5	1,438	1,339	2,986
1961	7,9	20,9	7,6	13,3	53,2	5,6	71,2	98,8	95,7	1,364	1,270	2,756
1962	7,5	19,8	7,9	11,9	54,2	5,9	78,9	98,8	95,8	1,305	1,216	2,506
1963	7,2	19,2	7,5	11,7	48,5	6,4	88,9	98,0	95,9	1,302	1,213	2,556
1964	7,4	18,1	7,6	10,5	47,2	6,7	90,5	98,9	95,9	1,242	1,157	2,387
1965	6,3	17,4	7,4	10,0	41,4	7,5	117,7	98,6	95,5	1,217	1,149	2,351
1966	7,1	16,7	7,3	9,4	38,6	7,7	107,9	98,7	95,4	1,174	1,128	2,277
1967	7,5	16,3	7,8	8,5	37,9	8,5	113,5	98,5	95,1	1,127	1,071	2,101
1968	8,0	16,2	7,6	8,6	33,4	9,1	114,0	98,5	95,1	1,084	1,044	2,147
1969	8,3	16,3	8,1	8,2	34,4	10,1	121,1	98,5	95,0	1,065	1,013	2,021
1970	8,5	16,6	8,1	8,5	33,4	10,6	123,3	98,5	95,0	1,064	1,011	2,046
1971	8,9	17,2	8,7	8,5	29,7	11,0	124,5	98,5	95,1	1,094	1,040	1,982
1972	9,3	17,4	8,0	9,4	28,6	11,3	121,6	98,6	95,1	1,082	1,034	2,171
1973	9,4	17,9	8,3	9,6	26,1	11,7	125,0	98,7	95,3	1,094	1,055	2,159

Значения избранных демографических признаков  
для городов Польши за период 1946—1973 гг.

Таблица 2

Год	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$
1946	13,8	25,2	9,9	15,3	108,3	7,7*	66,9 <sup>1</sup>	97,7 <sup>r</sup>	91,6 <sup>r</sup>	1,812 <sup>*</sup>	1,552 <sup>r</sup>	2,542
1947	14,8	27,8	10,3	17,5	103,8	7,9*	68,9 <sup>*</sup>	97,9 <sup>r</sup>	91,8 <sup>r</sup>	1,769 <sup>*</sup>	1,570	2,687
1948	14,8	28,5	10,3	18,2	98,6	8,0 <sup>r</sup>	70,9 <sup>*</sup>	97,7	91,3	1,726 <sup>*</sup>	1,487 <sup>r</sup>	2,771
1949	13,2	28,6	10,5	18,1	96,0	8,7	73,2	97,8	91,5	1,684 <sup>*</sup>	1,455	2,735
1950	12,6	30,0	10,9	19,1	102,6	8,8	75,1	98,0	91,7	1,558	1,300	2,761
1951	12,3	30,7	11,5	19,2	106,3	9,1	72,2	98,2	91,4	1,602	1,337	2,676
1952	11,5	29,7	10,4	19,3	88,2	9,8	83,6	98,2	91,7	1,550	1,369	2,863
1953	11,4	29,2	9,4	19,8	80,1	9,7	85,1	98,5	91,9	1,527	1,349	3,098
1954	11,2	28,7	9,5	19,2	75,5	9,0	81,0	98,6	92,4	1,517	1,341	3,032
1955	10,8	28,6	8,9	19,7	73,3	9,2	85,3	98,6	93,1	1,546	1,366	3,209
1956	10,2	26,8	8,3	18,5	64,8	9,3	90,8	98,7	93,8	1,455	1,287	3,229
1957	9,9	26,0	8,6	17,4	70,0	10,4	105,4	98,7	94,2	1,441	1,276	3,028
1958	9,7	24,5	7,7	16,8	64,2	10,1	103,6	98,7	94,4	1,381	1,260	3,166
1959	9,4	22,4	7,9	14,5	64,1	9,5	101,4	98,7	94,4	1,271	1,160	2,823
1960	8,8	19,9	7,0	12,9	49,7	8,8	99,9	98,7	94,8	1,168	1,098	2,845
1961	8,5	18,1	6,9	11,2	47,0	9,6	113,3	98,8	94,9	1,072	1,008	2,617
1962	7,9	16,9	7,0	9,8	47,8	10,3	129,7	98,8	95,0	1,013	0,952	2,376
1963	7,5	16,3	6,9	9,4	44,2	10,9	144,4	98,0	95,0	1,002	0,942	2,349
1964	7,6	15,5	7,0	8,5	41,6	11,3	148,2	98,9	95,0	0,959	0,901	2,226
1965	6,6	14,9	6,9	8,0	38,7	12,7	193,0	98,5	95,0	0,925	0,879	2,162
1966	7,3	14,4	7,0	7,4	35,2	13,0	179,1	98,6	94,4	0,891	0,847	2,070
1967	7,6	14,0	7,3	6,7	35,7	14,2	186,6	98,5	94,2	0,856	0,818	1,915
1968	8,0	13,8	7,2	6,6	31,4	13,2	189,2	98,4	94,1	0,828	0,791	1,924
1969	8,3	14,1	7,7	6,4	33,1	16,7	201,1	98,4	94,1	0,820	0,783	1,840
1970	8,7	14,7	7,7	7,0	31,6	17,3	201,7	98,4	94,3	0,832	0,794	1,909
1971	8,9	15,0	8,1	6,9	29,0	17,7	200,9	98,4	94,4	0,848	0,810	1,862
1972	9,3	15,6	7,7	7,9	28,5	17,9	193,9	98,5	94,6	0,842	0,807	2,027
1973	9,3	15,9	7,8	8,1	25,7	18,3	197,9	98,6	94,6	0,839	0,805	2,031

Значения избранных демографических признаков  
для сел Польши за период 1946—1973 гг.

Таблица 3

Год	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$	$x_{11}$	$x_{12}$
1946	10,9	26,7	10,3	16,4	125,3	1,3*	13,5*	98,2*	92,8*	2,156*	1,784*	2,589
1947	12,1	29,2	11,2	18,0	120,4	1,4*	13,9*	98,3*	92,9*	2,126*	1,777*	2,612
1948	12,5	29,8	11,6	18,2	119,2	1,6*	14,3*	98,5	92,8	2,096*	1,760*	2,572
1949	10,1	30,3	12,1	18,2	115,0	1,7	16,9	98,3	92,6	2,066	1,742*	2,512
1950	9,7	31,2	12,1	19,1	116,0	1,6	16,0	98,3	92,8	1,936	1,610	2,582
1951	9,7	31,2	12,9	18,3	124,6	1,4	14,9	98,3	93,1	1,960	1,630	2,415
1952	9,6	30,6	11,6	19,0	101,7	1,5	16,2	98,4	93,3	1,937	1,670	2,629
1953	9,0	30,0	10,7	19,3	94,0	1,5	16,6	98,6	93,1	1,918	1,655	2,816
1954	8,9	29,3	10,9	18,4	88,7	1,5	16,9	98,7	93,4	1,918	1,655	2,681
1955	8,4	29,5	10,1	19,4	88,9	1,5	17,9	98,7	94,1	1,941	1,675	2,916
1956	8,7	29,1	9,5	19,6	75,4	1,5	17,4	98,7	94,8	1,932	1,668	3,051
1957	8,4	29,0	10,3	18,7	82,6	1,5	17,9	98,7	95,0	1,937	1,673	2,823
1958	8,7	27,8	9,0	18,8	78,2	1,5	17,3	98,7	95,0	1,870	1,669	3,102
1959	9,5	26,8	9,3	17,5	76,6	1,5	15,9	98,7	95,5	1,835	1,638	2,895
1960	7,7	24,9	8,0	16,9	58,5	1,5	19,9	98,8	96,1	1,731	1,601	3,097
1961	7,3	23,6	8,2	15,4	57,7	1,8	25,1	98,8	96,3	1,691	1,564	2,865
1962	7,2	22,4	8,6	13,8	58,8	1,8	25,0	98,8	96,4	1,644	1,521	2,605
1963	6,8	21,9	8,0	13,9	51,6	2,0	29,2	99,0	96,5	1,660	1,536	2,729
1964	7,2	20,5	8,1	12,4	51,3	2,2	30,6	98,9	96,5	1,589	1,470	2,519
1965	6,1	19,7	7,8	11,9	43,4	2,3	38,0	98,7	96,2	1,582	1,487	2,514
1966	7,0	19,1	7,8	11,3	41,3	2,4	33,9	98,7	96,1	1,548	1,455	2,461
1967	7,3	18,6	8,2	10,4	39,5	2,6	35,8	96,6	95,9	1,493	1,421	2,266
1968	7,9	18,7	7,9	10,8	34,9	2,7	34,4	98,6	95,8	1,424	1,347	2,358
1969	8,3	18,7	8,5	10,2	35,4	3,1	37,3	98,6	95,7	1,403	1,328	2,192
1970	8,4	18,8	8,6	10,2	34,8	3,2	36,9	98,6	95,6	1,389	1,315	2,177
1971	8,9	19,5	9,3	10,2	30,3	3,5	39,7	98,6	95,7	1,447	1,370	2,098
1972	9,3	19,5	8,4	11,1	29,2	3,7	39,7	98,7	95,6	1,439	1,369	2,318
1973	9,6	20,4	8,9	11,5	62,4	4,0	41,8	98,8	95,9	1,509	1,449	2,291

В случае двойственных алгоритмов важное значение имеет выбор соответствующей меры отдаленности. Пользуясь, например, общепринятой мерой отдаленности точек в евклидовом пространстве

$$d_{ik} = \left[ \frac{\sum_{j=1}^n (x'_{ji} + x'_{jk})^2}{n} \right]^{1/2} \quad (ik = 1, 2, \dots, m), \quad (22)$$

где  $x'_{ji}$ ,  $x'_{jk}$  — стандартизованные на 0,1 значения  $i$ -го и  $k$ -го признаков для  $j$ -го периода,  $n$  — число анализируемых периодов,  $m$  — число признаков, принятых для исследования; можно легко заметить, что описанные ранее таксономические алгоритмы выбирают не признаки, минимально коррелирующие между собой, как следовало бы ожидать, а признаки, характеризующиеся максимальной степенью взаимной корреляции с отрицательным знаком.

Это вытекает из количественного соотношения<sup>5</sup> между коэффициентом линейной корреляции  $r_{ik}$  и евклидовым расстоянием  $d_{ik}$ :

$$d_{ik} = 2(1 - r_{ik}). \quad (23)$$

Легко доказать, что поиск признаков, максимально отдаленных один от другого ( $d_{ik}^2 \rightarrow 4$ ), равнозначен выделению признаков с максимальной отрицательной корреляцией ( $r_{ik} \rightarrow -1$ ). Поэтому необходимо обратиться к другим измерителям отдаленности признаков, например:

$$d_{ik}^{(1)} = 1 - |r_{ik}|; \quad (24)$$

$$d_{ik}^{(2)} = 1 - r_{ik}^2; \quad (25)$$

$$d_{ik}^{(3)} = \sqrt{1 - r_{ik}^2}. \quad (26)$$

Эти меры приобретают максимальные значения (равные единице), если соответствующий им коэффициент линейной корреляции  $r_{ik}$  равен нулю, и минимальные значения (равные нулю), если между признаками существует полная положительная или отрицательная корреляция ( $r_{ik} = -1$  или  $r_{ik} = 1$ ). Таким образом, признаки, наиболее отдаленные друг от друга в смысле метрик (24—26), — это вместе с тем признаки, минимально коррелирующие один с другим.

В эмпирических исследованиях мы воспользовались всеми тремя названными измерителями отдаленности. По-

<sup>5</sup> Более подробные замечания по этому поводу можно найти в работе [9].



Матрица отдаленности признаков для Польши  
(метрика (26))

$x_i$	1	2	3	4	5	6	7	8	9	10	11	12
<b>1</b>	0,0000	0,6654	0,5220	0,7480	0,5753	0,8731	0,7266	0,8237	0,4837	0,6017	0,6332	0,9554
<b>2</b>	0,6654	0,0000	0,4954	0,1734	0,3896	0,5559	0,3381	0,9475	0,5508	0,3215	0,3486	0,6638
3	0,5220	0,4954	0,0000	0,6385	0,4426	0,8044	0,6587	0,8278	0,3056	0,5537	0,6156	0,9461
<b>4</b>	0,7480	0,1734	0,6385	0,0000	0,4874	0,5044	0,3113	0,9764	0,6698	0,3701	0,3634	0,5251
5	0,5753	0,3896	0,4426	0,4874	0,0000	0,5183	0,3693	0,8409	0,4473	0,2297	0,2992	0,8225
6	0,8731	0,5559	0,8044	0,5044	0,5183	0,0000	0,3173	0,9688	0,8033	0,4593	0,4315	0,5959
7	0,7266	0,3381	0,6587	0,3133	0,3693	0,3173	0,0000	0,9524	0,6735	0,2803	0,2637	0,5965
8	0,8237	0,9475	0,8278	0,9764	0,8409	0,9688	0,9524	0,0000	0,7935	0,8932	0,9120	0,9915
9	0,4837	0,5508	0,3056	0,6698	0,4473	0,8033	0,6735	0,7935	0,0000	0,5308	0,5834	0,9495
10	0,6017	0,3215	0,5557	0,3701	0,2297	0,4593	0,2803	0,8932	0,5398	0,0000	0,0905	0,7138
11	0,6332	0,3486	0,6136	0,3634	0,2992	0,4313	0,2637	0,9120	0,5834	0,0905	0,0000	0,6726
12	0,9558	0,6634	0,9461	0,5251	0,8225	0,5059	0,5965	0,9915	0,9425	0,7138	0,6726	0,0000

**Матрица отдаленности для городов Польши**  
(метрика (24))

$x_i$	1	2	3	4	5	6	7	8	9	10	11	12
1	0,0000	0,1636	0,0898	0,2236	0,1011	0,4260	0,2280	0,2976	0,0867	0,1061	0,1239	0,4692
2	0,1636	0,0000	0,1378	0,0097	0,0783	0,2363	0,0810	0,6120	0,1722	0,0485	0,0553	0,1458
3	0,0898	0,1378	0,0000	0,2167	0,1043	0,4934	0,2891	0,3954	0,0419	0,1621	0,1977	0,5250
4	0,2236	0,0097	0,2167	0,0000	0,1150	0,2006	0,0675	0,7065	0,2474	0,0626	0,0611	0,0823
5	0,1011	0,0783	0,1043	0,1150	0,0000	0,1878	0,0848	0,3930	0,1097	0,0290	0,0164	0,3111
6	0,4260	0,2363	0,4984	0,2006	0,1878	0,0000	0,0664	0,7434	0,4643	0,1629	0,1506	0,1840
7	0,2280	0,0810	0,2891	0,0675	0,0848	0,0664	0,0000	0,6562	0,2911	0,0535	0,0481	0,1234
8	0,2976	0,6120	0,8354	0,7065	0,3930	0,7434	0,6562	0,0000	0,2781	0,4797	0,5121	0,9966
9	0,0867	0,1722	0,0419	0,2474	0,1097	0,4643	0,2911	0,2781	0,0000	0,1574	0,1865	0,5194
10	0,1061	0,0485	0,1621	0,0626	0,0290	0,1629	0,0535	0,4797	0,1574	0,0000	0,0030	0,1998
11	0,1239	0,0553	0,1977	0,0611	0,0464	0,1506	0,0181	0,5121	0,1865	0,0030	0,0000	0,1748
12	0,4692	0,1458	0,5250	0,0823	0,1131	0,1840	0,1234	0,9966	0,5194	0,1998	0,1748	0,0000

Матрица отдаленности для сел Польши  
(метрика (25))

$x_i$	1	2	3	4	5	6	7	8	9	10	11	12
1	0,0000	0,6579	0,4527	0,7816	0,5332	0,9498	0,7200	0,5031	0,3720	0,0028	0,7118	0,9990
2	0,6579	0,0000	0,2682	0,0495	0,1643	0,3517	0,1318	0,7568	0,3191	0,1373	0,2993	0,6581
3	0,4527	0,2682	0,0000	0,4833	0,2274	0,7376	0,4815	0,4320	0,1355	0,3923	0,5495	0,9932
4	0,7810	0,0495	0,4833	0,0000	0,2781	0,2591	0,1000	0,8855	0,4982	0,1682	0,1801	0,4385
5	0,5333	0,1643	0,2274	0,2782	0,0000	0,4528	0,1588	0,5106	0,1912	0,0800	0,1772	0,8483
6	0,9498	0,3517	0,7376	0,2591	0,3528	0,0000	0,1409	0,9230	0,7128	0,2634	0,1912	0,4322
7	0,7200	0,1218	0,4815	0,1000	0,1588	0,1409	0,0000	0,8010	0,4578	0,1007	0,1065	0,5361
8	0,5031	0,7569	0,4320	0,8855	0,5106	0,9230	0,8019	0,0000	0,3348	0,6989	0,8059	0,9483
9	0,3620	0,3191	0,1355	0,4982	0,1012	0,7128	0,4578	0,3348	0,0000	0,3225	0,4722	0,9352
10	0,6028	0,1373	0,3923	0,1682	0,0800	0,2634	0,1007	0,6989	0,3225	0,0000	0,0341	0,6762
11	0,7118	0,2092	0,5495	0,1801	0,1772	0,1912	0,1065	0,8059	0,5722	0,9841	0,0000	0,5570
12	0,9990	0,6581	0,9932	0,4385	0,8483	0,4322	0,5361	0,9488	0,9852	0,6762	0,5570	0,0000

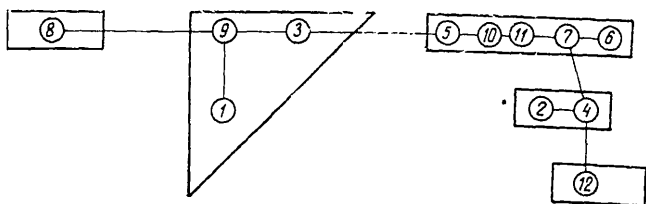


Рис. 1. Вроцлавский дендрит: 12 признаков, характеризующих демографическое развитие Польши. Масштаб: 1 см=0,2

сколько полученные результаты оказались близкими, в работе помещены только избранные, наиболее характерные результаты. Для признаков, описывающих развитие демографических явлений в городе, представлены результаты, полученные с помощью метрики (24), для села — с помощью метрики (25), и, наконец, для данных, относящихся ко всей Польше, выбраны результаты, полученные в соответствии с метрикой (26). Построенные на основе числовых данных, содержащихся в табл. 1, 2 и 3, симметрические квадратные матрицы отдаленности признаков приведены отдельно для каждого варианта исследований в табл. 4, 5 и 6.

Результаты вроцлавской таксономии представлены в виде кратчайших дендритов (рис. 1, 2 и 3), которые соединяют все рассматриваемые признаки. На рисунках обозначены также предлагаемые разбиения дендритов на группы сходных признаков, полученные путем последовательного отсечения нескольких наиболее длинных ветвей дендрита или на основе максимальных значений показателей силы разбиения Е. Перкаля.

Окончательные результаты разбиения признаков на более однородные группы, полученные с помощью вроцлавского территориального метода, содержатся в табл. 7.

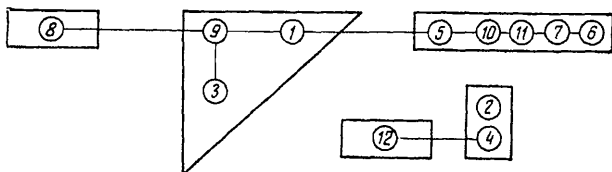


Рис. 2. Вроцлавский дендрит: 12 признаков, характеризующих демографическое развитие городов в Польше. Масштаб: 1 см=0,05

Последний столбец табл. 7 содержит значения радиусов гиперсфер, принятые в ходе дискриминантного анализа множества признаков и установленные на основе соотношения<sup>6</sup>

$$\rho = \bar{d} - 0,5 s_d, \quad (27)$$

где  $\bar{d}$  — средняя арифметическая минимальных значений измерителей отдаленности в отдельных строках (столбцах) матрицы отдаленности, а  $s_d$  — стандартное отклонение этих значений.

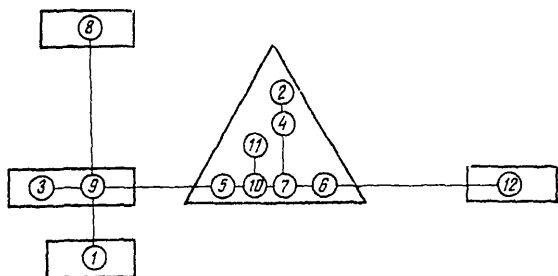


Рис. 3. Вроцлавский дендрит: 12 признаков, характеризующих демографическое развитие села в Польше. Масштаб: 1 см=0,1

Рассматривая полученные результаты, следует обратить внимание на большое сходство структуры признаков во всех трех вариантах исследования. Такое положение соответствует ожиданиям, поскольку нам представляется, что в первом приближении, если какой-либо признак обладает свойством диагностичности, то это свойство имеет общий характер, присущий разным составным частям данной социально-экономической системы. Определенные же различия должны проявляться только в ходе более детального изучения.

Наблюдаемое сходство структуры признаков — благоприятное обстоятельство для целей дальнейших исследований. В силу этого возникает возможность проведения полностью обоснованного сравнительного анализа закономерностей, возникающих в развитии демографических явлений, между городом и селом. Принятие для периодизации

<sup>6</sup> Ср. работу [2].

Результаты таксономии признаков, полученные  
с помощью вrocławского территориального метода

Вариант	Группы признаков	Радиус $\rho$
1. Польша	(1) (6) (8) (12) (2, 3, 4, 5, 7, 9, 10, 11)	0,5068
2. Город	(8) (12) (1, 3, 9) (2, 4, 5, 6, 7, 10, 11)	0,0943
3. Село	(1) (6) (8) (12) (2, 3, 4, 5, 7, 9, 10, 11)	0,2927

разных групп признаков серьезно затруднило бы такого рода сравнения.

Учитывая общие результаты таксономического анализа признаков, показанные на рис. 1—3 и в табл. 7, для дальнейших исследований было решено выбрать два набора признаков, общих для всех трех вариантов анализа. Набор *A* содержит признаки, обозначенные символами  $x_1, x_4, x_7, x_{12}$ , а набор *B* — признаки  $x_1, x_4, x_6, x_8, x_{12}$ . Признаки  $x_4$  и  $x_7$  представляют группы однородных признаков, состоящие из нескольких элементов, а остальные признаки в большинстве случаев представляют собой одноэлементные группы. Два несколько отличающихся набора диагностических признаков принимаются для того, чтобы избежать опасности получить случайные результаты анализа только для одного варианта.

Установление оптимальных векторов признаков завершает этап предварительных рассуждений и позволяет перейти к периодизации. Описанная ранее процедура выделения фаз развития требует, ввиду большой трудоемкости вычислений, применения электронной вычислительной техники. Все необходимые расчеты были проведены с помощью вычислительной машины «Минск-32» на основе специально составленной для этих целей программы<sup>7</sup> на языке Мифор<sup>8</sup>.

Поскольку оказалось, что общепринятый в социальных исследованиях уровень значимости  $\alpha = 0,05$  приводит в итоге к слишком детальному членению анализируемого периода, вычисления были повторены при измененном уровне значимости  $\alpha = 0,01$ . Это привело к сокращению числа фаз

<sup>7</sup> Точное описание алгоритма и программа приведены в работе [6].

<sup>8</sup> Язык Мифор — это упрощенный вариант стандартного языка Фортран 1900; он был разработан Академией Генерального Штаба Войска Польского для применения в машинах «Минск».

Фазы развития демографических процессов в Польше в период 1946—1973 гг.  
(для уровня значимости  $\alpha=0,01$ )

Год	Всего						Город						Село					
	Вариант А			Вариант В			Вариант А			Вариант В			Вариант А			Вариант В		
	$T^2$	$T'$	$R$	$T^2$	$T'$	$R$	$T^2$	$T'$	$R$	$T^2$	$T'$	$R$	$T^2$	$T'$	$R$	$T^2$	$T'$	$R$
1946																		
1947																		
1948																		
1949																		
1950																		
1951							26,64	2,01	3	28,05	1,86	2						
1952	28,40	2,14	3	37,81	2,50	3							25,40	1,91	2	29,93	1,09	2
1953																		
1954																		
1955										15,22	1,01	1						
1956							20,21	1,52	2									
1957																		
1958																		





Фазы развития демографических процессов в Польше в период 1946—1973 гг.  
(для уровня значимости  $\alpha=0,05$ )

Год	Всего						Город						Село					
	Вариант А			Вариант В			Вариант А			Вариант В			Вариант А			Вариант В		
	$T^2$	$T'$	R	$T^2$	$T'$	R	$T^2$	$T'$	R	$T^2$	$T'$	R	$T^2$	$T'$	R	$T^2$	$T'$	R
1946																		
1947							11,21	1,18	1									
1948	10,57	1,11	1	13,90	1,25	1				13,24	1,25	1	13,24	1,40	1			
1949																		
1950																		
1951							20,02	2,11	3	19,21	1,73	2						
1952	18,49	1,95	3	18,00	1,63	3										29,93	2,70	3
1953																		
1954													19,12	2,01	3			
1955				14,41	1,30	1	10,24	1,08	1	15,22	1,37	1						
1956	11,61	1,22	1															
1957																		
1958				14,57	1,32	1	9,77	1,03	1	11,56	1,24	1	10,59	1,12	1			



развития, причем каждая из них характеризовалась большей степенью значимости, чем прежде. Таким образом, при уровне значимости  $\alpha = 0,01$  получена картина наиболее общих закономерностей в демографическом развитии с разбивкой на город и село. Однако принятие уровня значимости  $\alpha = 0,05$  позволяет получить более детальные сведения.

Такой порядок действий близок к методике, применяемой при таксономическом анализе скоплений, так называемом кластерном анализе<sup>9</sup>. Отсюда следует, что уровень значимости и связанная с ним критическая величина статистики  $T^2$  определяют порог подобия смежных фаз развития, на основе которого принимаются решения о разбиении рассматриваемой совокупности периодов на более однородные подпериоды.

Сводные результаты периодизации для всех изучаемых совокупностей (Польши, города и села и обоих вариантов анализа) — *A*: признаки  $x_1, x_4, x_7, x_{12}$  и *B*: признаки  $x_1, x_4, x_6, x_8, x_{12}$  — представлены в табл. 8 (для уровня значимости  $\alpha = 0,01$ ) и в табл. 9 (для уровня значимости  $\alpha = 0,05$ ).

Каждая выделенная в процессе установления фаз развития граница характеризуется с помощью трех чисел. Первое из них — эмпирическая статистически значимая величина критерия  $T^2$ , полученная на последнем этапе периодизации, проверяющем правильность предварительно полученного разбиения всего периода времени на однородные части. Таким образом, статистики  $T^2$  определяют отдаленность в многомерном пространстве смежных подпериодов, в связи с чем их можно рассматривать как измерители надежности отдельных границ.

Величину  $T^2$  можно применять для оценки надежности соответствующей ей границы только в пределах данного варианта, так как абсолютные значения этого параметра зависят от принятого уровня значимости  $\alpha$  и от числа анализируемых признаков ( $m$ ). Одна и та же величина статистики  $T^2$ , равная, например, 17,00, может в одном случае означать границу, вполне значимую, если  $\chi_{0,05(4)}^2 = 9,48$ , а в другом — менее значимую, для  $\chi_{0,01(5)}^2 = 15,09$ . Поскольку в проведенном исследовании вариант *A* отличается от варианта *B* не только комплексом принятых для периодизации признаков, но и их числом, а к тому же анализ каждого

<sup>9</sup> Ср., например, работы [5] и [16].

из них проводился при разных уровнях значимости, то для того, чтобы сделать возможным сравнение надежности отдельных границ, следовало создать измерители, не зависящие от изменения числа признаков и уровня значимости.

Простейшую нормализованную меру надежности границ можно получить, разделив значения статистики  $T^2$  на соответствующее ей критическое значение  $m\chi_{\alpha}^2$ , тогда

$$T' = \frac{T^2}{m\chi_{\alpha}^2}. \quad (28)$$

Измерители  $T'$  приводятся в табл. 8 и 9 рядом со статистиками  $T^2$ , причем критические значения  $m\chi_{\alpha}^2$ , взятые из таблиц распределения  $\chi^2$ , составили для варианта *A* 9,488 (уровень значимости  $\alpha = 0,05$ ) или 13,277 (уровень значимости  $\alpha = 0,01$ ), а для варианта *B* — 11,070 ( $\alpha = 0,05$ ) или 15,086 ( $\alpha = 0,01$ ).

Последняя из трех величин, характеризующих выделенные границы, — это ранг, присвоенный им в зависимости от величины измерителя  $T'$ . Ранжирование величины параметра  $T'$  обосновано тем, что непрерывный характер этой меры может обещать большую точность оценки надежности границы. Однако, принимая во внимание стохастический характер алгоритма периодизации, возможные ошибки в исходном числовом материале или, наконец, то, что мог быть выбран не вполне оптимальный вектор признаков, мы полагаем, что о существенной дифференциации надежности выделенных границ может свидетельствовать только заметное расхождение в значениях нормализованных измерителей  $T'$ .

Итак, в последующем анализе принято, что из всех границ будут выделены три их вида. Границам малого значения, для которых величина  $T'$  не превышает 1,5, присвоен ранг 1. Границы со средней степенью надежности, заключенные в интервале  $T' \in (1,5 - 2,0)$ , обозначены рангом 2. И наконец, границам, весьма значимым (для величин  $T'$  свыше 2,0), присвоен ранг 3.

Переходя к анализу полученных результатов по существу, следует обратить внимание на следующие моменты.

1. В обоих вариантах исследования *A* и *B* получены близкие результаты, особенно для границ с высокой степенью значимости. Некоторые небольшие различия проявляются, скорее, для границ, мало значимых, и их можно объяснить различием принятых при периодизации наборов признаков. Обнаруженное общее соответствие фаз

развития, несомненно, укрепляет обоснованность выводов о закономерностях развития анализируемых демографических процессов.

2. В развитии демографических процессов в Польше можно выделить три отдельные фазы. Первая из них продолжалась до 1952 г., вторая закончилась на рубеже шестидесятых годов и, наконец, третья началась в 1966 г. Срединная граница в отличие от двух остальных выражена не очень резко и имеет характер переходного периода. Более детально анализируя отдельные фазы, можно выделить в первой из них два подпериода с граничной датой — 1948 г., а в третьей — 1971 г., в котором намечился новый период развития.

3. Интересно соотнести установленные фазы демографического развития Польши с результатами подобных исследований, касающихся экономического развития страны (см. [7], [6]). Оказывается, что общепринятые граничные даты — 1953, 1958 и 1964 гг. — совпадают с границами, полученными в данной работе, только в общих чертах. Это может свидетельствовать о довольно слабой связи между формированием демографических и экономических процессов, хотя наличие такой связи в свете нынешних исследований нельзя исключить. Быть может, это обусловлено отсутствием четкой согласованности изменений, происходящих в развитии экономических и демографических явлений.

4. При сравнении полученных фаз развития для города и села напрашивается вывод о возрастающем сходстве протекания исследуемых процессов в обоих совокупностях, особенно в последние 10—15 лет анализируемого периода. Некоторые различия можно установить только в первом пятнадцатилетии после войны. Возрастающее сходство в характере анализируемых демографических явлений, несомненно, связано с постепенным стиранием различий между городом и селом. Это вполне соответствует общепринятому представлению о том, что село перенимает все городские обычаи.

5. Сравнительно большое число существенно различающихся фаз развития, составляющее в среднем 7 при уровне значимости  $\alpha = 0,05$ , что дает в среднем по 4 года на каждую выделенную фазу, свидетельствует о динамичности изменений, происходящих в развитии демографических явлений в Польше после второй мировой войны. Такая ситуация создает серьезные трудности в области долгосрочного прогнозирования отдельных демографических пере-

менных. Прогноз формирования уровней этих переменных, сделанный путем экстраполяции прежней тенденции развития (тренда), даже в случае подтверждения ее плавности, в условиях частых, отчетливо вырисовывающихся изменений в развитии всей совокупности может быть чреват большой ошибкой. Такое положение вещей вызывает серьезные расхождения между величинами, установленными на основе демографических прогнозов, и фактическим состоянием.

6. Поскольку последняя отчетливая граница во всех рассматриваемых совокупностях отмечена в 1966 г., представляется, что при проведении краткосрочных прогнозов и другого рода исследований во временном аспекте достаточно и даже более оправданно опираться на данные, относящиеся к периоду после 1966 г. Обращение к более ранним материалам не рекомендуется, поскольку это означало бы, что рассматриваемая статистическая совокупность неоднородна. В связи с этим трудно было бы найти в ней какие-либо закономерности в силу существования многих разнонаправленно действующих систематических компонентов.

7. Принятие в исследовании двух уровней значимости  $\alpha = 0,01$  и  $\alpha = 0,05$ , с одной стороны, облегчило, а с другой — значительно обогатило проведенный анализ. Следует заметить, что детальные результаты периодизации, полученные на уровне значимости  $\alpha = 0,05$ , подтвердились и на более высоком уровне  $\alpha = 0,01$  с точки зрения как способа разбиения исследуемого периода на фазы развития, так и надежности отдельных границ.

8. Выделенные границы фаз развития должны быть подвергнуты более детальному и всестороннему качественному анализу как в демографическом, так и в социально-экономическом аспектах. Это имеет целью уловить те факторы, которые определяют изменения в формировании демографических явлений в каждой из выделенных фаз.

9. Дальнейшие исследования должны дать возможность определить продолжительность периодов опережения экономического развития, необходимых для того, чтобы проявились существенные изменения в демографическом развитии.

## ЛИТЕРАТУРА

1. Anderson T. W. An Introduction to Multivariate Statistical Analysis. New York — London, 1958.
2. Bukietyński W., Hellwig Z., Królik U., Smoluk A. Uwagi o dyskryminacji zbiorów skończonych. *Prace Naukowe WSE*, Wrocław, z. 21, (1969).

3. Fierich J. Metody taksonomiczne rejonizacji rolnictwa na przykładzie województwa Krakowskiego. *Mysł Gospodarcza*, № 1, (1957).
4. Florek K., Łukasiewicz J., Perkal J. Steinhauß H., Zubrzycki S. Taksonomia wrocławska. *Przegląd Antropologiczny*, t. XVII, Poznań, 1951.
5. Goower J. C., Ross G. J. S. Minimum Spanning Trees and Single Linkage Cluster Analysis. *Applied Statistics*, № 1, (1969).
6. Grabiński T. Algorytm klasyfikacji zbiorów liniowo uporządkowanych, Zeszyty Naukowe AE. Kraków, 1976.
7. Grabiński T. Periodyzacja rozwoju gospodarki narodowej oraz jej podstawowych sfer w okresie Polski Ludowej (studium statystyczne), Zeszyty Naukowe AE. Kraków, 1976.
8. Grabiński T., Zając K. Metody periodyzacji rozwoju demograficznego ludności miejskiej, referat na Sympozjum w Karpaczu, 16—21. IX 1974.
9. Jasiński R. Uwagi o współczynniku korelacji. *Prace Naukowe WSE*, Wrocław, № 33 (1972).
10. Kula W. Problemy i metody historii gospodarczej. Warszawa, 1963, s. 378.
11. Kuznets S. Statistical and Economic History. *Journal of Economic History*, vol. 1, 1941.
12. Morrison W. M. The Multivariate Analysis. New York, 1966.
13. Perkal J. Taksonomia wrocławska. *Przegląd Antropologiczny*, t. XIX, Poznań, 1953.
14. Pluta W. O pewnej metodzie klasyfikacji przedsiębiorstw. *Przegląd Statystyczny*, z. 1 (1972).
15. Rao C. R. Linear Statistical Inference and its Applications. New York — London — Sydney, 1964.
16. Shepard M. J., Willmat A. J. Cluster Analysis on the Atlas Computer. *Computer Journal*, № 11 (1968).
17. Stęczkowski J. Zasady i metody rejonizacji produkcji rolniczej. Warszawa, 1966.

*С. Митра, А. Романюк*

## КРИВАЯ ПИРСОНА ПЕРВОГО ТИПА И ВОЗМОЖНОСТЬ ЕЕ ПРИМЕНЕНИЯ ДЛЯ ПРОГНОЗА РОЖДАЕМОСТИ

S. Mitra, A. Romaniuk. Pearsonian Type I Curve and its Fertility Projection Potentialis. *Demography*, vol. 10, № 3, August 1973, p. 351—365.

### ВВЕДЕНИЕ

Статистики и демографы проявляют значительный интерес к кривой Пирсона первого типа, рассматривая ее в первую очередь как средство выравнивания возрастных коэффициентов рождаемости [12], [6], [4], [1]. Хотя некоторые авторы и указывают на возможность ее применения для перспективного расчета чисел рождений [10], [11], этот аспект, насколько нам известно, пока еще никем систематически не был изучен. Настоящая статья преследует две цели: во-первых, упрощение способа оценки параметров уравнения кривой Пирсона первого типа при выравнивании возрастных коэффициентов рождаемости; во-вторых, изучение возможности применения этой кривой в качестве основы при построении параметрической модели для перспективных расчетов чисел рождений. Эти две цели тесно связаны между собой; по сути дела, первая из них представляет собой предпосылку для второй. Однако модель, пригодная для целей выравнивания, может быть неподходящей для перспективных расчетов. Например, моменты более высокого порядка могут улучшить качество выравнивания, но они зачастую непригодны для прогнозов ввиду



трудности интерпретации их в демографическом смысле. Одна из первоочередных целей демографов состоит в разработке модели перспективного расчета с достаточно широкими аналитическими возможностями и вместе с тем удобной для вычислений. Однако для достижения этой цели существенно важна разработка методики, которая основывалась бы на ограниченном количестве демографически значимых параметров рождаемости и вместе с тем давала бы возможность удовлетворительно воспроизводить ряд возрастных коэффициентов рождаемости. Разработка методики, удовлетворяющей этим двум критериям, составляет основную задачу настоящей статьи.

В первом разделе статьи кратко описывается кривая Пирсона первого типа и обосновывается ее выбор в качестве средства для представления ряда возрастных коэффициентов рождаемости. Второй раздел содержит краткий обзор приемов, обычно применяемых для нахождения констант в уравнении кривой Пирсона первого типа, а также изложение новых методов, разработанных для оценки этих констант. Эти новые методы проверяются в третьем разделе, с привлечением исторических данных о рождаемости в Канаде, а в последнем разделе делается попытка исследовать возможные способы применения функции кривой Пирсона первого типа для перспективного расчета чисел рождений.

### ПРЕДСТАВЛЕНИЕ И ОБОСНОВАНИЕ КРИВОЙ ПИРСОНА ПЕРВОГО ТИПА

При начале отсчета от моды уравнений кривой Пирсона первого типа имеет вид

$$y = y_0 (1 - x/a_1)^{m_1} (1 - x/a_2)^{m_2}, \quad (1)$$

где  $y_0$  есть модальная ордината и  $-a_1 \leq x \leq a_2$ .

Кроме того,

$$m_1/a_1 = m_2/a_2. \quad (2)$$

В формуле (1)  $y_0$  есть модальный коэффициент рождаемости, который в человеческих популяциях лежит обычно между 20 и 30 годами жизни:  $a_1$  и  $a_2$  совместно определяют репродуктивный интервал;  $m_1$  и  $m_2$  определяют форму кривой рождаемости. Когда  $m_1$  и  $m_2$  приближенно равны, кривая приближается к нормальному распределению, но когда  $m_1 < m_2$ , она положительно асимметрична. Современные репродуктивные типы населения как раз и харак-

теризуются положительным асимметричным распределением; это демонстрирует табл. А 1 приложения, в которой представлены параметры распределения возрастных характеристик рождаемости, полученные на основе данных для Канады. Легко заметить в качестве первого признака асимметрии, что модальный возраст неизменно меньше среднего возраста и что мера асимметрии,  $\beta_1$ , положительна на протяжении периода 1926—1969 гг.

Однако выбор какой-либо конкретной кривой Пирсона для описания характера рождаемости той или иной страны не может основываться исключительно на субъективных взглядах и на суждениях, в основу которых положены упомянутые ранее параметры. Действительно, сам К. Пирсон [3] утверждал, что критерий такого рода выбора должен основываться на вычислении величины  $k$ :

$$k = \beta_1 (\beta_2 + 3)^2 / [4 (2\beta_2 - 3\beta_1 - 6) (4\beta_2 - 3\beta_1)].$$

Пирсон рекомендовал пользоваться кривой первого типа в случае, если  $k$  имеет отрицательное значение.

Из табл. А. 1 приложения видно, что величина  $k$  для Канады имеет отрицательное значение, и, таким образом, выбор этого типа кривой в данном случае обоснован. Необходимо, однако, отметить, что от года к году  $k$  последовательно принимает все более высокие отрицательные значения. Ввиду таких сдвигов в возрастных закономерностях рождаемости Канады стоило бы испробовать другие типы кривых из системы кривых Пирсона, например «нормальную кривую» или отвечающую ей кривую второго типа для представления возрастных закономерностей рождаемости в более ранние годы и кривую переходного третьего типа для представления возрастных закономерностей рождаемости в последние годы. Если рассматривать альтернативные варианты, то, за исключением моделей, основанных на так называемом гамма-распределении, в случае сравнительно высокой и менее асимметричной кривой рождаемости, оказалась применимой [11] только модель типа модели Мазура [5].

Сравнительное изучение подбора различных кривых не входит в задачу настоящей статьи. Вместо этого мы сосредоточим наше внимание на том типе кривой, который, будучи основанным на критерии  $k$ , по-видимому, больше всего подходит для рассматриваемого периода и, как можно ожидать, будет соответствовать типу кривой рождаемости для Канады также в ближайшие годы. Испытания, прове-

денные Митрой [6] и Эвери [1] на данных о рождаемости для ряда стран мира, показали, что из разработанных до сих пор математических функций функция кривой Пирсона первого типа относится к числу наиболее пригодных для аппроксимации наблюдаемых рядов возрастных характеристик рождаемости. К сожалению, ввиду недостатка данных подбор кривой в названных работах был ограничен коэффициентами рождаемости по пятилетним возрастным группам, и вследствие этого детальный анализ согласия подобранной кривой фактическим данным оказался невозможным.

Наша следующая задача — рассмотреть различные возможные методы нахождения констант в формуле (1).

### МЕТОДЫ НАХОЖДЕНИЯ КОНСТАНТ

Наиболее часто применяется для выравнивания фактических возрастных коэффициентов рождаемости по формуле (1) метод моментов. Однако число используемых моментов зависит от того, какая конкретная методика выбрана. Согласно методике Элдертона [3] константы  $a_1$ ,  $a_2$ ,  $m_1$  и  $m_2$  в уравнении (1) вычисляются по первым четырем моментам распределения частот, т. е. по средней, дисперсии и показателям асимметрии и эксцесса. Позднее Митра [6] разработал метод, который уменьшил требуемое число моментов до двух первых, однако этот метод предполагает постоянство возрастного интервала рождаемости. Так, при установлении интервала в возрастах 15—50 лет

$$a_1 + a_2 = 35. \quad (3)$$

Если выполняются условия, определяемые уравнениями (2) и (3), то требуемое число независимых параметров сводится всего лишь к двум и решения получаются из следующих уравнений:

$$m_1 + m_2 - \mu'_1 (a_1 + a_2 - \mu'_1) / \mu_2 = 3 \quad (4)$$

и

$$m_1 = [(m_1 + m_2 + 2) / (a_1 + a_2)] \mu'_1 - 1, \quad (5)$$

где  $\mu'_1$  — средняя, вычисленная при начальной точке в возрасте 15 лет,  $\mu_2$  — дисперсия.

Затем по уравнениям (2) и (3) можно вычислить значения  $a_1$  и  $a_2$ . Что касается  $y_0$ , то можно применить следующее выражение:

$$y_0 = N/[a_1 + a_2] B(m_1 + 1, m_2 + 1) \times \\ \times [m_1^{m_1} \cdot m_2^{m_2}] / [(m_1 + m_2)^{m_1 + m_2}]. \quad (6)$$

В уравнении (6)  $N$  есть сумма возрастных коэффициентов рождаемости, а  $B$  — бета-функция. По сути дела,  $y_0$  можно рассматривать как множитель, приравнивающий суммарные итоги наблюдаемого и выравнивающего распределений.

Рассмотренный метод основывается на приравнивании первых двух моментов наблюдаемого первым двум моментам теоретического распределения. При другом варианте метод вычисления констант можно упростить, воспользовавшись первым моментом распределения — средней и какой-либо другой мерой, например модой. Для оценки моды можно применить тот или иной метод или же просто принять в качестве моды среднюю точку одногодичного возрастного интервала с максимальным значением возрастного коэффициента рождаемости. В этом случае, если заданы значение моды или  $a_1$  и точка отсчета в начале кривой, то остальные параметры могут быть получены из следующих уравнений:

$$a_2 = (a_1 + a_2) - a_1; \quad (7)$$

$$m_2 = a_2 (a_1 + a_2 - 2\mu') / [(a_1 + a_2) (\mu' - a_1)] \quad (8)$$

и

$$m_1 = (a_1/a_2)m_2. \quad (9)$$

Хотя уменьшение числа моментов и может привести к несколько меньшему согласию подобранной кривой, оно дает некоторые аналитические и вычислительные преимущества. Эти преимущества будут более подробно рассмотрены дальше; сейчас же хотелось бы просто подчеркнуть, что параметры модели должны вести себя таким образом, чтобы это согласовалось с трендом коэффициентов рождаемости. При отсутствии такой согласованности нельзя будет соотнести вариацию распределений с вариацией параметров, и модель, таким образом, вряд ли будет иметь практическую ценность. При наличии множества возможных параметров чрезвычайно сложно соотнести, например, снижение суммарного коэффициента рождаемости с изменением модального возраста, детородного периода и т. д. Нало-

жение ограничений на значения возрастного интервала рождаемости во втором методе и на значения возрастного интервала рождаемости и модального возраста в третьем методе делает интерпретацию остальных параметров сравнительно простой и, возможно, более осмысленной.

Можно предложить и другие способы вывода или совершенствования оценки констант. Один из таких способов, основанный на значениях возрастного промежутка между началом и концом деторождения, модального возраста и модального коэффициента рождаемости, также, по-видимому, приемлем и заслуживает изучения. При начале отсчета в начальной точке кривой и пользуясь отношением

$$m_1/a_1 = m_2/a_2 = (m_1 + m_2)/(a_1 + a_2) = C, \quad (10)$$

уравнение (6) для модального коэффициента рождаемости можно переписать в виде

$$y_0 = \frac{N}{(a_1 + a_2) B(Ca_1 + 1, Ca_2 + 1)} \left[ \left( \frac{a_1}{a_1 + a_2} \right)^{a_1} \left( \frac{a_2}{a_1 + a_2} \right)^{a_2} \right]^C, \quad (11)$$

где  $N$  для одногодичного распределения эквивалентно суммарному коэффициенту рождаемости. Константа  $C$ , полученная путем некоторой итерации, может быть затем введена в уравнение (10) для оценки остальных параметров.

Кейфиц [4] предложил иной подход, который можно применить для улучшения согласия подобранной кривой, — способ итерации, основанный на методе спуска. Этот метод оказался полезным при выравнивании по кривой Гомперца данных о кумулятивной рождаемости для Канады [7], и после необходимой адаптации он может быть применен к данным о рождаемости, представленным в виде возрастных распределений.

Два последних способа были упомянуты лишь для того, чтобы показать дальнейшие возможности совершенствования подбора выравнивающих кривых. Сейчас мы не можем достаточно глубоко исследовать их пригодность для прогнозирования. Вместо этого мы сосредоточили свое внимание на первых трех методах и проверили их на основе эмпирических данных; полученные результаты представлены в следующем разделе.

Для упрощения записи в дальнейшем будем обозначать эти три метода с помощью следующих кодовых наименований в круглых скобках:

а) метод, основанный на первых четырех моментах (4М);

- б) метод, основанный на первых двух моментах (2M);  
в) метод, основанный на средней и моде (1M).

Числа 4, 2 и 1 означают порядок моментов, на которых основан конкретный метод.

### ЭМПИРИЧЕСКАЯ ПРОВЕРКА МЕТОДОВ И МОДЕЛИ

Каким бы ни было теоретическое обоснование того или иного конкретного метода, всякое решение относительно его практической пригодности должно основываться на некоторой эмпирической проверке, в частности на сравнении выведенных оценок с фактическими данными. Прежде чем перейти к проверке, рассмотрим вкратце данные, на которых она основывается.

Настоящая проверка производится на данных о рождаемости по одногодичным возрастным интервалам для Канады, которыми мы располагаем начиная с 1926 г. Выбор этой страны не объясняется никакими определенными причинами, если не считать того, что один из авторов — канадский гражданин. Предполагается, что данные о рождаемости для Канады доброкачественны, хотя и не вполне свободны от ошибок регистрации. Это, вероятно, справедливо особенно для более ранних лет, которые помимо всего прочего характеризуются значительными перепадами в изменении показателей модальной возрастной группы. Логического объяснения колебаний, наблюдаемых в старших возрастах в этой возрастной группе, по-видимому, нет. В какой степени такого рода ошибки повлияли на остальные возрасты, неизвестно; никаких попыток скорректировать данные, чтобы устранить такие смещения, не предпринималось. Однако нет оснований предполагать, что величина и направление этих смещений могли сколь-либо значительно исказить характеристики распределений, на которых основываются оценки параметров.

В настоящей работе применяются только показатели рождаемости для календарного периода. Данные по когортам представляют собой рекомбинацию данных по календарным периодам, дополненных путем экстраполяции для усеченных когорт. Хотя эти когортные данные и пригодны для общего анализа, сами по себе они не представляются достаточно достоверными, чтобы на их основе можно было оценивать надежность проверяемых в настоящей работе методик.

Для статистических сопоставлений были выбраны 1926, 1931, 1941, 1951, 1961 и 1969 гг., что объясняется соображениями объема статьи. На протяжении всего этого периода наблюдались значительные колебания как уровня, так и возрастных закономерностей рождаемости. В настоящей работе оценки для каждого года в период 1926—1969 гг. представлены только для метода (4М) (см. табл. А.2 приложения). Априори можно ожидать улучшения согласия подобранной кривой по мере увеличения числа принятых в расчет моментов. Из этого следует, что метод (4М) задает некоторые стандарты, с которыми можно сопоставлять оценки, полученные другими методами.

Сравним сначала меры рождаемости и параметры кривой, полученные путем оценки, с фактическими; затем исследуем для всего ряда повозрастных коэффициентов рождаемости согласие кривых, подобранных с помощью всех трех описанных в предыдущих разделах методов.

### Оценки и фактические меры рождаемости

Из всех параметров распределения первого типа, по-видимому, наиболее значим в данном контексте модальный возраст. Поскольку метод (2М) ограничен постоянным интервалом рождаемости, оценка модального возраста не будет однозначной, поскольку она зависит от конкретного выбора этого интервала. Результаты, основанные на двух альтернативных интервалах рождаемости (15—50 и 17—50 лет), полученные методом (2М), а также результаты, вычисленные методом (4М), сопоставляются в табл. 1 с фактическими модальными возрастными. Следует отметить, что наблюдавшиеся значения, представленные в таблице, были скорректированы, чтобы устранить перепады, вызванные, по-видимому, случайными колебаниями и неправильным указанием возраста. Эта корректировка была сделана проведенном от руки плавной кривой через точки наблюдавшихся модальных возрастов. Следует также напомнить, что метод (1М) не предусматривает в ходе расчета вычисления модального возраста и что значения последнего должны быть заданы, чтобы мы могли иметь возможность пользоваться этим методом для нахождения прочих параметров. Методы (4М) и (2М) при возрастном интервале 17—50 лет дают достаточно хорошую оценку модального возраста, тогда как метод (2М) при интервале рождаемости 15—50 лет имеет тенденцию преувеличивать его почти на год.

Оценки модального возраста для Канады  
с применением методов (4М), (2М) и (1М), 1926, 1931,  
1941, 1951, 1961 и 1969 гг.

Год	Модальный возраст, полученный с помощью метода				Суммарный коэффициент рождаемости (на 1000 женщин)
	(1М)	(2М)		(1М) (наб по дтская величина)	
		(15— 50 лет) <sup>1</sup>	(17— 50 лет) <sup>1</sup>		
(1)	(2)	(3)	(4)	(5)	(6)
1926	28,1 (17,0—18,5) <sup>1</sup>	28,9	27,9	28,0	3356
1931	27,6 (16,9—19,3) <sup>1</sup>	28,5	27,6	27,7	3201
1941	26,3 (17,2—43,6) <sup>1</sup>	27,6	26,1	27,1	2824
1951	25,5 (17,0—19,9) <sup>1</sup>	26,6	25,4	25,3	3480
1961	24,2 (17,2—50,4) <sup>1</sup>	25,7	24,3	24,0	3857
1969	24,0 (16,6—53,0) <sup>1</sup>	25,1	23,7	23,8	2410

<sup>1</sup> Постоянный интервал рождаемости

<sup>2</sup> Оценка пределов детородного периода, основанная на методе (1М)

Цифры в скобках в столбце 2 табл. 1 — это оценки пределов детородного периода, полученные с помощью метода (4М), и видно, что самый нижний (16,6 года) и самый верхний (53,0 года) пределы приходятся на 1969 г. Нижний предел никогда не превышал 17,2 года; минимальное значение верхнего предела — 48,5 года. Значения верхнего предела варьируют, по-видимому, больше, чем значения нижнего предела. В общем, выведенная мера репродуктивного периода не вполне согласуется с фактически наблюдавшейся. Такого рода несовместимости неизбежны, когда все параметры оцениваются на основе моментов распределения.

Что касается модального коэффициента рождаемости, представленного в табл. 2, то в различиях между оценками, полученными с помощью разных методов, не прослеживается ясной закономерности. Все они, по-видимому, достаточно хорошо оценивают модальный коэффициент рождаемости, за исключением метода (1М) при интервале 15—50 лет, имеющего тенденцию значительно его преуменьшать.



Оценки модальных коэффициентов рождаемости для Канады с применением методов (4М), (2М) и (1М), 1926, 1931, 1941, 1951, 1961 и 1969 гг.

Год	Модальный коэффициент рождаемости, полученный с помощью метода					Наблюдаемый модальный коэффициент рождаемости
	(4М)	(2М)		(1М)		
		(15—50 лет)*	(17—50 лет)*	(15—50 лет)*	(17—50 лет)*	
1926	178	181	181	152	184	188
1931	174	175	175	150	181	183
1941	159	159	160	147	173	168
1951	205	203	205	172	204	210
1961	239	231	236	195	231	257
1969	157	150	155	130	155	168

\* Постоянный интервал рождаемости.

Параметры  $m_1$  и  $m_2$ , найденные различными способами, представлены в табл. 3 и 4. Как можно было ожидать, значения этих параметров независимы (что, впрочем, и должно быть) от общей частоты, поскольку их вывод зависит в первую очередь от распределения относительных частот. Соответственно изменения значений параметров отражают изменения характера самих распределений. Па-

Таблица 3

Оценки величины  $m_1$  для Канады с применением методов (4М), (2М) и (1М), 1926, 1931, 1941, 1951, 1961 и 1969 гг.

Год	Величина $m_1$ , полученная с помощью метода				
	(4М)	(2М)		(1М)	
		(15—50 лет)*	(17—50 лет)*	(15—50 лет)*	(17—50 лет)*
1926	0,90	1,55	0,98	0,83	1,05
1931	0,95	1,55	0,97	0,89	1,09
1941	0,83	1,49	0,89	1,13	1,23
1951	0,85	1,45	0,82	0,75	0,81
1961	0,70	1,37	0,73	0,64	0,64
1969	0,95	1,39	0,72	0,76	0,74

\* Постоянный интервал рождаемости.

Оценки величины  $m_2$  для Канады  
с применением методов (4М), (2М) и (1М), 1926,  
1931, 1941, 1951, 1961 и 1969 гг.

Год	Величина $m_2$ , полученная с помощью метода				
	(4М)	(2М)		(1М)	
		(15— 50 лет)*	(17— 50 лет)*	(15— 50 лет)*	(17— 50 лет)*
1926	1,67	2,35	2,00	1,40	2,11
1931	1,93	2,46	2,08	1,57	2,28
1941	2,12	2,68	2,25	2,14	2,82
1951	2,43	2,92	2,43	1,81	2,40
1961	2,64	3,12	2,57	1,86	2,39
1969	3,71	3,42	2,80	2,25	2,84

\* Постоянный интервал рождаемости.

параметр  $m_2$ , по-видимому, особенно чувствителен к такого рода изменениям. Это находит выражение в высоких коэффициентах корреляции, представленных в табл. 5, которые показывают степень связи между значениями  $m_2$  и различными параметрами возрастной кривой рождаемости, такими, как средний и модальный возрасты, дисперсия, показатели асимметрии и эксцесса. На протяжении всего периода наблюдения параметр  $m_2$  обнаружил довольно устойчивую тенденцию к росту, что, по-видимому, просто отражает снижение рождаемости в Канаде и сдвиг ее возрастной кривой. В течение того же периода значение параметра  $m_1$ , если не принимать во внимание сравнительно незначительных колебаний (см. табл. А.2 приложения), оставалось относительно постоянным; очевидно, что изменение возрастных закономерностей рождаемости оказало на этот параметр меньшее воздействие. Однако его явно высокую (отрицательную) корреляцию с суммарным коэффициентом рождаемости, показанную в табл. 5, объяснить трудно.

В свете результатов, представленных в табл. 3 и 4, также вполне очевидно, что параметры  $m_1$  и  $m_2$  должны быть очень чувствительны к способу оценки. Оценки, полученные различными методами, отличаются одна от другой, и при сравнении оказывается, что величина этих различий довольно существенна.

Матрица корреляции

	$a_1$	$a_2$	$m_1$	$m_2$	Суммарный коэффициент рождаемости	Модальный коэффициент рождаемости	Модальный возраст	Средний возраст	Среднее квадратичное отклонение	Асимметрия	Экссесс	$\kappa$	$\Delta$
$a_1$	1,00	-0,94	0,84	-0,81	-0,49	-0,75	0,95	0,97	-0,93	-0,95	-0,92	-0,93	-0,67
$a_2$	-0,94	1,00	-0,63	0,96	0,24	0,57	-0,96	-0,99	-0,98	0,98	0,99	0,99	0,67
$m_1$	0,84	-0,63	1,00	-0,38	-0,72	-0,83	0,69	0,70	0,60	-0,68	-0,58	-0,63	-0,61
$m_2$	-0,81	0,96	-0,38	1,00	0,02	0,36	-0,88	-0,91	-0,95	0,92	0,97	0,95	0,59
Суммарный коэффициент рождаемости	-0,49	0,24	-0,72	0,02	1,00	0,93	-0,42	-0,36	-0,26	0,31	0,21	0,24	0,26
Модальный коэффициент рождаемости	-0,75	0,57	-0,83	0,36	0,93	1,00	-0,71	-0,66	-0,58	0,62	0,55	0,57	0,47
Модальный возраст	0,95	-0,96	0,69	-0,88	-0,42	-0,71	1,00	0,98	0,96	-0,96	-0,96	-0,96	-0,63
Средний возраст	0,97	-0,99	0,70	-0,91	-0,36	-0,66	0,98	1,00	0,98	-0,98	-0,98	-0,98	-0,65
Среднее квадратичное отклонение	0,93	-0,98	0,60	-0,95	-0,26	-0,58	0,96	0,98	1,00	-0,98	-0,99	-0,98	-0,61
Асимметрия	-0,95	0,98	-0,68	0,92	0,31	0,62	-0,96	-0,98	-0,98	1,00	0,98	0,99	0,70
Экссесс	-0,92	0,99	-0,58	0,97	0,21	0,55	-0,96	-0,98	-0,99	0,98	1,00	0,99	0,66
$\kappa$	-0,93	0,99	-0,63	0,95	0,24	0,57	-0,96	-0,98	-0,98	0,99	0,99	1,00	0,70
$\Delta$	-0,67	0,67	-0,61	0,59	0,26	0,47	-0,63	-0,65	-0,61	0,70	0,66	0,70	1,00

## Согласие подобранных кривых

До сих пор наше внимание было сосредоточено на согласии наблюдаемых и оцениваемых параметров рождаемости. Рассмотрим теперь, в какой степени согласуется ряд выведенных возрастных коэффициентов рождаемости с наблюдаемыми.

Классический подход к измерению степени согласия подобранной кривой — вычисление  $\chi^2$ , величина которого частично зависит от относительной разности наблюдаемых и выравненных значений. Однако для ситуации, характеризующейся наличием большого числа интервалов, для которых часть оценок, полученных методом (4М), а именно в начале и в конце кривой рождаемости, были равны нулю, мы сочли этот способ оценки согласия неподходящим. Сравнивая наблюдаемые и ожидаемые распределения, можно получить обобщающий это сравнение показатель разными способами. Один из таких обобщающих показателей, индекс неоднородности ( $\Delta$ ), получают путем сведения двух распределений к процентной форме и суммирования только положительных разностей между соответствующими процентными значениями. Та или иная величина этого индекса указывает, какой процент наблюдений должен быть перераспределен между интервалами, для того чтобы оба распределения стали одинаковыми. Легко видеть, что индекс неоднородности ( $\Delta$ ) может принимать значения в диапазоне от 0 до 100 и что его величина зависит от выбора интервалов группировки.

Судя по табл. 6, общая ошибка, по-видимому, сравнительно невелика. Как можно было ожидать, значения индекса неизменно наименьшие для метода (4М), но если учитывать простоту и логическую непротиворечивость других методов, то значения индекса и для них не очень велики. Таким образом, в общем наша модель, по крайней мере в качестве первого приближения, представляет собой вполне удовлетворительный способ выравнивания для коэффициентов рождаемости. Только метод (1М), с интервалом рождаемости 15—50 лет, дает последовательно более значительную ошибку.

Для того чтобы получить некоторое представление о характере связанных с возрастом смещений, порождаемых моделью, отклонения выведенных возрастных коэффициентов рождаемости от наблюдаемых были вычислены и отображены на рис. 1. Наиболее примечательная особенность

Значения  $\Delta$  для Канады,  
полученные с помощью методов (4M), (2M) и (1M),  
1926, 1931, 1941, 1951, 1961 и 1969 гг.

Год	Значения $\Delta$ , полученные с помощью метода				
	(4M)	(2M)		(1M)	
		(15— 50 лет)*	(17— 50 лет)*	(15— 50 лет)*	(17— 50 лет)*
1926	2,74	4,08	2,77	8,91	2,80
1931	2,66	3,78	2,59	7,15	2,27
1941	2,32	4,30	2,39	5,33	4,51
1951	2,28	4,03	2,41	9,42	2,41
1961	2,97	5,38	3,30	11,34	4,03
1969	3,48	5,23	3,99	10,39	4,13

\* Постоянный интервал рождаемости.

рисунка заключается в том, что для всех лет наблюдения существуют возрастные смещения совершенно определенного типа. Модель имеет тенденцию попеременно то уменьшать, то преувеличивать коэффициенты рождаемости для ряда последовательных возрастов. Сходные возрастные смещения циклического характера порождаются и двумя другими методами.

Таблица 7

Отношение оцениваемого и фактического годовых чисел  
рождений для Канады, вычисленное с помощью методов  
(4M), (2M) и (1M), 1926, 1931, 1941, 1951, 1961 и 1969 гг.

Год	Отношение оцениваемого и фактического годовых чисел рождений, вычисленное с помощью методов				
	(4M)	(2M)		(1M)	
		(15— 50 лет)*	(17— 50 лет)*	(15— 50 лет)*	(17— 50 лет)*
1926	1,001	0,986	0,999	1,005	1,000
1931	1,001	1,001	1,000	1,010	0,998
1941	0,998	0,997	0,998	1,000	0,996
1951	0,998	0,997	0,997	0,984	0,996
1961	1,003	1,005	1,001	1,013	1,002
1969	1,008	1,006	1,011	1,018	1,009

\* Постоянный интервал рождаемости.

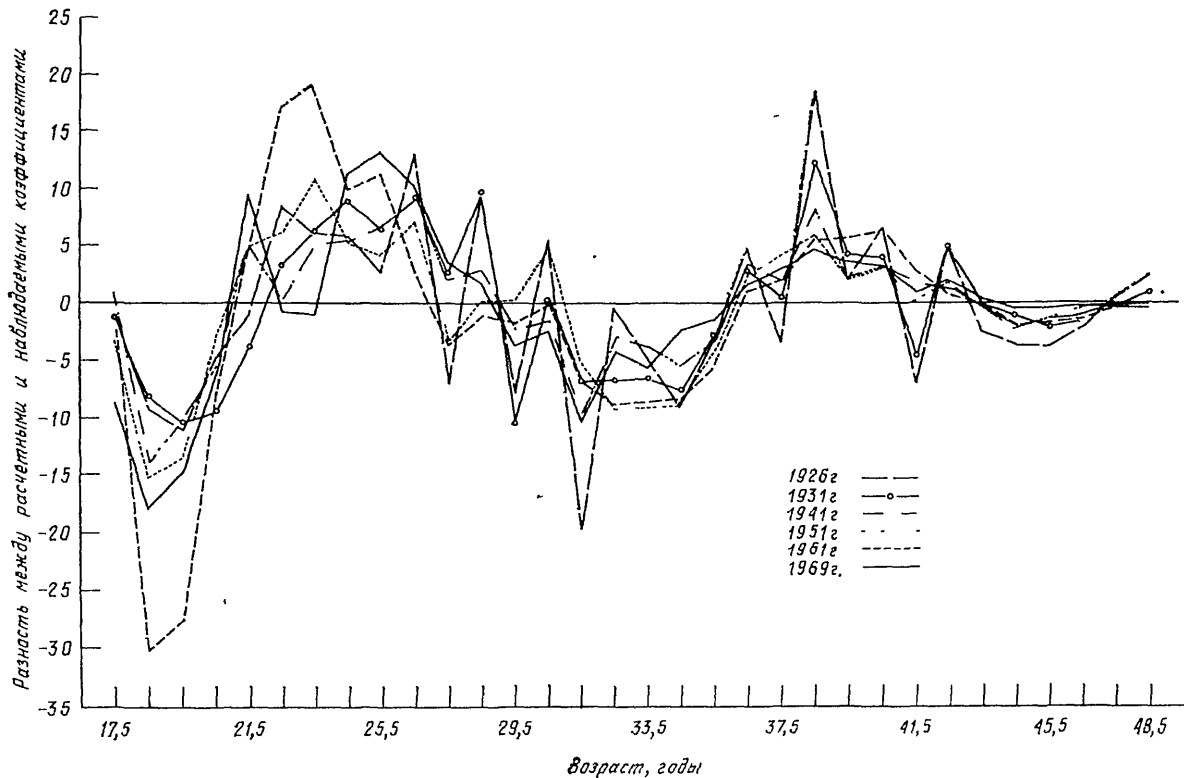


Рис 1 Абсолютное отклонение возрастных коэффициентов рождаемости, полученных методом [4M], от наблюдавшихся для населения Канады за несколько отдельных лет

Если модель предполагается применить для прогнозирования, то важно выяснить, каким структурным условиям рождаемости эта модель отвечает наилучшим образом. Для этой цели были вычислены коэффициенты корреляции между индексом неоднородности и параметрами возрастной кривой рождаемости. Как видно из табл. 5, все коэффициенты значимы, но самый высокий коэффициент (0,70) получен при корреляции с критерием  $k$ , который можно рассматривать как своего рода показатель совокупного влияния асимметрии и эксцесса на общую форму кривой рождаемости. Из этого следует, что при увеличении  $k$ , т. е. по мере того как его отрицательные значения все больше возрастают, ошибка, вызываемая моделью, становится все более значительной. Следовательно, после того как значение  $k$  достигает определенного уровня, уже другая кривая, например, упомянутая кривая переходного третьего типа, может обеспечивать лучшее выравнивание, чем кривая первого типа. Однако отрицательные значения критерия  $k$ , полученные на основе данных о рождаемости в Канаде, отнюдь не столь велики, чтобы оправдать применение кривой третьего типа. Ошибки в результатах, полученных каждым из трех методов, отличаются одна от другой по величине, но не по характеру. Таким образом, хотя рассмотренные коэффициенты и относятся к результатам, полученным с помощью метода (4М), выводы, которые можно сделать на их основе, справедливы для каждого из методов.

### ВОЗМОЖНОСТИ ПРИМЕНЕНИЯ КРИВОЙ ДЛЯ ПРОГНОЗА РОЖДАЕМОСТИ

В предыдущих разделах кривая Пирсона первого типа рассматривалась как средство выравнивания или воспроизведения наблюдаемых возрастных коэффициентов рождаемости. В настоящем разделе мы попытаемся распространить область применения этой кривой на прогнозирование. Сначала мы продемонстрируем, как это можно сделать, а затем покажем, какие именно вычислительные и аналитические преимущества модели, основанной на этой кривой, делают ее более эффективным средством прогнозирования чисел рождений по сравнению с обычными методами.

При попытке приспособить кривую Пирсона первого типа для целей прогнозирования полезно провести различие между параметрами, относящимися к *уровню* рождаемости

и к *возрастному распределению* показателей рождаемости. Напомним, что в формуле (1) первые представлены модельными коэффициентами рождаемости,  $y_0$ , а последние — параметрами  $a_1$ ,  $a_2$ ,  $m_1$  и  $m_2$ .

*Уровень рождаемости* можно прогнозировать одним из двух способов: либо прогнозируется модальный коэффициент рождаемости, для чего применяется модель в ее традиционной формулировке, предложенной Пирсоном; либо прогнозируется суммарный коэффициент рождаемости. Хотя между модальным и суммарным коэффициентами рождаемости существует очень высокая корреляция (0,93), по соображениям статистического и аналитического характера предпочтительнее прогнозировать суммарный коэффициент рождаемости. Это объясняется тем, что временные ряды модальных коэффициентов рождаемости не подвержены перепадам, вызванным, вероятно, ошибками в указаниях возраста и колебаниями случайного характера. Кроме того, применение суммарных коэффициентов рождаемости позволяет производить гораздо более глубокий анализ, чем применение модального коэффициента рождаемости. Например, для объяснения колебаний суммарного коэффициента рождаемости можно воспользоваться методом когорт, анализируя изменения величины семьи и сдвиги в календаре рождений [т. е. в размещении рождений по периоду деторождения. — *Ред.*] для последовательных когорт женщин [8]. Далее, будущие тенденции рождаемости можно прогнозировать путем анализа как данных о распределении женщин по числу рожденных ими детей, так и сведений, полученных из обследований относительно числа детей, которое намерены иметь опрашиваемые [9]. Эти преимущества, связанные с применением суммарного коэффициента рождаемости как основы для прогнозирования ее уровней, недостижимы при использовании модальных коэффициентов рождаемости, поскольку в последнем случае исследователь может полагаться только на прошлые тенденции, рассценивая их как ориентир для будущего.

Обратимся теперь к прогнозу возрастных коэффициентов рождаемости. В этом случае можно непосредственно прогнозировать либо *зависимые* параметры ( $a_1$ ,  $a_2$ ,  $m_1$  и  $m_2$ ), либо *независимые* параметры, т. е. моменты и прочие меры, связанные с распределением частот показателей рождаемости. Выбор одного из трех названных ранее методов как основы для вычисления зависимых параметров определяет выбор соответствующего независимого параметра.



Мы полагаем, что существуют два предварительных условия, соблюдение которых обязательно при выборе того или иного конкретного набора параметров для введения их в модель при перспективном расчете: параметры должны легко поддаваться анализу и допускать демографически осмысленное их толкование. Таким образом, модель, обеспечивающая хорошее согласие с эмпирическими данными, но не отвечающая одновременно двум упомянутым здесь условиям, может удовлетворять целям выравнивания, но совершенно не подходит для нужд прогнозирования. Согласно этим критериям *зависимые* параметры, обеспечивающие математическое описание, но не допускающие демографически осмысленного толкования, быть рекомендованы не могут. Здесь следует напомнить о трудностях, с которыми мы столкнулись ранее, пытаясь дать демографически осмысленное толкование поведению параметров  $m_1$  и  $m_2$ . В известной степени аналогичная судьба уготована и моментам более высокого порядка.

Из трех перечисленных ранее методов, по нашему мнению, критериям «рациональности» и демографически осмысленного толкования параметров в наибольшей степени отвечает метод (1M), требующий знания лишь двух простых и сравнительно легко интерпретируемых мер рождаемости, а именно возрастов средней и модальной рождаемости. Более того, поскольку коэффициент корреляции между этими двумя мерами высок (0,98), для чисто практических целей можно ограничиться анализом и прогнозированием только одной из них. Хотя метод (1M) и не обеспечивает такой высокой степени согласия, как другие методы, в частности, как (4M), однако недостаток точности первого метода с избытком компенсируется вычислительными и аналитическими преимуществами его применения.

Достоверность прогноза, основанного на каждом из рассмотренных методов, может быть определена только *ex post facto*. Однако в порядке дискуссии допустим на мгновение, что все независимые параметры были правильно спрогнозированы на период 1926—1970 гг. и что каждая из рассмотренных в настоящей работе моделей была применена для того, чтобы воспроизвести суммарные годовые числа рождений на протяжении данного периода. Если принять эти допущения, то всякие различия между фактическими и полученными с помощью моделей годовыми числами рождений будут, безусловно, зависеть от самих моделей. Отношение числа рождений, полученного с помощью каждой

из моделей, к фактическому числу рождений для Канады за отдельные годы приведено в табл. 7. Видно, что все значения этих отношений близки к единице. Это справедливо даже для метода (IM), основывающегося только на трех мерах рождаемости (а именно: на суммарном коэффициенте рождаемости, возрастах средней и модальной рождаемости). Следовательно, при вычислении суммарного годового числа рождений за указанный период каждая из трех моделей оказалась почти идеальной.

## ЗАКЛЮЧЕНИЕ

Прежде чем параметрическая модель прогноза рождаемости описанного здесь типа станет в полной мере пригодной для практических расчетов, необходимы дополнительные исследования. Тем не менее в настоящей работе содержатся существенно важные компоненты для создания такого рода модели. Показано, что возрастное распределение показателей рождаемости может быть выведено с помощью математических функций на основе ограниченного числа параметров, подлежащих прогнозу. В работе предлагаются альтернативные способы вычисления этих параметров, результаты вычислений проверяются путем сопоставления с данными о рождаемости для Канады; указываются дальнейшие пути развития и совершенствования прогнозов рождаемости. Сведение ряда возрастных коэффициентов рождаемости лишь к нескольким интерпретируемым параметрам дает возможность осуществить анализ с такой степенью глубины, которая была бы невозможна при традиционных методах, рассчитанных на прогноз рождаемости непосредственно с помощью возрастных коэффициентов рождаемости. Кроме того, модель дает значительную экономию усилий исследователей, поскольку при определении рождаемости посредством математической функции большое число расчетных операций может быть выполнено на ЭВМ.

## ЛИТЕРАТУРА

1. A very R. C. Graduation of age — specific fertility rates. A paper presented at the Third Conference on the Mathematics of Population, July 19—24, 1970, University of Chicago.
2. Canada. Dominion Bureau of Statistics. 1926—1969. Vital statistics, various years. Ottawa, Canada, Queen's Printer.
3. E l d e r t o n W. P. Frequency curves and correlation. Cambridge, Cambridge University Press, 1930.

4. Keyfitz N. Introduction to the mathematics of population. Reading, Mass., Addison — Wesley Publishing Company, 1968.
5. Mazur D. P. A demographic model for estimating age — order specific fertility rates. *Journal of the American Statistical Association*, 1963, vol. 58, p. 774—788.
6. Mitra S. The pattern of age — specific fertility rates. *Demography*, 1967, vol. 4, p. 894—906.
7. Murphy E. M. and Nagnur D. N. A Gompertz fit that fits: applications to Canadian fertility patterns. *Demography*, 1972, vol. 9, p. 35—50.
8. Ryder N. B. The emergence of a modern fertility pattern; United States, 1917—1966. In: S. J. Behrman, Leslie Corsa, Jr., and Ronald Freedman, Eds. *Fertility and family planning*. Ann Arbor, University of Michigan Press, 1969, p. 99—123.
9. Siegel J. S. and Akers D. S. Some aspects of the use of birth expectations data from sample surveys for population projections. *Demography*, 1969, vol. 6, p. 101—115.
10. Stone L. O. Parametric approaches of the age distribution of cohort total fertility for projections. Unpublished paper. Ottawa, Dominion Bureau of Statistics, 1970.
11. Tekse K. On demographic models of age — specific fertility rates. *Statistisk Tidskrift*, (Stockholm), Series III. *Statistical Review*, 1967, vol. 111 № 5, p. 189—207.
12. Wicksell S. D. Nuptiality, fertility and reproductivity. *Skandinavisk Aktuarietidskrift*, 1931, p. 125—157.

Приложение к статье см. на с. 148—150.

ПРИЛОЖЕНИЕ. Таблица А.1

Параметры распределения возрастных характеристик рождаемости для Канады, 1926 — 1969 гг.

Год	Суммарный коэффициент рождаемости (на 1000 женщин)	Фактическая модальная рождаемость	Модальный возраст*	Средний возраст матери при рождении ребенка	Среднее квадратичное отклонение	Мера асимметрии $\beta_1$	Мера эксцесса $\beta_2$	Критерий $\kappa$
1	2	3	4	5	6	7	8	9
1926	3360	188	27,98	30,13	6,61	0,059	2,285	-0,029
1927	3320	181	27,94	30,11	6,62	0,059	2,284	-0,029
1928	3300	196	27,86	30,04	6,62	0,062	2,286	-0,030
1929	3220	176	27,81	29,90	6,60	0,077	2,312	-0,037
1930	3280	182	27,75	29,87	6,58	0,082	2,321	-0,040
1931	3200	183	27,70	29,86	6,54	0,083	2,348	-0,042
1932	3090	175	27,65	29,96	6,57	0,075	2,341	-0,038
1933	2870	163	27,59	29,98	6,56	0,063	2,340	-0,032
1934	2800	154	27,52	30,08	6,54	0,054	2,331	-0,028
1935	2750	158	27,47	30,00	6,54	0,057	2,345	-0,030
1936	2700	156	27,42	29,96	6,52	0,058	2,337	-0,029
1937	2640	148	27,36	29,83	6,51	0,069	2,344	-0,035
1938	2700	152	27,30	29,67	6,52	0,084	2,351	-0,043
1939	2650	151	27,25	29,61	6,50	0,091	2,355	-0,046
1940	2760	159	27,19	29,36	6,47	0,115	2,388	-0,058
1941	2820	168	27,14	29,15	6,42	0,143	2,432	-0,072
1942	2950	171	26,94	29,12	6,39	0,141	2,446	-0,073
1943	3000	183	26,80	29,15	6,34	0,135	2,454	-0,071
1944	3000	173	26,60	29,30	6,37	0,114	2,415	-0,059
1945	3000	177	26,42	29,33	6,40	0,111	2,394	-0,056
1946	3360	209	26,25	29,01	6,33	0,155	2,464	-0,080
1947	3580	225	26,05	28,72	6,29	0,184	2,500	-0,095
1948	3420	211	25,87	28,67	6,32	0,181	2,498	-0,094
1949	3440	213	25,70	28,60	6,29	0,181	2,513	-0,095
1950	3430	213	25,50	28,64	6,36	0,109	2,523	-0,100
1951	3480	210	25,33	28,45	6,28	0,187	2,522	-0,099
1952	3620	222	25,15	28,35	6,26	0,197	2,536	-0,104
1953	3700	225	24,96	28,29	6,25	0,202	2,535	-0,106

Год	Суммарный коэффициент рождаемости (на 1000 женщин)	Фактическая модальная рождаемость	Модальный возраст*	Средний возраст матери при рождении ребенка	Среднее квадратичное отклонение	Мера асимметрии $\beta_1$	Мера эксцесса $\beta_2$	Критерий $\kappa$
1	2	3	4	5	6	7	8	9
1954	3810	242	24,78	28,25	6,27	0,211	2,518	-0,112
1955	3820	245	24,60	28,22	6,25	0,220	2,551	-0,111
1956	3850	254	24,40	28,12	6,23	0,232	2,581	-0,123
1957	3930	253	24,23	28,00	6,25	0,212	2,587	-0,126
1958	3880	247	24,14	27,93	6,21	0,249	2,605	-0,131
1959	3950	257	24,10	27,87	6,19	0,263	2,629	-0,140
1960	3910	255	24,06	27,81	6,17	0,270	2,636	-0,143
1961	3860	257	24,02	27,77	6,16	0,284	2,655	-0,151
1962	3770	255	24,00	27,75	6,11	0,291	2,669	-0,155
1963	3690	249	23,97	27,75	6,10	0,286	2,672	-0,155
1964	3520	236	23,94	27,80	6,11	0,284	2,681	-0,156
1965	3160	212	23,90	27,76	6,11	0,276	2,668	-0,151
1966	2820	193	23,86	27,62	6,13	0,296	2,700	-0,163
1967	2590	182	23,84	27,41	6,07	0,322	2,752	-0,182
1968	2440	170	23,80	27,28	5,98	0,341	2,837	-0,209
1969	2410	168	23,76	27,28	5,98	0,341	2,837	-0,208

\* Значения наблюдаемого модального возраста деторождения, выравненные от руки.

Источник. Данные для составления этой таблицы взяты из ежегодных отчетов, публикуемых в Vital Statistics, 1926—1969, Canada, Dominion Bureau of Statistics.

Таблица А.2

Параметры, оцененные с помощью метода Элдертон [3]  
1926—1969 гг.

Год	Модальный возраст	$\mu_0$	$\alpha_1$	$\alpha_2$	$m_1$	$m_2$
1	2	3	4	5	6	7
1926	28,08	177,90	11,10	20,50	0,90	1,67
1927	28,04	175,33	11,12	20,55	0,90	1,66
1928	27,92	174,35	10,99	20,64	0,89	1,66
1929	27,59	171,83	10,59	21,26	0,87	1,75

Год	Модальный возраст	$U_0$	$a_1$	$a_2$	$m_1$	$m_2$
1	2	3	4	5	6	7
1930	27,51	175,95	10,43	21,41	0,86	1,77
1931	27,60	173,79	10,68	21,68	0,95	1,93
1932	27,83	166,81	11,03	21,51	0,99	1,93
1933	28,09	161,01	11,61	21,20	1,09	1,99
1934	28,34	151,25	11,93	20,82	1,13	1,97
1935	28,26	149,00	11,99	21,08	1,17	2,05
1936	28,21	146,61	11,88	20,89	1,14	2,01
1937	27,86	143,78	11,30	21,24	1,06	1,99
1938	27,43	147,35	10,64	21,66	0,96	1,95
1939	27,24	144,93	10,34	21,79	0,92	1,93
1940	26,73	153,20	9,69	22,48	0,86	2,00
1941	26,30	159,39	9,14	23,27	0,83	2,12
1942	26,38	167,70	9,34	23,37	0,89	2,22
1943	26,58	172,31	9,62	23,25	0,96	2,33
1944	26,87	169,89	9,97	22,53	0,97	2,19
1945	26,84	168,40	9,88	22,27	0,92	2,07
1946	26,17	194,03	8,98	23,50	0,86	2,25
1947	25,62	209,85	8,34	24,09	0,79	2,29
1948	25,59	199,17	8,46	24,17	0,80	2,30
1949	25,63	201,82	8,65	24,30	0,86	2,41
1950	25,55	199,72	8,55	24,74	0,83	2,41
1951	25,45	205,07	8,54	24,43	0,85	2,43
1952	25,28	214,82	8,34	24,59	0,83	2,45
1953	25,14	219,89	8,15	24,55	0,80	2,41
1954	24,99	226,43	7,96	24,90	0,77	2,41
1955	24,88	228,13	7,76	24,88	0,74	2,38
1956	24,81	232,15	7,76	25,34	0,77	2,53
1957	24,55	236,81	7,51	25,45	0,73	2,47
1958	24,51	236,10	7,49	25,61	0,75	2,55
1959	24,39	241,98	7,35	25,96	0,74	2,62
1960	24,29	241,05	7,19	25,95	0,72	2,61
1961	24,17	239,29	7,01	26,24	0,70	2,64
1962	24,15	236,27	6,91	26,26	0,71	2,68
1963	24,25	231,65	7,07	26,27	0,74	2,75
1964	24,38	220,58	7,27	26,49	0,78	2,85
1965	24,36	196,56	7,38	26,40	0,79	2,81
1966	24,14	176,60	7,20	26,94	0,77	2,90
1967	23,91	165,30	7,01	27,54	0,79	3,09
1968	24,04	159,44	7,39	28,98	0,95	3,71
1969	24,04	157,39	7,39	28,97	0,95	3,71

Дональд Р. Мак-Нейл, Т. Джеймс Трассел,  
Джон С. Тернер

## ИНТЕРПОЛЯЦИЯ ДЕМОГРАФИЧЕСКИХ ДАННЫХ С ПОМОЩЬЮ СПЛАЙНОВ

Donald R. McNeil, T. James Trussell, John  
C. Turner. Spline Interpolation of Demographic  
Data. *Demography*, vol. 14, № 2, May 1977, p. 245—252.

Демографы часто сталкиваются с проблемами, связанными с интерполяцией, приближением и численным интегрированием и дифференцированием. Пожалуй, наиболее часто встречающаяся задача заключается в том, чтобы представить в виде некоторой плавной кривой ряд возрастных показателей рождаемости, для которого имеются только средние коэффициенты рождаемости по обычным пятилетним возрастным группам. Вторая по важности задача — это интерполяция столбца  $l_x$  (чисел доживающих) таблицы смертности.

Предположим, что имеется  $n + 2$  значения точек для  $n + 2$  возрастов. Пусть значениями нашей пока еще не определенной функции в этих возрастах будут  $f(x_0)$ ,  $f(x_1)$ , ...,  $f(x_{n+1})$ . Точки  $x_0$  и  $x_{n+1}$  называются *граничными* точками. Один из подходов мог бы состоять в том, чтобы соединить заданные точки прямыми линиями (т. е. многочленами первой степени); это решение действительно позволило бы получить непрерывную и *кусочно-гладкую* кривую\*, однако в заданных точках гладкость функции была бы нарушена. Далее можно думать о подборе к  $n + 2$

---

\* Гладкость — свойство функции иметь все возможные производные. — *Примеч. пер.*

точкам многочлена степени  $n + 1$ . Такая функция, несомненно, могла бы быть как непрерывной, так и гладкой. Значения неизвестных коэффициентов можно было бы найти, подобрав многочлен, равный  $f(x_0)$  при значении  $x_0$ ,  $f(x_1)$  — при значении  $x_1$  и т. д.

Основная сложность при таком решении состоит в том, что степень многочлена диктуется исключительно числом заданных точек; следовательно, по мере увеличения числа заданных точек увеличивается и размерность матрицы, которую следует обратить. Очевидно, из этого затруднения можно выйти, подобрав многочлены кусочно к парам заданных точек. Степень такого многочлена определяется произвольно. Для того чтобы обеспечить гладкость функции, можно было бы, например, подобрать многочлен второй степени. Такой многочлен в интервале от  $x_i$  до  $x_{i+1}$  имел бы коэффициенты  $a_i$ ,  $b_i$  и  $c_i$ . Поскольку эти многочлены должны перескаться, чтобы функция была непрерывной, и в точности соответствовать заданным точкам, мы имеем

$$a_i + b_i x_i + c_i x_i^2 = f(x_i) \quad (i = 0, 1, \dots, n) \quad (1)$$

и

$$a_i + b_i x_{i+1} + c_i x_{i+1}^2 = f(x_{i+1}) \quad (i = 0, 1, \dots, n). \quad (2)$$

Для того чтобы функция была гладкой в математическом смысле, производные каждой пары многочленов в точке их пересечения должны быть равны. Поэтому требуется, чтобы

$$b_i + 2c_i x_{i+1} = b_{i+1} + 2c_{i+1} x_{i+1} \quad (i = 0, 1, \dots, n-1). \quad (3)$$

Соотношения (1), (2) и (3) дают  $3n + 2$  уравнения с  $3n + 3$  неизвестными. Если, кроме того, можно задать значения производных в граничных точках  $x_0$  или  $x_{n+1}$ , то будет получено  $3n + 3$  линейных уравнения с  $3n + 3$  неизвестными коэффициентами; решение их хорошо известно. К сожалению, это решение требует обращения матрицы размерности  $(3n + 3) \times (3n + 3)$ . Ранее идея подбора многочлена степени  $n + 1$  казалась непривлекательной в силу того, что размерность обращаемой матрицы была столь же велика, сколь и число заданных точек. Альтернативное решение путем подбора кусочно-гладких многочленов степени  $m$  подразумевает обращение матриц размерности, в  $m$  раз большей, чем число заданных точек. Несмотря на трудность обращения большей матрицы, последний способ имеет



значительное преимущество перед предыдущим. Можно показать, что кусочные многочлены *более гладкие*, чем единственный многочлен степени  $n + 1$ . Кроме того, задачу подбора кусочных многочленов можно значительно упростить. При иной формулировке задачи число коэффициентов многочлена может быть сокращено с  $3n + 3$  до  $n + 3$ ; в общем виде число коэффициентов может быть сокращено до  $n + m + 1$ , где  $n$  — число заданных точек, за исключением граничных точек, и  $m$  — степень кусочно подбираемых многочленов; получаемый в результате многочлен называется *сплайном* или сплайн-функцией.

## СПЛАЙН-ФУНКЦИИ

Сплайн-функции известны на протяжении уже почти 30 лет (см. [5]), однако они все еще не нашли широкого применения, возможно, потому, что упоминаются лишь в очень немногих вводных пособиях по статистическим и численным методам. Исключение составляет работа Шемпина и Аллена [4]. Демографы могли получить некоторое представление о сплайнах из двухтомного труда Шрайока и Зигеля «Методы и материалы демографии» [7], однако никаких пояснений он не содержит\*.

Поэтому, в общем, мало известно, что сплайны — идеальное средство для решения многих практических задач, связанных с интерполяцией, приближением и численным интегрированием и дифференцированием. Более того, сплайн-функции просто определяются и легко вычисляются. Для того чтобы исправить создавшееся положение, мы приводим здесь достаточно доступное изложение методики интерполяции с помощью сплайнов, включающее определение и свойство гладкости, методику подбора и вычисление. В качестве иллюстрации приводится подбор сплайна к возрастным данным о рождаемости по Италии.

---

\* На русском языке опубликована работа: Алберг Д., Нильсон Э., Уолш Д. Теория сплайнов и ее приложения. М., Мир, 1972. Теория сплайн-функций отражена также в работах Н. П. Корнейчука, В. М. Тихомирова и Ю. Н. Субботина. — *Примеч. пер.*

## ОПРЕДЕЛЕНИЕ И СВОЙСТВО ГЛАДКОСТИ СПЛАЙНОВ

Сплайн степени  $m$  определяется как функция вида

$$s(x) = \sum_{j=0}^m d_j x^j + \sum_{i=1}^{n-1} c_i (x - x_i)_+^m, \quad (4)$$

где  $(x - x_i)_+ = x - x_i$  при  $(x - x_i) \geq 0$  и равно 0 в остальных случаях;  $x_1 < x_2 < \dots < x_n$ , а  $c_i$  и  $d_j$  — параметры, общее число которых составляет  $n + m + 1$ . Точки  $x_i$  называются *узловыми точками* сплайна; это точки, в которых нам известны значения функции. Из определения следует, что  $s(x)$  дифференцируема  $m$  раз и что производная степени  $m$  функции  $s(x)$ , обозначенная  $s(x)^{(m)}$ , есть ступенчатая функция с разрывами в узловых точках. Следует подчеркнуть, что для  $m = 2$  можно показать, что решение уравнения (4) идентично решению уравнений (1), (2) и (3) при условии, что идентичны граничные условия, рассматриваемые далее.

Сплайны обладают несколькими полезными свойствами (см., например, [3]). Можно показать, в частности, что среди всех функций  $f(x)$ , определенных на интервале  $(a, b)$ , значения которых заданы в точках  $x_1, x_2, \dots, x_n$ , ( $a < x_1; x_n < b$ ), существует некоторая единственная наиболее гладкая в математическом смысле функция, минимизирующая интеграл

$$\int_a^b \{f^{(k)}(x)\}^2 dx, \quad (5)$$

и что эта функция есть сплайн степени  $m = 2k - 1$ , удовлетворяющий дополнительным условиям:

$$\sum_{i=1}^n c_i x_i^j = 0 \quad (j=0, 1, \dots, k-1) \quad (6)$$

и

$$d_j = 0 \quad (j = k, k + 1, \dots, 2k - 1). \quad (7)$$

В случае, когда имеют силу соотношения (6) и (7), функция  $s(x)$  называется *нормальным сплайном* степени  $2k - 1$ . Поскольку (5) есть мера гладкости функции, нормальные сплайны применимы для гладкой интерполяции данных; более того, из выражения (5) следует, что для обеспечения

свойства гладкости должны применяться сплайны только нечетных степеней.

По условию задачи требуется, чтобы функция  $f(x)$  принимала значения, которые заданы в каждой из точек  $x_i$ . Если кроме этих интерполяционных условий функция  $f(x)$  должна удовлетворять граничным условиям:

$$\begin{aligned} f^{(i)}(a) &= u_i; \\ f^{(j)}(b) &= v_j, \end{aligned} \quad (8)$$

где  $i$  и  $j$  — элементы  $I$  и  $J$  — подмножеств целых чисел  $(0, 1, 2, \dots, k-1)$ , а  $u_i$  и  $v_j$  заданы, то оказывается, что остается единственная функция, которая минимизирует выражение (5), и что эта функция есть сплайн степени  $2k-1$ . В этом случае, однако, для того чтобы получить в итоге  $k$  ограничений, каждое в точках  $a$  и  $b$ , условия (6) и (7) заменяются условием (8) и вводятся дальнейшие условия относительно производных  $s(x)$  в точках  $a$  и  $b$ , начиная с производных  $k$ -й до  $m$ -й степени.

### МЕТОДИКА ПОДБОРА КРИВОЙ

Предположим, что у нас имеются данные  $y_1, y_2, \dots, y_n$  в точках решетки  $x_1, x_2, \dots, x_n$ . Имея в виду свойство гладкости, рассмотрим подбор к этим данным сплайнов лишь нечетных степеней 1, 3, 5, ... Подбор сплайна степени 1 тривиален: он заключается в простом соединении соседних точек прямыми линиями. В этом случае, как следует из сказанного ранее, невозможно удовлетворить каким бы то ни было граничным условиям относительно производных.

Далее рассмотрим задачу подбора сплайна степени 3, т. е. кубического сплайна. В этом случае  $k=2$ , тогда  $s(x)$  включает  $n+4$  параметра, которые определяются  $n$  интерполяционными условиями, а именно

$$s(x_i) = y_i, \quad (9)$$

и четырьмя уравнениями, получающимися из (6) и (7). Если обозначить  $\mathbf{c}' = (d_0, d_1, d_2, d_3, c_1, c_2, \dots, c_n)$  и  $\mathbf{y}' = (y_1, y_2, \dots, y_n, 0, 0, 0, 0)$ , то эту систему уравнений можно записать в матричной форме:

$$\mathbf{A}\mathbf{c} = \mathbf{y}, \quad (10)$$

где

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 & 0 & 0 & 0 & 0 \\ 1 & x_2 & x_2^2 & x_2^3 & (x_2 - x_1)^3 & 0 & 0 & 0 \\ 1 & x_n & x_n^2 & x_n^3 & (x_n - x_1)^3 & (x_n - x_2)^3 & \dots & (x_n - x_{n-1})^3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & & 1 & 1 \\ 0 & 0 & 0 & 0 & x_1 & x_2 & & x_{n-1} & x_n \\ 0 & 0 & 1 & 0 & 0 & 0 & & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & & 0 & 0 \end{bmatrix}$$

Таким образом, проблема подбора сплайн-функции сводится к решению системы уравнений (10).

Фактически нужно решить только  $n + 2$  уравнения, так как последние два уравнения в системе (10) в соответствии с условием (7) есть  $d_2 = d_3 = 0$ .

При интерполяции с помощью кубического сплайна существует возможность удовлетворить граничным условиям относительно значений данной функции и ее первой производной в начальной и конечной точках.

Предположим, что требуется удовлетворить условию (8) для  $J = I \{0, 1\}$ . Тогда нужно решить систему уравнений (10), в которой последние четыре элемента  $Y$  заменяются величинами  $u_0, u_1, v_0, v_1$  и последние четыре строки матрицы  $A$  заменяются следующими:

$$\begin{array}{ccccccc} 1 & a & a^2 & a^3 & 0 & \dots & 0 \\ 0 & 1 & 2a & 3a^2 & 0 & \dots & 0 \\ 1 & b & b^2 & b^3 & (b - x_1)^3 & \dots & (b - x_n)^3 \\ 0 & 1 & 2b & 3b^2 & 3(b - x_1)^2 & \dots & 3(b - x_n)^2 \end{array}$$

Первое и третье из этих уравнений указывают на то, что данная функция проходит через  $u_0$  и  $v_0$  в точках  $a$  и  $b$ ; второе и четвертое уравнения предполагают, что первые производные в точках  $a$  и  $b$  равны  $u_1$  и  $v_1$ . Аналогично производится интерполяция при помощи сплайнов 5, 7, ... степеней. В общем методический подход заключается в выборе величины  $n$ , на два меньшей, чем число интерполяционных точек; далее две крайние точки принимаются как гра-

ничные и их значения точно определяются указанием граничных условий.

## ВЫЧИСЛЕНИЕ

Как показано ранее, для того чтобы найти сплайн степени  $2k - 1$ , интерполирующий  $n$  заданных точек и удовлетворяющий до  $2k$  граничным условиям, нужно решить  $n + 2k$  линейных уравнений в виде (10). Параметры данного сплайна, таким образом, задаются при помощи уравнения

$$c = A^{-1}y. \quad (11)$$

Во многих случаях точки решетки и граничные условия определяются заранее характером решаемой задачи, поэтому подбор сплайна желательно производить после того, как получены данные. В этом случае полезно сначала вычислить  $A^{-1}$ , а затем получить  $c$ , повторно решая уравнение (11) при разных значениях  $y$ . Если  $A^{-1}$  рассчитано заранее, то эти вычисления можно делать вручную.

Таблицы в приложении содержат значения  $A^{-1}$  для решеток с равноудаленными точками  $x_i = i$  для  $n=4, 5, 6$  и для  $k = 2$  и  $3$ . Предполагается, что граничные условия  $f(0), f'(0), f(n+1), f'(n+1)$  для  $k = 2$  и  $f_0, f'(0), f''(0), f(n+1), f'(n+1), f''(n+1)$  для  $k = 3$  будут установлены самим исследователем; при пользовании указанной таблицей эти граничные условия обязательно должны задаваться именно в таком порядке, так как при вычислении таблицы была произведена замена условий (6) и (7) условием (8).

## ПРИМЕНЕНИЕ

Рассмотрим задачу нахождения наиболее близкой гладкой кривой, соответствующей ряду возрастных показателей рождаемости. Обычно имеются оценки коэффициентов рождаемости для возрастных интервалов 15—19, 20—24, 25—29, 30—34, 35—39, 40—44 и 45—49 лет, хотя иногда оба интервала 15—19 и 45—49 или один из них могут оказаться незаполненными. Если требуется отыскать

гладкую функцию (плотности распределения), площадь которой для каждого пятилетнего возрастного интервала соответствует заданным оценкам, то целесообразно искать некоторый сплайн, интерполирующий накопленные коэффициенты рождаемости к возрастам 15, 20, ..., 50 лет, и затем, продифференцировав этот сплайн, получить искомые значения функции возрастных коэффициентов рождаемости [6].

Без потери общности точки решетки могут быть взяты со значениями 1, 2, 3, 4, 5 и 6 с граничными точками при значениях 0 и 7, поскольку такая методика предполагает лишь перемену масштаба и начала координат по оси возраста. Желательно, чтобы подобранный ряд возрастных показателей рождаемости был непрерывным, а данный сплайн (т. е. функция, представляющая ряд накопленных коэффициентов рождаемости) имел бы нулевые производные в точках со значениями 0 и 7. Это требует применения, по меньшей мере, кубического сплайна. Практически желательно также, чтобы ряд возрастных показателей рождаемости имел нулевые производные в своих экстремальных точках, так чтобы значения плавно снижались до нуля в возрастах 15 и 50 лет. Хорошее приближение на крайних участках требует сплайна, по меньшей мере, пятой степени. Фактически может возникнуть необходимость в степенях, более высоких, чем пятая, так как часто случается, что первый и последний пятилетние коэффициенты рождаемости оказываются сравнительно небольшими. Если в этих случаях не применяются сплайны достаточно высоких степеней, то подобранная кривая возрастных коэффициентов рождаемости может иметь тенденцию принимать в некоторых точках отрицательные значения. Выбор величины  $k$  совершенно произволен. Как правило, выбирается наименьшее значение  $k$ , согласующееся с числом принятых ограничений. Если это значение дает неприемлемые результаты (например, отрицательные значения возрастных коэффициентов рождаемости), то выбирается значение  $k$ , на единицу большее. Однако в подобных случаях, может быть, предпочтительнее ослабить ограничение, касающееся соответствия площадей под кривой, и вместо интерполяционного сплайна подбирать к имеющимся данным условный сплайн [1].

Применение рассмотренной методики можно проиллюстрировать на следующем примере, содержащем ряд показателей рождаемости для женщин Италии за 1955 г. [2, с. 271]:

Возраст	Накопленные коэффициенты рождаемости
15	0
20	0,080
25	0,593
30	1,297
35	1,840
40	2,171
45	2,296
50	2,306

Согласно этим данным вектор  $y$  в уравнении (11) имеет следующие элементы: 0,080; 0,593; 1,297; 1,840; 2,171; 2,296; 0; 0; 0; 2,306; 0; 0; коэффициенты будут такими:

$j$	$d_j$	$c_j$
0	0	—
1	0	0,056073
2	0	—0,019144
3	0,020942	—0,006404
4	0,095030	0,008162
5	—0,035971	—0,002891
6	—	—0,005867

Подставив их в выражение (4), изменив масштаб и продифференцировав, находим

$$s(x) = \sum_{j=0}^5 d_j \left( \frac{x-15}{5} \right)^j + \sum_{i=1}^6 c_i \left( \frac{x-15}{5} - i \right)_+^5 \quad (12)$$

и

$$f(x) = s'(x) = \sum_{j=1}^5 \frac{j}{5} d_j \left( \frac{x-15}{5} \right)^{j-1} + \sum_{i=1}^6 c_i \left( \frac{x-15}{5} - i \right)_+^4. \quad (13)$$

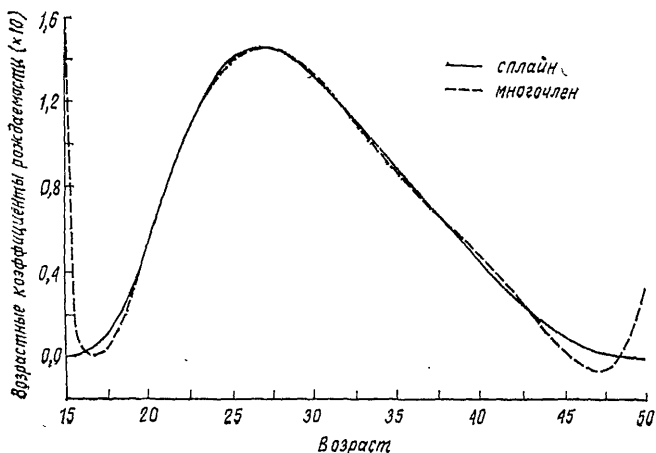


Рис. 1. Возрастная кривая рождаемости для женщин Италии, полученная при помощи сплайна и многочлена 7-й степени

На рис. 1 полученная таким способом возрастная кривая рождаемости сопоставлена с кривой, которая была бы получена путем подбора к тем же данным многочлена 7-й степени. Преимущества сплайна совершенно очевидны, так как кривая рождаемости, полученная с помощью многочлена 7-й степени, плохо выравнивает концы ряда. Некоторые значения на концах отрицательны. Можно предположить, что в этом случае целесообразно было бы ввести ограничения на концевые точки. Но если наложить ограничения, заключающиеся в том, что в возрастах 15 и 50 лет производные первой и второй степеней должны быть равны нулю, то по восьми заданным точкам нужно было бы подобрать многочлен 11-й степени с 12 неизвестными. Эта задача не легче, чем подбор сплайна 5-й степени, так как в обоих случаях должна быть обращена матрица размерности  $12 \times 12$ . При этом данный сплайн остается все же более гладким, чем многочлен 11-й степени.

#### ЛИТЕРАТУРА.

1. Anderssen Robert S., Bloomfield P. and McNeil D. R. Spline Functions in Data Analysis. Technica Report № 69, Series 2. Princeton, Department of Statistics, Princeton University, 1974.



2. Festy P. Évolution de la fécondité en Europe Occidentale Depuis la guerre. *Population*, 1970, 25, 229—274.
3. Greville T. N. E. Introduction to Spline Functions. In: Thomas N. E. Greville. *Theory and Applications to Spline Functions*. New York, Academic Press, 1969, p. 1—35.
4. Shampine Lawrence F. and Allen R. C. *Numerical Computing. An Introduction*. Philadelphia, W. B. Saunders Company, 1973.
5. Schoenberg I. J. Spline Functions and the Problem of Graduation. *Proceedings of the National Academy of Sciences*, 1946, 52, 947.
6. Schoenberg I. J. Splines and Histograms. In: A. Meir and A. Sharma. *Spline Functions and Approximation Theory*. Basel — Stuttgart, Birkhauser Verlag, 1973, p. 277—327.
7. Shryock Henry S., Siegel J. and Associates. *The Methods and Materials of Demography*. Washington, D. C., U. S. Government Printing Office, 1975.

Приложение к статье см. на с. 162—164.

Значения  $A^{-1}$  в уравнении (11) для некоторых  $n$  и  $k$   
 $k = 2; n = 4;$  сплайн 3-й степени

0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0
2,784688995	-0,746411483	0,200956938	-0,057416268	-2,196172249	-1,732057416	0,014354067	-0,004784689
-1,784688995	0,746411483	-0,200956938	0,057416268	1,196172249	0,732057416	-0,014354067	0,004784689
3,138755981	-1,985645933	0,803827751	-0,229665072	-1,784688995	-0,928229665	0,057416268	-0,019138756
-1,985645933	2,449760766	-1,813397129	0,803827751	0,746411483	0,248803828	-0,200956938	0,066985646
0,803827751	-1,813397129	2,449760766	-1,985645933	-0,200956938	-0,066985646	0,746411483	-0,248803828
-0,229665072	0,803827751	-1,985645933	3,138755981	0,057416268	0,019138756	-1,784688995	0,928229665

$k = 2; n = 5;$  сплайн 3-й степени

0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0
2,784615385	-0,746153846	0,2	-0,053846154	0,015384615	-2,196153846	-1,732051282	-0,003846154	0,001282051
-1,784615385	0,746153846	-0,2	0,053846154	-0,015384615	1,196153846	0,732051282	0,003846154	-0,001282051
3,138461538	-1,984615385	0,8	-0,215384615	0,061538462	-1,784615385	-0,928205128	-0,015384615	0,005128205
-1,984615385	2,446153846	-1,8	0,753846154	-0,215384615	0,746153846	0,248717949	0,053846154	-0,017948718
0,8	-1,8	2,4	-1,8	0,8	-0,2	-0,066666667	-0,2	0,066666667
-0,215384615	0,753846154	-1,8	2,446153846	-1,984615385	0,053846154	0,017948718	0,746153846	-0,248717949
0,061538462	-0,215384615	0,8	-1,984616385	3,138461538	-0,015384615	-0,005128205	-1,784615385	0,928205128

$k = 2; n = 6; \text{сплан 3-й степени}$

0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	1	0	0
2,7846101	-0,746135349	0,199931295	-0,053589832	0,014428032	-0,004122295	-2,196152525	-1,732050842	0,001030574	-0,000343525
-1,7846101	0,746135349	-0,199931295	0,053589832	-0,014428032	0,004122295	1,196152525	0,732050842	-0,001030574	0,000343525
3,138440398	-1,984541395	0,79972518	-0,214359327	0,057712126	-0,016469179	-1,7846101	-0,943203367	0,004122295	-0,001371018
-1,984541395	2,445894881	-1,799038131	0,750257643	-0,201992442	0,057712126	0,746135349	0,248711783	-0,014428032	0,004809344
0,79972518	-1,799038131	2,396427345	-1,786671247	0,750257643	-0,214359327	-0,199931295	-0,066643765	0,053589832	-0,017863277
-0,214359327	0,750257643	-1,786671247	2,396427345	-1,799038131	0,79972518	0,053589832	0,017863277	0,199931295	0,066643765
0,057712126	-0,201992442	0,750257643	-1,799038131	2,445894881	-1,984541395	-0,014428032	-0,004809344	0,746135349	-0,248711783
-0,016469179	0,057712126	-0,214359327	0,79972518	-1,984541395	3,138440398	0,004122295	0,001374098	-1,7846101	0,928203367

$k = 3; n = 4; \text{сплан 5-й степени}$

0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	1	0	0	0
4,900083998	-1,36051312	0,58056883	-0,331838873	-3,93271496	-3,237126876	-2,225201942	0,144414124	-0,078352503	0,012367419
-5,500320028	2,057372365	-0,915311084	0,525389666	4,061587152	3,049444314	1,616563438	-0,228718111	0,124096125	-0,019483126
1,60023603	-0,696859246	0,334742254	-0,193550794	-1,128872232	-0,812317438	-0,393271496	0,084063987	-0,045743622	0,007220706
-2,403104499	1,265642974	-0,711325425	0,442101714	1,60023603	1,100064006	0,4900084	-0,193550794	0,105077933	-0,016591944
1,265642974	-1,077085817	0,66488526	-0,711325425	-0,696859246	-0,411474473	-0,136051312	0,224742254	-0,165062217	0,029028441
-0,711325425	0,88488526	-1,077085817	1,265642974	0,334742254	0,183062217	0,038050003	-0,000000000	0,111474473	-0,068025656
0,442101714	-0,711325425	1,265642974	-2,403104499	-0,193550794	-0,105077933	-0,000000000	1,000000000	-1,100004006	0,2450042

$K = 3; \quad n = 5; \quad \text{сплайн } 5\text{-й степени}$

0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0
4,893080666	-1,348294657	0,552712385	-0,24842981	0,142743234	-2,929666518	-3,235472894	-2,224769795	-0,06214515	0,033716508	-0,00532257
-5,489227951	2,038020732	-0,871191245	0,39325389	-0,226070057	4,056758982	3,046824691	-1,617736447	0,098425649	-0,053403638	0,006429627
1,596147285	-0,689725575	0,31847525	-0,14485408	0,063326772	-1,12702404	-0,811351796	-0,392966652	-0,036280498	0,01965513	-0,00107258
-2,393711229	1,249255135	-0,57996268	0,390226795	-0,101282079	1,596147285	1,09784559	0,489308067	0,083326772	-0,045214011	0,007137164
1,249255135	-1,048494938	0,819699236	-0,516106274	0,330226795	-0,689725875	-0,407604146	-0,134829486	-0,14485408	0,078656778	-0,012421491
-0,57396268	0,819699236	-0,928430832	0,819699236	-0,97396268	0,31847886	0,174238249	0,055271238	0,31847886	-0,174235249	0,027635619
0,330226795	-0,516106274	0,819699236	-1,048494938	1,249255135	-0,14485406	-0,078656778	-0,024842981	-0,669725875	0,407604146	-0,067414745
-0,191282079	0,330226795	-0,67396268	1,249255135	-2,393711229	0,683326772	0,045214011	0,044274328	1,596147285	-1,09784559	0,214654033

$K = 3; \quad n = 6; \quad \text{сплайн } 5\text{-й степени}$

0	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0
4,891783385	-1,346037713	0,547705119	-0,236547566	0,106896035	-0,061452912	-3,929101718	-3,235166446	-2,224673051	0,02675537	-0,01451889	0,002291455
-5,487173306	2,034445847	-0,863260693	0,374464687	-0,16929486	0,097329201	4,055864447	3,046339335	-1,617583223	-0,042375324	0,022991964	-0,003629225
1,595389921	-0,688408134	0,315555574	-0,13791712	0,062398825	-0,035876289	-1,126762729	-0,811172889	-0,392910172	0,015619953	-0,008473065	0,001337769
-2,39197165	1,246228436	-0,667248226	0,31429339	-0,14321291	0,082397327	1,595389921	1,097434661	0,489178339	-0,035876289	0,01946384	-0,003072646
1,246228436	-1,043228774	0,808016733	-0,48838362	0,246589444	-0,14321291	-0,688408134	-0,406889169	-0,134603771	0,062398825	-0,033858972	0,005344802
-0,667248226	0,808016733	-0,902514205	0,758197475	-0,48838362	0,31429339	0,315555574	0,172652139	0,054770512	-0,13791712	0,074892937	-0,011827378
0,31429339	-0,48838362	0,758197475	-0,902514205	0,808016733	-0,667248226	-0,13791712	-0,074892937	-0,023654757	0,315555574	-0,172652139	0,027385256
-0,14321291	0,246589444	-0,48838362	0,808016733	-1,043228774	1,246228436	0,062398825	0,033858972	0,010689604	-0,062408134	0,00689169	-0,007301886
0,082397327	-0,14321291	0,31429339	-0,667248226	1,246228436	-2,39197165	-0,035876289	-0,01946584	-0,006145291	1,595389921	-1,097434661	0,244589169

## СОДЕРЖАНИЕ

Предисловие . . . . .	3
<i>Рафикул Хьюда Чаудхури.</i> Относительный доход и рождаемость. Перевод с английского Г. И. Ландсберг . . . . .	11
<i>Джеймс А. Уид.</i> Влияние возраста при вступлении в брак на различие коэффициентов разводимости по штатам. Перевод с английского О. Р. Пустыльник . . . . .	37
<i>Рональд Демос Ли.</i> Естественная рождаемость, демографические циклы и спектральный анализ рождений и браков. Перевод с английского А. Г. Минасовой . . . . .	64
<i>Тадеуш Грабиньский. Казимеж Заёнци.</i> Таксономические методы определения фаз развития демографических процессов. Перевод с польского Н. Н. Малютиной . . . . .	94
<i>С. Митра, А. Романюк.</i> Кривая Пирсона первого типа и возможность ее применения для прогноза рождаемости. Перевод с английского В. Л. Гопмана . . . . .	128
<i>Дональд Р. Мак-Нейл, Т. Джеймс Трассел, Джон С. Тернер.</i> Интерполяция демографических данных с помощью сплайнов. Перевод с английского В. А. Бирюкова . . . . .	151

С78

**Статистический анализ в демографии: Сб. статей/Под ред. А. Г. Волкова. — М.: Статистика, 1980. — 165 с., ил. — (Новое в зарубеж. демографии).**

1 р. 30 к.

Браки, разводы, рождения и смерти демография изучает как массовые явления. Узнать, как изменяются эти явления во времени и в пространстве, от чего зависят их различия, можно только с помощью статистических методов. Статьи сборника посвящены применению для изучения демографических явлений новых методов статистики — спектрального и дисперсионного анализа, методов таксономии, множественной регрессии и многомерной группировки.

Книга представляет интерес для всех, кто интересуется демографическими процессами в современном мире.

С  $\frac{10805^*-081}{008(01)-80}$  64—80 0703000000

ББК 60.7  
312

---

\* Второй индекс 10803.

## СТАТИСТИЧЕСКИЙ АНАЛИЗ В ДЕМОГРАФИИ

Редактор *Е. В. Крестьянинова*

Мл. редактор *И. Н. Горина*

Техн. редактор *Р. Н. Феоктистова*

Корректоры *Я. Б. Островский, В. Л. Долгова*

Худ. редактор *Э. А. Смирнов*

Обложка художника *Т. Н. Погореловой*

ИБ № 804

---

Сдано в набор 11.12.79. Подписано в печать 11.06.80  
Формат 84×108<sup>1</sup>/<sub>32</sub>. Бум. кн.-журн. Гарнитура «Литературная».  
Печать высокая. П. л. 5,25. Усл. п. л. 8,82. Уч.-изд. л. 8,92.  
Заказ 1528. Цена 1 р. 30 к. Тираж 2400 экз.  
Издательство «Статистика», Москва, ул. Кирова, 39.

---

Московская типография № 4 Союзполиграфпрома  
при Государственном комитете СССР  
по делам издательств, полиграфии и книжной торговли,  
129041 Москва, Б. Переяславская ул., 46

---

С 1981 г. издательство «Статистика» приступает к выпуску новой серии «Популярная демография», адресованной широкому кругу читателей. Одна из основных задач серии — в доступной форме освещать важнейшие демографические процессы. Серию готовят к печати ученые-демографы Ленинградского финансово-экономического института с привлечением широкого круга авторов из всех союзных республик.

В 1981 г. в серии выйдут книги:

Баздырев К. К. **Простое уравнение: муж + жена = семья**, 6 л.

Васильева Э. К. **Образ жизни городской семьи**, 6 л.

**Население Ленинграда** / Т. С. Баталлина, Г. Д. Платонов, Л. Е. Поляков и др., 6 л.

**Советская молодежь: демографический аспект** / В. Г. Асеев, Л. А. Горчаков, Л. Д. Константиновский и др., 6 л.

---