И. С. ПАЛЬМА, Л. Г. ЭЛЬГОРТ

ПРИМЕНЕНИЕ МЕТОДА КОРРЕЛЯЦИИ В СТРОИТЕЛЬСТВЕ

И. С. ПАЛЬМА, Л. Г. ЭЛЬГОРТ

ПРИМЕНЕНИЕ МЕТОДА КОРРЕЛЯЦИИ В СТРОИТЕЛЬСТВЕ

ПРИМЕНЕНИЕ МЕТОДА КОРРЕЛЯЦИИ В СТРОИТЕЛЬСТВЕ

СТАТИСТИКА МОСКВА 1971

В книге излагается методика корреляционного анализа экономики строительства на примерах парных связей и методы построения уравнений множественной регрессии (линейных и нелинейных), исследуется вопрос об использовании частных коэффициентов корреляции.

В работе приведены алгоритмы и блок-схемы программ выполнения корреляционного анализа с помощью электронно-вычислительных машин. В монографии использован большой фактический материал строительных организаций.

Книга рассчитана на экономистов и инженеров-строителей, работающих в научно-исследовательских и проектных институтах, в статистических и плановых органах, а также на студентов вузов.

ВВЕДЕНИЕ

ЦК КПСС и Совет Министров СССР 28 мая 1969 г. приняли постановление «О совершенствовании планирования капитального строительства и об усилении экономического стимулирования строительного производства»[1]. В свете постановления исследования в области экономики строительства должны доводиться до такой степени точности и конкретности, чтобы превратить результаты этих исследований в исходную базу для планомерного руководства строительством. В связи с этим повышается значение использования современных методов анализа деятельности строительных организаций.

В практике анализа работы отдельной строительной организации наиболее распространено сопоставление, которое сводится к выявлению отклонений величины какого-либо показателя от его планируемого, нормируемого или директивного значения за определенный отрезок времени (год, реже квартал) и определению экономии (перерасхода) денежных или материальных ресурсов, вызванных этими отклонениями[2]. Так, пользуясь сопоставительным методом, можно определить, как влияют на фактическую себестоимость строительно-монтажных работ отклонения от норм, цен, расценок, ставок, транспортных тарифов и других величин.

Однако, как только возникает необходимость установить закономерность влияния того или иного фактора на деятельность нескольких строительных организаций, указанный метод становится неприемлемым в силу несопоставимости исходных данных.

[1] «Правда», 20. VI. 1969 г.

[2] Описание метода приведено в пособии М. И. Балихина «Анализ производственно-хозяйственной деятельности строительных организаций», М., Изд-во литературы по строительству, 1967.

Для исследования влияния отдельных факторов на результаты экономической деятельности строительной организации широко применяется индексный метод. Анализ с помощью индексного метода ведется на основе сопоставления влияния факторов за два или более сравниваемых периода.

Ни метод группировок, ни индексный метод непригодны в тех случаях, когда необходимо выявить закономерности влияния нескольких факторов, действующих одновременно. Эти методы не дают возможности выявить и количественно измерить влияние на себестоимость ряда факторов, связанных с техническим прогрессом (рост производительности труда, сборности конструкций и др.), а также факторов, связанных с удорожанием строительства (неритмичное выполнение строительно-монтажных работ, разнотипность сооружений в жилищном строительстве и др.). Поэтому для экономического анализа деятельности строительных организаций должны быть применены иные методы.

За последние 15—20 лет в экономической науке особо важное значение приобретают математические методы для решения ряда практических задач на электронно-вычислительных машинах (ЭВМ).

Известно, что существуют два вида зависимостей. Одни из них характеризуются тем, что каждому значению переменной величины, называемой аргументом, соответствует определенное значение другой переменной величины, называемой функцией. Подобные зависимости называются функциональными. Так, в строительстве каждому изменению норм, расценок, ставок и других величин соответствует определенная величина себестоимости.

Другие зависимости характеризуются тем, что каждому значению переменной величины соответствует уже не одно, а несколько значений другой переменной величины. Такие зависимости называют корреляционными (от латинского correlatio — соотношение).

Связь между переменными величинами здесь проявляется уже не в каждом отдельном случае, а в совокупности однотипных случаев, причем она находит выражение в изменении обобщающих характеристик—средних величин. Так, связь между сметной стоимостью строительно-монтажных работ и фактической себестоимостью будет обнаружена только при сопоставлении средних (для ряда однотипных строек) значений сметной стоимости со средними значениями себестоимости работ.

Влияние множества факторов, которые находятся во взаимодействии, могут быть изучены с помощью корреляционного метода. Корреляционный метод анализа деятельности строительных организаций дает возможность оценить (количественно) влияние каждого фактора и всех в целом.

Впервые корреляционный метод был применен в работах Научно-исследовательского института строительного производства (НИИСП) Госстроя УССР в 1952—1954 гг. для расчетов нормативов накладных расходов в шахтном строительстве. Научными консультантами этой работы были проф. Б. С. Ястремский и проф. Я. И. Лукомский.

За прошедшие годы метод корреляционного анализа получил признание ученых-экономистов как надежный и в то же время тонкий инструмент проникновения в сущность сложных процессов, формирующихся под воздействием многих, часто противоположных тенденций.

В настоящее время корреляционный анализ применяется в работах научно-исследовательских институтов, в том числе НИИ экономики строительства Госстроя СССР.

Цель предлагаемого пособия — способствовать практическому использованию методов корреляции. Книга позволит получить общее представление о сути и значении применения методов корреляции в экономике и на практических примерах освоить технику расчетов или подготовки исходных материалов для обработки их на ЭВМ.

Процесс вычисления нескольких факторов с помощью счетных клавишных машин является весьма трудоемким, мы рекомендуем каждому, стремящемуся лучше освоить методы корреляции, пройти прежде этап ручной обработки. С этой целью в текст включен ряд упражнений с ответами.

В книге приводятся сведения по технике экономико-статистических исследований, анализ различных парных и множественных зависимостей и рекомендации по применению методов корреляции на различных уровнях строительной иерархии, по выбору факторов и порядку проведения исследований; методы построения нелинейных уравнений множественной регрессии, к которым все чаще приходится обращаться при анализе экономических процессов строительного производства, а также исследуются свойства частных коэффициентов корреляции, позволяющих полнее вскрывать механизм экономических явлений, и рекомендуются методы их определения.

Примеры, которыми иллюстрируется методика корреляционного анализа, построены на отчетных данных строительных организаций[1].

Перевод на новые условия планирования и экономического стимулирования всей строительной индустрии и введение с 1 января 1969 г. новых сметных норм и цен будут способствовать улучшению планирования в строительных организациях, повышению роли экономических методов управления, укреплению хозрасчетных взаимоотношений участников строительного процесса (заказчиков, генподрядчиков, субподрядчиков) и внутрихозяйственного расчета. Тем самым создадутся особенно благоприятные условия для более широкого применения экономико-математических методов, в том числе и методов корреляционного анализа в исследовании деятельности строительных организаций.

Важно отметить, что «Методические указания к составлению государственного плана развития народного хозяйства СССР» (М., «Экономика», 1969), утвержденные постановлением Госплана СССР от 21 февраля 1969 г. № 14, предусматривают применение в качестве дополнительного метода расчета в плане экономии затрат труда в строительстве корреляционного способа установления связи между отдельными факторами, влияющими на рост производительности труда. Для этой цели могут быть применены методы, изложенные в настоящем пособии.

Авторы глубоко признательны экономисту М. С. Гуревичу, оказавшему большую помощь в работе над книгой. Авторы выражают также благодарность математику Л. С. Воловик, выполнившей некоторые расчеты во 2—4-й главах книги; экономистам Л. А. Ковбасенко, И. Н. Бондаревой и Ю. Н. Подбело, которые приняли участие в сборе необходимой информации.

Отзывы о книге просим направлять по адресу: Киев-37, ул. Преображенская, 5/2, Научно-исследовательский институт автоматизированных систем планирования и управления в строительстве Госстроя УССР.

---

[1] Наименования организаций условны.

Глава I

СТАТИСТИЧЕСКИЕ ВЕЛИЧИНЫ, ИСПОЛЬЗУЕМЫЕ В КОРРЕЛЯЦИОННОМ АНАЛИЗЕ

§ 1. Ряд распределения

Объем выполненных в единицу времени (обычно за год) строительно-монтажных работ служит первой, самой общей характеристикой деятельности строительной организации.

Рассматривая однотипные, т. е. ведущие работы одного профиля и находящиеся в примерно одинаковых условиях строительно-монтажные управления (СМУ), можно сразу заметить, что объемы работ их весьма различны. Например, объемы работ, выполненных строительно-монтажными управлениями комбината[1] Донецкжилстрой (см. табл. 1), значительно различаются между собой: наибольшая их величина 3690 тыс. руб., а наименьшая — 210 тыс. руб. Естественно, что такие сведения об объеме работ не вскрывают причин их колебаний, от которых зависят важные технико-экономические показатели деятельности строительных организаций. Прежде всего необходимо систематизировать исходный материал. Самым простым в практике экономической работы способом систематизации является

[1] Комбинат — структурная единица системы управления строительными организациями, получившая распространение при возведении объектов тяжелой промышленности в Украинской ССР. Комбинат — территориальная хозрасчетная организация, объединяющая ряд общестроительных, некоторые специализированные тресты и промышленные предприятия строительной индустрии, осуществляющая планирование и регулирование деятельности подчиненных организаций; существует за счет отчислений первичных строительных подразделений, подчиняется непосредственно министерству.

Таблица 1

Исходные данные об объеме работ СМУ комбината Донецкжилстрой*

№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)	№ п/п	Объем работ (тыс. руб.)
1	550	12	310	23	950	34	1 050	45	1 040	56	1 250	67	1 280
2	1 460	13	930	24	1 580	35	1 340	46	260	57	860	68	1 560
3	1 530	14	2 440	25	860	36	1 540	47	910	58	1 720	69	1 300
4	2 830	15	560	26	930	37	1 070	48	1 490	59	1 160	70	900
5	790	16	790	27	1 270	38	3 690	49	920	60	1 450	71	1 810
6	1 370	17	930	28	1 820	39	790	50	2 120	61	2 460	72	1 080
7	1 750	18	2 260	29	580	40	270	51	1 190	62	390	73	1 280
8	1 350	19	1 470	30	1 440	41	1 550	52	1 360	63	2 000	74	1 760
9	1 480	20	1 650	31	1 550	42	1 750	53	1 500	64	2 290		
10	1 720	21	1 850	32	1 850	43	1 800	54	1 390	65	2 140		
11	1 910	22	2 000	33	2 090	44	1 950	55	2 230	66	210		

* Здесь и далее стоимостные показатели приведены в ценах, действовавших с 1 января 1961 г.

расположение исходных данных по возрастанию или убыванию признака (изучаемой переменной величины) с последующей разбивкой их на интервалы. Каждое значение показателя называют *вариантом* (в данном случае вариантов 74). В дальнейшем изучаемую переменную величину будем называть признаком.

От рационального выбора числа интервалов зависят результаты дальнейшего анализа. Возникает вопрос: на какое же число интервалов (l) следует разбивать исходный материал? Необходимо учитывать, что при использовании слишком малого числа интервалов могут не проявиться особенности признака; выбор чрезмерно большого числа интервалов препятствует упорядочению и обобщению данных, отдельные частности могут заслонить собой общую закономерность.

Исходный материал в 20—50 единиц рекомендуется разбивать на 4—6 интервалов, в 50—100 — на 6—8 интервалов и более 100— на 10—12 интервалов[1]. Каждый вариант

[1] С приведенной рекомендацией хорошо согласуются соотношения, которые приведены в ряде пособий: число интервалов принимается равным $1 + 3{,}2 \lg n$ (И. Г. Венецкий, Г. С. Кильдишев, «Ос-

(каждое наблюдение) должен входить только в один интервал, для этого в каждый интервал включаются варианты, значения которых больше нижней границы и меньше или равны верхней границе. И наконец, значения признака внутри каждого интервала по возможности не должны группироваться вблизи одной из границ; предполагается, что большинство вариантов располагается вокруг середины интервала.

Значения признака, расположенные в возрастающем или убывающем порядке, с указанием того, как часто они встречаются в пределах выбранных интервалов, представляют собой *ряд распределения*.

Ряды абсолютных или относительных частот, приведенные в табл. 2, являются *интервальными рядами распределения* годовых объемов работ строительно-монтажных управлений.

Таблица 2

Ряд распределения объема работ

№ интервала	Интервалы объемов работ (тыс. руб.)	Середина интервала	Абсолютные частоты	Относительные частоты
1	0—500	250	5	6,75
2	500—1 000	750	15	20,25
3	1 000—1 500	1 250	23	31,15
4	1 500—2 000	1 750	20	27,00
5	2 000—2 500	2 250	9	12,15
6	2 500—3 000	2 750	1	1,35
7	3 000—3 500	3 250	—	—
8	3 500—4 000	3 750	1	1,35
	Итого	—	74	100

Данные таблицы показывают определенную закономерность: интервал № 3 встречается с наибольшей частотой, т. е. наиболее часто на практике встречаются строительно-монтажные управления с годовым объемом работ 1000—

новы теории вероятностей и математической статистики», М. «Статистика», 1968, стр. 55; $1 + 3{,}322 \lg N$ (Ф. Миллс, «Статистические методы», Пер. с англ., М., Госстатиздат, 1958, стр. 49) и др.; последнее соотношение принято при составлении программы статистической обработки исходного материала на ЭВМ.

1500 тыс. руб. (31% общего числа управлений); по мере приближения к интервалу № 3 от границ ряда распределения частоты нарастают.

§ 2. Средняя арифметическая ($\bar{x}$)

Любая статистическая совокупность характеризуется прежде всего сводными, обобщающими показателями, в которых находит выражение закономерность изучаемого явления. Эти показатели называются *средними*.

Для корреляционного анализа из применяемых в статистике средних величин наибольшее значение имеет *средняя арифметическая* ($\bar{x}$). Для того, чтобы вычислить простую среднюю арифметическую из отдельных вариантов признака, нужно суммировать все варианты и полученную сумму разделить на их число:

$$\bar{x} = \frac{\sum_{1}^{n} x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}, \qquad (I.1)$$

где x_i — вариант переменной величины (признака);
n — число вариантов.

Средняя арифметическая годового объема строительно-монтажных работ (см. табл. 1) равна:

$$\bar{x} = \frac{550 + 1460 + 1530 + \ldots + 1080 + 1280 + 1760}{74} = 966 \text{ тыс. руб.}$$

Чаще всего приходится вычислять среднюю арифметическую рядов распределения, в которых учитывается частота вариантов, их «вес» в совокупности всех значений признака. В этом случае средняя арифметическая определяется по формуле:

$$\bar{x} = \frac{\sum_{1}^{n} x_i m_i}{\sum_{1}^{n} m_i} = \frac{x_1 m_1 + x_2 m_2 + \ldots + x_n m_n}{m_1 + m_2 + \ldots + m_n}, \qquad (I.2)$$

где m_i — абсолютные частоты или статистические веса вариантов. Эта величина называется *взвешенной средней арифметической*. Для рассмотренного нами примера (см. табл. 2).

$$\bar{x} = \frac{250 \cdot 5 + 750 \cdot 15 + 1250 \cdot 23 + 1750 \cdot 20}{74} +$$

$$+ \frac{2250 \cdot 9 + 2750 \cdot 1 + 3250 \cdot 0 + 3750 \cdot 1}{74} = 966 \text{ тыс. руб.}$$

Взвешивание по какому-либо признаку возможно только при условии соизмеримости весов. Например, если известно, что при возведении кирпично-блочных жилых зданий уровень механизации работ составляет: земляных 87,6%, монтажных 95,8%, штукатурных 54,7%, малярных 65% и прочих работ 45%, то определить средний (средневзвешенный) уровень механизации в целом — $\bar{y}_м$ можно только в том случае, если привести различные по характеру работы к единому измерителю, скажем, к трудоемкости. Если принять, что удельный вес указанных видов работ в общей трудоемкости равен соответственно 5,2; 32,4; 9,9; 13,8 и 38,7%, то средний уровень механизации составит:

$$\bar{y}_м = \frac{87,6 \cdot 5,2 + 95,8 \cdot 32,4 + 54,7 \cdot 9,9 + 65 \cdot 13,8 + 45 \cdot 38,7}{100} = 67,3\%.$$

Широкому использованию в расчетах именно средней арифметической способствуют два ее свойства: нулевое — сумма отклонений от средней арифметической равна нулю и минимальное — сумма квадратов отклонений от средней арифметической меньше, чем сумма квадратов отклонений от любого другого числа. Минимальное свойство выражается уравнением $f = \sum (x - c)^2 = \min$,
где c — число, которое обращает функцию f в минимум. Условие, выраженное этим уравнением, называется *требованием наименьших квадратов*.

§ 3. Среднее квадратическое отклонение σ, средние ошибки $\bar{x}$ и σ

Важнейшим качеством рядов распределения является возможность с их помощью давать количественную оценку *колеблемости* признака. Для примера рассмотрим данные по двум трестам комбината Донецкжилстрой (см. табл. 3).

В тресте № 1 выполнение программы работ по кварталам весьма незначительно отклонялось от среднего уровня

Таблица 3

Объемы работ, выполненные трестами комбината Донецкжилстрой

Кварталы 1968 г.	Трест № 1			Трест № 2		
	объем работ, выполненный собственными силами		отношение фактической себестоимости к сметной за год (%)	объем работ, выполненный собственными силами		отношение фактической себестоимости к сметной за год (%)
	тыс. руб.	%		тыс. руб.	%	
I	7 636	23,5	91,79	1 443	13,8	102,15
II	8 473	26,2		2 327	22,3	
III	8 344	25,8		2 641	25,2	
IV	7 836	24,5		4 048	38,7	

(25%) в квартал, в то время как в тресте № 2 аналогичные показатели характеризуются значительной колеблемостью, главным образом за счет снижения объема работ в I квартале и резкого повышения в IV. Сопоставление уровней себестоимости показывает, какое важное экономическое значение имеет меньшая колеблемость выполнения программы работ по кварталам. Резкие колебания объемов работ по кварталам года в тресте № 2 явились причиной повышения фактической себестоимости работ в сравнении со сметной на 2,15% за счет того, что рабочая сила и ведущие строительные машины использовались весьма неравномерно, нерационально. При более равномерной в течение года работе в тресте № 1 фактическая себестоимость была ниже сметной на 8,21%.

Явление колеблемости вариантов неизменно сопутствует изменению почти всех факторов, действующих в сфере экономики строительного производства. Измерение степени этой колеблемости имеет важное значение во всех случаях применения статистического метода к анализу экономики строительства.

Наиболее распространенным и удобным показателем характеристики колеблемости (рассеяния) признака является *среднее квадратическое отклонение* σ, которое определяется по формуле

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}, \qquad (I.3)$$

где σ, x, $\bar{x}$ имеют ту же размерность, что и варианты признака.

Вычислим σ по данным о производительности труда передвижных механизированных колонн Министерства сельского строительства УССР (1967 г.).

Таблица 4

Расчет среднего квадратического отклонения

№ п/п	Выработка на 1 работающего за год (руб.), x	$\lvert x-\bar{x}\rvert$	$(x-\bar{x})^2$	№ п/п	Выработка на 1 работающего за год (руб.) x	$\lvert x-\bar{x}\rvert$	$(x-\bar{x})^2$
1	4 330	416	173 056	17	5 070	324	104 976
2	6 810	2 064	4 260 096	18	4 830	84	7 056
3	3 364	1 382	1 909 924	19	4 713	33	1 089
4	4 605	141	19 881	20	4 194	552	304 704
5	4 566	180	32 400	21	5 620	874	763 876
6	5 297	551	303 601	22	4 961	215	46 225
7	4 349	397	151 609	23	4 270	476	226 576
8	5 124	378	142 884	24	5 113	367	134 689
9	4 525	221	48 841	25	4 168	578	334 084
10	5 500	754	568 516	26	4 607	139	19 321
11	4 322	424	179 776	27	4 930	184	33 856
12	4 238	508	258 064	28	4 089	657	431 649
13	5 560	814	662 596	29	4 137	609	370 881
14	4 782	36	1 296	30	4 309	437	190 969
15	4 154	592	350 464	31	5 202	406	207 936
16	5 387	641	410 881				
				Итого	$\Sigma = 147\,126$		$\Sigma = 12\,657\,772$

$$\bar{x} = \frac{147\,126}{31} = 4\,746; \qquad \sigma = \sqrt{\frac{12\,657\,772}{31}} = 638{,}9;$$

$\bar{x} \pm \sigma = 4746 \pm 638{,}9$ показывает величину средней колеблемости признака.

Данные табл. 4 не сгруппированы, и σ, вычисленная для этих данных по формуле (I.3), будет *простой*. Для сгруппированных данных (табл. 5), т. е. когда они представлены в виде рядов распределения, применяется вычисление *взвешенного* среднего квадратического отклонения по формуле

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2 \cdot m}{\sum m}}, \qquad (I.4)$$

где m — веса (частоты).

Колеблемость признака можно определять также с помощью показателя дисперсии σ^2. Такой показатель позволяет упростить ряд вычислений, связанных с определением колеблемости; термин «дисперсия» употребляется

Таблица 5

Расчет взвешенного среднего квадратического отклонения

Интервалы объемов работ (тыс. руб.)	Середины интервалов x	Частоты m	$x \cdot m$	$\lvert x-\bar{x}\rvert$	$(x-\bar{x})^2$	$(x-\bar{x})^2 \cdot m$
0— 500	250	5	1 250	1 142	1 304 160	6 520 800
500—1 000	750	15	11 250	642	412 164	6 182 460
1 000—1 500	1 250	23	28 750	142	20 164	463 772
1 500—2 000	1 750	20	35 000	358	128 164	2 563 280
2 000—2 500	2 250	9	20 250	858	736 164	6 625 476
2 500—3 000	2 750	1	2 750	1 358	1 844 416	1 844 416
3 000—3 500	3 250	—	—	1 858	3 452 160	—
3 500—4 000	3 750	1	3 750	2 358	5 560 160	5 560 160
Итого		74	103 000			29 760 364

$$\bar{x} = \frac{103\,000}{74} = 1\,392; \qquad \sigma = \sqrt{\frac{29\,760\,364}{74}} = 634.$$

также для обозначения самого свойства колеблемости.

Воспользуемся следующими свойствами $\bar{x}$ и σ:

I. Если все частоты (веса) умножить на одно и то же число, то $\bar{x}$ и σ не изменятся.

II. Если все варианты признака умножить на одно и то же число, то $\bar{x}$ и σ увеличатся во столько же раз.

III. Если ко всем вариантам признака прибавить одно и то же число «a», то $\bar{x}$ увеличится на «a», а σ останется без изменения.

IV. Дисперсия равна среднему квадрату минус квадрат средней:

для совокупности, представленной простым средним значением признака, $\sigma^2 = \frac{\sum x^2}{n} - (\bar{x})^2$;

для совокупности, представленной групповыми значениями признака, $\sigma^2 = \frac{\sum x^2 m}{\sum m} - (\bar{x})^2$.

Как будет показано далее, вычисление σ из уравнений

$$\sigma = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2} \tag{I.5}$$

или

$$\sigma = \sqrt{\frac{\sum x^2 m}{\sum m} - (\bar{x})^2} \tag{I.5a}$$

наиболее просто.

В экономических расчетах иногда необходимо объединять частные совокупности и уметь рассчитать показатели колеблемости объединенной совокупности. Например, имеются две группы рабочих 80 и 45 человек; среднее выполнение норм выработки рабочими первой группы 107,3%, второй — 121,5%; дисперсия по выполнению норм выработки первой группы 2,5 и второй группы 4,25. Обе группы рабочих объединены в одну численностью 125 человек. Надо найти среднее выполнение норм выработки $\bar{x}$ и дисперсию σ^2 объединенной совокупности. Показатели частных совокупностей назовем частными, искомые показатели объединенной совокупности — общими. Средняя общей совокупности:

$$\bar{x} = \frac{80 \cdot 107,3 + 45 \cdot 121,5}{80 + 45} = 112,2\%.$$

Общая дисперсия определяется уравнением

$$\sigma^2 = \bar{\sigma}_i^2 + \bar{\delta}_i^2, \tag{1.6}$$

где $\bar{\sigma}_i^2 = \dfrac{n_1 \sigma_1^2 + n_2 \sigma_2^2 + \ldots}{n_1 + n_2 + \ldots}$ — взвешенная средняя из частных дисперсий;

$\bar{\delta}_i^2 = \dfrac{n_1 \delta_1^2 + n_2 \delta_2^2 + \ldots}{n_1 + n_2 + \ldots}$ — дисперсия частных средних.

Следовательно, общая колеблемость складывается из колеблемости внутри частей и колеблемости между частями объединяемой совокупности.

В рассматриваемом примере:

$$\delta_1^2 = (x_1 - \bar{x})^2 = (107,3 - 112,2)^2 = 24,01;$$

$$\delta_2^2 = (x_2 - \bar{x})^2 = (121,5 - 112,2)^2 = 86,49;$$

$$\bar{\delta}_i^2 = \frac{80 \cdot 24,01 + 45 \cdot 86,49}{80 + 45} = 46,6;$$

$$\bar{\sigma}_i^2 = \frac{80 \cdot 2,5 + 45 \cdot 4,25}{80 + 45} = 3,13;$$

$$\sigma^2 = 46,6 + 3,13 = 49,73.$$

В данном случае общая дисперсия выполнения норм выработки, объединенной совокупностью рабочих (125 человек), характеризуется дисперсией частных средних, так как 46,6 : 49,73 = 0,94, т. е. на общую колеблемость показателей выполнения норм выработки в основном влияет колеблемость между частями объединенной совокупности.

Происходит это потому, что средний уровень признака в этих группах весьма различен (107,3% и 121,5%) и первая группа рабочих почти в два раза больше второй.

Среднее квадратическое отклонение — одна из наиболее важных статистических величин, с помощью которой можно судить о принадлежности того или иного наблюдения к известному уже ряду распределения.

Согласно теории вероятностей в пределах $\bar{x} \pm \sigma$ будет находиться 68,3% всего числа вариантов, в пределах $\bar{x} \pm 2\sigma$ — до 95,4% всех вариантов и в пределах $\bar{x} \pm 3\sigma$ практически уложится все количество вариантов — 99,7% (правило трех сигм).

В практике экономической работы в строительстве обычно оперируют данными о деятельности не всех однородных организаций (не всей совокупности организаций), а типичных для той или иной группы организаций. Отбор типичных объектов обеспечивает получение представительных средних и в то же время позволяет существенно сократить трудоемкость сбора исходных данных.

Измерить *все* значения вариантов какого-либо признака не всегда возможно. В этих случаях поступают следующим образом: в расчет включают дополнительную характеристику, которая позволяет по среднему значению, полученному на основании ограниченного числа наблюдений, судить об общей (истинной) величине средней всей совокупности. Такого рода характеристиками являются средние случайные ошибки, обозначаемые буквой μ с индексом, указывающим величину, к которой оно относится. Ошибки выражаются через σ и n или только n.

Так, средняя ошибка средней арифметической

$$\mu_{\bar{x}} = \frac{\sum |x - \bar{x}|}{n \cdot \sqrt{n}}. \qquad (I.7)$$

Средняя ошибка среднего квадратического отклонения

$$\mu_{\sigma} = \frac{\sqrt{\sum |x - \bar{x}|^2}}{n} = \frac{\sigma}{\sqrt{n}}. \qquad (I.8)$$

Для данных, приведенных в табл. 4,

$$\mu_{\bar{x}} = \frac{416 + 2064 + \ldots + 437 + 406}{31 \cdot \sqrt{31}} = 89{,}2 \text{ руб.};$$

$$\mu_{\sigma} = \frac{\sqrt{12657772}}{31} = \frac{638{,}9}{\sqrt{31}} = 114 \text{ руб.}$$

Отношение $\frac{\mu_\sigma}{\mu_{\bar{x}}}$ должно находиться в пределах 1,25÷1,30; такие пределы наблюдаются в тех случаях, когда случайные ошибки подчиняются закону нормального распределения. В данном случае

$$\frac{\mu_\sigma}{\mu_{\bar{x}}} = \frac{114}{89,2} = 1,28.$$

§ 4. Статистические показатели степени ритмичности строительного производства

Для предварительных расчетов колеблемости (величины рассеяния) признака используется также *размах вариации «р»* — разность между наибольшим и наименьшим значениями вариационного ряда. Показатели колеблемости σ, p являются размерными и измеряются в тех единицах, что и исходные данные. Для характеристики колеблемости признака применяются и относительные показатели, из которых в экономике строительства распространены коэффициенты вариации v[1] и коэффициент ритмичности R (для оценки равномерности процессов строительного производства)

$$v = \frac{\sigma}{\bar{x}} \cdot 100\%.$$

Так, для ряда, приведенного в табл. 4, характеризующего выработку на 1 работающего,

$$\text{при } \sigma = 638,9 \text{ руб.}$$

$$\bar{x} = 4746 \text{ руб.};$$

$$v = \frac{638,9}{4746} \cdot 100 = 13,45\%.$$

Коэффициент вариации v или, что то же — мера неритмичности, конечно, может служить для оценки колеблемости тех или иных процессов, а также в целом динамики строительного производства.

[1] По данным доктора эконом. наук В. И. Сиськова, приведенным в «Примерной методике сводной экономико-статистической оценки качества продукции массового производства» (М., «Статистика», 1967, стр. 30), значения коэффициента вариации указывают: от 0 до 17% на высокую степень однородности, от 17% до 33% — достаточную однородность, более 33% — на неоднородность признака.

Но обратим внимание на трудоемкость вычислений, связанных с определением только одного значения коэффициента вариации. Например, при определении v по четырем интервалам времени (например, по четырем кварталам года):

$$v=\frac{\sqrt{\frac{\sum(x-\bar{x})^2}{4}}}{\frac{\sum x}{4}}100\%=$$

$$=\frac{\sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+(x_4-\bar{x})^2}{4}}}{\frac{x_1+x_2+x_3+x_4}{4}}100\%.$$

Выполнение значительного количества вычислений, пусть даже элементарных (в нашем примере), для установления одного значения v делает вероятными арифметические ошибки и поэтому *малоприемлемо в практической* деятельности (т. е. в работе плановых отделов трестов, планово-экономических управлений главков, органов Госстроя и Госплана). Поэтому рекомендуется даже несложные вычисления коэффициента вариации заменять значительно более простым графическим методом, учитывая при этом возможность характеристики каждого явления при помощи одной из двух разнонаправленных — позитивной и негативной, либо дополняющих друг друга величин.

Поясним сказанное примером. Выражение «производительность труда на штукатурных работах равна 12 $м^2$ в смену» можно заменить иным — «трудоемкость штукатурных работ равна 1/12 $см/м^2$». Трудоемкость и производительность труда являются взаимно обратными величинами. Таково же соотношение величин коэффициента вариации v (или относительной степени неритмичности ρ) и дополняющей (или обратной ей) величиной коэффициента ритмичности (R):

$$v+R=1 \quad \text{или} \quad \rho+R=1.$$

Можно принять другие соотношения взаимного дополнения v и R; например, $R=\sqrt{1-v^2}$, $R=\frac{1}{1+v}$ и др. Однако такие соотношения сложнее и в данном случае неоправданны. Можно было принять v и R взаимно обратными. При этом если $v = 0{,}5$; 0,25; 0,1; 0, то соответственно $R = 2; 4; 10; +\infty$ и, таким образом, вполне ощутимому уменьшению v соответствует резкое увеличение R; это затрудняет восприятие показателя.

Целесообразно выражать степень неритмичности мерой вариации объемов работ (потребления ресурсов производства или величины сдачи объектов в эксплуатацию), при которой среднее квадратическое отклонение явится свободной абсолютной мерой неритмичности ρ. Степень ритмичности (R) определяется как дополнительная величина: $R = 1 - \rho$.

Приведенная точка зрения не общепринята; например, А. Л. Филахтов[1] считает, что вообще нельзя говорить о ритмичности, что имеет место только бо́льшая или меньшая степень неритмичности, а ритмичность — лишь частный случай неритмичности, при котором $v = 0$. Такая трактовка сужает понятие ритмичности, ограничивает ее только случаем, когда в течение всех отрезков рассматриваемого периода времени наблюдаются неизменные (средние арифметические) значения того или иного фактора.

Методика определения коэффициента ритмичности R, рассмотренная ниже, весьма проста и может быть (при разделении рассматриваемого периода времени на четыре отрезка, скажем, года — на четыре квартала) иллюстрирована следующим примером.

Таблица 6

Расчет коэффициента ритмичности R по абсолютным величинам

Кварталы	Объем работ (тыс. руб.) x	Отклонение от средней арифметической $x-\bar{x}$	$(x-\bar{x})^2$
I	499,5	— 9,5	90,25
II	547,4	38,4	1 474,56
III	536,5	27,5	756,25
IV	455,9	—53,1	2 819,61
Всего	2039,3	0	5 140,67

$$\bar{x} = \frac{2039{,}3}{4} = 509 \text{ тыс. руб.}; \quad \sigma = \sqrt{\frac{5\,140{,}67}{4}} = 35{,}81 \text{ тыс. руб.}$$

$$\rho = \frac{35{,}81}{509} = 0{,}0704; \quad R = (1 - 0{,}0704)\cdot 100 = 92{,}96\%.$$

[1] А. Л. Филахтов. Принципы поточного возведения гидроузлов в оптимальные сроки. Докторская диссертация, Киев, 1964.

Еще проще определяется R, если значения объемов работ (либо расходования ресурсов производства, либо сдачи готовой продукции в эксплуатацию) принимаются в процентном отношении.

Таблица 7

Расчет коэффициента ритмичности R по относительным величинам

Кварталы	Объем работ, выполненных по генподряду (тыс. руб.)	Отклонение от средней арифметической	$(x-\bar{x})^2$
	x	$x-\bar{x}$	
I	15	—10	100
II	20	— 5	25
III	40	15	225
IV	25	0	0
Всего	100	0	350

$$\bar{x}=\frac{100}{4}=25;\ \sigma=\sqrt{\frac{350}{4}}=9{,}35;\ \rho=\frac{9{,}35}{25}=0{,}374;$$

$$R=(1-0{,}374)\cdot 100=62{,}6\%.$$

Предлагается следующий способ определения R для четырех равных отрезков времени

$$R=(1-\rho)\cdot 100=\left(1-\frac{\sigma}{\bar{x}}\right)\cdot 100=$$

$$=\left(1-\frac{\sqrt{\Sigma(x-25)^2}}{2\cdot 25}\right)\cdot 100=2\left[50-\sqrt{\Sigma(x-25)^2}\right].\quad (I.9)$$

Таким образом, коэффициент ритмичности является функцией величины x. Принимаем

$$\frac{(x-25)^2}{25}=t$$

тогда

$$\sqrt{\Sigma t}=\frac{\sqrt{\Sigma(x-25)^2}}{5}$$

и

$$R=100\left(1-\frac{\sqrt{\Sigma(x-25)^2}}{2\cdot 5}\right)=$$

$$=100\left(1-\frac{\sqrt{\Sigma t}}{10}\right)=100-10\sqrt{\Sigma t},\qquad (I.10)$$

где

$$\sum t = t_{I} + t_{II} + t_{III} + t_{IV} = \frac{(x_1 - 25)^2}{25} + \frac{(x_2 - 25)^2}{25} + \frac{(x_3 - 25)^2}{25} + \frac{(x_4 - 25)^2}{25}.$$

Кривая $R = f(x)$ построена в декартовой системе координат, где по оси абсцисс отложены значения Σt, а по оси ординат — R (в %) (см. рис. 1). Величина t при значениях x от 0 до 50% приведена в табл. 8.

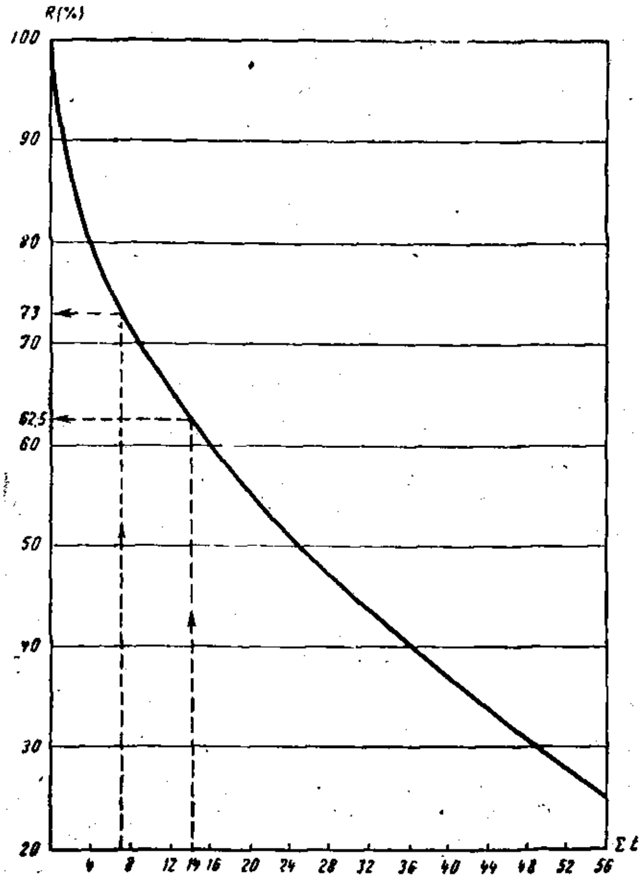


Рис. 1. Определение коэффициента ритмичности строительного производства (R).

Таблица 8

Значения t

x	t	x	t	x	t	x	t
0	25	13	5,8	26	0,04	39	7,8
1	23	14	4,8	27	0,2	40	9
2	21,2	15	4	28	0,4	41	10,2
3	19,4	16	3,3	29	0,6	42	11,6
4	17,6	17	2,6	30	1	43	12,9
5	16	18	2	31	1,4	44	14,4
6	14,4	19	1,4	32	2	45	16
7	12,9	20	1	33	2,6	46	17,6
8	11,6	21	0,6	34	3,3	47	19,4
9	10,2	22	0,4	35	4	48	21,2
10	9	23	0,2	36	4,8	49	23
11	7,8	24	0,04	37	5,8	50	25
12	6,8	25	0	38	6,8	—	—

Для определения величины R при любом распределении x по кварталам достаточно просуммировать значения t, соответствующие каждому x по табл. 8, затем на оси абсцисс отложить величину Σt и по ней найти на оси ординат значение R. Так, для $\Sigma t = 4 + 1 + 9 + 0 = 14$ (см. табл. 8) соответствующее значение R определяется по рис. 1, равным 62,5%, что лишь на 0,1% отличается от значения R, определенного аналитически.

При распределении по кварталам: $x_{\text{I}} = 17\%$; $x_{\text{II}} = 29\%$; $x_{\text{III}} = 20\%$; $x_{\text{IV}} = 34\%$, величина $\sum t = 2{,}6 + 0{,}6 + 1{,}0 + 3{,}3 = 7{,}5$ и $R = 73{,}0\%$.

При делении рассматриваемого периода времени на три отрезка (в исследовании ритмичности строительного производства в течение квартала помесячно или в течение месяца подекадно) предлагается пользоваться иной модификацией графического метода. Ход расчетов тот же, $\bar{x} = \frac{100}{3} = 33{,}3\%$.

$$R = \left(1 - \frac{\sqrt{\frac{\Sigma (x - 33{,}3)^2}{3}}}{33{,}3}\right) \cdot 100 = \left(1 - \frac{\sqrt{\Sigma (x - 33{,}3)^2}}{57{,}6}\right) \cdot 100\%.$$

Значения коэффициента ритмичности выполнения объемов работ в двух любых декадах месяца или двух любых месяцах квартала приведены в табл. 9. Предполагается, конечно, что $x_3 = [100 - (x_1 + x_2)] \cdot 100\%$. Например, коэффициент ритмичности выполнения работ в одном из месяцев, если в первой декаде было выполнено 15% месячного объема работ ($x_1 = 15$), во второй декаде — 30% ($x_2 = 30$), равен 50,7%. Такое же значение R получим, если его рассчитывать, например, по значениям $x_2 = 30\%$ и $x_3 = 100 - (15 + 30) = 55\%$ и по значениям x_1 и x_3. Значения R для величин x, не кратных 5, определяются интерполяцией. Так, для $x_1 = 17,5\%$ и $x_2 = 36\%$ $R = 62,8\%$.

Таблица 9

I интервал / II интервал	Определение R для трех интервалов времени. Значения R при объемах работ (%): 10	15	20	25	30	33,3	35	40	45	50	55
Значения R при объемах работ (%): 10	1,5	14,3	21,3	30,7	36,3	42,7	45	49	51,3	49	45
15	14,3	22,5	32,7	41,3	50,7	55,2	57	60,6	60,6	57	50,7
20	21,3	32,7	43,6	53,7	62,8	67,5	69,2	71,8	69,2	62,8	53,7
25	30,7	41,3	53,7	64,4	75,8	79,7	82,6	82,6	75,8	64,4	53,7
30	36,3	50,7	62,8	75,8	85,9	91,7	92,7	85,9	75,8	62,8	50,7
33,3	42,7	55,2	67,5	79,7	91,7	100,0	95,8	83,5	72,4	59,2	48,5
35	45	57	69,2	82,6	92,7	95,8	92,7	82,6	69,2	57	45
40	49	60,6	71,8	82,6	85,9	83,5	82,6	71,8	60,6	49	37,5
45	51,3	60,6	69,2	75,8	75,8	72,4	69,2	60,6	51,3	37,5	28
50	49	57	62,8	64,4	62,8	59,2	57	49	37,5	29,4	—
55	45	50,7	53,7	53,7	50,7	48,5	45	37,5	28	—	—

Определим L — предельные значения x по данным R. Задача сводится к нахождению максимальной величины (B) одной из переменных: x_1; x_2; x_3; x_4 из уравнения

$$\sqrt{\frac{\sum_{1}^{4} x_i^2}{4}} = B \qquad (1.11)$$

Геометрическая интерпретация степени ритмичности наглядна в том случае, когда в течение любых двух кварталов выполнение объемов работ равно нулю. Тогда искомые предельные значения располагаются на двух пересекающихся прямых, расположенных симметрично относительно нулевых отклонений от средних арифметических величин объемов работ (см. рис. 2).

Обозначая $x_1 = (25 - a)$; $x_2 = (25 - b)$; $x_3 = (25 - c)$; $x_4 = (25 - d)$, получим: $a + b + c + d = 100$. Так как переменные x_i равноправны, равновозможны, рассмотрим только одну их них — x_1. Пусть $x_{i\,\max} = k$, тогда

$$x_2 = x_3 = x_4 = 0 \ldots \tag{I.12}$$

Из уравнения (1.12) следует:

1) $b = c = d = 25$;

2) $a = 25$,

следовательно, и

$$x_1 = 0.$$

Положим,

$$x_3 = x_4 = 0, \quad \text{а} \quad x_1 = x_2 = x,$$

тогда уравнение (I.11) можно записать так:

$$\sqrt{\frac{2x^2}{4}} = \frac{x}{\sqrt{2}} = B,$$

откуда

$$x = \sqrt{2}\,B. \tag{1.13}$$

Из уравнения $R = 100 - \frac{B}{25}$ следует, что

$$B = 25\,(R - 100).$$

Из равенства (I.11) и (I.12)

$$x = 25\sqrt{2}\,(100 - R),$$

откуда

$$L = 25\left[\,1 \pm (100 - R)\sqrt{2}\,\right]\%.$$

При равномерном распределении x по кварталам наблюдается полная ритмичность процесса:

$$x_1 = x_2 = x_3 = x_4 = 25\,\%; \quad R = 100\,\%.$$

График значений L, представленный на рис. 2, наглядно показывает, что коэффициент ритмичности R характеризует степень отклонения величин x от средней величины. Для каждого значения R пределы, в которых могут находиться x, строго определены. Так, если $R = 73{,}6\%$, то значения x могут находиться только в пределах от ~ 13

до $\sim 37\%$; если $R = 36\%$, то эти значения могут колебаться от нуля до $\sim 54\%$.

Решение общей задачи о степени ритмичности как показателе предельных отклонений варьирующей величины при рассмотрении ее колеблемости в течение четырех пе-

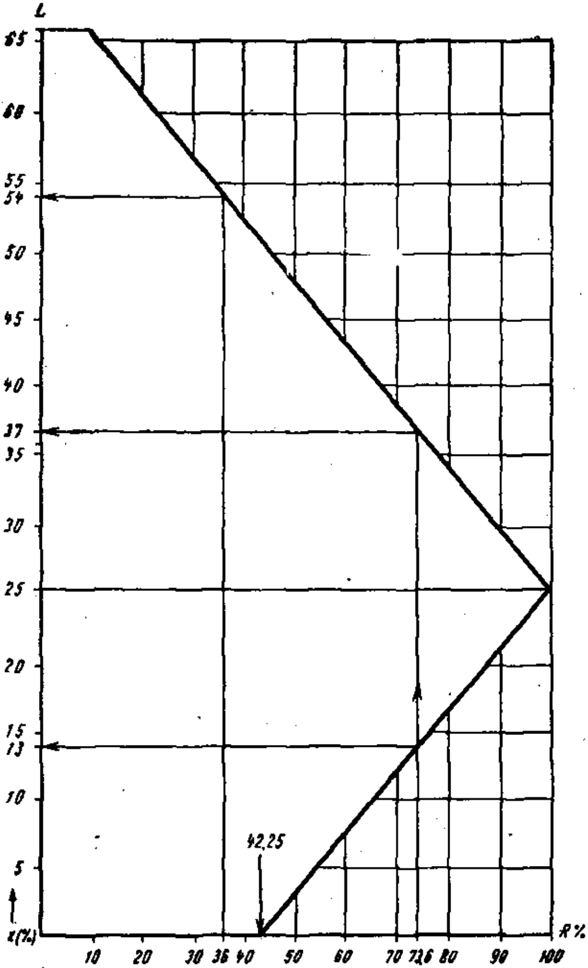


Рис. 2. Определение коэффициента ритмичности по отклонениям значений x от средней величины.

риодов времени сводится к следующему. Представим значение коэффициента ритмичности из уравнения (I.9) в долях единицы:

$$R = 1 - \frac{\sqrt{\sum (x_i - 25)^2}}{50},$$

где x_i — значение переменной величины, ритмичность которой определяется в один из четырех (i) периодов.

$$\sqrt{\sum (x_i - 25)^2} = 50(1 - R)$$

или

$$\sum (x_i - 25)^2 = 50^2 (1 - R)^2 \quad \text{(I.16)}$$

Обозначим $50^2(1-R)^2 = \tau^2$.

Из уравнения (I.16) имеем:

$$(x_1 - 25)^2 + (x_2 - 25)^2 + (x_3 - 25)^2 + (x_4 - 25)^2 = \tau^2$$

или

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 - 50(x_1 + x_2 + x_3 + x_4) + 4 \cdot 25^2 = \tau^2.$$

Но по условию

$$x_1 + x_2 + x_3 + x_4 = 100.$$

Поэтому

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = \tau^2 + 50^2. \quad \text{(I.17)}$$

Введем обозначение

$$\tau^2 + 50^2 = \Pi^2 \quad \text{(I.18)}$$

и перепишем (I.17) в виде

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = \Pi^2.$$

Из (I.16) $\Pi^2 = 50^2(1-R)^2 + 50^2$, откуда

$$\Pi = 50\sqrt{1 + (1 - R)^2}. \quad \text{(I.19)}$$

Определим, в каких пределах может находиться каждое из чисел x_i при данном значении R. Очевидно, что

$$x_{\min} \leqslant x_i \leqslant x_{\max} \quad (i = 1,\ 2,\ 3,\ 4).$$

Задачу достаточно решить только для одного x_i, например для x_1. Итак, необходимо найти наименьшее и наибольшее значения x_i при соотношениях:

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = \Pi^2 \quad \text{(I.20)}$$

$$x_1 + x_2 + x_3 + x_4 = 100 \quad \text{(I.21)}$$

Найдем x_1 из уравнения (I.20):

$$x_1=\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}.$$

Найдем x_1 из уравнения (I.21):

$$x_1=100-(x_2+x_3+x_4).$$

Таким образом, $x_1=f(x_2,\ x_3,\ x_4)$, которые связаны уравнением

$$\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}=100-(x_2+x_3+x_4). \qquad (I.22)$$

Найдем условный экстремум функции, выраженной уравнением (I.22), при условии:

$$\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}-100+(x_2+x_3+x_4)=0.$$

Составим вспомогательную функцию Лагранжа:

$$\Phi(x_2,\ x_3,\ x_4)=\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}-$$
$$-\lambda\left(100-x_2-x_3-x_4-\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}\right)$$

или

$$\Phi(x_2,\ x_3,\ x_4)=(1+\lambda)\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}-$$
$$-\lambda(100-x_2-x_3-x_4).$$

Приравниваем нулю частные производные этой функции:

$$\frac{\partial\Phi}{\partial x_2}=(1+\lambda)\frac{-x_2}{\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}}+\lambda=0;$$

$$\frac{\partial\Phi}{\partial x_3}=(1+\lambda)\frac{-x_3}{\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}}+\lambda=0;$$

$$\frac{\partial\Phi}{\partial x_4}=(1+\lambda)\frac{-x_4}{\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)}}+\lambda=0.$$

Откуда

$$x_2=\frac{\lambda}{1+\lambda}\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)};$$

$$x_3=\frac{\lambda}{1+\lambda}\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)};$$

$$x_4=\frac{\lambda}{1+\lambda}\sqrt{\Pi^2-(x_2^2+x_3^2+x_4^2)},$$

т. е. в экстремальных точках $x_2 = x_3 = x_4 = \omega$, и для экстремальных значений x_i справедливы равенства $x_1^2 + 3\omega^2 = \Pi^2$ (см. уравнение I.20); $x_1 + 3\omega = 100$ (см. уравнение I.21), откуда

$$\omega = \frac{100 - x_1}{3}$$

и

$$x_1^2 + 3\,\frac{(100 - x_1)^2}{9} = \Pi^2$$

или

$$3x_1^2 + 100^2 - 200\,x_1 + x_1^2 = 3\Pi^2.$$

Группируем:

$$4x_1^2 - 200\,x_1 - (3\Pi^2 - 100^2) = 0$$

или

$$2x_1^2 - 100\,x_1 - \frac{3\Pi^2 - 100^2}{2} = 0,$$

откуда

$$x_1 = \frac{100 \pm \sqrt{100^2 + 4\,(3\Pi^2 - 100^2)}}{4} = 25 \pm \frac{\sqrt{12\Pi^2 - 3\cdot 100^2}}{4}.$$

Подставляя Π^2 из уравнения (I.19), имеем:

$$x_1 = 25 \pm \frac{1}{4}\sqrt{12\cdot 50^2\,(1 - R)^2 + 12\cdot 50^2 - 3\cdot 100^2} =$$

$$= 25 \pm 25\,(1 - R)\sqrt{3}.$$

Окончательно:

$$x_{max} = 25 + 25\,(1 - R)\sqrt{3} = 25\left[1 + \sqrt{3}\,(1 - R)\right] \ldots \quad (I.23)$$

$$x_{min} = 25\left[1 - \sqrt{3}\,(1 - R)\right] \ldots \quad (I.24)$$

Наибольшее значение признак может иметь при $R = 0$; тогда

$$x_{max} = 25\cdot\left(1 + \sqrt{3}\right) = 25\cdot 2{,}73 = 68{,}25\,\%.$$

Из (I.24) очевидно, что наименьшее значение коэффициент ритмичности может иметь при $x_{min} = 0$; тогда

$$\sqrt{3}\,(1 - R) = 1$$

и

$$R = \frac{\sqrt{3} - 1}{\sqrt{3}} = 0{,}4225 = 42{,}25\,\%$$ (см. рис. 2).

Следовательно, графический метод определения показателя ритмичности дает положительные значения R в пределах от 42,25% до 100%. Это соответствует колебаниям объемов работ в отдельных кварталах от 0 до 68,25% и, таким образом, охватывает подавляющее большинство колебаний. При меньших значениях R нижний предел отклонений x_{min} принимается равным нулю.

Коэффициент ритмичности не фиксирует какое-либо одно определенное значение выполнения объемов работ (или расхода ресурсов) по кварталам года, а указывает предельные значения, в которых могут находиться эти величины. Например, значение $R = 0,7$ говорит о том, что величины объемов работ (или расхода каких-либо ресурсов) должны находиться в пределах не меньше 12% и не более 38% в любом из четырех кварталов года.

В строительном производстве следует рассматривать следующие виды ритмичности: использования рабочей силы, потребления материальных ресурсов, сдачи готовых объектов в эксплуатацию, выполнения объемов работ (в целом или по отдельным видам работ, в денежном или натуральном выражении).

Экономический эффект повышения уровня ритмичности строительного производства проявляется в росте производительности труда, снижении себестоимости. На основании корреляционного анализа установлено, что повышение на 1% ритмичности объемов работ, выполняемых собственными силами, приводит к снижению себестоимости возведения промышленных объектов от 0,12 до 0,23% и повышению выработки на одного работающего от 0,02 до 0,05 тыс. руб. в год.

§ 5. Кривая распределения

Наиболее наглядное представление о рядах распределения дает графический метод. В декартовой системе координат на оси абсцисс откладываем значения признака (варианты), на оси ординат — частоты.

Площади прямоугольников, пропорциональные высотам, изображающим частоты вариантов, называются *гистограммой распределения*. Ординаты гистограммы представляют собой частоту, приходящуюся на единицу измерения признака, и называются *плотностью частоты*. Плавная кривая, дающая изображение интервального ряда и характеризующая изменение плотности распределения, называется *кривой распределения.* Ряд распределения объемов работ строительно-монтажных управлений за год, приведенный в табл. 2, графически изображен на рис. 3.

Для каждой статистической совокупности характерен определенный, присущий лишь ей тип кривой распределения; два наиболее существенных показателя каждой совокупности — средняя арифметическая ($\bar{x}$) и среднее квадра-

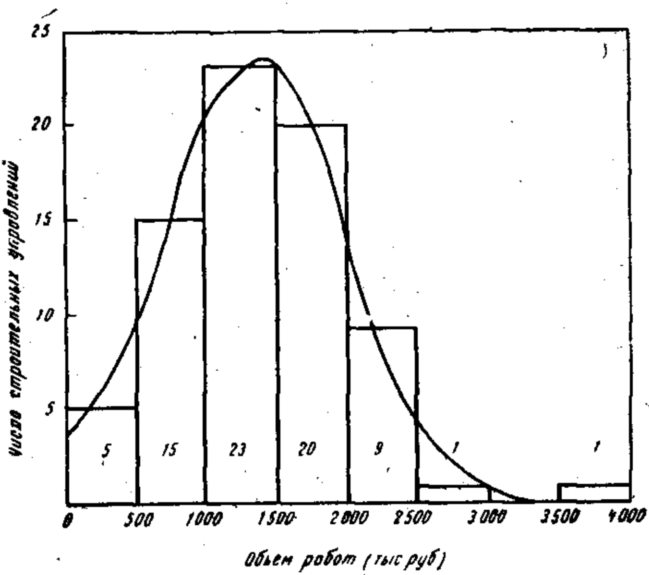


Рис. 3. Гистограмма распределения объемов работ строительно-монтажных управлений за год.

тическое *отклонение* (σ) — характеризуют положение кривой распределения по отношению к оси абсцисс и рассеяние кривой (дисперсию σ^2).

Любое искажение кривой указывает на нарушение условий проявления распределения: чаще всего на форму кривой оказывает влияние неоднородность исходного материала. О неоднородности материала сигнализирует появление второй вершины, случайных зигзагов кривой распределения.

Проиллюстрируем форму кривой распределения численности административно-хозяйственного персонала межколхозных строительных организаций УССР (см. рис. 4).

Для выявления закономерностей экономических явлений используются укрупнение интервалов и математическое выравнивание кривой распределения. Принцип укрупнения интервалов можно оценить, сопоставив рис. 4 и рис. 5 (см. кривую *б*), на котором тот же исходный материал разбит на более мелкие интервалы. Говоря о математическом выравнивании, имеют в виду приближение полу-

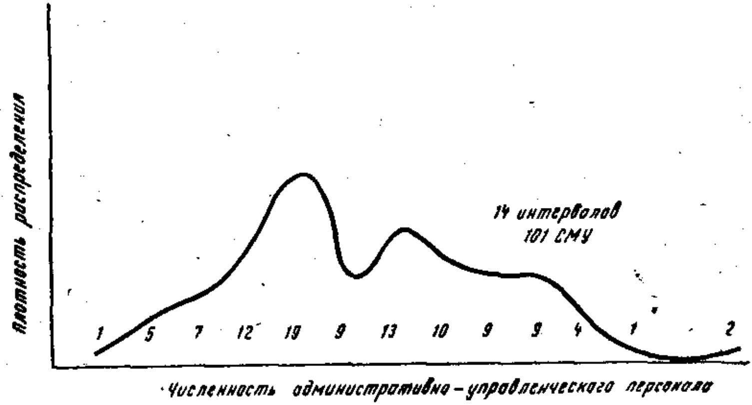


Рис. 4. Кривая распределения численности административно-хозяйственного персонала межколхозных строительных организаций УССР

ченной эмпирической кривой к кривой нормального распределения.

Закон нормального распределения характерен для большинства процессов экономики строительства. Исключение составляют, в частности, некоторые процессы эксплуатации и ремонта машин, подчиняющиеся экспоненциальному распределению.

А. Я. Хинчин [21] отмечал, что если интересующий нас признак можно рассматривать как результат суммарного действия факторов, каждый из которых мало связан с большинством остальных, и влияние каждого фактора на результат намного перекрывается суммарным влиянием всех остальных факторов, то при большом числе таких факторов распределение признака становится близким к закону нормального распределения.

Закон нормального распределения выражается формулой

$$\varphi(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)^2}, \qquad (1.32)$$

где x — вариант признака;

$\varphi(x)$ — соответствующая плотность вероятности, численно равная плотности частоты ряда;

$e = 2{,}71828...$ $\pi = 3{,}1416...$ — константы математического анализа;

$\bar{x}$ — средняя арифметическая величина признака.

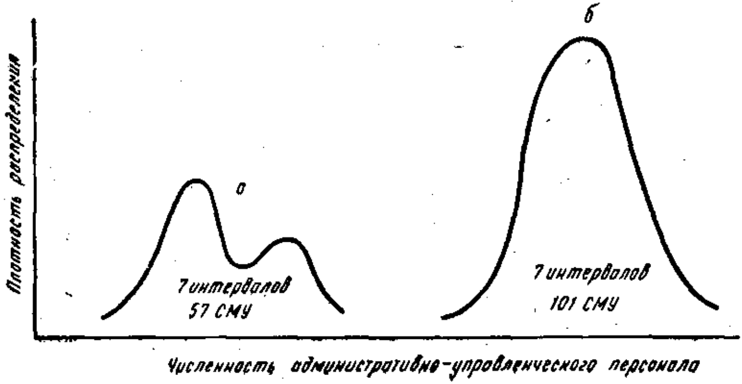


Рис. 5. Изменение формы кривой распределения численности административно-управленческого персонала при различном числе наблюдений.

При одновременном разнонаправленном воздействии на какой-либо признак нескольких факторов, каждый из которых подчиняется определенному закону распределения, суммарное их влияние также подчиняется закону нормального распределения.

Кривая нормального распределения (как сокращенно называют кривую закона нормального распределения) представлена на рис. 6. Это колоколообразная кривая, симметричная относительно вертикальной оси, проходящей через точку с абсциссой $x = \bar{x}$ (на рис. x_0) и максимальной ординатой в этой точке. По обе стороны от $\bar{x}$ откладываем отрезки, равные σ (на рис. σ_0); значения ординат для этих отрезков — 0,004; 0,054; 0,242; 0,399 приведены на рис. 6.

Укажем простейший способ построения нормальной кривой. Абсциссы ее: $x = \bar{x} \pm a\sigma$, где a — коэффициент, выражающий x в долях σ; ординаты ее:

$$y = h\ H,$$

где $H = \frac{0,3989 in}{\sigma}$ — максимальная ордината. Значения h принимаются по табл. 10, выборочно репродуцированной из книги Н. Л. Леонтьева «Техника статистических вычислений» (М., «Лесная промышленность», 1966, стр. 55).

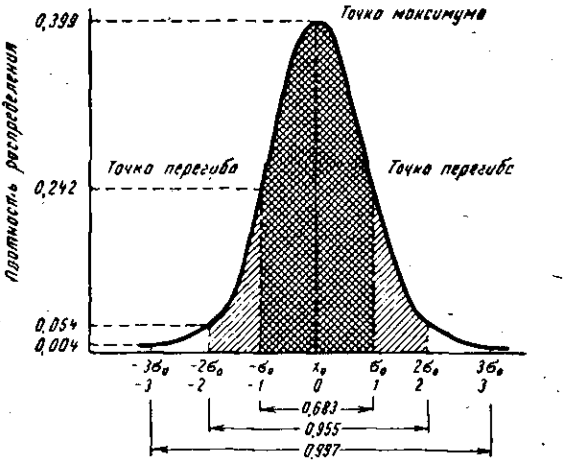


Рис. 6. Кривая нормального распределения.

При $i = 500$; $n = 74$; $\sigma = 634$; $H = \frac{0,3989 \cdot 74 \cdot 500}{634} = 23,3$, остальные ординаты соответственно равны 20,6; 14,2; 7,6; 3,2; 1,0; 0,2.

С изменением параметра $\bar{x}$ кривая перемещается по оси абсцисс, сохраняя свою форму; с изменением параметра σ кривая меняет форму относительно оси абсцисс (см. рис. 7).

В частном случае, когда абсцисса $\bar{x} = 0$ и $\sigma = 1$, уравнение (I.32) упрощается:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (I.33)$$

Таблица 10

Значения «h»

a в долях σ	h	a в долях σ	h
0,1	0,995	2,0	0,135
0,5	0,883	2,5	0,044
1,0	0,607	3,0	0,011
1,5	0,325		

и называется стандартным уравнением нормальной кривой. Уравнение (I.32) всегда можно привести к стандартной форме путем замены $\frac{x-\bar{x}}{\sigma}$ на t, которое показывает, на сколько сигм (σ) отклоняется данное значение x от средне-

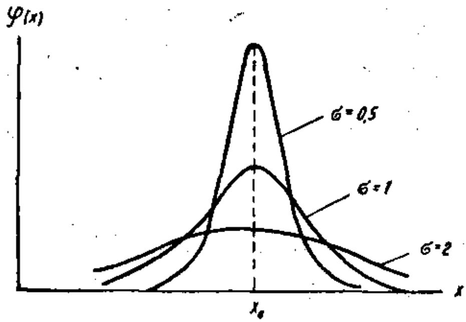


Рис. 7. σ как показатель рассеяния кривой распределения.

го $\bar{x}$. Таким образом, уравнение (I.32) преобразуется[1]:

$$\varphi(t)=\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}.$$

Построение нормальной кривой проследим на примере ряда распределения объемов работ, выполненных жилищно-строительными управлениями комбината Донецкжилстрой (табл. 11).

[1] $\varphi(t)$ для ряда значений t приводится в специальных таблицах; рекомендуются пособия [5], [6], [11].

Таблица 11

Расчет ординат кривой и гистограммы нормального распределения

Середины интервалов	Абсолютные частоты	Относительные частоты p_i	Нормированные середины интервалов h_i	Ординаты нормальной кривой	Ординаты гистограммы h'
1	2	3	4	5	6
250	5	0,068	—1,80	0,0790	0,086
750	15	0,203	—1,01	0,2396	0,257
1250	23	0,311	—0,22	0,3894	0,398
1750	20	0,270	0,57	0,3391	0,342
2250	9	0,122	1,36	0,1582	0,154
2750	1	0,013	2,15	0,0396	0,0164
3250	—	—	2,94	0,0053	—
3750	1	0,013	3,73	0,0036	0,0164
Итого	74	1,000			

Данные гр. 4 — нормированные середины интервалов находим по формуле

$$h_i = \frac{x_i - \bar{x}}{\sigma} = t_i.$$

Отсюда

$$h_1 = \frac{250 - 1392}{634} = -1{,}80.$$

h_2 вычисляем по формуле

$$h_2 = h_1 + \Delta h.$$

Для дальнейшего расчета необходимо вычислить Δh:

$$\Delta h = \frac{i}{\sigma} = \frac{500}{634} = 0{,}79;$$

$h_2 = -1{,}80 + 0{,}79 = -1{,}01$. По значениям $h_i = h_1 \ldots h_8$ в специальных таблицах[1], приведенных в пособиях по математической статистике, находим соответствующие величины функции $\varphi(t)$, показывающие ординаты нормальной кривой в нормированном масштабе и проставляем их в гр. 5 табл. 11.

[1] См., например, И. Г. Венецкий, Г. С. Кильдишев. Основы теории вероятностей и математической статистики. Приложение II, стр. 341.

Ординаты гистограммы в нормированном масштабе h' находим из уравнения $h' = \frac{p_i}{\Delta h}$, например

$$h'_3 = \frac{0,311}{0,79} = 0,398; \quad h'_5 = \frac{0,122}{0,79} = 0,154 \text{ и т. д.}$$

Для сопоставления кривых распределения реальной и нормальной служат показатели асимметрии и эксцесса.

Смещение вершины реальной кривой по отношению к нормальной вдоль оси абсцисс влево называется положительной асимметрией, а вправо — отрицательной. Для количественной оценки асимметрии применяется *показатель асимметрии*, или показатель скошенности — A.

Отрицательная асимметрия указывает на преобладание вариантов со сравнительно большими значениями.

При смещении вершины реальной кривой вверх по отношению к вершине нормальной кривой отмечается наличие *положительного эксцесса*; при смещении вершины вниз — наличие *отрицательного эксцесса*. Для количественной оценки эксцесса служит показатель эксцесса — E.

Положительный эксцесс свидетельствует о том, что большинство вариантов находятся в середине ряда распределения; при отрицательном эксцессе варианты скапливаются на границах крайних интервалов или распределены сравнительно равномерно и это придает кривой плосковершинность; при особенно резком отрицательном эксцессе наблюдается наличие двух вершин и даже разделение кривой на две самостоятельные ветви. Показатели асимметрии и эксцесса вычисляют по формулам:

$$A = \sqrt{\frac{M_3^2}{M_2^3}} = \frac{M_3}{\sqrt{M_2^3}} \tag{I.34}$$

$$E = \frac{M_4}{M_2^2} - 3, \tag{I.35}$$

где M_2, M_3, M_4 — центральные моменты соответствующих порядков.

Средняя ошибка показателя асимметрии

$$\mu_A = \pm\sqrt{\frac{6}{n}} \tag{I.36}$$

Средняя ошибка показателя эксцесса

$$\mu_E = \pm\sqrt{\frac{24}{n}} = \pm 2\mu_A. \qquad (1.37)$$

Если $\frac{A}{\mu_A}$ и $\frac{E}{\mu_E} < 3$, то асимметрия и эксцесс недостоверны и не имеют существенного значения, эмпирическая кривая распределения достаточно близка к нормальной. Превышение предела, равного 3, указывает на достоверность, стабильность такого положения, приводит к заключению, что такой характер кривой распределения отражает существо рассматриваемых явлений. При резком повышении A и E необходимо добиться намного большей однородности исходного материала, отбрасывая значения признака, особенно отличающиеся от средних значений, конечно обосновывая это качественным анализом существа дела. Строительно-монтажные управления, деятельность которых нами анализируется, специализированы на возведении жилых поселков для строящихся шахт, находящихся на значительных расстояниях друг от друга. Распыление работ по многочисленным мелким поселкам не способствовало, естественно, концентрации всех ресурсов (финансовых, рабочей силы и механизмов) в отдельных СМУ. Поэтому крупных подразделений было мало.

Статистическую обработку рядов распределения, т. е. определение $\bar{x}$ и σ, рекомендуется выполнять по вспомогательной таблице (см. табл. 12), в которой используется перевод реальных величин в условные и отсчет от условного нуля, значительно облегчающие все вычисления.

Табл. 12 расчерчивается на столько граф, сколько интервалов имеет ряд распределения (в нашем примере на 8); значения границ интервалов и середины их взяты из ряда распределения. Серединой условных интервалов (условным нулем) принимаем интервал с максимальной частотой; от него влево идут отрицательные, вправо — положительные значения. Строки имеют следующие значения:

а — абсолютные частоты n_x ряда распределения;

б — каждое значение n_x, умноженное на соответствующую величину условного интервала, т. е. $n_x \cdot x'$: $5(-2) = -10$; $15(-1) = -15$ и т. д.;

в — значения строки *б*, умноженные на соответствующие величины условных интервалов, т. е. $n_x \cdot (x')^2$;

Таблица 12

Схема расчета $\bar{x}$ и σ с помощью отсчета от условного нуля

Середины условных интервалов		−2	−1	0	1	2	3	4	5	Σ	Начальные моменты σ
Середины реальных интервалов		250	750	1250 C_0	1750	2250	2750	3250			
Реальные интервалы		500	1000	1500	2000	2500	3000	3500			
а	n_x	5	15	23	20	9	1	—	1	74	
б	$n_x \cdot x'$	−10	−15	0	20	18	3	—	5	21	$v_1 = \frac{21}{74}$
в	$n_x \cdot x'^2$	20	15	0	20	36	9	—	25	125	$v_2 = \frac{125}{74}$
г	$n_x \cdot x'^3$	−40	−15	0	20	72	27	—	125	189	$v_3 = \frac{189}{74}$
д	$n_x \cdot x'^4$	80	15	0	20	144	81	—	625	965	$v_4 = \frac{965}{74}$

г — значения строки *в*, умноженные на соответствующие величины условных интервалов, т. е. $n_x \cdot (x')^3$;

д — значения строки *г*, умноженные на соответствующие величины условных интервалов, т. е. $n_x \cdot (x')^4$.

В графе Σ приведены алгебраические суммы всех строк, которые входят в приведенные ниже формулы:

$$\bar{x} = \bar{x}' + C_0' = \frac{б}{n}\, i + C_0; \tag{I.38}$$

$$\bar{x} = \frac{21}{74} \cdot 500 + 1250 = 1392;$$

$$\sigma' = \sqrt{\frac{в}{n} - \left(\frac{б}{n}\right)^2}; \tag{I.39}$$

$$\sigma' = \sqrt{\frac{125}{74} - \left(\frac{21}{74}\right)^2} = 1{,}269.$$

Для перевода в реальные величины умножим σ' на i: $1{,}269 \cdot 500 = 634$. Сопоставив эти простейшие действия с громоздкими вычислениями, которые необходимо было выполнить для определения σ по табл. 5, отметим рациональность исчисления σ через σ^2 и использования II и III свойств $\bar{x}$ и σ (см. стр. 14). И наконец, в последней графе приведены начальные моменты v, определяемые по формулам:

$$v_1 = \frac{\sum б}{\sum a};$$
$$v_2 = \frac{\sum в}{\sum a};$$
$$v_3 = \frac{\sum г}{\sum a};$$
$$v_4 = \frac{\sum д}{\sum a}.$$

В рассматриваемом примере:

$$v_1 = \frac{21}{74} = 0{,}284; \quad v_2 = \frac{125}{74} = 1{,}69;$$
$$v_3 = \frac{189}{74} = 2{,}56; \quad v_4 = \frac{965}{74} = 13{,}05.$$

С помощью следующих формул переходим к определению центральных моментов:

второго порядка

$$M_2 = v_2 - v_1^2;$$

третьего порядка

$$M_3 = v_3 - 3 \cdot v_2 \cdot v_1 + 2v_1^3;$$

четвертого порядка

$$M_4 = v_4 - 4v_1 \cdot v_3 + 6v_1^2 \cdot v_2 - 3v_1^4.$$

Предварительно найдем степени v_1:

$$v_1^2 = 0{,}081; \quad v_1^3 = 0{,}022; \quad v_1^4 = 0{,}006.$$

Затем подставим их в формулы моментов:

$$M_2 = 1{,}69 - 0{,}081 = 1{,}609;$$
$$M_3 = 2{,}56 - 3 \cdot 1{,}69 \cdot 0{,}284 + 2 \cdot 0{,}022 = 1{,}164;$$
$$M_4 = 13{,}05 - 4 \cdot 0{,}284 \cdot 2{,}56 +$$
$$+ 6 \cdot 0{,}081 \cdot 1{,}69 - 3 \cdot 0{,}006 = 10{,}945.$$

Подставляя значения M_2, M_3, M_4 в формулы (1.34 — 1.37), получаем:

$$A = \sqrt{\frac{1,164^2}{1,609^3}} = 0,573; \quad E = \frac{10,928}{1,609^2} - 3 = 1,22;$$

$$\mu_A = \sqrt{\frac{6}{74}} = 0,285; \quad \mu_E = 2 \cdot 0,285 = 0,570;$$

$$\frac{A}{\mu_A} = \frac{0,573}{0,285} = 2,01 < 3; \quad \frac{E}{\mu_E} = \frac{1,22}{0,570} = 2,14 < 3.$$

Следовательно, эмпирическая кривая распределения объемов работ достаточно близка к нормальной.

В заключение рассчитаем исходные статистические параметры, необходимые для определения центральных моментов по данным, приведенным в известной монографии проффессора Я. И. Лукомского [11, стр. 354).

Ряд, характеризующий изменение суммы накладных расходов по 113 шахтостроительным управлениям комбината «Донецкшахтострой», имеет абсолютные частоты: 3, 25, 42, 29, 10, 2, 2.

Таблица 13

Расчет параметров при определении начальных моментов

	−2	−1	0	1	2	3	4	Σ	v
а	3	25	42	29	10	2	2	113	
б	−6	−25	0	29	20	6	8	32	$v_1 = 0,283$
в	12	25	0	29	40	18	32	156	$v_2 = 1,38$
г	−24	−25	0	29	80	54	128	242	$v_3 = 2,14$
д	48	25	0	29	160	162	512	936	$v_4 = 8,03$

$$v_1^2 = 0,080; \quad v_1^3 = 0,021; \quad v_1^4 = 0,06;$$

$$M_2 = 1,38 - 0,289^2 = 1,30;$$

$$M_3 = 2,14 - 3 \cdot 1,38 \cdot 0,283 + 2 \cdot 0,021 = 1,012;$$

$$M_4 = 8,03 - 4 \cdot 0,283 \cdot 2,14 + 6 \cdot 0,080 \cdot 1,38 - 3 \cdot 0,006 = 6,253;$$

$$A = \frac{1,012}{\sqrt{1,3^3}} = \frac{1,012}{1,485} = 0,672; \quad \mu_A = \sqrt{\frac{6}{113}} = 0,231;$$

$$\frac{A}{\mu_A} = \frac{0,672}{0,231} = 2,9 < 3; \quad E = \frac{6,253}{1,3^2} - 3 = 0,7;$$

$$\mu_E = 2 \cdot 0,231 = 0,462; \frac{E}{\mu_E} = \frac{0,7}{0,462} = 1,51 < 3.$$

§ 6. Критерии согласия

В некоторых случаях необходимо иметь обобщающую количественную оценку *близости* эмпирического распределения к теоретическому нормальному распределению (либо распределению Максвелла, Пуассона, экспоненциальному и др.). Чтобы исключить элемент субъективности в оценке близости, пользуются одним из *критериев согласия*, предложенных разными учеными: К. Пирсоном, А. Н. Колмогоровым, Б. С. Ястремским, В. И. Романовским.

Применим критерий А. Н. Колмогорова к оценке близости распределения, представленного в табл. 14, к нормальному распределению.

По критерию λ Колмогорова близость теоретического и эмпирического распределений устанавливается сравнением их интегральных (накопленных) распределений:

$$\lambda = \frac{D}{\sqrt{n}},$$

где D — максимальная величина разности накопленных теоретических M' частот и эмпирических M частот.

Для того чтобы эмпирическое и теоретическое распределения соответствовали друг другу, необходимо условие: $\lambda \leqslant 1{,}3 \div 1{,}4$.

В данном примере $\lambda = \frac{1}{\sqrt{74}} = 0{,}119 < 1{,}3$ указывает, что эмпирические частоты достаточно близки к теоретическим.

Проверим близость эмпирического и теоретического распределений по критерию Б. С. Ястремского:

$$I \leqslant 3\sqrt{2n - 4\theta},$$

где

$$I = |C - n|;\ C = \sum \frac{(m - m')^2}{m' \cdot q};\ \ q = 1 - p$$

p — относительные частоты $= \frac{m_i}{\sum m_i}$;

n — число интервалов (групп); при $n < 20$, $\theta = 0{,}6$.

Вычислим критерий согласия I (см. табл. 15).

Следовательно, данное эмпирическое распределение близко к нормальному. Рассмотрим еще пример применения кри-

Таблица 14

Расчет критерия согласия λ А. Н. Колмогорова

x	m	m'	M	M'	$\|M-M'\|$
1	2	3	4	5	6
250	5	5	5	5	0
750	15	14	20	19	1
1 250	23	23	43	42	1
1 750	20	20	63	62	1
2 250	9	9	72	71	1
2 750	1	2	73	73	0
3 250	—	—	73	73	0
3 750	1	—	74	73	1
Итого	74	73			

Примечание. Значения гр. 3 находим, умножая ординаты нормальной кривой (см. табл. 11, гр. 5) на $\frac{n \cdot i}{\sigma} = \frac{74 \cdot 500}{634} = 58,4$ с округлением до целых; в результате $\Sigma m' \neq \Sigma m$.

Таблица 15

Расчет критерия согласия Б. С. Ястремского

x	m	m'	p	q	$(m-m')^2$	$m' \cdot q$	$\frac{(m-m')^2}{m' \cdot q}$
1	2	3	4	5	6	7	8
250	5	5	0,066	0,934	0	0	0
750	15	14	0,184	0,816	1	0,816	1,225
1 250	23	23	0,302	0,698	0	0	0
1 750	20	20	0,264	0,736	0	0	0
2 250	9	10	0,132	0,868	1	0,868	1,152
2 750	1	2	0,026	0,974	1	0,974	1,027
3 250	—	—	—	—	—	—	—
3 750	1	2	0,026	0,974	1	0,974	1,027
Итого	74	76	1				4,431=C

$$I = 4,431 - 8 = 3,569; \qquad 3,569 < 3\sqrt{16 + 4 \cdot 0,6}.$$

терия согласия А. Н. Колмогорова с использованием геометрической интерпретации[1].

В 1962—1965 гг. нормативно-исследовательской станцией треста «Стальконструкция» проводились хронометражные наблюдения использования (по времени) башенных кранов на промышленном строительстве. Данные для построения ряда распределения по 525 наблюдениям приведены в табл. 16.

Таблица 16

Ряд распределения использования башенных кранов по времени

Простои (часов в смену), интервалы	Середины интервалов x	Абсолютные частоты m	Относительные частоты p
0,01—1	0,5	330	0,63
1—2	1,5	110	0,21
2—3	2,5	50	0,095
3—4	3,5	20	0,038
4—5	4,5	8	0,015
5—6	5,5	4	0,007
6—7	6,5	3	0,005
Итого		525	1

На рис. 8*а* пунктиром указана кривая, воспроизводящая изменение p в зависимости от x; это распределение простоев внешне близко к распределению Пуассона, отчетливо видно большое количество кратковременных простоев и сравнительно незначительное — длительных. Однако закон Пуассона может быть применен для характеристики варьирования только целочисленных, дискретных величин, в то время как в рассматриваемом случае величина простоев меняется непрерывно, смежные значения величины простоев могут отличаться на сколь угодно малую величину. Поэтому наиболее подходящим в данном случае оказался экспоненциальный закон распределения, характеризующийся уравнением связи типа $y = e^{ax}$.

Методом аппроксимации установлена зависимость $y = e^{-0,94x}$, интегральная форма распределения которой приведена на рис. 8*б*. Максимальная разница ординат накопленных, эмпирических (пунктирная кривая) и теоретиче-

---

[1] Работу выполнили сотрудники НИИСП Госстроя УССР А. А. Лившиц и Ж. В. Клейнер.

ских (сплошная кривая) относительных частот, измеренная по масштабу, равна $D = 0{,}0295$. Тогда $\lambda = D\sqrt{n} = 0{,}0295 \cdot \sqrt{525} = 0{,}68$; $0{,}68 < 1{,}3$ — распределение близко к нормальному. Для определения коэффициента использования парка башенных кранов треста определим вероятность простоев в интервалах до 2,5 часов, составляющих более 93% всех простоев: $\int\limits_{0,1}^{2,5} e^{-0,94x} = 0{,}82$. Таким образом, из 100 башенных кранов 82 имели простои до 2,5 часов в смену.

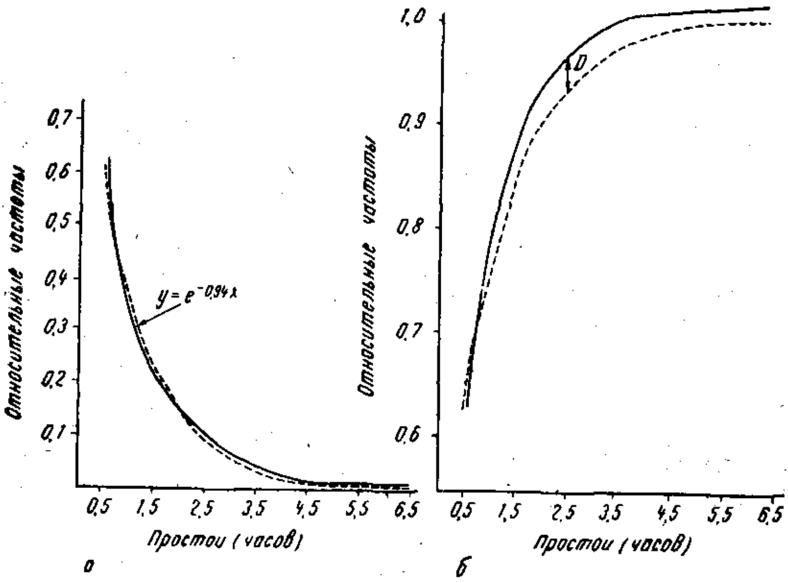


Рис. 8. Распределение простоев башенных кранов на промышленном строительстве:

а — распределение относительных частот по экспоненциальному закону; *б* — интегральная форма распределения относительных частот.

Один из критериев согласия χ^2 (*хи*-квадрат), предложен К. Пирсоном:

$$\chi^2 = \sum_{i=1}^{k} \frac{|n \cdot p_i - m_i|^2}{n \cdot p_i},$$

где m_i — текущие значения абсолютных частот;

n — объем совокупности;

p_i — разности последовательных значений нормальной функции распределения $\Phi_{h_{i+1}} - \Phi_{h_i}$.

Определение χ^2 проиллюстрируем данными уже рассмотренного выше примера (см. табл. 2).

Значения χ^2 находим по одному из пособий теории вероятностей (см. [5], [6]) в зависимости от величин h_i (графа 2, табл. 17), определяемым из выражения $h = \frac{i - \bar{x}}{\sigma}$,

$$h_1 = \frac{0 - 1392}{634} = -2{,}20; \quad h_2 = \frac{500 - 1392}{634} \ldots;$$

$$h_9 = \frac{4000 - 1392}{634} = 4{,}12.$$

Значения Φ_{h_i} для h_i равным 2,54 и 3,32, находим с помощью интерполяции — 0,9945 и 0,9996; $n = 74$.

Таблица 17

Расчет критерия χ^2

Границы интервалов	h_i	Φ_h	p_i	np_i	m_i	$\lvert np_i - m_i \rvert$	$\lvert np_i - m_i \rvert^2$	$\frac{\lvert np_i - m_i \rvert^2}{n \cdot p_i}$
1	2	3	4	5	6	7	8	9
0	—2,20	0,0139	0,0654	4,8	5	0,2	0,04	0,0083
500	—1,41	0,0793	0,1883	13,9	15	1,1	1,21	0,0870
1 000	—6,62	0,2676	0,2999	22,2	23	0,8	0,64	0,0288
1 500	0,17	0,5675	0,2640	19,5	20	0,5	0,25	0,0128
2 000	0,96	0,8315	0,1276	9,4	9	0,4	0,16	0,0170
2 500	1,74	0,9591	0,0354	2,6	1	1,3	1,69	0,6500
3 000	2,54	0,9945	0,0051	0,4	0	0,4	0,16	0,4000
3 500	3,32	0,9996	0,0004	0,0	1	1,0	1,00	0,0000
4 000	4,12	1,0000						$\chi^2 = 1{,}2039$

По одной из таблиц значений χ^2 (см. Е. С. Вентцель, стр. 297) по числу степеней свободы и χ^2 находим $p(\chi^2)$, т. е. вероятность χ^2; в данном случае число степеней свободы равно $8 - 3 = 5$; здесь 8 — число интервалов ряда распределения, 3 — число наложенных связей.

С помощью интерполяции находим $p(\chi^2) = 0{,}928$.

Таким образом, с вероятностью, близкой к 93%, можно утверждать, что распределение частот (5, 15, 23, 20, 9, 1, 0, 1) по закону нормального распределения можно считать правдоподобным.

Глава II

ИССЛЕДОВАНИЕ ПАРНЫХ КОРРЕЛЯЦИОННЫХ СВЯЗЕЙ

§ 1. Область применения корреляционного анализа и его методика

Мы рассмотрели вопросы, связанные с изменчивостью, варьированием какой-либо переменной величины, которую условились называть признаком. Такими величинами были: объем строительно-монтажных работ, процент выполнения норм выработки, размер выработки, доля участия общестроительных организаций в общем объеме работ при возведении группы однородных объектов металлургического производства. Мы научились выявлять закономерности в колеблемости признака, количественно измерять ее и оценивать степень близости эмпирического распределения вариантов к теоретическому. Но при этом мы абстрагировались от взаимосвязи явлений, от общеизвестного положения о том, что изменчивость любого признака неизбежно зависит от влияния на него множества причин. Описывая, например, ряд распределения, характеризующий выработку на одного работника, мы не упоминали о влиянии на ее величину условий организации труда, уровня специализации, механизации работ и ряда других факторов. Каждый из этих факторов можно рассматривать как отдельный признак, изменчивость которого следует характеризовать присущим ему рядом распределения. Для того чтобы выяснить, как все эти факторы влияют на выработку, вместе или порознь, дать количественную оценку и геометрическую интерпретацию этого влияния, необходимы методы корреляционного анализа.

Методы теории корреляции рекомендуется применять для установления зависимостей технико-экономических показателей от отдельных факторов при планировании, нормировании затрат ресурсов, а также в научно-исследовательских работах в качестве инструмента, позволяющего количественно оценивать взаимодействие экономических факторов в сфере строительного производства.

В строительном производстве низовой ячейкой, сохраняющей хозяйственную самостоятельность и имеющей четко выраженную специализацию, являются строительно-монтажные управления (СМУ) или соответствующие им управления начальника работ (УНР) в некоторых подрядных строительных трестах.

Применение методов корреляции для анализа работы одного треста встречает затруднения из-за недостатка исходных данных. Если минимальное количество однородных единиц статистического материала (т. е. число СМУ) принять равным 15—20, то при среднем их количестве (4—7 СМУ в общестроительных трестах) какие-либо выводы можно сделать только при наличии данных не менее чем за 3—4 года. Поэтому корреляционный метод целесообразен при анализе хозяйственной деятельности крупных строительных организаций — комбинатов, главков. При этом строительно-монтажные управления, тресты или иные подразделения, включаемые в наблюдения, должны быть однородными.

Однородные строительные подразделения характеризуются следующим.

1. Однотипным профилем технологической специализации, т. е. доля работы данного профиля в общей программе работ должна быть не менее 75%. Неправильно, например, включать в анализ 4 строительно-монтажных управления, ведущих жилищное строительство, и 2 СМУ этого же треста, специализирующихся на дорожных и отделочных работах.

2. Внешними условиями, которые проявляются в однотипности производственного профиля, действии одних и тех же штатных расписаний, сметных и финансовых нормативов, примерно равных условий выполнения работ и обеспечения ресурсами. Нельзя, например, признать однородными подразделения одного министерства, расположенные в различных климатических условиях.

При соблюдении самых жестких условий однородности все же неизбежны различия в значениях важнейших технико-экономических показателей деятельности строительных подразделений. Эти различия имеют вероятностный характер, являются следствием некоторых условий (часто предварительно неизвестных), которые призван вскрыть и количественно оценить корреляционный анализ.

При проведении корреляционного анализа деятельности таких крупных систем, как главк, министерство, низовые подразделения должны группироваться не по принципу административной подчиненности, а по принципу общности структуры работ, отраслевой специализации и другим факторам, обусловливающим цельность исследования. Так, при изучении корреляционных зависимостей между накладными расходами одного из шахтостроительных комбинатов Донбасса (за 4 года) и некоторыми определяющими их размер факторами 87 строительно-монтажных управлений были разделены на следующие шесть групп:

специализированные строительные управления по проходке шахтных стволов;

управления, ведущие строительство шахт после проходки стволов (частично эти управления строили также жилые и культурно-бытовые объекты);

управления, ведущие работы по возведению обогатительных и брикетных фабрик, машиностроительных и ремонтных заводов;

управления, ведущие строительство жилищ, объектов социально-культурного и бытового назначения (на примере материалов этой группы далее иллюстрируется техника собственно корреляционного анализа);

дорожно-строительные управления;

управления, ведущие все виды монтажных работ.

Таким образом, в данном случае важнейшим признаком группировки явилась специализация управлений на выполнении различных видов работ.

Следующий этап исследования — сбор необходимой информации. В зависимости от цели исследования исполнители, занимающиеся сбором информации, должны быть снабжены специальными однотипными формами и вопросниками. Если данные выбираются из бухгалтерской или статистической отчетности, следует указывать шифры соответствующих строк и граф, учитывая при этом, что в отдельные годы одни и те же показатели могут иметь различные шифры и нумерацию строк.

В некоторых случаях данные переносятся непосредственно из форм отчетности. При этом значения факторов сохраняют в дальнейшем свою размерность. В других же случаях более целесообразно использовать в исследовании относительные величины, т. е. отнесенные к какой-либо базе: к единице объема выполненных работ, к численности рабочих, к 100 $м^2$ жилой площади и т. д.

В тех случаях, когда значения факторов невозможно получить из форм отчетности, исходные данные располагают в специальных таблицах и выполняют вычисления.

Если для дальнейших расчетов предполагается использование счетно-вычислительных машин, то исходные данные кодируются. Кодирование ускоряет поиск необходимых материалов и способствует систематизации в размещении исходных данных и при ручной обработке. После первичной обработки и кодирования разрабатывается порядок расположения исходных данных на перфокарте (называемый макетом), составляется программа группировки данных по одному или нескольким признакам и устанавливается последовательность выполнения всех арифметических действий. Процесс механизированной обработки исходных данных предусматривает:

нанесение на перфокарты первичных данных;

контроль правильности перфорации, исправление возможных ошибок;

сортировку перфокарт по какому-либо признаку;

табуляцию перфокарт с необходимыми признаками и подсчет итогов.

По табуляграмме (полученной в результате обработки данных на машиносчетной станции), на которой значения признака размещены в порядке возрастания или убывания, можно сделать заключение о степени влияния того или иного фактора и уже на этой предварительной стадии исключить из дальнейшего анализа значения факторов, резко искажающие общую закономерность.

В качестве примера рассмотрим обработку данных об уровне сборности и выработке по 36 строительно-монтажным управлениям комбината «Химстрой», специализированным на возведении объектов химической промышленности.

Первичный материал (см. табл. 18, раздел а), не подвергшийся никакой обработке, не дает возможности сделать выводы о качестве представленных данных.

Систематизация данных по возрастанию уровня сборности (см. табл. 18, раздел б) позволяет выделить три показателя выработки на 1 работающего в год, резко отличающиеся от остальных: 6761 руб. (СМУ № 7), 2650 руб. (СМУ № 9) и 2584 руб. (СМУ № 29). Первый из этих показателей привлек внимание нелогичным сочетанием весьма низкого уровня сборности (11,1%) с наибольшей выработкой на одного работника. Проверкой установлена описка при выборке из документации; правильное значение выработки в данном случае — 3761 руб. СМУ № 9 и № 29, как установлено, выполняют работы, структура которых резко отличается от остальных и поэтому данные о деятельности этих СМУ должны быть исключены из дальнейшего анализа.

Систематизация данных по определенным интервалам уровня сборности может дать и более глубокую характеристику связи этого показателя с выработкой на одного работающего (см. табл. 18, раздел в). Как видим, при систематизации данных по однородным группам, тенденция к возрастанию выработки в связи с ростом уровня сборности проявляется особенно отчетливо и это дает основание считать исходный материал (по 34 СМУ) приемлемым для дальнейшего анализа.

Центральной задачей предварительной обработки исходного статистического материала является *выбор факторов*, влияющих на результативный признак. Как из массы факторов отобрать наиболее существенные, оказывающие на выбранный признак наибольшее влияние? Исследователь может руководствоваться некоторыми рекомендациями. Так, при возможности использования ЭВМ рекомендуется включать *все* факторы, связанные с исследованием, которые поддаются количественной оценке, и на основании математико-статистического анализа отбирать из них те, которые окажутся существенно влияющими на признак.

Однако здесь нам представляется уместным привести предостережение специалистов в области корреляционных исследований (см. М. Езекиэл, К. А. Фокс [8], стр. 201): «Исследователь не должен поддаваться искушению, в которое его может увлечь во вред делу легкость исчислений, выполняемых при помощи новейших счетных машин».

Исследователь, являясь специалистом в данной области, должен отобрать факторы, отражающие специфику изучаемого явления.

К факторам, отбираемым для исследования, предъявляются следующие требования: наличие связи с результативным признаком, однозначность количественного измерения, приемлемая трудоемкость получения исходных данных.

Таблица 18

Табуляграмма показателей уровня сборности и выработки по строительно-монтажным управлениям комбината «Луганскхимстрой»

Шифр СМУ	Уровень сборности (%)	Выработка в год на 1 работающего (руб.)	Шифр СМУ	Уровень сборности (%)	Выработка в год на 1 работающего (руб.)
		а			
1	17,3	3 456	19	22,4	3 100
2	35,2	4 199	20	20,9	3 619
3	24,5	3 100	21	57,3	6 033
4	15,6	3 896	22	26,3	3 544
5	8,3	2 498	23	36,2	4 688
6	30,8	4 741	24	32,3	4 751
7	11,1	6 761	25	43,8	4 598
8	22,7	3 547	26	24,3	3 462
9	35,4	2 650	27	36,0	4 318
10	23,5	3 924	28	60,0	6 280
11	51,7	4 923	29	45,7	2 584
12	50,5	4 326	30	69,5	4 811
13	27,1	3 693	31	25,3	3 426
14	26,8	3 477	32	8,5	3 280
15	19,5	3 761	33	28,6	3 751
16	51,1	4 750	34	29,4	3 969
17	16,5	3 558	35	15,9	3 306
18	24,2	3 048	36	88,0	7 033
		б			
5	8,3	2 498	13	27,1	3 693
32	8,5	3 280	33	28,6	3 751
7	11,1	6 761	34	29,4	3 969
4	15,6	3 896	6	30,8	4 741
35	15,9	3 306	24	32,3	4 751
17	16,5	3 558	2	35,2	4 199
1	17,3	3 456	9	35,4	2 650
15	19,5	3 761	27	36,0	4 318
20	20,9	3 619	23	36,2	4 688
19	22,4	3 100	25	43,8	4 598
8	22,7	3 547	29	45,7	2 584
10	23,5	3 924	12	50,5	4 326
18	24,2	3 048	16	51,1	4 750
26	24,3	3 462	11	51,7	4 923
3	24,5	3 100	21	57,3	6 033
31	25,3	3 426	28	60,0	6 280
22	26,3	3 544	30	69,5	4 811
14	26,8	3 477	36	88,0	7 033

Продолжение

Шифр СМУ	Уровень сборности (%)	Выработка в год на 1 работающего (руб.)	Шифр СМУ	Уровень сборности (%)	Выработка в год на 1 работающего (руб.)
			в		
5	8,3	2 498	6	30,8	4 741
32	8,5	3 280	24	32,3	4 751
7	11,1	3 761	2	35,2	4 199
4	15,6	3 896	27	36,0	4 318
35	15,9	3 306	23	36,2	4 688
17	16,5	3 558	25	43,8	4 598
1	17,3	3 456			
15	19,5	3 761		$\Sigma = 377,8$	$\Sigma = 49\,155$
20	20,9	3 619			
19	22,4	3 100	Средняя	31,48	4 096
8	22,7	3 547			
10	23,5	3 924	12	50,5	4 326
18	24,2	3 043	16	51,1	4 750
26	24,3	3 462	11	51,7	4 923
3	24,5	3 100	21	57,3	6 033
	$\Sigma = 587,2$	$\Sigma = 51\,316$	28	60,0	6 280
Средняя	18,35	3 421	30	69,5	4 811
31	25,3	3 426		$\Sigma = 340,1$	$\Sigma = 31\,123$
22	26,3	3 544	Средняя	56,68	5 187
14	26,8	3 477			
13	27,1	3 693	36	88,0	7 033
33	28,6	3 751			
34	29,4	3 969			

1. *Наличие связи факторов с результативным признаком*. В дальнейшем предположение о наличии связи может подтвердиться или может быть отвергнуто, но для предварительного решения о включении фактора в анализ оно необходимо.

Поясним это высказывание примерами.

При применении методов корреляции к анализу накладных расходов и их нормированию могут быть рекомендованы в качестве возможных факторов: объем строительно-монтажных работ, выполняемых собственными силами, численность работников, фонд заработной платы. Наблюдения показали, что включение иных факторов не оказывает существенного влияния на результат, но резко усложняет корреляционный анализ.

Один из обобщающих показателей деятельности строительных организаций — уровень производительности труда. На величину этого показателя влияет множество факторов, в числе которых одним из существенных является индустриализация строительства. Под индустриализацией строительства понимается механизированное и автоматизированное изготовление строительных деталей, конструкций и других изделий на специализированных предприятиях, комплексно-механизированное производство строительно-монтажных работ при возведении зданий и сооружений на строительных площадках.

Уровень индустриализации характеризуется двумя показателями: уровнем механизации и уровнем применения сборных железобетонных конструкций, а иногда и только последним из них. В отдельных случаях, например, при строительстве магистральных трубопроводов, где сборный железобетон вообще не применяется, об уровне индустриализации строительства следует судить, очевидно, по механо- и энерговооруженности труда и иным показателям. Можно рекомендовать исследование влияния на производительность труда таких факторов, как уровень специализации, степень ритмичности (как показатель организованности, налаженности работ), показатель сосредоточенности строительства, уровень квалификации рабочих, сборности строительства и механизации отдельных видов строительно-монтажных работ и других.

Из факторов, действующих в сфере производственно-хозяйственной деятельности строительных организаций, почти невозможно указать на такие, которые не влияли бы на уровень себестоимости. Разумеется, возможный выбор факторов, включаемых в корреляционный анализ, в каждом конкретном случае может меняться. Необходимо только, чтобы подход к выбору факторов, определяющих уровень себестоимости, как впрочем и любого иного признака, не был формальным. Например, неправильно было бы искать корреляционную зависимость между ростом производительности труда и объемом работ по генеральному подряду, потому что производительность труда в строительстве определяется по показателю объема работ, выполняемых собственными силами. Такой подход, в частности, имеет место в книге Г. Б. Полисюк [15]. Не следует искать корреляционные связи там, где заведомо существуют функциональные зависимости; скажем, лишены оснований поиски связи между размером прямых затрат и стоимостью израсходованных на строительство материалов, которая является их составной частью.

2. *Однозначность количественного измерения*. Нельзя вводить в анализ такие факторы, как «технический прогресс», «индустриализация», «уровень организации работ», ибо они являются показателями, не обладающими количественной определенностью. Стремление отразить в корреляционных уравнениях (моделях) влияние подобных показателей вынуждает расчленять их на более простые, поддающиеся количественному измерению, или условно принимать, что влияние их отражается на величине других показателей, близких к ним по значению, смыслу.

Так, иногда фактор «индустриализация строительства» заменяется двумя другими — (механо- и энерговооруженностью труда) или «степень ритмичности» интерпретируется как «уровень организации строительного производства». Такой метод (назовем его методом подмены факторов), конечно, не позволяет учесть влияния всех сторон многогранного фактора, но придает ему количественную определенность с помощью более простого фактора, без которого невозможно какое бы то ни было применение методов корреляции.

К выбору факторов, включаемых в корреляционный анализ, необходимо подходить не только с точки зрения экономической теории, но и практики работы строительных организаций.

Рассмотрим пример. Перерасход заработной платы и сверхнормативный расход материалов в управлениях, специализированных на санитарно-технических работах, часто являются следствием недоброкачественного выполнения работ. В отчетности не отражаются сведения о качестве санитарно-технических работ на сданных объектах, качество работ получает отражение только в актах сдачи объектов в эксплуатацию в виде оценок — удовлетворительных, хороших, отличных, причем две первые имеют место примерно на 96% всех объектов. Отсюда можно сделать ложный вывод, что качество работ не оказывает влияния на себестоимость их выполнения. Очевидно, дело в том, что подавляющее большинство объектов заносится в акт не с первого предъявления к сдаче, а лишь после устранения недоделок.

Таким образом, действительную связь между качеством работ и их себестоимостью следует искать в сравнительной оценке затрат на устранение недоделок и некачественно выполненных работ. Но здесь мы сталкиваемся с другой проблемой — затратами труда на выявление величин того или иного фактора.

3. *Трудоемкость, простота и доступность получения исходных данных.* Даже наиболее весомый с экономических позиций фактор, оказывающий несомненное влияние на важнейшие показатели работы строительных организаций, нельзя включить в корреляционный анализ, если его определение связано со значительными затратами труда. В лучшем случае, его величина может быть определена отдельными учеными-энтузиастами, но широкого использования в практике экономической работы анализ такого фактора не получит. Поэтому при решении вопроса о выборе факторов рекомендуется учитывать трудоемкость установления их величины.

Наконец, важно отметить два направления по вопросу о степени соответствия исходного материала закону нормального распределения. Одни авторы считают возможным использовать в корреляционном анализе только те исходные данные, ряды распределения которых безоговорочно следуют этому закону. Наиболее четко это требование выражено в работе В. В. Померанцева [16]. Представители другого направления не ставят столь жестких условий к исходному материалу и вводят в исследование данные, распределение которых в отдельных случаях резко отличается от нормального. Ф. Миллс [12], например, использует в корреляционном анализе данные, распределение которых не подчиняется закону нормального распределения.

Практика применения метода корреляции к анализу экономики строительства приводит к вполне репрезентативным выводам, если исходный статистический материал (после его классификации и группировки) удовлетворяет одному из рассмотренных нами критериев согласия.

Итак, высказаны некоторые рекомендации, которые помогут при выборе и обосновании факторов, включаемых в корреляционный анализ. Но универсальные, безотказные рекомендации по этим вопросам не могут быть даны. Отдельные схемы взаимодействия и влияния многочисленных факторов на производительность труда, себестоимость строительно-монтажных работ и другие важнейшие экономические показатели приведены во многих пособиях по экономике строительства.

§ 2. Линейные зависимости

Основы методики и техники корреляционного анализа целесообразно показать на наиболее простых зависимостях — линейных (прямолинейных).

Зависимость величины накладных расходов от объема работ, численности работников и фонда заработной платы

Экономическая постановка задачи. В строительстве производственные расходы делятся на прямые и накладные. Прямые затраты состоят из затрат на материалы, конструкции и детали, основной заработной платы рабочих на строительно-монтажных работах, расходов по эксплуатации строительных машин и механизмов и прочих затрат. Прямые затраты могут быть отнесены к конкретному объекту или даже виду работ. В состав накладных расходов включаются административно-хозяйственные расходы, дополнительная заработная плата рабочих, расходы, связанные с износом временных зданий и сооружений, а также ряд других статей затрат. Накладные расходы включаются в себестоимость объекта строительства методом распределения.

Накладные расходы нормируются в пределах $12 \div 17\%$ прямых затрат, или $10 \div 14\%$ себестоимости работ. Размеры накладных расходов могут меняться независимо от уровня прямых затрат. При анализе причин удорожания строительства устанавливаем, что вызывается оно большей частью превышением норм накладных расходов. Этим определяется внимание, уделяемое исследованиям зависимости накладных расходов от воздействующих на них факторов.

Основными факторами, влияющими на уровень накладных расходов, являются: сметная стоимость годового объема строительно-монтажных работ, выполненных собственными силами; численность рабочих, зависящая от уровня выработки; величина заработной платы рабочих. В величине годового объема работ косвенно отражается продолжительность возведения объектов.

Сокращение продолжительности строительства способствует увеличению объемов строительно-монтажных работ, выполняемых в единицу времени, что приводит к снижению условно постоянной части накладных расходов и, следовательно, к общему снижению накладных расходов.

От выработки рабочих также зависит условно постоянная часть накладных расходов, связанная с содержанием персонала (нормировщиков, табельщиков, счетоводов, учетчиков и др.). Жилищно-коммунальные расходы непосредственно связаны с численностью рабочих, и потому полностью зависят от выработки.

С уменьшением удельного веса основной заработной платы рабочих в общих затратах на строительно-монтажные работы достигается снижение накладных расходов. Разработку нормативов накладных расходов с помощью корреляционного анализа иллюстрируем на исходных данных 74 строительно-монтажных управлений жилищного строительства (табл. 19).

Для корреляционного анализа зависимости величины накладных расходов от объема работ необходима статистическая обработка данных. Обозначим: y — размер накладных расходов; x — объем работ.

Находим границы изменения каждой переменной.

y — меняется в пределах от 50 (СМУ № 16) до 570 тыс. руб. (СМУ № 74);

x — меняется в пределах от 210 (СМУ № 57) до 3690 тыс. руб. (СМУ № 74).

Определяем количество интервалов l, внутри которых изменениями y и x пренебрегаем; в данном случае для 74 СМУ принимаем для x $l = 8$, для y $l = 6$.

Находим i_x — шаг интервалов для x, равный разности между наибольшим и наименьшим значением переменной, деленной на количество интервалов без единицы:

$$i_x = \frac{x_{\max} - x_{\min}}{l-1} = \frac{3690-210}{8-1} \cong 500.$$

Находим верхние и нижние предельные значения интервала. Верхняя граница последнего интервал:

$$3690 + \frac{500}{2} = 4000 \text{ (с округлением).}$$

Нижняя граница первого интервала:

$$210 - \frac{500}{2} < 0 \text{ (но принимаем равной нулю).}$$

Получаем следующие интервалы x:

4000—3500, 3500—3000, 3000—2500, 2500—2000, 2000—1500, 1500—1000, 1000—500, 500—0, всего 8 интервалов.

Таблица 19

Исходные данные для корреляционных расчетов

№ СМУ	Накладные расходы (тыс. руб.)	Объем строительно-монтажных работ (тыс. руб.)	Численность работников	Фонд заработной платы (тыс. руб.)	№ СМУ	Накладные расходы (тыс. руб.)	Объем строительно-монтажных работ (тыс. руб.)	Численность работников	Фонд заработной платы (тыс. руб.)
1	100	550	308	265	38	320	2 260	990	767
2	170	860	424	231	39	390	2 460	937	799
3	140	910	359	230	40	100	270	210	207
4	230	930	438	396	41	140	580	290	207
5	200	930	468	473	42	140	920	416	298
6	160	1 050	509	399	43	200	1 160	506	431
7	190	1 070	534	377	44	240	1 300	517	510
8	260	1 460	638	504	45	250	1 350	657	528
9	270	1 530	716	476	46	200	1 440	578	505
10	260	1 580	726	494	47	280	1 450	694	591
11	330	1 720	719	445	48	270	1 470	613	546
12	360	1 760	723	515	49	260	1 480	608	602
13	340	2 270	1 000	704	50	260	1 500	583	512
14	380	2 440	957	752	51	320	1 550	764	706
15	470	2 830	1 086	917	52	290	1 650	750	659
16	50	260	44	31	53	340	1 720	911	815
17	110	310	182	161	54	300	1 810	791	564
18	100	560	169	178	55	360	1 950	886	678
19	210	790	239	237	56	260	2 140	702	714
20	90	790	287	244	57	150	210	39	48
21	70	790	309	291	58	160	390	258	426
22	140	860	294	344	59	170	1 080	412	503
23	260	900	484	405	60	250	1 190	534	487
24	140	930	401	341	61	300	1 280	664	543
25	160	950	442	293	62	210	1 360	523	740
26	280	1 040	597	501	63	310	1 390	730	673
27	220	1 250	601	505	64	360	1 750	725	678
28	150	1 280	446	316	65	330	1 800	817	874
29	210	1 270	561	403	66	290	1 850	678	713
30	290	1 340	662	617	67	330	1 850	836	743
31	170	1 370	598	498	68	280	1 910	749	700
32	280	1 490	699	672	69	340	2 000	792	763
33	280	1 540	624	525	70	340	2 000	782	743
34	320	1 560	732	571	71	390	2 090	923	1 012
35	240	1 750	677	524	72	290	2 230	666	699
36	260	1 820	641	558	73	320	2 290	772	972
37	340	2 120	912	693	74	570	3 690	1 335	1 378

Шаг интервалов для y:

$$i_y = \frac{570-50}{6-1} \cong 100.$$

Находим верхний интервал $570 + \frac{100}{2}$ (округлим до 600); нижний интервал $50 - \frac{100}{2} = 0$.

Интервалы равны: 600—500, 500—400, 400—300, 300—200, 200—100, 100—0.

После этого строим корреляционную таблицу (см. табл. 20, рис. 9). От начала координат (точка 0) по оси ординат откладываем интервальные значения y. Границы интервалов отмечены мелкими цифрами 0, 100, 200, 300, 400, 500, 600 тыс. руб. по стрелкам РИ (реальные интервалы).

По оси абсцисс откладываем интервальные значения x. Границами интервалов будут значения 0, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000 тыс. руб. Они также размещены по стрелкам РИ (реальные интервалы).

Значения y и x, приведенные в табл. 19, откладываем на рисунке в пределах назначенных интервалов.

Так все строительные управления с накладными расходами (y) от 200 до 300 тыс. руб. и объемом работ (x) от 500 до 1000 тыс. руб., будут обозначены точками, которые попадут в прямоугольник, находящийся на пересечении значений $y = 250$ и $x = 750$. Таких строительных управлений оказалось три.

Точки с нижними граничными значениями переносим в ближайщий прямоугольник координатной решетки, расположенный слева или внизу. Например, если $y = 200$, а x равно любому значению между 500 и 1000, точка (т. е. строительное управление) будет проставлена в вышеуказанном прямоугольнике; если y — равно любому значению от 200,1 до 300, а $x = 500$ точка будет проставлена в прямоугольнике слева от указанного (таких в данном примере не оказалось).

Строительные управления с одной или двумя верхними границами значений y и x оставляются в данном прямоугольнике. Значения y — 50, 150, 250, 350, 450, 550 и значения x — 250, 750, 1250, 1750, 2250, 2750, 3250, 3750 соответствуют центрам каждого интервала; они размещены по стрелкам ЦРИ (центры реальных интервалов).

Таким образом, число точек в каждом прямоугольнике соответствует количеству строительных управлений, кото-

Таблица 20

Корреляционная таблица											строительная организация шифр год				
	x'		-2	-1	0	1	2	3	4	5	ЦУИ	←	Итого		
y'	x / y		250	750	1250	1750	2250	2750	3250	3750	ЦРИ	←	и	к	л
		0	500	1000	1500	2000	2500	3000	3500	4000	← PU		n_y	$n_y y'$	$n_y (y')^2$
↓ ЦУИ	↓ ЦРИ	↓ PU													
3	550	600								^{3} 1			1	3	9
2	450	500						^{2} 1					1	2	4
1	350	400			^{1} 1	^{11} 11	^{7} 7						19	19	19
0	250	300		^{0} 3	^{0} 15	^{0} 9	^{0} 2						29	0	0
-1	150	200	$^{-2}$ 2	$^{-8}$ 8	$^{-7}$ 7								17	-17	17
-2	50	100	$^{-6}$ 3	$^{-8}$ 4									7	-14	28
а	n_x		5	15	23	20	9	1		1			n 74	-7	77
б	$n_x \cdot x'$		-10	-15	0	20	18	3		5			21		
в	$n_x (x')^2$		20	15	0	20	36	9		25			125		
г	$\Sigma n_i y'$		-8	-16	-6	11	7	2		3			-7		
д	$x' \Sigma n_i y'$		16	16	0	11	14	6		15			78		
е	$\bar{y}'$		-1,60	-1,06	-0,26	0,55	0,78	2,00		3,00			$i_y = 100$		
ж	$\bar{y}' i_y$		-160	-106	-26	55	78	200		300			$i_x = 500$		
з	$\bar{y}$		90	144	224	305	328	450		550			$y_0 = 250$ $x_0 = 1250$		

$\bar{y} = 53,4 + 0,134x$

$$r' = \frac{74 \cdot 78 - (-7)\,21}{\sqrt{74 \cdot 125 - 21^2}\,\sqrt{74 \cdot 77 - (-7)^2}} = \frac{5919}{93,9 \cdot 75,1} = 0,839$$

Рис. 9. Таблица расчета зависимости между накладными расходами и объемом строительно-монтажных работ (по данным 74 жилищно-строительных управлений).

рые попадают в интервалы с выбранными границами для признака y и фактора x. Количество точек в каждом прямоугольнике подсчитывается и записывается цифрой (см. на рис. — 3, 4, 2, 8, 7, 3, 15, 9, 2, 1, 11, 7, 1, 1).

Естественно, что общее число точек n должно равняться 74. Количество точек в вертикальных и горизонтальных рядах суммируется по осям и записывается в строке n_x и графе n_y.

В результате получаем два ряда цифр: горизонтальный ряд n_x — 5, 15, 23, 20, 9, 1, 1 и вертикальный ряд n_y — 7, 17, 29, 19, 1, 1. Эти ряды цифр, как мы уже знаем, являются рядами распределения.

Если отложить на корреляционной решетке ординаты, соответствующие в масштабе величинам цифр в рядах распределения, и концы их соединить плавными кривыми, получим колоколообразные кривые, изображенные на рис. 9 тонкими линиями.

Чем более однородны подразделения, включаемые в исследование, тем ближе реальные кривые к кривой нормального распределения.

Для оценки качества исходных данных достаточно внимательно проследить чтобы ряды распределения (ряды n_x, n_y) имели плавные нарастания и спад значений.

Далее действительные (реальные) значения y и x заменяем условными y' и x' с целью значительного упрощения и сокращения вычислительной работы. Для этого выбираем новые координатные оси. Вообще эти оси можно выбрать в любом месте, но удобнее всего — в центре наибольших значений рядов n_x и n_y. В нашем примере осью y' принята вертикаль, проходящая через значение $n_x = 23$, а осью x' — горизонталь, проходящая через значение $n_y = 29$.

На условных осях откладываем положительные (вправо и вверх) и отрицательные (влево и вниз) y' и x'. Значения y' (—2, —1, 0, 1, 2, 3) и x' (—2, —1, 0, 1, 2, 3, 4, 5) соответствуют серединам интервалов; они размещены по стрелкам ЦУИ (центры условных интервалов). В каждом прямоугольнике проставляются (справа вверху мелким шрифтом) произведения числа точек в данном прямоугольнике на значение условной ординаты.

Теперь корреляционная таблица готова и можно определить две важнейшие характеристики изучаемых зависимостей — тесноту и форму связи.

Теснота связи позволяет выяснить, как изменяется признак y в зависимости от изменения фактора x при условии, что влияние прочих факторов постоянно.

В действительности прочие факторы изменяются и своей изменчивостью искажают зависимость. При этом влияние их обнаруживается с большей или меньшей силой. Теснота связи определяет силу (надежность), с которой данная зависимость проявляется среди многообразно нарушающих ее воздействий.

В случае линейной связи эту величину характеризует коэффициент корреляции - r; при криволинейных связях (параболе, гиперболе и др.) корреляционное отношение — η.

Парный коэффициент корреляции r определяется по формуле:

$$r = \frac{n \cdot x'(\sum n_x \cdot x' - n_y \cdot y' \cdot n_x \cdot x'}{\sqrt{n \cdot n_x (x')^2 - (n_x x')^2} \sqrt{n \cdot n_y (y')^2 - (n_y\ y')^2}}, \quad (II.1)$$

где n — количество исследуемых строительных управлений;

n_x и n_y — распределение общего количества строительных организаций (или наблюдений, анализов) по осям y и x.

Значения остальных величин очевидны при рассмотрении корреляционной таблицы (см. рис. 9).

В строке $n_x\ x'$ записываются произведения n_x на соответствующие им условные ординаты:
$5\ (-2) = -10$; $15\ (-1) = -15$; $23 \cdot 0 = 0$; $20 \cdot 1 = 20$; $9 \cdot 2 = 18$; $1 \cdot 3 = 3$; $1 \cdot 5 = 5$.

В графе $n_y \cdot y'$ записываются произведения n_y на соответствующие условные ординаты:
$7 \cdot (-2) = -14$; $17 \cdot (-1) = -17$; $29 \cdot 0 = 0$; $19 \cdot 1 = 19$; $1 \cdot 2 = 2$; $1 \cdot 3 = 3$.

В строке $n_x \cdot (x')^2$ и графе $n_y \cdot (y')^2$ записываются произведения строк n_x и n_y на квадрат соответствующих условных ординат:
$n_x(x')^2 = 5 \cdot 4 = 20$; $15 \cdot 1 = 15$; $23 \cdot 0 = 0$; $20 \cdot 1 = 20$; $9 \cdot 4 = 36$; и т. д.
$n_y \cdot (y')^2 = 7 \cdot 4 = 28$; $17 \cdot 1 = 17$; $29 \cdot 0 = 0$; $19 \cdot 1 = 19$; $1 \cdot 4 = 4$; $1 \cdot 9 = 9$.

Затем в строке $\sum n_i \cdot y'$ записываем алгебраические суммы цифр, ранее проставленных в каждом прямоугольнике справа вверху: $-2 + (-6) = -8$; $0 + (-8) + (-8) =$

$= -16$; $1+0+(-7) = -6$; $11+0 = 11$; $7+0 = 7$; 2; 3. Суммы строк $\sum n_i \cdot y'$ и $\sum n_y \cdot y'$ должны быть равны, этим можно контролировать правильность подсчетов; в данном примере $\sum n_i \cdot y' = \sum n_y \cdot y' = -7$ (см. рис. 9). И наконец, в строке $x' \sum n_i \cdot y$ записываем произведения значений строки $\sum n_i \cdot y'$ на значения соответствующих условных ординат: $(-8) \cdot (-2) = 16$; $-16 \cdot (-1) = 16$; $-6 \cdot 0 = 0$; $11 \cdot 1 = 11$; $7 \cdot 2 = 14$; $2 \cdot 3 = 6$; $3 \cdot 5 = 15$.

Подставим найденные в таблице значения в формулу (II.1):

$$r = \frac{74 \cdot 78 - (-7) \cdot 21}{\sqrt{74 \cdot 125 - 21^2} \sqrt{74 \cdot 77 - (-7)^2}} = 0{,}839.$$

Коэффициент корреляции может принимать значения от -1 до $+1$. При $r = -1$ или $+1$ корреляционная зависимость превращается в функциональную, при которой каждому значению признака y соответствует вполне определенное значение фактора x.

Положительные значения r свидетельствуют о прямой связи, т. е. изменения признака и фактора направлены одноименно. Отрицательные значения r предполагают обратную связь, т. е. разнонаправленность изменений признака и фактора.

Значение $r = 0{,}839$ указывает, что в данном случае имеется надежная прямая корреляционная связь между x и y. Значения $r = 0{,}20 \div 0{,}25$ указывают на отсутствие надежной корреляционной зависимости.

Формула вычисления коэффициента корреляции для практических работников может быть записана в упрощенном виде:

$$r = \frac{n \cdot д - б \cdot к}{\sqrt{n \cdot в - б^2} \ \sqrt{nл - к^2}},$$

где $а$, б, $в$, $г$, $д$ — обозначения соответствующих строк на рис. 9; $к$, $л$ — граф; n — число включаемых в исследование строек, наблюдений.

Средняя квадратическая ошибка коэффициента парной корреляции определяется по формуле:

$$\mu_{or} = \frac{1 - r^2}{\sqrt{n}}. \qquad \text{(II.2)}$$

Показатель r считается достоверным, если критерий надежности

$$m_r = \frac{|r|}{\mu_{or}} \geqslant 2{,}6.$$

По данным корреляционной таблицы, представленной на рис. 9,

$$\mu_{or} = \frac{1-0{,}839^2}{\sqrt{74}} = 0{,}086 \text{ и } m_r = \frac{0{,}839}{0{,}087} = 9{,}65 > 2{,}6.$$

Форма связи между y и x определяется эмпирической, построенной на основании исходных данных, а затем теоретической линией регрессии.

Эмпирическая линия регрессии (обычно ломаная линия) показывает средние по группам значения y (или x) и дает таким образом первое приближенное понятие о характере средних изменений признака в зависимости от средних изменений фактора.

Теоретическая линия регрессии представляет собой такую математически правильную кривую (либо прямую) линию, которая проходит наиболее близко к точкам эмпирической линии регрессии и выражет общую закономерность средних изменений признака в связи со средними изменениями фактора.

Для расчета эмпирической линии в таблице служат последние три строки ($е$, $ж$, $з$). Строка $е = \bar{y}' = \frac{г}{д} = \frac{\sum n_i \cdot y'}{n_x}$ — средние значения y', соответствующие всем значениям x' в условных координатах. В данном примере $\bar{y}'$ имеет следующие значения:

$$\frac{-8}{5} = -1{,}60;\ \frac{-16}{15} = -1{,}06;\ \frac{-6}{23} = -0{,}26;\ \frac{11}{20} = 0{,}55;$$
$$\frac{7}{9} = 0{,}78;\ \frac{2}{1} = 2;\ \frac{3}{1} = 3.$$

Строка $ж = е\, i_y = \bar{y}' \cdot i_y$ — те же значения в действительных координатах; i_y — шаг интервала по вертикальной оси, в данном примере $i_y = 100$.

Последняя строка $з = \bar{y} = y_0 + \bar{y}'\ i_y$ — ординаты эмпирической линии регрессии в действительных координатах (x, y); y_0 — среднее значение y в действительных координатах, соответствующее $y' = 0$; в данном случае $y_0 = 250$.

Итак, значения последней строки ($\bar{y}$) находим следующим образом:

$250 + (-1{,}60 \cdot 100) = 90$; $250 + (-1{,}06 \cdot 100) = 144$; $250 + (-0{,}26 \cdot 100) = 224$; $250 + 0{,}55 \cdot 100 = 305$; $250 + 0{,}78 \cdot 100 = 328$; $250 + 2 \cdot 100 = 450$; $250 + 3 \cdot 100 = 550$;

Из центров интервалов перпендикулярно оси x восстанавливаем отрезки, равные (в принятом масштабе) 90; 144; 224 и т. д., соединяем концы этих отрезков пунктирной линией регрессии.

Расчет теоретической линии регрессии представляет интерес уже потому, что дает возможность весьма просто выразить в виде математической формулы зависимость средних изменений признака от средних изменений фактора. Характер размещения точек на корреляционном поле делает весьма вероятной гипотезу о линейной связи y от x. Параметры искомой прямой (a', b') находим из системы уравнений по способу наименьших квадратов:

$$\begin{cases} n \cdot a' + (n_x \cdot x') \cdot b' = \sum n_x \cdot x'; \\ (n_x \cdot x')\, a' + (n_x \cdot x'^2) \cdot b' = x' \sum n_x \cdot x'. \end{cases} \qquad \text{(II.4)}$$

Подставим в систему уравнений значения из табл. 20:

$$\begin{cases} 74\, a' + 21\, b' = -7; \\ 21\, a' + 125\, b' = 78. \end{cases}$$

Для определения коэффициентов a' и $в'$ разделим уравнения на коэффициенты при a':

$$a' + 0{,}284 \cdot b' = -0{,}095;$$
$$a' + 5{,}952 \cdot b' = 3{,}714.$$

Вычтем из второго уравнения первое:

$$5{,}668\, b' = 3{,}809,$$

$$\text{откуда } b' = \frac{3{,}809}{5{,}668} = 0{,}672.$$

Подставляя значение b' в одно из уравнений, находим:

$$a' = 3{,}714 - 5{,}952 \cdot 0{,}672 = -0{,}286.$$

Уравнение теоретической линии регрессии принимает вид:

$$y' = -0{,}286 + 0{,}672\, x'.$$

Для расчета параметров зависимости y от x может быть рекомендован и более простой метод.

Если известны средние арифметические $\overline{y}$ и $\overline{x}$ переменных, средние квадратические отклонения $\sigma_{\overline{x}}$ и $\sigma_{\overline{y}}$ и парный коэффициент корреляции r_{yx}, то параметры a и b прямой $y_x = a + bx$ можно определить из уравнений:

$$b = r_{yx}\frac{\sigma_y}{\sigma_x}; \tag{II.5}$$

$$a = \overline{y} - b\overline{x}. \tag{II.6}$$

Так, пользуясь корреляционной таблицей (рис. 9), определим $\overline{x} = \frac{21}{74} = 0{,}284$; $\overline{y} = \frac{-7}{74} = -0{,}095$.

Находим через дисперсию $\sigma_y = \sqrt{\frac{77}{74} - 0{,}095^2} = 1{,}016$;

$\sigma_x = \sqrt{\frac{125}{74} - 0{,}284^2} = 1{,}269$; напомним, что значение σ_x уже было получено при исследовании рядов распределения.

$$r_{yx} = 0{,}839. \text{ Тогда } b' = 0{,}839 \cdot \frac{1{,}016}{1{,}269} = 0{,}672;$$

$$a' = -0{,}095 - 0{,}672 \cdot 0{,}284 = -0{,}286 \text{ и}$$

$$\overline{y}' = -0{,}286 + 0{,}672\,x'.$$

Для перехода к реальным координатам определяем a и b из уравнений:

$$a = y_o + i_y a' - b' \cdot x_o \frac{i_y}{i_x}; \tag{II.7}$$

$$b = b'\frac{i_y}{i_x}. \tag{II.8}$$

Принимая значения i_y, i_x, y_0, x_0 из корреляционной таблицы (см. рис. 9), получаем:

$$a = 250 + 100\,(-0{,}286) - 0{,}672 \cdot 1250 \cdot \frac{100}{500} = 53{,}4;$$

$$b = 0{,}672 \cdot \frac{100}{500} = 0{,}134.$$

Окончательно: $\overline{y} = 53{,}4 + 0{,}134\ x$.

Это уравнение показывает, что с увеличением объема работ на 100 тыс. руб. накладные расходы жилищно-строительных организаций увеличиваются на 13,4 тыс. руб.; в то же время существует постоянная часть накладных расходов, не изменяющаяся с изменением объема работ, которая составляет 53,4 тыс. руб. в каждых 100 тыс. руб. накладных расходов.

Аналогичные уравнения могут найти применение в различных расчетах при решении экономических задач строительства и проектирования.

Более точный результат можно получить, если воспользоваться методом наименьших квадратов при обработке исходных данных табл. 21.

Таблица 21

Исходные данные для корреляционных расчетов

№ СМУ	Накладные расходы (тыс. руб.) y	Объем работ (тыс. руб.) x	x^2	$x \cdot y$
1	2	3	4	5
1	100	550	302 500	55 000
2	170	860	739 600	146 200
3	140	910	828 100	127 400
...	...	......	....	....
72	290	2 230	4 972 900	646 700
73	320	2 290	5 244 100	732 800
74	570	3 690	13 616 100	2 103 300
Итого	$\Sigma y = 18\,440$	$\Sigma x = 105\,680$	$\Sigma x^2 = 180\,104\,400$	$\Sigma xy = 30\,269\,300$

Значения параметров прямой определяем из уравнений:

$$74a + 105680b = 18440;$$

$$105680a + 180104400b = 30269300,$$

откуда $a = 56{,}5$ и $b = 0{,}134$.

Применение метода наименьших квадратов без предварительной группировки и усреднения исходных данных (прямым счетом), как отмечалось, дает наиболее точные результаты, но при отсутствии вычислительных машин более рационально использовать корреляционные таблицы.

Преимущества применения корреляционных таблиц следующие.

Отпадает необходимость вычислений, трудоемкость которых резко возрастает с увеличением числа наблюдений.

Графическое изображение явления на корреляционном поле сразу же позволяет судить о характере эмпирической кривой и в то же время является частью определения искомого уравнения регрессии, а при использовании метода наименьших квадратов без усреднения исходных данных эту работу необходимо выполнять дополнительно, чтобы иметь суждение о типе предполагаемой зависимости. В данном примере уравнение прямой, исчисленное по корреляционной таблице, имеет точность $\pm 1,76\%$.

Аналогично тому, как с помощью корреляционной таблицы (рис. 9) мы нашли уравнение, в котором среднее значение накладных расходов определяется в зависимости от изменения годовых объемов работ, можно найти уравнение связи между накладными расходами — y и среднегодовой численностью работников — x (число строительных управлений — 74).

Предоставляем читателю, выполнив все расчеты, проверить правильность уравнения:

$$\overline{y} = 4,7 + 0,0039x \text{ (при } r = 0,883).$$

Следует отметить, что обычно специальные таблицы используются только для определения коэффициента корреляции. Лишь в монографии Я. И. Лукомского [11] расчетная таблица включает также ряд граф для вычисления ординат эмпирической линии регрессии.

Предложенная форма корреляционной таблицы пригодна для выполнения всего комплекса операций: определения тесноты связи, формы связи и графического построения эмпирической и теоретической линий регрессии. Как показал опыт работы, такая форма таблицы наиболее рациональна, значительно сокращает затраты времени на вычислительные работы. Таблица полностью обрабатывается техником или лаборантом в течение 25—30 минут. Кроме того, для распределения исходных данных на корреляционной решетке на основе методики проф. Я. И. Лукомского авторами разработаны и применены специальные карточки (см. рис. 10).

Карточка содержит все данные, необходимые для исследования. Так, например, в карточке приведены данные, необходимые для изучения и нормирования накладных расходов.

При изучении производительности труда в зависимости от индустриализации, механизации и ряда других факторов карточка будет содержать, конечно, другие данные. Раскладывая в нашем примере 74 карточки по интервалам, выбранным для объемов работ (0—500—1000—1500 и т. д. тыс. руб.), автоматически получаем ряд распределения по этому фактору: 5—15—23—20—9—1—1 (всего 74).

Суммирование количества карточек по горизонтали (строка n_x) и по вертикали (графа n_y) дает ряд распределения по соответствующему фактору; таким образом, отпадает необходимость в разноске каждого управления в отдельно-

Накладные расходы (тыс. руб.)	Объем работ (тыс. руб.)	Среднесписочная численность работников (чел.)	Годовой фонд заработной платы работающих (тыс. руб.)	Средняя годовая выработка (на 1 работающего)	Средняя годовая заработная плата (на 1 работающего)
100	550	308	265	1,78	0,86

Шифры		
год	N	1
комбинат	N	1
трест	N	06
тип строительства	N	4
группа объема	N	01
строительное управление	N	106

Прямые затраты (тыс. руб.)	в том числе:				Всего затрат (тыс. руб.)
	Материалы	Заработная плата	Эксплуатация машин	Прочие затраты	
600	340	130	120	10	700

Рис. 10. Карточка для упрощения разноски исходных данных на корреляционные решетки.

сти по интервалам, чем значительно упрощается обработка данных.

Теснота и форма связи, установленные с помощью корреляционных таблиц, характеризуют парные зависимости, например между накладными расходами и сметной стоимостью выполненных работ. В дальнейшем парные зависимости явятся основой для расчета уравнений множественной регрессии, которые могут быть использованы при разработке нормативов накладных расходов.

Зависимость уровня производительности труда от квалификации рабочих

Экономическая постановка задачи. Систематическое повышение квалификации рабочих — непременное условие успешного внедрения новой техники в строительное производство. Как известно, признаком уровня квалификации рабочего является присвоенный ему тарифный разряд. Уровень квалификации звена или бригады рабочих характеризуется укрупненным показателем — средним тарифным разрядом. Средний тарифный разряд определяется как средняя из разрядов тарифной сетки, взвешенная по числу рабочих соответствующих разрядов с учетом среднего тарифа часовой ставки. Такой показатель дает возможность с приемлемой точностью соизмерить высококвалифицированный и малоквалифицированный труд и выявлять тем самым тенденции в изменении общего уровня квалификации рабочих как одного из основных факторов, определяющих производительность труда.

В качестве показателя производительности труда принята средняя выработка одного рабочего на строительно-монтажных работах.

Ряд примеров подтверждает общую тенденцию — зависимость выработки от среднего тарифного разряда.

Так, в тресте «Химстрой-1» одна из бригад каменщиков до повышения квалификации выполняла нормы выработки на 111%, а после учебы — на 230%; бригада бетонщиков соответственно на 141 и 167%; бригада кровельщиков в тресте «Химстрой-2» — на 148 и 157% и т. д.

В тресте № 8 (Харьков) в 1966 г. при среднем разряде рабочих, равном 2,5, выработка составила 4150 руб. в год, а в следующем году при среднем разряде 2,9 выработка достигла 4200 руб. В тресте № 9 при среднем разряде 2,8 выработка была 4190 руб. в год, а при повышении среднего разряда до 3 выработка повысилась до 4210 руб. при сохранении прежней структуры работ.

Зависимость увеличения выработки от повышения среднего разряда иллюстрируется материалами строительно-монтажных управлений одного из трестов Минспецстроя УССР, специализированных на выполнении санитарно-технических работ, за 1963—1966 гг.

Графическое изображение результатов наблюдения 36 строительно-монтажных управлений, т. е. корреляционное поле, делает вероятной гипотезу о линейной зависимости выработки (y) от уровня квалификации, выражаемой средним разрядом рабочих (x). Результаты приведены в табл. 22.

Расчет тесноты и формы связи выполним без применения корреляционной таблицы по следующим формулам:

Таблица 22

Исходные данные для корреляционных расчетов

Год	Шифр строительных управлений	Выработка на 1 рабочего в год, (тыс. руб.) y	Средний разряд рабочих, x	xy	x^2	y^2
1963	СУ-520	6 706	3,45	$231,4\cdot10^2$	11,90	$449,7\cdot10^5$
	СУ-521	6 387	3,93	$251,0\cdot10^2$	15,44	$307,9\cdot10^5$
	СУ-522	6 146	3,22	$197,9\cdot10^2$	10,37	$377,7\cdot10^5$
	СУ-523	6 586	3,67	$241,7\cdot10^2$	13,47	$433,8\cdot10^5$
	СУ-525	6 221	3,56	$221,5\cdot10^2$	12,67	$387,0\cdot10^5$
	СУ-526	5 550	3,58	$198,7\cdot10^2$	12,82	$308,0\cdot10^5$
	СУ-533	7 482	4,33	$324,0\cdot10^2$	18,75	$559,8\cdot10^5$
	СУ-535	5 234	2,82	$147,6\cdot10^2$	7,95	$279,9\cdot10^5$
	СУ-542	5 716	3,76	$214,9\cdot10^2$	14,14	$326,7\cdot10^5$
1964	СУ-520	6 690	4,70	$247,5\cdot10^2$	13,69	$447,6\cdot10^5$
	СУ-521	6 729	3,79	$255,0\cdot10^2$	14,36	$452,8\cdot10^5$
	СУ-522	6 337	2,99	$189,5\cdot10^2$	8,94	$401,6\cdot10^5$
	СУ-523	7 380	3,73	$275,3\cdot10^2$	13,91	$544,6\cdot10^5$
	СУ-525	6 293	3,34	$210,2\cdot10^2$	11,16	$396,0\cdot10^5$
	СУ-526	7 096	3,58	$254,0\cdot10^2$	12,82	$503,5\cdot10^5$
	СУ-533	7 893	4,37	$344,9\cdot10^2$	19,10	$623,0\cdot10^5$
	ПМК-33	8 798	3,64	$138,2\cdot10^2$	13,25	$144,2\cdot10^5$
1965	СУ-520	7 609	3,71	$282,3\cdot10^2$	13,76	$579,0\cdot10^5$
	СУ-521	7 025	3,66	$257,1\cdot10^2$	13,40	$493,5\cdot10^5$
	СУ-522	7 163	3,80	$272,2\cdot10^2$	14,44	$513,1\cdot10^5$
	СУ-523	7 951	4,23	$336,3\cdot10^2$	17,89	$632,2\cdot10^5$
	СУ-525	7 216	3,57	$257,6\cdot10^2$	12,74	$520,7\cdot10^5$
	СУ-526	6 553	3,61	$236,6\cdot10^2$	13,03	$429,4\cdot10^5$
	СУ-533	8 268	4,27	$353,0\cdot10^2$	18,23	$683,6\cdot10^5$
	ПМК-33	5 216	3,55	$185,2\cdot10^2$	12,60	$272,1\cdot10^5$
	ПМК-576	4 066	3,10	$126,0\cdot10^2$	9,61	$163,3\cdot10^5$
1966	СУ-520	7 570	3,71	$280,8\cdot10^2$	13,76	$573,0\cdot10^5$
	СУ-521	7 334	3,90	$286,0\cdot10^2$	15,21	$537,9\cdot10^5$
	СУ-522	7 452	3,60	$268,3\cdot10^2$	12,96	$555,3\cdot10^5$
	СУ-523	8 298	4,33	$359,3\cdot10^2$	18,75	$688,6\cdot10^5$
	СУ-525	7 979	2,66	$212,2\cdot10^2$	7,08	$636,6\cdot10^5$
	СУ-526	7 919	2,50	$198,0\cdot10^2$	6,25	$627,1\cdot10^5$
	СУ-533	8 511	3,46	$294,5\cdot10^2$	11,97	$724,4\cdot10^5$
	СУ-548	6 857	3,42	$234,5\cdot10^2$	11,70	$470,2\cdot10^5$
	ПМК-33	6 708	3,60	$241,5\cdot10^2$	12,96	$450,0\cdot10^5$
	ПМК-576	5 665	3,41	$193,2\cdot10^2$	11,63	$320,9\cdot10^5$
	$n=36$	$\Sigma y=$ $=243\,604$	$\Sigma x=$ $=129,55$	$\Sigma xy=$ $=8817,9\cdot10^2$	$\Sigma x^2=$ $=472,71$	$\Sigma y^2=$ $=16910,7\cdot10^5$

$$r = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\sigma_x \cdot \sigma_y}, \tag{II.9}$$

где $\overline{xy}$ — среднее значение произведения $x \cdot y$;

$\overline{y}$ и $\overline{x}$ — средние значения признака и фактора;

σ_y и σ_x — средние квадратические отклонения по x и y.

$$\overline{xy} = \frac{\Sigma xy}{n} = \frac{8817{,}9 \cdot 10^2}{36} = 24500; \quad \overline{x} = \frac{\Sigma x}{n} = \frac{129{,}55}{36} = 3{,}6.$$

$$\overline{y} = \frac{\Sigma y}{n} = \frac{243604}{36} = 6765; \quad \sigma_x^2 = \frac{\Sigma x^2}{n} -$$

$$- \overline{x}^2 = \frac{472{,}71}{36} - 3{,}6^2 = 0{,}2;$$

$$\sigma_x = \sqrt{0{,}2} = 0{,}447; \quad \sigma_y^2 = \frac{\Sigma y^2}{n} - \overline{y}^2 = \frac{16910{,}7 \cdot 10^5}{36} - 6765^2 =$$

$$= 1130000; \quad \sigma_y = 100\sqrt{113} = 1063.$$

Тогда $r = \dfrac{24500 - 3{,}6 \cdot 6765}{0{,}447 \cdot 1063} = \dfrac{200}{470} = 0{,}420;$

$$b = 0{,}420 \cdot \frac{1063}{0{,}447} = 100; \quad a = 6765 - 100 \cdot 3{,}6 = 3165;$$

$$\overline{y} = 3165 + 100x. \tag{II.10}$$

В табл. 22 имеются все данные для исследования корреляционной зависимости с помощью типизированной решетки (см. рис. 9). Предоставляем читателю самому выполнить все вычисления, проверив правильность уравнения (II.10). Незначительные отклонения могут возникнуть только в результате более или менее удачного выбора интервалов

Так же, как установлена количественная характеристика зависимости производительности труда от квалификации рабочих, могут быть определены связи с другими факторами. Например, связь с коэффициентом полезного фонда рабочего времени ($K_{пв}$) выражается уравнением $\overline{y} = 2100 + 23{,}7K_{пв}$; связь с уровнем механизации монтажа сборных железобетонных конструкций ($C_{жк}$, проц.) — уравнением $\overline{y} = 2450 + 16{,}3C_{жк}$ и т. д. Подобные уравнения могут быть использованы при анализе резервов и расчетах возможного повышения производительности труда в строительстве.

Выводы.

1. *Расчет и анализ линий (кривых) регрессии*, которыми завершается корреляционный анализ, настолько важен, что ряд авторов (например, см. работы [6], [12]) выделяет

их в самостоятельную область — регрессионный анализ.

Задача регрессионного анализа — вычисление параметров a и b и их статистических характеристик в уравнениях типа $y = a + bx$. Если экономические связи между явлениями строительного производства нелинейны относительно параметров, то используются те или иные методы преобразования таких кривых (нелинейных) регрессий в функции, линейные относительно параметров (см. [12], [13]). Наиболее важным в регрессионном анализе линейных зависимостей является рассмотрение сущности коэффициента b при переменной x. Этот коэффициент называют *угловым коэффициентом регрессии, или коэффициентом регрессии* y по x. Он показывает, на сколько единиц в среднем изменяется y, когда x изменяется на единицу.

Из уравнения $y = a + bx$ следует, что a показывает то значение y, которое он принимает при $x = 0$.

Оценка достоверности коэффициентов регрессии b выполняется следующим образом.

Определяется среднее квадратическое отклонение σ_b по формуле:

$$\sigma_b = \frac{\sigma_y}{\sigma_x} \cdot \frac{1-r^2}{\sqrt{n}}. \qquad (II.11)$$

В данном случае $\sigma_b = \frac{1,016}{1,269} \cdot \frac{1-0,839^2}{\sqrt{74}} =$

$$= \frac{1,016}{1,269} \cdot \frac{0,296}{8,61} = 0,0275.$$

Если мы хотим оценить значение b с доверительной вероятностью $p = 91\%$, то по таблице Стьюдента (приведенной, например, в пособии [5], стр. 293) находим значение $t = 1,70$. Тогда доверительные интервалы будут равны:

$$\sigma_y - 1,70\sigma_b < b < \sigma_y + 1,70\,\sigma_b;$$

$$1,016 - 1,70 \cdot 0,0275 < b < 1,016 + 1,70 \cdot 0,0275;$$

$$0,969 < b < 1,063.$$

Следовательно, при выбранной вероятности значение b находится в пределах от 0,969 до 1,063.

При $p = 0,95$ эти пределы увеличиваются от 0,962 до 1,070.

2. *Значения коэффициентов корреляции 0,839* (см. рис. 9) *и 0,883* (см. стр. 68) *являются показателями тесноты связи* между признаком (уровень накладных расходов) и выбранными факторами (объем работ и численность работников) и определены на основании только 74 наблюдений. Для того чтобы выявленную тесноту связи между исследуемыми показателями можно было распространить на всю генеральную совокупность аналогичных строительных организаций, необходимо определить доверительные интервалы для коэффициентов корреляции. С этой целью рекомендуем пользоваться графиком, представленным в работе [22, стр. 155], по которому для $n = 74$ и доверительной вероятности 0,95 значения r в генеральной совокупности находятся в пределах

$$|0{,}70| < |0{,}839| < |0{,}90|.$$

Доверительные пределы указывают, что в 95% всех случаев в генеральной совокупности теснота связи между уровнем накладных расходов и объемами выполняемых строительно-монтажных работ будет характеризоваться значением r, которое не меньше 0,70 и не выше 0,90.

3. *Для определения минимально необходимого количества наблюдений n_{min} рекомендуется табл. 23.*

Таблица 23

Определение числа наблюдений n_{min} по показателю вариации v

v% \ $\Phi(t)$	0,683	0,7	0,75	0,8	0,85	0,9	0,95
при $p = \pm 5\%$							
10	4	4	5	7	8	11	15
15	9	10	12	15	19	24	35
20	16	17	21	26	33	43	62
25	25	37	33	41	52	67	95
30	36	39	47	59	75	97	138
при $p = \pm 10\%$							
10	1	1	1	2	2	3	4
15	2	2	3	4	5	6	9
20	4	4	5	7	8	11	15
25	6	7	8	10	13	17	24
30	9	10	12	15	19	24	35

Значения n_{min} определяются в зависимости от принимаемых величин коэффициента вариации v, достоверности $\Phi(t)$ и точности результатов p.

Принимая, например, достоверность $\Phi(t) = 0{,}75$, мы заранее соглашаемся, что из 100 только в 75 случаях будет подтверждена правильность выводов.

4. *Необходимо помнить, что $y = f(x)$ и $x = \varphi(y)$ на корреляционных таблицах необратимы.* Если мы на рис. 9 поменяем местами оси и разместим те же исходные данные в координатных осях $x = f(y)$, то получим совершенно новую зависимость, не имеющую прямой связи с зависимостью $y = 53{,}4 + 0{,}134x$. В новых условных координатах (обозначим их x''):

$$\overline{x'} = \overline{y''} = 0{,}284;\ \overline{y'} = \overline{x''} = -0{,}095;$$

$$\sigma'_x = \sigma''_y = 1{,}016;\ \sigma'_y = \sigma''_x = 1{,}269;$$

$$r' = r'' \frac{74 \cdot 78 + 21 \cdot 7}{\sqrt{77 \cdot 74 - 7^2}\ \sqrt{125 \cdot 74 - 21^2}} = \frac{5919}{75{,}1 \cdot 93{,}9} = 0{,}839.$$

Используя формулы (II.5) и (II.6), имеем:

$$b'' = 0{,}839 \cdot \frac{1{,}269}{1{,}016} = 1{,}047;\ a' = 0{,}284 + 1{,}047 \cdot 0{,}095 = 0{,}293;$$

$$\overline{y''} = 0{,}293 + 1{,}047x''.$$

Вывод о необратимости функций $y = f(x)$ и $x = \varphi(y)$, иллюстрируемый этим примером, поможет читателю избежать ошибок при выборе осей для переменных величин на корреляционных таблицах. Признак (y) размещается только на оси ординат, фактор (x) — только на оси абсцисс, несмотря на то, что величина коэффициента корреляции при перемене осей не изменяется.

Отметим, что постоянство величины коэффициента корреляции при перемене координатных осей вытекает и из (II.5), по которой $r_{xy} = b' \frac{\sigma'_x}{\sigma'_y}$ и в то же время $r_{yx} = b'' \frac{\sigma''_y}{\sigma''_x}$, в данном примере: $0{,}672 \cdot \frac{1{,}269}{1{,}016} = 1{,}047 \cdot \frac{1{,}016}{1{,}269} = 0{,}839.$

5. *Вычисления парного коэффициента корреляции, параметров уравнения регрессии и другие корреляционные расчеты могут выполняться на счетно-перфорационных машинах.* Техника этих расчетов изложена в немногих пособиях, из которых рекомендуем (Б. Б. Розин, Р. С. Гейфман [18]).

§ 3. Параболические зависимости (парабола второго порядка)

Исследование зависимости использования мощности строительной организации от ритмичности ее работы

Экономическая постановка задачи. Под производственной мощностью строительно-монтажной организации понимается максимальный годовой объем строительно-монтажных работ в стоимостном выражении, который может быть выполнен данной организацией при полном использовании ее основных производственных фондов в соответствии с заданным режимом работы и применением передовой технологии и организации строительства.

Анализ производственно-хозяйственной деятельности строительных организаций различного профиля показал, что имеются резервы увеличения их мощности за счет улучшения использования основных производственных фондов, правильной организации труда рабочих, совершенствования технологии производства. Степень использования мощности строительной организации выражается коэффициентом $K_и$

$$K_и = \frac{P}{M},$$

где P — годовой объем работ;

M — производственная мощность.

Логично ожидать, что значения $K_и$ будут изменяться с изменением ритмичности выполнения строительно-монтажных работ, так как в ритмичности получает количественное выражение степень организованности, слаженности, бесперебойности работы строительных подразделений.

Для исследования связи между коэффициентом использования мощности $K_и$ и степенью ритмичности выполнения строительно-монтажных работ R были отобраны следующие данные за 1964 и 1965 гг. о 25 однотипных трестах, специализированных на возведении жилищно—гражданских объектов (см. табл. 24).

Значения коэффициента $K_и < 1$ свидетельствуют о недостаточном использовании основных фондов, значения $K_и > 1$ указывают на чрезмерное применение ручного труда при недостатке или нерациональном использовании основных средств.

Таблица 24

Исходные данные для корреляционных расчетов

№ п/п	Наименование трестов	1964 год		1965 год	
		R, %	$K_н$	R, %	$K_н$
1	Жилстрой-1	88,4	0,95	66,6	1,10
2	Жилстрой-2	85,1	0,84	68,3	0,95
3	Горстрой-1	88,3	1,15	86,0	1,10
4	Горстрой-3	87,7	0,87	70,0	0,97
5	Жилстрой-3	91,3	1,00	65,4	0,78
6	Жилстрой-4	92,3	0,77	67,2	0,68
7	Жилстрой-4	95,0	0,83	69,0	0,57
8	Горстрой-2	93,1	0,69	98,1	0,69
9	Горстрой-4	96,3	0,65	64,3	0,50
10	Горстрой-5	95,2	0,54	63,1	0,48
11	Жилстрой-5	83,4	1,27	81,3	0,84
12	Жилстрой-6	82,9	1,15	83,7	1,28
13	Жилстрой-7	82,1	0,98	84,2	1,43
14	Жилстрой-8	80,9	0,83	81,8	0,97
15	Жилстрой-9	82,2	1,10	62,7	0,38
16	Жилстрой-10	76,4	1,27	75,8	1,05
17	Жилстрой-11	78,1	1,20	77,1	1,08
18	Жилстрой-12	76,1	1,18	75,8	0,86
19	Жилстрой-13	79,8	1,38	78,7	0,95
20	Жилстрой-14	77,4	1,10	79,1	1,42
21	Жилстрой-15	78,2	1,40	76,8	1,00
22	Жилстрой-16	71,2	0,70	70,5	0,55
23	Жилстрой-17	73,1	0,90	72,7	1,15
24	Жилстрой-18	70,7	0,95	74,3	1,35
25	Жилстрой-19	71,3	1,20	73,8	1,10

Значения R и $K_н$, размещенные на типизированной корреляционной решетке, предположительно указывают на параболическую зависимость $K_н$ от R (см. табл. 25 на рис. 11).

Проверка исходного материала, представленного строкой a и столбцом u, на соответствие нормальному распределению по критерию Пирсона по оси y при $\chi^2 = 2{,}8$ и числе степеней свободы $k = 7 - 2 = 5$ дает искомую вероятность $p(\chi^2) = 0{,}7$; по оси x — соответственно ($\chi^2 = 2{,}26$; $k = 5$) $p(\chi^2) = 0{,}85$.

χ^2 определялся на ЭВМ. Нижний предел $p(\chi^2)$ принимается 0,001; значения $\chi^2 > 0{,}001$ указывают, что расхождения между теоретическими и эмпирическими частотами можно считать случайными.

Таблица 25

	x'		-3	-2	-1	0	1	2	3				ЦУУ	Итого		
y'	$y \backslash x$					80,40							ЦРУ	и	к	л
			59,75	65,65	71,55	77,45	83,35	89,25	95,15	100		R	РУ	n_y	$n_y y'$	$n_y (y')^2$
ЦУУ	ЦРУ	РУ														
		K_u														
3		1,5175			1^3	3^9	1^3							5	15	45
2		1,3425		1^2	2^4	1^2	2^4							6	12	24
1		1,1675		1^1	6^6	2^2	2^2							11	11	11
0	0,9058	0,9925		3	2	5	3	2						15	0	0
-1		0,8175	1^{-1}	2^{-2}				2^{-2}	2^{-2}					7	-7	7
-2	0,554	0,6425	2^{-4}	2^{-4}					1^{-2}					5	-10	20
-3		0,4675 0,2925	1^{-3}											1	-3	9
а	n'_x		4	9	11	11	8	4	3					$\boxed{n}$ 50	18	116
б	$n_x \cdot x'$		-12	-18	-11	0	8	8	9					-16		
в	$n_x (x')^2$		36	36	11	0	8	16	27					134		
г	$\Sigma n_i y'$		-8	-3	13	13	9	-2	-4					18		
д	$x' \Sigma n_i y'$		24	6	-13	0	9	-4	-12					10		
м	$n_x (x')^3$		-108	-72	-11	0	8	32	81					-70		
н	$n_x (x')^4$		324	144	11	0	8	64	243					794		
р	$x' n_i (y')^2$		-72	-12	13	0	9	-8	-36					-106		
е	$\bar{y}'$		-2	-0,333	1,182	1,182	1,125	-0,5	-1,333						$i_y = 0,175$	
ж	$\bar{y}' i_y$		[illegible]35	-0,06	0,21	0,21	0,20	-0,09	-0,23						$i_x = 5,9$	
з	$\bar{y}$		0,56	0,85	1,11	1,11	1,10	0,82	0,67						$y_0 = 0,905$ $x_0 = 80,40$	

Рис. 11. Корреляционная таблица расчета связи между использованием мощности строительных организаций и ритмичностью их работы (параболический тип зависимости).

В данном случае величины $p = (\chi^2)$ указывают на весьма высокое соответствие эмпирического распределения нормальному.

Общая характеристика рядов распределения для зависимости $K_и = f(R)$ приведена на рис. 12.

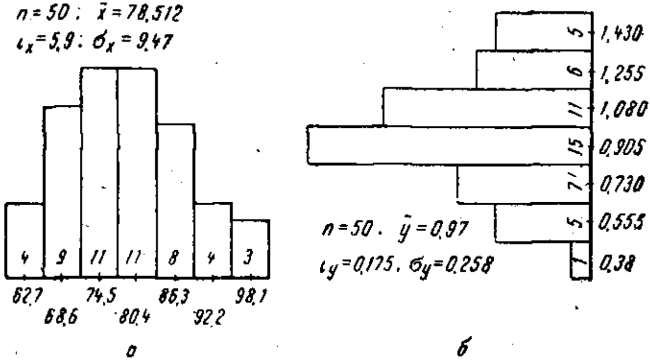


Рис. 12. Характеристика рядов распределения:
а — степени ритмичности; *б* — коэффициента использования мощности строительных организаций.

Кратко проследим ход расчета с помощью корреляционной табл. 25.

Строки *а*, *б*, *в*, *г*, *д* табл. 25 заполняются как при анализе линейной зависимости (см. рис. 9).

Значения строки *а*: 4, 9, 11, 11, 8, 4, 3, $\Sigma = 50$ представляют собой ряд распределения показателя ритмичности R. Значения последующих строк приведены в табл. 25. Новыми элементами табл. 25 являются строки *м*, *н*, *р*. Значения строки $(n_x x')^3$ получаем, умножая каждый член строки *в* на соответствующие условные ординаты.

$$n_x (x'^2)\, x' = n_x x'^3;$$

$36 \cdot (-3) = -108; 36 \cdot (-2) = -72;\ 11 \cdot (-1) = -11; 0 \cdot 0 = 0$
$8 \cdot 1 = 8;\ 16 \cdot 2 = 32;\ 27 \cdot 3 = 81;\ \Sigma = -70;$
Аналогично образуется строка *н*:

$$n_x \cdot (x'^3)\, x' = n_x x'^4;$$

$$-108 \cdot (-3) = 324;\ -72 \cdot (-2) = 144;\ -11 \cdot (-1) = 11;$$
$$0 \cdot 0 = 0;\ 8 \cdot 1 = 8;\ 32 \cdot 2 = 64;\ 81 \cdot 3 = 243;\ \Sigma = 794.$$

Значения строки $x'^2 \sum n_i x'$ получаем умножением каждого члена строки $x' \sum n_i x'$ на соответствующие условные ординаты:

$24\cdot(-3) = -72$; $6\cdot(-2) = -12$; $-13(-1) = 13$; $0\cdot 0 = 0$; $9\cdot 1 = 9$; $-4\cdot 2 = -8$; $-12\cdot 3 = -36$; $\Sigma = -106$.

Последние три строки таблицы образуются как в корреляционной таблице на рис. 9.

$$е = \frac{г}{а} \to \frac{-8}{4} = -2; \frac{-3}{9} = -0{,}333; \frac{13}{11} = 1{,}182;$$

$$\frac{13}{11} = 1{,}182; \frac{9}{8} = 1{,}125; \frac{-2}{4} = -0{,}500; \frac{-4}{3} = -1{,}333.$$

$$ж = е \cdot i_y \to -2\cdot 0{,}175 = -0{,}350;\ -0{,}333\cdot 0{,}175 = -0{,}058;$$

$$1{,}182\cdot 0{,}175 = 0{,}207;\quad 1{,}125\cdot 0{,}175 = 0{,}197;$$

$$-0{,}5\cdot 0{,}175 = -0{,}088;\quad -1{,}333\cdot 0{,}175 = -0{,}233.$$

$$з = ж + y_0 \to -0{,}350 + 0{,}905 = 0{,}555;$$

$$-0{,}058 + 0{,}905 = 0{,}847;\ 0{,}207 + 0{,}905 = 1{,}112;\ 1{,}112;$$

$$0{,}197 + 0{,}905 = 1{,}102;\ -0{,}088 + 0{,}905 = 0{,}817;$$

$$-0{,}233 + 0{,}905 = 0{,}672.$$

Графы *к*, *л* заполняются как и в корреляционной таблице 20 (см. рис. 9). Как и ранее, контролем правильности выполнения расчетов может служить выражение $\Sigma г = \Sigma к$, здесь равное 18.

Запишем (в условных координатах) систему уравнений для определения уравнения искомой параболы:

$$\begin{cases} na + n_x x' в + n_x x'^2 c = \sum n_i y'; \\ n_x x' a + n_x x'^2 в + n_x x'^3 c = x' \sum n_i y'; \\ n_x x'^2 a + n_x x'^3 в + n_x x'^4 c = x'^2 \sum n_i y'. \end{cases} \qquad \text{(II.12)}$$

$n_x x'^3 a + n_x x'^4 в + n_x x'^5 c =$

Или в принятых нами упрощенных обозначениях:

$$\begin{cases} na + \sum б \cdot b + \sum в \cdot c = \sum г; \\ \sum б \cdot a + \sum в \cdot b + \sum м \cdot c = \sum д; \\ \sum в \cdot a + \sum м \cdot b + \sum н \cdot c = \sum р. \end{cases}$$

Подставим в уравнения значения величин из табл. 25:

$$\begin{cases} 50a-16b+134c=18; \\ -16a+134b-70c=10; \\ 134a-70b+794c=-106. \end{cases}$$

Решив систему уравнений, найдем:

$$a=1{,}31713;\quad в=0{,}04835;\quad c=-0{,}35137.$$

Тогда

$$\bar{y}'=1{,}31713+0{,}04835x'-0{,}35137\,(x')^2. \qquad (II.13)$$

Для перехода к реальным координатам примем:

$$y'=\frac{y-C_y}{i_y}=\frac{y-0{,}905}{0{,}175} \text{ и } x'=\frac{x-C_x}{i_x}=\frac{x-80{,}4}{5{,}9}.$$

Подставим значения y' и x' в уравнение (II.13):

$$\frac{\bar{y}-0{,}905}{0{,}175}=1{,}31713+0{,}04835\,\frac{x-80{,}4}{5{,}9}-$$

$$-0{,}35137\left(\frac{x-80{,}4}{5{,}9}\right)^2,$$

откуда

$$\bar{y}=-10{,}39827+0{,}28548\,x-0{,}001766\,x^2$$

или

$$K_и=-10{,}3983+0{,}28548\,R-0{,}001766\,R^2. \qquad (II.13а)$$

При нелинейных (криволинейных) зависимостях теснота связи характеризуется корреляционным отношением — η. Величина корреляционного отношения меняется от 0 до 1 и, таким образом, может быть только положительной; при $\eta=1$ корреляционная зависимость превращается в функциональную.

Определение η значительно сложнее, чем r. Оно исчисляется по следующей формуле (для величин в условных координатах):

$$\eta=\sqrt{\frac{(\delta_i')^2}{(\sigma')^2}}, \qquad (II.14)$$

где

$$(\delta_i')^2=\frac{\Sigma\left(\bar{y}_x'-\bar{y}'\right)^2 n_x}{\Sigma n_x};$$

$$(\sigma')^2=\frac{\Sigma n_y\,(y')^2}{\Sigma n_x}-\left(\bar{y}'\right)^2. \qquad (II.15)$$

Таблица 26

Расчетные данные для определения $(\delta)^2$

n_x	$\bar{y}'_x = \frac{\bar{y}_x - y_0}{i_y}$	$\bar{y}'_x - \bar{y}'$	$(\bar{y}'_x - \bar{y}')^2$	$(\bar{y}'_x - \bar{y}')^2 \cdot n_x$
4	$\frac{0,557-0,905}{0,175} = -1,990$	$-1,990-0,36 = -2,350$	5,523	22,092
9	$\frac{0,873-0,905}{0,175} = -0,185$	$-0,185-0,36 = -0,545$	0,297	2,672
11	$\frac{1,066-0,905}{0,175} = 0,918$	$0,918-0,36 = 0,557$	0,311	3,419
11	$\frac{1,136-0,905}{0,175} = 1,317$	$1,317-0,36 = 0,957$	0,916	10,081
8	$\frac{1,083-0,905}{0,175} = 1,016$	$1,014-0,36 = 0,654$	0,428	3,425
4	$\frac{0,920-0,905}{0,175} = 0,006$	$0,085-0,36 = -0,275$	0,075	0,300
3	$\frac{0,608-0,905}{0,175} = -1,700$	$-1,700-0,36 = -2,060$	4,243	12,730
$\Sigma n_x = 50$			$\Sigma (\bar{y}_x - \bar{y}')^2 n_x =$	54,719

$$(\bar{\delta}')^2 = \frac{54,719}{50} = 1,094; \qquad \eta = \sqrt{\frac{1,094}{2,190}} = 0,7068$$

$\overline{y}'_x$ в условных ксординатах находится по формуле[1]

$$\overline{y}'_x = \frac{\overline{y}_x - y_0}{i_y}. \qquad (II.16)$$

$\overline{y}'$ и $(\sigma')^2$ определяются непосредственно по данным табл. 25:

$$\overline{y}' = \frac{\Sigma n_i \cdot y'}{\Sigma n} = \frac{18}{50} = 0{,}36; \quad (\sigma')^2 = \frac{\Sigma n_y y'}{\Sigma n} - \left(\frac{\Sigma n_i y'}{\Sigma n}\right)^2 =$$

$$= \frac{116}{50} - 0{,}36^2 = 2{,}190.$$

Весь расчет сведен в табл. 26.

Средняя квадратическая ошибка корреляционного отношения

$$\sigma_\eta = \frac{1 - 0{,}7068^2}{\sqrt{50}} = \frac{1 - 0{,}4996}{7{,}071} = 0{,}0708.$$

Критерий надежности $m = \frac{\eta}{\sigma_\eta} = \frac{0{,}7068}{0{,}0708} = \sim 10$. η^2 показывает, какая часть колеблемости признака y была обусловлена изменчивостью фактора x. В рассматриваемом примере величина $\eta^2 = 0{,}7068^2 = 0{,}5$ указывает, что коэффициент использования производственной мощности исследуемой группы трестов на 50% обусловлен колебаниями ритмичности их работы. В некоторых работах, в частности [5], мерой связи дисперсии признака с дисперсией фактора принимается не η^2, а непосредственно η.

В случае линейной связи корреляционное отношение[2] равно коэффициенту корреляции (абсолютной его величине). Выполненные расчеты показывают, что корреляционная связь между R и $K_и$ представляется вполне закономерной, отражающей экономический характер зависимости между ними. Аналогичные расчеты по трестам, специализированным на промышленном строительстве, показали: по 16 трестам, ведущим строительство объектов черной металлургии (трест — представитель «Металлургстрой»), 12 трестам, специализированным на возведении объектов химической промышленности (трест—представитель «Хим-

---

[1] Значения $\overline{y}_x$ — теоретические ординаты параболы, подсчитанные для середины интервалов по R, находим по уравнению (II.13)

[2] Имеется в виду теоретическое корреляционное отношение $\eta_т$, которое практически незначительно отличается от рассмотренного выше эмпирического η.

Таблица 27

Динамика объемов строительно-монтажных работ на возведении прокатных станов
(y, в % к общему объему работ)

Наименование объектов	Продолжительность строительства в месяцах															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Трубоэлектросварочный цех Южнотрубного завода	3,6	5,2	7,7	14,3	14,4	15,8	14,1	13,5	12,4							
Комплекс мелкосортного стана Криворожского металлургического завода им. Ленина	1,6	6,0	8,7	12,7	18,2	20,9	15,2	12,0	4,5	0,2						
Стан 600, Коммунарский завод	3,9	6,1	9,6	11,4	13,2	16,0	15,6	12,5	8,6	4,1						
Стан 250-2 Макеевского завода им. Кирова	3,2	6,2	9,8	10,5	10,4	10,7	11,8	11,8	11,1	9,9	4,6					
Стан 2800 цеха легированных сталей. Запорожсталь	1,4	3,9	7,1	9,8	10,4	10,9	10,3	10,8	9,8	10,1	9,4	6,1				
Цех холодной прокатки легированной стали, Запорожсталь	0,8	2,6	4,6	7,1	9,0	12,8	12,3	16,9	14,5	11,4	6,5	1,5				

строй») и 10 трестам, специализированным на объектах машиностроения (трест-представитель «Машстрой»), полученные значения η ниже, чем по трестам жилищного строительства, и соответственно равны 0,679; 0,571; 0,62. Причиной этого является бóльшая, чем в жилищном строительстве, неоднородность возводимых объектов. Тем не менее значения η позволяют утверждать, что ритмичность оказывает влияние также и на использование производственной мощности строительных организаций, ведущих промышленное строительство.

Исследование зависимости использования производственной мощности строительных организаций от ритмичности их деятельности с применением корреляционного анализа дало возможность ввести показатель ритмичности в нормативы использования мощности строительных организаций и тем самым получить более точную оценку резервов имеющихся мощностей.

Распределение объемов строительно-монтажных работ во времени для определения величины задела

При составлении перспективных планов капитального строительства важная роль принадлежит обоснованию величины переходящих заделов. Заделом строительно-монтажных работ на переходящих, незавершенных объектах (комплексах) промышленного строительства называется объем работ, который должен быть выполнен в предшествующем периоде для обеспечения своевременного ввода объектов (комплексов) в эксплуатацию и создания условий для ритмичной работы строительных организаций.

Специальными исследованиями (выполнены инженером М. С. Нейманом[1]) установлено, что объем работ на задельных объектах находится в определенной зависимости от продолжительности строительства. Исследуем зависимость распределения объемов работ от продолжительности возведения объектов, пользуясь данными о динамике объемов строительно-монтажных работ при строительстве группы прокатных цехов, приведенным в табл. 27.

Анализ данных табл. 27 приводит к следующему выводу: для всех объектов характерными являются нарастание и

[1] См. «Методика определения величины задела строительно-монтажных работ с использованием корреляционного анализа и ЭВМ». Научно-исследовательский институт строительного производства Госстроя УССР, Киев, 1964.

последующий спад объёмов работ в единицу времени, имеющие параболический характер. Параболический характер динамики объемов работ на отдельных объектах черной металлургии (см. рис. 13) дает возможность устанавливать величину задела при перспективном (в пределах 5—7 лет) планировании строительства; выявленные закономерности могут также использоваться при планировании расходования ресурсов в зависимости от объемов работ.

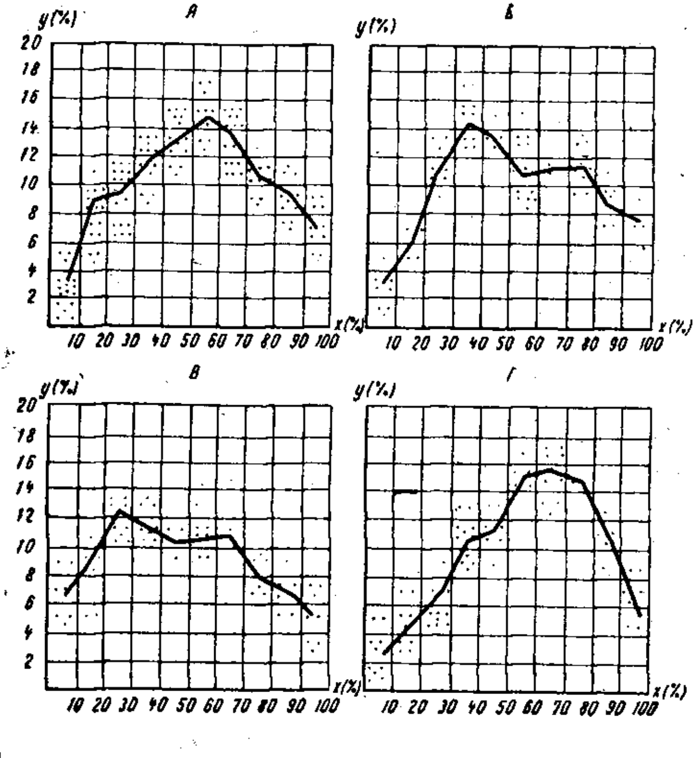


Рис. 13. Распределение объема строительно-монтажных работ (y) в зависимости от продолжительности возведения объектов (x).
А — сталепрокатные цехи; Б — ремонтно-вспомогательные цехи; В — горно-обогатительные комбинаты; Г — горнорудные шахты.

Исходные данные для решения системы нормальных уравнений приняты по данным табл. 27, сумма значений объемов работ, распределенных по месяцам, y (%); удельный вес месяцев в общей продолжительности возведения

Продолжение табл. 27

Наименование объектов	Продолжительность строительства в месяцах															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Комплекс блюминга Енакиевского металлургического завода	4,0	6,4	7,2	8,8	9,5	10,4	10,7	10,7	10,9	11,5	7,7	2,2				
Стан 650, Тагилстрой по графику	3,0	5,2	8,0	8,9	8,9	9,7	10,6	11,5	11,5	9,7	8,9	4,1				
фактически	2,7	4,9	7,2	8,2	8,2	7,2	9,8	9,9	10,9	12,2	10.8	8,0				
Комплекс цеха шарикоподшипниковых труб, з-д им. Либкнехта	3,3	6,5	9,5	11,5	13,5	15,0	14,1	11,8	8,8	3,7	1,5	0,8				
Листопрокатный стан 2800 Череповецкого металлургического завода	0,7	0,7	6,6	8,2	9,4	12,4	14,0	9,2	10,3	9,7	9,7	7,1	2,0			
Слябинг 1150 завода им. Ильича, г. Жданов	0,3	3,1	3,4	5,0	7,1	6,7	9,9	11,5	12,7	12,8	11,2	8,6	3,4	2,0	1,4	0,9

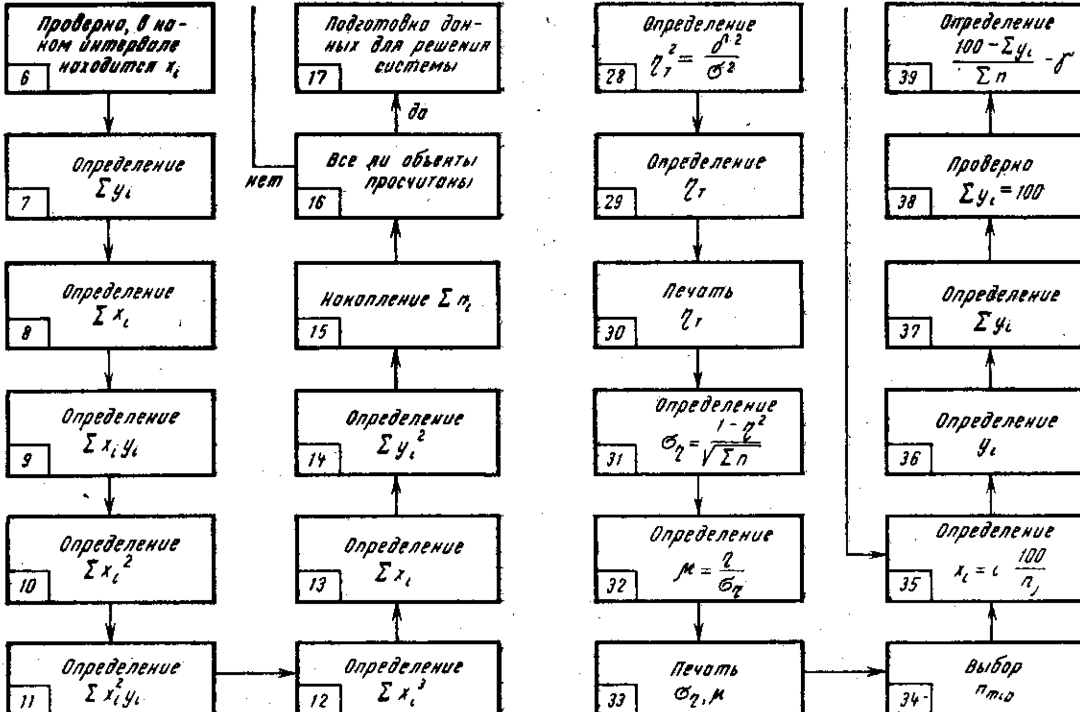


Рис. 14. Блок-схема программы решения системы нормальных уравнений для ЭВМ «Раздан-2».

объектов x (%)[1]:

$$\Sigma n = 141;\ \Sigma y = 1200;\ \Sigma x = 7650;\ \Sigma x^2 = 531{,}7 \cdot 10^3;$$

$$\Sigma x^3 = 415 \cdot 10^5;\ \ \Sigma x^4 = 3455 \cdot 10^6;\ \ \Sigma xy = 68{,}7 \cdot 10^3;$$

$$\Sigma x^2 y = 458{,}2 \cdot 10^4.$$

Решение системы уравнений по этим исходным данным приводит к зависимости:

$$\bar{y} = -2{,}228 + 0{,}496x - 0{,}00429x^2.$$

Корреляционное отношение $\eta = 0{,}79$, и, таким образом, связь между y и x является достаточно надежной.

Аналогичным путем рассчитаны уравнения для возведения наиболее характерных объектов черной металлургии: для горнообогатительных комбинатов

$$\bar{y} = 5{,}60 + 0{,}280x - 0{,}003x^2 \text{ (при } \eta = 0{,}60);$$

для доменных печей

$$\bar{y} = -2{,}439 + 0{,}560x - 0{,}005x^2 \text{ (при } \eta = 0{,}76);$$

для железорудных шахт

$$\bar{y} = 1{,}435 + 0{,}489x - 0{,}005x^2 \text{ (при } \eta = 0{,}71).$$

В целом для строительства в отрасли черной металлургии

$$\bar{y} = 0{,}84 + 0{,}47x - 0{,}004x^2 \text{ (при } \eta = 0{,}70).$$

На рис. 14 представлена блок-схема программы (для ЭВМ «Раздан-2») решения системы нормальных уравнений.

Наличие таких уравнений (или построенных на их основе таблиц, более удобных в применении) позволяет при планировании размеров капитальных вложений и определении величины заделов, в условиях отсутствия детальной технической документации по отдельным объектам, естественного при планировании на перспективу нескольких лет вперед, устанавливать объемы строительно-монтажных работ в зависимости от продолжительности возведения объектов.

На основании достаточно большого числа данных по проектам и строительству объектов-представителей аналогичные уравнения (или таблицы) могут быть рассчитаны для других видов промышленного строительства.

Следует предупредить читателя, что при решении систем нормальных уравнений пользование логарифмической ли-

[1] Для упрощения расчетов значения x уменьшены в 100 раз, итоговые результаты приведены в истинном масштабе.

нейкой может привести к результатам, резко отличающимся от правильных.

Для доказательства решим систему по данным, приведенным на стр. 87, пользуясь линейкой. Выравнивание коэффициентов при a приводит к системе:

$$a + 54{,}3\,b + 3{,}78 \cdot 10^3 c = 8{,}52;$$
$$a + 69{,}5\,b + 5{,}37 \cdot 10^3 c = 9{,}00;$$
$$a + 78{,}0\,b + 6{,}48 \cdot 10^3 c = 8{,}62.$$

Дальнейшие действия выполняются в обычном порядке:

$$15{,}2\,b + 1{,}59 \cdot 10^3 c = 0{,}48;$$
$$23{,}7\,b + 2{,}70 \cdot 10^3 c = 0{,}10;$$
$$b + 1{,}045 \cdot 10^2 c = 0{,}0316;$$
$$\underline{b + 1{,}140 \cdot 10^2 c = 0{,}0042}$$
$$9{,}5c = -0{,}0274; \quad c = -0{,}00285;$$

$$b = 0{,}3296; \; a = 1{,}42 \text{ и } \bar{y} = 1{,}42 + 0{,}3296x - 0{,}00285x^2.$$

Это уравнение, как видим, значительно отличается от уравнения, полученного с помощью ЭВМ.

Таким образом, пользование линейкой привело к резким изменениям (и даже изменению знаков) коэффициентов уравнения параболы; чувствительность линейки недостаточна для решения подобных задач. Указанные расчеты с успехом могут выполняться также при помощи счетно-перфорационных вычислительных машин (см. [18]).

§ 4. Гиперболические зависимости

Зависимость уровня себестоимости от объема работ

Экономическая постановка задачи. Выясним, как влияет на уровень себестоимости y такой важный показатель, как годовой объем работ x. С экономической точки зрения подобная зависимость представляется логичной.

В однотипных строительных подразделениях уровень стоимости обычно ниже там, где больше годовой объем работ. Эта тенденция объясняется тем, что в крупных подразделениях применяются более производительные механизмы, более прогрессивные методы организации труда. Немаловажное значение имеет и меньшая доля накладных расходов крупных строительных организациях, поскольку постоянная часть накладных расходов относится на больший объём работ.

Найдем математическое выражение этой связи и оценим ее существенность.

Исходным материалом являются данные 116 строительно-монтажных управлений, специализированных на выполнении санитарно-технических работ (см. табл. 28). Размещение точек на корреляционном поле указывает, что наиболее вероятной формой связи в данном случае будет гиперболическая связь признака y с фактором x.

Таблица 28

Исходные данные для корреляционных расчетов

№ СУ	Годовой объем работ (млн. руб.) x	Уровень себестоимости (%) y	№ СУ	Годовой объем работ (млн. руб.) x	Уровень себестоимости (%) y	№ СУ	Годовой объем работ (млн. руб.) x	Уровень себестоимости (%) y	№ СУ	Годовой объем работ (млн. руб.) x	Уровень себестоимости (%) y
1	0,95	0,97	30	1,80	0,94	59	1,62	0,787	88	2,70	0,79
2	0,91	0,935	31	1,78	0,92	60	1,68	0,80	89	2,60	0,76
3	1,00	0,95	32	1,85	0,92	61	1,80	0,805	90	2,90	0,82
4	1,10	0,94	33	1,65	0,88	62	2,10	0,875	91	3,00	0,83
5	0,98	0,92	34	1,70	0,875	63	2,25	0,88	92	3,10	0,825
6	0,99	0,925	35	1,75	0,88	64	2,30	0,89	93	2,90	0,785
7	1,00	0,89	36	1,90	0,885	65	2,10	0,845	94	3,00	0,79
8	1,20	0,86	37	1,95	0,90	66	2,15	0,85	95	3,10	0,805
9	1,115	0,858	38	1,87	0,875	67	2,20	0,857	96	2,90	0,77
10	1,350	0,97	39	1,90	0,89	68	2,25	0,86	97	2,95	0,76
11	1,27	0,94	40	1,850	0,887	69	2,30	0,865	98	3,10	0,755
12	1,40	0,95	41	1,78	0,89	70	2,35	0,855	99	3,30	0,83
13	1,50	0,92	42	1,65	0,845	71	2,37	0,85	100	3,27	0,79
14	1,45	0,91	43	1,68	0,85	72	2,18	0,86	101	3,35	0,80
15	1,58	0,92	44	1,70	0,86	73	2,115	0,815	102	3,80	0,805
16	1,23	0,875	45	1,72	0,867	74	2,20	0,82	103	3,70	0,79
17	1,37	0,88	46	1,75	0,85	75	2,225	0,83	104	3,27	0,78
18	1,45	0,875	47	1,80	0,855	76	2,30	0,835	105	3,30	0,77
19	1,50	0,89	48	1,88	0,845	77	2,20	0,785	106	2,60	0,83
20	1,58	0,88	49	1,90	0,85	78	2,30	0,80	107	4,10	0,79
21	1,21	0,845	50	1,95	0,847	79	2,28	0,76	108	2,65	0,80
22	1,40	0,86	51	1,98	0,865	80	2,47	0,845	109	2,78	0,805
23	1,38	0,85	52	1,87	0,86	81	2,60	0,85	110	4,10	0,77
24	1,50	0,865	53	1,67	0,815	82	2,70	0,86	111	2,67	0,76
25	1,27	0,82	54	1,70	0,82	83	2,50	0,865	112	4,40	0,82
26	1,35	0,83	55	1,75	0,825	84	2,58	0,82	113	2,80	0,83
27	1,40	0,835	56	1,81	0,83	85	2,60	0,83	114	4,36	0,79
28	1,55	0,838	57	1,92	0,835	86	2,70	0,835	115	3,70	0,80
29	1,47	0,77	58	1,88	0,827	87	2,50	0,80	116	2,54	0,76

Таблица 29

	x	0,910	1,346	1,782	2,218	2,654	3,090	3,526	3,962	4,398		
у		0,692	1,128	1,564	2,000	2,436	2,872	3,308	3,744	4,180	а	к
0,9705	0,957	0,9705 1	0,9705 1								2	1,9410
0,9435	0,930	2,8305 3	1,8870 2	0,9435 1							6	5,6610
0,9165	0,903	1,8330 2	1,8330 2	2,7485 3							7	6,4155
0,8895	0,876	0,8895 1	1,7790 2	7,1160 8	1,7790 2						13	11,5635
0,8625	0,849	0,8625 1	5,1750 6	8,6250 10	6,9000 8	2,5875 3					28	24,1500
0,8355	0,822		3,3420 4	5,8485 7	2,5065 3	4,1775 5	2,5065 3				22	18,3810
0,8085	0,795		0,8085 1	3,2340 4	3,2340 4	3,2340 4	1,6170 2	1,6170 2	0,8085 1	0,8085 1	19	15,3615
0,7815	0,768		0,7815 1	0,7815 1		0,7815 1	4,6890 6	0,7815 1	1,5630 2	0,7815 1	13	10,1595
0,7545	0,741				0,7545 1	2,2635 3	1,5090 2				6	4,5270
а	n_x	8	19	34	18	16	13	3	3	2	116	98,1600
б	$n_x x'$	7,280	25,574	60,588	39,924	42,464	40,170	10,578	11,886	8,796		
в	$n_x (x')^2$	6,625	34,423	107,968	88,551	112,699	124,125	37,298	47,092	38,685	$i_y = 0,027$	
г	$\frac{1}{x_i}$	1,099	0,743	0,561	0,450	0,377	0,322	0,283	0,252	0,227	$i_x = 0,436$	
д	$n_i \frac{1}{x_i}$	8,80	14,12	19,10	8,10	6,00	4,20	0,85	0,76	0,45	62,38	
е	$n_x \left(\frac{1}{x_i}\right)^2$	9,67	10,47	10,70	3,64	2,25	1,35	0,74	0,19	0,10	38,61	
ж	$\sum n_i y_i$	7,386	16,577	29,298	15,174	13,044	10,322	2,399	2,372	1,590	98,16	
з	$\frac{1}{x_i}\sum n_i y_i$	8,1	12,3	16,5	6,8	4,9	3,3	0,7	0,6	0,4	53,6	

Рис. 15. Корреляционная таблица расчета связи между уровнем себестоимости и годовым объемом работ (гиперболический тип зависимости).

Выполним подробный расчет уравнения зависимости y от x в корреляционной таблице (см. табл. 29 на рис. 15).

Приняв по y и по x 9 интервалов, получим два ряда распределения:

$$n_x \to 8,\ 19,\ 34,\ 18,\ 16,\ 13,\ 3,\ 3,\ 2;$$

$$n_y \to 6,\ 13,\ 19,\ 22,\ 28,\ 13,\ 7,\ 6,\ 2.$$

Проверим ряд распределения n_y на соответствие нормальному распределению по критерию согласия А. Н. Колмогорова (см. табл. 30):

Таблица 30

Расчет критерия согласия А. Н. Колмогорова

Частоты m	Накопленные частоты M	Плотность вероятности кривой распределения	Частота кривой распределения m'	Накопленные частоты M'	$\mid M-M'\mid$
1	2	3	4	5	6
6	—	0,075	5	—	—
13	19	0,174	11	16	3
19	38	0,301	19	35	3
22	60	0,390	26	61	1
28	88	0,378	24	85	3
7	108	0,148	9	111	3
6	114	0,060	4	115	1
2	116	0,018	1	116	0
$\Sigma m = 116$	—	1,817	116	—	—

Примечание. Значения графы 4 получаем, умножая значения графы 3 на отношение $\frac{116}{1,817}$, например,

$$0,378\ \frac{116}{1,817} = 0,378 \cdot 64 = \sim 24.$$

Подставим полученное значение в формулу λ (см. стр. 41):

$$\lambda = \frac{3}{\sqrt{116}} = 0,282 < 1,3; \quad p(\lambda) = 1,000.$$

Следовательно, выравненная кривая распределения совпадает с нормальной кривой.

Для того чтобы читатель мог самостоятельно проверить расчет, приводим параметры ряда распределения:

$$\bar{y} = 0,846; \ \sigma_y = 0,050; \ i_y = 0,0269; \ \Sigma n = 116.$$

Для ряда распределения по n_x имеем: $\bar{x} = 2,132$; $\sigma_x = 0,780$; $i_x = 0,436$; $\Sigma n = 116$; рекомендуем проверить неравенство $\lambda = 1,205 < 1,3$.

Строки *б*, *в* и графа *к* корреляционной таблицы заполняются так, как указывалось ранее (см. рис. 9).

Строка *г* содержит значения величин, обратных центрам интервалов:

$$г \to \frac{1}{0,910} = 1,099; \quad \frac{1}{1,346} = 0,743; \quad \frac{1}{1,782} = 0,561;$$

$$\frac{1}{2,228} = 0,450 \text{ и т. д.}$$

В строке *д* значения строки *г* умножаются на n_x: $д \to 1,099 \cdot 8 = 8,8; \; 0,743 \cdot 19 = 14,1; \; 0,561 \cdot 34 = 19,1; \; 0,450 \cdot 18 = 8,1$ и т. д.

В строке *е* значения строки *д* умножаются на значение строки *г*: $е \to 8,8 \cdot 1,099 = 9,67; \; 14,1 \cdot 0,743 = 10,47; \; 19,1 \cdot 0,561 = 10,7; \; 8,1 \cdot 0,45 = 3,64$ и т. д.

В строке *ж* суммируются все произведения количества точек в отдельных клетках корреляционного поля n_i на соответствующие им величины центров интервалов по оси y, т. е. на y_i:

$$ж \to 0,9705 + 2,8305 + 1,8330 + 0,8895 + 0,8625 = 7,386;$$

$$0,9705 + 1,8870 + 1,8330 + 1,7790 + 5,1750 + 3,3420 + 0,8085 + 0,7815 = 16,577;$$

. . . .

$$0,8085 + 0,7815 = 1,59.$$

И наконец, в строке *з* значения строки *ж* умножаются на значения строки *г*:

$$з \to 7,386 \cdot 1,099 = 8,1; \; 16,577 \cdot 0,743 = 12,3;$$

$$29,298 \cdot 0,561 = 16,5;$$

. . . .

$$1,59 \cdot 0,227 = 0,4.$$

Определение ординат эмпирической линии регрессии не требует пояснений.

Контролем правильности расчета является равенство $\Sigma к = \Sigma ж$. Система нормальных уравнений для определения постоянных параметров уравнения гиперболы:

$$\left\{ \begin{aligned} &\Sigma n a + n_i \frac{1}{x_i} b = \Sigma n_i \cdot y_i; \\ &n_i \frac{1}{x_i} a + n_i \left(\frac{1}{x_i}\right)^2 b = \frac{1}{x_i} \Sigma n_i y_i. \end{aligned} \right. \qquad \text{(II.17)}$$

Подставляя в уравнения значения строк табл. 29, получим:

$$\begin{cases} 116a + 62{,}\cancel{56}\,{}^{28}\,b = 98{,}16; \\ 62{,}\cancel{56}\,{}^{28}\,a + 38{,}61b = 53{,}\cancel{80}{,}60 \end{cases}$$

откуда $a = 0{,}754$; $b = 0{,}177$;

$$\bar{y} = 0{,}754 + \frac{0{,}177}{x}.$$

Корреляционное отношение в данном примере равно $\eta = 0{,}662$.

Систему уравнений можно записать и не строя корреляционную таблицу, пользуясь методом наименьших квадратов:

$$\begin{cases} \sum na + \sum \frac{1}{x} b = \sum y; \\ \sum \frac{1}{x} a + \sum \left(\frac{1}{x}\right)^2 b = \frac{1}{x} \sum y. \end{cases} \qquad \text{(II.17a)}$$

В табл. 31 выполнены расчеты $\sum \frac{1}{x}$, $\sum \left(\frac{1}{x}\right)^2$, $\sum \frac{1}{x} y$, для которых значения x, y взяты из табл. 28. Подставив расчетные величины в систему (II.17a), получим:

$$116a + 61{,}407b = 98{,}098;$$
$$61{,}407a + 36{,}809b = 52{,}675,$$

решив систему, найдем:

$$\bar{y} = 0{,}754 + \frac{0{,}173}{x},$$

что весьма близко к уравнению, полученному с помощью корреляционной таблицы[1].

Теперь, когда с помощью корреляционного анализа зависимость уровня себестоимости от объемов работ получила четкую количественную характеристику, вернемся к экономическому анализу. Установленная зависимость указывает характер влияния величины объемов работ на уровень

---

[1] С помощью подстановки $\frac{1}{x} = z$ уравнение гиперболы можно преобразовать в уравнение прямой $x = a + zx$ и, построив ее в новых координатных осях, найти параметры уравнения, а затем выразить z через x.

Таблица 31

Исходные данные для расчета уравнения гиперболы

N	x	$\frac{1}{x}$	$\left(\frac{1}{x}\right)^2$	y	$\frac{1}{x}y$
1	0,95	1,053	1,1088	0,970	1,021
2	0,91	1,099	1,2078	0,935	1,027
3	1,00	1,000	1,000	0,950	0,950
4	1,10	1,909	0,8263	0,940	0,855
5	0,98	1,020	1,0404	0,920	0,939
6	0,99	1,010	1,0201	0,925	0,934
7	1,00	1,000	1,000	0,890	0,890
...	...	...	...	...	...
64	2,30	0,435	0,1892	0,890	0,387
65	2,10	0,476	0,2266	0,845	0,402
66	2,15	0,465	0,2162	0,850	0,395
67	2,20	0,455	0,2070	0,857	0,590
68	2,25	0,444	0,1971	0,860	0,382
69	2,30	0,435	0,1892	0,890	0,387
70	2,35	0,426	0,1815	0,855	0,364
...	...	...	...	...	...
111	2,67	0,375	0,1406	0,760	0,285
112	4,40	0,227	0,0515	0,820	0,186
113	2,80	0,357	0,1274	0,830	0,296
114	4,36	0,229	0,0524	0,790	0,111
115	3,70	0,270	0,0729	0,800	0,216
116	2,54	0,394	0,1552	0,760	0,299

$$\Sigma x = 248,705; \ \Sigma \frac{1}{x} = 61,407; \ \Sigma \left(\frac{1}{x}\right)^2 = 36,809; \ \Sigma y = 98,098;$$

$$\Sigma \frac{1}{x} y = 52,675.$$

себестоимости. При достаточно больших объемах работ — x влияние их почти не отражается на уровне себестоимости — y.

Сравнительно малые x оказывают существенное влияние на y, повышая его. Так, по нашим данным при $x = 0,95$ влияние этого фактора возрастает вдвое по сравнению с влиянием величины x, близкой к среднему (напомним, $\bar{x} = 2,132$ млн. руб). В наибольшей мере снижение объемов работ сказывается на накладных расходах, входящих составной частью в себестоимость строительно-монтажных работ.

Наряду с уровнем себестоимости в экономике строительства большое значение имеет производительность труда. Поэтому изучение влияния факторов на этот показатель всегда представляет интерес. Связь производительности труда с одним из многочисленных факторов, влияющих на нее, рассмотрена ниже.

Зависимость производительности труда от текучести рабочей силы

Экономическая постановка задачи. Ликвидация текучести рабочей силы является важной проблемой.

В последние годы коэффициент оборота рабочей силы в строительных организациях Украинской ССР находился в пределах 1,5—1,8. Это значит, что общее число прибывавших и выбывавших рабочих значительно (в отдельные годы почти вдвое) превышало среднюю списочную численность рабочих. Основными причинами текучести рабочей силы в строительных СМУ, трестах являются недостатки в организации труда и заработной платы, в охране и нормировании труда, необеспеченность рабочих жильем и недостатки в культурно-бытовом обслуживании.

Можно ожидать, что повышение текучести рабочей силы снижает производительность труда. При замене выбывающих рабочих новыми неизбежно уменьшение выработки в связи с оформлением и устройством, ознакомлением с конкретными условиями труда на новом месте. Даже для квалифицированного рабочего необходим определенный период времени на адаптацию, приспособление к изменившимся условиям, в течение которого производительность его труда будет ниже свойственного ему среднего уровня. Кроме того, при пополнении состава трудящихся новыми, часто малоквалифицированными, рабочими производительность труда также снижается (частично это получает выражение при рассмотрении влияния изменения среднего разряда). Косвенным подтверждением этого может служить ряд примеров систематического повышения производительности труда в тех организациях, где имеются стабильные кадры рабочих. Показательны данные по тресту № 1 комбината «Промстрой» (табл. 32). При снижении за 9 лет коэффициента текучести более чем в 4 раза (при одновременном воздействии ряда других существенных факторов) выработка возросла в этом тресте на 70%.

Оценка влияния коэффициента текучести на производительность труда в строительно-монтажных управлениях

Таблица 32

Данные о выработке и текучести рабочей силы

Годы	Выработка на 1 работающего в год		Текучесть рабочей силы (%)		
	тыс. руб	в % к 1958 г.	выбыло по собственному желанию	уволено за нарушение трудовой дисциплины	коэффициент текучести (%)
1958	3,239	100	15	7,1	22,1
1959	3,845	118,6	10	6	16,0
1960	3,787	117	11,2	3	14,2
1961	3,812	118	9,7	2,2	11,9
1962	4,059	125	7,7	1,6	9,3
1963	4,297	132,8	5,1	1,3	6,4
1964	4,802	148,5	6,0	1,0	7,0
1965	5,237	161,5	5,2	0,7	5,9
1966	5,514	170,3	4,5	0,6	5,1

треста № 1 комбината «Промстрой» за 1963—1966 гг. с применением корреляционных методов показывает:

между коэффициентом текучести ($K_{тек}$) и выработкой существует обратная криволинейная зависимость; гипотеза о наличии гиперболической завиеимости типа $y = a + \frac{в}{x}$ подтверждается корреляционным отношением $\eta = 0{,}797$;

зависимость выработки от коэффициента текучести выражается уравнением: $y = 3618 + \frac{14070}{K_{тек}}$, представленным на рис. 16.

Это уравнение показывает, что сокращение текучести на 1% по сравнению со средним для ряда СМУ значением $K_{тек} = 21\%$ позволяет повысить выработку на 0,14%.

Такие уравнения позволяют оценивать резервы роста производительности труда, связанные с созданием более благоприятных условий работы, способствующих снижению текучести.

§ 5. Степенные зависимости. Исследование влияния выработки на уровень себестоимости

В оперативном планировании, учете, экономических исследованиях показатель производительности труда в натуральном выражении применяется только для характеристики отдельных процессов и видов работ.

Скажем, производительность труда штукатуров выражается квадратными метрами оштукатуренных поверхностей в единицу времени (в час, в смену).

В связи с неоднородностью строительной продукции невозможно определять выработку в натуральном измерении не только для всего народного хозяйства, но и для большинства отдельных подрядных строительных организаций.

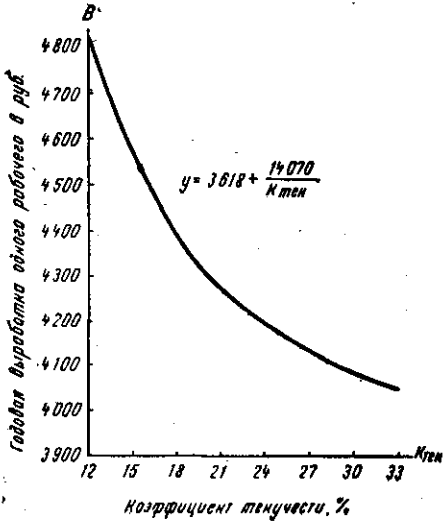


Рис. 16. Влияние сокращения текучести рабочих на рост производительности труда в строительстве.

Поэтому необходимо прибегать к стоимостному ее измерению — выработке.

Исходной базой для определения выработки является сметная стоимость строительно-монтажных работ, 2/3 величины которой зависит от стоимости материалов и конструкций, применяемых при возведении объектов. Логично предположить, что между отношением фактической себестоимости строительно-монтажных работ к сметной и выработкой должна наблюдаться корреляционная связь.

Значение этих переменных, нанесенное на поле корреляции (рис. 17), указывает на возможность описания этой связи уравнением типа

$$y = Q \cdot x^b. \qquad (II.18)$$

Методика вывода корреляционных уравнений при степенной связи иллюстрируется определением количественной оценки зависимости уровня себестоимости (y) от выработки (x) по данным 38 строительно-монтажных управлений комбината «Стальстрой» (1965—1966 гг.); y, x и их логарифмы представлены в табл. 33.

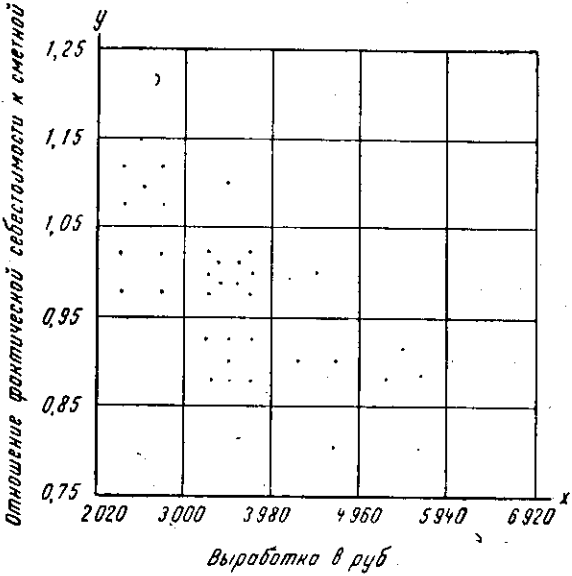


Рис. 17. Влияние выработки на уровень себестоимости (по данным 38 строительных управлений комбината «Стальстрой»; степенной тип зависимости).

Статистические характеристики переменных:

$$\overline{x} = 3712{,}421;\ \sigma_x = 991{,}453;\ \overline{y} = 0{,}706;\ \sigma_y = 0{,}044.$$

Логарифмируем уравнение (II.18):

$$\lg y = \lg Q + b \lg x.$$

Обозначая

$$\lg y = z;\ \lg Q = a;\ \lg x = u,$$

получаем:

$$z = a + b \cdot u,$$

Таблица 33

Исходные данные для определения степенной зависимости

y	x	$\lg y$	$\lg x$	y	x	$\lg y$	$\lg x$
0,960	3 300	—0,01773	3,51851	0,840	4 555	—0,07572	3,65849
1,090	3 001	0,03742	3,47727	0,850	5 855	—0,07058	3,76827
0,950	3 312	—0,02228	3,52009	1,050	2 609	0,02119	3,41647
1,050	2 654	0,02119	3,42390	0,901	3 668	—0,04526	3,56443
0,920	5 552	—0,03621	3,74444	1,00	2 514	0,00000	3,40037
1,150	2 758	0,06069	3,44059	0,920	3 433	—0,03621	3,53567
0,900	4 835	—0,04576	3,68440	1,080	2 701	0,03342	3,43152
0,995	3 200	—0,00218	3,50515	0,800	4 685	—0,09691	3,67071
0,980	3 362	—0,00877	3,52660	0,900	4 977	—0,04576	3,69697
0,970	3 632	—0,01323	3,56015	0,990	3 210	—0,00436	3,50651
1,030	4 145	0,01284	3,61752	0,921	3 528	—0,03574	3,54753
1,020	3 644	0,00860	3,56158	0,900	3 085	0,04576	3,48926
0,910	6 414	—0,04096	3,80713	0,950	3 621	—0,02228	3,55883
0,990	3 315	—0,00436	3,52018	1,120	2 884	0,04922	3,45100
0,880	5 013	—0,05552	3,70010	1,060	2 910	0,02531	3,46389
0,966	3 711	—0,01502	3,56949	0,920	4 809	—0,04096	3,68205
0,940	3 080	—0,02687	3,48856	1,120	2 929	0,04922	3,46672
0,980	3 181	—0,00877	3,50256	0,800	5 266	—0,09601	3,72148
1,040	2 670	0,01703	3,42651	0,981	3 044	—0,00833	3,48344

Составим систему нормальных уравнений:

$$\begin{cases} na + \Sigma ub = \Sigma z \\ \Sigma u \cdot a + \Sigma u^2 b = u\Sigma z. \end{cases} \qquad \text{(II.19)}$$

Величины, необходимые для решения системы нормальных уравнений:

$$z = -0{,}58632; \quad u\Sigma z = -2{,}20165; \quad \Sigma u = 135{,}11764;$$
$$\Sigma u^2 = 480{,}87934.$$

Решив систему (II.19), можем записать уравнение:

$$\overline{y} = 8{,}582 \cdot x^{-0,267}.$$

Рекомендуем читателю проверить правильность полученного уравнения.

§ 6. Другие формы зависимостей

В процессе анализа отчетных данных строительно-монтажных управлений, выполняющих монтаж технологического оборудования (табл. 34), установлено, что форма свя-

Таблица 34

Исходные данные для расчета уравнения параболы третьего порядка (данные в тыс. руб.)

Шифр строительной организации	Объем работ y	Оборотные средства x_1	Производственные основные фонды x_2
1	2	3	4
Промтехмонтаж-2			
СМУ-25—0[1]	3076,2	700,9	267,2
СМУ-25—1	3176,9	677,3	201,5
СМУ-25—2	3136,4	726,5	212,5
СМУ-31—0	3126,4	484,3	167,2
СМУ-31—1	3015,0	633,6	200,8
СМУ-31—2	2338,0	598,0	188,9
СМУ-30—0	3889,5	621,8	128,2
СМУ-30—1	3340,0	797,5	152,0
СМУ-30—2	3042,0	688,9	269,0
СМУ-26—0	1842,5	552,7	169,7
СМУ-26—1	1508,1	484,0	66,7
СМУ-26—2	1623,2	363,9	80,0
СМУ-23—0	5377,9	1134,4	258,6
СМУ-23—1	4450,8	243,0	251,6
СМУ-23—2	3815,0	768,4	255,5
СМУ-22—0	3240,2	560,2	487,1
СМУ-22—1	2320,7	584,2	286,7
СМУ-22—2	2231,4	484,9	181,0
СМУ-13—1	2011,0	195,8	93,3
СМУ-13—2	2274,8	401,9	88,8
Промтехмонтаж-1			
СМУ-1—0	3160,0	685,1	163,5
СМУ-1—1	2093,0	710,2	178,5
СМУ-1—2	2191,0	652,1	191,2
СМУ-4—0	2548,0	479,0	193,6
СМУ-4—1	1905,0	575,3	177,5
СМУ-4—2	1939,0	559,6	171,1
СМУ-17—1	506,0	95,1	49,3
СМУ-17—2	803,8	187,9	50,1
СМУ-14—0	1786,0	324,4	108,5
СМУ-14—1	1682,0	507,9	89,7
СМУ-14—2	1154,0	151,7	87,3
СМУ-18—1	2026,0	351,7	135,5
СМУ-18—2	1271,0	272,2	71,8
СМУ-5—0	1464,2	347,4	81,9
СМУ-5—1	1090,2	270,5	85,9

[1] Цифрами 0, 1, 2 указаны 1960, 1961, 1962 гг.

Продолжение

Шифр строительной организации	Объем работ y	Оборотные средства x_1	Производственные основные фонды x_2
1	2	3	4
СМУ-5—2	1075,4	282,5	86,1
СМУ-29—0	3880,0	568,7	479,9
СМУ-29—1	4257,0	584,2	504,3
СМУ-29—2	3659,0	568,2	377,9
СМУ-210—2	3062,0	708,0	304,9
СМУ-216—2	2634,0	460,0	212,2
СМУ-217—2	4170,0	181,7	148,5
Металлургмонтаж			
СУ-201—2	3810,0	508,7	348,2
СМУ-202—2	2818,0	959,8	388,4
СМУ-209—2	2536,0	452,8	287,3

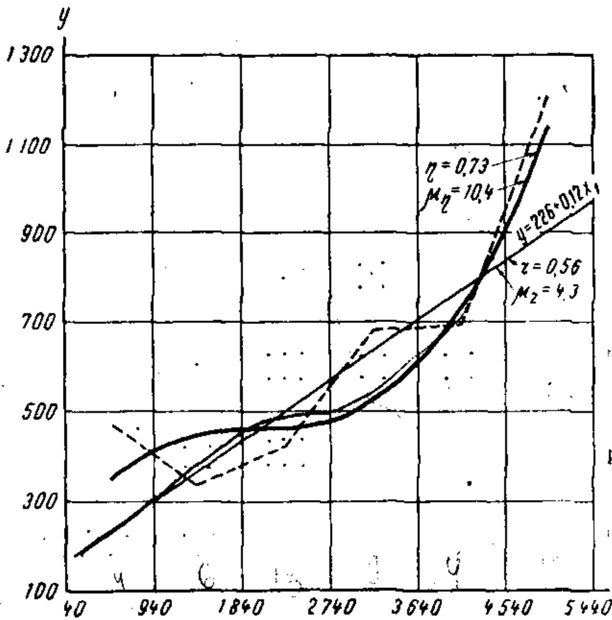


Рис. 18. Связь между объемами работ и оборотными средствами.

зи между объёмами строительно-монтажных работ, оборотными средствами и производственными основными фондами[1] не может с достаточной точностью быть описана одной из рассмотренных выше закономерностей.

Так, теснота связи между объемами работ (y) и оборотными средствами (x_1) при попытке описать эту связь уравнением прямой (рис. 18) характеризуется $r = 0{,}56$. Наличие перегиба в эмпирической линии регрессии указывает

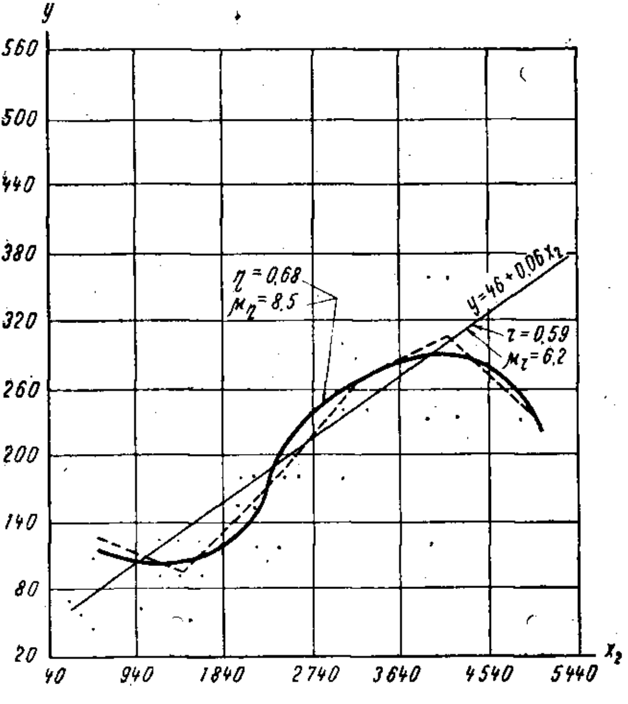


Рис. 19. Связь между объемами работ и размером производственных основных фондов.

на то, что в данном случае более приемлемой будет одна из кривых, уравнение которой выражается многочленом третьей степени (параболой третьего порядка).

---

[1] В строительно-монтажных управлениях типа «Металлургмонтаж» Минспецстроя УССР все строительные машины и механизмы находятся на балансе самих управлений.

Как показал расчет, в данном случае уравнение параболы третьего порядка имеет вид:

$$\bar{y} = 352{,}24 + 7{,}36x_1 - 26\cdot10^{-6}x_1^2 + 82\cdot10^{-10}x_1^3.$$

Отмечаем, что при этом корреляционное отношение $\eta = 0{,}73$ значительно выше приведенного коэффициента корреляции $r = 0{,}56$; критерий надежности $\mu_r = 4{,}3$, а $\mu_\eta = 10{,}4$, и, таким образом, надежность величины η более чем вдвое превышает надежность значения r.

Примерно такое же положение имеет место при установлении формы и тесноты связи между объемами работ (y) и размером производственных основных фондов (x_2). И в этом случае более близкой к реальной оказалась гипотеза не линейной связи, а зависимости, выражаемой параболой третьего порядка (см. рис. 19). Искомое уравнение связи:

$$\bar{y} = 191{,}67 - 0{,}21x_2 + 124\cdot10^{-6}x_2^2 - 16\cdot10^{-9}x_2^3.$$

Решение подобных уравнений вручную требует чрезмерных затрат труда. Поэтому оказалось рациональным передать эту работу на ЭВМ, именно так были получены уравнения, приведенные ранее. Описание и блок-схема программы для определения с помощью ЭВМ формы, тесноты и уравнений парных связей, наиболее часто встречающихся в экономике строительства, приведены в приложении 2.

Глава III

МНОЖЕСТВЕННЫЕ ЛИНЕЙНЫЕ КОРРЕЛЯЦИОННЫЕ ЗАВИСИМОСТИ

§ 1. Зависимость накладных расходов от двух и более факторов

В практике экономической работы одна из важнейших задач экономиста — выбор из множества факторов, формирующих величину показателя (признак), наиболее значительных, совместное влияние которых поддается количественному анализу.

Рассмотрим на примере, как влияет на признак ряд факторов одновременно.

Уравнение, в котором отражается совокупное влияние нескольких факторов на признак, называется уравнением множественной регрессии. Уравнения множественной регрессии могут быть выражены:

в обычном масштабе, т. е. входящие в уравнение переменные величины, имеют нормальную размерность (*т*, млн. руб., тыс. руб. в год на 1 рабочего, %, *квт* и т. д.)

и в сигмальном (нормированном) масштабе, т. е. все переменные величины выражены в сопоставимых единицах.

Применение сигмального масштаба позволяет выявлять и сравнивать влияние факторов, имеющих совершенно различную размерность, а также значительно сокращает и упрощает расчеты.

В дальнейшем изложении будем обозначать в уравнениях множественных зависимостей (в отличие от уравнений парных связей) признак x_1, факторы x_2, x_3, x_4, ..., x_n.

Уравнение множественной регрессии в общем виде можно записать так:

$$x_{1(234,\dots,n)} = b_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n. \qquad \text{(III.1)}$$

Для случая, когда все парные связи линейны, уравнение множественной регрессии в сигмальном масштабе также будет линейным:

$$t_{1(234,\ldots,n)} = \beta_2 t_2 + \beta_3 t_3 + \beta_4 t_4 + \ldots + \beta_n t_n, \quad \text{(III.2)}$$

где $t_{1(234,\ldots,n)}$ — значение признака t в зависимости от совместного влияния факторов 2, 3, 4, ..., n;

$t_2, t_3, t_4, \ldots, t_n$ — значения факторов;

$\beta_2, \beta_3, \beta_4, \ldots, \beta_n$ — коэффициенты, показывающие, как меняется признак под влиянием данного фактора при постоянных значениях остальных факторов.

Одно из замечательных свойств сигмального масштаба состоит в том, что коэффициенты β являются корнями системы уравнений, в которых постоянные члены — парные коэффициенты корреляции r. Применим это свойство при исследовании зависимости накладных расходов от нескольких факторов.

Как уже отмечалось, на накладные расходы (x_1) влияют многие факторы, помимо рассмотренного ранее фактического годового объема строительно-монтажных работ (x_2), выполняемого каждым СМУ собственными силами. Главные из них — производительность труда (отношение x_2 к численности работников, занятых на строительно-монтажных работах и в подсобных производствах, так называемая выработка в тысячах рублей на 1 работающего в год) и удельный вес основной заработной платы рабочих в общих затратах на строительно-монтажные работы.

Показатели x_1 и x_2 получаем непосредственно из отчетности; выработка на 1 работающего в год и удельный вес заработной платы рабочих в общем объеме работ — показатели относительные. Пользование прямыми величинами, получаемыми непосредственно из отчетности, упрощает вычисления. Поэтому целесообразно выработку на 1 работающего и удельный вес заработной платы рабочих заменять прямыми показателями — численностью работников, занятых на строительно-монтажных работах и в подсобных производствах (x_3), и фондом заработной платы рабочих (x_4). На последнем этапе расчета с помощью простейшей замены можно вновь перейти к выработке на 1 работающего и удельному весу заработной платы рабочих в общем объеме работ.

Прежде всего рассмотрим зависимость накладных расходов от двух факторов:

x_2 — сметной стоимости объема строительно-монтажных работ (тыс. руб.);

x_3 — численности работников.

Парный коэффициент корреляции между x_1 и x_2 уже рассчитан:

$r_{12} = 0{,}839$ (см. рис. 9).

Аналогичным способом для выявления связи между x_1 и x_3 можно определить $r_{13} = 0{,}886$. Если необходимо в одном уравнении выразить одновременное влияние x_2 и x_3 на x_1, то после определения r_{13} следует найти также парный коэффициент корреляции между факторами x_2 и x_3. ($r_{23} = 0{,}895$). Рекомендуем читателю самому определить величины r_{13} и r_{23}, построив корреляционные таблицы.

Для учета влияния двух факторов воспользуемся следующим уравнением множественной регрессии:

$$x_{1(23)} = b_1 + b_2 x_2 + b_3 x_3. \tag{III.3}$$

Выразим это же уравнение в сигмальном масштабе:

$$t_{1\,(23)} = \beta_2 t_2 + \beta_3 t_3. \tag{III.4}$$

Коэффициенты β находим из системы уравнений:

$$\begin{cases} \beta_1 + \beta_3 r_{23} = r_{12}; \\ r_{23}\beta_2 + \beta_3 = r_{13}. \end{cases}$$

Подставим в эту систему уравнений $r_{12} = 0{,}839$; $r_{13} = 0{,}886$ и $r_{23} = 0{,}895$.

$$\begin{cases} \beta_2 + \beta_3\, 0{,}895 = 0{,}839; \\ 0{,}895\beta_2 + \beta_3 = 0{,}886. \end{cases}$$

Решив систему, найдем: $\beta_2 = 0{,}230$ и $\beta_3 = 0{,}680$.

Следовательно,

$$t_{1\,(23)} = 0{,}23 t_2 + 0{,}68 t_3. \tag{III.5}$$

Знаки при коэффициентах β показывают направленность действия каждого фактора; в данном случае оба положительных знака при β_2 и β_3 показывают, что с увеличением x_2 и x_3 повышается и x_1. Конечно, в рассматриваемом примере элементарная логика подсказывает, что именно так и должно быть. Но в более сложных случаях это свойство β может оказаться не бесполезным.

Согласно свойству сигмального масштаба коэффициенты β показывают весомость, степень влияния каждого фактора. В данном случае следует вывод, что изменение численности работников влияло на уровень накладных расходов в $\left(\frac{0{,}680}{0{,}230}\right) \cong 3$ раза сильнее, чем изменение объема работ.

Сигмальный масштаб дает возможность наиболее просто определить совокупный коэффициент корреляции R, характеризующий степень тесноты связи между результативным признаком и совокупным влиянием нескольких факторов:

$$R = \sqrt{\beta_2 r_{12} + \beta_3 r_{13} + \beta_4 r_{14} + \ldots + \beta_n r_{1n}}. \qquad \text{(III.6)}$$

В рассматриваемом примере $R = \sqrt{\beta_2 r_{12} + \beta_3 r_{13}} = \sqrt{0{,}23 \cdot 0{,}839 + 0{,}68 \cdot 0{,}886} = \sqrt{0{,}796} = 0{,}89$, т. е. связь довольно тесная.

Для использования уравнения множественной регрессии в планировании, нормировании и некоторых проектных расчетах необходимо перейти от сигмального масштаба его выражения к обычному.

Для этого прежде всего значения переменных в условных координатах умножаем на шаг интервалов (i) и прибавляем действительные значения переменных, соответствующие центрам условных интервалов (C_0). По данным рассматриваемого примера эти значения определяются так (см. рис. 9):

в условных координатах	в действительных координатах
$\bar{x}'_1 = \frac{-7}{74} = -0{,}095;$	$\bar{x}_1 = -0{,}095 \cdot 100 + 250 = 240{,}5;$
$\bar{x}'_2 = \frac{21}{74} = 0{,}284.$	$\bar{x}_2 = 0{,}284 \cdot 500 + 1250 = 1392$[1].

---

[1] В тексте корреляционные таблицы для определения r_{13}, r_{23}, r_{34} не приведены. Укажем только выбранные интервалы для переменных x_3, x_4;

для $x_3 \to 0$ —200—400—600—800—1000—1200—1400; $i_{x_3} = 200$;
для $x_4 \to 0$—200—400—600—800—1000—1200—1400;

$i_{x_4} = 200$; $C_{0x_3} = 700$; $C_{0x_4} = 500$; при самостоятельном построении таблиц читателю рекомендуется выбрать и иные величины C_0 с тем, чтобы убедиться: величина r и действительные значения $\bar{x}$, σ не зависят от выбора C_0.

Затем определяем σ для каждой переменной величины:
в условных координатах в действительных координатах

$$\sigma_1 = \sqrt{\frac{77}{74} - \left(\frac{-7}{74}\right)^2} = 1{,}016; \quad \sigma_1 = 1{,}016 \cdot 100 = 101{,}6;$$

$$\sigma_2' = \sqrt{\frac{125}{74} - \left(\frac{21}{74}\right)^2} = 1{,}269. \quad \sigma_2 = 1{,}269 \cdot 500 = 634.$$

И наконец, находим коэффициенты регресии: b_2, b_3, b_1 из уравнений:

$$b_2 = \beta_2 \frac{\sigma_1}{\sigma_2}; \qquad \text{(III.7)}$$

$$b_3 = \beta_3 \frac{\sigma_1}{\sigma_3}; \qquad \text{(III.8)}$$

$$b_1 = \bar{x}_1 - b_2 \bar{x}_2 - b_3 \bar{x}_3. \qquad \text{(III.9)}$$

Из корреляционной таблицы для определения r_{13} (ее читатель построит сам) имеем: $x_3' = \frac{-35}{74} = -0{,}473$; $x_3 = -0{,}473 \cdot 200 + 700 = 605{,}4$;

$$\sigma_3' = \sqrt{\frac{115}{74} - \left(\frac{-35}{74}\right)^2} = 1{,}152; \; \sigma_3 = 1{,}152 \cdot 200 = 230{,}4.$$

Тогда

$$b_2 = 0{,}23 \cdot \frac{101{,}6}{634} = 0{,}0368;$$

$$b_3 = 0{,}68 \cdot \frac{101{,}6}{230{,}4} = 0{,}2998;$$

$$b_1 = 240{,}5 - 1392 \cdot 0{,}0368 - 605{,}4 \cdot 0{,}2998 = 7{,}775.$$

Таким образом, окончательное уравнение зависимости размера накладных расходов от объема работ и численности работников, т. е. уравнение множественной регрессии в обычном масштабе, будет следующим:

$$\bar{x}_{1(23)} = 7{,}775 + 0{,}0368 \bar{x}_2 + 0{,}2998 \bar{x}_3. \qquad \text{(III.10)}$$

Это уравнение дает возможность нормировать накладные расходы в зависимости от заданных объемов работ и численности работников. Так, для СМУ, выполняющего объем работ в 2000 тыс. руб. и имеющего численность работников 500 чел., норматив накладных расходов на жилищное строительство в Донбассе составит 230—235 тыс. руб.

Помимо этих факторов на размер накладных расходов существенно влияет величина фонда заработной платы рабочих (x_4). Введем в расчет дополнительно фактор (x_4); в этом случае необходимо уже не три, а шесть парных коэффициентов корреляции:

r_{12}, r_{13}, r_{14}—между признаком и каждым фактором;
r_{23}, r_{24}, r_{34}—между факторами.

Из соответствующих расчетных таблиц (исходные данные для этих расчетов имеются в тексте) получаем: $r_{14} = 0{,}838$; $r_{24} = 0{,}837$; $r_{34} = 0{,}869$. Все последующие расчеты аналогичны приведенным ранее.

Построим систему уравнений:

$$\begin{cases} \beta_2 + 0{,}895\,\beta_3 + 0{,}837\,\beta_4 = 0{,}839; \\ 0{,}895\,\beta_2 + \beta_3 + 0{,}869\,\beta_4 = 0{,}886; \\ 0{,}837\,\beta_2 + 0{,}869\,\beta_3 + \beta_4 = 0{,}838. \end{cases} \qquad \text{(III.11)}$$

Поделим каждый член в двух последних уравнениях на коэффициенты при β_2. Получим следующую систему уравнений:

$$\begin{cases} \beta_2 + 0{,}895\,\beta_3 + 0{,}837\,\beta_4 = 0{,}839; \\ \beta_2 + 1{,}117\,\beta_3 + 0{,}971\,\beta_4 = 0{,}990; \\ \beta_2 + 1{,}038\,\beta_3 + 1{,}195\,\beta_4 = 1{,}001. \end{cases}$$

Вычтем первое уравнение из второго и затем из третьего:

$$\begin{cases} 0{,}222\,\beta_3 + 0{,}134\,\beta_4 = 0{,}151; \\ 0{,}143\,\beta_3 + 0{,}358\,\beta_4 = 0{,}162. \end{cases}$$

Снова разделив уравнения на коэффициенты при β_3, получим:

$$\begin{array}{r} -\begin{cases} \beta_3 + 0{,}604\,\beta_4 = 0{,}680 \\ \beta_3 + 2{,}503\,\beta_4 = 1{,}134 \end{cases} \\ \hline -1{,}899\,\beta_4 = -0{,}454, \end{array}$$

откуда $\beta_4 = 0{,}239$.

Подставив значение β_4 в одно из двух последних уравнений, найдем $\beta_3 = 0{,}536$; и затем из уравнения (системы III.11) найдем $\beta_2 = 0{,}159$.

Проверим правильность найденных значений, подставив в одно из уравнений (III.11) значения β_2, β_3 и β_4:

$$0{,}895 \cdot 0{,}159 + 0{,}536 + 0{,}869 \cdot 0{,}239 = 0{,}896.$$

Таким образом, запишем для данного случая уравнение множественной регрессии в сигмальном масштабе:

$$t_{1\,(234)} = 0{,}159\,t_2 + 0{,}536\,t_3 + 0{,}239\,t_4. \qquad \text{(III.12)}$$

Сопоставив уравнения (III.5) и (III.12), обратим внимание на значения β-коэффициентов одноименных факторов. Во втором уравнении они существенно меньше: при t_2 — 0,159 против 0,230 и при t_3 — 0,536 против 0,680. Это объясняется тем, что в уравнение введен дополнительный фактор (x_4), влияние которого привело к уточнению относительных величин влияния факторов x_2 и x_3. Если бы β_2 и β_3 изменились незначительно, это указывало бы на нецелесообразность введения в анализ фактора x_4. В общем же случае с увеличением числа факторов, влияющих на признак, абсолютная величина соответствующих β — коэффициентов уменьшается; в результате снижаются и значения коэффициентов регрессии — b.

Значения σ_1, σ_2, σ_3 нами уже определены. Необходимо найти σ_4; данные для этого получим, если построим корреляционную таблицу зависимости x_1 от x_4:

$$\bar{x}'_4 = \frac{12}{74} = 0{,}162; \quad x_4 = 0{,}162 \cdot 200 + 500 = 532{,}4.$$

$$\sigma'_4 = \sqrt{\frac{94}{74} - 0{,}162^2} = 1{,}115; \quad \sigma_4 = 1{,}115 \cdot 200 = 223.$$

Откуда

$$b_2 = \frac{101{,}6}{634} \cdot 0{,}159 = 0{,}0255; \quad b_3 = \frac{101{,}6}{230{,}4} \cdot 0{,}536 = 0{,}236;$$

$$b_4 = \frac{101{,}6}{223} \cdot 0{,}239 = 0{,}1089;$$

$$b_4 = 240{,}5 - 1392 \cdot 0{,}0255 - 605{,}4 \cdot 0{,}236 - \\ - 532{,}4 \cdot 0{,}1089 = 4{,}152.$$

Следовательно, уравнение множественной регрессии в обычном масштабе будет таким:

$$\bar{x}_{1\,(234)} = 4{,}152 + 0{,}0255\,x_2 + 0{,}236\,x_3 + 0{,}1089\,x_4. \quad \text{(III.13)}$$

Во многих случаях целесообразно уравнение множественной регрессии представлять в виде зависимости относительных (удельных) величин, т. е. абсолютных величин, отнесенных к какой-либо базе. Примем за базу объем работ (x_2) и разделим на x_2 обе части уравнения (III.13).

Получим:

$$x_{1уд}=0{,}0255+\frac{4,152}{x_2}+0{,}236\,\frac{x_3}{x_2}+0{,}1089\,\frac{x_4}{x_2}. \qquad (III.14)$$

Замечаем, что $\frac{x_2}{x_3}=x_5$ представляет собой выработку на одного работающего, а $\frac{x_4}{x_2}=P$ — долю заработной платы рабочих в общей сметной стоимости объема работ.

Подставим замены в уравнение (III.14):

$$x_{1уд}=0{,}0255+\frac{4,152}{x_2}+\frac{0,236}{x_5}+0{,}1089\,P. \qquad (III.15)$$

Это уравнение дает возможность нормировать удельный вес накладных расходов в общем объеме работ с помощью подстановки в него определенных значений объемов работ, выработки на 1 работающего и фонда заработной платы рабочих. Приняв две любые переменные в уравнении (III.15) постоянными (средними), получим зависимость удельных накладных расходов от третьей переменной.

Например, предположив, что x_5 и P константы, определим $x_{1уд}$ в зависимости от x_2;

$$x_{1уд}=0{,}0255+\frac{4,152}{x_2}+0{,}236\cdot\frac{605,4}{1392}+0{,}1089\cdot\frac{532,4}{1392};$$
$$x_{1уд}=0{,}170+\frac{4,152}{x_2}. \qquad (III.16)$$

Подставляя различные значения объемов работ за год в уравнение (III.16), найдем ряд ординат точек и, соединив их вершины плавной кривой, получим гиперболу, наглядно показывающую зависимость $x_{1уд}$ от размера объемов работ (см. рис. 20).

Приведенная методика дает возможность измерить степень влияния объемов работ на уровень накладных расходов.

Влияние на накладные расходы объема работ аналогичным образом можно выявить и для других видов строительства: промышленного, дорожного, монтажных работ и т. д. На рис. 21 представлены кривые, отражающие уровень накладных расходов в зависимости от выработки на одного работника для строительных организаций различной специализации по материалам отчетности ряда строек комбината Шахтострой[1].

При изучении более сложных явлений может возникнуть необходимость учета влияния четырех, пяти и большего числа факторов. Так, при оценке коэффициента индустриализации строительства в расчет могут быть включены: уровень механизации, степень сборности строительства, заводская готовность деталей и конструкций, ритмичность производства работ и др.

При увеличении числа факторов резко возрастает количество таблиц, необходимых для определения парных ко-

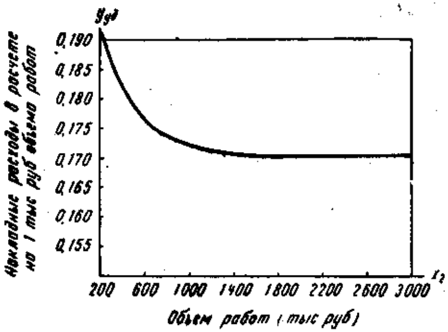


Рис. 20. Зависимость уровня накладных расходов от объема работ.

эффициентов корреляции: при двух факторах нужно рассчитать три таблицы, при трех факторах — шесть, при четырех — 10, при пяти — 14. Для учета влияния на выбранный признак n факторов необходимо найти значения β_n коэффициентов β_2, β_3, ..., β_n.

[1] Идея применения корреляционного анализа к исследованию накладных расходов в шахтном строительстве впервые была предложена доцентом А. В. Заруба в 1954 году.

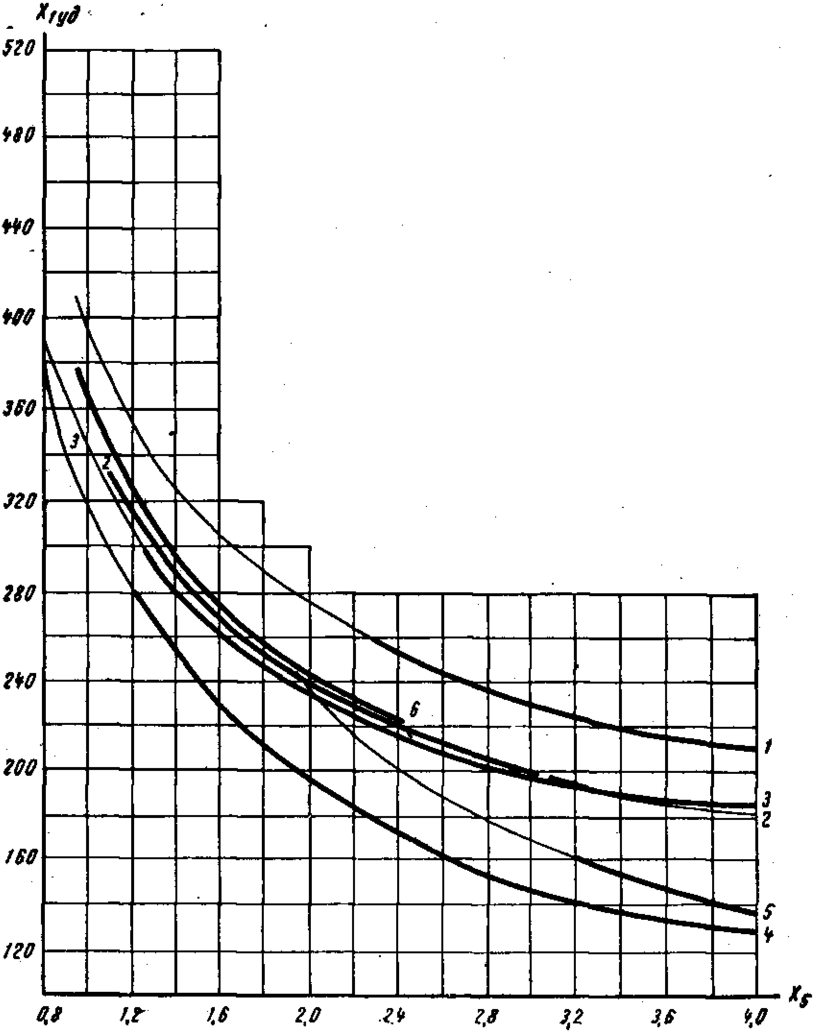


Рис. 21. Зависимость уровня накладных расходов от выработки на одного работающего для различных типов строительств

1. Шахтопроходческие управления. 2. Шахтостроительные управления. 3. Строительство обогатительных и брикетных фабрик. 4. Жилищно-строительные управления. 5. Дорожно-строительные управления. 6. Монтажные управления. Утолщенные линии показывают интервалы фактически встречающихся объемов работ.

Построим систему уравнений для определения β-коэффициентов при n факторах:

$$\left\{\begin{array}{l} r_{12} = \beta_2 + \beta_3 r_{32} + \beta_4 r_{42} + \beta_5 r_{52} + \dots \\ r_{13} = \beta_2 r_{23} + \beta_3 + \beta_4 r_{43} + \beta_5 r_{53} + \dots \\ r_{14} = \beta_2 r_{24} + \beta_3 r_{34} + \beta_4 + \beta_5 r_{54} + \dots \\ r_{15} = \beta_2 r_{25} + \beta_3 r_{35} + \beta_4 r_{45} + \beta_5 + \dots \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ r_{1n} = \beta_2 r_{2n} + \beta_3 r_{3n} + \beta_4 r_{4n} + \beta_5 r_{5n} + \dots \end{array}\right. \qquad (\text{III}.17)$$

Выпишем индексы при r:

$$\left\{\begin{array}{ccccc} 12 & 22 & 32 & 42 & 52 \\ 13 & 23 & 33 & 43 & 53 \\ 14 & 24 & 34 & 44 & 54 \\ 15 & 25 & 35 & 45 & 55 \\ \dots & \dots & \dots & \dots & \dots \\ 1n & 2n & 3n & 4n & 5n \end{array}\right. \qquad (\text{III}.18)$$

Индексы, перечеркнутые диагональю, указывают, что эти коэффициенты корреляции равны единице. Первая цифра индекса r указывает номер столбца, вторая — номер строки минус единица.

Так как $r_{ab} = r_{ba}$, то, например, для системы уравнений с признаком (1) и факторами (2, 3, 4, 5) необходимо вычислить четыре коэффициента корреляции между признаком и каждым фактором и шесть коэффициентов между факторами. Все расчеты уравнений с большим числом переменных отличаются от рассмотренных только объемом вычислений.

Достоверность уравнений множественной регрессии устанавливается проверочной таблицей (см. табл. 35), в которой сравниваются фактические значения (см. табл. 19) с расчетными значениями x_1, полученными в результате подстановки в уравнение (III.13) значений факторов x_2, x_3, x_4 (см. табл. 19).

Таблица 35

№ СМУ	$0{,}0255\ x_2$	$0{,}236\ x_3$	$0{,}1089\ x_4$	$x_{1\,(234)}$	$x_{1\,\text{ф}}$	$\frac{x_{1\,\text{ф}}}{x_{1\,(234)}}$ (%)
1	14,0	72,0	28,8	119,152	100	84
2	22,0	100,0	25,1	151,252	170	112,3
3	23,2	84,7	25,0	137,052	140	102,2
...	...	...	...	...	...	...
72	57,0	157,0	76,1	294,052	290	98,7
73	58,5	182,0	105,7	350,352	320	91,3
74	94,2	315,0	150,0	563,352	570	101,2

Значения $x_{1\,(234)}$ и $x_{1\text{ф}}$ коррелируем, откладывая на горизонтальной оси значения $x_{1\,(234)}$, а на вертикальной — $x_{1\text{ф}}$ (см. рис. 22). Коэффициент корреляции между $x_{1\text{ф}}$ и $x_{1\,(234)}$ $R_\text{ф} = 0{,}893$ будет фактическим совокупным коэффициентом корреляции.

Средняя квадратическая ошибка фактического коэффициента корреляции определяется по формуле:

$$\mu_{0R_\text{ф}} = \frac{1-R_\text{ф}^2}{\sqrt{n-p}}, \qquad \text{(III.19)}$$

где p — число факторов в уравнении множественной регрессии.

Следовательно,

$$\mu_{0R_\text{ф}} = \frac{1-0{,}893^2}{\sqrt{74-3}} = 0{,}0236.$$

$$\text{Критерий надежности } m_{R_\text{ф}} = \frac{0{,}893}{0{,}0236} = 37{,}8.$$

Таким образом, критерий надежности более чем в 12 раз превышает допускаемые пределы, и все наши выводы следует считать вполне достоверными.

В заключение важно отметить значение правильности выбора шага и положения интервалов переменных величин на координатных осях. В корреляционную зависимость между размером накладных расходов (x_1) и сметной стоимостью годового объема строительно-монтажных работ

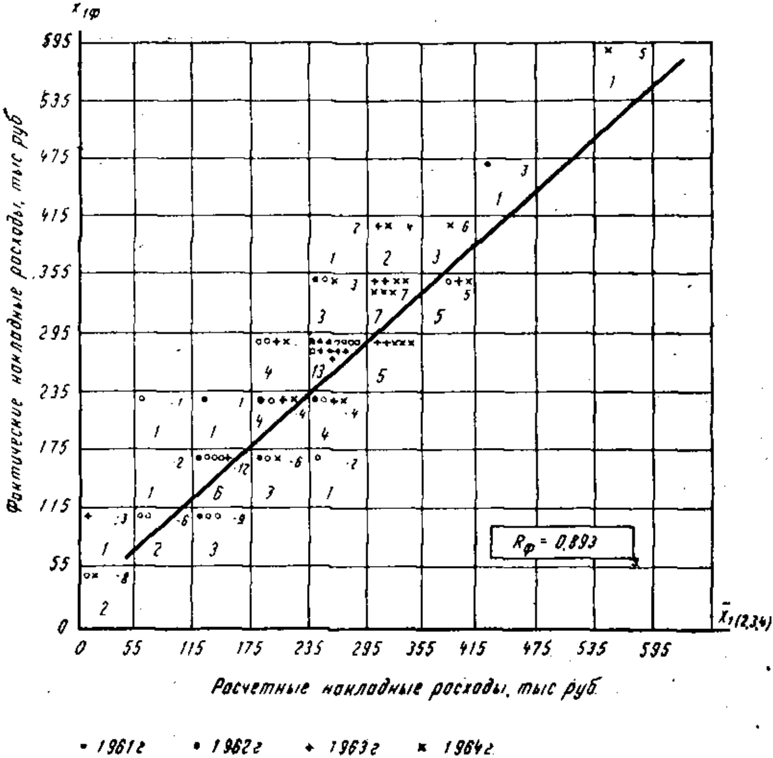


Рис. 22. Корреляция между теоретически исчисленными и фактическими значениями накладных расходов. · — 1961 г., ○ — 1962 г. △ — 1963 г. × — 1964 г.

(x_2) для той же совокупности 74 строительно-монтажных управлений внесем незначительные изменения:

1) шаг переменной по оси абсцисс оставим таким же ($i = 500$), а верхнюю границу интервалов сдвинем влево с 4000 до 3940 тыс. руб.;

2) шаг переменной по оси ординат увеличим со 100 до 104 тыс. руб.

Размещение тех же переменных на новом корреляционном поле (табл. 36 на рис. 23) приводит к изменению средних величин и, следовательно, всех остальных характеристик связи. Критерием целесообразности произведенных изменений является степень приближения средних величин, полученных в результате корреляционного расчета, к фактическим средним.

Таблица 37

Переменные	Фактическая величина x	По корреляционным расчетам		При изменении границ и шага интервалов	
		$\bar{x}$	в % к фактической величине	$\bar{x}$	в % к фактической величине
x_1	249,1	240,5	96,55	248,75	99,85
x_2	1428,1	1392	97,47	1413	98,94
x_3	606,1	605,4	99,88	—	—
x_4	531,8	532,4	100,11	—	—

Интервалы x_3 и x_4 выбраны весьма правильно; отличия x_3 и x_4 от фактических величин незначительны. Выбор же новых интервалов для x_1 и x_2 рационален, ибо x_1 и x_2 стали значительно ближе к фактическим величинам. Опуская промежуточные расчеты, найдем итоговое уравнение множественной регрессии из следующей системы нормальных уравнений:

$$\begin{cases} a_1 n + a_2 \sum x_2 + a_3 \sum x_3 + a_4 \sum x_4 = \sum x_1; \\ a_1 \sum x_2 + a_2 \sum x_2^2 + a_3 \sum x_3 x_2 + a_4 \sum x_4 x_2 = \sum x_2 x_1; \\ a_1 \sum x_3 + a_2 \sum x_2 x_3 + a_3 \sum x_3^2 + a_4 \sum x_4 x_3 = \sum x_3 x_1; \\ a_1 \sum x_4 + a_2 \sum x_2 x_4 + a_3 \sum x_3 x_4 + a_4 \sum x_4^2 = \sum x_4 x_1. \end{cases}$$

Подставив числовые значения величин[1], получим:

$$74\, a_1 + 1056{,}8\, a_2 + 448{,}5\, a_3 + 393{,}6\, a_4 = 184{,}4;$$
$$1096{,}8\, a_1 + 18010{,}4\, a_2 + 7472{,}6\, a_3 + 6601{,}4\, a_4 = 3041{,}9;$$
$$448{,}5\, a_1 + 7472{,}6\, a_2 + 3235{,}6\, a_3 + 2766{,}4\, a_4 = 1282{,}2;$$
$$393{,}6\, a_1 + 6601{,}4\, a_2 + 2766{,}4\, a_3 + 2493{,}3\, a_4 = 1130{,}9.$$

[1] Для упрощения расчетов значения в уравнениях уменьшены в 100 раз.

Таблица 36

корреляционная таблица

	x'		-3	-2	-1	0	1	2	3	4			ЦУИ ←	Итого		
y'	x / y		190	890	1190	1690	2190	2690	3190	3690			ЦРИ ←	и	к	л
		-60	440	940	1440	1940	2440	2940	3440	3940			РИ ←	n_y	$n_y y'$	$n_y y'^2$
ЦУИ ↓	ЦРИ ↓	РИ ↓														
		x_1 624														
3	572	520								1^{3}				1	3	9
2	468	416						1^{2}						1	2	4
1	364	312				8^{8}	9^{9}	1^{1}						18	18	18
0	260	208		3	9	16	2							30	0	0
-1	156	104	2^{-2}	7^{-7}	8^{-8}									17	-17	17
-2	52	0	3^{-6}	4^{-8}										7	-14	28
а	n_x		5	14	17	24	11	2		1				n 74	-8	76
б	$n_x x'$		-15	-28	-17	0	11	4		4				-41		
в	$n_x (x')^2$		45	56	17	0	11	8		16				153		
г	$\sum n_x y'$		-8	-15	-8	8	9	3		3				-8		
д	$x' \sum n_x y'$		24	30	8	0	9	6		12				80	r = 0,885	

$$\sigma_1' = \sqrt{\frac{76}{74} - \left(\frac{-8}{74}\right)^3} = 1{,}007 \qquad \bar{x}_1' = \frac{-8}{74} = -0{,}108$$

$$\sigma_2' = \sqrt{\frac{153}{74} - \left(\frac{-41}{74}\right)^2} = 1{,}327; \qquad \bar{x}_1 = -0{,}108 \cdot 104 + 266 = 248{,}75$$

$$\bar{x}_2' = \frac{-41}{74} = -0{,}554; \qquad x_2 = -0{,}554 \cdot 500 + 1690 = 1413$$

$i_y = 104$
$i_x = 500$
$C_y = 260$
$C_x = 1690$

Рис. 23. Корреляционная таблица расчета связи между накладными расходами и объемом работ при изменении размеров интервалов и шага интервалов.

Решив систему уравнений, определим: $a_1 = 21$; $a_2 = 0{,}00586$; $a_3 = 0{,}2792$; $a_4 = 0{,}09505$.

Таким образом,

$$x_1 = 21 + 0{,}00586\,x_2 + 0{,}2792\,x_3 + 0{,}09505\,x_4.$$

В рассмотренных нами примерах отклонения расчетных показателей от фактических, не превышающих $\pm 10\%$, наблюдались более чем у 70% строительных организаций (см. табл. 35). Это дает основание использовать уравнения множественной регрессии в планировании. Необходимым условием применения этих уравнений должна быть уверенность в том, что выявленные закономерности характерны не для кратковременного отрезка времени, а для длительного периода (нескольких лет).

Чаще всего исходным статистическим материалом являются данные нескольких строительных организаций за ряд лет (обычно 3—5 лет). Но, строго говоря, пользование данными одного СМУ (треста) за три-четыре года взамен данных трех-четырех СМУ довольно условно и несколько нарушает требование взаимной независимости показателей. Поэтому, размещая на корреляционном поле точки (значения), относящиеся к различным годам, рекомендуется отмечать их различными значками, как это сделано на рис. 22. Достаточно равномерное распределение точек каждого года вдоль теоретической линии регрессии дает основание распространить установленные закономерности на ближайшее будущее. Таким образом дополнительно обосновывается возможность использования уравнений множественной регрессии в планировании и прогнозировании.

§ 2. Зависимость производительности труда от трех и более факторов

Экономическая постановка задачи. Технический прогресс в строительстве обеспечивается использованием сборных конструкций и деталей заводского изготовления. Перенесение наиболее трудоемких процессов со строительных площадок на заводы превращает строительное производство в механизированный процесс монтажа зданий и сооружений.

Замена трудоемких укладочных процессов монтажом готовых сборных элементов, укрупнение размеров и повышение заводской готовности элементов приводят к значительному росту производительности труда в строительстве.

Предпосылкой технического прогресса является укрупнение строительно-монтажных организаций. Большие объемы работ позволяют строительным организациям создавать в районах сосредоточенного строительства мощные базы подсобных и вспомогательных производств, иметь механизированное складское хозяйство. Укрупнение строительных организаций создает необходимые условия для рационального использования рабочей силы в соответствии со специальностью и квалификацией работающих, создает возможность применения мощных и эффективных машин, внедрения передовой технологии.

Крупные строительно-монтажные организации позволяют создать нормальный задел в строительном производстве, без чего невозможно равномерное, ритмичное выполнение работ. Преимущества укрупнения строительных организаций способствуют росту производительности труда, совершенствованию использования основных производственных фондов.

Определение влияния применения сборных конструкций, укрупнения строительных организаций и интенсивности фондоотдачи на производительность труда в строительстве является важной экономической задачей.

Рассмотрим влияние этих факторов на рост производительности труда по данным отчетов 159 трестов, осуществлявших строительство главным образом промышленных объектов в УССР (1964—1967 гг.).

Расчет влияния указанных факторов на повышение производительности труда сведен к определению экономии затрат труда на 1 млн. руб. сметной стоимости строительно-монтажных работ.

Для данного примера примем следующие обозначения:

P — средняя годовая численность рабочих;

V — сметная стоимость годового объема строительно-монтажных работ с учетом изменения остатков незавершенного производства (млн. руб.);

$P_{\text{уд}} = \frac{P}{V}$ — удельный расход рабочей силы;

N — среднегодовая стоимость строительных машин и механизмов (млн. руб.);

Z — годовой объем смонтированных сборных железобетонных конструкций (тыс. m);

$M = \frac{V}{N}$ — интенсивность фондоотдачи;

$W = \frac{V}{Z}$ — степень эффективности применения сборных железобетонных конструкций.

Парная регрессия между признаком (P) и факторами (V, N, Z) и между факторами хорошо аппроксимируется прямыми. Определим уравнения парной регрессии и парные коэффициенты корреляции:

$$\begin{cases} P_N = 591,3 + 2055\,N; \quad r_{PN} = 0,872; \\ \qquad\qquad V_N = 3,20 + 8,20\,N; \quad r_{VN} = 0,801; \\ P_Z = 411,4 + 34,4\,Z; \quad r_{PZ} = 0,933; \\ \qquad\qquad V_Z = 2,23 + 0,13\,Z; \quad r_{VZ} = 0,884; \\ P_V = 301,4 + 208,3\,V; \quad r_{PV} = 0,902; \\ \qquad\qquad N_Z = 0,19 + 0,01\,Z; \quad r_{NZ} = 0,807. \end{cases} \qquad \text{(III.20)}$$

Уравнения связи (III.20) отражают следующие закономерности:

1) с возрастанием парка строительных машин на 1 млн. руб. численность рабочих треста увеличивается на 2055 человек;

2) внедрение 1 тыс. m сборных железобетонных конструкций требует увеличения численности рабочих на 32 человека;

3) с приростом объема строительно-монтажных работ на 1 млн. руб. численность рабочих увеличивается на 208 человек.

Уравнение связи для зависимости $V_N = f(N)$ показывает, что возрастание стоимости строительных машин и механизмов на 1 млн. руб. позволяет увеличить объем строительно-монтажных работ на 8,2 млн. руб.

Совокупное влияние V, N, Z на P отражено в уравнении множественной регрессии. В данном случае, т. е. в случае линейной зависимости, общий вид уравнения:

$$P_{VNZ} = b_0 + b_1 V + b_2 N + b_3 Z. \qquad \text{(III.21)}$$

Решение уравнения сводится к определению параметров b_0, b_1, b_2, b_3 при условии, что

$$\sum (P_{\text{факт}} - P_{\text{расч}})^2 = \min.$$

При построении уравнения выбран сигмальный масштаб, при котором, как нам уже известно, за начало отсчета

Для каждой переменной принимается значение средней арифметической ($\bar{x}$), а за единицу измерения — величина среднего квадратического отклонения (σ).

Так как все парные связи линейны, уравнение (III.21) в сигмальном масштабе:

$$t_{P\,(VNZ)} = \beta_V t_V + \beta_N t_N + \beta_Z t_Z. \qquad \text{(III.22)}$$

Значения коэффициентов определены в результате решения системы нормальных уравнений:

$$\begin{cases} \beta_V + 0{,}801\,\beta_N + 0{,}884\,\beta_Z = 0{,}902; \\ 0{,}801\,\beta_V + \beta_N + 0{,}807\,\beta_Z = 0{,}872; \\ 0{,}884\,\beta_V + 0{,}807\,\beta_N + \beta_Z = 0{,}933, \end{cases} \qquad \text{(III.23)}$$

откуда

$$\beta_V = 0{,}236; \quad \beta_N = 0{,}285; \quad \beta_Z = 0{,}495.$$

Как показали расчеты, на численность рабочих (P) наибольшее влияние оказывает объем внедрения сборных железобетонных конструкций; его влияние в 2,1 раза больше, чем объемов выполненных работ, и в 2,7 раза больше, чем влияние увеличения стоимости парка машин.

Подставим в уравнение (III.22) найденные значения:

$$t_{P\,(VNZ)} = 0{,}236\,t_V + 0{,}285\,t_N + 0{,}495\,t_Z.$$

Коэффициенты регрессии определялись из уравнений:

$$b_1 = \frac{\sigma_P}{\sigma_V}\beta_V; \quad b_2 = \frac{\sigma_P}{\sigma_N}\beta_N; \quad b_3 = \frac{\sigma_P}{\sigma_Z}\beta_Z;$$

$$b_0 = \bar{x}_1 - (b_1 x_V + b_2 x_N + b_3 x_Z),$$

откуда

$$b_1 = 54{,}5; \quad b_2 = 671; \quad b_3 = 17{,}2; \quad b_0 = 159{,}24.$$

Таким образом получим следующие уравнения связи в обычном масштабе:

$$P = 159{,}24 + 54{,}5\,V + 671\,N + 17{,}2\,Z. \qquad \text{(III.24)}$$

Последнее уравнение позволяет нормировать среднюю годовую численность рабочих по тресту в зависимости от объемов работ, среднегодовой стоимости строительных машин и механизмов и объема сборных железобетонных конструкций, подлежащих монтажу.

Достоверность уравнения (III.24) можно проверить, сопоставляя расчетные и фактические данные о численности рабочих, на основе следующей таблицы.

Таблица 38

Теоретически исчисленные и фактические данные о численности рабочих

Тресты	Фактические значения показателей			Расчетное число рабочих	Фактическое число рабочих	Отклонения (%)
	V	N	Z			
1	15,23	2,02	98,0	4 030	4 000	—0,75
2	11,52	1,67	43,8	2 661	2 600	—2,30
3	11,82	1,54	57,0	2 818	3 100	10,40
...	...	...	...	...	...	...
157	23,72	2,20	155,0	5 594	5 800	3,60
158	8,54	0,89	56,4	2 192	2 300	4,90
159	5,59	0,70	43,2	1 672	1 600	—4,50

Данные таблицы позволяют сделать вывод, что более чем у 75% всех трестов отклонение численности рабочих, рассчитанной методом корреляции, от фактической равно $\pm 15\%$. Это дает возможность использовать уравнение (III.24) в планировании численности рабочих на строительно-монтажных работах и в подсобных производствах с учетом рассмотренных факторов V, N и Z.

Чтобы перейти к удельным, т. е. отнесенным к 1 млн. руб. объема работ, значениям численности рабочих, делим обе части уравнения (III.24) на V:

$$\frac{P}{V}=\frac{159,24}{V}+\frac{54,5\,V}{V}+\frac{671\,N}{V}+\frac{17,2\,Z}{V}.$$

Затем, разделив числитель и знаменатель третьего члена уравнения на N, а числитель и знаменатель четвертого члена уравнения на Z, мы получаем уравнение множественной регрессии, характеризующее зависимость удельной численности рабочих $P_{уд}$ от интенсивности фондоотдачи — $M=\frac{V}{N}$ и от степени эффективности внедрения сборных железобетонных конструкций $W=\frac{V}{Z}$:

$$P_{уд}=54,5+\frac{159,2}{V}+\frac{671}{M}+\frac{17,2}{W}. \qquad \text{(III.25)}$$

Чтобы выявить степень влияния каждого фактора в уравнении (III.25), предполагаем неизмеными (средними) значения двух других факторов. Так, при средних значениях $V = 11{,}79$ млн. руб., $Z = 72{,}5$ тыс. m, $\overline{W} = \frac{11{,}79}{72{,}5} = 0{,}16$ млн. руб./1 тыс. m получаем уравнение, характеризующее чистое влияние интенсивности фондоотдачи на численность рабочих:

$$P_{\text{уд}} = 173{,}5 + \frac{671}{M}. \qquad \text{(III.26)}$$

Гиперболическая зависимость $P_{\text{уд}}$ от M представлена на рис. 24. Резкое сокращение удельной численности ра-

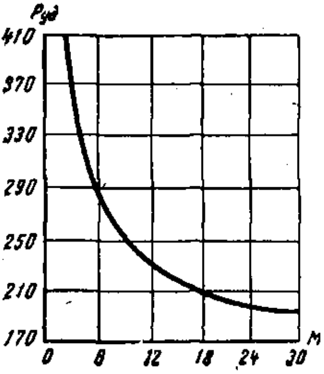


Рис. 24. Влияние интенсивности фондоотдачи на сокращение удельной численности рабочих.

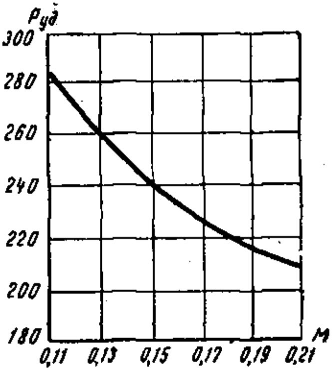


Рис. 25. Влияние степени эффективности внедрения сборных железобетонных конструкций на сокращение удельной численности рабочих.

бочих наблюдается при росте интенсивности фондоотдачи от 3 до 9%. Значительно оно (12%) и при увеличении M от 9 до 15%. Дальнейшее увеличение M ведет к незначительному уменьшению численности рабочих (всего на 5%).

Уравнение гиперболы, показывающее влияние эффективности внедрения сборных железобетонных конструкций на $P_{\text{уд}}$ при средних значениях $N = 1{,}054$ млн. руб., $M = \frac{11{,}79}{1{,}054} = 11{,}186$, имеет вид:

$$P_{\text{уд}} = 127{,}75 + \frac{17{,}2}{W}. \qquad \text{(III.27)}$$

Кривая этого уравнения (см. рис. 25) отражает существенное влияние на удельную численность рабочих эффективности внедрения сборных железобетонных конструкций. При росте W на 10% относительно среднего значения удельная численность рабочих снижается примерно на 4,3%.

Уравнение, связывающее $P_{уд}$ с V:

$$P_{уд} = 219{,}75 + \frac{159{,}2}{V}. \qquad \text{(III.28)}$$

В изучении резервов роста производительности труда в строительстве наряду с элементами индустриализации в настоящее время все большее значение приобретают организационно-технические факторы. Представляет интерес исследование роста производительности труда при одновременном влиянии на него сборности конструкций (как фактора технического прогресса), ритмичности выполнения строительно-монтажных работ, специализации и концентрации (как факторов прогресса в области организации и технологии строительства).

Из перечисленных факторов остановимся лишь на специализации, которая существенно влияет на рост производительности труда, но измерение которой еще не имеет общепринятых методов. В строительстве сложились две формы специализации: предметная, или отраслевая, при которой строительные организации специализируются на возведении объектов для определенных отраслей народного хозяйства, и технологическая, при которой они выполняют отдельные виды работ (работы нулевого цикла, отделочные и др.), являющиеся стадиями технологического процесса.

В специализированных организациях создаются наиболее благоприятные условия для лучшего использования высокопроизводительных машин, повышения квалификации рабочих. Поэтому производительность труда на одинаковых работах в специализированных организациях выше, чем в неспециализированных.

В качестве показателя специализации принимается отношение суммарной сметной стоимости видов работ, выполненных специализированными организациями (субподрядными и собственными подразделениями генподрядной организации), к общей сметной стоимости всего комплекса строительно-монтажных работ, выполненных по генподряду.

Допустим, что сметная стоимость объема работ, выполненных по генподряду, составляет 765 тыс. руб., из них работ, выполненных собственными силами,—4630 тыс. руб. (в том числе по внутреннему субподряду—475 тыс. руб.), сметная стоимость объема работ, выполненных субподрядными организациями, 7650—4630=3020 тыс. руб. Тогда показатель специализации будет равен: $\frac{475+3020}{7650}\cdot 100 = 45{,}7\%$.

Исследование роста производительности труда проводилось по данным 104 строительно-монтажных управлений, сооружающих предприятия машиностроения (см. приложение 1).

Примем следующие условные обозначения:

x_1 — выработка в год на 1 работающего (тыс. руб.);

x_2 — степень ритмичности выполнения работ по кварталам года (%);

x_3 — уровень сборности строительства (%);

x_4 — уровень специализации (%);

x_5 — объем работ, выполняемых своими силами (тыс. руб. в год).

Значения парных коэффициентов корреляции приведены в табл. 39.

Таблица 39

Парные коэффициенты корреляции

	x_1	x_2	x_3	x_4
x_2	0,669			
x_3	0,745	0,174		
x_4	0,459	0,101	0,177	
x_5	0,401	0,013	0,047	0,130

Система нормальных уравнений решается аналогично решению, приведенному на стр. 111:

$$\begin{cases} \beta_2 + 0{,}174\,\beta_3 + 0{,}101\,\beta_4 + 0{,}013\,\beta_5 = 0{,}669; \\ 0{,}174\,\beta_2 + \beta_3 + 0{,}177\,\beta_4 + 0{,}047\,\beta_5 = 0{,}745; \\ 0{,}101\,\beta_2 + 0{,}177\,\beta_3 + \beta_4 + 0{,}130\,\beta_5 = 0{,}459; \\ 0{,}013\,\beta_2 + 0{,}047\,\beta_3 + 0{,}130\,\beta_4 + \beta_5 = 0{,}401; \end{cases}$$

$$\begin{cases} \beta_2 + 0{,}174\,\beta_3 + \ \ 0{,}101\,\beta_4 + \ \ 0{,}013\,\beta_5 = \ \ 0{,}669; \\ \beta_2 + 5{,}747\,\beta_3 + \ \ 1{,}017\,\beta_4 + \ \ 0{,}270\,\beta_5 = \ \ 4{,}282; \\ \beta_2 + 1{,}752\,\beta_3 + \ \ 9{,}901\,\beta_4 + \ \ 1{,}287\,\beta_5 = \ \ 4{,}545; \\ \beta_2 + 3{,}615\,\beta_3 + 10{,}000\,\beta_4 + 76{,}923\,\beta_5 = 30{,}846; \end{cases}$$

$$\begin{cases} 5{,}573\,\beta_3 + 0{,}916\,\beta_4 + \ \ 0{,}257\,\beta_5 = \ \ 3{,}613; \\ 1{,}578\,\beta_3 + 9{,}800\,\beta_4 + \ \ 1{,}274\,\beta_5 = \ \ 3{,}876; \\ 3{,}441\,\beta_3 + 9{,}899\,\beta_4 + 76{,}910\,\beta_5 = 30{,}177. \end{cases}$$

$$\begin{cases} \beta_3 + 0{,}164\,\beta_4 + \ \ 0{,}046\,\beta_5 = 0{,}648; \\ \beta_3 = 6{,}210\,\beta_4 + \ \ 0{,}807\,\beta_5 = 2{,}456; \\ \beta_3 + 2{,}877\,\beta_4 + 22{,}351\,\beta_5 = 8{,}770. \end{cases}$$

$$\begin{cases} 6{,}046\,\beta_4 + \ \ 0{,}761\,\beta_5 = 1{,}808; \\ 2{,}713\,\beta_4 + 22{,}305\,\beta_5 = 8{,}122. \end{cases}$$

$$\begin{cases} \beta_4 + 0{,}126\,\beta_5 = 0{,}299; \\ \beta_4 + 8{,}222\,\beta_5 = 2{,}994. \end{cases}$$

$$8{,}096\,\beta_5 = 2{,}695; \quad \beta_5 = 0{,}333.$$

$$\beta_4 = 0{,}299 + 0{,}126 \cdot 0{,}333 = 0{,}257;$$

$$\beta_3 = 0{,}648 - 0{,}164 \cdot 0{,}257 - 0{,}046 \cdot 0{,}333 = 0{,}590;$$

$$\beta_2 = 0{,}669 - 0{,}174 \cdot 0{,}590 - 0{,}101 \cdot 0{,}257 - \\ - 0{,}013 \cdot 0{,}333 = 0{,}536.$$

Влияние ритмичности на выработку при элиминировании влияния остальных факторов отражено в уравнении:

$$\bar{x}_{12\,(345)} = 0{,}753 + 0{,}056 \cdot x_2.$$

При среднем значении $\bar{x}_2 = 88\%$ выработка составляет: $0{,}753 + 0{,}056 \cdot 88 = \sim 5{,}7$ тыс руб. в год на 1 работающего.

Повышение ритмичности на 1% способно привести к повышению выработки на 56 руб., что свидетельствует о существенном влиянии на производительность труда мер по улучшению организации производства и труда.

Аналогичные расчеты могут быть выполнены и для установления влияния на производительность труда остальных факторов.

§ 3. Зависимость уровня себестоимости от девяти факторов

Экономическая постановка задачи. Фактическая себестоимость строительно-монтажных работ является обобщающим показателем, величина которого зависит от различных факторов. Влияние одних приводит к снижению себестоимости, влияние других способствует ее повышению. Так, рост производительности труда приводит к уменьшению заработной платы рабочих на единицу продукции и, следовательно, к снижению себестоимости; увеличение объема строительно-монтажных работ в единицу времени приводит к снижению накладных расходов, являющихся частью себестоимости. С другой стороны, увеличение текучести рабочих, ведение работ на большом числе одновременно возводимых объектов приводят к повышению себестоимости.

В свою очередь объем строительно-монтажных работ, текучесть рабочих, величина накладных расходов также зависят от различных факторов. Влияние их на себестоимость имеет вероятностный характер и может быть изучено методами корреляции, построением корреляционных моделей.

Рассмотрим моделирование уровня себестоимости по данным 103 строительно-монтажных управлений, специализированных на выполнении санитарно-технических работ по семи трестам Минспецстроя УССР (1966—1967 гг.). В данном случае изменения уровня себестоимости (x_1) с достаточной полнотой отражаются моделью, включающей девять факторов, Большинство факторов, включенных в исследование, получены из существующих форм отчетности:

x_2 — сметная стоимость годового объема строительно-монтажных работ с учетом изменения остатков незавершенного производства (тыс. руб.);

x_3 — среднегодовая выработка на одного работающего (руб.);

x_4 — уровень накладных расходов (отношение фактических затрат по накладным расходам к сметной стоимости выполненных работ);

x_5 — уровень сборности строительства (удельный вес стоимости сборных конструкций и деталей заводского изготовления в общей стоимости материалов);

x_6 — удельный вес фонда заработной платы рабочих, занятых на строительно-монтажных работах и в подсобных производствах, в общем объеме строительно-монтажных работ;

x_7 — коэффициент текучести рабочих;

x_8 — интенсивность фондоотдачи;

x_9 — коэффициент структуры затрат;

x_{10} — показатель сосредоточения строительства.

Последние два фактора (x_9 и x_{10}) не могут быть получены непосредственно из форм отчетности для определения их величины необходимы дополнительные расчеты.

x_9 — коэффициент структуры работ характеризует размер трудовых затрат, приходящихся на единицу объема строительно-монтажных работ.

Введение этого коэффициента вызвано тем, что структура работ даже при сходном профиле специализации в отдельных СМУ в различные годы меняется, что оказывает существенное влияние на уровень себестоимости. Предлагаемый коэффициент структуры работ позволяет количественно оценить изменения комплексов работ. Определение x_9 приведем на примере данных по 103 СМУ санитарно-технологического профиля. Для исчисления его определяем средние взвешенные из затрат труда на каждый вид работ для всей совокупности 103 СМУ (см. табл. 40). Затем на основе исчисленных показателей подсчитываем трудовые затраты на тысячу рублей объема работ.

Средние затраты труда на тысячу рублей сметной стоимости выполненного объема санитарно-технических работ равны $\frac{25619900}{104556} = 245$ чел.-час.

Определим коэффициент структуры санитарно-технических работ (см. табл. 41).

В СМУ-504 средние затраты труда на тысячу рублей выполненного объема работ составили: $\frac{799\,000}{2876} = 277{,}8$ чел.-час.

Следовательно, коэффициент структуры работ СМУ-504 составил:

$$x_9 = \frac{277{,}8}{245} = 1{,}13.$$

Количественная оценка сосредоточения строительства x_{10} является совокупным показателем, в котором учтены расстояние от базы участка (или СМУ) до места строительства объекта (l, *км*), сметная стоимость объек-

Таблица 40

Затраты труда на санитарно-технические работы

Виды санитарно-технических работ	Сметная стоимость выполненного объема санитарно-технических работ (тыс. руб.)	Средневзвешенные затраты труда на тысячу рублей объема отдельных видов санитарно-технических работ (чел.-час)	Затраты на отдельные виды санитарно-технических работ (тыс.-чел. час.)
1	2	3	4
Монтаж отопительных котельных	7 347	160,5	1179,4
Монтаж промышленных котельных	1 680	708,2	1189,9
Монтаж центрального отопления	8 621	238,0	2052,4
Монтаж внутреннего водопровода	5 534	209,6	1159,9
Монтаж внутренней канализации	16 192	209,1	3386,2
Монтаж внутреннего газопровода	7 304	217,0	1584,7
Монтаж внутренних технологических трубопроводов	12 664	216,2	2738,2
Монтаж вентиляции	13 111	265,4	3480,0
» наружного водопровода	2 774	139,8	387,9
Монтаж наружной канализации	1 234	417,9	515,7
Монтаж наружного газопровода	4 944	168,7	834,3
Монтаж наружных технологических трубопроводов	3 873	152,2	589,4
Монтаж тепловых сетей	4 034	181,9	733,9
» металлоконструкций	2 242	192,8	432,3
Монтаж ливневой канализации	243	304,5	74,0
Монтаж сборного железобетона	913	123,7	112,9
Монтаж санитарно-технического оборудования	1 229	463,2	569,3
Монтаж котельных и газораспределительных станций	46	493,4	22,7
Прочие работы	10 571	437,5	4624,8
Всего	104 556		25619,9

Таблица 41

Расчет коэффициента структуры работ СМУ-504

Виды санитарно-технических работ	Сметная стоимость выполненного объема (тыс. руб.)	Принятые трудовые затраты на тысячу рублей объема отдельных видов работ (чел.-час, см. табл. 38)	Трудоемкость, подсчитанная по принятым средним трудовым затратам (тыс. чел.-час.)
1	2	3	4(гр. 2×гр. 3)
Монтаж отопительных котельных	160	160,5	25,7
Монтаж промышленных котельных	227	708,2	160,8
Монтаж центрального отопления	439	238,0	104,5
Монтаж внутреннего водопровода	130	209,6	27,2
Монтаж внутренней канализации	163	209,1	34,1
Монтаж внутреннего газопровода	108	217,0	23,4
Монтаж внутренних технологических трубопроводов	915	216,2	197,8
Монтаж вентиляции . . .	200	265,4	53,1
» наружного водопровода	41	139,8	5,7
Монтаж наружной канализации	22	417,9	9,2
Монтаж наружного газопровода	11	168,7	1,8
Монтаж наружных технологических трубопроводов	16	152,2	2,4
Монтаж тепловых сетей .	160	181,9	29,1
Прочие работы	284	437,5	124,2
Итого . . .	2 876		799,0

та (V, тыс. руб.) и количество объектов, возводимых СМУ (n). Рассмотрим две модификации такого показателя: в одной из них усилено влияние l путем возведения его в квадрат, в другой — влияние V.

Для выбора более приемлемого показателя сосредоточения строительства $\left(C_{\text{рас}} = \frac{\sum nl^2 V}{\sum nl}\right.$ или $\left.C_{\text{об}} = \frac{\sum nV^2 l}{\sum nV}\right)$ построены ряды нормального распределения этих показателей. Нами установлено, что более высокой вероятностью согласованности теоретического и эмпирического распределений обладает показатель сосредоточения строительства $C_{\text{рас}}$, равный 0,61 против $C_{\text{об}}$, равного 0,27.

Определим коэффициент сосредоточения строительства по строительно-монтажному управлению СМУ-544 треста «Сантехмонтаж».

Таблица 42

Расчет коэффициента сосредоточения

Место расположения объектов строительства	Количество объектов n	Объем работ на одном объекте (тыс. руб.) V	Среднее расстояние от объекта до базы участка l, *км*	l^2	nl	nl^2V
Днепропетровск . . .	74	3,2	22	484	1 628	114611,2
Запорожье	4	6,8	85	7 225	340	196520,0
Верхнеднепровск . . .	7	7,1	135	18 225	945	905782,5
Днепродзержинск . . .	75	3,2	29	841	2 175	201840,0
Кривой Рог	34	1,3	81	6 561	2 754	289996,2
Павлоград	16	1,8	93	8 649	1 488	249091,2
Кременчуг	3	4,0	280	78 400	840	940800,0
Новомосковск	4	2,3	35	1 225	140	11270,0
Никополь	20	9,4	76	5 776	1 52	1085888,0
Долинская	1	112,0	127	16 129	127	1806448,0
Ленинский район Кривого Рога	104	2,2	90	8 100	9 360	1853280,0
Октябрьский район Кривого Рога	117	2,8	116	13 456	13 572	4408185,6
Ингулецкий район Кривого Рога	91	2,9	117	13 689	10 647	3612527,1
Итого . . .	550	—	—	—	45 536	15676240,0

$$C_{\text{рас}} = \frac{15\,676\,240}{45\,536} = 344.$$

По такой же методике исчислены коэффициенты сосредоточения всех строительных управлений.

Для вывода уравнения множественной регрессии построим таблицу (см. табл. 43), которая содержит основные статистические характеристики уровня себестоимости (признака x_1) и каждого из девяти факторов, а также значения парных коэффициентов корреляции.

Рекомендуем читателю проверить, пользуясь данными этой таблицы, значения β-коэффициентов и вывести уравнение множественной регрессии вида $x_1 = f(x_2, \ldots, x_{10})$:

$$\begin{aligned} x_1 = 0{,}56763 - 0{,}586 \cdot 10^{-5} x_2 - 0{,}782 \cdot 10^{-5} x_3 + \\ + 0{,}44957 x_4 + 0{,}53383 x_5 + 0{,}00051 x_6 + \\ + 0{,}03251 x_7 - 0{,}00113 x_8 + 0{,}02458 x_9 - 0{,}00155 x_{10}. \end{aligned} \qquad \text{(III.29)}$$

Совокупный коэффициент корреляции $R = 0{,}829$ указывает на достаточную тесную связь между уровнем себестоимости и принятыми факторами.

Знаки при коэффициентах членов уравнения (III.29) не противоречат логике, и это обстоятельство позволяет получать уравнения чистой регрессии, характеризующие зависимость уровня себестоимости от одного фактора при элиминировании влияния всех остальных факторов.

Большие колебания величин среднеквадратических отклонений (σ) при факторах x_8 и x_{10} объясняются тем, что для строительных организаций, выполняющих санитарно-технические работы, характерна нестабильность производства, так как они проводят работы на многих объектах строительства, в большом диапазоне расстояний.

Приводим уравнения чистой регрессии для связи (x_1) с каждым из девяти факторов:

$$x_1 = 0{,}909 - 0{,}586 \cdot 10^{-5} x_2;$$
$$x_1 = 0{,}952 - 0{,}782 \cdot 10^{-3} x_3;$$
$$x_1 = 0{,}843 + 0{,}44957 x_4;$$
$$x_1 = 0{,}574 + 0{,}53383 x_5;$$
$$x_1 = 0{,}889 + 0{,}00051 x_6;$$
$$x_1 = 0{,}888 + 0{,}03251 x_7;$$
$$x_1 = 0{,}918 - 0{,}00113 x_8;$$
$$x_1 = 0{,}872 + 0{,}02458 x_9;$$
$$x_1 = 0{,}899 - 0{,}00155 x_{10}.$$

Уравнения чистой регрессии могут быть использованы для выявления резервов снижения себестоимости.

Так, рост производительности труда на 1% дает возможность строительным управлениям по производству санитарно-технических работ снизить себестоимость на 0,078%. По данным Минспецстроя УССР средний рост производительности труда в СМУ санитарно-технического профиля составляет около 5% в год, следовательно, снижение за счет этого фактора может выразиться в пределах 0,3—0,35% в год, рост величины фондоотдачи на 1% может обеспечить снижение себестоимости на 0,11%.

Каждая строительная организация, составляя перспективный и текущий планы организационно-технических мероприятий, располагает такими данными, как рост основных производственных фондов, изменения структуры работ, рост годовых объемов работ, и др. Пользуясь уравнениями чистой регрессии, можно количественно оценить резервы снижения себестоимости как основного показателя производственно-хозяйственной деятельности строительных организаций, и на этой основе определить уровень прибыли.

Полученная модель себестоимости (III.29) позволяет также определить влияние отдельных факторов на уровень себестоимости в их общей совокупности. Для этого подставляем в уравнение модели средние значения факторов x_2, ..., x_8, полученное среднее значение признака x_1 принимаем за 100% (см. табл. 44).

Средняя ошибка аппроксимации (ε)[1], характеризующая степень соответствия расчетных значений фактическим, составила 1,5% что подтверждает достаточно точное воспроизведение изучаемой совокупности.

Отклонения расчетных значений от фактических в пределах до 1% наблюдались в 44,7% общего числа исследуемых строительных подразделений; отклонения от 1 до 2% — в 29,1%; от 2 до 3% — в 16,5% подразделений. Отклонения более 4% вообще отсутствовали. Таким образом, модель уровня себестоимости характеризуется высокой точностью.

Особенно наглядно преимущество прогнозирующей способности уравнений множественной регрессии в сравнении с обычными методами планирования при сопоставлении отклонений фактического уровня себестоимости от расчетного и планируемого.

Сравним фактический уровень себестоимости с расчетным и планируемым по данным строительных управлений треста «Сантехмонтаж-1» за ряд лет (см. табл. 45).

---

[1] $\varepsilon = \frac{1}{n}\sum\frac{|y-\overline{y}|}{y}\,100\%.$

Таблица 43

Расчет парных коэффициентов корреляции и β-коэффициентов

Факторы	Статистические характеристики переменных		r	Парные коэффициенты корреляции									β-коэффициенты
	$\bar{x}$	σ		r_1	r_2	r_3	r_4	r_5	r_6	r_7	r_8	r_9	
x_1	0,897	0,048	r_1	—									
x_2	2051	631	r_2	−0,35746	—								β_2 −0,07597
x_3	7099	1262	r_3	−0,52277	0,37511	—							β_3 −0,20278
x_4	0,121	0,018	r_4	0,47256	−0,24349	−0,62197	—						β_4 −0,16617
x_5	0,608	0,046	r_5	0,50976	−0,13781	0,03163	−0,05549	—					β_5 0,50423
x_6	15,4	2,35	r_6	0,44795	−0,32066	−0,66082	0,50054	0,01869	—				β_6 0,02410
x_7	0,271	0,148	r_7	0,21384	−0,30020	−0,21518	0,20303	−0,02473	0,04549	—			β_7 0,09876
x_8	18,9	12,18	r_8	−0,45706	0,09464	0,28819	−0,32534	−0,01881	−0,35671	−0,01794	—		β_8 −0,28334
x_9	1,01	0,18	r_9	0,52624	−0,29137	−0,86916	0,66670	−0,01871	0,68186	0,22190	−0,32636	—	β_9 0,08984
x_{10}	250	120	r_{10}	−0,41951	0,21229	0,29697	−0,23441	−0,37733	−0,24976	−0,06218	0,20154	−0,24158	β_{10} 0,02395

Таблица 44

Данные об удельном влиянии факторов

Показатели	Обозначения факторов	Направленность влияния (полож. +, отриц. −)	Удельный вес влияния факторов (%)
Годовой объем работ . . .	x_2	—	5,17
Выработка на одного работающего	x_3	—	13,80
Уровень накладных расходов	x_4	+	11,31
Удельный вес стоимости сборных конструкций и деталей заводского изготовления в общей стоимости материалов	x_5	+	34,32
Удельный вес фонда заработной платы рабочих в себестоимости	x_6	+	1,64
Коэффициент текучести .	x_7	+	6,72
Размер фондоотдачи . . .	x_8	—	19,28
Коэффициент структуры работ	x_9	+	6,11
Коэффициент сосредоточения строительства . . .	x_{10}	—	1,65
Итого . . .			100

Данные табл. 45 показывают, что отклонения между фактическим и плановым уровнем (*б*) намного выше, чем между фактическим и расчетным уровнем себестоимости (*а*).

Это происходит потому, что до сего времени плановые задания составляются без учета объективных возможностей строительных организаций, т. е. без учета влияния факторов, определяющих уровень тех или иных показателей производственно-хозяйственной деятельности строительных организаций. Используемый повсеместно так называемый «базисный метод» планирования основан на том, что строительные организации, работающие лучше, получают более жесткие задания, в то время как отстающим организациям устанавливаются задания на более низком уровне, что не стимулирует «подтягивание» их до уровня передовых. Кроме того, плановые показатели корректируются много раз в течение года.

Таблица 45

Сравнение фактического уровня себестоимости с расчетным и планировавшимся

№ СУ	Отклонения (%)									
	1963 г.		1964 г.		1965 г.		1966 г.		1967 г.	
	а*	б**	а	б	а	б	а	б	а	б
520	2,20	3,32	1,49	8,43	0,22	1,10	1,20	3,10	1,16	2,30
521	0,09	2,90	2,93	4,30	0,22	0,11	1,40	3,70	0,30	1,30
522	0,70	0,70	3,12	6,50	0,11	1,70	0,47	3,80	2,12	2,81
523	0,50	3,31	0,64	1,30	0,79	2,80	1,30	4,10	0,30	0,30
525	2,71	4,31	0,66	1,60	0,78	0,34	1,60	0,36	1,18	1,17
526	3,23	0,35	2,21	1,90	0,31	3,80	1 30	0,36	2,14	2,87
533	2,38	3,50	1,41	2,09	0,23	1,40	1,60	4,30	1 20	1,60
548	—	—	—	—	1,10	5,80	0,95	0,20	0,90	1,90
Средние отклонения	1,68	2,63	1,78	3,71	0,47	2,10	1,22	2,49	1,16	1,78

* отклонения фактического уровня себестоимости от расчетного;
** отклонения фактического уровня себестоимости от планировавшегося.

Исследования показывают, что выявленная закономерность влияния принятых факторов на уровень себестоимости по данным 1963—1967 гг. может быть распространена и на ближайшие годы. Модель достаточно точно воспроизводит экономический процесс, характеризующий образование уровня себестоимости.

В заключение следует сказать, что материалы по моделированию уровня себестоимости [1] были рассмотрены и одобрены техническими советами трестов, главным планово-экономическим управлением Министерства монтажных и специальных строительных работ УССР.

Как видно из приведенного примера, корреляционные модели (или уравнения множественной регрессии) способны учитывать систему многомерных разнонаправленных факторов и представлять результат их действия в достаточно простой форме. Важно отметить, что модели, получаемые с помощью методов корреляции, позволяют рассчитывать

[1] Моделирование уровня себестоимости выполнялось под руководством кандидата технических наук В. А. Михельса.

изменения показателей, характеризующих экономику строительного производства, и таким образом определять необходимые параметры для обеспечения заданных планом результатов. Благодаря этим свойствам корреляционные модели могут быть использованы в планировании — текущем и перспективном. Сравнение фактических данных с результатами, полученными предварительно с помощью моделей, дает возможность судить о качестве прогнозирующей способности корреляционных моделей. Для прогнозирования корреляционные уравнения могут использоваться лишь в тех пределах изменения факторов, для которых они были получены. В этой связи важно при нахождении уравнений регрессии обеспечить достаточное количество значений фактора в пределах выбранного диапазона.

§ 4. О количестве наблюдений, числе факторов и форме связей

Исходными данными для корреляционного анализа экономики строительного производства являются показатели выборочной совокупности. Поэтому коэффициенты корреляции и параметры линий регрессии достоверны с определенной степенью вероятности.

Искажения в расчетах тем меньше, чем больше отношение числа наблюдений (n) к количеству факторов (N), включенных в исследование. Обычно для исследований следует считать достаточным, если

$$\frac{n}{N+1} \geqslant 8.$$

Такое соотношение принято на основании множества исследований экономики строительства с применением корреляционного метода. Именно для этих случаев связи, когда соблюдалось указанное соотношение, корреляционные уравнения обладали наибольшей аппроксимирующей способностью.

В практике исследований для увеличения числа наблюдений данные отчетности какого-либо объекта за год часто рассматриваются как отдельные наблюдения (см. стр. 121). Против такого приема выдвигаются возражения. Главное из них — наличие автокорреляции, т. е. корреляции между показателями одного объекта, взятыми за ряд лет. Автокорреляция действительно приводит к искажению среднеквадратических ошибок коэффициентов корреляции. Но при исследовании процессов в строительстве эти искажения не столь значительны, как в промышленном производстве, в связи с большей разнотипностью продукции и более резкими изменениями структуры работ из года в год.

Нередко в практике экономико-статистического анализа прибегают к исключению в ходе исследования отдельных наблюдений, явно отличающихся от основной их массы в выборочной совокупности. Возражения против исключения отдельных наблюдений основаны на том, что оно производится только по результатам рассмотрения полей парной корреляции. При совокупном же рассмотрении влияния ряда переменных, т. е. в случае множественной корреляции, случайное отклонение от общей закономерности может оказаться причинно обусловленным. Если все-таки исследователь исключает отдельные «пиковые» точки (наблюдения) на данном поле парной корреляции, он должен исключить эти же точки из всех прочих полей корреляции, даже в тех случаях, когда они не нарушают общей закономерности. В этом должна проявиться объективность, являющаяся необходимым качеством любого исследования. Какой же критерий может быть рекомендован для выбора разумных границ в исключении факторов?

В практике исследований в уравнении множественной регрессии чаще всего может быть исключен тот фактор, для которого коэффициент регрессии (*в*) по абсолютной величине меньше его среднеквадратической ошибки:

$$\mu_{\sigma_{b_2, b_3, \ldots, b_n}} > (b_2, b_3, \ldots, b_n), \tag{III.30}$$

где $b_2, b_3, \ldots, b_n$ — параметры.

Для парной регрессии:

$$\mu_{\sigma_b} = \frac{\sigma_y \sqrt{1-r^2}}{\sigma_x \sqrt{n}}. \tag{III.31}$$

Для множественной регрессии;

$$\mu_{\sigma_{b_k}} = \frac{\sigma_1 \sqrt{1-R^2_{1,\,2,\,3,\,\ldots,\,k,\,\ldots,\,n}}}{\sigma_k \sqrt{1-R^2_{k,\,\ldots,\,n}}\sqrt{n}}, \tag{III.32}$$

где

$$R_{1,\,2,\,3,\,\ldots,\,n} = \sqrt{\beta_2 r_{12} + \beta_3 r_{13} + \beta_k r_{1k} + \ldots + \beta_n r_{1n}}.$$

Значения (b) будут достоверными в 95 случаях из 100, если

$$\frac{|b|}{\mu_{\sigma_b}} > 1{,}96.$$

Исключение из анализа тех или иных факторов необходимо и в том случае, если значения коэффициентов корреляции между признаком и каждым фактором и между факторами превышают по абсолютной величине 0,85—0,90, что указывает на тесную линейную связь между ними. Наличие такой линейной связи между двумя переменными факторами называется коллинеарностью и между несколькими факторами — мультиколлинеарностью. Связи при явлениях коллинеарности (мультиколлинеарности) будут неустойчивыми, что находит выражение в больших ошибках среднеквадратических отклонений μ_{σ}.
В этих случаях некоторые ученые рекомендуют исключать из анализа один или несколько коллинеарно связанных факторов (см. К. К. Вальтух [4], Г. В. Розанов, А. А. Френкель [17]). Однако в указанных случаях не всегда прибегают к пересмотру состава факторов (см. Я. И. Лукомский [11], Г. А. Пановский [13]). Влияние коллинеарности на устойчивость связи нами не рассматривалась, и, ссылаясь на опыт М. Езекиела и К. А. Фокса [8], Ф. Миллса [12], мы не используем в настоящей работе исключение факторов в связи с их коллинеарностью. В работе приведены примеры, в которых факторы практически взаимно независимы, что видно из значений парных коэффициентов корреляции: от 0,01 до 0,1 и несколько больше. Наряду с ними приведены примеры, в которых коэффициенты весьма высоки и превышают значения 0,7÷0,8. В расчетах характеристик деятельности строительных организаций чаще всего наблюдаются следующие пределы значений парных коэффициентов корреляции:

для характеристики тесноты связей между признаком и основными факторами, влияющими на него: $r = 0{,}3 \div 0{,}5$ и больше;

для характеристики тесноты связей между факторами, которые взаимно независимы: $r = 0{,}2 \div 0{,}1$ и меньше.

Указанные пределы значений r весьма ориентировочны и значительно отличаются от рекомендуемых другими авторами. Так, В. В. Померанцев [16] отмечает, что коэффициенты парной корреляции характеризуют надежную корреляционную связь при значениях не менее 0,5; Я. И. Лукомский [11] оперирует значениями и менее 0,25.

В качестве приближенного решения вопроса можно использовать соотношение $\frac{r}{\sigma_r} \geqslant t$ ($t=2$ или $t=3$). После необходимых преобразований имеем: $\frac{r\sqrt{n}}{1-r^2} = 2$ (или 3). Отсюда $r=0{,}2$ при $n=100$ и т. д.

Таким образом, расчет тесноты связи обусловливается прежде всего техническими трудностями (значительным объемом вычислительных работ). Определение же формы корреляционных зависимостей невозможно без отчетливого представления о существе изучаемых процессов. В тех случаях, когда, несмотря на кажущуюся линейную связь между двумя переменными, логический анализ сущности экономического явления подсказывает, что более приемлема гипотеза о зависимости, подчиняющейся закону гиперболы, следует исследовать именно последнюю зависимость. Первоначальное суждение на основе визуального рассмотрения поля корреляции есть следствие недостаточного объема выборки, в результате которого часть гиперболы ошибочно принимается за отрезок прямой. Однако всегда следует учитывать реальные диапазоны изменения какой-либо величины.

Критерием правильности выбора той или иной формы зависимости является величина совокупного коэффициента корреляции (R) и соответствие формы кривой экономическим соображениям.

Использование формы зависимости, противоречащей сути рассматриваемого экономического явления, в отдельных случаях может привести к повышению совокупного коэффициента корреляции. Но эти удачи иллюзорны, и к подобному способу «улучшения» результатов анализа не следует прибегать. Если вопрос о форме зависимости между переменными величинами решить трудно, можно пользоваться следующим критерием:

$$k = 0{,}742\sqrt{n}\sqrt{\eta^2 - r^2}, \qquad \text{(III.33)}$$

где η — корреляционное отношение, меньшее из двух возможных $\eta_{x/y}$ и $\eta_{y/x}$;

r — парный коэффициент корреляции.

При $k > 2{,}5$ зависимость следует считать криволинейной.

При статистической обработке на ЭВМ выбор типа кривой происходит согласно программе, подробная блок-схема которой и инструкция к ней приведены в приложении 2.

Глава IV

МЕТОДЫ ПОСТРОЕНИЯ НЕЛИНЕЙНЫХ УРАВНЕНИЙ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

§ 1. Степенная модель

Мы рассмотрели множественные корреляционные зависимости, которые могут быть выражены уравнением прямой и применяются для анализа явлений при совокупном действии множества факторов. Однако в некоторых случаях экономические процессы аппроксимируются более точно *степенными зависимостями*[1].

Исследуем зависимость уровня, себестоимости строительно-монтажных работ (y) от выработки на одного работающего (x_1), коэффициента структуры работ (x_2) и годового объема строительно-монтажных работ, выполняемых своими силами (x_3).

Уравнение множественной регрессии в виде степенной функции:

$$y = a_0 x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}. \qquad \text{(IV.1)}$$

Прологарифмируем правую и левую части уравнения (IV.1):

$$\lg y = \lg a_0 + a_1 \lg x_1 + a_2 \lg x_2 + \dots + a_n \lg x_n.$$

Для рассматриваемого примера введем следующие замены:

$$\lg y = z; \quad \lg a_0 = a_0'; \quad \lg x_1 = u_1; \quad \lg x_2 = u_2; \quad \lg x_3 = u_3$$

---

[1] В корреляционном анализе экономики промышленного производства получил применение оказавшийся весьма плодотворным метод замены степенных моделей логарифмически линейными. Рекомендуем читателю ознакомиться с интересной, по нашему мнению, брошюрой [20].

и подставим их в уравнение (IV.1):

$$z = a_0' + a_1 u_1 + a_2 u_2 + a_3 u_3.$$

Параметры этого уравнения определим с помощью метода наименьших квадратов. В результате получим систему нормальных уравнений:

$$\begin{cases} \sum z = na_0' + a_1 \sum u_1 + a_2 \sum u_2 + a_3 \sum u_3; \\ \sum zu_1 = a_0' \sum u_1 + a_1 \sum u_1^2 + a_2 \sum u_2 u_1 + a_3 \sum u_3 u_1; \\ \sum zu_2 = a_0' \sum u_2 + a_1 \sum u_1 u_2 + a_2 \sum u_2^2 + a_3 \sum u_3 u_2; \\ \sum zu_3 = a_0' \sum u_3 + a_1 \sum u_1 u_3 + a_2 \sum u_2 u_3 + a_3 \sum u_3^2. \end{cases} \quad \text{(IV.2)}$$

Значения переменных, входящих в систему (IV.2), приведены в табл. 46.

Таблица 46

№ п/п	$\lg y$	$\lg x_1$	$\lg x_2$	$\lg x_3$	$\hat{y}$	y
1	2	3	4	5	6	7
1	—0,0309	0,7825	0,0374	0,3242	0,9246	0,933
2	—0,0942	0,8811	0,0170	0,3473	0,8561	0,805
3	—0,0424	0,8119	0,0414	0,2581	0,8916	0,907
4	—0,0297	0,7257	—0,0706	0,1217	0,9433	0,934
5	—0,0390	0,7569	—0,0506	0,1008	0,9185	0,914
...	...	...	...	...	...	...
99	—0,0315	0,8984	—0,0706	0,4047	0,8475	0,930
100	—0,0273	0,7677	0,0000	0,0950	0,9114	0,939
101	—0,0362	0,7892	0,0828	0,1183	0,9004	0,920
102	—0,0590	0,8300	0,0755	0,2238	0,8813	0,873
103	—0,0491	0,8035	—0,0969	0,3896	0,9120	0,893

Находим корни системы уравнений[1]:

$a_0' = 0{,}2221$ (откуда, потенцируя, находим $a_0 = 1{,}667$); $a_1 = -0{,}3475$; $a_2 = 0{,}0131$; $a_3 = -0{,}0476$.

Подставив значения корней в уравнение (IV.1), получим модель уровня себестоимости:

$$y = 1{,}667 x_1^{-0{,}3475} \cdot x_2^{0{,}0131} \cdot x_3^{-0{,}0476}. \quad \text{(IV.3)}$$

[1] В тексте все вычисления для упрощения расчетов выполнены с точностью до четырех знаков.

Проверим расчет, подставив значение логарифмов переменных y, x_1, x_2, x_3 для любого из 103 строительно-монтажных управлений в уравнение (IV.1). Например, для СУ-1:

$$\lg \hat{y} = 0{,}2221 - 0{,}3475 \cdot 0{,}7825 + 0{,}0131 \cdot 0{,}0374 - \\ - 0{,}0476 \cdot 0{,}3242 = -0{,}0309.$$

$\hat{y} = 0{,}9246$ отличается от $y = 0{,}933$ на 1%.

$\sigma^2_{\text{ост}}$ (остаточная дисперсия) $= 0{,}000882$;
$\eta_{\text{м}}$ (множественное корреляционное отношение) $= 0{,}91$;
ε (средняя ошибка аппроксимации) $= 2{,}7\%$.

Вычисления параметров уравнения множественной регрессии, как видим, довольно громоздки. Для выполнения их на ЭВМ необходимо представить систему нормальных уравнений (IV.2) в матричной записи:

$$(X^* X) A = X^* У,$$

где А — матрица-столбец искомых коэффициентов;
Х — матрица значений отобранных факторов для всей совокупности объектов (в нашем примере число их равно 103);
У — матрица-столбец фактических значений признака;
Х* — матрица, транспонированная к матрице Х:

$$A = (X^* X)^{-1} (X^* У).$$

Программа разработки уравнения множественной регрессии, представленного в виде степенной функции, состоит из двух частей: первая вводит числовой материал, формирует константы, определяет данные, необходимые для расчета парных связей ($\bar{y}$, x_i, σ_y, σ_{x_i}, a_0, η), и записывает исходный материал на магнитный барабан машин; вторая выдает множественное уравнение регрессии.

В программе используются шесть стандартных подпрограмм: перевода $10 \to 2$; перевода $2 \to 10$ с выдачей результата на печать; определения $\ln n$; обращения матрицы; умножения матрицы на вектор; определения y^x.

Числовой материал записывается следующим образом:

№ ячейки	Содержание
0011	Число точек
0022	Число факторов
1700	Массив для y, x_i

Программа начинается с 6000-й ячейки и занимает 640 ячеек с формированием и константами. Расчет возможен для $n = 110$, $k = 20$; после ввода нужно сделать пуск с 6000-й ячейки. На печать выдаются:

Характеристики парных уравнений связей $y'' = a_0 x_i^{a_1}$:

$$y, y_i, a_0, a_1, \sigma_y, \sigma_{x_i}.$$

Коэффициенты регрессии степенной функции;

$$a_0, a_1, a_2, \ldots a_n.$$

Остаточная дисперсия $\sigma^2_{\text{ост}}$

$$\eta_{\text{м}}, \mu_{\eta_{\text{м}}}, t_{\eta_{\text{м}}}.$$

Средняя ошибка аппроксимации ε (показатель адекватности расчетных значений фактическим).

Блок-схема этой программы представлена на рис. 26.

§ 2. Экономико-статистическое моделирование различных форм связи

Исследования сложных процессов в экономике строительства невозможны без использования уравнений множественной регрессии высших порядков. Если связи между признаком и факторами выражены различными формами зависимостей (линейными и криволинейными), а спрямление нелинейных зависимостей невозможно или нецелесообразно, может быть рекомендован так называемый «метод Брандона» [3]. Этот метод называется также графическим, поскольку вид зависимости может быть установлен с помощью графиков функций (полей корреляции).

Применение метода проиллюстрируем, используя исходные данные, которые были приведены в § 1. Однако в данном случае y с x имеет прямолинейную связь, а с x_2 и x_3 — параболическую.

Представим y в виде произведения функций, каждая из которых зависит только от одного из факторов:

$$y = c f_1(x_1) f_2(x_2) f_3(x_3), \tag{IV.4}$$

где c — постоянная величина, соответствующая $\bar{y}$.

Значения c и функции $f_i(x_i)$ определяются в следующей последовательности:

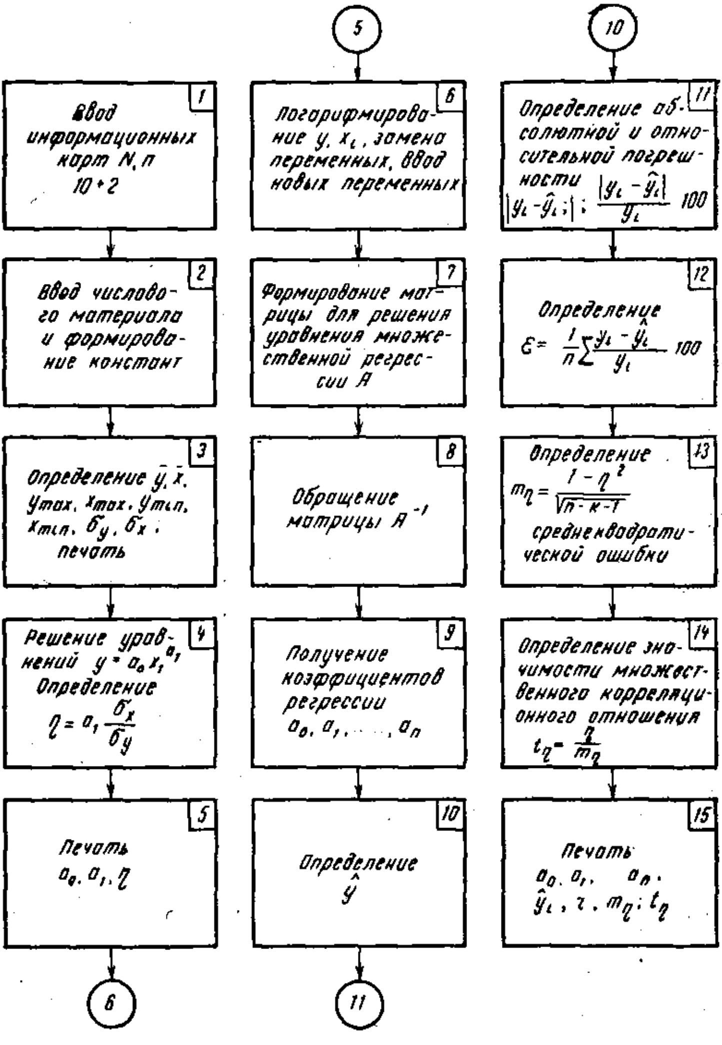


Рис. 26. Блок-схема вывода уравнения множественной регрессии для степенной модели.

заменяем y нормализованным его значением:

$$y_0 = \frac{y}{\bar{y}};$$

исчисляем условные показатели:

$$y_1 = \frac{y_0}{f_1(x_1)} = \frac{y_0}{cf_1(x_1)} = f_2(x_2);$$

$$y_2 = \frac{y_1}{f_2(x_2)} = \frac{y}{cf_1(x_1) f_2(x_2)} = f_3(x_3);$$

находим корреляционные зависимости:

$$y_0 \text{ от } x_1;\ y_1 \text{ от } x_2;\ y_2 \text{ от } x_3;$$

получаем уравнения регрессий:

$$f_1(x_1), \quad f_2(x_2), \quad f_3(x_3).$$

Каждая последующая зависимость условного показателя характеризует остаточный результат, т. е. результат, зависящий не от данного, а от последующих факторов.

Исходные и расчетные данные для выполнения расчета по графическому методу приведены в табл. 47.

Таблица 47

Расчет уравнения регрессии по методу Брандона

№ СУ	$y_{факт}$	x_1	$y_0 = \frac{y}{\bar{y}}$	$f_1(x_1)$	$y_1 = \frac{y_0}{f_1(x_1)}$	x_2	$f_2(x_2)$	$y_2 = \frac{y_1}{f_2(x_2)}$	x_3	$f_3(x_3)$
1	2	3	4	5	6	7	8	9	10	11
1	0,933	6 061	1,04	1,012	1,027	1,09	0,948	1,083	2 109,7	1,000
2	0,805	7 605	0,89	0,963	0,924	1,04	0,938	0,985	2 225,0	1,018
3	0,907	6 591,4	1,01	1,095	0,923	1,10	0,938	0,984	1 811,8	0,985
4	0,934	5 317	1,04	1,036	1,00	0,85	0,950	1,052	1 321,8	1,059
5	0,914	5 713	1,02	1,023	0,997	0,89	0,941	1,059	1 261,3	1,079
6	0,844	6 598	0,95	1,095	0,867	0,84	0,953	0,909	1 706,7	0,990
7	0,928	5 866	1,03	1,018	1 012	0,76	0,984	1,028	1 962,4	0,987
...	...	...	...	...	...	...	...	...	...	...
99	0,930	7 914	0,92	0,956	0,962	0,85	0,950	1,148	2 539,0	1,100
100	0,939	5 858	1,05	1,019	1,030	1,00	0,934	1,103	1 244,6	1,081
101	0,920	6 154	1,03	1,009	1,021	1,21	0,997	0,024	1 313,1	1,061
102	0,873	6 762	0,97	0,996	0,974	1,19	0,985	0,989	1 674,0	0,993
103	0,893	6 361	1,00	1,002	0,998	0,80	0,968	1,031	2 452,4	1,073
	$\bar{y} = 0,897$									

Для СУ № 1 $y_0 = \frac{0,933}{0,897} = 1,04$; для СУ № 7 $y_0 = \frac{0,928}{0,897} = 1,03$ и т. д.

Размещение y_0 и x_1 на корреляционном поле показывает, что зависимость между ними может быть аппроксимирована уравнением прямой:

$$f_1(x_1) = 1,206 - 0,000032\, x_1. \qquad \text{(IV.5)}$$

Подставив в уравнение (IV.5) реальные величины $(y, \bar{y})$, запишем в гр. 5 табл. 47 значения $f_1(x_1)$, затем исчислим y_1 для всех строительно-монтажных управлений

$$\text{СУ № 1 } y_1 = \frac{1,04}{1,012} = 1,027;$$

$$\text{СУ № 99 } y_1 = \frac{0,92}{0,956} = 0,962$$

и т. д.

Определив все значения y_1, впишем их в гр. 6.

Размещаем y_1 и реальные значения x_2 на корреляционном поле. Влияние коэффициента структуры работ (x_2) на остаточное значение уровня себестоимости (y_1) может быть аппроксимировано уравнением параболы второго порядка.

$$f_2(x_2) = 2,0125 - 2,219\, x_2 + 1,14\, x_2^2. \qquad \text{(IV.6)}$$

Получаемые значения после подстановки в уравнение (IV.6) величин x_2 внесем в гр. 8. Так, для СУ № 1

$$f_2(x_2) = 2,0125 - 2,219 \cdot 1,09 + 1,14 \cdot 1,09^2 = 0,948$$

и т. д.

Найдем условный показатель y_2, который уже будет зависеть только от x_3 (но не от x_2 и x_1):

$$\text{для СУ № 1 } y_2 = \frac{y_1}{f_2(x_2)} = \frac{1,027}{0,948} = 1,083;$$

$$\text{для СУ № 5 } y_2 = \frac{0,997}{0,941} = 1,059$$

и т. д.

Результаты отражаются в гр. 9.

В общем случае расчет продолжается, пока не определены все функции $f_i(x_i)$, в данном примере следующий этап является последним.

Аналитическое значение $f_3(x_3)$ определяем по виду корреляционной зависимости y_2 от x_3 которая выразилась уравнением:

$$f_3(x_3)=1{,}865-0{,}00095x_3+0{,}254\cdot10^{-6}\,x_3^2. \qquad \text{(IV.7)}$$

Подставляя реальные значения x_3 (из гр. 10) в уравнение (IV.7), получаем величины $f_3(x_3)$ и проставляем в гр. 11 табл. 47.

Так, для СУ № 1

$$f_3(x_3)=1{,}865-0{,}00095\cdot2109{,}7+0{,}254\cdot10^{-6}\cdot2109{,}7^2=1{,}000$$

и т. д.

Таким образом, модель уровня себестоимости в зависимости от факторов x_1, x_2, x_3 выразится уравнением:

$$y=0{,}897\,(1{,}206-0{,}000032x_1)\cdot(2{,}0125-2{,}219x_2+1{,}14x_2^2)\times$$
$$\times\left(1{,}865-0{,}00095x_3+0{,}254\cdot10^{-6}\,x_3^2\right). \qquad \text{(IV.8)}$$

Сопоставление всех фактических значений y с исчислен ными по уравнению (IV.8) показывает, что отклонения не превышают $\pm10\%$.

Для определения общей тесноты связи между признаком y и факторами x_1, x_2, x_3 найдем множественное корреляционное отношение:

$$\eta_{\text{м}}=\sqrt{\frac{\sigma^2_{x_1\,x_2\,x_3}}{\sigma^2_y}}, \qquad \text{(IV.9)}$$

где σ_y^2 — общая дисперсия y;

$\sigma^2_{x_1\,x_2\,x_3}=\dfrac{\sum(\tilde{y}-\tilde{\tilde{y}})^2}{n-1}$ — дисперсия теоретической поверхности регрессии, характеризующая влияние на y только отобранных факторов x_1, x_2, x_3;

$\tilde{y}$ — значение уровня себестоимости, рассчитанное для каждого строительно-монтажного управления по модели;

$\tilde{\tilde{y}}$ — среднее значение расчетного показателя уровня себестоимости; $\tilde{\tilde{y}}=0{,}8974$ (значение $\tilde{\tilde{y}}$ весьма мало отличается от $\bar{y}=0{,}897$, что свидетельствует о высокой точности расчета);

n — число строительно-монтажных управлений.

Для определени $\sigma^2_{x_1 x_2 x_3}$ необходимо рассчитать отклонения y от $\bar{y}$ и $\tilde{y}$ (см. табл. 48).

Таблица 48

Сравнение фактических значений y с расчетными

№ CУ	Фактические значения y	$(y-\bar{y})$	$(y-\bar{y})^2$	Расчетные значения $\tilde{y}$	$\tilde{y}-\bar{\tilde{y}}$	$(\tilde{y}-\bar{\tilde{y}})^2$
1	1	3	4	5	6	7
1	0,933	0,036	0,001300	0,861	−0,036	0,001300
2	0,805	−0,092	0,008460	0,825	−0,072	0,005180
3	0,907	0,010	0,000100	0,907	0,010	0,000100
4	0,934	0,037	0,001370	0,935	0,038	0,001440
5	0,914	0,017	0,000289	0,942	0,045	0,002030
6	0,844	−0,053	0,002810	0,927	0,030	0,000900
7	0,928	0,031	0,000961	0,887	−0,040	0,000100
...	...	...	...	...	...	...
97	0,918	0,021	0,000441	0,855	−0,042	0,001760
98	0,821	−0,076	0,005780	0,844	−0,058	0,002810
99	0,930	0,033	0,001090	0,893	−0,004	0,000016
100	0,939	0,042	0,001760	0,923	0,026	0,000676
101	0,920	0,023	0,000529	0,957	0,060	0,003600
102	0,873	−0,024	0,000576	0,874	−0,023	0,000529
103	0,893	−0,004	0,000016	0,934	0,037	0,001370
Итого $\Sigma\bar{y}=92,391$			$\Sigma(y-\bar{y})^2= =0,18746;$	$\Sigma\tilde{y}= =92,432$		$\Sigma(\tilde{y}-\bar{\tilde{y}})^2= =0,17136$

$$\bar{y}=\frac{92,301}{103}=0,8970;\quad \sigma^2_y=\frac{0,18746}{103}=0,00182;$$

$$\bar{\tilde{y}}=\frac{92,432}{103}=0,8974;\quad \sigma^2_{x_1 x_2 x_3}=\frac{0,17136}{103-1}=0,00168;$$

$$\eta_м=\sqrt{\frac{0,00168}{0,00182}}=0,9606.$$

Значимость множественного корреляционного отношения может быть проверена по критерию t (критерий Стьюдента):

$$t=\frac{\eta}{\mu_\eta}, \tag{IV.10}$$

где μ_η — ошибка множественного корреляционного отношения, определяемая по формуле

$$\mu_\eta=\frac{1-\eta^2}{\sqrt{n-k-1}}; \tag{IV.11}$$

k — количество факторов, в нашем случае $k=3$; Находим:

$$\mu_\eta = \frac{1-0{,}923}{\sqrt{103-3-1}} = 0{,}0077.$$

Отсюда

$$t = \frac{0{,}9606}{0{,}0077} = 124{,}1,$$

что во много раз превышает табличное значение $t = 2{,}58$, указанное для уровня значимости 0,01. Следовательно, показатель достаточно значим.

С увеличением в модели количества факторов, имеющих как линейные, так и нелинейные парные связи с признаком, точность рассмотренного метода возрастает.

Читатель, возможно, обратит внимание на то, что на основании одних и тех же исходных данных делаются якобы разноречивые выводы: в одном случае связи уровня себестоимости с коэффициентом структуры работ и объемом строительно-монтажных работ принимаются линейными (гл. III, § 4), а в другом случае — нелинейными (гл. IV, § 1). Это противоречие кажущееся, так как в первом случае рассматриваются связи факторов с натуральным признаком, во втором — признаком являются условные показатели y_0, y_1, и y_2, выражающие остаточное значение признака, формирующиеся под влиянием последовательно рассматриваемых факторов.

Аппроксимировать уровень себестоимости в зависимости от включенных в исследование факторов (x_1, x_2, x_3) следует по уравнению множественной регрессии; только это конечное уравнение имеет реальный экономический смысл.

Выбор форм связей в подобных исследованиях может быть передан ЭВМ. Блок схема вывода уравнения множественной регрессии при использовании графического метода приведена на рис. 27.

§ 3. Применение многошагового регрессионного анализа

Метод корреляционного анализа позволяет количественно оценивать связи между величинами при условии действия большого числа факторов, ряд из которых неизвестны. Его применение дает возможность проверить различные гипотезы о наличии и тесноте связи между некоторыми переменными или группами переменных, а также гипотезы о форме связи. Корреляционный анализ помогает установить, как в среднем изменяется случайная величина[1] с изменением одной или нескольких других случайных величин при фиксированном (среднем) значении неучтенных факторов. Однако не все факторы — случайные величины. Поэтому в экономических исследованиях часто приходится рассматривать связи между случайными и неслучайными величинами. Такие связи называются регрессионными. Регрессионный анализ представляет собой метод определения степени влияния каждого фактора (случайной величины) на изучаемый признак (неслучайную величину) и оценки степени этого влияния с помощью различных критериев.

При многофакторном регрессионном анализе совокупность факторов, характеризующих экономические явления, состоит как из факторов, являющихся случайными величинами (использование рабочего времени, текучесть рабочей силы и т. д.), так и из факторов, являющихся неслучайными величинами (размер производства, фондовооруженность, специализация и т. п.). Такое деление условно и справедливо только в рамках построения моделей.

Оба метода — корреляционный и регрессионный — тесно связаны; регрессионный метод в известной мере является продолжением корреляционного, дополняя его главным образом анализом и интерпретацией коэффициентов множественной регресии.

Многофакторные экономические модели показывают, *как* каждый из факторов *влияет* на признак и как *изменяется* последний при изменении каждого фактора на единицу измерения его величины. Количественная оценка влияния, его удельный вес в совокупном влиянии всех факторов, выделение из ряда факторов наиболее существенных и выбор наиболее удачных форм аппроксимации — эти задачи решаются совместно корреляционным и регрессионным анализом.

Суть многошагового регрессионного анализа состоит в том, чтобы последовательными приближениями (шагами) найти такой полином (уравнение множественной регрессии),

[1] Случайной величиной называется величина, принимающая в результате испытаний или в процессе наблюдений числовое значение, которое принципиально нельзя предсказать, исходя из условий испытания или наблюдений: любая другая величина будет неслучайной.

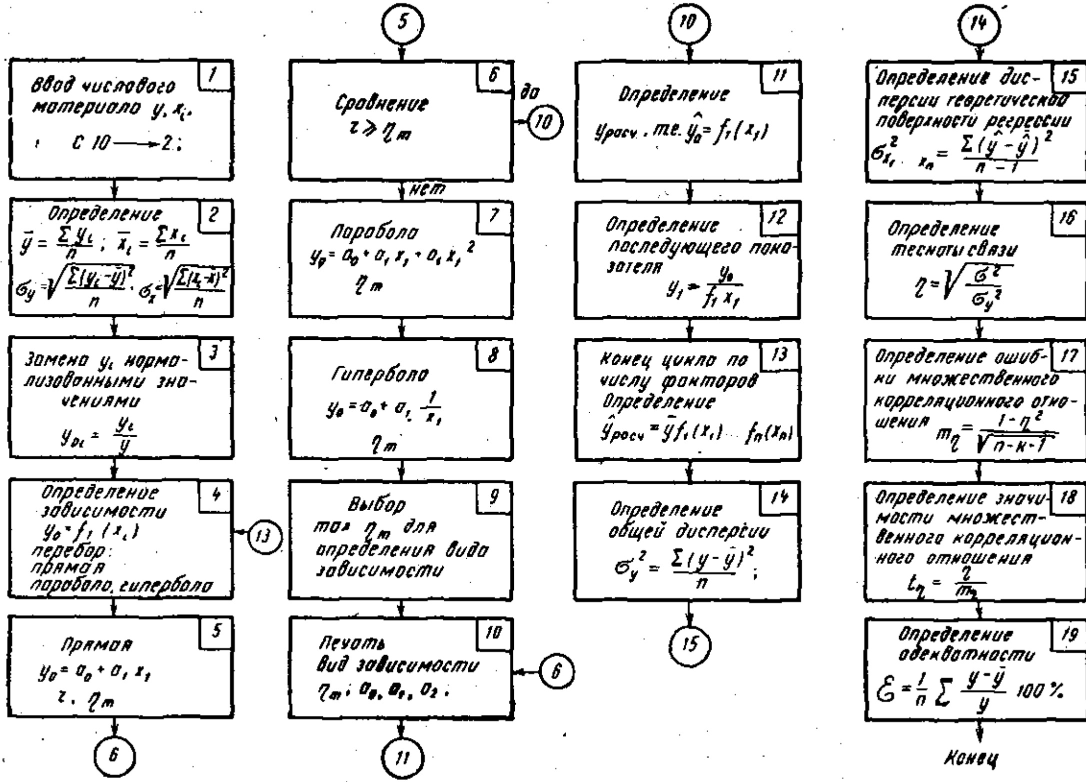


Рис. 27. Блок-схема вывода уравнения множественной регрессии при использовании графического метода.

который в наибольшей мере был бы адекватен реальному изменению признака под влиянием фактора. Адекватность проверяется по F-критерию (критерию Фишера) и по показателю $\bar{\varepsilon}$.

Построение экономико-статистической модели с помощью метода многошагового регрессионного анализа иллюстрируется следующим примером (данные приведены в § 1 стр. 146). Сначала в модель (полином) включаются все факторы в первой степени:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n.$$

В нашем примере

$$\hat{y} = 1{,}0793 + 0{,}117 \cdot 10^{-5} x_1 - 0{,}177 \cdot 10^{-4} x_2 - 0{,}227 \cdot 10^{-2} x_3. \qquad \text{(IV.12)}$$

Для определения F-критерия необходимо найти отношение дисперсии фактических значений $\sigma^2_{\text{ср}}$ к остаточной дисперсии $\sigma^2_{\text{ост}}$ признака (в данном случае уровня себестоимости):

$$F = \frac{\sigma^2_{\text{ср}}}{\sigma^2_{\text{ост}}} > 1. \qquad \text{(IV.13)}$$

Остаточная дисперсия

$$\sigma^2_{\text{ост}} = \frac{\sum (y_{\text{ф}} - \hat{y}_{\text{р}})^2}{n - (p - 1)}, \qquad \text{(IV.14)}$$

где $\hat{y}_{\text{р}}$ — расчетные значения уровня себестоимости, получаемые при подстановке в уравнение (IV.12) факторов для i-го строительно-монтажного управления;

$y_{\text{ф}}$ — фактические значения уровня себестоимости;

n — число управлений в рассматриваемой совокупности;

p — число коэффициентов регрессии (с учетом свободного члена).

Знаменатель формулы (IV.14) представляет собой выражение числа степеней свободы. Исчислим остаточную дисперсию по данным, приведенным в табл. 49.

$$\sigma^2_{\text{ост}} = \frac{0{,}14219}{103 - (4 - 1)} = 0{,}00142.$$

Таблица 49

Расчет остаточной дисперсии $\sigma^2_{ост}$ для полинома первого порядка

№ СУ	Значения факторов			$y_ф$	y по уравнению (IV.12)	К расчету $\sigma^2_{ост}$		К расчету $\sigma^2_{ср}$	
	x_1 X2	x_2 X1	x_3 Y2			$\lvert y_ф-\hat{y}_р\rvert$	$(y_ф-\hat{y}_р)^2$	$\lvert y_ф-\bar{y}_ф\rvert$	$(y_ф-\bar{y}_ф)^2$
1	2	3	4	5	6	7	8	9	10
1	2 110	6 061	1,09	0,933	0,911	0,022	0,00048	0,036	0,00129
2	2 225	7 605	1,04	0,805	0,869	0,063	0,00397	0,092	0,00847
3	1 812	6 591	1,10	0,907	0,896	0,011	0,00012	0,010	0,00010
4	1 322	5 317	0,85	0,834	0,920	0,003	0,00001	0,037	0,00137
5	1 261	5 713	0,89	0,914	0,896	0,006	0,00004	0,017	0,00029
6	1 707	6 598	0,84	0,844	0,917	0,052	0,00270	0,053	0,00381
...	...	...	...	...	...	...	...	...	...
96	1 962	5 866	0,76	0,928	0,917	0,011	0,00012	0,031	0,00096
97	2 268	6 659	0,92	0,918	0,895	0,023	0,00053	0,021	0,00044
98	2 383	7 829	1,02	0,821	0,862	0,042	0,00176	0,076	0,00578
99	2 539	7 914	0,85	0,930	0,861	0,069	0,00476	0,033	0,00109
100	1 245	5 858	1,00	0,939	0,915	0,023	0,00053	0,042	0,00176
101	1 313	6 154	1,21	0,920	0,907	0,013	0,00017	0,023	0,00053
102	1 674	6 762	1,19	0,873	0,891	0,018	0,00032	0,024	0,00057
103	2 452	6 361	0,80	0,893	0,904	0,011	0,00012	0,004	0,00002
				$\bar{y}_ф = 0,897$			$\sum(y_ф-\hat{y}_р)^2 = 0,14219$		$\sum(y_ф-\bar{y}_ф)^2 = 0,19915$

Дисперсия фактических значений относительно среднего значения переменной y:

$$\sigma^2_{ср} = \frac{\sum(y_ф-\bar{y}_ф)^2}{n-1} = \frac{0,19915}{103-1} = 0,00196.$$

Откуда

$$F = \frac{0,00196}{0,00142} = 1,378.$$

Как видим, условие (IV.13) выдержано.

Далее необходимо сопоставить F с F_q (критическим значением).

F_q находим для уровня значимости, который принят равным $q = 5\%$, по одной из таблиц значений F_q в зависимости от числа степеней свободы (k_1 и k_2)[1]. В данном случае $k_1 = n - 1 = 103 - 1 = 102$; $k_2 = n - (p - 1) = 103 - (4 - 1) = 100$. При этом $F_{00,5} = 1{,}39$. Так как полученное значение ниже критического ($1{,}378 < 1{,}39$), то уравнение (IV.12) нуждается в дальнейшем исследовании. Суть его состоит в повышении степени полинома, аппроксимирующего уровень себестоимости:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \ldots + a_k x_k + a_{11} x_1^2 + a_{12} x_2^2 + \\ + \ldots + a_{1k} x_{1k}^2 + a_{21} x_1 x_2 + a_{22} x_1 x_3 + \ldots + a_{2k} x_{k-1} x_k. \quad \text{(IV.15)}$$

С повышением степени полинома растет число его членов, количество коэффициентов полинома определяется по выражению $p = C_{k+d}^d$. В данном случае число независимых переменных $k = 3$ (x_1, x_2, x_3), степень полинома $d = 2$, тогда число коэффициентов $p = C_{3+2}^2 = 10$. Несколько упростим уравнение (IV.15), заменив члены второго порядка линейными:

$$x_1^2 = x_{k+1}; \; x_2^2 = x_{k+2}; \; x_3^2 = x_{k+3} \ldots x_k^2 = x_{2k};$$
$$x_1 x_2 = x_{2k+1} \ldots x_{k-1} x_k = x_k';$$

В данном случае

$$x_1^2 = x_4; \; x_2^2 = x_5; \; x_3^2 = x_6; \; x_1 x_2 = x_7; \; x_1 x_3 = x_8; \; x_2 x_3 = x_9.$$

Подставив в уравнение (IV.15) замены, получим:

$$\hat{y} = b_0 x_0 + b_1 x_1 \ldots + b_{k'} x_{k'}, \quad \text{где } k' = C_{k+d-1}^d,$$

Минимизируем сумму квадратов отклонений:

$$\sum_{u=1}^{n} (y_u - b_0 x_{ou} - b_1 x_{1u} - \ldots - b_k x_{ku})^2 = \min.$$

Приравнивая нулю частные производные от этой квадратичной формы, взятые по переменным b_0, $b_1, \ldots, b_k$, получим систему нормальных уравнений:

$$b_0 (00) + b_1 (01) + \ldots + b_k (0k) = (0k);$$
$$b_0 (01) + b_1 (11) + \ldots + b_k (1k) = (1y);$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$b_0 (k0) + b_1 (k1) + \ldots + b_k (kk) = (ky).$$

[1] Рекомендуем Н. В. Смирнов, И. В. Дунин-Барковский «Курс теории вероятностей и математической статистики», М., «Наука», 1965. Приложение, табл. VI.

Чтобы найти нужные коэффициенты регрессии, следует решить систему нормальных уравнений относительно b_0, b_1, ..., b_k.

Это делается при помощи матрицы, посредством которой записываются нормальные уравнения:

$$A=\begin{vmatrix}(00) & (01)\dots(0k)\\ (10) & (11)\dots(1k)\\ \dots & \dots\dots\\ (k0) & (k1)\quad(kk)\end{vmatrix}\quad Y=\begin{vmatrix}(0y)\\ (1y)\\ \dots\\ (ky)\end{vmatrix}\quad b=\begin{vmatrix}b_0\\ b_1\\ \dots\\ b_k\end{vmatrix}.$$

Затем находим матрицу A^{-1}, обратную матрице A. Умножив A^{-1} на матрицу-столбец $У$, получим матрицу-столбец B, элементами которой являются коэффициенты регрессии b_0, b_1, ..., b_k.

В результате расчетов получили матрицу и матрицу-столбец, записанные в табл. 50.

Матричная таблица A квадратна: число строк и граф равно числу неизвестных, в данном случае десяти. Составляется система нормальных уравнений, коэффициентами которых являются члены матрицы. Уравнения составлены по единой схеме; приводим в качестве примера седьмое уравнение (см. табл. 50)

$$13{,}407+23{,}738\,x_1+87{,}827\,x_2+13{,}5244\,x_3+48{,}2295\,x_1x_2+\\+13{,}3959\,x_1x_3+165{,}174\,x_2x_3+165{,}174\,x_1^2+24{,}723\,x_2^2+\\+88{,}720\,x_3^2=11{,}882.$$

Решив систему уравнений, получим следующее уравнение уровня себестоимости:

$$\hat{y}=0{,}84137+0{,}01721\,x_1-0{,}06757x_2+0{,}081692\,x_3+\\+0{,}02689\,x_1x_2-0{,}23251\,x_1x_3-0{,}04377\,x_2x_3+\\+0{,}01852\,x_1^2+0{,}00004\,x_2^2-0{,}03913x_3^2.\qquad\text{(IV.16)}$$

Второй этап исследования заключается в определении F и F_q уравнения (IV.16). Для этого находим остаточную дисперсию нелинейных членов уравнения по формуле:

$$\sigma^2_{\text{ост}}=\frac{\sum(y_{\text{ф}}-\hat{y}_p)^2}{n-(p'-m)},$$

где p' — число коэффициентов регрессии (без учета свободного члена);

m — число линейных коэффициентов.

Таблица 50

Матричная таблица для определения уравнения зависимости уровня себестоимости от факторов x_1, x_2, x_3

A										*У*
103	26,273	91,288	13,560	52,1474	603,5123	13,4070	174,8044	25,2409	88,6494	12,559
26,273	52,1474	174,8044	25,2409	108,4051	1179,013	24,7383	353,0901	49,6735	168,284	23,4804
91,288	174,8044	603,611	88,6494	353,0901	87,827	1179,094	1179,094	168,294	587,160	81,6016
13,560	25,2409	88,6484	13,407	49.6735	13,5244	168,294	168,294	24,738	87,827	12,151
52,1474	108,4051	353,094	49,8735	233,683	48,2295	744,759	744,759	102,4075	336,971	46,473
603,611	1179,013	4046,19	5,8716	2423,072	582,904	8064,3759	8064,375	1136,83	3942,84	537,539
13,407	23,738	87,827	13,5244	48,2295	13,3959	165,174	165,174	24,723	88,720	11,882
174,804	359,090	1179,013	168,294	744,759	165,174	2423,072	2423,072	336,971	1136,830	155,656
25,2409	49,6735	169,290	24,7383	102,4075	24,723	336,971	336,971	48,2295	165,174	22,558
88,6494	168,290	584,16	87,827	336,571	88,72	1136,830	1136,830	165,174	582,9048	79,651

А — матрица сумм; *У* — матрица-столбец.

В уравнении (IV.16) $p' = 10 - 1 = 9$; $m = 3$ и число степеней свободы: $k_1 = 102$; $k_2 = 103 - (9 - 3) = 97$.

$$\sigma^2_{\text{ост}} = \frac{0,09473}{97} = 0,00098.$$

Расчет $\sum(y_\text{ф} - \hat{y}_\text{р})^2$ приведен в табл. 51.

Таблица 51

Расчет остаточной дисперсии $\sigma^2_{\text{ост}}$ для полинома второго порядка

№ СУ	$y_\text{ф}$	$\hat{y}_\text{р}$	$\lvert y_\text{ф} - \hat{y}_\text{р} \rvert$	$(y_\text{ф} - \hat{y}_\text{р})^2$
1	2	3	4	5
1	0,933	0,916	0,017	0,00029
2	0,805	0,838	0,033	0,00109
3	0,907	0,881	0,026	0,00068
4	0,834	0,834	0,000	0,00000
5	0,914	0,909	0,005	0,00003
6	0,844	0,866	0,022	0,00048
...	...	...	...	...
96	0,928	0,917	0,011	0,00012
97	0,918	0,899	0,019	0,00036
98	0,821	0,841	0,020	0,00040
99	0,930	0,883	0,047	0,00221
100	0,939	0,925	0,014	0,00020
101	0,920	0,935	0,015	0,00023
102	0,873	0,873	0,000	0,00000
103	0,893	0,939	0,046	0,00212
				$\sum(y_\text{ф} - \tilde{y}_\text{р})^2 = 0,09473$

Для уравнения (IV.16) $F = \frac{0,00196}{0,00098} = 2,0062$ и $F_{0,05} = 1,41$. Таким образом, уравнение регрессии уровня себестоимости в виде полинома второго порядка удовлетворяет условию (IV.13). Для выявления значимых членов уравнения (IV.16) производим оценку коэффициентов регрессии с помощью критерия t_{b_i}:

$$t_{b_i} = \frac{b_i}{\sigma_{\text{ост}} \sqrt{C_{ii}}}, \tag{IV.17}$$

где b_i — коэффициенты регрессии;

C_{ii} — диагональный элемент обратной матрицы $A \to (A^{-1})$.

Оценка выполняется поэтапно. Сначала определяются элементы обратной матрицы (A^{-1})[1].

Найдем матрицу A^{-1}, обратную матрице

$$A = \begin{vmatrix} a_{1,} \dots a_{1n} \\ a_{n1} \dots a_{nn} \end{vmatrix}.$$

Для этого:

вычисляем определитель этой матрицы $Д$;

определяем алгебраическое дополнение A_{ij} каждого элемента определителя матрицы a_{ij};
составляем из чисел A матрицу.

$$\widetilde{A} = \begin{Vmatrix} A_{11} & A_{12} \dots A_{1n} \\ A_{21} & A_{22} \dots A_{2n} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ A_{n1} & A_{n2} \dots A_{nn} \end{Vmatrix}$$

транспонируем матрицу $\widetilde{A}$ и получаем:

$$\widetilde{A}' = \begin{Vmatrix} A_{11} & A_{21} \dots A_{n1} \\ A_{12} & A_{22} \dots A_{n2} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \\ A_{1n} & A_{2n} \dots A_{nn} \end{Vmatrix}$$

составляем матрицу $A^{-1} = \dfrac{\widetilde{A}'}{Д}$ и обозначаем элементы этой обратной матрицы через C_{ij}:

$$A^{-1} = \begin{Vmatrix} C_{00} & C_{01} & C_{02} \dots C_{0n} \\ C_{10} & C_{11} & C_{12} \dots C_{1n} \\ \cdot & \cdot & \cdot \quad \cdot \quad \cdot \\ C_{n0} & C_{n1} & C_{n2} \dots C_{nn} \end{Vmatrix}.$$

Диагональные элементы матрицы A^{-1} представляют собой последовательность $C_{00}, C_{11}, C_{22}, \dots, C_{nn}$.

---

[1] См. Д. К. Фадеев, В. Н. Фадеева. Вычислительные методы линейной алгебры. М.—Л., Гос. издательство физико-математической литературы, 1963.

Значения диагональных элементов равны:

$$\begin{aligned} C_{00} &= 257,227; & C_{55} &= 40,422; \\ C_{11} &= 5,590; & C_{66} &= 8,699; \\ C_{22} &= 309,526; & C_{77} &= 1,637; \\ C_{33} &= 109,043; & C_{88} &= 1,304; \\ C_{44} &= 0,00692; & C_{99} &= 47,597. \end{aligned}$$

Члены уравнения, для которых значения t_{b_i} минимальны, признаем несущественными и исключаем из уравнения (IV.17).

Например,

$$t_{b_i} = \frac{|b_{33}|}{\sigma_{\text{ост}} \sqrt{C_{33}}} = \frac{0,01721}{\sqrt{0,00098} \cdot \sqrt{109,043}} = 0,05273.$$

Значения t_{b_i} для всех коэффициентов регрессии приведены в табл. 52, из которой следует, что достаточно минимальными значениями t_{b_3} (0,05273) и t_{b_4} (0,01551) можно пренебречь; поэтому исключаем b_3 и b_4 из уравнения (IV.16).

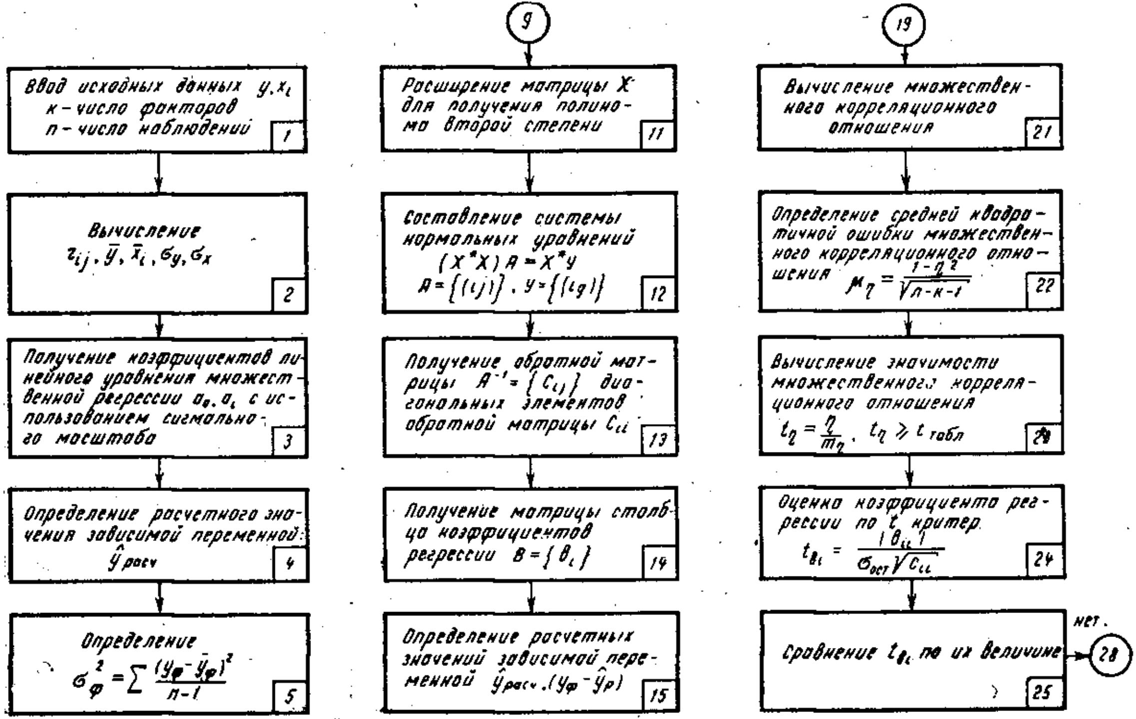
Далее повторяем указанную операцию, каждый раз исключая t_{b_i}, имеющие значения на порядок ниже остальных, одновременно наблюдая характер изменения $\sigma^2_{\text{ост}}$. Окончательным вариантом расчета будет тот, при котором наблюдается минимальное значение $\sigma^2_{\text{ост}}$. В данном случае таким будет четвертый вариант, ибо в дальнейшем значение $\sigma^2_{\text{ост}}$ вновь повышается. Окончательно уравнение регрессии имеет вид:

$$\check{y} = 0,95707 - 0,08601\, x_1 + 0,74374\, x_2 - 0,04822\, x_1 x_2 + \\ + 0,03822\, x_1 x_3 - 0,21706\, x_2 x_3. \qquad \text{(IV.18)}$$

Все расчеты и оценки F, η, $\sigma^2_{\text{ост}}$ приведены в табл. 52. Статистические характеристики уравнения (IV.18):

$\sigma^2_{\text{ост}} = 0,00067$; $\eta = 0,832$; $F = 2,92$ (при табличном $F_{0,05} = 1,39$).

Основным критерием оценки достоверности полученного уравнения множественной корреляции является остаточная дисперсия теоретической поверхности регрессии — $\sigma^2_{\text{ост}}$, характеризующая влияние на признак только отобранных факторов: x_1, x_2, и x_3.



Ввод исходных данных y, x_i к – число факторов n – число наблюдений
1
Вычисление $r_{ij}, \bar{y}, \bar{x}_i, \sigma_y, \sigma_x$
2
Получение коэффициентов линейного уравнения множественной регрессии a_0, a_i с использованием сигмального масштаба
3
Определение расчетного значения зависимой переменной $\hat{y}_{расч}$
4
Определение $\sigma_\varphi^2 = \sum \frac{(y_\varphi - \bar{y}_\varphi)^2}{n-1}$
5
9
Расширение матрицы X для получения полинома второй степени
11
Составление системы нормальных уравнений $(X^*X)A = X^*y$ $A = \{(ij)\}, y = \{(ig)\}$
12
Получение обратной матрицы $A^{-1} = \{C_{ij}\}$ диагональных элементов обратной матрицы C_{ii}
13
Получение матрицы столбца коэффициентов регрессии $B = \{b_i\}$
14
Определение расчетных значений зависимой переменной $\hat{y}_{расч}, (y_\varphi - \hat{y}_p)$
15
19
Вычисление множественного корреляционного отношения
21
Определение средней квадратичной ошибки множественного корреляционного отношения $\mu_\eta = \frac{1-\eta^2}{\sqrt{n-k-1}}$
22
Вычисление значимости множественного корреляционного отношения $t_\eta = \frac{\eta}{m_\eta}$, $t_\eta \geqslant t_{табл}$
23
Оценка коэффициента регрессии по t критер. $t_{b_i} = \frac{|b_{ii}|}{\sigma_{ост}\sqrt{C_{ii}}}$
24
Сравнение t_{b_i} по их величине
25
нет
28

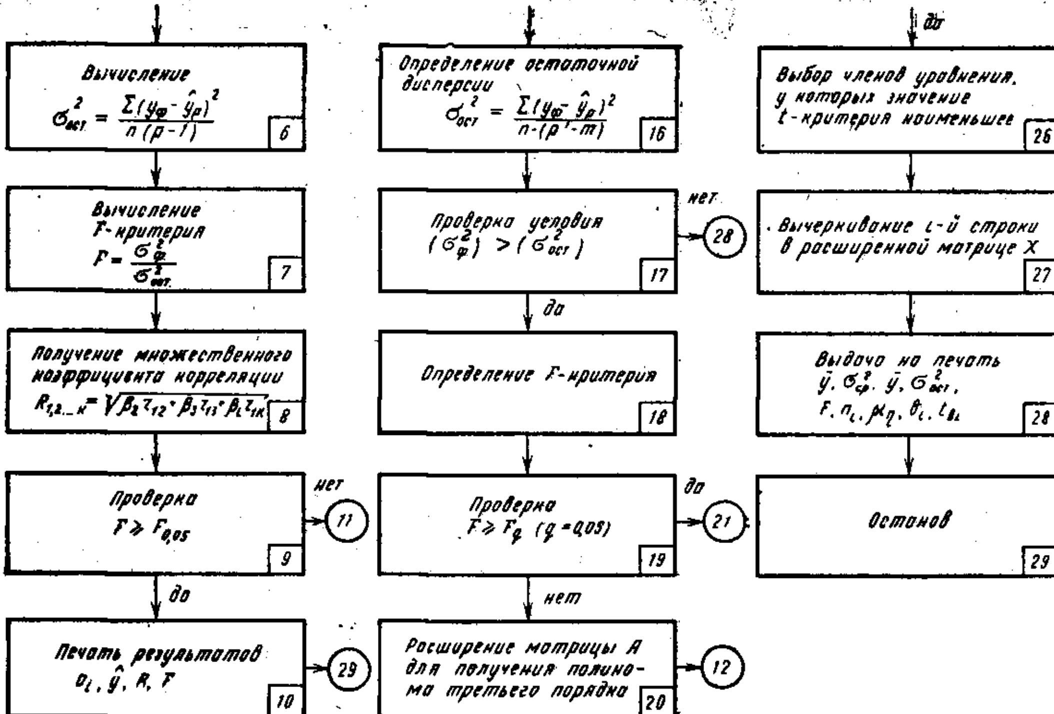


Рис. 28. Блок-схема алгоритма многошагового регрессионного анализа.

Таблица 52

Данные по различным вариантам расчета

Номер коэффициента	Обозначение величин	Варианты расчета $\sigma^2_{ост}$				
		I	II	III	IV	V
1	b_0	0,84135	0,90117	0,90343	0,95707	—
	b_1	—0,06757	—0,06864	—0,06965	—0,08601	—
	t_1	0,88652	0,96740	1,02000	0,97620	—
2	b_2	0,81693	0,73078	0,72603	0,74374	—
	t_2	1,48588	0,65240	2,32700	2,59490	—
3	b_3	0,01721	—	—	—	—
	t_3	0,05273	—	—	—	—
4	b_4	0,00004	—	—	—	—
	t_4	0,01551	—	—	—	—
5	b_5	—0,03913	—0,00224	—	—	—
	t_5	0,19693	0,00438	—	—	—
6	b_6	0,01853	0,01959	0,01947	—	—
	t_6	0,20102	0,36370	0,39880	—	—
7	b_7	—0,04377	—0,04393	—0,04393	—0,64822	—
	t_7	1,09488	1,12990	1,19180	1,47755	—
8	b_8	0,02690	0,02778	0,02780	0,03822	—
	t_8	0,75387	0,94180	0,89650	2,45610	—
9	b_9	—0,23251	—0,22389	—0,22374	—0,21706	—
	t_9	1,07844	1,84580	2,04180	2,16460	—
	F	2,0068	2,256	2,507	2,920	—
	$\sigma^2_{ост}$	0,00098	0,00087	0,00078	0,00067	0,00106
	$F_{0,05}$	1,410	1,392	1,390	1,390	—
	$\sigma^2_{ср}$	0,00196	0,00196	0,00196	0,00196	—
	η	0,803	0,833	0,833	0,832	—

Значения $\sigma^2_{ост}$ для модели уровня себестоимости, полученной различными методами, равны:

при графическом методе (различные формы связей)	0,00168
при степенной функции .	0,00088
при многошаговом регрессионном анализе . . .	0,00067

Блок-схема алгоритма многошагового регрессионного анализа приведена на рис. 28.

Глава V

ЧАСТНЫЕ КОЭФФИЦИЕНТЫ КОРРЕЛЯЦИИ

§ 1. Определение частных коэффициентов корреляции

На величину парных коэффициентов корреляции могут влиять, помимо исследуемых двух переменных, и так называемые «скрытые», или неучтенные, факторы. Неучтенными могут быть как факторы, включенные в исследование, так и факторы, которые еще неизвестны.

Совместное влияние неучтенных факторов в некоторых случаях может быть так велико, что способно резко изменить силу связи, выражаемую величиной парного коэффициента корреляции. Для исследования действия неучтенных факторов пользуются так называемыми частными коэффициентами корреляции, которые позволяют измерить силу связи между двумя величинами при условии «закрепления» влияния ряда остальных факторов. Если при этом принято во внимание влияние только одной величины из числа неучтенных, то частный коэффициент корреляции называют коэффициентом первого порядка; если двух — коэффициент называют частным коэффициентом корреляции второго порядка и т. д. Следовательно, парный коэффициент корреляции можно назвать частным коэффициентом корреляции нулевого порядка, ибо в нем локализируется попеременно влияние только одной из двух величин, сила взаимодействия которых и характеризуется парным коэффициентом корреляции.

Введем следующие обозначения:

частный коэффициент корреляции нулевого порядка (парный коэффициент корреляции) — ^0r

частный коэффициент корреляции первого порядка — ^1r

» » » второго » — ^2r

частный коэффициент корреляции третьего порядка — ${}^3r$
» » » четвертого » — ${}^4r$
и т. д. Частные коэффициенты корреляции могут быть исчислены по следующей общей формуле:

$$ {}^1r = r_{12\cdot 3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}, \tag{V.1} $$

из которой следует, что частный коэффициент корреляции первого порядка определяется через частные коэффициенты корреляции нулевого порядка (т. е. парные коэффициенты корреляции). Аналогично определяются частные коэффициенты корреляции n-го порядка из коэффициента ближайшего n-1-го порядка:

$$ r_{ab\cdot cd\ldots m} = \frac{r_{ab\cdot cd\ldots(m-1)} - r_{am\cdot cd\ldots(m-1)}\cdot r_{bm\cdot cd\ldots(m-1)}}{\sqrt{1-r_{am\cdot cd\ldots(m-1)}^2}\sqrt{1-r_{bm\cdot cd\ldots(m-1)}^2}}, \tag{V.2} $$

где a, b, c, d, ..., m — факторы.

Так,

$$ {}^2r_{12\cdot 34} = \frac{r_{12\cdot 3} - r_{14\cdot 3}\cdot r_{24\cdot 3}}{\sqrt{1-r_{14\cdot 3}^2}\sqrt{1-r_{24\cdot 3}^2}}; $$

$$ {}^3r_{12\cdot 345} = \frac{r_{12\cdot 34} - r_{15\cdot 34}\cdot r_{25\cdot 34}}{\sqrt{1-r_{15\cdot 34}^2}\sqrt{1-r_{25\cdot 34}^2}}; $$

$$ {}^4r_{12\cdot 3456} = \frac{r_{12\cdot 345} - r_{16\cdot 345}\cdot r_{26\cdot 345}}{\sqrt{1-r_{16\cdot 345}^2}\sqrt{1-r_{26\cdot 345}^2}}. $$

В индексе r слева от точки указываются величины, для которых определяется r, справа — величины, влияние которых элиминируется, остается постоянным. Так как слева от точки всегда проставляются только две цифры, точку можно опускать. Например, выражение $r_{26.345}$ следует читать: коэффициент корреляции между величинами 2 и 6 при локализации величин 3, 4 ,5.

Определение частных коэффициентов корреляции весьма трудоемкая операция и поэтому редко выполняется несмотря на важность информации, которую в результате ее получает экономист. Исчисление частных коэффициентов корреляции, имеющих порядок $n > 4$, сопряжено с решением нескольких сот формул и поэтому должно быть передано на электронно-вычислительные машины. Для этого имеются специальные программы, в частности рекомендуются программы, разработанные Киевским институтом автоматики.

В исследованиях экономики строительства обычно рассматривается действие трех, четырех и значительно реже пяти факторов. Для этих случаев рекомендуются таблицы, пользование которыми значительно уменьшает трудоемкость исчисления частных коэффициентов корреляции без использования ЭВМ.

Проследим ход выполнения расчетов частных коэффициентов корреляции до четвертого порядка (см. табл. 53), для упрощения в ней приведены только символы при r. Из таблицы следует, что для ^4r необходимо знать все парные коэффициенты корреляции ($^\circ r$) между признаком и каждым из пяти факторов, а также семь парных коэффициентов корреляции между факторами (r_{23}, r_{24}, r_{25}, r_{26}, r_{34}, r_{35}, r_{36}). Они должны быть известны с самого начала исследования.

Для исчисления первых двух ^4r необходимо рассчитать: 10 коэффициентов корреляции первого порядка, 6 — второго, 3 — третьего и 1 — четвертого порядка, что с помощью рекомендуемых[1] таблиц представляется не чрезмерно трудоемкой задачей.

Обратим внимание, что для $r_{14.2356}$ количество вычислений сокращается на 25% (см. табл. 53), для $r_{15.2346}$ — на 60%, а последний $r_{16.2345}$ вычисляется только по одной формуле; в среднем для вычисления одного ^4r приходится выполнить 14 операций.

Определим ^1r и ^2r по уже приведенным данным о накладных расходах (см. стр. 108). При воздействии на уровень накладных расходов двух факторов — объема работ (x_2) и численности работников (x_3) $^\circ r$ имели следующие значения: $r_{12} = 0{,}839$; $r_{13} = 0{,}886$ и $r_{23} = 0{,}895$. Найдем частные коэффициенты корреляции ^1r при элиминировании влияния каждого из двух остальных факторов:

$$r_{12.3} = \frac{0{,}839 - 0{,}886 \cdot 0{,}895}{\sqrt{1 - 0{,}886^2} \cdot \sqrt{1 - 0{,}895^2}} = \frac{0{,}046}{0{,}467 \cdot 0{,}445} = 0{,}221;$$

$$r_{13.2} = \frac{0{,}886 - 0{,}839 \cdot 0{,}895}{\sqrt{1 - 0{,}839^2} \cdot \sqrt{1 - 0{,}895^2}} = \frac{0{,}135}{0{,}544 \cdot 0{,}445} = 0{,}556.$$

---

[1] См. приложение 3.

Таблица 53

Порядок расчета частных коэффициентов корреляции

Исходная информация	⁰r признак-фактор	⁰r фактор-фактор	Рассчитываемые частные коэффициенты корреляции	¹r признак-фактор	¹r фактор-фактор	²r признак-фактор	²r фактор-фактор	³r признак-фактор	³r фактор-фактор	⁴r
	12 13 14 15 16	23 34 35 36 46 56		12.3 14.3 15.3 16.3	24.3 25.3 26.3 45.3 46.3 56.3	12.34 15.34 16.34	25.34 26.34 56.34	12.345 16.345	26.354	$r_{12.3456}$
	12 13 14 15 16	23 24 25 26 46 56		13.2 14.2 15.2 16.2	34.2 35.2 36.2 45.2 46.2 56.2	13.24 15.24 16.24	35.24 36.24 56.24	13.245 16.245	36.245	$r_{13.2456}$
	12 13 14 15 16	23 24 25 26 46 56		13.2 14.2 15.2 16.2	34.2 35.2 36.2 46.2 56.2	14.23 15.23 16.23	45.23 46.23 56.23	14.235 16.235	46.235	$r_{14.2356}$
	12 13 14 15 16	23 24 25 26 46 56		13.2 14.2 15.2 16.2	34.2 35.2 36.2 46.2 56.2	14.23 15.23 16.23	45.23 46.23 56.23	15.234 16.234	56.234	$r_{15.2346}$
	12 13 14 15 16	23 24 25 26 46 56		13.2 14.2 15.2 16.2	34.2 35.2 36.2 46.2 56.2	14.23 15.23 16.23	45.23 46.23 56.23	15.234 16 234	56.234	$r_{16.2345}$

При исследовании зависимости накладных расходов от трех факторов: объема работ, численности работников и фонда заработной платы рабочих (x_4), необходимо дополнительно исчислить:

$$r_{14}=0,838;\quad r_{24}=0,837;\quad r_{34}=0,869.$$

Частные коэффициенты корреляции ^1r при элиминировании влияния каждого их трех факторов определяются аналогично ($r_{12.3}$ и $r_{13.2}$ уже известны):

$$r_{14.2}=\frac{0,838-0,839\cdot0,837}{\sqrt{1-0,839^2}\cdot\sqrt{1-0,837^2}}=\frac{0,136}{0,544\cdot0,547}=0,456;$$

$$r_{34.2}=\frac{0,869-0,895\cdot0,837}{\sqrt{1-0,895^2}\cdot\sqrt{1-0,837^2}}=\frac{0,120}{0,445\cdot0,547}=0,494;$$

$$r_{14.3}=\frac{0,838-0,886\cdot0,869}{\sqrt{1-0,886^2}\cdot\sqrt{1-0,869^2}}=\frac{0,068}{0,463\cdot0,495}=0,296;$$

$$r_{24.3}=\frac{0,837-0,895\cdot0,869}{\sqrt{1-895^2}\cdot\sqrt{1-0,869^2}}=\frac{0,060}{0,445\cdot0,495}=0,268;$$

$$r_{12.4}=\frac{0,839-0,838\cdot0,837}{\sqrt{1-0,838^2}\cdot\sqrt{1-0,837^2}}=\frac{0,138}{0,545\cdot0,546}=0,464;$$

$$r_{13.4}=\frac{0,886-0,838\cdot0,869}{\sqrt{1-0,838^2}\cdot\sqrt{1-0,869^2}}=\frac{0,158}{0,545\cdot0,495}=0,586;$$

$$r_{23.4}=\frac{0,895-0,837\cdot0,869}{\sqrt{1-837^2}\cdot\sqrt{1-0,869^2}}=\frac{0,168}{0,547\cdot0,495}=0,621.$$

Из полученных значений коэффициентов первого порядка (^1r), пользуясь рекуррентной формулой, получаем:

$$r_{14.23}=\frac{0,456-0,556\cdot0,494}{\sqrt{1-0,556^2}\cdot\sqrt{1-0,494^2}}=\frac{0,184}{0,830\cdot0,871}=0,254;$$

$$r_{13.24}=\frac{0,556-0,456\cdot0,494}{\sqrt{1-0,456^2}\cdot\sqrt{1-0,494^2}}=\frac{0,331}{0,890\cdot0,871}=0,429;$$

$$r_{12.34}=\frac{0,221-0,296\cdot0,268}{\sqrt{1-0,296^2}\cdot\sqrt{1-0,268^2}}=\frac{0,141}{0,955\cdot0,962}=0,153.$$

Выпишем рассмотренные коэффициенты корреляции между признаком и факторами в табл. 54.

Обратим внимание, что в обоих случаях, т. е. при действии двух и трех факторов величины коэффициентов ^1r и ^2r снижаются в сравнении с ^0r. Особенно заметно, до 0,221, снижается $r_{12.3}$ при элиминировании влияния фактора x_3; если исключается еще влияние и x_4, то сила взаимодействия между накладными расходами и объемом работ характеризуется $r_{12.34}=0,153$.

Очевидно, на величину парного коэффициента корреляции между величиной накладных расходов и объемами

Таблица 54

Частные коэффициенты корреляции

При воздействии на признак двух факторов			При воздействии на признак трех факторов		
${}^{0}r$	$r_{12} = 0{,}839$	$r_{13} = 0{,}886$	$r_{12} = 0{,}839$	$r_{13} = 0{,}886$	$r_{14} = 0{,}838$
${}^{1}r$	$r_{12.3} = 0{,}221$ —	$r_{12.3} = 0{,}556$ —	$r_{12.3} = 0{,}221$ $r_{12.4} = -0{,}464$	$r_{13.2} = 0{,}556$ $r_{13.4} = 0{,}586$	$r_{14.2} = 0{,}456$ $r_{14.3} = 0{,}296$
${}^{2}r$	—	—	$r_{12.34} = 0{,}153$	$r_{13.24} = 0{,}429$	$r_{14.23} = 0{,}254$

выполняемых строительно-монтажных работ ($r_{12} = 0{,}839$) косвенно оказывает влияние численность рабочих x_3, которая непосредственно связана с выработкой и объемами работ.

Но без расчета частных коэффициентов корреляции это явление невозможно установить. Аналогичные заключения возможны и в отношении других переменных, связанных с накладными расходами.

Пример определения частного коэффициента корреляции ($r_{12.3456}$) приведен (см. приложение 3) по данным анализа изменения себестоимости (x_1) строительных управлений, специализированных на выполнении монтажных работ, в зависимости от следующих факторов объем работ — x_2, основные фонды — x_3, выработка — x_4, уровень ритмичности — x_5 и оборотные фонды — x_6. Нами рассчитан только $r_{12.3456} = -0{,}464$. Рекомендуем читателю выполнить расчеты остальных частных коэффициентов корреляции четвертого порядка, проверив значения: $r_{13.2456} = -0{,}139$; $r_{14.2356} = -0{,}315$; $r_{15.2346} = -0{,}531$ и $r_{16.2345} = 0{,}042$ (необходимые данные имеются в приложении 3). При этом важно проследить, как снижается абсолютная величина частных коэффициентов корреляции при постепенном включении в исследование новых факторов.

Рассматривая влияние фактора x_2, видим, что коэффициенты корреляции имели следующие значения:
${}^{0}r = -0{,}850$; ${}^{1}r = -0{,}740$; ${}^{2}r = -0{,}600$; ${}^{3}r = -0{,}512$; ${}^{4}r = -0{,}464$. А как поведут себя остальные факторы? На рис. 29 приведена динамика частных коэффициентов

корреляции в рассматриваемом примере при условии, что значения парных коэффициентов корреляции приняты равными 100%.

В рассматриваемых исследованиях влияние факторов на признак было однонаправленным: каждый последующий фактор усиливал связь между признаком и первым фактором. Поэтому постепенное, поочередное исключение факторов вызывало одинаковый по характеру результат: сокращение абсолютной величины частных коэффициентов

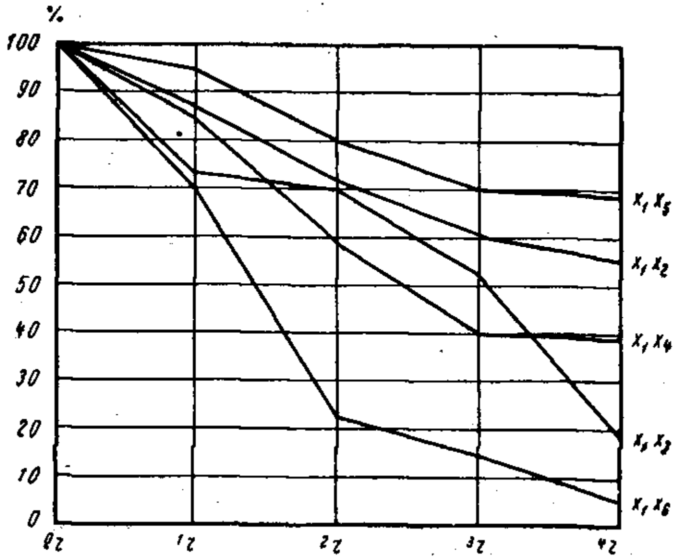


Рис. 29. Динамика частных коэффициентов корреляции.

корреляции. Однако возможно и иное положение, при котором элиминирование влияния какого-либо одного или нескольких факторов приводит к повышению абсолютной величины частного коэффициента корреляции.

Итак, соотношение между парными и частными коэффициентами корреляции представляется в следующем виде.

Парные коэффициенты корреляции устанавливаются на основании внешних проявлений связей между переменными величинами, участвующими в формировании того или иного процесса; они представляют первую (неполную, а порой и неточную) информацию о тесноте связей между рассматриваемыми величинами.

Частные коэффициенты корреляции исчисляются на основе более глубокого анализа явлений и дают более точное представление о характере взаимодействия и тесноте связей между переменными величинами, помогают вскрывать внутренний механизм изучаемых процессов.

Такова, нам кажется, связь между этими коэффициентами — они дополняют друг друга.

Однако встречаются и иные точки зрения. Так, К. К. Вальтух [4] отмечает: «Некоторые авторы предлагают пользоваться для установления наличия или отсутствия той или иной связи эмпирическими полями корреляции (т. е. парными коэффициентами корреляции. — *И. П.*, *Л. Э.*) Этот путь представляется нам совершенно неприемлемым».

Но поставим себя на место экономиста, производящего анализ. Как он должен сформулировать первоначальное суждение о силе и характере связей между величинами? Разумеется, для такого суждения в распоряжении экономиста должны быть простейшие, доступные каждому методы построения парных связей. Направление, характер зависимостей экономист предвидит, полагаясь на знание существа рассматриваемых процессов; с помощью же парных коэффициентов корреляции он получает первоначальную количественную оценку тесноты ? связей. В дальнейшем при построении уравнений множественной регрессии первоначальные наметки могут подтвердиться, либо будут скорректированы, а для первых выводов нет иного пути, чем использование эмпирических полей корреляции и исчисление по ним парных коэффициентов корреляции. Это дает возможность экономисту исключить из дальнейшего исследования связи, характеризуемые парными коэффициентами корреляции, меньшими $0{,}2 \div 0{,}3$. Конечно, возможны и такие весьма сложные связи, где действительно трудно вначале безошибочно установить направления взаимодействия между переменными величинами. Тем важнее первоначально предположить более приемлемый знак коэффициентов корреляции и проверить это предположение последующим анализом.

Парные коэффициенты корреляции характеризуют тесноту связи между двумя величинами на основании первичной информации, на основании внешних проявлений различных экономических процессов. И как уже отмечалось, на их величине косвенно отражается влияние прочих неучтенных величин. Если суммарное влияние этих величин значительно, и если, в то же время, каждая из них не имеет существенного значения, либо влияние разнонаправленно, в этом случае парный коэффициент корреляции будет достаточно точно характеризовать тесноту связи между признаком и одним из факторов. В противном случае действительная теснота связи между двумя исследуемыми переменными может не только существенно измениться по абсолютной величине, но даже изменить направление, что получит выражение в перемене знака парного коэффициента корреляции.

Поэтому так велико значение частных коэффициентов корреляции, с помощью которых вскрывается существо процесса, возникает более полное представление о механизме связей между отдельными экономическими показателями.

Совершенно точного представления получить невозможно, так как в любое исследование мы не включаем (не можем включить) все действующие переменные. Объясняется это тем, что с введением каждого нового фактора трудоемкость вычислений возрастает в геометрической прогрессии, кроме того, из множества факторов оказывают существенное, решающее, определяющее влияние только некоторые.

Исследование влияния величины частных коэффициентов корреляции позволяет дать оценку действительной тесноты связей и отобрать для дальнейшего использования в планировании, нормировании либо для других целей только наиболее существенные факторы.

Как отмечают М. Езекиэл и К. Фокс [8], частные коэффициенты корреляции «... измеряют значение каждой из нескольких независимых переменных, устанавливая, насколько понижается определяемая часть вариации *после того, как исключено влияние всех других независимых переменных*» (курсив наш. — *И. П., Л. Э.*). Укажем, еще на одну область применения коэффициентов корреляции. С помощью частных коэффициентов корреляции и остаточной дисперсии ($\sigma^2_{\text{ост}}$) целесообразно определять параметры в уравнениях множественной регрессии. Общие формулы остаточной дисперсии:

$$\sigma^2_{\text{ост } 1234\ldots n} = \sigma_1^2\left(1-r_{12}^2\right)\left(1-r_{13.2}^2\right)\times \times\left(1-r_{14.23}^2\right)\ldots\left[1-r^2_{1n234\ldots(n-1)}\right] \qquad (V.3)$$

либо

$$\sigma^2_{\text{ост } 234\ldots(n-1)n1} = \sigma_1^2\left(1-r_{23}^2\right)\left(1-r_{24.3}^2\right)\times \times\left[1-r^2_{2(n-1)34\ldots n1}\right]\left[1-r^2_{21\ldots 34(n-1)n}\right].$$

Принимая во внимание пределы, для которых целесообразно использование частных коэффициентов корреляции и остаточной дисперсии, при расчетах вручную получим:

$$\sigma^2_{\text{ост } 12} = \sigma_1^2\left(1 - r_{12}^2\right);$$
$$\sigma^2_{\text{ост } 1.23} = \sigma_1^2\left(1 - r_{12}^2\right)\left(1 - r_{13.2}^2\right);$$
$$\sigma^2_{\text{ост } 1.234} = \sigma_1^2\left(1 - r_{12}^2\right)\left(1 - r_{13.2}^2\right)\left(1 - r_{14.23}^2\right);$$
$$\sigma^2_{\text{ост } 1.23456} = \sigma_1^2\left(1 - r_{12}^2\right)\left(1 - r_{13.2}^2\right)\left(1 - r_{14.23}^2\right) \times$$
$$\times \left(1 - r_{15.234}^2\right)\left(1 - r_{16.2345}^2\right).$$

Значение β—коэффициентов, найденные путем решения системы уравнений с помощью частных r и $\sigma^2_{\text{ост}}$, должны быть одинаковыми.

Так, β_2 в уравнениях (III.11) могут быть определены из формулы:

$$\beta_2 = r_{12.34} \frac{\sqrt{\sigma^2_{\text{ост } 1}}}{\sqrt{\sigma^2_{\text{ост } 2}}} = \frac{\sqrt{\left(1 - r_{12}^2\right)\left(1 - r_{13.2}^2\right)\left(1 - r_{14.23}^2\right)}}{\sqrt{\left(1 - r_{23}^2\right)\left(1 - r_{24.3}^2\right)\left(1 - r_{21.34}^2\right)}}\, r_{12.34}; \tag{V.4}$$

$$\beta_2 = 0,153 \frac{\sqrt{(1 - 0,839^2)(1 - 0,556^2)(1 - 0,254^2)}}{\sqrt{(1 - 0,895^2)(1 - 0,268^2)(1 - 0,153^2)}} = 0,159$$

и далее $b_2 = 0,159 \dfrac{101,6}{634} = 0,0255.$

Значения $\sqrt{1 - r^2}$ приведены в приложении 4.

Для определения b_3 делаем следующие расчеты:

$$b_3 = \frac{\sigma_1}{\sigma_3} r_{13.24} \frac{\sqrt{\left(1 - r_{13}^2\right)\left(1 - r_{14.3}^2\right)\left(1 - r_{12.34}^2\right)}}{\sqrt{\left(1 - r_{34}^2\right)\left(1 - r_{23.4}^2\right)\left(1 - r_{13.24}^2\right)}} =$$
$$= \frac{101,6}{230,4} 0,429 \frac{\sqrt{(1 - 0,886^2)(1 - 0,296^2)(1 - 0,153^2)}}{\sqrt{(1 - 0,869^2)(1 - 0,621^2)(1 - 0,429^2)}} = 0,236.$$

Метод дает наилучшие результаты при использовании ЭВМ, так как в этом случае можно исследовать влияние неограниченного количества факторов. Постепенно вводя новые факторы, анализируют поведение $\sigma^2_{\text{ост}}$ (или совокупного коэффициента корреляции R). Если при дальнейшем увеличении числа факторов не происходит существенного снижения $\sigma^2_{\text{ост}}$(либо повышения R), то очевидно, введение их нецелесообразно. Особенно эффективен такой предварительный анализ в технических исследованиях, в которых определение значений какого-либо фактора связано порой со значительными затратами (установкой датчиков, систем предоставления информации и т. д.).

В экономике строительства также нежелательно усложнять анализ рассмотрением факторов, которые существенно не меняют результаты исследования.

Таким образом, в дополнение к измерению влияния на признак всех комбинированных независимых переменных определение показателей частного воздействия каждой переменной (каждого фактора) может оказаться весьма действенным средством проникновения в сущность экономических явлений.

В заключение отметим, что некоторые авторы, в частности Ф. Миллс [12], для определения частных коэффициентов корреляции рекомендуют матричный метод. Все возможные ${}^{0}r$ сводятся в матрицу, имеющую i строк и k столбцов Так как $r_{ik} = r_{ki}$, то матрица симметрична:

$$M = \overbrace{\begin{vmatrix} 1 & r_{12} & r_{13} & \dots & r_{1m} \\ r_{21} & 1 & r_{23} & \dots & r_{2m} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ r_{m1} & r_{m2} & r_{m3} & \dots & 1 \end{vmatrix}}^{k} \left.\vphantom{\begin{matrix}1\\1\\1\\1\end{matrix}}\right\} i. \qquad (V.5)$$

Тогда

$${}^{n}r = \frac{-D_{ik}}{\sqrt{D_{ii} D_{kk}}}, \qquad (V.6)$$

где D_{ik} — определитель, образованный вычеркиванием i-й строки и k-го столбца:

D_{ii}, D_{kk} — определители, образованные вычеркиванием i-й строки и i-го столбца, k-й строки и k-го столбца соответственно.

Простота формулы (V.6) иллюзорна; покажем это на числовом примере. Пусть $r_{12} = -0,632$; $r_{13} = -0,435$; $r_{23} = 0,390$ и исходная матрица имеет вид:

$$\begin{array}{cc|ccc|} & & \multicolumn{3}{c}{\overbrace{\qquad\qquad\qquad\qquad}^{k}} \\ & & 1 & 2 & 3 \\ & 1 & 1 & -0,632 & -0,439 \\ i & 2 & -0,632 & 1 & 0,390 \\ & 3 & -0,439 & 0,390 & 1 \end{array}, \text{ а } r_{12.3} = \frac{-D_{12}}{\sqrt{D_{11} D_{22}}}.$$

Определитель $D_{ik} = D_{12}$ находим, пользуясь правилом Саррюса:

$$D_{12} = 1 + 2(-0{,}632)\cdot 0{,}390\cdot(-0{,}439) - (-0{,}439)\cdot 1 \times \times(-0{,}439) - 0{,}390^2 - 0{,}632^2 = 1 + 0{,}108\cdot 2 - -0{,}193 - 0{,}152 - 0{,}400 = 0{,}471.$$

$D_{ii} = D_{11}$ находим, вычеркивая 1-ю строку и 1-й столбец и решая матрицу:

$$D_{11} = \begin{vmatrix} 1 & 0{,}390 \\ 0{,}390 & 1 \end{vmatrix} = 1 - 0{,}390^2 = 0{,}848.$$

$D_{kk} = D_{22}$ находим, вычеркивая 2-ю строку и 2-й столбец и решая матрицу:

$$D_{22} = \begin{vmatrix} 1 & -0{,}439 \\ -0{,}439 & 1 \end{vmatrix} = 1 - 0{,}439^2 = 0{,}807;$$

$$r_{12.3} = \frac{-0{,}471}{0{,}848\cdot 0{,}807} = -0{,}568.$$

Для исчисления одного частного коэффициента второго порядка объем вычислительной работы возрастает более чем в четыре раза. Поэтому мы не считаем возможным рекомендовать матричный метод для определения частных коэффициентов корреляции в условиях выполнения вычислений вручную.

§ 2. Исследование корреляционных связей способом обратного решения

В практике корреляционного анализа ошибки исчисления в тех или иных промежуточных расчетах, накапливаясь, могут привести к искажению оценок реально существующих связей между исследуемыми переменными. В результате могут последовать выводы о неприемлемости самого корреляционного анализа. Поэтому были предложены методы направленные на устранение возможных ошибок вычислений, упрощение вычислительных операций, введение элементов автоматики в технику корреляционного анализа, выполняемого без использования вычислительных машин.

В зарубежной практике получил применение автоматический контроль решения нормальных уравнений по способу «обратного решения», известного также как метод Дулитла. Простота применения этого метода, не снижающая, однако, строгости его математического построения, делает целесообразной рекомендацию метода Дулитла нашему читателю.

В качестве примера используем приведенные ранее данные о зависимости уровня накладных расходов в жилищном строительстве (x_1) от объема работ (x_2), численности рабочих (x_3) и фонда заработной платы (x_4).

Построение и решение системы нормальных уравнений при постоянном самоконтроле выполняется с помощью специальных таблиц (см. приложение 5).

В табл. 1 приложения 5 данные первых четырех граф взяты из табл. 19 с уменьшением значений каждой переменной в 100 раз для упрощения вычислительных операций. Графа 5 (Σx) является суммой первых четырех и называется контрольной суммой. Сложение итоговых сумм первых четырех граф, естественно, должно быть равным итоговой сумме:

$$\sum x_2 + \sum x_3 + \sum x_4 = \sum (\sum x)$$

или

$$1056{,}80 + 448{,}50 + 393{,}55 + 184{,}40 = 2083{,}25.$$

Эта простейшая операция является первым шагом контроля с помощью обратного решения.

Затем следуют графы 6—19, в которых каждая переменная умножается на одну из них; такие действия именуются алгебраическими расширениями. Каждому такому расширению справа придается графа контрольной суммы по данной переменной, подобная уже рассмотренной нами графе Σx; таковы графы $x_2\Sigma x$, $x_3\Sigma x$, $x_4\Sigma x$ и $x_1\Sigma x$. Аналогично проводятся и проверки «обратным решением», например:

$$18010{,}44 + 7472{,}60 + 6601{,}42 + 3041{,}93 = 35126{,}39.$$

В табл. 2 приложения 5 изложен способ вычисления средних величин с помощью контрольных сумм, а также вычисление произведений этим сумм на каждую переменную, корректированных отклонениями от средних величин. В первой строке табл. 2 показаны суммы значений каждой переменной, включая контрольную сумму. Деление их на число наблюдений (в данном случае $n = 74$) дает среднюю величину для каждой переменной, записанную во второй строке табл. 2. При этом равенство сумм в первых четырех графах суммы графы Σx контролирует деление:

$$14{,}28 + 6{,}06 + 5{,}32 + 2{,}49 = 2083{,}25 : 74 = 28{,}15.$$

В третью строку табл. 2 записываются итоги расширений с x_2 из табл. 1. Затем суммы значений каждой переменной (строка 1 табл. 2) умножаются на средние значения $\bar{x}_2$ (14,28108), а произведения записываются в соответствующие графы строки 4 табл. 2. Значения строки 5 табл. 2 получаем, вычитая из строки 3 строку 4. Полученные величины являются алгебраическими расширениями, выраженными в виде отклонений от средних значений $\bar{x}_2$. Суммы отклонений от средних значений $\bar{x}_3$, $\bar{x}_4$, $\bar{x}_1$ представлены соответственно в строках 8, 11, 14. Значения строк 5, 8 и 11 табл. 2 используются для определения коэффициентов регрессии системы уравнений:

$$\begin{cases} \sum (x_2)^2 b_{12.34} + \sum (x_2x_3) b_{13.24} + \sum (x_2x_4) b_{14.23} = \sum (x_1x_2); \\ \sum (x_2x_3) b_{12.34} + \sum (x_3)^2 b_{13.24} + \sum (x_3x_4) b_{14.23} = \sum (x_1x_3); \\ \sum (x_2x_4) b_{12.34} + \sum (x_3x_4) b_{13.24} + \sum (x_4)^2 b_{14.23} = \sum (x_1x_4). \end{cases}$$

Решение системы уравнений с использованием метода обратного решения и контрольной суммы приведено в табл. 3. Метод обратного решения есть последовательное определение коэффициентов переменных множественной регрессии b_4, b_3 и b_2 по данным строк 11, 6 и 2 табл. 3.

Так по строке 11 находим коэффициент b_4 (при x_4), как частное $\frac{x_1}{x_4}$ с обратным знаком. Значения x_1 и x_4 принимаются по строке 10, табл. 3; $x_1 = 5{,}69$; $x_4 = 59{,}86$;

$$b_4 = -\frac{5{,}69}{59{,}86} = -0{,}09505.$$

Подобным образом по данным строки 6 табл. 3 определяем коэффициент b_3.

$$-b_3 + (0{,}47490)\, b_4 = x_1;$$
$$-b_3 + (-0{,}47490)\,(0{,}09505) = -0{,}32429,$$

откуда $b_3 = 0{,}32429 - 0{,}47490 \cdot 0{,}09505 = 0{,}2792$.

Коэффициент b_2 находим по показателям строки 2 табл. 3.

$$-b_2 + (-0{,}36582)\, b_3 + (-0{,}33620)\, b_4 = x_1;$$
$$-b_2 + (-0{,}36582) \cdot 0{,}2792 + (-0{,}33620) \times$$
$$\times 0{,}09505 = -0{,}13998, \text{ откуда } b_2 = 0{,}00586.$$

Проверим расчет путем подстановки полученных значений коэффициентов регрессии (b) в уравнение, учитывающее значения переменных, соответствующие каждому коэффициенту регрессии.

Так, для определения x_1 пользуемся следующим уравнением системы нормальных уравнений:

$$\sum (x_2x_4)\, b_2 + \sum (x_3x_4)\, b_3 + \sum (x_4)^2\, b_4 = \sum (x_1x_4).$$

В это уравнение подставляем значения строки 7 табл. 3. $(981,1)\cdot(0,00586) + (381,14)\cdot(0,2792) + (400,26)\cdot(0,09505) = 150,24$, что соответствует значению x_1, полученному из расчета отклонений от средних величин табл. 2 строки 11.

Значение свободного члена уравнения регрессии a определяем по формуле:

$$a_{1.234} = \bar{x}_1 - b_2\bar{x}_2 - b_3\bar{x}_3 - b_4\bar{x}_4 = 2,49189 - 0,00586 \times \times 14,281 - 0,2792\cdot 6,06081 + 0,09505\cdot 5,31842 = 0,211.$$

Таким образом, уравнение $\bar{x}_1 = f(x_2, x_3, x_4)$ при увеличении в 100 раз расчетных значений для перехода к реальным величинам имеет вид:

$$\bar{x}_1 = 21,1 + 0,00586x_2 + 0,2792x_3 + 0,09505x_4.$$

Это уравнение почти полностью соответствует уравнению, полученному с помощью метода наименьших квадратов (см. стр. 121).

Определим коэффициент совокупной корреляции:

$$R^2_{1.234} = \frac{b_2\left(\sum x_1 x_2\right) + b_3\left(\sum x_1 x_3\right) + b_4\left(\sum x_1 x_4\right)}{\sum (x_1)^2}.$$

Величины, соответствующие коэффициентам неизвестных $(\Sigma x_1 x_2)$, $(\Sigma x_1 x_3)$, $(\Sigma x_1 x_4)$ принимаем из табл. 2, графа x_1, строки 5, 8 и 11. Величину $\Sigma (x_1)^2$ принимаем по той же таблице, графа x_1, строка 14.

$$R^2_{1.234} = \frac{(0,00586)(408,50) + (0,2792)(164,62) + (0,09505)(150,24)}{70} = 0,8948.$$

Таблицы, приведенные в приложении 5, удобны для расчетов. По предлагаемой методике весьма просто определяются не только коэффициенты множественной регрессии (R) и, следовательно, уравнения множественной регрессии, но и частные коэффициенты корреляции (через значения коэффициентов совокупной корреляции).

Для этого в отдельных таблицах 4, 5, 6 определяем коэффициенты совокупной корреляции между зависимой переменной (x_1) и независимыми переменными (x_2, x_3, x_4) при последовательном исключении отдельных независимых переменных из каждого их ряда, т. е. находим величины $R^2_{1.23}$, $R^2_{1.24}$, $R^2_{1.34}$. Все расчеты, выполненные в табл. 3, не повторяются заново, так как отдельные графы можно перестраивать в необходимом порядке с незначительными дополнениями расчетами для поочередного исключения каждой независимой переменной.

Проследим методику образования, например, $R^2_{1.23}$ (см. табл. 4). В первой части таблицы представлена вся исходная информация, которая используется для определения $R^2_{1.23}$ методом обратного решения. Во второй части табл. 4 иллюстрируем образование коэффициентов $b_{12.3}$ и $b_{13.2}$. В третьей части табл. 4 рассчитывается числитель для определения $R^2_{1.23}$.

В таком же порядке в табл. 5 представлено определение коэффициента $R^2_{1.24}$.

Для определения коэффициента $R^2_{1.34}$ необходимо произвести полностью новые прямые решения при исключении x_2.

Эта операция изложена в табл. 6.

Общая формула для определения частных коэффициентов корреляции

$$R_{1.n\ldots23(n-1)} = 1 - \sqrt{\frac{1 - R^2_{1.23\ldots n}}{1 - R^2_{1.23\ldots(n-1)}}}.$$

Предварительно находим выражения $1 - R^2$ по вспомогательной табл. 55.

Таблица 55

Определение выражений $1 - R^2$

Переменные	R^2	$1 - R^2$
1.234	0,894800	0,105200
1.23	0,887244	0,112756
1.24	0,873008	0,126992
1.34	0,892214	0,107786

Тогда

$$r_{12.34} = \sqrt{1 - \frac{0,105200}{0,107786}} = 0,153;$$

$$r_{13.24} = \sqrt{1 - \frac{0,105200}{0,126992}} = 0,429;$$

$$r_{14.23} = \sqrt{1 - \frac{0,105200}{0,112756}} = 0,254.$$

Метод обратных решений особенно перспективен при применении ЭВМ, так как действия, которые он предусматривает, легко формализуются и программируются. Этот метод может быть рекомендован для тех случаев, когда известна, хотя бы в общих чертах, значимость действия факторов на исследуемый показатель, т. е. когда нет необходимости вычисления парных коэффициентов корреляции.

Выводы. Применение корреляционного анализа предполагает следующий порядок операций: отбор факторов, которые существенно влияют на изучаемый экономический процесс, состав этих факторов определит объем и программу сбора отчетных и иных исходных данных; построение рядов распределения, вычисление средних ошибок показателей колеблемости, определение критериев согласия; исследование парных корреляционных связей между выбранным признаком и всеми факторами, а также между самими факторами; построение уравнений множественных корреляционных зависимостей; оценка аппроксимирующей способности уравнений множественных корреляционных зависимостей (так называемых моделей).

Этого достаточно для решения практических задач в строительных организациях.

В научно-исследовательские работы рекомендуется включать определение также и частных коэффициентов корреляции. В книге приведены методы корреляции, примененные в научных работах институтов Госстроя СССР и Госстроя УССР. Перечисление их свидетельствует о широте диапазона задач, в которых был рационально применен метод корреляции: разработка нормативов накладных расходов, установление мощности строительных организаций, определение величины задела, выявление резервов роста производительности труда и т. д.

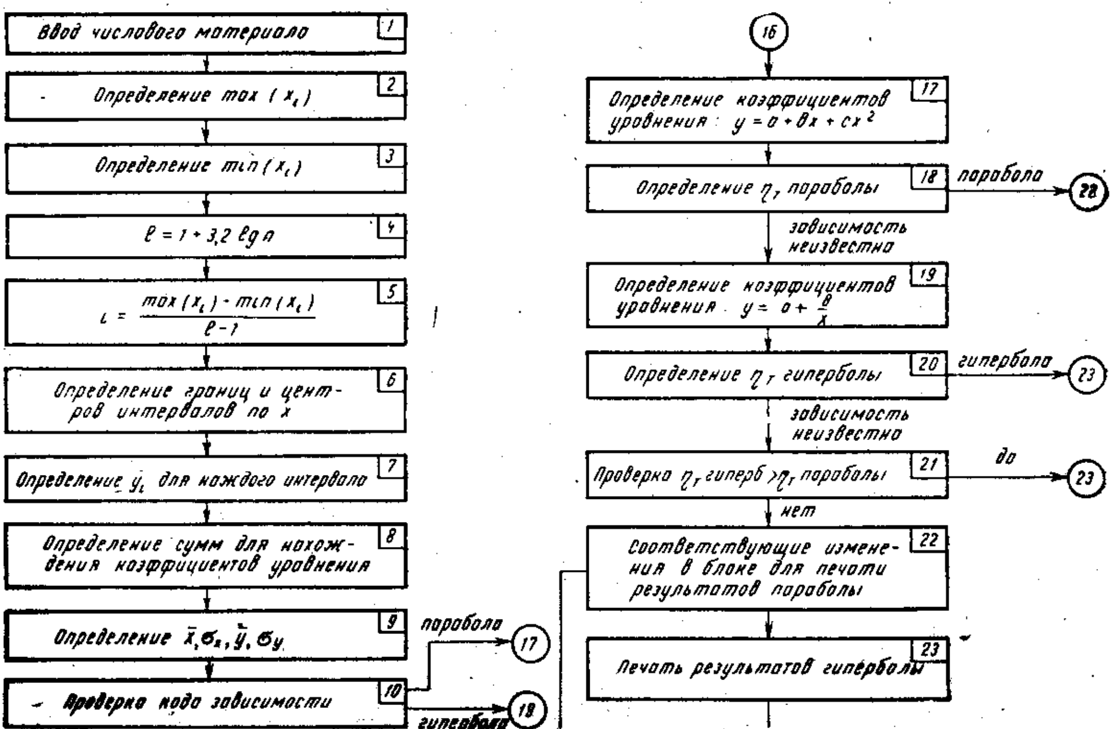
Широкие возможности имеются для применения этого метода в экономической работе проектных институтов.

Приложение 1

Данные по строительно-монтажным управлениям, ведущим строительство предприятий машиностроения

№ п/п	Шифр СУ	Фактическая себестоимость (в % к сметной)	Выработка	R объема	R рабочей силы	Сборность (%)	Уровень специализации (%)	Объем работ собственными силами за год (тыс. руб.)
1	2	3	4	5	6	7	8	9
1	864 261	96,3	4,89	89	87	73,1	52,7	2 693
2	864 262	97,8	4,81	93,7	86,5	70,2	59,3	2 341
3	864 271	103,7	4,17	86	92,6	38,6	35,4	2 177
4	864 272	99,3	4,50	92,1	86,0	51,7	41,7	1 839
5	864 301	108,0	3,16	89,5	99,0	35,8	51,1	1 660
6	864 302	96,5	3,84	91	89,7	36,1	54,0	1 421
7	864 351	97,5	3,88	83,5	92,0	37,8	57,4	1 117
8	864 352	95,7	4,54	93,4	89,0	24,8	42,3	1 004
9	864 481	96,9	4,29	92,5	—	40,6	47,4	1 931
10	864 482	101,0	5,46	91,0	98,0	71,1	49,9	2 035
11	864 601	103,3	3,89	78,1	98,0	21,2	45,2	1 802
12	864 602	101,7	3,96	90,6	91,0	32,6	47,5	1 931
13	894 450	94,0	3,24	91,8	—	44,4	54,4	2 220
14	894 451	93,1	3,31	92	—	64,7	48,7	1 893
15	894 452	92,7	3,36	85,9	99,0	39,5	51,1	1 650
16	894 470	105,5	2,63	90,5	—	32,1	43,7	1 572
17	894 471	103,0	2,34	91,0	—	31,7	53,6	989
18	894 472	98,1	2,75	89,0	98,5	30,2	38,8	878
19	894 610	110,0	3,31	87,4	—	23,4	—	1 235
20	894 611	109,7	2,70	86,1	—	22,7	49,0	1 874
21	894 612	119,5	2,68	99,3	91,0	21,2	—	836
22	894 680	127,0	2,57	67,1	—	20,4	—	—
23	894 681	137,5	1,98	63,9	—	15,7	32,0	—
24	894 682	107,5	2,61	83,0	99,1	36,5	—	1 448
25	874 392	94,8	3,20	92,0	—	48,7	47,5	1 655
26	874 652	94,0	3,34	90,0	—	36,6	36,4	1 905
27	874 341	95,5	—	98,5	—	42,8	48,6	2 854
28	874 342	98,7	3,85	85,5	—	38,1	48,4	3 100
29	874 321	87,3	3,76	87,6	—	36,4	45,9	2 310
30	874 322	91,5	3,97	93,1	—	40,0	43,0	2 321
31	874 331	116,1	3,56	93,3	—	34,2	56,8	2 290
32	874 332	93,3	3,47	91,3	—	33,8	63,3	2 101
33	ДПС3СО	87,0	4,78	95,0	99,1	64,5	50,8	2 440
34	ДПС3С1	98,2	4,76	91	95	40,5	55,5	2 071
35	ДПС3С2	95,1	5,41	87	94	40	47,5	2 089
36	ДПСМШ2	102,0	2,87	94	98	22,4	68,5	943
37	ДПСГ2	115,0	2,03	77	71,7	31,0	48,0	804
38	ДТПС3СО	85,0	3,51	92	94	45	47	1 654
39	ДТПС3С1	90,0	3,42	87	98	36	50,5	1 306
40	ДТПС3С2	90,0	3,29	89	95	36	75	1 274
41	ДТПСНМ0	95,8	3,60	85,2	91	25	41,5	1 441
42	ДТПСНМ1	97,5	3,65	79,5	87	31	62	1 714
43	ДТПСНМ2	100,1	3,43	89,5	97	45	56,5	1 271
44	ДТПСПП0	91,5	3,30	86,0	96	28	40,5	1 306
45	ДТПСПП1	92,4	3,45	86,0	95	20,6	35	1 028
46	ДТПСПП2	105,0	3,05	97,0	98	31,5	51,5	986
47	ДСПСО	101,5	3,66	98,1	87	493	—	—
48	ДСПС1	97,3	3,86	77,1	99,0	46,7	—	—
49	ДСПС2	95,2	3,46	85,3	94	41,3	—	—
50	6 260	84,3	4,14	97,3	—	75	36,9	1 667
51	6 261	95,8	3,45	87	—	50,5	48,2	1 243

№ п/п	Шифр СУ	Фактическая себестоимость (в % к сметной)	Выработка	R объема	R рабочей силы	Сборность (%)	Уровень специализации (%)	Объем работ своими силами за год (тыс. руб.)
1	2	3	4	5	6	7	8	9
52	6 262	109,7	3,03	89,6	—	50,9	45,7	1 185
53	6 271	98,7	3,27	98,1	—	35,7	58,3	1 300
54	6 272	102	3,01	83,7	—	39,4	51,1	—
55	6 281	106,2	3,57	82	—	51,1	49,5	1 062
56	6 282	96,9	3,75	88	—	19,2	42,8	1 000
57	1 710	91	4,11	88,6	93	62,3	53	1 500
58	1 710	90,6	4,5	92	93,5	27,4	58	1 400
59	1 712	91	4,7	96	80,1	62,7	51,5	1 368
60	1 720	92	3,21	86	83	65,1	73	2 434
61	1 721	96,6	3,56	92	87	35	72	1 180
62	1 722	99	4,15	56,3	59,3	51,3	57	1 249
63	1 730	83	4,1	91	83,6	65	70	2 022
64	1 731	85,4	4,58	95	96	16,5	60,5	2 165
65	1 732	85,1	4,82	90	86,5	—	48	2 359
66	1 740	88,3	4,64	87	95	55,4	41	2 083
67	1 741	91,5	4,63	87	98,5	—	34	2 169
68	1 742	88	5,38	98	91	59,1	30,5	2 095
69	17 121	101	3,89	80	88	37,1	49,5	1 715
70	17 122	98	3,96	86,5	92	34,6	58	2 104
71	1 770	86,6	4,22	81	80,7	47,5	40,2	2 053
72	1 771	95	3,93	93	97	38	46,6	2 011
73	1 772	92,5	4,45	88,5	96,5	36,6	51,3	1 868
74	1 790	114,1	6,19	66	79,8	55	48,5	1 936
75	1 791	84	5,97	94	95	64	37	1 720
76	1 792	89	5,39	99,1	98,7	62	—	1 770
77	ДПСПС0	98	3,9	94	98,7	29,1	45,3	1 650
78	ДПСПС1	101	3,79	99,2	98	55,5	28,8	1 471
79	ДПСПС2	103	4,08	90	98,5	22,4	46,6	1 404
80	3 301	89,6	3,34	84	—	61,0	37,4	862
81	3 302	105,5	3,41	86	—	20	33,1	633
82	3 311	100,1	4,06	81	—	18,1	55,1	1 698
83	3 312	122,5	4,08	63	—	30,3	46,0	1 846
84	3 321	96,6	4,10	85,5	—	37,1	38,8	2 104
85	3 322	102,5	4,30	94	—	42,3	42,4	2 358
86	3 341	91,0	3,36	83	—	57,7	38,9	2 885
87	3 442	95,0	3,51	85	—	31,8	30,7	3 097
88	3 361	105,7	3,73	88,5	—	32,5	30,2	1 822
89	3 362	110,5	3,80	92	—	33,1	39,6	2 505
90	3 110	86,9	4,47	90	—	52,7	66	1 733
91	3 111	89,2	5,13	91	—	36,3	52,3	2 196
92	3 112	89,3	5,56	93	—	33,9	64,7	2 390
93	3 120	88,4	4,55	92	—	39,4	61,2	1 790
94	3 121	88,7	4,28	96	—	40,1	61,7	2 026
95	3 122	91,3	4,58	98	—	46,5	65,6	1 791
96	3 130	103,6	5,78	71	—	29,1	37,4	3 297
97	3 131	90,7	5,88	96,1	—	28,1	42,2	3 026
98	3 132	89,9	6,03	91,4	—	42,0	47,7	3 381
99	3 140	95,9	3,29	87	—	39,7	77,2	3 013
100	3 141	91,3	3,74	96	—	24,4	54,3	3 151
101	3 142	89,0	4,46	97	—	53,6	38,0	3 336
102	3 391	118,0	2,77	54,1	—	23,6	51,1	863
103	3 392	115,5	3,22	60,0	—	12,6	61,5	1 174
104	3 341	91,0	3,36	89	—	57,7	38,9	2 885


Ввод числового материала 1
Определение max (x_i) 2
Определение min (x_i) 3
ℓ = 1 + 3,2 lg n 4
i = (max (x_i) - min (x_i)) / (ℓ - 1) 5
Определение границ и центров интервалов по x 6
Определение y_i для каждого интервала 7
Определение сумм для нахождения коэффициентов уравнения 8
Определение x̄, σ_x, ȳ, σ_y 9
Проверка вида зависимости 10
парабола
17
гипербола
18
16
Определение коэффициентов уравнения: y = a + bx + cx² 17
Определение η_т параболы 18
парабола
22
зависимость неизвестна
Определение коэффициентов уравнения y = a + b/x 19
Определение η_т гиперболы 20
гипербола
23
зависимость неизвестна
Проверка η_т гиперб > η_т параболы 21
да
23
нет
Соответствующие изменения в блоке для печати результатов параболы 22
Печать результатов гиперболы 23

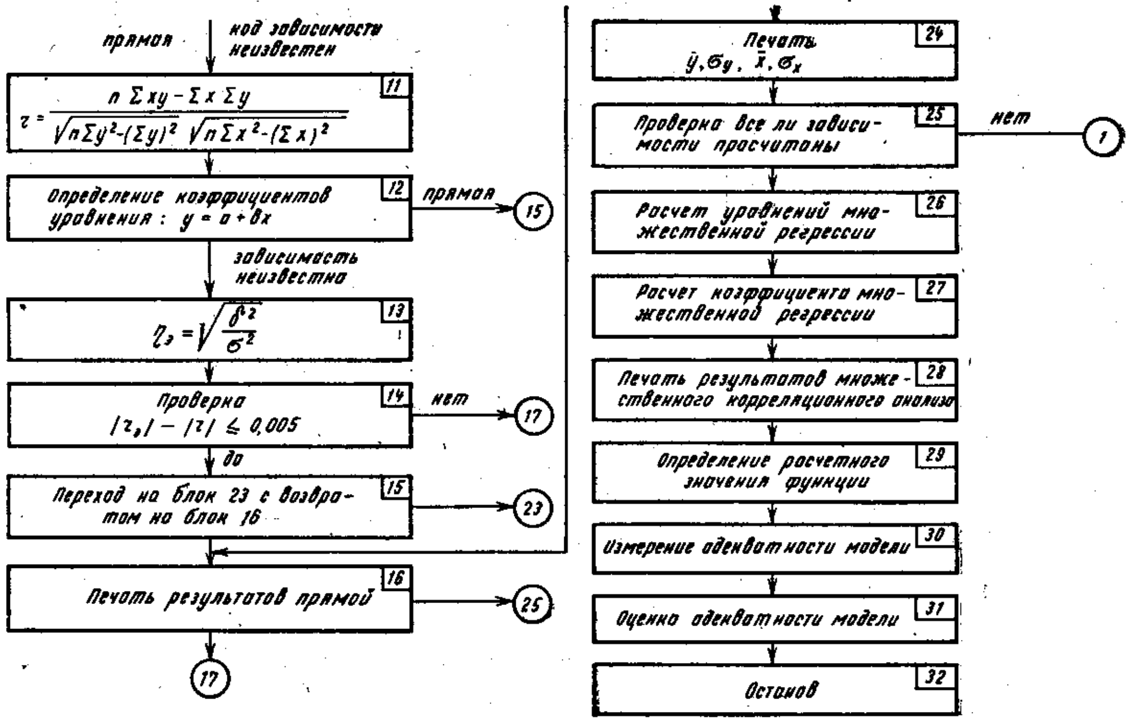


Рис. 30. Блок-схема программы корреляционного анализа для БЭСМ-3.

Программа состоит из двух частей. Первая часть позволяет осуществить парный, а вторая — множественный корреляционный анализ.

Первая часть данной программы предполагает по исходным данным (ряд значений независимой величины — x_i и ряд соответствующих значений зависимой величины — y_i) определить вид корреляционной зависимости, тесноту связи, коэффициенты уравнения связи. Рассматриваются зависимости: прямая, парабола 2-го порядка и гипербола. Если заранее известен вид зависимости, то, записав соответствующий код по определенному адресу (в ячейку 1076), заставляем работать программу по нужному участку и на печать выдается результат. Если же вид зависимости неизвестен, то в ячейку (1076) ничего не записывается, а программа работает по перебору всех видов зависимостей (прямой, параболы, гиперболы). Программа выдает на печать:

r — коэффициент корреляции или η_T в случае криволинейной зависимости;

$\bar{y}$ — среднее значение признака;

σ_y — среднеквадратическое отклонение по y;

$\bar{x}_i$ — среднее значение фактора;

σx_i — среднеквадратическое отклонение по x.

Первая часть программы состоит из двух разделов: основной программы, собственно программы определения корреляционной зависимости, которая занимает 0230 — 1107 ячейки МОЗУ (она используется для просчета только одного варианта); и управляющей программы, занимающей 0200—0216 ячейки МОЗУ.

Если нужно просчитать более одного варианта, употребляются обе программы. Управляющая программа ставится первой. Она осуществляет запись основной программы на магнитный барабан, затем после просчета каждого варианта очистку в МОЗУ места для программы и запись самой программы (так как после просчета программа вновь использоваться не может).

В *программе используются следующие стандартные подпрограммы*: перевода $10 \to 2$; $2 \to 10$ с выдачей результата на печать; определения $\lg n$; обращения матрицы; умножения матрицы на вектор.

Инструкция к I части программы

Числовой материал записывается следующим образом:

№ ячейки	Содержимое
1075	№ варианта — (десятичное число)
1076	Код зависимости[1]
1077	n — число точек (десятичное число)
$1100-1100+(n-1)$	Массив для y_i
$4100-4100+(n-1)$	Массив для x_i

[1] Коды зависимости имеют вид.

1076	0000	0000	0001	Прямая
1076	0000	0000	0002	Парабола
1076	0000	0000	0000	Гипербола
1076	0	0	0	Вид зависимости неизвестен

Если используется только основная программа без управляющей, то пуск нужно делать с ячейки 0230. Если же используются обе программы, пуск делается с ячейки 0200.

На печать выдается:

N просчитываемого варианта;

код предполагаемой зависимости;

код зависимости, которую машина выдала на печать;

коэффициент корреляции r (или η_T);

коэффициенты уравнения данного вида зависимости:

$$\bar{y},\ \sigma_y,\ \bar{x},\ \sigma_x.$$

Инструкция ко II части программы

Числовой материал записывается следующим образом:

Ячейки	Содержание
0011	Число факторов n_{10}
0012—0035	Матрица—столбец свободных членов уравнения—парные коэффициенты корреляции признака с факторами
0037—0656	Матрица парных коэффициентов корреляции между факторами
0660	$\bar{y}$—среднее значение признака
0661—0704	$\bar{x}_i$ —средние значения факторов
0705	σ_y—среднеквадратичное отклонение признака
0706—0731	σ_{x_i}—среднеквадратические отклонения факторов

После ввода программы сделать пуск с ячейки 1200.

Порядок печати результатов II раздела программы:

β_i — стандартизованный коэффициент;

R — совокупный коэффициент корреляции;

b_0 — свободный член уравнения;

$b_1,\ b_2\ \ldots\ b_n$ — коэффициенты при соответствующих факторах;

$\hat{y}_{i\ \text{расч}}$ — по каждому СУ;

$|y_{i\ \text{факт}} - \hat{y}_{i\ \text{расч}}|$ — абсолютная погрешность по каждому СУ;

$\dfrac{y_{i\ \text{ф}} - \hat{y}_{i\text{р}}}{y_{i\text{ф}}}$ — относительная погрешность по каждому СУ (%);

$$\varepsilon = \frac{1}{n}\sum\left|\frac{y_\text{ф} - \hat{y}_\text{ф}}{y_\text{ф}}\right| 100\% \text{ — ошибка аппроксимации.}$$

Приложение 8

Вычисление частных коэффициентов корреляции

а) первого порядка

Формула частного коэффициента корреляции — r	r нулевого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r первого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{13.2} = \frac{r_{13} - r_{12} \cdot r_{23}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{23}^2}}$	13 12 23						$r_{13.2}$	
$r_{14.2} = \frac{r_{14} - r_{12} \cdot r_{24}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{24}^2}}$	14 12 24						$r_{14.2}$	
$r_{15.2} = \frac{r_{15} - r_{12} \cdot r_{25}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{25}^2}}$	15 12 25						$r_{15.2}$	
$r_{16.2} = \frac{r_{16} - r_{12} \cdot r_{26}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{26}^2}}$	16 12 26						$r_{16.2}$	
$r_{34.2} = \frac{r_{34} - r_{23} \cdot r_{24}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{24}^2}}$	34 23 24						$r_{34.2}$	

$r_{35.2}=\dfrac{r_{35}-r_{23}\cdot r_{25}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{25}^2}}$	35 23 25						$r_{35.2}$	
$r_{36.5}=\dfrac{r_{36}-r_{23}\cdot r_{26}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{26}^2}}$	36 23 26						$r_{36.2}$	
$r_{45.2}=\dfrac{r_{45}-r_{24}\cdot r_{25}}{\sqrt{1-r_{24}^2}\sqrt{1-r_{25}^2}}$	45 24 25						$r_{45.2}$	
$r_{46.2}=\dfrac{r_{46}-r_{24}\cdot r_{26}}{\sqrt{1-r_{24}^2}\sqrt{1-r_{26}^2}}$	46 24 26						$r_{46.2}$	
$r_{56.2}=\dfrac{r_{56}-r_{25}\cdot r_{26}}{\sqrt{1-r_{25}^2}\sqrt{1-r_{26}^2}}$	56 25 26						$r_{56.2}$	
$r_{12.3}=\dfrac{r_{12}-r_{13}\cdot r_{23}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{23}^2}}$	12 13 23	—0,850 —0,758 0,622	0,652 0,783	—0,472	—0,378	0,511	$r_{12.3}$	—0,740
$r_{14.3}=\dfrac{r_{14}-r_{13}\cdot r_{34}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{34}^2}}$	14 13 34	—0,610 —0,758 0 229	0,652 0,973	—0,174	—0,436	0,635	$r_{14.3}$	—0,587

Продолжение

Формула частного коэффициента корреляции — r	r нулевого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r первого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{15.3}=\frac{r_{15}-r_{13}\cdot r_{35}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{35}^2}}$	15 13 35	—0,779 —0,758 0,455	0,652 0,891	—0,345	—0,434	0,589	$r_{15.3}$	—0,738
$r_{16.3}=\frac{r_{16}-r_{13}\cdot r_{36}}{\sqrt{1-r_{13}^2}\sqrt{1-r_{36}^2}}$	16 13 36	0,836 —0,758 —0,784	0,652 0,621	0,594	0,242	0,405	$r_{16.3}$	0,598
$r_{24.3}=\frac{r_{24}-r_{23}\cdot r_{34}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{34}^2}}$	24 23 34	0,557 0,622 0,229	0,783 0,973	0,142	0,415	0,761	$r_{24.3}$	0,545
$r_{25.3}=\frac{r_{25}-r_{23}\cdot r_{35}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{35}^2}}$	25 23 35	0,681 0,622 0,455	0,783 0,891	0,283	0,398	0,698	$r_{25.3}$	0,571
$r_{26.3}=\frac{r_{26}-r_{23}\cdot r_{36}}{\sqrt{1-r_{23}^2}\sqrt{1-r_{36}^2}}$	26 23 36	—0,823 0,622 —0,784	0,783 0,621	—0,487	—0,336	0,487	$r_{26.3}$	—0,690

$r_{45.3}=\frac{r_{45}-r_{34}\cdot r_{35}}{\sqrt{1-r_{34}^2}\sqrt{1-r_{35}^2}}$	45 34 35	0,619 0,229 0,455	0,973 0,891	0,104	0,515	0,867	$r_{45.3}$	0,594
$r_{46.3}=\frac{r_{46}-r_{34}\cdot r_{36}}{\sqrt{1-r_{34}^2}\sqrt{1-r_{36}^2}}$	46 34 36	—0,467 0,229 —0,784	0,973 0,621	—0,179	—0,288	0,604	$r_{46.3}$	—0,477
$r_{56.3}=\frac{r_{56}-r_{35}\cdot r_{36}}{\sqrt{1-r_{35}^2}\sqrt{1-r_{36}^2}}$	56 35 36	—0,639 0,455 —0,784	0,891 0,621	—0,357	—0,282	0,554	$r_{56.3}$	—0,510
$r_{12.4}=\frac{r_{12}-r_{14}\cdot r_{24}}{\sqrt{1-r_{14}^2}\sqrt{1-r_{24}^2}}$	12 14 24						$r_{12.4}$	
$r_{13.4}=\frac{r_{13}-r_{14}\cdot r_{34}}{\sqrt{1-r_{14}^2}\sqrt{1-r_{34}^2}}$	13 14 34						$r_{13.4}$	
$r_{15.4}=\frac{r_{15}-r_{14}\cdot r_{45}}{\sqrt{1-r_{14}^2}\sqrt{1-r_{45}^2}}$	15 14 45						$r_{15.4}$	
$r_{16.4}=\frac{r_{16}-r_{14}\cdot r_{46}}{\sqrt{1-r_{14}^2}\sqrt{1-r_{46}^2}}$	16 14 46						$r_{16.4}$	

Продолжение

Формула частного коэффициента корреляции — r	r нулевого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r первого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{23.4}=\dfrac{r_{23}-r_{24}\cdot r_{34}}{\sqrt{1-r_{24}^2}\sqrt{1-r_{34}^2}}$	23 24 34						$r_{23.4}$	
$r_{25.4}=\dfrac{r_{25}-r_{24}\cdot r_{45}}{\sqrt{1-r_{24}^2}\sqrt{1-r_{45}^2}}$	25 24 45						$r_{25.4}$	
$r_{26.4}=\dfrac{r_{26}-r_{24}\cdot r_{46}}{\sqrt{1-r_{24}^2}\sqrt{1-r_{46}^2}}$	26 24 46						$r_{26.4}$	
$r_{35.4}=\dfrac{r_{35}-r_{34}\cdot r_{45}}{\sqrt{1-r_{34}^2}\sqrt{1-r_{45}^2}}$	35 34 45						$r_{35.4}$	
$r_{36.4}=\dfrac{r_{36}-r_{34}\cdot r_{46}}{\sqrt{1-r_{34}^2}\sqrt{1-r_{46}^3}}$	36 34 46						$r_{36.4}$	

$r_{56.4}=\dfrac{r_{56}-r_{45}\cdot r_{46}}{\sqrt{1-r_{45}^2}\sqrt{1-r_{46}^3}}$	56 45 46						$r_{56.4}$	
$r_{12.5}=\dfrac{r_{12}-r_{15}\cdot r_{25}}{\sqrt{1-r_{15}^2}\sqrt{1-r_{25}^2}}$	12 15 25						$r_{12.5}$	
$r_{13.5}=\dfrac{r_{13}-r_{15}\cdot r_{35}}{\sqrt{1-r_{15}^2}\sqrt{1-r_{35}^2}}$	13 15 35						$r_{13.5}$	
$r_{14.5}=\dfrac{r_{14}-r_{15}\cdot r_{45}}{\sqrt{1-r_{15}^2}\sqrt{1-r_{45}^2}}$	14 15 45						$r_{14.5}$	
$r_{16.5}=\dfrac{r_{16}-r_{15}\cdot r_{56}}{\sqrt{1-r_{15}^2}\sqrt{1-r_{56}^2}}$	16 15 56						$r_{16.5}$	
$r_{23.5}=\dfrac{r_{23}-r_{25}\cdot r_{35}}{\sqrt{1-r_{25}^2}\sqrt{1-r_{35}^2}}$	23 25 35						$r_{23.5}$	
$r_{24.5}=\dfrac{r_{24}-r_{25}\cdot r_{45}}{\sqrt{1-r_{25}^2}\sqrt{1-r_{45}^2}}$	24 25 45						$r_{24.5}$	

Продолжение

Формула частного коэффициента корреляции — r	r нулевого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac}\times r_{bc}$	Числитель	Знаменатель	r первого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{26.5}=\frac{r_{26}-r_{25}\cdot r_{56}}{\sqrt{1-r_{25}^2}\sqrt{1-r_{56}^2}}$	26 25 56						$r_{26.5}$	
$r_{34.5}=\frac{r_{34}-r_{35}\cdot r_{45}}{\sqrt{1-r_{35}^2}\sqrt{1-r_{45}^2}}$	34 35 45						$r_{34.5}$	
$r_{36.5}=\frac{r_{36}-r_{35}\cdot r_{56}}{\sqrt{1-r_{35}^2}\sqrt{1-r_{56}^2}}$	36 35 56						$r_{36.5}$	
$r_{46.5}=\frac{r_{46}-r_{45}\cdot r_{56}}{\sqrt{1-r_{45}^2}\sqrt{1-r_{56}^2}}$	46 45 56						$r_{46.5}$	
$r_{12.6}=\frac{r_{12}-r_{16}\cdot r_{26}}{\sqrt{1-r_{16}^2}\sqrt{1-r_{26}^2}}$	12 16 26						$r_{12.6}$	

$r_{13.6}=\dfrac{r_{13}-r_{16}\cdot r_{36}}{\sqrt{1-r_{16}^2}\sqrt{1-r_{36}^2}}$	13 16 36						$r_{13.6}$	
$r_{14.6}=\dfrac{r_{14}-r_{16}\cdot r_{46}}{\sqrt{1-r_{16}^2}\sqrt{1-r_{46}^2}}$	14 16 46						$r_{14.6}$	
$r_{15.6}=\dfrac{r_{15}-r_{16}\cdot r_{56}}{\sqrt{1-r_{16}^2}\sqrt{1-r_{56}^2}}$	15 16 56						$r_{15.6}$	
$r_{23.6}=\dfrac{r_{23}-r_{25}\cdot r_{36}}{\sqrt{1-r_{25}^2}\sqrt{1-r_{36}^2}}$	23 25 36						$r_{23.6}$	
$r_{24.6}=\dfrac{r_{24}-r_{26}\cdot r_{46}}{\sqrt{1-r_{26}^2}\sqrt{1-r_{46}^2}}$	24 26 46						$r_{24.6}$	
$r_{25.6}=\dfrac{r_{25}-r_{26}\cdot r_{56}}{\sqrt{1-r_{26}^2}\sqrt{1-r_{56}^2}}$	25 26 56						$r_{25.6}$	
$r_{34.6}=\dfrac{r_{34}-r_{36}\cdot r_{46}}{\sqrt{1-r_{36}^2}\sqrt{1-r_{46}^2}}$	34 36 46						$r_{34.6}$	

Продолжение

Формула частного коэффициента корреляции — r	r нулевого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r первого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{35.6} = \dfrac{r_{35} - r_{36} \cdot r_{56}}{\sqrt{1-r_{36}^2}\sqrt{1-r_{56}^2}}$	35 36 56						$r_{35.6}$	
$r_{45.6} = \dfrac{r_{45} - r_{46} \cdot r_{56}}{\sqrt{1-r_{46}^2}\sqrt{1-r_{56}^2}}$	45 46 56						$r_{45.6}$	

б) второго порядка

Формула частного коэффициента корреляции — r	r первого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r второго порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{14.23} = \dfrac{r_{14.2} - r_{13.2} \cdot r_{34.2}}{\sqrt{1-r_{13.2}^2}\sqrt{1-r_{34.2}^2}}$	14.2 13.2 34.2						$r_{14.23}$	

$r_{15.23}=\dfrac{r_{15.2}-r_{13.2}\cdot r_{35.2}}{\sqrt{1-r_{13.2}^2}\sqrt{1-r_{35.2}^2}}$	15.2 13.2 35.2						$r_{15.23}$	
$r_{16.23}=\dfrac{r_{16.2}-r_{13.2}\cdot r_{36.2}}{\sqrt{1-r_{13.2}^2}\sqrt{1-r_{36.2}^2}}$	16.2 13.2 36.2						$r_{16.23}$	
$r_{13.24}=\dfrac{r_{13.2}-r_{14.2}\cdot r_{34.2}}{\sqrt{1-r_{14.2}^2}\sqrt{1-r_{34.2}^2}}$	13.2 14.2 34.2						$r_{13.24}$	
$r_{15.24}=\dfrac{r_{15.2}-r_{14.2}\cdot r_{45.2}}{\sqrt{1-r_{14.2}^2}\sqrt{1-r_{45.2}^2}}$	15.2 14.2 45.2						$r_{15.24}$	
$r_{16.24}=\dfrac{r_{16.2}-r_{14.2}\cdot r_{46.2}}{\sqrt{1-r_{14.2}^2}\sqrt{1-r_{46.2}^2}}$	16.2 14.2 46.2						$r_{16.24}$	
$r_{13.25}=\dfrac{r_{13.2}-r_{15.2}\cdot r_{35.2}}{\sqrt{1-r_{15.2}^2}\sqrt{1-r_{35.2}^2}}$	13.2 15.2 35.2						$r_{13.25}$	
$r_{14.25}=\dfrac{r_{14.2}-r_{15.2}\cdot r_{45.2}}{\sqrt{1-r_{15.2}^2}\sqrt{1-r_{45.2}^2}}$	14.2 15.2 45.2						$r_{14.25}$	

Продолжение

Формула частного коэффициента корреляции — r	r первого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r второго порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{16.25} = \frac{r_{16.2} - r_{15.2} \cdot r_{56.2}}{\sqrt{1-r_{15.2}^2}\sqrt{1-r_{56.2}^2}}$	16.2 15.2 56.2						$r_{16.25}$	
$r_{13.26} = \frac{r_{13.2} - r_{16.2} \cdot r_{36.2}}{\sqrt{1-r_{16.2}^2}\sqrt{1-r_{36.2}^2}}$	13.2 16.2 36.2						$r_{13.26}$	
$r_{14.26} = \frac{r_{14.2} - r_{16.2} \cdot r_{46.2}}{\sqrt{1-r_{16.2}^2}\sqrt{1-r_{46.2}^2}}$	14.2 16.2 46.2						$r_{14.26}$	
$r_{15.26} = \frac{r_{15.2} - r_{16.2} \cdot r_{56.2}}{\sqrt{1-r_{16.2}^2}\sqrt{1-r_{56.2}^2}}$	15.2 16.2 56.2						$r_{15.26}$	
$r_{12.34} = \frac{r_{12.3} - r_{14.3} \cdot r_{24.3}}{\sqrt{1-r_{14.3}^2}\sqrt{1-r_{24.3}^2}}$	12.3 14.3 24.3	—0,740 —0,687 0,545	0,727 0,838	—0,374	—0,366	0,610	$r_{12.34}$	—0,600

$r_{15.34}=\frac{r_{15.3}-r_{14.3}\cdot r_{45.3}}{\sqrt{1-r_{14.3}^2}\sqrt{1-r_{45.3}^2}}$	15.3 14.3 45.3	−0,738 −0,687 0,594	0,727 0,804	−0,408	−0,330	0,585	$r_{15.34}$	−0,564
$r_{16.34}=\frac{r_{16.3}-r_{14.3}\cdot r_{46.3}}{\sqrt{1-r_{14.3}^2}\sqrt{1-r_{46.3}^2}}$	16.3 14.3 46.3	0,598 −0,687 −0,477	0,727 0,879	0,328	0,270	0,639	$r_{16.34}$	0,423
$r_{12.35}=\frac{r_{12.3}-r_{15.3}\cdot r_{25.3}}{\sqrt{1-r_{15.3}^2}\sqrt{1-r_{25.3}^2}}$	12.3 15.3 25.3						$r_{12.35}$	
$r_{16.35}=\frac{r_{16.3}-r_{15.3}\cdot r_{56.3}}{\sqrt{1-r_{15.3}^2}\sqrt{1-r_{56.3}^2}}$	16.3 15.3 56.3						$r_{16.35}$	
$r_{12.36}=\frac{r_{12.3}-r_{16.3}\cdot r_{26.3}}{\sqrt{1-r_{16.3}^2}\sqrt{1-r_{26.3}^2}}$	12.3 16.3 26.3						$r_{12.36}$	
$r_{14.36}=\frac{r_{14.3}-r_{16.3}\cdot r_{46.3}}{\sqrt{1-r_{16.3}^2}\sqrt{1-r_{46.3}^2}}$	14.3 16.3 46.3						$r_{14.36}$	
$r_{15.36}=\frac{r_{15.3}-r_{16.3}\cdot r_{56.3}}{\sqrt{1-r_{16.3}^2}\sqrt{1-r_{56.3}^2}}$	15.3 16.3 56.3						$r_{15.36}$	

Продолжение

Формула частного коэффициента корреляции — r	r первого порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r второго порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{12.45}=\frac{r_{12.4}-r_{15.4}\cdot r_{25.4}}{\sqrt{1-r_{15.4}^2}\sqrt{1-r_{25.4}^2}}$	12.4 15.4 25.4						$r_{12.45}$	
$r_{13.45}=\frac{r_{13.4}-r_{15.4}\cdot r_{35.4}}{\sqrt{1-r_{15.4}^2}\sqrt{1-r_{35.4}^2}}$	13.4 15.4 35.4						$r_{13.45}$	
$r_{16.45}=\frac{r_{16.4}-r_{15.4}\cdot r_{56.4}}{\sqrt{1-r_{15.4}^2}\sqrt{1-r_{56.4}^2}}$	16.4 15.4 56.4						$r_{16.45}$	
$r_{12.46}=\frac{r_{12.4}-r_{16.4}\cdot r_{26.4}}{\sqrt{1-r_{16.4}^2}\sqrt{1-r_{26.4}^2}}$	12.4 16.4 26.4						$r_{12.46}$	
$r_{13.46}=\frac{r_{13.4}-r_{16.4}\cdot r_{36.4}}{\sqrt{1-r_{16.4}^2}\sqrt{1-r_{36.4}^2}}$	13.4 16.4 36.4						$r_{13.46}$	

$r_{15.46}=\dfrac{r_{15.4}-r_{16.4}\cdot r_{56.4}}{\sqrt{1-r_{16.4}^2}\sqrt{1-r_{56.4}^2}}$	15.4 16.4 56.4						$r_{15.46}$	
$r_{12.56}=\dfrac{r_{12.5}-r_{16.5}\cdot r_{26.5}}{\sqrt{1-r_{16.5}^2}\sqrt{1-r_{26.5}^2}}$	12.5 16.5 26.5						$r_{12.56}$	
$r_{13.56}=\dfrac{r_{13.5}-r_{16.5}\cdot r_{36.5}}{\sqrt{1-r_{16.5}^2}\sqrt{1-r_{36.5}^2}}$	13.5 16.5 36.5						$r_{13.56}$	
$r_{14.56}=\dfrac{r_{14.5}-r_{16.5}\cdot r_{46.5}}{\sqrt{1-r_{16.5}^2}\sqrt{1-r_{46.5}^2}}$	14.5 16.5 46.5						$r_{14.56}$	

в) третьего порядка

Формула частного коэффициента корреляции — r	r второго порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac}\times r_{bc}$	Числитель	Знаменатель	r третьего порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{15.234}=\dfrac{r_{15.23}-r_{14.23}\cdot r_{45.23}}{\sqrt{1-r_{14.23}^2}\sqrt{1-r_{45.23}^2}}$	15.23 14.23 45.23						$r_{15.234}$	

Продолжение

Формула частного коэффициента корреляции — r	r второго порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r третьего порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{16.234}=\frac{r_{16.23}-r_{14.23}\cdot r_{46.23}}{\sqrt{1-r_{14.23}^2}\sqrt{1-r_{46.23}^2}}$	16.23 14.23 46.23						$r_{16.234}$	
$r_{14.235}=\frac{r_{14.23}-r_{15.23}\cdot r_{45.23}}{\sqrt{1-r_{15.23}^2}\sqrt{1-r_{45.23}^2}}$	14.23 15.23 45.23						$r_{14.235}$	
$r_{16.235}=\frac{r_{16.23}-r_{15.23}\cdot r_{56.23}}{\sqrt{1-r_{15.23}^2}\sqrt{1-r_{56.23}^2}}$	16.23 15.23 56.23						$r_{16.235}$	
$r_{14.236}=\frac{r_{14.23}-r_{16.23}\cdot r_{46.23}}{\sqrt{1-r_{16.23}^2}\sqrt{1-r_{46.23}^2}}$	14.23 16.23 46.23						$r_{14.236}$	
$r_{15.236}=\frac{r_{15.23}-r_{16.23}\cdot r_{56.23}}{\sqrt{1-r_{16.23}^2}\sqrt{1-r_{56.23}^2}}$	15.23 16.23 56.23						$r_{15.236}$	

$r_{13.245}=\dfrac{r_{13.24}-r_{15.24}\cdot r_{35.24}}{\sqrt{1-r_{15.24}^2}\sqrt{1-r_{35.24}^2}}$	13.24 15.24 35.24						$r_{13.245}$	
$r_{16.245}=\dfrac{r_{16.24}-r_{15.24}\cdot r_{56.24}}{\sqrt{1-r_{15.24}^2}\sqrt{1-r_{56.24}^2}}$	16.24 15.24 56.24						$r_{16.245}$	
$r_{13.246}=\dfrac{r_{13.24}-r_{16.24}\cdot r_{36.24}}{\sqrt{1-r_{16.24}^2}\sqrt{1-r_{36.24}^2}}$	13.24 16.24 36.24						$r_{13.246}$	
$r_{15.246}=\dfrac{r_{15.24}-r_{16.24}\cdot r_{56.24}}{\sqrt{1-r_{16.24}^2}\sqrt{1-r_{56.24}^2}}$	15.24 16.24 56.24						$r_{15.246}$	
$r_{13.256}=\dfrac{r_{13.25}-r_{16.25}\cdot r_{36.25}}{\sqrt{1-r_{16.25}^2}\sqrt{1-r_{36.25}^2}}$	13.25 16.25 36.25						$r_{13.256}$	
$r_{14.256}=\dfrac{r_{14.25}-r_{16.25}\cdot r_{46.25}}{\sqrt{1-r_{16.25}^2}\sqrt{1-r_{46.25}^2}}$	14.25 16.25 46.25						$r_{14.256}$	
$r_{12.345}=\dfrac{r_{12.34}-r_{15.34}\cdot r_{25.34}}{\sqrt{1-r_{15.34}^2}\sqrt{1-r_{25.34}^2}}$	12.34 15.34 25.34	−0,600 −0,564 0,367	0,826 0,930	−0,207	−0,393	0,768	$r_{12.345}$	−0,512

Продолжение

Формула частного коэффициента корреляции — r	r второго порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r третьего порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{16.345}=\frac{r_{16.34}-r_{15.34}\cdot r_{56.34}}{\sqrt{1-r_{15.34}^2}\sqrt{1-r_{56.34}^2}}$	16.34 15.34 56.34	0,423 —0,564 —0,321	0,826 0,947	0,181	0,242	0,782	$r_{16.345}$	0,309
$r_{12.346}=\frac{r_{12.34}-r_{16.34}\cdot r_{26.34}}{\sqrt{1-r_{16.34}^2}\sqrt{1-r_{26.34}^2}}$	12.34 16.34 26.34						$r_{12.346}$	
$r_{15.346}=\frac{r_{15.34}-r_{16.34}\cdot r_{56.34}}{\sqrt{1-r_{16.34}^2}\sqrt{1-r_{56.34}^2}}$	15.34 16.34 56.34						$r_{15.346}$	
$r_{12.356}=\frac{r_{12.35}-r_{16.35}\cdot r_{26.35}}{\sqrt{1-r_{16.35}^2}\sqrt{1-r_{26.35}^2}}$	12.35 16.35 26.35						$r_{12.356}$	
$r_{14.356}=\frac{r_{14.35}-r_{16.35}\cdot r_{46.35}}{\sqrt{1-r_{16.35}^2}\sqrt{1-r_{46.35}^2}}$	14.35 16.35 46.35						$r_{14.356}$	

$r_{12.456}=\frac{r_{12.45}-r_{16.45}\cdot r_{26.45}}{\sqrt{1-r^2_{16.45}}\sqrt{1-r^2_{26.45}}}$	12.45 16.45 26.45						$r_{12.456}$	
$r_{13.456}=\frac{r_{13.45}-r_{16.45}\cdot r_{36.45}}{\sqrt{1-r^2_{16.45}}\sqrt{1-r^2_{36.45}}}$	13.45 16.45 36.45						$r_{13.456}$	

г) четвертого порядка

Формула частного коэффициента корреляции—r	r третьего порядка		$\sqrt{1-r^2_{ab}}$	$r_{ac}\times r_{bc}$	Числитель	Знаменатель	r четвертого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{12.3456}=\frac{r_{12.345}-r_{16.345}\cdot r_{26.345}}{\sqrt{1-r^2_{16.345}}\sqrt{1-r^2_{26.345}}}$	12.345 16.345 26.345	—0,512 0,309 —0,799	0,951 0,601	—0,247	—0,265	0,571	$r_{12.3456}$	—0,464
$r_{13.2456}=\frac{r_{13.245}-r_{16.245}\cdot r_{36.245}}{\sqrt{1-r^2_{16.245}}\sqrt{1-r^2_{36.245}}}$	13.245 16.245 36.245						$r_{13.2456}$	—0,139
$r_{14.2356}=\frac{r_{14.235}-r_{16.235}\cdot r_{46.235}}{\sqrt{1-r^2_{16.235}}\sqrt{1-r^2_{46.235}}}$	14.235 16.235 46.235						$r_{14.2356}$	—0,315

Продолжение

Формула частного коэффициента корреляции — r	r третьего порядка		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r четвертого порядка	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{15.2346} = \frac{r_{15.234} - r_{16.234} \cdot r_{56.234}}{\sqrt{1-r_{16.234}^2}\sqrt{1-r_{56.234}^2}}$	15.234 16.234 56.234						$r_{15.2346}$	—0,531
$r_{16.2345} = \frac{r_{16.234} - r_{15.234} \cdot r_{56.234}}{\sqrt{1-r_{15.234}^2}\sqrt{1=r_{56.234}^2}}$	16.234 15.234 56.234						$r_{16.2345}$	0,042

д) вычисление вспомогательных коэффициентов корреляции

Формула частного коэффициента корреляции — r	Вспомогательный r		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r разных порядков	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{26.345} = \frac{r_{26.34} - r_{25.34} \cdot r_{56.34}}{\sqrt{1-r_{25.34}^2}\sqrt{1-r_{56.34}^2}}$	26.34 25.34 56.34	—0,585 —0,367 —0,321	0,930 0,947	0,118	—0,703	0,880	$r_{26.345}$	—0,799

$r_{36.245} = \dfrac{r_{36.24} - r_{35.24} \cdot r_{56.24}}{\sqrt{1 - r^2_{35.24}}\sqrt{1 - r^2_{56.24}}}$	36.24 35.24 56.24						$r_{36.245}$	
$r_{46.235} = \dfrac{r_{46.23} - r_{45.23} \cdot r_{56.23}}{\sqrt{1 - r^2_{45.23}}\sqrt{1 - r^2_{56.23}}}$	46.23 45.23 56.23						$r_{46.235}$	
$r_{56.234} = \dfrac{r_{56.23} - r_{45.23} \cdot r_{46.23}}{\sqrt{1 - r^2_{45.23}}\sqrt{1 - r^2_{46.23}}}$	56.23 45.23 46.23						$r_{56.234}$	
$r_{25.34} = \dfrac{r_{25.3} - r_{24.3} \cdot r_{45.3}}{\sqrt{1 - r^2_{24.3}}\sqrt{1 - r^2_{45.3}}}$	25.3 24.3 45.3	0,571 0,545 0,594	0,838 0,804	0,324	0,247	0,674	$r_{25.34}$	0,376
$r_{56.34} = \dfrac{r_{56.3} - r_{45.3} \cdot r_{46.3}}{\sqrt{1 - r^2_{45.3}}\sqrt{1 - r^2_{46.3}}}$	56.3 45.3 46.3	—0,510 0,594 —0,477	0,804 0,879	—0,287	—0,227	0,706	$r_{56.34}$	—0,321
$r_{35.24} = \dfrac{r_{35.2} - r_{34.2} \cdot r_{45.2}}{\sqrt{1 - r^2_{34.2}}\sqrt{1 - r^2_{45.2}}}$	35.2 34.2 45.2						$r_{35.24}$	
$r_{56.24} = \dfrac{r_{56.2} - r_{45.2} \cdot r_{46.2}}{\sqrt{1 - r^2_{45.2}}\sqrt{1 - r^2_{46.2}}}$	56.2 45.2 46.2						$r_{56.24}$	
$r_{45.23} = \dfrac{r_{45.2} - r_{43.2} \cdot r_{35.2}}{\sqrt{1 - r^2_{43.2}}\sqrt{1 - r^2_{35.2}}}$	45.2 43.2 35.2						$r_{45.23}$	

Продолжение

Формула частного коэффициента корреляции — r	Вспомогательный r		$\sqrt{1-r_{ab}^2}$	$r_{ac} \times r_{bc}$	Числитель	Знаменатель	r разных порядков	
	символ	величина					символ	величина
1	2	3	4	5	6	7	8	9
$r_{56.23}=\frac{r_{56.2}-r_{35.2}\cdot r_{36.2}}{\sqrt{1-r_{35.2}^2}\sqrt{1-r_{36.2}^2}}$	56.2 35.2 36.2						$r_{56.23}$	
$r_{45.23}=\frac{r_{45.2}-r_{34.2}\cdot r_{35.2}}{\sqrt{1-r_{34.2}^2}\sqrt{1-r_{35.2}^2}}$	45.2 34.2 35.2						$r_{45.23}$	
$r_{46.23}=\frac{r_{46.2}-r_{34.2}\cdot r_{35.2}}{\sqrt{1-r_{34.2}^2}\sqrt{1-r_{36.2}^2}}$	46.2 34.2 36.2						$r_{46.23}$	
$r_{26.34}=\frac{r_{26.3}-r_{24.3}\cdot r_{46.3}}{\sqrt{1-r_{24.3}^2}\sqrt{1-r_{46.3}^2}}$	26.3 24.3 46.3	—0,690 0,545 —0,477	0,838 0,879	—0,260	—0,430	0,736	$r_{26.34}$	—0,585
$r_{36.24}=\frac{r_{36.2}-r_{34.2}\cdot r_{46.2}}{\sqrt{1-r_{34.2}^2}\sqrt{1-r_{46.2}^2}}$	36.2 34.2 46.2	—0,612 —0,168 —0,018	0,985 0,999	0,003	—0,615	0,986	$r_{36.24}$	—0,625

Значения $\sqrt{1-r^2}$ для r от 0,001 до 0,999

	0	1	2	3	4	5	6	7	8	9
0,01	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999	0,999
02	999	999	999	999	999	999	999	999	999	999
03	999	999	999	999	999	999	999	999	999	999
04	999	999	999	999	999	999	999	999	999	999
05	999	999	999	999	999	998	998	998	998	998
06	998	998	998	998	998	998	998	998	998	998
07	998	997	997	997	997	997	997	997	997	997
08	997	997	997	997	996	996	996	996	996	996
09	996	996	996	996	996	995	995	995	995	995
0,10	995	995	995	995	995	994	994	994	994	994
11	994	994	994	994	993	993	993	993	993	993
12	993	993	992	992	992	992	992	992	992	992
13	992	991	991	991	991	991	991	991	991	990
14	990	990	990	990	990	989	989	989	989	989
15	989	989	988	998	998	998	998	998	987	987
16	987	987	987	987	986	986	986	986	986	986
17	985	985	985	985	985	985	984	984	984	984
18	984	983	983	983	983	983	983	982	982	982
19	982	982	981	981	981	981	981	980	980	980
0,20	980	980	979	979	979	979	979	978	978	978
21	978	978	977	977	977	977	976	976	976	976
22	976	975	975	975	975	974	974	974	974	973
23	973	973	973	972	972	972	972	971	971	971
24	971	971	970	970	970	970	969	969	969	969
25	968	968	968	967	967	967	967	966	966	966
26	966	965	965	965	965	964	964	964	963	963
27	963	963	962	962	962	961	961	961	961	960
28	960	960	959	959	959	959	958	958	958	957
29	957	957	956	956	956	956	955	955	955	954
0,30	954	954	953	953	953	952	952	952	951	951
31	951	950	950	950	949	949	949	948	948	948
32	947	947	947	946	946	946	945	945	945	944
33	944	944	943	943	943	942	942	942	941	941
34	940	940	940	939	939	939	938	938	937	937
35	937	936	936	936	935	935	935	934	934	934
36	933	933	932	932	931	931	931	930	930	930
37	929	929	928	928	927	927	927	926	926	926
38	925	925	924	924	923	923	922	922	922	921
39	921	920	920	920	919	919	918	918	917	917

	0	1	2	3	4	5	6	7	8	9
0,40	0,917	0,916	0,916	0,915	0,915	0,914	0,914	0,913	0,913	0,913
41	912	912	911	911	911	910	910	909	909	908
42	908	907	907	906	906	905	905	904	904	903
43	903	902	902	901	901	900	900	899	899	898
44	898	897	897	897	896	896	895	895	894	894
45	893	893	892	892	891	891	890	890	889	889
46	888	888	887	887	886	886	885	885	884	884
47	883	883	882	882	881	880	879	879	878	878
48	877	877	876	876	875	875	874	873	873	872
49	872	871	871	870	869	869	868	868	867	867
0,50	866	866	865	864	863	863	863	862	862	861
51	860	860	859	858	858	857	856	856	855	855
52	854	854	853	852	852	851	850	850	849	849
53	848	847	847	846	845	845	844	844	843	842
54	842	841	840	840	839	838	838	837	836	836
55	835	835	834	833	833	832	831	831	830	829
56	828	828	827	826	826	825	825	824	823	822
57	822	821	820	820	819	818	817	817	816	815
58	815	814	813	812	812	811	810	809	808	808
59	807	807	806	805	804	804	803	802	801	801
0,60	800	799	799	797	797	796	796	795	794	793
61	792	792	791	790	789	789	788	787	786	785
62	785	784	783	782	781	780	780	779	778	777
63	776	776	775	774	773	773	772	771	770	769
64	768	767	767	766	765	764	764	762	762	761
65	760	759	758	758	756	756	755	754	753	752
66	751	750	750	748	748	747	746	745	744	743
67	743	742	740	740	739	738	737	736	735	734
68	734	732	731	731	729	727	727	727	726	725
69	724	723	722	721	720	719	718	717	716	715

	0	1	2	3	4	5	6	7	8	9
0,70	0,714	0,713	0,712	0,711	0,710	0,7 9	0,709	0,707	0,706	0,705
71	704	703	702	701	700	699	698	697	696	695
72	694	693	692	691	690	668	688	686	686	685
73	683	683	681	680	679	678	677	676	675	674
74	672	672	670	669	668	667	666	665	663	663
75	661	660	659	658	657	656	654	653	652	651
76	650	649	647	647	645	644	643	642	640	640
77	638	637	636	634	633	632	631	629	628	627
78	626	625	623	622	620	620	618	617	616	614
79	613	612	611	609	608	607	605	604	602	602
0,80	600	598	597	596	595	593	592	591	589	588
81	587	585	584	582	581	580	578	577	575	574
82	573	571	569	568	566	565	564	562	560	559
83	558	556	555	553	551	550	549	547	546	544
84	542	541	539	538	537	535	533	532	530	528
85	526	525	523	522	521	519	517	516	514	512
86	510	509	507	505	504	502	500	498	497	495
87	493	491	490	488	486	484	483	481	479	476
88	475	473	471	469	468	466	464	462	459	458
89	456	454	452	450	448	446	444	442	440	438
0,90	436	434	431	430	428	425	423	421	420	417
91	415	412	410	407	406	404	201	399	396	394
92	392	390	387	385	382	379	378	376	373	370
93	367	365	362	359	358	355	352	349	346	344
94	341	339	336	333	330	327	324	321	318	315
95	311	310	307	303	300	297	293	290	286	283
96	279	276	274	270	266	263	259	255	251	247
97	243	239	235	230	226	221	217	212	210	205
98	200	195	190	184	178	173	167	161	155	148
99	141	134	126	118	110	100	089	077	063	045

с применением контрольной суммы

расширения

x_3				x_4			x_1	
x_3^2	$x_3 \cdot x_4$	$x_3 \cdot x_1$	Контрольная сумма $x_3 \Sigma x$	x_4^2	$x_4 \cdot x_1$	Контрольная сумма $x_4 \Sigma x$	x_1^2	Контрольная сумма $x_1 \cdot \Sigma x$
11	12	13	14	15	16	17	18	19
9,4864	8,1620	3,080	37,6684	7,0225	2,650	32,4095	1,00	12,230
17,9776	9,7944	7,208	71,4440	5,3361	3,927	38,9235	2,89	28,645
12,8881	8,2570	5,026	58,8401	5,2900	3,220	37,6970	1,96	22,946
19,1844	17,3448	10,074	87,3372	15,6816	9,108	78,9624	5,29	45,862
21,9024	22,1364	9,360	96,9228	22,3729	9,460	97,9583	4,00	41,420
25,9081	20,3091	8,144	107,8062	15,9201	6,384	84,5082	2,56	33,888
28,5156	20,1318	10,146	115,9314	14,2129	7,163	81,8467	3,61	41,249
40,7044	32,1552	16,588	182,5956	25,4016	13,104	144,2448	6,76	74,412
51,2656	34,0816	19,332	214,2272	22,6576	12,852	142,4192	7,29	80,784
......	......	......	......	......	......	......	......	......
45,8329	35,4748	16,248	216,0307	27,4576	12,576	167,2084	5,76	76,584
41,0881	35,7678	16,666	210,1839	31,1364	14,508	182,9682	6,76	85,254
83,1744	63,2016	31,008	370,7280	48,0249	23,562	281,7045	11,56	138,210
98,0100	75,9390	31,680	429,3630	58,8289	24,544	332,6479	10,24	138,784
87,7969	74,8663	36,543	429,7082	63,8401	31,161	366,4214	15,21	178,854
4,4100	4,3470	2,100	16,5270	4,2849	2,070	16,2909	1,00	7,870
8,4100	6,0030	4,060	38,2930	4,2849	2,898	25,1919	1,96	17,038
17,3056	12,3968	5,824	73,7984	8,8804	4,172	52,8652	1,96	24,836
25,6036	21,8086	10,120	116,2282	18,5761	8,620	99,0007	4,00	45,940
......	......	......	......	......	......	......	......	......
45,9684	48,3414	19,662	239,4018	50,8369	20,677	251,7603	8,41	102,399
69,8896	62,1148	27,558	314,2524	55,2049	24,519	279,2937	10,89	124,047
56,1001	52,4300	20,972	272,5611	49,0000	19,600	254,7300	7,84	101,892
62,7264	60,4296	26,928	308,4840	58,2169	25,942	297,1885	11,56	132,430
61,1524	58,1026	26,588	302,2430	55,2049	25,262	287,1695	11,56	131,410
85,1929	93,4076	35,997	407,5045	102,4144	39,468	446,7980	15,21	172,185
44,3556	46,5534	19,314	258,7410	48,8601	20,271	271,5615	8,41	112,665
59,5984	75,0384	24,704	336,1288	94,4784	31,104	423,2088	10,24	139,328
178,2225	183,9630	76,095	930,8955	189,8884	78,546	960,8794	32,49	397,461
3155,6042	2766,3683	1282,235	14676,8035	2493,2557	1130,920	12991,269	529,500	5984,585

Вычисление алгебраических расширений

№ СМУ	Факторы: Годовой объем строительно-монтажных работ — x_2	Факторы: Численность работников x_3	Факторы: Фонд заработной платы, (тыс. руб.) x_4	Признак: Накладные расходы, (тыс. руб.) x_1	Контрольная сумма Σx	Алгебраические: x_2: x_2^2	$x_2 \cdot x_3$	$x_2 \cdot x_4$	$x_2 \cdot x_1$	Контрольная сумма $x_2 \Sigma x$
А	1	2	3	4	5	6	7	8	9	10
1	5,50	3,08	2,65	1,00	12,23	30,25	16,940	14,575	5,50	67,265
2	8,60	4,24	2,31	1,70	16,85	73,96	36,464	19,866	14,62	144,910
3	9,10	3,59	2,30	1,40	16,39	82,81	32,669	20,930	12,74	149,149
4	9,30	4,38	3,96	2,30	19,94	86,49	40,734	36,828	21,39	185,442
5	9,30	4,68	4,73	2,00	20,71	86,49	43,524	43,989	18,60	192,603
6	10,50	5,09	3,99	1,60	21,18	110,25	53,445	41,895	16,80	222,390
7	10,70	5,34	3,77	1,90	21,71	114,49	57,135	40,339	20,33	232,297
8	14,60	6,38	5,04	2,60	28,62	213,16	93,148	73,584	37,96	417,852
9	15,30	7,16	4,76	2,70	29,92	234,09	109,548	72,828	41,31	457,776
..	....	....	....	....	.....	......	......	......	......	......
35	17,50	6,77	5,24	2,40	31,91	306,25	118,475	91,700	42,00	558,425
36	18,20	6,41	5,58	2,60	32,79	331,24	116,662	101,556	47,32	596,778
37	21,20	9,12	6,93	3,40	40,65	449,44	193,344	146,916	72,08	861,780
38	22,60	9,90	7,67	3,20	43,37	510,76	223,740	173,342	72,32	980,162
39	24,60	9,37	7,99	3,90	45,86	605,16	230,502	196,554	95,94	1128,156
40	2,70	2,10	2,07	1,00	7,87	7,29	5,670	5,589	2,70	21,249
41	5,80	2,90	2,07	1,40	12,17	33,64	16,820	12,006	8,12	70,586
42	9,20	4,16	2,98	1,40	17,74	84,64	38,272	27,416	12,88	163,208
43	11,60	5,06	4,31	2,00	22,97	134,56	58,696	49,996	23,20	266,452
..	....	....	....	....	.....	......	......	......	......	......
66	18,50	6,78	7,13	2,90	35,31	342,25	125,430	131,905	53,65	653,235
67	18,50	8,36	7,43	3,30	37,59	342,25	154,660	137,455	61,05	695,415
68	19,10	7,49	7,00	2,80	36,39	364,81	143,059	133,700	53,48	695,049
69	20,00	7,92	7,63	3,40	38,95	400,00	158,400	152,600	68,00	779,000
70	20,00	7,82	7,43	3,40	38,65	400,00	156,400	148,600	68,00	773,000
71	20,90	9,23	10,12	3,90	44,15	436,81	192,907	211,508	81,51	922,735
72	22,30	6,66	6,99	2,90	38,85	497,29	148,518	155,877	64,67	866,355
73	22,90	7,72	9,72	3,20	43,54	524.41	176,788	222,588	73,28	997,066
74	36,90	13,35	13,78	5,70	69,73	1361,61	492,615	508,482	210,33	2573,037
$\Sigma n = 74$	1056,80	448,50	393,55	184,40	2083,25	18010,440	7472,596	6601,425	3041,93	35126,391

Таблица 2

Вычисления средних величин с помощью контрольных сумм и произведения сумм, корректированных отклонениями от средних

Номера строк		x_2	x_3	x_4	x_1	Σx
1	Сумма значений каждой переменной: Σx_2; Σx_3; Σx_4; Σx_1 (табл. 1)	1056,80	448,50	393,55	184,40	2083,25
2	Средние значения переменных: $\bar{x}_2$; $\bar{x}_3$; $\bar{x}_4$; $\bar{x}_1$ при $n=74$	14,28108	6,06081	5,31824	2,49189	28,15
3	Сумма произведения переменных на x_2; Σx_2^2; $\Sigma x_3 \quad x_2$; ... (табл. 1)	18010,44	7472,60	6601,42	3041,93	35126,39
4	Сумма значений каждой переменной (строка 1), умноженная на среднее значение $\bar{x}_2$ (14,28108)	15092,25	6405,07	5620,32	2633,43	29751,07
5	Сумма отклонений от средних значений x_2 (строка 3 — строка 4)	2918,19	1067,53	981,10	408,50	5375,32
6	Σx_3^2; $\Sigma x_4 \quad x_3$; $\Sigma x_1 \cdot x_3$; $\Sigma x \cdot x_3$ (табл. 1)	—	3155,60	2766,37	1282,23	14676,80
7	Сумма значений переменных Σx_3; Σx_4; Σx_1; Σx (строка 1), умноженная на среднее значение $\bar{x}_3$ (6,06081)	—	2718,27	2385,23	1117,61	12626,18
8	Сумма отклонений от средних значений x_3 (строка 6—строка 7)	—	437,33	381,14	164,62	2050,62
9	Сумма произведения переменных на x_4; Σx_4^2; $\Sigma x_1 x_4$; $\Sigma x \cdot x_4$ (табл. 1)	—	—	2493,25	1130,92	12991,97
10	Сумма значений переменных $\Sigma x_4, \Sigma x_1$; Σx (строка 1), умноженная на среднее значение $\bar{x}_4$ (5,31824)	—	—	2092,99	980,68	11079,22
11	Сумма отклонений от средних значений $\bar{x}_4$ (строка 9—строка 10)	—	—	400,26	150,24	1912,76
12	Σx_1^2; $\Sigma x \cdot x_1$ (табл. 1)	—	—	—	529,50	5984,58
13	Суммы значений переменных Σx_1, Σx (строка 1), умноженные на среднее значение $\bar{x}_1$ (2,49189)	—	—	—	459,50	5191,23
14	Сумма отклонений от средних значений $\bar{x}_1$ (строка 12—строка 13)	—	—	—	70,00	793,35

Таблица 3

Вычисление коэффициентов переменных уравнения множественной регрессии с использованием метода обратного решения и контрольных сумм

Номера строк		x_2	x_3	x_4	x_1	Σx
1	Данные строки 5 табл. 2	2918,19	1067,53	981,10	408,50	5375,32
2	Частные, полученные от деления данных строки 1 на первое число 1 строки с обратным знаком (—2918,19)	—1,00000	—0,36582	—0,33620	—0,13998	—1,84200
3	Данные строки 8 табл. 2, начиная с графы x_3	(1067,53)	437,33	381,14	164,62	2050,62
4	Результаты произведения данных строки 1 на величину x_3 строки 2 (—0,36582)	—(1067,53)	—390,52	—358,90	—149,44	—1966,39
5	Сумма строк 3 и 4	—	46,81	22,23	15,18	84,22
6	Частные от деления данных строки 5 на величину x_3 строки 5 с обратным знаком (—46,81)	—	—1,00000	—0,47490	—0,32429	—1,79919
7	Данные строки 11 табл. 2, начиная с графы x_4	(981,10)	(381,14)	400,26	150,24	1912,75
8	Результаты произведения данных строки 1 на величину x_4 строки 2 (—0,33620)	(—981,10)	(—358,90)	—329,84	—138,34	—1807,18
9	Результаты произведения данных строки 5 на величину x_4 строки 6 (—0,4740)	—	(—22,23)	—10,56	—7,21	—40,00
10	Сумма строк 7 + 8 + 9	—	—	59,86	5,69	65,55
11	Частные от деления данных строки 10 на значение x_4 с обратным знаком (—59,86)	—	—	—1,00000	—0,09505	—1,09505

Примечание. Величины в скобках указаны для контроля графы Σx, например для строки 3:
(1067,53)+437,33+381,14+164,62=2050,62.

Таблица 4

Определение $R^2_{1.23}$

Номера строк	Исходные данные	x_3	x_1
1	Табл. 2, строка 5	—	408,5
2	Табл. 3, строка 2	—0,36582	—0,13998
3	Табл. 2, строка 8	—	164,62
4	Табл. 3, строка 6	—	—0,32429
5	Табл. 2, строка 14	—	70
	Определение значений коэффициентов методом обратных решений при исключении x_4	**Частные коэффициенты корреляции**	
		$b_{12.3}$	$b_{13.2}$
6	Показатели графы x_1 строк 2 и 4 с обратным знаком	0,13998	0,32429
7	—0,36582 (—0,32429)	0,11863	—
8	Значение коэффициентов (строка 6 — строка 7)	0,02135	0,32429
	Определение значения числителя $R^2_{1.23}$		
9	$b_{13.2} \times$ строку $3 = 0,32429 \times 164,62 = 53,386$		
10	$b_{12.3} \times$ строку $1 = 0,02135 \times 408,5 = 8,721$		
	$R^2_{1.23} = \dfrac{53,386 + 8,721}{70} = 0,887244$		

Таблица 5

Определение $R^2_{2.24}$

Номера строк	Исходные данные	x_4	x_1
1	Табл. 2, строка 5	—	408,5
2	Табл. 3, строка 2	—0,33620	—0,13998
3	Табл. 3, строка 7	400,26	150,24
4	Табл. 3, строка 8	—329,84	—137,34
5	Сумма строк 3+4	70,42	12,90
6	Частные от деления строки 5 на 70,42 с обратным знаком	—1,00000	—0,1832
7	Табл. 2, строка 14	—	70
	Определение значений коэффициентов методом обратных решений при исключении x_3	**Частные коэффициенты корреляции**	
		$b_{12.4}$	$b_{14.2}$
8	Показатели графы x_1 строк 2 и 6 с обратным знаком	0,13998	0,1832
9	(—0,33620)·(—0,1832)	0,06159	—
10	Значения коэффициентов (строка 8—строка 9)	0,07839	0,1832
	Определение значения $R^2_{1.24}$		
11	$b_{14.2} \times$ строку $3 = 0,1832 \times 150,24 = 27,524$		
12	$b_{12.4} \times$ строку $1 = 0,07839 \times 408,50 = 33,586$		
	$R^2_{1.24} = \dfrac{27,524 + 33,586}{70} = 0,873008$		

Таблица 6

Определение $R^2_{1.34}$

Номера строк	Исходные данные	x_3	x_4	x_1
1	Табл. 3, строка 3	437,33	381,14	164,62
2	Частные от деления строки 1 на 437,33 с обратным знаком	—1,00000	—0,87151	—0,37642
3	Табл. 3, строка 7	—	400,26	150,24
4	Данные x_4, x_1 строки 1, умноженные на x_4 строки 2 (—0,87154)	—	—332,17	—143,47
5	Сумма строк 3 и 4	—	68,09	6,77
6	Частные от деления строки 5 на 68,09 с обратным знаком	—	—1,00000	—0,09943
7	Табл. 2, строка 14	—	—	70

	Определение значений коэффициентов методом обратных решений при исключении x_2	Частные коэффициенты корреляции	
		$b_{13.4}$	$b_{14.3}$
8	Показатели графы x_1 строк 2 и 6 с обратным знаком	0,37642	0,09943
9	—0,87151·(—0,09943)	0,08665	—
10	Значения коэффициентов (строка 8—строка 9)	0,28977	0,09943

Определение значения числителя $R^2_{1.34}$

11 $b_{14.3} \times$ строку $3 = 0{,}0943 \times 150{,}24 = 14{,}938$

12 $b_{13.4} \times$ строку $1 = 0{,}28977 \times 164{,}62 = 47{,}517$

$$R^2_{1.34} = \frac{47{,}517 + 14{,}938}{70} = 0{,}892214$$

ЛИТЕРАТУРА

1. Авилов В. А. Математико-статистические методы технико-экономического анализа производства, М., «Экономика», 1967.

2. Бородкин Ф. М. Некоторые вопросы применения корреляционного анализа в экономике и планировании. — В сб. «Статистические модели и методы в экономическом анализе и планировании». Вып. VII, Новосибирск, 1963.

3. Brandon D. B., Hooke R. and Van Nice R. I. Developing Mathematical Models for Computer Control. «ISA journal», july 1959, v. 6.

4. Вальтух К. К. О применении методов множественной корреляции для анализа и планирования эффективности использования фондов промышленных предприятий. — В сб. «Статистические модели и методы в экономическом анализе и планировании». Вып. VII, Новосибирск, 1963.

5. Венецкий И. Г., Кильдишев Г. С. Основы теории вероятностей и математической статистики, М., «Статистика», 1968.

6. Вентцель Е. С. Теория вероятностей, издание 3, исправленное. М., «Наука», 1964.

7. Веселая Г. Н., Френкель А. А. Многошаговый регрессионный анализ при построении статистических моделей производительности труда. — «Экономика и математические методы». Т. I, вып. 4, 1965.

8. Езекиэл М., Фокс К., Методы анализа корреляций и регрессий. Пер. с англ. М., «Статистика», 1966.

9. Куприянова З. В. Опыт применения метода множественной корреляции для анализа факторов роста производительности труда в строительных организациях. — В сб. «Статистические модели и методы в экономическом анализе и планировании». Вып. VII, Новосибирск, 1963.

10. Леонтьев Н. Л. Техника статистических вычислений, М., «Лесная промышленность», 1966.

11. Лукомский Я. И. Теория корреляции и ее применение к анализу производства. Изд. 2, М., Госстатиздат, 1961.

12. Миллс Ф. Статистические методы. Пер. с англ. М., Госстатиздат, 1958.

13. Пановский Г. А., Брайер Г. В. Статистические методы в метеорологии. Пер. с англ. М., Метеоиздат, 1967.

14. Перегудов В. Н. Метод наименьших квадратов и его применение в исследованиях, М., «Статистика», 1965.

15. Поянсюк Г. Б. Анализ себестоимости строительно-монтажных работ методом корреляции, М., Стройиздат, 1967,
16. Померанцев В. В. Расчеты в перспективном планировании, М., «Экономика», 1966.
17. Розанов Г. В., Френкель А. А. Корреляционный и регрессионный анализ в экономических исследованиях. — «Экономика и математические методы». Том III, вып. 3, 1967.
18. Розин Б. Б., Гейфман Р. С., Математические методы и счетная техника в организации металлургического производства, М., 1962. Металлургиздат, 1962.
19. Френкель А. А. Графический способ построения экономико-статистических моделей. — «Экономика и математические методы». Т. I, вып. 2, 1965.
20. Хайкин В. П., Найденов В. С., Галуза С. Г. Корреляция и статистическое моделирование в экономических расчетах, М., «Экономика», 1964.
21. Хинчин Л. Я. Предельные законы для сумм независимых случайных величин, М.—Л., Редакция технико-теоретической литературы, 1938.
22. Янко Я. Математико-статистические таблицы. Пер. с чешского. М., Госстатиздат, 1961.

ОГЛАВЛЕНИЕ

Исаак Соломонович Пальма, Лиза Гершовна Эльгорт

Применение метода корреляции в строительстве

Научные редакторы: *В. А. Садовская, В. И. Сиськов*
Редактор *В. Н. Третьякова*
Техн. редактор *Р. Н. Феоктистова* Корректор *А. А. Лопатина*
Худож. редактор *Т. В. Стихно*
Переплет художницы *Т. Н. Погореловой*

Сдано в набор 4/VIII 1970 г. Подписано к печати 10/XII 1970 г.
Формат бумаги 84×108 1/32 Бумага № 2 Объем 7,0 печ. л. Уч.-изд. л. 10,89
Тираж 3250 экз. А11914 (Тематич. план 1970 г. № 19
Зак. 1307 Цена 97 коп.
Издательство «Статистика», Москва, ул. Кирова, 39.

Московская типография № 4 Главполиграфпрома
Комитета по печати при Совете Министров СССР
Б. Переяславская, 46

Замеченные опечатки

Страница	Строка	Напечатано	Следует читать
26	3 снизу	x_l	x_1
96, табл. 31	4 сверху, 3-я графа	1,909	0,909
135	12 сверху	$+0{,}53383\ x_5$	$-0{,}53383\ x_5$
	7 снизу	$+0{,}53383\ x_5$	$-0{,}53383\ x_5$
138, табл. 44	4 сверху, 2-я графа	+	—
150	7 сверху	$(y,\ \bar{y})$	(x_1)
163	9 сверху	(IV.17)	(IV.16)

Зак. 1307