Л. Л. ТЕРЕХОВ
В. А. КУЦЕНКО
С. П. СИДНЕВ

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В ПЛАНИРОВАНИИ И УПРАВЛЕНИИ

Л. Л. ТЕРЕХОВ,
В. А. КУЦЕНКО,
С. П. СИДНЕВ

ЭКОНОМИКО-МАТЕМАТИЧЕСКИЕ МЕТОДЫ И МОДЕЛИ В ПЛАНИРОВАНИИ И УПРАВЛЕНИИ

Рекомендовано Министерством высшего и среднего специального образования УССР для слушателей факультетов повышения квалификации руководящих работников и специалистов народного хозяйства

КИЕВ
ГОЛОВНОЕ ИЗДАТЕЛЬСТВО
ИЗДАТЕЛЬСКОГО ОБЪЕДИНЕНИЯ
«ВИЩА ШКОЛА»
1984

65.9(2)21
Т35

УДК 330. 115 (07)

Экономико-математические методы и модели в планировании и управлении. Терехов Л. Л., Куценко В. А., Сиднев С. П.— К.: «Вища школа». Головное изд-во, 1984.— 231 с.

В учебном пособии рассмотрены наиболее универсальные и широко применяемые в планировании и управлении, в АСУ различного уровня экономико-математические методы и модели. Изложены методы корреляционно-регрессионного анализа, балансовые методы, методы имитационного моделирования, типовые модели оптимального и сетевого планирования, методы и модели математического прогнозирования, применяемые в отраслях народного хозяйства и производственных объединениях. Показана необходимость построения АСУ, использующих систему экономико-математических методов.

Для слушателей факультетов повышения квалификации руководящих работников и специалистов народного хозяйства.

Табл. 26. Ил. 20. Библиогр.: 64 назв.

Рецензенты: кандидаты экономических наук *В. М. Гранатуров* (Одесский институт народного хозяйства) и *О. Е. Блинов* (Московский институт управления)

Редакция литературы по кибернетике, электронике и энергетике
Зав. редакцией *М. С. Хойнацкий*

Т $\frac{1502000000-014}{М211\ (04)-84}$ 110—84

ВВЕДЕНИЕ

Совершенствование управления народным хозяйством относится к числу важнейших научных и практических задач на современном этапе коммунистического строительства. Широкие возможности для совершенствования управления, повышения его эффективности, оперативности, действенности открывает использование электронной вычислительной техники в сочетании с современными математическими и кибернетическими методами. Практический результат применения ЭВМ заключается в разработке и внедрении в народное хозяйство автоматизированных систем управления (АСУ), являющихся комплексными человеко-машинными системами сбора, передачи, хранения, обработки и выдачи информации.

Автоматизированные системы управления разрабатываются во всех звеньях народного хозяйства. В низовом звене — производственных объединениях и предприятиях — создаются и внедряются АСУ двух видов: автоматизированные системы управления технологическими процессами (АСУТП) и автоматизированные системы организационно-экономического управления производством (АСУП). В отраслях народного хозяйства (в территориальном разрезе — союзных и республиканских) создаются и внедряются отраслевые автоматизированные системы (ОАСУ). В народном хозяйстве разрабатываются и частично функционируют автоматизированная система плановых расчетов (АСПР) Госплана СССР и госпланов союзных республик, автоматизированная система государственной статистики (АСГС), АСУ материально-техническим снабжением и др. В перспективе эти системы будут интегрироваться в общегосударственную автоматизированную систему сбора и обработки информации для учета, планирования и управления народным хозяйством — ОГАС.

Комплекс задач, решаемых в рамках АСУ, имеет большое значение с точки зрения обеспечения ее высокой экономической эффективности.

Потенциальная эффективность автоматизированных систем резко возрастает, когда в сферу АСУ широко вовлекаются задачи активного характера, т. е. планирование, организация, оперативное управление производством. Эти задачи многовариантны. Тем не менее только на базе активных задач можно в полной мере использовать возможности ЭВМ для совершенствования процессов управления. Однако одного включения в АСУ задач планирования и управления недостаточно. Трудно ожидать существенного эффекта, когда постановка и методы решения этих задач машиной фактически не отличаются от сложившейся практики их ручного решения. Иное дело, когда при создании АСУ максимально возможное число задач планирования и управления формулируется в виде оптимизационных моделей, когда широко используется весь арсенал экономико-математических и кибернетических методов, когда в АСУ разрабатываются и решаются такие задачи, которые раньше не ставились или ставились слишком упрощенно. Лишь при таком подходе АСУ не просто облегчают техническую вычислительную работу и упорядочивают обмен информацией, но и поднимают на новый, более высокий уровень всю систему управления производством. Таким образом, эффективность АСУ во многом зависит от их насыщенности задачами оптимального планирования и управления.

Многовариантные оптимизационные задачи относятся к наиболее сложным из числа решаемых в АСУ. Их труднее сформулировать и решить, чем задачи прямого счета, например учетные. Проблема сводится к тому, чтобы адекватно построить модель задачи и подобрать эффективный вычислительный метод ее решения. Наличие модели задачи и метода ее решения — необходимые и чаще всего достаточные условия для реализации экономической задачи на ЭВМ в АСУ или вне ее.

Математическая модель экономической задачи представляет собой совокупность уравнений, неравенств, функционалов, логических условий и других соотношений, отражающих реальные взаимосвязи и зависимости экономических показателей. Подлежащая решению экономическая задача вначале предстает, как правило, в неформальном виде. Человек в состоянии решать как точно сформулированные, так и неформально поставленные задачи. Но для решения на ЭВМ любые задачи должны быть предварительно формализованы, строго интерпретированы в логико-математическом смысле. В этом и состоит основная задача моделирования: довести первоначально нечетко формулируемую задачу до строгого формально-математического выражения.

Являясь всегда в той или иной мере упрощением реальной действительности, модель имеет и неоспоримые достоинства: в ней отсеиваются всевозможные незначащие частности, посторонние и случайные факторы, но сохраняется все наиболее существенное, что характерно для поставленной задачи; на модели можно проводить разнообразные многовариантные эксперименты, что далеко не всегда возможно на самих экономических объектах; модель можно расчленять на составные части или укрупнять с тем, чтобы колоссальная размерность реальных экономических задач не стала непреодолимым препятствием для их количественного анализа и оптимизации. Разумеется, всякие упрощения допустимы лишь в той мере, в какой конечные результаты расчетов по модели могут быть отнесены к реальной задаче с практически необходимой точностью.

Для разработанной модели экономической задачи должен быть определен метод ее численного решения. Математическая модель может быть построена так, что метод ее решения неизвестен либо известен, но требует непомерно большого объема вычислений даже при использовании самых быстродействующих ЭВМ. Поэтому процессы построения модели и выбора метода решения взаимосвязаны. Нередко приходится даже подгонять модель под известные эффективные численные методы. Отсюда понятна роль экономико-математических методов в обеспечении создания и функционирования эффективных автоматизированных систем управления.

Наука и практика располагают арсеналом математических методов для решения многообразных экономических задач. Многие их них были созданы практически одновременно с появлением первых ЭВМ. Это относится прежде всего к методам линейного, нелинейного, динамического программирования, сетевого планирования, теории игр, теории расписаний. Получили значительное развитие и ранее известные методы, в частности корреляционно-регрессионный анализ, анализ межотраслевых связей, теория массового обслуживания.

В нашей стране широкое развитие экономико-математических исследований началось на рубеже 50-х и 60-х годов. С тех пор накоплен уже немалый опыт постановки и решения экономико-математических задач. На уровне народного хозяйства страны, союзных республик, регионов регулярно проводятся аналитические и плановые расчеты по модели межотраслевого баланса производства и распределения продукции. Экономический эффект, полученный в результате оптимизации с применением ЭВМ планов развития и размещения производства в отраслях народного хозяйства, в том числе таких ведущих, как черная и цветная металлургия, добыча и переработка нефти, химическая промышленность, достигает многих миллионов рублей. Большое число оптимизационных задач решено на транспорте. В промышленных и производственных объединениях, на предприятиях экономико-математические методы позволяют определять оптимальный состав и структуру производимой продукции, наиболее рационально загружать оборудование, минимизировать отходы производства, совершенствовать оперативно-календарное планирование. В строительных отраслях и предприятиях повсеместно и эффективно применяются сетевые графики.

В начале развития и построения экономико-математических моделей отсутствовал какой-либо системный подход. Это приводило к тому, что оптимальное решение разрозненных задач в отдельных отраслях промышленности входило в противоречие с неоптимальностью планирования в смежных отраслях. Поэтому моделирование экономико-математических задач в настоящее время осуществляется в большинстве случаев в рамках основных подсистем АСУ, прежде всего — подсистем перспективного, текущего и оперативно-производственного планирования. Это вносит системность, комплексность в экономико-математические исследования в целом.

В настоящем пособии дается описание наиболее глубоко разработанных теоретически и широко применяемых в практических расчетах экономико-математических методов и моделей. В шести главах пособия последовательно изложены: методы корреляционно-регрессионного анализа; матричные методы и модели, включая основы построения межотраслевого баланса; методы и модели математического (преимущественно линейного) программирования; основы сетевого планирования и управления; некоторые методы и модели исследования операций, включая календарное планирование, задачи массового обслуживания и управления запасами; методы имитационного и игрового моделирования. По каждому из рассматриваемых методов дается общая характеристика, а затем анализируются сущность и особенности соответствующих прикладных экономико-математических задач.

Глава 1

КОРРЕЛЯЦИОННО-РЕГРЕССИОННЫЙ АНАЛИЗ

1.1. ОСНОВЫ ТЕОРИИ КОРРЕЛЯЦИЙ И РЕГРЕССИЙ

Построение корреляционных моделей позволяет дать количественную характеристику связи, зависимости и взаимной обусловленности экономических показателей. Хотя модель является упрощенным отражением действительности, она обеспечивает строго математический подход к исследованию экономических взаимосвязей. Вследствие математической завершенности, количественной определенности своих характеристик корреляционно-регрессионная модель служит не только средством анализа предшествующего экономического развития, но и становится важным инструментом прогнозных и плановых расчетов.

Несколько показателей могут быть связаны функциональной или корреляционной зависимостью. Функциональная зависимость проявляется определенно и точно в каждом отдельном случае, в каждом отдельном наблюдении. Например, закон Ома устанавливает функциональную зависимость между напряжением, приложенным к проводнику, его сопротивлением и силой тока. Этот закон строго соблюдается в каждом отдельном опыте независимо от того, какой длины и сечения взят проводник, из какого материала он изготовлен, большое или малое напряжение к нему приложено. Знание функциональных зависимостей позволяет абсолютно точно предсказывать события. Так, возможно на много лет вперед предсказать солнечные и лунные затмения с точностью до минут и даже секунд.

Корреляционная зависимость в отличие от функциональной проявляется лишь в среднем и только в массе наблюдений. Предположим, что исследуется зависимость между ростом людей и их массой. Ясно, что такая зависимость существует: если, скажем, взять наугад 100 человек и поставить их по росту, то по мере его увеличения масса людей в среднем тоже будет увеличиваться (50 более высоких наверняка перевесят 50 более низких). Но если взять только двух человек, то нет никакой гарантии, что более высокий из них будет весить больше. Вот почему корреляционная зависимость проявляется только в массе наблюдений.

Если подобрать математическую формулу, отражающую общую закономерность изменения массы с увеличением роста, то по ней, очевидно, нельзя будет определить массу какого-либо конкретного человека, рост которого известен. Тем не менее, исследование указанной зависимости и получение соответствующей математической формулы представляет большой интерес, хотя бы потому, что позволяет устанавливать, какую массу при данном росте следует считать нормальной. В обиходе распространена простейшая формула определения нормальной массы взрослого человека: $y = x - 100$, где x — рост в сантиметрах, а y — масса в килограммах. Зная норму, можно в каждом конкретном случае определить величину и характер отклонений от нее, что служит основанием для практических выводов и рекомендаций. Таковы особенности корреляционной зависимости по сравнению с функциональной.

Экономические величины складываются обычно под влиянием множества факторов. Закономерности не проявляются в сфере экономики с той точностью и неизменностью, как в мире неживой природы. Поэтому при изучении взаимосвязей экономических показателей чаще всего приходится прибегать к корреляционному анализу.

В простейшем случае корреляционного анализа исследуется связь между двумя показателями, из которых один рассматривается как независимый фактор (x), а второй — как зависимая переменная (y). Наличие самой зависимости между этими показателями устанавливается в результате качественного анализа, позволяющего вскрыть внутреннюю сущность изучаемого явления и порождающих его причин. Сам же корреляционный анализ предназначен для количественного измерения выявленной связи, хотя он нередко способствует и уточнению выводов самого качественного анализа. Таким образом, еще до математического расчета считается установленным, что связь между x и y может существовать и характеризуется корреляционной функцией $y = f(x)$.

Одной из первых задач корреляционного анализа является установление вида этой функции, т. е. отыскание такого корреляционного уравнения — уравнения регрессии, которое наилучшим образом соответствует характеру изучаемой связи. Уравнение регрессии — важнейшая составная часть корреляционных моделей, и его правильный подбор и расчет относятся к наиболее ответственным этапам моделирования.

Простейшим уравнением, которое может характеризовать зависимость между двумя переменными, является уравнение прямой вида

$$y = a_0 + a_1 x,$$

где a_0 и a_1 — постоянные коэффициенты (константы, параметры уравнения).

Уравнение прямой описывает такую связь между двумя переменными, при которой с изменением независимой переменной на какую-либо постоянную величину зависимая переменная изменяется на другую постоянную величину (в частности, при изменении x на одну единицу величина y изменяется на a_1 единиц).

Если качественный анализ изучаемой зависимости допускает прямолинейный характер связи двух переменных, то это предположение проверяется затем непосредственно на количественных данных. Для этого необходимо иметь ряд фактических значений переменной x и соответствующих ей величин зависимой переменной y. Поскольку корреляционная связь с достаточной четкостью и полнотой проявляется лишь в массе случаев, количество наблюдений, на основании которых строится модель, должно быть достаточно велико. Считается, что число наблюдений должно по меньшей мере в 5—6 раз превышать количество параметров уравнения.

Вывод о прямолинейном характере связи проверяется вначале путем простого сопоставления по имеющимся данным вариации зависимой и независимой переменных, а также графическим способом. При графическом способе каждое наблюдение отмечается в виде точки в прямоугольной системе координат, в которой по оси абсцисс отсчитываются значения x, по оси ординат — значения y. При достаточно большом количестве наблюдений расположение точек на графике позволяет определить правильность или ошибочность представления о линейном характере связи между изучаемыми переменными.

Следующим этапом является выявление уравнения прямой при данной конкретной зависимости между x и y. Для этого необходимо определить численные значения постоянных величин уравнения (a_0 и a_1), при которых прямая будет наилучшим образом соответствовать имеющимся фактическим данным. Критерий, по которому отыскивается наилучшая прямая, в известной мере условен. В качестве такого критерия принято брать минимум суммы квадратов отклонений фактических значений y от вычисленных по уравнению прямой. Минимуму квадратов отклонений соответствует единственная прямая, коэффициенты которой отыскиваются так называемым методом наименьших квадратов.

Таким образом, если связь между x и y характеризуется прямой вида $y = a_0 + a_1x$, то первой задачей корреляционного исчисления является определение таких значений a_0 и a_1, при которых сумма квадратов отклонений наблюдаемых значений $y_ф$ от расчетных $y_р$ будет минимальной, т. е.

$$\sum (y_ф - y_р)^2 = \min.$$

Уравнение прямой можно записать следующим образом:

$$y_р = a_0 + a_1 x_ф.$$

Подставляя значение $y_р$ в условие минимизации суммы квадратов, получим

$$\sum (y_ф - a_0 - a_1 x_ф)^2 = \min.$$

Рассматриваемая сумма квадратов представляет собой функцию, в которой $x_ф$ и $y_ф$ являются известными величинами (исходными данными), а a_0 и a_1 — неизвестными (искомыми) величинами.

В точке минимума функции первая производная равна нулю. Поэтому для расчета искомых коэффициентов прямой необходимо приравнять нулю частные производные данной функции по a_0 и a_1. Обозначив функцию в целом через f и отбросив индексы у x и y, имеем

$$\frac{\partial f}{\partial a_0} = -2 \sum (y - a_0 - a_1 x) = 0;$$

$$\frac{\partial f}{\partial a_1} = -2 \sum (y - a_0 - a_1 x)\, x = 0.$$

После упрощений и почленного суммирования получаем:

$$\begin{aligned} \sum y &= N a_0 + a_1 \sum x; \\ \sum xy &= a_0 \sum x + a_1 \sum x^2. \end{aligned} \qquad (1.1)$$

Это так называемая система нормальных уравнений, решение которой приводит к определению величины коэффициентов a_0 и a_1, все остальные входящие в систему величины рассчитываются на основании фактических исходных данных; N — обозначает число наблюдений.

Рассмотрим расчет уравнения прямой для характеристики корреляционной зависимости экономических показателей на следующем примере. В табл. 1.1 приведены данные по строительному тресту за 1972—1981 гг. о фондовооруженности труда и годовой выработке на одного работающего. Производительность труда (годовая выработка) выступает здесь в качестве анализируемого показателя. Естественно поставить в связь рост выработки с ростом фондовооруженности труда. Простое сопоставление 2-й и 3-й граф табл. 1.1 подтверждает наличие такой связи. На производительность труда влияет

Таблица 1.1. **Данные расчета зависимости производительнсcти труда от фондовооруженнсcти**

Годы	Фондовооруженность труда, тыс. руб. x	Годовая выработка на одного работающего $y_ф$	Расчетная выработка (по уравнению прямой) $y_р$	Отклонение $y_ф - y_р$
1972	0,90	3,95	4,04	—0,09
1973	0,98	4,15	4,15	0,00
1974	1,24	4,56	4,52	+0,04
1975	1,35	4,72	4,68	+0,04
1976	1,52	4,96	4,92	+0,04
1977	1,70	5,24	5,18	+0,06
1978	2,27	5,90	6,00	—0,10
1979	2,48	6,25	6,29	—0,04
1980	2,65	6,52	6,54	—0,02
1981	2,80	6,82	6,75	+0,07
Итого	17,89	53,07	—	0,00

множество факторов, но фондовооруженность занимает среди них ведущее место.

Для выявления характера изучаемой зависимости представим исходные данные примера графически. На графике (рис. 1.1) каждому году соответствует точка, координаты которой равны по оси абсцисс фондовооруженности x, а по оси ординат — выработке y. Расположение точек на графике свидетельствует о линейном характере зависимости этих показателей за исследуемый 10-летний период.

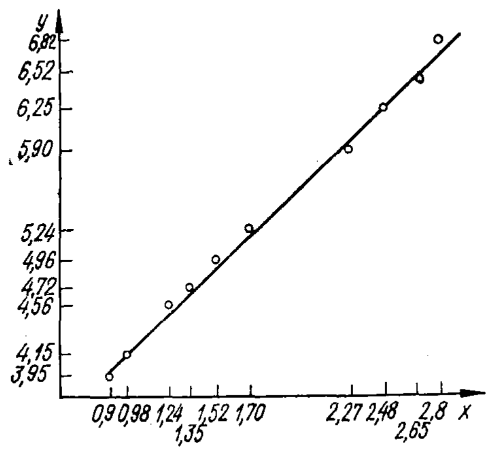

Рис. 1.1. Эмпирическая и теоретическая зависимость между производительностью труда и фондовооруженностью

Следующим этапом корреляционного анализа является расчет параметров a_0 и a_1 линейного уравнения $y = a_0 + a_1x$ путем решения системы нормальных уравнений (1.1). В рассматриваемом примере $N = 10$; $\Sigma x = 17,89$; $\Sigma y = 53,07$.

Для исчисления величины Σxy нужно за каждый год перемножить x и y и просуммировать полученные произведения; величина Σx^2 определяется в результате суммирования квадратов значений x. В итоге имеем следующую систему уравнений:

$$53{,}07 = 10a_0 + 17{,}89a_1;$$
$$101{,}36 = 17{,}89a_0 + 36{,}5a_1.$$

Решение системы дает следующий результат: $a_0 = 2{,}75$; $a_1 = 1{,}43$. Таким образом, искомое уравнение связи производительности труда и фондовооруженности труда имеет вид

$$y = 2{,}75 + 1{,}43x.$$

Если сами по себе исходные статистические данные, даже упорядоченные и охватывающие ряд периодов, остаются все же простым перечнем значений показателей, то уравнение связи указывает общий закон, по которому на данном предприятии изменялась выработка с изменением фондовооруженности труда. С помощью уравнения связи можно получить оценку прогнозного характера величины выработки и на будущие периоды, если известно, как будет изменяться фондовооруженность труда. Полученная оценка будет достаточно надежной, хотя с действительной цифрой будущей производительности труда она, вероятно, и не совпадет.

Рассмотрим подробнее вопрос о надежности оценок зависимой переменной y в соответствии с величиной независимого показателя x. Если мы имеем ряд значений признака и основываем оценки только на этих данных, не принимая в расчет каких-либо факторов, то лучшей обобщающей характеристикой ряда является обычно средняя арифметическая. Насколько же надежна средняя арифметическая для суждения об отдельных членах ряда?

Если основная масса индивидуальных значений тесно группируется вокруг средней арифметической, то она более надежна, если же индивидуальные значения признака колеблются в широких пределах, то оценки с помощью средней арифметической окажутся слишком грубыми. Для характеристики колеблемости индивидуальных значений определяют сумму квадратов отклонений индивидуальных значений признака от средней арифметической и делят эту сумму на количество членов ряда; полученный показатель называют дисперсией. Для перехода к первой степени из величины дисперсии извлекается квадратный корень; получается величина, называемая средним квадратическим отклонением.

Показатели дисперсии и среднего квадратического отклонения и служат для численного выражения надежности средней арифметической. В данном примере среднегодовая выработка составляет 5,307 тыс. руб. а среднее квадратическое отклонение — 0,96 тыс. руб. Вариация весьма велика; по отношению к величине самой средней арифметической она составляет более 18 %. Следовательно, среднегодовая выработка не дает надежной оценки выработки за отдельные годы.

С помощью уравнения прямой показана связь между изменением выработки и изменением фондовооруженности труда. Таким образом, получена новая возможность для оценки производительности труда. Насколько же оценки по уравнению надежнее оценок с помощью средней?

Подставляя в уравнение действительные значения x по каждому году, можно рассчитать соответствующие теоретические значения y и сравнить их с фактическими. Вычисленные по уравнению значения показаны в графе 4 табл. 1.1., а в графе 5 приведены отклонения фактических значений от расчетных.

Определим, насколько все отклонения от средней величины превышают отклонения от прямой. Воспользуемся, как и прежде, квадратами отклонений. Если возвести в квадрат каждое отклонение от средней арифметической, найти сумму этих квадратов и разделить полученную сумму на число членов ряда, то получим величину дисперсии. Поскольку это дисперсия значений y, обозначим ее s_y^2. Для нашего примера $s_y^2 = 0{,}92$. Среднее квадратическое отклонение $s_y = 0{,}96$ (тыс. руб.).

Такие же операции необходимо проделать и с отклонениями от прямой. Новую дисперсию обозначим s_{yx}^2. Для данного примера $s_{yx}^2 = 0{,}0033$. Среднее квадратическое отклонение (его называют обычно стандартным отклонением) $s_{yx} = 0{,}058$ тыс. руб. Это означает, что действительная выработка отклоняется от вычисленной по уравнению в среднем на 58 руб. Как видим, оценки по уравнению в шестнадцать с лишним раз точнее оценок с помощью средней арифметической.

Чтобы определить, насколько сократилась сумма квадратов отклонений при переходе от средней арифметической к уравнению прямой, определим так называемый коэффициент детерминации:

$$\frac{s_y^2 - s_{yx}^2}{s_y^2} = \frac{0{,}92 - 0{,}0033}{0{,}92} = 0{,}996.$$

Извлекая квадратный корень из приведенного выше выражения, получим коэффициент корреляции

$$r = \pm\sqrt{\frac{s_y^2 - s_{yx}^2}{s_y^2}} = \sqrt{0{,}996} = 0{,}998.$$

Уравнение связи и коэффициент корреляции — важнейшие обобщающие характеристики корреляционной зависимости между изучаемыми признаками. Уравнение в конкретной количественной форме показывает, какая существует зависимость между двумя переменными, а коэффициент корреляции позволяет судить о силе этой зависимости, о тесноте связи.

Если в действительности никакой связи между двумя изучаемыми переменными нет, то никакая прямая не даст лучшей характеристики величины y, чем средняя арифметическая. Тогда числитель формулы для коэффициента корреляции равен нулю и $r = 0$. Если между переменными существует очень тесная связь, то коэффициент корреляции близок к единице. Тесная связь между изучаемыми признаками наблюдается и в нашем примере.

Коэффициент корреляции по абсолютной величине может принимать значения от нуля до единицы и иметь положительный или отрицательный знак, тот же, который имеет коэффициент a_1 в уравнении связи.

При $r < 0$ имеет место отрицательная корреляция (т. е. с увеличением независимой переменной величина зависимой переменной уменьшается), при $r > 0$ — положительная.

Часто связь между переменными по своей сущности не может быть охарактеризована уравнением прямой. Так, при изучении связи между размером предприятий и себестоимостью единицы их продукции логично предположить, что до известного предела увеличение размера предприятия способствует снижению себестоимости, а дальнейший его рост приведет к росту себестоимости в связи с усилением действия отрицательных особенностей предприятий-гигантов (сложности в организации и управлении производством, увеличение дальности перевозок сырья и готовой продукции).

Указанная зависимость отражается кривой, имеющей минимум. Отыскание минимума, т. е. оптимального размера предприятия, становится важной задачей самого моделирования. Простейшей кривой, описывающей подобного рода зависимости, является парабола второго порядка, характеризуемая уравнением

$$y = a_0 + a_1 x + a_2 x^2.$$

Вообще любая зависимость между двумя переменными может быть описана с помощью параболы n-го порядка:

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n.$$

Кроме того, часто используются следующие кривые:

$$y = a_0 x^{a_1},\quad y = a_0 + \frac{a_1}{x},\quad y = a_0 + \frac{a_1}{x} + \cdots + \frac{a_n}{x_n},\quad y = a_0 a_1^x.$$

При использовании любой формы криволинейной корреляционной зависимости теснота связи между переменными может быть определена с помощью индекса корреляции (i), который рассчитывается аналогично коэффициенту корреляции, т. е. по формуле

$$i = \sqrt{1 - \frac{s_{yx}^2}{s_y^2}},$$

где s_{yx}^2 — средний квадрат отклонений фактических значений y от значений, вычисленных по уравнению кривой; s_y^2 — средний квадрат отклонений фактических значений y от их средней арифметической величины. Индекс корреляции по величине изменяется от 0 до 1.

Индекс корреляции — условная величина, рассчитываемая лишь по отношению к конкретной форме кривой, и ее абсолютное значение всегда может быть доведено до единицы. Например, если в качестве кривой взять параболу, в уравнении которой количество постоянных коэффициентов равно числу исходных единиц наблюдения, то кривая пройдет через все точки графика, а величина индекса корреляции достигнет единицы. Однако было бы ошибочно считать это признаком выявления кривой, наилучшим образом характеризующей изучаемую зависимость. Слишком сложные уравнения регрессии, как правило, лишены реального экономического содержания, так как в них теряется различие между нетипичным и существенным, а случайность возводится в закономерность. Поэтому усложнение уравнения допустимо лишь в определенных пределах. Параметры уравнения должны поддаваться определенному экономическому толкованию.

Корреляционный анализ основывается обычно на достаточно большой совокупности исходных данных, которая, однако, не охватывает все аналогичные, однородные в качественном отношении данные. При исследовании зависимости между экономическими показателями однотипных предприятий ограничиваются данными лишь части таких предприятий, при условии, что эта часть действительно является представительной по отношению ко всей массе предприятий. Совокупность всех единиц, качественно однородных в отношении изучаемых признаков, называется генеральной совокупностью, а отобранная для анализа и расчета группа единиц наблюдения называется выборочной совокупностью.

Предположим, что в ходе исследования связи между двумя показателями на основании выборочных данных рассчитаны уравнение регрессии и коэффициент корреляции. В дальнейшем оказалось возможным привлечь для анализа все единицы генеральной совокупности и на основании этой более широкой группы данных вновь определить уравнение связи и коэффициент корреляции. Возникает вопрос: совпадут ли по величине параметры и коэффициенты корреляции, вычисленные для выборочной и генеральной совокупности? Полное совпадение может произойти лишь случайно. В целом же характеристики, вычисленные для выборочной совокупности, содержат определенную ошибку по сравнению с соответствующими характеристиками генеральной совокупности. Исчисление вероятных ошибок и оценка параметров генеральной совокупности по данным выборочного расчета являются необходимой составной частью разработки модели.

Рассмотрим вопрос об ошибке вычисленных по уравнению связи величин зависимой переменной y. В качестве меры этой ошибки принимается среднее квадратическое отклонение фактических значений y (s_{yx}). Предположим, что отклонения действительных значений y от прямой линии близки к нормальному распределению. Нормальное распределение, имеющее важное значение в теории статистики, графически представляет собой колоколообразную кривую, которая симметрична относительно своей максимальной ординаты. От точки максимума кривая плавно опускается к оси абсцисс и асимптотически приближается к ней, простираясь в обе стороны в бесконечность. Если провести на графике две параллельные прямые, расположенные на расстоянии s_{yx} по обе стороны от линии регрессии, то в соответствии со свойствами нормального распределения между этими прямыми будет находиться 68 %. на расстоянии $\pm 2s_{yx}$ будет заключено приблизительно 95 %, а в пределах $\pm 2{,}58\,s_{yx}$ — 99 % всех фактических значений y.

Тем же соотношениям между фактическими и расчетными величинами может быть дана иная интерпретация. Если для какой-либо единицы, которая входит в генеральную совокупность, но не входит в выборочную, известно значение $x_{\text{ф}}$, то по уравнению регрессии можно определить для этой единицы расчетное значение — $y_{\text{р}}$. Нельзя ожидать, что величина $y_{\text{р}}$ в точности совпадает с фактическим значением $y_{\text{ф}}$. Однако границы, в которых заключается $y_{\text{ф}}$, могут быть определены достаточно точно.

С вероятностью 0,68 $y_{\text{ф}}$ находится в пределах $y_{\text{р}} \pm s_{yx}$; с вероятностью 0,95 $y_{\text{ф}}$ — в пределах $y_{\text{р}} \pm 2s_{yx}$; наконец, с вероятностью 0,99 справедливо

$$y_{\text{р}} - 2{,}58s_{yx} \leqslant y_{\text{ф}} \leqslant y_{\text{р}} + 2{,}58s_{yx}.$$