Глава 7

Понятие стохастической связи.

Задачи корреляционного анализа.

Способы изучения парной корреляции.

Методика множественного корреляционного анализа.

Методика оценки и практического применения результатов корреляционного анализа.

СПОСОБЫ ИЗУЧЕНИЯ СТОХАСТИЧЕСКИХ (КОРРЕЛЯЦИОННЫХ) СВЯЗЕЙ В АНАЛИЗЕ ХОЗЯЙСТВЕННОЙ ДЕЯТЕЛЬНОСТИ

7.1. Понятие стохастической связи и задачи корреляционного анализа

Сущность стохастических взаимосвязей между показателями. Отличия стохастических связей от функциональных. Способы исследования зависимостей в стохастическом факторном анализе. Условия применения и задачи корреляционного анализа.

В предыдущих главах рассматривалась методика решения задач детерминированного факторного анализа. Однако на практике далеко не все экономические явления и процессы могут изучаться с помощью этой методики, так как в большинстве случаев их нельзя свести к функциональным зависимостям, когда величине факторного показателя соответствует единственная величина результативного показателя.

Чаще в экономических исследованиях встречаются стохастические зависимости, которые отличаются приблизительностью, неопределенностью. Они проявля-

ются только в среднем по значительному количеству объектов (наблюдений). Здесь каждой величине факторного показателя (аргумента) может соответствовать несколько значений результативного показателя (функции). Например, увеличение фондовооруженности труда рабочих дает разный прирост производительности труда на разных предприятиях даже при очень выравненных прочих условиях. Это объясняется тем, что все факторы, от которых зависит производительность труда, действуют в комплексе, взаимосвязанно. В зависимости от того, насколько оптимально сочетаются разные факторы, будет неодинаковой степень воздействия каждого из них на величину результативного показателя.

Взаимосвязь между исследуемыми факторами и результативным показателем проявится, если взять для исследования большое количество наблюдений (объектов) и сравнить их значения. Тогда в соответствии с законом больших чисел влияние других факторов на результативный показатель сглаживается, нейтрализуется. Это дает возможность установить связь, соотношения между изучаемыми явлениями.

Значит, ***корреляционная (стохастическая) связь — это неполная, вероятностная зависимость между показателями, которая проявляется только в массе наблюдений.*** Отличают парную и множественную корреляцию.

Парная корреляция — это связь между двумя показателями, один из которых является факторным, а другой — результативным. ***Множественная корреляция*** возникает от взаимодействия нескольких факторов с результативным показателем.

Для исследования стохастических зависимостей используются следующие способы экономического анализа, с которыми мы уже знакомились в предыдущих главах: сравнение параллельных и динамических рядов, аналитические группировки, графики. Однако они позволяют выявить только общий характер и направление связи. Основная же задача факторного анализа — определить степень влияния каждого фактора на уровень результативного показателя. Для этой цели применяются *способы корреляционного, дисперсионного, компонентного, дискриминантного, современного многомерного факторного анализа* и т.д.

Наиболее широкое применение в АХД нашли приемы корреляционного анализа, которые позволяют количественно выразить взаимосвязь между показателями.

Необходимые условия применения корреляционного анализа.

1. Наличие достаточно большого количества наблюдений о величине исследуемых факторных и результативных показателей (в динамике или за текущий год по совокупности однородных объектов).

2. Исследуемые факторы должны иметь количественное измерение и отражение в тех или иных источниках информации.

Применение корреляционного анализа позволяет решить следующие задачи:

1) определить изменение результативного показателя под воздействием одного или нескольких факторов (в абсолютном измерении), то есть определить, на сколько единиц изменяется величина результативного показателя при изменении факторного на единицу;

2) установить относительную степень зависимости результативного показателя от каждого фактора.

Исследование корреляционных зависимостей имеет огромное значение в АХД. Это проявляется в том, что значительно углубляется факторный анализ, устанавливаются место и роль каждого фактора в формировании уровня исследуемых показателей, углубляются знания об изучаемых явлениях, определяются закономерности их развития и как итог — точнее обосновываются планы и управленческие решения, более объективно оцениваются итоги деятельности предприятий и более полно определяются внутрихозяйственные резервы.

7.2. Использование способов парной корреляции для изучения стохастических зависимостей

Формы стохастической связи. Приемы обоснования уравнения связи. Порядок расчета параметров уравнения прямой, параболы, гиперболы. Методика расчета коэффициентов корреляции при прямолинейной и криволинейной формах зависимости. Интерпретация результатов корреляционно-регрессионного анализа.

Одной из основных задач корреляционного анализа является определение влияния факторов на величину результативного показателя (в абсолютном измерении). Для решения этой задачи подбирается соответствующий тип математического уравнения, которое наилучшим образом отражает характер изучаемой связи (прямолинейной, криволинейной и т.д.). Это играет важную роль в корреляционном анализе, потому что от правильного выбора уравнения регрессии зависит ход решения задачи и результаты расчетов.

Обоснование уравнения связи делается с помощью сопоставления параллельных рядов, группировки данных и линейных графиков. Размещение точек на графике покажет, какая зависимость образовалась между изучаемыми показателями: прямолинейная или криволинейная.

Наиболее простым уравнением, которое характеризует прямолинейную зависимость между двумя показателями, является уравнение прямой:

$$Y_x = a + bx, \tag{7.1}$$

где x — факторный показатель; Y — результативный показатель; a и b — параметры уравнения регрессии, которые требуется отыскать.

Это уравнение описывает такую связь между двумя признаками, при которой с изменением факторного показателя на определенную величину наблюдается равномерное возрастание или убывание значений результативного показателя. В качестве примера для иллюстрации корреляционного анализа прямолинейной зависимости могут быть использованы сведения об изменении урожайности зерновых культур (Y) в зависимости от качества пахотной земли (x) (см. табл. 4.7).

Значения коэффициентов a и b находят из системы уравнений, полученных по способу наименьших квадратов. В данном случае система уравнений имеет следующий вид:

$$\begin{cases} na + b\sum x = \sum y; \\ a\sum x + b\sum x^2 = \sum xy, \end{cases} \tag{7.2}$$

где n — количество наблюдений (в нашем примере — 20).

Значения $\sum x$, $\sum y$, $\sum xy$, $\sum x^2$ рассчитываются на основе фактических исходных данных (табл. 7.1).

Таблица 7.1

Расчет производных величин для определения параметров уравнения связи и коэффициента корреляции

n	x	y	xy	x^2	y^2	Y_x
1	32	19,5	624	1024	380,25	19,8
2	33	19,0	627	1089	361,00	20,2
3	35	20,5	717	1225	420,25	21,0
...	...	...	...	...	...	...
20	60	33,0	1980	3600	1089,00	31,0
Итого	900	500,0	22 900	41 500	12 860,00	500,0

Подставив полученные значения в систему уравнений, получим

$$\begin{cases} 20a + 900b = 500; \\ 900a + 41\,500b = 22\,900. \end{cases}$$

Умножив все члены первого уравнения на 45 (900/20), получим следующую систему уравнений:

$$\begin{cases} 900a + 40\,500b = 22\,500; \\ 900a + 41\,500b = 22\,900. \end{cases}$$

Отнимем от второго уравнения первое. Отсюда $1000b = 400$; $b = 0,4$,

$$a = \frac{500 - (900 \times 0,4)}{20} = 7,0.$$

Таким образом, уравнение связи, которое описывает зависимость урожайности от качества почвы, будет иметь вид

$$Y_x = 7,0 + 0,4x.$$

Коэффициент a — постоянная величина результативного показателя, которая не связана с изменением данного фактора. *Параметр b* показывает среднее изменение результативного показателя с повышением или понижением величины фактора на единицу его измерения. В данном примере с увеличением качества почвы на один балл урожайность зерновых культур повышается в среднем на 0,4 ц/га.

Подставив в уравнение регрессии соответствующие значения x, можно определить выравненные (теоретические) значения результативного показателя (Y) для каждого хозяйства. Например, чтобы рассчитать урожайность зерновых культур для первого хозяйства, где качество почвы оценивается 32 баллами, необходимо это значение подставить в уравнение связи:

$$Y_x = 7 + 0,4 \times 32 = 19,8 \text{ ц/га}.$$

Полученная величина показывает, какой была бы урожайность при качестве почвы 32 балла, если бы данное хозяйство использовало свои производственные возможности в такой степени, как в среднем все хозяйства района. Аналогичные расчеты сделаны для каждого хозяйства. Данные приведены в последней графе табл. 7.1. Сравнение фактического уровня урожайности с расчетным позволяет оценить результаты работы отдельных предприятий.

По такому же принципу решается уравнение связи ***при криволинейной зависимости между изучаемыми явлениями***. Если при увеличении одного показателя значения другого возрастают до определенного уровня, а потом начинают снижаться (например, зависимость производительности труда рабочих от их возраста), то для записи такой зависимости лучше всего подходит парабола второго порядка:

$$Y_x = a + bx + cx^2. \tag{7.3}$$

В соответствии с требованиями метода наименьших квадратов для определения параметров a, b и c необходимо решить следующую систему уравнений:

$$\begin{cases} na + b\sum x + c\sum x^2 = \sum y; \\ a\sum x + b\sum x^2 + c\sum x^3 = \sum xy; \\ a\sum x^2 + b\sum x^3 + c\sum x^4 = \sum x^2 y. \end{cases} \tag{7.4}$$

Значения $\sum x, \sum y, \sum xy, \sum x^2y, \sum x^2, \sum x^3, \sum x^4$ находят на основании исходных данных (табл. 7.2).

Таблица 7.2

Зависимость производительности труда (y) от возраста работников (x)

Средний возраст по группе (x)	Среднемесячная выработка (y)	$x/10$	xy	x^2	x^2y	x^3	x^4	y_x
20	4,2	2.0	8,4	4,00	16,8	8,00	16	3,93
25	4.8	2.5	12.0	6,25	30,0	15,62	39	4,90
30	5,3	3.0	15,9	9,00	47,7	27,00	81	5,55
35	6.0	3.5	21,0	12,25	73,5	42,87	150	5,95
40	6,2	4.0	24,8	16,00	99,2	64,00	256	6,05
45	5,8	4.5	26,1	20,25	117,4	91,13	410	5,90
50	5.3	5.0	26.5	25,00	132,5	125,00	625	5,43
55	4,4	5.5	24.2	30,25	133,1	166,40	915	4,78
60	4.0	6.0	24,0	36,00	144,0	216,00	1296	3,70
Всего	46.0	36.0	183,0	159,00	794,0	756,00	3788	46,00

Подставив полученные значения в систему уравнений, получим

$$\begin{cases} 9a + 36b + 159c = 46; \\ 36a + 159b + 756c = 183; \\ 159a + 756b + 3788c = 794. \end{cases}$$

Параметры a, b и c находят способом определителей или способом исключения. Используем способ определителей.

Сначала найдем общий определитель:

$$\Delta = \begin{vmatrix} 9 & 36 & 159 \\ 36 & 159 & 756 \\ 159 & 756 & 3788 \end{vmatrix} =$$

$$= 9 \times 159 \times 3788 + 36 \times 756 \times 159 + 36 \times 756 \times 159 - 159^3 - 36^2 \times 3788 - 756^2 \times 9 = 2565;$$

затем частные определители Δa, Δb и Δc:

$$\Delta a = \begin{vmatrix} 46 & 36 & 159 \\ 183 & 159 & 756 \\ 794 & 756 & 3788 \end{vmatrix} = -6846;$$

$$\Delta b = \begin{vmatrix} 9 & 46 & 159 \\ 36 & 183 & 756 \\ 159 & 794 & 3788 \end{vmatrix} = 11349;$$

$$\Delta c = \begin{vmatrix} 9 & 36 & 46 \\ 36 & 159 & 183 \\ 159 & 756 & 794 \end{vmatrix} = -1440.$$

Отсюда

$$a = \frac{\Delta a}{\Delta} = \frac{-6846}{2565} = -2{,}67; \quad b = \frac{\Delta b}{\Delta} = \frac{11349}{2565} = 4{,}424;$$

$$c = \frac{\Delta c}{\Delta} = \frac{-1440}{2565} = -0{,}561.$$

Уравнение параболы будет иметь следующий вид:

$$Y_x = -2{,}67 + 4{,}424x - 0{,}561x^2.$$

Параметры полученного уравнения экономического смысла не имеют. Если подставить в данное уравнение соответствующие значения x, то получим выравненные значения производительности труда в зависимости от возраста рабочих. Результаты приведены в последней графе табл. 7.2.

Из таблицы видно, что производительность труда рабочих повышается до 40-летнего возраста, после чего начинает снижаться. Значит, те предприятия, которые имеют больше работников 30–40-летнего возраста, будут иметь и более высокие показатели производительности труда при прочих равных условиях. Этот фактор необходимо учитывать при планировании уровня производительности труда и при подсчете резервов ее роста.

Довольно часто в экономическом анализе для записи криволинейных зависимостей используется гипербола

$$Y_x = a + \frac{b}{x}. \qquad (7.5)$$

Для определения ее параметров необходимо решить следующую систему уравнений:

$$\begin{cases} na + b\sum \frac{1}{x} = \sum y; \\ a\sum \frac{1}{x} + b\sum \left(\frac{1}{x}\right)^2 = \sum \left(\frac{1}{x}\right) y. \end{cases} \quad (7.6)$$

Гипербола описывает такую зависимость между двумя показателями, когда при увеличении одной переменной значения другой увеличиваются до определенного уровня, а потом прирост снижается, например, зависимость урожайности от количества внесенного удобрения, продуктивности животных от уровня их кормления, себестоимости продукции от объема производства и т.д.

При более сложном характере зависимости между изучаемыми явлениями используются более сложные параболы (третьего, четвертого порядка и т.д.), а также квадратические, степенные, показательные и другие функции.

Таким образом, используя тот или иной тип математического уравнения, можно определить степень зависимости между изучаемыми явлениями, т.е. узнать, на сколько единиц в абсолютном измерении изменяется величина результативного показателя с изменением факторного на единицу. Однако регрессионный анализ не дает ответа на вопрос: тесная это связь или нет, решающее воздействие оказывает данный фактор на величину результативного показателя или второстепенное?

Для измерения тесноты связи между факторными и результативными показателями определяется коэффициент корреляции.

В случае ***прямолинейной формы связи*** между изучаемыми показателями коэффициент корреляции рассчитывается по следующей формуле:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \times \left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \text{ или} \quad (7.7)$$

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{\left(\sum x^2 n - \left(\sum x\right)^2\right) \times \left(\sum y^2 n - \left(\sum y\right)^2\right)}}. \quad (7.8)$$

Подставляя значения $\sum xy, \sum x, \sum y, \sum x^2$ и $\sum y^2$ в формулу (7.7), получаем

$$r = \frac{22\,900 - \frac{900 \times 500}{20}}{\sqrt{\left(41\,500 - \frac{900^2}{20}\right) \times \left(12\,860 - \frac{500^2}{20}\right)}} = 0{,}66.$$

Коэффициент корреляции может принимать значения от 0 до ±1. Чем ближе его величина к 1, тем более тесная связь между изучаемыми явлениями, и наоборот. В данном случае величина коэффициента корреляции является существенной (r = 0,66). Это позволяет сделать вывод о том, что плодородие почвы — один из основных факторов, от которого в данном районе зависит уровень урожайности зерновых культур.

Если коэффициент корреляции возвести в квадрат, получим коэффициент детерминации (d = 0,435). Он показывает, что урожайность зерновых культур на 43,5% зависит от качества почвы, а на долю других факторов приходится 56,5% ее прироста.

Что касается измерения тесноты связи при ***криволинейной форме зависимости***, то здесь используется не линейный коэффициент корреляции, а корреляционное отношение:

$$\eta = \sqrt{\frac{\sigma_y^2 - \sigma_{y_x}^2}{\sigma_y^2}}, \quad (7.9)$$

где $\sigma_y^2 = \frac{\sum (y - \bar{y})^2}{n}$; $\quad \sigma_{y_x}^2 = \frac{\sum (y - y_x)^2}{n}$.

Показатель (7.9) является универсальным. Его можно применять при любой форме зависимости. Однако для определения его величины вначале необходимо решить уравнение регрессии и рассчитать выравненные значения результативного показателя (y_x), для чего в полученное уравнение нужно подставить значения x по каждой возрастной группе (табл. 7.3).

Таблица 7.3

Расчет исходных данных для определения корреляционного отношения при криволинейных зависимостях

y	y_x	$y-\overline{y}$	$(y-\overline{y})^2$	$y-y_x$	$(y-y_x)^2$
4.[illegible]	3.93	−0.9	0.81	+0,27	0,073
4.[illegible]	4.90	−0.3	0.09	−0,10	0,010
5.[illegible]	5.55	+0.2	0.04	−0,25	0,062
6.[illegible]	5.95	+0.9	0.81	+0,05	0,003
6.[illegible]	6.05	+1.1	1.21	+0,15	0,022
5.[illegible]	5.90	+0.7	0.49	−0,10	0,010
5.[illegible]	5.43	+0.2	0.04	−0,13	0,017
4.[illegible]	4.78	−0.7	0.49	−0,38	0,144
4.[illegible]	3.70	−1.1	1.21	+0,30	0,090
4[illegible]	46.0	–	5.19	–	0,431

Подставив полученные значения в формулу (7.9), определим величину корреляционного отношения:

$$\eta = \sqrt{\frac{5{,}19 / 9 - 0{,}431 / 9}{5{,}19 / 9}} = \sqrt{\frac{0{,}576 - 0{,}049}{0{,}576}} = 0{,}956.$$

В заключение необходимо отметить, что мы рассмотрели использование способов парной корреляции только на двух примерах. Однако эта методика может быть использована для исследования соотношений между разными экономическими показателями, что позволяет значительно углубить знания об изучаемых явлениях, определить место и роль каждого фактора в изменении уровня исследуемого показателя.

7.3. Методика множественного корреляционного анализа

Необходимость применения многофакторного корреляционного анализа. Этапы многофакторного корреляционного анализа. Правила отбора факторов для корреляционной модели. Обоснование необходимого объема выборки данных для корреляционного анализа. Сбор и статистическая оценка исходной информации. Способы обоснования уравнения связи. Основные показатели связи в корреляционном анализе и их интерпретация. Сущность парных (общих), частных и множественных коэффициентов корреляции и детерминации. Оценка значимости коэффициентов корреляции. Порядок расчета уравнения множественной регрессии. Интерпретация его параметров. Назначение коэффициентов эластичности и стандартизированных бетта-коэффициентов.

Экономические явления и процессы хозяйственной деятельности предприятий зависят от большого количества факторов. Как правило, каждый фактор в отдельности не определяет изучаемое явление во всей полноте. Только комплекс факторов в их взаимосвязи может дать более или менее полное представление о характере изучаемого явления.

Многофакторный корреляционный анализ состоит из нескольких этапов. ***На первом этапе*** определяются факторы, которые оказывают воздействие на изучаемый показатель, и отбираются наиболее существенные для корреляционного анализа.

На втором этапе собирается и оценивается исходная информация, необходимая для корреляционного анализа.

На третьем этапе изучается характер и моделируется связь между факторами и результативным показателем, то есть подбирается и обосновывается математическое уравнение, которое наиболее точно выражает сущность исследуемой зависимости.

На четвертом этапе проводится расчет основных показателей связи корреляционного анализа.

На пятом этапе дается статистическая оценка результатов корреляционного анализа и практическое их применение.

Отбор факторов для корреляционного анализа является очень важным моментом в экономическом анализе. От того, насколько правильно он сделан, зависит точность выводов по итогам анализа. Главная роль при отборе факторов принадлежит теории, а также практическому опыту анализа. При этом необходимо придерживаться следующих правил.

1. При отборе факторов в первую очередь следует учитывать причинно-следственные связи между показателями, так как только они раскрывают сущность изучаемых явлений. Анализ же таких факторов, которые находятся только в математических соотношениях с результативным показателем, не имеет практического смысла.

2. При создании многофакторной корреляционной модели необходимо отбирать самые значимые факторы, которые оказывают решающее воздействие на результативный показатель, так как охватить все условия и обстоятельства практически невозможно. Факторы, которые имеют критерий надежности по Стьюденту меньше табличного, не рекомендуется принимать в расчет.

3. Все факторы должны быть количественно измеримы, т.е. иметь единицу измерения, и информация о них должна содержаться в учете и отчетности.

4. В корреляционную модель линейного типа не рекомендуется включать факторы, связь которых с результативным показателем имеет криволинейный характер.

5. Не рекомендуется включать в корреляционную модель взаимосвязанные факторы. Если парный коэффициент корреляции между двумя факторами больше 0,85, то по правилам корреляционного анализа один из них необходимо исключить, так как это приведет к искажению результатов анализа.

6. Нежелательно включать в корреляционную модель факторы, связь которых с результативным показателем носит функциональный характер.

Большую помощь при отборе факторов для корреляционной модели оказывают аналитические группировки, способ сопоставления параллельных и динамических рядов, линейные графики. Благодаря им можно определить наличие, направление

и форму зависимости между изучаемыми показателями. Отбор факторов можно производить также в процессе решения задачи корреляционного анализа на основе оценки их значимости по критерию Стьюдента, о котором будет сказано ниже.

Исходя из перечисленных выше требований и используя названные способы отбора факторов, ***для многофакторной корреляционной модели уровня рентабельности (Y) подобраны следующие факторы***, которые оказывают наиболее существенное влияние на ее уровень:

x_1 — материалоотдача, руб.;

x_2 — фондоотдача, коп.;

x_3 — производительность труда (среднегодовая выработка продукции на одного работника), тыс. руб.;

x_4 — продолжительность оборота оборотных средств предприятия, дни;

x_5 — удельный вес продукции высшей категории качества, %.

Поскольку корреляционная связь с достаточной выразительностью и полнотой проявляется только в массе наблюдений, объем выборки данных должен быть достаточно большим, так как только в массе наблюдений сглаживается влияние других факторов. Чем большая совокупность объектов исследуется, тем точнее результаты анализа.

Учитывая это требование, влияние перечисленных факторов на уровень рентабельности исследуется на примере 40 предприятий.

Следующим этапом анализа является ***сбор и статистическая оценка исходной информации,*** которая будет использоваться в корреляционном анализе. Собранная исходная информация должна быть проверена на достоверность, однородность и соответствие закону нормального распределения.

В первую очередь необходимо убедиться в ***достоверности информации,*** насколько она соответствует объективной действительности. Использование недостоверной, неточной информации приведет к неправильным результатам анализа и выводам.

Одно из условий корреляционного анализа — ***однородность исследуемой информации*** относительно распределения ее около среднего уровня. Если в совокупности имеются группы объектов, которые значительно отличаются от среднего уровня, то это говорит о неоднородности исходной информации.

Критерием однородности информации служит среднеквадратическое отклонение и коэффициент вариации, которые рассчитываются по каждому факторному и результативному показателю.

Среднеквадратическое отклонение показывает абсолютное отклонение индивидуальных значений от среднеарифметической. Оно определяется по формуле

$$\sigma = \sqrt{\frac{\sum\left(x-\bar{x}\right)^2}{n}}. \qquad (7.10)$$

Коэффициент вариации характеризует относительную меру отклонения отдельных значений от среднеарифметической. Он рассчитывается по формуле

$$V = \frac{\sigma}{\bar{x}} \times 100. \qquad (7.11)$$

Чем больше коэффициент вариации, тем относительно больший разброс и меньшая выравненность изучаемых объектов. Изменчивость вариационного ряда принято считать незначительной, если вариация не превышает 10 %, средней — если составляет 10–20 %, значительной — если она больше 20 %, но не превышает 33 %. Если же вариация выше 33 %, то это говорит о неоднородности информации и необходимости исключения нетипичных наблюдений, которые обычно бывают в первых и последних ранжированных рядах выборки.

В нашем примере (табл. 7.4) самая высокая вариация по x_5 (V = 22,98), но она не превышает 33 %. Значит, исходная информация является однородной и ее можно использовать для дальнейших расчетов.

На основании самого высокого показателя вариации можно определить ***необходимый объем выборки данных*** для корреляционного анализа по следующей формуле:

$$n = \frac{V^2 \times t^2}{m^2} = \frac{22{,}98^2 \times 1{,}96^2}{8^2} = 32, \qquad (7.12)$$

где n — необходимый объем выборки данных; V — вариация, %; t — показатель надежности связи, который при уровне вероятности P = 0,05 равен 1,96; m — показатель точности расчетов (для экономических расчетов допускается ошибка 5–8 %).

Значит, принятый в расчет объем выборки (40 предприятий) является достаточным для проведения корреляционного анализа.

Таблица 7.4

Показатели статистической характеристики исходной информации

Номер переменной	Среднеарифметическое значение	Среднеквадратическое отклонение	Вариация, %	Асимметрия	Эксцесс	Ошибка	
						асимметрии	эксцесса
Y	27,15	2,85	10,5	0,20	−1,16	0,37	0,73
x_1	2,77	0,28	10,08	0,36	−0,81	0,37	0,73
x_2	92,57	8,70	9,39	0,24	−0,69	0,37	0,73
x_3	8,46	0,59	7,00	0,10	−0,52	0,37	0,73
x_4	17,77	2,76	15,55	0,72	−0,08	0,37	0,73
x_5	31,68	7,28	22,98	0,63	−0,13	0,37	0,73

Следующее требование к исходной информации — ***соответствие ее закону нормального распределения***. Согласно этому закону, основная масса исследуемых сведений по каждому показателю должна быть сгруппирована около ее среднего значения, а объекты с очень маленькими значениями или с очень большими должны встречаться как можно реже. График нормального распределения информации имеет следующий вид (рис. 7.1).

Для количественной оценки степени отклонения информации от нормального распределения служит отношение показа-