

ГРАФИЧЕСКИЙ СПОСОБ ПОСТРОЕНИЯ ЭКОНОМИКО-СТАТИСТИЧЕСКИХ МОДЕЛЕЙ

А. А. ФРЕНКЕЛЬ

(МОСКВА)

За последние годы большое значение приобрело применение методов математической статистики в анализе и планировании экономических показателей. Произведены исследования по оценке текущих и по составлению перспективных планов основных качественных показателей работы предприятий: производительности труда, себестоимости и т. д. Построен ряд экономико-статистических моделей для анализа и планирования этих показателей в различных отраслях [1, 2].

В настоящей статье рассматривается графический метод построения экономико-статистических моделей на примере построения модели производительности труда для предприятий станкостроения (по отчетным данным за 1962 г. 52 станкостроительных заводов)*.

Были отобраны следующие шесть факторов, влияющих на уровень производительности труда.

1. Стоимость производственных фондов, рассчитываемая как сумма среднегодовой стоимости основных фондов и среднегодовой стоимости нормируемых оборотных фондов в производстве. Для определения последней из среднегодовой стоимости всех нормируемых оборотных фондов вычиталась среднегодовая стоимость готовой продукции. Готовая продукция вычитается потому, что она представляет собой оборотные средства в сфере обращения.

Показатель стоимости производственных фондов характеризует размер производства.

2. Коэффициент фондовооруженности определяют как отношение среднегодовой стоимости основных фондов (за вычетом среднегодовой стоимости зданий и сооружений) к числу промышленно-производственных рабочих, работающих в наибольшую смену. Число рабочих, работающих в наибольшую смену, находят как отношение среднегодовой численности промышленно-производственных рабочих к коэффициенту сменности.

3. Коэффициент электровооруженности, получаемый как отношение всей потребленной электроэнергии к количеству отработанных за год человеко-часов.

Коэффициенты фондовооруженности и электровооруженности отражают уровень технического развития предприятий.

4. Средний размер серии (предложен В. П. Хайкиным [2]) рассчитывают по формуле

$$K = \frac{\Sigma C}{\Sigma C/q}, \quad (1)$$

*В работе принимал участие Д. Ф. Шутко.

где K — показатель среднего размера серии; C — стоимость данного вида продукции, произведенной на предприятии; q — количество произведенной продукции данного вида.

В расчет не принимают продукцию, по которой нельзя выделить количество изделий в единицах (литье, поковки и т. д.).

Показатель среднего размера серии характеризует специализацию производства.

5. Удельный вес стоимости поступивших в порядке кооперирования полуфабрикатов и изделий во всей стоимости сырья и основных материалов указывает на уровень кооперирования.

6. Удельный вес несравнимой (новой) товарной продукции во всем объеме произведенной предприятием товарной продукции характеризует ассортиментные сдвиги в номенклатуре выпускаемой продукции.

Вышеприведенные шесть факторов были отобраны потому, что, во-первых, они оказывают значительное влияние на уровень производительности труда и, во-вторых, их можно получить из существующей отчетности предприятий.

Введем следующие обозначения: y — выработка на одного работающего в руб.; x_1 — стоимость производственных фондов в тыс. руб.; x_2 — коэффициент фондовооруженности; x_3 — коэффициент электровооруженности; x_4 — средний размер серий в шт.; x_5 — удельный вес стоимости поступивших в порядке кооперирования полуфабрикатов и изделий во всей стоимости сырья и основных материалов в %; x_6 — удельный вес несравнимой товарной продукции во всей стоимости товарной продукции в %.

Математически задачу можно сформулировать следующим образом. Требуется найти аналитическое выражение, показывающее, как величина y зависит от величин x_1, x_2, \dots, x_6 (факторов, ее определяющих), т. е. необходимо найти функцию

$$y = f(x_1, x_2, \dots, x_6). \quad (2)$$

Естественно, что имеется бесконечное множество функций, которые при значениях x_i будут представлять значение функции y . Поэтому тип функции можно выбирать до некоторой степени произвольно (в определенном классе).

Предположим, что функция y может выражаться произведением некоторых функций, каждая из которых зависит только от одного фактора:

$$y = c f_1(x_1) f_2(x_2), \dots, f_6(x_6), \quad (3)$$

где y, x_1, x_2, \dots, x_6 — параметры, упомянутые выше; c — константа, соответствующая среднему значению параметра y .

Уравнение (3) охватывает большой класс зависимостей величины y от факторов x_1, x_2, \dots, x_6 . Поэтому, учитывая некоторую ограниченность диапазона изменений показателей x_1, x_2, \dots, x_6 , это уравнение можно широко применять для аппроксимации реальных связей в изучаемом явлении.

Каждая из функций $f_i(x_i)$ может быть как линейной, так и нелинейной.

Для решения поставленной задачи требуется найти величину c и аналитическое выражение всех функций $f_i(x_i)$.

Методы нахождения этих параметров были описаны американским статистиком Д. Брандоном [3].

Сначала для упрощения расчетов заменим y его нормализованным значением y_0 , определяемым как

$$y_0 = y / \bar{y}, \quad (4)$$

где \bar{y} — среднее значение y .

\bar{y} в общей модели соответствует c . В некоторых работах [4, 5] сначала находятся все значения $f_i(x_i)$, а затем уже c , определяемая тоже как средняя арифметическая величина. Другими словами, рассматриваемый метод и метод, примененный в работах [4, 5], совершенно эквивалентны.

После того как найдено значение y_0 для каждого завода, строят график зависимости между y_0 и x_1 . На корреляционном поле проводят эмпирическую линию регрессии, которую выравнивают теоретической линией регрессии. Параметры последней определяют методом наименьших квадратов.

На основании полученных данных находят аналитическое выражение зависимости y_0 от x_1 , т. е. $f_1(x_1)$ для каждого завода. После этого рассчитывают условный показатель (для каждого завода):

$$y_1 = \frac{y_0}{f_1(x_1)}. \quad (5)$$

Этот остаточный результат уже будет зависеть не от изменения x_1 , а только от величин x_2, \dots, x_6 .

$$y_1 = \frac{y}{cf_1(x_1)} = f_2(x_2), \dots, f_6(x_6). \quad (6)$$

Затем строят график зависимости этого условного показателя y_1 от x_2 и определяют уравнение регрессии $f_2(x_2)$. Находят новый условный показатель

$$y_2 = \frac{y_1}{f_2(x_2)} = \frac{y}{cf_1(x_1)f_2(x_2)} = f_3(x_3), \dots, f_6(x_6). \quad (7)$$

Такой расчет будет продолжаться до тех пор, пока не определены все функции $f_i(x_i)$. В нашем примере последним рассчитывается условный показатель

$$y_5 = \frac{y_4}{f_5(x_5)} = \frac{y}{cf_1(x_1)f_2(x_2)f_3(x_3)f_4(x_4)f_5(x_5)} \quad (8)$$

и находится аналитическое значение $f_6(x_6)$.

Ниже приводится методика построения модели производительности труда для предприятий станкостроения.

Сначала находим $\bar{y} = 3573$. Для каждого завода определяем значение y_0 .

Строим график зависимости y_0 от стоимости производственных фондов (x_1) (рис. 1).

На графике эта зависимость выражается уравнением регрессии

$$f_1(x_1) = 0,972 + 0,000004x_1. \quad (9)$$

Теоретическая линия регрессии изображена сплошной линией. Пунктиром проведена эмпирическая линия регрессии.

На остальных графиках эти обозначения будут сохраняться.

Рассчитываем значения $f_1(x_1)$ для всех заводов, подставляя в уравнение регрессии вместо x_1 его конкретные значения.

После этого находим для каждого завода $y_1 = \frac{y_0}{f_1(x_1)}$ и строим гра-

фик зависимости y_1 от x_2 (рис. 2).

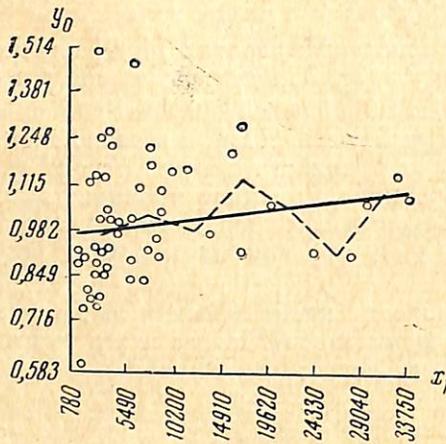


Рис. 1. График зависимости y_0 от стоимости производственных фондов

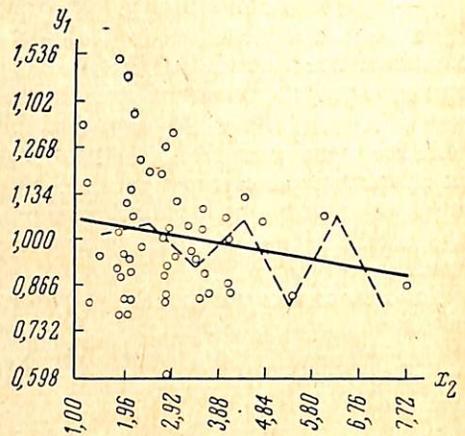


Рис. 2. График зависимости y_1 от коэффициента фондовооруженности

Теоретическая линия регрессии имеет вид

$$f_2(x_2) = 1,068 - 0,022x_2. \quad (10)$$

Расчет $f_i(x_i)$ продолжают до тех пор, пока не будет найдено значение $f_6(x_6)$.

Ниже приведены графики зависимостей условных показателей y_2, y_3, y_4, y_5 от факторов x_3, x_4, x_5, x_6 (рис. 3—6).

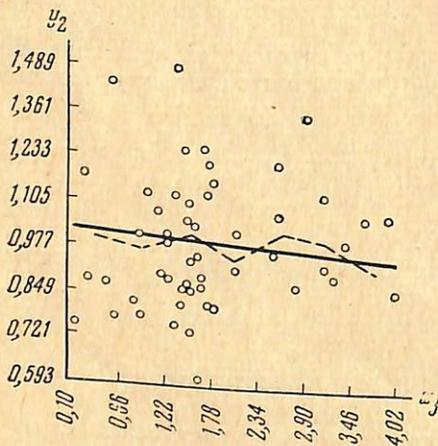


Рис. 3. График зависимости y_2 от коэффициента электровооруженности

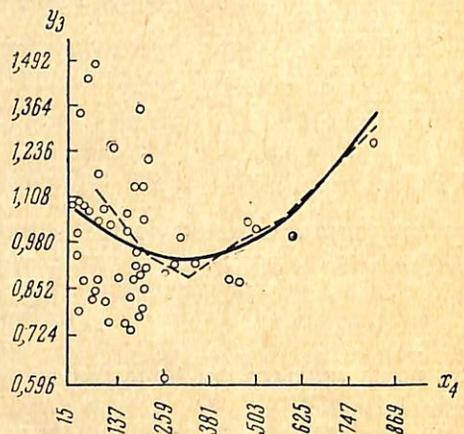


Рис. 4. График зависимости y_3 от среднего размера серии

Уравнение теоретической линии регрессии (рис. 3) равно

$$f_3(x_3) = 0,997 - 0,0016x_3. \quad (11)$$

Уравнение теоретической линии регрессии (рис. 4) описано параболой

$$f_4(x_4) = 1,105 - 0,001x_4 + 0,000016x_4. \quad (12)$$

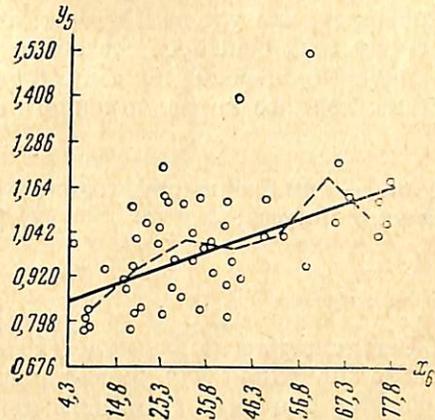
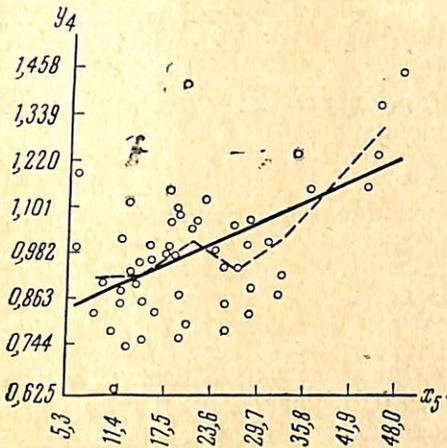


Рис. 5. График зависимости y_4 от удельного веса стоимости поступивших в порядке кооперирования полуфабрикатов и изделий во всей стоимости сырья и основных материалов

Рис. 6. График зависимости y_5 от удельного веса несравнимой товарной продукции во всей стоимости товарной продукции

Уравнение теоретической линии регрессии (рис. 5) имеет вид

$$f_5(x_5) = 0,796 + 0,009x_5. \tag{13}$$

Уравнение теоретической линии регрессии (рис. 6) описано прямой

$$f_6(x_6) = 0,859 + 0,004x_6. \tag{14}$$

Модель производительности труда для предприятий станкостроения имеет следующий вид:

$$y = 3573 (0,972 + 0,000004x_1) \times (1,068 - 0,022x_2) \times (0,996 - 0,0016x_3) \times (1,105 - 0,001x_4 + 0,000016x_4^2) \times (0,796 + 0,009x_5) \times (0,859 + 0,004x_6). \tag{15}$$

Для определения тесноты связи между выработкой на одного работающего и факторами, влияющими на ее уровень, было вычислено множественное корреляционное отношение по формуле

$$\eta^2 = \sqrt{\frac{S_{x_1, x_2, \dots, x_6}^2}{\sigma_y^2}}, \tag{16}$$

где σ_y^2 — общая дисперсия y , показывающая влияние на y всех факторов, в том числе и отобранных; $S_{x_1, x_2, \dots, x_6}^2$ — дисперсия теоретической поверхности регрессии, характеризующая влияние на y только отобранных факторов x_1, x_2, \dots, x_6 .

$S_{x_1, x_2, \dots, x_6}^2$ определяют по формуле:

$$S_{x_1, x_2, \dots, x_6}^2 = \frac{\sum (\tilde{y} - \bar{\tilde{y}})^2}{n - 1}, \tag{17}$$

где \tilde{y} — значение выработки, рассчитанное по модели; $\bar{\tilde{y}}$ — среднее значение расчетной выработки; $n - 1$ — число степеней свободы (n — число заводов).

Множественное корреляционное отношение оказалось равным 0,843, что указывает на достаточно высокую степень тесноты связи между выработкой и отобранными факторами.

Квадрат множественного корреляционного отношения, равный 0,7119, показывает, что удельный вес влияния отобранных факторов на выработку среди всех факторов составляет 71,19%.

Была проверена по t -критерию (критерию Сьюдента) значимость множественного корреляционного отношения

$$t = \eta / m_{\eta}, \quad (18)$$

где m_{η} — ошибка множественного корреляционного отношения, найденная из формулы:

$$m_{\eta} = \frac{1 - \eta^2}{\sqrt{n - k - 1}}. \quad (19)$$

Вычисленное по формуле (18) значение t оказалось равным 36,1. Табличное значение t для уровня значимости 0,01 равно 2,58. Таким образом, расчетное значение t намного превышает табличное. Отсюда можно сделать вывод, что множественное корреляционное отношение значимо.

Адекватность модели измеряется соответствием между расчетными и фактическими значениями выработки, т. е. точностью аппроксимации. Для оценки адекватности можно пользоваться показателем средней ошибки аппроксимации [6], определяемой по формуле

$$\bar{\varepsilon} = \frac{1}{n} \sum \frac{|y - \tilde{y}|}{y} \cdot 100\%. \quad (20)$$

Для построенной модели показатель ε равен 6,24%. Это говорит о том, что модель достаточно точно воспроизводит изучаемую совокупность

Таблица 1

y	$u_0 = \frac{y}{y}$	x_1	$f_1(x_1)$	$u_1 = \frac{y_0}{f_1(x_1)}$	x_2	$f_2(x_2)$	$u_2 = \frac{u_1}{f_2(x_2)}$	x_3	$f_3(x_3)$	$u_3 = \frac{u_2}{f_3(x_3)}$
3391	0,949	3917	0,989	0,959	3,3	0,996	0,962	2,5	0,993	0,968
4139	1,158	33749	1,107	1,046	4,7	0,965	1,084	3,9	0,991	1,094
3931	1,100	9143	1,009	1,091	2,3	1,018	1,071	1,1	0,996	1,075
3172	0,886	6962	1,000	0,886	1,7	1,031	0,859	2,8	0,993	0,863
3211	0,898	4161	0,988	0,909	2,3	1,018	0,892	2,1	0,994	0,897
4063	1,137	1778	0,979	1,162	1,1	1,044	1,113	1,5	0,995	1,118
3159	0,884	780	0,975	0,906	1,4	1,038	0,872	0,5	0,596	1,042
3301	0,923	2484	0,982	0,940	2,0	1,024	0,918	1,2	0,995	0,992
3791	1,060	20863	1,055	1,004	3,9	0,982	1,022	2,1	0,994	1,028
3207	0,897	956	0,976	0,919	1,9	1,026	0,895	1,3	0,995	0,899

Таблица 1 (продолжение)

x_4	$f_4(x_4)$	$u_4 = \frac{u_3}{f_4(x_4)}$	x_5	$f_5(x_5)$	$u_5 = \frac{u_4}{f_5(x_5)}$	x_6	$f_6(x_6)$	\tilde{y}	Отклонения в %
181	0,976	0,991	27,8	1,046	0,947	18,2	0,932	3326	-1,9
19	1,086	1,007	15,3	0,934	1,078	64,2	1,116	4277	3,3
155	0,988	1,088	18,4	0,962	1,130	73,1	1,154	4002	1,8
465	0,986	0,875	11,2	0,897	0,988	35,1	0,999	3234	2,0
187	0,974	0,921	17,2	0,951	0,968	33,7	0,994	3287	2,4
205	0,967	1,156	43,2	1,185	0,975	27,8	0,970	4037	-0,6
166	0,978	1,064	28,6	1,053	1,010	4,3	0,876	3248	2,8
183	0,975	0,945	14,9	0,930	1,016	35,4	1,001	3248	-1,6
496	1,002	1,025	11,8	0,901	1,136	67,0	1,127	3748	-1,3
262	0,953	0,943	27,6	1,044	0,903	16,2	0,924	3273	2,1

(в экономических исследованиях модель можно считать адекватной, если средняя ошибка аппроксимации не превышает 10%).

Вместе с тем была предпринята попытка улучшить адекватность модели путем построения линейной модели зависимости отклонений $y - \hat{y}$ от отобранных факторов.

Однако расчет показал, что в пределах точности, допускаемой программой для «Урал-2», матрица системы нормальных уравнений, решаемая методом наименьших квадратов, оказалась вырожденной. Это может служить подтверждением того, что поверхность множественной линейной регрессии тождественна 0, а отклонения не зависят от отобранных факторов.

Модель (15) может быть применена для проверки существующих планов и перспективного планирования производительности труда в станкостроении. Для этого надо в модель подставить вместо неизвестных соответствующие плановые показатели текущего или будущего периода.

Достоинством модели является легкость ее построения ручным способом в короткие сроки. Это, однако, не исключает использование ЭВМ для построения теоретических линий регрессии.

Модель (15) была использована для проверки планов по выработке на 1963 г. ряда станкостроительных заводов. Например, на Стерлитамакском станкостроительном заводе им. В. И. Ленина установленный совнархозом план составлял 3294 руб. на одного работающего. План, рассчитанный по модели (15), — 3327 руб. Фактическая выработка за 1963 г. на заводе равнялась 3352 руб.

Пример построения модели (15) для десяти предприятий см. в табл. 1.

ЛИТЕРАТУРА

1. Статистические модели и методы в экономическом анализе и планировании. Сб. научн. тр. Новосибирск, 1963.
2. В. П. Хайкин, В. С. Найденов, С. Г. Галуза. Корреляция и статистическое моделирование в экономических расчетах. М., «Экономика», 1964.
3. D. V. Brandon. Developing mathematical models for computer control. ISA J. 1959, No. 7.
4. Б. В. Вольтер, Е. В. Маркова. Исследование процесса полимеризации этилена методом множественной корреляции. Химическая промышленность, 1961, № 4.
5. А. А. Френкель. Некоторые вопросы применения математико-статистических методов и ЭВМ для построения моделей производительности труда. В сб. Применение электронной вычислительной техники в учетно-плановых и статистических работах. Моск. эконом.-статист. ин-т (ротапринт), 1963.
6. В. Коржевич, И. Хавин. Применение математических методов и электронной вычислительной техники для анализа себестоимости продукции. Вестн. статистики, 1964, № 12.

Поступила в редакцию
4 IX 1964