Министерство образования и науки Украины
ДОНБАССКАЯ ГОСУДАРСТВЕННАЯ МАШИНОСТРОИТЕЛЬНАЯ АКАДЕМИЯ

Л.В.Васильева, Е.А.Клеваник

ЭКОНОМЕТРИКА: НАЧАЛЬНЫЙ КУРС

ПОСТРОЕНИЕ ЛИНЕЙНЫХ И НЕЛИНЕЙНЫХ МОДЕЛЕЙ. СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

Учебное пособие
для студентов высших учебных заведений

Утверджено
на заседании
Ученого совета ДГМА.
Протокол № 6 от 24.02.05г.

Краматорск 2005

ББК 60.6
УДК 330.43(075.8)
В 19

Рецензенты:

А.А.Каргин, д.т.н., проф. (Донецкий национальный университет);
П.И.Сагайда, к.т.н, доц. (Донбасская государственная машиностроительная академия).

Васильева Л.В., Клеваник Е.А.

В19 Эконометрика: начальный курс. Построение линейных и нелинейных моделей. Системы одновременных уравнений: Учебное пособие для студентов высших учебных заведений – Краматорск: ДГМА, 2005. – 100 с.

ISBN 966-7851-87-7

Учебное пособие содержит теоретические сведения и практическую часть по следующим разделам эконометрики: линейная и нелинейная однофакторная регрессия, проверка адекватности модели, доверительный интервал и доверительная область для линейной и нелинейной регрессии, прогноз по выбранной модели; модель многофакторной регрессии, коллинеарность и мультиколлинеарность факторов; эластичность модели; системы одновременных уравнений, эндогенные и экзогенные переменные.

Пособие рассчитано на студентов и аспирантов экономических специальностей, а также будет полезно лицам, желающим самостоятельно освоить эконометрические расчеты.

ББК 60.6

ISBN 966-7851-87-7

СОДЕРЖАНИЕ

ЧАСТЬ 1

1 ПРЕДМЕТ ЭКОНОМЕТРИКИ

Потребность в способах статистического анализа данных в экономической практике очень большая. Для успешного функционирования в условиях жесткой конкуренции предприятия, банки, страховые компании испытывают потребность в анализе имеющейся информации и получении обоснованных выводов. Анализ такой информации осуществляется с помощью методов, объединенных в дисциплину «Эконометрика».

Буквальный перевод слова «эконометрика» означает «измерение экономики».

Эконометрика – это наука, которая изучает количественные закономерности и взаимосвязи экономических объектов и процессов с помощью математико-статистических методов и моделей. Т.е. эконометрика восстанавливает неизвестные экономико-математические зависимости по статистическим данным и рассматривает возможность использования этих моделей в экономических исследованиях.

Модель – это искусственное воссоздание некоторого экономического процесса для исследований. В эконометрии под моделью подразумевают математическую модель, т.е. описание экономического процесса с помощью математических формул.

Эконометрические модели количественно описывают связь между входными факторами экономической системы X и результирующим показателем (откликом) Y плюс влияние случайной компоненты ε.

По модели получают прогноз.

Прогноз – это расчет неизвестного показателя по заданным факторам на основе модели.

1.1 Основные задачи эконометрики

Эконометрика должна решать пять *основных задач*:

1 *Выбор конкретного вида функции для некоторого экономического процесса.* Например, зависимость между доходом и расходом можно описать так: $y = \beta_1 x + \beta_0$.

2 *Сбор и подготовка экономической информации.* Важно выбрать правильные обозначения для переменных и правильные единицы измерения. Например, если речь идет об изменении дохода с течением времени, то функция будет иметь вид y = f(t), где y – доход, t – время. Если речь идет о национальном доходе, то в качестве единиц измерения принимаем: для y – млн. грн., для t – год. Если речь идет о предприятии, то для y – грн., для t – месяц.

3 *Оценка на основании имеющихся статистических данных значений параметров модели.*

4 *Проверка модели на адекватность, оценка качества выбранной модели, ее простоты, точности описания данных.*

5 *Экономический анализ модели.*

1.2 Этапы эконометрического анализа

Чтобы провести эконометрический анализ, нужно:

1 Выдвинуть гипотезу о виде зависимости по статистическим данным в соответствии с набором факторов.

2 Провести оценку неизвестных параметров модели.

3 Проверить модель на адекватность.

4 Использовать модель в экономических прогнозах и исследованиях.

1.3 Классификация эконометрических моделей

1 Однофакторные $y = f(x)$.

a) Линейные вида $y = b_0 + b_1 x$.

b) Нелинейные:

1) сводящиеся к линейным;

2) существенно нелинейные.

2 Многофакторные $y = f(x_1, x_2, ..., x_p)$.

a) Линейные вида $y = b_0 + b_1 x_1 + ... b_p x_p$.

b) Нелинейные:

1) сводящиеся к линейным;

2) существенно нелинейные.

1.4 Информационная база эконометрики

Решение задач эконометрики проводится на базе статистических данных. *Статистические данные* – это данные, собранные на реальных экономических объектах.

В эконометрике статистические данные можно подразделить на 2 типа: динамические (временные) и вариационные ряды.

Временные ряды – это последовательность наблюдений за одним и тем же процессом или явлением в различные промежутки времени. Например, данные о динамике уровня инфляции за определенный период.

Вариационные ряды – последовательность наблюдений по какому-либо экономическому показателю для разных однотипных объектов. Все замеры производятся в одно и то же время. Значения вариационного ряда располагают в порядке возрастания.

1.5 Обработка информационных данных

Совокупность данных динамических и вариационных рядов обрабатывается по правилам, разработанным в математической статистике.

Генеральная совокупность – все возможные реализации интересующего нас показателя. На практике наблюдаем случайно выбранные значения этого показателя (*выборка*). По генеральной совокупности можно получить точные значения параметров, по выборке – приближенные, или оценки.

Объем выборки n – суммарное количество наблюдений. Объемы выборок могут быть небольшими ($n \le 10$), большими ($n \approx 100$) и очень большими ($n \approx 10^4$). На практике чаще всего приходится иметь дело с большими и очень большими выборками, поэтому расчет проводится с помощью компьютера.

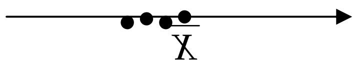
Во всех случаях всю совокупность выборочных данных x_i ($i=1\ldots n$) стараются охарактеризовать некоторыми усредненными параметрами, которые учитывают особенности выборки. По выборкам производится расчет *основных статистических характеристик*:

1 Среднее значение $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

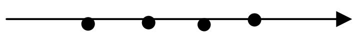
2 Вариация (дисперсия) $Var(x) = D(x) = \sigma_x^2 = \frac{1}{n}\sum (x_i - \bar{x})^2$.

Дисперсии характеризуют, как сильно рассеяны значения выборки относительно среднего значения:

D(X) малая



D(X) большая



3 Среднеквадратическое отклонение $\sigma_x = \sqrt{Var(x)}$ или стандартное отклонение.

Эта величина характеризует отклонение выборочных значений в среднем от $\bar{x}$.

2 ОДНОФАКТОРНАЯ ЛИНЕЙНАЯ РЕГРЕССИЯ

Изучение зависимостей экономических показателей начинают со случая двух переменных – X и Y: Y = f (X). Этот метод наиболее прост и может быть представлен графически.

Для начала нужно установить, существует ли функциональная зависимость между фактором X и откликом Y, и если существует, то определить формулу связи.

Для анализа данные представляют в виде таблицы:

X	**Y**
x_1	y_1
x_2	y_2
...	...
x_n	y_n

По таблице строится корреляционное поле (диаграмма рассеивания).

Корреляционным полем называют систему точек (x_i, y_i), $(i = 1, …, n)$, изображенную на координатной плоскости XOY (рис.1).

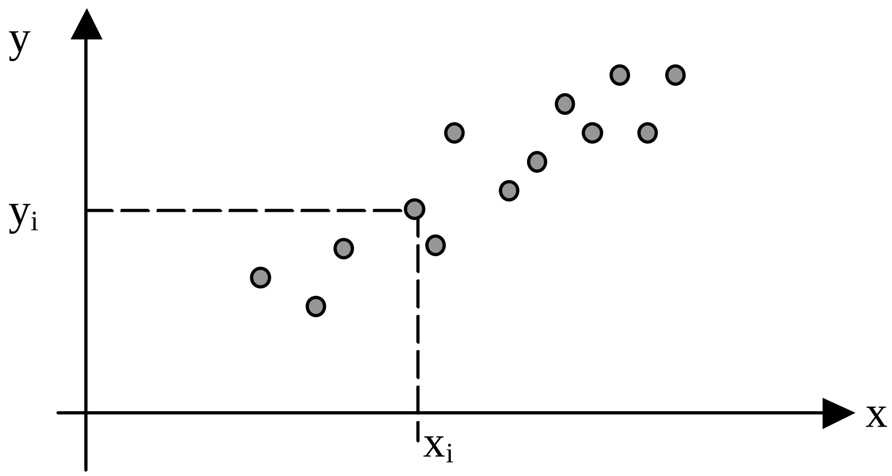


Рисунок 1

Точка с координатами $(\bar{x}, \bar{y})$ называется центром рассеяния.

По виду корреляционного поля выдвигается предположение, является ли зависимость между y и x линейной или нелинейной.

Значения σ_x (большие или малые) еще не дают характеристику того, есть ли связь между *x* и *y*. На рис.2, 3, 4 показаны ситуации, когда σ_x, σ_y малы, но в случае рис.2 зависимости вида y = f(x) нет, в случае рис.3 зависимость есть, и она линейная, в случае рис.4 есть явно нелинейная зависимость.

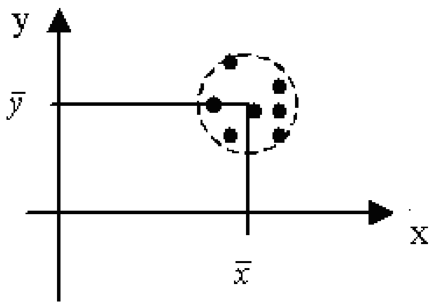


Рисунок 2

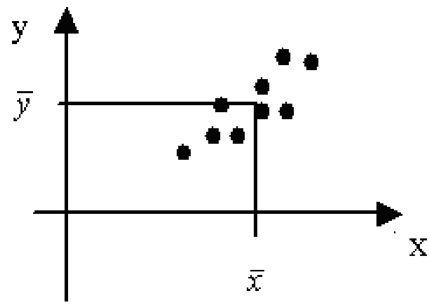


Рисунок 3

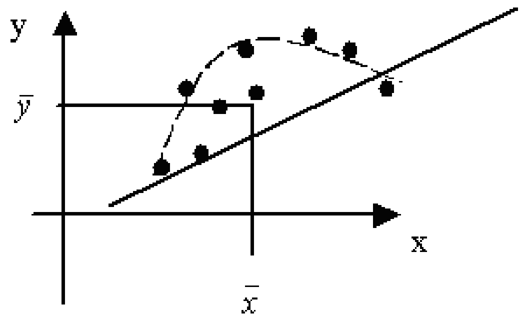


Рисунок 4

Поэтому вводится еще одна статистика – *коэффициент корреляции*. Вначале считается *ковариация x, y* – cov(x,y) (совместная вариация):

$$\mathrm{cov}(x,y)=\frac{1}{n}\sum(x_i-\bar{x})(y_i-\bar{y}).$$

Ковариация обладает тем свойством, что для случаев рис.2 и рис.4 равна 0, а для случая рис.3 не равна 0, и тем больше по модулю, чем ближе корреляционное поле к прямой.

Если корреляционное поле начинает размываться (рис.5), ковариация уменьшается.

Для удобства работы ковариацию делят на произведение σ_x, σ_y и называют *коэффициентом корреляции*. Обозначают $\boldsymbol{r_{xy}}$.

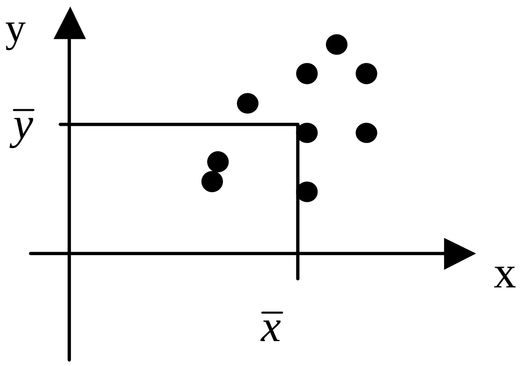


Рисунок 5

Коэффициент корреляции между переменными x и y вычисляется по формуле

$$r_{xy}=\frac{\mathrm{cov}(x,y)}{\sigma_x\sigma_y}.$$

Коэффициент корреляции является показателем плотности *линейной* взаимосвязи.

Свойства коэффициента корреляции:

1 $-1 \le r_{xy} \le 1$.

2 Если $r_{xy} > 0$, то зависимость между фактором x и y прямая, т.е. с ростом x показатель y также возрастает.

3 Если $r_{xy} < 0$, то зависимость между фактором x и y обратная.

4 Если $|r_{xy}| \approx 1$, связь между x и y – почти линейная (см. рис. 3).

5 Если $|r_{xy}| \approx 0$, либо связи нет (см. рис.2), либо связь резко нелинейная (см. рис.4).

Плотность линейной взаимосвязи оценивают по следующей таблице:

Значение $\lvert r_{xy} \rvert$	Плотность линейной связи
0,9…1,0	Тесная
0,6…0,9	Достаточная
0,3…0,6	Слабая
< 0,3	Нет связи

Следует иметь в виду, что величина линейного коэффициента корреляции r_{xy} оценивает тесноту *линейной* связи. Поэтому близость к нулю $|r_{xy}|$ не означает отсутствия связи между признаками. При $|r_{xy}| \to 0$ отсутствует *линейная* связь между х и у.

Обычно строят корреляционную таблицу (корреляционную матрицу) связи между переменными x и y.

Она имеет вид:

	x	y
x	1	r_{yx}
y	r_{xy}	1

Свойства корреляционной матрицы:

1 Корреляция фактора с самим собой равна 1: $r_{xx} = r_{yy} = 1$.

2 Матрица симметрична относительно главной диагонали: $r_{xy} = r_{yx}$.

3 ПОДБОР ПАРАМЕТРОВ ПРЯМОЙ РЕГРЕССИИ ПО МЕТОДУ НАИМЕНЬШИХ КВАДРАТОВ (МНК)

Парной (однофакторной) линейной регрессией называется линейная зависимость $y = b_0 + b_1 x$ между зависимым показателем Y и независимым фактором X.

Линейную связь между x и y описывают зависимостью

$$\hat{y} = b_0 + b_1 x. \tag{1}$$

В силу случайных влияний показатель y_i является случайным и может быть записан:

$$y_i = b_0 + b_1 x_i + e_i, \text{ i=1..n}, \tag{2}$$

где e_i – случайное отклонение (рис.6).

Отклонение (ошибка) исходных данных y_i от рассчитанных по модели значений $\hat{y}_i$=y(x_i) вычисляется по формуле

$$e_i = \hat{y}_i - y_i = b_0 + b_1 x_i - y_i.$$

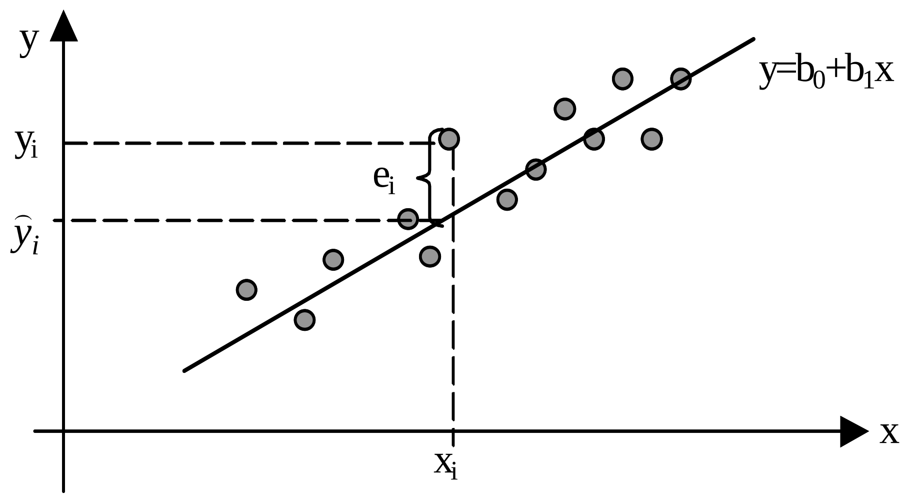


Рисунок 6

3.1 Метод наименьших квадратов

Суть МНК состоит в том, чтобы минимизировать отклонения e_i в совокупности путем правильного подбора коэффициентов b_0, b_1.

Т.к. отклонение может иметь случайный знак (+ или –), то рассматривают квадраты отклонений и минимизируют сумму квадратов отклонений:

$$S = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} \left(b_0 + b_1 x_i - y_i\right)^2 . \qquad (3)$$

Сумма S является функцией двух неизвестных параметров b_0, b_1. Необходимое условие минимума функции S – равенство нулю ее частных производных:

$$\begin{cases} \dfrac{\partial S}{\partial b_0} = 0; \\ \dfrac{\partial S}{\partial b_1} = 0. \end{cases} \qquad (4)$$

Получилась система двух линейных уравнений от двух неизвестных. Если ее определитель отличен от нуля, то такая система имеет единственное решение.

Выразив коэффициенты b_0 и b_1 и сделав арифметические преобразования, получим выражения для определения этих коэффициентов:

$$b_1 = r_{xy} \frac{\sigma_y}{\sigma_x}; \qquad (5)$$
$$b_0 = \bar{y} - b_1 \bar{x}.$$

Подставляя выражения (5) в уравнение $y = b_0 + b_1 x$, получим формулу линии регрессии

$$y - \bar{y} = r_{xy} \frac{\sigma_y}{\sigma_x} \left(x - \bar{x}\right) . \qquad (6)$$

Параметр b_1 называется *коэффициентом регрессии*. Его величина показывает среднее изменение результата при изменении фактора на одну единицу. Так, если в функции издержек y(x) = 3 000 + 2x (y, тыс. грн. – издержки, x – количество единиц продукции), то с увеличением объема продукции x на одну единицу издержки производства возрастают в среднем на 2 тыс. грн., то есть дополнительный прирост продукции на одну единицу потребует увеличения затрат в среднем на 2 тыс. грн.

Возможность четкой экономической интерпретации коэффициента регрессии b_1 объясняет распространенность линейного уравнения регрессии в эконометрических исследованиях.

Формально b_0 – это значение y при x = 0. Но удовлетворительно интерпретировать можно лишь знак при параметре b_0. Если b_0>0, то относительное изменение показателя y происходит медленнее, чем изменение фактора. Если b_0<0, то изменение результата опережает изменение фактора.

3.2 Свойства линейной регрессии

1 Сравним уравнение (6) с уравнением прямой, проходящей через точку (x_0, y_0)(рис.7):

$$y - y_0 = k(x - x_0). \tag{7}$$

Из сравнения уравнений (6) и (7) видно, что прямая регрессии всегда проходит через центр рассеивания корреляционного поля, т.е. через точку ($\bar{x}$, $\bar{y}$).

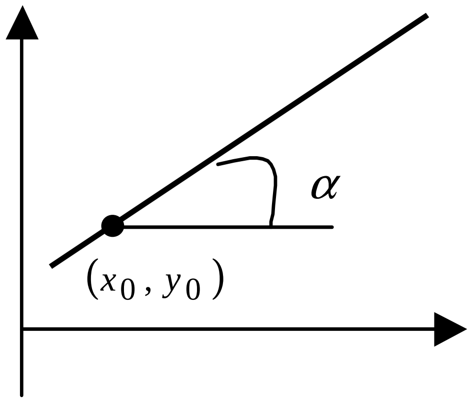


Рисунок 7

2 Из выражения $b_1 = r_{xy}\dfrac{\sigma_y}{\sigma_x}$ следует, что угловой коэффициент b_1 выражается через коэффициент корреляции r_{xy} и среднее квадратичное отклонение фактора и отклика, т.е. знак b_1 совпадает со знаком коэффициента корреляции (т.к. $\sigma_x, \sigma_y > 0$ всегда).

Если $r_{xy} > 0$, то $b_1 > 0$, угол α острый (рис.8), связь между *x* и *y* – прямая, т.е. с ростом *x* возрастает *y*.

Если $r_{xy} < 0$, то $b_1 < 0$, α тупой, связь между *x* и *y* обратная (рис.9).

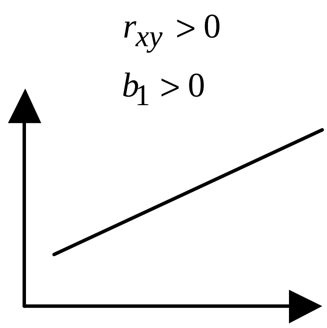


Рисунок 8

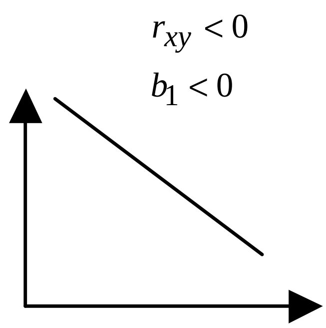


Рисунок 9

4 СТАТИСТИЧЕСКИЕ КРИТЕРИИ. СВЕДЕНИЯ ИЗ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Статистическая гипотеза – это предположение либо о виде распределения случайной величины, либо о значении числовой характеристики случайной величины.

Например:

1 Выдвигается гипотеза: случайные отклонения e_i выборочных значений y_i от расчетных значений $\hat{y}_i = b_0 + b_1 x_i$ распределены по нормальному закону. Это гипотеза о виде распределения.

2 Гипотеза: две выборочные дисперсии $D_1 = \sigma_1^2$ и $D_2 = \sigma_2^2$ равны между собой, т.е. $\dfrac{\sigma_1^2}{\sigma_2^2} \approx 1 \Rightarrow D_1 \approx D_2$. Это гипотеза о числовых характеристиках.

4.1 Нулевая и конкурирующая гипотезы

Гипотеза, выдвинутая первой, называется *нулевой* и обозначается H_0.

Например, H_0: $b_1 = 0$ означает, что $y(x) = b_0$, т.е. между *y* и *x* нет зависимости.

Гипотеза, противоположная гипотезе H_0, называется *конкурирующей*, или альтернативной, и обозначается H_1.

Например, H_1: $b_1 \neq 0$.

4.2 Ошибки 1 и 2 рода

При проверке выполнения гипотез возникают две ошибки.

Ошибка 1 рода состоит в том, что будет отвергнута правильная гипотеза. Вероятность отвергнуть правильную гипотезу обозначают α и называют *уровнем значимости гипотезы*. Обычно принимают $\alpha = 0{,}01 \div 0{,}05$.

Например: α = 0,05 означает, что в 5 случаях из 100 будет отвергнута правильная гипотеза.

Величина $\gamma = (1 - \alpha)$ называется *уровнем доверия*.

Ошибка 2 рода состоит в том, что будет принята неправильная гипотеза.

4.3 Статистические критерии проверки нулевой гипотезы

Статистический критерий – это специально сконструированная случайная величина.

Например, для проверки гипотезы о равенстве двух дисперсий используют критерий Фишера.

Критерий Фишера – это специально сконструированная случайная величина, равная отношению двух дисперсий:

$$F = \frac{D_1}{D_2} = \frac{\sigma_1^2}{\sigma_2^2}.$$

График плотности распределения вероятности этой случайной величины приведен на рис.10.

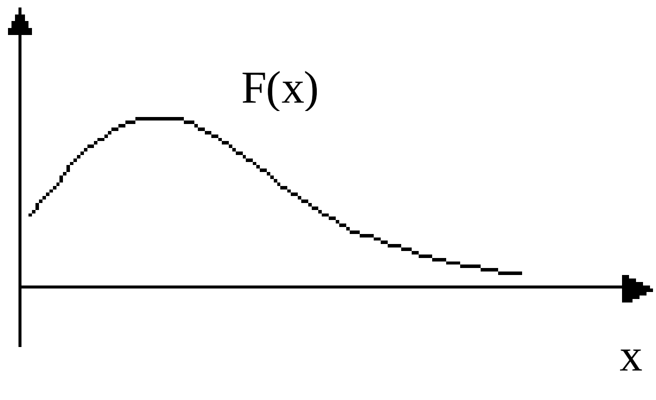


Рисунок 10

При этом важную роль играет понятие числа степеней свободы.

4.4 Число степеней свободы

Число степеней свободы – это разница между объемом выборки, по которой вычисляется выборочная численная характеристика, и числом связей, наложенных на выборочные значения.

Пример. Имеется выборка объема n: $x_1, x_2, \ldots x_n$.

По ней вычисляется среднее значение $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. Эта величина имеет n степеней свободы.

Рассмотрим выборочную дисперсию

$$D = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2. \qquad (8)$$

Выборочные значения x_i можно изменить (уменьшить или увеличить), причем так, что дисперсия не изменится, но на изменение значений x_i наложена одна связь – это выборочное среднее $\bar{x}$. Оно входит в формулу для вычисления дисперсии (8), и значит x_i должно меняться так, чтобы $\bar{x}$ не изменялось $\Rightarrow$ дисперсия D имеет (n – 1) степеней свободы. Обычно выборочную дисперсию вычисляют по формуле

$$D = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Т.к. в критерий Фишера входит 2 дисперсии – D_1 и D_2 – и каждая имеет свою степень свободы k_1 и k_2, то критерий Фишера зависит от двух степеней свободы – k_1 и k_2.

Получаем функцию $F(x, k_1, k_2)$.

При увеличении k_1 и k_2 распределение приближается к нормальному.

4.5 Наблюдаемые значения критерия

Наблюдаемое значение критерия вычисляется по имеющимся данным. Предположим, что проверяется нулевая гипотеза H_0:

$$H_0 : D_1 = D_2 .$$

По выборкам находятся дисперсии D_1 и D_2 и соответствующие им степени свободы k_1 и k_2. Наблюдаемое значение критерия Фишера: $F_{набл}(x, k_1, k_2)$.

4.6 Критические точки

Для того чтобы принять или отвергнуть гипотезу H_0, необходимо знать наблюдаемое и критическое значения статистического критерия.

Критическое значение – заранее рассчитанное значение критерия с определенным уровнем значимости. Это значение определяется как абсцисса на графике плотности распределения с заданным уровнем значимости α и степенями свободы k_1 и k_2.

Критическая точка имеет следующий смысл:

$$P\{F > F_{кр}\} = \alpha .$$

Приведем рисунок, поясняющий критическое значение критерия (рис.11).

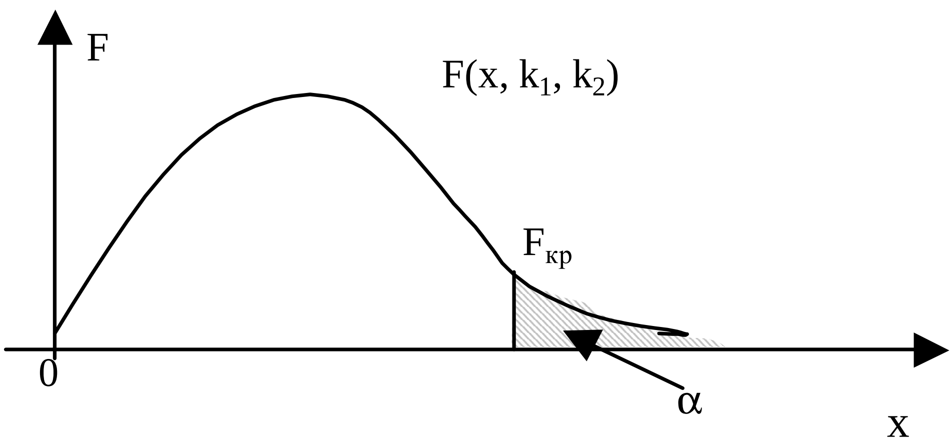


Рисунок 11

На графике α – это площадь заштрихованной области. Для критерия Фишера рассчитаны таблицы критических точек $F_{кр}$.

Каждому значению α соответствует своя таблица. Например, для $\alpha = 0,05$ таблица имеет вид:

$k_2 \backslash k_1$	1	…	10	…	∞
1	161		242		254
…					
10	4,97		2,47		2,54
…					
∞	3,84		1,83		1,00

Эти таблицы приводятся в учебниках по математической статистике.

4.7 Критерий принятия гипотезы

Принять или отвергнуть гипотезу H_0 можно, сравнив критическое и наблюдаемое значения критерия. Если наблюдаемое значение *меньше* критического, то гипотеза H_0 *принимается*. Если наблюдаемое значение *больше* критического, то гипотеза H_0 *отвергается*.

Для этого нужно знать уровень значимости α. В экономике, как правило, принимают $\alpha = 0,05$. Если уменьшать уровень значимости, то возрастает вероятность совершить ошибку второго рода.

5 ПРОВЕРКА ЛИНЕЙНОЙ РЕГРЕССИИ НА АДЕКВАТНОСТЬ

После того как была построена модель линейной регрессии $y = b_0 + b_1x$, необходимо проверить ее на адекватность, т.е. соответствует ли построенная модель имеющимся статистическим данным.

Вначале рассмотрим вариацию (разброс) зависимого показателя Y относительно своего среднего значения. Отклонение равно $y_i - \bar{y}$. Можно записать: $y_i - \hat{y}_i + \hat{y}_i - \bar{y}$, где $\hat{y}_i = b_0 + b_1 x_i$ – расчетные значения. Т.е. вариацию зависимого показателя Y вокруг своего среднего значения можно разделить на два слагаемых: $\hat{y}_i - \bar{y}$ – вариация расчетных значений вокруг среднего; $y_i - \hat{y}_i$ – вариация расчетных значений вокруг фактических.

Обозначим:

$\sigma_p^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$ – вариация, объясняемая регрессией, с числом степеней свободы $k_1 = 1$;

$\sigma_e^2 = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$ – остатки, необъясненный разброс, с числом степеней свободы $k_2 = n - 2$;

$\sigma_y^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2$ – общая вариация с числом степеней свободы $k_3 = n - 1$.

5.1 Коэффициент детерминации

Для анализа общего качества оцененной линейной регрессии обычно используют коэффициент детерминации:

$$R^2 = 1 - \frac{\sigma_e^2}{\sigma_y^2}; \qquad 0 \le R^2 \le 1.$$

В числителе стоит сумма квадратов отклонений линии регрессии от фактических значений, в знаменателе – от среднего значения. Значит, чем меньше отклонение расчетных значений от фактических, тем меньше дробь и тем ближе значение коэффициента детерминации к 1. Поэтому считается, что чем ближе значение коэффициента детерминации к 1, тем лучше модель описывает статистические данные.

Обычно в экономике для вариационных рядов величина коэффициента детерминации не превышает 0,6…0,7. Считается, что общее качество такой модели хорошее. Ответ на вопрос об адекватности модели R^2 не даёт.

5.2 Проверка модели на адекватность с помощью критерия Фишера

Если рассматривается линейная зависимость y от фактора x вида у = $b_0 + b_1 x$, то могут встретиться ситуации, показанные на рис.12.

Проверка линейной регрессии на адекватность означает выяснение наличия зависимости у от х. В случае рис.12, *г* такой зависимости нет. Фактически это означает, что угловой коэффициент $b_1 = 0$. В случае рис.12, а,б,*в*, $b_1 \neq 0$.

Постановка задачи

Выдвигаем гипотезу: H_0: ($b_1 = 0$). Уравнение регрессии будет иметь вид $y = b_0 = \bar{y}$. Т.е. функциональной зависимости между x и y нет.

Для проверки этой гипотезы сравниваются между собой две дисперсии:

$$D_1 = \frac{1}{k_1}\sigma_p^2 = \frac{1}{k_1}\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 \text{ и } D_2 = \frac{1}{k_2}\sigma_e^2 = \frac{1}{k_2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 .$$

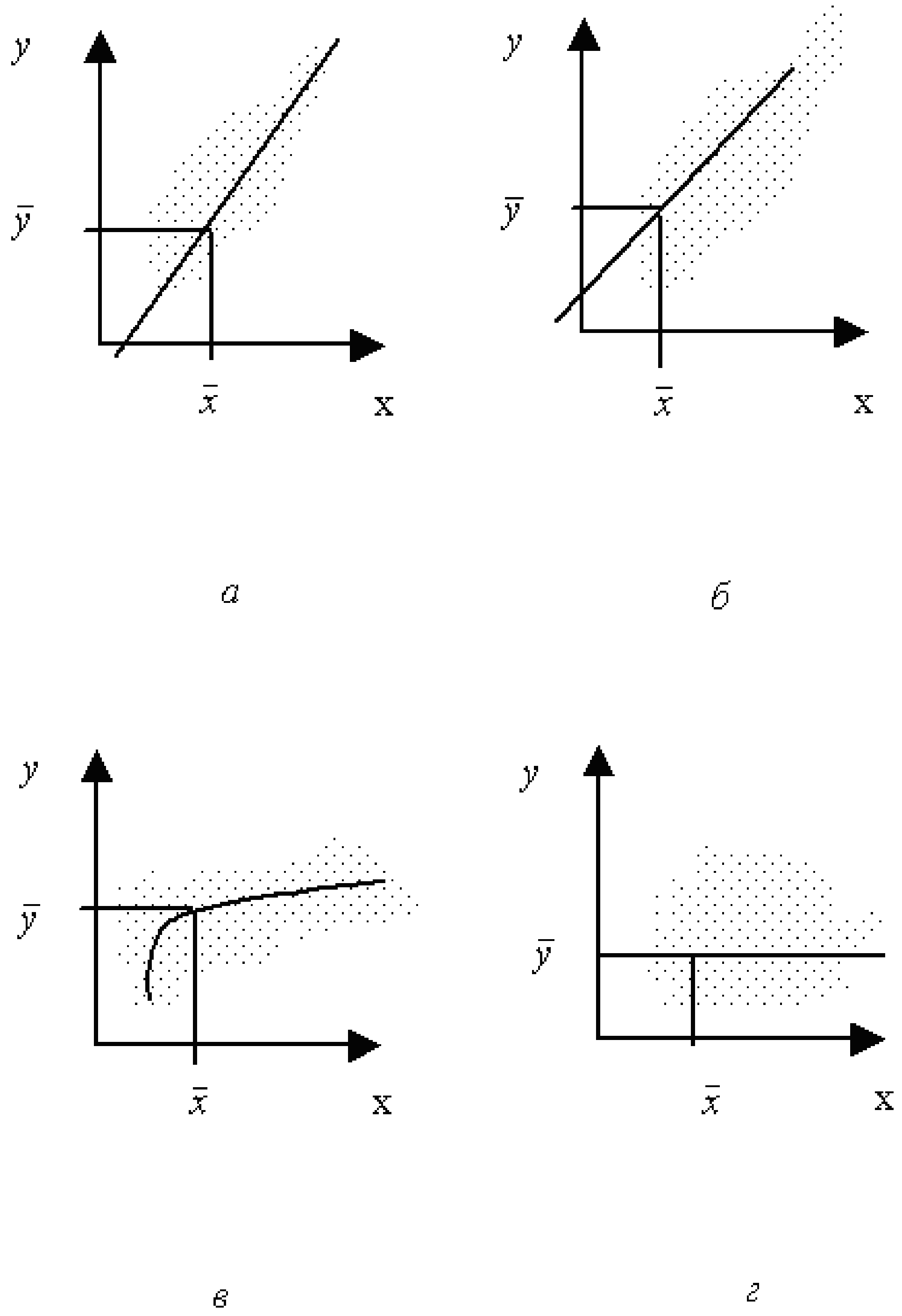


Рисунок 12

Т.е. вычисляем дисперсию остатков e_i и дисперсию расчетных значений $\hat{y}_i$, взятых с регрессионной прямой (рис.13).

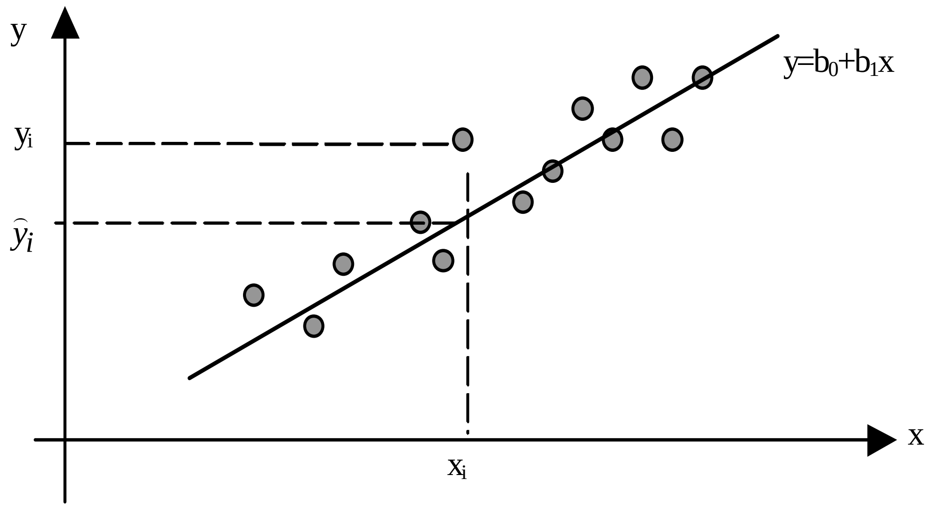


Рисунок 13

Вычисляем k_1, k_2 – количество степеней свободы для статистик D_1 и D_2. Число степеней свободы дисперсии D_2 равно $k_2 = n - 2$ (n – объем выборки).

Число степеней свободы статистики D_1 для однофакторной регрессии всегда равно 1, т.к. прямая регрессии всегда обязана проходить через центр регрессии, для нее можно только слегка изменить угол наклона прямой.

Следовательно, отношение введенных дисперсий представляет собой случайную величину, распределенную по закону Фишера со степенями свободы k_1, k_2:

$$\frac{D_1}{D_2} = F(x, k_1, k_2) = F(x, 1, n-2).$$

Проанализируем, что дает отношение дисперсий в случае рис.12,*г*.

Так как $\hat{y}_i$ берется с регрессионной прямой, которая в случае рис.12, *г* – горизонтальная, то $\hat{y}_i = \bar{y}$, т.е. все слагаемые в D_1 равны 0, и наблюдаемое значение критерия Фишера тоже равно 0:

$$F_{набл.} = \frac{D_1}{D_2} = 0\ .$$

В случае рис.12, *в* $F_{набл.} \approx 0$.

Переход от случая, когда можно признать $F_{набл} = 0$ (а, следовательно, $b_1 = 0$, и зависимость *y* от *x* отсутствует), к случаю, когда следует признать $F_{набл} \neq 0$ ($b_1 \neq 0$, т.е. есть зависимость *y* от *x*), производят, сравнивая $F_{набл}$ с теоретически вычисленным критическим значением для критерия Фишера $F_{кр}$ (см. п. 4.6).

Рассчитывают точку $F_{кр}$ при некотором уровне значимости гипотезы α. Если $F_{набл} < F_{кр}$, то делаем заключение, что $b_1 = 0$, значит, *y* от *x* не зависит, следовательно, модель неадекватна.

Если же $F_{набл} > F_{кр}$, то гипотеза H_0 отвергается, значит, $b_1 \neq 0$, *y* зависит от *x*, следовательно, модель $y = b_0 + b_1 x$ адекватна (с гарантией $(1 - \alpha)*100\%$).

Наблюдаемое значение критерия Фишера можно записать через коэффициент детерминации:

$$F_{набл.} = \frac{R^2}{1 - R^2}(n - 2).$$

Для многофакторной регрессии:

$$F_{набл.} = \frac{R^2}{1 - R^2} \cdot \frac{n - p - 1}{p},$$ где p – количество факторов.

5.3 Статистическая значимость коэффициентов модели

Найденные по МНК параметры модели b_0 и b_1 являются не точными, а случайными величинами. При изменении даже одной точки выборки получим другие коэффициенты b_0' и b_1' и т. д.

МНК гарантирует, что найденные с его помощью параметры:

- *не смещенные*. Это означает, что b_0 и b_1 случайны, так как найдены по выборке, но в среднем они такие, как если бы они были найдены по генеральной совокупности;

- *эффективные*. МНК обеспечивает быструю сходимость параметров модели к точным значениям, которые можно было бы рассчитать по генеральной совокупности;

- *состоятельные*. С увеличением объема выборки увеличивается точность рассчитанных параметров.

Так как b_0 и b_1 являются случайными величинами, то необходимо проверить их статистическую значимость. Это можно сделать с помощью специально сконструированной статистики, распределенной по закону Стьюдента.

Распределение Стьюдента T(x, k) возникает каждый раз, когда сравниваются два математических ожидания (два средних). Распределение Стьюдента симметрично относительно начала координат (рис.14).

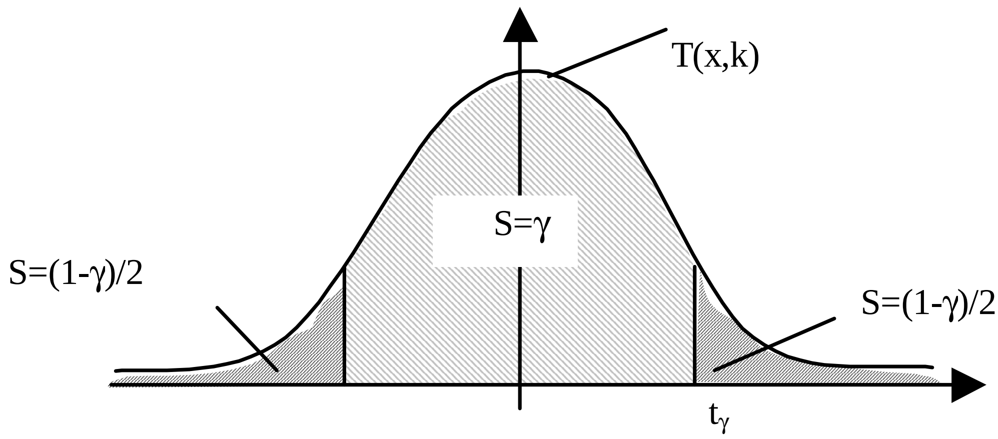


Рисунок 14

Число степеней свободы для критерия Стьюдента k = n – 2.

Рассчитывают для каждого коэффициента наблюдаемое значение $t_{набл} = \frac{b_i}{\sigma_{b_i}}$, для всех i = 0 ... p и сравнивают с критическим значением критерия $t_{кр}$. Если $t_{набл} > t_{кр}$, то соответствующий коэффициент статистически значим. Если $t_{набл} < t_{кр}$, то соответствующий коэффициент статистически не значим (является статистическим “нулем”). В модель включаются только статистически значимые параметры.

Для рассматриваемой однофакторной регрессии:

$$t_{b_1} = \frac{b_1}{\sigma_{b_1}}, \text{ где } \sigma_{b_1} = \sqrt{\frac{\sigma_e^2 /(n-2)}{\sum (x_i - \bar{x})^2}},$$

$$t_{b_0} = \frac{b_0}{\sigma_{b_0}}, \text{ где } \sigma_{b_0} = \sqrt{\frac{\sigma_e^2}{n-2} . \frac{\sum x_i^2}{n \cdot \sum (x_i - \bar{x})^2}}.$$

5.4 Проверка статистической значимости коэффициента корреляции

Коэффициент корреляции r_{xy}, рассчитанный по выборке, сам является случайной величиной. Значит, необходимо проверить статистическую значимость коэффициента корреляции. Эту проверку выполняют аналогично проверке статистической значимости параметров модели b_0 и b_1 с помощью критерия Стьюдента. Фактическое значение t – критерия Стьюдента определяется по формуле

$$t_r = \frac{r_{xy}}{\sqrt{1 - r_{xy}^2}} . \sqrt{n-2}$$

если выборка большая и $|r_{xy}|$ не близок к 0, либо по формуле

$$t_r = \frac{\sqrt{n-3}}{2} . \ln\left(\frac{1+r}{1-r}\right), \text{ если } |r_{xy}| \to 1.$$

Если $t_r > t_{kp}(\alpha, n-2)$, то коэффициент корреляции r_{xy} статистически значим.

6 ПРОГНОЗ НА ОСНОВАНИИ ЛИНЕЙНОЙ РЕГРЕССИИ

Если построенная модель $y = b_0 + b_1 x$ адекватна исходным статистическим данным, то по этой модели можно рассчитать прогноз в любой точке $x_{пр}$ из области прогнозов. *Областью прогнозов* называется отрезок прямой, заключенный между x_{min} и x_{max} (рис.15).

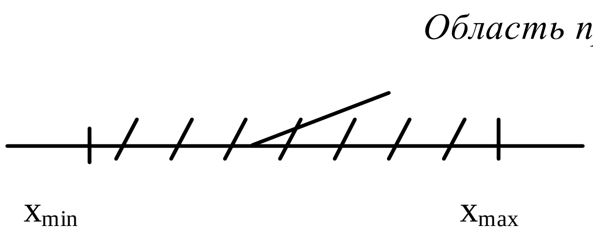


Рисунок 15

Такой прогноз $\hat{y}(x_{пр}) = b_0 + b_1 x_{пр}$ называется *точечным*.

6.1 Понятие о доверительном интервале

Если бы имелись сведения по всей генеральной совокупности (X), то можно было бы довольно точно найти статистические характеристики, например $\bar{x}^*$. Но, как правило, имеется выборка, в которой порядка десятка точек. По выборке рассчитывают выборочное среднее $\bar{x}$.

Истинное значение $\bar{x}^*$ может быть как больше, так и меньше выборочного $\bar{x}$, т.е. точное значение $\bar{x}^*$ попадает в некоторый интервал, центром которого является выборочное значение $\bar{x}$.

Если задаться вероятностью γ (например: 0,9; 0,99; 0,95) попадания $\bar{x}^*$ в интервал, то чем больше будет задана вероятность, тем шире будет получаться интервал. Если начать уменьшать γ, интервал будет сужаться.

Описанный интервал называется *доверительным интервалом*, а γ – коэффициентом доверия. Чаще всего на практике берут γ = 0,95. Это означает, что в 95% случаев точное значение параметра попадет в интервал.

Доверительный интервал – это интервал, в который с заданной вероятностью попадет истинное значение неизвестного параметра.

Коэффициент доверия – это вероятность, с которой доверительный интервал накроет неизвестный параметр.

6.2 Алгоритм нахождения полуширины доверительного интервала

По генеральной совокупности для конкретного x можно было бы довольно точно найти прогноз $y(x) = \beta_0 + \beta_1 x$. По выборке строится линейная регрессия $\hat{y} = b_0 + b_1 x$, и за $y(x)$ принимают $\hat{y}(x)$, снятое с прямой регрессии.

Доверительный интервал, в который попадает неизвестное $y(x)$ с некоторым коэффициентом доверия γ, в случае линейной регрессии оказывается симметричным относительно $\hat{y}(x)$ (рис.16). Поэтому достаточно найти полуширину доверительного интервала δ.

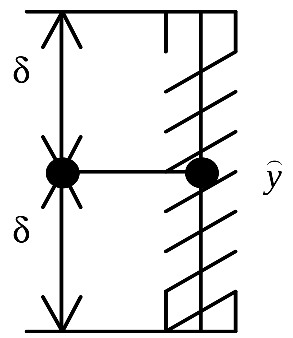


Рисунок 16

Полуширина доверительного интервала в точке прогноза $x_{пр}$ вычисляется по формуле

$$\delta = \sigma_e \cdot t_\gamma \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_{пр} - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}},$$

где σ_e – среднеквадратическое отклонение выборочных точек от линии регрессии $\sigma_e = \sqrt{\frac{1}{n-2} \sum e_i^2}$, здесь $e_i = y_i - \hat{y}_i$;

t_γ – критическая точка распределения Стьюдента с числом степеней свободы k = n – 2;

n – объем выборки;

$x_{пр}$ – точка из области прогнозов.

Прогнозируемый доверительный интервал для любого x из области прогнозов записывается: $(\hat{y} - \delta, \hat{y} + \delta)$.

Совокупность доверительных интервалов для всех x из области прогнозов образует доверительную область. Для линейной однофакторной регрессии она симметрична относительно линии регрессии (рис.17). Наиболее узкое место доверительной области в точке $(\bar{x}, \bar{y})$.

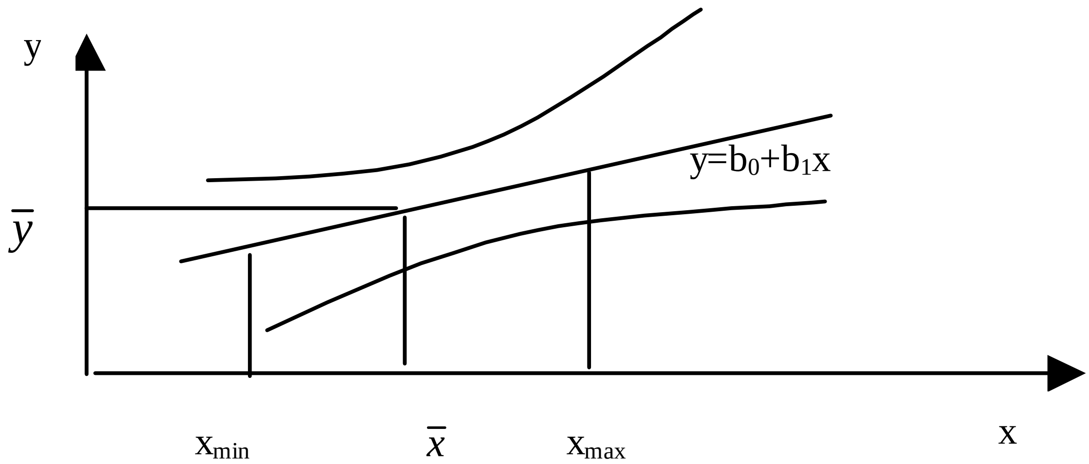


Рисунок 17

Прогноз для произвольного x дает интервал, в который с вероятностью γ попадает неизвестное $y(x)$. Т.е. прогноз при заданном x составит от $\hat{y}-\delta$ до $\hat{y}+\delta$ с надежностью $\gamma \cdot 100\%$. Это прогноз с учетом доверительного интервала.

7 НЕЛИНЕЙНАЯ ОДНОФАКТОРНАЯ МОДЕЛЬ

Многие экономические процессы не могут быть адекватно описаны линейной зависимостью вида $y = b_0 + b_1x$.

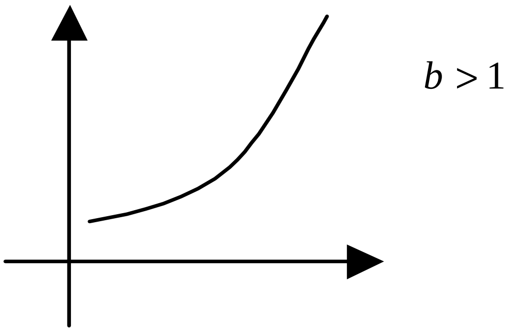
Примером таких экономических процессов могут служить: жизненный цикл товаров, процесс накопления капитала, маркетинговые усилия фирм и др.

Наиболее часто используются 5 нелинейных зависимостей, которые предпочтительней других зависимостей тем, что их удается линеаризовать (свести к линейным).

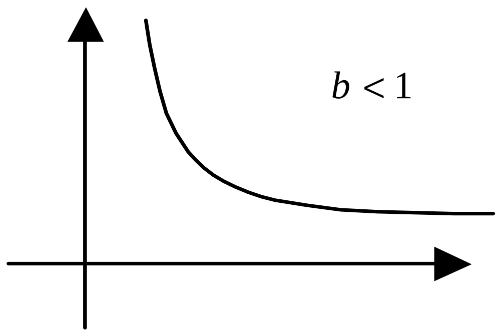
7.1 Виды нелинейных зависимостей

1 *Степенная зависимость:* $y = Ax^b$.

Кривые могут иметь вид, показанный на рис. 18, *а* и 18, *б*.



а



б

Рисунок 18

Прологарифмируем $y = Ax^b$:

$$\mathrm{Ln}y=\ln(Ax^b) \Rightarrow \mathrm{Ln}y=\mathrm{Ln}A+\mathrm{Ln}x^b \Rightarrow \mathrm{Ln}y=\mathrm{Ln}A+b\mathrm{Ln}x.$$

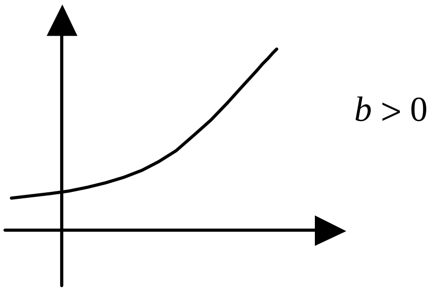
Обозначим $V = Lny$; $u = Lnx$; $b_0 = LnA$ и $b_1 = b$. Получим линейную модель от новых переменных: $V = b_0 + b_1u$.

Обратное преобразование: $V = Lny \Rightarrow y = e^v \Rightarrow e^{b0+b1u} \Rightarrow e^{b0}e^{b1u} \Rightarrow e^{b0}(e^{lnx})^{b1} \Rightarrow e^{b0}x^{b1}$.

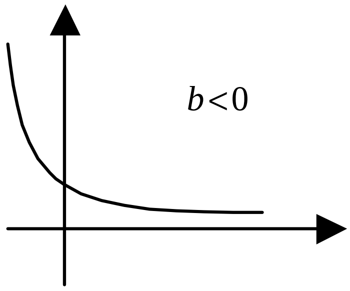
Значит, $A = e^{b0}$, $b = b_1$, $y=e^v$.

2 *Экспоненциальная зависимость*: $y = A{\cdot}e^{bx}$.

Экспоненциальные кривые могут иметь вид, показанный на рисунках 19, *а* и 19, *б*.



а



б

Рисунок 19

Чтобы ее линеаризовать, прологарифмируем уравнение $y = Ae^{bx}$.

$Lny = ln(e^{bx}) \Rightarrow Lny = LnA + Lne^{bx} \Rightarrow Lny = LnA + bxLne \Rightarrow Lny = LnA + bx$.

Обозначим $V = Lny$; $u = x$; $b_0 = LnA$ и $b_1 = b$. Получим $V=b_0+ b_1u$.

Обратное преобразование:

$$V = Lny \Rightarrow y = e^v \Rightarrow e^{b0+b1u} \Rightarrow e^{b0}e^{b1u} \Rightarrow e^{b0}e^{xb1}.$$

Значит, $A = e^{b0}$, $b = b_1$, $y = e^v$.

3 *Логарифмическая зависимость*: $y = A + Blnx$.

Кривая может иметь вид, показанный на рис.20.

Сделаем замену: $V = y$; $u = Lnx$; $b_0 = A$ и $b_1 = B$. Получили $V = b_0 + b_1u$.

Примером использования такой функции может служить взаимосвязь доли расходов на товары длительного пользования и общих сумм расходов или доходов.

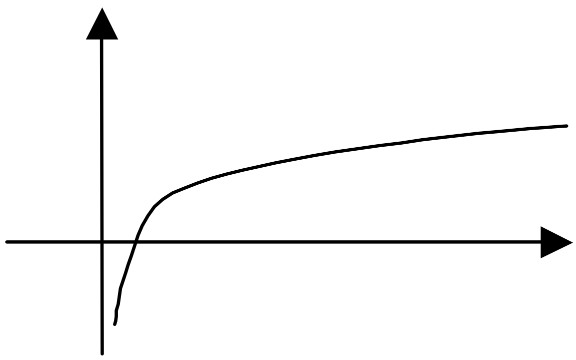


Рисунок 20

Математическое описание подобного рода взаимосвязей получило название кривых Энгеля.

4 *Обратная зависимость*: $y = A + B\frac{1}{x}$.

Кривая может иметь вид, показанный на рис. 21.

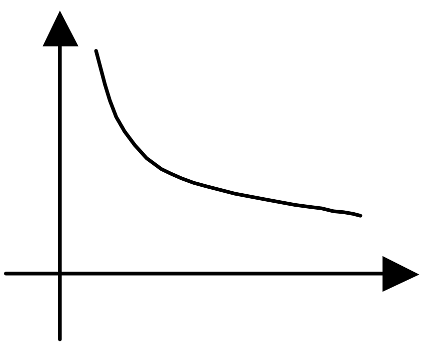


Рисунок 21

Сделаем замену: V = y; u $=\frac{1}{x}$; b_0 = A и b_1 = B. Получили $V=b_0 + b_1u$.

Эта модель достаточно известна в эконометрике. Классическим ее примером является кривая Филлипса, характеризующая нелинейное соотношение между нормой безработицы x и процентом прироста заработной платы y: $y(x) = 0{,}00679 + \dfrac{0{,}1842}{x}$. Величина параметра $b_0 = 0{,}00679$ означает, что с ростом уровня безработицы темп прироста заработной платы в пределе стремится к нулю.

5 *Логистическая кривая*: $y = \dfrac{1}{A + Be^{-x}}$.

Кривая может иметь вид, показанный на рис.22.

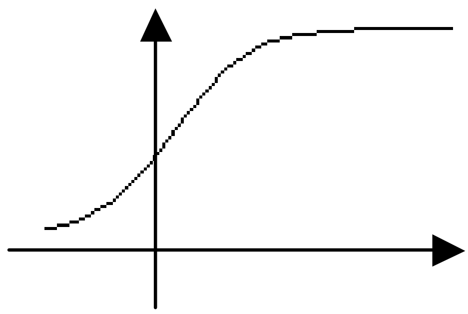


Рисунок 22

Сделаем замену: $V = \dfrac{1}{y}$; $u = e^{-x}$; $b_0 = A$ и $b_1 = B$. Получили $V = b_0 + b_1 u$.

Нелинейные зависимости можно разделить на:

- регрессии, нелинейные по оцениваемым параметрам: $y = Ax^b$; $y = A \cdot e^{bx}$; $y = \dfrac{1}{A + Be^{-x}}$;

- регрессии, нелинейные по оцениваемым факторам: $y = A + B\ln x$; $y = A + B\dfrac{1}{x}$.

7.2 Алгоритм построения нелинейных эконометрических моделей

1 Имеется выборка (таблица 1), относительно которой есть экономические соображения о виде зависимости между *x* и *y*: y=f(x).

Таблица 1

X	x_1	x_2	…	x_n
Y	y_1	y_2	…	y_n

2 Если зависимость известна, то, используя соответствующую замену, пересчитываем значения выборки и получаем новую выборку (таблица 2), по которой можно построить линейную модель:

$$v = b_0 + b_1 u. \tag{9}$$

Таблица 2

U	u_1	u_2	…	u_n
V	v_1	v_2	…	v_n

3 Найденную модель (9) проверяем на адекватность. Если она адекватна, то и исходная нелинейная модель адекватна. Если линеаризованная модель неадекватна, то исходная модель выбрана неверно, и нужно подобрать другую нелинейную модель (например, вместо степенной модели попробовать экспоненциальную).

4 Если линеаризованная модель (9) адекватна, то в тех точках, в которых нужно посчитать прогноз, рассчитываем полуширину доверительного интервала δ_u для линеаризованной модели.

5 Пересчитываем прогноз и доверительный интервал для точки прогноза из линеаризованного вида в исходный нелинейный. Для этого находим границы доверительного интервала для линеаризованной модели (v_{min}, v_{max}) и для них, а также для значений $\hat{v}$, с помощью обратного преобразования находим y_{min}, y_{max} и $\hat{y}$.

Для нелинейной регрессии доверительный интервал может быть несимметричен относительно линии регрессии.

8 ГЕТЕРОСКЕДАСТИЧНОСТЬ

Рассмотрим линию регрессии, построенную по корреляционным полям различного вида (рис.23).

На рис.23, *а* дисперсия остатков e_i не зависит от x, т.е. $D(e_i) = const$. Такое распределение называется *гомоскедастичным*. В случаях, представленных на рис. 23, *б*...*г*, дисперсия остатков непостоянна. Такое явление называется *гетероскедастичностью*.

Основным следствием гетероскедастичности является очень широкий доверительный интервал и, как следствие, широкая доверительная область, что приводит к потере экономического смысла прогнозирования на основании такой модели (рис.23).

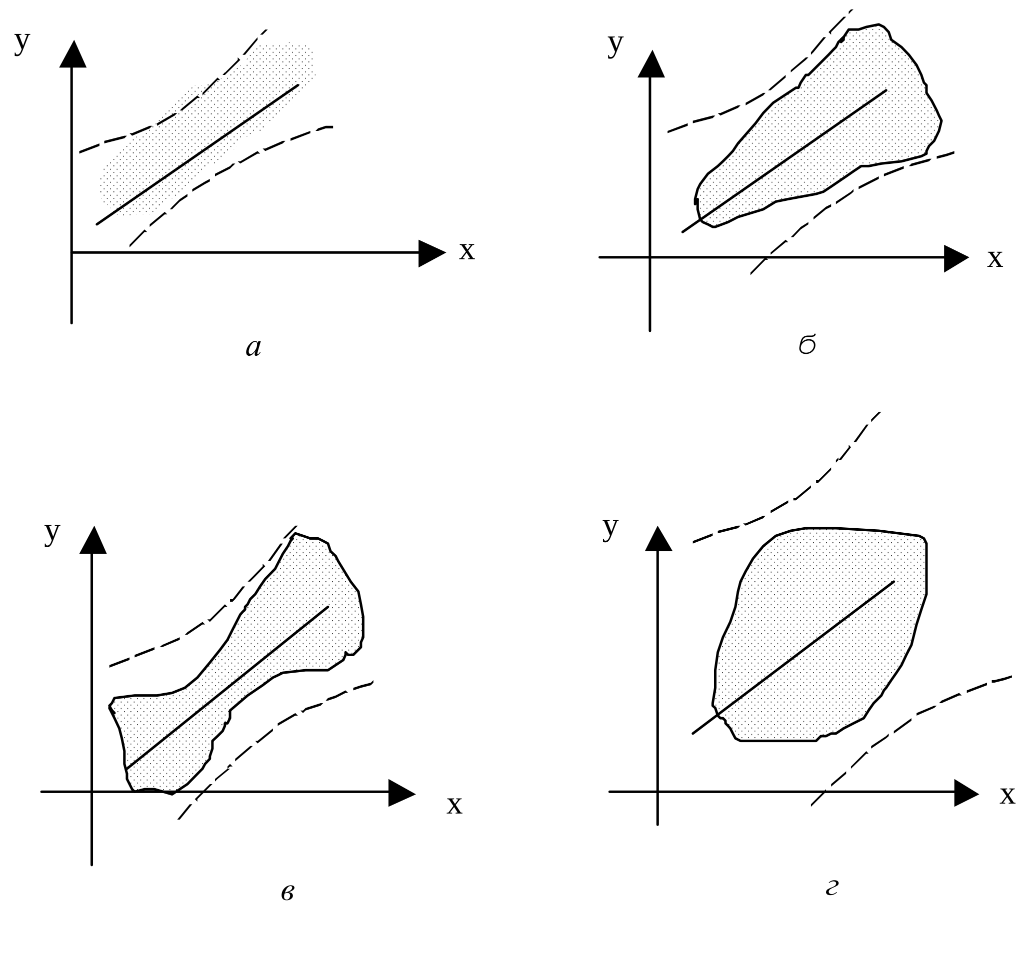


Рисунок 23

9 ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

При наличии гетероскедастичности в остатках рекомендуется традиционный метод наименьших квадратов (МНК) заменять обобщенным методом наименьших квадратов (ОМНК).

Будем предполагать, что:

- среднее значение остаточных величин равно нулю;
- дисперсия остаточных величин не остается неизменной для различных значений фактора, а пропорциональна некоторой величине K_i, т.е.

$$\sigma^2_{\varepsilon_i} = \sigma^2 \cdot K_i,$$

где $\sigma^2_{\varepsilon_i}$ – дисперсия ошибки на конкретном (i-м) значении фактора;

σ^2 – постоянная дисперсия ошибки при соблюдении предпосылки о гомоскедастичности остатков;

K_i – коэффициент пропорциональности, меняющийся с изменением величины фактора, что и обуславливает неоднородность дисперсии.

При этом полагается, что величина σ^2 неизвестна, а в отношении величины K_i выдвигаются определенные гипотезы, характеризующие структуру гетероскедастичности.

В общем виде уравнение регрессии примет вид

$$\frac{y_i}{\sqrt{K_i}} = \frac{\alpha}{\sqrt{K_i}} + \beta \frac{x_i}{\sqrt{K_i}} + \varepsilon_i.$$

Исходные данные для этого уравнения будут иметь вид

$$y=\begin{vmatrix}\dfrac{y_1}{\sqrt{K_1}}\\ \dfrac{y_2}{\sqrt{K_2}}\\ \dots\\ \dfrac{y_n}{\sqrt{K_n}}\end{vmatrix},\quad x=\begin{vmatrix}\dfrac{x_1}{\sqrt{K_1}}\\ \dfrac{x_2}{\sqrt{K_2}}\\ \dots\\ \dfrac{x_n}{\sqrt{K_n}}\end{vmatrix}.$$

По отношению к обычной регрессии уравнение с новыми, преобразованными переменными представляет собой взвешенную регрессию, в которой переменные x и y взяты с весами $1/\sqrt{K}$.

Оценка параметров нового уравнения с преобразованными переменными приводит к методу наименьших квадратов, для которого необходимо минимизировать сумму квадратов отклонений вида

$$S=\sum_{i=1}^{n}\frac{1}{K_i}\cdot\left(y_i-a-b\cdot x_i\right)^2\to\min.$$

10 МНОГОФАКТОРНАЯ РЕГРЕССИЯ

10.1 Понятие многофакторной модели и этапы ее построения

Реальные экономические процессы, как правило, зависят не от одного, а от нескольких факторов:

$$y = f(x_1, x_2, ,\dots, x_p). \qquad (10)$$

Процесс нахождения зависимости (10) и получения полезных выводов из этой зависимости таков:

1 Знакомство с экономической теорией, выдвижение гипотез о виде взаимосвязи. Постановка задачи.

2 Спецификация модели – теоретические представления и принятие гипотезы в виде математических уравнений, которые устанавливают связь между независимыми переменными.

3 Формирование массивов входной информации согласно задачам исследования.

4 Исследование исходных статистических данных на наличие зависимости между факторами (автокорреляция), влияния отклика y на факторы x_i (авторегрессия), эффектов, связанных с запаздыванием реакции рынка (лаговые эффекты).

5 Оценка параметров модели на основании статистических данных.

6 Проверка полученной модели на адекватность.

7 Прогноз на основании уравнения регрессии и расчет доверительного интервала для прогноза.

10.2 Спецификация модели

Эконометрическая модель основана на объединении двух аспектов – теоретического, качественного анализа и опытной информации. Теоретическая информация находит свое отображение в спецификации модели. Спецификация модели – это аналитическая форма эконометрической модели.

Из опыта эконометрических исследований можно привести примеры функций, которые могут описывать взаимосвязь между показателем У и факторами X_i:

1 Линейная функция $Y = b_0 + b_1 X_1 + \cdots + b_p X_p$.

2 Степенная функция

$$Y = a_0 X_1^{a_1} \cdot \cdots \cdot X_p^{a_p} \rightarrow \ln y = \ln a_0 + a_1 \ln X_1 + \cdots + a_p X_p \text{ или}$$

$$V = b_0 + b_1 U_1 + \cdots + b_p U_p.$$

3 Гиперболическая функция

$$Y = a_0 + \frac{a_1}{X_1} + \cdots + \frac{a_p}{X_p} \rightarrow y = a_0 + a_1 U_1 + \cdots + a_p U_p, U_i = \frac{1}{X_i}.$$

4 Квадратическая функция

$$Y = b_0 + b_1 X_1^2 + \cdots + b_p X_p^2 \rightarrow Y = b_0 + b_1 U_1 + \cdots + b_p U_p, \quad U_i = X_i^2.$$

Во всех моделях Y – зависимая переменная, показатель X_i – независимые переменные, факторы, b_i – параметры модели.

Линейные функции являются достаточно распространенными, и приведенные выше нелинейные функции также сводятся к линейным, поэтому далее будем рассматривать построение линейной модели.

Имея в виду, что выбор аналитической формы эконометрической модели не может рассматриваться без перечисления конкретных факторов, спецификация модели предусматривает отбор факторов для эконометрического анализа.

На этом этапе требуется хорошее понимание экономического процесса. Предварительно анализируется, от чего зависит Y. Например, объем продаж может зависеть от рекламы, имиджа, среднего заработка людей в регионе и т.д. Собираются статистические данные. Выборочные сведения оформляются в виде таблицы 3. После выявления всех факторов X_i нужно исключить все факторы, для которых мало наблюдений.

Таблица 3

Номер набл.	Y	X_1	X_2	…	X_p
1	y_1	x_{11}	x_{12}		x_{1p}
2	y_2	x_{21}	x_{22}		x_{2p}
			…		
n	y_n	x_{n1}	x_{n2}		x_{np}

Ошибки спецификации могут быть следующими:

1) игнорирование важного фактора при построении эконометрической модели;

2) ввод в модель фактора, который существенно не влияет на показатель;

3) выбор несоответствующей математической формы зависимости.

Ввиду четкой интерпретации параметров наиболее широко используются линейная и степенная функции. В линейной многофакторной модели $Y = b_0 + b_1 X_1 + \cdots + b_p X_p$ параметры $\boldsymbol{b_i}$ $\boldsymbol{(i = 1 \ldots p)}$ называют коэффициентами “чистой” регрессии. Они характеризуют среднее изменение результата с изменением соответствующего фактора на единицу при неизменном значении других факторов, закрепленных на среднем уровне.

*Пример***.** Пусть зависимость расходов на продукты питания по совокупности семей характеризуется следующим уравнением:

$$\hat{y}(x_1, x_2) = 0.5 + 0.35 x_1 + 173 x_2,$$

где у – расходы семьи за месяц на продукты питания, грн.;

x_1 – месячный доход на 1 члена семьи, грн.;

x_1 x_2 – размер семьи (человек).

Анализ этого уравнения позволяет сделать выводы:

- с ростом дохода на одного члена семьи на 100 грн. расходы на питание возрастут в среднем на 35 грн. при этом же среднем составе семьи.

- увеличение размера семьи при тех же её доходах предполагает дополнительный рост расходов на питание на 173 грн.

Степенные функции $\hat{Y} = a_0 X_1^{a_1} \cdot \cdots \cdot X_p^{a_p}$ получили наибольшее распространение в производственных функциях, в исследованиях спроса и потребления.

Пример. Пусть при исследовании спроса на мясо получена модель: $\hat{y} = 0.82\, x_1^{-2.63} x_2^{1.11}$, где $\hat{y}$ – количество спрашиваемого мяса, x_1 – цена, x_2 – доход. Отсюда следует, что рост цен на 1% при том же доходе вызывает снижение спроса в среднем на 2,63%. Увеличение дохода на 1% при неизменной цене обуславливает рост спроса на 1,11%.

Если эти функции не устраивают исследователя, то современные компьютерные программы позволяют перебирать другие линеаризуемые функции и выбирать наиболее соответствующую из них. При этом следует помнить, что, чем сложнее функция, тем менее интерпретируемы её параметры.

10.3 Анализ факторов на мультиколлинеарность

Одним из условий применения МНК является то, что факторы должны быть независимыми друг от друга. Однако очевидно, что в экономике очень трудно выделить такой массив факторов, которые были бы совсем независимы друг от друга. Поэтому каждый раз необходимо выяснять, не влияет ли зависимость факторов на оценку параметров модели.

Факторы X_1, X_2, …, X_p интерпретируются как векторы размерности n:

$$\vec{X}_i = \begin{pmatrix} x_{1i} \\ x_{2i} \\ \cdots \\ x_{ni} \end{pmatrix}, \ i = 1...p.$$

Два вектора коллинеарны, если их координаты пропорциональны:

$$\vec{X}_1 \mathrm{II} \vec{X}_2 \Leftrightarrow \frac{x_{11}}{x_{12}} = \frac{x_{21}}{x_{22}} = \cdots \frac{x_{n1}}{x_{n2}}. \qquad (11)$$

Если выполняется условие (11), то коэффициент корреляции между факторами X_1 и X_2 равен 1:

$$\vec{X}_1 \text{II} \vec{X}_2 \Leftrightarrow r_{X_1 X_2} = 1. \qquad (12)$$

Факторы X_1 и X_2 строго коллинеарны.

На практике строгая коллинеарность встречается редко. Более распространен случай, когда коллинеарность не строгая:

$$r_{X_i X_J} \approx 1.$$

Поскольку одним из условий построения множественной регрессии является независимость действия факторов, т.е. $r_{X_i X_J} = 0$, коллинеарность факторов нарушает это условие. Если факторы явно коллинеарны, то они дублируют друг друга и один из них рекомендуется исключить из регрессии. Предпочтение при этом отдается не фактору, более тесно связанному с результатом, а тому фактору, который при достаточно тесной связи с результатом имеет наименьшую тесноту связи с другими факторами. В этом требовании проявляется специфика множественной регрессии как метода исследования комплексного воздействия факторов в условиях их независимости друг от друга.

Пусть, например, при изучении зависимости $Y = f(X_1, X_2, X_3)$ матрица парных коэффициентов корреляции оказалась следующей:

	У	X_1	X_2	X_3
У	1			
X_1	0,8	1		
X_2	0,7	0,8	1	
X_3	0,6	0,5	0,2	1

Очевидно, что факторы X_1 и X_2 дублируют друг друга. В дальнейший анализ целесообразно включить фактор X_2, а не фактор X_1, так как корреляция X_2 с третьим фактором X_3 слабее: $r_{X_2 X_3} < r_{X_1 X_3}$. Поэтому в данном случае в уравнение множественной регрессии включаются факторы X_2, X_3: $Y = f(X_2, X_3)$.

По величине парных коэффициентов корреляции обнаруживается лишь явная коллинеарность факторов. Наибольшие трудности в использовании аппарата множественной регрессии возникают при наличии мультиколлинеарности факторов, когда имеет место совокупное воздействие факторов друг на друга.

Мультиколлинеарностью называется линейная зависимость между факторами X_1, X_2, …, X_p. Вектор является линейно зависимым, если его можно представить в виде линейной комбинации остальных векторов:

$$\vec{X}_1 = \lambda_2 \vec{X}_2 + \lambda_3 \vec{X}_3 + \cdots + \lambda_p \vec{X}_p, \qquad (13)$$

причем хотя бы один множитель $\lambda_i \neq 0$.

Коллинеарность является частным случаем мультиколлинеарности.

Одной из возможностей учета межфакторной корреляции является переход к совмещенным уравнениям регрессии, которые отражают не только влияние факторов, но и их взаимодействие, например, не

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3,$$

а

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{123} x_1 x_2 x_3.$$

Это уравнение учитывает взаимодействие всех факторов. Часть слагаемых может быть отброшена, если взаимодействие между некоторыми факторами окажется несущественным.

10.4 Последствия мультиколлинеарности

Если в условии задачи присутствует совершенная мультиколлинеарность, т.е. формула (13) выполняется точно, то при определении коэффициентов уравнения линейной регрессии

$$\widehat{Y} = b_0 + b_1 X_1 + \cdots + b_p X_p$$

по МНК возникает неопределенность типа $\frac{0}{0}$, т.е. коэффициенты определить невозможно – задача не имеет решения.

Если мультиколлинеарность не совершенная, т.е. $\vec{X}_1 \approx \lambda_2 \vec{X}_2 + \lambda_3 \vec{X}_3 + \cdots + \lambda_p \vec{X}_p$, то неопределенность $\frac{0}{0}$ не возникает. Коэффициенты $b_0, b_1, \ldots b_p$ по МНК возможно определить, но среднеквадратические погрешности очень велики. Как следствие: доверительные интервалы очень широкие, и прогнозы теряют практическую ценность.

Наиболее простой способ устранения мультиколлинеарности в эконометрической модели – исключение одного из факторов. Однако если вслепую устранять мультиколлинеарность по предложенному алгоритму, не задумываясь над экономическим смыслом задачи, можно выхолостить модель настолько, что она потеряет ценность.

***Пример*.** Пусть Y – потребление, X_1 – богатство, X_2 – доход. Эти сведения собираются случайным образом с большого количества семей. Понятно, что потребление зависит как от богатства, так и от дохода, т.е. $Y = b_0 + b_1 X_1 + b_2 X_2 + e$, где e – случайное малое слагаемое. Понятно, что доход пропорционален богатству: $\vec{X}_1$ почти коллинеарен $\vec{X}_2$. Если же исключить один из факторов, то модель потеряет экономический смысл.

10.5 Способы устранения мультиколлинеарности

Наиболее простой способ устранения мультиколлинеарности – это преобразование данных:

1 Взять $u_i = x_i - \bar{x}$ – отклонение от среднего.

2 Вместо абсолютных значений взять относительные.

3 Нормализовать фактор $u_{ik} = \frac{x_{ik} - \bar{x}_k}{\sigma_{x_k}}, i = 1 \ldots n.$

4 Заменить одну переменную на другую.

Более сложные способы устранения мультиколлинеарности см. [5].

Если ни один из способов не дает возможности устранить мультиколлинеарность, то МНК использовать не рекомендуется.

10.6 Нахождение регрессионной модели

Линейная многофакторная модель имеет вид:

$$Y = b_0 + b_1 X_1 + \cdots + b_p X_p . \tag{14}$$

Параметры линейной многофакторной модели определяются по МНК аналогично тому, как определяют параметры линейной однофакторной модели. Только получаем систему p уравнений. Эту систему решают специальными методами матричной алгебры.

На этом этапе анализа появляются следующие *сведения*:

1 Коэффициент множественной корреляции

$$R = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}} . \tag{15}$$

С помощью этого коэффициента оценивается практическая значимость уравнения многофакторной регрессии. Он оценивает тесноту совместного влияния факторов на результирующий показатель.

2 Коэффициент детерминации R^2.

3 Уточненный коэффициент детерминации с учетом степеней свободы

$$\boldsymbol{R^2_{adj} = 1 - \left(1 - R^2\right)\frac{(n-1)}{(n-p)}} .$$

4 Наблюдаемое значение и степени свободы критерия Фишера.

Критерий Фишера в случае многофакторной регрессии имеет тот же смысл, что и в случае однофакторной регрессии, т.е. отношение двух дисперсий, однако число степеней свободы меняется: $k_1 = p, k_2 = n - p - 1$. Здесь p – количество факторов в модели, n – объем выборки. Для многофакторной регрессии H_0: все коэффициенты $b_i = 0$, кроме b_0.

5 Статистическая значимость коэффициентов b_i.

Получается модель:

$$Y = b_0 + b_1 X_1 + \cdots + b_p X_p .$$

В полученной модели $Y = b_0 + b_1 X_1 + \cdots + b_p X_p$ коэффициенты b_i – случайные величины. Их математические ожидания при выполнении некоторых условий равны, соответственно, точным значениям β_i. При этом оценки тем надежнее, чем меньше их разброс вокруг точных значений, т.е. дисперсия. Можно доказать, что

$$D(b_i) = \sigma_{b_i}^2 = \frac{S^2}{\sum\limits_i (x_i - \bar{x})^2}, i = 1, \ldots, p, \; D(b_0) = \sigma_{b_0}^2 = \frac{S^2 \sum\limits_i x_i^2}{n \sum\limits_i (x_i - \bar{x})^2},$$

где $S^2 = \dfrac{\sum\limits_i e_i^2}{n - p - 1}$ – дисперсия остатков.

Формально значимость (отличие от нуля) коэффициента b_i может быть проверена с помощью критерия Стьюдента. Вычисляют $t_{набл} = \dfrac{b_i}{\sigma_b}$. Если $t_{набл} < t_{кр}$, то коэффициент b_i является статистическим нулем.

Во многих пакетах программ статистическая значимость коэффициентов считается автоматически.

10.7 Прогноз на основании линейной модели

Одна из важнейших целей моделирования заключается в прогнозировании поведения исследуемого объекта. Обычно термин “прогнозирование” используется в тех ситуациях, когда требуется предсказать состояние системы в будущем. Для регрессионных моделей он имеет более широкое значение. Данные, как в рассматриваемом случае, могут не иметь временной структуры, но может ставиться задача оценить значение зависимой переменной для некоторого набора независимых переменных, которых нет в исходных наблюдениях.

Так же, как и в случае однофакторной регрессии, различают точечное и интервальное прогнозирование. В первом случае оценка – это конкретное число, во втором – это интервал, в котором с заданным уровнем доверия содержится истинное значение искомой переменной.

Прогноз делается по уравнению регрессии

$$\widehat{Y} = b_0 + b_1 X_1 + \cdots + b_p X_p . \qquad (16)$$

Точка прогноза $\vec{X}$ с координатами $(x_1;\ x_2;\ \ldots\ ;x_p)$ из p-мерного пространства выбирается из области прогноза, которая определяется системой неравенств

$$\begin{cases} x_{1_{\min}} \le x_1 \le x_{1_{\max}}; \\ x_{2_{\min}} \le x_2 \le x_{2_{\max}}; \\ \ldots \\ x_{p_{\min}} \le x_p \le x_{p_{\max}}. \end{cases}$$

Если, например, модель двухфакторная $\widehat{Y} = b_0 + b_1 X_1 + b_2 X_2$, то область прогноза определяется прямоугольником, показанным на рис. 24.

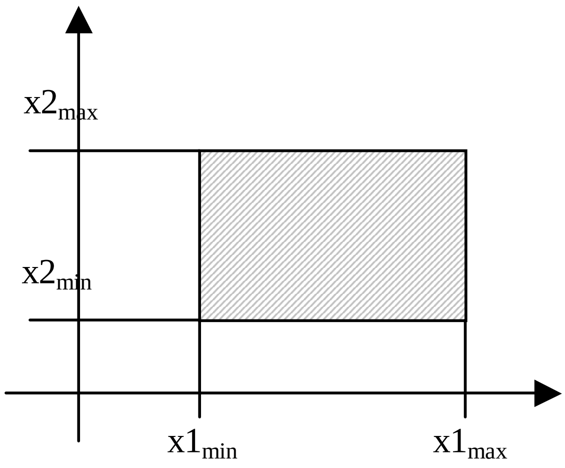


Рисунок 24

11 ПОНЯТИЕ ОБ ЭЛАСТИЧНОСТИ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

Эластичностью в экономике называют способность показателя откликаться на изменение того или иного фактора.

11.1 Коэффициент эластичности для однофакторной модели

Пусть в точке x_0 фактор изменится на величину Δx_0. Соответственно, y изменится на величину $y_0+\Delta y_0$. Относительное изменение *x* равно $\frac{\Delta x_0}{x_0}$. В процентах (темп роста) $T_{x_0} = \frac{\Delta x_0}{x_0} \cdot 100\%$. Соответственно, для y: $T_{y_0} = \frac{\Delta y_0}{y_0} \cdot 100\%$. Чтобы узнать, на сколько процентов изменится *y* при изменении *x* на величину *Δx₀*, возьмем отношение $\frac{T_{y_0}}{T_{x_0}} = \frac{\frac{\Delta y_0}{y_0} \cdot 100\%}{\frac{\Delta x_0}{x_0} \cdot 100\%} = \frac{\Delta y_0}{\Delta x_0} \cdot \frac{x_0}{y_0}$.

Так как (x_0, y_0) – произвольная точка, то можно записать: $\frac{T_y}{T_x} = \frac{\Delta y}{\Delta x} \cdot \frac{x}{y}$.

При $\Delta x \to 0$ $\frac{\Delta y}{\Delta x} = y'_x$.

Отсюда $E_x = \frac{x}{y} y'_x$ – формула для расчета коэффициента эластичности.

Коэффициент эластичности показывает, на сколько процентов увеличится (если $E_x > 0$) или уменьшится (если $E_x < 0$) показатель *y*, если фактор *x* изменится на 1% .

Вычислим коэффициент эластичности для некоторых моделей:

1 Линейная:

$$y = b_0 + b_1 x \Rightarrow y' = b_1 \Rightarrow E_x = \frac{b_1 x}{b_0 + b_1 x}.$$

Так как коэффициент эластичности для линейной функции не является величиной постоянной, а зависит от значения фактора x, то обычно рассчитывают средний показатель эластичности по формуле

$$E_x = b_1 \cdot \frac{\bar{x}}{\bar{y}}.$$

2 Степенная:

$$y = Ax^b \Rightarrow y' = A \cdot bx^{b-1} \Rightarrow E_x = \frac{xAbx^{b-1}}{Ax^b} = b(const)$$

Для степенной модели коэффициент эластичности постоянный и равен показателю степени.

3 Экспоненциальная:

$$y = Ae^{bx} \Rightarrow y' = A \cdot be^{bx} \Rightarrow E_x = \frac{xAbe^{bx}}{Ae^{bx}} = bx.$$

Так как коэффициенты эластичности представляют экономический интерес, приведем формулы расчета коэффициентов эластичности для наиболее распространенных типов уравнений регрессии:

Вид функции y(x)	Производная $y'(x)$	Коэффициент эластичности $E_x = \frac{x}{y(x)} y'_x$
$y = b_0 + b_1 x$	b_1	$\frac{b_1 x}{y(x)}$
$y = a\sqrt{x} + b$	$\frac{a}{2\sqrt{x}}$	$\frac{a\sqrt{x}}{2 \cdot y(x)}$
$y = ax^b$	abx^{b-1}	b
$y = \frac{a}{x} + b$	$\frac{-a}{x^2}$	$\frac{-a}{a + bx}$
$y = \frac{1}{ax + b}$	$\frac{-a}{(ax+b)^2}$	$-ax \cdot y(x)$
$y = ax^2 + b$	$2ax$	$\frac{2ax^2}{y(x)}$
$y = be^{ax}$	abe^{ax}	ax
$y = a\ln x + b$	$\frac{a}{x}$	$\frac{a}{y(x)}$

Несмотря на широкое использование в эконометрике коэффициентов эластичности, возможны случаи, когда их расчет экономического смысла не имеет. Это происходит в тех случаях, когда для рассматриваемых признаков не имеет смысла определение изменения значений в процентах (например, если x – стаж работы, измеряемый в годах, или x – качество почвы, измеряемое в баллах).

11.2 Коэффициент эластичности для многомерных моделей

Если показатель зависит от нескольких факторов, то, используя коэффициент эластичности, можно определить степень влияния каждого фактора на показатель.

$E_{x_i} = \frac{x_i}{y} y'_{x_i}$ – формула для расчета коэффициента частной эластичности.

Коэффициент частной эластичности показывает, на сколько процентов изменится y при изменении фактора x_i на 1% при прочих неизменных факторах.

Например: для моделирования зависимости между объемом выпущенной продукции y, трудозатратами x_1 и объемом основных средств производства x_2 используется функция Кобба-Дугласа $y = Ax_1^{\alpha_1} x_2^{\alpha_2}$. Так как эта функция степенная, то коэффициенты частной эластичности равны степени соответствующего фактора: $E_{x_1} = \alpha_1$, $E_{x_2} = \alpha_2$. Если $\alpha_1 < \alpha_2$, то x_2 больше влияет на y, т.е. если увеличивать основные средства, то объем выпускаемой продукции будет расти быстрее, чем при увеличении трудозатрат.

12 СИСТЕМЫ УРАВНЕНИЙ

При моделировании сложных экономических объектов часто приходится вводить не одно, а несколько связанных между собой уравнений, то есть описывать модель системой уравнений.

Например, простейшая макроэкономическая кейсианская модель потребления может быть представлена в следующем виде:

$$\begin{cases} C = a + by + \varepsilon, \\ y = C + I, \end{cases}$$

где C – личное потребление в постоянных ценах, y – национальный доход в постоянных ценах, I – инвестиции в постоянных ценах, ε – случайная составляющая.

Наличие связи между C и y, входящими в оба уравнения, требует корректировки метода наименьших квадратов для оценивания параметров модели a и b.

Есть два вида таких моделей: система внешне не связанных между собой уравнений (они связаны только наличием корреляции между ошибками) и система одновременных уравнений.

Система совместных, одновременных уравнений обычно содержит эндогенные и экзогенные переменные.

Эндогенные переменные – это зависимые переменные, которые определяются внутренней структурой исследуемого экономического процесса.

Экзогенные переменные – это переменные, которые не зависят от внутренней структуры исследуемого экономического процесса. Они влияют на эндогенные переменные, но не зависят от них.

Разделение переменных на эндогенные и экзогенные зависит от теоретической концепции принятой модели. Экзогенные переменные могут выступать в одних моделях как эндогенные, а в других как экзогенные переменные.

Система одновременных, взаимозависимых уравнений вида

$$\begin{cases} y_1 = b_{12}y_2 + b_{13}y_3 + \ldots + b_{1n}y_n + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m + \varepsilon_1, \\ y_2 = b_{21}y_1 + b_{23}y_3 + \ldots + b_{2n}y_n + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m + \varepsilon_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ y_n = b_{n1}y_1 + b_{n2}y_2 + b_{n3}y_3 + \ldots + b_{nn-1}y_{n-1} + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{cases} \qquad (17)$$

называется также *структурной формой модели.*

В системе одни и те же переменные (y) одновременно рассматриваются как зависимые переменные в одних уравнениях и как независимые в других. Каждое уравнение системы не может рассматриваться самостоятельно, поэтому традиционный МНК не применим.

Коэффициенты b_{ij} и $a_{ik}(i = 1..n,\ j = 1..n, k = 1..m)$ называют *структурными коэффициентами модели.*

Использование МНК для оценивания структурных коэффициентов модели дает смещенные и несостоятельные оценки. Поэтому обычно для определения структурных коэффициентов модель преобразуется в приведенную форму модели, которая представляет собой систему линейных функций эндогенных от экзогенных:

$$\begin{cases} \hat{y}_1 = \delta_{11}x_1 + \delta_{12}x_2 + \ldots + \delta_{1m}x_m, \\ \hat{y}_2 = \delta_{21}x_1 + \delta_{22}x_2 + \ldots + \delta_{2m}x_m, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \hat{y}_n = \delta_{n1}x_1 + \delta_{n2}x_2 + \ldots + \delta_{nm}x_m, \end{cases} \tag{18}$$

где $\delta_{ik}\,(i = 1...n, k = 1...m)$ – коэффициенты приведенной модели.

Приведенная форма модели по своему виду сходна с системой независимых уравнений, параметры которой определяются по МНК. Коэффициенты модели (18) представляют собой нелинейные функции коэффициентов структурной формы модели.

Приступать к оцениванию коэффициентов того или иного уравнения в системе (17) имеет смысл лишь после того, как установлена его идентифицируемость.

Идентифицируемость – это единственность соответствия между приведенной и структурной формами модели.

Для этого необходимо, чтобы выполнялось неравенство

$$n_s - 1 \le m - m_s, \tag{19}$$

где n_s – количество эндогенных переменных, которые входят в s-е уравнение структурной формы модели;

m – общее количество экзогенных переменных модели, которые входят в s-е уравнение структурной формы модели.

Неравенство (19) называется *порядковым (счетным) условием* и является необходимым условием идентифицируемости уравнения.

Если в выражении (19) стоит знак равенства, то модель точно *идентифицируема*, ее структурные коэффициенты определяются однозначно, единственным образом по коэффициентам приведенной формы модели.

Число параметров структурной модели равно числу параметров приведенной формы модели.

Если неравенство (19) выполняется как строгое, то модель *сверхидентифицируема*. В этом случае число коэффициентов в приведенной модели больше числа коэффициентов структурной модели, то есть один и тот же структурный коэффициент допускает разные выражения через коэффициенты приведенной формы. Это сужает область его применения как с теоретической, так и с практической точки зрения.

Если неравенство (19) не выполняется, то модель *неидентифицируема* и структурные коэффициенты не могут быть оценены через коэффициенты приведенной формы модели.

Выполнение условия идентификации проверяется для каждого уравнения системы. Для оценки параметров структурной модели система должна быть идентифицируема или сверхидентифицируема.

Порядковое условие (19) есть необходимое, но не достаточное условие идентификации. Уравнение идентифицируемо, если по отсутствующим в нем переменным (эндогенным и экзогенным) можно из коэффициентов при них в других уравнениях системы получить матрицу, определитель которой отличен от нуля, а ранг матрицы не меньше, чем число эндогенных переменных в системе без одного. Это условие называется *ранговым* и является достаточным условием идентификации. Подробнее об идентификации см. [7].

Коэффициенты структурной модели могут быть оценены разными способами в зависимости от вида системы одновременных уравнений. Наиболее распространенными являются следующие два метода:

1 Косвенный (непрямой) метод наименьших квадратов. При идентифицируемости уравнения оценки структурных коэффициентов можно найти, оценив методом наименьших квадратов приведенную форму модели. Затем коэффициенты приведенной формы преобразуются в параметры структурной модели.

2 Для оценки параметров сверхидентифицируемой модели используется двухшаговый метод наименьших квадратов (2МНК). Основная идея 2МНК – на основе приведенной формы модели получить для сверхидентифицируемого уравнения теоретические значения эндогенных переменных, содержащихся в правой части уравнения. Далее, подставив их вместо фактических значений, применить МНК к структурной форме сверхидентифицируемого уравнения.

ЧАСТЬ 2

ВВЕДЕНИЕ

При обработке выборок каждого вида используется специфический математический аппарат (методы математической статистики и методы анализа случайных процессов). Но в любом случае такая обработка сопряжена с громоздкими и трудоёмкими вычислениями. Поэтому необходимо использовать математические пакеты, специально предназначенные для обработки статистических данных. В среде Excel for Windows есть специальные возможности для таких расчетов.

1 КРАТКИЕ СВЕДЕНИЯ ОБ ЭКОНОМЕТРИЧЕСКОМ АНАЛИЗЕ В ПАКЕТЕ EXCEL

1.1 Настройка пакета анализа

Для проведения эконометрического анализа в пакете Excel должен быть установлен «Пакет анализа» (рис.25). Путь: Сервис – Надстройки – Пакет анализа – Ок.

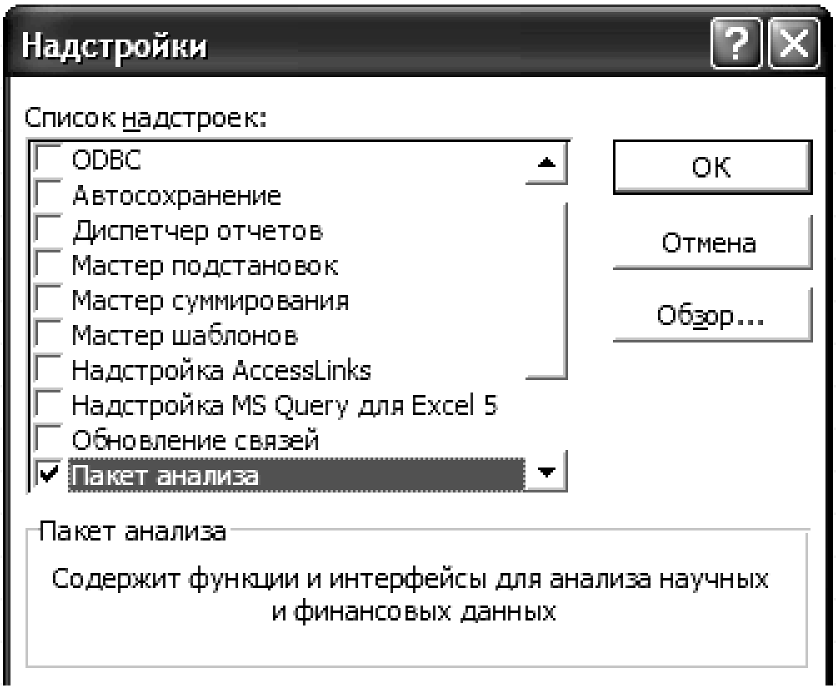


Рисунок 25

После этого в меню «Сервис» добавится строка «Анализ данных».

1.2 Ввод данных

Исходные данные вводятся на рабочий лист пакета Excel (табл. 4).

Таблица 4

Строки Excel	**Столбцы Excel**	
	A	**B**
1	x	y
2	8540	38,34
3	2911	44,69
4	6630	39,4
5	8492	38,93
6	2901	46,96
7	5410	39,48
8	1920	46,05
9	2569	43,5
10	3520	56,11
11	2340	42,79
12	6921	40,15
13	7671	40,44
14	1586	69,76
15	3223	42,99
16	7224	40,69

1.3 Построение диаграммы рассеивания (корреляционного поля)

По исходным данным строится диаграмма рассеивания с помощью «Мастера диаграмм», тип диаграммы - точечная.

Диаграмма форматируется таким образом, чтобы наиболее ясно представлялись исходные данные (рис. 26).

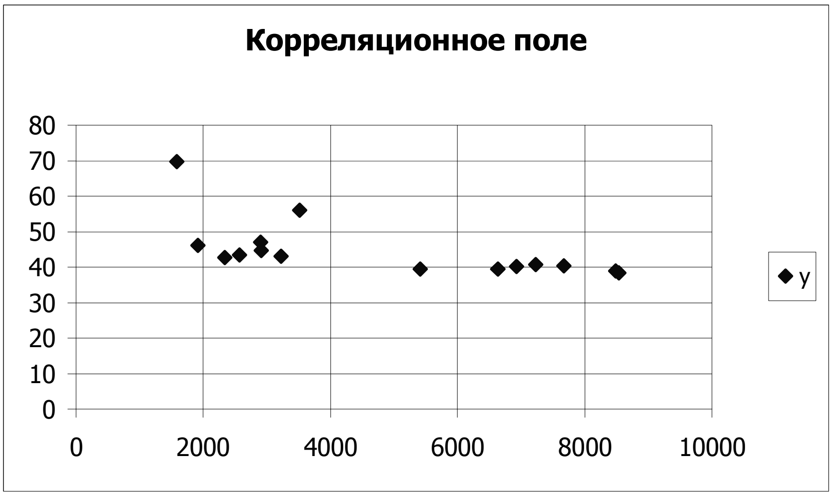


Рисунок 26

1.4 Нахождение коэффициента корреляции

Выбирается пункт меню Сервис - Анализ данных – Корреляция.

Задается входной интервал для X и Y – A1:B16 (группирование данных – по столбцам), устанавливается флажок в окошке «Метки» (это означает, что в первой строке – метки (имена данных) – X и Y), «Выходной диапазон» – на новый лист или указывается выходной интервал на исходном листе (рис. 27).

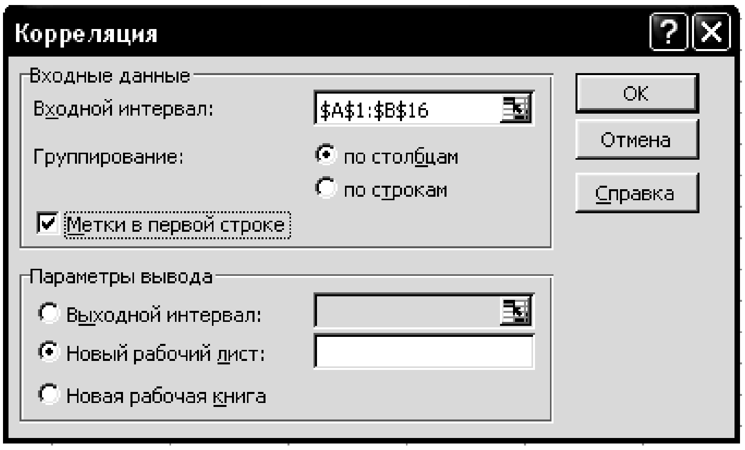


Рисунок 27

Замечания:

1 Для многофакторной регрессии выделяется весь диапазон данных (X1, X2, Y).

2 Для однофакторной регрессии «Выходной диапазон» – выделить блок 3 на 3, для двухфакторной – 4 на 4 ячейки.

3 Полученная матрица симметрична относительно главной диагонали.

Получаем матрицу следующего вида для однофакторной регрессии:

	X	Y
X	1	
Y	-0,61975	1

Соответствие:

	X	Y
X	r_{xx}	
Y	r_{xy}	r_{yy}

1.5 Нахождение основных числовых характеристик

Чтобы найти основные числовые характеристики, выбираем пункт меню Сервис – Анализ данных – Описательная статистика.

Здесь: задаем входной интервал для X и Y – A1:B16; устанавливаем флажок в окошках «Метки» и «Итоговая статистика»; «Выходной диапазон» – на новый лист или указать выходной интервал (блок из 15 строк и 4 столбцов для однофакторной регрессии) на исходном листе (рис. 28).

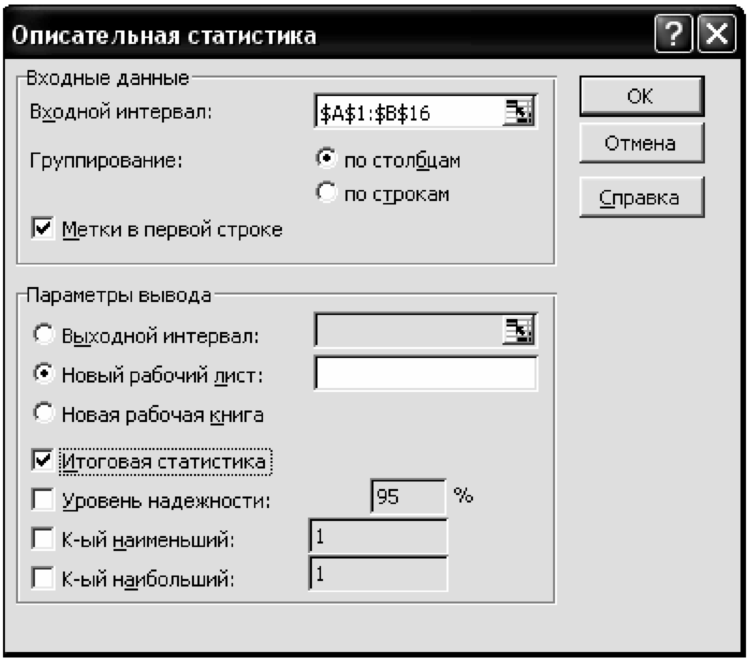


Рисунок 28

Получается следующая таблица для однофакторной регрессии:

Таблица 5

	X	*Y*	*Пояснения*
Среднее	4790,533333	44,68533333	Среднее значение
Стандартная ошибка	657,9484194	2,134904897	
Медиана	3520	42,79	
Мода	#N/A	#N/A	
Стандартное отклонение	2548,223271	8,268451113	Среднеквадратическое отклонение
Дисперсия выборки	6493441,838	68,36728381	Дисперсия выборки
Эксцесс	-1,712674833	6,049645013	
Асимметричность	0,280614344	2,374906795	
Интервал	6954	31,42	
Минимум	1586	38,34	Минимальное значение
Максимум	8540	69,76	Максимальное значение
Сумма	71858	670,28	
Счет	15	15	Объем выборки

1.6 Нахождение параметров линейной регрессии

Чтобы найти параметры регрессии, выбираем пункт меню Сервис - Анализ данных - Регрессия. Здесь задаем диапазоны отдельно для Y, отдельно – для X (для многофакторной регрессии в поле «Входной интервал X» выделяем все значения факторов), устанавливаем флажок в окошке «Метки», «Выходной диапазон» – на новый лист. Ок (рис. 29).

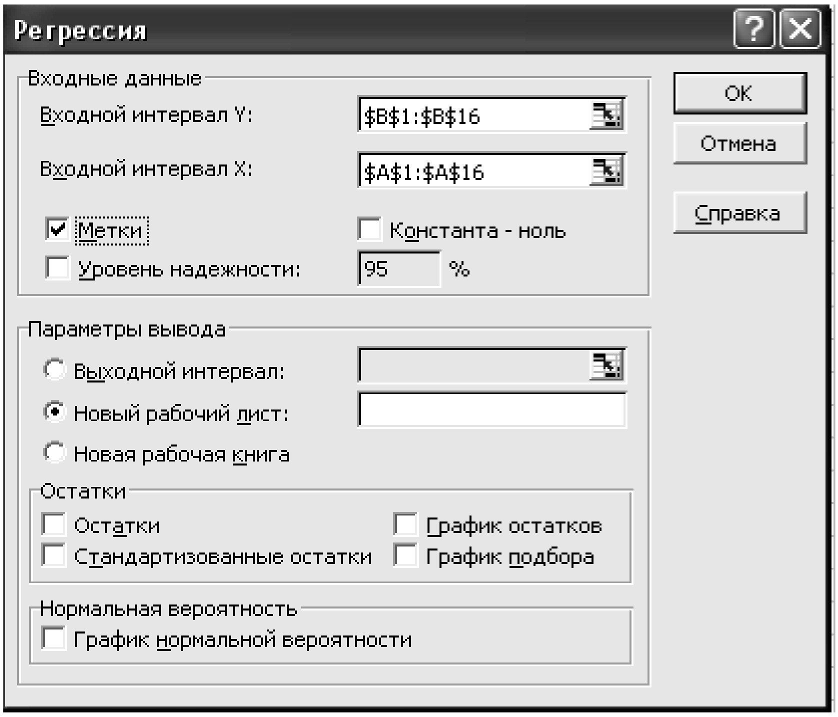


Рисунок 29

Результат получили в виде таблицы:

ВЫВОД ИТОГОВ								
Регрессионная статистика								
Множественный R	0,61974714							
R-квадрат	0,384086518							
Нормированный R-квадрат	0,336708557							
Стандартная ошибка	6,734050364							
Наблюдения	15							
Дисперсионный анализ								
	df	*SS*	*MS*	*F*	*Значимость F*			
Регрессия	1	367,6253273	367,6253273	8,106860574	0,013729018			
Остаток	13	589,516646	45,34743431					
Итого	14	957,1419733						
	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*	*Нижние 95,0%*	*Верхние 95,0%*
Y-пересечение	54,31885511	3,804055967	14,27919452	2,53059E-09	46,10069341	62,5370168	46,1006934	62,537017
x	-0,00201095	0,000706277	-2,847254919	0,013729018	-0,00353677	-0,00048513	-0,0035368	-0,0004851

Из получившегося окна «ВЫВОД ИТОГОВ» выбираем следующие величины (табл.6).

Таблица 6

Название в Excel	**Смысл**	*Для данного примера*
Y - пересечение	Коэффициент b0	54,31885511
X	Коэффициент b1	-0,00201095
R - квадрат	Коэффициент детерминации R^2	0,384086518
Множественный R	Коэффициент корреляции (по модулю)	0,61974714
Стандартная ошибка в регрессионной статистике	Среднее квадратическое отклонение остатков	6,734050364
Наблюдения	Объем выборки	15
F	Fнабл	8,106860574
Значимость F	Уровень значимости для критерия Фишера	0,013729018
Df регрессия остаток	Число степеней свободы: k_1 k_2	 1 13
Стандартная ошибка (рядом со значением коэффициентов)	Дисперсия коэффициентов	3,804055967 0,000706277
Столбец t - статистика	Наблюдаемое значение критерия Стьюдента	14,27919452 -2,847254919
Столбец P - значение	Значимость коэффициентов по критерию Стьюдента	$2,53059 \cdot 10^{-9}$ 0,013729018

1.7 Нахождение критической точки распределения Стьюдента

Выбираем команду «Вставка функции», категорию «Статистические», функцию СТЬЮДРАСПОБР. Вводим требуемую вероятность (0,05) и число степеней свободы (k2 = n – 2). Получим для однофакторной регрессии 2,16.

2 ПОСЛЕДОВАТЕЛЬНОСТЬ ПОСТРОЕНИЯ МОДЕЛЕЙ

2.1 Задание № 3(1)

План построения линейной однофакторной модели

В первом задании строим линейную модель вида $y = b_0 + b_1 x$ для показателя y и первого фактора x.

Последовательность действий:

1 Вводим данные. Определяем основные статистики (см. табл.5).

2 Строим диаграмму рассеивания (корреляционное поле).

3 Определяем тесноту линейной связи по коэффициенту корреляции.

4 Записываем линейную модель вида $y = b_0 + b_1 x$ (см. табл.6).

5 Определяем общее качество модели по коэффициенту детерминации R^2. Проверяем полученную модель на адекватность по критерию Фишера. Все дальнейшие расчеты выполняются только при условии адекватности модели исходным статистическим данным.

6 Проверяем статистическую значимость коэффициентов модели.

7 По полученной модели рассчитываем значения показателя y для всех точек выборки и в точке прогноза (точку прогноза выбираем произвольно из области прогноза).

8 Рассчитываем полуширину доверительного интервала

$$\delta = \sigma_e \cdot t_\gamma \cdot \sqrt{1 + \frac{1}{n} + \frac{(x_{np} - \bar{x})^2}{(n-1)D(x)}},$$

где σ_e– среднеквадратическое отклонение выборочных точек от линии регрессии (см. табл.6);

t_γ – критическая точка распределения Стьюдента для надежности γ = 0,95 и k2 = 13;

n =15 – объем выборки;

D(x) – дисперсия выборки (см.табл.5);

$\bar{x}$ – среднее значение;

$x_{пр}$ – точка из области прогнозов (от x_{min} до x_{max}).

9 Рассчитываем доверительный интервал для всех точек выборки и в точке прогноза: (y–δ, y+δ).

10 Рассчитываем коэффициент эластичности

$$E_x = \frac{x}{y(x)} \cdot y'_x .$$

Для линейной модели $y'_x = b_1$. Получим

$$E_x = \frac{b_1 x}{y(x)},$$

где y(x) – рассчитанное по модели значение показателя.

11 Строим доверительную область – диаграмму вида (рис. 30):

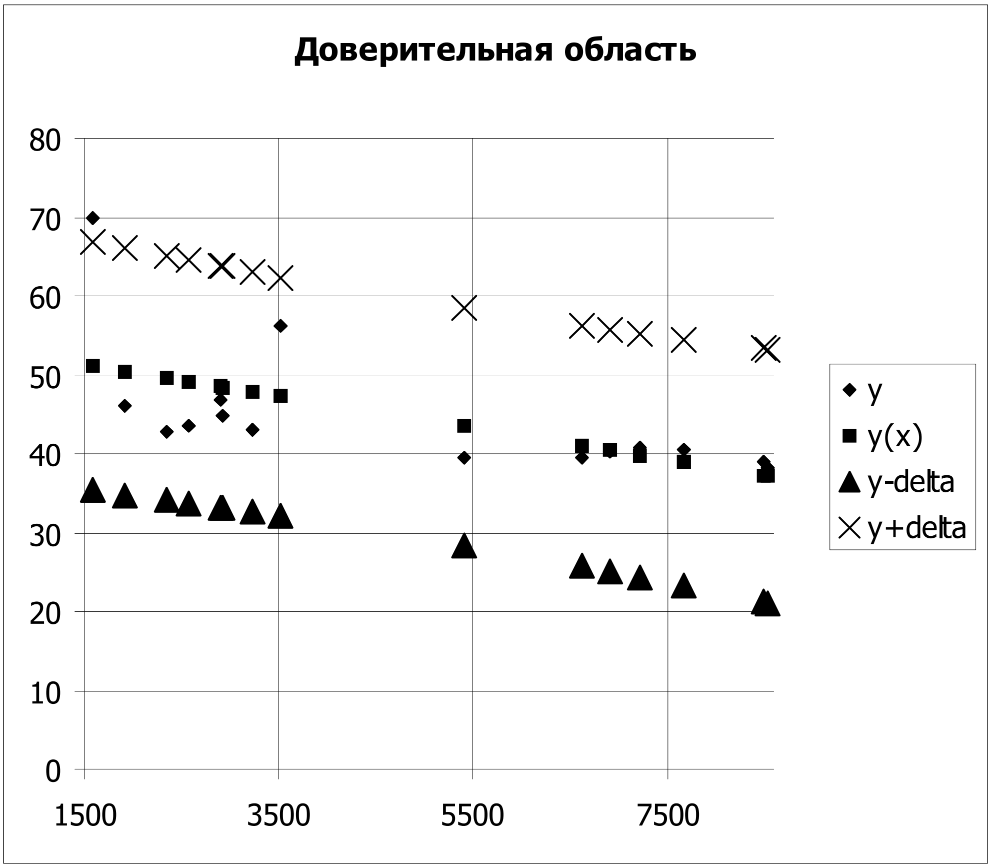


Рисунок 30

12 Используя полученные данные и теоретические сведения, делаем эконометрический анализ – описываем процесс построения модели и все сопутствующие расчеты.

2.2 Задание № 3(2)

План построения нелинейной однофакторной модели

Во втором задании требуется построить нелинейную модель зависимости показателя y от фактора x. Нелинейная функция задана для каждого варианта.

Последовательность действий:

1 Вводим данные. Определяем основные статистики (см. табл.5). Строим корреляционное поле. По его виду выдвигаем гипотезу о нелинейной зависимости между x и y.

2 Линеаризуем нелинейную модель с помощью формул перехода (таблица 7).

Таблица 7

Вид зависимости	Линеаризующая подстановка		Обратное преобразование	
	$u=$	$v=$	$a=$	$b=$
$y=\frac{a}{x}+b$	$\frac{1}{x}$	y	b_1	b_0
$y=a\sqrt{x}+b$	$\sqrt{x}$	y	b_1	b_0
$y=ax^b$	$\ln x$	$\ln y$	e^{b_0}	b_1
$y=a\cdot\ln x+b$	$\ln x$	y	b_1	b_0
$y=e^{ax}\cdot b$	x	$\ln y$	b_1	e^{b_0}
$y=a\cdot x^2+b$	x^2	y	b_1	b_0

Получаем линейную модель относительно новых переменных $v = b_0 + b_1 u$.

3 Определяем тесноту линейной связи по коэффициенту корреляции.

4 Записываем линейную модель вида $v = b_0 + b_1 u$ (см. табл.6).

5 Определяем общее качество модели по коэффициенту детерминации R^2. Проверяем полученную модель на адекватность по критерию Фишера. Все дальнейшие расчеты выполняются только при условии адекватности модели исходным статистическим данным.

6 Проверяем статистическую значимость коэффициентов модели.

7 По полученной модели рассчитываем значения показателя v для всех точек выборки и в точке прогноза (точку прогноза выбираем произвольно из области прогноза).

8 Рассчитываем полуширину доверительного интервала

$$\delta = \sigma_e \cdot t_\gamma \cdot \sqrt{1 + \frac{1}{n} + \frac{(u_{np} - \bar{u})^2}{(n-1)D(u)}},$$

где σ_e – среднеквадратическое отклонение выборочных точек от линии регрессии (см. табл.6);

t_γ – критическая точка распределения Стьюдента для надежности γ = 0,95 и k2 = 13 (см. табл.6);

n =15 – объем выборки;

D(u) – дисперсия выборки (см. табл.5);

$\bar{u}$ – среднее значение;

u_{np} – точка из области прогнозов (от u_{min} до u_{max}).

9 Рассчитываем доверительный интервал для всех точек

выборки и в точке прогноза: (v–δ, v+δ).

10 Если линеаризованная модель $v = b_0 + b_1 u$ адекватна (по критерию Фишера), то и исходная нелинейная модель будет адекватна.

11 По формулам обратного перехода (см. табл.7) пересчитываем значения y, y_{min} (левая граница доверительного интервала), y_{max} (правая граница доверительного интервала).

12 Рассчитываем коэффициент эластичности

$$E_x = \frac{x}{y(x)} \cdot y'_x.$$

13 Строим доверительную область.

14 Используя полученные данные и теоретические сведения, делаем эконометрический анализ – описываем процесс построения модели и все сопутствующие расчеты.

2.3 Задание № 3(3)

План построения линейной двухфакторной модели

В третьем задании требуется построить линейную двухфакторную модель вида $y = b_0 + b_1 x_1 + b_2 x_2$ для показателя y и факторов x_1 и x_2.

Последовательность действий:

1 Вводим данные.

2 Определяем основные статистики (см. табл.5).

3 По корреляционной таблице проверяем факторы на коллинеарность.

4 Записываем линейную модель вида $y = b_0 + b_1 x_1 + b_2 x_2$ (см. табл.6).

5 Определяем общее качество модели по коэффициенту детерминации R^2. Проверяем полученную модель на адекватность по критерию Фишера. Все дальнейшие расчеты выполняются только при условии адекватности модели исходным статистическим данным.

6 Проверяем статистическую значимость коэффициентов модели.

7 По полученной модели рассчитываем значения показателя у для всех точек выборки и в точке прогноза (точку прогноза выбираем произвольно из области прогноза).

8 Рассчитываем частичные коэффициенты эластичности:

- по фактору x_1:

$$E_{x_1} = \frac{x_1}{y(x)} \cdot y'_{x_1} = \frac{b_1 x_1}{y(x)};$$

- по фактору x_2:

$$E_{x_2} = \frac{x_2}{y(x)} \cdot y'_{x_2} = \frac{b_2 x_2}{y(x)}.$$

9 Используя полученные данные и теоретические сведения, делаем эконометрический анализ – описываем процесс построения модели и все сопутствующие расчеты.

ЧАСТЬ 3

1 ВЫБОР ВАРИАНТА

Выбор варианта задания производится по таблице 8. В ней: 1-е число – номер варианта задания 1; 2-е – номер варианта задания 2; 3-е – номер варианта задания 3.

Таблица 8

Последние цифры зачетки	Задания	Последние цифры зачетки	Задания	Последние цифры зачетки	Задания	Последние цифры зачетки	Задания
00	1,1,1	25	6,4,26	50	11,7,21	75	16,10,16
01	2,2,2	26	7,5,27	51	12,8,22	76	17,11,17
02	3,3,3	27	8,6,28	52	13,9,23	77	18,12,18
03	4,4,4	28	9,7,29	53	14,10,24	78	19,1,19
04	5,5,5	29	10,8,30	54	15,11,25	79	20,2,20
05	6,6,6	30	11,9,1	55	16,12,26	80	1,3,21
06	7,7,7	31	12,10,2	56	17,1,27	81	2,4,22
07	8,8,8	32	13,11,3	57	18,2,28	82	3,5,23
08	9,9,9	33	14,12,4	58	19,3,29	83	4,6,24
09	10,10,10	34	15,1,5	59	20,4,30	84	5,8,25
10	11,11,11	35	16,2,6	60	1,6,1	85	6,9,26
11	12,12,12	36	17,4,7	61	2,7,2	86	7,10,27
12	13,2,13	37	18,5,8	62	3,8,3	87	8,11,28
13	14,3,14	38	19,6,9	63	4,9,4	88	9,12,29
14	15,4,15	39	20,7,10	64	5,10,5	89	10,1,30
15	16,5,16	40	1,8,11	65	6,11,6	90	11,2,1
16	17,6,17	41	2,9,12	66	7,12,7	91	12,3,2
17	18,7,18	42	3,10,13	67	8,1,8	92	13,4,3
18	19,8,19	43	4,11,14	68	9,2,9	93	14,5,4
19	20,9,20	44	5,12,15	69	10,3,10	94	15,6,5
20	1,10,21	45	6,1,16	70	11,4,11	95	16,7,6
21	2,11,22	46	7,2,17	71	12,5,12	96	17,9,7
22	3,12,23	47	8,3,18	72	13,7,13	97	18,10,8
23	4,1,24	48	9,5,19	73	14,8,14	98	19,11,9
24	5,3,25	49	10,6,20	74	15,9,15	99	20,12,10

2 ТРЕБОВАНИЯ К ВЫПОЛНЕНИЮ КОНТРОЛЬНОЙ РАБОТЫ

1 Контрольная работа выполняется в тетради или на листах формата А4, жестко скрепленных между собой.

2 Каждое задание должно содержать условие, эконометрический анализ представленных данных, распечатки листов пакета Excel с расчетами и формулами.

3 Эконометрический анализ (допускается печатный и рукописный варианты) включает в себя подробное описание построения модели, проверку ее адекватности и нахождение прогнозов с использованием произведенных расчетов на основе исходных данных. В анализе используются необходимые определения и формулы, как в общем (теоретическом) виде, так и с конкретными данными.

Внимание! Все три задания выполняются по одной таблице данных (свой вариант). Для первого задания нужно взять первый фактор и показатель; для второго – второй фактор и показатель; для третьего – вся таблица.

3 ЗАДАНИЯ К КОНТРОЛЬНОЙ РАБОТЕ

3.1 Задание 1

Ответить на *теоретический вопрос*.

1 Основные задачи эконометрики. Этапы эконометрического анализа

2 Классификация эконометрических моделей. Информационная база эконометрики

3 Генеральная совокупность. Выборка. Объем выборки.

Среднее значение. Дисперсия. Среднеквадратическое отклонение

4 Корреляционное поле. Центр рассеивания

5 Коэффициент корреляции и его свойства

6 Метод наименьших квадратов для однофакторной линейной регрессии

7 Свойства линейной регрессии

8 Статистическая гипотеза. Нулевая и конкурирующая гипотезы. Ошибки 1 и 2 рода

9 Критерий Фишера. Наблюдаемое и критическое значения критерия

10 Проверка линейной регрессии на адекватность

11 Коэффициент детерминации

12 Проверка модели на адекватность с помощью критерия Фишера

13 Область прогноза для однофакторной и двухфакторной модели. Доверительный интервал. Коэффициент доверия

14 Прогноз по линейной однофакторной модели с учетом доверительного интервала

15 Коэффициент эластичности для однофакторной модели

16 Виды нелинейных однофакторных моделей. Способ их линеаризации

17 Алгоритм построения нелинейных эконометрических моделей

18 Понятие многофакторной модели и этапы ее построения

19 Коллинеарность и мультиколлинеарность

20 Коэффициент эластичности для многомерных моделей

3.2 Задание 2

Найти коэффициент эластичности для указанной модели в заданной точке x (табл.9). Сделать экономический вывод.

Таблица 9

Номер варианта	Модель	x
1	$y = \frac{2}{x} + 5$	0,2
2	$y = \frac{1}{2x+1}$	1
3	$y = 3x^2 + 1$	1
4	$y = 6x^5$	1
5	$y = 2\sqrt{x} + 4$	4
6	$y = 3e^{2x}$	2
7	$y = \frac{2e^{5x}}{5}$	1
8	$y = 3\ln x + 2$	1
9	$y = 2x^3 + 1$	1
10	$y = \frac{e^x}{2}$	2
11	$y = -\frac{x}{4}$	1
12	$y = \frac{1}{x} + 1$	1
13	$y = \frac{e^{x+1}}{6}$	3
14	$y = \ln\frac{x}{4} + 1$	4
15	$y = \sqrt{2x+4}$	1
16	$y = 5x + 5$	1
17	$y = \frac{x^5}{2}$	1
18	$y = \frac{\sqrt{3x}}{3}$	2

Продолжение табл.9

Номер варианта	**Модель**	**x**
19	$y=\frac{e^{3x}}{3}$	1
20	$y=\frac{4}{4x+3}$	1
21	$y=5x^3+1$	1
22	$y=\sqrt{x^3}+1$	1
23	$y=\frac{3}{x^3}$	2
24	$y=-3x+3$	2
25	$y=\frac{\sqrt{x}}{2}$	1
26	$y=\frac{2x^2+2}{3}$	1
27	$y=\frac{2}{3x^3}$	2
28	$y=x^4+1$	1
29	$y=3\ln 2x+1$	0,5
30	$y=\frac{4}{x^2+1}$	1

3.3 Задание 3

Для представленных данных выполнить следующее *задание*:

1 Провести эконометрический анализ *линейной зависимости показателя от первого фактора*. Сделать прогноз для любой точки из области прогноза, построить доверительную область. Найти коэффициент эластичности в точке прогноза.

2 Провести эконометрический анализ *нелинейной зависимости показателя от второго фактора*, воспользовавшись подсказкой. Сделать прогноз для любой точки из области прогноза, построить доверительную область. Найти коэффициент эластичности в точке прогноза.

3 Провести эконометрический анализ *линейной зависимости показателя от двух факторов*. Сделать точечный прогноз для

любой точки из области прогноза. Найти частичные коэффициенты эластичности в точке прогноза.

Варианты

1 Производительность труда, фондоотдача и уровень рентабельности по плодоовощным консервным заводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Производительность труда, грн	Фондоотдача, грн	
1	8540	1,24	38,34
2	2911	0,63	44,69
3	6630	1,18	39,4
4	8492	1,12	38,93
5	2901	0,44	46,96
6	5410	1,19	39,48
7	1920	0,48	46,07
8	2569	0,65	43,5
9	3520	0,26	56,11
10	2340	0,75	42,79
11	6921	1,03	40,15
12	7671	0,89	40,44
13	1586	0,16	69,76
14	3223	0,67	42,99
15	7224	0,90	40,69

Нелинейную зависимость принять $y = \frac{a}{x} + b$.

2 Известны следующие данные об убыточности производства говядины по КСП административных районов области за год:

Номер района	Фактор		Уровень убыточности, %
	Среднесуточный прирост, грн	Себестоимость 1 ц, грн	
1	249	138,99	37,7
2	231	105,86	29,7
3	245	114,19	26,8
4	242	131,73	28,4
5	250	139,86	43,2
6	190	141,52	48
7	283	118,9	33,9
8	273	163,26	29,1
9	290	143,7	29,8
10	150	221,88	66
11	294	102,4	19,6
12	196	149,06	48,8
13	241	135,5	27,4
14	214	178,17	53,6
15	188	229,36	62,1

Нелинейную зависимость принять $y = a\sqrt{x} + b$.

3 В таблице приведены данные об удельном весе пашни, лугов и пастбищ в сельскохозяйственных угодьях и уровне рентабельности производства сельскохозяйственной продукции по районам области за год.

Номер района	Фактор		Уровень рентабельности всей сельскохозяйственной продукции, %
	Удельный вес пашни в сельскохозяйственных угодьях, %	Удельный вес лугов и пастбищ, %	
1	80,00	20,0	2,0
2	87,20	12,8	1,8
3	90,80	9,2	1,1
4	84,70	15,3	3,5
5	81,40	18,6	10,1
6	91,30	10,8	3,3
7	71,30	28,7	24,2
8	86,20	13,8	1,9
9	71,40	28,6	20,8
10	77,70	22,9	19,2
11	86,00	14,0	3,4
12	87,00	13,0	2,7
13	87,20	12,8	1,4
14	75,00	25,0	20,1
15	86,20	13,8	7,8

Нелинейную зависимость принять $y = ax^b$.

4 Производительность труда, фондоотдача и уровень рентабельности по хлебозаводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда в расчете на 1 работника, грн	
1	20,1	4322	12,2
2	64,2	13381	17,6
3	61,1	14181	17,5
4	13,3	3363	10,3
5	10,8	5177	12,8
6	17,2	3720	13,1
7	34,1	9900	16,9
8	32,3	8931	14,4
9	27,8	6740	16,0
10	24,2	6980	16,4
11	35,6	14333	18,3
12	17,1	3930	10,8
13	13,9	2500	10,0
14	25,5	5342	14,0
15	31,1	6743	16,1

Нелинейную зависимость принять $y = a \ln x + b$.

5 В таблице приведены данные о затратах на 1 грн. товарной продукции, удельном весе простоев оборудования и уровне рентабельности

по молокозаводам области за год.

Номер завода	Фактор		Уровень рентабельности, %
	Затраты на 1 грн товарной продукции, грн	Удельный вес простоев оборудования, %	
1	0,59	8,1	15,45
2	2,25	11,8	20,33
3	0,36	7,4	14,67
4	1,37	9,4	16,05
5	5,44	17,8	37,39
6	2,02	12,1	22,19
7	1,74	10,2	17,01
8	3,10	14,1	26,24
9	1,73	10,1	16,74
10	4,59	16,7	33,83
11	6,76	19,4	43,58
12	1,84	10,4	17,24
13	4,73	16,2	30,62
14	4,58	16	30,1
15	3,66	15,1	28,81

Нелинейную зависимость принять $y = ax^b$.

6 Производительность труда, фондоотдача и уровень рентабельности по плодоконсервным заводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда, грн	
1	1,08	7343	20,1
2	1,05	3991	12,9
3	0,99	5760	18,0
4	1,02	3000	11,7
5	0,98	5241	17,9
6	1,04	4500	16,8
7	1,03	4300	15,6
8	1,10	3210	14,3
9	1,03	6743	18,1
10	0,89	5234	17,8
11	0,78	2500	13,0
12	0,99	3930	14,2
13	1,43	14333	24,2
14	1,03	6980	20,0
15	1,05	6740	19,3

Нелинейную зависимость принять $y = a\ln x + b$.

7 Производительность труда, фондоотдача и уровень рентабельности по хлебозаводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда в расчете на 1 работника, грн	
1	33,4	3447	12,3
2	29,1	3710	14,7
3	25,3	2827	10,9
4	27,1	2933	16,1
5	43,3	5428	22,3
6	47,2	5001	21,1
7	49,3	6432	24,3
8	35,7	4343	13,3
9	45,8	7321	27,6
10	43,4	6432	28,3
11	42,1	6003	25,1
12	40,1	5342	20,2
13	33,3	4341	13,7
14	41,2	5040	19,9
15	39,7	4343	14,2

Нелинейную зависимость принять $y = a\ln x + b$.

8 В таблице приведены данные об удельном весе рабочих со специальной технической подготовкой, удельном весе механизированных работ и производительности труда по плодоовощным заводам области за год.

Номер завода	Фактор		Производительность труда, грн
	Удельный вес рабочих с технической подготовкой, %	Удельный вес механизированных работ, %	
1	64	84	4300
2	61	83	4150
3	47	47	3000
4	46	55	3420
5	49	69	3300
6	54	78	4300
7	53	73	3420
8	61	81	4100
9	57	77	3700
10	54	72	3500
11	60	70	4000
12	67	85	4450
13	63	83	4270
14	50	70	3300
15	67	81	4500

Нелинейную зависимость принять $y = e^{ax} \cdot b$.

9 Известны следующие данные об убыточности производства говядины по КСП административных районов области за год:

Номер района	Фактор		Уровень убыточности, %
	Затраты на 1 ц прироста, человеко-часов	Среднесуточный прирост, грн	
1	76,8	249	37,7
2	76	271	23,7
3	74,6	245	26,8
4	79	242	28,4
5	76,6	250	43,2
6	93,4	190	48
7	71,8	283	33,9
8	93	223	49,1
9	66,6	290	29,8
10	112	150	69
11	66,9	304	19,6
12	94,6	196	53,8
13	70	241	27,4
14	92,2	214	53,6
15	89,2	188	62,1

Нелинейную зависимость принять $y = ax^b$.

10 В таблице приведены данные об уровне технической подготовки рабочих, стаже их работы и уровне заработной платы по сахарным заводам области за год.

Номер завода	Фактор		Заработная плата за месяц, грн
	Удельный вес рабочих с технической подготовкой, %	Удельный вес рабочих со стажем свыше 10 лет, %	
1	40	35	192,20
2	33	40	202,33
3	37	43	204,20
4	39	47	199,95
5	37	42	204,37
6	41	42	199,80
7	49	44	220,11
8	38	48	218,33
9	55	67	263,30
10	43	49	222,72
11	56	63	239,39
12	47	46	217,01
13	44	47	223,40
14	55	62	237,87
15	54	62	234,20

Нелинейную зависимость принять $y = e^{ax} \cdot b$.

11 Производительность труда, фондоотдача и уровень рентабельности по плодоконсервным заводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда, грн	
1	5,46	3842,9	37,6
2	5,53	3457,7	28,9
3	7,05	3066,4	32,1
4	7,29	3011,9	32,1
5	7,40	3013,3	31,9
6	7,10	3164,3	33,4
7	6,25	3289,1	31,3
8	8,64	4320,3	39,3
9	5,18	2829,3	24,8
10	1,81	2562,2	20
11	2,30	2402,6	25,5
12	5,53	3336,7	26,4
13	2,22	2227,8	20,3
14	3,54	2725,8	29,1
15	3,23	2710,8	27,7

Нелинейную зависимость принять $y = a\ln x + b$.

12 В таблице приведены данные об удельном весе пашни, лугов и пастбищ в сельскохозяйственных угодьях и уровне убыточности продукции животноводства по районам области за год.

Номер района	Фактор		Уровень убыточности продукции животноводства, %
	Удельный вес пашни в сельскохозяйственных угодьях, %	Удельный вес лугов и пастбищ, %	
1	80,0	20,0	20,0
2	87,2	12,8	37,5
3	90,8	9,2	43,4
4	94,7	11,3	45,6
5	81,4	18,6	23,4
6	79,2	10,8	25,0
7	71,3	28,7	17,2
8	86,2	13,8	33,3
9	71,4	28,6	15,0
10	77,7	22,9	18,7
11	75,4	14,0	24,8
12	77,9	13,0	34,5
13	87,2	12,8	33,1
14	68,1	25,0	19,2
15	86,2	13,8	31,8

Нелинейную зависимость принять $y = \frac{a}{x} + b$.

13 В таблице приведены данные об удельном весе в товарообороте потребительской кооперации продукции собственного производства, удельном весе переработанной продукции и уровне рентабельности предприятий области за год.

Номер предприятия	Фактор		Уровень рентабельности, %
	Удельный вес продукции собственного производства, %	Удельный вес переработанной продукции, %	
1	25,2	20,5	11,8
2	58,2	28,4	19,8
3	42,2	20,4	14,8
4	46,8	29,1	19,4
5	60,5	30,9	21,4
6	66,1	31,4	20,4
7	26,5	24,1	15,4
8	59,9	28,1	20,7
9	43,2	24,6	16,4
10	47,8	25,7	18,4
11	61,8	28,7	19,7
12	68,1	32,4	22,4
13	32,0	20,1	13,7
14	60,2	27,1	22,4
15	44,2	23,4	16,7

Нелинейную зависимость принять $y = ax^2 + b$.

14 Производительность труда, фондоотдача и уровень рентабельности по мясокомбинатам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда, грн	
1	1,25	5396	9,2
2	2,32	10583	14,7
3	1,71	8675	10,3
4	1,64	7392	10,0
5	1,38	3088	7,9
6	1,18	5138	9,1
7	1,44	5867	9,8
8	1,17	4154	6,4
9	1,72	13182	13,0
10	2,21	12351	13,8
11	1,64	13000	13,2
12	1,73	9519	11,4
13	1,17	4286	8,1
14	1,39	5000	9,0
15	1,07	7419	11,1

Нелинейную зависимость принять $y = a \ln x + b$.

15 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов (сбор овощей с 1 га и себестоимость 1 ц), ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Сбор овощей с 1га, ц	Себестоимость 1 ц, грн	
1	52,8	31,84	31,4
2	72,6	32,30	30,9
3	50,4	32,21	37,1
4	33,4	48,95	45,7
5	31,5	42,48	57,7
6	54,6	35,38	46,7
7	54,3	29,11	33,3
8	36,6	67,06	63,8
9	15,6	65,52	68,8
10	73,2	21,26	29,8
11	65,9	31,29	39,4
12	44,6	33,63	46,2
13	23,7	73,35	68,8
14	64,6	40,12	34,0
15	25,6	43,63	47,6

Нелинейную зависимость принять $y = ax^b$.

16 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов (сбор овощей с 1 га, ц, и затраты труда, человеко-часов на 1 ц), ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Сбор овощей с 1 га, ц	Затраты труда на 1 ц, человеко-часов	
1	93,2	2,3	8,8
2	65,9	26,8	39,4
3	44,6	22,8	26,2
4	18,7	56,6	78,8
5	64,6	16,4	34
6	25,6	26,5	47,6
7	47,2	26	43,7
8	48,2	12,4	23,6
9	64,1	10	19,9
10	30,3	41,7	50
11	28,4	47,9	63,1
12	47,8	32,4	44,2
13	101,3	20,2	11,2
14	31,4	39,6	52,8
15	67,6	18,4	20,2

Нелинейную зависимость принять $y = e^{ax} \cdot b$.

17 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов, ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Сбор овощей с 1га, ц	Себестоимость 1 ц, грн	
1	52,8	31,84	31,4
2	72,6	32,30	20,9
3	50,4	32,21	37,1
4	33,4	48,95	45,7
5	31,5	42,48	57,7
6	54,6	35,38	46,7
7	54,3	29,11	33,3
8	36,6	67,06	63,8
9	15,6	65,52	68,8
10	73,2	21,26	12,8
11	65,9	31,29	39,4
12	44,6	33,63	26,2
13	23,7	73,35	68,8
14	64,6	40,12	34
15	25,6	43,63	47,6

Нелинейную зависимость принять $y = a\sqrt{x} + b$.

18 Уровень убыточности выращивания овощей в сельскохозяйственных предприятиях и факторы, ее формирующие, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Себестоимость 1 ц, грн	Сбор овощей с 1га, ц	
1	21,26	73,2	10,8
2	31,29	65,9	29,4
3	33,63	44,6	26,2
4	73,35	23,7	68,8
5	40,12	64,6	31,1
6	43,63	25,6	47,6
7	32,2	47,2	43,7
8	49,85	38,2	43,6
9	39,02	64,1	25,9
10	41,7	30,3	50
11	49,53	28,4	43,1
12	38	47,8	34,2
13	17,14	101,3	8,2
14	44,17	41,4	52,8
15	31,4	67,6	20,2

Нелинейную зависимость принять $y = \frac{a}{x} + b$.

19 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов, ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Себестоимость 1 ц, грн	Затраты на 1 га посевов, грн	
1	31,84	1549	31,4
2	32,3	1694	40,9
3	32,21	1807	37,1
4	48,95	1615	45,7
5	42,48	1926	57,7
6	35,38	1542	46,7
7	29,11	1309	13,3
8	67,06	2093	63,8
9	63,52	1836	68,8
10	21,26	1649	12,8
11	31,29	1601	39,4
12	33,63	1560	26,2
13	73,35	2213	68,8
14	40,12	2028	34
15	65,52	2136	68,8

Нелинейную зависимость принять $y = a\ln x + b$.

20 Уровень рентабельности и показатели хозяйственной деятельности торговых предприятий характеризуются следующими данными за год:

Номер предприятия	Фактор		Уровень рентабельности, %
	Товарооборот на душу населения, грн	Относительный уровень издержек обращения, %	
1	27	17,4	3,62
2	29	17,35	3,8
3	21	17,33	2,77
4	21	21,2	2,01
5	33	16,96	4,33
6	28	17,01	4,01
7	23	19,77	2,12
8	28	18,4	3,73
9	30	15,35	3,92
10	22	18,34	2,87
11	22	22,2	2,11
12	34	16,06	4,39
13	31	16,01	4,11
14	22	18,7	2,13
15	29	17,4	3,87

Нелинейную зависимость принять $y = ax^b$.

21 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов, ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Себестоимость 1 ц, грн	Цена реализации 1 ц, грн	
1	31,84	21,83	31,4
2	32,3	19,09	40,9
3	32,21	20,26	37,1
4	48,95	20,57	45,7
5	42,48	17,96	57,7
6	35,38	15,32	46,7
7	29,11	29,19	13,3
8	67,06	11,26	63,8
9	65,52	10,47	68,8
10	21,26	29,67	12,8
11	31,29	18,95	39,4
12	33,63	24,81	26,2
13	73,35	12,92	68,8
14	40,12	26,49	34
15	43,63	22,83	47,6

Нелинейную зависимость принять $y = e^{ax} \cdot b$.

22 Убыточность выращивания овощей в сельскохозяйственных предприятиях и уровни факторов, ее формирующих, характеризуются следующими данными за год:

Номер района	Фактор		Уровень убыточности, %
	Себестоимость 1 ц, грн	Цена реализации 1 ц, грн	
1	21,26	31,67	18,8
2	31,29	18,95	39,4
3	33,63	24,81	36,2
4	73,35	14,92	68,8
5	40,12	26,49	34,7
6	43,63	22,83	47,6
7	32,2	18,13	43,7
8	49,85	20,14	43,6
9	39,02	23,47	39,9
10	41,7	20,85	50
11	49,53	21,17	43,1
12	38	21,2	44,2
13	22,14	28,87	21,2
14	44,17	20,83	52,8
15	31,4	30	20,2

Нелинейную зависимость принять $y = \frac{a}{x} + b$.

23 Производительность труда, фондоотдача и уровень рентабельности по хлебозаводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда, грн	
1	38,9	3742	10,7
2	33,3	2983	11,3
3	37,7	3000	12,2
4	31,1	2537	12,4
5	29,4	2421	10,9
6	37,2	3047	11,3
7	35,6	3002	11,1
8	34,1	2887	14,0
9	16,1	2177	6,8
10	22,8	2141	7,1
11	21,7	2005	8,9
12	26,8	1843	4,2
13	23,3	2031	7,4
14	24,5	2340	11,4
15	19,9	1933	4,8

Нелинейную зависимость принять $y = a \ln x + b$.

24 Известны следующие данные об убыточности производства говядины по КСП административных районов области за год:

Номер района	Фактор		Уровень убыточности, %
	Затраты на 1 голову крупного рогатого скота, грн	Затраты на 1 ц прироста, грн	
1	283	309,95	37,7
2	214	260	23,7
3	246	264,03	26,8
4	265	306,74	28,4
5	262	288,72	43,2
6	213	287,5	38
7	243	267,34	33,9
8	360	444,84	79,1
9	248	287,77	29,8
10	301	456,84	62
11	210	196,8	19,6
12	305	413,8	53,8
13	234	271,71	27,4
14	279	351,94	53,6
15	361	499,39	62,1

Нелинейную зависимость принять $y = a\sqrt{x} + b$.

25 В таблице приведены данные об относительном уровне издержек обращения, производительности труда и уровне рентабельности по магазинам промышленных товаров за год.

Номер магазина	Фактор		Уровень рентабельности, %
	Относительный уровень издержек обращения, %	Производительность труда, грн	
1	7,89	17646	8,9
2	14,41	10177	4,3
3	6,01	19343	10,2
4	9,17	14789	4,9
5	6,78	18172	8,3
6	8,91	17477	7,8
7	6,17	22110	13,1
8	10,11	14331	4,9
9	5,98	24111	13,3
10	6,10	19393	10,7
11	5,90	25445	13,7
12	8,13	17010	5,6
13	9,01	13137	4,7
14	6,00	21100	11,1
15	6,13	19378	10,8

Нелинейную зависимость принять $y = a\ln x + b$.

26 Производительность труда, фондоотдача и уровень рентабельности по плодоовощным консервным заводам области за год характеризуются следующими данными:

Номер завода	Фактор		Уровень рентабельности, %
	Производительность труда, грн	Фондо-отдача, грн	
1	8540	1,24	38,34
2	2911	0,63	44,69
3	6630	1,18	39,4
4	8492	1,12	38,93
5	2901	0,44	46,96
6	5410	1,19	39,48
7	1920	0,48	46,07
8	2569	0,65	43,5
9	3520	0,26	56,11
10	2340	0,75	42,79
11	6921	1,03	40,15
12	7671	0,89	40,44
13	1586	0,16	69,76
14	3223	0,67	42,99
15	7224	0,90	40,69

Нелинейную зависимость принять $y = \dfrac{a}{x} + b$.

27 Известны следующие данные об убыточности производства говядины по КСП административных районов области за год:

Номер района	Фактор		Уровень убыточности, %
	Среднесуточный прирост, грн	Себестоимость 1 ц, грн	
1	249	138,99	37,7
2	231	105,86	29,7
3	245	114,19	26,8
4	242	131,73	28,4
5	250	139,86	43,2
6	190	141,52	48
7	283	118,9	33,9
8	273	163,26	29,1
9	290	143,7	29,8
10	150	221,88	66
11	294	102,4	19,6
12	196	149,06	48,8
13	241	135,5	27,4
14	214	178,17	53,6
15	188	229,36	62,1

Нелинейную зависимость принять $y = a\sqrt{x} + b$.

28 В таблице приведены данные о затратах на 1 грн. товарной продукции, удельном весе простоев оборудования и уровне рентабельности по молокозаводам области за год.

Номер завода	Фактор		Уровень рентабельности, %
	Затраты на 1 грн. товарной продукции, грн	Удельный вес простоев оборудования, %	
1	0,59	8,1	15,45
2	2,25	11,8	20,33
3	0,36	7,4	14,67
4	1,37	9,4	16,05
5	5,44	17,8	37,39
6	2,02	12,1	22,19
7	1,74	10,2	17,01
8	3,10	14,1	26,24
9	1,73	10,1	16,74
10	4,59	16,7	33,83
11	6,76	19,4	43,58
12	1,84	10,4	17,24
13	4,73	16,2	30,62
14	4,58	16	30,1
15	3,66	15,1	28,81

Нелинейную зависимость принять $y = ax^b$.

29 Производительность труда, фондоотдача и уровень рентабельности по плодоконсервным заводам области за год характеризуются следующими данными.

Номер завода	Фактор		Уровень рентабельности, %
	Фондоотдача, грн	Производительность труда, грн	
1	1,08	7343	20,1
2	1,05	3991	12,9
3	0,99	5760	18,0
4	1,02	3000	11,7
5	0,98	5241	17,9
6	1,04	4500	16,8
7	1,03	4300	15,6
8	1,10	3210	14,3
9	1,03	6743	18,1
10	0,89	5234	17,8
11	0,78	2500	13,0
12	0,99	3930	14,2
13	1,43	14333	24,2
14	1,03	6980	20,0
15	1,05	6740	19,3

Нелинейную зависимость принять $y = a\ln x + b$.

30 Известны следующие данные об убыточности производства говядины по КСП административных районов области за год:

Номер района	Фактор		Уровень убыточности, %
	Затраты на 1 ц прироста, человеко-часов	Среднесуточный прирост, грн	
1	76,8	249	37,7
2	76	271	23,7
3	74,6	245	26,8
4	79	242	28,4
5	76,6	250	43,2
6	93,4	190	48
7	71,8	283	33,9
8	93	223	49,1
9	66,6	290	29,8
10	112	150	69
11	66,9	304	19,6
12	94,6	196	53,8
13	70	241	27,4
14	92,2	214	53,6
15	89,2	188	62,1

Нелинейную зависимость принять $y = ax^b$.

4 САМОСТОЯТЕЛЬНАЯ РАБОТА

Тема: «СИСТЕМА ОДНОВРЕМЕННЫХ УРАВНЕНИЙ»

Задание

Статическая модель Кейнса для описания народного хозяйства страны в наиболее простом варианте имеет следующий вид:

$$\begin{cases} C = a + b \cdot y + \varepsilon, \\ y = C + I, \end{cases}$$

где C – личное потребление в постоянных ценах,

y – национальный доход в постоянных ценах,

I – инвестиции в постоянных ценах,

ε – случайная составляющая.

По данным, приведенным в таблице, постройте систему одновременных уравнений. Определите параметры уравнений с помощью непрямого метода наименьших квадратов (НМНК). Проанализируйте получившуюся модель.

Пример выполнения

Дано:

Уровень произв. и дохода y, млрд. долл.	Потребление C, млрд. долл.	Инвестиции I, млрд. долл.
370	365	5
415	397	18
430	409	21
456	430	26
486	450	36
490	455	35
505	467	38
520	479	41
546	500	46
567	516	51

По смыслу задачи, y и C – эндогенные переменные, а I – экзогенная переменная. Определим идентифицируемость 1-го уравнения:

$n_1 = 2$ – число эндогенных переменных (y и C), входящих в 1-е уравнение,

$m = 1$ – общее количество экзогенных переменных в системе,

$m_1 = 0$ – количество экзогенных переменных, входящих в 1-е уравнение.

Проверяем идентифицируемость с помощью счетного правила:

$$n_s - 1 \le m - m_s.$$

Получим $2-1=1-0$, $1=1$. Первое уравнение точно идентифицируемо. Следовательно, система точно идентифицируема, и для определения параметров уравнений можно использовать НМНК.

Для записи приведенной формы модели преобразуем уравнения структурной формы модели.

Подставим значение y из 2-го уравнения в 1-е. Получим:

$$C = \frac{a}{1-b} + \frac{b}{1-b} \cdot I + \frac{\varepsilon}{1-b}$$

Теперь подставим значение C из 1-го уравнения во 2-е. Получим:

$$y = \frac{a}{1-b} + \frac{1}{1-b} \cdot I + \frac{\varepsilon}{1-b}.$$

Обозначим эндогенные переменные C и y через $Y1$и $Y2$, а экзогенную переменную I через $X1$. Тогда систему одновременных уравнений в приведенной форме можно записать в виде

$$Y1 = b_{10} + b_{11} X1 + \varepsilon_1,$$
$$Y2 = b_{20} + b_{21} X1 + \varepsilon_2.$$

Здесь $b_{10} = b_{20} = \dfrac{a}{1-b}$, $b_{11} = \dfrac{b}{1-b}$, $b_{21} = \dfrac{1}{1-b}$.

Параметры каждого из уравнений этой системы находим с помощью МНК.

Для первого уравнения получим следующие результаты:

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	**0,993323**
R-квадрат	0,986692
Нормированный R-квадрат	**0,985028**
Стандартная ошибка	5,771513
Наблюдения	10

Дисперсионный анализ

	df	*SS*	*MS*	**F**	Значимость F
Регрессия	1	19757,12	19757,12	593,1223	**8,62366E-09**
Остаток	8	266,4829	33,31036		
Итого	9	20023,6			

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*
Y-пересечение	**341,1915**	4,704804	72,51982	**1,46E-12**
X1	**3,331498**	0,136794	24,3541	**8,62E-09**

Анализ полученных результатов проводится таким же образом, как и при построении линейной однофакторной модели.

Множественный R = 0,993323, это значит, что между $X1$ и $Y1$ существует тесная линейная связь. Коэффициент детерминации $R^2 = 0{,}985$.

Коэффициенты для 1-го уравнения $b_{10} = 341{,}19$, $b_{11} = 3{,}33$. При проверке их статистической значимости по критерию Стьюдента, получим следующие значения уровня значимости: $\alpha_{b10} = 1{,}46 \cdot 10^{-12}$, $\alpha_{b11} = 8{,}62 \cdot 10^{-9}$. Т.к. эти значения меньше 0,05, то с вероятностью 0,95 можно утверждать, что оба коэффициента статистически значимы и могут быть включены в модель.

Таким образом, первое уравнение приведенной формы модели примет вид: $Y1 = 341{,}19 + 3{,}33X1 + \varepsilon_1$. При проверке этого уравнения на адекватность по критерию Фишера получим $\alpha_{F_{НАБЛ}} = 8{,}62 \cdot 10^{-9}$. Т.к. $\alpha_{F_{НАБЛ}} < 0{,}05$, то с вероятностью 0,95 можно утверждать, что уравнение адекватно исходным данным.

Выполним аналогичный расчет для второго уравнения приведенной формы модели.

ВЫВОД ИТОГОВ

Регрессионная статистика	
Множественный R	**0,9960342**
R-квадрат	0,9920842
Нормированный R-квадрат	**0,9910947**
Стандартная ошибка	5,771513
Наблюдения	10

Дисперсионный анализ

	df	*SS*	*MS*	*F*	***Значимость F***
Регрессия	1	33398,02	33398,02	1002,631	**1,07703E-09**
Остаток	8	266,4829	33,31036		
Итого	9	33664,5			

	Коэффициенты	*Стандартная ошибка*	*t-статистика*	***P-Значение***
Y-пересечение	**341,19151**	4,704804	72,51982	**1,46E-12**
X1	**4,3314982**	0,136794	31,66436	**1,08E-09**

Множественный R = 0,9960342, это значит, что между $X1$ и $Y1$ существует тесная линейная связь. Коэффициент детерминации $R^2 = 0{,}991$.

Коэффициенты для 2-го уравнения $b_{20} = 341{,}19$, $b_{21} = 4{,}33$. При проверке их статистической значимости по критерию Стьюдента получим следующие значения уровня значимости: $\alpha_{b20} = 1{,}46 \cdot 10^{-12}$, $\alpha_{b21} = 1{,}08 \cdot 10^{-9}$. Так как эти значения меньше 0,05, то с вероятностью 0,95 можно утверждать, что оба коэффициента статистически значимы и могут быть включены в модель.

Получили второе уравнение приведенной формы модели $Y2 = 341{,}19 + 4{,}33X1 + \varepsilon_2$. При проверке этого уравнения на адекватность по критерию Фишера получим $\alpha_{F_{НАБЛ}} = 1{,}077 \cdot 10^{-9}$. Т.к. $\alpha_{F_{НАБЛ}} < 0{,}05$, то с вероятностью 0,95 можно утверждать, что уравнение адекватно исходным данным.

Запишем приведенную форму модели:

$$Y1 = 341{,}19 + 3{,}33x1 + \varepsilon_1,$$

$$Y2 = 341{,}19 + 4{,}33x1 + \varepsilon_2.$$

Перейдем от приведенной формы модели к структурной.

Как было записано ранее: $b_{10} = b_{20} = \frac{a}{1-b}$, $b_{11} = \frac{b}{1-b}$, $b_{21} = \frac{1}{1-b}$.

Подставим полученные значения b_{10}, b_{20}, b_{11} и b_{21}:

$$341{,}19 = \frac{a}{1-b},$$

$$4{,}33 = \frac{1}{1-b},$$

$$3{,}33 = \frac{b}{1-b}$$

Отсюда:

$$a = 78{,}81,$$

$$b = 0{,}769.$$

Запишем структурную форму модели:

$$\begin{cases} C = 78{,}81 + 0{,}769 \cdot y + \varepsilon, \\ y = C + I. \end{cases}$$

В ней структурный коэффициент b характеризует предельную склонность к потреблению. Здесь $b = 0{,}769$, значит, из каждой дополнительной тысячи дохода на потребление расходуется 769 д.е., а 231 д.е. инвестируется.

Инвестиционный мультипликатор потребления – это коэффициент $b_{11} = \frac{b}{1-b} = M_c = 3{,}33$. Эта величина означает, что дополнительные вложения в размере 1 тыс. д.е. приведут при прочих равных условиях к дополнительному увеличению потребления на 3,33 тыс. д.е.

Инвестиционный мультипликатор национального дохода – это коэффициент $b_{21} = \frac{1}{1-b} = M_y = 4{,}33$. Эта величина означает, что дополнительные инвестиции в размере 1 тыс. д.е. приведут при прочих равных условиях к дополнительному доходу на 4,33 тыс. д.е.

Варианты

Вариант 1

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
370	362	8
415	396	18
430	409	21
456	430	26
486	450	33
490	455	34
505	467	37
520	479	40
546	500	46
567	516	50

Вариант 2

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
370	365	5
415	380	35
430	410	20
456	428	28
486	435	51
490	455	35
505	465	40
520	490	30
546	500	46
567	523	44

Вариант 3

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
381	376	5
426	391	35
441	421	20
467	439	28
497	446	51
501	466	35
516	476	40
531	501	30
557	511	46
578	534	44

Вариант 4

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
381	375	6
426	390	36
441	420	21
467	438	29
497	445	52
501	465	36
516	475	41
531	500	31
557	510	47
578	533	45

Вариант 5

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
370	360	10
415	396	19
430	407	23
456	428	28
486	451	35
490	454	36
505	466	39
520	478	42
546	498	48
567	514	53

Вариант 6

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
370	360	10
415	395	20
430	405	25
456	428	28
486	450	36
490	454	36
505	465	40
520	478	42
546	497	49
567	514	53

Вариант 7

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
370	360	10
415	395	20
430	405	25
455	428	27
485	450	35
490	454	36
505	465	40
520	478	42
545	497	48
570	514	56

Вариант 8

Уровень произв. и дохода, млрд. дол. (y)	Потребление, млрд. дол. (C)	Инвестиции, млрд. дол. (I)
384	372	11
429	407	22
444	417	26
469	440	29
499	462	38
504	466	39
519	477	42
534	490	45
559	509	50
584	526	59

Вариант 9

Уровень произв. и дохода, млрд. дол. (у)	Потребление, млрд. дол. (С)	Инвестиции, млрд. дол. (I)
391	375	16
436	390	46
451	420	31
477	438	39
507	445	62
511	465	46
526	475	51
541	500	41
567	510	57
588	533	55

Вариант 10

Уровень произв. и дохода, млрд. дол. (у)	Потребление, млрд. дол. (С)	Инвестиции, млрд. дол. (I)
306	311	0
351	346	5
366	358	8
392	378	14
422	401	20
426	405	21
441	418	25
456	429	28
482	449	33
503	465	38

СПИСОК ЛИТЕРАТУРЫ

1 Боровиков В.П. STATISTICA/ В.П.Боровиков, И.П.Боровиков. – М., 1997. – 592с.

2 Доугерти К. Введение в эконометрику. – М., 2001. – 402с.

3 Лук'яненко І. Економетрика: Практикум / І.Лук'яненко, Л.Краснікова. – Київ: Знання, 1998. – 217с.

4 Лук'яненко І. Економетрика/ І.Лук'яненко, Л.Краснікова. – Київ: Знання, 1998. – 493с.

5 Магнус Я.Р. Эконометрика. Начальный курс: Учебник/ Я.Р.Магнус, П.К.Катышев, А.А.Пересецкий. – 4-е изд. – М.: Дело, 2000. – 400с.

6 Эконометрика: Учебник/ Под ред. И.И.Елисеевой. – М.: Финансы и статистика, 2002. – 344с.

7 Фишер Ф. Проблемы идентификации в эконометрии. – М.: Статистика, 1978. – 245 с.

Людмила Владимировна Васильева,
Елена Анатольевна Клеваник

ЭКОНОМЕТРИКА: НАЧАЛЬНЫЙ КУРС

ПОСТРОЕНИЕ ЛИНЕЙНЫХ И НЕЛИНЕЙНЫХ МОДЕЛЕЙ. СИСТЕМЫ ОДНОВРЕМЕННЫХ УРАВНЕНИЙ

Учебное пособие
для студентов высших учебных заведений

Редактор Дудченко Елена Александровна

Подп. в печать Формат 60x84 1/16
Ризограф. печать. Усл.печ.л. Уч.-изд. л.
Тираж экз. Зак. №

ДГМА. 84313, г. Краматорск, ул. Шкадинова, 72