Построение модели парной регрессии

Оценка параметров модели по методу наименьших квадратов.

При эмпирическом (экспериментальном) изучении функциональной зависимости одной величины у от другой величины х производят ряд измерений величины у при различных значениях х. Результаты могут быть представлены в виде таблицы:

x	x_1	x_2	…	x_κ	…	x_n
y	y_1	y_2	…	y_k	…	y_n

(1)

Метод, основанный на требовании минимизации суммы квадратов отклонений, называется методом наименьших квадратов (1 МНК).

С его помощью изображают статистическую функциональную зависимость в виде аналитической зависимости и отыскиваются такие оценки параметров уравнения регрессии, которые сводят к минимуму выбранную меру разброса. В результате происходит выравнивание эмпирических значений в одну линию регрессии. При этом, для однозначного нахождения в качестве меры разброса используют один из показателей рассеяния случайностей величины – дисперсию.

МНК применяется для решения задач, связанных с обработкой результатов испытаний. Этот метод не решает вопроса о выборе вида аналитической функции, а дает возможность, в эмпирически подобранной функции, определить наиболее вероятные значения для параметров аппроксимирующей функции – в этом и заключается основная задача метода наименьших квадратов.

Пусть после экономического анализа с учетом характера скопления точек $M_i(x_i, y_i),\ i=\overline{1,n}$, взятых из таблицы (1), на плоскости в системе координат ХОУ получена диаграмма рассеяния, по которой и подбирается эмпирическая
(аппроксимическая) функция.

1. Предположим, что диаграмма рассеяния такова, что зависимость между переменными x и y "наилучшим" образом может быть представлена в виде прямой линии $\widehat{y} = ax + b$ Это означает, что отклонения фактических значений функции от "подобранной" прямой $Z_i = y_i - \bar{y}_i$ должны быть минимальными, т.е. прямая подбирается так, чтобы сумма квадратов отклонений была наименьшей.

$$z_1^2 + z_2^2 + \ldots + z_i^2 + \ldots z_n^2 \to \min \qquad (2)$$

Пусть

$$\widehat{y} = ax + b \qquad (3)$$

есть уравнение "подобранной" прямой. Тогда, согласно (2), должно выполняться равенство

$$z = \sum_{i=1}^{n} (y_i - ax_i - b)^2 = \min \qquad (4)$$

Требуется определить параметры a и b так, чтобы Z достигла минимума. Известно, что необходимое условие существования минимума состоит в том, чтобы

$$\frac{\partial z}{\partial a} = 0, \frac{\partial z}{\partial b} = 0, \frac{\partial z}{\partial c} = 0 \qquad (5)$$

После дифференцирования (4) и упрощений, получим систему уравнений

$$\begin{cases} a \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a \sum_{i=1}^{n} x_i + bn = \sum_{i=1}^{n} y_i \end{cases} \qquad (6)$$

Для решения системы уравнений (6) относительно a и b, составляют расчетную таблицу, которая может иметь следующий вид:

i	x_i	y_i	x_i^2	$x_i y_i$
1 2 . . n				
Σ	$\sum\limits_{i=1}^{n} x_i$	$\sum\limits_{i=1}^{n} y_i$	$\sum\limits_{i=1}^{n} x_i^2$	$\sum\limits_{i=1}^{n} x_i y_i$

Полученные в последней строке таблицы суммы подставляем в систему (6) и решаем ее любым известным способом.

Система (6) называется системой нормальных уравнений в случае выбора эмпирической функции в виде линейной зависимости. Определив параметры a и b, подставляют их в уравнение

$$\bar{y} = ax + b$$

1. Если диаграмма рассеяния такова, что эмпирическую зависимость целесообразно выбрать в виде квадратичной функции

$$\hat{y} = ax^2 + bx + c,$$

тогда, согласно МНК, будем иметь

$$z = \sum_{i=1}^{n} (y_i - ax_i^2 - bx_i - c)^2 = \min \qquad (7)$$

В этой функции искомыми величинами являются параметры a, b и c, поэтому что, согласно необходимых условий экстремума функции, нужно, чтобы

$$\frac{\partial z}{\partial a} = 0, \frac{\partial z}{\partial b} = 0, \frac{\partial z}{\partial c} = 0$$

Дифференцируя (7), после упрощений, получим

$$\begin{cases} a\sum\limits_{i=1}^{n} x_i^4 + b\sum\limits_{i=1}^{n} x_i^{\ 3} + c\sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} x_i^2 y_i \\ a\sum\limits_{i=1}^{n} x_i^{\ 3} + b\sum\limits_{i=1}^{n} x_i^2 + c\sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} x_i y_i \\ a\sum\limits_{i=1}^{n} x_i^{\ 2} + b\sum\limits_{i=1}^{n} x_i + cn = \sum\limits_{i=1}^{n} y_i \end{cases} \qquad (8)$$

Это система нормальных уравнений в случае выбора квадратичной функции в качестве эмпирической функции.

Составим расчетную таблицу:

i	x_i	y_i	x_i^2	x_i^3	x_i^4	$x_i y_i$	$x^2{}_i y_i$
1 2 . . n							
$\sum$	$\sum\limits_{i=1}^{n} x_i$	$\sum\limits_{i=1}^{n} y_i$	$\sum\limits_{i=1}^{n} x^2{}_i$	$\sum\limits_{i=1}^{n} x^3{}_i$	$\sum\limits_{i=1}^{n} x^4{}_i$	$\sum\limits_{i=1}^{n} x_i y_i$	$\sum\limits_{i=1}^{n} x^2{}_i y_i$

Полученные в последней строке таблицы суммы подставляем в систему (8) и решаем ее любым известным способом. Определив параметры a , b и c, подставляем их в уравнение

$$\hat{y} = ax^2 + bx + c\,.$$

3. Возможен случай, если эмпирическая квадратичная функция задана в виде:

$$\hat{y} = a_0 + a_1 x + a_2 x^2\,,$$

то систему нормальных уравнений можно записать в таком виде:

$$\begin{cases} a_0 n + a_1 \sum\limits_{i=1}^{n} x_i + a_2 \sum\limits_{i=1}^{n} x_i^2 = \sum\limits_{i=1}^{n} y_i \\ a_0 \sum\limits_{i=1}^{n} x_i + a_1 \sum\limits_{i=1}^{n} x_i^2 + a_2 \sum\limits_{i=1}^{n} x_i^3 = \sum\limits_{i=1}^{n} x_i y_i \\ a_0 \sum\limits_{i=1}^{n} x_i^{\ 2} + a_1 \sum\limits_{i=1}^{n} x_i^3 + a_2 \sum\limits_{i=1}^{n} x_i^4 = \sum\limits_{i=1}^{n} x_i^{\ 2} y_i \end{cases}$$

когда диаграмма рассеяния такова, что эмпирическую зависимость целесообразно выбрать в виде гиперболической зависимости:

$$\hat{y} = a + \frac{b}{x}, \tag{9}$$

Согласно идее МНК нужно, чтобы

$$z = \sum_{i=1}^{n} (y_i - a - \frac{b}{x_i})^2 = \min \tag{10}$$

Нужно подобрать параметры a и b так, чтобы выполнялось условие (10), а для этого нужно, чтобы

$$\frac{\partial z}{\partial a} = 0 \text{ и } \quad \frac{\partial z}{\partial b} = 0 \tag{11}$$

Продифференцировав (10) и упростив, получим

$$\begin{cases} an + b\sum\limits_{i=1}^{n} \frac{1}{x_i} = \sum\limits_{i=1}^{n} y_i \\ a\sum\limits_{i=1}^{n} \frac{1}{x_i} + b\sum\limits_{i=1}^{n} \frac{1}{x_i^2} = \sum\limits_{i=1}^{n} \frac{y_i}{x_i} \end{cases} \tag{12}$$

Это система нормальных уравнений в случае выбора гиперболической функции в качестве эмпирической функции.

Составим расчетную таблицу:

i	x_i	y_i	x_i^2	$\frac{1}{x_i}$	$\frac{1}{x^2{}_i}$	$\frac{y_i}{x_i}$
1 2 . n						
$\sum$	$\sum_{i=1}^{n} x_i$	$\sum_{i=1}^{n} y_i$	$\sum_{i=1}^{n} x^2{}_i$	$\sum_{i=1}^{n} \frac{1}{x_i}$	$\sum_{i=1}^{n} \frac{1}{x^2{}_i}$	$\sum_{i=1}^{n} \frac{y_i}{x_i}$

Полученные в последней строке таблицы суммы подставляем в систему (12) и решаем ее любым известным способом. Определив параметры a и b, подставляем их в уравнение (9), найдем эмпирическую функцию.

4. Может встретиться случай, когда опытные точки $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$ образуют некоторую линию похожую на график показательной функции и эта линия наилучшим образом отражает зависимость между x и y, тогда искомое аппроксимирующее уравнение записывают в виде:

$$\widehat{y} = a \cdot b^x \tag{13}$$

Программируем обе части (13), получим

$$\log \widehat{y} = \log a + x \log b \tag{14}$$

Введем обозначения: log a = A, log b = B,

Тогда получим уравнение

$$\log \widehat{y} = A + Bx \tag{15}$$

Отсюда следует, что функция $\bar{y}$, представленная на графике в системе координат XOY, где ось ординат разделена по логарифмической шкале, а ось абсцисс – по нормальной шкале, дает прямую с угловым коэффициентом B.

Искомыми параметрами в (15) являются A и B. По МНК нужно, чтобы

$$z = \sum_{i=1}^{n} (\log y_i - A - Bx)^2 = \min \quad (16)$$

Требуется подобрать параметры А и В так, чтобы выполнялось условие (16), а для этого нужно, чтобы

$$\frac{\partial z}{\partial a} = 0 \text{ и } \quad \frac{\partial z}{\partial b} = 0 \quad (17)$$

Продифференцировав (16) и упростив полученное выражение, получим:

$$\begin{cases} An + B\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \log y_i \\ A\sum_{i=1}^{n} x_i + B\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i \log y_i \end{cases} \quad (18)$$

Эта система нормальных уравнений в случае выбора функции (13) в качестве эмпирической функции.

Составим расчетную таблицу:

i	x_i	y_i	x_i^2	$\log y_i$	$x_i \log y_i$
1 2 . . n					
$\sum$	$\sum_{i=1}^{n} x_i$		$\sum_{i=1}^{n} x^2{}_i$	$\sum_{i=1}^{n} \log y_i$	$\sum_{i=1}^{n} x_i \log y_i$

Полученные в последней строке таблицы суммы подставляем в систему (18) и решаем ее любым известным способом. Определив параметры А и В, подставляем их в уравнение (15). Затем по параметрам А и В пересчитываем интересующие нас параметры a и b.

Следует иметь в виду, что параметры А и В обращают в минимум сумму квадратов отклонений значений преобразованных величин, а не сумму квадратов отклонений измеренных величин у от соответствующих расчетных и могут служить только в качестве первого приближения к наилучшим оценкам отыскиваемых параметров. Поэтому они должны быть уточнены.

Этим способом можно исследовать и такие функции: $\widehat{y} = ax^b$; $\bar{y} = ae^{bx}$; в функции $y = \dfrac{1}{ax+b}$ применяют преобразование $\widehat{y} = \dfrac{1}{y} = ax + b$ (или $x = \dfrac{1}{x}$, $y = \dfrac{1}{y}$); в функции $\widehat{y} = \dfrac{1}{a + be^{-x}}$ применяют преобразование $y = \dfrac{1}{y}$, $x = e^{-x}$.

Эти преобразования необходимы, так как искомые параметры входят в эмпирическую функцию нелинейно.

3.2 МНК в матричной форме.

Включение в регрессию нескольких переменных (или факторов) усложняет расчетные формулы МНК. Поэтому при использовании множественной регрессии с m переменными будем пользоваться аппаратом теории матриц.

Уравнение линейной многофакторной регрессии запишем в виде:

$$y = a_0x_0 + a_1x_1 + a_2x_2 + \ldots + a_mx_m = \sum_{k=0}^{m} a_k x^k , \tag{19}$$

в которой фактор x_0 является фиктивным фактором и включен в регрессию для симметрии. Обычно $x_0 = 1$.

Эта линейная регрессия в матричной форме запишется так:

$$Y = XA + e , \tag{20}$$

Здесь результаты наблюдений $y_1, y_2, \ldots, y_n$ записываем в форме матрицы – столбца (вектор-столбца) размерности (n*1):

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix},$$

где n – количество наблюдений. Значения факторов (или переменных) записываем в виде матрицы X размерности (n*(m*1)):

$$X=\begin{pmatrix} x_{10} & x_{11} & x_{12}\ldots & x_{1m} \\ x_{20} & x_{21} & x_{22}\ldots & x_{2m} \\ . & . & . & . \\ x_{n0} & x_{n1} & x_{n2}\ldots & x_{nm} \end{pmatrix}=\begin{pmatrix} 1 & x_{11} & x_{12}\ldots & x_{1m} \\ 1 & x_{21} & x_{22}\ldots & x_{2m} \\ . & . & . & . \\ 1 & x_{n1} & x_{n2}\ldots & x_{nm} \end{pmatrix},$$

где m – число независимых переменных, а всем x_{10} , x_{20} ,…, x_{n0} присвоим значения, равные 1.

Оценки параметров регрессии образуют матрицу-столбец (вектор-столбец) A размерности ((m+1)*1):

$$A=\begin{pmatrix} a_0 \\ a_1 \\ \ldots \\ a_m \end{pmatrix}.$$

Остатки (или ошибки, или отклонения) функции регрессии запишем в виде матрицы-столбца (вектор-столбца) e размерности (n*1):

$$e=\begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix}$$

Теперь (20) можно записать в развернутом виде следующим образом:

$$\begin{pmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{pmatrix}=\begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1m} \\ 1 & x_{21} & x_{22} & \cdots & x_{2m} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nm} \end{pmatrix}\cdot\begin{pmatrix} a_0 \\ a_1 \\ \ldots \\ a_m \end{pmatrix}+\begin{pmatrix} e_1 \\ e_2 \\ \ldots \\ e_n \end{pmatrix} \qquad (21)$$

Для оценки неизвестных параметров a_0 , a_1 , a_2 ,…, a_m функции регрессии применяем обычный метод наименьших квадратов, но применению его в матричной форме должны предшествовать некоторые предпосылки. Они касаются прежде всего случайной переменной e и имеют общий характер. Они не связаны ни с объемом выборки, ни с числом включенных в регрессию переменных. Если эти предпосылки выполняются, то оценки параметров вектора А должны быть несмещенными, эффективными и состоятельными.

Пусть для нахождения оценок вектора А, которые будем обозначать через $\widehat{A}$, осуществлена выборка, удовлетворяющая предпосылкам МНК. Тогда, согласно критерия этого метода, сумма квадратов остатков должна быть минимальной, что в матричной форме выглядит так:

$$\sum_{i=1}^{n} e_j^2 = e' \cdot e = \min \qquad (22)$$

Так как

$$e = Y - X\widehat{A}\text{, то}$$
$$\left(Y - X\widehat{A}\right)'\left(Y - X\widehat{A}\right) = min \qquad (23)$$

Для нахождения значения $\widehat{A}$, которое минимизирует сумму квадратов остатков, продифференцируем (23) по $\widehat{A}$, приравняем к нулю, получим

$$X'X\widehat{A} = X'Y \qquad (24)$$

Это система нормальных уравнений в матричной форме. Если матрица $X'X$ обратима, т.е. для нее существует матрица $\left(X'X\right)^{-1}$, то получим в качестве решения системы (24) вектор-столбец искомых оценок параметров регрессии:

$$\widehat{A} = \left(X'X\right)^{-1} X'Y \qquad (25)$$

Выражение (25) является основным результатом процедуры оценивания параметров функции регрессии МНК.

При этом, матрица

$$X'Y = \begin{pmatrix} \sum y_i \\ \sum x_{i1} y_i \\ \dots \\ \sum x_{im} y_i \end{pmatrix}.$$

называется матрицей моментов, а матрица $\left(X'X\right)^{-1}$, где

$$X'X = \begin{pmatrix} n & \sum x_{i_1} & \dots & \sum x_{i_m} \\ \sum x_{i_1} & \sum x_{i_1}^2 & \dots & \sum x_{i_1} x_{i_m} \\ \sum x_{i_m} & \sum x_{i_m} x_{i_1} & \dots & \sum x_{i_m}^2 \end{pmatrix}$$

называется матрицей ошибок.

После вычисления выражения, стоящего в правой части равенства (25) и сравнивания с матрицей $\widehat{A}$, найдем оценки параметров функции регрессии, т.е. найдем $\widehat{a}_0, \widehat{a}_1, \widehat{a}_2, \dots, \widehat{a}_n$, которые и подставляют в уравнение (19).

3.3 Порядок выполнения работы.

Порядок выполнения работы рассмотрим на следующем примере.

Пример 1.

Методом наименьших квадратов найти значения параметров эмпирической функции, если опытные данные о значениях X и Y представлены в таблице:

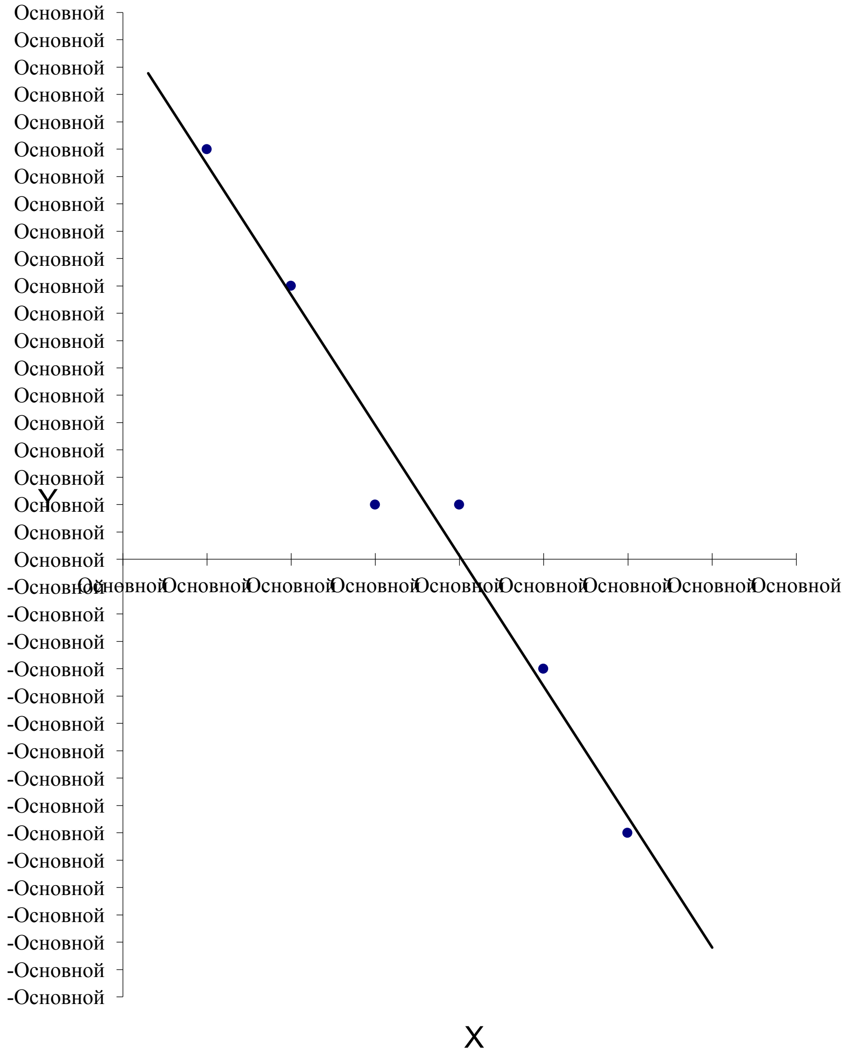
X	1	2	3	4	5	6
Y	15	10	2	2	-4	-10

(26)

Решение

По выборке наблюдений (26) построим в системе координат XOY диаграмму рассеяния, т.е. на плоскости в XOY нанесем точки:

(1,15) (2,10) (3,2) (4,2) (5,-4) (6,-10)



Анализ опытных данных показывает, что в качестве эмпирической (подобранной) функции можно использовать линейную функцию

$$\hat{y} = a \cdot x + b \qquad (27)$$

В выражении (27) необходимо найти параметры a и b, для чего применяем МНК. Тогда для нахождения параметров будем иметь систему нормальных уравнений:

$$\begin{cases} a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \\ a\sum_{i=1}^{n} x_i + b \cdot n = \sum_{i=1}^{n} y_i \end{cases}$$

Для удобства вычислений составим следующую расчетную таблицу (из условия задачи известно, что n = 6).

i	x_i	y_i	x_i^2	$x_i \cdot y_i$
1	1	15	1	15
2	2	10	4	20
3	3	2	9	6
4	4	2	16	8
5	5	-4	25	-20
6	6	-10	36	-60
$\sum$	$\sum_{i=1}^{6} x_i = 21$	$\sum_{i=1}^{6} y_i = 15$	$\sum_{i=1}^{6} x_i^2 = 91$	$\sum_{i=1}^{6} x_i \cdot y_i = -31$

Подставим данные последней строки таблицы в нормальную систему уравнений (6):

$$\begin{cases} 91 \cdot a + 21 \cdot b = -31 \\ 21 \cdot a + 6 \cdot b = 15 \end{cases}$$

Решая эту систему любым известным способом, получим:

a = -4,76; b = 19,2

Подставляя эти значения параметров в (27), получим эмпирическую функцию:

$$\hat{y} = -4{,}76 \cdot x + 19{,}2$$

Если эмпирическая функция будет записана в виде $\hat{y} = a + b \cdot x$, то система нормальных уравнений имеет вид:

$$\begin{cases} a \cdot n + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a \cdot \sum_{i=1}^{n} x_i + b \cdot \sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases} \qquad (28)$$

Если в условии задачи эмпирическая функция задана, то диаграмму рассеяния строить не нужно.

Рассмотрим пример применения МНК в матричной форме.

Пример 2.

Для оценки параметров функции $y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + e$ пусть имеем совокупность наблюдений, которая приведена в таблице (29):

№	y	x_1	x_2
1	10	2	1
2	12	2	2
3	17	8	10
4	13	2	4
5	15	6	8
6	10	3	4
7	14	5	7
8	12	3	3
9	16	9	10
10	18	10	11
$\sum$	137	50	60

(29)

Решение

Найдем оценки параметров a_0 , a_1 , a_2 , применяя МНК в матричной форме. По данным таблицы (29) составим систему нормальных уравнений (24), для чего запишем следующие матрицы:

$$X = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 2 \\ 1 & 8 & 10 \\ 1 & 2 & 4 \\ \dots & \dots & \dots \\ 1 & 10 & 11 \end{pmatrix}; \qquad X' = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 2 & 2 & 8 & 2 & \dots & 10 \\ 1 & 2 & 10 & 4 & \dots & 11 \end{pmatrix};$$

$$Y = \begin{pmatrix} 10 \\ 12 \\ 17 \\ \cdots \\ 18 \end{pmatrix}; \qquad \widehat{A} = \begin{pmatrix} a_0 \\ \widehat{a}_1 \\ \widehat{a}_2 \end{pmatrix}$$

Найдем матрицу моментов:

$$B = \left(X' \cdot X \right) = \begin{pmatrix} 10 & 50 & 60 \\ 50 & 336 & 398 \\ 60 & 398 & 480 \end{pmatrix}$$

Вычислим $\left(X' \cdot Y \right) = \begin{pmatrix} 137 \\ 756 \\ 908 \end{pmatrix}$. Найдем B^{-1}, для которой $detB = 7160$:

$$B^{-1} = \left(X' \cdot X \right)^{-1} \cdot = \begin{pmatrix} 0{,}40168 & -0{,}01676 & -0{,}03631 \\ -0{,}01676 & 0{,}16760 & -0{,}13687 \\ -0{,}03631 & -0{,}13687 & 0{,}12011 \end{pmatrix}$$

$$B^{-1} = \left(X' \cdot X \right)^{-1} \cdot X' \cdot Y = \begin{pmatrix} 9{,}39 \\ 0{,}13 \\ 0{,}61 \end{pmatrix} = \begin{pmatrix} \widehat{a}_0 \\ \widehat{a}_1 \\ \widehat{a}_2 \end{pmatrix},$$

откуда $\widehat{a}_0 = 9{,}39$; $\widehat{a}_1 = 0{,}13$; $\widehat{a}_2 = 0{,}61$

Тогда $\widehat{y} = 9{,}39 + 0{,}13 \cdot x_1 + 0{,}61 \cdot x_2 + e$

3.4 Варианты заданий.

Цель работы – ознакомление с одношаговым методом наименьших квадратов и приобретение навыка нахождения оценок параметров эмпирической функции.

ЗАДАНИЕ

По заданным статистическим данным подобрать эмпирическую функцию, если она не задана и:

1. Построить диаграмму рассеяния.

2. Записать эмпирическую функцию.
3. Записать систему нормальных уравнений.
4. Составить расчетную таблицу.
5. Решить полученную систему и записать эмпирическую функцию с найденными параметрами.

(№ задания или варианта указывает преподаватель)

1. Считая, что зависимость между переменными X и Y имеет вид y=ax+в, найти оценки параметров по следующим выборам :

1)

x	54	63	74	90	112	140	190
y	8	10	11	13	15	17	19

2)

x	100	120	110	115	125	130	125	140	140	150
y	12	13	18	19	20	20	25	30	31	35

3)

x	1	3	4	2	5	7	8	9
y	80	90	120	100	110	150	160	130

4)

x	5	4	6	7	3	4	6	7	4	3
y	6,3	6,0	7,5	8,5	3,5	6,2	7,5	8,7	6,0	3,7

5)

x	152	116	100	108	129	141	147	156	163
y	47	34	31	32	38	42	45	47	49

6)

x	90	110	120	130	180	200	280
y	25	28	31	32	36	42	55

7)

x	2	4	3	5	2	2	5
y	13	15	12	16	15	11	14

8)

x	6,0	6,1	6,8	7,2	7,4	7,9	8,2	8,5	8,6	9,1
y	2	3	6	4	3	3	4	5	6	8

9)

x	6	8	9	9	10	11	11	13	14	15
y	4	4	5	7	5	6	8	7	9	10

10)

x	8	9	10	11	12	13	14	16	17	19
y	9	8,5	9,2	9,6	9,4	10,5	11,2	10,8	11,0	11,5

11)

x	2	2	3	4	5	6	7	8	9	11
y	2,5	3,1	3,0	3,5	4,2	5,1	5,5	6,0	6,2	6,4

12)

x	66	70	75	80	82	85	90	92	95	98
y	60	68	65	78	74	70	78	85	88	90

13)

x	13	14	15	17	18	18	18	19	22	25
y	7	9	10	12	11	14	15	15	16	18

14)

x	0	4	10	15	21	29	36	51	68
y	66,7	71,0	76,3	80,6	85,7	92,9	99,4	113,6	125,1

15)

x	0,30	0,91	1,50	2,00	2,20	2,62	3,00	3,30
y	0,20	0,43	0,35	0,52	0,81	0,68	1,15	0,85

16)

x	37	47	49	51	61	75	80	92	102	117	120	122
y	53	42	30	24	22	22	26	31	35	38	38	36

17)

x	1	2	3	4	5	6	7	8	9	10	11	12	13
y	7,6	7,2	6,2	8,3	8,2	7,6	7,9	7,5	8,5	8,7	7,0	8,8	8,5

18)

x	1	5	6	7	8	9	10	11
y	100	156	170	184	194	205	220	229

19)

x	1	2	3	4	5	6	7	8	9	10	11	12
y	100	113	121	148	183	194	219	260	277	304	338	352

20)

x	2	5	8	10	14	15	4	12	3	7	6
y	14,39	9,45	7,05	5,32	16,94	1,97	8,75	3,41	13,37	8,22	9,39

21)

x	2,7	4,6	6,3	7,8	9,2	10,6	12,0	13,4	14,7
y	17.0	16,2	13,3	13,0	9,7	9,9	6,2	5,8	5,7

22)

x	7,9	11,6	12.8	14,9	16,3	18,6	20.3	21,9	23,6
y	13.0	22,8	24,8	28,6	31,6	38,7	40,0	44,9	43,0

23)

x	1	2	3	4	5	6	7	8	9	10	11	12
y	0,21	0,32	0,58	1,02	1,76	2.68	3.75	5,07	6,62	8,32	10,21	12,33

24)

x	2	4	6	8	10
y	4,5	7	8	7,5	9

2. Считая, что зависимость между переменными x и y имеет вид $y = ax^2 + вx + c$, найти оценки параметров по следующим выборкам:

1)

x	2,0	3,5	4,0	4,5	5,5	6,0
y	1,9	1,7	1,8	1,6	1,5	1,4

2)

x	5,0	6.0	6,5	7,0	8,0
y	25	2,8	31	35	40

3)

x	2	3	4	4	5	6	6	6	7	8
y	8	10	7	6	5	5	4	3	4	5

4)

x	40	55	64	75	82	94	104	110	115	120
y	2,8	4,3	4,6	4,9	5,6	6,4	7,7	7,9	10,2	9,8

5)

x	10	20	30	40	50	60	70	80	90
y	4,2	12,6	14,8	16,8	21,0	22,2	22,8	21,8	19,4

6)

x	7	12	17	22	27	32	37
y	83,7	72,9	63,2	54,7	47,5	41,4	36,3

7)

x	12,0	13,1	14,0	16,1	17,4	18,0	
y	54	59	67	76	85	97	
x	20,0	21,4	21,9	24,1	25,0	26,8	28,1
y	107	118	127	139	153	160	178

8)

x	-2	-1	0	1	2	3
y	-3	-4	-3	0	5	12

9)

x	1	2	3	4	5	6	7
y	0,5	0,5	1	2	3	5	8

В последующих примерах взять эмпирическую функцию в виде:

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

10)

x	0	2	4	6	8	10
y	5	-1	-0,5	1,5	4,5	8,5

11)

x	0,07	0,31	0,61	0,99	1,29	1,78	2,09
y	1,34	1,08	0,94	1,06	1,25	2,01	2,6

12)

x	26	30	34	38	42	46	50
y	3,94	4,60	5,67	6,93	8,25	7,73	10,55

13)

x	-2	-1	0	1	2
y	4,8	0,4	-3,4	0,8	3,2

14)

x	-3	-2	-1	0	1	2	3
y	-10	0	4	5	4	2	-2

15)

x	-3	-2	-1	0	1	2	3
y	-5	-2	-1,5	-1	0	3	14

16)

x	-3	-2	-1	0	1	2	3
y	6	0	-1	-1	1	5	12

17)

x	-2	-1	0	1	2
y	3	0	3	6	9

18)

x	-3	-2	-1	0	2	3
y	-6	-4	-2	-1	1	0

19)

x	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0
y	0,4	0,3	1,0	1,7	2,1	3,4	4,1	5,8	7,7	9,4	11,4	13,6

20)

x	0,4	0,8	1,2	1,6	2	2,4	2,8
y	0,43	0,94	1,91	3,01	4	4,56	6,45
x	3,2	3,6	4,0	4,4	4,8	5,2	
y	8,59	11,15	13,88	16,93	20,47	24,15	

21)

x	0	0,05	0,10	0,15	0,20	0,25	0,30
y	25	26	4	7	6	13	30
x	0,35	0,40	0,45	0,50	0,55	0,60	
y	26	32	40	32	21	11	

22)

x	1,0	1,5	2,0	2,5	3,0	3,5
y	0,22	0,23	0,31	0,43	0,56	0,82
x	4,0	4,5	5,0	5,5	6,0	6,5
y	1,06	1,25	1,72	2,28	2,67	3,26

23)

x	5	10	15	20	25
y	50	59,3	59,8	64,9	70,2

24)

x	87,5	84,0	77,8	63,7	46,7	36,9
y	292	283	270	235	197	181

3. Считая, что зависимость X и Y имеет вид $y = a + \frac{b}{x}$, найти оценки параметров по следующим выборкам:

1)

x	1	2	3	4	5
y	16,50	13,75	13,31	12,50	13,52
x	6	7	8	9	10
y	12,75	12,30	12,83	12,28	1 2,34

2)

x	1	2	3	4	5
y	1,25	1,15	1,00	0,80	0,65
x	10	20	30	50	100
y	0,41	0,36	0,20	0,15	0,1

3)

x	54	63	74	90	112	140	190
y	0,50	0,70	0,80	1,00	1,40	2,20	2,50

4)

x	40	55	64	75	82	94	104	110	115	120
y	2,8	4,3	4,6	4,9	5,6	6,4	7,7	7,9	10,2	9,8

5)

x	54	63	74	90	112	140	190
y	0,15	0,20	0,35	0,55	0,95	1,40	1,60

6)

x	1,0	0,5	0,07	0,3	0,25	0,34	0,13	0,08	0,22	0,58
y	1,6	1,0	8,5	5,0	4,4	2,0	6,0	7,5	3,8	1,4

7)

x	75	90	120	150	180	220	300	450	600	800
y	10	9,2	8,1	7,8	7,9	7,0	6,1	5,8	5,3	5,0

8)

x	5,67	4,45	3,84	3,74	3,73	2,18
y	6,8	8,5	10,5	10,2	6,8	11,8

9)

x	2	4	6	12
y	8	5,25	3,50	3,25

10)

x	1	2	3	4	5
y	16,50	13,75	13,31	12,50	13,52
x	6	7	8	9	10
y	12,75	12,30	12,83	12,28	12,34

11)

x	2,6	5,2	7,8	15,6
y	8	5,25	3,50	3,25

12)

x	5,67	4,45	3,84	3,74	3,73	2,18
y	6,8	8,5	10,5	10,2	6,8	11,8

13)

x	3	6	9	18
y	8	5,25	3,50	3,25

14)

x	1	2	3	5	10	20	30	50	100	200
y	10,15	5,52	4,08	2,85	2,11	1,62	1,41	1,30	1,21	1,15

15)

x	8,8	11	13,2	14,85	15,4
y	80	72	65	70	68

16)

x	8,1	16,1	21,8	43,9	65,8	87,6	96,5
y	0,330	0,271	0,242	0,183	0,158	0,142	0,138

17)

x	8	10	12	13,5	14
y	80	72	65	70	68

18)

x	0,125	0,25	0,5	1	2	4	8
y	170	90	50	30	20	15	12,5

19)

x	5	7	9	11	13	15	17	19
y	3,6	4,0	4,4	4,8	5,2	5,6	6,0	6,4

20)

x	5	7	9	11	13	15	17	19
y	6,16	5,31	4,83	4,53	4,32	4,17	4,05	3,96

21)

x	37,8	44,1	51,8	63	78,4	98	133
y	0,15	0,20	0,35	0,55	0,95	1,40	1,60

22)

x	37,5	45	60	75	90	110	150	225	300	400
y	10	9,2	8,1	7,8	7,9	7,0	6,1	5,8	5,3	5,0

23)

x	3,5	7	10,5	14	17,5
y	59,3	59,8	60,1	64,9	70,2

24)

x	37,8	44,1	51,8	63,0	78,4	98	133
y	0,50	0,70	0,80	1,0	1,40	2,20	2,50

4.Считая, что зависимость между переменными x_1, x_2, и у имеет вид $y = a_0 + a_1x_1 + a_2x_2 + e$, найти оценки параметров МНК в матричной форме.

Эти задания составлены в девяти вариантах, поэтому выбор варианта зависит от начальной буквы фамилии студента.

Начальная буква фамилии студента	Номер варианта
А, Б, В	первый
Г, Д, Е	второй
Ж, З, И	третий
К, Л, М	четвертый
Н, О, П	пятый
Р, С, Т	шестой
У, Ф, Х	седьмой
Ц, Ч, Ш	восьмой
Щ, Э, Ю, Я	девятый