Московский университет им. С.Ю. Витте
Рейтинговая работа
Эконометрика
Вариант 5

Задание

В соответствие со своим вариантом выбрать исходные данные. Выполнить следующие расчеты:

1. Построить модель парной линейной регрессии $y = a + b \cdot x + e$.

2. Изобразить на графике исходные и модельные значения.

3. Рассчитать коэффициенты корреляции и эластичности, коэффициенты эластичности сопоставить с коэффициентами регрессии.

4. Сделать прогноз на следующий шаг.

Таблица 1

Исходные данные

x_i	33,0	24,9	34,1	30,0	28,6	30,8	37,1	19,6	34,1	39,2	41,0
y_i	67	52	82	40	70	84	88	43	76	91	95

Решение

1. Для построения модели парной линейной регрессии построим вспомогательную таблицу (табл.2)

Таблица 2

Вспомогательные расчеты

i	x	y	x^2	y^2	xy
1	33	67	1089	4489	2211
2	24,9	52	620,01	2704	1294,8
3	34,1	82	1162,81	6724	2796,2
4	30	40	900	1600	1200
5	28,6	70	817,96	4900	2002
6	30,8	84	948,64	7056	2587,2
7	37,1	88	1376,41	7744	3264,8
8	19,6	43	384,16	1849	842,8
9	34,1	76	1162,81	5776	2591,6
10	39,2	91	1536,64	8281	3567,2
11	41	95	1681	9025	3895
Сумма	352,4	788	11679,44	60148	26252,6
Среднее	32,036	71,636	1061,767	5468	2386,6

Перейдем к нахождению коэффициентов регрессии.

$$b = \frac{\overline{xy} - \overline{x} \cdot \overline{y}}{\overline{x^2} - \left(\overline{x}\right)^2}, \qquad a = \overline{y} - b \cdot \overline{x},$$

где,

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i =$$

$$= \frac{33 + 24{,}9 + 34{,}1 + 30 + 28{,}6 + 30{,}8 + 37{,}1 + 19{,}6 + 34{,}1 + 39{,}2 + 41}{11} = 32{,}036;$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{67 + 52 + 82 + 40 + 70 + 84 + 88 + 43 + 76 + 91 + 95}{11} = 71{,}636;$$

$$\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2 = \frac{1}{11} \cdot \left(33^2 + 24{,}9^2 + 34{,}1^2 + 30^2 + 28{,}6^2 + 30{,}8^2 + \right.$$

$$\left. + 37{,}1^2 + 19{,}6^2 + 34{,}1^2 + 39{,}2^2 + 41^2\right) = 1061{,}767;$$

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i = \frac{1}{11}\cdot(33\cdot 67 + 24{,}9\cdot 52 + 34{,}1\cdot 82 + 30\cdot 40 + 28{,}6\cdot 70 + 30{,}8\cdot 84 +$$
$$+37{,}1\cdot 88 + 19{,}6\cdot 43 + 34{,}1\cdot 76 + 39{,}2\cdot 91 + 41\cdot 95) = 2386{,}6.$$

Тогда,

$$b = \frac{\overline{xy} - \overline{x}\cdot\overline{y}}{\overline{x^2} - (\overline{x})^2} = \frac{2386{,}6 - 32{,}036\cdot 71{,}636}{1061{,}767 - (32{,}036)^2} = 2{,}5856,$$

$$a = 71{,}636 - 2{,}5856\cdot 32{,}036 = -11{,}1979.$$

Для проверки правильности расчетов воспользуемся функцией Регрессия из Пакета анализа Excel. (рис.1)

ВЫВОД ИТОГОВ						
Регрессионная статистика						
Множественный R	0,839					
R-квадрат	0,705					
Нормированный R-квадрат	0,672					
Стандартная ошибка	11,017					
Наблюдения	11					
Дисперсионный анализ						
	df	*SS*	*MS*	*F*	*Значимость F*	
Регрессия	1	2606,1768	2606,1768	21,4722	0,0012	
Остаток	9	1092,3687	121,3743			
Итого	10	3698,5455				
	Коэффициенты	*Стандартная ошибка*	*t-статистика*	*P-Значение*	*Нижние 95%*	*Верхние 95%*
Y-пересечение	-11,198	18,182	-0,616	0,553	-52,329	29,933
Переменная X 1	2,586	0,558	4,634	0,001	1,323	3,848

Рис. 1 – Результаты регрессии в Excel

Из рисунка: Y-пересечение соответствует значению параметра a, Переменная X1 – значение параметра b.

Таким образом, получим уравнение парной линейной регрессии:

$$\hat{y} = -11{,}1979 + 2{,}5856\cdot x. \qquad (1)$$

2. Построим таблицу с расчетными данными регрессии:

Таблица 3

Расчет модельных значений регрессии

i	x	$\hat{y}$
1	33	74,12797
2	24,9	53,18435

3	34,1	76,97217
4	30	66,37108
5	28,6	62,75119
6	30,8	68,43958
7	37,1	84,72907
8	19,6	39,48049
9	34,1	76,97217
10	39,2	90,1589
11	41	94,81304

Изобразим на графике исходные и модельные значения.

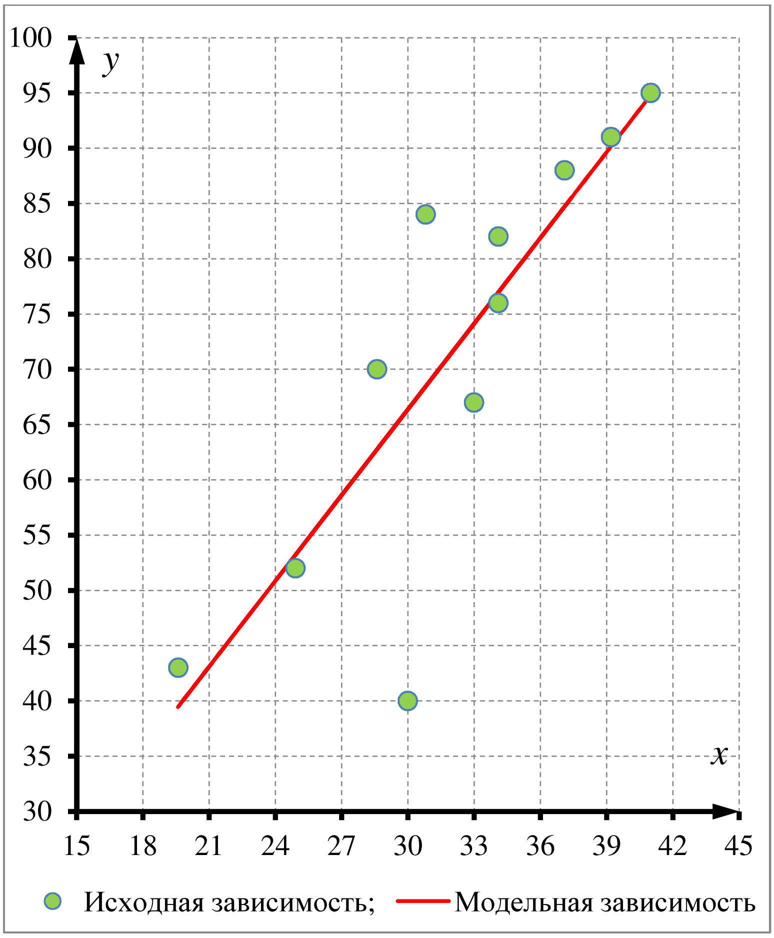

Рис. 2

3. Построим вспомогательную таблицу.

Таблица 4

Вспомогательные расчеты

i	x	y	y^2	$\hat{y}$	$e = y - \hat{y}$	e^2
1	33	67	4489	74,128	-7,128	50,808

2	24,9	52	2704	53,184	-1,184	1,403
3	34,1	82	6724	76,972	5,028	25,279
4	30	40	1600	66,371	-26,371	695,434
5	28,6	70	4900	62,751	7,249	52,545
6	30,8	84	7056	68,440	15,560	242,127
7	37,1	88	7744	84,729	3,271	10,699
8	19,6	43	1849	39,480	3,520	12,387
9	34,1	76	5776	76,972	-0,972	0,945
10	39,2	91	8281	90,159	0,841	0,707
11	41	95	9025	94,813	0,187	0,035
Сумма	352,4	788	60148			1092,369
Среднее	32,036	71,636	5468			

Найдем статистические суммы:

Из таблицы 4 видно, что $Q_e = \sum_{i=1}^{n} e_i^2 = 1092{,}369$, $\sum_{i=1}^{n} y_i^2 = 60148$.

$$Q_y = \sum_{i=1}^{n} y_i^2 - n \cdot \left(\overline{y}\right)^2 = 60148 - 11 \cdot 71{,}636^2 = 3698{,}545.$$

Коэффициент детерминации равен:

$$R^2 = 1 - \frac{Q_e}{Q_y} = 1 - \frac{3698{,}545}{60148} = 0{,}705,$$

что соответствует значению из регрессии (рис.1) полю R-квадрат.

Выборочный коэффициент корреляции равен:

$$r_{xy} = \sqrt{R^2} = \sqrt{0{,}705} = 0{,}839,$$

что соответствует значению из регрессии (рис.1) полю Множественный R.

Коэффициент корреляции показывает очень сильную связь между x и y.

Рассчитаем коэффициент эластичности:

$$Э_{yx} = b \cdot \frac{\overline{x}}{\overline{y}} = 2{,}5856 \cdot \frac{32{,}036}{71{,}636} = 1{,}156.$$

Таким образом, при изменении x на 1% от своей средней величины, показатель y изменится в среднем на 1,156%.

4. Сделаем прогноз на следующий шаг. Подставим в формулу (1) прогнозное значение $x_{прогн.} = 43$ и получим прогнозное значение у:

$\hat{y}_{прогн.} = -11{,}1979 + 2{,}5856 \cdot 43 = 99{,}984$.