

Лабораторная работа №3

Множественная линейная регрессия

Карпенко Дмитрий МП-403

Цель работы: Построение и статистический анализ множественной линейной регрессионной модели, описывающей зависимость издержек производства от основных производственных фондов и численности занятых в производстве.

Исходные данные:

Для исследования используется выборка объемом $n = 50$ наблюдений со следующими переменными:

- Y —издержки производства (тыс. руб.)
- X_1 —основные производственные фонды (тыс. руб.)
- X_2 —численность занятых в производстве (чел.)

Построение регрессионной модели:

Модель множественной линейной регрессии имеет вид: $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$

Оценка параметров модели выполнена методом наименьших квадратов в EViews с использованием уравнения: $Y = C(1) + C(2)*X_1 + C(3)*X_2$.

Получены следующие оценки коэффициентов:

- $\hat{\beta}_0 = 204.1017$
- $\hat{\beta}_1 = 1.200829$
- $\hat{\beta}_2 = 1.547608$

Оценочное уравнение регрессии: $\hat{y} = 204.1017 + 1.200829 \cdot x_1 + 1.547608 \cdot x_2$

Статистический анализ модели

Значимость коэффициентов регрессии:

Для проверки значимости коэффициентов использован t-критерий Стьюдента с уровнем значимости $\alpha = 0.05$ и числом степеней свободы $n-k-1 = 50-2-1 = 47$.

Коэффициент	Оценка	Std. Error	t-Statistic	Prob	Значимость
$\hat{\beta}_0$	204.1017	3.135214	65.099	0.0000	значим
$\hat{\beta}_1$	1.200829	0.002943	408.104	0.0000	значим
$\hat{\beta}_2$	1.547608	0.024008	64.462	0.0000	значим

Все коэффициенты регрессии статистически значимы ($p < 0.05$), что подтверждает их существенное влияние на издержки производства.

Качество модели:

Коэффициент детерминации: $R^2 = 0.999744$

Значение R^2 близко к 1, что свидетельствует о высокой согласованности по строенной модели с исходными данными. Модель объясняет 99.97% вариации издержек производства.

Значимость уравнения регрессии

Проверка значимости уравнения в целом выполнена с использованием F-критерия Фишера:

- F-statistic: 95433.71
- Prob(F-statistic): 0.000000

Уравнение статистически значимо, так как Prob(F-statistic) < 0.05.

Доверительные интервалы коэффициентов

Для построения 95% доверительных интервалов использовано табличное значение t критерия Стьюдента: $t_{\alpha} = 2.011741$

Формула для расчета доверительных интервалов: $\hat{\beta}_i \in (\hat{\beta}_i - t_{\alpha} \cdot s.e.(\hat{\beta}_i); \hat{\beta}_i + t_{\alpha} \cdot s.e.(\hat{\beta}_i))$.

Результаты расчетов:

$$\hat{\beta}_0 \in (197.794; 210.4089)$$

$$\hat{\beta}_1 \in (1.19491; 1.2067)$$

$$\hat{\beta}_2 \in (1.4993; 1.59591)$$

С вероятностью 95% можно утверждать, что истинные значения коэффициентов регрессии находятся в пределах вычисленных интервалов.

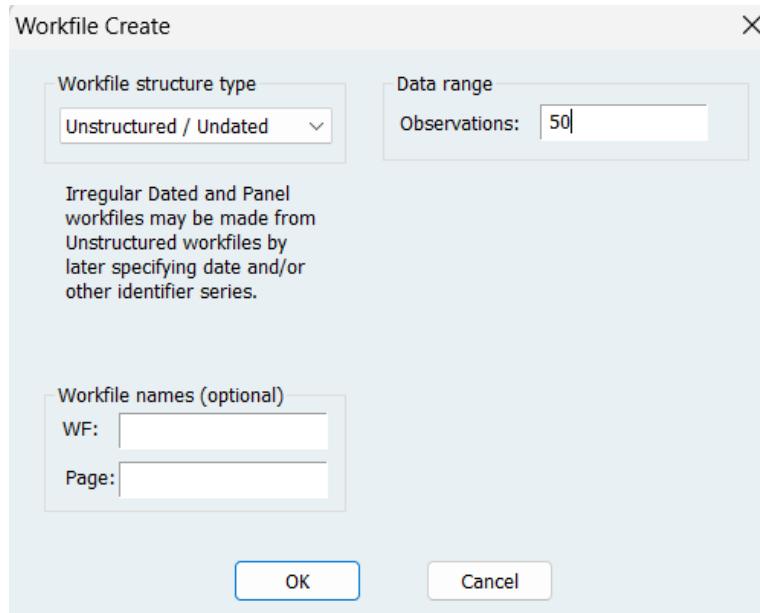
Прогнозирование

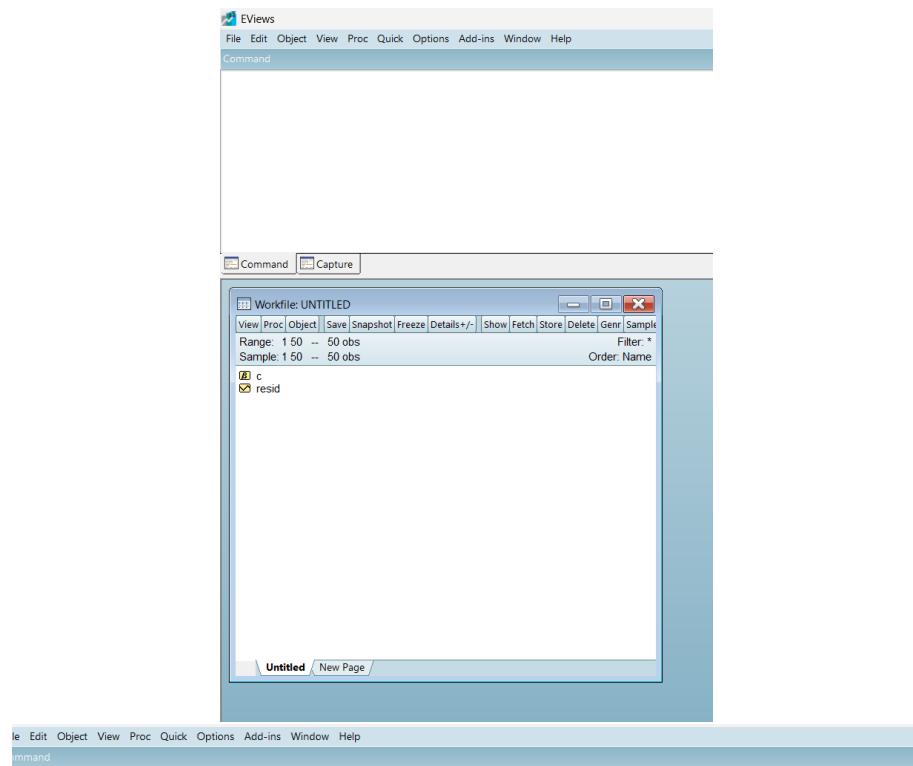
Определим уровень издержек производства при $X_1 = 1500$ тыс. руб. и $X_2 = 12$ чел:

$$Y_T = 204.1017 + 1.200829 \cdot 1500 + 1.547608 \cdot 12 = 2023.918 \text{ тыс. руб.}$$

При заданных условиях ожидаемые издержки производства составят 2023.918 тыс. руб.

Скриншоты в процессе работы:





Workfile: UNTITLED			
		Y	X1
30	30	1411.25	933.46
31	31	763.83	338.25
32	32	1417.02	921.36
33	33	1417.17	825.42
34	34	1393.94	890.51
35	35	1377.13	785.45
36	36	1071.33	604.52
37	37	1958.56	1325.41
38	38	1945.18	1282.41
39	39	960.49	515.4
40	40	1500.46	910.42
41	41	1927.52	1254.89
42	42	1043.73	575.27
43	43	1313.84	812.55
44	44	1392.15	785.3
45	45	820.61	408.16
46	46	925.17	449.92
47	47	1103.44	661.43
48	48	891.7	514.35
49	49	1006.11	582.08
50	50	1301.78	835.23

Command | Capture

Workfile: UNTITLED

View Proc Object Save Snapshot Freeze Details+/- Show Fetch Store Delete Gen Sample

Range: 1 50 -- 50 obs Filter: *

Sample: 1 50 -- 50 obs Order: Name

(B) c
(B) mn_regres
resid
x1
x2
y

Group: MN_REGRES Workfile: UNTITLED:Untitled\

View Proc Object Print Name Freeze Default Sort Edit+/- Smpl+/- Compare+/-

	Y	X1	X2
30	1411.25	933.46	62
31	763.83	338.25	99
32	1417.02	921.36	68
33	1417.17	825.42	140
34	1393.94	890.51	74
35	1377.13	785.45	150
36	1071.33	604.52	95
37	1958.56	1325.41	102
38	1945.18	1282.41	128
39	960.49	515.4	87
40	1500.46	910.42	135
41	1927.52	1254.89	143
42	1043.73	575.27	96
43	1313.84	812.55	90
44	1392.15	785.3	159
45	820.61	408.16	80
46	925.17	449.92	120
47	1103.44	661.43	63
48	891.7	514.35	50
49	1006.11	582.08	66
50	1301.78	835.23	63

Equation: EQ1 Workfile: UNTITLED:Untitled\

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 11/27/25 Time: 13:45
Sample: 1 50
Included observations: 50
 $Y=(C1+C2*X1+C3*X2)$

Coefficient	Std. Error	t-Statistic	Prob.
C(1)	204.1017	3.135214	65.09976 0.0000
C(2)	1.209829	0.002943	408.0967 0.0000
C(3)	1.547608	0.024008	64.46287 0.0000

R-squared: 0.999744 Mean dependent var: 1330.611
Adjusted R-squared: 0.999733 S.D. dependent var: 356.5675
S.E. of regression: 5.827774 Akaike info criterion: 6.421272
Sum squared resid: 1596.259 Schwarz criterion: 6.535993
Log likelihood: -157.5318 Hannan-Quinn criter.: 6.464958
F-statistic: 91692.29 Durbin-Watson stat: 1.908620
Prob(F-statistic): 0.000000

Command Capture

G Group: MN_REGRES Workfile: UNTITLED:Untitled

View Proc Object Save Snapshot * Range: 1 50 — 50 obs Sample 1 50 — 50 obs

		Y	X1	X2	YT
1	1	650.88	246.95	99	2023.916496
2	2	1915.48	1123.56	167	2023.916496
3	3	807.53	442.16	46	2023.916496
4	4	1462.61	1038.71	20	2023.916496
5	5	1307.49	791.07	101	2023.916496
6	6	1710.4	1137.65	89	2023.916496
7	7	1430.44	908.92	84	2023.916496
8	8	1774.42	681.17	89	2023.916496
9	9	1309.36	663.25	43	2023.916496
10	10	1464.13	965.68	81	2023.916496
11	11	1058.87	607.24	83	2023.916496
12	12	1010.98	541.3	105	2023.916496
13	13	1515.41	1069.7	22	2023.916496
14	14	1746.04	1108.75	138	2023.916496
15	15	2190.04	1490.49	97	2023.916496
16	16	709.53	253.78	126	2023.916496
17	17	1581.01	968.32	139	2023.916496
18	18	1648.89	1094.29	80	2023.916496
19	19	1158.94	682.56	92	2023.916496
20	20	1924.64	1268.93	130	2023.916496
21	21	**n/a**	***n/a***	n/a	2023.916496
22					

Untitled New Page

Equation: EQ1 Workfile: UNTITLED:Untitled

View Proc Object Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y
Method: Least Squares (Gauss-Newton / Marquardt steps)
Date: 11/27/25 Time: 13:45
Sample: 1 50
Included observations: 50
Y=C(1)+C(2)*X1+C(3)*X2

Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-204.1017	3.135214	65.09978 0.0000
C(2)	1.200829	0.002943	408.0967 0.0000
C(3)	1.547608	0.024008	64.46287 0.0000

R-squared: 0.999744 Mean dependent var: 1320.611
Adjusted R-squared: 0.999733 S.D. dependent var: 356.5675
S.E. of regression: 5.827774 Akaike info criterion: 6.421272
Sum squared resid: 1596.259 Schwarz criterion: 6.535993
Log likelihood: -157.5318 Hannan-Quinn criter: 6.464958
F-statistic: 91692.29 Durbin-Watson stat: 1.908620
Prob(F-statistic): 0.000000