

Раздел 1. Парная и множественная регрессия и корреляция в эконометрических исследованиях

Задача 1. Построить уравнение множественной линейной регрессии формирования сбыта товара в зависимости от затрат на рекламу и цены единицы товара (табл. 1.1).

Таблица 1.1. Исходные данные для построения уравнения регрессии

№ п.п.	Объем сбыта товара, у. д. е. (y)	Цена единицы товара, у. д. е. (x ₁)	Затраты на рекламу товара, у. д. е. (x ₂)	x ₁ ²	x ₂ ²	x ₁ x ₂	x ₁ y	x ₂ y
1	115	4	12	16	144	48	115	1380
2	130	2	13	4	169	26	130	1690
3	100	7	6	49	36	42	100	600
4	90	9	4	81	16	36	90	360
5	115	4	13	16	169	52	115	1495
6	125	4	12	16	144	48	125	1500
7	125	3	14	9	196	42	125	1750
8	85	9	4	81	16	36	85	340
9	115	3	14	9	196	42	115	1610
10	110	5	8	25	64	40	110	880
11	120	3	14	9	196	42	120	1680
12	100	6	7	36	49	42	100	700
13	90	8	5	64	25	40	90	450
14	115	4	12	16	144	48	115	1380
15	120	5	15	25	225	75	120	1800
16	125	4	15	16	225	60	125	1875
17	115	3	10	9	100	30	115	1150
18	110	5	10	25	100	50	110	1100
19	105	5	8	25	64	40	105	840
20	95	7	6	49	36	42	95	570
21	105	4	8	16	64	32	105	840
22	105	17	9	289	81	153	105	945
Сумма	2415	121	219	885	2459	1066	2415	24935
Среднее	109,8	5,5	10,0	40,2	111,8	48,5	110	1133

На основе приведенной информации необходимо:

1) дать количественную характеристику качественному признаку, если он присутствует в корреляционной модели (КМ).

Если качественный признак присутствует в опыте, определяем его количественно – ставим единицу, если отсутствует – ставим нуль. В КМ вводим столько дискретных величин, сколько качественных признаков;

2) используя формулы асимметрии А и эксцесса Э, проверить данные на соответствие их требованиям закона нормального распределения:

$$A^* = \frac{\sum_{i \in I_0} (x_i - \bar{x})^3}{n \sigma_x^3}; \quad \mathcal{E}^* = \frac{\sum_{i \in I_0} (x_i - \bar{x})^4}{n \sigma_x^4},$$

где А*, Э* – данные формулы распространяются и на результирующий показатель y;

i – номер варианта опыта (хозяйства);

I₀ – множество вариантов опытов (хозяйств);

x_i – фактическое значение факторного показателя;

\bar{x} – среднее значение факторного показателя;

n – число вариантов опыта;

σ_x – среднее квадратическое отклонение;

Среднее квадратическое отклонение определяем по одной из формул:

$$\sigma_x = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} = \sqrt{s^2 - \bar{x}^2},$$

Информация не противоречит требованиям закона нормального распределения, если фактические значения А и Э равны нулю или не противоречат условиям:

$$|A| \leq 3\sigma_A, \quad |\mathcal{E}| \leq 5\sigma_{\mathcal{E}},$$

где $\sigma_A, \sigma_{\mathcal{E}}$ – средние квадратические отклонения, или стандартные ошибки асимметрии и эксцесса.

Стандартные ошибки асимметрии и эксцесса определяются по следующим формулам:

$$\sigma_A = \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)'}}$$

$$\sigma_{\varepsilon} = \sqrt{\frac{24n(n-1)^2}{(n-3)(n-2)(n+3)(n+5)}}$$

Если А и Э выходят за допустимые границы, проверяем, нет ли среди данных резко выделяющихся вариантов. Данные принадлежат выборке, если $|x_i - \bar{x}| \leq 3\sigma_x$.

В случае невыполнения условия для какого-то варианта (вариантов) информацию об этих исключаем из выборки и вновь проверяем оставшиеся данные на соответствие их требованиям закона нормального распределения;

3) определить вид корреляционной модели одним из известных способов: аналитическим, графическим, логическим;

4) используя метод наименьших квадратов, рассчитать параметры корреляционной модели;

5) определить коэффициенты парной корреляции по формуле

$$r_{xy} = \frac{\bar{xy} - \bar{x}\bar{y}}{\sigma_x \sigma_y};$$

6) определить коэффициент множественной корреляции, используя одну из следующих формул:

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}}{1 - r_{x_1x_2}^2}};$$

$$R = \sqrt{1 - \frac{\sum_{i \in I_0} (y_x - y_i)^2}{\sum_{i \in I_0} (y_i - \bar{y})^2}},$$

где y_x – расчетное значение результативного показателя;

y_i – фактическое значение результативного показателя;

\bar{y} – среднее значение результативного показателя;

7) определить коэффициент существенности коэффициента парной (множественной) корреляции:

$$t_r = \frac{r}{\mu_r}, \quad t_R = \frac{R}{\mu_R}$$

где μ_r (μ_R) – ошибка коэффициента парной (множественной) корреляции.

Ошибка коэффициента парной (множественной) корреляции определяется по формулам:

$$\mu_r = \frac{1 - r^2}{\sqrt{n - 1}}, \quad \mu_R = \frac{1 - R^2}{\sqrt{n - k - 1}}$$

где k – количество факторов;

8) определить существенность коэффициентов регрессии:

$$t_{a_j} = \frac{a_j}{\mu_{a_j}}$$

где a_j – значение коэффициента регрессии для j-го фактора;

μ_{a_j} – ошибка коэффициента регрессии.

Ошибка коэффициента регрессии определяется по формуле

$$\mu_{a_j} = \frac{\sigma_{y_x y_i}}{\sigma_x \sqrt{n}}$$

где $\sigma_{y_x y_i}$ – среднее квадратическое отклонение по данным КМ.

Среднее квадратическое отклонение по данным КМ определяется по формуле

$$\sigma_{y_x y_i} = \sqrt{\frac{\sum_{i \in I_0} (y_x - y_i)^2}{n}}$$

Сравнить табличное значение коэффициента существенности коэффициента регрессии t_r с расчетным значением.

В случае если $t_{a_j} < t_r$, то несущественный фактор нужно исключить из КМ и ее параметры рассчитать вновь, но без фактора, который исключен;

9) дать экономическую интерпретацию коэффициентам регрессии. Определить коэффициенты эластичности (ε_j) и β_j -коэффициенты и сравнить влияние факторных показателей на формирование результативного показателя:

$$\varepsilon_j = a_j \frac{\bar{x}_j}{\bar{y}_i}, \quad \beta_j = \frac{\sigma_{x_j}}{\sigma_y};$$

10) сравнить фактические и расчетные значения результативного показателя и сделать выводы об особенностях использования ресурсов на каждом предприятии;

11) рассчитать параметры и характеристики КМ на персональном компьютере (ПК).

Задача 2. Проверить данные столбца X на соответствие их требованиям закона нормального распределения, если $n = 30$, $\sum(x_i - \bar{x})^2 = 35 + N$; $\sum(x_i - \bar{x})^3 = 90 + N$; $\sum(x_i - \bar{x})^4 = 350 + N$; $\sigma_A = 0,43$; $\sigma_3 = 0,83$.

Задача 3. Рассчитать коэффициент парной корреляции, если $n=40$, $\sum x_1 = 25 + N$; $\sum y = 30$; $\sum xy = 24$; $\sum(x_i - \bar{x})^2 = 31 + N$; $\sum(y_i - \bar{y})^2 = 29$. Пояснить полученное значение.

Задача 4. Рассчитать коэффициент множественной корреляции и сделать вывод, если $(y_i - y_x)^2 = 20$; $\sum(y_i - \bar{y})^2 = 140 + N$. Определить, на сколько процентов включенные в регрессионную модель факторы объясняют изменение резульативного показателя.

Задача 5. Рассчитать критерий Фишера и сделать вывод, если $\sum(y_i - y_x)^2 = 100$; $\sum(y_i - \bar{y})^2 = 360 + N$.

Задача 6. Рассчитать критерий существенности коэффициента регрессии $a_1 = 1,03$, если известно, что $n = 20$; $\sigma_x = 2,5$; $\sum(y_i - y_x)^2 = 24 + 0,1N$. Сделать вывод.

Задача 7. Рассчитать коэффициенты β и эластичности, сделать вывод, если известно: $\bar{x}_1 = 48 + 0,1N$; $\bar{x}_2 = 4,6 + 0,1 + N$; $\bar{y} = 76$; $\sigma_{x1} = 24,3$; $\sigma_{x2} = 22,6$; $\sigma_y = 70,4$, а модель имеет вид: $y = 153 + 1,8x_1 + 2,8x_2$.

Задача 8. Найти коэффициент множественной корреляции и сделать вывод, если критерий Фишера (F) = $5,9 + 0,3N$.

Задача 9. Определить расчетное значение себестоимости картофеля (ден. ед/ц) в хозяйстве, если урожайность культуры равна $(300 + 10 + N)$ ц/га, расход удобрений – $3,3$ ц д. в/га, а регрессионная модель формирования себестоимости картофеля имеет вид

$$y = 1458 - 1,2x_1 + 2,5x_2,$$

где y – себестоимость картофеля, ден. ед/ц;

x_1 – урожайность картофеля, ц/га.;

x_2 – расход удобрений, ц д. в/га.

Задача 10. По информации сельскохозяйственных организаций рассчитана регрессионная модель:

$$y = 190,6 + 4,82x_1 - 1,42x_2 + 7,38x_3;$$
$$R = 0,92, D = 0,83, F = 34, t_{a1} = 2,35, t_{a2} = -5,98, t_{a3} = 0,12;$$
$$\beta_1 = 0,25, \beta_2 = 0,16, \beta_3 = 0,31.$$

где y – стоимость валовой продукции, у. д. е.;

x_1 – численность работников, чел.;

x_2 – стоимость основных производственных фондов, у. д. е.;

x_3 – наличие энергетических мощностей, л. с.

Необходимо: дать экономическую интерпретацию коэффициента регрессии при факторах x_1 , x_2 , x_3 и пояснить значение всех приведенных характеристик модели в целом и отдельных факторов.

Задача 11. Проверить данные столбца X на соответствие их требованиям закона нормального распределения, если $n = 20 + N$, $\sum(x_i - \bar{x})^2 = 40 + N$; $\sum(x_i - \bar{x})^3 = 75$; $\sum(x_i - \bar{x})^4 = 210$; $\sigma_A = 0,51$, $\sigma_3 = 0,99$.

Задача 12. Рассчитать коэффициент парной корреляции, если $n = 50$; $\sum x_1 = 30 + N$; $\sum y = 35$; $\sum xy = 25 + N$; $\sum(x_i - \bar{x})^2 = 28 + N$; $\sum(y_i - \bar{y})^2 = 18$. Пояснить полученное значение.

Задача 13. Рассчитать коэффициент множественной корреляции и сделать вывод, если $\sum(y_i - y_x)^2 = 180 + 10N$; $\sum(y_i - \bar{y})^2 = 230 + 10N$. Определить, на сколько процентов включенные в регрессионную модель факторы объясняют изменение резульативного показателя.

Задача 14. Рассчитать критерий Фишера и сделать вывод, если $\sum(y_i - y_x)^2 = 120 + N$; $\sum(y_i - \bar{y})^2 = 140 + N$.

Задача 15. Рассчитать критерий существенности коэффициента регрессии $a_1 = 1,85$, если известно, что $n = 20$; $\sigma_x = 0,2$; $\sum(y_i - y_x)^2 = 50 + N$. Сделать вывод.

Задача 16. Рассчитать коэффициенты β и эластичности и сделать вывод, если известно: $\bar{x}_1 = 24$; $\bar{x}_2 = 2,3$; $\bar{y} = 38$; $\sigma_{x1} = 12,3$; $\sigma_{x2} = 11,6$; $\sigma_y = 35,3$, а модель имеет следующий вид: $y = 153 + 1,8x_1 + 2,8x_2$.

Задача 17. Найти коэффициент множественной корреляции и сделать вывод, если критерий Фишера (F) = $1,6 + 0,2N$.

Задача 18. Определить расчетное значение себестоимости зерновых в хозяйстве (ден. ед/ц), если урожайность равна 50 ц/га, расход удобрений – $2,1$ ц д. в/га, а корреляционно-регрессионная модель формирования себестоимости имеет вид:

$$y = 283 - 0,18x_1 + 1,2x_2,$$

где y – себестоимость зерновых, ден. ед/ц;

x_1 – урожайность зерновых, ц/га;

x_2 – расход удобрений, ц д. в/га.

Задача 19. Дать экономическую интерпретацию параметрам и характеристикам корреляционно-регрессионной модели формирования стоимости товарной продукции (y , тыс. руб.):

$$y = 1259,3 + 366,3x_1 + 150,4x_2 + 120,9x_3,$$

$$R = 0,9, D = 0,91, F = 23,8, t_{a1} = 1,99, t_{a2} = -0,98, t_{a3} = 2,1;$$
$$\beta_1 = 0,24, \beta_2 = 0,31, \beta_3 = 0,45,$$

где x_1 – стоимость ОПФ, тыс. руб.;

x_2 – численность работников, чел.;

x_3 – наличие оборотных фондов, тыс. руб. Т