

Расчетный файл выложен на странице https://www.matburo.ru/ex_ec.php?p1=ecexcel

Эконометрика. Множественная регрессия. Решение задачи

Задание. Построить требуемое уравнение регрессии. Вычислить коэффициент детерминации, частные коэффициенты эластичности, частные бета коэффициенты и дать их смысловую нагрузку в терминах задачи. Проверить адекватность уравнения с помощью F теста. Найти оценку матрицы ковариаций оценок параметров регрессии и 95% доверительные интервалы для параметров регрессии. Проверить наличие мультиколлинеарности в модели. Данные взять из таблицы.

Построить уравнение линейной регрессии себестоимости единицы товара (в сотнях руб.) от величины энерговооруженности (кВт) и производительности труда (тов/час).

Энерговооруженность	8,7	8,7	6,5	8	8,4	8,4	7,5
Производительность труда	9	8,5	6,1	9,5	6,5	7,1	8,3
Себестоимость	14,6	14,4	14,4	14,8	14,1	14,1	15,2

Решение.

Введем в рассмотрение матрицы

$$Y = \begin{pmatrix} 14,6 \\ 14,4 \\ 14,4 \\ 14,8 \\ 14,1 \\ 14,1 \\ 15,2 \end{pmatrix} \quad \theta = \begin{pmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \\ \hat{\theta}_2 \end{pmatrix} \quad X = \begin{pmatrix} 1 & 8,7 & 9 \\ 1 & 8,7 & 8,5 \\ 1 & 6,5 & 6,1 \\ 1 & 8 & 9,5 \\ 1 & 8,4 & 6,5 \\ 1 & 8,4 & 7,1 \\ 1 & 7,5 & 8,3 \end{pmatrix}$$

Тогда матрица $X^T X$ примет вид

$$X^T X = \begin{pmatrix} 7 & 56 & 55 \\ 56 & 455 & 444 \\ 55 & 444 & 442 \end{pmatrix}$$

обратная к ней равна

$$(X^T X)^{-1} = \begin{pmatrix} 17,5777 & -1,9407 & -0,2360 \\ -1,9407 & 0,3323 & -0,0926 \\ -0,2360 & -0,0926 & 0,1246 \end{pmatrix}$$

отсюда оценка вектора параметров модели примет вид

$$\hat{\theta} = (X^T X)^{-1} X^T Y = \begin{pmatrix} 15,0867 \\ -0,3473 \\ 0,2820 \end{pmatrix}$$

В результате искомое уравнение имеет вид

$$Y = 15,0867 - 0,3473X_1 + 0,282X_2$$

Найдем векторы прогноза и остатков регрессии

$$\widehat{Y}_x = \begin{pmatrix} 14,6034 \\ 14,4624 \\ 14,5496 \\ 14,9875 \\ 14,0026 \\ 14,1718 \\ 14,8227 \end{pmatrix} \quad \varepsilon = \begin{pmatrix} -0,0034 \\ -0,0624 \\ -0,1496 \\ -0,1875 \\ 0,0974 \\ -0,0718 \\ 0,3773 \end{pmatrix}$$

Оценка дисперсии ошибок:

$$s^2 = \frac{\varepsilon^T \varepsilon}{n-2-1} = 0,0546$$

Матрица отклонений Y от своего среднего значения:

$$y = Y_i - \bar{Y} = \begin{pmatrix} 0,0857 \\ -0,1143 \\ -0,1143 \\ 0,2857 \\ -0,4143 \\ -0,4143 \\ 0,6857 \end{pmatrix}$$

Вычислим коэффициент детерминации

$$R^2 = 1 - \frac{\varepsilon^T \varepsilon}{y^T y} = 0,7648 \text{ или } 76,48\%$$

это показывает, что изменение себестоимости на 76,48% объясняется изменением энерговооруженности и производительности труда.

Вычислим частные коэффициенты эластичности

$$E_1 = \hat{\theta}_1 \cdot \frac{\overline{X_1}}{\overline{Y}} = -0,1921$$

$$E_2 = \hat{\theta}_2 \cdot \frac{\overline{X_2}}{\overline{Y}} = 0,1527$$

это показывает, что при изменении энерговооруженности в среднем на 1%, себестоимость снизится в среднем на 0,1921%, а при изменении производительности труда в среднем на 1%, себестоимость вырастет в среднем на 0,1527%,

Найдем средне квадратические отклонения переменных

$$\sigma_{x_1} = 0,7952$$

$$\sigma_{x_2} = 1,2985$$

$$\sigma_y = 0,3934$$

Тогда частные бета коэффициенты равны

$$\beta_{y/x_1} = \hat{\theta}_1 \cdot \frac{\sigma_{x_1}}{\sigma_y} = -0,702$$

$$\beta_{y/x_2} = \hat{\theta}_2 \cdot \frac{\sigma_{x_2}}{\sigma_y} = 0,9309$$

это показывает, что при изменении разброса энерговооруженности на 0,7952 кВт. разброс себестоимости изменится на $0,3934 \cdot 0,702 = 0,2762$ сот. руб.

При изменении разброса производительности труда на 1,2985 тов/час. разброс себестоимости изменится на $0,3934 \cdot 0,9309 = 0,3662$ сот. руб.

Проверка адекватности построенного уравнения проведем с помощью F теста проверки гипотезы $H_0: \theta_1 = \theta_2 = 0$ на уровне значимости $\alpha = 0;05$. Расчетное значение F теста равно

$$F = \frac{R^2}{1-R^2}(n-2) = 6,5028$$

критическое

$$F_{кр} = F_{0,05}(2;4) = 6,9443$$

расчетное значение меньше критического, следовательно, гипотеза H_0 : $\theta_1 = \theta_2 = 0$ принимается и уравнение нельзя признать адекватным изучаемому процессу.

Оценка матрицы ковариаций оценок параметров регрессии примет вид:

$$\widehat{D(\hat{\theta})} = S^2 (X^T X)^{-1} = \begin{pmatrix} 0,9598 & -0,1060 & -0,0129 \\ -0,1060 & 0,0181 & -0,0051 \\ -0,0129 & -0,0051 & 0,0068 \end{pmatrix}$$

из этой матрицы получаем, что

$$S_0^2 = 0,9598$$

$$S_1^2 = 0,0181$$

$$S_2^2 = 0,0068$$

$$t_{0,05}(4) = 2,7764$$

Тогда доверительные интервалы примут вид:

$$\hat{\theta}_0 : (12,3667 \quad 17,8068)$$

$$\hat{\theta}_1 : (-0,7213 \quad 0,0267)$$

$$\hat{\theta}_2 : (0,0530 \quad 0,5111)$$

Проверим наличие мультиколлинеарности в модели

$$r_{X_1 X_2} = \frac{\overline{X_1 X_2} - \overline{X_1} \overline{X_2}}{\sigma_{X_1} \sigma_{X_2}} = 0,4549 < 0,7$$

то в модели корреляционная связь между переменными X_1 и X_2 - отсутствует, и следовательно, в модели отсутствует явление мультиколлинеарности.