

Корреляционно- регрессионные МОДЕЛИ

Представление исходных данных

$Y_{t1}, Y_{t2}, Y_{t3} \dots Y_{tn}$

Представление исходных данных

- y_1
- y_2
- ...
- y_n

$$\begin{array}{cccc} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{array}$$

$$n > k \sim 3$$

Представление исходных данных

Объект	Время	Признаки			
Объект 1	$t=1$	Y_{11}	X_{11}	Z_{11}
	$t=2$	Y_{21}	X_{21}	Z_{21}

	$t=T$	Y_{1t}	X_{1t}	Z_{1t}
Объект 2	$t=1$	Y_{21}	X_{21}	Z_{21}

Построение корреляционно-регрессионной модели

$$Y = b_0 + b_1 * X_1 + b_2 * X_2 + \dots + b_k X_k + \varepsilon$$

Оценка параметров модели

- Оценка параметра называется *несмещенной*, если ее математическое ожидание равно оцениваемому параметру
- Оценка параметра называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру при возрастании количества наблюдений
- Оценка параметра называется *эффективной*, если она имеет наименьшую дисперсию среди возможных несмещенных оценок параметра, вычисленных по выборкам одного и того же объема n

Требования к исходным данным

- Объясняющие переменные x_1, x_2, \dots, x_k рассматриваются как неслучайные величины
- Величины x_1, x_2, \dots, x_k не связаны между собой линейной функциональной зависимостью

Требования к регрессионным остаткам

- Регрессионные остатки ε_i есть взаимонезависимые случайные величины с нулевым математическим ожиданием
- Регрессионные остатки ε_i имеют постоянную остаточную дисперсию
- Вектор регрессионных остатков подчиняется n -мерному нормальному закону распределения вероятностей

Исходная информация

- $B = (X^T X)^{-1} X^T Y$

- y_1
- y_2
- ...
- y_n

$$\begin{array}{ccc} 1 & X_{11} & X_{12} \\ 1 & X_{21} & X_{22} \\ \dots & \dots & \dots \\ 1 & X_{n1} & X_{n2} \end{array}$$

Используемые функции

- X^T транспонированная матрица.
- Копировать, специальная вставка, выбрать окошко «транспонировать», ввод
- МУМНОЖ
- МОБР
- Shift + Alt + Enter

Оценка значимости уравнения в целом

- $F_{\text{расч}} = \frac{Q_r / (k + 1)}{Q_{\text{ост}} / (n - k - 1)}$

Сравнение расчетного и табличного значения

- $F_{\text{табл}} = F_{\text{РАСПОБР}}$
- Вероятность = вероятности ошибки
- $V_1 = K+1$
- $V_2 = n-k-1$
- $F_{\text{расч}} > F_{\text{табл}}$ Уравнение значимо

Оценка значимости регрессоров

- $T_{bj} = b_j/s_{bj}$
- Ковариационная матрица по b_j
- $$S^*(X^T X)^{-1}$$
 - $S^2 = Q_{ост}/(n-k-1)$
- По диагонали этой матрицы находим S_{bj} в квадрате

Определение значимости регрессоров

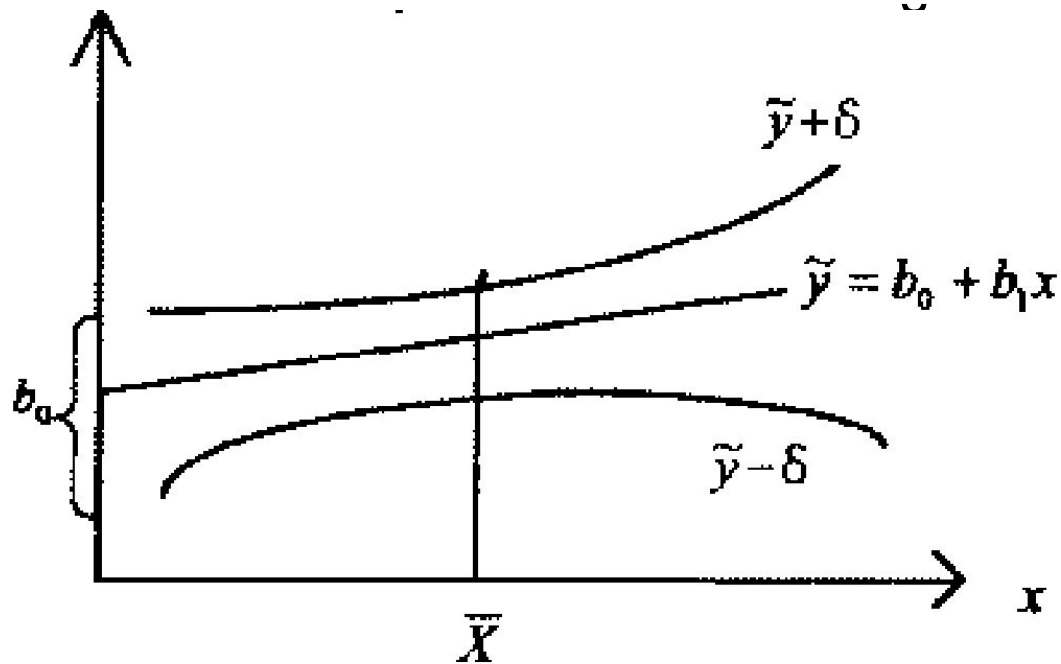
- $t_{\text{табл}} = \text{СТЪЮДРАСПОБР}$
- Вероятность ошибки
- $V_1 = n - k - 1$

- $t_{\text{расч}} > t_{\text{табл}}$ по абсолютной величине,
то регрессор значим

Множественный коэффициент детерминации

$$R_y^2 = 1 - \frac{\sum_{i=1}^n (y_i - y_{mod})^2}{\sum_{i=1}^n (y_i - y_{cp})^2}$$

Точечный и интервальный прогноз



IM

Расчет интервального прогноза для простейшей модели

- $Y_{n+1} \in Y_{n+1} \pm t_T S \sqrt{\frac{1}{n} + \frac{(x_{n+1} - x_{cp})^2}{\sum_{i=1}^n (x_i - x_{cp})^2} + 1}$

Расчет интервального прогноза для множественной модели

- $Y_{n+1} \in Y_{n+1} \pm t_T S \sqrt{(X^0)^T (X^T X)^{-1} X^0 + 1}$