Корреляционно-регрессионные модели

Представление исходных данных

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

Представление исходных данных

- $У_1$ $X_{11}\ X_{12}\ \ldots\ldots\ldots\ldots X_{1k}$
- $У_2$ $X_{21}\ X_{22}\ \ldots\ldots\ldots\ldots X_{2k}$
- … ………… ………
- $У_n$ $X_{n1}\ X_{n2}\ \ldots\ldots\ldots\ldots X_{nk}$

$n > k \sim 3$

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

$$У=в_0+в_1*x_1+в_2*x_2 + _{…} В_k X_k +ε$$

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…..$x_к$ рассматриваются как неслучайные величины
- Величины x_1, x_2…..$x_к$ 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 регресионных остатков подчиняется n-мерному нормальному закону распределения вероятностей

Исходная информация

- $B=(X^{т}X)^{-1}X^{т}У$

- $У_1$ $1\ X_{11}\ X_{12}$
- $У_2$ $1\ X_{21}\ X_{22}$
- … …………
- $У_n$ $1\ X_{n1}\ X_{n2}$

Используемые функции

- X^T транспонированная матрица.
- Копировать, специальная вставка, выбрать окошко «транспонировать», ввод

- МУМНОЖ
- МОБР

- Shift + Alt + Enter

Оценка значимости уравнения в целом

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

Сравнение расчетного и табличного значения

- Fтабл =FРАСПОБР
- Вероятность = вероятности ошибки
- V1 = K+1
- V2 = n-k-1
- Fрасч > Fтабл   Уравнение значимо

Оценка значимости регрессоров

- $Tbj = bj/sbj$

- Ковариационная матрица по bj
- $S^*(X^T X)^{-1}$

  - $S2 = Q_{ост}/(n-k-1)$

- По диагонали этой матрицы находим Sbj в квадрате

Определение значимости регрессоров

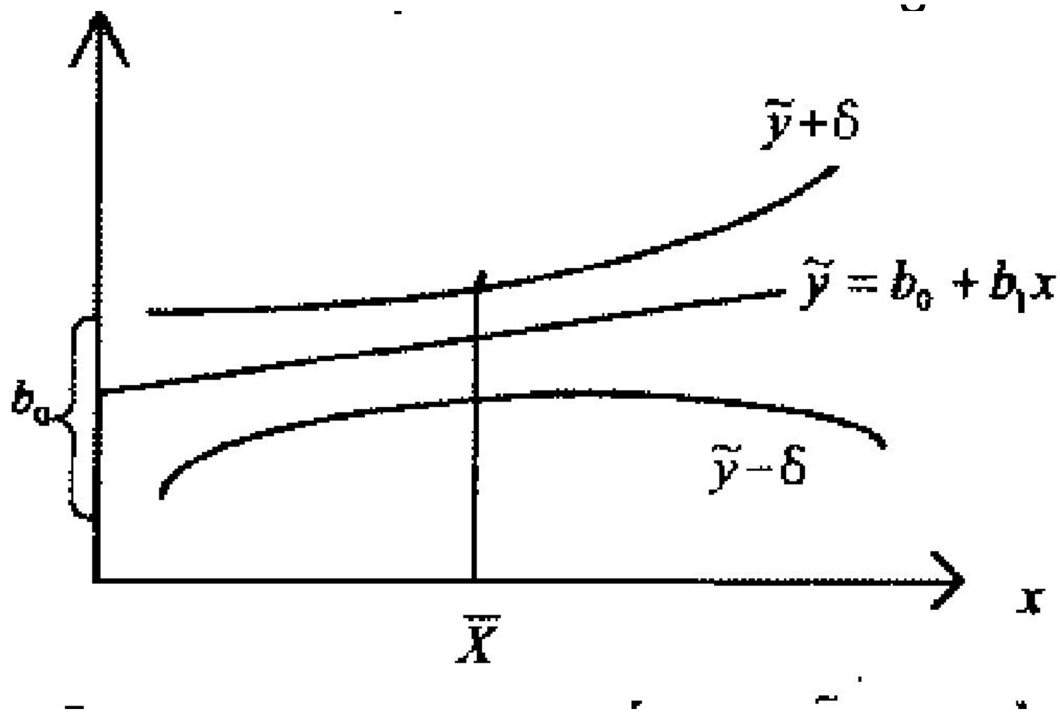
- tтабл = СТЬЮДРАСПОБР
- Вероятность ошибки
- V1=n-k-1

- tрасч > tтабл по абсолютной величине, то регрессор значим

Множественный коэффициент детерминации

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

Точечный и интервальный прогноз



Расчет интервального прогноза для простейшей модели

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

Расчет интервального прогноза для множественной модели

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