ДЗ 4

№1

a) $Y_i = \alpha \exp(\beta X_i) u_i \xrightarrow{\ln} \underbrace{\ln Y_i}_{Y_i'} = \underbrace{\ln \alpha}_{\alpha'} + \beta X_i + \underbrace{\ln u_i}_{\varepsilon_i}$

$Y_i' = \alpha' + \beta X_i + \varepsilon_i$ → линейная модель, можно преобр.

b) $Y_i = \alpha \exp(\beta X_i) + u_i$

Модель аддитивна по ошибке

$\ln Y_i = \ln(\alpha e^{\beta X_i} + u_i)$ – линейная часть не выделяется

c) $Y_i = \exp(\alpha + \beta X_i + u_i)$

$\ln Y_i = \alpha + \beta X_i + u_i$

$Y_i' = \alpha + \beta X_i + u_i$ → линейная модель

d) $Y_i = \frac{\alpha}{1 - X_i} + u_i \xrightarrow[(Z_i = 1 - X_i)]{} Y_i = \alpha Z_i + u_i$

$X_i \neq 1$

линейная модель

e) $Y_i = \frac{1}{1 + \exp(\beta_0 + \beta_1 X_i + u_i)}$ $Y_i \neq 0$

$\frac{1}{Y_i} = 1 + \exp(\beta_0 + \beta_1 X_i + u_i)$

$\frac{1 - Y_i}{Y_i} = \exp(\beta_0 + \beta_1 X_i + u_i)$

$Y_i' = \ln\left(\frac{1 - Y_i}{Y_i}\right) = \beta_0 + \beta_1 X_i + u_i$ линейная модель

Для МНК оценки в случае линейности ∃ аналит. решение $\hat{\beta} = (X^T X)^{-1} \ldots$

№2

Пусть есть $Y = X\beta + \varepsilon$, X – матр. $n \times k$ где 1й столбец это единицы

Для такой модели известно решение: $\hat{\beta} = (X^T X)^{-1} X^T Y$

и $\hat{Y} = X\hat{\beta} = P_x Y$, $P_x = X(X^T X)^{-1} X^T$

и $R^2 = \dfrac{\sum(\bar{Y}_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})}$

тогда как $\widehat{corr}^2(\hat{Y}, Y) = r^2(\hat{Y}, Y) = \dfrac{[\sum(\hat{Y}_i - \bar{\hat{Y}})(Y_i - \bar{Y})]^2}{\sum(\hat{Y}_i - \bar{\hat{Y}})^2 \sum(Y_i - \bar{Y})^2}$

но в X 1й столбец – единицы $\Rightarrow$ если $j = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$ и

$\Rightarrow$ ~~P_x~~ P_x – матрица проекции на простр-во X, то

$P_x j = j$ т.к. $j \in col(X)$ – уже лежит в пр-ве столбцов X

$\left(j = X \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right)$

$j^T \hat{Y} = j^T P_x Y = j^T Y \Rightarrow \sum \hat{Y}_i = \sum Y_i \Rightarrow \bar{\hat{Y}} = \bar{Y}$

$r^2(\hat{Y}, Y) = \dfrac{[\sum(\hat{Y}_i - \bar{Y})(Y_i - \bar{Y})]^2}{\sum(\hat{Y}_i - \bar{Y})^2 \sum(Y_i - \bar{Y})^2}$.

$\sum(\hat{Y}_i - \bar{Y})(Y_i - \bar{Y}) = \sum(\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i + \hat{Y}_i - \bar{Y}) =$

$= \underbrace{\sum(\hat{Y}_i - \bar{Y})(Y_i - \hat{Y}_i)} + \sum(\hat{Y}_i - \bar{Y})^2$

$\underbrace{\sum \hat{Y}_i (Y_i - \hat{Y}_i)}_{\hat{Y}^T (Y - \hat{Y}) = \hat{Y}^T \varepsilon = 0} - \bar{Y} \underbrace{\sum (Y_i - \hat{Y}_i)}_{= 0 \text{ по св-ву МНК остатков}} = 0$

Тогда $r^2(\hat{Y}, Y) = \dfrac{(\sum(\hat{Y}_i - \bar{Y})^2)^2}{\cancel{\sum(\hat{Y}_i - \bar{Y})} \sum(Y_i - \bar{Y})^2} = \dfrac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = R^2$

№3 $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

a) $S(\beta_0, \beta_1) = \sum (Y_i - \beta_0 - \beta_1 X_i)^2$ — квадрат. ошибка, найдём min по β_0, β_1

$$\frac{\partial S}{\partial \beta_0} = -2\sum (Y_i - \beta_0 - \beta_1 X) = 0$$

$$\frac{\partial S}{\partial \beta_1} = -2\sum X_i(Y_i - \beta_0 - \beta_1 X_i) = 0$$

$$\begin{cases} n\beta_0 + \beta_1 \sum X_i = \sum Y_i \\ \beta \sum X_i + \beta_1 \sum X_i^2 = \sum X_i Y_i \end{cases} \Rightarrow \begin{cases} \hat{\beta}_1 = \dfrac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2} \\ \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \end{cases}$$

Теперь матр. форма:

$$X = \begin{pmatrix} 1 & X_1 \\ 1 & X_2 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix} \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$$

$$Y = X\beta + \varepsilon \qquad \hat{\beta} = (X^T X)^{-1} X^T Y$$

$$X^T X = \begin{pmatrix} n & \sum X_i \\ \sum X_i & \sum X_i^2 \end{pmatrix} \qquad X^T Y = \begin{pmatrix} \sum Y_i \\ \sum X_i Y_i \end{pmatrix}$$

$$(X^T X)^{-1} = \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \cdot \begin{pmatrix} \sum X_i^2 & -\sum X_i \\ -\sum X_i & n \end{pmatrix}$$

$$\hat{\beta} = (X^T X)^{-1} X^T Y = \frac{1}{n\sum X_i^2 - (\sum X_i)^2} \begin{pmatrix} \sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i \\ -\sum X_i \sum Y_i + n\sum X_i Y_i \end{pmatrix}$$

$$\sum X_i Y_i - n\bar{X}\bar{Y} = \sum (X_i - \bar{X})(Y_i - \bar{Y})$$

$$\sum X_i^2 - n\bar{X}^2 = \sum (X_i - \bar{X})^2$$

$$\sum X_i = n\bar{X} \qquad \sum Y_i = n\bar{Y}$$

$$\hat{\beta}_0 = \frac{\bar{Y}\sum X_i^2 - \bar{X}\sum X_i Y_i}{\sum X_i^2 - n\bar{X}^2} = \frac{\bar{Y}(\sum X_i^2 - n\bar{X}^2) + n\bar{Y}\bar{X}^2 - \bar{X}\sum X_i Y_i}{\sum X_i^2 - n\bar{X}^2} =$$

$$= \bar{Y} + \frac{n\bar{Y}\bar{X}^2 - \bar{X}\sum X_i Y_i}{\sum X_i^2 - n\bar{X}^2} = \bar{Y} - \frac{\bar{X}\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

Сканировано с CamScanner

$$\hat{\beta}_1 = \frac{n\sum X_iY_i - \sum X_i \sum Y_i}{n\sum X_i^2 - (\sum X_i)^2} = \frac{\sum X_iY_i - n\bar{X}\bar{Y}}{\sum X_i^2 - n\bar{X}^2} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (X_i - \bar{X})^2}$$

совпало

$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$ совпало

b) $P_X = X(X^T X)^{-1} X^T$ надо пок-ть, что $P_X^2 = P_X$ и $P_X^T = P_X$

$$P_X^2 = X(X^T X)^{-1} X^T \cdot X(X^T X)^{-1} X^T = X(X^T X)^{-1}(X^T X)(X^T X)^{-1} X^T =$$

$$= X(X^T X)^{-1} X^T = P_X$$

$$P_X^T = (X(X^T X)^{-1} X^T)^T = X\underbrace{((X^T X)^{-1})^T}_{\text{симметр}} X^T = X(X^T X)^{-1} X^T = P_X$$

$$Q_X = I - P_X$$

$$Q_X^2 = (I - P_X)(I - P_X) = I - 2P_X + P_X^2 = I - P_X = Q_X$$

$$Q_X^T = (I - P_X)^T = I^T - P_X^T = I - P_X = Q$$

№4

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 d_i + \beta_3 X_i d_i + \varepsilon_i$$

При $d_i = 0$: $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

При $d_i = 1$: $Y_i = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) X_i + \varepsilon_i$

Пусть n_0 – число набл с $d_i = 0$, n_1 –//– с $d_i = 1$

$$X = \begin{pmatrix} 1 & X_1 & 0 & 0 \\ 1 & X_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n_0} & 0 & 0 \end{pmatrix}$$

$$X^T X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}, \quad A = \sum \begin{pmatrix} 1 & X_i \\ X_i & X_i^2 \end{pmatrix}$$

$$X^T X = \underbrace{\begin{pmatrix} 1 & 1 & \dots & 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_{n_0} & X_{n_0+1} & \dots & X_n \\ 0 & 0 & \dots & 0 & 1 & \dots & 1 \\ 0 & 0 & \dots & 0 & X_{n_0+1} & \dots & X_n \end{pmatrix}}_{4\times n} \underbrace{\begin{pmatrix} 1 & X_1 & 0 & 0 \\ 1 & X_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n_0} & 0 & 0 \\ 1 & X_{n_0+1} & 1 & X_{n_0+1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_n & 1 & X_n \end{pmatrix}}_{n\times 4} =$$

$$= X_0^T X_0 + X_1^T X_1, \text{ где } X_0 = \begin{pmatrix} 1 & X_1 & 0 & 0 \\ 1 & X_2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_{n_0} & 0 & 0 \end{pmatrix}, \quad X_1 = \begin{pmatrix} 1 & X_{n_0+1} & 1 & X_{n_0+1} \\ 1 & X_{n_0+2} & 1 & X_{n_0+2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & X_n & 1 & X_n \end{pmatrix}$$

$$X_0^T X_0 = \left(\begin{array}{cc|c} n_0 & \sum^{n_0} X_i & \\ \sum^{n_0} X_i & \sum^{n_0} X_i^2 & 0 \\ \hline & 0 & 0 \end{array}\right)$$

$$X_1^T X_1 = \sum_{i=n_0+1}^{n} \begin{pmatrix} 1 & X_i & 1 & X_i \\ X_i & X_i^2 & X_i & X_i^2 \\ 1 & X_i & 1 & X_i \\ X_i & X_i^2 & X_i & X_i^2 \end{pmatrix} = \begin{pmatrix} n_1 & S_x & n_1 & S_x \\ S_x & S_{xx} & S_x & S_{xx} \\ n_1 & S_x & n_1 & S_x \\ S_x & S_{xx} & S_x & S_{xx} \end{pmatrix}$$

$$X^T Y\Big|_{d_i=0} = \sum_{i=1}^{n_0} \begin{pmatrix} Y_i \\ X_i Y_i \\ 0 \\ 0 \end{pmatrix} \qquad X^T Y\Big|_{d_i=1} = \begin{pmatrix} Y_i \\ X_i Y_i \\ Y_i \\ X_i Y_i \end{pmatrix}$$

$$X^T X \hat{\beta} = X^T Y \Rightarrow (X_0^T X_0 + X_1^T X_1)\hat{\beta} = X_0^T Y_0 + X_1^T Y_1$$

тяжело считать/дел-ть в чисто матричном виде

$$\begin{cases} \dfrac{\partial S}{\partial \beta_0} = -2\sum (Y_i - \beta_0 - \beta_1 X_i - \beta_2 d_i - \beta_3 X_i d_i) = 0 & (1) \\ \dfrac{\partial S}{\partial \beta_1} = -2\sum X_i (Y_i - \beta_0 - \beta_1 X_i - \beta_2 d_i - \beta_3 X_i d_i) = 0 & (2) \end{cases}$$

$\cancel{\sum_{i:d_i=0}}$

$$\begin{cases} \sum_{i:d_i=0} (Y_i - \beta_0 - \beta_1 X_i) = 0 & (1) \\ \sum_{i:d_i=0} X_i (Y_i - \beta_0 - \beta_1 X_i) = 0 & (2) \end{cases}$$

при $d_i = 0$

получилась система норм. ур-ий для МНК в подвыб. $d_i = 0$ ←

Сканировано с CamScanner

$$\begin{cases} \sum\limits_{i: d_i=1} (Y_i - \beta_0 - \beta_1 X_i - \beta_2 - \beta_3 X_i) = 0 & (1) \\ & \text{при } d_i = 1 \\ \sum\limits_{i: d_i=1} X_i (Y_i - \beta_0 - \beta_1 X_i - \beta_2 - \beta_3 X_i) = 0 & (2) \end{cases}$$

при $d_i = 1$ слагаемые в ур-ниях содержатся и β_2, β_3, но

они не влияют на решение для β_0, β_1 при $d_i = 0$