Круглов А.И. Б05-204 АВ2

2 CUPED: $y(x) = \theta_0 + \theta_1 x$, x – ковариата, y – таргет. Показать, что $\hat{\theta}_1$ по МНК $\equiv$ коэф.-т CUPED преобр.-я.

В векторном виде: $Y = X\theta$, где $Y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix}$, $X = \begin{pmatrix} x_1 & 1 \\ \dots & \dots \\ x_n & 1 \end{pmatrix}$, $\theta = \begin{pmatrix} \theta_1 \\ \theta_0 \end{pmatrix}$. Оценка $\hat{\theta}$ по МНК:

$$\hat{\theta} = (X^T X)^{-1} X^T Y = \left[\begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix} \right]^{-1} \begin{pmatrix} x_1 & \cdots & x_n \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum_1^n x_i^2 & \sum_1^n x_i \\ \sum_1^n x_i & n \end{pmatrix}^{-1} \cdot \begin{pmatrix} \sum_1^n x_i y_i \\ \sum_1^n y_i \end{pmatrix} = \frac{1}{n\sum_1^n x_i^2 - (\sum_1^n x_i)^2} \cdot$$

$$\cdot \begin{pmatrix} n & -\sum_1^n x_i \\ -\sum_1^n x_i & \sum_1^n x_i^2 \end{pmatrix} \cdot \begin{pmatrix} \sum_1^n x_i y_i \\ \sum_1^n y_i \end{pmatrix} = \frac{1}{n\sum_1^n x_i^2 - (\sum_1^n x_i)^2} \cdot \begin{pmatrix} n\sum_1^n x_i y_i - (\sum_1^n x_i)(\sum_1^n y_i) \\ (\sum_1^n x_i^2)(\sum_1^n y_i) - (\sum_1^n x_i)(\sum_1^n x_i y_i) \end{pmatrix} \Rightarrow \hat{\theta}_1 = \frac{n\sum_1^n x_i y_i - (\sum_1^n x_i)(\sum_1^n y_i)}{n\sum_1^n x_i^2 - (\sum_1^n x_i)^2} \Big| : n^2 =$$

$$= \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i}{n} \cdot \frac{\sum y_i}{n}}{\frac{\sum x_i^2}{n} - \frac{\sum x_i}{n} \cdot \frac{\sum x_i}{n}} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - \bar{x}^2} = \boxed{\frac{\operatorname{cov}(x, y)}{Dx}}$$

, что совпало с коэффициентом из CUPED, ч.т.д.

4 K страт, $k \in \{1, \dots, K\}$, $w_1 + \dots + w_K = 1$, выб. размера n $(P(k) = w_k)$. Посчитать $E\frac{1}{n_k}$ до $o(\frac{1}{n^2})$

Разложим $f(x) = \frac{1}{x}$ в $x_0 \neq 0$ в ряд Тейлора:

$$\frac{1}{x} = \frac{1}{x_0} - \frac{x - x_0}{x_0^2} + \frac{(x - x_0)^2}{x_0^3} + o((x - x_0)^2)$$

(wolfram). Теперь подставим $x = n_k$, $x_0 = En_k$ $\Rightarrow$

$$\frac{1}{n_k} = \frac{1}{En_k} - \frac{n_k - En_k}{(En_k)^2} + \frac{(n_k - En_k)^2}{(En_k)^3} + o((n_k - En_k)^2) \Rightarrow E\frac{1}{n_k} \approx \frac{1}{En_k} - \underbrace{\frac{E(n_k - En_k)}{(En_k)^2}}_{0} + \frac{E(n_k - En_k)^2}{(En_k)^3} = \frac{1}{En_k} + \frac{Dn_k}{(En_k)^3} =$$

$$= \frac{1}{nw_k} + \frac{n \cdot w_k \cdot (1 - w_k)}{(nw_k)^3} \quad (\text{т.к. } n_k \sim Bin(n, w_k)) = \frac{1}{nw_k} + \frac{1 - w_k}{n^2 w_k^2} = \boxed{\frac{nw_k + 1 - w_k}{n^2 w_k^2}}$$

1 2 независ. группы –//– размера, $y(t) = \theta_0 + \theta_1 t$, $t \in \{0, 1\}$ – № группы, y – таргет. Эквивалентность крит.-ю Вальда с HCO?

Гипотеза о незначимости θ_1: $H_0 (\theta_1 = 0)$. Введем в векторном виде: $Y = X\theta$, где

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_{2n} \end{pmatrix}, \quad X = \begin{pmatrix} 1 & 0 \\ \dots & \dots \\ 1 & 0 \\ 1 & 1 \\ \dots & \dots \\ 1 & 1 \end{pmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n \text{ шт} \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} n \text{ шт} \end{matrix}, \quad \theta = \begin{pmatrix} \theta_0 \\ \theta_1 \end{pmatrix}.$$

Статистика $F = \dfrac{\frac{1}{2-1} |X\hat{\theta} - 2n\theta_0|^2}{\frac{1}{2n-2} |Y - X\hat{\theta}|^2} = (2n-2) \dfrac{n^2 \theta_1^2}{|Y - X\hat{\theta}|^2} =$

$$= (2n-2) \frac{n^2 \theta_1^2}{\sum_{i=1}^n (y_i - \theta_0)^2 + \sum_{i=n+1}^{2n} (y_i - \theta_0 - \theta_1)^2} \sim F(1, 2n-2)$$

(отвергаем, если $F > F_{1-\alpha}(1, 2n-2)$)

HCO: $(X^T X)^{-1} (X^T \operatorname{diag}(\hat{\varepsilon}_1^2 \dots \hat{\varepsilon}_{2n}^2) X) (X^T X)^{-1} \overset{(*)}{=} \dfrac{S}{n^2} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \frac{\sum_{i=1}^n y_i - n\theta_0}{S} + 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = \dfrac{S}{n^2} \begin{pmatrix} \frac{\Sigma - n\theta_0}{S} & -\frac{\Sigma - n\theta_0}{S} \\ -\frac{\Sigma - n\theta_0}{S} & \frac{\Sigma - n\theta_0}{S} + 1 \end{pmatrix}$

$(*)$: $(X^T X)^{-1} = \left[\begin{pmatrix} 1 \cdots 1 & 1 \cdots 1 \\ 0 \cdots 0 & 1 \cdots 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dots & \dots \\ 1 & 1 \end{pmatrix} \right]^{-1} = \begin{bmatrix} 2n & n \\ n & n \end{bmatrix}^{-1} = \dfrac{1}{2n^2 - n^2} \begin{pmatrix} n & -n \\ -n & 2n \end{pmatrix} = \dfrac{1}{n} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$

?

$$X^T \operatorname{diag}(\dots) X = \begin{pmatrix} 1 \cdots 1 & 1 \cdots 1 \\ 0 \cdots 0 & 1 \cdots 1 \end{pmatrix} \begin{pmatrix} y_1 - \theta_0 & & & & \\ & \ddots & & & \\ & & y_n - \theta_0 & & \\ & & & y_{n+1} - \theta_0 - \theta_1 & \\ & & & & \ddots \\ & & & & & y_{2n} - \theta_0 - \theta_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dots & \dots \\ 1 & 0 \\ 1 & 1 \\ \dots & \dots \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} y_1 - \theta_0 & \cdots & y_n - \theta_0 & y_{n+1} - \theta_0 - \theta_1 \cdots y_{2n} - \theta_0 - \theta_1 \\ 0 & \cdots & 0 & y_{n+1} - \theta_0 - \theta_1 \cdots y_{2n} - \theta_0 - \theta_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \dots & \dots \\ 1 & 0 \\ 1 & 1 \\ \dots & \dots \\ 1 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=1}^{2n} y_i - 2n\theta_0 - n\theta_1 & \sum_{i=n+1}^{2n} y_i - n\theta_0 - n\theta_1 \\ \sum_{i=n+1}^{2n} y_i - n\theta_0 - n\theta_1 & \sum_{i=n+1}^{2n} y_i - n\theta_0 - n\theta_1 \end{pmatrix} = \Big/ \sum_{i=n+1}^{2n} y_i - n\theta_0 - n\theta_1 =: S \Big/ = \begin{pmatrix} \sum_{i=1}^n y_i - n\theta_0 + S & S \\ S & S \end{pmatrix} = S \begin{pmatrix} \frac{\sum_{i=1}^n y_i - n\theta_0}{S} + 1 & 1 \\ 1 & 1 \end{pmatrix}$$