ДЗ 5

№1 $f(a) = \mathbb{E}|X-a|$

Предположим, $f_\alpha(X,\alpha) = (X-a)(\alpha - \mathbb{1}_{X<\alpha})$

т.е. $f_\alpha(X,\alpha) = \begin{cases} \alpha(X-\alpha) & X \geq \alpha \\ (1-\alpha)(\alpha - X) & X < \alpha \end{cases}$

Тогда $L(a) = \mathbb{E}[f_\alpha(X,\alpha)] = \alpha \int\limits_a^{+\infty} (x-a)\,dF(x) + (1-\alpha)\int\limits_{-\infty}^{a} (a-x)\,dF(x)$

$\frac{dL}{da} = -\alpha \int\limits_a^{+\infty} dF(x) + (1-\alpha)\int\limits_{-\infty}^{a} dF(x) = -\alpha(1-F(a)) + (1-\alpha)F(a) =$

$= -\alpha + \alpha F(a) + F(a) - \alpha F(a) = F(a) - \alpha = 0 \Rightarrow F(a) = \alpha$

то есть $a = q_\alpha$ – α-квантиль по определению

$\frac{d^2L}{da^2} = f(a) > 0 \Rightarrow L(a) = \mathbb{E}[f_\alpha(X,\alpha)]$ минимизируется в $a = q_\alpha$

Сканировано с CamScanner

№2

$\mathbb{E}(Y_i|X_i) = \beta X_i$, $\hat{\beta} = \frac{\sum Y_i X_i^3}{\sum X_i^4}$

$\hat{\beta} = \frac{\frac{1}{n}\sum Y_i X_i^3}{\frac{1}{n}\sum X_i^4}$ по ЗБЧ: $\frac{1}{n}\sum X_i^4 \xrightarrow{\mathbb{P}} \mathbb{E}[X_i^4]$

$\frac{1}{n}\sum Y_i X_i^3 \xrightarrow{\mathbb{P}} \mathbb{E}[Y_i X_i^3]$

$\mathbb{E}[Y_i X_i^3] = \mathbb{E}[\mathbb{E}[Y_i X_i^3 | X_i]] = \mathbb{E}[X_i^3 \cdot \mathbb{E}[Y_i|X_i]] =$
$= \mathbb{E}[X_i^3 \cdot \beta X_i] = \beta \mathbb{E}[X_i^4]$

Тогда $\hat{\beta} \xrightarrow{\mathbb{P}} \frac{\beta \mathbb{E}[X_i^4]}{\mathbb{E}[X_i^4]} = \beta$ => оценка состоятельна

предел по вер-сти

Пусть $\varepsilon_i = Y_i - \beta X_i$ — ошибка регрессии, $\mathbb{E}[\varepsilon_i | X_i] = 0$

$\hat{\beta} - \beta = \frac{\sum Y_i X_i^3 - \beta \sum X_i^4}{\sum X_i^4} = \frac{\sum (Y_i - \beta X_i) X_i^3}{\sum X_i^4} = \frac{\sum \varepsilon_i X_i^3}{\sum X_i^4}$

$\sqrt{n}(\hat{\beta} - \beta) = \frac{\frac{1}{\sqrt{n}}\sum \varepsilon_i X_i^3}{\frac{1}{n}\sum X_i^4}$, $\frac{1}{n}\sum X_i^4 \xrightarrow[\text{по ЗБЧ}]{\mathbb{P}} \mathbb{E}[X_i^4]$

$\frac{1}{\sqrt{n}}\sum \varepsilon_i X_i^3 \to ?$ $\mathbb{E}[\varepsilon_i X_i^3 | X_i] = X_i^3 \mathbb{E}[\varepsilon_i | X_i] = 0$

$\Downarrow$

$\mathbb{E}[\varepsilon_i X_i^3] = 0$

$\mathbb{D}(\varepsilon_i X_i^3) = \mathbb{E}[\varepsilon_i^2 X_i^6] - (\underbrace{\mathbb{E}[\varepsilon_i X_i^3]}_{0})^2 = \mathbb{E}[\mathbb{E}[\varepsilon_i^2 | X_i] \cdot X_i^6]$

Если $\sigma^2(X_i) = \mathbb{E}[\varepsilon_i^2 | X_i] = \mathbb{D}[Y_i | X_i]$ то $\mathbb{D}[\varepsilon_i X_i^3] =$
$= \sigma^2 \mathbb{E}[X_i^6]$

и по ЦПТ $\frac{1}{\sqrt{n}}\sum \varepsilon_i X_i^3 \xrightarrow{d} \mathcal{N}(0, \mathbb{E}[\varepsilon_i^2 X_i^6])$

и по т. Слуцкого: $\sqrt{n}(\hat{\beta} - \beta) \to \mathcal{N}\left(0, \frac{\mathbb{E}[\varepsilon_i^2 X_i^6]}{(\mathbb{E}[X_i^4])^2}\right)$

Сканировано с CamScanner

№3

$\mathbb{E}(Y_i | X_i) = \alpha + \beta X_i \qquad X'_i = X_i + \xi_i$

$Y_i = \tilde{\alpha} + \tilde{\beta} X'_i + u_i$

$$\hat{\beta}_{МНК} = \frac{\sum (Y_i - \bar{Y})(X'_i - \bar{X}')}{\sum (X'_i - \bar{X}'_{\xi})^2}$$

Пусть $\varepsilon_i = Y_i - \alpha - \beta X_i$ — истинная ошибка регрессии с $\mathbb{E}(\varepsilon_i | X_i) = 0$

числитель $\hat{\beta}_{МНК}$:

$\frac{1}{n} \sum (Y_i - \bar{Y})(X'_i - \bar{X}') = \frac{1}{n} \sum (\beta X_i + \varepsilon_i - \beta \bar{X} - \bar{\varepsilon})(X_i + \xi_i - \bar{X}')$

$= \frac{1}{n} \sum (\beta (X_i - \bar{X}) + (\varepsilon_i - \bar{\varepsilon}))((X_i - \bar{X}) + (\xi_i - \bar{\xi})) =$

$= \beta \overbrace{\frac{1}{n} \sum (X_i - \bar{X})^2}^{\xrightarrow{P} \mathbb{D}(X_i) = \sigma_x^2} + \beta \frac{1}{n} \cancel{\sum (X_i - \bar{X})(\xi_i - \bar{\xi})}^{\to 0} +$

$+ \frac{1}{n} \cancel{\sum (\varepsilon_i - \bar{\varepsilon})(X_i - \bar{X})}^{\to 0} + \frac{1}{n} \cancel{\sum (\varepsilon_i - \bar{\varepsilon})(\xi_i - \bar{\xi})}^{\to 0} =$

$= \beta \frac{1}{n} \sum (X_i - \bar{X})^2 \xrightarrow[ЗБЧ]{P} \beta \sigma_x^2$

знаменатель $\hat{\beta}_{МНК}$:

$\frac{1}{n} \sum (X'_i - \bar{X}')^2 = \frac{1}{n} \sum ((X_i - \bar{X}) + (\xi_i - \bar{\xi}))^2 =$

$= \frac{1}{n} \sum (X_i - \bar{X})^2 + 2 \frac{1}{n} \cancel{\sum (X_i - \bar{X})(\xi_i - \bar{\xi})}^{\to 0} + \frac{1}{n} \sum (\xi_i - \bar{\xi})^2 =$

$= \frac{1}{n} \sum (X_i - \bar{X})^2 + \frac{1}{n} \sum (\xi_i - \bar{\xi})^2 \xrightarrow{P} \sigma_x^2 + \sigma_\xi^2$

По т. Слуцкого:

$$\hat{\beta}_{МНК} \xrightarrow{P} \frac{\beta \sigma_x^2}{\sigma_x^2 + \sigma_\xi^2} = \beta \underbrace{\frac{\sigma_x^2}{\sigma_x^2 + \sigma_\xi^2}}_{<1} < \beta$$

т.е. оценка смещена $\Rightarrow$ не состоятельна

№4

$y_i = \alpha x_i + \beta w_i + e_i$

$\frac{1}{n}\sum X_i X_i^T \xrightarrow{P} \mathbb{E}[X_i X_i^T] = Q$

$\frac{1}{\sqrt{n}}\sum X_i e_i \xrightarrow{d} \mathcal{N}(0, \sigma_e^2 Q)$

$X_i = (x_i, w_i)^T$

$$Q = \mathbb{E}\left[\begin{pmatrix} x_i \\ w_i \end{pmatrix}(x_i \; w_i)\right] = \begin{pmatrix} q_{xx} & q_{xw} \\ q_{xw} & q_{ww} \end{pmatrix}$$

$$Q^{-1} = \frac{1}{q_{xx}q_{ww} - q_{xw}^2}\begin{pmatrix} q_{ww} & -q_{xw} \\ -q_{xw} & q_{xx} \end{pmatrix}$$

Полная регрессия (коэф. при w_i значим):

$$\begin{pmatrix} \hat{\alpha}_f \\ \hat{\beta}_f \end{pmatrix} = \left(\sum_{i=1}^{n} X_i X_i^T\right)^{-1} \sum X_i y_i \quad , \; y_i = \alpha x_i + \beta w_i + e_i$$

$$\begin{pmatrix} \hat{\alpha}_f - \alpha \\ \hat{\beta}_f - \beta \end{pmatrix} = \left(\frac{1}{n}\sum X_i X_i^T\right)^{-1}\left(\frac{1}{n}\sum X_i e_i\right)$$

$$\sqrt{n}\begin{pmatrix} \hat{\alpha}_f - \alpha \\ \hat{\beta}_f - \beta \end{pmatrix} = \left(\frac{1}{n}\sum X_i X_i^T\right)\left(\frac{1}{\sqrt{n}}\sum X_i e_i\right) \xrightarrow{d} \mathcal{N}(0, \sigma_e^2 Q^{-1})$$

Простая регрессия: $y_i = \hat{\alpha} x_i + u_i$

$$\hat{\alpha}_s = \frac{\sum x_i y_i}{\sum x_i^2} = \alpha + \beta\frac{\sum x_i w_i}{\sum x_i^2} + \frac{\sum x_i e_i}{\sum x_i^2}$$

$$\sqrt{n}(\hat{\alpha}_s - \alpha) = \beta\sqrt{n}\,\frac{\frac{1}{n}\sum x_i w_i}{\frac{1}{n}\sum x_i^2} + \frac{\frac{1}{\sqrt{n}}\sum x_i e_i}{\frac{1}{n}\sum x_i^2} \quad , \text{но } \beta = \beta_n = \frac{b}{\sqrt{n}}$$

$$\sqrt{n}(\hat{\alpha}_s - \alpha) = b\,\frac{\frac{1}{n}\sum x_i w_i}{\frac{1}{n}\sum x_i^2} + \frac{\frac{1}{\sqrt{n}}\sum x_i e_i}{\frac{1}{n}\sum x_i^2} \xrightarrow{d} b\cdot\frac{q_{xw}}{q_{xx}} + \frac{z_x}{q_{xx}}$$

$$\sqrt{n}(\hat{\alpha}_s - \alpha) \xrightarrow{d} \mathcal{N}\left(b\cdot\frac{q_{xw}}{q_{xx}}, \frac{\sigma_e^2}{q_{xx}}\right)$$

Используем гипотезы для проверки значимости β:

$H_0: \beta = 0$ в полной регрессии

$$T_n = \frac{\hat{\beta}_f}{se(\hat{\beta}_f)} = \frac{\sqrt{n}\,\hat{\beta}_f}{\sqrt{\sigma_e^2 (Q^{-1})_{22}}} \approx \frac{\sqrt{n}\,\beta_n + Z_\beta}{\sqrt{\sigma_e^2 (Q^{-1})_{22}}} \quad \text{где } Z_\beta \sim N(0, \sigma_e^2 (Q^{-1})_{22})$$

$$T_n \Big|_{\beta_n = \frac{b}{\sqrt{n}}} \xrightarrow{d} \frac{b + Z_\beta}{\sqrt{\sigma_e^2 (Q^{-1})_{22}}} = \frac{b}{\sigma_e \sqrt{(Q^{-1})_{22}}} + Z \quad \text{где } Z \sim N(0,1)$$

Тогда при $|T_n| > c$ (крит. знач., константа) то $\hat{\alpha}_f$, иначе $\hat{\alpha}_s$

то есть
$$\sqrt{n}(\hat{\alpha} - \alpha) = \begin{cases} \sqrt{n}(\hat{\alpha}_f - \alpha), & |T_n| > c \\ \sqrt{n}(\hat{\alpha}_s - \alpha), & T_n \leq c \end{cases}$$

Если $p = P\left(\left|Z + \frac{b}{\sigma_e \sqrt{(Q^{-1})_{22}}}\right| > c\right)$ – асимпт. вер-сть отклонения H_0, то

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{d} \begin{cases} N(0, \sigma_e^2 (Q^{-1})_{11}), & p \\ N\left(b \cdot \frac{q_{xw}}{q_{xx}}, \frac{\sigma_e^2}{q_{xx}}\right), & 1-p \end{cases}$$