Решение семинара 2.

Задание 1.

X_i, Y_i – исходные показатели

$x_i = X_i - \bar{X}$, $y_i = Y_i - \bar{Y}$ – центрированные

$$k_i = \frac{x_i}{\sum_{j=1}^{n} x_j^2}$$

a) $$\sum_{i=1}^{n} k_i = \sum_i \frac{x_i}{\sum_j x_j^2} = \frac{1}{\sum_j x_j^2} \sum_i x_i = \frac{1}{\sum_j x_j^2} \underbrace{\sum_i (X_i - \bar{X})}_{=0} = 0.$$

в) $$\sum_i k_i x_i = \sum k_i X_i = 1$$

$$\sum_i k_i x_i = \sum_i \frac{x_i \cdot x_i}{\sum_j x_j^2} = \frac{1}{\sum_j x_j^2} \sum_i (X_i - \bar{X})(X_i - \bar{X}) = \frac{1}{\sum_j x_j^2} \sum_i (X_i - \bar{X}) X_i =$$

$$= \sum_i \frac{x_i \cdot X_i}{\sum_j x_j^2} = \sum_i k_i \cdot X_i$$

$$\sum_i k_i x_i = \sum_i \frac{x_i^2}{\sum_j x_j^2} = \frac{1}{\sum_j x_j^2} \sum_i x_i^2 = 1.$$

c) $$\sum_i k_i^2 = \sum_i \left(\frac{x_i}{\sum_j x_j^2} \right)^2 = \frac{1}{(\sum_j x_j^2)^2} \sum_i x_i^2 = \frac{1}{\sum_i x_i^2}$$

d) $$\sum_i k_i y_i = \sum_i \frac{x_i y_i}{\sum_j x_j^2} = \sum_i \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_j x_j^2} = \frac{\sum (X_i - \bar{X}) Y_i}{\sum_j x_j^2} = \sum_i \frac{x_i \cdot Y_i}{\sum_j x_j^2} =$$

$$= \sum_i k_i Y_i$$

Задание 2.

$y_i = \beta_1 x_i + \varepsilon_i$

а) МНК оценка:

$$RSS = \sum_{i=1}^{n} e_i^2 \to \min_{\hat{\beta}_1}$$

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_1 x_i$$

$$RSS = \sum_i (y_i - \hat{\beta}_1 x_i)^2 \to \min_{\hat{\beta}_1}$$

$$\frac{\partial RSS}{\partial \hat{\beta}_1} = 2 \sum_i (y_i - \hat{\beta}_1 x_i) \cdot (-x_i) = 0$$

$$\sum_i (y_i x_i - \hat{\beta}_1 x_i^2) = 0$$

$$\sum_i y_i x_i = \hat{\beta}_1 \sum_i x_i^2$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}.$$

в) $$Var(\hat{\beta}_1) = Var\left(\frac{\sum_i y_i x_i}{\sum_i x_i^2}\right) = \left\{\begin{matrix} x_i \text{ - детерм.} \\ y_i \text{ - незав. с.в.} \end{matrix}\right\} =$$

$$= \frac{1}{\left(\sum_i x_i^2\right)^2} \sum_i x_i^2 \underbrace{Var(y_i)}_{Var(\varepsilon_i) = \sigma_\varepsilon^2} = \frac{\sigma_\varepsilon^2}{\left(\sum_i x_i^2\right)^{\cancel{2}}} \cdot \cancel{\sum_i x_i^2} = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n} x_i^2}.$$

Задание 3.

$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$

МНК оценки: $\hat{\beta}_1 = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2} = \frac{\sum_i x_i y_i}{\sum_i x_i^2} = \sum_i k_i y_i = \sum k_i Y_i$

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

a) $Var(\hat{\beta}_1) = Var(\sum_i k_i Y_i) = \left(\begin{array}{l} Y_i \text{ - незав.} \\ Var(Y_i) = Var(\varepsilon_i) = \sigma_\varepsilon^2, \ i = \overline{1,n} \end{array} \right) =$

$= \sum_i k_i^2 \cdot Var(Y_i) = \sum_i k_i^2 \cdot \sigma_\varepsilon^2 = \sigma_\varepsilon^2 \sum_i k_i^2 = \frac{\sigma_\varepsilon^2}{\sum_i x_i^2} = \frac{\sigma_\varepsilon^2}{\sum_i (X_i - \bar{X})^2}.$

$Var(\hat{\beta}_0) = Var(\bar{Y} - \hat{\beta}_1 \bar{X}) = Var(\bar{Y} - \sum_i k_i Y_i \bar{X}) = Var(\frac{1}{n}\sum_i Y_i - \bar{X}\sum_i k_i Y_i) =$

$= Var(\sum_i \frac{1}{n} Y_i - \bar{X}\sum k_i Y_i) = Var(\sum_i (\frac{1}{n} - \bar{X} k_i) Y_i) =$

$= \sum_i (\frac{1}{n} - \bar{X} k_i)^2 \underbrace{Var(Y_i)}_{= \sigma_\varepsilon^2} = \sigma_\varepsilon^2 \underbrace{\sum_i (\frac{1}{n} - \bar{X} k_i)^2}_{*} =$

$* \quad \sum_i (\frac{1}{n} - \bar{X} \cdot k_i)^2 = \sum_i (\frac{1}{n^2} - \frac{2}{n}\bar{X} k_i + (\bar{X})^2 k_i^2) =$

$= \frac{1}{n} - \frac{2}{n}\bar{X} \underbrace{\sum_i k_i}_{= 0} + (\bar{X})^2 \sum_i k_i^2 = \frac{1}{n} + \frac{(\bar{X})^2}{\sum_i x_i^2} = \frac{\overbrace{\sum x_i^2 + n(\bar{X})^2}^{**}}{n \sum_i x_i^2} =$

$** \quad \sum_i x_i^2 = \sum_i (X_i - \bar{X})^2 = \sum_i (X_i^2 - 2\bar{X} X_i + (\bar{X})^2) = \sum_i X_i^2 - 2\bar{X} \underbrace{\sum_i X_i}_{= n\bar{X}} + n(\bar{X})^2 =$

$= \sum_i X_i^2 - n(\bar{X})^2 \Rightarrow \sum_i x_i^2 + n(\bar{X})^2 = \sum_i X_i^2$

$= \frac{\sum X_i^2}{n \sum x_i^2}$

$= \frac{\sigma_\varepsilon^2 \sum X_i^2}{n \sum x_i^2} = \frac{\sigma_\varepsilon^2 \sum_i X_i^2}{n \sum_i (X_i - \bar{X})^2}.$

в) Естественно, $\bar{Y}$ – случайная величина!
и $\mathbb{E}(\bar{Y}) \neq \bar{Y}$

А теперь верное решение:

$$\mathrm{cov}(\hat{\beta}_0, \hat{\beta}_1) = \mathrm{cov}(\bar{Y} - \hat{\beta}_1 \bar{X}, \hat{\beta}_1) = \mathrm{cov}(\bar{Y}, \hat{\beta}_1) - \bar{X} \cdot \mathrm{cov}(\hat{\beta}_1, \hat{\beta}_1)$$

$$= \underbrace{\mathrm{cov}(\bar{Y}, \hat{\beta}_1)}_{*} - \bar{X} \cdot \mathrm{Var}(\hat{\beta}_1) \boxed{=}$$

$$* \quad \mathrm{cov}(\bar{Y}, \hat{\beta}_1) = \mathrm{cov}\left(\frac{1}{n}\sum_i Y_i, \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_i (X_i - \bar{X})^2}\right) =$$

$$= \frac{1}{n\sum_i (X_i - \bar{X})^2} \mathrm{cov}\left(\sum_i Y_i, \sum_i (X_i - \bar{X}) Y_i\right) =$$

$$= \frac{1}{n\sum_i (X_i - \bar{X})^2} \cdot \mathrm{cov}(Y_1 + \ldots + Y_n, (X_1 - \bar{X}) Y_1 + \ldots + (X_n - \bar{X}) Y_n) =$$

$$= \frac{1}{n\sum_i (X_i - \bar{X})^2} \sum_i (X_i - \bar{X}) \underbrace{\mathrm{cov}(Y_i, Y_i)}_{\mathrm{cov}(\varepsilon_i, \varepsilon_i) = \mathbb{E}(\varepsilon_i \varepsilon_i) = V(\varepsilon_i) = \sigma_\varepsilon^2} =$$

$$= \frac{\sigma_\varepsilon^2}{n\sum_i (X_i - \bar{X})^2} \cdot \underbrace{\sum_i (X_i - \bar{X})}_{= 0} = 0.$$

$$\boxed{=} \; 0 - \bar{X} \cdot \mathrm{Var}(\hat{\beta}_1) = -\bar{X} \cdot \frac{\sigma_\varepsilon^2}{\sum_i x_i^2} = -\frac{\bar{X} \cdot \sigma_\varepsilon^2}{\sum_{i=1}^{n} (X_i - \bar{X})^2}.$$

"Оценка дисперсии случайного члена ε."

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i$$

$$\overline{Y} = \beta_1 + \beta_2 \overline{X} + \overline{\varepsilon}$$

$$Y_i - \overline{Y} = \beta_2 (X_i - \overline{X}) + (\varepsilon_i - \overline{\varepsilon})$$

$$y_i = \beta_2 x_i + (\varepsilon_i - \overline{\varepsilon})$$

$$e_i = y_i - \hat{y}_i = \beta_2 x_i + (\varepsilon_i - \overline{\varepsilon}) - \hat{\beta}_2 x_i = (\beta_2 - \hat{\beta}_2) x_i + (\varepsilon_i - \overline{\varepsilon})$$

$$\sum_i e_i^2 = \sum_i \left[(\beta_2 - \hat{\beta}_2)^2 x_i^2 + 2(\beta_2 - \hat{\beta}_2) x_i (\varepsilon_i - \overline{\varepsilon}) + (\varepsilon_i - \overline{\varepsilon})^2 \right] =$$

$$= (\beta_2 - \hat{\beta}_2)^2 \sum_i x_i^2 + 2(\beta_2 - \hat{\beta}_2) \sum x_i (\varepsilon_i - \overline{\varepsilon}) +$$

$$+ \sum_i (\varepsilon_i - \overline{\varepsilon})^2$$

$$\mathbb{E}\left(\sum_i e_i^2\right) = \sum_i x_i^2 \underbrace{\mathbb{E}(\beta_2 - \hat{\beta}_2)^2}_{Var(\hat{\beta}_2)} + \underbrace{2 \sum_i x_i \mathbb{E}\left[(\beta_2 - \hat{\beta}_2) \cdot (\varepsilon_i - \overline{\varepsilon})\right]}_{I} + \underbrace{\sum_i \mathbb{E}(\varepsilon_i - \overline{\varepsilon})^2}_{II} =$$

$$\beta_2 = \mathbb{E}(\hat{\beta}_2)$$

$$\text{I. } 2\,\mathbb{E}\sum_i(\beta_2-\hat\beta_2)x_i(\varepsilon_i-\bar\varepsilon)=$$

$$=-2\,\mathbb{E}\sum_i(\hat\beta_2-\beta_2)x_i(\varepsilon_i-\bar\varepsilon)=$$

$$=-2\,\mathbb{E}\sum_i\Big[\Big(\beta_2+\sum_i k_i\varepsilon_i-\beta_2\Big)x_i(\varepsilon_i-\bar\varepsilon)\Big]=$$

$$=-2\,\mathbb{E}\sum_i\Big[\Big(\sum_i k_i\varepsilon_i\Big)\cdot x_i(\varepsilon_i-\bar\varepsilon)\Big]=$$

$$=-2\overbrace{\sum_i\mathbb{E}\Big[\Big(\sum_i k_i\varepsilon_i\Big)\cdot x_i\varepsilon_i\Big]}^{(*)}+2\overbrace{\sum_i\mathbb{E}\Big[\Big(\sum_i k_i\varepsilon_i\Big)x_i\bar\varepsilon\Big]}^{(**)}=$$

$$=-2\sum_i\mathbb{E}(k_i\varepsilon_i^2x_i)+2\sum_i\mathbb{E}\Big(k_i\frac{\varepsilon_i^2}{n}x_i\Big)=$$

$$=-2\sum_i k_ix_i\underbrace{\mathbb{E}(\varepsilon_i^2)}_{=\sigma_\varepsilon^2}=-2\sigma_\varepsilon^2\underbrace{\sum k_ix_i}_{=1}=-2\sigma_\varepsilon^2.$$

$$(*)\ \mathbb{E}\sum_i\Big[\Big(\sum_i k_i\varepsilon_i\Big)x_i\varepsilon_i\Big]=\mathbb{E}\Big((k_1\varepsilon_1+k_2\varepsilon_2+\ldots+k_n\varepsilon_n)x_1\varepsilon_1+\ldots+$$

$$+(k_1\varepsilon_1+\ldots+k_n\varepsilon_n)x_n\varepsilon_n\Big)=\Big|\,\mathbb{E}(\varepsilon_i\varepsilon_j)=0,\ i\neq j\,\Big|=\sum_i\mathbb{E}(k_ix_i\varepsilon_i^2)=\sigma_\varepsilon^2$$

$$(**)\ \mathbb{E}\sum_i\Big[\Big(\sum k_i\varepsilon_i\Big)x_i\bar\varepsilon\Big]=\mathbb{E}\sum_i\Big[\Big(\sum_i k_i\varepsilon_i\Big)x_i\Big(\frac{\varepsilon_1}{n}+\ldots+\frac{\varepsilon_n}{n}\Big)\Big]=$$

$$=\mathbb{E}\Big((k_1\varepsilon_1+\ldots+k_n\varepsilon_n)x_1\Big(\frac{\varepsilon_1}{n}+\ldots+\frac{\varepsilon_n}{n}\Big)+\ldots+(k_1\varepsilon_1+\ldots+k_n\varepsilon_n)x_n\cdot$$

$$\cdot\Big(\frac{\varepsilon_1}{n}+\ldots+\frac{\varepsilon_n}{n}\Big)\Big)=\mathbb{E}\Big(\frac{k_1\varepsilon_1^2}{n}x_1+\ldots+\frac{k_n\varepsilon_n^2}{n}x_1+\ldots+\frac{k_1\varepsilon_1^2}{n}\cdot x_n+\ldots+\frac{k_n\varepsilon_n^2x_n}{n}\Big)=$$

$$=\frac{\sigma_\varepsilon^2}{n}\cdot x_1\underbrace{\sum_i k_i}_{=0}+\ldots+\frac{\sigma_\varepsilon^2}{n}x_n\cdot\underbrace{\sum_i k_i}_{=0}=0$$

II. $$\sum_i \mathbb{E}(\xi_i-\bar{\xi})^2 = \mathbb{E}\sum_i\left(\xi_i^2 - 2\xi_i\bar{\xi} + (\bar{\xi})^2\right) =$$

$$= \mathbb{E}\left(\sum_i \xi_i^2 - 2\bar{\xi}\overbrace{\sum_i \xi_i}^{=n\cdot\bar{\xi}} + \sum_i(\bar{\xi})^2\right) =$$

$$= \mathbb{E}\left(\sum_i \xi_i^2 - 2\bar{\xi}\cdot n\bar{\xi} + n(\bar{\xi})^2\right) =$$

$$= \mathbb{E}\left(\sum_i \xi_i^2 - 2n(\bar{\xi})^2 + n(\bar{\xi})^2\right) =$$

$$= \mathbb{E}\left(\sum_i \xi_i^2 - n(\bar{\xi})^2\right) = \sum_i \underbrace{\mathbb{E}(\xi_i^2)}_{=\sigma_\xi^2} - n\,\mathbb{E}(\bar{\xi}^2) =$$

$$= n\cdot\sigma_\xi^2 - n\cdot\mathbb{E}\left(\frac{\sum_i \xi_i}{n}\right)^2 =$$

$$= n\cdot\sigma_\xi^2 - \frac{n}{n^2}\mathbb{E}\left(\sum_i \xi_i\right)^2 = n\cdot\sigma_\xi^2 - \frac{1}{n}\mathbb{E}\left(\xi_1^2+\xi_2^2+\ldots+\right.$$

$$\left.+\xi_n^2 + 2\xi_1\xi_2 + \ldots\right) = n\cdot\sigma_\xi^2 - \frac{1}{n}\underbrace{\sum_i \mathbb{E}(\xi_i^2)}_{\text{т.к. } \mathbb{E}(\xi_i\xi_j)=0,\ i\neq j} =$$

$$= n\cdot\sigma_\xi^2 - \frac{1}{n}\cdot n\cdot\sigma_\xi^2 = n\cdot\sigma_\xi^2 - \sigma_\xi^2 = (n-1)\sigma_\xi^2.$$

$\Rightarrow$ ③ $Var(\hat{\beta}_2) \cdot \sum_i x_i^2 - 2\sigma_\varepsilon^2 + (n-1)\sigma_\varepsilon^2 =$

$$= \frac{\sigma_\varepsilon^2}{\sum_i x_i^2} \cdot \sum_i x_i^2 - 2\sigma_\varepsilon^2 + (n-1)\sigma_\varepsilon^2 =$$

$$= \sigma_\varepsilon^2 (1 - 2 + n - 1) = (n-2)\sigma_\varepsilon^2$$

Получили, что

$$\mathbb{E}\left(\sum_i e_i^2\right) = (n-2)\sigma_\varepsilon^2$$

$$\Rightarrow \hat{\sigma}_\varepsilon^2 = \frac{\sum_i e_i^2}{n-2}$$ — оценка дисперсии случайного члена ε.

число наблюдений — число параметров "β" в модели

④ $\mathbb{E}(\hat{\sigma}_\varepsilon^2) = \mathbb{E}\left(\frac{\sum_i e_i^2}{n-2}\right) = \frac{1}{n-2} \cdot (n-2)\sigma_\varepsilon^2 = \sigma_\varepsilon^2$ —

<u>несмещенная</u>!