Конспект по метрике к КР-1

*Посвящается Зехову Матвею, лучшему ассистенту и
просто прекрасному и отзывчивому человеку, в честь 21-летия*

1) Парная регрессия: $\mathbb{E}(Y \mid X = x) = \beta_0 + \beta_1 x$

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \begin{matrix} Y_i - \text{зависимая;} \\ X_i - \text{регрессор;} \\ \varepsilon_i - \text{навязка} \end{matrix}$$

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \underset{\text{остаток}}{e_i} \qquad\qquad RSS = \sum_1^n e_i^2 = \sum_1^n (Y_i - \hat{Y}_i)^2$$

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

2) МНК: $RSS \to min$

$$FOC : \begin{cases} n\hat{\beta}_0 + \sum_1^n X_i \hat{\beta}_1 = \sum_1^n Y_i \\[2mm] \sum_1^n X_i \hat{\beta}_0 + \sum_1^n X_i^2 \hat{\beta}_1 = \sum_1^n X_i Y_i \end{cases} \Rightarrow \begin{aligned} \hat{\beta}_1 &= \frac{\sum_1^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_1^n (X_i - \bar{X})^2} = \frac{sCov(X,Y)}{sVar(X)} \\[2mm] \hat{\beta}_0 &= \bar{Y} - \hat{\beta}_1 \bar{X} \end{aligned}$$

3) Леммы:

1. Линия регрессии проходит через точку $(\bar{X}, \bar{Y})$, т.е. $\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$

   Док-во: $\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \Rightarrow \boxed{\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}}$

2. $\sum_1^n e_i = 0$

   Док-во:

   $$e_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i$$

   $$\sum_1^n e_i = \sum_1^n Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_1^n X_i$$

   $$\frac{1}{n}\sum_1^n e_i = \bar{Y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{X} = 0 \Rightarrow \boxed{\sum_1^n e_i = 0}$$

3. $\sum_1^n Y_i = \sum_1^n \hat{Y}_i$

   Док-во:
   $$\sum_1^n e_i = \sum_1^n Y_i - \sum_1^n \hat{Y}_i = 0 \Rightarrow \boxed{\sum_1^n Y_i = \sum_1^n \hat{Y}_i}$$

   Следствие:
   $$\frac{1}{n}\sum_1^n e_i = \bar{Y} - \bar{\hat{Y}} = 0 \Rightarrow \boxed{\bar{Y} = \bar{\hat{Y}}}$$

4. $\sum_1^n X_i e_i = 0$

   Док-во:

   $$\frac{\partial RSS}{\partial \hat{\beta}_1} = -2\sum_1^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)X_i = 0 \Rightarrow \boxed{\sum_1^n X_i e_i = 0}$$

5. $\sum_1^n \hat{Y}_i e_i = 0$

   Док-во:

   $$\sum_1^n \hat{Y}_i e_i = \sum_1^n (\hat{\beta}_0 + \hat{\beta}_1 X_i)e_i = \hat{\beta}\sum_1^n e_i + \hat{\beta}_1 \sum_1^n X_i e_i = 0 \Rightarrow \boxed{\sum_1^n \hat{Y}_i e_i = 0}$$

4) $Y_i = \hat{Y}_i + e_i$; $\bar{Y} = \bar{\hat{Y}} \Rightarrow Y_i - \bar{Y} = \hat{Y}_i - \bar{\hat{Y}} + e_i$

$$\sum_1^n (Y_i - \bar{Y})^2 = \sum_1^n (\hat{Y}_i - \bar{\hat{Y}})^2 + 2\left(\sum_1^n \hat{Y}_i e_i - \bar{\hat{Y}} \sum_1^n e_i\right) + \sum_1^n e_i^2$$

$$\sum_1^n (Y_i - \bar{Y})^2 = \sum_1^n (\hat{Y}_i - \bar{\hat{Y}})^2 + \sum_1^n e_i^2$$

$$TSS = ESS + RSS$$

5) Коэффициент детерминации R^2:

$R^2 = \frac{ESS}{TSS}$; $x_i = X_i - \bar{X}$; $y_i = Y_i - \bar{Y}$; $\bar{\hat{Y}} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$

$$R^2 = \frac{\sum_1^n (\hat{Y}_i - \bar{\hat{Y}})^2}{\sum_1^n y_i^2} = \frac{\sum_1^n (\hat{\beta}_0 + \hat{\beta}_1 X_i - \hat{\beta}_0 - \hat{\beta}_1 \bar{X})^2}{\sum_1^n y_i^2} = \frac{\sum_1^n (\hat{\beta}_1 x_i)^2}{\sum_1^n y_i^2} = \frac{\hat{\beta}_1^2 \sum_1^n x_i^2}{\sum_1^n y_i^2} = \frac{\sum_1^n (x_i y_i)^2 \cdot \sum_1^n x_i^2}{(\sum_1^n x_i^2)^2 \cdot \sum_1^n y_i^2} =$$

$$= \frac{\sum_1^n (x_i y_i)^2}{\sum_1^n x_i^2 \cdot \sum_1^n y_i^2} = \frac{sCov^2(X,Y)}{s\mathrm{Var}(X) \cdot s\mathrm{Var}(Y)} = sCorr^2(X,Y)$$

6) Регрессия без свободного члена:

$Y_i = \beta_1 X_i + \varepsilon_i$

$\hat{Y}_i = \hat{\beta}_1 X_i \qquad RSS \to \min : \hat{\beta}_1 = \frac{\sum_1^n x_i y_i}{\sum_1^n x_i^2}$

$e_i = Y_i - \hat{\beta}_1 X_i$

a) $\sum_1^n e_i \neq 0$

b) $\sum_1^n Y_i \neq \sum_1^n \hat{Y}_i$

c) $\bar{Y} \neq \bar{\hat{Y}}$

R^2 неприменим

7) Теорема Гаусса-Маркова для парной регрессии:

Если:

1. Модель правильно специфицирована

2. Все X_i детерминированы и не равны между собой

3. Ошибки не носят систематического характера, т.е. $\mathbb{E}(\varepsilon_i) = \forall i$

4. $\mathrm{Var}(\varepsilon_i) = \sigma_\varepsilon^2 \forall i$

5. Ошибки некоррелированы

Тогд оценки МНК $\hat{\beta}_0$, $\hat{\beta}_1$ оптимальны (имеют наименьшую дисперсию) в классе линейных несмещенных оценок – являются BLUE.

8) МНК-оценки - случайные величины:

$$\hat{\beta}_1 = \frac{sCov(X,Y)}{s\mathrm{Var}(X)} = \frac{sCov(X, \beta_0 + \beta_1 X + \varepsilon)}{s\mathrm{Var}(X)} = \beta_1 \frac{s\mathrm{Var}(X)}{s\mathrm{Var}(X)} + \frac{sCov(X,\varepsilon)}{s\mathrm{Var}(X)} = \beta_1 + \frac{sCov(X,\varepsilon)}{s\mathrm{Var}(X)}$$

$$\mathbb{E}(\hat{\beta}_1) = \mathbb{E}\left(\beta_1 + \frac{sCov(X,\varepsilon)}{s\mathrm{Var}(X)}\right) = \beta_1 + \frac{1}{s\mathrm{Var}(X)} \cdot \mathbb{E}\left(\frac{1}{n-1}\sum_1^n (X_i - \bar{X})(\varepsilon_i - \bar{\varepsilon})\right) =$$

$$= \beta_1 + \frac{1}{s\mathrm{Var}(X)} \cdot \frac{1}{n-1} \cdot \sum_1^n (X_i - \bar{X})(\mathbb{E}(\varepsilon_i) - \mathbb{E}(\bar{\varepsilon})) \stackrel{\mathbb{E}(\varepsilon)=\mathbb{E}(\bar{\varepsilon})}{=} \beta_1$$

2

Утверждение 1: $\mathrm{Var}(\hat{\beta}_1) = \dfrac{\sigma_\varepsilon^2}{\sum\limits_1^n x_i^2}$; $\mathrm{Var}(\hat{\beta}_0) = \dfrac{\sigma_\varepsilon^2 \sum\limits_1^n x_i^2}{n \sum\limits_1^n x_i^2}$

$$\hat{\beta}_1 = \frac{\sum\limits_i x_i y_i}{\sum\limits_j x_j^2} = \sum_i \frac{x_i}{\sum\limits_j x_j} y_i = \sum_i w_i y_i$$

$$a) \sum_i w_i = 0 \qquad\qquad b) \sum_i w_i x_i = 1 \qquad c) \sum_i w_i^2 = \frac{1}{\sum\limits_j x_j^2}$$

$$\sum_i w_i = \frac{\sum\limits_i x_i}{\sum\limits_j x_j^2} = \frac{\sum\limits_i (X_i - \bar{X})}{\sum\limits_j x_j^2} = \frac{n\bar{X} - n\bar{X}}{\sum\limits_j x_j^2} = 0 \qquad \frac{\sum\limits_i x_i \cdot x_i}{\sum\limits_j x_j^2} = 1 \qquad \sum_i w_i^2 = \frac{\sum\limits_i x_i^2}{(\sum\limits_j x_j^2)^2} = \frac{1}{\sum\limits_j x_j^2}$$

$$\mathrm{Var}(\hat{\beta}_1) = \mathrm{Var}\left(\sum_i w_i y_i\right) = \sum_i w_i^2 \,\mathrm{Var}(y_i) + \sum_i \sum_j w_i w_j Cov(y_i, y_j) = \sum_i w_i^2 \sigma_\varepsilon^2 = \boxed{\frac{\sigma_\varepsilon^2}{\sum\limits_j x_j^2}}$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = \frac{1}{n} \sum_1^n Y_i - \sum_1^n w_i(Y_i - \bar{Y})\cdot\bar{X} = \sum_1^n \left(\frac{1}{n} - w_i\bar{X}\right) Y_i + \bar{X}\bar{Y} \sum_1^n w_i = \sum_1^n \left(\frac{1}{n} - w_i\bar{X}\right) Y_i$$

$$\mathrm{Var}(\hat{\beta}_0) = \sum_1^n \left(\frac{1}{n^2} - \frac{2}{n}w_i\bar{X} + w_i^2 \bar{X}^2\right)\sigma_\varepsilon^2 = \sigma_\varepsilon^2 \left(\frac{1}{n} - \frac{2}{n}\bar{X}\sum_1^n w_i + \bar{X}^2 \sum_1^n w_i^2\right) = \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{\bar{X}^2}{\sum\limits_1^n x_i^2}\right) = $$

$$= \sigma_\varepsilon^2 \frac{\sum\limits_1^n x_i^2 + n\bar{X}^2}{n \sum\limits_1^n x_i^2} = \boxed{\sigma_\varepsilon^2 \frac{\sum\limits_1^n X_i^2}{n \sum\limits_1^n x_i^2}}$$

$Cov(\hat{\beta}_0, \hat{\beta}_1) - ?$

$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}$; $\mathrm{Var}(\bar{Y}) = \dfrac{\sigma_\varepsilon^2}{n}$

$$\mathrm{Var}(\hat{\beta}_0 + \hat{\beta}_1 \bar{X}) = \mathrm{Var}(\hat{\beta}_0) + 2\bar{X} Cov(\hat{\beta}_0, \hat{\beta}_1) + \bar{X}^2 \,\mathrm{Var}(\hat{\beta}_1)$$

$$\frac{\sigma_\varepsilon^2}{n} = \sigma_\varepsilon^2 \frac{\sum\limits_1^n X_i^2}{n \sum\limits_1^n x_i^2} + 2\bar{X} Cov(\hat{\beta}_0, \hat{\beta}_1) + \bar{X}^2 \frac{\sigma_\varepsilon^2}{\sum\limits_j x_i^2}$$

$$Cov(\hat{\beta}_0, \hat{\beta}_1) = \frac{\sigma_\varepsilon^2 \left(\frac{1}{n} - \frac{1}{n} - \frac{\bar{X}^2}{\frac{n}{\sum\limits_1 x_i^2}} - \frac{\bar{X}^2}{\frac{n}{\sum\limits_1 x_i^2}}\right)}{2\bar{X}} = \frac{-\sigma_\varepsilon^2 \cdot 2\bar{X}^2}{2\bar{X}\sum\limits_1^n x_i^2} = \boxed{-\frac{\sigma_\varepsilon^2 \cdot \bar{X}}{\sum\limits_1^n x_i^2}}$$

Утверждение 2: при выполнении условий ТГМ оценки $\hat{\beta}_0$, $\hat{\beta}_1$ являются лучшими, т.е. имеют наименьшую дисперсию в классе всех линейных несмещенных оценок

Пусть $\tilde{\beta}_1 = \sum\limits_1^n \tilde{w}_i Y_i$ – другая несмещенная оценка, т.е. $\mathbb{E}(\tilde{\beta}_1) = \beta_1$

$$\mathbb{E}(\tilde{\beta}_1) = \sum\limits_1^n \tilde{w}_i \mathbb{E}(Y_i) = \sum\limits_1^n \tilde{w}_i \mathbb{E}(\beta_0 + \beta_1 X_i + \varepsilon_i) = \beta_0 \sum\limits_1^n \tilde{w}_i + \beta_1 \sum\limits_1^n \tilde{w}_i X_i \equiv \beta_1 \Rightarrow \quad \begin{aligned} &(1) \sum\limits_1^n \tilde{w}_i = 0 \\ &(2) \sum\limits_1^n \tilde{w}_i x_i = 1 \end{aligned}$$

Т.е. необходимо решить задачу: $\text{Var}(\tilde{\beta}_1) = \sigma_\varepsilon^2 \sum\limits_1^n \tilde{w}_i^2 \to \min$ при ограничениях (1) и (2)

$$\sum\limits_1^n \tilde{w}_2^2 = \sum\limits_1^n (\tilde{w}_i - w_i + w_i)^2 = \sum\limits_1^n (\tilde{w}_i - w_i)^2 + 2\sum\limits_1^n (\tilde{w}_i - w_i)w_i + \sum\limits_1^n w_i^2$$

$$\sum\limits_i (\tilde{w}_i - w_i)w_i = \sum\limits_i \tilde{w}_i \cdot w_i - \sum\limits_i w_i^2 = \sum\limits_i \tilde{w}_i \frac{(X_i - \bar{X})}{\sum\limits_j x_j^2} - \sum\limits_i w_i^2 = \frac{1}{\sum\limits_j x_j^2}\left(\sum\limits_i \tilde{w}_i X_i - \bar{X}\sum\limits_1^n \tilde{w}_i\right) - \frac{1}{\sum\limits_j x_j^2} =$$

$$= \frac{1}{\sum\limits_j x_j^2} - \frac{1}{\sum\limits_j x_j^2} = 0 \Rightarrow \sum\limits_i \tilde{w}_i^2 = \sum\limits_1^n (\tilde{w}_i - w_i)^2 + \frac{1}{\sum\limits_j x_j^2}$$

$\Rightarrow \sum\limits_1^n \tilde{w}_i^2$ достигает минимума при $\tilde{w}_i = w_i$, т.е. для оценок МНК.

Для β_0 доказательство аналогично

9) МНК – линейны поY

$$\hat{\beta}_1 = \frac{sCov(X,Y)}{sVar(X)} \qquad \text{Если } \tilde{Y} = nY, \qquad \hat{\tilde{\beta}} = \frac{sCov(X,\tilde{Y})}{sVar(X)} = n\hat{\beta}_1$$

10) Гипотезы о конкретном значении:

В рамках линейной модели усилим ТГМ утверждением $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$

Утверждение 1: Если $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$, то $\hat{\beta}_0 \sim N(\beta_0, \sigma_{\hat{\beta}_0}^2)$, $\hat{\beta}_1 \sim N(\beta_1, \sigma_{\hat{\beta}_1}^2)$

$H_0 : \beta_1 = \beta_1^0$

Тогда статистика $z = \dfrac{\hat{\beta}_1 - \beta_1^0}{\sigma_{\hat{\beta}_1}} \sim N(0,1)$

В $\sigma_{\hat{\beta}_1}^2$ входит σ_ε^2 и она неизвестна. Заменим оценкой $\hat{\sigma}_\varepsilon^2$

$z = \dfrac{\hat{\beta}_1 - \beta_1^0}{\hat{\sigma}_{\hat{\beta}_1}} \sim ?$

Утверждение 2: $\hat{\sigma}_\varepsilon^2 = \dfrac{RSS}{n-2}$ – несмещенная оценка σ_ε^2

Утверждение 3: $\hat{\sigma}_\varepsilon^2 = \dfrac{RSS}{n-k}$ – общий случай для k регрессоров (включая свободный член), $n - k$ – число независимых слагаемых RSS, степени свободы

Утверждение 4: $\dfrac{RSS}{\sigma_\varepsilon^2} \sim \chi_{n-k}^2$, где k - количество регрессоров

Утверждение 5: Оценки $\hat{\beta}_1$ и $\hat{\sigma}_\varepsilon^2$ независимы

$$\hat{\sigma}_{\hat{\beta}_0} = \sqrt{\hat{\sigma}_{\varepsilon^2} \cdot \frac{\sum\limits_1^n X_i^2}{n \sum\limits_1^n x_i^2}}, \qquad \hat{\sigma}_{\hat{\beta}_1} = \sqrt{\frac{\hat{\sigma}_{\varepsilon^2}}{\sum\limits_1^n x_i^2}}$$

$$\frac{\hat{\beta}_1 - \beta_1^0}{\hat{\sigma}_{\hat{\beta}_1}} = \frac{(\hat{\beta}_1 - \beta_1^0)/\sigma_{\hat{\beta}_1}}{\hat{\sigma}_{\hat{\beta}_1}/\sigma_{\hat{\beta}_1}} = \frac{(\hat{\beta}_1 - \beta_1^0)/\sigma_{\hat{\beta}_1}}{\sqrt{\frac{\hat{\sigma}_{\hat{\beta}_1}^2 (n-k)}{\sigma_{\hat{\beta}_1}^2}/(n-k)}} = \frac{(\hat{\beta}_1 - \beta_1^0)/\sigma_{\hat{\beta}_1}}{\sqrt{\frac{RSS}{\sigma_{\hat{\beta}_1}^2}/(n-k)}}$$

Так как $\dfrac{\hat{\beta}_1 - \beta_1^0}{\sigma_{\hat{\beta}_1}} \sim N(0,1)$, $\dfrac{RSS}{\sigma_{\hat{\beta}_1}^2} \sim \chi_{n-k}^2$ и они независимы, то

$$\frac{\hat{\beta}_1 - \beta_1^0}{\hat{\sigma}_{\hat{\beta}_1}} \sim t_{n-k} \qquad \boxed{\frac{\hat{\sigma}_{\hat{\beta}_1}}{\sigma_{\hat{\beta}_1}} = \frac{\hat{\sigma}_\varepsilon^2/\sum\limits_1^n x_i^2}{\sigma_\varepsilon^2/\sum\limits_1^n x_i^2} = \frac{\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2} = \frac{RSS/(n-2)}{\sigma_\varepsilon^2} = \frac{RSS/\sigma_\varepsilon^2}{n-2} \sim \chi_{n-k}^2}$$

<mark style="background:red">Штуку сверху надо полностью перепроверить и исправить</mark>

ДИ для коэффициентов: $\left|\dfrac{\hat{\beta}_1 - \beta_1^0}{\hat{\sigma}_{\hat{\beta}_1}}\right| \leq t_{n-k}^{cr} \Rightarrow |\hat{\beta}_1 - \beta_1^0| \leq t_{n-k}^{cr} \cdot \hat{\sigma}_{\hat{\beta}_1}$

$$-t_{n-k}^{cr} \cdot \hat{\sigma}_{\hat{\beta}_1} - \hat{\beta}_1 \leq -\beta_1^0 \leq t_{n-k}^{cr} - \hat{\sigma}_{\hat{\beta}_1} - \hat{\beta}_1 \Rightarrow [\hat{\beta}_1 - t_{n-k}^{cr} \cdot \hat{\sigma}_{\hat{\beta}_1}; \hat{\beta}_1 + t_{n-k}^{cr} \cdot \hat{\sigma}_{\hat{\beta}_1}]$$

11) Прогнозирование:

a) Точечный прогноз

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

$Y_{n+1} = \hat{\beta}_0 + \hat{\beta}_1 X_{n+1}$ – точечный прогноз

b) Интервальный прогноз (ИП) индивидуальный ($=$ предикативный)

$Y_{n+1} = \beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1}$

$e_{n+1}^{ind} = Y_{n+1} - \hat{Y}_{n+1} = \beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1} - \hat{\beta}_0 - \hat{\beta}_1 X_{n+1} = \varepsilon_{n+1} + \beta_0 - \hat{\beta}_0 + (\beta_1 - \hat{\beta}_1)X_{n+1}$

$$\mathrm{Var}(e^{ind}) = \mathrm{Var}(\varepsilon_{n+1}) + \mathrm{Var}[(\beta_0 - \hat{\beta}_0) + (\beta_1 - \hat{\beta}_1)X_{n+1}] =$$
$$= \sigma_\varepsilon^2 + \mathrm{Var}(\hat{\beta}_0) + X_{n+1}^2 \mathrm{Var}(\hat{\beta}_1) + 2X_{n+1} Cov[(\beta_0 - \hat{\beta}_0), (\beta_1 - \hat{\beta}_1)] =$$
$$= \sigma_\varepsilon^2 + \mathrm{Var}(\hat{\beta}_0) + X_{n+1}^2 \mathrm{Var}(\hat{\beta}_1) + 2X_{n+1} Cov(\hat{\beta}_0, \hat{\beta}_1) =$$
$$= \sigma_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{\bar{X}^2}{\sum\limits_1^n x_i^2} + \frac{X_{n+1}^2}{\sum\limits_1^n x_i^2} - \frac{2X_{n+1}\bar{X}}{\sum\limits_1^n x_i^2}\right) =$$
$$= \sigma_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{\bar{X}^2 + X_{n+1}^2 - 2X_{n+1}\bar{X}}{\sum\limits_1^n x_i^2}\right) = \sigma_\varepsilon^2 \left(1 + \frac{1}{n} + \frac{(X_{n+1} - \bar{X})^2}{\sum\limits_1^n x_i^2}\right)$$

$$\frac{\hat{Y}_{n+1} - Y_{n+1}}{\sqrt{\hat{\text{Var}}(e_{n+1}^{ind})}} \sim t_{n-k}$$

ДИ для индивидуального прогноза:

$$[\hat{\beta}_0 + \hat{\beta}_1 X_{n+1} - t_{n-k}^{cr} \cdot \sqrt{\hat{\text{Var}}(e_{n+1}^{ind})} \leq Y_{n+1} \leq \hat{\beta}_0 + \hat{\beta}_1 X_{n+1} + t_{n-k}^{cr} \cdot \sqrt{\hat{\text{Var}}(e_{n+1}^{ind})}]$$

c) Индивидуальный прогноз для среднего

$$\mathbb{E}(Y_{n+1} \,|\, X) = \mathbb{E}(\beta_0 + \beta_1 X_{n+1} + \varepsilon_{n+1} \,|\, X = X_{n+1})$$

Ошибка прогноза:

$$e_{n+1}^{mean} = \mathbb{E}(Y_{n+1} \,|\, X = X_{n+1}) - \hat{Y}_{n+1} = \beta_0 + \beta_1 X_{n+1} - \hat{\beta}_0 - \hat{\beta}_1 X_{n+1} = (\beta_0 - \hat{\beta}_0) + X_{n+1}(\beta_1 - \hat{\beta}_1)$$

$$\text{Var}(e_{n+1}^{mean}) = \text{Var}(\beta_0 - \hat{\beta}_0) + X_{n+1}^2 \text{Var}(\beta_1 - \hat{\beta}_1) + 2X_{n+1} Cov(\beta_0 - \hat{\beta}_0,\, \beta_1 - \hat{\beta}_1) = \cdots = \sigma_\varepsilon^2 \left(\frac{1}{n} + \frac{(X_{n+1} - \bar{X})^2}{\sum\limits_1^n x_i^2} \right)$$

ДИ для среднего прогноза:

$$[\hat{\beta}_0 \hat{\beta}_1 X_{n+1} - t_{n-k}^{cr} \cdot \sqrt{\hat{\text{Var}}(e_{n+1}^{mean})};\ \hat{\beta}_0 \hat{\beta}_1 X_{n+1} + t_{n-k}^{cr} \cdot \sqrt{\hat{\text{Var}}(e_{n+1}^{mean})}]$$