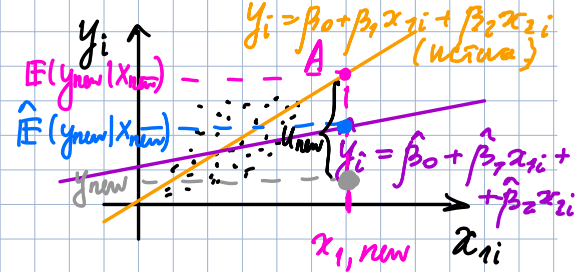
Прогнозирование



$\rightarrow y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$

$\rightarrow \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i}$

$(x_{1,new}, x_{2,new})$ – новое набл.

$\boxed{\text{ТГМ}}$

для среднест. набл.

Выборка

$A = \mathbb{E}(y_{new} | x_{1new}, x_{2new}) =$

$= \underbrace{\beta_0 + \beta_1 x_{1new} + \beta_2 x_{2new}}$

→ Точечн. → Интерв.

$\hat{\mathbb{E}}(y_{new} | X_{new})$
(подстав. и посч. $\hat{y}_{new}$)

$Ош = \hat{y}_{new} - \mathbb{E}(y_{new} | X_{new})$

$\hat{Var}(\hat{y}_{new} - \overset{const}{\mathbb{E}(y_{new} | X_{new})}) =$

$= \hat{Var}(\hat{y}_{new}) = \hat{Var}(\hat{\beta}_0 + \hat{\beta}_1 x_{1new} + \hat{\beta}_2 x_{2new}) =$

$= \hat{Var}(\hat{\beta}_0) + x_{1new}^2 \cdot \hat{Var}(\hat{\beta}_1) +$

конкрет. набл.

Выборка + u_{new}

$y_{new} = \underbrace{\beta_0 + \beta_1 x_{1new} + \beta_2 x_{2new}}_{\mathbb{E}(y_{new} | X_{new})} + u_{new}$

→ Интерв.

$Ош = \hat{y}_{new} - y_{new}$

$\hat{Var}(\hat{y}_{new} - y_{new}) =$

$= \hat{Var}(\hat{y}_{new} - \underbrace{\mathbb{E}(y_{new} | X_{new})}_{const} - u_{new}) =$

$= \hat{Var}(\hat{y}_{new} - u_{new}) =$

$= \hat{Var}(\hat{\beta}_0 + \hat{\beta}_1 x_{1new} + \hat{\beta}_2 x_{2new} - u_{new}) = \hat{Var}(\hat{y}_{new}) + \hat{Var}(u_{new})$

$\hat{\beta} = (X^T X)^{-1} X^T (\ldots + u)$

$\hat{y} = X \hat{\beta}$

$+ \ldots + \widehat{cov}(\hat{\beta}_1, \hat{\beta}_2) \cdot x_{1new}, x_{new} +$

$+ \ldots$

$$\frac{\hat{y}_{new} - \mathbb{E}(y_{new} \mid X_{new})}{\sqrt{\widehat{Var}(\hat{y}_{new} - \mathbb{E}(y_{new} \mid X_{new}))}} \sim t_{n-k} \xrightarrow[n\to\infty]{} N(0,1)$$

$\mathbb{E}(y_{new} \mid X_{new}) \in [\hat{y}_{new} \pm se(\hat{y}_{new}) \cdot z_{cr} / t_{cr}]$

$se(M) := \sqrt{\widehat{Var}(M)}$

standard error

Доверит. инт.

$$\frac{\hat{y}_{new} - y_{new} - 0}{se(\hat{y}_{new} - y_{new})} \sim t_{n-k} \underset{n\to\infty}{\sim} N(0,1)$$

$y_{new} \in [\hat{y}_{new} \pm se(\hat{y}_{new} - y_{new}) \cdot z_{cr} / t_{cr}]$

Предиктивный инт.

(и) $y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$

ТЛМ, $u \sim N(0, \sigma^2 I)$

$\hat{y}_i = 2 + 2x_{1i} + 1 \cdot x_{2i}$

$x_{1,6} = 10$, $x_{2,6} = 7$

" $x_{1,new}$ " $x_{2,new}$

$\hat{y}_6 = 2 + 2 \cdot 10 + 1 \cdot 7 = 29$

(к) 95% CI для $\mathbb{E}(y_6 \mid X_6)$

confidence interval

$\mathbb{E}(y_{new} \mid X_{new}) \in [\hat{y}_{new} \pm se(\hat{y}_{new}) \cdot z_{cr} / t_{cr}]$

$$\hat{Var}(\hat{y}_{new}) = \hat{Var}(\hat{\beta}_0 + \hat{\beta}_1 \cdot 10 + \hat{\beta}_2 \cdot 7) =$$

$$= \hat{Var}(\hat{\beta}_0) + 100 \cdot \hat{Var}(\hat{\beta}_1) + \ldots + 7 \cdot 10 \cdot \hat{cov}(\hat{\beta}_1, \hat{\beta}_2) +$$

$$+ \ldots$$

$$\hat{Var}(\hat{\beta}) = \begin{array}{c} \hat{Var}(\hat{\beta}_0) \\ \begin{pmatrix} 1/3 & -1/3 & 0 \\ -1/3 & 4/3 & -1 \\ 0 & -1 & 2 \end{pmatrix} \end{array} \ \hat{cov}(\hat{\beta}_1, \hat{\beta}_2)$$

$$\mathbb{E}(y_{new} | X_{new}) \in [29 \pm se(\hat{y}_{new}) \cdot 4.3]$$

(л) 95% PI для $y_6 | X_6$

predictive interval

$$y_{new} \in [\underset{\underset{29}{\parallel}}{\hat{y}_{new}} \pm \underset{\underset{\sqrt{\ldots}}{\parallel}}{se(\hat{y}_{new} - y_{new})} \cdot \underset{\underset{4.3}{\parallel}}{z_{cr} / t_{cr}}]$$

$$\hat{Var}(\hat{y}_{new} - y_{new}) = \underset{\underset{\text{знаем}}{\parallel}}{\hat{Var}(\hat{y}_{new})} + \underset{\underset{\hat{\sigma}^2 = \frac{RSS}{n-k} = 1}{\parallel}}{\hat{Var}(u_{new})}$$

с прошл. сем.

Распределения в лин. регр.

№3 $u \in \mathbb{R}^n$, $u \sim N(0, I)$

Def $V \subset \mathbb{R}^n$, $\dim V = k$

$$\| \hat{u}_V \|_2^2 \sim \chi^2_k$$

$\hat{u}_V$ – проекция u на V

(а) $u \in \mathbb{R}^3$, $u \sim N(0, I)$

$$V = \{ (x_1, x_2, x_3 \mid x_3 = 2x_1 + x_2) \}$$

$\hat{u}_V$ – ? $\quad \|\hat{u}_V\|^2 \sim$?

Выберем базис в V: $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

$$X = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 1 \end{pmatrix} \qquad \hat{u}_V = X \underbrace{(X^T X)^{-1} X^T}u$$

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \qquad \text{МНК: } \hat{y} = X \cdot \hat{\beta}, \quad \hat{\beta} = (X^T X)^{-1} X^T y$$

$$\hat{u}_V = \text{Матрица} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

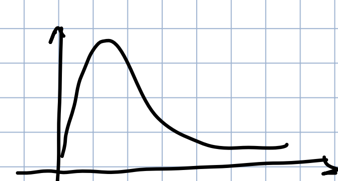
$$\|\hat{u}_V\|^2 \sim \chi^2_2$$

(б) $\dim(V^\perp) = n - k = 3 - 2 = 1$

$$\|\hat{u}_{V^\perp}\|^2 \sim \chi^2_1$$

№4 $\quad X \sim \chi^2_a, \quad Y \sim \chi^2_b, \quad X \perp\!\!\!\perp Y$

$$Z = \frac{X/a}{Y/b} \sim F_{a,b}$$

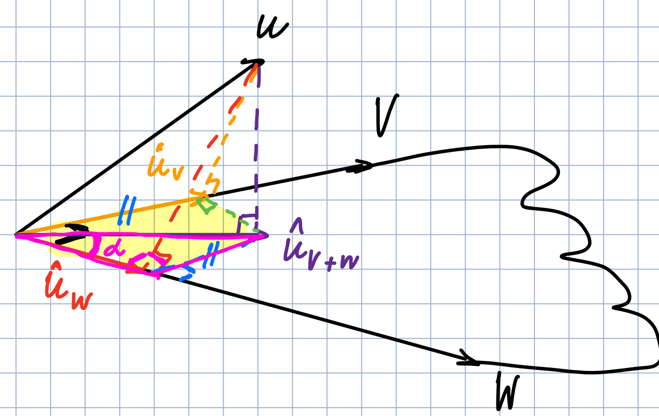


$u \in \mathbb{R}^n$, $u \sim \mathcal{N}(0, I)$

$V \perp W$

$\dim V = k$, $\dim W = m$

(а) $\hat{u}_V$, $\hat{u}_W$



$\|\hat{u}_V\|^2 \sim \chi^2_k$

$\|\hat{u}_W\|^2 \sim \chi^2_m$

(б) $\hat{u}_{V+W}$

(в) $tg^2 \alpha = \dfrac{\|\hat{u}_V\|^2}{\|\hat{u}_W\|^2}$ $\qquad tg\, \alpha = \dfrac{\|\hat{u}_V\|}{\|\hat{u}_W\|}$

(г) $$\frac{m}{k} tg^2 \alpha = \frac{m}{k} \cdot \frac{\|\hat{u}_V\|^2}{\|\hat{u}_W\|^2} = \frac{\|\hat{u}_V\|^2 / k}{\|\hat{u}_W\|^2 / m} = \frac{\|\hat{u}_V\|^2 / \dim V}{\|\hat{u}_W\|^2 / \dim W}$$

$\sim F_{k,m}$

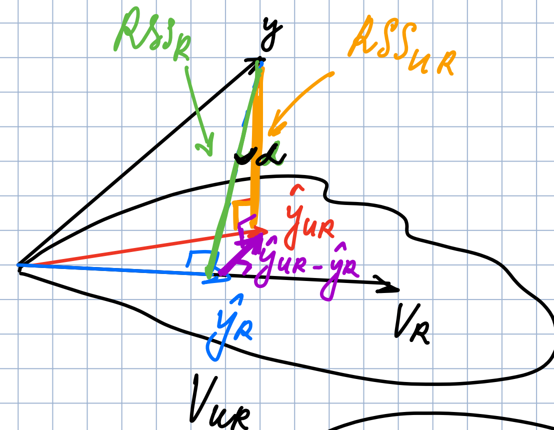
№5 $\begin{cases} H_0: & UR, R \text{ верна} \\ H_1: & UR \text{ верна}, R \text{ неверна} \end{cases}$

$$F = \frac{(RSS_R - RSS_{UR}) / (k_{UR} - k_R)}{RSS_{UR} / (n - k_{UR})} \sim F_{(k_{UR} - k_R;\, n - k_{UR})}$$

UR: $y_i = \beta_0 + \beta_1 x_i + \beta_2 z_i + u_i$

R: $y_i = \beta_0 + \beta_1 (x_i + z_i) + u_i$

$u_i \sim N(0, 1)$ и незав. ; ТГМ



(a) V_{UR}, V_R

(б) $RSS_R - RSS_{UR} =$

$= \| \hat{y}_{UR} - \hat{y}_R \|^2$

$$\underbrace{\| \hat{y}_{UR} - \hat{y}_R \|^2}_{\text{катет}} = \underbrace{\| y - \hat{y}_R \|^2}_{\text{гипот.}} - \underbrace{\| y - \hat{y}_{UR} \|^2}_{\text{катет}} = RSS_R - RSS_{UR}$$

$\| y - \hat{y}_R \|^2 = RSS_R$, $\| y - \hat{y}_{UR} \|^2 = RSS_{UR}$

(в) $$tg^2 \alpha = \frac{\| \hat{y}_{UR} - \hat{y}_R \|^2}{\| y - \hat{y}_{UR} \|^2}$$

(г) $y - \hat{y}_{UR}$ (оранж. вектор)

$y - \hat{y}_{UR}$ — проекция y на $(V_{UR})^\perp$

$\dim (V_{UR}^\perp) = n - k_{UR}$

$\hat{y}_{UR} - \hat{y}_R$ (фиол.) — проекц. y на $(V_{UR} \cap V_R^{\perp})$

$\dim(V_{UR} \cap V_R^{\perp}) = k_{UR} - k_R$

(f) $$F = \frac{\|\hat{y}_{UR} - \hat{y}_R\|^2 / (k_{UR} - k_R)}{\|y - \hat{y}_{UR}\|^2 / (n - k_{UR})} \sim F(k_{UR} - k_R; n - k_{UR})$$

$$\parallel$$

$$\frac{(RSS_R - RSS_{UR}) / (k_{UR} - k_R)}{RSS_{UR} / (n - k_{UR})}$$