① Rozkład Studenta

$X \sim N(0,1)$; $Z \sim \chi^2(n)$ – niezależne

$$\Rightarrow T = \frac{X}{\sqrt{\frac{Z}{n}}} = \sqrt{n}\,\frac{X}{\sqrt{Z}} \sim t(n)$$

rozkład Studenta z n st. sw.

$X \sim N(0,1)$
$Y_1 \dots Y_n \sim N(0,1)$
niezależne

$$\Rightarrow T = \sqrt{n}\,\frac{X}{\sqrt{Y_1^2 + \dots + Y_n^2}} \sim t(n)$$

Tw. Studenta

$X_1 \dots X_n \sim N(\mu, \sigma^2)$, niezależne

$\bar{X} = \frac{1}{n} \sum X_i$ $\qquad S^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$

$$\Rightarrow T = \sqrt{n}\,\frac{\bar{X} - \mu}{S} \sim t(n-1)$$

Rozkład Fishera (Fishera-Snedecora)

$X \sim \chi^2(m)$; $Y \sim \chi^2(n)$ – niezależne

$$\Rightarrow Z = \frac{X/m}{Y/n} \sim F(m,n)$$ – rozkład Fishera z m, n st. sw.

$X_1 \dots X_m \sim N(0,1)$
$Y_1 \dots Y_n \sim N(0,1)$
niezależne

$$\Rightarrow Z = \frac{n}{m}\,\frac{X_1^2 + \dots + X_m^2}{Y_1^2 + \dots + Y_n^2} \sim F(m,n)$$

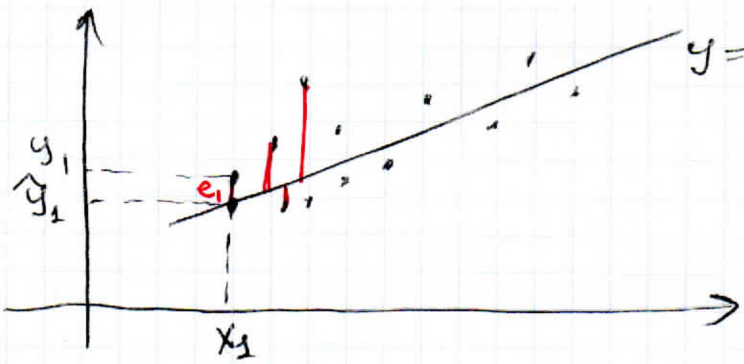
$T^2 \sim$???

② $y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ $\qquad \varepsilon_i \sim N(0, \sigma^2)$

$\beta_0, \beta_1, \sigma^2$ - ???

X_i - dane

y_i - zm. los.



$$\hat\beta_1 = \frac{\sum (X_i - \bar X)(y_i - \bar y)}{\sum (x_i - \bar x)^2}$$

$$\hat\beta_0 = \bar y - \hat\beta_1 \bar x$$

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

SST, SSE, SSM

$(y_1 \ldots \quad y_n)$

$(\hat y_1 \ldots \quad \hat y_n)$

$\bar y = \frac{1}{n}\sum y_i = \frac{1}{n}\sum \hat y_i$

$\ulcorner \sum \hat y_i = \sum \hat\beta_0 + \sum \hat\beta_1 x_i =$

$= \sum(\bar y - \hat\beta_1 \bar x) + \hat\beta_1 \underbrace{\sum x_i}_{n\bar x} =$

$= n\bar y - n\hat\beta_1 \bar x + \hat\beta_1 \cdot n\bar x = n\bar y$

$\Rightarrow \bar y = \frac{1}{n}\sum \hat y_i \quad \lrcorner$

$SST = \sum (y_i - \bar y)^2 \qquad /\ \mathrm{var}\, y.$

$SSE = \sum (y_i - \hat y)^2 \qquad / \sum e_i^2$

$SSM = \sum (\hat y_i - \bar y)^2 \qquad / \mathrm{var}\, \hat y$

$$\begin{array}{ccccc} SST & = & SSE & + & SSM \\ \downarrow & & \downarrow & & \downarrow \\ n-1 & & n-2 & & 1 \\ \shortparallel & & \shortparallel & & \shortparallel \\ dfT & = & dfE & + & dfM \\ MST & & MSE & & MSM \end{array}$$

$\sigma^2 \chi^2.$

$$MS_{..} = \frac{SS_{..}}{df_{..}}$$

$$SSM = \sum (\hat{y}_i - \bar{y})^2 = \sum (\underbrace{\hat{\beta}_0}_{\bar{y} - \hat{\beta}_1 \bar{x}} + \hat{\beta}_1 x_i - \bar{y})^2 =$$

$$= \sum (\hat{\beta}_1 x_i - \hat{\beta}_1 \bar{x})^2 = \hat{\beta}_1^2 \sum (x_i - \bar{x})^2$$

$$\hat{\beta}_1 \sim N\left(0, \frac{\sigma^2}{\sum (x_i - \bar{x})^2}\right)$$

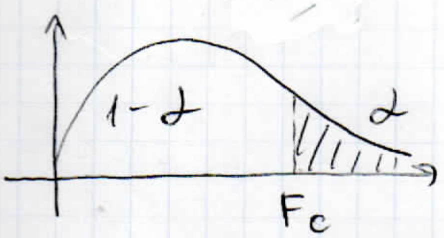
$$s^2 = \frac{SSE}{dfE} = \frac{1}{n-2} \sum (y_i - \hat{y}_i)^2$$ - estymator σ^2

Testowanie $\beta_1 = 0$?

$H_0: \beta_1 = 0$ $\quad H_1: \beta_1 \neq 0$

$$F = \frac{MSM \sim \chi^2(1)/dfM}{MSE \sim \chi^2(n-2)/dfE} \sim F(1, n-2) \quad (H_0 +)$$

$F > F_c$ - odrzucamy H_0



$F_c = qf(1-\alpha, df1 = 1, df2 = n-2)$

P - wartość:

$p = P(Z > F)$ $\quad Z \sim F(1, n-2)$

$F = \frac{MSM}{MSE}$ - obliczone na podstawie danych

$$R^2 = \frac{SSM}{SST} = 1 - \frac{SSE}{SST}$$ współczynnik determinacji

$= r^2 = cor^2(X, Y)$ ← random

$r = \pm\sqrt{R^2}$ (sign of β_1)

obowiązk.

predict (______ , ______ , se.fit = F ,

object klasy "lm" — data.frame które zmienne przewidywać — czy wymagane błędy są?

interval = , level = 0,95 , ...)

interval: "none" "confidence" "prediction" — level: poziom ufności

DF = data.frame (...)

X	y
x_1	y_1
⋮	⋮
x_n	y_n

LinM = lm(Y~X, DF) X0 <- data.frame(size = c(x0))

predict (LinM) – predykcja dla $x_1 \dots x_n$

predict (LinM, X0) – predykcja dla x_0

predict (LinM, X0, interval = "confidence") – predykcja dla x_0 + przedział ufności dla Ey_h

predict (LinM, X0, interval = "prediction") – – " – + przedział predykcyjny dla y_h.

predict (LinM, X0, se.fit = T) – " – + błąd standardowy dla $Ey_h = \hat{y}_h$