# PART 2: LINEAR REGRESSION

TMA4267 L10
17.02.2017

[F32]

## Distribution of $\hat{\varepsilon}$ (residuals)

Residuals: $\hat{\varepsilon} = Y - \hat{Y} = Y - X\hat{\beta} = Y - \overbrace{X\underbrace{(X^T X)^{-1} X^T Y}_{\hat{\beta}}}^{H}$

$$= Y - HY = (I - H)Y$$

$$E(\hat{\varepsilon}) = E\big((I-H)Y\big) = (I-H)\underbrace{E(Y)}_{X\beta}$$

$Y = X\beta + \varepsilon$, $E(\varepsilon) = 0$

$E(Y) = X\beta$

$\mathrm{Cov}(Y) = \mathrm{Cov}(\varepsilon) = \sigma^2 I$

$$= \underbrace{(I-H)}_{\text{project onto}} X\beta = \underline{\underline{0}}$$

project onto the space orth. to column space of $X$

or $(I-H)X\beta = (X - \overbrace{HX}^{X})\beta = 0$

$$HX = X\underbrace{\underbrace{(X^T X)^{-1}}\underbrace{X^T X}}\ = X$$

$$\mathrm{Cov}(\hat{\varepsilon}) = \mathrm{Cov}\big((I-H)Y\big) = (I-H)\underbrace{\mathrm{Cov}(Y)}_{\sigma^2 I}(I-H)^T$$

$$= \sigma^2 (I-H) I (I-H) = \underline{\underline{\sigma^2 (I-H)}}$$

Assume $\varepsilon \sim N_n(0, \sigma^2 I) \Rightarrow \underline{\hat{\varepsilon} \sim N_n(0, \sigma^2 (I-H))}$

NB: $\mathrm{rank}(H) = p$, $\mathrm{rank}(I-H) = n-p$, which means $(I-H)^{-1}$ does not exist and we use the singular version of the normal pdf.

## Distribution of SSE and $\hat{\sigma}^2$

$$SSE = \hat{\varepsilon}^T \hat{\varepsilon} = Y^T (I-H)(I-H) Y$$

$$\uparrow$$

$$\sum_{i=1}^{n} (y_i - \hat{y})^2 \qquad \hat{\varepsilon} = (I-H)Y$$

$$\underline{SSE = Y^T (I-H) Y}$$

remember rank $(I-H) = n-p$.

RecEx3.P3 looks at the distribution of $\frac{1}{\sigma^2} Y^T(I-H)Y = \frac{SSE}{\sigma^2}$ by using the result on quadratic forms from Part 1 (see slide)

$$Y \sim N_n(X\beta, \sigma^2 I)$$

$$Y^* = \frac{1}{\sigma}(Y - X\beta) \sim N_n(0, I)$$

$$\underbrace{Y^{*T} \underbrace{(I-H) Y^*}_{\substack{(I-H)(Y-X\beta) \\ (I-H)Y - \underbrace{(I-H)X\beta}_{0}}}}_{\chi^2_{n-p}} = \frac{1}{\sigma^2} Y^T (I-H) Y$$

$$\parallel$$

$$\frac{SSE}{\sigma^2} \sim \chi^2_{n-p}$$

Now: $\hat{\sigma}^2 = \frac{1}{n-p} \underset{\substack{\hat{\varepsilon}^T\hat{\varepsilon} \\ \parallel}}{SSE} \iff SSE = (n-p)\hat{\sigma}^2$

$$V = \frac{SSE}{\sigma^2} = \frac{(n-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}$$

$E(V) = n-p \qquad Var(V) = 2(n-p)$

Is $\hat{\sigma}^2$ an unbiased estimator?

$$E(\hat{\sigma}^2) = E\Big(\frac{1}{n-p} \underbrace{SSE}_{\sigma^2 \cdot V}\Big) = \frac{1}{n-p} E(\sigma^2 V)$$

$$= \frac{\sigma^2}{n-p} \underbrace{E(V)}_{n-p} = \underline{\underline{\sigma^2}}$$ unbiased.

This is true when we assume $\varepsilon \sim N$.

If we do not assume $\varepsilon \sim N$, then we can use the trace-formula

$$E(SSE) = E\big(Y^T (I-H) Y\big) \qquad E(Y) = X\beta, \quad Cov(Y) = \sigma^2 I$$

$$= tr\big((I-H)\sigma^2 I\big) + (X\beta)^T \underbrace{(I-H) X\beta}_{0}$$

$$= (n-p)\sigma^2 + 0$$

$$E(\hat{\sigma}^2) = E\Big(\frac{SSE}{n-p}\Big) = \underline{\underline{\sigma^2}}$$

## Inference about one $\beta_j$

Ex: Acid rain $\quad \beta_1$ = effect of $SO_4$ on pH of lake

$$\hat{\beta}_1 = -0.315$$

$$SD(\hat{\beta}_1) = \sqrt{\sigma^2 \left[(X^T X)^{-1}\right]_{[\text{diagonal elem. corresp. to } SO_4\ x_1]}}$$

$$\hat{SD}(\hat{\beta}_1) = \sqrt{\hat{\sigma}^2 \left[(X^T X)^{-1}_{[x_1, x_1]}\right]}$$

$\Downarrow$

St. Error $\Rightarrow$ 0.0587 in printout

$\hat{\sigma}$ : "Residual standard error" = 0.1165

$$\hat{\sigma} = \sqrt{\frac{SSE}{n-p}} \qquad \left.\begin{matrix} n = 26 \\ p = 8 \end{matrix}\right\} \; n - p = 18$$

"on 18 degrees of freedom".

To find a confidence interval for $\beta_j$ – or to test hypotheses about $\beta_j$ we need to know the distribution of a statistic involving $\hat{\beta}_j$ and $\beta_j$ – with no other unknown parameters.

$$\hat{\beta}_j \sim N_1\left(\beta_j \, , \, \sigma^2 \underbrace{(X^T X)^{-1}_{[j,j]}}_{c_{jj}}\right)$$

and $\hat{\sigma}^2 = \frac{SSE}{n-p}$ where $\frac{SSE}{\sigma^2} \sim \chi^2_{n-p}$.

Then:

$$\frac{\hat{\beta}_j - \overbrace{\beta_j}^{E(\hat{\beta}_j)}}{\underbrace{\sqrt{C_{jj}\sigma^2}}_{SD(\hat{\beta}_j)}} \sim N(0,1)$$

$\hat{SD}(\hat{\beta}_j) = \sqrt{C_{jj}} \cdot \hat{\sigma}$ and $\hat{\beta}_j$ and $\hat{\sigma}^2$ are independent (to be shown)

$$T_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}}\,\hat{\sigma}} \sim t_{n-p}$$

General result:

$$\frac{N(0,1)}{\sqrt{\frac{\chi^2_q}{q}}} \sim t_q$$

We have:

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}^{-1}}\ \sigma} \sim N(0,1)$$

and

$$\frac{(n-p)\,\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p}$$

$$\frac{\dfrac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}^{-1}}\cdot\cancel{\sigma}}}{\sqrt{\dfrac{\cancel{(n-p)}\,\hat{\sigma}^2}{\cancel{\sigma^2}} \Big/ \cancel{n-p}}} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}^{-1}}\ \hat{\sigma}} \sim t_{n-p}$$

$\hat{\beta}_j$ and $\hat{\sigma}^2$ need to be independent in order that this holds. $\Rightarrow$ Rec Ex 3.P3 + slides

$$\Downarrow$$

Use $T_j = \dfrac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}^{-1}}\ \hat{\sigma}} \sim t_{n-p}$ for inference.

a) Find a 95% confidence interval (CI) for $\beta_j$

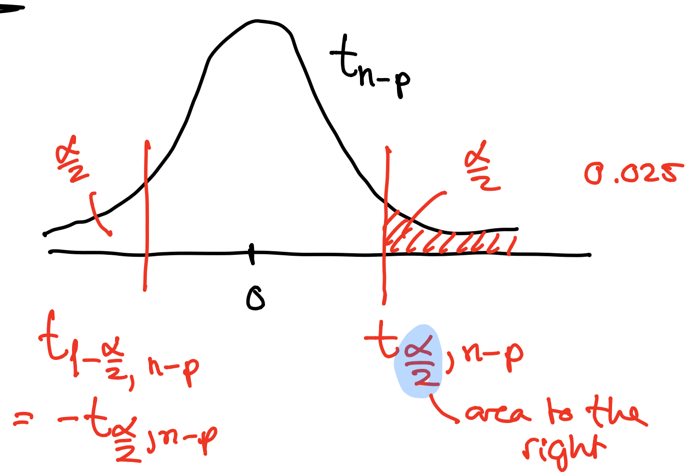
b) Test:

$$H_0: \beta_j = 0 \quad vs \quad H_1: \beta_j \neq 0$$

at significance level 0.05.

Ex: Acid rain $\underline{\beta_1}$ = effect of $SO_4$ on pH.

a) $\underbrace{95\%}_{(1-\alpha)\cdot 100}$ CI :




$$P\Big(-\underset{0.025}{t_{\frac{\alpha}{2}, n-p}} < T_j < \underset{0.025}{t_{\frac{\alpha}{2}, n-p}}\Big) = \underset{0.95}{1-\alpha}$$

$$\uparrow$$

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{C_{jj}}\cdot\hat{\sigma}}$$

L = lower
U = upper

$$P\Big(\overbrace{\hat{\beta}_j - t_{\frac{\alpha}{2}, n-p}\cdot\sqrt{C_{jj}}\cdot\hat{\sigma}}^{\hat{\beta}_L} < \beta_j < \underbrace{\hat{\beta}_j + t_{\frac{\alpha}{2}, n-p}\sqrt{C_{jj}}\,\hat{\sigma}}_{\hat{\beta}_U}\Big) = 1-\alpha$$

Ex:

$$\left.\begin{aligned} \hat{\beta}_1 &= -0.815 \\ \sqrt{C_{jj}}\,\hat{\sigma} &= 0.058 \\ n = 26&, \; p = 8 \\ t_{0.025,\,18} &= 2.1 \end{aligned}\right\} \quad \begin{aligned} &-0.815 \pm 2.1 \cdot 0.058 \\ &= \underline{\underline{[-0.44, -0.19]}} \end{aligned}$$

How do you interpret this interval?

→ strong belief (95%) that $\beta_j$ is in interval:

We see that 0 is not in the interval – what does this mean? ⇒ Reject $H_0: \beta_j = 0$ vs $H_1: \beta_j \neq 0$ at sign. level 5%