M2: Multiple linear regression (2nd week) 04.09.2017

So far: What to remember

Model: $\underset{n\times 1}{Y} = \underset{n\times p}{X}\underset{p\times 1}{\beta} + \underset{n\times 1}{\varepsilon}$

Where X (design matrix) has full rank (rank p) $n \gg p$

Normal model: $\varepsilon \sim N_n(0, \sigma^2 I) \Rightarrow Y \sim N_n(X\beta, \sigma^2 I)$

$\varepsilon_1, \dots, \varepsilon_n$ i.i.d $N(0, \sigma^2)$

ML estimation (LS): $\hat{\beta} = (X^T X)^{-1} X^T Y$

$\sim N_p(\beta, \sigma^2 (X^T X)^{-1})$

$\hat{\varepsilon} = Y - \hat{Y} = Y - X\hat{\beta}$

REML estimation: $\hat{\sigma}^2 = \frac{1}{n-p} SSE$

$\hat{\varepsilon}^T\hat{\varepsilon} = \sum_{i=1}^{n} \hat{\varepsilon}_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

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

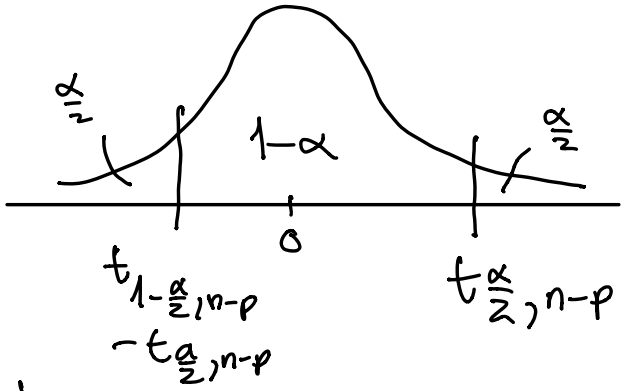
Elements of $\hat{\beta}$: $\hat{\beta}_j \sim N_1(\beta_j, \sigma^2 \underbrace{(X^T X)^{-1}_{[j,j]}}_{c_{jj}})$

$$T_j = \frac{\frac{\hat{\beta}_j - \beta_j}{\sqrt{c_{jj}}\cdot\sigma} \sim N(0,1)}{\sqrt{\frac{\hat{\sigma}^2 (n-p)}{\sigma^2} \Big/ (n-p)}} = \frac{\hat{\beta}_j - \beta_j}{\sqrt{c_{jj}}\cdot\hat{\sigma}} \sim t_{n-p}$$

$\frac{N(0,1)}{\sqrt{\frac{\chi^2_{n-p}}{n-p}}} \sim t_{n-p}$ ← general rule

Confidence interval (CI): β_j

$T_j \sim t_{n-p}$



$$P(-t_{\frac{\alpha}{2},n-p} \leq T_j \leq t_{\frac{\alpha}{2},n-p}) = 1-\alpha$$

and solve for β_j in the middle

Interpretation: 95% CI

Fix X

Collect Y

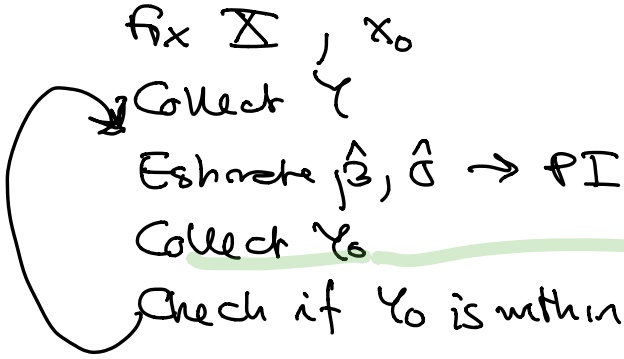
Construct CI from $\hat{\beta}$ and $\hat{\sigma}$

Check if true β_j in CI

On average 95% of the CI will contain the true β_j.

Prediction interval (PI)

Interpretation : 95% PI



On average 95% of the PI's will contain the collected Y_0.

Single hypothesis testing

$H_0: \beta_j = 0$ vs $H_1: \beta_j \neq 0$ (think area)

	H_0 true $\beta_j = 0$	H_0 false $\beta_j \neq 0$
Reject H_0 say "$\beta_j \neq 0$"	Type I error	Correct
Not reject H_0 say "$\beta_j = 0$"	Correct	Type II error

Linear hypotheses

$H_0: C\beta = d \quad vs \quad C\beta \neq d$ (C: $r \times p$, β: $p \times 1$, d: $r \times 1$)

A : large model

B : small model (H_0 true)

SSE ↑ A $\qquad$ SSE_{H_0} ↑ B

$$f_{obs} = \frac{\frac{1}{r}(SSE_{H_0} - SSE)}{\frac{SSE}{n-p}} \sim f_{r,\, n-p}$$

Anova tables: sequential change in SSE

(β_0 always in here)

$SSE(\beta_1, \beta_2, \dots, \beta_k)$ ← SSE full model

$SSE(\beta_1)$ = SSE when only β_1 (with x_1) in

$SSE(\beta_1, \beta_2)$ = SSE when both β_1 and β_2 (x_1 and x_2) in

$$SSE(\beta_2 | \beta_1) = SSE(\beta_1, \beta_2) - SSE(\beta_1)$$

↑ added effect of β_2 compared to model with β_1

$$F = \frac{\frac{1}{r} SSE(\beta_2|\beta_1)}{\frac{SSE(\beta_1, \beta_2, \dots, \beta_k)}{n-p}}$$

← added effect

full model

r = difference in number of parameters between model with (β_1, β_2) and (β_1)

This is used in 'anova(fit)' in R.

Called type 1.

(Can not – as far as I understand – be written as an hypothesis test with $C\beta = d$.)