Resum Càlculs Matrius

Model: $Y = X \cdot \beta + e$ amb

$E[Y] = X \cdot \beta$ i $Var(Y) = Var(e) = I_n \sigma^2$

- $\hat{\beta} = (X^t X)^{-1} X^t Y$
- $\hat{Y} = X\hat{\beta} = X(X^t X)^{-1} X^t Y = H \cdot Y$

  amb $H = X(X^t X)^{-1} X^t = (h_{ij})$

  i $var\left(\hat{Y}\right) = \sigma^2 diag(H) = \sigma^2 (h_{ii})$
- $\hat{r} = Y - \hat{Y} = (I_n - H) Y$
- $\hat{\sigma}^2 = \frac{SQE}{GLE} = \frac{\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}{n-p} = SE^2$

continuació

- $Y_0 = X_0\hat{\beta}$
- $r_{stand.} = \left(\frac{\hat{r}_i}{\sqrt{1_n - h_{ii}}}\right)\frac{1}{SE}$ limits habituals: ± 2
- $r_{student} = \left(\frac{\hat{r}_i}{\sqrt{1_n - h_{ii}}}\right)\frac{1}{SE_{(i)}}$ limits habituals: ± 2
- $Leverage = (h_{ii}) = diag\,(H)$ limits habituals: $0,\ 2 \cdot \overline{H}$
- $Cook's\ distance = \left(\frac{r^2_{stand.i}}{p}\frac{h_{ii}}{1-h_{ii}}\right)$ limits habituals: $0,\ \frac{4}{n}$
- $Dffits = \left(\frac{\hat{Y}_i - \hat{Y}_{i(i)}}{\sqrt{h_{ii}}}\right)\frac{1}{SE_{(i)}}$ limits habituals: $\pm 2\sqrt{\frac{p}{n}}$

Nota: Els límits habituals també es posen amb 2 i 3

Càlculs Matrius

$Y = X \cdot \beta + e$

- $E[Y] = X \cdot \beta$
- $Var(Y) = Var(e) = I_n \sigma^2$

$\hat{\beta}$

$X^t Y = X^t X \hat{\beta}$

$$\hat{\beta} = (X^t X)^{-1} X^t Y$$

- $E\left[\hat{\beta}\right] = (X^t X)^{-1} X^t E[Y] = (X^t X)^{-1} X^t X \beta = \beta$
- $Var\left(\hat{\beta}\right) = (X^t X)^{-1} X^t (I_n \sigma^2) \left((X^t X)^{-1} X^t\right)^t = \sigma^2 (X^t X)^{-1} X^t I_n X (X^t X)^{-1} = \sigma^2 (X^t X)^{-1}$
- $var\left(\hat{\beta}\right) = \sigma^2 diag\left((X^t X)^{-1}\right)$

$\hat{Y}$

$$\hat{Y} = X\hat{\beta} = X\left(X^tX\right)^{-1}X^tY = (hat)\,Y$$

- $E\left[\hat{Y}\right] = X\left(X^tX\right)^{-1}X^tE\left[Y\right] = X\left(X^tX\right)^{-1}X^tX\beta = X\beta$
- $Var\left(\hat{Y}\right) = X\left(X^tX\right)^{-1}X^tVar\left(Y\right)\left(X\left(X^tX\right)^{-1}X^t\right)^t = \sigma^2X\left(X^tX\right)^{-1}X^tX\left(X^tX\right)^{-1}X^t = \sigma^2\left(HAT\right)$
- $var\left(\hat{Y}\right) = \sigma^2 diag\left(HAT\right) = \sigma^2 hat$

$\hat{Y}_0$

$X_0 = (X_{0,1}, \ldots, X_{0,p}) \Rightarrow \hat{Y}_0 = X_0 \hat{\beta}$

- $E\left[\hat{Y}_0\right] = X_0 E\left[\hat{\beta}\right] = X_0 \beta$
- $Var\left(\hat{Y}_0\right) = X_0 Var\left(\hat{\beta}\right) X_0^t = \sigma^2 X_0 \left(X^t X\right)^{-1} X_0^t$

$\hat{r}$

$$\hat{r} = Y - \hat{Y} = \left(I_n - X\left(X^tX\right)^{-1}X^t\right)Y$$

- $E\left[\hat{r}\right] = \left(I_n - X\left(X^tX\right)^{-1}X^t\right)E\left[Y\right] = \left(X - X\left(X^tX\right)^{-1}X^tX\right)\beta = 0$
- $Var\left(\hat{r}\right) = \left(I_n - X\left(X^tX\right)^{-1}X^t\right)\left(I_n\sigma^2\right)\left(I_n - X\left(X^tX\right)^{-1}X^t\right)^t = \sigma^2\left(I_n - 2X\left(X^tX\right)^{-1}X^t + X\left(X^tX\right)^{-1}X^tX\left(X^tX\right)^{-1}X^t\right) = \sigma^2\left(I_n - X\left(X^tX\right)^{-1}X^t\right) = \sigma^2\left(I_n - HAT\right)$
- $var\left(\hat{r}\right) = \sigma^2 diag\left(I_n - HAT\right) = \sigma^2\left(1_n - hat\right)$

Variàncies mostrals

$$\hat{\sigma}^2 = \frac{1}{n-p}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 = \frac{1}{n-p}\sum_{i=1}^{n}\left(\hat{r}_i\right)^2 = \frac{\hat{r}^t\hat{r}}{n-p} =$$
$$= \frac{SQE}{GLE} = MQE \text{ i } \hat{\sigma} = \sqrt{MQE} = SE$$

- $S^2_{\hat{\beta}} = MQE \cdot diag\left(\left(X^tX\right)^{-1}\right)$
- $S^2_{\hat{Y}} = MQE \cdot hat$
- $S^2_{\hat{Y}_0} = MQE \cdot X_0^t\left(X^tX\right)^{-1}X_0$
- $S^2_{\hat{r}} = MQE \cdot \left(1_n - hat\right)$

Residuals i hat

$$r_{stand.} = \frac{\hat{r} - E[\hat{r}]}{S_{\hat{r}}} = \frac{\hat{r}}{SE \cdot \sqrt{1_n - hat}}$$

- limits habituals: ± 2

$$r_{student} = \frac{\hat{r}}{SE_{(i)} \cdot \sqrt{1 - hat}}$$

- limits habituals: ± 2

$$hat = diag\left(X\left(X^t X\right)^{-1} X^t\right) = diag\left(HAT\right)$$

- limits habituals: $0,\ 2 \cdot \overline{hat}$

Influència

$$Cook's\ distance = \frac{\hat{r}^2}{MQE}\frac{hat}{p\cdot(1_n - hat)^2} = \frac{r^2_{stand.}}{p}\frac{hat}{1-hat}$$

- limits habituals: 0, $\frac{4}{n}$

$$Dffits = \frac{\hat{Y} - \hat{Y}_{(i)}}{SE_{(i)}\sqrt{hat}} = \frac{\hat{Y} - X_i\hat{\beta}_{(i)}}{SE_{(i)}\sqrt{hat}}$$

- limits habituals: $\pm 2\sqrt{\frac{p}{n}}$

$$Dfbetas = \frac{\hat{\beta} - \hat{\beta}_{(i)}}{S_{\hat{\beta}_{(i)}}}$$

- limits habituals: $\pm\frac{2}{\sqrt{n}}$