a) Suponga que $\mathbf{x}_0 = [1, x_{01}, \cdots, x_{0k}]$ es un punto en el que no se comete extrapolación, luego $\mathbf{x}_0(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{x}_0^T < 1$.

$$X(X^TX)^{-1}X^T = H \quad \text{hat}$$

$$¿\, h_{00} = x_0(X^TX)^{-1}x_0^T < 1 \,?$$

un punto de intrapolación

$$\boxed{h_{00} < \max\{h_{ii}\}} < 1$$

$\max\{h_{ii}\}$ no puede ser un punto de balanceo

b) Considere a la entrada h_{ii} de la matriz $n \times n$ definida como: $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$, se tiene que $\sum_{i=1}^{n} h_{ii}$ es igual al número de covariables en el modelo.

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} & \\ h_{21} & h_{22} & & \\ & & \ddots & \\ & & & \end{bmatrix} \qquad \sum_{i=1}^{n} h_{ii} = k$$

$$X_{n\times p} = \begin{bmatrix} 1 & x_1 & x_2 & x_3 & & x_k \\ \vdots & & & & & \\ \vdots & & & & & \end{bmatrix}_{n\times p}$$

$$x^T_{p\times n}$$

$$Tr(H) = Tr(\underbrace{X}_{A}\,\underbrace{(X^TX)^{-1}X^T}_{B})$$

$$Tr(AB) = Tr(BA)$$

$$Tr(H) = Tr((X^TX)^{-1}X^TX) = Tr(I_{p\times p}) = p$$

$$\left(\left(X^T_{p\times n} X_{n\times p}\right)^{-1}_{p\times p} \left(X^T X\right)_{p\times p}\right)_{p\times p}$$

$$I_{p\times p} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & & & & & \\ 0 & & 1 & & & & \\ 0 & & & 1 & & & \\ 0 & & & & \ddots & & \\ & & & & & & 1 \end{bmatrix}_{p\times p}$$

$p-1$

$$\sum_{i=1}^{n} h_{ii} = Tr(H) = p$$

c) En un modelo de regresión suponga que $2p > n$, luego el criterio para halllar puntos de balanceo es si para el dato x_i su $h_{ii} > 2\frac{p}{n}$. F

$$h_{ii} > \frac{2p}{n} \qquad \frac{2p}{n} > 1$$

$$2\frac{p}{n} < 1$$

$$h_{ii} < 1$$

d) Una observación es influencial si $|\mathrm{DFFITS}_i| > 2\sqrt{\frac{k}{n}}$.

$$|\mathrm{Dffits}_i| > 2\sqrt{\frac{p}{n}}$$

$$\overset{Y}{altura_i} = \beta_0 + \beta_1 \overset{X_1}{peso_i} + \beta_2 \overset{X_2}{muñeca_i} + \beta_3 \overset{X_3}{biceps_i} + \varepsilon_i$$

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

$$\hat\beta_0 \quad \hat\beta_1 \quad \hat\beta_2 \quad \hat\beta_3$$

$\alpha = 0{,}05$

Se plantea la

$$\begin{cases} H_0: \boxed{\beta_2 = \beta_3 = 0} \\ H_a: \text{Alguno} \neq 0 \end{cases}$$ ✓

$$F_0 = \frac{\dfrac{SSE(MR) - SSE(MF)}{2}}{\dfrac{SSE(MF)}{32}} \sim F_{2,32}$$

MF: $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$

MR: $Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i$

	Sum_of_Squares	DF	Mean_Square	F_Value	P_value
Model	2469.74	3	823.2461	18.7654	3.32021e-07
Error	1403.85	32	43.8703		

SSE(MF) ; $MSE(MF) = \dfrac{SSE(MF)}{32}$ MF

	Sum_of_Squares	DF	Mean_Square	F_Value	P_value
Model	2462.85	1	2462.8465	59.3565	5.85272e-09
Error	1410.74	34	41.4924		

MR

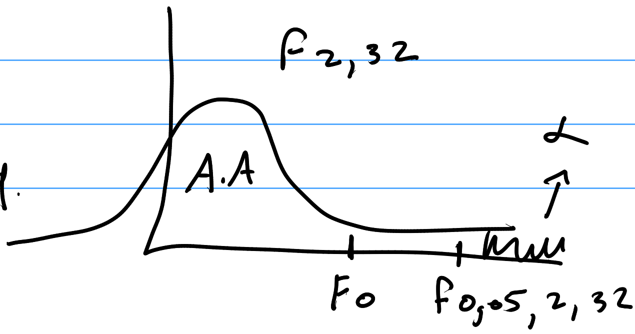
SSE(MR)

$$F_0 = \frac{\dfrac{1410{,}74 - 1403{,}85}{2}}{43{,}8703}$$

$F_0 = 1{,}5\ F$

$$= 0{,}0785$$

$$F_{0{,}05,2,32} = 3{,}29$$



$|Dffits_i|$

vif

Cook.D

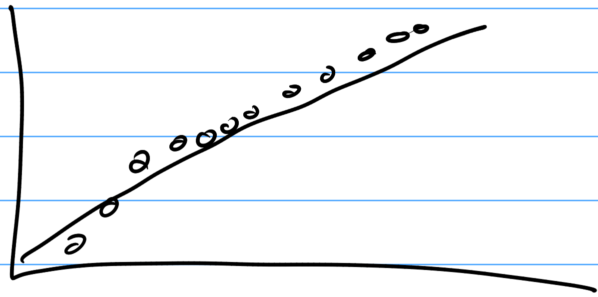
$\hat{\beta} \rightarrow \hat{\beta}_1 = 45$

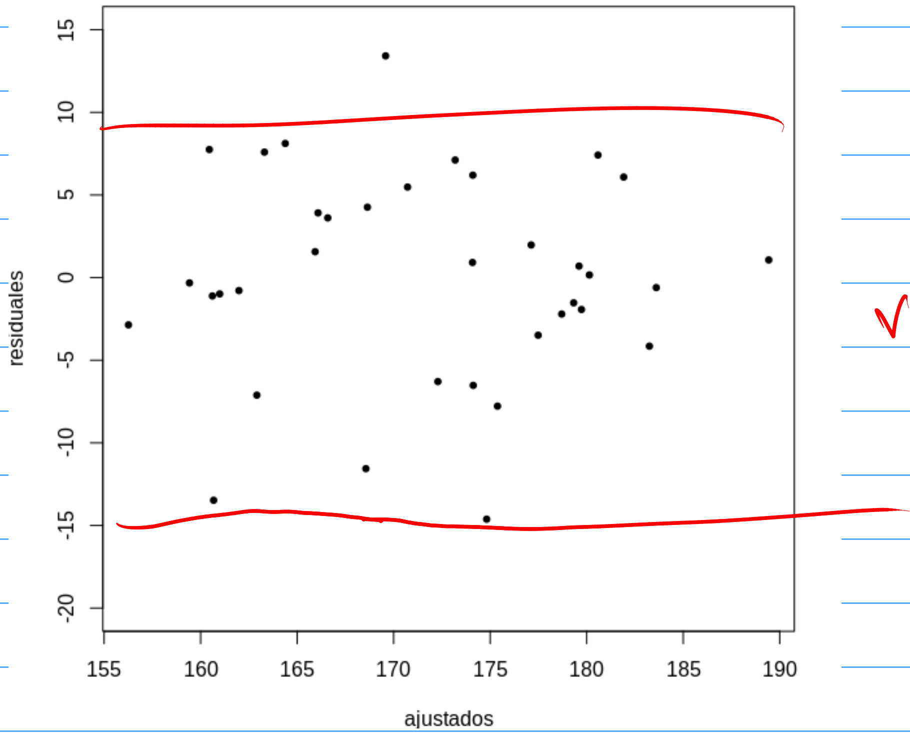
$\hat{\beta}_1 = 36$

R^2

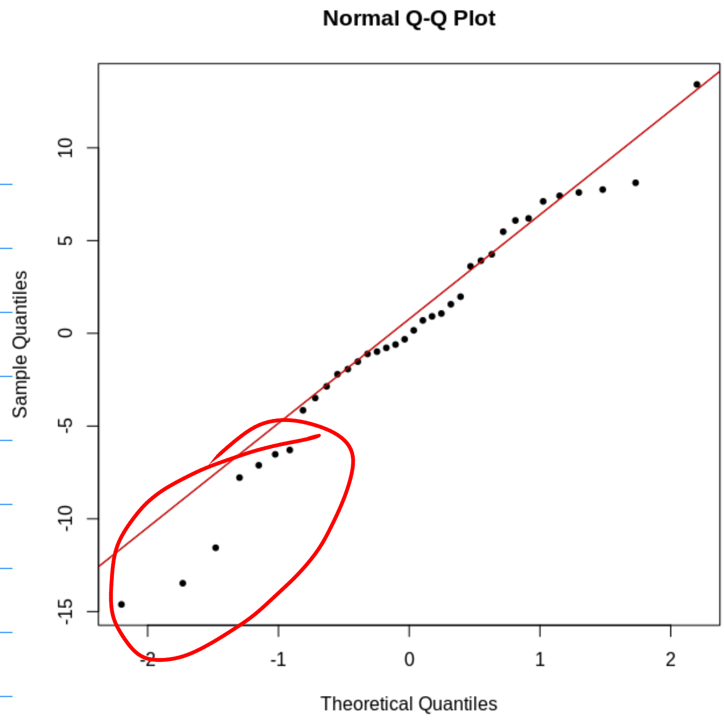
No normalidad

h_{ii}





Var cte



No Normal

var cte

No normalidad → No es válido

$$h_{00} < \underbrace{\max\{h_{ii}\}}_{\text{No puede ser } > \frac{2p}{n}} = 0{,}16$$

no puede ser punto de balanceo

$$h_{00} < \max\{h_{ii}\}$$