# DFBETAS

De manera similar a la distancia de Cook, permite evaluar cuánto cambia el j-ésimo coeficiente de regresión si se elimina la i-ésima observación. Se tiene que:

matriz del modelo

i-ésima fila de X

i-ésimo error

$$\hat{\beta} - \hat{\beta}_{(i)} = \frac{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_i\epsilon_i}{1 - h_{ii}}$$

i-ésimo valor hat

i-ésimo valor de la diagonal de la matriz hat

donde:

- $\hat{\beta}$ es la estimación de $\beta$ en el conjunto de datos completo
- $\hat{\beta}_{(i)}$ cuando se elimina la i-ésima observación

$$
\begin{array}{cc|c}
 & \text{cook} & \\
\text{Obs} & D_i & \begin{array}{ccc} \text{DFBeta}_0 & \text{DFBeta}_1 & \text{DFBeta}_2 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ \vdots \\ n \end{array} & \begin{array}{c} D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_n \end{array} & \left[\begin{array}{ccc} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ & & \\ \bullet & \bullet & \bullet \end{array}\right]
\end{array}
$$

$n \times K$

$K$: # coeficientes

$n$: # observaciones

| | $\beta_0$ (Intercept) | $\beta_1$ x |
|---|---|---|
| 1 | -0.275691211 | 0.1929583403 |
| 2 | -0.123523176 | 0.0672754784 |
| 3 | -0.006071150 | -0.0102087577 |
| 4 | -0.203638872 | 0.1288329898 |
| 5 | -0.077603339 | 0.0194540553 |
| 6 | -0.075649967 | 0.0321127060 |
| 7 | -0.018349996 | -0.0004230407 |
| 8 | 4.053966371 | -1.7208709421 |
| 9 | -0.119122513 | 0.0648787078 |
| 10 | 0.639178659 | -0.9865042898 |
| 11 | -0.018923427 | 0.0065534903 |
| 12 | 0.003665428 | 0.0156715670 |

$12 \times 2$

DFBeta0 más alejado de 0 : 8

DFBeta1 más alejado de 0 : 8

# DFFITS

¿Qué sucede con el valor de $\hat{y}$ si se elimina la i-ésima observación?

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{\hat{\sigma}^2_{(i)} h_{ii}}} = t_i \times \sqrt{\frac{h_{ii}}{1 - h_{ii}}}$$

i-ésimo valor diagonal haT

CME al quitar la i-ésima obs

| Obs | $\hat{y}$ |
|---|---|
| 1 | $\hat{y}_1$ |
| 2 | $\hat{y}_2$ |
| 3 | $\hat{y}_3$ |
| $\vdots$ | |
| $n$ | $\hat{y}_n$ |

| Obs | |
|---|---|
| • | $\hat{y}_{(1)}$ |
| 2 | |
| 3 | |
| $\vdots$ | |
| $n$ | |

$$DFFITS_i = \frac{\hat{y}_i - \hat{y}_{(i)}}{\sqrt{\hat{\sigma}^2_{(i)} h_{ii}}} = t_i \times \sqrt{\frac{h_{ii}}{1 - h_{ii}}}$$

```
modelo |> predict() -> y_pred
x_sin1 = x[-1]
y_sin1 = y[-1]
modelo_sin1 = lm(y_sin1 ~ x_sin1, data.frame(x_sin1,y_sin1))
modelo_sin1 |> predict(data.frame(x_sin1 = x[1])) -> y_pred_sin
(modelo_sin1 |> sigma())**2 -> cme_sin
modelo |> hatvalues() -> h
(y_pred[1]-y_pred_sin[1])/sqrt(cme_sin*h[1])
```

```
         1
-0.2832984
```

```
modelo |> rstudent() -> t
t[1]*sqrt(h[1]/(1-h[1]))
```

```
         1
-0.2832984
```

# COVRATIO

$$COVRATIO_i = \frac{\left|\hat{\sigma}^2_{(i)}(\mathbf{X}'_{(i)}\mathbf{X}_{(i)})^{-1}\right|}{|\hat{\sigma}^2(\mathbf{X}'\mathbf{X})^{-1}|} = \left(\frac{\hat{\sigma}^2_{(i)}}{\hat{\sigma}^2}\right)^k \left(\frac{1}{1-h_{ii}}\right)$$

¿Qué sucede si COVRATIO>1? ¿Y si es menor a 1? ¿Qué papel cumplen los leverages en los COVRATIOs?

$COVRATIO > 1 \Rightarrow \left|\hat{\sigma}^2_{(i)}(X'_{(i)}X_{(i)})^{-1}\right| > \left|\hat{\sigma}^2(X'X)^{-1}\right|$ Al retirar la i-ésima observación, la variabilidad de $\hat{\beta}$ aumenta

$COVRATIO < 1 \Rightarrow \left|\hat{\sigma}^2_{(i)}(X'_{(i)}X_{(i)})^{-1}\right| < \left|\hat{\sigma}^2(X'X)^{-1}\right|$ Al retirar la i-ésima observación, la variabilidad de $\hat{\beta}$ disminuye

leverage $(h_{ii}) \uparrow$ COVRATIO $\uparrow$

Resumen proporcionado por R:

```r
modelo |> influence.measures()
```



```
Influence measures of
     lm(formula = y ~ x) :

       dfbeta0          dfbeta1                covratio                hii
     dfb.1_      dfb.x    dffit  cov.r   cook.d    hat inf
1  -0.27569  0.192958 -0.2833 1.32982 0.042527 0.1554
2  -0.12352  0.067275 -0.1390 1.33788 0.010548 0.1088
3  -0.00607 -0.010209 -0.0281 1.36340 0.000439 0.0960
4  -0.20364  0.128833 -0.2163 1.32335 0.025104 0.1292
5  -0.07760  0.019454 -0.1122 1.31126 0.006896 0.0859
6  -0.07565  0.032113 -0.0938 1.33808 0.004838 0.0944
7  -0.01835 -0.000423 -0.0337 1.34310 0.000628 0.0833
8   4.05397 -1.720871  5.0243 0.00175 0.502575 0.0944   *
9  -0.11912  0.064879 -0.1340 1.34113 0.009821 0.1088
10  0.63918 -0.986504 -1.0389 7.79559 0.586971 0.8482   *
11 -0.01892  0.006553 -0.0251 1.35387 0.000350 0.0894
12  0.00367  0.015672  0.0339 1.37800 0.000637 0.1060
```