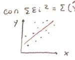
reg Lin

Sunday, March 13, 2022 11:13 PM

Para verificar q' tan buena es una regresión usamos R^2

$$\underbrace{\sum(y_i-\bar{y})^2}_{TSS} = \underbrace{\sum(\hat{y}_i-\bar{y})}_{ESS} + \underbrace{\sum \varepsilon_i^2}_{RSS} \quad \text{con } \sum\varepsilon_i^2 = \sum(\hat{y}_i - y_i)^2$$



$$R^2 = 1 - \frac{\sum\varepsilon_i^2}{TSS}$$

$R^2 = 1$: Puedo explicar en un 100% la variabilidad de y solo con x.

$R^2 = 0$: Cuando mi recta de regresión tiene $\beta_1 = 0$ (pendiente 0) x no me dice nada sobre y.

¿Qué pasa si todos los $\beta_i = 0$?

Aunque incluya muchas variables q' no tienen nada que ver, a más variables, aumenta el R^2 aunq' sea ruido.

↳ Necesito ver q' todas mis vars sean significativas.

Distribución muestral de $\hat{\beta}$ y $\hat{\varepsilon}$

$\hat{\varepsilon} \sim N(0, \sigma^2 \mathbb{I})$

$\hat{\beta} = (\mathbb{Z}'\mathbb{Z})^{-1}\mathbb{Z}'\mathbb{Y} \sim N_{r+1}(\beta, \sigma^2(\mathbb{Z}'\mathbb{Z})^{-1})$ → $r = \#$ vars.

↳ $\hat{\beta}$ es buena con respecto a β

Región de confianza de β.

$(\beta - \hat{\beta})'(\mathbb{Z}'\mathbb{Z})(\beta - \hat{\beta}) \leq (r+1)\, S^2\, F_{r+1,\, n-r-1}$ → $\#n$: obs.

$S^2 \to \dfrac{\hat{\varepsilon}'\hat{\varepsilon}}{n-r-1}$

$\mathbb{Z}$ tiene q' ser de rango completo

Las direcciones de los ejes son los vecs propios de $\mathbb{Z}'\mathbb{Z}$.

$\hat{\beta}_i \pm \sqrt{\widehat{var}(\hat{\beta}_i)}\sqrt{(r+1)F_{r+1,\, n-r-1}(\alpha)}$ → simultáneos

↳ estimación de la varianza de $\hat{\beta}_i$ = i'ésimo elem de la diagonal de $S^2(\mathbb{Z}'\mathbb{Z})^{-1}$

Si en este intervalo está 0, no rechazo la hipótesis de q' $\beta_i = 0$ y por tanto la puedo sacar de mi regresión pues no es significativo.

$\hat{\beta}_i \pm t_{n-1}\left(\frac{\alpha}{2}\right)\sqrt{\widehat{var}(\hat{\beta}_i)}$ → uno a uno.