

RECTA DE REGRESIÓN

Secuencia de cálculo

(1°) Generación de la tabla de frecuencias: $x_i, y_i, x_i y_i, x_i^2, y_i^2 \rightarrow \sum_{\forall i} \dots$

(2°) Cálculo de las "sumas de cuadrados":

$$SS_{XX} = \sum_{\forall i} (x_i - \bar{x})^2 = \sum_{\forall i} x_i^2 - \frac{1}{n} \left(\sum_{\forall i} x_i \right)^2 \quad [1]$$

$$SS_{YY} = \sum_{\forall i} (y_i - \bar{y})^2 = \sum_{\forall i} y_i^2 - \frac{1}{n} \left(\sum_{\forall i} y_i \right)^2 \quad [2]$$

$$SS_{XY} = \sum_{\forall i} (x_i - \bar{x})(y_i - \bar{y}) = \sum_{\forall i} x_i y_i - \frac{1}{n} \sum_{\forall i} x_i \sum_{\forall i} y_i \quad [3]$$

(3°) Cálculo de los parámetros de la recta:

$$\hat{\beta}_1 = \frac{SS_{XY}}{SS_{XX}} \quad [4] \quad \leftarrow \hat{\beta}_1 = \frac{dy}{dx}$$

$$\hat{\beta}_0 = \frac{1}{n} \left(\sum_{\forall i} y_i - \hat{\beta}_1 \sum_{\forall i} x_i \right) \quad [5]$$

$\underbrace{\hspace{10em}}_{\hat{\beta}_0 = y - \hat{\beta}_1 x}$

(4°) Cálculo de la "suma de los cuadrados de los errores":

$$SS_E = \sum_{\forall i} (y_i - \hat{y}_i)^2 = SS_{YY} - \hat{\beta}_1 SS_{XY} \quad [6]$$

(5°) Cálculo de los "coeficientes de correlación (r) y de determinación (r al cuadrado)":

$$r = \frac{SS_{XY}}{\sqrt{SS_{XX}} \sqrt{SS_{YY}}} \quad (\text{adimensional}) \quad [7]$$

Notar que, por el contrario, $\hat{\beta}_1$ no es adimensional:

$$r^2 = \frac{SS_{YY} - SS_E}{SS_{YY}} = \frac{SS_{XY}^2}{SS_{XX} SS_{YY}} \quad [8]$$

(6°) Cálculo de la "varianza" de los errores ($\varepsilon_i = y_i - \hat{y}_i$) y de la variable aleatoria ($\hat{\beta}_1$):

$$\sigma_\varepsilon^2 = \frac{SS_E}{n-2}$$

$$\sigma_{\hat{\beta}_1}^2 = \frac{\sigma_\varepsilon^2}{SS_{XX}} = \frac{1}{n-2} \frac{SS_E}{SS_{XX}} \quad [9]$$