Fórmulas para regresión Lineal Simple

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} y_i x_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\text{donde} \quad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} \quad \text{y} \quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

$$S_{xy} = \sum_{i=1}^{n} y_i (x_i - \bar{x}) = \sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n}$$

Entonces con esta notación podemos escribir $\hat{\beta}_1 = \dfrac{S_{xy}}{S_{xx}}$

$$\hat{\sigma}^2 = \frac{S_{yy} - \dfrac{S_{xy}^2}{S_{xx}}}{n-2}$$

$$\frac{\hat{\beta}_0 - \beta_0}{\sqrt{\hat{\sigma}^2\left(\dfrac{1}{n} + \dfrac{\bar{x}^2}{S_{xx}}\right)}} \sim t_{\text{n-2}} \qquad \text{y} \qquad \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\dfrac{\hat{\sigma}^2}{S_{xx}}}} \sim t_{\text{n-2}}$$

Intervalos de confianza para β_1

$$\left[\hat{\beta}_1 - t_{\frac{\alpha}{2},n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}};\ \hat{\beta}_1 + t_{\frac{\alpha}{2},n-2}\sqrt{\frac{\hat{\sigma}^2}{S_{xx}}}\right]$$

INTERVALO DE CONFIANZA PARA β_0

$$\left[\hat{\beta}_0 - t_{\frac{\alpha}{2},n-2}\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)};\ \hat{\beta}_0 + t_{\frac{\alpha}{2},n-2}\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{\bar{x}^2}{S_{xx}}\right)}\right]$$

- Intervalo de confianza para la respuesta media

$$\left[\hat{\beta}_0+\hat{\beta}_1 x_0-t_{\frac{\alpha}{2},n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]};\ \hat{\beta}_0+\hat{\beta}_1 x_0+t_{\frac{\alpha}{2},n-2}\sqrt{\hat{\sigma}^2\left[\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]}\right]$$

Intervalos de predicción para PREDICCIÓN de futuras observaciones

$$\left[\hat{Y}_0-t_{\frac{\alpha}{2},n-2}\,\hat{\sigma}\,\sqrt{\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]};\hat{Y}_0+t_{\frac{\alpha}{2},n-2}\,\hat{\sigma}\,\sqrt{\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{S_{xx}}\right]}\right]$$

coeficiente de determinación.

$$R^2=1-\frac{SS_R}{S_{YY}}$$

coeficiente de correlación muestral

$$\frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i-\bar{X})^2\sum_{i=1}^{n}(Y_i-\bar{Y})^2}}=\frac{S_{XY}}{\sqrt{S_{XX}S_{YY}}}=R$$

También se puede calcular como

$$\frac{n\left(\sum xy\right)-\left(\sum x\right)\left(\sum y\right)}{\sqrt{\left[n\left(\sum x^2\right)-\left(\sum x\right)^2\right]\left[n\left(\sum y^2\right)-\left(\sum y\right)^2\right]}}$$