Ejercicio 3

Equipo

2024-03-26

3. Expresión alternativa para R^2

Considere el coeficiente de correlación muestral o de Pearson para dos variables X y Y :

$$r_{xy}=\frac{\sum_{i=1}^{n}\left(x_i-\bar{X}\right)\left(y_i-\bar{Y}\right)}{\left(\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2\sum_{i=1}^{n}\left(y_i-\bar{Y}\right)^2\right)^{1/2}}$$

.

Considere el modelo de regresión

$$y_i=\beta_0+\beta_1 x_i+\varepsilon_i$$

.

a. Demuestre que:

$$R^2=r_{xy}^2$$

b. Demuestre que $t^*=t$, donde t es la estadística usada para contrastar " $H_0:\beta_1=0$ vs $H_1:\beta_1\neq 0$ ":

$$t=\frac{\widehat{\beta}_1}{\sqrt{\frac{\widehat{\sigma}^2}{\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2}}}.$$

Por otra parte, $t^*=\frac{r_{xy}\sqrt{n-2}}{\sqrt{1-r_{xy}^2}}$ es la estadística usada para contrastar " $H_0:\rho=0$ vs $H_a:\rho\neq 0$ " cuando (X,Y) sigue una distribución normal bivariada con coeficiente de correlación $\rho=\rho_{xy}$.

SOLUCIÓN a.

Recordemos que: $R^2=\frac{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}=\frac{SCR}{SCT}$

Y por la expresión (77) de las notas sabemos que:

$$\begin{aligned}SCR=\hat{\beta}_1^2 SS_x&=\left(\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\right)^2\cdot\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\\&=\frac{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\end{aligned}$$

Entonces tenemos:

$$\begin{aligned}
R^2 &= \frac{SCR}{SCT} \\
&= \frac{\frac{\left(\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\right)^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2} \\
&= \frac{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}
\end{aligned}$$

Por tanto tenemos que:

$$R^2 = \frac{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}$$

Ahora, notemos que:

$$\begin{aligned}
r_{xy}^2 &= \left(\frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2}}\right)^2 \\
&= \frac{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\right)^2}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2 \cdot \sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2} \\
&= R^2
\end{aligned}$$

$$\therefore R^2 = r_{xy}^2$$

SOLUCIÓN b.

Primero notemos lo siguiente:

$$\begin{aligned}
t &= \frac{\widehat{\beta}_1}{\sqrt{\frac{\widehat{\sigma}^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}} \\
&= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\sqrt{\frac{\frac{\sum_{i=1}^{n}(y_i-\widehat{y_i})^2}{n-2}}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}} \\
&= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\frac{\sqrt{\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2}}{\sqrt{n-2}\cdot\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}} \\
&= \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\cdot\sqrt{n-2}\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}}{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\cdot\sqrt{\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2}} \\
&= \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\cdot\sqrt{n-2}}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2}} \\
&= \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\cdot\sqrt{n-2}}{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\cdot\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2\right)^{\frac{1}{2}}}
\end{aligned}$$

Así: $t=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\cdot\sqrt{n-2}}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\cdot\sum_{i=1}^{n}(y_i-\widehat{y_i})^2\right)^{\frac{1}{2}}}$

Por otro lado tenemos que:

$$
\begin{aligned}
t^* &= \frac{r_{xy}\cdot\sqrt{n-2}}{\sqrt{1-r_{xy}^2}} \\
&= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\left(\sum_{i=1}^{n}(x_i-\bar{x})^2\cdot\sum_{i=1}^{n}(y_i-\bar{y_i})^2\right)^{\frac{1}{2}}}\cdot\sqrt{n-2}}{\sqrt{1-\frac{\sum_{i=1}^{n}(\widehat{y_i}-\bar{y_i})^2}{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}} \\
&= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\cdot\sqrt{n-2}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}}{\sqrt{\frac{\sum_{i=1}^{n}(y_I-\bar{y_i})^2-\sum_{i=1}^{n}(\widehat{y_i}-\bar{y_i})^2}{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}}
\end{aligned}
$$

Ahora, por la expresión (68) de las notas de clase sabemos que:

$$
\sum_{i=1}^{n}(y_i-\bar{y_i})^2=\sum_{i=1}^{n}(\widehat{y_i}-\bar{y_i})^2+\sum_{i=1}^{n}(y_i-\widehat{y_i})^2 \Longrightarrow \sum_{i=1}^{n}(y_i-\bar{y_i})^2-\sum_{i=1}^{n}(\widehat{y_i}-\bar{y_i})^2=\sum_{i=1}^{n}(y_i-\widehat{y_i})^2
$$

Así:

$$
\begin{aligned}
t^* &= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\cdot\sqrt{n-2}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}}{\sqrt{\frac{\sum_{i=1}^{n}(y_i-\widehat{y_i})^2}{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}} \\
&= \frac{\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})\cdot\sqrt{n-2}}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}}{\frac{\sqrt{\sum_{i=1}^{n}(y_i-\widehat{y_i})^2}}{\sqrt{\sum_{i=1}^{n}(y_i-\bar{y_i})^2}}} \\
&= \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\cdot\sqrt{n-2}}{\sqrt{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\widehat{y_i})^2}} \\
&= \frac{\sum_{i=1}^{n}\left(x_i-\bar{x}\right)\left(y_i-\bar{y}\right)\cdot\sqrt{n-2}}{\left(\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2\cdot\sum_{i=1}^{n}(y_i-\widehat{y_i})^2\right)^{\frac{1}{2}}} \\
&= t
\end{aligned}
$$

$$
\therefore t^*=t
$$