Súper nota

estadística descriptiva

Bryant reyes robles

02-08-24

4.4 covarianza

$$Covarianza = \frac{\sum(\bar{X}-X)*(\bar{Y}-Y)}{n-1} = \frac{290.8}{19} = 15.30$$

$$r = \frac{covarianza}{S_x * S_y} = \frac{15.30}{8.087*2.137} = 0.885$$

S_x = Desviación típica x = 8.087

S_y = Desviación típica y = 2.137

4.5 .- test de hipótesis de r

Tras realizar el cálculo del coeficiente de correlación de Pearson (r) debemos determinar si dicho coeficiente es estadísticamente diferente de cero. Para dicho calculo se aplica un test basado en la distribución de la t de student.

$$Error\ estandard\ de\ r = \sqrt{\frac{1-r^2}{n-2}}$$

Si el valor del r calculado (en el ejemplo previo r = 0.885) supera al valor del error
estándar multiplicado por la t de Student con n-2 grados de libertad, diremos que el coeficiente de correlación es significativo.
El nivel de significación viene dado por la decisión que adoptemos al buscar el valor en la tabla de la t de Student.
En el ejemplo previo con 20 niños, los grados de libertad son 18 y el valor de la tabla de la t de student para una seguridad del 95% es de 2.10 y para un 99% de seguridad el valor es 2.88.

$$Error\ estandard\ de\ r = \sqrt{\frac{1-0.885^2}{20-2}} = 0.109$$

4.6.- interpretación de la correlación

El coeficiente de correlación como previamente se indicó oscila entre -1 y + l encontrándose en medio el valor 0 que indica que no existe asociación lineal entre las dos variables a estudio.
Un coeficiente de valor reducido no indica necesariamente que no exista correlación ya que las variables pueden presentar una relación no lineal como puede ser el peso del recién nacido y el tiempo de gestación. En este caso el r infra estima la asociación al medirse linealmente.
Los métodos no paramétricos estarían mejor utilizados en este caso para mostrar si las variables tienden a elevarse conjuntamente o a moverse en direcciones diferentes.
La significancia estadística de un coeficiente debe tenerse en cuenta conjuntamente con la relevancia clínica del fenómeno que estudiamos ya que coeficientes de 0.5 a 0.7 tienden ya a ser significativos como muestras pequeñas. Es por ello muy útil calcular el intervalo de confianza del r ya que en muestras pequeñas tenderá a ser amplio.

$$r_s = \frac{n\sum r_x r_y - \sum r_x \sum r_y}{\sqrt{\left[n\sum r_x^2 - \left(\sum r_x\right)^2\right]\left[n\sum r_y^2 - \left(\sum r_y\right)^2\right]}}$$

$$\sum r_x = \sum r_y = 55 \qquad \sum r_x^2 = \sum r_y^2 = 385$$

$$\sum r_x r_y = 2\ (8) + 5\ (3) + 8\ (6) + \dots + 4\ (1) = 325$$

$$r_s = \frac{10(325) - 55\ (55)}{\sqrt{\left[10\ (385) - 55^2\right]\left[10\ (385) - 55^2\right]}} = 0.27$$

Bibliografía

https://plataformaeducativauds.com.mx/assets/docs/libro/LNU/2e38faf807e4310316facdc1b7d23494-LC-LNU302%20ESTADISTICA%20DESCRIPTIVA%20EN%20NUTRICION.pdf