



SUPER NOTA

Nombre del alumno: Miguel Angel Perez Gordillo

Unidad: Cuarto Parcial

Materia: Estadística Descriptiva

Docente:

**Licenciatura: Administración Y Estrategia De
Negocios**

Tercer Cuatrimestre

COVARIANZA

COVARIANZA

$$\text{Cov}(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x}) * (y_i - \bar{y})}{n - 1}$$

ES EL VALOR A TRAVÉS DEL CUAL SE REFLEJA EN QUÉ CUANTÍA DON VARIABLES CUALESQUIERA VARÍAN DE FORMA CONJUNTA RESPECTO DE SUS MEDIAS ARITMÉTICAS

$$X = (x_1, x_2, x_3) = (0, 4, 8) \\ Y = (y_1, y_2, y_3) = (5, 5, 8)$$

$$\bar{x} = \frac{0 + 4 + 8}{3} = 4$$

$$\bar{y} = \frac{5 + 5 + 8}{3} = 6$$

$$\text{Cov}(X, Y) = \frac{(0-4) \times (5-6) + (4-4) \times (5-6) + (8-4) \times (8-6)}{3}$$

$$\text{Cov}(X, Y) = 4$$

TEST DE HIPOTESIS 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. 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.

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

Como quiera que $r = 0.885 > 2.10 * 0.109 = 2.30$ podemos asegurar que el coeficiente de correlación es significativo ($p < 0.05$). Si aplicamos el valor obtenido en la tabla de la t de Student para una seguridad del 99% ($t = 2.88$) observamos que como $r = 0.885$ sigue siendo $> 2.88 * 0.109 = 0.313$ podemos a su vez asegurar que el coeficiente es significativo ($p < 0.001$). Este proceso de razonamiento es válido tanto para muestras pequeñas como para muestras grandes. En esta última situación podemos comprobar en la tabla de la t de student que para una seguridad del 95% el valor es 1.96 y para una seguridad del 99% el valor es 2.58.

INTERPRETACION DE LA CORRELACION

EL COEFICIENTE DE CORRELACIÓN COMO PREVIAMENTE SE INDICÓ OSCILA ENTRE -1 Y +1 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

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Rango de valores de r_{xy}	Interpretación
$0.00 \leq r_{xy} < 0.10$	Correlación nula
$0.10 \leq r_{xy} < 0.30$	Correlación débil
$0.30 \leq r_{xy} < 0.50$	Correlación moderada
$0.50 \leq r_{xy} < 1.00$	Correlación fuerte