

TEMATICA: Análisis de correlación

ESTUDIANTE:

Códigos:

Fecha de entrega:

## ***Afirmaciones extraordinarias requieren pruebas extraordinarias***

***David Hume (1711-1776)***

### **1. CONTEXTUALIZACIÓN.**

El valor máximo de la presión arterial<sup>1</sup> cuando el corazón se contrae, se denomina presión arterial sistólica, la cual depende del volumen del latido y de la rigidez de los vasos, cuya unidad de medida está dada por mililitros de mercurio. (Agudelo, et al 2012) evaluó la presión arterial, edad, género, talla y estrato socioeconómico de 2,807 estudiantes de diferentes instituciones del área urbana de la ciudad de Cali, donde concluyo que la presión arterial está determinada por diversos factores, entre ellos la talla, la edad y el género.

Teniendo en cuenta lo anterior se desea observar una posible relación lineal, entre la presión sistólica, edad y peso. Para evidenciar esta posible correlación se simularán datos en software estadístico R, teniendo en cuenta edades de 7 a 9 años, el peso y las presiones sistólicas comunes en estas edades.

Considerando lo anterior responda.

1. ¿Cuál es la población de interés?

2. ¿Cuáles son las variables de interés?

Estudiantes de 7 a 9 años en la ciudad de Cali

Peso, edad, presión sistólica

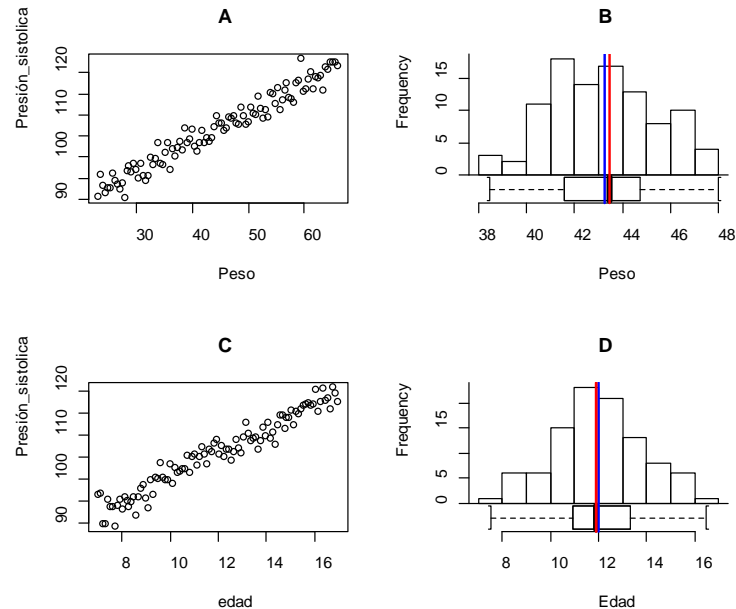
### **2. INSPECCIÓN DE GRAFICAS E IDENTIFICACIÓN DE TENDENCIAS.**

---

<sup>1</sup> La presión arterial es la presión que ejerce la sangre contra las paredes arteriales

Teniendo en cuenta que se desea observar la relación entre un par de variables, la dispersión y la simetría de los datos, se analizarán gráficos de dispersión y de barras.

**Teniendo en cuenta los gráficos, responde las preguntas del 1 al 5.**



- ¿Por qué cuando el peso o la edad aumenta la presión sistólica aumenta?  
*Con la edad, los vasos sanguíneos se endurecen y se vuelven más rígidos, lo que hace que la presión arterial aumente, La obesidad y el sobrepeso son factores de riesgo para la hipertensión arterial.*
- ¿Se podría afirmar que existe una relación entre las variables en los gráficos A y C?  
*Sí, la edad y el peso son síntomas que aumentan la presión sistólica según la gráfica*

- ¿Existe evidencia de datos atípicos en A y C?

*No se aprecian datos atípicos.*

- Teniendo en cuenta el gráfico A y C complete.

Variable predictora o regresora.

*Edad y peso*

Variable de respuesta.

*Presión sistólica*

El coeficiente de correlación de Pearson mide la relación lineal entre variables el cual varía en un intervalo  $[-1,1]$ , cuando el coeficiente es igual a cero indica que no existe asociación entre las variables, si el coeficiente se acerca a -1 indica una correlación negativa y cuando se acerca a uno indica una correlación positiva. Se calcula el coeficiente de correlación de Pearson como:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

5. Escriba y compile el siguiente código en el software R, cambie los valores de la medida de dispersión por 1, 4, 8, 10

```
#####Simulación edad vs presión sistólica#####

par(mfrow=c(2,2)) #dividimos el área en 4 partes

sigma=

n=100

edad=seq(7, 17, length=n)

Presión_sistolica= 72.303 + 2.790*edad + rnorm(n,0,sigma)

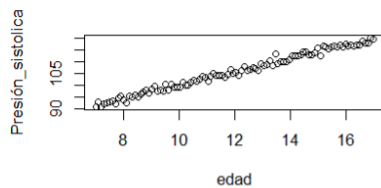
plot(edad,Presión_sistolica)

cor(edad,Presión_sistolica) #Coeficiente de Pearson
```

6. Registre los valores de correlación.

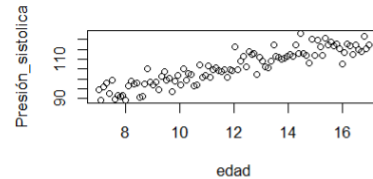
**Coeficiente de correlación cuando  $\sigma = 1$**

**$r = 0.9927131$**



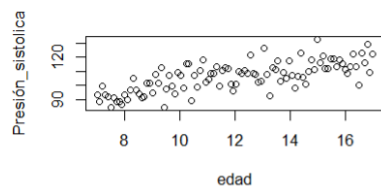
**Coeficiente de correlación cuando  $\sigma = 4$**

**$r = 0.9025731$**



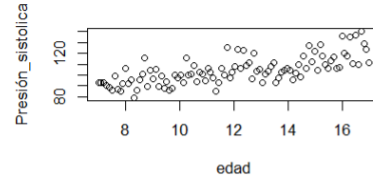
**Coeficiente de correlación cuando  $\sigma = 8$**

**$r = 0.7119522$**



**Coeficiente de correlación cuando  $\sigma = 10$**

**$r = 0.695623$**



7. Teniendo en cuenta los coeficientes de correlación lineal anteriores ¿Que se puede concluir con respecto a la edad y presión sistólica?

*Entre más se aumenta el coeficiente de correlación los datos se vuelven muy dispersos, respecto a la proporcionalidad inicial.*

Para identificar si la relación entre la presión sistólica y la edad o la presión sistólica y el peso es producto de una casualidad en la muestra o están relacionadas en la población, se realizará la prueba de hipótesis

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Generalmente se usa un test  $t$  denominado correlación del producto de los momentos.

$$r_t = r \sqrt{\frac{n-2}{1-r^2}} \sim t_{n-2},$$

8. Plantee la prueba de hipótesis para la correlación de la presión sistólica y el peso.

$H_0$ : No existe una relación o asociación entre la presión sistólica y la edad

$H_1$ : Existe una relación o asociación entre la presión sistólica y la edad

9. Escriba y compile el siguiente código en software R.

```
sigma=2  
  
n=100  
  
Peso=seq(23.26, 66.03,length=n)  
  
Presión_sistolica= 75.6096 + 0.6975*Peso + rnorm(n,0,sigma)  
  
cor.test(Peso,Presión_sistolica)
```

10. Con un nivel de significancia  $\alpha = 0.005$  el p-valor es de **2. 2e- 16**.

*Se rechaza la hipótesis nula  $H_0$ , se acepta la hipótesis alternativa  $H_1$ , es decir, que existe una relación o asociación entre el peso y la presión sistólica.*