

<b>Instituto Tecnológico de Costa Rica</b> <b>Escuela de Computación</b>  Programa en Ciencias de Datos <b>Curso: Estadística</b>  Profesor: Ph. D. Saúl Calderón Ramírez	<b>QUIZ 2</b> Entrega: Lunes 23 de Octubre, a través del TEC digital Debe subir un <i>pdf</i> con la respuesta.  Valor: 100 pts. Puntos Obtenidos: _____  Nota: _____
Nombre del (la) estudiante: <u>Gabriel Valentine Fonseca</u>	
Carné: _____	

1. Su equipo de ciencias de datos desea comparar un nuevo algoritmo de detección de grietas en piezas de acero de una línea de producción (**algoritmo A**). El sistema en uso (**algoritmo B**), en un conjunto de datos de prueba de  $N = 20$  logra las siguientes tasas de aciertos, por imagen:

0.72020	0.77719	0.72291	0.91679
0.79565	0.71765	0.69226	0.67864
0.88172	0.83181	0.83718	0.91055
0.70773	0.73465	0.80328	0.78082
1.0	0.78594	0.62688	0.87664

Mientras que el nuevo algoritmo de clasificación, reporta, en ese mismo conjunto de imágenes de prueba, los siguientes resultados:

0.66578	0.74194	0.65410	0.88453
1.0	0.65613	0.72290	0.84243
0.78553	0.83343	0.80881	0.74326
0.85252	0.78917	0.99126	0.83146
0.78544	0.72603	0.66458	0.78727

El líder del equipo plantea usar un ANOVA para verificar si existe una mejora estadísticamente significativa del algoritmo B sobre el A. Es por ello que se plantea la hipótesis nula de que ambos tratamientos tienen medias iguales, la alternativa, de que son diferentes.

- (a) **(60 puntos)** Calcule la intra e inter-varianza de ambos conjuntos de datos. Muestre todos los pasos intermedios, y corrobore el calculo con la implementación del código respectivo en pytorch. Adjunte el código y muestre el resultado.

**Conjunto A:**

- Número de observaciones en A ( $n$ ): 20
- Media de A ( $\mu_A$ ):

$$\mu_A = \frac{\sum_{i=1}^{20} x_i}{20} \approx \frac{15.79849}{20} \approx 0.78992$$

- Varianza de A ( $\sigma_A^2$ ):

$$\sigma_A^2 = \frac{\sum_{i=1}^{20} (x_i - \mu_A)^2}{19} \approx \frac{0.02125}{19} \approx 0.00112$$

**Conjunto B:**

- Número de observaciones en B ( $n$ ): 20
- Media de B ( $\mu_B$ ):

$$\mu_B = \frac{\sum_{i=1}^{20} y_i}{20} \approx \frac{15.76657}{20} \approx 0.78833$$

- Varianza de B ( $\sigma_B^2$ ):

$$\sigma_B^2 = \frac{\sum_{i=1}^{20} (y_i - \mu_B)^2}{19} \approx \frac{0.02211}{19} \approx 0.00116$$

**Intra-varianza ( $SS_{\text{error}}$ ):**

$$SS_{\text{error}} = (20 - 1) \cdot \sigma_A^2 + (20 - 1) \cdot \sigma_B^2 \approx 0.02212$$

**Inter-varianza ( $SS_{\text{treatments}}$ ):**

- Media global ( $\mu_{\text{global}}$ ):

$$\mu_{\text{global}} = \frac{\sum_{i=1}^{40} z_i}{40} \approx \frac{31.56506}{40} \approx 0.78913$$

- Suma de cuadrados entre grupos ( $SS_{\text{treatments}}$ ):

$$SS_{\text{treatments}} = 20 \cdot (\mu_A - \mu_{\text{global}})^2 + 20 \cdot (\mu_B - \mu_{\text{global}})^2 \approx 0.00102$$

### Código:

```
import torch

#datos
A = [0.72020, 0.77719, 0.72291, 0.91679, 0.79565, 0.71765, 0.69226, 0.67864,
      0.88172, 0.83181, 0.83718, 0.91055, 0.70773, 0.73465, 0.80328, 0.78082,
      1.0, 0.78594, 0.62688, 0.87664]

B = [0.66578, 0.74194, 0.65410, 0.88453, 1.0, 0.65613, 0.72290, 0.84243,
      0.78553, 0.83343, 0.80881, 0.74326, 0.85252, 0.78917, 0.99126, 0.83146,
      0.78544, 0.72603, 0.66458, 0.78727]

n = len(A)

#datos a tensores de PyTorch
tensor_A = torch.tensor(A)
tensor_B = torch.tensor(B)

suma = torch.sum(tensor_B)
meanmano = suma / 20

#medias
mean_A = tensor_A.mean()
mean_B = tensor_B.mean()
global_mean = (mean_A + mean_B) / 2

#varianzas muestrales
var_A = tensor_A.var(unbiased=True)
var_B = tensor_B.var(unbiased=True)

#intravarianza
ss_error_A = (n - 1) * var_A
ss_error_B = (n - 1) * var_B
ss_error = ss_error_A + ss_error_B

#intervarianza
ss_treatments = (n * (mean_A - global_mean)**2) + (n * (mean_B - global_mean)**2)

print("Varianza intra-grupos (SSerror):", ss_error.item())
print("Varianza inter-grupos (SSTreatments):", ss_treatments.item())

print("\nPrueba de resultados:")
print(f"Varianza Total = Intravarianza + Intervarianza = {ss_error.item() + ss_treatments.item()}")
```

### Output:

Intravarianza: 0.3528668284416199

Intervarianza: 2.5471328626736067e-05

Prueba de resultados:

Varianza Total = Intravarianza + Intervarianza = 0.3528922997702466

- (b) **(20 puntos)** Calcule el estadístico  $F_0$  y verifique si la hipótesis nula se acepta o rechaza. Verifique el resultado usando la función de `scipy.stats.f_oneway`.

- Varianza Intra-Grupos ( $SS_{\text{error}}$ ): 0.02212
- Varianza Inter-Grupos ( $SS_{\text{treatments}}$ ): 0.00102

El estadístico  $F_0$  se calcula como:

$$F_0 = \frac{SS_{\text{treatments}}}{SS_{\text{error}}} = \frac{0.00102}{0.02212} \approx 0.0459$$

Código:

```
import scipy.stats as stats

#estadístico F y p-value
F_statistic, p_value = stats.f_oneway(A, B)

#estadístico F manualmente
F_manual = 0.00102 / 0.02212

print(f"Estadístico F calculado manualmente: {F_manual:.5f}")
print(f"Estadístico F según f_oneway: {F_statistic:.5f}")
print(f"Valor p: {p_value:.5f}")

#nivel de significancia
alpha = 0.05
if p_value < alpha:
    print("Rechazamos la hipótesis nula: Hay una diferencia significativa entre los grupos.")
else:
    print("Aceptamos la hipótesis nula: No hay una diferencia significativa entre los grupos.")
```

Output:

```
Estadístico F según f_oneway: 0.00274
Valor p: 0.95850
Aceptamos la hipótesis nula: No hay una diferencia significativa entre los grupos.
```

- (c) **(20 puntos)** Usando estadísticos como la media, mediana y moda, explique si el líder del equipo tomó una decisión adecuada al proponer usar un ANOVA para tomar la decisión si el nuevo algoritmo presenta una mejora estadísticamente significativa frente al algoritmo en uso.

Basándonos en los estadísticos como la media, mediana y moda, podemos observar que:

Media de A: 0.7883

Media de B: 0.7899

- Ambas medias son muy cercanas, indicando una similitud en el rendimiento promedio entre los dos algoritmos.

Mediana:

- Ambos grupos tienen una mediana de aproximadamente 0.79, lo que sugiere que la mayoría de las observaciones están cerca de este valor en ambos casos.

Moda:

- Moda de A: 0.78594
- Moda de B: 0.66578, 0.65613 y 0.78553
- Se observa una moda mayor para el nuevo algoritmo A

Dado que la media, mediana y moda muestran similitudes entre los dos algoritmos, el líder del equipo tomó una decisión adecuada al proponer el uso de ANOVA para evaluar si existe una mejora significativa en la detección de grietas del nuevo algoritmo sobre el antiguo. Sin embargo, considero que al mostrar estadísticos tan similares no hay necesidad de gastar recursos en verificar si vale la pena usar el nuevo algoritmo.