OCE 313
TÉCNICAS DE ANÁLISIS NO PARAMÉTRICO

CLASE 5 – PERMUTACIÓN Y CORRELACIÓN

Dr. José Gallardo

Abril 2021

PONTIFICIA
UNIVERSIDAD
CATÓLICA DE
VALPARAÍSO

PLAN DE LA CLASE

- **Concepto de permutación:**

  ¿Qué es? ¿Por qué es importante? ¿Qué haremos?

- Permutación : caso 1, 2 y 3
- **Solución de un problema de permutación.**
- Prueba de correlación no paramétrica.
- **Tarea: Prácticas DataCamp**

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Permutación

¿Qué es?
Una **permutación** es una combinación **ordenada** de elementos.

¿Por qué es importante?
El concepto matemático de permutación está subyacente a muchos métodos de análisis no paramétricos.

¿Qué haremos?
Calcularemos las posibles permutaciones de los elementos de un conjunto de datos de un experimento aleatorio.

Permutación sin repetición

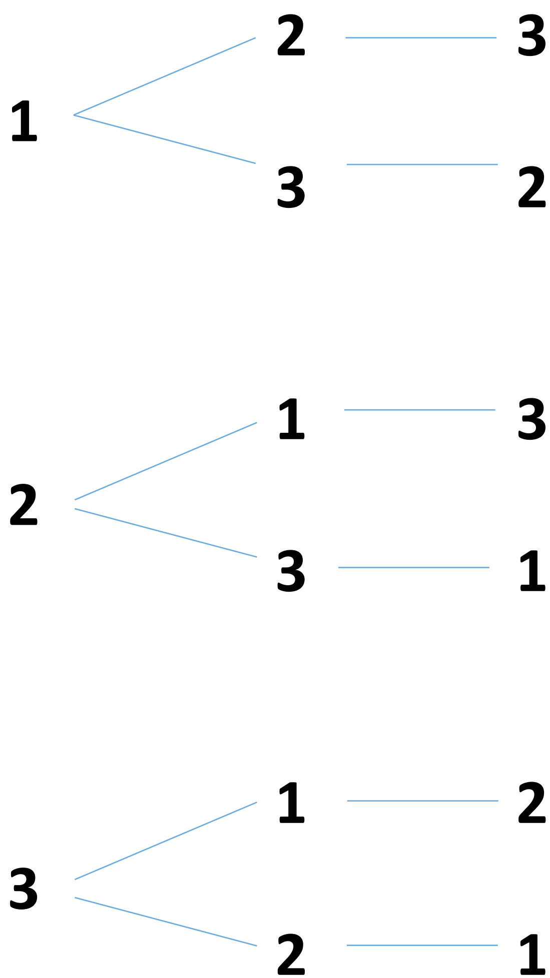
1.- Si para el conjunto {1,2} existen 2 permutaciones 1-2 y 2-1

¿Cuántas permutaciones de los elementos {1,2,3} existen?

R: Complete la siguiente tabla sin repetir números en cada fila.

A	B	C
1	2	3

Permutación sin repetición (Pn): caso 1



$P_n = n! =$ **n factorial**

$P_3 = 3! = 3 \times 2 \times 1 = 6$

¿Cuántas permutaciones existen para P_4?

$P_4 = \{1,2,3,4\}$

$= 4! = 24$

Permutación con repetición – caso 2

¿Cuántas claves diferentes existen en el candado?

Si cuando el candado tiene 3 filas con 10 números {0.. 9} existen 10x10x10 permutaciones o 1000 claves diferentes.

¿Cuántas permutaciones existen para los candados con 4 filas?

Respuesta

$n^r = 10^4 = 10.000$ permutaciones

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Permutación con repetición – caso 3

¿Cuántas palabras diferentes se pueden formar con las letras de la palabra GATA? Ej. {TAGA, AGAT, GTAA …}

Si, G = 1 vez; T = 1 vez y A = 2 veces

Entonces, existen

4! / 1! x 1! X 2! = 24 / 2 = 12

¿Cuántas palabras se pueden formar con la palabra BANANA?

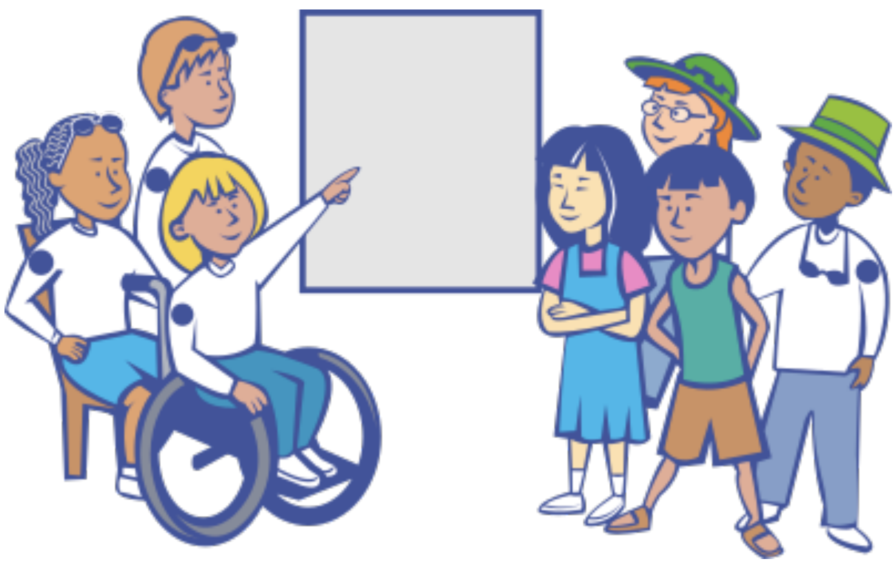
6! / 1! x 2! X 3! = 720 / 12 = 60

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Problema: Dividir una sala en grupos.

Tiene 7 personas en una sala y las desea dividir en 2 grupos, uno tendrá 3 y otro 4.

¿Cuántas formas de agrupar a los alumnos son posibles?



1.- ¿A qué caso se parece 1, 2 o 3?
2.- Esquematice el problema.
3.- Resuelva

Problema: Dividir una sala en grupos.

Esquema

Total alumnos						
1	2	3	4	5	6	7

$t_a! = 7! = 5040$

Grupo 1		
1	2	3

Grupo 2			
4	5	6	7

$g_1! = 3! = 6$

$g_2! = 4! = 24$

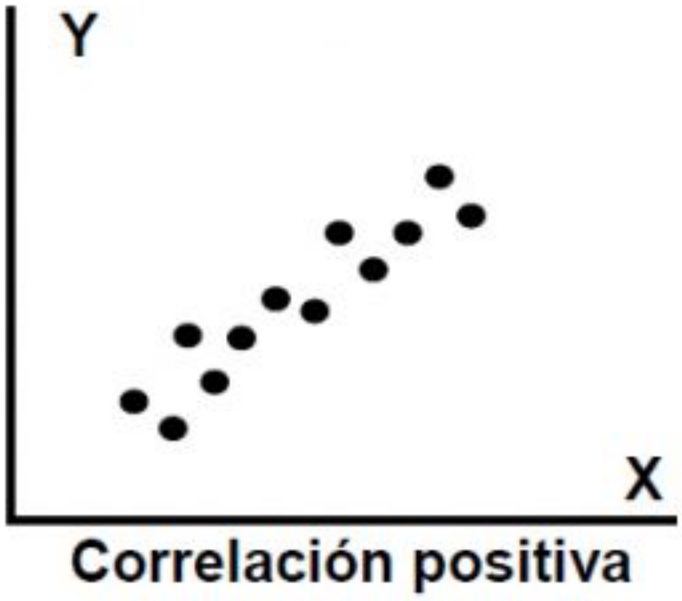
Solución

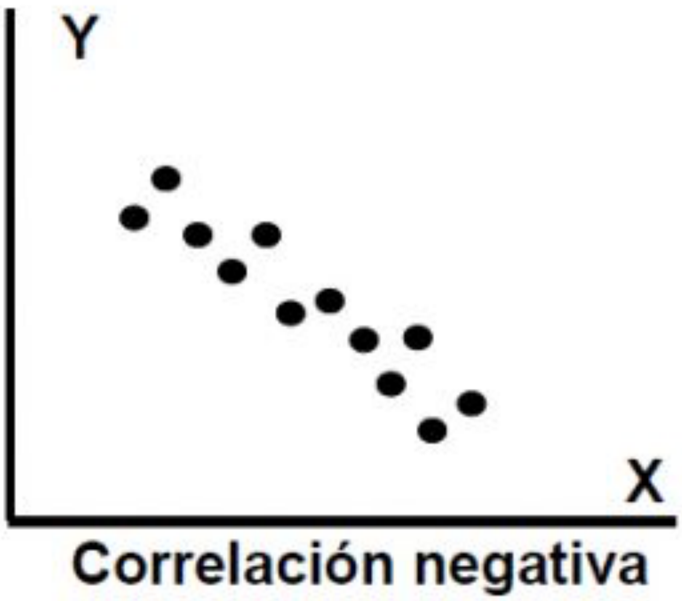
$7! / 3! \times 4! = 5040 / 144 = 35$

PRUEBA DE CORRELACIÓN NO PARAMÉTRICA

Correlación de Pearson

- **La correlación estudia el comportamiento recíproco de dos variables *X* e *Y***





PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Coeficiente de correlación de Pearson (*r*)

$$r = \frac{\sum XY - \frac{\sum X \sum Y}{N}}{\sqrt{\left(\sum X^2 - \frac{\left(\sum X\right)^2}{N}\right)}\sqrt{\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{N}\right)}}$$

$-1 <= r <= 1$

Hipótesis

H_0 : r = 0 ausencia de correlación

H_1 : r ≠ 0 existencia de correlación

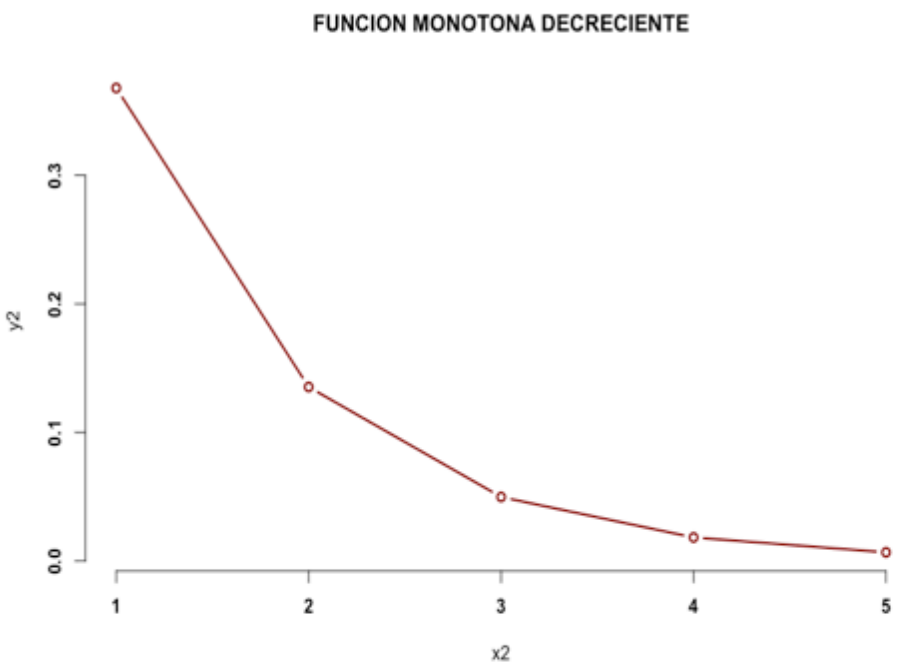
Supuestos:

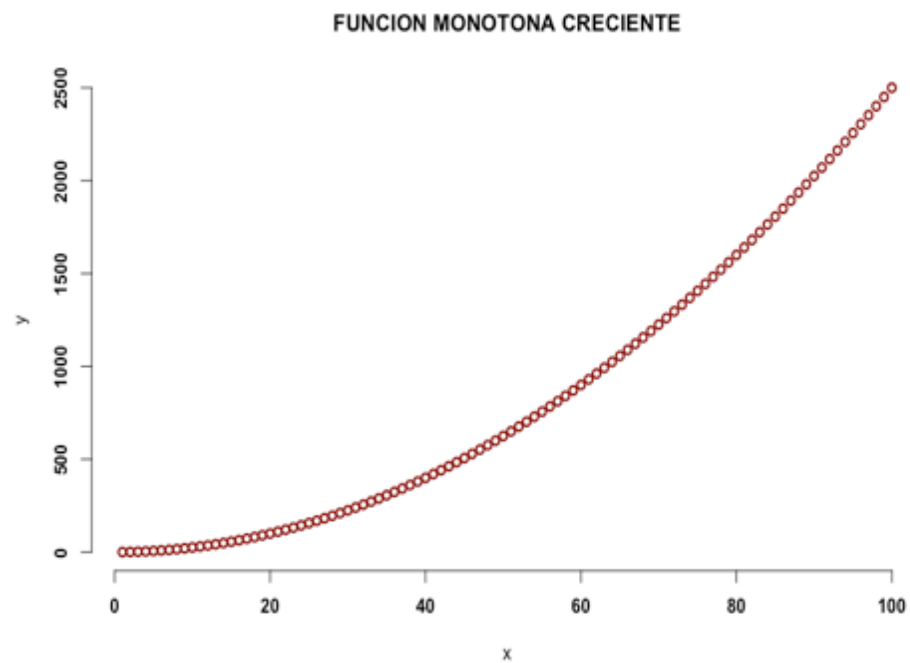
- Las variables *X* e *Y* son continuas.
- Existe relación lineal.
- La distribución conjunta de (X, Y) es una distribución Bivariable normal.

Coeficiente de correlación de Pearson no aplica para funciones monótonas.

Supuestos test correlación pearson que no se cumple:

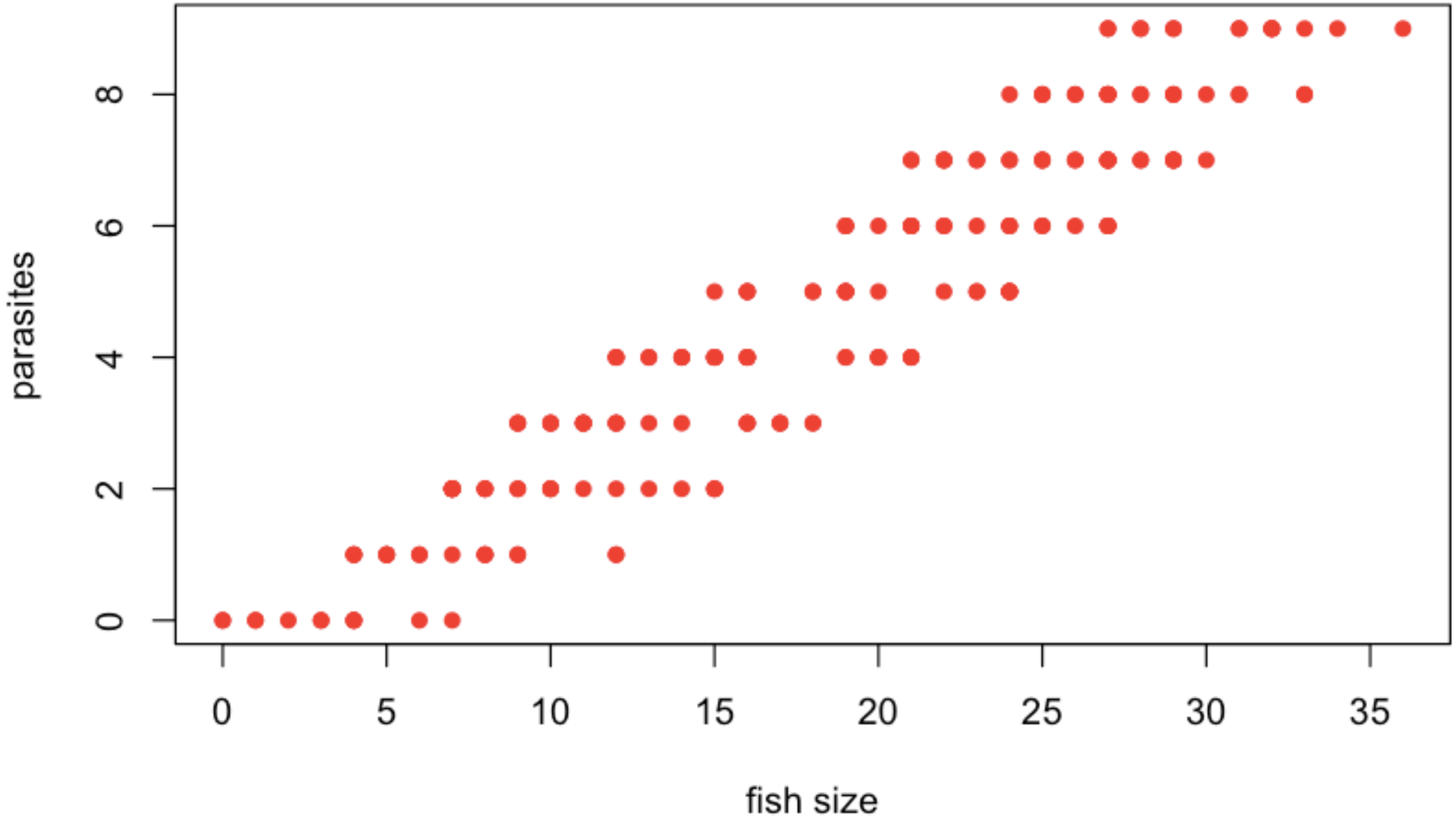
Existe una relación lineal.





PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Coeficiente de correlación de Pearson no aplica para variables discretas u ordinales.



PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Coeficiente de correlación de Spearman (ρ = rho)

Hipótesis

H_0 : X e Y son mutuamente independientes.

H_1 : X e Y no son mutuamente independientes.

Datos originales

Fish	Fish size (x)	Parasites (y)
1	942	13
2	101	14
3	313	18
4	800	10

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

d : diferencia entre ranking

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

2 RESULTADOS POSIBLES DE CORRELACIÓN

Rx	Ry
4	1
1	4
2	2
3	3

ρ =	= - 1

Rx	Ry
4	4
1	1
2	2
3	3

ρ =	= 1

¿Cuántas permutaciones / correlaciones son posibles?

$P_4 = 4! = 4 \times 3 \times 2 \times 1 = 24$ permutaciones posibles

Coeficiente de correlación de Spearman (ρ = rho)

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

Datos originales

Fish	Fish size (x)	Parasites (y)
1	942	13
2	101	14
3	313	18
4	800	10

Rangos

Sujeto	Rx	Ry
1		
2		
3		
4		

Diferencia entre rangos

d	d^2
Σ = S	

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

Coeficiente de correlación de Spearman (ρ = rho)

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

Hipótesis

H_0 : X e Y son mutuamente independientes.
H_1 : X e Y no son mutuamente independientes.

Datos originales

Fish	Fish size	Parasites
1	942	13
2	101	14
3	313	18
4	800	10

Rangos

Fish	Fish size	Parasites
1	4	2
2	1	3
3	2	4
4	3	1

Diferencia entre rangos

d	d^2
2	4
-2	4
-2	4
2	4
Σ = S	**16**

ρ =	$1 - ((6*16) / 4 (4^2 - 1)$	= - 0,6

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

DISTRIBUCIÓN MUESTRAL DE POSIBLE VALORES DE CORRELACIÓN

-1.00	- 0.80	- 0.80	- 0.80	- 0.60	- 0.40	- 0.40	- 0.40	- 0.40	- 0.20	- 0.20	0.00
0.00	0.20	0.20	0.40	0.40	0.40	0.40	0.60	0.80	0.80	0.80	1.00

¿Cómo someto a prueba la hipótesis?

Hipótesis	Verdadera cuando
H_0 : X e Y son mutuamente independientes	Rho cerca de cero
H_1 : X e Y no son mutuamente independientes.	Rho lejos de cero, ya sea positivo o negativo.

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

PRUEBA DE HIPÓTESIS CORRELACION

-1.00	- 0.80	- 0.80	- 0.80	- 0.60	- 0.40	- 0.40	- 0.40	- 0.40	- 0.20	- 0.20	0.00
0.00	0.20	0.20	0.40	0.40	0.40	0.40	0.60	0.80	0.80	0.80	1.00

p = 10 / 24

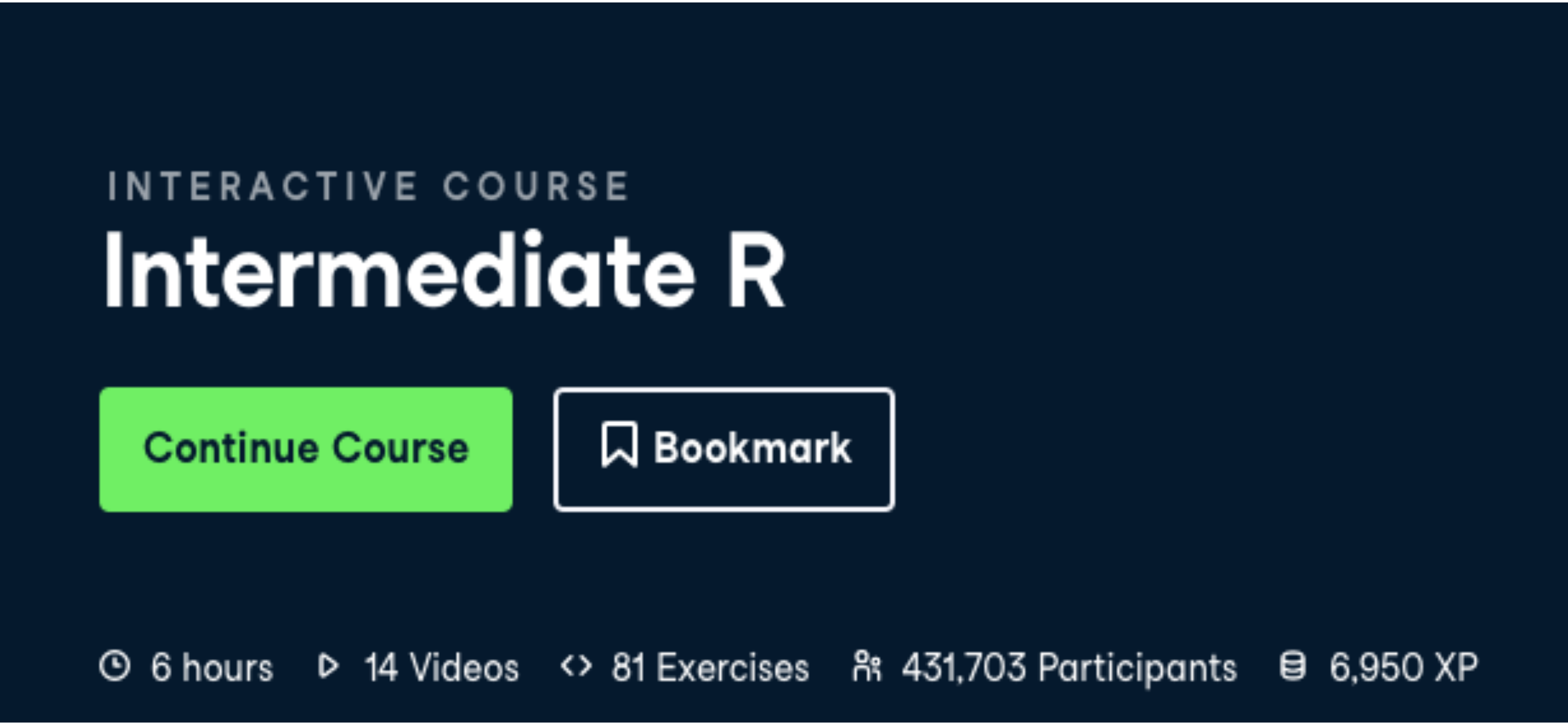
p = 0.4167

No se rechaza H_0 porque p = 0,416 es mayor a 0,05

Note que para el experimento de 4 peces nunca podré aceptar H_0 porque aun con el valor de Rho = 1, la probabilidad es de 0,08.

PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

AUTOAPRENDIZAJE CON DATACAMP



RESUMEN DE LO APRENDIDO

1. **Revisión de conceptos de permutación.**
2. Recordatorio correlación Pearson.
3. **Funciones monótonas.**
4. Aplicación de concepto de permutación a prueba de correlación de Spearman.