

Clase 10 Pruebas no paramétricas paramétricas

Diplomado en Análisis de datos con R para la acuicultura.

Dr. José A. Gallardo | jose.gallardo@pucv.cl | Pontificia
Universidad Católica de Valparaíso

19 July 2021

PLAN DE LA CLASE

1.- Introducción

- ▶ ¿Qué son las pruebas no paramétricas?
- ▶ Test de Correlación
- ▶ Prueba de asociación Chi cuadrado
- ▶ Pruebas de contraste no paramétrico
- ▶ Elaboración e interpretación pruebas no paramétricas con R

2.- Práctica con R y Rstudio cloud

- ▶ Realizar pruebas no paramétricas.
- ▶ Realizar gráficas avanzadas con ggplot2.
- ▶ Elaborar un reporte dinámico en formato pdf.

MÉTODOS NO PARAMÉTRICOS

Conjunto diverso de pruebas estadísticas.

El concepto de “no paramétrico” a veces es confuso, pues los métodos no paramétricos si estiman y someten a prueban hipótesis usando parámetros, pero no los de distribución normal.

Se aplican usualmente para variables cuantitativas que no cumplen con el supuesto de normalidad y para variables cualitativas.

SUPUESTOS: MÉTODOS NO PARAMÉTRICOS

El principal supuesto de los métodos no paramétricos es que las variables aleatorias son independientes y con idéntica distribución.

Usualmente no tienen supuestos acerca de la distribución de la variable, alternativamente se conocen como métodos de distribución libre.

El concepto matemático de permutación está subyacente a muchos métodos no paramétricos y se utiliza para someter a prueba una hipótesis.

PRUEBA DE CORRELACIÓN NO PARAMÉTRICA

¿Cuándo aplicar prueba de correlación no paramétrica?

Cuando no se cumple uno o varios supuestos de la correlación paramétrica:

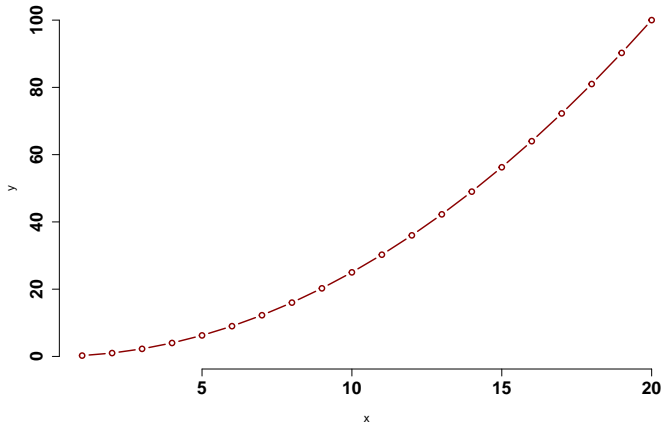
- Las variables X e Y no son continuas.
- No existe relación lineal.
- La distribución conjunta de (X, Y) no es una distribución Bivariable normal.

EJEMPLO FUNCIÓN MONÓTONA

¿Cuál es el supuesto que no se cumple?

No existe una relación lineal

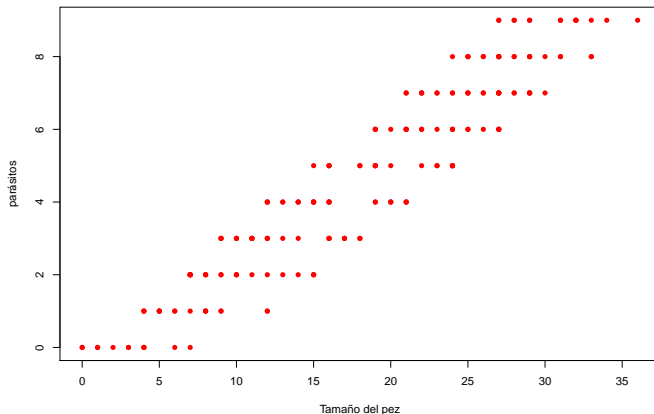
Relación no lineal



EJEMPLO VARIABLES DISCRETAS U ORDINALES

¿Cuál es el supuesto que no se cumple?

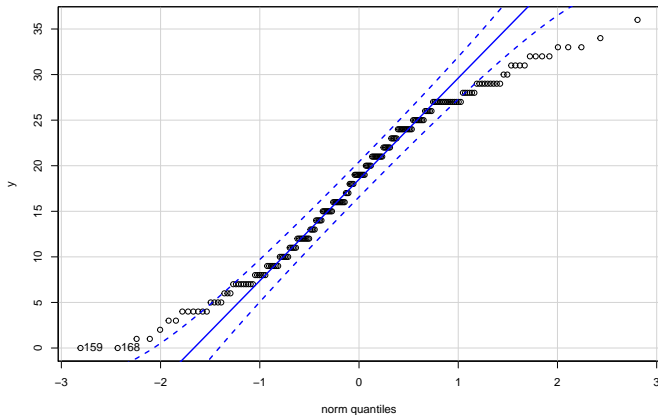
Parásitos es variable discreta.



EJEMPLO VARIABLES DISCRETAS U ORDINALES 2

¿Cuál es el supuesto que no se cumple?

Parásitos no tiene distribución normal



[1] 159 168

COEFICIENTE DE CORRELACIÓN DE SPEARMAN

¿Cómo se calcula?

Coeficiente de correlación de Spearman ($\rho = \text{rho}$)

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

Fish size (X)	Parásitos (Y)	Ranking X	Ranking Y	d	d ²
942	13	4	2	2	4
101	14	1	3	-2	4
313	18	2	4	-2	4
800	10	3	1	2	4

$$\sum d^2$$

$$= 16 ; \rho = -0,6$$

¿CUÁNTAS CORRELACIONES SON POSIBLES?

Permutación: combinación ordenada de elementos.

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

Ejemplo

Ranking X	Ranking Y	Ranking X	Ranking Y
4	1	4	4
1	4	1	1
2	3	2	2
3	2	3	3
$\rho = -1$			$\rho = 1$

DISTRIBUCIÓN MUESTRAL DE CORRELACIÓN

```
## [1] -1.0 -0.8 -0.8 -0.8 -0.6 -0.4 -0.4 -0.4 -0.2 -0.2
## [16]  0.4  0.4  0.6  0.6  0.6  0.8  0.8  0.8  1.0
```

PRUEBA DE HIPÓTESIS DE CORRELACIÓN

Hipótesis	Verdadera cuando
H_0 : X e Y mutuamente independientes	$\rho = 0$
H_1 : X e Y no son mutuamente independientes	$\rho \neq 0$

$$p = 10 / 24 \quad p = 0.4167$$

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

PRUEBA DE CORRELACIÓN CON R

```
# Crea objetos X e Y  
X <- c(942,101,313,800)  
Y <- c(13,14,18,10)
```

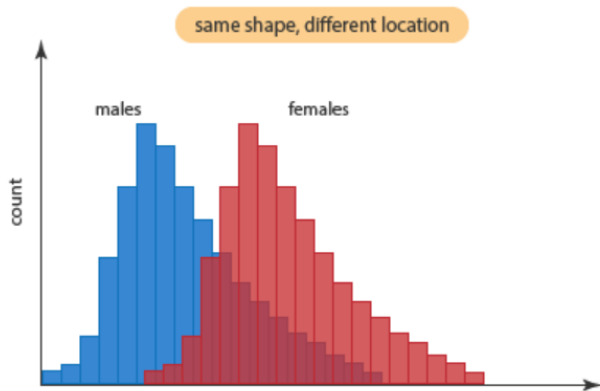
```
# Realiza test de correlación  
cor.test(X,Y, method = "spearman", alternative = "two.sided")
```

```
##  
## Spearman's rank correlation rho  
##  
## data: X and Y  
## S = 16, p-value = 0.4167  
## alternative hypothesis: true rho is not equal to 0  
## sample estimates:  
## rho  
## -0.6
```

PRUEBA DE MANN-WHITNEY

¿**Cuándo usarla?** Para comparar dos muestras con idéntica distribución, con diferentes medianas y sin normalidad.

Usualmente para variables discretas.



PRUEBA DE MANN-WHITNEY

¿Cómo se calcula el estadístico W? Como la diferencia de los ranking entre tratamiento y control

Tratamiento (T)	Control (C)	Ranking T	Ranking C
9	0	4	1
12	4	5	2
13	6	6	3
		$\sum = 15$	$\sum = 6$

$$W = 15 - 6 = 9$$

PRÁCTICA ANÁLISIS DE DATOS

- ▶ Guía de trabajo práctico disponible en drive y Rstudio.cloud.

Clase_10

- ▶ El trabajo práctico se realiza en Rstudio.cloud.

Clase 10 - Estadística no paramétrica

RESUMEN DE LA CLASE

- ▶ **Teoría**

- ▶ Supuestos de los análisis paramétricos.
- ▶ Consecuencias de la violación de los supuestos.
- ▶ Interpretación de métodos gráficos, análisis de residuos y pruebas de hipótesis para evaluar supuestos.

- ▶ **Evaluación de supuestos**

- ▶ Independencia.
- ▶ Homocedasticidad.
- ▶ Normalidad.