

CLASE 07 - PRUEBAS DE CONTRASTE DE HIPÓTESIS

OCE 313 - Técnicas de análisis no paramétricos.

Dr. José Gallardo Matus

Pontificia Universidad Católica de Valparaíso

01 May 2022

PLAN DE LA CLASE

1.- Introducción

- ▶ Comparación de 2 muestras independientes.
- ▶ Comparación de muestras independientes pareadas.
- ▶ Comparación de múltiples muestras independientes.
- ▶ Interpretación test no paramétricos con R.

2). Práctica con R y Rstudio cloud.

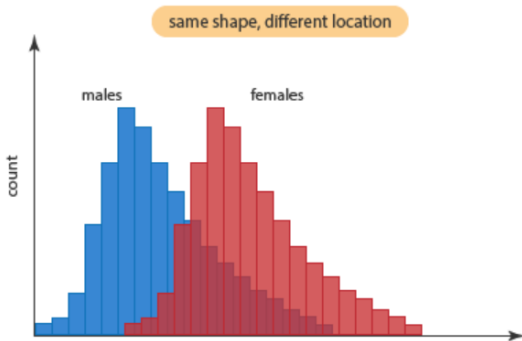
- ▶ Prácticas pruebas de contraste de hipótesis no paramétricas.
- ▶ Realizar gráficas avanzadas con ggplot2.

COMPARACIÓN DE DOS MUESTRAS INDEPENDIENTES

¿Para qué sirve?

Para comparar dos muestras con idéntica distribución, con diferentes medianas y sin normalidad.

Usualmente para variables discretas.



PRUEBA DE MANN-WHITNEY (W)

Estudio de caso: Comparación del reclutamiento de ostras en el intermareal entre dos sitios (S1 y S2) **Adaptado de Deck, 2007.**

Nº ostras S1	Nº ostras S2
9	0
12	4
13	6

CÁLCULO DE ESTADÍSTICO MANN-WHITNEY (W)

¿Cómo se calcula el estadístico W?

Como la diferencia de los ranking entre tratamiento y control

S1	S2	Ranking S1	Ranking S2
9	0	4	1
12	4	5	2
13	6	6	3
		$\Sigma = 15$	$\Sigma = 6$

$$W = 15 - 6 = 9$$

Máxima diferencia posible entre S1 y S2.

¿CUÁNTAS COMBINACIONES SON POSIBLES?

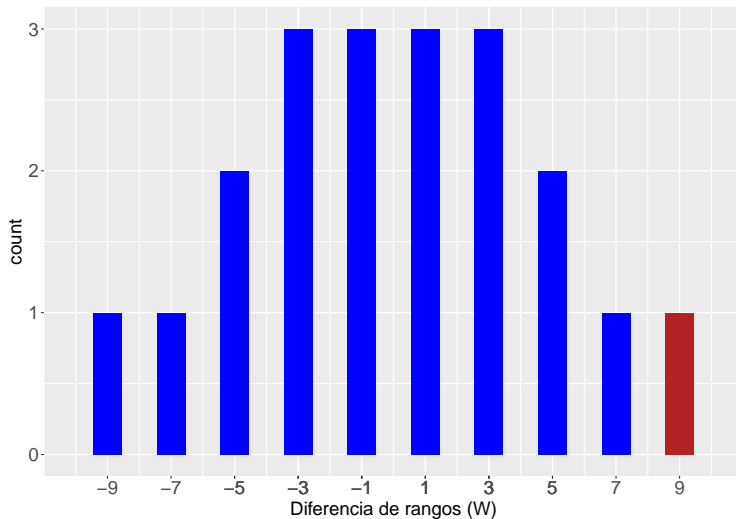
¿Cuántas combinaciones son posibles?

$$6! / 3! \times 3! = 720 / 36 = 20$$

S1	S2
1	4
2	5
3	6
6	15
$W = -9$	

S1	S2
2	1
5	3
6	4
13	8
$W = 5$	

DISTRIBUCIÓN MUESTRAL DE W



PRUEBA DE HIPÓTESIS DE MANN-WHITNEY

Hipótesis

$$H_0: S1 = S2$$

$$H_1: S1 > S2$$

Resultado obtenido $W=9$.

$$p = 1/20$$

$$p = 0.05$$

No se rechaza H_0 porque $p = 0,05$

PRUEBA DE MANN-WHITNEY CON R

```
# Crea objetos tratamiento y control
```

```
t <- c(9, 12, 13)
```

```
c <- c(0, 4, 6)
```

```
# Realiza prueba de Mann-Whitney
```

```
wilcox.test(t, c, alternative = "g",  
            paired = FALSE)
```

```
##
```

```
## Wilcoxon rank sum exact test
```

```
##
```

```
## data: t and c
```

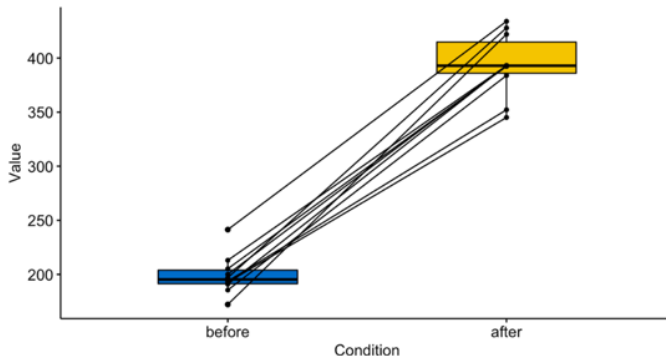
```
## W = 9, p-value = 0.05
```

```
## alternative hypothesis: true location shift is greater t
```

COMPARACIÓN DE MUESTRAS PAREADAS

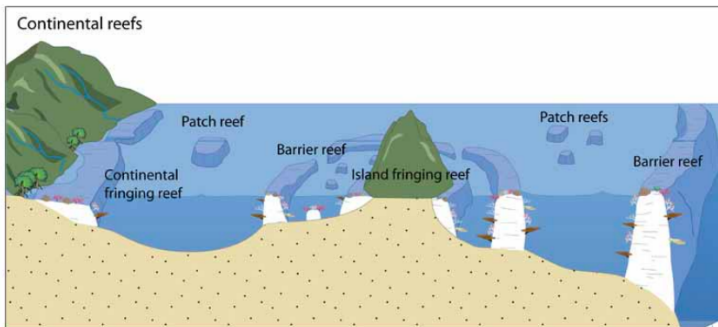
¿Para que sirve?

Para comparar dos muestras *pareadas* con idéntica distribución, con diferentes medianas y sin normalidad.



PRUEBA DE WILCOXON MUESTRAS PAREADAS

Estudio de caso: Comparación de productividad béntica (Coral Path Reef) entre dos tiempos. **Adaptado de Toro-Farmer et al. 2016**



DATOS SIMULADOS DE PRODUCTIVIDAD

¿Aumenta la productividad?

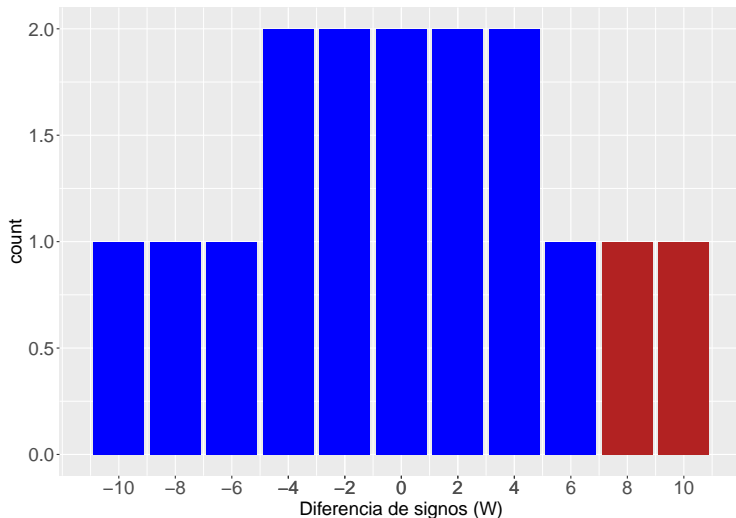
Arrecife de coral	Mes 1	Mes 2	d	Ranking con signo
1	45	49	4	2
2	41	50	9	4
3	47	52	5	3
4	52	50	2	-1

W = suma de los ranking = 8

V = suma de casos positivos (aumenta) = 9

DISTRIBUCIÓN MUESTRAL DE W

¿Cuántas combinaciones de signos (+ o -) son posibles? $2^4 = 16$



PRUEBA DE HIPÓTESIS DE WILCOXON

Hipótesis

H₀: $d = 0$

H₁: $d > 0$

$$p = 2/16$$

$$p = 0,125$$

No se rechaza **H₀** porque $p = 0,125$ es mayor a $0,05$

PRUEBA DE WILCOXON PAREADAS CON R

```
# Crea objetos pre y post  
pre <- c(45, 41, 47, 52)  
post <- c(49, 50, 52, 50)  
# Realiza prueba de Wilcoxon  
wilcox.test(post - pre, alternative = "greater")
```

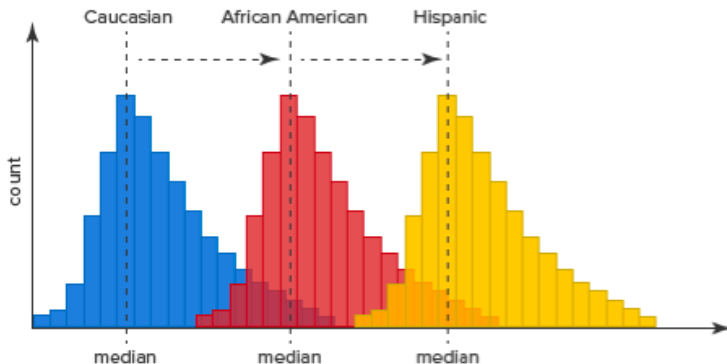
```
##  
## Wilcoxon signed rank exact test  
##  
## data: post - pre  
## V = 9, p-value = 0.125  
## alternative hypothesis: true location is greater than 0
```

```
# no es necesario indicar muestras pareadas  
# pues estamos haciendo la resta en la función.
```

COMPARACIÓN DE MÚLTIPLES MUESTRAS INDEPENDIENTES

¿Para que sirve?

Para comparar múltiples muestras con idéntica distribución, con diferentes medianas y sin normalidad.



PRUEBA DE KRUSKAL - WALLIS

Estudio de caso: Comparación de abundancia de zooplancton entre estaciones **Adaptado de Torreblanca et al. 2016.**

Verano	Otoño	Invierno	Primavera
900	600	500	50
800	650	450	40
700	550	400	75
750	500	475	100

Hipótesis

H_0 : La distribución de los k grupos son iguales.

H_1 : Al menos 2 grupos son distintos.

PRUEBA DE KRUSKAL - WALLIS CON R

```
V <- c(900,800,700,750) # Verano
O <- c(600,650,550,500) # Otoño
I <- c(500,450,400,475) # Invierno
P <- c(50,40,75,100) # Primavera
kt <- kruskal.test(list(V, O, I,P))
kt[["statistic"]]
```

```
## Kruskal-Wallis chi-squared
##                13.96723
```

```
kt[["parameter"]]
```

```
## df
##   3
```

```
kt[["p.value"]]
```

```
## [1] 0.002950098
```

RESUMEN DE LA CLASE

- ▶ Comparación de 2 muestras independientes.
- ▶ Comparación de muestras independientes pareadas.
- ▶ Comparación de múltiples muestras independientes.
- ▶ Interpretación test no paramétricos con R.