

Colegio de
Postgraduados

Estadística no paramétrica

Introducción

Humberto Vaquera Huerta

2024-05-28

Contenido

Prueba de Kruskal-Wallis

Pruebas para comparaciones múltiples de tratamientos

- Prueba de Dunn
- Prueba de Conover
- Prueba de Nemenyi

Prueba de Friedman

Prueba de Quade

Prueba de Jonckheere-Terpstrata

Prueba Q Cochran

Prueba de Kruskal-Wallis

La Prueba de **Kruskal-Wallis (Prueba H)** es el equivalente no paramétrico del **ANOVA unidireccional**. La prueba de Kruskal-Wallis nos permite probar diferencias entre **dos o mas muestras**. Se basa en el calculo de rangos. Tiene supuestos menos restrictivos que los de ANOVA.

Las suposiciones iguales que en la prueba U de Mann-Whitney:

- La variable respuesta se mide en escalas *ordinales, de intervalo o de razón*.
- Las observaciones de ambos grupos son *independientes* entre sí.
- Las *distribuciones tienen formas similares*: las distribuciones en cada grupo deben tener una forma similar.

Prueba de Kruskal-Wallis

La Prueba de **Kruskal-Wallis (Prueba H)** se utiliza para probar la hipótesis nula de que el muestras independientes provienen de poblaciones con medianas iguales.

- $H_0 : Med_1 = Med_2 = \dots = Med_k$: Las muestras provienen de poblaciones con medianas iguales.
- H_1 : Las muestras provienen de poblaciones con medianas que no todas son iguales

Para realizar una prueba de Kruskal-Wallis, usamos los rangos de los datos para calcular el estadístico de prueba, H, dado por

Prueba de Kruskal-Wallis

Estadístico H

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

Donde N es el tamaño total de la muestra, k es el número de grupos que estamos comparando, R_i es la suma de los rangos del grupo i y n_i es el tamaño de la muestra del grupo.

Se H con un valor crítico determinado por la distribución de chi-cuadrada (con tamaño de muestra de cada grupo ≥ 5).

Se rechaza H_0 si $H \geq \chi^2_{(k-1),\alpha}$.

Prueba de Kruskal-Wallis

Calculo del estadístico H :

- Paso 1: Indique su hipótesis – Hipótesis nula: las medianas (media de los rangos) son iguales en todas las muestras – Hipótesis alternativa: al menos una mediana es diferente
- Paso 2: Prepare y clasifique sus datos – Organice los datos de todos los grupos en una lista en orden ascendente – Asigne un rango a cada una de las entradas de datos
- Paso 3: Sume los rangos de cada grupo
- Paso 4: Calcular el estadístico de prueba, H
- Paso 5: Compárelo con el límite crítico, determinado por el valor crítico de chi-cuadrado
- Paso 6: Interpreta tus resultados

Ejemplo; Experimento de Maiz

Datos de un diseño completamente aleatorio donde cuatro métodos de cultivo de maiz diferentes, la variable respuesta fue el **rendimiento** por acre en parcelas donde se probaron los **cuatro métodos**.

##	trat	rend
## 1	1	83
## 2	1	91
## 3	1	94
## 4	1	89
## 5	1	89
## 6	1	96
## 7	1	91
## 8	1	92
## 9	1	90
## 10	2	91
## 11	2	90
## 12	2	81
## 13	2	83
## 14	2	84
## 15	2	83
## 16	2	88
## 17	2	91
## 18	2	89
## 19	2	84
## 20	3	101
## 21	3	100
## 22	3	91
## 23	3	93
## 24	3	96
## 25	3	95
## 26	3	94
## 27	4	78
## 28	4	82
## 29	4	81
## 30	4	77
## 31	4	79
## 32	4	81
## 33	4	80
## 34	4	81

Ejemplo; Experimento de Maiz

Queremos determinar cómo funcionan los 4 métodos comparándolos entre sí. Básicamente, buscamos determinar si los datos de rendimientos para en cada método de cultivo se originan en la misma distribución. Tenemos tamaños de muestra relativamente pequeños, por lo que no podemos determinar bien si los datos se distribuyen normalmente, por lo que usamos la prueba de Kruskal-Wallis.

Paso 1:

- **Hipótesis Nula** H_0 :los métodos de cultivo hacen que se produzca el mismo rendimiento (los 4 grupos se originan de la misma distribución y tienen la misma mediana)
- **Hipótesis alternativa** H_1 =Al menos uno de los métodos provoca un rendimiento diferente (al menos un grupo se origina en una distribución diferente y tiene una mediana diferente)

Ejemplo; Experimento de Maiz

Paso 2:

Aquí organizamos nuestros datos en orden ascendente y luego le damos a cada uno un rango.

```
##      trat rend rango
## 30      4   77   1.0
## 27      4   78   2.0
## 31      4   79   3.0
## 33      4   80   4.0
## 12      2   81   6.5
## 29      4   81   6.5
## 32      4   81   6.5
## 34      4   81   6.5
## 28      4   82   9.0
## 1       1   83  11.0
## 13      2   83  11.0
## 15      2   83  11.0
## 14      2   84  13.5
## 19      2   84  13.5
## 16      2   88  15.0
## 4       1   89  17.0
## 5       1   89  17.0
## 18      2   89  17.0
## 9       1   90  19.5
## 11      2   90  19.5
## 2       1   91  23.0
## 7       1   91  23.0
## 10      2   91  23.0
## 17      2   91  23.0
## 22      3   91  23.0
## 8       1   92  26.0
## 23      3   93  27.0
## 3       1   94  28.5
## 26      3   94  28.5
## 25      3   95  30.0
## 6       1   96  31.5
## 24      3   96  31.5
## 21      3  100  33.0
## 20      3  101  34.0
```

Ejemplo; Experimento de Maiz

Paso 3:

Ahora volvemos a colocar nuestros datos en sus grupos originales y sumamos los rangos para cada grupo.

##	trat	rend	rango
## 1	1	83	11.0
## 2	1	91	23.0
## 3	1	94	28.5
## 4	1	89	17.0
## 5	1	89	17.0
## 6	1	96	31.5
## 7	1	91	23.0
## 8	1	92	26.0
## 9	1	90	19.5
## 10	2	91	23.0
## 11	2	90	19.5
## 12	2	81	6.5
## 13	2	83	11.0
## 14	2	84	13.5
## 15	2	83	11.0
## 16	2	88	15.0
## 17	2	91	23.0
## 18	2	89	17.0
## 19	2	84	13.5
## 20	3	101	34.0
## 21	3	100	33.0
## 22	3	91	23.0
## 23	3	93	27.0
## 24	3	96	31.5
## 25	3	95	30.0
## 26	3	94	28.5
## 27	4	78	2.0
## 28	4	82	9.0
## 29	4	81	6.5
## 30	4	77	1.0
## 31	4	79	3.0
## 32	4	81	6.5
## 33	4	80	4.0
## 34	4	81	6.5

Prueba de Kruskal-Wallis

Sumas de Rangos por tratamiento \scriptsize

trat	SumaR	ni
1	196.5	9
2	153.0	10
3	207.0	7
4	38.5	8

Paso 4:

Calculo de $H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$ con $N = 34$, $k = 4$,

$R_i = 196.5, 153, 207, 38.5$, $n_i = 9, 10, 7, 8$

substituyendo

$$H = \frac{12}{34(34+1)} \left[\frac{196.5^2}{9} + \frac{153^2}{10} + \frac{207^2}{7} + \frac{38.5^2}{8} \right] - 3(34+1) = 25.46$$

Prueba de Kruskal-Wallis

Paso 5:

Comparar H con el valor crítico de chi-cuadrado correspondiente.

Encontramos los grados de libertad restando 1 de k :

$$gl = k - 1 = 4 - 1 = 3$$

Usando este valor y con $\alpha = 0.05$ encontramos

$$\chi^2_{(3)} = 7.814728$$

La comparación entre $H = 25.46$ y $\chi^2_{(3)} = 7.814728$ da $H > \chi^2(3)$.

Paso 6:

Dado que H es mayor que el límite crítico $\chi^2(3)$, rechazamos la hipótesis nula; las medianas no son las mismas en los cuatro grupos, al menos uno de ellos tiene una mediana diferente a los demás.

Esto significa que con un 95% de confianza los datos nos indican que los métodos de cultivo no producen el mismo rendimiento, al menos uno hace que las parcelas produzcan un rendimiento de maíz diferente que las demás.

Es importante tener en cuenta que Kruskal-Wallis solo puede decirnos que al menos uno de los grupos se origina en una distribución diferente. No puede decirnos cuál del grupo(s) es(son).

Kruskal-Wallis en R

```
library(agricolae)
data(corn)
corn$method=as.factor(corn$method)
library(tidyverse)
maiz2=corn %>%
  rename(rend = observation) %>%
  rename(trat = method) %>%
  rename(rango = rx)
kruskal.test(rend ~trat, data = maiz2)
```

```
##
##      Kruskal-Wallis rank sum test
##
## data:  rend by trat
## Kruskal-Wallis chi-squared = 25.629, df = 3, p-value = 1.141e-05
```

Kruskal-Wallis en SAS

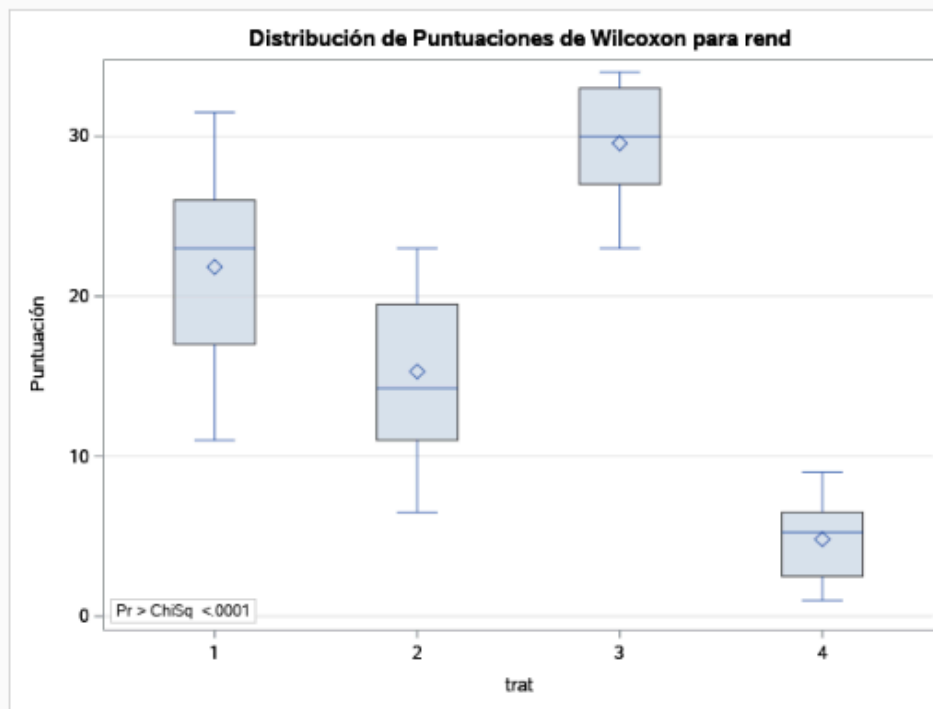
```
data fert;
input trat rend @@;
cards;
1 83 2 89
1 91 2 84
1 94 3 101
1 89 3 100
1 89 3 91
1 96 3 93
1 91 3 96
1 92 3 95
1 90 3 94
2 91 4 78
2 90 4 82
2 81 4 81
2 83 4 77
2 84 4 79
2 83 4 81
2 88 4 80
2 91 4 81
;
Kruskal-Wallis;
proc npar1way data=fert wilcoxon dscf;
class trat;
var rend;
run;
```

Kruskal-Wallis en SAS

Procedimiento NPAR1WAY

Puntuaciones de Wilcoxon (Sumas de rango) para variable rend Clasificado por variable trat					
trat	N	Suma de puntuaciones	Esperado debajo de H0	Desv. est. debajo de H0	Puntuación media
1	9	196.50	157.50	25.535050	21.833333
2	10	153.00	175.00	26.372486	15.300000
3	7	207.00	122.50	23.403260	29.571429
4	8	38.50	140.00	24.551448	4.812500
Se utilizaron puntuaciones media para valores repetidos.					

Test de Kruskal-Wallis		
Chi-cuadrado	DF	Pr > ChiSq
25.6286	3	<.0001



Prueba para comparaciones múltiples de tratamientos

Prueba para comparaciones múltiples de tratamientos

Si la prueba de Kruskal-Wallis es significativa, se puede realizar un análisis *post-hoc* para determinar qué grupos difieren entre sí.

Probablemente la prueba post-hoc más popular para la prueba de Kruskal-Wallis es la prueba de Dunn. También usan la prueba de Conover y la prueba de Nemenyi.

Debido a que la prueba post-hoc producirá múltiples p – *valores*, se pueden hacer ajustes a los valores p para evitar inflar la posibilidad de cometer un error de *tipo I*. Hay una variedad de métodos para controlar la tasa de error familiar o para controlar la tasa de descubrimiento falso FDR.

Cuando hay muchos valores de p para evaluar, es útil condensar una tabla de valores de p en un formato de visualización de letras compacto. En la salida, los grupos están separados por letras.

Prueba de Dunn

Dunn (1964) ha propuesto una prueba para comparaciones múltiples de sumas de rango basadas en la estadística z de la distribución normal estándar. La hipótesis nula (H_0), la probabilidad de observar un valor seleccionado al azar del primer grupo que es mayor que un valor seleccionado al azar del segundo grupo es igual a la mitad, se rechaza, si se excede una diferencia absoluta crítica de sumas de rango medio:

$$|\bar{R}_i - \bar{R}_j| > Z_{(1-\alpha^*/2)} \sqrt{\frac{n(n+1)}{12} \left(\frac{1}{n_i} - \frac{1}{n_j} \right)} \quad (1)$$

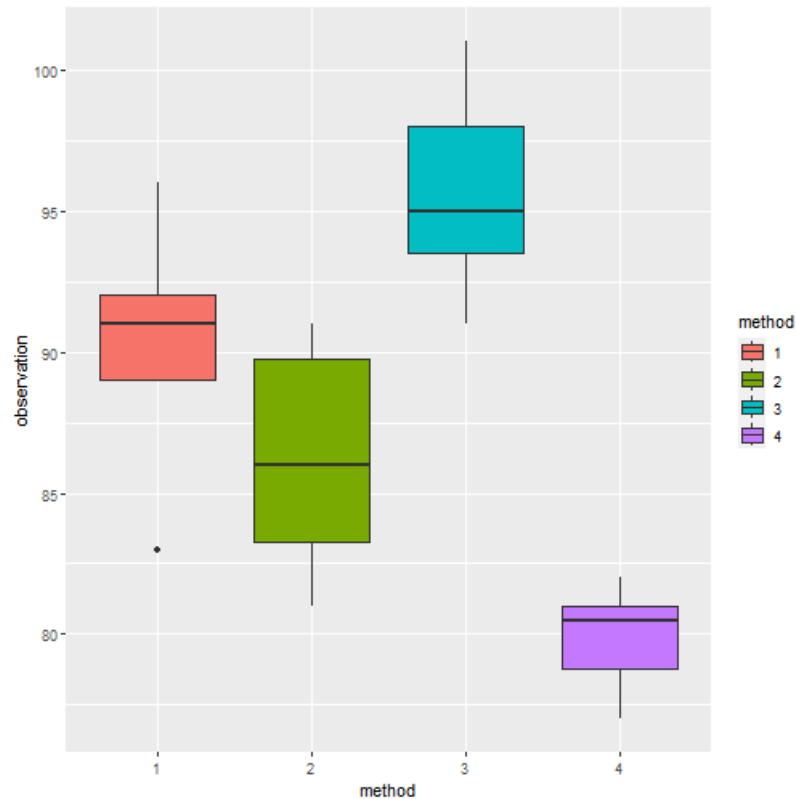
$Z_{(1-\alpha^*/2)}$ el valor de la distribución normal estándar para un nivel $\alpha^* = \frac{\alpha}{k(k-1)}$ (ajuste de Holm) ajustado dependiendo del número de pruebas realizadas o por ajuste de Bonferroni $\alpha^* = \frac{\alpha}{k(k-1)/2}$.

Ejemplo Prueba de Dunn para Maiz

```
library(PMCMRplus)
library(agricolae)
data(corn)
corn$method=as.factor(corn$method)
library(tidyverse)
rend1=corn$observation
trat1=as.factor(corn$method)
kwAllPairsDunnTest(x=rend1, g=trat1,dist="Tukey",
p.adjust.method = "holm" )
library(tidyverse)
ggplot(corn, aes(x =method, y = observation, fill = method)) +
geom_boxplot()
```

Ejemplo Prueba de Dunn para Maiz

```
##      1      2      3
## 2 0.2438 -      -
## 3 0.2438 0.0141 -
## 4 0.0021 0.0778 8.6e-06
```



Prueba de Conover-Iman para comparaciones múltiples de tratamientos

El procedimiento de Conover-Iman es simplemente el método de diferencia mínima significativa de Fisher (LSD) realizado en rangos.

$$|\bar{R}_i - \bar{R}_j| < t_{1-\alpha/2, n-k} \sqrt{s^2 \left[\frac{N-1-H}{N-k} \right] \times \left[\frac{1}{n_i} + \frac{1}{n_j} \right]}$$

Con:

$$s^2 = \frac{1}{N-1} \times \left[\sum R_i^2 - N \times \frac{(N+1)^2}{4} \right]$$

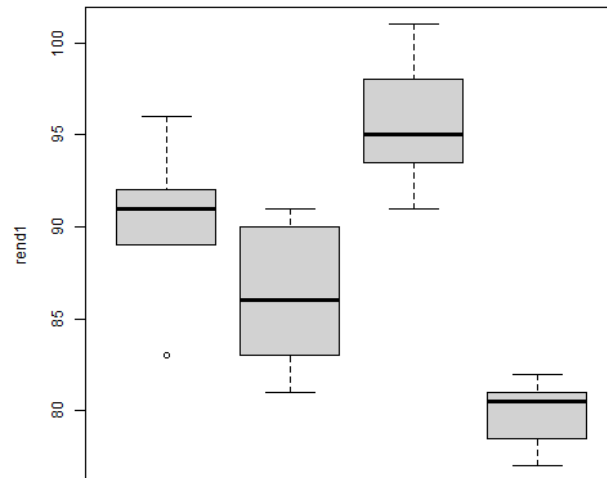
donde t es un cuantil de la distribución t de Student en $N - k$ grados de libertad.

Una alternativa a Kruskal-Wallis es realizar un ANOVA unidireccional en los rangos de las observaciones y luego aplicar la prueba LSD.



Prueba de Conover en R

```
##
## Conover's test of multiple comparisons : holm
##
##      mean.rank.diff      pval
## 2-1      -6.533333 0.00794 **
## 3-1       7.738095 0.00794 **
## 4-1     -17.020833 3.2e-07 ***
## 3-2      14.271429 7.7e-06 ***
## 4-2     -10.487500 0.00029 ***
## 4-3     -24.758929 5.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```





Prueba de Conover en SAS

```
data fert;
input trat rend @@;
cards;
1 83 2 89
1 91 2 84
1 94 3 101
1 89 3 100
1 89 3 91
1 96 3 93
1 91 3 96
1 92 3 95
1 90 3 94
2 91 4 78
2 90 4 82
2 81 4 81
2 83 4 77
2 84 4 79
2 83 4 81
2 88 4 80
2 91 4 81
;
prueba de Conover ;
proc rank;
  var rend;
  ranks rango_rend;
run;
proc print; run;
proc anova;
  class trat;
  model rango_rend= trat;
  means trat/lsd alpha=0.05 lines ;
run;
```

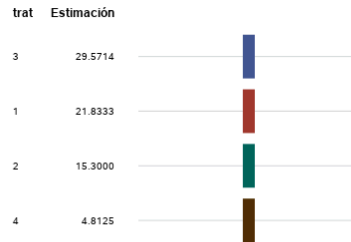

Prueba de Conover en SAS

Alpha	0.05
Grados de error de libertad	30
Error de cuadrado medio	24.20943
Valor crítico de t	2.04227
Diferencia menos significativa	4.9175
Media armónica de tamaño de celdas	8.351284

Note: Cell sizes are not equal.

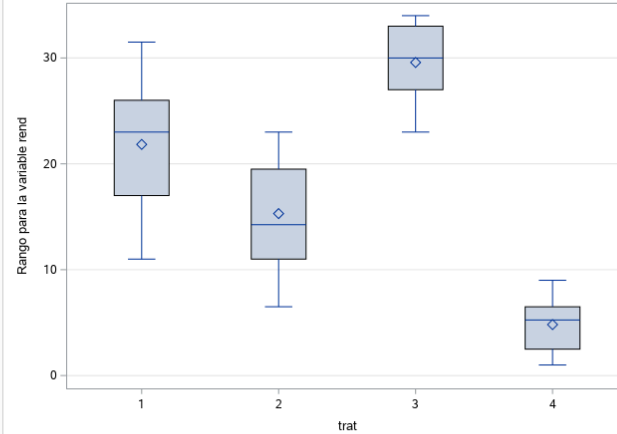
rango_rend t Grouping for Means of trat (Alfa = 0.05)

Means cubiertas por la misma barra no son significativamente diferentes.



Procedimiento ANOVA

Distribución de rango_rend



Prueba de Nemenyi

Si rechazamos H_0 y aceptamos H_a : "al menos una de las muestras k no pertenece a la misma población", uno puede estar interesado en identificar qué grupo o grupos son significativamente diferentes. El número de contrastes por pares o pruebas posteriores que deben realizarse es $m = k(k - 1)/2$ para detectar las diferencias entre cada grupo. Nemenyi propuso una prueba basada en sumas de rango. El enfoque se basa en el de Tukey y Kramer y utiliza sumas de rango medio. La hipótesis nula $H_0: \bar{R}_i = \bar{R}_j$ se rechaza, si se excede una diferencia absoluta crítica de sumas de rango medio.

$$|\bar{R}_i - \bar{R}_j| = \frac{q_{(\infty; k; \alpha)}}{\sqrt{2}} \sqrt{\frac{n(n+1)}{12} \left(\frac{1}{n_i} - \frac{1}{n_j} \right)} \quad (2)$$

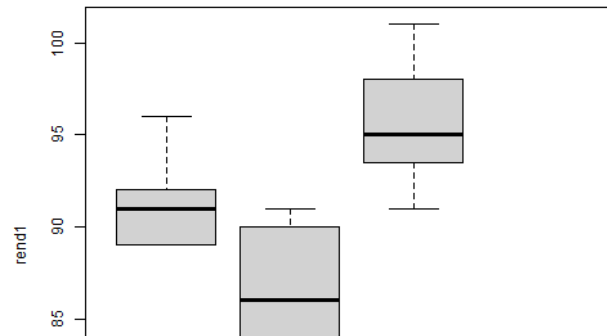
donde $q_{(\infty, k, \alpha)}$ es un valor de tablas de rangos estudentizados de Tukey.

Prueba de Nemenyi en R

```
library(agricolae)
data(corn)
library(PMCMRplus)
rend1=corn$observation
trat1=as.factor(corn$method)
kwAllPairsNemenyiTest(x=rend1, g=trat1,dist="Tukey")
```

```
##      1      2      3
## 2 0.4818 -      -
## 3 0.4123 0.0191 -
## 4 0.0025 0.1177 9.3e-06
```

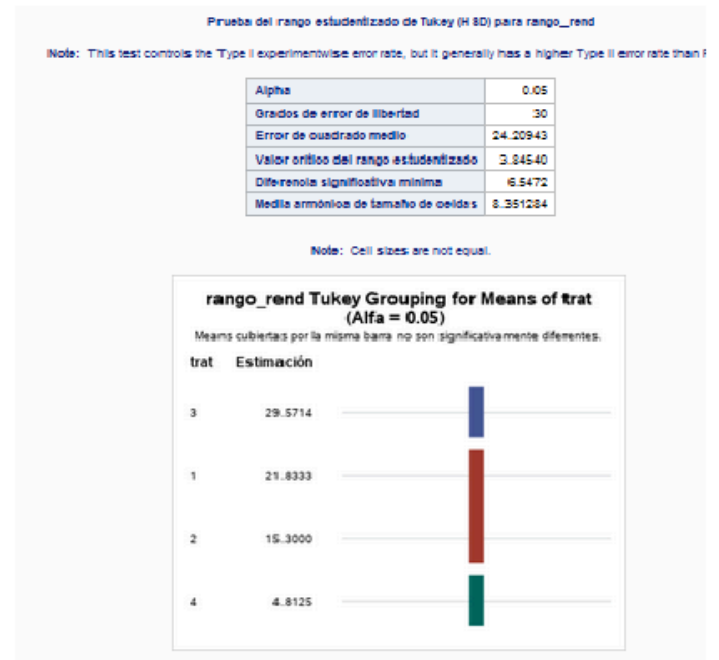
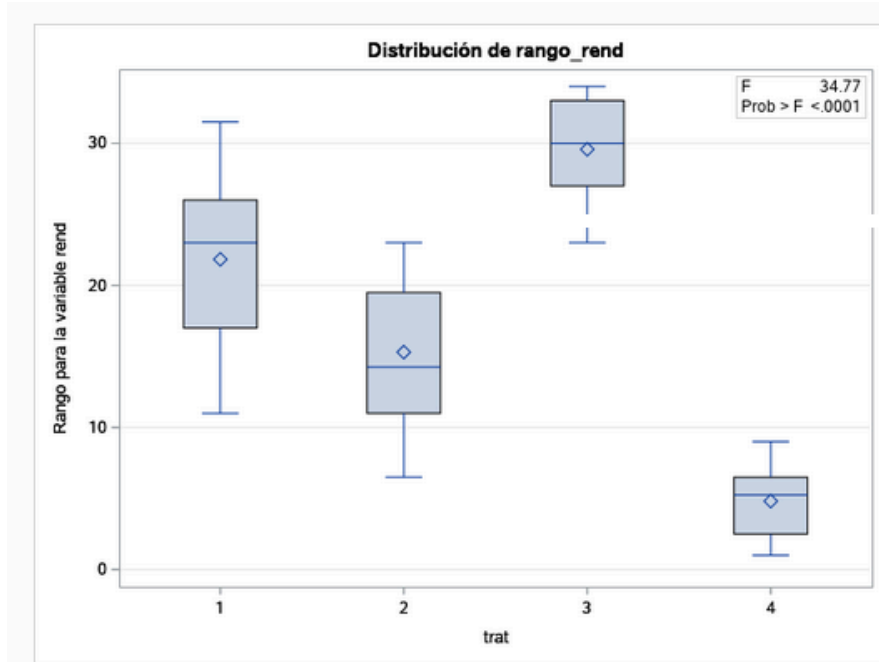
```
boxplot(rend1 ~ trat1)
```



Prueba de Nemenyi en SAS

```
data fert;
input trat rend @@;
cards;
1 83 2 89
1 91 2 84
1 94 3 101
1 89 3 100
1 89 3 91
1 96 3 93
1 91 3 96
1 92 3 95
1 90 3 94
2 91 4 78
2 90 4 82
2 81 4 81
2 83 4 77
2 84 4 79
2 83 4 81
2 88 4 80
2 91 4 81
;
prueba de Connover ;
proc rank;
  var rend;
  ranks rango_rend;
run;
proc print; run;
proc anova;
  class trat;
  model rango_rend= trat;
  means trat/tukey alpha=0.05 lines ;
run;
```

Prueba de Nemenyi en SAS



Ejemplo2; Contaminación del aire

Lecturas diarias de los siguientes valores de calidad del aire en Nueva York del 1 de mayo de 1973 (un martes) al 30 de septiembre de 1973.

- Ozono: Ozono medio en partes por billón de 1300 a 1500 horas en Roosevelt Island
- Solar.R: Radiación solar en Langleys en la banda de frecuencia 4000-7700 Angstroms de 08:00 a 12:00 horas en Central Park
- Viento: Velocidad promedio del viento en millas por hora a las 07:00 y 10:00 horas en el Aeropuerto LaGuardia
- Temp: Temperatura máxima diaria en grados Fahrenheit en el Aeropuerto La Guardia.

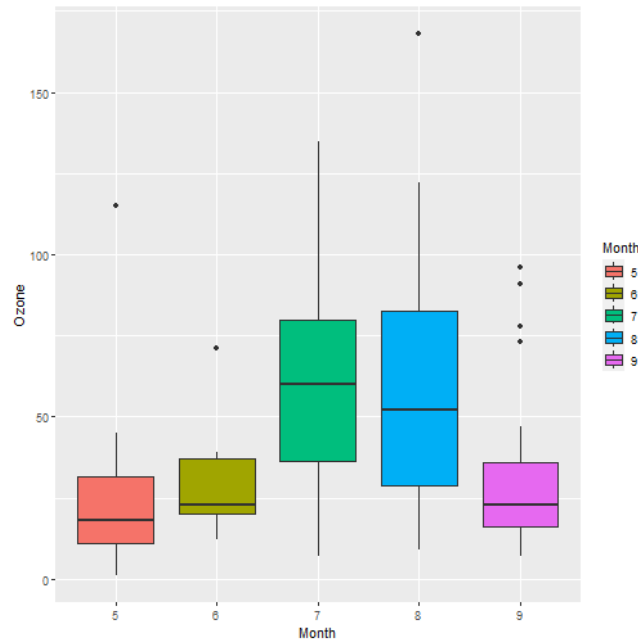
```
head(airquality,10)
```

##		Ozone	Solar.R	Wind	Temp	Month	Day
##	1	41	190	7.4	67	5	1
##	2	36	118	8.0	72	5	2
##	3	12	149	12.6	74	5	3

Ejemplo2; Contaminación del aire

Usando Mediciones diarias de la calidad del aire en Nueva York se desea investigar si los niveles de ozono se mantienen de manera similar o no entre los meses.

```
library(tidyverse)
airquality$Month=as.factor(airquality$Month)
ggplot(airquality, aes(x = Month, y = Ozone, fill = Month)) +
  geom_boxplot()
```



Ejemplo2; Contaminación del aire

Para datos de contaminación deseamos probar

- $H_0 : Med_5 = Med_6 = Med_7 = Med_8 = Med_9 :$
- H_1 :al menos una Med_i difiere del resto.

Usando Prueba de Kruskal Wallis en R

```
#head(airquality)
library(knitr)
kruskal.test(Ozone ~ Month, data = airquality)
```

```
##
##      Kruskal-Wallis rank sum test
##
## data:  Ozone by Month
## Kruskal-Wallis chi-squared = 29.267, df = 4, p-value = 6.901e-06
```

note que el $pvalue = 6.901e - 06$ es menor $\alpha = 0.05$ por lo que rechazo H_0

Con un 95% de confianza los datos nos indican que las concentraciones de ozono no son similares entre meses.

Ejemplo2; Contaminación del aire

Cual es el peor mes con altos niveles de ozono? Usando la prueba de Nemenyi

```
library(PMCMRplus)
Ozone=airquality[,1]
Month=as.factor(airquality[,5])
kwAllPairsNemenyiTest(x=Ozone, g=Month,
                      dist="Tukey")
```

```
##      5      6      7      8
## 6 0.88737 -      -      -
## 7 9.7e-05 0.16373 -      -
## 8 0.00035 0.24773 0.99853 -
## 9 0.67819 1.00000 0.01136 0.02867
```

```
library(tidyverse)
airquality$Month=as.factor(airquality$Month)
ggplot(airquality, aes(x = Month, y = Ozone, fill = Month)) +
  geom_boxplot()
```





ANOVA (Análisis de varianza)

El diseño de dos vías de clasificación

- Dos factores A (factor de tratamiento) y B (factor de bloqueo).
 - A tiene niveles de k .
 - B tiene n niveles.
 - el interés primario está en el factor A .
- El interés principal es sobre las ubicaciones relativas (medianas) de tres o más poblaciones (poblaciones (k)) dentro de cada uno de los grupos homogéneos.
- Datos: diseño de bloques completos al azar (RCBD), diseño de bloques al azar (RBD), sin observaciones en algunos bloques (diseño de bloques incompletos balanceados), al menos una observación en cada celda.

El diseño de dos vías de clasificación

<i>Bloques</i>	<i>Tratamientos</i>			
	1	2	\dots	k
1	X_{111}	X_{121}	\dots	X_{1k1}
	\vdots	\vdots		\vdots
	$X_{11c_{11}}$	$X_{12c_{12}}$	\dots	$X_{1kc_{1k}}$
2	X_{211}	X_{221}	\dots	X_{2k1}
	\vdots	\vdots		\vdots
	$X_{21c_{21}}$	$X_{22c_{22}}$	\dots	$X_{2kc_{2k}}$
n	X_{n11}	X_{n21}	\dots	X_{nk1}
	\vdots	\vdots		\vdots
	$X_{n1c_{n1}}$	$X_{n2c_{n2}}$	\dots	$X_{nkc_{nk}}$

El diseño de dos vías de clasificación

- $N = \sum_{i=1}^n \sum_{j=1}^k c_{ij}$ observaciones.
- c_{ij} - observaciones de la combinación del bloque i -ésimo con el tratamiento j -ésimo.
- $X_{ijc_{ij}}$ - c_{ij} -ésima respuesta de la combinación del bloque i -ésimo con el tratamiento j -ésimo.
- Supuestos:
 - Las variables aleatorias N son independientes entre sí.
 - Las variables aleatorias c_{ij} son una muestra aleatoria de una distribución continua F_{ij} .
 - $F_{ij}(u) = F(u - \beta_i - \tau_j)$

El diseño de dos vias de clasificación

- El modelo de diseño de dos vias de clasificación es;

$$X_{ijt} = \theta + \beta_i + \tau_j + e_{ijt},$$

con $i = 1, \dots, n, j = 1, \dots, k, t = 1, \dots, c_{ij}$

	<i>Tratamientos</i>			
<i>Bloques</i>	1	2	\dots	k
1	X_{11}	X_{12}	\dots	X_{1k}
2	X_{21}	X_{22}	\dots	X_{2k}
	\vdots	\vdots		\vdots
n	X_{n1}	X_{n2}	\dots	X_{nk}

- $H_0 : [\tau_1 = \tau_2 = \dots = \tau_k]$ versus $H_A : [\tau_1, \tau_2, \dots, \tau_k \text{ no todos son iguales}]$.

Diseño de bloques completos al azar (prueba de Friedman)

- Friedman (1937) estadístico S

$$S = \frac{12n}{k(k+1)} \sum_{j=1}^k \left(R_{.j} - \frac{k+1}{2} \right)^2,$$

donde $R_{.j} = \frac{R_j}{n}$ y $R_j = \sum_{i=1}^n r_{ij}$, suma de los rangos en el tratamiento j .

Prueba de Friedman

Prueba de Friedman

La prueba de Friedman se utiliza cuando un grupo de n individuos son medidos k veces (k repeticiones), por ejemplo: cuando se evalúa la aplicación de k tratamientos a n individuos o cuando n individuos son evaluados por k observadores. En agronomía en diseños en bloques al azar.

Hipótesis.

Hipótesis nula H_0 : Las k mediciones de los n individuos tiene igual distribución.

- $H_0 : F_1(x) = F_2(X) = \dots = F_k(X)$ vs
- $H_a : F_i(X) \neq F_j(X)$ para $i \neq j$

La estadística de prueba es:

$$\chi_c^2 = \frac{12 \sum_{i=1}^k R_i^2}{nk(k+1)} - 3n(k+1) \sim \chi_{k-1}^2$$

Si $\chi_c^2 > \chi_{k-1, \alpha}^2$ se rechaza la hipótesis nula.

Ejemplo BDO en Agua

Una compañía de agua buscó evidencia de que las medidas tomadas para limpiar un río fueron efectivas. Se comparó la demanda biológica de oxígeno (DBO) en 12 sitios del río antes de la limpieza y 1 mes y 1 año después de la limpieza. Los resultados se dan a continuación.

Site	BOD		
	before	after 1 month	after 1 year
1	17.4	13.6	13.2
2	15.7	10.1	9.8
3	12.9	10.3	9.7
4	9.8	9.2	9.0
5	13.4	11.1	10.7
6	18.7	20.4	19.6
7	13.9	10.4	10.2
8	11	11.4	11.5
9	5.4	4.9	5.2
10	10.4	8.9	9.2
11	16.4	11.2	11.0
12	5.6	4.8	4.6

Ejemplo BDO en Agua

Asignación de rangos dentro de Bloque

Site	BOD		
	before	after 1 month	after 1 year
1	3	2	1
2	3	2	1
3	3	1.5	1.5
4	3	2	1
5	3	2	1
6	1	3	2
7	3	2	1
8	1	2	3
9	3	1	2
10	3	1	2
11	3	2	1
12	3	2	1

Site	before	after 1 month	after 1 year
Sum of ranks	32	22.5	17.5
(Sum of ranks) ²	1024	506.25	306.25
No of Columns, k	3	$(= 1024 + 506.25 + 306.25)$ $(= 12/12 \times 3 \times 4)$ $(= 3 \times 12 \times 4)$	
No of Rows, n	12		
ΣR^2	1836.5		
$12/nk(k+1)$	0.083		
$3n(k + 1)$	144		

$$Fr = 9.042 \text{ con } (= 0.083 \times 1836.5 - 144)$$

Ejemplo BDO en Agua



Upper Critical Values for the Friedman Test (k treatments and b blocks)

- Notes 1. In the table below, the critical values give significance levels as close as possible to, but not exceeding the nominal α .
2. For values of k and b beyond the range of the table below, various approximations are available.

	$k = 3$		$k = 4$		$k = 5$		$k = 6$	
b	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01
2	-	-	6.000	-	7.600	8.000	9.143	9.714
3	6.000	-	7.400	9.000	8.533	10.13	9.857	11.76
4	6.500	8.000	7.800	9.600	8.800	11.20	10.29	12.71
5	6.400	8.400	7.800	9.960	8.960	11.68	10.49	13.23
6	7.000	9.000	7.600	10.20	9.067	11.87	10.57	13.62
7	7.143	8.857	7.800	10.54	9.143	12.11		
8	6.250	9.000	7.650	10.50	9.200	12.30		
9	6.222	9.556	7.667	10.73	9.244	12.44		
10	6.200	9.600	7.680	10.68				
11	6.545	9.455	7.691	10.75				
12	6.500	9.500	7.700	10.80				
13	6.615	9.385	7.800	10.85				
14	6.143	9.143	7.714	10.89				
15	6.400	8.933	7.720	10.92				
16	6.500	9.375	7.800	10.95				
17	6.118	9.294	7.800	11.05				
18	6.333	9.000	7.733	10.93				
19	6.421	9.579	7.863	11.02				
20	6.300	9.300	7.800	11.10				
21	6.095	9.238	7.800	11.06				
22	6.091	9.091	7.800	11.07				
23	6.348	9.391						
24	6.250	9.250						

De tablas de Friedman con 12 bloques y 3 tratamientos El valor critico es 6.5 Por lo tanto rechazo H_0

Usando tablas de chi cuadrada con 2 grados de libertad el valor critico es 5.99 Por lo tanto rechazo H_0

Prueba de Friedman: ejemplo2

Nueve estudiantes de agronomía fueron evaluados en tres materias distintas obteniendo las notas que se presentan a continuación.

\scriptsize

```
library(readr)
ej1 <- read_csv("Comparaciones-Multiples_files/ejemplo1.csv")
knitr::kable( ej1)
```

Alumno	Matematicas	Fisiologia Vegetal	Entomologia
1	98	95	77
2	95	71	79
3	76	80	91
4	95	81	84
5	83	77	80
6	99	70	93
7	82	80	87

Prueba de Friedman en R

Hipotesis:

- $H_0 : Me_{mat} = Me_{fis} = Me_{ent}$
- H_1 al menos hay una materia Me diferente

```
calif<-matrix(c(98, 95, 76, 95, 83, 99, 82, 75, 88,  
               95, 71, 80, 81, 77, 70, 80, 72, 81,  
               77, 79, 91, 84, 80, 93, 87, 81, 83),  
             nrow=9,ncol=3,  
             dimnames = list(1:9,c("Mat","Fis","Entom")))  
friedman.test(calif)
```

```
##  
##      Friedman rank sum test  
##  
## data:  calif  
## Friedman chi-squared = 8.6667, df = 2, p-value = 0.01312
```

Rechazamos H_0 con $\alpha = 0.05$ por lo hay diferencia entre calif entre las materias.

Resultados de prueba de Nemenyi despues de Friedman

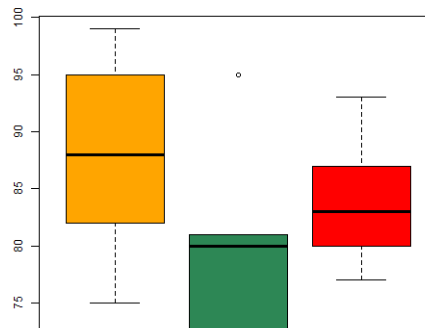
Cual es materia mas dificil?

se puede aplicar la prueba de Dunn o Nemenyi..

```
library(PMCMRplus)
frdAllPairsNemenyiTest(calif,p.adjust.method="bonferroni")
```

```
##          Mat    Fis
## Fis      0.013 -
## Entom 0.759 0.086
```

```
boxplot(calif, col=c("orange", "seagreen", "red"))
```



Prueba de Friedman en SAS

```
data cursos;
input alumno materia $ calif @@;
cards;
1 Mat 98 1 Fisiol 95 1 Ento 77
2 Mat 95 2 Fisiol 71 2 Ento 79
3 Mat 76 3 Fisiol 80 3 Ento 91
4 Mat 95 4 Fisiol 81 4 Ento 84
5 Mat 83 5 Fisiol 77 5 Ento 80
6 Mat 99 6 Fisiol 70 6 Ento 93
7 Mat 82 7 Fisiol 80 7 Ento 87
8 Mat 75 8 Fisiol 72 8 Ento 81
9 Mat 88 9 Fisiol 81 9 Ento 83
;
proc freq data=cursos;
    tables alumno*materia*calif /
    cmh2 scores=rank noprint;
run;
```


Prueba de Friedman en SAS

Procedimiento FREQ

Estadísticos de sumarización para materia por calif
Control para alumno

Estadísticos de Cochran-Mantel-Haenszel (Basado en puntuaciones de rango)				
Estadístico	Hipótesis alternativa	DF	Valor	Prob
1	Correlación nonzero	1	0.5000	0.4795
2	Diferencia de puntuaciones de la media de la fila	2	8.6667	0.0131

Tamaño total de la muestra = 27

Prueba de Nemenyi Friedman en SAS

```
data cursos;
input alumno materia $ calif @@;
cards;
1 Mat 98 1 Fisiol 95 1 Ento 77
2 Mat 95 2 Fisiol 71 2 Ento 79
3 Mat 76 3 Fisiol 80 3 Ento 91
4 Mat 95 4 Fisiol 81 4 Ento 84
5 Mat 83 5 Fisiol 77 5 Ento 80
6 Mat 99 6 Fisiol 70 6 Ento 93
7 Mat 82 7 Fisiol 80 7 Ento 87
8 Mat 75 8 Fisiol 72 8 Ento 81
9 Mat 88 9 Fisiol 81 9 Ento 83
;
proc rank;
  by alumno;
  var calif;
  ranks rango_calif;
run;
proc print; run;
proc anova;
  class alumno materia;
  model rango_calif = alumno materia;
  means materia/tukey lines;
run;
```

Resultados de prueba de Nemenyi Friedman en SAS

Procedimiento ANOVA

Prueba del rango estudentizado de Tukey (HSD) para rango_calif

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

Alpha	0.05
Grados de error de libertad	16
Error de cuadrado medio	0.583333
Valor crítico del rango estudentizado	3.64914
Diferencia significativa mínima	0.929

rango_calif Tukey Grouping for Means of materia (Alfa = 0.05)

Means cubiertas por la misma barra no son significativamente diferentes.

materia Estimación

Mat 2.5556

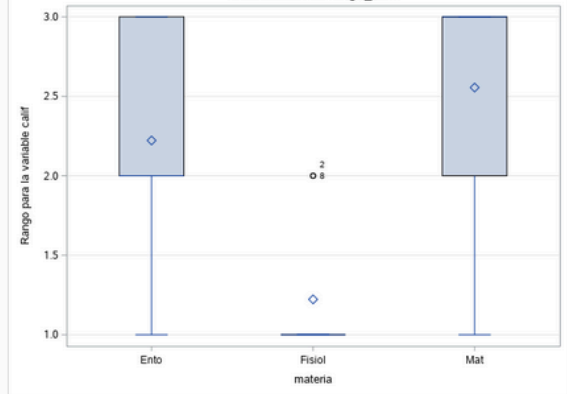
Ento 2.2222

Fisiol 1.2222



Procedimiento ANOVA

Distribución de rango_calif



Procedimiento ANOVA

Prueba del rango estudentizado de Tukey (HSD) para rango_calif

Prueba de Quade

Prueba de Quade

La prueba Quade se utiliza para datos e hipótesis similares a la prueba de Friedman, es decir, para diseños de bloques completos no replicados.

Quade es más potente que la prueba de Friedman en el caso de $k < 5$. La prueba de Quade es una generalización de la prueba de rango con signo de dos muestras. La prueba de Quade es apropiada para datos ordinales, pero se puede emplear en datos de intervalo o razón.

$H_0 : \theta_1 = \theta_2 = \dots = \theta_k$ vs H_1 : existe al menos un par (i, j) tal que $\theta_i \neq \theta_j$.

Primero, los datos se clasifican dentro de cada bloque (es decir, fila) para producir R_{ij} . Entonces el rango en cada bloque (máximo menos el valor mínimo de los datos) necesita ser calculado

Prueba de Quade

y ordenado, Q_i . Los scores son $S_{ij} = Q_i \times (R_{ij} - (k + 1)/2)$ y $S_j = \sum_{i=1}^b S_{ij}$ y el estadístico de prueba es $F_c = \frac{(b-1)B}{A-B}$ con $A = \sum_{i=1}^b \sum_{j=1}^k S_{ij}^2$ y $B = \frac{1}{b} \sum_{j=1}^k S_j^2$. Se rechaza H_0 si F_c es menor que un valor de la tabla F con un α dado y con $(k - 1)$ y $(b - 1) \times (k - 1)$ grados de libertad. La prueba de Quade es equivalente a realizar un ANOVA a los datos transformados a rangos dentro de cada bloque. Para comparación de Medias se puede aplicar la prueba de Nemenyi o Dunn.

Prueba de Quade: ejemplo

Utilizando los mismos datos que en la Prueba de Friedman.

$$H_0 = Me_{mat} = Me_{fis} = Me_{ent}$$

```
calif<-matrix(c(98, 95, 76, 95, 83, 99, 82, 75, 88,  
               95, 71, 80, 81, 77, 70, 80, 72, 81,  
               77, 79, 91, 84, 80, 93, 87, 81, 83),  
             nrow=9,ncol=3,  
             dimnames = list(1:9,c("Mat","Fis","Entom")))  
quade.test(calif)
```

```
##  
##      Quade test  
##  
## data:  calif  
## Quade F = 4.1765, num df = 2, denom df = 16, p-value = 0.03472
```

Prueba de Jonckheere-Terpstrata

Prueba de Jonckheere-Terpstrata

Para analizar k muestras independientes frente a una alternativa ordenada (p. Ej., grupos con dosis crecientes de tratamiento), se puede emplear la prueba según Jonckheere (1954). Ambas, la hipótesis nula y alternativa se pueden formular comparando las medianas (θ_i) de k poblaciones ($k > 2$). Así, la hipótesis nula, $H_0 : \theta_1 = \theta_2 = \dots = \theta_k$ contra la alternativa, $H_a : \theta_1 \leq \theta_2 \leq \dots \leq \theta_k$, con al menos una desigualdad estricta.

El estadístico es T_{JT} se define: $J = \sum U_{xy}$

Donde U_{xy} es el número de observaciones en el grupo y que son mayores que cada observación en el grupo x . La estadística de prueba estandarizada se calcula como

$$Z_c = \frac{J - E(J)}{\sqrt{\text{Var}(J)}} \quad \text{con} \quad E(J) = \frac{N^2 - \sum_{j=1}^k n_j^2}{4} \quad \text{y}$$

$$\text{Var}(J) = N^2(2N + 3) - \sum_{j=1}^k n_j^2(2n_j + 3)$$

Prueba de Jonckheere-Terpstrata

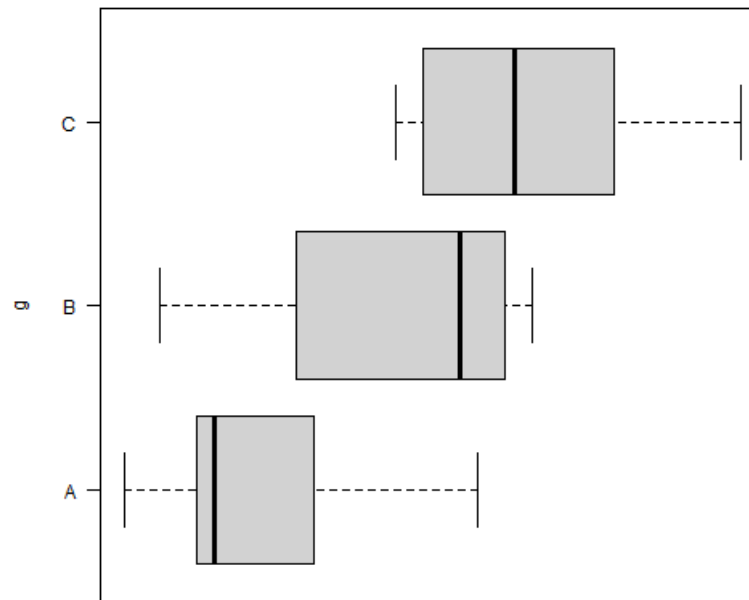
Dónde, N es el número total de observaciones en todos los grupos, n_j es la observación en el grupo j y j es el número de grupo y k es el número total de grupos.

Se Rechaza H_0 si $Z_c > Z_\alpha/2$ de tablas de Z.

Prueba de Jonckheere-Terpstrata Ejemplo

\scriptsize

```
library(PMCMR)
x <- c(106, 114, 116, 127, 145, 110, 125, 143, 148, 151,
      136, 139, 149, 160, 174)
g <- as.factor(c(rep(1,5), rep(2,5), rep(3,5)))
levels(g) <- c("A", "B", "C")
boxplot(x~g, horizontal=TRUE, las=1,)
```



Prueba de Jonckheere-Terpstrata Ejemplo

\scriptsize

```
library(PMCMRplus)
x <- c(106, 114, 116, 127, 145, 110, 125, 143, 148, 151,
      136, 139, 149, 160, 174)
g <- as.factor(c(rep(1,5), rep(2,5), rep(3,5)))
levels(g) <- c("A", "B", "C")
jonckheereTest(x , g)
```

```
##
##      Jonckheere-Terpstra test
##
## data:  x and g
## z = 2.2716, p-value = 0.02311
## alternative hypothesis: two.sided
## sample estimates:
## JT
## 59
```

Prueba de Jonckheere-Terpstrata, ejemplo r

$H_0 : \theta_1 = \theta_2 = \theta_3$ vs, $H_a : \theta_1 \leq \theta_2 \leq \theta_3$

\scriptsize

```
jonckheereTest(x , g)
```

Test statistic	P value	Alternative hypothesis	JT
----------------	---------	------------------------	----

2.272	0.02311 *	two.sided	59
-------	-----------	-----------	----

Table: Jonckheere-Terpstra test: x and g

se acepta hipotesis de tendencia creciente

Prueba de Jonckheere-Terpstrata, ejemplo SAS

```
data ejem1;  
input trat resp @@;  
cards;  
1 106 2 110 3 136  
1 114 2 125 3 139  
1 116 2 143 3 149  
1 127 2 148 3 160  
1 145 2 151 3 174  
;  
proc print; run;  
Jonckheere-Terpstra Trend Test for Continuous Data";  
proc freq data=ejem1;  
tables trat*resp/jt;  
run;
```

Prueba de Jonckheere-Terpstrata, ejemplo SAS

Estadísticos para la tabla de trat por resp

Test Jonckheere-Terpstra	
Estadístico	59.0000
Z	2.2716
Pr de un lado $\geq Z$	0.0116
Pr de dos lados $\geq Z $	0.0231

Tamaño de la muestra = 15

Prueba Q Cochran

Prueba Q Cochran

La prueba Q de Cochran, se utiliza cuando un grupo de n individuos son medidos k veces (k repeticiones), y la variable de interés tienen respuesta dicotómica (generalmente 1 denota el “éxito” y 0 el “fracaso”). Hipótesis nula: Las proporciones de aciertos son iguales en todas las repeticiones.

$$H_0 : P_1 = P_2 = \dots = P_k$$

Hipótesis alternativa: Por lo menos alguna proporción de acierto es diferente entre dos repeticiones.

$$H_a : P_i \neq P_j \text{ para algun } i \neq j.$$

Estadístico de prueba.

$$Q = \frac{k(k-1) \sum_{j=1}^k C_j^2 - (k-1)N^2}{kN - \sum_{i=1}^n R_i^2}$$

Prueba Q Cochran

Datos de X_{ij} son 0 o 1.

<i>Individuos</i>	1	2	...	k	$\sum_{j=1}^k x_{ij}$
1	X_{11}	X_{12}	...	X_{1k}	R_1
2	X_{21}	X_{22}	...	X_{2k}	R_2
.
n	X_{n1}	X_{n2}	...	X_{nk}	R_n
$\sum_{i=1}^n x_{ij}$	C_1	C_2	...	C_k	N

Estadístico de prueba.

$$Q = \frac{k(k-1) \sum_{j=1}^k C_j^2 - (k-1)N^2}{kN - \sum_{i=1}^n R_i^2}$$

Rechazo H_0 si $Q \geq \chi^2_{(k-1), \alpha}$.

Ejemplo Q Cochran

En un curso de Agroecología se realiza una encuesta a alumnos sobre su disposición a adoptar buenas prácticas de cuidados de césped ecológico: pruebas de suelo, medición de la cantidad de riego, aumento de la altura de corte y devolución de los cortes de césped al césped. Se quiere saber qué práctica los estudiantes estaban más dispuestos a adoptar (1=sí, 0=no).

Ejemplo Q Cochran

Sujeto	Clippings	Irrigation	MowHeight	SoilTest
a	1	1	1	1
b	0	0	0	0
c	0	0	1	1
d	1	0	1	1
e	1	0	1	0
f	0	0	1	1
g	0	0	1	1
h	1	0	1	1
i	1	1	0	0
j	1	0	1	1
k	1	0	1	1
l	0	0	1	1
m	1	0	1	1

Ejemplo Q Cochran

```
library(RVAideMemoire)
```

```
## Warning: package 'RVAideMemoire' was built under R version 4.2.3
```

```
respuesta<-c(1,0,0,1,1,0,0,1,1,1,1,0,1,0,1,0,0,0,0,0,1,
             ,0,0,0,0,0,1,
             0,1,1,1,1,1,1,0,1,1,1,1,1,1,0,1,1,0,1,1,1,0,1,
             ,1,1,1,1)
tratamientos<-c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,
                ,2,2,2,2,2,2,2,
                3,3,3,3,3,3,3,3,3,3,3,3,3,3,4,4,4,4,4,
                ,4,4,4,4,4,4,4,4,4)
sujeto<-c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,1,2,3,4,5,6,7,
          ,8,9,10,11,12,13,
          14,1,2,3,4,5,6,7,8,9,10,11,12,13,14,1,2,3,4,5,6,
          ,7,8,9,10,11,12,13,14)
```



Ejemplo Q Cochran

\scriptsize

```
cochran.qtest(respuesta~tratamientos|sujeto)
```

```
##
##      Cochran's Q test
##
## data:  respuesta by tratamientos, block = sujeto
## Q = 16.9535, df = 3, p-value = 0.0007225
## alternative hypothesis: true difference in probabilities is not equal to 0
## sample estimates:
## proba in group          <NA>          <NA>          <NA>
##      0.5714286      0.1428571      0.8571429      0.7857143
```

Si hay alguna preferencia;

Ejemplo Q Cochran SAS

\scriptsize

```
data agro;
input Sujeto$ Clippings Irrigation MowHeight SoilTest;
cards;
a 1 1 1 1
b 0 0 0 0
c 0 0 1 1
d 1 0 1 1
e 1 0 1 0
f 0 0 1 1
g 0 0 1 1
h 1 0 1 1
i 1 1 0 0
j 1 0 1 1
k 1 0 1 1
l 0 0 1 1
m 1 0 1 1
n 0 0 1 1
;
  Cochran Q test";
proc freq data=agro;
tables clippings*Irrigation*MowHeight*SoilTest / agree noprint;
run;
```

Ejemplo Q Cochran SAS

Procedimiento FREQ

Estadísticos de sumarización para MowHeight por SoilTest
Control para Clippings y Irrigation

Q de Cochran para Clippings por Irrigation por MowHeight por SoilTest		
Chi-cuadrado	DF	Pr > ChiSq
16.9535	3	0.0007

Tamaño total de la muestra = 14