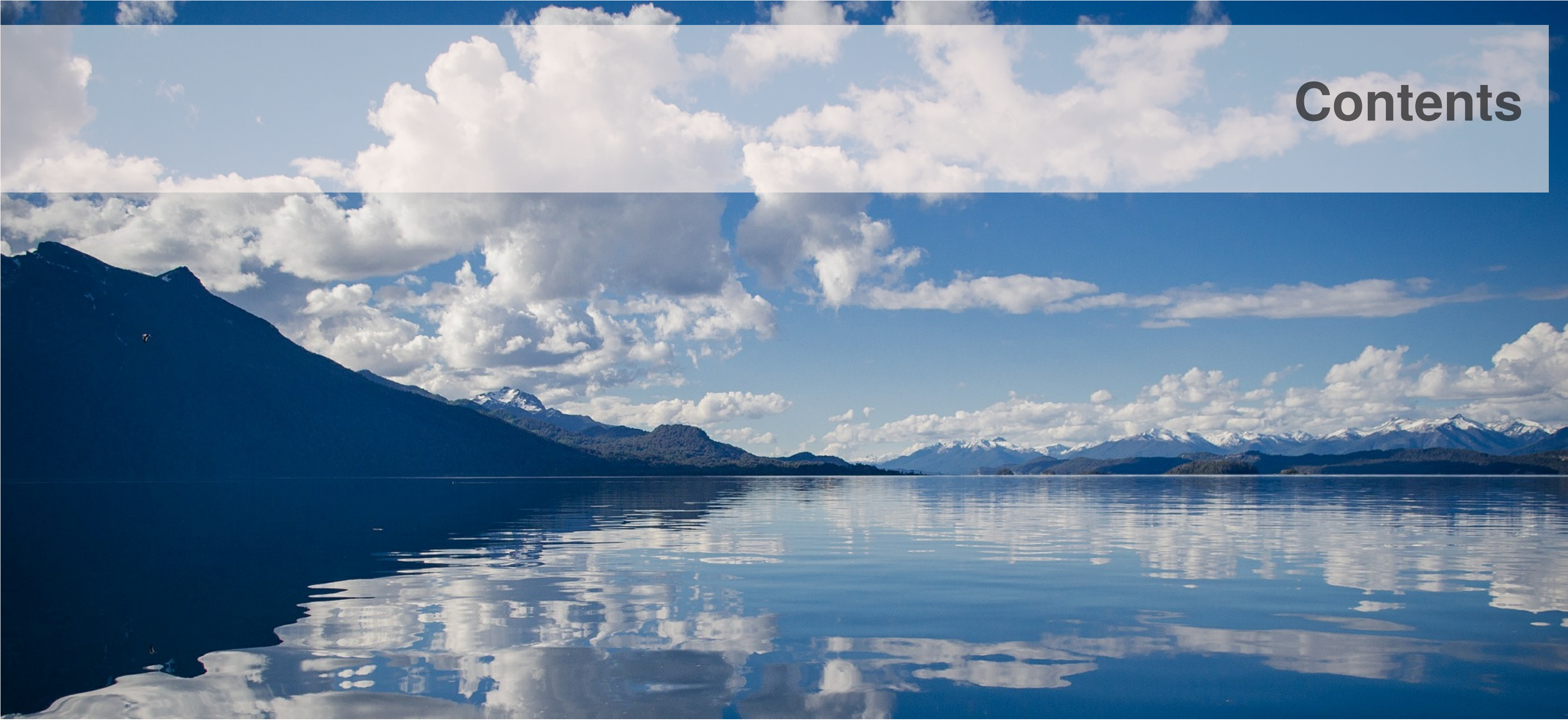
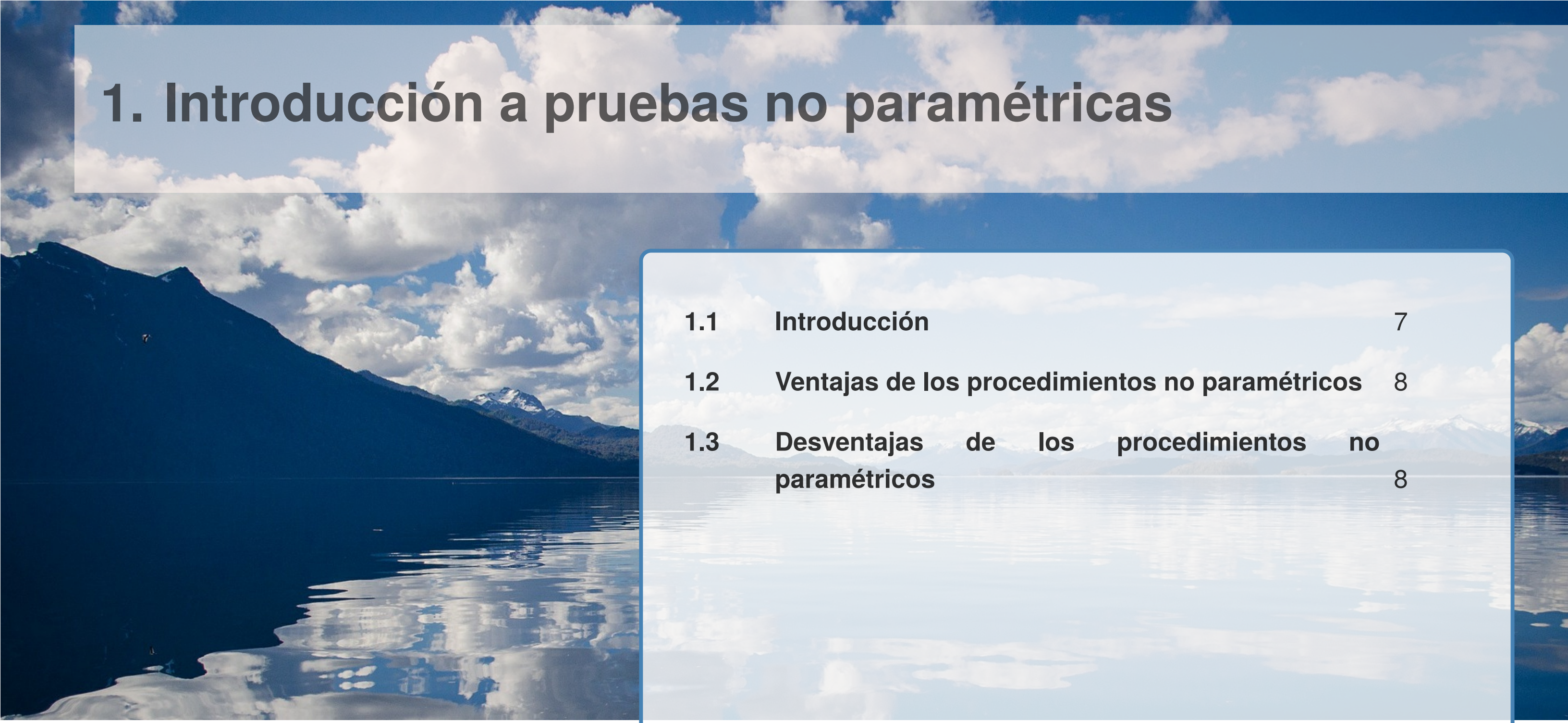
PRUEBAS NO PARAMÉTRICAS CLÁSICAS

Alondra Elizabeth Matos Mendoza
Eduardo Vilchis Marín

CIMAT



Contents



1. Introducción a pruebas no paramétricas

Una de las finalidades de la estadística es realizar inferencias con respecto a una población a partir de la información contenida en una muestra y proporcionar una medida correspondiente para la exactitud de la inferencia. Las inferencias se expresan como estimaciones de los parámetros, o bien, como pruebas de hipótesis respecto a valores de los parámetros. Esta primera sección introduce el concepto de pruebas no paramétricas.

1.1 Introducción

Las pruebas de hipótesis se aplican en todos los campos en los que la teoría se puede confrontar con la observación. Una vez establecida la hipótesis estadística, se recaban los datos que permiten decidir acerca de esa hipótesis y se realiza un procedimiento objetivo para rechazar o no tal hipótesis, asumiendo un cierto grado de riesgo de que la decisión tomada sea incorrecta.

Definition 1.1 **(Hipótesis estadística)** Una hipótesis estadística es una suposición que se hace acerca de los valores de uno o más parámetros de la distribución de la variable aleatoria de estudio en una o más poblaciones.

Definition 1.2 **(Prueba estadística de una hipótesis)** Una prueba estadística de una hipótesis es un procedimiento en el cual se usa una (varias) muestra(s) con el fin de determinar cuándo podemos rechazar o no rechazar la hipótesis propuesta.

En resumen, una hipótesis estadística es una aseveración sobre un modelo probabilístico y una prueba de hipótesis es un método para determinar sobre la plausibilidad de esa aseveración, usando la muestra como guía.

Definition 1.3 Los **métodos paramétricos** son aquellos que se aplican a problemas donde la(s) distribución(es) de la(s) cual(es) se toman las muestras están especificadas excepto para los valores de un número finito de parámetros. Los **métodos no paramétricos** se aplican en todos los otros casos.

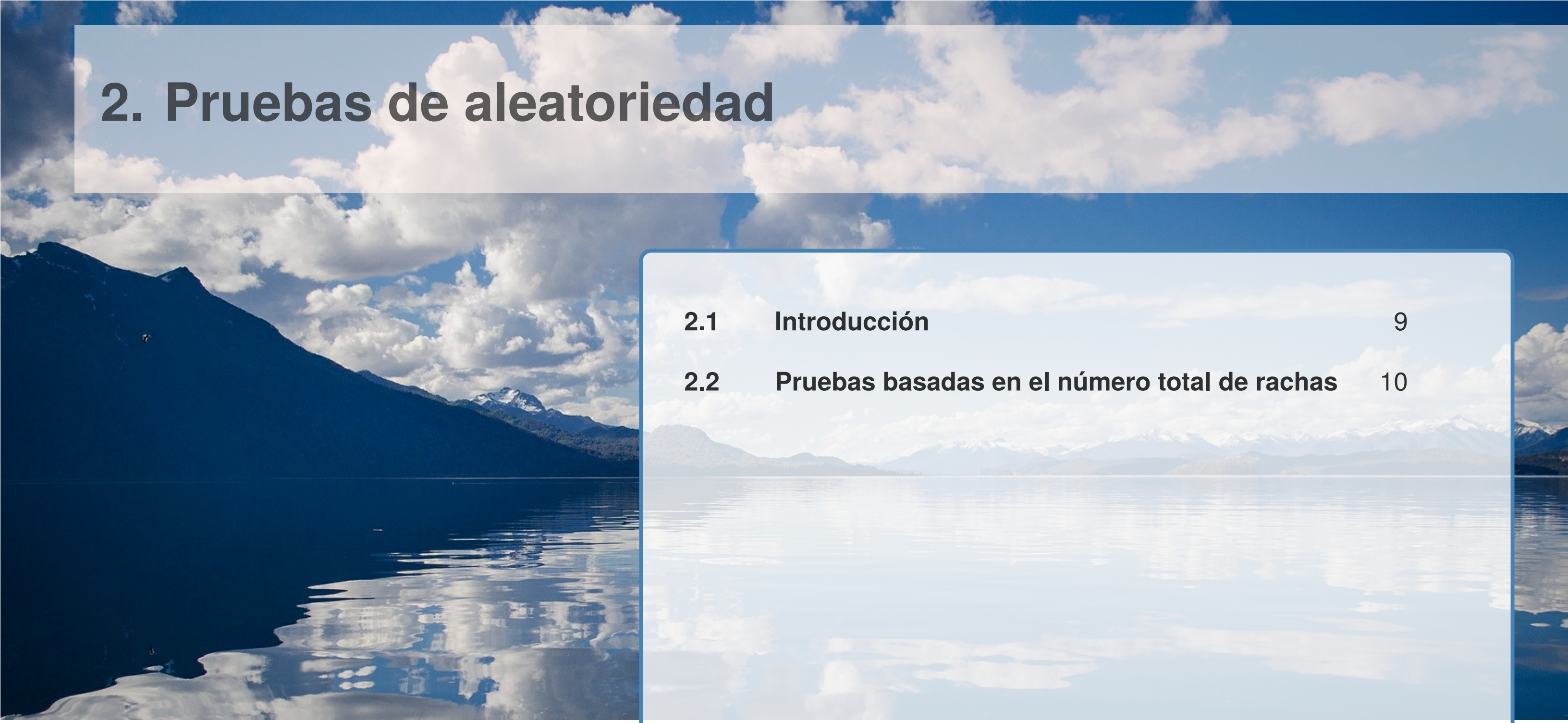
Al no especificar la distribución de la cual se selecciona la muestra, los métodos no paramétricos suelen también denominarse como *métodos de distribución libre*.Estas técnicas son útiles principalmente para analizar datos en las escalas nominal y ordinal. O bien, cuando no se cumplen los supuestos de alguna prueba paramérica y existe el método no paramétrico correspondiente.

1.2 Ventajas de los procedimientos no paramétricos

1. Si el tamaño de la muestra es muy pequeña y se desconoce la distribución de la población a la que proviene la muestra, usar una prueba no paramétrica sería una opción.
2. Las pruebas no paramétricas hacen menos suposiciones acerca de los datos.
3. Las pruebas no paramétricas son adecuadas para analizar datos que son inherentes a los rangos, así como datos cuyas puntuaciones numéricas tienen aparentemente la fuerza de los rangos.
4. Las pruebas no paramétricas también son apropiadas para analizar datos que son simplemente clasificatorios o categóricos.

1.3 Desventajas de los procedimientos no paramétricos

- El uso de procedimientos no paramétricos con datos que pueden manejarse con un procedimiento paramétrico conduce a un desperdicio de información.
- La aplicación de algunas de las pruebas no paramétricas puede ser laboriosa para muestras grandes.
- Si en los datos a analizar se cumplen todos los supuestos de un modelo estadístico paramétrico y se utiliza uno no paramétrico, el resultado sería deficiente. Este grado de falta de eficiencia es expresado por la potencia-eficacia de la prueba no paramétrica



2. Pruebas de aleatoriedad

En esta sección se detallará una técnica no paramétrica que se relaciona con inferencias sobre una propiedad de la distribución conjunta de un grupo de observaciones muestrales que están idénticamente distribuidas pero que posiblemente son dependientes.

2.1 Introducción

Imagina que hay una fila de 6 personas esperando comprar su boleto para una obra de teatro, supongamos que observamos un arreglo de 3 mujeres y 3 hombres en la fila de la siguiente manera: M,H,M,H,M,H. ¿Podría considerarse un arreglo aleatorio por género? De manera intuitiva, uno podría pensar que no, pues pareciera que la secuencia sigue un patrón intencinal de alternar por parejas. En situaciones menos extremas, la aleatoriedad puesde ser probada estadísticamente usando la teoría de rachas.

Definition 2.1 Dada una secuencia ordenada de uno o más tipos de símbolos, una **racha** *(run)* es definida como una sucesión de uno o más tipos de símbolos que son seguidos y precedido por un símbolo diferente o ningún símbolo en absoluto.

Example 2.1 (Racha) Considera el siguiente escenario: M,M,M,H,H,M,H,H,M,M. Se observa que el número total de rachas es 5, las cuales se muestran a continuación bajos distinto color: M,M,M,H,H,M,H,H,M,M

La falta de aleatoriedad se debe a la tendencia de los símbolos a mostrar un patrón definido en la secuencia, el cual se refleja a partir del número y la longitud de las rachas (que están correlacionados). Por lo tanto, las pruebas de aleatoriedad se basan en estos criterios. Muy pocas rachas, demasiadas rachas, una racha de longitud grande, etc, se pueden usar como criterio estadístico para el rechazo de la hipótesis nula de aleatoriedad, pues dichos escenarios raramente ocurren en una secuencia aleatoria.

Retomando el ejemplo del teatro, una secuencia con muchas rachas con tendencia de alternar el género, podría sugerir que la obra es popular entre jovénes y adultos que asisten en pareja.

2.2 Pruebas basadas en el número total de rachas

Hipótesis

$$H_0 : La\ secuencia\ es\ aleatoria$$

$$VS$$

$$H_1 : La\ secuencia\ no\ es\ aleatoria$$

Estadístico de Prueba

Considera una secuencia ordenada de n elementos de 2 tipos, n_1 del primero y n_2 del segundo, donde $n = n_1 + n_2$. Sea R_1 el número de rachas con elementos del tipo 1 y R_2 el número de rachas correspondientes al tipo 2, el total de rachas es $R = R_1 + R_2$. Entonces, para realizar la prueba de aleatoriedad basada en R, necesitamos la distribución de R cuando la hipótesis nula es cierta.

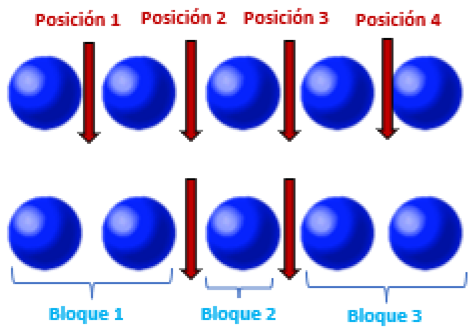
Distribución nula exacta de R

Primero se determinará la distribución conjunta de R_1 y R_2, para luego calcular la distribución de su suma R. Bajo la hipótesis nula de aleatoriedad, cada posible arreglo de $n_1 + n_2$ elementos es equiprobable, la probabilidad de que $R_1 = r_1$ y $R_2 = r_2$ es igual al número de arreglos distintos de $n_1 + n_2$ con r_1 y r_2 rachas del tipo 1 y tipo 2 respectivamente, dividido por el número total de arreglos distintos (todas las diferentes formas de ordenar los n elementos), que es $\frac{n!}{n_1!n_2!} = \binom{n_1+n_2}{n_1}$

Para determinar la cantidad del numerador, consideremos lo siguiente:

Supongamos que tenemos 5 bolas azules acomodadas en una fila y queremos dividirlas en 3 bloques insertando bloques de bolas rojas entre cualesquiera 2 bolas blancas. Como se ilustra a continuacón, hay 4 posiciones en la que cada bloque de bolas rojas puede ser puesta y como solo se necesitan 2 bloques de bolas rojas para determinar 3 bloques, el número de maneras distintas de distribuir 5 bolas en 3 bloquess (sin celdas vacías) es $\binom{4}{2}$

De manera general, el número de maneras distintas de distribuir n objetos iguales en r bloques (sin celdas vacías) es $\binom{n-1}{r-1}, n \geq r, r \geq 1$.



Por lo tanto, para obtener una secuencia con r_1 rachas del tipo 1, los n_1 objetos deben ser colocados en r_1 bloques, lo cual ocurre de diferentes $\binom{n_1-1}{r_1-1}$ maneras. Análogamente, con respecto a los objetos del tipo 2, hay $\binom{n_2-1}{r_2-1}$ maneras de colocar n_2 objetos en r_2 bloques. En consecuencia, el total de arreglos distintos comenzando con una racha del tipo 1 es $\binom{n_1-1}{r_1-1}\binom{n_2-1}{r_2-1}$. Similarmente, para una secuencia que inicia con una racha del tipo 2, los bloques de objetos del tipo 1 y tipo 2 se deben intercambiar, y por ende, $r_1 = r_2 \pm 1$ o $r_1 = r_2$.

- Si $r_1 = r_2 + 1$, la secuencia debe comenzar con una racha del tipo 1
- Si $r_1 = r_2 - 1$, la secuencia debe comenzar con una racha del tipo 2
- Si $r_1 = r_2$, la secuencia puede comenzar con una racha de cualquier tipo, y el número de arreglos distintos debe ser el doble.

Entonces, la distribución conjunta de R_1 y R_2 es de la siguiente forma:

Theorem 2.1 **(Distribución conjunta de R_1 y R_2)** Sea R_1 y R_2 número de rachas de los n_1 objetos del tipo 1 y n_2 objetos del tipo 2 de una muestra de tamaño $n = n_1 + n_2$. La probabilidad conjunta de R_1 y R_2 es

$$f_{R_1,R_2}(r_1,r_2) = c\frac{\binom{n_1-1}{r_1}\binom{n_2}{r_2-1}}{\binom{n_1+n_2}{n_1}}, r_1 = 1,2,...,n_1; r_2 = 1,2,...,n_2; r_1 = r_2 \ o \ r_1 = r_2 \pm 1$$

donde $c = 2$ si $r_1 = r_2$ y $c = 1$ si $r_1 = r_2 \pm 1$

Ahora bien, consideremos que r es par, entonces debe haber el mismo número de rachas para ambos tipos, por lo que los únicos valores posibles de r_1 y r_2 son $r_1 = r_2 = r/2$. De la distribución conjunta de R_1 y R_2, se obtiene que la función de densidad de R cuando $R = r$ es par viene dada por $2\frac{\binom{n_1-1}{r/2-1}\binom{n_2-1}{r/2-1}}{\binom{n_1+n_2}{n_1}}$. Por otro lado, r es impar si $r_1 = r_2 \pm 1$, por lo que r_1 y r_2 toman los siguientes pares de valores: $\left(r_1 = (r-1)/2,\ r2 = (r+1)/2\right)$ y $\left(r_1 = (r+1)/2,\ r2 =\right.$

$(r-1)/2\big)$; por lo tanto, la función de densidad de R cuando $R=r$ es impar está definido por $\frac{\binom{n_1-1}{((r-1)/2)-1}\binom{n_2-1}{(r+1)/2-1}+\binom{n_1-1}{(r+1)/2-1}\binom{n_2-1}{(r-1)/2-1}}{\binom{n_1+n_2}{n_1}}$.

Theorem 2.2 (Distribución de R) La distribución de R, el total de rachas de $n=n_1+n_2$ objetos, n_1 del tipo 1 y n_2 del tipo 2, es

$$f_R(r)=\begin{cases} 2\dfrac{\binom{n_1-1}{r/2-1}\binom{n_2-1}{r/2-1}}{\binom{n_1+n_2}{n_1}} & si\ r\ es\ par \\ \\ \dfrac{\binom{n_1-1}{(r-3)/2}\binom{n_2-1}{(r-1)/2}+\binom{n_1-1}{(r-1)/2}\binom{n_2-1}{(r-3)/2}}{\binom{n_1+n_2}{n_1}} & si\ r\ es\ impar \end{cases}$$

Para $r=2,3,\ldots,n_1+n_2$

Aproximación para muestras grandes

Los cálculos bajo la distribución exacta se vuelven laboriosos a menos que n_1 y n_2 sean ambos pequeños. Sin embargo, para muestras grandes $(n_1,n_2\geq 10)$, se puede usar la aproximación a la distribución nula.

Considerando que

$$\mu_R=\frac{2n_1n_2}{n_1+n_2}+1,\qquad \sigma_R=\sqrt{\frac{(2n_1n_2)(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

Estandarizando la variable aleatoria

$$Z=\frac{R-\mu_R}{\sigma_R}=\frac{R-\frac{2n_1n_2}{n_1+n_2}-1}{\sqrt{\frac{(2n_1n_2)(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}}$$

La función de propabilidad límite de Z es la densidad normal estándar.

Región de rechazo

El valor crítico es determinado por las tablas de probabilidades de R para n_1 y n_2. El mínimo valor crítico es el valor más chico que se encuentra en la tabla de probabilidades, mientras que el máximo es el valor crítico más grande de dicha tabla.

Para una prueba de dos colas con un nivel de significancia α, mediante la aproximación normal y considerando una corrección de continuidad del 0.5, la hipótesis nula de aleatoriedad será rechazada cuando

$$\frac{R - 0.5 - 1 - \frac{2n_1 n_2}{n_1+n_2}}{\sqrt{\frac{(2n_1 n_2)(2n_1 n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}} \geq z_{\alpha/2}$$

donde z_γ es el número que satisface $\phi(z_\gamma) = 1 - \gamma$ o, de manera equivalente, z_γ es el punto $(1-\gamma)$th cuantil de la distribución normal estándar.

Ejemplo

Exercise 2.1 La temperatura alta registrada en una ciudad turística de Florida para cada uno de 10 consecutivos días durante el mes de enero de este año se compara con el promedio histórico alto para el mismo día en años anteriores y señalado como por encima del histórico promedio (A) o por debajo (B). Para los datos A A B A B B A A A B, pruebe la hipótesis nula de dirección aleatoria de desviación de temperatura alta promedio contra la alternativa de no aleatoriedad, usando el nivel 0.05.

Example 2.2 (Solución:) Dado que los datos constan de seis A y cuatro B, haciendo B el tipo 1 se tiene que haga $n_1 = 4, n_2 = 6$. El número total de corridas observadas es $R = 6$ (A A B A B B A A B).

La tabla de probabilidades muestra que $P(R \leq 2) = 0.010$ y $P(R \geq 9) = 0.024$, y estos son las mayores probabilidades respectivas que no superan 0.025; la región de rechazo es entonces $R \leq 2$ o $R \geq 9$ con nivel exacto 0.034. Como $R = 6$ no cae en esta región, no rechazamos la hipótesis nula de aleatoriedad a un nivel de significancia del 0.05.

Ejemplo

Exercise 2.2 Durante 52 lunes consecutivos, se registró si llovió o no en Boston. Los lunes sin lluvia son representados como S(para seco), mientras que los lunes con lluvia son representados como L. Bajo esta notación, la siguiente secuencia representa los 52 lunes consecutivos. ¿Hay suficiente evidencia para sustentar la aseveración de que la lluvia de los lunes es aleatoria?

S S S S L S L S S L S S L S S S L S S L L L S S S S
L S L S L L L S L S S S L S S S L S L S S L S S S L

Example 2.3 (Solución:)

Hipótesis

H_0 : La secuencia es aleatoria (la lluvia de los lunes es aleatoria)

$$VS$$

H_1 :La secuencia no es aleatoria (la lluvia de los lunes no es aleatoria)

Estadístico de prueba

R =número de rachas

Cálculo del estadístico de prueba

Es fácil examinar la secuencia para ver que n_1 =número de S=33 n_2 =número de L=19
R =número de rachas =30

Como las muestras son grandes, se puede usar la aproximación normal.

Considerando que

$$\mu_R = \frac{2n_1 n_2}{n_1 + n_2} + 1 = \frac{2(33)(19)}{33+19} + 1 = 25.115$$

$$sigma_R = \sqrt{\frac{(2n_1 n_2)(2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}} = 3.306$$

Entonces, el valor del estadístico de prueba es:

$$z = \frac{R - \mu_R}{\sigma_R} = \frac{30 - 25.115}{3.306} = 1.48$$

Región de rechazo

Como $P(Z > 1.96) = 0.05$, entonces la región de rechazo es:

$$RR : Z > 1.96 \;\; o \;\; Z < -1.96$$

El estadístico de prueba $z = 1.48$ no cae dentro de la región crítica

Cálculo del Valor P

$$ValorP = 2P(Z \geq 1.48) = 0.13$$

Decisión estadística con $\alpha = 0.05$

Como $Valor\ P = 0.13 > 0.05$, no se rechaza la hipótesis nula de aleatoriedad. A un nivel de confianza del 95%, la secuencia parece ser aleatoria.

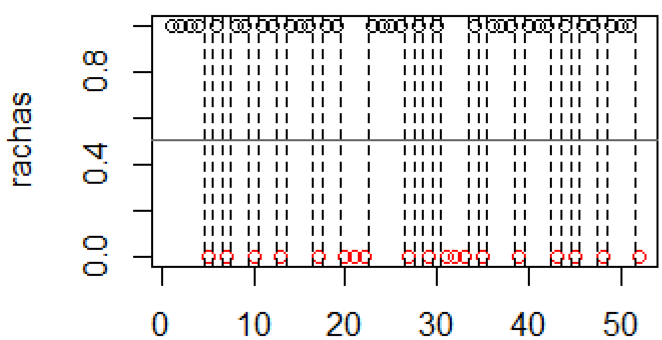
Código

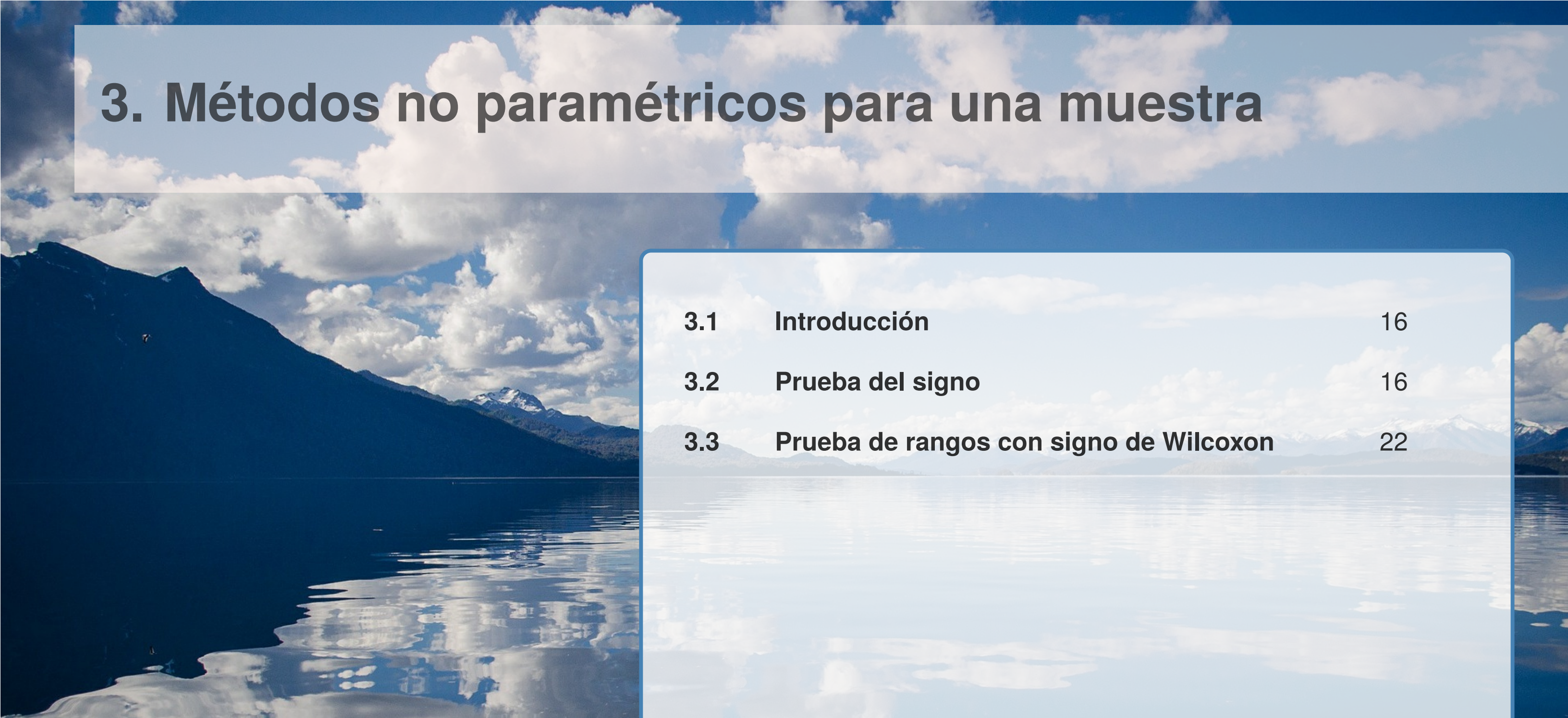
```r
library(randtests)

secuencia<-c("S","S", "S","S", "L", "S", "L", "S", "S", "L", "S", "S", "L",
             "S", "S", "S", "L", "S", "S", "L", "L", "L", "S", "S", "S", "S",
             "L", "S", "L", "S", "L", "L", "L", "S", "L", "S", "S", "S", "L",
             "S", "S", "S", "L", "S", "L", "S", "S", "L", "S", "S", "S", "L")

rachas<-ifelse(secuencia =="S",1,0)

runs.test(rachas,alternative = "two.sided",threshold = 0.5,pvalue = "normal",
    plot=T)
```

Listing 2.1: Prueba de rangos por aproximación en R

Salida

```
        Runs Test

data:  rachas
statistic = 1.4775, runs = 30, n1 = 33, n2 = 19, n = 52, p-value = 0.1396
alternative hypothesis: nonrandomness
```





3. Métodos no paramétricos para una muestra

En esta sección se detallarán técnicas no paramétricas para la realización de inferencias en problemas que incluyan una muestra.

3.1 Introducción

Las hipótesis que se muestran en esta sección se enfocan en probar si existe una diferencia significativa en la localizacón (tendencia central) entre la muestra y la población. Las hipótesis se refieren a la mediana o algún otro cuantil en lugar de la media como parámetro de ubicación, pero tanto la media como la mediana son buenos índices de tendencia central y coinciden para poblaciones simétricas. En cualquier población, la mediana siempre existe (lo que no es cierto para la media) y es más robusta como una estimación de la ubicación.

3.2 Prueba del signo

Supuestos:

- Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con mediana (M) desconocida, $F_X^{-1}(0.5) = M$.
- La variable de interés es medida en escala al menos ordinal.

Hipótesis:

a) $H_0 : M = M_0$, $H_1 : M \neq M_0$ (alternativa de dos colas)

b) $H_0 : M \leq M_0$, $H_1 : M > M_0$(alternativa de cola superior)

c) $H_0 : M \geq M_0$, $H_1 : M < M_0$ (alternativa de cola inferior) donde M_0 denota un valor de la mediana.

Estadístico de Prueba

Para $H_0 : M = M_0$, $H_1 : M \neq M_0$, si los datos de la muestra son consistentes con el valor hipotético de la mediana, en promedio la mitad de las observaciones de la muestra estará por encima de M_0 y la mitad por debajo. Por lo tanto, sea $K = min\{$número de observaciones mayores que M_0, número de observaciones menores que $M_0\}$,es decir, el mínimo entre el número de signos positivos y el número de signos negativos obtenidos de las diferencias $X_i - M_0,\ i = 1, 2, ..., n$; K se puede utilizar para probar la validez de la hipótesis nula, pues un número suficientemente pequeño de signos positivos, o bien, de signos negativos, es evidencia para rechazar H_0 ; pues si $H_0 : M = M_0$ es cierta, cualquier diferencia $X_i - M_0$ tiene la misma probabilidad de ser positiva que de ser negativa, es decir, 0.5.

Al ser dicotomizadas las observaciones muestrales, estas constituyen un conjunto de n variables aleatorias independientes que se distribuyen Bernoulli con parámetro $\theta = P(X > M_0) = P(X < M_0) = 0.5$. En consecuencia, $K \sim Binomial(n, \theta = 0.5)$.

Procediendo en forma análoga para la otras hipótesis alternativas planteadas, el estadístico de prueba para cada alternativa es:

a) $K = min\{K_-, K_+\}$

b) $K = número\ de\ signos\ negativos\ \ (K_-)$

c) $K = número\ de\ signos\ positivos\ \ (K_+)$

En resumen $K =$ número de signos de interés bajo la hipótesis alternativa dada.

Diferencias cero (empates)

El procedimiento para manejar los ceros (empates) es eliminarlos del análisis, lo cual reduce el tamaño de la muestra n, es decir, la prueba del signo se aplica al resto de los datos.

Región de rechazo

a) Para $H_1 : M \neq M_0$, la región de rechazo (de dos colas) es $K \geq k_{\alpha/2}$ o $K \leq k'_{\alpha/2}$, donde $k_{\alpha/2}$ es el entero más pequeño que satisface $P(K \geq k_{\alpha/2}|H_0) = \sum_{i=k_{\alpha/2}}^{N} \binom{N}{i}(0.5)^N \leq \alpha/2$ y k'_α es el entero más grande que satisface $P(K \leq k'_\alpha|H_0) = \sum_{i=0}^{k'_{\alpha/2}} \binom{N}{i}(0.5)^N \leq \alpha/2$

b) Para $H_1 : M > M_0$, la región de rechazo (de cola superior) es $K \geq k_\alpha$, donde k_α es el entero más pequeño que satisface $P(K \geq k_\alpha|H_0) = \sum_{i=k_\alpha}^{N} \binom{N}{i}(0.5)^N \leq \alpha$.

c) Para $H_1 : M < M_0$, la región de rechazo (de cola inferior) es $K \leq k'_\alpha$, donde k'_α es el entero más grande que satisface $P(K \leq k'_\alpha|H_0) = \sum_{i=0}^{k'_\alpha} \binom{N}{i}(0.5)^N \leq \alpha$.

Cálculo del valor P

Como la distribución binomial es simétrica cuando $\theta = 0.5$,para las hipótesis alternativas planteadas en $b)$ y $c)$:

$$Valor\ P = P(K \geq k|n, p = 0.5) \quad k\ es\ un\ valor\ del\ estad\acute{\imath}stico\ de\ prueba\ K$$

En el caso de la hipótesis de dos colas $a)$:

$$Valor\ P = 2P(K \geq k|n, p = 0.5)$$

Aproximación Normal

Sabemos que la aproximación normal a la binomial es buena cuando $\theta = 0.5$, entonces la aproximación normal a la binomial se puede utilizar. Una regla empíirica para aproximar la distribución binomial mediante la distribución normal es $np > 5$ y $nq > 5$.

Considerando que se aproxima una distribución discreta mediante una distribución continua, se utiliza un factor de corrección de continuidad de 0.5. Por lo tanto, el estadístico de prueba es:

$$Z = \frac{(K + 0.5) - (0.5)n}{\sqrt{0.25n}} = \frac{(K + 0.5) - (0.5)n}{(0.05)\sqrt{n}} \dot{\sim} N(0, 1)$$

El cálculo del $Valor\ p$ se realiza como en las pruebas estadísticas basadas en la distribución normal estándar.

Potencia de la prueba

Por definición, el poder de la prueba del signo frente a la alternativa H_1 es la probabilidad:

$$Pw(\theta) = P(K \geq k_\alpha|H_1)$$

Bajo H_1, la distribución de K es binomial con parámetros n y $\theta = P(X_i > M_0|H_1)$, así que potencia se puede expresar como:

$$Pw(\theta) = \sum_{i=k_\alpha}^{n} \binom{n}{i}(\theta^i)(1-\theta)^{n-i}$$

donde K_α es el entero más pequeño que satisface $\sum_{i=k_\alpha}^{n} \binom{n}{i}(0.5)^n \leq \alpha$.

Para evaluar la función de potencia para la prueba de signo, primero encontramos el valor crítico k_a para un nivel de significancia α dado, digamos 0.05. Entonces se calcula la probabilidad $\theta = P(X_i > M_0|H_1)$.

Ejemplo

Exercise 3.1 Los investigadores desean saber si al instruir en cuidados y aseo personal a una muestra de niñas con retraso mental mejoraría su apariencia. Se eligió aleatoriamente a 10 niñas de una escuela para niños con retraso mental, para que recibieran educación especial sobre cuidado y aseo personal. Dos semanas después de concluir el curso de instrucción, las niñas fueron entrevistadas por una enfermera y una trabajadora social, quienes asignaron a cada niña una calificación basada en su apariencia general. Los investigadores creían que, como máximo, las calificaciones alcanzarían el nivel de una escala ordinal. Creían que aunque una calificación de, digamos 8, representaba una apariencia mejor que una de 6, no podían decir que la diferencia entre las calificaciones de 6 y 8 era igual a la diferencia entre las calificaciones 8 y 10. Las calificaciones se muestran en la tabla 1. Se desea saber si es posible concluir que la calificación mediana de la población de la que se supone se extrajo la muestra es diferente de 5.

Table 3.1: Calificaciones de apariencia general de 10 niñas con retraso mental

Niña	Calificación
1	4
2	5
3	8
4	8
5	9
6	6
7	10
8	7
9	6
10	6

Example 3.1 (Solución:)

Hipótesis

$H_0 : M = 5 \ vs \ H_1 : M \neq 5$

Estadístico de prueba

$K = min\{$número de signos negativos, número de signos positivos$\}$

Cálculo de estadístico de prueba

Para determinar el valor del estadístico de prueba es necesario obtener los signos de las diferencias $(X_i - 5 < 0)$,como se muestra a continuacíon.

Nia	$Calificación$	$X_i - 5$
1	4	$-$
2	5	0
3	8	$+$
4	8	$+$
5	9	$+$
6	6	$+$
7	10	$+$
8	7	$+$
9	6	$+$
10	6	$+$

Así, $K = min\{1, 8\} = 1$

Cálculo del Valor P

$P = 2P(K \leq k|n, p = 0.5) = 2P(K \leq 1|n = 9, p = 0.5) = 2(0.0195) = 0.0390$. Note que $n = 9$ y no 10, esto se debe a que se obtuvo un empate.

Decisión estadística con $\alpha = 0.05$

Como $P = 0.0390 < 0.05$ se rechaza la hipótesis nula y se decide por la alternativa, es decir, la calificación mediana de apariencia de las niñas resultó significativamente diferente a 5.

Código

```r
library(BSDA)
datos = c(4,5,8,8,9,6,10,7,6,6)
SIGN.test(datos,md=5 ,alternative = "two.sided", conf.level = 0.95)
```

Listing 3.1: Prueba del signo en R

Salida

Otra manera de realizar la prueba:

```r
p.signo.bi <- function(muestra, H.null, a){
  #  "+" Mayor a la H. nula
  #  "-" Menor a la H. nula
  #  "0" igual a la H. nula
  signos = c()
  for (valor in datos){
    if (valor > H.null){

      signos <- c(signos, '+')
```

```
                One-sample Sign-Test

data:  datos
s = 8, p-value = 0.03906
alternative hypothesis: true median is not equal to 5
95 percent confidence interval:
 5.324444 8.675556
sample estimates:
median of x
        6.5

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.8906 6.0000 8.0000
Interpolated CI       0.9500 5.3244 8.6756
Upper Achieved CI     0.9785 5.0000 9.0000
```

```

    }else if (valor < H.null){

      signos <- c(signos, '-')

    }else{
      signos <- c(signos, '0')
    }
  }
  # Frecuencia de signos
  frecuencia <- table(signos)
  k = frecuencia[["-"]]
  n = frecuencia[["-"]] + frecuencia[["+"]]

  # Calcular P-valor
  p.value = pbinom(k, n, 0.5) * 2

  print(paste("Número de signos negativos:", frecuencia[["-"]]))
  print(paste("Número de signos positivos:", frecuencia[["+"]]))
  print(paste("P-valor:", p.value))
  print(paste("Mediana:", median(muestra)))
  if (p.value < a){
    print(paste("Se rechaza la hipótesis nula y se decide por la alternativa, la mediana es diferente a", H.null))

  }else{
    print(paste("No hay suficiente evidencia estadística para indicar que la mediana no es", H.null))
  }
}

# Muestra
```

```
datos = c(4,5,8,8,9,6,10,7,6,6)

# Hipótesis
H.null = 5

p.signo.bi(muestra = datos, H.null = H.null, a = 0.05)
```

Listing 3.2: Función creada para la prueba del signo

```
# Muestra
datos = c(4,5,8,8,9,6,10,7,6,6)

# Hipótesis
H.null = 5

p.signo.bi(muestra = datos, H.null = H.null, a = 0.05)
```

Listing 3.3: Evaluación de la función creada

```
## [1] "Número de signos negativos: 1"
## [1] "Número de signos positivos: 8"
## [1] "P-valor: 0.0390625"
## [1] "Mediana: 6.5"
## [1] "Se rechaza la hipótesis nula y se decide por la alternativa, la mediana e
```

3.3 Prueba de rangos con signo de Wilcoxon

Si los datos son medidos al menos en una escala de intervalo, utilizar la prueba del signo podría desperdiciar información contenida en los datos. Un método estadístico no paramétrico más apropiado es la prueba de rangos con signo de Wilcoxon, que desperdicia menos información, ya que toma en consideración las magnitudes de las diferencias $X_i - M_0$ en lugar de sólo los signos de dichas diferencias.

Supuestos:

- Sea $X_1, X_2, ..., X_n$ una muestra aleatoria de una población con mediana (M) desconocida.
- La variable de interés es medida en escala al menos de intervalo.
- La población muestreada es simétrica, pues se asume que la pobación es simétrica sobre M.

Hipótesis:

a) $H_0 : M = M_0$, $H_1 : M \neq M_0$ (alternativa de dos colas)

b) $H_0 : M \leq M_0$, $H_1 : M > M_0$(alternativa de cola superior)

c) $H_0 : M \geq M_0$, $H_1 : M < M_0$ (alternativa de cola inferior)

donde M_0 denota un valor de la mediana.

Estadístico de Prueba

Bajo $H_0 : M = M_0$ cierta, las diferencias $D_i = X_i - M_0$ son simétricamente distribuidas sobre 0. Supongamos que ordenamos las diferencias absolutas $|D_1|, |D_2|, ..., |D_n|$ de manera ascendente y les asignamos rangos $1, 2, ..., N$, haciendo un seguimiento de los signos originales de la diferencias D. Si M_0 es la verdadera mediana de la población simétrica, se esperaría que la suma de los rangos de las diferencias positiva (denotado por T^+) sea igual a la suma de los rangos de las diferencias negativas, (denotado por T^-).

Example 3.2 Supongamos que $H_0 : M = 53$, la siguiente tabla muestra un ejemplo del cálculo para determinar el estadístico de prueba.

Datos	$d_i = x_i - 53$	$\|d_i\|$	Rango de $\|d_i\|$	Rango de d_i positivas	Rango de d_i negativas
67	14	14	5	5	
45	-8	8	3		3
58	5	5	2	2	
43	-10	10	4		4
49	-4	4	1		1
				T^+ =7	$T^+ = 8$

Como se observa, el rango i es asociado con un signo de $+$ o de $-$ de acuerdo con el signo de $D_j = X_j - M_0$, donde D_j ocupa la i-ésima posición en el arreglo ordenado del valor absoluto de las diferencias $|D_j|$. Si definimos $r(\dot{})$ como el rango de una variable aleatoria, el estadístico puede expresarse como

$$T^+ = \sum_{i=1}^{N} Z_i r(|D_1|) \quad T^- = \sum_{i=1}^{N} (1 - Z_i) r(|D_1|)$$

donde

$$Z_i = \begin{cases} 1 & si\ D_i > 0 \\ 0 & si\ D_i \leq 0 \end{cases}$$

Con la hipótesis nula cierta, las Z_i son variables aleatorieas Bernoulli independientes e idénticamente distribuidas con $P(Z_i = 1) = P(Z_i = 0) = 1/2$, así que $E(Z_i) = 1/2$ y $Var(Z_i) =$

$1/4$. Recordando que $\sum_{i=0}^{n} i = \frac{n(n+1)}{2}$ y $\sum_{i=0}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$, se tiene que

$$E(T^+|H_0) = \sum_{i=1}^{n} \frac{r(|D_1|)}{2} = \frac{n(n+1)}{4}$$

Además, considerando que Z_i es independiente de $r(|D_i|)$ bajo H_0 cierta, se puede demostrar que

$$Var(T^+|H_0) = \sum_{i=1}^{n} \frac{(r(|D_1|))^2}{4} = \frac{n(n+1)(2n+1)}{24}$$

Además, se tiene la siguiente relación entre las sumas de rangos de las diferencias:

$$T^+ + T^- = \frac{n(n+1)}{2}$$

Grandes valores de T^+ corresponden a pequeños valores de T^-.

Nota: *Por simetría los resultados son válidos si se sustituye T^- en vez de T^+*

Si todas las diferencias tienen el mismo signo, positivo o negativo, los valores extremos de T^+ son 0 y $\frac{n(n+1)}{2}$.Como T^+ es completamente determinado por las variables indicadoras Z_i, el espacio muestral puede considerarse como un conjunto de todas las tuplas $(z_1, z_2, ..., z_n)$ posibles (y equiprobables bajo h_0) con componentes 1 o 0. Por lo tanto, la cardinalidad del espacio muestral es 2^n. Entonces la distribución de T^+ está dada por

$$P(T^+ = t) = \frac{u(t)}{2^n}$$

donde $u(t)$ es el número de maneras de asignar signos de $+$ y $-$ a los primeros n enteros, tal que la suma de los enteros positivos es igual a t.

T^+ y T^- son idénticamente distribuidos bajo H_0

$$\begin{aligned}
P(T^+ \geq c) &= P(T^+ - \frac{n(n+1)}{4}) \geq c - \frac{n(n+1)}{4}) \\
&= P(\frac{n(n+1)}{4}) - T^+ \geq c - \frac{n(n+1)}{4}) \\
&= P(\frac{n(n+1)}{2}) - T^+ \geq c) \\
&= P(T^- \geq c)
\end{aligned}$$

El estadístico de prueba depende del tipo de la hipótesis alternativa.

i) Si $H_0 : M = M_0$ vs $H_1 : M \neq M_0$, entonces $T = min\{T^+, T^-\}$.

ii) Si $H_0 : M \geq M_0$ vs $H_1 : M < M_0$,entonces $T = T^+$. Nota que si $H_0 : M \geq M_0$ es cierta, se espera un valor grande de T^+. Por lo tanto, cuando la hipótesis alternativa es $H_1 : M < M_0$, un valor suficientemente pequeño de T^+ causa el rechazo de H_0.

iii Si $H_0 : M \leq M_0$ vs $H_1 : M > M_0$, entonces $T = T^-$. Nota que un valor un valor suficientemente pequeño de T^- causa el rechazo de H_0.

Región de rechazo

Sea t_α es el número tal que $P(T \leq t_\alpha) = \alpha$,

a) Para $H_1 : M \neq M_0$, la región de rechazo (de dos colas) es $T^+ \leq t_{\alpha/2}$ o $T^- \leq t_{\alpha/2}$

b) Para $H_1 : M > M_0$, la región de rechazo (de cola superior) es $T^- \leq t_\alpha$

c) Para $H_1 : M < M_0$, la región de rechazo (de cola inferior) es $T^+ \leq t_\alpha$

Cálculo del valor P

Como la distribución binomial es simétrica cuando $\theta = 0.5$,para las hipótesis alternativas planteadas en $b)$ y $c)$:

$$ValorP = P(K \geq k|n, p = 0.5) \;\; k\ es\ un\ valor\ del\ estad\acute{i}stico\ de\ prueba\ K$$

En el caso de la hipótesis de dos colas $a)$:

$$ValorP = 2P(K \geq k|n, p = 0.5)$$

Aproximación Normal

Del teorema del límite central, T se distribuye asintóticamente normal. Entonces,

$$Z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \dot{\sim} N(0,1)\ cuando\ n \longrightarrow \infty$$

Para las alternativas de una cola, sustituya T con $T = T^+$ o $T = T^-$.

El cálculo del $valor\ p$ se realiza como en las pruebas estadísticas basadas en la distribución normal estándar.

Correción por empates

Cuando dos o más valores absolutos de diferencias son iguales, es decir, $|d_i| = |d_j|$ para al menos un $i \neq j$, las observaciones esán empatadas y el rango que se les asigna es el rango medio (promedio de los rangos que tendrían si fueran distintos). La varianza de los rangos se reduce.

$$Var(T^+|H_0) = \frac{N(N+1)(2n+1)}{24} - \frac{\sum t(t^2-1)}{48}$$

Por lo tanto, el denominador de Z se reemplaza por $\sqrt{\frac{N(N+1)(2n+1)}{24} - \frac{\sum t(t^2-1)}{48}}$.

Potencia de la prueba

La distribución de T^+ es aproximadamente normal para tamaños de muestra grandes, independientemente de si la hipótesis nula es verdadera. Por lo tanto, se puede calcular una aproximación de muestra grande a la potencia.

La eficiencia relativa asintótica de la prueba de rango con signo de Wilcoxon (relativa a la prueba t) es al menos 0.8644 para cualquier distribución continua y simétrica alrededor de cero, $es0.955$ para la distribución normal, y es 1.5 para el doble distribución exponencial.

Cabe señalar que la distribución de probabilidad de T^+ no es simétrica cuando la hipótesis nula no es cierta. Además, T^+ y T^- no son idénticamente distribuidos cuando la hipótesis nula no es cierta. Sin embargo, todavía podemos encontrar la distribución de probabilidad de T a partir de la de T^+ usando la relación:

$$P((T^- = k) = P(\frac{N(N+1)}{2}) - T^+ = k$$

Ejemplo

Exercise 3.2 En un estudio sobre abuso de drogas en un área urbana, los investigadores encontraron que la mediana del IQ de los detenidos de 16 años de edad o más que abusan de las drogas fue 107. Suponga que un investigador desea saber en otra área urbana, si la mediana del IQ de los detenidos que abusan de las drogas con edad mayor o igual a 16 años es diferente de 107. La siguiente tabla contiene los QI de una muestra aleatoria de 15 personas de la población de interés. ¿Qué puede concluir el investigador? Utilice un nivel de significación de 0.05

Table 3.2: QI de 15 personas con edad mayor o igual a 16 años que abusan de las drogas en cierta área urbana.

99	100	90	94	135	108	107	111	119	104	127	109	117	105	125

Example 3.3 (Solución:)

Hipótesis

$H_0 : M = 107 \ vs \ H_1 : M \neq 107$

Estadístico de prueba

$K = min\{T^+, T^-\}$

Cálculo del estadístico de prueba

La sigiente tabla resume los cálculos para determinar el valor del estaístico de prueba con los datos de IQ.

IQ	$d_i = x_i - 107$	$\|d_i\|$	$Rango\ de\ \|d_i\|$	$Rango\ de\ d_i\ positiva$	$Rango\ de\ d_i\ negativas$
99	-8	8	7		7
100	-7	7	6		6
90	-17	17	11		11
94	-13	13	10		10
135	$+28$	28	14	14	
108	$+1$	1	1	1	
107	0	0	$-----$	$------$	$------$
111	$+4$	4	5	5	
119	$+12$	12	9	9	
104	-3	3	4		4
127	$+20$	20	13	13	
109	$+2$	2	2.5	2.5	
117	$+10$	10	8	8	
105	-2	2	2.5		2.5
125	$+18$	18	12	12	
				$T^- = 64.5$	$T^- = 40.5$

Así, $T = min\{64.5, 40.5\} = 40.5$

Región de rechazo

la Tabla tabulada de probabilidades para el estadístico con $n = 14$,para $\frac{\alpha}{2} = 0.025$, se obtiene que la región de rechazo es:

$$RR : T \leq T_{n,\alpha/2} = T_{14,0.025} = 21$$

Como $T = 40.5 > 21$, no se rechaza la hipótesis nula con nivel de significación 0.05, es decir, la mediana del IQ no difirió significativamente de 107.

Cálculo del Valor P

Con base en la Tabla tabulada de probabilidades para el estadístico con $n = 14$, se observa que el valor del estadístico de prueba $T = 40.5$ se encuentra entre los valores tabulados de $40(P = 0.2316)$ y $41(P = 0.2508)$.Por lo tanto, se tiene que el $valorP$ de la prueba se encuentra en el siguiente intervalo:

$$2(0.2316) < P < 2(0.2508)$$

$$0.4632 < P < 0.5016$$

Decisión estadística con $\alpha = 0.05$

Como $P > 0.05$ no se rechaza la hipótesis nula, es decir, la mediana del IQ de los detenidos que abusan de las drogas con edad mayor o igual a 16 años de cierta área urbana no difirió significativamente de 107.

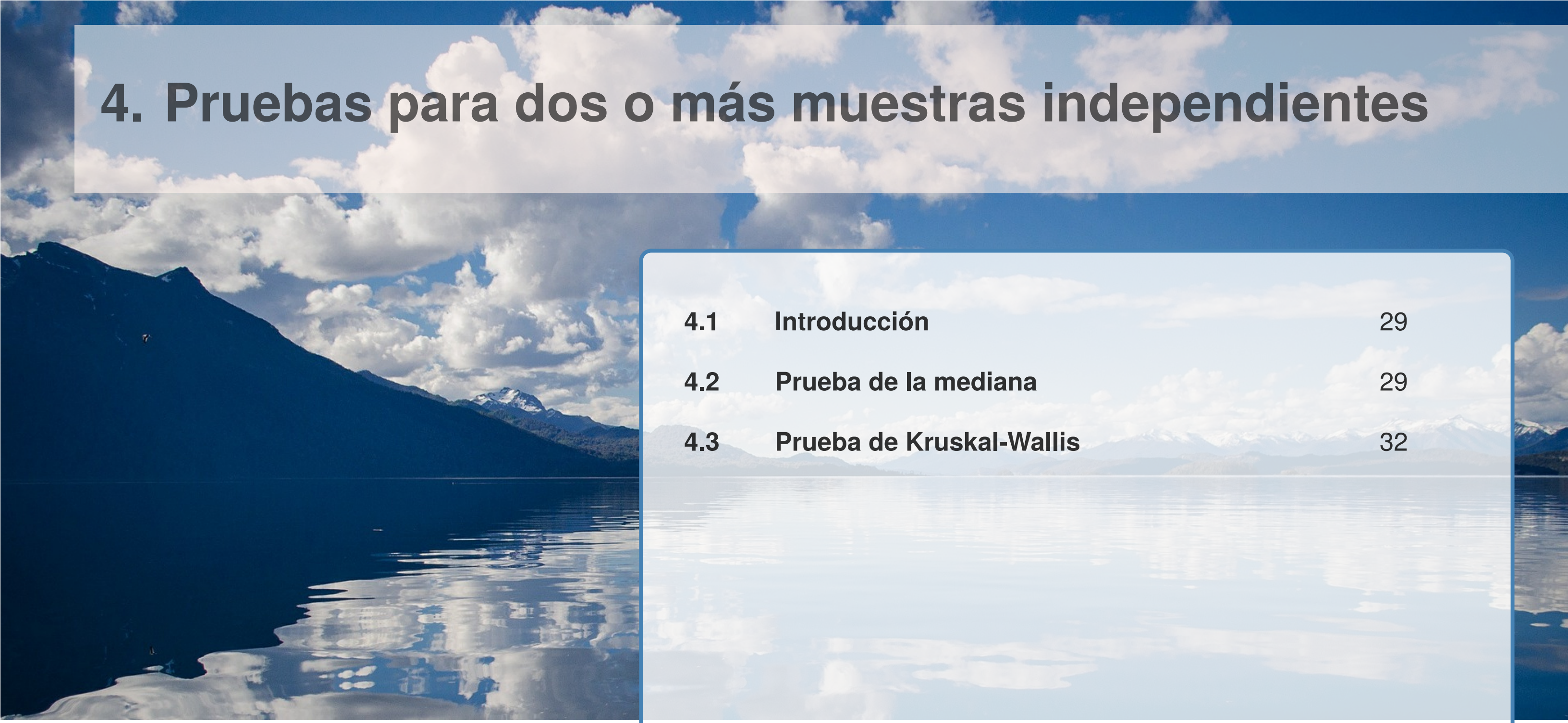
Código

```
muestra<-c(99, 100, 90, 94, 135, 108, 107, 111, 119, 104, 127, 109, 117, 105,
    125)
wilcox.test(muestra,mu=107,exact=F,correct = T, alternative =  "two.sided",
    conf.int = 0.95)
```

Listing 3.4: ”Prueba del Wilcoxon en R ”

Salida

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  muestra
## V = 64.5, p-value = 0.4702
## alternative hypothesis: true location is not equal to 107
## 95 percent confidence interval:
##  102 118
## sample estimates:
## (pseudo)median
##          109.5
```



4. Pruebas para dos o más muestras independientes

En esta sección se detallarán técnicas no paramétricas métodos no paramétricos basados en dos o más muestras independientes.

4.1 Introducción

Las muestras independientes son muestras en las cuales la medición de una variable en cierta población no ejerce efecto sobre la medición de esa variable en otra población.

A continuación, se presentarán procedimientos para evaluar diferencias entre dos o más grupos a partir de muestras independientes. Para 2 muestras, se presentará la prueba de la mediana y para 3 0 más pruebas, la prueba de Kruskal-Wallis.

4.2 Prueba de la mediana

La prueba de la mediana es una técnica no paramétrica para probar si dos muestras independientes fueron extraídas de poblaciones con medianas iguales.

Supuestos:

i) Se tiene dos muestras aleatorias independientes.

ii) Se tiene dos muestras aleatorias independientes.

Hipótesis:

$$H_0 : M_1 = M_2 \quad VS \quad H_1 = M_1 \neq M_2$$

Estadítico de prueba:

Si las dos muestras provienen de poblaciones con la misma mediana, se puede esperar que aproximadamente la mitad de los datos en cada muestra este arriba de la mediana combinada y la otra mitad por debajo. Así, el primer paso es calcular la mediana combinada de las muestras, luego por medio de una tabla de contingencia $2x2$ se determina para cada muestra el número de observaciones que caen por encima y por debajo de la mediana combinada.

	Muestra 1	Muestra 2
$>$ mediana combinada	n_{11}	n_{12}
$\leq$ mediana combinada	n_{21}	n_{22}

Si H_0 es cierta, el estadístico de prueba es:

$$T = X^2 = \sum_{i=1}^{r=2}\sum_{j=1}^{c=2}\frac{(n_{ij}-E(n_{ij}))^2}{E(n_{ij})} = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(n_{ij}-nP_{ij})^2}{nP_{ij}} \dot{\sim} \chi_1^2$$

donde $E(n_{ij}) = nP_{ij} = \frac{(n_{i.})(n_{.j})}{n}$.Es decir, cuando n es grande X^2 tiene una distribución aproximadamente ji-cuadrada con 1 g.l. (grados de libertad).

En cuanto a n grande, X^2 se distribuye ji cuadrada si las $E(n_i) \geq 1$ y no más del 20% son menores que 5

Cuando $g.l. = 1$, se recomienda utilizar la corrección de Yates para continuidad.

$$X_c^2 = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{(|n_{ij}-E(n_{ij})|-0.5)^2}{E(n_{ij})}$$

Región de rechazo

Rechace H_0 si $X^2 > \chi_{\alpha,1}$,o bien, si el valor $P(\chi_1^2 > X^2) < \alpha$.

Ejemplo

Exercise 4.1 A una muestra aleatoria de 12 estudiantes varones de una escuela preparatoria rural y a una muestra aleatoria de 16 varones de una escuela preparatoria urbana, se les aplicí una prueba para medir su madurez mental.Los resultados se muestran en la tabla siguiente.¿Difiere la mediana para las calificaciones de la madurez mental de estudiantes varones de la preparatoria urbana y rural?

Urbana		Rural	
35	25	29	50
26	27	50	37
27	45	43	34
21	46	22	31
27	33	42	
38	26	47	
23	46	42	
25	41	32	

Example 4.1 (Solución:)

Hipótesis

$H_0 : M_U = M_R \; vs \; M_U \neq M_R$

Estadístico de prueba

Si H_0 es cierta, el estadístico de prueba es:

$$X_c^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(|n_{ij} - E(n_{ij})| - 0.5)^2}{E(n_{ij})} \dot{\sim} \chi_1^2$$

Cálculo del estadístico de prueba

Tabla de frecuencias observadas:

	$Urbana$	$Rural$	$Total$
> 33.5	6	8	14
≥ 33.5	10	4	14
$Total$	16	12	28

Tabla de frecuencias esperadas:

	$Urbana$	$Rural$
> 33.5	$E(n_{11}) = 26\frac{14}{28} = 8$	$E(n_{12}) = 12\frac{14}{28} = 6$
≥ 33.5	$E(n_{21}) = 26\frac{14}{28} = 8$	$E(n_{22}) = 12\frac{14}{28} = 6$

Así, $X_c^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} \frac{(|n_{ij} - E(n_{ij})| - 0.5)^2}{E(n_{ij})} \dot{\sim} \chi_1^2 = 1.3125$.

Región de rechazo

Como $P(\chi_1^2 > 3.841) \approx 0.05$, entonces la región de rechazo es:

$$RR : X_c^2 > \chi_{1,0.05}^2 = 3.841$$

Cálculo del Valor P

$valor\ P = P(\chi^2_1 > 1.3125) = 0.251.$

Decisión estadística con $\alpha = 0.05$

Como $1.3125 < \chi^2_{1,0.05} = 3.841$ no se rechaza H_0.Con nivel de significación 5%, se concluye que las muestras fueron extraídas de poblaciones con medianas no significativamente diferentes.

Por lo tanto, no difiere significativamente $(P > 0.05)$ la mediana de las calificaciones de madurez mental entre estudiantes de preparatoria urbana y rural $(X^2_c = 1.3125, P = 0.2519, g.l. = 1)$.

Código

```r
#Lectura de datos
datos <- data.frame(
  Preparatoria= c(rep("Urbana", 16), rep("Rural", 12)),
  Calificacion = c
    (35,25,26,27,27,45,21,46,27,33,38,26,23,46,25,41,29,50,50,37,43,34,22,31,42,47,42,32)
    )

#Aplicación de la prueba
library(RVAideMemoire)
mood.medtest(datos$Calificacion,datos$Preparatoria,exact = TRUE)
```

Listing 4.1: "Prueba de la mediana en R "

Salida

```
         Mood's median test

data:  datos$Calificacion by structure(c(2L, 2L, 2L, 2L,
 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, 2L, datos$Calificacion
 by 2L, 2L, 2L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L, 1L,
 1L, 1L), levels = c("Rural", datos$Calificacion by "Urb
ana"), class = "factor")
p-value = 0.2519
```

4.3 Prueba de Kruskal-Wallis

También conocido como análisis de varianza en un sentido por rangos de Kruskal-Wallis y puede verse como una generalización del método propuesto por Wilcoxon . Es una alternativa a la prueba F del ANOVA (Análisis de Varianza) de una vía paramétrico cuando el supuesto de normalidad no se cumple o cuando los datos son de escala ordinal.

Supuestos:

i) K muestras aleatorias independientes.

ii) La escala de medición es al menos ordinal.

Los datos se arreglan de la manera siguiente:

Grupos(Tratamientos)			
1	2	...	K
Y_{11}	Y_{21}	...	Y_{k1}
Y_{11}	Y_{21}	...	Y_{k1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Y_{1n_1}	Y_{2n_2}	...	Y_{kn_k}

donde $Y_i j$ es el dato para la j-ésima observación en el i-ésimo grupo y n_i es el número de observación en el i- ésimo grupo.

Hipótesis:

Se prueba la H_0 de que las k muestras provienen de poblaciones idénticas con la misma mediana.

$$H_0 : \theta_1 = \theta_2 = ... = \theta_k \quad vs \quad H_1 : \theta_i \neq \theta_j \; para \; al \; menos \; un \; par \; \; (i,j) \; con \; i \neq j$$

donde θ_i representa la mediana de la población para el i-ésimo grupo(tratamiento).

Las hipótesis también pueden enunciarse como

H_0 : Las k distribuciones poblacionales son idénticas

$$vs$$

H_1 : Al menos una de las poblaciones tiende a dar observaciones más grandes que las otras. O bien, al menos dos de las distribuciones poblacionales difieren en localización.

Estadítico de prueba:

Las $n_1, n_2, ..., n_k$ observaciones de los k grupos se combinan en una sola serie de tamaño n y se arreglan en orden de magnitud desde el más pequeño hasta el más grande. Entonces las observaciones se reemplazan por rangos desde 1, que es el asignado a la observación menor, hasta N, que se asigna a la observación mayor. Cuando dos o más observaciones tienen el mismo valor (empate), a cada observación se le asigna la media de los rangos en los cuales empató.

Posteriormente, se debe encontrar la suma de rangos en cada muestra (columna) . A partir de estas sumatorias de rangos, podemos calcular los rangos promedio para cada muestra. Ahora, si las muestras provienen de la misma o de idénticas poblaciones, los rangos promedio deberian ser (aproximadamente) los mismos. La prueba de Kruskal-Wallis evalúa la diferencia entre los rangos promedios para determinar si son lo suficientemente dispares, de tal suerte que no sea probable que las muestras hayan sido extraídas de la misma población.

Sea R_j la suma de los rangos correspondientes a las n_j observaciones de la muestra j-ésima, y sea $\bar{R_j} = R_j/n_j$ el valor medio de los rangos en esa muestra. Bajo H_0 se espera que $\bar{R_j}$ sea próximo a $\frac{N+1}{2}$, para todo j.

El estadístico Kruskal-Wallis(KW) es:

$$KW = \frac{\sum_{j=1}^{k} n_j(\bar{R_j} - \frac{N+1}{2})^2}{\frac{N(N+1)}{12}} = \frac{12}{N(N+1)}\sum_{j=1}^{k} n_j(\bar{R_j} - \bar{R})^2 = \frac{12}{N(N+1)}\sum_{j=1}^{k} n_j\bar{R_j}^2 - 3(N+1)$$

donde

k =número de muestras o grupos.
n_j =número de casos en la j-ésima muestra.
N =número de casos en la muestra combinada(la suma del número total de observaciones).
R_j =sumatoria de los rangos en el j-ésimo grupo.
$\bar{R} = \frac{N+1}{2}$ =promedio de los rangos en la muestra combinada.

Bajo H_0, los rangos se asignan aleatoriamente a los tratamientos. As í, los n_i rangos del tratamiento son tratados como una asignación aleatoria. El n úmero de posibles asignaciones a los tratamientos es

$$\frac{n!}{\prod_{i=1}^{k} n_i!}$$

Enumerando las posibles asignaciones que generan un valor particular del estadístico

$$N(w) = KW_n = w$$

La distribución de probabilidad es descrita por:

$$f_{KW_n}(w) = \frac{N(w)}{\frac{n!}{\prod_{i=1}^{k} n_i!}}$$

La cual está computada en Tablas y y es utilizada si el tamaño de la muestra es pequeño.

Región de rechazo

Rechace H_0 si $KW > k_{\alpha;n1,\dots,nk}$ donde la constante $KW > k_{\alpha;n1,\dots,nk}$ satisface $KW > k_{\alpha;n1,\dots,nk} = \alpha$.

Aproximación para muestras grandes

Si los n_i son grandes $(n_i \geq 5)$ y bajo H_0 verdadera, la distribución del estadístico de prueba KW es aproximádamente ji-cuadrada con $k-1$ g.l., es decir, $KW \dot{\sim} \chi^2_{k-1}$. Por lo tanto, rechace H_0 si $H > \chi^2_{k-1}$

En caso de que la prueba determine diferencias estadísticas significativas entre los tratamientos con respecto a la variable de estudio o de respuesta, deben efectuarse comparaciones múltiples para determinar qué pares de tratamientos difieren.

Observaciones empatadas

Cuando ocurren empates entre dos o más puntuaciones (al margen del grupo), a cada puntuación se le asigna el promedio de los rangos en los que ocurrieron los empates.

La varianza de la distribución muestral de KW es influida por los empates. Para realizar la corrección del efecto de los empates, el KW debe calcularse como:

$$KW = \frac{\frac{12}{N(N+1)} \sum_{j=1}^{k} n_j \bar{R_j}^2 - 3(N+1)}{1 - \frac{\sum_{i=1}^{g}(t_i^3 - t_i)}{N^3 - N}}$$

donde

k =número de muestras o grupos.
n_j =número de casos en la j-ésima muestra.
N =número de casos en la muestra combinada(la suma del número total de observaciones).
R_j =sumatoria de los rangos en el j-ésimo grupo.
$\bar{R} = \frac{N+1}{2}$ =promedio de los rangos en la muestra combinada.
g =número de grupos de rangos empatados.
t_i =número de rangos empatados en el i-ésimo grupo.

Potencia de la prueba

Comparada con la prueba paramétrica más poderosa, la prueba F, y en condiciones donde se cubren los supuestos asociados con el modelo estadístico del aálisis de varianza paramétrico, la prueba de Kruskal-Wallis tiene una eficacia asintótica de $3/\pi = 95.5$

Ejemplo

Exercise 4.2 Un gerente de mercadotecnia de una cadena de una línea de productos de cómputo, le interesa saber si hay diferencias en las ventas de sus productos en tres ciudades. Elige al azar 15 tiendas similares (5 por ciudad) entre las que integran la cadena. Las variables bajo control directo de la compañía, como precio y publicidad, se mantuvieron al mismo nivel en los 30 días del experimento y se registraron las ventas (en miles) para dicho periodo. Utilice un nivel de significancia de 0.05.

Ventas (en miles)		
Ciudad 1	**Ciudad 2**	**Ciudad 3**
10	16	15
14	18	12
18	22	8
15	18	10
12	15	13

Example 4.2 (Solución:)

Hipótesis

$H_0 : \theta_1 = \theta_2 = \theta_3 \quad vs \quad H_1 : \theta_i \neq \theta_j \; para\ al\ menos\ un\ par\ (i,j)\ con\ i \neq j$ donde θ_i representa la mediana de las ventas para la i-ésima ciudad.

Estadístico de prueba

Si H_0 es cierta, el estadístico de prueba es:

$$KW = \frac{\frac{12}{N(N+1)} \sum_{j=1}^{k} n_j \bar{R}_j^{\,2} - 3(N+1)}{1 - \frac{\sum_{i=1}^{g}(t_i^3 - t_i)}{N^3 - N}}$$

Cálculo del estadístico de prueba

Rangos asignados, totales y promedios para los datos del ejemplo.

$Ciudad\ 1$	**Rango**	$Ciudad\ 2$	**Rango**	$Ciudad\ 3$	**Rango**
10	2.5	16	11	15	9
14	7	18	13	12	4.5
18	13	22	15	8	1
15	9	18	13	10	2.5
12	4.5	15	9	13	6
$Suma\ de\ rangos$	$R_{1.} = 36$		$R_{2.} = 61$		$R_{3.} = 23$
$Rango\ promedio$	$\bar{R}_{1.} = 7.2$		$\bar{R}_{2.} = 12.2$		$\bar{R}_{3.} = 4.6$

Así, $KW = \dfrac{\frac{12}{N(N+1)}\sum_{j=1}^{k} n_j \bar{R_j}^2 - 3(N+1)}{1-\frac{\sum_{i=1}^{g}(t_i^3-t_i)}{N^3-N}} = 7.59$

Región de rechazo

Como $P(\chi_2^2 > 5.991) \approx 0.05$, entonces la región de rechazo es:

$$RR : KW > \chi^2_{2,0.05} = 5.991$$

Cálculo del Valor P

$valor\ P = P(\chi_2^2 > 1.3125) = 0.0224$.

Decisión estadística con $\alpha = 0.05$

Como $7.59 > \chi^2_{2,0.05} = 5.991$, se rechaza H_0.Con un nivel de significación del 5%, se concluye que las muestras fueron extraídas de poblaciones con medianas significativamente diferentes.

Difiere significativamente las ventas de los productos de cómputo en las ciudades (KW=7.5956, P=0.0224<0.05, g.l.=2).

Como se determinaron diferencias estadísticas significativas entre las ciudades, deben efectuarse comparaciones múltiples para determinar qué pares de ciudades difieren entre sí.

Código

```
datos <- data.frame(
  ciudad= c(rep("Ciudad 1", 5), rep("Ciudad 2", 5), rep("Ciudad 3", 5)),
  Ventas = c(10,14,18,15,12,16,18,22,18,15,15,12,8,10,13))

head(datos,6)
```

Listing 4.2: Lectura de datos

Salida

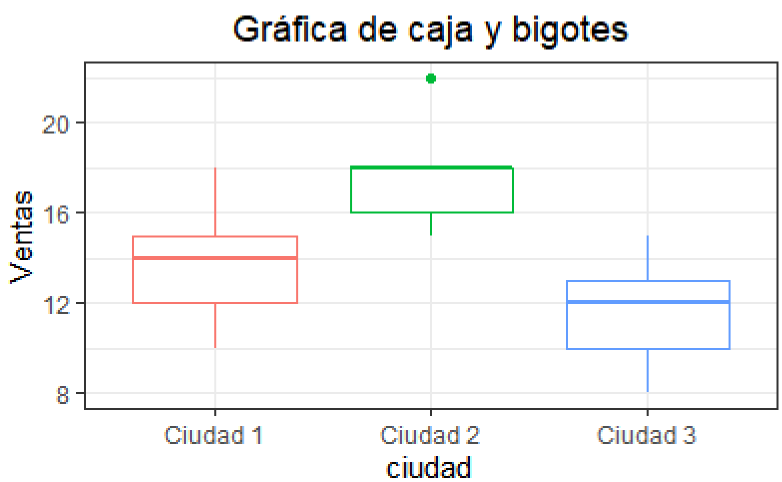
```
##     ciudad Ventas
## 1 Ciudad 1     10
## 2 Ciudad 1     14
## 3 Ciudad 1     18
## 4 Ciudad 1     15
## 5 Ciudad 1     12
## 6 Ciudad 2     16
```

```
library(ggplot2)
```

```
ggplot(data = datos, mapping = aes(x =ciudad, y = Ventas, colour = ciudad)) +
  geom_boxplot() +
  theme_bw() +
  theme(legend.position = "none",plot.title = element_text(hjust = 0.5)) +
  labs(title="Gráfica de caja y bigotes")
```

Listing 4.3: Código del boxplot

Salida



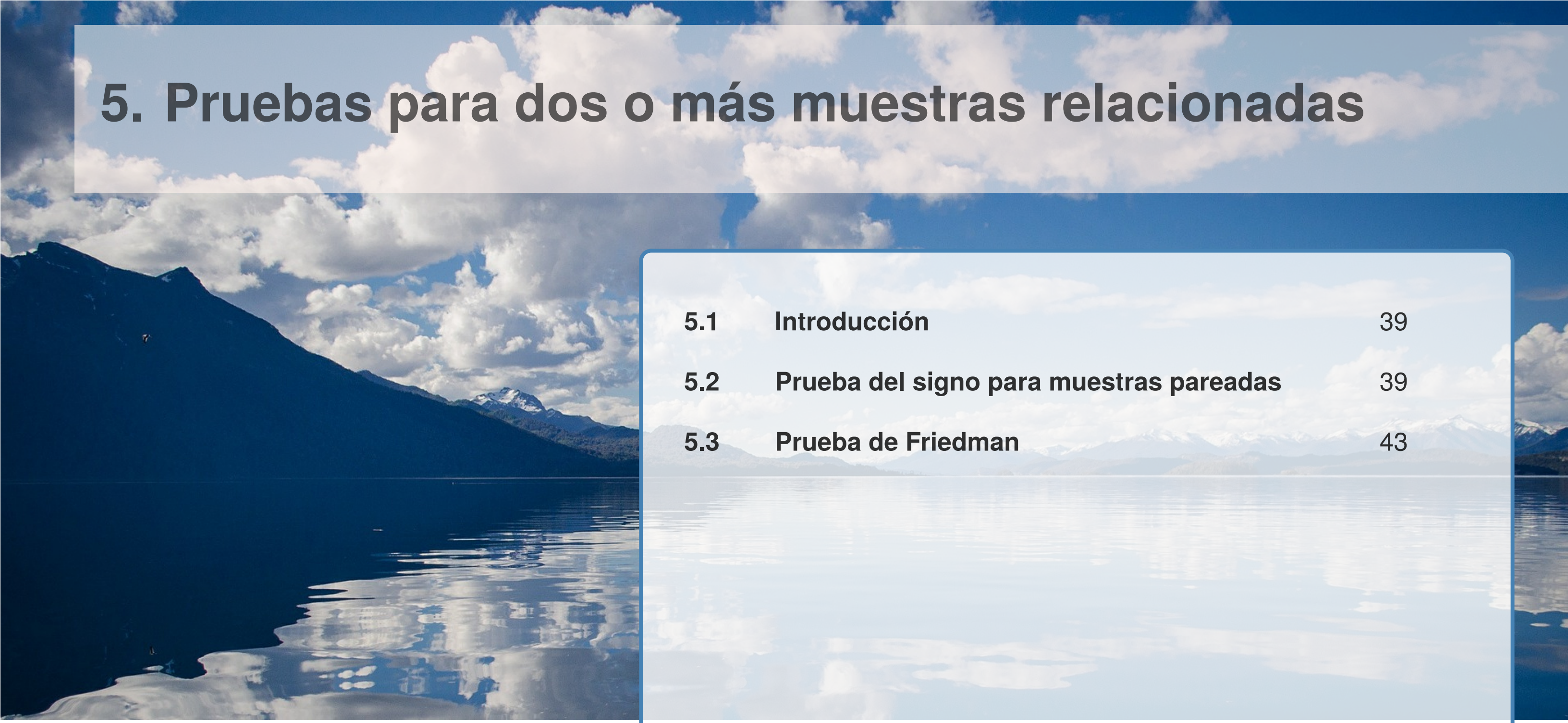
El box-plot sugiere que las ventas podrían diferir, al menos, entre la ciudad 2 y 3.

```
#Medianas de los grupos
aggregate(Ventas ~ ciudad, data = datos, FUN = median)

#Resultado de la prueba
kruskal.test(Ventas ~ ciudad, data = datos)
```

Listing 4.4: Test de Kruskal-Wallis

Salida

```
##     ciudad Ventas
## 1 Ciudad 1     14
## 2 Ciudad 2     18
## 3 Ciudad 3     12
##
##  Kruskal-Wallis rank sum test
##
## data:  Ventas by ciudad
## Kruskal-Wallis chi-squared = 7.5956, df = 2, p-value = 0.02242
```



5. Pruebas para dos o más muestras relacionadas

En esta sección se detallarán técnicas no paramétricas para la realización de inferencias en problemas que incluyan dos o más muestras relacionadas.

5.1 Introducción

Las muestras relacionadas son muestras en las cuales la medición de una variable en la población influye en la medición de la variable en otra población. Cuando los mismos sujetos son expuestos a dos o más tratamientos, las comparaciones entre tratamientos no son independientes debido a que en todas las condiciones se usan los mismos sujetos.

A continuación, se presentarán procedimientos para evaluar diferencias entre dos o más grupos a partir de muestras relacionadas. Para 2 muestras, se presentará la prueba del signo y para 3 0 más muestras, la prueba de Friedman.

5.2 Prueba del signo para muestras pareadas

Supuestos:

- Sea $(X_1, Y_1), (X_2, Y_2), .., (X_n, Y_n)$ una muestra aleatoria de n pares de mediciones (dos muestras relacionadas) de una población con mediana (M_D) desconocida.
- La variable de interés es $D_i = X_i - Y_i$ para $i = 1, ..., n$.
- La variable de respuesta es medida en escala al menos ordinal.

Hipótesis:

a) $H_0 : M_D = 0$, $H_1 : M_D \neq 0$ (alternativa de dos colas)

b) $H_0 : M_D \leq M_0$, $H_1 : M_D > 0$(alternativa de cola superior)

c) $H_0 : M_D \geq M_0$, $H_1 : M_D < 0$ (alternativa de cola inferior)

donde M_0 denota un valor de la mediana.

Estadístico de Prueba

Si $H_0 : M_D = 0$ es cierta, cualquier diferencia tiene la misma probabilidad de ser positiva que de ser negativa, es decir, 0.5

Consideremos las hipótesis planteadas en $b)H_0 : M_D \leq 0, H1 : M_D > 0$, note que un número suficientemente pequeño de signos negativos es evidencia para el rechazo de $H_0 : M_D \leq 0$. Por lo que la estadística de prueba es $K = número\ de\ signos\ menos$, así, su distribución muestral es $B(n, p = 0.5)$, es decir, $K\ B(n, p = 0.5)$.

Procediendo en forma análoga para las otras dos hipótesis alternativas planteadas, la estadística de prueba para cada alternativa es:

a) $K = min\{K_-, K_+\}$

b) $K = número\ de\ signos\ negativos\ \ (K_-)$

c) $K = número\ de\ signos\ positivos\ \ (K_+)$

En resumen $K =$ número de signos de interés bajo la hipótesis alternativa dada.

Diferencias cero (empates)

El procedimiento para manejar los ceros (empates) es eliminarlos del análisis, lo cual reduce el tamaño de la muestra n, es decir, la prueba del signo se aplica al resto de los datos.

Región de rechazo

a) Para $H_1 : M \neq M_0$, la región de rechazo (de dos colas) es $K \geq k_{\alpha/2}$ o $K \leq k'_{\alpha/2}$, donde $k_{\alpha/2}$ es el entero más pequeño que satisface $P(K \geq k_{\alpha/2}|H_0) = \sum_{i=k_{\alpha/2}}^{N} \binom{N}{i}(0.5)^N \leq \alpha/2$ y k'_α es el entero más grande que satisface $P(K \leq k'_\alpha|H_0) = \sum_{i=0}^{k'_{\alpha/2}} \binom{N}{i}(0.5)^N \leq \alpha/2$

b) Para $H_1 : M > M_0$, la región de rechazo (de cola superior) es $K \geq k_\alpha$, donde k_α es el entero más pequeño que satisface $P(K \geq k_\alpha|H_0) = \sum_{i=k_\alpha}^{N} \binom{N}{i}(0.5)^N \leq \alpha$.

c) Para $H_1 : M < M_0$, la región de rechazo (de cola inferior) es $K \leq k'_\alpha$, donde k'_α es el entero más grande que satisface $P(K \leq k'_\alpha|H_0) = \sum_{i=0}^{k'_\alpha} \binom{N}{i}(0.5)^N \leq \alpha$.

Cálculo del valor P

Como la distribución binomial es simétrica cuando $\theta = 0.5$,para las hipótesis alternativas planteadas en $b)$ y $c)$:

$$Valor\ P = P(K \geq k|n, p = 0.5)\ \ k\ es\ un\ valor\ del\ estad\acute{\imath}stico\ de\ prueba\ K$$

En el caso de la hipótesis de dos colas $a)$:

$$Valor\ P = 2P(K \geq k|n, p = 0.5)$$

Aproximación Normal

Sabemos que la aproximación normal a la binomial es buena cuando $\theta = 0.5$, entonces la aproximación normal a la binomial se puede utilizar. Una regla empíirica para aproximar la distribución binomial mediante la distribución normal es $np > 5$ y $nq > 5$.

Considerando que se aproxima una distribución discreta mediante una distribución continua, se utiliza un factor de corrección de continuidad de 0.5. Por lo tanto, el estadístico de prueba es:

$$Z = \frac{(K + 0.5) - (0.5)n}{\sqrt{0.25n}} = \frac{(K + 0.5) - (0.5)n}{(0.05)\sqrt{n}} \dot{\sim} N(0, 1)$$

El cálculo del $Valor\ p$ se realiza como en las pruebas estadísticas basadas en la distribución normal estándar.

Ejemplo

Exercise 5.1 Para determinar la eficacia de un nuevo sistema de control de tránsito, se observó el número de accidentes que ocurrieron en 12 intersecciones peligrosas durante cuatro semanas antes y cuatro semanas después de la instalación del nuevo sistema, y se obtuvieron los siguientes datos:

3 y 1	5 y 2	2 y 0	3 y 2	3 y 2	3 y 0
0 y 2	4 y 3	1 y 3	6 y 4	4 y 1	1 y 0

¿El nuevo sistema de control de tránsito es eficaz para reducir el número de accidentes en intersecciones peligrosas? Utilice 5% de nivel de significación.

Example 5.1 (Solución:) Sea $D_i = X_i - Y_i$ la variable de las diferencias entre el número de accidentes durante cuatro semanas antes (x_i) y los que ocurrieron durante cuatro semanas después (yi) de la instalación del nuevo sistema, $i = 1, 2, , 12$.

Hipótesis

$H_0 : M_D = M_{antes} - M_{despues} \leq 0 \;\; vs \;\; H_1 : M > 0$

Estadístico de prueba

K = número de signos menos (k_1)

Cálculo de estadístico de prueba

Para determinar el valor de la estadística de prueba es necesario obtener los signos de las diferencias $D_i = X_i - Y_i$, las cuales se encuentran en la tabla 1

$Intersección$	$NAA(X_i)$	$NAD(Y_i)$	$D_i = X_i - Y_i$
1	3	1	+
2	5	2	+
3	2	0	+
4	3	2	+
5	3	2	+
6	3	0	+
7	0	2	−
8	4	3	+
9	1	3	−
10	6	4	+
11	4	1	+
12	1	0	+

Así, $K = k_= 2$

Cálculo del Valor P

$P = P(K \leq k|n, p = 0.5) = P(K \leq 2|n = 12, p = 0.5) = 0.0193$

Decisión estadística con $\alpha = 0.05$

Como $P = 0.0193 < 0.05$ se rechaza H_0 y se decide por H_1, es decir, la mediana del número de accidentes antes de la instalación del nuevo sistema de control de tránsito resultó significativamente mayor que después de la instalación.

El nuevo sistema de control de tránsito es eficaz para reducir el número de accidentes en intersecciones peligrosas $(K = 2, P = 0.0193)$.

Código

```
#Lectura de datos
antes<-c(3,5,2,3,3,3,0,4,1,6,4,1)
despues<-c(1,2,0,2,2,0,2,3,3,4,1,0)

#Aplicación de la muestra
library(BSDA)
```

```
SIGN.test(antes,despues,alternative = "greater",conf.level = 0.95)
```

Listing 5.1: Prueba del signo para muestras pareadas en R

Salida

```
        Dependent-samples Sign-Test

data:  antes and despues
S = 10, p-value = 0.01929
alternative hypothesis: true median difference is greater than 0
95 percent confidence interval:
   1 Inf
sample estimates:
median of x-y
          1.5

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.9270      1    Inf
Interpolated CI       0.9500      1    Inf
Upper Achieved CI     0.9807      1    Inf
```

5.3 Prueba de Friedman

La prueba de Friedman es una ampliación de Prueba de Wilcoxon de los rangos con signo para muestras pareadas. Determina la probabilidad de que diferentes columnas de rangos (muestras) provengan de la misma población, es decir, que las k variables tengan la misma mediana.

Supuestos:

i) Los resultados dentro de un bloque no influyen los resultados dentro de otro bloque.

ii) Los tratamientos se asignan al azar a unidades experimentales dentro de los bloques.

iii) Las mediciones se pueden clasificar (asignarles rangos) dentro de los bloques.

Hipótesis:

Las hipótesis pueden ser enunciadas en algunas de las siguientes formas equivalentes:

a) $H_0 : \theta_1 = \theta_2 = ... = \theta_k \quad vs \quad H_1 : \theta_i \neq \theta_j \ para\ al\ menos\ dos\ condiciones\ i\ y\ \ con\ i\ \neq j$ donde θ_i representa la mediana poblacional en la i-ésimo condición(tratamiento).

b) H_0 : Las k distribuciones poblacionales son idénticas

$$vs$$

H_1 : Al menos una de las poblaciones tiende a dar observaciones más grandes que las otras. O bien, al menos dos de las distribuciones poblacionales difieren en localización.

c) H_0 : Los tratamientos tienen efecto idénticos

$$vs$$

H_1 : Al menos uno de los tratamientos tiende a producir valores observados mas grandes que al menos otro tratamiento.

Estadístico de Prueba

Los datos se arreglan de la manera siguiente:

1	2	...	K
Y_{11}	Y_{21}	...	Y_{k1}
Y_{11}	Y_{21}	...	Y_{k1}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
Y_{1n_1}	Y_{2n_2}	...	Y_{kn_k}

Los renglones representan los sujetos o conjuntos de sujetos igualados, y las columnas, las distintas condiciones.

Se calcula el rango de cada medición en relación con su posición dentro de su mismo bloque, es decir, dentro de cada bloque se asignan los rangos a las medidas con base a sus magnitudes relativas.Esto es, estudiando k condiciones, los rangos en cualquier renglón varían de 1 a k.

Bajo H_0 cierta, la distribución de los rangps en cada columna sería cuestíon de oportunidad y se esperarı que los rangos $1, 2, ..., k$ aparecieran en cada columna con, aproximádamente, igual frecuencia. Si los datos fueran aleatorios, la suma de los rangos en cada columna sería $\frac{N(k+1)}{2}$; de lo contrario, el promedio de los rangos en las distintas columnas sería, aproximadamente, el mismo.

El estadístico de prueba es:

$$F_r = \frac{\sum_{j=1}^{k}(R_j - \frac{N(K+1)}{2})^2}{\frac{N(N+1)}{12}} = \frac{12}{N1(k+1)}\sum_{j=1}^{k} R_j^2 - 3N(k+1)$$

donde

N =número de renglones (sujetos).

k =número de columnas (variables o condiciones).

R_j = suma de los rangos en la j-ésima columna (variable).

$\sum_{j=1}^{k}$ =la sumatoria de los cuadrados de los rangos de todas las condiciones.

El cálculo de la distribución exacta se realiza teniendo en cuenta que cada asignación de rangos es una de las $(k!)^N$ equiprobables.

Observaciones empatadas

Cuando existan empates entre los rangos en cualquier bloque (renglón), el estadístico F_r debe ser corregido para modificar la distribución muestral:

$$F_r = \frac{12\sum_{j=1}^{k} R_j^2 - 3N^2k(k+1)^2}{Nk(k+1) - (\frac{1}{k-1})\sum_{i=1}^{N}\{\sum_{j=1}^{g_i} t_{ij}^3\}}$$

donde

N =número de renglones (sujetos).
k =número de columnas (variables o condiciones).
R_j = suma de los rangos en la j-ésima columna (variable).
g_i=Número de grupos de rangos empatados en el i-ésimo grupo.
t_{ij}=Tama n del j-ésimo conjunto de rangos empatados en el iésimo grupo.

Región de rechazo

Rechace H_0 si $F_r = \chi_r^2 > \chi_{r,\alpha,a,b}^2$ donde $\chi_{r,\alpha,a,b}^2$ es tal que $P(\chi_{r,\alpha,a,b}^2) = \alpha$ y se obtiene de la tabla de valores críticos de la distribución χ_r^2 de Friedman.

Aproximación para muestras grandes

La distribución del estadístico F_r de Friedman cuando H_0 es verdadera, se puede aproximar mediante una distribución ji-cuadrada con $k-1$ grados de libertad mientras n sea grande. La evidencia empírica indica que la aproximación es adecuada si n o k es mayor que 5.

Comparaciones múltiples

Cuando el valor obtenido de F_r, es significativo, este resultado refleja que al menos una de las condiciones difiere con respecto a otra de las condiciones. Se realizan comparaciones múltiples para saber cuál grupo es el diferente y cuántos de los grupos Análisis de Friedman difieren entre sí.

Pueden consultarse comparaciones múltiples en: a) Siegel y Castellan (2001) para muestras grandes con o sin tratamiento control, b) Conover (1999), c) Zar (2010) con o sin tratamiento control, y en d) Hollander et al. (2014) para muestras pequeñ ñas y grandes con o sin tratamiento control.

Potencia de la prueba

La potencia-eficacia de la prueba de Friedman para datos normalmente distribuidos cuando se comparan con su contraparte normal (la prueba F), es $\frac{2}{\pi} = 0.64$ cuando $k = 2$, y es mayor conforme incrementa $k(k = 5, eficacia = 0.80; k = 10, eficacia = 0.87; k = 20, eficacia = 0.91)$. Cuando se comparan muestras con distribución uniforme o exponencial, la eficacia es mayor.

Ejemplo

Exercise 5.2 Una empresa opera 24 horas al día, cinco días a la semana. Los trabajadores cambian de turno cada semana. La gerencia está interesada en saber si hay alguna diferencia en el número de unidades producidas cuando los empleados laboran en diversos turnos. Se seleccioná cinco obreros y se registró su producción en cada turno. Cada empleado es un bloque porque se espera que las mediciones del mismo empleado sean más parecidas entre sí que las mediciones de varios empleados. Al nivel de significancia 0.05, ¿puede concluirse diferencias en la producción por turno?

Unidades producidas			
Empleado	**Mañana**	**tarde**	**noche**
1	31	25	35
2	33	26	33
3	28	24	30
4	30	29	28
5	28	26	27

Example 5.2 (Solución:)

Hipótesis

$H_0 : \theta_1 = \theta_2 = \theta_3 \quad vs \quad H_1 : \theta_i \neq \theta_j \ para \ al \ menos \ un \ par \ (i, j) \ con \ i \neq j$ donde θ_i representa la mediana de la porducción para el i-ésimo turno.

Cálculo del estadístico de prueba

De la salida de R, se obtiene que $F_rc = 5.1579$.

Región de rechazo

Como $P(\chi^2_2 > 5.991) \approx 0.05$, entonces la región de rechazo es:

$$RR : F_r > \chi^2_{2,0.05} = 5.991$$

Cálculo del Valor P

$valor\ P = P(\chi^2_2 > F_rc = 5.1579) = 0.0759.$

Decisión estadística con $\alpha = 0.05$

Como $5.1579 < \chi^2_{2,0.05} = 5.991$, no se rechaza H_0.

Con un nivel de significación del 5%, se concluye que no difiere significativamente la producción por turno ($Fr = 5.1579, P = 0.0759 > 0.05$).

Código

```
Sujeto <- factor( rep( 1:5, each = 3 ) )
Turno <- factor( rep( c( "mañana", "tarde", "noche" ), 5 ) )
Productos<- c( 31, 25, 35,
              33, 26, 33,
              28, 24, 30,
              30, 29, 28,
              28, 26, 27)

datos <- data.frame( Sujeto, Turno, Productos )
head(datos,6)
```

Listing 5.2: Lectura de datos

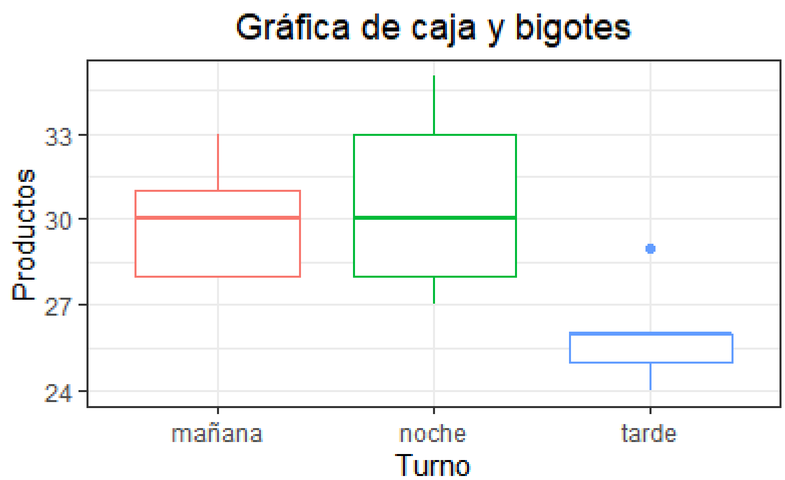
Salida

```
##   Sujeto  Turno Productos
## 1      1 mañana        31
## 2      1  tarde        25
## 3      1  noche        35
## 4      2 mañana        33
## 5      2  tarde        26
## 6      2  noche        33
```

```
library(ggplot2)
ggplot(data = datos, mapping = aes(x = Turno, y = Productos, colour =Turno)) +
  geom_boxplot() +
  theme_bw() +
  theme(legend.position = "none",plot.title = element_text(hjust = 0.5)) +
  labs(title="Gráfica de caja y bigotes")
```

Listing 5.3: Código para graficar

Salida



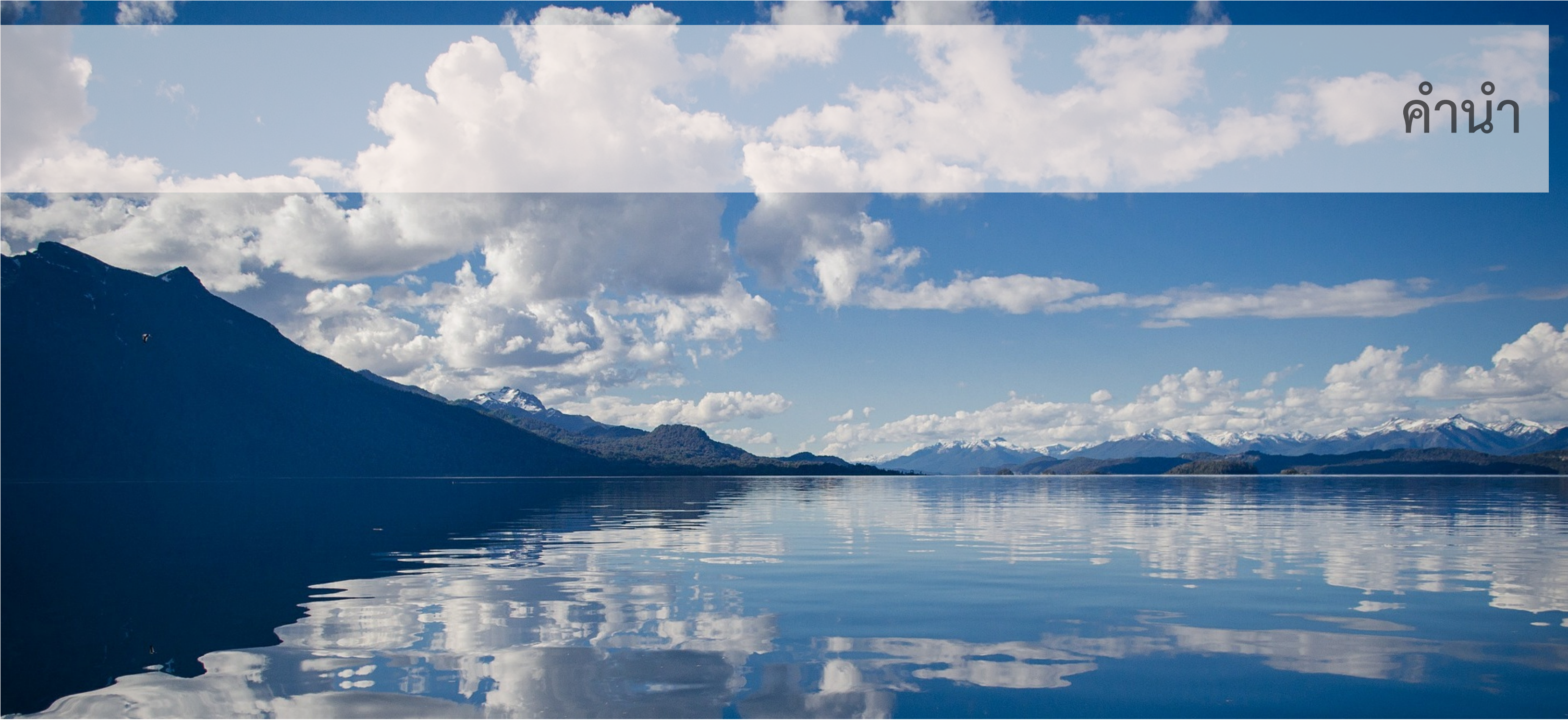
El box-plot parece indicar que la productividad no difiere por horario de trabajo.

```
#Medianas no los grupo
by(data = datos$Productos, INDICES = datos$Turno, FUN = median)

#Resultado de la prueba
friedman.test(Productos ,Turno, Sujeto)
```

Listing 5.4: Test de Friedman

Salida

```
## datos$Turno: mañana
## [1] 30
## ------------------------------------------------------------
## datos$Turno: noche
## [1] 30
## ------------------------------------------------------------
## datos$Turno: tarde
## [1] 26
##
##  Friedman rank sum test
##
## data:  Productos, Turno and Sujeto
## Friedman chi-squared = 5.1579, df = 2, p-value = 0.07585
```



Bibliografía

1. *Gibbons, J. D., Chakraborti, S. (2014). Nonparametric statistical inference. CRC press.*
2. *Siegel, S. y Castellan, J.N. (2007). Estadística no paramétrica: aplicada a las ciencias de la conducta. (4a. ed.). México: Trillas*
3. *Conover, W. (1999). Practical Nonparametric Statistics. (3a. ed.). New York, NY: Wiley. (Clásico)*
4. *Daniel, W.W. (1990). Applied Nonparametric Statistics. (2a. ed.). Pacific Grove, CA: Duxbury. (Clásico)*