

No paramétrica

Carlos Javier Rincón

versión:2019

Contents

Introducción	3
Pruebas de bondad y ajuste - Kolmogorov Smirnov	5
Aplicación	6
Pruebas de normalidad	12
Gráficos para evaluar normalidad	14
Prueba del signo	19
Aplicación	20
Prueba de rangos de signos de Wilcoxon	24
Aplicación	25
Prueba U de Mann Whitney - Wilcoxon	29
Aplicación	30
Prueba de rachas de Wald - Wolfowitz	32
Aplicación	33
Prueba de McNemar	36
Casos y controles pareados 1 a 1	37
Aplicación	38
Prueba de la mediana	40
Aplicación	40
Prueba de Kruskal-Wallis	42
Aplicación	43
Prueba de Friedman	45
Aplicación	46
Prueba Q Cochran	48
Aplicación	48
Medidas de Asociación	50
Coefficiente tau de Kendall	51
Aplicación	53
Coefficiente de correlación de rangos de Spearman	56
Aplicación	57
Prueba de independencia - Distribución X^2	63
Aplicación	64

Prueba de Homogeneidad - Distribución X^2	67
Aplicación	67
Prueba de bondad y ajuste - Distribución X^2	70
Aplicación	71
Bootstrap (en construcción)	72
Suavizamientos (en construcción)	77
Bibliografía	79



Notas de clase *Estadística no paramétrica*. Se incluye una revisión de las pruebas no paramétricas junto con su aplicación utilizando el programa R. Se requiere un conocimiento básico por parte del lector en concepto básicos de inferencia y en el uso del programa R y RStudio.

Introducción

La estadística clásica se puede clasificar en descriptiva e inferencial (inductiva). La estadística descriptiva busca conocer el comportamiento general de un grupo de sujetos a partir de la medición de sus características y la posterior aplicación de medidas de resumen (medidas de tendencia central o dispersión, frecuencias absolutas o relativas) y métodos de visualización (tablas y gráficos). En este caso no se realizan supuestos sobre la distribución de una población de la cual provienen los datos observados, pues no se busca generalizar los resultados que se obtienen más allá de los sujetos observados.

La estadística inferencial, parte de la selección aleatoria de una muestra que proviene de una población de referencia, con el fin de realizar una estimación (intervalos de confianza) o evaluar una afirmación (pruebas de hipótesis) sobre un **parámetro**; para que los resultados pueden representar a la población, generalmente se requiere realizar supuesto sobre la distribución de los valores de la variable en la población de estudio. En el campo de la estadística inferencial, cuando los supuestos no se cumplen se pueden utilizar los métodos **no paramétricos**. En estos métodos, se encuentran las pruebas *no paramétricas* que son pruebas cuya hipótesis no corresponde a una afirmación sobre un parámetro, y las pruebas de *libre distribución* donde su aplicación no depende de la distribución de la variable de interés en la población de estudio. En este contexto, las pruebas dentro de la estadística inferencial clásica se denominan pruebas **paramétricas**.

Cuando la hipótesis a probar corresponde a una afirmación que no está relacionada con un parámetro particular, la estadística clásica no ofrece una alternativa para su evaluación, por consiguiente los métodos no paramétricos se presentan como única alternativa de elección. Cuando la hipótesis a probar corresponde a una afirmación sobre un parámetro, tenemos dos alternativas (paramétricos y no paramétricos) en cuyo caso debemos evaluar cual de los dos métodos es el recomendable. Por lo anterior, se han identificado algunas ventajas generales de los métodos no paramétricos que se listan a continuación:

1. Son más rápidos y fáciles de aplicar (cálculos aritméticos simples).
2. Con frecuencia son más fáciles de entender.
3. Son relativamente insensibles a datos atípicos.
4. Los supuestos requeridos son en general más fáciles de cumplir.
5. Se pueden aplicar en muestras pequeñas donde no se pueden verificar los supuestos de la estadística inferencial clásica.
6. Resuelven preguntas en nuevos escenarios, por ejemplo cuando se trabaja con variables medidas en escalas nominales.

Ahora bien, en favor de las pruebas paramétricas, estos métodos son los recomendables cuando se cumplen los supuestos sobre las distribuciones de las variables en la población de estudio.

En este documento se realiza la revisión de distintas pruebas no paramétricas buscando cumplir con los siguientes objetivos:

1. Probar hipótesis relacionadas con la comparación entre dos o más muestras independientes o emparejadas.
2. Evaluar la relación y correlación entre dos variables.
3. Evaluar si una muestra proviene de una población con una cierta distribución.

El listado de temas que se presentan son los siguientes:

1. Pruebas de Kolmogorov-Smirnov.
2. Pruebas de Shapiro Wills|Francia.
3. Gráficas para evaluar bondad y ajuste.
4. Prueba del signo.
5. Prueba de rangos de signos de Wilcoxon.
6. Prueba de Mann Whitney|suma de rangos de Wilcoxon.
7. Prueba de rachas Wald-Wolfowitz.
8. Prueba de McNemar.
9. Prueba de la mediana.
10. Prueba Kruskal-Wallis.

11. Prueba de Friedman.
12. Prueba Q de Cochran.
13. Coeficiente τ de Kendall.
14. Coeficiente de correlación de rangos de Spearman- ρ .
15. Aplicación de la distribución χ^2 (Prueba de independencia - Coeficiente de asociación).
16. Aplicación de la distribución χ^2 (Prueba de homogeneidad).
17. Aplicaciones de la distribución χ^2 (Prueba de bondad de ajuste).
18. Bootstrap.
19. Métodos de suavizamiento.

Para cada prueba bajo estudio, se presentan los siguientes aspectos:

1. Supuestos necesarios.
2. Planteamiento de la hipótesis a evaluar.
3. Cálculo y evaluación del estadístico de prueba (distribución exacta y asintótica) y ajustes por la presencia de empates y ceros.

Al final de cada sección, se incluye un caso o escenario donde se aplica cada prueba utilizando el programa R.

Este documento fue elaborado a partir de la revisión de los libros de: Jean Dickinson Gibbons y Subhabrata Chakraborti (Gibbons 2003b), W.J. Conover (Conover 1999) y Wayne W. Daniel (Daniel 2008).

Pruebas de bondad y ajuste - Kolmogorov Smirnov

Las pruebas de bondad y ajuste evalúan si las observaciones o datos recolectados presentan un comportamiento específico siguiendo un patrón; por ejemplo, estas prueba permiten evaluar si los datos obtenidos provienen de una población con una cierta función de distribución de probabilidades. Una prueba de bondad y ajuste que se utiliza en este contexto, es la prueba de Kolmogorov - Smirnov, la cual se basa en la comparación entre la distribución de probabilidades acumuladas de la muestra, sea $S_n(X)$ que se denomina como la distribución empírica acumulada, y la función de distribución de probabilidades acumulada de la población, sea $F_0(X)$.

- Las hipótesis a evaluar son:

Hipótesis nula $H_0 : F(X) = F_0(X)$

Hipótesis alternativas H_a :

1. $H_a : F(X) \neq F_0(X)$
2. $H_a : F(X) > F_0(X)$
3. $H_a : F(X) < F_0(X)$

- Estadística de prueba.

Sea $S_n(X)$ la distribución empírica acumulada, el estadístico de prueba es igual a:

$$D = \underbrace{\max}_X (|S_n(X) - F_0(X)|)$$

Los estadísticos de prueba para las pruebas unilaterales correspondientes a las hipótesis alternativas 2 y 3 son respectivamente:

1. $D^+ = \underbrace{\max}_X (S_n(X) - F_0(X))$
2. $D^- = \underbrace{\min}_X (S_n(X) - F_0(X))$

Las diferencias entre $S_n(X)$ y $F_0(X)$ sólo se pueden obtener para los valores de la variable X observados en la muestra. Lo anterior implica que se debe considerar que las distancias pueden ser mayores en los puntos donde no se observó X . Por lo anterior, la estadística de prueba toma la siguiente expresión:

$$D = \underbrace{\max}_X (|S_n(X_i) - F_0(X_i)|; |S_n(X_{i-1}) - F_0(X_i)|)$$

Aunque en estas notas no se presenta la función de distribución de probabilidades del estadístico de prueba, se ha encontrado que a partir de 35 observaciones se puede utilizar una aproximación por su distribución asintótica descrita en gibbons (pág 110).

La aplicación de esta prueba esta sujeta a la completa determinación de la función de distribución de probabilidades de la población en la hipótesis nula, es decir que los parámetros de esta función no se deben estimar de la muestra sino que deben ser definidos por el investigador.

Adicionalmente, con la prueba de kolmogorov Smirnov también se pueden probar si las funciones de distribución de dos poblaciones de donde se obtuvieron las dos muestra, son iguales. Es decir, probar la hipótesis nula:

$$H_0 : F_x = F_y$$

En este caso, la estimación del estadístico de prueba se realiza comparando las dos funciones de distribuciones empíricas acumuladas.

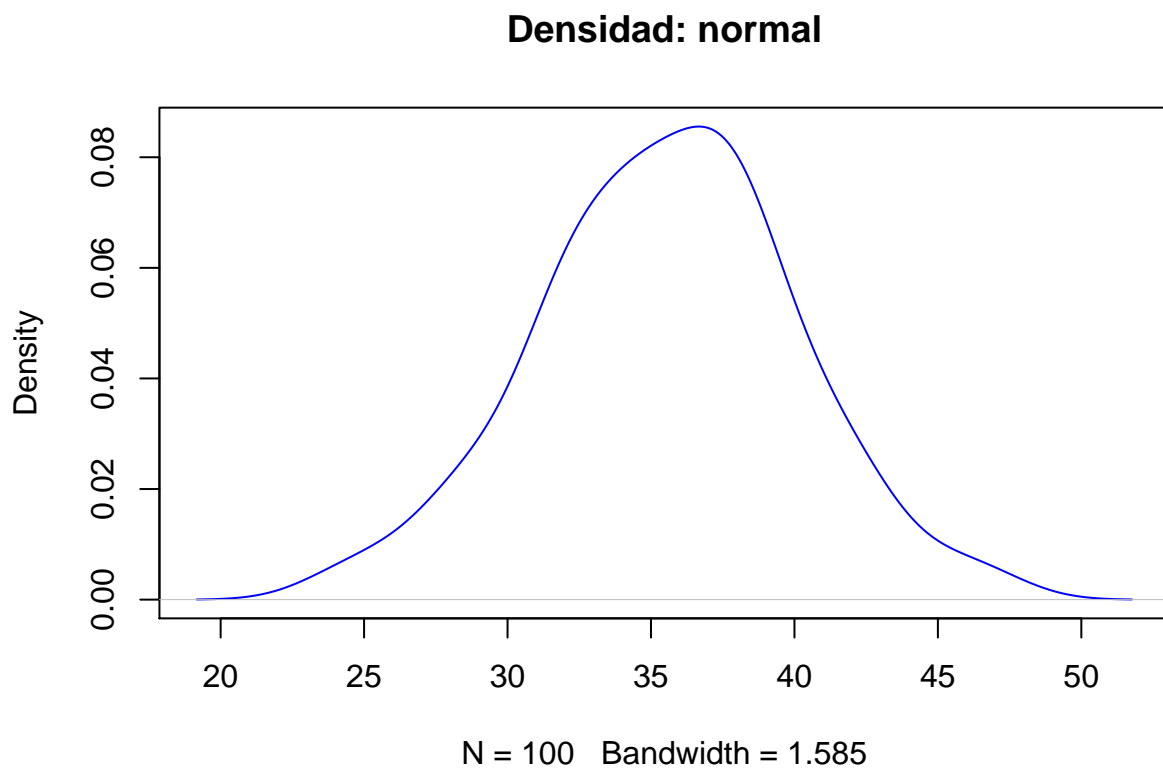
Aplicación

- Datos simulados:

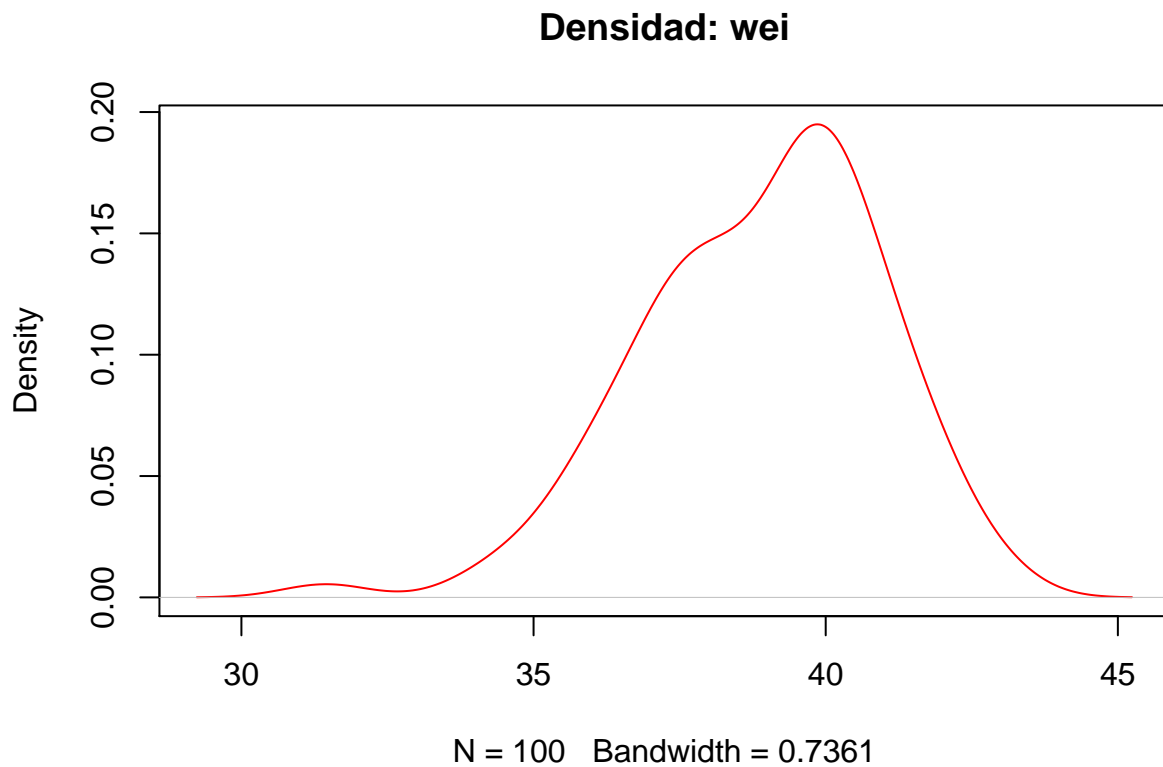
Primero se generan 100 valores aleatorios de una distribución normal con $\mu = 35$ y $\sigma^2 = 25$ y otros 100 de una distribución weibull con $\alpha = 20$ (shape) y $\theta = 40$ (scale).

```
set.seed(1)
normal<-rnorm(100,mean=35,sd=5)
set.seed(1)
wei<-rweibull(100,shape=20,scale = 40)
datos<-data.frame(normal,wei)

plot(density(datos$normal),main = "Densidad: normal",col="blue")
```



```
plot(density(datos$wei),main = "Densidad: wei",col="red")
```

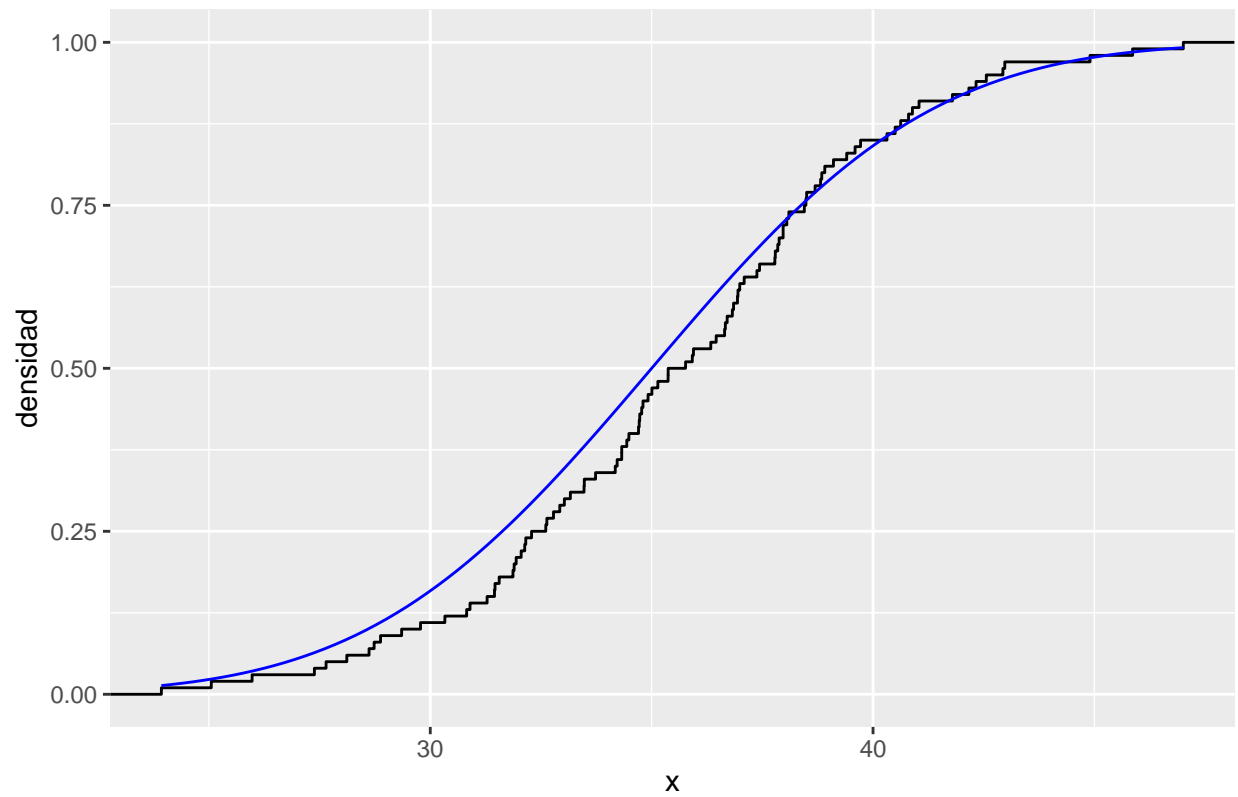


con estas dos “muestras” evaluaremos si las dos variables provienen de una población con distribución normal $\mu = 35$ y $\sigma = 5$. Las funciones de distribución empírica acumulada de las dos muestras versus la distribución normal se presentan a continuación:

```
library(tidyverse)

ggplot(datos, aes(x=normal)) +
  stat_ecdf(geom="step") +
  stat_function(fun=pnorm, color="blue", args = list(35,5)) +
  labs(title="Empírica_normal vs teórica", y="densidad", x="x")
```

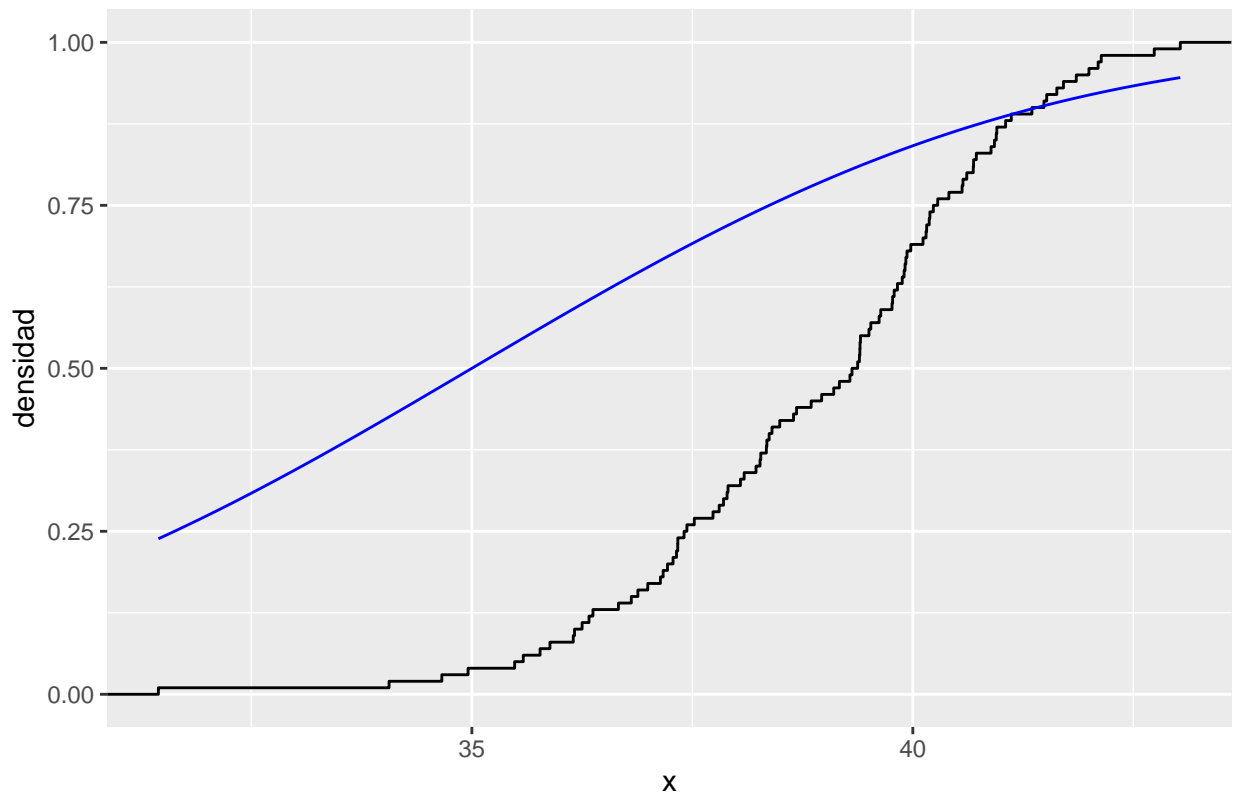
Empírica_normal vs teórica



Ahora lo mismo para las observaciones wei:

```
ggplot(datos, aes(x=wei))+  
  stat_ecdf(geom="step")+  
  stat_function(fun=pnorm, color="blue", args = list(35,5))+  
  labs(title="Empírica_wei vs teórica", y="densidad", x="x")
```


Empírica_wei vs teórica



Utilizando el comando `ks.test` ejecutamos la prueba de Kolmogorov - Smirnov, así:

```
ks.test(datos$normal,"pnorm",35,5,alternative = "t",exact = FALSE)
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: datos$normal  
## D = 0.094659, p-value = 0.3317  
## alternative hypothesis: two-sided
```

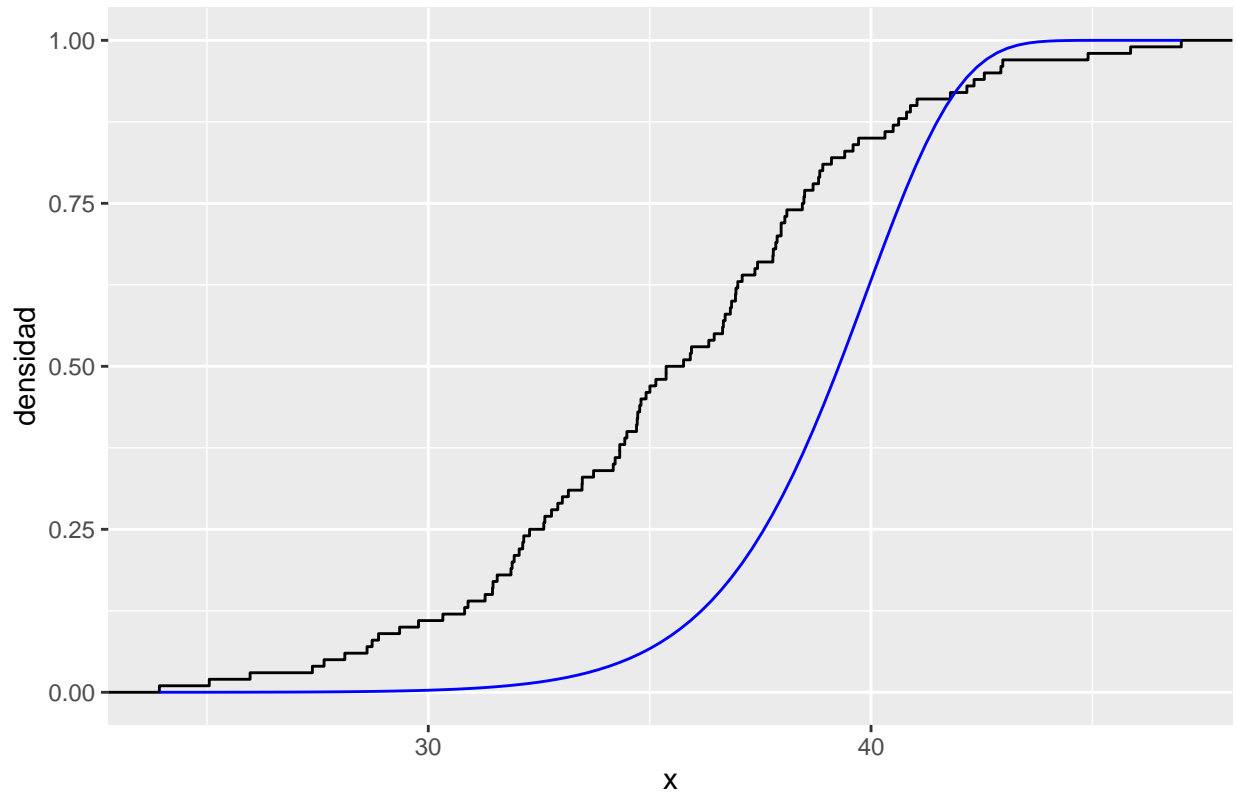
```
ks.test(datos$wei,"pnorm",35,5,alternative = "t",exact = FALSE)
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: datos$wei  
## D = 0.51091, p-value < 2.2e-16  
## alternative hypothesis: two-sided
```

Ahora, evaluaremos si las dos muestras simuladas provienen de una distribución weibull con parámetros $\alpha = 20$ y $\theta = 40$. La visualización de las distribuciones empíricas acumuladas con relación a la distribución de la población son las siguientes:

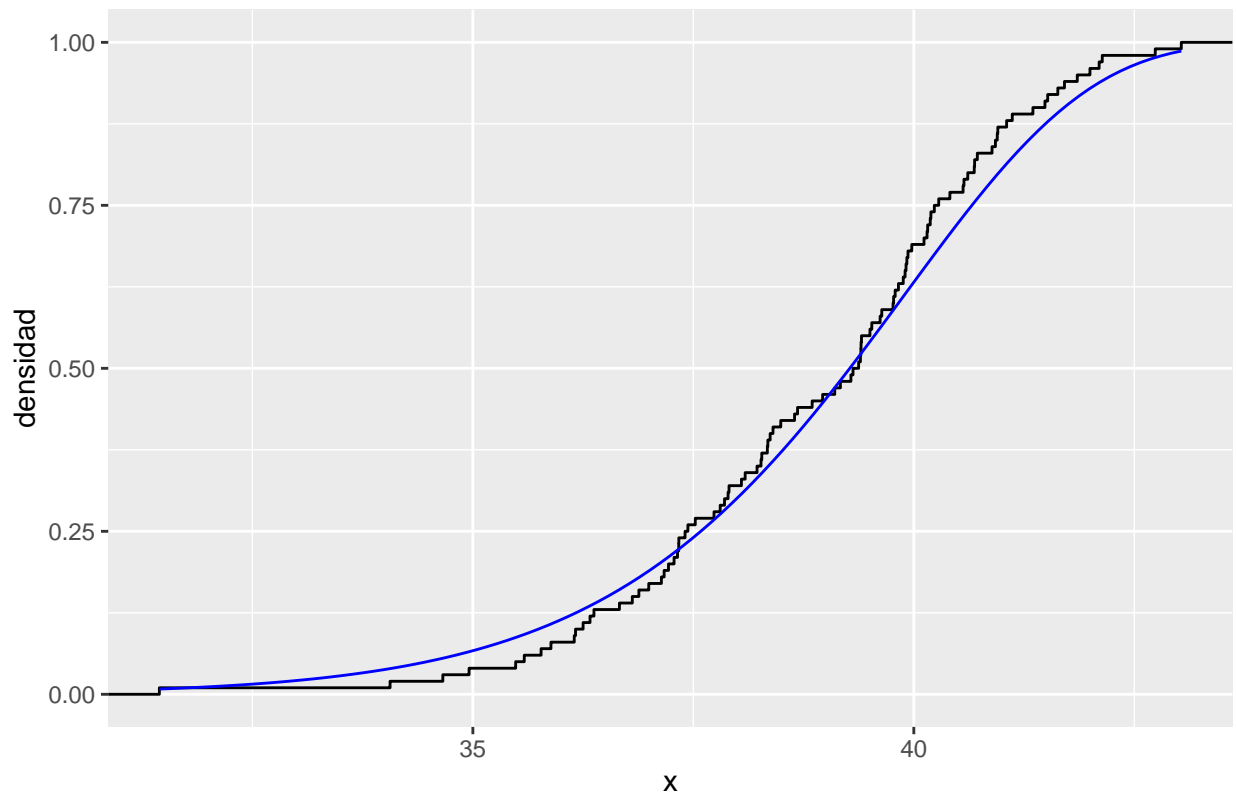
```
ggplot(datos,aes(x=normal))+  
  stat_ecdf(geom="step")+  
  stat_function(fun=pweibull,color="blue",args = list(20,40))+  
  labs(title="Empírica_normal vs teórica", y="densidad",x="x")
```

Empírica_normal vs teórica



```
ggplot(datos,aes(x=wei))+  
  stat_ecdf(geom="step")+  
  stat_function(fun=pweibull,color="blue",args = list(20,40))+  
  labs(title="Empírica_wei vs teórica", y="densidad",x="x")
```

Empírica_wei vs teórica



La aplicación de las pruebas de kolmogorov Smirnov obtienen los siguientes resultados:

```
ks.test(datos$normal,"pweibull",20,40,alternative = "t",exact = FALSE)
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: datos$normal  
## D = 0.44193, p-value < 2.2e-16  
## alternative hypothesis: two-sided
```

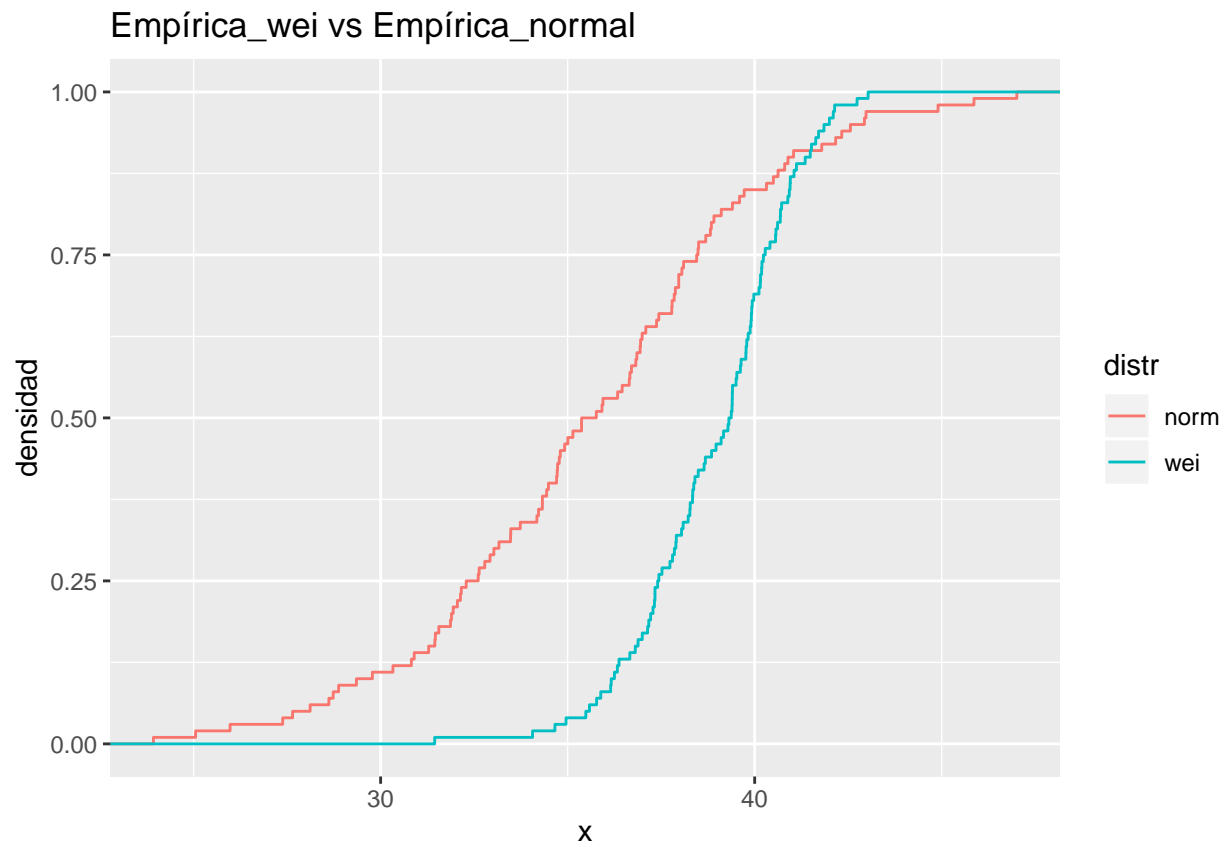
```
ks.test(datos$wei,"pweibull",20,40,alternative = "t",exact = FALSE)
```

```
##  
## One-sample Kolmogorov-Smirnov test  
##  
## data: datos$wei  
## D = 0.076272, p-value = 0.6058  
## alternative hypothesis: two-sided
```

Finalmente, utilicemos las dos muestras para evaluar si las distribuciones de las dos poblaciones son iguales entre sí.

```
d1<-data.frame(rep("norm",100),normal)  
colnames(d1) <- c("distr", "valores")  
d2<-data.frame(rep("wei",100),wei)  
colnames(d2) <- c("distr", "valores")  
datos2<-rbind(d1,d2)
```

```
ggplot(datos2,aes(x=valores,color=distr))+
  stat_ecdf(geom="step")+
  labs(title="Empírica_wei vs Empírica_normal", y="densidad",x="x")
```



```
ks.test(datos$normal,datos$wei,alternative = "t",exact = FALSE)
```

```
##
## Two-sample Kolmogorov-Smirnov test
##
## data: datos$normal and datos$wei
## D = 0.47, p-value = 5.099e-10
## alternative hypothesis: two-sided
```

Pruebas de normalidad

Al evaluar si la muestra obtenida proviene de una distribución normal, y no se conocen los parámetros μ y σ^2 , se puede utilizar una corrección de la prueba de kolmogorov Smirnov que se denomina la prueba de lilliefors. El estadístico de prueba es el mismo pero se realiza una corrección en su distribución ya que esta cambia debido a que se deben estimar estos parámetros a partir de la muestra.

Utilizando los datos generados en la sección anterior, el código para correr esta prueba, que se encuentra en el paquete `nortest` (Gross and Ligges 2015), es el siguiente:

```
#install.packages("nortest")
```

```
library(nortest)
lillie.test(datos$normal)
```

```
##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  datos$normal
## D = 0.047014, p-value = 0.8479
```

```
lillie.test(datos$wei)
```

```
##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  datos$wei
## D = 0.093511, p-value = 0.03111
```

Ahora, se debe mencionar que la prueba de kolmogorov Smirnov no se recomienda para evaluar normalidad ya que se han construido otras pruebas específicas para esta distribución. Entre las más conocidas se encuentra la prueba de Shapiro - Wilk donde la expresión del estadístico es igual a:

$$W = \frac{1}{d} \sum_{i=1}^k a_i (X^{(n-i+1)} - X^{(i)})^2$$

donde

$$d = \sum_{i=1}^n (X_i - \bar{X})^2$$

$X^{(i)}$ es el valor de la variable en la posición i organizando los valores de la variable de menor a mayor, y k es aproximadamente $\frac{n}{2}$. Los coeficientes a_i son constantes que se encuentran tabuladas en libros de estadística no paramétrica.

La evaluación de la prueba de Shapiro Wilk se puede realizar utilizando el comando `shapiro.test` como se presenta a continuación:

```
shapiro.test(datos$normal)
```

```
##
## Shapiro-Wilk normality test
##
## data:  datos$normal
## W = 0.9956, p-value = 0.9876
```

```
shapiro.test(datos$wei)
```

```
##
## Shapiro-Wilk normality test
##
## data:  datos$wei
## W = 0.97503, p-value = 0.05424
```

Una prueba adicional que también se utiliza para evaluar normalidad, es la prueba de Shapiro - Francia. Esta prueba se puede ejecutar utilizando el comando `sf.test`.

```
sf.test(datos$normal)
```

```
##  
## Shapiro-Francia normality test  
##  
## data: datos$normal  
## W = 0.99582, p-value = 0.9719
```

```
sf.test(datos$wei)
```

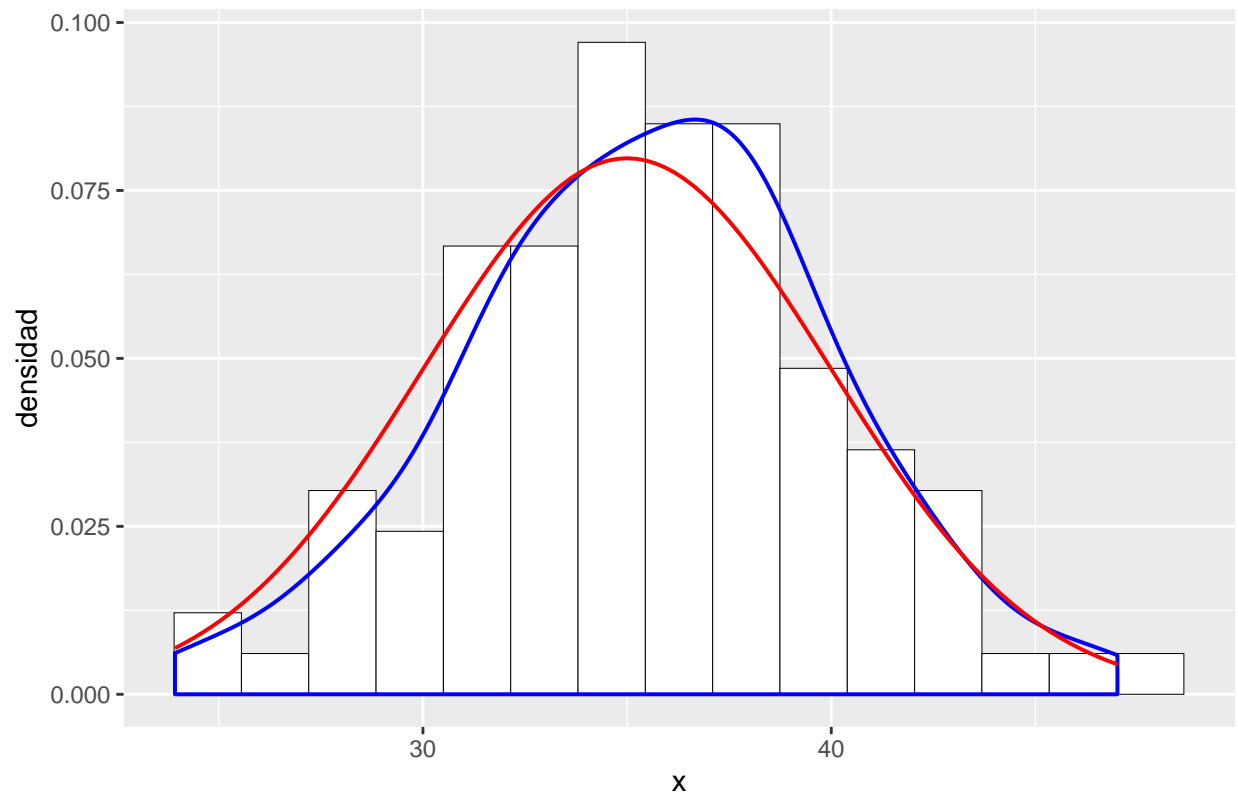
```
##  
## Shapiro-Francia normality test  
##  
## data: datos$wei  
## W = 0.97298, p-value = 0.03767
```

Gráficos para evaluar normalidad

Se recomienda que al aplicar una prueba de bondad y ajuste, también se realicen gráficos que representen la distribución de la variable de interés. Con los datos simulados en la sección anterior, inicialmente podemos obtener el histograma y el gráfico de cajas y bigotes para la variable **normal**:

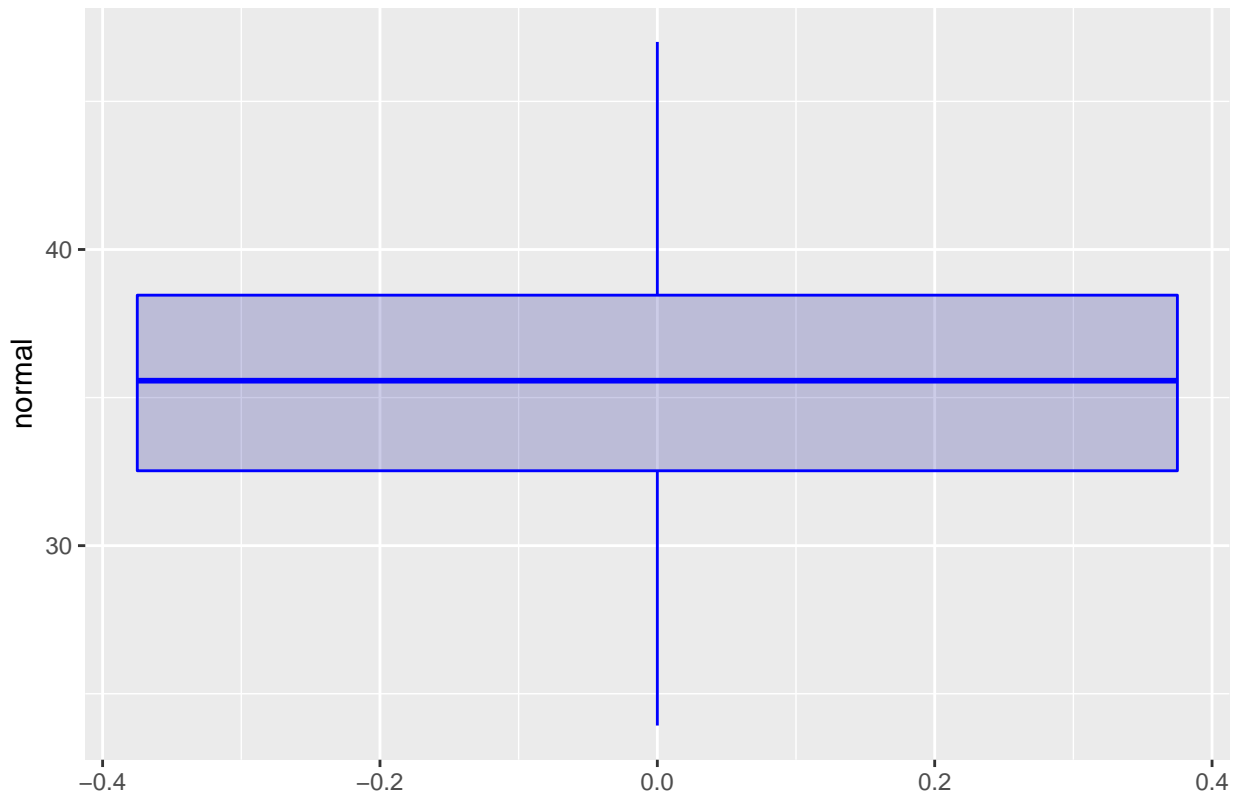
```
ggplot(datos,aes(x=normal))+  
  geom_histogram(bins=15,aes(y = ..density..),fill="white", color = "black",  
                size=0.1)+  
  geom_density(col="blue",size=0.7)+  
  stat_function(fun=dnorm,color="red",args = list(35,5),size=0.7)+  
  labs(title="Histograma", y="densidad",x="x")
```

Histograma



```
ggplot(datos, aes(y=normal)) +  
  geom_boxplot(alpha=0.2, fill="darkblue", color="blue", outlier.color="red") +  
  labs(title = "Gráfico de cajas y bigotes")
```

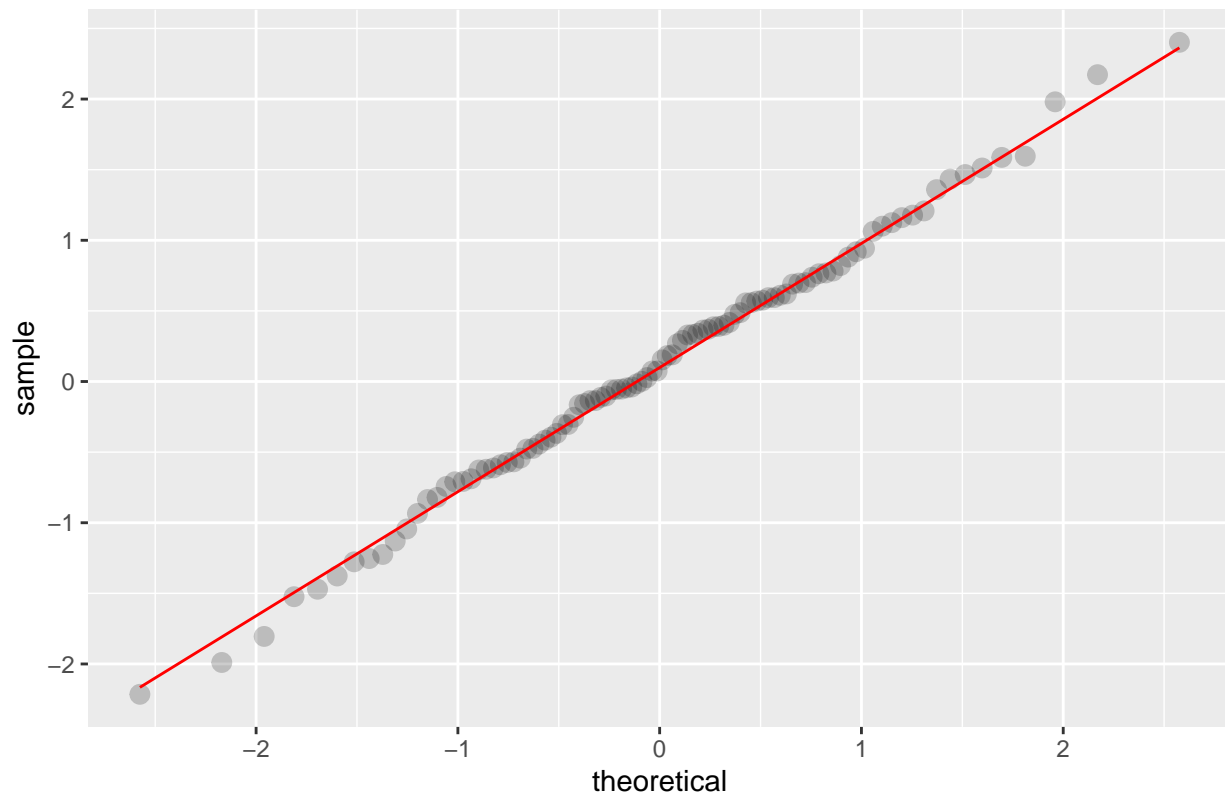
Gráfico de cajas y bigotes



Un gráfico para observar si una variable proviene de una distribución normal es el **qqplot**. Utilizando un gráfico de dispersión que compara los percentiles de la variable de interés (estandarizada, es decir: $z = \frac{x_i - \bar{x}}{s}$) versus los percentiles de una variable con distribución normal estándar. Si los puntos obtenidos se encuentran sobre una recta que pasa por las coordenadas (0,0) a (1,1), indica que la variable tiene una distribución normal. Este gráfico se puede obtener de la siguiente forma:

```
ggplot(datos, aes(sample=((normal-35)/5))) +  
  stat_qq(alpha=0.2, col="black", size=3) +  
  geom_qq_line(col="red") +  
  labs(title="QQplot")
```


QQplot



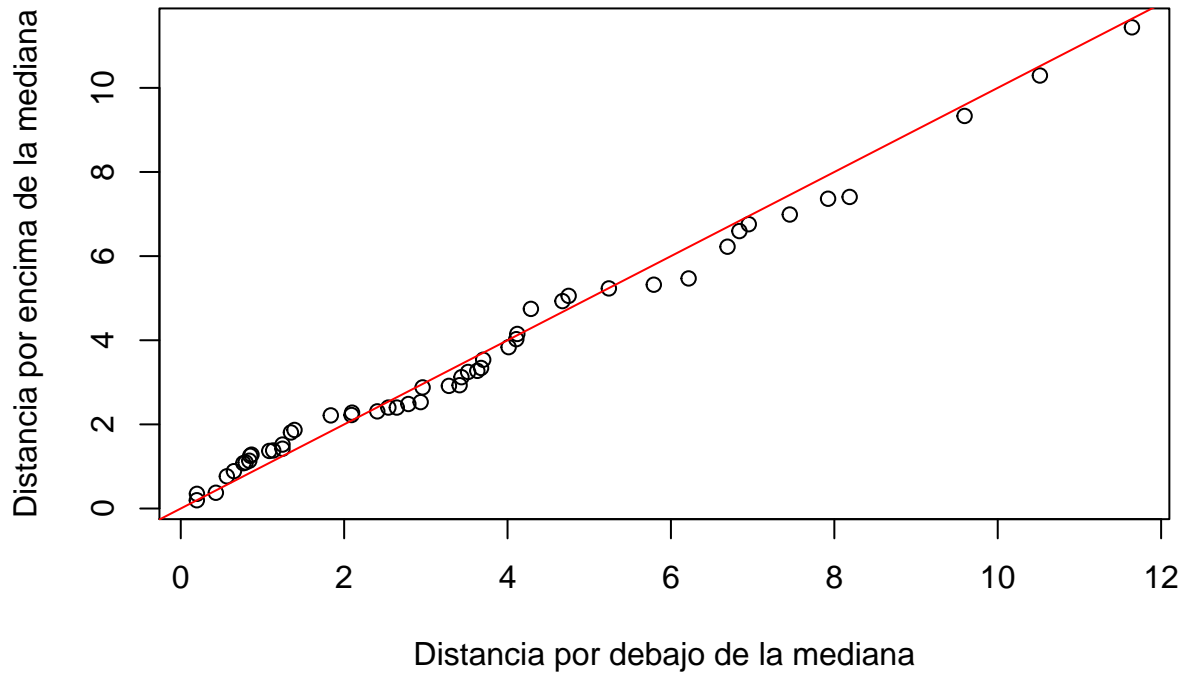
Finalmente se presenta el gráfico de simetría. Este gráfico se construye a partir de un gráfico de dispersión donde en el eje x se colocan las diferencias con relación a la mediana, de todas las observaciones por debajo de la mediana y en el eje y se colocan las mismas diferencia de las observaciones por encima de la mediana. Cada punto en el gráfico refleja las diferencias de las dos observaciones que se ubicada en la misma posición desde la mediana hasta la observación.

El gráfico se puede obtener utilizando la siguiente función:

```
symplot<-function(x)
{n = length(x)
n2 = n %% 2
sx = sort(x)
mx = median(x,na.rm = TRUE)
plot(mx - sx[1:n2], rev(sx)[1:n2] - mx,
      xlab = "Distancia por debajo de la mediana",
      ylab = "Distancia por encima de la mediana", main = "Gráfico de simetría")
abline(a = 0, b = 1, lty = "solid",col="red")
}

symplot(datos$normal)
```

Gráfico de simetría



Prueba del signo

- Supuestos.

Se asume que la variable es continua, aunque la prueba se puede aplicar en variables medidas al menos en una escala ordinal y donde tenga “sentido” la interpretación de la mediana.

- Hipótesis.

Hipótesis nula: La mediana es igual a un cierto valor Me_0

$$H_0 : Me = Me_0$$

Hipótesis alternativas:

1. $H_a : Me \neq Me_0$
2. $H_a : Me > Me_0$
3. $H_a : Me < Me_0$

Si cada valor de la variable se clasifica con signo “-” cuando es menor que Me_0 o con signo “+” cuando es mayor a Me_0 , las tres hipótesis alternativas anteriores se pueden escribir de la siguiente forma, respectivamente:

1. $H_a : P(+)\neq P(-)$
2. $H_a : P(+)> P(-)$
3. $H_a : P(+)< P(-)$

Donde $P(+)$ es la probabilidad de obtener una observación mayor a Me_0 y $P(-)$ es la probabilidad de obtener una observación menor a Me_0 .

- Estadístico de prueba.

Corresponde al número de valores observados por encima de Me_0 (observaciones con signo +) que se denota por \mathbf{k} . La distribución de \mathbf{k} asumiendo que H_0 es verdadera, es *Binomial*($n, 0.5$) donde n es el número de observaciones y 0.5 la probabilidad de acierto (que una observación esté por encima de Me_0). A partir de esta distribución, el valor p para evaluar la hipótesis nula para las tres hipótesis alternativas es igual, respectivamente, a:

1. $2\hat{u}P(x \leq k)$ si $k < \frac{n}{2}$ ó $2\hat{u}P(x \geq k)$ si $k > \frac{n}{2}$
2. $P(x \geq k)$
3. $P(x \leq k)$

- Presencia de ceros.

Los cálculos anteriores se basan en que la variable bajo estudio es de naturaleza continua, es decir que la probabilidad de que una observación sea exactamente igual a Me_0 es cero. Pero en la práctica se pueden presentar observaciones exactamente iguales a Me_0 (si el instrumento de medición no recoge la naturaleza continua de la variable) lo que ocasiona que no se pueda asignar un signo “+” o “-” a la observación. La anterior situación se puede denominar como la presencia de ceros, para lo cual se plantean cuatro alternativas de solución (Gibbons 2003c):

1. Retirar las observaciones con ceros (valores iguales a la Me_0).
2. Colocar la mitad de los ceros con el signo + y la otra mitad con el signo -.
3. Asignarle signos a los ceros de forma aleatoria.
4. Asignar a los ceros el signo que incline la decisión a no rechazar la hipótesis nula.

R implementa la primera alternativa.

- Distribución asintótica.

Para muestras de tamaño moderado, 12 o más observaciones (Gibbons 2003d), realizando el ajuste por continuidad, la distribución binomial se puede aproximar a partir de la distribución normal. De tal forma que:

$$z = \frac{k \pm 0.5 - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

entonces, para un valor de k observado, a partir de la distribución normal estándar se pueden obtener el valor p para cada hipótesis alternativa descrita anteriormente. El uso de “+” o “-” en el numerador de la expresión anterior dependerá de la hipótesis alternativa que se planteó.

- Aplicación de la prueba en datos emparejados.

La prueba del signo también se puede utilizar cuando se obtiene una muestra de observaciones pareadas, por ejemplo cuando se aplican dos evaluaciones al mismo individuo. En este caso la prueba del signo permite evaluar si hay diferencias entre estas dos observaciones. Simplemente se aplica la prueba del signo sobre la diferencia de las dos observaciones definiendo la hipótesis nula con $M_{e0} = 0$, es decir que se evalúa si la mediana de la diferencia es igual a cero.

Aplicación

Caso (Una muestra): Nueve laboratorios evaluaron la efectividad de un cierto medicamento obteniendo los siguientes resultados: 0.41, 0.68, 0.52, 0.82, 0.45, 0.78, 0.96, 0.91 y 0.75. Se quiere evaluar si el medicamento tiene una efectividad de 0.9.

El paquete BSDA (Arnholt 2012) incluye la prueba `SIGN.test` que permite evaluar la prueba del signo para una muestra o para muestras emparejadas. A continuación se carga el paquete y se crea un vector con los datos de efectividad observados:

```
library(BSDA)
efectividad<-c(0.41,0.68,0.52,0.82,0.45,0.78,0.96,0.91,0.75)
```

Se corre la prueba del signo para las tres hipótesis alternativas y se obtienen los siguientes resultados:

```
SIGN.test(efectividad,md=0.9,alternative = "two.sided",conf.level = 0.95)
```

```
##
## One-sample Sign-Test
##
## data:  efectividad
## s = 2, p-value = 0.1797
## alternative hypothesis: true median is not equal to 0.9
## 95 percent confidence interval:
##  0.4554444 0.9030000
## sample estimates:
## median of x
##      0.75
##
## Achieved and Interpolated Confidence Intervals:
##
##                Conf.Level L.E.pt U.E.pt
## Lower Achieved CI    0.8203 0.5200  0.820
## Interpolated CI      0.9500 0.4554  0.903
## Upper Achieved CI    0.9609 0.4500  0.910
```

```
SIGN.test(efectividad,md=0.9,alternative = "greater",conf.level = 0.95)
```

```
##
## One-sample Sign-Test
##
```

```

## data: efectividad
## s = 2, p-value = 0.9805
## alternative hypothesis: true median is greater than 0.9
## 95 percent confidence interval:
## 0.4803333      Inf
## sample estimates:
## median of x
##      0.75
##
## Achieved and Interpolated Confidence Intervals:
##
##           Conf.Level L.E.pt U.E.pt
## Lower Achieved CI    0.9102 0.5200   Inf
## Interpolated CI      0.9500 0.4803   Inf
## Upper Achieved CI    0.9805 0.4500   Inf
SIGN.test(efectividad,md=0.9,alternative = "less",conf.level = 0.95)

```

```

##
## One-sample Sign-Test
##
## data: efectividad
## s = 2, p-value = 0.08984
## alternative hypothesis: true median is less than 0.9
## 95 percent confidence interval:
## -Inf 0.871
## sample estimates:
## median of x
##      0.75
##
## Achieved and Interpolated Confidence Intervals:
##
##           Conf.Level L.E.pt U.E.pt
## Lower Achieved CI    0.9102 -Inf  0.820
## Interpolated CI      0.9500 -Inf  0.871
## Upper Achieved CI    0.9805 -Inf  0.910

```

Otra alternativa es utilizar la prueba exacta de la binomial introduciendo los valores de **k**, **n** y **p** y utilizando el comando `binom.test`

```

binom.test(2,9,p=0.5,alternative="two.sided",conf.level=0.95)

##
## Exact binomial test
##
## data: 2 and 9
## number of successes = 2, number of trials = 9, p-value = 0.1797
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
## 0.02814497 0.60009357
## sample estimates:
## probability of success
##      0.2222222

```

Aunque en el caso anterior, el n es menor a 12; podemos observar como se utilizaría la aproximación a la distribución binomial a partir de la distribución normal. El valor del estadístico de prueba z y su

correspondiente valor p para las tres hipótesis alternativas se presentan a continuación:

```
# Hipótesis alternativa bilateral
```

```
z<-(2+0.5-0.5*9)/(0.5*sqrt(9))
```

```
z
```

```
## [1] -1.333333
```

```
pnorm(z)*2
```

```
## [1] 0.1824224
```

```
# Hipótesis alternativa >
```

```
z<-(2-0.5-0.5*9)/(0.5*sqrt(9))
```

```
z
```

```
## [1] -2
```

```
1-pnorm(z)
```

```
## [1] 0.9772499
```

```
# Hipótesis alternativa <
```

```
z<-(2+0.5-0.5*9)/(0.5*sqrt(9))
```

```
z
```

```
## [1] -1.333333
```

```
pnorm(z)
```

```
## [1] 0.09121122
```

Caso (datos pareados): Se piensa que la susceptibilidad a la hipnosis disminuye con un cierto entrenamiento. Se aplicó una escala para medir la susceptibilidad antes y después del entrenamiento a seis sujetos encontrando los siguientes resultados:

```
##      Antes Despues
## sujeto 1     18     10
## sujeto 2     19     16
## sujeto 3     11      7
## sujeto 4      3      4
## sujeto 5      5      7
## sujeto 6      3      2
```

Se crean dos vectores para los datos observados antes y después del entrenamiento y se corre la prueba del signo para las tres hipótesis alternativas.

```
antes<-c(18,19,11,3,5,3)
```

```
despues<-c(10,16,7,4,7,2)
```

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

```
##
```

```
## Dependent-samples Sign-Test
```

```
##
```

```
## data: antes and despues
```

```
## S = 4, p-value = 0.6875
```

```
## alternative hypothesis: true median difference is not equal to 0
```

```
## 95 percent confidence interval:
```

```

## -1.9 7.6
## sample estimates:
## median of x-y
##      2
##
## Achieved and Interpolated Confidence Intervals:
##
##           Conf.Level L.E.pt U.E.pt
## Lower Achieved CI   0.7812  -1.0   4.0
## Interpolated CI    0.9500  -1.9   7.6
## Upper Achieved CI   0.9688  -2.0   8.0
SIGN.test(antes,despues,alternative = "g",conf.level = 0.95)

##
## Dependent-samples Sign-Test
##
## data:  antes and despues
## S = 4, p-value = 0.3437
## alternative hypothesis: true median difference is greater than 0
## 95 percent confidence interval:
## -1.633333      Inf
## sample estimates:
## median of x-y
##      2
##
## Achieved and Interpolated Confidence Intervals:
##
##           Conf.Level  L.E.pt U.E.pt
## Lower Achieved CI   0.8906 -1.0000  Inf
## Interpolated CI    0.9500 -1.6333  Inf
## Upper Achieved CI   0.9844 -2.0000  Inf
SIGN.test(antes,despues,alternative = "l",conf.level = 0.95)

##
## Dependent-samples Sign-Test
##
## data:  antes and despues
## S = 4, p-value = 0.8906
## alternative hypothesis: true median difference is less than 0
## 95 percent confidence interval:
##      -Inf 6.533333
## sample estimates:
## median of x-y
##      2
##
## Achieved and Interpolated Confidence Intervals:
##
##           Conf.Level L.E.pt U.E.pt
## Lower Achieved CI   0.8906  -Inf 4.0000
## Interpolated CI    0.9500  -Inf 6.5333
## Upper Achieved CI   0.9844  -Inf 8.0000

```

Prueba de rangos de signos de Wilcoxon

- Supuestos.

La variable es de naturaleza continua pero la prueba se puede aplicar en variables medidas en escala al menos de intervalo. Bajo el supuesto de que la población tiene una distribución simétrica, la mediana y la media serían iguales y la hipótesis sobre la mediana aplicaría para la media. A diferencia de la prueba del signo, esta no sólo considera la posición de la observación (rangos) sino también su magnitud.

- Hipótesis.

Hipótesis nula: La mediana es igual a un cierto valor Me_0

$$H_0 : Me = Me_0$$

Hipótesis alternativa:

1. $H_a : Me \neq Me_0$
2. $H_a : Me > Me_0$
3. $H_a : Me < Me_0$

- Estadístico de prueba.

Se clasifica cada observación x_i con signo “-” si toma valores menores a Me_0 o con signo “+” si toma valores mayores a Me_0 . Para cada observación se calcula el valor absoluto de la diferencia entre x_i menos Me_0 lo que se denota por $|D_i|$. Este resultado se ordena de menor a mayor y se asigna la posición (rango r_i) que ocupa dentro del total de n observaciones, es decir el $|D_i|$ más pequeño tendrá la primera posición que equivale a un rango $r_i = 1$, el segundo $|D_i|$ más pequeño tendrá la segunda posición que equivale a un rango $r_i = 2$ y así sucesivamente. El estadístico de prueba T^+ corresponde a la suma de los rangos asignados a las observaciones clasificadas con el signo “+”, es decir:

$$T^+ = \sum_{i=1}^n Z_i \hat{u} r_i$$

donde $Z_i = 1$ si $|D_i| > 0$ o $Z_i = 0$ si $|D_i| < 0$.

Ahora, si H_0 es verdadera entonces

$$E(T^+ | H_0) = \frac{n(n+1)}{4}$$

y

$$Var(T^+ | H_0) = \frac{n(n+1)(2n+1)}{24}$$

Para muestra de tamaño moderados ($n \geq 15$) (Gibbons 2003e), se tiene que:

$$z = \frac{4T^+ - n(n+1)}{\sqrt{\frac{2n(n+1)(2n+1)}{3}}} \sim N(0, 1)$$

A partir de la expresión anterior se pueden calcular el valor p asociado a cada hipótesis alternativa listada anteriormente. Se puede realizar el ajuste por continuidad, sumando o restando 0.5 al numerador de z ; si z es negativo se suma y si z es positivo se resta.

- Empates y ceros.

En la evaluación del estadístico de prueba descrito anteriormente, hay dos situaciones que se pueden presentar en la práctica: la primera es que distintas observaciones tomen el mismo valor lo que se denomina empates, ocasionando que no se pueda asignar directamente un rango a cada observación, y la segunda es que x_i es igual a Me_0 lo que se denomina presencia de ceros, ocasionando que no se le pueda colocar un signo “+” o “-”. Cuando se presenta la primera situación, una alternativa es asignar el promedio de los rangos iniciales a cada una de las observaciones que presentaron el empate.

Cuando se presentan las dos situaciones anteriores se deben realizar ajustes sobre el valor esperado y la varianza de T^+ . Al valor esperado se le debe restar la mitad de la suma de los rangos asignados a las observaciones con D_i iguales a cero, y a la varianza se le deben restar el resultado de las dos expresiones siguientes:

1. $\frac{\sum t(t^2-1)}{48}$ donde t es el número de empates y la sumatoria se realiza sobre el número de veces que se presentan estos empates.
 2. $\frac{n_0(n_0+1)(2n_0+1)}{24}$ donde n_0 es el número de observaciones con D_i iguales a cero.
- Distribución exacta.

En distintos libros de estadística se encuentran tablas con la función de distribución de probabilidades de la estadística T^+ , la cual aplican cuando se tienen menos de 15 observaciones. En este documento no se incluyen estas tablas, al contrario presentaremos los comandos en R que permiten calcular el valor exacto del valor p para esta prueba.

- Datos emparejados.

Al igual que la prueba del signo, la prueba de rango de signos de Wilcoxon se puede aplicar para una muestra de datos pareados (x_i, y_i) , donde D_i es igual a $x_i - y_i$ y $Me_0 = 0$.

Aplicación

Caso (Una muestra): A partir de una escala se midió el estado de salud auto-reportado (VES) por un grupo de 18 individuos. La escala toma valores entre 0 y 100 puntos, donde valores bajos indican un peor estado de salud y puntajes altos indican un mejor estado de salud. Se quiere saber si la mediana del VES es igual a 80 puntos. Los valores obtenidos de los 18 sujetos fueron: {80 78 78 77 76 76 88 89 89 90 95 65 60 60 56 56 50 45}.

En los paquete `MASS` y `exactRankTests` se encuentra los comandos `wilcox.test` y `wilcox.exact`, respectivamente, para realizar la prueba de rangos del signo de Wilcoxon. Tres observaciones sobre los dos comandos:

1. `wilcox.exact` permite calcular el valor p exacto ante la presencia de empates, mientras `wilcox.test` no.
2. Los dos comandos calculan la estadística T^+ eliminando los valores de $D_i = 0$.
3. `wilcox.test` permite realizar el ajuste por continuidad para el cálculo del valor p aproximado por la distribución normal, mientras que `wilcox.exact` no lo permite.

A continuación se incluyen los valores del estado de salud auto-reportado en un vector nombrado `VES` y se presenta la aplicación del comando `wilcox.test` incluida en el paquete `MASS` (Venables and Ripley 2002):

```
library(MASS)
VES<-c(80,78,78,77,76,76,88,89,89,90,95,65,60,60,56,56,50,45)

wilcox.test(VES,mu=80,exact=T,alternative = "t",conf.int = 0.95)

## Warning in wilcox.test.default(VES, mu = 80, exact = T, alternative =
## "t", : cannot compute exact p-value with ties
```

```

## Warning in wilcox.test.default(VES, mu = 80, exact = T, alternative =
## "t", : cannot compute exact confidence interval with ties

## Warning in wilcox.test.default(VES, mu = 80, exact = T, alternative =
## "t", : cannot compute exact p-value with zeroes

## Warning in wilcox.test.default(VES, mu = 80, exact = T, alternative =
## "t", : cannot compute exact confidence interval with zeroes

##
## Wilcoxon signed rank test with continuity correction
##
## data: VES
## V = 40.5, p-value = 0.09258
## alternative hypothesis: true location is not equal to 80
## 95 percent confidence interval:
## 63.50004 82.50006
## sample estimates:
## (pseudo)median
## 72.49996

```

```
wilcox.test(VES,mu=80,exact=F,correct = T, alternative = "t", conf.int = 0.95)
```

```

##
## Wilcoxon signed rank test with continuity correction
##
## data: VES
## V = 40.5, p-value = 0.09258
## alternative hypothesis: true location is not equal to 80
## 95 percent confidence interval:
## 63.50004 82.50006
## sample estimates:
## (pseudo)median
## 72.49996

```

Ahora, se presenta la aplicación del comando `wilcox.exact` (Hothorn and Hornik 2015):

```
library(exactRankTests)

wilcox.exact(VES,mu=80,exact = T,alternative = "t",conf.int = 0.95)
```

```

##
## Exact Wilcoxon signed rank test
##
## data: VES
## V = 40.5, p-value = 0.09045
## alternative hypothesis: true mu is not equal to 80
## 95 percent confidence interval:
## 63.5 82.5
## sample estimates:
## (pseudo)median
## 72.25

```

```
wilcox.exact(VES,mu=80,exact = F,alternative = "t",conf.int = 0.95)
```

```

##
## Asymptotic Wilcoxon signed rank test
##

```

```
## data: VES
## V = 40.5, p-value = 0.08808
## alternative hypothesis: true mu is not equal to 80
## 95 percent confidence interval:
## 63.99998 82.50004
## sample estimates:
## (pseudo)median
## 72.49998
```

Caso (Muestra pareada): Seis estudiantes fueron medidos en la mañana y en la noche con el fin de probar si la estatura de las personas en estos dos momentos son diferentes (ejemplo tomado de Daniel). Las tallas observadas fueron las siguientes:

```
##           mañana noche
## sujeto 1    201    203
## sujeto 2    182    178
## sujeto 3    191    186
## sujeto 4    188    183
## sujeto 5    188    181
## sujeto 6    174    165
```

La aplicación de la prueba de rangos del signo de wilcoxon con el comando `wilcox.test` se presenta a continuación.

```
# vectores con información:
manana<-c(201,182,191,188,188,174)
noche<-c(203,178,186,183,181,165)

# evaluación de la prueba:
wilcox.test(manana,noche,paired = T,exact = T, correct = F, conf.int = 0.95)
```

```
## Warning in wilcox.test.default(manana, noche, paired = T, exact = T,
## correct = F, : cannot compute exact p-value with ties
```

```
## Warning in wilcox.test.default(manana, noche, paired = T, exact = T,
## correct = F, : cannot compute exact confidence interval with ties
```

```
##
## Wilcoxon signed rank test
##
```

```
## data: manana and noche
## V = 20, p-value = 0.0458
## alternative hypothesis: true location shift is not equal to 0
## 95 percent confidence interval:
## 0.9999852 8.0000384
## sample estimates:
## (pseudo)median
## 4.99998
```

```
wilcox.test(manana,noche,paired = T,exact = F,correct = T,conf.int = 0.95)
```

```
## Warning in wilcox.test.default(manana, noche, paired = T, exact = F,
## correct = T, : requested conf.level not achievable
```

```
##
## Wilcoxon signed rank test with continuity correction
##
## data: manana and noche
```

```
## V = 20, p-value = 0.05848
## alternative hypothesis: true location shift is not equal to 0
## 90 percent confidence interval:
## 1.499973 7.000015
## sample estimates:
## (pseudo)median
## 4.99998
```

y la aplicación con el comando `wilcox.exact` sería:

```
wilcox.exact(manana,noche,paired = T,exact = T,conf.int = 0.95)
```

```
##
## Exact Wilcoxon signed rank test
##
## data: manana and noche
## V = 20, p-value = 0.0625
## alternative hypothesis: true mu is not equal to 0
## 95 percent confidence interval:
## -2 9
## sample estimates:
## (pseudo)median
## 4.75
```

```
wilcox.exact(manana,noche,paired = T,exact = F,conf.level = 0.95)
```

```
##
## Asymptotic Wilcoxon signed rank test
##
## data: manana and noche
## V = 20, p-value = 0.0458
## alternative hypothesis: true mu is not equal to 0
```

Prueba U de Mann Whitney - Wilcoxon

- Supuestos.

Se asume que la variable es de naturaleza continua, pero la prueba se puede aplicar en variables medidas al menos en escala ordinal.

- Hipótesis.

Hipótesis nula: La distribución de las dos variables son iguales.

$$H_0 : F_X = F_Y$$

Hipótesis alternativa:

1. $H_a : F_X \neq F_Y$
2. $H_a : F_X > F_Y$
3. $H_a : F_X < F_Y$

- Estadístico de prueba:

Para dos muestras independientes sean X_i con $i = 1, 2, \dots, n_1$ y Y_j con $j = 1, 2, \dots, n_2$, se presentan dos formas para calcular el estadístico de prueba que se denota por U :

Forma 1:

Tenemos que $C_{ij} = 1$ si $Y_j < X_i$ y $C_{ij} = 0$ si $Y_j > X_i$, entonces:

$$U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} C_{ij}$$

Forma 2:

$$U = R_1 - \frac{n_1(n_1 + 1)}{2}$$

donde R_1 es la suma de los rangos de X_i obtenidos de ordenar las dos muestras de menor a mayor.

Ahora, si H_0 es verdadera,

$$E(U|H_0) = \frac{n_1 n_2}{2}$$

y

$$V(U|H_0) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12}$$

Para muestras n_1 y $n_2 > 10$ (Canavos 1994), la distribución de U se puede aproximar a partir de la distribución normal estándar donde:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0, 1)$$

- Distribución exacta

En este documento no se presenta el desarrollo de la distribución exacta de U , pero se presentan los comando en R para obtener la probabilidad exacta del valor p .

- Empates

Con relación a la primera forma de cálculo del estadístico de prueba, el manejo de los empates se realiza sobre la definición de C_{ij} donde se presentan dos alternativas:

la primera es:

$C_{ij} = .$	si:
1	$Y_j < X_i$
0.5	$Y_j = X_i$
0	$Y_j > X_i$

y la segunda:

$C_{ij} = .$	si:
1	$Y_j < X_i$
0.5	$Y_j = X_i$
-1	$Y_j > X_i$

R utiliza la primera alternativa.

Con relación a la segunda forma de cálculo del estadístico de prueba, el manejo de los empates se realiza de la misma forma que se ha trabajado en las pruebas anteriores: asignado el promedio de los rangos a los empates.

Adicional a los ajuste sobre la estadística de prueba, al utilizar la aproximación a partir de la distribución normal estándar, se debe realizar un ajuste sobre la varianza U multiplicándola por la siguiente expresión:

$$1 - \frac{\sum t^3 - t}{(n_1 + n_2)((n_1 + n_2)^2 - 1)}$$

donde t es el número de empates y la sumatoria se realiza sobre el número de veces que se presentan los empates.

Aplicación

Caso: Se dictó una asignatura a siete estudiantes utilizando un método tradicional, mientras que a seis estudiantes se dictó la misma asignatura utilizando un método nuevo. La nota obtenidas al final del curso para los dos grupos fueron: tradicional:{68 69 72 78 79 80 84} y nuevo:{60 64 68 70 72 73}.

Para obtener esta prueba, a continuación cargamos el paquete `exactRankTests` (Hothorn and Hornik 2015) y creamos dos vectores con las notas de cada grupo.

```
library(exactRankTests)
actual<-c(68,69,72,78,79,80,84)
nuevo<-c(60,64,68,70,72,73)
```

Utilizando el comando `wilcox.exact` corremos la prueba U de Mann-Whitney a partir de la distribución exacta y de la aproximación por la distribución normal estándar.

```
wilcox.exact(actual,nuevo,exact = T, alternative = "t",conf.int = 0.95)
```

```
##
## Exact Wilcoxon rank sum test
##
## data: actual and nuevo
## W = 34, p-value = 0.0676
## alternative hypothesis: true mu is not equal to 0
## 95 percent confidence interval:
## -1 15
```

```
## sample estimates:
## difference in location
##           7.5
wilcox.exact(actual,nuevo,exact = F, alternative = "t",conf.int = 0.95)

##
## Asymptotic Wilcoxon rank sum test
##
## data:  actual and nuevo
## W = 34, p-value = 0.06257
## alternative hypothesis: true mu is not equal to 0
## 95 percent confidence interval:
## -0.0000027749 15.0000532804
## sample estimates:
## difference in location
##           7.999994
```

Prueba de rachas de Wald - Wolfowitz

- Supuesto.

Aplica para variable de naturaleza continua, medida en una escala al menos ordinal.

- Hipótesis.

Hipótesis nula: La distribución de la población del grupos X es igual a la distribución de la población de grupo Y .

$$H_0 : F_X = F_Y$$

Hipótesis alternativa:

$$H_a : F_X \neq F_Y$$

- Estadística de prueba

El estadístico de prueba se conoce como *rachas* y se denota por r . Se obtiene de ordenar de menor a mayor las observaciones de los datos de los dos grupos combinados. A partir de este ordenamiento, r es el número de veces que se encuentran valores consecutivos del mismo grupo. Por ejemplo, si se observan los siguientes valores para dos grupos sean A y B:

$$A: \{1,5,9,11\} \quad B: \{2,4,6,7\}$$

Al ordenar las observaciones de menor a mayor de los dos grupos obtenemos el siguiente arreglo:

$$1 \ 2 \ 4 \ 5 \ 6 \ 7 \ 9 \ 11$$

Donde se pueden identificar cinco “rachas” ($r = 5$) o cinco secuencias de números consecutivos del mismo grupo.

La función de distribución de la estadística r tiene la siguiente expresión:

$f(r) = .$	<i>si:</i>
$\frac{{}^2\binom{n_1 - 1}{r/2 - 1} \binom{n_2 - 1}{r/2 - 1}}{\binom{n_1 + n_2}{n_1}}$	r es par
$\frac{\binom{n_1 - 1}{(r - 1)/2} \binom{n_2 - 1}{(r - 3)/2} + \binom{n_1 - 1}{(r - 3)/2} \binom{n_2 - 1}{(r - 1)/2}}{\binom{n_1 + n_2}{n_1}}$	r es impar

Nota: $\binom{a}{b} = 0$ si $b > a$

Al construir el arreglo de datos de los dos grupos, ordenados de menor a mayor, si las distribuciones de los dos grupos fueran iguales, es decir si H_0 fuera verdadera, se observaría un número alto de rachas ya que se verían los datos “muy mezclados”, pero si las dos distribuciones fueran diferentes los valores de un grupos se ubicarían en un cierto lugar del arreglo de datos ordenados obteniendo un número bajo de rachas. Por lo anterior, la hipótesis nula se rechaza para valores bajos de la estadística r , por lo tanto el valor p para la hipótesis alternativa de diferencia de las dos distribuciones se calcula obteniendo la probabilidad menor o igual al valor de rachas observado.

- Distribución asintótica.

Para n_1 y $n_2 > 10$ (Gibbons 2003f), la distribución de r se puede aproximar a partir de la distribución normal estándar, con la siguiente expresión:

$$Z = \frac{r - E(r|H_0)}{\sqrt{V(r|H_0)}}$$

donde

$$E(r|H_0) = 1 + \frac{2n_1n_2}{n_1 + n_2}$$

y

$$V(r|H_0) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1 + n_2)^2(n_1 + n_2 - 1)}$$

- Empates

Cuando el supuesto de continuidad en la práctica no se cumple, es posible que se presenten empates en las variable que estamos comparado dando distintas formas de organización de las dos muestras. Cuando los empates se presentan dentro de uno de los dos grupos no hay problema ya que las diferentes organizaciones que se realicen no van a alterar el número de rachas que se presentan, por lo tanto el valor p siempre va a ser el mismo. ahora, si los empates se presentan en observaciones de los dos grupos, el número de rachas dependiendo la organización que se realice si puede cambiar. Un manejo conservador de esta situación es realizar todas las posibles organizaciones de los datos, calcular el número de rachas para cada ordenamiento y seleccionar el mayor número de rachas como estadística de prueba. El número total de posibles ordenamiento se puede calcular de la siguiente forma:

$$\prod_{i=1}^k \binom{s_i + t_i}{s_i}$$

donde s_i es el número de observaciones empatadas del grupo 1, t_i es el número de observaciones empatadas del grupo 2 y k es el número de empates con observaciones de los dos grupos que se presentaron.

- Otra aplicación.

La prueba de *rachas* también se utiliza para evaluar *aleatoriedad*, es decir, se quiere saber si la ocurrencia de un evento sigue un cierto patrón. Por ejemplo, se quiere saber si la llegada de los pacientes a la consulta médica ocurre de forma aleatoria con relación al sexo, o llegan más hombre o mujeres al comienzo del día, o al final del día, de manera intercalada u otro posible patrón de llegada. Para este escenario, el estadístico de prueba (número de rachas) se obtiene de la misma forma que en la aplicación anterior, pero el ordenamiento no se realiza por los valores observado de una variable, sino por el orden de llegada. La diferencia que se presenta para este escenario es el cálculo del valor p, es decir, la hipótesis nula de *aleatoriedad* se rechazaría no solo cuando se presente un número bajo de rachas sino también un número alto.

En nuestro ejemplo del orden de llegada de los pacientes, si se presentan dos rachas (todas las mujeres/hombre llegan primero - caso extremo) o el orden de llegada es de manera intercalada (hombre,mujer, hombre, mujer - caso extremo), esto sería evidencia para rechazar la hipótesis nula. Por consiguiente el cálculo del valor p se realizaría a dos colas ya que un número elevado de rachas también sería evidencia de la presencia de un cierto patrón. En este escenario no debemos preocuparnos por la presencia de empates ya que el ordenamiento es único.

Aplicación

Caso: En un grupo de 12 hombres y 11 mujeres estudiantes de medicina se midió el tiempo (horas por semana) de estudio extra clase con el fin de determinar si existen diferencias entre los dos grupos. Se encontraron los siguientes resultados: Hombres {5.6 4.8 5.7 7.0 5.4 3.4 5.4 8.0 3.4 4.7 4.8 5.0} y Mujeres {6.5 7.8 7.5 8.1 7.9 8.1 6.5 8.0 4.3 5.6 7.1}

Al ordenar las observaciones de menor a mayor y marcar a los hombres con 1 y las mujeres con cero, se obtiene un vector de la siguiente forma:

```
{1 1 0 1 1 1 1 1 1 1 0 1 0 0 1 0 0 0 0 1 0 0 0}
```

En el arreglo anterior se observan 10 rachas.

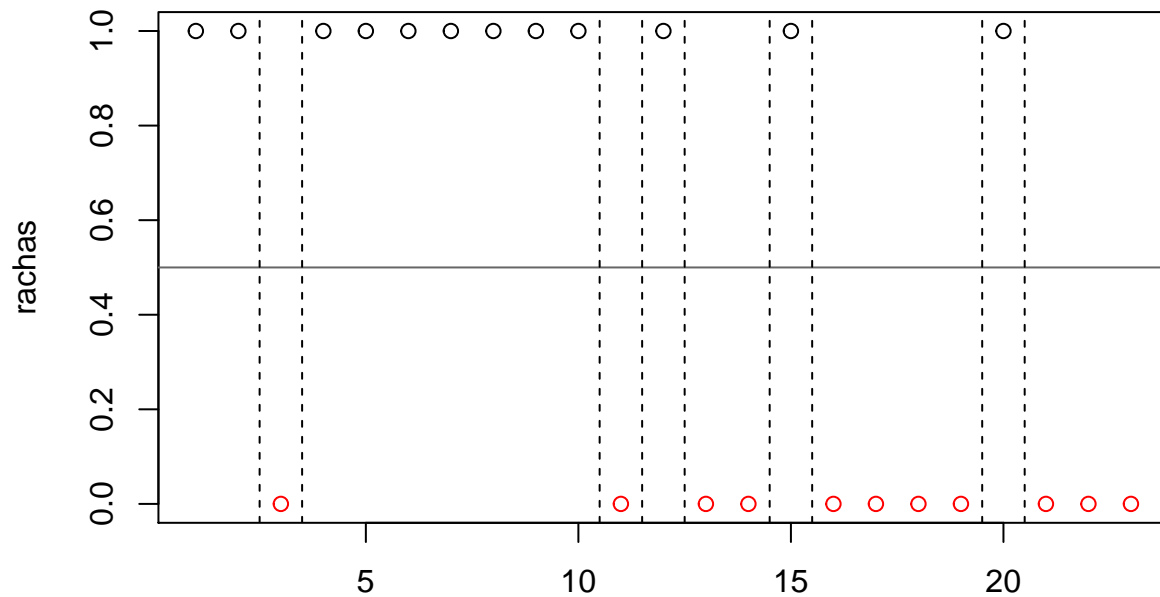
cargando el paquete `randtests` (Caeiro and Mateus 2014) y creando un vector de 1 y 0 que contiene el arreglo de los datos observados, se obtiene la prueba de Wald-Wolfowitz de la siguiente forma:

```
library(randtests)
rachas<-c(1,1,0,1,1,1,1,1,1,1,0,1,0,0,1,0,0,0,0,1,0,0,0)

runs.test(rachas,alternative = "left.sided",threshold = 0.5,
          pvalue = "exact",plot=F)
```

```
##
## Runs Test
##
## data:  rachas
## statistic = -1.0599, runs = 10, n1 = 12, n2 = 11, n = 23, p-value
## = 0.2017
## alternative hypothesis: trend
```

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



```
##
## Runs Test
##
## data:  rachas
## statistic = -1.0599, runs = 10, n1 = 12, n2 = 11, n = 23, p-value
```

```
## = 0.1446  
## alternative hypothesis: trend
```

Prueba de McNemar

Esta prueba permite comparar un desenlace dicotómico medido dos veces en el mismo sujeto. En general, los resultados de la variable desenlace se describen como éxito (1) o fracaso (0). La prueba se desarrolla a partir de los resultados que se organizan de la siguiente forma:

		Medición 2		
Medición 1	Éxito	Fracaso	Total	
Éxito	$X(11)$	$X(12)$	$X(1.)$	
Fracaso	$X(21)$	$X(22)$	$X(2.)$	
Total	$X(.1)$	$X(.2)$	N	

donde $X(ij)$ corresponde al número de sujetos que en la primera medición obtuvieron como respuesta i y en la segunda medición obtuvieron como respuesta j ; para $i, j = 1, 2$. Si en lugar de colocar el número de sujetos en cada una de las celdas de la tabla anterior, se coloca la proporción de individuos con relación al total de los N sujetos, se obtendría una tabla como la siguiente:

		Medición 2		
Medición 1	Éxito	Fracaso	Total	
Éxito	$P(11)$	$P(12)$	$P(1.)$	
Fracaso	$P(21)$	$P(22)$	$P(2.)$	
Total	$P(.1)$	$P(.2)$	P	

donde $P(ij)$ corresponde a la proporción de individuos con una primera medición igual a i y la segunda medición igual a j ; es decir $P(ij) = \frac{X(ij)}{N}$.

- Hipótesis.

Hipótesis nula: La proporción de éxitos en la primera medición es igual a la proporción de éxitos en la segunda medición.

$$H_0 : P(1.) = P(.1)$$

La anterior afirmación, también se puede expresar como

$$H_0 : P(12) = P(21)$$

Hipótesis alternativa:

1. $H_a : P(12) \neq P(21)$
2. $H_a : P(12) > P(21)$
3. $H_a : P(12) < P(21)$

Una tercera forma de plantear la misma hipótesis nula es a partir de la proporción $p = X(12)/X(12) + X(21)$. Si $P(12) = P(21)$ entonces la H_0 se expresa en términos de una disparidad de la siguiente forma:

$$H_0 : Odds = \frac{P(12)}{1-P(12)} = 1$$

- Estadístico de prueba.

El estadístico de prueba es $X(12)$, es decir el número de sujetos que obtuvieron éxito en la primera medición y fracaso en la segunda medición. Si asumimos que la H_0 es verdadera, $X(12)$ tiene una distribución $Binomial(X(12) + X(21); 0.5)$. A partir de la estadística de prueba $X(12)$ y su función de distribución, se evalúa la hipótesis nula frente a las tres hipótesis alternativas.

- Distribución asintótica.

Para muestras grandes

$$z = \frac{X(12) - X(21)}{\sqrt{X(12) + X(21)}} \sim N(0, 1)$$

Ahora, incluyendo la corrección por continuidad

$$z = \frac{|X(12) - X(21)| - 1}{\sqrt{X(12) + X(21)}} \sim N(0, 1)$$

Con base en la estadística z , se puede evaluar la hipótesis nula con relación a las tres hipótesis alternativas descritas anteriormente.

Si se eleva al cuadrado la estadística z , se puede evaluar la hipótesis bilateral, a partir de la siguiente expresión:

$$\chi^2 = \frac{[X(12) - X(21)]^2}{[X(12) + X(21)]} \sim \chi^2_{(1)}$$

Incluyendo la corrección por continuidad:

$$\chi^2 = \frac{[|X(12) - X(21)| - 1]^2}{[X(12) + X(21)]} \sim \chi^2_{(1)}$$

Casos y controles pareados 1 a 1

Una aplicación de la prueba de McNemar es en los diseños de **casos y controles emparejados** donde se evalúa si estar expuesto a un cierto factor se relaciona con presentar un desenlace. En este diseño, se eligen un grupo de sujetos que presentaron el desenlace (casos) y a cada uno se le asigna un sujeto (control) que no presentó el desenlace pero que tiene las mismas características del caso con relación a posibles variables de confusión (emparejamiento).

La hipótesis nula a evaluar en este diseño se plantea en términos de la razón de disparidades (Odds Ratio:OR) de estar expuesto entre los casos versus los controles:

$$H_0 : OR = 1$$

indicando, en caso de rechazar la hipótesis nula, que existe una relación entre la exposición y el desenlace.

Para evaluar esta hipótesis, la tabla en la cual se basa el cálculo del estadístico de prueba se construye de la siguiente forma:

	Control		
Caso	Expuesto	No Expuesto	Total
Expuesto	$X(11)$	$X(12)$	$X(1.)$
No expuesto	$X(21)$	$X(22)$	$X(2.)$
Total	$X(.1)$	$X(.2)$	N

donde $X(i, j)$ para $i, j = Expuesto, Noexpuesto$ corresponde al número de parejas donde el caso toma un valor de i y el control de j . De la tabla anterior se obtiene el estadístico de prueba ($X(12)|H_0 \sim Binomial(X(12) + X(21), 0.5)$) a partir del cual se evalúa H_0 .

De la tabla anterior se puede obtener una estimación del **OR** a partir de la expresión de Mantel-Haenszel para estimar el **OR ajustado** para datos estratificados, donde cada estrato en este caso, corresponde a cada pareja conformada por un caso y un control. La expresión de Mantel-Haenszel bajo este escenario es la siguiente:

$$OR_{MH}^{\hat{}} = \frac{\sum_i^N (a_i d_i)}{\sum_i^N (c_i b_i)}$$

donde a_i, b_i, c_i y d_i se obtiene de construir la siguiente tabla para cada pareja de caso y control:

Expuestos	Caso	Control
Si	a_i	b_i
No	c_i	d_i

es decir, que se construyen N tablas donde cada una sólo puede tomar una de las siguientes cuatro opciones:

	Opción A		Opción B		Opción C		Opción D	
Expuestos	Caso	Control	Caso	Control	Caso	Control	Caso	Control
Si	1	1	1	0	0	1	0	0
No	0	0	0	1	1	0	1	1

en consecuencia la expresión anterior se simplifica en el numerador al número de tablas de la opción B (número de parejas con caso expuesto y el control no expuesto), y en el denominador al número de tablas de la opción C (número de parejas donde el caso no esta expuesto y el control si).

$$OR_{MH}^{\hat{}} = \frac{X_{(12)}}{X_{(21)}}$$

Aplicación

Caso: Se quiere probar si una nueva terapia física ayuda a disminuir el dolor en un grupo de pacientes que sufrieron cierta lesión. Se aplicó una escala que clasifica al paciente en si presenta un alto grado o bajo grado de dolor, antes y después de la terapia. Fueron evaluados 20 pacientes y los datos se presentan en la siguiente tabla donde 1 corresponde a alto grado de dolor y 0 a bajo grado.

Med	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
antes	1	1	1	1	0	0	1	0	1	0	1	0	0	1	1	0	1	0	0	1
después	1	0	1	0	0	0	1	0	0	1	1	0	0	0	0	1	1	1	0	0

Utilizando el paquete `exact2x2` (Fay 2010) y el comando `mcnemar.exact` se puede evaluar la prueba bajo la distribución binomial.

A continuación se carga el paquete y se definen dos vectores que incluyen los resultados observados

```
library(exact2x2)
antes<-c(1,1,1,1,0,0,1,0,1,0,1,0,0,1,1,0,1,0,0,1)
despues<-c(1,0,1,0,0,0,1,0,0,1,1,0,0,0,0,1,1,1,0,0)
```

Los resultados de la prueba son los siguientes:

```
mcnemar.exact(despues,antes,conf.level = 0.95)
```

```
##
## Exact McNemar test (with central confidence intervals)
##
## data:  despues and antes
## b = 6, c = 3, p-value = 0.5078
## alternative hypothesis: true odds ratio is not equal to 1
## 95 percent confidence interval:
##  0.4271342 12.3592268
## sample estimates:
## odds ratio
##          2
```

Para evaluar la prueba con base en la distribución asintótica (χ^2) se puede utilizar el comando `mcnemar.test` incluido en el paquete `stats` que permite realizar la prueba con o sin la corrección por continuidad incluyendo la opción `correct=FALSE` O `TRUE`. A continuación se presenta la aplicación de este comando:

```
mcnemar.test(antes,despues,correct=FALSE)
```

```
##
## McNemar's Chi-squared test
##
## data:  antes and despues
## McNemar's chi-squared = 1, df = 1, p-value = 0.3173
```

```
mcnemar.test(antes,despues,correct=TRUE)
```

```
##
## McNemar's Chi-squared test with continuity correction
##
## data:  antes and despues
## McNemar's chi-squared = 0.44444, df = 1, p-value = 0.505
```

Prueba de la mediana

- Supuesto.

Aplica para variable de naturaleza continua medidas al menos en escala ordinal.

- Hipótesis.

Hipótesis nula: Las k variables tienen la misma distribución.

$$H_0 : F_1(x) = F_2(x) = \dots = F_k(x)$$

Hipótesis alternativa:

$$H_a : F_i(x) \neq F_j(x) \text{ por lo menos para un } i \neq j.$$

- Estadístico de prueba:

Para k muestras independientes, a partir de la mediana de todos los valores observados en las k muestras independientes, sea Me_{global} , se puede construir la siguiente tabla:

	X_1	X_2	...	X_k	Total
$< Me_{global}$	n_{11}	n_{12}	...	n_{1k}	$n_{1\bar{u}}$
$> Me_{global}$	n_{21}	n_{22}	...	n_{2k}	$n_{2\bar{u}}$
Total	$n_{.1}$	$n_{.2}$...	$n_{.k}$	n

donde n_{1j} es el número de observaciones de la muestra j ($j = 1, 2, \dots, k$) que toma valores menores a Me_{global} y n_{2j} es el número de observaciones de la muestra j que toma valores mayores a Me_{global} . Si la hipótesis nula fuera cierta, se encontrarían proporciones de observaciones por encima y por debajo de Me_{global} parecidas en las k muestras. En base a lo anterior, el estadístico de prueba es igual a:

$$\sum_{i=1}^2 \sum_{j=1}^k \frac{(n_{ij} - E_{ij})^2}{E_{ij}} \sim \chi_{(k-1)}^2$$

donde

$$E_{ij} = \frac{n_{i\bar{u}} n_{\bar{u}j}}{n}$$

- Prueba exacta de Fisher.

Cuando el número de individuos en cada celda es muy pequeño, se debe utilizar la prueba exacta de Fisher. Una de las reglas que se utilizan para determinar si se debe utilizar Fisher, es si el 20% o más de las celdas presentan $E_{ij} < 5$.

Aplicación

Caso: Se aplicó una escala para medir el nivel de estrés laboral a personas de cinco ocupaciones distintas: médicos, economistas, arquitectos, abogados e ingenieros. La escala toma valores de 0 a 100 puntos donde valores altos indican altos niveles de estrés laboral. Los resultados de la escala se presentan a continuación: médicos: {82,72,88,75,95,80,75,86,84}
 economistas: {60,50,66,75,80,66,70,75}
 ingenieros: {45,56,68,64,70,65,48,49,50}
 arquitectos: {45,35,45,26,30,45,46,39}
 abogados: {22,10,25,35,26,29,16,48,22,19}

Primero se deben cargar las observaciones organizadas en dos vectores, el primero registra el grupo al cual pertenece la observación, y el segundo incluye los valores observados de la variable de interés.

Una forma de cargar los datos es utilizando el comando `import` del paquete `rio`.

```
library(rio)
```

```
## Warning: package 'rio' was built under R version 3.6.1
```

```
datos<-import("estres.xlsx")
```

Después de la importación de los datos, la prueba se realiza con el comando `mood.medtest` incluido en el paquete `RVAideMemoire` (Hervé 2016). Los resultados de la prueba bajo la distribución χ^2 o bajo la prueba exacta de Fisher se presentan a continuación:

```
library(RVAideMemoire)
```

```
## *** Package RVAideMemoire v 0.9-73 ***
```

```
mood.medtest(datos$estres,datos$profesion,exact = FALSE)
```

```
##
```

```
## Mood's median test
```

```
##
```

```
## data: datos$estres by structure(c(5L, 5L, 5L, 5L, 5L, 5L, 5L, 5L, 5L, 4L, 4L, 4L, 4L, datos$estres
```

```
## X-squared = 31.585, df = 4, p-value = 2.325e-06
```

```
mood.medtest(datos$estres,datos$profesion,exact = TRUE)
```

```
##
```

```
## Mood's median test
```

```
##
```

```
## data: datos$estres by structure(c(5L, 5L, 5L, 5L, 5L, 5L, 5L, 5L, 5L, 4L, 4L, 4L, 4L, datos$estres
```

```
## p-value = 3.048e-08
```

Prueba de Kruskal-Wallis

- Supuestos.

Aplica para variables de naturaleza continua medidas al menos en una escala ordinal.

- Hipótesis.

Hipótesis nula: Las k variables tienen la misma distribución.

$$H_0 : F_1(x) = F_2(x) = \dots = F_k(x)$$

Hipótesis alternativa:

$$H_a : F_i(x) \neq F_j(x) \text{ por lo menos para un } i \neq j.$$

- Estadístico de prueba.

Para la construcción del estadístico de prueba, se toman las n observaciones de los k grupos, y se ordenan de menor a mayor; a partir de esta organización se asignan los rangos correspondientes donde R_i denota la suma de los rangos del grupo i , con $i = 1, 2, \dots, k$. Asumiendo que H_0 es verdadera, la suma de rangos esperados para cada grupo sería igual a $\frac{n_i(n+1)}{2}$, donde n_i es el número de observaciones del grupo i . Con base en lo anterior la estadística de prueba es igual a:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^k \frac{1}{n_i} \left[R_i - \frac{n_i(n+1)}{2} \right]^2 \sim \chi_{(k-1)}^2$$

Ahora, si se rechaza la hipótesis nula, la pregunta que surge es ¿cuál de las parejas de comparaciones entre los k grupos es la que presenta diferencias?. Para responder esta pregunta, se obtiene para las $\binom{k}{2}$ posibles diferentes parejas de grupos, la siguiente estadística (Dunn 1964):

$$Z_{ij} = \frac{|\bar{R}_i - \bar{R}_j|}{((n(n+1)/12)(1/n_i + 1/n_j))^{1/2}} \sim N(0, 1)$$

donde $\bar{R}_i = \frac{R_i}{n_i}$. Con base en la estadística Z_{ij} se evalúa la hipótesis nula $H_0 : F_i = F_j$ para un i, j específicos.

Cuando se evalúa la estadística Z_{ij} , se deben ajustar los valores de p . Existen distintos métodos de ajustes, una alternativa es definir un nivel de significancia para evaluar la estadística H por ejemplo de $\alpha = 0.2$ y el nivel de significancia para evaluar la estadística Z_{ij} sería igual a $\frac{\alpha}{k(k-1)}$. Otra alternativa es utilizar el método de Bonferroni; en este método se multiplica el valor p de cada Z_{ij} por el número de comparaciones entre parejas que se pueden realizar $(\frac{k(k-1)}{2})$.

Finalmente se presenta el método propuesto por Holm que se desarrolla a partir del siguiente algoritmo.

Primero se calcula el valor p de cada comparación a partir de la estadística Z_{ij} , a continuación, se organizan los valores p de menor a mayor y a cada uno se le asigna su posición l (denótese $p_l = 1, 2, \dots, k(k-1)/2$).

Cada valor p_l se multiplica por $\frac{k(k-1)}{2} - l + 1$ y se denota por pH_l . Finalmente, el valor p que se utiliza para evaluar si la distribución es igual entre los dos grupos específicos es:

$$p(\text{Holm})_l = \min(1, \max(pH_l; pH_{l-1}))$$

- Empates.

Al igual que en pruebas anteriores ante la presencia de empates a cada observación correspondiente se le asigna el promedio de los rangos iniciales. Adicionalmente, se realiza el ajuste de la estadística H dividiéndola por la siguiente expresión:

$$1 - \frac{\sum t(t^2 - 1)}{n(n^2 - 1)}$$

De igual forma, el ajuste sobre la estadística Z_{ij} se realiza a partir de la siguiente expresión:

$$Z_{ij} = \frac{|\bar{R}_i - \bar{R}_j|}{((n(n+1)/12 - B)(1/n_i + 1/n_j))^{1/2}}$$

donde $B = \frac{\sum (t^3 - t)}{12(n-1)}$

Aplicación

Caso: Se aplicó una escala para medir el estado de salud (VES) auto-reportado por un grupo de personas de distintos grupos étnicos. La escala toma valores entre 0 y 100 donde a valores altos indican mejor estado de salud. Los resultados se presentan a continuación: indígenas: {78 79 89 79 81 83 95} gitanos: {67 78 78 79 80} negros: {95 99 89 87 90 92} otros: {78 80 81 78 67 70 71 73}

El comando `kruskal.test` del paquete `stats` incluye la prueba. Primero se crean cuatro vectores con las observaciones de cada grupo, y a partir de estos se ejecuta el comando:

```
indigenas<-c(78,79,89,79,81,83,95)
gitanos<-c( 67, 78, 78, 79,80)
negros<-c(95,99,89,87,90,92)
otros<-c(78,80,81,78,67,70,71,73)
kruskal.test(list(indigenas,gitanos,negros,otros))
```

```
##
## Kruskal-Wallis rank sum test
##
## data: list(indigenas, gitanos, negros, otros)
## Kruskal-Wallis chi-squared = 15.88, df = 3, p-value = 0.0012
```

Para evaluar la seis posibles parejas a comparar, se puede utilizar el comando `posthoc.kruskal.dunn.test` incluido en el paquete `PMCMR` (Pohlert 2014). El comando incluye tres opciones de ajuste del valor p (no ajuste, Holm y Bonferroni) que se presentan a continuación:

```
library(PMCMR)
posthoc.kruskal.dunn.test(list(indigenas,gitanos,negros,otros),
                           p.adjust.method = "none")
```

```
## Warning in posthoc.kruskal.dunn.test.default(list(indigenas, gitanos,
## negros, : Ties are present. z-quantiles were corrected for ties.
##
## Pairwise comparisons using Dunn's-test for multiple
## comparisons of independent samples
##
## data: list(indigenas, gitanos, negros, otros)
##
## 1      2      3
## 2 0.10967 -      -
```

```

## 3 0.12066 0.00295 -
## 4 0.03367 0.77547 0.00028
##
## P value adjustment method: none
posthoc.kruskal.dunn.test(list(indigenas,gitanos,negros,otros),
                           p.adjust.method = "holm")

## Warning in posthoc.kruskal.dunn.test.default(list(indigenas, gitanos,
## negros, : Ties are present. z-quantiles were corrected for ties.

##
## Pairwise comparisons using Dunn's-test for multiple
## comparisons of independent samples
##
## data: list(indigenas, gitanos, negros, otros)
##
## 1      2      3
## 2 0.3290 -      -
## 3 0.3290 0.0148 -
## 4 0.1347 0.7755 0.0017
##
## P value adjustment method: holm
posthoc.kruskal.dunn.test(list(indigenas,gitanos,negros,otros),
                           p.adjust.method = "bonferroni")

## Warning in posthoc.kruskal.dunn.test.default(list(indigenas, gitanos,
## negros, : Ties are present. z-quantiles were corrected for ties.

##
## Pairwise comparisons using Dunn's-test for multiple
## comparisons of independent samples
##
## data: list(indigenas, gitanos, negros, otros)
##
## 1      2      3
## 2 0.6580 -      -
## 3 0.7239 0.0177 -
## 4 0.2020 1.0000 0.0017
##
## P value adjustment method: bonferroni

```

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.

- Hipótesis.

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

$$H_0 : F_1(x) = F_2(X) = \dots = F_k(X)$$

Hipótesis alternativa:

$$H_a : F_i(X) \neq F_j(X) \text{ por lo menos para un } i \neq j.$$

- Estadístico de prueba.

Al evaluar la variable X , en n individuos k veces, los resultados se pueden organizar en la siguiente matriz:

Individuos	1	2	...	k
1	X11	X12	...	X1k
2	X21	X22	...	X2k
:	:	:	...	:
n	Xn1	Xn2	...	Xnk

donde X_{ij} es el valor observado del sujeto i en la repetición j . A partir de la matriz anterior, para cada individuo (filas) se asignan rangos de menor a mayor dependiendo el valor de la variable que tomó en la repetición j , con $j = 1, 2, \dots, k$. Realizando lo anterior se obtiene una matriz con la siguiente información:

Individuos	1	2	...	k	$\sum_{j=1}^k r_{.j}$
1	r11	r12	...	r1k	$k(k+1)/2$
2	r21	r22	...	r2k	$k(k+1)/2$
:	:	:	...	:	:
n	rn1	rn2	...	rnk	$k(k+1)/2$
$\sum_{i=1}^n r_{i.}$	r_1	r_2	...	r_k	$nk(k+1)/2$

Con base en la tabla anterior, el estadístico de prueba Q es igual a:

$$Q = \frac{12S}{nk(k+1)} = \frac{12 \sum_{j=1}^k r_j^2}{nk(k+1)} - 3n(k+1) \sim \chi_{k-1}^2$$

donde

$$S = \sum_{j=1}^k \left[r_j - \frac{n(k+1)}{2} \right]^2 = \left[\sum_{j=1}^k r_j^2 \right] - \frac{n^2 k(k+1)^2}{4}$$

Si a partir de Q se rechaza la hipótesis nula, lo anterior indica que existe por lo menos una pareja de repeticiones que provienen de distribuciones diferentes. para evaluar las comparaciones de las posibles parejas, se utiliza la estadística

$$z_{ij} = \frac{|r_i - r_j|}{\sqrt{\frac{nk(k+1)}{6}}}$$

que tiene una distribución normal estándar. Una alternativa para evaluar la hipótesis nula global ($H_0 : F_1(x) = F_2(X) = \dots = F_k(X)$) es definir un nivel de significancia igual a $\alpha = 0.2$ y para las pruebas de comparación de parejas utilizar un nivel de significación de $\frac{\alpha}{k(k-1)}$.

- Empates.

Ante la presencia de empates, la estadística de prueba Q se ajusta de la siguiente forma:

$$Q = \frac{12(k-1)S}{nk(k^2-1) - \sum \sum t(t^2-1)}$$

donde t es el número de observaciones empatadas, y la doble sumatoria corresponde a la suma de todos los empates que se presentan tanto por repeticiones ($j=1,2,\dots,k$) como por individuos ($i=1,2,\dots,n$). Como en los casos anteriores, la asignación de rangos de las observaciones empatadas se realiza colocando el promedio de los rangos correspondientes.

Aplicación

Caso: Nueve estudiantes de medicina fueron evaluados en tres asignaturas distintas obteniendo las notas que se presentan a continuación.

ID	Básica	Fisiología	Anatomía
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
8	75	72	81
9	88	81	83

Primero se crea una matriz con las observaciones de los 9 individuos y las tres repeticiones (asignaturas)

```
notas<-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("basicas","fisilogia","anatomia")))
print(notas)
```

```
## basics fisilogia anatomia
## 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
## 8      75         72         81
```

```
## 9      88      81      83
```

Utilizando el comando `friedman.test` se aplica la prueba.

```
friedman.test(notas)
```

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

dos opciones de ajustes sobre el valor p que se encuentran implementados en *R* son Conover y Nemenyi. Los comandos se encuentran en el paquete PMCMR (Pohlert 2014).

```
library(PMCMR)  
posthoc.friedman.conover.test(notas,p.adjust.method = "bonferroni")
```

```
##  
## Pairwise comparisons using Conover's test for a two-way  
##           balanced complete block design  
##  
## data:  notas  
##  
##           basicas fisilogia  
## fisilogia 8.5e-05 -  
## anatomia  0.5036  0.0015  
##  
## P value adjustment method: bonferroni
```

```
posthoc.friedman.nemenyi.test(notas,p.adjust.methods="none")
```

```
##  
## Pairwise comparisons using Nemenyi multiple comparison test  
##           with q approximation for unreplicated blocked data  
##  
## data:  notas  
##  
##           basicas fisilogia  
## fisilogia 0.013 -  
## anatomia  0.759  0.086  
##  
## P value adjustment method: none
```

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 “acierto” y 0 el “fracaso”).

- Hipótesis.

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 una proporción de acierto es diferente entre dos repeticiones.

$$H_a : P_i \neq P_j \text{ por lo menos para un } i \neq j.$$

- Estadístico de prueba.

La evaluación de una variable dicotómica en n individuos y k repeticiones, se pueden organizar en la siguiente tabla:

Individuos	1	2	...	k	$\sum_{j=1}^k X_{ij}$
1	X11	X12	...	X1k	R1
2	X21	X22	...	X2k	R2
:	:	:	...	:	:
n	Xn1	Xn2	...	Xnk	Rn
$\sum_{i=1}^n X_{ij}$	C · 1	C · 2	...	C · k	N

donde X_{ij} es 1 (acierto) o 0 (fracaso) en el individuo i de la repetición j , R_i es el total de aciertos en el individuo i y C_j es el total de aciertos en la repetición j . Con base en la tabla anterior, la estadística de prueba Q es igual a:

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

que tiene una distribución χ_{k-1}^2 .

Aplicación

Caso: Se aplicaron tres tratamientos a 12 pacientes con el fin de establecer cual de los tratamientos era más efectivo. Si el paciente respondía al tratamiento se marcaba con 1 mientras que si el paciente no respondía se marcaba con 0. En la siguiente tabla se observan los resultados obtenidos.

ID	Tratamiento 1	Tratamiento 2	Tratamiento 3
1	0	1	1
2	0	1	1
3	0	1	0
4	1	1	0
5	0	0	0
6	1	1	1
7	1	1	1
8	1	1	0
9	0	0	1
10	0	1	0

ID	Tratamiento 1	Tratamiento 2	Tratamiento 3
11	0	1	1
12	1	1	1

Primero, se carga el paquete `RVAidememore` (Hervé 2016) y se crean tres vectores donde el primero incluye los resultados de cada tratamiento para cada individuo (vectores de 1 y 0), el segundo identifica el tratamiento aplicado para cada individuo, y el tercero es el identificador de cada sujeto. Los tres vectores deben tener el mismo número de datos.

```
library(RVAideMemoire)

respuesta<-c(0,0,0,1,0,1,1,1,0,0,0,1,1,1,1,1,0,1,1,1,0,1,1,1,1,0,0,0,1,
            1,0,1,0,1,1)

tratamientos<-c(1,1,1,1,1,1,1,1,1,1,1,2,2,2,2,2,2,2,2,2,2,2,2,3,3,3,3,
               3,3,3,3,3,3,3,3)

ID<-c(1,2,3,4,5,6,7,8,9,10,11,12,1,2,3,4,5,6,7,8,9,10,11,12,1,2,3,4,5,6,
      7,8,9,10,11,12)
```

A partir de los tres vectores anteriores, la prueba se ejecuta a partir del comando `cochran.qtest` de la siguiente forma:

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

##
## Cochran's Q test
##
## data: respuesta by tratamientos, block = ID
## Q = 4.75, df = 2, p-value = 0.09301
## alternative hypothesis: true difference in probabilities is not equal to 0
## sample estimates:
## proba in group          <NA>          <NA>
##          0.4166667          0.8333333          0.5833333
```

Medidas de Asociación

Sean X y Y dos variables medidas en n individuos, donde $(X_1, Y_1); (X_2, Y_2); \dots (X_n, Y_n)$ corresponde a las n parejas de observaciones. Se puede definir que dos parejas tienen una relación “directa” si para $X_j > X_i$ también se observa que $Y_j > Y_i$ o si para $X_j < X_i$ también se observa que $Y_j < Y_i$. También se puede definir que dos parejas tienen una relación “indirecta” si para $X_j > X_i$ se observa que $Y_j < Y_i$ o si para $X_j < X_i$ se observa que $Y_j > Y_i$. Una medida de asociación (φ) evalúa cual de las dos relaciones anteriores es la que más se presenta, sobre el total de $\binom{n}{2}$ posibles comparaciones entre parejas de observaciones.

A continuación se presentan siete criterios que son deseables en una medida de asociación (Gibbons 2003a):

1. $\varphi = 1$ si todas las comparaciones entre dos parejas tienen una relación directa.
2. $\varphi = -1$ si todas las comparaciones entre dos parejas tienen una relación indirecta.
3. Si no se cumplen 1 y 2, la medida debe tomar valores entre -1 y 1 y cuando se presentan más relaciones directas φ se acerca a 1 y cuando se presentan más relaciones indirectas φ se acerca a -1.
4. $\varphi = 0$ si las dos variables X y Y son independientes o no se relacionan.
5. $\varphi_{X,Y} = \varphi_{Y,X} = \varphi_{-X,-Y} = \varphi_{-Y,-X}$.
6. $\varphi_{-X,Y} = \varphi_{X,-Y} = -\varphi_{X,Y}$.
7. $\varphi_{X,Y} = \varphi_{h(X,Y)}$ para cualquier función $h(\cdot)$ que no altere el ordenamiento de cada una de las variables.

Dos medidas que cumplen con los siete criterios anteriores son: el coeficiente τ de Kendall y el coeficiente de correlación de rangos de Spearman ρ .

Coeficiente tau de Kendall

- Cálculo del coeficiente.

Para una muestra de n parejas de la forma

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$$

el coeficiente τ de Kendall es igual a

$$\tau = p_c - p_d$$

donde $p_c = P[(X_j - X_i)(Y_j - Y_i) > 0]$ es la probabilidad de parejas con relación directa y $p_d = P[(X_j - X_i)(Y_j - Y_i) < 0]$ es la probabilidad de parejas con relación indirecta.

La estimación de τ se obtiene de la siguiente forma: Se construyen dos matrices U y V de tamaño $n \times n$. Si se define $signo(u)$ como:

$signo(u) = .$	$si :$
1	$u > 0$
0	$u = 0$
-1	$u < 0$

y en la matriz U y V , se colocan los resultados de $signos(X_i - X_j)$ y $signo(Y_i - Y_j)$ respectivamente, como se denota a continuación:

Matriz U:

-	X_1	X_2	\dots	X_n
X_1	0	$signo(X_1 - X_2)$	\dots	$signo(X_1 - X_n)$
X_2	$signo(X_2 - X_1)$	0	\dots	$signo(X_2 - X_n)$
:	:	:	:	:
X_n	$signo(X_n - X_1)$	$signo(X_n - X_2)$	\dots	0

Matriz V:

-	Y_1	Y_2	\dots	Y_n
Y_1	0	$signo(Y_1 - Y_2)$	\dots	$signo(Y_1 - Y_n)$
Y_2	$signo(Y_2 - Y_1)$	0	\dots	$signo(Y_2 - Y_n)$
:	:	:	:	:
Y_n	$signo(Y_n - Y_1)$	$signo(Y_n - Y_2)$	\dots	0

A partir de U y V se construye la matriz A de tamaño $n \times n$, donde $a_{ij} = signos(X_i - X_j) \hat{u} signo(Y_i - Y_j)$ como se denota a continuación:

-	a_1	a_2	\dots	a_n
a_1	0	a_{12}	\dots	a_{1n}
a_2	a_{21}	0	\dots	a_{2n}
:	:	:	:	:
a_n	a_{n1}	a_{n2}	\dots	0

$$\frac{1}{2n(n-1)} \left[\sum u(u-1) \right] \left[\sum v(v-1) \right]$$

La estadística Z tiene una distribución normal estándar. Finalmente, la estadística Z (con corrección de continuidad) es igual a:

$$Z = \frac{|S| - 1}{\sqrt{\text{Var}(S|H_0)}} \sim N(0, 1)$$

Aplicación

Caso: Se quiere comparar el resultados de dos escalas que miden el coeficiente intelectual. Las dos escalas toman valores de 0 a 100 puntos, donde a mayores valores mayor nivel de inteligencia. Se aplicaron las dos escalas en 13 personas obteniendo los siguientes resultados: Escala 1: {83 78 79 81 68 79 89 90 83 91 86 89 77} Escala 2: {85 76 89 78 90 77 69 77 85 89 82 78 77}

Para obtener el coeficiente τ de Kendall se crean dos vectores con las observaciones respectiva de cada escala, y se puede utilizar el paquete `Kendall` (McLeod 2011).

```
library(Kendall)
escala1<-c(83,78,79,81,68,79,89,90,83,91,86,89,77)
escala2<-c(85,76,89,78,90,77,69,77,85,89,82,78,77)
```

utilizando el comando `Kendall` se obtiene el coeficiente τ_b , la estadística S (score), el denominador ajustado por empates para el cálculo de τ y el valor p para probar la hipótesis $H_0 : \tau = 0$ con base a la estadística S (algunos programas la denomina **score**) incluyendo la corrección por continuidad.

```
Kendall(escala1,escala2)

## tau = -0.0544, 2-sided pvalue =0.85219

summary(Kendall(escala1,escala2))

## Score = -4 , Var(Score) = 259.2308
## denominator = 73.4847
## tau = -0.0544, 2-sided pvalue =0.85219
```

Otro comando que se puede utilizar para obtener el coeficiente τ de Kendall es `cor.test`. A partir de los dos vectores creados anteriormente, el comando permite obtener el valor de τ ajustado por empates, el valor p bajo la distribución exacta para probar H_0 cuando no hay presencia de empates, e incluir o no la corrección por continuidad para la estadística

$$Z = \frac{|S|}{\sqrt{\text{Var}(S|H_0)}} \sim N(0, 1)$$

La ejecución de este comando se presenta a continuación.

```
cor.test(escala1,escala2,method ="kendall", exact = TRUE,continuity = FALSE)

## Warning in cor.test.default(escala1, escala2, method = "kendall", exact =
## TRUE, : Cannot compute exact p-value with ties

##
## Kendall's rank correlation tau
##
## data:  escala1 and escala2
## z = -0.24844, p-value = 0.8038
```

```
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## -0.05443311
```

```
cor.test(escala1,escala2,method ="kendall", exact = FALSE,continuity = FALSE)
```

```
##
## Kendall's rank correlation tau
##
## data:  escala1 and escala2
## z = -0.24844, p-value = 0.8038
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## -0.05443311
```

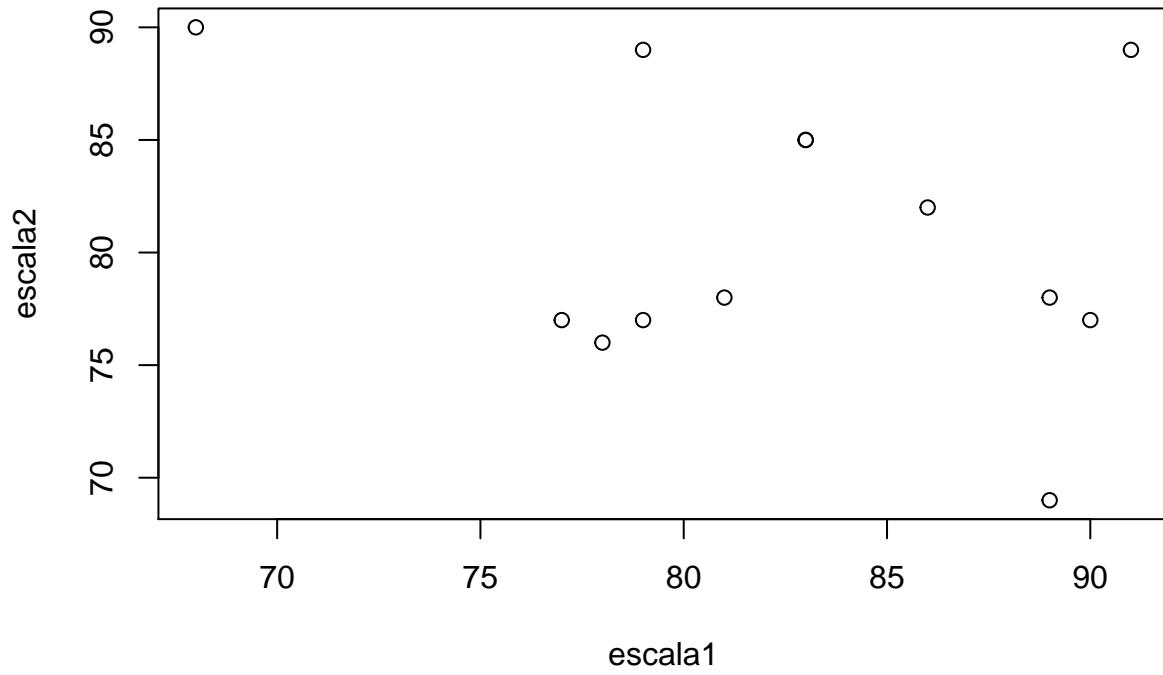
```
cor.test(escala1,escala2,method ="kendall", exact = FALSE,continuity = TRUE)
```

```
##
## Kendall's rank correlation tau
##
## data:  escala1 and escala2
## z = -0.18633, p-value = 0.8522
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## -0.05443311
```

En ocasiones, un gráfico de dispersión aporta en la descripción del tipo de relación que se presenta entre las dos variables.

```
plot(escala1,escala2, main = "Gráfico de dispersión entre la escala 1
y la escala 2")
```

Gráfico de dispersión entre la escala 1 y la escala 2



Coeficiente de correlación de rangos de Spearman

Al igual que el coeficiente τ de Kendall, el coeficiente de correlación de rangos de Spearman cumple con los siete criterios enunciados en la sección **Medidas de asociación**.

- Cálculo del coeficiente.

Para una muestra de n parejas de la forma:

$$(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$$

el coeficiente de correlación de rangos de Spearman (ρ) se puede estimar a partir de la siguiente expresión:

$$\hat{\rho} = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{[\sum_{i=1}^n (R_i - \bar{R})^2 \sum_{i=1}^n (S_i - \bar{S})^2]^{1/2}}$$

donde R_i y S_i corresponde respectivamente a los rangos asignados de las variables X_i y Y_i cuando estas han sido ordenada de menor a mayor, y $\bar{R} = \bar{S} = \frac{n+1}{2}$.

Simplificando la expresión anterior, se presentan dos fórmulas para obtener $\hat{\rho}$:

1.

$$\hat{\rho} = \frac{12 \sum_{i=1}^n R_i S_i}{n(n^2 - 1)} - \frac{3(n+1)}{(n-1)}$$

2.

$$\hat{\rho} = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)}$$

donde $D_i = R_i - S_i$

- Hipótesis.

Hipótesis nula: el coeficiente ρ es igual a cero.

$$H_0 : \rho = 0$$

Para evaluar la hipótesis nula, se puede utilizar la estadística que se presenta a continuación:

$$Z = \rho \sqrt[3]{n-1}$$

que tiene una distribución normal estándar. Programas como Stata y R evalúan la hipótesis nula a partir de la estadística:

$$t = \frac{|\rho| \sqrt[3]{n-2}}{[1 - \rho^2]^{1/2}}$$

que tiene una distribución t -Student con $n - 2$ grados de libertad.

- Intervalo de confianza

Una forma de obtener un intervalo de confianza para estimar ρ es utilizando la transformación z de fisher. Primero se construye z , así:

$$z_{fisher}(\hat{\rho}) = z = 0.5 * \ln\left(\frac{1 + \hat{\rho}}{1 - \hat{\rho}}\right)$$

y sus respectivos intervalos de confianza $(1 - \frac{\alpha}{2})$ de la siguiente forma (supuesto de normalidad):

$$LI_z = z - z_{1-\alpha/2} * \frac{1}{\sqrt{n-3}}$$

y

$$LS_z = z + z_{1-\alpha/2} * \frac{1}{\sqrt{n-3}}$$

Con los resultados de los dos límites anteriores se construyen los límites para un intervalo de confianza para ρ , así:

$$LI_\rho = \frac{e^{2LI_z} - 1}{e^{2LI_z} + 1}$$

y

$$LS_\rho = \frac{e^{2LS_z} - 1}{e^{2LS_z} + 1}$$

De igual forma, basado en la transformación z de fisher, se pueden evaluar pruebas de hipótesis (bivariadas y univariadas) para distintos valores de $\rho = 0$. Por ejemplo, para la hipótesis nula

$$H_0 : \rho = \rho_0$$

el estadístico de prueba es igual a:

$$\frac{z_{fisher}(\hat{\rho}) - z_{fisher}(\hat{\rho}_0)}{\sqrt{1/(n-3)}} \sim N(0, 1)$$

- Presencia de empates.

Ante la presencia de empates ya sea en la variable X o en la variable Y , el rango que se asigna corresponde al promedio de los rangos originales como se ha trabajado en las pruebas anteriores, y al aplicar la primera fórmula que se presentó en esta sección, se obtiene el cálculo de ρ ajustado por empates. Por el contrario las dos últimas fórmulas ya no se pueden aplicar.

Otra expresión de ρ que incluye el ajuste por empates es la siguiente:

$$\rho = \frac{12 \sum_{i=1}^n R_i S_i - n(n+1)^2/4}{\sqrt{[n(n^2-1) - 12t_x][n(n^2-1) - 12u_y]}}$$

donde $t_x = \frac{\sum t(t^2-1)}{12}$ y $u_y = \frac{\sum u(u^2-1)}{12}$ siendo t y u el número de empates en la variable X y en la variable Y respectivamente, realizando las sumatorias sobre el número de veces que se presentaron los empates.

La evaluación de la hipótesis nula con la presencia de empates se realiza a partir de la estadística Z o la estadística t reemplazando por el ρ ajustado por empates.

Aplicación

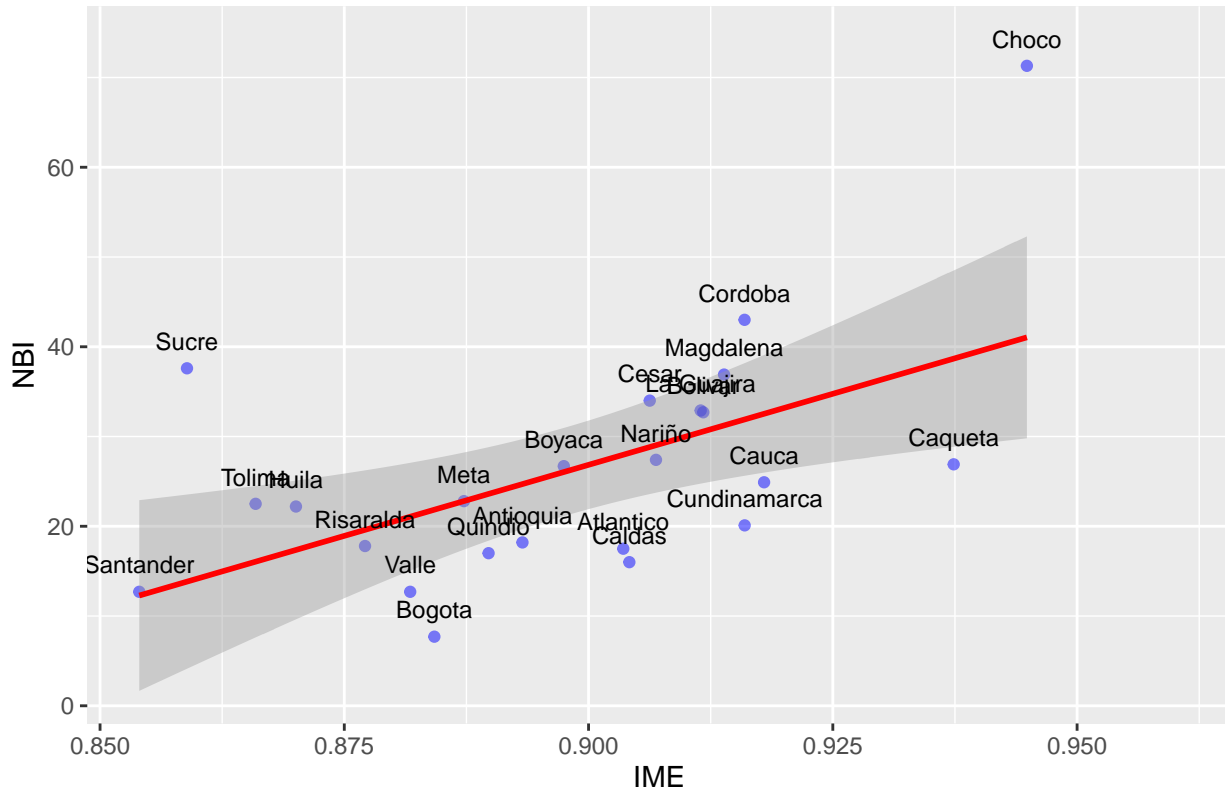
Caso: Se buscó evaluar la relación entre los indicadores: inequidad de la mortalidad evitable (IME), necesidades básicas insatisfechas (NBI) y línea de pobreza (LP) evaluados en los departamentos de Colombia.

Para obtener el coeficiente ρ , primero cargamos la información, obtenemos un gráfico de dispersión y ejecutamos el comando `cor.test` con la opción `method="spearman"`.

```
datos<-import("datos.xlsx")

ggplot(data=datos, aes(x=IME,y=NBI))+
  geom_point(alpha=0.5,color="blue")+
  geom_smooth(color="red",method = lm,se=TRUE)+
  labs(title="IME vs. NBI")+
  geom_text(aes(label=Departamentos), size=3,nudge_y = 3)
```

IME vs. NBI



```
cor.test(datos$IME,datos$NBI,method = "spearman",exact = FALSE,
  continuity = FALSE)
```

```
##
## Spearman's rank correlation rho
##
## data: datos$IME and datos$NBI
## S = 954.74, p-value = 0.009561
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
## rho
## 0.5282926
```

Para obtener el coeficiente de correlación para las 3 $\binom{3}{2}$ posibles parejas de indicadores a comparar, tenemos:

```
#install.packages("corrplot")
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
ind<-transmute(datos,IME,NBI,LP)
```

```
rho<-cor(ind,use = "pairwise.complete.obs", method = "spearman");rho
```

```
##           IME      NBI      LP
## IME 1.0000000 0.5282926 0.2816206
## NBI 0.5282926 1.0000000 0.6389919
## LP  0.2816206 0.6389919 1.0000000
```

```
corrplot(rho,method = "number",title = "",tl.cex = 0.8)
```



```
#method = c("circle", "square", "ellipse", "number", "shade", "color", "pie")
```

El cálculo del intervalo de confianza se puede obtener de las siguientes dos formas:

```
p<-cor.test(datos$IME,datos$NBI,method = "spearman",exact = FALSE,
            continuity = FALSE)
```

```
# a mano:
```

```
# Transformación z (Fisher)
```

```
z=0.5*log((1+p$estimate)/(1-p$estimate))
```

```
Z1<-z-qnorm(0.975)*sqrt(1/20)
```

```

Z2<-z+qnorm(0.975)*sqrt(1/20)

LI<-(exp(2*Z1)-1)/(exp(2*Z1)+1)
LS<-(exp(2*Z2)-1)/(exp(2*Z2)+1)

cat("IC95%=",LI,LS)

## IC95%= 0.1484083 0.7723132
# 0 utilizando la siguiente función
#install.packages("DescTools")

library(DescTools)
SpearmanRho(ind$IME,ind$NBI,use="pairwise.complete.obs",conf.level = 0.95)

##      rho      lwr.ci      upr.ci
## 0.5282926 0.1484083 0.7723132
# Una opción por bootstrap:
#library(RVAideMemoire)
#spearman.ci(ind$IME,ind$NBI,nrep = 1000, conf.level = 0.95)

```

Evaluar una prueba de hipótesis, por ejemplo que $H_0 : \rho \leq 0.2$, tenemos:

```

# H0: rho <= 0.2 vs Ha: rho > 0.2

z0=0.5*log((1+0.2)/(1-0.2))
E<-(z-z0)/sqrt(1/20)

valorp<-1-pnorm(E);valorp

##      rho
## 0.04253869

```

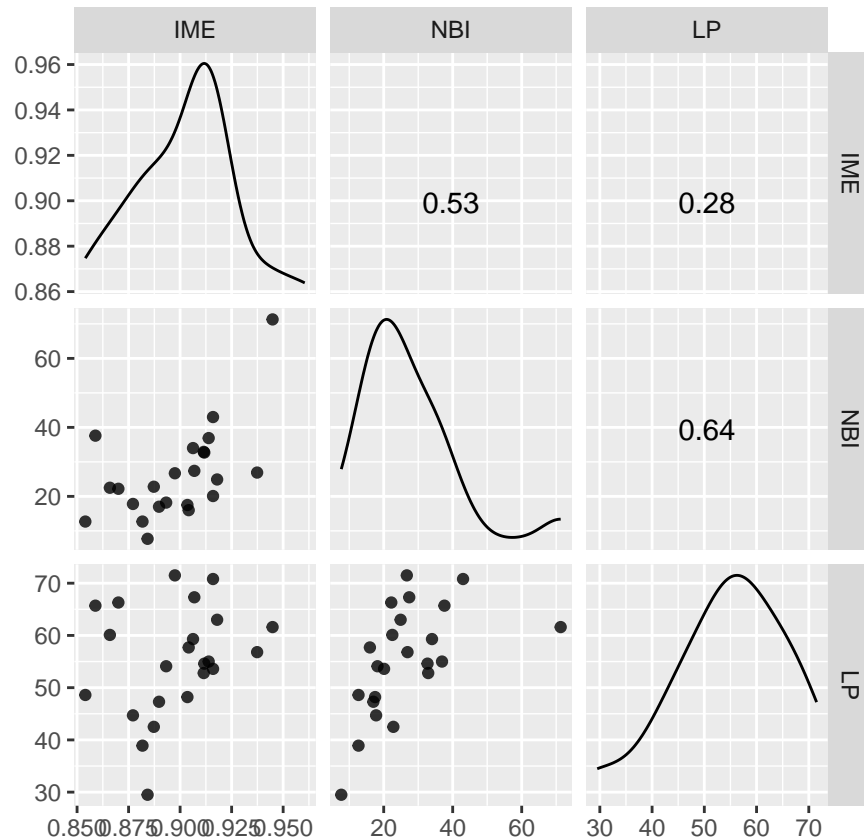
y ahora, una matriz con gráficos de dispersión, distribución de cada indicador y el coeficiente de correlación:

```

#install.packages("GGally")

library(GGally)
ggscatmat(datos, columns = 2:4, alpha=0.8,corMethod="spearman")

```



Finalmente, otra alternativa es utilizar el paquete `pspearman` (Savicky 2014) y el comando `spearman.test` como se presenta a continuación.

```
library(pspearman)
spearman.test(datos$IME,datos$NBI,alternative = "t",approximation = "exact")
```

```
## Warning in spearman.test(datos$IME, datos$NBI, alternative = "t",
## approximation = "exact"): Cannot compute exact p-values with ties
```

```
##
## Spearman's rank correlation rho
##
## data: datos$IME and datos$NBI
## S = 954.74, p-value = 0.01054
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.5282926
```

```
spearman.test(datos$IME,datos$NBI,alternative = "t",
              approximation = "t-distribution")
```

```
##
## Spearman's rank correlation rho
##
## data: datos$IME and datos$NBI
## S = 954.74, p-value = 0.009561
## alternative hypothesis: true rho is not equal to 0
```

```
## sample estimates:  
##      rho  
## 0.5282926
```

Prueba de independencia - Distribución χ^2

Sean A y B dos variables medidas en una escala nominal u ordinal, y r y c corresponden al número de categorías de respuesta de las dos variables respectivamente, los valores observados de estas dos variables en n sujetos se puede representar en una matriz de dimensiones $r \times c$ que se conoce como **tabla de contingencia de dos vías**. Esta tabla se puede expresar de la siguiente forma:

	b_1	b_2	\dots	b_c	total
a_1	x_{11}	x_{12}	\dots	x_{1c}	$x_{1.}$
a_2	x_{21}	x_{22}	\dots	x_{2c}	$x_{2.}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
a_r	x_{r1}	x_{r2}	\dots	x_{rc}	$x_{r.}$
Total	$x_{.1}$	$x_{.2}$	\dots	$x_{.c}$	n

donde x_{ij} es el número de individuos que presentaron la categoría i de la variable A y la categoría j de la variable B , $x_{i.}$ corresponde al número de individuos que presentaron la categoría i de la variable A y $x_{.j}$ es el número de individuos que presentaron la categoría j de la variables B .

- Hipótesis.

Hipótesis nula.

H_0 : A y B son independientes.

Hipótesis alternativa.

H_a : A y B se relacionan.

- Estadística de prueba.

Si se asume que la hipótesis nula es verdadera la probabilidad de observar un sujeto que presenta la categoría i de la variable A y la categoría j de la variable B es igual a la probabilidad de que el sujeto presente la categoría i de la variable A multiplicado por la probabilidad de que el sujeto presente la categoría j de la variable B , es decir:

$$P(a_i \cap b_j) = P(a_i) \cdot P(b_j)$$

Con base en lo anterior, el número de sujetos que presentan la categoría i de la variable A y la categoría j de la variable B si la hipótesis nula es verdadera se conoce como la **frecuencia esperada** y es igual a

$$E_{ij} = n \cdot P(a_i \cap b_j) = n [P(a_i) \cdot P(b_j)] = n \frac{x_{i.}}{n} \frac{x_{.j}}{n} = \frac{x_{i.} \cdot x_{.j}}{n}$$

Si se denomina la frecuencia observada, sea O_{ij} , como el número de sujetos que presentaron la categoría i de la variable A y la categoría j de la variable B , el estadístico de prueba para evaluar H_0 se basa en la comparación entre E_{ij} y O_{ij} . La expresión es:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

que tiene una distribución χ^2 con $(r - 1)(c - 1)$ grados de libertad.

- Test Exacto de Fisher.

Cuando el número de observaciones es muy pequeño se debe utilizar el test exacto de Fisher para obtener el valor exacto de p. La regla más utilizada para determinar que el número de observaciones es bajo es la siguiente: *Si el 20% o más de las celdas de la tabla a dos vías presentan frecuencias esperadas menores a 5 (Cochran).*

- Coeficientes de asociación.

Si la estadística χ^2 se acerca a cero indica que las frecuencias observadas se acercan a las frecuencias esperadas es decir se apoyan la hipótesis nula, mientras que si la estadística χ^2 toma valores muy grandes, indican que las frecuencias observadas se aleja mucho de las frecuencias esperadas por lo tanto se apoya la hipótesis alternativa. Por lo anterior, se puede utilizar la estadística χ^2 como una medida del grado de asociación entre las dos variables. Se han propuesto varios coeficientes basados en la estadística χ^2 los cuales se definen a continuación:

1. Coeficiente de contingencia: $C = \sqrt{\frac{\chi^2}{\chi^2 + n}}$
2. Razón C/C_{\max} : $\frac{C}{\sqrt{\frac{t-1}{t}}}$ donde $t = \min(r, c)$
3. Coeficiente φ : $\varphi = \sqrt{\chi^2/n}$
4. Cramer V: $V = \sqrt{\frac{\chi^2}{n(t-1)}}$ donde $t = \min(r, c)$

Estos indicadores toman valores entre 0 y 1, donde a mayor valor mayor grado de asociación.

Aplicación

Caso: En 141 adultos se evaluó la variable nivel educativo y problemas de alcohol. Nivel educativo se categorizó en: Ninguno/Primaria, Secundaria, Técnico/Tecnólogo y Universitario, mientras que problemas de alcohol se categorizó en: probable dependencia al alcohol, bebedores en riesgo, consumo excesivo y ninguno de los anteriores.

Para evaluar la hipótesis de independencia primero se cargan la información en *R*, y se pueden utilizar dos comandos `chi.test` y `fisher.test`

```
adultos<-import("independencia.xlsx")
chisq.test(adultos$educacion,adultos$alcohol)
```

```
## Warning in chisq.test(adultos$educacion, adultos$alcohol): Chi-squared
## approximation may be incorrect
```

```
##
## Pearson's Chi-squared test
##
## data: adultos$educacion and adultos$alcohol
## X-squared = 5.5083, df = 9, p-value = 0.7879
```

```
fisher.test(adultos$educacion,adultos$alcohol)
```

```
##
## Fisher's Exact Test for Count Data
##
## data: adultos$educacion and adultos$alcohol
## p-value = 0.6334
## alternative hypothesis: two.sided
```

Para obtener el coeficiente V de Cramer, se puede utilizar el comando `cramersV` del paquete `{lsr}` (Navarro 2015).


```
library(lsr)
cramersV(adultos$educacion,adultos$alcohol)
```

```
## Warning in chisq.test(...): Chi-squared approximation may be incorrect
## [1] 0.1141141
```

Otro comando que permite obtener la prueba de independencia es `CrossTable` y se encuentra en el paquete `gmodels` (Warnes et al. 2015). Con este comando se pueden obtener valores adicionales de la tabla de contingencia como se muestra a continuación.

```
library(gmodels)
```

```
## Registered S3 method overwritten by 'gdata':
## method from
## reorder.factor DescTools
```

```
CrossTable(adultos$educacion,adultos$alcohol,expected = TRUE,
            prop.r = TRUE,prop.c = TRUE,fisher = TRUE)
```

```
## Warning in chisq.test(t, correct = FALSE, ...): Chi-squared approximation
## may be incorrect
```

```
##
##
## Cell Contents
## |-----|
## |                N |
## |      Expected N |
## | Chi-square contribution |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
```

```
##
##
## Total Observations in Table: 141
##
```

```
##          | adultos$alcohol
## adultos$educacion | Beb ries | Cons Exc | Ning | Prob dep | Row Total |
## -----|-----|-----|-----|-----|-----|
##      Nin_Pri |      7 |      1 |    33 |      1 |      42 |
##              |  5.957 |  4.468 | 30.681 |  0.894 |          |
##              |  0.182 |  2.692 |  0.175 |  0.013 |          |
##              |  0.167 |  0.024 |  0.786 |  0.024 |      0.298 |
##              |  0.350 |  0.067 |  0.320 |  0.333 |          |
##              |  0.050 |  0.007 |  0.234 |  0.007 |          |
## -----|-----|-----|-----|-----|
##          Sec |     10 |     12 |     55 |      2 |      79 |
##              | 11.206 |  8.404 | 57.709 |  1.681 |          |
##              |  0.130 |  1.538 |  0.127 |  0.061 |          |
##              |  0.127 |  0.152 |  0.696 |  0.025 |      0.560 |
##              |  0.500 |  0.800 |  0.534 |  0.667 |          |
##              |  0.071 |  0.085 |  0.390 |  0.014 |          |
## -----|-----|-----|-----|-----|
##          Tec |      2 |      1 |      9 |      0 |      12 |
```

```

##          |      1.702 |      1.277 |      8.766 |      0.255 |          |
##          |      0.052 |      0.060 |      0.006 |      0.255 |          |
##          |      0.167 |      0.083 |      0.750 |      0.000 |      0.085 |
##          |      0.100 |      0.067 |      0.087 |      0.000 |          |
##          |      0.014 |      0.007 |      0.064 |      0.000 |          |
## -----|-----|-----|-----|-----|-----|
##          Univ |          1 |          1 |          6 |          0 |          8 |
##          |      1.135 |      0.851 |      5.844 |      0.170 |          |
##          |      0.016 |      0.026 |      0.004 |      0.170 |          |
##          |      0.125 |      0.125 |      0.750 |      0.000 |      0.057 |
##          |      0.050 |      0.067 |      0.058 |      0.000 |          |
##          |      0.007 |      0.007 |      0.043 |      0.000 |          |
## -----|-----|-----|-----|-----|
##      Column Total |          20 |          15 |          103 |          3 |          141 |
##          |      0.142 |      0.106 |      0.730 |      0.021 |          |
## -----|-----|-----|-----|-----|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## -----
## Chi^2 =  5.50832      d.f. =  9      p =  0.7879391
##
##
## Fisher's Exact Test for Count Data
## -----
## Alternative hypothesis: two.sided
## p =  0.6334377
##
##

```

Prueba de Homogeneidad - Distribución χ^2

En la prueba de independencia se selecciona una muestra de tamaño n y se miden dos características para observar si estas se encuentran relacionadas entre sí o se consideran independientes. En ese escenario, el número total de individuos que se obtienen por categorías de cada variable es aleatorio, es decir que no dependen del investigador. Un escenario distinto, es cuando los investigadores definen un número fijo de sujetos a seleccionar de cada grupo o población de interés, y se busca evaluar si la distribución en cada grupo es homogénea en relación a una segunda variable. Esta situación corresponde a la aplicación de la estadística χ^2 como prueba de homogeneidad.

- Hipótesis:

Hipótesis nula.

H_0 : Los grupos se distribuyen igual con relación a la variable de interés.

- Estadística de prueba.

La evaluación de la prueba de homogeneidad utiliza el estadístico χ^2 visto en la sección anterior. La expresión es:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

que tiene una distribución χ^2 con $(r - 1)(c - 1)$ grados de libertad, donde r es el número de grupos seleccionados y c es el número de categorías de la variable de interés. Al igual que en la prueba de independencia, si más del 20% de las celdas presentan frecuencias esperadas menores a 5, se utiliza el test exacta de Fisher.

Aplicación

Caso: Con el objetivo de evaluar si el consumo de alcohol es diferente entre las poblaciones de indígenas, afrodescendientes y blancos, se seleccionaron 100 sujetos de cada población y se clasificaron a partir de una escala en: probable dependencia al alcohol, bebedores en riesgo, consumo excesivo y ninguno de los anteriores.

Los comandos utilizando para evaluar la prueba de independencia aplican directamente para evaluar la prueba de homogeneidad. Los comandos son `chi.test` y `fisher.test` y su aplicación se presenta a continuación.

```
alcohol<-import("homogeneidad.xlsx")
chisq.test(alcohol$etnia,alcohol$audit)
```

```
## Warning in chisq.test(alcohol$etnia, alcohol$audit): Chi-squared
## approximation may be incorrect
```

```
##
## Pearson's Chi-squared test
##
## data: alcohol$etnia and alcohol$audit
## X-squared = 1.5651, df = 6, p-value = 0.9551
```

```
fisher.test(alcohol$etnia,alcohol$audit)
```

```
##
## Fisher's Exact Test for Count Data
##
```

```
## data: alcohol$etnia and alcohol$audit
## p-value = 0.9533
## alternative hypothesis: two.sided
```

De igual forma, la aplicación del comando `CrossTable` (Warnes et al. 2015) para este caso es la siguiente.

```
library(gmodels)
CrossTable(alcohol$audit,alcohol$etnia, expected = TRUE,
           prop.r = TRUE,prop.c = TRUE,fisher = TRUE)
```

```
## Warning in chisq.test(t, correct = FALSE, ...): Chi-squared approximation
## may be incorrect
```

```
##
##      Cell Contents
## |-----|
## |                N |
## |      Expected N |
## | Chi-square contribution |
## |      N / Row Total |
## |      N / Col Total |
## |      N / Table Total |
## |-----|
##
##
## Total Observations in Table: 300
##
##
##      | alcohol$etnia
## alcohol$audit |      Afro |      Blan |      Ind | Row Total |
## -----|-----|-----|-----|-----|
##      Beb ries |      13 |      13 |      9 |      35 |
##      |      11.667 |      11.667 |      11.667 |      |
##      |      0.152 |      0.152 |      0.610 |      |
##      |      0.371 |      0.371 |      0.257 |      0.117 |
##      |      0.130 |      0.130 |      0.090 |      |
##      |      0.043 |      0.043 |      0.030 |      |
## -----|-----|-----|-----|
##      Cons Exc |      9 |      10 |      10 |      29 |
##      |      9.667 |      9.667 |      9.667 |      |
##      |      0.046 |      0.011 |      0.011 |      |
##      |      0.310 |      0.345 |      0.345 |      0.097 |
##      |      0.090 |      0.100 |      0.100 |      |
##      |      0.030 |      0.033 |      0.033 |      |
## -----|-----|-----|-----|
##      Ning |      76 |      75 |      80 |      231 |
##      |      77.000 |      77.000 |      77.000 |      |
##      |      0.013 |      0.052 |      0.117 |      |
##      |      0.329 |      0.325 |      0.346 |      0.770 |
##      |      0.760 |      0.750 |      0.800 |      |
##      |      0.253 |      0.250 |      0.267 |      |
## -----|-----|-----|-----|
##      Prob dep |      2 |      2 |      1 |      5 |
##      |      1.667 |      1.667 |      1.667 |      |
##      |      0.067 |      0.067 |      0.267 |      |
```

```

##          |      0.400 |      0.400 |      0.200 |      0.017 |
##          |      0.020 |      0.020 |      0.010 |           |
##          |      0.007 |      0.007 |      0.003 |           |
## -----|-----|-----|-----|-----|
## Column Total |      100 |      100 |      100 |      300 |
##          |      0.333 |      0.333 |      0.333 |           |
## -----|-----|-----|-----|-----|
##
##
## Statistics for All Table Factors
##
##
## Pearson's Chi-squared test
## -----
## Chi^2 = 1.565069    d.f. = 6    p = 0.9550557
##
##
## Fisher's Exact Test for Count Data
## -----
## Alternative hypothesis: two.sided
## p = 0.9532724
##
##

```

Prueba de bondad y ajuste - Distribución χ^2

Una aplicación más que tiene la distribución χ^2 es evaluar si una muestra proviene de una población con una distribución de probabilidades específica, por ejemplo: normal, binomial o poisson. En esta sección describiremos como realizar esta aplicación para la distribución normal.

- Hipótesis.

Hipótesis nula.

H_0 : La variable proviene de una población con distribución normal.

- Estadística de prueba.

Se parte de clasificar los valores de la variable de interés en r categorías de tal forma que cada valor observado de la variable X_i se puede clasificar en una de estas categorías obteniendo una tabla de frecuencias observadas como la que se presenta a continuación:

Grupos	Frecuencias
$G_1 : I_1 - S_1$	O_1
$G_2 : I_2 - S_2$	O_2
:	:
$G_r : L_r - S_r$	O_r

Ahora, asumiendo que la hipótesis nula es verdadera, es decir que la población de la cual se extrajo la variable tiene una distribución normal, se puede estimar el número de sujetos (frecuencias esperadas) en cada una de las categorías. Para ello, primero se estima la media y la desviación estándar, sean \bar{X} y s , y con base en la función de distribución acumulada de la distribución normal $\phi(X)$, se calcula la probabilidad de que un sujeto caiga en cada categoría, es decir:

$$P(I_i < X < S_i) = \phi(S_i, \bar{x}, s) - \phi(I_i, \bar{x}, s)$$

La frecuencia esperada de la categoría i , E_i , es igual a la probabilidad de estar en la categoría multiplicado por n . Incluyendo la frecuencia esperada, la tabla anterior se puede completar de la siguiente forma:

Grupos	O_i	E_i
$G_1 : I_1 - S_1$	O_1	E_1
$G_2 : I_2 - S_2$	O_2	E_2
:	:	:
$G_r : L_r - S_r$	O_r	E_r

Con base en la tabla anterior, el estadístico de prueba es igual a:

$$\chi^2 = \sum_{i=1}^r \frac{(O_i - E_i)^2}{E_i}$$

que tiene una distribución $\chi^2_{(r-3)}$.

Si se conoce uno de los dos parámetros de la distribución normal que se define en la hipótesis nula, los grados de libertad son iguales a $r - 2$, y si se conocen los dos parámetros, los grados de libertad son iguales a $r - 1$.

Aplicación

Caso: Se registró la información de la estatura de 157 mujeres de distintas edades. Se quiere evaluar si esta variable proviene de una población con distribución normal.

Para evaluar la prueba de bondad y ajuste utilizando la estadística χ^2 , primero se carga la información de la estatura de las 157 mujeres, y se estiman los dos parámetros.

```
estatura<-import("bondadyajuste.xlsx")
media<-mean(estatura$Talla)
desv<-sd(estatura$Talla)
media
```

```
## [1] 84.4293
```

```
desv
```

```
## [1] 11.70881
```

Ahora, se definen los puntos de corte para construir las categorías, y se calculan las frecuencias observadas en cada categoría

```
talla.cut<-cut(estatura$Talla,breaks = c(0,70,80,90,100,111)) # Puntos de corte
table(talla.cut)
```

```
## talla.cut
## (0,70] (70,80] (80,90] (90,100] (100,111]
##      25      31      42      47      12
```

```
obs<-vector()
for(i in 1:5) obs[i]<-table(talla.cut)[[i]] # Frecuencias observadas
obs
```

```
## [1] 25 31 42 47 12
```

A continuación se obtienen las frecuencias esperadas

```
f1<-157*(pnorm(70,media,desv))
f2<-157*(pnorm(80,media,desv)-pnorm(70,media,desv))
f3<-157*(pnorm(90,media,desv)-pnorm(80,media,desv))
f4<-157*(pnorm(100,media,desv)-pnorm(90,media,desv))
f5<-157*(1-pnorm(100,media,desv))
esp<-c(f1,f2,f3,f4,f5)#definir frecuencias esperadas.
esp
```

```
## [1] 17.09887 38.26065 51.85278 35.37706 14.41063
```

A partir de las frecuencias observadas y las frecuencias esperadas se calcula el estadístico de prueba y su valor p asociado.

```
X2<-sum(((obs-esp)^2)/esp) # chi^2
X2
```

```
## [1] 11.12291
```

```
g1<-5-3 # grados de libertad
1-pchisq(X2,g1) # valor p
```

```
## [1] 0.003843174
```

Bootstrap (en construcción)

```
costos<-c(30.6,27.6,69.5,73.3,22.0,238.5,40.5,110.7,200.9,55.9,147.3,63.0,21.0,
          79.9,29.7,50.3,129.3,107.5,42.4,72.6,146.5,22.2,22.6,54.6,88.8,89.1,
          162.5,41.5,45.7,31.6,185.9,50.8,75.9,44.7)

data<-data.frame(costos)

# Gráficos:

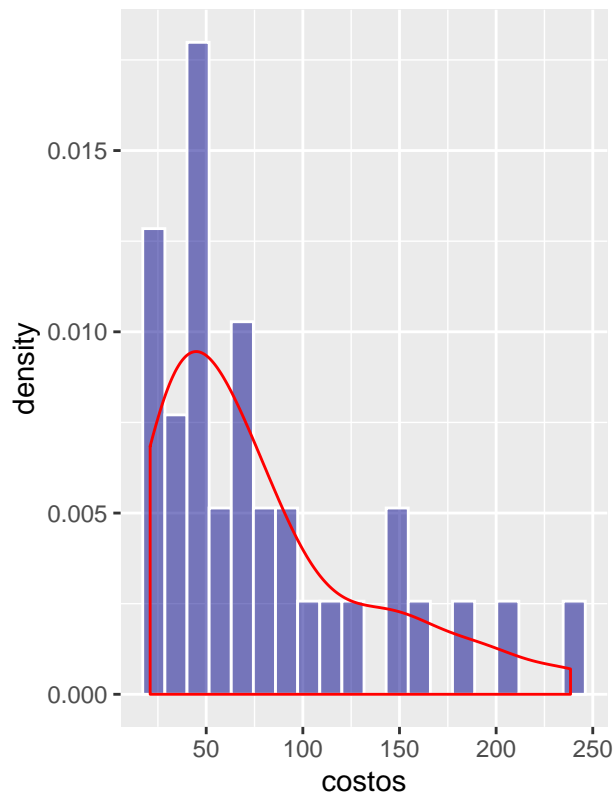
g1<-ggplot(data,aes(x=costos))+
  geom_histogram(bins=20,aes(y = ..density..), alpha=0.5,
                fill="darkblue", color = "white")+
  geom_density(col="red",size=0.5)+
  labs(title="Histograma")

g2<-ggplot(data,aes(y=costos))+
  geom_boxplot(alpha=0.5,fill="darkblue", color="blue",outlier.color="red")+
  labs(title="Boxplot")

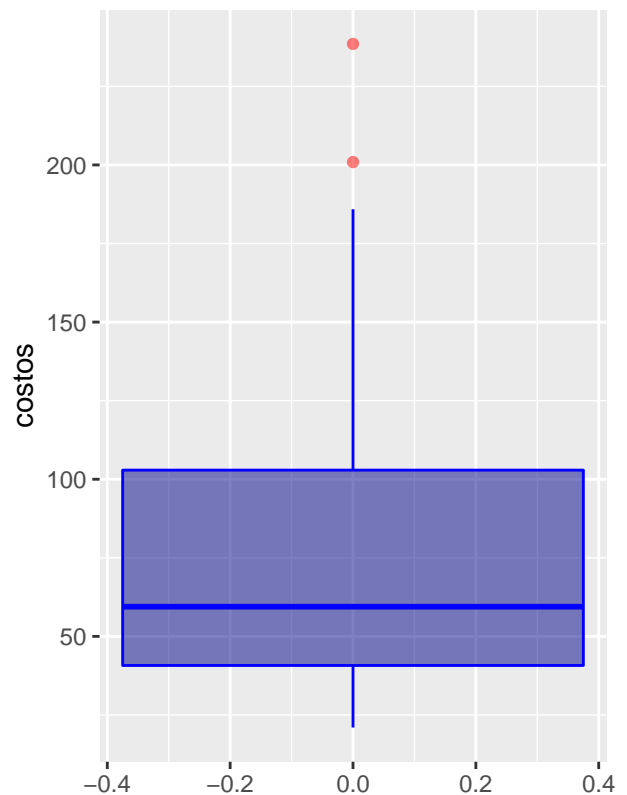
library(grid)

grid.newpage()
pushViewport(viewport(layout = grid.layout(1, 2)))
vlayout <- function(x, y) viewport(layout.pos.row = x, layout.pos.col = y)
print(g1, vp = vlayout(1, 1))
print(g2, vp = vlayout(1, 2))
```


Histograma



Boxplot



```
# Bootstrap IC
#install.packages("boot")
library(boot)
```

```
##
## Attaching package: 'boot'
## The following object is masked from 'package:lattice':
##
##   melanoma
B <- boot(data$costos,function(x,i) median(x[i]), R=1999)
print(B)
```

```
##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
## Call:
## boot(data = data$costos, statistic = function(x, i) median(x[i]),
##      R = 1999)
##
##
## Bootstrap Statistics :
##   original bias   std. error
## t1*    59.45 1.87066   10.19772
```

```
mean(B$t) # media
```

```
## [1] 61.32066
```

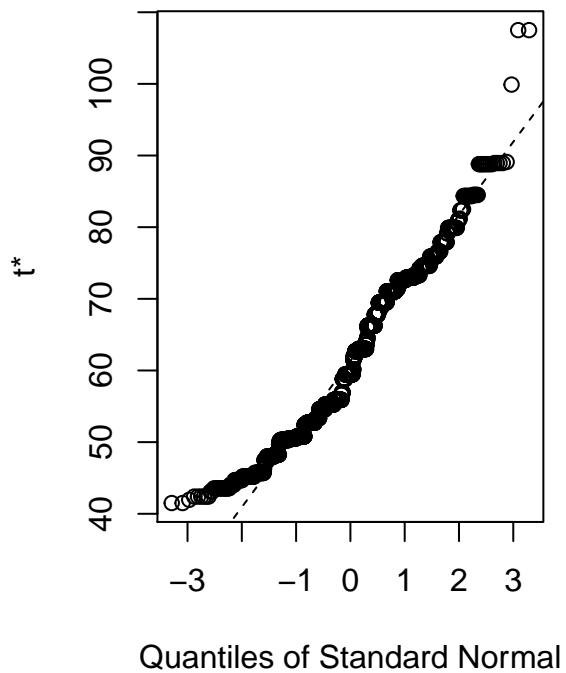
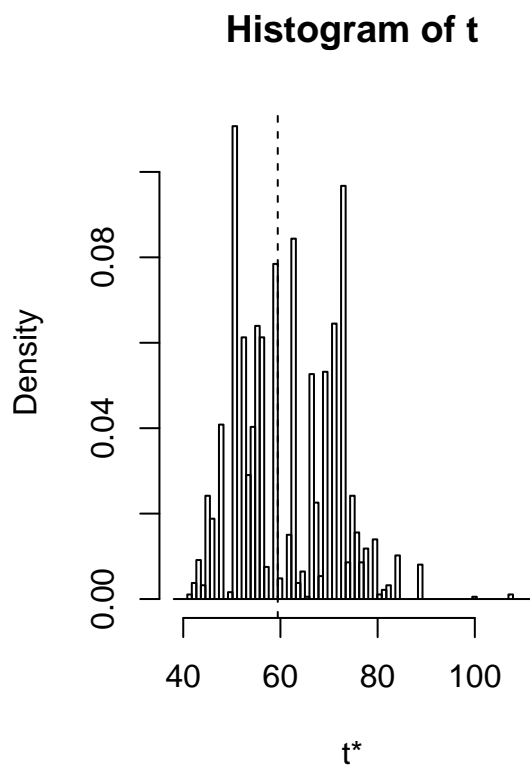
```
bias<-mean(B$t)-B$t0 # sesgo  
bias
```

```
## [1] 1.87066
```

```
se<-sd(B$t) # error estándar  
se
```

```
## [1] 10.19772
```

```
plot(B)
```



```
shapiro.test(B$t)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: B$t  
## W = 0.96283, p-value < 2.2e-16
```

```
# Normal
```

```
boot.ci(B,conf = 0.95,type = "norm")
```

```
## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS  
## Based on 1999 bootstrap replicates
```

```

##
## CALL :
## boot.ci(boot.out = B, conf = 0.95, type = "norm")
##
## Intervals :
## Level      Normal
## 95%      (37.59, 77.57 )
## Calculations and Intervals on Original Scale
LI<-B$t0-bias-1.96*se
LS<-B$t0-bias+1.96*se
cat(round(LI,2),round(LS,2))

## 37.59 77.57
# Basic
boot.ci(B,conf = 0.95,type = "basic")

## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 1999 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = B, conf = 0.95, type = "basic")
##
## Intervals :
## Level      Basic
## 95%      (39.0, 73.7 )
## Calculations and Intervals on Original Scale
quantile(B$t0-(B$t-B$t0),c(0.025,0.975))

## 2.5% 97.5%
## 39.0 73.7
# Percentiles
boot.ci(B,conf = 0.95,type = "perc")

## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 1999 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = B, conf = 0.95, type = "perc")
##
## Intervals :
## Level      Percentile
## 95%      (45.2, 79.9 )
## Calculations and Intervals on Original Scale
quantile(B$t,c(0.025,0.975))

## 2.5% 97.5%
## 45.2 79.9
# BC - bias corrected
boot.ci(B,conf = 0.95,type = "bca")

```

```

## BOOTSTRAP CONFIDENCE INTERVAL CALCULATIONS
## Based on 1999 bootstrap replicates
##
## CALL :
## boot.ci(boot.out = B, conf = 0.95, type = "bca")
##
## Intervals :
## Level      BCa
## 95%      (44.05, 77.90 )
## Calculations and Intervals on Original Scale

Best<-data.frame(B$t)
Best<-mutate(Best,p=ifelse(B.t<B$t0,1,0))
p<-table(Best$p)[2]

Best<-arrange(Best,B.t) # estimadores ordenados

k<-qnorm(p/B$R)

# Límite superior:

Q2<-(B$R+1)*pnorm(2*k+1.96)
a2<-floor(Q2)
b2<-a2+1

LSup<-Best[a2,1]+(qnorm(Q2/(B$R+1))-qnorm(a2/(B$R+1)))/
  (qnorm(b2/(B$R+1))-qnorm(a2/(B$R+1)))*(Best[b2,1]-Best[a2,1])

# Límite inferior:

Q1<-(B$R+1)*pnorm(2*k-1.96)
a1<-floor(Q1)
b1<-a1+1

LInf<-Best[a1,1]+(qnorm(Q1/(B$R+1))-qnorm(a1/(B$R+1)))/
  (qnorm(b1/(B$R+1))-qnorm(a1/(B$R+1)))*(Best[b1,1]-Best[a1,1])

cat(round(LInf,2),round(LSup,2))

## 44.05 77.9

```

Suavizamientos (en construcción)

```
ID<-import("ID.xlsx")

ggplot(ID,aes(x=M5,y=MM))+
  geom_point(col="blue",alpha=0.3)

#plot(ID$M5,ID$MM)
#lines(supsmu(ID$M5,ID$MM,span="cv"),col="red")

suav<-supsmu(ID$M5,ID$MM,span="cv")
id2<-merge(ID,suav,by.x="M5",by.y="x")
id2$e<-id2$MM-id2$y
plot(id2$y,id2$e)

#plot(id2$y,id2$e)
#lines(supsmu(id2$y,id2$e,span="cv"),col="red")

# loess
#install.packages("fANCOVA")

w<-function(x){(1-abs(x)^p)^p}
p<-1
plot(w,-1,1,col="red")

ggplot(ID,aes(x=M5,y=MM))+
  geom_point(col="blue",alpha=0.2)

library(fANCOVA)

loess.as(ID$M5,ID$MM,degree = 1,family ="symmetric",user.span = 0.5,plot = TRUE)
suav2<-loess.as(ID$M5,ID$MM,degree = 2,criterion = "gcv" ,
  family ="symmetric",plot = TRUE)

## criterion="gcv" o "aicc"
##family="gaussian","symmetric"

loess.as(suav2$fitted,suav2$residuals,degree = 1,criterion = "gcv",
  family = "symmetric",plot=TRUE)

# Kernel

install.packages("np")
library(np)

ID<-arrange(ID,M5)

# Distintos kernel: ckertype="gaussian", "epanechnikov" or "uniform".

kernel1<-npreg(bws=10,ID$M5,ID$MM,ckertype="epanechnikov",residuals=TRUE)
```

```
summary(kernel1)
plot(kernel1)
points(ID$M5, ID$MM, cex=.5, col="red")

# h bwmethod="cv.aic" or "cv.ls"

kernel2<-npreg(ID$M5, ID$MM, ckertype="epanechnikov", bwmethod="cv.ls")
summary(kernel2)
plot(kernel2)
```

Bibliografía

- Arnholt, Alan T. 2012. *BSDA: Basic Statistics and Data Analysis*.
<https://CRAN.R-project.org/package=BSDA>.
- Caeiro, Frederico, and Ayana Mateus. 2014. *Randtests: Testing Randomness in R*.
<https://CRAN.R-project.org/package=randtests>.
- Canavos. 1994. “Métodos No Paramétricos.” In *Probabilidad Y Estadística - Aplicaciones Y Metodos*, 575. México etc.: McGraw-Hill Companies.
- Conover, W. J. 1999. *Practical Nonparametric Statistics*. 3rd edition. TBS.
- Daniel, Wayne W. 2008. *Biostatistics: A Foundation for Analysis in the Health Sciences*. 9 edition. Hoboken, NJ: Wiley.
- Fay, Michael P. 2010. “Two-Sided Exact Tests and Matching Confidence Intervals for Discrete Data.” *R Journal* 2 (1): 53–58. <http://journal.r-project.org/>.
- Gibbons, Jean Dickinson Gibbons; Subhabrata Chakraborti; Gibbons Dickinson. 2003a. “Measures of Association for Bivariate Samples.” In *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*, 4 edition, 401. CRC Press.
- . 2003b. *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*. 4 edition. CRC Press.
- . 2003c. “One-Sample and Paired-Sample Procedures.” In *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*, 4 edition, 171. CRC Press.
- . 2003d. “One-Sample and Paired-Sample Procedures.” In *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*, 4 edition, 170. CRC Press.
- . 2003e. “One-Sample and Paired-Sample Procedures.” In *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*, 4 edition, 204. CRC Press.
- . 2003f. “Tests of Randomness.” In *Nonparametric Statistical Inference, Fourth Edition: Revised and Expanded*, 4 edition, 83. CRC Press.
- Gross, Juergen, and Uwe Ligges. 2015. *Nortest: Tests for Normality*.
<https://CRAN.R-project.org/package=nortest>.
- Hervé, Maxime. 2016. *RVAideMemoire: Diverse Basic Statistical and Graphical Functions*.
<https://CRAN.R-project.org/package=RVAideMemoire>.
- Hothorn, Torsten, and Kurt Hornik. 2015. *exactRankTests: Exact Distributions for Rank and Permutation Tests*. <https://CRAN.R-project.org/package=exactRankTests>.
- McLeod, A. I. 2011. *Kendall: Kendall Rank Correlation and Mann-Kendall Trend Test*.
<https://CRAN.R-project.org/package=Kendall>.
- Navarro, Daniel. 2015. *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners. (Version 0.5)*. Adelaide, Australia: University of Adelaide. <http://ua.edu.au/ccs/teaching/lsr>.
- Pohlert, Thorsten. 2014. *The Pairwise Multiple Comparison of Mean Ranks Package (PMCMR)*.
<http://CRAN.R-project.org/package=PMCMR>.
- Savicky, Petr. 2014. *Pspearman: Spearman’s Rank Correlation Test*.
<https://CRAN.R-project.org/package=pspearman>.
- Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. Fourth. New York: Springer.
<http://www.stats.ox.ac.uk/pub/MASS4>.
- Warnes, Gregory R., Ben Bolker, Thomas Lumley, Randall C. Johnson Contributions from Randall C. Johnson are Copyright SAIC-Frederick, Inc Funded by the Intramural Research Program, of the NIH,

National Cancer Institute, and Center for Cancer Research under NCI Contract NO1-CO-12400. 2015.
Gmodels: Various R Programming Tools for Model Fitting. <https://CRAN.R-project.org/package=gmodels>.