



**VNiVERSiDAD
DSALAMANCA**

CAMPUS DE EXCELENCIA INTERNACIONAL



FACULTAD DE CIENCIAS

GRADO EN ESTADÍSTICA

TRABAJO FIN DE GRADO

PROGRAMACIÓN Y DESARROLLO DE UNA INTERFAZ CON EL PAQUETE “SHINY” DE R PARA HERRAMIENTAS DE ESTADÍSTICA NO PARAMÉTRICA

Autor: Alfredo del Río Moldes

Tutor: Dr. José Manuel Sánchez Santos



VNIVERSIDAD
D SALAMANCA
CAMPUS DE EXCELENCIA INTERNACIONAL



800 AÑOS
VNIVERSIDAD
D SALAMANCA

FACULTAD DE CIENCIAS

GRADO EN ESTADÍSTICA

TRABAJO FIN DE GRADO

PROGRAMACIÓN Y DESARROLLO DE
UNA INTERFAZ CON EL PAQUETE
“SHINY” DE R PARA HERRAMIENTAS
DE ESTADÍSTICA NO PARAMÉTRICA

Autor: Alfredo del Río Moldes

Tutor: Dr. José Manuel Sánchez Santos

Salamanca, 2021

ÍNDICE DE CONTENIDOS

1.	INTRODUCCIÓN	1
2.	MATERIAL Y METODOS	3
2.1	Inferencia no paramétrica	3
2.1.1	Conceptos previos	3
2.1.2	Pruebas no paramétricas de aleatoriedad	5
2.1.3	Pruebas no paramétricas de bondad de ajuste	6
2.1.4	Pruebas no paramétricas de localización de una muestra (Mediana).	8
2.1.5	Pruebas no paramétricas para dos muestras	9
2.1.6	Pruebas no paramétricas para k muestras	11
2.1.7	Pruebas no paramétricas de independencia.	13
2.2	R y RStudio	15
2.3	Shiny	16
2.3.1	Reactividad.....	16
2.3.2	Inputs y Outputs	18
2.3.3	Interfaz de usuario	19
2.3.4	Servidor	21
2.4	Shinyapps.io	22
3.	RESULTADOS.....	22
3.1	App desarrollada.....	22
3.1.1	Esquema general.....	22
3.1.2	Inicio	23
3.1.3	Aprende.....	24
3.1.4	Datos	35
3.1.5	Análisis.....	37
3.2	Shinyapps.io	45
4.	CONCLUSIONES	46
5.	BIBLIOGRAFÍA.....	47
6.	SUMMARY	49
7.	ANEXOS	58
7.1	Tablas pruebas no paramétricas	59
7.2	Interfaz de usuario NPS.....	73
7.3	Esquemas reactivos NPS	95
7.4	Código de la aplicación NPS	115

1. INTRODUCCIÓN

En la actualidad, cada vez es más demandada la utilización por parte del investigador y/o estadístico de distintos lenguajes de programación orientados a la ciencia de datos como Python o R. Estos lenguajes de programación son cada vez más utilizados, existiendo una amplia comunidad de programadores que pueden ayudar a resolver ciertas dudas a los principiantes y a los no tan noveles. Además, son lenguajes de alto nivel, es decir, sencillos de entender por el programador. (Boschetti & Massaron, 2015)

El desarrollo de aplicaciones web con estos lenguajes puede ser de gran utilidad para el propio investigador (para ayudarle en procesos de la investigación o para utilizarla como herramienta para transmitir conocimiento sobre un tema o investigación concreta) o para el público al que vaya dirigida (facilitar aprendizaje a estudiantes, investigadores de otros campos que quieran aplicar ciertos procesos a sus datos sin conocimientos de programación). La utilidad de estas aplicaciones no se reduce exclusivamente al campo de la inferencia estadística, pues pueden ser útiles en múltiples ámbitos como la demografía, la medicina, la biología o la economía. En la galería de Shiny pueden consultarse distintas aplicaciones publicadas, con complejidades y temáticas muy variadas. (RStudio, s. f.-a; Satyahadewi & Perdana, 2021)

En el lenguaje R existe la posibilidad de desarrollar estas aplicaciones web con la librería Shiny. De una forma sencilla se pueden crear aplicaciones simples y útiles. Con esta librería pueden desarrollarse aplicaciones algo más complejas combinando el lenguaje R con otros lenguajes como HTML (contenido) (*HyperText Markup o Lenguaje de Marcas de Hipertexto*) o CSS (estilo, forma, posición) (*Cascading Style Sheets* o Hojas de estilo en cascada) llegando a obtener aplicaciones web muy vistosas, lo cual puede ser muy interesante si el desarrollador quiere presentar ciertos resultados a un público. Con el fin de dar visibilidad al desarrollo de este tipo de aplicaciones y su utilidad, se decidió desarrollar una aplicación web que pudiese ser usada por la comunidad científica y universitaria. (Wickham, 2021)

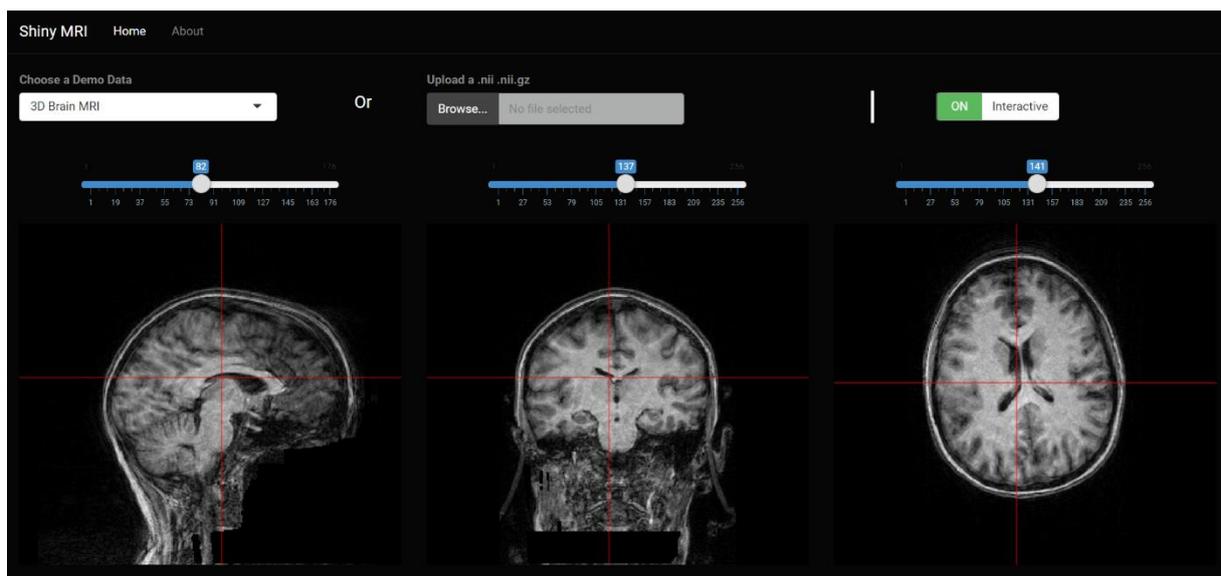


Ilustración 1: Interfaz de usuario de Shiny MRI.

Fuente: Shiny MRI

En la Ilustración 1 (Zhu et al., s. f.) se muestra la interfaz de una de las aplicaciones que aparecen en la galería de Shiny. En esta aplicación se permite al usuario visualizar imágenes de una cabeza humana obtenidas por resonancia magnética. La aplicación ofrece la posibilidad de obtener distintas vistas manipulando los valores de los botones deslizantes. Además, es posible importar en formato “.nii” un archivo con las imágenes por resonancia magnética para poder visualizarlas en la aplicación.

El enlace a la aplicación web mostrada en la Ilustración 1 aparece en la referencia correspondiente de la bibliografía. Además, puede accederse mediante el siguiente código QR:



La temática elegida para el desarrollo de la aplicación fueron las pruebas no paramétricas, que pueden ser muy útiles en ciertos contextos de aplicación de inferencia estadística. Aunque las pruebas paramétricas sean mucho más habituales, estas no siempre son la mejor opción, ya que, a pesar de que suelen ser contrastes con mayor potencia o sensibilidad (el error tipo II, β , es menos probable), éstas son menos robustas que las no paramétricas, es decir, suelen exigir ciertas condiciones previas como la homocedasticidad (igualdad de varianzas) o la normalidad de los datos, siendo válidas en un menor número de situaciones. Además, si se quiere realizar algún contraste sobre algún parámetro como la mediana o la moda, contrastar la aleatoriedad de la muestra o realizar una prueba de bondad de ajuste, los contrastes paramétricos no ofrecen una solución, al contrario que los no paramétricos. Si el investigador no está seguro del tipo de distribución de probabilidad de la que provienen los datos de la muestra, las pruebas no paramétricas son altamente recomendables. (Lehmann, 1951; Siegel, 1957)

Por tanto, el objetivo general de este trabajo es desarrollar una aplicación web que ofrezca al usuario un entorno sencillo e intuitivo en el que poder utilizar las principales pruebas no paramétricas.

Los objetivos específicos son:

- Crear una sección en la que el usuario pueda aprender los conceptos básicos y el procedimiento de cada una de las pruebas que aparezcan en la aplicación.
- Crear otra sección en la que poder aplicar las pruebas no paramétricas a cualquier conjunto de datos.

La estructura principal de este trabajo se divide en material y métodos (2), donde se detalla la teoría de las pruebas no paramétricas utilizadas, se ofrece una breve explicación sobre R y RStudio y se abordan los conceptos, elementos y funciones básicas que contiene y utiliza una aplicación Shiny, y resultados (3), donde se muestra la aplicación web obtenida, su interfaz y la explicación del funcionamiento, la utilidad, usabilidad, excepciones y los esquemas reactivos de cada apartado. Finalmente, en el apartado de conclusiones (4), se ofrece una breve reflexión sobre la importancia de Shiny y de las pruebas no paramétricas, la utilidad que pretende alcanzar la aplicación y algunas perspectivas de mejora para ésta.

En los anexos pueden consultarse la mayoría de las imágenes de la interfaz de usuario (7.2), esquemas reactivos (7.3) o código de la aplicación (7.4) que no se muestran en el cuerpo del trabajo.

2. MATERIAL Y METODOS

2.1 Inferencia no paramétrica

2.1.1 Conceptos previos

2.1.1.1 Rachas

(Gibbons & Chakraborti, 1992)

Dada una variable X en la que se puede definir una dicotomía con dos símbolos diferentes (+, -), llamamos racha a una secuencia de uno o más símbolos iguales. Denotamos con R al número total de rachas, con $R +$ al número de rachas formadas con los valores del primer símbolo de la dicotomía y con $R -$ al número de rachas formadas con los valores del segundo símbolo de esta. A la cantidad de valores que conforman una racha la llamamos longitud de la racha.

En la práctica, dependiendo de la tipología de la variable aleatoria X se genera la dicotomía de distinta manera:

- Variables cualitativas: Se asignan al primer símbolo una o varias categorías y al segundo símbolo las categorías restantes.
- Variables cuantitativas discretas: Se asignan al primer símbolo uno o varios valores y al segundo símbolo los valores restantes.
- Variables cuantitativas continuas: Se suele dicotomizar la variable respecto al valor de la mediana. Se asignan al primer símbolo los valores mayores o iguales a la mediana y al segundo símbolo los valores menores a la misma.

Propiedades:

- 1 Una racha y la siguiente se diferencian porque contienen símbolos distintos.
- 2 $R \geq 2$
- 3 Si R es par ($R = 2K$), entonces: $K = R + = R -$
- 4 Si R es impar ($R = 2K + 1$), entonces: $\begin{cases} K + 1 = R +, K = R - & \text{si primera racha} + \\ K + 1 = R -, K = R + & \text{si primera racha} - \end{cases}$
- 5 Sean $n_1 = \#\{+\}$, $n_2 = \#\{-\}$ y $n = n_1 + n_2$ se verifica que la función de densidad del número de rachas R es de la forma:

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

$$6 \quad \begin{cases} 2 \leq R \leq 2 \cdot \min\{n_1, n_2\} + 1 & \text{si } n_1 \neq n_2 \\ 2 \leq R \leq 2 \cdot n_1 & \text{si } n_1 = n_2 \end{cases}$$

$$7 \quad E[R] = \frac{2n_1n_2}{n} + 1$$

$$8 \quad Var[R] = \frac{2n_1n_2(2n_1n_2-n)}{n^2(n-1)}$$

- 9 La distribución asintótica de R ($n \rightarrow \infty$) es:

$$R \approx N(E[R], Var[R]) \quad \text{por tanto:} \quad Z = \frac{R-E[R]}{\sqrt{Var[R]}} \approx N(0,1)$$

En la práctica se aplica una corrección por continuidad de 0.5.

2.1.1.2 Rangos de un conjunto de observaciones

(Gibbons & Chakraborti, 1992)

Sea x_1, x_2, \dots, x_n una muestra aleatoria simple de tamaño n de una variable ordinal o cuantitativa:

Ordenando de menor a mayor los datos obtenemos la correspondiente muestra ordenada, que denotamos: $\{x_{(1)}, x_{(2)}, \dots, x_{(n)}\}$. Cuando un valor se repite diremos que la muestra tiene observaciones empatadas o empates.

Definiciones:

- Si todos los datos son distintos \rightarrow No hay observaciones empatadas \rightarrow El rango R_i del dato x_i es el número de orden que ocupa en la muestra ordenada: $x_{(1)} < x_{(2)} < \dots < x_{(n)} \rightarrow R_i = i$
- Si un valor se repite r veces (es decir, tiene r empates), se le asigna como rango el rango medio correspondiente a los r datos empatados entre sí:

$$\dots x_{(k-1)} < x_{(k)} = x_{(k+1)} = x_{(k+2)} = \dots = x_{(k+r-1)} < x_{(k+r)} \dots$$

$$R_k = \frac{k + (k+1) + (k+2) + \dots + (k+r-1)}{r} = k + \frac{r-1}{2}$$

siendo k la posición del primer elemento empatado en la muestra ordenada y r el número de elementos que forman el empate.

- Llamamos g al número de grupos con datos empatados y t_j al número de datos empatados en el grupo j . ($j = 1, \dots, g$)

Propiedades:

1. Tanto si hay datos con empates como si los hay sin empates, la suma de los rangos de una muestra de tamaño n vale: $\sum_{i=1}^n R_i = \frac{n(n+1)}{2}$
2. La suma de los cuadrados de los rangos sin empates es: $(\sum_{i=1}^n R_i^2)_{SE} = \frac{n(n+1)(2n+1)}{6}$
3. La suma de los cuadrados de los rangos con empates cuando hay g grupos de empates de tamaños t_j es: $(\sum_{i=1}^n R_i^2)_{CE} = \frac{n(n+1)(2n+1)}{6} - \frac{1}{12} \sum_{j=1}^g t_j (t_j^2 - 1)$

El elemento añadido respecto a la suma de cuadrados sin empates se conoce como factor de corrección por empates.

4. La media aritmética de la variable de los rangos con o sin empates es: $\bar{R} = \frac{n+1}{2}$
5. La varianza de la variable de los rangos sin empates es: $Var(R)_{SE} = \frac{n^2-1}{12}$
6. La varianza de la variable de los rangos con empates es:

$$Var(R)_{CE} = \frac{n^2-1}{12} - \frac{1}{12n} \sum_{j=1}^g t_j (t_j^2 - 1)$$

2.1.2 Pruebas no paramétricas de aleatoriedad

Cuando nuestro objetivo es determinar la aleatoriedad de una muestra obtenida de una distribución realizamos un contraste basado en rachas con los datos de la muestra en el orden obtenido.

Una de las técnicas no paramétricas más habituales es la prueba de las rachas.

2.1.2.1 Prueba de las rachas de Wald-Wolfowitz

(Gibbons & Chakraborti, 1992)

Dada una muestra de tamaño n las hipótesis del contraste son: $\begin{cases} H_0: \text{Existe aleatoriedad} \\ H_1: \text{No existe aleatoriedad} \end{cases}$

Suponiendo (como en 2.1.1.1) que es posible dicotomizar la muestra con dos símbolos (+, -) para contabilizar el número de rachas, si se encuentra un número muy elevado o reducido de ellas es probable que no exista aleatoriedad.

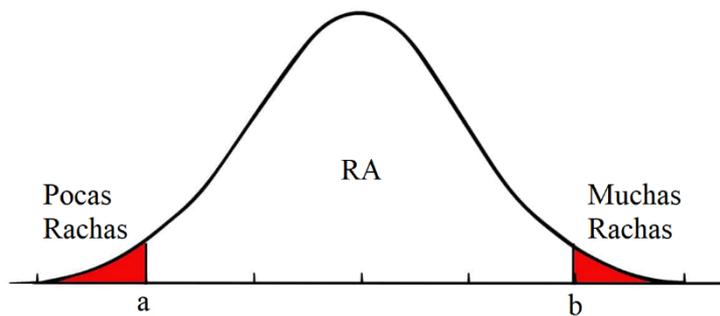


Ilustración 2: Región de aceptación y región crítica en una prueba de las rachas.

Fuente: Elaboración Propia

Procedimiento:

- Se calcula el número total de rachas en la muestra: r_{obs}
- Si se tienen muestras pequeñas, (en la práctica $n_1, n_2 \leq 20$) se busca en la tabla 7 (R) (Anexo 7.1) (Swed & Eisenhart, 1943) los valores a y b que forman un intervalo cerrado que se corresponde con la región de aceptación: $[a, b] = RA$
 - Si $a \leq r_{obs} \leq b \rightarrow$ Se acepta H_0 , existe aleatoriedad
 - Si $r_{obs} < a$ o $r_{obs} > b \rightarrow$ Se rechaza H_0 , no existe aleatoriedad.
- Para muestras grandes (en la práctica n_1 o $n_2 > 20$) se utilizará la distribución asintótica de R:

$$z_{obs} = \frac{r_{obs} - E[R]}{\sqrt{Var[R]}} \approx N(0,1)$$

Se compara el valor z_{obs} con el cuantil de la distribución normal estándar $Z_{1-\frac{\alpha}{2}}$

- Si $|z_{obs}| < Z_{1-\frac{\alpha}{2}} \rightarrow$ Se acepta H_0 , existe aleatoriedad
- Si $|z_{obs}| \geq Z_{1-\frac{\alpha}{2}} \rightarrow$ Se rechaza H_0 , no existe aleatoriedad.

Una de las funciones en R que realiza la prueba de las rachas es *runs.test* de la librería *randtests*.

2.1.3 Pruebas no paramétricas de bondad de ajuste

Las pruebas de bondad de ajuste son útiles para contrastar si un conjunto de observaciones se ajusta o no a una distribución de probabilidad dada o si dos muestras provienen de la misma distribución de probabilidad. En este trabajo se abordará la prueba de Kolmogorov-Smirnov para ambas situaciones.

2.1.3.1 Prueba de Kolmogorov-Smirnov

(Gibbons & Chakraborti, 1992; Massey Jr, 1951)

Distribución conocida F_0

Para el caso en el que se tiene una distribución conocida F_0 , suponiendo que se obtiene una muestra aleatoria simple de tamaño n de una variable con distribución desconocida F y se quiere contrastar si la muestra proviene o no de la distribución F_0 :

$$\begin{cases} H_0: F = F_0 \\ H_1: F \neq F_0 \end{cases} \quad \text{Contraste unilateral superior.}$$

Procedimiento:

- Se calcula la función de distribución empírica $F_n^*(x) = \frac{\sum_{i=1}^n I\{x_i \leq x\}}{n}$
donde $I\{x_i \leq x\} = \begin{cases} 1 & \text{si } x_i \leq x \\ 0 & \text{si } x_i > x \end{cases}$
Cuando $n \rightarrow \infty$ entonces $F_n^*(x) \rightarrow F(x)$
- Para medir la similitud entre F y F_0 se utiliza el estadístico de Kolmogorov-Smirnov que compara $F_n^*(x)$ con $F_0(x)$: $D = D(F_n^*(x), F_0) = \max\{|F_n^*(x) - F_0(x)|, \forall x \in \mathbb{R}\}$
- Para aceptar H_0 las diferencias deben ser pequeñas, próximas a cero. La región de aceptación es $\{D_{obs} < D_{n;\alpha}\}$ y la región crítica $\{D_{obs} \geq D_{n;\alpha}\}$. $D_{n;\alpha}$ es un cuantil del estadístico D que se encuentra en la tabla 6 (KS) (Anexo 7.1)
- Fijado el valor de α se compara el valor de D_{obs} con el cuantil $D_{n;\alpha}$:
 - Si $D_{obs} < D_{n;\alpha} \rightarrow$ Se acepta H_0
 - Si $D_{obs} \geq D_{n;\alpha} \rightarrow$ Se rechaza H_0

Distribuciones F y G

Si se tienen dos muestras aleatorias simples $\{x_1, x_2, \dots, x_n\}$ y $\{y_1, y_2, \dots, y_m\}$ obtenidas de dos poblaciones independientes con distribuciones F y G desconocidas y se quiere contrastar si ambas poblaciones coinciden o no:

$$\begin{cases} H_0: F = G \\ H_1: F \neq G \end{cases} \quad \text{Contraste unilateral superior.}$$

Procedimiento:

- Se unen todas las observaciones en una única muestra ordenada de menor a mayor $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n+m)}$ donde cada $z_{(k)}$ es un valor x_i o y_j
- Se calculan las funciones de distribución empíricas en cada valor $z_{(k)}$:

$$F_n^*(z_{(k)}) = \frac{\sum_{i=1}^n I\{x_i \leq z_{(k)}\}}{n} \quad G_m^*(z_{(k)}) = \frac{\sum_{j=1}^m I\{y_j \leq z_{(k)}\}}{m}$$

$$\text{donde } I\{x_i \leq z_{(k)}\} = \begin{cases} 1 & \text{si } x_i \leq z_{(k)} \\ 0 & \text{si } x_i > z_{(k)} \end{cases} \quad I\{y_j \leq z_{(k)}\} = \begin{cases} 1 & \text{si } y_j \leq z_{(k)} \\ 0 & \text{si } y_j > z_{(k)} \end{cases}$$

- Se comparan las dos distribuciones $F_n^*(z_{(k)})$ y $G_m^*(z_{(k)})$ con el estadístico de Kolmogorov-Smirnov:

$$D = D(F_n^*(z_{(k)}), G_m^*(z_{(k)})) = \text{máx}\{|F_n^*(z_{(k)}) - G_m^*(z_{(k)})|\}$$

- Para aceptar H_0 las diferencias deben ser pequeñas, próximas a cero. La región de aceptación es $\{D_{obs} < D_{n;\alpha}\}$ y la región crítica $\{D_{obs} \geq D_{n;\alpha}\}$. $D_{n;\alpha}$ es un cuantil del estadístico D que se encuentra en la tabla 6 (KS) (Anexo 7.1)
- Fijado el valor de α se compara el valor de D_{obs} con el cuantil $D_{n;\alpha}$:
 - Si $D_{obs} < D_{n;\alpha} \rightarrow$ Se acepta H_0
 - Si $D_{obs} \geq D_{n;\alpha} \rightarrow$ Se rechaza H_0

Aplicación de las técnicas en RStudio:

Una de las funciones en R que realiza la prueba de bondad de ajuste para ambas situaciones es *ks.test* de la librería *stats*.

2.1.3.2 Gráficos Cuantil-Cuantil

(Marden, 2004)

Una prueba gráfica muy habitual cuando se quieren comparar dos distribuciones es el gráfico cuantil-cuantil o Q-Q plot. Estos gráficos comparan los cuantiles de la distribución empírica muestral con los cuantiles teóricos de una distribución conocida ($F = F_0$) o con los cuantiles de otra distribución empírica muestral en el caso $F = G$.

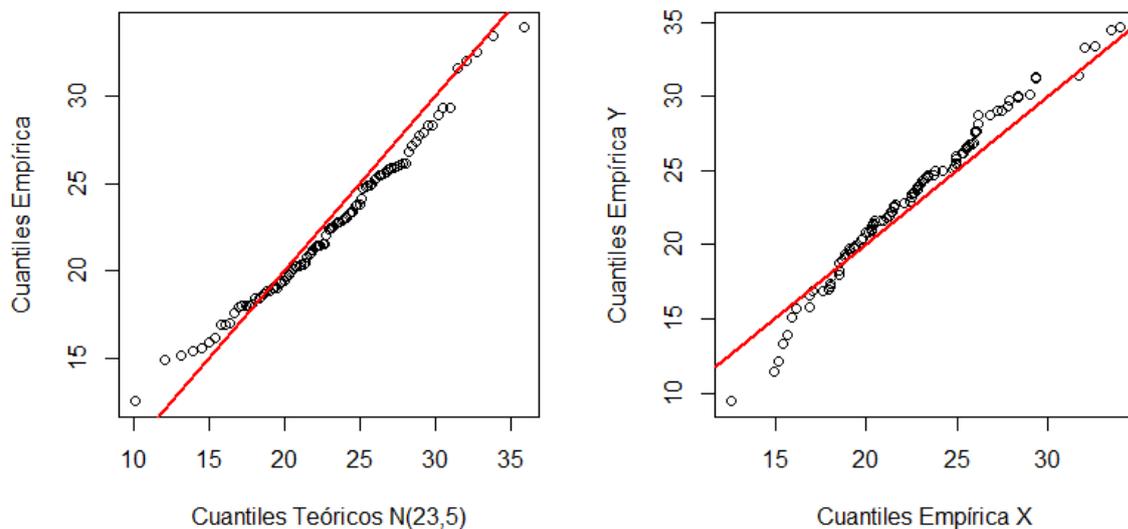


Ilustración 3: Gráficos Cuantil-Cuantil. A la izquierda para el caso $F=F_0$ y a la derecha para el caso $F=G$.

Fuente: Elaboración Propia en RStudio

2.1.4 Pruebas no paramétricas de localización de una muestra (Mediana).

(Hollander et al., 2013)

Los contrastes no paramétricos de localización consisten en estudiar el valor de la mediana θ de la población de la que proviene la muestra.

Prueba de los rangos con signo de Wilcoxon

Dada una muestra aleatoria simple X de una población con mediana θ se quiere contrastar:

$\begin{cases} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{cases}$	$\begin{cases} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{cases}$	$\begin{cases} H_0: \theta \leq \theta_0 \\ H_1: \theta > \theta_0 \end{cases}$
Bilateral	Unilateral inferior	Unilateral superior

Procedimiento:

- Se construye una variable $D = X - \theta_0$
La mediana de la nueva variable sería: $\theta_D = \theta - \theta_0$
Con lo que las hipótesis se transforman en:

$\begin{cases} H_0: \theta_D = 0 \\ H_1: \theta_D \neq 0 \end{cases}$	$\begin{cases} H_0: \theta_D \geq 0 \\ H_1: \theta_D < 0 \end{cases}$	$\begin{cases} H_0: \theta_D \leq 0 \\ H_1: \theta_D > 0 \end{cases}$
Bilateral	Unilateral inferior	Unilateral superior

- Se calculan los valores absolutos de los valores de la nueva variable $|D_i|$ y se calculan los rangos R_i .
- Se obtiene el estadístico de los rangos con signo de Wilcoxon T^+ :

$$T^+ = \sum_{D_i > 0} R_i$$

- Si el tamaño muestral es pequeño (en la práctica $n \leq 12$), se busca en la tabla 8 (T^+) (Anexo 7.1) la significación muestral correspondiente a cada hipótesis:

- $H_1: \theta_D < 0 \rightarrow H_1: \theta < \theta_0 \rightarrow P_I = P(T^+ \leq T_{obs}^+) \rightarrow \begin{cases} \text{si } P_I > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_I \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D > 0 \rightarrow H_1: \theta > \theta_0 \rightarrow P_S = P(T^+ \geq T_{obs}^+) \rightarrow \begin{cases} \text{si } P_S > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_S \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D \neq 0 \rightarrow H_1: \theta \neq \theta_0 \rightarrow P = 2 \cdot \min\{P_I, P_S\} \rightarrow \begin{cases} \text{si } P > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$

- Si el tamaño muestral es grande (en la práctica $n > 12$), se utiliza la distribución asintótica de T^+ :

$$z_{obs} = \frac{T_{obs}^+ - E[T^+]}{\sqrt{Var[T^+]}} \approx N(0,1)$$

donde:

$$E[T^+] = \frac{n(n+1)}{4} \quad Var[T^+] = \frac{n(n+1)(2n+1)}{24}$$

Se compara el valor z_{obs} con el cuantil de la distribución normal estándar correspondiente según el tipo de contraste:

- $H_1: \theta_D < 0 \rightarrow H_1: \theta < \theta_0 \rightarrow \begin{cases} \text{si } z_{obs} > z_\alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \leq z_\alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$

- $H_1: \theta_D > 0 \rightarrow H_1: \theta > \theta_0 \rightarrow \begin{cases} \text{si } z_{obs} < z_{1-\alpha} \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \geq z_{1-\alpha} \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D \neq 0 \rightarrow H_1: \theta \neq \theta_0 \rightarrow \begin{cases} \text{si } |z_{obs}| < z_{1-\frac{\alpha}{2}} \rightarrow \text{Se acepta } H_0 \\ \text{si } |z_{obs}| \geq z_{1-\frac{\alpha}{2}} \rightarrow \text{Se rechaza } H_0 \end{cases}$

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de localización de una muestra es *wilcox.test* de la librería *stats*, en la que un parámetro “mu” hace las veces del valor de θ_0 de la hipótesis nula.

2.1.5 Pruebas no paramétricas para dos muestras

2.1.5.1 Prueba para datos apareados de Wilcoxon

(Hollander et al., 2013)

La prueba para datos apareados de Wilcoxon se utiliza para comparar el rango medio de dos muestras relacionadas (X e Y) para comprobar si existen diferencias significativas. Es una alternativa al test paramétrico t-Student para datos relacionados.

$\begin{cases} H_0: \theta_x = \theta_y \\ H_1: \theta_x \neq \theta_y \end{cases}$	$\begin{cases} H_0: \theta_x \leq \theta_y \\ H_1: \theta_x > \theta_y \end{cases}$	$\begin{cases} H_0: \theta_x \geq \theta_y \\ H_1: \theta_x < \theta_y \end{cases}$
Bilateral	Unilateral inferior	Unilateral superior

Procedimiento:

- Se calcula una nueva variable $D = Y - X$
Las hipótesis se transforman en:

$\begin{cases} H_0: \theta_D = 0 \\ H_1: \theta_D \neq 0 \end{cases}$	$\begin{cases} H_0: \theta_D \geq 0 \\ H_1: \theta_D < 0 \end{cases}$	$\begin{cases} H_0: \theta_D \leq 0 \\ H_1: \theta_D > 0 \end{cases}$
Bilateral	Unilateral inferior	Unilateral superior

- Se calculan los rangos R_i de los valores absolutos $|D_i|$
- Se obtiene el estadístico de los rangos con signo de Wilcoxon T^+ :

$$T^+ = \sum_{D_i > 0} R_i$$

- Si el tamaño muestral es pequeño (en la práctica $n \leq 12$), se busca en la tabla 8 (T^+) (Anexo 7.1) la significación muestral correspondiente a cada hipótesis:

- $H_1: \theta_D < 0 \rightarrow H_1: \theta_x > \theta_y \rightarrow P_I = P(T^+ \leq T_{obs}^+) \rightarrow \begin{cases} \text{si } P_I > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_I \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D > 0 \rightarrow H_1: \theta_x < \theta_y \rightarrow P_S = P(T^+ \geq T_{obs}^+) \rightarrow \begin{cases} \text{si } P_S > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_S \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D \neq 0 \rightarrow H_1: \theta_x \neq \theta_y \rightarrow P = 2 \cdot \min\{P_I, P_S\} \rightarrow \begin{cases} \text{si } P > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$

- Si el tamaño muestral es grande (en la práctica $n > 12$), se utiliza la distribución asintótica de T^+ :

$$z_{obs} = \frac{T_{obs}^+ - E[T^+]}{\sqrt{Var[T^+]}} \approx N(0,1)$$

donde:

$$E[T^+] = \frac{n(n+1)}{4} \quad \text{Var}[T^+] = \frac{n(n+1)(2n+1)}{24}$$

Se compara el valor z_{obs} con el cuantil de la distribución normal estándar correspondiente según el tipo de contraste:

- $H_1: \theta_D < 0 \rightarrow H_1: \theta_x > \theta_y \rightarrow \begin{cases} \text{si } z_{obs} > z_\alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \leq z_\alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D > 0 \rightarrow H_1: \theta_x < \theta_y \rightarrow \begin{cases} \text{si } z_{obs} < z_{1-\alpha} \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \geq z_{1-\alpha} \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: \theta_D \neq 0 \rightarrow H_1: \theta_x \neq \theta_y \rightarrow \begin{cases} \text{si } |z_{obs}| < z_{1-\frac{\alpha}{2}} \rightarrow \text{Se acepta } H_0 \\ \text{si } |z_{obs}| \geq z_{1-\frac{\alpha}{2}} \rightarrow \text{Se rechaza } H_0 \end{cases}$

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de datos apareados de Wilcoxon es *wilcox.test* de la librería *stats*, en la que un parámetro “*paired=TRUE*” indica que las muestras son relacionadas/pareadas.

2.1.5.2 Prueba para dos muestras independientes de Mann-Whitney

(Gibbons & Chakraborti, 1992)

La prueba de Mann-Whitney se utiliza para comparar dos muestras independientes (X e Y). La muestra X tiene un tamaño de n_1 y la muestra Y de n_2 .

$\begin{cases} H_0: X = Y \\ H_1: X \neq Y \end{cases}$	$\begin{cases} H_0: X \leq Y \\ H_1: X > Y \end{cases}$	$\begin{cases} H_0: X \geq Y \\ H_1: X < Y \end{cases}$
Bilateral	Unilateral inferior	Unilateral superior

Procedimiento:

- Se calculan los rangos de los $n = n_1 + n_2$ datos.
- Se obtienen los valores de R_1 y R_2 :

$$R_1 = \sum_{i=1}^{n_1} \text{Rango}(x_i) \quad R_2 = \sum_{j=1}^{n_2} \text{Rango}(y_j)$$

- Se calculan los estadísticos U_1 y U_2 :

$$U_1 = R_1 - \frac{n_1(n_1 + 1)}{2} \quad U_2 = R_2 - \frac{n_2(n_2 + 1)}{2}$$

- El valor del estadístico de Mann-Whitney se obtiene como:

$$U_{obs} = \min\{U_1, U_2\}$$

- Si el tamaño muestral es pequeño (en la práctica $n_1, n_2 \leq 8$), se busca en la tabla 9 (U) (Anexo 7.1) el valor de $P(U \leq U_{obs})$ y se procede según el tipo de contraste:

- $H_1: X > Y \rightarrow P_1 = P(U \leq U_{obs}) \rightarrow \begin{cases} \text{si } P_1 > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_1 \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$
- $H_1: X < Y \rightarrow P_S = P(U \geq U_{obs}) \rightarrow \begin{cases} \text{si } P_S > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P_S \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$

$$\circ H_1: X \neq Y \rightarrow P = 2 \cdot \min\{P_I, P_S\} \rightarrow \begin{cases} \text{si } P > \alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } P \leq \alpha \rightarrow \text{Se rechaza } H_0 \end{cases}$$

La tabla solo ofrece los valores de $P(U \leq U_{obs})$. Si se necesita el valor de $P(U \geq U_{obs})$ utilizar:
 $P(U \geq U_{obs}) = 1 - P(U < U_{obs}) = 1 - P(U \leq U_{obs} - 1)$

- Si el tamaño muestral es grande (en la práctica $n_1, n_2 > 8$), se utiliza la distribución asintótica de U:

$$z_{obs} = \frac{U_{obs} - E[U]}{\sqrt{Var[U]}} \approx N(0,1)$$

donde:

$$E[U] = \frac{n_1 \cdot n_2}{2} \quad Var[U] * = \frac{n_1 \cdot n_2(n_1 + n_2 + 1)}{12}$$

Se compara el valor z_{obs} con el cuantil de la distribución normal estándar correspondiente según el tipo de contraste:

$$\begin{aligned} \circ H_1: X > Y &\rightarrow \begin{cases} \text{si } z_{obs} > z_\alpha \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \leq z_\alpha \rightarrow \text{Se rechaza } H_0 \end{cases} \\ \circ H_1: X < Y &\rightarrow \begin{cases} \text{si } z_{obs} < z_{1-\alpha} \rightarrow \text{Se acepta } H_0 \\ \text{si } z_{obs} \geq z_{1-\alpha} \rightarrow \text{Se rechaza } H_0 \end{cases} \\ \circ H_1: X \neq Y &\rightarrow \begin{cases} \text{si } |z_{obs}| < z_{1-\frac{\alpha}{2}} \rightarrow \text{Se acepta } H_0 \\ \text{si } |z_{obs}| \geq z_{1-\frac{\alpha}{2}} \rightarrow \text{Se rechaza } H_0 \end{cases} \end{aligned}$$

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba para dos muestras independientes de Mann-Whitney es *wilcox.test* de la librería *stats*, en la que un parámetro “*paired=FALSE*” indica que las muestras son independientes.

2.1.6 Pruebas no paramétricas para k muestras

2.1.6.1 Prueba de Kruskal-Wallis

(Gibbons & Chakraborti, 1992; Hollander et al., 2013)

La prueba de Kruskal-Wallis es útil si se quiere generalizar para $k > 2$ muestras independientes la prueba de Mann-Whitney. Es la versión no paramétrica del análisis de la varianza de un factor en poblaciones normales (ANOVA paramétrico).

Sea x_{ij} el dato j de la muestra i, n_i el tamaño de la muestra i y N el tamaño muestral total $N = \sum n_i$, siendo $i = (1, 2, \dots, k)$ y $j = (1, 2, \dots, n_i)$ se podría escribir el modelo de la forma:

$$x_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

donde μ es el efecto común, τ_i el efecto de la muestra i y ε_{ij} el error.

Las hipótesis del contraste serían:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k & \text{(muestras iguales)} \\ H_1: \tau_{i_1} \neq \tau_{i_2} \text{ para algún } i_1 \neq i_2 & \text{(muestras diferentes)} \end{cases}$$

Procedimiento:

- Se calculan los rangos de los N datos: $R_{ij} = \text{rango}(x_{ij})$
- Para cada muestra i , se calculan los valores de $R_{i.} = \sum_{j=1}^{n_i} R_{ij}$ (suma de los rangos de la muestra i -ésima)
- Se obtiene el estadístico de Kruskal-Wallis como:

$$H_{obs} = \left(\frac{12}{N(N+1)} \sum_{i=1}^k \frac{(R_{i.})^2}{n_i} \right) - 3(N+1)$$

- Se busca en la Tabla 10 (H) (Anexo 7.1) el valor de H_{crit} y se compara con el valor de H_{obs} :
 - Si $H_{obs} < H_{crit}$ se Acepta H_0
 - Si $H_{obs} \geq H_{crit}$ se Rechaza H_0
- Si el valor de H_{crit} no se encuentra en la tabla o las muestras tienen empates se utiliza la distribución asintótica del estadístico H :

$$H \approx \chi_{k-1}^2$$

Se comparan los cuantiles de la distribución χ_{k-1}^2 con el valor de H_{obs} :

- Si $H_{obs} < \chi_{k-1;1-\alpha}^2$ se Acepta H_0
- Si $H_{obs} \geq \chi_{k-1;1-\alpha}^2$ se Rechaza H_0

Si el contraste resulta significativo se realiza un análisis post-hoc comparando las muestras dos a dos con la técnica para dos muestras independientes de Mann-Whitney.

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de Kruskal-Wallis es *kruskal.test* de la librería *stats*.

2.1.6.2 Prueba de Friedman (Hollander et al., 2013)

(Hollander et al., 2013)

La prueba de Friedman es útil si se quiere generalizar para $k > 2$ muestras relacionadas la prueba de dos muestras pareadas de Wilcoxon. Es la versión no paramétrica del análisis de la varianza de medidas repetidas en poblaciones normales.

Sea x_{ij} el valor de la variable j en el individuo i , n el número de individuos y $N = n \cdot k$ el tamaño muestral total, siendo $i = (1, 2, \dots, n)$ y $j = (1, 2, \dots, k)$ se podría escribir el modelo de la forma:

$$x_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

donde μ es el efecto común, τ_j el efecto de la muestra j , β_i el efecto del individuo i y ε_{ij} el error.

En este caso únicamente se mide la influencia del efecto de la muestra pareada τ_j

Las hipótesis del contraste serían:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k & (\text{muestras iguales}) \\ H_1: \tau_{i_1} \neq \tau_{i_2} \text{ para algún } i_1 \neq i_2 & (\text{muestras diferentes}) \end{cases}$$

Procedimiento:

- Se calculan los rangos R_{ij} de las variables en cada individuo.
- Se obtienen los valores de $R_{.j} = \sum_{i=1}^n R_{ij}$

- Se calcula el valor del estadístico de Friedman:

$$S_{obs} = \left(\frac{12}{nk(k+1)} \sum_{j=1}^k R_{.j}^2 \right) - 3n(k+1)$$

- Se busca en la Tabla 11 (S) (Anexo 7.1) el valor de S_{crit} y se compara con el valor de S_{obs} :
 - Si $S_{obs} < S_{crit}$ se Acepta H_0
 - Si $S_{obs} \geq S_{crit}$ se Rechaza H_0
- Si $k > 5$ o $n > 15$ o si hay muchos empates se utiliza la distribución asintótica de S:

$$S \approx \chi_{k-1}^2$$

Se comparan los cuantiles de la distribución χ_{k-1}^2 con el valor de S_{obs} :

- Si $S_{obs} < \chi_{k-1;1-\alpha}^2$ se Acepta H_0
- Si $S_{obs} \geq \chi_{k-1;1-\alpha}^2$ se Rechaza H_0

Si el contraste resulta significativo se realiza un análisis post-hoc comparando las muestras dos a dos con la técnica para dos muestras pareadas de Wilcoxon.

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de Friedman es *friedman.test* de la librería *stats*.

2.1.7 Pruebas no paramétricas de independencia.

Si se quiere estudiar la dependencia de dos variables que tengan distintas unidades sólo podemos compararlas dos a dos.

2.1.7.1 Prueba de Kendall

(Gibbons & Chakraborti, 1992)

En la prueba de Kendall se estudia el posible comportamiento conjunto o concordancia de dos variables X e Y.

Las hipótesis del contraste son: $\begin{cases} H_0: \text{Variables Independientes} \\ H_1: \text{Variables Dependientes} \end{cases}$

Concordancia

- Una pareja de observaciones (x_i, y_i) y (x_j, y_j) es concordante si $(x_i - x_j)(y_i - y_j) > 0$
- Una pareja de observaciones (x_i, y_i) y (x_j, y_j) es discordante si $(x_i - x_j)(y_i - y_j) < 0$
- Si $(x_i - x_j)(y_i - y_j) = 0$ la pareja de observaciones no es concordante ni discordante.

Procedimiento:

- Para cada $i < j$, se calcula $k((x_i, y_i), (x_j, y_j)) = \begin{cases} 1 & \text{si } (x_i - x_j)(y_i - y_j) > 0 \\ 0 & \text{si } (x_i - x_j)(y_i - y_j) = 0 \\ -1 & \text{si } (x_i - x_j)(y_i - y_j) < 0 \end{cases}$
- Se calcula el valor del estadístico de Kendall:

$$k_{obs} = \sum_{i=1}^{n-1} \sum_{j=i+1}^n k((x_i, y_i), (x_j, y_j)) = \#\{concordantes\} - \#\{discordantes\}$$

- Se obtiene el coeficiente de correlación τ de Kendall

$$\tau_{obs} = \frac{2k}{n(n-1)} \quad -1 \leq \tau \leq +1$$

- Se busca en la Tabla 12 (τ) (Anexo 7.1) el valor de τ_{crit} y se compara con el valor de τ_{obs} :
 - Si $\tau_{obs} < \tau_{crit}$ se Acepta H_0
 - Si $\tau_{obs} \geq \tau_{crit}$ se Rechaza H_0

- Si el tamaño muestral es grande (en la práctica $n > 30$), se utiliza la distribución asintótica de τ :

$$z_{obs} = \frac{\tau_{obs} - E[\tau]}{\sqrt{Var[\tau]}} \approx N(0,1)$$

donde:

$$E[\tau] = 0 \quad Var[\tau] = \frac{2(2n+5)}{9n(n-1)}$$

Se compara el valor de $|z_{obs}|$ con el cuantil $z_{1-\frac{\alpha}{2}}$:

- Si $|z_{obs}| < z_{1-\frac{\alpha}{2}}$ se Acepta H_0
 - Si $|z_{obs}| \geq z_{1-\frac{\alpha}{2}}$ se Rechaza H_0
- Si se rechaza la hipótesis nula el valor de τ_{obs} nos indica el grado de dependencia:
 - Si $\tau_{obs} < 0$ la dependencia es inversa, X e Y son discordantes.
 - Si $\tau_{obs} > 0$ la dependencia es directa, X e Y son concordantes.

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de Kendall es *cor.test* de la librería *stats*, en la que un parámetro “*method='kendall'*” selecciona esta prueba.

2.1.7.2 Prueba de Spearman

(Hollander et al., 2013)

En la prueba de Spearman se estudia la posible correlación entre dos variables ordinales o cuantitativas. Las variables no necesitan seguir una distribución normal como en el caso de la correlación de Pearson. Esta prueba se basa en reemplazar los valores por sus rangos y medir la correlación monótona entre X e Y.

Las hipótesis del contraste son: $\begin{cases} H_0: \text{Variables Independientes} & r_{xy} = 0 \\ H_1: \text{Variables Dependientes} & r_{xy} \neq 0 \end{cases}$

Procedimiento:

- El primer paso es calcular los rangos de ambas variables por separado $R_i = \text{rango}(x_i)$; $S_i = \text{rango}(y_i)$
- Se calcula el estadístico de Spearman:

$$r_{s_{obs}} = \frac{12}{n^3 - n} \cdot \sum_{i=1}^n \left(R_i - \frac{n+1}{2} \right) \left(S_i - \frac{n+1}{2} \right) \quad -1 \leq r_s \leq +1$$

- Se busca en la tabla 13 (RS) (Anexo 7.1) el valor de r_{crit} y se compara con el valor de $r_{s_{obs}}$:
 - Si $r_{s_{obs}} < r_{crit}$ se Acepta H_0
 - Si $r_{s_{obs}} \geq r_{crit}$ se Rechaza H_0
- Si el tamaño muestral es grande (en la práctica $n > 30$), se utiliza la distribución *t-Student* (r_s no sigue una distribución asintótica conocida):

$$t_{obs} = r_{s_{obs}} \sqrt{\frac{n-2}{1-r_{s_{obs}}^2}} \approx t_{n-2}$$

Se compara el valor de $|t_{obs}|$ con el cuantil $t_{n-2; 1-\frac{\alpha}{2}}$:

- Si $|t_{obs}| < t_{n-2; 1-\frac{\alpha}{2}}$ se Acepta H_0
 - Si $|t_{obs}| \geq t_{n-2; 1-\frac{\alpha}{2}}$ se Rechaza H_0
- Si se rechaza la hipótesis nula el valor de $r_{s_{obs}}$ nos indica el grado de correlación:
 - Si $r_{s_{obs}} < 0$ la correlación es inversa.
 - Si $r_{s_{obs}} > 0$ la correlación es directa.

Aplicación de la técnica en RStudio:

Una de las funciones en R que realiza la prueba de Spearman es *cor.test* de la librería *stats*, en la que un parámetro “*method=’spearman’*” selecciona esta prueba.

2.2 R y RStudio

R

R es un lenguaje de programación orientado principalmente al análisis estadístico. Fue desarrollado inicialmente en la universidad de Auckland, Nueva Zelanda por Robert Gentleman y Ross Ihaka en 1993. (Ihaka, s. f.) Puede decirse que sus bases provienen del lenguaje de programación S, desarrollado por John Chambers, Rick Becker y otros colaboradores desde finales de la década de los setenta.

La versión 1.0.0 de R fue presentada en el año 2000. (Dalgaard, 2000) A la fecha de realización de este trabajo la última versión disponible de R es la 4.1.0 (mayo 2021). Como se puede ver es un lenguaje bastante moderno.

En este lenguaje se pueden aplicar multitud de pruebas, algoritmos y gráficos estadísticos además de permitir, al tratarse de un lenguaje de programación, la definición de funciones propias que nos ayuden a obtener los resultados deseados.

RStudio

RStudio es un entorno de desarrollo integrado (IDE) para R, que ofrece una consola y un editor, así como herramientas de depuración o gestión del directorio de trabajo, facilitando el aprendizaje y el uso de este lenguaje. (RStudio, s. f.)

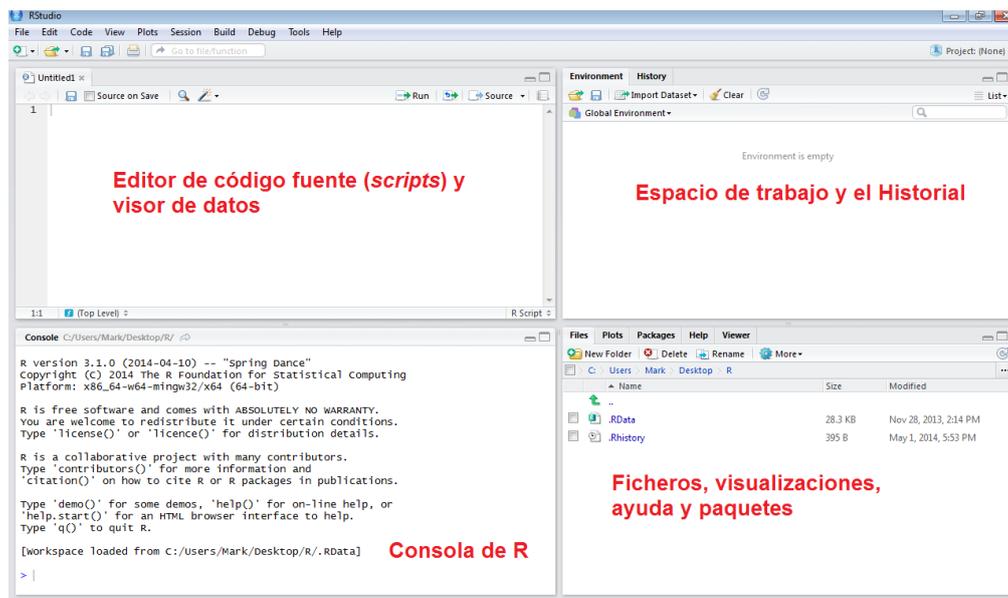


Ilustración 4: Interfaz del IDE RStudio.

Fuente: (Pérez González & Colebrook Santamaría, 2014)

2.3 Shiny

(Chang et al., 2017; RStudio, s. f.-b)

Shiny es una librería de R que facilita el desarrollo de aplicaciones web interactivas. Se pueden crear aplicaciones útiles y simples de manera sencilla y sin conocimientos de desarrollo web. En la galería de Shiny se pueden encontrar multitud de aplicaciones desarrolladas con esta librería, algunas sencillas y otras algo más complejas que se apoyan en herramientas como JavaScript, HTML o CSS.

El interés principal de estas aplicaciones es su interactividad, que permite al usuario utilizar R sin necesidad de crear o modificar código. Además, pueden ser utilizadas desde el ordenador, tablet o móvil.

Estas aplicaciones pueden ser muy útiles en el ámbito investigador o profesional, ya que pueden facilitar el uso de herramientas estadísticas a usuarios no expertos y ahorrar mucho tiempo de desarrollo de código. También puede ser útil en el ámbito docente el desarrollo de aplicaciones que permitan repasar conceptos y practicar lo aprendido de manera interactiva.

En general es una buena herramienta en cualquier proceso que pueda automatizarse con código de R.

2.3.1 Reactividad

(Mulero, s. f.; Wickham, 2021; Hollander et al., 2013)

La programación reactiva es un paradigma que se basa en la propagación de cambios en valores de datos que afectan al valor de otros datos. Al contrario del entorno de programación imperativo en el cual $a = b + c$ quiere decir que a a se le asigna el valor de $b + c$ en el momento de ejecución de esa sentencia y el posterior cambio de los valores de b o c no afecta al valor de a ; en el entorno de programación reactiva el valor de a se actualiza de manera automática al variar el valor de b o c , sin la necesidad de ejecutar de nuevo de forma explícita la sentencia $a = b + c$.

Debido a la interactividad buscada en las aplicaciones Shiny este paradigma de programación es muy adecuado.

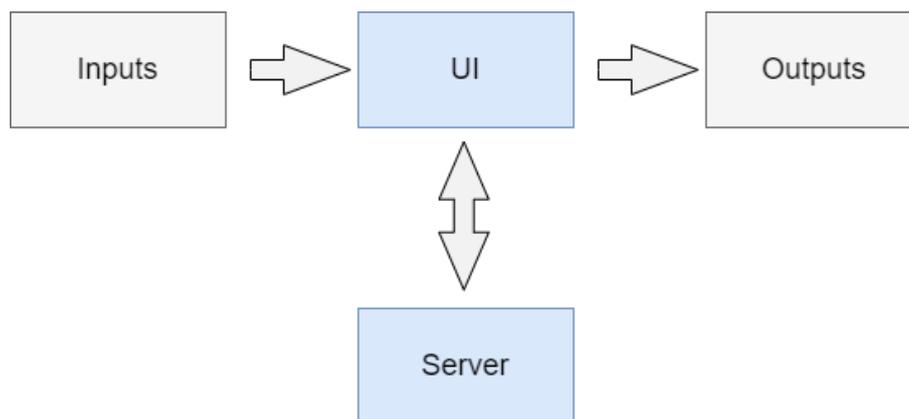


Ilustración 5: Esquema básico de una aplicación Shiny.

Fuente: Elaboración propia (diagrams.net)

En la Ilustración 5 se muestra el esquema básico de una aplicación Shiny, donde UI es la interfaz de usuario (recibe los *inputs* y muestra los *outputs*), Server el servidor (realiza los cálculos) e Inputs y Outputs las entradas y salidas respectivamente.

En las aplicaciones Shiny hay tres tipos de elementos relacionados con la programación reactiva:

- Fuentes Reactivas: Inputs introducidos en UI que se envían al Server.
- Conductores Reactivos: Transformación de Inputs que se utilizan en Server.
- Puntos Finales Reactivos: Outputs obtenidos en Server que se envían a UI.

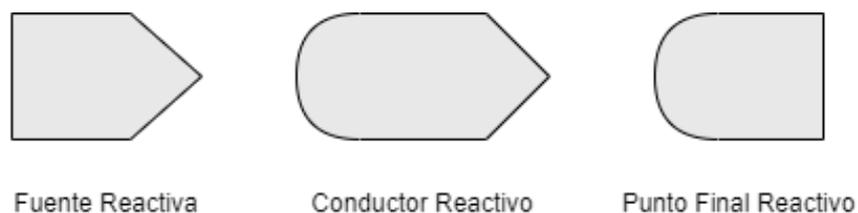


Ilustración 6: Elementos reactivos básicos.

Fuente: Elaboración propia (diagrams.net)

Los esquemas reactivos se utilizan para representar el diagrama de flujo de datos en las aplicaciones interactivas.

El esquema más simple es aquel con una fuente reactiva como entrada y un punto final reactivo como resultado obtenido a partir de la entrada.



Ilustración 7: Esquema reactivo más simple

Fuente: Elaboración propia (diagrams.net)

La librería Shiny ofrece distintas opciones a la hora de controlar la reactividad en las aplicaciones. Algunas de las funciones que ofrece son:

- `reactive()` : Construye un objeto reactivo, que actualiza su valor si alguno de los elementos que lo forma cambia. Dentro de esta función se escribe entre llaves el código que forma el objeto. Este código es el que se ejecuta cada vez que se actualiza su valor.
- `isolate()` : Construye un objeto no reactivo, es decir, que no actualiza su valor aunque haya cambios posteriores en los valores de los elementos que lo forman.
- `observeEvent()` : Desencadena la ejecución de cierto código en Server si se modifica un valor reactivo concreto. El resto de valores reactivos que se encuentren en ese bloque de código no afectarán a la re-ejecución del mismo.
- `eventReactive()` : Construye un objeto reactivo que únicamente actualiza su valor si varía un valor reactivo específico. El resto de valores reactivos no influyen en la actualización del objeto.

Estas expresiones reactivas se llaman como si fuesen una función, escribiendo los paréntesis. Ej: `data ← reactive({rnorm(input$n)}) ; hist(data())`

2.3.2 Inputs y Outputs

(RStudio, 2015, 2016; Wickham, 2021)

Inputs

Los inputs son valores que el usuario introduce mediante el uso de uno o varios widgets que se muestren en la UI. Dependiendo del tipo de input que se quiera obtener se utiliza un determinado tipo de widget. (Más información sobre los principales tipos de widgets en el apartado de Interfaz de usuario.)

El uso de estos widgets ofrece al usuario una selección de inputs sencilla e intuitiva, además de evitar múltiples errores que pudiese proceder de la errónea introducción de valores por el usuario, ya que, por lo general, los widgets limitan las opciones que permite seleccionar.

Uno de los parámetros que se necesitan para un widget es el nombre o id con el que se va a identificar el valor introducido por el usuario. Dentro del Server los valores de estos inputs se utilizarán para obtener los valores de los outputs que se necesiten con la lista `input` y el nombre o id del input que se precise. Ej: `input$id`

Las operaciones realizadas en Server con los valores de inputs son de tipo reactivo, al menos que se utilicen funciones como `isolate()` o `eventReactive()`. Por lo tanto, es necesario incluir estas operaciones en funciones reactivas que evalúen el posible cambio de los inputs y actualicen los resultados.

Outputs

Una vez realizadas las operaciones con los inputs se obtienen los valores de uno o varios outputs que pueden ser de distinto tipo y por lo tanto necesitan estar incluidos en una función de renderizado diferente. Las distintas opciones de estas funciones se detallan en el apartado Servidor (2.3.4)

Estas funciones de renderizado se asignan a objetos en la lista `output`.

Ej: `output$salida1 ← renderText(paste0("El valor del input elegido es:", input$id))`

Estos outputs se muestran en UI a través de distintas funciones dependiendo del tipo de objeto.

Dentro de estas funciones debe ir escrito el nombre o *id* que se le asigno previamente al *output* en la lista *output*. Ej: `textOutput("salida1")`

Function	Inserts
<code>dataTableOutput()</code>	an interactive table
<code>htmlOutput()</code>	raw HTML
<code>imageOutput()</code>	image
<code>plotOutput()</code>	plot
<code>tableOutput()</code>	table
<code>textOutput()</code>	text
<code>uiOutput()</code>	a Shiny UI element
<code>verbatimTextOutput()</code>	text

Ilustración 8: Tabla de funciones Output relacionadas con el objeto que insertan en la interfaz.

Fuente: Documentación RStudio, Shiny

2.3.3 Interfaz de usuario

(Mulero, s. f.; RStudio, 2015; Wickham, 2021)

El elemento principal de una aplicación Shiny interactiva es la interfaz de usuario, ésta es el medio por el cual el usuario de la aplicación puede interactuar con el servidor para obtener resultados. Uno de los objetivos principales de esta interfaz es conseguir que la comunicación del usuario con el servidor sea sencilla e intuitiva.

Estructura

La estructura de una interfaz Shiny suele estar compuesta por una función `fluidPage()`, que contiene distintas opciones según el estilo y funcionalidad que se quiera conseguir.

Una de las estructuras más habituales es utilizar `sidebarLayout()` junto con las funciones `sidebarPanel()` y `mainPanel()`. En este panel lateral suelen ir los *widgets* que manipula el usuario para enviar los valores de los *inputs* y en el panel principal suelen ir los resultados elaborados por el servidor.

Además, existen muchas otras opciones de estructuración de la interfaz, por ejemplo:

- `navbarPage()` + `tabpanel()`: Barra de navegación para acceder a distintos bloques de interfaz.
- `conditionalPanel()`: Permite la posibilidad de que algunos bloques de interfaz de usuario solo se muestren si se cumple alguna condición.
- `tabsetPanel()` + `tabpanel()`: Barra de navegación para acceder a distintos bloques de interfaz. Distinto estilo que `navbarPage()`.
- `navlistPanel()` + `tabpanel()`: Similar a `tabsetPanel()` pero en horizontal.

Si se quiere estructurar la interfaz de una forma más precisa se puede segmentar en paneles con la ayuda de ciertas funciones como `fluidRow()` + `column()`, indicando el ancho y otros parámetros.

Contenido

Dentro de los bloques de la estructura de la interfaz se puede añadir distintos tipos de contenido, estos se incluyen principalmente con la ayuda de las funciones incluidas en la lista *tags*, que ayudan al desarrollador a incluir contenido HTML sin conocer este lenguaje.

Algunas de las funciones de la lista *tags* más habituales son:

- `h1()`, `h2()` ...: Títulos, subtítulos

- `a()`: Hipervínculos
- `p()`: Nuevo párrafo
- `em()`: Texto en cursiva
- `strong()`: Texto en negrita
- `br()`: Retorno de carro
- `hr()`: Línea horizontal
- `img()`: Añadir imagen (El archivo debe estar en la carpeta “www” dentro del directorio *App*)

Con el comando `names(tags)` pueden consultarse las funciones incluidas en esta lista.

Si queremos incluir texto plano podemos incluirlo directamente en el código entre comillas sin ninguna función. También pueden incluirse los *widgets* con las funciones para *inputs* (2.3.2), añadiendo distintos parámetros a estos según el tipo de *input* para configurar su apariencia y utilidad.

Shiny ofrece además la posibilidad de incluir segmentos de código HTML directamente con la función `HTML()`.

Si el desarrollador maneja con facilidad el lenguaje HTML tiene la posibilidad de construir la interfaz de usuario completamente con este lenguaje, incluyendo en el directorio de la aplicación dentro de la carpeta “www” un archivo *.html* que deberá unir en el código de la aplicación con el servidor, apoyándose en la función `shinyApp()` y `htmlTemplate()`.

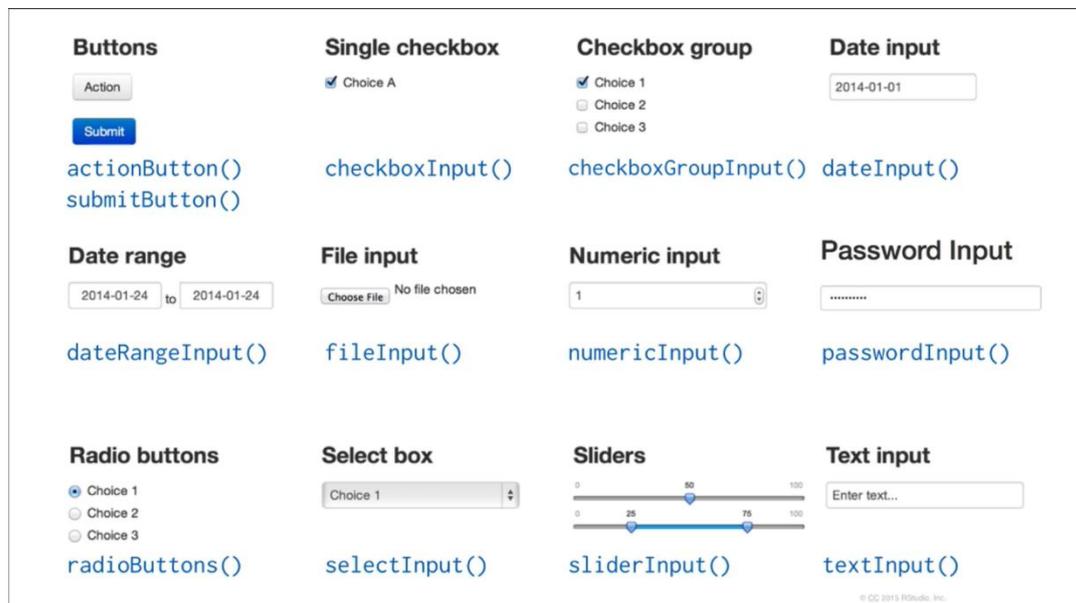


Ilustración 9: Widgets para inputs.

Fuente: Documentación RStudio, Shiny

Apariencia

Si queremos cambiar la apariencia de la aplicación de manera sencilla es posible utilizar la librería *shinythemes*, dándole el valor `shinytheme("estilo")` al parámetro *theme* dentro de una función como `fluidPage()` o `navbarPage()` en el bloque de código de la interfaz de usuario. Algunos de los distintos estilos que ofrece esta librería pueden consultarse en su documentación. (Chang et al., 2015)

Si se desea controlar más el estilo de la aplicación, Shiny permite utilizar CSS (*Cascading Style Sheets*) para customizar la apariencia de los elementos de la aplicación web. (Lie & Bos, 2005)

Lenguaje Matemático

En el contexto de las matemáticas y la estadística puede ser de gran utilidad incluir lenguaje matemático en la interfaz de las aplicaciones. Esto puede solucionarse con la función `withMathJax()` de la propia librería Shiny.

Llamando a esta función al comienzo de la interfaz, podremos escribir fórmulas con la misma sintaxis que en LaTeX, teniendo en cuenta que las expresiones deben encerrarse entre paréntesis con doble barra invertida y que los comandos de LaTeX deben llevar también doble barra invertida en lugar de una.

Ej: `\\(\\frac{ e^{x} }{ n_{1} } \\)` $\equiv \frac{e^x}{n_1}$

2.3.4 Servidor

(RStudio, 2015; Wickham, 2021)

El segundo bloque imprescindible en una aplicación Shiny interactiva es el servidor, donde se reciben los valores de los *inputs* elegidos por el usuario mediante los *widgets* en la interfaz de usuario y se elaboran los *outputs*, que son enviados a la interfaz para mostrárselos al cliente.

Los valores de los *inputs* son utilizados en el servidor con la lista *input* y el nombre que se le dio al definir el *widget* del *input* correspondiente. Ej: `input$n`

Estos valores de *inputs* se utilizan dentro de funciones reactivas que se ejecutan cada vez que alguno de los *inputs* o funciones reactivas utilizadas en su interior cambia su valor. Para utilizar el valor de una función reactiva dentro del propio servidor, esta debe ser llamada incluyendo los paréntesis, ya que se trata de una función.

function	creates
<code>renderDataTable()</code>	An interactive table <small>(from a data frame, matrix, or other table-like structure)</small>
<code>renderImage()</code>	An image (saved as a link to a source file)
<code>renderPlot()</code>	A plot
<code>renderPrint()</code>	A code block of printed output
<code>renderTable()</code>	A table <small>(from a data frame, matrix, or other table-like structure)</small>
<code>renderText()</code>	A character string
<code>renderUI()</code>	a Shiny UI element

Ilustración 10: Tabla de funciones Render relacionadas con el tipo de objeto que crea.

Fuente: Documentación RStudio, Shiny

Cuando se quiera crear un *output* se guarda en la lista *output* con el nombre que se vaya a utilizar en la función de salida de la interfaz de usuario. Estos *outputs* tienen que haberse renderizado con una función de este tipo (Ilustración 10), explicadas en el apartado de *outputs* (2.3.2). Dentro de estas funciones se incluye entre llaves el código necesario para obtener la salida deseada.

En el servidor también se pueden renderizar *outputs* que incluyan lenguaje matemático, hipervínculos... es decir, elementos propios de la interfaz de usuario, utilizando la función de renderizado `renderUI()` combinada con la función `uiOutput()` utilizada en la interfaz. Para el lenguaje matemático es necesario utilizar la función `withMathJax()` en cada uno de los *outputs* que se precise, a pesar de haberla invocado previamente en el bloque de código de la interfaz de usuario.

Otra opción interesante es obtener en el servidor ciertos valores que sirvan como parámetros para algún *input* de la interfaz. Por ejemplo, si se quiere que en un *input* aparezcan los valores de las variables de un conjunto de datos que importa el usuario. Esto se consigue incluyendo el parámetro *session* en la función *server* y utilizando las funciones `observe()` y `updateSelectInput(session, ...)`.

2.4 Shinyapps.io

(RStudio, 2015)

El desarrollo de estas aplicaciones para uso particular puede suponer un elevado ahorro de tiempo en la realización de ciertas tareas laborales o de investigación. Sin embargo, muchas veces se quiere compartir la aplicación con compañeros de trabajo o con cierta comunidad a la que la aplicación web pueda ser de utilidad, especialmente con usuarios que no conozcan el lenguaje R pero les interese utilizar la aplicación.

La mejor solución para esto es subir la aplicación a un servidor en la nube que ejecute la aplicación para cualquier usuario que acceda a ella.

RStudio ofrece una plataforma llamada *shinyapps.io* a la que subir las aplicaciones web de forma gratuita con ciertas limitaciones, como el número máximo de aplicaciones que puedes subir o la cantidad de horas que tu aplicación puede estar activa mensualmente. Pagando una cuota mensual o anual puedes obtener menos limitaciones.

Existen otras opciones de servidores online que den soporte a las aplicaciones, incluso puedes subir las aplicaciones a un servidor propio.

3. RESULTADOS

3.1 App desarrollada

Se ha desarrollado una aplicación web en RStudio con la librería Shiny. Esta aplicación, que recibió el nombre de NPS (*Non Parametric Statistics*), está centrada en la divulgación de las pruebas no paramétricas descritas en el apartado 2.1. y en servir como herramienta a estudiantes, empresas, investigadores y todo aquel al que pueda ser de utilidad, para aplicar estas técnicas a un conjunto de datos propio de manera rápida y sencilla, sin necesidad de aplicar conocimientos de programación. Además de estas pruebas no paramétricas, la aplicación también genera distintos gráficos relacionados con estas pruebas y ofrece la posibilidad de variar distintos parámetros.

3.1.1 Esquema general

NPS consta de cuatro secciones principales: Inicio, Aprende, Datos y Análisis. Es posible navegar a través de los distintos apartados mediante la barra de navegación principal en la zona superior de la interfaz.

La sección Análisis despliega un menú en el que poder seleccionar la prueba no paramétrica deseada.

El aspecto de la aplicación se genera con la librería *shinythemes*, concretamente con el tema “*cosmo*”.

La estructura de ficheros de la aplicación se muestra en la Ilustración 11. En la carpeta principal “*NPSApp*” se encuentran los dos scripts de R, el principal con la interfaz de usuario y el servidor (*App.R*) y uno auxiliar (*funciones.R*) con las librerías necesarias (comprueba que las librerías están instaladas y las instala si es necesario además de cargarlas) y funciones propias que son útiles en ciertos bloques del código. Además, se incluye un fichero “*empleados2.txt*” con el conjunto de datos por defecto, un fichero “*README.txt*” con una pequeña guía de usuario y una carpeta “*www*” donde se incluyen las tablas de

las pruebas no paramétricas en PDF (Anexo 7.1) y el logo de la Universidad de Salamanca en formato PNG.

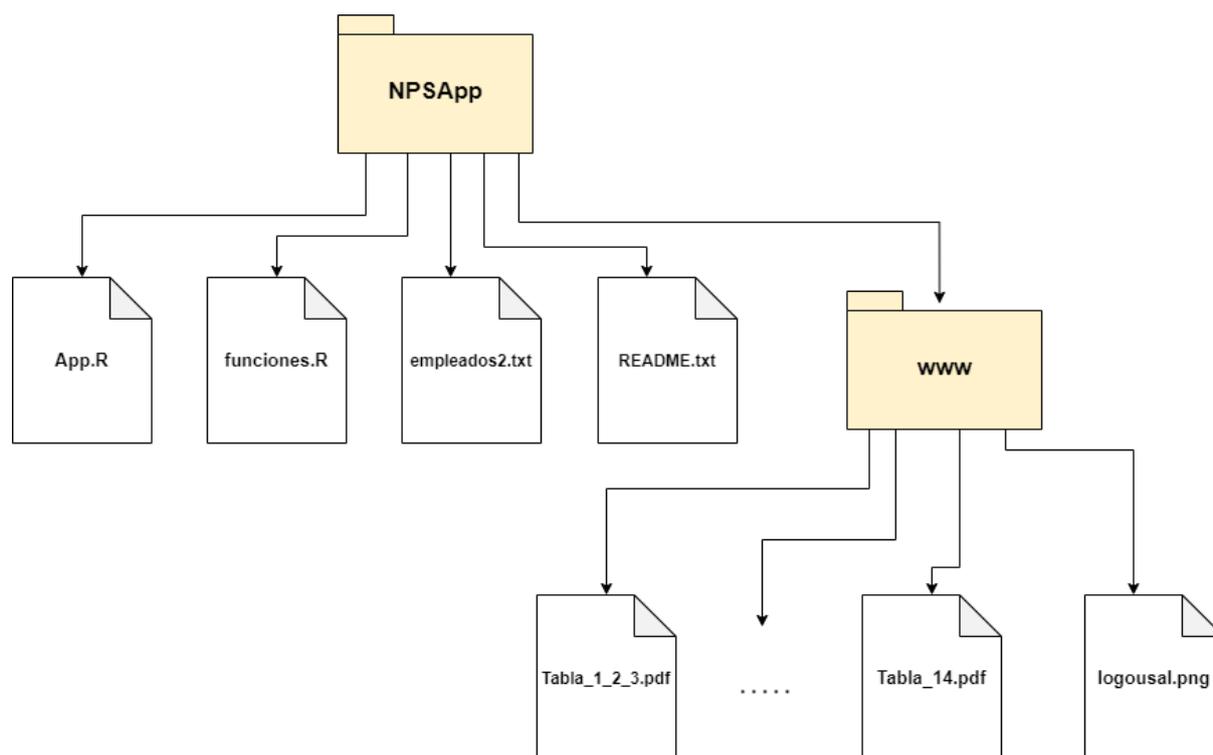


Ilustración 11: Esquema de la estructura de ficheros de NPS

Fuente: Elaboración propia (diagrams.net)

3.1.2 Inicio

La sección Inicio es la primera que se muestra al usuario al abrir NPS. En ella se describe el funcionamiento general de las distintas secciones, así como otros datos de interés como el contacto del desarrollador de la aplicación.

Esta sección no presenta un esquema reactivo ya que no contiene ni inputs ni outputs.

La interfaz de usuario puede consultarse en la Ilustración 12

NPS

¡Bienvenido a NPS!

Esta aplicación web ofrece herramientas para aprender sobre algunas técnicas no paramétricas así como la posibilidad de importar un set de datos y aplicar test no paramétricos a estos datos importados.

A continuación se describe el funcionamiento general de cada uno de los apartados principales de la aplicación.

Aprende

En la sección Aprende se explica el proceso de cada una de las pruebas no paramétricas con un ejemplo práctico, incluyendo las fórmulas utilizadas y enlaces a las tablas de cuantiles necesarias en cada caso. Es posible practicar estos ejercicios las veces que se necesite, ocultando las soluciones y generando nuevos datos. Después, se muestran las soluciones para comprobar que se han realizado las cuentas correctamente y que se ha comprendido el proceso del test no paramétrico en cuestión. Además, este apartado ofrece enlaces en los que podemos encontrar más información sobre las pruebas.

Los tests no paramétricos que se incluyen en este apartado son:

- **Aleatoriedad: Rachas W-W** → Contraste de aleatoriedad: Prueba de las rachas de Wald-Wolfowitz.
- **Bondad de Ajuste: K-S $F = F_0$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = F_0$.
- **Bondad de Ajuste: K-S $F = G$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = G$.
- **Localización de una muestra: W** → Contraste de localización de una muestra: Prueba de los rangos con signo de Wilcoxon.
- **Dos muestras independientes: U M-W** → Contraste para dos muestras independientes: Prueba U de Mann-Whitney.
- **Dos muestras relacionadas: Apareados W** → Contraste para dos muestras relacionadas: Prueba para datos apareados de Wilcoxon.
- **k muestras independientes: K-W** → Contraste para k muestras independientes: Prueba de Kruskal-Wallis.
- **k muestras relacionadas: Friedman** → Contraste para k muestras relacionadas: Prueba de Friedman.
- **Independencia: Kendall** → Contraste de independencia: Prueba de Kendall.
- **Independencia: Spearman** → Contraste de independencia: Prueba de Spearman.

Ilustración 12: Interfaz de NPS, Sección Inicio

Fuente: NPS

3.1.3 Aprende

El objetivo principal de esta sección es ofrecer al usuario un entorno en el que adquirir conocimientos acerca de las principales técnicas no paramétricas, mostrando una breve descripción de la utilidad de la prueba, las hipótesis del contraste, hipervínculos con más información y una breve explicación del funcionamiento de cada apartado.

Se puede acceder a los apartados para cada prueba mediante la barra de navegación de la zona superior de la sección. En cada uno de estos apartados se permite generar una o varias muestras aleatorias seleccionando algunas opciones en el panel lateral, como el tipo de variable o el tamaño muestral. Estas muestras aleatorias son generadas en R con funciones como $rnorm()$, $runif()$ o $rexp()$, que generan una cantidad de datos pseudoaleatorios concreta de una distribución, especificando los valores de los parámetros según el tipo de distribución aleatoria. Estos valores son pseudoaleatorios, ya que el ordenador los produce a partir de un algoritmo y un valor inicial (semilla). Si el valor de la semilla no varía, los valores obtenidos serían los mismos, ya que el ordenador es determinístico y obtiene siempre la misma salida para unos valores de entrada fijos.

En el panel principal se puede ver la muestra aleatoria obtenida y el desarrollo de la prueba no paramétrica correspondiente paso a paso, con los resultados intermedios. Si se quisieran ocultar estos resultados para, por ejemplo, practicar resolviendo la técnica a mano, pueden deshabilitarse las casillas del panel lateral.

Además, en el desarrollo de cada técnica en el panel principal, se ofrecen hipervínculos con las tablas necesarias para el proceso de las pruebas y algunos gráficos que suelen acompañar a cada tipo de contraste.

Las técnicas (apartados) incluidas en esta sección son las siguientes:

- Aleatoriedad: Rachas W-W
- Bondad de Ajuste: K-S ($F=F_0$)
- Bondad de Ajuste: K-S ($F=G$)
- Localización de una muestra: W
- Dos muestras independientes: U M-W
- Dos muestras relacionadas: Apareados W
- K muestras independientes: K-W
- K muestras relacionadas: Friedman

- Independencia: Kendall
- Independencia: Spearman

Ilustración 13: Interfaz de NPS, Sección Aprende

Fuente: NPS

3.1.3.1 Aleatoriedad: Rachas W-W

El apartado *Aleatoriedad: Rachas W-W* ofrece una breve introducción a la prueba de las rachas de *Wald-Wolfowitz* (2.1.2.1), las hipótesis del contraste en cuestión y un hipervínculo con más información acerca del test. Parte de la interfaz de este apartado puede verse en la Ilustración 13.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variable aleatoria entre el lanzamiento de una moneda con la opción “Lanzar una Moneda” (Variable cualitativa nominal) y una secuencia numérica que proviene de una distribución uniforme $U(-45,45)$ con la opción “Secuencia Numérica” (Variable cuantitativa continua). También es posible decidir el tamaño de la muestra aleatoria, seleccionando un valor en el botón deslizante “Tamaño de la muestra”, entre 5 y 25.

Una vez elegidos los valores deseados de la nueva muestra aleatoria se acciona el botón “Nueva Muestra” y se mostrará en el panel principal la nueva muestra aleatoria obtenida. Debajo de la muestra se encuentra el procedimiento de la prueba de las rachas de *Wald-Wolfowitz* con los datos de la muestra obtenida, es decir, se pueden generar cuantas veces sea necesario ejercicios resueltos de este test. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 7 (R), 1, 2 y 3 (Normal) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un gráfico de dispersión de la muestra ordenada con una línea roja que muestra el valor de la mediana calculada para dicotomizar la muestra en el caso de que la variable sea cuantitativa. Para la variable cualitativa se asigna a una categoría el valor uno y a las demás categorías el valor cero.

Si el usuario desea resolver por sí mismo la prueba para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas del panel lateral. De nuevo, este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tipo de variable y/o el tamaño muestral.

El esquema reactivo de este apartado, utilizando la notación de la Ilustración 6, se muestra en la Ilustración 14.

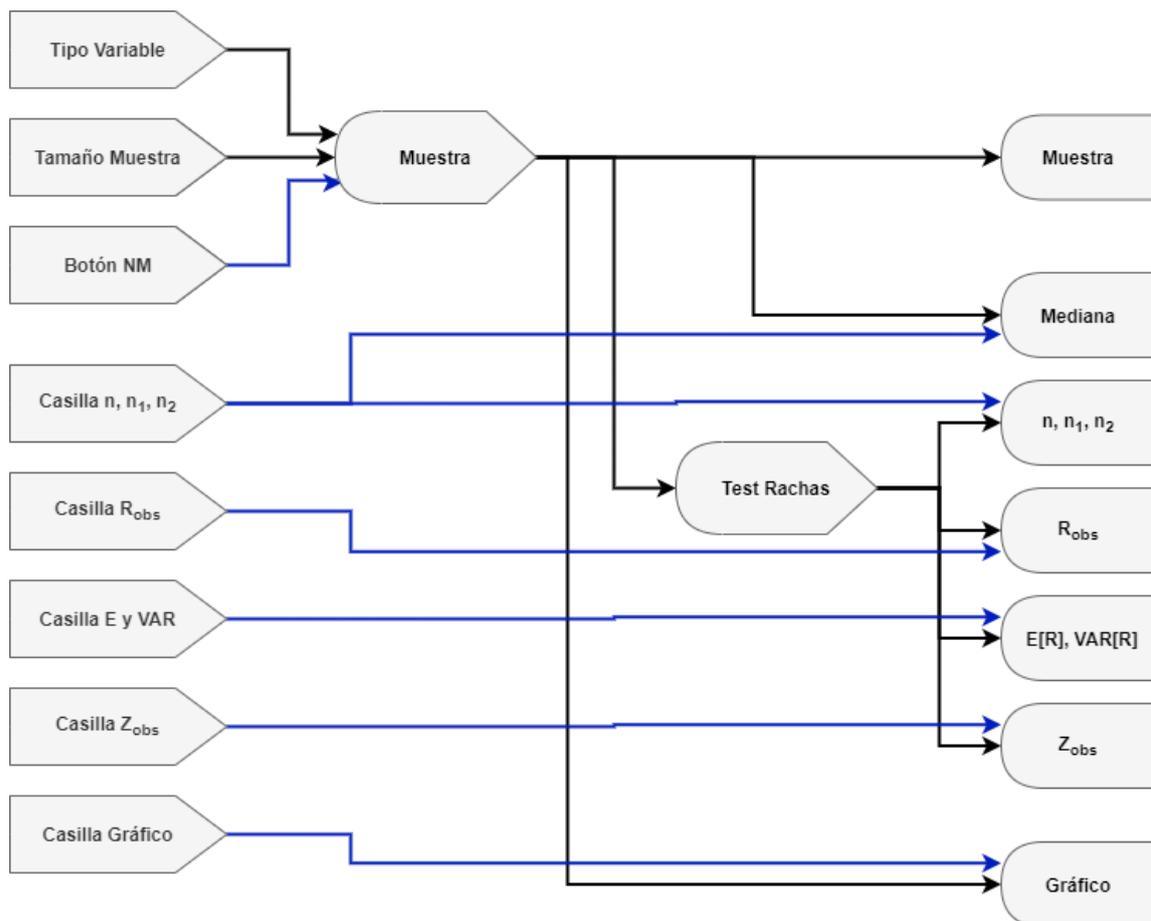


Ilustración 14: Esquema Reactivo del apartado Aleatoriedad: Rachas W-W de la sección Aprende.

Fuente: Elaboración propia (diagrams.net)

Las flechas azules en el esquema reactivo indican que el valor del elemento reactivo del que proceden afecta a la reactividad del elemento al que llega, es decir, para el caso de la fuente reactiva “Botón Nueva Muestra”, si ésta no es alterada, los cambios en las fuentes reactivas “Tipo Variable” y “Tamaño Muestra” no provocan un cambio en el conductor reactivo Muestra ni en los elementos reactivos que dependan de éste. Para el caso de las fuentes reactivas “Casilla ...”, la alteración de cada una afecta a distintos puntos finales reactivos, controlando la ejecución y obtención de los resultados parciales y finales de la prueba y por tanto controlando también si estos se muestran o no en el panel principal.

Algunos de estos puntos finales reactivos están además controlados por otras condiciones, por ejemplo, en el caso del punto final reactivo “Mediana”, éste solo se ejecuta si los valores del conductor reactivo “Muestra” son numéricos, es decir, si se trata de una muestra aleatoria cuantitativa.

3.1.3.2 Bondad de Ajuste: K-S ($F = F_0$)

El apartado *Bondad de Ajuste: K-S ($F = F_0$)* ofrece una breve introducción al test de bondad de ajuste de *Kolmogorov-Smirnov* (2.1.3.1), las hipótesis del contraste en cuestión, la definición de distribución empírica y dos hipervínculos con más información acerca del test y la distribución empírica. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.3 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tamaño de la muestra aleatoria, seleccionando un valor en el botón deslizante “Tamaño de la muestra”, entre 6 y 20.

Una vez elegido el tamaño muestral deseado de la nueva muestra aleatoria se acciona el botón “Nueva Muestra” y se mostrará en el panel principal la nueva muestra aleatoria obtenida, acompañada de un enunciado. Esta muestra es generada mediante una función propia, que genera un valor de λ , de μ y de σ , donde $\lambda \in U(1,5)$, $\mu \in U(-50,50)$ y $\sigma \in U(1,8)$, los valores obtenidos de λ , μ y σ se redondean con cero decimales. Estos valores se utilizan para generar la cantidad de datos que el usuario haya elegido, procedentes de una distribución concreta al azar entre:

- $x \in Exp(\lambda)$
- $x \in Exp(\lambda + 3)$
- $x \in N(\mu, \sigma)$
- $x \in N(\mu + 5, \sigma + 2)$

donde x es la muestra aleatoria de tamaño elegido por el usuario. (Las cuatro opciones son equiprobables).

Los valores obtenidos se redondean con dos decimales para facilitar los posibles cálculos al usuario.

El enunciado se genera dependiendo del tipo de distribución que siga la muestra generada. Si la distribución es Exponencial, el enunciado generado pide comprobar si los datos de la muestra provienen de una distribución $E(\lambda)$, donde λ toma el valor obtenido al generar la muestra. Del mismo modo, si la distribución es Normal, el enunciado generado pide comprobar si los datos de la muestra provienen de una distribución $N(\mu, \sigma)$, donde μ y σ toman el valor obtenido al generar la muestra.

De este modo, un 50% de las veces se generarán contrastes que la mayor parte de las veces sean no significativos ($x \in Exp(\lambda)$ y $x \in N(\mu, \sigma)$), ya que coinciden los valores de los parámetros de las distribuciones utilizadas para generar la muestra aleatoria con los valores de los parámetros de la distribución del enunciado. El otro 50% de las veces, se generarán pruebas que sean mayoritariamente significativas ($x \in Exp(\lambda + 3)$ y $x \in N(\mu + 5, \sigma + 2)$), ya que los valores de los parámetros utilizados para generar la muestra no coinciden con los valores de los parámetros utilizados para generar el enunciado.

Debajo de la muestra y del enunciado se encuentra el procedimiento de la prueba de bondad de ajuste de *Kolmogorov-Smirnov* para el caso $F = F_0$ con los datos de la muestra obtenida y el enunciado asociado, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de este contraste. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a la tabla 6 (KS) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un gráfico Cuantil-Cuantil (2.1.3.2) donde se comparan los cuantiles de la muestra con los cuantiles de la distribución teórica conocida y un histograma de la muestra con una línea roja que muestra la distribución de probabilidad de la distribución teórica conocida.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas del panel lateral. De nuevo, este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño muestral y, de forma aleatoria, obteniendo distintas distribuciones y enunciados.

El esquema reactivo es similar al del apartado anterior y al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.2.

3.1.3.3 Bondad de Ajuste: K-S ($F = G$)

El apartado *Bondad de Ajuste: K-S ($F=G$)* ofrece una breve introducción a la prueba de bondad de ajuste de *Kolmogorov-Smirnov* (2.1.3.1), las hipótesis del contraste en cuestión, la definición de distribución

empírica y dos hipervínculos con más información acerca de la prueba y la distribución empírica. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.4 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tamaño de las muestras aleatorias, seleccionando un valor en cada uno de los botones deslizantes “Tamaño de la muestra X” y “Tamaño de la muestra Y”, ambos entre 6 y 20.

Una vez elegidos los tamaños de ambas muestras aleatorias se acciona el botón “Nuevas Muestras” y se mostrará en el panel principal las dos nuevas muestras aleatorias obtenidas. Estas muestras se obtienen mediante la misma función propia del apartado anterior cambiando el parámetro de entrada “c” por un dos (Dos muestras). Los parámetros utilizados para generar las muestras son los mismos que en el caso de una muestra, $\lambda \in U(1,5)$, $\mu \in U(-50,50)$ y $\sigma \in U(1,8)$, redondeando estos valores con cero decimales. Estos valores se utilizan para generar dos muestras aleatorias del tamaño seleccionado por el usuario procedentes de unas distribuciones de probabilidad concretas al azar entre:

- $x \in N(\mu, \sigma)$, $y \in N(\mu, \sigma)$
- $x \in Exp(\lambda)$, $y \in Exp(\lambda)$
- $x \in N(\mu + 5, \sigma + 2)$, $y \in N(\mu, \sigma)$

donde x e y son la primera y la segunda muestra aleatoria respectivamente, ambas del tamaño elegido por el usuario. (Las tres opciones son equiprobables).

Los valores obtenidos en ambas muestras se redondean con dos decimales para facilitar los posibles cálculos al usuario.

De este modo, dos tercios de las veces se generarán contrastes que la mayor parte de las veces sean no significativos ($x \in N(\mu, \sigma)$ e $y \in N(\mu, \sigma)$ o $x \in Exp(\lambda)$ e $y \in Exp(\lambda)$), ya que en ambas muestras coinciden las distribuciones de procedencia, así como los valores de los parámetros de estas distribuciones. El otro tercio de las veces, se generarán pruebas que sean mayoritariamente significativas ($x \in N(\mu + 5, \sigma + 2)$ e $y \in N(\mu, \sigma)$), ya que en las dos muestras coinciden las distribuciones de procedencia, pero no los valores de los parámetros de estas distribuciones.

Debajo de las muestras se encuentra el procedimiento de la prueba de bondad de ajuste de *Kolmogorov-Smirnov* para el caso $F = G$, con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a la tabla 6 (KS) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un gráfico Cuantil-Cuantil (2.1.3.2) donde se comparan los cuantiles de la muestra x con los cuantiles de la muestra y .

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas del panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño muestral de ambas muestras y, de forma aleatoria, obteniendo distintas distribuciones.

El esquema reactivo es similar al del apartado anterior y al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.3.

3.1.3.4 Localización de una muestra: W

El apartado *Localización de una muestra: W* ofrece una breve introducción a la prueba de los rangos con signo de *Wilcoxon* (2.1.4), los tres posibles tipos de hipótesis del contraste y un hipervínculo con más información acerca de la prueba. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.5 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tamaño de la muestra aleatoria, seleccionando un valor en el botón deslizante “Tamaño de la muestra”, entre 10 y 20.

Una vez elegido el tamaño muestral deseado de la nueva muestra aleatoria se acciona el botón “Nueva Muestra” y se mostrará en el panel principal la nueva muestra aleatoria obtenida, acompañada de un enunciado. Esta muestra proviene de una distribución uniforme $x \in U(-45,45)$. Posteriormente se redondean estos valores muestrales obtenidos con dos decimales para facilitar los posibles cálculos del usuario.

El enunciado se genera de forma aleatoria:

- Contraste bilateral $H_0: \theta = \theta_0$
- Contraste unilateral superior $H_0: \theta \leq \theta_0$
- Contraste unilateral inferior $H_0: \theta \geq \theta_0$

donde $\theta_0 = Me(x) + U(-10,10)$, siendo $U(-10,10)$ un valor de una distribución uniforme discreta y redondeando con cero decimales el valor de la mediana calculada. (Las tres opciones son equiprobables).

De esta manera, un tercio de las veces se generan contrastes bilaterales, otro tercio contrastes unilaterales superiores y el otro tercio contrastes unilaterales inferiores. La forma de calcular θ_0 se debe a la finalidad de que los contrastes salgan no significativos con mayor facilidad.

Debajo de la muestra y el enunciado se encuentra el procedimiento de la prueba de los rangos con signo de *Wilcoxon* con los datos de la muestra obtenida y el valor de θ_0 , es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de este contraste. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 8 (T^+), 1, 2 y 3 (Normal) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un Box-Plot de la muestra con una línea en rojo que señala el valor de θ_0 .

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas del panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño muestral y, de forma aleatoria, obteniendo nuevos enunciados con distintos tipos de contraste y diferentes valores de θ_0 .

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.4.

3.1.3.5 Dos muestras independientes: U M-W

El apartado *Dos muestras independientes: U M-W* ofrece una breve introducción a la prueba *U* de *Mann-Whitney* (2.1.5.2), los tres posibles tipos de hipótesis del contraste, y un hipervínculo con más información acerca del test. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.6 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variables aleatorias entre dos variables cuantitativas con la opción “Variables Cuantitativas” o dos cualitativas ordinales con la opción “Variables Ordinales”. También es posible seleccionar el tamaño de las muestras aleatorias, eligiendo un valor en cada uno de los botones deslizantes “Tamaño de la muestra X_1 ” y “Tamaño de la muestra X_2 ”, ambos entre 6 y 20.

Una vez elegidos los valores deseados de las nuevas muestras aleatorias se acciona el botón “Nuevas Muestras” y se mostrará en el panel principal las nuevas muestras aleatorias obtenidas. Si el tipo de variables elegidas fue cuantitativo, las muestras provienen de distribuciones uniformes redondeando sus valores con un decimal $x_1 \in U(8,25)$ y $x_2 \in U(6,21)$, estas son distintas para que el contraste sea significativo con mayor frecuencia. Si el tipo de variables elegidas fue ordinal, las muestras provienen de distribuciones uniformes discretas que representan con cada número una categoría ordenada $x_1 \in U(1,4)$ y $x_2 \in U(1,7)$, estas son distintas para que el contraste sea significativo con mayor

frecuencia, sin embargo, al tratarse de muestras independientes de la misma variable ambas muestran tienen siete categorías, aunque en la primera muestra solo aparezcan cuatro de estas siete.

Debajo de las muestras se encuentra el procedimiento de la prueba para dos muestras independientes U de *Mann-Whitney* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 9 (U), 1, 2 y 3 (Normal) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestran dos Box-Plot, uno para cada muestra.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de ambas muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.5.

3.1.3.6 Dos muestras relacionadas: Apareados W

El apartado *Dos muestras relacionadas*: W ofrece una breve introducción a la prueba de los rangos con signo de *Wilcoxon* (2.1.5.1), los tres posibles tipos de hipótesis del contraste, y un hipervínculo con más información acerca del test. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.7 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variables aleatorias entre dos variables cuantitativas con la opción “Variables Cuantitativas” o dos cualitativas ordinales con la opción “Variables Ordinales”. También es posible seleccionar el tamaño de las muestras aleatorias, eligiendo un valor entre 6 y 20 en el botón deslizante “Tamaño de la muestra”. Este tamaño será el mismo para las dos muestras, ya que se trata de muestras relacionadas/pareadas.

Una vez elegidos los valores deseados de las nuevas muestras aleatorias se acciona el botón “Nuevas Muestras” y se mostrará en el panel principal las nuevas muestras aleatorias obtenidas. Si el tipo de variables elegidas fue cuantitativo, las muestras provienen de distribuciones uniformes redondeando sus valores con un decimal $x_1 \in U(8,25)$ y $x_2 \in U(6,21)$, estas son distintas para que el contraste sea significativo con mayor frecuencia. Si el tipo de variables elegidas fue ordinal, las muestras provienen de distribuciones uniformes discretas que representan con cada número una categoría ordenada $x_1 \in U(1,4)$ y $x_2 \in U(1,7)$, estas son distintas para que el contraste sea significativo con mayor frecuencia, sin embargo, al tratarse de muestras relacionadas de la misma variable ambas muestran tienen siete categorías, aunque en la primera muestra solo aparezcan cuatro de estas siete.

Debajo de las muestras se encuentra el procedimiento de la prueba para dos muestras relacionadas de los rangos con signo de *Wilcoxon* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 8 (T^+), 1, 2 y 3 (Normal) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestran dos Box-Plot, uno para cada muestra.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de las muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.6.

3.1.3.7 K muestras independientes: K-W

El apartado *k muestras independientes: K-W* ofrece una breve introducción a la prueba de *Kruskal-Wallis* (2.1.6.1), las hipótesis del contraste y un hipervínculo con más información acerca del test. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.8 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variable aleatoria entre variable cuantitativa con la opción “Variables Cuantitativas” o cualitativa ordinal con la opción “Variables Ordinales”. También es posible seleccionar el tamaño de las muestras aleatorias, eligiendo un valor en cada uno de los botones deslizantes “Tamaño de la muestra X_1 ”, “Tamaño de la muestra X_2 ” y “Tamaño de la muestra X_3 ”, todos entre 3 y 20.

Una vez elegidos los valores deseados de las nuevas muestras aleatorias se acciona el botón “Nuevas Muestras” y se mostrará en el panel principal las nuevas muestras aleatorias obtenidas. Si el tipo de variables elegidas fue cuantitativo, las muestras provienen de distribuciones uniformes redondeando sus valores con un decimal $x_1 \in U(8,25)$, $x_2 \in U(6,21)$ y $x_3 \in U(11,27)$, estas son distintas para que el contraste sea significativo con mayor frecuencia. Si el tipo de variables elegidas fué ordinal, las muestras provienen de distribuciones uniformes discretas que representan con cada número una categoría ordenada $x_1 \in U(1,4)$, $x_2 \in U(1,7)$ y $x_3 \in U(1,5)$, estas son distintas para que el contraste sea significativo con mayor frecuencia, sin embargo, al tratarse de muestras independientes de la misma variable, todas las muestras tienen siete categorías, aunque en la primera muestra solo aparezcan cuatro y en la tercera únicamente cinco de las siete categorías posibles.

Debajo de las muestras se encuentra el procedimiento de la prueba para *k* muestras independientes de *Kruskal-Wallis* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 10 (H) y 4 (χ^2) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestran tres Box-Plot, uno para cada muestra.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de las muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.7.

3.1.3.8 K muestras relacionadas: Friedman

El apartado *k muestras relacionadas: Friedman* ofrece una breve introducción a la prueba para *k* muestras relacionadas de *Friedman* (2.1.6.2), las hipótesis del contraste y un hipervínculo con más información acerca de la prueba. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.9 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variable aleatoria entre variable cuantitativa con la opción “Variables Cuantitativas” o cualitativa ordinal con la opción “Variables Ordinales”. También es posible seleccionar el tamaño de las muestras aleatorias, eligiendo un valor entre 3 y 20 en el botón deslizante “Tamaño de la muestra”. Este tamaño será el mismo para las tres muestras, ya que se trata de muestras relacionadas/pareadas.

Una vez elegidos los valores deseados de las nuevas muestras aleatorias se acciona el botón “Nuevas Muestras” y se mostrará en el panel principal las nuevas muestras aleatorias obtenidas. Si el tipo de variables elegidas fue cuantitativo, las muestras provienen de distribuciones uniformes redondeando sus valores con un decimal $x \in U(8,25)$, $y \in U(6,21)$ y $z \in U(11,27)$, estas son distintas para que el

contraste sea significativo con mayor frecuencia. Si el tipo de variables elegidas fue ordinal, las muestras provienen de distribuciones uniformes discretas que representan con cada número una categoría ordenada $x \in U(1,4)$, $y \in U(1,7)$ y $z \in U(1,5)$, estas son distintas para que el contraste sea significativo con mayor frecuencia, sin embargo, al tratarse de muestras relacionadas de la misma variable, todas las muestras tienen siete categorías, aunque en la primera muestra solo aparezcan cuatro y en la tercera únicamente cinco de las siete categorías posibles.

Debajo de las muestras se encuentra el procedimiento de la prueba para k muestras relacionadas de *Friedman* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 11 (S) y 4 (χ^2) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestran tres Box-Plot, uno para cada muestra.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de las muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.8.

3.1.3.9 Independencia: Kendall

El apartado *Independencia: Kendall* ofrece una breve introducción al contraste de independencia de *Kendall* (2.1.7.1), las hipótesis del contraste en cuestión y un hipervínculo con más información acerca del test. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.10 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variables entre: dos variables cuantitativas con la opción “Variables Cuantitativas”, dos variables cualitativas ordinales con la opción “Variables Ordinales” y una variable cuantitativa y otra cualitativa ordinal con la opción “Variable Cuantitativa y Ordinal”. También es posible elegir el tamaño de la muestra aleatoria, seleccionando un valor en el botón deslizante “Tamaño de la muestra”, entre 6 y 15. Las dos muestras serán del mismo tamaño.

Una vez elegido el tipo de variables y el tamaño muestral deseado de las nuevas muestras aleatorias, se acciona el botón “Nuevas Muestras” y se mostrarán en el panel principal las nuevas muestras aleatorias obtenidas. Estas muestras se generan dependiendo del tipo de variables elegidas y del azar de la siguiente manera:

Si las variables elegidas fueron cuantitativas:

- $x \in U(5,20)$, $y \in U(3,7)$
- $x \in U(1,12)$, $y = \frac{x^2}{5} + \varepsilon$ donde $\varepsilon = U(-3,3)$

los valores finales de x e y son redondeados con un decimal para facilitar los posibles cálculos al usuario. Ambos casos son equiprobables. En el segundo caso se fuerza la concordancia entre las variables para aumentar la frecuencia de contrastes significativos, calculando los valores de la muestra y con los valores de la muestra x , añadiendo un error ε para que la concordancia no sea perfecta y no se pierda la sensación de aleatoriedad.

Si las variables elegidas fueron cualitativas ordinales:

- $x \in U(1,4)$, $y \in U(1,7)$
- $x \in U(1,4)$, $y = x + 1 + \varepsilon$ donde $\varepsilon = U(-1,1)$

las distribuciones utilizadas en este caso para obtener las muestras son uniformes discretas. Ambas opciones son equiprobables. En el segundo caso se fuerza la concordancia de las variables calculando los valores de la muestra y y con los valores de la muestra x , añadiendo un error ε para que la concordancia no sea perfecta y no se pierda la sensación de aleatoriedad.

Si las variables elegidas fueron una cualitativa ordinal y otra cuantitativa:

- $x \in U(5,20)$, $y \in U(1,4)$
- $x \in U(5,20)$, $y = \frac{x}{5} + 1 + \varepsilon$ donde $\varepsilon = U(-1,1)$

las distribuciones utilizadas en este caso para obtener las muestras son uniformes continuas para las muestras x , y uniformes discretas para la muestra y del primer caso y para ε . Los valores de las muestras cuantitativas se redondean con un decimal para facilitar los posibles cálculos al usuario. Ambas opciones son equiprobables. En el segundo caso se fuerza la dependencia de las variables calculando los valores de la muestra y y con los valores de la muestra x , añadiendo un error ε para que la concordancia no sea perfecta y no se pierda la sensación de aleatoriedad. Los valores de y en el segundo caso se redondean con cero decimales.

De esta manera, independientemente del tipo de variables elegidas, la mitad de las veces se generarán contrastes que mayoritariamente sean no significativos. La otra mitad de las veces se obtendrán contrastes mayoritariamente significativos, ya que al generar las muestras, una se formó a partir de la otra.

Debajo de las muestras se encuentra el procedimiento de la prueba de independencia de *Kendall* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 12 (τ), 1, 2 y 3 (Normal) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un gráfico de dispersión de las dos muestras.

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de las muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.9.

3.1.3.10 Independencia: Spearman

El apartado *Independencia: Spearman* ofrece una breve introducción al contraste de independencia de *Spearman* (2.1.7.2), las hipótesis del contraste en cuestión y un hipervínculo con más información acerca del test. La interfaz de este apartado es similar a la del resto de apartados de la sección Aprende. Ésta puede consultarse en el Anexo 7.2.11 o en la propia aplicación web.

En el panel lateral se ofrece la posibilidad de seleccionar el tipo de variables entre: dos variables cuantitativas con la opción “Variables Cuantitativas”, dos variables cualitativas ordinales con la opción “Variables Ordinales” y una variable cuantitativa y otra cualitativa ordinal con la opción “Variable Cuantitativa y Ordinal”. También es posible elegir el tamaño de la muestra aleatoria, seleccionando un valor en el botón deslizante “Tamaño de la muestra”, entre 6 y 25. Las dos muestras serán del mismo tamaño.

Una vez elegido el tipo de variables y el tamaño muestral deseado de las nuevas muestras aleatorias, se acciona el botón “Nuevas Muestras” y se mostrarán en el panel principal las nuevas muestras aleatorias obtenidas. Estas muestras se generan dependiendo del tipo de variables elegidas y del azar de la siguiente manera:

Si las variables elegidas fueron cuantitativas:

- $x \in U(5,20)$, $y \in U(3,7)$
- $x \in U(5,20)$, $y = -14 \cdot \cos x + 3x + \varepsilon$ donde $\varepsilon = U(-5,5)$

los valores finales de x e y son redondeados con un decimal para facilitar los posibles cálculos al usuario. Ambos casos son equiprobables. En el segundo caso se fuerza la relación entre las variables para aumentar la frecuencia de contrastes significativos, calculando los valores de la muestra y con los valores de la muestra x , añadiendo un error ε para que la correlación no sea perfecta y no se pierda la sensación de aleatoriedad.

Si las variables elegidas fueron cualitativas ordinales:

- $x \in U(1,4)$, $y \in U(1,7)$
- $x \in U(1,4)$, $y = x + 1 + \varepsilon$ donde $\varepsilon = U(-1,1)$

las distribuciones utilizadas en este caso para obtener las muestras son uniformes discretas. Ambas opciones son equiprobables. En el segundo caso se fuerza la correlación de las variables calculando los valores de la muestra y con los valores de la muestra x , añadiendo un error ε para que la correlación no sea perfecta y no se pierda la sensación de aleatoriedad.

Si las variables elegidas fueron una cualitativa ordinal y otra cuantitativa:

- $x \in U(5,20)$, $y \in U(1,7)$
- $x \in U(5,20)$, $y = \frac{x}{5} + 1 + \varepsilon$ donde $\varepsilon = U(-1,1)$

las distribuciones utilizadas en este caso para obtener las muestras son uniformes continuas para las muestras x , y uniformes discretas para la muestra y del primer caso y para ε . Los valores de las muestras cuantitativas se redondean con un decimal para facilitar los posibles cálculos al usuario. Ambas opciones son equiprobables. En el segundo caso se fuerza la dependencia de las variables calculando los valores de la muestra y con los valores de la muestra x , añadiendo un error ε para que la correlación no sea perfecta y no se pierda la sensación de aleatoriedad. Los valores de y en el segundo caso se redondean con cero decimales.

De esta manera, independientemente del tipo de variables elegidas, la mitad de las veces se generarán contrastes que mayoritariamente sean no significativos. La otra mitad de las veces se obtendrán contrastes mayoritariamente significativos, ya que, al generar las muestras, una se formó a partir de la otra.

Debajo de las muestras se encuentra el procedimiento de la prueba de independencia de *Spearman* con los datos de las muestras obtenidas, es decir, como en el resto de los apartados de la sección Aprende, se pueden generar cuantas veces sea necesario ejercicios resueltos de esta prueba. Además, a lo largo del desarrollo hay enlaces a las tablas necesarias, en este caso a las tablas 13 (RS) y 14 (*t-Student*) y las fórmulas necesarias para obtener los resultados parciales de la prueba. También se muestra un gráfico de dispersión de las dos muestras. El cuantil que aparece de la distribución *t-Student* se actualiza cuando varía el tamaño muestral y se vuelve a activar el botón “Nueva Muestra” (ya que el tamaño muestral afecta a sus grados de libertad).

Si el usuario desea resolver por sí mismo el ejercicio para posteriormente comprobar los resultados en la aplicación, puede deshabilitar las soluciones en el panel lateral en la sección “Mostrar Soluciones”, desmarcando las casillas correspondientes. Una vez resuelto el ejercicio, puede comprobar si el procedimiento es correcto marcando de nuevo las casillas en el panel lateral. Este proceso puede repetirse cuantas veces crea necesario el usuario, generando nuevas muestras/ejercicios en el panel lateral, variando el tamaño de las muestras y el tipo de variables.

El esquema reactivo es similar al de los demás apartados de la sección Aprende. Éste puede consultarse en el Anexo 7.3.10.

3.1.4 Datos

El objetivo principal de la sección Datos es permitir al usuario seleccionar el conjunto de datos que quiere analizar en la sección Análisis.

En el panel lateral el usuario puede seleccionar en el menú “Dataset” si quiere utilizar el conjunto de datos por defecto llamado “Empleados” o quiere importar un archivo desde su dispositivo.

La base de datos de “Empleados” consta de doce variables medidas en 64 individuos. Estos valores fueron obtenidos con una encuesta realizada por una empresa a sus trabajadores. Las doce variables de este set de datos son:

- sexo: Género del trabajador (Hombre, Mujer) (V. Cualitativa Nominal)
- ecivil: Estado civil del trabajador (Casado, Soltero) (V. Cualitativa Nominal)
- edad: Edad del trabajador (V. Cuantitativa Discreta)
- antigüedad: Antigüedad en años del trabajador en la empresa (V. Cuantitativa Discreta)
- salario2006: Salario anual del trabajador en 2006 (V. Cuantitativa Continua)
- salario2008: Salario anual del trabajador en 2008 (V. Cuantitativa Continua)
- salario2010: Salario anual del trabajador en 2010 (V. Cuantitativa Continua)
- evaluacion: Valoración de la empresa sobre el trabajador (Mala, Regular, Buena, Muy buena, Excelente) (V. Cualitativa Ordinal)
- dias_trabajados: Días trabajados (V. Cuantitativa Discreta)
- zona_trabajo: Zona de la empresa en la que trabaja el empleado (Norte, Sur, Este, Oeste) (V. Cualitativa Nominal)
- estudios: Nivel de estudios del empleado (Primarios, Bachillerato, Universitarios) (V. Cualitativa Ordinal)
- satisfaccion: Nivel de satisfacción del trabajador en la empresa (Bajo, Medio, Alto) (V. Cualitativa Ordinal)

además, se incluye una variable ID con el orden en el que se registraron los datos de los empleados.

Esta base de datos es muy variada y posibilita al usuario poder aplicar todas las técnicas de la sección Análisis a un conjunto de datos sin necesidad de importar uno propio.

Si el usuario selecciona la opción de importar un archivo, debe incluir éste en el widget “Subir Archivo”, clicando el botón “Examinar...” y seleccionando el archivo que quiera importar.

Los tipos de archivos que admite la aplicación son:

- Archivos de texto (.txt)
- Archivos de valores delimitado por comas (.csv)
- Archivos de Excel (.xls o .xlsx)
- Archivos de datos de SPSS (.sav)

en el menú desplegable “Tipo de Archivo” se debe especificar el tipo de archivo que se haya importado.

Si los nombres de las variables se encuentran en la primera fila se debe marcar la casilla correspondiente en el panel lateral.

Además, para los archivos de texto .txt se permite seleccionar:

- Tipo de separador: coma, tabulador o punto y coma
- Tipo de decimal: coma o punto

Una vez seleccionados correctamente todos los parámetros necesarios en el panel lateral, se mostrará en el panel principal el conjunto de datos importado, ofreciendo al usuario la posibilidad de ordenar los datos por los valores de las distintas variables, aplicar algún filtro o limitar el número de entradas que se muestran en pantalla. Al final del panel principal se muestran las dimensiones del archivo importado.

Al haberse importado un conjunto de datos, se puede volver a importar otro para trabajar con él o cambiar el valor del menú desplegable “Dataset” para utilizar el conjunto de datos por defecto.

El conjunto de datos seleccionado, que siempre es el que aparece en el panel principal de la sección Datos, será el que estará disponible para usar en la sección Análisis.

La aplicación reconoce las variables tipo carácter como factores, con lo cual, los valores de las variables que sean cualitativas deben ser caracteres o números entre comillas.

Si la importación de datos provoca un error, se muestra un mensaje en el panel principal alertando al usuario de que hay algún error. Si se produce un *warning* al importar el archivo, la aplicación también muestra un mensaje avisando al usuario.

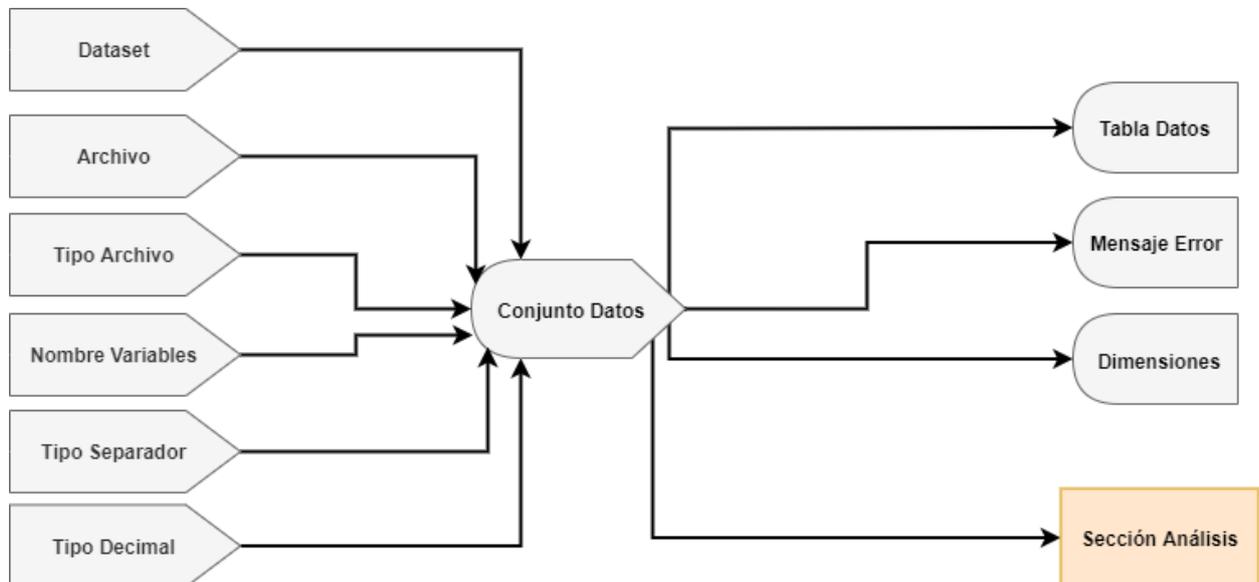


Ilustración 15: Esquema Reactivo de la sección Datos.

Fuente: Elaboración Propia (diagrams.net)

En el esquema reactivo de esta sección (Ilustración 15), se observa que, al contrario que en la sección Aprende, no hay flechas azules, ya que la reactividad no está controlada directamente por ninguna fuente reactiva. El conductor reactivo “Conjunto Datos” varía instantáneamente al modificarse alguno de los inputs. Puede que algunas modificaciones no provoquen cambios, como en el caso de las fuentes reactivas “Tipo Separador” o “Tipo Decimal”, que solo afectan al conjunto de datos importado si el input “Tipo Archivo” toma el valor “TXT”. También, si el input “Dataset” toma el valor “Empleados”, se utilizará el conjunto de datos por defecto y el resto de las entradas de la sección no afectarán al conductor reactivo.

El conductor reactivo “Conjunto Datos” se utiliza en la elaboración de los tres puntos finales reactivos y en la sección Análisis (Cuadro naranja) para obtener los valores de algunos inputs y los resultados de las pruebas aplicadas a estos datos. Si se utiliza el conjunto de datos por defecto o el archivo se importó correctamente, se mostrarán las salidas “Tabla Datos” y “Dimensiones” (Mensaje Error tendría el valor *NULL*), mientras que, si el set de datos se importó incorrectamente, sólo se muestra en el panel principal la salida “Mensaje Error”, que notifica al usuario de que se produjo un error o un *warning* al importar el archivo (Las otras dos salidas tendrían valor *NULL*).

¿No tiene datos? Puede utilizar el dataset de ejemplo 'Empleados':

Dataset: Empleados

Importar Datos:

Subir Archivo: Examinar... Sin Datos

Tipo de Archivo: CSV

¿Están en la primera fila los nombres de las variables?

Tipo de Separador:*
 Coma: ;
 Tabulador
 Punto y Coma: ;'

Tipo de Decimal:*
 Coma: ;
 Punto: ;'

* El tipo de separador y el tipo de decimal sólo funcionan con archivos .txt

Archivo importado:

Show 10 entries

ID	sexo	ecivil	edad	antigüedad	salario2006	salario2008	salario2010	evaluacion	dias_trabajados	zona_trabajo	et
1	Mujer	Casado	29	2	18689.00	19170.00	21751.00	Excelente	119	Norte	Ur
2	Mujer	Casado	33	11	21282.00	22000.00	23654.00	Excelente	101	Norte	Ba
3	Mujer	Casado	36	6	21216.00	21512.00	22369.00	Muy buena	98	Norte	Ba
4	Mujer	Casado	44	17	25014.00	26000.00	27589.00	Regular	102	Norte	Ba
5	Hombre	Casado	31	2	17928.19	18536.15	19382.17	Muy buena	115	Norte	Ur
6	Hombre	Casado	35	9	19034.05	19529.58	20461.82	Muy buena	118	Norte	Ba
7	Hombre	Casado	45	11	18469.10	18908.47	19908.79	Muy buena	122	Norte	Pri
8	Hombre	Casado	44	19	25645.19	26232.54	27118.86	Buena	118	Norte	Ba
9	Hombre	Casado	35	11	22249.47	22863.13	23616.30	Buena	106	Norte	Ba
10	Hombre	Casado	35	12	22748.31	23162.60	24085.45	Regular	113	Norte	Ba

Showing 1 to 10 of 64 entries

Las dimensiones del archivo importado son: 64 x 13

Ilustración 16: Interfaz de NPS, Sección Datos

Fuente: NPS

En la Ilustración 16 se muestra la interfaz de usuario de la sección Datos, donde está seleccionado el set de datos por defecto “Empleados”, que se muestra en el panel principal.

3.1.5 Análisis

El objetivo principal de esta sección es ofrecer al usuario un entorno en el que poder aplicar las principales técnicas no paramétricas a un conjunto de datos propio o, en su defecto, al set de datos Empleados que ofrece NPS.

Al acceder mediante la barra de navegación principal a la sección Análisis se despliega un menú en el que seleccionar a qué prueba/apartado acceder. En cada uno de estos apartados se ofrece al usuario la posibilidad de seleccionar, en el panel lateral izquierdo, la variable o variables a las que se quiere aplicar el test y variar otro tipo de valores del contraste específico como, por ejemplo, el tipo de contraste (Bilateral, Unilateral superior, Unilateral inferior), si se desea el estadístico y el p-valor exactos o si se quiere realizar una corrección por continuidad.

En el panel principal se muestra el nombre de la prueba/apartado en el que se encuentra el usuario, las hipótesis del contraste en cuestión, los valores elegidos por el usuario en el panel lateral y los resultados de la prueba, que principalmente son el valor del estadístico, el p-valor asociado y un gráfico que acompaña al test.

Las pruebas disponibles en esta sección son exactamente las mismas que las de la sección Aprende, aunque las pruebas de Kendall y Spearman en esta sección aparecen en el mismo apartado.

- Aleatoriedad: W-W

- Bondad de Ajuste: K-S ($F=F_0$)
- Bondad de Ajuste: K-S ($F=G$)
- Localización Muestra: W
- Dos muestras independientes: U M-W
- Dos muestras relacionadas: Apareados W
- K muestras independientes: K-W
- K muestras relacionadas: Friedman
- Independencia: Kendall y Spearman

En la Ilustración 17 se muestra la interfaz de usuario del apartado Aleatoriedad: W-W de la sección Análisis, donde se aplicó la prueba de las rachas de Wald-Wolfowitz a la variable “satisfacción” del conjunto de datos por defecto Empleados. También se muestra el menú desplegable con los distintos apartados de la sección.

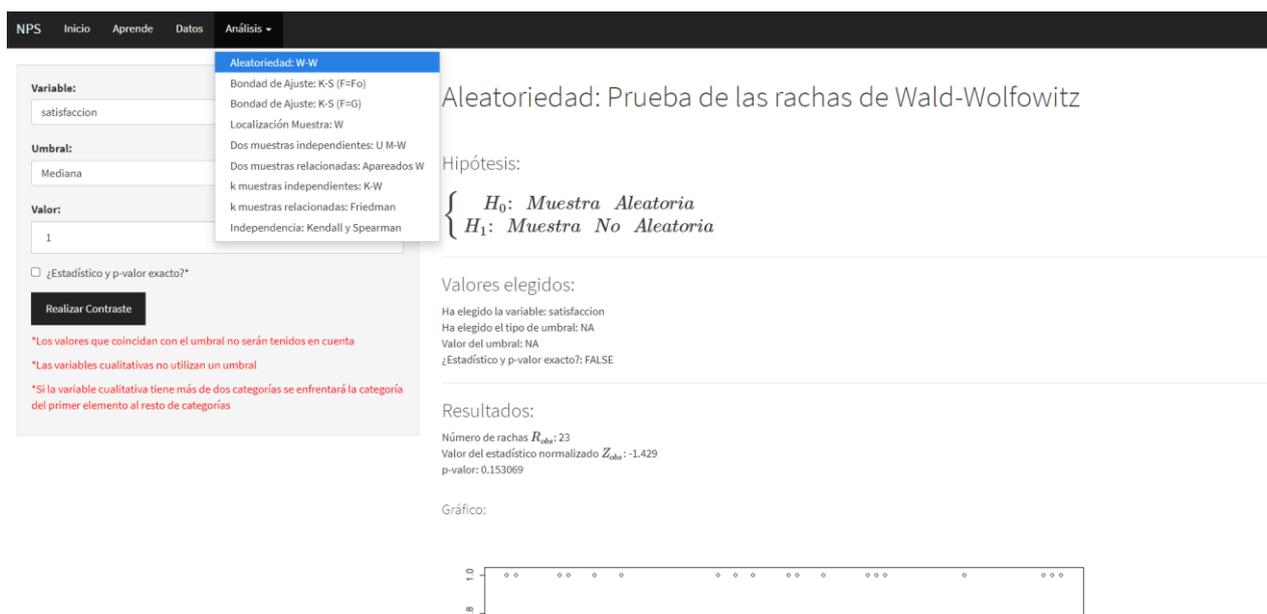


Ilustración 17: Interfaz de NPS, Sección Análisis

Fuente: NPS

3.1.5.1 Aleatoriedad: W-W

En el apartado *Aleatoriedad: W-W* se permite al usuario realizar el contraste de aleatoriedad de la prueba de las rachas de *Wald-Wolfowitz* (2.1.2.1). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Ésta puede consultarse en la Ilustración 17 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable a la que se quiere aplicar la prueba en el menú desplegable “Variable”. En este menú aparecerán las variables del conjunto de datos que se haya seleccionado en la sección Datos. En este panel puede elegirse también el tipo de umbral entre los siguientes:

- Mediana: La mediana de la muestra
- Media: La media de la muestra
- Valor: Valor elegido por el usuario

este tipo de umbral solo se aplica si la variable es cuantitativa. (Si la variable es cualitativa no se necesita ningún umbral para dicotomizar la muestra).

También puede elegirse el valor del umbral (si el tipo de umbral seleccionado es “Valor”) y si se desean los valores exactos del estadístico y del p-valor.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable, Tipo de umbral, Valor del umbral, Valores exactos) y los resultados del test:

- Número de rachas R_{obs}
- Estadístico estandarizado Z_{obs}
- p-valor
- Gráfico Dispersión

en el gráfico de dispersión se muestra una línea roja con el valor del umbral elegido si la variable es cuantitativa. Si la variable cualitativa tiene más de dos categorías se asigna al valor de la primera categoría el valor uno y al resto de categorías el valor 0.

El usuario puede aplicar a los conjuntos de datos que necesite la prueba de las rachas de *Wald-Wolfowitz*, cambiando las variables a las que aplica el test y los tipos y valores del umbral en el caso de las variables cuantitativas.

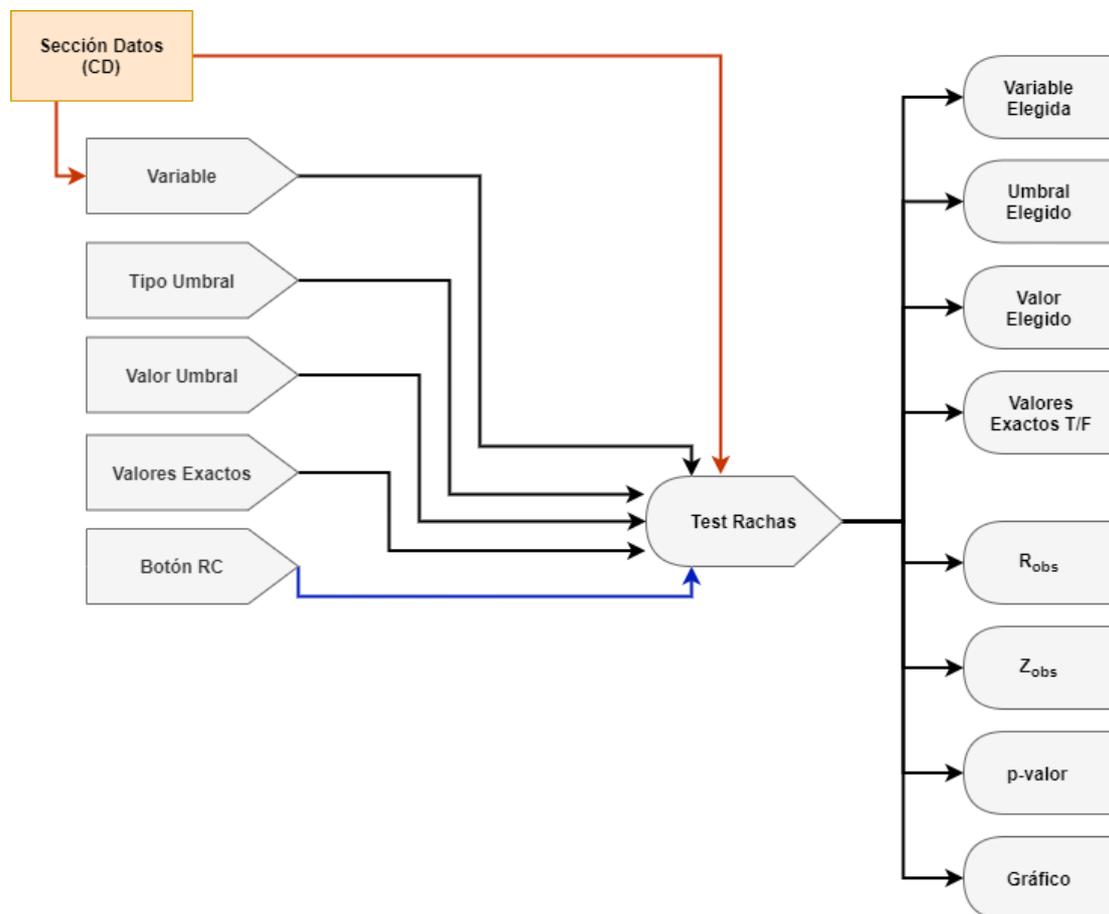


Ilustración 18: Esquema Reactivo del apartado Aleatoriedad: W-W de la sección Análisis.

Fuente: Elaboración propia (diagrams.net)

En el esquema reactivo de este apartado (Ilustración 18), las flechas naranjas indican que la información proviene del conductor reactivo “Conjunto Datos” de la sección Datos, representado en este diagrama con un recuadro naranja.

Se puede apreciar como la fuente reactiva “Variable” se actualiza según el valor del conductor reactivo “Conjunto Datos”. Las opciones de selección de esta entrada dependerán de los nombres de las variables del conjunto de datos elegido en la sección Datos.

De manera similar a los apartados de la sección Aprende, el “Botón RC” controla la reactividad del conductor reactivo “Test Rachas” (y por lo tanto controla también el cambio de todas las salidas), actualizándose los valores de este únicamente si se activa el botón, independientemente de que se alteren los valores de otras entradas.

Los principales resultados de la prueba de este apartado se obtienen con la función *runs.test()* de la librería *randtests*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.2 Bondad de Ajuste: K-S ($F=F_0$)

En el apartado *Bondad de Ajuste: K-S ($F=F_0$)* se permite al usuario realizar la prueba de bondad de ajuste de *Kolmogorov-Smirnov* para el caso $F=F_0$ (2.1.3.1). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Ésta puede consultarse en el Anexo 7.2.14 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable a la que se quiere aplicar la prueba en el menú desplegable “Variable”. En este menú aparecerán únicamente las variables cuantitativas del conjunto de datos que se haya seleccionado en la sección “Datos”. En este panel puede elegirse también la distribución F_0 en el menú desplegable “Distribución Conocida” entre:

- Normal: $F_0 = N(\mu, \sigma)$
- Exponencial: $F_0 = Exp(\lambda)$
- Chi-Cuadrado: $F_0 = \chi^2(k)$
- t-Student: $F_0 = t(k)$

donde μ es la media y σ la desviación típica de la distribución normal, λ el parámetro de la distribución exponencial y k los grados de libertad de las distribuciones χ^2 y *t-Student*.

Los valores de estos parámetros pueden elegirse también en el panel lateral en las entradas “Parámetro 1” (μ, λ, k) y “Parámetro 2” (σ).

Además, puede pedirse el valor de la media, mediana y desviación típica en la casilla “Calcular Media, Mediana y Desviación Típica” y los valores exactos del estadístico y p-valor en la casilla “Estadístico y p-valor exacto”.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable, Distribución Conocida, Valor del parámetro 1, Valor del parámetro 2, Media, Mediana y Desviación, Valores exactos) y los resultados del test:

- Valor del estadístico D_{obs}
- p-valor
- Tabla con Media, Mediana y Desviación (Si se pide)
- Gráfico Cuantil-Cuantil y un Histograma

en el histograma se muestra una línea roja que representa la función de densidad de F_0 . En el gráfico cuantil-cuantil se comparan los cuantiles de la distribución empírica de los datos de la variable seleccionada con los cuantiles de la distribución teórica F_0 , la línea azul representa la recta $y = x$

El usuario puede aplicar a los conjuntos de datos que necesite la prueba de bondad de ajuste de *Kolmogorov-Smirnov*, cambiando la variable a la que aplica el test, la distribución teórica F_0 y los valores de sus parámetros.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección “Análisis”, este puede consultarse en el Anexo 7.3.13.

Los principales resultados de la prueba de este apartado se obtienen con la función *ks.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.3 Bondad de Ajuste: K-S ($F=G$)

En el apartado *Bondad de Ajuste: K-S ($F=G$)* se permite al usuario realizar la prueba de bondad de ajuste de *Kolmogorov-Smirnov* para el caso $F=G$ (2.1.3.1). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.15 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable de agrupación que forme los dos grupos que se quieren comparar y la variable que se quiera comparar en ambos grupos en los menús “Variable Agrupación” y “Variable X”. En el primer menú aparecerán únicamente las variables cualitativas con dos categorías del conjunto de datos que se haya seleccionado en la sección Datos. En el segundo menú solo aparecerán las variables cuantitativas. En este panel lateral puede también especificarse si se prefiere el estadístico y p-valor exacto.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable de agrupación, Variable X, Valores exactos) y los resultados del test:

- Valor del estadístico D_{obs}
- p-valor
- Gráfico Cuantil-Cuantil

la línea roja del gráfico cuantil-cuantil es la recta $y = x$.

El usuario puede aplicar a los conjuntos de datos que necesite la prueba de bondad de ajuste de *Kolmogorov-Smirnov*, cambiando la variable de agrupación y la variable a la que aplica el test.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.14.

Los principales resultados de la prueba de este apartado se obtienen con la función *ks.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.4 Localización Muestra: W

En el apartado *Localización Muestra: W* se permite al usuario realizar la prueba de los rangos con signo de *Wilcoxon* (2.1.4). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.16 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable a la que se quiere aplicar la prueba en el menú desplegable “Variable”. En este menú aparecerán únicamente las variables cuantitativas del conjunto de datos que se haya seleccionado en la sección Datos. En este panel puede elegirse también el tipo de contraste:

- Bilateral $H_0: \theta = \theta_0$
- Unilateral Inferior $H_0: \theta \geq \theta_0$
- Unilateral Superior $H_0: \theta \leq \theta_0$

donde θ_0 es el valor que se quiere contrastar, como hipótesis nula.

Este valor se puede elegir también en el panel lateral en la entrada “ θ_0 ”.

Además, también puede pedirse una tabla con el valor de la media, mediana y desviación típica de los datos marcando la casilla correspondiente.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable, Tipo Contraste, Valor de θ_0 , Media, Mediana y Desviación) y los resultados del test:

- Valor del estadístico T_{obs}^+
- p-valor
- Tabla con Media, Mediana y Desviación (Si se pide)
- Box-Plot

en el box-plot aparece una línea roja que representa el valor de θ_0 .

El usuario puede aplicar a los conjuntos de datos que necesite la prueba de localización muestral para el valor de la mediana de los rangos con signo de *Wilcoxon*, cambiando la variable a la que aplica el test, el tipo de contraste y el valor de θ_0 .

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.15.

Los principales resultados de la prueba de este apartado se obtienen con la función *wilcox.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.5 Dos muestras independientes: U M-W

En el apartado *Dos muestras independientes: U M-W* se permite al usuario realizar la prueba *U* de *Mann-Whitney* para dos muestras independientes (2.1.5.2). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.17 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable de agrupación que forme las dos muestras que se quieran comparar y la variable que se quiera comparar en ambos grupos en los menús “Variable Agrupación” y “Variable X”. En el primer menú aparecerán únicamente las variables cualitativas con dos categorías del conjunto de datos que se haya seleccionado en la sección Datos. En el segundo menú solo aparecerán las variables cuantitativas (También las cualitativas ordinales si están expresadas con números en la base de datos.). En este panel lateral puede también especificarse si se prefiere el estadístico y p-valor exacto o realizar una corrección por continuidad marcando las casillas correspondientes, así como el tipo de contraste entre:

- Bilateral $H_0: X_1 = X_2$
- Unilateral Inferior $H_0: X_1 \leq X_2$
- Unilateral Superior $H_0: X_1 \geq X_2$

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable de agrupación, Variable X, Tipo Contraste, Valores exactos, Corrección Continuidad) y los resultados del test:

- Valor del estadístico U_{obs}
- p-valor
- Box-Plot, uno para cada muestra

El usuario puede aplicar a los conjuntos de datos que necesite la prueba para dos muestras independientes U de *Mann-Whitney*, cambiando la variable de agrupación, la variable a la que aplica el test y el tipo de contraste.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.16.

Los principales resultados de la prueba de este apartado se obtienen con la función *wilcox.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.6 Dos muestras relacionadas: Apareados W

En el apartado *Dos muestras relacionadas: Apareados W* se permite al usuario realizar la prueba de los rangos con signo de *Wilcoxon* para dos muestras relacionadas (2.1.5.1). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.18 o en la propia aplicación web.

En el panel lateral izquierdo se pueden seleccionar las variables a las que se quiere aplicar la prueba en los menús desplegados “Variable X” y “Variable Y”. En estos menús aparecerán únicamente las variables cuantitativas del conjunto de datos que se haya seleccionado en la sección Datos. La aplicación no permite seleccionar la misma variable en ambos menús. En este panel lateral puede también especificarse si se prefiere el estadístico y p-valor exacto o realizar una corrección por continuidad marcando las casillas correspondientes, así como el tipo de contraste entre:

- Bilateral $H_0: X = Y$
- Unilateral Inferior $H_0: X \leq Y$
- Unilateral Superior $H_0: X \geq Y$

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable X, Variable Y, Tipo Contraste, Valores exactos, Corrección Continuidad) y los resultados del test:

- Valor del estadístico T_{obs}^+
- p-valor
- Box-Plot, uno para cada muestra

El usuario puede aplicar a los conjuntos de datos que necesite la prueba de los rangos con signo de *Wilcoxon* para dos muestras relacionadas, cambiando las variables que se quieren comparar y el tipo de contraste.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.17.

Los principales resultados de la prueba de este apartado se obtienen con la función *wilcox.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.7 K muestras independientes: K-W

En el apartado *K muestras independientes: K-W* se permite al usuario realizar la prueba de *Kruskal-Wallis* para k muestras independientes (2.1.6.1). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.19 o en la propia aplicación web.

En el panel lateral izquierdo se puede seleccionar la variable de agrupación que forme las k muestras y la variable que se quiera comparar en ambos grupos en los menús “Variable Agrupación” y “Variable X”. En el primer menú aparecerán únicamente las variables cualitativas con dos o más categorías del

conjunto de datos que se haya seleccionado en la sección Datos. En el segundo menú solo aparecerán las variables cuantitativas (También las cualitativas ordinales si están expresadas con números en la base de datos.). En este panel lateral puede también especificarse si quiere realizar un análisis post-hoc para comparar las muestras dos a dos marcando la casilla correspondiente.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable de agrupación, Variable X, Post-hoc) y los resultados del test:

- Valor del estadístico H_{obs}
- p-valor
- Box-Plot, uno para cada muestra
- Tabla con los p-valores del análisis post-hoc (Si se pide)

el cálculo de los p-valores ajustados se realiza con la función *pairwise.wilcox.test()* con el método “*fdr*” (Benjamini & Hochberg, 1995)

El usuario puede aplicar a los conjuntos de datos que necesite la prueba para k muestras independientes de *Kruskal-Wallis*, cambiando la variable de agrupación, la variable a la que aplica el test y solicitando o no el análisis post-hoc.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.18.

Los principales resultados de la prueba de este apartado se obtienen con la función *kruskal.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.8 K muestras relacionadas: Friedman

En el apartado *K muestras relacionadas: Friedman* se permite al usuario realizar la prueba de *Friedman* para k muestras relacionadas (2.1.6.2). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.20 o en la propia aplicación web.

En el panel lateral izquierdo se debe seleccionar el número de variables relacionadas que se quieren comparar en el botón deslizante “Número de Variables” seleccionando un valor entre tres y seis. Después, aparecen tantos menús como variables se seleccionaron en el botón deslizante y se podrán seleccionar las variables a las que se quiere aplicar la prueba en los menús desplegables “Variable 1”, “Variable 2”, “Variable 3” ... etc. En estos menús aparecerán únicamente las variables cuantitativas del conjunto de datos que se haya seleccionado en la sección Datos. La aplicación no permite seleccionar la misma variable en más de un menú. En este panel lateral puede pedirse también un análisis post-hoc para comparar las muestras dos a dos marcando la casilla correspondiente.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variables elegidas, Post-hoc) y los resultados del test:

- Valor del estadístico S_{obs}
- p-valor
- Box-Plot, uno para cada muestra
- Tabla con los p-valores del análisis post-hoc (Si se pide)

el cálculo de los p-valores ajustados se realiza con la función *pairwise.wilcox.test()* con el método “*fdr*” (Benjamini & Hochberg, 1995)

El usuario puede aplicar a los conjuntos de datos que necesite la prueba para k muestras relacionadas de *Friedman*, cambiando las variables a las que aplica el test y solicitando o no el análisis post-hoc.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.19.

Los principales resultados de la prueba de este apartado se obtienen con la función *friedman.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.1.5.9 Independencia: Kendall y Spearman

En el apartado *Independencia: Kendall y Spearman* se permite al usuario realizar las pruebas de independencia de *Kendall* y *Spearman* (2.1.7). La interfaz de este apartado es similar a la del resto de apartados de la sección Análisis. Esta puede consultarse en el Anexo 7.2.21 o en la propia aplicación web.

En el panel lateral izquierdo se pueden seleccionar las variables a las que se quiere aplicar la prueba en los menús desplegables “Variable 1” y “Variable 2”. En estos menús aparecerán únicamente las variables cuantitativas del conjunto de datos que se haya seleccionado en la sección Datos (Y las cualitativas ordinales que se hayan introducido como números). La aplicación no permite seleccionar la misma variable en ambos menús. En este panel lateral puede también especificarse si se prefiere el estadístico y p-valor exacto o realizar una corrección por continuidad marcando las casillas correspondientes.

Dependiendo de si se quiere realizar la prueba de *Kendall* o la de *Spearman* se debe elegir en este panel lateral en el menú desplegable “Método” cuál de las dos técnicas se quiere aplicar.

Una vez elegidos los valores deseados para realizar el contraste se debe accionar el botón “Realizar Contraste”.

En el panel principal se muestra: el título del apartado en el que se encuentra el usuario, las hipótesis del contraste, los valores elegidos en el panel lateral (Variable 1, Variable 2, Método, Valores exactos, Corrección Continuidad) y los resultados del test:

- Valor de asociación τ_{obs} o r_{obs} (Kendall o Spearman respectivamente)
- Valor del estadístico z_{obs} o t_{obs} (Kendall o Spearman respectivamente)
- p-valor
- Diagrama de dispersión de las dos variables

El usuario puede aplicar a los conjuntos de datos que necesite las pruebas de independencia de *Kendall* o *Spearman*, cambiando las variables a las que aplica el test y el método utilizado.

El esquema reactivo de este apartado es similar al de los demás apartados de la sección Análisis, este puede consultarse en el Anexo 7.3.20.

Los principales resultados de la prueba de este apartado se obtienen con la función *cor.test()* de la librería *stats*. El código utilizado puede consultarse en el Anexo 7.4.

3.2 Shinyapps.io

La aplicación NPS fue subida al servidor online *Shinyapps.io*, para que cualquier usuario pudiese acceder a ella sin necesidad de tener instalado R.

Ésta puede consultarse y utilizarse accediendo mediante el siguiente código QR o enlace:



<https://a-est-gal.shinyapps.io/npsapp/>

4. CONCLUSIONES

Sin duda, el desarrollo de aplicaciones web en R con la librería Shiny puede ser de gran utilidad en muchos ámbitos, especialmente para la investigación y el ámbito empresarial, acortando en gran medida el tiempo de análisis de datos o de aplicación de distintos algoritmos. Al contrario que las habituales aplicaciones como SPSS o MATLAB, las aplicaciones desarrolladas con Shiny, al basarse en el software gratuito R, no requieren de ningún tipo de pago, además de ser totalmente personalizables por el desarrollador, adaptándose a las necesidades del usuario.

El desarrollo de aplicaciones con esta librería puede ser útil también para presentar informes estadísticos de manera dinámica o como herramienta de divulgación sobre distintos temas de interés.

La aplicación desarrollada NPS pretende ser una herramienta útil para el aprendizaje de las principales técnicas no paramétricas, así como para su aplicación en distintos conjuntos de datos, con el objetivo de ayudar a investigadores, estudiantes y a cualquier usuario que lo necesite a realizar contrastes no paramétricos de una forma sencilla, rápida y gratuita, además de no requerir conocimientos sobre el lenguaje de programación R.

Las técnicas no paramétricas fueron elegidas en esta aplicación por su relevancia en la inferencia estadística, siendo una buena alternativa a las pruebas paramétricas en algunas situaciones, como tamaños muestrales pequeños o ausencia de normalidad, debido a que no se requiere que los datos sigan una distribución concreta ni que exista homocedasticidad ni otros requisitos habituales en los test paramétricos. Además, algunas pruebas no paramétricas admiten variables ordinales y estudian la localización muestral de la mediana o la moda. En definitiva, las pruebas no paramétricas presentan, en general, una robustez mayor que las pruebas paramétricas.

La aplicación NPS ofrece la sección “Aprende” con la finalidad de divulgar conceptos básicos sobre las técnicas utilizadas en la aplicación, ya que, por lo general, las pruebas no paramétricas son menos habituales y conocidas que las paramétricas.

Como perspectivas de mejora para NPS, podría realizarse un estudio de opinión del usuario objetivo, con el fin de añadir y/o modificar funcionalidades que pudiesen aumentar la utilidad, usabilidad y valor de la aplicación.

5. BIBLIOGRAFÍA

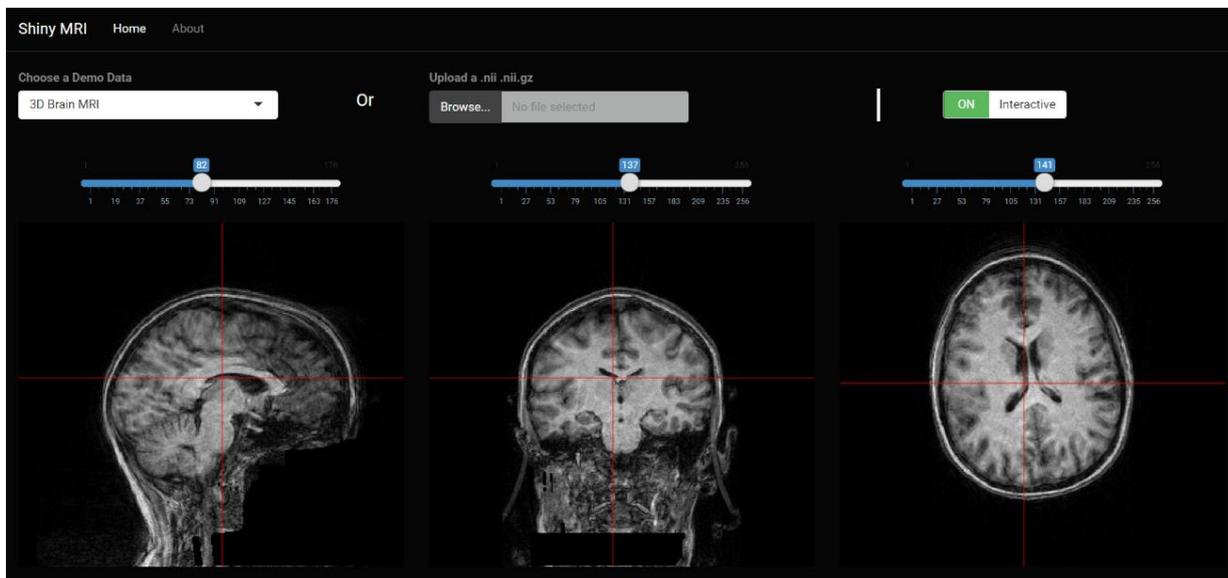
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1), 289-300.
- Boschetti, A., & Massaron, L. (2015). *Python data science essentials*. Packt Publishing Ltd.
- Caeiro, F., & Mateus, A. (2014). *Reference Guide: Package «randtests»*. Disponible en: <https://cran.r-project.org/web/packages/randtests/randtests.pdf>
- Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2017). Shiny: Web application framework for R. *R package version*, 1(5), 2017.
- Chang, W., Park, T., Dziedzic, L., Willis, N., & McInerney, M. (2015). shinythemes: Themes for Shiny. *R package version*, 1(1), 144.
- Dalgaard, P. (2000). *R-1.0.0 is released*. Recuperado de: <https://stat.ethz.ch/pipermail/r-announce/2000/000127.html>
- Gibbons, J. D., & Chakraborti, S. (1992). *Nonparametric statistical inference*. CRC press.
- Hollander, M., Wolfe, D. A., & Chicken, E. (2013). *Nonparametric statistical methods* (Vol. 751). John Wiley & Sons.
- Ihaka, R. (s. f.). *R: Past and Future History*. Recuperado de: <http://www.stat.auckland.ac.nz/~ihaka/downloads/Interface98.pdf>
- Lehmann, E. L. (1951). Consistency and unbiasedness of certain nonparametric tests. *The annals of mathematical statistics*, 165-179.
- Lie, H. W., & Bos, B. (2005). *Cascading style sheets: Designing for the web, Portable Documents*. Addison-Wesley Professional.
- Marden, J. I. (2004). Positions and QQ plots. *Statistical Science*, 19(4), 606-614.
- Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American statistical Association*, 46(253), 68-78.
- Mulero, J. (s. f.). *Aplicaciones interactivas diseñadas con Shiny*. Recuperado de: <https://rua.ua.es/dspace/bitstream/10045/54325/1/shiny.pdf>

- Pérez González, C., & Colebrook Santamaría, M. (2014). *Curso introductorio de Introducción a la interfaz RStudio*. Recuperado de:
[http://mcolebrook.github.io/CursoRStudio/RStudio.html#\(1\)](http://mcolebrook.github.io/CursoRStudio/RStudio.html#(1))
- RStudio. (s. f.-a). *Shiny Gallery*. Disponible en: <https://shiny.rstudio.com/gallery/>
- RStudio, Inc. (s. f.-b). *Documentación RStudio, Shiny*.
- RStudio, Inc. (2012). *Tutorial: Building «Shiny» Applications with R*. Disponible en:
<https://rstudio.github.io/shiny/tutorial/>
- RStudio, Inc. (2015). *How to Start, Shiny tutorial*. Disponible en:
<https://shiny.rstudio.com/tutorial/>
- RStudio, Inc. (2016). *Shiny: Cheat Sheet*. Recuperado de:
<https://shiny.rstudio.com/articles/cheatsheet.html>
- Satyahadewi, N., & Perdana, H. (2021). Web Application Development for Inferential Statistics using R Shiny. *1st International Conference on Mathematics and Mathematics Education (ICMMEd 2020)*, 50, 425-429.
- Siegel, S. (1957). Nonparametric statistics. *The American Statistician*, 11(3), 13-19.
- *Stackoverflow*. (s. f.). Disponible en: <https://es.stackoverflow.com/>
- Swed, F. S., & Eisenhart, C. (1943). Tables for testing randomness of grouping in a sequence of alternatives. *The Annals of Mathematical Statistics*, 14(1), 66-87.
- Wickham, H. (2021). *Mastering Shiny: Build Interactive Apps, Reports & Dashboards Powered by R*. Disponible en: <https://mastering-shiny.org/index.html>
- Zhu, H., Mahaveer Chand, N., & Trivison, T. (s. f.). *Shiny MRI App*. Disponible en:
https://haozhu233.shinyapps.io/shinyMRI-contest/?_ga=2.177540589.958378940.1625094513-947499331.1612178125

6. SUMMARY

The development of web applications with data science-oriented languages such as Python or R can be very useful for the researcher who develops them or for the user to whom they are addressed. The usefulness of these applications is not limited exclusively to the field of statistical inference, as they can be useful in many fields such as demography, medicine, biology or economics. In Shiny's gallery you can consult different published applications, with a wide variety of complexities and themes.

In the R language, it is possible to develop these web applications with the Shiny library. With this library you can develop very simple applications and more complex applications combining the R language with other languages such as HTML or CSS, achieving very attractive applications.



The illustration above (Zhu et al., n. d.) shows the interface of one of the applications in the Shiny gallery. This application allows the user to visualise images of a human head obtained by magnetic resonance imaging.

In order to give visibility to the development of this type of applications and their usefulness, it was decided to develop a web application that could be used by the scientific and university community. The theme chosen for the development of the application was non-parametric tests, which can be very useful in certain contexts of statistical inference applications.

Therefore, the general objective of this work is to develop a web application that offers the user a simple and intuitive environment in which to learn and use the main non-parametric tests.

The main non-parametric tests used in the application were: Wald-Wolfowitz runs test, Kolmogorov-Smirnov goodness-of-fit tests for the cases $F=F_0$ and $F=G$, Wilcoxon signed-rank test for the location of the median and for comparing two related samples, the Mann-Whitney U test for two independent samples, the Kruskal-Wallis test for $k>2$ independent samples, the Friedman test for $k>2$ related samples and Kendall and Spearman contrasts of independence.

Non-parametric tests are more advisable than parametric tests in certain contexts such as to test a parameter such as median or mode, to test the randomness of a sample, to perform a goodness-of-fit test or if the researcher is unsure about the type of probability distribution from which the sample data come.

Wald-Wolfowitz runs test

The Wald-Wolfowitz runs test consists of determining the randomness of a sample by calculating the total number of runs in the sample and checking whether or not this falls within an interval obtained from the Table 7 (R) (Annex 7.1) if $n_1, n_2 \leq 20$.

Otherwise, the asymptotic distribution of R is used:

$$z_{obs} = \frac{r_{obs} - E[R]}{\sqrt{Var[R]}} \approx N(0,1)$$

And the value of z_{obs} is compared with the quantile of the standard normal distribution $Z_{1-\frac{\alpha}{2}}$

Kolmogorov-Smirnov test

Kolmogorov-Smirnov goodness-of-fit tests measure the similarity between two probability distributions. For the case $F=F_0$ the empirical distribution of F is compared with the corresponding quantiles of the known theoretical distribution F_0 using the Kolmogorov-Smirnov statistic which compares $F_n^*(x)$ with $F_0(x)$:

$$D = D(F_n^*(x), F_0) = \max\{|F_n^*(x) - F_0(x)|, \forall x \in \mathbb{R}\}$$

For the case $F=G$ the empirical distributions of F and G are compared with the Kolmogorov-Smirnov statistic:

$$D = D(F_n^*(z_{(k)}), G_m^*(z_{(k)})) = \max\{|F_n^*(z_{(k)}) - G_m^*(z_{(k)})|\}$$

For both cases the value of D_{obs} is compared with the quantile $D_{n,\alpha}$ obtained in Table 6 (KS) (Annex 7.1)

Wilcoxon signed-rank test (Median)

The Wilcoxon signed-rank test for the location of the median consists of constructing a new variable $D = X - \theta_0$, where X is the sample and θ_0 is the value of the median in the null hypothesis. Calculate the ranks of the absolute values of the new variable D. And obtain the value of the Wilcoxon signed ranks statistic:

$$T^+ = \sum_{D_i > 0} R_i$$

If the sample size is small (in practice $n \leq 12$), the value of the sampling significance corresponding to each hypothesis is looked up in Table 8 (T^+) (Annex 7.1):

- $H_1: \theta_D < 0 \rightarrow H_1: \theta < \theta_0 \rightarrow P_I = P(T^+ \leq T_{obs}^+) \rightarrow \begin{cases} \text{If } P_I > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P_I \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$
- $H_1: \theta_D > 0 \rightarrow H_1: \theta > \theta_0 \rightarrow P_S = P(T^+ \geq T_{obs}^+) \rightarrow \begin{cases} \text{If } P_S > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P_S \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$
- $H_1: \theta_D \neq 0 \rightarrow H_1: \theta \neq \theta_0 \rightarrow P = 2 \cdot \min\{P_I, P_S\} \rightarrow \begin{cases} \text{If } P > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$

If the sample size is large (in practice $n > 12$), the asymptotic distribution of T^+ is used:

$$z_{obs} = \frac{T_{obs}^+ - E[T^+]}{\sqrt{Var[T^+]}} \approx N(0,1)$$

And the value of z_{obs} is compared with the quantile of the corresponding standard normal distribution according to the type of contrast.

Wilcoxon signed-rank test (two samples)

The Wilcoxon test for paired data follows almost the same procedure as the test for the location of the median, in this case the new variable is obtained by subtracting the values of the two samples to be compared.: $D = Y - X$ The procedure from this point on is identical to the median location test.

Mann-Whitney U test

The Mann-Whitney U test for two independent samples consists of calculating the joint sample ranks of the two samples, obtaining the values of R_1 y R_2 :

$$R_1 = \sum_{i=1}^{n_1} \text{Rank}(x_i) \quad R_2 = \sum_{j=1}^{n_2} \text{Rank}(y_j)$$

the values of U_1 and U_2 :

$$U_1 = R_1 - \frac{n_1(n_1 + 1)}{2} \quad U_2 = R_2 - \frac{n_2(n_2 + 1)}{2}$$

And the value of Mann-Whitney statistic:

$$U_{obs} = \min\{U_1, U_2\}$$

If the sample size is small (in practice $n_1, n_2 \leq 8$), the value of $P(U \leq U_{obs})$ is searched in Table 9 (U) (Annex 7.1) and proceed according to the type of contrast:

- $H_1: X > Y \rightarrow P_I = P(U \leq U_{obs}) \rightarrow \begin{cases} \text{If } P_I > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P_I \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$
- $H_1: X < Y \rightarrow P_S = P(U \geq U_{obs}) \rightarrow \begin{cases} \text{If } P_S > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P_S \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$
- $H_1: X \neq Y \rightarrow P = 2 \cdot \min\{P_I, P_S\} \rightarrow \begin{cases} \text{If } P > \alpha \rightarrow H_0 \text{ is accepted} \\ \text{If } P \leq \alpha \rightarrow H_0 \text{ is rejected} \end{cases}$

If the sample size is large (in practice $n_1, n_2 > 8$), the asymptotic distribution of U is used:

$$z_{obs} = \frac{U_{obs} - E[U]}{\sqrt{\text{Var}[U]}} \approx N(0,1)$$

And the value of z_{obs} is compared with the corresponding standard normal distribution quantile, according to the type of contrast.

Kruskal-Wallis test

The Kruskal-Wallis test consists of calculating the ranks of the pooled sample, for each sample i , the values of $R_i = \sum_{j=1}^{n_i} R_{ij}$ are calculated and the Kruskal-Wallis statistic is obtained as:

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

where N is the aggregate sample size.

The value of H_{crit} is searched in Table 10 (H) (Annex 7.1) and compared with the value of H_{obs} :

- If $H_{obs} < H_{crit}$ H_0 is accepted
- If $H_{obs} \geq H_{crit}$ H_0 is rejected

If the value of H_{crit} is not in the table or the samples have ties the asymptotic distribution of the statistic is used:

$$H \approx \chi_{k-1}^2$$

And the quantiles of the distribution χ_{k-1}^2 are compared with the value of H_{obs} . If the contrast is significant, a post-hoc analysis is performed comparing the samples two by two.

Friedman test

The Friedman test consists of calculating the ranks of the variables in each individual R_{ij} , calculating the values of $R_{.j} = \sum_{i=1}^n R_{ij}$ and the Friedman statistic value:

$$S_{obs} = \left(\frac{12}{nk(k+1)} \sum_{j=1}^k R_{.j}^2 \right) - 3n(k+1)$$

The value of S_{crit} is searched on Table 11 (S) (Annex 7.1) and compared with the value of S_{obs} :

- If $S_{obs} < S_{crit}$ H_0 is accepted
- If $S_{obs} \geq S_{crit}$ H_0 is rejected

If $k > 5$ or $n > 15$ or if there are many ties, the asymptotic distribution of S is used:

$$S \approx \chi_{k-1}^2$$

And the quantiles of χ_{k-1}^2 are compared with the value of S_{obs} . If the contrast is significant, a post-hoc analysis is performed comparing the samples two by two.

Kendall test

Kendall's test for concordance of two variables consists of, for each $i < j$, calculating:

$$k((x_i, y_i), (x_j, y_j)) = \begin{cases} 1 & \text{si } (x_i - x_j)(y_i - y_j) > 0 \\ 0 & \text{si } (x_i - x_j)(y_i - y_j) = 0 \\ -1 & \text{si } (x_i - x_j)(y_i - y_j) < 0 \end{cases}$$

The Kendall statistic is then obtained as:

$$k_{obs} = \sum_{i=1}^{n-1} \sum_{j=i+1}^n k((x_i, y_i), (x_j, y_j)) = \#\{\text{concordant}\} - \#\{\text{discordant}\}$$

And the τ Kendall's correlation coefficient is calculated:

$$\tau_{obs} = \frac{2k}{n(n-1)} \quad -1 \leq \tau \leq +1$$

The value of τ_{crit} is searched in the Table 12 (τ) (Annex 7.1) and is compared with the value of τ_{obs} :

- If $\tau_{obs} < \tau_{crit}$ H_0 is accepted
- If $\tau_{obs} \geq \tau_{crit}$ H_0 is rejected

If the sample size is large (in practice $n > 30$), the asymptotic distribution of τ is used:

$$z_{obs} = \frac{\tau_{obs} - E[\tau]}{\sqrt{Var[\tau]}} \approx N(0,1)$$

And the value of $|z_{obs}|$ is compared with the quantile $z_{1-\frac{\alpha}{2}}$

Spearman test

Finally, Spearman's test consists of calculating the ranks of the two samples separately $R_i = rango(x_i)$; $S_i = rango(y_i)$ and obtain Spearman's statistic:

$$r_{s_{obs}} = \frac{12}{n^3 - n} \cdot \sum_{i=1}^n \left(R_i - \frac{n+1}{2} \right) \left(S_i - \frac{n+1}{2} \right) \quad -1 \leq r_s \leq +1$$

The value of r_{crit} is searched in Table 13 (RS) (Annex 7.1) and compared with the value of $r_{s_{obs}}$:

- If $r_{s_{obs}} < r_{crit}$ H_0 is accepted
- If $r_{s_{obs}} \geq r_{crit}$ H_0 is rejected

If the sample size is large (in practice $n > 30$), the *t-Student* distribution is used (r_s does not follow a known asymptotic distribution)

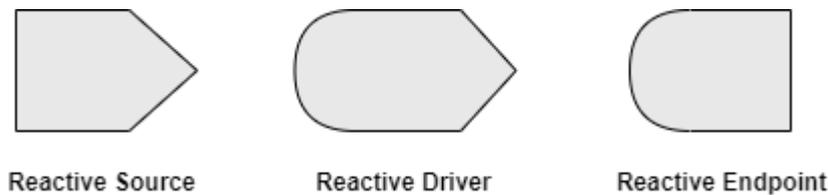
$$t_{obs} = r_{s_{obs}} \sqrt{\frac{n-2}{1-r_{s_{obs}}^2}} \approx t_{n-2}$$

The value of $|t_{obs}|$ is compared with the quantile $t_{n-2; 1-\frac{\alpha}{2}}$

Shiny

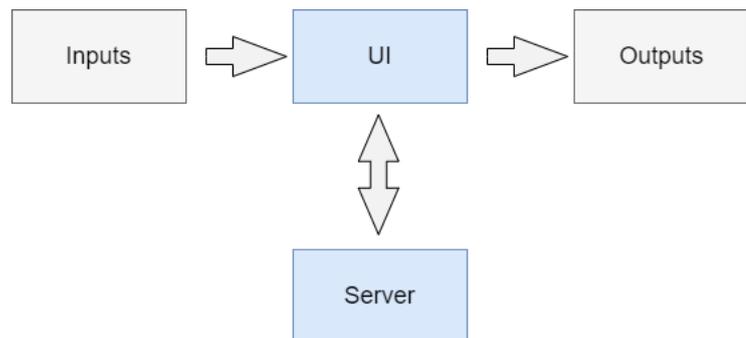
Shiny is an R library that facilitates the development of interactive web applications. The main interest of these applications is their interactivity, which allows the user to use R without having to create or modify code. In addition, they can be used from the computer, tablet or mobile phone.

In Shiny applications there are three types of elements related to reactive programming: reactive sources, reactive drivers and reactive endpoints.



The main element of an interactive Shiny application is the user interface, which is the means by which the user of the application can interact with the server to obtain results. One of the main objectives of this interface is to make the user's communication with the server simple and intuitive.

The second essential block in an interactive Shiny application is the server, where the values of the inputs chosen by the user are received through the widgets in the user interface and the outputs are elaborated and sent to the interface to be shown to the client.



If you want to share a Shiny app, you can upload it to the cloud on your own server or on a server such as Shinyapps.io

NPS

The application developed, which received the name NPS (Non Parametric Statistics), is focused on the dissemination of non-parametric tests and on serving as a tool for students, companies, researchers and anyone else to whom it may be useful, to apply these techniques to their own dataset quickly and easily, without the need to apply programming knowledge.

NPS consists of four main sections: Home, Learn, Data and Analysis.

The "Inicio" section describes the general functioning of the different sections, as well as other information of interest such as the contact details of the application's developer. The user interface is shown in the image below.

NPS

¡Bienvenido a NPS!

Esta aplicación web ofrece herramientas para aprender sobre algunas técnicas no paramétricas así como la posibilidad de importar un set de datos y aplicar test no paramétricos a estos datos importados.

A continuación se describe el funcionamiento general de cada uno de los apartados principales de la aplicación.

Aprende

En la sección Aprende se explica el proceso de cada una de las pruebas no paramétricas con un ejemplo práctico, incluyendo las fórmulas utilizadas y enlaces a las tablas de cuantiles necesarias en cada caso. Es posible practicar estos ejercicios las veces que se necesite, ocultando las soluciones y generando nuevos datos. Después, se muestran las soluciones para comprobar que se han realizado las cuentas correctamente y que se ha comprendido el proceso del test no paramétrico en cuestión. Además, este apartado ofrece enlaces en los que podemos encontrar más información sobre las pruebas.

Los tests no paramétricos que se incluyen en este apartado son:

- **Aleatoriedad: Rachas W-W** → Contraste de aleatoriedad: Prueba de las rachas de Wald-Wolfowitz.
- **Bondad de Ajuste: K-S $F = F_0$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = F_0$.
- **Bondad de Ajuste: K-S $F = G$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = G$.
- **Localización de una muestra: W** → Contraste de localización de una muestra: Prueba de los rangos con signo de Wilcoxon.
- **Dos muestras independientes: U M-W** → Contraste para dos muestras independientes: Prueba U de Mann-Whitney.
- **Dos muestras relacionadas: Apareados W** → Contraste para dos muestras relacionadas: Prueba para datos apareados de Wilcoxon.
- **k muestras independientes: K-W** → Contraste para k muestras independientes: Prueba de Kruskal-Wallis.
- **k muestras relacionadas: Friedman** → Contraste para k muestras relacionadas: Prueba de Friedman.
- **Independencia: Kendall** → Contraste de independencia: Prueba de Kendall.
- **Independencia: Spearman** → Contraste de independencia: Prueba de Spearman.

The "Aprende" section offers the user an environment in which to acquire knowledge about the main non-parametric techniques, showing a brief description of the usefulness of the test, the hypotheses of the contrast, hyperlinks with more information and a brief explanation of the functioning of each subsection. In each of the subsections of this section it is possible to generate one or several random samples by selecting some options in the side panel, such as the type of variable or the sample size. In the main panel you can see the random sample obtained and the development of the corresponding non-parametric test step by step, with the intermediate results. If you would like to hide these results in order, for example, to practice solving the technique by hand, you can disable the boxes in the left side panel. In addition, in the development of each technique in the main panel, hyperlinks are provided to the tables necessary for the testing process and some graphs that usually accompany each type of contrast.

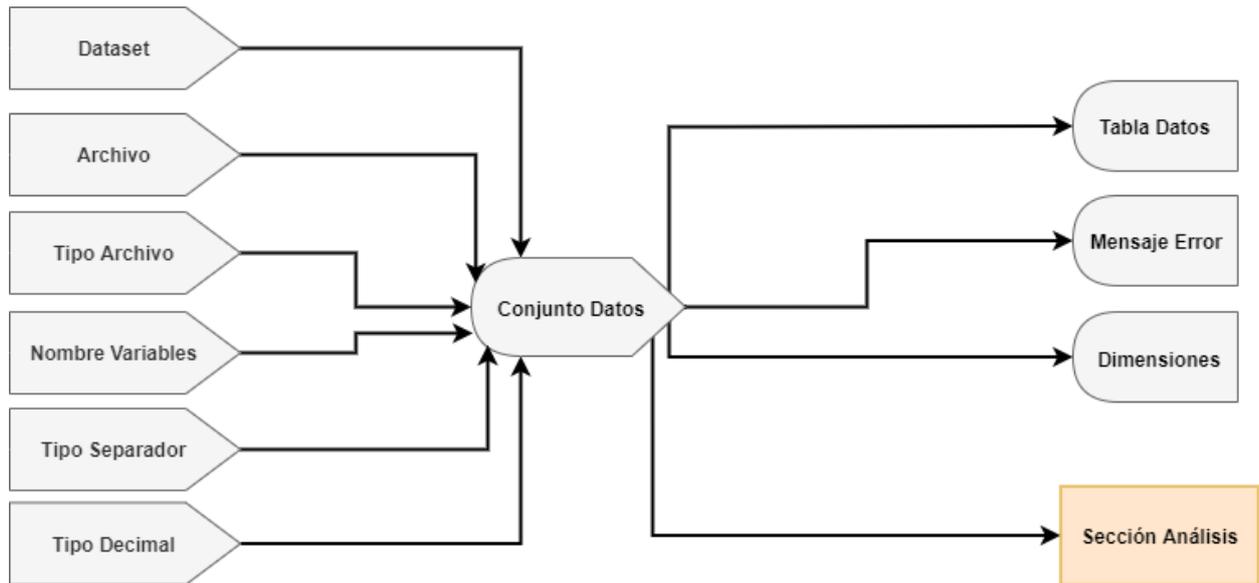
The reactive scheme and user interface for each of these subsections of the "Aprende" section can be found in the Annex 7.3 and Annex 7.2.

In the "Datos" section, is given to the user the possibility to select the dataset to be analysed in the "Análisis" section. In the side panel it is possible to select in the "Dataset" menu whether to use the default dataset called "Empleados" or to import a file.

The "Empleados" dataset consists of twelve variables measured on 64 individuals. These values were obtained from a survey conducted by a company on its employees. This database is very diverse and allows the user to apply all the techniques of the "Análisis" section to a dataset without the need to import one of his own.

The application supports four file types: text (.txt), comma-separated values (.csv), Excel files (.xls or .xlsx) and SPSS data files (.sav). In addition, certain values of the imported file can be specified such as whether variables are in the first row or the type of separator and decimal in text files.

The reactive scheme and the user interface of this section is as follows:



NPS Inicio Aprende Datos Análisis ▾

¿No tiene datos? Puede utilizar el dataset de ejemplo 'Empleados':

Dataset:
Empleados

Importar Datos:

Subir Archivo:
Examinar... Sin Datos

Tipo de Archivo:
CSV

¿Están en la primera fila los nombres de las variables?

Tipo de Separador:*
 Coma: ','
 Tabulador
 Punto y Coma: ';'

Tipo de Decimal:*
 Coma: ','
 Punto: '.'

* El tipo de separador y el tipo de decimal sólo funcionan con archivos .txt

Archivo importado:

Show 10 entries Search:

ID	sexo	ecivil	edad	antiguedad	salario2006	salario2008	salario2010	evaluacion	dias_trabajados	zona_trabajo
1	Mujer	Casado	29	2	18689.00	19170.00	21751.00	Excelente	119	Norte
2	Mujer	Casado	33	11	21282.00	22000.00	23654.00	Excelente	101	Norte
3	Mujer	Casado	36	6	21216.00	21512.00	22369.00	Muy buena	98	Norte
4	Mujer	Casado	44	17	25014.00	26000.00	27589.00	Regular	102	Norte
5	Hombre	Casado	31	2	17928.19	18536.15	19382.17	Muy buena	115	Norte
6	Hombre	Casado	35	9	19034.05	19529.58	20461.82	Muy buena	118	Norte
7	Hombre	Casado	45	11	18469.10	18908.47	19908.79	Muy buena	122	Norte
8	Hombre	Casado	44	19	25645.19	26232.54	27118.86	Buena	118	Norte
9	Hombre	Casado	35	11	22249.47	22863.13	23616.30	Buena	106	Norte
10	Hombre	Casado	35	12	22748.31	23162.60	24085.45	Regular	113	Norte

Showing 1 to 10 of 64 entries

Las dimensiones del archivo importado son: 64 x 13

Previous 1 2 3 4 5 6 7 Next

The "Análisis" section provides the user an environment in which to apply the main non-parametric techniques to one's own dataset or, failing that, to the NPS "Empleados" dataset.

In each of the subsections of this section, is offered to the user the possibility of selecting, in the left-hand side panel, the variable or variables to which the test is to be applied and to vary other types of values of the specific contrast such as, the type of contrast (Bilateral, Upper Unilateral, Lower Unilateral), whether the exact statistic and p-value are desired or whether a continuity correction is desired. The main panel shows the name of the test/subsection the user is in, the hypotheses of the contrast, the values chosen by the user in the side panel and the results of the test, which are mainly the value of the statistic, the associated p-value and a graph accompanying the test.

The non-parametric tests available in this section are the same as those in the "Aprende" section. The user interface of this section is as follows:

The reactive scheme and the user interface of each of the subsections of the "Análisis" section can be found in the Annex 7.3 and Annex 7.2.

The NPS application was uploaded to the Shinyapps.io online server, so that any user could access it without needing to have R installed.

It can be consulted and used by accessing the following QR code or link:

<https://a-est-gal.shinyapps.io/npsapp/>



Undoubtedly, the development of web applications in R with the Shiny library can be very useful in many areas, especially for research and business, greatly shortening the time of data analysis or application of different algorithms.

The developed NPS application aims to be a useful tool for learning the main non-parametric techniques, as well as for their application to different data sets, with the aim of helping researchers, students and any user who needs it to perform non-parametric tests in a simple, fast and free way, as well as not requiring knowledge of the R programming language.

The NPS application offers a "Aprende" section with the aim of disseminating basic concepts about the techniques used in the application, since non-parametric tests are generally less common and less well-known than parametric tests.

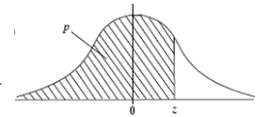
As an improvement perspective for NPS, a study of the target user's opinion could be carried out, with the aim of adding and/or modifying functionalities that could increase the usefulness, usability and value of the application.

7. ANEXOS

7.1 Tablas pruebas no paramétricas

Tabla 1 (N): Probabilidad acumulada de la distribución Normal estándar N(0,1)

$$p = \Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-x^2/2} dx. \text{ Para } z < 0 \text{ utilizar: } \Phi(z) = 1 - \Phi(-z). \text{ Para } z \geq 5 \text{ usar } \Phi(z) \approx 1$$



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000000	0,50398936	0,50797831	0,51196647	0,51595344	0,51993881	0,52392218	0,52790317	0,53188137	0,53585639
0,1	0,53982784	0,54379531	0,54775843	0,55171679	0,55567000	0,55961769	0,56355946	0,56749493	0,57142372	0,57534543
0,2	0,57925971	0,58316616	0,58706442	0,59095412	0,59483487	0,59870633	0,60256811	0,60641987	0,61026125	0,61409188
0,3	0,61791142	0,62171952	0,62551583	0,62930002	0,63307174	0,63683065	0,64057643	0,64430875	0,64802729	0,65173173
0,4	0,65542174	0,65909703	0,66275727	0,66640218	0,67003145	0,67364478	0,67724189	0,68082249	0,68438630	0,68793305
0,5	0,69146246	0,69497427	0,69846821	0,70194403	0,70540148	0,70884031	0,71226028	0,71566115	0,71904269	0,72240468
0,6	0,72574688	0,72906910	0,73237111	0,73565271	0,73891370	0,74215389	0,74537309	0,74857110	0,75174777	0,75490291
0,7	0,75803635	0,76114793	0,76423750	0,76730491	0,77035000	0,77337265	0,77637271	0,77935005	0,78230456	0,78523612
0,8	0,78814460	0,79102991	0,79389195	0,79673061	0,79954581	0,80233746	0,80510548	0,80784980	0,81057035	0,81326706
0,9	0,81593987	0,81858875	0,82121362	0,82381446	0,82639122	0,82894387	0,83147239	0,83397675	0,83645694	0,83891294
1,0	0,84134475	0,84375235	0,84613577	0,84849500	0,85083005	0,85314094	0,85542770	0,85769035	0,85992891	0,86214343
1,1	0,86433394	0,86650049	0,86864312	0,87076189	0,87285685	0,87492806	0,87697560	0,87899952	0,88099989	0,88297680
1,2	0,88493033	0,88686055	0,88876756	0,89065145	0,89251230	0,89435023	0,89616532	0,89795768	0,89972743	0,90147467
1,3	0,90319952	0,90490208	0,90658249	0,90824086	0,90987733	0,91149201	0,91308504	0,91465655	0,91620668	0,91773556
1,4	0,91924334	0,92073016	0,92219616	0,92364149	0,92506630	0,92647074	0,92785496	0,92921912	0,93056338	0,93188788
1,5	0,93319280	0,93447829	0,93574451	0,93699164	0,93821982	0,93942924	0,94062006	0,94179244	0,94294657	0,94408260
1,6	0,94520071	0,94630107	0,94738386	0,94844925	0,94949742	0,95052853	0,95154277	0,95254032	0,95352134	0,95448602
1,7	0,95543454	0,95636706	0,95728378	0,95818486	0,95907049	0,95994084	0,96079610	0,96163643	0,96246202	0,96327304
1,8	0,96406968	0,96485211	0,96562050	0,96637503	0,96711588	0,96784323	0,96855724	0,96925809	0,96994596	0,97062102
1,9	0,97128344	0,97193339	0,97257105	0,97319658	0,97381016	0,97441194	0,97500210	0,97558081	0,97614824	0,97670453
2,0	0,97724987	0,97778441	0,97830831	0,97882173	0,97932484	0,97981778	0,98030073	0,98077383	0,98123723	0,98169110
2,1	0,98213558	0,98257082	0,98299698	0,98341419	0,98382262	0,98422239	0,98461367	0,98499658	0,98537127	0,98573788
2,2	0,98609655	0,98644742	0,98679062	0,98712628	0,98745454	0,98777553	0,98808937	0,98839621	0,98869616	0,98898934
2,3	0,98927589	0,98955592	0,98982956	0,99009692	0,99035813	0,99061329	0,99086253	0,99110596	0,99134368	0,99157581
2,4	0,99180246	0,99202374	0,99223975	0,99245059	0,99265637	0,99285719	0,99305315	0,99324435	0,99343088	0,99361285
2,5	0,99379033	0,99396344	0,99413226	0,99429687	0,99445738	0,99461385	0,99476639	0,99491507	0,99505998	0,99520120
2,6	0,99533881	0,99547289	0,99560351	0,99573076	0,99585470	0,99597541	0,99609297	0,99620744	0,99631889	0,99642740
2,7	0,99653303	0,99663584	0,99673590	0,99683328	0,99692804	0,99702024	0,99710993	0,99719719	0,99728206	0,99736460
2,8	0,99744487	0,99752293	0,99759882	0,99767260	0,99774432	0,99781404	0,99788179	0,99794764	0,99801162	0,99807379
2,9	0,99813419	0,99819286	0,99824984	0,99830519	0,99835894	0,99841113	0,99846180	0,99851100	0,99855876	0,99860511
3,0	0,99865010	0,99869376	0,99873613	0,99877723	0,99881711	0,99885579	0,99889332	0,99892971	0,99896500	0,99899922
3,1	0,99903240	0,99906456	0,99909574	0,99912597	0,99915526	0,99918365	0,99921115	0,99923781	0,99926362	0,99928864
3,2	0,99931286	0,99933633	0,99935905	0,99938105	0,99940235	0,99942297	0,99944294	0,99946226	0,99948096	0,99949906
3,3	0,99951658	0,99953352	0,99954991	0,99956577	0,99958111	0,99959594	0,99961029	0,99962416	0,99963757	0,99965054
3,4	0,99966307	0,99967519	0,99968689	0,99969821	0,99970914	0,99971971	0,99972991	0,99973977	0,99974929	0,99975849
3,5	0,99976737	0,99977595	0,99978423	0,99979222	0,99979994	0,99980738	0,99981457	0,99982151	0,99982820	0,99983466
3,6	0,99984089	0,99984690	0,99985270	0,99985829	0,99986368	0,99986888	0,99987389	0,99987872	0,99988338	0,99988787
3,7	0,99989220	0,99989637	0,99990039	0,99990426	0,99990799	0,99991158	0,99991504	0,99991838	0,99992159	0,99992468
3,8	0,99992765	0,99993052	0,99993327	0,99993593	0,99993848	0,99994094	0,99994331	0,99994558	0,99994777	0,99994988
3,9	0,99995190	0,99995385	0,99995573	0,99995753	0,99995926	0,99996092	0,99996253	0,99996406	0,99996554	0,99996696
4,0	0,99996833	0,99996964	0,99997090	0,99997211	0,99997327	0,99997439	0,99997546	0,99997649	0,99997748	0,99997843
4,1	0,99997934	0,99998022	0,99998106	0,99998186	0,99998263	0,99998338	0,99998409	0,99998477	0,99998542	0,99998605
4,2	0,99998665	0,99998723	0,99998778	0,99998832	0,99998882	0,99998931	0,99998978	0,99999023	0,99999066	0,99999107
4,3	0,99999146	0,99999184	0,99999220	0,99999254	0,99999288	0,99999319	0,99999350	0,99999379	0,99999407	0,99999433
4,4	0,99999459	0,99999483	0,99999506	0,99999529	0,99999550	0,99999571	0,99999590	0,99999609	0,99999627	0,99999644
4,5	0,99999660	0,99999676	0,99999691	0,99999705	0,99999719	0,99999732	0,99999744	0,99999756	0,99999768	0,99999778
4,6	0,99999789	0,99999799	0,99999808	0,99999817	0,99999826	0,99999834	0,99999842	0,99999849	0,99999857	0,99999863
4,7	0,99999870	0,99999876	0,99999882	0,99999888	0,99999893	0,99999898	0,99999903	0,99999908	0,99999912	0,99999917
4,8	0,99999921	0,99999925	0,99999928	0,99999932	0,99999935	0,99999938	0,99999941	0,99999944	0,99999947	0,99999950
4,9	0,99999952	0,99999954	0,99999957	0,99999959	0,99999961	0,99999963	0,99999965	0,99999967	0,99999968	0,99999970

Tablas 2 y 3 (N): Cuantiles inferiores z_p de la distribución Normal estándar N(0,1)

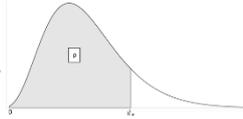
Para $0 < p < 0.5$ usar: $z_p = -z_{1-p}$

p	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,5	0,0000	0,0251	0,0502	0,0753	0,1004	0,1257	0,1510	0,1764	0,2019	0,2275
0,6	0,2533	0,2793	0,3055	0,3319	0,3585	0,3853	0,4125	0,4399	0,4677	0,4959
0,7	0,5244	0,5534	0,5828	0,6128	0,6433	0,6745	0,7063	0,7388	0,7722	0,8064
0,8	0,8416	0,8779	0,9154	0,9542	0,9945	1,0364	1,0803	1,1264	1,1750	1,2265
0,9	1,2816	1,3408	1,4051	1,4758	1,5548	1,6449	1,7507	1,8808	2,0537	2,3263

Especiales	
p	z_p
0,975	1,95996398
0,995	2,57582930
0,999	3,09023231
0,9999	3,71901649

Tabla 4 (Chi2): Cuantiles inferiores $\chi^2_{n,p}$ de la distribución Chi-cuadrado de Pearson

$$p = P(\chi_n^2 \leq \chi_{n,p}^2) = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})} \int_{-\infty}^{\chi_{n,p}^2} x^{(n/2)-1} e^{-x/2} dx.$$



Si $n > 50$ usar: $\chi_{n,p}^2 \approx n \left\{ 1 - \frac{2}{9n} + z_p \cdot \sqrt{\frac{2}{9n}} \right\}^3$ con $z_p =$ cuantil p de $N(0,1)$

$n \setminus p$	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995
1	0,00004	0,00016	0,00098	0,00393	0,01579	0,10153	0,45494	1,32330	2,70554	3,84146	5,02389	6,63490	7,87944
2	0,01003	0,02010	0,05064	0,10259	0,21072	0,57536	1,38629	2,77259	4,60517	5,99146	7,37776	9,21034	10,5966
3	0,07172	0,11483	0,21580	0,35185	0,58437	1,21253	2,36597	4,10834	6,25139	7,81473	9,34840	11,3449	12,8382
4	0,20699	0,29711	0,48442	0,71072	1,06362	1,92256	3,35669	5,38527	7,77944	9,48773	11,1433	13,2767	14,8603
5	0,41174	0,55430	0,83121	1,14548	1,61031	2,67460	4,35146	6,62568	9,23636	11,07050	12,8325	15,0863	16,7496
6	0,67573	0,87209	1,23734	1,63538	2,20413	3,45460	5,34812	7,84080	10,6446	12,59159	14,4494	16,8119	18,5476
7	0,98926	1,23904	1,68987	2,16735	2,83311	4,25485	6,34581	9,03715	12,0170	14,06714	16,0128	18,4753	20,2777
8	1,34441	1,64650	2,17973	2,73264	3,48954	5,07064	7,34412	10,2189	13,3616	15,50731	17,5345	20,0902	21,9550
9	1,73493	2,08790	2,70039	3,32511	4,16816	5,89883	8,34283	11,3888	14,6837	16,91898	19,0228	21,6660	23,5894
10	2,15586	2,55821	3,24697	3,94030	4,86518	6,73720	9,34182	12,5489	15,9872	18,30704	20,4832	23,2093	25,1882
11	2,60322	3,05348	3,81575	4,57481	5,57778	7,58414	10,3410	13,7007	17,2750	19,67514	21,9200	24,7250	26,7568
12	3,07382	3,57057	4,40379	5,22603	6,30380	8,43842	11,3403	14,8454	18,5493	21,02607	23,3367	26,2170	28,2995
13	3,56503	4,10692	5,00875	5,89186	7,04150	9,29907	12,3398	15,9839	19,8119	22,36203	24,7356	27,6882	29,8195
14	4,07467	4,66043	5,62873	6,57063	7,78953	10,1653	13,3393	17,1169	21,0641	23,68479	26,1189	29,1412	31,3193
15	4,60092	5,22935	6,26214	7,26094	8,54676	11,0365	14,3389	18,2451	22,3071	24,99579	27,4884	30,5779	32,8013
16	5,14221	5,81221	6,90766	7,96165	9,31224	11,9122	15,3385	19,3689	23,5418	26,29623	28,8454	31,9999	34,2672
17	5,69722	6,40776	7,56419	8,67176	10,0852	12,7919	16,3382	20,4887	24,7690	27,58711	30,1910	33,4087	35,7185
18	6,26480	7,01491	8,23075	9,39046	10,8649	13,6753	17,3379	21,6049	25,9894	28,86930	31,5264	34,8053	37,1565
19	6,84397	7,63273	8,90652	10,1170	11,6509	14,5620	18,3377	22,7178	27,2036	30,14353	32,8523	36,1909	38,5823
20	7,43384	8,26040	9,59078	10,8508	12,4426	15,4518	19,3374	23,8277	28,4120	31,41043	34,1696	37,5662	39,9968
21	8,03365	8,89720	10,2829	11,5913	13,2396	16,3444	20,3372	24,9348	29,6151	32,67057	35,4789	38,9322	41,4011
22	8,64272	9,54249	10,9823	12,3380	14,0415	17,2396	21,3370	26,0393	30,8133	33,92444	36,7807	40,2894	42,7957
23	9,26042	10,1957	11,6886	13,0905	14,8480	18,1373	22,3369	27,1413	32,0069	35,17246	38,0756	41,6384	44,1813
24	9,88623	10,8564	12,4012	13,8484	15,6587	19,0373	23,3367	28,2412	33,1962	36,41503	39,3641	42,9798	45,5585
25	10,5197	11,5240	13,1197	14,6114	16,4734	19,9393	24,3366	29,3389	34,3816	37,65248	40,6465	44,3141	46,9279
26	11,1602	12,1981	13,8439	15,3792	17,2919	20,8434	25,3365	30,4346	35,5632	38,88514	41,9232	45,6417	48,2899
27	11,8076	12,8785	14,5734	16,1514	18,1139	21,7494	26,3363	31,5284	36,7412	40,11327	43,1945	46,9629	49,6449
28	12,4613	13,5647	15,3079	16,9279	18,9392	22,6572	27,3362	32,6205	37,9159	41,33714	44,4608	48,2782	50,9934
29	13,1211	14,2565	16,0471	17,7084	19,7677	23,5666	28,3361	33,7109	39,0875	42,55697	45,7223	49,5879	52,3356
30	13,7867	14,9535	16,7908	18,4927	20,5992	24,4776	29,3360	34,7997	40,2560	43,77297	46,9792	50,8922	53,6720
31	14,4578	15,6555	17,5387	19,2806	21,4336	25,3901	30,3359	35,8871	41,4217	44,98534	48,2319	52,1914	55,0027
32	15,1340	16,3622	18,2908	20,0719	22,2706	26,3041	31,3359	36,9730	42,5847	46,19426	49,4804	53,4858	56,3281
33	15,8153	17,0735	19,0467	20,8665	23,1102	27,2194	32,3358	38,0575	43,7452	47,39988	50,7251	54,7755	57,6484
34	16,5013	17,7891	19,8063	21,6643	23,9523	28,1361	33,3357	39,1408	44,9032	48,60237	51,9660	56,0609	58,9639
35	17,1918	18,5089	20,5694	22,4650	24,7967	29,0540	34,3356	40,2228	46,0588	49,80185	53,2033	57,3421	60,2748
36	17,8867	19,2327	21,3359	23,2686	25,6433	29,9730	35,3356	41,3036	47,2122	50,99846	54,4373	58,6192	61,5812
37	18,5858	19,9602	22,1056	24,0749	26,4921	30,8933	36,3355	42,3833	48,3634	52,19232	55,6680	59,8925	62,8833
38	19,2889	20,6914	22,8785	24,8839	27,3430	31,8146	37,3355	43,4619	49,5126	53,38354	56,8955	61,1621	64,1814
39	19,9959	21,4262	23,6543	25,6954	28,1958	32,7369	38,3354	44,5395	50,6598	54,57223	58,1201	62,4281	65,4756
40	20,7065	22,1643	24,4330	26,5093	29,0505	33,6603	39,3353	45,6160	51,8051	55,75848	59,3417	63,6907	66,7660
41	21,4208	22,9056	25,2145	27,3256	29,9071	34,5846	40,3353	46,6916	52,9485	56,94239	60,5606	64,9501	68,0527
42	22,1385	23,6501	25,9987	28,1440	30,7654	35,5099	41,3352	47,7663	54,0902	58,12404	61,7768	66,2062	69,3360
43	22,8595	24,3976	26,7854	28,9647	31,6255	36,4361	42,3352	48,8400	55,2302	59,30351	62,9904	67,4593	70,6159
44	23,5837	25,1480	27,5746	29,7875	32,4871	37,3631	43,3352	49,9129	56,3685	60,48089	64,2015	68,7095	71,8926
45	24,3110	25,9013	28,3662	30,6123	33,3504	38,2910	44,3351	50,9849	57,5053	61,65623	65,4102	69,9568	73,1661
46	25,0413	26,6572	29,1601	31,4390	34,2152	39,2197	45,3351	52,0562	58,6405	62,82962	66,6165	71,2014	74,4365
47	25,7746	27,4158	29,9562	32,2676	35,0814	40,1492	46,3350	53,1267	59,7743	64,00111	67,8206	72,4433	75,7041
48	26,5106	28,1770	30,7545	33,0981	35,9491	41,0794	47,3350	54,1964	60,9066	65,17077	69,0226	73,6826	76,9688
49	27,2493	28,9406	31,5549	33,9303	36,8182	42,0104	48,3350	55,2653	62,0375	66,33865	70,2224	74,9195	78,2307
50	27,9907	29,7067	32,3574	34,7643	37,6886	42,9421	49,3349	56,3336	63,1671	67,50481	71,4202	76,1539	79,4900
60	35,5345	37,4849	40,4817	43,1880	46,4589	52,2938	59,3347	66,9815	74,3970	79,08194	83,2977	88,3794	91,9517
70	43,2752	45,4417	48,7576	51,7393	55,3289	61,6983	69,3345	77,5767	85,5270	90,53123	95,0232	100,4252	104,2149
80	51,1719	53,5401	57,1532	60,3915	64,2778	71,1445	79,3343	88,1303	96,5782	101,87947	106,6286	112,3288	116,3211
90	59,1963	61,7541	65,6466	69,1260	73,2911	80,6247	89,3342	98,6499	107,5650	113,14527	118,1359	124,1163	128,2989
100	67,3276	70,0649	74,2219	77,9295	82,3581	90,1332	99,3341	109,1412	118,4980	124,34211	129,5612	135,8067	140,1695
z_p	-2,5758	-2,3263	-1,9599	-1,6448	-1,2815	-0,6745	0,0000	0,6745	1,2815	1,6448	1,9599	2,3263	2,5758

Tabla 5 (B): Probabilidad acumulada de la distribución Binomial $B(n, p_0)$

$p = P(B(n, p_0) \leq k)$ Si $n > 12$ utilizar que $B(n, p_0) \approx N(np_0, np_0(1 - p_0))$

n	p_0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7183	0.6480	0.5748	0.5000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094
	2	0.9978	0.9841	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438
	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6563
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625
	2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164	0.2266
	3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083	0.5000
	4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471	0.7734
	5	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643	0.9375
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352
	2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201	0.1445
	3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770	0.3633
	4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396	0.6367
	5	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115	0.8555
	6	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819	0.9648
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020

10	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195
	2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495	0.0898
	3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614	0.2539
	4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214	0.5000
	5	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342	0.7461
	6	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502	0.9102
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9986	0.9962	0.9909	0.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992	0.9980
11	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
12	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
	1	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0606	0.0302	0.0139	0.0059
	2	0.9848	0.9104	0.7788	0.6174	0.4552	0.3127	0.2001	0.1189	0.0652	0.0327
	3	0.9984	0.9815	0.9306	0.8389	0.7133	0.5696	0.4256	0.2963	0.1911	0.1133
	4	0.9999	0.9972	0.9841	0.9496	0.8854	0.7897	0.6683	0.5328	0.3971	0.2744
	5	1.0000	0.9997	0.9973	0.9883	0.9657	0.9218	0.8513	0.7535	0.6331	0.5000
	6	1.0000	1.0000	0.9997	0.9980	0.9924	0.9784	0.9499	0.9006	0.8262	0.7256
	7	1.0000	1.0000	1.0000	0.9998	0.9988	0.9957	0.9878	0.9707	0.9390	0.8867
	8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9980	0.9941	0.9852	0.9673
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9993	0.9978	0.9941
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002
	1	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0424	0.0196	0.0083	0.0032
	2	0.9804	0.8891	0.7358	0.5583	0.3907	0.2528	0.1513	0.0834	0.0421	0.0193
	3	0.9978	0.9744	0.9078	0.7946	0.6488	0.4925	0.3467	0.2253	0.1345	0.0730
	4	0.9998	0.9957	0.9761	0.9274	0.8424	0.7237	0.5833	0.4382	0.3044	0.1938
	5	1.0000	0.9995	0.9954	0.9806	0.9456	0.8822	0.7873	0.6652	0.5269	0.3872
	6	1.0000	0.9999	0.9993	0.9961	0.9857	0.9614	0.9154	0.8418	0.7393	0.6128
	7	1.0000	1.0000	0.9999	0.9994	0.9972	0.9905	0.9745	0.9427	0.8883	0.8062
	8	1.0000	1.0000	1.0000	0.9999	0.9996	0.9983	0.9944	0.9847	0.9644	0.9270
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9992	0.9972	0.9921	0.9807
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9968
11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	

Tabla 6 (KS): Valores críticos D_{crit} de la prueba de Kolmogorov-Smirnov

(Rechazar H_0 cuando $D_{obs} \geq D_{crit}$)

$n \setminus \alpha$	0,01	0,02	0,05	0,1	0,15	0,2
1	0,99500	0,99000	0,97500	0,95000	0,92500	0,90000
2	0,92929	0,90000	0,84189	0,77639	0,72600	0,68377
3	0,82900	0,78456	0,70760	0,63604	0,59700	0,56481
4	0,73424	0,68887	0,62394	0,56522	0,52500	0,49265
5	0,66853	0,62718	0,56328	0,50945	0,47400	0,44698
6	0,61661	0,57741	0,51926	0,46799	0,43600	0,41037
7	0,57581	0,53844	0,48342	0,43607	0,40500	0,38148
8	0,54179	0,50654	0,45427	0,40962	0,38100	0,35831
9	0,51332	0,47960	0,43001	0,38746	0,36000	0,33910
10	0,48893	0,45662	0,40925	0,36866	0,34200	0,32260
11	0,46770	0,43670	0,39122	0,35242	0,32600	0,30829
12	0,44905	0,41918	0,37543	0,33815	0,31300	0,29577
13	0,43247	0,40362	0,36143	0,32549	0,30200	0,28470
14	0,41762	0,38970	0,34890	0,31417	0,29200	0,27481
15	0,40420	0,37713	0,33760	0,30397	0,28300	0,26588
16	0,39201	0,36571	0,32733	0,29472	0,27400	0,25778
17	0,38086	0,35528	0,31796	0,28627	0,26600	0,25039
18	0,37062	0,34569	0,30936	0,27851	0,25900	0,24360
19	0,36117	0,33685	0,30143	0,27136	0,25200	0,23735
20	0,35241	0,32866	0,29408	0,26473	0,24600	0,23156
21	0,34400	0,32100	0,28700	0,25900	0,24832	0,22600
22	0,33700	0,31400	0,28100	0,25300	0,24261	0,22100
23	0,33000	0,30700	0,27500	0,24700	0,23728	0,21600
24	0,32300	0,30100	0,26900	0,24200	0,23228	0,21200
25	0,32000	0,30349	0,27000	0,24000	0,22759	0,21000
26	0,31100	0,29000	0,25900	0,23300	0,22317	0,20400
27	0,30500	0,28400	0,25400	0,22900	0,21900	0,20000
28	0,30000	0,27900	0,25000	0,22500	0,21505	0,19700
29	0,29500	0,27500	0,24600	0,22100	0,21131	0,19300
30	0,29000	0,27704	0,24000	0,22000	0,20776	0,19000
31	0,28500	0,26600	0,23800	0,21400	0,20438	0,18700
32	0,28100	0,26200	0,23400	0,21100	0,20116	0,18400
33	0,27700	0,25800	0,23100	0,20800	0,19809	0,18200
34	0,27300	0,25400	0,22700	0,20500	0,19516	0,17900
35	0,27000	0,25649	0,23000	0,21000	0,19235	0,18000
36	0,26500	0,24700	0,22100	0,19900	0,18966	0,17400
37	0,26200	0,24400	0,21800	0,19600	0,18708	0,17200
38	0,25800	0,24100	0,21500	0,19400	0,18460	0,17000
39	0,25500	0,23800	0,21300	0,19100	0,18222	0,16800
40	0,25000	0,23993	0,21000	0,19000	0,18000	0,17000
45	0,24263	0,22621	0,20245	0,18244	0,16964	0,15992
50	0,23000	0,21460	0,19000	0,17000	0,16000	0,15000
$n > 50$	$\frac{1,62762}{\sqrt{n}}$	$\frac{1,51743}{\sqrt{n}}$	$\frac{1,35810}{\sqrt{n}}$	$\frac{1,22385}{\sqrt{n}}$	$\frac{1,13795}{\sqrt{n}}$	$\frac{1,07275}{\sqrt{n}}$

Tabla 7 (R): Intervalos para la prueba de las rachas ($\alpha=0.05$)

(Aceptar H_0 : aleatoriedad, cuando $R_{obs} \in [r_{low}, r_{up}]$)

$n = \max(n_1, n_2)$, $m = \min(n_1, n_2)$

n	m	$[r_{low}, r_{up}]$													
20	20	15	27	18	7	7	15	15	4	4	9	11	2	0	5
20	19	14	26	18	6	6	13	15	3	4	7	10	10	7	15
20	18	14	26	18	5	6	11	15	2	3	5	10	9	6	15
20	17	14	25	18	4	5	9	14	14	10	20	10	8	6	14
20	16	13	24	18	3	4	7	14	13	10	19	10	7	6	13
20	15	13	24	18	2	3	5	14	12	9	19	10	6	5	12
20	14	12	23	17	17	12	24	14	11	9	18	10	5	4	11
20	13	11	22	17	16	12	23	14	10	8	17	10	4	4	9
20	12	11	21	17	15	12	22	14	9	8	16	10	3	3	7
20	11	10	20	17	14	11	22	14	8	7	15	10	2	0	5
20	10	10	19	17	13	11	21	14	7	6	14	9	9	6	14
20	9	9	19	17	12	10	20	14	6	6	13	9	8	6	13
20	8	8	16	17	11	10	19	14	5	5	11	9	7	5	13
20	7	7	15	17	10	9	18	14	4	4	9	9	6	5	12
20	6	7	13	17	9	8	19	14	3	3	7	9	5	4	11
20	5	6	11	17	8	8	16	14	2	3	5	9	4	4	9
20	4	5	9	17	7	7	15	13	13	9	19	9	3	3	7
20	3	4	7	17	6	6	13	13	12	8	19	9	2	0	5
20	2	3	5	17	5	5	11	13	11	8	18	8	8	5	13
19	19	14	26	17	4	5	9	13	10	8	17	8	7	5	12
19	18	14	25	17	3	4	7	13	9	7	16	8	6	4	11
19	17	13	25	17	2	3	5	13	8	7	15	8	5	4	10
19	16	13	24	16	16	12	22	13	7	6	14	8	4	4	9
19	15	12	23	16	15	11	22	13	6	6	13	8	3	3	7
19	14	12	22	16	14	11	21	13	5	5	11	8	2	0	5
19	13	11	21	16	13	10	21	13	4	4	9	7	7	4	12
19	12	11	21	16	12	10	21	13	3	3	7	7	6	4	11
19	11	10	20	16	11	9	19	13	2	3	5	7	5	4	10
19	10	9	19	16	10	9	18	12	12	8	18	7	4	3	9
19	9	9	19	16	9	8	19	12	11	8	17	7	3	3	7
19	8	8	16	16	8	7	16	12	10	8	16	7	2	0	5
19	7	7	15	16	7	7	15	12	9	7	15	6	6	4	10
19	6	7	13	16	6	6	13	12	8	7	15	6	5	4	9
19	5	6	11	16	5	5	11	12	7	6	13	6	4	3	8
19	4	5	9	16	4	5	9	12	6	5	12	6	3	3	7
19	3	4	7	16	3	4	7	12	5	5	11	6	2	0	5
19	2	3	5	16	2	3	5	12	4	4	9	5	5	3	9
18	18	13	25	15	15	11	21	12	3	3	7	5	4	3	8
18	17	13	24	15	14	10	21	12	2	3	5	5	3	0	7
18	16	12	24	15	13	10	20	11	11	8	16	5	2	0	5
18	15	12	23	15	12	9	19	11	10	7	16	4	4	0	8
18	14	11	22	15	11	9	18	11	9	7	15	4	3	0	7
18	13	11	21	15	10	8	17	11	8	6	14	4	2	0	5
18	12	10	21	15	9	8	19	11	7	6	13	3	3	0	6
18	11	10	19	15	8	7	15	11	6	5	12	3	2	0	5
18	10	9	18	15	7	7	14	11	5	5	11	2	2	0	4
18	9	9	19	15	6	6	13	11	4	4	9				
18	8	8	16	15	5	5	11	11	3	3	7				

Tabla 9 (U): Probabilidades inferiores del estadístico U de la prueba de la suma de rangos de Mann-Whitney

$m = \max \{n_1, n_2\}$ $n = \min \{n_1, n_2\}$ $u = u_{obs}$ / Para probabilidades superiores: $P(U \geq u) = 1 - P(U < u) = 1 - P(U \leq u - 1)$

m	n	u	$P(U \leq u)$
3	1	0	0,250
3	1	1	0,500
3	1	2	0,750
3	2	0	0,100
3	2	1	0,200
3	2	2	0,400
3	2	3	0,600
3	3	0	0,050
3	3	1	0,100
3	3	2	0,200
3	3	3	0,350
3	3	4	0,500
3	3	5	0,650

m	n	u	$P(U \leq u)$
5	1	0	0,167
5	1	1	0,333
5	1	2	0,500
5	1	3	0,667
5	2	0	0,047
5	2	1	0,095
5	2	2	0,190
5	2	3	0,286
5	2	4	0,429
5	2	5	0,571
5	3	0	0,018
5	3	1	0,036
5	3	2	0,071
5	3	3	0,125
5	3	4	0,196
5	3	5	0,286
5	3	6	0,393
5	3	7	0,500
5	3	8	0,607
5	4	0	0,008
5	4	1	0,016
5	4	2	0,032
5	4	3	0,056
5	4	4	0,095
5	4	5	0,143
5	4	6	0,206
5	4	7	0,278
5	4	8	0,365
5	4	9	0,452
5	4	10	0,548
5	5	0	0,004
5	5	1	0,008
5	5	2	0,016
5	5	3	0,028
5	5	4	0,048
5	5	5	0,075
5	5	6	0,111
5	5	7	0,155
5	5	8	0,210
5	5	9	0,274
5	5	10	0,345
5	5	11	0,421
5	5	12	0,500
5	5	13	0,579

m	n	u	$P(U \leq u)$
6	1	0	0,143
6	1	1	0,286
6	1	2	0,428
6	1	3	0,571
6	2	0	0,036
6	2	1	0,071
6	2	2	0,143
6	2	3	0,214
6	2	4	0,321
6	2	5	0,429
6	2	6	0,571
6	3	0	0,012
6	3	1	0,024
6	3	2	0,048
6	3	3	0,083
6	3	4	0,131
6	3	5	0,190
6	3	6	0,274
6	3	7	0,357
6	3	8	0,452
6	3	9	0,548
6	4	0	0,005
6	4	1	0,010
6	4	2	0,019
6	4	3	0,033
6	4	4	0,057
6	4	5	0,086
6	4	6	0,129
6	4	7	0,176
6	4	8	0,238
6	4	9	0,305
6	4	10	0,381
6	4	11	0,457
6	4	12	0,545
6	5	0	0,002
6	5	1	0,004
6	5	2	0,009
6	5	3	0,015
6	5	4	0,026
6	5	5	0,041
6	5	6	0,063
6	5	7	0,089
6	5	8	0,123
6	5	9	0,165
6	5	10	0,214
6	5	11	0,268

6	5	12	0,331
6	5	13	0,396
6	5	14	0,465
6	5	15	0,535
6	6	0	0,001
6	6	1	0,002
6	6	2	0,004
6	6	3	0,008
6	6	4	0,013
6	6	5	0,021
6	6	6	0,032
6	6	7	0,047
6	6	8	0,066
6	6	9	0,090
6	6	10	0,120
6	6	11	0,155
6	6	12	0,197
6	6	13	0,242
6	6	14	0,294
6	6	15	0,350
6	6	16	0,409
6	6	17	0,469
6	6	18	0,531

7	3	7	0,258
7	3	8	0,333
7	3	9	0,417
7	3	10	0,500
7	3	11	0,583
7	4	0	0,003
7	4	1	0,006
7	4	2	0,012
7	4	3	0,021
7	4	4	0,036
7	4	5	0,055
7	4	6	0,082
7	4	7	0,115
7	4	8	0,158
7	4	9	0,206
7	4	10	0,264
7	4	11	0,324
7	4	12	0,394
7	4	13	0,464
7	4	14	0,538
7	5	0	0,001
7	5	1	0,003
7	5	2	0,005
7	5	3	0,009
7	5	4	0,015
7	5	5	0,024
7	5	6	0,037
7	5	7	0,053
7	5	8	0,074
7	5	9	0,101
7	5	10	0,134
7	5	11	0,172
7	5	12	0,216
7	5	13	0,265
7	5	14	0,319
7	5	15	0,378
7	5	16	0,438
7	5	17	0,500
7	5	18	0,562
7	6	0	0,001
7	6	1	0,001
7	6	2	0,002
7	6	3	0,004
7	6	4	0,007
7	6	5	0,011
7	6	6	0,017
7	6	7	0,026

m	n	u	$P(U \leq u)$
4	1	0	0,200
4	1	1	0,400
4	1	2	0,600
4	2	0	0,067
4	2	1	0,133
4	2	2	0,267
4	2	3	0,400
4	2	4	0,600
4	3	0	0,028
4	3	1	0,057
4	3	2	0,114
4	3	3	0,200
4	3	4	0,314
4	3	5	0,429
4	3	6	0,571
4	4	0	0,014
4	4	1	0,029
4	4	2	0,057
4	4	3	0,100
4	4	4	0,171
4	4	5	0,243
4	4	6	0,343
4	4	7	0,443
4	4	8	0,557

m	n	u	$P(U \leq u)$
7	1	0	0,125
7	1	1	0,250
7	1	2	0,375
7	1	3	0,500
7	1	4	0,625
7	2	0	0,028
7	2	1	0,056
7	2	2	0,111
7	2	3	0,167
7	2	4	0,250
7	2	5	0,333
7	2	6	0,444
7	2	7	0,556
7	3	0	0,008
7	3	1	0,017
7	3	2	0,033
7	3	3	0,058
7	3	4	0,092
7	3	5	0,133
7	3	6	0,192

Tabla 9 (U) cont.:

7	6	8	0,037
7	6	9	0,051
7	6	10	0,069
7	6	11	0,090
7	6	12	0,117
7	6	13	0,147
7	6	14	0,183
7	6	15	0,223
7	6	16	0,267
7	6	17	0,314
7	6	18	0,365
7	6	19	0,418
7	6	20	0,473
7	6	21	0,527
7	7	0	0,000
7	7	1	0,001
7	7	2	0,001
7	7	3	0,002
7	7	4	0,003
7	7	5	0,006
7	7	6	0,009
7	7	7	0,013
7	7	8	0,019
7	7	9	0,027
7	7	10	0,036
7	7	11	0,049
7	7	12	0,064
7	7	13	0,082
7	7	14	0,104
7	7	15	0,130
7	7	16	0,159
7	7	17	0,191
7	7	18	0,228
7	7	19	0,267
7	7	20	0,310
7	7	21	0,355
7	7	22	0,402
7	7	23	0,451
7	7	24	0,500
7	7	25	0,549

<i>m</i>	<i>n</i>	<i>u</i>	<i>P(U ≤ u)</i>
8	1	0	0,111
8	1	1	0,222
8	1	2	0,333
8	1	3	0,444

8	1	4	0,556
8	2	0	0,022
8	2	1	0,044
8	2	2	0,089
8	2	3	0,133
8	2	4	0,200
8	2	5	0,267
8	2	6	0,356
8	2	7	0,444
8	2	8	0,556
8	3	0	0,006
8	3	1	0,012
8	3	2	0,024
8	3	3	0,042
8	3	4	0,067
8	3	5	0,097
8	3	6	0,139
8	3	7	0,188
8	3	8	0,248
8	3	9	0,315
8	3	10	0,387
8	3	11	0,461
8	3	12	0,539
8	4	0	0,002
8	4	1	0,004
8	4	2	0,008
8	4	3	0,014
8	4	4	0,024
8	4	5	0,036
8	4	6	0,055
8	4	7	0,077
8	4	8	0,107
8	4	9	0,141
8	4	10	0,184
8	4	11	0,230
8	4	12	0,285
8	4	13	0,341
8	4	14	0,404
8	4	15	0,467
8	4	16	0,533
8	5	0	0,001
8	5	1	0,002
8	5	2	0,003
8	5	3	0,005
8	5	4	0,009
8	5	5	0,015
8	5	6	0,023
8	5	7	0,033
8	5	8	0,047
8	5	9	0,064

8	5	10	0,085
8	5	11	0,111
8	5	12	0,142
8	5	13	0,177
8	5	14	0,217
8	5	15	0,262
8	5	16	0,311
8	5	17	0,362
8	5	18	0,416
8	5	19	0,472
8	5	20	0,528
8	6	0	0,000
8	6	1	0,001
8	6	2	0,001
8	6	3	0,002
8	6	4	0,004
8	6	5	0,006
8	6	6	0,010
8	6	7	0,015
8	6	8	0,021
8	6	9	0,030
8	6	10	0,041
8	6	11	0,054
8	6	12	0,071
8	6	13	0,091
8	6	14	0,114
8	6	15	0,141
8	6	16	0,172
8	6	17	0,207
8	6	18	0,245
8	6	19	0,286
8	6	20	0,331
8	6	21	0,377
8	6	22	0,426
8	6	23	0,475
8	6	24	0,525
8	7	0	0,000
8	7	1	0,000
8	7	2	0,001
8	7	3	0,001
8	7	4	0,002
8	7	5	0,003
8	7	6	0,005
8	7	7	0,007
8	7	8	0,010
8	7	9	0,014
8	7	10	0,020
8	7	11	0,027
8	7	12	0,036
8	7	13	0,047

8	7	14	0,060
8	7	15	0,076
8	7	16	0,095
8	7	17	0,116
8	7	18	0,140
8	7	19	0,168
8	7	20	0,198
8	7	21	0,232
8	7	22	0,268
8	7	23	0,306
8	7	24	0,347
8	7	25	0,389
8	7	26	0,433
8	7	27	0,478
8	7	28	0,522
8	8	0	0,000
8	8	1	0,000
8	8	2	0,000
8	8	3	0,001
8	8	4	0,001
8	8	5	0,001
8	8	6	0,002
8	8	7	0,003
8	8	8	0,005
8	8	9	0,007
8	8	10	0,010
8	8	11	0,014
8	8	12	0,019
8	8	13	0,025
8	8	14	0,032
8	8	15	0,041
8	8	16	0,052
8	8	17	0,065
8	8	18	0,080
8	8	19	0,097
8	8	20	0,117
8	8	21	0,139
8	8	22	0,164
8	8	23	0,191
8	8	24	0,221
8	8	25	0,253
8	8	26	0,287
8	8	27	0,323
8	8	28	0,360
8	8	29	0,399
8	8	30	0,439
8	8	31	0,480
8	8	32	0,520

Tabla 10 (H): Valores críticos H_{crit} de la prueba de Kruskal-Wallis

(Rechazar H_0 cuando $H_{obs} \geq H_{crit}$)

n_1	n_2	n_3	n_4	n_5	$\alpha = 0.05$	$\alpha = 0.01$
3	2	2			4,714	
3	3	1			5,143	
3	3	2			5,361	
3	3	3			5,600	
4	2	2			5,333	
4	3	1			5,208	
4	3	2			5,444	6,444
4	3	3			5,791	6,745
4	4	1			4,967	6,667
4	4	2			5,455	7,036
4	4	3			5,598	7,144
4	4	4			5,692	7,654
5	2	1			5,000	
5	2	2			5,160	6,533
5	3	1			4,960	
5	3	2			5,251	6,909
5	3	3			5,648	7,079
5	4	1			4,985	6,955
5	4	2			5,273	7,205
5	4	3			5,656	7,445
5	4	4			5,657	7,760
5	5	1			5,127	7,309
5	5	2			5,338	7,338
5	5	3			5,705	7,578
5	5	4			5,666	7,823
5	5	5			5,780	8,000
6	2	1			4,822	
6	2	2			5,345	6,982
6	3	1			4,855	
6	3	2			5,348	6,970
6	3	3			5,615	7,410
6	4	1			4,947	7,106
6	4	2			5,340	7,340
6	4	3			5,610	7,500
6	4	4			5,681	7,795
6	5	1			4,990	7,182
6	5	2			5,338	7,376
6	5	3			5,602	7,590
6	5	4			5,661	7,936
6	5	5			5,729	8,028
6	6	1			4,945	7,121
6	6	2			5,410	7,467
6	6	3			5,625	7,725
6	6	4			5,724	8,000
6	6	5			5,765	8,124
6	6	6			5,801	8,222

n_1	n_2	n_3	n_4	n_5	$\alpha = 0.05$	$\alpha = 0.01$
7	7	7			5,819	8,378
8	8	8			5,805	8,465
2	2	2	1		5,679	
2	2	2	2		6,167	6,667
3	2	2	1		5,833	
3	2	2	2		5,333	7,133
3	3	1	1		6,333	
3	3	2	1		6,244	7,200
3	3	2	2		6,527	7,636
3	3	3	1		6,600	7,400
3	3	3	2		6,727	8,105
3	3	3	3		7,000	8,538
4	2	1	1		5,833	
4	2	2	1		6,133	7,000
4	2	2	2		6,545	7,391
4	3	1	1		6,178	7,067
4	3	2	1		6,309	7,455
4	3	2	2		6,621	7,871
4	3	3	1		6,545	7,758
4	3	3	2		6,795	8,333
4	3	3	3		6,984	8,659
4	4	1	1		5,945	7,909
4	4	2	1		6,386	7,886
4	4	2	2		6,731	8,346
4	4	3	1		6,635	8,231
4	4	3	2		6,874	8,621
4	4	3	3		7,038	8,876
4	4	4	1		6,725	8,588
4	4	4	2		6,957	8,871
4	4	4	3		7,142	9,075
4	4	4	4		7,235	9,287
2	2	2	1	1	6,750	
2	2	2	2	1	7,133	7,533
2	2	2	2	2	7,418	8,291
3	2	1	1	1	6,583	
3	2	2	1	1	6,800	7,600
3	2	2	2	1	7,309	8,127
3	2	2	2	2	7,682	8,682
3	3	1	1	1	7,111	
3	3	2	1	1	7,200	8,073
3	3	2	2	1	7,591	8,576
3	3	2	2	2	7,910	9,115
3	3	3	1	1	7,576	8,424
3	3	3	2	1	7,759	9,051
3	3	3	2	2	8,044	9,505
3	3	3	3	1	8,000	9,451
3	3	3	3	2	8,200	9,876
3	3	3	3	3	8,333	10,200

Tabla 11 (S): Valores críticos S_{crit} de la prueba de Friedman

(Rechazar H_0 cuando $S_{obs} \geq S_{crit}$)

<i>k</i>	<i>n</i>	<i>α bilateral</i>			
		<i>0,1</i>	<i>0,05</i>	<i>0,025</i>	<i>0,01</i>
<i>k variables (columnas)</i>	<i>n individuos (filas)</i>				
2	2	2,0000	2,0000	2,0000	2,0000
2	3	3,0000	3,0000	3,0000	3,0000
2	4	4,0000	4,0000	4,0000	4,0000
2	5	1,8000	5,0000	5,0000	5,0000
2	6	2,6667	2,6667	2,6667	2,6667
2	7	3,5714	3,5714	3,5714	7,0000
2	8	2,0000	4,5000	4,5000	4,5000
2	9	2,7778	2,7778	5,4444	5,4444
2	10	3,6000	3,6000	3,6000	6,4000
2	11	2,2727	4,4545	4,4545	7,3636
2	12	3,0000	3,0000	5,3333	5,3333
2	13	1,9231	3,7692	3,7692	6,2308
2	14	2,5714	4,5714	4,5714	7,1429
2	15	3,2667	3,2667	5,4000	5,4000
3	2	4,0000	4,0000	4,0000	4,0000
3	3	4,6667	4,6667	6,0000	6,0000
3	4	4,5000	6,0000	6,5000	6,5000
3	5	4,8000	5,2000	6,4000	7,6000
3	6	4,3333	6,3333	7,0000	8,3333
3	7	4,5714	6,0000	7,1429	8,0000
3	8	4,7500	5,2500	7,0000	7,7500
3	9	4,6667	6,0000	6,8889	8,6667
3	10	4,2000	5,6000	7,4000	8,6000
3	11	4,9091	5,6364	7,0909	8,9091
3	12	4,6667	6,1667	7,1667	8,6667
3	13	4,5841	5,9469	7,2920	9,0796
3	14	4,4286	5,5714	7,0000	9,0000
3	15	4,8000	5,7333	6,9333	8,5333
4	2	5,4000	5,4000	6,0000	6,0000
4	3	5,8000	7,0000	7,4000	8,2000
4	4	6,0000	7,5000	8,1000	9,3000
4	5	6,1200	7,3200	8,2800	9,7200
4	6	6,2000	7,4000	8,6000	10,0000
4	7	6,2571	7,6286	8,6571	10,3714
4	8	6,1500	7,5000	8,8500	10,3500
4	9	6,0667	7,5333	8,7333	10,4667
4	10	6,2400	7,5600	8,8800	10,6800
4	11	6,1636	7,5818	8,8909	10,6364
4	12	6,1000	7,6000	9,0000	10,7000
4	13	6,0462	7,6154	9,0000	10,7539
4	14	6,2571	7,6286	9,0000	10,8857
4	15	6,2000	7,5600	9,0800	10,7600
5	2	6,8000	7,2000	7,6000	7,6000
5	3	7,2000	8,2667	9,3333	9,8667
5	4	7,4000	8,6000	9,6000	11,0000
5	5	7,5200	8,8000	10,0800	11,5200
5	6	7,6000	8,9333	10,2667	11,8667
5	7	7,6571	9,0286	10,5143	12,0043
5	8	7,7000	9,2000	10,6400	12,2000
5	9	7,6444	9,1556	10,5778	12,3556
5	10	7,6800	9,2000	10,6400	12,4000
5	11	7,7091	9,2363	10,7636	12,5818
5	12	7,6667	9,2667	10,7333	12,5333
5	13	7,6923	9,2923	10,7692	12,7385
5	14	7,7143	9,3143	10,8000	12,7429
5	15	7,6800	9,3333	10,8267	12,7467

Tabla 12 (Tau): Valores críticos τ_{crit} de la prueba de Kendall

(Rechazar H_0 cuando $|\tau_{obs}| \geq \tau_{crit}$)

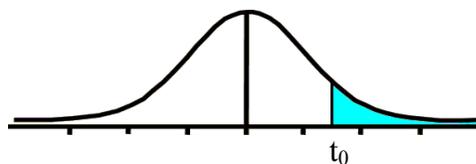
$n \setminus \alpha$	0,2	0,1	0,05	0,02	0,01	0,002
4	1,000	1,000	--	--	--	--
5	0,800	0,800	1,000	1,000	--	--
6	0,600	0,733	0,867	0,867	1,000	--
7	0,524	0,619	0,714	0,810	0,905	1,000
8	0,429	0,571	0,643	0,714	0,786	0,857
9	0,389	0,500	0,556	0,667	0,722	0,833
10	0,378	0,467	0,551	0,600	0,644	0,778
11	0,345	0,418	0,491	0,564	0,600	0,709
12	0,303	0,394	0,455	0,545	0,576	0,667
13	0,308	0,359	0,436	0,513	0,564	0,641
14	0,275	0,363	0,407	0,473	0,516	0,604
15	0,276	0,333	0,390	0,467	0,505	0,581
16	0,250	0,317	0,383	0,433	0,483	0,567
17	0,250	0,309	0,368	0,426	0,471	0,544
18	0,242	0,294	0,346	0,412	0,451	0,529
19	0,228	0,287	0,333	0,392	0,439	0,509
20	0,221	0,274	0,326	0,379	0,421	0,495
21	0,210	0,267	0,314	0,371	0,410	0,486
22	0,203	0,264	0,307	0,359	0,394	0,472
23	0,202	0,257	0,296	0,352	0,391	0,455
24	0,196	0,246	0,290	0,341	0,377	0,449
25	0,193	0,240	0,287	0,333	0,367	0,440
26	0,188	0,237	0,280	0,329	0,360	0,428
27	0,179	0,231	0,271	0,322	0,356	0,419
28	0,180	0,228	0,265	0,312	0,344	0,413
29	0,172	0,222	0,261	0,310	0,340	0,404
30	0,172	0,218	0,255	0,301	0,333	0,393

Tabla 13 (RS): Valores críticos r_{crit} de la prueba de Spearman

(Rechazar H_0 cuando $|r_{obs}| \geq r_{crit}$)

$n \setminus \alpha$	0,2	0,1	0,05	0,02	0,01	0,002
4	1,000	1,000	--	--	--	--
5	0,800	0,900	1,000	1,000	--	--
6	0,657	0,829	0,886	0,943	1,000	--
7	0,571	0,714	0,786	0,893	0,929	1,000
8	0,524	0,643	0,738	0,833	0,881	0,952
9	0,483	0,600	0,700	0,783	0,833	0,917
10	0,455	0,564	0,648	0,745	0,794	0,879
11	0,427	0,536	0,618	0,709	0,755	0,845
12	0,406	0,503	0,587	0,678	0,727	0,818
13	0,385	0,484	0,560	0,648	0,703	0,791
14	0,367	0,464	0,538	0,626	0,679	0,771
15	0,354	0,446	0,521	0,604	0,654	0,750
16	0,341	0,429	0,503	0,582	0,635	0,729
17	0,328	0,414	0,488	0,566	0,618	0,711
18	0,317	0,401	0,472	0,550	0,600	0,692
19	0,309	0,391	0,460	0,535	0,584	0,675
20	0,299	0,380	0,447	0,522	0,570	0,662
21	0,292	0,370	0,436	0,509	0,556	0,647
22	0,284	0,361	0,425	0,497	0,544	0,633
23	0,278	0,353	0,416	0,486	0,532	0,621
24	0,271	0,344	0,407	0,476	0,521	0,609
25	0,265	0,337	0,398	0,466	0,511	0,597
26	0,259	0,331	0,390	0,457	0,501	0,586
27	0,255	0,324	0,383	0,449	0,492	0,576
28	0,250	0,318	0,375	0,441	0,483	0,567
29	0,245	0,312	0,368	0,433	0,475	0,558
30	0,240	0,306	0,362	0,425	0,467	0,549

Tabla t-Student



Grados de libertad	0.25	0.1	0.05	0.025	0.01	0.005
1	1.0000	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.8165	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.7649	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.7407	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.7267	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.7176	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.7111	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.7064	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.7027	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.6998	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.6974	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.6955	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.6938	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.6924	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.6912	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.6901	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.6892	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.6884	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.6876	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.6870	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.6864	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.6858	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.6853	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.6848	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.6844	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.6840	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.6837	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.6834	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.6830	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.6828	1.3104	1.6973	2.0423	2.4573	2.7500
31	0.6825	1.3095	1.6955	2.0395	2.4528	2.7440
32	0.6822	1.3086	1.6939	2.0369	2.4487	2.7385
33	0.6820	1.3077	1.6924	2.0345	2.4448	2.7333
34	0.6818	1.3070	1.6909	2.0322	2.4411	2.7284
35	0.6816	1.3062	1.6896	2.0301	2.4377	2.7238
36	0.6814	1.3055	1.6883	2.0281	2.4345	2.7195
37	0.6812	1.3049	1.6871	2.0262	2.4314	2.7154
38	0.6810	1.3042	1.6860	2.0244	2.4286	2.7116
39	0.6808	1.3036	1.6849	2.0227	2.4258	2.7079
40	0.6807	1.3031	1.6839	2.0211	2.4233	2.7045
41	0.6805	1.3025	1.6829	2.0195	2.4208	2.7012
42	0.6804	1.3020	1.6820	2.0181	2.4185	2.6981
43	0.6802	1.3016	1.6811	2.0167	2.4163	2.6951
44	0.6801	1.3011	1.6802	2.0154	2.4141	2.6923
45	0.6800	1.3007	1.6794	2.0141	2.4121	2.6896
46	0.6799	1.3002	1.6787	2.0129	2.4102	2.6870
47	0.6797	1.2998	1.6779	2.0117	2.4083	2.6846
48	0.6796	1.2994	1.6772	2.0106	2.4066	2.6822
49	0.6795	1.2991	1.6766	2.0096	2.4049	2.6800

50	0.6794	1.2987	1.6759	2.0086	2.4033	2.6778
51	0.6793	1.2984	1.6753	2.0076	2.4017	2.6757
52	0.6792	1.2980	1.6747	2.0066	2.4002	2.6737
53	0.6791	1.2977	1.6741	2.0057	2.3988	2.6718
54	0.6791	1.2974	1.6736	2.0049	2.3974	2.6700
55	0.6790	1.2971	1.6730	2.0040	2.3961	2.6682
56	0.6789	1.2969	1.6725	2.0032	2.3948	2.6665
57	0.6788	1.2966	1.6720	2.0025	2.3936	2.6649
58	0.6787	1.2963	1.6716	2.0017	2.3924	2.6633
59	0.6787	1.2961	1.6711	2.0010	2.3912	2.6618
60	0.6786	1.2958	1.6706	2.0003	2.3901	2.6603
61	0.6785	1.2956	1.6702	1.9996	2.3890	2.6589
62	0.6785	1.2954	1.6698	1.9990	2.3880	2.6575
63	0.6784	1.2951	1.6694	1.9983	2.3870	2.6561
64	0.6783	1.2949	1.6690	1.9977	2.3860	2.6549
65	0.6783	1.2947	1.6686	1.9971	2.3851	2.6536
66	0.6782	1.2945	1.6683	1.9966	2.3842	2.6524
67	0.6782	1.2943	1.6679	1.9960	2.3833	2.6512
68	0.6781	1.2941	1.6676	1.9955	2.3824	2.6501
69	0.6781	1.2939	1.6672	1.9949	2.3816	2.6490
70	0.6780	1.2938	1.6669	1.9944	2.3808	2.6479
71	0.6780	1.2936	1.6666	1.9939	2.3800	2.6469
72	0.6779	1.2934	1.6663	1.9935	2.3793	2.6458
73	0.6779	1.2933	1.6660	1.9930	2.3785	2.6449
74	0.6778	1.2931	1.6657	1.9925	2.3778	2.6439
75	0.6778	1.2929	1.6654	1.9921	2.3771	2.6430
76	0.6777	1.2928	1.6652	1.9917	2.3764	2.6421
77	0.6777	1.2926	1.6649	1.9913	2.3758	2.6412
78	0.6776	1.2925	1.6646	1.9908	2.3751	2.6403
79	0.6776	1.2924	1.6644	1.9905	2.3745	2.6395
80	0.6776	1.2922	1.6641	1.9901	2.3739	2.6387
81	0.6775	1.2921	1.6639	1.9897	2.3733	2.6379
82	0.6775	1.2920	1.6636	1.9893	2.3727	2.6371
83	0.6775	1.2918	1.6634	1.9890	2.3721	2.6364
84	0.6774	1.2917	1.6632	1.9886	2.3716	2.6356
85	0.6774	1.2916	1.6630	1.9883	2.3710	2.6349
86	0.6774	1.2915	1.6628	1.9879	2.3705	2.6342
87	0.6773	1.2914	1.6626	1.9876	2.3700	2.6335
88	0.6773	1.2912	1.6624	1.9873	2.3695	2.6329
89	0.6773	1.2911	1.6622	1.9870	2.3690	2.6322
90	0.6772	1.2910	1.6620	1.9867	2.3685	2.6316
91	0.6772	1.2909	1.6618	1.9864	2.3680	2.6309
92	0.6772	1.2908	1.6616	1.9861	2.3676	2.6303
93	0.6771	1.2907	1.6614	1.9858	2.3671	2.6297
94	0.6771	1.2906	1.6612	1.9855	2.3667	2.6291
95	0.6771	1.2905	1.6611	1.9852	2.3662	2.6286
96	0.6771	1.2904	1.6609	1.9850	2.3658	2.6280
97	0.6770	1.2903	1.6607	1.9847	2.3654	2.6275
98	0.6770	1.2903	1.6606	1.9845	2.3650	2.6269
99	0.6770	1.2902	1.6604	1.9842	2.3646	2.6264
100	0.6770	1.2901	1.6602	1.9840	2.3642	2.6259
∞	0.6745	1.2816	1.6449	1.9600	2.3263	2.5758

7.2 Interfaz de usuario NPS

7.2.1 INICIO

NPS Inicio Aprende Datos Análisis ▾

NPS

¡Bienvenido a NPS!

Esta aplicación web ofrece herramientas para aprender sobre algunas técnicas no paramétricas así como la posibilidad de importar un set de datos y aplicar test no paramétricos a estos datos importados.

A continuación se describe el funcionamiento general de cada uno de los apartados principales de la aplicación.

Aprende

En la sección Aprende se explica el proceso de cada una de las pruebas no paramétricas con un ejemplo práctico, incluyendo las fórmulas utilizadas y enlaces a las tablas de cuantiles necesarias en cada caso. Es posible practicar estos ejercicios las veces que se necesite, ocultando las soluciones y generando nuevos datos. Después, se muestran las soluciones para comprobar que se han realizado las cuentas correctamente y que se ha comprendido el proceso del test no paramétrico en cuestión. Además, este apartado ofrece enlaces en los que podemos encontrar más información sobre las pruebas.

Los tests no paramétricos que se incluyen en este apartado son:

- **Aleatoriedad: Rachas W-W** → Contraste de aleatoriedad: Prueba de las rachas de Wald-Wolfowitz.
- **Bondad de Ajuste: K-S $F = F_0$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = F_0$.
- **Bondad de Ajuste: K-S $F = G$** → Contraste de bondad de ajuste: Prueba de Kolmogorov-Smirnov $F = G$.
- **Localización de una muestra: W** → Contraste de localización de una muestra: Prueba de los rangos con signo de Wilcoxon.
- **Dos muestras independientes: U M-W** → Contraste para dos muestras independientes: Prueba U de Mann-Whitney.
- **Dos muestras relacionadas: Apareados W** → Contraste para dos muestras relacionadas: Prueba para datos apareados de Wilcoxon.
- **k muestras independientes: K-W** → Contraste para k muestras independientes: Prueba de Kruskal-Wallis.
- **k muestras relacionadas: Friedman** → Contraste para k muestras relacionadas: Prueba de Friedman.
- **Independencia: Kendall** → Contraste de independencia: Prueba de Kendall.
- **Independencia: Spearman** → Contraste de independencia: Prueba de Spearman.



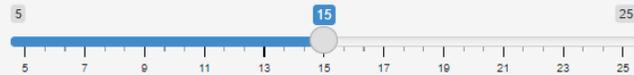
7.2.2 APRENDE: ALEATORIEDAD: RACHAS W-W

Generar Datos:

Variable:

Lanzar una Moneda

Tamaño de la muestra:



Nueva Muestra

Mostrar Soluciones:

- Mostrar n , n_1 y n_2
- Mostrar R_{obs}
- Mostrar $E[R]$ y $Var[R]$
- Mostrar Z_{obs}
- Mostrar Gráfico

Aleatoriedad: Rachas W-W

Bondad de Ajuste: K-S (F=Fo)

Bondad de Ajuste: K-S (F=G)

Localización de una muestra: W

Dos muestras independientes: U M-W

Dos muestras relacionadas: Apareados W

k muestras independientes: K-W

k muestras relacionadas: Friedman

Independencia: Kendall

Independencia: Spearman

Las pruebas de aleatoriedad se suelen utilizar para estudiar la aleatoriedad de una distribución, tomando los datos de la muestra en el orden dado.

En este apartado se utiliza la prueba de las rachas de Wald-Wolfowitz para estudiar la aleatoriedad de una muestra cualitativa (Moneda) o cuantitativa (Secuencia Numérica).

Hipótesis:

$$\begin{cases} H_0: \text{Muestra Aleatoria} \\ H_1: \text{Muestra No Aleatoria} \end{cases}$$

+ Info: [Test de las Rachas de Wald-Wolfowitz](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variable y el tamaño muestral.

Al clicar en 'Nueva Muestra' se genera una muestra con los valores elegidos.

Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

X X C X X C C C X X X C X X C

Se contabiliza el número de símbolos del primer tipo n_1 , del segundo tipo n_2 y el tamaño total de la muestra n (si la variable es cuantitativa se dicotomiza con la mediana)*:

* Los valores que coinciden con la mediana no se tienen en cuenta en el test.

7.2.3 APRENDE: BONDAD DE AJUSTE: K-S (F=F₀)

Generar Datos:

Tamaño de la muestra:



Nueva Muestra

Mostrar Soluciones:

- Mostrar Datos Ordenados
- Mostrar Distribución Empírica F^*
- Mostrar Distribución Conocida F_0
- Mostrar Diferencias en valor absoluto
- Mostrar D_{obs}
- Mostrar Gráfico Q-Q
- Mostrar Histograma y Densidad

[Aleatoriedad: Rachas W-W](#)

[Bondad de Ajuste: K-S \(F=F₀\)](#)

[Bondad de Ajuste: K-S \(F=G\)](#)

[Localización de una muestra: W](#)

[Dos muestras independientes: U M-W](#)

[Dos muestras relacionadas: Apareados W](#)

[k muestras independientes: K-W](#)

[k muestras relacionadas: Friedman](#)

[Independencia: Kendall](#)

[Independencia: Spearman](#)

Las pruebas de bondad de ajuste sirven para contrastar si un conjunto de observaciones se ajusta a una distribución de probabilidad. En este apartado se utiliza el test de Kolmogorov-Smirnov para el caso de una distribución teórica conocida F_0 , que se compara con la distribución empírica de la muestra F^*

Distribución empírica:

$$F_n^*(x) = \frac{\#\{x_i \leq x\}}{n}$$

Hipótesis:

$$\begin{cases} H_0: F = F_0 \\ H_1: F \neq F_0 \end{cases}$$

+ Info: [Distribución Empírica](#)

+ Info: [Prueba de Kolmogorov-Smirnov](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tamaño muestral.

Al clicar en 'Nueva Muestra' se genera una muestra del tamaño elegido y un nuevo enunciado.

Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

7.2.4 APRENDE: BONDAD DE AJUSTE: K-S (F=G)

Generar Datos:

Tamaño de la muestra X:



Tamaño de la muestra Y:



Nuevas Muestras

Mostrar Soluciones:

- Mostrar Datos Ordenados
- Mostrar Distribución Empírica de $X F^*$
- Mostrar Distribución Empírica de $Y G^*$
- Mostrar Diferencias en valor absoluto
- Mostrar D_{obs}
- Mostrar Gráfico Q-Q

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) **Bondad de Ajuste: K-S (F=G)** Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Las pruebas de bondad de ajuste también sirven para contrastar si dos conjuntos de observaciones provienen de la misma distribución de probabilidad. En este apartado se utiliza el test de Kolmogorov-Smirnov para el caso de dos muestras.

Distribución empírica:

$$F_n^*(x) = \frac{\#\{x_i \leq x\}}{n}$$

Hipótesis:

$$\begin{cases} H_0: F = G \\ H_1: F \neq G \end{cases}$$

+ Info: [Distribución Empírica](#)

+ Info: [Prueba de Kolmogorov-Smirnov](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tamaño de ambas muestras.

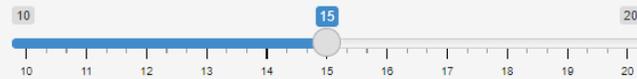
Al clicar en 'Nuevas Muestras' se generan dos muestras del tamaño elegido.

Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

7.2.5 APRENDE: LOCALIZACIÓN DE UNA MUESTRA: W

Generar Datos:

Tamaño de la muestra:



Nueva Muestra

Mostrar Soluciones:

- Mostrar D
- Mostrar $|D|$ y Rangos $|D|$
- Mostrar Estadístico T_{obs}^+
- Mostrar $E[T^+]$ y $Var[T^+]$
- Mostrar Estadístico tipificado Z_{obs}
- Mostrar Box-Plot

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Las pruebas no paramétricas de localización de una muestra consisten en estudiar el valor de la mediana de la población de la que proviene la muestra. En este apartado se aplica la prueba de los rangos con signo de Wilcoxon a muestras cuantitativas.

Hipótesis:

$$\left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: \theta \leq \theta_0 \\ H_1: \theta > \theta_0 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{array} \right.$$

+ Info: [Prueba de los rangos con signo de Wilcoxon](#)

Funcionamiento:

En la sección lateral se puede elegir el tamaño muestral. Al clicar en 'Nueva Muestra' se genera una muestra del tamaño elegido y un nuevo enunciado. Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

X:

-27.15 -25.03 7.23 35.06 17.54 -40.88 -7.14 -43.47 44.37 -17.76 25.13 8.68 7.27 -18.21 -35.64

7.2.6 APRENDE: DOS MUESTRAS INDEPENDIENTES: U M-W

Generar Datos:

Variable:

VARIABLES CUANTITATIVAS

Tamaño de la muestra X_1 :



Tamaño de la muestra X_2 :



Nuevas Muestras

Mostrar Soluciones:

- Mostrar tamaño de las muestras
- Mostrar rangos de X_1 y X_2
- Mostrar R_1 y R_2
- Mostrar U_1 , U_2 y U_{obs}
- Mostrar $E[U]$ y $Var[U]$
- Mostrar Z_{obs}

[Aleatoriedad: Rachas W-W](#) [Bondad de Ajuste: K-S \(F=Fo\)](#) [Bondad de Ajuste: K-S \(F=G\)](#) [Localización de una muestra: W](#)

[Dos muestras independientes: U M-W](#) [Dos muestras relacionadas: Apareados W](#) [k muestras independientes: K-W](#) [k muestras relacionadas: Friedman](#)

[Independencia: Kendall](#) [Independencia: Spearman](#)

Si se quiere comparar la distribución de una variable X entre dos muestras independientes X_1, X_2 , se puede utilizar la prueba para muestras independientes de Mann Whitney

Hipótesis:

$$\begin{cases} H_0: X_1 = X_2 \\ H_1: X_1 \neq X_2 \end{cases} \quad \begin{cases} H_0: X_1 \leq X_2 \\ H_1: X_1 > X_2 \end{cases} \quad \begin{cases} H_0: X_1 \geq X_2 \\ H_1: X_1 < X_2 \end{cases}$$

+ Info: [Prueba para dos muestras independientes de Mann-Whitney](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variable y el tamaño muestral de ambas muestras. Al clicar en 'Nuevas Muestras' se generan dos muestras con los valores elegidos. Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

X_1 :

7.2.7 APRENDE: DOS MUESTRAS RELACIONADAS: APAREADOS W

Generar Datos:

Variable:

Variables Cuantitativas ▾

Tamaño de la muestra:



Nuevas Muestras

Mostrar Soluciones:

- Mostrar Diferencias $Y_i - X_i$
- Mostrar $|Y_i - X_i|$
- Mostrar Rangos
- Mostrar T_{obs}^+
- Mostrar $E[T^+]$ y $Var[T^+]$
- Mostrar Estadístico tipificado Z_{obs}
- Mostrar Box-Plots

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W **Dos muestras relacionadas: Apareados W** k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Si se necesita comparar dos variables que están relacionadas se utiliza la prueba para datos apareados de Wilcoxon. Esta prueba es una alternativa al test t-student para datos relacionados. Es necesario que las variables estén medidas en los mismos individuos y tengan las mismas unidades.

Este tipo de contraste es muy utilizado cuando se compara una variable antes y después de un tratamiento, evento o experimento.

Hipótesis:

$$\left\{ \begin{array}{l} H_0: X = Y \\ H_1: X \neq Y \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X \leq Y \\ H_1: X > Y \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X \geq Y \\ H_1: X < Y \end{array} \right.$$

+ Info: [Datos apareados de Wilcoxon](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variables y el tamaño de las muestras.

Al clicar en 'Nuevas Muestras' se generan dos muestras con los valores elegidos.

Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

X:

7.2.8 APRENDE: K MUESTRAS INDEPENDIENTES: K-W

Generar Datos:

Variable:

Variables Cuantitativas ▾

Tamaño de la muestra X_1 :



Tamaño de la muestra X_2 :



Tamaño de la muestra X_3 :



Nuevas Muestras

Mostrar Soluciones:

- Mostrar tamaño de las muestras
- Mostrar rangos de X_1, X_2, X_3 (R_{ij})
- Mostrar R_1, R_2 , y R_3 .
- Mostrar Estadístico H_{obs}

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Si se quieren comparar $k > 2$ muestras independientes del mismo o distinto tamaño de una variable se utiliza el test de Kruskal-Wallis. Esta prueba es equivalente al ANOVA paramétrico de 1 factor.

Hipótesis:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k \\ H_1: \tau_i \neq \tau_j \text{ para algun } i \neq j \end{cases}$$

$$\begin{cases} H_0: \text{No hay diferencias entre las } k \text{ muestras} \\ H_1: \text{Al menos dos muestras son distintas} \end{cases}$$

+ Info: [Prueba de Kruskal-Wallis](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variables y el tamaño de las muestras. Al clicar en 'Nuevas Muestras' se generan tres muestras con los valores elegidos. Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

7.2.9 APRENDE: K MUESTRAS RELACIONADAS: FRIEDMAN

Generar Datos:

Variable:

Variables Cuantitativas ▾

Tamaño de la muestra:



Nuevas Muestras

Mostrar Soluciones:

- Mostrar tamaño muestral
- Mostrar rangos individuales
- Mostrar $R_1, R_2, y R_3$.
- Mostrar Estadístico S_{obs}
- Mostrar Box-Plots

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Si se quieren comparar $k > 2$ muestras relacionadas se utiliza la prueba de Friedman.

Hipótesis:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k \\ H_1: \tau_i \neq \tau_j \text{ para algun } i \neq j \end{cases}$$

$$\begin{cases} H_0: \text{No hay diferencias entre las } k \text{ muestras} \\ H_1: \text{Al menos dos muestras son distintas} \end{cases}$$

+ Info: [Prueba de Friedman](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variables y el tamaño de las muestras. Al clicar en 'Nuevas Muestras' se generan tres muestras con los valores elegidos. Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones.'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

7.2.10 APRENDE: INDEPENDENCIA: KENDALL

Generar Datos:

Variable:

Variables Cuantitativas

Tamaño de la muestra:



Nuevas Muestras

Mostrar Soluciones:

- Mostrar Tabla de concordancias
- Mostrar Parejas Concordantes y Discordantes
- Mostrar k
- Mostrar τ_{obs}
- Mostrar $E[\tau]$ y $Var[\tau]$
- Mostrar Estadístico tipificado Z_{obs}
- Mostrar Gráfico Dispersión

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Si se quiere estudiar la relación entre dos variables pareadas con unidades distintas sólo podemos compararlas dos a dos. En este apartado se utiliza la prueba de Kendall para contrastar la concordancia de dos muestras.

Hipótesis:

$$\begin{cases} H_0: \text{Variables Independientes} \\ H_1: \text{Variables Dependientes} \end{cases}$$

+ Info: [Coeficiente de Kendall](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variables y el tamaño de las muestras.

Al clicar en 'Nuevas Muestras' se generan dos muestras con los valores elegidos.

Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

X:

7 5.9 5.8 10.1 3 2.7 11.8 3.3 3.4 10.3

7.2.11 APRENDE: INDEPENDENCIA: SPEARMAN

Generar Datos:

Variable:

VARIABLES CUANTITATIVAS

Tamaño de la muestra:



Nuevas Muestras

Mostrar Soluciones:

- Mostrar Rangos (R_i, S_i)
- Mostrar Estadístico r_{obs}
- Mostrar t_{obs} y grados libertad t
- Mostrar Gráfico Dispersión

Aleatoriedad: Rachas W-W Bondad de Ajuste: K-S (F=Fo) Bondad de Ajuste: K-S (F=G) Localización de una muestra: W

Dos muestras independientes: U M-W Dos muestras relacionadas: Apareados W k muestras independientes: K-W k muestras relacionadas: Friedman

Independencia: Kendall Independencia: Spearman

Si se quiere estudiar la relación entre dos variables pareadas con unidades distintas sólo podemos compararlas dos a dos. En este apartado se utiliza la prueba de Spearman para contrastar la asociación de dos muestras. (Las variables no necesitan seguir una distribución normal como en el cálculo de la correlación de Pearson.)

Hipótesis:

$$\begin{cases} H_0: r_{xy} = 0 \\ H_1: r_{xy} \neq 0 \end{cases}$$

+ Info: [Coeficiente de Spearman](#)

Funcionamiento:

En la sección lateral se puede seleccionar el tipo de variables y el tamaño muestral. Al clicar en 'Nuevas Muestras' se generan dos muestras con los valores elegidos. Es posible ocultar los resultados deshabilitando las casillas en 'Mostrar Soluciones:'

Datos:

*Las categorías de las variables ordinales están expresadas como números enteros (Ej: 'Mal'= 1, 'Regular'= 2, 'Bien'= 3)

X:

10.6 10.6 19 15.7 10.9 19.4 19 16.2 16.1 5.8 11.2 13.5 14.5 10.4 15.7

7.2.12 DATOS

¿No tiene datos? Puede utilizar el dataset de ejemplo 'Empleados':

Dataset:

Archivo Importado ▾

Importar Datos:

Subir Archivo:

Examinar...

diamonds.txt

Upload complete

Tipo de Archivo:

TXT ▾

¿Están en la primera fila los nombres de las variables?

Tipo de Separador:*

Coma: ','

Tabulador

Punto y Coma: ';'

Tipo de Decimal:*

Coma: ','

Punto: '.'

* El tipo de separador y el tipo de decimal sólo funcionan con archivos .txt

* Los archivos csv utilizan la coma ',' como separador y el punto '.' para

Archivo importado:

Show 10 ▾ entries

Search:

X	carat	cut	color	clarity	depth	table	price	x	y	z
1	0.23	Ideal	E	SI2	61.5	55	326	3.95	3.98	2.43
2	0.21	Premium	E	SI1	59.8	61	326	3.89	3.84	2.31
3	0.23	Good	E	VS1	56.9	65	327	4.05	4.07	2.31
4	0.29	Premium	I	VS2	62.4	58	334	4.20	4.23	2.63
5	0.31	Good	J	SI2	63.3	58	335	4.34	4.35	2.75
6	0.24	Very Good	J	VVS2	62.8	57	336	3.94	3.96	2.48
7	0.24	Very Good	I	VVS1	62.3	57	336	3.95	3.98	2.47
8	0.26	Very Good	H	SI1	61.9	55	337	4.07	4.11	2.53
9	0.22	Fair	E	VS2	65.1	61	337	3.87	3.78	2.49
10	0.23	Very Good	H	VS1	59.4	61	338	4.00	4.05	2.39

X carat cut color clarity depth table price x y z

Showing 1 to 10 of 14,351 entries

Previous 1 2 3 4 5 ... 1436 Next

Las dimensiones del archivo importado son: 14351 x 11

7.2.13 ANÁLISIS: ALEATORIEDAD: W-W

Variable:

satisfaccion

Umbral:

Mediana

Valor:

1

¿Estadístico y p-valor exacto?*

Realizar Contraste

*Los valores que coincidan con el umbral no serán tenidos en cuenta

*Las variables cualitativas no utilizan un umbral

*Si la variable cualitativa tiene más de dos categorías se enfrentará la categoría del primer elemento al resto de categorías

Aleatoriedad: Prueba de las rachas de Wald-Wolfowitz

Hipótesis:

$$\begin{cases} H_0: \text{Muestra Aleatoria} \\ H_1: \text{Muestra No Aleatoria} \end{cases}$$

Valores elegidos:

Ha elegido la variable: satisfaccion

Ha elegido el tipo de umbral: NA

Valor del umbral: NA

¿Estadístico y p-valor exacto?: FALSE

Resultados:

Número de rachas R_{obs} : 23

Valor del estadístico normalizado Z_{obs} : -1.429

p-valor: 0.153069

Gráfico:



7.2.14 ANÁLISIS: BONDAD DE AJUSTE: K-S ($F=F_0$)

Variable:

salario2008

Distribución Conocida:

Normal

Parámetro 1:

20900

Parámetro 2:

2250

¿Calcular Media, Mediana y Desviación Típica?

¿Estadístico y p-valor exacto?*

*Si hay empates no se podrán obtener los valores exactos

Realizar Contraste

*Parámetro 1: Media/Landa/Grados Libertad

*Parámetro 2: Desviación Estándar

Bondad de Ajuste: Prueba de Kolmogorov-Smirnov $F = F_0$

Hipótesis:

$$\begin{cases} H_0: F = F_0 \\ H_1: F \neq F_0 \end{cases}$$

Valores elegidos:

Ha elegido la variable: salario2008

La distribución elegida es: Normal

Valor del primer parámetro: 20900

Valor del segundo parámetro: 2250

¿Calcular Media, Mediana y Desviación Típica?: TRUE

¿Estadístico y p-valor exacto?: FALSE

Resultados:

Valor del estadístico D_{obs} : 0.076

p-valor: 0.854327

Media	Mediana	Desviación
20949.9192	20743.2633	2256.2537

Gráfico:

7.2.15 ANÁLISIS: BONDAD DE AJUSTE: K-S (F=G)

Variable Agrupación:

sexo ▾

Variable X:

salario2008 ▾

¿Estadístico y p-valor exacto?*

*Si hay empates no se podrán obtener los valores exactos

Realizar Contraste

Bondad de Ajuste: Prueba de Kolmogorov-Smirnov $F = G$

Hipótesis:

$$\begin{cases} H_0: F = G \\ H_1: F \neq G \end{cases}$$

Valores elegidos:

Ha elegido la variable de agrupación: sexo. Variable X: salario2008

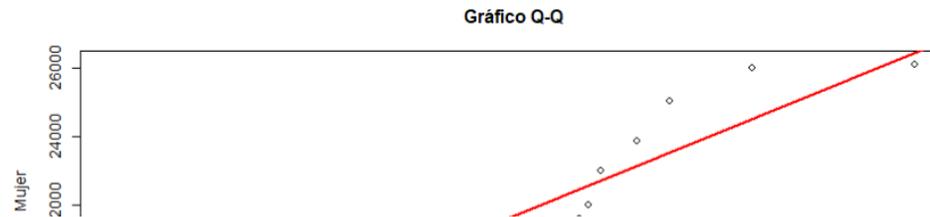
¿Estadístico y p-valor exacto?: FALSE

Resultados:

Valor del estadístico D_{obs} : 0.23

p-valor: 0.387651

Gráfico:



7.2.16 ANÁLISIS: LOCALIZACIÓN MUESTRA: W

Variable:

salario2006 ▾

Tipo de contraste:

Bilateral (H0: =) ▾

θ_0

20300

¿Calcular Media, Mediana y Desviación Típica?

Realizar Contraste

Localización Muestra: Rangos con signo de Wilcoxon

Hipótesis:

$$\left\{ \begin{array}{l} H_0: \theta = \theta_0 \\ H_1: \theta \neq \theta_0 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: \theta \leq \theta_0 \\ H_1: \theta > \theta_0 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: \theta \geq \theta_0 \\ H_1: \theta < \theta_0 \end{array} \right.$$

Valores elegidos:

Ha elegido la variable: salario2006

El tipo de contraste elegido es: Bilateral (H0: =)

El valor elegido de la mediana en las hipótesis θ_0 es: 20300

¿Calcular Media, Mediana y Desviación Típica?: TRUE

Resultados:

Valor del estadístico T_{obs}^+ : 1056

p-valor: 0.917442

Media	Mediana	Desviación
20444.7604	20257.1130	2161.9408

Gráfico:

En rojo aparece el valor de la mediana en las hipótesis (θ_0)

7.2.17 ANÁLISIS: DOS MUESTRAS INDEPENDIENTES: U M-W

Variable Agrupación:

sexo

Variable X:

salario2008

Tipo de contraste:

Bilateral (H0: X1=X2)

¿Estadístico y p-valor exacto?*

*Si hay empates no se podrán obtener los valores exactos

¿Corrección por continuidad?

Realizar Contraste

Dos muestras independientes: U de Mann-Whitney

Hipótesis:

$$\left\{ \begin{array}{l} H_0: X_1 = X_2 \\ H_1: X_1 \neq X_2 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X_1 \leq X_2 \\ H_1: X_1 > X_2 \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X_1 \geq X_2 \\ H_1: X_1 < X_2 \end{array} \right.$$

Valores elegidos:

Ha elegido como variable de agrupación: sexo, y como variable X: salario2008

El tipo de contraste elegido es: Bilateral (H0: X=Y)

¿Estadístico y p-valor exacto?: FALSE

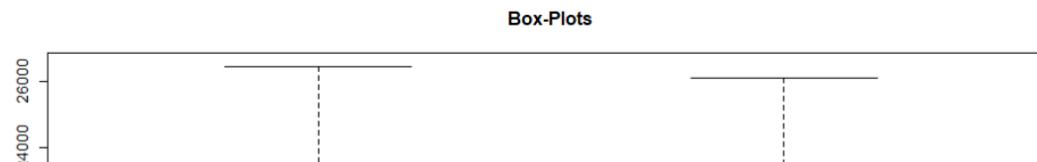
¿Corrección por continuidad?: FALSE

Resultados:

Valor del estadístico U_{obs} : 546

p-valor: 0.360058

Gráfico:



7.2.18 ANÁLISIS: DOS MUESTRAS RELACIONADAS: APAREADOS W

Variable X:

salario2006

Variable Y:

salario2008

Tipo de contraste:

Unilateral superior (H1: X<Y)

¿Estadístico y p-valor exacto?*

*Si hay empates no se podrán obtener los valores exactos

¿Corrección por continuidad?

Realizar Contraste

Dos muestras relacionadas: Datos apareados de Wilcoxon

Hipótesis:

$$\left\{ \begin{array}{l} H_0: X = Y \\ H_1: X \neq Y \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X \leq Y \\ H_1: X > Y \end{array} \right. \quad \left\{ \begin{array}{l} H_0: X \geq Y \\ H_1: X < Y \end{array} \right.$$

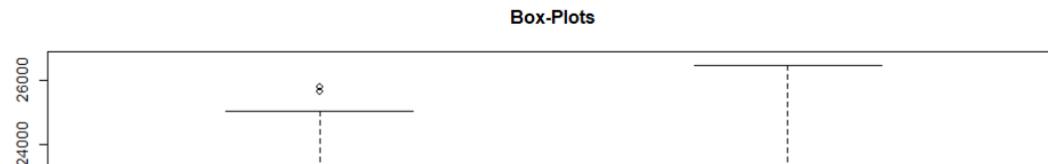
Valores elegidos:

Ha elegido las variables: salario2006 y salario2008
El tipo de contraste elegido es: Unilateral Superior (H1: X<Y)
¿Estadístico y p-valor exacto?: FALSE
¿Corrección por continuidad?: FALSE

Resultados:

Valor del estadístico T_{obs}^+ : 64
p-valor: 0

Gráfico:



7.2.19 ANÁLISIS: K MUESTRAS INDEPENDIENTES: K-W

Variable Agrupación:

zona_trabajo

Variable X:

días_trabajados

¿Incluir Análisis post-hoc?

Realizar Contraste

k muestras independientes: Prueba de Kruskal-Wallis

Hipótesis:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k \\ H_1: \tau_i \neq \tau_j \text{ para algun } i \neq j \end{cases}$$

Valores elegidos:

Ha elegido como variable de agrupación: zona_trabajo, y como variable X: días_trabajados

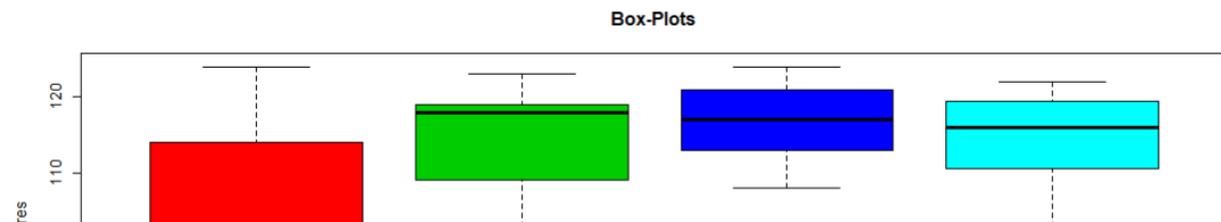
¿Incluir Análisis post-hoc?: FALSE

Resultados:

Valor del estadístico H_{obs} : 12.11

p-valor: 0.007004

Gráfico:



7.2.20 ANÁLISIS: K MUESTRAS RELACIONADAS: FRIEDMAN

Número de Variables:

3 6

Variable 1:

salario2006

Variable 2:

salario2008

Variable 3:

salario2010

¿Incluir Análisis post-hoc?

Realizar Contraste

k muestras relacionadas: Prueba de Friedman

Hipótesis:

$$\begin{cases} H_0: \tau_1 = \tau_2 = \dots = \tau_k \\ H_1: \tau_i \neq \tau_j \text{ para algun } i \neq j \end{cases}$$

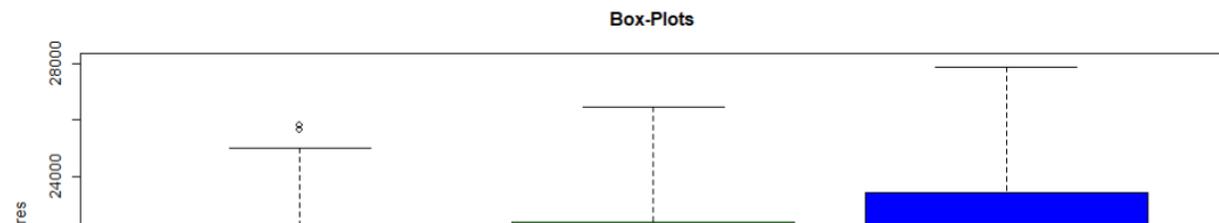
Valores elegidos:

Ha elegido las variables: salario2006, salario2008 y salario2010
¿Incluir Análisis post-hoc?: TRUE

Resultados:

Valor del estadístico S_{obs} : 122.09
p-valor: 0

Gráfico:



7.2.21 ANÁLISIS: INDEPENDENCIA: KENDALL Y SPEARMAN

Variable 1:

antigüedad

Variable 2:

salario2006

Método:

Spearman

¿Estadístico y p-valor exacto?*

*Si hay empates no se podrán obtener los valores exactos

¿Corrección por continuidad?

Realizar Contraste

Independencia: Pruebas de Kendall y Spearman

Hipótesis:

$$\begin{cases} H_0: \text{Variables Independientes} \\ H_1: \text{Variables Dependientes} \end{cases} \quad \begin{cases} H_0: r_{xy} = 0 \\ H_1: r_{xy} \neq 0 \end{cases}$$

Valores elegidos:

Ha elegido las variables: antigüedad y salario2006

Ha elegido el método: spearman

¿Estadístico y p-valor exacto?: FALSE

¿Corrección por continuidad?: FALSE

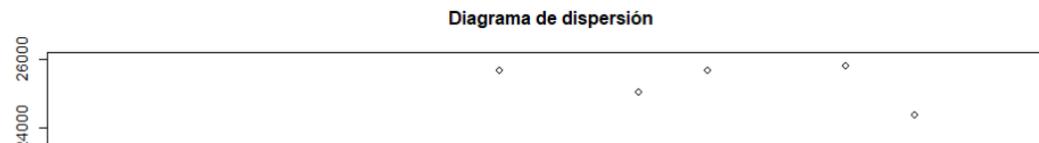
Resultados:

Valor de asociación (Kendall: τ_{obs} , Spearman: r_{obs}): 0.738

Valor del estadístico (Kendall: Z_{obs} , Spearman: t_{obs}): 8.619

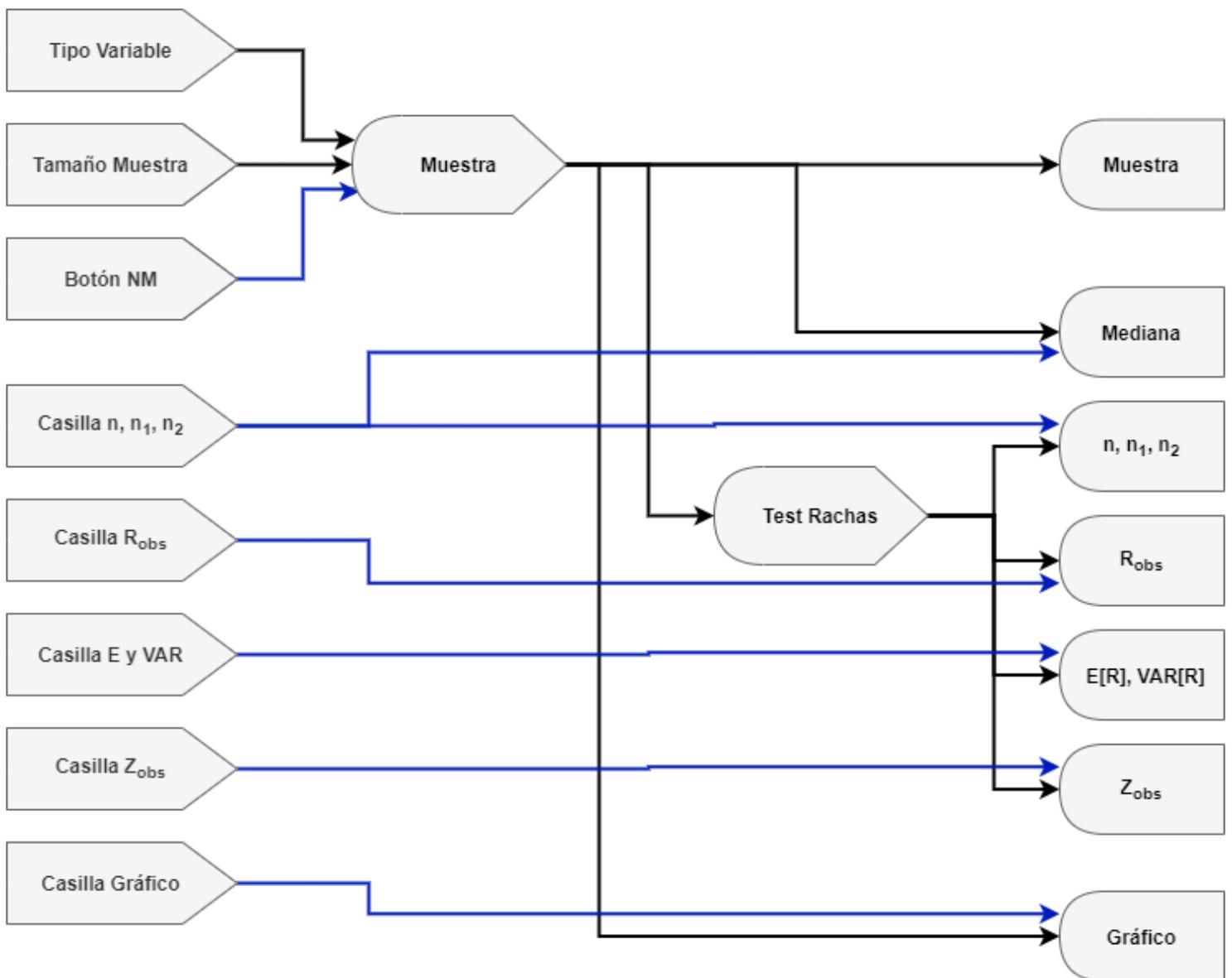
p-valor: 0

Gráfico:

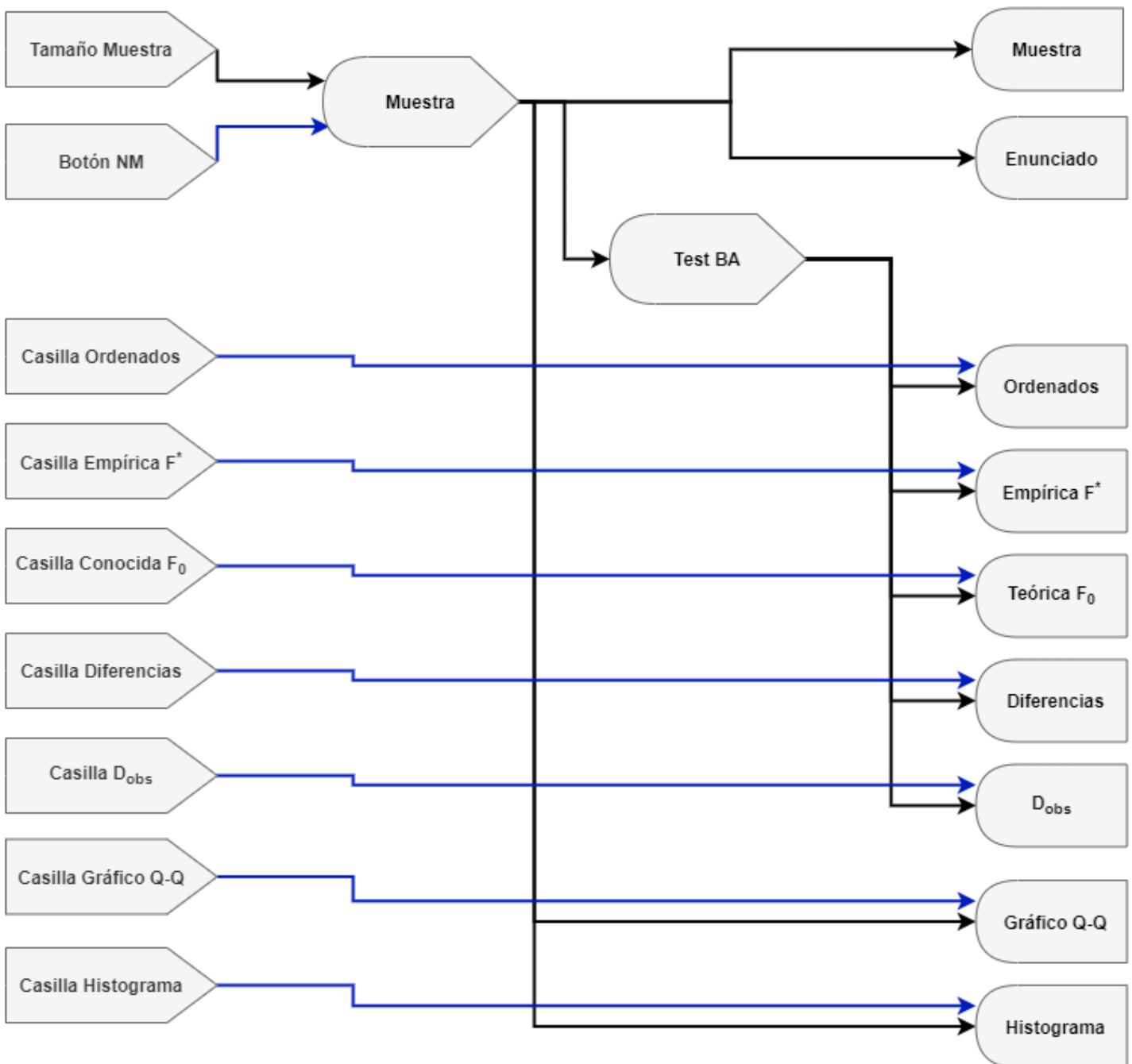


7.3 Esquemas reactivos NPS

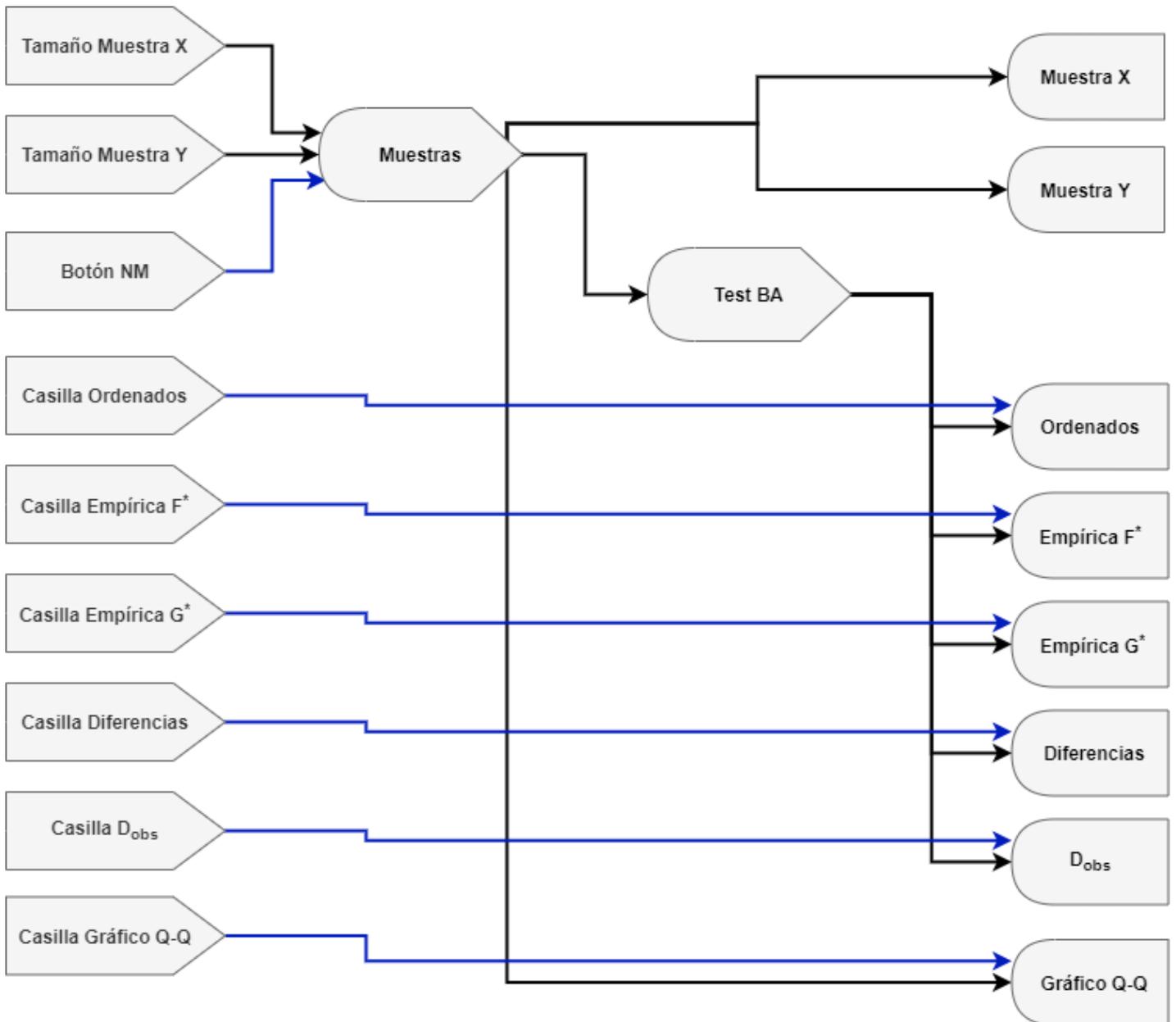
7.3.1 APRENDE: ALEATORIEDAD: RACHAS W-W



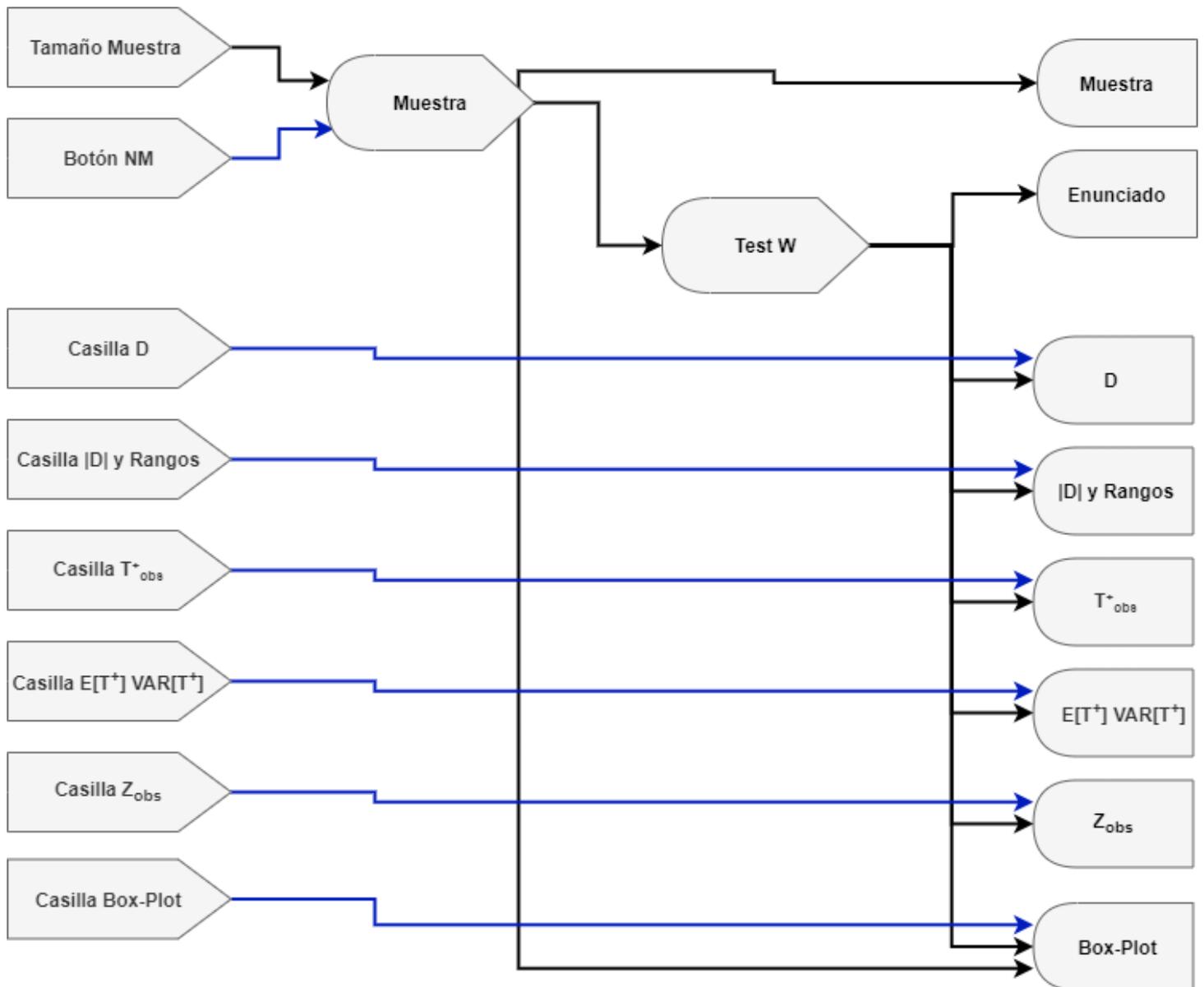
7.3.2 APRENDE: BONDAD DE AJUSTE: K-S ($F=F_0$)



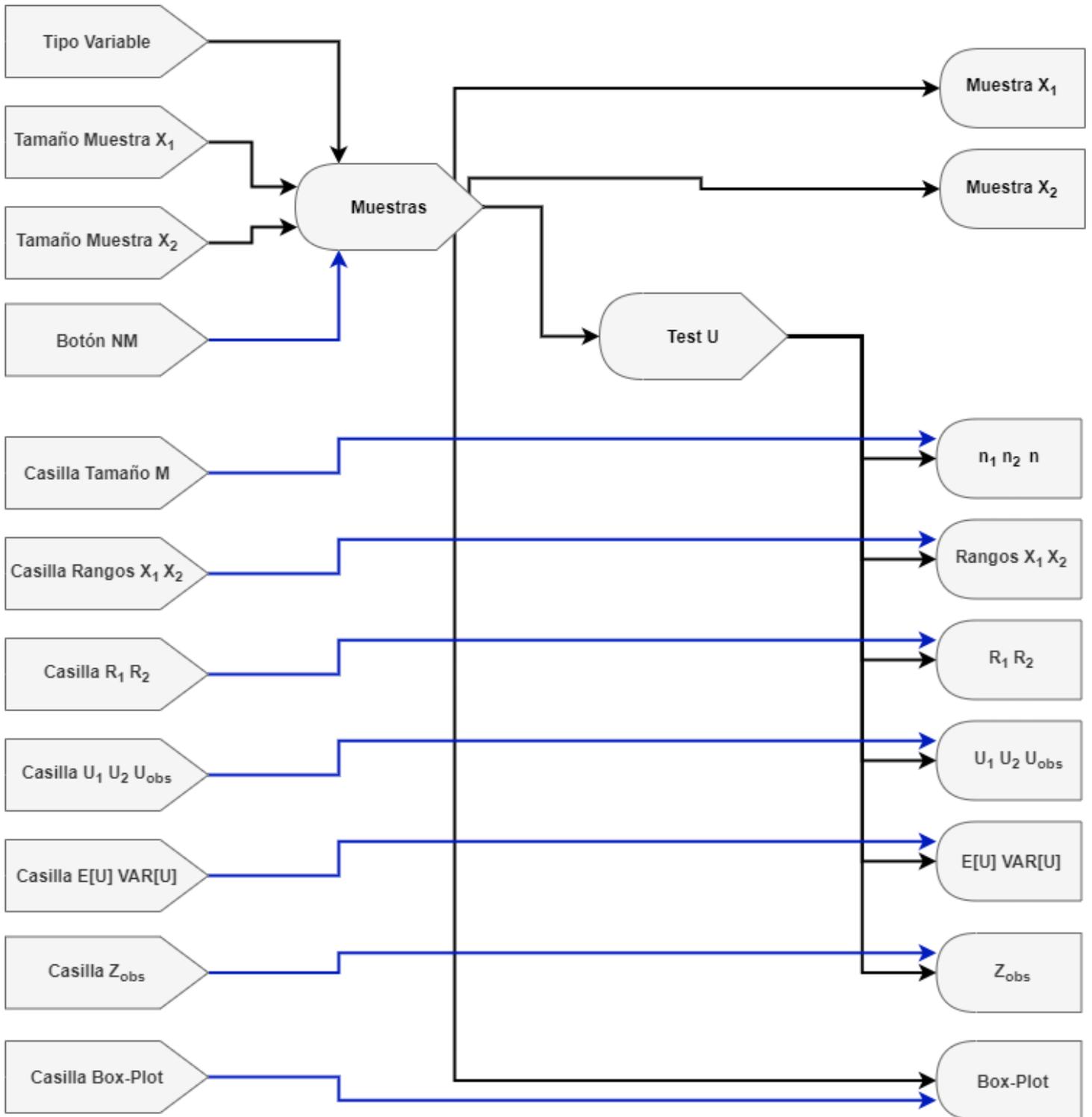
7.3.3 APRENDE: BONDAD DE AJUSTE: K-S (F=G)



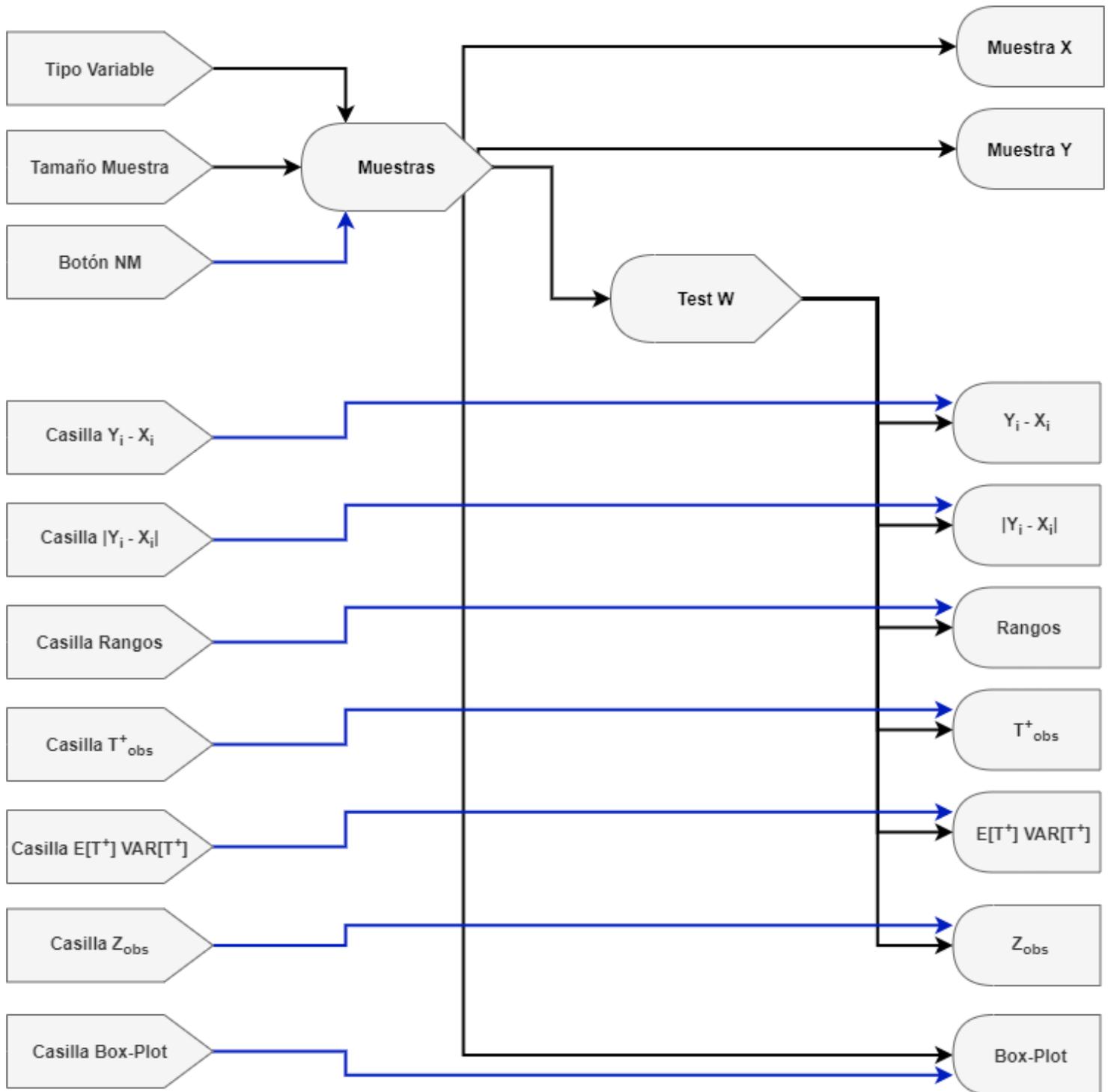
7.3.4 APRENDE: LOCALIZACIÓN DE UNA MUESTRA: W



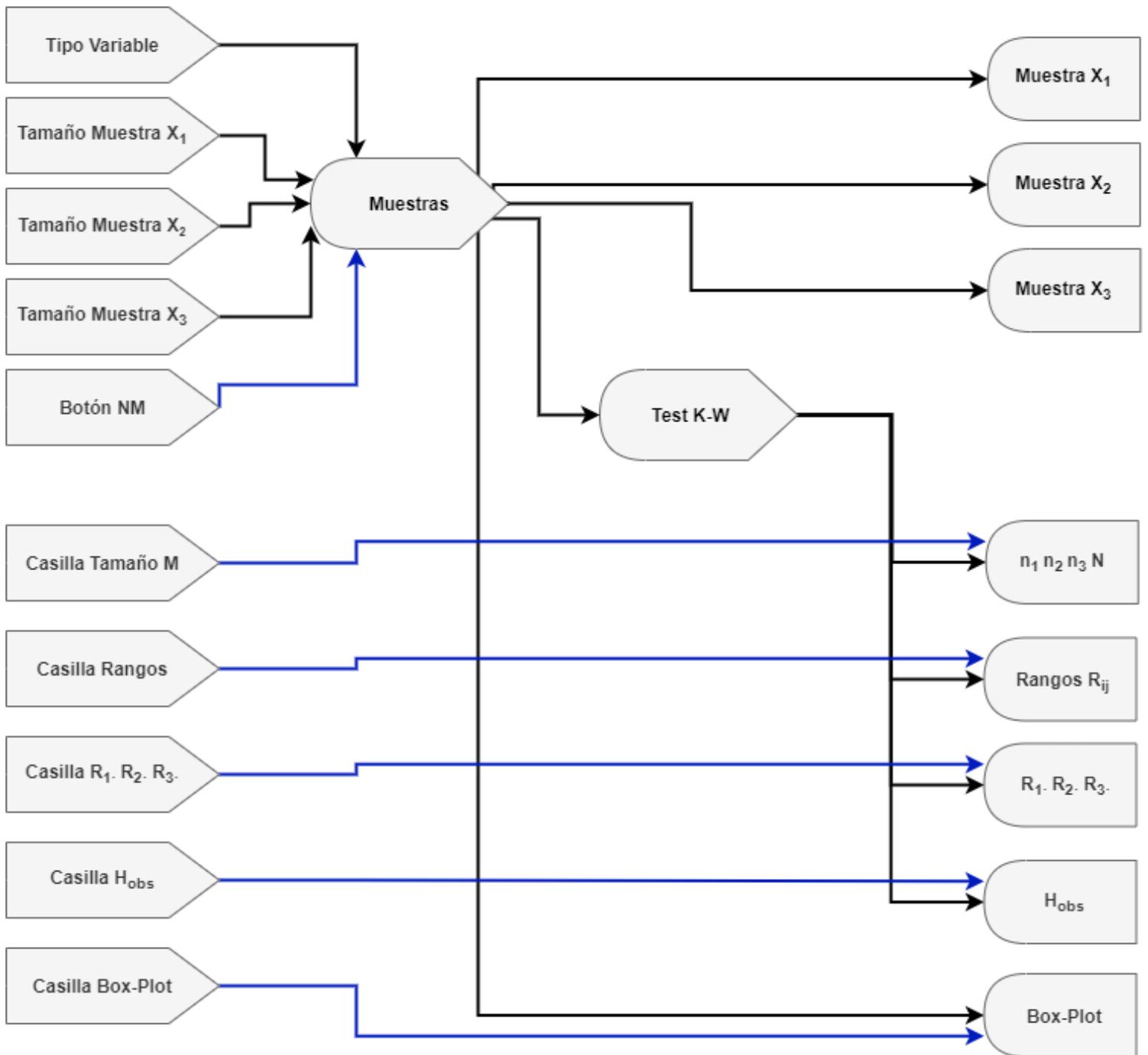
7.3.5 APRENDE: DOS MUESTRAS INDEPENDIENTES: U M-W



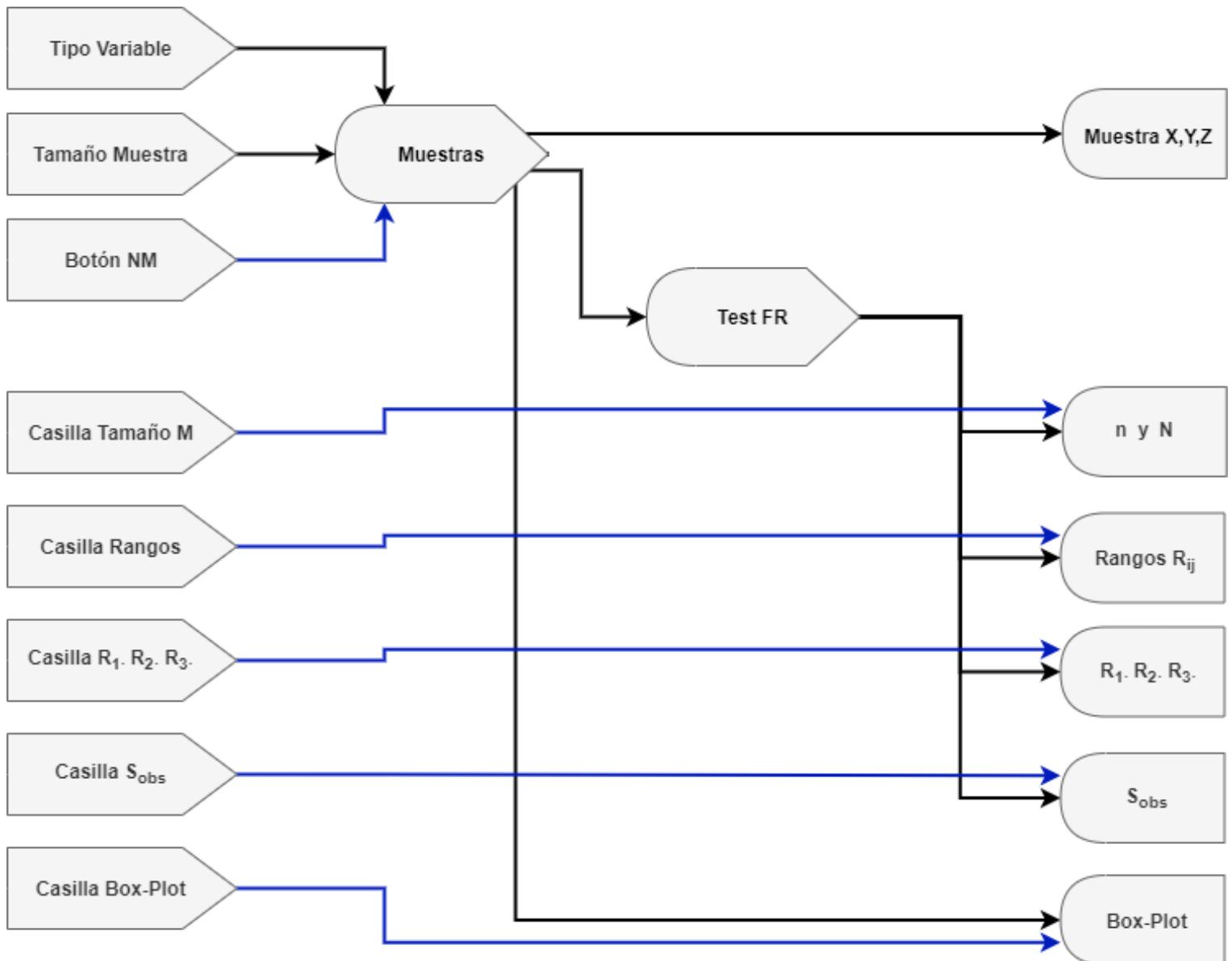
7.3.6 APRENDE: DOS MUESTRAS RELACIONADAS: APAREADOS W



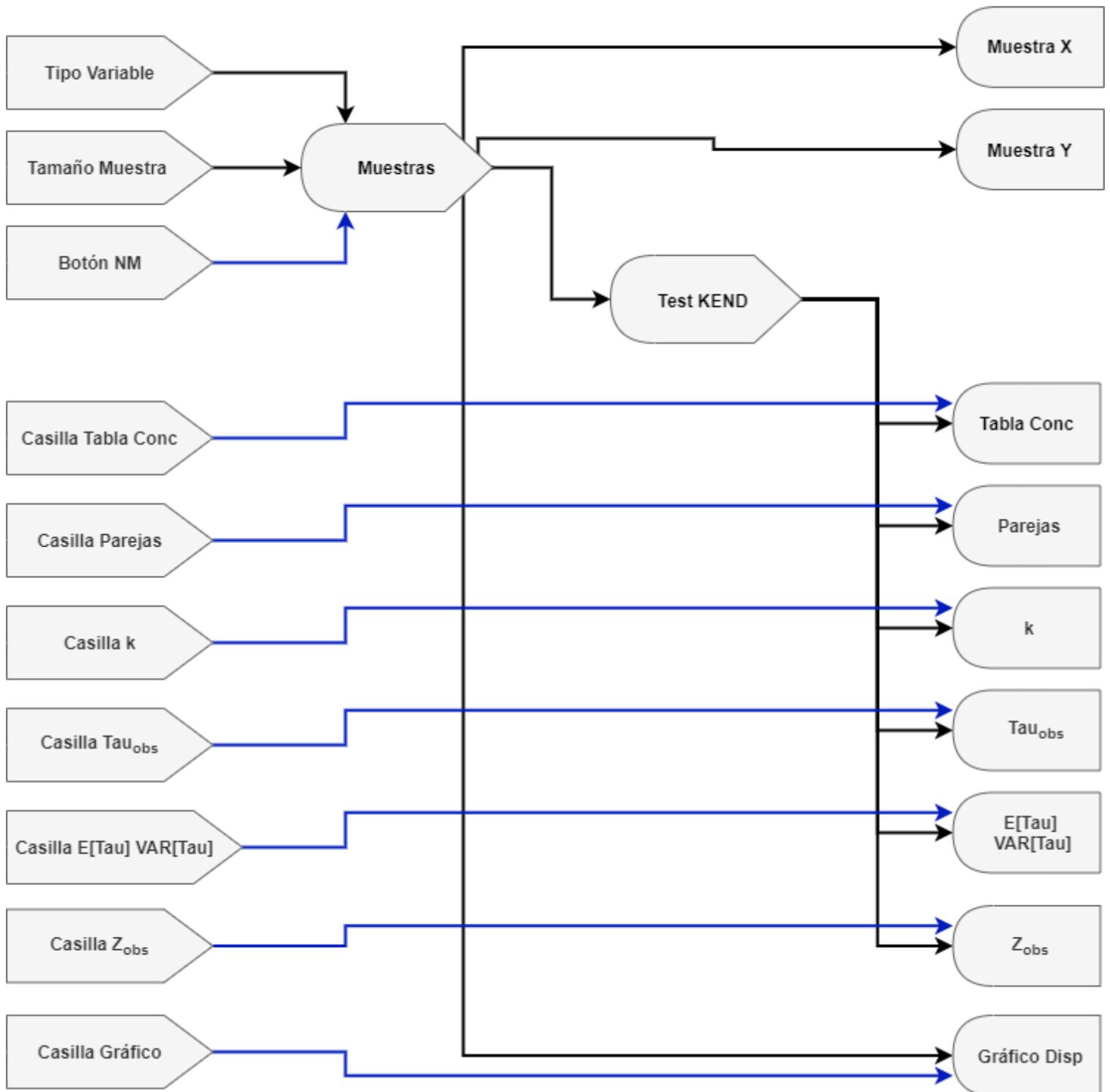
7.3.7 APRENDE: K MUESTRAS INDEPENDIENTES: K-W



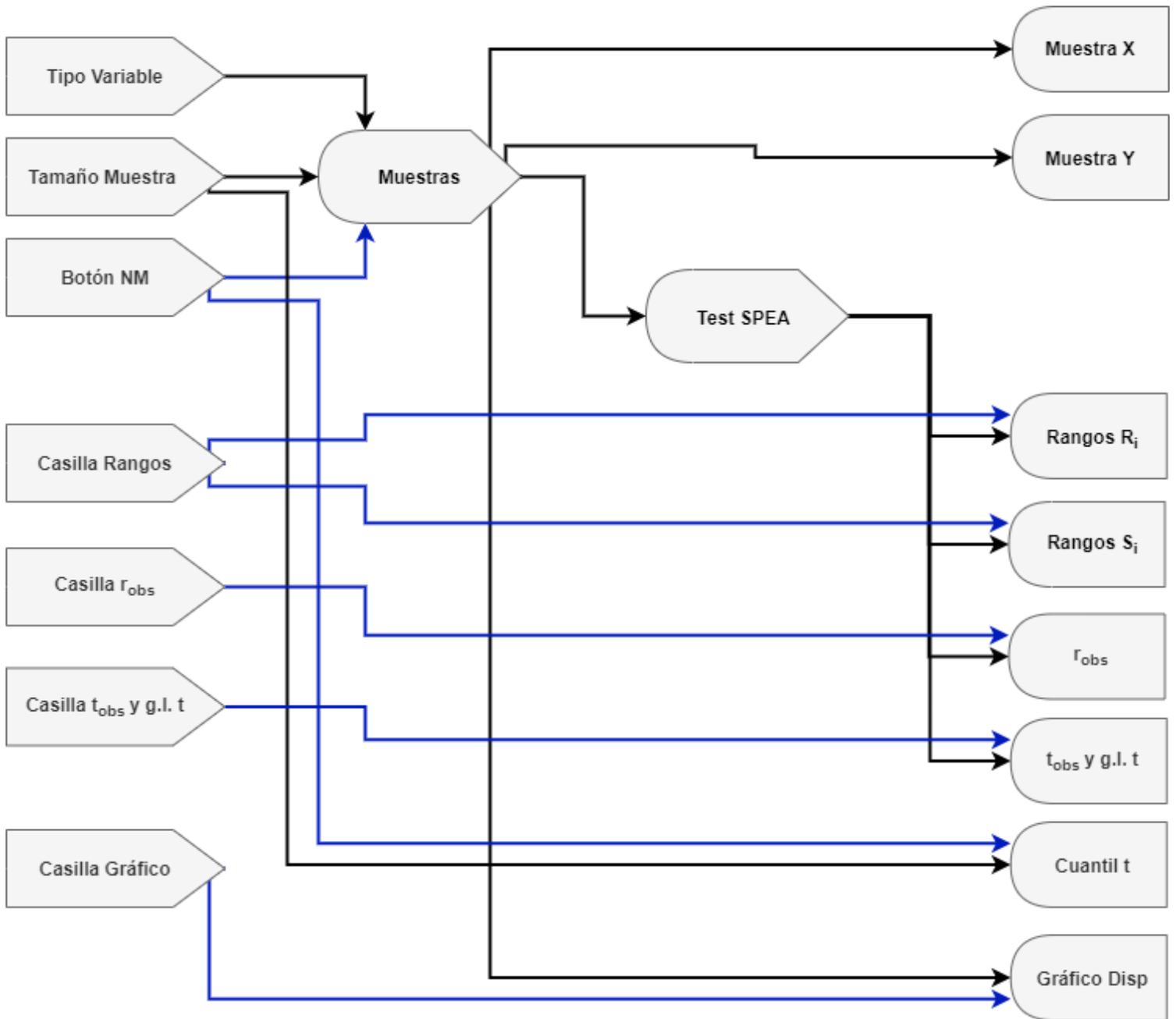
7.3.8 APRENDE: K MUESTRAS RELACIONADAS: FRIEDMAN



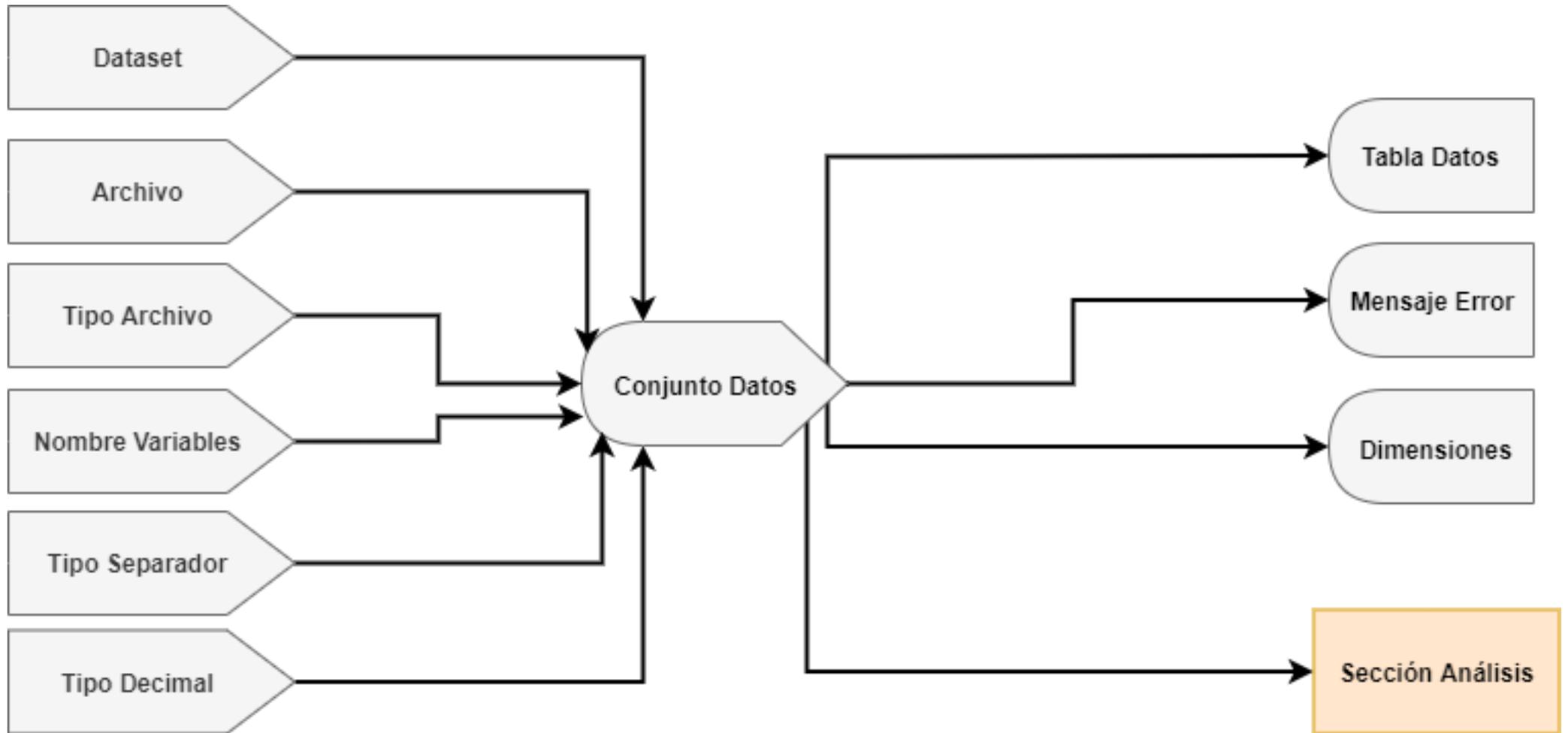
7.3.9 APRENDE: INDEPENDENCIA: KENDALL



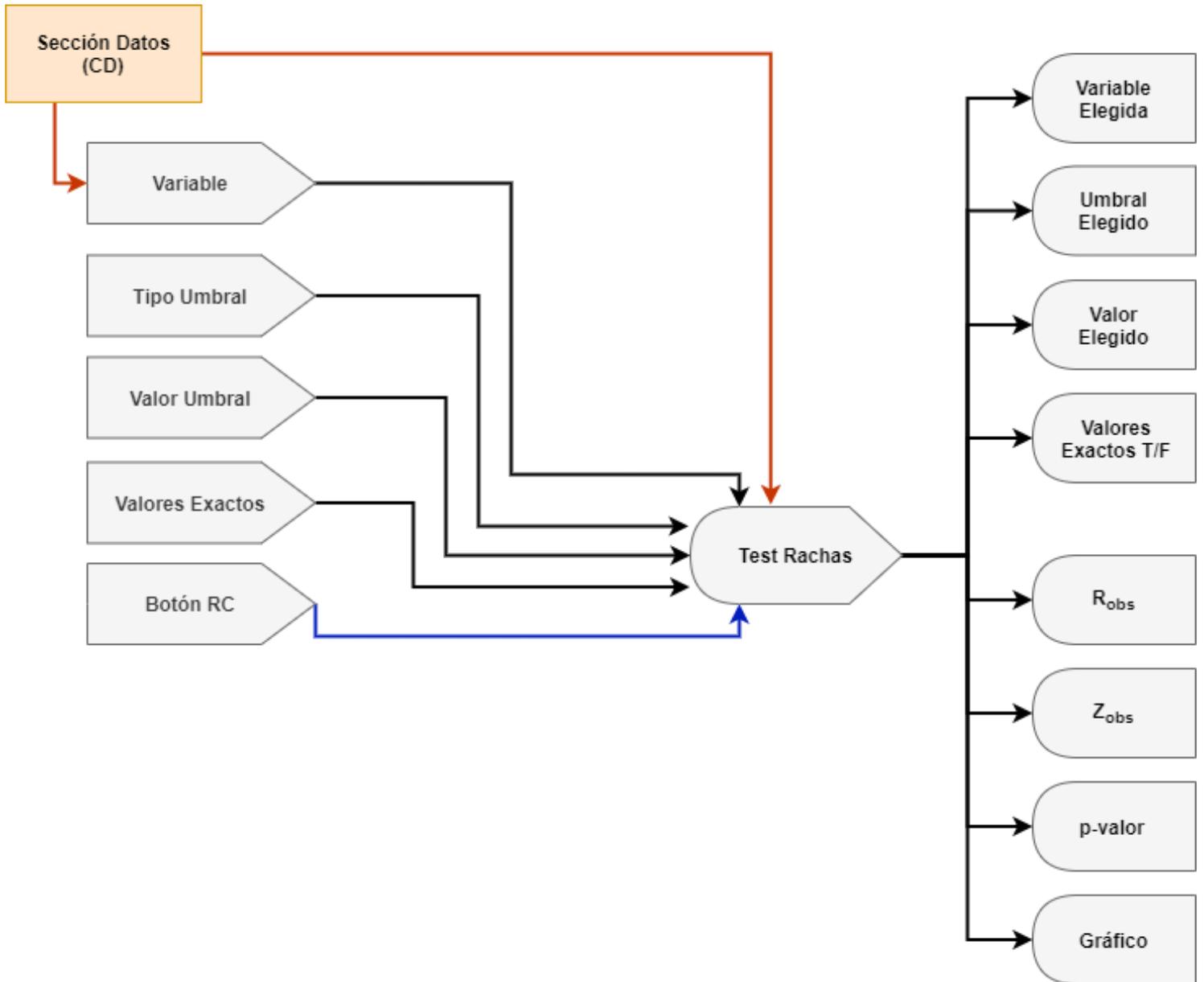
7.3.10 APRENDE: INDEPENDENCIA: SPEARMAN



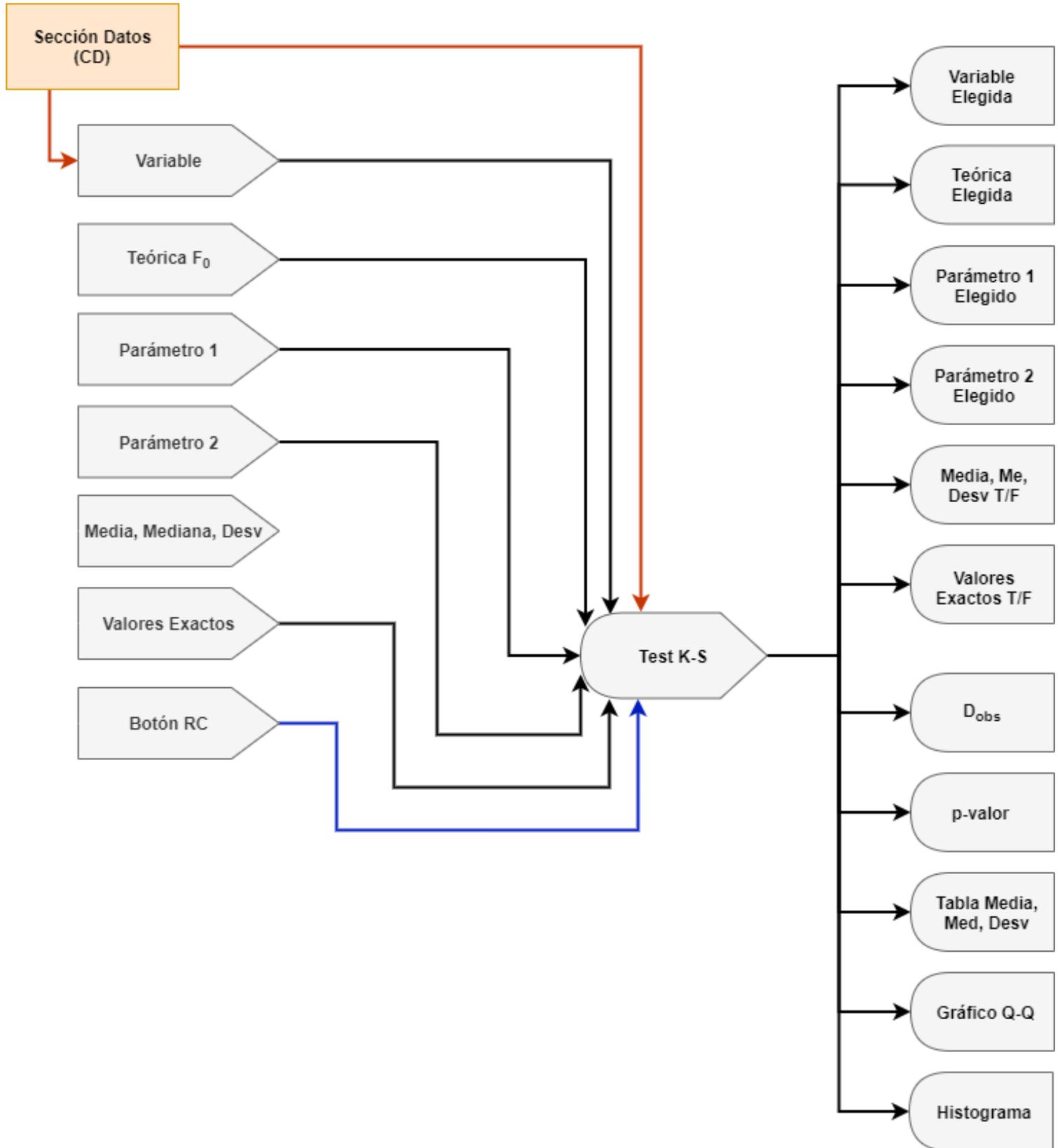
7.3.11 DATOS



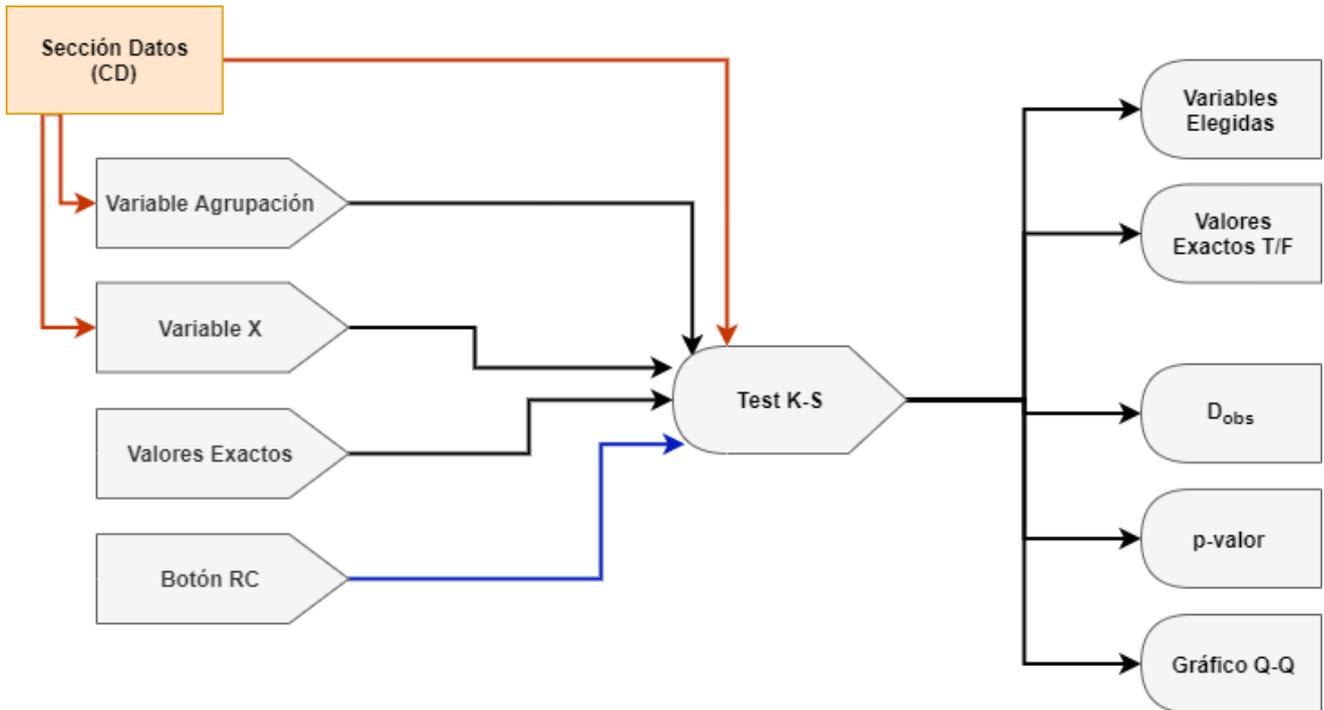
7.3.12 ANÁLISIS: ALEATORIEDAD: W-W



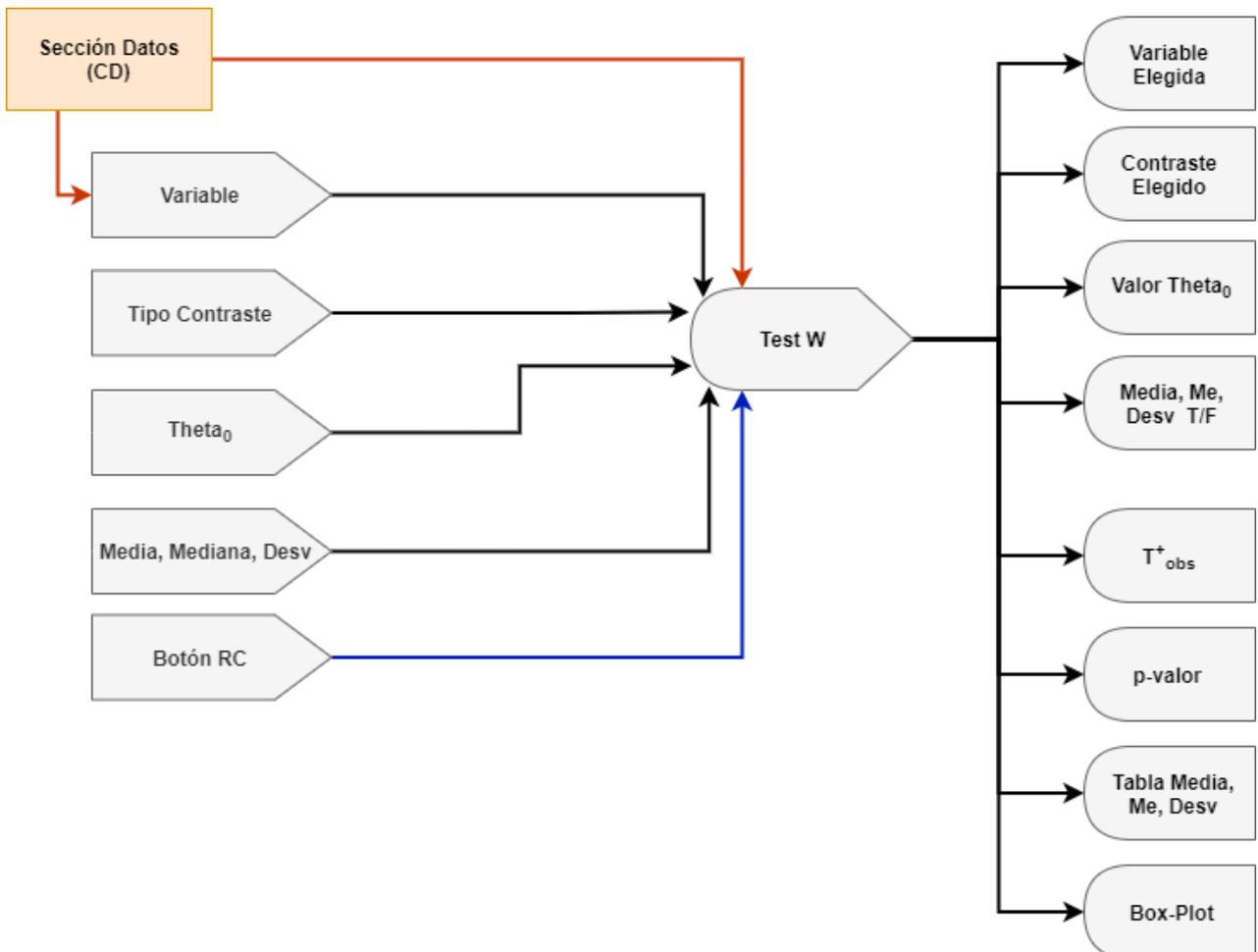
7.3.13 ANÁLISIS: BONDAD DE AJUSTE: K-S ($F=F_0$)



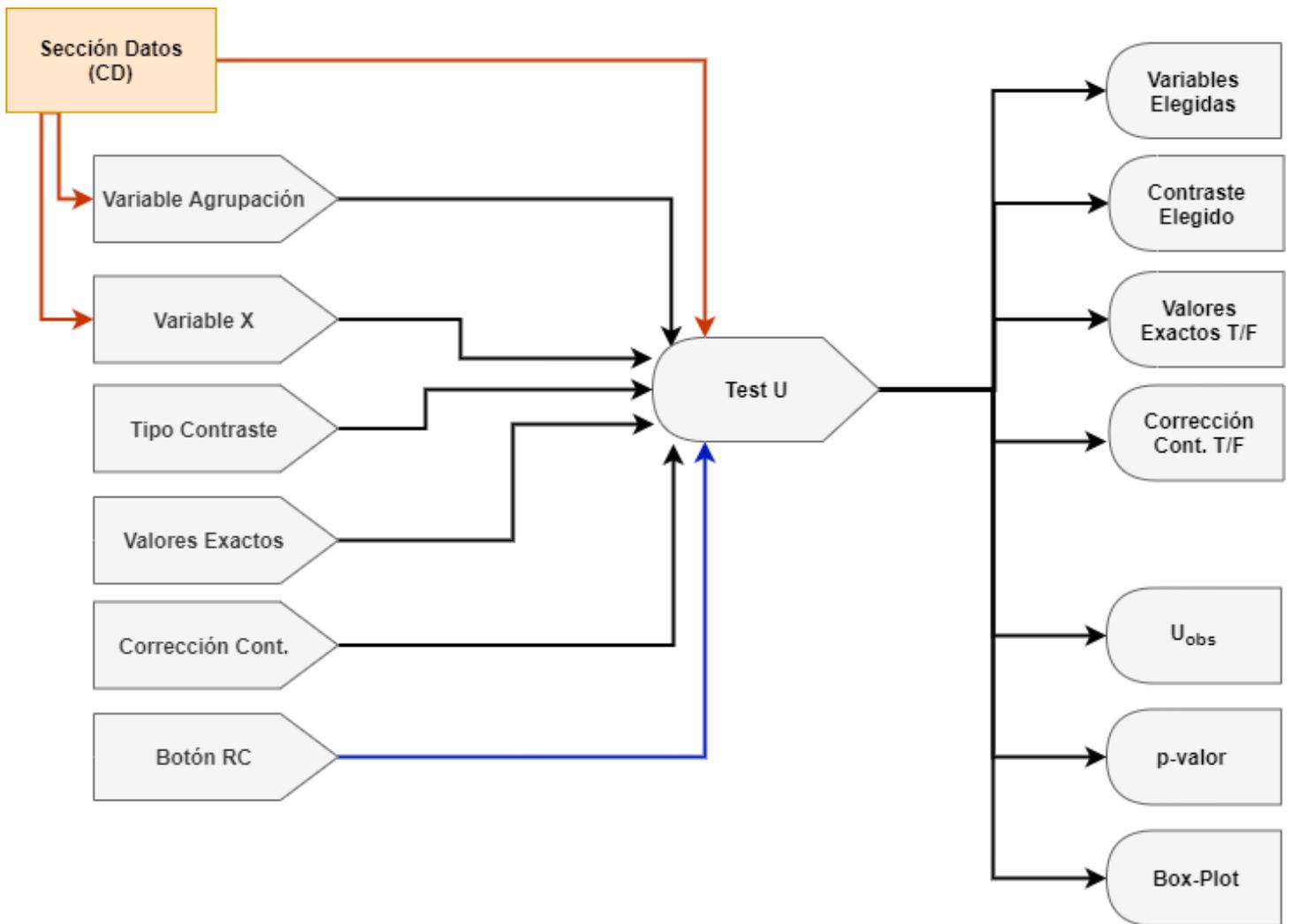
7.3.14 ANÁLISIS: BONDAD DE AJUSTE: K-S (F=G)



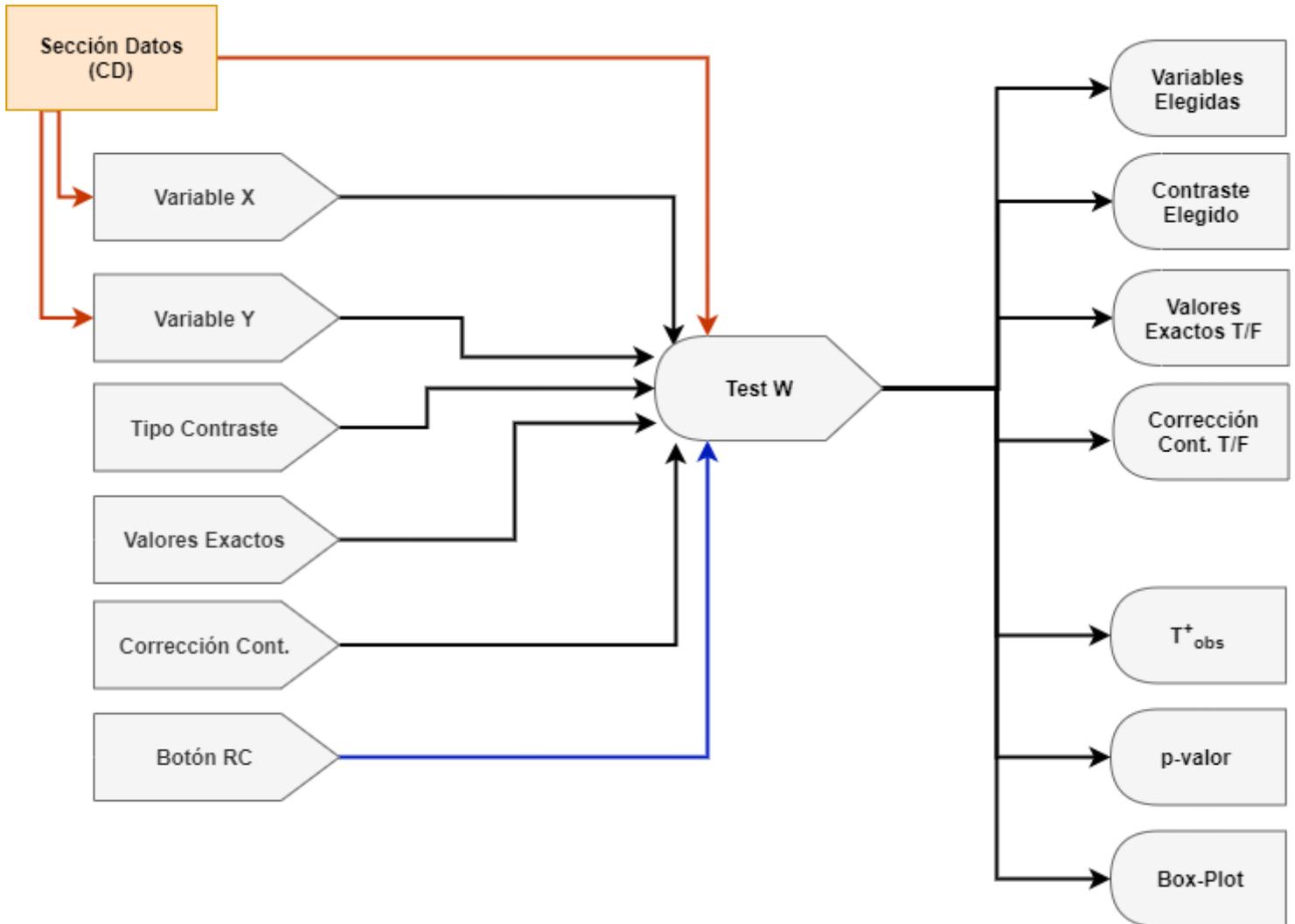
7.3.15 ANÁLISIS: LOCALIZACIÓN MUESTRA: W



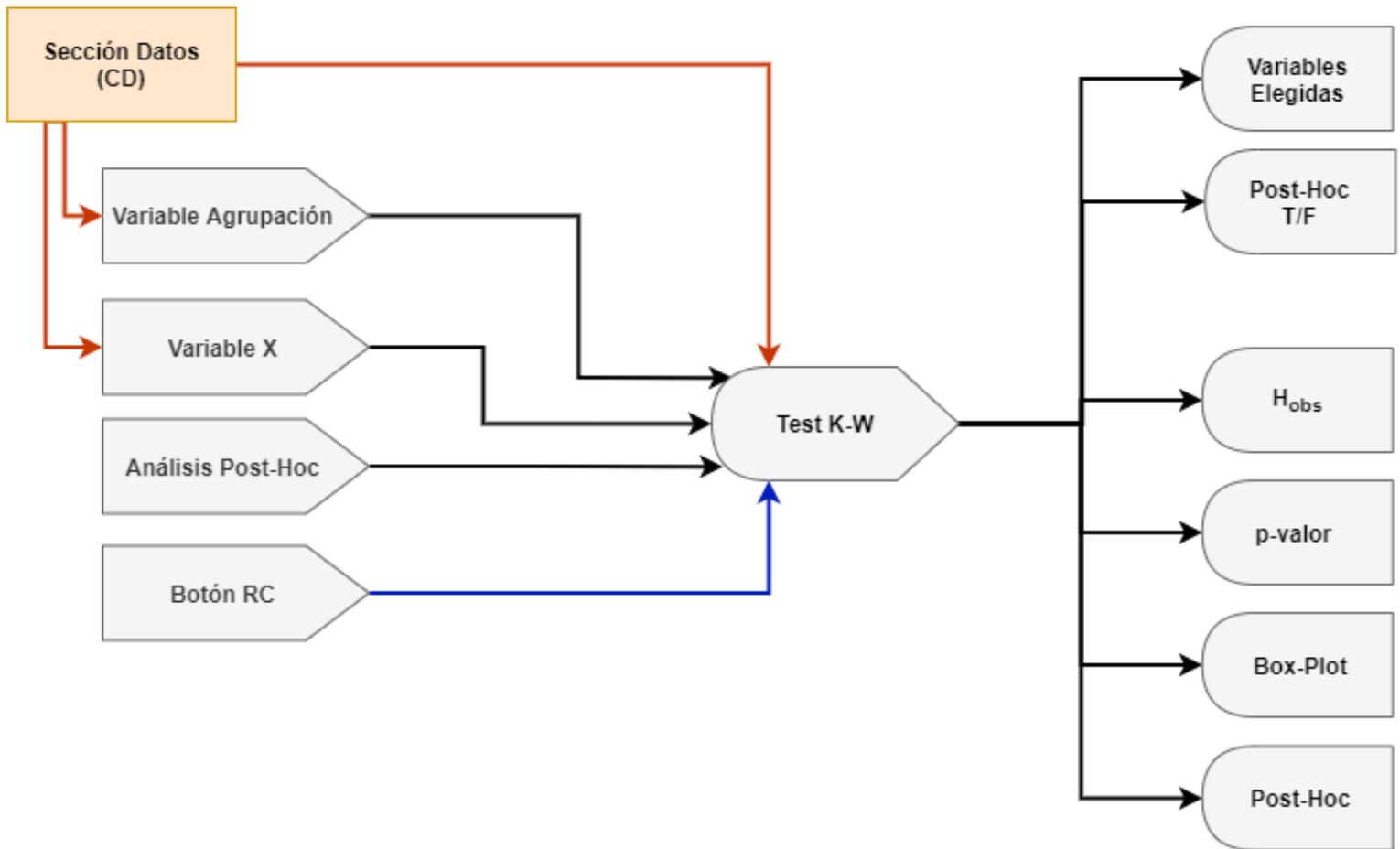
7.3.16 ANÁLISIS: DOS MUESTRAS INDEPENDIENTES: U M-W



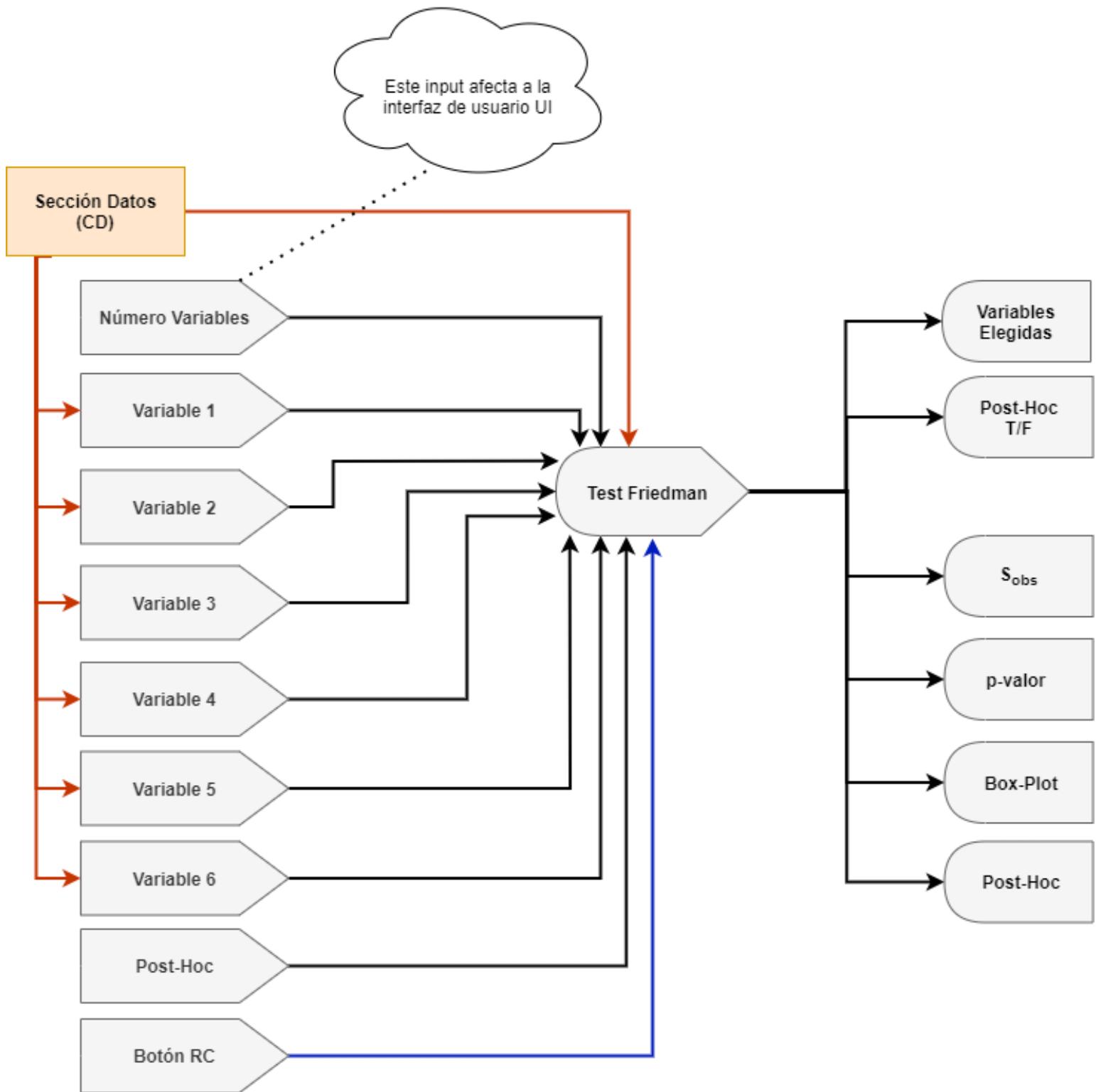
7.3.17 ANÁLISIS: DOS MUESTRAS RELACIONADAS: APAREADOS W



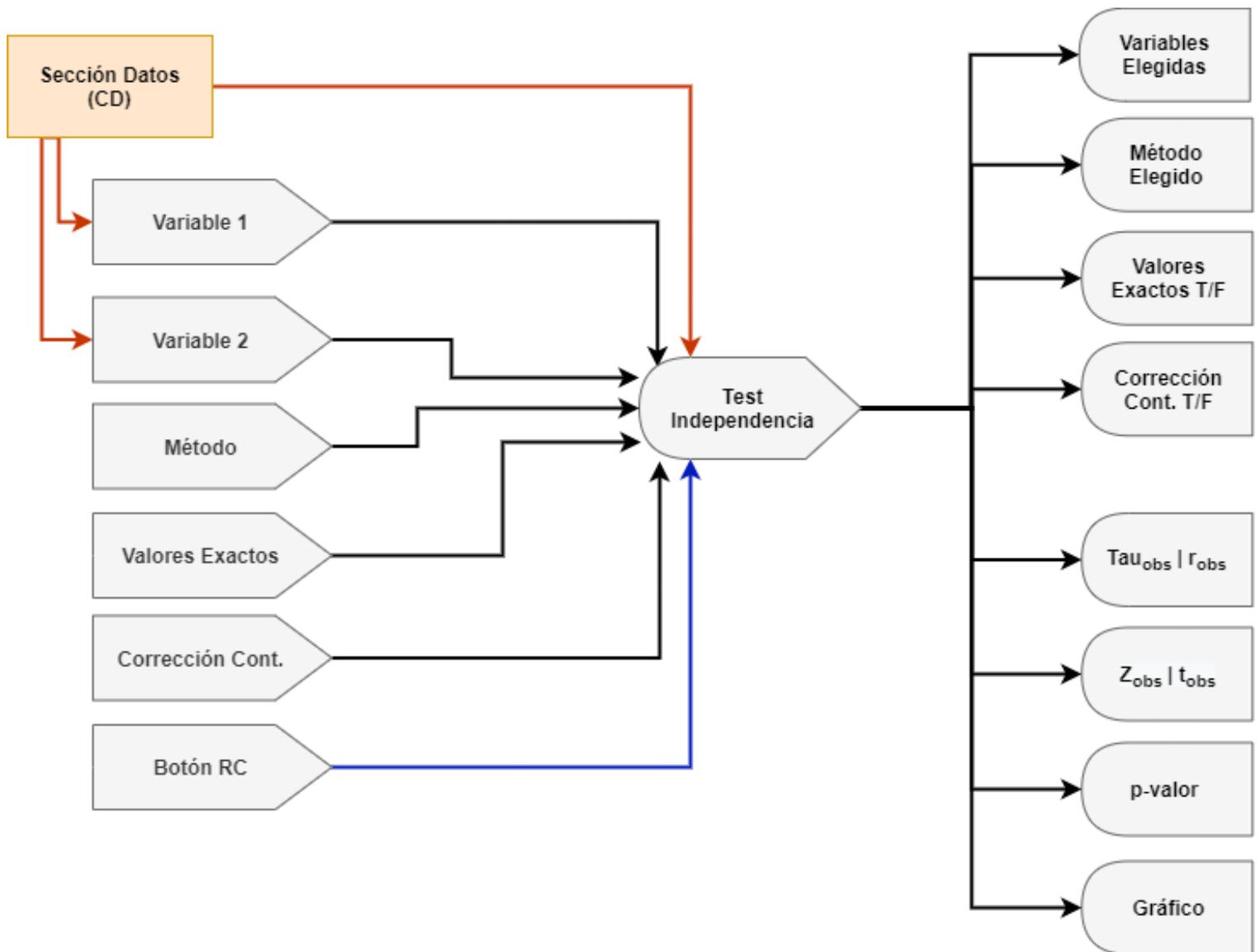
7.3.18 ANÁLISIS: K MUESTRAS INDEPENDIENTES: K-W



7.3.19 ANÁLISIS: K MUESTRAS RELACIONADAS: FRIEDMAN



7.3.20 ANÁLISIS: INDEPENDENCIA: KENDALL Y SPEARMAN



7.4 Código de la aplicación NPS

El código de los *scripts* de R, el archivo *README.txt* y el archivo *empleados2.txt* pueden consultarse en el siguiente enlace o código QR:



<https://github.com/Alfredo318/NPS>