UNIVERSIDAD METROPOLITANA

DEPARTAMENTO DE MATEMÁTICAS

ESTADÍSTICA 2

PROFESORA: JUAN CARLOS TRABUCO FERRO

Informe Proyecto 1:

Pruebas no paramétricas

Briceño, Paula

Camacho, Jesús

Martin, Andrés

Tralci, Vicente

Caracas, 22 de noviembre del 2024

Índice

Índice **2**
Introducción **3**
Objetivo **4**
Descripción de pruebas y algoritmos **4**
Kruskal-Wallis 4
Mann-Whitney 5
Wilcoxon 6
Comparación de Resultados **6**
Kruskal-Wallis 6
Mann-Whitney 9
Wilcoxon 12
Conclusión **15**

Introducción

En estadística, los contrastes de hipótesis no paramétricos se utilizan para comparar muestras y evaluar hipótesis sin asumir una distribución específica de los datos. En este proyecto, nos enfocamos en implementar y analizar tres contrastes de hipótesis no paramétricos ampliamente utilizados: Kruskal-Wallis, Mann-Whitney y Wilcoxon, siguiendo los conceptos y ejemplos presentados en el libro de texto de Wackerly, Mendenhall y Sheaffer.

El objetivo de este proyecto es desarrollar un código en R que realice estos contrastes paso a paso, evitando el uso de funciones automáticas que puedan oscurecer la comprensión del proceso subyacente y facilitar la lectura de archivos de Excel mediante scripts para probar diferentes conjuntos de datos, permitiendo así una aplicación práctica y versátil de los métodos estudiados.

El siguiente informe tiene como propósito demostrar la comprensión de pruebas no paramétricas y de su resolución. Se espera obtener un conjunto de scripts que puedan ser utilizados para realizar los contrastes de Kruskal-Wallis, Mann-Whitney y Wilcoxon en diversos conjuntos de datos. Además, el informe proporcionará una comprensión de estos métodos no paramétricos y sus aplicaciones prácticas, destacando las fortalezas y limitaciones de cada enfoque.

Objetivo

Implementar en R los contrastes de Kruskal-Wallis, Mann-Whitney y Wilcoxon, siguiendo los conceptos y ejemplos del texto de Wackerly, Mendenhall, y Sheaffer.

Descripción de pruebas y algoritmos

Kruskal-Wallis

Esta prueba de hipótesis se utiliza para comparar tres o más grupos independientes. Este test es la extensión del test de Wilcoxon para más de dos grupos y es útil cuando no se puede asumir que los datos siguen una distribución normal. Su aplicación consiste en la implementación de los siguientes pasos:

1. **Formulación de las Hipótesis**:
    - **Hipótesis nula (H_0)**: Las medianas de los grupos son iguales.
    - **Hipótesis alternativa (H_1)**: Al menos una de las medianas de los grupos es diferente.
2. **Ordenar los Datos**:
    - Combina los datos de todos los grupos y ordénalos en orden ascendente.
3. **Asignar Rangos**:
    - Asigna rangos a los datos combinados. Si hay datos iguales, se les asigna el rango promedio de las posiciones que ocupan.
4. **Calcular la Suma de Rangos para Cada Grupo**:
    - Calcula la suma de rangos (Ri) para cada grupo i.
5. **Calcular el Estadístico de Kruskal-Wallis (H)**:
    - Utiliza la siguiente fórmula:

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

donde:

- N es el tamaño total de la muestra.
- k es el número de grupos.
- ni es el tamaño de la muestra del grupo i.
- Ri es la suma de rangos del grupo i.

6. **Determinar el Valor p**:
    - El valor H se compara con la distribución chi-cuadrado con k-1 grados de libertad para determinar el valor p.
7. **Conclusión**:
    - Si el valor p es menor que el nivel de significancia (α), se rechaza la hipótesis nula.
    - Si el valor p es mayor que α, no se puede rechazar la hipótesis nula.

Mann-Whitney

Esta prueba de hipótesis se utiliza para comparar dos muestras independientes. Es una alternativa al test t para muestras independientes cuando los datos no siguen una distribución normal. Se basa en 8 pasos:

1. **Formulación de las Hipótesis**:
   - **Hipótesis nula (H_0)**: Las distribuciones de las dos muestras son iguales.
   - **Hipótesis alternativa (H_1)**: Las distribuciones de las dos muestras no son iguales (una muestra tiende a tener valores más altos o más bajos que la otra).
2. **Combinar las Muestras**:
   - Unir las dos muestras en una única lista.
3. **Asignar Rangos**:
   - Ordenar los datos combinados en orden ascendente y asignar rangos. Si hay empates, se les asigna el rango promedio.
4. **Calcular las Sumas de Rangos para Cada Muestra**:
   - Calcular la suma de rangos (R1 y R2) para cada muestra.
5. **Calcular el Estadístico U**:
   - Se calcula el estadístico U para ambas muestras:

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

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

   - Donde n1 y n2 son los tamaños de las dos muestras.
6. **Seleccionar el Menor de los Dos Valores U**:
   - El valor del estadístico U es el menor de U1 y U2.
7. **Determinar el Valor p**:
   - Usar una tabla de distribución de Mann-Whitney o software estadístico para encontrar el valor p correspondiente al estadístico U calculado.
8. **Conclusión**:
   - Si el valor p es menor que el nivel de significancia (α), se rechaza la hipótesis nula, indicando que las distribuciones de las dos muestras no son iguales.
   - Si el valor p es mayor que α, no se puede rechazar la hipótesis nula.

Wilcoxon

Esta prueba de hipótesis se basa en comparar dos muestras emparejadas o relacionadas. Es una alternativa al test t de muestras dependientes, especialmente cuando los datos no siguen una distribución normal. Se sigue 9 pasos para la realización de la misma que consisten en

1. **Formulación de las hipótesis**:
    - **Hipótesis nula (H_0)**: No hay diferencia en las medianas de las dos muestras emparejadas.
    - **Hipótesis alternativa (H_1)**: Hay una diferencia en las medianas de las dos muestras emparejadas.
2. **Calcular las diferencias**: Restar cada par de datos relacionados (muestra A - muestra B) para obtener las diferencias.
3. **Ignorar las diferencias cero**: Eliminar las diferencias que sean exactamente cero.
4. **Calcular los rangos**: Asignar rangos a las diferencias absolutas ignorando el signo.
5. **Asignar signos a los rangos**: Asignar a cada rango el signo de su diferencia original.
6. **Sumar los rangos positivos y negativos**: Calcular la suma de los rangos positivos (T^+) y la suma de los rangos negativos (T^-).
7. **Seleccionar el menor de los dos rangos**: T=min (T^+,T^−).
8. **Determinar el valor p**: Utilizar tablas de Wilcoxon o software estadístico para encontrar el valor p correspondiente al estadístico de Wilcoxon obtenido con.

$$Z = \frac{T^{+/-} - E(T^{+/-})}{\sqrt{V(T^{+/-})}}$$

9. **Conclusión**:
    - Si el valor p es menor que el nivel de significancia (α), se rechaza la hipótesis nula.
    - Si el valor p es mayor que α, no se puede rechazar la hipótesis nula.

Comparación de Resultados

Kruskal-Wallis

Para la realización de esta prueba, se utilizó un archivo de Excel con datos simulados que representan un estudio sobre el efecto de tres tipos de dietas en el peso final de ratones de laboratorio. Las dietas se clasificaron en tres grupos:

1. Dieta alta en proteínas (grupo A).
2. Dieta alta en carbohidratos (grupo B).
3. Dieta estándar o control (grupo C).

Los datos consisten en los pesos finales de los ratones después de 4 semanas de tratamiento. Se recogieron muestras de tamaño nA=5, nB=5, y nC=5 respectivamente.

Se planteó el siguiente contraste de hipótesis para evaluar si las dietas tienen un efecto significativo sobre el peso final de los ratones:

- *H0 : Mediana (A) = Mediana (B) = Mediana (C) Las dietas de los ratones no afectaron significativamente en el peso obtenido.*
- *H1 : Mediana (i) != Mediana (j) Al menos una de las dietas tuvo un efecto sobre el peso de los ratones*

Los resultados fueron los siguientes:

En primer lugar, al ejecutar el código de la librería `kruskal.test()`, se observó que el estadístico arrojado (H=1.359857) era ligeramente diferente al estadístico obtenido con el código programado manualmente (H=1.355). Esta discrepancia puede atribuirse a diferencias en la implementación de los cálculos de rangos y precisión interna en los algoritmos.

Luego de realizar varias pruebas, se verificó que ambos métodos calculan correctamente los rangos y sumas por grupo. En ambos casos, el p-valor fue obtenido a partir de la distribución chi-cuadrado con los grados de libertad correspondientes. Sin embargo, se observó una leve discrepancia entre los p-valores:

- **Código manual**:
  - Estadístico H: 1.355
  - p-valor calculado: 0.5078851
- **Librería `kruskal.test()`**:
  - Estadístico H: 1.359857
  - p-valor librería: 0.5066533

Aunque los resultados son muy similares, la discrepancia en el p-valor podría deberse a diferencias en la precisión de los cálculos internos de la librería en comparación con la implementación manual.

En conclusión, ambas implementaciones coinciden en el sentido de la prueba: el p-valor es mayor al nivel de significancia (α=0.05), por lo que no se rechaza la hipótesis nula. Esto indica que no hay evidencia estadística suficiente para afirmar que las dietas producen diferencias significativas en el peso final de los ratones.

```
Kruskal-Wallis Test (propio):
$`suma rangos`
   A    B    C 
35.0 35.5 49.5 

$gl
[1] 2

$`Estadísitco H`
[1] 1.355

$p_valor
[1] 0.5078851


Kruskal-Wallis Test (librerías):
$H
Kruskal-Wallis chi-squared 
                  1.359857 

$p_valor
[1] 0.5066533
```

Código

```
#Funcion K-W sin el uso de librerías

kruskal_wallis_test <- function(datos, variable, grupo) {

  valores <- datos[[variable]]

  grupos <- datos[[grupo]]

  datos_unidos <- data.frame(valores, grupos)

  datos_unidos$rangos <- rank(datos_unidos$valores)

  suma_rangos <- tapply(datos_unidos$rangos, datos_unidos$grupos, sum)

 n_por_grupo   <-   tapply(datos_unidos$valores,   datos_unidos$grupos,
length)

  N <- length(valores)

  k <- length(unique(grupos))

  gl <- k - 1

  H <- (12 / (N * (N + 1))) * sum((suma_rangos^2) / n_por_grupo) - 3 *
(N + 1)

  p_valor <- pchisq(H,gl,lower.tail = FALSE)

  return(list("H" = H, "p_valor" = p_valor))

}
```

```r
#Funcion K-W con el uso de librerías

kruskal_wallis_test_con_librerias <- function(datos, variable, grupo){

  resultado  <-  kruskal.test(as.formula(paste(variable,  "~",  grupo)),
data = datos)

  H <- resultado$statistic

  p_valor <- resultado$p.value

  return(list("H" = H, "p_valor" = p_valor))

}
```

Mann-Whitney

Para la realización de esta prueba se utilizó un archivo de excel con datos basados en el libro de Newbold, Carlson y Thorne, los cuales representan el número de horas que dedican a estudiar los estudiantes de Administración y Banca. Para los estudiantes de Administración se utilizó una muestra de tamaño *n = 10* y para los estudiantes de Banca, una de tamaño *n = 12* .

Se planteó el contraste de hipótesis, cuya hipótesis nula es que las posiciones centrales (medianas) de las dos distribuciones poblacionales son idénticas:

H0 : Mediana (1) = Mediana (2) Los estudiantes dedican la misma cantidad de tiempo a estudiar las asignaturas de economía financiera y de contabilidad

para luego, llevar los datos al código programado por el equipo y al código de la librería “*wilcox.test*” .

Los resultados fueron los siguientes:

- En primer lugar, al ejecutar el código de la librería se observó que el estadístico arrojado (*U = 38,5)* era muy distinto al estadístico del ejemplo del libro (*U = 81,5)* : donde se define cuál será la población 1 y en base a eso, se calcula “U”.
- Luego de hacer varias pruebas en el código, se calcularon ambos estadísticos para entender el funcionamiento de la librería. Es decir, uno con el rango de la población 1 y otro con el rango de la población 2. Allí se descubrió que la librería arroja el menor de ellos.
- Sin embargo, al ajustar el código programado para escoger el menor estadístico y calcular el p-valor, se observó una discrepancia con el resultado de la librería:
    - p-valor calculado: 0,15628
    - p-valor librería: 0,15453

```
Mann-Whitney Test (propio):
$Rango_Grupo1
[1] 93.5

$Rango_Grupo2
[1] 159.5

$U1
[1] 81.5

$U2
[1] 38.5

$`Estadístico U`
[1] 38.5

$Z
[1] -1.417668

$p_valor
[1] 0.1562877


Mann-Whitney Test (librerías):
$U
   W
38.5

$p_valor
[1] 0.1545304
```

Código

```
# Función para realizar el contraste de Mann-Whitney sin librerías

mann_whitney_test <- function(datos, variable, grupo) {

  valores <- datos[[variable]]

  grupos <- datos[[grupo]]

  datos_unidos <- data.frame(valores, grupos)

  datos_unidos$rangos <- rank(datos_unidos$valores)

  grupos_unicos <- unique(grupos)

   rangos_grupo1    <-    datos_unidos$rangos[datos_unidos$grupos    ==
grupos_unicos[1]]

  rangos_grupo2    <-    datos_unidos$rangos[datos_unidos$grupos    ==
grupos_unicos[2]]

  R1 <- sum(rangos_grupo1)

  R2 <- sum(rangos_grupo2)
```

```
  n1 <- length(rangos_grupo1)

  n2 <- length(rangos_grupo2)

  U1 <- (n1 * n2) + ((n1 * (n1 + 1)) / 2) - R1

  U2 <- n1 * n2 + (n2 * (n2 + 1)) / 2 - R2

  U <- min(U1, U2)

  if (n1 >= 10 & n2 >= 10) {

    mu_U <- (n1 * n2) / 2

    sigma_U <- sqrt((n1 * n2 * (n1 + n2 + 1)) / 12)

     Z <- (U - mu_U) / sigma_U

    p_valor <- 2 * pnorm(-abs(Z)) # Prueba a dos colas

  } else {

    Z <- NA

    p_valor <- NA

  }

  return(list(

    "Rango_Grupo1" = R1,

    "Rango_Grupo2" = R2,

    "U1" = U1,

    "U2" = U2,

    "Estadístico U" = U,

    "Z" = Z,

    "p_valor" = p_valor

  ))

}

# Función con librerías
```

```r
mann_whitney_test_con_librerias <- function(datos, variable, grupo) {

  resultado <- wilcox.test(as.formula(paste(variable, "~", grupo)), data
= datos, exact = FALSE, correct = FALSE)

  U <- resultado$statistic

  p_valor <- resultado$p.value

  return(list("U" = U, "p_valor" = p_valor))

}
```

Wilcoxon

Esta prueba se realizó con ejemplos de muestras de datos aleatorios generados por Excel, como el siguiente:

> *200 Estudiantes fueron evaluados en aptitud matemática en dos turnos (sea X el turno de la mañana y sea Y el turno de la noche). Las evaluaciones se hacen en una escala entre 100 y 1000 puntos.*
>
> *Se desea saber si existe diferencia entre las evaluaciones en ambos turnos.*

Por un lado, se utilizó el código programado por el equipo, y por el otro, con distintas librerías como "*wilcox.test*" obteniéndose siempre los mismos resultados para los valores de Suma de Rangos positivos T^+, Suma de rangos negativos T^-, Estadístico Wilcoxon T, Valor esperado E(T), Varianza V(T) y estadístico Z.

```
> datos <- leer_excel("datos_prueba_wilcoxon.xlsx")

> wilcoxon_test(datos,"X","Y")
$T
[1] 9608

$T_plus
[1] 9608

$T_minus
[1] 10492

$Z
[1] -0.5393152

$p_valor
[1] 0.5896694
```

```
> cat("Suma de rangos positivos (T+):", T_plus, "\n")
Suma de rangos positivos (T+): 9608
> cat("Suma de rangos negativos (T-):", T_minus, "\n")
Suma de rangos negativos (T-): 10492
> cat("Estadístico de Wilcoxon (T):", T, "\n")
Estadístico de Wilcoxon (T): 9608
> # Determinar el valor p usando una tabla de Wilcoxon o software estadístico
> # Para esto, se puede usar la función wilcox.test para comparar resultados
> resultado <- wilcox.test(A, B, paired = TRUE, exact = FALSE)
> cat("Valor p:", resultado$p.value, "\n")
Valor p: 0.5900894
```

Existe una ligera discrepancia en el valor del p-valor, esto se debe a aproximaciones numéricas en pasos intermedios del cálculo manual, lo cual puede introducir pequeñas diferencias debido a precisión; otra razón por la cual se nota una ligera diferencia en el cálculo del p-valor es debido a que las librerías en R están optimizadas y utilizan métodos estadísticos más avanzados en el cálculo, incluyendo ajustes por valores atípicos y correcciones de continuidad.

Código

```r
#Función sin librerías

wilcoxon_test <- function(datos, variable_x, variable_y) {

 x <- datos[[variable_x]]

 y <- datos[[variable_y]]

 if (length(x) != length(y)) {

   stop("Las muestras X e Y deben tener el mismo tamaño para un test pareado.")

 }

 diferencias <- x - y

 diferencias <- diferencias[diferencias != 0]

 rangos <- rank(abs(diferencias))

 rangos_signed <- rangos * sign(diferencias)

 T_plus <- sum(rangos_signed[rangos_signed > 0])

 T_minus <- -sum(rangos_signed[rangos_signed < 0])

 T <- min(abs(T_plus), abs(T_minus))

 n <- length(diferencias)

 mu_T <- n * (n + 1) / 4

 sigma_T <- sqrt(n * (n + 1) * (2 * n + 1) / 24)

 Z <- (T - mu_T) / sigma_T

 p_valor <- 2 * pnorm(-abs(Z))

 return(list(

   "T" = T,

   "T_plus" = T_plus,
```

```
    "T_minus" = T_minus,

    "Z" = Z,

    "p_valor" = p_valor

  ))

}
#Función con librerias

wilcox_lib <- function(datos) {

  datos$X <- as.numeric(datos$X)

  datos$Y <- as.numeric(datos$Y)

  if (any(is.na(datos$X)) || any(is.na(datos$Y))) {

    stop("Las columnas contienen valores no numéricos o valores faltantes.")

  }

  resultado <- wilcox.test(x = datos$X, y = datos$Y, paired = TRUE)

  print(resultado)

}
```

Conclusión

En este proyecto se realizó una comparación exhaustiva de los resultados obtenidos mediante pruebas estadísticas Wilcoxon, Mann-Whitney y Kruskal-Wallis, implementadas tanto con el uso de librerías en R como mediante un código desarrollado sin dichas librerías. Los resultados obtenidos de ambas metodologías mostraron una concordancia significativa en los valores críticos de las estadísticas de prueba, validando la precisión y la solidez del enfoque manual al replicar estos tests bien establecidos.

Sin embargo, se observaron ligeras discrepancias en los p-valores obtenidos entre ambas metodologías. Estas diferencias pueden atribuirse a la precisión y optimización intrínsecas de las funciones integradas en las librerías de R, que están diseñadas para manejar grandes conjuntos de datos y cálculos complejos de manera más eficiente.

Es importante destacar que las librerías de R, no especifican con claridad en qué estudios o publicaciones se fundamenta el cálculo de los estadísticos ni el procedimiento para la obtención del p-valor. Lo cual, resulta preocupante cuando se compara con autores del libro, quienes emplean metodologías bien documentadas y ampliamente aceptadas. Tampoco se encuentra información explícita sobre la fuente de la tabla Z utilizada por la librería para calcular los p-valores.

A pesar de estas diferencias mínimas en los p-valores, los resultados de ambas metodologías llevaron a las mismas conclusiones estadísticas en términos de aceptación o rechazo de las hipótesis nulas respectivas.