Diseño de experimentos I Ciclo 2025
Ci0131
Laboratorio 3

# Experimentos comparativos simples

# Prueba T para comparar dos medias independientes

Basado en el ejercicio de Joaquín Amat Rodrigo

https://www.cienciadedatos.net/documentos/12_t-test

**Introducción**

Una de las opciones cuando se quiere comparar una variable continua entre dos grupos consiste en comparar los resultados promedio obtenidos para cada uno. El hecho de que los valores promedio de cada grupo no sean iguales no implica que haya evidencias de una diferencia significativa. Dado que cada grupo tiene su propia variabilidad, aunque el tratamiento no sea eficaz, las medias muestrales no tienen por qué ser exactas.

Con el fin de estudiar si la diferencia observada entre las medias de dos grupos es significativa, se puede recurrir a métodos paramétricos como el basado en Z-scores o en la distribución T de Student. En ambos casos se pueden calcular tanto intervalos de confianza para saber entre qué valores se encuentra la diferencia real de las medias poblacionales, como test de hipótesis para determinar si la diferencia es significativa.

La distribución T de Student se asemeja en gran medida a la distribución normal. Tiene como parámetros la media, la varianza y además incorpora a través de los grados de libertad una modificación que permite flexibilizar las colas en función del tamaño que tenga la muestra. A medida que se reduce el tamaño muestral, la probabilidad acumulada en las colas aumenta, siendo así menos estricta de lo cabría esperar en una distribución normal. Una distribución T de Student con 30 o más grados de libertad es prácticamente igual a una distribución normal.

El número de grados de libertad de la distribución T de Student se calcula:

• Para estudiar una sola muestra: $df = tamaño\ muestra - 1$
• Cuando se comparan dos muestras: existen varios métodos, uno de los utilizados es emplear los grados de libertad de la muestra de menor tamaño $mínimo\ ((n1 - 1), (n2 - 1))$.

En los programas informáticos se emplean métodos para ajustar de forma más precisa los grados de libertad.

**Condiciones de una prueba t de Student para muestras independientes**

Las condiciones para calcular intervalos de confianza o aplicar una prueba de hipótesis basados en la distribución T de Student son:

• Independencia: Las observaciones tienen que ser independientes unas de las otras. Para ello el muestreo debe ser aleatorio y el tamaño de la muestra inferior al 10% de la población.

• Normalidad: Las poblaciones que se comparan tienen que distribuirse de forma normal. A pesar de que la condición de normalidad recae sobre las poblaciones, normalmente no se dispone de información sobre ellas por lo que las muestras (dado que son reflejo de la población) tiene que distribuirse de forma aproximadamente normal.

En caso de cierta asimetría las pruebas T de Student muchos autores consideran que las pruebas siguen siendo razonablemente robustas cuando el tamaño de las muestras es mayor o igual a 30.

En general el hecho de contar con una muestra grande se podría usar si estuviéramos realizando una prueba $Z_0$ para la que necesitaríamos conocer también las varianzas poblacionales (que no es este caso). Aun en el caso de una prueba $Z_0$ la no normalidad debe considerarse con mucha precaución, especialmente si la distribución de los datos subyacentes es muy asimétrica o presenta colas importantes. Además, este hecho es necesario mencionarlo en las conclusiones.

• Igualdad de varianza (homocedasticidad): la varianza de ambas poblaciones comparadas debe de ser igual o muy similar. Tal como ocurre con la condición de normalidad, si no se dispone de información de las poblaciones, esta condición se ha de asumir a partir de las muestras. En caso de no cumplirse esta condición se puede emplear una prueba T de Welch. Está corrección se incorpora a través de los grados de libertad permitiendo compensar la diferencia de varianzas, pero con el inconveniente de que pierde un poco de precisión.

**Prueba de hipótesis**

Los pasos a seguir para realizar una prueba para comparar dos medias son:

1. Establecer las hipótesis.
2. Calcular el estadístico (parámetro estimado) que se va a emplear.
3. Validar los supuestos para aplicar la prueba (puede ser T de Student, T de Welch, u otra prueba no paramétrica).
4. Determinar el tipo de prueba, una o dos colas.
5. Determinar el nivel de significancia $\alpha$.
6. Calcular el p-value y comparar con el nivel de significancia establecido.
7. Calcular el tamaño del efecto (opcional pero recomendado).
8. Calcular la potencia de la prueba (opcional pero recomendado).
9. Analizar los resultados y emitir las conclusiones.

**1. Establecer las hipótesis**

Hipótesis nula ($H0$): es la hipótesis escéptica que considera que no hay diferencia o cambio. Suele contener en su definición el símbolo =. En el caso de comparar dos medias independientes la hipótesis nula considera que $\mu1 = \mu2$.

Hipótesis alternativa ($HA$): considera que el valor real de la media poblacional es mayor, menor o distinto del valor que establece la Ho. Suele contener los símbolos >, <, ≠. En el caso de comparar dos medias independientes la hipótesis alternativa considera que $\mu1 \neq \mu2$.

**2. Calcular el estadístico (parámetro estimado)**

El estadístico es el valor que se calcula a partir de la muestra y que se quiere extrapolar a la población de origen. En este caso es la diferencia de las medias muestrales ($X1 - X2$).

**3. Validar los supuestos para aplicar una prueba T**

Validar las condiciones de independencia, normalidad y homogeneidad de varianzas que permitan aplicar la prueba T.

**4. Determinar el tipo de test, una o dos colas**

Las pruebas de hipótesis pueden ser de una cola o de dos colas. Si la hipótesis alternativa emplea ">" o "<" se trata de una prueba de una cola, en el que solo se analizan desviaciones en un sentido. Si la hipótesis alternativa es del tipo "diferente de" se trata de una prueba de dos colas, en el que se analizan posibles desviaciones en las dos direcciones. Solo se emplean test de una cola cuando se sabe con seguridad que las desviaciones de interés son en un sentido y únicamente si se ha determinado antes de observar la muestra, no a posteriori.

**5. Determinar el nivel de significancia $\boldsymbol{\alpha}$**

El nivel de significancia $\alpha$ determina la probabilidad de error que se quiere asumir a la hora de rechazar la hipótesis nula. Se emplea como punto de referencia para determinar si el valor de p-value obtenido en la prueba de hipótesis es suficientemente bajo como para considerar significativas las diferencias observadas y por lo tanto rechazar $H0$.

A menor valor de alfa, menor probabilidad de rechazar la hipótesis nula. Por ejemplo, si se considera $\alpha = 0.05$, se rechazará la hipótesis nula en favor de la hipótesis alternativa si el p-value obtenido es menor que 0.05, y se tendrá una probabilidad del 5% de haber rechazado $H0$ cuando realmente es cierta.

En nivel de significancia debe establecerse en función de qué error sea más costoso:
• Error tipo I: Error de rechazar la $H0$ cuando realmente es cierta.
• Error tipo II: Error de no rechazar $H0$ cuando realmente es falsa.

**6. Cálcular p-value y comparar con el nivel de significancia**

Si las condiciones mencionadas en el punto 3 se cumplen, se puede aplicar la prueba **T de Student**.

En caso de que no se cumpliera el supuesto de igualdad de varianzas, se podría aplicar la variante **T de Welch**, que hace un ajuste a los grados de libertad y no asume igualdad de varianzas.

En caso de que no se cumpliera el supuesto de normalidad, se deberá aplicar una prueba no paramétrica, como la prueba de **Wilcoxon** o la de **Mann-Whitney**, que utilizan las posiciones de los datos. Estas pruebas no paramétricas no necesitan que las varianzas sean iguales tampoco.

**7. Identificar el tamaño del efecto**

El tamaño del efecto es la diferencia neta observada entre los grupos de un estudio. No se trata de una medida de inferencia estadística ya que no se pretende identificar si las poblaciones son significativamente diferentes, sino que simplemente indica la diferencia observada entre muestras, independientemente de la varianza que tengan.

Se trata de un parámetro que siempre debe acompañar a los p-values, ya que un p-value solamente indica si hay evidencias significativas para rechazar la hipótesis nula pero no dice nada de si la diferencia es importante o práctica. Esto último se averigua mediante el tamaño del efecto.

En el caso de las pruebas T de medias independientes, existen dos medidas posibles del tamaño del efecto: la **d de Cohen** y la **r de Pearson**. Ambas son equivalentes y pueden transformarse de una a otra. Cada una de estas medidas tiene unas magnitudes recomendadas para considerar el tamaño del efecto como pequeño, mediano o grande. La función cohen.d() del paquete *effsize* permite calcular el tamaño del efecto de la diferencia de medias independientes.

**D de Cohen para muestras independientes**

$$d = \frac{|diferencia\ de\ medias\ entre\ los\ grupos|}{sd}$$

Existen dos formas distintas se utilizan para calcular la sd conjunta de ambas muestras:

$$sd = \sqrt{\frac{(n-1)sd_1^2 + (n-2)sd_2^2)}{n1 + n2 - 2}}$$

$$sd = \sqrt{\frac{sd_1^2 + sd_2^2}{n1 + n2}}$$

Los límites más utilizados para clasificar el tamaño del efecto con d-Cohen son:

Si d es aproximadamente 0.2 se considera un efecto pequeño
Si d está alrededor de 0.5 se considera mediano
Si d está cerca de 0.8 se considera un efecto grande

**R de Pearson para t-test independiente:**

$$r = \sqrt{t^2}(t^2 + gl)$$

t = estadístico t obtenido en la prueba
gl = grados de libertad de la prueba

Los límites más utilizados para clasificar el tamaño del efecto con r son:

r aproximado a 0.1, pequeño
r aproximado a 0.3, mediano
r aproximado a 0.5, grande

**8. Calcular la potencia de la prueba**

Es importante calcular la potencia de la prueba, que determina la probabilidad de rechazar $H0$ cuando realmente es falsa. Este valor estará relacionado con la diferencia de las poblaciones (o en su defecto, la diferencia observada de las dos muestras) y el tamaño de la muestra.

En este punto también hay que considerar si las muestras de los dos niveles de estudio son iguales, y si las varianzas son conocidas (similares o no) o desconocidas. En caso de no saberlas, se podría utilizar las varianzas muestrales.

**9. Interpretar los resultados**

Si el p-value obtenido en la prueba utilizada es menor que el valor de $\alpha$ seleccionado, existen evidencias suficientes para rechazar $H0$ en favor de $HA$.

## Ejemplo utilizando RStudio (no se entrega en el Reporte)

Inicialmente realizaremos un ejercicio guiado que **no es necesario incluir en el Reporte** de Laboratorio. Luego, se proponen dos ejercicios que usted deberá realizar y que formarán el cuerpo del Reporte.

**Conjunto de datos**

El datset “**red2025aleatorio.csv**” contiene información sobre 250 observaciones de transmisiones de paquetes entre 2 nodos. El tiempo de transmisión en milisegundos está registrado en la variable tiempo.

También se cuenta con información de la topología (columna topología = estrella/red), si los paquetes eran grandes o pequeños y si la distancia entre nodos era larga o corta. El conjunto de datos también registra la hora y minutos en que se realizó la observación (para efectos de eventualmente clasificar el momento del día) y los porcentajes de uso del CPU y de la memoria registrados en el momento de la observación.

**Ejercicio guiado**

Se quiere determinar si existen evidencias significativas de que la topología de red (variable topología, con valores malla y estrella) afecta el tiempo de transferencia entre nodos..

Cargue la biblioteca tidyverse, el conjunto de datos **red2025aleatorio.csv** y verifique que los datos se cargaron con el comando head.

```
library(tidyverse)

# Si diera problemas, pueden instalarlo

install.packages("tidyverse")

redes = (read.csv(file.choose(), header=TRUE, encoding = "UTF-8"))
attach(redes)

head(redes, 10)
```

```
> head(redes, 10)
   Random1 Registro uso_cpu uso_memoria distancia hora minuto tiempo paquetes topologia
1        1      126      NA          18     largo    7     40   1.69 pequenos  estrella
2        1       73      29          30     corto   15     44   6.38  grandes  estrella
3        3       50      NA          32     corto    8     30   6.81 pequenos  estrella
4        3       13      25          19     corto   13     47   8.38  grandes  estrella
5        4       99      34          31     corto   14     44   7.50 pequenos     malla
6        4       90      32          25     corto   13     50   7.75 pequenos     malla
7        5       72      27          21     corto    8     26   5.94 pequenos  estrella
8        6        2      34          36     corto    5     35   7.69  grandes     malla
9        8       43      35          28     corto   10     25   6.88  grandes  estrella
10       9       80      39          37     corto    5     45   5.13  grandes     malla
```

**NOTA:** Para todo el ejercicio supondremos un nivel de significación de 0.05 (el Alpha!)

### 1. Establecer la Hipótesis

H0: no hay diferencia entre las medias poblacionales: µ(estrella) = µ(malla)

HA: si hay diferencia entre las medias poblacionales: µ(estrella) ≠ µ(malla)

### 2. Parámetro estimado (estadístico)

Se calcula la diferencia entre las medias muestrales. Primero se crea top_est con los datos de tiempos filtrados por la topología estrella, y luego top_malla para la topología de malla. Luego se calcula la diferencia de las medias de ambas:

```
top_est    <- redes %>% filter(topologia == "estrella") %>% pull(tiempo)

top_malla <- redes %>% filter(topologia == "malla") %>% pull(tiempo)

mean(top_malla) - mean(top_est)
```

```
> mean(top_malla) - mean(top_est)
[1] 0.3419935
```

Hay una diferencia de 0.34 milisegundos entre ambas medias, pero todavía no sabemos si esta diferencia es significativa.

### 3. Validar las condiciones para aplicar una prueba T

#### 3.1. Independencia:

Se trata de un muestreo aleatorio. Se puede afirmar que los eventos son independientes. Sin embargo, podemos también realizar un gráfico de las observaciones respecto del tiempo en que se tomaron:

```
dev.off()  # Cierra dispositivos gráficos activos

plot(redes$tiempo, axes = TRUE , ylab = "Tiempo observado", xlab = "Obser
vaciones")
abline(h = mean(redes$tiempo), col = "blue")
```

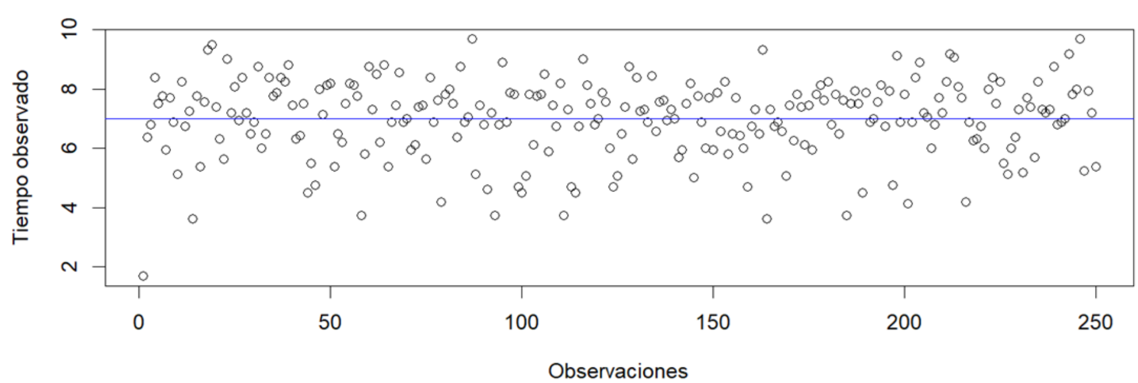


**Figura 1. Gráfico de observaciones en el tiempo**

Puede verse que no se observan patrones aparentes que comprometan la independencia de las observaciones.

**3.2. Normalidad**:

Se pueden crear histogramas para ver la forma de las observaciones.

```
ggplot(redes,aes(x = tiempo)) +
  geom_histogram(aes(y = ..density.., colour = topologia)) +
  facet_grid(.~ topologia) +
  theme_bw() + theme(legend.position = "none") +
  theme  (text = element_text(size=16))
```

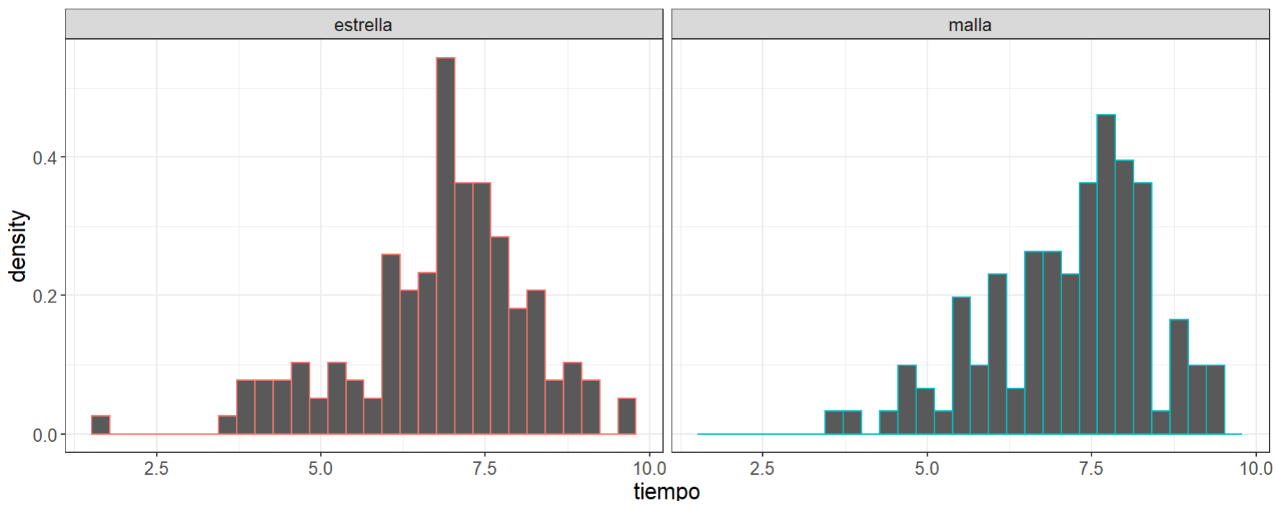


**Figura 2. Histogramas de la variable tiempo por topología**

Las formas de los histogramas no muestran claramente que las observaciones sigan una distribución normal. En ambos casos parece que hay una asimetría hacia la izquierda.

Otro gráfico que se puede utilizar es la gráfica de probabilidad normal (que en R puede hacerse con qqnorm y qqline). Por un lado, qqnorm grafica las observaciones ordenadas de la muestra contra sus respectivas frecuencias acumuladas, mientras que qqline muestra una línea de referencia de cómo deberían alinearse los datos si siguen una distribución normal.

```
par(mfrow = c(1, 2))

qqnorm(top_malla, xlab = "", ylab = "",
       main = "Topología malla", col = "firebrick")
qqline(top_malla)

qqnorm(top_est, xlab = "", ylab = "",
       main = "Topología estrella", col = "springgreen4")
qqline(top_est)
```

La instrucción par(mfrow = c(1,2)) permite graficar ambos conjuntos de datos uno al lado del otro.

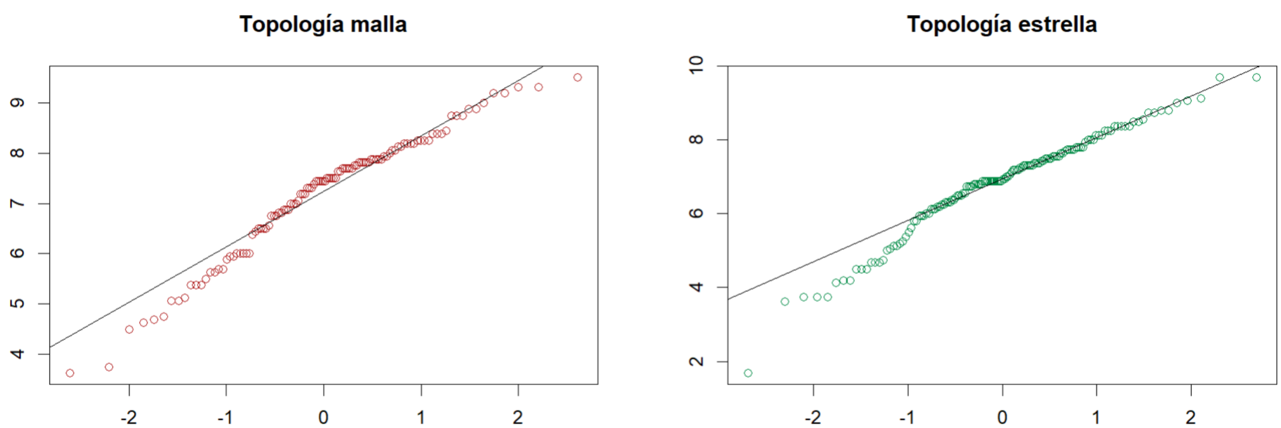


**Figura 3. Gráficos de cuantil-cuantil por topología**

Este gráfico **QQ** compara los **cuantiles** de los datos observados con los cuantiles teóricos de la distribución normal. Si los puntos del gráfico caen sobre la línea recta, los datos se distribuyen aproximadamente de forma normal. En ambos casos, las observaciones más pequeñas no caen sobre la línea teórica.

En este caso no es claro que los datos sigan la distribución normal, ni para la topología de malla ni para la de estrella.

Se puede entonces ejecutar el estadístico de prueba Shapiro-Wilks, primero para los datos de topología estrella y luego para los de malla.

```
shapiro.test(top_est)
```

```
> shapiro.test(top_est)

        Shapiro-Wilk normality test

data:  top_est
W = 0.9612, p-value = 0.0005352
```

Esta prueba utiliza la estrategia de hipótesis nula (los datos provienen de una población normal) y alternativa (los datos no provienen de una población normal). Si el valor p obtenido es menor que el alfa escogido con antelación, se rechaza la hipótesis nula.

Puede verse que el p-value es mucho menor que 0.05 (el Alpha que seleccionamos para todo el ejercicio), por lo que **se rechaza la hipótesis nula** de que las observaciones de la topología estrella siguen una distribución normal.

Aplicamos la prueba en top_malla:

```
shapiro.test(top_malla)
```

```
> shapiro.test(top_malla)

        Shapiro-Wilk normality test

data:  top_malla
W = 0.96282, p-value = 0.003693
```

En este caso, el p-value también es menor que 0.05, por lo que **se rechaza** la hipótesis nula de que las observaciones de la topología malla siguen una distribución normal.

Entonces, Ambas pruebas encuentran evidencias significativas de que los datos **no** proceden de poblaciones con distribución normal para ambas topologías. Este es un supuesto necesario para usar las pruebas T.

Una prueba no paramétrica, por ejemplo la prueba Mann-Withney, que se basa en el orden o ranking de las observaciones, sería más adecuados que la prueba T. Otra opción sería estudiar si los datos anómalos son excepciones que se pueden excluir del análisis.

### 3.3. Igualdad de varianzas:

Inicialmente se puede revisar visualmente si las rectas teóricas qq line creadas para estos datos se muestran paralelas. Para ello hay que estar seguros que los ejes tienen la misma escala para ambos gráficos. En el gráfico que se usó arriba para ver normalidad los **ejes Y no coinciden**, por lo que no se pueden comparar así como los graficamos. Aunque se puede construir en R un gráfico fijando los ejes, vamos a realizar la prueba con estadísticos.

También podemos graficar boxplots con los datos para observar visualmente las distribuciones:

```
ggplot(data = redes) +
  geom_boxplot(aes(x = topologia, y = tiempo, colour = topologia)) +
  theme_bw() + theme(legend.position = "none") +
  theme (text = element_text(size=16))
```

El resultado es el siguiente gráfico:

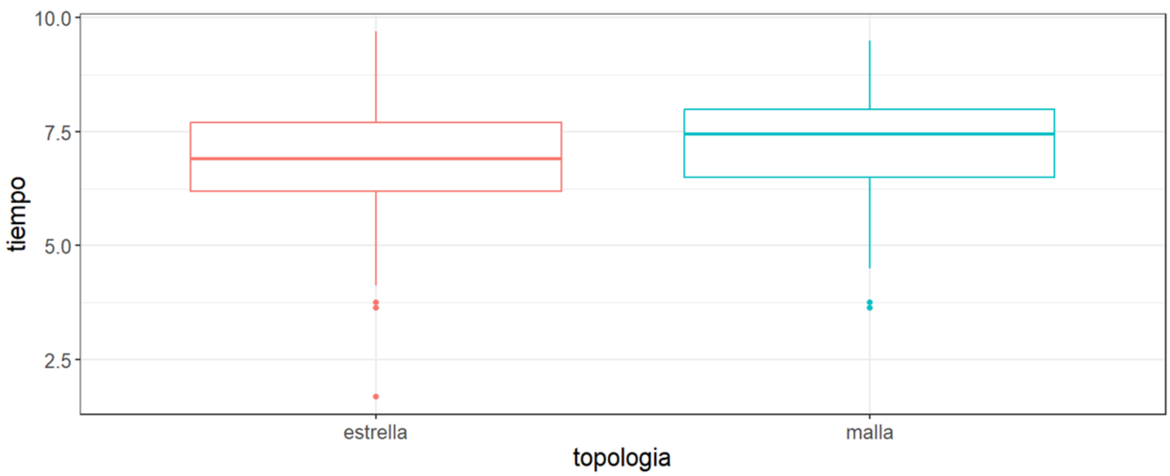


**Figura 4. Gráficos de boxplot de la variable tiempo por topología**

Viendo estos boxplots parecería que los datos para ambas topologías sí presentan varianzas similares. Vamos a revisarlo con estadísticos para estar seguros.

Existen varios estadísticos que permiten comparar varianzas. Dado que **no** se cumple el criterio de normalidad, uno de los recomendados es el test **Levene** o el test no paramétrico de **Fligner-Killeen**.

Si sí se cumpliera el criterio de normalidad, se podría usar la prueba de **Bartlett** (en R bartlett.test)

Nuevamente estas pruebas utilizan la estrategia de hipótesis nula (los datos poseen varianzas iguales) y alternativa (los datos no poseen varianzas iguales). Si el valor p obtenido es menor que el alfa escogido con antelación, se rechaza la hipótesis nula.

Ejecutamos las pruebas estadísticas:

```
library(car)

fligner.test(tiempo ~ topologia, data = redes)
```

```
> fligner.test(tiempo ~ topologia, data = redes)

        Fligner-Killeen test of homogeneity of variances

data:  tiempo by topologia
Fligner-Killeen:med chi-squared = 0.19, df = 1, p-value =
0.6629
```

La prueba de Flinger-Killeen genera un p-value de 0.6629, mayor que el Alpha establecido de 0.05, por lo que no hay razones para rechazar la hipótesis nula de que las varianzas son iguales. Entonces sí concluimos que las varianzas son iguales.

Si ejecutamos la prueba de Levene:

```
leveneTest(tiempo ~ topologia, data = redes, center = "median")
```

```
> leveneTest(tiempo ~ topologia, data = redes, center = "median")
Levene's Test for Homogeneity of Variance (center = "median")
       Df F value Pr(>F)
group   1  0.1904  0.663
      248
Warning message:
In leveneTest.default(y = y, group = group, ...) : group coerced to factor.
```

Esta prueba genera un p-value de 0.663, que tampoco encuentra evidencia significativa (para alpha = 0.05) para rechazar la hipótesis nula de que las variancias sean distintas entre ambas poblaciones.

**Atención**: Si las varianzas no fuesen iguales (y sí se cumplen los otros supuestos, particularmente el de normalidad) se podría usar la prueba T con la corrección de **Welch**.

#### 4. Determinar el tipo de test

Se trata de una prueba de dos colas, dado que la hipótesis nula es con igualdad.

Además, sabemos que las varianzas son iguales.

También sabemos que las observaciones no siguen una distribución normal.

**5. Determinar el nivel de significancia**

Como se indicó al inicio del ejercicio, se usará α = 0.05.

**6. Cálculo de p-value**

R tiene una función integrada que permite realizar la prueba T para una o dos muestras (llamada T de Student), tanto con corrección de Welch (en caso de que las varianzas no sean iguales) como sin ella.

Esta función devuelve tanto el p-value del test como el intervalo de confianza para la verdadera diferencia de medias.

1) Aunque no contamos con normalidad de los datos, vamos a ejecutar la prueba T para efectos pedagógicos de este laboratorio, y así practicar el utilizar este estadístico tan conocido.

Realizamos entonces la prueba T para los datos de las dos topologías, indicando mu = 0, dado que estamos usando la prueba para comparar ambas medias, o lo que es igual, que la diferencia de las medias es 0.

También indicamos que las varianzas son iguales y que el nivel de confianza es 0.95, ó 1 – 0.005 (el nivel de significación):

```
t.test(
  x           = top_est,
  y           = top_malla,
  alternative = "two.sided",
  mu          = 0,
  var.equal   = TRUE,
  conf.level  = 0.95
)
```

```
> t.test(
+   x           = top_est,
+   y           = top_malla,
+   alternative = "two.sided",
+   mu          = 0,
+   var.equal   = TRUE,
+   conf.level  = 0.95
+ )

        Two Sample t-test

data:  top_est and top_malla
t = -2.0544, df = 248, p-value = 0.04098
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.66986317 -0.01412384
sample estimates:
mean of x mean of y
 6.835643  7.177636
```

Esta prueba muestra un p-value de 0.0408, menor que el nivel de significancia de 0.05, por lo que sí habría evidencia para rechazar la hipótesis nula. Recordemos que usamos esta prueba con fines académicos.

2) Ahora bien, dado que **no** es cierto el supuesto de **normalidad**, vamos a ejecutar la prueba no paramétrica de Mann-Whitney, que en R forma parte de la función wilcox.test(). Esta función permite hacer la prueba Wilcoxon para datos pareados o la prueba Mann-Whitney para datos no pareados, que es este caso.

```
wilcox.test(
  x = top_est,
  y = top_malla,
  alternative = "two.sided",
  mu = 0,
  paired = FALSE,
  conf.int = 0.95)
```

El resultado es:

```
> wilcox.test(
+   x = top_est,
+   y = top_malla,
+   alternative = "two.sided",
+   mu = 0,
+   paired = FALSE,
+   conf.int = 0.95)

        Wilcoxon rank sum test with continuity correction

data:  top_est and top_malla
W = 6456.5, p-value = 0.02846
alternative hypothesis: true location shift is not equal to 0
95 percent confidence interval:
 -6.299955e-01 -4.662828e-05
sample estimates:
difference in location
            -0.3699581
```

Puede verse que el p-value de 0.028 es menor que el Alpha 0.05. Sí se dispone de evidencia para rechazar la hipótesis nula de que las medias de los tiempos de transferencia son iguales para las dos topologías.

**7. Cálculo tamaño de efecto**

Utilizamos la biblioteca effsize, con la función cohen.d:

```
library(effsize)

cohen.d(formula = tiempo ~ topologia, data = redes, paired = FALSE)
```

```
> cohen.d(formula = tiempo ~ topologia, data = redes, paired = FALSE)

Cohen's d

d estimate: -0.2617574 (small)
95 percent confidence interval:
      lower       upper
-0.51376149 -0.00975327
```

Puede verse que el d estimado está alrededor de 0.2 (en magnitud), por lo que se considera que el efecto de la topología en los tiempos de transmisión es pequeño.

**8. Potencia de la prueba**

Este valor estará relacionado con la diferencia de las poblaciones (o en su defecto, la diferencia observada de las dos muestras) y el tamaño de la muestra. También debemos considerar si las muestras de los dos niveles de estudio son iguales, y si las varianzas son conocidas (similares o no) o desconocidas.

En este caso particular, sabemos que las varianzas muestrales son similares. No contamos con información sobre las varianzas poblacionales. Como las varianzas son similares, podríamos tomar cualquiera de ellas para hacer el cálculo de la potencia (que necesita la desviación estándar), o también calcular la media de ambas y usar esta media. Entonces:

```
sd_grupo1 <- sd(top_est)
sd_grupo2 <- sd(top_malla)
#
# Se puede usar cualquiera de las dos dado que son similares
# o un promedio entre ambas

sd_comun <- mean(c(sd_grupo1, sd_grupo2))

sd_comun
```

```
> sd_comun
[1] 1.298417
```

**Atención:** Si las **varianzas no fueran iguales** o similares, se debe modificar el cálculo para el sd_comun.

```
# Desviaciones estándar de los grupos
sd_grupo1 <- sd(grupo1)
sd_grupo2 <- sd(grupo2)

# Tamaños de muestra
n1 <- length(grupo1)
n2 <- length(grupo2)

# Calcular la desviación estándar combinada
sd_combinada <- sqrt(((n1 - 1) * sd_grupo1^2 + (n2 - 1) * sd_grupo2^2) /
(n1 + n2 - 2))
#
sd_comun
```

También se necesita el **delta** que es la diferencia entre las medias de las poblaciones. Como no tenemos información, usaremos la diferencia entre las medias muestrales, que ya habíamos calculado en el punto 2:

```
delta_observado <- mean(top_malla) - mean(top_est)

delta_observado
```

```
> delta_observado
[1] 0.3419935
```

Además, debemos revisar si las muestras son de igual tamaño:

```
length(top_est)
length(top_malla)
```

```
> length(top_est)
[1] 140
> length(top_malla)
[1] 110
```

En este caso hay 140 observaciones para la topología estrella y 110 para la topología de malla. Para el cálculo de la potencia utilizaremos el valor más pequeño de los dos.

Seguidamente ejecutamos el comando power.t.test() indicando:

```
 power.t.test(n = 110, delta = delta_observado, sd = sd_comun, sig.level
= 0.05, type = "two.sample")
```

El resultado es el siguiente:

```
> power.t.test(n = 110, delta = delta_observado, sd = sd_comun, sig.level = 0.05, type
= "two.sample")

     Two-sample t test power calculation 

              n = 110
          delta = 0.3419935
             sd = 1.298417
      sig.level = 0.05
          power = 0.4939359
    alternative = two.sided

NOTE: n is number in *each* group
```

De este resultado se puede ver que con este conjunto de datos se obtuvo una potencia de prueba de 0.49, baja si se compara con 0.80 que sería más apropiada. Esto está determinado por el delta que se encontró: diferencias grandes entre las medias son más fácilmente identificables con muestras menores, mientras que diferencias pequeñas necesitan muestras más grandes.

Se puede utilizar esta misma función para saber de cuánto debería ser la muestra para llegar a una potencia de 0.80. Se modifica la invocación de la siguiente forma:

```
power.t.test(power = 0.8, delta = delta_observado, sd = sd_comun, sig.le
vel = 0.05, type = "two.sample")
```

```
> power.t.test(power = 0.8, delta = delta_observado, sd = sd_comun, sig.level = 0.05,
type = "two.sample")

     Two-sample t test power calculation

              n = 227.2362
          delta = 0.3419935
             sd = 1.298417
      sig.level = 0.05
          power = 0.8
    alternative = two.sided

NOTE: n is number in *each* group
```

El resultado muestra que para obtener una potencia de prueba de 0.80 que esté relacionada con el delta encontrado (0.34) se necesitan al menos 227 observaciones para cada topología.

**9. Análisis de resultados y conclusiones**

Puede verse que el p-value obtenido con la prueba Mann-Whitney de 0.028 es menor que el Alpha 0.05. Sí se dispone de evidencia para rechazar la hipótesis nula de que las medias de los tiempos de transferencia son iguales para las dos topologías.

A esta conclusión también llegamos utilizando la prueba T de Student, que arrojó un p-value de 0.0408, también menor a 0.05. Como no contamos con normalidad de los datos, este resultado lo tomamos con precaución.

A su vez, el tamaño de efecto atribuible a la topología, medido por d-Cohen, es pequeño (-0.26).

Finalmente, la potencia de la prueba es baja, de 0.49. Para obtener una prueba con una potencia de 0.80 se necesitaría tener una muestra de 227 observaciones para cada topología.

## Trabajo a entregar

Para esta segunda parte seguiremos utilizando el conjunto de datos red2025aleatorio.csv.

A. Realice un análisis similar sobre los tiempos de transferencia, pero en lugar de considerar la topología ahora analice la variable **paquetes**. Es decir, evalúe si los tamaños de los paquetes están relacionados con el tiempo de transferencia.

B. Haga lo mismo para la variable **distancia**. Es decir, evalúe si la distancia entre los nodos incide en el tiempo de transferencia.

**Atención**:

1. Para ambos apartados A y B debe realizar los 9 puntos que se efectuaron para la variable **topología**.
2. Para el punto 6. Cálculo del p-value, ejecute la prueba T de Student (con o sin la corrección de Welch, dependiendo de las varianzas) sin importar si se cumple el supuesto de normalidad.
   Luego, si no se cumple normalidad, aplique también la prueba de Mann-Whitney.

**Indicaciones de la entrega**

- El trabajo debe realizarse en grupos de dos personas.
- En ambos casos (puntos A y B) debe realizar los pasos 1 al 9, incluyendo en el reporte el código R utilizado y los resultados de la ejecución del código en R (que pueden ser valores o gráficos).
- Use la prueba T pesar de que no se cuente con normalidad. Sin embargo, notifique en su reporte si esta situación se presenta. Adicionalmente, ponga atención a los resultados de los estadísticos que usa para verificar supuestos, por si eventualmente debe aplicar ciertas correcciones a la hora de aplicar otros estadísticos o la misma prueba T.
  Además, si no se cumple normalidad, ejecute también la prueba Mann-Whitney.
- El reporte es un documento con un formato uniforme y coherente, en el que se incluye una pequeña introducción (al menos una línea de texto) para cada gráfico o código R que se incluye. Debe poder leerlo una persona que no cuente con el enunciado.
- Los gráficos deben ser apropiadamente presentados, con etiquetas correctas y debidamente rotulado con la información que presenta.

- El código R debe ser también incluido apropiadamente en el reporte, en forma de imágenes o de texto debidamente formateado e indentado para que se distinga del resto del texto.
- Las salidas o los gráficos de código R que deba incluir deben aparecer junto al código que los genera. No es válido incluir varias instrucciones de código y luego una serie de respuestas y gráficos. Cada instrucción que produzca una salida o un gráfico debe estar asociada directamente con el resultado (el resultado inmediatamente después de la instrucción que lo generó).
- Las entregas se realizarán en formato PDF. La letra debe ser Times New Roman, Arial o Cambria. Los tamaños permitidos son 10-12.
- Debe incluir una portada de página completa en la que incluya el nombre del curso, del laboratorio, así como los nombres completos de los estudiantes. Esta página será utilizada por el profesor para apuntar la nota y también para incluir comentarios en caso de tener que incluir generales del informe.
- El reporte se debe entregar en Mediación Virtual, en un archivo .pdf que pueda ser leído en programas comerciales de uso habitual. Debe verificar que el .pdf que subió a Mediación Virtual contiene los ejercicios resueltos y que el archivo puede abrirse correctamente. En caso de problemas con el archivo .pdf (no abre correctamente, está corrupto, etc.) se considerará que no entregó la tarea.
- Las entregas tardías se penalizarán con un 10% de la nota luego de vencida la fecha y hora de entrega, más un 10% adicional por cada hora de retraso.