

Estadística y R para Ciencias de la Salud

---

## Tema 6. Bioestadística

# Índice

## Esquema

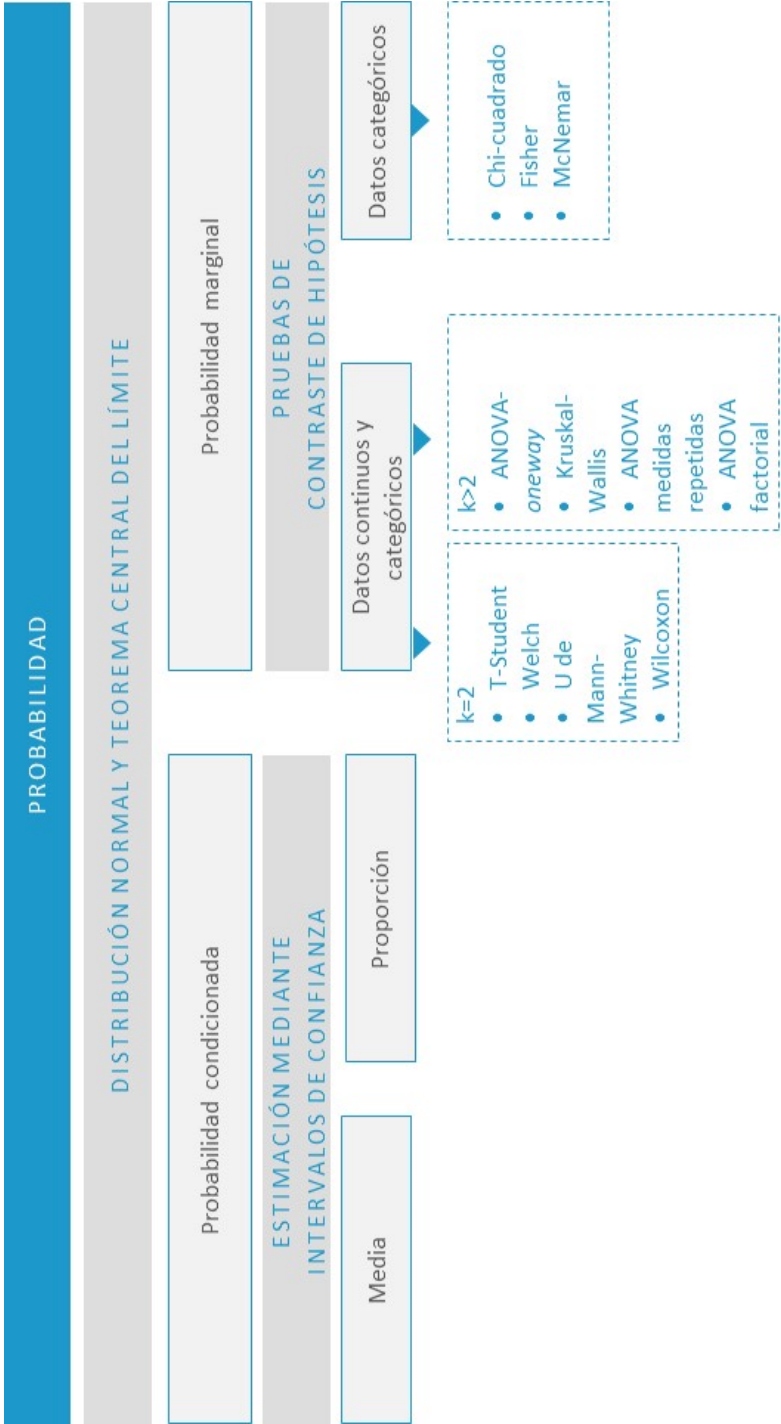
### Ideas clave

- 6.1. Introducción y objetivos
- 6.2. Probabilidad: base de la bioestadística
- 6.3. Estimación mediante intervalos de confianza
- 6.4. Pruebas de hipótesis
- 6.5. Aplicaciones de las pruebas de contrastes de hipótesis
- 6.6. Referencias bibliográficas

### A fondo

- ¿Cuál es la forma correcta de aplicar la prueba de comparación múltiple?
- Análisis post-hoc ANOVA
- Ejemplos de contrastes de hipótesis con R Markdown

### Test



## 6.1. Introducción y objetivos

La bioestadística inferencial es una rama de la estadística aplicada que se utiliza en el ámbito de la biología y las ciencias de la salud para realizar inferencias y tomar decisiones basadas en muestras de datos [1]. A diferencia de la estadística descriptiva, que se enfoca en describir y resumir características de una muestra, la bioestadística inferencial se centra en hacer **generalizaciones** y sacar **conclusiones** acerca de una población más amplia utilizando información limitada de una muestra representativa [2].

En la investigación biomédica y de la salud, a menudo, no es factible o práctico estudiar a toda una población debido a limitaciones de tiempo, recursos y accesibilidad. En cambio, se toman **muestras representativas** de la población y se utilizan **métodos estadísticos inferenciales** para obtener conclusiones sobre la población en su conjunto. Estos métodos se basan en la teoría de la probabilidad y en modelos estadísticos para estimar parámetros desconocidos y realizar pruebas de hipótesis.

La bioestadística inferencial se aplica en una amplia gama de áreas, como ensayos clínicos, estudios epidemiológicos, genética, estudios de salud pública y análisis de datos en ciencias biológicas. Permite a los investigadores obtener conclusiones más allá de los datos observados y proporciona herramientas para evaluar la certeza de las conclusiones obtenidas. Al comprender los principios y métodos de la bioestadística inferencial, los investigadores pueden tomar decisiones informadas basadas en la evidencia y contribuir al avance de la investigación biomédica y de la salud.

Por lo tanto, el objetivo de este tema se centrará en:

- ▶ Comprender los conceptos fundamentales de la probabilidad incluyendo la aplicación de distribuciones de probabilidad en el contexto de la bioestadística inferencial.
- ▶ Adquirir habilidades para realizar estimaciones mediante intervalos de confianza utilizando técnicas adecuadas para determinar la precisión y confiabilidad de las estimaciones en estudios biomédicos y de salud.
- ▶ Desarrollar la capacidad para realizar pruebas de hipótesis aplicando métodos estadísticos para evaluar afirmaciones y tomar decisiones basadas en la evidencia obtenida de muestras de datos.
- ▶ Familiarizarse con las aplicaciones prácticas de las pruebas estadísticas en la bioestadística inferencial, incluyendo su uso en estudios epidemiológicos, ensayos clínicos y análisis de datos en ciencias biológicas.
- ▶ Explorar los métodos de computación intensiva y comparaciones múltiples utilizados en la bioestadística inferencial, comprendiendo sus fundamentos teóricos y su aplicación en la interpretación y análisis de datos de gran escala.

## 6.2. Probabilidad: base de la bioestadística

La probabilidad es un concepto fundamental en la bioestadística, ya que proporciona las **bases teóricas** necesarias para el **análisis** y la **interpretación** de los datos en la investigación biomédica y de la salud [3].

### Conceptos básicos de probabilidad

La probabilidad es una **medida numérica** que asigna una posibilidad a un evento específico. En el contexto de la bioestadística, la probabilidad se utiliza para **cuantificar** la incertidumbre asociada a los resultados de un experimento o estudio.

Los conceptos clave en probabilidad incluyen:

- ▶ **Espacio muestral:** es el conjunto de todos los **posibles resultados** de un experimento [4]. Por ejemplo, al lanzar una moneda el espacio muestral está compuesto por dos posibles resultados: cara o cruz.
- ▶ **Evento:** es un **subconjunto** del espacio muestral que contiene uno o varios resultados [5]. Por ejemplo, en el caso del lanzamiento de una moneda el evento de obtener cara es un subconjunto que contiene solo un resultado.
- ▶ **Probabilidad:** es una medida numérica que asigna una posibilidad a un **evento específico** [4]. Se expresa como un número entre 0 y 1 donde 0 representa una probabilidad nula (evento imposible) y 1 representa una probabilidad total (evento seguro).

### Conceptos básicos de probabilidad condicionada y marginal

La **probabilidad condicionada** se refiere a la probabilidad de que ocurra un evento dado que otro evento ya ha ocurrido [6]. Se denota en lenguaje matemático como  $P(A|B)$ , donde A y B son eventos. La fórmula para calcular la probabilidad condicionada se representa en la Figura 1.

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Figura 1. Probabilidad condicionada del evento A y B simultáneamente, donde  $P(A \cap B)$  representa la probabilidad de que ocurran ambos eventos, A y B simultáneamente, y  $P(B)$  es la probabilidad del evento B. Fuente: elaboración propia.

Sin embargo, la **probabilidad marginal** se refiere a la probabilidad de que ocurra un evento sin tener en cuenta ningún otro evento específico [7]. Es la probabilidad de un evento considerado de forma **independiente**.

Por ejemplo, si se tienen dos eventos A y B, la probabilidad marginal de A se calcula sumando todas las probabilidades conjuntas de A con cada posible resultado de B. Esto se expresa como aparece en la Figura 2.

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \dots + P(A \cap B_n)$$

Figura 2. Probabilidad marginal del evento A, donde  $B_1, B_2, \dots, B_n$  son los posibles resultados de B. Fuente: elaboración propia.

La probabilidad marginal se utiliza para analizar la **distribución de probabilidades** de un evento individual sin considerar otros eventos relacionados.

A continuación, se muestra un ejemplo de **código de R** para el cálculo de la probabilidad condicional y marginal (en color morado se muestra la salida de R, para su reproducibilidad en el programa no copiar esa parte).

```
# Ejemplo 1: Generamos distribución no normal/normal

# Generación de datos biomédicos inventados
set.seed(123) # Establecer una semilla para reproducibilidad

# Creación de datos de género (0: masculino, 1: femenino)
genero <- sample(0:1, 1000, replace = TRUE, prob = c(0.6, 0.4))

# Creación de datos de edad (entre 20 y 60 años)
edad <- sample(20:60, 1000, replace = TRUE)

# Creación de datos de enfermedad (0: no enfermo, 1: enfermo)
enfermedad <- rbinom(1000, 1, prob = ifelse(genero == 1, 0.3, 0.2))

# Creación de datos de resultado de prueba (0: negativo, 1: positivo)
resultado_prueba <- rbinom(1000, 1, prob = ifelse(enfermedad == 1, 0.9, 0.05))

# Cálculo de la probabilidad marginal de tener la enfermedad
prob_enfermedad <- mean(enfermedad)

# Cálculo de la probabilidad condicionada de tener la enfermedad dado el género
femenino
prob_enfermedad_genero_femenino <- mean(enfermedad[genero == 1])

# Cálculo de la probabilidad condicionada de tener la enfermedad dado un resultado de
prueba positivo
prob_enfermedad_resultado_positivo <- mean(enfermedad[resultado_prueba == 1])

# Imprimir los resultados
cat("Probabilidad marginal de tener la enfermedad:", prob_enfermedad, "\n")
Probabilidad marginal de tener la enfermedad: 0.243

cat("Probabilidad condicionada de tener la enfermedad dado el género femenino:",
prob_enfermedad_genero_femenino, "\n")
Probabilidad condicionada de tener la enfermedad dado el género femenino: 0.2977099

cat("Probabilidad condicionada de tener la enfermedad dado un resultado de prueba
positivo:", prob_enfermedad_resultado_positivo, "\n")
Probabilidad condicionada de tener la enfermedad dado un resultado de prueba positivo:
0.851711
```

Figura 3. Probabilidad condicional y marginal. Fuente: elaboración propia.

## Distribución normal: distribución de probabilidad continua

La **distribución normal**, también conocida como distribución de Gauss o distribución de campana de Gauss, es una de las distribuciones de probabilidad más importantes y ampliamente **utilizadas** en la bioestadística. Es una distribución continua que se caracteriza por su forma de **campana simétrica** alrededor de su media y por tener una forma específica definida por su media y desviación estándar.



## Características de la distribución normal [8]:

- ▶ **Forma de campana:** la distribución normal tiene una **forma simétrica** en la que los valores se agrupan en torno a su media. La curva de la distribución es suave y continua.
- ▶ **Media y desviación estándar:** la media de una distribución normal indica el **punto central** de la distribución, mientras que la desviación estándar mide la **dispersión** de los datos alrededor de la media. La media determina la posición central de la campana y la desviación estándar determina la amplitud de la campana.
- ▶ **Propiedad de la campana:** la distribución normal tiene una propiedad clave conocida como la **regla empírica** o la regla del 68–95–99,7. Según esta regla, aproximadamente el 68 % de los datos caen dentro de **una desviación estándar** de la media, alrededor del 95 % caen dentro de **dos desviaciones estándar** y cerca del 99.7 % caen dentro de **tres desviaciones estándar**.
- ▶ **Parámetros de la distribución:** la distribución normal se caracteriza completamente por su media y su desviación estándar. Estos parámetros determinan la **forma** y la **posición** de la curva de la distribución normal.

La distribución normal se aplica en numerosas **áreas** de la bioestadística debido a su amplio uso y propiedades estadísticas. Algunas **aplicaciones** comunes incluyen:

- ▶ **Modelado de datos biométricos:** La distribución normal se utiliza para **modelar características biológicas**, como el peso, la altura, la presión arterial, los niveles de colesterol, entre otros. Estos datos suelen tener una distribución que se asemeja a una curva normal.
- ▶ **Inferencia estadística:** La distribución normal es fundamental en la inferencia estadística, donde se utiliza para calcular **intervalos de confianza** y realizar **pruebas de hipótesis**. Estas técnicas permiten hacer inferencias sobre parámetros poblacionales utilizando muestras de datos.

- ▶ **Estimación de probabilidades:** La distribución normal se utiliza para estimar la probabilidad de que un **evento ocurra** dentro de un **rango específico** de valores. Por ejemplo, se puede utilizar para estimar la probabilidad de que un individuo tenga un nivel de glucosa en sangre dentro de un rango determinado.

### Teorema central del límite

El teorema central del límite establece que, bajo ciertas condiciones, la distribución de la **suma** o promedio de un gran número de **variables aleatorias independientes** tiende a aproximarse a una distribución **normal**, independientemente de la forma de las distribuciones originales [9].

Este teorema es fundamental en la estadística y tiene una estrecha relación con la distribución normal [9]. El teorema central del límite permite explicar por qué la distribución normal es tan común en la naturaleza y se encuentra en muchos **fenómenos biológicos y sociales**.

La relación entre el **teorema central del límite** y la **distribución normal** se basa en los siguientes puntos:

- ▶ **Suma de variables aleatorias:** el teorema central del límite establece que la suma de un gran número de variables aleatorias independientes, independientemente de su distribución original, convergen hacia una distribución normal (Figura 4). Esto implica que, cuando se suman o promedian muchas observaciones independientes de una misma variable aleatoria, la distribución de los resultados se aproximará a una distribución normal [10].

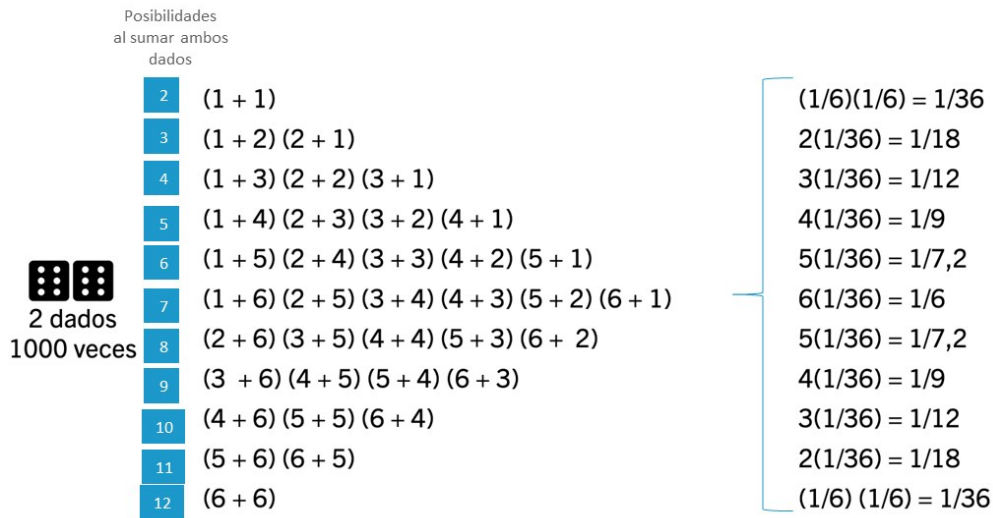


Figura 4. Distribución normal: suma de variables aleatorias. Fuente: elaboración propia.

- **Independencia y tamaño muestral:** para que se aplique el teorema central del límite se requiere que las variables aleatorias sean independientes y que el tamaño de la muestra sea lo suficientemente grande [10]. Esta **independencia y tamaño muestral** adecuado permite que las fluctuaciones aleatorias se cancelen entre sí, dando lugar a una distribución aproximadamente normal.

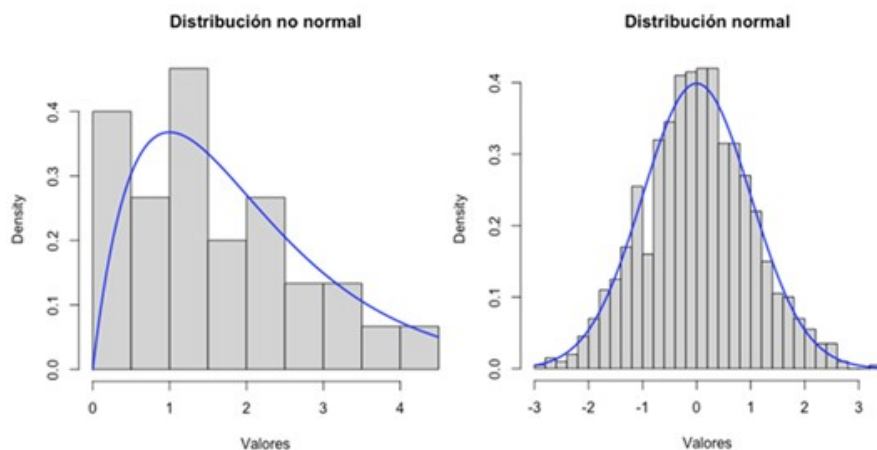


Figura 5. Distribución normal: tamaño muestral. Fuente: elaboración propia

```
# Ejemplo 2 de Figura 4: Generamos distribución no normal/normal

# Establecer una semilla de aleatorización para reproducibilidad

set.seed(123)

# Generar una distribución no normal con tamaño muestral pequeño

n_small <- 30

data_nonnormal <- rgamma(n_small, shape = 2, rate = 1)

# Graficar la distribución no normal

hist(data_nonnormal,

breaks = 10,

main = "Distribución no normal",

xlab = "Valores",

prob = TRUE)

curve(dgamma(x, shape = 2, rate = 1), add = TRUE, col = "blue", lwd = 2)

# Generar una distribución normal con tamaño muestral grande

n_large <- 1000

data_normal <- rnorm(n_large, mean = 0, sd = 1)

# Graficar la distribución normal

hist(data_normal,

breaks = 30,
```

```
main = "Distribución normal",  
  
xlab = "Valores",  
  
prob = TRUE)  
  
curve(dnorm(x, mean = 0, sd = 1), add = TRUE, col = "blue", lwd = 2)
```

- **Forma de campana:** la distribución normal es una distribución simétrica en forma de campana [10]. El teorema central del límite explica por qué se observa esta forma de campana en muchos conjuntos de datos reales. Al sumar o promediar un gran número de observaciones las **fluctuaciones aleatorias individuales** se **suavizan** y se obtiene una **distribución simétrica** alrededor de la media.

### 6.3. Estimación mediante intervalos de confianza

La estimación mediante intervalos de confianza es una técnica fundamental en la estadística inferencial que nos permite obtener una **estimación** de un **parámetro desconocido** de una población a partir de una muestra de datos [11]. En lugar de proporcionar una estimación puntual única, los intervalos de confianza nos brindan un rango de valores, dentro del cual es probable que se encuentre el verdadero valor del parámetro.

El proceso de construcción de un intervalo de confianza implica **dos elementos** clave: el nivel de confianza y la estimación puntual. El **nivel de confianza** se refiere a la **probabilidad** de que el intervalo de confianza capture el verdadero valor del parámetro [12]. Comúnmente se utiliza un nivel de confianza del 95 % lo que implica que, en ese porcentaje de las muestras que podríamos obtener, el intervalo de confianza contendría el verdadero valor del parámetro.

La estimación puntual se obtiene a través de una **estadística descriptiva**, como la media muestral o la proporción muestral, que brinda una **aproximación** del valor del parámetro en la población [13]. Luego, se utiliza la **distribución de probabilidad** correspondiente (como la distribución t-Student) para calcular el tamaño del intervalo de confianza [13].

Es importante tener en cuenta que un intervalo de confianza no nos dice nada acerca de la probabilidad de que un valor individual en particular se encuentre dentro del intervalo. Más bien, nos da una idea de la **precisión** de nuestra estimación y la **confianza** que podemos tener en ella.

Los intervalos de confianza son especialmente útiles en la **toma de decisiones** y la **interpretación de resultados** [14]. Nos permiten evaluar la precisión de nuestras estimaciones y **comparar** diferentes grupos o condiciones en un estudio. Además,

nos brindan una medida de incertidumbre y nos ayudan a comunicar los resultados de manera más completa y precisa [14].

Intervalos de confianza al 95 % para una media

Para calcular un intervalo de confianza al 95 % para una media en R puedes utilizar la función `t.test()`. A continuación, se muestra un ejemplo del caso:

```
# Ejemplo 3: calcular un intervalo de confianza al 95% para una media

# Datos de muestra
datos <- c(23, 18, 21, 16, 25, 20, 22, 19, 24, 17)

# Cálculo del intervalo de confianza para la media
resultado <- t.test(datos, conf.level = 0.95)
intervalo_confianza <- resultado$conf.int

# Imprimir el intervalo de confianza
print(intervalo_confianza)

[1] 18.33415 22.66585 # IC 95% = (18.33 a 22.66)
```

Figura 6. Intervalo de confianza para una media. Fuente: elaboración propia.

## Intervalos de confianza al 95 % para una proporción

Para calcular un intervalo de confianza al 95 % para una proporción en R puedes utilizar la función `prop.test()`. A continuación, se muestra un ejemplo del caso:

```
# Ejemplo 4: calcular un intervalo de confianza al 95% para una
proporción

# Datos de muestra
 exitos <- 35
 total <- 100

# Cálculo del intervalo de confianza para la proporción
 resultado <- prop.test(exitos, total, conf.level = 0.95)
 intervalo_confianza <- resultado$conf.int

# Imprimir el intervalo de confianza
 print(intervalo_confianza)

[1] 0.2591235 0.4525560 # IC 95% = (0.26 a 0.45)
```

Figura 7. Intervalo de confianza para una proporción. Fuente: elaboración propia.



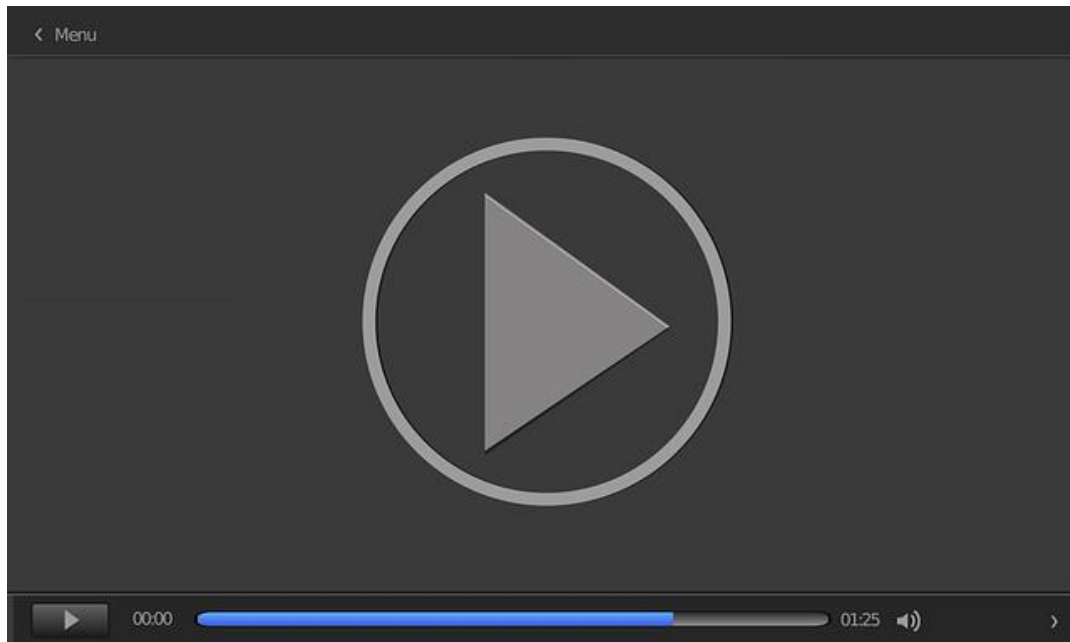
## 6.4. Pruebas de hipótesis

En estadística y dentro de las ciencias biomédicas la hipótesis es una **afirmación** o **suposición** acerca de una característica desconocida de una **población** o de un **fenómeno** en estudio. Las **pruebas de hipótesis** son procedimientos utilizados para evaluar la evidencia en contra o a favor de una hipótesis mediante el análisis de los datos de una muestra [15]. Estas pruebas permiten **tomar decisiones** basadas en la evidencia empírica obtenida de los datos.

En una prueba de hipótesis, se plantean dos hipótesis: la **hipótesis nula** ( $H_0$ ) y la **hipótesis alternativa** ( $H_1$ ). La hipótesis nula establece una **afirmación** sobre el valor del parámetro o la característica en estudio que se asume como verdadera, a menos que se disponga de suficiente evidencia en contra [16]. Por otro lado, la hipótesis alternativa propone una **afirmación contraria** a la hipótesis nula, es decir, sostiene que el valor del parámetro es diferente, mayor o menor al valor establecido por la hipótesis nula [16].

La prueba de hipótesis se realiza mediante el **cálculo** de una **estadística de prueba**, que es una medida obtenida de los datos de la muestra y se compara con un **valor crítico** o un rango de valores críticos [17]. La elección del valor crítico depende del **nivel de significancia** establecido, que es la probabilidad de cometer un error sistemático (error llamado tipo I para este caso) al rechazar incorrectamente la hipótesis nula cuando es verdadera [17].

A continuación, puedes acceder al vídeo *Curvas de potencia estadística para la realización de ensayos*.



Curvas de potencia estadística para la realización de ensayos

---

Accede al vídeo:

<https://unir.cloud.panopto.eu/Panopto/Pages/Embed.aspx?id=cb182555-8f73-42cb-a8e1-b08900db981c>

---

Si la estadística de prueba cae dentro de la **región de rechazo** determinada por los valores críticos se rechaza la hipótesis nula en favor de la hipótesis alternativa. En cambio, si la estadística de prueba cae dentro de la **región de no rechazo** no se tienen suficientes pruebas para rechazar la hipótesis nula.

Es importante destacar que las pruebas de hipótesis no pueden probar la verdad absoluta de una hipótesis, sino que proporcionan **evidencia** para aceptar o rechazar una afirmación sobre el parámetro en estudio [18]. Además, los resultados de las pruebas de hipótesis están sujetos a cierto grado de **incertidumbre** debido a la **variabilidad inherente** de los datos muestrales [18].

### Pruebas de contrastes de hipótesis entre datos continuos y categóricos

Las pruebas de contrastes de hipótesis son herramientas fundamentales en el campo de la estadística inferencial. Estas pruebas nos permiten **evaluar** la **evidencia** en contra o a favor de una afirmación o suposición acerca de un **parámetro poblacional** desconocido. A través de las pruebas de contrastes podemos tomar **decisiones** basadas en la información obtenida de una muestra de datos y determinar si existe suficiente evidencia para rechazar una hipótesis nula en favor de una hipótesis alternativa. Cada prueba asume determinadas **condiciones** en función del número de grupos a comparar. Cuando el número de grupos son **2 ( $k=2$ )** se utilizan determinadas pruebas específicas (Figura 8), mientras que cuando el número de grupos son  **$>2$  ( $k>2$ )** se utilizan otras pruebas específicas (Figura 9).

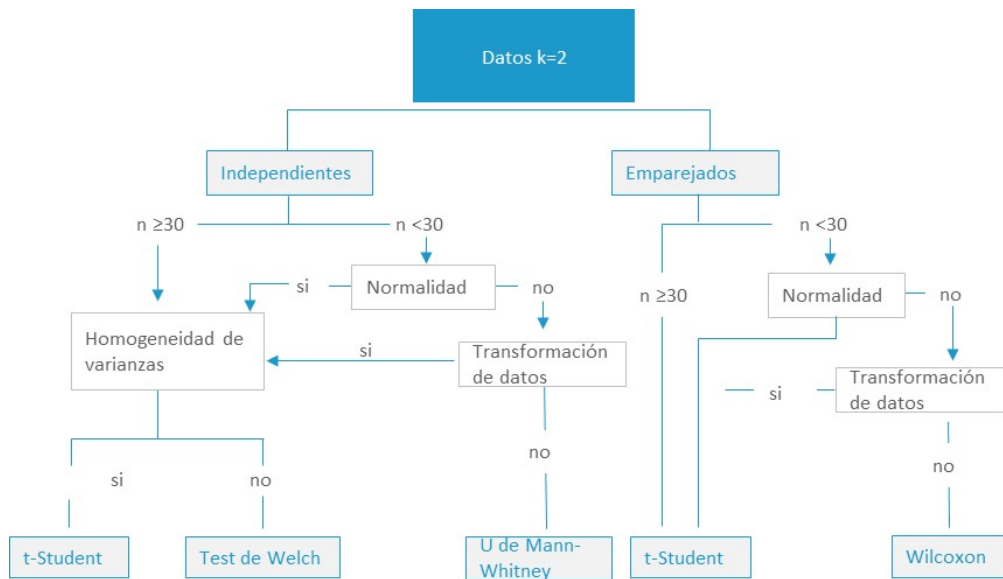


Figura 8. Test de contraste de hipótesis para k=2. Fuente: elaboración propia

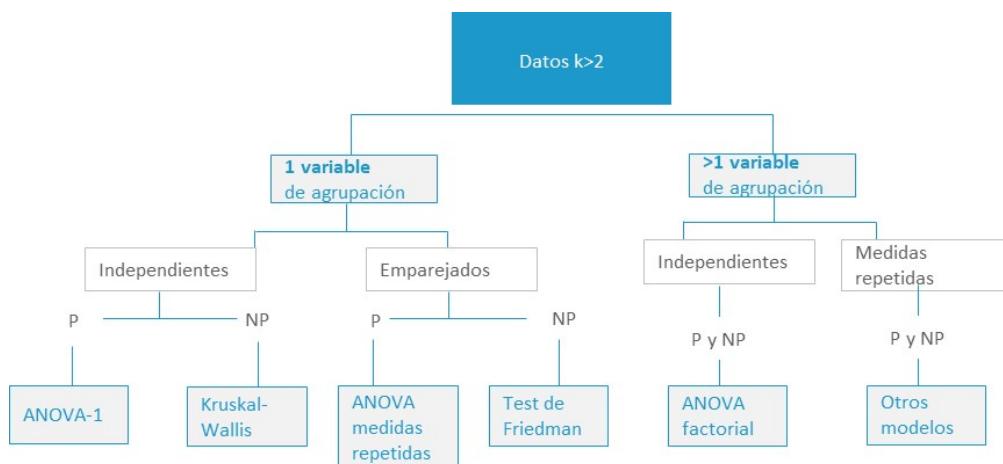


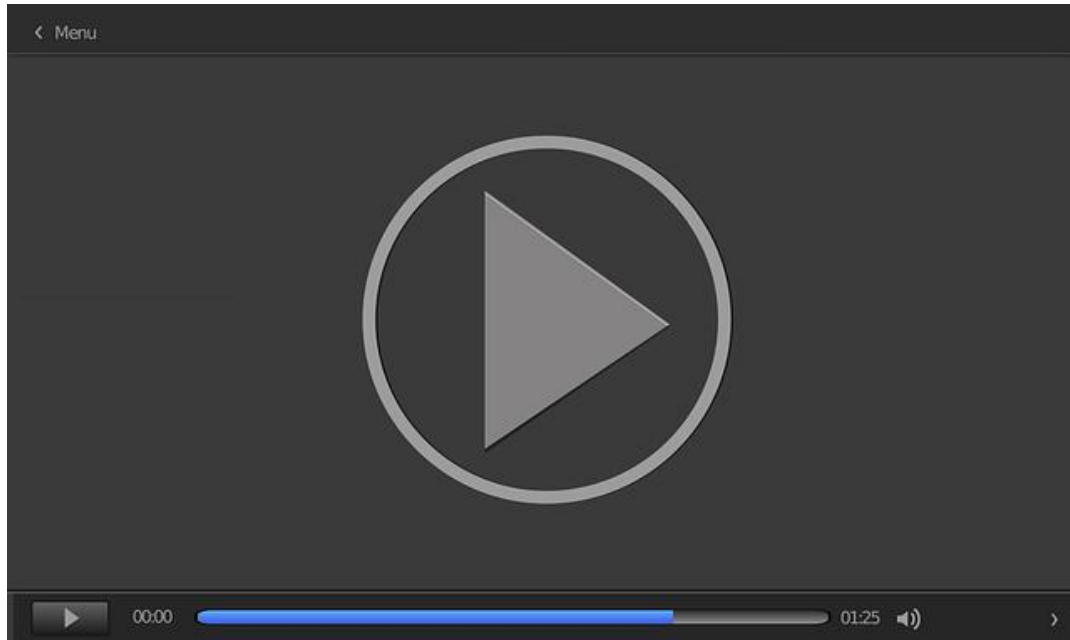
Figura 9. Test de contraste de hipótesis para k>2. Donde las abreviaturas son P: paramétrico y NP: no paramétrico. Fuente: elaboración propia.

La normalidad de los datos es un supuesto importante en muchos contrastes de hipótesis [19]. Para evaluar la **normalidad** de los datos se pueden utilizar diferentes **métodos gráficos y estadísticos**. Uno de los enfoques más comunes es mediante gráficos, como el histograma, el gráfico de probabilidad normal y el gráfico cuantil-cuantil (Q-Q *plot*). Estos gráficos permiten visualizar la **distribución** de los datos y

**compararla** con una distribución normal teórica. Si los datos siguen, aproximadamente, una distribución normal (métodos paramétricos, P) se considera que el supuesto de normalidad está satisfecho.

En cuanto a la homogeneidad de varianzas, se refiere a la **igualdad** de las **varianzas** en **diferentes grupos** o condiciones [20]. Este supuesto es importante en pruebas, como el análisis de varianza (ANOVA) y la prueba t de Student. Para evaluar la homogeneidad de varianzas se puede utilizar el **gráfico de dispersión** (*scatterplot*) y el **gráfico de residuos**. Estos gráficos permiten examinar si la dispersión de los datos es similar en todas las condiciones o grupos. Además, existen pruebas estadísticas específicas, como la prueba de Levene o la prueba de Bartlett, que permiten evaluar formalmente la homogeneidad de varianzas.

A continuación, puedes acceder al vídeo *Contrastes post hoc de 3 o más grupos*.



Contrastes post hoc de 3 o más grupos

---

Accede al vídeo:

<https://unir.cloud.panopto.eu/Panopto/Pages/Embed.aspx?id=5e734924-a928-43f7-9d1a-b08900db97f2>

---

En caso de que los datos no cumplan con el supuesto de normalidad (métodos no paramétricos, NP) o de homogeneidad de varianzas es posible realizar **transformaciones** en los **datos** para lograr que se ajusten a estos supuestos [21]. Algunas de las transformaciones comunes incluyen la **transformación logarítmica**, la **transformación exponencial**, la **transformación de raíz cuadrada**, entre otras. Estas transformaciones se aplican a los datos antes de realizar el contraste de hipótesis, lo que permite cumplir con los supuestos necesarios.

### Pruebas de contrastes de hipótesis entre datos categóricos

Para realizar pruebas de **contrastos de hipótesis** con datos categóricos se pueden utilizar diferentes métodos dependiendo del **tipo de datos** y del **diseño del estudio**. Las pruebas de contraste de hipótesis más comunes son la prueba de chi-cuadrado, exacta de Fisher y la prueba de McNemar.

La **prueba de chi-cuadrado** se utiliza para evaluar la **asociación** entre **dos variables categóricas**. Compara, en una tabla de contingencia, las frecuencias observadas con las frecuencias esperadas bajo la hipótesis nula de independencia. La **prueba exacta de Fisher** se utiliza cuando las **frecuencias esperadas** en una tabla de contingencia son demasiado **bajas** para aplicar la prueba de chi-cuadrado. Es especialmente útil en tablas de 2x2. Por último, la **prueba de McNemar** se utiliza para evaluar la **asociación** entre **dos variables categóricas** en un diseño de datos emparejados. Este modelo compara las diferencias en las proporciones de respuesta entre los niveles de las variables.

## 6.5. Aplicaciones de las pruebas de contrastes de hipótesis

Comencemos primero por saber cuáles son los **comandos** para la aplicación de las **pruebas de normalidad** (Tabla 1), **homogeneidad de varianzas** (Tabla 2), **transformación de datos** (Tabla 3), y pruebas de **contrastos de hipótesis** en función del número de grupos (Tabla 4).

Prueba	Comando	Interpretación <i>p-value</i>
Shapiro-Wilk	<code>shapiro.test(x)</code>	$>0.05$ = normalidad
Anderson-Darling	<code>library(nortest)</code> <code>ad.test(x)</code>	$>0.05$ = normalidad
Kolmogorov-Smirnov	<code>ks.test(x, "pnorm", mean = mean(x), sd = sd(x))</code>	$>0.05$ = normalidad
Gráfico QQ	<code>qqnorm(x)</code>	Los puntos se ajustan a la recta = normalidad
Gráfico de densidad	<code>library(car)</code> <code>densityPlot(x)</code>	Los puntos se ajustan a la recta = normalidad

Tabla 1. Comandos para las pruebas de normalidad. Fuente: elaboración propia.

Prueba	Comando	Interpretación <i>p-value</i>
Levene	<code>library(car)</code> <code>leveneTest(x, y)</code>	$>0.05$ = homogeneidad
F-test	<code>var.test(x1, x2)</code>	$>0.05$ = homogeneidad

Tabla 2. Comandos para la homogeneidad de varianzas. Fuente: elaboración propia.



Transformación	Comando
Logarítmica	<code>x_transformada &lt;- log(x)</code>
Raíz cuadrada	<code>x_transformada &lt;- sqrt(x)</code>
Inversa	<code>x_transformada &lt;- 1/x</code>

Tabla 3. Comandos para la transformación de datos. Fuente: elaboración propia.

Prueba	Nº <i>k</i>	Comando	Interpretación <i>p-value</i>
T-Student de muestra independientes	2	<code>t.test(x,y)</code>	<0.05 = diferencias entre grupos
Test de Welch	2	<code>t.test(x,y,var.equal = FALSE)</code>	<0.05 = diferencias entre grupos
U de Mann-Whitney	2	<code>wilcox.test(x,y)</code>	<0.05 = diferencias entre grupos
T-Student de muestras emparejadas	2	<code>t.test(x, y, paired = TRUE)</code>	<0.05 = diferencias entre grupos
Wilcoxon	2	<code>wilcox.test(x, y, paired = TRUE)</code>	<0.05 = diferencias entre grupos
ANOVA one-way	>2	<code>aov(x,y)</code>	<0.05 = diferencias entre grupos
Kruskal-Wallis	>2	<code>kruskal.test(x,y)</code>	<0.05 = diferencias entre grupos
ANOVA de medias repetidas	>2	<code>library(rstatix)</code> <code>anova_test(data = data, dv = y, wid = id, within = time)</code>	<0.05 = diferencias entre grupos
Test de Friedman	>2	<code>friedman.test(x,y)</code>	<0.05 = diferencias entre grupos
ANOVA factorial	>2	<code>model &lt;- lm(y ~ factor1 * factor2, data = data)</code> <code>anova_result &lt;- anova(model)</code>	<0.05 = interacción entre grupos
Chi-cuadrado	>1x>1	<code>chisq.test(factor1, factor2)</code>	<0.05 = diferencias entre factores
Fisher	>1x>1	<code>fisher.test(factor1, factor2)</code>	<0.05 = diferencias entre factores
McNemar	>1x>1	<code>mcnemar.test(factor1, factor2)</code>	<0.05 = diferencias entre factores

Tabla 4. Comandos para el contraste de hipótesis. Fuente: elaboración propia.

Ahora, veamos un ejemplo de cómo se podrían aplicar algunas de las siguientes pruebas (en color morado se muestra la salida de R, para su reproducibilidad en el programa no copiar esa parte y en subrayador amarillo se resaltan los valores importantes a fijarse en la pantalla de resultados).

**Pruebas de contrastes de hipótesis de datos no paramétricos:  
diferencias significativas de la variable X entre las 2 categorías de la  
variable Y**

```
# Ejemplo 5 de datos no paramétricos
# Generar datos inventados
set.seed(123)
x_group1 <- rnorm(50, mean = 10, sd = 2) # Grupo 1 de la variable x
x_group2 <- rnorm(50, mean = 15, sd = 2) # Grupo 2 de la variable x
y <- c(x_group1, x_group2) # Variable y que combina los grupos

# Prueba de normalidad (Shapiro-Wilk)
shapiro.test(y) # p<0.05 = No paramétrico

Shapiro-Wilk normality test

data: y
W = 0.97383, p-value = 0.04372

# Transformamos datos
y_transformada <- log(y)
shapiro.test(y_transformada) # p<0.05 = No paramétrico

Shapiro-Wilk normality test

data: y_transformada
W = 0.92173, p-value = 1.741e-05

y_transformada <- sqrt(y)
shapiro.test(y_transformada) # p<0.05 = No paramétrico

Shapiro-Wilk normality test

data: y_transformada
W = 0.97172, p-value = 0.03

y_transformada <- 1/y
shapiro.test(y_transformada) # p<0.05 = No paramétrico

Shapiro-Wilk normality test

data: y_transformada
W = 0.92173, p-value = 1.741e-05

# Prueba de U de Mann-Whitney para comparar los grupos
wilcox.test(y ~ rep(c("Grupo 1", "Grupo 2"), each = 50))

Wilcoxon rank sum test with continuity correction

data: y by rep(c("Grupo 1", "Grupo 2"), each = 50)
W = 51, p-value < 2.2e-16
alternative hypothesis: true location shift is not equal to 0
```

Figura 10. Datos no paramétricos. Fuente: elaboración propia.

**Pruebas de contrastes de hipótesis de datos paramétricos: diferencias significativas de la variable X entre las 2 categorías de la variable Y**

```
# Ejemplo 6 de datos paramétricos
# Generar datos inventados
set.seed(123)
x_group1 <- rnorm(1000, mean = 12, sd = 2) # Grupo 1 de la variable x
x_group2 <- rnorm(1000, mean = 15, sd = 2) # Grupo 2 de la variable x
y <- c(x_group1, x_group2) # Variable y que combina los grupos

# Prueba de normalidad (Shapiro-Wilk)
shapiro.test(y) # p>0.05 = Paramétrico

Shapiro-Wilk normality test

data: y
W = 0.99882, p-value = 0.1957

# Prueba de homogeneidad de varianzas (Levene's test)
library(car)
leveneTest(y ~ rep(c("Grupo 1", "Grupo 2"), each = 1000)) # p>0.05 = homogeneidad

Levene's Test for Homogeneity of Variance (center = median)
      Df F value    Pr(>F)
group  1  0.8609 0.3536
      1998

# Prueba de t-Student para comparar los grupos
t.test(y ~ rep(c("Grupo 1", "Grupo 2"), each = 1000))

Welch Two Sample t-test

data: y by rep(c("Grupo 1", "Grupo 2"), each = 1000)
t = -22.933, df = 1997.4, p-value < 2.2e-16
alternative hypothesis: true difference in means between group Grupo 1 and group Grupo 2 is not equal to 0
95 percent confidence interval:
 -2.228213 -1.877137
sample estimates:
mean in group Grupo 1    mean in group Grupo 2
      10.03226           12.08493
```

Figura 11. Datos paramétricos. Fuente: elaboración propia.

## Pruebas de contrastes de hipótesis de datos emparejados

```

# Ejemplo 7 de datos emparejados
# Generar datos inventados
set.seed(123)
x_before <- rnorm(50, mean = 10, sd = 2) # var x antes del trat
x_after <- x_before + rnorm(50, mean = 2, sd = 1) # var x después del trat
y <- x_after - x_before # Diferencia entre las variables x antes y después

# Prueba de normalidad (Shapiro-Wilk)
shapiro.test(y) # p>0.05 = Paramétrico

      Shapiro-Wilk normality test

data:  y
W = 0.99073, p-value = 0.9618

# Prueba de homogeneidad de varianzas (F-test)
var.test(x_before, x_after) # p>0.05 = homogeneidad

      F test to compare two variances

data:  x_before and x_after
F = 0.83056, num df = 49, denom df = 49, p-value = 0.5182
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.4713202 1.4635947
sample estimates:
ratio of variances
 0.8305551

# Prueba t de Student para datos emparejados
t.test(x_before, x_after, paired = TRUE)

      Paired t-test

data:  x_before and x_after
t = -16.762, df = 49, p-value < 2.2e-16
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 -2.403734 -1.889083
sample estimates:
mean difference
 -2.146408

```

Figura 12. Datos emparejados. Fuente: elaboración propia.

## Pruebas de contrastes de hipótesis de datos paramétricos $k > 2$ grupos: ANOVA-*oneway*

```
# Ejemplo 8: ANOVA-oneway
# Generar datos inventados
set.seed(123)
group1 <- rnorm(50, mean = 10, sd = 2) # Grupo 1
group2 <- rnorm(50, mean = 12, sd = 2) # Grupo 2
group3 <- rnorm(50, mean = 15, sd = 2) # Grupo 3

# Combinar los datos en un dataframe
data <- data.frame(
  y = c(group1, group2, group3), # Variable de respuesta
  group = rep(c("Grupo 1", "Grupo 2", "Grupo 3"), each = 50)
)

# Prueba de normalidad (Shapiro-Wilk)
shapiro.test(data$y) #  $p > 0.05$  = Paramétrico

      Shapiro-Wilk normality test

data: data$y
W = 0.99476, p-value = 0.8694

# Realizar el ANOVA-oneway opción A
modelaov <- aov(y ~ group, data = data)
summary(modelaov) #  $< 0.05$  = diferencias entre grupos

      Df Sum Sq Mean Sq F value Pr(>F)
group      2  489.2   244.58   69.07 <2e-16 ***
Residuals 147   520.5     3.54
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Realizar el ANOVA-oneway opción B
model <- lm(y ~ group, data = data)
anova_result <- anova(model)
anova_result #  $p < 0.05$  = diferencias entre grupos

      Df Sum Sq Mean Sq F value Pr(>F)
group      2  489.2   244.58   69.07 <2e-16 ***
Residuals 147   520.5     3.54
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 13. ANOVA-*oneway*. Fuente: elaboración propia.

## Pruebas de contrastes de hipótesis de datos paramétricos $k > 2$ grupos: ANOVA-*oneway*

```
# Ejemplo 8: ANOVA-oneway
# Generar datos inventados
set.seed(123)
group1 <- rnorm(50, mean = 10, sd = 2) # Grupo 1
group2 <- rnorm(50, mean = 12, sd = 2) # Grupo 2
group3 <- rnorm(50, mean = 15, sd = 2) # Grupo 3

# Combinar los datos en un dataframe
data <- data.frame(
  y = c(group1, group2, group3), # Variable de respuesta
  group = rep(c("Grupo 1", "Grupo 2", "Grupo 3"), each = 50)
)

# Prueba de normalidad (Shapiro-Wilk)
shapiro.test(data$y) #  $p > 0.05$  = Paramétrico

      Shapiro-Wilk normality test

data: data$y
W = 0.99476, p-value = 0.8694

# Realizar el ANOVA-oneway opción A
modelaov <- aov(y ~ group, data = data)
summary(modelaov) #  $< 0.05$  = diferencias entre grupos

      Df Sum Sq Mean Sq F value Pr(>F)
group      2  489.2   244.58   69.07 <2e-16 ***
Residuals 147   520.5     3.54
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# Realizar el ANOVA-oneway opción B
model <- lm(y ~ group, data = data)
anova_result <- anova(model)
anova_result #  $p < 0.05$  = diferencias entre grupos

      Df Sum Sq Mean Sq F value Pr(>F)
group      2  489.2   244.58   69.07 <2e-16 ***
Residuals 147   520.5     3.54
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Figura 14. Kruskal-Wallis. Fuente: elaboración propia.

## Pruebas de contrastes de hipótesis de datos paramétricos $k > 2$ grupos: ANOVA-factorial

```
# Ejemplo 10: ANOVA factorial
# Generar datos inventados
set.seed(123)
group1 <- rnorm(20, mean = 10, sd = 2) # Grupo 1
group2 <- rnorm(20, mean = 12, sd = 2) # Grupo 2
group3 <- rnorm(20, mean = 15, sd = 2) # Grupo 3

# Generar factores
factor1 <- rep(c("A", "B"), each = 10) # Factor 1: A y B
factor2 <- rep(c("X", "Y"), times = 10) # Factor 2: X y Y

# Combinar los datos en un dataframe
data <- data.frame(
  y = c(group1, group2, group3), # Variable de respuesta
  factor1 = factor(factor1), # Factor 1
  factor2 = factor(factor2) # Factor 2
)

# Realizar el ANOVA factorial
model <- lm(y ~ factor1 * factor2, data = data)
anova_result <- anova(model)
anova_result # p>0.05 = No interacción
```

Analysis of Variance Table

Response: y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
factor1	1	8.23	8.2288	1.0523	0.3094
factor2	1	0.41	0.4114	0.0526	0.8194
factor1:factor2	1	0.01	0.0066	0.0008	0.9770
Residuals	56	437.93	7.8201		

Figura 15. ANOVA factorial. Fuente: elaboración propia.



## Pruebas de contrastes de hipótesis de datos categóricos

```
# Ejemplo 11: Chi-cuadrado
# Generar datos inventados
set.seed(123)
group1 <- sample(c("A", "B", "C"), size = 100, replace = TRUE) # Grupo 1
group2 <- sample(c("X", "Y", "Z"), size = 100, replace = TRUE) # Grupo 2

# Combinar los datos en un dataframe
data <- data.frame(
  group1 = factor(group1), # Variable categórica Grupo 1
  group2 = factor(group2) # Variable categórica Grupo 2
)

# Crear tabla de contingencia
table <- table(data$group1, data$group2)
table
  X Y Z
A 10 12 11
B 11 8 13
C 9 18 8

# Realizar la prueba de Chi-cuadrado
chi2_result <- chisq.test(table)
chi2_result # p>0.05 = No hay diferencias

      Pearson's Chi-squared test

data:  table
X-squared = 5.2033, df = 4, p-value = 0.2671
```

Figura 16. Chi-cuadrado. Fuente: elaboración propia.

## 6.6. Referencias bibliográficas

1. Upton G y Cook I. A Dictionary of Statistics. OUP Oxford; 2008.
2. Hazra A y Gogtay N. Biostatistics series module 1: Basics of biostatistics. Indian J Dermatol. 2016 en.; 61(1): 10.
3. Stuart A y Ord JK. Kendall's Advanced Theory of Statistics, Volume 1, Distribution Theory. Sexta edición. Wiley; 2010.
4. Stark H y Woods JW. Probability and Random Processes with Applications to Signal Processing. Tercera edición. Pearson; 2002.
5. Leon-García A. Probability, statistics and random processes for electrical engineering. Pearson; 2008.
6. Dekking FM, Kraaikamp C, Lopuhaä HP, Meester LE. A Modern Introduction to Probability and Statistics. Londres: Springer London; 2005.
7. Study. Study [Internet]. s. f. Marginal & Conditional Probability Distributions: Definition & Examples [citado 2023 sept. 13]. Disponible en: <https://study.com/academy/lesson/marginal-conditional-probability-distributions-definition-examples.html>
8. Lukacs E. A Characterization of the Normal Distribution. Ann. Math. Statist. 1942 mzo.; 13(1): 91–3.
9. Fischer H. A history of the central limit theorem: from classical to modern probability theory. Buchwald JZ, Berggren JL y Lutzen J. Springer; 2012.
10. Grinstead CM y Snell JL. Central Limit Theorem: Introduction to Probability. Segunda edición. AMS Bookstore; 1997.

11. Agresti A y Coull BA. Approximate Is Better than “Exact” for Interval Estimation of Binomial Proportions. *Am Stat.* 1998 may.; 52(2): 119.
12. Agresti A. *Categorical Data Analysis*. Wiley; 2002.
13. Casella, G y Berger RL. *Statistical Inference*. Segunda edición. Duxbury Press; 2002.
14. Altman DG. Confidence intervals for the number needed to treat. *BMJ.* 1998 nov. 7; 317(7168): 1309–12.
15. Armitage P; Berry G y Matthews JNS. *Statistical Methods in Medical Research*. Cuarta edición. Blackwell Science; 2002.
16. Hulley SB; Cummings SR; Browne WS; Grady DG y Newman TB. *Designing Clinical Research*. Cuarta edición. Lippincott Williams & Wilkins; 2013.
17. Zar JH. *Biostatistical Analysis*. Quinta edición. Pearson; 2010.
18. Sullivan LM y Dukes KA. *Introduction to Biostatistics*. Segunda edición. Wiley; 2012.
19. Altman DG. *Practical Statistics for Medical Research*. Chapman and Hall/CRC; 1991.
20. Montgomery DC, Peck EA y Vining GG. *Introduction to Linear Regression Analysis* Quinta edición. Wiley; 2012.
21. Conover WJ. *Practical Nonparametric Statistics*. Tercera edición. Wiley; 1999.

## ¿Cuál es la forma correcta de aplicar la prueba de comparación múltiple?

Lee S, Lee DK. What is the proper way to apply the multiple comparison test? Korean J Anesthesiol. 2018 oct.; 71(5): 353-360. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/30157585/>

Las comparaciones múltiples son necesarias para controlar el error tipo I y explorar diferentes grupos o condiciones, en algunos estudios es necesario comparar múltiples grupos o condiciones para analizar diferencias o similitudes entre ellos. Al realizar comparaciones múltiples, se pueden identificar patrones o relaciones más complejas que no serían evidentes al analizar cada comparación de forma aislada. También, maximizar la información, es decir, realizar múltiples comparaciones, permite analizar diferentes aspectos de un problema o fenómeno de interés. Esto puede proporcionar una visión más completa y detallada de los datos, lo cual es importante para una comprensión más precisa y exhaustiva de los resultados

## Análisis post-hoc ANOVA

Amat-Rodríguez J. Rpubs [Internet]. ANOVA análisis de varianza para comparar múltiples medias [citado 2023 sept. 13]. Disponible en: [https://rpubs.com/Joaquin\\_AR/219148](https://rpubs.com/Joaquin_AR/219148)

El análisis *post hoc* se refiere a un tipo de análisis estadístico que se realiza después de realizar una prueba de significancia global, como un análisis de varianza (ANOVA) o una prueba de efecto principal, y, así, realizar comparaciones específicas entre los grupos o condiciones involucradas en el estudio. Cuando se realiza una prueba global y se encuentra una diferencia significativa, el análisis *post hoc* ayuda a identificar qué grupos o condiciones específicas difieren entre sí. Este análisis permite realizar comparaciones múltiples y determinar qué diferencias son estadísticamente significativas.

## Ejemplos de contrastes de hipótesis con R Markdown

Alberich R. Rstudio [Internet]. 2014 oct. 12. Ejemplos de contrastes de hipótesis con R Markdown [citado 2023 sept. 23]. Disponible en: [https://rstudio-pubs-static.s3.amazonaws.com/65042\\_a1784120e81a430f9de400ed9b899b0b.html](https://rstudio-pubs-static.s3.amazonaws.com/65042_a1784120e81a430f9de400ed9b899b0b.html)

En esta página se pueden ver otros ejemplos de cómo calcular los contrastes de hipótesis con el programa R. Existen múltiples comandos para poder realizar contrastes de hipótesis, aunque los más utilizados son los vistos en este tema.

1. ¿Cuál de las siguientes afirmaciones describe mejor la probabilidad en el contexto de la bioestadística?
  - A. Es la certeza absoluta de que un evento o resultado ocurrirá.
  - B. Es la frecuencia con la que ocurre un evento o resultado en una muestra de datos.
  - C. Es la posibilidad de que un evento o resultado ocurra en una población.
  - D. Es la variabilidad aleatoria de los datos recopilados en una muestra.
  
2. ¿Cuál de las siguientes opciones describe mejor la estimación mediante intervalos de confianza?
  - A. Es una técnica para probar hipótesis acerca de los parámetros de una población.
  - B. Es un método para estimar la probabilidad de que un evento ocurra en una muestra de datos.
  - C. Es una manera de resumir los datos recopilados en una muestra utilizando medidas de tendencia central.
  - D. Es un enfoque para estimar un parámetro desconocido y proporcionar un rango de valores plausibles.
  
3. ¿Cuál de las siguientes afirmaciones describe mejor las pruebas de hipótesis?
  - A. Son técnicas utilizadas para estimar un parámetro desconocido en una población.
  - B. Son métodos para resumir los datos recopilados en una muestra utilizando medidas de dispersión.
  - C. Son enfoques para comparar las diferencias entre las medias de dos o más grupos.
  - D. Son procedimientos para evaluar la evidencia en contra de una afirmación acerca de un parámetro poblacional.

4. ¿Cuál de las siguientes opciones describe mejor una aplicación de las pruebas de contrastes de hipótesis?

- A. Determinar el intervalo de confianza para una estimación puntual del parámetro de una población.
- B. Comparar las diferencias entre las medias de dos grupos independientes.
- C. Calcular la varianza de una muestra para estimar la dispersión de los datos.
- D. Realizar una regresión lineal para predecir una variable de respuesta.

5. En una prueba de hipótesis, ¿cuál es la hipótesis nula ( $H_0$ )?

- A. Es la hipótesis que se desea aceptar o probar.
- B. Es la hipótesis que se desea rechazar o refutar.
- C. Es la hipótesis que asume la existencia de una diferencia o efecto.
- D. Es la hipótesis que asume la igualdad o ausencia de una diferencia o efecto.

6. En una prueba de hipótesis, ¿cuál es el nivel de significancia ( $\alpha$ )?

- A. Es la probabilidad de cometer un error de tipo II.
- B. Es la probabilidad de obtener un resultado significativo.
- C. Es la probabilidad de cometer un error de tipo I.
- D. Es la probabilidad de obtener un resultado no significativo.



7. ¿Cuál de las siguientes afirmaciones es cierta acerca del valor  $p$  ( $p$ -value)?
  - A. Es la probabilidad de obtener un resultado no significativo.
  - B. Es la probabilidad de obtener un resultado significativo.
  - C. Es el nivel de significancia utilizado en la prueba de hipótesis.
  - D. Es la probabilidad de obtener un resultado igual o más extremo que el observado, asumiendo que la hipótesis nula es verdadera.
  
8. ¿Cuál de las siguientes pruebas estadísticas se utiliza para comparar las medias de dos grupos independientes cuando los datos siguen una distribución normal?
  - A. Prueba  $t$  de Student para muestras independientes.
  - B. Prueba  $U$  de Mann-Whitney.
  - C. Prueba de Wilcoxon.
  - D. Análisis de varianza (ANOVA).
  
9. ¿Cuál de las siguientes pruebas estadísticas se utiliza para comparar las medianas de dos grupos independientes sin asumir una distribución particular de los datos?
  - A. Prueba  $t$  de Student para muestras independientes.
  - B. Prueba  $U$  de Mann-Whitney.
  - C. Prueba de Wilcoxon.
  - D. Análisis de varianza (ANOVA).
  
10. ¿Cuál de las siguientes pruebas estadísticas se utiliza para comparar las medianas de dos grupos relacionados o pareados?
  - A. Prueba  $t$  de Student para muestras independientes.
  - B. Prueba  $U$  de Mann-Whitney.
  - C. Prueba de Wilcoxon.
  - D. Prueba de McNemar.