

# ESTADÍSTICA APLICADA: Análisis de Datos

VOLUMEN III: AVANZADO

Domina el análisis de datos con un método sencillo y eficaz que te permite obtener resultados rápidos y autonomía.

- 1 Conceptos esenciales simplificados
- 2 Análisis de datos con pasos ilustrados
- 3 El mejor software gratuito para tus análisis

MUESTRA



APRENDER ESTADÍSTICA FÁCIL



## MÁS INFORMACIÓN SOBRE EL VOLUMEN III

La iniciativa del **Volumen III** es única.

Nos permite incorporar temas sugeridos por nuestros seguidores en cada nueva edición.

Y aquellos que ya adquirieron el paquete tendrán acceso vitalicio a estas actualizaciones.

El Volumen III aborda temas más avanzados que los tratados en los dos primeros volúmenes.

Como planeamos actualizar regularmente el Volumen III con nuevos temas, te invitamos a enviarnos tus sugerencias a través de nuestro perfil en Instagram.



**@StatisticsEasily**



**OLVIDA TODO**

**LO QUE HAS**

**APRENDIDO**

**SOBRE**

**ESTADÍSTICA**

**PARA ANALIZAR**

**TUS DATOS**



De ~~49,00~~  
por solo **19,00**  
\*LOS 3 VOLÚMENES INCLUIDOS

**ESTADÍSTICA APLICADA: Análisis de Datos**  
VOLUMEN: ANÁLISIS

Domina el análisis de datos con el método sencillo y fácil que te permitirá obtener resultados rápidos y efectivos.

1. Descripción y análisis de los datos  
2. Análisis de los datos para su interpretación  
3. Análisis de los datos para su interpretación

5 estrellas

 **Aprender Estadística Fácil**  
@LearnStatisticsEasily

## Vas a aprender a analizar tus datos de forma rápida, sencilla e indiscutible.

Metodología de enseñanza SIMPLE que en poco tiempo te permitirá analizar tus datos por tu cuenta.

Abordamos TODAS las etapas y solo lo que realmente es necesario para analizar tus datos.

Partimos del principio, ya validado, de que es perfectamente posible analizar datos de forma correcta sin necesidad de entender conceptos o fórmulas complejas.

Así, este libro te será útil incluso si no tienes ningún conocimiento previo de estadística.

*Los 3 volúmenes están incluidos en esta oferta.*

**¡ADQUIERE EL TUYO HACIENDO CLIC AQUÍ!**



Estadística Aplicada: Análisis de Datos (Volumen III - Avanzado) / Alves, Ana. - Aprender Estadística Fácil, 2025. 133 p. 1ª ed.

1. Estadística. 2. Bioestadística. 3. Aprender Estadística Fácil. 4. Estadística Aplicada. 5. Análisis de Datos. 6. Estadística Inferencial. 7. Estadística Descriptiva. 8. Gráficos. 9. Tamaño Muestral.

© 2025 Aprender Estadística Fácil

Todos los derechos reservados. Esta publicación o cualquier parte de ella no puede ser reproducida ni utilizada sin el permiso expreso del titular de los derechos de autor.



Aprender Estadística Fácil  
[es.statisticseasily.com](https://es.statisticseasily.com)



# PRÓLOGO

Bienvenido a su guía definitiva para desbloquear el poder del análisis de datos.

Este libro presenta un método de enseñanza innovador que le permite analizar sus datos de forma independiente, con rapidez y precisión.

Destilamos lo esencial, proporcionando únicamente la información necesaria para aprender análisis de datos sin perderse en complejidades.

Diga adiós a conceptos intimidantes, fórmulas y tablas. Esta guía ha sido diseñada para beneficiarlo, incluso si su conocimiento en estadística es limitado.

Nuestra innovadora metodología para "*aprender análisis de datos rápidamente, fácilmente, de forma independiente y con confianza*" diferencia este libro de los demás.



# DOMINANDO NUESTRA METODOLOGÍA

**(a)** Destilamos únicamente los conceptos más vitales, haciéndolos fácilmente comprensibles.

**(b)** Ejemplos claros y diagramas dan vida a cada concepto.

**(c)** Nuestro algoritmo para la selección de análisis y gráficos es simple y directo.

**(d)** Abordamos los análisis estadísticos más comunes, cubriendo el 99 % de los escenarios del mundo real.

**(e)** Nuestras instrucciones, con pasos detallados e ilustraciones, hacen que el análisis de datos sea fácilmente comprensible.

**(f)** Experimente lo que hay de más amigable, completo e intuitivo en software estadístico gratuito.





# ÍNDICE

- 1 COMENZANDO:  
CONOCIMIENTO ESENCIAL**  
Comprenda los conceptos clave de manera simplificada y accesible.
- 2 PRINCIPALES SOFTWARES ESTADÍSTICOS GRATUITOS**  
Descubra herramientas para el análisis de datos, gráficos, y cálculos de tamaño muestral.
- 3 ESTADÍSTICA DESCRIPTIVA:  
MEDIDAS RESUMEN**  
Profundice en las medidas más importantes para resumir y presentar sus datos.
- 4 ESTADÍSTICA INFERENCIAL:  
ANÁLISIS DE DATOS**  
Aprenda a elegir el análisis adecuado y aplicarlo con precisión.
- 5 ELIGIENDO EL GRÁFICO PERFECTO (VOL. II)**  
Siga una guía paso a paso para seleccionar y crear el gráfico ideal para sus datos.
- 6 CONTENIDO EXTRA Y TEMAS AVANZADOS (VOL. III)**  
Profundice en consejos adicionales y explore temas un poco más sofisticados.



# ÍNDICE - VOLUMEN III

## 1 CÁLCULO DEL TAMAÑO MUESTRAL

1. ¿Por qué calcular el tamaño de la muestra? .....11
2. Software para calcular el tamaño de la muestra .....14
3. Parámetros clave del G\*Power .....15
4. Cálculo del tamaño de la muestra .....20
5. Comparación de grupos no pareados .....24
6. Comparación de grupos pareados .....41
7. Búsqueda de relación entre variables .....69

## 2 CÓMO REALIZAR UN MUESTREO

1. ¿Por qué usar un método de muestreo adecuado? .....96
2. Conceptos básicos .....97
3. Tipos de muestreo .....100
4. Consideraciones finales sobre el muestreo .....104
5. Cómo hacerlo en la práctica .....107

## 3 TEMAS AVANZADOS EXTRA

1. Valor de p: el regreso .....112
2. Tamaño del efecto .....118
3. Intervalo de confianza .....122
4. Potencia del test .....130

## 4 ANÁLISIS INFERENCIAL AVANZADO

1. Modelo lineal generalizado .....próximamente



# ÍNDICE - VOLUME III

## CAPÍTULO 1: CÁLCULO DEL TAMAÑO MUESTRAL

### **BUSCANDO DIFERENCIAS**

#### **ENTRE GRUPOS NO PAREADOS**

1. Prueba t - muestras independientes (U de Mann-Whitney) .25
2. ANOVA de un factor (Prueba H de Kruskal-Wallis) .....30
3. ANOVA de dos factores (factorial) .....35

### **BUSCANDO DIFERENCIAS**

#### **ENTRE GRUPOS PAREADOS**

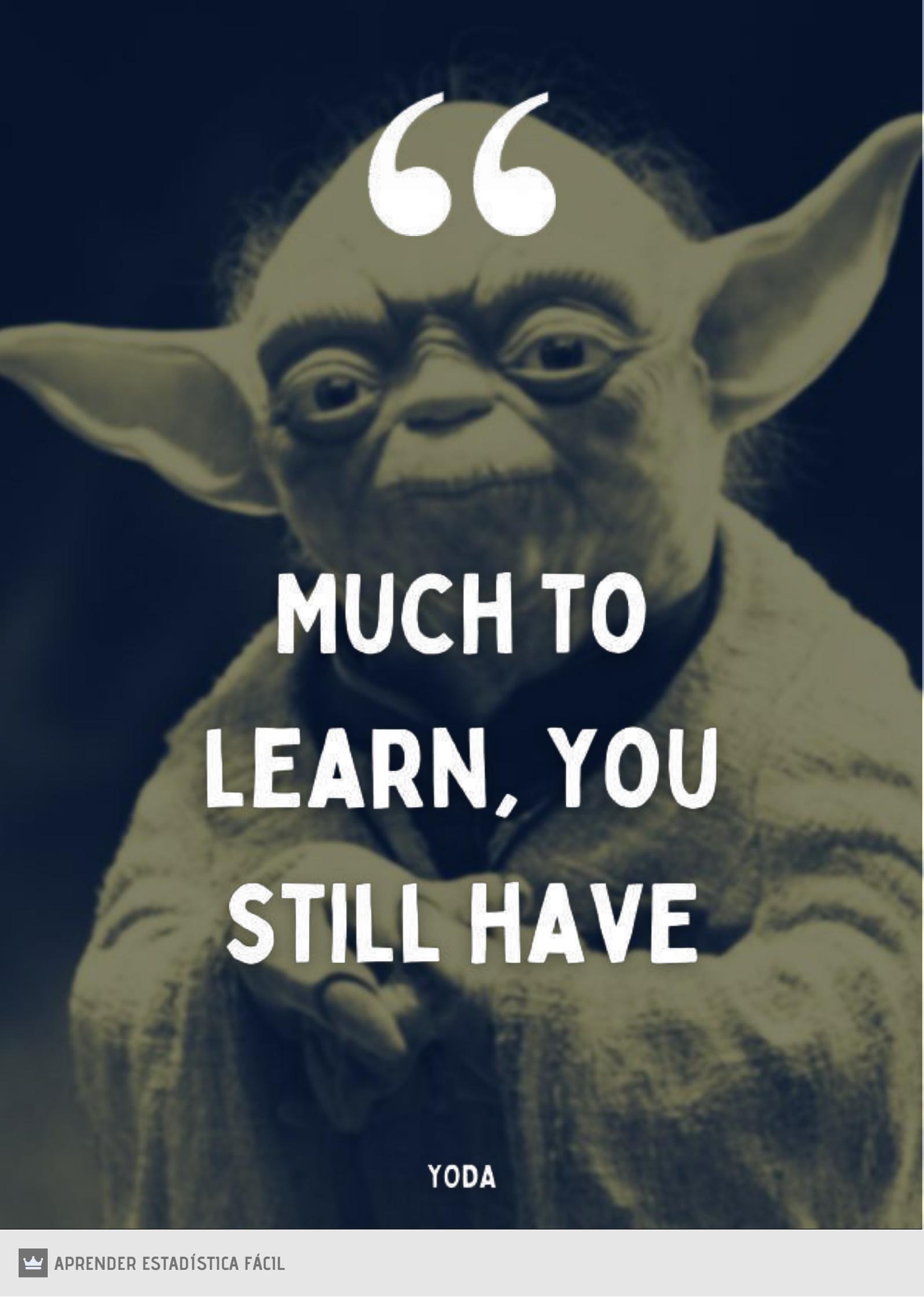
4. Prueba t para muestras pareadas (Prueba de Wilcoxon) ....42
5. ANOVA - un factor con medidas repetidas (Friedman) .....48
6. ANOVA de dos factores con medidas repetidas .....55
7. ANOVA de efecto mixto .....62

### **BUSCANDO RELACIONES**

#### **ENTRE VARIABLES**

8. Correlación de Pearson (Kendall & Spearman) .....70
9. Regresión lineal simple .....74
10. Regresión lineal múltiple .....78
11. Regresión logística binaria simple .....84
12. Regresión logística binaria múltiple .....91





“

**MUCH TO  
LEARN, YOU  
STILL HAVE**

**YODA**



aprender estadística fácil



**CAPÍTULO 1**

# **CÁLCULO DEL TAMAÑO MUESTRAL**

# 1. ¿POR QUÉ CALCULAR EL TAMAÑO MUESTRAL?

Al realizar un **muestreo**, recolectamos datos solo de una parte de los elementos que componen la población.

Esto se debe a que recolectar datos de todos los elementos de la población sería **demorado** y **costoso**.

Además, recolectar todos estos datos sería **innecesario** si seguimos las etapas descritas aquí.

## ¿Por qué es importante definir un tamaño de muestra adecuado?

1

Un tamaño muestral adecuado reduce el error de muestreo.

2

Muestras excesivamente grandes resultan en desperdicio de tiempo y dinero.

3

Muestras excesivamente pequeñas producen resultados no confiables.

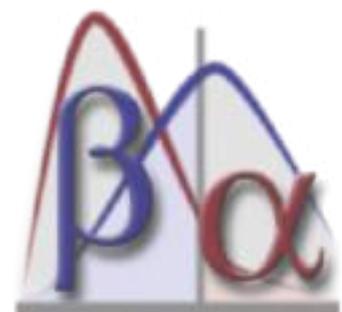
## 2. SOFTWARE PARA EL CÁLCULO DEL TAMAÑO DE LA MUESTRA

**G\*Power:** Statistical Power Analyses es una herramienta estadística poderosa y gratuita para el cálculo del tamaño de la muestra.

Además de otras funciones, como las relacionadas con el poder de la prueba y el tamaño del efecto, G\*Power es el software ideal para calcular el **tamaño de la muestra** de las pruebas que abordamos en este libro, incluyendo pruebas t, ANOVA, regresión lineal, entre otras.

Puedes **descargar** G\*Power en el [site oficial](#).

En el sitio, desplázate hacia abajo hasta encontrar la sección de **Download**, elige la versión deseada, ya sea Windows o macOS, y descarga e instala G\*Power.



Cada cálculo del tamaño muestral en G\*Power puede dividirse en tres pasos simples:

1. Selección de la prueba adecuada
2. Inserción de los parámetros
3. Estimación del tamaño del efecto

**t tests** – Means: Difference between two independent means (two groups)

**Analysis:** A priori: Compute required sample size

**Input:**

Tail(s)	= One
Effect size d	= 2.0000000
$\alpha$ err prob	= 0.05
Power (1- $\beta$ err prob)	= 0.95
Allocation ratio N2/N1	= 1

**Output:**

Noncentrality parameter $\delta$	= 3.7416574
Critical t	= 1.7822876
Df	= 12
Sample size group 1	= 7
Sample size group 2	= 7
Total sample size	= 14
Actual power	= 0.9695815

Test family: t tests

Statistical test: Means: Difference between two independent means (two groups)

Type of power analysis: A priori: Compute required sample size – given  $\alpha$ , power, and effect size

**Input Parameters**

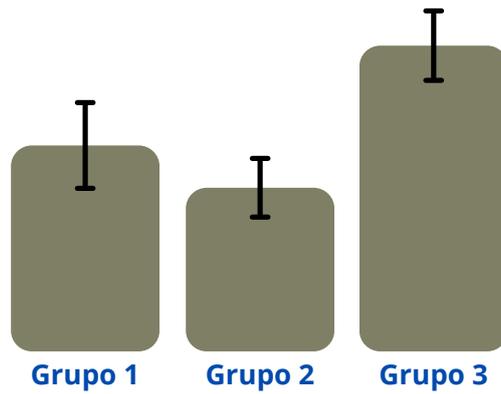
Tail(s)	One
Determine => 3 Effect size d	2.0000000
$\alpha$ err prob	0.05
Power (1- $\beta$ err prob)	0.95
Allocation ratio N2/N1	1

**Output Parameters**

Noncentrality parameter $\delta$	3.7416574
Critical t	1.7822876
Df	12
Sample size group 1	7
Sample size group 2	7
Total sample size	14
Actual power	0.9695815

X-Y plot for a range of values

Calculate



## CALCULANDO EL TAMAÑO MUESTRAL

¿ESTOS GRUPOS INDEPENDIENTES SON DIFERENTES?

5

# COMPARANDO GRUPOS INDEPENDIENTES

5.1 PRUEBA T PARA MUESTRAS INDEPENDIENTES

5.2 ANOVA DE UN FACTOR

5.3 ANOVA DE DOS FACTORES (FACTORIAL)



## 5.1 PRUEBA T PARA MUESTRAS INDEPENDIENTES

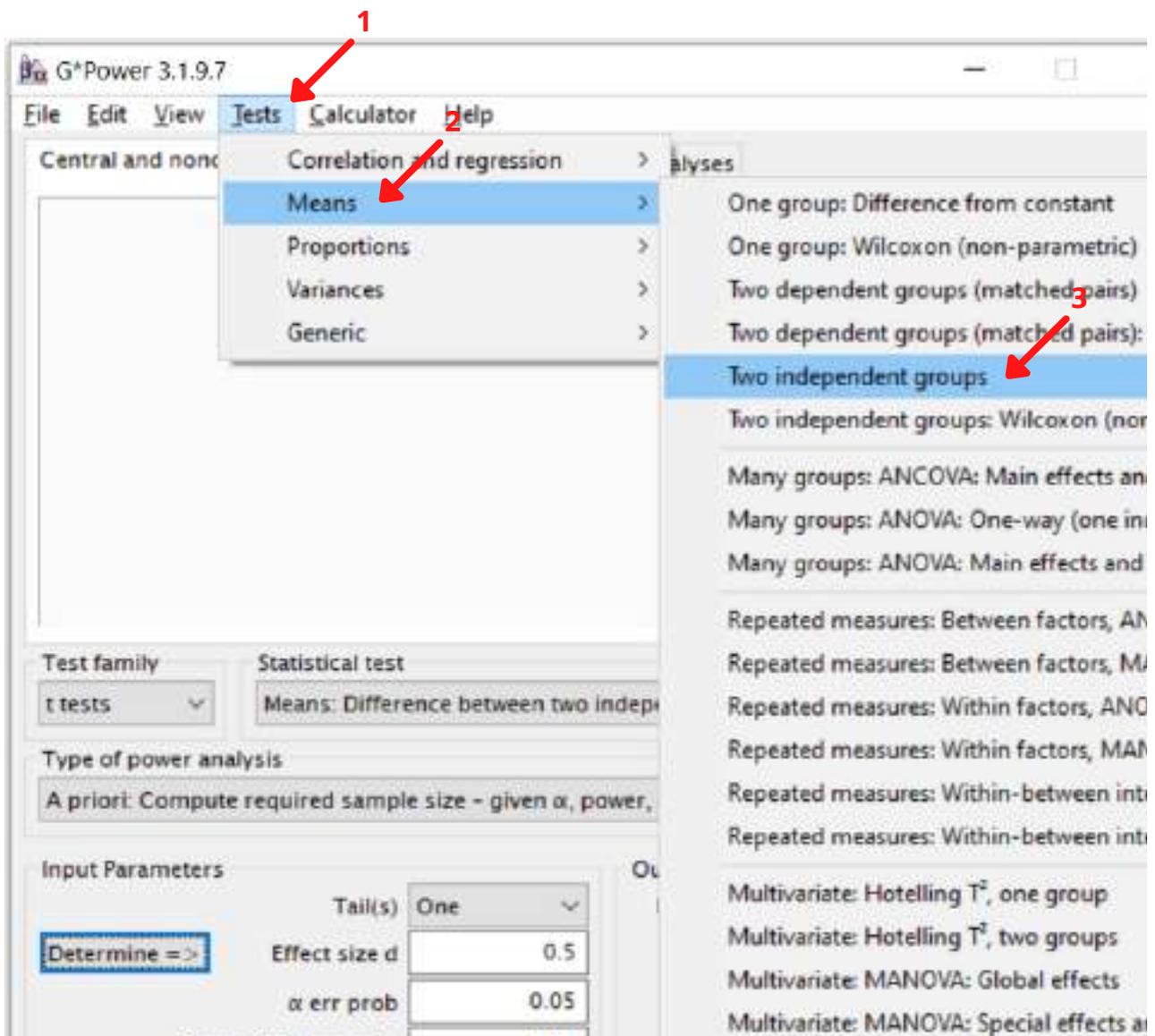
### Ejemplo:

Probar si existe una diferencia en la altura entre individuos de sexo masculino y femenino en una tribu indígena.

### Cálculo del Tamaño Muestral:

A partir de un estudio piloto, obtuvimos 4 observaciones de cada grupo y estimamos su media y desviación estándar.

1. Al abrir G\*Power, haga clic en **Tests: Means: Two independent groups**.



- 2. Ahora completaremos los Parámetros de Entrada (**Input Parameters**) de acuerdo con la siguiente información:

Input Parameters		Output Parameters
Tail(s)	Two	Noncentr
Effect size d		
α err prob	0.05	
Power (1 - β err prob)	0.95	Sar
Allocation ratio N2/N1	1	Sar

- **Tail(s) = Cola(s):** Marque **One** (Una) si la prueba es unicaudal o **Two** (Dos) si es bicaudal. Pero, ¿cuándo se debe usar una prueba unicaudal o bicaudal?

Si la hipótesis alternativa ( $H_1$ ) es específica, por ejemplo, "la media del grupo 1 es mayor que la media del grupo 2", usamos la prueba unicaudal. Si la hipótesis alternativa ( $H_1$ ) es general — sin una distinción inicial de mayor o menor — por ejemplo, "las medias son diferentes entre los grupos", usamos la prueba bicaudal.

Esta hipótesis debe basarse en el conocimiento existente en el campo de estudio. Si no está seguro, mantener la opción **Two** (bicaudal) es la mejor decisión.

- **α err prob (nivel de significación):** La probabilidad de rechazar la hipótesis nula cuando es verdadera (Error Tipo I).

Generalmente, los valores utilizados son **0.05** o **0.01**. Un nivel de significación de 0.05, por ejemplo, indica un riesgo del 5% de concluir que hay una diferencia cuando en realidad no la hay.

- **Power ( $1 - \beta$  err prob) (potencia de la prueba):** La probabilidad de rechazar la hipótesis nula si es falsa, es decir, cuánto la prueba controla el Error Tipo II.

Generalmente, un valor aceptable está entre **0.80** y **0.99**.

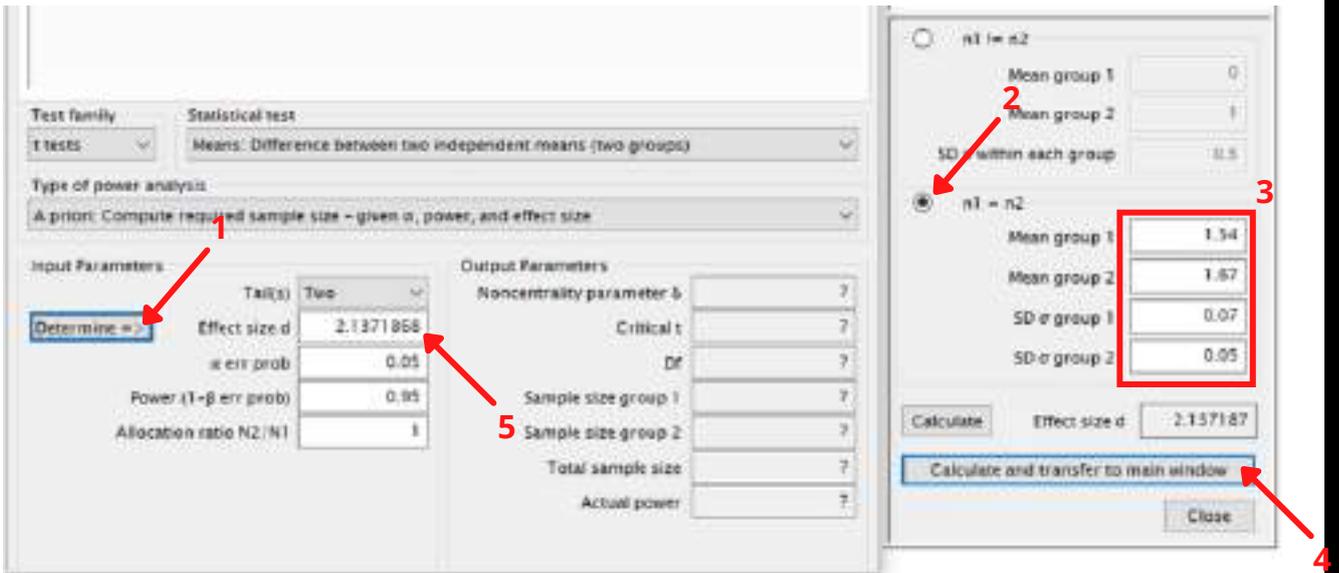
Cuanto mayor sea la potencia de la prueba, mejor, pero a medida que aumenta, el tamaño de la muestra también crece.

- **Allocation ratio N2/N1 (Proporción de asignación entre los grupos N2/N1):** Si desea que el tamaño de la muestra de ambos grupos sea equivalente, ingrese el valor **1** (uno). Introduzca el valor requerido si desea tener otra proporción de asignación entre los grupos.

- **3. Effect size d (Tamaño del efecto d de Cohen):** Representa la magnitud de la diferencia detectada entre los grupos analizados. Convencionalmente,  $d = 0.20$  se considera un efecto pequeño,  $d = 0.50$  se considera un efecto mediano y  $d = 0.80$  se considera un efecto grande.

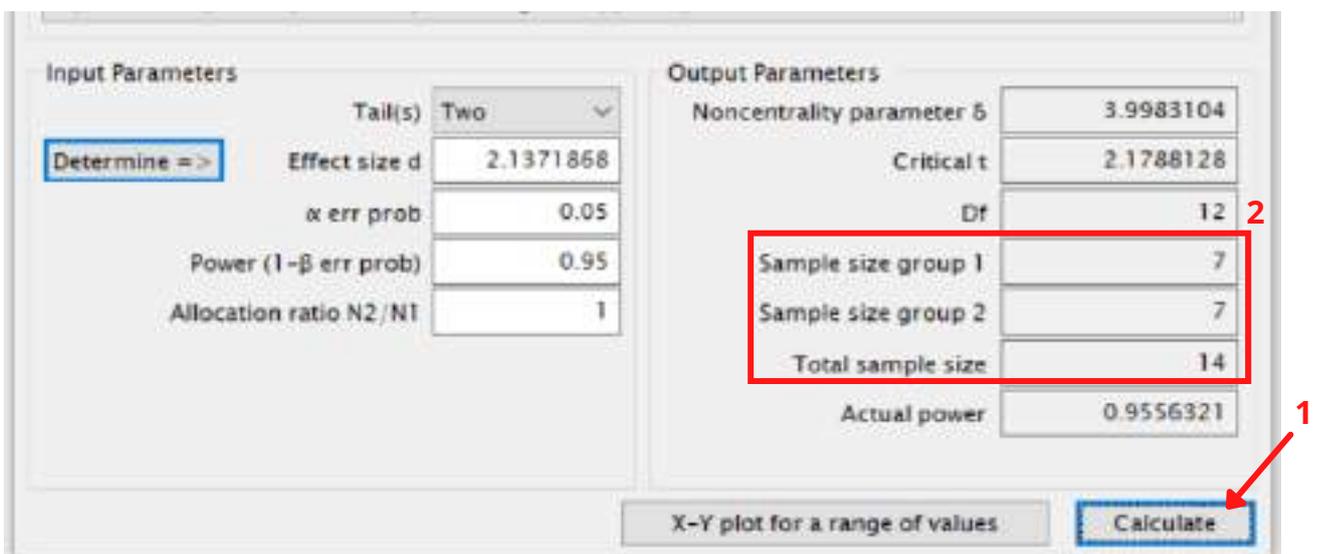
Para estimar  $d$  en nuestro cálculo del tamaño de muestra, haga clic en **Determine =>** (Determinar =>), y se abrirá una nueva pestaña. En esta nueva sección, debemos ingresar valores para **Mean** (Media) y **SD  $\sigma$**  (Desviación estándar  $\sigma$ ), que obtuvimos del estudio piloto para cada uno de los dos grupos. Estos valores resumen pueden calcularse fácilmente según las instrucciones en **Volumen I: Capítulo 3: Tópico 5**.

Después de completar la información, haga clic en **Calculate and transfer to the main window**. (Calcular y transferir a la ventana principal). El valor de  $d$  será calculado y el campo **Effect size d** será llenado automáticamente.



4. Con todos los **Input Parameters** (Parámetros de Entrada) completados, debemos hacer clic en **Calculate** (Calcular) en la esquina inferior derecha para obtener el número de muestras a recolectar.

En este caso, el tamaño total de la muestra calculado fue de 14 elementos, 7 en el grupo 1 y 7 en el grupo 2, que debemos obtener aleatoriamente de la población. Así, necesitaremos este tamaño de muestra para detectar este tamaño de efecto; y, suponiendo que sea correcto, con una probabilidad del 5% de cometer un error Tipo I y 95% de probabilidad de no cometer un error Tipo II.



# NOTAS

## 1. ¿Cómo podemos obtener los valores de la media y la desviación estándar de los dos grupos para estimar el tamaño del efecto?

Comúnmente, estos valores se **estiman** de dos maneras:

**(1)** A través de una recolección piloto con algunas muestras o

**(2)** A través de otros estudios con poblaciones, especies o condiciones similares.

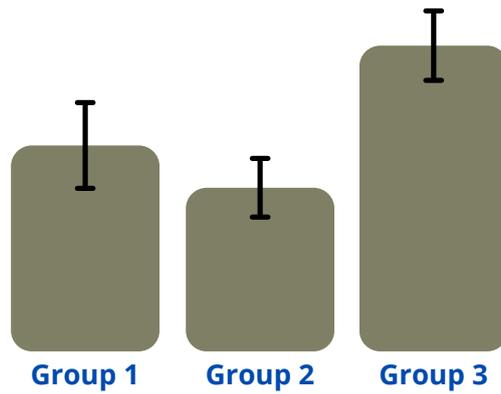
A veces, podemos tener razones para **suponer** directamente el tamaño del efecto esperado sin estimar los estadísticos resumen de la media y la desviación estándar.

## 2. Para calcular el tamaño de la muestra para la prueba U de Mann-Whitney, equivalente no paramétrica a la prueba t:

Haga clic en **Tests: Means: Two independent groups: Wilcoxon (non-parametric)** en el primer paso.

Mantenga la opción **Parent distribution** (Distribución de origen) como **Normal**.

El resto de los pasos es el mismo.



## CALCULANDO EL TAMAÑO MUESTRAL

¿SON DIFERENTES ESTOS GRUPOS PAREADOS?



# COMPARANDO GRUPOS PAREADOS

**6.1 PRUEBA T PARA MUESTRAS PAREADAS**

**6.2 ANOVA DE UN FACTOR CON MEDIDAS REPETIDAS**

**6.3 ANOVA DE DOS FACTORES CON MEDIDAS REPETIDAS**

**6.4 ANOVA DE EFECTOS MIXTOS**



## 6.1 PRUEBA T PARA MUESTRAS PAREADAS

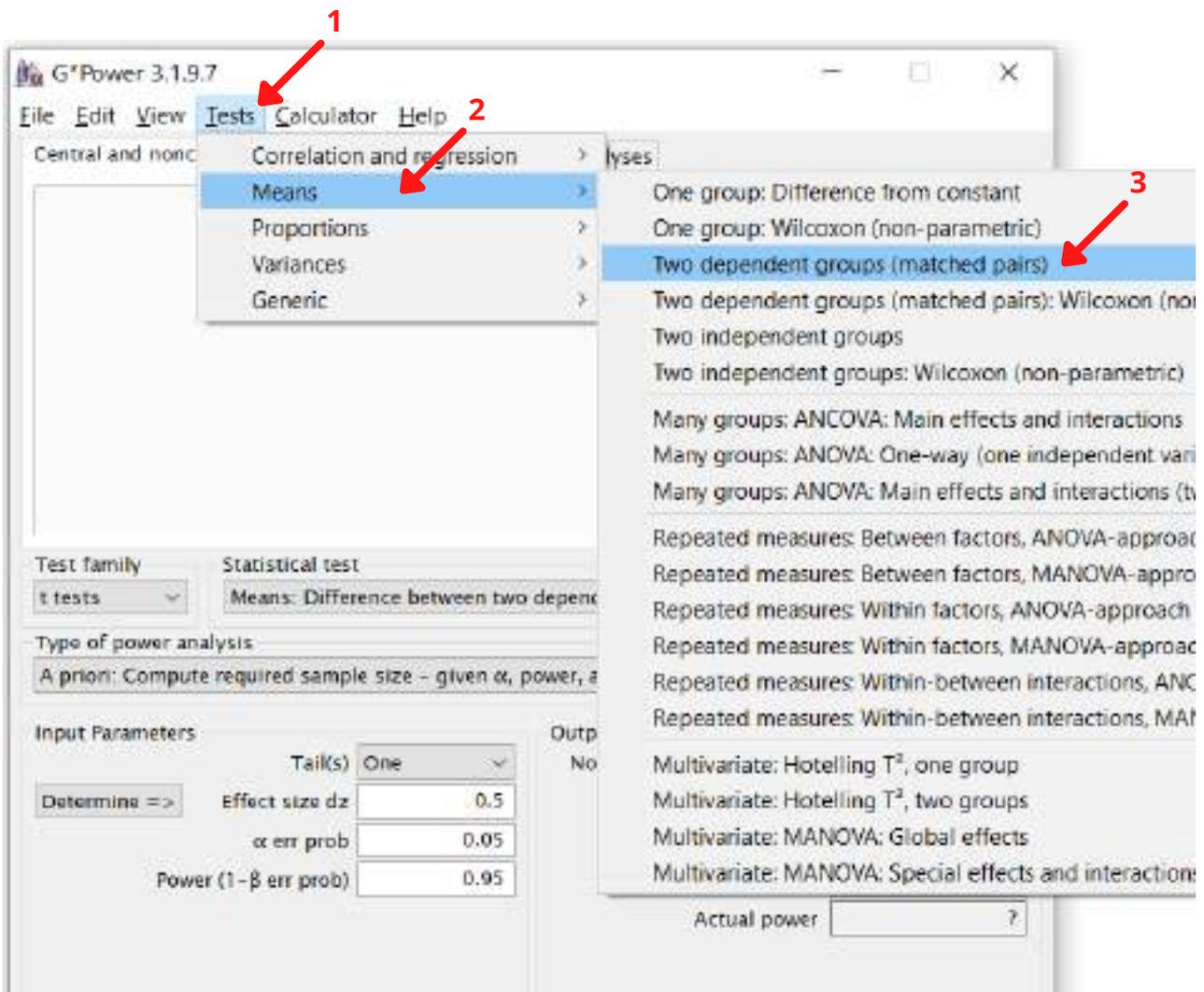
### Ejemplo:

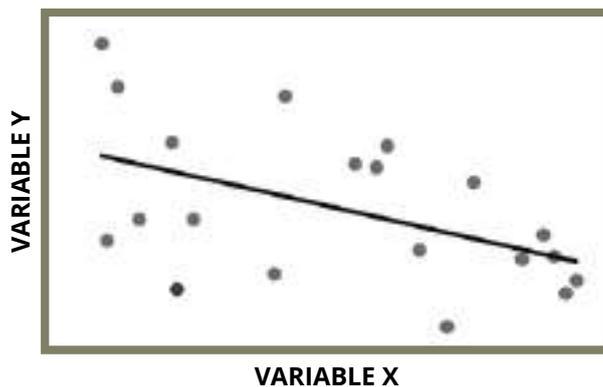
¿El uso del medicamento afecta la presión arterial (PA) de mujeres ancianas? Se medirá la PA dos veces: una antes y otra después de la intervención (medicación).

### Cálculo del Tamaño Muestral:

En un estudio piloto, medimos la PA de 4 mujeres antes y después de la intervención. Luego, calculamos medias, desviaciones estándar y el coeficiente de correlación.

1. Para comenzar, abra G\*Power y haga clic en **Tests: Means: Two dependent groups (matched pairs)**.





## CALCULANDO EL TAMAÑO MUESTRAL

¿ESTAS VARIABLES ESTÁN RELACIONADAS?

7

# BUSCANDO RELACIONES ENTRE VARIABLES

## 7.1 CORRELACIÓN DE PEARSON

### REGRESIÓN LINEAL

## 7.2 SIMPLE Y 7.3 MÚLTIPLE

### REGRESIÓN LOGÍSTICA

## 7.4 SIMPLE Y 7.5 MÚLTIPLE



## 7.1 CORRELACIÓN DE PEARSON

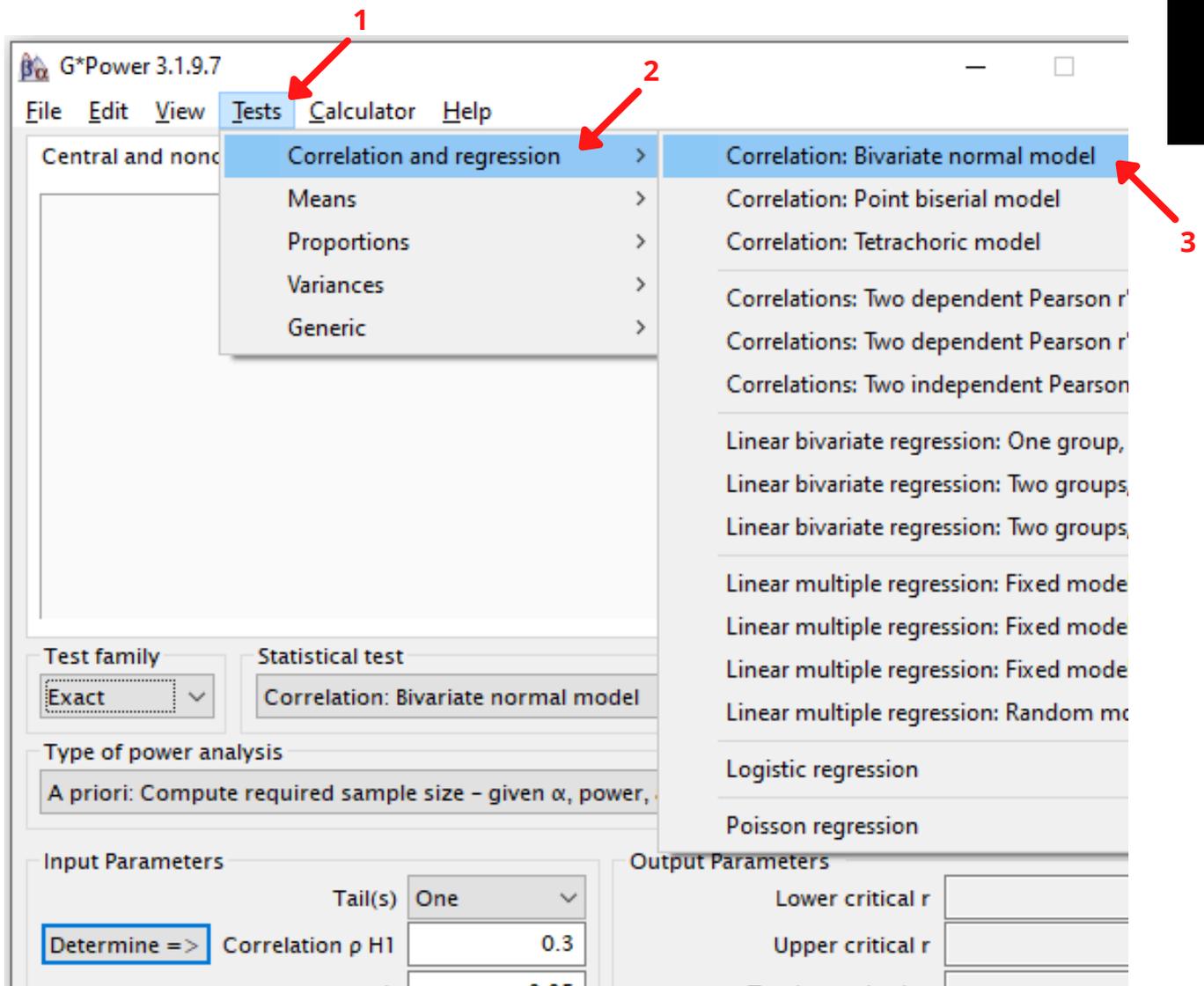
### Ejemplo:

Comprobar si existe una correlación entre el peso y la altura de los miembros de una tribu indígena.

### Cálculo del Tamaño Muestral:

Obtenemos el peso y la altura de 6 individuos a partir de un estudio piloto y estimamos el coeficiente de correlación entre estas dos variables.

1. Abra G\*Power y haga clic en **Tests: Correlation and regression: Correlation: Bivariate normal model**.



De ~~49,00~~  
por solo **19,00**  
\*LOS 3 VOLÚMENES INCLUIDOS

**ESTADÍSTICA APLICADA: Análisis de Datos**  
VOLUMEN: ANÁLISIS

Domina el análisis de datos con el método sencillo y fácil que te permitirá obtener resultados rápidos y efectivos.

1. Descripción y análisis de los datos  
2. Análisis de los datos para la toma de decisiones  
3. Análisis de los datos para la toma de decisiones

5 estrellas

 **Aprender Estadística Fácil**  
@LearnStatisticsEasily

## Vas a aprender a analizar tus datos de forma rápida, sencilla e indiscutible.

Metodología de enseñanza SIMPLE que en poco tiempo te permitirá analizar tus datos por tu cuenta.

Abordamos TODAS las etapas y solo lo que realmente es necesario para analizar tus datos.

Partimos del principio, ya validado, de que es perfectamente posible analizar datos de forma correcta sin necesidad de entender conceptos o fórmulas complejas.

Así, este libro te será útil incluso si no tienes ningún conocimiento previo de estadística.

*Los 3 volúmenes están incluidos en esta oferta.*

**¡ADQUIERE EL TUYO  
HACIENDO CLIC AQUÍ!**



aprender estadística fácil

**CAPÍTULO 2**

**¿CÓMO  
REALIZAR UN  
MUESTREO?**

# 1. ¿POR QUÉ DEBEMOS USAR UN MÉTODO DE MUESTREO APROPIADO?

Vimos en el capítulo anterior que calcular el tamaño de la muestra es esencial para obtener resultados más **confiables** en nuestros análisis, con el mejor costo-beneficio posible.

Sin embargo, supongamos que el muestreo se realiza **inadecuadamente**. En este caso, no basta con calcular el tamaño de la muestra y obtener un número adecuado de elementos.

Un muestreo **mal realizado** introduce sesgos en la muestra, aumenta el error de muestreo y nos lleva a sacar conclusiones erróneas.

El muestreo representa un proceso en el cual **seleccionamos** una parte de los elementos que componen una población para caracterizarla.

Este **proceso** puede realizarse de múltiples maneras.

Aun así, existen formas más **adecuadas** de llevarlo a cabo según la situación.

Por lo tanto, discutiremos los **métodos de muestreo** y cómo debemos aplicarlos correctamente.

### 3.3 MUESTREO ALEATORIO ESTRATIFICADO

El **muestreo estratificado** es un método de muestreo probabilístico utilizado cuando uno o más criterios deben considerarse en la población antes de seleccionar elementos. Este método implica dos etapas:

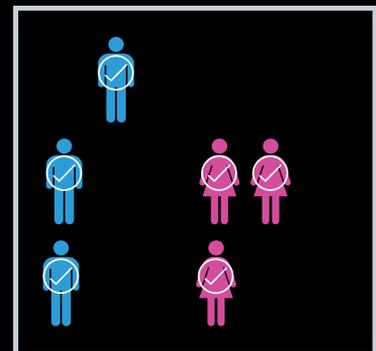
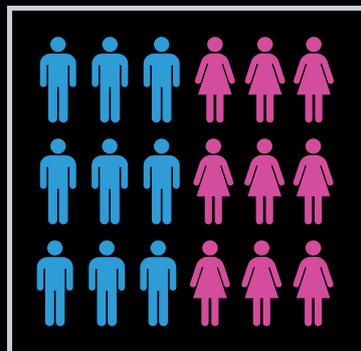
1. Dividir la población en dos o más grupos (estratos) según la característica de interés.
2. Seleccionar aleatoriamente elementos de cada estrato de manera proporcional a su tamaño.

*Este método permite un mejor **control** de la representatividad de diferentes elementos dentro de la población.*

## MUESTREO ALEATORIO ESTRATIFICADO

**EJEMPLO:** Seleccionar 6 personas de una población de 18, dividiendo la población en 2 estratos según el sexo. Luego, elegir aleatoriamente 3 hombres y 3 mujeres de sus respectivos estratos, proporcionalmente a su tamaño.

**POBLACIÓN** → **ESTRATO** → **MUESTRA**



# Y NUNCA OLVIDE EL SESGO DE MUESTREO

SACAR CONCLUSIONES A PARTIR  
DE DATOS NO REPRESENTATIVOS  
DE LA POBLACIÓN OBJETIVO.



En 1948, un periódico de Chicago predijo erróneamente al próximo presidente de Estados Unidos basándose en una encuesta telefónica. No consideraron que, en esa época, solo la clase alta podía tener un teléfono.

La gran lección es que siempre se debe garantizar que la muestra represente a la población.

## 5. CÓMO HACERLO EN LA PRÁCTICA

Podríamos realizar los **sorteos** a la antigua, como usar una bolsa o una urna en un juego de bingo. Pero, obviamente, hoy en día tenemos posibilidades más prácticas y rápidas para hacerlo.

Entonces, primero necesitamos que cada elemento de la población tenga una **identificación** única y esté organizado en una hoja de cálculo. Esta identificación puede ser cualquier número, nombre, código, DNI, etc.

Recuerde que cada **elemento** puede ser una persona, un animal, un registro médico, un cuadrante, una ciudad, un bosque, una laguna, una placa de Petri, una casa, una compra, etc.

Para nuestro **ejemplo**, supongamos que tenemos 200 nombres de pacientes atendidos en una clínica y queremos seleccionar 20 para la investigación.

**1.** Aquí tenemos este ejemplo con los nombres de los pacientes en orden alfabético.

	A	B	C	D	E	F	G
1	<b>NOMBRE</b>						
2	Adrián						
3	Alejandro						
4	Alicia						
5	Ana						
6	Andrea						
7	Andrés						
8	Ángel						
9	Antonio						
10	Beatriz						
11	Carlos						

**aprender estadística fácil**

**CAPÍTULO 3**

**TÓPICOS  
AVANZADOS  
EXTRAS**



# 1. VALOR DE P: EL REGRESO

En el **Volumen I**, presentamos una definición didáctica del valor  $p$ , pero esta no refleja con precisión su significado.

Por lo tanto, para un primer contacto con el tema, esta **simplificación** contribuye a su comprensión.

Así que, como prometimos, ahora presentamos una definición **precisa** del valor  $p$ , pero que requiere mayor atención y abstracción.

---

Cuando realizamos una prueba de hipótesis inferencial, como chi-cuadrado, prueba  $t$ , ANOVA, correlación, regresión, etc., básicamente tenemos dos hipótesis:

## HIPÓTESIS NULA ( $H_0$ )

El patrón más simple, que establece que "*no hay diferencias entre los grupos*" o "*no hay relación entre las variables*".

## HIPÓTESIS ALTERNATIVA ( $H_1$ )

Un estado alternativo, complementario a  $H_0$ , que indica que "*existen diferencias entre los grupos*" o "*existe una relación entre las variables*".

Así, el objetivo básico de cualquier prueba de hipótesis es determinar si rechazamos o no la hipótesis nula ( $H_0$ ), y esta decisión dependerá de **dos factores fundamentales**.

## 1. NIVEL DE SIGNIFICACIÓN ( $\alpha$ )

Representa un **valor de corte**, un criterio que definimos para rechazar o no  $H_0$ . La definición de su valor — generalmente 1 % o 5 % — debe realizarse antes de la prueba.

## 2. VALOR DE P ( $p$ )

El valor  $p$  representa una **probabilidad**, que se obtiene cuando realizamos una prueba de hipótesis inferencial.

Así, cuando realizamos nuestro análisis y obtenemos el valor  $p$ , lo **comparamos** con el nivel de significación ( $\alpha$ ) que definimos previamente.

Como **ejemplo**, supongamos que establecemos un nivel de significación ( $\alpha$ ) de 0.05 (o 5 %). Entonces, al comparar  $\alpha$  con el valor  $p$  obtenido en la prueba, tenemos dos posibilidades:

**1.** Si el valor  $p$  es menor o igual al nivel de significación  $\alpha$  ( **$p \leq 0.05$** ), **debemos rechazar** la hipótesis nula ( $H_0$ ). En este caso, decimos que nuestra prueba fue estadísticamente significativa.

**2.** Si el valor  $p$  es mayor que el nivel de significación  $\alpha$  ( **$p > 0.05$** ), **no debemos rechazar** la hipótesis nula ( $H_0$ ). En este caso, decimos que nuestra prueba no fue estadísticamente significativa.

Así, en términos técnicos, el **valor de p** puede definirse con precisión como:



*El valor-p representa la probabilidad de obtener un resultado igual o (más extremo que) el obtenido a partir de nuestros datos, asumiendo que la hipótesis nula es verdadera.*

**Si mi prueba arrojó, por ejemplo, un valor-p del 2 %, ¿qué significa esto?**

Si consideramos que  $H_0$  es verdadera, la probabilidad de obtener resultados iguales o (más extremos que) los nuestros sería de solo un 2 %.

Sin embargo, como es inferior a  $\alpha = 5 \%$ , rechazamos  $H_0$ . A continuación, se presenta una explicación más detallada.

**$p < 0,05$**

***"We Are The Champions"***





Era una tarde de verano en Harpenden, **Inglaterra**, a principios de la década de 1920.

Un grupo de **científicos** — un estadístico, una psicóloga y un bioquímico — estaban sentados a una mesa para la hora del té en la Estación Experimental de Rothamsted.

La psicóloga, **Muriel** Bristol, insistía en que el té vertido sobre la leche tenía un sabor diferente al que se obtenía cuando la leche se vertía sobre el té.

Todos **cuestionaron** su afirmación — *¿Cuál podría ser la diferencia?* No podían concebir que realmente hubiera alguna **diferencia** en el sabor.

El estadístico **Ronald Fisher** ideó un experimento en el que la señora sería servida con una secuencia de ocho tazas, la mitad con leche vertida sobre el té y la otra mitad con té vertido sobre la leche, sin que ella pudiera ver la preparación.

Entonces, le ofrecieron la **primera taza**.

Ella tomó un pequeño sorbo y **declaró** que, en esa, la leche había sido vertida sobre el té.

Fisher anotó la respuesta y, sin comentar nada, le entregó la **segunda taza...**

**Sorprendentemente**, identificó correctamente todas las 8 tazas al final del experimento.



Ahora vamos a **asociar** este ejemplo con la definición del valor p discutida en este tema.

***Suponiendo que la hipótesis nula ( $H_0$ ) es verdadera — es decir, que Muriel no puede identificar las diferentes infusiones — la probabilidad de clasificar correctamente las 8 tazas, como realmente lo hizo, es de solo un 1,43%.***

Dado que este valor p de 1,43% es menor que el nivel de significancia ( $\alpha$ ) del 5%, la prueba debe considerarse **estadísticamente significativa**.

**Rechazaríamos** la hipótesis nula ( $H_0$ ).

Entonces, "**aceptaríamos**" la hipótesis alternativa ( $H_1$ ) de que Muriel puede distinguir las diferentes infusiones.

### **Nota:**

La importancia de definir el nivel de significancia ( $\alpha$ ) antes de realizar la prueba queda clara en este ejemplo.

Dado que  $\alpha$  suele ser 1% o 5%, cuando obtenemos un valor p entre estos dos valores, por ejemplo, tenderíamos a seleccionar el  $\alpha$  que haría que la diferencia fuera significativa, es decir, 5%.



## 2. TAMAÑO DEL EFECTO

El tamaño del efecto es una medida descriptiva que sirve como un **complemento** a la prueba de significancia.

Representa la **magnitud** — la fuerza — del efecto experimental.

En otras palabras, cuanto mayor sea el tamaño del efecto, más fuerte será la "*diferencia entre grupos*" o la "*relación entre variables*". Además, será más fácil detectarlo, se requerirá un tamaño de muestra menor y el poder estadístico será mayor.

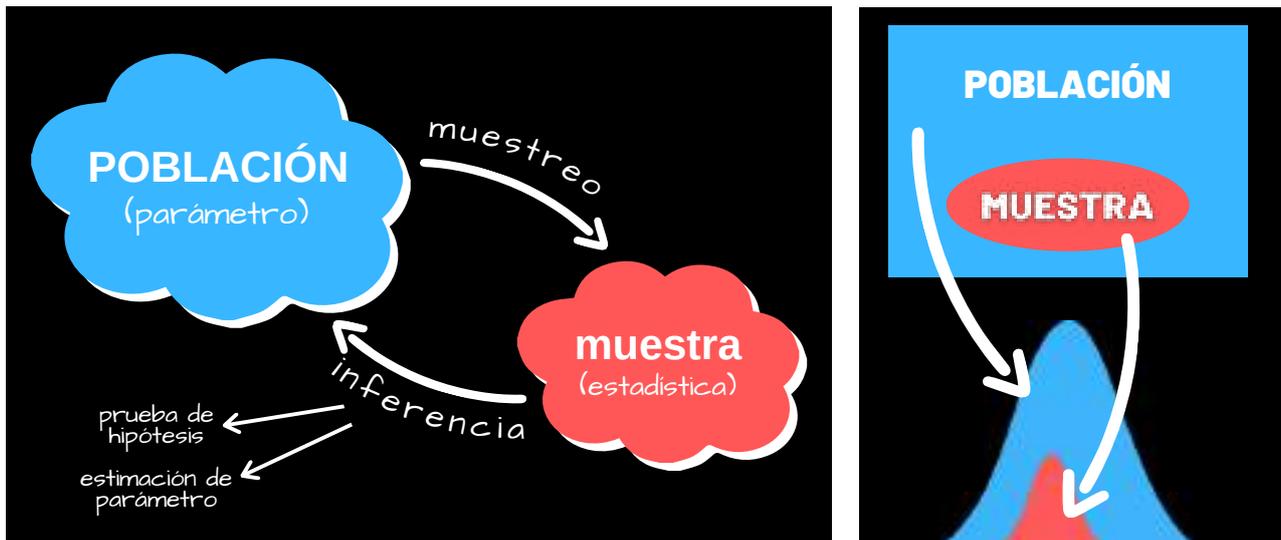


**Comparando** la estatura media entre las poblaciones anteriores, por ejemplo, tenemos:

- **Menor tamaño del efecto:** Población 1 vs. Población 2
- **Mayor tamaño del efecto:** Población 1 vs. Población 3

### 3. INTERVALO DE CONFIANZA

Algunos conceptos ya discutidos en el **Volumen I** son esenciales para entender el intervalo de confianza, como población, muestra y error de muestreo.



Una **población** engloba todos los elementos de un conjunto. Por otro lado, una **muestra** incluye solo una parte de esos elementos.

La **estadística inferencial** nos permite sacar conclusiones sobre la población a partir de la muestra.

Sin embargo, como una muestra representa solo una parte de la población, **nunca** será una representación perfecta.

Así, surge el **error muestral**, representado por la diferencia entre el valor verdadero de la población y el valor obtenido en el muestreo.

## 4. PODER DE LA PRUEBA

El **poder de la prueba** está representado por  $1 - \beta$ , donde  $\beta$  es la probabilidad de cometer un Error de Tipo II.

Cuanto mayor sea el poder de la prueba, mejor, ya que es más probable que podamos **detectar un efecto**. Sin embargo, manteniendo un nivel de significación constante ( $\alpha$ ), solo podemos aumentarlo a medida que incrementamos el tamaño de la muestra.

Además, cuanto mayor sea el **tamaño del efecto** inherente a nuestros datos, mayor tiende a ser el poder de la prueba con un tamaño de muestra más pequeño. Y, obviamente, cuanto menor sea el tamaño del efecto, mayor tendrá que ser nuestra muestra para alcanzar un poder de prueba más alto.

Un valor aceptable para el parámetro de poder de la prueba generalmente está entre **0,80** y **0,99**.

Así, el poder de la prueba ( $1 - \beta$ ) representa la **probabilidad** asumida de:

- (1) Controlar el error de Tipo II, falso negativo.
- (2) Rechazar la hipótesis nula ( $H_0$ ) si realmente es falsa.
- (3) Detectar una diferencia cuando realmente existe.

**¡ESPERA!**  
**AÚN NO HA**  
**TERMINADO**

**MÁS CONTENIDO SERÁ INCLUIDO EN ESTE VOLUMEN. ¡SE ACTUALIZARÁ MÁS VECES! VISITA NUESTRO PERFIL EN INSTAGRAM Y HAZ SUGERENCIAS DE TEMAS.**



**@StatisticsEasily**



De ~~49,00~~  
por solo **19,00**  
\*LOS 3 VOLÚMENES INCLUIDOS

**ESTADÍSTICA APLICADA: Análisis de Datos**  
VOLUMEN: ANÁLISIS

Domina el análisis de datos con el método sencillo y fácil que te permitirá obtener resultados rápidos y efectivos.

1. Descripción y análisis de los datos  
2. Análisis de los datos para su interpretación  
3. Análisis de los datos para su interpretación

UNIVERSIDAD COMPLUTENSE DE MADRID

 **Aprender Estadística Fácil**  
@LearnStatisticsEasily

## Vas a aprender a analizar tus datos de forma rápida, sencilla e indiscutible.

Metodología de enseñanza SIMPLE que en poco tiempo te permitirá analizar tus datos por tu cuenta.

Abordamos TODAS las etapas y solo lo que realmente es necesario para analizar tus datos.

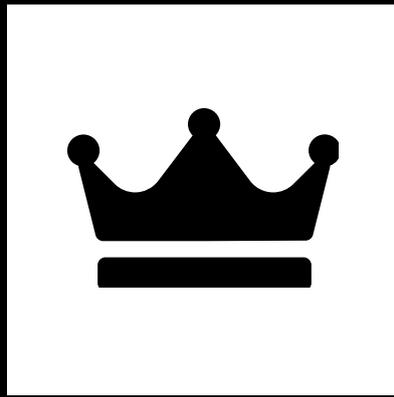
Partimos del principio, ya validado, de que es perfectamente posible analizar datos de forma correcta sin necesidad de entender conceptos o fórmulas complejas.

Así, este libro te será útil incluso si no tienes ningún conocimiento previo de estadística.

*Los 3 volúmenes están incluidos en esta oferta.*

**¡ADQUIERE EL TUYO  
HACIENDO CLIC AQUÍ!**





**Desbloquea los secretos para analizar tus datos de manera rápida, fácil y con confianza.**

**Nuestro enfoque SIMPLE se centra en enseñarte exactamente lo que necesitas saber para dominar el análisis de datos.**

**Deja atrás las complejidades de conceptos, fórmulas y tablas: este curso demuestra que un análisis de datos preciso es posible para todos.**

**Este recurso accesible está diseñado para quienes tienen poco o ningún conocimiento previo de estadística.**

**Descubre nuestro método incomparable para la "análisis de datos rápida, fácil y confiable", un cambio de juego que no encontrarás en ningún otro lugar.**



**APRENDER ESTADÍSTICA FÁCIL**