Test de hipótesis para una población normal

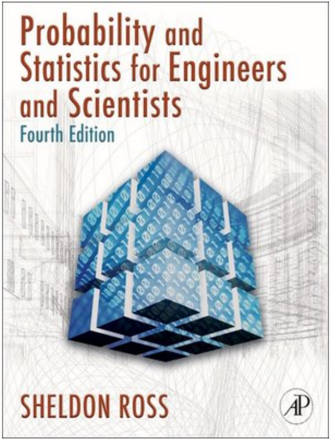
Estadística, Grado en Sistemas de Información

Constantino Antonio García Martínez

Universidad San Pablo Ceu

Contenidos

References



S. Ross. Introduction to Probability and Statistics for Engineers and Scientists.
Chapter 8.

Introducción y test de hipótesis de dos colas

Ejemplo: Homeopatía y pérdida de peso

Un producto homeopático afirma que **"gracias a su uso, perderás 2 Kg en dos semanas"**.

Escéptico ante esta afirmación, reclutas a 50 personas de tu ciudad para participar en un experimento. Las personas usan el producto homeopático durante dos semanas y reportan su pérdida de peso (por ejemplo, $x_i = 3$ significaría que se han perdido 3 Kg, mientras que $x_i = -3$ significaría que se han ganado 3). Datos en "homeo_weight_loss.csv".

En base a los datos, ¿es creible la afirmación del producto homeopático? Por sencillez, asume que la pérdida de peso tiene desviación típica poblacional $\sigma = 2.5$.

Ejemplo: Homeopatía y pérdida de peso

Un producto homeopático afirma que **"gracias a su uso, perderás 2 Kg en dos semanas"**... asume que la pérdida de peso tiene desviación típica poblacional $\sigma = 2.5$.

Ejemplo: Homeopatía y pérdida de peso

Un producto homeopático afirma que **"gracias a su uso, perderás 2 Kg en dos semanas"**... asume que la pérdida de peso tiene desviación típica poblacional $\sigma = 2.5$.

Los datos no parecen apoyar la afirmación del fabricante... ($\bar{x} = 0.037$). ¿Cómo podemos argumentarlo?

1. Como en un juicio, asumimos la "inocencia" del producto homeopático y nos creemos (por el momento) que el producto permite perder 2 Kgs en dos semanas. Asumiendo normalidad, entonces

$$X_i \sim \mathcal{N}(2, \sigma^2 = 2.5^2) \to \bar{X} \sim \mathcal{N}(2, \sigma^2 = \frac{2.5^2}{50}).$$

   LLamaremos a esto, la **hipótesis nula**.

Ejemplo: Homeopatía y pérdida de peso

Un producto homeopático afirma que **"gracias a su uso, perderás 2 Kg en dos semanas"**... asume que la pérdida de peso tiene desviación típica poblacional $\sigma = 2.5$.

Los datos no parecen apoyar la afirmación del fabricante... ($\bar{x} = 0.037$). ¿Cómo podemos argumentarlo?

1. Como en un juicio, asumimos la "inocencia" del producto homeopático y nos creemos (por el momento) que el producto permite perder 2 Kgs en dos semanas. Asumiendo normalidad, entonces
$$X_i \sim \mathcal{N}(2, \sigma^2 = 2.5^2) \rightarrow \bar{X} \sim \mathcal{N}(2, \sigma^2 = \frac{2.5^2}{50}).$$
LLamaremos a esto, la **hipótesis nula**.
2. Bajo este hipótesis, la probabilidad de perder 0.037 Kg (o menos) es muy baja:
$$P(\bar{X} \leq 0.037) \approx 0$$

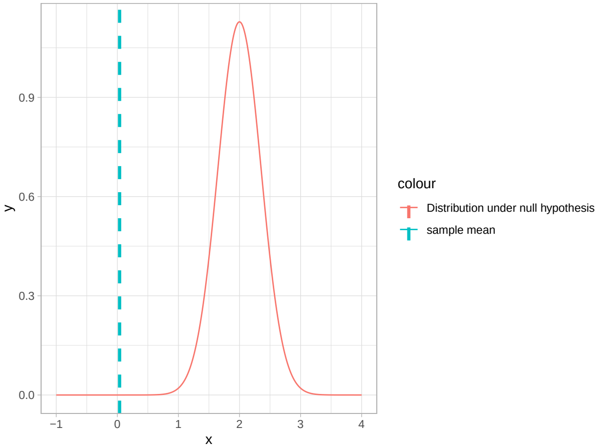
Ejemplo: Homeopatía y pérdida de peso

Un producto homeopático afirma que **"gracias a su uso, perderás 2 Kg en dos semanas"**... asume que la pérdida de peso tiene desviación típica poblacional $\sigma = 2.5$.

Los datos no parecen apoyar la afirmación del fabricante... ($\bar{x} = 0.037$). ¿Cómo podemos argumentarlo?

1. Como en un juicio, asumimos la "inocencia" del producto homeopático y nos creemos (por el momento) que el producto permite perder 2 Kgs en dos semanas. Asumiendo normalidad, entonces
$$X_i \sim \mathcal{N}(2, \sigma^2 = 2.5^2) \rightarrow \bar{X} \sim \mathcal{N}(2, \sigma^2 = \frac{2.5^2}{50}).$$
LLamaremos a esto, la **hipótesis nula**.
2. Bajo este hipótesis, la probabilidad de perder 0.037 Kg (o menos) es muy baja:
$$P(\bar{X} \leq 0.037) \approx 0$$
3. La probabilidad es tan baja que tenemos una **evidencia muy fuerte** contra la hipótesis inicial.

Test de hipótesis



Podemos resumir el procedimiento como sigue:

1. Formulamos una hipótesis. Generalmente aceptamos como hipótesis de partida (**hipótesis nula**, H_0) lo contrario de lo que queremos probar (**hipótesis alternativa**, H_a) [Pensad en la presunción de inocencia].
2. Buscamos un **estadístico de contraste** T que nos permita testear la veracidad de H_0. Es fundamental discutir las **asunciones** bajo las que la distribución de este estadístico es conocido.
3. Calculamos el **p-valor**: bajo la H_0, la probabilidad de observar un estadístico tan extremo como el realmente observado. *Cuanto más pequeño sea el p-valor, más pequeña es la evidencia a favor de* H_0.
4. Comparamos el p-valor con un umbral llamado **nivel de significancia** α. Si p-valor $< \alpha$, descartamos la hipótesis nula y diremos que la hipótesis ha sido **rechazada** o que los **resultados son estadísticamente significativos**. Los valores más típicos de α son 0.05 y 0.01.

Test de hipótesis

Ejercicio: Homeopatía y pérdida de peso
Realiza el test de hipótesis para el problema de la pérdida de peso. ¿Cuál sería tu conclusión?

```
## [1] "p-value =  2.8209e-08"
```

Ejercicio: Homeopatía y pérdida de peso: T-test
En el ejercicio anterior hemos asumido que σ es conocida. Como sabemos, esto ocurre muy rara vez en la práctica.

Repite el test de hipótesis para el problema de la pérdida de peso si σ es desconocida. ¿Cuál sería tu conclusión?

```
## [1] "p-value =  3.21e-06"
```

Hipótesis y errores

Confusiones típicas:

1. La hipótesis relevante es la **hipótesis alternativa**. ¡No rechazar H_0 no implica que sea cierta!

Ejercicio: H_0 no demuestra nada

Interpreta el siguiente fragmento de código...

```
set.seed(42)
x <- rnorm(30, mean = 0, sd = 5)
print(c(
  t.test(x, mu = 1)$p.value,
  t.test(x, mu = -1)$p.value
))

## [1] 0.5707162 0.2506672
```

2. ¡Como en un juicio, es posible cometer errores!

```
set.seed(4)
x <- rnorm(30, mean = 0, sd = 5)
print(
  t.test(x, mu = 0)$p.value
)

## [1] 0.002541535
```

Ejercicio: Interpretación de α

Completa el siguiente fragmento de código para verificar la siguiente interpretación de α: Si repetimos el test muchas veces, **rechazaremos la H_0 siendo esta correcta un α % de veces**.

```
N <- 5000
alpha <- 0.05

sims <- replicate(N, {
  # H0: mu = 0
  x <- rnorm(100, mean = 0)            # Genera muestras de H0 (H0 es correcta)
  test <- t.test(x, mu = 0)$p.value    # Testea H0
  # Añade condición de éxito para calcular la probabilidad de
  # "Rechazar H_0 a pesar de ser correcta"
  ??
})
print(paste("alpha =", alpha, "| p(incorrectly reject H0) = ", mean(sims)))
```

Errores tipos I y II

	H_0 es cierta	H_a es cierta
Se escogió H_0	No hay error	Error de tipo II/Falso negativo (β)
Se escogió H_a	Error de tipo I/Falso positivo (α)	No hay error

Tenemos que

$$P(\text{escoger } H_a | H_0 \text{ es cierta}) = \alpha$$

$$P(\text{escoger } H_0 | H_a \text{ es cierta}) = \beta$$

Se denomina **potencia del contraste** al valor $1 - \beta$, esto es, a la probabilidad de escoger H_a cuando esta es cierta.

Test de hipótesis de una sola cola y tamaño del efecto

Test de hipótesis de una sola cola y tamaño del efecto

Test de hipótesis de una sola cola

Resuelva la siguiente versión del ejercicio de la homeopatía…

Ejercicio: Test de una sola cola
Un producto homeopático afirma que "gracias a su uso, perderás **al menos** 2 Kg en dos semanas"… ¿Es creíble esta afirmación?

```
## [1] "p-value =  1.6027e-06"
```

Tests de hipótesis de una sola cola

Resuelva la siguiente versión del ejercicio de la homeopatía...

Ejercicio: Test de una sola cola
Un producto homeopático afirma que "gracias a su uso, perderás **al menos** 2 Kg en dos semanas"... ¿Es creíble esta afirmación?

```
## [1] "p-value =  1.6027e-06"
```

A este tipo de problemas se les conoce como **hipótesis de una sola cola** (Vs. **hipótesis de dos colas**).

Test de hipótesis de una sola cola y tamaño del efecto

Tamaño del efecto

Ejercicio: Fármaco para pérdida de peso

Gracias a tu éxito con el análisis del producto homeopático, una farmacéutica interesada en desarrollar un fármaco para la pérdida de peso te contrata. La empresa quiere comercializar su (carísimo) producto con un eslogan del tipo "Hay evidencia científica de nuestro producto te hará perder peso si lo usas dos meses".

Te facilitan los datos de "pharma_weight_loss.csv". ¿Hay suficiente evidencia de que el fármaco te hace perder peso? (Usa $\alpha = 0.05$.)

```
## [1] "P-value =  0.00031"
```

Ejercicio: Fármaco para pérdida de peso

Gracias a tu éxito con el análisis del producto homeopático, una farmacéutica interesada en desarrollar un fármaco para la pérdida de peso te contrata. La empresa quiere comercializar su (carísimo) producto con un eslogan del tipo "Hay evidencia científica de nuestro producto te hará perder peso si lo usas dos meses".

Te facilitan los datos de "pharma_weight_loss.csv". ¿Hay suficiente evidencia de que el fármaco te hace perder peso? (Usa $\alpha = 0.05$.)

```
## [1] "P-value =  0.00031"
```

¿Tú crees que alguien comprará el producto?

Ejercicio: Fármaco para pérdida de peso

Gracias a tu éxito con el análisis del producto homeopático, una farmacéutica interesada en desarrollar un fármaco para la pérdida de peso te contrata. La empresa quiere comercializar su (carísimo) producto con un eslogan del tipo "Hay evidencia científica de nuestro producto te hará perder peso si lo usas dos meses".

Te facilitan los datos de "pharma_weight_loss.csv". ¿Hay suficiente evidencia de que el fármaco te hace perder peso? (Usa $\alpha = 0.05$.)

```
## [1] "P-value =  0.00031"
```

¿Tú crees que alguien comprará el producto? **No hay que confundir la significancia estadística con la relevancia práctica**. Para esto último debemos usar **ICs** o **el tamaño del efecto**.

Tamaño del efecto e ICs

Ejercicio: Tamaño del efecto
Calcula el tamaño del efecto para el problema del fármaco. ¿Cómo lo interpretas? Usa la siguiente tabla.

```
## [1] "Cohens d =  0.109"
```

Ejercicio: ICs unilaterales
Calcula el **IC unilateral** asociado al test.

```
## [1] 0.05542461        Inf
## attr(,"conf.level")
## [1] 0.95
```

Siempre que hagas un test de medias, **usa el tamaño del efecto o ICs para complementar los p-valores**.

Potencia de un test

Ejercicio: Test para la varianza

Los test de cociente intelectual (CI) están diseñados para que la desviación típica poblacional sea de 15 puntos. Sin embargo, en los procesos de traducción de un test "oficial" de CI pueden surgir desajustes.

Por ejemplo, "iq_spanish.csv" tiene los resultados de un test de CI traducido del inglés al español. ¿Hay evidencia de que la desviación típica es distinta de 15 y de que, por tanto, debe revisarse la traducción? Usa un nivel de significación del 5 %.

```
## [1] "P-value = 0.018"
## [1] "Confidence interval = (215.3795, 859.0724)"
```

Ejercicio: Test para la varianza

Los test de cociente intelectual (CI) están diseñados para que la desviación típica poblacional sea de 15 puntos. Sin embargo, en los procesos de traducción de un test "oficial" de CI pueden surgir desajustes.

Por ejemplo, "iq_spanish.csv" tiene los resultados de un test de CI traducido del inglés al español. ¿Hay evidencia de que la desviación típica es distinta de 15 y de que, por tanto, debe revisarse la traducción? Usa un nivel de significación del 5 %.

```
## [1] "P-value = 0.018"
## [1] "Confidence interval = (215.3795, 859.0724)"
```

Lo razonable sería repetir el experimento con más muestras. Ahora bien, ¿con cuántas?

Potencia de un test

Ejercicio: Potencia de un test

Como $\hat{s}^2 = 388.6902$, escribe una función que calcule la probabilidad de rechazar H_0 si $\sigma^2 = 388.6902$ para un número de muestras n. Completa el siguiente código:

```
power_var_test <- function(n, H0_sigma2 = 15 ^ 2, true_sigma2 = 388.6902,
                           significance = 0.05, N = 5000) {
  sims <- replicate(N, {
    data <- rnorm(n, sd = sqrt(true_sigma2))
    var_stat <- ??
    p_value <- ??
    p_value < significance
  })
}
```

Ejercicio: Potencia de un test: tamaño de la muestra

Halla ahora el número de muestras necesarias para obtener una potencia de test del 90 %.

```
## [1] "Se necesitan sobre 70 muestras"
```

Ejercicio: Potencia del T-test: número de muestras

Según la estadísticas oficiales, la media de peso de las mujeres de cierto país es de 63.5 Kg (con desviación típica 4.1). Sin embargo, un equipo de investigadores cree que debido a cambios en la alimentación la media se ha incrementado. ¿Cuántas muestras necesitarán para poder detectar un incremento de medio Kg con un nivel de significación del 1 % y una potencia del 90 %? Usa *power.t.test*

```
## [1] "Se necesitarán 221.531 muestras"
```