Fundamentos de Estadística

Silvia N. Pérez
Especialización en Ciencia de Datos - UNO

Tests o pruebas de hipótesis

Biblio:

- Clásicos de PyE: Devore, Mendenhall, Meyer… etc

-

Ensayos o pruebas de hipótesis para un parámetro

Si tenemos una variable aleatoria X de la que no conocemos uno de sus parámetros, podemos hacer “inferencia” acerca de este.. cómo?

- Estimando un valor ✓
- Encontrando un intervalo de confianza ✓
- Decidiendo si este valor cumple alguna condición

Ejemplo

Se tiene interés en el tiempo medio para embalar un producto. El objetivo es que este sea de 10 minutos, por lo que el encargado de planta debe controlarlo.

Decide que, a menos que haya evidencia sustancial de que esto no se esté cumpliendo, seguirá el proceso como hasta ahora.
¿Cómo hace el control?
Y cómo decide?

Organizando el problema

- Defino la variable de interés

X: tiempo de embalado (en min)

- ¿Qué sabemos de X?

nos informan que es una v.a. Normal pero no sabemos su media…
aunque se sabe que su desvío es 3 minutos.

-Estamos interesados en saber si $\mu = 10$ ó $\mu \neq 10$

… cómo decidimos? Con datos!

$\overline{X} = \frac{\sum_{1}^{n} X_i}{n}$

- Tomamos una m.a. de tamaño n y consideramos la variable promedio.

media de la muestra = promedio

Análisis

Llamamos hipótesis a las afirmaciones sobre las que queremos decidir

$H0{:}\ \mu = 10$ HIPÓT. NULA

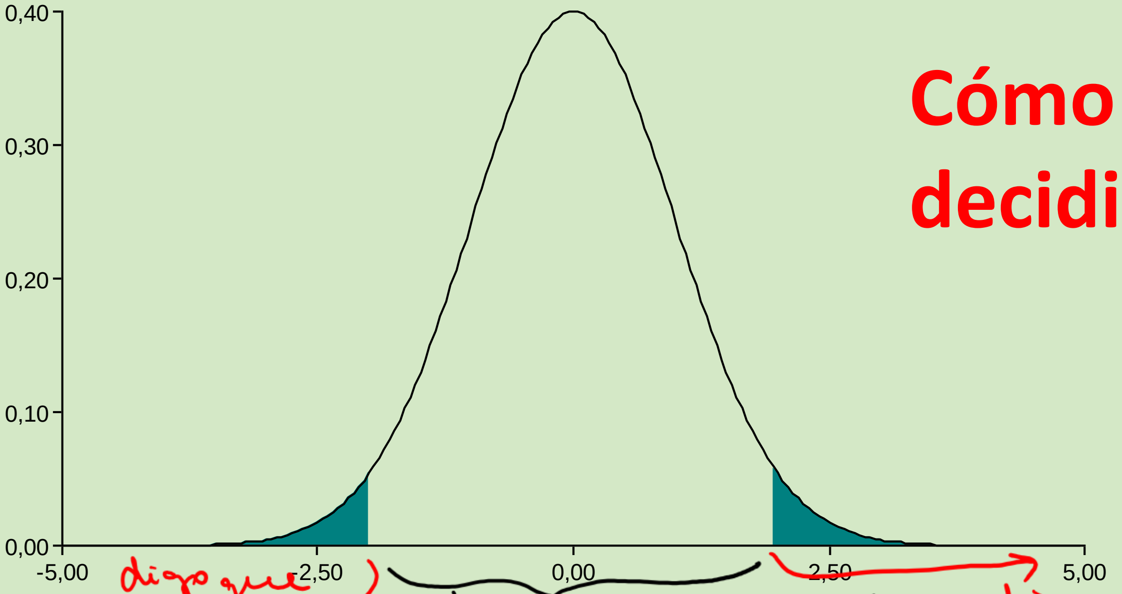
$H1{:}\ \mu \neq 10$ HIPÓT ALTERNATIVA

Veamos lo que nos puede decir la muestra: $\bar{X} \sim \mathcal{N}(\mu\,; \frac{\sigma}{\sqrt{n}})$

si se está cumpliendo Ho
o sea $\mu = 10$

$\bar{X}$ estandarizada

$$\frac{\bar{X} - 10}{\sigma/\sqrt{n}} \sim \mathcal{N}(0; 1)$$



digo que $\mu \neq 10$ digo 'parece q' $\mu = 10$' digo $\mu \neq 10$

Cómo decidimos????

Test de hipótesis

Se quiere decidir acerca de dos afirmaciones:

H_0 versus H_1

Hay pruebas de distinto tipo:

- Acerca de un parámetro → media, varianza, proporción
- Acerca de una distribución → si X es normal ó si X es Binomial ó, ...
- Acerca de comparación en dos o más poblaciones
- Etc!

Se requiere construir una regla de decisión a partir de observar la muestra y luego tomaremos una decisión en base a la muestra observada.

El planteo será siempre:

"¿hay suficiente evidencia para rechazar H_0 ?"

Pruebas de hipótesis para un parámetro

Si queremos decidir acerca del parámetro θ de una variable X, proponemos dos afirmaciones:

H0: θ está en un conj de valores

$\boxed{\textit{H1:}\ \theta \text{ no está en esos valores}}$

Es la hipótesis de interés!

Hacer un test de hipótesis es diseñar una regla de decisión:

- Si en la muestra veo alguna condición -> digo que no vale H0 (se dice "rechazo H0")

Esta regla se puede definir poniendo condiciones para el riesgo de equivocarnos.

Riesgos o errores en un ensayo de hipótesis

Lo básico:

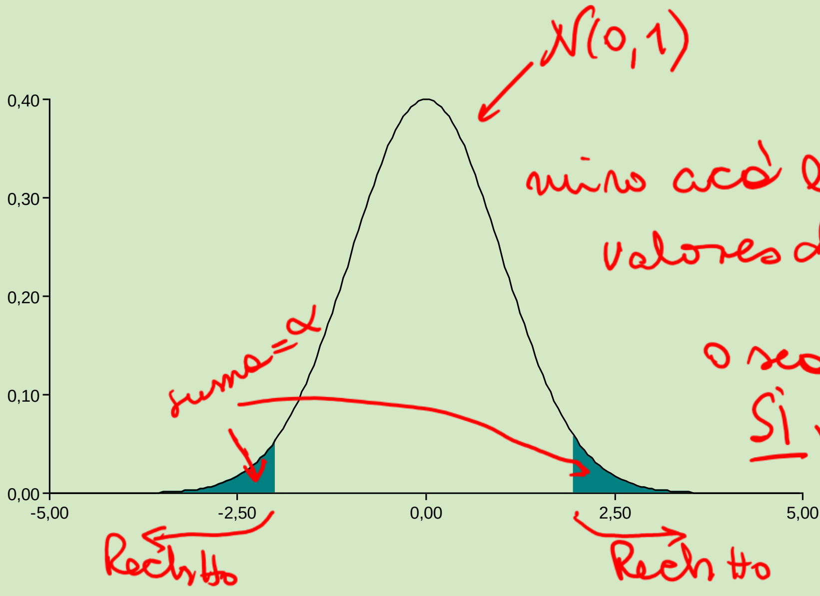
Realidad \ Acción	NO rechazo H0	Rechazar H0
H0 es verdadera	ok	Error tipo I
H0 es falsa	Error tipo II	ok

riesgo a controlar es P(error tipo I)

α= max P(Rechazar H_0 / H_0 es V) = max P(error tipo I)   nivel de significación

$\beta = \beta(\mu)$ = P(no Rechazar H_0 / H_0 es Falso)

Volviendo al ejemplo



miro acá los valores de $\frac{\bar{X}-10}{\sigma/\sqrt{n}}$

o sea SI vale Ho

- Hipótesis $H0\text{: } \mu = 10$

  $H1\text{: } \mu \neq 10$ ← es lo que quisiera decir

- Consideramos una m.a. de tamaño *n*
- Fijamos un riesgo o probabilidad de equivocarnos. Se llama nivel de significación:

$$\alpha = \max P(\text{Rechazar } H_0 \,/\, H_0 \text{ es V})$$

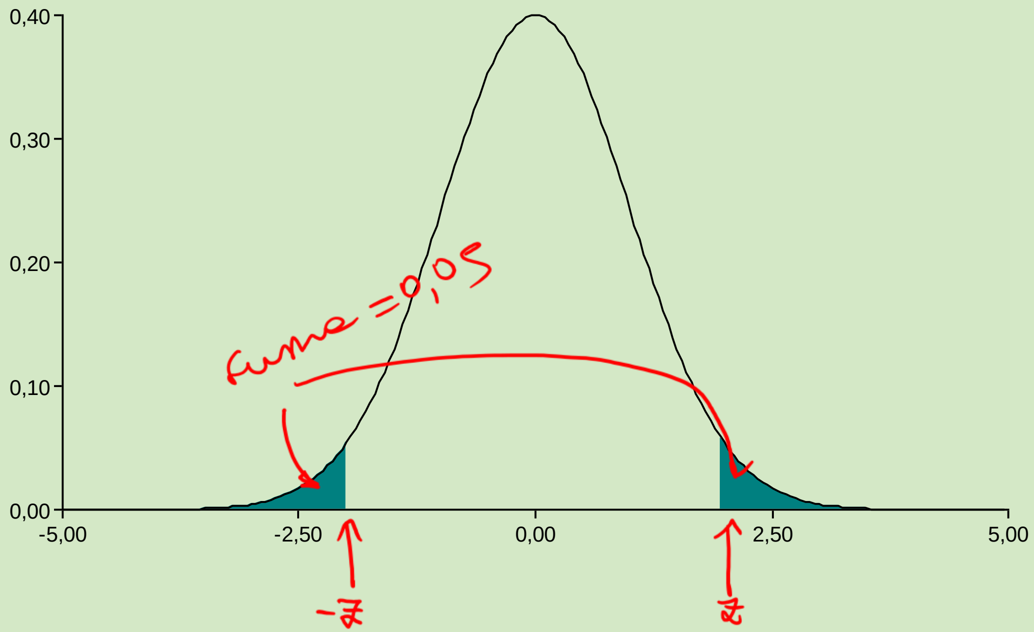
α mide el riesgo de "decir H_1" cuando vale Ho

- Construímos una regla de decisión a partir de una variable a observar en la muestra, que dependa del parámetro y que tenga una distribución conocida (estadístico de prueba)

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0;1)$$

- Tomamos una decisión en base a la muestra observada.

En el ejemplo



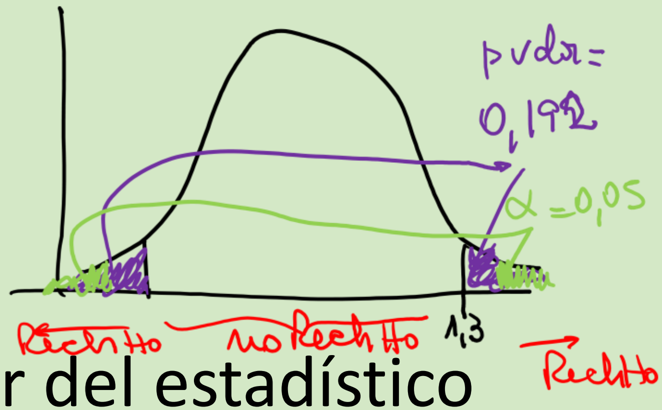
- Hipótesis $H0{:}\ \mu = 10$
  $H1{:}\ \mu \neq 10$

- Por ejemplo, si la m.a. es de tamaño *n= 25* , $\sigma=3$

- Fijamos un riesgo α=0,05

- Regla de decisión: si $\frac{\bar{X}-10}{3/5} < -z$ ó $\frac{\bar{X}-10}{3/5} > z$ entonces diremos que hay evidencia para decir que el tiempo medio de embalaje es distinto de 10 (rechazamos H_0)

P-valor: otra forma de decidir



Bajo el mismo planteo,

$$H0: \mu = 10$$
$$H1: \mu \neq 10$$

Una forma diferente de decidir surge de observar el valor del estadístico de prueba calculado en los datos:

Esa prob, podemos decir que

$$\text{Zobs} = \frac{\bar{X}obs - 10}{3/5} = 1,3$$

0 ← pv
da evidencia p/ Rech Ho

pv →
1 pvalor
no hay evidencia p/ Rechazar

Se define p-valor de la prueba como el mínimo nivel de significación que me llevaría a rechazar la hipótesis nula.
En el ejemplo: pvalor = 2*P(Z > 1,3) = 0,192 alto!! Refuerza la idea de no rechazar H0.

Ensayo de hipótesis para media de una Normal

¡Necesitamos un estadístico de prueba! es la variable q' observo en la muestra para decidir

$X \sim \mathcal{N}(\mu\,;\sigma^2)$	Varianza $\boldsymbol{\sigma}^2$ conocida	$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0;1)$ bajo Ho
$X \sim \mathcal{N}(\mu\,;\sigma^2)$ condición!	Varianza $\boldsymbol{\sigma}^2$ desconocida	$\dfrac{\bar{X}-\mu}{s/\sqrt{n}} \sim t_{(n-1)}$ bajo Ho

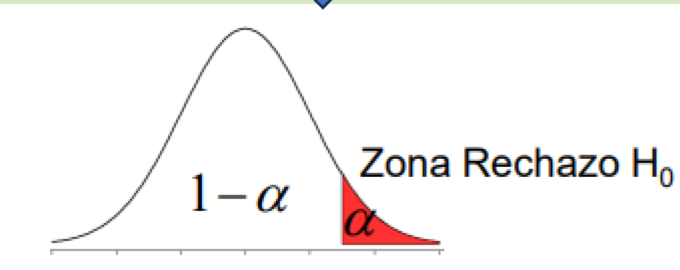
Tipos de pruebas para la media

$$H_0: \mu = \mu_0$$

Pruebas unilaterales

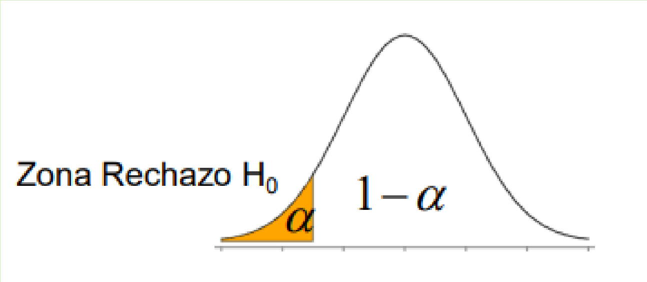
unilateral derecha

$H1: \mu > \mu_0$



unilateral izq

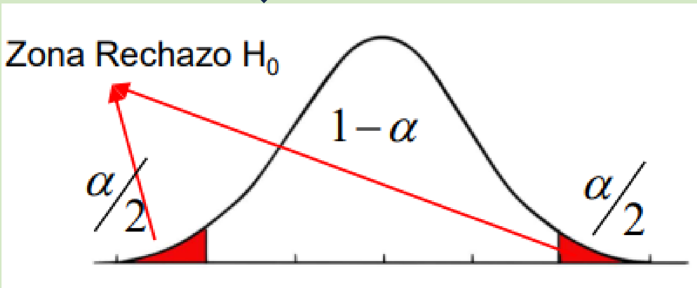
$H1: \mu < \mu_0$



Prueba bilateral

a 2 colas

$H1: \mu \neq \mu_0$



Ensayo de hipótesis para μ de Normal con varianza desconocida

- Plantear las hipótesis nula y alternativa (Ho y H1).
- n?
- Estadístico de prueba (será la t…)
- Establecer el nivel de significación (α) y la regla de decisión (regiones de rechazo y de no rechazo)
- Ó ir por el pvalor!!
- Calcular el valor del estadístico observado y el pvalor
- Decidir!

¿Porqué hablamos sólo de “no rechazo H0”?

$H0$: *no es culpable*
$H1$: *es culpable*

Si la justicia encuentra pruebas suficientes	→	**CULPABLE**
Si la justicia NO encuentra pruebas suficientes	→	**no hay evidencia para declararlo culpable!**

Pruebas de normalidad

En muchas aplicaciones se requiere cumplir condiciones de Normalidad para la variable.

¿Cómo decidir si eso se cumple?

- Método gráfico: Qqplot (para tener una idea..)
- Pruebas estadísticas de normalidad: H_0: X se distribuye Normal H_1: X no se distribuye normal
    - Shapiro-Wilk: para contrastar normalidad cuando el tamaño de la muestra es menor de 50.
    - Kolmogorov-Smirnov con modificación de Lilliefors (para muestras grandes): asume que la media y varianza son desconocidas,especialmente desarrollado para contrastar la normalidad.
    - Jarque- Bera: no requiere estimaciones de los parámetros. Recomendable sumar a los otros.

Resumiendo: para hacer un test de hipótesis para μ (con varianza desconocida)

- Plantear las hipótesis nula y alternativa (Ho y H1).
- Revisar si la variable puede considerarse con distribución Normal.
- Estadístico de prueba → t
- Establecer la regla de decisión (regiones de rechazo y de no rechazo)
- Calcular el valor del estadístico observado
- Analizar el p-valor
- Decidir expresando "bien" las conclusiones

OBS: - el supuesto de normalidad no necesita cumplirse estrictamente.
- si n es grande, aún cuando no se cumple (o sea regular) la normalidad, igual vale el test t.

Ejemplo en R

A partir de los datos de estudiantes, se quiere decidir si la nota media es:

1) distinta de 6.6

2) Mayor a 6.3

3) Menor a 6.7

Previamente, tendremos que decidir sobre la normalidad de la variable X: nota de un estudiante en la población.

Prueba de normalidad

Lilliefors (Kolmogorov-Smirnov) normality test

data: estudiantes$nota ← variable X

H_0: X es Normal H_1: X no es Normal

D = 0.043198, p-value = 0.7255 → alto! ⇒ no hay evidencia para Rechazar H_0

⇒ " " " pero decir X no es normal

One Sample t-test

data: nota   todos

t = -0.78667, df = 145, $\boxed{\text{p-value = 0.4328}}$

alternative hypothesis: true mean is not equal to 6.6 → $H_0: \mu = 6.6$ $H_1: \mu \neq 6.6$

95 percent confidence interval:
[6.306496; 6.726381] IC con nivel 0,95=1−0,05

6.6 ∈ IC entonces no hay evidencia p/ Rech H_0

sample estimates:
mean of x
 6.516438

$\hat{\mu}$ = estimado de μ = $\bar{x}$ = 6,5164

Relación entre prueba de hipótesis e intervalos de confianza

Dado un parámetro θ para una variable X:

- El IC de confianza es un intervalo [LI, LS] tal que si el parámetro a estimar es θ, entonces:

$$P(LI \leq \theta \leq LS) = 1\text{-}\alpha$$

Decimos entonces que para intervalo observado en la muestra tenemos **mucha confianza** de que contenga al parámetro.

- Si planteamos las hipótesis H0: $\theta = \theta_0$ y H0: $\theta \neq \theta_0$ entonces, podemos decir que

$$\text{Rechazo H0 a un nivel } \alpha \Longleftrightarrow \text{el IC no contiene a } \theta_0$$

Relación entre IC y el test de medias

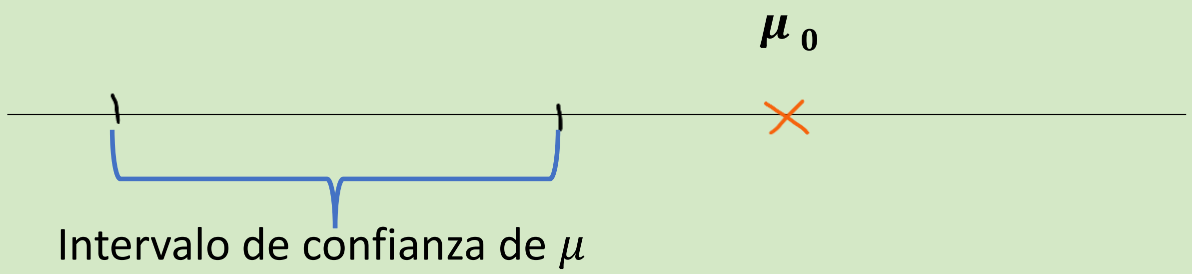
Tenemos un intervalo de confianza IC para la media μ

y queremos probar

$$H_0: \mu = \mu_0 \qquad H_1: \mu \neq \mu_0$$

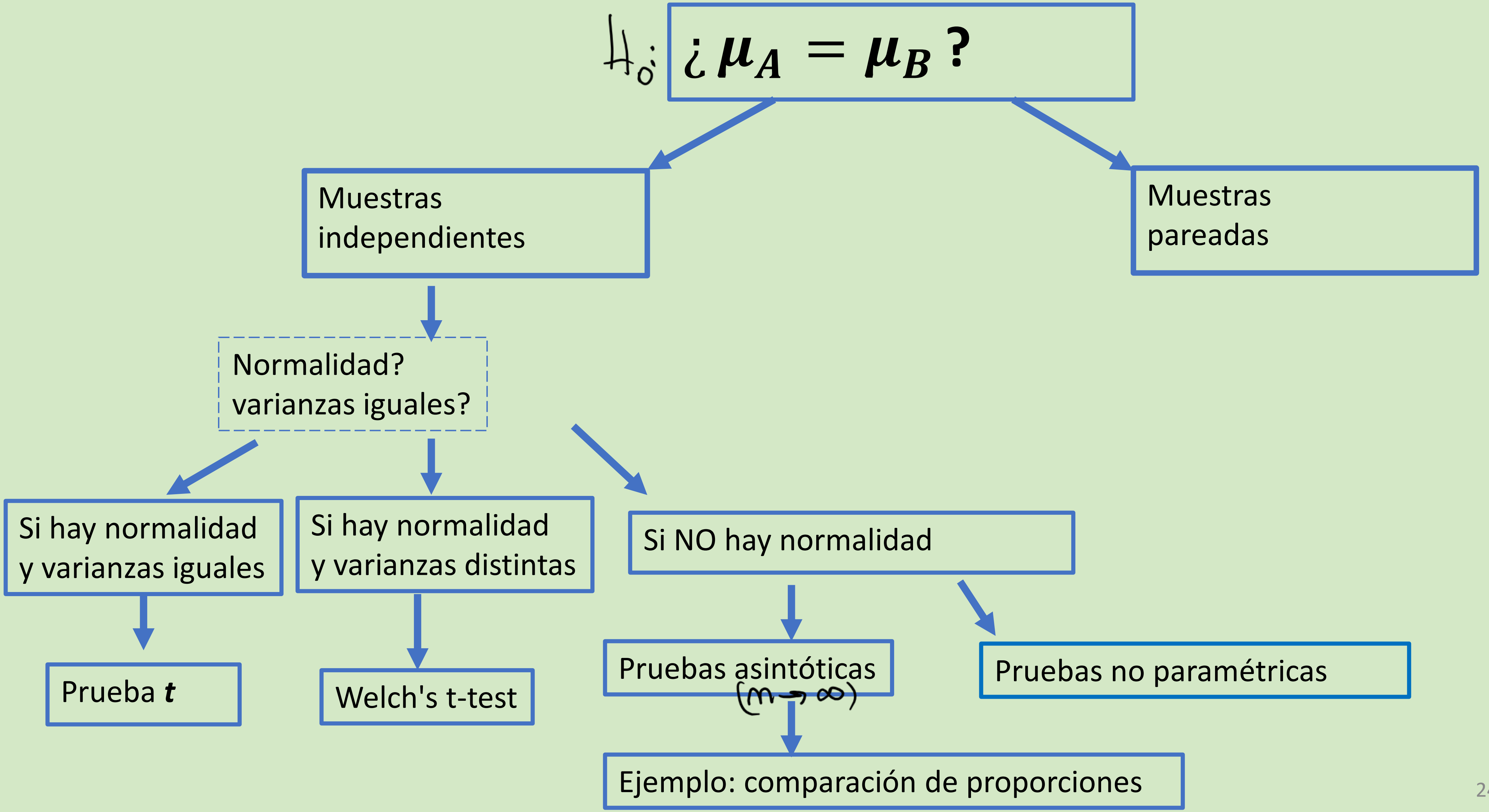
$\boxed{\text{Rechazo } H_0}$ ⟺ $\boxed{\mu_0 \text{ NO pertenece al IC}}$ de nivel $(1-\alpha)$

con nivel de signif α



Pruebas de comparación de medias

Pruebas de comparación de medias



Comparación de medias. Caso Normales Independientes

Se quiere comparar si

H0: $\boldsymbol{\mu_A} = \boldsymbol{\mu_B}$ $\Longleftrightarrow$ $\mu_A - \mu_B = 0$

H1: $\boldsymbol{\mu_A} \neq \boldsymbol{\mu_B}$

$H_1: \mu_A > \mu_B \longrightarrow \mu_A - \mu_B > 0$

$H_1 \quad <$

Condiciones: normalidad de las variables y homogeneidad de varianzas.

Estadístico de prueba: $$\frac{\bar{X}_A - \bar{X}_B}{\tilde{s}\sqrt{\frac{1}{nA} + \frac{1}{nB}}} \sim t_{(nA+nB-2)}$$

Pruebas de comparación de varianzas (homogeneidad de varianzas)

$H_0: \sigma_A^2 = \sigma_B^2$ $\qquad$ $H_1: \sigma_A^2 \neq \sigma_B^2$

Test de Levene: sirve para comparar 2 o más poblaciones con muestras de igual tamaño o similar.

Test de Bartlett: es más sensible que el anterior a la falta de normalidad.

Ejemplo en R

A partir de los datos de estudiantes, se quiere decidir si las nota medias de estudiantes de privados y públicas son:

1) iguales

2) La de privados es menor que la de públicos

3) Puedo decir que la diferencia entre públicos y privados es mayor a 0.4??

Previamente, tendremos que decidir sobre si las varianzas pueden suponerse iguales. p-value = 0.9383 entonces puedo asumir que las varianzas son iguales
Si no se cumple, R puede calcular el test de Welch

1) $H_0: \mu_{Priv} = \mu_{pub} \rightsquigarrow \mu_{PRIV} - \mu_{PUB} = 0$ $\quad H_1: \mu_{Priv} - \mu_{pub} \neq 0$

Two Sample t-test

data: privados$nota and publicos$nota

t = -1.9896, df = 144, p-value = 0.04853 → $p_v < 0{,}05$ ⇒ hay evidencia p/ Rech Ho (pvalor chico)

a un nivel de signif 0,05...

o sea hay evidencia para decir q' las medias son ≠

alternative hypothesis: true difference in means is not equal to 0 } H_1

95 percent confidence interval:

[-0.834437057;-0.002746041] IC [ ] 0 no contiene al 0 !!

sample estimates:

mean of x mean of y

$\underbrace{6.301408}_{\bar{x}}$ 6.720000

$\bar{x}$ nota prom priv

$\bar{y}$ nota prom pub

2) $H_0: \mu_{priv} = \mu_{pub}$ $\quad - \quad = 0$

$H_1: \mu_{priv} < \mu_{pub} \rightsquigarrow \mu_{priv} - \mu_{pub} < 0 \leftarrow$ esto es lo q' interesa

Two Sample t-test

data: privados$nota and publicos$nota
t = -1.9896, df = 144, p-value = 0.02426 → hay evidencia p/ Rech Ho a nivel de signif
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
-Inf -0.07029515

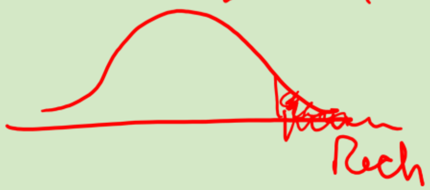
-0,07

sample estimates:
mean of x mean of y
6.301408 6.720000

3) la diferencia entre públicos y privados es mayor a 0.4??

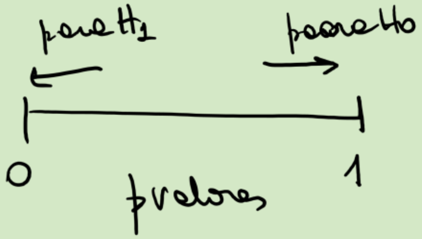
α = max $\mathbb{P}$(Rech Ho | Ho es V) quiero q' sea chico

si



$H_0: \mu_{PUB} - \mu_{PRIV} = 0,4$

$H_1: \mu_{pub} - \mu_{priv} > 0.4$



Two Sample t-test

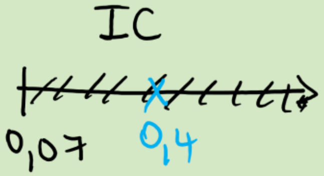
data: publicos$nota and privados$nota

t = 0.088368, df = 144, p-value = 0.4649 → es alto! ⇒ no hay evidencia p/ Rechazar Ho

(" " " p/ decir q' μ_{pub} supera en más de 0,4 a la de priv)

alternative hypothesis: true difference in means is greater than 0.4] $H_1: \mu_{pub} - \mu_{priv} > 0,4$

95 percent confidence interval:
 0.07029515 Inf



0,4 ∈ IC ⇒ no se puede Rechazar H_0

sample estimates:

mean of x mean of y

 6.720000 6.301408

Comparación de medias: Caso apareadas

Se quiere comparar si

$H_0: \mu_A = \mu_B$ ⇨ $\mu_D = \mu_B - \mu_A = 0$

$H_1: \mu_A \neq \mu_B$

El estadístico de prueba resulta de plantear $D=Y-X$ y con esto, se vuelve un caso univariado

es una variable en 1 dimensión!

$$\frac{\overline{\boldsymbol{D}} - \mathbf{0}}{\boldsymbol{s}/\sqrt{\boldsymbol{n}}} \sim \boldsymbol{t}_{(\boldsymbol{n}-\mathbf{1})}$$

Ejemplo en R

A partir de los datos de presión antes y después de ejercicio físico, se quiere decidir si :

1) La media antes es menor que la media después

2) La diferencia entre media después y media antes es mayor a….

Ej: datos antes - después

Paired t-test

data: datos$antes and datos$despues
t = -1.3979, df = 19, p-value = 0.08913
alternative hypothesis: true mean difference is less than
0
95 percent confidence interval:
 -Inf 0.9241517
sample estimates:
mean difference
 -3.9

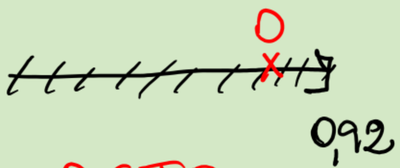
$H_0: \mu_{antes} - \mu_{después} = 0$

$H_1: \mu_{antes} - \mu_{después} < 0$

con nivel de signif $0{,}05 = \alpha$, no Rechazo H_0

pero con nivel $\alpha = 0{,}1$ el $pv = 0{,}089 < 0{,}1$

$\Rightarrow$ hay evidencia p/ Rech H_0



$0 \in IC \Rightarrow$ al nivel $\alpha = 0{,}05$ no rechazo H_0