

## PRECISIÓN Y VERACIDAD. EVALUACIÓN DE DATOS ANALÍTICOS

- Introducción
- Planteamiento de las pruebas
- Hipótesis nula y alternativa
- Prueba o test de una y dos colas
- Errores de tipo I y II
- Resumen de tests o pruebas de significación

## PRECISIÓN Y VERACIDAD. EVALUACIÓN DE DATOS ANALÍTICOS

Se determina el  $\text{CaCO}_3$  de una muestra de un MRC (contiene 50,00 mg)  
Se hacen 4 determinaciones ( $n=4$ ). La media  $\bar{x} = 49,30$  y  $\sigma$  (estimada previamente) vale 0,82

### Objetivo:

Conocer si las medias obtenidas por nosotros son significativamente diferentes del valor verdadero o teórico esperado.

Así sabremos si nuestras medidas llevan asociados un (bias) además del error aleatorio inevitable.

$$\begin{array}{c} \text{Repetitividad ó Reproducibilidad intermedia} \\ \downarrow \\ e_{ij} = (x_{ij} - \mu_j) + (\mu_j - \mu_0) \\ \uparrow \\ \text{bias (método + laboratorio)} \end{array}$$

Método

**Se trata de ver si las diferencias son achacables exclusivamente al puro azar (o aleatoriedad) o existe un sesgo o bias que causa dichas diferencias**

punto de vista químico)

## HIPÓTESIS NULA Y ALTERNATIVA

### Hipótesis nula

"No hay diferencias entre  $\mu$  y  $\mu_0$ "

$$H_0: \mu = \mu_0$$

### Hipótesis alternativa

Necesaria por si  $H_0$  no puede mantenerse

"Hay diferencias entre  $\mu$  y  $\mu_0$ "

$$H_0 \equiv H_0$$
$$H_1 \equiv H_1$$

$$H_1: \mu \neq \mu_0$$

**¡Solo puede elegirse una!**

$$H_1: \mu > \mu_0$$

$$H_1: \mu < \mu_0$$

## PLANTEAMIENTO DE LAS PRUEBAS

### Método 1



Plantear  $H_0$  y  $H_1$

Elegir el nivel de significación  $\alpha$

Calcular el intervalo de confianza alrededor de  $\bar{x}$  a un nivel de  $100-\alpha$  % = 95%

Investigar si  $\mu_0$  está dentro de ese intervalo

Si la respuesta es positiva, se acepta  $H_0$ . Si es negativa se rechaza  $H_0$  (y se acepta  $H_1$ )



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

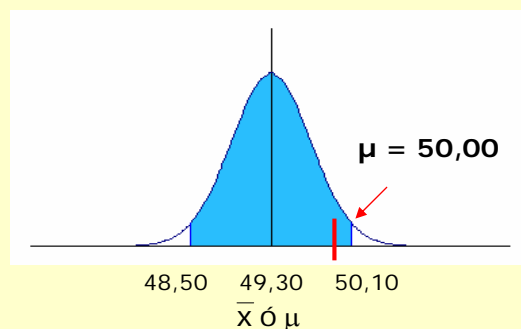
**Trabajamos con medias**

Nivel  $\alpha = 0,05$

$\mu = \bar{x} \pm 1,96\sigma / \sqrt{n} = 49,30 \pm 1,96 * 0,82 / \sqrt{4} = 49,30 \pm 0,80$  , es decir:  $48,50 \leq \mu \leq 50,10$

$\mu_0$  (50,00) está dentro de ese intervalo

Se acepta  $H_0$  y  $\mu = \mu_0$



**¡No hemos probado que  $H_0$  sea cierta, sino sólo que **no** tenemos evidencias para rechazarla!**

## PLANTEAMIENTO DE LAS PRUEBAS

### Método 2

Ahora se supone que  $H_0$  es cierta, y se calcula la **máxima distancia** a la que llega el intervalo de confianza de  $\mu_0$  en unidades  $z$  o  $t$  (valor crítico). Si la distancia de nuestra media (valor experimental) es menor que la crítica, entonces nuestra media pertenecerá a la población de  $\mu_0$  y la  $H_0$  será cierta.

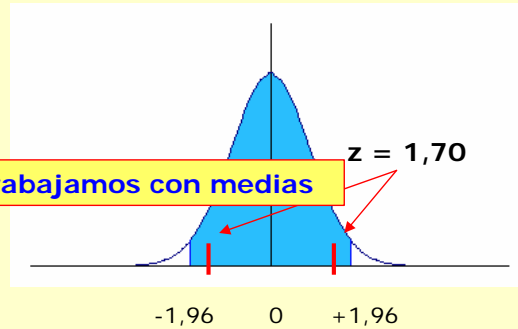
- ➡ Plantear  $H_0$  y  $H_1$
- ➡ Elegir el nivel de significación  $\alpha$
- ➡ Calcular el valor  $z_{crit}$  del estadístico correspondiente (distancia máxima)
- ➡ Calcular el valor  $z_{exp}$  (o valor  $z_{cal}$ ) del estadístico (distancia de nuestra media)
- ➡ Comparar ambos valores
- ➡ Si  $|z_{exp}| < z_{crit}$  se acepta  $H_0$ . En caso contrario se rechaza  $H_0$  (y se acepta  $H_1$ )

➡  $H_0 : \mu = \mu_0 ; H_1 : \mu \neq \mu_0$   
 ➡ Nivel  $\alpha = 0,05$

➡  $z_{crit} = 1,96$   
 ➡  $|z|_{exp} = \frac{|\bar{x} - \mu_0|}{\sigma/\sqrt{n}} = \frac{|49,3 - 50,0|}{0,82/\sqrt{4}} = 1,70$

➡  $|z_{exp}| < z_{crit}$   
 ➡ Se acepta  $H_0$  y  $\mu = \mu_0$

Trabajamos con medias



¡No hemos probado que  $H_0$  sea cierta, sino sólo que **no tenemos evidencias para rechazarla!**

## PLANTEAMIENTO DE LAS PRUEBAS

Se determina el  $\text{CaCO}_3$  de una muestra de un MRC (contiene 50,00 mg)  
 Se hacen 6 determinaciones ( $n=6$ ). La media  $\bar{x} = 49,10$  y  $s = 0,80$

Caso 2

### Método 1°

$H_0 : \mu = \mu_0 ; H_1 : \mu \neq \mu_0$   
 Nivel  $\alpha = 0,05$

$\mu = \bar{x} \pm t_{\alpha/2, (n-1)} (s/\sqrt{n}) = 49,10 \pm 2,57 * 0,80 / \sqrt{6} = 49,10 \pm 0,84$  es decir:  $48,26 \leq \mu \leq 49,94$

$\mu_0$  (50,00) no está dentro de ese intervalo

Se rechaza  $H_0$  y se acepta  $H_1$  y concluiremos que  $\mu \neq \mu_0$

Determinada por nosotros

### Método 2°

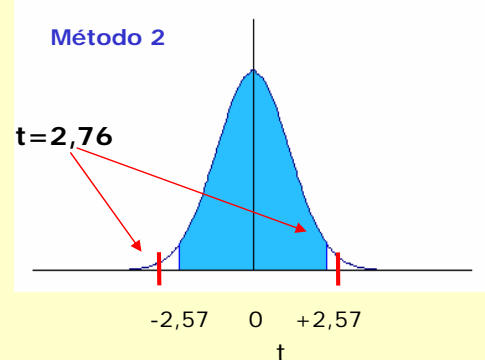
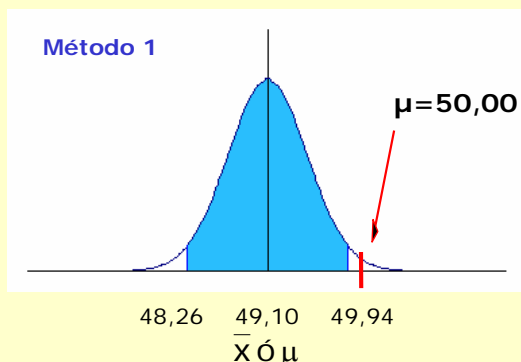
$H_0 : \mu = \mu_0 ; H_1 : \mu \neq \mu_0$   
 Nivel  $\alpha = 0,05$

$t_{crit} = 2,57$  (2 colas, 5 g.d.l.)

$|t|_{exp} = \frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} = \frac{|49,1 - 50,0|}{0,80/\sqrt{6}} = 2,76$

Se rechaza  $H_0$  y se acepta  $H_1$  y concluiremos que  $\mu \neq \mu_0$

Trabajamos con medias



## PLANTEAMIENTO DE LAS PRUEBAS

### Probabilidad "a posteriori"

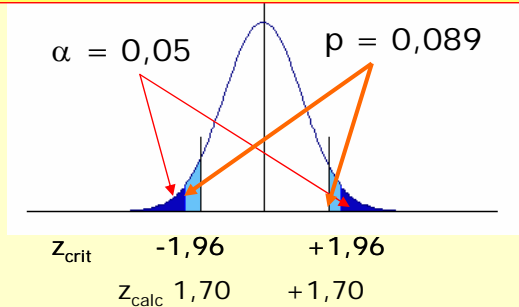
Hasta ahora la prueba de significación sólo nos indica si  $H_0$  se mantiene o no.

Para ver con qué fuerza se mantiene (o rechaza)  $H_0$  se calcula el **nivel de significación a posteriori** (denominado **p** para diferenciarlo del nivel de significación  $\alpha$  a priori).

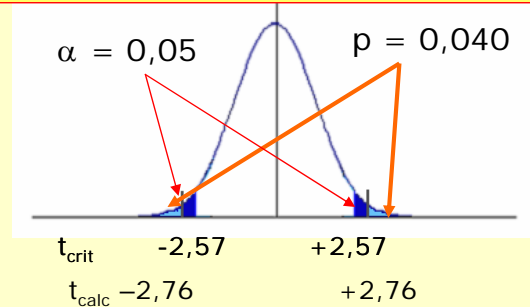
p es la probabilidad de obtener un valor de z o de t igual al  $|\text{valor}|_{\text{exp}}$  o  $|\text{valor}|_{\text{cal}}$ . Por tanto, es el **nivel de significación hasta el que se puede mantener  $H_0$**

**Cuanto mayor es p respecto de  $\alpha$  con más fuerza se mantiene  $H_0$**

**Cuanto menor es p respecto de  $\alpha$  con más fuerza se rechaza  $H_0$  y se acepta  $H_1$**



**Si  $p > 0,05$  se mantiene  $H_0$**

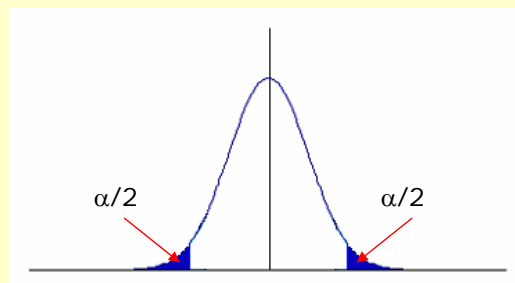


**Si  $p < 0,05$  se rechaza  $H_0$**

## ERRORES DE TIPO I Y II

### Error de tipo I o error alfa

- Es la probabilidad de rechazar la  $H_0$  cuando es cierta.
- Esto es lo mismo que afirmar equivocadamente que  $\bar{x}$  no pertenece a la población de  $\mu_0$  cuando sí que pertenece a ella.
- El error de tipo I (error  $\alpha$ ) coincide con el nivel  $\alpha$  de significación a priori.



- Dada una situación en la que  $H_0$  es cierta y con un nivel de significación  $\alpha$  a priori de 0,05, de cada 100 veces que hagamos la prueba de significación, en 5 de ellas rechazaremos la  $H_0$  a pesar de ser cierta y **NOS EQUIVOCAREMOS.**

$H_0$  : Ensayo negativo

$H_1$  : Ensayo positivo



**Error de tipo I : Falso positivo**

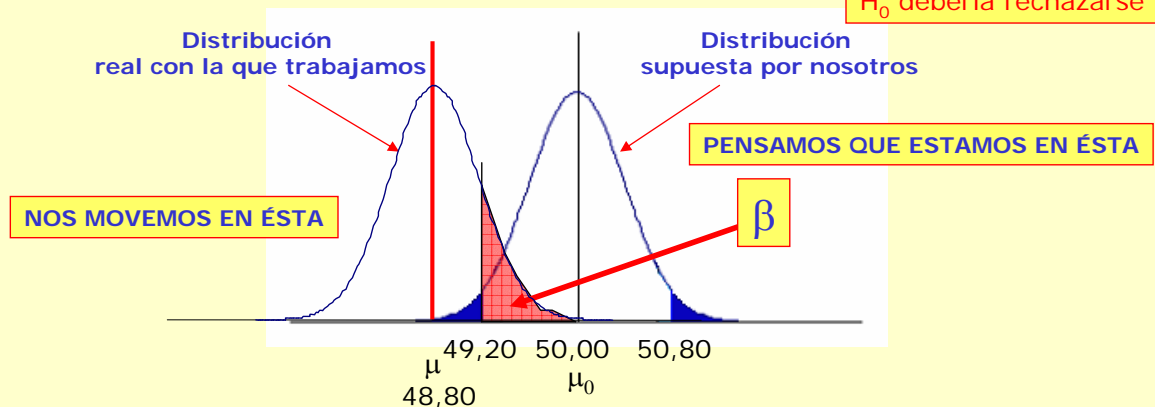
## ERRORES DE TIPO I Y II

### Error de tipo II o error beta

- Es la probabilidad de aceptar la  $H_0$  cuando es falsa
- Sea el Caso 1 en que  $H_0$  es cierta. En ese caso, los valores de  $\bar{x}$  que deberíamos obtener deberían estar (con un 95% de probabilidades) dentro del intervalo

$$\mu_0 \pm \frac{1,96\sigma}{\sqrt{n}} = 50,00 \pm \frac{1,96 \cdot 0,82}{\sqrt{4}} = 50,00 \pm 0,80$$

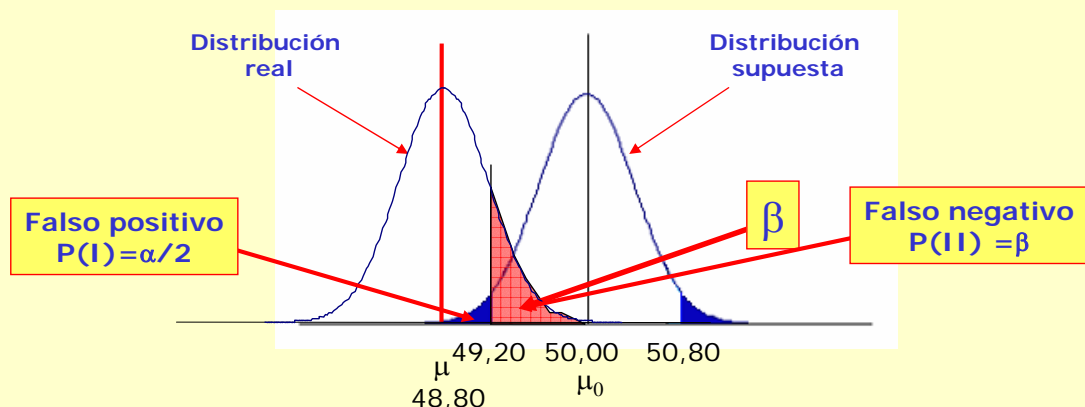
- Cualquier valor de  $\bar{x}$  que obtuviéramos fuera del intervalo 49,20 – 50,80, haría que rechazáramos  $H_0$  (aceptando  $H_1$ ) cometiendo por tanto un error de tipo I o error  $\alpha$ .
- Supongamos que nuestro método de medida **tiene un bias**, desconocido previamente por nosotros, de -1,20 de manera que la media de nuestras determinaciones esté realmente centrada en  $50,00 - 1,20 = 48,80$



- Si encontramos un valor de  $\bar{x}$  mayor de 49,20 aceptaremos  $H_0$  **a pesar de que es falsa** (hay un bias). Cometemos un **Error de tipo II o Error  $\beta$ : Falso negativo**

## ERRORES DE TIPO I Y II

### Error de tipo II o error beta



Para calcular  $P(II) = \beta$ , hay que conocer la cuantía del bias (-1,20) y calcular la probabilidad en la distribución con bias, centrada en  $(50,00 - 1,20) = 48,80$ , de encontrar un valor mayor de 49,20 (límite inferior de la distribución sin bias).

$$z = \frac{(z_{\alpha/2} - \mu_{\text{bias}})}{\sigma / \sqrt{n}} = \frac{(49,20 - 48,80)}{0,82 / \sqrt{4}} = 0,976$$

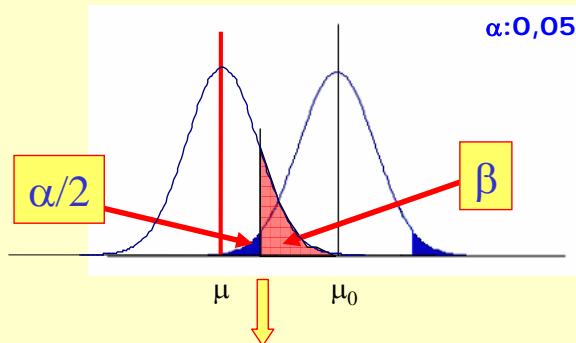
EJEMPLO L3

La probabilidad de encontrar un  $z > 0,976$  se determina en la tabla 3.3 o por medio de EXCEL:  $1 - \text{DISTR.NORM.ESTAND}(0,976) = 0,165$ . Luego  $P(II) = \beta = 16,5 \%$

## ERRORES DE TIPO I Y II

Los errores de tipo I y II están relacionados entre sí.

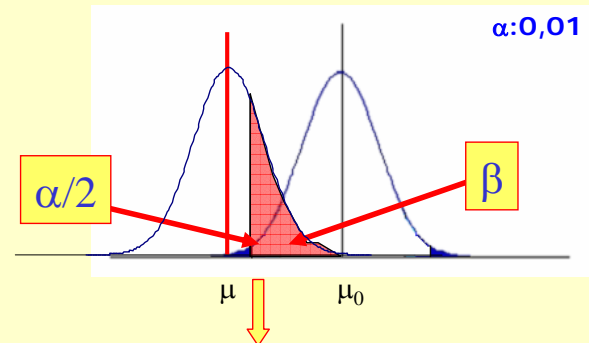
Para un mismo bias, una disminución en  $\alpha$  origina un aumento en  $\beta$ .



$$z = \frac{(z_{\alpha/2} - \mu_{\text{bias}})}{\sigma / \sqrt{n}} = \frac{(49,20 - 48,80)}{0,82 / \sqrt{4}} = 0,976$$

$$P(\text{II}) = \beta = 0,165$$

$$1 - \text{DISTR.NORM.ESTAND}(0,976)$$



$$z = \frac{(z_{\alpha/2} - \mu_{\text{bias}})}{\sigma / \sqrt{n}} = \frac{(48,95 - 48,80)}{0,82 / \sqrt{4}} = 0,336$$

$$P(\text{II}) = \beta = 0,357$$

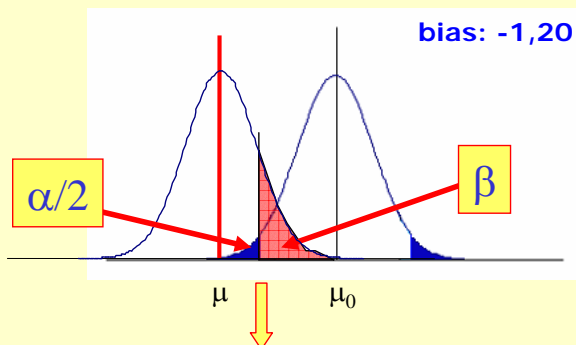
$$1 - \text{DISTR.NORM.ESTAND}(0,336)$$

Se necesitan muy buenas razones para disminuir  $\alpha$  ya que ello aumenta la probabilidad de cometer un error  $\beta$ , es decir de que un bias pase **desapercibido**.

## ERRORES DE TIPO I Y II

Los errores de tipo I y II están relacionados entre sí.

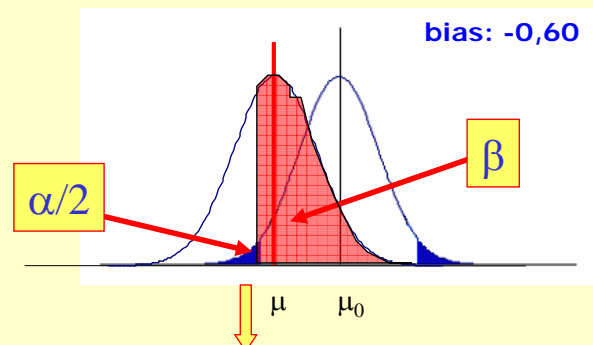
Para un mismo alfa, cuanto menor es el bias (diferencia entre  $\mu_1$  y  $\mu_0$ ) mayor es el error  $\beta$ , y viceversa.



$$z = \frac{(z_{\alpha/2} - \mu_{\text{bias}})}{\sigma / \sqrt{n}} = \frac{(49,20 - 48,80)}{0,82 / \sqrt{4}} = 0,976$$

$$P(\text{II}) = \beta = 0,165$$

$$1 - \text{DISTR.NORM.ESTAND}(0,976)$$



$$z = \frac{(z_{\alpha/2} - \mu_{\text{bias}})}{\sigma / \sqrt{n}} = \frac{(49,20 - 49,40)}{0,82 / \sqrt{4}} = -0,488$$

$$P(\text{II}) = \beta = 0,687$$

$$1 - \text{DISTR.NORM.ESTAND}(-0,488)$$

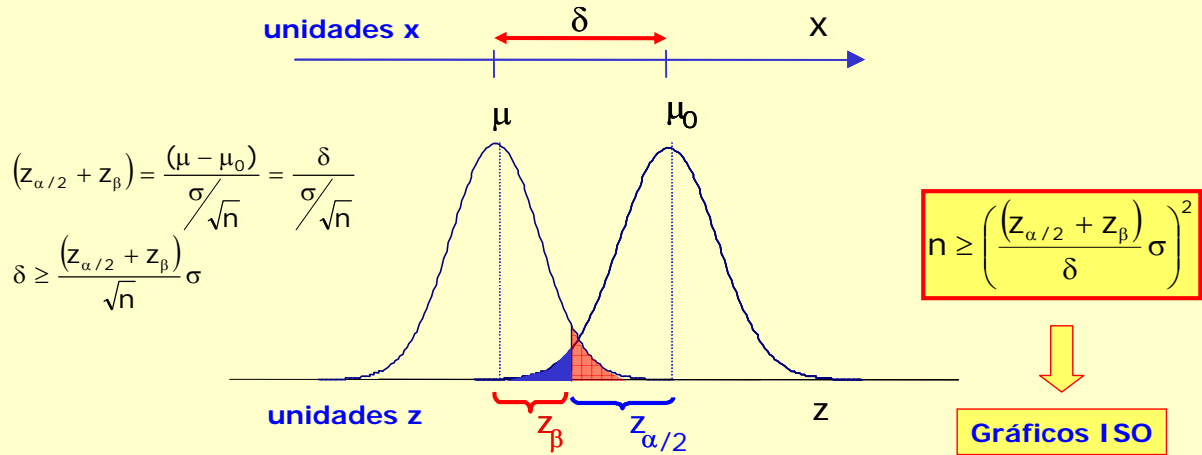
Por tanto, cuanto **menor es un bias, más difícil es su detección** y más posibilidades tenemos de confundirnos y cometer un error  $\beta$ .

## ERRORES DE TIPO I Y II

### Importancia del tamaño de muestra

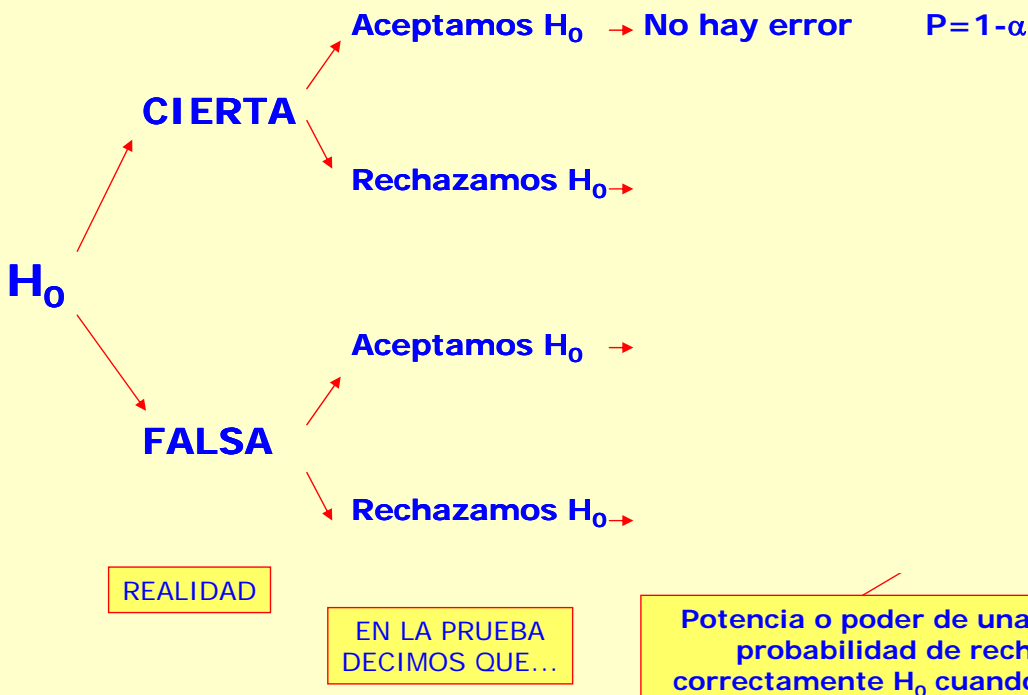
- El intervalo de confianza de cualquier media puede disminuirse aumentando el tamaño de la muestra  $\mu = \bar{x} \pm 1,96\sigma / \sqrt{n}$
- Se puede calcular el tamaño  $n$  de la muestra que debe ser estudiado para llevar a cabo una prueba de significación con la suficiente confianza, es decir para poder detectar una diferencia (bias)  $|\mu - \mu_0|$  mediante un procedimiento que tiene una precisión  $\sigma$  (o  $s$ ), de manera que haya no más de un  $\alpha\%$  de probabilidades de decidir que hay diferencia (bias) cuando no la hay (error tipo I), y al mismo tiempo una probabilidad de no más que  $\beta\%$  de no detectar la diferencia (bias) cuando sí que existe (error tipo II)
- Por tanto, se deben conocer el bias a detectar,  $\delta = \mu - \mu_0$ , así como  $\sigma$

### NOS PROTEGEMOS SIMULTÁNEAMENTE CONTRA AMBOS TIPOS DE ERROR



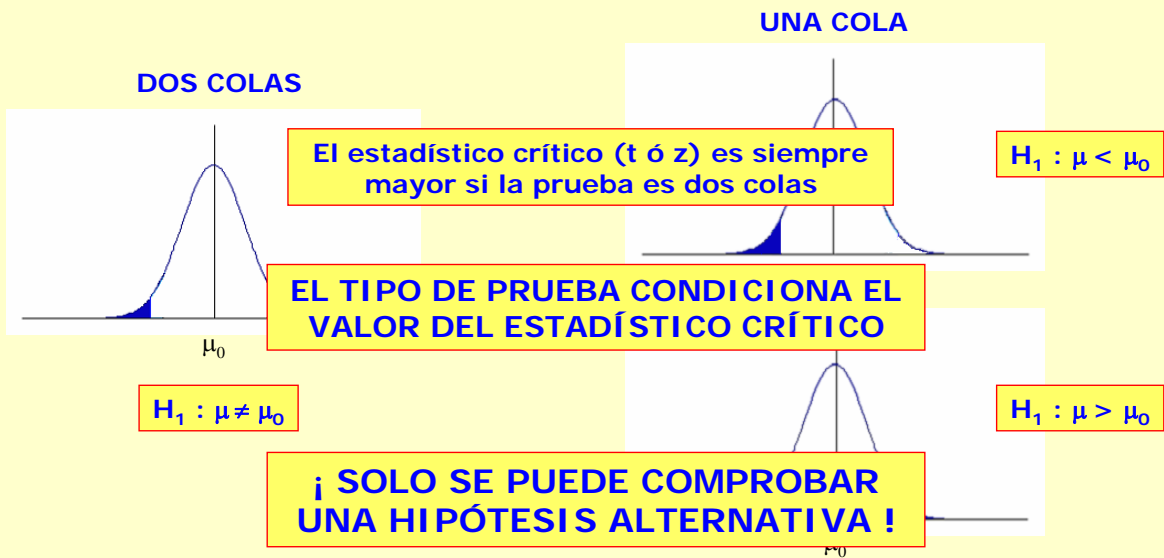
## ERRORES DE TIPO I Y II

### Resumen de posibilidades



## PRUEBA (TEST) DE UNA O DOS COLAS

- Hasta ahora,  $H_0$  debe rechazarse tanto si  $\mu$  es significativamente mayor que  $\mu_0$ , como si  $\mu$  es significativamente menor. La hipótesis alternativa es pues  $H_1 : \mu \neq \mu_0$ .
- La prueba de significación (o test de hipótesis) es de **dos colas** ya que se exploran ambos lados de la distribución
- En ocasiones, solo estamos interesados en saber si  $\mu$  es significativamente mayor (o significativamente menor) que  $\mu_0$  o no. En ese lado la prueba de significación (o test de hipótesis) es de una cola pues solo se explora un lado de la distribución.
- Por tanto, la hipótesis nula siempre es  $H_0 : \mu = \mu_0$ . Para  $H_1$  hay más posibilidades:



## PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

### Pruebas para evaluar la exactitud (detección de bias)

Comparación de una media experimental con un valor conocido			
$H_0: \bar{X} = \mu_0$	Estadístico	$H_1: \bar{X} \neq \mu_0$	$H_1: \bar{X} > \mu_0$ Ó $\bar{X} < \mu_0$
$\sigma$ conocida	$z = \frac{ \bar{X} - \mu_0 }{\sigma/\sqrt{n}}$	$z_{\text{crit}}\left(1-\frac{\alpha}{2}\right)$	$z_{\text{crit}(1-\alpha)}$
$\sigma$ desconocida $n > 30$	$z = \frac{ \bar{X} - \mu_0 }{s/\sqrt{n}}$	$z_{\text{crit}}\left(1-\frac{\alpha}{2}\right)$	$z_{\text{crit}(1-\alpha)}$
$\sigma$ desconocida $n < 30$	$t = \frac{ \bar{X} - \mu_0 }{s/\sqrt{n}}$	$t_{\text{crit}}\left(1-\frac{\alpha}{2}, v\right)$ $v = n - 1$	$t_{\text{crit}(1-\alpha, v)}$ $v = n - 1$

#### Dos colas

DISTR.NORM.ESTAND.INV( $\alpha/2$ )  
DISTR.T.INV( $\alpha; v$ )

*p a posteriori*  
 $2 * (1 - \text{DISTR.NORM.ESTAND}(z))$   
DISTR.T( $z; v; 2$ )

#### Una cola

DISTR.NORM.ESTAND.INV( $\alpha$ )  
DISTR.T.INV( $2\alpha; v$ )

*p a posteriori*  
 $1 - \text{DISTR.NORM.ESTAND}(z)$   
DISTR.T( $z; v; 1$ )

## PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

### Pruebas para evaluar la exactitud (detección de bias)

Comparación de dos medias				
$H_0: \bar{X}_1 = \bar{X}_2$	<b>Estadístico</b>	$H_1: \bar{X}_1 \neq \bar{X}_2$	$H_1: \bar{X}_1 < \bar{X}_2$ ó $\bar{X}_1 > \bar{X}_2$	
$\sigma_1, \sigma_2$ conocidas	$Z = \frac{ \bar{X}_1 - \bar{X}_2 }{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	$Z_{\text{crit}}\left(1-\frac{\alpha}{2}\right)$	$Z_{\text{crit}}$	<b>Dos colas</b> DISTR.NORM.ESTAND.INV( $\alpha/2$ ) DISTR.T.INV( $\alpha; v$ ) <i>p a posteriori</i> 2*(1-DISTR.NORM.ESTAND(z)) DISTR.T(t; v; 2)
$\sigma_1, \sigma_2$ desconocidas $n_1, n_2 > 30$	$Z = \frac{ \bar{X}_1 - \bar{X}_2 }{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$Z_{\text{crit}}\left(1-\frac{\alpha}{2}\right)$	$Z_{\text{crit}}$	<b>Una cola</b> DISTR.NORM.ESTAND.INV( $\alpha$ ) DISTR.T.INV( $2\alpha; v$ ) <i>p a posteriori</i> 1-DISTR.NORM.ESTAND(z) DISTR.T(t; v; 1)
$\sigma_1, \sigma_2$ desconocidas $S_1, S_2$ comparables $n_1, n_2 < 30$	$t = \frac{ \bar{X}_1 - \bar{X}_2 }{\sqrt{S^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}}$	$t_{\text{crit}}\left(1-\frac{\alpha}{2}, v\right)$	$t_{\text{crit}}$	$(n_1 - 1) + (n_2 - 1)$
$\sigma_1, \sigma_2$ desconocidas $S_1, S_2$ no comparables $n_1, n_2 < 30$	$t = \frac{ \bar{X}_1 - \bar{X}_2 }{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	$t_{\text{crit}}\left(1-\frac{\alpha}{2}, v\right)$	Test de Cochran $t_{\text{crit}}^* = \frac{t_1 \frac{S_1^2}{n_1} + t_2 \frac{S_2^2}{n_2}}{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ $t_1 = t_{\text{crit}}(v_1 = n_1 - 1) ; t_2 = t_{\text{crit}}(v_2 = n_2 - 1)$	

## PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

### Pruebas para evaluar la exactitud (detección de bias)

Se corresponden una a una

Comparación de pares de valores (medidas apareadas)				
$d_i = X_{1i} - X_{2i}$	$H_0: \mu_d = 0$	$H_1: \mu_d \neq 0$	$H_1: \mu_d < 0$ ó $\mu_d > 0$	
$n > 30$	$Z = \frac{ \bar{d} - \mu_d }{s_d / \sqrt{n}} = \frac{ \bar{d} }{s_d / \sqrt{n}}$	$Z_{\text{crit}}\left(1-\frac{\alpha}{2}\right)$	$Z_{\text{crit}}(1-\alpha)$	
$n < 30$	$t = \frac{ \bar{d} - \mu_d }{s_d / \sqrt{n}} = \frac{ \bar{d} }{s_d / \sqrt{n}}$	$t_{\text{crit}}\left(1-\frac{\alpha}{2}, v\right)$	$t_{\text{crit}}(1-\alpha, v)$	$v = n - 1$

**Dos colas**  
 DISTR.NORM.ESTAND.INV( $\alpha/2$ )  
 DISTR.T.INV( $\alpha; v$ )  
*p a posteriori (dos colas)*  
 2\*(1-DISTR.NORM.ESTAND(z))  
 DISTR.T(t; v; 2)

**Una cola**  
 DISTR.NORM.ESTAND.INV( $\alpha$ )  
 DISTR.T.INV( $2\alpha; v$ )  
*p a posteriori*  
 1-DISTR.NORM.ESTAND(z)  
 DISTR.T(t; v; 1)

## PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

### Pruebas para evaluar la precisión

A veces se usa como paso previo a la prueba de comparación de medias

Comparación de dos varianzas: Prueba F de Fischer			
$H_0: \sigma_1^2 = \sigma_2^2$ $H_1: \sigma_1^2 > \sigma_2^2$ (test de 1 cola)			
$S_1^2 \geq S_2^2$	$F = \frac{S_1^2}{S_2^2}$	$F_{\text{crit}(1-\alpha, v_1, v_2)}$	$v_1 = n_1 - 1$ $v_2 = n_2 - 1$

*Tabla F de una cola*  
**DISTR.F.INV( $\alpha; v_1; v_2$ )**  
*p a posteriori*  
**DISTR.F(F;  $v_1; v_2$ )**

## PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

### Prueba de rechazo de valores discrepantes (outliers)

Prueba Q de Dixon		
$H_0$ : el valor sospechoso $\in N(\mu, \sigma^2)$		
$H_1$ : el valor sospechoso $\notin N(\mu, \sigma^2)$ (test de 2 colas)		
$Q = \frac{ X_{\text{sospechoso}} - X_{\text{próximo}} }{X_{\text{mayor}} - X_{\text{menor}}}$	$Q_{\text{crit}(1-\frac{\alpha}{2}, v)}$	$v = n$

# PRUEBAS DE SIGNIFICACION ESTADISTICA (TESTS DE HIPOTESIS)

## Prueba de normalidad

**Sirve para cualquier tipo de distribuciones**

<b>Prueba chi-cuadrado (<math>\chi^2</math>)</b>		
<b>H<sub>0</sub>:</b> $X \in N(\mu, \sigma^2)$	<b>H<sub>1</sub>:</b> $X \notin N(\mu, \sigma^2)$	(test de 2 colas)
$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$	$\chi^2_{crit(1-\alpha, v)}$	$v = k - 1 - h$ k, número de clases h, número de parámetros de la distribución normal (2: $\mu$ y $\sigma^2$ )
Calcular la media y la desviación estándar de los datos experimentales Dividir los datos experimentales en k clases, siendo $k \approx n/5$ Calcular las frecuencias $O_i$ de cada clase ( $\sum O_i = n$ ). Transformar los valores límite de cada clase en valores z: $z = \frac{LS - \bar{x}}{s}$ Calcular la frecuencia acumulada y la frecuencia relativa de cada clase. Calcular las frecuencias esperadas $E_i$ para cada clase. Calcular para cada clase el valor $\frac{(O_i - E_i)^2}{E_i}$ Calcular $\chi^2_{exp}$ como suma de los valores anteriores		

**Tabla  $\chi^2$**   
**PRUEBA.CHI.INV( $\alpha; v$ )**  
*p a posteriori*  
**DISTR.CHI ( $\chi^2; v$ )**