

# Unidad IV

## Prueba de Hipótesis Parte I

¿¿¿Con cuál me quedo????

**Hipótesis Nula:**  
No hay diferencias

**Hipótesis Alternativa:**  
Hay diferencias significativas



Pistonesi Silvana

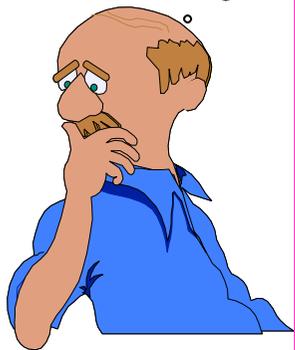
## Prueba de hipótesis

---

En la practica con frecuencia tenemos que tomar decisiones relativas a una población sobre la base de la información proveniente de una muestra

Para ello es útil plantear hipótesis estadísticas (conjeturas) sobre la población implicada y estas pueden o no ser ciertas

Que decido????



Una **hipótesis estadística** es una proposición que puede o no ser verdadera pero que se adopta provisionalmente hasta recabar información que sugiera lo contrario. Si hay inconsistencia, se rechaza la hipótesis. Las **pruebas de hipótesis** se usan precisamente para evaluar el **grado de esa inconsistencia**.

# Prueba de hipótesis

## Def:

**Prueba de hipótesis** es una metodología estadística (**inferencial**) que sirve para probar la validez de una afirmación acerca del valor de uno ó más parámetros de una población. Esta metodología permite rechazar o no la hipótesis propuesta en base a una muestra aleatoria de la población.

## Ejemplo 1.

Un empresario está considerando la posibilidad de ampliar su negocio mediante la adquisición de un pequeño bar. El dueño actual del bar afirma que el **ingreso promedio diario** del establecimiento es **\$15375**. Pero el empresario en realidad sospecha que el ingreso promedio diario *es inferior* al que afirma el actual dueño. Para comprobar la validez de la sospecha del empresario, se decide tomar una muestra de **treinta días**. Si por ejemplo, observa un ingreso diario promedio muestral,  $\bar{x} = \$15250$ , es probable que el empresario no rechace la posibilidad

de adquirir el bar, teniendo en cuenta la afirmación supuesta. Mientras que si la media muestral fuese **\$13900** estaría predispuesto a rechazar oportunidad.



## Prueba de hipótesis

---

En un **test de hipótesis** se establecen **2 teorías o hipótesis estadísticas**, afirmaciones cuantitativas acerca del o los parámetros poblacionales, siendo ambas incompatibles, antagónicas (si una es cierta, la otra necesariamente ha de ser falsa).

- **hipótesis nula** y se nota  $H_0$ , a aquella hipótesis que se supone cierta hasta que la evidencia muestral haga pensar lo contrario. Se tratará de determinar hasta qué grado las observaciones registradas son coherentes con  $H_0$ . Es la hipótesis en donde **se establece la igualdad** del parámetro al **valor propuesto**.

- **hipótesis alternativa**, también llamada **hipótesis de investigación** debido a que se trata de la hipótesis que el investigador desea comprobar con base en la información contenida en una muestra. La hipótesis alternativa representa alguna forma de **negación** de la **hipótesis nula**. Esta será admitida cuando  $H_0$  sea rechazada. Generalmente se nota  $H_1$  ó  $H_a$ .

## Prueba de hipótesis



**OJO!!!!!!**

El planteo de las hipótesis estadísticas se efectúa siempre respecto a un **parámetro** (población), y no respecto a un **estadístico** (muestra).

### Ejemplo 2.

El **número promedio** de smart tv en los hogares de Bahía Blanca es mayor a **2**:

$$H_0: \mu = 2$$

$$H_1: \mu > 2$$

~~$$H_0: \bar{X} = 2$$~~

~~$$H_1: \bar{X} > 2$$~~

**Mal !!!**

## Prueba de hipótesis

---

Sólo en caso de que haya fuertes indicios de incompatibilidad entre el supuesto de que  $H_0$  sea cierta y los datos obtenidos empíricamente, descartaremos  $H_0$  como hipótesis de trabajo y en su lugar tomaremos como cierta la hipótesis alternativa  $H_1$ .

En nuestro **Ejemplo 1**, las hipótesis serían:

Eje

$$H_0: \mu = \$15375 \text{ (ó } \mu \geq \$15375\text{).}$$

$$H_1: \mu < \$15375.$$



# Prueba de hipótesis

## Tipos de pruebas

- **Bilaterales** es una prueba donde la **hipótesis alternativa** establece que el valor del parámetro de interés es **distinto** del valor indicado por  $H_0$ .
- **Unilateral** son aquellas pruebas cuya **hipótesis alternativa** establece que el parámetro poblacional es **sólo mayor o sólo menor** que el valor especificado en la hipótesis nula.

### Prueba Bilateral

$$H_0: \theta = \theta_0$$

$$H_1: \theta \neq \theta_0$$

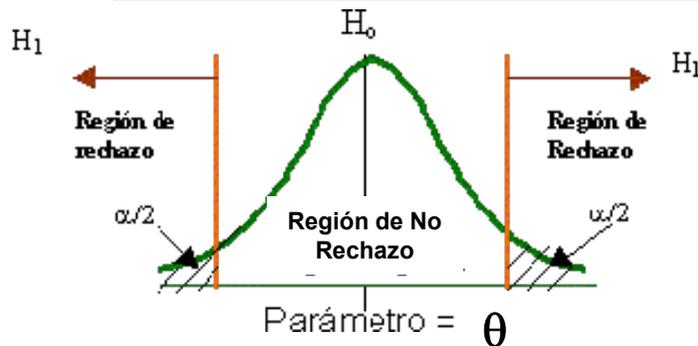
### Pruebas Unilaterales

$$H_0: \theta = \theta_0 \text{ (ó } \theta \leq \theta_0)$$

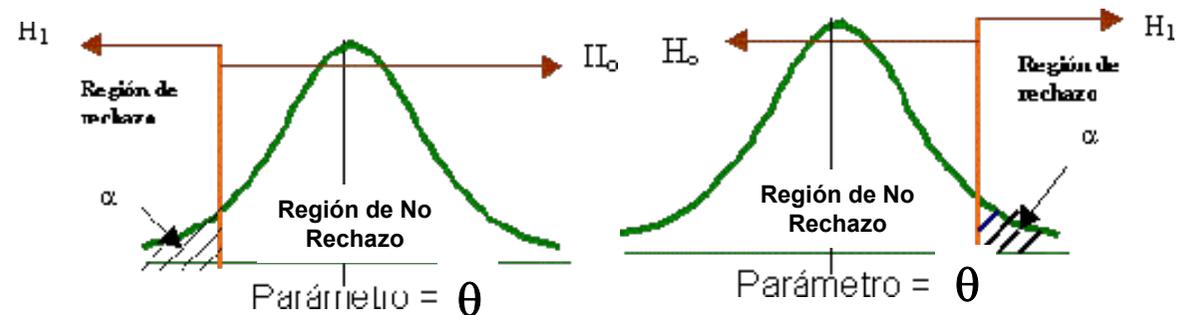
$$H_1: \theta > \theta_0$$

$$H_0: \theta = \theta_0 \text{ (ó } \theta \geq \theta_0)$$

$$H_1: \theta < \theta_0$$



Prueba Bilateral



Pruebas Unilaterales

## Prueba de hipótesis

Debido a que toda decisión estará basada sobre la información brindada de una **muestra aleatoria** de la población habrá siempre una posibilidad de tomar una **decisión incorrecta**.

**Situaciones posibles que pueden surgir en un test de hipótesis**

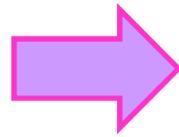
Realidad \ Decisión Posible	$H_0$ es Verdadera	$H_0$ es Falsa
	No rechazar $H_0$	Decisión correcta
Rechazar $H_0$	Decisión incorrecta <b>Error de Tipo I</b>	Decisión correcta

▪ Si  $H_0$  es verdadera y tomo la decisión de rechazar  $H_0$  se comete el **error tipo I**

▪ El **error tipo II** ocurriría si siendo falsa  $H_0$  se toma la decisión de no rechazarla.

La idea es que sean **poco probables** de darse. Serían **nulos** si la decisión se basase en un examen de la **población completa**, lo que resulta impráctico.

Si este fuese el planteo



$H_0$ : el acusado es **inocente**  
 $H_1$ : el acusado es **culpable**

Decisión Posible (a través de información muestral)	Realidad	$H_0$ es Verdadera (es inocente)	$H_0$ es Falsa (es culpable)
	No rechazar $H_0$	Decisión correcta	Decisión incorrecta <b>Error de Tipo II</b>
Rechazar $H_0$	Decisión incorrecta <b>Error de Tipo I</b>	Decisión correcta	

## Prueba de hipótesis: tipos de errores

**Error tipo I** = rechazar  $H_0$  /  $H_0$  es verdadera = suponer que el acusado es **culpable**, cuando en realidad es **inocente**.

**Error tipo II** = No rechazar  $H_0$  /  $H_0$  es falsa = suponer que el acusado es **inocente**, cuando en realidad es **culpable**.

## Prueba de hipótesis: errores

---

Dado que descartaremos o no la hipótesis nula a partir de datos muestrales (es decir, **no dispondremos de información completa sobre la población**), **no** será posible **garantizar** que la **decisión** tomada sea la **correcta**.

Lo que **sí podremos** hacer es **controlar** la **probabilidad de cometer un error**.

La **probabilidad de cometer un Error de Tipo I**, se denota con  $\alpha$ ,  
y  
la **probabilidad de cometer un Error de Tipo II** por  $\beta$ .

Simbólicamente se puede expresar:

$$P(\text{Error de Tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ verdadera}) = \alpha$$

$$P(\text{Error de Tipo II}) = P(\text{No rechazar } H_0 / H_0 \text{ falsa}) = \beta$$

## Prueba de hipótesis: errores

$$P(\text{Error de Tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ verdadera}) = \alpha$$

$$P(\text{Error de Tipo II}) = P(\text{No rechazar } H_0 / H_0 \text{ falsa}) = \beta$$

En el Ejemplo 1, respecto de la adquisición del bar:

Planteo



$$H_0: \mu = \$15375 \text{ (ó } \mu \geq \$15375\text{)}.$$

$$H_1: \mu < \$15375.$$

- Se comete un **error de Tipo I**: si se **rechaza la hipótesis nula** de  $\mu = \$15375$  de acuerdo a los resultados de la prueba, pero realmente el valor medio es de  $\$15375$ , ie, **suponer que el ingreso promedio diario es inferior al que afirma el actual dueño, cuando en realidad es de  $\$15375$ .**
- Se comete un **error de Tipo II**: si la hipótesis  $H_0: \mu = \$15375$  **no es rechazada** según los datos de la prueba, pero realmente el valor medio es menor a  $\$15375$ , ie, **suponer que el ingreso promedio diario es de  $\$15375$ , cuando en realidad es inferior al que afirma el actual dueño.**

## Prueba de hipótesis: errores

---

### Ejemplo 3.

Estudios pasados han determinado que la duración media de supervivencia de los pacientes afectados por cierta enfermedad es de 3,4 meses. Un investigador afirma que una nueva droga prolonga la vida de esos pacientes. Se desea comprobar si la droga es efectiva.

$H_0: \mu = 3,4$  meses (droga no es efectiva)

$H_1: \mu > 3,4$  meses (droga es efectiva)

**Error tipo I** = suponer que la duración media de supervivencia en los pacientes que padecen esta enfermedad es **superior a 3,4 meses**, cuando en realidad es de **a lo sumo 3.4 meses**.

**Error tipo II** = suponer que la duración media de supervivencia en los pacientes que padecen esta enfermedad es de **a lo sumo 3.4 meses** cuando en realidad es **superior a 3.4 meses**.

## Prueba de hipótesis: errores

---

### Ejemplo 4.

El gerente de una empresa de servicio de internet establece que la velocidad de descarga (downstream) media del nuevo plan que ofrecen es de 50Mbps. Un usuario que contrató el servicio afirma que la velocidad de descarga media del nuevo plan es menor que 50Mbps. Se desea comprobar si es veraz la opinión del usuario.

$$H_0: \mu = 50 \text{ Mbps}$$

$$H_1: \mu < 50 \text{ Mbps}$$

**Error tipo I** = suponer que la velocidad de descarga media es **menor de 50Mbps**, cuando en realidad es de **50Mbps**.

**Error tipo II** = suponer que la velocidad de descarga media es **de 50Mbps** cuando en realidad es **inferior a 50 Mbps**.

$$\alpha = P(\text{error de Tipo I})$$

se conoce también como **nivel de significación**. Es la máxima probabilidad que un investigador esta dispuesto a correr el riesgo de cometer el **error tipo I, si se rechaza  $H_0$** . Valores habituales de  $\alpha$ : **0.05, o 0.01, o 0.10**.

El enfoque general en Prueba de Hipótesis estadísticas, **es aceptar** la premisa de que el **error de Tipo I es más serio que el error de Tipo II** y por lo tanto construir la prueba de modo que  $\alpha$  sea pequeña.

El complemento de  $\beta$  es la probabilidad:

$$1 - \beta = P(\text{Rechazar } H_0 / H_0 \text{ falsa}) = \gamma$$

Es la **potencia de la prueba**, se **denota** con la letra griega gamma:  $\gamma$ . Representa la capacidad que tiene la prueba de reconocer correctamente que la hipótesis nula es falsa, ie, probabilidad de no equivocarse al rechazar  $H_0$  cuando es falsa. Poder que tiene una prueba para detectar diferencias existentes entre el valor verdadero del parámetro de la población y el valor supuesto. Es deseable que una prueba tenga una potencia grande (cercana a uno).

## Relación entre $\alpha$ , $\beta$ y $\gamma$

$\alpha \uparrow$	$\beta \downarrow$	$\gamma \uparrow$
$\alpha \downarrow$	$\beta \uparrow$	$\gamma \downarrow$
$n \uparrow$ , fijado $\alpha$	$\beta \downarrow$	$\gamma \uparrow$

Para cualquier  $\alpha$  dado, el aumento del **tamaño de la muestra** siempre produce una mayor **potencia** de la prueba estadística, por lo tanto, existe una disminución de  $\beta$ .

## Estadístico de prueba

Planteadas las hipótesis nula y alternativa, la decisión de rechazar o no  $H_0$  se basa en algún **estadístico** apropiado el cual recibe el nombre de **estadístico de prueba** y cuya **distribución de probabilidad es conocida** bajo  $H_0$ . Mide la **discrepancia** entre el valor de **estimador puntual** y el valor hipotético del **parámetro poblacional**.

## Prueba de hipótesis para la media de una población normal con varianza conocida

En el **Ejemplo 1**. del empresario que está considerando la posibilidad de ampliar su negocio mediante la adquisición de un pequeño bar, tomó una muestra de treinta días y ésta reveló un ingreso diario promedio de **\$13125**. Suponiendo que el ingreso diario del bar es v.a. **distribuida normalmente** con un **desvío estándar de \$3700** y utilizando un **nivel de significación del 5 %**, ¿Hay evidencia de que el ingreso diario promedio sea menor del que afirma el presente dueño?

El **parámetro** al cual se refiere la hipótesis nula es la **media poblacional  $\mu$** . El mejor **estimador puntual** de  $\mu$  es la media muestral  $\bar{X}$ .

## Prueba de hipótesis para la media de una población normal con varianza conocida

---

Los valores del **estadístico de prueba  $Z$** , para los cuales se **rechaza la hipótesis nula** constituyen lo que se conoce como la **región crítica o región de rechazo de  $H_0$**  de la prueba.

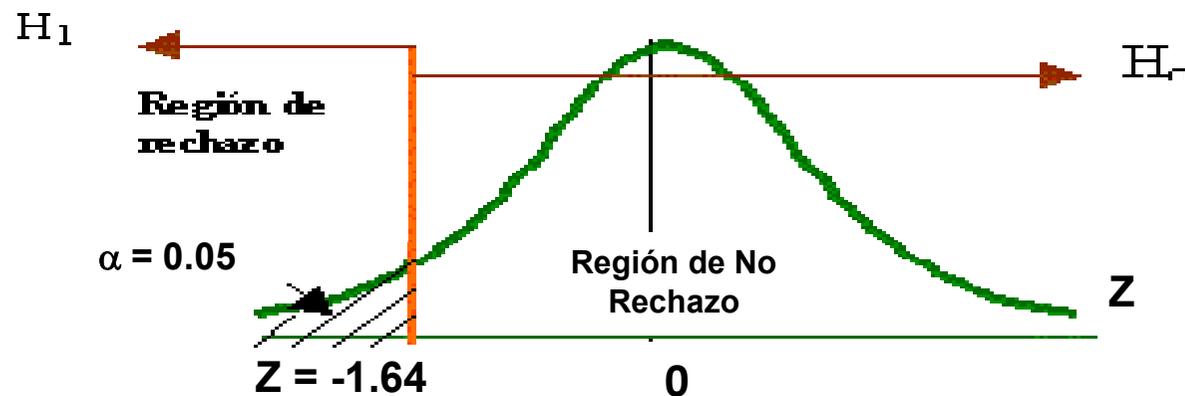
El **valor crítico** es el punto que separa la **región crítica** del conjunto de valores para los cuales **no se rechaza  $H_0$** .

Para establecer **la región crítica o zona de rechazo de  $H_0$** , se elige un valor para  **$\alpha$** . El valor de  **$\alpha$**  indica la importancia que se atribuye a las consecuencias asociadas con un rechazo incorrecto de  **$H_0$** .

En el Ejemplo 1.,  $\alpha = 0.05$ , la **región crítica** incluye el **5% del área bajo la función de densidad del estadístico** que es, en este caso, la distribución normal estándar.

Como la prueba de hipótesis es **unilateral** ( $H_1: \mu < \$15375$ ), la **región crítica** de área  $\alpha = 0.05$ , se ubica según la hipótesis  $H_1$ , en la **cola izquierda de la distribución**.

En la **tabla de la distribución Normal Estándar**, el valor de **Z** que deja un área de  $\alpha = 0.05$  a izquierda es  $z_{0.05} = -1.64$ .



Establecida la **región crítica** el próximo paso es, a partir de los datos muestrales, tomar una **decisión** (Rechazo o no rechazo  $H_0$ ) y determinar, si es posible, la probabilidad del tipo de error que se ha cometido.

## La prueba de hipótesis para la media poblacional con varianza conocida

Sea  $X$  =« Ingreso diario del establecimiento »  $X \sim N(\mu, \sigma^2 = \$^23700^2)$ , con  $\sigma^2$  conocida.

**Paso 1:** Planteo de lo que se desea probar :

$H_0: \mu = \$15375$  ( ó  $\mu \geq \$15375$  ) .

$H_1: \mu < \$15375$

**Paso 2:**  $\alpha = 0.05$  (nivel de significación fijado)

$$X \sim N(\mu, \sigma^2 = \$^23700^2), \quad \Rightarrow \quad \bar{X} \sim N(\mu, \sigma^2/n)$$

**Paso 3:** El estadístico de prueba a usar (El estadístico de prueba que se basa en la media muestral  $\bar{X}$ )

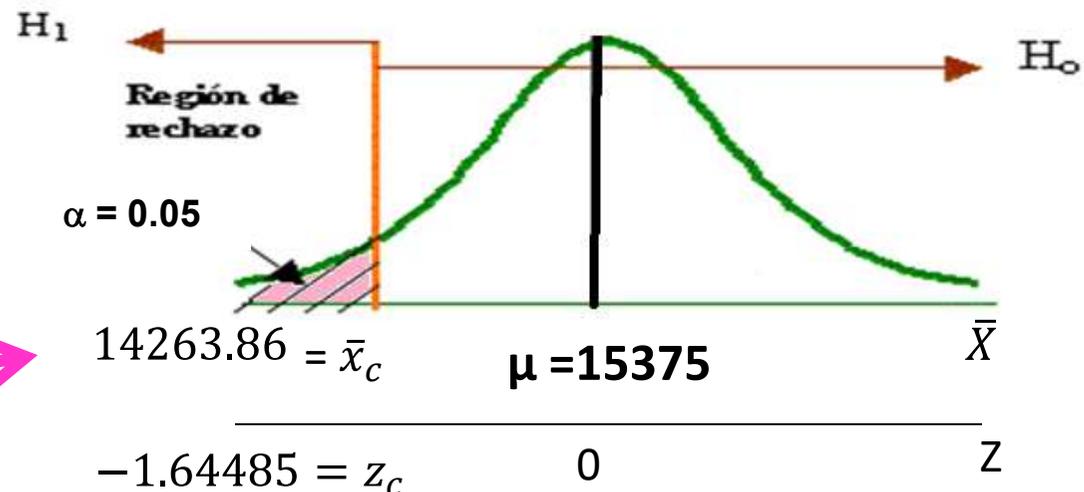
$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - 15375}{3700/\sqrt{30}} \sim N(0, 1), \text{ Bajo } H_0 \text{ cierta.}$$

**Paso 4:** Determinar la región crítica gráficamente .  
Puede hacerse en la escala de  $Z$  o de  $\bar{X}$ .

$$-z_c = -1.64485 = \frac{\bar{X} - 15375}{3700/\sqrt{30}}$$

$$14263.86 = \bar{x}_c$$

Pistonesi Silvina



## La prueba de hipótesis para la media poblacional con varianza conocida

**Paso 5:** La región crítica o de rechazo analíticamente,

$$RC = \{ z / z < z_{\alpha} = -1.64485 \}$$

$$RC = \{ \bar{X} : \bar{X} < \bar{x}_c = 14263.86 \}$$

**Paso 6 :** Valor del estadístico ( calculo del valor observado).

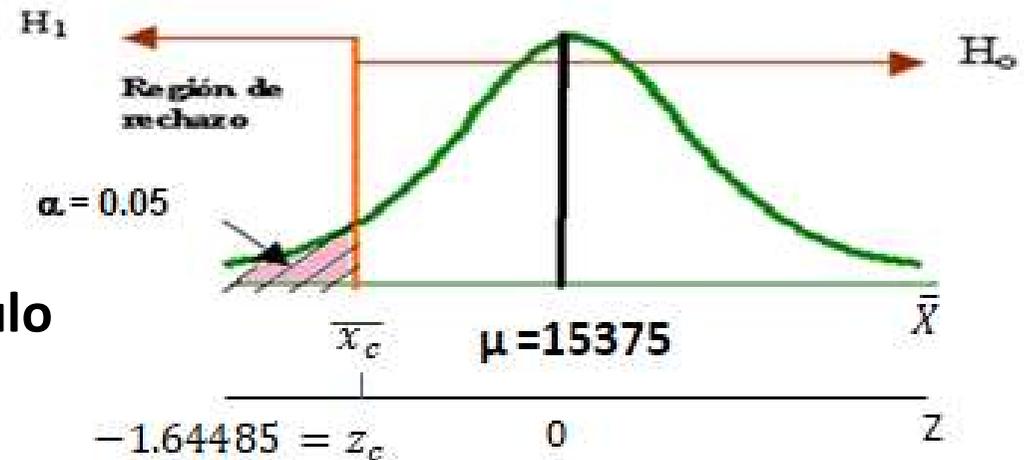
$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} = \frac{13125 - 15375}{3700 / \sqrt{30}} = \frac{-2250}{3700 / \sqrt{30}} = -3.34$$

**Paso 7 :** comparo el valor del estadístico con el valor crítico.

Como  $z_{\text{obs}} = -3.34 < -1.64485 = z_c$  entonces Rechazo  $H_0$  (la decisión tomada con la información muestral)

**Paso 8 : Conclusión:**

Tengo evidencias suficientes para afirmar que el ingreso promedio diario del bar sea inferior a \$15375, tengo pruebas que el dueño le este mintiendo al comprador, con probabilidad de error 0.05.



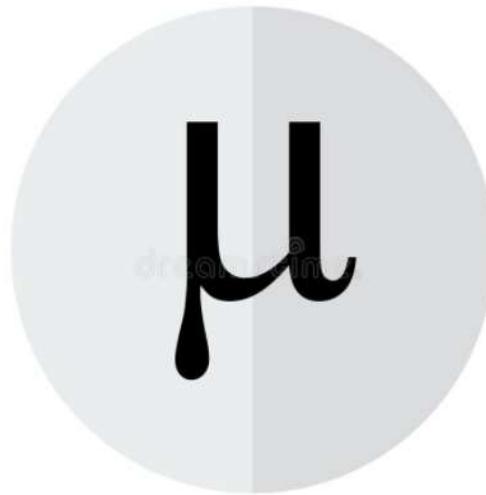
## Procedimiento estándar en la prueba de hipótesis

---

La prueba de hipótesis puede resumirse en los siguientes pasos, una vez identificado el parámetro de interés:

1. Plantear las hipótesis nula y alternativa.
2. Seleccionar un nivel de significación  $\alpha$ .
3. Determinar el estadístico de prueba apropiado y su distribución de probabilidad bajo  $H_0$ .
4. Determinar la región crítica.
5. Calcular a partir de los datos de la muestra el valor del estadístico utilizado en la prueba.
6. Comparar el/los valor/es crítico/s con el observado en la muestra. Decidir si se rechaza o no  $H_0$ .
7. Concluir interpretando los resultados en términos de la variable bajo estudio e indicando, si es posible, la probabilidad del error cometido en la decisión tomada.

# Prueba de Hipotesis para la media de una población normal con varianza conocida



# La prueba de hipótesis para la media poblacional con varianza conocida

Sea  $X \sim N(\mu, \sigma^2)$ , con  $\sigma^2$  conocida.

El estadístico de prueba que se basa en la media muestral  $\bar{X}$  es:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$$

Se desea probar  $H_0: \mu = \mu_0$  frente a una de las siguientes hipótesis alternativas posibles:

1.  $H_1: \mu \neq \mu_0$
2.  $H_1: \mu > \mu_0$
3.  $H_1: \mu < \mu_0$ .

El estadístico de prueba:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1), \text{ Bajo } H_0 \text{ cierta.}$$

La región crítica o de rechazo:

1.  $\{ z / z < z_{\alpha/2} \text{ o } z > z_{1-\alpha/2} \}$
2.  $\{ z / z > z_{1-\alpha} \}$
3.  $\{ z / z < z_{\alpha} \}$

Asociada a cada hipótesis alternativa

# Prueba de hipótesis para la media poblacional con varianza conocida

---

## Ejemplo 5.

Se realizó un estudio respecto de los salarios mensuales (en pesos) que gozan los ingenieros en computación de una empresa de desarrollo de software según su antigüedad en la misma. Para ello se eligieron al azar 10 ingenieros de la empresa y la información recabada arrojó un salario promedio de \$33100.

Se sabe que el salario mensual de un ingeniero en computación es una v.a. normalmente distribuida con un desvío estándar de \$2998.2.

A un nivel de significación del 0.05, ¿se puede concluir que el salario promedio mensual de un ingeniero difiere de 35000 pesos?



la v.a.  $X$  = “salario mensual de un ingeniero en computación de una empresa de desarrollo de software según su antigüedad en la misma ” (en pesos) ,

$$X \sim N(\mu, \sigma^2 = 2998.2^2). \quad n = 10,$$

Como  $X \sim N(\mu, \sigma^2 = 2998.2^2)$  entonces  $\bar{X} \sim N(\mu, \sigma^2 / n = 2998.2^2 / 10)$

1. Planteo:  $H_0: \mu = 35000$  pesos

$H_1: \mu \neq 35000$  pesos

2.  $\alpha = 0.05$

3. Estadístico de Prueba bajo  $H_0$  cierta:

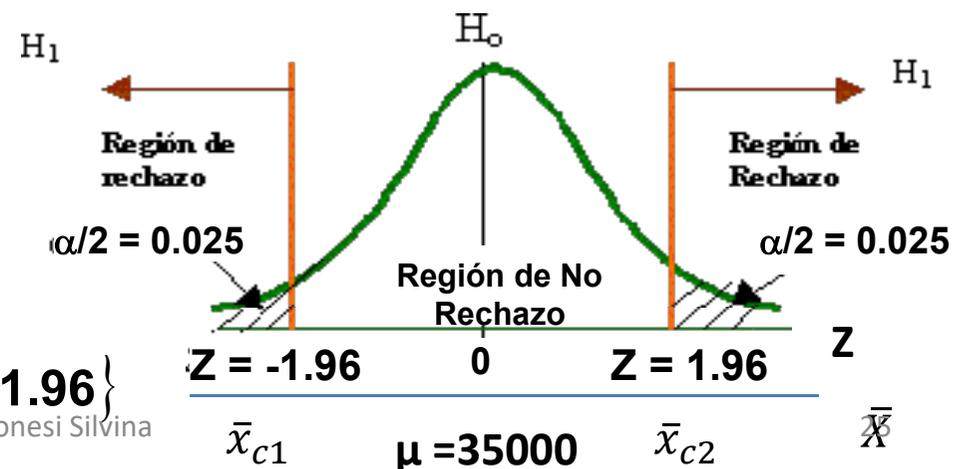
$$Z = \frac{\bar{X} - 35000}{2998.2 / \sqrt{10}} \sim N(0, 1)$$

Como  $\alpha = 0.05$  entonces  $\alpha / 2 = 0.025$ , por lo tanto

$$z_{\alpha/2} = -1.96 \quad \text{y} \quad z_{1-\alpha/2} = 1.96$$

4. Región Crítica:

$$RC = \left\{ z / z < z_{\alpha/2} = -1.96 \quad \text{o} \quad z > z_{1-\alpha/2} = 1.96 \right\}$$



**5. El valor del estadístico es:**

Datos:  $n=10$  y  $\bar{x} = 33100$ .

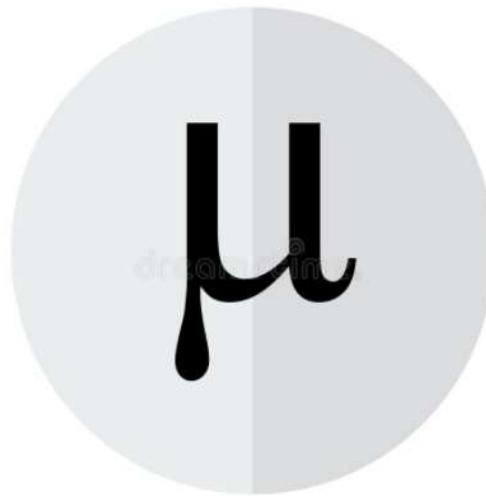
$$z_{\text{obs}} = \frac{33100 - 35000}{2998.2/\sqrt{10}} = \frac{-1900}{948.11} = -2.004$$

**6. Comparación: Como  $z_{\text{obs}} = -2.004 < -1.96$  entonces **Rechazo  $H_0$** .**

**7. Conclusión:**

Con una probabilidad de error de **0.05**, existen evidencias suficientes para afirmar que el salario promedio mensual de un ingeniero en computación que trabaja en esa empresa difiere de los 35000 pesos.

# Prueba de Hipotesis para la media de una población normal con varianza desconocida



# Prueba de hipótesis para la media de una población normal con varianza desconocida

Sea  $X \sim N(\mu, \sigma^2)$ , con  $\sigma^2$  desconocida. El estadístico de prueba que se basa en la media muestral,  $\bar{X}$ , es

$$t_{n-1} = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \quad \text{donde } S \text{ es la v.a. desvío muestral}$$

**Planteo:** Se quiere probar  $H_0: \mu = \mu_0$  frente a alguna de las hipótesis alternativas planteadas

1.  $H_1: \mu \neq \mu_0$
2.  $H_1: \mu > \mu_0$
3.  $H_1: \mu < \mu_0$ .

**Estadístico de prueba:**  $t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$  Si  $H_0$  es verdadera

**Las respectivas regiones críticas son:**

$$1. \{ t_{n-1} / t_{n-1} < -t_{n-1, \alpha/2} \text{ o } t_{n-1} > t_{n-1, \alpha/2} \}$$

$$2. \{ t_{n-1} / t_{n-1} > t_{n-1, \alpha} \}$$

$$3. \{ t_{n-1} / t_{n-1} < -t_{n-1, \alpha} \}$$

## Ejemplo

La secretaría de control y fiscalización ambiental desea investigar si una empresa textil cumple con la normativa vigente y las aguas residuales que emite son tratadas convenientemente antes de ser vertidas al exterior. La normativa exige que el contenido de cierta sustancia tóxica en las aguas residuales, no supera en promedio las 55 ppm. Para verificarlo, se tomaron 25 medidas de aguas residuales vertidas por la empresa al medio en diferentes días, obteniéndose contenido promedio de esa sustancia toxica de 60 ppm, con una desviación estándar de 10 ppm.

Si se supone que el contenido de cierta sustancia tóxica es una v.a. con distribución normal, a un nivel de significación del 5% ¿Proporciona esta información muestral evidencia suficiente de que empresa textil no cumple con la normativa vigente ?



Sea la v.a.  $X$  = "contenido de cierta sustancia tóxica que vierte al medio la empresa por día", (ppm)  $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  es desconocido.

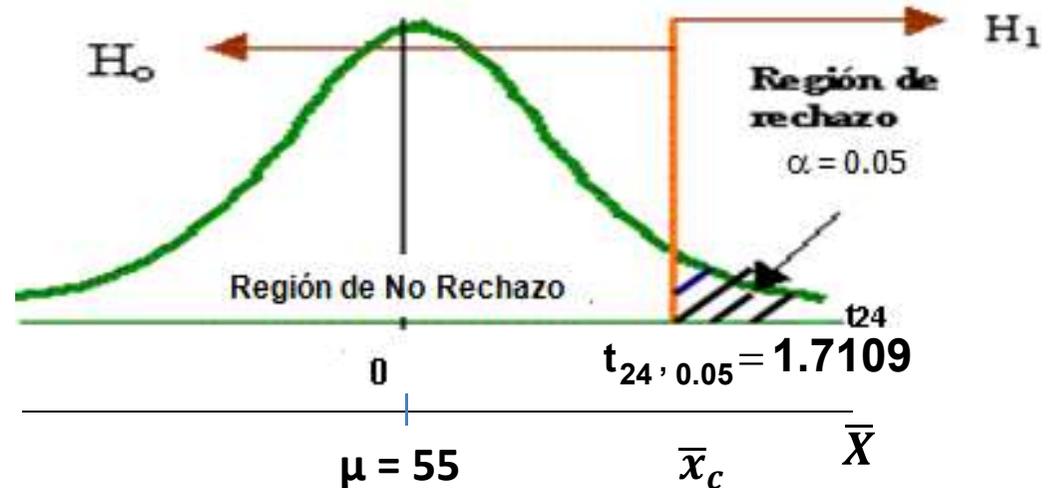
En consecuencia,  $\bar{X} \sim N(\mu, \sigma^2/n)$ ,

1. Planteo

$H_0: \mu \leq 55$  (es equivalente a  $\mu = 55$ )

$H_1: \mu > 55$

2.  $\alpha = 0.05$



3. Estadístico de Prueba:

$$t = \frac{\bar{X} - 55}{S/\sqrt{25}} \sim t_{24} \quad \text{bajo } H_0 \text{ cierta}$$

Como  $\alpha = 0.05$ , por lo tanto  $t_{24, 0.05} = 1.7109$

4. Región Crítica:  $RC = \{ t_{24} / t_{24} > t_{24, 0.05} = 1.7109 \}$

5. El valor del estadístico es:

Datos:  $n = 25$ ,

$\bar{x} = 60\text{ppm}$      $s = 10\text{ppm}$

$$t_{\text{obs}} = \frac{60 - 55}{10/\sqrt{25}} = \frac{5}{2} = 2.5$$

6. Comparación: como  $t_{\text{obs}} = 2.5 > 1.7109$  entonces **Rechazo  $H_0$** .

## 7. Conclusión:

Con una probabilidad de error de **0.05**, existen evidencias suficientes para afirmar que el contenido promedio de cierta sustancia tóxica en las aguas vertidas por la empresa superan los 55 ppm. La empresa no está cumpliendo con la normativa vigente.

## Ejemplo

El Centro de Atención Telefónica del banco Business debe tardar a lo sumo 4 minutos en resolver cualquier trámite de un usuario, para cumplir con lo establecido por el Departamento de Atención al Cliente. Una muestra de 25 llamados arrojó una duración promedio de 4.5 minutos, con una desviación estándar de 1.4 minutos.

Si se supone que la duración de una llamada telefónica para resolver cualquier trámite es una v.a. con distribución normal, a un nivel de significación del 1% ¿el Centro de atención al Cliente cumple con lo fijado por el Departamento del Atención al Cliente?

Sea la v.a.  $X$  = “duración de una llamada telefónica para resolver cualquier trámite”,  $X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  es desconocido.

### 1. Planteo

$H_0: \mu \leq 4$  (es equivalente a  $\mu = 4$ )

$H_1: \mu > 4$

### 2. $\alpha = 0.01$

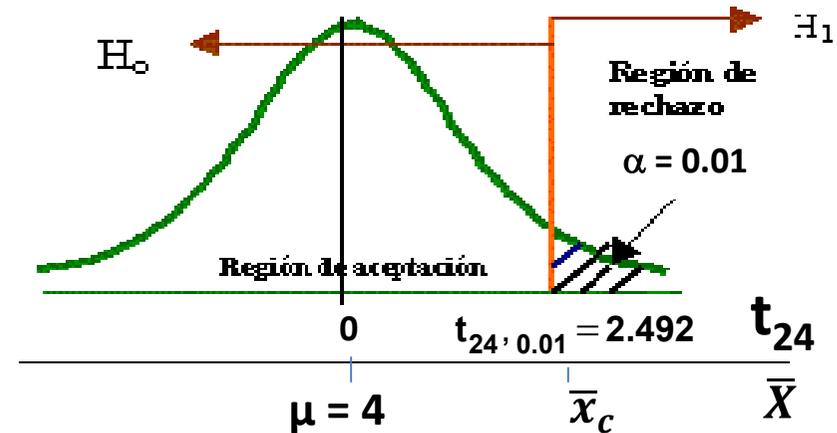


$X \sim N(\mu, \sigma^2)$ ,  $\sigma^2$  es desconocido en consecuencia,  $\bar{X} \sim N(\mu, \sigma^2/n)$ ,

3. Estadístico de Prueba:  $t = \frac{\bar{X} - 4}{S/\sqrt{25}} \sim t_{n-1=24}$  bajo  $H_0$  cierta

Como  $\alpha = 0.01$ , por lo tanto

$$t_{24, 0.01} = 2.492$$



4. Región Crítica:

$$RC = \{ t_{24} / t_{24} > t_{24, 0.01} = 2.492 \}$$

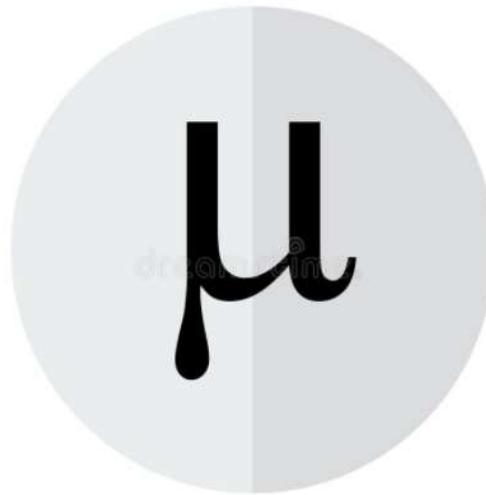
5. El valor del estadístico es:

Datos:  $n = 25$   $\bar{x} = 4.5$  minutos y  $s = 1.4$  minutos.  $t = \frac{4,5 - 4}{1.4/\sqrt{25}} = \frac{0.5}{0.28} = 1.79$

6. Comparación: como  $t = 1.79 < 2.492$  entonces **No rechazo  $H_0$ .**

7. Conclusión: Con un nivel de significación de **0.01**, no existen evidencias suficientes para afirmar que la duración promedio de una llamada telefónica para resolver cualquier trámite sea superior a los 4 minutos. El Centro de atención al Cliente cumple con lo fijado por el Departamento del Atención al Cliente.

# Prueba de Hipotesis para la media de una población cualquiera con varianza conocida o desconocida



# Prueba de hipótesis para la media de una población

---

## Observación importante!!!

Cuando se prueban hipótesis sobre la media  $\mu$ :

- Si la v.a.  $X$  **no está normalmente** distribuida pero  $\sigma^2$  **conocida** y el tamaño de la muestra es  $n \geq 30$ , el **teorema central del límite** garantiza de todos modos la distribución normal estándar del **estadístico de prueba** es

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

- Si la v.a.  $X$  **no está normalmente** distribuida pero  $\sigma^2$  **desconocida** y el tamaño de la muestra es  $n \geq 30$ , la varianza muestral  $S^2$  se acercará al verdadero valor de  $\sigma^2$  y entonces sigue siendo aplicable el **teorema central de límite** proporcionando el siguiente **estadístico de prueba**:

$$Z \cong \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0, 1)$$

## Ejemplo:

El gerente de una planta industrial química afirma que la producción promedio diaria es de 880 toneladas. Para comprobar dicha afirmación se registró la producción de 50 días elegidos al azar y se obtuvo una media de 871 toneladas y una desviación de 21 toneladas. A un nivel de 0.01, ¿esta de acuerdo con lo que establece el gerente de la empresa?

Sea la v.a.  $X$  = “producción diaria de la planta”,  $X \sim ?$ ,  $\sigma^2$  es desconocido.

### 1. Planteo:

$$H_0: \mu = 880$$

$$H_1: \mu \neq 880$$

2.  $\alpha = 0.01$



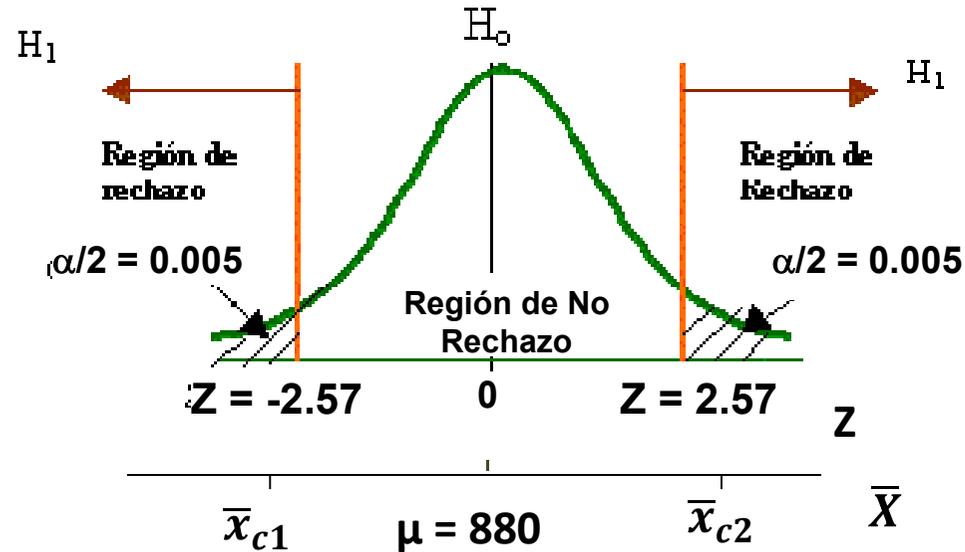
Como  $n = 50 \geq 30$ ,  $S^2$  se acercará al verdadero valor de  $\sigma^2$  y entonces por el teorema central de límite.

### 3. Estadístico de Prueba bajo $H_0$ cierta:

$$Z \cong \frac{\bar{X} - 880}{S/\sqrt{50}} \sim N(0, 1)$$

Como  $\alpha = 0.01$  entonces  $\alpha/2 = 0.005$  y  $1-\alpha/2 = 0.995$ , por lo tanto

$$z_{\alpha/2} = -2.57 \text{ y } z_{1-\alpha/2} = 2.57$$



#### 4. Región Crítica:

$$RC = \{ z / z < z_{\alpha/2} = -2.57 \text{ o } z > z_{1-\alpha/2} = 2.57 \}$$

#### 5. El valor del estadístico es:

Datos:  $n = 50$ ,  $\bar{x} = 871$  toneladas y un  $s = 21$  toneladas

$$z = \frac{871 - 880}{21/\sqrt{50}} = \frac{-9}{2.97} = -3.03$$

6. Comparación: Como  $z = -3.03 < -2.57$  entonces Rechazo  $H_0$ .

#### 7. Conclusión:

Con una probabilidad de error de **0.01**, existen evidencias suficientes para afirmar que la producción promedio diaria difiere de 880 toneladas.

Se realizaron **cuatro Pruebas de Hipótesis**. A continuación se muestran las decisiones tomadas según el valor de  $\alpha$ , completar en cuadro convenientemente:

**OJO!!! Tomo con  $\alpha$  de referencia al 0.05.**

Prueba	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	Decisión	¿Probabilidad de error o nivel de Significación?	Valor de $\alpha$
<b>1</b>	<b>Rechazo <math>H_0</math></b>	<b>Rechazo <math>H_0</math></b>	<b>Rechazo <math>H_0</math></b>	<b>Rechazo <math>H_0</math></b>	<b>Probabilidad de error</b>	<b><math>\alpha = 0.01</math></b>

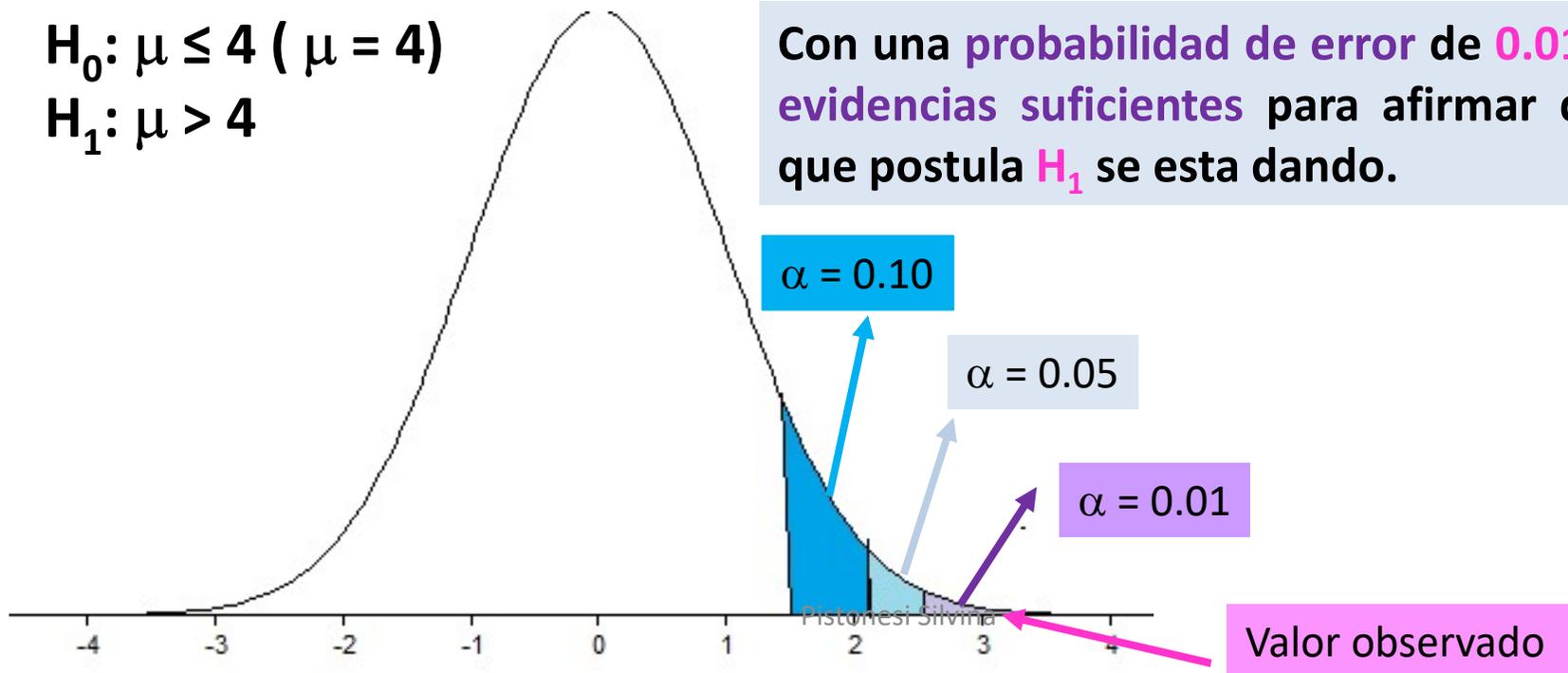
1..

$$P(\text{Error de Tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ verdadera}) = \alpha$$

$$H_0: \mu \leq 4 \quad (\mu = 4)$$

$$H_1: \mu > 4$$

Con una probabilidad de error de **0.01**, existen evidencias suficientes para afirmar que la lo que postula  $H_1$  se esta dando.



Dados los siguientes resultados de cuatro Pruebas de Hipótesis, completar el cuadro:

Prueba	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	Decisión	¿Probabilidad de error o nivel de Significación?	Valor de $\alpha$
2	Rechazo $H_0$	Rechazo $H_0$	No rechazo $H_0$	Rechazo $H_0$	Probabilidad de error	$\alpha = 0.05$

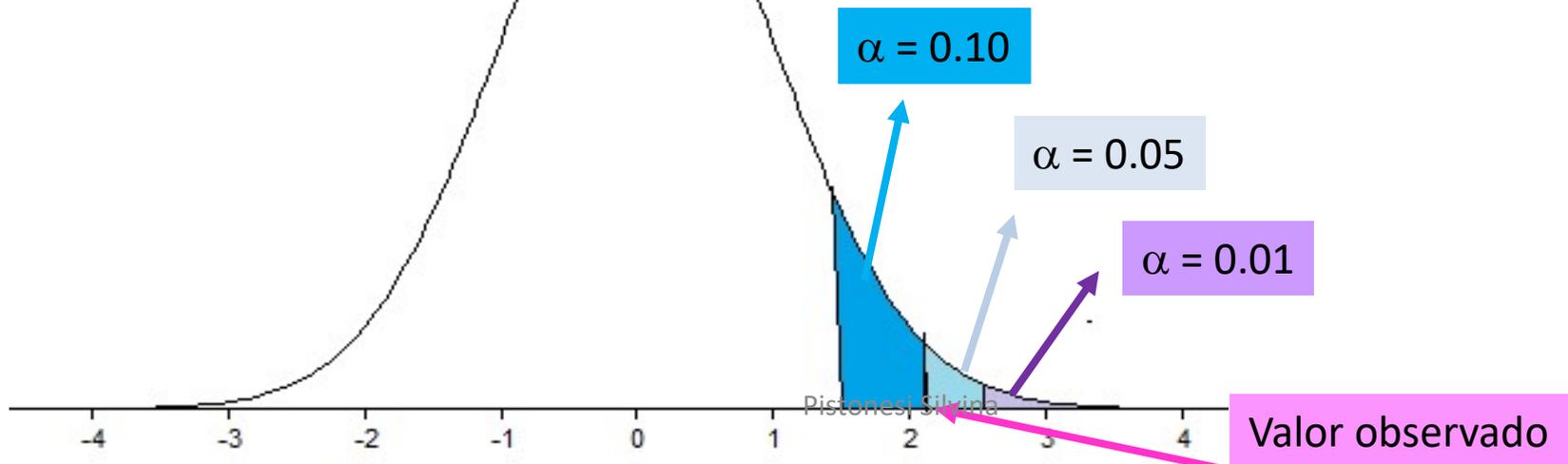
2.

$$P(\text{Error de Tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ verdadera}) = \alpha$$

$$H_0: \mu \leq 4 \quad (\mu = 4)$$

$$H_1: \mu > 4$$

Con una probabilidad de error de **0.05**, existen evidencias suficientes para afirmar que la lo que postula  $H_1$  se esta dando.



Dados los siguientes resultados de cuatro Pruebas de Hipótesis, completar el cuadro:

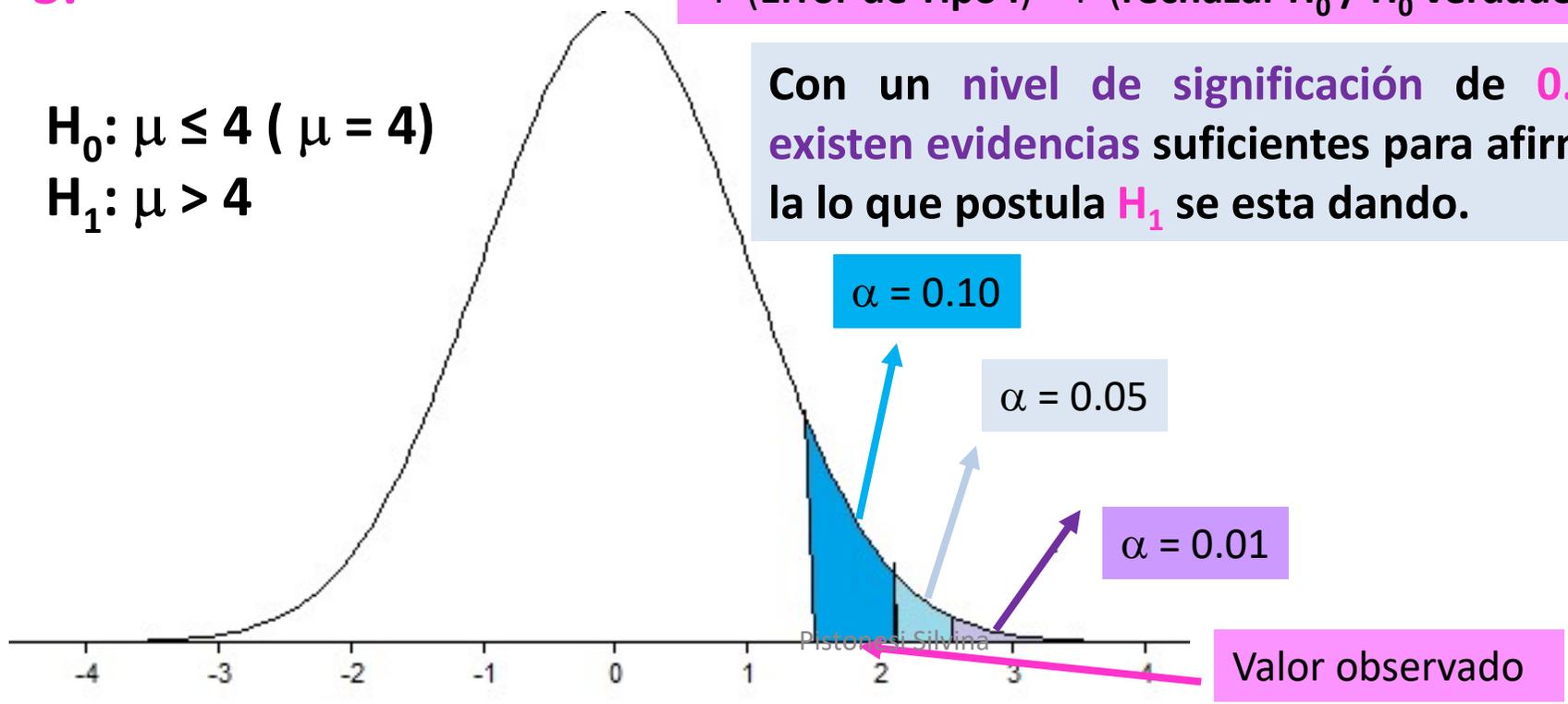
Prueba	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	Decisión	¿Probabilidad de error o nivel de Significación?	Valor de $\alpha$
3	Rechazo $H_0$	No Rechazo $H_0$	No Rechazo $H_0$	No Rechazo $H_0$	Nivel de significación	$\alpha = 0.05$

3.

$H_0: \mu \leq 4$  ( $\mu = 4$ )  
 $H_1: \mu > 4$

$P(\text{Error de Tipo I}) = P(\text{rechazar } H_0 / H_0 \text{ verdadera}) = \alpha$

Con un nivel de significación de 0.05, No existen evidencias suficientes para afirmar que la lo que postula  $H_1$  se esta dando.



Dados los siguientes resultados de cuatro Pruebas de Hipótesis, completar el cuadro:

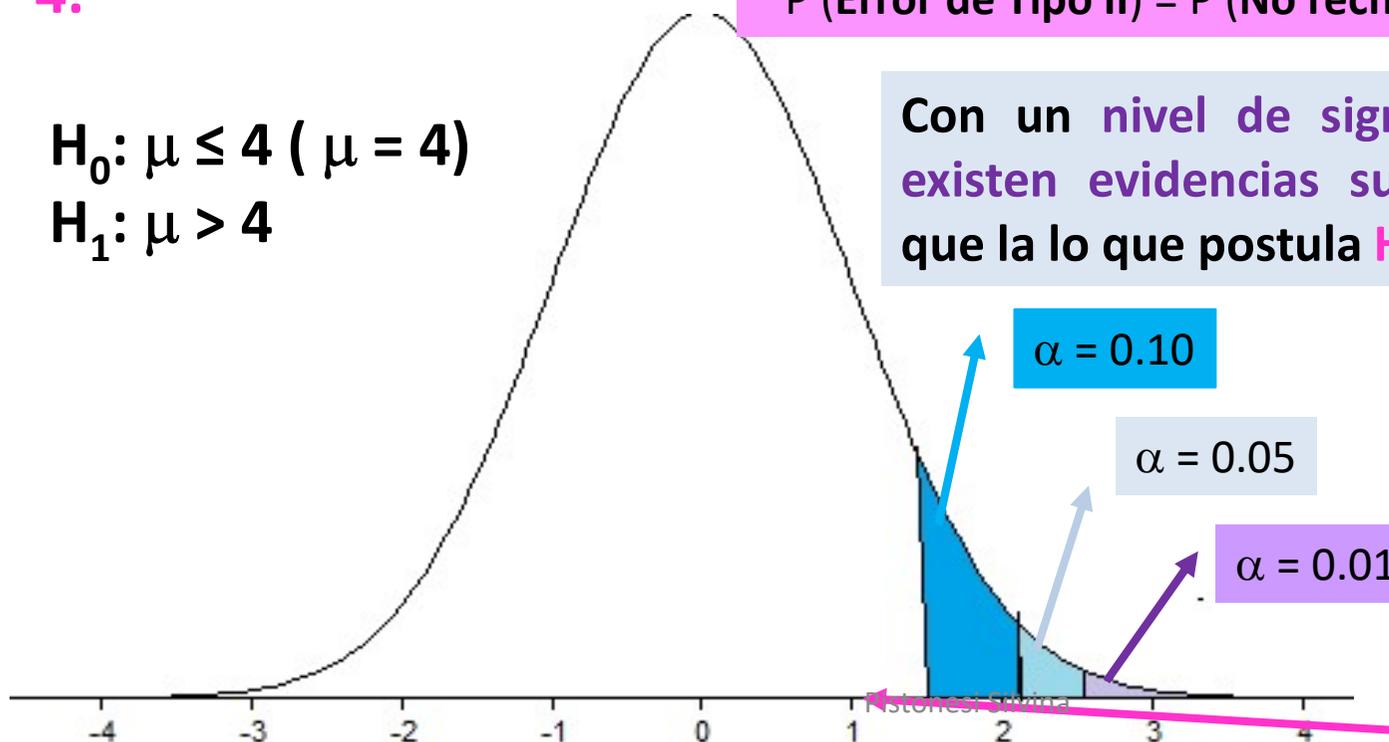
Prueba	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	Decisión	¿Probabilidad de error o nivel de Significación?	Valor de $\alpha$
4	No Rechazo $H_0$	No Rechazo $H_0$	No Rechazo $H_0$	No Rechazo $H_0$	Nivel de significación	$\alpha = 0.10$

4.

$H_0: \mu \leq 4$  ( $\mu = 4$ )  
 $H_1: \mu > 4$

$P(\text{Error de Tipo II}) = P(\text{No rechazar } H_0 / H_0 \text{ falsa}) = \beta$

Con un nivel de significación de **0.10**, No existen evidencias suficientes para afirmar que la lo que postula  $H_1$  se esta dando.



Valor observado <sup>41</sup>

## Relación entre $\alpha$ , $\beta$ y $\gamma$

$\alpha \uparrow$	$\beta \downarrow$	$\gamma \uparrow$
$\alpha \downarrow$	$\beta \uparrow$	$\gamma \downarrow$
$n \uparrow$ , fijado $\alpha$	$\beta \downarrow$	$\gamma \uparrow$

# Estadísticos de Prueba de la Prueba de Hipótesis para $E(X) = \mu$

