

Unidad IV

Prueba de Hipótesis Parte I

¿¿¿Con cuál me quedo???

Hipótesis Nula:
No hay diferencias

Hipótesis Alternativa:
Hay diferencias significativas



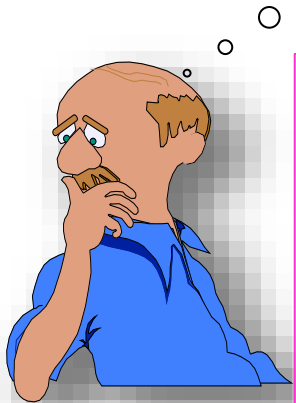
Pistonesi Silvina

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 **\$1537500**. 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, $\bar{x} = \$1518000$, 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 **\$739000** 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 .

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
 $H_0: \mu = \$1537500$ (ó $\mu \geq \$1537500$).

$H_1: \mu < \$1537500$.



Prueba de hipótesis

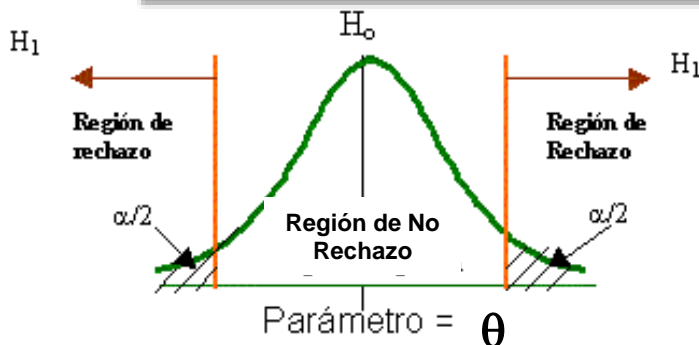
Tipos de pruebas

- **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

$$H_0: \theta = \theta_0$$

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

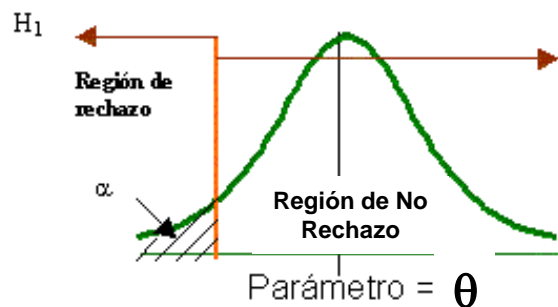


Prueba Bilateral

Pruebas Unilaterales

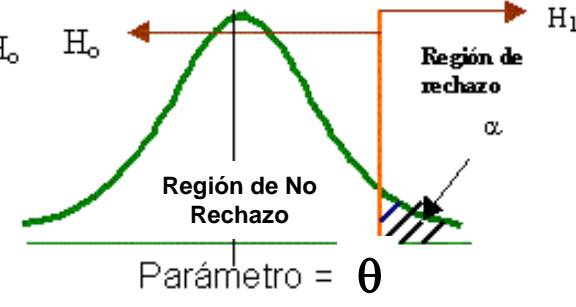
$$H_0: \theta = \theta_0 \text{ (ó } \theta \leq \theta_0)$$

$$H_1: \theta > \theta_0$$



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

$$H_1: \theta < \theta_0$$



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**.

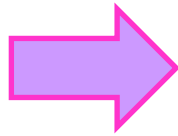
Situaciones posibles que pueden surgir en un test de hipótesis

Realidad Decisión Posible		
	H_0 es Verdadera	H_0 es Falsa
No rechazar H_0	Decisión correcta	Decisión incorrecta Error de Tipo II
Rechazar H_0	Decisión incorrecta Error de Tipo I	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 Error de Tipo II
Rechazar H_0	Decisión incorrecta Error de Tipo I	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**.

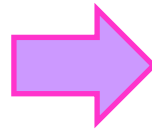
Prueba de hipótesis: errores

Error de Tipo I = Rechazar H_0 cuando H_0 es verdadera

Error de Tipo II = No rechazar H_0 cuando H_0 es falsa

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

Planteo



$H_0: \mu = \$1537500$ (ó $\mu \geq \$1537500$).

$H_1: \mu < \$1537500$.

- Se comete un **error de Tipo I**: si se **rechaza la hipótesis nula** de $\mu = \$1537500$ de acuerdo a los resultados de la prueba, pero realmente el valor medio es de $\$1537500$, ie, **suponer que el ingreso promedio diario es inferior al que afirma el actual dueño**, cuando en realidad es de $\$1537500$.
- Se comete un **error de Tipo II**: si la hipótesis $H_0: \mu = \$1537500$ **no es rechazada** según los datos de la prueba, pero realmente el valor medio es menor a $\$1537500$, ie, **suponer que el ingreso promedio diario es de $\$1537500$** , 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 sobrevida 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 sobrevida 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 sobrevida en lo 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**.

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 α ,
y
la **probabilidad de cometer un Error de Tipo II** por β .

Simbólicamente se puede expresar:

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

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

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

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 α : **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 α sea pequeña.

El complemento de β 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: γ . 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 α , β y γ

$\alpha \uparrow$

$\beta \downarrow$

$\gamma \uparrow$

$\alpha \downarrow$

$\beta \uparrow$

$\gamma \downarrow$

$n \uparrow$, fijado α

$\beta \downarrow$

$\gamma \uparrow$

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

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 **\$1523000**. Suponiendo que el ingreso diario del bar es v.a. **distribuida normalmente** con un **desvío estándar de \$70000** 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 μ** . El mejor **estimador puntual** de μ 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.

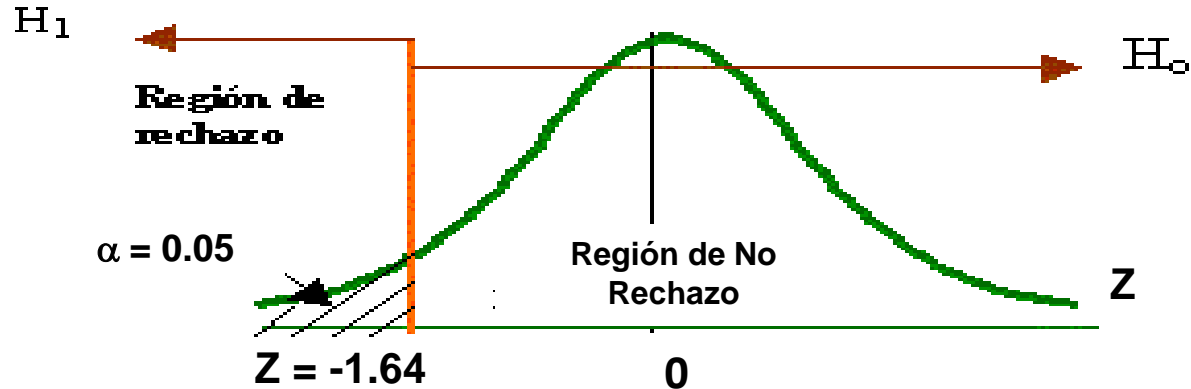
El **valor crítico** es el punto que separa la **región crítica** del conjunto de valores para los cuales **no se rechaza H_0** .

Para establecer **la región crítica o zona de rechazo de H_0** , se elige un valor para **α** . El valor de **α** 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** ($H_1: \mu < \$1537500$), 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 » $X \sim N(\mu, \sigma^2 = \$^270000^2)$, con σ^2 conocida.

Paso 1: Planteo de lo que se desea probar :

$H_0: \mu = \$1537500$ (ó $\mu \geq \$1537500$) .
 $H_1: \mu < \$1537500$

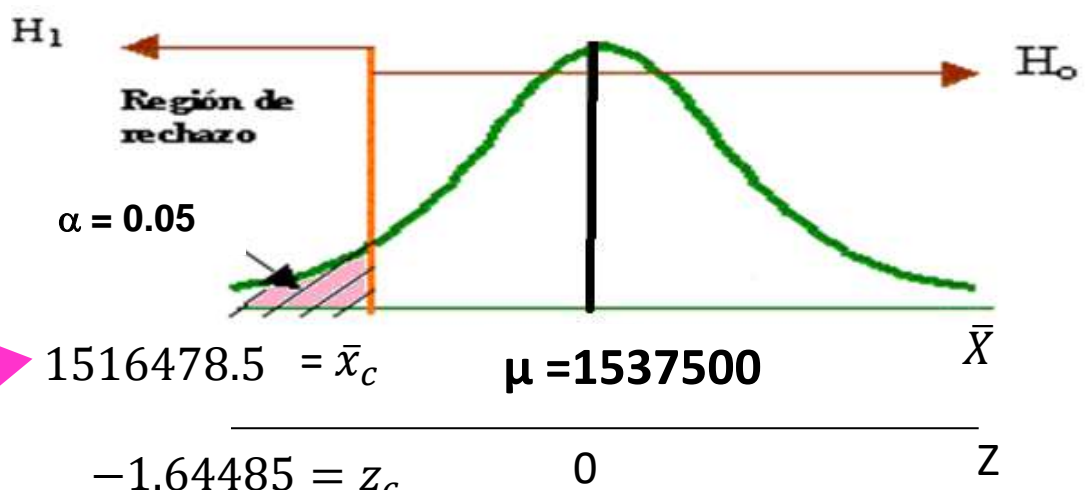
Paso 2: $\alpha = 0.05$ (nivel de significación fijado) $X \sim N(\mu, \sigma^2 = \$^270000^2)$, $\Rightarrow \bar{X} \sim N(\mu, \sigma^2/n)$

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

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{\bar{X} - 1537500}{70000/\sqrt{30}} \sim N(0, 1), \text{ Bajo } H_0 \text{ cierta.}$$

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

$$-z_c = -1.64485 = \frac{\bar{X} - 1537500}{70000/\sqrt{30}}$$
$$1516478.5 = \bar{x}_c$$



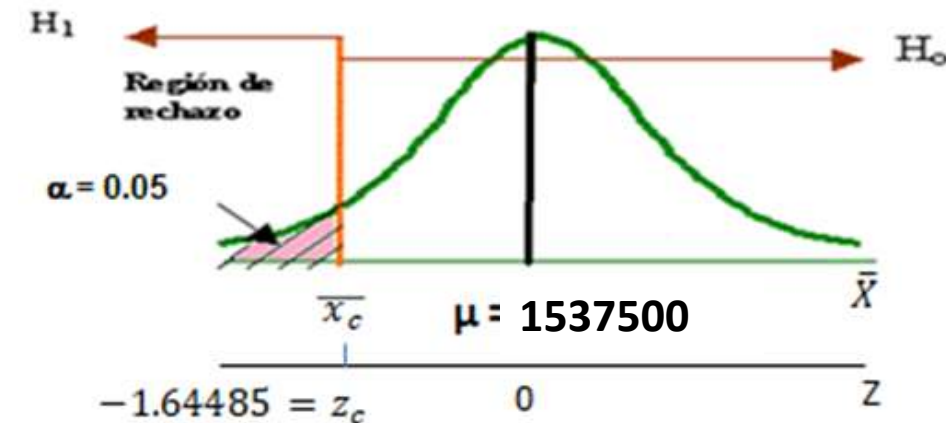
La prueba de hipótesis para la media poblacional con varianza conocida

Paso 5: La región crítica o de rechazo analíticamente,

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

$$RC = \{ \bar{X} : \bar{X} < \bar{x}_c = 1516478.5 \}$$

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



$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{1523000 - 1537500}{70000 / \sqrt{30}} = -1.13$$

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

Como $z_{\text{obs}} = -1.13 > -1.64485 = z_c$ entonces No Rechazo H_0 (la decisión tomada con la información muestral)

Paso 8 : Conclusión:

No Tengo evidencias suficientes para afirmar que el ingreso promedio diario del bar sea inferior a \$1537500, tengo pruebas que el dueño le este mintiendo al comprador, con nivel de significación de 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 α .
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\os 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.

Parámetro

Una **v.a.** o **población X** está caracterizada por su función de densidad ó de distribución de probabilidad,

Un **parámetro** es aquella constante que informa de manera sintética de una propiedad relevante o característica de la población o de la variable aleatoria.

Características:

- Es un valor único.
- Generalmente desconocido.
- Para hallar su valor se necesita de todos los elementos de la población.

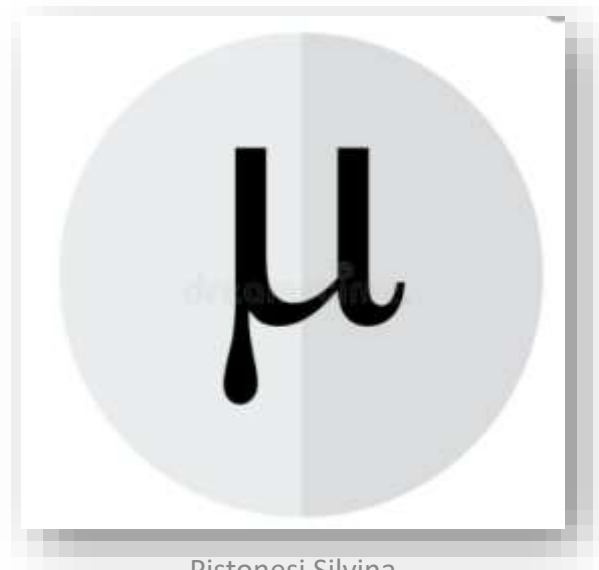
Los parámetros clásicos son:

μ = $E(X)$ el valor medio

σ^2 = $V(X)$ la varianza

σ = $V(X)^{1/2}$ el desvío estándar

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 σ^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.

Un proveedor de servicios de internet afirma que la velocidad media de transferencia de datos de su nueva red de fibra óptica es de 120 Mbps. Un grupo de usuarios sospecha que la velocidad promedio real es diferente a la que se anuncia. Para comprobarlo, se tomaron al azar 40 mediciones de velocidad, obteniendo una velocidad media transferencia de datos de 117.8 Mbps.

Se sabe la velocidad de transferencia de datos de la nueva red de fibra óptica es una v.a. normalmente distribuida con un desvío estándar de 5 Mbps.

A un nivel de significación del 0.05, ¿se puede concluir que la velocidad media real transferencia de datos de su nueva red difiere de 120 Mbps?



X = “la velocidad media de transferencia de datos de su nueva red de fibra óptica ” (en Mbps)

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

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

1. Planteo:
 $H_0:$
 $H_1:$

2. $\alpha =$

3. Estadístico de Prueba bajo H_0 cierta:

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

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

4. Región Crítica:

$$RC = \{ \quad \}$$

5. El valor del estadístico es:

Datos: $n =$ y $\bar{x} =$

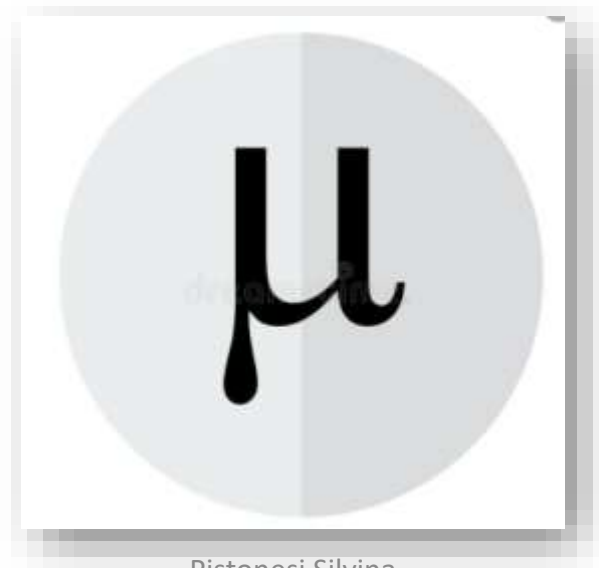
$$= \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} =$$

6. Comparación: Como entonces H_0 .

7. Conclusión:

Con de **0.05**, existen evidencias suficientes para afirmar que la velocidad media de transferencia de datos de la nueva red de fibra óptica difiere .

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 $\mathbf{X} \sim \mathbf{N}(\mu, \sigma^2)$, con σ^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 sospecha que la empresa textil no cumple con la normativa vigente y las aguas residuales que emite no 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) $X \sim N(\mu, \sigma^2)$, σ^2 es desconocido.

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

1. Planteo

$$H_0: \mu$$

$$H_1: \mu$$

2. $\alpha =$

3. Estadístico de Prueba:
$$t = \frac{\bar{X} - 55}{S/\sqrt{25}}$$

Como $\alpha = 0.05$, por lo tanto

4. Región Crítica: $RC = \{ \quad \quad \quad \}$

5. El valor del estadístico es:

Datos: $n = 25$,

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

$$t_{\text{obs}} = \frac{\bar{x} - 55}{s/\sqrt{n}} =$$

6. Comparación: como entonces H_0 .

7. Conclusión:

Con de **0.05**, existen evidencias suficientes
para afirmar que el contenido promedio de cierta sustancia tóxica en las
aguas vertidas por la empresa ppm.

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”, $X \sim N(\mu, \sigma^2)$, σ^2 es desconocido.

1. Planteo

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

$H_1: \mu > 4$

2. $\alpha = 0.01$

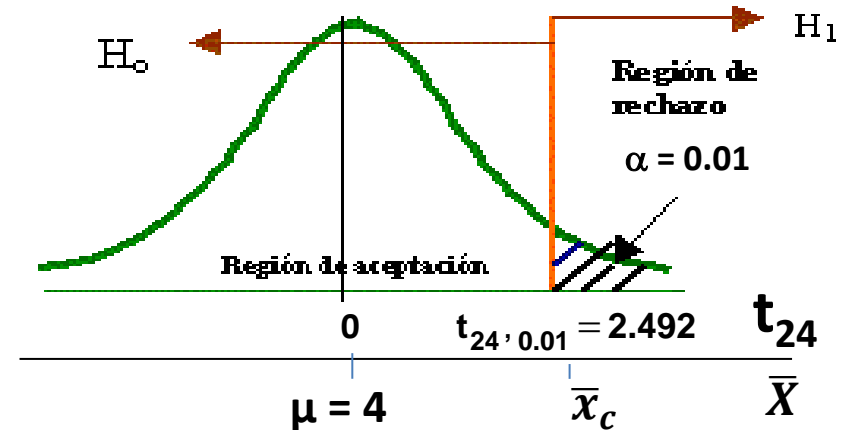


$X \sim N(\mu, \sigma^2)$, σ^2 es desconocido en consecuencia, $\bar{X} \sim N(\mu, \sigma^2/n)$,

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:

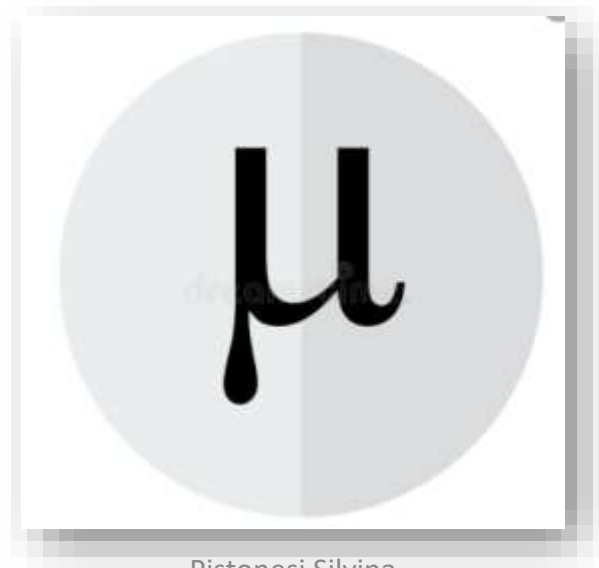
Datos: $n = 25$ $\bar{x} = 4.5$ minutos y $s = 1.4$ minutos.

$$t = \frac{4.5 - 4}{1.4/\sqrt{25}} = \frac{0.5}{0.28} = 1.79$$

6. Comparación: como $t = 1.79 < 2.492$ entonces **No rechaza 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 μ :

- Si la v.a. X **no está normalmente** distribuida pero σ^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 σ^2 **desconocida** y el tamaño de la muestra es **$n \geq 30$** , la varianza muestral S^2 se acercará al verdadero valor de σ^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:

Un ingeniero en ciencias de la computación ha desarrollado un **nuevo algoritmo de ordenamiento** que, según él, mejora significativamente el rendimiento en tiempo de ejecución respecto al algoritmo actualmente en uso. El algoritmo actual tiene, un **tiempo de ejecución promedio de 2.5 segundos** sobre conjuntos de datos de tamaño estándar. Para validar la afirmación del ingeniero, se ejecuta el nuevo algoritmo sobre una muestra aleatoria de 100 conjuntos de datos de tamaño estándar y se obtiene un tiempo de ejecución promedio de 2.3 segundos y un desvío de 0.6 segundos. A un nivel de 0.01, ¿esta de acuerdo con lo que establece el ingeniero?

Sea la v.a. **X** = “tiempo de ejecución del algoritmo de ordenamiento sobre un conjunto de datos”, **X ~ ?**, **σ^2 es desconocido**.

1. Planteo:

$$H_0: \mu = 2.5$$

$$H_1: \mu < 2.5$$

2. $\alpha = 0.01$

3. Estadístico de Prueba bajo H_0 cierta:

$$Z \cong \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

Como $\alpha =$ entonces $\alpha/2 =$ y $1-\alpha/2 =$



$$z_{\alpha} =$$

4. Región Crítica:

$$\mathbf{RC} = \{ \quad \quad \quad \}$$

5. El valor del estadístico es:

$$= \frac{-}{\sqrt{\quad}} =$$

Datos: $n =$, $\bar{x} =$ y un $s =$

6. Comparación:

Como

entonces

H_0 .

7. Conclusión:

Con
para afirmar que

de **0.01**,

existen evidencias suficientes

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

OJO!!! Tomo con α 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 α
1	Rechazo H_0	Rechazo H_0	Rechazo H_0	Rechazo H_0	Probabilidad de error	$\alpha = 0.01$

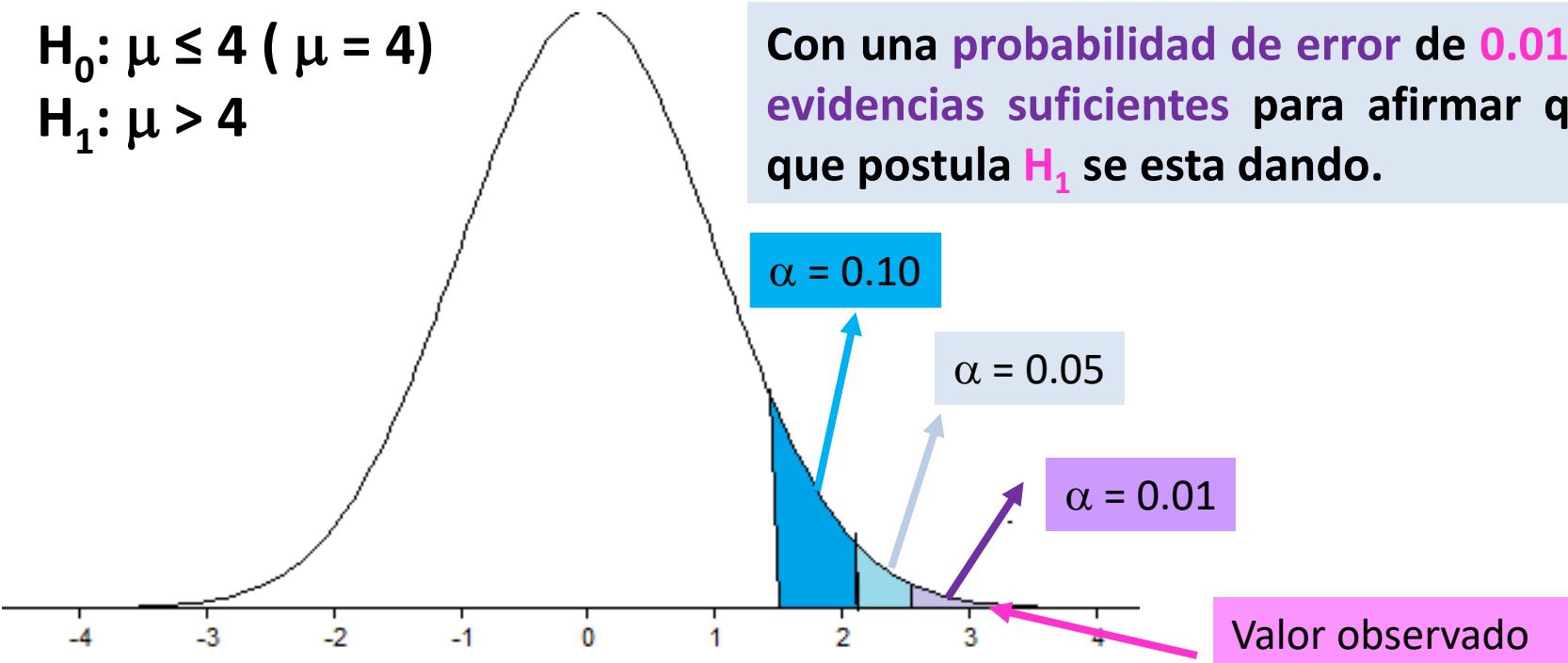
1..

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

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

$$H_1: \mu > 4$$

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 α
2	Rechazo H_0	Rechazo H_0	No rechazo H_0	Rechazo H_0	Probabilidad de error	$\alpha = 0.05$

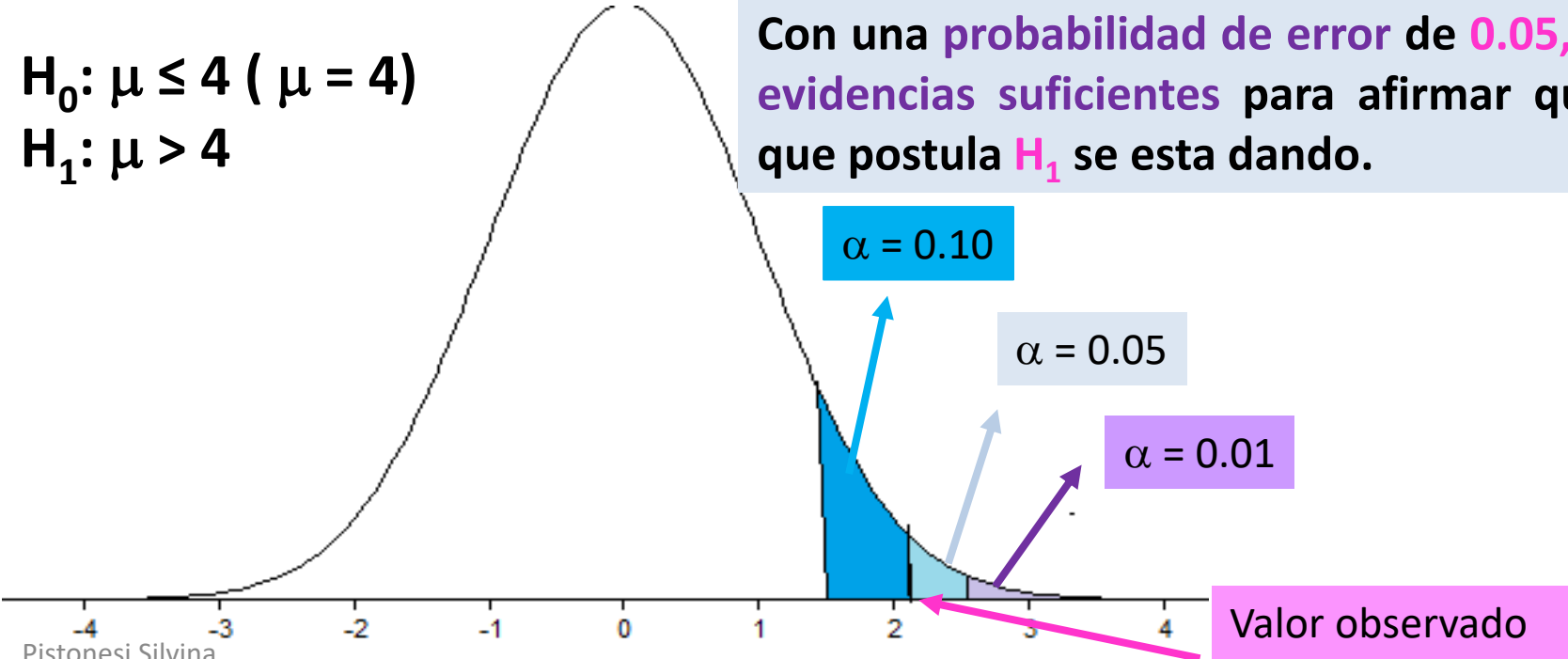
2.

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

$$H_0: \mu \leq 4 \text{ (} \mu = 4 \text{)}$$

$$H_1: \mu > 4$$

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

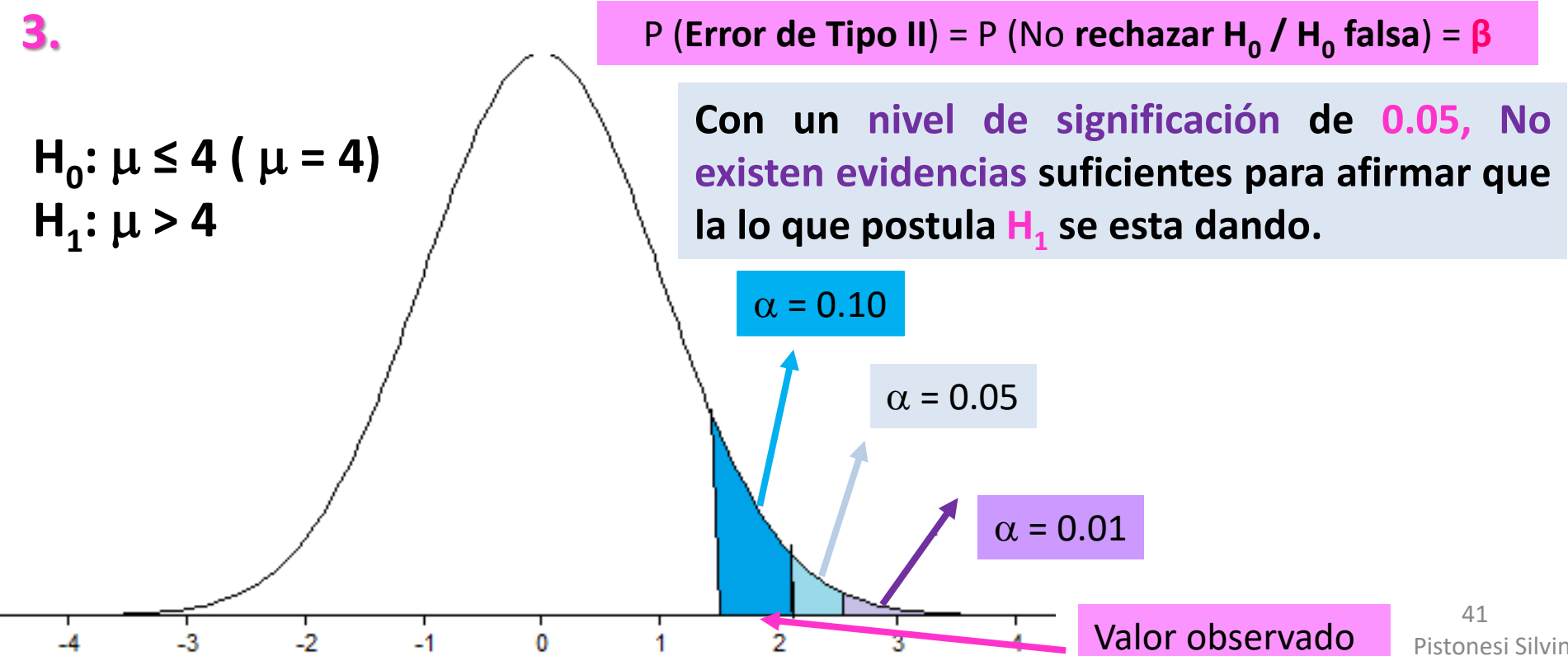


Dados los siguientes resultados de cuatro Pruebas de Hipótesis, completar el cuadro:

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

3.

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

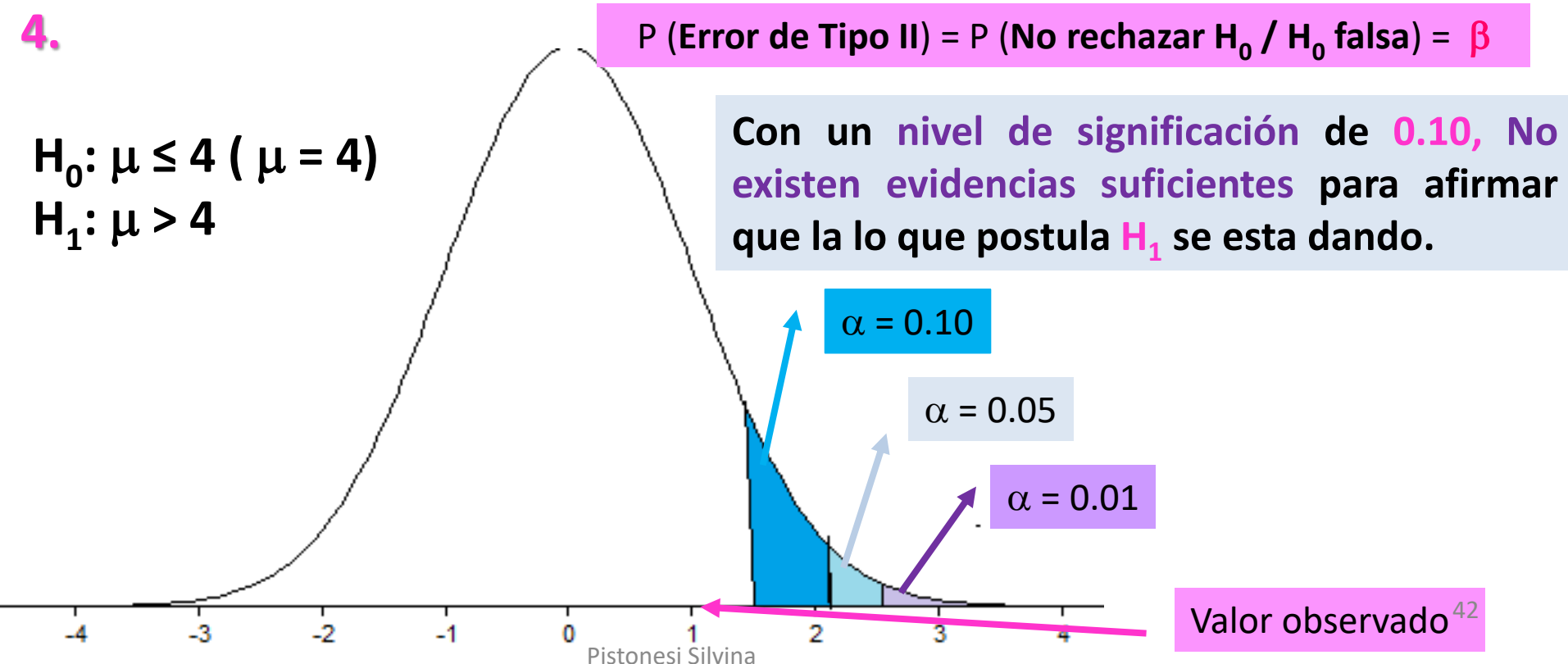


Dados los siguientes resultados de cuatro Pruebas de Hipótesis, completar el cuadro:

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

4.

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



Relación entre α , β y γ

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

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

σ^2 es conocido $\bar{X} \sim N(\mu, \sigma^2/n)$

Si $X \sim N(\mu, \sigma^2)$

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

σ^2 es desconocido $\bar{X} \sim N(\mu, \sigma^2/n \text{ ?})$

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

Si $X \sim \text{????}$

σ^2 es conocido

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

$n \geq 30$, por el
Teo. Central de L.

σ^2 es desconocido

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