Ingeniería Industrial.
Asignatura: Estadística I **Año: 3er.** **Modalidad: CE.**
Tema 4. Pruebas de Hipótesis
Sumario.
4.1. Conceptos básicos. Hipótesis nula e hipótesis alternativa. Tipos de errores: tipo I y tipo II. Probabilidad de cometer ambos tipos de errores. Pvalor ó valor-p.
4.2 Procedimiento lógico para realizar una P.H.
4.3. Pruebas de hipótesis para medias de una muestra conociendo y desconociendo la varianza poblacional (σ^2).
4.4. Pruebas de hipótesis para comparación de medias de dos muestras conociendo y desconociendo la varianza poblacional (σ^2).
4.5. Pruebas de hipótesis para las proporciones de una y dos muestras de poblaciones.
Objetivos:
Caracterizar los diferentes tipos de error que se pueden cometer al realizar una Prueba de Hipótesis.
Aplicar Pruebas de Hipótesis para la solución de problemas afines a su perfil profesional
Método:
- Expositivo

Medios:
- Tiza y pizarra, Libro de texto, Tabla Estadísticas dispositivos digitales

Forma de enseñanza: Conferencia
Bibliografía:
- Estadística Dra Caridad Bustillo.
- Materiales Didácticos digitales entregados por el profesor.

Es muy frecuente en la vida tener que tomar decisiones sobre los parámetros de una población conociendo solo los valores de muestra.
Ej: Si el empleo de un proceso productivo reduce los costos promedios más que otro; si un método de gestión es más eficiente que otro, si un sistema de pago mejora los salarios medios de los trabajadores.

4.1 Conceptos básicos. Hipótesis nula e hipótesis alternativa. Tipos de errores: tipo I y tipo II. Probabilidad de cometer ambos tipos de errores. Pvalor.

Prueba de Hipótesis (P.H.): conocida también como Dócima, es un proceso mediante el cual, a partir de los valores de una muestra aleatoria extraída de una población, se decide si se rechaza o no el supuesto que plantea el investigador para el parámetro o parámetros de esa población o poblaciones bajo estudio, pero con cierta probabilidad de error (riesgo) alr tomar una decisión.
La hipótesis nula Ho: es el "status quo", lo convencional, lo que sabemos de la población, lo aceptado hasta el momento, la igualdad.
La hipótesis alternativa Hi: es una alternativa a la hipótesis nula, implica cambio, a la hipótesis nula.

En la toma de decisiones se puede cometer dos tipos de errores.

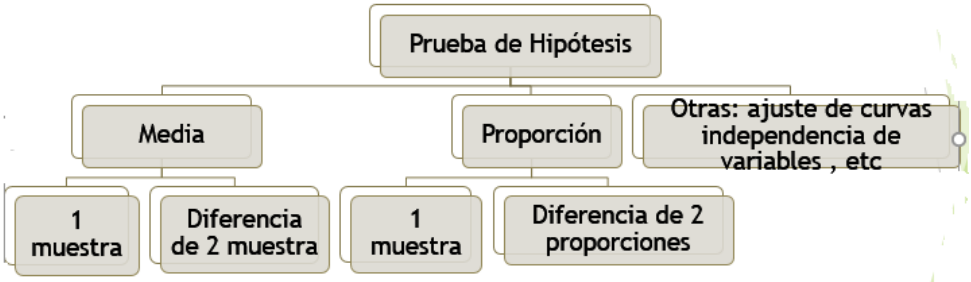
Toma de Decisión	Verdadera H_0	Falsa H_0
Rechazo H_0	**Error Tipo I (α)**	Correcto (1- β) Potencia
No Rechazo H_0	Correcto (1-α)	**Error Tipo II (β)**

Error Tipo I: es el error que se comete cuando Rechazamos la hipótesis nulas (Ho) siendo esta hipótesis verdadera.

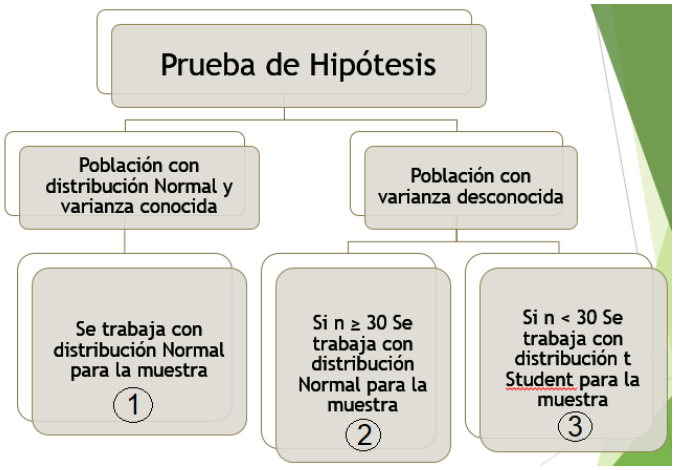
Error tipo II: es el error que se comete cuando No Rechazamos la hipótesis nula (Hi) siendo esta hipótesis falsa.
Nivel de significación (α): se denota por la letra griega alfa, es la probabilidad de cometer el error tipo I. Rechazamos la hipótesis nulas (Ho) siendo esta hipótesis verdadera.
Un nivel de significación del 5% significa que, en promedio, 5 de cada 100 veces que rechacemos la hipótesis nula estaremos equivocados (verdadero/falso). Justifique.
(β): Es la probabilidad de cometer el error tipo II o sea No Rechazar Ho siendo esta Falsa
Potencia de una prueba: es la probabilidad de rechazar Ho cuando realmente es Falsa P= 1-β.
El (P-valor) ó (valor-p) o (sig) es la probabilidad, de que el valor del estadístico calculado sea menor que el valor del estadístico crítico o tabulado. Informa sobre la compatibilidad de los datos con la Hipótesis Nula Ho.
Entendemos que a menor valor-p, se rechaza la Hipótesis Nula /H_o) mayor es la evidencia de los datos en contra de Ho Hipótesis Nula
No Rechazo no significa Aceptar Ej. Se afirma que dentro de una caja hay un animal y suponemos que es un paloma
Ho: Paloma
Si nos dicen que tiene alas nosotros "No Rechazamos" que sea una Paloma, pero no podemos "Aceptar" que sea una paloma, porque pudiera ser una mariposa, hay que aportar suficientes evidencias más y aun así siempre habrá probabilidad de equivocarnos.

4.2 Procedimiento lógico para realizar una P.H.

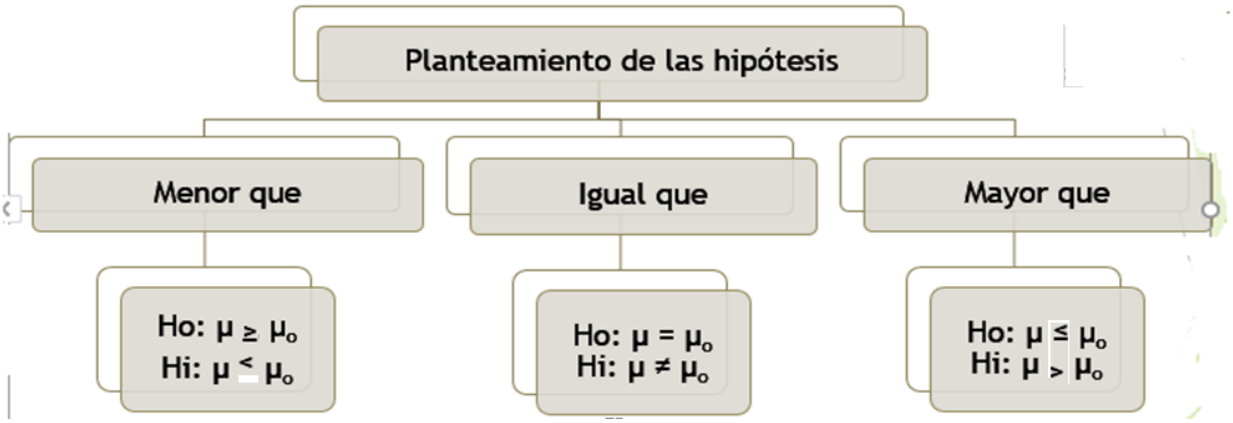
a) Identificar la medida descriptiva o la prueba que se va a contrastar.



b) Analizar los datos que se tienen de la población para saber con que distribución trabajar.



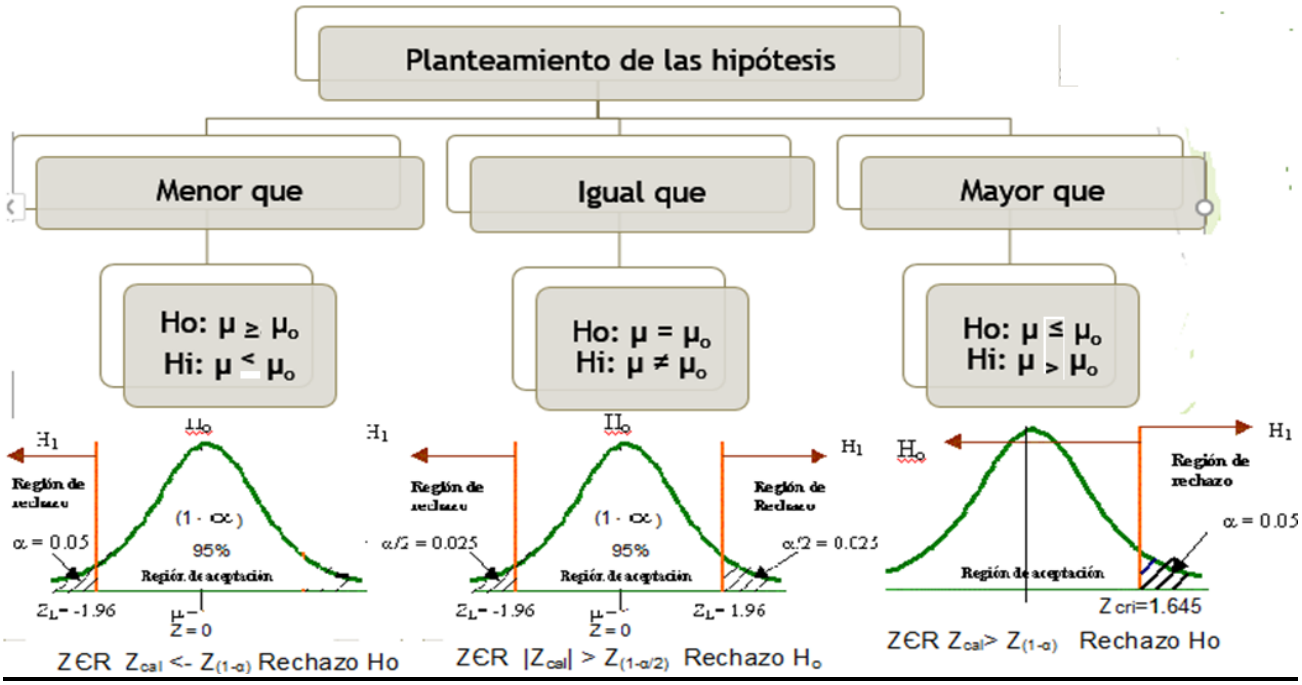
1er Paso. Plantear el tipo de hipótesis que se va contrastar.



2do Paso. Determinar el nivel de significación (α) probabilidad de cometer el error tipo I.

En las ciencias económicas generalmente se trabaja con un nivel de confianza (1-α) del 95%, lo cual implica que **α = 0.05**

3er Paso. Establecer la región crítica y la regla de decisión según las hipótesis planteadas.



4to Paso. Calcular el estadístico de prueba (estadístico calculado).

	Media		Proporción
	Varianza (σ^2) conocida	**Varianza (σ^2) desconocida y n<30**	**Para n suficientemente grande**
Una muestra	$Z=(\overline{X} - \mu_o) / (\sigma / \sqrt{n})$	$t = (\overline{X} - \mu_o) / (s / \sqrt{n})$	$Z_{cal}=(X-np_o)/(\sqrt{np_o(1-p_o)})$
Dos muestras	$Z = \frac{\lvert\bar{x}_1 - \bar{x}_2\rvert}{\sqrt{\left[\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right]}}$	$t = \frac{\lvert x^1 - \bar{x}^2\rvert}{\sqrt{\frac{(n_1 - 1)s_1^2 (n_2 - 1)s_2^2}{n_1 + n_2 - 2}\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$	$z_{cal} = \frac{\bar{p}_1 - \bar{p}_2 - D}{\sqrt{\frac{\bar{p}_1(1-\bar{p}_1)}{n_1} + \frac{\bar{p}_2(1-\bar{p}_2)}{n_2}}}$

5to. Paso. Tomar la decisión comparando el estadístico calculado con la regla de decisión o región crítica.

Si el estadístico calculado cae dentro de la región crítica No se Rechaza Ho.
Si el estadístico calculado cae fuera de la región crítica se Rechaza Ho.

4.3. Pruebas de hipótesis para medias de una muestra conociendo y desconociendo la varianza poblacional (σ^2) a partir del estadístico crítico y del P-valor.

Problema Inicial

Ejemplo 1. Los costos de un producto son como promedio $10.20. Se decide variar su método de producción y después se selecciona una muestra de 15 productos al azar, obteniendo un costo promedio de $10.32. Suponiendo que la población sigue una distribución Normal, con una varianza poblacional de 1.21$ ^{2} se podrá afirmar que han variado los costos después del cambio de método de producción.

1) Datos:
N(μ,σ^2) Distribución Normal con media poblacional de $10.20 y varianza poblacional Conocida σ^2 de 1.21 ^{2} **Caso 1**
μ= $10.20.
n= 15
$\overline{X}$ **= $10.32**
σ^2= 1.21 de donde $\sigma = \sqrt{1.21} = \$1.1$

PRUEBA DE HIPÓTESIS UTILIZANDO EL ESTADÍSTICO CRÍTICO

1er Paso. Planteamiento de la Hipótesis.

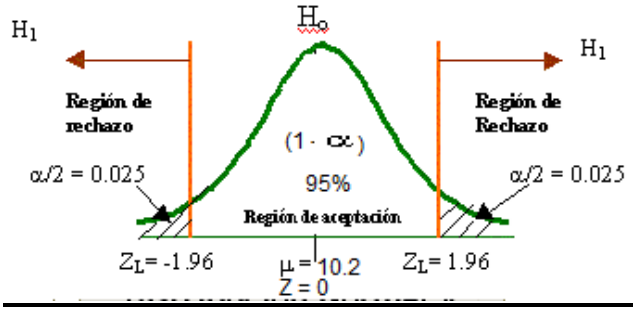
H_o: μ = $10.20 iguales
H_1: $\mu \neq$ $10.20 desiguales

2do Paso. Fijar el Nivel de significación α ó fijar el Nivel de confianza (1- α)

α= 0.05

3er Paso. Determinar la Zona Crítica

Si $|Z_{cal}| > Z_{(1-\alpha/2)}$ Rechazo H_o



Prueba de dos colas

4to Paso: Calcular el Estadístico $Z_{calculado}$ y $Z_{crítico}$

$Z_{cal} = (\overline{X} - \mu) / (\sigma / \sqrt{n})$

$Z_{cal} = (10.32 - 10.20) / (1.1/\sqrt{15}) = 0.42$ $Z_{cal} = 0.42$ $Z_{crítico} = Z_{(1-\alpha/2)} = Z_{(0.975)} = 1.96$

5to Paso: Toma de decisión.

Si $|Z_{cal}| > Z_{(crítico)}$ Rechazo H_o

$0.42 \leq 1.96$ Entonces No Rechazo H_o

R/ No se puede afirmar que variaron los costos después del cambio de forma de elaboración con un nivel de significación del 5% o con una confianza del 95%

Analicemos el problema con otro enfoque

Ejemplo 2 Los costos de un producto son como promedio $10.20 Se decide variar su forma de producción y después se selecciona una muestra de 15 productos al azar obteniendo un costo promedio de $10.32 y una desviación estándar 1.1$. Se podrá afirmar que han aumentado los costos después del cambio de método de producción.

Datos:

(μ, S^2) media poblacional de 10.20

μ= $10.20

n= 15 como n es menos que 30 y no se conoce la varianza poblacional (σ^2) se emplea la distribución t student **Caso 3**

$\overline{X}$ = $10.32

S = $1.1

p-valor = 0.062

PRUEBA DE HIPÓTESIS UTILIZANDO EL ESTADÍSTICO CRÍTICO

1er Paso: Planteamiento de la Hipótesis.

H_o: $\mu \leq$ $10.20 Iguales o menores

H_1: $\mu >$ $10.20 mayores

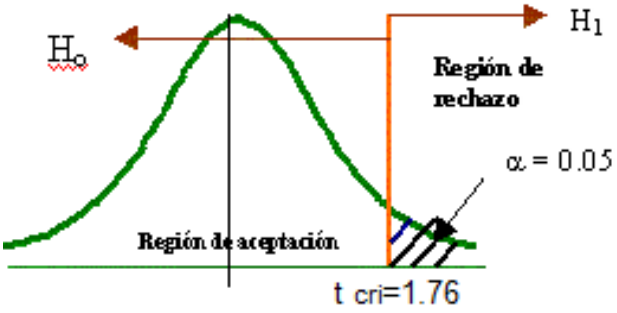
2do Paso: Fijar el Nivel de significación α

α= 0.05

3er Paso: Determinar la Zona Crítica y Regla de decisión

Como no se conoce la varianza poblacional y n es menor que 30 entonces se está en el caso 2 hay que trabajar con la distribución t student

Si $t_{cal} > t_{(n-1);(1-\alpha)}$ Rechazo H_o



4to Paso: Calcular el Estadístico t_{cal}

$t_{cal} = (\overline{X} - \mu) / (S/\sqrt{n})$

$t_{cal} = (10.32 – 10.20) / (1.1/\sqrt{15})$ $t_{cal} = 0.42$ $t_{crítico} = t_{(0.95;\ 14)} = 1.76$

5to Paso: Toma de decisión.

Si $t_{ca} > t_{(0.95;\ 14)}$ Rechazo H_o

$0.42 \leq 1.76$ Entonces No Rechazo H_o

R/ No se puede afirmar que los costos aumentaron después del cambio de forma de elaboración con un nivel de significación del 5%

PRUEBA DE HIPÓTESIS UTILIZANDO EL P-valor

1er Paso. Planteamiento de la Hipótesis.

H_o: $\mu \leq \$10.20$ Iguales o menores

H_1: $\mu > \$10.20$ mayores

2do Paso. Fijar el Nivel de significación α

α= 0.05

3er Paso. Regla de decisión Empleando el P-valor

Regla de decisión Si Pvalor < α Se rechaza Ho.

4to paso. Tomar la decisión (comparar Pvalor con α)

Si P-valor < α entonces Rechazar Ho

0.062 > 0.05 entonces No Rechazar Ho

R/ No se puede afirmar que variaron los costos después del cambio de forma de elaboración con un nivel de significación del 5% o con una confianza del 95%

4.4. Pruebas de hipótesis para comparación de medias de dos muestra conociendo y desconociendo la varianza poblacional (σ^2).

Ejemplo 3 Se desea comparar el peso (kg), de un tipo de artículos producidos en dos máquinas diferentes (1 y 2), entonces se seleccionan muestras al azar de 10 artículos de cada máquina obteniendo los siguientes resultados.

Máquina 1	**Máquina 2**
$\overline{X}$ **= 3.5 kg**	$\overline{X}$ **= 4.5 kg**
σ = 1.2	**σ = 1.1**
n = 10	**n = 10**

Existirá diferencias significativas con una confianza del 95% en el peso del articulo dependiendo del tipo de máquina que lo produce?

1er Paso: Planteamiento de las Hipótesis.

H_o: $\mu_1 = \mu_2$ (SON IGUALES LOS PESOS)

H_1: $\mu_1 \neq \mu_2$ (NO SON IGUALES LOS PESOS)

2do Paso: Fijar el Nivel de significación (α)

α= 0.05

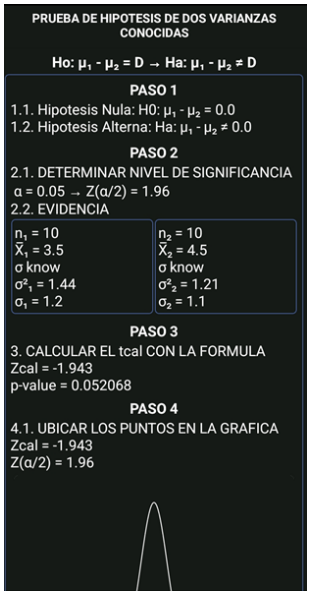
3er Paso: Regla de decisión Empleando el P-valor

Regla de decisión Si Pvalor < α Se rechaza Ho.

Con la Aplicación para moviles ADB







4to paso Tomar la decisión (comparar Pvalor con α)

Si P-valor < α entonces Rechazar Ho

0.052 > 0.05 entonces No Rechazar Ho

R/ No se puede afirmar que hay diferencias significativas entre los pesos de los artículos producidos por las maquinas 1 y 2, con un nivel de significación del 5% o con una confianza del 95%

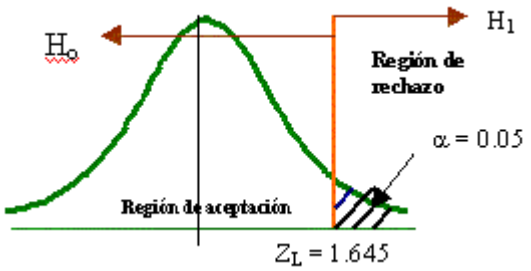
4.5 Pruebas de hipótesis para proporción de una población

Veamos otra Prueba de Hipótesis para la proporción de una población.
Ejemplo 4 Se afirma que un lote de piezas contiene menos o igual al 30% de piezas defectuosas. Para comprobarlo se revisaron 50 piezas seleccionadas al azar entre las que se detectaron 10 defectuosas ¿hay razón para mantener la afirmación con un nivel de significación del 0.05?
Datos:
p_o= 30% = 0.30 proporción de la población
n= 50
X= 10
α= 0.05

1er Paso: Planteamiento de la Hipótesis.
H_o: p ≤ 0.30 se mantienen los resultados
H_i: p > 0.30 no se mantienen los resultados

2do Paso: Fijar el Nivel de significación α
α= 0.05

3er Paso: Determinar la Zona Crítica
Si ZЄR | $Z_{cal} > Z_{(1-\alpha)}$ Rechazo H_o



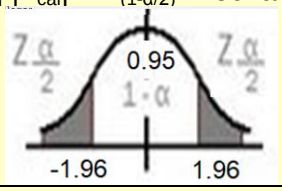
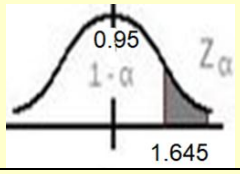
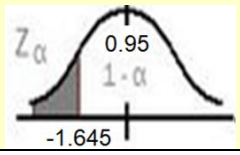
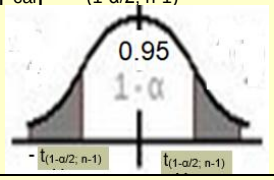
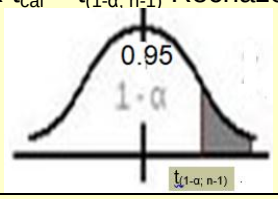
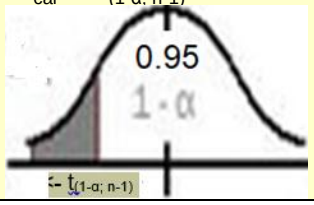
4to Paso: Calcular el Estadístico Z_{cal}
$Z_{cal} = (X - np_o) / (\in np_o(1-p_o))$
$Z_{cal} = (10 – 50(0.30)) / (\in 50(0.3)0.7)$ $Z_{cal} = -1.54$ $Z_{crit.} = Z_{(0.95)} = 1.645$

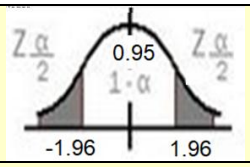
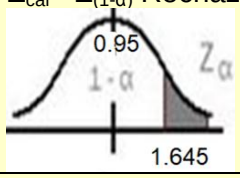
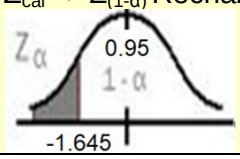
5to Paso: Toma de decisión.
Si $Z_{cal} > Z_{(0.95)}$ Rechazo H_o
-1.54 < 1.645 Entonces No rechazo H_o
Hay razón para considerar que se mantienen los resultados

CONCLUSIONES

El (P-valor) ó (valor-p) o (sig) es la probabilidad, calculada bajo el supuesto que Ho es verdadera. Informa sobre la compatibilidad de los datos con la Hipótesis Nula Ho.
Entendemos que a menor valor-p, se rechaza la Hipótesis Nula /H_o) mayor es la evidencia de los datos en contra de Ho Hipótesis Nula

1. TABLA RESUMEN

<table>
<tr><th colspan="4">Prueba de hipótesis de media con Distribución Normal y Varianza (σ) conocida o n> 30</th></tr>
<tr><th colspan="2">Se tiene el P-valor</th><th colspan="2">No se tiene el P-valor</th></tr>
<tr><th>Hipótesis</th><th>Regla de decisión</th><th>Región Crítica</th><th>Estadígrafo</th></tr>
<tr><td>H_o:μ=$μ_o$
H_i:μ≠$μ_o$</td><td rowspan="3">Si Pvalor < α rechazar Ho</td><td>Z∈R| $|Z_{cal}| > Z_{(1-α/2)}$ Rechazo H_o
</td><td rowspan="3">$Z=(\bar{X} - μ_o) / (σ/\sqrt{n})$</td></tr>
<tr><td>H_o:μ≤$μ_o$
H_i:μ>$μ_o$</td><td>Z∈R $Z_{cal} > Z_{(1-α)}$ Rechazo Ho
</td></tr>
<tr><td>H_o:μ≥$μ_o$
H_i:μ<$μ_o$</td><td>Z∈R| $Z_{cal} < -Z_{(1-α)}$ Rechazo Ho
</td></tr>
<tr><th colspan="4">Prueba de hipótesis de media con Varianza (σ) desconocida y n < 30</th></tr>
<tr><td>H_o:μ=$μ_o$
H_i:μ≠$μ_o$</td><td rowspan="3">Si Pvalor < α rechazar Ho</td><td>t Є R| $|t_{cal}| > t_{(1-α/2;\ n-1)}$ Rechazo H_o
</td><td rowspan="3">$t_{cal} = (\bar{X} - μ_o) / (s/\sqrt{n})$</td></tr>
<tr><td>H_o:μ≤$μ_o$
H_i:μ>$μ_o$</td><td>t Є R $t_{cal} > t_{(1-α;\ n-1)}$ Rechazo Ho
</td></tr>
<tr><td>H_o:μ≥$μ_o$
H_i:μ<$μ_o$</td><td>t Є R $t_{cal} < -t_{(1-α;\ n-1)}$ Rechazo Ho
</td></tr>
<tr><th colspan="4">Prueba de hipótesis de proporciones para n significativamente grande</th></tr>
<tr><td>H_o: p=p_o
H_i: p≠p_o</td><td>Si Pvalor < α rechazar Ho</td><td>Z∈R| $|Z_{cal}| > Z_{(1-α/2)}$ Rechazo H_o</td><td>$Z=(X-np_o)/(\sqrt{np_o(1-p_o)})$</td></tr>
</table>

			
H_o: $p \leq p_o$ H_i: $p > p_o$		ZЄR $Z_{cal} > Z_{(1-\alpha)}$ Rechazo Ho 	
H_o: $p \geq p_o$ H_i: $p < p_o$		ZЄR| $Z_{cal} < -Z_{(1-\alpha)}$ Rechazo Ho 	

ESTUDIO INDEPENDIENTE:

Descarque la Apk para moviles **ADB Estadística** e intente realizar las P.H. de los ejemplos desarrollados en la Conferencia con esta aplicación, interpretando los resultados de sus salidas.