Intro a la PH

- Decidir si una afirmación sobre una población parece verdadero o falso, basándonos en una muestra.

Ej:

← Alguien dice

Hipótesis nula (H_0)

- Afirmación que se asume verdadera hasta que haya evidencia en contra.

Hipótesis Alternativo (H_1)

- Lo contrario a H_0, la que se prueba.

Ejemplo 1:

Afirmación $\mu = ₡1300$

$H_0: \mu = ₡1300$

$H_1: \mu \neq 1300$ v $H_1: \mu < 1300$

$H_1: \mu > 1300$

Estimación con dos poblaciones

σ : Desviación estandar poblacional
S : Desviación estandar muestral

$\hat{p}_1$, $\hat{p}_2$ } proporciones muestrales

Diferencia de promedios

Para sacar el IC

CASO 1:

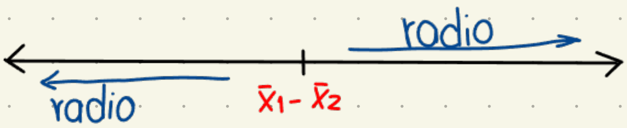
Fórmula	Uso	¿Cuándo?
$\overline{x_1} - \overline{x_2} \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$	IC: $\mu_1 - \mu_2$	$\overline{x_1}, \overline{x_2}$ se distribuye normalmente, $n_1, n_2 \geq 30$. σ_1, σ_2 se conocen.
$n \geq \frac{Z_{\alpha/2}^2(\sigma_1^2 + \sigma_2^2)}{r^2}$	Tamaño de la muestra.	$\overline{x_1}, \overline{x_2}$ se distribuye normalmente. σ_1, σ_2 se conocen.

Grupo 1 y Grupo 2: queremos comparar sobre si se pueden considerar si los grupos son iguales o parecidos a partir de los promedios

Nota: si a nivel de promedio son iguales entonces los grupos son iguales. A nivel estadística

centro: $\boxed{\overline{x_1} - \overline{x_2}}$ radio: $\pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$

$$\boxed{\overline{x_1} - \overline{x_2}} \pm Z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$



Nota: si $n_1, n_2 \geq 30$ y σ_1, σ_2 NO se conocen, se puede asumir $\sigma_1^2 = s_1^2$ y $\sigma_2^2 = s_2^2$

$$n \geq \frac{Z_{\alpha/2}^2(\sigma_1^2 + \sigma_2^2)}{r^2}$$

lo d l Cru l s D e m lirse lg a

CASO 2:

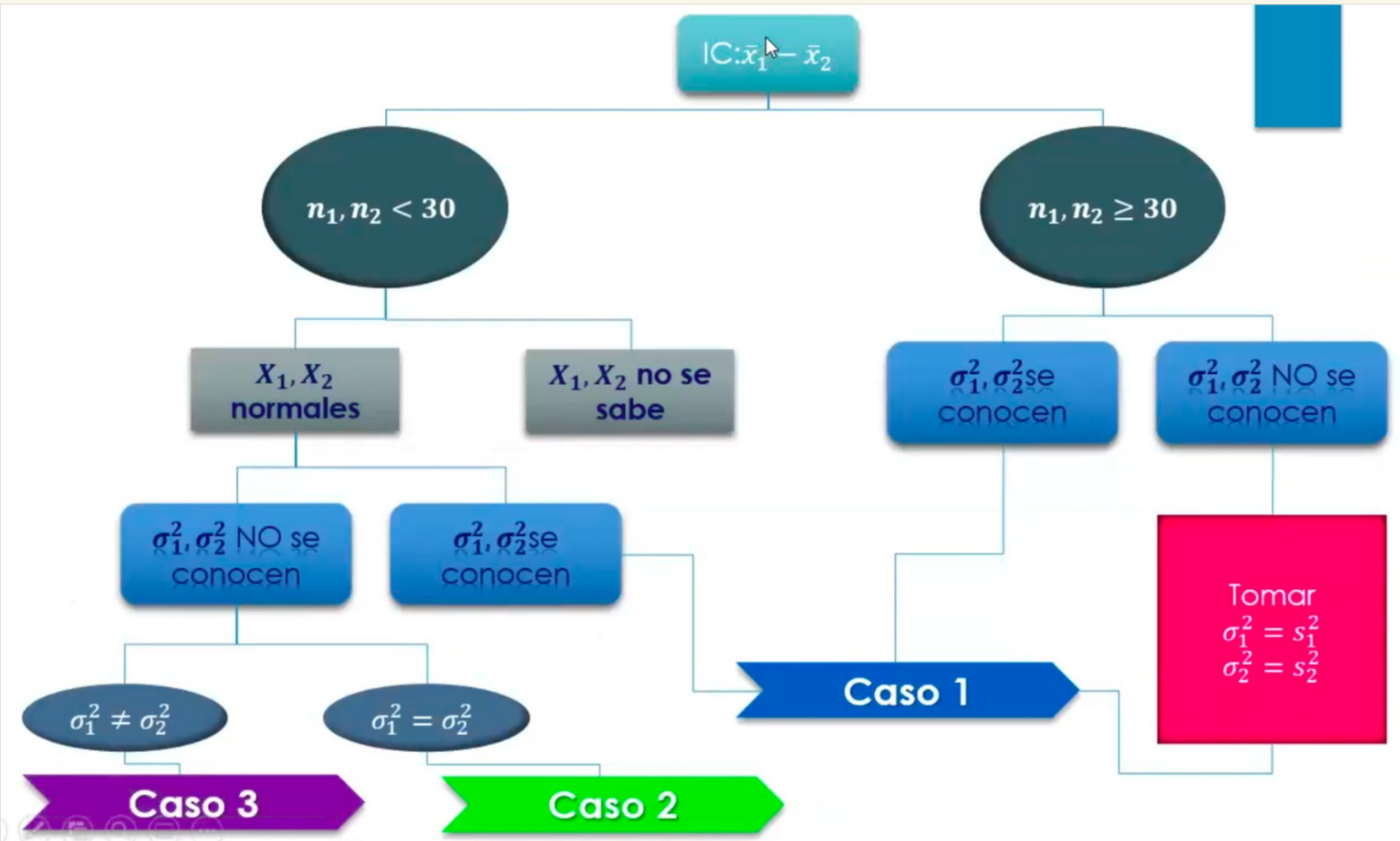
Fórmula	Uso	¿Cuándo?
$(\overline{x_1} - \overline{x_2}) \pm t_{\alpha/2,v} \sqrt{\frac{s_p{}^2}{n_1} + \frac{s_p{}^2}{n_2}}$ **Con:** $v = n_1 + n_2 - 2\ gl$	IC: $\mu_1 - \mu_2$	Poblaciones normales, $n_1 < 30$ o $n_2 < 30$, σ_1 y σ_2 no se conocen, pero se asumen $\sigma_1 = \sigma_2$ $s_p{}^2 = \frac{(n_1 - 1)s_1{}^2 + (n_2 - 1)s_2{}^2}{n_1 + n_2 - 2}$

t ratio

varianza muestral promedio

CASO 3:

Fórmula	Uso	¿Cuándo?
$\overline{x_1} - \overline{x_2} \pm t_{\alpha/2,v} \sqrt{\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}}$ **Donde:** $v = \frac{\left(\frac{s_1{}^2}{n_1} + \frac{s_2{}^2}{n_2}\right)^2}{\frac{\left(\frac{s_1{}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2{}^2}{n_2}\right)^2}{n_2 - 1}}$	IC: $\mu_1 - \mu_2$	Poblaciones normales, $n_1 < 30$ o $n_2 < 30$, σ_1 y σ_2 NO se conocen y se asume $\sigma_1 \neq \sigma_2$



IC Para el cociente de dos varianzas

Fórmula	Uso	¿Cuándo?
$\frac{s_2^2 f_{\alpha/2, v_1, v_2}}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{s_2^2}{s_1^2 f_{\alpha/2, v_2, v_1}}$	IC: $\frac{\sigma_2^2}{\sigma_1^2}$	$v_1 = n_1 - 1\ gl$ $v_2 = n_2 - 1\ gl$
$\frac{s_1^2 f_{\alpha/2, v_2, v_1}}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2 f_{\alpha/2, v_1, v_2}}$	$\overline{}\,2$	Poblaciones normales.

#Remplazar de forma adecuada

- Varianzas iguales si : $1 \in IC$
- Var ¬ Iguales si : $1 \notin IC$

IC Para 2 poblaciones

Ejemplo 10:

Un mismo examen se aplica a varios grupos de un curso. Para investigar la diferencia de rendimiento entre los grupos matutinos y los nocturnos se seleccionan al azar 45 estudiantes que reciben el curso en la mañana, y 38 que lo reciben en la noche. Entre los estudiantes de la mañana la nota promedio es 71.3 con una desviación estándar de 4.1, y entre los de la noche el promedio es 68.2 con una desviación estándar de 5.8. Llamemos población M al grupo de la mañana, y población N al de la noche. Como las muestras son suficientemente grandes, suponemos que las distribuciones de $\bar{X}_M$ y $\bar{X}_N$ son normales, y que $\sigma_M = 4.1$ y $\sigma_N = 5.8$. Determine IC de 95% para la diferencia de promedios. *R/]0.9 , 5.3[*

Datos:

$n_M = 45$ $\quad \bar{X}_M = 71,3$ $\quad S_M = 4.1$ $\quad \sigma_M = 4.1$

$n_N = 38$ $\quad \bar{X}_N = 68.2$ $\quad S_N = 5.8$ $\quad \sigma_N = 5.8$

$\alpha = 0.05$ $\quad Z_{0.025} = 1,95996$

$\alpha/2 = 0.025$

caso 1: $n_1 \wedge n_2 \geq 30$
o se conocen

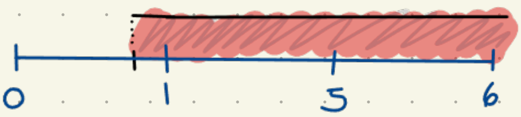
$$\bar{X}_1 - \bar{X}_2 \pm Z_{\alpha/2} \cdot \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\bar{X}_1 - \bar{X}_2 = \bar{X}_M - \bar{X}_N = 71.3 - 68.2 = 3.1$$

$$3.1 \pm (1{,}95996) \cdot \sqrt{\frac{4.1^2}{45} + \frac{5.8^2}{38}}$$

$$= [\,0.90\,,\ 5.299\,]$$

¿será válido asumir que los promedios de ambos grupos son iguales?



- En este caso, no, para que tenga validez $0 \in IC$.

¿Qué podemos concluir?

- Al ser ambos extremos positivos, se intuye que el grupo de la mañana tiene mejor promedio que la noche.

Ejemplo 11:

En el ejemplo 10, ¿de qué tamaño deben ser las muestreas para que el IC de 95% tenga radio no mayor que 2? *R/49 estudiantes de la mañana y 49 estudiantes de la noche*

Tamaño de la muestra. $\mathbf{n} \geq \frac{Z^2_{\alpha/2}\,(\sigma_1{}^2 + \sigma_2{}^2)}{r^2}$

Tenemos: $r = 2$ ← fines prácticos
$Z_{\alpha/2} = 1{,}95996$

$$n \geq \frac{(1{,}95996) \cdot \left((4.5)^2 + (5.8)^2\right)}{(2)^2} \Rightarrow n \geq 48{,}4520$$

$$n \geq 49$$

Ejemplo 15:

En el ejemplo 10, ¿Cuál es un IC de 95% para el cociente de las variancias en los dos grupos? ¿Puede afirmarse que la variación en la noche es mayor que en la mañana?

R/]1.0780 , 3.7798[

Grupo Mañana

$S_M = 4.1$

$V_M = 45-1 = 44$

$\alpha = 0.05$

Grupo Noche

$S_N = 5.8$

$V_n = 38-1 = 37$

$\alpha/2 = 0.025$

Calcular el IC de 95% para $\frac{\sigma_M^2}{\sigma_N^2}$

Fórmula: $\frac{S_M^2}{S_N^2} \cdot f_{0.025,37,44} < \frac{\sigma_M^2}{\sigma_N^2} < \frac{S_M^2}{S_N^2} \cdot \frac{1}{f_{0.025,44,37}}$

$f_{0.025,37,44} = 0.52954$

$f_{0.025,44,37} = 0.5387$

$$\left[\frac{4.1^2}{5.8^2} \cdot 0.52954 \,,\; \frac{4.1^2}{5.8^2} \cdot \frac{1}{0{,}5387}\right]$$

$$= [0{,}2646 \,,\; 0{,}9276]$$

Análogamente, si se desea calcular el IC para $\frac{\sigma_N^2}{\sigma_M^2}$

Izquierda: $f_{0.025, v_M, v_N} = f_{0.025, 44, 37} = 0.5387$

Derecha: $f_{0.025, v_N, v_M} = f_{0.025, 37, 44} = 0.52954$

$$\left]\frac{5.8^2}{4.1^2} \cdot (0.5387) \,,\; \frac{5.8^2}{4.1^2} \cdot \frac{1}{0.52954}\right[$$

$$= \left] 1.078 \,,\; 3.7791 \right[$$

Relación entre ambos:

IC Para $\frac{\sigma_M^2}{\sigma_N^2}$

$]0{,}2646 \,,\; 0{,}9276[$

IC Para $\frac{\sigma_N^2}{\sigma_M^2}$

$]1{,}078 \,,\; 3{,}7791[$

$\frac{1}{0{,}2646} = 3{,}7791$

$\frac{1}{0.9276} = 1{,}078$

PH con dos promedias

Población	PARÁMETROS		Tamaño de la muestra	ESTADÍSTICA	
	Media poblacional	Varianza poblacional		Media muestral	Varianza muestral
X_1	μ_1	σ_1^2	n_1	$\bar{X}_1$	s_1^2
X_2	μ_2	σ_2^2	n_2	$\bar{X}_2$	s_2^2

$\bar{X}_1 - \bar{X}_2$ es un estimador insesgado de $\mu_1 - \mu_2$, además $Var(\bar{X}_1 - \bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$.
La hipótesis nula para este tipo de pruebas es de la forma $\boldsymbol{H_0: \mu_1 - \mu_2 = d_0}$, donde d_0 se conoce como ***diferencia nula***. Además, se sabe que:

1) Si $\bar{X}_1$ y $\bar{X}_2$ siguen una distribución normal para muestras de tamaño n_1 y n_2 respectivamente, y se conocen σ_1 y σ_2 entonces

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - d_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

→ No e necesario conoce σ_1 y σ_2
pues: $\sigma_1 \approx S_1$

Se utiliza si:

- $n_1 \wedge n_2 \geq 30$
- $\bar{X}_1$ y $\bar{X}_2$ siguen una distribución normal.

2) Si las poblaciones X_1 y X_2 siguen una distribución normal y se desconocen las desviaciones poblacionales σ_1 y σ_2 pero **se suponen iguales**, entonces

$$T = \frac{\bar{X}_1 - \bar{X}_2 - d_0}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

sigue una distribución t *de Student* con $v = n_1 + n_2 - 2$ grados de libertad, donde s_p^2 es una estimación dada por el estadístico:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

3) Si las poblaciones X_1 y X_2 siguen una distribución normal y se desconocen las desviaciones poblacionales σ_1 y σ_2 pero **no se suponen iguales**, entonces

$$T = \frac{\bar{X}_1 - \bar{X}_2 - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

sigue una distribución $t\ de\ Student$ con υ grados de libertad, donde

$$\upsilon = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}}$$

AGREGAR REGLA DE DECISIÓN:

Si es bilateral:

Si $Z < Z_c = -K$ v
$Z > Z_c = K$
} se rechaza H_0

Estos son donde empiezan las áreas de rechazo.

PH Dif promedio: Muestras grandes

Se desea verificar si es verdad que en promedio las mujeres ganan $750 más por mes que los hombres en cierto trabajo. Los datos de las muestras que se tomaron para realizar el estudio, se muestran en la tabla:

Género	Tamaño de la muestra	Salario Mensual ($)	Desviación típica ($)
Mujer	75	10,550	880
Hombre	60	9,550	817

Recursos Humanos, por su parte, afirma que, en promedio, la diferencia en salarios mensuales entre las mujeres y los hombres es menor a $750. Con un nivel de significancia de 0.01, determine si lo dicho por Recursos Humanos tiene sustento estadístico o no.

Más que

Ejemplo 2

1. **Establecer H_0 y H_1 :**

H_0: Los sueldos mensuales de las mujeres son mayores en $750.

$$\mu_M = \mu_H + 750$$

$$\mu_M - \mu_H = 750$$

H_1: Los sueldos mensuales de las mujeres son mayores en menos de $750.

$$\mu_M < \mu_H + 750$$

$$\mu_M - \mu_H < 750$$

Se desea verificar si es verdad que en promedio las mujeres ganan $750 más por mes que los hombres en cierto trabajo. Los datos de las muestras que se tomaron para realizar el estudio, se muestran en la tabla:

Género	Tamaño de la muestra	Salario Mensual ($)	Desviación típica ($)
Mujer	75	10,550	880
Hombre	60	9,950	817

Recursos Humanos, por su parte, afirma que en promedio, la diferencia en salarios mensuales entre las mujeres y los hombres es menor a $750. Con un nivel de significancia de 0.01, determine si lo dicho por Recursos Humanos tiene sustento estadístico o no.

Ejemplo 2

2. Definir tipo de prueba y α:

- Prueba de extremo inferior
- $\alpha = 0.01$

3. Estadístico de prueba:

$$Z = \frac{(\overline{x}_1 - \overline{x}_2) - D}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \left\{ \begin{array}{l} D = \$\,750 \\ \overline{x}_1 = \$10{,}550 \\ \sigma_1 = s_1 = \$880 \\ n_1 = 75 \\ \overline{x}_2 = \$9{,}950 \\ \sigma_2 = s_2 = \$817 \\ n_2 = 60 \end{array} \right.$$

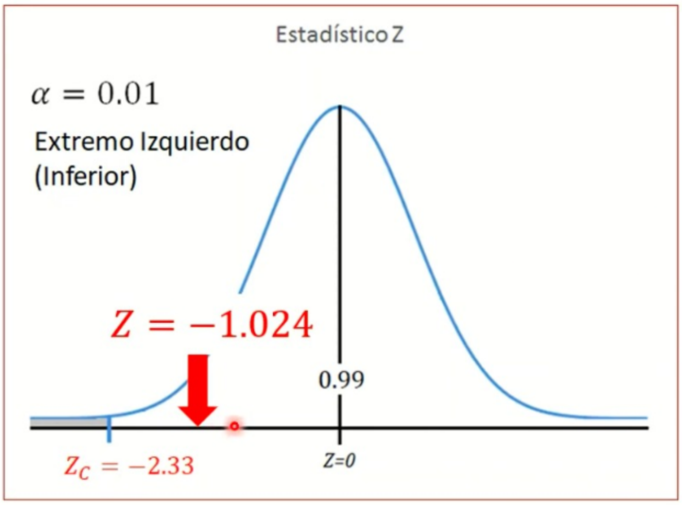
TEMA: PRUEBAS DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS. MUESTRAS GRANDES

$$Z = \frac{(10{,}550 - 9{,}950) - 750}{\sqrt{\frac{880^2}{75} + \frac{817^2}{60}}} = \frac{-150}{146.46}$$

$Z = -1.024$

4. Regla de decisión: Si $Z < Z_c = -2.33$ la H_0 deberá rechazarse.

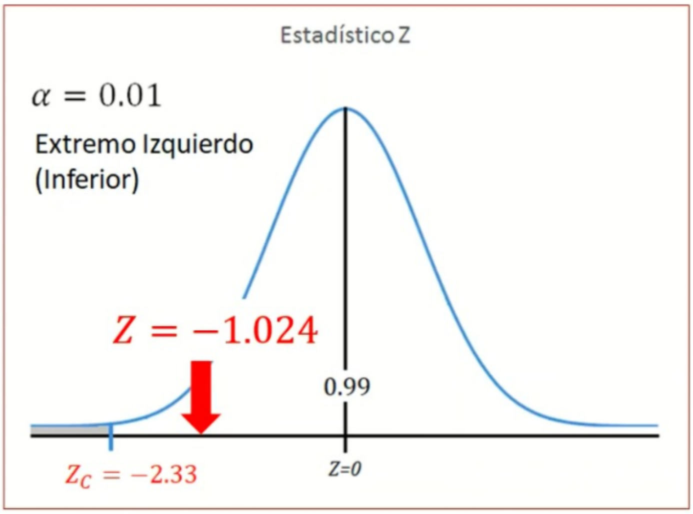
$$(Z = -1.024) > (Z_C = -2.33)$$



TEMA: PRUEBAS DE HIPÓTESIS PARA DIFERENCIA DE MEDIAS. MUESTRAS GRANDES

5. Decisión:

La hipótesis nula se acepta, ya que HAY EVIDENCIA ESTADÍSTICA, con una significancia del 1 % ($\alpha = 0.01$), de que la diferencia entre los sueldos mensuales promedio de las mujeres y de los hombres si es de $750 a favor de las mujeres.



Para muestras pequeñas

$\nu = n_1 + n_2 - 2 \leftarrow gl$

Ejemplo 1

El director del departamento de captación de alumnos para una universidad privada, está interesado en saber si existe una diferencia significativa entre el número de alumnos que inscriben semanalmente los promotores hombres y las promotoras mujeres. El equipo de promotores afirma que no hay diferencia en el número promedio de inscritos que logran tanto los hombres como las mujeres.

Con la finalidad de establecer una posición respecto a este tema, el director de captación decide hacer una prueba de hipótesis para soportar su postura ante su equipo de promotores. Para tal efecto, selecciona una muestra de los resultados de inscritos logrados en 9 semanas por los promotores hombres y de 8 semanas por las promotoras mujeres, obteniendo los resultados siguientes.

Promotores	Número de Inscritos (semanal)								
Hombres	12	10	14	12	13	12	13	12	14
Mujeres	13	11	11	13	13	14	10	12	

Ejemplo 1

Como no tenemos los datos de las medias y desviaciones estándar muestrales, debemos calcularlos:

Prom.	Número de Inscritos (semanal)								
Hom.	12	10	14	12	13	12	13	12	14
Muj.	13	11	11	13	13	14	10	12	

Variables Muestrales	Hombres	Mujeres
Tamaño de la muestra	9	8
Inscritos promedio	12.44	12.13
Desviación Estándar	1.24	1.36

1. **Establecer H_0 y H_1 :**

H_0: El número de inscritos semanales logrado por los promotores hombres y las promotoras mujeres es igual.

$$\mu_H = \mu_M$$

$$\mu_H - \mu_M = 0$$

H_1: El número de inscritos semanales logrado por los promotores hombres y las promotoras mujeres es diferente.

$$\mu_H \neq \mu_M$$

$$\mu_H - \mu_M \neq 0$$

Ejemplo 1

Variables Muestrales	Hombres	Mujeres
Tamaño de la muestra	9	8
Inscritos promedio	12.44	12.13
Desviación Estándar	1.24	1.36

2. **Definir tipo de prueba y α:**

- Prueba de dos extremos
- $\alpha = 0.05$

3. **Estadístico de prueba:**

$$S_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$S_p = \sqrt{\frac{(9-1)\cdot 1.24^2 + (8-1)\cdot 1.36^2}{9+8-2}}$$

$$S_p = \sqrt{\frac{(8)\cdot 1.5376 + (7)\cdot 1.8496}{15}}$$

$$S_p = \sqrt{\frac{25.248}{15}} = 1.297$$

$$S_p = 1.297$$

Ejemplo 1

Variables Muestrales	Hombres	Mujeres
Tamaño de la muestra	9	8
Inscritos promedio	12.44	12.13
Desviación Estándar	1.24	1.36

2. Definir tipo de prueba y α:

- Prueba de dos extremos
- $\alpha = 0.05$

3. Estadístico de prueba:

$S_p = 1.297$

$$t = \frac{(\overline{x}_1 - \overline{x}_2) - D}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \left\{ \begin{array}{l} D = 0 \\ \overline{x}_1 = 12.44 \\ n_1 = 9 \\ \overline{x}_2 = 12.13 \\ n_2 = 8 \end{array} \right.$$

$$t = \frac{(12.44 - 12.13) - 0}{1.297 \cdot \sqrt{\frac{1}{9} + \frac{1}{8}}} = \frac{0.31}{1.297 \cdot 0.4859}$$

$$t = 0.492$$

Ejemplo 1

4. Regla de decisión:

$$t_c = ?$$

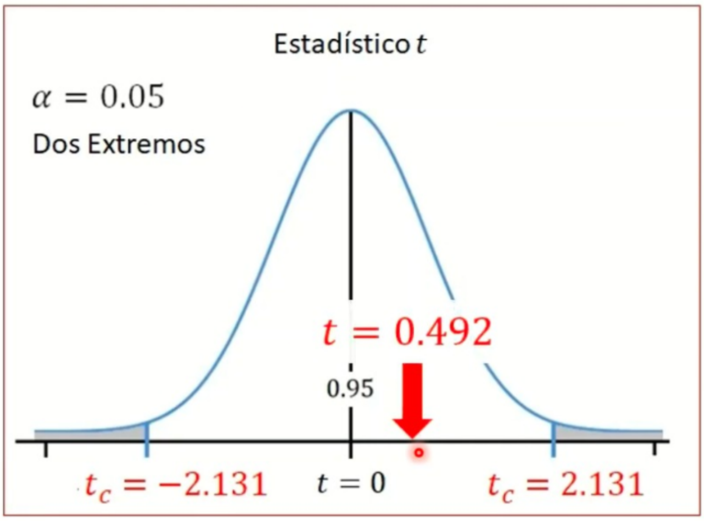
$$\alpha = 0.05$$

$$gl = 9 + 8 - 2 = 15$$

$$t_c = \pm 2.131$$

Si $t < t_c = -2.131 \vee t > t_c = 2.131$ la H_0 deberá rechazarse.

$$-2.131 < t = 0.492 < 2.131$$



Ejemplo 2

Un analista de automóviles hace un estudio para determinar el rendimiento del combustible, en $\frac{Km}{l}$, de dos marcas de automóviles pertenecientes a un mismo segmento. Tras hacer el ensayo en 14 automóviles de la marca R, encontró que el rendimiento promedio fue de $11.95\,\frac{Km}{l}$, con una desviación estándar de $1.49\,\frac{Km}{l}$. Por otro lado, para la marca M, los 12 automóviles estudiados arrojaron un rendimiento promedio de $13.65\,\frac{Km}{l}$, con una desviación estándar de $1.83\,\frac{Km}{l}$.

El fabricante de la marca M, le aseguró al analista que su marca ofrece un rendimiento promedio de $2.5\,\frac{Km}{l}$ más alto que la marca R. Por su parte, el fabricante de la marca R afirma que la diferencia es menor, lo cual se compensa con la comodidad de su marca.

Establecer una prueba de hipótesis que permita aceptar o rechazar lo que afirmó el fabricante de la marca M.

Ejemplo 2

Variables Muestrales	Marca del Automóvil	
	R	M
Tamaño de la muestra	14	12
Rendimiento Promedio (km/l)	11.95	13.65
Desviación Estándar (km/l)	1.49	1.83

1. **Establecer H_0 y H_1 :**

H_0: La diferencia de rendimiento entre la marca M y la marca R es de 2.5 Km/l

$$\mu_M - \mu_R = 2.5\ Km/l$$

H_1: La diferencia de rendimiento entre la marca M y la marca R es menor de 2.5 Km/l

$$\mu_M - \mu_R < 2.5\ Km/l$$

Mas qu

Ejemplo 2

Variables Muestrales	Marca del Automóvil	
	R	M
Tamaño de la muestra	14	12
Rendimiento Promedio (km/l)	11.95	13.65
Desviación Estándar (km/l)	1.49	1.83

2. **Definir tipo de prueba y α:**
   - Prueba de extremo inferior
   - $\alpha = 0.05$
3. **Estadístico de prueba:**

$$S_p = \sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}}$$

$$S_p = \sqrt{\frac{(12-1)\cdot 1.83^2 + (14-1)\cdot 1.49^2}{12+14-2}}$$

$$S_p = \sqrt{\frac{(11)\cdot 3.3489 + (13)\cdot 2.2201}{24}}$$

$$S_p = \sqrt{\frac{65.6992}{24}} = 1.655$$

$$S_p = 1.655$$

Mas qu

Ejemplo 2

Variables Muestrales	Marca del Automóvil	
	R	M
Tamaño de la muestra	14	12
Rendimiento Promedio (km/l)	11.95	13.65
Desviación Estándar (km/l)	1.49	1.83

2. **Definir tipo de prueba y α:**
   - Prueba de extremo inferior
   - $\alpha = 0.05$
3. **Estadístico de prueba:**

$S_p = 1.655$

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - D}{S_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \begin{cases} D = 2.5 \\ \bar{x}_1 = 13.65 \\ n_1 = 12 \\ \bar{x}_2 = 11.95 \\ n_2 = 14 \end{cases}$$

$$t = \frac{(13.65 - 11.95) - 2.5}{1.655 \cdot \sqrt{\frac{1}{12} + \frac{1}{14}}} = \frac{-0.8}{1.655 \cdot 0.3934}$$

$$t = -1.229$$

Ejemplo 2

4. Regla de decisión:

$t_c = ?$

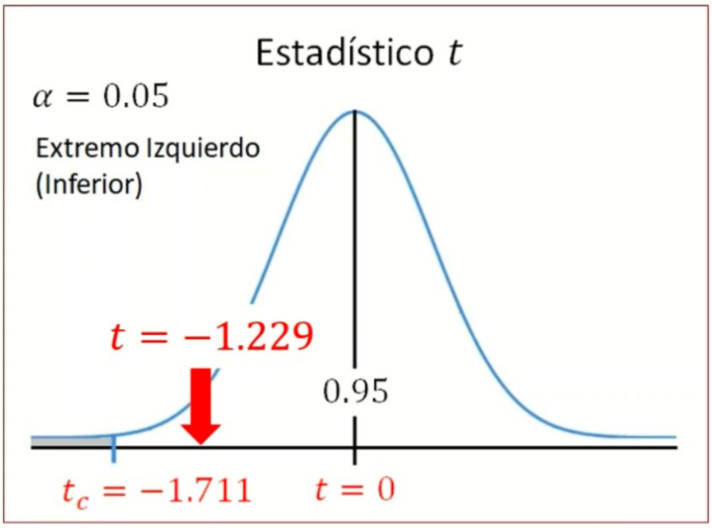
$\alpha = 0.05$

$gl = 12 + 14 - 2 = 24$

$t_c = -1.711$

Si $t < t_c = -1.711$ la H_0 deberá rechazarse.

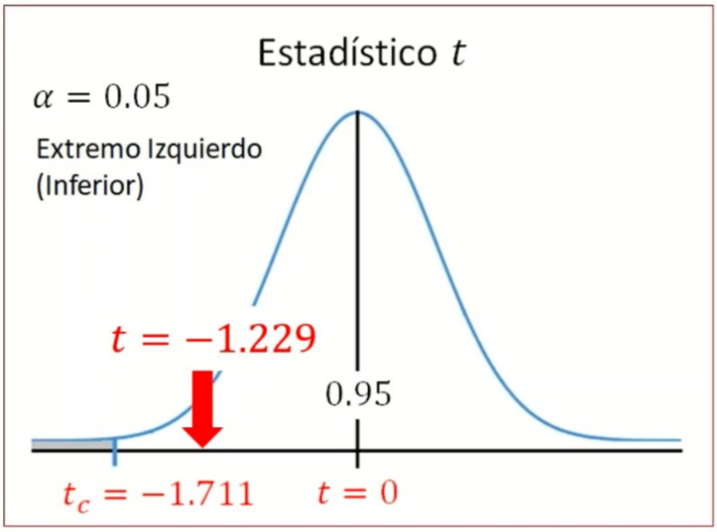
$-1.711 < (t = -1.229)$



Ejemplo 2

5. Decisión:

La hipótesis nula se acepta, ya que HAY EVIDENCIA ESTADÍSTICA, con una significancia del 5 % ($\alpha = 0.05$), de que la diferencia en el rendimiento de combustible de la marca M con respecto a la marca R es 2.5 Km/l.



PH con DOS varianzas

Ejemplo 10: comparación de varianzas

Se obtuvieron las estaturas de 21 mujeres y 31 hombres seleccionados aleatoriamente de una población de alumnos de cierta escuela. En la siguiente tabla se resumen los datos encontrados.

n_M

Sean σ_m^2 la varianza de las estaturas de las mujeres y σ_h^2 la varianza de las estaturas de los hombres. Asumiendo que las estaturas para ambos grupos se distribuyen de forma normal, determine, mediante una prueba de hipótesis y con nivel de significancia del 1%, si se debe suponer que $\sigma_m^2 = \sigma_h^2$

	Nº Alumnos	Media	Desv.Std
Mujeres	21	63,8	2,18
Hombres	31	69,8	1,92

Tenemos:

$\alpha = 0.01$

Planteamiento de hipótesis:

$H_0: \sigma^2_m = \sigma^2_n$ $\quad$ $H_1: \sigma^2_m \neq \sigma^2_n$

$\vee$ $\quad$ $\vee$

$H_0: \dfrac{\sigma^2_m}{\sigma^2_m} = 1$ $\quad$ $H_1: \dfrac{\sigma^2_m}{\sigma^2_m} \neq 1$

$r_0 = 1$ $\quad$ $r_0 = 1$

Tipo de prueba: DOS colas de Distribución F.

$\alpha/2 = 0.005 \rightarrow$ RR | RA | RR $\leftarrow \alpha/2 = 0.005$

f_{c1} 0,32016 $\quad$ 1,28 f_{obs} $\quad$ f_{c2} 2,82304

$f_{c1} = f_{0.005,20,30} = 0,32016$

$f_{c2} = f_{0.995,20,30} = 2,82304$

$\therefore$ RA $=]\,0,32016 , 2.82304\,[$

RR $=]\,0 , 0.32016\,[\cup]\,2,82304 , +\infty\,[$

Estandarizar el valor observado a terminos de la distribución F:

$F = \dfrac{s_1^{\,2}}{\underbrace{r_0 s_2^{\,2}}}; \ v_1 = n_1 - 1; \ v_2 = n_2 - 1$	Aplica $\underbrace{H_0: \dfrac{\sigma_1^{\,2}}{\sigma_2^{\,2}} = r_o}$

Como el $r_0 = 1$, entonces:

$$f_{obs} = \frac{S_m^2}{S_h^2 \cdot r} = \frac{(2.18)^2}{(1.92)^2 \cdot 1} = 1,2892$$

Como:

$$f_{c1} = 0,32016 < f_{obs} = 1.2892 < f_{c2} = 2,82304$$

Se concluye que $f_{obs} \in A$

Por lo tanto, no se encontró suficiente evidencia en contra para rechazar Ho, por lo que es posible asumir homocedasticidad de varianzas

Un vendedor de accesorios de computadoras afirma que los discos duros Modelo A tienen una vida útil promedio que supera a la vida útil promedio del Modelo B en más de un año. Para investigar se obtuvo la siguiente información:

Disco duro	n	$\bar{x}$	s
Modelo A	15	6,5	2,3
Modelo B	17	5,4	1,6

Suponga que la vida útil de los ambos modelos se comportan de forma normal.

1. Prueba la hipótesis de que las varianzas de las duraciones de ambos modelos son iguales, a un nivel de significancia del 5%.
2. ¿Es aceptable la información del vendedor?

Sugerencia; hacer una prueba de diferencia de medias, pero, como las muestras son pequeñas, se requiere distribución T, la cuál tiene dos fórmulas distintas para estandarizar que dependen de si las varianzas son iguales o diferentes, por tanto, antes de hacer la prueba de3 hipótesis para la diferencia de medias, se debe hacer primero una prueba de hipótesis para las varianzas que permita concluir si se pueden asumir como iguales o distintas.

CASO 2:

$$T_c = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

Con: $v = n_1 + n_2 - 2\ gl$

Poblaciones normales o $n_1, n_2 \geq 30$ o 40, σ_1, σ_2 no se conocen, pero se asumen $\sigma_1 = \sigma_2$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

CASO 3:

$$T_c = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Donde: $v = \dfrac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2 - 1}}$

Poblaciones normales o $n_1, n_2 \geq 30$ o 40 σ_1, σ_2 NO se conocen y se asume $\sigma_1 \neq \sigma_2$