

Resumen de materia

IC para I población

Sí no dan σ en
cualquier tipo de
ejercicio, se asume
 $\sigma = 0.05$

Pueden pedir:

- ① Intervalo de Confianza (IC)
- ② Tamaño de muestra
- ③ Confianza utilizada
- ④ Casos especiales

Los procedimientos a pesar de ser
similares, tienen algunas diferencias

Para I proporción solo Z

Para I varianza solo χ^2

IC: Intervalo de Confianza

D) Intervalo de Confianza (IC)

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

n = cantidad total

$$\hat{p} = \frac{\text{Todos los que cumplen}}{\text{Cantidad total}} \quad \text{Ej: "De } 30, 20 \text{ cumplen"}$$

$$\hat{p}_i = \frac{20}{30}$$

$$\hat{q} = 1 - \hat{p}$$

$\alpha = 1 - \text{Confianza}$, debe dividirse entre 2

"Determine IC para 95%"

$\rightarrow \alpha = 1 - 0,95 \rightarrow \alpha = 0,05 \rightarrow \frac{\alpha}{2} = 0,025$ y luego
en la app no importa el simbolo por ser IC

Ejemplo:

En una ciudad de 50 amas de casa 18 no utilizan el detergente X. Determine el IC de 95% para la verdadera proporción de amas de casa que no utilizan el detergente X. R/]0.22695, 0.49305[

$$n = 50 \quad 18 \text{ no utilizan} \rightarrow \frac{18}{50} = 0,36$$

$$\hat{p} = 0,36$$

$$\hat{q} = 0,64 \quad (1 - \hat{p}) \quad \alpha = 0,05 \quad (1 - \text{Confianza})$$

$$a = 0,36 - 1,95996 \cdot \underbrace{0,36 \cdot 0,64}_{50} = 0,22695$$

$$b = 0,36 + 1,95996 \cdot \underbrace{0,36 \cdot 0,64}_{50} = 0,49305$$

El IC para 95% corresponde a
]0.22695, 0.49305[

2) Tamaño de muestra

Hay 2 maneras, teniendo o no p y q

Caso 1: Teniendo \hat{p} y \hat{q}

Solo se aplica la formula

$$n \geq \frac{z_{\alpha/2}^2 pq}{r^2}$$

Caso 2: Sin tener p y q

Se asume $pq = 0,25$

Ejemplo:

En una ciudad de 50 amas de casa 18 no utilizan el detergente X .

Considerando el ejemplo 5.a, dado que el IC determinado es muy grande, ¿de qué tamaño debe ser la muestra si se desea tener un confianza de al menos 95% de que el error estimado al estimar la proporción sea menor que 0.02, sin importar el verdadero valor de p ? ¿es viable para la compañía obtener esta muestra? R/ 2401

Caso 1: Utilizando buenas estimaciones de p y q

$$n=50 \quad p=\frac{18}{50}=0,36 \quad q=0,64 \quad r=0,02 \quad \alpha=0,05$$

$$\frac{\alpha}{2}=0,025 \rightarrow Z_{0,025}= \pm 1,95996$$

$$n = \left(\frac{1,95996 \cdot \sqrt{0,36 \cdot 0,64}}{0,02} \right)^2$$

$$n \geq \frac{z_{\alpha/2}^2 pq}{r^2}$$

$$n \geq 2212,671288$$

n ≥ 2213 → Redondeo arriba hacia arriba

Caso 2: No importa valor de $p \rightarrow pq=0,25$

$$n = \left(\frac{1,95996 \cdot \sqrt{0,25}}{0,02} \right)^2 \quad \Rightarrow \quad n \geq 2400,90$$

$$\boxed{n \geq 2401}$$

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}} \quad E = \frac{b-a}{2}$$

Centro Radio

En estos problemas mas bien dan un intervalo $[a, b]$ entonces hay que igualar el radio a E y despejar el $Z_{\frac{\alpha}{2}}$, luego ese resultado se mete con 2 colas en la app en la X y luego se hace $1 - \text{resultado}$.

Ejemplo

Seguidamente se presenta una muestra de notas obtenidas en el examen de Admisión 2010 de una universidad:

$$72, 87, 28, 55, 92, 75, 83, 70, 30, 60, 53, 91, 90, 70, 70, 70, 55, 85 \quad \begin{matrix} 12 \text{ aprobados} \\ 18 \text{ totales} \end{matrix}$$

Con base en los datos el rector de la universidad determinó de manera correcta un IC para el porcentaje de estudiantes que aprobaron el examen (se aprueba con al menos 70), obteniendo: 0.448889, 0.884444[

Determine aproximadamente el nivel de confianza del IC que halló el rector. R/ 95%

$$h = 18 \quad p = \frac{12}{18} \quad q = \frac{2}{3} \quad \text{No se usa } z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

$$E = \frac{0,884444 - 0,448889}{2} = 0,217775 \quad (\text{calulando } E)$$

$$Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\frac{2}{3} \cdot \frac{1}{3}}{18}} = 0,217775 \quad \begin{matrix} \text{Reemplazando y} \\ \text{igualando a } E \end{matrix}$$

$$Z_{\frac{\alpha}{2}} = \frac{0,217775}{\sqrt{\frac{\frac{2}{3} \cdot \frac{1}{3}}{18}}} \quad \begin{matrix} \text{Despejando } Z_{\frac{\alpha}{2}} \\ \text{y} \end{matrix}$$

$$Z_{\frac{\alpha}{2}} = 1,9599975 \quad \begin{matrix} \text{Esto en la app de} \\ \text{la app con 2 colas} \end{matrix}$$

$$1 - 0,05 = 0,95 \rightarrow 95\%$$

$\mu = 0$	$\sigma = 1$
$x = 1.95996$	$2P(X > x) = 0.05000$

④ Caso especial:

Piden cual fue el tamaño de muestra UTILIZADO

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

$$p = \frac{q+a}{2}$$

$$E = \frac{b-a}{2} \quad \alpha = 1 - \text{Confianza} \quad \text{luego hacer } \frac{\alpha}{2}$$

$$q = 1 - p$$

Aquí suelen dar solo un IC $[a, b]$ y un nivel de confianza, entonces usando las fórmulas de arriba se debe despejar el n del radio, el centro NO se utiliza, se redondea hacia arriba

Ejemplo

20. Si un IC de 95% para la proporción de jóvenes de una cierta ciudad que no piensa tener hijos es $[0.1608, 0.2392]$, ¿cuál fue el tamaño de la muestra que se usó?

$$z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{pq}{n}} = E$$

$$p = \frac{q+a}{2}$$

$$z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

$$p = \frac{0.1608 + 0.2392}{2}$$

$$\sqrt{\frac{pq}{n}} = \frac{E}{z_{\frac{\alpha}{2}}}$$

$$p = 0.20 \quad q = 0.80$$

$$\frac{pq}{n} = \left(\frac{E}{z_{\frac{\alpha}{2}}} \right)^2$$

$$E = \frac{b-a}{2}$$

$$0.2392 - 0.1608$$

$$n = \frac{pq}{\left(\frac{E}{z_{\frac{\alpha}{2}}} \right)^2}$$

$$= 0.0392$$

$$\alpha = 0.05$$

$$\frac{\alpha}{2} = 0.025$$

$$z_{0.025} = 1.95996$$

$$n = \frac{0.20 \cdot 0.80}{\left(\frac{0.0392}{1.95996} \right)^2}$$

$$n = 399.98$$

$$n = 400$$

no importa el símbolo, pues se eleva a la 2

IC: I Varianza

II) Intervalo de Confianza (IC)

$$n = \text{Cantidad total} \quad \text{Varianza} \left\{ \begin{array}{l} \frac{(n-1)s^2}{\chi_{1-\alpha/2,\nu}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{\alpha/2,\nu}^2} \\ \text{con } \nu = n-1 \end{array} \right.$$

$$S^2 = \text{Varianza muestral} \quad \text{Desviación estandar} \left\{ \begin{array}{l} \sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha/2,\nu}^2}} < \sigma < \sqrt{\frac{(n-1)s^2}{\chi_{\alpha/2,\nu}^2}} \\ \text{con } \nu = n-1 \end{array} \right.$$

$$\chi_{C_1}^2 = \chi^2_{1-\frac{\alpha}{2}, \nu} \quad \chi_{C_2}^2 = \chi^2_{\frac{\alpha}{2}, \nu} \quad \text{el símbolo en la app siempre}$$

Si dan S , se eleva al cuadrado para obtener s^2

Si dan datos, se mete en la calcu

Ejemplo para Varianza

5. Para controlar el buen embolsado de sus productos, un productor de fertilizantes toma una muestra de 15 bolsas del mismo, obteniendo una desviación estándar de 0.50 kg. Suponga que los pesos de las bolsas siguen una distribución normal.

(a) Determine un intervalo de confianza del 98% para la variancia de los pesos de las bolsas de fertilizante.

$$R/ [0.120105, 0.751004[\quad \frac{(n-1)s^2}{\chi_{1-\alpha/2,\nu}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{\alpha/2,\nu}^2}$$

$$n = 15 \quad S = 0.50 \quad \alpha = 0.02 \quad \frac{\alpha}{2} = 0.01$$

$$\nu = 14 \quad s^2 = 0.25 \quad 1 - \frac{\alpha}{2} = 0.99$$

$$\chi^2_{0.99, 14} = 29,14129 \quad \chi^2_{0.01} = 9,66073$$

$$a = \frac{14 \cdot 0.25}{29,14129} = 0.120105 \quad b = \frac{14 \cdot 0.25}{9,66073} = 0.751004$$

R/ El IC para 98% corresponde a

$$[0.120105, 0.751004[$$

Ejemplo para desviación estandar

Un medio afirma que los pesos de niñas de 3 años de edad de cierta provincia son muy similares. Para analizar esta información se toma el peso en kilogramos de 10 niñas:

14.5, 11.6, 12.8, 15.1, 14.2, 13.7, 12.9, 13.8, 14.1, 11.9

Suponga que el peso de las niñas de tres años sigue una distribución normal. Determine el IC de 95% para la desviación estándar de los pesos de las niñas de 3 años. R/]0.779006, 2.06759 [

$$n = 10 \quad gl = 9 (n-1) \quad \alpha = 0,05 \\ s^2 = 1,28 \quad (\text{calculadora})$$

para σ : $\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}} \right]$

34, 6 [

$$\frac{\alpha}{2} = \frac{0,05}{2} = 0,025$$

Para $a = 1 - 0,025 = 0,975$
 $\chi^2_{0,975, 9} = 19,02277$

Para $b = 0,025$
 $\chi^2_{0,025, 9} = 2,70039$

para σ : $\left[\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2, n-1}}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2, n-1}}} \right]$

$$\left[\sqrt{\frac{q, 1,28}{19,02277}}, \sqrt{\frac{q, 1,28}{2,70039}} \right]$$

[30,779, 2,0675 [

$X \sim \text{ChiSq}(v)$

$v=9$

$x=2,70039$

$P(X < x) = 0,025$

$X \sim \text{ChiSq}(v)$

$v=9$

$x=19,02277$

$P(X < x) = 0,975$

NO existe para IC
② Tamaño de muestra de la varianza

③ Confianza utilizada

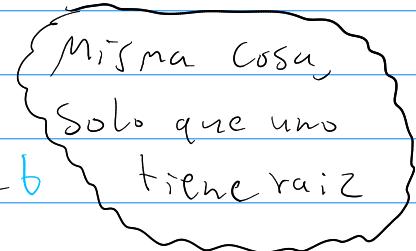
Aquí dan un intervalo $[a, b]$ pero la fórmula tiene 2 partes entonces hay 2 maneras de hacerlo:

$$\left[\begin{array}{l} \left\{ \frac{(n-1)s^2}{\chi_{1-\alpha/2,\nu}^2} < \sigma^2 < \frac{(n-1)s^2}{\chi_{\alpha/2,\nu}^2} \right\} \\ \text{con } \nu = n-1 \end{array} \right]$$

$$\left[\begin{array}{l} \left\{ \sqrt{\frac{(n-1)s^2}{\chi_{1-\alpha/2,\nu}^2}} < \sigma^2 < \sqrt{\frac{(n-1)s^2}{\chi_{\alpha/2,\nu}^2}} \right\} \\ \text{con } \nu = n-1 \end{array} \right]$$

Para Varianzas

$$\begin{aligned} (n-1).s^2 = a & \quad \vee \quad (n-1).s^2 = b \quad \text{Ojo!!!} \\ \chi^2_{1-\frac{\alpha}{2}, \nu} & \quad \vee \quad \chi^2_{\frac{\alpha}{2}, \nu} \end{aligned}$$



Para desviaciones estandar

$$\begin{aligned} \sqrt{(n-1).s^2} = a & \quad \vee \quad \sqrt{(n-1).s^2} = b \\ \chi^2_{1-\frac{\alpha}{2}, \nu} & \quad \vee \quad \chi^2_{\frac{\alpha}{2}, \nu} \end{aligned}$$

Luego de elegir la forma, se debe despejar el $\chi^2_{1-\frac{\alpha}{2}, \nu}$ o $\chi^2_{\frac{\alpha}{2}, \nu}$ y el resultado meterlo en la app con en la χ con cola izquierda <

Antes de hacer I - resultado de meter el despeje de χ^2 en la app hay que despejar también el $1-\frac{\alpha}{2}$ o $\frac{\alpha}{2}$

$$I - \frac{\alpha}{2} = \text{resultado}$$

$$\frac{\alpha}{2} = \text{resultado}$$

$$\frac{\alpha}{2} = I - \text{resultado}$$

$$\alpha = 2 \cdot \text{resultado}$$

$$\alpha = 2(I - \text{resultado})$$

y al final hacer $I - \alpha$ para obtener la confianza

Ejemplos abajo

Ejemplo de desviación estandar

[4 puntos] La duración de las llamadas que entran a una central de servicio al cliente sigue una distribución normal. En una muestra aleatoria de 8 de estas llamadas se observan duraciones de 151, 153, 175, 134, 170, 172, 156 y 114 segundos. Con estos datos, se ha calculado un intervalo de confianza I para la desviación estándar poblacional.

Determine el nivel de confianza asociado al intervalo $I =]14.6725, 37.38[$

Usando a

$$n=8 \quad v=7 \quad s^2 = 482,69673$$

a b

a b

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,v}}} < \sigma^2 < \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,v}}}$$

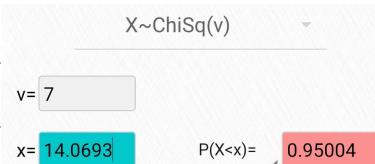
con $v = n - 1$

$$\sqrt{\frac{7 \cdot 482,69673}{\chi^2_{1-\frac{\alpha}{2},7}}} = 18,6725 \quad \left. \right\} \text{Igualando a}$$

$$\frac{7 \cdot 482,69673}{\chi^2_{1-\frac{\alpha}{2},7}} = 18,6725^2 \quad \left. \right\} \text{Despejando el } \chi^2$$

$$\chi^2_{1-\frac{\alpha}{2},7} = \frac{7 \cdot 482,69673}{18,6725}$$

$$\chi^2_{1-\frac{\alpha}{2},7} = 14,0693 \quad \left. \right\} \text{Esto en la app con } \leftarrow$$



$$\begin{aligned} 1 - \frac{\alpha}{2} &= 0,95007 \\ \frac{\alpha}{2} &= 1 - 0,95007 \\ \alpha &= 2(1 - 0,95007) \\ \alpha &= 0,10 \end{aligned} \quad \left. \right\} \begin{array}{l} \text{Igualar el resultado y} \\ \text{despejar } \alpha \end{array}$$

$$1 - 0,10 = 0,90 \rightarrow \boxed{90\%} \quad \left. \right\} \begin{array}{l} \text{Haciendo} \\ 1 - \alpha \end{array}$$

Mismo ejemplo pero usando b

[4 puntos] La duración de las llamadas que entran a una central de servicio al cliente sigue una distribución normal. En una muestra aleatoria de 8 de estas llamadas se observan duraciones de 151, 153, 175, 134, 170, 172, 156 y 114 segundos. Con estos datos, se ha calculado un intervalo de confianza I para la desviación estándar poblacional.

Determine el nivel de confianza asociado al intervalo $I =]14.6725, 37.38[$

Usando b

$$n=8 \quad v=7 \quad s^2 = 482,69673$$

a b

a b

$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2, \nu}}} < \sigma^2 < \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2, \nu}}}$$

con $\nu = n - 1$

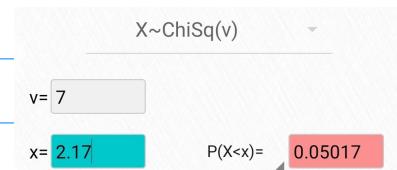
$$\sqrt{\frac{7 \cdot 482,69673}{\chi^2_{1-\frac{\alpha}{2}, 7}}} = 37,38 \quad \} \text{ Igualando b}$$

$$\frac{7 \cdot 482,69673}{\chi^2_{1-\frac{\alpha}{2}, 7}} = (37,38)^2$$

$$\chi^2_{1-\frac{\alpha}{2}, 7} = \frac{7 \cdot 482,69673}{(37,38)^2}$$

} Despejando
el χ^2

$$\chi^2_{1-\frac{\alpha}{2}, 7} = 2,17 \quad \} \text{ Esto en la
ape con } \leftarrow$$



$$\begin{aligned} \frac{\alpha}{2} &= 0,05 \\ \alpha &= 0,10 \end{aligned} \quad \} \text{ Igualar el
resultado y
despejar } \alpha$$

$$1 - 0,10 = 0,90 \rightarrow \boxed{90\%} \quad \} \text{ Haciendo } 1 - \alpha$$

IC para 2 poblaciones

Pueden pedir:

- 1) Intervalo de Confianza (IC)
- 2) Interpretar el IC
- 3) Tamaño de muestra
- 4) Confianza utilizada
- 5) Casos especiales

Los procedimientos a pesar de ser similares, tienen algunas diferencias

Para 2 medias se utilizan Z y t -student

Para 2 proporciones solo se utilizar Z

Para todos los IC de 2 poblaciones

Si Todo el IC $\subset [0, 1]$ o $[0.92, 30]$

Si Todo el IC $\subset [-3.5, -1]$

Si hay > 0 en C_0 , No se sube $30, -2$

Para 2 varianzas solo se usa F

Para IC de 2 varianzas

Se asumen varianzas iguales

Si $I \in IC$, diferentes si $I \notin IC$

Iguales $\rightarrow IC: [0.73, 1.13], I \in IC$

Diferentes $\rightarrow IC: [0.50, 0.70], I \notin IC$

El símbolo \supset en la app es irrelevante para cualquier distribución en IC

IC; 2 medias

Existen 3 casos y 3 formulas distintas

Caso 1: $n_1, n_2 > 30$ y se conocen las desviaciones estándar poblacionales σ_1 y σ_2 , mejor dicho Si no dicen si son iguales o diferentes

② Intervalo de Confianza (IC) para Caso 1

Se sabe que el peso de los sacos de cemento de dos marcas A y B siguen distribuciones normales con medias desconocidas y desviaciones estándar poblacionales $\sigma_A = 4\text{kg}$. y $\sigma_B = 5\text{kg}$. respectivamente. Pese a que ambas marcas reportan en sus sacos el mismo peso, un investigador piensa que los de la marca A tienen mayor peso que los de la marca B, por lo que hace un estudio tomando muestras aleatorias de ambos tipos de sacos. Segundo su estudio, en una muestra de 50 sacos de cemento de la marca A, se obtuvo un peso promedio de 49,8 kg; mientras que en una muestra de 40 sacos de la marca B el peso promedio fue de 47,5 kg.

Se
Conoce
la
desv.
estándar

a) [3 puntos] Determine un intervalo de confianza del 95% para la diferencia entre el peso medio de los sacos de cemento de la marca A y los de la marca B.

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

$$n_1 = 50 \quad n_2 = 40 \quad \alpha = 0,05 \quad \bar{x}_1 - \bar{x}_2 \geq 3,0$$

$$\bar{x}_1 = 49,8 \quad \bar{x}_2 = 47,5 \quad \frac{\sigma}{\sqrt{n}} = 0,625$$

$$\sigma_1 = 4 \quad \sigma_2 = 5 \quad Z_{0,025} = \pm 1,95996$$

$$\alpha = 49,8 - 47,5 - 1,95996. \sqrt{\frac{4^2}{50} + \frac{5^2}{40}} = 0,3997$$

$$\alpha = 49,8 - 47,5 + 1,95996. \sqrt{\frac{4^2}{50} + \frac{5^2}{40}} = 9,2052$$

[ET] IC para 95% corresponde a $[0,3997, 9,2052]$

③ Interpretar el IC

b) [1 punto] ¿Considera usted que la evidencia respalda (con un nivel de confianza del 95%) lo que pensaba el investigador? Justifique su respuesta.

Si pues todo el IC > 0, $A > B$

↙ Solo el Caso
1 tiene

2) Tamaño de muestra

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 $\sigma_N = 5.8$. ~~desviación~~
~~de los datos~~

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

$$\begin{array}{ll} n_1 = 45 \geq 30 & n_2 = 38 \geq 30 \\ \bar{x}_1 = 71,3 & \bar{x}_2 = 68,2 \\ \sigma_1 = 4,1 & \sigma_2 = 5,8 \end{array}$$

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

$$r = 0,05 \rightarrow z = 0,025 \rightarrow Z_{0,025} = \pm 1,95996 \quad r = 2$$

$$n \geq \frac{(1,95996)^2 \cdot [(4,1)^2 + (5,8)^2]}{2^2}$$

$$n \geq 48,4520$$

$$n \geq 49$$

$$\boxed{n_M \geq 49 \quad \wedge \quad n_N \geq 49}$$

3) Caso especial.

Ejemplo 62. Un IC del 90 % para la diferencia de promedios ($\mu_1 - \mu_2$) es $[165.5, 192.9]$. Suponga que σ_1 y σ_2 son conocidos. Si las muestras utilizadas en el cálculo del IC son ambas de tamaño 50, determine el valor $\bar{x}_1 - \bar{x}_2$ observado en las muestras y el valor aproximado de $\sigma_1^2 + \sigma_2^2$. → Despejar ese nada mas

$$\boxed{\bar{x}_1 - \bar{x}_2 = \frac{165.5 + 192.9}{2} = 179.2}$$

$$n_1 = n_2 = 50 \quad \alpha = 0.20$$

$$\frac{\alpha}{2} = 0.05 \quad z_{0.05} = 1.69485$$

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

$$E = \frac{b-a}{2}$$

$$E = \frac{192.9 - 165.5}{2} = 13.7$$

$$1.69485 \cdot \sqrt{\frac{\sigma_1^2}{50} + \frac{\sigma_2^2}{50}} = 13.7$$

$$\sqrt{\frac{\sigma_1^2}{50} + \frac{\sigma_2^2}{50}} = \frac{13.7}{1.69485}$$

$$\frac{\sigma_1^2 + \sigma_2^2}{50} = \left(\frac{13.7}{1.69485} \right)^2$$

$$\sigma_1^2 + \sigma_2^2 = \left(\frac{13.7}{1.69485} \right)^2 \cdot 50$$

$$\boxed{\sigma_1^2 + \sigma_2^2 = 3468.63}$$

Caso 2: $h_1 < 30$ y $h_2 > 30$ y no se conocen las desviaciones estándares poblacionales σ_1 y σ_2 , mejor dicho si se dice que las varianzas poblacionales son iguales y con ver que el n_1 o el n_2 sea < 30

Un profesor considera que el rendimiento promedio (nota promedio) de los estudiantes de Computación en el curso de Matemática Elemental es superior en al menos 9 puntos al rendimiento promedio de los estudiantes de otras carreras. Para analizar esto se tomó una muestra de estudiantes que cursaron el curso el año pasado, obteniendo los siguientes datos:

Estudiantes	tamaño de muestra	Rendimiento promedio observado (\bar{x})	Desviación estándar (s)
De computación:	19	78 puntos	4.3 puntos
De otras carreras:	17	65 puntos	4.7 puntos

Suponga que el rendimiento promedio en el curso de Matemática Elemental, tanto en Computación como en otras carreras, se distribuye normalmente.

Encuentre un intervalo de confianza de 95% para la diferencia entre los promedios de las notas de ambos tipos de estudiantes. Asuma varianzas poblacionales iguales. (5 puntos)

muestra

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, \nu} \sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

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

y $\nu = n_1 + n_2 - 2$

$$h_1 = 19 \quad h_2 = 17 \quad V = 19 + 17 - 2 = 34$$

$$\bar{x}_1 = 78 \quad \bar{x}_2 = 65 \quad \alpha = 0.05$$

$$s_1 = 4.3 \quad s_2 = 4.7 \quad \frac{s}{\sqrt{2}} = 0.025 \rightarrow t_{0.025, 34} = 2.03229$$

$$S_p^2 = \frac{(19 - 1) \cdot 4.3^2 + (17 - 1) \cdot 4.7^2}{19 + 17 - 2} \approx 20.1891$$

$$a = 78 - 65 - 2.03229 \cdot \sqrt{\frac{20.1891 + 20.1891}{19 + 17}} = 9.9519$$

$$b = 78 - 65 - 2.03229 \cdot \sqrt{\frac{20.1891 + 20.1891}{19 + 17}} = 16.0981$$

RI El IC para 95% corresponde a

$$[9.9519, 16.0981]$$

Caso 3: n_1 \neq n_2 y no se conocen las desviaciones estándares poblacionales σ_1 y σ_2 , mejor dicho Si le dicen que las varianzas poblacionales son diferentes y con ver que el n_1 o el n_2 sea <math><30</math>

Ejemplo 65. La universidad Bienestar Seguro tiene dos fórmulas para examen de admisión que pretende utilizar durante los próximos tres años. Sin embargo, un profesor de estadística de esa institución afirma que la fórmula A va a tener un mejor rendimiento promedio que la fórmula B. Ante esto, la universidad aplicó las fórmulas a un grupo de estudiantes, con lo cual obtuvo los siguientes resultados:

Fórmula	Tamaño de la muestra	Media muestral	Desviación muestral
A	21	65	24
B	17	63	15

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2, \nu} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

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

1. Si se supone que las variancias no son iguales, encuentre un intervalo de confianza del 95 % para la diferencia de promedios de notas entre la fórmula A y la fórmula B.

$$n_1 = 21 \quad n_2 = 17 \quad \alpha = 0.05$$

$$\bar{x}_1 = 65 \quad \bar{x}_2 = 63 \quad \frac{s}{2} = 0.025$$

$$s_1 = 24 \quad s_2 = 15$$

$$v = \left(\frac{24^2}{21} + \frac{15^2}{17} \right)^2$$

$$\frac{\left(\frac{24^2}{21} \right)^2}{21-1} + \frac{\left(\frac{15^2}{17} \right)^2}{17-1} = 34.0789$$

$$+_{0.025, 34.0789} = \pm 2.03274$$

$$a = 65 - 63 - 2.03274 \cdot \sqrt{\frac{24^2}{21} + \frac{15^2}{17}} = -10.9586$$

$$b = 65 - 63 + 2.03274 \cdot \sqrt{\frac{24^2}{21} + \frac{15^2}{17}} = 19.9586$$

IV El IC para 95% corresponde a]-10.9586, 19.9586[

② Interpretar el IC

2. ¿Es aceptable la afirmación del profesor?

IC es positivo y negativo, no

IC; 2 Proportiones

$n_i = \text{Cantidad total}$

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

$\hat{p}_i = \frac{\text{Todos los que cumplen}}{\text{Cantidad total}}$

$$\hat{q}_i = 1 - \hat{p}_i$$

El orden en el que se elige el p_1 y p_2 no es relevante siempre y cuando se interprete bien el IC

D) Intervalo de Confianza (IC)

Para un modelo de tostadores se encontró que 31 de 126 unidades vendidas fueron devueltas durante el primer mes de uso por encontrarse defectuosas. Para otro modelo la proporción fue 17 unidades devueltas de 138 vendidas.

Encontrar un IC de 90% para la diferencia de proporciones. R/]0.04473 , 0.20095[

$$\begin{aligned} p_1 &= \frac{31}{126} & \hat{p}_1 &= \frac{17}{138} & q &= 0,10 \\ q_1 &= \frac{95}{126} & \hat{q}_1 &= \frac{121}{138} & \alpha &= 0,05 \\ n_1 &= 126 & n_2 &= 138 & Z_{0,05} &= \pm 1,64985 \end{aligned}$$

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1 \hat{q}_1}{n_1} + \frac{\hat{p}_2 \hat{q}_2}{n_2}}$$

$$\frac{31}{126} - \frac{17}{138} - 1,64985 \cdot \sqrt{\frac{\frac{31}{126} \frac{95}{126}}{126} + \frac{\frac{17}{138} \frac{121}{138}}{138}} \approx 0,04473$$

$$\frac{31}{126} - \frac{17}{138} + 1,64985 \cdot \sqrt{\frac{\frac{31}{126} \frac{95}{126}}{126} + \frac{\frac{17}{138} \frac{121}{138}}{138}} \approx 0,20095$$

El IC para 90% corresponde a]0.04473 , 0.20095[

2) Interpretar el IC

Se puede asumir que las proporciones de tostadores defectuosos son iguales en ambas poblaciones?

NO se puede asumir igualdad de proporciones, puesto que el valor cero NO está en el IC.

En este caso al ser ambos extremos del IC positivos, se puede inferir que el primer grupo tiene una proporción de tostadores defectuosos mayor que el segundo grupo.

3) Tamaño de muestra

Hay 2 maneras, si hay p_1, q_1 y p_2, q_2

O si no hay se asume $p_1, q_1 = 0,25$

$$n \geq \frac{z_{\alpha/2}^2(p_1q_1 + p_2q_2)}{r^2}$$

95%

Si se desea encontrar un IC de 95% para la diferencia de proporciones de tostadores devueltos entre los dos modelos, con radio no mayor que 0.05. Determine el tamaño de la muestra. R/

Caso 1: Usando $p \neq q$

$$\hat{p}_1 = \frac{31}{126}$$

$$\hat{p}_2 = \frac{17}{138}$$

$$\hat{q}_1 = \frac{95}{126}$$

$$\hat{q}_2 = \frac{121}{138}$$

$$n_1 = 126$$

$$n_2 = 138$$

$$\alpha = 0,05 \quad \frac{\alpha}{2} = 0,025 \quad Z_{0,025} = \pm 1,95996 \quad r = 0,05$$

$$n \geq \frac{(1,95996)^2 \cdot \left(\frac{31}{126} \cdot \frac{95}{126} + \frac{17}{138} \cdot \frac{121}{138} \right)}{(0,05)^2}$$

$$n \geq 450,999 \rightarrow n \geq 451$$

Caso 2: Sin importar $p, q, p_1, q_1 = 0,25$

$$n \geq \frac{(1,95996)^2 (0,25 + 0,25)}{(0,05)^2} \quad n \geq \frac{z_{\alpha/2}^2(p_1q_1 + p_2q_2)}{r^2}$$

$$n \geq 769,2886 \rightarrow n \geq 769$$

7) Confianza utilizada

$$\text{Centro} \quad \widehat{p}_1 - \widehat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\widehat{p}_1 \widehat{q}_1}{n_1} + \frac{\widehat{p}_2 \widehat{q}_2}{n_2}}$$

Aquí igualmente es igualar el radio a \bar{E} y despejar el

ese resultado a la app con cola derecha

y el resultado de la app se hace esto

$$\frac{\alpha}{2} = \text{resultado}$$

$$\lambda = 2 \cdot \text{resultado}$$

Y al final $1 - \lambda$

2. [4 puntos] Se construyó un intervalo de confianza para la diferencia de proporciones de personas hombres (p_h) y mujeres (p_m) que prefieren una cierta marca de café, en un determinado lugar. Se utilizaron muestras de 800 hombres y 800 mujeres. En la muestra se obtuvo que 625 hombres gustan de la marca de café. Suponga que $n_1 \widehat{p}_h \geq 5$, $n_1 \widehat{q}_h \geq 5$, $n_2 \widehat{p}_m \geq 5$ y $n_2 \widehat{q}_m \geq 5$. Si el intervalo de confianza para $p_h - p_m$, obtenido de manera correcta, es $[-0.0122196, 0.07471958]$, determine cuántas mujeres en la muestra gustan de la marca de café y el nivel de confianza utilizado. $X = 600$ (Se resuelve igual al de la siguiente pagina)

$$n_1 = 800$$

$$n_2 = 800$$

$$\cancel{z_{\alpha/2}} \sqrt{\frac{\widehat{p}_1 \widehat{q}_1}{n_1} + \frac{\widehat{p}_2 \widehat{q}_2}{n_2}} = \bar{E}$$

$$p_1 = 0,78125$$

$$p_2 = 0,75$$

$$\bar{E} = \underline{5 - \alpha} = \underline{0,07471958} - \underline{-0,0122196}$$

$$q_1 = 0,21875$$

$$q_2 = 0,25$$

$$\frac{\alpha}{2} \quad \frac{\alpha}{2}$$

$$= 0,09396959$$

$$Z \frac{\alpha}{2} \cdot \sqrt{\frac{0,78125 \cdot 0,21875 + 0,75 \cdot 0,25}{800}} = 0,09396959$$

$$Z \frac{\alpha}{2} = \frac{0,09396959}{\sqrt{\frac{0,78125 \cdot 0,21875 + 0,75 \cdot 0,25}{800}}} = 2,05375$$

$$P(Z > 2,05375) = 0,02 \rightarrow \frac{\alpha}{2} = 0,02$$

$$\lambda = 0,09$$

$$1 - 0,09 = 0,96 \rightarrow \boxed{96\%}$$

▷ Caso especial I

Encontrar la cantidad que cumplen por ejemplo ^ de p_1 , & no comien^, encontrar es & x es en si encontrar el centro, para esto dan un intervalo $[a, b]$ & se usa la formula

$$p_1 - p_2 = \frac{a+b}{2}, \text{ se despeja el que falte}$$

Actualmente existe un proyecto de ley sobre ajuste tributario. En una encuesta, en la ciudad A se encuentra que 164 de 250 ciudadanos están a favor del proyecto, y en la ciudad B, que x de 240 lo apoyan. Con estos datos, se obtuvo un IC para la diferencia de proporciones de apoyo al proyecto entre las dos ciudades (la proporción en A menos la proporción en B), este es

$$]0.0709767, 0.216023[$$

Determine el valor de x . El Centro

$$\begin{aligned} n_1 &= 250 & n_2 &= 240 \\ p_1 &= \frac{164}{250} & p_2 &= \frac{x}{240} \end{aligned}$$

$$\frac{164}{250} - \frac{x}{240} = 0,0709767 + 0,216023$$

$$x = \left(-0,0709767 + 0,216023 + \frac{164}{250} \right) \cdot 240$$

$$x \approx 123$$

IC: 2 Varianzas

con $\nu_i = n_i - 1$

$$\frac{s_2^2 f_{\alpha/2, \nu_2, \nu_1}}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{s_2^2}{s_1^2 f_{\alpha/2, \nu_1, \nu_2}}$$

Para Comparar Varianza 1
Con varianza 2

$$\frac{s_2^2 f_{\alpha/2, \nu_1, \nu_2}}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{s_2^2}{s_1^2 f_{\alpha/2, \nu_2, \nu_1}}$$

Para Comparar Varianza 2
Con varianza 1

Como se cual usar?

Si el enunciado dice cual, se hace en ese orden

Y si no dice nada? Se pone el que tenga mayor S arriba

Y si lo pongo mal? No importa mientras se interprete bien el IC

Para 2 Varianzas solo se usa F

Para IC de 2 Varianzas

Se asumen varianzas iguales

Si $I \in IC$, diferentes si $I \notin IC$

Iguales $\rightarrow IC: [0.73, 1.13]$, $I \in IC$

Diferentes $\rightarrow IC: [0.50, 0.70]$, $I \notin IC$

Para Todos los IC de 2 poblaciones

Si Todo el IC $(0, 1)$, $A \geq B \Rightarrow 0.82, 30 \%$

Si Todo el IC $(0, 1)$, $A \leq B \Rightarrow -1.5, -9 \%$

Si hay $> 10 \%$, No se sube $[0, -2]$

Caso 1: Comparando σ_1^2 con σ_2^2

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 $\sigma_N = 5.8$. ~~desviación~~

~~se calcula la diferencia de variancias~~

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?

~~se calcula la diferencia de variancias~~

Comparar 1 con 2

$$n_1 = 45 \quad n_2 = 38 \quad \alpha = 0.05 \\ V_1 = 47 \quad V_2 = 37 \quad \frac{\alpha}{2} = 0.025$$

$$S_1 = 7.1 \quad S_2 = 5.8$$

$$\frac{s_1^2 f_{\alpha/2, v_2, v_1}}{s_2^2} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{s_1^2}{s_2^2 f_{\alpha/2, v_1, v_2}}$$

$$f_{0.025, 37, 47} = 0.52957$$

$$f_{0.025, 47, 37} = 0.53870$$

$$a = \frac{7.1^2 \cdot 0.52957}{5.8^2} = 0.2646$$

$$b = \frac{7.1^2}{5.8^2 \cdot 0.53870} = 0.9276$$

El IC para 95% corresponde a $[0.2646, 0.9276]$ y al ser todo positivo si se puede afirmar que el de la noche es mayor

Como I & IC se asumen varianzas diferentes

Caso 2: Comparando σ_2^2 con σ_1^2

Un profesor considera que el rendimiento promedio (nota promedio) de los estudiantes de Computación en el curso de Matemática Elemental es superior en al menos 9 puntos al rendimiento promedio de los estudiantes de otras carreras. Para analizar esto se tomó una muestra de estudiantes que cursaron el curso el año pasado, obteniendo los siguientes datos:

Estudiantes	tamaño de muestra	Rendimiento promedio observado (\bar{x})	Desviación estándar (s)
De computación:	19	78 puntos	4.3 puntos
De otras carreras:	17	65 puntos	4.7 puntos

Suponga que el rendimiento promedio en el curso de Matemática Elemental, tanto en Computación como en otras carreras, se distribuye normalmente.

- (a) Encuentre un intervalo de confianza de 90% para el cociente de las varianzas de las notas de ambos tipos de estudiantes. (5 puntos)

Como no indican cual se pone la mayor s arriba

$$n_1 = 19 \quad n_2 = 17 \quad \bar{x} = 0,70$$

$$v_1 = 18 \quad v_2 = 16 \quad \frac{s}{\bar{x}} = 0,05$$

$$s_1 = 4,3 \quad s_2 = 4,7$$

$$\frac{s_2^2 f_{\alpha/2, \nu_1, \nu_2}}{s_1^2} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{s_2^2}{s_1^2 f_{\alpha/2, \nu_2, \nu_1}}$$

$$f_{0.05, 18, 16} = 0,77953$$

$$f_{0.05, 16, 18} = 0,93447$$

$$a = \frac{4,7^2 \cdot 0,77953}{4,3^2} = 0,53107$$

$$b = \frac{4,7^2}{4,3^2 \cdot 0,93447} = 2,79978$$

MIC para 90%, corresponde a
]0,53107, 2,79978[

Como I E IC, se asumen
varianzas iguales

Ph para una Población

Como plantear bien cualquier H_0 y H_1 ?

Esto sirve para cualquier prueba de hipótesis 1 y 2 poblaciones

Las desigualdades NO estrictas $\leq, \geq, =$

van Siempre en H_1 por lo que el H_0 sería el contrario

$$H_0: \leq \rightarrow H_1: >$$

$$H_0: \geq \rightarrow H_1: <$$

$$H_0: = \rightarrow H_1: \neq$$

→ Ojo, el \leq o \geq viene implícito por que H_0 siempre usa $=$

Las desigualdades estrictas $<, >, \neq$

van Siempre en H_0 por lo que el H_1 sería el contrario

$$H_1: < \rightarrow H_0: \geq$$

$$H_1: > \rightarrow H_0: \leq$$

$$H_1: \neq \rightarrow H_0: =$$

→ Ojo, el \leq o \geq viene implícito por que H_0 siempre usa $=$

Las Ph se pueden resolver de varias maneras

- 1) Normal
- 2) Valor P
- 3) RI y RA (Regiones de aceptación y rechazo)
- 4) Otros métodos específicos

Cuando se acepta o rechaza H_0 ?

El V_{obs} es el valor estandarizado

El v_{crit} es el valor critico metido en la app

V es una distribucion χ^2, z, F, t , etc

Cola izquierda <

Si se cumple

$V_{obs} > v_c$, NO se rechaza H_0

Si se cumple

$V_{obs} < v_c$, Se rechaza H_0

Cola derecha

Si se cumple

$V_{obs} < v_c$, NO se rechaza H_0

Si se cumple

$V_{obs} > v_c$, Se rechaza H_0

2 colas

Si se cumple

$v_{c1} < V_{obs} < v_{c2}$, NO se rechaza H_0

Con valor P

Si $P > \alpha$, NO se rechaza H_0

Si $P < \alpha$, se rechaza H_0

$P = P(V > V_{obs})$ en la app con cola derecha

Ph I Proporción

$n = \text{Cantidad total}$

$\hat{P} = \frac{\text{Los que cumplen}}{\text{Cantidad total}}$

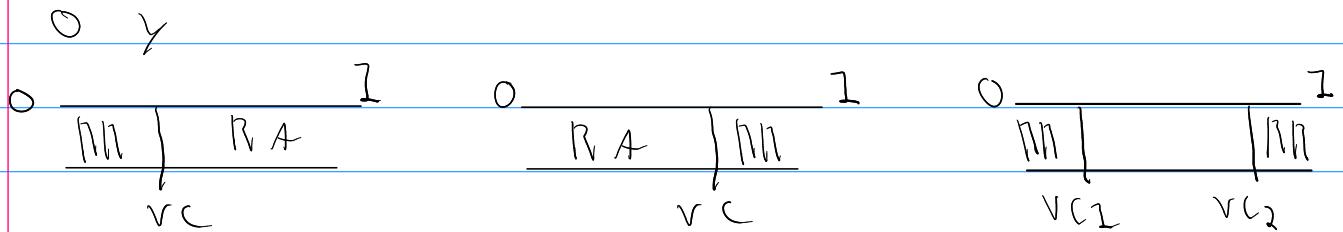
$$p_0 = H_0$$

$$q_0 = I - p_0$$

Formula de
estandarización

$$Z = \frac{\hat{P} - p_0}{\sqrt{p_0 q_0 / n}}$$

Recordar que las varianzas están entre



Cola izquierda $H_0;]0, VC[$
 $H_A;]VC, 1[$

Cola derecha $H_0;]VC, 1[$
 $H_A;]0, VC[$

Dos colas $H_0;]0, VC_1[\cup]VC_2, 1[$
 $H_A;]VC_1, VC_2[$

Ph I proporción

$n = \text{Cantidad total}$

$$p_0 = H_0$$

Formula de
estandarización

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

$$\hat{p} = \frac{\text{los que cumplen}}{\text{Cantidad total}}$$

$$q_0 = 1 - p_0$$

Ejemplo 4: una proporción

$$H_0: p \leq$$

Por estadísticas que se tienen, se ha establecido que a lo sumo 80% de los fanáticos de cierto equipo, apoyan la idea de que se vendan acciones del equipo a los fanáticos. Una muestra aleatoria de 500 aficionados, reveló que 435 de ellos están de acuerdo con la propuesta. A un nivel de significancia del 5%. ¿Tendrán razón la afirmación inicial?

Normal

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

$$H_0: p = 0,80 (\leq)$$

$$n = 500$$

$$p_0 = 0,80$$

$$H_1: p > 0,80$$

$$\hat{p} = \frac{435}{500} = 0,87$$

$$q_0 = 0,20$$

$$\alpha = 0,05$$

Estandarizando

$$z_{obs} = \frac{0,87 - 0,80}{\sqrt{0,80 \cdot 0,20}} = 3,9131$$

$$\begin{array}{c|c} 114 & | 171 \\ \hline 171 & \end{array}$$

3,9131

$$z_c = z_{0,05} = 1,67785$$

$$1,69$$

Como $z_{obs} = 3,9131 > z_c = 1,67785$
se rechaza H_0

Tip mental: Si se cumple el símbolo del mismo tipo de cola, se rechaza, por ejemplo aquí era cola derecha y se cumplió el $>$ entonces se rechaza

Con valor P

$$H_0: p = 0,80 (\leq) \quad n = 500 \quad p_0 = 0,80$$
$$H_1: p > 0,80 \quad \hat{p} = \frac{435}{500} = 0,87 \quad q_0 = 0,20$$
$$\alpha = 0,05$$

Estandarizando

$$Z_{655} = \frac{0,87 - 0,80}{\sqrt{\frac{0,80 \cdot 0,20}{500}}} = 3,9131$$

$$P(Z > 3,9131) = 0,00005$$

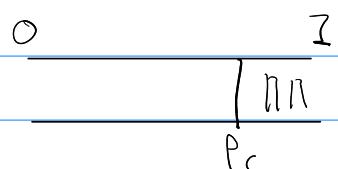
II/ Como $0,00005 < 0,05$
se rechaza H_0

Con RR y RA en términos de p

$$p_0 \pm z_{\alpha/2} \sqrt{\frac{p_0 q_0}{n}} \quad p_0 = 0,80$$
$$H_0: p = 0,80 (\leq) \quad n = 500 \quad q_0 = 0,20$$
$$H_1: p > 0,80 \quad \hat{p} = \frac{435}{500} = 0,87 \quad \alpha = 0,05$$
$$Z_{0,05} = 1,64985$$

$$b = 0,8 + 1,64985 \cdot \sqrt{\frac{0,80 \cdot 0,20}{500}} = 0,8294 \leftarrow P_c$$

RR: $[0,8294, 1]$
RA: $[0,8294]$



II/ Como $\hat{p} = 0,87 \in RR$, se rechaza

$H_0: p \leq$
Tamaño de muestra $\leq 50\%$

4. Un profesor de la Universidad Futuro Garantizado cree que más de la mitad de los estudiantes matriculados se retiran de al menos una materia. En una muestra de 80 estudiantes, se observó que 45 se habían retirado de al menos una materia.

- (c) ¿De qué tamaño debe ser una muestra para que la prueba tenga una significancia de 2% y una potencia de 90% cuando el verdadero porcentaje de estudiantes que se retiran de al menos una materia es del 60%?

R/ ~~273~~ 279

$$H_0: p = 0.5 (\leq) \quad p_0 = 0.5 \quad p_1 = 0.60 \quad n \geq \frac{(|z_{\alpha/2}| \sqrt{p_0 q_0} + |z_{\beta}| \sqrt{p_1 q_1})^2}{(p_1 - p_0)^2}$$

$$H_1: p > 0.5 \quad q_0 = 0.5 \quad q_1 = 0.40 \quad \text{donde } k \text{ es el número de colas}$$

$$H_1: p = 0.60 \quad \alpha = 0.02 \quad \beta = 0.10$$

$$z_{0.02} = 2.05375 \quad z_{0.10} = 1.28155$$

$$n \geq \frac{(2.05375 \cdot \sqrt{0.5 \cdot 0.5} + 1.28155 \cdot \sqrt{0.60 \cdot 0.40})^2}{(0.60 - 0.5)^2}$$

$$n \geq 273.80 \rightarrow n \geq 274$$

Error tipo I (Básicamente valor p)

Ejemplo 101. Se toma una muestra aleatoria de 300 habitantes de cierta ciudad y se les pregunta si están a favor de que el país C realice un tratado de libre comercio con China. Al realizar la prueba de hipótesis se ha determinado que si menos de 100 personas responden de manera afirmativa, se concluye que a lo sumo el 30 % de los habitantes está a favor del tratado.

$H_0: p \leq$

2. Encuentre la probabilidad del error tipo I

$$H_0: p = 0.30 (\leq) \quad \hat{P} = \frac{2}{3} \quad n = 300 \quad Z = \frac{\hat{P} - p_0}{\sqrt{p_0 q_0 / n}}$$

$$H_1: p > 0.30 (\hat{P} > 0.30) \quad p_0 = 0.30 \quad q_0 = 0.70$$

$$\alpha = 0.05 \quad z_{0.05} = 1.67985$$

$$z_{0.05} = \frac{0.33 - 0.30}{\sqrt{0.30 \cdot 0.70 / 300}} = 1.25988$$

$$P(Z > 1.25988) = 0.10386$$

Error tipo 2

Reemplazar el p_0 y q_0 por el \hat{p}_1 y q_1

$H_1: p = H_1'$ $Z = \frac{\hat{p}_1 - p_0}{\sqrt{pq/n}}$ y eso se mete en la app con el símbolo contrario

3. ¿Cuál es la probabilidad del error tipo II si solo el 40 % de los habitantes está a favor del tratado?

$$H_0: p = 0,30 (\leq)$$

$$\hat{p} = \frac{1}{3}$$

$$n = 360$$

$$H_1: p > 0,30 (\hat{p} > 0,30)$$

$$p_0 = 0,30 \quad q_0 = 0,70$$

$$\alpha = 0,05 \quad Z_{0,05} = 1,67985$$

$$H_2: p = 0,40$$

$$\hat{p}_C = \frac{1}{3} \quad p_C = 0,70 \quad q_C = 0,60$$

Estandarizar para error 2

$$Z = \frac{\hat{p}_C - p_0}{\sqrt{pq/n}}$$

$$Z = \frac{\frac{1}{3} - 0,70}{\sqrt{\frac{0,70 \cdot 0,60}{360}}} = -2,35702$$

$$P(Z < -2,35702) = 0,009137$$

Contrario a

$$H_1: p > 0,30$$

Tamaño de muestra

4. Una casa encuestadora desea realizar la prueba con un nivel de significancia del 5 % y una potencia del 2 % si la verdadera proporción es del 40 %. ¿Qué tamaño de muestra debe tomar?

$$H_0: p = 0,30$$

$$\hat{p} = \frac{2}{3}$$

$$n = 300$$

$$H_1: p > 0,30$$

$$p_0 = 0,30 \quad q_0 = 0,70$$

$$\beta I = 0,90 \quad \alpha I = 0,60$$

$$H_2: p = 0,40$$

$$\alpha = 0,05$$

$$\beta = 0,02$$

$$n \geq \frac{(|z_{\alpha/k}| \sqrt{p_0 q_0} + |z_{\beta}| \sqrt{p_1 q_1})^2}{(p_1 - p_0)^2}$$

donde k es el número de colas

$$z_{0,05} = 1,675$$

$$z_{0,02} = 2,05$$

$$n \geq \frac{(1,675 \cdot \sqrt{0,30 \cdot 0,70} + 2,05 \cdot \sqrt{0,70 \cdot 0,60})^2}{(0,40 - 0,30)^2}$$

$$n \geq 309,1 \rightarrow \boxed{n \geq 310}$$

Ph: I Variancia

Esto se puede resolver de 7 maneras

Usando χ^2

Usando s^2

Usando s

Usando valor P , lo mismo de siempre

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \text{ con } \nu = n-1$$

Solo es despejar la formula

$$\chi^2_{\text{obs}} = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

$$s_c^2 = \frac{\chi^2_c \cdot \sigma_0^2}{n-1}$$

$$s_c = \sqrt{\frac{\chi^2_c \cdot \sigma_0^2}{n-1}}$$

$$\chi^2_c = \chi^2_{\alpha, \nu}$$

$$s_{\text{obs}}^2 = s^2$$

$$s_{\text{obs}} = s$$

Aqui es

al revés el

s_c^2 es el

estandarizado y

el $s_{\text{obs}}^2 = s^2$

Aqui es

al revés el

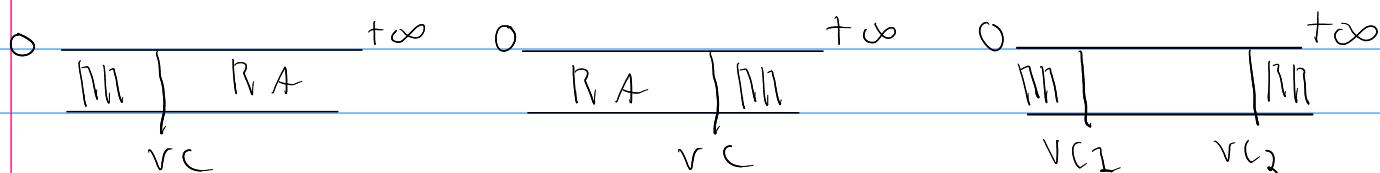
s_c es el

estandarizado y

el $s_{\text{obs}} = s$

(ual usar? el enunciado dice

Recordar que las varianzas están entre
0 y $+\infty$



Cola izquierda

RA:]0, VC [RA:]VC, $+\infty$ [

Cola derecha

RA:]VC, $+\infty$ [RA:]0, VC [

Dos colas

RA:]0, VC₁ [∪]VC₂, $+\infty$ [

RA:]VC₁, VC₂ [

Ejemplo 8: una varianza

ES con s_c pero
hagamos las 7
maneras para
aprender

Se estima que el tiempo que requieren los estudiantes de estadística para resolver el primer parcial se distribuye normalmente. Si este año una muestra de 20 estudiantes dio una desviación estándar de 5 min y considerando una significancia del 10%, ¿hay evidencia de que en realidad la desviación estándar de las notas es menor que los 8 min?

$$\text{Usando } \chi^2_C \quad n = 20 \quad s_0 = 8 \quad \chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \text{ con } \nu = n-1$$

$$H_0: \sigma = 8 \quad (\leq) \quad H_A: \sigma < 8$$

$$H_1: \sigma < 8$$

$$\chi^2_{0.95} = \frac{19 \cdot 5^2}{8^2} = 7.7278$$

$$\chi^2_C = \chi^2_{0.10, 19} = 11.65091$$

[R] Como $\chi^2_{0.95} = 7.7278 < \chi^2_C = 11.65091$
Se rechaza H_0

$$\text{Usando } s^2_C \quad \text{Aqui el } s^2_{0.95} = s^2$$

$$\chi^2_C = \frac{(n-1) \cdot s^2}{s_0^2} \quad \text{Aqui el } s^2_{0.95} = s^2$$

$$n = 20 \quad s_0 = 8 \quad \nu = 19 \quad s = 5 \quad \alpha = 0.10$$

$$s^2 = \frac{\chi^2_C \cdot s_0^2}{n-1} \quad \chi^2_C = \chi^2_{0.10, 19} = 11.65091$$

$$s^2_C = \frac{11.65091 \cdot 8^2}{19} = 39.2952 \quad s^2_{0.95} = s^2 = 25$$

[R] Como $s^2_{0.95} = 25 < s^2_C = 39.2952$
Se rechaza H_0

Usando S_C Aquí el $S_{0.05} = S$

$$\chi^2_C = \frac{(n-1) \cdot s^2}{\sigma_0^2} \quad n=20 \quad \sigma_0 = 8$$
$$v=19 \quad s=5 \quad \alpha=0.10$$

$$S_C = \frac{\chi^2_C \cdot \sigma_0^2}{n-1} \quad \chi^2_C = \chi^2_{0.10, 19} = 11.65091$$

$$S_C = \sqrt{\frac{\chi^2_C \cdot \sigma_0^2}{n-1}}$$

$$S_C = \sqrt{11.65091 \cdot 8^2} = 6.2696 \quad S_{0.05} = 5$$

R/ Como $S_{0.05} = S < S_C = 6.2696$
Se rechaza H_0

Usando valor P $n=20 \quad \sigma_0 = 8 \quad \chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ con $v=n-1$
 $v=19 \quad s=5 \quad \alpha=0.10$

$$\chi^2 = \frac{19 \cdot 5^2}{8^2} = 7.423875$$

$$P(\chi^2 < 7.423875, 19) = 0.00890$$

R/ Como $P=0.00890 < 0.10$
Se rechaza H_0

Error tipo 2

Normalmente hay que buscar el s^2 despejando con los valores normales y luego usar la $\sigma_1^2 = h_1^{-1}$ en vez de σ_0^2

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \text{ con } \nu = n-1$$

La empresa de llantas Mundiales ha sacado un nuevo modelo de llantas para automóviles al mercado asegurando que su duración tiene una desviación estándar menor a 5000 km. Suponga que la duración de las llantas se distribuye normalmente. Se quiere contrastar la afirmación a un nivel de significancia del 10% con una muestra aleatoria de 25 llantas del nuevo modelo.

Acote la probabilidad del error tipo II de la prueba si duración de las llantas del nuevo modelo tiene una desviación estándar de 4700 km

$$H_0: \sigma = 5000 \quad (\geq) \quad n=25 \quad \sigma_0 = 5000$$

$$H_1: \sigma < 5000 \quad V=24 \quad S=?$$

$$\lambda = 0,10 \quad \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \text{ con } \nu = n-1$$

$$H_1: \sigma = 4700$$

Primero se debe encontrar S^2 con los valores normales

$$\chi^2_C = \frac{(h-1) - S^2}{\sigma_0^2}$$

$$\chi^2_C = \chi^2_{0,010,24} = 15,65868$$

$$S = \frac{\chi^2_C \cdot \sigma_0^2}{h-1}$$

$$\sigma_0^2 = 5000^2$$

$$h-1 = 24$$

$$S^2 = \frac{15,65868 \cdot 5000^2}{24} \rightarrow S^2 = 16311125$$

Ahora se debe estandarizar con los nuevos datos en este caso con χ^2 y luego ese resultado en la app

$$\chi^2 = \frac{24 \cdot 16311125}{4700^2} \quad n=25 \quad \sigma_0 = 4700$$

$$V=24 \quad S=16311125$$

Contrario de <

$$\chi^2 = 17,72 \rightarrow P(\chi^2 > 17,72, 24)$$

$$\rightarrow 0,816307$$

RR v RA

2. Un ingeniero de computación, se ha quejado ante la Oficina de Trabajo por las grandes diferencias salariales que se le paga a un ingeniero en computación de un empresa a otra. Ante esto, el presidente de la Cámara de Empresas ha señalado que la desviación estandar del salario mensual de un ingeniero en computación es menor a 100 dólares. Para analizar esta afirmación, la Oficina de Trabajo en una muestra de 40 ingenieros observó un salario promedio de 2000 dólares con una desviación estandar de 90 dólares. Se quiere contrastar la afirmación a un nivel de significancia del 10%.

(a) Determine las regiones de aceptación y rechazo de la prueba para el estadístico S_c^2 .

$$H_0: \sigma = 100 (\geq) \quad n = 40 \quad s = 90 \quad \chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \text{ con } \nu = n - 1$$

$$H_1: \sigma < 100 \quad V = 39 \quad \sigma_0 = 100$$

$$\alpha = 0.10$$

$$S_c^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

$$S_c^2 = \chi^2_{0.10, 39} = 28,795.79$$

$$\sigma^2 = 100^2$$

$$S_c^2 = \frac{\chi^2_{0.10} \cdot \sigma^2}{n-1}$$

$$n-1 = 39$$

$$\text{Aqui: } S_{obs}^2 = s^2$$

$$S_c^2 = 28,795.79 \cdot 100^2 = 7229,69 \leftarrow$$

$$39 \quad 0 \quad \frac{100}{S_c^2} \quad +\infty$$

RR:]0, 7229.69[

RA:]7229.69, +\infty[

R / Como $S_{obs}^2 = 90^2 = 8100 \in RA$
 NO se rechaza H_0

Notese como piden RR y RA con s^2 o s , pero no con χ^2 por lo de que con s^2 o s estandarizados son el valor critico pero el χ^2 es el observado

Tamaño mínimo de s^2 o s

Para una determinada población se afirma que $\sigma^2 \geq 24$. Un estudiante de computación realizó correctamente el contraste de hipótesis tomando una muestra de tamaño 40 al nivel de significancia de 0.05, obteniendo que no existe evidencia en contra de la afirmación. ¿Qué tan pequeño puede ser la desviación estándar de los datos de la muestra?

$$H_0: \sigma^2 = 24 (\geq)$$

$$n = 40 \quad \alpha = 0.05$$

$$H_1: \sigma^2 < 24$$

$$gl = 39 \quad \sigma_0^2 = 24$$

$$\chi^2_{0.05, 39} = 25,69539$$

↳ Esto está en términos de χ^2 pero la pregunta es de σ

Recordando

$$\chi^2_C = \frac{(n-1) \cdot s^2}{\sigma_0^2}$$

$$\chi^2_C \cdot \sigma_0^2 = (n-1) \cdot s^2$$

$$s^2 = \frac{\chi^2_C \cdot \sigma_0^2}{n-1}$$

$$s = \sqrt{\frac{\chi^2_C \cdot \sigma_0^2}{n-1}} \quad n = 40 \quad \alpha = 0.05$$
$$gl = 39 \quad \sigma_0^2 = 24$$

$$s = \sqrt{\frac{25,69539 \cdot 24}{39}}$$

$$s = 3,9765$$

Despejar s , si fuera s^2 , no aplicar la raíz.

Ph para 2 poblaciones

Aqui sigue aplicando todo lo de 1 cohaciòn

Como plantear bien cualquier H_0 y H_1 ?

Esto sirve para cualquier prueba de hipótesis 1 y 2 poblaciones

Las desigualdades NO estrictas $\leq, \geq, =$

Van Siempre en H_0 por lo que el H_1 Seria el contrario

$$H_0: \leq \rightarrow H_1: >$$

$$H_0: \geq \rightarrow H_1: <$$

$$H_0: = \rightarrow H_1: \neq$$

↳ Ojo, el \leq o \geq viene implícito por que H_0 siempre usa $=$

Las desigualdades estrictas $<, >, \neq$

Van Siempre en H_1 por lo que el H_0

Seria el contrario

$$H_1: < \rightarrow H_0: \geq$$

$$H_1: > \rightarrow H_0: \leq$$

$$H_1: \neq \rightarrow H_0: =$$

↳ Ojo, el \leq o \geq viene implícito por que H_0 siempre usa $=$

Las Ph se pueden resolver de varias maneras

① Normal

② Valor P

③ RRI y RA (Regiones de aceptación y rechazo)

④ Otros métodos específicos

¿Cuando se acepta o rechaza H_0 ?

El V_{obs} es el valor estandarizado

El V_{crit} es el valor critico metido en la app

V es una distribucion χ^2, Z, F, t , etc

Cola izquierda <

Si se cumple

$V_{obs} > V_c$, NO se rechaza H_0

Si se cumple

$V_{obs} < V_c$, Se rechaza H_0

Cola derecha

Si se cumple

$V_{obs} < V_c$, NO se rechaza H_0

Si se cumple

$V_{obs} > V_c$, Se rechaza H_0

2 colas

Si se cumple

$V_{c1} < V_{obs} < V_{c2}$, NO se rechaza H_0

Con valor P

Si $P > \alpha$, NO se rechaza H_0

Si $P < \alpha$, se rechaza H_0

$P = P(V > V_{obs})$ en la app con cola derecha

Siempre que diga "diferencia de \neq medias,
proporciones o varianzas, es de 2 colas
entonces $H_0: \mu_1 = \mu_2$ y $H_1: \neq$

Ph para 2 medias

Existen 3 casos y 3 formulas
diferentes (Iguales a los de IC para 2 medias)

Sobre el δ_0 (Viene en las 3 formulas)

Si es de 2 colas, $\delta_0 = 0$

Si es de 1 cola, $\delta_0 = H_0$ (Cuece que H_0 sea 0)

Si NO es de colas, el H_0 suele plantearse
como $H_0: \mu_1 - \mu_2 = (\leq)$, como una resta
 $H_0: \mu_1 - \mu_2 = (\geq)$

Caso 1: $n_1, n_2 > 30$ y se conocen las desviaciones estándares poblacionales σ_1 y σ_2 , mejor dicho. Si no dicen si son iguales o diferentes, solo este caso tiene tamaño de muestra.

Ejemplo 5: diferencia de medias

Un gerente aplicó el mismo test de capacitación a 2 grupos. El primero de 50 empleados obtuvo una media de 65 pts con una desviación de 10 pts. El segundo grupo de 40 empleados arrojó una media de 62 pts con una desviación estándar de 8 pts.

¿Existe diferencia significativa entre las medias de los dos grupos a un nivel de significancia del 5%? $H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

Siempre que diga "diferencia de" ya sea de media, proporción o varianzas, es de 2 colas $\rightarrow H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$

$H_0: \mu_1 = \mu_2 \rightarrow$ No hay diferencia

$H_1: \mu_1 \neq \mu_2 \rightarrow$ Hay diferencia

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

Estandarizando

$$Z_{\text{obs}} = \frac{65 - 62 - 0}{\sqrt{\frac{10^2}{50} + \frac{8^2}{40}}} \approx 1,5811$$

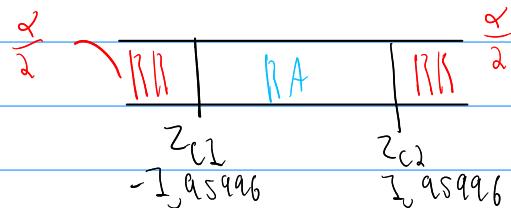
$$n_1 = 50 \quad n_2 = 40 \quad \alpha = 0,05$$

$$\bar{x}_1 = 65 \quad \bar{x}_2 = 62 \quad \frac{\alpha}{2} = 0,025$$

$$\sigma_1 = 10 \quad \sigma_2 = 8 \quad d_0 = 0$$

$$Z_C = 0,025 = \pm 1,95996$$

Como es de 2 colas, hace falta ver si el Z_C está en RECHAZO



Como

$$Z_{C1} < Z_{\text{obs}} < Z_{C2}$$

$$-1,95996 < 1,5811 < 1,95996 \quad \text{No se rechaza } H_0$$

por lo que se asume ambos tienen mismo promedio

Enfoque de valor P

Como es de 2 colas

Valor $P = 2 \cdot P(Z > Z_{\text{obs}})$ en la app

Estandarizando

$$Z_C = \frac{65 - 62}{\sqrt{\frac{10^2}{50} + \frac{8^2}{40}}} \approx 1,5811$$

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

↙ Valor P

$$2 \cdot P(Z > 1,5811) = 0,11386$$

[Como valor $P = 0,11386 > 0,05$, no se rechaza H_0]

Recordar que con valor P indiferentemente del tipo de cola, siempre se hace la comparación
Valor $P \geq \alpha$, No se rechaza
Valor $P \leq \alpha$, Se rechaza

Caso 2: $n_1 < 30$ y $n_2 > 30$ y no se conocen las desviaciones estándar poblacionales σ_1 y σ_2 , mejor dicho si se dice que las varianzas poblacionales son iguales y con ver que el n_2 o el n_1 sea < 30

Ejemplo 106. Se obtuvieron las estaturas de 20 mujeres y 30 hombres tomados aleatoriamente de la población de alumnos de cierta escuela. En la siguiente tabla se resumen los datos encontrados

	Número	Promedio	Desviación
Hombres (m):	30	69.8	1.92
Mujeres (f):	20	63.8	2.18

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

$$\text{con } S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$\text{y } \nu = n_1 + n_2 - 2$$

Suponga que las estaturas de las mujeres y las estaturas de los hombres se distribuyen normalmente y tienen desviaciones estándar similares.

¿Puede concluirse, con un nivel de significancia del 4 %, que el promedio de estaturas de los hombres de la escuela supera en más de 3 cm el promedio de las mujeres?

Normal

$$H_0: \mu_1 - \mu_2 \leq 3 \quad h_1 = 30 \quad h_2 = 20 \quad d_0 = 3$$

$$H_1: \mu_1 - \mu_2 > 3 \quad \bar{X}_1 = 69.8 \quad \bar{X}_2 = 63.8 \quad V = 78$$

$$S_1 = 1.92 \quad S_2 = 2.18 \quad \alpha = 0.04$$

$$S_p^2 = \frac{(30-1) \cdot 1.92^2 + (20-1) \cdot 2.18^2}{78} = 7.108$$

$$T_{obs} = \frac{69.8 - 63.8 - 3}{\sqrt{\frac{7.108}{30} + \frac{7.108}{20}}} = 5.127$$

$$T_C = +_{0.04} 78 = 1.788$$

Como $Z_{0.04} = 5.127 > 1.788$
Se rechaza H_0

Valor p

$$P(T > 5.127) = 0.00000$$

Como $p = 0.00000 < 0.04$
Se rechaza H_0

Caso 3: $H_1: \mu_1 < \mu_2$ y $H_0: \mu_1 \geq \mu_2$ y no se conocen las desviaciones estandar poblacionales σ_1 y σ_2 , mejor dicho si se dice que las varianzas poblacionales son diferentes y con ver que el n_1 o el n_2 sea < 30

Un investigador desea comparar dos modelos de baterías de teléfonos celulares para determinar cuál de ellos tiene mayor duración de carga. Como desconoce la media y la varianza poblacional de los dos modelos, decide tomar una muestra de 26 baterías del modelo X y de 21 del modelo Z. En ambos casos, el tiempo de duración de la carga se distribuyó de forma normal, con una media muestral de 26 horas y una desviación muestral de 2,3 horas para el modelo X y con una media muestral de 24 horas y una desviación muestral de 1,7 horas para el modelo Z.

- b) [4 puntos] Realice una prueba de hipótesis con un nivel de significancia del 4% para determinar si la diferencia entre el tiempo de duración de carga de ambos modelos de batería es menor a 3 horas. (Suponga que las varianzas son distintas).

Normal

$$H_0: \mu_1 - \mu_2 = 3 (\geq) \quad H_1: \mu_1 - \mu_2 < 3 \quad \bar{X}_1 = 26 \quad \bar{X}_2 = 24 \quad d_0 = 3$$

$$S_1 = 2,3 \quad S_2 = 1,7 \quad v = 49,69$$

$$T = \frac{\bar{X}_1 - \bar{X}_2 - d_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$\text{con } v = \frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}$$

$$T_{0.04} = \frac{26 - 24 - 3}{\sqrt{\frac{2,3^2}{26} + \frac{1,7^2}{21}}} = -1,712267$$

$$v = \frac{\left(\frac{(2,3)^2}{26} + \frac{(1,7)^2}{21} \right)^2}{\frac{\left((2,3)^2 \right)^2}{26-1} + \frac{\left((1,7)^2 \right)^2}{21-1}} = 49,6930$$

$$t_{0.04, 49,6930} = -1,79147$$

|| Como $t_{0.04} = -1,712267 < t_c = -1,79147$
No se rechaza H_0 .

Ph 2 Proporciones

$$n_i = \text{Cantidad total} \quad \hat{p}_i = \frac{x_i}{n_i}$$

$x_i = \text{los que cumplen}$

9. De una muestra de 100 acciones de la bolsa de valores A, 32 tuvieron ganancia el martes pasado. Una muestra de 100 acciones de la bolsa de valores B indica que 27 obtuvieron ganancia ese mismo día. Ante estos resultados, una persona afirma que hay una mayor proporción de acciones que obtuvieron ganancia en la bolsa A, con respecto a la otra bolsa, el martes pasado. Utilice $\alpha = 0.04$.

- (a) ¿Hay evidencia que respalde la afirmación de esta persona? R/ Valor $p = 0.223627$. No hay evidencia a favor de la afirmación

$$\begin{aligned} H_0: p_A &= p_B (\leq) & n_1 &= 100 & n_2 &= 100 & Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\hat{P}\hat{Q}}{n_1} + \frac{\hat{P}\hat{Q}}{n_2}}} \text{ con } \hat{P} = \frac{X_1 + X_2}{n_1 + n_2} \\ H_1: p_A &> p_B & p_1 &= 0,32 & p_2 &= 0,27 & \\ & & x_1 &= 32 & x_2 &= 27 & \alpha = 0,04 \end{aligned}$$

$$\begin{aligned} \hat{p} &= \frac{32+27}{100+100} = 0,295 & \hat{q} &= 0,705 & Z &= \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\hat{P}\hat{Q}}{n_1} + \frac{\hat{P}\hat{Q}}{n_2}}} \text{ con } \hat{P} = \frac{X_1 + X_2}{n_1 + n_2} \end{aligned}$$

$$Z_{0.05} = \frac{0,32 - 0,27}{\sqrt{\frac{0,295 \cdot 0,705}{100} + \frac{0,295 \cdot 0,705}{100}}} = 0,7752$$

$$Z_c = Z_{0.09} = 1,75069$$

R/ Como $Z_{0.05} = 0,7752 < Z_c = 1,75069$
No se rechaza H_0

$$Z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\frac{\hat{P}\hat{Q}}{n_1} + \frac{\hat{P}\hat{Q}}{n_2}}} \text{ con } \hat{P} = \frac{X_1 + X_2}{n_1 + n_2}$$

Con valor P

$$\alpha = 0.09$$

$$Z_{0.05} = \frac{0.32 - 0.27}{\sqrt{\frac{0.295 \cdot 0.705}{100} + \frac{0.295 \cdot 0.705}{150}}} = 0.7752$$

$$P(Z > 0.7752) = 0.22$$

R/ Como $P = 0.22 < 0.09$
No se rechaza H_0

Tamaño de muestra

- (b) ¿De qué tamaño deberán haberse tomado las muestras si se desea hacer el contraste de hipótesis con un nivel de significancia de $\alpha = 0.04$ y una potencia de 80% cuando el 30% de las personas obtuvo ganancia de la bolsa A y el 25% en B?

$$R/ n \geq 1068.85 \rightarrow 1071$$

$$\begin{array}{lll} p'_1 = 0.30 & p'_2 = 0.25 & \alpha = 0.09 \\ q'_1 = 0.70 & q'_2 = 0.75 & \beta = 0.20 \end{array}$$

$$n \geq \frac{(|z_{\alpha/k}| \sqrt{\frac{1}{2}(p'_1 + p'_2)(q'_1 + q'_2)} + |z_{\beta}| \sqrt{p'_1 q'_1 + p'_2 q'_2})^2}{(p'_1 - p'_2)^2}$$

donde k es el número de colas Cola derecha

$$Z_{0.04} = 1.75069 \quad Z_{0.20} = 0.89162$$

$$n \geq \frac{(1.75069 \cdot \sqrt{\frac{1}{2} \cdot (0.30 + 0.25) \cdot (0.70 + 0.75)} + 0.89162 \cdot \sqrt{0.30 \cdot 0.70 + 0.25 \cdot 0.75})^2}{(0.30 - 0.25)^2}$$

$$n \geq 1070.75 \rightarrow n \geq 1071$$

```

import math

# Datos
p1 = 0.30
p2 = 0.25
z_alpha = 1.75069    # z0.04
z_beta = 0.84162   # z0.20

# Proporción promedio
p_bar = (p1 + p2) / 2

# Numerador
num = (
    z_alpha * math.sqrt(2 * p_bar)
    * (1 - p_bar) +
    z_beta * math.sqrt(p1 * (1 -
    p1) + p2 * (1 - p2))
) ** 2

# Denominador
den = (p1 - p2) ** 2

# Tamaño de muestra
n = num / den

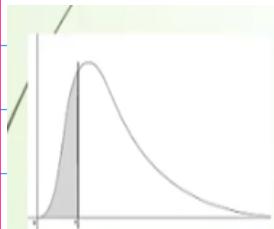
print(n) #1070.7598979728687

```

Ph 2 Varianzas

Se usa la distribución F

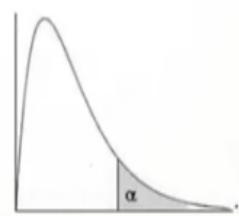
Izquierda



$$f_c = f_{\alpha, n_1-1, n_2-1}$$

$$P(F < f_{obs})$$

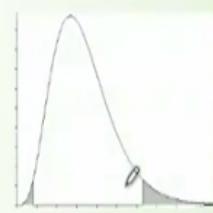
Derecha



$$f_c = f_{1-\alpha, n_1-1, n_2-1}$$

$$\text{Valor } P \text{ es } P(F > f_{obs}).$$

Dos colas



$$f_{c1} = f_{\alpha/2, n_1-1, n_2-1}, f_{c2} = f_{1-\alpha/2, n_1-1, n_2-1}$$

$$\begin{array}{ll} f_{obs} > 1 & 2P(F > f_{obs}) \\ f_{obs} < 1 & 2P(F < f_{obs}) \end{array}$$

Aquí si

$$P(X < x)$$

Aquí si

$$P(X > x)$$

Realmente se puede
meter siempre
un C en la app

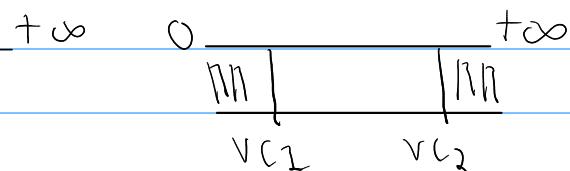
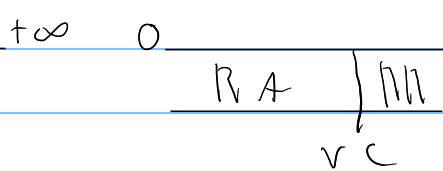
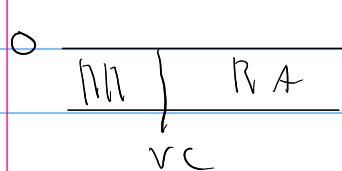
$$r_0 = H_0$$

$$r_0 = H_0$$

$$r_0 = 1$$

Recordar que las varianzas están entre

$$0 \leq +\infty$$



Cola izquierda

$$RA:]0, vc[\cup [RA; +\infty[$$

Cola derecha

$$RA:]vc, +\infty[\cup [RA; 0, vc[$$

Dos colas

$$RA:]0, vc_1[\cup]vc_2, +\infty[$$

$$RA:]vc_1, vc_2[$$

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.

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 $\sigma_m^2 = \sigma_h^2$.

$\sigma_m^2 = 2$ Colas

$$r_0 = 1$$

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

$$H_0: \frac{\sigma_m^2}{\sigma_h^2} = 1$$

$$n_1 = 21 \quad n_2 = 31$$

$$v_1 = 20 \quad v_2 = 30$$

$$S_1 = 2,18 \quad S_2 = 1,92 \quad \alpha = 0,01$$

$$F = \frac{S_1^2 / S_2^2}{\frac{v_1 - 1}{v_2 - 1}} \text{ con } v_i = n_i - 1$$

$$H_1: \frac{\sigma_m^2}{\sigma_h^2} \neq 1$$

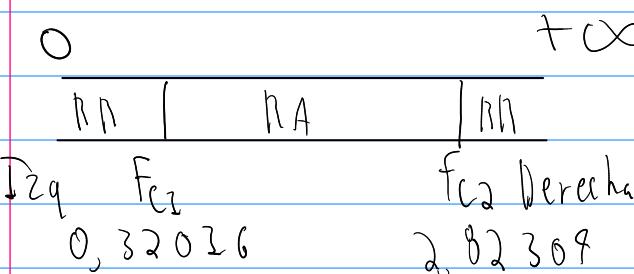
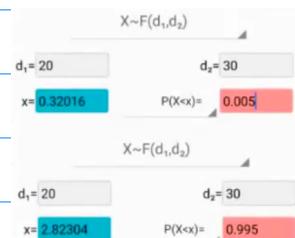
$$\frac{\alpha}{2} = 0,005$$

$$1 - \frac{\alpha}{2} = 0,995$$

$$F_{0,05} = \frac{2,18^2}{1,92^2} = 1,2892$$

$$f_{C1} = f_{0,005, 20, 30} = 0,32016$$

$$f_{C2} = f_{0,995, 20, 30} = 2,82309$$



$$RA:]0,032016; +\infty[$$

$$RA:]0,32016, 2,82309[$$

RA como

$$f_{C1} = 0,32016 < F_{0,05} = 1,2892 < f_{C2} = 2,82309$$

NO se rechaza H_0 .

con valor P

$$F = \frac{S_1^2/S_2^2}{\nu_1} \text{ con } \nu_i = n_i - 1$$

$$F_{0.95} = \frac{2.78^2}{1.92^2} = 1.2892 \quad \alpha = 0.05$$

$$v_1 = 20 \quad v_2 = 30$$

$$2P(F > 1.2892)$$

$$= 2 \cdot 0.25858$$

$$= 0.51716$$

Se mete con α en la app
y se multiplica por 2

manualmente

[V] Como $P = 0.51716 > 0.01$
NO se rechaza H_0 .

Bondad de ajuste

Lo que se busca probar es si los datos se comportan o no de cierta manera

Todos los de χ^2 son de cola derecha, en si la formula del χ^2_c de cola derecha

"use 0,05 de significancia" $\rightarrow \chi^2_{0,05, v}$ con v en la app

n = cantidad de muestra

k = cantidad de eventos $v = k - 1$

O_i = Valores observados P_i = Probabilidad

El χ^2_c se mete con \leq
y el valor P con $>$

e_i = Valores esperados $n \cdot P_i$ $\chi^2_{\text{obs}} = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i}$ con $v = k - 1$

$\chi^2_c = \chi^2$ significancia dada, v

Si un valor e_i da < 5 , ese e_i y el o_i deben combinarse con el anterior y ademas se debe actualizar el k y el v

Ejemplo: $O_1 = 7, O_2 = 3, O_3 = 6$
 $e_1 = 7, e_2 = 6, \underline{e_3 = 9} < 5$

$$e_2 = 9 \rightarrow e_2 = 6 + 9 = 10 \quad k = 3 \rightarrow k = 2 \\ O_2 = 3 \rightarrow O_2 = 3 + 6 = 9 \quad v = 2 \rightarrow v = 1$$

Hay 2 casos, dan o no distribucion

Caso 1: Dan una distribución

5. [5 puntos] Los siguientes datos se refieren a la cantidad de accidentes laborales menores por semana en una empresa de manufactura registrados durante 36 semanas.

	Accidentes	1	2	3	4	5	$\rightarrow n = 36$
	Semanas	10	10	7	5	4	$V = 9$

Si se define X como la frecuencia de accidentes por semana, ¿existe evidencia en contra de suponer que $X \sim P(2)$ con significancia 0.03?

$$\alpha = 0.03$$

Recuerde $P(2)$ se refiere a una distribución Poisson con media igual a 2 y

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$$

Paso 1: Plantear H_0 y H_1 que casi siempre son decir que se comportan y no se comportan de la manera esperada respectivamente

$$H_0: X \sim P(2)$$

$$H_1: X \neq P(2)$$

Paso 2: Identificar n, k, V, α

$$n = 36 \quad k = 5 \quad \lambda = 2$$

$$\alpha = 0.03 \quad V = 9$$

Paso 3: Calcular χ^2 ; o sea χ^2_{cal} , donde χ^2_{cal} es el resultado de remplazar los valores esperados en la distribución dada

Accidentes	1	2	3	4	5
------------	---	---	---	---	---

$$\chi^2_1: 36, \frac{e^{-2} \cdot 2^2}{1!} = 9.74 \geq S \checkmark$$

$$\chi^2_2: 36, \frac{e^{-2} \cdot 2^2}{2!} = 9.74 \geq S \checkmark$$

$$\chi^2_3: 36, \frac{e^{-2} \cdot 2^3}{3!} = 6.99 \geq S \checkmark$$

$$P_3: 36, \frac{e^{-2} \cdot 2^3}{3!} = 6,99 \geq 5 \checkmark$$

$$P_4: 36, \frac{e^{-2} \cdot 2^4}{4!} = 3,29 \text{ } \checkmark \quad \left. \begin{array}{l} \text{Como estos son } \leq 5 \\ \text{Hay que combinar} \end{array} \right\}$$

$$P_5: 36, \frac{e^{-2} \cdot 2^5}{5!} = 1,29 \text{ } \checkmark$$

Paso 4: Actualizar si es necesario

$$e_3 = 6,99 \rightarrow e_3 = 6,99 + 3,29 + 1,29 = 11,02$$

$$k=5 \rightarrow k=3$$

$$V=4 \rightarrow V=2$$

$$O_3 = 7 \rightarrow O_3 = 7 + 5 + 7 = 16$$

Paso Si calcular el χ^2_{obs}

Observados: 10, 10, 16

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \text{ con } \nu = k - 1$$

esperados: 9,78, 9,74, 11,02

$$\chi^2_{\text{obs}} = \frac{(10-9,78)^2}{9,74} + \frac{(10-9,74)^2}{9,74} + \frac{(16-11,02)^2}{11,02} = 2,26$$

$$\chi^2_{\text{obs}} = 2,26 \quad \chi^2_c = \chi^2_{0,03,2} = 7,01332$$

1) Como $\chi^2_{\text{obs}} = 2,26 < \chi^2_c = 7,01332$

No se rechaza H_0 , entonces
no existe evidencia de lo propuesto

Valor ℓ :

$$\chi^2_{\text{obs}} = \frac{(10 - 9,79)^2}{9,79} + \frac{(10 - 9,79)^2}{9,79} + \frac{(16 - 17,02)^2}{17,02} = 2,26$$

$$\chi^2_{\text{obs}} = 2,26 \quad \alpha = 0,03 \quad V = 2$$

$$P(\chi^2 > 2,26) = 0,32303$$

R/ Como $\ell = 0,32303 > \alpha = 0,03$
No se rechaza H_0

Casos que NO dan una distribución

Aquí se define el $P_i = \frac{1}{k}$, lo demás igual

Hay seis tipos diferentes de colores en un paquete de chocolates M&Ms: azul, rojo, naranja, amarillo, verde y café. Se quiere saber si la distribución de estos 6 colores sucede en igual proporción. En la fábrica, antes de empacarlos, se toma una muestra simple de 600 chocolates M&Ms y se observan las distribuciones que se muestran en el Cuadro 1. Constraste la hipótesis de que los colores de los M&Ms se adaptan a una distribución uniforme.

(5 puntos)

Color	1	2	3	4	5	6
Cantidad	212	147	103	50	46	42

Dan eso pero no dan una
fórmula como en el pasado

$$H_0: X \sim U \quad n=600 \quad k=6 \quad P_i = \frac{1}{6}$$

$$H_1: X \sim U \quad \lambda=0.05 \quad V=5$$

$$P_i = 600 \cdot \frac{1}{6} = 100 \quad (\text{Todos son iguales})$$

Observados: 212, 147, 103, 50, 46, 42

Esperados: 100, 100, 100, 100, 100, 100

$$\chi^2 = \sum_{i=1}^k \frac{(o_i - e_i)^2}{e_i} \text{ con } \nu = k - 1$$

$$\chi^2_{0.05} = \frac{(212-100)^2}{100} + \frac{(147-100)^2}{100} + \frac{(103-100)^2}{100}$$

$$+ \frac{(50-100)^2}{100} + \frac{(46-100)^2}{100} + \frac{(42-100)^2}{100}$$

$$\chi^2_{\text{obs}} = 235.42 \quad \chi^2_C = \chi^2_{0.05, 5} = 11.07050$$

Como $\chi^2_{\text{obs}} = 235.42 > \chi^2_C = 11.07050$
Se rechaza H_0

$$P = P(\chi^2 > \chi^2_{\text{obs}}) = P(\chi^2 > 235.42) \approx 0 \quad [P = 0 < \alpha = 0.05, \text{ se rechaza } H_0]$$