Resumen de materia

IC para 1 población

Si no dan α en cualquier tipo de ejercicio, se asume $\alpha = 0{,}05$

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 1 proporción solo Z

Para 1 varianza solo χ^2

IC;1 proporcion

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

1) Intervalo de confianza (IC)

n_i = Cantidad total

$\hat{p} = \dfrac{\text{Todos los que cumplen}}{\text{Cantidad total}}$, Ej: "De 30 (Totales (n)), 20 (cumplen) comen"

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

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

α = 1 − 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[

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

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

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

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

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

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

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

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

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

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

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

$n \geq 2212,671284$

$\boxed{n \geq 2213}$ → Redondeando hacia arriba

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

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

$n \geq 2400,90$

$\boxed{n \geq 2401}$

Centro Radio

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

3) Confianza utilizada

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-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

$\frac{12}{18}$ aprobados / totales

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%

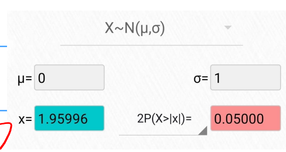
$n = 18 \quad p = \frac{12}{18} \quad q = \frac{2}{3}$ No se usa $\cancel{\hat{p} \pm}\, z_{\alpha/2}\sqrt{\frac{pq}{n}}$

$$E = \frac{0{,}884444 - 0{,}448889}{2} = 0{,}2177775$$ Calculando E

$$Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\frac{2}{3}\cdot\frac{1}{3}}{18}} = 0{,}2177775$$ Remplazando y igualando a E

$$Z_{\frac{\alpha}{2}} = \frac{0{,}2177775}{\sqrt{\frac{\frac{2}{3}\cdot\frac{1}{3}}{18}}}$$ Despejando $Z_{\frac{\alpha}{2}}$

$$Z_{\frac{\alpha}{2}} = 1{,}9599975$$ ← Esto en la x de la app con 2 colas



$$\boxed{1 - 0{,}05 = 0{,}95 \rightarrow 95\ \%}$$

4) Caso especial:

Piden cual fue el tamaño de muestra UTILIZADO

Centro Radio

$\cancel{\hat{p}} \pm z_{\alpha/2}\sqrt{\frac{pq}{n}}$ $p = \frac{a+b}{2}$ $E = \frac{b-a}{2}$ $\alpha = 1 -$ Confianza y luego hacer $\frac{\alpha}{2}$

$q = 1 - p$

Aqui suelen dar solo un IC $]a, b[$ y un nivel de confianza, entonces usando las formulas 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?

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

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

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

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

$$p = \frac{0{,}1608 + 0{,}2392}{2}$$

$$p = 0{,}20 \quad q = 0{,}80$$

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

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

$$\frac{0{,}2392 - 0{,}1608}{2}$$

$$= 0{,}0392$$

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

$$\alpha = 0{,}05$$

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

$$z_{0{,}025} = 1{,}95996$$

$$n = \frac{0{,}20 \cdot 0{,}80}{\left(\frac{0{,}0392}{1{,}95996}\right)^2}$$

No importa el Simbolo, pues se eleva a la 2

$$n = 399{,}98$$

$$\boxed{n = 400}$$

IC i 1 Varianza

I) Intervalo de confianza (IC)

n = Cantidad total

s^2 = Varianza muestral

$V = n-1$ $\quad \alpha = 1 - \text{Confianza}$

Varianza $\left\{ \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}} < \sigma^2 < \dfrac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}} \right.$ con $\nu = n-1$

Desviación estandar $\left\{ \sqrt{\dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}} < \sigma^2 < \sqrt{\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}} \right.$ con $\nu = n-1$

$X^2_{C_1} = X^2_{1-\frac{\alpha}{2}, V}$ $\quad X^2_{C_2} = X^2_{\frac{\alpha}{2}, V}$, el símbolo en la app siempre con C

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/* $]0.120\,105\,, 0.751\,004[$

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

$n = 15$ $\quad S = 0{,}50$ $\quad \alpha = 0{,}02$ $\quad \frac{\alpha}{2} = 0{,}01$

$V = 14$ $\quad S^2 = 0{,}25$ $\quad 1 - \frac{\alpha}{2} = 0{,}99$

$X^2_{0.99, 14} = 29{,}14124$ $\quad X^2_{0{,}01} = 4{,}66043$

$a = \dfrac{14 \cdot 0{,}25}{29{,}14124} = 0{,}120105$ $\quad b = \dfrac{14 \cdot 0{,}25}{4{,}66043} = 0{,}751004$

$$\boxed{\text{R/ El IC para 98\% corresponde a }]0.120105,\ 0.751004[}$$

Ejemplo para desviacion estandar

Un medio afirma que los peso 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 [

$n = 10 \qquad gl = 9\ (n-1) \qquad \alpha = 0{,}05$

$s^2 = 1{,}28$ (calculadora)

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

$]a, b[$

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

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 : $\left]\sqrt{\dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}}, \sqrt{\dfrac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}}\right[$

$$\left]\sqrt{\frac{9 \cdot 1{,}28}{19{,}02277}}, \sqrt{\frac{9 \cdot 1{,}28}{2{,}70039}}\right[$$

$$\boxed{]0{,}779,\ 2{,}0675[}$$

2) ~~Tamaño de muestra~~ NO existe para IC de 1 varianza

3) Confianza utilizada

Aqui dan un un intervalo]a, b[pero la formula tiene 2 partes entonces hay 2 maneras de hacerlo:

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

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

Para varianzas

$$\frac{(n-1)\cdot s^2}{\chi^2_{1-\frac{\alpha}{2},v}} = a \quad \vee \quad \frac{(n-1)\cdot s^2}{\chi^2_{\frac{\alpha}{2},v}} = b$$

Ojo!!!

Misma cosa, solo que uno tiene raiz

Para desviaciones estandar

$$\sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{1-\frac{\alpha}{2},v}}} = a \quad \vee \quad \sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{\frac{\alpha}{2},v}}} = b$$

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

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

$$1-\frac{\alpha}{2} = \text{resultado} \qquad\qquad \frac{\alpha}{2} = \text{resultado}$$
$$\frac{\alpha}{2} = 1 - \text{resultado} \qquad\qquad \alpha = 2\cdot\text{resultado}$$
$$\alpha = 2(1-\text{resultado})$$

Y al final hacer $1-\alpha$ para obtener la confianza

Ejemplos abajo

Ejemplo de desviacion estandart

[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 $I =]\underbrace{14.6725}_{a} \, , \underbrace{37.38}_{b}[$

Usando a

$n=8 \quad v=7 \quad s^2 = 432{,}69643$ a   b

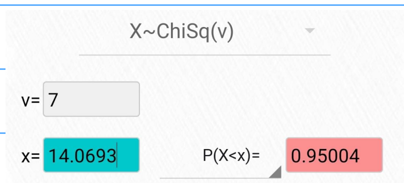
$$\sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2,\nu}}} < \sigma^2 < \sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2,\nu}}}$$
con $\nu = n-1$

$$\sqrt{\frac{7 \cdot 432{,}69643}{\chi^2_{1-\frac{\alpha}{2},7}}} = 14{,}6725$$ Igualando a

$$\frac{7 \cdot 432{,}69643}{\chi^2_{1-\frac{\alpha}{2},7}} = 14{,}6725^2$$

$$\chi^2_{1-\frac{\alpha}{2},7} = \frac{7 \cdot 432{,}69643}{14{,}6725^2}$$ Despejando el χ^2

$$\chi^2_{1-\frac{\alpha}{2},7} = 14{,}0693$$ Esto en la app con <



$$1-\frac{\alpha}{2} = 0{,}95004$$
$$\frac{\alpha}{2} = 1-0{,}95004$$
$$\alpha = 2(1-0{,}95004)$$
$$\alpha = 0{,}10$$
Igualar el resultado y despejar α

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

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 $I =]14.6725\ ,\ 37.38[$

$\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad a \quad\quad b$

Usando b

$n=8 \quad v=7 \quad s^2 = 432{,}69643 \qquad a \quad b$

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

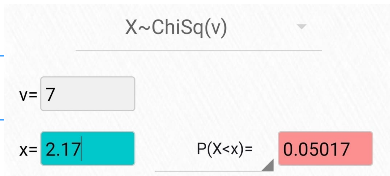
$$\sqrt{\frac{7 \cdot 432{,}69643}{\chi^2_{1-\frac{\alpha}{2},7}}} = 37{,}38 \quad \Big\} \text{ Igualando b}$$

$$\frac{7 \cdot 432{,}69643}{\chi^2_{1-\frac{\alpha}{2},7}} = (37{,}38)^2$$

$$\chi^2_{1-\frac{\alpha}{2},7} = \frac{7 \cdot 432{,}69643}{(37{,}38)^2}$$

Despejando el χ^2

$\chi^2_{1-\frac{\alpha}{2},7} = 2{,}17$ } Esto en la app con <



$$\frac{\alpha}{2} = 0{,}05$$
$$\alpha = 0{,}10$$

} Igualar el resultado y despejar α

$1 - 0{,}10 = 0{,}90 \rightarrow \boxed{90\%}$ } 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 utiliza Z

Para todos los IC de 2 poblaciones
Si todo el IC > 0, A > B] 0.82, 80 [
Si todo el IC < 0, A < B] -3.5, -4 [
Si hay > 0 ∧ < 0, No se sabe] 0, -2 [

Para 2 varianzas solo se usa F

Para IC de 2 varianzas
Se asumen varianzas iguales
si $1 \in IC$, diferentes si $1 \notin IC$

Iguales → IC:] 0,73, 1.33 [, $1 \in IC$
Diferentes → IC:] 0,50, 0,70 [, $1 \notin IC$

El símbolo > < en la app es irrelevante para cualquier distribución en IC

IC: 2 medias

Existen 3 casos y 3 formulas distintas

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

I) Intervalo de confianza (IC) para caso I

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 $\sigma_A = 4kg.$ y $\sigma_B = 5kg.$ 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. Según 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. estandar

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.

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

$n_1, n_2 \geq 30$

$n_1 = 50 \qquad n_2 = 40 \qquad \alpha = 0{,}05$

$\bar{x}_1 = 49{,}8 \qquad \bar{x}_2 = 47{,}5 \qquad \frac{\alpha}{2} = 0{,}025$

$\sigma_1 = 4 \qquad \sigma_2 = 5 \qquad z_{0{,}025} = \pm 1{,}95996$

$$a = 49{,}5 - 47{,}5 - 1{,}95996 \cdot \sqrt{\frac{4^2}{50} + \frac{5^2}{40}} = 0{,}3947$$

$$a = 49{,}5 - 47{,}5 + 1{,}95996 \cdot \sqrt{\frac{4^2}{50} + \frac{5^2}{40}} = 4{,}2052$$

$$\boxed{\text{El IC para 95\% corresponde a }]0.3947\,;\,4.2054[}$$

2) 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 σ_M= 4.1 y σ_N= 5.8. [redacted]

de [redacted]

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*

$n_1 = 45 \geq 30$ $\quad n_2 = 38 \geq 30$

$\bar{x}_1 = 71,3$ $\quad \bar{x}_2 = 68,2$

$\sigma_1 = 4,1$ $\quad \sigma_2 = 5,8$

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

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

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

$n \geq 48,4520$

$n \geq 49$

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

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 $\overline{x}_1 - \overline{x}_2$ observado en las muestras y el valor aproximado de $\sigma_1^2 + \sigma_2^2$. → Despejar eso nada mas

$$\boxed{\overline{x}_1 - \overline{x}_2 = \frac{165{,}5 + 192{,}9}{2} = 179{,}2}$$

$n_1 = n_2 = 50 \qquad \alpha = 0{,}10$

$\frac{\alpha}{2} = 0{,}05 \qquad z_{0{,}05} = 1{,}64485$

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

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

$$E = \frac{192{,}9 - 165{,}5}{2} = 13{,}7$$

$$1{,}64485 \cdot \sqrt{\frac{\sigma_1^2}{50} + \frac{\sigma_2^2}{50}} = 13{,}7$$

$$\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{50}} = \frac{13{,}7}{1{,}64485}$$

$$\frac{\sigma_1^2 + \sigma_2^2}{50} = \left(\frac{13{,}7}{1{,}64485}\right)^2$$

$$\sigma_1^2 + \sigma_2^2 = \left(\frac{13{,}7}{1{,}64485}\right)^2 \cdot 50$$

$$\boxed{\sigma_1^2 + \sigma_2^2 = 3468{,}63}$$

Caso 2: $n_1 < 30$ v $n_2 < 30$ y no se conocen las Desviaciones estandart poblacionales σ_1 y σ_2, mejor dicho Si le dicen 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 datoss:

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

< 30

Suponga que el redimiento 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)

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

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

y $\nu = n_1 + n_2 - 2$

$n_1 = 19$ $\quad n_2 = 17$ $\quad v = 19 + 17 - 2 = 34$

$\bar{x}_1 = 78$ $\quad \bar{x}_2 = 65$ $\quad \alpha = 0{,}05$

$s_1 = 4{,}3$ $\quad s_2 = 4{,}7$ $\quad \frac{\alpha}{2} = 0{,}025 \rightarrow t_{0{,}025,\,34} = 2{,}03224$

$$S_p^2 = \frac{(19-1)\cdot 4{,}3^2 + (17-1)\cdot 4{,}7^2}{19+17-2} \approx 20{,}1891$$

$$a = 78 - 65 - 2{,}03224 \cdot \sqrt{\frac{20{,}1891}{19} + \frac{20{,}1891}{17}} = 9{,}9519$$

$$b = 78 - 65 - 2{,}03224 \cdot \sqrt{\frac{20{,}1891}{19} + \frac{20{,}1891}{17}} = 16{,}0481$$

R/ $\boxed{\text{El IC para 95\% corresponde a }]9.9519,\ 16.0481[}$

Caso 3: $n_1 < 30$ V $n_2 < 30$ y no se conocen las Desviaciones estandart 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 < 30

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 varianzias 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.

$n_1 = 21$ $n_2 = 17$ $\alpha = 0{,}05$

$\bar{x}_1 = 65$ $\bar{x}_2 = 63$ $\frac{\alpha}{2} = 0{,}025$

$S_1 = 24$ $S_2 = 15$

$$V = \frac{\left(\frac{24^2}{21} + \frac{15^2}{17}\right)^2}{\frac{\left(\frac{24^2}{21}\right)^2}{21-1} + \frac{\left(\frac{15^2}{17}\right)^2}{17-1}} = 34{,}0484$$

$t_{0{,}025,\,34{,}0484} = \pm 2{,}03234$

$$a = 65 - 63 - 2{,}03234 \cdot \sqrt{\frac{24^2}{21} + \frac{15^2}{17}} = -10{,}9586$$

$$b = 65 - 63 + 2{,}03234 \cdot \sqrt{\frac{24^2}{21} + \frac{15^2}{17}} = 14{,}9586$$

R/ El IC para 95% corresponde a $]-10.9586, 14.9586[$

2) Interpretar el IC

2. ¿Es aceptable la afirmación del profesor?

IC es positivo y negativo, no

IC: 2 Proporciones

n_i = Cantidad total

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}$$

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

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

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

1) 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[*

$\hat{p}_1 = \frac{31}{126}$ $\quad \hat{p}_2 = \frac{17}{138}$ $\quad \alpha = 0,10$

$\hat{q}_1 = \frac{95}{126}$ $\quad \hat{q}_2 = \frac{121}{138}$ $\quad \frac{\alpha}{2} = 0,05$

$n_1 = 126$ $\quad n_2 = 138$ $\quad Z_{0,05} = \pm 1,64485$

$$\widehat{p_1} - \widehat{p_2} \pm Z_{\alpha/2}\sqrt{\frac{\widehat{p_1}\,\widehat{q_1}}{n_1} + \frac{\widehat{p_2}\,\widehat{q_2}}{n_2}}$$

$$\frac{31}{126} - \frac{17}{138} - 1,64485 \cdot \sqrt{\frac{\frac{31}{126}\cdot\frac{95}{126}}{126} + \frac{\frac{17}{138}\cdot\frac{121}{138}}{138}} \approx 0,04473$$

$$\frac{31}{126} - \frac{17}{138} + 1,64485 \cdot \sqrt{\frac{\frac{31}{126}\cdot\frac{95}{126}}{126} + \frac{\frac{17}{138}\cdot\frac{121}{138}}{138}} \approx 0,20095$$

$\boxed{\text{R/ El IC para 90\% corresponde a }]0.04473, 0.20095[}$

2) Interpretar el IC

¿Se puede asumir que las proporciones de tostadores defectuorsos 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 $p_i q_i = 0{,}25$

$$n \geq \frac{z^2_{\alpha/2}(p_1 q_1 + p_2 q_2)}{r^2}$$

Si se desea encontrar un IC de ~~90%~~ 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/ [redacted]

Caso 1: Usando p y q

$\hat{p}_1 = \frac{31}{126}$ $\hat{p}_2 = \frac{17}{138}$

$\hat{q}_1 = \frac{95}{126}$ $\hat{q}_2 = \frac{121}{138}$

$n_1 = 126$ $n_2 = 138$

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

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

$$n \geq 450{,}9943 \rightarrow \boxed{n \geq 451}$$

Caso 2: Sin importar pq, $p_i q_i = 0{,}25$

$$n \geq \frac{(1{,}95996)^2 (0{,}25 + 0{,}25)}{(0{,}05)^2} \qquad n \geq \frac{z^2_{\alpha/2}(p_1 q_1 + p_2 q_2)}{r^2}$$

$$n \geq 768{,}2886 \rightarrow \boxed{n \geq 769}$$

Centro Radio

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

4) Confianza utilizada

Aqui igualmente es igualar el radio a E y despejar el

ese resultado a la app con cola derecha

y el resultado de la app se hace esto

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

$\alpha = 2 \cdot$ resultado

Y al final $1-\alpha$

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 $n_1\hat{p}_h \geq 5$, $n_1\hat{q}_h \geq 5$, $n_2\hat{p}_m \geq 5$ y $n_2\hat{q}_m \geq 5$. 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)

$n_1 = 800$ $\quad$ $n_2 = 800$

$p_1 = 0{,}78125$ $\quad$ $p_2 = 0{,}75$

$q_1 = 0{,}21875$ $\quad$ $q_2 = 0{,}25$

$$z_{\alpha/2}\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}} = E$$

$$E = \frac{b-a}{2} = \frac{0{,}07471958 - -0{,}0122196}{2} = 0{,}04346959$$

$$Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{0{,}78125 \cdot 0{,}21875}{800} + \frac{0{,}75 \cdot 0{,}25}{800}} = 0{,}04346959$$

$$Z_{\frac{\alpha}{2}} = \frac{0{,}04346959}{\sqrt{\frac{0{,}78125 \cdot 0{,}21875}{800} + \frac{0{,}75 \cdot 0{,}25}{800}}} = 2{,}05375$$

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

$\alpha = 0{,}04$

$1 - 0{,}04 = 0{,}96 \rightarrow \boxed{96\%}$

5) Caso especial I

Encontrar la cantidad que cumplen, por ejemplo "De 90, x no comen", encontrar ese x es en sí encontrar el centro, para esto dan un intervalo]a, b[y se usa la formula

$$p_1 - p_2 = \frac{a+b}{2}$$, 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

$n_1 = 250$ $n_2 = 240$

$p_1 = \frac{164}{250}$ $p_2 = \frac{x}{240}$

$$\frac{164}{250} - \frac{x}{240} = \frac{0,0709767 + 0,216023}{2}$$

$$x = \left(-\frac{0,0709767 + 0,216023}{2} + \frac{164}{250} \right) \cdot 240$$

$$\boxed{x \approx 123}$$

IC: 2 Varianzas

con $\nu_i = n_i - 1$

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

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 $1 \in IC$, diferentes si $1 \notin IC$

Iguales → IC:]0,73, 1,13[, $1 \in IC$
Diferentes → IC:]0,50, 0,70[, $1 \notin IC$

Para Todos los IC de 2 Poblaciones
Si Todo el IC >0, $A>B$]0.82, 80[
Si Todo el IC <0, $A<B$]-3,5, -7[
Si hay $>0 \land <0$, No se sabe]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 σ_M= 4.1 y σ_N= 5.8. [redacted]

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?

[redacted] Comparar 1 con 2

$n_1 = 45 \quad n_2 = 38 \quad \alpha = 0{,}05$

$v_1 = 44 \quad v_2 = 37 \quad \frac{\alpha}{2} = 0{,}025$

$s_1 = 4{,}1 \quad s_2 = 5{,}8$

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

$f_{0.025,37,44} = 0{,}52954$

$f_{0.025,44,37} = 0{,}53870$

$$a = \frac{4{,}1^2 \cdot 0{,}52954}{5{,}8^2} = 0.2646$$

$$b = \frac{4{,}1^2}{5{,}8^2 \cdot 0{,}53870} = 0{,}9276$$

R/ 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 $1 \notin IC$ 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 datoss:

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

Suponga que el redimiento 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

$n_1 = 19 \quad n_2 = 17 \quad \alpha = 0,10$

$v_1 = 18 \quad v_2 = 16 \quad \frac{\alpha}{2} = 0,05$

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

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

$f_{0.05,18,16} = 0,44953$

$f_{0.05,16,18} = 0,43447$

$$a = \frac{4,7^2 \cdot 0,44953}{4,3^2} = 0.53107$$

$$b = \frac{4,7^2}{4,3^2\, 0,43447} = 2.74978$$

R/ El IC para 90%, corresponde a $]0.53107, 2.74978[$

Como $1 \in$ IC, se asumen varianzas iguales

Ph para una poblacion

Como plantear bien cualquier H_0 y H_1?
Esto sirve para cualquier prueba de hipotesis 1 y 2 poblaciones

Las desigualdades NO estrictas $\leq, \geq, =$ van siempre en H_0 por lo que el H_0 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 implicito por que H_0 siempre usa $=$

Las desigualdades estrictas $<, >, \neq$ van siempre en H_0 por lo que el H_1 seria el contrario

$H_1: < \rightarrow H_0: \geq$
$H_1: > \rightarrow H_0: \leq$
$H_1: = \rightarrow H_0: \neq$

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

Las Ph se pueden resolver de varias maneras

1) Normal
2) Valor P
3) RR y RA (Regiones de aceptacion y rechazo)
4) Otros metodos especificos

Cuando se acepta o rechaza H_0?

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

V es una distribucion x^2, z, F, z, etc

<u>Cola izquierda</u> $<$

si se cumple
$V_{obs} > V_C$, NO se rechaza H_0

Si se cumple
$V_{obs} < V_C$, Se rechaza H_0

<u>Cola derecha</u>

si se cumple
$V_{obs} < V_C$, NO se rechaza H_0

Si se cumple
$V_{obs} > V_C$, Se rechaza H_0

<u>2 colas</u>

si se cumple
$V_{C1} < V_{obs} < V_{C2}$, NO se rechaza H_0

<u>Con valor P</u>

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 1 Proporcion

n = Cantidad total

$\hat{p} = \dfrac{\text{Los que cumplen}}{\text{Cantidad total}}$

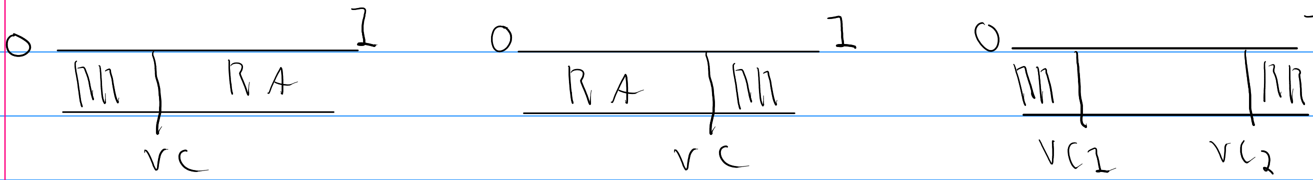
$p_0 = H_0$

$q_0 = 1 - p_0$

Formula de estandarizacion

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

Recordar que las varianzas estan entre 0 y 1



Cola izquierda RR:]0, VC[

RA:]VC, 1[

Cola derecha RR:]VC, 1[

RA:]0, VC[

Dos colas RR: $]0, VC_1[\cup]VC_2, 1[$

RA: $]V_1, V_2[$

Ph 1 proporcion

Formula de estandarizacion

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

n = Cantidad total

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

$p_0 = H_0$

$q_0 = 1 - p_0$

Ejemplo 4: una proporción

$H_0: \leq$

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

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

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

$H_1: p > 0{,}80$

$n = 500$

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

$p_0 = 0{,}80$

$q_0 = 0{,}20$

$\alpha = 0{,}05$

Estandarizando

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

3,9131

RA	RR

1,64

$Z_c = Z_{0{,}05} = 1{,}64485$

$\boxed{\text{RR Como } Z_{obs} = 3{,}9131 > Z_c = 1{,}64485 \text{ se rechaza } H_0}$

Tip mental: Si se cumple el simbolo del mismo tipo de cola, se rechaza, por ejemplo aqui era cola derecha y se cumplio 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

$$Z_{0,05} = \frac{0{,}87 - 0{,}80}{\sqrt{\frac{0{,}80 \cdot 0{,}20}{500}}} = 3{,}9131$$

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

$\boxed{\text{R/ Como } 0{,}00005 < 0{,}05 \text{ se rechaza } H_0}$

Con RR y RA

en terminos de p

$\hat{p}_0 \pm z_{\alpha/2}\sqrt{\frac{p_0 q_0}{n}}$

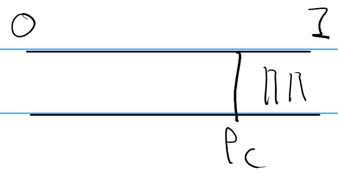
$H_0: p = 0{,}80 \ (\leq)$ $\quad n = 500$ $\quad p_0 = 0{,}80$

$H_1: p > 0{,}80$ $\quad \hat{p} = \frac{435}{500} = 0{,}87$ $\quad q_0 = 0{,}20$

$\alpha = 0{,}05$

$Z_{0,05} = 1{,}64485$

$$b = 0{,}8 + 1{,}64485 \cdot \sqrt{\frac{0{,}80 \cdot 0{,}20}{500}} = 0{,}8294 \leftarrow P_c$$



RR: $]0.8294, 1[$

RA: $]0, 0.8294[$

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

$H_0: \leq$

50%

Tamaño de muestra

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.

Redondeo

(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~~ 274

$H_0: p = 0{,}5 (\leq)$ $\quad p_0 = 0{,}5 \quad p_1 = 0{,}60$

$$n \geq \frac{(|z_{\alpha/k}|\sqrt{p_0q_0} + |z_\beta|\sqrt{p_1q_1})^2}{(p_1 - p_0)^2}$$

donde k es el número de colas

$H_1: p > 0{,}5$ $\quad q_0 = 0{,}5 \quad q_1 = 0{,}40$

$H'_1: p = 0{,}60 \quad \alpha = 0{,}02 \quad \beta = 0{,}10$

$Z_{0.02} = 2{,}05375 \quad Z_{0.10} = 1{,}28155$

$$n \geq \frac{(2{,}05375 \cdot \sqrt{0{,}5 \cdot 0{,}5} + 1{,}28155 \cdot \sqrt{0{,}60 \cdot 0{,}40})^2}{(0{,}60 - 0{,}5)^2}$$

$n \geq 273{,}80 \longrightarrow \boxed{n \geq 274}$

Error tipo I (Basicamente 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.

$H_0: \leq$

2. Encuentre la probabilidad del error tipo I

$H_0: P = 0{,}30 (\leq)$ $\quad \hat{P} = \frac{1}{3} \quad n = 300$

$$Z = \frac{\hat{P} - p_0}{\sqrt{p_0q_0/n}}$$

$H_1: p > 0{,}30 \; (\hat{P} > 0{,}30)$ $\quad p_0 = 0{,}30 \quad q_0 = 0{,}70$

$\alpha = 0{,}05 \quad Z_{0.05} = 1{,}64485$

$$Z_{obs} = \frac{0{,}33 - 0{,}30}{\sqrt{\frac{0{,}30 \cdot 0{,}70}{300}}} = 1{,}25988$$

$P(Z > 1{,}25988) = \boxed{0{,}10386}$

Error tipo 2

Remplazar el p_0 y q_0 por el p_1 y q_1

$q_1 = H_1'$ $\quad Z = \dfrac{\hat{P}_c - p_1}{\sqrt{p_1 q_1/n}}$ y eso se mete en la app con el simbolo contrario

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

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

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

$\hat{p} = \frac{1}{3}$ $\quad n = 300$

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

$\alpha = 0{,}05$ $\quad Z_{0{,}05} = 1{,}64485$

$H_2: p = 0{,}40$

$P_c = \frac{1}{3}$ $\quad p_1 = 0{,}40$ $\quad q_1 = 0{,}60$

Estandarizar para error 2 $\quad Z = \dfrac{\hat{P}_c - p_1}{\sqrt{p_1 q_1/n}}$

$$Z = \frac{\frac{1}{3} - 0{,}40}{\sqrt{\frac{0{,}40 \cdot 0{,}60}{300}}} = -2{,}35702$$

$$P(Z < -2{,}35702) = \boxed{0{,}009137}$$

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?

$H_0: p = 0,30$

$H_1: p > 0,30$

$H_2: p = 0,40$

$\hat{p} = \frac{1}{3}$ $\quad n = 300$

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

$p_1 = 0,40 \quad q_1 = 0,60$

$\alpha = 0,05$

$\beta = 0,02$

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

donde k es el número de colas

$Z_{0.05} = 1,645$

$Z_{0.02} = 2,05$

$$n \geq \frac{(1,645 \cdot \sqrt{0,30 \cdot 0,70} + 2,05 \cdot \sqrt{0,40 \cdot 0,60})^2}{(0,40 - 0,30)^2}$$

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

Ph: 1 Varianza

Esta se puede resolver de 4 maneras

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

Usando X^2
Usando S^2
Usando S

} Solo es despejar la formula

Usando valor P, lo mismo de siempre

$$X^2_{obs} = \frac{(n-1) \cdot s^2}{\sigma^2_0} \qquad s^2_C = \frac{X^2_C \cdot \sigma^2_0}{n-1} \qquad s_C = \sqrt{\frac{X^2_C \cdot \sigma^2_0}{n-1}}$$

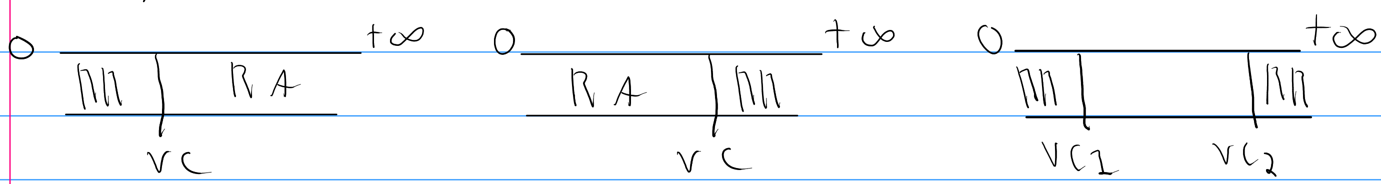
$$X^2_C = X^2_{\alpha, V} \qquad S^2_{obs} = S^2 \qquad S_{obs} = S$$

(Para S^2) Aqui es al reves, el s^2_C es el estandarizado y el $s^2_{obs} = s^2$

(Para S) Aqui es al reves, el s^2_C es el estandarizado y el $s_{obs} = s$

Cual usar? el enunciado dice

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



Cola izquierda RR: $]0, VC[$ RA: $]VC, +\infty[$
Cola derecha RR: $]VC, +\infty[$ RA: $]0, VC[$
Dos colas RR: $]0, VC_1[\cup]VC_2, +\infty[$
RA: $]V_1, V_2[$

Ejemplo 8: una varianza

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?.

Es con S_c pero hagamos las 4 maneras para aprender

Usando χ^2_c $\quad n=20 \quad \sigma_0=8$ $\quad \chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ con $\nu = n-1$

$H_0: \sigma = 8 \ (\geq) \ 8>5$ $\quad v=19 \quad S=5 \quad \alpha=0{,}10$

$H_1: \sigma < 8$

$$\chi^2_{obs} = \frac{19 \cdot 5^2}{8^2} = 7{,}4218$$

$$\chi^2_c = \chi^2_{0.10,\,19} = 11{,}65091$$

$$\boxed{\text{R/ Como } \chi^2_{obs} = 7{,}4218 < \chi^2_c = 11{,}65091 \text{ Se rechaza } H_0}$$

Usando S^2_c $\quad \chi^2_c = \frac{(n-1)S^2}{\sigma_0^2}$ con $\nu = n-1$ $\quad \sigma^2_{obs}$

Aquí el $S^2_{obs} = 5^2$

$n=20 \quad \sigma_0=8$

$$\chi^2_c = \frac{(n-1)\cdot S^2}{\sigma_0^2}$$ $\quad v=19 \quad S=5 \quad \alpha=0{,}10$

$$S^2 = \frac{\chi^2_c \cdot \sigma_0^2}{n-1} \qquad \chi^2_c = \chi^2_{0.10,\,19} = 11{,}65091$$

S^2_c

$$S^2_c = \frac{11{,}65091 \cdot 8^2}{19} = 39{,}2452 \qquad S^2_{obs} = 5^2 = 25$$

S^2_{obs}

$$\boxed{\text{R/ Como } S^2_{obs} = 25 < S^2_c = 39{,}2452 \text{ se rechaza } H_0}$$

Usando S_c Aquí el $S_{obs} = 5$

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

$n = 20$ $\sigma_0 = 8$

$\nu = 19$ $s = 5$ $\alpha = 0,10$

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

$$\chi_c^2 = \chi^2_{0.10,\,19} = 11,65091$$

$$S_c = \sqrt{\frac{\chi_c^2 \cdot \sigma_0^2}{n-1}}$$

$$S_c = \sqrt{\frac{11,65091 \cdot 8^2}{19}} = 6,2646 \qquad S_{obs} = 5$$

$$\boxed{\text{R/ Como } S_{obs} = 5 < S_c = 6,2646 \text{ Se rechaza } H_0}$$

Usando valor P

$n = 20$ $\sigma_0 = 8$ $\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$ con $\nu = n - 1$

$\nu = 19$ $S = 5$ $\alpha = 0,10$

$$\chi^2 = \frac{19 \cdot 5^2}{8^2} = 7,421875$$

$$P(\chi^2 < 7,421875,\ 19) = 0,00840$$

$$\boxed{\text{R/ Como } P = 0,00840 < 0,10 \text{ Se rechaza } H_0}$$

Error tipo 2

Normalmente hay que buscar el s^2 despejando con los valores normales y luego usar la $\sigma_1^2 = H_1'$ en vez de σ_0^2

$$\chi^2 = \frac{(n-1)S^2}{\sigma_{0_1}^2} \text{ con } \nu = n-1$$

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*

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

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

$\alpha = 0{,}10$

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

$H_1': \sigma = 4700$

Primero se debe encontrar S^2 con los valores normales

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

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

$$S = \frac{\chi_c^2 \cdot \sigma_0^2}{n-1}$$

$\sigma_0^2 = 5000^2$

$n-1 = 24$

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

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

$$\chi^2 = \frac{(n-1)S^2}{\sigma_{0_1}^2} \text{ con } \nu = n-1$$

$$\chi^2 = \frac{24 \cdot 16311125}{4700^2}$$

$n = 25 \quad \sigma_0 = 4700$

$V = 24 \quad S = 16311125$

Contrario de <

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

$\rightarrow \boxed{0{,}816307}$

RR y 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 estandár del salario mensual de un ingenierio 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 estándar 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^2_c.**

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

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

$\alpha = 0{,}10$

$\chi^2_c = \frac{(n-1) \cdot s^2}{\sigma^2_0}$

$\chi^2_c = \chi^2_{0{,}10,\,39} = 28{,}19579$

$\sigma^2 = 100^2$

$s^2_c = \frac{\chi^2_c \cdot \sigma^2_0}{n-1}$ $\quad n - 1 = 39$ $\qquad$ Aqui el $S^2_{obs} = S^2$

s^2_c

$s^2_c = \frac{28{,}19579 \cdot 100^2}{39} = 7229{,}69 \leftarrow$

0 ______________ $+\infty$

RR | RA

s^2_c

RR: $]0,\ 7229.69[$

RA: $]7229.69,\ +\infty[$

$$\boxed{\text{R/ Como } S_{obs} = 90^2 = 8100 \in RA \quad \text{No se rechaza } H_0}$$

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?

$H_0: \sigma^2 = 24 \ (\geq)$ $\qquad n = 40 \qquad \alpha = 0{,}05$

$H_1: \sigma^2 < 24$ $\qquad gl = 39 \qquad \sigma_0^2 = 24$

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

↳ Esto esta en terminos de χ^2 pero la pregunta es de σ

Recordando

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

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

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

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

$n = 40 \qquad \alpha = 0{,}05$

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

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

$$\boxed{s = 3{,}9765}$$

Despejar s, si fuera s^2, no aplicar la raiz

Ph para 2 poblaciones
Aqui sigue aplicando todo lo de 1 poblacion

¿Como plantear bien cualquier H_0 y H_1?
Esto sirve para cualquier prueba de hipotesis 1 y 2 poblaciones

Las desigualdades NO estrictas $\leq, \geq, =$ van siempre en H_0 por lo que el H_0 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 implicito por que H_0 siempre usa $=$

Las desigualdades estrictas $<, >, \neq$ van siempre en H_0 por lo que el H_1 seria el contrario

$H_1: < \rightarrow H_0: \geq$
$H_1: > \rightarrow H_0: \leq$
$H_1: = \rightarrow H_0: \neq$

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

Las Ph se pueden resolver de varias maneras
1) Normal
2) Valor P
3) RR y RA (Regiones de aceptacion y rechazo)
4) Otros metodos especificos

Cuando se acepta o rechaza H_0?

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

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

Cola izquierda <

Si se cumple
$V_{obs} > VC$, NO se rechaza H_0

Si se cumple
$V_{obs} < VC$, Se rechaza H_0

Cola derecha

Si se cumple
$V_{obs} < VC$, NO se rechaza H_0

Si se cumple
$V_{obs} > VC$, Se rechaza H_0

2 Colas

Si se cumple
$VC_1 < V_{obs} < VC_2$, 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" medias, proporciones o varianzas, es de 2 colas entonces $H_0: =$ 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 d_0 (viene en las 3 formulas)

Si es de 2 colas, $d_0 = 0$
Si es de 1 cola, $d_0 = H_0$ (Puede que H_0 sea 0)

Si NO es de colas, el H_0 suele plantearse como $H_0: u_1 - u_2 = (\leq)$, como una resta
$H_0: u_1 - u_2 = (\geq)$

Caso 1: $n_1, n_2 > 30$ y se conocen las Desviaciones estandart 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.

Siempre que diga "diferencia de" ya sea de media, proporcion o varianzas, es de 2 colas → H_0: =, H_1: ≠

¿Existe diferencia significativa entre las medias de los dos grupos a un nivel de significancia del 5%? Lo son diferentes o iguales?

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

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

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

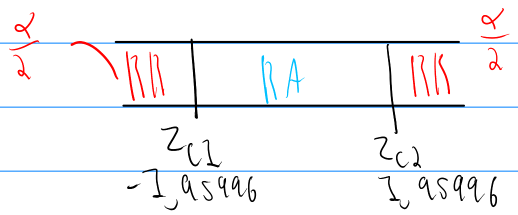
Estandarizando

$n_1 = 50$	$n_2 = 40$	$\alpha = 0{,}05$
$\bar{x}_1 = 65$	$\bar{x}_2 = 62$	$\frac{\alpha}{2} = 0{,}025$
$\sigma_1 = 10$	$\sigma_2 = 8$	$d_0 = 0$

$$Z_{0{,}05} = \frac{65 - 62 - 0}{\sqrt{\frac{10^2}{50} + \frac{8^2}{40}}} \approx 1{,}5811$$

$Z_c = 0{,}025 = \pm 1{,}95996$

Como es de 2 colas, hace falta ver si el Z_c esta en RR u RA



Como

$Z_{c1} < Z_{0{,}05} < Z_{c2}$

$-1{,}95996 < 1{,}5811 < 1{,}95996$ 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 $= 2 \cdot P(Z > Z_{obs})$ 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

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

$\boxed{\text{Como valor P} = 0{,}11386 > 0{,}05, \text{ no se rechaza } H_0}$

Recordar que con valor P indiferentemente del tipo de cola, siempre se hace la comparación

Valor P $> \alpha$, No se rechaza

Valor P $\leq \alpha$, Se rechaza

Caso 2: $n_1 < 30$ v $n_2 < 30$ y no se conocen las Desviaciones estandart poblacionales σ_1 y σ_2, mejor dicho si le dicen que las varianzas poblacionales son iguales y con ver que el n_1 o el n_2 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

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

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

y $\nu = n_1 + n_2 - 2$

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

$H_0: \mu_1 - \mu_2 \leq 3$ $\quad n_1 = 30 \quad n_2 = 20 \quad d_0 = 3$

$H_1: \mu_1 - \mu_2 > 3$ $\quad \bar{x}_1 = 69{,}8 \quad \bar{x}_2 = 63{,}8 \quad V = 48$

$S_1 = 1{,}92 \quad S_2 = 2{,}18 \quad \alpha = 0{,}04$

$$S_p^2 = \frac{(30-1)\cdot 1{,}92^2 + (20-1)\cdot 2{,}18^2}{48} = 4{,}108$$

$$T_{obs} = \frac{69{,}8 - 63{,}8 - 3}{\sqrt{\frac{4{,}108}{30} + \frac{4{,}108}{20}}} = 5{,}127$$

$$T_c = t_{0{,}04,\,48} = 1{,}788$$

R/ Como $Z_{obs} = 5{,}127 > 1{,}788$ se rechaza H_0

Valor P

$P(T > 5{,}127) = 0{,}00000$

Como $P = 0{,}0000 < 0{,}04$ se rechaza H_0

Caso 3: $n_1 < 30$ v $n_2 < 30$ y no se conocen las Desviaciones estandart 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 < 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

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

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

$H_0: u_1 - u_2 = 3 \ (\geq)$ $\quad n_1 = 26 \quad n_2 = 21 \quad d_0 = 3$

$H_1: u_1 - u_2 < 3$ $\quad \bar{x}_1 = 26 \quad \bar{x}_2 = 24 \quad V = 44{,}69$

$S_1 = 2{,}3 \quad S_2 = 1{,}7 \quad \alpha = 0{,}04$

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

$$V = \frac{\left(\frac{(2{,}3)^2}{26} + \frac{(1{,}7)^2}{21}\right)^2}{\frac{\left(\frac{(2{,}3)^2}{26}\right)^2}{26-1} + \frac{\left(\frac{(1{,}7)^2}{21}\right)^2}{21-1}} = 44{,}6930$$

$t_{0.04,\ 44.6930} = -1{,}79147$

$$\boxed{\text{R/ Como } t_{obs} = -1{,}712267 < t_c = -1{,}79147 \quad \text{NO se rechaza } H_0}$$

Ph 2 Proporciones

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

$H_0: P_A = P_B \ (\leq)$ $n_1 = 100$ $n_2 = 100$

$H_1: P_A > P_B$ $P_1 = 0{,}32$ $P_2 = 0{,}27$

$X_1 = 32$ $X_2 = 27$ $\alpha = 0{,}04$

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

$\hat{p} = \frac{32+27}{100+100} = 0{,}295$ $\hat{q} = 0{,}705$

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

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

$Z_c = Z_{0.04} = 1{,}75069$

$\boxed{\text{R/ Como } Z_{obs} = 0{,}7752 < Z_c = 1{,}75069 \text{ No se rechaza } H_0}$

Con valor P

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

$\alpha = 0{,}04$

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

$P(Z > 0{,}7752) = 0{,}22$

$\boxed{\text{R/ Como } P = 0{,}22 < 0{,}04 \text{ No se rechaza } H_0}$

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/* $n \geq \cancel{1068{,}85}$ 1071

$p'_1 = 0{,}30 \quad p'_2 = 0{,}25 \quad \alpha = 0{,}04$

$q'_1 = 0{,}70 \quad q'_2 = 0{,}75 \quad \beta = 0{,}20$

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

donde k es el número de colas Cola derecha

$Z_{0,04} = 1{,}75069 \quad Z_{0,20} = 0{,}84162$

$$n \geq \frac{\left(1{,}75069 \cdot \sqrt{\frac{1}{2} \cdot (0{,}30 + 0{,}25)(0{,}70 + 0{,}75)} + 0{,}84162 \cdot \sqrt{0{,}30 \cdot 0{,}70 + 0{,}25 \cdot 0{,}75}\right)^2}{(0{,}30 - 0{,}25)^2}$$

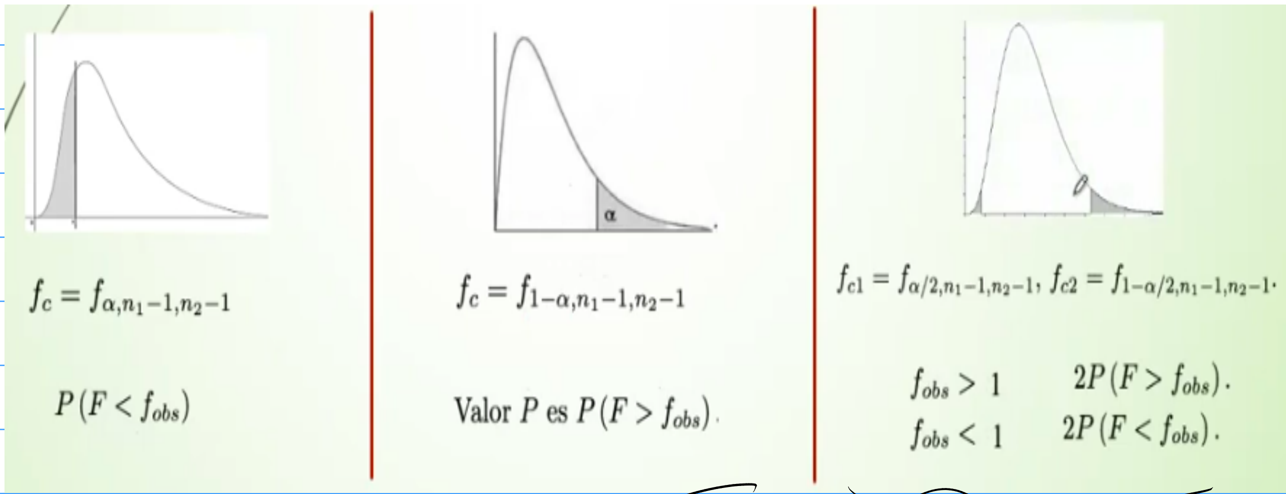
$n \geq 1070{,}75 \rightarrow \boxed{n \geq 1071}$

```
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 | Derecha | Dos colas

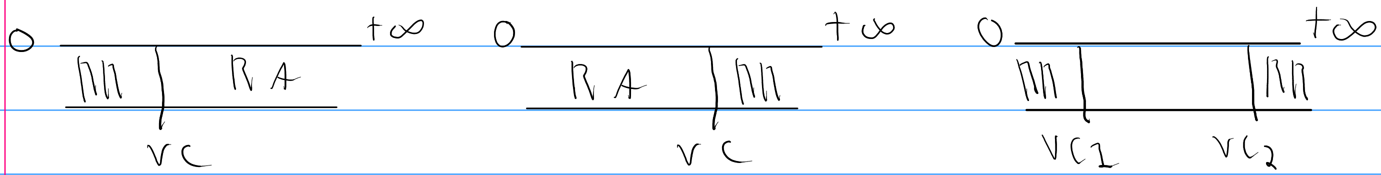


Aquí Si | Aquí si | Realmente se puede meter siempre con < en la app

$P(X < x)$ | $P(X > x)$

$r_0 = H_0$ | $r_0 = H_0$ | $r_0 = 1$

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



Cola izquierda RR: $]0, VC[$ RA: $]VC, +\infty[$
Cola derecha RR: $]VC, +\infty[$ RA: $]0, VC[$
Dos colas RR: $]0, VC_1[\cup]VC_2, +\infty[$
RA: $]V_1, V_2[$

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 $\sigma_m^2 = \sigma_h^2$

2 Colas

$r_0 = 1$

	N° Alumnos	Media	Desv.Std
Mujeres	21	63,8	2,18
Hombres	31	69,8	1,92

$H_0: \frac{\sigma_m^2}{\sigma_H^2} = 1$ $\quad n_1 = 21 \quad n_2 = 31$ $\qquad F = \frac{S_1^2/S_2^2}{\cancel{r_0}\ 1}$ con $\nu_i = n_i - 1$

$v_1 = 20 \quad v_1 = 30$

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

$H_1: \frac{\sigma_m^2}{\sigma_H^2} \neq 1 \qquad \frac{\alpha}{2} = 0,005$

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

$$f_{obs} = \frac{2,18^2}{1,92^2} = 1,2892$$

$f_{c1} = f_{0,005,\,20,\,30} = 0,32016$

$f_{c2} = f_{0,995,\,20,\,30} = 2,82304$

0 $\qquad\qquad$ $+\infty$

RR | RA | RR

Izq $\quad f_{c1}$ $\qquad\qquad f_{c2}$ Derecha

0,32016 $\qquad\qquad$ 2,82304

RR: $]0, 0.32016[\cup]2.82304, +\infty[$

RA: $]0.32016, 2.82304[$

$\boxed{\text{R/ Como } f_{c1} = 0,32016 < f_{obs} = 1,2892 < f_{c2} = 2,82304 \text{ NO se rechaza } H_0}$

Con valor P

$$F = \frac{S_1^2/S_2^2}{\cancel{\sigma_0} 1} \text{ con } \nu_i = n_i - 1$$

$$F_{obs} = \frac{2{,}18^2}{1{,}92^2} = 1{,}2892 \qquad \alpha = 0{,}01$$

$$\nu_1 = 20 \qquad \nu_1 = 30$$

$$2P(F > 1{,}2892)$$
$$= 2 \cdot 0{,}25858$$
$$= 0{,}51716$$

→ se mete con > en la app y se multiplica por 2 manualmente

$\boxed{\text{R/ Como } P = 0{,}51716 > 0{,}01 \text{ No se rechaza } H_0}$

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" -> $\chi^2_{0.05, v}$ con > en la app

n = cantidad de muestra

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

O_i = valores observados P_i = probabilidad

e_i = valores esperados $n \cdot P_i$

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

$\chi^2_c = \chi^2_{\text{significancia dada}, v}$

El χ^2_c se mete con > y el valor p con >

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

Ejemplo: $O_1 = 4$, $O_2 = 3$, $O_3 = 6$
$e_1 = 7$, $e_2 = 6$, $e_3 = 4$ < 5

$e_2 = 4 \to e_2 = 6 + 4 = 10$ $k = 3 \to k = 2$
$O_2 = 3 \to O_2 = 3 + 6 = 9$ $v = 2 \to v = 1$

Hay 2 casos, dan o no distribucion

Caso 1: Dan una distribucion

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

1 2 3 4 5 → $k = 5$

$v = 4$

	Accidentes	1	2	3	4	5
e	**Semanas**	10	10	7	5	4
o						

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?

$\alpha = 0{,}03$ $\quad \lambda = 2$

Recuerde $P(2)$ se refiere a una distribucion 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

$H_0: X \sim P(2)$

$H_1: X \nsim P(2)$

Paso 2: Identificar n, k, v, α

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

$\alpha = 0{,}05 \quad k = 4$

Paso 3: Calcular $n \cdot P_i$ osea $n \cdot e_i$ donde e_i es el resultado de remplazar los valores esperados en la distribucion dada

Accidentes	1	2	3	4	5

$e_1: 36 \cdot \frac{e^{-2} \cdot 2^1}{1!} = 9{,}74 \geq 5$ ✓

$e_2: 36 \cdot \frac{e^{-2} \cdot 2^2}{2!} = 9{,}74 \geq 5$ ✓

$e_3: 36 \cdot \frac{e^{-2} \cdot 2^3}{3!} = 6{,}49 \geq 5$ ✓

$e_3: 36 \cdot \frac{e^{-2} \cdot 2^3}{3!} = 6{,}49 \geq 5$ ✓

$e_4: 36 \cdot \frac{e^{-2} \cdot 2^4}{4!} = 3{,}24 < 5$

$e_5: 36 \cdot \frac{e^{-2} \cdot 2^5}{5!} = 1{,}29 < 5$

Como estos son < 5
Hay que combinar

Paso 4: Actualizar si es necesario

$e_3 = 6{,}49 \longrightarrow e_3 = 6{,}49 + 3{,}24 + 1{,}29 = 11{,}02$

$k = 5 \rightarrow k = 3$

$V = 4 \rightarrow V = 2$

$O_3 = 7 \rightarrow O_3 = 7 + 5 + 4 = 16$

Cola derecha
Si se cumple
$\chi^2_{obs} < \chi^2_c$, NO se rechaza H_0
Si se cumple
$\chi^2_{obs} > \chi^2_c$, Se rechaza H_0

Paso 5: Calcular el χ^2_{obs}

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

Observados: 10, 10, 16
esperados: 9,74, 9,74, 11,02

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

$\chi^2_{obs} = 2{,}26 \qquad \chi^2_c = \chi^2_{0.03,2} = 7{,}01312$

R/ Como $\chi^2_{obs} = 2{,}26 < \chi^2_c = 7{,}01332$
No se rechaza H_0, entonces
no existe evidencia de lo propuesto

Valor P:

$$\chi^2_{obs} = \frac{(10-9{,}79)^2}{9{,}74} + \frac{(10-9{,}79)^2}{9{,}74} + \frac{(16-11{,}02)^2}{11{,}02} = 2{,}26$$

$\chi^2_{obs} = 2{,}26$ $\alpha = 0{,}03$ $v = 2$

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

$\boxed{\text{R/ Como } P = 0{,}32303 > \alpha = 0{,}03 \text{ No se rechaza } H_0}$

Caso 2: NO dan una distribucion

Aqui se define el $p_i = \frac{1}{k}$, lo demas 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)**

	1	2	3	4	5	6
Color	Azul	Naranja	Verde	Rojo	Amarillo	Café
Cantidad	212	147	103	50	46	42

Dan eso pero no dan una formula como en el pasado

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

$H_1: X \nsim U \qquad \alpha = 0{,}05 \qquad v = 5$

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

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

esperados: 100, 100, 100, 100, 100, 100

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

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

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

$\chi^2_{obs} = 235{,}42 \qquad \chi^2_c = \chi^2_{0{,}05,5} = 11{,}07050$

$\boxed{\text{Como } \chi^2_{obs} = 235{,}42 > \chi^2_c = 11{,}07050 \text{ Se rechaza } H_0}$

$P = P(\chi^2 > \chi^2_{obs}) = P(\chi^2 > 235.42) \approx 0$, $\boxed{P = 0 < \alpha = 0{,}05 \text{, se rechaza } H_0}$