# Proporciones

Sea $p$ la proporción de determinados éxitos en una población. Dada proporción $\hat{P} = \frac{B}{n}$ para muestras de tamaño $n$ (estadístico asociado al parámetro $p$)donde $B$ es la variable que indica el número de éxitos en la muestra. Se tiene lo siguiente:

1. $B$ sigue una distribución binomial: $B \sim B(n, p)$
2. $E(\hat{P}) = p$, es decir, $\hat{p}$ es insesgado.
3. $Var(\hat{P}) = \frac{pq}{n}$, donde $q = 1 - p$

En el curso de probabilidad, se estudia el caso para la distribución binomial cuando $n$ es grande, haciendo una aproximación a la normal.

**TEOREMA:** Para $n$ suficientemente grande se cumple que $\hat{p}$ sigue una distribución Normal como sigue:

$$\hat{P} \sim N\left(p, \frac{pq}{n}\right)$$

## Caso1: Muestras grandes

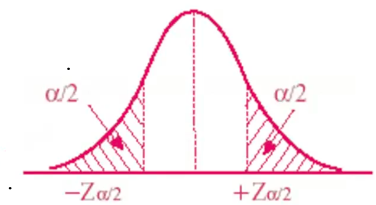
Considere una población dada:

**Proporción poblacional:** $p$

**Estadístico**: $\hat{P}$, para muestras de tamaño $n$

**Valor de $\hat{P}$ para una muestra de tamaño $n$**: $\hat{p}$

**Condición**: $n\hat{p} \geq 5 \wedge n\hat{q} \geq 5$




1. Entonces, un IC $(1-\alpha)100\%$ para $p$ sería: $\hat{p} \pm \underbrace{z_{\frac{\alpha}{2}} \cdot \left.\sqrt{\frac{\hat{p}\hat{q}}{n}}\right\}}_{\text{Radio}}$ Desviación estandart
2. Tamaño de muestra para encontrar un IC para $\mu$ con un radio menor o igual a $r$: $n \geq \left(\frac{z_{\alpha/2} \cdot \sqrt{pq}}{r}\right)^2$

$$p = \frac{\text{Los que cumplen}}{\text{Totales}}$$

$p = \dfrac{\text{Los que cumplen}}{\text{Totales}}$, Aquí $= \dfrac{18}{50}$

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 $\rightarrow \dfrac{18}{50} = 0{,}36$

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

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

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

$$\hat{P} = \frac{B}{n}$$

**Condición**: $n\hat{p} \geq 5 \wedge n\hat{q} \geq 5$

1. $B$ sigue una distribución binomial: $B \sim B(n, p)$
2. $E(\hat{P}) = p$, es decir, $\hat{p}$ es insesgado.
3. $Var(\hat{P}) = \frac{pq}{n}$, donde $q = 1 - p$

$50 \cdot 0{,}36 \geq 5$ ✓ $\wedge$ $50 \cdot 0{,}64 \geq 5$ ✓

Entonces $]a, b[$

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

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

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

$$\boxed{\text{El IC para 95\% corresponde a } ]0.22695, 0.49305[}$$

NOTA: Observe que en este caso el IC es muy grande, pues la proporción de amas de casa que no utilizan el detergente X podría ser tan pequeña que se acerca el 25% o tan grande que se acerca al 50%. Difundir estos resultados, puede que no sea conveniente debido a la falta de precisión, por lo que se requiere un IC más pequeño.

¿Cómo se logra un IC más pequeño?: ampliando el tamaño de la muestra.

¿Qué tan grande debe ser este tamaño de muestra?:

$$n \geq \left(\frac{z_{\alpha/2} \cdot \sqrt{pq}}{r}\right)^2$$

# Tamaño de la muestra

NOTA: Para el tamaño de muestra, dado que $p$ y $q$ son desconocidos, se puede realizar lo siguiente:

1. Utilizar buenas estimación es $\hat{p}$ y $\hat{q}$ de $p$ y $q$ , obtenidas previamente para un valor de $n$, es decir, que cumplan que $n\hat{p} \geq 5 \wedge n\hat{q} \geq 5$, así la estimación sería : $n \geq \left(\frac{z_{\alpha/2} \cdot \sqrt{\hat{p}\hat{q}}}{r}\right)^2$
2. Otra forma útil es tomar $pq \leq 0.25$, puesto que $pq = p - p^2$ es una parábola cóncava hacia abajo, cuyo valor máximo está en el vértice que es $(0.5, 0.25)$, así la estimación de muestra se haría para el valor que maximiza a $pq$ y como $pq$ es parte de la estimación del IC, se estaría buscando un tamaño de muestra apropiado que se ajuste al que podría generar el IC más grande. Decir $pq = 0{,}25$ basicamente

$$n \geq \left(\frac{z_{\frac{\alpha}{2}} \cdot \sqrt{0{,}25}}{r}\right)^2$$

## 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 usar $pq = 0{,}25$

Caso 1: Utilizando buenas estimaciones de $p$ y $q$

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

$\frac{\alpha}{2} = 0{,}025 \rightarrow z_{0{,}025} = \pm 1{,}95996$ $\qquad n \geq \left(\frac{z_{\alpha/2} \cdot \sqrt{\hat{p}\hat{q}}}{r}\right)^2$

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

$n \geq 2212{,}673284$

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

Caso 2: No importa valor de $p$ → $pq = 0{,}25$

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

→ $n \geq 2400{,}90$

$\boxed{n \geq 2401}$

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%

Recordar, aquí es igualar sin el $\hat{p} \pm$ al E y despejar $Z_{\frac{\alpha}{2}}$, luego ese resultado se mete en la app $Z_{resultado}$ y al final se hace $1 - Z_{resultado}$ con 2 colas en la app

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

$$Z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

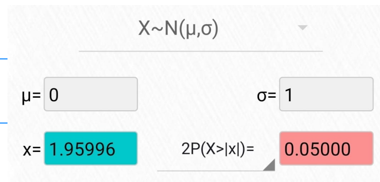
$n = 18 \quad p = \frac{12}{18} = \frac{2}{3} \quad q = \frac{1}{3}$

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

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

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

$$Z_{\frac{\alpha}{2}} = 1{,}9599975$$




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

El IC que halló el rector es muy grande, ¿de qué tamaño debe ser la muestra si se desea tener una confianza de 90% para estimar el porcentaje de estudiantes que aprobaron el examen con un error de estimación menor que 0.05? R/ Usando la estimación n=241, sin importar $p$: n=271

$n = 18$ $\quad p = \frac{12}{18} = \frac{2}{3}$ $\quad q = \frac{1}{3}$ $\quad r = 0{,}05$ $\quad n \geq \left(\frac{z_{\alpha/2} \cdot \sqrt{\hat{p}\hat{q}}}{r}\right)^2$

$\alpha = 0{,}10$

Caso 1: Usando estimacion

$\frac{\alpha}{2} = \frac{0{,}10}{2} = 0{,}05 \rightarrow Z_{0{,}05} = 1{,}64485$

$n \geq \left(\frac{1{,}64485 \cdot \sqrt{\frac{2}{3} \cdot \frac{1}{3}}}{0{,}05}\right)^2 \rightarrow n \geq 240{,}49$

$\boxed{n \geq 241}$

Caso 2: Sin importar p

$n \geq \left(\frac{1{,}64485 \cdot \sqrt{0{,}25}}{0{,}05}\right)^2 \rightarrow n \geq 270{,}55$

$\boxed{n \geq 271}$

Una muestra aleatoria de 50 estudiantes del Tec, da los siguientes resultados:

| | Mujeres | Hombres |
|---|---|---|
| De San José | 12 | 9 |
| De otras provincias | 10 | 19 |

Encuentre un IC de 95% para la proporción de estudiantes de otras provincias. R/ ]044319,0.71681[

$n = 50$ $\quad p = \frac{10+19}{50} = 0{,}58$ $\quad q = 0{,}42$ $\quad \hat{p} \pm z_{\frac{\alpha}{2}} \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}$

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

$a = 0{,}58 - 1{,}95996 \cdot \sqrt{\frac{0{,}58 \cdot 0{,}42}{50}} = 0{,}044319$

$b = 0{,}58 + 1{,}95996 \cdot \sqrt{\frac{0{,}58 \cdot 0{,}42}{50}} = 0{,}7168$

$\boxed{\text{IC } ]0.044319 \,,\, 0.7168[}$

Un profesor asegura que el promedio de calificación de los estudiantes en el primer parcial del curso de Estadística es 85 y presenta los siguientes datos (tomados de una muestra aleatoria de 16 estudiantes).

| *Calificaciones* | | | |
|---|---|---|---|
| 87 | 90 | 88 | 76 |
| 74 | 70 | 100 | 95 |
| 65 | 92 | 78 | 87 |
| 82 | 50 | 89 | 89 |

$\frac{9}{16}$

**[3 puntos]** Determine un intervalo de confianza del 98 % para la proporción poblacional de los estudiantes del curso que obtuvieron una nota superior a 85.

$n = 16 \qquad p = \frac{9}{16} \qquad q = \frac{7}{16}$

$\alpha = 0{,}02$

**Condición**: $n\hat{p} \geq 5 \wedge n\hat{q} \geq 5$

$16 \cdot \frac{9}{16} \geq 5 \checkmark \qquad 16 \cdot \frac{7}{16} \geq 5 \checkmark$

$\frac{\alpha}{2} = \frac{0{,}02}{2} = 0{,}01 \rightarrow z_{0{,}01} = \pm 2{,}32635$

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

$$a = \frac{9}{16} - 2{,}32635 \cdot \sqrt{\frac{\frac{9}{16} \cdot \frac{7}{16}}{16}} = 0{,}273987$$

$$a = \frac{9}{16} + 2{,}32635 \cdot \sqrt{\frac{\frac{9}{16} \cdot \frac{7}{16}}{16}} = 0{,}852013$$

$\boxed{\text{El IC para 98\% corresponde a } ]0.273987,\ 0.852013[}$

Para Varianzas

## Distribución Chi *Cuadrado*

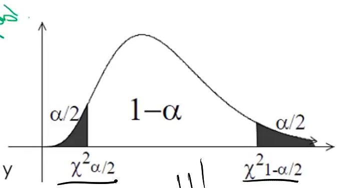
Considere una población dada por la variable aleatoria $X$ que sigue una distribución normal con varianza poblacional $\sigma^2$. Dada una muestra aleatoria de esta población $(X_1, X_2, \ldots, X_n)$, entonces la variable:

Chi Cuadrado

gl — varianza muestral — var. pobl

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

Sigue una distribución $\chi^2$ con $v = n - 1$ *grados de libertad* $(gl)$.



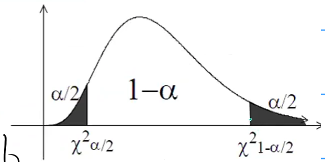

o,o!!!

La expresión $\chi^2_{\alpha/2,n-1}$ deja un área a la izquierda de $\frac{\alpha}{2}$ y $\chi^2_{1-\alpha/2,n-1}$ deja un área de a la derecha $\frac{\alpha}{2}$.

## IC para la varianza

Considere una población dada:

- **Variable aleatoria:** $X$
- **Varianza poblacional:** $\sigma^2$
- **Muestra aleatoria de esta población:** $(X_1, X_2, \ldots, X_n)$
- **Varianza de la muestra:** $S^2$




Entonces:

Varianza — a , b

**IC de $(1-\alpha)100\%$ para $\sigma^2$:** $\left]\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right[$ → $1 - \frac{\alpha}{2}$ , $n-1$ NO $\frac{1-\alpha}{2}$

Desviación estandart

**IC de $(1-\alpha)100\%$ para $\sigma$ :** $\left]\sqrt{\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}}, \sqrt{\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}}\right[$

a , b

Para calcular a y b en la app, ya sea para $\sigma^2$ y $\sigma$, siempre se mete con $\underbrace{P(X<x)}_{\text{menor que}}$

La varianza poblacional $\sigma^2$ y la varianza de la muestra $s^2$ ambas se sacan en la calcu menu → 6 → 2 → datos → opt n → 3

Si dan $S_x$, hay que elevar al cuadrado

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 [

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

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

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

$]a, b[$

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

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

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

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

Para el buen funcionamiento de una empresa tamalera, sus productos terminados deben tener un peso similar. En la tamalera **Ta' bien** se tomó una muestra para determinar si la producción se encuentra bajo control (tamales con pesos similares).
Los pesos (en gramos) obtenidos son 125, 123, 122, 124, 127, 126, 127, 130, 128, 122, 134, 129, 124, 122, 129. Con un nivel de confianza de 90 %, ¿puede sostenerse que los tamales producidos en **Ta' bien** tienen pesos similares? (8 puntos)

Según la definición de la desviación estandart, que dice que tan alejados estan los datos respecto al valor promedio, por eso hay que calcular el IC para σ

$n = 15 \quad s^2 = 11,1238$

$gl = 14 \quad \alpha = 0,10$

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

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

Para $a = 1 - 0,05 = 0,95$ $\quad \chi^2_{0.95,14} = 23,68479$

Para $b = 0,05$ $\quad \chi^2_{0.05,14} = 6,57063$

$$\left]\sqrt{\frac{14 \cdot 11,1238}{23,68479}}, \sqrt{\frac{14 \cdot 11,1238}{6,57063}}\right[$$

$$\boxed{]\ 2,677,\ 5,0825\ [}$$

Dado que la mayor diferencia media posible, con confianza de 90 %, es de poco más de 5 gramos; se concluye que los pesos de los tamales de **Ta'bien** son similares.

**[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 $I = ]\overset{a}{14.6725}\ ,\ \overset{b}{37.38}[$

En estos como la formula tiene 2 partes se debe compara la parte a de la formula con la parte a del intervalo o b con b, no importa, luego despejar el $\chi^2_{1-\frac{\alpha}{2},n-1}$ si se escoge a o $\chi^2_{\frac{\alpha}{2},n-1}$ si se escoge b, ese resultado meterlo en la app con $P(X<x)$ para cualquier caso y a ese resultado multiplicarlo por 2 y al final hacer 1 − eso, 1 − 2·resultado pues

$n = 8 \qquad S^2 = 432{,}69643$

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

Usando Formula b con intervalo b

$$\sqrt{\frac{7\cdot 432{,}69643}{\chi^2_{\frac{\alpha}{2},n-1}}} = 37{,}38$$ Paso 1: despejar

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

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

$$\chi^2_{\frac{\alpha}{2},7} = 2{,}17$$

$$\chi^2_{2{,}17,7} = 0{,}05$$ Paso 2: meter resultado en la app

$$1 - 2\cdot 0{,}05 = 0{,}90 \rightarrow \boxed{90\%}$$ Paso 3: hacer $1 - 2\cdot$resultado