Fórmulas para calcular la muestra

Fórmula que aplica en caso que no se conozca con precisión el tamaño de la población:

$$\mathbf{n} = \frac{\boldsymbol{Z}^2 \mathbf{p\ q}}{\boldsymbol{E}^2}$$

donde:

n es el tamaño de la muestra;

Z es el nivel de confianza;

p es la variabilidad positiva o probabilidad de éxito;

q es la variabilidad negativa o probabilidad de fracaso;

E es la precisión o error.

En caso que sí se conozca el tamaño de la población entonces se aplica la siguiente fórmula:

$$\mathrm{n} = \frac{\boldsymbol{Z}^2 \mathbf{pqN}}{\boldsymbol{E}^2 \boldsymbol{x}(\mathbf{N}-\mathbf{1}) + \boldsymbol{Z}^2 \boldsymbol{xPxq}}$$

donde:

n es el tamaño de la muestra;

Z es el nivel de confianza;

p es la variabilidad positiva o probabilidad de éxito;

q es la variabilidad negativa o probabilidad de fracaso;

N es el tamaño de la población;

E es la precisión o error.

Para ejemplificar lo anterior se supone el siguiente caso:

¿Cuántas familias se tendrían que estudiar para conocer la preferencia del mercado en cuanto a una marca de arroz, si se desconoce la población total?

Nivel de confianza = 95%;

Error de precisión = 3%;

Probabilidad de éxito = asumamos que puede ser próxima al 5%; si no tuviésemos ninguna idea de dicha proporción utilizaríamos el valor p = 0,5 (50%) que maximiza el tamaño de la muestra.

Entonces:

Za2 = 1,962 (ya que el nivel de confianza es del 95%)

p = proporción esperada (en este caso 5% = 0,05)

q = 1 - p (en este caso 1 – 0,05 = 0,95)

d = error de precisión (en este caso deseamos un 3%)

$$n = \frac{1{,}96^2 \text{ x } 0{,}05 \text{ x } 0{,}95}{0{,}03^2} = \mathbf{203}$$

El resultado de esta operación es 203, que es el tamaño de la muestra.

Se requeriría encuestar a no menos de 203 familias para poder tener una seguridad del 95%.

¿Cómo varía el ejemplo anterior, si se desconoce la proporción esperada?

Cuando se desconoce la proporción esperada, se tiene que utilizar el criterio conservador (p = q = 0,5), lo cual maximiza el tamaño de la muestra de la siguiente manera:

- Za2 = 1,962 (ya que la seguridad es del 95%)
- p = proporción esperada (en este caso 50%= 0,5)
- q = 1 - p (en este caso 1 – 0,5 =0,5)
- d = precisión (en este caso deseamos un 3%), quedando como resultado:

$$n = \frac{1{,}96^2 \text{ x } 0{,}5 \text{ x } 0{,}5}{0{,}03^2} = 1068$$

Se requeriría encuestar a no menos de 1068 familias para poder tener una seguridad del 95%.

El mismo ejemplo, pero cuando se conoce el tamaño de la población, daría el siguiente resultado:

¿A cuántas familias se tendría que estudiar para conocer la preferencia del mercado en cuanto a una marca de arroz, si se conoce que el número de familias es de 15.000?
Seguridad = 95%; Precisión = 3%;

Probabilidad de éxito = asumimos que puede ser próxima al 5%; si no tuviese ninguna idea de dicha proporción utilizaríamos el valor p = 0,5 (50%), que maximiza el tamaño de la muestra.

$$n = \frac{15000\, x\, 1{,}96^2 \text{ x } 0{,}05 \text{ x } 0{,}95}{0{,}03^2 \text{x}(15000-1) + 1{,}96^2 \text{x } 0.05 \text{ x } 0{,}95} = 200$$

¿Cómo hubiera cambiando el ejemplo anterior, si se desconoce la proporción esperada?

Si se desconoce la proporción esperada, se tendría que utilizar el criterio conservador (p = q = 0,5), lo cual maximiza el tamaño de muestra de la siguiente manera:

- Z α 2 = 1,962 (ya que la seguridad es del 95%)
- p = proporción esperada (en este caso 50%= 0,5)
- q = 1 – p (en este caso 1 — 0,5 = 0,5)
- d = precisión (en este caso deseamos un 3%), quedando como resultado:

$$n = \frac{15000\, x\, 1{,}96^2 \text{ x } 0{,}5 \text{ x } 0{,}5}{0{,}03^2 \text{x}(15000-1) + 1{,}96^2 \text{x } 0{,}5 \text{ x } 0{,}5} = \mathbf{997}$$

Tamaño necesario de la muestra: ¿cuántas personas necesitamos?

Tamaño de la población	Nivel de confianza α =0,05 (z = 1,96)	
	Para e = 0,05	Para e = 0,03
N = 100	n = 80	n = 92
N = 150	n = 108	n = 132
N = 200	n = 132	n = 169
N = 250	n = 152	n = 203
N = 500	n = 217	n = 341
N = 1000	n = 278	n = 516
N = 2500	n = 333	n = 748
N = 5000	n = 357	n = 879
N = 10000	n = 370	n = 964
N = 100000	n = 383	n = 1056
N = 1000000	n = 384	n = 1066
N = 2000000	n = 384	n = 1066

Cuando la población es muy pequeña y el error tolerado muy pequeño, prácticamente hay que tomar a toda o casi toda la población.

Tamaño de la población	Error tolerado	
	Para e = 0,05	Para e = 0,03
N = 32	n = 30	n = 31
N = 31	n = 29	n = 30
N = 30	n = 28	n = 29
N = 29	n = 27	n = 28
N = 28	n = 26	n = 27
N = 27	n = 25	n = 26
N = 26	n = 24	n = 25
N = 25	n = 24	n = 24

Con un error tolerado del 5% y poblaciones entre 25 y 15 personas, la muestra debe ser N-1 (podemos prescindir de una persona), y con menos de 15 personas debemos incluir a toda la población.

TABLA DE APOYO AL CÁLCULO DEL TAMAÑO DE UNA MUESTRA POR NIVELES DE CONFIANZA				
Certeza	95%	94%	93%	92%
Z	1,96	1,88	1,81	1,75
Z (al cuadrado)	3,84	3,53	3,28	3,06