

TAMAÑO DE MUESTRA

Es uno de los elementos centrales en el diseño de cualquier investigación estadística, ya que determina la precisión, confiabilidad y representatividad de las estimaciones realizadas a partir de los datos muestrales. El tamaño de la muestra debe elegirse cuidadosamente para garantizar que los resultados obtenidos sean suficientemente cercanos a los valores reales de la población, sin requerir un número innecesariamente grande de observaciones que incrementen los costos y el tiempo del estudio.

Para determinar el tamaño adecuado, se consideran factores como el nivel de confianza deseado (comúnmente 90%, 95% o 99%), el margen de error aceptable, la variabilidad de la población (representada por la desviación estándar), y el tipo de parámetro que se desea estimar (media o proporción). A mayor variabilidad o menor margen de error, mayor deberá ser la muestra. Asimismo, la fórmula cambia dependiendo de si la población es muy grande (aproximadamente infinita) o si debe aplicarse la corrección por población finita.

En el ámbito de los negocios internacionales, determinar correctamente el tamaño de muestra es crucial para estudios de mercados globales, análisis logísticos, mediciones de competitividad, evaluación de tratados comerciales y encuestas a empresas exportadoras. Un tamaño inadecuado puede conducir a conclusiones equivocadas, decisiones comerciales erróneas o diagnósticos imprecisos sobre el desempeño empresarial en el comercio exterior. Por ello, el cálculo del tamaño de muestra se utiliza como una herramienta estratégica para asegurar la validez de los estudios que sirven de base para políticas públicas, decisiones corporativas o estudios académicos en comercio internacional.

Selección del nivel de confianza

Es una decisión fundamental en el cálculo del tamaño de la muestra, ya que determina el grado de certeza con el que se desea que los resultados muestrales reflejen el verdadero valor del parámetro poblacional, representa la probabilidad de que el intervalo estimado mediante la muestra contenga el valor real del parámetro en un número muy grande de repeticiones del estudio. Los niveles de confianza más utilizados en investigación aplicada son 90%, 95% y 99%, los cuales corresponden a valores críticos de la distribución normal estándar (z) de aproximadamente 1.645, 1.96 y 2.576, respectivamente.

La selección del nivel de confianza implica un balance entre precisión y riesgo, un nivel de confianza más alto (99%) reduce el riesgo de error, pero aumenta considerablemente el tamaño de la muestra, lo que eleva los costos y el tiempo de la investigación. Por su parte un nivel de confianza más bajo (90%) requiere una muestra más pequeña, pero incrementa la posibilidad de que el parámetro poblacional quede fuera del intervalo estimado. El nivel de confianza estándar (95%) se considera un equilibrio óptimo, ya que proporciona una protección razonable frente a la incertidumbre sin demandar muestras excesivamente grandes.

En investigaciones de negocios internacionales, la selección del nivel de confianza depende del propósito del estudio, el impacto de las decisiones y la disponibilidad de recursos. Por ejemplo:

- Estudios estratégicos sobre entrada a nuevos mercados, riesgo logístico o evaluación de tratados comerciales suelen utilizar 95% o 99%, ya que las decisiones derivadas tienen consecuencias económicas significativas.
- Estudios exploratorios, encuestas rápidas o diagnósticos preliminares en ferias internacionales suelen trabajar con 90%, pues buscan rapidez y orientación inicial más que exactitud absoluta.

- Organismos internacionales como la OECD, UNCTAD y el Banco Mundial generalmente utilizan 95%, por ser el estándar que equilibra precisión, comparabilidad y eficiencia estadística.

Ejemplo: Una investigadora de Negocios Internacionales quiere estimar la proporción de empresas exportadoras de una región que utilizan plataformas digitales (marketplaces B2B, e-commerce, etcétera) para vender al extranjero.

Desea:

- Nivel de confianza: **95 %**
- Margen de error máximo: **± 4 %** ($E = 0.04$)
- Estimación preliminar de la proporción: $p = 0.60$ (es decir, se cree que el 60 % las usa)
- Entonces $q = 1 - p = 0.40$

Se utiliza la fórmula para proporciones:

$$n = \frac{z^2 p q}{E^2}$$

Con:

- $z = 1.96$ (para 95 % de confianza)
- $p = 0.60$
- $q = 0.40$
- $E = 0.04$

$$n = \frac{(1.96)^2(0.60)(0.40)}{(0.04)^2}$$

Paso a paso:

1. $1.96^2 \approx 3.8416$
2. $0.60 \times 0.40 = 0.24$
3. Numerador: $3.8416 \times 0.24 \approx 0.9220$
4. Denominador: $0.04^2 = 0.0016$
5. $n \approx \frac{0.9220}{0.0016} \approx 576.25$

redondea hacia arriba para trabajar con valores enteros, con lo que la investigadora necesitaría:

$n = 577$ empresas exportadoras.

Si se hiciera el mismo estudio con los mismos datos pero distinto nivel de confianza, el tamaño de muestra cambiaría (mismo $p = 0.60$, $E = 0.04$):

- Con **90 %** ($z \approx 1.645$) $\rightarrow n \approx 406$ empresas
- Con **95 %** ($z \approx 1.96$) $\rightarrow n \approx 577$ empresas
- Con **99 %** ($z \approx 2.576$) $\rightarrow n \approx 996$ empresas

Conclusión:

Si se desea mayor seguridad (**99 %**), paga el precio de casi mil encuestas.

Si se conforma con **90 %**, puede trabajar con unas 400 empresas que quizá resulten ser pocas para el estudio.

Margen de error (E)

Es la cantidad máxima en la que se permite que la estimación muestral (como una proporción o una media) difiera del valor real de la población. Representa la “tolerancia” del estudio, es decir, la precisión deseada en los resultados. Cuanto más pequeño sea el margen de error, más precisa será la estimación, pero también más grande deberá ser la muestra.

El margen de error se calcula de manera diferente según el tipo de parámetro que se quiera estimar:

Para una proporción

$$E = z \sqrt{\frac{pq}{n}}$$

Donde:

- z = valor crítico de la normal (de acuerdo con el nivel de confianza).
- p = proporción estimada.
- $q = 1 - p$
- n = tamaño de la muestra.

Para una media

$$E = z \frac{\sigma}{\sqrt{n}}$$

Donde:

- z = valor crítico de la normal (de acuerdo con el nivel de confianza).
- σ = desviación estándar de la población (o estimación).
- n = tamaño de la muestra.

Dependiendo del objetivo del estudio, del costo y del nivel de riesgo que se puede asumir, existen diferentes recomendaciones para elegir el nivel de error:

- $E = 5\%$ (0.05) → Estándar más común, se utiliza en estudios de mercado, comercio exterior, logística, educación, salud, etcétera.
- $E = 3\%$ (0.03) → Más preciso, recomendable cuando las decisiones tienen consecuencias económicas importantes.
- $E = 1-2\%$ → Muy exigente; usado solo cuando los recursos son amplios o la estimación debe ser extremadamente precisa (por ejemplo, encuestas nacionales o estudios regulatorios).

Es importante además seleccionar el parámetro con el que se desea trabajar para el nivel de error, esto porque puede entregar valores diferentes y en ocasiones no suficientes para el estudio en cuestión. Existen algunos criterios o recomendaciones para seleccionar trabajar con la media o bien con una proporción.

Para la proporción (p, %):

Cuando la variable es **dicotómica**, es decir, solo tiene dos posibles resultados, por ejemplo:

- Exporta / no exporta.
- Cumple / no cumple con un tratado.
- Tiene certificación / no la tiene.

Esta forma es fácil de interpretar y es adecuada para estudios de comercio exterior y logística.

Para la media (μ):

Cuando la variable es **numérica continua**, y se piden resultados como:

- Tiempo promedio de exportación.
- Costos logísticos.
- Valor monetario de exportaciones.
- Kilogramos transportados.

Esta forma es útil para análisis financieros, tiempos de entrega, rendimiento logístico, precios, etcétera.

Ejemplo aplicado

Se desea estimar la proporción de empresas exportadoras de un país que utilizan plataformas digitales para vender al extranjero.

Desea:

- Nivel de confianza $\rightarrow 95\%$
- Margen de error $\rightarrow \pm 5\%$ ($E = 0.05$)
- Proporción estimada (basada en estudios previos) $\rightarrow p = 0.55$
- Entonces: $q = 0.45$

Utilizando la fórmula para tamaño de muestra en proporciones:

$$n = \frac{z^2 pq}{E^2}$$

Sustituyendo:

$$z = 1.96$$

$$p = 0.55, q = 0.45$$

$$E = 0.05$$

$$n = \frac{(1.96)^2(0.55)(0.45)}{0.05^2}$$

Se realizan los cálculos correspondientes para cada término de la fórmula:

$$1.96^2 = 3.8416$$

$$0.55 \times 0.45 = 0.2475$$

$$\text{Numerador: } 3.8416 \times 0.2475 = 0.9503$$

$$\text{Denominador: } 0.0025$$

Finalmente se sustituyen en la fórmula para la proporción los valores obtenidos en el paso anterior.

$$n = \frac{0.9503}{0.0025} = 380.12$$

Tamaño mínimo de muestra: 381 empresas exportadoras.

Interpretación:

- Con 95% de confianza y un margen de error del 5%, necesitará encuestar 381 empresas para obtener una estimación confiable.
- Si quisiera un margen de error del 3%, la muestra subiría a 1,055 empresas, con lo que se demuestra cómo el margen de error controla el esfuerzo de investigación.

Referencia:

Lind, D. A., Marchal, W. G., & Wathen, S. A. (2021). Statistical Techniques in Business & Economics (18th ed.). McGraw-Hill Education.