

Tamaño de la muestra:
Ejemplo 1
(Variable CUANTITATIVA
-población INFINITA)

$$n = \frac{Z^2 \times S^2}{E^2}$$

Caso 7.1 Introducción de una nueva tarjeta de crédito- Banco JTRQP

Usted(es) ha(n) sido contratado(a)s por una reconocida entidad financiera (Banco JTRQP) con casa matriz en Estados Unidos, que quiere iniciar operaciones de banca en Colombia con motivo de la firma del TLC. El respectivo banco quiere realizar un estudio de mercados que le permita evaluar el gasto promedio mensual (en \$) que actualmente realizan las personas (hombres y mujeres con edades entre 25 y 40 años) por utilización de tarjeta de crédito. Su función es apoyar al gerente de Marketing para estimar el tamaño de muestra representativo con el fin de realizar el respectivo estudio que tendrá lugar en la ciudad de BGTX. Por favor, estime(n) el tamaño de la respectiva muestra representativa para el grupo de población objetivo por encuestar.

Procedimiento

1. Se identifica si la población es infinita o finita (para el caso se estimará la muestra en las dos condiciones) y se define la variable objeto de medición (X_i) (en este caso, esa variable se denomina “Gasto promedio mensual por uso de tarjeta de crédito”), la cual es una variable cuantitativa y se define la fórmula del muestreo que se va a utilizar.
2. Se determina el valor de la desviación estándar, para el caso, valor que se obtendrá mediante muestra piloto: se mide la variable objeto del estudio en por lo menos 30 personas con las características de la respectiva población sobre la cual se realizará el estudio. En este caso, para estimar la desviación estándar se ha decidido realizar una muestra piloto a 57 personas con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito. A continuación, se muestra la tabla de frecuencias que se obtuvo de esa muestra piloto.

Tamaño de la muestra:
Ejemplo 1
(Variable CUANTITATIVA -población INFINITA)

TABLA 7.4 Gasto promedio mensual por uso de tarjeta de crédito en personas con edades entre 25 y 40 años en la ciudad de BGTX

No. de encuestas	x_i = gasto promedio/ mes (miles de \$)	No. de encuestas	x_i = gasto promedio/ mes (miles de \$)
1	200	30	1 300
2	450	31	900
3	370	32	1 200
4	1 000	33	350
5	800	34	250
6	400	35	1 100
7	300	36	700
8	250	37	500
9	900	38	700
10	1 100	39	800
11	750	40	500
12	500	41	300
13	300	42	450
14	500	43	800
15	800	44	700
16	750	45	300
17	350	46	250
18	1 200	47	300
19	250	48	750
20	600	49	900
21	750	50	600
22	900	51	1 200
23	450	52	900
24	1 100	53	1 100
25	650	54	1 250
26	350	55	900
27	500	56	700
28	900	57	300
29	450	Total: $n_0 = 57$	$\bar{X} = \sum x_i / n_0 = 663,51$

Media: f(x) MEDIA.ACOTADA
matriz: rango
porcentaje: 0

Desviación f(x) DESVEST.M
Estándar: número 1: rango

3. Se estima el valor de la respectiva desviación estándar mediante la siguiente fórmula:

$$S = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / n_0 - 1} \implies S = \$312,12$$

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

Donde:
 S : desviación estándar de la muestra piloto de personas con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito.
 X_i : gasto promedio de cada una de las personas encuestadas en la muestra piloto.
 $\bar{X}$: gasto promedio del conjunto de las personas encuestadas de la muestra piloto.
 n_0 : número de personas entrevistadas en la muestra piloto.

Homogeneidad o heterogeneidad de la muestra = $\frac{S}{\bar{X}} \implies \frac{312,12}{663,51} > 30\%$

En este caso, debido a que el valor de la relación $S/\bar{X}$ es superior al 30%, se considera que la población tiende a ser heterogénea y el valor que se asigne a E (error de estimación) se sugiere que debe ser pequeño, en lo posible menor que 5% (el valor máximo de E es 10%).

4. Estimación del tamaño de la muestra

a. Población infinita

En este caso, se supone que **no** se conoce el número total de personas (N) con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito; entonces, para estimar el tamaño de la muestra se utiliza la fórmula de muestreo aleatorio simple para población infinita, como se muestra a continuación:

$$n = \frac{Z^2 \times S^2}{E^2}$$

Donde:
 n : tamaño necesario de la muestra de personas con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito.
 Z : nivel de confianza o margen de confiabilidad (para este caso: 96% de confianza, $Z = 2,054$).
 S : desviación estándar de la población (estimada mediante una muestra piloto aplicada a 57 personas con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito) $\implies S = 312,12$.

$$\bar{X} = \frac{\sum X}{n}$$

Tamaño de la muestra:
Ejemplo 1
(Variable CUANTITATIVA -población INFINITA)

E : error de estimación (en este caso, se tomó un error de estimación pequeño debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a gasto a través del uso de tarjeta de crédito. Entonces, $E = 0,03$ del promedio del gasto en el respectivo uso de tarjeta).

$$n = \frac{Z^2 \times S^2}{E^2}, \Rightarrow n = \frac{(2,054)^2 \times (312,12)^2}{(0,03 \times 663,51)^2} \Rightarrow n = 1\,038 \text{ personas}$$

En conclusión, utilizando un nivel de confianza del 96% y un error de estimación del 3% se requiere entrevistar a por lo menos 1 038 personas de entre 25 y 40 años, residentes en la ciudad de *BGTX*, usuarios de tarjeta de crédito, para conocer el gasto promedio mensual en tarjeta de crédito de las personas en el respectivo rango de edad.

Tamaño de la muestra:

Ejemplo 2

(Variable CUANTITATIVA -población FINITA)

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

b. Tamaño de muestra cuando la población es finita (se conoce N)

En el caso que se conozca el total de la población, es decir, el número total de personas con edades entre 25 y 40 años, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito, para estimar el tamaño de la muestra se utiliza la fórmula de muestreo aleatorio simple, como se muestra a continuación:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Donde:

N : número total de personas entre 25 y 40 años de edad, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito. En este caso se supone un número total de 780 000 personas con el respectivo perfil.

n : tamaño necesario de la muestra de personas de entre 25 y 40 años de edad, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito.

Z : nivel de confianza o margen de confiabilidad (en este caso: 96% de confianza, $Z = 2,054$).

S : desviación estándar de la población (estimada mediante una muestra piloto aplicada a 57 personas de entre 25 y 40 años de edad, residentes en la ciudad de BGTX y usuarios de tarjeta de crédito) ==> $S = 312,12$.

E : error de estimación (en este caso, se tomó un error de estimación pequeño debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a gasto a través del uso de tarjeta de crédito: $E = 0,03$ del promedio del gasto en el respectivo uso de tarjeta).

El procedimiento para la estimación del respectivo tamaño de muestra es el mismo realizado para estimar tamaño de muestra para la población infinita, pero la fórmula es diferente y como sigue:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} \quad n = \frac{(312,12)^2}{\frac{(0,03 \times 663,51)^2}{(2,054)^2} + \frac{(312,12)^2}{780\,000}} \quad n = 1\,036 \text{ personas}$$

En conclusión, utilizando un nivel de confianza del 96% y un error de estimación del 3% para una población de 780 000 personas con edades de entre 25 y 40 años, residentes en la ciudad de BGTX, usuarios de tarjeta de crédito, se requiere entrevistar a por lo menos 1 036 personas para conocer el gasto promedio mensual en tarjeta de crédito por parte de las personas en el respectivo rango de edad en la mencionada ciudad.

Tamaño de la muestra:
Ejemplo 3
(Variable **CUALITATIVA**-población **FINITA**)

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Caso 7.2 Percepción sobre la gestión del conocimiento

Suponga que usted tiene interés en realizar un estudio cuyo objetivo es conocer la percepción que tienen las personas que laboran en una reconocida compañía del sector farmacéutico en esta ciudad, sobre la forma como en ella se gestiona el conocimiento. Su función es estimar el tamaño de muestra representativa de trabajadores (directivos y no directivos) de la respectiva compañía a quienes para tal propósito se les aplicará una encuesta conformada por 58 enunciados estructurados en escala Likert, donde el encuestado califica cada enunciado según su percepción, en puntuaciones de 1 a 5, siendo:

1 = Pésima; 2 = Mala; 3 = Regular; 4= Buena; 5 = Excelente.

Procedimiento

1. Se identifica si la población es infinita o finita. Para el caso, se puede saber el número total de personas que trabajan en la compañía en la ciudad, y por ello se dirá que es una población finita. La variable objeto de medición (X_p) –en este caso se denomina “Percepción sobre la forma de gestionar el conocimiento”– es una variable cualitativa y es un criterio clave para definir el error de estimación”.
2. Se determina el valor de la desviación estándar, para el caso, valor que se obtendrá mediante muestra piloto: se mide la variable objeto del estudio en por lo menos 30 personas con las características de la respectiva población sobre la cual se realizará el estudio. En este caso, para estimar la desviación estándar se aplica la encuesta a una muestra piloto de 40 personas que laboren en la compañía. El siguiente es el respectivo cuadro de frecuencias de la muestra piloto sobre la percepción de las personas encuestadas sobre la forma en que se gestiona el conocimiento en la compañía. En este caso se utilizó la distribución unidimensional, pero el CD que acompaña a este libro ilustra un ejemplo más completo.

Tamaño de la muestra:

Ejemplo 3

(Variable CUALITATIVA - población FINITA)

TABLA 7.5 Percepción sobre gestión del conocimiento. Compañía farmacéutica

No. de encuestas	X _i Percepción	No. de encuestas	X _i Percepción
1	3	24	2
2	2	25	1
3	5	26	1
4	3	27	2
5	5	28	1
6	1	29	4
7	2	30	4
8	5	31	4
9	4	32	3
10	3	33	2
11	2	34	2
12	2	35	5
13	2	36	4
14	1	37	4
15	3	38	4
16	2	39	3
17	5	40	3
18	4	41	4
19	4	42	4
20	3	43	1
21	2	44	2
22	2	45	2
23	5	Total n ₀ = 45	$\bar{X} = \sum x_i / n_0 = 2,92$

La fórmula para la estimación de la desviación estándar es la siguiente:

$$S = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / n_0 - 1} \implies S = 1,29$$

Donde:

S : desviación estándar de la percepción sobre la forma como se gestiona el conocimiento en la compañía por las personas de la muestra piloto.

X_i : percepción de cada una de las personas encuestadas en la muestra piloto sobre la forma como se gestiona el conocimiento en la compañía.

$\bar{X}$: percepción promedio del conjunto de las personas encuestadas de la muestra piloto sobre la forma como se gestiona el conocimiento en la compañía.

n₀ : número de personas entrevistadas en la muestra piloto.

$$\text{Homogeneidad o heterogeneidad de la muestra} = \frac{S}{\bar{X}} \implies \frac{1,29}{2,92} > 30\%$$

En este caso, debido a que el valor de la relación $S/\bar{X}$ es superior al 30%, se considera que la población tiende a ser heterogénea y el valor que se asigne a E (error de estimación) se sugiere que debe ser pequeño, en lo posible menor que 5% (el valor máximo de E es 10%).

- Se estima el tamaño representativo de la muestra. Se sabe que la compañía tiene una nómina de 7 584 trabajadores, es decir, este es el número total de personas que laboran en la compañía; por tanto, es una población finita y la fórmula de muestreo aleatorio simple, que se debe utilizar para estimar el tamaño de muestra representativa de personas que laboran en la compañía por encuestar para conocer la percepción sobre la forma como se gestiona el conocimiento, se presenta a continuación:

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}}$$

Donde:

N : número total de personas que laboran en la compañía (7 584).

n : tamaño necesario de la muestra representativa de personas que se debe encuestar para determinar la verdadera percepción que tienen las personas en la compañía sobre la forma como se gestiona el conocimiento en la misma.

Z : nivel de confianza o margen de confiabilidad (para este caso a las directivas de la compañía les interesa ser rigurosas y, por ello, deciden utilizar un nivel de confianza alto: 95% de confianza, para un valor de Z = 1,96).

S : desviación estándar de la población (estimada mediante la muestra piloto aplicada a 45 personas $\implies S = 1,29$)

E : error de estimación (en este caso se tomó un error de estimación un valor pequeño, debido a que la población objeto del estudio tiende a ser heterogénea en cuanto a la percepción de la variable objeto de medición; entonces: E = 0,05*).

$$n = \frac{S^2}{\frac{E^2}{Z^2} + \frac{S^2}{N}} = \frac{(1,29)^2}{\frac{(0,05)^2}{(1,96)^2} + \frac{(1,29)^2}{7\,584}} = 1\,912 \text{ personas}$$

Tamaño de la muestra:
Ejemplo 4
(proporciones – población INFINITA)

$$n = \frac{Z^2 \times P \times Q}{E^2}$$

Caso 7.4 Mujeres cabeza de familia cotizantes a salud

Una organización encargada de la protección social en la ciudad está interesada en estimar el tamaño de muestra representativo que le permita conocer la proporción de personas mujeres cabeza de familia de estrato 1 y 2, que son cotizantes para los servicios de salud de los miembros de su familia.

Procedimiento para estimar el tamaño de la muestra representativa para una población infinita

El siguiente es el procedimiento para la estimación de la respectiva muestra cuando la población es infinita y no se tiene referencia de una proporción de mujeres cabeza de familia cotizantes de servicios de salud para los miembros de su familia en su ciudad.

1. Se construye la tabla de distribución de frecuencias para las mujeres cabeza de familia de los estratos 1 y 2, que son cotizantes para los servicios de salud para los miembros de su familia en la ciudad. Para ello, se realizará una muestra piloto de mujeres de la población objeto del estudio para saber la proporción de mujeres cabeza de familia que cotizan servicios de salud. Después se les pregunta a las mujeres seleccionadas si cotizan (1) o no (2) como cabeza de familia.

Tamaño de la muestra:

Ejemplo 4

(proporciones – población INFINITA)

TABLA 7.6

Proporción de mujeres cabeza de familia en los estratos 1 y 2 que cotizan servicio de salud para los miembros de su familia

No. de encuestas	Cotiza salud	No. de encuestas	Cotiza salud
1	1	32	2
2	1	33	1
3	2	36	2
4	1	35	1
5	2	36	1
6	1	37	1
7	1	38	1
8	1	39	1
9	1	40	1
10	2	41	2
11	2	42	2
12	1	43	2
13	1	44	1
14	1	45	1
15	2	46	1
16	2	47	1
17	1	48	1
18	2	49	1
19	1	50	1
20	1	51	1
21	1	52	2
22	1	53	1
23	1	54	1
24	1	55	1
25	2	56	2
26	2	57	2
27	2	58	2
28	2	59	1
29	2	60	1
30	2	61	1
31	2	62	1
63	2	70	1
64	2	71	1
65	1	72	1
66	1	73	2
67	2	74	2
68	2	75	1
69	2		
		Total No. = 75	P = $\frac{\text{Total enc. con 1}}{\text{Total de encuestas}}$ P = 50/ 75 = 0,666

2. Se recurre a la fórmula para estimar el tamaño de muestra.

$$n = \frac{Z^2 \times P \times Q}{E^2}$$

Donde:

n : tamaño de muestra por estimar (número de mujeres cabeza de familia de estrato 1 y 2 que es necesario encuestar para conocer la proporción representativa de mujeres del estrato 1 y 2 que cotizan a salud como cabeza de familia).

Z : nivel de confianza o margen de confiabilidad (97%, es decir, *Z* = 2,17).

P : proporción de mujeres del estrato 1 y 2 de la muestra piloto que cotizan a salud como cabeza de familia.

Q = 1- *P* : proporción de mujeres del estrato 1 y 2 de la muestra piloto que no cotizan a salud como cabeza de familia.

E : error de estimación (diferencia máxima entre la proporción muestral y la proporción poblacional que el equipo investigador está dispuesto a aceptar en función del nivel de confianza definido para el estudio). En este caso, *E* = 0,03 ó 4%. Entonces 3%.

$$n = \frac{Z^2 \times P \times Q}{E^2} \quad n = \frac{2,17^2 \times 0,666 \times 0,334}{0,03^2} \quad n = 1\,164$$

En conclusión, con un nivel de confianza del 97%, un error de estimación del 3% es necesario encuestar un total de 1 164 mujeres de estrato 1 y 2 cotizantes a salud para conocer la proporción representativa de mujeres de los respectivos estratos, en la ciudad que cotizan a salud como cabeza de familia.

Tamaño de la muestra:

Ejemplo 5

(proporciones – población FINITA)

$$n = \frac{Z^2 \times P \times Q \times N}{E^2(N-1) + Z^2 \times P \times Q}$$

Procedimiento para estimar el tamaño de la muestra representativa para una población finita

El procedimiento para estimar el tamaño de muestra representativo es igual que para las poblaciones infinitas, pero la fórmula es diferente, así:

$$n = \frac{Z^2 \times P \times Q \times N}{E^2(N-1) + Z^2 \times P \times Q}$$

Donde:

n : tamaño de muestra por estimar (número de mujeres cabeza de familia de estrato 1 y 2 que es necesario encuestar para conocer la proporción representativa de mujeres del estrato 1 y 2 que cotizan a salud como cabeza de familia).

Z : nivel de confianza o margen de confiabilidad (97%, es decir, $Z = 2,17$).

P : proporción de mujeres del estrato 1 y 2 de la muestra piloto que cotizan a salud como cabeza de familia.

$Q = 1 - P$: proporción de mujeres del estrato 1 y 2 de la muestra piloto que no cotizan a salud como cabeza de familia.

N : número total de mujeres del estrato 1 y 2 que cotizan a salud en la ciudad. 580.000

E : error de estimación (diferencia máxima entre la proporción muestral y la proporción proporcional que el equipo investigador está dispuesto aceptar en función del nivel de confianza definido para el estudio). En este caso $E = 0,03$ ó 3%.

$$n = \frac{Z^2 \times P \times Q \times N}{E^2(N-1) + Z^2 \times P \times Q} \quad n = \frac{2,17^2 \times 0,666 \times 0,334 \times 580\,000}{0,03^2(580\,000 - 1) + 2,17^2 \times 0,666 \times 0,334}$$

$n = 1\,162$ mujeres de los estratos 1 y 2 que cotizan a salud en la respectiva ciudad.

En conclusión, para una población de 580 000 mujeres de los estratos 1 y 2, con un nivel de confianza del 97% y un error de estimación del 3%, es necesario encuestar un total de 1 162 mujeres de estrato 1 y 2 cotizantes a salud para conocer la proporción representativa de mujeres de los respectivos estratos en la ciudad que cotizan a salud como cabeza de familia.