Documentos Educativos

Probability distributions III: Normal distribution and determination of sample size (sampling by proportions).

Distribuciones de probabilidad III: Distribución normal y determinación de tamaño de muestra (muestreo por proporciones).

Juan Luis Soto Espinosa [1] [0000-0003-2600-9292](https://orcid.org/0000-0003-2600-9292).

[1] *FES Zaragoza, UNAM*

Dirección (autor principal): Batalla de 5 de Mayo esq. Fuerte de Loreto, Col Ejército de Oriente, Alcaldía Iztapalapa, C.P.09230, Ciudad de México.
Correo electrónico de contacto: *soej@unam.mx*

DE

Fecha de envío: 27-abril-2025
Fecha de aprobación: 07-mayo-2025

Abstract

The STANDARD NORMAL DISTRIBUTION is a probability distribution with broad applications in different disciplines such as biology, medicine, economics, sociology, demography, and psychology, due to its characteristics that reproduce the behavior of different variables in the natural environment. This property of the distribution makes it an extremely important tool for generating and understanding models of variables and making decisions with a percentage of confidence while considering a generally very low percentage of maximum permissible error. One of the most commonly used applications in disciplines that conduct studies on populations is the determination of sample size, ensuring that this size best reflects the behavior of the study variable within the target population.

There are different procedures to determine the sample size based on the use of the normal distribution. In this delivery, we will review the procedure for determining sample size based on known proportions, so that by knowing the proportion of the population that possesses the characteristic under study, we can proceed. A proportion is a simple statistic that allows us to relate two variables to each other and represents the ratio (quotient) of a subset of elements from the population divided by the total number of elements, expressed as a fraction, percentage, or proportion. By defining a confidence percentage and a maximum permissible error, it is possible to determine the number of participants, samples, or elements that must be considered in the sample for it to be representative with the required degree of confidence.

Keywords: Standard Normal Distribution, Z Value, Proportion, Population, Sample, Representative Sample.

Resumen

La DISTRIBUCIÓN NORMAL ESTÁNDAR es una distribución de probabilidad con amplias aplicaciones en diferentes disciplinas como la biología, medicina, economía, sociología, demografía, psicología, debido a que sus características reproducen el comportamiento de diferentes variables en el entorno natural.

Esta propiedad de la distribución la convierte en una herramienta de suma importancia para generar y entender modelos de variables y tomar decisiones con un porcentaje de confianza y considerando un porcentaje, generalmente muy bajo, de error máximo permisible. Uno de los usos más utilizados en disciplinas que realizan estudios sobre poblaciones es la determinación del tamaño de muestra y que este tamaño refleje lo mejor posible el comportamiento de la variable de estudio dentro de la población objetivo

Existen diferentes procedimientos para determinar el tamaño de muestra a partir del uso de la distribución normal, en esta entrega se revisará el procedimiento de determinación de tamaño de muestra basado en proporciones conocidas, de forma que conociendo la proporción de la población que posee la característica de estudio. Una proporción es un estadístico simple que permite asociar dos variables entre sí y representa la razón (cociente) entre un subconjunto de elementos de la población dividido entre el total de elementos, expresada como una fracción, porcentaje o proporción. Definiendo un porcentaje de confianza y un error máximo permisible es posible determinar el numero de participantes, ejemplares o elementos que deben considerarse en la muestra para que esta sea representativa con el grado de confianza requerido.

Palabras clave: Distribución Normal Estándar, Valor Z, Proporción, Población, Muestra, Muestra representativa

Los estudios que permiten analizar el comportamiento de diferentes variables dentro de un conjunto de elementos, individuos o procesos con el fin de entender su dinámica, optimizar su rendimiento o disminuir los efectos nocivos que pudieran tener son muy comunes en el quehacer científico cotidiano.

La necesidad de realizar investigaciones sobre grandes conjuntos de elementos se presenta en múltiples disciplinas y campos de conocimiento como biología, psicología, sociología, antropología, medicina, procesos productivos, agricultura, control de calidad, entre muchos, muchos otros ejemplos.

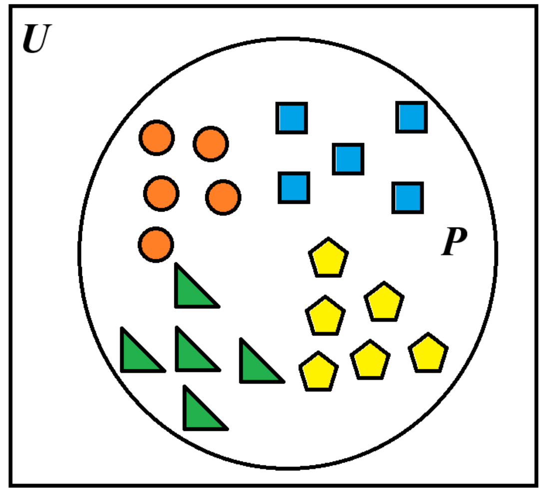
Las investigaciones en el mundo real se realizan partiendo del Universo. Universo es un concepto estadístico que podemos entender como la totalidad de elementos que existen y que presentan la o las características de estudio.

Del universo se selecciona una población, a población es un subconjunto del universo que está conformado por elementos que presentan las características de interés en el estudio en un tiempo y espacio determinados y que son accesibles para que el investigador tome datos de ellos a partir de los instrumentos de medición o colecta de información adecuados.

Es posible que una población esté conformada por todos los elementos del Universo, en cuyo caso los dos términos serían equivalentes, ejemplo:

Universo = Figuras geométricas

Población = Figuras geométricas con color



U = P

Dónde:

U = Universo

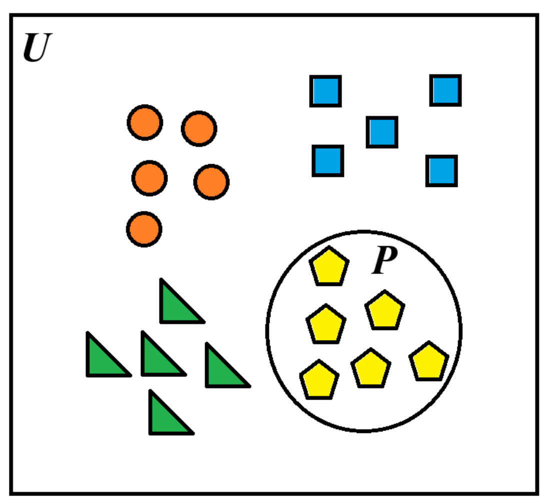
P = Población

En este caso, todos los elementos del universo (U) están dentro de la población (P).

Sin embargo, en muchos casos el universo está conformado por un número mucho mayor de elementos que la población, en estos casos, la población es un subconjunto del universo, como se muestra a continuación:

Universo = Figuras geométricas

Población = Pentágonos amarillos



$$\mathbf{U \neq P}$$

$$\mathbf{P \subset U}$$

$$\mathbf{P \in U}$$

Dónde:

U = Universo

P = Población

En este ejemplo, la población no contiene todos los elementos del universo, por lo que Universo y población son diferentes (U ≠ P), la población es un subconjunto del universo (P ⊂ U) y la población pertenece al universo (P ∈ U).

De acuerdo a la cantidad de elementos que la conforman una población puede ser finita o infinita.

Cuando es posible contabilizar e identificar cada uno de los elementos que forman la población, asignando algún identificador (etiqueta que permita diferenciar elementos), y se conoce el valor que adquiere la característica de interés (variable) en este el elemento, se tiene una Población Finita donde el tamaño de la población (N) corresponde al número de elementos que la forman.

Cuando el tamaño de la población es extremadamente grande, no es posible contabilizar e identificar cada uno de sus elementos o se desconoce el número exacto de elementos que la conforman, se considera a este conjunto de elementos una Población Infinita.

Cuando se requiere llevar a cabo una investigación sobre una población que está conformada por un gran número de individuos, con una amplia distribución geográfica y con condiciones de acceso variables, donde algunos elementos son de más fácil acceso que otros, resulta complicado acceder a todos y cada uno de los miembros de la población y pueden presentarse algunos de los siguientes problemas:

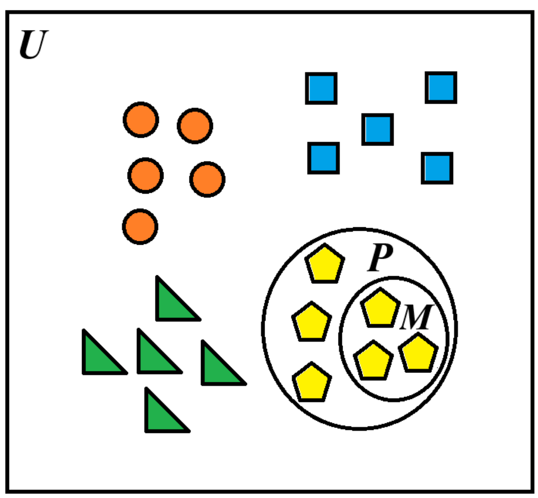
1) El costo para realizar un levantamiento en el total de la población suele ser muy elevado.
2) El tiempo invertido en el proceso de colecta de información es largo.
3) Se requiere una gran cantidad de participantes y recursos para realizar el levantamiento de información.
4) El volumen de datos puede ser muy grande, lo que dificulta su procesamiento.
5) El momento de presentar los resultados puede estar cronológicamente alejado del inicio del estudio, por lo que las condiciones pueden haber cambiado durante los diferentes periodos del proceso, por lo que las conclusiones y recomendaciones pueden ser poco pertinentes.

Para realizar este tipo de estudios, los investigadores suelen trabajar con subconjuntos de la población, los cuales deben reproducir lo más fielmente posible las características de la población que se desea analizar, a este subconjunto de la población se le llama **Muestra.** La muestra cumple con lo siguiente:

Universo = Figuras geométricas

Población = Pentágonos amarillos

Muestra = Pentágonos amarillos (accesibles)



$$\mathbf{M \neq P}$$

$$\mathbf{M \subset P}$$

$$\mathbf{M \in P}$$

Donde:

P = Población

M = Muestra

Una muestra no contiene NECESARIAMENTE a todos los elementos de la población, por lo que Población y muestra son diferentes ($P \neq M$), la muestra es un subconjunto de la población ($M \subset P$) y la muestra pertenece a la población ($M \in P$).

En estudios que se realizan sobre todos los miembros de la población no existe una muestra como tal ya que se consideran la totalidad de elementos, cuando se aplica una técnica de muestreo siempre hay algún factor limitante que impide tomar información de todos los elementos, puede ser la dificultad espacio temporal de acceder a todos, el gran número de elementos que conforman la población, la cantidad de recursos necesarios o la gran cantidad de tiempo requerido.

Al conjunto de técnicas y herramientas utilizadas para elegir los elementos del subconjunto de la población sobre el que se realizará la colecta de información necesaria para realizar un estudio se le conoce como MUESTREO.

De acuerdo con lo anterior, las técnicas de muestreo deben asegurar que los elementos que componen la muestra representen de la manera más fiel posible las variables de estudio de la población objeto de análisis, esto es, deben compartir las características esenciales, tanto en términos de similitudes como de diferencias y reflejar lo mejor posible la variabilidad que existe en la población.

Cuando la escala de medición de las variables de estudio es cuantitativa, es posible calcular el tamaño adecuado de una muestra, considerando algunos estadísticos de la población de estudio.

Existen dos tipos de muestreo: el muestreo probabilístico y el muestreo no probabilístico.

Muestreo probabilístico

Este tipo de muestreo parte de una selección aleatoria y asegura que todos los elementos integrantes de una población deben tener la misma probabilidad de ser elegidos e integrados en la muestra; de la misma manera, debe garantizar que no existan elementos con probabilidad nula de ser elegidos.

La componente de aleatoriedad en la elección de la muestra garantiza que la esta última refleje con precisión la diversidad que existe dentro de la población objetivo, además de que reduce el sesgo en el proceso de selección y permite que las inferencias hechas por los investigadores sobre los resultados obtenidos a partir de la muestra puedan ser extrapolados para obtener conclusiones acerca de la población.

Para que un muestreo probabilístico sea exitoso debe cumplir tres requisitos:

Todos los elementos que conforman la población deben tener la misma probabilidad de ser electos dentro de la muestra.

La probabilidad de que un elemento sea electo deber ser conocida de antemano, por ejemplo, en una población de 100 personas, la probabilidad de que cada persona sea seleccionada es de 1 en 100, o bien 0.01.

El muestreo debe ser aleatorio para garantizar que la muestra sea representativa de la población en su conjunto, para ello se puede recurrir a herramientas como tablas de número aleatorios, rutinas para generación automática de números aleatorios desarrolladas en Excel, Google Sheets o cualquier otra hoja de cálculo, le realización de un sorteo con ayuda de una tómbola, entre otros métodos.

Muestreo no probabilístico

En este tipo de muestreo los elementos tienen una probabilidad diferente se ser electos como parte de la muestra, esto es, hay elementos que tienen una mayor probabilidad de ser incorporados en un estudio que otros, por lo que los resultados no pueden ser generalizados a toda la población debido a que se carece de la certeza de representatividad de la muestra debido al sesgo que pueden presentar los datos.

El muestreo no probabilístico es más fácil de aplicar, es menos costoso y permite reunir los elementos de la muestra de manera más rápida, sin embargo, carece de la representatividad del muestreo probabilístico.

En estas técnicas de muestreo se elige a los elementos de la población que cumplan con los criterios que plantea el investigador, tratando, en la medida de lo posible, la muestra sea lo más representativa posible.

Determinación del tamaño de muestra

Existen diversos métodos para determinar el tamaño de muestra más adecuado para una población, en este artículo se aborda un método para el cálculo considerando las propiedades de la distribución normal, que permiten a los investigadores calcular el tamaño (número de elementos) para que una muestra refleje lo mejor posible las características de la población con cierto grado de confianza, es decir, que sea representativa

El primer paso es determinar si la población con la que se trabaja es una población finita o una población infinita, ya que la forma de calcular el tamaño de la muestra varía dependiendo de esta característica. Una vez identificada la población, se procede a determinar el grado de confianza que se requiere tener en el estudio.

La confianza estadística se define como la probabilidad, expresada con un valor entre 0 y 1, de que los elementos seleccionados reflejen las características de la población, es por ello que, durante las pruebas y herramientas estadísticas, se considera un valor cercano a 1.

Así como existe una probabilidad de contar con una muestra representativa, existe una probabilidad de contar con una muestra que presente una desviación o sesgo respecto al comportamiento de la variable poblacional, esta probabilidad, representa el margen de error que el investigador está dispuesto a asumir en la determinación del tamaño de la muestra.

La confianza y el margen de error son complementarios entre sí, por lo que cumplen con los siguientes postulados:

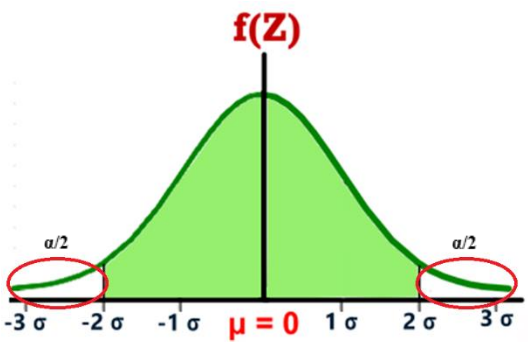
a) Consideran un evento dicotómico, donde la confianza estadística corresponde a la probabilidad de que la muestra sea significativa y su complemento corresponde a la probabilidad de que la muestra no sea significativa.
b) Tanto la confianza estadística como su complemento adquieren un valor comprendido entre 0 y 1
c) La suma de ambas probabilidades es siempre igual a uno, representando la probabilidad total.
d) Tanto la confianza como a su complemento deben ser definidos para poder realizar el cálculo de tamaño de muestra.
e) El valor de la confianza estadística siempre es un valor cercano a uno, mientras que el complemento tiene un valor cercano a cero.
f) Los valores de confianza estadística más utilizados corresponden a 90%, 95%, 99%, 99.50%, 99.90% y 99.99%; mientras que los valores de error máximo permisible más utilizados son 10%, 5%, 1%, 0.5%, 0.1% y 0.01%, respectivamente.

De acuerdo con lo anterior, si se decide que la confianza en la determinación del tamaño de muestra es del 95%, el margen de error corresponderá al 5%

Una vez definida la confianza que se ha de considerar en la determinación del tamaño de muestra, se procede a calcular o a consultar en tablas el valor Z que le corresponde.

Recuerde que el valor Z representa la posición de un punto dentro de la curva de distribución normal debajo del cual se presenta la probabilidad acumulada de ocurrencia de un valor.

La posible discrepancia de valores del tamaño de muestra puede darse en dos sentidos, que la muestra sea mayor que el valor definido o que la muestra sea menor, por lo que el margen de error es bilateral en una distribución normal. Por lo tanto, la mitad del margen de error se encuentra hacia la cola izquierda de la distribución y la otra mitad se localiza hacia la cola izquierda de la curva, como se muestra en la siguiente figura:



De acuerdo al valor de confianza estadística que se considere, los valores Z más usuales son:

Confianza	Margen de Error (α)	1-α/2	Valor Z
0.90	0.1	0.95	1.65
0.95	0.05	0.975	1.96
0.99	0.01	0.995	2.58
0.999	0.001	0.9995	3.28

Si se define un valor de confianza diferente a los expresados en la tabla, basta buscar en la tabla el valor Z que corresponde a la confianza definida. Para ello:

a) Se expresa la probabilidad en formato decimal, por ejemplo, si se define una confianza del 80%, se expresa este valor como 0.80.
b) Se busca en la tabla de determinación de valores Z el valor para 1-α/2, como α=0.20 tenemos: 1-0.20/2 = 0.90.
c) Se lee el valor que corresponde a la fila en la columna de valores zeta, a la extrema izquierda de la tabla, en este ejemplo, el valor que corresponde es 1.2
d) Se lee el valor en el encabezado de la columna donde se ubicó el valor deseado, en este ejemplo, el encabezado de la columna corresponde al valor de 0.08
e) Se suman ambos valores y el resultado es el valor Zeta buscado. En el ejemplo, tenemos que 0.08 sumado a 1.2 da un total de 1.28.

Por lo tanto, el valor Z para una confianza del 80% es igual a 1.28.

Cabe recordar que una probabilidad igual a cero corresponde a los eventos imposibles, es decir, a la ocurrencia de un valor fuera del espacio muestral de la variable; mientras que el valor 1 corresponde a la al evento seguro, es decir, la ocurrencia de un valor esperado, independientemente de las condiciones en que se realice la medición o determinación del valor de una variable.

Las probabilidades de ocurrencia de las modalidades (valores que puede adquirir una variable, consideradas dentro de su espacio muestral) son, generalmente, mayores que cero y menores o iguales a uno.

Una vez determinados el tipo de población, la confianza deseada y el margen de error, se procede a realizar el cálculo del tamaño de muestra considerando lo siguiente:

Muestreo para Proporciones conocidas

La proporción es el estadístico más simple que permite asociar dos variables entre sí y en estadística se define como la razón (cociente) entre un subconjunto de elementos de la población dividido entre el total de elementos, expresada como una fracción, porcentaje o proporción. La finalidad de las proporciones es expresar la probabilidad de ocurrencia de una característica dentro de los individuos de una población.

Considere una población de 300 trabajadores de una empresa metal mecánica dentro de la cual 45 de ellos presentan problemas de lumbalgia. Partiendo de este escenario, es posible calcular la proporción de trabajadores que presentan lumbalgia a partir del conocimiento del tamaño de la población y el número de trabajadores con lumbalgia. La proporción en este caso representa la probabilidad de ocurrencia de una característica dentro de la población; en otras palabras, representa la probabilidad de que si tomamos un individuo al azar de esta población este presente el problema de lumbalgia.

La probabilidad de ocurrencia de la característica (variable) de estudio en la población puede ser determinada a partir de registros o conteos realizados con anterioridad, utilizando para su cálculo, la fórmula de cálculo de probabilidad a priori dado por:

$$p = \frac{No\ de\ casos\ favorables}{No.\ de\ casos\ totales}$$

Mientras que q se calcula:

$$q = 1 - p$$

En el ejemplo propuesto, se tiene que:

Casos favorables = 45 (trabajadores con la característica que se desea estudiar)

Casos totales = 300 (tamaño de la población)

Sustituyendo valores en la fórmula, tenemos:

$$p = \frac{No\ de\ casos\ favorables}{No.\ de\ casos\ totales}$$

$$p = \frac{45}{300} = 0.15$$

La proporción de individuos con lumbalgia es de 45 entre 300; también puede expresarse con su valor decimal 0.15 o bien, multiplicar el resultado por 100 y afirmar que la proporción es del 15%.

A esta proporción, que representa la cantidad de individuos que presentan una determinada enfermedad o factor de riesgo en un momento determinado en el tiempo se le conoce como PREVALENCIA.

El cálculo del tamaño de la muestra para proporciones se utiliza cuando se estudia la presencia de una característica en una población y tiene como objetivo fundamental estimar la prevalencia de una característica en una población o comparar las proporciones de la variable de estudio entre dos o más grupos.

Poblaciones infinitas

Cuando se tiene una característica cuya frecuencia puede ser determinada (contada) entre los individuos de una población, ser expresada numéricamente y se tiene una población cuyo tamaño es extremadamente grande o imposible de contar (infinita), es posible calcular el tamaño de muestra utilizando la siguiente fórmula:

$$n = \frac{Z^2 pq}{e^2}$$

Dónde:

n = Tamaño de muestra

Z = Variable estandarizada, correspondiente a la confianza estadística deseada

p = Probabilidad de ocurrencia de la característica de estudio en la población

q = Probabilidad de no ocurrencia (ausencia) de la característica de estudio en la población, también se representa como 1-p, ya que se trata del complemento de la probabilidad p en una población.

e = error máximo permisible en el estudio (definido por el investigador), expresado como número decimal, en algunos trabajos, se le representa también con la letra d.

Pongamos un ejemplo:

S ha encontrado que 4 de cada 10 mazorcas de maíz son atacadas por el hongo *Ustilago maydis* (huitlacoche).en el centro del país ¿Cuál es la probabilidad de encontrar una mazorca infectada?,

Para obtener la probabilidad de ocurrencia p, es necesario identificar los casos favorables y los casos totales. En este ejemplo se habla de que 4 de cada 10 mazorcas son infectadas. Como lo que se desea estudiar son las mazorcas infectadas, los casos favorables estarían representados por los 4 casos que se presentan dentro de 10 casos evaluados, por lo tanto, se tiene que:

$$p = \frac{No\ de\ casos\ favorables}{No.de\ casos\ totales} = \frac{4}{10} = 0.4$$

Por lo tanto, la probabilidad de ocurrencia de una infección por este hongo en mazorcas de maíz es de 0.4.

Para calcular el complemento (la probabilidad de encontrar mazorcas sanas) se tiene:

$$q = 1 - p = 1 - 0.4 = 0.6$$

Por lo tanto, la probabilidad de encontrar una mazorca sana es de 0.6.

Si se desea calcular el tamaño de muestra para una población infinita con una confianza del 95% (Z=1.96) y se establece un error máximo permisible del 10% (0.1), el tamaño de muestra requerido es:

$$n = \frac{Z^2pq}{e^2} = \frac{(1.96)^2 * 0.4 * 0.6}{(0.1)^2}$$

$$n = \frac{0.9220}{0.01} = 92.2$$

Cabe señalar que el valor obtenido tiene una parte entera y una decimal; en un muestreo no es posible tomar solamente 0.2 (la quinta parte) de una mazorca para el estudio. Una pregunta interesante es ¿se redondea hacia el valor superior (93) hacia el valor inferior (92)?

Si analizamos la fórmula podemos notar que del lado izquierdo del signo igual aparece el tamaño de muestra como numerador, mientras que en el lado derecho aparece el error máximo permisible en el denominador, lo que implica que son INVERSAMENTE proporcionales, es decir, que si aumenta el tamaño de muestra disminuye el error y viceversa.

De acuerdo a este razonamiento, estadísticamente es más conveniente redondear hacia el valor superior que hacia el inferior. Poor lo tanto es posible afirmar que el tamaño de muestra necesario para realizar un estudio representativo, con una confianza del 95% es igual a 93 mazorcas de maíz.

Pero ¿qué sucede en los casos en los que se desconocen los casos favorables de una variable así como la probabilidad de ocurrencia de la característica de estudio en la población?

Es posible calcular el tamaño de muestra considerando que los valores de p y de q son iguales a 0.5. Si bien esta decisión puede parecer arbitraria, tiene un fundamento matemático. Revisemos los resultados del producto de 10 posibles combinaciones de valores de p y q:

p	q	Multiplicación	Resultado
0.1	0.9	0.1 * 0.9	0.09
0.2	0.8	0.2 * 0.8	0.16
0.3	0.7	0.3 * 0.7	0.21
0.4	0.6	0.4 * 0.6	0.24
0.5	0.5	0.5 * 0.5	0.25
0.6	0.4	0.6 * 0.4	0.24
0.7	0.3	0.7 * 0.3	0.21
0.8	0.2	0.8 * 0.2	0.16
0.9	0.1	0.1 * 0.9	0.09

Nótese que el valor más alto obtenido en la multiplicación de p*q se obtiene cuando p y q tienen un valor de 0.5, lo que implica que esta combinación de valores garantiza el TAMAÑO DE MUESTRA MÁS GRANDE, y como al incrementar el tamaño de muestra disminuimos el error, es posible utilizar estos valores como regla empírica en los casos en que los valores de p y q se desconozcan.

Poblaciones finitas

En el caso de contar con una proporción y una población de tamaño conocido, es posible determinar el y tamaño de muestra utilizando la fórmula:

$$n = \frac{NZ^2pq}{e^2(N-1) + Z^2pq}$$

Donde:

N= Tamaño de la población

n = Tamaño de muestra

Z = Variable estandarizada, correspondiente a la confianza estadística deseada

p = Probabilidad de ocurrencia de la característica de estudio en la población

q = Probabilidad de no ocurrencia (ausencia) de la característica de estudio en la población.

e = error máximo permisible en el estudio (definido por el investigador).

Ejemplo:

En una población de 3,000 alumnos de nivel primaria se desea realizar un estudio acerca de sobrepeso, las estadísticas nacionales señalan que 35% de los niños tienen sobrepeso u obesidad en el país. ¿Qué tamaño de muestra se requiere para que esta sea representativa con un 90% de confianza?

En este ejemplo se conoce el valor de p (probabilidad de ocurrencia de sobrepeso u obesidad en la población), por lo que se procede a calcular el valor de q:

$$p = 1 - q$$

Despejando q:

$$q = 1 - p$$

$$q = 1 - 0.35$$

$$q = 0.65$$

Una vez conocidos los valores de p y q, se procede a identificar el valor de Z, que para una confianza del 90% corresponde a

$$Z_{1-\frac{0.1}{2}} = Z_{0.95} = 1.65$$

Si se define que el error máximo permisible es de 10%, se procede a sustituir los valores en la fórmula:

$$n = \frac{NZ^2pq}{e^2(N-1) + Z^2pq}$$

$$n = \frac{3{,}000 * 1.65^2 * 0.35 * 0.65}{(0.1)^2(3{,}000-1) + 1.65^2 * 0.35 * 0.65}$$

$$n = \frac{1{,}858.106}{29.99 + 0.6193}$$

$$n = \frac{1{,}858.106}{30.6093}$$

$$n = \frac{1{,}858.106}{30.6093}$$

$$n = 60.70$$

Redondeamos hacia el entero superior:

$$n \approx 61$$

Por lo tanto, el tamaño de muestra recomendado es de 61 escolares.

Si el investigador optara por considerar las probabilidades de ocurrencia como p=0.5 y q=0.5, el cálculo quedaría como sigue:

DE

$$n = \frac{3{,}000 * \ 1.65^2 * 05 * 0.5}{(0.1)^2(3{,}000 - 1) + \ 1.65^2 * 0.5 * 0.5}$$

$$n = \frac{2{,}041.875}{29.99 + \ 0.6806}$$

$$n = \frac{2{,}041.875}{29.99 + \ 0.6806}$$

$$n = \frac{2{,}041.875}{30.6706}$$

$$n = 66.57$$

Redondeamos hacia el entero superior:

$$n \approx 67$$

Como puede observarse, los valores 0.5 para p y q aseguran un tamaño de muestra mayor, por lo que algunos investigadores optan por calcular el tamaño de muestra siguiendo este criterio.

Próxima entrega: Distribuciones de probabilidad IV: Distribución normal y determinación de tamaños de muestra (muestreo para variables cuantitativas).

Referencias:

Baird, D. (1991). Propiedades matemáticas de la distribución normal o de Gauss.

Condori-Ojeda, Porfirio (2020). Universo, población y muestra. Curso Taller. https://www.aacademica.org/cporfirio/18

Departamento de Didáctica de la Matemática. (2011). Estadística con proyectos. (C. Batanero, & C. Díaz, Eds.) Granada, España: Facultad de Ciencias de la Educación, Universidad de Granada.

García Pérez, A. (2008). Estadística aplicada: conceptos básicos (2a edición ed.). Madrid, España: Educación permanente / Universidad Nacional de Educación a Distancia.

Pértegas Díaz, S., & Pita Fernández, S. (2001). La distribución normal. Cad Aten Primaria, 8, 268-274.

Rev. chil. anest. Vol. 43 Número 2 pp. 116-121|https://doi.org/10.25237/revchilanestv43n02.08

Vargas Velasco, C. (2018). Distribución normal estándar.

DE

Declaración de conflicto de intereses

Los autores de este artículo expresan que no tuvieron ningún conflicto de intereses durante la preparación de este documento ni para su publicación.

Tabla 1: Distribución Normal Estándar. Áreas bajo la curva

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999