

Inferencia estadística.

A partir de datos estadísticos, tales como por ejemplo el promedio, la desviación estándar, la proporción,... de la muestra se obtiene el valor real de la población, es decir del universo.

Definiciones:

Error muestral es la cantidad de error que aceptamos a priori, antes de comenzar a realizar el estudio, es decir, el error que tendrán los resultados que obtendremos.

Después de hacer la encuesta se vuelve a calcular el error porque puede tener variaciones, puede aumentar o disminuir, dependiendo si hemos conseguido completar el número de encuestas que pretendíamos.

Intervalo de confianza: es el margen que damos de acertar, de que los datos son correctos.
Es el porcentaje de que el valor de la población esté dentro del intervalo definido por el promedio más la desviación estándar.

$$X_{real} = \bar{X} \pm \sigma$$

Cálculo del tamaño muestral:

Con variables cuantitativas:

Es decir, con variables numéricas y el valor que buscamos es el promedio.

$$n = \frac{N \cdot k^2 \cdot \sigma^2}{N \cdot e^2 + k^2 \cdot \sigma^2}$$

Donde: n es el tamaño de la muestra
N es el tamaño del universo
k es el nivel de confianza
 σ^2 es la varianza
e es el error muestral

Con variables cualitativas:

Es decir, con variables categóricas y el valor que buscamos es una proporción.

$$n = \frac{N \cdot k^2 \cdot p \cdot (1 - p)}{(N - 1) \cdot e^2 + k^2 \cdot p \cdot (1 - p)}$$

Donde: n es el tamaño de la muestra
N es el tamaño del universo
k es el nivel de confianza
p es la proporción
e es el error muestral

Varianza y proporción:

En ambos casos la varianza y la proporción se deberían conocer de estudios anteriores para que la muestra sea más pequeña. Dicho de otra forma, si ya disponemos de información de cómo ha sido anteriormente el dato que buscamos necesitamos una muestra más pequeña. Si no conocemos como se ha comportado la variable que buscamos entonces el criterio de máxima indeterminación.

a) para la varianza:

Consideramos los valores máximo y mínimo, y el promedio de esos valores.

Por ejemplo si la variable puede tomar valores de 0 a 10 entonces

Mínimo: 0

Máximo: 10

$$\text{Promedio: } \bar{X} = \frac{X_{\text{máximo}} + X_{\text{mínimo}}}{2} = \frac{10 + 0}{2} = 5$$

La varianza la calculamos:

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{n} = \frac{(0 - 5)^2 + (10 - 5)^2}{2} = 25$$

b) para la proporción:

el criterio de máxima dispersión es que tenga igual proporción para cada posible valor, si son dos el 50 % para cada uno.

Consideraremos $p = 0.5$

Nivel de confianza:

Si consideramos un intervalo de confianza del 95 %, entonces $k = 1.96$

Esto significa que el 95 % de los valores están dentro del intervalo del promedio y la desviación estándar.

Si consideramos un intervalo de confianza del 68 %, entonces $k = 1$

El error muestral:

Este está a criterio de la persona que hace la investigación.

Por ejemplo si aceptamos un error del 5 %, entonces $e = 0.05$

Si aceptamos un error del 3 %, entonces $e = 0.03$

Si aceptamos un error del 10 %, entonces $e = 0.10$

Si aceptamos un error del 12 %, entonces $e = 0.12$

Aunque no son recomendables errores superiores al diez por ciento, lo usual es el 5 %.

Actividades:

- 1.) Calcule el tamaño de una muestra con el 5 % de error y un nivel de confianza del 95%, para una variable cuantitativa que tiene un valor máximo de 200 y un valor mínimo de 25, en:
 - a. En todo el departamento de Nueva Segovia.
 - b. En cada uno de los municipios del departamento de Nueva Segovia
- 2.) Calcule el tamaño de una muestra con el 3 % de error y un nivel de confianza del 95%, para una variable cuantitativa que tiene un valor máximo de 20 y un valor mínimo de 5, en:
 - a. En todo el departamento de Nueva Segovia.
 - b. En cada uno de los municipios del departamento de Nueva Segovia
- 3.) Calcule el tamaño de una muestra con el 5 % de error y un nivel de confianza del 95%, para una variable cualitativa, en:
 - a. En todo el departamento de Nueva Segovia.
 - b. En cada uno de los municipios del departamento de Nueva Segovia
- 4.) Calcule el tamaño de una muestra con el 5 % de error y un nivel de confianza del 68%, para una variable cualitativa, en:
 - a. En todo el departamento de Nueva Segovia.
 - b. En cada uno de los municipios del departamento de Nueva Segovia
- 5.) Observando los datos que ha encontrado, que conclusiones obtiene. Piense en las variaciones del tamaño de la muestra al variar el error, el nivel de confianza y el tamaño del universo.

NUEVA SEGOVIA		208,523
1	Jalapa	54,491
2	Murra	14,847
3	El Jicaro	25,901
4	San Fernando	8,549
5	Monzonte	6,795
6	Diplito	5,207
7	Macuelizo	6,076
8	Santa María	4,404
9	Ocotal	34,580
10	Ciudad Antigua	4,868
11	Quilali	26,461
12	Wiwili (NS)	16,344

NOTA: al calcular el tamaño de la muestra por municipio tiene que realizar 12 cálculos. 12 muestras diferentes porque el universo es diferente.