TECNICA DE MUESTREO ESTADISTICO II

Mg. Sc. Bustamante Cabrera Gladys¹

RESUMEN

Las investigaciones científicas, se valen de diferentes métodos investigativos, que siguen una secuencia lógica de pensamiento para lograr la obtención de información.

En el intento de delimitar la cantidad de población a ser estudiada, el investigador encuentra conflictos a la hora de fijar el número exacto de muestra a estudiar. Es así que el diseño de la investigación, determinará, cuánto de muestra es la ideal para el trabajo científico, manteniendo la confiabilidad del mismo, teniendo el menor error deseable en el momento de emitir conclusiones.

Actualmente, muchos programas estadísticos, han reemplazado el cálculo matemático manual, disminuyendo el tiempo de ejecución de las fórmulas estadísticas, sin embargo, no todos los interesados en investigación manejan a la perfección los paquetes informáticos en mención, por lo que el artículo, intenta orientar con claridad, la forma de ejecución de un muestreo ante posibilidades diferentes de objetivos planteados por un investigador en un estudio de prevalencias.

Reiterativamente se insistirá en la necesidad de un diseño metodológico claro, para elegir la fórmula ideal de muestreo que permita encontrar un número adecuado de observaciones para dicho diseño.

¹ Médico Internista. Docente Emérito Facultad de Odontología UMSA. Mg.Sc. Psicopedagogía y Educación Superior. Mgs. Planificación, Gestión, y Evaluación de proyectos. MBL. Dirección de Desarrollo Local.

PALABRAS CLAVE

Prevalencia. Muestreo. Estadística

INTRODUCCION

Las técnicas de muestreo estadístico deben orientar al investigador a tomar un número representativo de la población para que el análisis resultante permita extrapolar con algún grado de confiabilidad los resultados a todo el grupo poblacional de estudio.

En ocasiones, el muestreo estadístico puede generar conflictos en la decisión del investigador al desconocer que fórmula se debe utilizar para obtener una muestra ideal. En este sentido, varios programas estadísticos (Epi. Info. Statistics, SPSS, Excel, etc.) han resuelto el problema de uso de fórmulas manuales, las cuales sin embargo deben ser conocidas para orientación en la resolución o llenado de información en informática.

De este modo es necesario tener en cuenta, que para la aplicación del muestreo estadístico, el tipo de estudio planificado por el investigador será prioritario. Es así, que el muestreo en estudios de prevalencia toma algunos parámetros diferentes a los estudios de cohorte o caso control, y por supuesto a aquellos que son de carácter experimental.^{1,2}

ESTUDIOS PARA DETERMINAR PREVALENCIAS

Un estudio que pretende determinar prevalencia, es aquel en el que se analizan relaciones entre enfermedades o características relacionadas con la salud. Es así que la relación entre una variable y la enfermedad en estudio, puede ser observada en diferentes subgrupos de población en presencia o

3. Determinar la muestra poblacional y la selección de la misma.
4. Validar los instrumentos.
5. Asegurar la posibilidad de comparación en los diferentes grupos.
6. Determinar el tipo de análisis epidemiológico.
7. Establecer los resultados y conducta esperada con los mismos.

Prevalencia esperada	Precisión o error						
	25%	20%	10%	5%	3%	1%	0,5%
5%	3	5	19	73	292	1.825	7.300
10%	6	9	35	139	554	3.458	13.830
15%	8	13	49	196	784	4.899	19.593
20%	10	16	62	246	984	6.147	24.587
25%	12	19	73	289	1.153	7.203	28.812
30%	13	21	81	323	1.291	8.068	32.270
35%	14	22	88	350	1.399	8.740	34.959
40%	15	24	93	369	1.476	9.220	36.880
45%	16	24	96	381	1.522	9.508	38.032
50%	16	25	97	385	1.537	9.604	38.416

Tabla 1. Tamaño de muestra necesario para determinar la prevalencia en una población grande y con un nivel de confianza del 95%

Fuente: Mateu Enric. *Tamaño de la Muestra*

ausencia de la variable, o en presencia o ausencia de la variable en individuos enfermos comparándola con los sanos.

Los estudios de prevalencia se caracterizan por ser de carácter transversal y no determinan incidencias ni relación causa efecto.

El diseño de un estudio de prevalencia debe:¹⁻⁷

1. Definir con claridad la población a ser estudiada.
2. Determinar si el estudio se realizará en toda la población o en una muestra, dependiendo de la cantidad de observaciones existentes.

El muestreo para la determinación de la prevalencia, pretende estimar la cantidad presente de alguna enfermedad en el grupo poblacional elegido, por lo tanto, requiere tener los siguientes datos:

- a) Frecuencia esperada de enfermedad (p), lo que explica que el investigador, no puede ir a buscar un resultado al azar, sino más bien a tratar de comprobar que la hipótesis que se plantea de un porcentaje esperado es cierto o falso. Para ello, el marco teórico de un estudio, puede orientar con información, sobre las frecuencias encontradas en poblaciones similares, de donde se podrán estimar valores de p .

b) Tamaño de la población: tomar en cuenta que se muestreará a poblaciones grandes, no a aquellas cuyo número es pequeño y fácilmente investigable. Es así que se considerará *población infinita* a aquellas mayores a 10.000, o a las que se suponen menores al 5% de la población total. Mientras que una *población finita*, será aquella cuyo registro poblacional es menor a 10.000, o corresponde a una población cerrada cuyos valores no serán modificados en el momento de la investigación. (*Fórmulas para acceso, vol. 10 serie 1*).

c) Nivel de confianza, o precisión (Z_α) que determina que la cantidad elegida de la población permitirá resultados aceptables, en los cuales la enfermedad que se está estudiando esté comprendida en los intervalos de confianza. Regularmente se utiliza el 95%.⁶

MUESTREO POR ESTIMACION DE PROPORCIONES

Los estudios de prevalencia, pueden a su vez:

1. Delimitar la estimación de una proporción de una población, con precisión absoluta, para lo cual se utilizan las fórmulas de muestreo finito o infinito según el número poblacional que se tenga.
2. Pero si el objetivo del estudio es probar que la proporción de la población de estudio difiere de un valor establecido en la literatura o en estudios previos (p) se debe establecer una *hipótesis* que describa esta diferencia y que permita que el investigador intente demostrar la misma. De este modo, este tipo de muestreo deberá tener los siguientes datos:

- a) Valor de la proporción bajo la hipótesis nula (P_0)
- b) Valor anticipado de la proporción o lo que se desea comparar (P_a)
- c) Confiabilidad esperada del estudio (Z_α)
- d) Potencia del test $100(1-\beta)\%$
- e) Hipótesis alternativa: $P_a > P_0$, $P_a < P_0$ ó $P_a = P_0$

Tomando en cuenta que los valores para Z_α y Z_b son de:

Tabla 2. Valores de Z_α y Z_β más frecuentemente utilizados

Z_α		
α	Test unilateral	Test bilateral
0.200	0.842	1.282
0.150	1.036	1.440
0.100	1.282	1.645
0.050	1.645	1.960
0.025	1.960	2.240
0.010	2.326	2.576
Potencia		
β	($1-\beta$)	Z_β
0.01	0.99	2.326
0.05	0.95	1.645
0.10	0.90	1.282
0.15	0.85	1.036
0.20	0.80	0.842
0.25	0.75	0.674
0.30	0.70	0.524
0.35	0.65	0.385
0.40	0.60	0.253
0.45	0.55	0.126
0.50	0.50	0.000

Fuente: *Tamaño de la muestra*⁶

La fórmula a ser utilizada para una cola será:

$$n = \frac{[z_{1-\alpha}\sqrt{P_0(1-P_0)} + z_{1-\beta}\sqrt{P_a(1-P_a)}]^2}{(P_0 - P_a)^2}$$

$$n = \frac{1,96^2 \cdot 0,55(1-0,55) + 1,645^2 \cdot 0,43(1-0,43)}{(0,55-0,43)^2}$$

n=124,09

Para pruebas de dos colas la formula tiene una variación pequeña, Al tomarse un segundo valor de referencia.

$$n = \frac{[z_{1-\alpha/2}\sqrt{P_0(1-P_0)} + z_{1-\beta}\sqrt{P_a(1-P_a)}]^2}{(P_0 - P_a)^2}$$

Ejemplo:

Si un investigador plantea una pregunta donde establece que la proporción descrita en literatura (Po) sobre la enfermedad X de 55% (0,55), es solo de 43% (Pa=0,43), teniendo un 95% de confianza, se reemplazan los valores en mención en la fórmula descrita, obteniéndose de esta manera la muestra a ser estudiada.

$$n = \frac{1,96/2^2 \cdot 0,55(1-0,55) + 1,645^2 \cdot 0,43(1-0,43)}{(0,55-0,43)^2}$$

n=90,24

3. Si el investigador tiene como fin la estimación de una diferencia entre dos proporciones poblacionales, los datos necesarios para el muestreo deben ser:

- Proporciones esperadas de ambos grupos poblacionales (P1-P2).
- Confiabilidad del estudio (Zα).
- Precisión absoluta requerida (d).

Debiéndose aplicar la siguiente fórmula:

$$n = \frac{z_{1-\alpha/2}^2 [P_1(1-P_1) + P_2(1-P_2)]}{d^2}$$

Ejemplo: El investigador pretende estimar una diferencia de riesgo de una situación X con un nivel de confianza de 95% teniendo una precisión de 0,05%, sin que tenga datos de las proporciones de los dos grupos poblacionales en estudio. Por lo tanto P1=P2= 0,5 y d 5%

$$n = \frac{1,645^2/2 (0,5(1-0,5)+0,5(1-0,5))}{0,05^2}$$

n=100,67

4. De otra forma, si el investigador tiene la intención de probar que dos proporciones de población son diferentes, debe tener la siguiente información:

- Proporción esperada de las poblaciones P1 y P2.
- Confiabilidad del estudio (Zα)
- Potencia del test (Z β)
- Hipótesis nula: P1-P2=0
- Hipótesis alternativa P1-P2 mayor a 0, P1-P2 menor a 0 o P1-P2 igual o desigual a 0

En caso de hacer una prueba de una cola la fórmula a ser utilizada es:

$$n = \frac{[z_{1-\alpha}\sqrt{2P_m(1-P_m)} + z_{1-\beta}\sqrt{V}]^2}{(P_1 - P_2)^2}$$

$$P_m = (P_1 + P_2)/2$$

$$V = P_1(1-P_1) + P_2(1-P_2)$$

En el supuesto que la información bibliográfica mostrara una p1 de 15%(0,015) y p2 45% (0,045), y se realizara un estudio con 95% de confiabilidad con una potencia de 90%, se tendría:

$$P_m = 0,015 + 0,045/2$$

$$P_m = 0,3$$

$$V = 0,015(1-0,015) + 0,045(1-0,45)$$

$$V = 0,057$$

$$n = \frac{1,645^2 \times 2 \times 0,3 \times 0,7 + 1,282^2 \times 0,057}{(0,015 - 0,045)^2}$$

$$n = 2101,9$$

Cuando se realiza una prueba de dos colas, la formula será:

$$n = \frac{[z_{1-\alpha/2} \sqrt{2P_m(1-P_m)} + z_{1-\beta} \sqrt{V}]^2}{(P_1 - P_2)^2}$$

$$n = \frac{1,645/2^2 \times 2 \times 0,3 \times 0,7 + 1,282^2 \times 0,057}{(0,015 - 0,045)^2}$$

$$n = 785,37$$

PRUEBAS DE UNA Y DOS COLAS

La formulación de hipótesis en la investigación no es obligatoria y su declaración responde a un paradigma deductivo, que no es único en la generación investigativa, pudiendo los resultados proponer el inicio de hipótesis para confirmarse con estudios posteriores.⁸

Sin embargo, algunos estudios requieren el planteamiento de hipótesis,

basada en estudios y publicaciones previas.

La aceptación o rechazo de las hipótesis, requiere de pruebas o contrastes estadísticos que llevan a dicha decisión.

Una prueba de cualquier hipótesis estadística con alternativa unilateral como:

$$H_0: O = O_0$$

$$H_1: O > O_0 \text{ ó } O < O_0$$

Se denomina prueba de una cola, donde la región crítica para la hipótesis alternativa $O > O_0$ generalmente yace en la derecha de la distribución del estadístico, mientras que la región crítica de la hipótesis alternativa $O < O_0$, está a la izquierda de la distribución del estadístico.

Cuando la alternativa de la hipótesis es bilateral como:

$$H_0: O = O_0$$

$$H_1: O \neq O_0$$

Recibe en nombre de prueba de dos colas, donde la región crítica se divide en dos partes con la misma probabilidad en estadístico de prueba.⁹⁻¹⁰

BIBLIOGRAFIA

1. Anónimo. Determinación del tamaño muestral. 1-15 .
2. Mateu E., Casal J. Tamaño de la muestra. Rev. Epidem. Med. Prev. 2003;1;8-14
3. Pértega Díaz S, Pita Fernández S. La distribución normal. Cad Aten Primaria 2001; 8: 268-274
4. Sales Carmona V., Orozco Montero J.A. Calculo del tamaño de la muestra. URL disponible en: <http://www.facmed.unam.mx/deptos/salud/censenanza/spiii/spiii/para%20remplazar/sales2.pdf> Accedido en fecha. 20 de mayo del 2013.

5. Estudios de Prevalencia. Departamento de Estadística. Universidad Carlos III de Madrid. Apuntes 1-15.
6. Tamaño de la muestra. URL disponible en: <http://cursos.aiu.edu/Probabilidad%20y%20Estadística.html> Accedido en fecha 20 de mayo del 2013.
7. Dawson-Saunders B, Trapp RG. Bioestadística Médica. 2^{da} edición. México: Editorial el Manual Moderno; 1996
8. Morales P- Sobre las Hipótesis direccionales y no direccionales. URL disponible en: <http://www.upcomillas.es/personal/peter/investigacion/Hipotesis.pdf>. Accedido en fecha: 20 de mayo del 2013.
9. Escobar M. La prueba estadística. URL disponible en: <http://web.usal.es/~modesto/pdf/A6.-%20La%20prueba%20estadística.pdf> Accedido en fecha: 20 de mayo del 2013.
10. Montero R. Contraste de hipótesis. URL disponible en: <http://www.ugr.es/~montero/matematicas/contrastos.pdf> Accedido en fecha: 20 de mayo del 2013