APROXIMACION AL MUESTREO ESTADISTICO EN INVESTIGACIONES CIENTIFICAS.

Mgs. Dra. Bustamante C. Gladys¹

RESUMEN

El muestreo estadístico es posiblemente una de las áreas con mayor conflicto para el investigador joven, quien a falta de orientación complejiza el muestreo estadístico al desconocer el uso de cada fórmula en cada caso dado.

Es en este sentido que el presente artículo, intenta aproximar al lector a métodos simples de muestreo, los cuales pueden ser aplicados en las investigaciones de carácter cuantitativo, permitiendo de este modo la fiabilidad de resultados en la investigación.

PALABRAS CLAVE

Muestreo, método probabilístico, cálculo muestral

INTRODUCCION

Durante el proceso de investigación científica, el muestreo de población se constituye en un área conflictiva para los jóvenes que se introducen en el proceso de la investigación científica. Es así que se puede mencionar que el muestreo no es más que el uso de una herramienta, donde se utilizan poblaciones en números pequeños para lograr hacer inferencias en poblaciones con las mismas características del grupo elegido. En realidad, el proceso de selección de la muestra, parte primero de la elección de una población, de la cual se extrae un número racional de objetos que sean representativos de la población.

TERMINOLOGIA COMUN

Para poder comprender el término muestreo, se debe dejar en claro otros términos asociados como son:

- a) Población (N): se define como población al grupo de personas u objetos elegidos para el estudio investigativo.
- b) Unidad de muestreo. Cantidad de elementos de una población que se pretende investigar.
- c) Unidad de Análisis: persona u objeto único, del cual se obtendrá información durante la investigación
- d) Muestra (n): es una porción de la población que cumpla con las características comunes de ésta y represente al total de los objetos que la integran
- e) Error de muestreo: es la falla producida por el investigador por error en la selección de personas u objetos de la población, habiendo elegido solo una parte que no es la representativa del total
- f) Desviación estándar: Es una medida de centralización o dispersión, equivalente a la raíz cuadrada de la varianza
- g) Varianza: es la media aritmética (evada al cuadrado), de las desviaciones con respecto a la media.

TIPOS DE MUESTREO

Los tipos de muestreo pueden ser

A. **PROBABILÍSTICOS:** El muestreo probabilístico se caracteriza por que su metodología es rígida, permitiendo la reproducción de resultados y la replicación de los mismos, disminuyendo el error muestral y el sesgo de selección. Se utiliza en investigaciones de carácter cuantitativo. El muestreo probabilístico puede ser:

- a. **Aleatorio simple:** El procedimiento del muestreo aleatorio simple se caracteriza por que todos los miembros de la población tienen las mismas oportunidades de ser elegidos. Se utiliza para ello la tabla de aleación simple, la cual asigna un número a cada sujeto u objeto a investigar, y por medios manuales o mecánicos se procede a la elección de sujetos u objetos de acuerdo al número necesario para la investigación. Si se requiriesen 12 unidades de muestra, estas serán elegidas de la tabla de aleación utilizando la fila o la columna con números pares, y en forma ordenada,

¹ Médico Internista. Docente Emérito UMSA.
Mgs.Sc. Psicopedagogía y Educación Superior.

elegirá los números que se le presenten haciendo comparación con el listado previo de unidades de análisis, hasta completar los 12 seleccionados ⁽¹⁾.

Ejemplo: De una lista de 80 personas se elegirán 12, que serán los números marcados (**en negrilla**) en el modelo de tabla de números aleatorios siguiente.

1. **4251** 5149 4751 4847 4249
2. **4849** 5051 5046 4756 4738
3. **5692** 9870 3583 8997 1533
4. **2080** 3828 7880 0586 8482
5. **1039** 3382 7600 1077 4455
6. **6477** 5289 4092 4223 6454
7. **4554** 6146 4846 4647 5034
8. **0772** 2160 7236 0812 4195
9. 0092 1629 0377 3590 2209
10. **7315** 3365 7203 1231 0546
11. **5775** 7517 8974 3961 2183
12. **5500** 2276 6307 2346 1285
13. **3251** 8902 8843 2112 8567
14. 4675 1435 2192 0874 2897
15. 3521 8772 6612 0721 3899

Es decir que los números elegidos corresponden a las personas de la lista previa que se encuentran en el puesto 42, 48, 56, 20, 10, etc. Las tablas de números aleatorios se encuentran de manera accesible en libros de estadística, páginas web, o programas de análisis estadístico donde se pueden generar números de elección aleatoria de manera infinita.

- b. **Aleatorio estratificado:** se caracteriza por la división de la población en grupos de características comunes ⁽¹⁻²⁻³⁾. Por *ejemplo:*
- Varones
  - Mujeres

Este tipo de muestreo tiene la ventaja de que se selecciona a la población de acuerdo a las variables que se están estudiando, pero puede generar problemas en el momento del análisis de los resultados, si la elección de las variables fue inadecuada, o se deben ponderar pesos de participación de cada una de las variables.

- c. **Muestreo polietápico :** Es la obtención de la muestra en varias etapas sucesivas, utilizada cuando la población a estudiar es muy grande.

Ejemplo: Muestra de 10000 personas para estudio sobre calidad del servicio telefónico en Bolivia. Primera etapa, muestreo en ciudades principales. Segunda etapa, muestreo en provincias, etc.

- d. **Muestreo por conglomerados:** donde la población se divide en grupos más pequeños de los cuales el comportamiento interno es muy variable, pero es similar al de los otros grupos ⁽⁵⁾.
Ejemplo: zonas geográficas de la ciudad de La Paz (zona norte, zona sur, zona este, zona oeste, etc.)
- e. **Muestreo sistemático:** El este tipo de muestreo se seleccionan los elementos a ser estudiados, se define el número de elementos de la muestra, definiendo un intervalo resultante de la división de la población y la muestra:

Ejemplo: Población: 1000 sujetos, muestra 10 sujetos Intervalo de salto= $100/10 = 10$. Este resultado de intervalo, es la elección cada 10 sujetos de la lista hasta llegar al final de la población.

- Este muestreo puede al mismo tiempo ser:
- *De Asignación simple.*, donde en cada estrato habrán el mismo número de unidades muestrales.
  - *De asignación proporcional*, donde el tamaño de la muestra está en relación proporcional con el estrato correspondiente, en relación a la población total ⁽⁴⁾
  - *Asignación óptima:* en la cual uno de los estratos tiene mayor numero muestral a mayor variabilidad interna, es decir es aquel estrato más heterogéneo.

Tiene la ventaja de ser un método sencillo, sin embargo puede generar conflictos si no se tiene una lista bien establecida ⁽¹⁻²⁻⁴⁾.

- B. **NO PROBABILÍSTICOS:** El muestreo no probabilístico es aquel usado frecuentemente en investigaciones cualitativas, donde la posibilidad de generalización de los resultados es poco factible. En este caso la muestra no es representativa de toda la población y la

elección depende de los criterios del investigador, tendiendo a generar sesgos en la selección poblacional.

Este tipo de muestreo se divide a su vez en los siguientes:

- a) Muestreo por cuotas: o “accidental”, donde el investigador identifica sujetos u objetos con características comunes, pero sin el carácter aleatorio de la muestra probabilística: Por ejemplo: universitarios que trabajan por la noche.
- b) Muestreo opinático o intencional: Es la intención del investigador de observar características comunes en un grupo heterogéneo, por elección y análisis del mismo investigador.
- c) Muestreo casual o incidental: donde el investigador elige en forma intencional, quienes ingresaran al estudio, dependiendo de la posibilidad de acceso a los mismos, cercanía, amistad, etc.

CALCULO DE LA MUESTRA PROBABILISTICA:

Existen numerosas posturas sobre el número de muestra ideal. Algunos autores mencionan que la muestra ideal mínima es de 30 unidades de análisis. Otros mencionan que el número ideal es el correspondiente al 30% de la población.

En el momento de realizar el muestreo, debemos definir qué tipo de estudio pretendemos llevar a cabo.

Si el estudio que se pretende ejecutar requiere precisión rígida y criterios de certeza, la elección de la muestra deberá ejecutar fórmulas que permitan la extrapolación de resultados a partir del cálculo de “la confiabilidad el estudio” y del “error esperado”.

En este sentido si se desea un estudio con un nivel de confianza que permita la extrapolación de resultados en forma óptima, se utilizará el nivel comprendido entre el 95% al 99.9%

De este modo tenemos que el nivel de confianza esperado, se apropia de un valor de Z_{α} , mismo que será posteriormente

reemplazado en las fórmulas de cálculo muestral ⁽¹⁻³⁻⁶⁾.

El Nivel de confianza es aquel porcentaje de seguridad de generalización de los hallazgos del estudio de investigación. El valor comúnmente utilizado es el de 95%, que explica esta posibilidad porcentual de que nuestros hallazgos tienen una posibilidad de un error de 5%. Sin embargo el nivel de confianza que se menciona como valor porcentual, no es un porcentaje en sí, éste nivel, se obtiene de la distribución normal estándar de un área simétrica de una curva normal (área de confianza) donde se busca el Valor Z de la variable aleatoria de ésta área.

Nivel de confianza	Z_{α}
99.7%	3
99%	2,58
98%	2,33
96%	2,05
95%	1,96
90%	1,645
80%	1,28
50%	0,674

Fuente: Elaboración propia

Luego de identificar el nivel de confianza y su valor Z_{α} , se identifica el “error o porcentaje de error de la muestra” (d), vale decir, la elección de una probabilidad de aceptar una hipótesis que sea falsa como si fuera verdadera. Es común aceptar errores que van del 4% al 6%. A medida que la muestra es mayor, el error tiende a reducirse

Ejemplo: Nivel de confianza 95%, error de 5%
Nivel de confianza de 98%, error 2%

Otro elemento a tomar en cuenta en el momento de aplicar la formula de muestreo es la “variabilidad positiva” (p) y la “variabilidad negativa” (q), ambas son complementarias, es decir, que su suma es igual a la unidad

$$p+q=1$$

Entonces asumimos que $q=1-p$

A partir de esta información se pueden definir dos tipos de muestreo:

1.- Muestreo de población infinita: En la cual el investigador no conoce el número exacto de la población a estudiar, y este número seguramente se modificará en el tiempo de estudio. (población mayor a 100.000)

Ejemplo: Se requiere estudiar la valoración de la calidad de los servicios de salud en pacientes que acuden a hospitales públicos en los siguientes 5 años.

La fórmula a ser aplicada en este caso debe contar con la siguiente información:

- Nivel de confianza: 95% = $Z_{\alpha/2}$ 1,96
- Error esperado (d): 5% = 0,05
- p= 96%= 0.06
- q= (1-p) 4%=0.04
- n= muestra

Se aplica la siguiente formula, donde se reemplazaran los valores arriba descritos.

$$n = \frac{Z^2 \times p \times q}{d^2}$$

Por lo tanto, se tendrá lo siguiente:

$$n = 1.96^2 \times 0,96 \times 0.04 / 0,05^2$$

n= 59

2.- Muestreo de población finita: Se utiliza cuando el investigador conoce el número exacto de la población a estudiar (Población menor a 100.000), misma que no se modificará durante el estudio:

Ejemplo: Se requiere estudiar valoración de la calidad de los servicios de salud en pacientes que acudieron al servicio de ginecología del Hospital XXX el mes de junio del 2011. En este caso se conoce el número de pacientes durante ese periodo. Los datos que se deben consignar son:

- Población (N) = 3000 pacientes
- Nivel de confianza: 98% = $Z_{\alpha/2}$ 1,96
- Error esperado (d): 5% = 0,05
- p= 96%= 0.06
- q= (1-p) 4%=0.04
- n= muestra

Se aplica la siguiente formula, donde se reemplazaran los valores arriba descritos.

$$n = \frac{Z^2 \times p \times q \times N}{(N-1)d^2 + Z^2 \times p \times q}$$

Por lo tanto, se tendrá lo siguiente:

$$n = \frac{1.96^2 \times 0,96 \times 0.04 \times 3000}{0,05^2 \times 2999 + 0.1475}$$

n= 57,9

SI SE CONOCE LA DESVIACIÓN ESTANDAR

Cuando la desviación estándar es conocida la fórmula debe utilizar esta información para hacer una estimación con mayor precisión utilizando la siguiente fórmula:

$$n = \frac{Z^2 \times s^2}{d^2}$$

Si se conoce la prevalencia y es mayor al 50%, se puede utilizar la siguiente tabla para estimar el tamaño de la muestra necesaria.

Ejemplo: Si deseo calcular una muestra con una prevalencia esperada del 60% (40% de la tabla) con una precisión del 0.5% y un nivel de confianza del 95% el tamaño es 36.880

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 E.

Sin embargo el muestreo estadístico tiene muchas aristas de elección de acuerdo al tipo de estudio planteado ⁽⁶⁾.

Esta aproximación inicial orientará a los jóvenes investigadores al uso inicial de formulas de aplicación estadística que les permita calcular la muestra con facilidad.

BIBLIOGRAFIA

1. Fernández M, Virtudes A. Ruiz Fuentes N. Muestreo estadístico, 1era edición. Editorial Oviedo, 2004: 15-200
2. Muestreo. URL disponible en: <http://www.chospab.es/calidad/archivos/M>

- [etodos/Muestreo.pdf](#). Fecha de acceso , 30 de junio del 2011
3. Martínez Bencardino C. Estadística y muestreo. Edición 12ª. Colección de textos universitarios. 2005; 455-889
  4. Introducción a la Inferencia estadística. URL disponible http://www.uv.es/~rmartine/inferencia_AD E/tipos%20de%20muestreo.PDF. Fecha de acceso 1 de julio del 2011
  5. Curso de estadística Aplicada a las Ciencias Sociales. Tema 8 Poblaciones y Muestras. Manual (tema 19) y Agresti (Cap. 2) URL disponible <http://www.uclm.es/profesorado/jmezo/es tadistica/t8.pdf>. Fecha de acceso 1 de julio del 2011
  6. Mateu E. Casal J. Tamaño de la muestra. Rev.Epidem.Med.Prev.(2003),1:8-14
  7. Metodología de la investigación Capitulo V. Universidad Tecnológica El Salvador URL disponible en <http://biblioteca.utec.edu.sv/siab/virtual/au prides/39064/capitulo%205.pdf>. Fecha de acceso 1 de julio 2011
  8. Saravia Gallardo M.A. Metodología de la Investigación Científica. URL disponible en : <http://www.cienciaytecnologia.gob.bo/con vocatorias/publicaciones/Metodologia.pdf> . Fecha de acceso 1 de julio del 2011-07-02
  9. Días Narváez. V.P. Metodología de la Investigación científica y Bioestadística: Para médicos, odontólogos y estudiantes de Ciencias de la Salud. 2da Ed. Chile. 2009:50-111