

UNIVERSO Y MUESTRA

(Elaborado por Dra. Cristina Ludewig)

OBJETIVOS:

- Identificar el universo y la muestra en investigaciones y ejemplos que se le presenten.
- Reconocer las muestras representativas y no representativas.
- Identificar los distintos tipos y clases de muestra.
- Seleccionar la muestra apropiada para ejemplos de investigación y describir el procedimiento a seguir para conformar esa muestra.
- Mencionar los factores de los cuales depende el tamaño de la muestra.

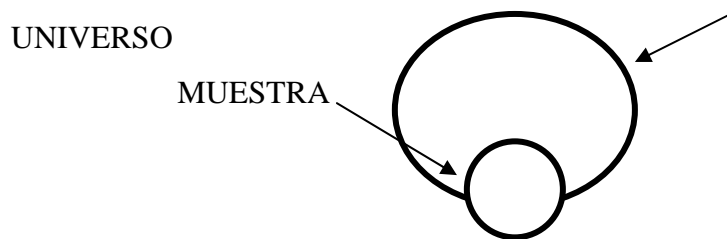
En la mayoría de las situaciones de investigación no es posible estudiar todos los elementos o sujetos a los cuales se refiere el problema, sino que se trabaja con un grupo de ellos para luego generalizar los resultados a la totalidad, en un proceso que se conoce como inferencia estadística. Para poder hacer esta inferencia es necesario que la cantidad de sujetos y la forma como son seleccionados, sean adecuadas. A continuación se desarrollan algunos aspectos básicos para facilitar a los alumnos de Investigación en Salud, el manejo apropiado de esta importante fase de la investigación.

DEFINICIÓN DE TÉRMINOS

Dos conceptos que interesa definir son los de población y muestra. **Población** (o universo) es cualquier colección finita o infinita de elementos o sujetos (población de pacientes que acuden al Hospital Central “Antonio María Pineda”, población de habitantes del área de influencia del ambulatorio urbano tipo III Cabudare, población de familias de una comunidad, población de viviendas en el municipio Iribarren). Algunos autores establecen diferencias entre los términos universo y población, indicando con el primero un conjunto de personas, seres u objetos y con el segundo, un conjunto de números obtenidos midiendo o contando cierta característica de los mismos, de allí que un universo puede contener varias poblaciones. Por ejemplo, del universo de pacientes del Hospital Central Antonio María Pineda, una población estaría dada por el conjunto de las edades de esas personas, sus tallas, sus pesos, etc. Cuando no haya posibilidad de confusión, se usa el término población como sinónimo de universo, siendo ese el criterio que se usará en esta lectura.

Se habla de que una población es *finita* cuando consta de un número limitado de elementos, ejemplo: todos los habitantes de una comunidad. Una población es *infinita* cuando no se pueden contabilizar todos sus elementos pues existen en número ilimitado, como por ejemplo, la población de insectos en el mundo.

Una *muestra* es un subconjunto de la población, que se obtiene para averiguar las propiedades o características de esta última, por lo que interesa que sea un reflejo de la población, que sea *representativa* de ella, concepto al que volveremos más adelante.



Puesto que el fin que perseguimos al hacer una investigación basada en el estudio de una muestra, es inferir los resultados a la población que nos interesa, es recomendable distinguir entre dos tipos de población: la población objetivo y la población muestreada. La *población objetivo* es aquella sobre la cual el investigador desea establecer una conclusión, por ejemplo, si deseamos determinar la frecuencia de hipertensión arterial en la población adulta de Barquisimeto, la población objetivo está representada por todas las personas adultas que residen en esta ciudad. La *población muestreada* es aquella a partir de la cual se extrajo la muestra y sobre la que puede establecerse la conclusión. Para el ejemplo anterior, supóngase que se decidió extraer la muestra de personas adultas de la ciudad entre los asistentes a los establecimientos de salud del sector público (ambulatorios y hospitales de Barquisimeto); en este caso la población muestreada está constituida por todos los usuarios adultos de estos establecimientos de salud. Los métodos de la inferencia estadística permiten al investigador sacar conclusiones sobre la población muestreada, no sobre la población objetivo, por lo que es conveniente que ambas coincidan, sin embargo en ocasiones esto no es factible y la población muestreada es más restringida que la objetivo, en cuyo caso es necesario que el investigador esté consciente de lo expuesto anteriormente. En el ejemplo que estamos manejando, la conclusión de la investigación será la frecuencia de hipertensión arterial en los usuarios de los ambulatorios y hospitales de la ciudad de Barquisimeto.

VENTAJAS Y LIMITACIONES DEL USO DE MUESTRAS

Ventajas:

- Costo reducido: resulta obvio que si no se estudia la totalidad de sujetos sino una muestra de ellos, los recursos financieros, materiales, personal, etc necesarios para hacer la investigación serán menores.
- Mayor rapidez: de igual forma, la recolección de la información se hará en menos tiempo.
- Mayor exactitud: al estudiar una muestra se reduce el volumen de trabajo, por lo cual es posible entonces emplear personal más capacitado, supervisar con mayor cuidado las actividades de campo, el procesamiento de los datos, y de esta forma obtener resultados más exactos que los que obtendríamos de estudiar toda la población.
- Mayores posibilidades: Existen casos en los cuales no es posible estudiar toda la población, como por ejemplo, cuando ésta es infinita o muy grande o cuando el proceso de medida para estudiar la característica deseada es destructivo (determinar la dosis letal de una droga por ejemplo, o al consumir un artículo para juzgar sobre su calidad). En ese caso solo es posible estudiar una muestra.

Limitaciones:

- No se debe emplear muestras cuando la población es muy pequeña
- La teoría del muestreo es compleja y no es del dominio de la mayoría de los investigadores, por lo que con frecuencia deben buscar apoyo en especialistas en la materia.

CARACTERÍSTICAS DE UNA BUENA MUESTRA

Una muestra debe ser adecuada en cantidad y en calidad. En relación con el primer aspecto, existen procedimientos estadísticos para saber cuál es el número mínimo de elementos que debemos incluir en el estudio para obtener resultados válidos. La calidad involucra el concepto de ***representatividad*** de la muestra. Se dice que una muestra es representativa de la población cuando es un reflejo de ella, es decir cuando reúne las características principales de la población en relación con la variable en estudio.

Si deseamos determinar cuál es la prevalencia de desnutrición en la población infantil de Barquisimeto y estudiamos una muestra de niños obtenida de la zona este de

la ciudad (donde se encuentra la mayoría de las urbanizaciones de clase media y alta), esa no sería una muestra representativa para dicha investigación y la prevalencia de desnutrición que obtendríamos subestimaría la cifra real para la ciudad. Si nuestro objetivo es determinar la duración promedio de la hospitalización de los pacientes del Hospital Central Antonio María Pineda, para lograr una muestra representativa deberíamos incluir pacientes de los departamentos de Medicina, Cirugía, Obstetricia y Pediatría, dado que la estancia hospitalaria difiere entre ellos. Si estudiamos una gran proporción de pacientes de obstetricia, estaríamos subestimando la duración de la estancia hospitalaria en dicho centro asistencial. La representatividad de la muestra es pues un aspecto de gran importancia en la investigación y para lograrla es necesario seleccionar el tipo y clase de muestreo que garantice esta condición y trabajar con un tamaño de muestra adecuado.

TIPOS DE MUESTREO

Se conoce como muestreo el proceso de obtención de la muestra. Puede ser probabilístico y no probabilístico. Hablamos de un *muestreo probabilístico* cuando los integrantes de la muestra se escogen al azar y por lo tanto, puede calcularse con antelación la probabilidad de obtener cada una de las muestras que pueden formarse de esa población o la probabilidad que tiene cada elemento de la población de ser incluido en la muestra.

La selección de los elementos puede hacerse por el método de la lotería, la tabla de números aleatorios o con paquetes automatizados que forman números al azar. El método de la lotería consiste en asignarle un número a cada integrante de la población y luego seleccionar tantos números como sea necesario para completar la muestra. Esto puede hacerse con un biombo (como en las loterías), o con papeles numerados introducidos en una bolsa de la cual se extraen. La tabla de números aleatorios consta de una gran cantidad de números distribuidos en filas y columnas de la cual podemos extraer tantos como necesitemos para formar la muestra. Si tenemos una población compuesta por 800 personas y queremos seleccionar aleatoriamente 30, los pasos serían:

Tabla de números aleatorios

	00-05	06-11	12-17	18-23	24-29
01	1 0 8 5 8 1	0 3 3 7 6 7	6 3 7 8 4 1	6 5 3 1 5 4	1 0 2 2 1 5
02	4 1 9 4 2 0	6 8 4 3 2 4	6 9 6 6 1 4	6 5 6 1 5 9	2 3 6 6 6 2
03	7 7 8 6 4 4	5 8 2 2 8 2	8 6 8 2 8 1	6 0 5 5 1 5	8 8 5 3 4 4
04	2 4 2 4 7 2	0 6 5 5 1 1	3 6 7 6 0 1	5 4 1 7 0 6	7 3 1 3 9 3
05	6 6 2 7 5 8	8 6 1 9 8 8	4 1 5 6 5 7	6 8 1 6 8 6	4 7 0 1 0 1
06	1 6 3 5 2 2	7 1 6 5 7 7	5 0 1 5 8 6	5 2 2 0 7 1	3 6 4 4 3 8
07	1 8 6 1 2 6	2 7 1 7 1 4	0 6 1 3 5 9	1 3 9 2 6 8	1 8 7 7 3 1
08	1 3 8 1 0 2	6 4 8 6 0 2	8 6 6 1 3 0	2 8 3 4 2 1	7 9 7 2 8 9
09	9 9 1 8 8 5	6 5 5 1 6 5	7 4 1 5 3 2	2 1 7 5 2 8	7 4 6 6 4 5
10	1 6 4 5 1 9	3 1 7 9 6 5	6 9 3 1 3 8	2 8 3 8 1 6	1 9 7 3 8 9
11	8 1 2 6 7 7	4 4 3 3 2 0	2 8 4 0 3 8	3 7 1 2 5 3	6 3 7 4 6 6
12	6 8 1 3 5 6	3 1 3 5 7 9	5 4 4 1 0 7	8 3 6 6 6 6	7 5 4 4 8 2
13	4 6 8 0 4 7	6 0 6 6 2 3	6 5 6 3 6 6	3 6 4 4 1 4	1 7 5 8 8 3
14	5 4 5 4 9 8	6 3 1 6 6 5	1 3 3 5 7 4	8 5 6 9 7 2	0 4 5 6 1 6
15	6 6 1 1 0 5	5 2 5 1 1 3	3 7 6 0 3 6	9 2 0 6 0 1	5 2 2 0 3 7
16	3 9 1 4 3 0	6 6 9 8 1 1	3 3 4 7 6 6	7 6 4 5 6 3	3 3 3 3 2 5
17	5 4 2 8 2 2	3 1 1 4 1 3	5 7 0 8 2 3	1 5 1 8 2 4	1 5 9 6 4 9
18	2 8 5 3 5 5	5 5 7 5 5 6	4 7 1 7 8 9	4 5 1 7 6 3	8 8 8 7 1 4
19	6 8 5 2 6 1	1 7 2 4 6 8	4 8 6 8 2 8	4 6 8 1 1 7	1 2 7 7 3 1
20	1 6 3 0 3 0	4 3 6 3 7 6	8 4 0 1 9 6	6 4 7 8 4 3	7 3 7 7 7 0

1. Obtener un listado de las personas o elementos que forman la población, luego enumerarla.
2. Determinar el orden que va a seguir para seleccionar los números en la tabla. Puede hacerlo en sentido vertical, horizontal, diagonal, etc. Lo importante es seguir siempre este orden hasta completar la muestra. Suponga que decidió hacerlo en sentido vertical.
3. Determine la fila y columna por la cual iniciará la selección. Por ejemplo suponga que se decidió iniciar por la fila 10, columna 03.
4. Inicie la selección por esa fila y columna teniendo el cuidado de constituir números de 3 dígitos (recuerde que la población es de 800 personas y todas ellas deben tener

la oportunidad de ser escogidas). De esta manera el primer número a ser incluido en la muestra es el 519.

5. Continúe la selección (en sentido vertical como se decidió). El segundo número es el 677, el tercer el 356, y así sucesivamente hasta tener los 30 números (cada uno de ellos representa una persona).
6. Los números que sobrepasen al 800, no son tomados en cuenta pues no corresponden a ningún miembro de la población.

Es recomendable trabajar con muestras probabilísticas puesto que permiten que los resultados obtenidos en ellas puedan ser extrapolados a la población con un margen de confianza determinado.

En relación con las muestras *no probabilísticas*, llamadas también muestras por conveniencia, los elementos son escogidos con base en la opinión del investigador y se desconoce la probabilidad que tiene cada elemento de ser elegido para la muestra. En este tipo de muestreo existen el *intencional* (o deliberado) y los *accidentales* (o por comodidad). En el primero el investigador escoge aquellos elementos que considera típicos de la población. En los segundos, se toman los casos que estén disponibles en el momento. Otro tipo es el *muestreo por cuotas* en el cual el investigador establece una cuota o cantidad de elementos según algunas características de la población, ejemplo, sexo, estado civil y edad, luego escoge los sujetos que encuentra hasta cubrir la cuota establecida. Este último se usa frecuentemente en las encuestas de opinión pública.

En ocasiones se trabaja combinando una elección al azar con una no probabilística: es el caso del muestreo *semiprobabilístico superior* en el cual se conoce la probabilidad de escoger un segmento de la población más no la de un elemento dentro de él (Ejemplo: se seleccionan aleatoriamente las manzanas de una urbanización, dejando a la decisión del entrevistador la elección de las viviendas dentro de las manzanas seleccionadas). En el muestreo *semiprobabilístico inferior* se hace lo contrario (Ejemplo: se seleccionan las manzanas que nos parezcan más típicas de la urbanización y en ellas se escogen al azar las viviendas a estudiar) .

CLASES DE MUESTREO

Existen varias clases de muestreo y la decisión sobre cuál utilizar depende de las características de la población y de la posibilidad de manejar los aspectos técnicos del

diseño de la muestra. A continuación se presenta la descripción del diseño, las indicaciones, ventajas y desventajas de cada clase de muestreo.

Muestreo aleatorio simple (MAS):

A) Descripción: consiste en seleccionar “n” elementos de los “N” que conforman la población de forma que todos ellos tengan igual posibilidad de ser escogidos para integrar la muestra. (convencionalmente se emplea la letra “n” minúscula para indicar el tamaño de la muestra y “N” mayúscula para designar el de la población). El diseño comprende:

- Disponer de una lista numerada de los “N” integrantes de la población.
- Seleccionar en forma aleatoria (por el método de la lotería, la tablas de números aleatorios o por computadora) cada uno de los integrantes de la muestra.

B) Usos: esta clase de muestreo está indicado cuando la población es bastante homogénea en lo que respecta a la variable en estudio (la varianza tiende a cero) y es posible obtener el listado de los elementos de la población. Ejemplo: si deseamos conocer la opinión de los alumnos de esta sección de Investigación en Salud acerca de la dinámica empleada en las clases, es bastante factible que esta opinión sea parecida entre la mayoría de los alumnos, por lo que se puede considerar una población homogénea. Por otra parte, el listado de los alumnos existe, por lo que esta clase de muestreo es aplicable.

C) Ventajas: la sencillez del diseño y de los cálculos estadísticos.

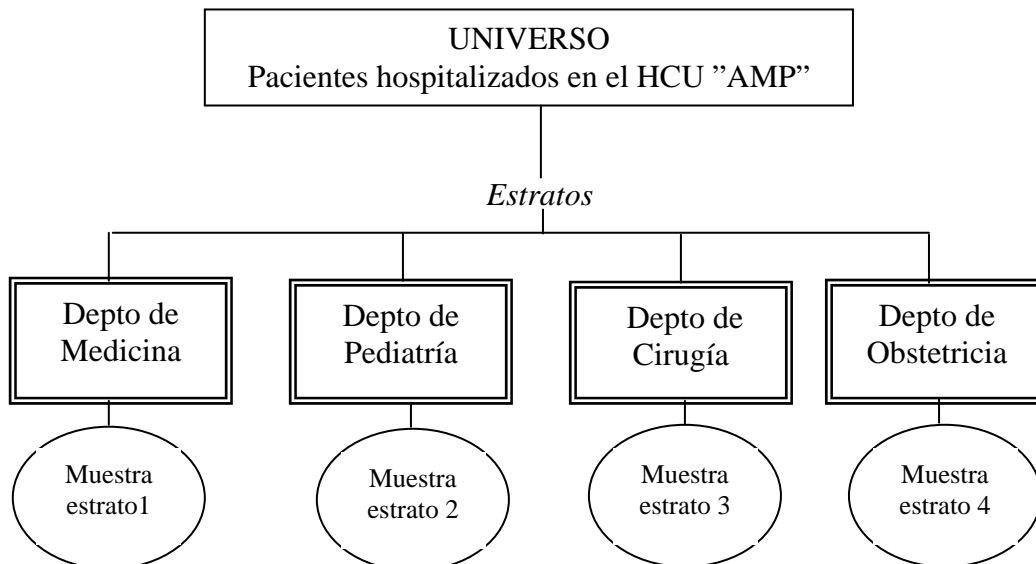
D) Limitaciones: no se puede practicar cuando es imposible obtener la lista con todos los integrantes de la población.

Muestreo aleatorio estratificado (MAE):

A) Descripción: consiste en dividir el conjunto “N” de elementos de la población en varios subconjuntos o estratos, de tal forma que cada estrato formado sea internamente homogéneo, es decir, que sus integrantes se parezcan mucho entre sí en lo que se refiere la variable a estudiar, mientras que los estratos difieren unos de otros. Una vez formados los estratos, se escogen aleatoriamente los elementos a estudiar en cada uno de ellos.

Ejemplo: si deseamos conocer la opinión de los alumnos de Investigación en Salud acerca de la dinámica empleada en las clases, cada sección sería un estrato puesto que los docentes emplean diferentes dinámicas. Dentro de cada sección la variable opinión acerca de la dinámica empleada se supone homogénea, pero entre las distintas secciones (estratos) posiblemente hay diferencias. De igual forma, para conocer el promedio de

estancia hospitalaria en el Hospital Antonio María Pineda”, cada departamento sería un estrato, con un tiempo de hospitalización parecido entre sus pacientes pero que difiere grandemente del de los pacientes de otros departamentos (los pacientes de obstetricia son parturientas que duran pocos días hospitalizadas, a diferencia de los pacientes de cirugía quienes tienen que guardar postoperatorio).



El número de elementos a seleccionar en cada estrato puede establecerse de dos formas: una vez calculado el total de sujetos a estudiar (tamaño de la muestra), se divide este total entre el número de estratos, estudiándose igual cantidad de sujetos en cada uno de ellos (*asignación uniforme*). También puede repartirse este total de la muestra de manera proporcional al tamaño de los estratos, de esa forma los estratos más grandes tendrán una muestra más grande (*asignación proporcional*).

B) Usos: En poblaciones heterogéneas en lo que se refiere a la variable en estudio.

C) Ventajas:

- Se logra mayor precisión en los resultados para un tamaño de muestra dado. Si en el ejemplo anterior en lugar de usar un MAE se hubiese empleado un MAS, el error sería muy grande pues la muestra podría quedar integrada en su mayoría por pacientes de un solo departamento.
- Se obtienen valores por separado para cada estrato.
- Ventajas de índole administrativa (menor costo).

D) Desventajas: Se debe poseer un listado de todos los integrantes de la población.

Muestreo sistemático:

A) Descripción: Consiste en seleccionar aleatoriamente un cierto número “i” que designará, en una lista o población de N elementos, al primero que va a formar parte de la muestra; a continuación, de manera rígida o sistemática, se van tomando los elementos:

$$i + K$$

$$i + 2K$$

$$i + 3K, \text{ así hasta agotar todos los elementos disponibles de la lista, lo}$$

que ocurrirá cuando se llegue al que ocupa el lugar:

$$i + (n - 1)K$$

El primer elemento a estudiar se selecciona aleatoriamente entre el primero de la lista y el número K. Este número K resulta de dividir N/n ($K= N/n$). Por ejemplo, si $N= 1000$ y deseamos estudiar una muestra de $n = 25$ personas, $K= 1000/25$; es decir $K= 40$. Suponga que aleatoriamente escogemos un número entre 1 y 40 y obtenemos el número 20 ($i = 20$), la persona que tenga ese número en la lista será la primera que formará parte de la muestra. La segunda persona a incluir en la muestra será la $i+2K$, ($20+2*40$), es decir, la persona número 100. La tercera será la $i+3K$, ($20+3*40$) = 140, y así se continúa hasta completar las 25 personas de la muestra.

B) Usos:

- Solo en poblaciones heterogéneas y siempre y cuando no exista relación entre la variable a estudiar y la forma como se encuentra distribuida la población.

C) Ventajas:

- Facilidad para extraer la muestra y hacerlo sin errores.
- Es más preciso que el muestreo aleatorio simple cuando la población es heterogénea. Si es homogénea la información se repite de unidad a unidad.

D) Desventajas:

- Poca precisión cuando existe una periodicidad insospechada. En este caso la representatividad de la muestra depende del valor de K. Si K es múltiplo o igual al valor del período, todas las observaciones serán similares. Por ejemplo, tenemos una población con $N = 80$ elementos que repiten los valores 1, 3, 0, 8 de la siguiente forma:

1, 3, 0, 8, 1, 3, 0, 8, 1, 3, 0, 8,1, 3, 0, 8

Si tomamos una muestra sistemática de $n = 20$ elementos, $K = 80/20 = 4$; cualquiera que sea el número i que elijamos entre 1 y 4 para comenzar la selección, la muestra estará constituida siempre por el mismo valor.

- En poblaciones con tendencia lineal, si no se corrige esta tendencia, el muestreo sistemático tiene menor precisión que el estratificado. Ejemplo, si una población está ordenada según su coeficiente de inteligencia y se desea estudiar los resultados de un test de razonamiento abstracto, los resultados probablemente estarán sobreestimados.
- En poblaciones autocorrelacionadas la precisión del muestreo sistemático es poca. Si se desea determinar los efectos de la contaminación provocada por una empresa en la salud de los habitantes de una ciudad, se trata de una población autocorrelacionada pues en la medida en que estén más cerca de la empresa, habrá mayor contaminación.

Muestreo por conglomerados:

A) Descripción:

Consiste en dividir el conjunto de elementos en subconjuntos llamados conglomerados, cuya característica es que internamente son heterogéneos en lo que se refiere a la variable en estudio pero si se comparan varios conglomerados, se nota que son parecidos entre sí. Una vez dividida la población en “N” conglomerados, se escoge en forma aleatoria “n” de ellos y se estudian todos sus elementos. Ejemplo de este tipo de muestreo es la selección de manzanas de viviendas para estudiar las características de los grupos familiares que residen en una urbanización, de escuelas para estudiar el conocimiento de los alumnos sobre algún aspecto que nos interesa, etc.

B) Usos:

Se recomienda en el caso que se desee estudiar localidades más o menos grandes por lo que se le conoce también como muestreo de áreas.

C) Ventajas:

- No se requiere del listado de los elementos de la población, sino solamente de los conglomerados.
- Los costos son mucho menores pues las entrevistas están concentradas.
- Se controla mejor la calidad de los datos.

D) Desventajas:

- Lo complicado de los cálculos.
- Las inferencias que se extraen de esta clase de muestreo no son tan confiables como las de un estudio hecho con muestreo aleatorio.

Muestreo por etapas:

En ocasiones se selecciona una muestra de conglomerados (primera etapa) y estos no son estudiados todos sino que en ellos se conforman otros conglomerados de los cuales se selecciona una muestra que es la estudiada (segunda etapa). Por ejemplo, cuando se seleccionan barrios de una ciudad y en los barrios seleccionados se escogen varias manzanas en las cuales se lleva a cabo la investigación. Esta clase de muestreo se conoce como bietápico. Puede hacerse también en tres etapas o más (muestreo polietápico).

Métodos mixtos:

En la práctica es frecuente la combinación de varias clases de muestreo, como comenzar por formar estratos de la población, dividir cada estrato en conglomerados a seleccionar y en ellos seleccionar las unidades a estudiar por muestreo sistemático. Esto se hace para lograr una mayor representatividad de la muestra a un menor costo. Así, si se desea determinar la opinión de los estudiantes universitarios del país sobre la constituyente universitaria, se pueden formar dos estratos (universidades públicas y universidades privadas), a continuación se selecciona una muestra aleatoria de “n” universidades (conglomerados) en cada estrato y en ellas se escoge sistemáticamente, de la lista de alumnos, una muestra de ellos.

TAMAÑO DE LA MUESTRA

El cálculo del tamaño de la muestra exige una variedad tal de información que excede las posibilidades de este curso, pues para cada clase de muestra y dentro de ellas, para cada tipo de medición, existe un procedimiento particular. Sin embargo, es necesario que la persona que realiza investigaciones conozca los factores de los cuales depende el tamaño de la muestra para que pueda comunicarse con el especialista capacitado para hacerlo. También, existen en la actualidad paquetes como el EPI INFO (3) que nos ayudan a calcular este tamaño. La información que proporcionamos a continuación le ayudará a entenderse con ambos.

Los factores que intervienen en el tamaño de la muestra son:

- a) **El error máximo admisible** (“d” o “e”): determina la precisión de los resultados. Es la máxima diferencia que podemos tolerar entre el valor de la variable obtenido en la muestra y el verdadero valor de ésta en el universo. Para establecer el valor de “e” debemos preguntarnos cuán precisos deseamos que sean los resultados de la investigación, es decir, qué tan exactamente deseamos conocer el valor de la variable. Por ejemplo, para una investigación que tiene como objetivo determinar el porcentaje de trabajadores agrícolas del Valle de Quíbor que desconoce las normas de seguridad para el manejo de plaguicidas y deseamos un error no mayor de 3%, entonces, si en la muestra se obtiene que el porcentaje de trabajadores agrícolas que desconoce las normas de seguridad resulta ser de 19%, estaríamos aceptando que realmente el verdadero valor esté entre 19 ± 3 , es decir, entre 16% y 22%. El error máximo admisible está inversamente relacionado con el tamaño de la muestra (a mayor error, muestras más pequeñas, a menor error muestras más grandes).
- b) **Coefficiente de confianza de la estimación**: para hacer la afirmación anterior sobre los límites dentro de los cuales se encuentra el verdadero valor poblacional, debe conocerse la probabilidad de que estos resultados sean ciertos y esto lo da el coeficiente o nivel de confianza, el cual es la medida probabilística de que el intervalo fijado con “e”, contenga el valor poblacional. En el ejemplo de los trabajadores agrícolas el investigador decidió trabajar con un nivel de confianza de 95%, con lo cual podríamos afirmar con un 95% de certeza que el porcentaje de trabajadores agrícolas del Valle de Quíbor está entre 16% y 22%. Claro está, queda 5% de posibilidades de equivocarnos al establecer esa conclusión. Si trabajamos con 99% de confianza, nos quedaría una posibilidad de equivocarnos de 1%. Por supuesto que nos gustaría tener un gran nivel de confianza (de hecho, en el área de la salud se acostumbra trabajar con 95 o 99% de confianza), pero resulta que el tamaño de la muestra se relaciona directamente con este nivel, en el sentido que mientras más confianza queremos tener en los resultados, el tamaño de la muestra será mayor.

Los valores de “e” y del nivel de confianza los establece el investigador dependiendo de:

- Las características del estudio: ¿qué resultados espera obtener?, ¿para qué se utilizarán los resultados?.
- Del número de sujetos que realmente puede estudiar con los recursos disponibles.

c) **La homogeneidad de la población en lo que respecta a la variable en estudio:** si la variable se encuentra presente en un elevado porcentaje de la población, decimos que ésta es homogénea y bastaría una muestra pequeña para realizar la investigación. Las fórmulas para calcular el tamaño de la muestra contienen una estimación del valor de la variable en la población, el cual tenemos que conjeturar a partir del conocimiento que tengamos de la población, de estudios realizados en poblaciones parecidas a la nuestra, o de estudios pilotos.

A continuación se presentan varios cálculos de tamaño muestral efectuados con el paquete EPIINFO versión 6.04b para que pueda apreciar cómo cambia el tamaño de la muestra dependiendo del nivel de confianza y el error máximo admisible. El ejemplo se refiere a la investigación del porcentaje de trabajadores agrícolas del Valle de Quíbor que desconocen las normas de seguridad para el manejo de plaguicidas. Supongamos que el tamaño de la población es de 8000 trabajadores agrícolas, que la frecuencia esperada es de 20% (pensamos por estudios previos que alrededor de 20% de los trabajadores de esa zona desconocen las normas de seguridad para el manejo de plaguicidas) y que podemos tolerar un error de 3% (23% sería el máximo valor aceptable por error). Con esos datos el paquete EPIINFO nos calcula el tamaño de muestra que debemos estudiar para diferentes niveles de confianza, para que decidamos con cuál trabajar. El reporte que ofrece este programa es:

Population Survey or Descriptive Study Using Random (Not Cluster) Sampling

Population Size	:	8,000
Expected Frequency	:	20.00 %
Worst Acceptable	:	23.00 %
Confidence Level		Sample Size
-----		-----
80 %		282
90 %		454
95 %		629
99 %		1,028
99.9 %		1,552
99.99 %		2,014

Aquí podemos observar que el tamaño de la muestra aumenta en la medida que lo hace el nivel de confianza.

Por otra parte, si en lugar de aceptar un error máximo de 3%, aceptamos un error de 10% (30% es el máximo valor aceptable por error), vemos cómo se modifica el tamaño de la muestra para diferentes niveles de confianza, siendo inferiores a los obtenidos

Population Survey or Descriptive Study Using Random (Not Cluster) Sampling

Population Size	:	8,000
Expected Frequency	:	20.00 %
Worst Acceptable	:	30.00 %
Confidence Level		Sample Size
-----		-----
80 %		26
90 %		43
95 %		61
99 %		105
99.9 %		170
99.99 %		235

El paquete EPIINFO permite calcular el tamaño de la muestra cuando se hace una investigación descriptiva para estimar una proporción o porcentaje en la población (porcentaje de asmáticos en un área, proporción de alumnos con calificaciones superiores a 16 puntos, porcentaje de personas de una comunidad que tienen ingresos inferiores a la canasta básica, etc). También permite calcular el tamaño de la muestra para investigaciones de casos y controles, de cohortes y de corte transversal. Estos son cálculos válidos cuando se usa muestreo aleatorio simple.

EJERCICIOS

1. Para los siguientes casos, diga si se empleó una muestra representativa. Justifique su respuesta. ¿Cómo seleccionaría Ud. una muestra representativa?. Describa el procedimiento.
 - a) “Para determinar qué porcentaje de la población diabética de Barquisimeto, tiene ingresos familiares inferiores a la canasta básica, se revisaron 200 historias médicas seleccionadas aleatoriamente entre todas las historias de pacientes hospitalizados con diagnóstico de diabetes en el Hospital Central Antonio María Pineda”.
 - b) “Los investigadores deseaban determinar el ingreso mensual de las familias de Barquisimeto, para lo cual entrevistaron 15.000 usuarios de los ambulatorios de esta ciudad, seleccionados por muestreo combinado”.

2. Describa el procedimiento de muestreo a seguir en los dos casos siguientes:
 - Una investigación para determinar cuáles son las causas de hospitalización más frecuentes en los distintos departamentos de Hospital Antonio María Pineda.
 - Una investigación para determinar cuál es la cobertura de vacunación en la población menor de 5 años del área de influencia del ambulatorio de Cerrito Blanco, es decir, qué porcentaje de niños tiene el esquema de vacunación completo para su edad .

3. Se desea determinar a través de una muestra, cuál es la frecuencia de caries dental entre los alumnos de educación básica de Barquisimeto.
 - Diga cuál es el universo de esa investigación:
 - Mencione la clase de muestra que se debe emplear para que sea representativa:
 - Describa brevemente cómo se seleccionaría esa muestra.

BIBLIOGRAFÍA CONSULTADA

- 1) Azorín F. Curso de muestreo y aplicaciones. 2ª ed. Caracas: Instituto de Investigaciones Económicas, Dirección de Publicaciones;1970.
- 2) Cochran W. Técnicas de muestreo. México D.F: Compañía Editora Continental, S.A; 1980.
- 3) Dean AG, Dean LA, Burton AH, Dicker RC. Epiinfo, versión 6.04b: a word processing, database, and statistics program for epidemiology on microcomputers. Stone Mountain (Georgia): USD, Incorporated; 1997.
- 4) Ludewig C, Rodríguez A, Zambrano A. Taller de metodología de la investigación. Material de Trabajo. Barquisimeto: Ediciones FUNDAEDUCO; 1998.
- 5) Scheaffer R, Mendenhall W, Ott L. Elementos de muestreo. México D.F: Grupo Editorial Iberoamericano SA de CV; 1986.