

## Diseños de muestreo probabilísticos (I)

Comenzaremos a abordar en este número los procedimientos de muestreo probabilística, en concreto, el muestreo aleatorio simple y el muestreo sistemático.

Los muestreos probabilísticos, como comentamos en el número anterior, son aquellos en los que se utiliza algún sistema de selección aleatoria para garantizar que cada unidad de la población tenga una probabilidad específica de ser seleccionada, cumpliéndose así dos requisitos:

1. Toda unidad tiene una probabilidad de ser elegida
2. Esa probabilidad es conocida de antemano.

Las principales técnicas probabilísticas son:

- Muestreo aleatorio simple
- Muestreo sistemático
- Muestreo estratificado
- Muestreo por conglomerados

### Muestreo Aleatorio Simple.

Es el método conceptualmente más sencillo de muestreo, sin embargo es el menos utilizado ya que requiere contar a priori con un listado de todas las unidades de la población accesible (lo que se denomina marco muestral) a partir del cual se seleccionará la muestra. El procedimiento para realizar un muestreo aleatorio simple consiste en:

1. Disponer de la lista de todos los elementos de la población y a continuación numerarlos consecutivamente.
2. Mediante la utilización de una tabla de números aleatorios o mediante programas informáticos diseñados al efecto obtendremos aleatoriamente los "n" números que componen nuestra muestra en función del tamaño deseado.
3. Iremos al listado donde tenemos numerados los elementos de la población y tomaremos cada uno de los "n" elementos de la población, que correspondan a los números obtenidos anteriormente de forma aleatoria y esos elementos serán los que constituirán nuestra muestra.

El muestreo aleatorio se trata de un método equiprobabilístico pues todos los elementos tienen una probabilidad de ser elegidos, esa probabilidad es conocida de antemano y además esa probabilidad es la misma para todos los elementos.

Por ejemplo si deseáramos conocer la incidencia de caries en alumnos de 1º de Educación Primaria de una Zona Básica de Salud, a la hora de decidir tomar una muestra de esa población podemos optar por un muestreo aleatorio simple. Para ello necesitaremos un listado de los alumnos matriculados en ese curso en los colegios de la zona, posteriormente numeramos consecutivamente a todos los alumnos obteniendo así el marco muestral, es decir el listado de todos los niños de 1º de Educación Primaria (que en nuestro ejemplo tiene un tamaño de  $N = 200$ ), mediante la utilización de una tabla de números aleatorios, o de programas de informática que nos generan aleatoriamente números, obtendremos "n" números aleatorios entre el 1 y el 200 (en nuestro ejemplo  $n = 50$ ) iremos posteriormente al listado donde tenemos numerados a los niños matriculados en 1º y cogeremos los 50 niños que correspondan con los números obtenidos aleatoriamente y ellos constituirán nuestra muestra.

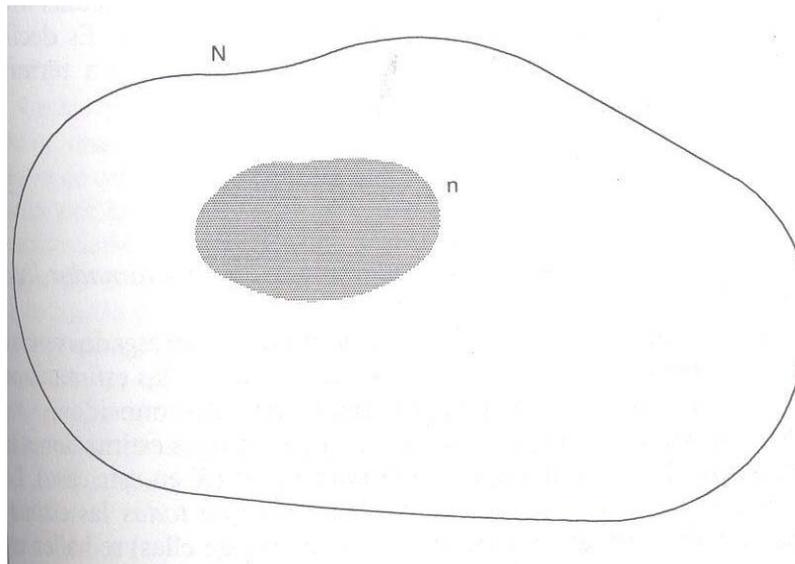


Fig. 1 Muestreo aleatorio simple. Fuente: Silva LC Muestreo para la investigación en ciencias de la salud Edic. Díaz de Santos, Madrid, 1993.

El Muestreo Aleatorio Simple es el método que produce muestras más representativas puesto que sólo interviene el azar y, por lo tanto, no está la muestra sujeta a sesgos del investigador, además de ser el método más simple tanto conceptual como estadísticamente.

Sin embargo cuenta con ciertas desventajas que hace que no sea muy utilizado:

1. Es necesario contar con una lista enumerada de todas las unidades de la población.
2. Los sujetos seleccionados pueden estar muy dispersos por lo que contactar con cada uno de ellos puede resultar costoso en tiempo y dinero.
3. Algunos subgrupos de la población, especialmente aquellos más minoritarios, pueden prácticamente no estar representados en la muestra si ésta es pequeña.

### Muestreo sistemático.

Consiste en seleccionar a los individuos según una regla o proceso periódico. Para ello en primer lugar se debe calcular la **constante de muestreo** (K), dividiendo el total de la población elegible por el tamaño de la muestra deseado ( $N/n = K$ )

La primera unidad (r) se extrae tomando un número al azar entre 1 y la constante de muestreo (K) y a partir de ahí se va sumando la constante de muestreo consecutivamente hasta completar el tamaño de la muestra, (siendo el primer individuo r, el segundo  $r + K$ , el tercero  $r + 2K$  y así sucesivamente hasta completar los "n" individuos).

Este tipo de muestreo tiene la ventaja de no necesitar tener la lista de la población cerrada de antemano para seleccionar la muestra, por ejemplo si deseáramos conocer como se recogen los hábitos tóxicos en las historias que se abrirán durante el próximo año, teniendo un cálculo aproximado del total de historias que se van a abrir podemos calcular la constante de muestreo y el punto de arranque e ir seleccionando las historias que correspondan para nuestra muestra según se vayan abriendo.

Sin embargo este tipo de muestreo no es aconsejable en las situaciones en que las unidades de muestreo están ordenadas por algún criterio periódico y la constante de muestreo puede coincidir con ello. Por ejemplo si queremos conocer la demanda en un centro de salud y elegimos como unidad de muestreo el día de la semana, puede ocurrir que obtengamos como constante de muestreo (K) el número 7, en ese caso estaríamos describiendo siempre la demanda que se produce un día determinado, por ejemplo el sábado, teniendo una elevada probabilidad que la demanda que se produce el sábado sea diferente de la demanda que se produce a lo largo de toda la semana, no siendo por tanto la muestra obtenida representativa de la población que queremos estudiar.

Si en el ejemplo que desarrollamos anteriormente sobre la incidencia de caries en alumnos de 1º de Educación Primaria de una Zona Básica de Salud optásemos por un muestreo sistemático, necesitaríamos calcular primero la constante de muestreo (si  $N = 200$  y  $n = 50$ ) la constante de muestreo K será 4 ( $200/50 = 4$ ) cogeremos aleatoriamente un número entre 1 y el 4 (supongamos que es el 3,  $r = 3$ ) y ese será nuestro punto de arranque. Tomaremos el listado de los alumnos matriculados en ese curso en los colegios de la zona numerados consecutivamente y el primer niño que tomaremos será el que ocupa el puesto 3 ( $r = 3$ ), el segundo 7 ( $r + K$ , es decir  $3 + 4$ ), el tercero 11 ( $r + 2K$ , es decir  $3 + 2 \times 4$ ), el cuarto 15 ( $r + 3K$  es decir  $3 + 3 \times 4$ ) y así sucesivamente hasta tomar los 50 niños que constituyen nuestra muestra.

La figura 2 nos representa gráficamente otro ejemplo de muestreo sistemático con un tamaño de población de 71 ( $N = 71$ ) un tamaño de la muestra de 12 ( $n = 12$ ) la constante de muestreo (K) es de 6 ( $N/n = K$ ) y el punto de arranque un número aleatorio entre 1 y 6 en nuestro ejemplo 3 ( $r = 3$ ), el primer elemento de la muestra será el que ocupa el puesto 3 (r), el segundo el que ocupa el puesto 9 ( $r + K$ ) el tercero el que ocupa el puesto 15 ( $r + 2K$ ) y así sucesivamente hasta completar el último elemento de la muestra que será el que ocupa el puesto 69 ( $r + 11K$ ).

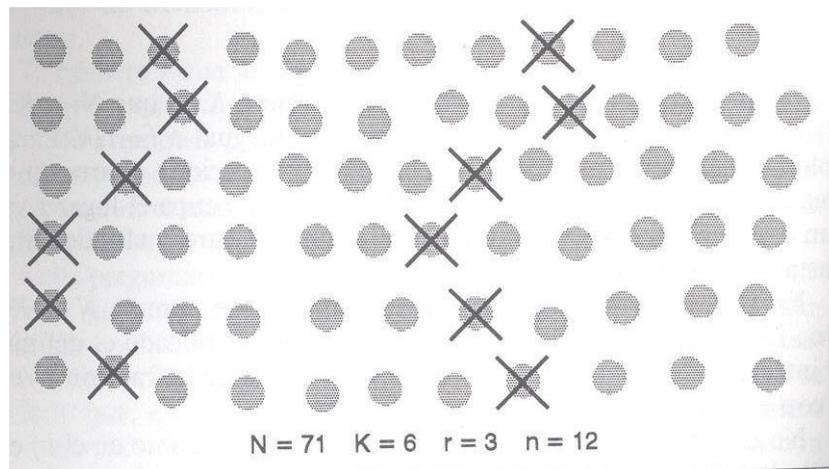


Fig. 2 Muestreo Sistemático

Fuente: Silva LC Muestreo para la investigación en ciencias de la salud Edic. Díaz de Santos, Madrid, 1993.

**Bibliografía consultada:**

1. Polit, Hungler, Investigación científica en ciencias de la salud, 5ª edición, Edit. McGraw-Hill Interamericana, México 2002.
2. Rebagliato M, Ruiz I, Arranz M, Metodología de investigación en epidemiología. Edit. Díaz de Santos, Madrid, 1996.
3. Silva LC, Muestreo para la investigación en ciencias de la salud. Edit. Díaz de Santos, Madrid, 1993.