

El paradigma del 10%: rompiendo mitos con respecto al muestreo en investigación en ciencias médicas

Érika Valverde Altamirano^{1*} y Juan José Romero Zúñiga².

- 1 Programa de Investigación en Medicina Poblacional, Escuela de Medicina Veterinaria, Universidad Nacional, Costa Rica.
- 2 Programa de Investigación en Medicina Poblacional, Escuela de Medicina Veterinaria, Universidad Nacional, Costa Rica. E-mail: jromero@medvet.una.ac.cr.

ABSTRACT

Sampling is defined as the process of selecting a sample from a population of concern. The objective of sampling is to yield knowledge about that population and, in certain cases, about the universe; concepts related to the validity of the study and that will be addressed in the content. The necessity to make a sample, for determining the percentage of items experiencing a disease in a population, frequently arises. How many individuals should be sampled? In many opportunities the answer heard, without too many calculations is: 10%! However, only in few occasions is 10% the sample fraction required to obtain the results that prove a hypothesis. Thus, which is the quantity of items to sample? That number is specific for each particular situation. Sampling comes up as an option to achieve objective investigations in a population, without having to draw samples from all of the units that are part of it; in consequence, the results obtained are good estimates of what happens in the population of concern. This practice economizes resources, facilitates the recollection of information and, with the considerations that each step merits, offers results that are representative for the population; especially for research in medical sciences, where the utilization of a great number of individuals is unlikely. For this reason, focusing on the own aspects of sampling is fundamental, minding the sample size and the design of the study. This document offers several basic elements of sampling for achieving investigations that derive in valid results, representative and applicable in the everyday tasks of those who work in health sciences.

Keywords: *sampling, medical sciences, representativeness, sampling fraction.*

RESUMEN

Se define como muestreo la técnica de seleccionar una muestra a partir de una población de origen. El objetivo de muestrear es que los resultados obtenidos sean extrapolables a la población de origen y, en ocasiones, a la población externa; conceptos relativos a la validez del estudio, que se tratarán con mayor profundidad en el texto.

Con frecuencia, surge la necesidad de hacer un muestreo para establecer el porcentaje de sujetos que presentan un evento o característica dentro de una población particular. ¿A cuántos individuos se debe muestrear? En un considerable número de oportunidades, la respuesta obtenida, sin hacer mayores cálculos matemáticos es: ¡el 10%! Sin embargo, sólo en algunas ocasiones, el 10% es la fracción de muestreo requerida para obtener los resultados que permitan comprobar una hipótesis de estudio. Entonces, ¿cuál es la cantidad de sujetos que deben ser muestreados? Esta, es específica para cada situación particular.

El muestreo surge como una opción para realizar una investigación objetiva en una población, sin tener que tomar muestras a todas las unidades que la conforman, de modo que los resultados obtenidos sean buenos estimadores de lo que sucede en la población de origen. Esta práctica permite economizar recursos, facilita la recolección de la información y, con las consideraciones que cada

Recibido: 20 de diciembre del 2007

Aceptado: 26 de marzo del 2009

* Autora para correspondencia: telefax: (506) 2260-2155, apdo. postal: 304-3000, Heredia., Costa Rica, e-mail: erikava@medvet.una.ac.cr.

caso amerita, especialmente en ciencias médicas, donde muchos estudios no permiten utilizar gran número de sujetos. Por esta razón, es primordial prestar especial atención a los aspectos propios del muestreo, cuidando tanto su diseño, como el tamaño de la muestra.

Se presentan en este artículo, algunos elementos básicos de muestreo para lograr estudios que deriven en resultados válidos, representativos y aplicables en el quehacer diario de quienes se desempeñan en cualquiera de las áreas de la salud, ofreciendo conceptos necesarios y contundentes que derriben el paradigma de un muestreo "universal" del 10% de la población de origen.

Palabras claves: muestreo, ciencias médicas, representatividad, fracción de muestreo.

INTRODUCCIÓN

Es de suma importancia para un investigador, o para un usuario del conocimiento generado mediante la investigación, conocer de estadística en el momento de desarrollar o interpretar un estudio. Independientemente del área de investigación, la aplicación de ciertas herramientas metodológicas le dan solidez y objetividad; además, para que esta sea correcta desde el punto de vista estadístico y metodológico, es imprescindible el manejo de una serie de conceptos básicos. Este artículo tiene como objetivo demostrar, de manera muy concisa, la relevancia del muestreo para realizar un estudio objetivo.

Conceptos básicos

Para iniciar una investigación, es necesario conocer y aplicar una serie de conceptos. Según Noordhuizen *et al.* (1997):

- población externa es cada una de las posibles poblaciones más amplias a las que se desea generalizar los resultados;

- población objetivo es el conjunto de individuos o elementos de los que se requiere evaluar una característica específica;
- población de estudio es aquella de donde se obtiene la información;
- unidad de muestreo es el grupo de elementos no superpuestos que conforman la población completa;
- elemento de muestreo es el individuo en el que se toman las mediciones.

La población de estudio debe delimitarse cuidadosa y completamente antes de recolectar la muestra y, al definirla, se deben describir tanto los elementos que la conforman como los parámetros o características a considerar. Este proceso puede resultar dificultoso, pues en ocasiones la población que se tiene como objetivo no siempre puede estar disponible y el investigador deberá, entonces, reunir información adicional para obtener las respuestas necesarias. La dificultad es mayor si el número de elementos es infinito, si se trabaja con registros y algunos han sido destruidos o tienen información incompleta o inconsistente, si al tomar mediciones

los individuos sufren daños o son lastimados, si están muy dispersos o si la realización del estudio es muy costosa (Scheaffer *et al.*, 1987; Hernández, 1997). Como solución a esto, se toman mediciones solamente a una parte de la población, la cual se denomina muestra; que es un subconjunto de casos o individuos seleccionados con la intención de inferir propiedades fiables y comparables a la población completa, por lo que deben ser representativas de esta; así, el muestreo es la técnica utilizada para seleccionarlas (Calero, 1978).

Representatividad

Un aspecto crucial para lograr resultados objetivos en el estudio, es la representatividad de las muestras, definida por su precisión y validez. El azar no es una garantía de esto; así, el muestreo pretende asegurar la representación de cada grupo en la muestra. La precisión es la carencia del error aleatorio, que se refiere a la reducción del error debido al azar. Puede ser mejorada incrementando el tamaño de la muestra y modificando el diseño del estudio para aumentar la eficiencia con que se obtiene la información. La validez es la carencia del error sistemático (sesgo), lo que hace referencia al grado de similitud entre los tres tipos de poblaciones. La representatividad se puede desagregar en dos componentes: la validez interna, extraída de las conclusiones provenientes de la población de estudio (población objetivo), y la validez externa

o generalizabilidad, que ocurre cuando estas conclusiones se aplican a los sujetos fuera de la población de estudio, es decir, la capacidad de generalizar los resultados hacia la población externa (universo). No se puede hablar de validez externa si no existe validez interna (Rothman, 1987; Romero, 2007).

Tamaño de la muestra

Cada investigación tiene un tamaño de muestra ideal que permite confirmar la hipótesis con una seguridad aceptable. Para calcularlo, existen fórmulas estadísticas apropiadas y específicas para cada tipo de estudio y elementos que se deben considerar:

- **Proporción:** es la relación entre dos variables, en la que el numerador está incluido en el denominador, y es normalmente expresada en notación fraccionaria. Cuando es expresada en fracción de 100, se denomina porcentaje (Romero, 2006).
Por ejemplo, si en una población de 25,000 habitantes se diagnostican 1,800 con diabetes, se dice que la proporción de diabetes en esa población es de $1,800/25,000 = 0.072$ (7.2%).
- **Prevalencia:** es el porcentaje de individuos que expresan una característica o evento (infección, enfermedad, accidente, etc.) determinado dentro de una población. Se puede describir como la posibilidad de que un elemento escogido al azar, sufra, padezca u obtenga dicha característica

o evento en un momento específico (Colimón, 1990). La prevalencia no se utiliza con fines predictivos y su determinación es apropiada para la medición de procesos de carácter crónico, pero no para aquellos de carácter agudo o transitorio.

Así, de una muestra aleatoria de 250 habitantes de la población mencionada en el ejemplo anterior, se diagnosticaron 19 con diabetes. En este caso, la prevalencia de diabetes para esa población es de $19/250 = 0.076$ (7.6%).

- Media (media aritmética): es el promedio de una cantidad finita de números, es decir, la suma de todos ellos dividida entre la cantidad de unidades muestreadas. Es uno de los principales estadísticos muestrales. El símbolo μ es utilizado para representar la media aritmética de una población y \bar{X} para la de una muestra (Wikipedia, 2007).
Por ejemplo, si en una habitación hay 5 personas, con estaturas (metros) de 1.72, 1.68, 1.75, 1.70 y 1.77, la media de sus estaturas sería la suma de todas (8.62) dividida entre 5 ($\bar{X} = 1.72$ m).
- Error (muestral, de estimación, estándar): es producto de la aleatoriedad del proceso de muestreo y, no obstante la experiencia y las buenas técnicas de investigación aplicadas, este siempre existirá. La meta que se debe perseguir en todo estudio es la exactitud o precisión en la medición, en otras palabras, estimar el valor del parámetro objetivo con

poco error (Rothman, 1987; Noordhuizen *et al.*, 1997). Esto se debe a que el estudio de la muestra, aunque sea representativa de la población objetivo, no proporciona resultados idénticos al estudio de la población completa. Así, al seleccionar aleatoriamente dos muestras del mismo tamaño, de una misma población, para realizar una encuesta de forma simultánea, se notará que los resultados, aunque sumamente parecidos, no serán idénticos.

El error muestral se debe contemplar a la hora de calcular el tamaño de la muestra.

- Confianza: es la probabilidad de que la evaluación generada se ajuste a la realidad, asegurando que la medición se acerque al valor real del parámetro en estudio. Este concepto está íntimamente ligado al intervalo de confianza, que se define como la probabilidad de que al repetir el muestreo un determinado número de veces en la misma población, las estimaciones de un porcentaje de los muestreos se encontrarán dentro de un rango de valores que, sin ser idénticos, no son estadísticamente diferentes. Tal rango representa un grupo de valores que es consecuente con los datos observados. La amplitud del intervalo de confianza depende de la cantidad de variabilidad que exista en los datos, pero también de un valor seleccionado subjetivamente, que especifica el grado de relación entre los límites del intervalo y dichos datos (Rothman, 1987). Así, por ejemplo, si se realizan 100 muestreos de una

población y el indicador que se obtiene en 95 de ellos está dentro de ese rango, se hablaría, en este caso, de un intervalo de confianza del 95%. En general, el nivel de confianza mínimo aceptable es 90% y el nivel máximo real es 99.5%.

Rol de las hipótesis en el proceso de muestreo

Antes de iniciar un análisis se deben establecer hipótesis, que son proposiciones acerca de la población. Estas conjeturas motivan la investigación, pues a menudo, los estudios se llevan a cabo para determinar si pueden o no sostenerse una vez sometidas al proceso de prueba que exige el método científico. La hipótesis nula es la teoría de una fingida igualdad entre características que se suponen verdaderas dentro de la población. Es de tipo estadística, pues puede ser evaluada mediante técnicas adecuadas y, a pesar de sus limitaciones, el término “estadísticamente significativo” se percibe como una etiqueta de “garantía de calidad”. En general, se establece con el propósito de ser rechazada al concluir el estudio; si esta no se rechaza, se dice que los datos sobre los que se basó la prueba no proporcionaron evidencia suficiente (Daniel, 2002). Con respecto al proceso de prueba de la hipótesis nula se pueden cometer dos errores:

- Error tipo I (α): rechazar una igualdad real o aceptar una diferencia inexistente.

- Error tipo II (β): aceptar una igualdad falsa o rechazar una diferencia existente.

El proceso de aceptación o rechazo de la hipótesis lleva, entonces, implícito el riesgo de cometer alguno de los errores mencionados. La significancia estadística es la probabilidad de cometer un error tipo I (α), que normalmente es fijada en 5%. Además, es importante señalar que ella depende de dos componentes fundamentales: la magnitud de la diferencia entre las dos variables en estudio y el tamaño de la muestra (Daniel, 2002). Se reitera, así, la importancia de la precisión a la hora del muestreo y se insta, conjuntamente, a considerar como esencial la dimensión del contraste de las variables, para asegurar una investigación estadísticamente correcta y fiable.

Otro elemento para calcular el tamaño deseable de un estudio consiste en utilizar los valores de la potencia (poder) que es una contraparte del error tipo II (β), pues es la probabilidad de detectar u observar, en la muestra, un efecto que existe en la población, en otras palabras, el poder, es la probabilidad de rechazar una hipótesis nula, falsa (Noordhuizen *et al.*, 1997; Fariñas *et al.*, 2001).

Se desprende, entonces, hasta este punto, que estudios basados en hipótesis particulares requieren de tamaños de muestra específicos para cada caso, de modo que los resultados sean fiables

La desviación estándar se calcula utilizando la fórmula, $SD = \sqrt{P(1-P)}$ donde P es la prevalencia esperada.

Ejemplo: Considerando el ítem anterior y estableciendo la prevalencia en 50% (para que el tamaño de la muestra sea el máximo), se tiene que $SD = \sqrt{0.5 * 0.5} = 0.5$. El valor de t de Student para el nivel de confianza señalado (0.95) es 1.96 y el error aceptado 5%.

$$\text{En este caso, } n = \left(\frac{1.96 * 0.5}{.05} \right)^2 = 384.$$

Esta fórmula es válida si el valor de n resulta menor al 10% de la población (N), que suele suceder cuando la población es mayor a 1,000 individuos, pero, cuando la población es pequeña (con menos de 1,000 sujetos), se debe hacer un ajuste, de modo que:

$$n_a = \frac{n}{1+f}$$

En la que:

n_a = tamaño de muestra ajustado
 n = tamaño de muestra calculado originalmente utilizando la fórmula anterior

$$f = \text{fracción de muestreo} \left(\frac{n}{N} \right)$$

Ejemplo: Si el caso anterior necesitara

$$\text{ajuste, entonces } n_a = \frac{384}{1 + \left(\frac{384}{1,000} \right)} = 277$$

3. Determinar diferencia entre proporciones

Es una práctica común hacer comparaciones entre poblaciones, especialmente para indicadores globales de la salud de las poblaciones, como son las prevalencias. Una simple comparación de prevalencias entre dos poblaciones, que deriva en una frecuencia notablemente mayor en una de ellas, puede dar origen a la formulación de hipótesis causales en busca de factores asociados a ese evento. Sin embargo, para tener certeza de que tal diferencia existe, se debe muestrear una cantidad mínima que permita comprobarla. Entonces:

$$n = \left(\frac{Z_\alpha * Z_\beta}{p_1 - p_2} \right)^2 * (p_1q_1 + p_2q_2)$$

En la que:

Z_α, Z_β = valor de la distribución normal estándar a un determinado nivel de confianza y de poder

p_1, p_2 = proporciones estimadas en las poblaciones 1 y 2

$q_1, q_2 = 1 - p_1$ y $1 - p_2$, respectivamente

Ejemplo: Si en la segunda población, el porcentaje de estudiantes fumadores equivale al 42.5%, así $n =$

$$\left(\frac{1.96 * 0.84}{0.385 - 0.425} \right)^2 * (0.385 * 0.615) + (0.425 * 0.575) = 2,363.$$

* En este caso se utiliza 1.96 que es el valor en la tabla de distribución normal de probabilidades de Z, a una confianza de 95%, la potencia establecida es 80% y a dos colas.

4. Comprobar un promedio

Si se desea obtener o comprobar la media aritmética de una variable, se debe tomar en cuenta el tamaño de muestra necesario para obtener un valor significativo de esa media. Así,

- cuando se conoce el tamaño de la población:

$$n = \frac{N * t_{\alpha}^2 * SD^2}{(N - 1) * L^2 + t_{\alpha}^2 * SD^2}$$

Donde:

n = tamaño de muestra necesario
 t_{α} = valor de la t de Student para el nivel de confianza especificado
 SD = desviación estándar esperada
 N = tamaño de la población
 L = error absoluto esperado (o precisión requerida)

Ejemplo: Tomando los 1000 estudiantes universitarios y los valores establecidos y obtenidos en los ejemplos anteriores, si se indaga sobre la edad a la que los estudiantes iniciaron el fumado. En este caso se señala 5 como SD:

$$n = \frac{1000 * 1.96^2 * S^2}{(999 * S^2) + (1.96^2 * S^2)}$$

(redondeado a 4)

Si el tamaño de la población es desconocido:

$$n = \left(\frac{t_{\alpha} * SD}{L} \right)^2$$

Donde:

n = tamaño de muestra necesario
 t_{α} = valor de la t de Student para el nivel de confianza especificado
 SD = desviación estándar esperada
 L = error absoluto esperado (o precisión requerida)

Ejemplo: Si el tamaño de la población de estudiantes universitarios fuese desconocido, entonces:

$$n = \left(\frac{1.96 * S}{S} \right)^2 = 3.84 \text{ (redondeado a 4)}$$

5. Comprobar diferencia de promedios

El investigador puede estar interesado en estudiar diferencias o similitudes entre grupos, comparando las medias de un parámetro. En este caso:

$$n = \frac{2 * (Z_{\alpha} + Z_{\beta})^2 * SD^2}{\delta^2}$$

En la que:

Z_{α}, Z_{β} = valores especificados de Z para confianza y poder
 SD = desviación estándar de la variable
 δ = diferencia estimada entre medias

Ejemplo: Si en el segundo grupo de estudiantes que formaron parte de la investigación, se quiere saber la edad a la que iniciaron el fumado los estudiantes de sexo masculino en comparación con fumadores de sexo femenino es, así utilizando los valores,

$$n = \left(\frac{2 * (1.96 + 0.842)^2 * 25}{3^2} \right) = 43.$$

Por lo tanto, se requiere de 43 sujetos por grupo para hacer la comparación.

$$n = \frac{\left(t_{\alpha} \sqrt{\left(1 + \frac{1}{c}\right) pm(1 - pm)} + t_{\beta} \sqrt{p_1(1 - p_1) + \frac{p_0(1 - p_0)}{c}} \right)^2}{(p_1 - p_0)^2}$$

En la que:

$$p_1 = \frac{p_0 * OR}{1 + p_0 * (OR - 1)} \text{ y } pm = \frac{p_1 + cp_0}{1 + c}$$

Donde:

n = número de muestra necesario

t_{α} = valor de t de Student a un nivel de confianza dado

t_{β} = valor de t de Student a un nivel de poder dado

p_0 = proporción expuesta en el grupo control

p_1 = proporción calculada expuesta en el grupo caso

6. Comprobar asociación entre evento y determinantes (factores) de enfermedad

Esta se utiliza en los casos para los que se desea determinar la relación entre la exposición a un factor y la manifestación de un evento, padecimiento, enfermedad o característica. La fórmula que permite obtener el tamaño adecuado de muestra es:

pm = proporción expuesta en la población
 1-pm = proporción no expuesta en la población
 c = número de controles por caso
 OR = OR estimado

Ejemplo: Si se tratara de relacionar el factor "fumado" con el evento "cáncer de pulmón" en los estudiantes universitarios que han formado parte de los ejemplos anteriores, estableciendo el porcentaje de estudiantes no fumadores expuestos (grupo control) a 15%, en relación con controles por caso de 2 y OR 2:

$$p_1 = \frac{0.15 * 2}{1 + 0.15 * (2 - 1)} = 0.26 \text{ y } pm = \frac{0.26 + (2 * 0.15)}{1 + 2} = 0.186$$

$$n = \frac{\left(1.96\sqrt{\left(1 + \frac{1}{2}\right)0.186(1 - 0.186)} + 0.842\sqrt{0.26(1 - 0.26) + \frac{0.15(1 - 0.15)}{2}}\right)^2}{(0.26 - 0.15)^2} = 152$$

Métodos de selección de muestras

Tan importante como el tamaño de la muestra, es el método de muestreo. Un error de muestreo usualmente ocurre cuando no se lleva a cabo la encuesta de la población completa y la muestra seleccionada no es representativa. El error sistemático es medido en términos de probabilidad y para evitarlo se deben utilizar métodos aleatorios a la hora de hacer el muestreo. Cabe destacar que respuestas inconsistentes, incompletas o no determinadas, no son consideradas errores muestrales. Existen cuatro tipos de muestreo aleatorio, según Rodas *et al.* (2006):

- a. **Muestreo aleatorio simple:** al utilizar este método, las unidades incluidas en la muestra son seleccionadas al azar, por lo que cada una de ellas tiene la misma probabilidad de ser seleccionada. Este método, aunque asegura que la muestra sea representativa, es difícil de realizar, debido a que, a menudo, no todos los elementos de la población objetivo están incluidos en el marco muestral.
- b. **Muestreo sistemático:** en este método, los elementos son seleccionados de manera ordenada. Esto depende del número de elementos en la población y el tamaño de la muestra. Se elige al azar un número entre 1 y

100, por ejemplo, 25; por lo tanto, los individuos escogidos serían: 25, 125, 225, y así sucesivamente. Por supuesto, todos los individuos tendrán que estar debidamente identificados y habrá mejores resultados si se encuentran ordenados al azar.

- c. **Muestreo estratificado:** para llevarlo a cabo, se debe dividir la población en grupos, llamados estratos, que son más homogéneos que la población entera. Luego, los elementos de una muestra son seleccionados al azar mediante un muestreo sistemático de cada estrato. Usualmente con este método se tiene más precisión (o menos error muestral) que si se aplicara el método aleatorio simple. El número de individuos seleccionados puede o no ser proporcional al tamaño del estrato en relación con la población.
- d. **Muestreo de conglomerados:** se debe dividir la población en grupos de individuos que son aptos para el muestreo. Luego se selecciona una porción de los grupos mediante un método sistemático y posteriormente se hace lo mismo con la porción seleccionada para obtener la muestra. Aunque no todos los grupos son muestreados, cada uno tiene la misma probabilidad de ser elegido, lo que hace a la muestra aleatoria. Con este tipo de muestreo, se corre el riesgo de

aumentar el error muestral, debido a que todos los individuos dentro de cada grupo suelen ser iguales, pero si el tamaño de la muestra se aumenta y si dentro de cada grupo la variación es tan grande como la de la población entera, se puede obtener la misma precisión que con una muestra aleatoria simple.

Para el estudio de los determinantes de un evento, se podría utilizar un método no probabilístico. Estos no dependen del azar sino del juicio personal. Suelen ser subjetivos, basados en la experiencia o en la certeza de que los individuos de la población cooperarán con el estudio. Se suele referir a él como “muestreo a conveniencia” o “muestreo de juicio”. Está basado en la teoría de que la probabilidad no puede ser empleada para medir el error muestral. Las ventajas de este método son la facilidad con que se obtiene y los bajos costos de su aplicación (Rodas *et al.*, 2006).

CONCLUSIONES

- En pocos casos, el 10% es la fracción de muestreo necesaria para comprobar una hipótesis de trabajo en un estudio diseñado con rigor científico.
- El tamaño de muestra de un estudio es específico para cada caso particular. No existe una regla empírica que garantice que el 10% corresponde a la cantidad requerida de sujetos o unidades para un estudio.
- Un tamaño de muestra adecuado no es suficiente por sí solo para garantizar la validez y la representatividad de un estudio; se requiere además de un diseño de muestreo que garantice la representatividad de las unidades muestrales.
- Los estudios científicos, en sus diversas modalidades, deben ser sometidos a escrutinio, valorando los alcances de sus resultados y conclusiones, teniendo como referencia el tamaño de la muestra y el diseño del muestreo, verificando la presencia del error aleatorio y el sistemático en su mínima expresión.

REFERENCIAS BIBLIOGRÁFICAS

Calero, A. 1978. *Técnicas de Muestreo*. Editorial Pueblo y Educación, La Habana.

Colimón, K. 1990. *Fundamentos de epidemiología*. 2a. ed. Ediciones Díaz de Santos, Madrid.

Daniel, W. 2002. *Bioestadística. Base para el análisis de las ciencias de la salud*. Limusa, México.

Domenech, J.M. 1999. *Métodos estadísticos en Ciencias de la Salud*. Signo, Barcelona.

Fariñas, C.; M.C. Fariñas; J. Llorca & M. Delgado. 2001. Factores de riesgo de sepsis nosocomial: un estudio de

- casos y controles. *Med. Clin.* 116: 765-769.
- Hernández, R. 1997. *Metodología de la Investigación*. McGraw-Hill, México.
- Jekel, J.F.; J.G. Elmore & D.L. Katz. 1996. *Epidemiology Biostatistics and Preventive Medicine*. W.B. Saunders Company, Philadelphia.
- Noordhuizen, J.P.T.M.; K. Frankena; C.M. van der Hoofd & E.A.M. Graat. 1997. *Application of quantitative methods in veterinary epidemiology*. Wageningen Pers, The Netherlands.
- Rodas, O.; C. Bolaños; S. Barrientos; J. Muñoz & M. Catalán. 2006. Teoría básica del muestreo [en línea]. www.monografias.com. (Consulta: julio del 2007).
- Romero, J.J. 2007. Elementos para análisis de datos. No publicado.
- Romero, M. 2006. Proporcionalidad [en línea]. www.encyclopedia.us.es. (Consulta: julio del 2007).
- Rothman, K. 1987. *Epidemiología moderna*. Ediciones Díaz de Santos, Madrid.
- Sánchez, R. & J. Torres. 1989. *Estadística Elemental 7*. Editorial Pueblo y Educación, La Habana.
- Scheaffer, R.; W. Mendenhall & L. Ott. 1987. *Elementos de muestreo*. Grupo Editorial Iberoamérica, México.
- Taro, Y. 1989. *Elementary Sampling Theory*. Editorial Pueblo y Educación, La Habana.