

SOFTWARE LIBRE QUE CALCULA EL TAMAÑO DE MUESTRA MEDIANTE MÉTODOS DE MUESTREO PROBABILÍSTICO

Juan Ruiz Ramírez¹

Gabriela E. Hernández Rodríguez²

Christian Pérez Salazar³

RESUMEN

Un problema recurrente en los estudios observacionales es la determinación del tamaño de muestra y la estimación puntual o por intervalo de confianza de los parámetros de la población objetivo, es por ello que se elaboró un software que calcula el tamaño de muestra mediante métodos de muestreo probabilístico. Los métodos empleados son el muestreo irrestricto aleatorio, sistemático, estratificado y de conglomerados. Este software se desarrolló con el lenguaje de programación JAVA. Adicionalmente a la determinación del tamaño de muestra para cada caso, se implementaron los métodos para seleccionar los elementos de muestreo que conformarán la muestra obtenidos de manera aleatoria. Se concluye que este software permite calcular el tamaño de muestra de manera fácil, precisa y rápida y es de utilidad en los estudios en los cuales se requiere obtener información primaria, principalmente en los estudios descriptivos.

Palabras clave: Software libre, la innovación tecnológica, métodos de muestreo, tamaño de muestra, estudios descriptivos u observacionales.

JEL: A120 - Relation of Economics to Other Disciplines;

¹ Dr. en Ciencias, profesor de la Facultad de Economía de la Universidad Veracruzana. Av. Xalapa s/n, esq. Manuel Ávila Camacho, 91020, Xalapa, Veracruz, México. jruizuv@gmail.com. (228) 8-17-42-00, ext. 14210

² Lic. en Economía, estudiante del Doctorado en Finanzas Públicas en la Facultad de Economía de la Universidad Veracruzana. gabyrendera@yahoo.com.mx.

³ Maestro en Redes y Sistemas Integrados. Técnico Académico del Instituto de Investigaciones Multidisciplinarias de la Universidad Veracruzana. chperez@uv.mx.

ABSTRACT

A recurring problem in observational studies is the sample size determination and the point estimate or confidence interval of the parameters of the target population, which is why we developed software that calculates the sample size by probability sampling methods.

The methods used are unrestricted random sampling, systematic, stratified and conglomerates. This software was developed in Java programming language. In addition to determining the sample size in each case, the methods were implemented to select sampling elements that make up the sample obtained at random. We conclude that this software can calculate the sample size easily, accurate and fast and is useful in studies in which primary information is required, mainly in descriptive studies.

Keywords: Free software, technological innovation, sampling methods, sample size, descriptive or observational studies.

INTRODUCCIÓN

La educación está asociada directamente al nivel de desarrollo económico y social de un país y para crear las condiciones para su óptima realización implica una fuerte inversión en recursos financieros, tecnológicos y humanos, aunado a una estrecha vinculación con el sector productivo de la economía para determinar e identificar áreas estratégicas que potencializan su valor agregado en la producción en un contexto internacional cada vez más complejo y competitivo.

Es por ello que la educación que se imparte en las Instituciones de Educación Superior (IES) debe ser de calidad y esto se logra a través de la calidad del empleo de sus egresados, debido a esto, se requiere poner cuidado en sus técnicas de enseñanza para mejorar la situación actual del sistema educativo y buscar una total vinculación entre la educación y el campo laboral (Ruiz-Ramírez y Hernández-Rodríguez, 2011).

En la estrategia de enseñanza-aprendizaje se emplean los medios tecnológicos, entre ellos, el uso de software estadístico, los cuales, en su mayoría no determinan los tamaño de muestra, ni realizan la estimación de los parámetros, motivo por el cual se tienen que consultar algunos software de uso libre, como son Muestreo (Rodríguez, 1999); Muestreo estadístico, tamaño de muestra y estimación de parámetros (Montesinos *et al.*, 2012, <http://www.ucol.mx/docencia/facultades/economia/muestreo.php>), SOTAM (<http://www.aloj.us.es/vmanzano/programas/dos/>); Epidat 3.1 (<http://aprendeenlinea.udea.edu.co/lms/moodle/course/view.php?id=322>) y también se citan los que requieren un pago por licencia anual como es el software POSDEM (<http://posdem.com/>), o los que requieren un costo como son por ejemplo: ForSampling 1.3 (<http://www.forsampling.com/>).

Con respecto al software Muestreo que realizó (Rodríguez, 1999) en la facultad de estadística e informática de la Universidad Veracruzana, sólo se puede acceder bajo MS-DOS y en computadoras antiguas, motivo por el cual se requieren softwares elaborados en lenguajes de programación actuales como es el JAVA.

En el curso de la Experiencia Educativa Estadística I que se imparte en la facultad de Economía de la Universidad, se tiene en su programa de estudios, el tema de Muestreo el cual comúnmente no se explica a detalle en el Periodo académico, y aunado a que se hacen los cálculos del tamaño de muestra al sustituir los valores requeridos en las fórmulas para cada estimador y método de muestreo empleado, esto no lo hace atractivo para los estudiantes.

Al mencionar las fuentes de información empleadas en las investigaciones, se puede utilizar información secundaria que proviene de fuentes oficiales como son el INEGI, el Banco de México, SIAP-SAGARPA, etc., sin embargo cuando la información requerida no existe o no está disponible, entonces se debe generar, y esto comúnmente se realiza a través de la aplicación de encuestas y el problema que se presenta es la determinación del tamaño de muestra y la elección de los elementos que conformarán la muestra. Una manera de resolverlo es a través de la aplicación de las fórmulas correspondientes, en función al método probabilístico y al estimador requerido. Sin embargo, en ocasiones se dificulta realizar las operaciones aritméticas por lo que emplear un software ex-profeso es otra opción para resolver este problema.

El objetivo que se plantea es presentar un software libre que calcula el tamaño de muestra mediante métodos de muestreo probabilístico y realiza la selección aleatoria de los elementos que conformarán la muestra y una vez aplicada la encuesta, se realiza la estimación de los parámetros.

Métodos de Muestreo mediante el Software Muestreos Estadísticos

El presente programa estadístico computacional se elaboró con la plataforma Java™, Standard Edition 7, Development Kit (Java, 2013), que es un entorno de desarrollo para crear aplicaciones y componentes utilizando el lenguaje de programación Java; apoyados con el entorno de desarrollo integrado (IDE por sus

siglas en inglés) de código abierto, NetBeans, en su versión 7.2.1, el cual se encuentra disponible para Windows, Mac, Linux y Solaris (NetBeans, 2013).

Para utilizar el programa, es necesario tener instalado el ambiente de ejecución de Java en su computadora (también se hace referencia al software de Java como: Java Runtime, Runtime Environment, Runtime, JRE, máquina virtual de Java, máquina virtual, Java VM, JVM, VM, plugin de Java, complemento de Java o descarga de Java), el cual puede ser descargado de forma gratuita de su página oficial (<http://www.java.com/es/download/>). Después de instalar java y extraer los archivos del software “Muestreos Estadísticos” en alguna carpeta se procede a la ejecución del programa de Muestreos Estadísticos haciendo doble clic sobre el archivo “MuestreosEstadisticos.jar”. Independiente de la plataforma, puede ser usado en Windows, Mac, Linux, etc.

En la imagen 1 se muestra la portada del software donde se describe brevemente el software y en esta versión del programa se considera el cálculo de cuatro diferentes tipos de muestreo probabilístico: Muestreo Aleatorio Simple, Muestreo Aleatorio Estratificado, Muestreo Sistemático y Muestreo por Conglomerados. Adicionalmente, para cada muestreo se permite realizar la estimación de la media poblacional, total poblacional y proporción poblacional, así como el cálculo de los tamaños de muestra para cada caso. Cuenta además con un apartado para generar números aleatorios (ordenados y sin ordenar) en base al total de población y tamaño de muestra especificados. Este software libre, se puede solicitar a los autores y además se cuenta con el Registro público de Derechos de Autor en México, número 03-2014-040910331100-01 del 19 de junio de 2014.



Imagen 1. Presentación del software “Muestreos Estadísticos”.

En la figura 1 se muestra el diagrama de flujo del Muestreo Aleatorio Simple para realizar el cálculo de la media poblacional, total poblacional y proporción poblacional. Algo similar es para los casos de los demás muestreos (Ruiz, Pérez y Hernández, 2014).

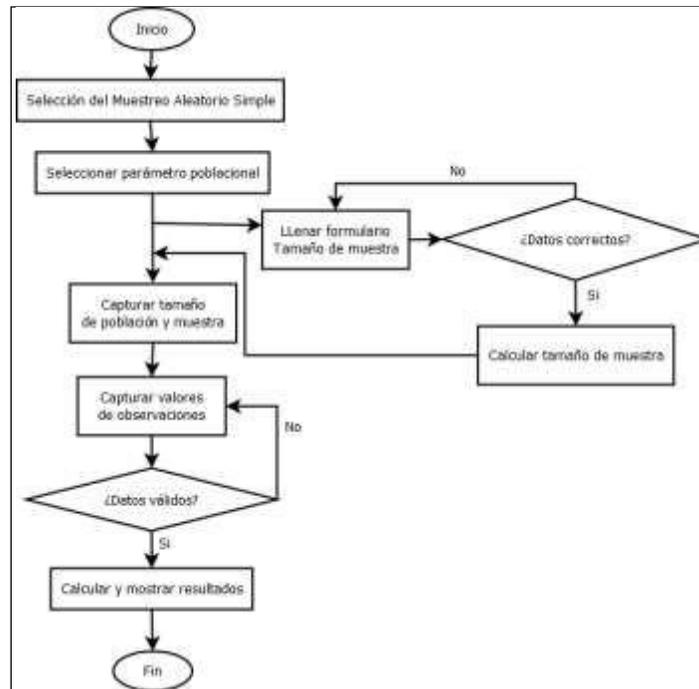


Figura 1. Diagrama de Flujo del Muestreo Aleatorio Simple

Las fórmulas empleadas en la determinación del tamaño de muestra o del intervalo de confianza para los diferentes estimadores y para el muestreo irrestricto aleatorio o estratificado, se emplearon las fórmulas de los Cuadros 1 y 2, respectivamente.

Cuadro 1. Fórmulas para calcular el tamaño de muestra cuando se estiman μ, τ y p en el muestreo irrestricto aleatorio.

Parámetro	Estimador	Varianza Estimada	Tamaño Muestral
μ	$\hat{\mu} = \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$	$\hat{V}(\bar{y}) = \frac{S^2}{n} \frac{N-n}{N}$	$n = \frac{N\sigma^2}{\left(\frac{N-1}{4}\right) B^2 + \sigma^2}$
τ	$\hat{\tau} = Ny$	$\hat{V}(\hat{\tau}) = N^2 \frac{S^2}{n} \frac{N-n}{N}$	$n = \frac{N\sigma^2}{\left(\frac{N-1}{4N^2}\right) B^2 + \sigma^2}$
p	$\hat{p} = \frac{1}{n} \sum_{i=1}^n y_i$	$\hat{V}(\hat{p}) = \frac{\hat{p}(1-\hat{p})}{n-1} \frac{N-n}{N}$	$n = \frac{Np(1-p)}{\left(\frac{N-1}{4}\right) B^2 + p(1-p)}$

Cuadro 2. Fórmulas para calcular el tamaño de muestra cuando se estiman μ, τ y p en el muestreo estratificado.

	Estimador	Varianza Estimada	Tamaño Muestral
μ	$\bar{y}_{est} = \frac{1}{N} \sum_{i=1}^L N_i \bar{y}_i$	$\hat{V}(\bar{y}_{est}) = \frac{1}{N^2} \sum_{i=1}^L N_i^2 \frac{S_i^2}{n_i} \frac{N_i - n_i}{N_i}$	$n = \frac{\sum_{i=1}^L N_i^2 \sigma_i^2}{\sum_{i=1}^L N_i \sigma_i^2 + N^2 D}$
τ	$\hat{\tau}_{est} = N \bar{y}_{est}$	$\hat{V}(\hat{\tau}_{est}) = N^2 \hat{V}(\bar{y}_{est})$	$D = \begin{cases} \frac{B^2}{4} & \text{para la media} \\ \frac{B^2}{4N^2} & \text{para el total} \end{cases}$ $\omega = \frac{n_i}{N \sigma} = \frac{\frac{c_i}{N}}{\frac{\sum_{i=1}^L N_i \sigma_i}{N \sigma}} = \frac{c_i}{\sum_{i=1}^L N_i \sigma_i}$
p	$\hat{p}_{est} = \frac{1}{N} \sum_{i=1}^L N_i \hat{p}_i$	$\hat{V}(\hat{p}_{est}) = \frac{1}{N^2} \sum_{i=1}^L N_i^2 \frac{\hat{p}_i(1-\hat{p}_i)}{n_i-1} \frac{N_i - n_i}{N_i}$	$n = \frac{\sum_{i=1}^L N_i p_i (1-p_i)}{\sum_{i=1}^L N_i p_i (1-p_i) + N^2 \frac{B^2}{4}}$ $\omega = \frac{n_i}{N} = \frac{\frac{c_i}{N}}{\frac{\sum_{i=1}^L N_i \sqrt{p_i(1-p_i)}}{N}} = \frac{c_i}{\sum_{i=1}^L N_i \sqrt{p_i(1-p_i)}}$

Determinación del tamaño de muestra

A continuación se muestran dos ejemplos del cálculo del tamaño de muestra y la estimación de la media poblacional para el muestreo aleatorio simple. La solución se obtuvo a través de sus fórmulas correspondientes y también con el software "Muestreos Estadísticos". Si se desea revisar los siguientes ejemplos, se sugiere consultar el libro de "Elementos de Muestreo" de Sheaffer, Mendenhall y Ott (1987).

Ejemplo 1: La cantidad promedio de dinero M para las cuentas por cobrar de un hospital debe ser estimada. Aunque no se cuenta con datos anteriores para estimar la varianza poblacional, se sabe que la mayoría de las cuentas caen dentro de una amplitud de variación de \$100. Existen N=1000 cuentas abiertas. Encuentre el tamaño de muestra necesario para estimar M con un límite para el error de estimación de B=\$3.

Solución con el uso de fórmulas:

A continuación se sustituyen los valores del ejemplo 1 en las fórmulas del Cuadro 1.

El primer paso es realizar la estimación de la varianza poblacional σ^2 .

$$\sigma^2 = \frac{100}{4} = 25$$

Al sustituirla en la ecuación del tamaño de muestra del muestreo simple aleatorio:

$$D = \frac{(3)^2}{4} = 2.25$$

$$n = \frac{1000(625)}{(999(2.25) + 625)} = 217.56$$

Solución con el software “Muestreos Estadísticos”

En la figura 2 se muestran los resultados del cálculo del tamaño de muestra para el muestreo aleatorio simple.

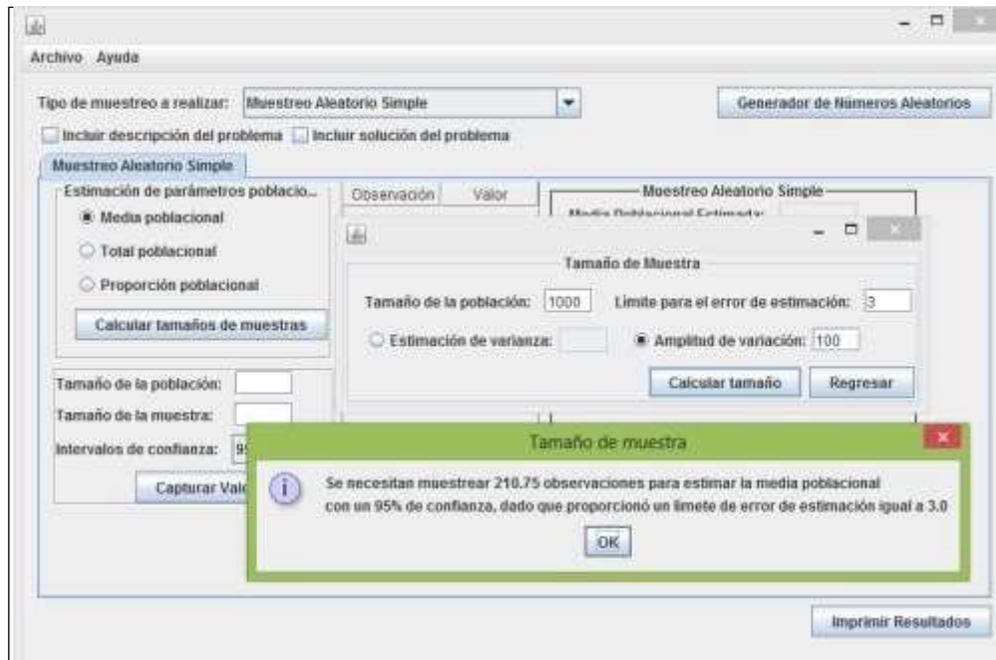


Figura 2. Captura de pantalla para el cálculo del tamaño de muestra.

El hospital debe muestrear 211 cuentas por cobrar para estimar la cantidad promedio de dinero, con un límite para el error de estimación igual a 3 y una confiabilidad del 95%.

Estimación de una media poblacional

Ejemplo 2: Una muestra irrestricta aleatoria de $n=9$ registrados del hospital es seleccionada para estimar la cantidad promedio de una deuda sobre $N=484$ cuentas abiertas. Los valores de la muestra para estos nueve registros están listados en el cuadro 3. Estime la cantidad μ (cantidad promedio de la deuda) y establezca un límite para el error de estimación.

Cuadro 3. Cantidad de dinero adeudado de 9 registros seleccionados aleatoriamente

Observación	Deuda	Observación	Deuda
1	33.50	6	41.00
2	32.00	7	45.00
3	52.00	8	42.50
4	43.00	9	39.00
5	40.00		

Fuente: Sheaffer, Mendenhall y Ott (1987).

Partiendo de esta información, se realizó el cálculo de la media poblacional y otros valores con el programa "Muestreos Estadísticos". En la figura 3 se muestran los resultados obtenidos.

Solución con el software “Muestreos Estadísticos”

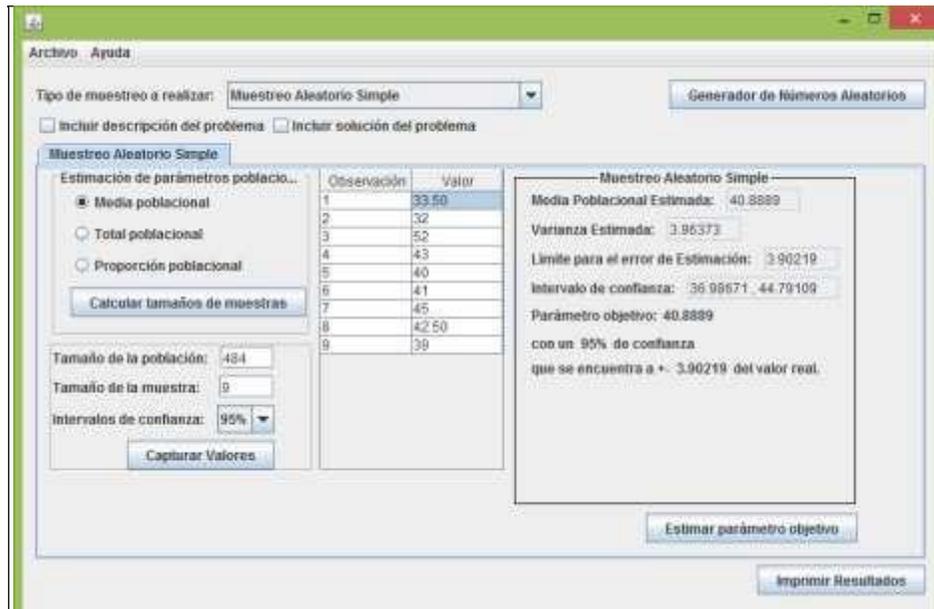


Figura 3. Captura de pantalla del programa de muestreos con los resultados del Muestreo Aleatorio Simple.

Se estima que el hospital tiene una deuda promedio (μ) de \$40.89, con un intervalo de confianza al 95% que se encuentre a \pm \$3.90 de la media poblacional real.

Lo anterior muestra la facilidad del software “Muestreos Estadísticos” para realizar el cálculo del tamaño de muestra y la estimación de la media poblacional en el Muestreo Irrestringido Aleatorio, lo cual es similar con los restantes métodos de muestreo.

CONCLUSIONES

Se concluye que el software libre “Muestreos Estadísticos” es una aportación a las Tecnologías de la Información y la Comunicación (TIC) y puede ser utilizado como apoyo en las estrategias de enseñanza-aprendizaje, así como en el campo profesional en la solución de problemas laborales y de investigación; aunado a que los softwares libres permiten optimizar los recursos económicos de las Universidades públicas o privadas; y debido a que el presente software es multiplataforma, libre, lo cual permite mejorar la versión, al disponer del programa fuente y también al ser gratis, lo cual contribuye al conocimiento y permite disminuir las brechas computacionales por el uso del software comercial y software libre, entre los países desarrollados y subdesarrollados.

El programa para calcular los tamaños de muestra mediante el software muestreos estadísticos se concibe como un complemento para el desarrollo académico de los estudiantes de Licenciaturas en Estadística, economía, entre otras, dónde se imparta el tema de métodos de muestreo, sin distinción alguna del tipo de Institución de Educación Superior.

En esta primera versión se sugiere se agreguen otros métodos de muestreo estadístico probabilístico como es el de conglomerados en dos o más etapas y también se desarrollen los muestreos no-probabilísticos.

REFERENCIAS

Java SE Development Kit 7 Downloads.

<http://www.oracle.com/technetwork/java/javase/downloads/jdk7-downloads-1880260.html>. Consultado 12/08/2013.

Montesinos López O.A., Luna Espinoza I., Hernández Suárez C.M., Tinoco Zermeño M.A. (2009). Muestreo estadístico. Tamaño de muestra y estimación de parámetros. Textos Técnicos Universitarios. Universidad de Colima.

NetBeans IDE 7.2.1 Release Information.

<https://netbeans.org/community/releases/72/> Consultado 12/08/2013.

Rodríguez Pérez, Guillermo. (1999). "Sistema computacional que facilita el cálculo de las diversas técnicas de muestreo". Modalidad Práctico-Técnico. Facultad de Estadística e Informática. Universidad Veracruzana.

Ruiz-Ramírez. Juan y Hernández-Rodríguez Gabriela Eréndira. (2011). "Calidad del empleo y educación superior de los economistas de la Universidad Veracruzana". Contribuciones a la Economía, noviembre: 1-22. Puede consultarse en: <http://www.eumed.net/ce/2011b/>

Ruiz Ramírez Juan, Pérez Salazar Christian y Hernández Rodríguez Gabriela E. (2014). Aplicaciones de software libre en la enseñanza de la estadística en las Instituciones de Educación Superior. I Conferencia Científica Internacional UCIENCIA 2014. 14 p.

Sheaffer L. Richard, Mendenhall William y Ott Lyman. (1987). "Elementos de Muestreo". Tercera edición (Trad. Rendón Sánchez Gilberto y Gómez Aguilar José Roberto). Grupo Editorial Iberoamérica.