

# STIMA 1.1 para estimar parámetros aplicando el método autodocimante bootstrap en investigaciones agropecuarias

## *STIMA 1.1 to estimate parameters applying the method bootstrap in agricultural investigations*

MSc. José Antonio Pino Roque<sup>1</sup>

Dra.C. Mayra Arteaga Barrueta<sup>2</sup>

MSc. Yasmany Vicente Felipe<sup>3</sup>

MSc. Yolanda Rosa Jiménez Álvarez<sup>1</sup>

<sup>1</sup> Departamento Matemática Física, Facultad de Ciencias Técnica

<sup>2</sup> Departamento Química, Facultad de Agronomía.

<sup>3</sup> CUM Jaruco, Facultad de Cultura Física.

Universidad Agraria de La Habana “Fructuoso Rodríguez Pérez”. Autopista Nacional, carretera Tapaste, km 23 <sup>1</sup>/<sub>2</sub>, San José de Las Lajas, Mayabeque.

Autores para correspondencia: pino@unah.edu.cu, mayra@unah.edu.cu, yasmanyv@unah.edu.cu, yolandar@unah.edu.cu

### Resumen

La estimación de parámetros de poblaciones se considera uno de los problemas fundamentales de la Estadística. En ocasiones, los investigadores cuentan con pequeños tamaños de muestras para sus experimentos, es aquí donde el método bootstrap aporta una solución a la problemática planteada para la estimación de parámetros de la población. Es objetivo del colectivo de investigadores diseñar un software que cumpla con los requerimientos del método autodocimante y con el formato de sistema de fácil manejo sobre ambiente Windows, el cual aporta una solución para estimar parámetros bootstrap en investigaciones agropecuarias, reflejando una adecuada exactitud en las estimaciones que se realizan por esta vía. La introducción de estos conocimientos en las actividades docentes, a través de la asignatura de Estadística, que se imparte en diferentes facultades de la UNAH y en cursos de postgrados, han aportado nuevas opciones para estudiantes, profesores e investigadores. Los usuarios de esta herramienta digital han expresado un 100 % de satisfacción con su empleo.

**Palabras claves:** Estimación, Bootstrap, Investigaciones Agropecuarias, estadística inferencial, método intensivo por ordenador.

## Abstract

The estimate of populations' parameters to consider one of the fundamental problems of the Statistic. In occasions, the investigators have small sizes of samples for their experiments, it is here where the method bootstrap contributes a solution to the problem outlined for the estimate of the population's parameters. It is objective of the community of investigators to design a software that fulfills the requirements of the method and with the format of system of easy handling on ambient Windows, which contributes a solution to estimate parameters bootstrap in agricultural investigations, reflecting an appropriate accuracy in the estimates that are carried out for this road. The introduction of these knowledge in the educational activities, through the subject of Statistic that is imparted in different abilities of the UNAH and in courses of graduate degrees, they have contributed new options for students, professors and investigators. The users of this digital tool have expressed 100 % of satisfaction with their employment.

**Keywords:** Estimate, Bootstrap, Investigations Agricultural, statistical inferencial, on-line intensive method.

Recibido: 2 de octubre de 2021

Aprobado: 26 de octubre de 2021

## Introducción

Muchos de los métodos estadísticos que usualmente se utilizan en nuestros días, se concibieron y se desarrollaron entre 1800 y 1930, época en que los trabajos de cómputo eran lentos y costosos, cosa que hoy es rápida y económica. Durante todos estos años han surgido nuevos métodos estadísticos que sacan partido de la gran velocidad de los ordenadores digitales.

Uno de los problemas fundamentales de la Estadística es la estimación de parámetros de poblaciones. En ocasiones, los investigadores cuentan con pequeños tamaños de muestras para sus experimentos. El objetivo fundamental de la estimación es obtener un conocimiento de cierto aspecto de la población sobre el cual se desconocen las características fundamentales (Cue *et al.*,1987). A menos que se conozcan los parámetros, no se puede conocer la bondad de la estimación que alcanza un parámetro con un estadígrafo muestral; hay que conformarse con saber lo bueno que es ese estimativo en promedio, esto es, saber lo bien que se comporta en un muestreo repetitivo, o saber cuántos valores muestrales se pueden esperar que caigan dentro de un intervalo dado en torno al parámetro.

En 1977 Bradley Efron (1993) crea un método

## Introduction

Many of the statistical methods commonly used today were conceived and developed between 1800 and 1930, a time when computer work was slow and expensive, which is fast and inexpensive today. During all these years, new statistical methods have emerged that take advantage of the high speed of digital computers.

One of the fundamental problems of Statistics is the estimation of population parameters. Researchers sometimes have small sample sizes for their experiments. The fundamental objective of the estimation is to obtain a knowledge of a certain aspect of the population for which the fundamental characteristics are unknown (Cue *et al.*, 1987). Unless the parameters are known, the goodness of the estimate that a parameter achieves with a sample statistic cannot be known; you have to settle for knowing how good that estimate is on average, that is, knowing how well it performs in repetitive sampling, or knowing how many sample values can be expected to fall within a given interval around the parameter.

denominado Bootstrap, traducido como “autodocimante”, de muy sencilla descripción, pero imposible de utilizar 35 años atrás por la carencia de máquinas computadoras como las de hoy en día, y el gran volumen de cálculo que se debe utilizar.

No hay elementos para afirmar que en Cuba se hace uso extensivo del método intensivo por ordenador Bootstrap para estimar parámetros en investigaciones porcinas, siendo escasa la bibliografía al respecto; sin embargo a nivel internacional, es más creciente cada día su uso y aplicación en las diferentes investigaciones científicas. Es por ello, que este trabajo tiene como objetivo diseñar un software que cumpla con los requerimientos del Método Intensivo por Ordenador Bootstrap, con el formato de sistema de fácil manejo sobre ambiente Windows, el cual aporte una solución para estimar parámetros bootstrap  $(\bar{x}_{BOOT}, s_{BOOT}^2, s_{BOOT}, cv_{BOOT})$  en investigaciones.

## Materiales y métodos

Para facilitar el procesamiento de la información a través del método Bootstrap se confeccionó el sistema STIMA (versión 1.1). Se tuvo en cuenta la descripción gráfica del método autodocimante que propone Pino *et al.* (2007, 2008) y un diagrama de transición de estados para la estructura (Gráfico 1). La programación se realizó en lenguaje Borland Delphi sobre plataforma Windows, lenguaje de amplia generación con un compilador que genera aplicaciones de tipo cliente/servidor de 32/64 bits que brinda grandes facilidades para el desarrollo de propósitos generales, utilizando una programación visual basada en una metodología orientada (Copplien, 1998).

La presentación del sistema fue diseñada utilizando el procesador gráfico sobre plataforma Windows que brinda Image Compouser. En estos momentos, el software se encuentra en fase de perfeccionamiento (versión 2.0) en lenguaje de programación Java, con extensiones del método a pruebas de hipótesis bootstrap.

El sistema consta de tres ficheros: Stima.hlp (help

In 1977 Bradley Efron (1993) created a method called Bootstrap, translated as “bootstrap”, with a very simple description, but impossible to use 35 years ago due to the lack of computer machines like those of today, and the large volume of calculation to be used.

There are no elements to affirm that in Cuba extensive use is made of the intensive Bootstrap computer method to estimate parameters in pig research, and the bibliography in this regard is scarce; However, at an international level, its use and application in different scientific investigations is growing every day. That is why this work aims to design a software that meets the requirements of the Bootstrap Computer Intensive Method, with the easy-to-use system format on a Windows environment, which provides a solution to estimate bootstrap parameters in research.

## Materials and methods

To facilitate the processing of information through the Bootstrap method, the STIMA system (version 1.1) was developed. The graphic description of the bootstrap method proposed by Pino *et al.* (2007, 2008) and a state transition diagram for the structure (Graph 1). The programming was carried out in the Borland Delphi language on a Windows platform, a wide-generation language with a compiler that generates 32/64-bit client / server applications that provides great facilities for general-purpose development, using visual programming based on a oriented methodology (Copplien, 1998).

The presentation of the system was designed using the graphics processor on Windows platform provided by Image Compouser. At the moment, the software is in the refinement phase (version 2.0) in the Java programming language, with extensions of the method to

file) con 811 729 Kb; Stima.GID (Gid file) con 10 834 Kb; y Stima.exe (application) con 1 338 880 Kb.

Para la confección del sistema STIMA 1.1 se desarrolló el siguiente algoritmo:

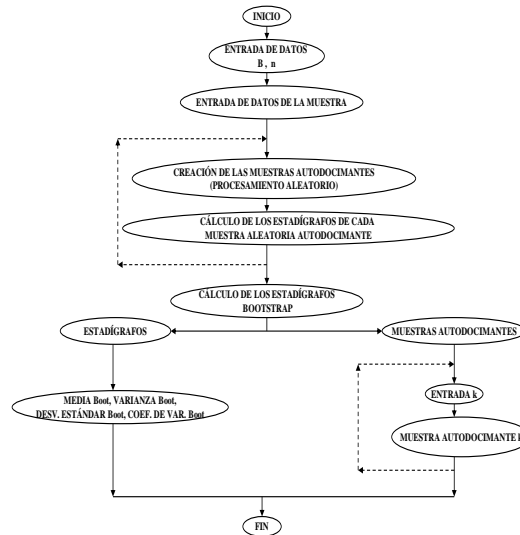
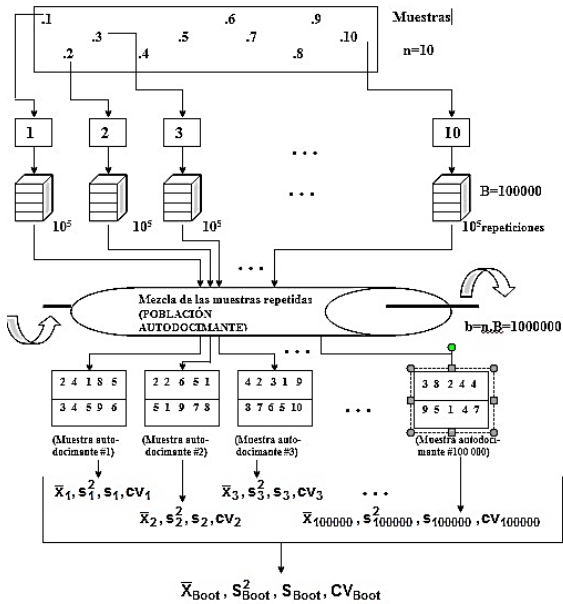
1. Introducir el valor Bootstrap (B) y el tamaño de muestra (n).
  2. Entrar los datos que conforman la muestra a analizar de acuerdo al tamaño de muestra declarado y presentación en pantalla para su posible modificación antes de ser procesada.
  3. Crear un ciclo para seleccionar muestras autodocimantes con tamaño n, de forma aleatoria. Al final quedará una cantidad de B muestras autodocimantes. (1er. Paso del método Bootstrap).
  4. Calcular los estadígrafos  $\bar{x}_K$ ,  $S_K$ ,  $S_K^2$ ,  $CV_K$  de cada muestra autodocimante, con  $K=1, 2, 3, \dots, B$ . (2do. paso del método Bootstrap).
  5. Cálculo de los estadígrafos bootstrap  $\bar{x}_{BOOT}$ ,  $S_{BOOT}^2$ ,  $S_{BOOT}$ ,  $CV_{BOOT}$ . (3er. Paso del método Bootstrap).
  6. Mostrar la muestra  $K$  de la población autodocimante con el objetivo de que el usuario las conozca.
  7. Mostrar los estadígrafos bootstrap (selección opcional) para que el usuario haga uso de los mismos.
- El punto 6 del algoritmo sirvió para validar el programa confeccionado. Del punto 3 al 5 son pasos del algoritmo que no se muestra en el display, ya que forman parte de la estructura del programa para que el ordenador ejecute las operaciones.
8. Salvar las bases de datos utilizadas, con la extensión .btp.
  9. Imprimir los cálculos realizados en el punto 5 del algoritmo.
  10. Repetir el algoritmo utilizando una nueva muestra.

bootstrap hypothesis tests.

The system consists of three files: Stima.hlp (help file) with 811 729 Kb; Stima.GID (Gid file) with 10 834 Kb; and Stima.exe (application) with 1,338,880 Kb.

For the preparation of the STIMA 1.1 system, the following algorithm was developed:

1. Enter the Bootstrap value (B) and the sample size (n).
  2. Enter the data that make up the sample to be analyzed according to the declared sample size and presentation on the screen for possible modification before being processed.
  3. Create a loop to select bootstrap samples of size n, randomly. At the end there will be a quantity of B bootstrap samples left. (1st. Step of Bootstrap method).
  4. Calculate the statistics for each bootstrap sample, with  $K = 1, 2, 3, \dots, B$ . (2nd step of the Bootstrap method).
  5. Calculation of bootstrap statistics (3rd. Step of Bootstrap method).
  6. Show the sample of the bootstrap population so that the user knows them.
  7. Show bootstrap statistics (optional selection) for the user to use.
- Point 6 of the algorithm served to validate the prepared program. Points 3 to 5 are steps of the algorithm that are not shown on the display, since they are part of the program structure for the computer to carry out operations.
8. Save the databases used, with the extension .btp.
  9. Print the calculations made in point 5 of the algorithm.
  10. Repeat the algorithm using a new sample.



**Gráfico 1:** Descripción gráfica y diagrama de transición de estados.

**Graph 1:** Graphic description and state transition diagram.

En el algoritmo anterior no aparece el punto correspondiente a la ayuda, el cual si se incluye, ya que el usuario puede hacer uso de un help, confeccionado para el manejo del sistema, donde se brindan orientaciones, contenidos de interés sobre el método y bibliografías actualizadas para su estudio.

Como requerimientos mínimos para su ejecución se debe contar con 2 161 443 KB y se recomienda como mínimo el empleo de un ordenador Pentium con 32 Mb RAM (memoria activa interna) y 166 Mhz (velocidad del procesador). El cumplimiento de estos requerimientos garantizará una óptima ejecución del mismo.

Todos estos aspectos se tuvieron en cuenta para introducir conocimientos en las actividades docentes, a través de la asignatura Estadística que se imparte en la carrera de Medicina Veterinaria en la UNAH y cursos de postgrados, para capacitar a estudiantes y profesores en nuevas opciones para ser utilizadas en el procesamiento de datos de investigaciones que las requieran.

In the previous algorithm, the point corresponding to the help does not appear, which is included, since the user can make use of a help, made for the management of the system, where guidance, content of interest about the method and bibliographies are provided updated for your study.

As minimum requirements for its execution, it must have 2 161 443 KB and the use of a Pentium computer with 32 Mb RAM (internal active memory) and 166 Mhz (processor speed) is recommended as a minimum. The fulfillment of these requirements will guarantee an optimal execution of the same.

All these aspects were taken into account to introduce knowledge in teaching activities, through the Statistics subject taught in the Veterinary Medicine career at UNAH and postgraduate courses, to train students and teachers in new options to be used in the processing of research data that require them.

## Resultados y discusión

El Sistema STIMA 1.1, brinda la posibilidad de obtener los estimadores más importantes de los parámetros poblacionales, utilizando el Método Intensivo por Ordenador Bootstrap. El software es de muy fácil manejo por estar programado sobre ambiente Windows, lo que permite una buena interacción usuario-máquina.

Se ha empleado en diferentes investigaciones (Pino *et al.*; 2003; 2019) que requerían estimar parámetros poblacionales con pequeños tamaños de muestra:

- Tiempo (en minutos) de llenado de un trailer con cañas de azúcar según el rendimiento del campo para 14 t/ha; 53 t/ha y 58 t/ha.
- Rendimiento de descarga de la combinada KTP-2 para 49 t/ha; 53 t/ha y 63 t/ha.
- Peso al destete (PD) y peso final (PF), en kilogramos, para la producción porcina correspondiente a razas de tres genotipos maternos terminales: Yorkshire x Landrace (YL), Yorkshire x Duroc (YD) y Duroc x Hampshire (DH); alimentadas con dietas de alta proporción de productos y subproductos nacionales.
- Pesaje de leche (en litros) en la primera lactancia (producción bovina) correspondiente a vacas de la raza Holstein Negro.
- Altura y diámetro de espigas, número de flores, y diámetro de los tabacos de las flores en gladiolos de diferentes colores a los cuales se les aplicó un bioestimulante nacional (Liplant).

Los estudiantes, profesores e investigadores que recibieron preparación y aplicaron el sistema STIMA 1.1 en sus investigaciones plantearon un 100 % de satisfacción con el empleo del mismo. Consideran que es una herramienta digital de fácil uso y brinda una ayuda (help) muy completa para neófitos en esta materia.

Se trabaja en una aplicación amboide para ser empleada en el procesamiento de datos

## Results and Discussion

The STIMA 1.1 System offers the possibility of obtaining the most important estimators of the population parameters, using the Bootstrap Computer-Intensive Method. The software is very easy to use because it is programmed on a Windows environment, which allows good user-machine interaction.

It has been used in different investigations (Pino *et al.* ; 2003; 2019) that required estimating population parameters with small sample sizes:

Time (in minutes) to fill a trailer with sugar cane according to field yield for 14 t / ha; 53 t / ha and 58 t / ha.

- Unloading performance of the combined KTP-2 for 49 t / ha; 53 t / ha and 63 t / ha.
- Weaning weight (PD) and final weight (PF), in kilograms, for pig production corresponding to breeds of three terminal maternal genotypes: Yorkshire x Landrace (YL), Yorkshire x Duroc (YD) and Duroc x Hampshire (DH) ; fed with diets with a high proportion of national products and by-products.
- Milk weighing (in liters) in the first lactation (bovine production) corresponding to Black Holstein cows.
- Height and diameter of spikes, number of flowers, and diameter of the flower tobacco in gladioli of different colors to which a national biostimulant (Liplant) was applied.

Students, professors and researchers who received training and applied the STIMA 1.1 system in their research reported 100% satisfaction with its use. They consider that it is an easy-to-use digital tool and provides a very complete help (help) for neophytes in this matter.

para estimar parámetros por el método bootstrap.

### Conclusiones

La informatización en un software, del algoritmo que utiliza el método intensivo por ordenador Bootstrap, en el sistema STIMA 1.1, cumple con los requerimientos del método creado por Bradley Efron, siendo un sistema de fácil manejo con ambiente Window, que puede ser utilizado para estimar parámetros  $(\bar{x}_{BOOT}, S_{BOOT}^2, S_{BOOT}, CV_{BOOT})$  en investigaciones.

Las actividades docentes que introducen los contenidos sobre Bootstrap, a través de la asignatura Bioestadística que se imparte en Medicina Veterinaria en la UNAH y cursos de postgrados, han aportado nuevas opciones para estudiantes, profesores e investigadores. Los usuarios de esta herramienta digital han expresado un 100 % de satisfacción con su empleo.

We are working on an amboid application to be used in data processing to estimate parameters by the bootstrap method.

### Conclusions

The computerization in a software, of the algorithm that uses the intensive Bootstrap computer method, in the STIMA 1.1 system, meets the requirements of the method created by Bradley Efron, being an easy-to-use system with a Window environment, which can be used to estimate parameters in investigations.

The teaching activities that introduce the contents on Bootstrap, through the Biostatistics subject taught in Veterinary Medicine at UNAH and postgraduate courses, have provided new options for students, teachers and researchers. Users of this digital tool have expressed 100% satisfaction with their employment.

### Bibliografía / References

- Copplien, J. O. 1998. *Advanced Borland Delphi programming styles and idioms*. Addison-Wesley. 468 p.
- Cue, J; Castell, E; Hernández, J. M. 1987. Estadística. 1<sup>RA</sup>. Parte. Empresa Nacional de Producción del Ministerio de Educación Superior Imprenta “Andre Voisin”. Ciudad de la Habana: 243 p.
- Cue, J; Castell, E; Hernández, J. M. 1987. Estadística. 2<sup>DA</sup>. Parte. Empresa Nacional de Producción del Ministerio de Educación Superior Imprenta “Andre Voisin”. Ciudad de la Habana: 357 p.
- Efron, B; Tibshirani, R. 1993. *An Introduction to the bootstrap*. New York: Chapman and Hall: 436 p.
- Guerra, C. W; Menéndez, E; Barrero, R; Egaña, E. 1990. Estadística. Editorial Pueblo y Educación. 376 p.
- Jöckel, K. H; Rothe, G; Sandler, W. 1992. Bootstrapping and related techniques. Proceedings of an International Conference Held in Trier, FRG, June 4 – 8, 1990. Springer-Verlag.
- Kendall & Kendall. 1991. Análisis y Diseño de Sistemas. Prentice-Hall Hispano-americano. 345p.

- Pino, J. A; Arteaga, M; Sabín, Y. 2003. Sistema Stima 1.0 para la estimación de parámetros en investigaciones relacionadas con la mecanización agropecuaria utilizando el método Bootstrap. Revista Ciencias Técnicas Agropecuarias. Vol. 12 / ISSN-1010-2760, No. 2/2003.
- Pino, J. A; Arteaga, M; Sabín, Y; Toledo, V; García, L. 2007. *Método Intensivo por ordenado Bootstrap*. Revista Ciencia e Ingeniería Aplicada. Año I. Volumen I. Rioacha. Centro de Investigaciones de la Universidad de la Guajira, Colombia.
- Pino Roque, J. A; Arteaga, M; Sabín, Y. 2008. *STIMA (versión 1.0): Software para la estimación de parámetros aplicando el método intensivo por ordenador Bootstrap*. CD-ROM XVI Congreso Científico Internacional del INCA. Noviembre 2008.
- Pino, J. A; Arteaga, M; Fernández, L; Jiménez, Y. R. 2019. Estimación de parámetros en investigaciones de producción animal (porcina) utilizando el software STIMA 1.0. Memorias PORCICULTURA TROPICAL 2019. CD-ROM. ISBN: 978-959-7208-36-5.
- Rodríguez, L. F; Bermúdez, L.T. 2005. Usos y aplicaciones de la simulación en la investigación agropecuaria. Agronomía colombiana, Vol. XII, No. 1: 198- 204.