

TÉCNICAS DE SIMULACIÓN PARA EL ANÁLISIS ESTADÍSTICO DE DATOS DE MEDICIÓN

Enrique Villa Diharce
Centro de Investigación en Matemáticas A.C.
Callejón de Jalisco s/n, Mineral de Valenciana,
Guanajuato, Gto.
Tel.: (473) 732 71 55, Fax: (473) 732 57 49
Correo electrónico: villadi@cimat.mx

Resumen: Las técnicas de simulación en estadística, como son los métodos de Monte Carlo, y los procedimientos de remuestreo conocidos como bootstrap, son de gran utilidad cuando no tenemos expresiones cerradas para calcular medidas de incertidumbre como son la desviación estándar de estimadores y los intervalos de confianza. Estos métodos de simulación permiten obtener estimaciones con menores supuestos que los métodos analíticos, a cambio de un trabajo computacional mas intenso. La disponibilidad creciente de los recursos computacionales, hacen de las técnicas de simulación una herramienta de uso creciente. En este trabajo se discuten estas técnicas de simulación, y se ilustran con ejemplos sencillos.

INTRODUCCIÓN

En el contexto estadístico, entendemos por simulación, la técnica de muestreo estadístico controlado, que se utiliza conjuntamente con un modelo, para obtener respuestas aproximadas a preguntas que surgen en problemas complejos de tipo probabilístico. En metrología, el proceso de medición es de naturaleza probabilística y los modelos de medición con frecuencia son complejos [1]. Estas dos características del proceso de medición, complejidad y aleatoriedad, hacen del análisis de datos de medición un área de oportunidad natural para los métodos de simulación.

Dos métodos que son aplicables en el análisis de datos de medición, son el método de simulación de Monte Carlo y el método de remuestreo de bootstrap. Estos son de gran utilidad para determinar propiedades estadísticas de las mediciones. A continuación presentamos estos dos procedimientos de análisis, acompañados con ejemplos sencillos para ilustrar su aplicación.

MONTECARLO

De acuerdo a la Guía para la expresión de incertidumbre de medición [2], un sistema de medición se modela por una relación funcional entre las magnitudes de entrada X_1, X_2, \dots, X_p , y el mensurando Y , que es la magnitud de salida o respuesta del sistema,

$$Y=f(X_1, X_2, \dots, X_p) \quad (1)$$

En este sistema de medición, los valores de entrada son las variables de influencia que determinan el valor de la medición del mensurando. Las variables de entrada son de naturaleza aleatoria y su comportamiento está determinado por las correspondientes funciones de distribución $G_1(x_1)$, $G_2(x_2)$, ..., $G_p(x_p)$. La información de estas distribuciones se transmite hacia el mensurando Y a través de la función de medición f . Este proceso se conoce como propagación de distribuciones, y tiene una mayor generalidad que el proceso de propagación de incertidumbres, debido a que maneja una cantidad mayor de información sobre la medición del mensurando.

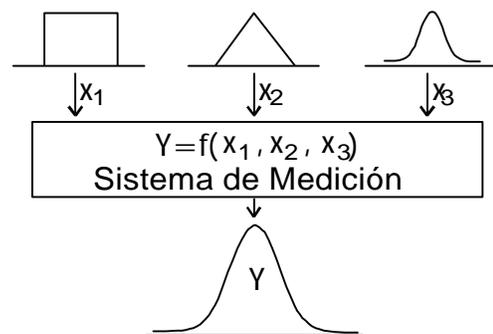


Fig. 1 Propagación de distribuciones en un sistema de medición.

En la Figura 1 se ilustra el proceso de propagación de distribuciones. Cuando conocemos las distribuciones de las variables de entrada X_1, X_2, \dots, X_p , generamos un valor al azar de cada una de ellas y obtenemos el valor de Y , evaluando el modelo de medición f en estas entradas aleatorias. Este proceso se repite un número grande M de veces, y así con las realizaciones de Y obtenemos una función de distribución empírica, a partir de la cual, podemos obtener cualquier estadística del mensurando, acompañada por un intervalo de confianza. Este tipo de simulación, se denomina simulación de Monte Carlo.

El modelo de un sistema de medición es una función que mapea las variables de influencia en una estimación del mensurando bajo medición. La interpretación de este proceso como una propagación de distribuciones (de las variables de entrada a la salida que corresponde a la medición), es de una mayor generalidad que el enfoque de propagación de incertidumbre, ya que en este último, solo le damos seguimiento a las medias y desviaciones estándar, a través de la aproximación de Taylor de la función de medición. Este proceso de aproximación no es adecuado en el caso de funciones de medición altamente no lineales. El proceso de muestrear las variables de influencia y la transformación de sus realizaciones en una medición, es justamente el proceso de propagación de distribuciones. Y esto no es otra cosa que una simulación de Monte Carlo.

Ejemplo 1

Para ilustrar el método de Monte Carlo, consideramos un modelo de medición sumamente simple. Supongamos que en la medición de una longitud determinada utilizamos una cinta metálica que tiene una resolución de un milímetro, en donde se obtiene la lectura $L_0 = 49\text{mm}$. Esta medición debe corregirse tomando en cuenta que la escala fue calibrada a una temperatura de 20°C , y que con el cambio de temperatura, la cinta sufre un cambio por dilatación. Los factores que intervienen en el proceso de dilatación son el valor del coeficiente de dilatación térmica C y la temperatura T a la cual se hace la medición. Supongamos además que de acuerdo al fabricante, el coeficiente de dilatación térmica de la cinta es $12.65 \times 10^{-6} \text{ }^\circ\text{C}^{-1} \pm 5 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$, mientras que la temperatura se mide con un termómetro cuya escala tiene una resolución de un grado, y la lectura arroja el valor de 35°C .

El modelo de medición en este caso es,

$$L = L_0(1 + C(T - 20)) \tag{2}$$

De acuerdo a la información sobre las variables de entrada, asignamos a L_0 la distribución UNIFORME(48.5, 49.5), a C una distribución UNIFORME(12.15×10^{-6} , 13.15×10^{-6}), y finalmente a T una distribución UNIFORME(34.5, 35.5). Para obtener una estimación de L y un intervalo de confianza por el método de Monte Carlo, simulamos valores de L_0 , C y T de acuerdo a las distribuciones mencionadas y calculamos el valor correspondiente de Y , utilizando el modelo de medición (2). Repetimos este proceso 10 000 veces y obtenemos la distribución empírica de la medición del mensurando, a partir de la cual obtenemos una estimación, así como un intervalo de 95% de confianza. En la Figura 2 se muestra un histograma correspondiente a las estimaciones Monte Carlo del mensurando.

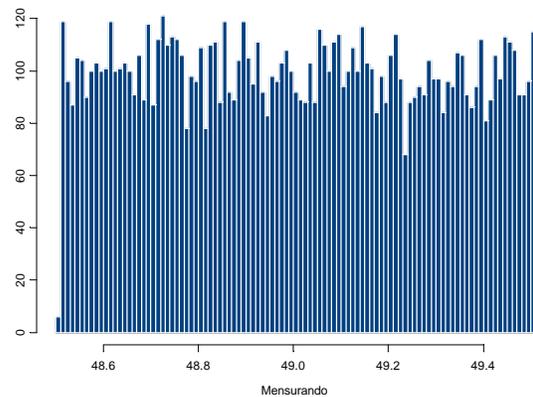


Fig. 2 Histograma de las 10 000 realizaciones del valor del mensurando obtenidas por simulación de Monte Carlo.

Variable	Estimación	Desv. Estándar
L_0	49.00 mm	.289 mm
C	$12.65 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$	$.289 \times 10^{-6} \text{ }^\circ\text{C}^{-1}$
T	35°C	$.289^\circ\text{C}$
L	49.009 mm	0.291 mm

Tabla 1. Estimaciones e incertidumbre de estimación para las variables de entrada y el mensurando.

De la distribución empírica de Y obtenemos un intervalo de confianza de probabilidad .95 para el mensurando, tomando como límites inferior y superior del intervalo a los cuantiles simétricos .025 y .975 respectivamente. Estos cuantiles se obtienen fácilmente después de ordenar en forma ascendente las realizaciones de Y, eligiendo las que ocupan la posición número 250 (=0.025x10 000), y 9750 (= .975x10 000). En la Tabla 1 tenemos estimaciones y medidas de incertidumbre para las variables de entrada y el mensurando, calculadas de acuerdo a sus distribuciones.

BOOTSTRAP

La calidad de un estimador está relacionada con la varianza de dicho estimador. Esto implica que un estimador puntual de una cantidad de interés, debe ir acompañado por una estimación de su varianza, o una medida de su incertidumbre como puede ser un intervalo de confianza de dicha cantidad de interés. Algunas veces es difícil calcular la incertidumbre de un estimador, debido a que no hay una expresión cerrada que permita calcular la varianza del estimador, este es por ejemplo el caso de la mediana.

Existe una metodología, que fue desarrollada en 1979 por Efron [3], para calcular la variabilidad de un estimador, conocida como bootstrap, que consiste en tomar muestras aleatorias de la muestra que tenemos originalmente, y para cada una de ellas calculamos nuestro estimador. Este proceso lo repetimos un número grande B de veces y encontramos la distribución empírica de las estimaciones, que utilizamos para estimar la incertidumbre del estimador, ya sea como desviación estándar o intervalo de confianza.

El paradigma de esta técnica surge de una analogía en la cual la muestra observada asume el papel de la población de la cual proviene. Esto es, el procedimiento consiste en tomar muestras de una muestra. A este tipo de procedimiento también se le conoce como remuestreo. En la Figura 3 ilustramos este procedimiento.

Algoritmo de remuestreo

El punto de partida es una muestra de tamaño n, $\mathbf{x}=(x_1, x_2, \dots, x_n)$ tomada de la población. De esta muestra tomamos una muestra de tamaño n con reemplazo, $\mathbf{x}_b=(x_1, x_2, \dots, x_n)$. Para esta muestra,

calculamos la estimación \hat{q}_b^* de la cantidad de interés θ . Repetimos lo anterior B veces, y obtenemos la distribución empírica de las estimaciones bootstrap \hat{q}_b^* , $b=1,2,\dots,B$. La distribución muestral de \hat{q} , se aproxima por la distribución bootstrap de \hat{q}^* . De esta distribución empírica, calculamos finalmente la desviación estándar, y un intervalo de confianza, medidas de incertidumbre del estimador \hat{q} de θ . Además de estas medidas de incertidumbre, la información que tenemos en las muestras bootstrap generadas nos permiten conocer la magnitud de sesgo de un estimador.

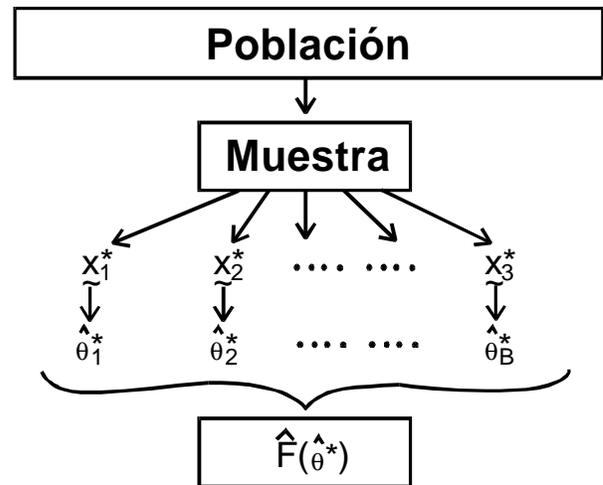


Fig. 3 Determinación de la variabilidad por remuestreo (bootstrap).

La expresión para calcular la estimación bootstrap de la desviación estándar es:

$$DE_b = \left\{ \frac{1}{B-1} \sum_{b=1}^B \left[\hat{q}_b^* - (1/B) \sum_{b=1}^B \hat{q}_b^* \right]^2 \right\}^{1/2} \quad (3)$$

La estimación bootstrap de la medida de sesgo del estimador \hat{q} de θ , es la diferencia entre la media de las estimaciones bootstrap \hat{q}_b^* , y la estimación \hat{q} de θ , y está dada por:

$$\text{Sesgo}_b = (1/B) \sum_{b=1}^B \hat{q}_b^* - \hat{q} \quad (4)$$

La incertidumbre de la estimación en forma de intervalo de $100(1-2\alpha)\%$ de confianza, se obtiene tomando los cuantiles simétricos α y $(1-\alpha)$ de la distribución bootstrap de \hat{q}^* , con esto evitamos la suposición de normalidad de la distribución bootstrap de \hat{q}^* . Este intervalo de confianza simétrico es simplemente

$$\left[F_{\hat{q}^*}^{-1}(\alpha), F_{\hat{q}^*}^{-1}(1-\alpha) \right]. \quad (5)$$

Ilustramos la metodología bootstrap de estimación de incertidumbre con un ejemplo sencillo.

Ejemplo 2

La Tabla 2 contiene una muestra de mediciones de diámetros de componentes mecánicos generados por un proceso. Nos interesa saber si la media y la mediana de los diámetros de los componentes generados por el proceso se pueden usar en forma equivalente como valores representativos del proceso. Podemos comparar estos parámetros a través del estudio de su diferencia. En este caso, el conocimiento de un intervalo de confianza de la diferencia de parámetros, nos permite determinar si hay una diferencia significativa entre ellos.

4.64	4.47	4.46	4.54	4.50	4.57
4.50	4.42	4.33	4.52	4.49	4.54
4.50	4.38	4.48	4.49	4.57	4.51

Tabla 2. Mediciones de diámetro en mm de componentes mecánicos generados por un proceso.

La metodología bootstrap es de gran utilidad para el cálculo de un intervalo de confianza de la diferencia media-mediana, ya que no contamos con una expresión para la incertidumbre del estimador de dicha diferencia.

En la Figura 3 tenemos el histograma y una estimación de la densidad para las estimaciones bootstrap de la diferencia de media y mediana calculadas para 10 000 muestras extraídas con reemplazo de la muestra de la Tabla 2. En la Tabla 3 tenemos valores de la estimación de la diferencia, basada en la muestra de la Tabla 2. También se da la media de las estimaciones bootstrap de la diferencia, la desviación estándar, el sesgo, y un intervalo de 95% de confianza de la diferencia, basadas en las 10 000 muestras bootstrap.

Como el cero se encuentra dentro del intervalo de confianza, concluimos que los datos no tienen evidencia en contra de la hipótesis de equivalencia entre media y mediana de los diámetros. Este resultado implica que podemos usar en forma indistinta a la media y la mediana como valores representativos de los diámetros de los componentes generados por el proceso.

Est.	Sesgo	Media	Des.Est.	Int. de Conf.
-0.005	.0014	-0.0035	.0114	(-.0266,.0183)

Tabla 3. Estimación de la diferencia media-mediana, y la media de las 10 000 estimaciones bootstrap de la diferencia. Además, se presentan medidas de incertidumbre y sesgo para la media y la mediana de los datos del Ejemplo 2.

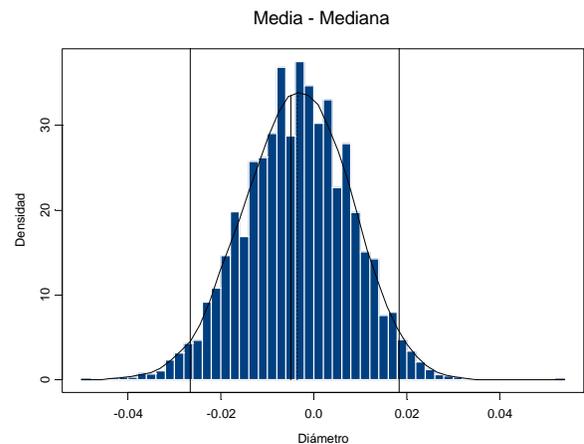


Fig. 3 Histograma y función de densidad de las 10000 realizaciones de las estimaciones bootstrap de la diferencia media - mediana. También se muestra un intervalo de 95% de confianza de la diferencia.

APLICACIONES

Las áreas de oportunidad de aplicación para las técnicas de simulación aquí expuestas son muy diversas, como lo muestra C. E. Lunnenborg [4], en el caso de remuestreo, y G. S. Fishman [5], en el tema general de Monte Carlo. Ambos hacen una amplia presentación de ejemplos y algoritmos para llevar a cabo experimentos de muestreo en una computadora. En un apéndice de Efron y Tibshirani [6] se dan funciones para S-PLUS [7], para calcular intervalos de confianza bootstrap. Las últimas versiones del lenguaje de programación S-PLUS incluyen funciones para hacer inferencia bootstrap.

En el campo de la metrología, P. Ciarlini [8] propone el uso de la metodología bootstrap para analizar datos de comparaciones interlaboratorios, en donde las muestras son de tamaño limitado, y los datos son de naturaleza heterogénea. E. Pardo, y M. G. Cox [9], utilizan el remuestreo bootstrap, para calcular los coeficientes de la mediana total, que proponen como estimador robusto de la media. M. G. Cox, et al [10], discuten el uso de de simulación en la evaluación de la incertidumbre de medición.

CONCLUSIONES

El cálculo de incertidumbres de estimadores por simulación, es de gran utilidad cuando: 1) No tenemos una expresión cerrada para obtener un índice de incertidumbre, 2) Nos interesa comparar varios estimadores propuestos para una cantidad de interés. Existe una aplicación creciente de técnicas de simulación en el análisis estadístico de datos de medición. En este trabajo, hemos presentado la metodología de muestreo de Monte Carlo, y el procedimiento de remuestreo bootstrap, con ejemplos que ilustran su aplicación en el análisis de datos de medición.

REFERENCIAS

- [1] W. Bich, The ISO Guide to the Expression of Uncertainty in Measurement: A Bridge Between Statistics and Metrology, in P., Ciarlini, M. G. Cox, F. Pavese and D. Richter, editors, Advanced Mathematical Tools in Metrology III. Series on Advances in Mathematics for Applied Sciences, Vol. 45, World Scientific, 1997, pp. 1-11.
- [2] ISO, Guide to the expression of Uncertainty in Measurement, Switzerland, First ed. (1993).

- [3] B. Efron. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7, pp.1-26.
- [4] C.E. Lunnenborg, *Data Analysis by Resampling: Concepts and Applications*, Duxbury Press, 2000.
- [5] G. S. Fishman, *Monte Carlo: concepts, algorithms, and applications*, Springer-Verlag, 1996.
- [6] Efron, B., and Tibshirani, R. J. (1993), *An introduction to the Bootstrap*, New York: Chapman and Hall.
- [7] S-PLUS, Statistical Sciences, Inc. (1996), *S-PLUS User's Manual, Volumes 1-2, Version 3.4*, Seattle: Statistical Sciences.
- [8] P. Ciarlini (1999). Advantages of the bootstrap technique in solving interlaboratory comparison problems, *Workshop on Statistical Analysis of interlaboratory comparisons*, 11-12 November, 1999, NPL.Teddington, U. K.
- [9] E. Pardo and M. Cox, *The Total Median*, *Workshop on Statistical Analysis of interlaboratory comparisons*, 11-12 November, 1999, NPL.Teddington, U. K.
- [10] M. G. Cox, M. P. Dainton, A. B. Forbes, P. M. Harris, P. M. Schwenke, B. R. L. Seibert, and W. Woger, *Use of Monte Carlo Simulation for Uncertainty Evaluation in Metrology*, in P. Ciarlini, M. G. Cox, E. Filipe, F. Pavese, and D. Richter, editors, *Advanced Mathematical Tools in Metrology V. Series on Advances in Mathematics for Applied Sciences*, Vol. 57, World Scientific, 2001, pp. 93-104.