

Efecto del tamaño de muestra y el número de réplicas bootstrap

Effect of sample size and number of bootstrap samples

Isabel C. Ramírez*§, Carlos J. Barrera*· **, Juan C. Correa*

**Facultad de Ciencias, Escuela de Estadística, Universidad Nacional de Colombia Sede Medellín, Medellín, Colombia*

***Facultad de Ciencias Exactas y Aplicadas, Programa de Física, Instituto Tecnológico Metropolitano, Medellín, Colombia*

§*iscramirezgu@unal.edu.co, carlosbarrera@itm.edu.co, jccorrea@unal.edu.co*

(Recibido: Marzo 05 de 2012 - Aceptado: Abril 08 de 2013)

Resumen

Un problema difícil de responder es determinar el tamaño de muestra mínimo y el número de réplicas bootstrap que se debe utilizar para obtener distribuciones muestrales que se aproximen bien a la verdadera distribución muestral utilizando la metodología bootstrap. Se evalúa el efecto del tamaño de muestra y el número de réplicas bootstrap en las estimaciones de la media, varianza y algunos percentiles para una distribución de probabilidad bajo diferentes medidas de asimetría. Presentamos algunas recomendaciones para el uso del bootstrap no paramétrico.

Palabras clave: *Asimetría, bootstrap, estimación, muestras bootstrap, tamaño de muestra.*

Abstract

A difficult problem to address is to determine the minimum sample size and the numbers of bootstrap samples to be used for obtain sampling distribution that makes a better fit to the true sampling distribution using the bootstrap methodology. We evaluate the effect of sample size and the number of bootstrap samples in the estimates of the mean, variance and some quantiles for probability distribution under different measures of skewness. We give some recommendations for using the non-parametric bootstrap.

Keywords: *bootstrap, bootstrap samples, estimation, sample size, skewness*

1. Introducción

Un problema difícil de responder es determinar el tamaño de muestra mínimo y el número de muestras bootstrap que se debe utilizar para obtener distribuciones muestrales apropiadas utilizando la metodología bootstrap no paramétrica. En el presente artículo se dan algunas recomendaciones para el uso del bootstrap no paramétrico. Se evalúa el efecto del tamaño de muestra y el número de muestras bootstrap en las estimaciones de la media, varianza y algunos percentiles para una distribución gamma, considerando diferentes valores de los parámetros de dicha distribución, lo cual proporciona diferentes medidas de asimetría de la distribución de estudio.

La técnica bootstrap no paramétrica fue propuesta por Efron (1979), y se basa en el remuestreo de una muestra. La idea básica es que si se toma una muestra aleatoria (x_1, x_2, \dots, x_n) entonces la muestra puede ser utilizada para obtener más muestras. El procedimiento es un remuestreo aleatorio (con reemplazo) de la muestra original tal que cada punto x_i tiene igual e independiente oportunidad de ser seleccionado como elemento de la nueva muestra bootstrap. El proceso completo son repeticiones independientes de muestreo hasta obtener un número grande de muestras bootstrap. Múltiples estadísticos pueden calcularse para cada muestra bootstrap y, por lo tanto, sus distribuciones pueden ser estimadas (Efron & Tibshirani, 1993).

Chernick (2008) comenta algunas variantes del bootstrap, por ejemplo el suavizado que se utiliza cuando resulta razonable reemplazar la función de distribución empírica con una versión suavizada que posiblemente es basada en métodos kernel. También presenta el bootstrap paramétrico el cual asume que la distribución de los datos tiene una forma paramétrica. Es usual asumir en este caso que ésta sigue una distribución normal con el fin de llegar a estimaciones bootstrap que son los estimadores de máxima verosimilitud. Chernick (2008) presenta otras variantes, como por ejemplo el bootstrap bayesiano y el bootstrap doble.

En este artículo se utiliza el bootstrap no paramétrico debido a su amplio uso en diferentes aplicaciones, lo cual es impulsado por su sencillez en la implementación además de diferentes argumentos a su favor como lo que expresa Chernick (2008) quien afirma que una motivación para su uso es que la distribución empírica es el estimador de máxima verosimilitud de la distribución de los datos (F) cuando no hay supuestos acerca de esta. Consecuentemente, se pueden ver las estimaciones bootstrap de los parámetros de F como los estimadores de máxima verosimilitud no paramétricos de dichos parámetros. El autor comenta que el uso del bootstrap no paramétrico se ve impulsado por su facilidad y eficiencia computacional comparada con algunas de sus variantes. Además, Davison & Hinkley (1997) justifican el bootstrap no paramétrico como una forma de validar los resultados obtenidos con las variaciones de dicha metodología.

Algunos autores estudian las desventajas del bootstrap no paramétrico, por ejemplo Andrews (2000) comenta que en modelos con coeficientes aleatorios, a menudo las varianzas estimadas de algunos de los coeficientes aleatorios son pequeñas y por lo tanto las varianzas verdaderas de algunos de estos coeficientes podrían ser cero. En este caso el bootstrap no paramétrico no es consistente. Hesterberg et al. (2003) realizan un análisis del bootstrap no paramétrico basado en muestras pequeñas y anota que la distribución bootstrap puede ser muy variable, su forma y la dispersión reflejan las características de la muestra y pueden no estimar con precisión la forma y la dispersión de la distribución muestral. También afirma que el bootstrap no es confiable para estadísticos como la mediana y los cuartiles cuando el tamaño de muestra es pequeño.

Se puede encontrar diversas aplicaciones de la metodología bootstrap en múltiples áreas de la ingeniería, Fethke et al. (2007) realizan un análisis bootstrap para examinar la precisión de ciertas mediciones a medida que la duración del muestreo se incrementa gradualmente. La precisión la estiman calculando el coeficiente de variación

de la distribución bootstrap en cada incremento de la duración del muestreo. Tung-I. Tsai & Der-Chiang Li (2008) estudian el tamaño de las corridas en las pruebas pilotos en los procesos de producción donde normalmente se obtienen pocos datos. Proponen utilizar la metodología bootstrap para estos casos y muestran que los errores de predicción disminuyen significativamente aplicando dicha metodología.

La metodología bootstrap también ha sido utilizada para estudiar problemas en el sector financiero, Kosowski et al. (2006) examinan la significancia estadística de los rendimientos de los fondos mutuos de inversión utilizando un procedimiento bootstrap. Concluyen que la ventaja de la metodología que aplican es que no es necesario establecer una distribución de probabilidad exacta para modelar los retornos. Halkos & Tzeremes (2010) utilizan varios modelos de análisis envolvente de datos (DEA) combinando múltiples medidas financieras en una única medida utilizando la técnica bootstrap. Plantean cuestiones relativas a la influencia de factores no financieros que se interrelacionan con la elección de las medidas financieras adoptadas y cómo los errores en la estimación de la eficiencia se pueden evitar con el uso de técnicas bootstrap.

En la literatura se encuentran aplicaciones de la metodología bootstrap en el control estadístico de calidad. Chatterjee & Qiu (2009) proponen utilizar una secuencia de límites de control para la gráfica de control de sumas acumuladas (CUSUM), donde los límites son determinados con la distribución condicional del estadístico CUSUM y su estimación se realiza por bootstrap. Comentan que una de las ventajas del procedimiento propuesto es que no tiene supuestos sobre la distribución y es robusto para diferentes elecciones de distribuciones en control y fuera de control. Chou et al. (2006) realizan una estimación de intervalos de confianza bootstrap para el índice de capacidad de proceso C_{pp} . Adicionalmente Castillo & Hadi (1995), Heierman et al. (2005) y Bigerelle et al. (2006) realizan estudios de confiabilidad haciendo uso de la metodología bootstrap.

Efron (1987) discute el número de simulaciones necesarias para la estimación del sesgo y los cuantiles. Mientras que Young & Daniels (1990) muestran mediante un estudio de simulación que el bootstrap es sesgado para muestras pequeñas considerando la estimación de la distribución de un pivote para dos situaciones utilizando bootstrap.

Shao (1994) estudió el estimador bootstrap de la distribución de muestreo de un estadístico en el caso en que no se cumplen las condiciones de regularidad donde el estadístico dado es no suavizado. Encontró que el método bootstrap cuando no se cumplen estas condiciones basado en muestras bootstrap del mismo tamaño de los datos originales produce estimadores inconsistentes bootstrap, pero cuando se extraen muestras bootstrap más pequeñas bajo las mismas condiciones, no hay problemas de consistencia en los estimadores.

Según Davison et al. (2003), desde sus inicios el método bootstrap ha sido una herramienta poderosa para dar soluciones a la práctica estadística, y una rica fuente de problemas teóricos y metodológicos. Efron (2003) opina que una combinación de estadística Bayesiana con la frecuentista surgirá para solucionar problemas complejos y que los métodos computacionales tales como el bootstrap facilitarán esta combinación.

Götze & Rackauskas (2001) consideran secuencias de estadísticos $T_n(\hat{P}_n, P)$ de una muestra de tamaño n , donde analizan datos basados en el procedimiento propuesto por otros autores para seleccionar el tamaño de muestra $m = m_n \leq n$ para el muestreo bootstrap de tipo m tomadas de n tal que la secuencia bootstrap T_m^* de estos estadísticos es consistente y el error es comparable con el mínimo error en la selección conociendo la distribución P .

Chung & Lee (2001) comentan que algunos estudios han mostrado que utilizando un tamaño de muestra bootstrap diferente de n algunas veces provee una solución más satisfactoria, ver Swanepoel (1986) y Bickel et al. (1997). Los

autores utilizan este último enfoque para corregir el error de cobertura en la construcción de límites de confianza bootstrap. Por su parte, Geluk & Haan (2002) estudiaron el tamaño de muestra bootstrap en la teoría de valores extremos, encontrando resultados similares cuando usan el bootstrap para estimar un cuantil intermedio.

Algunos artículos reportan aplicaciones del método bootstrap a variables industriales con distribución gamma, por ejemplo, Marazzi (2001) utiliza el procedimiento bootstrap para probar la igualdad de medias robustas en una, dos, y múltiples muestras en problemas para datos distribuidos asimétricamente. Utilizan datos que corresponden a tiempos de estancia de pacientes en un hospital. Chang et al. (2011) aplican una variante del bootstrap paramétrico para una prueba de diferencias de medias para distribuciones gamma. Realizan una aplicación con datos de un estudio de seguimiento de la disponibilidad de agua subterránea para pozos domésticos.

2. Metodología

Para analizar el efecto del tamaño de muestra y el número de muestras bootstrap en las distribuciones de algunos estimadores, se simula por el método de Monte Carlo una muestra aleatoria de tamaño n de una distribución $gamma(\alpha, \beta)$, de esta muestra se toman m muestras de tamaño n con reemplazo y a cada una de estas m muestras se les calcula el estadístico de interés, por lo tanto se van a tener m valores de dicho estadístico con los cuales se estima su distribución muestral. Se realizan 1000 simulaciones de este procedimiento obteniendo 1000 distribuciones bootstrap del estadístico de interés. Luego se compara la verdadera distribución muestral del estadístico con cada una de las 1000 distribuciones bootstrap. La comparación se hace mediante el uso de la distancia máxima de Kolmogorov-Smirnov. Finalmente se hace un análisis del promedio de las distancia máximas para cada uno de los casos.

Los tamaños de muestras n , que se evalúan son 5, 10, 20, 40, 80, 150 y 300. Para cada uno de estos tamaños de muestras se extraen m réplicas

bootstrap de tamaños 50, 100, 250, 500 y 1000. Esta selección se hace teniendo en cuenta los valores más frecuentes que son utilizados por algunos autores, por ejemplo ver Chung & Lee (2001) y Martin (2007). Las muestras aleatorias $gamma$ que se generan tienen parámetros (1,2), (2,2), (3,2), (5,1) y (9,0.5). Las distribuciones $gamma(1,2)$ y $gamma(2,2)$ corresponden a las distribuciones exponencial con parámetro 2 y chi-cuadrado con 1 grado de libertad, respectivamente. Se seleccionan estos valores para los parámetros con el fin de evaluar el desempeño del método bootstrap no paramétrico cuando se tienen diferentes formas distribucionales no simétricas.

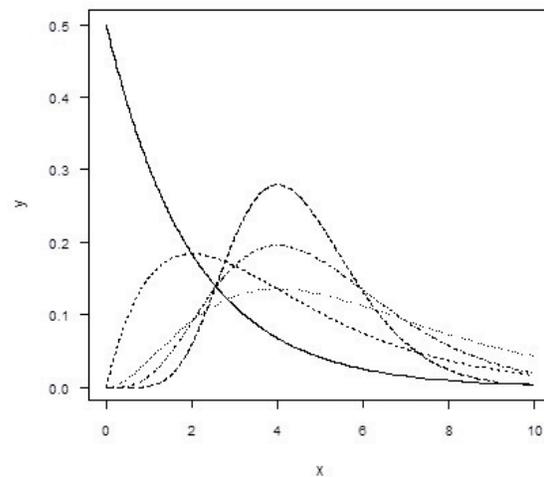


Figura 1. Distribuciones gamma con parámetros (α, β) respectivamente iguales a —Gamma (1,2), --- Gamma (2,2), Gamma (3,2), - - - Gamma (5,1) y Gamma (9,0.5)

Se realiza el análisis de la distribución muestral de la media, varianza, percentil 5, 25, 50, 75 y 95. En la figura 1 se observan las diferentes distribuciones $gamma$ estudiadas. Se puede observar que representan diferentes medidas de asimetría.

3. Resultados y discusión

En las Figuras 2, 3, 4 y 5 se presentan los perfiles para el promedio de los estadísticos de Kolmogorov-Smirnov, es decir, la distancia máxima promedio (DMP) para las distribuciones

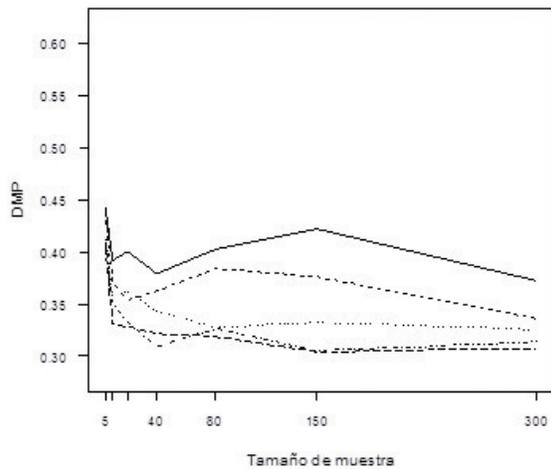


Figura 2. Promedio de los estadísticos de Kolmogorov-Smirnov para la media de una gamma (1,2) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

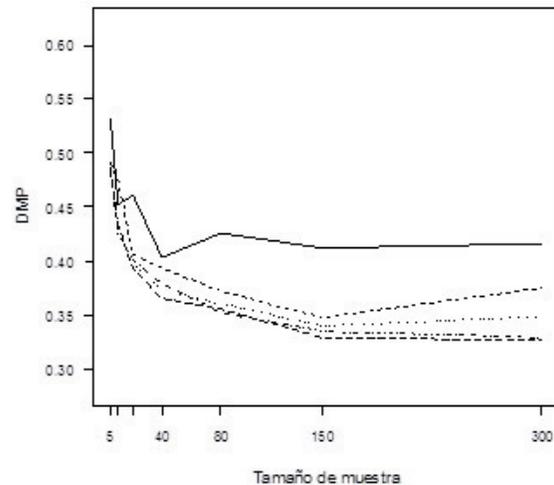


Figura 3. Promedio de los estadísticos de Kolmogorov-Smirnov para la varianza de una gamma (2,2) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

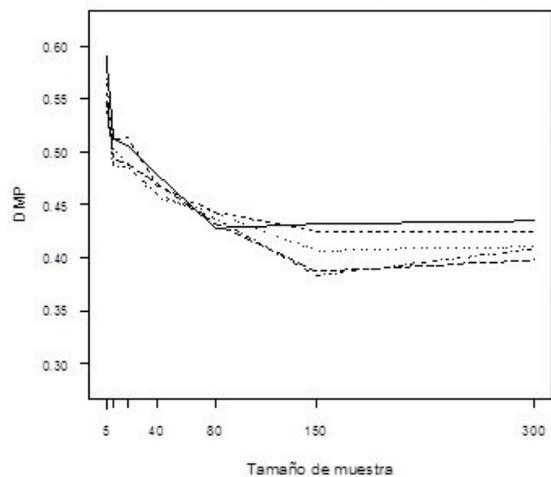


Figura 4. Promedio de los estadísticos de Kolmogorov-Smirnov para el percentil 5 de una gamma (9,0.5) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

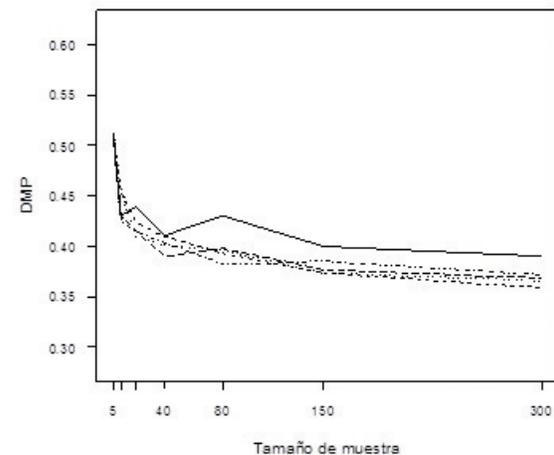


Figura 5. Promedio de los estadísticos de Kolmogorov-Smirnov para el percentil 25 de una gamma (5,1) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

empíricas del estadístico de interés con respecto a las distribuciones de cada estadístico vía bootstrap, mostrándose en el eje x los diferentes tamaños de muestras utilizados y en el eje y la DMP. En cada gráfico se muestra una curva para los diferentes número de muestras bootstrap.

En los gráficos de perfiles para el estadístico de la media observamos que existen dos patrones, unos para las distribuciones *gamma* (1,2) y (2,2), y el otro para las otras distribuciones *gamma* estudiadas, por lo tanto se explica el comportamiento general para cada patrón. En la

figura 2 vemos que para la $gamma(1,2)$, cuando se tienen tamaños de muestra pequeños la DMP es mayor y tiende a estabilizarse a medida que el tamaño de muestra aumenta, aproximadamente en 40 muestras. Podemos notar que si comparamos el número de muestras bootstrap pequeños (50, 100) contra grandes (250, 500 y 1000), la DMP es inferior para estos últimos casos. Para el otro patrón, se encontró que el aumento del tamaño de muestra provoca una disminución de la DMP , pero únicamente hasta un tamaño de muestra de aproximadamente 80, luego de esto, el valor de la DMP se estabiliza. Por su parte, el número de muestras bootstrap tiene un comportamiento similar al explicado en la figura 2.

Los gráficos de perfiles para el estadístico de la varianza también se observan dos patrones: uno para las distribuciones $gamma(1,2)$ y $(2,2)$ y otro para las otras distribuciones. En la figura 3, que corresponde a la $gamma(2,2)$, se observa que a medida que aumenta el tamaño de la muestra la DMP disminuye, pero a partir del tamaño de muestra 150 la DMP tiende a estabilizarse. El comportamiento de la DMP cuando las muestras bootstrap son 50 es inestable y presenta DMP mayores comparadas con los otros casos. Para los casos de la distribuciones $gamma(3,2)$, $(5,1)$ y $(9,0.5)$ la DMP se estabiliza aproximadamente en un tamaño de muestra 80 y no se encontró un efecto del número de muestras bootstrap.

En el caso de los percentiles 5, 25, 50 y 75 se observó un comportamiento similar de los gráficos de perfiles para todas las distribuciones estudiadas. En la figura 4 se muestra el caso del percentil 5 para la $gamma(9,0.5)$, se observa que a medida que aumenta el tamaño de muestra la DMP disminuye y tiende a estabilizarse a partir de 150 muestras. No se presenta un efecto del número de muestras bootstrap. Para el percentil 50 la DMP presenta el mismo comportamiento del percentil 5.

La DMP para el percentil 25 de una $gamma(5,1)$, figura 5, disminuye a medida que aumenta el tamaño de muestra pero a partir del tamaño de muestra 80 la DMP es aproximadamente

constante. En general, las diferentes curvas para los diferentes número de muestras bootstrap se superponen en todos los gráficos, pero el caso de 50 muestras bootstrap la DMP es mayor a través de la mayoría de tamaños de muestra evaluados.

En el caso del percentil 75 el comportamiento al que se observa para el percentil 25, existe un efecto del tamaño de muestra en el DMP hasta cuando el tamaño de muestra es 80, para los valores mayores el DMP tiende a estabilizarse. No se presenta un efecto del número de bootstrap en el DMP para este percentil.

Nuevamente, para el caso de percentil 95 observamos dos patrones en el comportamiento del gráfico de perfiles, uno para las distribuciones $gamma(1,2)$ y $(2,2)$ y otro para las otras distribuciones. Como se observa en las figuras 6 y 7 la DMP para la distribución del percentil 95 se encuentra entre 0.35 y 0.65. En general, a medida que aumenta el tamaño de muestra el DMP disminuye, pero, en casi todos los casos, cuando se aumenta el tamaño de muestra de 150 a 300 el DMP aumenta. Cuando se cambia el número de muestras bootstrap, con una misma distribución, el comportamiento de las diferentes curvas es muy similar. Sin embargo, en el caso de la $gamma(2,2)$ (ver figura 6) cuando el número de muestras bootstrap es 50 la DMP es mayor para la mayoría de los tamaños de muestra estudiados.

Con el análisis de los gráficos de perfiles para las diferentes distribuciones estudiadas recomendamos tamaños de muestra y número de muestras bootstrap que se debería utilizar para la estimación de la distribución de los estadísticos que se analizaron en este trabajo. Se realiza esta recomendación teniendo en cuenta el coeficiente de asimetría de cada una de las distribuciones.

Las recomendaciones dadas en las tablas 1 y 2 deben tenerse en cuenta de manera conjunta. Por ejemplo, si se va a aplicar la metodología bootstrap para obtener la distribución de la media de una distribución con coeficiente de asimetría 2, recomendamos tomar 250 muestras bootstrap de tamaño 40.

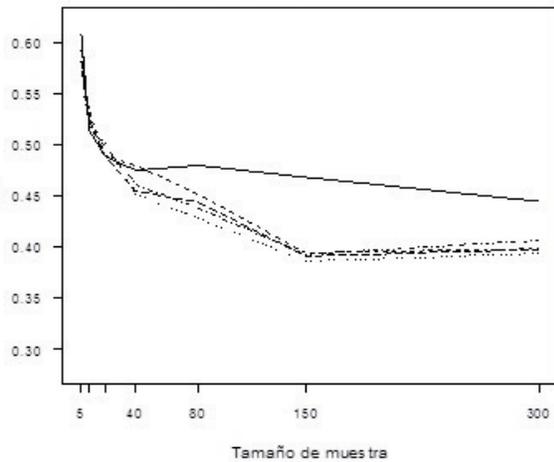


Figura 6. Promedio de los estadísticos de Kolmogorov-Smirnov para el percentil 95 de una gamma (2,2) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

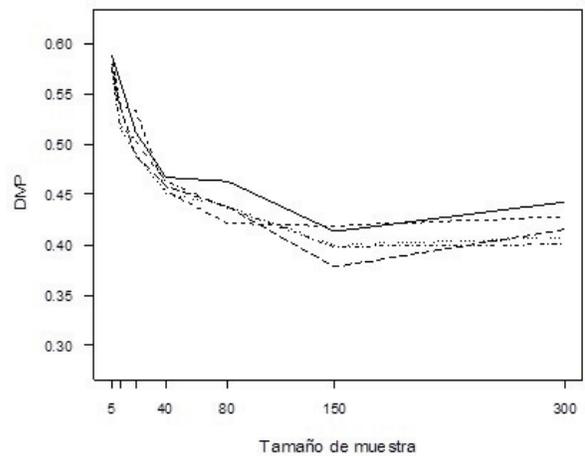


Figura 7. Promedio de los estadísticos de Kolmogorov-Smirnov para el percentil 95 de una gamma (5,1) con diferentes tamaños de muestra bootstrap (Nboot): — Nboot=50, - - - Nboot=100, Nboot=250, - · - · Nboot=500, - - - Nboot=1000.

Tabla 1. Tamaño de muestra recomendado bajo diferentes coeficientes de asimetría para la estimación de la distribución de algunos estadísticos

Coeficiente de asimetría	Media	Tamaño de muestra recomendado					
		Varianza	P. 5	P.25	P.50	P.75	P.95
2	40	150	150	80	150	80	--
1,414213562	40	150	150	80	150	80	--
1,154700538	80	80	150	80	150	80	--
0,894427191	80	80	150	80	150	80	--
0,666666667	80	80	150	80	150	80	--

P.5, P.25, P.50, P.75 y P.95 corresponden a los percentiles 5, 25, 50, 75 y 95 respectivamente.

Tabla 2. Número de muestras bootstrap recomendado bajo diferentes coeficientes de asimetría para la estimación de la distribución de algunos estadísticos

Coeficiente de asimetría	Media	Varianza	Número de muestras bootstrap recomendado				
			P. 5	P.25	P.50	P.75	P.95
2	250	250	100	100	100	--	100
1,414213562	250	250	100	100	100	--	100
1,154700538	250	250	100	100	100	--	--
0,894427191	250	250	100	100	100	--	--
0,666666667	250	250	100	100	100	--	--

4. Conclusiones

En este trabajo se realiza un estudio de simulación para determinar el efecto del tamaño de muestra y el número de muestras bootstrap en las estimaciones de la media, varianza, percentil 5,

25, 50, 75 y 95, bajo distribuciones con diferentes medidas de asimetría.

Se encontró que el tamaño de muestra tiene un efecto en las estimaciones de las distribuciones de los estadísticos de interés. En todos los casos

la *DMP* disminuye a medida que se aumenta el tamaño de muestra hasta un valor dado, pero en el caso del percentil 95, este comportamiento se puede apreciar. Por otro lado, se notó que la *DMP* disminuye dependiendo del número de muestras bootstrap utilizado, pero los valores recomendados en nuestros casos estudiados son de 100 o 200 muestras bootstrap.

5. Referencias bibliográficas

- Andrews, D. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, 68(2), 399-405.
- Bickel, P. J., Götze, F. & van Zwet, W R. (1997). Resampling fewer than n observations: gains, losses, and remedies for losses. *Statist. Sinica*, 7, 1-31.
- Bigerelle, M., Najjar, D., Fournier, B., Rupin, N., & Lost, A. (2006). Application of lambda distributions and bootstrap analysis to the prediction of fatigue lifetime and confidence intervals. *International Journal of Fatigue*, 28 (3), 223-236.
- Castillo, E., & Hadi, A.S. (1995). Modeling lifetime data with application to fatigue models. *Journal of the American Statistical Association*, 90(431), 1041-1054.
- Chatterjee, S., & Qiu, P. (2009) Distribution-free cumulative sum control charts using bootstrap-based control limits. *The Annals of Applied Statistics*, 3(1), 349-369.
- Chernick, M. (2008). *Bootstrap method: A guide for practitioners and researchers*. Second edition, John Wiley & Sons, Inc. New Jersey.
- Ching-Hui Chang, Jyh-Jiuan Lin, & Pal, N. (2011), Testing the equality of several gamma means: a parametric bootstrap method with applications, *Comput Stat.* 26, 55–76.
- Chou, C., Lin, Y., Chang, C. & Chen, C. (2006). On the bootstrap confidence intervals of the process incapability index C_{pp} . *Reliability Engineering & System Safety*, 91(4), 452-459.
- Chung, K., & Lee, S. (2001). Optimal bootstrap sample size construction of percentile confidence bounds. *Scandinavian Journal of Statistics*, 28, 225-239.
- Davison, A. C, & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press, Cambridge.
- Davison, A., Hinkley, D., & Young, G. (2003). Recent developments in bootstrap methodology. *Statistical Science*, 18(2), 141-157.
- Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26.
- Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82(397), 171-185.
- Efron, B. (2003). Second thoughts on the bootstrap. *Statistical Science*, 18(2), 135-140.
- Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman and Hall.
- Fethke, N. B., Anton, D., Cavanaugh, J. E., Gerr, F., & Cook, T. M. (2007) Bootstrap exploration of the duration of surface electromyography sampling in relation to the precision of exposure estimation. *Scandinavian Journal of Work, Environment & Health*, 33 (5), 358-367.
- Geluk, J., & Haan, L. (2002). On bootstrap sample size in extreme value theory. *Publicacions de L'Institut Matemàtic*, 71(85), 21-25.
- Götze, F., & Rackauskas, A. (2001). Adaptive choice of bootstrap sample sizes. *Lecture Notes-Monograph Series*, 36, 286-309.
- Halkos, G., & Tzeremes, N. (2010). Performance evaluation using bootstrapping DEA techniques: Evidence from industry ratio analysis. *International Journal of Fatigue*, 28 (3), 223-236.

Heiermann, K., Riesch-Opperman, H., & Huber, N. (2005). Reliability confidence intervals for ceramic components as obtained from bootstrap methods and neural networks. *Computational Materials Science*, 34(1), 1-13.

Hesterberg, T., Monaghan, S., Moore, D., Clipson, A., and Epstein, R. (2003). *The practice of business statistics*. W. H. Freeman and Company, New York.

Kosowski, R., Timmermann, A., Wermers, R., & White, H. (2006). Can mutual fund “stars” really pick stocks? New evidence from a bootstrap Analysis. *The Journal of Finance*, 61 (6), 2551-2595.

Marazzi, A. (2002). Bootstrap tests for robust means of asymmetric distributions with unequal shapes. *Computational Statistics and Data Analysis*. 39, 503-528.

Martin, M. (2007). Bootstrap hypothesis testing for some common statistical problems: A critical evaluation of size and power properties. *Computational Statistics & Data Analysis*. 51, 6321-6342.

Shao, J. (1994). Bootstrap sample size in nonregular cases. *Proceedings of the American Mathematical Society*, 122(4), 1251–1262.

Swanepoel, J. W H. (1986). A note on proving that the (modified) bootstrap works. *Comm. Statist. A* 15, 3193-3203.

Tung-I. Tsai & Der-Chiang Li (2008). Utilize bootstrap in small data set learning for pilot run modeling of manufacturing systems. *Expert Systems with Applications*, 35(3), 1293-1300.

Young, G., & Daniels., H. (1990). Bootstrap bias. *Biometrika*, 77(1), 179-185.