Anuario de Psicología
1992, nº 55, 143-154

Bootstrap: fundamentos e introducción a sus aplicaciones*

Antonio Solanas
Vicenta Sierra
Universidad de Barcelona

La técnica bootstrap *proporciona estimaciones del error estadístico, imponiendo escasas restricciones sobre las variables aleatorias analizadas y estableciéndose como un procedimiento de carácter general, independientemente del estadístico considerado. En este trabajo se realiza una presentación de los fundamentos teóricos de la técnica* bootstrap, *más desde una perspectiva divulgadora que estrictamente teórica, además de realizarse un breve estudio donde se podrá comparar la eficacia de la técnica frente a otras más establecidas.*
Palabras clave: Bootstrap, *error estadístico, remuestreo Monte Carlo, tasa de error tipo I, técnica percentil.*

The bootstrap is a general approach to estimate statistical error in situations where the usual statistical assumptions are not tenable. This article concerns basic bootstrap theoretical fundamentals and computer simulation data to compare the bootstrap and parametric approaches to estimating confidence intervals and type I error rates of the mean.
Key words: *Bootstrap, Statistical Error, Monte Carlo Resampling, Type I Error Rate, Percentil Method.*

Aunque puedan contemplarse las técnicas estadísticas desde una perspectiva independiente de las diferentes disciplinas científicas, la estadística aplicada debe considerarse desde el conocimiento del ámbito sustantivo sobre el cual se

Dirección de los autores: Antonio Solanas, Vicenta Sierra. Departamento de Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Universidad de Barcelona. Adolf Florensa s/n. 08028 Barcelona.
* Este trabajo ha sido realizado con la ayuda de una beca predoctoral de formación de profesorado en áreas deficitarias, otorgada por el Departament d'Ensenyament de la Generalitat de Catalunya a Vicenta Sierra.

realiza el análisis, básicamente para utilizar aquellos modelos estadísticos que se ajusten a las características de los datos de cada disciplina. En términos más precisos, los modelos estadísticos incorporan distintos supuestos (p. ej.: normalidad e independencia) que establecen restricciones sobre las variables aleatorias analizadas. A este respecto, las denominadas técnicas paramétricas incorporan, comparativamente a las pruebas no paramétricas, un conjunto de supuestos más restrictivos, implicando que las primeras puedan utilizarse en un menor número de contextos analíticos; pero las técnicas no paramétricas conllevan pérdida de potencia, en general, frente a las paramétricas. Este inconveniente, en cualquier caso, no justifica la elección indiscriminada de técnicas paramétricas para realizar análisis estadísticos, pues la violación de los supuestos por éstas requeridos incide sobre la estimación de la probabilidad asociada al estadístico, fundamento de la decisión estadística. La pérdida de potencia y la infraestimación de la probabilidad de error tipo I constituyen aspectos fundamentales en la aplicación de la estadística. Sobre este tópico, ha sido práctica habitual obviar el análisis de supuestos y utilizar las técnicas estadísticas sin considerar la adecuación o no de las mismas. Este uso puede considerarse crítico, especialmente cuando las técnicas no son robustas; por su parte, las técnicas no paramétricas han alcanzado niveles de difusión que no se corresponden con las amplias posibilidades analíticas que proporcionan.

Esta problemática ha sido contemplada por distintos investigadores en el campo de la estadística desde mediados del siglo XX (Quenouille, 1956; Tukey, 1958, 1977), poniéndose, en algunos casos, más énfasis en los aspectos descriptivos, mientras, en otros, se destacan los problemas relacionados con la inferencia estadística. Maurice Quenouille, en 1949, inició el estudio no paramétrico del error estadístico, desarrollando la técnica denominada *Jackknife*, siendo John W. Tukey quien, en 1950, perfeccionó esta estrategia analítica. La denominación *Jackknife* se debe a Tukey, quien pretendía recoger con el término la idea de una técnica de utilidad general. La técnica *Bootstrap*, propuesta por Bradley Efron (1979a) constituye otra estrategia no paramétrica para estimar el error estadístico, aunque pueden hallarse versiones paramétricas de la técnica, aspectos a los que posteriormente haremos referencia.

El interés en las técnicas no paramétricas, máximo en los ámbitos sustantivos en los cuales se desconoce la distribución de probabilidad de las variables aleatorias mesuradas, como ocurre en distintos contextos psicológicos, justifica el estudio de las aportaciones que pueda ofrecer la técnica *bootstrap* en las ciencias del comportamiento. Paralelamente, consideramos muy importante destacar la relevancia de la estimación del error estadístico, aspecto fundamental en la decisión estadística, que, contrariamente a lo que pudiera suponerse, no puede ser obtenido para la totalidad de los estimadores y, cuando es factible determinarlo, conlleva un elevado número de restricciones. En este punto debe notarse la relevancia de las técnicas que nos permiten obtener estimaciones del error estadístico (p. ej.: error estándar, error de predicción, sesgo, etc.) con el menor número de supuestos. Estos motivos justifican el presente trabajo, cuyo objetivo es divulgar la técnica *bootstrap*.

Fundamentos de la técnica *bootstrap*

Aspectos generales

Sea $\Theta(F)$ un parámetro dependiente de una desconocida función de distribución F y, por otro lado, considérense las variables aleatorias $X_1, X_2, ..., X_n$, tales que,

$$X_1, X_2, ..., X_n \underset{iid}{\sim} F$$

o sea, son variables aleatorias con función de distribución idéntica e independiente. Por otro lado, representaremos mediante el conjunto $\{x_1, x_2, ..., x_n\}$ una muestra correspondiente a extracciones aleatorias sobre las referidas variables. Nótese que los valores correspondientes a la muestra nos permiten obtener la distribución empírica $\hat{F}_n$, que constituye la estimación no paramétrica de máxima verosimilitud de la función de distribución F. En apoyo de esta estimación, podemos referirnos al Teorema de Glivenko-Cantelli, que establece una convergencia casi segura, cuando $n \to \infty$, entre las distribuciones F y $\hat{F}_n$, aunque debe notarse que es una convergencia asintótica.

$$\sup_x | \hat{F}_n(x) - F(x) | \underset{c.s}{\to} 0$$

Por tanto, es factible establecer $F = \hat{F}_n$, significando que la función de distribución de la variable aleatoria de interés se estima a partir de la distribución empírica, constituyendo el aspecto fundamental del denominado *Bootstrap no paramétrico.*

En ocasiones, el investigador conoce la función de distribución correspondiente a la variable aleatoria objeto de estudio, aunque se desconozcan los parámetros de la misma, punto que nos conduce al *Bootstrap paramétrico.* Supongamos que $\hat{\Theta} = \hat{\Theta}(x)$ es un estimador de Θ, siendo éste un parámetro o vector de parámetros de F, obteniéndose $\hat{\Theta}$ a partir de datos muestrales extraídos de la función de distribución F, que depende del parámetro Θ, por lo cual puede expresarse $F = F_\Theta$. Estimado el parámetro o vector de parámetros, podemos recurrir a la estimación $\hat{F} = F_{\hat{\Theta}}$.

El estadístico puede considerar la distribución empírica $\hat{F}_n$, retomando la expresión utilizada por Efron (1979b), *una cruda estimación de F.* La mencionada expresión pretende reflejar que la función de distribución empírica es escalonada, o sea, la estimación obtenida de F presenta discontinuidades, siendo este hecho, precisamente, el punto crítico cuando el estadístico supone la continuidad de F (función suave, derivable en todos los puntos) o posee información adicional que apoya la continuidad de la función de distribución F. En estos casos, es aconsejable suponer que F posee cierto grado de suavidad, o sea, se asemeja parcialmente a una función de distribución derivable en todos sus puntos, que denominaremos $\hat{F}^s$; donde *s* representa las posibles funciones de distribución suaves (por ejemplo: normal, uniforme, etc.). Puede establecerse que la mejor estimación de F proviene de una convolución de las funciones de distribución $\hat{F}_n$ y $\hat{F}^s$, obteniéndose, de esta forma, $\hat{F}^c$. Este procedimiento de estimación de F se ha denominado *bootstrap suavizado*, sobre el cual no entraremos en mayor detalle, hallándose descrito con amplitud en los trabajos de Efron (1979b, 1982).

Remuestreo mediante Monte Carlo y muestras bootstrap

Aunque la técnica Monte Carlo no puede considerarse un elemento inherente a la estrategia *bootstrap*, es preciso realizar una breve mención en la medida de su relevancia como técnica para obtener muestras aleatorias, extraídas a partir de un conjunto inicial de datos (remuestreo) o correspondientes a realizaciones de variables aleatorias cuya función de distribución de probabilidad ha sido prefijada (simulación de variables estocásticas). En concreto, dado que la técnica *bootstrap* genera diferentes muestras a partir de la muestra de datos recogida por el investigador (muestra original), e, independientemente de si se ha optado por un *bootstrap* paramétrico, no paramétrico o suavizado, es necesaria una estrategia para garantizar el remuestreo aleatorio, punto éste en el cual puede apreciarse el nexo entre la estrategia *bootstrap* y la técnica Monte Carlo.

La técnica Monte Carlo requiere la especificación de una función de distribución de probabilidad F(X), evidentemente, siendo conocidos sus parámetros o, en algún caso, estimados. El proceso para la obtención de distintas realizaciones de la variable aleatoria X, puede resumirse en la siguiente forma:

a) Especificar la función de distribución de probabilidad F(X).

b) Seleccionar un número aleatorio α_i tal que $\alpha_i \sim U(0,1)$. Se obtiene un número aleatorio (con mayor propiedad cabe considerar que se trata de un número pseudoaleatorio) comprendido entre 0 y 1 donde la probabilidad de obtener un valor perteneciente a un intervalo de amplitud prefijada es constante para cualquiera de los intervalos en que se ha dividido la recta real en el dominio comprendido entre 0 y 1 (función uniforme continua estándar).

c) Considerando F(X), obtener $x_i = F^{-1}(\alpha_i)$, siendo x_i el dato de la variable aleatoria X correspondiente al valor de probabilidad α_i. Nótese que en la función de distribución de una variable aleatoria continua existe una correspondencia biyectiva entre los valores de la variable aleatoria y la función de distribución de la misma, o sea, dado x_i, $F(x_i) = \alpha_i$ y $F^{-1}(\alpha_i) = x_i$, donde F^{-1} corresponde a la función inversa de F. Esto permite que, cuando se ha obtenido un número pseudoaleatorio, de forma inequívoca se pueda determinar el valor de la variable aleatoria correspondiente al mismo. Un razonamiento similar puede realizarse para variables aleatorias discretas.

d) Iterando los pasos b y c, pueden obtenerse cuantos valores de la variable aleatoria, cuya función de distribución ha sido especificada, sean necesarios.

Es importante, dada su relevancia en el proceso, que el generador de números pseudoaleatorios sea sometido a diferentes pruebas de aleatoriedad, a fin de garantizar, si bien es un problema de difícil solución a nivel práctico, un comportamiento estocástico de la sucesión de datos simulados. Este aspecto se aleja de la temática, por tanto el lector puede dirigirse a trabajos sobre la generación de números pseudoaleatorios, donde hallará diferentes estrategias, además de distintas pruebas de comportamiento aleatorio; pero, en cualquier caso, puede optarse por *software* comercializado.

Las técnicas de remuestreo se caracterizan por la obtención de submuestras a partir de los datos que constituyen la muestra original, permitiendo evaluar diferentes propiedades de los estimadores. En este sentido, la técnica *bootstrap* debe ser considerada como un plan de remuestreo, posibilitando estudiar el error estadístico, ya sea en cuanto a sesgo, error estándar o tasa de error en una predicción. Ahora bien, la forma en la cual la estrategia *bootstrap* realiza el remuestreo queda sintetizada en los siguientes pasos, englobados dentro del proceso de estimación *bootstrap*:

a) El investigador selecciona un estadístico de interés, sea $\hat{\theta}$ (media, mediana, correlación, tasa de error —en análisis discriminante, por ejemplo—, etc.).

b) Obtiene una muestra correspondiente a la realización de *n* variables aleatorias idéntica e independientemente distribuidas (puede conceptualizarse como *n* realizaciones de una variable aleatoria, sin pérdida de generalidad). Representaremos esta muestra original mediante $\{x_1, x_2, ..., x_n\}$.

c) El investigador debe decidir cómo estimará la distribución F, ya sea mediante *bootstrap* no paramétrico, paramétrico o suavizado. Este paso está ligado con el apartado a), cuando hacíamos mención a la técnica Monte Carlo.

d) Mediante la técnica Monte Carlo (apartados b y c), se obtiene un valor de la variable aleatoria *X*. Iterando este proceso *n* veces, se obtiene un conjunto de datos que constituyen la denominada *muestra bootstrap*, que representaremos mediante la notación $\{x^*_{1i}, x^*_{2i}, ..., x^*_{ni}\}$, denotando *i* que son los datos correspondientes a la i-ésima muestra *bootstrap*. Para cada una de estas muestras se obtiene el estadístico de interés que, para diferenciarlo del calculado sobre los valores de la muestra original, denotaremos mediante $\hat{\theta}^{*i} = \hat{\theta}\,(x^*_1, x^*_2, ..., x^*_n)$.

e) Repitiendo el apartado anterior el número de veces que se determine, supongamos que sean *b* ocasiones, se obtendrá la cantidad correspondiente de estimaciones del estadístico.

Una vez generadas las distintas muestras *bootstrap*, es factible realizar diferentes estimaciones del error estadístico, además de obtener estimaciones sobre parámetros de interés, aspectos que desarrollaremos en los apartados siguientes.

Estimaciones bootstrap *del error estadístico*

Puesto que se dispone de distintas estimaciones *bootstrap* del estadístico $\hat{\theta}$, el conjunto de estimaciones $\hat{\theta}^{*i}$ constituye la estimación *bootstrap* de la distribución muestral del estadístico de interés, que denominaremos *distribución muestral bootstrap de la distribución muestral de* $\hat{\theta}$.

Entre los errores estadísticos de mayor relevancia en el ámbito de la estadística, se halla el error estándar o desviación estándar de la distribución muestral de un estadístico. A este respecto, la estimación *bootstrap* de este error estadístico se obtiene mediante la expresión

$$\hat{\sigma}_{BOOT} = \left(\sum_{i=1}^{b} (\hat{\theta}^{*i} - \hat{\theta}^{*\cdot})^2 / (b - 1) \right)^{1/2}$$

donde $\hat{\theta}^{*\cdot}$ corresponde a la estimación *bootstrap* de la media de la distribución muestral *bootstrap*, obtenida mediante

$$\hat{\theta}^{*\cdot} = \sum_{i=1}^{b} \hat{\theta}^{*i}/b$$

También es factible obtener estimaciones del sesgo del estadístico de interés. Si $\hat{\theta}$ denota la estimación del estadístico objetivo sobre los datos de la muestra original, la estimación *bootstrap* del sesgo del estadístico se obtiene mediante

$$\widehat{\text{sesgo}}_{BOOT}(\hat{\theta}) = \hat{\theta}^{*\cdot} - \hat{\theta}$$

Otra interesante aplicación de la estimación *bootstrap* reside en los intervalos de predicción de los modelos de regresión, habitualmente obtenidos considerando que el término error del modelo posee determinadas características distribucionales. La técnica *bootstrap*, en este caso, estima la distribución del término error a partir de la estimación de la distribución *bootstrap* del componente residual. Así, los valores $\hat{G}^{-1}(\alpha/2)$ y $\hat{G}^{-1}(1-\alpha/2)$ determinan los límites del intrervalo de predicción, correspondiendo $\hat{y}+\hat{G}^{-1}(\alpha/2)$ e $\hat{y}+\hat{G}^{-1}(1-\alpha/2)$ al límite inferior y superior, respectivamente. En el siguiente apartado se expondrán con mayor detalle los aspectos fundamentales referentes a la distribución *bootstrap* del estadístico.

Puesto que no pretendemos realizar una revisión exhaustiva de la técnica *bootstrap*, el lector podrá hallar en las referencias bibliográficas distintas posibilidades de la técnica *bootstrap* para estimar el error estadístico (Efron y Tibshirani, 1986; Stine, 1985).

Intervalos de confianza bootstrap

Anteriormente nos hemos referido a la estimación *bootstrap* de la distribución muestral del estadístico, concepto fundamental cuando nos centramos en la elaboración de intervalos de confianza no paramétricos mediante la estrategia *bootstrap*. Debe matizarse que, aunque la distribución muestral *bootstrap* puede derivarse mediante procedimientos analíticos o mediante una expansión en series de Taylor (Efron, 1979a), generalmente es necesario recurrir a la técnica Monte Carlo, refiriéndonos en la exposición a ésta última posibilidad.

Representaremos mediante $\hat{G}(t)$ la estimación de la distribución *bootstrap* del estadístico $\hat{\theta}$, definida mediante

$$\hat{G}(t) = \text{card}\,\{\hat{\theta}^{*i} \mid \hat{\theta}^{*i} \leq t\}\,/b$$

Como puede notarse la estimación *bootstrap* de la distribución muestral se estima a partir de los estadísticos obtenidos en las diferentes muestras *bootstrap*; en concreto, la expresión anterior proporciona la estimación *bootstrap* de la función de distribución del estadístico. Por tanto, fijado un valor t, se estima la probabilidad de obtener un valor del estadístico inferior o igual al especificado mediante la proporción de valores que cumplan la referida desigualdad en la estimación de la distribución muestral *bootstrap*. Si el objetivo es determinar un valor de la distribución muestral del estadístico, la inversa de la función anterior, representada mediante $\hat{G}^{-1}(\alpha)$, permite obtenerlo.

Aunque el punto de referencia común sea la distribución muestral *bootstrap*, existen diversas estrategias para elaborar un intervalo de confianza, entre las cuales cabe destacar el *método percentil, percentil corregido para el sesgo* y *percentil corregido para el sesgo acelerado*. De hecho, esta enumeración recoge las estrategias desarrolladas por Bradley Efron, a las cuales deben añadirse otras como el *método percentil-t, percentil automático, iteración pivotal*, etc. Entrar en profundidad sobre las diferentes estrategias de elaboración de intervalos de confianza no es abordable en el presente trabajo, aunque, siempre desde una perspectiva introductoria, nos referiremos con mayor detalle al método percentil; si bien no está exento de consideraciones más o menos complejas de estadística matemática, hemos optado por una presentación simplificada al máximo.

El estadístico aplicado está familiarizado en la construcción de intervalos de confianza, proceso en el cual impone determinados supuestos sobre la variable aleatoria analizada, y, tras obtener una estimación del estadístico de interés (habitualmente se trata de la estimación de máxima verosimilitud), obtiene una estimación del error estándar de la distribución muestral del estadístico. Fijado un nivel de confianza 1-α, el intervalo de confianza (IC) se obtiene mediante la conocida expresión

$$\hat{\Theta} \mp Z_{\alpha/2} \cdot \hat{\sigma}_{\hat{\Theta}}$$

donde $\hat{\sigma}_{\hat{\Theta}}$ es la estimación del error estándar de la distribución muestral del estadístico y $Z_{\alpha/2}$, bajo ciertas suposiciones, se corresponde con los percentiles asociados al nivel de confianza establecido. Este procedimiento exige un conocimiento sobre las variables aleatorias objeto de estudio que, en ocasiones (quizá, en muchos ámbitos de investigación, en la mayoría de los casos), no se posee. Si el tamaño muestral es escaso la anterior construcción del intervalo de confianza se modifica recurriendo a la distribución *t*, que se ha mostrado eficaz, utilizándose el término en el sentido de concordancia entre tasa de error nominal y empírica para cualquier tamaño muestral y siempre que las desviaciones de la normalidad no sean excesivas.

Pero, ¿qué puede aportar la técnica *bootstrap*? En primer lugar, es posible que la función de distribución de la variable aleatoria sea tal que la eficacia de los procedimientos habituales pueda mejorarse. A este respecto, ¿qué ocurre si la variable aleatoria posee distribución uniforme? Si el tamaño muestral es elevado no parece factible mejorar la estimación estándar, pues es conocido que la suma de *n* variables aleatorias con distribución uniforme se aproxima con rapidez a la distribución normal; pero, en cualquier caso, con escaso número de datos la distribución *t* proporciona una adecuada aproximación a la distribución muestral del estadístico. La prueba fundamental de la técnica *bootstrap* consiste en que se comporte con precisión ante una amplia variedad de situaciones, motivo por el cual es práctica habitual contrastar la estrategia *bootstrap* en diversas condiciones y realizar análisis comparativos con otras técnicas. Así, para un procedimiento que se pretende general, sin restricciones distribucionales, es también necesario demostrar su adecuación en contextos analíticos sobre los cuales ya poseemos técnicas que han manifestado su adecuado funcionamiento. Otras perspectivas se han encaminado hacia la posibilidad de corregir el sesgo del estimador y, por tanto, lograr intervalos que se ajusten a los límites teóricos. Pero, la mayor aportación de la estrategia *bootstrap*, en el contexto de la construcción de intervalos de confianza, reside en la posibilidad de disponer de un procedimiento general, solventando las dificultades que se desprenden del desconocimiento de la distribución muestral de los estadísticos. Pero, puesto que no se trata de una técnica analítica, en el sentido de derivar teóricamente los intervalos exactos, no es factible determinar su adecuación sin recurrir al análisis comparativo.

La construcción de intervalos mediante el procedimiento percentil obtiene los límites de los mismos en base a la inversa de la estimación *bootstrap* de la distribución muestral del estadístico. De esta forma, los percentiles $\hat{G}^{-1}(\alpha/2)$ y $\hat{G}^{-1}(1\text{-}\alpha/2)$ determinan el límite inferior y superior, respectivamente. A nivel operativo, una vez han sido generadas todas las muestras *bootstrap*, y mediante el estadístico de orden, puede obtenerse la estimación *bootstrap* de la distribución muestral del estadístico, coincidiendo los límites inferior y superior con los valores $x_{(b(\alpha/2))}$ y $x_{(b(1-\alpha/2))}$, donde b es el número de muestras *bootstrap* obtenido y el paréntesis mayor representa la ordenación de los datos.

Breve estudio sobre el procedimiento percentil

Para ejemplificar el procedimiento de elaboración de un intervalo de confianza *bootstrap* mediante la técnica percentil, se generaron 10 realizaciones de una variable aleatoria con distribución de probabilidad N(10,25), consistiendo el objetivo en estimar el valor de la media. La utilización de la técnica percentil no sólo se justifica en la elección de un procedimiento simple, a fin de mostrar el proceso, puesto que también puede considerarse como una estrategia adecuada cuando el estimador carece de sesgo. La muestra original, obtenida mediante la simulación, se halla en la Tabla 1, además de otros datos de interés. Mientras el intervalo exacto del estadístico, fijado $1-\alpha=0.1$, es $[7.4 \div 12.6]$, la estimación normal proporciona un intervalo $[8.963 \div 11.247]$; pero, dado que el tamaño muestral es escaso, puede considerarse más adecuada la construcción del intervalo en base a la distribución de Student, obteniéndose $[8.832 \div 11.378]$. Aunque puede notarse que en ambas estrategias de estimación el parámetro poblacional pertenece al intervalo, cabe destacar que la reproducción del mismo no es muy precisa, pues la amplitud de los intervalos es inferior a la esperada y la distancia a los límites esperados bastante elevada. Sobre este punto, el intervalo obtenido mediante la estimación *bootstrap* del error estándar, cuyo valor es 0.639, presenta la misma característica que los anteriores: excesiva amplitud del mismo. En este

TABLA 1. VALORES CORRESPONDIENTES A LA MEDIA Y ERROR ESTÁNDAR DE LA DISTRIBUCIÓN MUESTRAL *BOOTSTRAP*, OBTENIDA A PARTIR DE 100 MUESTRAS *BOOTSTRAP* GENERADAS POR EL ALGORITMO MONTE CARLO A PARTIR DE LOS DATOS MUESTRALES (10.25379, 9.049, 13.16632, 8.042605, 8.1233337, 13.24509, 12.63261, 9.524437, 9.741493, 7.276687), CUYA MEDIA Y VARIANCIA SON, RESPECTIVAMENTE, 10.10554 Y 4.822998

Muestra	Media	Variancia	Muestra	Media	Variancia
1	10.205640	4.101766	51	10.381200	2.822320
2	9.827989	4.171061	52	10.350760	5.156318
3	11.458920	4.589573	53	10.624370	3.166775
4	9.979952	2.873359	54	10.789744	4.847006
5	10.290950	5.123088	55	9.278691	3.204468
6	10.391720	4.172241	56	9.809025	4.316515
7	9.556492	2.056552	57	9.698462	2.135790
8	11.062950	5.398397	58	10.425840	5.417725
9	11.100500	4.918295	59	10.096740	6.495023
10	9.250974	2.871189	60	9.912020	6.382725
11	9.213295	2.717082	61	9.862844	5.071222
12	10.395910	4.149604	62	10.352020	3.881063
13	9.694277	4.221307	63	9.920154	4.902012
14	10.598280	5.398763	64	9.632731	3.095724
15	11.427160	5.565308	65	11.195120	4.555515
16	9.791852	5.496623	66	10.540590	5.326362
17	9.382947	2.650689	67	9.943721	5.215766
18	10.524420	7.235026	68	10.466360	3.927246
19	8.860502	3.250054	69	10.214530	5.018365
20	10.215710	7.084893	70	10.002760	4.816447
21	9.935846	5.158576	71	10.513050	5.899584
22	10.129760	2.869710	72	9.554697	2.763048
23	9.928306	5.553230	73	10.036290	4.926982
24	10.030060	4.626425	74	9.934134	5.973348
25	10.097460	4.859213	75	10.654240	5.867636
26	11.495790	4.180732	76	10.303660	2.462240
27	9.499108	4.969503	77	9.664118	0.425049
28	10.600200	5.212606	78	10.705470	5.789972
29	9.075514	2.474413	79	11.098420	4.600179
30	10.589900	4.395020	80	10.005090	5.070069
31	10.657210	3.774034	81	10.295340	5.010742
32	10.788920	6.622749	82	9.797809	5.206591
33	9.684062	3.904507	83	10.659250	4.750163
34	11.080500	4.101427	84	10.472570	4.135973
35	10.494050	6.319865	85	9.695092	3.468567
36	9.914664	3.627625	86	9.286694	2.724162
37	9.990919	4.091397	87	8.989834	2.598667
38	10.955940	7.172133	88	9.268716	2.791301
39	9.316626	4.066040	89	10.275950	2.787326
40	10.420920	6.872342	90	10.926390	5.061388
41	10.196760	2.980388	91	9.374389	4.221626
42	9.471998	2.221761	92	9.705834	3.990533
43	10.509640	3.365248	93	10.227160	6.070774
44	10.739080	6.532281	94	9.834112	3.334642
45	11.113290	3.288289	95	9.902759	3.470418
46	9.686462	4.017965	96	9.953394	4.564691
47	10.375010	5.344944	97	10.626510	3.507962
48	11.669870	3.131144	98	9.218829	2.692010
49	9.259055	3.075745	99	8.584343	2.699782
50	11.774470	3.074436	100	9.878884	5.272332

sentido, ¿qué ocurre cuando se elabora un intervalo *bootstrap* mediante el procedimiento percentil? En este caso, con 100 muestras *bootstrap*, se obtuvo [9.21 ÷ 11.2], reflejando una disminución de la amplitud del intervalo comparativamente a los anteriores. Ciertamente, el intervalo *bootstrap* contiene el parámetro, lo cual podría sugerir una primera valoración optimista y considerar que mejora las anteriores estimaciones en la medida que proporciona un rango de valores más estrecho; pero esta primera impresión no puede descontextualizarse del incremento en la tasa de error tipo I que conlleva no reproducir exactamente el intervalo teórico cuando éste no se recubre totalmente.

Puesto que los anteriores resultados se obtuvieron a partir de una única muestra original, constituyendo más un ejemplo ilustrativo que un intento de aproximarse al problema de la estimación de parámetros mediante intervalos, realizamos un breve experimento de simulación. En una primera condición fueron generadas 200 muestras originales, especificándose las mismas características anteriores sobre la variable aleatoria. En la segunda condición, idéntica en cuanto al número de muestras generadas, se especificó que la variable aleatoria tuviera distribución de probabilidad $U(10-5\cdot3^{1/2}, 10+5\cdot3^{1/2})$, por lo cual coincidían media y variancia de ambas variables aleatorias. El número de muestras *bootstrap* generadas a partir de los datos de cada muestra original fue únicamente de 100, muy escaso si consideramos las recomendaciones realizadas para otros estadísticos sobre los cuales se han elaborado intervalos de confianza *bootstrap*, situándose los valores aconsejados entre 500 y 2000. A partir de los datos de cada una de las muestras originales se obtuvieron los intervalos de confianza siguientes: normal, student, percentil y studentboot. En el último de éstos se estimaba el error estándar mediante los datos de la estimación de la distribución muestral del estadístico, utilizándose este dato para construir el intervalo de confianza considerando la distribución de Student. Los resultados del experimento se muestran en la Tabla 2.

TABLA 2. TASA DE ERROR EMPÍRICA Y AMPLITUD PROMEDIO DEL INTERVALO DE CONFIANZA DEL ESTADÍSTICO MEDIA. SE GENERARON 200 MUESTRAS ORIGINALES DE TAMAÑO 10, FIJÁNDOSE EL NÚMERO DE MUESTRAS *BOOTSTRAP* Y LA TASA DE ERROR NOMINAL EN 100 Y α=0.1, RESPECTIVAMENTE. LAS DIFERENTES ESTIMACIONES SE INDICAN EN LA TABLA, ADEMÁS DE LAS CONDICIONES DEL EXPERIMENTO

	Uniforme		Normal	
	Tasa de error	Amplitud	Tasa de error	Amplitud
Normal	0.140	5.118	0.160	4.806
Student	0.090	5.703	0.140	5.355
Studentboot	0.100	5.386	0.145	5.049
Percentil	0.160	4.961	0.170	4.662

Los resultados obtenidos ponen de manifiesto que el procedimiento percentil muestra una mayor tasa de error tipo I, aunque el valor empírico no dista excesivamente de la tasa nominal. Esta mayor tasa de error es consecuencia de la menor amplitud de los intervalos, por lo cual no recubre adecuadamente el intervalo exacto, incrementándose los errores en la estimación. Por el contrario, como era esperado, la elaboración de intervalos de confianza fundamentados en la distribución t se muestra aceptable, en general. Cabe destacar que la estimación *bootstrap* del error estándar permite elaborar intervalos más estrechos, por tanto más informativos, dato relevante cuando se mantiene una tasa de error tipo I aceptable y se reproduce con bastante precisión la amplitud del intervalo exacto.

Si la tasa de error tipo I del procedimiento percentil se mostrara algo menor, podría considerarse la ventaja que supone la escasa amplitud del intervalo de confianza percentil; pero, como puede notarse, la tasa de error se mantiene en niveles similares al intervalo estándar. Estos resultados, en principio contrarios a la utilización del procedimiento percentil en la elaboración del IC del estadístico media, mejorarían si el número de muestras *bootstrap* hubiera sido mayor, puesto que la estimación de la distribución muestral *bootstrap* es excesivamente cruda. Incorporar un mayor número de muestras *bootstrap* sería deseable y, de la suavización resultante, seguramente se mejoraría la tasa obtenida en la condición de la variable aleatoria normal. Con menor plausibilidad ésta sería una explicación de los datos obtenidos en la condición de la variable aleatoria uniforme.

Discusión

La técnica *bootstrap* se propone como un procedimiento general para estimar el error estadístico en cualquiera de sus formas, aunque esta amplia aplicabilidad conlleva que existan técnicas que, bajo determinadas condiciones, mejoran las estimaciones del error estadístico. La breve investigación realizada es un ejemplo de este hecho, pues como hemos podido comprobar, para el estadístico media, existen los intervalos de confianza elaborados mediante la distribución t, que reproducen con mayor precisión la tasa de error nominal. A este respecto, posiblemente, los resultados obtenidos al utilizar el procedimiento percentil hubieran mejorado si el número de muestras *bootstrap* hubiera sido mayor; pero, en cualquier caso, no hemos podido observar un mejor comportamiento de la técnica *bootstrap* cuando los datos procedían de poblaciones con distribución uniforme.

En el estudio se ha evaluado el procedimiento percentil frente a otras técnicas que sabíamos previamente que mostraban niveles de reproducción de la tasa nominal muy aceptables, pero es necesario probar otras condiciones de simulación. Entre éstas, puede modificarse la distribución de la variable aleatoria, el estadístico objetivo, la presencia de dependencia entre las observaciones, etc. Esta evaluación sistemática permitirá, progresivamente, determinar si la técnica *bootstrap* puede ser adoptada frente a otras estrategias.

Un punto favorable al procedimiento percentil consiste en que elabora intervalos de confianza estrechos, lo cual permite considerar más informativos los intervalos elaborados. Pero este aspecto destacable necesita combinarse con la ya mencionada reproducción de la tasa nominal de error tipo I; en caso de no producirse este hecho, el estadístico cometerá un mayor número de errores en su decisión.

En el ámbito psicológico la técnica *bootstrap* tiene el mismo campo de aplicación que en cualquier disciplina, únicamente destacar que los datos procedentes de la medición del comportamiento humano no están exentos de características que violan los supuestos de diversos modelos estadísticos, hecho que incide sobre la tasa de error tipo I de las técnicas clásicas, especialmente en investigaciones donde se obtienen medidas repetidas. En otro tópico psicológico, la técnica *bootstrap* puede ser utilizada en el proceso de selección de ítems, proporcionando estimaciones del error estándar de los estadísticos utilizados para obtener los ítems adecuados.

Referencias

Efron, B. (1979a). The 1977 Rietz lecture. *The Annals of Statistics, 7*(1), 1-26.

Efron, B. (1979b). Computers and the theory of statistics: thinking the unthinkable. *Siam Review, 21*(4), 460-480.

Efron, B. (1982). *The Jackknife, the bootstrap and other resampling plans.* Philadelphia: Society for Industrial and Applied Mathematics.

Efron, B. & Tibshirani (1986). Bootstrap methods for standar errors, confidence intervals, and other measures of satatistical accuracy. *Statistical Science, 1*(1), 54-77.

Quenouille, M.H. (1956). Notes on bias estimation. *Biometrika, 43*, 353-360.

Stine, R.A. (1985). Bootstrap prediction intervals for regression. *Journal of the American Statistical Association, 392*(80), 1026-1031.

Tukey, J. (1958). Bias and confidence in not quite large samples abstract. *Annals of Mathematic Statistics, 29*, 614.

Tukey, J. (1977). *Exploratory data analysis.* Reading, M.A.: Addison-Wesley.