

Introducción al Bootstrap. Desarrollo de un ejemplo acompañado de software de aplicación

Rubén Ledesma

CONICET / Universidad Nacional de Mar del Plata, Argentina

El bootstrap es un tipo de técnica de remuestreo de datos que permite resolver problemas relacionados con la estimación de intervalos de confianza o la prueba de significación estadística. Este enfoque puede resultar de interés para los investigadores en Psicología, no solo porque es menos restrictivo que el enfoque estadístico clásico, sino también porque es más general en su formulación y más simple de comprender en lo referente al procedimiento básico que subyace al método. En lugar de fórmulas o modelos matemáticos abstractos, el bootstrap simplemente requiere un ordenador capaz de simular un proceso de muestreo aleatorio de los datos. Sin embargo, y debido quizás a la escasa difusión de la técnica, los investigadores aún no han incorporado el bootstrap al repertorio habitual de herramientas de análisis de datos. En este trabajo realizamos una presentación conceptual del bootstrap, ilustramos la técnica mediante un ejemplo y revisamos algunas opciones disponibles en materia de software estadístico. El trabajo incluye además un programa para correr el ejemplo dentro de ViSta "The Visual Statistics System", un sistema estadístico gratuito y abierto.

The Bootstrap is a data resampling technique that can be used to aid in different problems related to confidence intervals (CI) estimation and significance hypothesis testing. This approach can be of interest for researchers in Psychology; this is because the bootstrap is less restrictive than the classic statistical approach, but also because the underlying procedure is more general and easier to understand. In fact, instead of complex formulas and mathematical models, bootstrap simply requires a personal computer capable of simulating a random sampling process. Nevertheless, the bootstrap is not incorporated yet as a data analysis tool in Psychology, perhaps due to the little diffusion of this technique. This paper: a) briefly introduces the basis of the bootstrap, b) presents a heuristic example that illustrates the bootstrap procedure, and c) describes some statistical software useful in carrying out bootstrap applications. We also include a bootstrap computer module that can be used to run the example into ViSta "The Visual Statistics System", a free and open statistical system.

Las computadoras modernas han hecho posible ciertas formas de manipulación y análisis de datos antes inconcebibles. La visualización dinámica de datos, las técnicas de minería, la simulación y los métodos de remuestreo son algunos ejemplos de enfoques que, apoyados en las capacidades de los ordenadores modernos,

han enriquecido el trabajo de exploración y análisis estadístico de datos. La relación entre ordenadores y análisis de datos es patente para el caso de las llamadas técnicas de *remuestreo de datos* ('data resampling'), entre las que encontramos el Jackknife, los test de aleatorización y permutación, la validación cruzada y el *bootstrap* (Yu, 2003).

Estas se basan, justamente, en el uso intensivo del cálculo o cómputo por ordenador como estrategia para resolver problemas de análisis estadístico. En cierta forma, el remuestreo es un enfoque revolucionario, ya que aprovechando las capacidades de las nuevas tecnologías ofrece al investigador una alternativa sustancialmente diferente al enfoque estadístico clásico y, además, conlleva algunas ventajas comparativas que comentaremos más adelante.

No obstante, estas nuevas técnicas son aún poco conocidas y utilizadas en la práctica, en parte debido a su escasa difusión y también a que la formación estadística de los investigadores sigue centrada mayormente en la estadística clásica. Este es el caso del bootstrap (Efron, 1979; Efron y Tibshirani, 1993), el método de remuestreo más prototípico y general, así como el más desarrollado y estudiado por los expertos, aunque aún escasamente utilizado por el investigador en Psicología. En este contexto, el presente trabajo se propone: a) realizar una presentación conceptual del método bootstrap como técnica de remuestreo de datos y analizar sus posibles ventajas en el contexto de la investigación psicológica, b) ilustrar la lógica y funcionamiento general del método mediante una aplicación concreta: la estimación de intervalos de confianza para el coeficiente de correlación r de Pearson, y c) mostrar como este tipo de análisis puede realizarse mediante dos programas gratuitos: *Resampling* (Howell, 2001) y *ViSta* (Young, 1996). En el último caso, se trata de un software que hemos desarrollado específicamente para este trabajo y que funciona como módulo del sistema libre y gratuito *ViSta*. La presentación carece de referencias matemáticas o analíticas, en su lugar, enfatizamos los conceptos básicos subyacentes al enfoque y su aplicación mediante la asistencia de programas de cálculo apropiados para este fin.

Introducción Conceptual al Bootstrap

El bootstrap como método fue conceptualizado y descrito sistemáticamente por Efron (1979) y Efron y Tibshirani (1993), aunque se ha señalado que algunas ideas básicas relacionadas con este método pueden encontrarse en trabajos de autores previos (Hall, 2003). Se trata más que de una técnica o modelo específico, de un método general a partir del cual pueden cubrirse diferentes objetivos de análisis de datos, incluyendo la estimación de intervalos de confianza (IC) o el Test de significación estadística. En esencia, veremos que el método permite aproximar la distribución de muestreo de un estadístico y de sus propiedades mediante un procedimiento muy simple: Crear un gran número de muestras con reposición de los datos observados. Recordemos que la distribución de muestreo de un estadístico es clave a la hora de realizar tareas de inferencia estadística. Por ejemplo, para crear un IC

alrededor de un parámetro poblacional, como la media, es necesario disponer de una estimación del error típico, que es una propiedad de su distribución de muestreo. Pues bien, el bootstrap permite estimar estas propiedades aproximando empíricamente la distribución de muestreo del estadístico en cuestión. Esta capacidad general del enfoque ha demostrado ser eficiente en una gran variedad de situaciones, en temas tan diversos como el análisis de correlación y regresión, el análisis factorial exploratorio y confirmatorio, el análisis discriminante y el análisis de propiedades de pruebas e items, entre otros (Fan, 2003). Además, también se ha sugerido que el mismo procedimiento posee utilidad si solo se persiguen propósitos descriptivos o exploratorios (Thompson, 1995).

A nivel conceptual, el bootstrap puede incluirse dentro del enfoque mayor del *remuestreo* de datos o *resampling*, el cual comprende una variedad de técnicas que implementan procesos de simulación por computadora para estimar probabilidades empíricas a partir de experimentos con los propios datos (Rudner y Shafer, 1992). La denominación de *remuestreo* se debe a que los métodos se basan, esencialmente, en la extracción de un gran número de muestras repetidas de los propios datos, y sobre esta base se realizan posteriormente descripciones e inferencias estadísticas. Se trata pues, de una estrategia general para resolver problemas de probabilidad y estadística aplicada (Simon, 1997), dentro del cual se pueden diferenciarse algunos métodos o procedimientos más específicos, entre los cuales el bootstrap resulta el más estudiado y extendido.

En efecto, el bootstrap comparte con otras técnicas de remuestreo el procedimiento esencial de usar simulación por computadora o cómputo intensivo como medio para responder preguntas de análisis de datos, lo que diferencia a este tipo de técnicas de los procedimientos estadísticos clásicos que encuentran su apoyo en modelos teóricos desarrollados de forma analítica. Por ello, algunos autores se refieren a estas técnicas como *métodos intensivos de cómputo* y las incluyen dentro un *enfoque estadístico moderno*, en oposición al enfoque matemático clásico (Simon, 1997). Por otro lado, en español también se suele utilizar la expresión *métodos autosuficientes* de estimación. Justamente, la expresión en inglés 'pull oneself up by one's bootstraps', de donde se toma el término bootstrap, quiere significar el hecho de 'salir adelante sin ayuda' o ser auto-suficiente. En el contexto de análisis de datos esta expresión alude a que solo se utiliza la información disponible en los propios datos, prescindiendo de la ayuda externa de modelos o supuestos teóricos.

Como hemos mencionado, el bootstrap puede ser aplicado en la práctica para estimar IC y realizar Test de Hipótesis, aunque también con fines exploratorios y descriptivos, como diagnosticar modelos o evaluar la

replicabilidad de los resultados de un estudio. En general, podemos decir que los problemas que pueden resolverse mediante este nuevo enfoque son similares a los que pueden tratarse mediante los métodos convencionales, aunque varios autores coinciden en señalar algunas ventajas comparativas del primero (Hesterberg, Monaghan, Moore, Clipson, y Epstein, 2003; Rudner y Shafer, 1992; Simon, 1997). Entre los beneficios comparativos hay acuerdo en señalar que el bootstrap es *más general* en su formulación, resulta *más accesible y simple de comprender*, y requiere *menos supuestos* que los métodos clásicos.

Efectivamente, la *generalidad* es una de las principales ventajas del enfoque bootstrap, ya que la misma metodología básica puede usarse en una gran variedad de situaciones, sin necesidad de acudir a modelos, expresiones o fórmulas específicas para cada problema (Efron y Tibshirani, 1993; Hesterberg, Monaghan, Moore, Clipson y Epstein, 2003; Simon, 1997). En el próximo apartado daremos una descripción más detallada del procedimiento, pero a modo de introducción podemos decir que, más allá de la aplicación concreta de que se trate, el procedimiento básico resulta siempre el mismo. Partiendo de una muestra de datos para los cuales se calcula un estadístico de interés (por ejemplo, una media o un coeficiente de correlación), el método consiste en: 1) crear un gran número de sub-muestras con reposición de los mismos datos, por ejemplo, 2000 muestras, y 2) calcular para cada muestra resultante el valor del estadístico en cuestión. Se obtiene así una aproximación a la distribución de muestreo del estadístico, a partir de la cual podemos construir un IC para dicho estadístico o realizar una prueba de significación. Como se observa, el procedimiento es relativamente sencillo y proporciona una estrategia bastante general para enfrentar diferentes tipos de situaciones, independientemente del tipo de datos y variables que son objeto de análisis.

Por otro lado, el procedimiento bootstrap resulta relativamente *simple de comprender*, aspecto que se ha señalado como otra ventaja importante de las técnicas de remuestreo, especialmente en el ámbito de la docencia. Entre otros autores, Braun (1995), Hesterberg (1998), Ricketts y Berry (1994), Simon (1997), Wood (2005) y Yu (2003) han planteado que este enfoque puede resultar beneficioso en el contexto de la enseñanza, ya que los estudiantes parecerían comprender mejor ciertos conceptos si se utiliza esta perspectiva en lugar del método clásico. Esto se explicaría porque la descripción de los métodos no requiere fundamentos ni expresiones matemáticas que pueden resultar complejos o abstractos para muchos estudiantes, simplemente se necesita poder simular un proceso de remuestreo de datos, lo que en la actualidad puede hacerse fácilmente mediante un software apropiado. En síntesis, podemos decir que el remuestreo proporciona una analogía

más concreta de los conceptos estadísticos y favorece un aprendizaje más activo, ya que los estudiantes pueden controlar y experimentar por ellos mismos los resultados de la simulación (Wood, 2005).

Por último, el bootstrap no es exigente en cuanto al cumplimiento de supuestos teóricos para su aplicación y es, en este sentido, *menos restrictivo* que las técnicas convencionales. Para empezar, es posible prescindir de supuestos relativos a las distribuciones, ya que en vez de asumir a priori una determinada distribución teórica, se utiliza la muestra original y se generan un gran número de sub-muestras que sirven de base para estimar inductivamente la forma de la distribución muestral de los estadísticos. De esta manera, pueden analizarse datos provenientes de distribuciones desconocidas o incluso abordarse problemas complejos, frente a los cuales no hay una solución analítica conocida (Efron y Tibshirani, 1993). Esto último sucede, por ejemplo, cuando trabajamos con estadísticos para los cuales no se conoce la forma de su distribución muestral, como puede ser el caso de la diferencia entre dos medianas. La derivación analítica de fórmulas para estos casos resulta muy compleja, por eso, como señalan Efron y Tibshirani (1993), la estadística clásica se ha desarrollado principalmente para estadísticos simples y para un conjunto limitado de distribuciones. El bootstrap, en cambio, reemplaza el trabajo analítico por un procedimiento empírico, de modo que ofrece soluciones al investigador cuando no hay 'fórmulas' disponibles para resolver un problema.

En cuanto al las técnicas no-paramétricas, que podrían ser una alternativa cuando no se cumplen ciertos supuestos, hay que señalar que estas generalmente requieren reducir las variables en sus niveles de medida - por ejemplo, puntuaciones de test a posiciones o rangos -, perdiendo así una parte de la información disponible en los datos (Gil-Flores, 2005). Esta reducción del tipo de escala de medida no es necesaria en el bootstrap. Además, también es conveniente mencionar que no hay alternativas no-paramétricas desarrolladas para todos los correspondientes modelos paramétricos.

Se ha señalado que la validez del bootstrap se basa en un único supuesto o principio que subyacente al procedimiento en sí mismo. Puesto que el enfoque se apoya en la capacidad de la muestra para reflejar o representar los aspectos relevantes de la población de la cual fue extraída, la calidad de la muestra resulta crucial, sea porque no ha sido extraída por medio de un procedimiento que asegure cierta representatividad o sea porque su tamaño es demasiado pequeño (Gil-Flores, 2005). En estas situaciones, puede ser dudoso que la información ofrecida por la muestra permita estimar las propiedades desconocidas de la población. No obstante, estas limitaciones son igualmente aplicables a las

técnicas clásicas, ya que también se basan en la idea de selección aleatoria y dependen en su eficiencia del tamaño de la muestra. Así, se ha afirmado que los métodos bootstrap permitirían ‘extraer lo máximo a partir de la poca información disponible’ (Chernick, 1999, p. 149, en Gil-Flores, 2005).

Algunas barreras y limitaciones en el uso del bootstrap

Las razones esgrimidas en favor del bootstrap hacen suponer que se trataría de un enfoque potencialmente útil para el análisis de datos en Psicología, así como para la enseñanza de la estadística en el área. Sin embargo, esta metodología no ha sido incorporada en toda su extensión al repertorio de herramientas que se aplican en la investigación psicológica (Fan, 2003). Una razón de este hecho sería el escaso conocimiento que los investigadores tienen de las técnicas de remuestreo, puesto que a pesar de su relativa simplicidad, no se incluyen como contenido en la formación de grado o posgrado. Por otro lado, su uso tampoco ha sido impulsado por las asociaciones o las normas editoriales de nivel internacional, como si ha sucedido con otras metodologías. Por ejemplo, el esfuerzo realizado por la APA para incentivar la inclusión de medidas del tamaño de efecto en todas las publicaciones psicológicas (Wilkinson y APA, 1999).

Otra barrera añadida pareciera ser que las técnicas bootstrap no estén incorporadas de forma accesible a los programas estadísticos más populares en Psicología (Fan, 2003), como SPSS. En esto nos sumamos a la opinión de Sánchez-Bruno y Borges del Rosal (2005), quienes han manifestado que lo no está en SPSS “parece no existir” para el investigador en Psicología. Aquí, es bueno señalar que en algunos aspectos el software libre y gratuito puede resultar superior a los programas de origen comercial (Ledezma, 2004). Para el caso del bootstrap veremos en el próximo apartado que existe software de este tipo que puede utilizarse para realizar los análisis de forma eficiente.

Por otro lado, Yu (2003) manifiesta que además de los factores mencionados, existiría una dosis de ‘conservadurismo metodológico’ por el cual los investigadores preferirían los métodos clásicos, aún incluso existiendo evidencia en favor de la superioridad de las nuevas técnicas. Este autor entiende que los investigadores se muestran reticentes a aplicar procedimientos que pueden resultar ‘marginales’ y agrega: “Traditional procedures are perceived as founded on solid theoretical justification and empirical substantiation, while newer techniques face harsh criticisms and seem to be lacking theoretical support” (Yu, 2003, s.p).

En todo caso, parece conveniente insistir en la difusión de este tipo de técnicas y remarcan sus posibles beneficios en la investigación psicológica. Asimismo, entendemos que

es necesario facilitar el acceso a las tecnologías informáticas que pueden asistirnos en su aplicación, bien sea con fines científicos o educativos.

Desarrollo De Un Ejemplo

Para aclarar mejor el funcionamiento del método de bootstrap plantearemos un ejemplo tomando del propio manual de Efron y Tibshirani (1993), comentaremos los pasos en la aplicación del procedimiento y mostraremos como puede resolverse el caso mediante dos programas gratuitos diferentes: *Resampling* y *ViSta*.

Planteo del Caso: IC para un Coeficiente de Correlación

El coeficiente de correlación lineal r de Pearson es una de las medidas más utilizadas en las investigaciones psicológicas, sin embargo, son muy pocos los trabajos que informan IC para este coeficiente, aún cuando las normas editoriales y las recomendaciones de la APA lo plantean como una necesidad (Sánchez-Bruno y Borges del Rosal, 2005). Esto se debe en parte a que el cálculo no es muy directo, y además, a que programas como SPSS no lo incluyen como una opción de análisis. Tomaremos entonces este caso particular para ilustrar como puede usarse bootstrap para estimar empíricamente un IC.

Supongamos que tenemos los datos que se presentan en la Tabla 1 (‘The law school data’ [Efron y Tibshirani, 1993, p.19]) y que también representamos gráficamente en la Figura 1. Se trata de una muestra de 15 estudiantes de Leyes obtenidos de una población de 82 estudiantes de USA. Se dispone de dos medidas: LSAT, que representa puntajes en un test nacional de leyes, y GPA, que es una medida de promedio académico general (average undergraduate Grade-Point Average).

En esta muestra, el coeficiente de correlación arroja un valor de 0,776, indicando una relación fuerte y positiva entre las variables LSAT y GPA. Si se asume la normalidad bivariada de los datos y se aplica la transformación Z de Fisher se puede obtener un IC usando el enfoque estadístico ‘clásico’ (para más detalles, ver Beaulieu-Prevost, 2006). Siguiendo este procedimiento se obtiene que los límites inferior y superior al 95 % son 0,44 y 0,92 respectivamente, lo que se interpreta como un rango de valores plausibles para el coeficiente de correlación en la población. Podemos alternativamente interpretar este resultado en términos del test de hipótesis convencional, diciendo que el valor observado es estadísticamente significativo, puesto que el IC no incluye el valor “0”.

Una primera cuestión práctica con el procedimiento anterior es que la obtención del IC no es tan directa como en otros casos, por ejemplo el IC para un promedio, ya que requiere una serie de transformaciones que se aplican a los valores del coeficiente de correlación. Si bien este cálculo no

Tabla 1. 'Law school data'. Tomados de Efron y Tibshirani, 1993, p.19.

Sujeto	LSAT	GPA
1	576	3.39
2	635	3.3
3	558	2.81
4	578	3.03
5	666	3.44
6	580	3.07
7	555	3.0
8	661	3.43
9	651	3.36
10	605	3.13
11	653	3.12
12	575	2.74
13	545	2.76
14	572	2.88
15	594	2.96

es computacionalmente complejo, se trata de un procedimiento que no está incluido en la mayoría de los programas, incluyendo SPSS. En nuestro caso hemos utilizado ViSta para desarrollar una función de cálculo que se puede integrar al sistema.

No obstante, también es verdad que no hay fórmulas para crear IC para todos los casos e incluso aún si las hay, los programas solo proporcionan un conjunto limitado de opciones. En esta línea, Efron y Tibshirani (1993) señalan como limitación del enfoque tradicional la dificultad intrínseca del trabajo analítico necesario para derivar este tipo de soluciones, lo cual es especialmente difícil cuando se trata de estadísticos más complejos o de distribuciones menos 'tratables' que la distribución normal. Desde luego, otra limitación importante concierne a los supuestos sobre los que se basan estas fórmulas. En el caso anterior, el supuesto de distribución normal bivariada, sobre el cual se apoya el método. Algunos autores han estudiado el comportamiento de la transformación Z de Fisher y han expresado que el investigador tiene que asegurarse que se respetan las condiciones de normalidad, en caso contrario no debería emplearse (Sánchez-Bruno y Borges del Rosal, 2005). Así, cuando no hay "fórmulas" disponibles o los supuestos sobre los que se basan los procedimientos son poco realistas, el

bootstrap se plantea como una alternativa. Entonces, ¿cómo funciona la alternativa del bootstrap en este caso?, ¿cómo puede estimarse por este medio un IC para r ?

IC Bootstrap para el Coeficiente de Correlación

La Figura 2 muestra esquemáticamente cual es el procedimiento general y cómo se aplica en este caso. Se parte de los datos muestrales, representados para nuestro ejemplo mediante el diagrama de dispersión en la parte superior del esquema. El bootstrap trata a estos datos como si fueran la población y procede extrayendo un gran número de muestras con reposición de los mismos datos (por ejemplo, 1000 o 2000 muestras). Los diagramas de dispersión en el segundo nivel del esquema representan ejemplos de muestras con reposición de los datos originales. En este tipo de muestras un individuo puede ser seleccionado varias veces, puesto que luego de cada extracción aleatoria el elemento es reincorporado. Los puntos de mayor volumen en estos diagramas representan individuos que han sido seleccionados más de una vez en la misma muestra.

A continuación, se calcula para cada una de las muestras resultantes el valor del estadístico de interés; en nuestro ejemplo, se obtiene el coeficiente de correlación r de Pearson para cada caso. Tomando las muestras del ejemplo, estos valores son 0.92, 0.73, 0.60 (...) y 0.99. Como resultado de este proceso se dispone de un gran número de valores de r , que constituyen la distribución empírica bootstrap del estadístico y que se usa como una aproximación a la verdadera distribución de probabilidad. En el esquema, el histograma ubicado en la parte inferior muestra dicha distribución. Se observa que la media de estos valores es 0.766, con un error estándar de 0.133.

Por último, en el esquema también se informan los percentiles 2.5 y 97.5 de la distribución, que se utilizan como una aproximación a los límites superior e inferior del IC al 95% para el estadístico en cuestión. Este es en realidad el método más simple de estimación de un IC y se denomina *método percentilar*, ya que se basa justamente en tomar los valores que caen en determinadas posiciones de la distribución. En nuestro caso, tenemos entonces que el 95% de los valores caen entre los valores 0.46 - 0.96; lo que constituye el IC buscado para el coeficiente r .

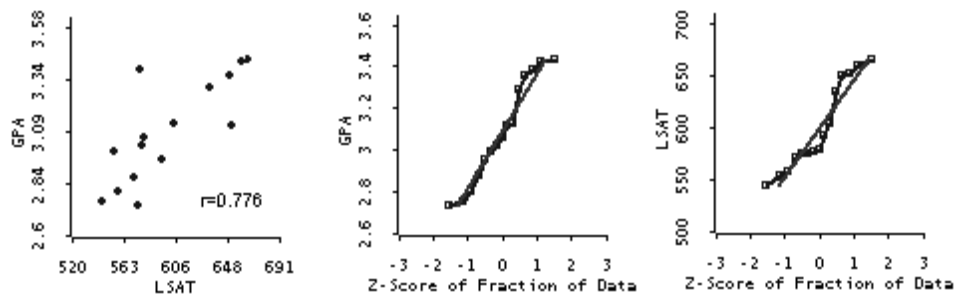


Figura 1. Diagrama de dispersión de los datos del ejemplo y gráficos de probabilidad normal para cada variable.

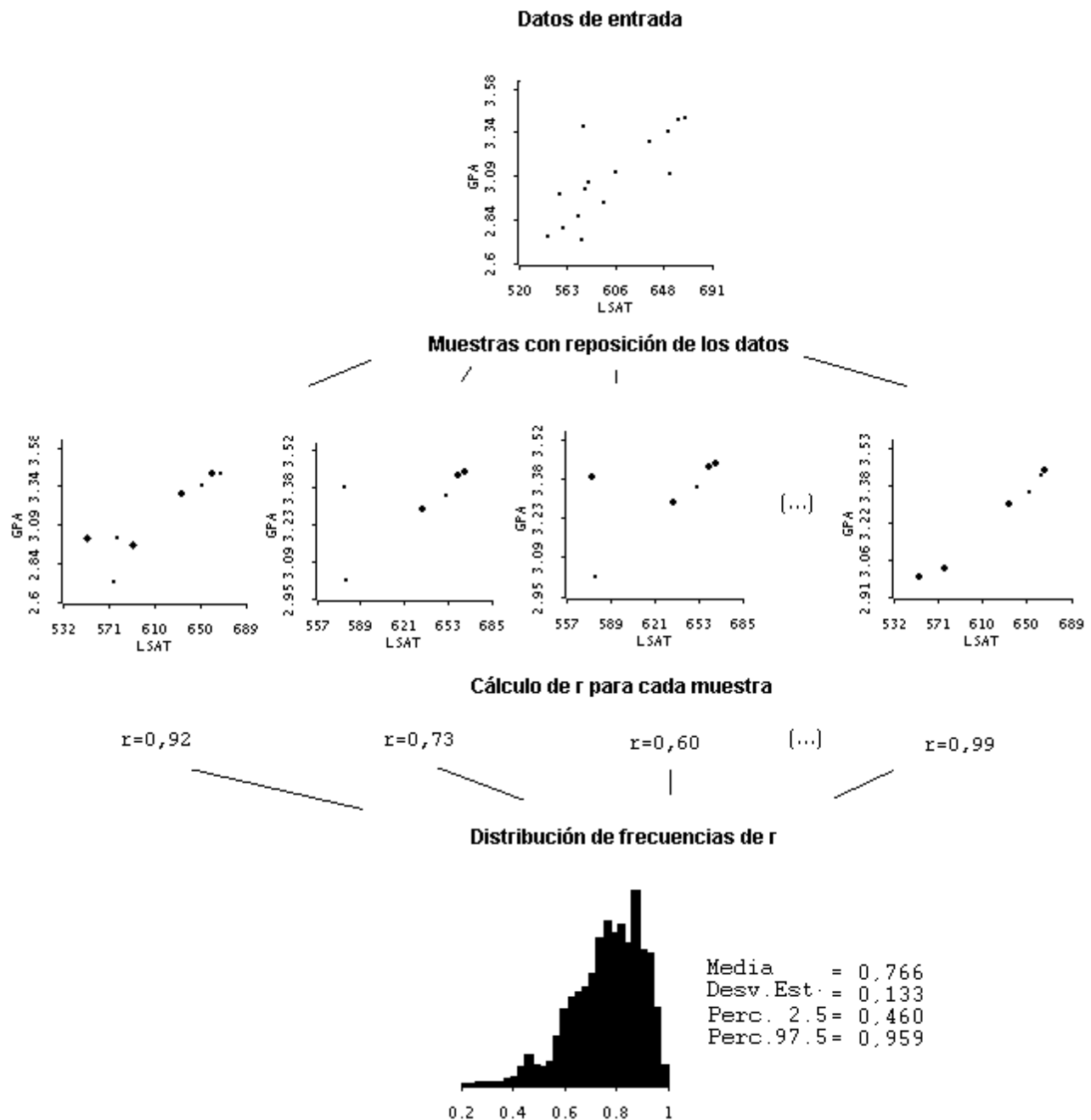


Figura 2. Representación esquemática del procedimiento bootstrap para obtener el IC en el caso del coeficiente de correlación r de Pearson.

Solución del ejemplo mediante Resampling y ViSta

Obviamente, sin un ordenador y un software apropiado resultaría prácticamente imposible realizar este tipo de análisis. A continuación, ilustraremos la resolución del caso mediante los programas *Resampling* y *ViSta*. En la sección final del artículo se comentan y analizan comparativamente otras opciones disponibles de software capaz de realizar un análisis bootstrap.

Resampling

Resampling es un programa gratuito desarrollado por Howell (2001), resulta sencillo de utilizar y proporciona varios métodos de remuestreo, incluyendo bootstrap y técnicas de aleatorización de datos. El software incluye entre sus opciones la posibilidad de crear IC para un coeficiente de correlación. La Figura 3 muestra una imagen de cómo deben definirse los datos del ejemplo para ser analizados con este programa. Luego de definir y guardar los datos, simplemente basta con ejecutar la opción del menú

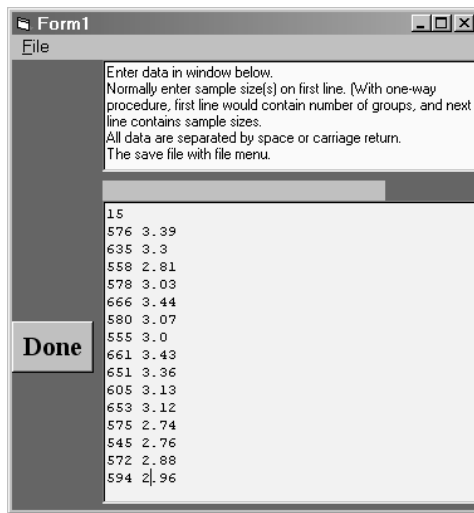


Figura 3. Definición de los datos de entrada en el programa *Resampling*.

denominada “*Bootstrapping Correlation*”. Esta opción abre la ventana que aparece en la Figura 4. Obsérvese que el usuario puede definir los parámetros de la simulación, correr el análisis y visualizar los resultados obtenidos en la misma ventana, lo que simplifica la ejecución del bootstrap y la interpretación de la salida. En nuestro caso, podemos ver que el programa informa el coeficiente de correlación observado (‘Obtained r ’); los límites inferior y superior del IC obtenidos por el método percentilar (‘Lower Con. Limit’ y ‘Upper Conf. Limit’) y el Error Estándar de la distribución (‘St. Error’). Los valores obtenidos en cada caso se corresponden con los que hemos descrito en el apartado anterior.

ViSta “The Visual Statistics System”

El mismo análisis puede realizarse utilizando otro software que se caracteriza por su sencillez y capacidad gráfica: el programa *ViSta*. *ViSta* también es un sistema estadístico gratuito, pero además es abierto, de modo que los usuarios avanzados pueden modificar, extender o adaptar el programa a sus necesidades. Esto es en buena medida posible gracias a *LispStat* (Tierney, 1990), el lenguaje de programación estadística en que se basa *ViSta*. Mediante *LispStat*, pueden programarse e implementarse métodos o modelos estadísticos complejos y, especialmente, desarrollar poderosas herramientas de visualización de datos (Ledesma, Molina, Young y Valero-Mora, 2007; Young, Valero-Mora y Friendly, 2006). En nuestro caso hemos aprovechado estas capacidades para crear un nuevo módulo bootstrap dentro de *ViSta* con la capacidad de obtener un IC bootstrap para el coeficiente de correlación r de forma simple y directa.

La figura 5 muestra una imagen de *ViSta* con la planilla de datos en donde se han definido las variables del ejemplo y el cuadro de diálogo de la opción de análisis

correspondiente al módulo bootstrap que estamos describiendo. El cuadro también permite advertir que puede aplicarse el mismo procedimiento al caso del coeficiente Rho de Spearman. Aunque aquí no nos ocupemos del tema, mencionaremos que el procedimiento bootstrap es esencialmente el mismo, salvo porque para cada muestra generada se calcula el estadístico Rho en lugar de r .

Luego de correr el análisis, y como es propio de todos los métodos de *ViSta*, los resultados pueden presentarse en formato de texto (ver Figura 6) o mediante una visualización múltiple (ver Figura 7). Estas salidas pueden seleccionarse desde la opción ‘Model’ del menú principal de *ViSta* (comandos ‘Report Model’ y ‘Visualize Model’).

La visualización proporciona, simultáneamente y en una misma ventana, varios gráficos específicos para explorar la aplicación bootstrap. Los tres gráficos de la parte superior son representaciones de los datos empíricos; se incluye un diagrama de dispersión y dos histogramas que representan la distribución de frecuencias de cada variable en el análisis. Por otro lado, los tres gráficos en la parte inferior de la visualización ofrecen representaciones del resultado bootstrap. El gráfico ubicado abajo y a la izquierda de la ventana muestra el valor del coeficiente r obtenido a través de las diferentes muestras bootstrap. La línea central en este diagrama representa la media de todos los valores y las líneas por encima y por debajo indican los límites superior e inferior del IC (es decir los valores correspondientes a los percentiles 2,5 y 97,5). El siguiente gráfico, abajo y al centro de la visualización, también es una representación del proceso bootstrap, pero tiene un sentido más ‘evolutivo’.

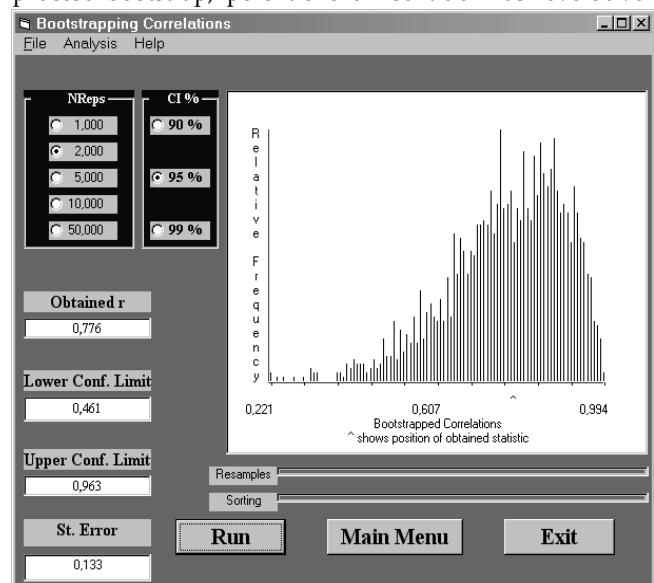


Figura 4. Interfaz gráfica del comando *Bootstrapping Correlation* del programa *Resampling*. Permite al usuario definir los parámetros de la simulación y visualizar los resultados en la misma ventana.

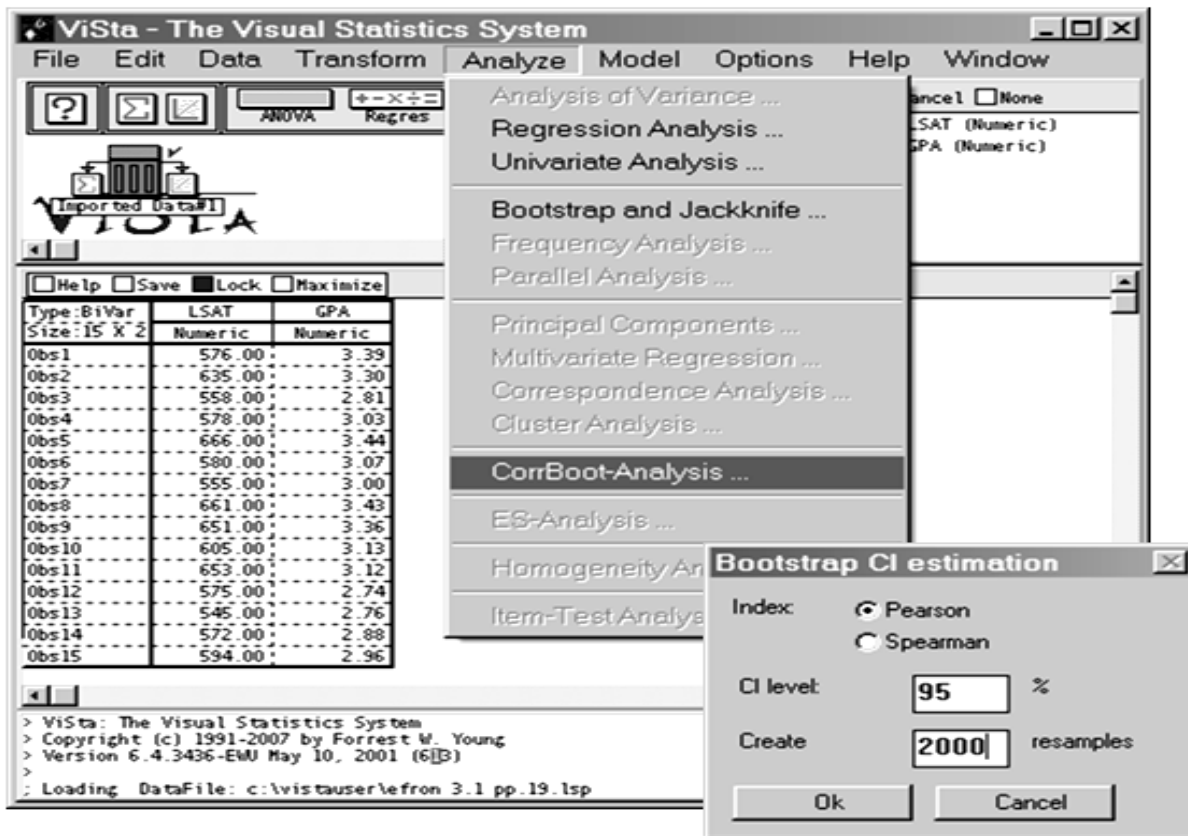


Figura 5. Imagen de ViSta y cuadro de diálogo del módulo bootstrap para el coeficiente de correlación.

Muestra el cambio en la media de la distribución bootstrap en la medida en que aumenta el número de muestras acumuladas en la simulación. Permite inspeccionar la estabilidad de la solución final y alerta sobre la eventual necesidad de aumentar el número de muestras en el bootstrap. En nuestro ejemplo, se observa que el número parece ser suficiente, ya que la estimación final se ha estabilizado en torno a un valor que prácticamente no presenta fluctuaciones. Por último, el gráfico ubicado abajo a la derecha de la visualización muestra el histograma con la distribución de remuestreo del estadístico r . En conjunto, estos gráficos pueden ser de utilidad para el investigador que aplica el método, pero también para el docente que desee introducir a sus estudiantes en la metodología bootstrap. En este sentido, la visualización puede ser un apoyo importante al momento de lograr una comprensión más intuitiva de este tipo de métodos.

Disponibilidad del programa

El sistema ViSta (versión 6.4) pueden obtenerse gratuitamente accediendo a la página URL: www.mdp.edu.ar/psicologia/vista/ El interesado debe instalar ViSta 6.4 en primer lugar, y luego instalar el plug-in ViSta-CorrBoot.exe que anexamos a este trabajo y que añade el módulo bootstrap al programa.

Comentarios Finales

El bootstrap es una técnica de remuestreo que puede resultar de gran interés en Psicología, tanto por sus capacidades como enfoque de análisis de datos como por su potencial en el ámbito de la enseñanza. Como herramienta de análisis estadístico, proporciona un enfoque general que permite afrontar problemas diversos, sin la necesidad de asumir supuestos teóricos que pueden resultar poco realistas, como supuestos relativos a la distribución de los datos. En su lugar, el bootstrap usa la información de la muestra para estimar, mediante remuestreo de los propios datos, propiedades de los estimadores estadísticos y así poder realizar tareas tales como crear un IC o realizar un test de hipótesis. Obviamente, la validez del enfoque depende fundamentalmente de la calidad o capacidad de la muestra para contener o representar apropiadamente las características de la población.

Por su sencillez y generalidad, el enfoque también resulta atractivo para ser implementado en la docencia. Este método parece más fácil de comprender que las técnicas de inferencia clásica, ya que en lugar de explicaciones o formulaciones matemáticas que pueden resultar abstractas para los estudiantes de Psicología, se utiliza una metodología basada principalmente en realizar experiencias


```

Data: Imported Data
Bootstrapped Statistic: R-PEARSON
Number of bootstrap samples: 2000
CI-level: 95
95% CI based on Fisher's Z: (0.4385 0.9220)
CI based on bootstrap percentile method: (0.4503 0.9579)

ObsStat  BootStat  Bias  SE
Obs 1    0.7764    0.7665 -0.0028 0.1338

```

Figura 6. Informe de la salida de resultados del bootstrap en ViSta.

directas con los datos. En lo fundamental, podemos decir que este enfoque solo requiere que el estudiante sea capaz de comprender y aplicar nociones básicas de estadística descriptiva, tales como distribuciones de frecuencias, histogramas, promedios, varianzas y percentiles.

No obstante las posibilidades de este método en el contexto de nuestra disciplina, su utilización en la práctica es casi inexistente, y reservada mayormente a los especialistas en metodología o análisis de datos. Esta situación se ha atribuido a la escasa difusión de las técnicas bootstrap y a cierto conservadurismo por parte de los investigadores, quienes se mostrarían reticentes a incorporar nuevas técnicas para realizar tareas que pueden llevar a cabo mediante los procedimientos estadísticos convencionales. Asimismo, también se ha mencionado que el software de uso popular no incluye opciones bootstrap al usuario. No obstante, como hemos puesto de manifiesto en este trabajo, existe software estadístico alternativo y eficiente para realizar este tipo de aplicaciones. En nuestro caso, hemos ilustrado la resolución del ejemplo mediante Resampling y ViSta, pero existen otros programas de distribución libre y comercial que pueden utilizarse para realizar aplicaciones bootstrap similares o incluso más complejas.

Un programa más general, pero de distribución comercial, es Resampling Stats (Simon y Bruce, 1991), que originalmente funcionaba como un software autónomo y actualmente se ha diseñado como "add-in" del programa Excel. Este programa está bien documentado y se relaciona con los varios trabajos de los autores en el área. En cuanto a los paquetes estadísticos generales, los sistemas Stata, Statistica y SimStat incluyen opciones bootstrap entre sus análisis. Este último es relativamente accesible y proporciona varios métodos de remuestreo integrados a los comandos tradicionales. Asimismo, existen algunos macros elaborados para SPSS, aunque hay que decir que el uso de macros no siempre resulta tan simple para los usuarios familiarizados con el uso de programas con interfaces completamente gráficas. Por otro lado, otra opción la ofrecen

los módulos bootstrap desarrollados en lenguajes como R o LispStat. Estos proporcionan la opción más avanzada y flexible, aunque no resulten tan amigables en términos de interacción de usuario. Una ventaja de los desarrollos en LispStat es que pueden acoplarse a la interfaz gráfica de ViSta. Nosotros hemos aplicado este enfoque para integrar opciones bootstrap al sistema, entre las que se cuenta un módulo general para estimar IC para un rango importante de estadísticos univariados. Este módulo resulta simple de usar y, como es característico de ViSta, incluyen técnicas gráficas para visualizar los resultados de los análisis.

Esperamos que el presente trabajo contribuya a un mayor conocimiento y difusión del bootstrap en la investigación psicológica. Asimismo, confiamos en que el software anexo a este artículo pueda alentar a los docentes y estudiantes a experimentar con el método en el aula, sea para replicar el ejemplo o para experimentar con nuevos

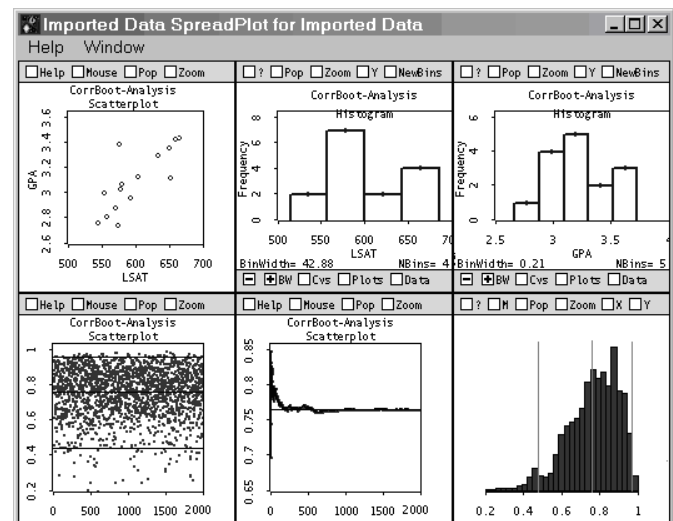


Figura 7. Estructura de gráficos dinámicos para visualizar los resultados del bootstrap en ViSta.

datos. Por último, es importante aclarar que esta introducción no ha contemplado algunos temas importantes que deberían ser profundizados por el lector interesado. Por ejemplo, existen formas alternativas de obtener IC que pueden ser más eficientes en ciertos casos (como el método BCa). Por otro lado, también debemos mencionar que cuando se trata de estructuras de datos más complejas, como medidas correlacionadas o datos basados en muestras estratificadas, puede ser necesario adaptar el procedimiento bootstrap básico, ya que el muestreo con reposición debe realizarse respetando otras exigencias de la información y del problema que se está analizando. Un ejemplo de este tipo, es el bootstrap de bloques móviles, más apropiado cuando se trabaja con datos o modelos de series temporales. Además, también existen algunas formas de modificación del esquema de muestreo destinadas a mejorar la eficiencia del método es sus aspectos computacionales (p.e., el

bootstrap equilibrado o 'balanced bootstrap' que fuerza a que cada observación sea seleccionada un determinado número de veces en todo el proceso bootstrap). Todas estas cuestiones pueden revisarse y profundizarse consultando el manual de Efron y Tibshirani (1993). Alternativamente, Davison y Hinkley (1999) y Manly (1997) proporcionan textos con revisiones de estos temas. Recomendamos también el manual de Simon (1997), un material sumamente claro y orientado a la docencia disponible a través de Internet.

Referencias

- Beaulieu-Prevóst, D. (2006) Confidence Intervals: from test of statistical significance to confidence intervals, range hypotheses and substantial effects. *Tutorials in Quantitative Methods for Psychology*, 2, 11-19
- Braun, W. J. (1995) An Illustration of Bootstrapping Using Video Lottery Terminal Data. *Journal of Statistics Education*, 3, 2. Disponible: www.amstat.org/publications/jse/v3n2/datasets.braun.html
- Davison, A.C., y Hinkley, D.V. (1999). *Bootstrap Methods and their Applications*. Cambridge, Inglaterra: Cambridge University Press.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7, 1-26.
- Efron, B., y Tibshirani, R. (1993) *An Introduction to the Bootstrap*. Chapman and Hall, New York, London.
- Fan, X. (2003) Using commonly available software for bootstrapping in both substantive and measurement analyses. *Educational and Psychological Measurement*, 63, 24-50
- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10, 507-521
- Gil-Flores, J. (2005) Aplicación del método Bootstrap al contraste de hipótesis en la investigación educativa. *Revista de Educación*, 336, 251-265.
- Hall, P. (2003) A Short Prehistory of the Bootstrap. *Statistical Science*, 18, 158-167
- Hesterberg, T. C. (1998). Simulation and Bootstrapping for Teaching Statistics. *Proceedings of the Statistical Education Section*, American Statistical Association, 44-52.
- Howell, D. (2001) Resampling [programa informático]. Disponible: <http://www.uvm.edu/~dhowell/StatPages/Resampling/Resampling.html>
- Ledesma, R. (2004). Sistemas estadísticos de propósitos múltiples: una revisión de programas gratuitos. *Metodología de Encuestas*, 6, 2, 105-117
- Ledesma, R., Molina, G., Young, F., y Valero-Mora, P. (2007) Desarrollo de técnicas de visualización múltiple en el programa ViSta: ejemplo de aplicación al Análisis de Componentes Principales. *Psicothema*, 19, 3, 497-505
- Manly, B.(1997) *Randomization, Bootstrap, and Monte Carlo Methods in Biology* (2nd edition). London: Chapman & Hall.
- Ricketts, C., y Berry, J. (1994). Teaching statistics through resampling. *Teaching Statistics*, 16(2), 41-4.
- Rudner, L. M., y Shafer, M. M. (1992). Resampling: a marriage of computers and statistics. *Practical Assessment, Research & Evaluation*, [Online], 3(5). Disponible: <http://PAREonline.net/getvn.asp?v=3&n=5>
- Sánchez-Bruno, A y Borges del Rosal, A. (2005) Transformación Z de Fisher para la determinación de intervalos de confianza del coeficiente de correlación de Pearson. *Psicothema*, 17, 148-153
- Simon, J. (1997) *Resampling: The New Statistics*. Arlington, VA: Resampling Stats, Inc. Disponible: www.resample.com
- Simon, J. L., y Bruce, P. (1991) Resampling: A Tool for Everyday Statistical Work, *Chance*, 4, 22-32.
- R Developer Core Team (2002). *R1.7.1* [programa informático] [en línea], Disponible: URL: <http://www.r-project.org/>
- Tierney, L. (1990). *Lisp-Stat An Object-Oriented Environment for Statistical Computing and Dynamic Graphics*. NY: John Wiley & Sons.
- Wood, M. (2005) The Role of Simulation Approaches in Statistics. *Journal of Statistics Education*, 13, 3. Disponible: www.amstat.org/publications/jse/v13n3/wood.html
- Young, F. W., Valero-Mora, P. M., y Friendly, M. (2006). *Visual statistic: Seeing data with dynamic interactive graphics*. Hoboken, NJ: John Wiley & Sons.
- Young, F.W. (2003). *ViSta "The Visual Statistics System"*. [programa informático] [en línea] Available: URL: <http://forrest.psych.unc.edu/research/index.html>
- Yu, Ch. H. (2003). Resampling methods: Concepts, Applications, and Justification. *Practical Assessment, Research & Evaluation*, [Online] 8 (19). Disponible: <http://PAREonline.net/getvn.asp?v=8&n=19>

Manuscript received January 17th, 2008

Manuscript accepted March 20th, 2008