

De acuerdo con el principio de *INTERCIENCIA* de adelantar la discusión libre de opiniones e ideas, dentro de un tono de altura, nuestras páginas están abiertas a las personas e instituciones que deseen expresar puntos de vista aunque no necesariamente coincidan con los que se publican en la revista.

/CARTAS AL DIRECTOR/
/LETTERS TO THE EDITOR/
CARTAS AO DIRETOR/

El Editor

ANÁLISIS DE RUTAS EN BIOLOGÍA

Distinguido Director,

La lectura del artículo “Análisis de Rutas en Biología: Estadística para Sistemas Multicausales” por Raúl Ortiz-Pulido (*Interciencia* 25(7): 329-336) me lleva a hacer algunas observaciones. Mis comentarios se refieren a la última sección, subtitulada “Herramientas que mejoran la aplicación de un análisis de rutas” que contiene un conjunto de afirmaciones de carácter estadístico, en mi opinión discutible. Esta sección no tiene relación directa con el resto del artículo que es de gran calidad didáctica.

En la sección comentada se presentan estos conceptos definitorios (p. 335): “Las iteraciones son básicamente un re-muestreo de los datos con que se cuenta”... “A través de la técnica de iteraciones podemos muestrear nuestros datos muchas veces”... “bootstrap lo que hace es un muestreo al azar con reemplazamiento (es decir, que toma de la muestra n número de valores -en este caso cinco, pero pudiendo repetir un valor hasta n veces en esa muestra)”. Tomando como base estos conceptos, se puede afirmar, como punto conceptual básico, que a partir del muestreo -suponemos- aleatorio, sobre los datos originales de una muestra aleatoria, no es posible obtener ningún bit extra adicional de información.

INTERCIENCIA
ANÁLISIS DE RUTAS EN BIOLOGÍA
HONGKIN AND THE MIDDLE ORINOCO SEQUENCE
ANÁLISIS DE ORO EN MUESTRAS GEOLÓGICAS



Otras afirmaciones resaltantes (p. 336):

“Otro de los usos de las iteraciones es para determinar la distribución exacta de nuestros datos y no suponerla, como se hace comúnmente cuando se usa una prueba paramétrica que presume una distribución normal de los datos”.

“Con la técnica bootstrap se resuelve, cuando menos más convincentemente que antes, dos problemas: el bajo

tamaño de la muestra y el supuesto de una distribución que no es (normal, binomial, Poisson, etc.)”.

“Al realizar las iteraciones puede obtenerse n^{10} poblaciones de datos, lo que permite a la vez modelar una distribución más “real” de los datos a través de ellos mismos”.

“Eso es lo que hace esta técnica: construye la distribución de los datos a partir de ellos mismos”.

“Para cada iteración podemos obtener, por ejemplo, una media, que en este caso podemos llamar media de la iteración (p. 335)”.

Nótese que si fuese viable registrar todas las muestras (iteraciones) posibles (n^{10}) de tamaño $n_1 = 10$ tomadas con reposición de una muestra madre original de tamaño $n_0 = n$, lo que al final obtendremos, tal como lo establece la teoría estadística, será una distribución de medias, perfectamente normal con la misma media de la muestra original y con una varianza igual a la décima parte de la varianza de la muestra original.

Max Contasti
Universidad Nacional Abierta
Caracas, Venezuela.
mcontas@reacciun.ve

RESPUESTA de Raúl Ortiz-Pulido

El Profesor Contasti supone que al hacer el re-muestreo de los datos con la técnica bootstrap “no es posible obte-

ner ningún bit extra adicional de información”. Su afirmación puede ser tomada como parcialmente cierta. Es cierta

porque no se creará ningún valor nuevo que pueda ser integrado a la población original de datos. No es cierta porque se

creará nueva información, como la determinación del espacio probabilístico en que se encuentra un coeficiente de ruta o se podrá saber si las rutas planteadas en un modelo *a priori* son comúnmente obtenidas a partir del re-muestreo de los datos. Además, como lo indiqué en el artículo, la técnica de bootstrap aplicada a sistemas multicausales permite la elaboración de modelos nuevos no considerados *a priori*. La creación de esos modelos es información nueva, que sin la utilización de la iteración no existiría. De hecho, en la definición de iteración dada por la Enciclopedia Británica encontramos: "un método iterativo puede ser descrito como uno en el cual el conocimiento de una solución exitosa aproximada es usada para computar una mejor solución". Por lo anterior, concuerdo parcialmente con la observación del Prof. Contasti, pues sí es posible obtener información adicional con la aplicación de bootstrap.

La segunda, observación es: "...si fuese viable registrar todas las muestras

(iteraciones)... lo que al final obtendremos... será una distribución de medias, perfectamente normal con la misma media de la muestra original y con una varianza igual a la décima parte de la varianza de la muestra original". Primero, sí es viable obtener todas las iteraciones. Existen varios programas de cómputo que facilitan la tarea. Para el caso específico del Análisis de Rutas puede ser usado el programa EPA. Segundo, si consideramos una sola variable, ciertamente se obtendrían datos con esa distribución. Tercero, en el caso del Análisis de Rutas, donde no consideramos una sola variable, lo que se obtiene es el equivalente a la media, es decir modelos nuevos no considerados *a priori*, que son modelos de rutas muy cercanos al modelo verdadero. Previo a la introducción de las iteraciones en los Análisis de Rutas, el proceso de descubrimiento de modelos verdaderos era más azaroso. En Análisis de Rutas la viabilidad de la técnica bootstrap para obtener modelos cercanos al verdadero ha sido

demostrada a través de pruebas de Montecarlo. Llamo modelos verdaderos a modelos preconcebidos en los cuales se conoce la relación precisa entre variables. De estos modelos se obtienen series de datos, predichas por el mismo modelo, que son re-muestreadas con la técnica bootstrap y, a través de lo indicado se obtienen los nuevos modelos predichos por los datos. Debido a las limitantes de espacio, remito a los interesados en el tema a Shipley, 1997 (*American Naturalist* 149:1113-1138).

Agradezco al Prof. Contasti sus interesantes observaciones. Concuerdo parcialmente con su primera observación y me permití aclarar la segunda, que tiene aristas previamente no consideradas cuando se le aplica al Análisis de Rutas.

Raúl Ortiz-Pulido
Instituto de Ecología, A.C., México.
pulidos1@infosel.net.mx