II Jornadas de Enseñanza e Investigación Educativa en el campo de las Ciencias Exactas y Naturales *Actas*, II (2): 6-13, 2009. La Plata.

ANÁLISIS DE LA EVOLUCIÓN DE LOS RESULTADOS DE LOS ALUMNOS Y LA METODOLOGÍA BOOTSTRAP PARA DETECCIÓN DE CAMBIOS

CALANDRA, M. V[1]*.; ARGERI, J. G.*[1]

[1] Grupo de Aplicaciones Matemáticas y Estadísticas de la Facultad de Ingeniería (GAMEFI), Área Departamental Ciencias Básicas, Facultad de Ingeniería. Universidad Nacional de La Plata, Calles 1 y 47, La Plata (1900), Prov. Buenos Aires.
mava@mate.unlp.edu.ar

RESUMEN

En este trabajo nos enfocamos en el problema del punto de cambio aplicado al control de calidad del proceso enseñanza-aprendizaje. Para ello se tomo en cuenta la evolución temporal de la proporción de alumnos promocionados, por cuatrimestre, de la asignatura Estadística de la Facultad de Ingeniería de la UNLP, desde el año 2001 al 2008. El objetivo es analizar la posible aparición de cambios en dicha proporción no detectados por las cartas de control convencionales. Se trata de establecer las posibles causas de esos cambios en el marco de las transformaciones ocurridas a partir de la acreditación de las carreras de Ingeniería de la UNLP, usando estas herramientas de estudio. El análisis de punto de cambio es una novedosa herramienta utilizada con el fin de determinar la existencia o no de cambios en procesos de diferente índole. Para su aplicación se emplea un test de hipótesis y la metodología Bootstrap.

Palabras clave: cambios, bootstrap, cartas control, evaluaciones

INTRODUCCIÓN

El objetivo final de nuestro trabajo, es el de tratar de estimar cambios en los parámetros de los indicadores, que se utilizan para la evaluación de los alumnos que se inscriben en asignaturas de un plan de estudios, en nuestro caso dicho parámetro será la proporción de alumnos promocionados por cuatrimestre. Se utilizaron datos obtenidos del desempeño de alumnos correspondientes a asignaturas del área matemática aplicada, en particular de la cátedra Estadística, de la Facultad de Ingeniería de la Universidad Nacional de La Plata (UNLP). Existen distintas metodologías empleadas para detectar cambios sutiles en dichas proporciones, los cuales fueron presentados con anterioridad en otros trabajos. El análisis de esta temática la venimos desarrollando a partir del año 2006 (Calandra y Vericat, 2008) y entre las herramientas utilizadas para encontrar puntos de cambio se propusieron oportunamente las metodologías de sumas acumuladas, mínimos cuadrados y el Criterio de Información de Schwarz (SIC) (Chen y Gupta, 2000; Schwarz, 1978), en esta oportunidad se pretende usar un test de hipótesis basado en el criterio de Cociente de Verosimilitud y una herramienta estadística muy usada actualmente, debido a los avances de la informática, llamada Bootstrap paramétrico. La metodología de punto de cambio es utilizada en control de calidad: en un proceso contínuo, la calidad de un producto es esperable que sea estable. Sin embargo, por diversas razones, el proceso puede a partir de un determinado momento fallar y ó cambiar . Uno quisiera por lo tanto encontrar en que punto o período (punto de cambio) ha ocurrido dicho cambio para poder buscar causas asignables, si es que las hay.

Características del estimador

Los primeros estudios realizados sobre el estimador de punto de cambio datan de 1950. En general en estos casos uno tendería a hacerse las siguientes preguntas: ¿Ha ocurrido un cambio en el proceso? ¿Ha ocurrido más de un cambio? ¿Cuándo han ocurrido dichos cambios? ¿Con qué nivel de significación? y desde luego finalmente ¿Cuáles serán las causas asignables a dichos cambios?. Todas estas preguntas podrían responderse mediante la aplicación del estimador de punto de cambio. Este método puede ser aplicado a todo tipo de datos ordenados en el tiempo. Tradicionalmente la metodología de las cartas de control es usada para detectar cambios en procesos, la diferencia fundamental entre el estimador de punto de cambio y las cartas de control, es que ésta última metodología se puede aplicar a medida que se van recolectando los datos, en cambio para realizar un análisis mediante el estimador de punto de cambio se debe contar con toda la muestra de datos. El estimador de punto de cambio, además, es capaz de encontrar cambios no detectados por las metodologías habituales de cartas de control.

Este trabajo tiene como objetivo mostrar la aplicación de la metodología Bootstrap paramétrica para la detección de posibles puntos de cambio en una secuencia de parámetros asociados a variables aleatorias discretas, en nuestro caso binomiales. Usualmente, la inferencia estadística acerca de punto de cambio tiene dos aspectos que abordaremos: el primero es detectar si hay algún cambio en la secuencia, y el segundo es estimar el número de cambios y su correspondiente localización.

METODOLOGÍA

Para mostrar la aplicación de esta metodología en el análisis de datos correspondientes a la evolución de resultados de alumnos se planteara el problema considerando que tendremos variables aleatorias x_i = número de alumnos promocionados en el período i, y llamaremos p_i a la proporción de alumnos promocionados en dicho período i, y denotaremos con n_i al número total de alumnos inscriptos en cada periodo *i*, siendo *c* el número total de períodos a evaluar. Por lo tanto nuestro caso lo podemos plantear como sigue. Supongamos *c* variables aleatorias con distribución binomial, $x_i \sim b_i(n_i, p_i)$, y que $x_i = m_i$ (número de alumnos promocionados en el cuatrimestre i) para $i = 1, \ldots, c$, donde $x_i = \#$ de éxitos en los n_i ensayos. El objetivo es testear las siguientes hipótesis:

$H_0 : p_1 = p_2 = \ldots = p_c = p$ (desconocido) *vs.* $H_1 : p_1 = \ldots = p_k = p \neq p_{k+1} = \ldots = p_c = p'$

La hipótesis H_0 sería equivalente a que el proceso es estable y la hipótesis H_1 es equivalente a que el proceso es inestable y que hay un punto de cambio en el período k.

En caso de rechazar la hipótesis H_0, se dirá que hay un cambio en el período *k*. Para este test se plantea el estadístico basado en Cociente de Verosimilitud, sea:

$M_k = \sum_{i=1}^{k} m_i$, $N_k = \sum_{i=1}^{k} n_i$ ($k=1, \ldots , c$) y $M \equiv M_c$, $N \equiv N_c$, $M_k^{'} \equiv M - M_k$, $N_k^{'} \equiv N - N_k$.

Entonces, bajo H_0, la función de máxima verosimilitud es $L_0(p) = \prod_{i=1}^{c} \binom{n_i}{m_i} p^{m_i} (1-p)^{n_i - m_i}$

Luego el estimador de máxima verosimilitud de p bajo H_0 resulta ser $\hat{p} = \frac{M}{N}$, (proporción general de promocionados en un proceso estable).

Bajo H_1, la función de máxima verosimilitud es:

$$L_1(p, p') = \prod_{i=1}^{k} \binom{n_i}{m_i} p^{m_i} (1-p)^{n_i - m_i} \prod_{j=k+1}^{c} \binom{n_j}{m_j} (p')^{m_j} (1-p')^{n_j - m_j}$$

y luego los estimadores de máxima verosimilitud de p y p' resultan ser los siguientes:

$$\hat{p} = \frac{M_k}{N_k}, \quad \text{y} \quad \hat{p}' = \frac{M - M_k}{N - N_k} = \frac{M'_k}{N'_k}$$

Luego el logaritmo del cociente de máxima verosimilitud se obtiene como sigue

$$\log \frac{L_0(\hat{p})}{L_1(\hat{p}, \hat{p}')} = \log L_0(\hat{p}) - \log L_1(\hat{p}, \hat{p}')$$

$$= M \log \frac{M}{N} + (N - M) \log\left(\frac{N-M}{N}\right) - M_k \log \frac{M_k}{N_k} - (N_k - M_k) \log\left(\frac{N_k - M_k}{N_k}\right)$$

$$-(M - M_k) \log \frac{M_k^{'}}{N_k^{'}} - [N - M - (N_k - M_k)] \log\left(\frac{N_k^{'} - M_k^{'}}{N_k^{'}}\right)$$

Definiendo: $l(n,m) = m \log m + (n-m) \log(n-m) - n \log n$

Sea $L_k = -2 \log \frac{L_0(\hat{p})}{L_1(\hat{p}, \hat{p}')} = 2\left[l(N_k, M_k) + l(N_k^{'}, M_k^{'}) - l(N, M)\right]$

entonces la posición del punto de cambio es aquel valor de $\hat{k}$ que hace máximo L_k, esto es $L_{\hat{k}} = \max_{1 \leq k \leq c-1} L_k$, a su vez este procedimiento de prueba requiere rechazar H_0 si la probabilidad de obtener un valor del estadístico $L_{\hat{k}}$ tan grande como el observado es poco probable, cuando H_0 es verdadera. Para nuestro caso usaremos Bootstrap paramétrico para estimar dicha probabilidad que es llamada p-valor ó tasa de error del test.

Bootstrap paramétrico

La metodología Bootstrap es una técnica de remuestreo que permite estimar la distribución de un estadístico, en nuestro caso se desea estimar la distribución de probabilidad del estadístico $L_{\hat{K}}$ .En primera instancia se dijo que bajo H_0, el parámetro $p = \hat{p} = \frac{M}{N}$. Luego si la hipótesis H_0 es verdadera, se puede suponer que las variables aleatorias x_i tienen distribución Binomial con parámetros n_i y $\hat{p}$ (dado que todas las proporciones de alumnos promocionados serían iguales para todos los cuatrimestres). Por lo tanto para inferir la distribución del estadístico $L_{\hat{K}}$, bajo H_0, simularemos 999 muestras de dichas distribuciones Binomiales, y calcularemos para cada una de estas simulaciones los estadísticos L_K, que lo denotaremos L_K^*, luego calculamos $L_{\hat{K}*}^* = \max_{1 \le k \le c-1} L_K^*$ y después analizamos si rechazamos o no la hipótesis H_0 estimando el p-valor del test (Davidson y Hinkley, 1997). que es la probabilidad de observar un valor del estadístico $L_{\hat{K}}$ tan extremo como el observado.

El p-valor se estima como sigue: $p - valor = \frac{1 + \#\left(L_{\hat{k}*}^* \ge L_{\hat{k}}\right)}{1000}$, (siendo #(A) es el número de veces que ocurre el suceso A).Luego si el p-valor es un valor muy chico, es decir, menor o igual a 0,05 se rechaza H_0, y por lo tanto hay un cambio es significativo en el período $\hat{k}$.

Este criterio tiene como objetivo la identificación de un solo punto de cambio k. En función de detectar todos los puntos de cambio y su ubicación en un proceso aleatorio, Vostrikova (1981) propuso un método, conocido como el procedimiento de segmentación binario. El mérito que tiene este procedimiento es que permite detectar todos los posibles puntos de cambio. Esta técnica se puede describir como sigue:

Paso 1: en la prueba entre ningún punto de cambio y un punto de cambio, se testea , $H_0 : p_1 = p_2 = \ldots = p_c = p$ vs $H_1 : p_1 = \ldots = p_k = p \neq p_{k+1} = \ldots = p_c = p'$, donde k es la ubicación de un punto de cambio en esta etapa. Si H_0 no es rechazada, entonces se detiene el proceso, y no hay punto de cambio. Caso contrario, existe un punto de cambio y se sigue al Paso 2.

Paso 2: Se testan en forma separada las dos subsecuencias, previa y posterior, al punto de cambio encontrado en el Paso 1 en busca de un punto de cambio.

Paso 3: Se repite el proceso hasta que no haya puntos de cambio en las posteriores subsecuencias.

Paso 4: la colección de puntos de cambio halladas en los Pasos 1-3 puede ser denotada como $\{\hat{k}_1, \hat{k}_2, .., \hat{k}_q\}$, y el número estimado de puntos de cambio encontrado será q.

Datos empleados

Los datos se obtuvieron a partir de las actas de clase de la asignatura Estadística, recolectados desde el 2do Cuatrimestre del año 2001, hasta el 2do. Cuatrimestre del año 2008. Cabe aclarar que las asignaturas tienen una duración cuatrimestral. A partir del proceso de acreditación de las carreras de Ingeniería en el país (CONEAU), se ha establecido un nuevo plan de estudios (en el año 2002) para todas las carreras que se dictan en la Facultad (Ing. Aeronáutica, Electrónica, Eléctrica, Electromecánica, Mecánica, Hidráulica, Civil, Construcciones, Materiales, Química), a excepción de las carreras de Agrimensura e Ingeniería Industrial que acreditaron a partir del año 2007. Dicho cambio fue implementado en ésta asignatura desde el año 2004. En el relevamiento de datos se consideró la cantidad de alumnos inscriptos en la asignatura y la proporción de alumnos que promocionaron la materia (según la reglamentación, aquellos alumnos que obtuvieron un promedio de nota mayor o igual a 6 (seis) aprueban la materia sin necesidad de rendir un examen final). Los datos suministrados para el 2do. Cuatrimestre del año 2001 no siguen estrictamente las reglas indicadas precedentemente, puesto que corresponden a una etapa de transición. En la Tabla 1 se presentan dichos datos.

Posición	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cuatrimestre	2do. 2001	1ero. 2002	2do. 2002	1ero. 2003	2do. 2003	1ero.2004	2do. 2004	1ero. 2005	2do. 2005	1ero. 2006	2do. 2006	1ero. 2007	2do. 2007	1ero. 2008	2do. 2008
Nº Alumnos	391	311	548	330	475	483	424	498	478	336	212	248	326	227	414
Proporción Promoción	0,28	0,49	0,35	0,42	0,42	0,4	0,45	0,38	0,36	0,21	0,61	0,37	0,50	0,33	0,52

Tabla 1. Proporciones de la variable a analizar.

RESULTADOS

En la Tabla 2 se presentan los puntos de cambio detectados según el método Bootstrap paramétrico. Además se presentan los intervalos de segmentación encontrados según Vostrikova y el p-valor según Bootstrap. En el caso que el p-valor de la Tabla 2 de un valor superior a 0,05 se considera que en esa posición no hay cambio significativo, como se ve en las posiciones 3 y 5. Los cambios significativos corresponden a las posiciones 10, 1, 11, 14, 2 y 7.

Bootstrap paramétrico		Alumnos Promocionados	
Puntos de cambio	Intervalo de Estudio	Posición	p-valor
1	[1, 15]	10	0
2	[1, 9]	1	0,00001
3	[11, 15]	11	0,00004
4	[12, 15]	14	0,0011
5	[2, 9]	2	0,0075
6	[3, 9]	3	0,0862
7	[4, 9]	7	0,0315
8	[4, 6]	5	0.669

Tabla 2. Puntos de cambio obtenidos de la aplicación del método.

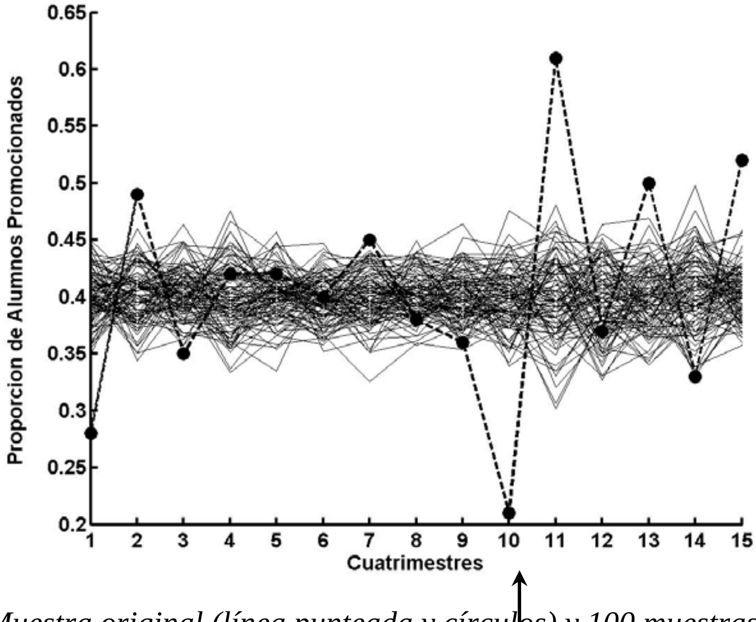

Figura 1. Muestra original (línea punteada y círculos) y 100 muestras Bootstrap paramétricas en la detección del primer punto de cambio (posición 10, ver Tabla 2).

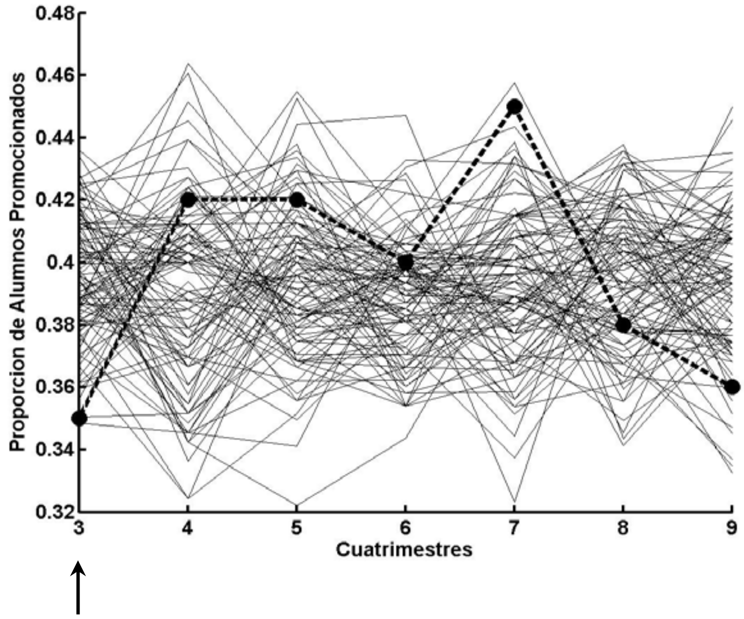

Figura 2 . Muestra original (línea punteada y círculos) y 100 muestras Bootstrap paramétricas en la detección del punto de cambio en la posición 3 (ver Tabla 2).

Se puede observar que si el proceso fuera estable no aparecerían puntos ubicados fuera la banda delimitada por la mayoría de las series temporales simuladas mediante Bootstrap; y a su vez se observarían algunos valores de las series Bootstrap que superen el valor correspondiente al de la muestra original observada (línea punteada). Para la posición 10 (ver Fig.1) se detecta un punto de cambio, se ve claramente que la línea punteada queda muy afuera de la banda delimitada por la mayoría de las simulaciones Bootstrap y a su vez no hay ningún valor de la simulación que lo supere. Por cuestiones de visualización solo se presentan 100 de las 999 muestras Bootstrap. En este caso, se aprecia la posición del primer cambio encontrado, el p-valor da cero (ver Tabla 2), con lo cual el cambio es muy significativo ó evidente, es decir es verdadera la hipótesis H_1. En la Fig.2 se observa la gráfica de las muestras Bootstrap y la muestra original (línea punteada), para el caso en el cual se detecta un posible punto de cambio en la posición 3, correspondiente al 2do. Cuatrimestre de 2002 (ver Tabla 2), habiéndose encontrado un p-valor de 0,0862, que al superar el valor 0,05 indica que no es un cambio significativo, es decir no se puede decir que sea verdadera H_1. Esto se nota en la Fig.2 al encontrarse valores correspondientes a las series Bootstrap que superan al valor original, para cualquier período y la línea punteada queda dentro de las bandas de la simulación.

CONCLUSIONES

La ventaja de la técnica presentada en este trabajo, es que no sólo detecta si hay cambio en el parámetro y su ubicación, sino que también nos da una medida de la evidencia de dicho cambio dada por el p-valor del test, si el p-valor es chico el cambio es muy significativo o evidente. Tradicionalmente para la detección de cambios en procesos que dependen del tiempo se emplean las metodologías de cartas de control, las cuales fundamentan su análisis en gráficos que nos muestran la evolución del parámetro estudiado en función del tiempo. La problemática del empleo de estas técnicas se encuentra en la necesidad de poseer una cierta experiencia para la detección de posibles cambios mediante la simple observación, la cual da buenos resultados siempre y cuando el observador tenga cierta experiencia. La metodología presentada fundamenta la detección de los posibles puntos de cambio en métodos inferenciales basados en test de hipótesis lo cual nos libera de la necesidad de una experiencia previa en el análisis de los resultados desde el punto de vista observacional. Se desprende del presente trabajo una técnica aplicable a otros test paramétricos, en los cuales el estadístico de prueba tenga una distribución desconocida o engorrosa.

En lo que respecta a los cambios encontrados se observa que el más importante, por su nivel de significación está en la posición 10 y corresponde al 1er. Cuatrimestre de 2006, lo cual no parece ser asignable directamente al cambio de planes de estudio, por su ubicación en el tiempo, como así tampoco los cambios de las posiciones 2,11 y 14. El cambio detectado en la posición 1 corresponde al 2do Cuatrimestre de 2002, dicho cambio en este período es asignable a la manera en que se asentaban las notas previo a las transformaciones realizadas. Por otro lado el cambio en la posición 3 (2do Cuatrimestre de 2002) y en la posición 5 (2do Cuatrimestre de 2003) se presentan como no significativos, es decir que no son tenidos en cuenta. Además se observa un cambio de nivel significativo en la posición 7 (2do. Cuatrimestre de 2004). En este año 2004 se comienzan a aplicar las modificaciones en los planes de estudio, aparece la posibilidad por parte de los alumnos de optar por la modalidad promoción o boleta de trabajos prácticos con examen final, este cambio encontrado podría asignarse a esta variación en la evaluación de los alumnos. De los resultados obtenidos no se desprende, aún, un elemento de análisis concreto para definir los efectos que ha tenido la implementación de los nuevos planes en la evolución de los alumnos. Es de esperar que prosiguiendo con este análisis los próximos años se podrá establecer en forma fehaciente los resultados obtenidos de dichas transformaciones, y por lo tanto el establecimiento de nuevos parámetros que nos permitan determinar si se deberán realizar reajuste importantes en esta modificaciones impuestas.

Podemos afirmar, como conclusión final, que la aplicación de esta metodología permite un análisis exhaustivo de la evolución de alumnos en el tiempo, para cualquier nivel de la enseñanza, otorgando una útil herramienta para el análisis de los resultados obtenidos en las distintas asignaturas.

BIBLIOGRAFÍA

Calandra, M. V.; Vericat, F. (2008) "Una aplicación del criterio de información de Schwarz para modelos discretos", *Octavo Congreso Latinoamericano de Sociedades de Estadística, Universidad de la República, Montevideo, Uruguay*, 7-10 de octubre de 2008.

Chen, J.; Gupta, A.K. (2000), *Parametric statistical change point analysis*, Ed. Birkhäuser, Boston, USA.

Davison, A.C.; Hinkley, D.V. (1997). *Bootstrap Methods and Their Applications*. Nueva York: Cambridge University Press.

Schwarz, G. (1978), Estimating the Dimension of a Model, *Annals of Statistics,* 6, 461–464.

Vostrikova, L..J., (1981), "Detecting "disorder" in multidimensional random processes", *Soviet Mathematics Doklady,* 24, 55-59.