

SEGUIMIENTO Y CONTROL DE LAS PREDICCIONES: EL CASO DE LAS EPIDEMIAS DE ASMA EN BARCELONA

Carles Murillo / Marc Sáez

Universidad de Barcelona. Departament d'Econometria, Estadística i Economia Espanyola

Resumen

En el trabajo se discuten procedimientos para la evaluación de los modelos de series temporales y la gestión práctica de una base de datos relativa a un sistema de información sanitaria para la vigilancia epidemiológica y la detección de valores inusuales que podrían corresponder a períodos epidémicos.

La información estadística tratada corresponde a la serie temporal construida con el número de casos atendidos en los servicios de urgencias de varios hospitales de la ciudad con objeto de analizar los fenómenos epidémicos de asma, entre el 1 de enero de 1985 y el 7 de septiembre de 1987.

De los tres métodos analizados, el basado en las correlaciones de los errores un período en adelante (el segundo) es el de mayor especificidad y, además, no exige en su puesta en funcionamiento de la definición previa de día epidémico. Por otra parte, en términos de coste computacional y dificultad analítica, este segundo procedimiento es el más asequible.

Palabras clave: Series temporales. Predicción. Control y seguimiento. Autocorrelación. Asma.

MONITORING AND SURVEILLANCE OF FORECASTING: THE CASE OF ASTHMA EPIDEMIA IN BARCELONA

Summary

In this work we discuss several procedures of evaluation of time series models and analyze practical aspects of the management of a data base relative to a health information system devoted to the epidemiological surveillance and the detection of unusual values that could correspond to epidemic events.

The analyzed statistical information corresponds to the time series of the number of emergency room admissions for asthma in several hospitals of Barcelona city, between the first of January of 1985 and the seventh of september of 1987.

Among the three analyzed methods, the one based on the correlations of the one-step-ahead residuals (the second one) is the most specific and, furthermore, it doesn't require the previous definition of an epidemic day. In addition, this second procedure has less analytical and computational difficulty.

Key words: Time series. Forecasting. Monitoring and surveillance. Autocorrelation. Asthma.

Introducción

El análisis de series temporales se fundamenta en el supuesto que toda colección de observaciones periódicas de una variable es la resultante de cuatro componentes no medibles directamente. Estos componentes son la tendencia, el ciclo, la estacionalidad y un elemento irregular y aleatorio. Los métodos estadísticos han desarrollado procedimientos para conseguir aislar dichos componentes a partir del estudio de las

pautas de comportamiento observadas en la serie. Para ello, es preciso incorporar algunas hipótesis estadísticas acerca de la naturaleza probabilística de las leyes que gobiernan cada componente y de la forma funcional del propio proceso de agregación. Los modelos utilizados son de carácter determinista o estocástico según cual sea la importancia atribuible a aquellas leyes. Desde una perspectiva estadística, los modelos de series temporales, a diferencia de lo que ocurre con los modelos para datos individuales de corte transversal, suponen la existencia de de-

Correspondencia: Carles Murillo. Departament d'Econometria, Estadística i Economia Espanyola. Avgda Diagonal 690. 08034 Barcelona

Este artículo fue recibido el 11 de febrero de 1992 y fue aceptado tras revisión el 15 de enero de 1993.

pendencias entre los sucesivos valores de la variable o, mejor dicho, entre las distribuciones de probabilidad que definen los procesos estocásticos -de parámetros poblacionales desconocidos- que generan los datos. Precisamente, el conocimiento de la forma e intensidad de dichas relaciones es el instrumento que permite establecer las extrapolaciones necesarias para inferir las conductas temporales de la variable más allá de la muestra observada.

La utilización del análisis de series temporales en el ámbito de la salud pública conoce sus orígenes en la aplicación de los modelos de regresión con datos de series históricas con el objetivo de realizar proyecciones a largo plazo¹⁻⁴. Más recientemente se han incorporado⁵ modelos de series temporales basados en la metodología sugerida por Box y Jenkins⁶ tanto para cubrir objetivos de predicción en el corto plazo como para evaluar el impacto de intervenciones y, complementariamente, en la mejora de los métodos para el análisis de la existencia de relaciones de dependencia entre variables⁷. El análisis de series temporales se ha venido aplicando generalmente a los estudios de carácter ecológico⁸ aunque también existe otra vía de investigación en el dominio del comportamiento temporal de los individuos.

Al observador minucioso no se le escapa el hecho de que la validez de los modelos para datos históricos debería, en buena lógica y en consonancia con la propia temporalidad de los mismos, ser revisada y actualizada ante posibles cambios estructurales. El objeto de este trabajo consiste en la propuesta de un procedimiento para la evaluación de los modelos de series temporales y la gestión práctica de una base de datos relativa a un sistema de información sanitaria para la vigilancia epidemiológica y la detección de valores inusuales que podrían corresponder a períodos epidémicos. El instrumento es aplicable sin modificaciones sustanciales a otras finalidades como, por ejemplo, el seguimiento y control de las predicciones de indicadores de actividad y la evaluación de programas.

Material y métodos

La información estadística tratada corresponde a la serie temporal construida con el número de casos atendidos en los servicios de urgencias de varios hospitales de la ciudad con objeto de analizar los fenómenos epidémicos de asma a partir de la propuesta del Grupo de Colaboración⁹ constituido al efecto. La serie histórica es de periodicidad diaria abarcando el tiempo comprendido entre el primer

día del año 1985 y el 7 de septiembre de 1987. Esta última fecha coincide con la modificación de las condiciones en las que se producen las descargas de soja en el puerto de Barcelona y marca el inicio de una nueva época en la que ya no se han detectado nuevos brotes¹⁰.

La metodología seguida ha consistido, en primer lugar, en la construcción de un modelo estadístico para ajustar el comportamiento de la serie temporal observada. El procedimiento utilizado ha sido el de la especificación de un modelo lineal tipo ARIMA. Un modelo ARIMA se representa en la forma:

$$\phi(L) X_t = \mu + \theta(L) \varepsilon_t$$

o, lo que es lo mismo,

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p) X_t = \mu + (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

y, más detalladamente:

$$X_t = \mu + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (I)$$

en donde $\phi(L)$ es un polinomio en L de grado p que incluye todos los parámetros de la denominada parte autorregresiva del modelo, $\theta(L)$ es un polinomio en L de grado q que incluye todos los parámetros de la estructura calificada de media móviles, L es un operador de retardos definido de tal forma que $L^h X_t = X_{t-h}$ ($h = 0, 1, 2, \dots$), μ sirve para determinar el nivel medio de la serie y, ε_t representa un término de error aleatorio que se supone distribuido idéntica e independientemente según una ley normal de media nula y varianza constante

$$\varepsilon_t \text{ iid} - N(0, \sigma_\varepsilon^2) \quad (II)$$

Una vez elegido el modelo que mejor se adapta a la serie temporal que estamos estudiando, se realiza la estimación de los parámetros desconocidos por el método de la máxima verosimilitud no condicional⁶. En este procedimiento, los elementos que corresponden a valores de la variable o del término de error referidos a períodos de tiempo anteriores al primer dato y que intervienen en la función de verosimilitud deducida según (I), son considerados como parámetros desconocidos y estimados también en el mismo proceso inferencial. A partir de estas estimaciones es posible obtener las predicciones óptimas varios períodos en adelante.

Para la modelización de la serie de casos de asma se especificó un modelo lineal multiplicativo de series temporales tipo ARIMA en el que no aparecen

explicitados, pero sí incluidos, los elementos tendenciales y estacionales de la serie. Para nuestro caso la especificación propuesta es la que mejor ajusta la serie observada bajo el criterio del error cuadrático medio mínimo de las predicciones un período en adelante. Como criterio adicional de correcta especificación se utilizó el estadístico de Ljung-Box⁶ que contrasta la hipótesis nula de inexistencia de una conducta temporal sistemática en los residuos del modelo.

La tercera etapa del proceso constituye la detección de valores anómalos mediante un proceso de seguimiento y control del modelo. Para ello se han ensayado los tres procedimientos siguientes:

a) Deducción de valores atípicos ("outliers") en la serie temporal ajustada (M1). Para este caso se utiliza toda la muestra tanto para la estimación de los parámetros como para la detección de los valores anómalos. En la práctica, tanto el momento concreto de ocurrencia del outlier como los parámetros del modelo son desconocidos. De hecho se puede demostrar¹¹ que los estimadores de los parámetros están sesgados por la existencia de outliers. Por ello, se sugiere utilizar un procedimiento iterativo en el que tras estimar inicialmente el modelo, se contrasta para cada observación la hipótesis nula de existencia de outliers. Debe señalarse que se considera que el valor observado en un período determinado es atípico, únicamente si excede significativamente del comportamiento esperado de la serie en ese período.

Los outliers detectados en esa primera etapa deben ser incorporados en el modelo mediante el recurso a variables ficticias o artificiales ("dummy") de manera similar a cómo se introducen en el modelo en la regresión lineal. A continuación se estiman los parámetros del modelo completo y se contrastan de nuevo las hipótesis formuladas. El proceso se repite hasta que no quede ningún outlier por detectar. El procedimiento descrito se ha aplicado utilizando el paquete informático S.C.A.¹².

b) Determinación de un límite de control para la monitorización de los errores de predicción observados (M2). Se trata de comparar los valores ajustados con los reales y los pronósticos estadísticos con los datos reservados a tal efecto y que no fueron utilizados para la especificación ni para la estimación del modelo. El procedimiento que se sugiere está fundamentado en el cálculo de las funciones de autocorrelación temporal de los errores de predicción un período en adelante. Cualquier persistencia en la magnitud y signo de los errores ilustrará el supuesto de un proceso generador de los errores que no sea como el imaginado para los términos de perturbación, es decir como en (II).

El mecanismo para el cálculo de estas correlaciones es el siguiente:

$$I_t = \frac{\text{Cov}_t}{\text{MSE}_t}$$

con

$$\text{Cov}_t = e_t e_{t-1} + \beta \text{Cov}_{t-1}$$

y

$$\text{MSE}_t = e_{t-1}^2 + \beta \text{MSE}_{t-1} \quad (0 \leq \beta \leq 1)$$

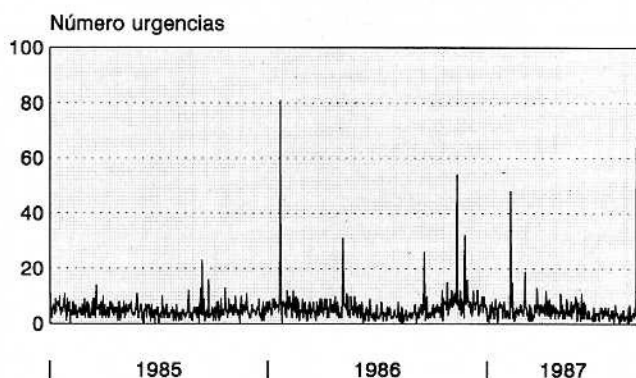
donde, e_t , es el error de predicción un período en adelante.

El procedimiento de cálculo es recursivo y arranca con valores nulos para e_0 y Cov_0 , una estimación del error estándar de la estimación para MSE_0 y fijando $\beta = 0,1$. A partir de estas correlaciones se obtienen los valores anómalos como aquellos días en los que el número de casos registrados se sitúa por encima del valor umbral, al 5% de significación, que es igual a 0,35¹³. La puesta en funcionamiento de este método se ha realizado con ayuda del programa informático MONITOR¹⁴ dispuesto sobre una hoja de cálculo por lo cual su manejo y utilización está abierta a cualquier usuario de este tipo de instrumentos.

c) En este tercer procedimiento (M3) el límite de control se define a partir de la banda superior del intervalo de confianza del predictor de error cuadrático medio mínimo al 5% de significación, una vez se ha eliminado la influencia de los valores atípicos¹⁵. Para eliminar la influencia de los valores atípicos se han sustituidos los datos correspondientes a los días epidémicos por la predicción que resultaría con un modelo ARIMA construido con toda la información disponible hasta la fecha anterior a la epidemia. Este proceso se ha seguido para todos y cada uno de los 10 días definidos como epidémicos. El valor umbral se obtiene del valor máximo entre el valor sobre la banda de confianza para cada período, y el inmediatamente anterior y posterior. Se define un día como anómalo si el número de casos observados está situado por encima del valor umbral.

Estos tres métodos se ponen a prueba en relación con la calificación de día epidémico¹⁶. Un día inusual de asma se define como aquel en que el número de admisiones en los servicios de urgencias por asma es mayor que el esperado con una probabilidad de 0,025 o menor. El número esperado de admisiones en los servicios de urgencias es igual al ajustado, cuando se utiliza una media móvil de orden

Figura 1. Evolución de la serie de urgencias por asma en Barcelona



Número de urgencias por asma (1 enero 1985 - 8 septiembre 1987).

15 en la que se ha supuesto que la serie observada es una variable aleatoria que sigue una distribución de Poisson. Un día epidémico de asma se define como un día inusual con un cluster horario. Éste ocurre cuando en un período de cuatro horas el número observado de admisiones es mayor que el esperado con una probabilidad de 0,05 o menor.

Resultados

La serie ASMA presenta un nivel sostenido con ausencia aparente de variaciones al alza o a la baja en su valor promedio (Fig. 1). Sin embargo, su evolución diaria está salpicada en el período de tiempo considerando por valores anómalamente altos en relación a la media global. Estos valores atípicos de la serie han sido calificados, con arreglo a otros criterios estadísticos y epidemiológicos, como días epidémicos e inusuales^{4,16}. El modelo propuesto es un modelo ARIMA de características similares al presentado en otros trabajos¹⁰. Para toda la muestra el modelo especificado contiene una estructura autorregresiva de primer orden, otra de medias móviles con parámetros de orden 3, 11 y 13, además de un término constante para fijar el nivel de la serie en un valor distinto de cero. La ecuación A de la tabla 1 refleja los resultados de la estimación realizada así como un resumen de los estadísticos de prueba usuales. La serie ASMA es estacionaria en media puesto que no es preciso realizar diferenciaciones para alcanzar correlaciones temporales decrecientes. La estructura de medias móviles refleja la presencia de valores atípicos más que una verdadera pauta estacional de naturaleza

Tabla 1. Estimaciones de los modelos de series temporales para la serie ASMA

Ecuación A. Todo el período muestral 1/I/1985-7/IX/1987
 $(1 - \phi_1 L) ASMA_t = Cte + (1 - \theta_3 L^3) (1 - \theta_{11} L^{11}) (1 - \theta_{13} L^{13}) \epsilon_t$

Parámetro	Coficiente	Error estándar	t Student	p
Cte	4,3195	0,2781	15,53	0,0001
AR1 ϕ_1	0,1913	0,0332	5,76	0,0001
MA3 θ_3	-0,1745	0,0326	-5,36	0,0001
MA11 θ_{11}	-0,0500	0,0367	-1,36	0,1000
MA13 θ_{13}	-0,0813	0,0366	-2,22	0,0250

Varianza residual = 25,9120

LBQ (1) = 0,0 LBQ (4) = 2,1 LBQ (12) = 4,7 LBQ (36) = 18,4 (*)

Ecuación B. 1/I/1985-31/XII/1985

$(1 - \phi_1 L) ASMA_t = Cte + (1 - \theta_5 L^5) (1 - \theta_7 L^7) (1 - \theta_{11} L^{11}) \epsilon_t$

Parámetro	Coficiente	Error estándar	t Student	p
Cte	4,3179	0,2938	14,70	0,0001
AR1 ϕ_1	0,0994	0,0526	1,89	0,0500
MA5 θ_5	0,1280	0,0520	2,46	0,0100
MA7 θ_7	-0,1078	0,0522	-2,07	0,0250
MA11 θ_{11}	-0,1018	0,0525	-1,94	0,0500

Varianza residual = 7,24348

LBQ (1) = 0,1 LBQ (4) = 4,3 LBQ (12) = 11,4 LBQ (36) = 41,9 (*)

Ecuación C. Todo el período muestral 1/I/1985-7/IX/1987. Sustituyendo los valores atípicos

$(1 - \phi_1 L) ASMA_t = Cte + (1 - \theta_3 L^3) (1 - \theta_7 L^7) (1 - \theta_{11} L^{11}) (1 - \theta_{17} L^{17}) (1 - \theta_{24} L^{24}) \epsilon_t$

Parámetro	Coficiente	Error estándar	t Student	p
Cte	4,1993	0,2057	20,41	0,0001
AR1 ϕ_1	0,1350	0,0318	4,25	0,0001
MA3 θ_3	-0,0516	0,0320	-1,61	0,1000
MA7 θ_7	-0,0886	0,0319	-2,78	0,0050
MA11 θ_{11}	-0,0705	0,0319	-2,21	0,0250
MA17 θ_{17}	-0,0790	0,0320	-2,47	0,0100
MA24 θ_{24}	-0,1428	0,0314	-4,54	0,0001

Varianza residual = 8,03493

LBQ (1) = 0,0 LBQ (4) = 1,0 LBQ (12) = 9,0 LBQ (36) = 47,8 (*)

(*) LBQ Estadístico Ljung-Box, contrasta la hipótesis nula de inexistencia de una conducta temporal sistemática en los residuos del modelo.

Fuente: Grupo colaborador ASMA, Barcelona; y elaboración propia.

estable. Todos los coeficientes son estadísticamente significativos al 10% y todos, excepto el factor media móvil de orden 11, lo son también al 5% de significación. La varianza residual del modelo estimado toma el valor 25,9. Teniendo en cuenta que la varianza total de la variable explicada es igual a 354,8 puede afirmarse que se trata de un ajuste débil, aunque

Tabla 2. Número de días detectados como epidémicos e inusuales con distintos procedimientos estadísticos

Primer procedimiento (M1)

	Alerta	No alerta	Total
Día epidémico	9	1	10
Día no epidémico	12	593	605
Total	21	594	615

Segundo procedimiento (M2)

	Alerta	No alerta	Total
Día epidémico	8	2	10
Día no epidémico	0	605	605
Total	8	607	615

Tercer procedimiento (M3)

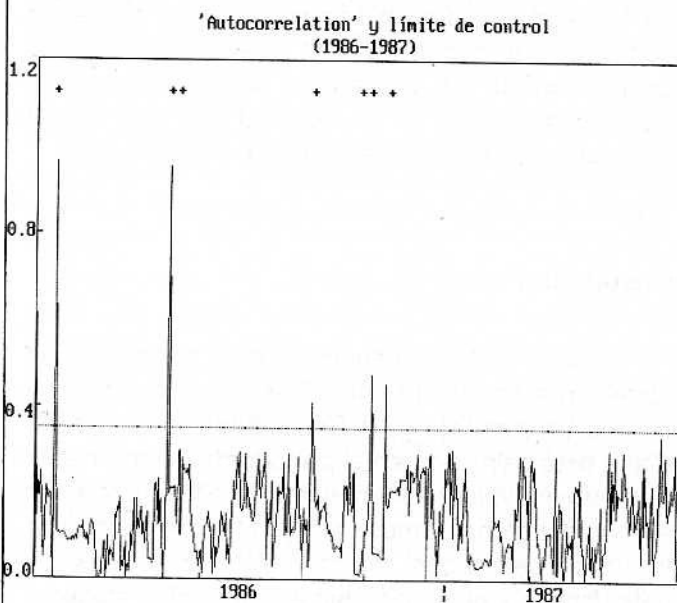
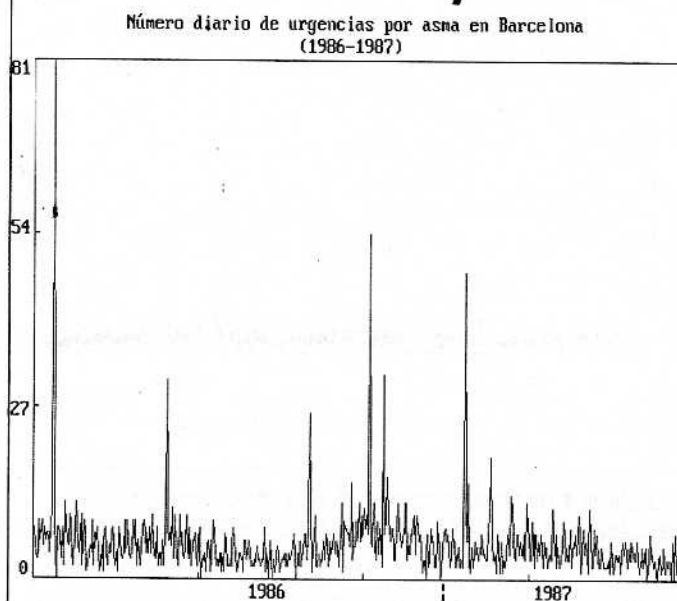
	Alerta	No alerta	Total
Día epidémico	9	1	10
Día no epidémico	13	592	605
Total	22	593	615

Días detectados como epidémicos e inusuales con distintos procedimientos estadísticos

Días epidémicos	Outliers M1	Monitor M2	Watier et al M3
21-I-1986	21-I-1986	21-I-1986	21-I-1986
	1-II-1986		1-II-1986
	1-IV-1986		
6-VI-1986	6-VI-1986	6-VI-1986	6-VI-1986
7-VI-1986	7-VI-1986	7-VI-1986	7-VI-1986
17-IX-1986	17-IX-1986	17-IX-1986	17-IX-1986
	17-X-1986		17-X-1986
	26-X-1986		26-X-1986
			2-XI-1986
11-11-1986	11-XI-1986	11-XI-1986	11-XI-1986
	16-XI-1986		16-XI-1986
24-XI-1986	24-XI-1986	24-XI-1986	24-XI-1986
25-XI-1986	25-XI-1986	25-XI-1986	25-XI-1986
	28-XI-1986		28-XI-1986
	8-II-1987		8-II-1987
	11-II-1987		11-II-1987
	4-IV-1987		4-IV-1987
11-IV-1987			23-III-1987
	23-III-1987		24-III-1987
7-V-1987	7-IV-1987		7-IV-1987
	3-IX-1987		3-IX-1987
4-IX-1987	4-IX-1987	4-IX-1987	4-IX-1987
	7-IX-1987		7-IX-1987
10	21	8	22

estadísticamente significativo, aspecto que es debido precisamente a la presencia de outliers. La validación de los modelos se ha efectuado siguiendo los procedimientos habituales¹⁷ y, fundamentalmente, la significación estadística de los parámetros estimados, la minimización de la varianza residual y el

Figura 2. Evolución de la serie de urgencias por Asma en Barcelona y 'Autocorrelación' (segundo procedimiento)

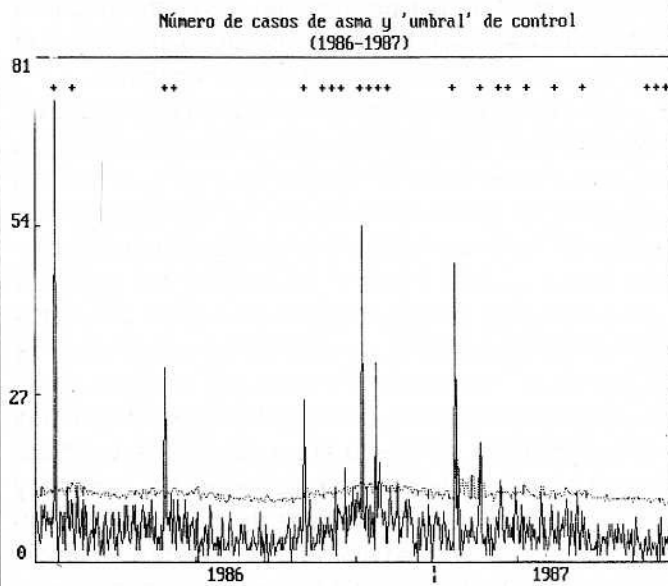


Leyenda: en línea continua (—) se representan los coeficientes de autocorrelación de primer orden calculados según el segundo procedimiento; en línea discontinua (- - -) el límite de control; con '+' se representan las ocho señales de alerta detectadas por dicho procedimiento.

análisis de los residuos, incluido el contraste de Ljung-Box (LBQ)^{6,17} que no permite rechazar la hipótesis nula de ausencia de correlaciones en la serie de los residuos.

Con el primero de los procedimientos enunciados (M1) para la detección de períodos de alerta se obtienen 21 observaciones que calificamos como outliers (Tabla 2). De estos 21 días, solamente nueve se corresponden con días epidémicos. Además, una observación no se califica de outlier y es en

Figura 3. Evolución de la serie de urgencias por asma en Barcelona y umbral de alerta (tercer procedimiento)



Leyenda: en línea continua (—) se representa el número diario de casos de Asma; en línea discontinua (- - -) el umbral de control calculado según el tercer procedimiento; con '+' se representan las veintidós señales de alerta detectadas por dicho procedimiento

realidad uno de los días epidémicos.

Tanto en el caso del método basado en las correlaciones de los errores de predicción un período en adelante (M2), como en el de las predicciones después de sustituir los valores atípicos (M3) la muestra total disponible se ha particionado de modo que su parte final se ha reservado para la fase de evaluación y de predicción ex-post. Así, para la serie diaria del asma con 980 observaciones se utilizaron los 365 datos correspondientes al año 1985 para la especificación y estimación de los parámetros del modelo seleccionado y las 515 restantes, desde el primer día del año 1986 hasta el 7 de septiembre, para la evaluación de los resultados.

El modelo estimado para el método de las autocorrelaciones de los errores de predicción es parecido al que resuelve al tener en consideración toda la muestra disponible, tal como se comprueba a analizar los resultados que aparecen en la ecuación B de la tabla 1. La parte media móvil se ha modificado debido a la estructura de correlaciones que han generado los días epidémicos que corresponden a los años 1986 y 1987. Los resultados que refleja la figura 2, donde se incluyen los gráficos de las autocorrelaciones y el límite de control, muestran los ocho valores que superan este límite y que constituyen las alertas detectadas con este procedimiento (Tabla 2). En ningún caso se señaló como alerta un

día que no fuera epidémico pero se dejaron de detectar dos observaciones de días epidémicos.

Finalmente, se muestran los resultados del tercer método, que calcula las predicciones una vez corregida la serie de la presencia de valores atípicos. La ecuación C de la tabla 1 señala el resultado de la estimación para toda la muestra una vez se sustituyeron esos valores en los días epidémicos, con el método de ajuste descrito en el apartado anterior. Una vez más el modelo resultante tiene un factor autorregresivo, además de componentes multiplicativos de medias móviles, con coeficientes estadísticamente significativos tanto individual como conjuntamente. La varianza residual es de 8,03, es decir que el ajuste es mucho mejor que en la ecuación A puesto que se han corregido los datos de la presencia de valores atípicos. El estadístico LBQ y el análisis de los residuos confirman la conveniencia de mantener el modelo estimado. La figura 3 presenta los valores de las predicciones que sirven para construir la banda que constituye la señal de alerta para 1986 y 1987, y muestra los días en los que se supera el límite (22, tabla 2). Uno de los días epidémicos no fue captado por este tercer método y en 13 días no epidémicos se detectó una alerta, que terminó siendo falsa.

Discusión

Los tres métodos propuestos para obtener señales de alerta se basan en el análisis de modelos de series temporales y en la gestión de las predicciones, dentro o fuera de la muestra. El primero y tercero de ellos utilizan las estimaciones de la muestra completa, con la diferencia de que en el primer caso (M1) los valores atípicos se incorporan en el propio modelo ajustado mediante la presencia de factores en la estructura media móvil, en tanto que en el tercer método (M3) se estima el modelo después de controlar sus efectos. El método de las autocorrelaciones (M2) utiliza una parte de la muestra y predice un período en adelante independientemente de la presencia de valores extraños en la muestra. En términos de coste computacional y dificultad analítica, este segundo procedimiento es más asequible, aunque la generalización de esta afirmación a otros casos debe tomarse con precaución puesto que depende de la tasa de ocurrencia de valores extraños.

El segundo de los métodos analizados (M2) detecta en nuestro caso, y de modo relativamente simple, los momentos del tiempo que se corresponden con aquéllos en los que el proceso aleatorio

generador de los datos pronostica presumiblemente valores destacadamente por debajo de los efectivamente observados. Los valores del instrumento empleado para el seguimiento de la serie que sobrepasen los denominados límites de control de las predicciones permiten identificar señales de alerta de interés en la vigilancia epidemiológica y, más generalmente, en la gestión de modelos predictivos. En general cuando estos valores anómalos se repiten de manera agudizada sin explicación plausible, el modelo utilizado debe ser actualizado. Debe resaltarse que la naturaleza estadística del mecanismo de seguimiento y control exige complementarse con una discusión epidemiológica de la plausibilidad e interpretabilidad de los resultados.

Los días epidémicos no detectados como anómalos por este método corresponden a fechas que se sitúan inmediatamente a continuación de algún día epidémico. Esta misma situación se produce en el primero y tercero de los procedimientos en una sola ocasión. El método basado en las correlaciones de los errores un período en adelante (M2) es el de mayor especificidad y, además, no exige en su puesta en funcionamiento de la definición previa de día epidémico, como ocurre en el tercero de los procedimientos presentados (M3). Aunque es cierto que el método de detección de outliers (M1) en la serie temporal tampoco requiere una definición pre-

via de día epidémico, adolece de una importante complejidad analítica. Pero además, necesita imponer un conjunto de hipótesis previas sobre la naturaleza concreta del presunto outlier. Es evidente que la definición de esas hipótesis no tiene por que ser siempre tan precisa como sería conveniente.

En este trabajo se ha extendido la propuesta de Watier et al¹⁵ relativo a la utilización de señales de alerta basadas en la bondad de las predicciones con modelos de series temporales. La aplicabilidad de estos métodos de control y seguimiento de las predicciones puede resultar de interés en aquellas situaciones en las que mediante modelos con datos de series temporales se realizan predicciones de futuro para indicadores de salud, de usos de servicios sanitarios, de recursos productivos y de precios de los mismos, así como el pronóstico de variaciones en los indicadores de actividad de centros y servicios sanitarios. La validez de los modelos utilizados, la detección de valores anómalos y la vigilancia de la capacidad predictiva de los instrumentos empleados en cualquier proyección pueden seguirse periódicamente con el recurso del método sugerido en el presente estudio. La puesta en funcionamiento del procedimiento de seguimiento y control que aquí se presenta en una aplicación real requiere un soporte informático en forma de hoja de cálculo. Sobre este aspecto se está investigando en la actualidad.

Bibliografía

1. Spierce CC. The mathematical modelling of influential epidemics. *Br Med Bull* 1979; 35: 23-8.
2. H. M. Stationery Office. *Acute Hospital Services in London: A Profile by the London Health Planning Consortium* London: 1979.
3. Lew EA. Mortality and the business cycle: How far can we push an association? *Am J Public Health* 1979; 69: 783.
4. Hawaka M. Projection of cancer incidence: experience and some results in Finland. *World Healt Statistics Quarterly* 1990; 33: 228-40.
5. Kahn HA, Sempos C. *Statistical Methods in Epidemiology* Oxford: Oxford University Press, 1989.
6. Box GEP, Jenkins GM. *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day, 1976.
7. Antó JM, Murillo C. Aplicaciones sanitarias del Análisis de Series Temporales. *Gac Sanit* 1991; 22: 3-4.
8. Catalano RA, Serxner S. Time series designs of potential interest to epidemiologist. *Am J Epidemiol* 1987; 126: 724-31.
9. Antó JM, Sunyer J, Asthma Collaborative Group. A point source asthma outbreak. *Lancet* 1986; 1: 900-3.
10. Sunyer J, Antó JM, Sabriá J, et al. Risk Factors of Soybean Epidemic Asthma: the Role of Smoking and Atopy. *Am Rev Respir Dis* (en prensa).
11. Tiao GC. Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series En Hannan EJ, Krishnaiah PR, Rao MM (eds). *Statistics, Volume 5, Time Series in the Time Domain*. Amsterdam: Nort-Holland, 1985; 85: 118.
12. Scientific Computing Associates. *The SCA Statistical System. Reference Manual for Forecasting and Time Series Analysis*. DeKalb, Illinois: SCA, 1985.
13. Gardner ES. Automatic Monitoring in Forecast Errors. *J Forecasting* 1987; 2: 1-21.
14. Saez M. *Control Automàtic dels Errors de Predicció*. Barcelona: Departament d'Econometria, Estadística i Economia Espanyola, 1990; Document de Treball 9001.
15. Watier L, Richardson S, Hubert B. A Time Series Construction of an Alert Treshold with Application to S. Bovismorbificans in France. *Stat Medicine* 1991; 10: 1493-509.
16. Antó JM, Sunyer J. Epidemiologic Studies of Asthma Epidemics in Barcelona. *Chest* 1990; 98,5: 185s-90s.
17. Murillo C. *Análisis de Series Temporales*. Barcelona: Departament d'Econometria, Estadística i Economia Espanyola, Universitat de Barcelona, 1992; 294-300.

