Boletín Científico CIOH No. 24, ISSN 0120-0542, 60-63 (2006)

ARTÍCULO

Persistencia en series binarias de precipitación

Persistency in binary series of precipitation

Fecha recepción: 2006-08-08 / **Fecha aceptación**: 2006-08-24

Juan Carlos Anduckia
E-mail: janducki@yahoo.com
Centro de Investigaciones Oceanográficas e Hidrográficas CIOH,
Isla Manzanillo, Cartagena de Indias, D. T. y C.

CIOH
www.cioh.org.co

Resumen

Con base en series binarias para la ocurrencia de lluvias para el año 2005 en Cartagena se analiza el parámetro de persistencia de los estados del tiempo meteorológico local y se propone representarlo como un proceso de Markov de primer orden.

Palabras claves: Tiempo meteorológico, modelos de Markov, persistencia.

Abstract

Relying on binary series for precipitation along 2005 at Cartagena, an analysis of local weather "persistence" is done, and a first-order Markov process is thus applied.

Key words: Weather, Markov models, persistency.

Introducción

Es usual hacer uso del denominado "método de la persistencia" en la elaboración de pronósticos meteorológicos del tiempo, método a veces conocido como "pronóstico inercial", y que consiste en pronosticar para el día de mañana la misma ocurrencia del día de hoy. El presente artículo se propone investigar el grado de persistencia que es la base de un pronóstico de este tipo.

Cuando los eventos en cuestión en un proceso probabilístico dado son independientes, la ocurrencia de eventos en el pasado no afecta la predicción del estado siguiente de un sistema: Por tanto, los resultados son siempre los mismos y ocurren con la misma probabilidad. En 1907, el matemático ruso A. Markov inició el estudio de procesos aleatorios para los cuales el resultado anterior sí es relevante para la predicción del estado siguiente. A estos procesos se les llama "procesos de Markov". Su utilidad en el campo de pronóstico está ampliamente soportada por un gran número de aplicaciones en la literatura científica, y se propone aquí describir el más simple de ellos.

Breve descripción de un proceso de Markov

Si X es una variable aleatoria discreta, entonces una cadena de Markov para X es un proceso estocástico con un conjunto de estados y un conjunto de probabilidades $_{j}$ para las transiciones entre estados [1]. A estas probabilidades $\{p_{ij}\}$ se les denomina probabilidades de transición.

El modelo de Markov más simple es el de primer orden, el cual es equivalente a un modelo autorregresivo discreto de orden 1. En aquél, las probabilidades $\{p_{ij}\}$ dependen solamente del estado actual, i, pero no de estados anteriores, es decir, no dependen del camino por el cual se llegó al estado i. Matemáticamente esto se expresa así:

$$P\left\{X_{t+1}\ X_{t,} X_{t\ 1}, X_{t\ 2}\ X_{t\ 3}, \ldots\right\} = P\left\{X_{t+1}\ X_{t}\right\}$$

Esta propiedad refleja nuestra primera intuición acerca del comportamiento del tiempo y del la inercia de los procesos atmosféricos: mañana depende de hoy, pero no de días anteriores. Por supuesto, las probabilidades son diferentes según sea diferente el estado i. La idea, entonces, es poder capturar la correlación serial o la persistencia en una serie de observaciones meteorológicas de ocurrencia de precipitación.

Si la propiedad del estado futuro fuera siempre la misma, independiente del estado actual, entonces la serie temporal contendría valores independientes para las probabilidades condicionales, lo que indicaría que, en caso de existir persistencia, la probabilidad de permanencia en el estado i sería mayor que la probabilidad de alcanzar ese estado a partir de otros estados. Si los $\{p_{ij}\}$no cambiaran en el tiempo, la serie sería estacionaria. Pero por lo general las probabilidades $\{p_{ij}\}$dependen del tiempo, y en cada caso particular deben estimarse a partir de las observaciones existentes.

El propósito del presente artículo es investigar la persistencia en series binarias de ocurrencia de precipitación mediante la aplicación de un modelo de Markov de primer orden.

Metodología

Las series binarias objeto del análisis corresponden a la ocurrencia de precipitación en la ciudad de Cartagena registradas en tablas de contingencia entre los meses de junio y diciembre del año 2005 [2]. Estos datos también están disponibles para otras ciudades, pero no se van a analizar aquí.

Si en estas series los ceros y unos tienden a aparecer en grupos, es porque el proceso subyacente no es binomial o puramente aleatorio, sino que existe una dependencia serial tal que, por ejemplo:

$$p(0|0) > p(0|1)$$
$$p(1|1) > p(1|0)$$

Gráficamente el proceso se vería así:

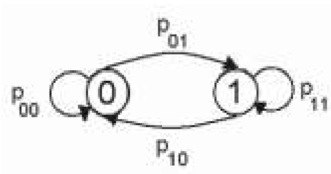

Figura 1. Cadena de Markov de primer orden para eventos binarios (0,1)**.**

Se cumple

$$p_{00} + p_{01} = 1, p_{10} + p_{11} = 1,$$

donde

$$p_{00} = P\{X_{t+1} = 0 | X_t = 0\}$$

$$p_{00} = P\{X_{t+1} = 1 | X_t = 0\}$$

$$p_{00} = P\{X_{t+1} = 0 | X_t = 1\}$$

$$p_{00} = P\{X_{t+1} = 1 | X_t = 1\}$$

Se pueden estimar entonces, sobre la base de los registros mensuales:

$$p_{01} = \text{\# 1's en seguida de 0's / \# total de 0's} = n_{01} / (n_{00} + n_{01})$$

$$p_{11} = \text{\# 1's en seguida de 1's / \# total de 1's} = n_{11} / (n_{10} + n_{11})$$

Cada uno de estos procesos por separado es binomial y en los conteos los extremos no se contabilizan.

En términos de las probabilidades de transición, la autocorrelación de rezago 1 de la serie binaria, llamada también *parámetro de persistencia* [3], es:

$$r_1 = p_{11} - p_{01}.$$

Si r_1 aumenta, la diferencia es mayor y un 1 sigue a un 1 con alta probabilidad, lo que quiere decir que los ceros y unos aparecen en grupos, lo que es la base del pronóstico inercial. Cuando $r_1 = 0$, $p_{11} = p_{01}$, lo que es característico de un proceso binomial con $N = 1$, es decir un proceso de Markov de orden cero.

Debido a la propiedad de Markov, se cumple:

$$r_k = r_1^k.$$

Para aplicar el modelo de Markov de primer orden para el análisis de las series diarias de precipitación construimos las tablas siguientes:

X_t \ X_{t+1}		
	n_{00}	n_{01}
	n_{10}	n_{11}

observadas

X_t \ X_{t+1}		
	e_{00}	e_{01}
	e_{10}	e_{11}

esperadas bajo H_O

Figura 2. Tablas para la puesta a prueba de la hipótesis nula *Ho* (independencia).

Los $\{n_{ij}\}$ se determinan a partir de las series por conteo. Los $\{e_{ij}\}$ se determinan de acuerdo con la suposición de eventos independientes, que es la hipótesis nula H_o:

$$e_{ij} = nP\{X_t = i\}P\{X_{t+1} = j\} = n(n_{i0} + n_{i1})/n(n_{j0} + n_{j1})/n$$

La estadística de prueba es:

$$\chi^2 = \sum_{i=0}^{1}\sum_{j=0}^{1} \frac{(n_{ij} - e_{ij})^2}{e_{ij}},$$

donde el número de grados de libertad es 1.

Resultados

Una prueba de hipótesis de independencia vs. modelo de Markov de primer orden se puede construir con las cantidades observadas vs. esperadas bajo el supuesto de la hipótesis nula (no persistencia o independencia). El análisis se realiza para cada mes del segundo semestre del 2005 y los resultados se muestran en la tabla 1.

La fila r_1 de la tabla 1 es el parámetro de persistencia; la fila indica el valor de la estadística de prueba que compara el modelo de primer orden con un proceso de orden cero, binomial y aleatorio. En los meses de septiembre, octubre y noviembre se puede rechazar la hipótesis nula a favor del modelo de persistencia de primer orden con una

Tabla 1. Análisis de persistencia para series mensuales binarias de ocurrencia de precipitación en Cartagena, año 2005.

	JUN *	JUL	AGO	SEP	OCT	NOV	DIC **
n_{00}	4	13	11	11	11	9	10
n_{01}	5	7	7	4	4	3	1
n_{10}	4	7	7	3	5	4	1
n_{11}	3	3	5	11	10	13	1
e_{00}	9*8/16	20*20/30	18*18/30	15*14/29	15*16/30	12*13/29	11*11/13
e_{01}	9*8/16	20*10/30	18*12/30	15*15/29	15*14/30	12*16/29	11*2/13
e_{10}	7*8/16	10*20/30	12*18/30	14*14/29	15*16/30	17*13/29	2*11/13
e_{11}	7*8/16	10*10/30	12*12/30	14*14/29	15*14/30	17*16/29	2*2/13
p_{00}	4/9	13/20	11/18	11/15	11/15	9/12	10/11
p_{01}	5/9	7/20	7/18	4/15	4/15	3/12	1/11
p_{10}	4/7	7/10	7/12	3/14	5/15	4/17	1/2
p_{11}	3/7	3/10	5/12	11/14	10/15	13/17	1/2
r_1	-0.13	-0.05	0.03	0.52	0.4	0.51	0.41
χ^2	0.25	0.07	0.02	8.52	4.82	7.53	2.17

(* n = 17. ** n = 14. Las demás variables se explican en el texto)

confiabilidad del 95%, y también del 99% excepto en septiembre. Se observa que el parámetro de persistencia alcanza valores elevados en estos meses, incluso superiores al 50%.

En el presente análisis se distinguen bien los meses en los cuales hay un buen grado de confianza para suponer la persistencia (septiembre, octubre y noviembre) sobre la base de un modelo de Markov de primer orden.

Conclusión

El análisis llevado a cabo está basado en observaciones para la estimación de probabilidades con las cuales se pueden realizar pronósticos de ocurrencia de eventos que pueden exhibir persistencia, como las lluvias. Puede ser complementado con un seguimiento mayor para incluir los meses no considerados (en especial la primera estación lluviosa del año, entre el mes de y el mes de julio), y para poder acumular datos para un análisis multianual.

Con la ayuda de una prueba de hipótesis se muestra en qué circunstancias está justificado el uso de la persistencia como un elemento de ayuda para la elaboración de pronósticos. Puesto que hay meses en los que la correlación serial es débil, y la ocurrencia de lluvias parece ser aleatoria, en estos casos el empleo de otras herramientas de predicción es deseable.

Referencias bibliográficas

[1] Laurie Snell J. An introduction to probability. 1era. Ed. NY: Random House; 1988. 478 p.

[2] Anduckia J. Asesoría para el pronóstico meteomarino del CIOH - Informe Final. CIOH; 2005 dic.

[3] Wilks D. S., Statistical methods in the atmospheric sciences, 467 p., International Geophysics Series vol. 59, Academic Press 1995.