

Actas del V Coloquio de Geografía Cuantitativa  
Universidad de Zaragoza  
1992, Zaragoza

## APLICACION DE LAS CADENAS DE MARKOV EN EL ANALISIS DE CONVERSION LLUVIA- ESCORRENTIA

Carmelo CONESA GARCIA  
Francisco ALONSO SARRIA

*Universidad Murcia.  
Campus de La Merced  
30001 MURCIA*

### Introducción. Teoría de las matrices de Markov

Considerando un sistema S que puede existir en un número finito de estados y que puede cambiar de estado sólo en puntos discretos en el tiempo, un proceso de Markov discreto es un proceso estocástico para el cual el sistema S pasa de un estado i en un tiempo t a un estado j en el tiempo t+1 con una probabilidad  $m_{ij}$ .

Asociada a todo proceso estocástico puede definirse una matriz de transición  $M=(m_{ij})$ , independiente del tiempo, que cumple las condiciones de las matrices de Markov, o sea:

$$m_{ij} \geq 0$$

$$\sum m_{ij} = 1$$

$$j=1,2,3,\dots,N$$

siendo el vector probabilístico aquél cuyas componentes  $x_i$  satisfacen:

Puesto que el estado de un sistema S en un tiempo t es una variable aleatoria que puede tomar los valores  $i=1,2,\dots,N$ ; pueden introducirse N funciones de t definidas como

$x_i(t)$ =probabilidad de que el sistema esté en el estado i en el tiempo t

En  $t=0$   $x_i(0)=\delta_{ik}$  siendo k el estado inicial de S y  $\delta_{ik}$  el símbolo delta de Kronecker definido por:

$$\delta_{ik}=0 \text{ si } i \neq j$$

$$\delta_{ik}=1 \text{ si } i = j$$

El paso del sistema de un estado a otro viene dado por

$$x(t+1)=Mx(t)$$

$$t=0,1,2,\dots$$

Puede demostrarse que:

$$\lim_{t \rightarrow \infty} x(t)=y$$

$$t \rightarrow \infty$$

donde y es un vector de probabilidad independiente de  $x(0)$ , y es además un autovector de la matriz de transición M asociado al autovalor 1.

En España MARTIN VIDE (1983), y posteriormente en colaboración con MORENO GARCIA (1985), ha utilizado la cadena de Markov de dos estados para hallar la probabilidad de ocurrencia de un día lluvioso o seco condicionada al día anterior (primer orden) o a los dos días precedentes (segundo orden). En la misma línea se enmarcan los estudios de probabilidad de secuencias secas y lluviosas de PEREZ MANRIQUE *et al.* (1984).

### Metodología

En este trabajo se ha ampliado esta metodología, aumentando el número de estados posibles del sistema a analizar, mediante dos vías:

-Combinación de distintas variables

**-Establecimiento de umbrales**

Se ha aplicado al estudio de la conversión lluvia-escorrentía en la Rambla de Algeciras (Cuenca del río Guadalentín) utilizando tres variables: precipitación en dos estaciones diferentes y caudal en el tramo final de la Rambla de Algeciras. Esta rambla reúne condiciones idóneas para la utilización del análisis markoviano, ya que se trata de un sistema de drenaje de modestas dimensiones (44 Km<sup>2</sup>), de régimen esporádico, dotado de observatorios meteorológicos que disponen de series pluviométricas suficientemente largas (uno en cabecera -H.España- y el otro en el tramo inferior -Librilla-), y de una estación de aforos, con datos válidos en el período escogido (1958-1978).

Debido a la multiplicación de estados, y a la disminución consiguiente de la probabilidad de cada uno de ellos tomados por separado, se ha tenido en cuenta la necesidad de discernir si la probabilidad de ocurrencia de un estado, asociada a un estado previo, puede considerarse significativamente diferente de la probabilidad independiente del tiempo para ese mismo estado.

Para ello se han generado cien series aleatorias de la misma longitud que los datos a analizar y con la misma probabilidad empírica de ocurrencia para cada estado. De estas series se han extraído dos matrices de transición; una con los valores máximos para cada una de las transiciones ( $MAX_{ij}$ ) y otra con los valores mínimos ( $MIN_{ij}$ ). Y se han considerado como no significativamente diferentes aquellas transiciones que cumplen:

$$MIN_{ij} < m_{ij} < MAX_{ij}$$

Finalmente se ha considerado el modelo de Markov de segundo orden para cadenas homogéneas de tiempo discreto y de dos estados, en distintas situaciones.

**Cadena de Markov de 8 estados**

En primer lugar se han definido 8 posibles estados del sistema, dando el valor 0 a los días seco y 1 a los lluviosos o con escorrentía. Estos estados aparecen, junto a su probabilidad empírica de ocurrencia y la calculada por el modelo, en el cuadro 1.

Cuadro 1

Estado	Esc.	Alh.	Lib.	P. Empírica	P. Markov
1	0	0	0	0.817	0.831
2	0	0	1	0.039	0.041
3	0	1	0	0.045	0.042
4	0	1	1	0.070	0.066
5	1	0	0	0.002	0.001
6	1	0	1	0.005	0.003
7	1	1	0	0.006	0.003
8	1	1	1	0.016	0.012

A partir de este cuadro se puede también obtener información acerca de una sola de las variables sumando las probabilidades asociadas a su estado 0 por un lado y a su estado 1 por otro, así por ejemplo la posibilidad total de que se produzca escorrentía en un día concreto es 0.019 y la de que no se produzca 0.98.

En el cuadro 2 se muestra la matriz de transición ligada a este sistema, que debe leerse como la probabilidad de pasar del estado fila al estado columna. Los espacios en blanco indican que la probabilidad calculada no es significativamente diferente de la empírica. El vector y de probabilidades independientes del estado inicial  $x(0)$  al que, como se ha visto anteriormente, tiende asintóticamente el sistema, aparece como columna de probabilidades calculadas en el cuadro anterior.

Cuadro 2. Matriz de transición del sistema para 8 estados

Estado	1	2	3	4	5	6	7	8
1	0.897	0.036	0.021	0.039				0.004
2	0.403	0.090	0.215	0.236			0.009	0.045
3	0.699	0.084	0.075	0.113		0.023		0.006
4	0.485		0.159	0.247	0.002		0.004	0.060
5	0.600	0.200	0.000	0.000				0.200
6	0.217		0.261	0.130	0.043		0.261	0.087
7	0.609	0.130	0.130	0.000			0.870	0.430
8	0.333	0.029	0.118	0.167	0.069	0.020	0.088	0.176

La probabilidad total calculada de día con escorrentía es de 0.019. Este valor puede descomponerse en las probabilidades de escorrentía asociadas a diversos estados en el día anterior. Esta probabilidad para un estado anterior  $i$  es igual a:

$$P_i = \sum y_i * m_{ij}$$

para aquellos estados  $j$  con escorrentía=1

Se obtienen así los siguientes resultados:

Cuadro 3. Probabilidades de día con escorrentía precedido por cuatro estados diferentes.

Estado	Probabilidad	Escorrentía	No escorrentía
Precipitación en Librilla	0.0034	0.371	0.057
Precipitación en Alhama	0.0016	0.130	0.029
Precipitación en ambas	0.0089	0.353	0.066
Sin precipitación	0.0052	0.200	0.006

La columna denominada probabilidad indica la probabilidad de que aparezca una racha formada por un día caracterizado por los estados pluviométricos descritos, seguido de un día con escorrentía.

Las dos últimas columnas de este cuadro muestran la probabilidad de que, tras un día caracterizado por la pluviometría indicada, con y sin escorrentía, siga un día con escorrentía.

El que aparezcan probabilidades relativamente altas de día con escorrentía, precedido de día sin lluvia, se debe a la frecuencia de escorrentías generadas por precipitaciones caídas el mismo día. Es destacable la mayor probabilidad de escorrentía tras precipitación en Librilla que tras precipitación en Alhama.

Por otro lado aunque la probabilidad independiente del estado inicial de los estados 6 y 7 (escorrentía con precipitación en Librilla y en Alhama respectivamente) es la misma, la probabilidad de no escorrentía es mayor ligada a una precipitación en Alhama (estado 3) que a una precipitación en Librilla (estado 2).

Parece por tanto que son las precipitaciones caídas en Librilla más que las caídas en Alhama las que producen precipitación. La explicación de este fenómeno habría que buscarla en las condiciones ambientales, totalmente diferentes, en que se enmarcan ambas estaciones. Alhama se encuentra en Sierra Espuña, en un área con un tapiz vegetal considerable, Librilla está en un área abarrancada sin apenas vegetación. Por tanto parece lógico pensar que los coeficientes de escorrentía en el sector de Librilla son mucho mayores. Mientras en el área de Alhama, una mayor capacidad de infiltración puede dificultar la producción de escorrentía.

Por otro lado se observa una marcada diferencia para cualquier estado pluviométrico, en función de que el día precedente registre o no escorrentía.

### **Cadena de Markov de 18 estados**

En segundo lugar se ha definido un umbral de precipitación para cada estación pluviométrica, de modo que tendremos 3 posibles estados para cada una y 2 para el aforador. Los umbrales se han situado en el valor de la precipitación media diaria (sin contar los días con  $p=0$ ) de cada estación: 6.73 mm para Librilla y 9.77 mm para Alhama- Huerta

España. Por debajo de dicho umbral se asigna el valor 1, y por encima el valor 2.

De este modo el número de posibles estados del sistema aumenta a 18. En el cuadro 4 se expresan dichos estados, junto a su probabilidad empírica de ocurrencia y la calculada por el modelo.

Cuadro 4

Estado	Esc.	Alh.	Lib.	P.empír.	P.Markov
1	0	0	0	0,817	0,831
2	0	0	1	0,035	0,035
3	0	0	2	0,004	0,006
4	0	1	0	0,037	0,034
5	0	1	1	0,042	0,036
6	0	1	2	0,005	0,006
7	0	2	2	0,008	0,008
8	0	2	1	0,009	0,010
9	0	2	2	0,014	0,012
10	1	0	0	0,002	0,001
11	1	0	1	0,002	0,001
12	1	0	2	0,003	0,002
13	1	1	0	0,002	0,001
14	1	1	1	0,002	0,002
15	1	1	2	0,004	0,002
16	1	2	0	0,003	0,002
17	1	2	1	0,001	0,001
18	1	2	2	0,10	0,08

Este enfoque permite añadir algo más de información, e intentar descubrir si el modelo tiene un comportamiento lógico.



Cuadro 5. Matriz de transición del sistema para 18 estados

	1	2	3	4	5	6	7	8	9
1	0.897		0.006	0.019	0.026	0.003	0.002	0.005	0.007
2	0.449	0.095	0.007	0.170	0.124	0.025	0.035	0.025	0.042
3	0.154	0.019		0.077	0.038	0.038	0.192	0.212	0.058
4	0.693	0.071	0.014	0.064	0.060	0.018	0.014	0.018	0.025
5	0.565		0.000	0.117	0.120	0.013	0.026	0.032	
6	0.283	0.022		0.196		0.065	0.087	0.109	0.087
7	0.730	0.063	0.016	0.063	0.016	0.032	0.000		0.032
8	0.438	0.075		0.200	0.112	0.025	0.025	0.038	0.025
9	0.366	0.050		0.069	0.168	0.010	0.030	0.059	0.089
10	0.600	0.100	0.100	0.000			0.000		
11	0.500	0.000	0.000						
12	0.067				0.067	0.067	0.333		
13	0.714			0.143	0.000		0.143		
14	0.667		0.067	0.133		0.000	0.000	0.000	0.000
15	0.308			0.154				0.077	0.000
16	0.563	0.063	0.125	0.000	0.000	0.000	0.063	0.000	
17	0.333	0.000		0.222	0.111		0.111		
18	0.262	0.015	0.015		0.169	0.031	0.031	0.000	0.000

	10	11	12	13	14	15	16	17	18
1	0.000				0.001	0.001	0.000	0.000	0.003
2					0.007	0.004	0.007	0.000	0.011
3						0.019	0.019	0.038	0.015
4		0.004			0.000	0.000			
5					0.000		0.000		0.032
6							0.043	0.022	0.043
7			0.032						0.016
8									0.063
9	0.010			0.000	0.020	0.020	0.000		0.099
10						0.100			0.100
11							0.250		0.125
12						0.067	0.267		0.000
13						0.000			
14					0.067				
15				0.154			0.154		0.077
16				0.125	0.063				
17		0.111				0.111			
18		0.115		0.046	0.046	0.015	0.031	0.062	0.108

La probabilidad de día con escorrentía es ahora de 0.02. Si, como se ha hecho antes, se descompone esta probabilidad en probabilidades dependientes del estado de las precipitaciones del sistema en el día anterior, se obtienen entonces los resultados del siguiente cuadro.

Cuadro 6.

Librilla	Alhama	Probabilidad	Esc	No esc.
0	0	0.00601	0.200	0.007
0	1	0.0009	0.000	0.026
0	2	0.00076	0.188	0.222
1	0	0.00151	0.500	0.029
1	1	0.00135	0.067	0.032
1	2	0.00085	0.222	0.063
2	0	0.00192	0.334	0.210
2	1	0.00141	0.385	0.108
2	2	0.00523	0.431	0.149

De nuevo parece ser que la precipitación en Librilla es más determinante sobre la existencia o no de escorrentía en la cuenca.

En los casos en que la precipitación en Librilla es cero, no parece influir la existencia o no de escorrentía en el día precedente.

### Bondad de la cadena de Markov de segundo orden

El modelo de la cadena de Markov de segundo orden permite un mejor ajuste respecto a las probabilidades observadas.

Las cadenas de 8 y de 18 estados empleadas ofrecen unos resultados aceptables para el conjunto de las series analizadas; sin embargo, para un mismo número de estados, la cadena de segundo orden presenta mayor bondad que las de primero. En varios casos se ha reducido el número de estados a dos (cuadro 7 y 8) a fin de realizar un primer ensayo, y se ha logrado verificar dicho supuesto.

En la cadena de Markov de dos estados y de segundo orden se considera que la probabilidad de que un día presente uno de los dos estados, sea, en nuestro caso, la aparición de un día con escorrentía, depende de lo ocurrido el día precedente -como la cadena de Markov de primer orden- y, además, de lo ocurrido el día anterior a éste. La expresión probabilística de la cadena es:

$$q_n = p_{011} \cdot p_{111}^{n-2} \cdot p_{110}$$

$$q_1 = p_{010}$$

siendo  $p_{011}$  la probabilidad empírica de ocurrencia de un día con escorrentía tras uno también con escorrentía y el anterior a éste con lecho seco;  $p_{111}$  la correspondiente probabilidad de día con escorrentía después de otros dos días de avenidas;  $p_{110}$ , la de uno seco tras dos días con escorrentía; y  $p_{010}$ , la de un día de avenida entre dos días con lecho seco.

Utilizando la serie de aforos diarios del presente análisis,

$$q_n = 0.3832 \cdot 0.1667^{n-2} \cdot 0.833$$

$$q_1 = 0.6168$$

Cuadro 7. Probabilidades observadas y calculadas mediante las cadenas de Markov de primero y segundo orden para diferentes  $n$  (número de días de escorrentia).

$n$	Mv1 <sup>o</sup>	$\Delta$	Mv2 <sup>o</sup>	$\Delta$	Obs.
1	0.697	0.080	0.617	0.000	0.617
2	0.211	-0.107	0.319	0.001	0.318
3	0.064	-0.001	0.053	-0.012	0.065
4	0.019	0.019	0.009	0.009	0.000

$\Delta$  = Diferencia respecto a la probabilidad observada.

La cadena de segundo orden se ha aplicado también para calcular la probabilidad de generación de escorrentía, dependiendo del estado de lluvia generalizada o no en toda la cuenca, durante el día o los dos días anteriores al inicio del flujo (cuadro 8), dando como resultado:

$$q_n = 0.4321 - 0.2708^{n-2} - 0.7292$$

$$q_1 = 0.5679$$

Cuadro 8. Probabilidades de escorrentía observadas y calculadas mediante las cadenas de Markov de primero y segundo orden para diferentes  $n$  (número de días de lluvia generalizada en la cuenca).

$n$	Mv1 <sup>o</sup>	$\Delta$	Mv2 <sup>o</sup>	$\Delta$	Obs.
1	0.636	0.068	0.568	0.000	0.568
2	0.231	-0.065	0.315	0.019	0.296
3	0.084	-0.014	0.085	-0.013	0.098
4	0.031	0.021	0.023	0.011	0.012

$\Delta$  = Diferencia respecto a la probabilidad observada.

- **Aplicación a umbrales.** Para analizar la probabilidad markoviana de primero y segundo orden de que se genere escorrentía superficial a partir de determinados umbrales, se ha dado a  $n$  pasos una dimensión espacial y temporal. Como en los casos anteriores, se verifica entonces para  $n \geq 2$

$$p_{ij} = \sum_{k=1}^n p_{ik}^{(h)} p_{kj}^{(n-h)}$$

$$1 \leq h < n$$

siendo aquí  $n$  la ocurrencia de un mismo intervalo de precipitación, en cada una de las estaciones consideradas, durante los días que presenten continuidad pluviométrica. Este análisis incluye sólo las secuencias pluviométricas con escorrentía, utilizando unos umbrales previamente establecidos por histogramas de frecuencias e intensidades de precipitación: valores de precipitación diaria en Librilla de 10-30 mm, 30-50 mm ó  $> 50$  mm, precedida de la ocurrencia de intervalos análogos en la estación de H.Espuña ese mismo día y, en ambas estaciones o sólo en una de ellas, durante los días anteriores (cuadro 9). El mismo modelo se aplica para estimar la probabilidad de escorrentía a partir de la precipitación caída en un día en H.Espuña, tras una determinada longitud de  $n$ .

Cuadro 9. Probabilidades de generación de escorrentía estimadas por las cadenas de Markov (1º y 2º orden) y probabilidades observadas, para diferentes umbrales pluviométricos y valores de n.

(a) Base recuento pluviométrico: LIBRILLA												
11-30 mm			31-50 mm				> 50 mm					
n	Mv 1º	Mv 2º	Obs.	n	Mv 1º	Mv 2º	Obs.	n	Mv 1º	Mv 2º	Obs.	
1	0.414	0.207	0.203	1	0.556	0.300	0.300	1	0.750	0.667	0.667	
2	0.243	0.445	0.407	2	0.247	0.656	0.600	2	0.188	0.333	0.333	
3	0.142	0.195	0.170	3	0.110	0.041	0.100	3	0.047	0.000	0.000	
4	0.083	0.086	0.186	4	0.048	0.003	0.000	4	0.012	0.000	0.000	
5	0.049	0.037	0.034	5	0.022	0.000	0.000	5	0.003	0.000	0.000	

  

(b) Base recuento pluviométrico: H.ESPUNA												
11-30 mm			31-50 mm				> 50 mm					
n	Mv 1º	Mv 2º	Obs.	n	Mv 1º	Mv 2º	Obs.	n	Mv 1º	Mv 2º	Obs.	
1	0.397	0.146	0.146	1	0.556	0.300	0.300	1	0.875	0.857	0.857	
2	0.239	0.480	0.458	2	0.247	0.612	0.600	2	0.109	0.143	0.143	
3	0.144	0.210	0.167	3	0.110	0.077	0.100	3	0.014	0.000	0.000	
4	0.087	0.092	0.167	4	0.048	0.009	0.000	4	0.002	0.000	0.000	
5	0.053	0.040	0.042	2	0.022	0.001	0.000	2	0.000	0.000	0.000	



En todos los casos se obtiene un excelente ajuste de la cadena de segundo orden, mucho mejor que el que proporciona la de primero. La interpretación de los datos pone de manifiesto:

(1) Que para el umbral 10-30 mm la mayor probabilidad de generación de escorrentía (0.445) se produce cuando dicho umbral se alcanza durante un mismo día en las dos estaciones del análisis ( $n=2$ ).

(2) Que para el intervalo 30-50 mm se da la misma circunstancia, aumentando la probabilidad de escorrentía, para  $n=2$ , a un 0.656, si la secuencia de  $n$  valores pluviométricos diarios comienza en Librilla, y a un 0.612, si lo hace en H.España.

(3) Que la probabilidad de generación de escorrentía con la lluvia caída en un solo día y en una única estación ( $n=1$ ) aumenta, como es lógico, con el incremento del umbral de precipitación, de manera que en el umbral superior a 50 mm se obtienen probabilidades para  $n=1$  muy altas (0.667-0.857), mientras que en el umbral inferior (10-30 mm)  $q_1$  apenas vale 0.2.

## Conclusiones

De la distribución del sistema  $p^{(n)}$  y la matriz de transición  $P$  de Markov obtenidas para 8 estados se infiere una probabilidad total de día con escorrentía de 0.019, de la que aproximadamente el 37 % corresponde a precipitación en Librilla, el 13 % a precipitación en Alhama H.España y el 35 % a lluvia en ambas estaciones.

Estos datos están en concordancia con las características medioambientales (litología, suelos, pendientes, cobertera vegetal...) de las áreas en que se ubican los observatorios, si bien hay que tener presente el posible efecto que supone la mayor proximidad de la estación pluviométrica de Librilla respecto al aforo de Algeciras.

Un rasgo también bastante común en el Sureste Peninsular, puesto de relieve por el modelo de la cadena de Markov de segundo orden, para el que se ha obtenido un buen ajuste, es la escasa duración de la escorrentía superficial ( $q_1 = 0.617$ ). En torno a un 62 % de las escorrentías producidas durante el período de análisis duran teóricamente un día (presumiblemente su duración real sea inferior) y sólo el 5 % alcanza una duración igual o superior a tres días. Asimismo, se ha detectado una

importante relación entre el inicio de la escorrentía en el punto de aforo y la superación del umbral de los 30 mm/día de lluvia generalizada en la cuenca o de los 50 mm/día en una única estación, especialmente en la de Librilla.

---

**Bibliografía**

MARTIN VIDE, J. (1983): "La aceptación del modelo estocástico de la cadena de Markov homogénea de tiempo discreto y de dos estados en los cálculos de la probabilidad de la precipitación diaria". *VIII Coloquio de Geógrafos Españoles. Comunicaciones*. A.G.E., Secc. Geografía, Univ. Barcelona, pp. 24-31.

MARTIN VIDE, J. y MORENO GARCIA, M<sup>a</sup>.C. (1985): "El estudio de las sequías mediante el análisis probabilístico de las secuencias secas. El caso de Almería". *IX Coloquio de Geografía. Ponencias, I*, A.G.E.

PEREZ MANRIQUE, C. *et al.* (1984): "Estudio de rachas secas y lluviosas en Gijón y San Sebastián". *Revista de Geofísica*, 40, C.S.I.C., Madrid, pp. 73-80.