

SOBRE CADENAS DE MARKOV Y SU POTENCIAL EN LAS APLICACIONES

On Markov chains and their potential in applications

EPISTEMUS
ISSN: 2007-8196 (electrónico)
ISSN: 2007-4530 (impresa)

Luz del Carmen Rosas Rosas¹
Jovan Israel Segundo Rosas²

Recibido: 12 de septiembre de 2016,
Aceptado: 30 de noviembre del 2016

Autor de Correspondencia:
Luz del Carmen Rosas Rosas
Correo: lcrosas@mat.uson.mx

Resumen

En este artículo se presenta una muestra sobre la relevante herramienta proporcionada por la teoría de probabilidad, específicamente a través de las cadenas de Markov. Considerando la importancia de dicho concepto matemático, el propósito fundamental de este trabajo es el de ofrecer un breve panorama en el que se contemplan, tanto los elementos generales que caracterizan a tales procesos, concretamente en el caso de tiempo discreto, así como la inclusión de algunos ejemplos ilustrativos de su amplia gama de aplicaciones.

Palabras clave: Cadena de Markov, Espacio de estados, Probabilidad de Transición, Matriz de transición, Distribución inicial.

Abstract

In this paper we present a sample about the relevance of the tool provided by probability theory, specifically through the Markov chains. According to the importance of this mathematical concept, the main purpose of this work is to offer a brief overview of the general elements that characterize such processes in the specific case of discrete time, as well as include some examples of their wide range of applications.

Keywords: Markov chain, States space, Transition probability, Transition matrix, Initial distribution.

¹ Departamento de Matemáticas, Universidad de Sonora, México. Correo: lcrosas@mat.uson.mx

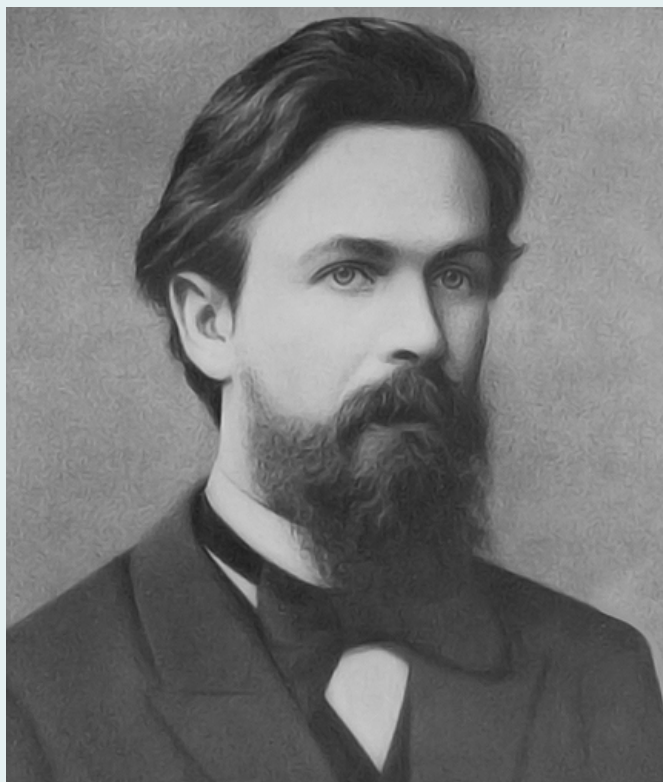
² Departamento de Matemáticas, Universidad de Sonora, México. Correo: jovan.kakashi@gmail.com



INTRODUCCIÓN

En el análisis de diversos fenómenos reales que de alguna forma se ven impactados por la incertidumbre, con frecuencia se requiere trabajar un esquema de variables aleatorias (v.a.'s) que dependen del tiempo, como podría ser por ejemplo:

- ▶ la demanda semanal de una clase específica de artículo,
- ▶ el número de pólizas reclamadas anualmente ante cierta aseguradora,
- ▶ la concentración diaria de un contaminante dado en determinado acuífero,
- ▶ el número de denuncias mensuales por violencia de género registradas en cierta región,
- ▶ el estado del tiempo en determinada región diariamente,
- ▶ el número de niñas que reciben la dosis correspondiente de la vacuna de VPH por año, entre otros. De hecho, en nuestro caso particular las cadenas de Markov proporcionan modelos estocásticos de gran utilidad para el análisis y la predicción del comportamiento a corto o a largo plazo, para cierta clase de sistemas que evolucionan en el tiempo de forma no determinista, en torno por supuesto a un conjunto de estados fijo y a lo más numerable.



Andrei A. Markov 1856-1922

Con fundamento en lo publicado por J. J. O'Connor y E. F. Robertson [1], cabe destacar que alrededor de 1905

el matemático de origen ruso Andrei Andreyevich Markov (1856-1922) desarrolló la teoría formal de cadenas de Markov desde un punto de vista matemático, lo cual dio pie a una nueva rama de la probabilidad, originándose así la teoría de procesos estocásticos, tal como lo comenta L. Rincón [2]. Y de hecho, las cadenas de Markov pueden aplicarse al estudio y resolución de problemas asociados a fenómenos aleatorios reales que surgen de áreas tan diversas como Economía, Meteorología, Epidemiología, Genética e Ingenierías, etc. Por otra parte, en un terreno menos informal podemos decir que las cadenas de Markov forman una clase de procesos estocásticos cuyo espacio de estados es numerable, y que como todo proceso de Markov, poseen la característica general de que la probabilidad del futuro, dado el presente y el pasado, depende únicamente de la información su el estado presente.

DESCRIPCIÓN

Preliminares

Básicamente, un proceso estocástico (PE) consiste de una familia $\{X_t: t \in T\}$ de v.a.'s parametrizadas por el factor tiempo y definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) fijo, tal como lo presentan Z. Brzezniak y T. Zastawniak [3], de manera que toman valores en un conjunto común S llamado *espacio de estados* del PE, tal como lo define O. Hernández-Lerma [4]. A fin de ilustrar este útil concepto podríamos citar por ejemplo, la clase de sistemas dinámicos en que intervienen elementos aleatorios cuyo propósito consiste en representar cierto tipo de incertidumbre, misma que influye de forma determinante en el comportamiento del sistema en cuestión; o bien, algunos problemas relacionados con la toma secuencial de decisiones en diversas etapas, las cuales dependen de que ocurran o no eventos aleatorios específicos durante el desarrollo del proceso en el tiempo, situación que hace evidente su relación con el factor incertidumbre. Tal es el caso por ejemplo, en: predicciones de contextos económicos, patrones de compra por parte de la clientela, estimación de indicadores, pronósticos de votaciones, administración de inventarios, proyecciones demográficas, solo por mencionar algunos escenarios que admiten dicha clase de ambiente.

De hecho, para el estudio de ese tipo de procesos es conveniente clasificarlos, ya sea dependiendo de su espacio de estados, o bien, de su conjunto de parámetros (generalmente el tiempo en que se observa), etc. En particular, aquí consideramos cadenas de Markov a tiempo discreto, como es el caso que incluye O. Hernández-Lerma [5], y que no es más que una clase de PE's de la forma

$$\mathbf{X} := \{X_n: n \in \mathbb{N}_0\}, \quad \mathbb{N}_0 := \{0, 1, \dots\}, \quad (1)$$

con espacio de estados discreto dado por $S := \{x_0, x_1, \dots\}$ y, tal que se caracteriza por registrar una *memoria a corto plazo* llamada propiedad de Markov, esto es, para cada colección de estados $x_0, \dots, x_{n-1}, x, y \in S$ se verifica la

siguiente identidad entre probabilidades condicionales

$$P(X_{n+1}=y | X_n = x, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) = P(X_{n+1} = y | X_n = x) \quad \forall n \in \mathbb{N}_0. \quad (2)$$

De hecho, $X_n = x$ significa que al tiempo de la n -ésima etapa el estado del sistema es x . Además, se dice que la cadena de Markov \mathbf{X} es finita siempre que \mathbb{S} sea finito.

Elementos de una cadena de Markov

Dada una cadena de Markov \mathbf{X} a tiempo discretos, para cada pareja de estados $x, y \in \mathbb{S}$, la probabilidad condicional

$$P(x, y) := P(X_{n+1} = y | X_n = x)$$

denota la (función de) **probabilidad de transición** de ir de x a y **en un paso**. En general, las probabilidades de transición no sólo dependen de los estados sino también del instante en el cual se efectúa la transición misma; no obstante, en adelante trataremos únicamente con cadenas de Markov **homogéneas** (en el tiempo), de modo que

$$P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x) \quad \forall n \in \mathbb{N}_0;$$

luego, si todas las probabilidades de transición son de este tipo diremos que la cadena de Markov (CM) respectiva es **homogénea**.

Definición 1. Se define la **matriz de transición (en un paso)** de la CM \mathbf{X} , como

$$\mathbf{P} := [P(x, y)] := \begin{matrix} & \begin{matrix} x_0 & x_1 & \dots \end{matrix} \\ \begin{matrix} x_0 \\ x_1 \\ \vdots \end{matrix} & \begin{bmatrix} P(x_0, x_0) & P(x_0, x_1) & \dots \\ P(x_1, x_0) & P(x_1, x_1) & \dots \\ \vdots & \vdots & \ddots \end{bmatrix} \end{matrix}$$

Observación 1. De lo anterior se desprenden las propiedades siguientes:

$$P(x, y) \geq 0 \quad \forall x, y \in \mathbb{S}.$$

$\sum_{y \in \mathbb{S}} P(x, y) = 1 \quad \forall x \in \mathbb{S}$, es decir, cada fila representa una distribución de probabilidad.

En caso de que \mathbb{S} sea finito por contener únicamente k estados ($k \in \mathbb{N} := \mathbb{N}_0 - \{0\}$), entonces \mathbf{P} es una matriz cuadrada de orden k , esto es, con k filas y k columnas.

A manera de visualización supongamos que día a día en determinada región del país para el pronóstico del tiempo solamente se manejan dos estados de interés: *mayormente soleado* (H) siempre que el sol únicamente luce radiante al menos 8 horas al día, o bien, en caso contrario como *nublado* (N), de manera que $\mathbb{S} := \{H, N\}$ y, para lo cual asumiremos que: *i*) en cualquier día el estado del tiempo para el día siguiente depende exclusivamente de las condiciones de hoy, tal como la propiedad (2); *y*, *ii*) para constantes fijas $p, q \in (0,1)$, la matriz de transición y la representación gráfica correspondientes están dadas como sigue

$$\mathbf{P} = \begin{matrix} & \begin{matrix} H & N \end{matrix} \\ \begin{matrix} H \\ N \end{matrix} & \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \end{matrix}$$

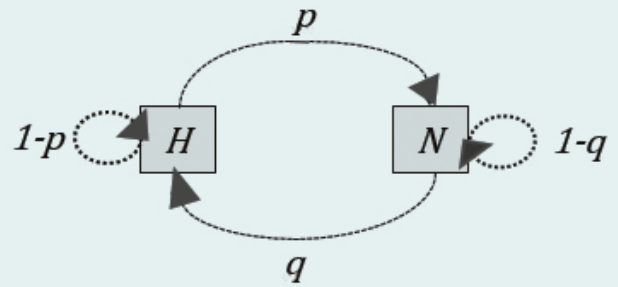


Figura2.

Si para cada $n \in \mathbb{N}_0$, X_n representa el estado del tiempo al n -ésimo día, nótese que:

$$P(X_{n+1} = N | X_n = H) = p = 1 - P(X_{n+1} = H | X_n = H) \quad y$$

$$P(X_{n+1} = H | X_n = N) = q = 1 - P(X_{n+1} = N | X_n = N)$$

Por otra parte, cuando se requiere determinar probabilidades tales como las conjuntas o quizá las marginales, de acuerdo al Teorema de Probabilidad Total así como al de Multiplicación de Probabilidades en su caso, resultan imprescindibles tanto el concepto de probabilidad de transición en n pasos, como el de distribución inicial. De modo que, para una CM \mathbf{X} (a tiempo discreto), específicamente la **probabilidad de transición en n etapas** de ir de x a y (justo después de realizar n transiciones) denotada por $P_n(x, y)$ para cada par de estados $x, y \in \mathbb{S}$ queda definida por

$$P_n(x, y) := P(X_n = y | X_0 = x) \quad \forall n \in \mathbb{N}_0$$

bajo las dos restricciones: $P_1(x, y) = P(x, y)$ y $P_0(x, y) := \delta(x, y)$, donde

$$\delta(x, y) := \begin{cases} 1 & \text{si } y = x, \\ \vdots & \\ 0 & \text{si } y \neq x. \end{cases}$$

Asimismo, la suposición sobre homogeneidad sustenta que para cada $x, y \in \mathbb{S}$

$$P_n(x, y) := P(X_{n+m} = y | X_m = x) \quad \forall n, m \in \mathbb{N}_0,$$

y en tal contexto, la **matriz de transición en n pasos** queda expresada como

$$\mathbf{P}_n := [P_n(x, y)] \quad \forall n \in \mathbb{N}_0,$$

donde \mathbf{P}_0 coincide con la matriz identidad ($\mathbf{P}_0 := I$), mientras que, $\mathbf{P}_1 \equiv \mathbf{P}$. Y en consecuencia, como lo presenta por ejemplo M.E. Caballero [6], para cada $n, m \in \mathbb{N}_0$ sin mayor dificultad se obtiene la relación $\mathbf{P}_{n+m} = \mathbf{P}_n \cdot \mathbf{P}_m$, generando un valioso resultado conocido como **ecuaciones de Chapman-Kolmogorov** y cuya expresión detallada (término a término) establece que

$$P_{n+m}(x, y) = \sum_{z \in \mathbb{S}} P_n(x, z) P_m(z, y) \quad \forall x, y \in \mathbb{S}; \quad n, m \in \mathbb{N}_0,$$

de donde, \mathbf{P}_n coincide con \mathbf{P}^n (n -ésima potencia de \mathbf{P}). Por otra parte, para cada $n \in \mathbb{N}$ el vector-renglón $\boldsymbol{\pi}^{(n)} := \{\pi_n(x), x \in \mathbb{S}\}$ representa la distribución en la n -ésima etapa, ya que

$$\pi_n(x) := P(X_n = x), x \in \mathbb{S};$$

de lo cual, para cada $n \in \mathbb{N}$ se desprende que $\sum_{x \in \mathbb{S}} \pi_n(x) = 1$ y $\pi_n(x) \geq 0, x \in \mathbb{S}$. Y en particular, $\boldsymbol{\pi}^{(0)}$ representa la **distribución inicial** para la mencionada CMX. De modo que, tal como lo muestra por ejemplo O. Hernández-Lerma [5], o bien, Sh. M. Ross [7], la evolución de dicha CM está determinada completamente por su matriz de transición y su distribución inicial, denotado esto por CM $(\boldsymbol{\pi}^{(0)}, \mathbf{P})$ Presentamos en la siguiente proposición el resultado correspondiente, y cuya demostración se enmarca en un proceso inductivo (sobre n), con sustento en el Teorema de Multiplicación de Probabilidades y la propiedad de Markov (2) por parte de \mathbf{X} .

Proposición 1. Un PE $\mathbf{X} := \{X_n: n \in \mathbb{N}_0\}$ es una CM $(\boldsymbol{\pi}^{(0)}, \mathbf{P})$ si y solo si para cada $n \in \mathbb{N}_0$ y para cada n estados $x_0, x_1, \dots, x_n \in \mathbb{S}$ se satisface la identidad siguiente

$$P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) = \pi_0(x_0)P(x_0, x_1) \cdots P(x_{n-1}, x_n)$$

Siendo un resultado que permite conocer la distribución conjunta de las $k (\in \mathbb{N} - \{1\})$ v.a.'s incluidas en el vector (X_0, X_1, \dots, X_k) . Además, otra consecuencia útil incluida es la siguiente.

Proposición 2. $\boldsymbol{\pi}^{(n)} = \boldsymbol{\pi}^{(0)} \mathbf{P}^n$.

Clasificación de estados

Sean dos estados $x, y \in \mathbb{S}$ se dice que x se comunica con y si para algún $n \in \mathbb{N}_0, P_n(x, y) > 0$, lo cual se denota por $x \rightarrow y$, y que están intercomunicados si además es posible la comunicación recíproca, en cuyo caso se simboliza por $x \leftrightarrow y$. De hecho, tal como lo presenta por ejemplo L. Rincón [2], no es difícil demostrar que dicho concepto de intercomunicación constituye una relación de equivalencia, de donde puede observarse que las clases de equivalencia entre estados que ella genera, son precisamente los subconjuntos de estados (en \mathbb{S}) que se comunican entre sí. Y entonces en particular, una CM con solo una clase es aquella en la que todos sus estados están intercomunicados, en cuyo caso a tal CM se le llama **irreducible**. Por otra parte, a un estado $x \in \mathbb{S}$ se le denomina **absorbente**, si $P(x, x) = 1$, nótese que en tal situación x forma una sola clase.

Definición 2. Para cada estado $x \in \mathbb{S}$ se define su período, denotado por $\delta(x)$, mediante

$$\delta(x) := \text{mcd}\{n \in \mathbb{N}_0: P_n(x, x) > 0\}$$

(mcd =máximo común divisor). Y se dice que x es **aperiódico** si en particular $\delta(x) = 1$.

Observación 2. De lo anterior es posible notar que:

$$x \leftrightarrow y \Rightarrow \delta(x) = \delta(y).$$

Todos los estados intercomunicados tienen el mismo período.

Definición 3. Se dice que un estado $x \in \mathbb{S}$ es **recurrente** si para algún $n \in \mathbb{N}$ ocurre que $P_n(x, x) = 1$, mientras que, en caso contrario se le llama estado **transitorio**.

Para cada $x, y \in \mathbb{S}$ y $n \in \mathbb{N}$, sea y y $f_{xy}^n := P(X_n = y, X_{n-1} \neq y, \dots, X_2 \neq y, X_1 \neq y | X_0 = x)$

Entonces, la probabilidad de que partiendo del estado x la CM visite alguna vez el estado y está dada por

$$f_{xy} := \sum_{n=1}^{\infty} f_{xy}^n.$$

A consecuencia de lo anterior se obtienen dos caracterizaciones: (i) $x \in \mathbb{S}$ es recurrente si y solo si $f_{xx} = 1$, y (ii) $x \in \mathbb{S}$ es transitorio si y solo si $f_{xx} < 1$. Además, para cada estado recurrente $x \in \mathbb{S}$ suele ser de interés analizar su **tiempo medio de recurrencia** (o de retorno) denotado por μ_x (número promedio de transiciones para regresar a x) y expresado por

$$\mu_x = \sum_{n=1}^{\infty} n f_{xx}^n.$$

Por otra parte, un estado $y \in \mathbb{S}$ se llama recurrente positivo si $\mu_y < \infty$, y en caso contrario se le conoce como nulo. Cabe señalar que, específicamente se clasifica como **ergódico** a todo estado aperiódico recurrente positivo, y como CM **ergódica** a aquella tal que todos sus estados son de ese tipo. En ese sentido el resultado a continuación proporciona una importante herramienta.

Proposición 3. Sea una CM ergódica irreducible \mathbf{X} con matriz de transición \mathbf{P} . Entonces, independientemente de $x \in \mathbb{S}$,

$$\pi(y) := \lim_{n \rightarrow \infty} P_n(x, y)$$

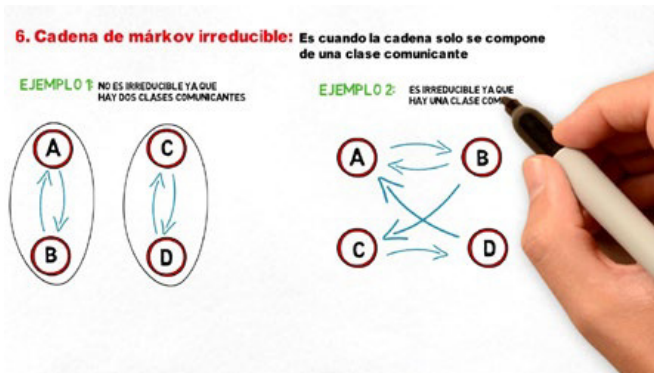
existe, y constituye la única solución no negativa del sistema de ecuaciones definido por las expresiones

$$\pi(y) = \sum_{x \in \mathbb{S}} \pi(x) P(x, y), x \in \mathbb{S}$$

$$\text{y} \quad \sum_{y \in \mathbb{S}} \pi(y) = 1.$$

Modelos ilustrativos

Caminata aleatoria. Como antecedente, considérese una sucesión $\{Z_n: n \in \mathbb{N}_0\}$ de v.a.'s independientes e idénticamente distribuidas (i.i.d.), cuya función de probabilidad para alguna constante dada $p \in (0,1)$ establece que $P(Z_j = +1) = p = 1 - P(Z_j = -1)$. Tal como en particular lo describe R. Romera [8], el PE (1) se le llama caminata



aleatoria siempre y cuando la v.a. X_n se defina como $X_0 := 0$, y para cada $n \in \mathbb{N}$ mediante la expresión

$$X_n := \sum_{j=0}^n Z_j \quad (1)$$

De donde es posible observar que su espacio de estados está dado por $\mathbb{S} = \mathbb{Z}$ (los enteros), y además que para cada $n \in \mathbb{N}_0$ satisface la relación $X_{n+1} = X_n + Z_{n+1}$, lo cual implica que la v.a. X_{n+1} es independiente de la información histórica X_0, \dots, X_{n-1} para cada $n \in \mathbb{N}_0$, ya que depende únicamente de X_n de acuerdo al desarrollo siguiente

$$\begin{aligned} P(X_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0) \\ &= P(X_n + Z_{n+1} = y | X_n = x, X_{n-1} = x_{n-1}, \dots, X_1 = x_1, X_0 = x_0) \\ &= P(X_n + Z_{n+1} = y | X_n = x) = P(Z_{n+1} = y - x), \end{aligned}$$

de donde por (2) el PE $\mathbf{X} := \{X_n : n \in \mathbb{N}_0\}$ en (3) es una CM ($\mathbb{S} \equiv \mathbb{Z}$), tal que para cada $x, y \in \mathbb{S}$

$$P(x, y) = \begin{cases} p & \text{si } y = x + 1 \\ 1 - p & \text{si } y = x - 1 \\ 0 & \text{en otro caso.} \end{cases}$$

Modelo de Enrhenfest (Difusión de moléculas de fluido a través de una membrana). Se consideran $h \in \mathbb{N}$ esferas (idénticas) etiquetadas de 1 a h y distribuidas aleatoriamente entre dos urnas A y B como incluye L. Rincón [2] (véase pág. 36). Luego se elige al azar un número entre 1 y h , acto seguido se observa en qué urna está la esfera correspondiente y se cambia de urna, de modo que la v.a. de interés se define como X_n : número de esferas contenidas en A en la etapa n ; de aquí, cada selección genera una transición del proceso, y entonces se trata de una caminata aleatoria como (3), pero con espacio de estado finito $\mathbb{S} = \{0, 1, \dots, h-1, h\}$, tal que $P(0, 1) = 1 = P(h, h-1)$, en cuyo caso 0 y h se llaman estados reflejantes, mientras que

$$P(x, y) = \begin{cases} \frac{x}{h} & \text{si } x > 0, y = x - 1, \\ \frac{h-x}{h} & \text{si } x < h, y = x + 1, \\ 0 & \text{en otro caso.} \end{cases}$$

Por otra parte y dado que, como es el caso en

muchos de los escenarios, las CM's suelen implementarse para modelar sistemas dinámicos (estocásticos) que evolucionan en el tiempo, resulta entonces importante señalar que entre los rubros de interés sobresalientes en el ambiente de las aplicaciones se encuentran por ejemplo, pronosticar el estado en el que con mayor probabilidad el sistema se encontrará en determinada etapa futura, así como también, conocer acerca del comportamiento límite del sistema cuando aumenta el número de etapas a observar, es decir, si $n \rightarrow \infty$; o bien, para la posible toma de decisiones por etapa, averiguar ya sea:

- ▶ la probabilidad de que en un tiempo finito el sistema alcance algún estado particular, como pudiera ser, la probabilidad de extinción para alguna población dada, por ejemplo;
- ▶ el tiempo promedio transcurrido hasta que el sistema alcance algún estado específico;
- ▶ el número de visitas promedio que el sistema hace a determinado estado, entre otros posibles.

COMENTARIOS

Entre las diversas publicaciones sobre trabajos que se apoyan en la teoría de CM's aplicadas a otras áreas, podemos mencionar por ejemplo: el trabajo doctoral con el que, en 2013, M. Jiménez-López [9] modeló el comportamiento del cáncer de mama mediante una CM a tiempo discreto considerando 3 estados: sin recaída (1), recaída (2), y muerte (3); siendo 1 el estado inicial para todo paciente justo después de la cirugía, mientras que 3 se caracteriza por ser absorbente; asimismo, enfatizamos aquí la interesante y novedosa estrategia que recién (en el año en curso) E. Delgado-Moya y A. Marrero-Severo [10] acaban de proponer sobre control en lo que corresponde a la dinámica de transmisión del dengue, basada precisamente en una CM que incluyó 5 estados: susceptible (S), infectado (I), inmune (R), muerte por dengue (E) y muerte por causas naturales (M). Por otra parte, resulta notable la aplicación que para efecto de predicción, en cuanto a la dinámica del comportamiento de pacientes ingresados a una unidad cardiológica de cuidado intensivo, publicaron en 2006 V. Albornoz V, M. Hinrichsen, P. Miranda y P. Peña [11], considerando para ello una CM con un total de 5 estados: riesgo bajo (A), riesgo medio (B), riesgo alto (C), riesgo grave (D) y el paciente abandonó la unidad (E); asimismo, es imposible dejar de citar a M. Paegelow, T. Camacho y J. Menor [12], quienes dedicados al mundo del paisajismo, en 2003 publicaron resultados obtenidos al analizar de forma paralela las series cronológicas de usos de suelo respectivas para dos regiones europeas, Valle de los Garrotxes y Alta Alpujarra, donde como parte de su herramienta de trabajo implementaron la teoría de CM's asumiendo 8 y 10 estados, respectivamente, y teniendo como objetivo, tanto el reconstruir la evolución correspondiente, como de precisar las tendencias de cambio o de estabilidad paisajística a futuro; siendo esto solo una pequeña muestra, ya que



también los hay en Genética, Ciencias Computacionales, Movilidad Humana, etc..

CONCLUSIONES

De acuerdo a los múltiples trabajos sustentados en CM's que al momento han sido publicados, es posible constatar que, de hecho, éstas estructuras proporcionan una ingeniosa y elegante herramienta matemática que ha venido facilitando el manejo de información acumulada, apoyando con ello la formulación de políticas dirigidas a resolver problemas reales en tan diversas áreas ajenas a las Matemáticas, lo cual enfatiza el potencial de dicho instrumento.

BIBLIOGRAFÍA

[1] J. J. O'Connor, E. F. Robertson. Mac Tutor History of Mathematics Archive. <http://www-history.mcs.st-and.ac.uk/Biographies/Markov.html> [Último acceso: 2018]

[2] L. Rincón. Introducción a los Procesos Estocásticos. Facultad de Ciencias, UNAM. <http://lya.fciencias.unam.mx/lars/libros/procesos2012.pdf>; [Último acceso: 2018]

[3] Z. Brzezniak, T. Zastawniak. Basic Stochastic Processes: A Course Through Exercises. 4ta. ed. Springer Undergraduated Mathematics Series ISSN 1615-2085, Edit. Springer-Verlag London Berlin Heidelberg, año 2002.

[4] O. Hernández-Lerma. Probabilidad y Procesos Estocásticos. Vol. I; Departamento de Matemáticas-CINVESTAV, IPN; año 2005.

[5] O. Hernández-Lerma. Probabilidad y Procesos Estocásticos. Vol. II; Departamento de Matemáticas-CINVESTAV, IPN; año 2005.

[6] M.E. Caballero. Procesos Estocásticos I: Cadenas de Markov. Serie: Vínculos Matemáticos, Núm. 224; Publicaciones Matemáticas, Facultad de Ciencias, UNAM; año 1999.

[7] Sh. M. Ross. Introduction to Probability Models, 10a. edic., Edit. Academic Press, Elsevier; Oxford, UK; pp. 191-290; año 2010.

[8] R. Romera. Cadenas de Markov con Parámetro Discreto. Año 2009. http://ocw.uc3m.es/estadistica/procesos-estocasticos-con-aplicaciones-al-ambito-empresarial/lecturas/Tema2_Cadenas_de_Markov_de_parametro_discreto.pdf [Último acceso: 2018]

[9] M. Jiménez López. Procesos de Markov discretos en supervivencia: estudio del cáncer de mama. (Tesis Máster) Depto. de Estadíst. e Investig. Operativa, Univ. de Granada; año 2013. http://masteres.ugr.es/moea/pages/tfm-1213/tfm_jimenezlopezmnieves/I, [Último acceso: 2018]

[10] E. Delgado-Moya, A. Marrero-Severo. Estudio estocástico con el uso de cadenas de Markov para la transmisión del dengue. Rev. UNICIENCIA Vol. 32, No. 1, pp. 108-117; año 2018. ISSN: 2215-3470 DOI: <http://dx.doi.org/10.15359/ru.32-1.7>, año 2018. [Último acceso: 2018]

[11] V. Albornoz, M. Hinrichsen, P. Miranda, P. Peña. Uso de cadenas de Markov para la predicción de la dinámica del comportamiento de pacientes en una unidad de cuidado intensivo cardiológica. Ingeniare - Revista Chilena de Ingeniería, 14:2 pp. 153-158, año 2016. https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-33052006000100009 [Último acceso: 2018]

[12] M. Paegelow, T. Camacho, J. Menor. Cadenas de Markov, evaluación multicriterio y evaluacdión multiobjetivo para la modelización prospectiva del paisaje. GeoFocus: Revista Internacional de Ciencia y Tecnología de la Información Geográfica, Núm. 3, p.22-44, ISSN: 1578-5157, año 2003. http://geofocus.rediris.es/docPDF/Articulo2_2003.pdf [Último acceso: 2018]