**UNIVERSIDAD CENTROCCIDENTAL**

**“LISANDRO ALVARADO”**

**Decanato de Ciencias y Tecnología**

**Licenciatura en Ciencias Matemáticas**

# Análisis de la evolución de la epidemia de VIH-SIDA en España durante el año 2004 a través del uso de una Cadena de Markov Homogénea con tiempos de observación anuales

Trabajo Especial de Grado presentado por

Br. Evelyn Arenas Carrillo

COMO REQUISITO FINAL

PARA OBTENER EL TÍTULO DE LICENCIADO

EN CIENCIAS MATEMÁTICAS

ÁREA DE CONOCIMIENTO: Estadística

TUTOR: Lic. Maria Teresa Esteves

Barquisimeto,Venezuela Abril de 2017

Universidad centroccidental
"Lisandro Alvarado"
Decanato de Ciencias y Tecnología
Licenciatura en Ciencias Matemáticas

ACTA
TRABAJO ESPECIAL DE GRADO

Los suscritos miembros del Jurado designado por el Jefe del Departamento de Matemáticas del Decanato de Ciencias y Tecnología de la Universidad Centroccidental "Lisandro Alvarado", para examinar y dictar el veredicto sobre el Trabajo Especial de Grado titulado:

"Análisis de la evolución de la epidemia de VIH-SIDA en España durante el año 2004 a través del uso de una Cadena de Markov Homogénea con tiempos de observación anuales"

Presentado por la ciudadana Br. Evelyn Arenas Carrillo titular de la Cédula de Identidad No. 18.799.959, con el propósito de cumplir con el requisito académico final para el otorgamiento del título de Licenciado en Ciencias Matemáticas.

Luego de realizada la Defensa y en los términos que imponen los Lineamientos para el Trabajo Especial de Grado de la Licenciatura en Ciencias Matemáticas, se procedió a discutirlo con la interesada habiéndose emitido el veredicto que a continuación se expresa

[1]___________________

Con una calificación de _________________ puntos.

En fe de lo expuesto firmamos la presente Acta en la Ciudad de Barquisimeto a los ______ dias del mes de ____________ de ____________

| | |
|---|---|
| ___________________ | ___________________ |
| TUTOR | FIRMA |
| ___________________ | ___________________ |
| PRINCIPAL | FIRMA |
| ___________________ | ___________________ |
| PRINCIPAL | FIRMA |

OBSERVACIONES:

_______________________________________________________________________________

_______________________________________________________________________________

_______________________________________________________________________________

[1] Aprobado ó Reprobado

*A mis Padres y a mi hija Alana*

# Agradecimientos

Primeramente a Dios Todopoderoso por llenarme de fortaleza para seguir adelante y conducirme en el camino indicado para así alcanzar esta meta tan deseada.

Seguidamente a mis padres Altamira Carrillo y Adolfo Arenas los cuales fueron pilares fundamentales para lograr esta meta, pero en especial a ti (mamá) por estar allí en los momentos más indicado de mi vida, por siempre creer en mí, y darme la confianza de que si podía lograrlo. Eternamente agradecida con ustedes los amo a los dos.

Al regalito más hermoso de mi vida mi Alana Valentina, la que termino de darme ese impulso ya que mami se esfuerza para darte el mejor de los ejemplo , eres mi personita favorita y Te AMO inmensamente, así que esta meta lograda mami te la dedica a ti mi princesa.

A mi esposo, mi amor Jean Carlos (Catire) por apoyarme ya que fuiste pilar fundamental para que lograra esta meta, gracias mil gracias te amo.

Mis hemanos vilklys, veykler y paola personas que adoro, gracias por siempre creer en mi, este triunfo también es de ustedes.

Mi otra personita especial mi sobrino Veycker Adrian tia te ama y se que vas a estar orgulloso de tia te Amo mi rey.

Y a mis amigos ya que fueron gratos los momentos que hemos recorrido juntos, entre ellos karla, Mafer, Gene, Shaday, Joelviz, eme, entre otros, así que mil gracias a todos por su amistad, paciencia y fidelidad los quiero.

A mi tutora Profesora Maria Teresa gracias por su paciencia y enseñanza agradecida con usted profe...

# Resumen

En este trabajo se estudiará la teoría básica de los procesos estocásticos, haciendo particular énfasis en el estudio de las propiedades de las cadenas de Markov en tiempo discreto, posteriormente estudiaremos el artículo Modelos de Markov aplicados a la investigación en ciencias de la salud [1], en donde se utilizo un modelo Markoviano para estudiar la evolución de los casos de VIH-SIDA en España durante el año 2004.

# Índice general

# Introducción

En ciencias de la salud, muchas variables de interés muestran cambios en el tiempo, con el fin de predecir qué valor futuro alcanzará una variable bajo determinadas condiciones. Esto constituye una importante fuente de información para la investigación básica y aplicada.

El análisis de procesos estocásticos permite predecir el estado en que se encontrará el proceso en el futuro a partir de la información disponible sobre su pasado. La complejidad de estas predicciones dependen en gran medida del tipo de proceso; sin embargo las características de algunos de ellos facilitan su análisis e interpretación.

Desde comienzos de los años 80, el Sida se ha convertido en una pandemia de nuestra época. En la actualidad, España es uno de los paises del oeste de Europa más afectado por el virus [1], por lo que el conocimiento de la evolución futura de la epidemia es un objetivo prioritario para establecer políticas sanitarias adecuadas.

Además, el proceso estocástico que mejor se adapta a la evolución de los casos de VIH-SIDA es una cadena de Markov discreta.

# Capítulo 1

# Preliminares

Para el desarrollo de este capitulo seguiremos los conceptos desarrollados Athreya K, Lahiri S *Measure Theory and Probability Theory*

**Definición 1.1 (Espacio de probabilidad)** *Un espacio de probabilidad es una terna* $(\Omega, F, \mathbb{P})$ *donde* $\Omega$ *es el espacio muestral de todos los posibles resultados de un experimento aleatorio;* $F$ *es una* $\sigma$*-algebra de conjuntos, es decir,* $F$ *es una colección que sastiface*

**1.** $A \in F \Rightarrow A^c \in F$

**2.** $A_1, A_2 \in F \Rightarrow A_1 \cup A_2 \in F$

*Así los elementos* $A \in F$*, pertenecientes a la* $\sigma$*-algebra son denominados eventos, por ello,* $F$ *suele ser llamada también la* $\sigma$*-algebra de eventos. En este orden de ideas se define* $\omega$ *como los elementos en* $\Omega$*, así pues un evento* $A \subset \Omega$*, es una colección de puntos muestrales.* $\mathbb{P}$ *es una medida de probabilidad en* $(\Omega, F)$*, es decir,* $\mathbb{P}$ *es una función definida en* $\Omega$ *tal que:*

- $0 \leq \mathbb{P}(A) \leq 1$
- $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$
- $A_i \in F \Rightarrow \cup_{i=1}^{\infty} A_i \in F$

**3.** $A_1, A_2 \in F, A_1 \cap A_2 = \emptyset \Rightarrow P(A_1 \cup A_2) = P(A_1) + P(A_1)$

**Definición 1.2** *Un conjunto de Borel es un elemento de la llamada* $\sigma$*-álgebra de Borel, la cual no es más que la mínima* $\sigma$*-álgebra que contiene a una topología dada. Como caracterización alternativa, también se puede decir que un conjunto de Borel es cualquier conjunto obtenido mediante uniones e intersecciones numerables de conjuntos cerrados o abiertos en la topología considerada.*

**Definición 1.3 (Función Medible)** *Sea* $(\Omega, F, \mathbb{P})$ *un espacio de probabilidad y* $X : \Omega \to \mathbb{R}$ *es* $(F, B(\mathbb{R}))$*-medible, esto es* $X^{-1}(A) \in F$ *para todo* $A \in B(\mathbb{R})$*, entonces* $X$ *es una variable aleatoria en* $(\Omega, F, \mathbb{P})$

*Recordemos que* $X : \Omega \to \mathbb{R}$ *es* $(F, B(\mathbb{R}))$*-medible, si y solo si,* $x \in \mathbb{R}$*,* $\{\omega : X(\omega) \leq x\} \in F$*.*

**Definición 1.4 (Función Distribución)** *Sea $X$ una variable aleatoria en $(\Omega, F, \mathbb{P})$. Sea*

$$F_x(x) \equiv P(\omega \in \Omega : X(\omega) \leq x),\ x \in \mathbb{R} \tag{1.1}$$

*Entonces $F_x(.)$ es conocida como la función distribución acumulada o simplemente función distribución respecto a la variable $X$.*

**Definición 1.5** *Sea $X$ una variable aleatoria en $(\Omega, F, \mathbb{P})$. Sea*

$$P_x(A) \equiv P(X^{-1}(A))\ para\ todo,\ A \in B(\mathbb{R}) \tag{1.2}$$

*Entonces la medida de probabilidad $\mathbb{P}_x$ es llamada distribución de probabilidad de $X$.*

*La definición anterior nos da la posibilidad de definir la probabilidad de que $x$ pertenezca a un intervalo $(a, b]$, es decir,*

$$P(X(\omega) \in (a, b]) = F_x(b) - F_x(a),\ a < b$$

*Puesto que se ha asumido que $X$ es una función medible, el evento $(\omega \in \Omega : X(\omega \in B))$ está también en $F$, de manera que su medida de probabilidad $\mathbb{P}$ está bien definida y por tanto $F_x$. La función $F_x$ es una función no decreciente tal que*

$$\lim_{z \to -\infty} F(x) = 0,\ \lim_{z \to +\infty} F(x) = 1,$$

*y es además una función continua por la derecha.*

*Si $F$ es absolutamente continua, su derivada*

$$\frac{d}{dx} F_x(x) = f_x(x)$$

*donde $f_x$ es llamada una función de densidad, la cual es una función no negativa Lebesgue Integrable cuya integral en la recta real es igual a uno. Dicho de una manera práctica, si $F_x(x)$ es la probabilidad de que una variable aleatoria tome valores menores o iguales a $x$, la cantidad $f_x(x)dx$ puede ser interpretada como la probabilidad de que la variable aleatoria tome valores en el intervalo infinitesimal $[x, x + dx)$.En este caso,*

$$F_x(x) = \int_{-\infty}^{x} F_x(u) du,\ x \in \mathbb{R}$$

*donde $f_x$ es la función de densidad tal que $f_x(x) \geq 0$ para todo $x \in \mathbb{R}$ y $\int_{-\infty}^{\infty} F(x)dx = 1$.*

*Si la variable aleatoria toma valores solamente en un conjunto numerable, entonces se dice que es discreta y su densidad en el punto $x$ se define por $P(X = x)$. En el caso de que $X$ sea una variable continua, $P(X = x) = 0$ para todo $x$*

$$\lim_{h \to 0} F_x(x + h) = F_x(x),\ para\ todo\ x \tag{1.3}$$

**Definición 1.6** *Sea $(\Omega, F, \mathbb{P})$ un espacio de probabilidad, $k \in \mathbb{N}$ y $X : \Omega \to \mathbb{R}^k$ es $(F, B(\mathbb{R})^k)$-medible, es decir, $X^{-1}(A) \in F$ para todo $A \in B(\mathbb{R})^k$. Entonces $X$ es llamado un vector aleatorio $k$-dimensional en $(\Omega, F, \mathbb{P})$.*

*Sea $X = (X_1, \ldots, X_k)$ un vector aleatorio con componentes $X_i$, $i = 1, 2, \ldots, k$. Entonces cada $X_i$, es una variable aleatoria en $(\Omega, F, \mathbb{P})$.*

## 1.1. Valor esperado, Varianza y Momentos

**Definición 1.7** *Sea* $X_1, \ldots, X_k$ *un vector aleatorio de* $(\Omega, F, \mathbb{P})$. *Luego, para cada* $i = 1, \ldots, k$, *la función de distribución* $f_{X_i}$ *y la distribución de probabilidad* $P_{X_i}$, *de la variable aleatoria* $X_i$ *es llamada función de distribución marginal y distribución de probabilidad marginal de* $X_i$, *respectivamente.*

*Es evidente que la distribución de* $X$ *determina la distribución marginal* $P_{X_i}$ *de* $X_i$ *para cada* $i = 1, \ldots, k$.

**Definición 1.8** *Sea* $X$ *una variable aleatoria en* $(\Omega, F, \mathbb{P})$. *El valor esperado de* $X$, *se denota por* $\mathbb{E}X$ *o* $\mathbb{E}(X)$ *y se define como*

$$\mathbb{E}(X) = \int_{\Omega} X d\mathbb{P}, \tag{1.4}$$

*siempre que la integral este bien definida.*

*Si* $X$ *es una variable aleatoria en* $(\Omega, F, \mathbb{P})$ *y* $h : \mathbb{R} \to \mathbb{R}$ *es una medición de Borel, entonces* $Y = h(X)$ *es tambien una variable aleatoria en* $(\Omega, F, \mathbb{P})$.

**Definición 1.9** *Para cualquier entero positivo* $n$ *el momento* $\mu_n$ *de una variable aleatoria* $X$ *es definida por*

$$\mu_n \equiv \mathbb{E}(X^n), \tag{1.5}$$

*siempre que la expresión este bien definida.*

**Definición 1.10** *La varianza de una variable aleatoria* $X$ *se define como*

$$\mathbb{V}ar(X) = \mathbb{E}(X - \mathbb{E}(X))^2, \tag{1.6}$$

*siempre que* $\mathbb{E}(X)^2 < \infty$.

**Definición 1.11** *La función generadora de momento de una variable aleatoria* $X$ *se define como*

$$M_X(t) \equiv \mathbb{E}(e^{tX}) \text{ para todo, } t \in \mathbb{R} \tag{1.7}$$

*Por lo tanto* $e^{tX}$ *siempre es no negativo,* $\mathbb{E}(e^{tX})$ *esta bien definido, pero puede ser infinito.*

**Definición 1.12** *La matriz de covarianzas de* $X$ *denotada por* $\sigma_X$, *se define por*

$$\Sigma_X = (\mathbb{C}ov(X_i, X_j) : i, j = 1, \ldots, n) \tag{1.8}$$

*donde* $\mathbb{C}ov(X_i X_j) = \mathbb{E}\{(X_i - \mathbb{E}(X_i))(X_j - \mathbb{E}(X_j))\}$ *es la covarianza de* $X_i$ *y* $X_j$. *Si* $i = j$, *entonces* $\mathbb{C}ov(X_i, X_j) = \sigma_X^2$

## 1.2. Independencia

**Definición 1.13** *Dos variables aleatorias* $X$ *e* $Y$ *son independientes si*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B) \tag{1.9}$$

*para cualquiera dos conjuntos $A$ y $B$ en $\mathbb{R}$*

*Alternativamente se puede definir la independencia mediante las funciones de distribución y las funciones de densidad. Las variables aleatorias $X_1$ y $X_2$ son independientes si y sólo si:*

$$F_{X_1,X_2}(x_1,x_2) = F_{X_1}(x_1)F_{X_2}(x_2) \;\; x_1, x_2 \in \mathbb{R}$$

*Si $(X_1, X_2)$ tiene función de densidad $f_{X_1,X_2}$ con funciones de densidad marginales $f_{X_1}$, y $f_{X_2}$. Entonces las variables aleatorias $X_1$ y $X_2$ son independientes si y sólo si:*

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2) \;\; x_1, x_2 \in \mathbb{R}$$

*Una consecuencia importante de la independencia de variables aleatorias es la propiedad siguiente:*

*Si $X_1, ...., X_n$ son variables aleatorias independientes, entonces para cualquier funciones reales $g_1, ..., g_n$*

$$E[g_1(X_1), ..., g(X_n)] = E[g_1(X_1)], ..., E[g_n(X_n)] \tag{1.10}$$

*siempre que la esperanzas consideradas esten bien definidas.*

**Definición 1.14** *Se define la correlación de $X_1$ y $X_2$ como*

$$\mathbb{C}orr(X_1, X_2) = Cov\frac{(X_1, X_2)}{\sqrt{Var(X_1)Var(X_2)}} \tag{1.11}$$

*como se puede ver, es el resultado de estandarizar la covarianza de las variables $X_1$ y $X_2$, por la que su valor está comprendido entre $-1$ y $1$.*

*En particular, se puede concluir que las variables aleatorias independientes $X_1$ y $X_2$ son no correlacionadas esto es, $\mathbb{C}orr(X_1, X_2) = 0$. En general, el recíproco no es cierto, es decir, las variables aleatorias no correlacionadas no son necesariamente independientes.*

## 1.3. Teorema de Convergencia Monótona

**Definición 1.15 (Suceción creciente)** *Una sucesión $(a_n)_{n\in\mathbb{N}}$ en $\mathbb{R}$ se llama creciente si $a_n \leq a_{n+1}$ para todo $n \in \mathbb{N}$.*

**Definición 1.16 (Sucesión estrictamente creciente )** *Una sucesión $(a_n)_{n\in\mathbb{N}}$ en $\mathbb{R}$ se llama estrictamente creciente si $a_n < a_{n+1}$ para todo $n \in \mathbb{N}$.*

*El concepto de sucesión creciente de funciones se define punto a punto.*

**Definición 1.17 (Sucesión crecientes de funciones)** *Sea $(f_n)n \in \mathbb{N}$ una sucesión de funciones $X \longrightarrow \mathbb{R}$. Decimos que esta sucesión de funciones es creciente si para cada punto $x$ en $X$ la sucesión numérica $(f_n(x))n \in \mathbb{N}$ es creciente.*

**Teorema 1.1 (Teorema de lebesgue))** *Sea $(X, F, \mu)$ un espacio de medida y sea $(fn)_{n\in\mathbb{N}}$ una sucesión creciente en $M(X, F, [0, +\infty])$. Denotemos por $g : X \longrightarrow [0, +\infty]$ a la función límite:*

$$g(x) := \lim_{n\longrightarrow\infty} f_n(x). \tag{1.12}$$

*Entonces* $g \in M(X, F, [0, +\infty])$ *y*

$$\int_X g d\mu = \lim_{n \longrightarrow \infty} \int_X f_n d\mu. \tag{1.13}$$

## 1.4. Ley de los grandes números

Esto establece que, bajo ciertas condiciones, el promedio de variables aleatorias converge a una constante cuando el número de sumandos tiende a infinito.

Existen dos versiones de esta afirmación, la primera basada en el concepto de convergencia la cual se conoce como la *ley débil* y la segunda, basada en el concepto de convergencia casi segura se llama *ley fuerte*. La ley fuerte implica entonces la ley débil.

**Definición 1.18 (Ley débil de los grandes números.)** *Si* $X_1, X_2, \ldots$ *son variables independientes idénticamente distribuida con media* $\mu$ *y varianza finita* $\sigma^2$*, entonces se da la convergencia en probabilidad,*

$$\frac{1}{n} \sum_{i=1}^{n} X_i \longrightarrow \mu. \tag{1.14}$$

**Definición 1.19 (Ley fuerte de los grandes números.)** *Sean* $X_1, X_2, \ldots$ *independientes e idénticamente distribuidas con media* $\mu$*. Entonces se da la convergencia casi segura*

$$\frac{1}{n} \sum_{i=1}^{n} X_i \longrightarrow \mu. \tag{1.15}$$

## 1.5. Teorema de la Probabilidad Total

**Teorema 1.2** *Sea* $A_1, \ldots, A_n$ *una partición sobre el espacio muestral y sea* $B$ *un suceso cualquiera del que se conocen las probabilidades condicionales* $P(B|A_i)$*, entonces la probabilidad del suceso* $B$ *viene dada por la expresión:*

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

## 1.6. Procesos estoscásticos

**Definición 1.20** *Sea* $(\Omega, F, \mathbb{P})$ *un espacio de probabilidades. Un proceso estocástico es una colección de variables aleatorias* $\{X_t : t \in T\}$ *parametrizada por un conjunto* $T$*, llamado espacio parametral, y con valores en un conjunto* $S$ *llamado espacio de estados.*

*Para* $T = N$*, tenemos un proceso en tiempo discreto y para* $T \subset \mathbb{R}$ *tenemos un proceso en tiempo continuo.*

Un proceso estocástico puede considerarse como una función de dos variables $X : T \times \Omega \longrightarrow S$ tal que a la pareja $(t, \omega)$, se le asocia el estado $X(t, \omega)$, lo cual también puede escribirse como $X_t(\omega)$. Para cada valor de $t$ en $T$, el mapeo $\omega \longmapsto X_t(\omega)$ es una variable aleatoria, mientras que para cada $\omega$ en $\Omega$ fijo, la función $t \longmapsto X_t(\omega)$ es llamada una trayectoria o realización del proceso. Es decir, a cada $\omega$ del espacio muestral le corresponde una trayectoria del proceso. Es por ello que a veces se define un proceso estocástico como una función aleatoria.

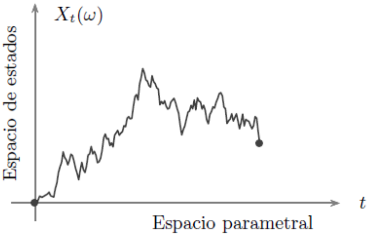


Figura 1.1: Paseo Estocastico

**Definición 1.21 (Momentos de un Proceso Estocástico)** *El valor esperado y varianza de un proceso estocástico son definidos por*

$$\mathbb{E}(X_t) = \int_{\Omega} X(\omega, t) d\mathbb{P}(\omega), 0, 2cmT \in [0, T] \tag{1.16}$$

$$\mathbb{V}ar(X_t) = \mathbb{E}[X_t - \mathbb{E}(X_t)]^2 0, 2cmT \in [0, T] \tag{1.17}$$

*El $k$-ésimo momento de $X_t$, $k \geq 1$, se define, para $t \in [0, T]$, como $\mathbb{E}(X_t^k)$. Esas cantidades están bien definidas cuando las correspondientes integrales son finitas.*

# Capítulo 2

# Cadenas de Markov

Este capitulo se desarrollará siguiendo el libro de Introducción a los PROCESOS ESTOCÁSTICOS de Luis Rincón. Vamos a considerar procesos estocásticos a tiempo discreto $\{X_n\}$ que cumplen la propiedad de Markov. Para describir, a esta propiedad y varias de sus condiciones equivalentes de una manera simple, a la probabilidad $P(X_n = x_n)$ se le escribirá como $p(x_n)$ es decir, el subíndice indica también la variable a la que se hace referencia. El significado de la probabilidad condicional $p(x_{n+1} \mid x_n)$ es análogo.

**Definición 2.1 (Propiedad de Markov)** *Una cadena de Markov a tiempo discreto es una sucesión de variables aleatorias $X_n$, $n \geq 1$ que toman valores en un conjunto finito o numerable $\xi$, conocido como espacio de estados, y que satisface la siguiente propiedad*

$$P(X_{n+1} = j|X_0 = i_0, \ldots, X_{n-1, X_n = i_n}) = P(X_{n+1} = j|X_n = i_n) \tag{2.1}$$

*para todo $n$ y cualesquiera estados $i_0, \ldots, i_n$, $j$ en $\xi$. (reftlf) se conoce como la propiedad de Markov*

## 2.1. Cadenas de Markov

**Definición 2.2** *Una cadena de Markov es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, \ldots\}$, con espacio de estados discreto, y que sastiface la propiedad de Markov, esto es, para cualquier entero $n \geq 0$, y para cualesquiera estados $x_0, \ldots, x_{n+1}$, se cumple*

$$p(x_{n+1} \mid x_0, \ldots, x_n) = p(x_{n+1} \mid x_n) \tag{2.2}$$

*Si el tiempo $n + 1$ se considerá como un tiempo futuro, el tiempo $n$ como el presente y los tiempos $0, 1, \ldots, n - 1$ como el pasado, entonces la condicion (2.2) anterior establece que la distribución de probabilidad del estado del proceso al tiempo futuro $n + 1$ depende unicamente del estado del proceso al tiempo $n$, y no depende de los estados en los tiempos pasados $0, 1, \ldots, n - 1$.*

*Sin perdida de generalidad tomaremos como espacio de estados de una cadena de Markov al conjunto discreto $\{0, 1, \ldots\}$, o cualquier subconjunto finito que conste de los primeros elementos de este conjunto. Cuando el espacio de estados de una cadena de Markov es un conjunto finito se dice que la cadena es finita.*

**Definición 2.3 (Probabilidades de Transición)** *A la probabilidad $P(X_{n+1} = j \mid X_n = i)$ se le denota por $p_{ij}(n, n+1)$, y representa la probabilidad de transición del estado $i$ en el tiempo $n$, al estado $j$ en el tiempo $n+1$. Estas probabilidades se conocen como las probabilidades de transición en un paso. Cuando los números $p_{ij}(n, n+1)$ no dependen de $n$ se dice que la cadena es estacionaria u homogénea en el tiempo. Por simplicidad se asume tal situación de modo que las probabilidades de transición en un paso se escriben simplemente como $p_{ij}$.*

*Variando los índices $i$ y $j$, por ejemplo sobre el conjunto de estados $\{0, 1, \ldots\}$, se obtiene la matriz (2.1) de probabilidades de transición en un paso. La entrada $(i, j)$ de esta matriz es la probabilidad de transición $p_{ij}$, es decir, la probabilidad de pasar del estado $i$ al estado $j$ es una unidad de tiempo. En general, al escribir estas matrices omitiremos escribir la identificación de los estados en los renglones y columnas como aparece en la figura, tal identificación será evidente a partir de conocer el espacio de estados del proceso. El índice $i$ se refiere a la fila de la matriz, y el índice $j$ a la columna.*

**Matriz 2.1**

$$p = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

**Proposición 2.1** *La matriz de probabilidades de transición $P = (p_{ij})$ cumple las siguientes dos propiedades.*

1. $p_{ij} \geq 0$.
2. $\sum_j p_{ij} = 1$.

**Demostración**

La primera condición es evidente a partir del hecho de que estos números son probabilidades. Para la segunda propiedad se tiene que para cualquier estado $i$ y cualquier entero $n \geq 0$,

$$\begin{aligned} 1 &= P(X_{n+1} \in \{0, 1, \ldots\}) \\ &= P(X_{n+1} \in \{0, 1, \ldots\} \mid X_n = i) \\ &= P(\bigcup_j (X_{n+1} = j) \mid X_n = i) \\ &= \sum_j P(X_{n+1} = j \mid X_n = i) \\ &= \sum_j p_{ij}. \end{aligned}$$

Esto último significa que a partir de cualquier estado $i$ con probabilidad uno la cadena pasa necesariamente a algún elemento del espacio de estados al siguiente momento. En general toda matriz cuadrada que cumpla estas dos propiedades se dice que es una matriz estocástica. Debido a la propiedad de Markov, esta matriz captura la esencia del proceso y determina el comportamiento de la cadena en cualquier tiempo futuro. Si además la matriz sastiface la condición $\sum_i p_{ij} = 1$, es decir, cuando la suma por columnas también es uno, entonces se dice que es doblemente estocástica.

**Distribución de probabilidad inicial:**

En general puede considerarse que una cadena de Markov inicia su evolución partiendo de un estado $i$ cualquiera, o más generalmente considerando una distribución de probabilidad inicial sobre el espacio de estados. Una distribución inicial para una cadena de Markov con espacio de estados $\{0, 1, \ldots\}$ es simplemente una distribución de probabilidad sobre este conjunto, es decir, es una colección de números $p_0, p_1, p_2, \ldots$ que son no negativos y que suman uno. El número $p_i$ corresponde a la probabilidad de que la cadena inicie en el estado $i$. En general, la distribución inicial juega un papel secundario en el estudio de las cadenas de Markov.

**Existencia:**

La propiedad de Markov es equivalente a la igualdad $p(x_0, x_1, \ldots, x_n) = p(x_0)p(x_1 \mid x_0) \ldots p(x_n \mid x_{n-1})$. Esta identidad establece que las distribuciones conjuntas $p(x_0, x_1, \ldots, x_n)$ se encuentran completamente especificadas por la matriz de probabilidades de transición y una distribucion inicial.

**Probabilidades de transición en n pasos:**

La probabilidad $P(X_{n+m} = j \mid X_m = i)$ corresponde a la probabilidad de pasar del estado $i$ al tiempo $m$, al estado $j$ al tiempo $m + n$. Dado que hemos supuesto la condición de homogeneidad en el tiempo, esta probabilidad no depende realmente de $m$, por lo tanto coincide con $P(X_n = j \mid X_0 = i)$, y se le denota por $p_{ij}(n)$. A esta probabilidad también se le denota como $p_{ij}^{(n)}$; en donde el número de pasos $n$ se escribe entre paréntesis para distinguirlo de algún posible exponente, y se le llama *probabilidad de transición* en $n$ pasos. Cuando $n = 1$ simplemente se omite su escritura, a menos que se quiera hacer énfasis en ello. También cuando $n = 0$ es natural definir

$$p_{ij}(0) = \delta_{ij} = \begin{cases} 0, & \mathrm{si, i \neq j;} \\ 1, & \mathrm{si, i = j.} \end{cases}$$

Es decir, después de realizar cero pasos la cadena no puede estar en otro estado mas que en su lugar de origen. Esta es la función delta de Kronecker.

**Ejemplo 2.1 (Cadena de dos estados.)** *Considere una cadena de Markov con espacio de estados* $\{0, 1\}$*, con matriz y diagrama de transición como aparece en la figura (2.1), en donde* $0 \leq a \leq 1$*, y* $0 \leq b \leq 1$*. Suponga que la distribución inicial esta dada por* $p_0 = P(X_0 = 0)$ *y* $p_1 = P(X_0 = 1)$*.*

*Aunque de aspecto sencillo, esta cadena es susceptible de muchas aplicaciones pues es común encontrar situaciones en donde se presenta siempre la dualidad de ser o no ser, estar o no estar, tener o no tener, siempre en una constante alternancia entre un estado y el otro. Cuando* $a = 1 - b$*, las variables* $X_1, X_2, \ldots$ *son independientes e idénticamente distribuidas con* $P(X_n = 0) = 1 - a$ *y* $P(X_n = 1) = a$*, para cada* $n \geq 1$*. Cuando* $a \neq 1 - b$*,* $X_n$ *depende de* $X_{n-1}$*.*

$$p = \begin{pmatrix} p_{00}(n) & p_{01}(n) \\ p_{10}(n) & p_{11}(n) \end{pmatrix} = \frac{1}{a+b} \begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b} \begin{pmatrix} a & -a \\ -b & b \end{pmatrix}$$

$$p = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

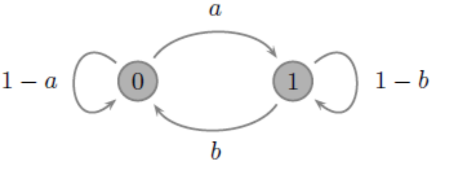


Figura 2.1: Cadena de dos estados

## 2.2. Ecuación de Chapman-Kolmogorov

Esta ecuación es una fórmula sencilla y muy útil que permite descomponer la probabilidad de pasar del estado $i$ al estado $j$ en $n$ pasos, en la suma de probabilidades de las trayectorias que van de $i$ a $j$, y que atraviesan por un estado $k$ cualquiera en un tiempo intermedio $r$.

**Proposición 2.2 (Ecuación de Chapman-Kolmogorov)** *Para cualesquiera números enteros $r$ y $n$ tales que $0 \leq r \leq n$, y para cualesquiera estados $i$ y $j$ se cumple*

$$p_{ij}(n) = \sum_k p_{ik}(r) p_{kj}(n-r).$$

**Demostración**
Por el teorema de probabilidad total y la propiedad de Markov,

$$\begin{aligned}
p_{ij}(n) &= P(X_n = j \mid X_0 = i) \\
&= \sum_k P(X_n = j, X_r = k \mid X_0 = i) \\
&= \sum_k \frac{P(X_n = j, X_r = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_k \frac{P(X_n = j, X_r = k, X_0 = i)}{P(X_r = k, X_0 = i)} \frac{P(X_r = k, X_0 = i)}{P(X_0 = i)} \\
&= \sum_k P(X_n = j \mid X_r = k, X_0 = i) P(X_r = k \mid X_0 = i) \\
&= \sum_k P(X_n = j \mid X_r = k) P(X_r = k \mid X_0 = i) \\
&= \sum_k p_{kj}(n-r) p_{ik}(r) \\
&= \sum_k P_{ik}^{(r)} P_{kj}^{(n-r)}.
\end{aligned}$$

Esta ecuación es importante para hacer ciertos cálculos, y se usa con regularidad en el estudio de las cadenas de Markov. En particular, la siguiente desigualdad será utilizada más adelante: Para cualquier estado $k$ y para $0 \leq r \leq n$, se tiene que $p_{ij}(n) \geq p_{ik}(r)p_{kj}(n-r)$. Como una consecuencia importante de la ecuación de Chapman-Kolmogorov se tiene el siguiente resultado.

**Proposición 2.3** *La probabilidad de transición en $n$ pasos, $p_{ij}(n)$, está dada por la entrada $(i,j)$ de la $n$-ésima potencia de la matriz $P$, es decir,*

$$p_{ij}(n) = (P^n)_{ij}.$$

**Demostración**

*Esta identidad es consecuencia de la ecuación Chapman-Kolmogorov aplicada $n-1$ veces. La suma que aparece abajo corresponde a la entrada $(i,j)$ de la matriz resultante de multiplicar $P$ consigo misma $n$ veces.*

$$\begin{aligned}
p_{ij}(n) &= \sum_{i_1} p_{i,i_1}(1)p_{i_1,j}(n-1) \\
&= \sum_{i_1,i_2} p_{i,i_1}(1)p_{i_1,i_2}(1)p_{i_2,j}(n-2) \\
&\vdots \\
&= \sum_{i_1,\ldots,i_{n-1}} p_{i,i_1}(1)p_{i_1,i_2}(1)\ldots p_{i_{n-1},j}(1) \\
&= (P^n)_{ij}.
\end{aligned}$$

*En palabras este resultado establece que el problema de calcular las probabilidades de transición en $n$ pasos se transforma en obtener la $n$-ésima potencia de la matriz de probabilidades de transición en un paso, es decir,*

$$p = \begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \end{pmatrix}^n$$

*Si se conoce esta matriz y si $p_i = P(X_0 = i)$ es una distribución inicial, entonces la distribución de $X_n$ es $P(X_n = j) = \sum_i p_i p_{ij}(n)$.*

*Cuando una matriz estocástica $P$ es diagonalizable, es decir, cuando puede ser escrita en la forma $QDQ^{-1}$ en donde $D$ es una matriz diagonal, las potencias de $P$ se calculan fálcimente pues $P^n = QD^nQ^{-1}$. Como $D$ es diagonal, $D^n$ es la matriz con cada elemento de la diagonal elevado a la $n$-ésima potencia.*

**Ejemplo 2.2** *Toda matriz estocástica con probabilidad de transición en $n$ pasos, $p_{ij}(n)$, está dada por la entrada $(i,j)$ de la $n$-ésima potencia de la matriz $P$, es decir,*

$$p_{ij}(n) = (P^n)_{ij}.$$

A continuación se va a ilustrar el proceso de diagonalización de una matriz estocástica de dimensión dos. Consideremos la cadena general de dos estados

$$p = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

Los autovalores de esta matriz están dados por la ecuación $|P - \lambda I| = 0$, y resultan ser $\lambda_1 = 1$ y $\lambda_2 = 1 - a - b$. Los correspondientes autovectores escritos como vectores fila son $(1, 1)$ y $(a, -b)$, respectivamente. La matriz $Q$ esta compuesta por los autovectores como columnas, y si se cumple que $a + b > 0$, entonces es invertible, es decir,

$$Q = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}, \qquad Q^{-1} = \frac{1}{a+b}\begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}$$

La matriz $D$ es la matriz diagonal que contiene a los dos autovalores. Puede entonces comprobarse la identidad $P = QDQ^{-1}$, es decir

$$\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} = \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1-a-b \end{pmatrix} \frac{1}{a+b}\begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}.$$

Por lo tanto,

$$\begin{aligned} P^n &= QD^nQ^{-1} \\ &= \begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & (1-a-b)^n \end{pmatrix}\begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}\frac{1}{a+b} \\ &= \frac{1}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix} + \frac{(1-a-b)^n}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix} \end{aligned}$$

## 2.3. Comunicación

Definiremos a continuación a la comunicación entre dos estados de una cadena de Markov como la posibilidad de pasar de un estado a otro en algún número finito de transiciones.

**Definición 2.4** *Se dice que el estado $j$ es accesible desde el estado $i$ si existe un entero $n \geq 0$ tal que $p_{ij}(n) > 0$, esto se escribe simplemente como $i \to j$. Se dice además que los estados $i$ y $j$ son comunicantes, y se escribe $i \leftrightarrow j$, si se cumple que $i \to j$ y $j \to i$.*

*Observemos que siempre se cumple que $i \to i$, pues por definición $p_{ii}(0) = 1$. Además observe que si ocurre que $i \leftrightarrow j$, la accesibilidad de $i$ a $j$ puede darse en un número de pasos distinto que la accesibilidad de $j$ a $i$.*

Es sencillo verificar que la comunicación es una relación de equivalencia, es decir, cumple las siguientes propiedades.

- Es reflexiva: $i \leftrightarrow i$
- Es simétrica: si $i \leftrightarrow j$, entonces $j \leftrightarrow i$.
- Es transitiva: si $i \leftrightarrow j$ y $j \leftrightarrow k$, entonces $i \leftrightarrow k$.

En consecuencia, la comunicación induce una partición del espacio de estados de una cadena de Markov dada por los subconjuntos de estados comunicantes, es decir, dos estados pertenecen al mismo elemento de la partición si, y sólo si, tales estados se comunican. De este modo el espacio de estados de una cadena de Markov se subdivide en clase de comunicación. A la clase de comunicación de un estado $i$ se le denotará por $C(i)$. Por lo tanto, $i \leftrightarrow j$ si, y sólo si, $C(i) = C(j)$.


Figura 2.2: Accesibilidad


Figura 2.3: Comunicación

**Definición 2.5** *Se dice que una cadena de Markov es irreducible si todos los estados se comunican entre sí.*

*En otras palabras, una cadena de Markov es irreducible si existe sólo una clase de comunicación, es decir, si la partición generada por la relación de comunicación es trivial. Por ejemplo, la cadena de racha de éxitos o la cadena de la caminata aleatoria son cadenas irreducibles pues todos los estados se comunican entre sí.*

## 2.4. Periodo

El periodo es un número entero no negativo que se calcula para cada estado de una cadena. Una interpretación de este número será mencionado más adelante y aparecerá también dentro de los enunciados generales sobre el comportamiento límite de cadenas de Markov.

**Definición 2.6** *El periodo de un estado $i$ es un número entero no negativo denotado por $d(i)$, y definido como sigue:*

$$d(i) = m.c.d.\{n \geq 1 : p_{ii}(n) > 0\},$$

*en donde m.c.d. significa máximo común divisor. Cuando $p_{ii}(n) = 0$ para toda $n \geq 1$, se define $d(i) = 0$. En particular, se dice que un estado $i$ es aperiódico si $d(i) = 1$. Cuando $d(i) = k \geq 2$ se dice que $i$ es periódico de periodo $k$.*

**Proposición 2.4** *Si los estados $i$ y $j$ pertenecen a la misma clase de comunicación, entonces tienen el mismo periodo.*

**Demostración**

Claramente el resultado es válido para $i = j$. Suponga entonces que $i$ y $j$ son distintos. Como los estados $i$ y $j$ están en la misma clase de comunicación, existen enteros $n \geq 1$ y $m \geq 1$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Sea $s \geq 1$ un entero cualquiera tal que $p_{ii}(s) > 0$. Tal entero existe pues $p_{ii}(n+m) \geq p_{ij}(n)p_{ji}(m) > 0$. Esto quiere decir que $d(i)|s$. Además $p_{jj}(n+m+s) \geq p_{ji}(m)p_{ii}(s)p_{ij}(n) > 0$. Análogamente, $p_{jj}(n+m+2s) \geq p_{ji}(m)p_{ii}(2s)p_{ij}(n) > 0$. Por lo tanto $d(j)|(n+m+s)$ y $d(j)|(n+m+2s)$. Entonces $d(j)$ divide a la diferencia $(n+m+2s)-(n+m+s) = s$. Por lo tanto, todo entero $s \geq 1$ tal que $p_{ii}(s) > 0$ cumple $d(j)|s$. Pero $d(i)$ divide a $s$ de manera máxima, por lo tanto $d(i) \geq d(j)$. De manera análoga, escribiendo $i$ por $j$, y $j$ por $i$ se obtiene $d(j) \geq d(i)$. Se concluye entonces que $d(i) = d(j)$.

El recíproco del resultado anterior es en general falso, es decir, dos estados pueden tener el mismo periodo y sin embargo no ser comunicantes.

**Proposición 2.5** *Para cada estado $i$, existe un entero $N$ tal que para toda $n \geq N$, se cumple $p_{ii}(nd(i)) > 0$.*

**Demostración**

Si $p_{ii}(n) = 0$ para cada $n \geq 1$, entonces $d(i) = 0$ y por lo tanto la afirmación es válida pues $p_{ii}(0) = 1$, sin embargo la interpretación de recurrencia periódica no se aplica en este caso. Suponga entonces que $n_1, \ldots, n_k$ son enteros tales que $p_{ii}(n_1) > 0, \ldots, p_{ii}(n_k)$. Sea $d = \text{m.c.d.}\{n_1, \ldots, n_k\} \geq d(i)$. Como $d(i)$ es divisor de cada entero $n_1, \ldots, n_k$, se tiene que $d(i)|d$, y por lo tanto existe un entero $q$ tal que $qd(i) = d$.

> *Sea $n_1, \ldots, n_k$ enteros no negativos y sea . $d = m.c.d\{n_1, \ldots, n_k\}$ Entonces existe un entero $M$ tal que para cada $m \geq M$ existen enteros no negativos $c_1 \ldots c_k$ tales que $M = c_1n_1 + \ldots + c_kn_k$.*

Entonces existe un entero no negativo $M$ tal que para cada $m \geq M$, $md = c_1n_1 + \ldots + c_kn_k$, para algunos enteros $c_1, \ldots, c_k$, y por lo tanto

$$p_{ii}(md) = p_{ii}(c_1n_1+, \ldots, +c_kn_k) \geq p_{ii}(c_1n_1) \ldots p_{ii}(c_kn_k) > 0.$$

Por lo tanto, para cada $m \geq M, p_{ii}(md) = p_{ii}(mqd(i)) > 0$. Defina $N = Mq$. Se puede entonces concluir que para toda $n \geq N, p_{ii}(nd(i)) > 0$.

Como corolario de la proposición anterior se tiene que el siguiente resultado: Si es posible pasar de $i$ a $j$ en $m$ pasos, entonces también es posible tal transición en $m + nd(j)$ pasos con $n$ suficientemente grande, suponiendo $d(j) > 0$.

**Proposición 2.6** *Si $p_{ij}(m) > 0$ para algún entero $m$, entonces existe un entero $N$ tal que para toda $n \geq N$ se cumple $p_{ij}(m + nd(j)) > 0$.*

**Demostración**

Por el resultado anterior, para $n$ suficientemente grande, se tiene que $p_{ij}(m + nd(j)) \geq p_{ij}(m)p_{jj}(nd(j)) > 0$.

**Definición 2.7 (Estados Absorbentes)** *Se dice que un estado es absorbente si es cero la probabilidad de hacer una transición fuera de ese estado. Por tanto, una vez que el sistema hace una transición hacia un estado absorbente, permanece en el siempre*

## 2.5. Primeras visitas

En ocasiones interesa el primer momento en el que una cadena de Markov visita un estado particular o un conjunto de estados. Definiremos a continuación este tiempo y después demostraremos una fórmula útil que lo relaciona con las probabilidades de transición.

**Definición 2.8** *Sea $A$ un subconjunto del espacio de estados de una cadena de Markov $X_n$. El tiempo de primera visita al conjunto $A$ es la variable aleatoria*

$$\tau_A = \begin{cases} \min\{n \geq 1 : X_n \in A\} & \text{si} \quad X_n \in A \, \text{para algun n} \geq \ 1 \\ \\ \infty & \text{otro caso} \end{cases}$$

Es decir, $\tau_A$ es el primer momento positivo en el cual la cadena toma un valor dentro de la colección de estados $A$, si ello eventualmente sucede. Estaremos interesados principalmente en el caso cuando el conjunto $A$ consta de un solo estado $j$, y si suponemos que la cadena inicia en $i$ entonces el tiempo de primera visita al estado $j$ se escribe $\tau_{ij}$. Cuando los estados $i$ y $j$ coinciden se escribe simplemente $\tau_i$.

**Definición 2.9** *Para cada $n \geq 1$, el número $f_{ij}(n)$ denota la probabilidad de que una cadena inicia en el estado $i$, llegue al estado $j$ por primera vez en exactamente $n$ pasos, es decir,*

$$f_{ij}(n) = P(\tau_{ij} = n).$$

*Adicionalmente se define $f_{ij}(0) = 0$, incluyendo el caso $i = j$.*

Otra forma equivalente de escribir a la probabilidad de primera visita es a través de la expresión $f_{ij}(n) = P(X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i)$. En particular, observe que $f_{ii}(n)$ es la probabilidad de regresar por primera vez al mismo estado $i$ en el $n$-ésimo paso, y que $f_{ii}(1)$ es simplemente $p_{ii}$. El uso de la letra $f$ para esta probabilidad proviene el término en inglés first para indicar que es la probabilidad de primera visita, saber esto ayuda a recordar su significado.

El siguiente resultado establece que la probabilidad de visitar el estado $j$, a partir de $i$, en $n$ pasos, puede descomponerse en las probabilidades de los eventos disjuntos en los que se presenta la primera visita, la cual puede efectuarse en el primer paso, o en el segundo paso, y así sucesivamente, hasta el último momento posible $n$.

**Proposición 2.7** *Para cada $n \geq 1$,*

$$p_{ij}(n) = \sum_{k=1}^{n} f_{ij}(k) p_{jj}(n-k) \tag{2.3}$$

**Demostración**
La prueba se basa en la siguiente descomposición que involucra eventos disjuntos. Se usa además la propiedad de Markov.

$$\begin{aligned}
P_{ij}(n) &= P(X_n = j \mid X_0 = i) \\
&= \sum_{k=1}^{n} P(X_n = j, X_k = j, X_{k-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i) \\
&= \sum_{k=1}^{n} P(X_n = j \mid X_k = j) P(X_k = j, X_{k-1} \neq j, \ldots, X_1 \neq j \mid X_0 = i) \\
&= \sum_{k=1}^{n} p_{jj}(n-k) f_{ij}(k).
\end{aligned}$$

## 2.6. Recurrencia y transitoriedad

Veremos a continuación que los estados de una cadena de Markov pueden ser clasificados, en una primera instancia, en dos tipos, dependiendo si la cadena es capaz de regresar con certeza al estado de partida.

**Definición 2.10 (I)** *Se dice que un estado $i$ es recurrente si la probabilidad de eventualmente regresar a $i$, partiendo de $i$ es uno, es decir, si*

$$P(X_n = i \ \ para \ \ alguna \ \ n \geq 1 \mid X_0 = i) = 1.$$

*Un estado que no es recurrente se llama transitorio, y en tal caso la probabilidad anterior es estrictamente menor a uno.*

De manera intuitiva, un estado es recurrente si con probabilidad uno la cadena es capaz de regresar eventualmente a ese estado, y cuando ello ocurre en algún momento finito, por la propiedad de Markov, se puede regresar a él una y otra vez con probabilidad uno. Debido a este comportamiento es que al estado en cuestión se le llama recurrente. En cambio, el estado se llama transitorio si existe una probabilidad positiva de que la cadena, iniciando en él, ya no regrese nunca a ese estado. En términos de las probabilidades de primera visita, la probabilidad de una eventual visita al estado $j$, a partir del estado $i$, es el número

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n).$$

Recordemos que $f_{ij}(n)$ es la probabilidad de que la primera visita al estado $j$, a partir de $i$, se efectúe exactamente en el paso $n$. Siendo estos eventos disjuntos para valores distintos de n, esta suma representa la probabilidad de una posible visita al estado $j$. De este modo la definición de recurrencia y transitoriedad puede enunciarse de manera equivalente de la siguiente forma.

**Definición 2.11 (II)** *Un estado $i$ es recurrente si $f_{ii} = 1$, es decir, si la probabilidad de regresar a él en un tiempo finito es uno. Análogamente, un estado $i$ es transitorio si $f_{ii} < 1$.*

Además de la definición, tenemos el siguiente criterio útil para determinar si un estado es recurrente o transitorio.

**Proposición 2.8 (Criterio para la recurrencia)** *El estado $i$ es*

1. *recurrente si, y sólo si* $\displaystyle\sum_{n=1}^{\infty} p_{ii}(n) = \infty$
2. *transitorio si, y sólo si* $\displaystyle\sum_{n=1}^{\infty} p_{ii}(n) < \infty$.

**Demostración**

Sea $N_i$ la variable aleatoria que cuenta el número de veces que el proceso regresa al estado $i$ a partir del primer paso, es decir, $N_i = \sum_{n-1}^{\infty} 1(_{X_n=i})$, cuando $X_0 = i$. Entonces $N_i$ tiene una distribución geométrica pues para $k \geq 1$,

$$
\begin{aligned}
P(N_i \geq k \mid X_0 = i) &= \sum_{m=1}^{\infty} P(N_i \geq k, X_m = i, X_{m-1} \neq i, \ldots, X_1 \neq i \mid X_0 = i) \\
&= \sum_{m=1}^{\infty} P(N_i \geq k \mid X_m = i, X_{m-1} \neq i, \ldots, X_1 \neq i, X_0 = i) \\
&\qquad P(X_m = i, X_{m-1} \neq i, \ldots, X_1 \neq i, X_0 = i) \\
&= P(N_i \geq k-1 \mid X_0 = i) f_{ii}(m) \\
&\vdots \\
&= P(N_i \geq 1 \mid X_0 = i)(f_{ii})^{k-1} \\
&= (f_{ii})^k.
\end{aligned}
$$

La esperanza de $N_i$, posiblemente infinita, puede calcularse de la siguientes dos formas. Primero,

$$
\begin{aligned}
E(N_i \mid X_0 = i) &= \sum_{k=1}^{\infty} P(N_i \geq k \mid X_0 = i) \\
&= \sum_{k=1}^{\infty} (f_{ii})^k \\
&= \begin{cases} \frac{f_{ii}}{1-f_{ii}} & \text{si} \quad 0 \leq f_{ii} < 1, \\ \infty & \text{si} \quad f_{ii} = 1 \end{cases}
\end{aligned}
$$

Por otro lado, por el teorema de convergencia monótona,

$$
\begin{aligned}
E(N_i \mid X_0 = i) &= \sum_{n=1}^{\infty} E(1_{(X_n=i)} \mid X_0 = i) \\
&= \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = i) \\
&= \sum_{n=1}^{\infty} p_{ii}(n).
\end{aligned}
$$

El resultado se sigue de igualar a estas dos expresiones.

Siguiendo con la notación de la demostración anterior, observe que la esperanza $E(N_i \mid X_0 = i)$ es el número promedio de retornos al estado $i$, de modo que un estado $i$ es recurrente si, y sólo si, el número promedio de retornos a él es infinito. En contraparte, un estado $i$ es transitorio si, y sólo si, el número promedio de retornos a él es finito.

Desmostraremos a continuación que la recurrencia y la transitoriedad son propiedades de clase, es decir, si dos estados están en una misma clase de comunicación, entonces ambos estados son recurrentes o ambos estados son transitorios.

**Proposición 2.9** *La recurrencia es una propiedad de clase, es decir,*

1. *Si $i$ es recurrente e $i \leftrightarrow j$, entonces $j$ es recurrente.*
2. *Si $i$ es transitorio e $i \leftrightarrow j$, entonces $j$ es transitorio.*

**Demostración**
Como $i \leftrightarrow j$, existen enteros $n \geq 1$ y $m \geq 1$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Entonces $p_{jj}(m+n+r) \geq p_{ji}(m)p_{ii}(r)p_{ij}(n)$. De modo que, por la ecuación de Chapman-Kolmogorov,

$$\sum_{r=1}^{\infty} p_{jj}(m+n+r) \geq p_{ji}(m) \sum_{r=1}^{\infty} p_{ii}(r)p_{ij}(n).$$

Si $i$ es recurrente, la suma del lado derecho es infinita. Se sigue entonces que la suma del lado izquierdo también lo es, es decir, $j$ es recurrente. La segunda afirmación se demuestra fácilmente usando el primer resultado.

En consecuencia, cuando una cadena es irreducible y algún estado es recurrente, todos los estados lo son, y se dice que la cadena es recurrente. También puede presentarse la situación en donde el espacio de estados conste de varias clases de comunicación recurrentes, en tal caso la cadena también se llama recurrente. En contraparte, una cadena es transitoria si todos los estados lo son, ya sea conformando una sola clase de comunicación de estados transitorios o varias de ellas. De este modo el espacio de estados de una cadena de Markov puede descomponerse en dos subconjuntos ajenos de estados, aquellos que son transitorios y aquellos que son recurrentes.

**Proposición 2.10** *Sea $j$ un estado transitorio. Para cualquier estado inicial $i$, se cumple que $\sum_{n=1}^{\infty} p_{ij}(n) < \infty$. En consecuencia, $\lim_{n\to\infty} p_{ij}(n) = 0$.*

**Demostración** Usando (2.3) y el teorema de Fubini,

$$\begin{aligned}
\sum_{n=1}^{\infty} p_{ij}(n) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} f_{ij}(n-k)p_{jj}(k) \\
&= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} f_{ij}(n-k)p_{jj}(k) \\
&= \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} f_{ij}(m)p_{jj}(k). \\
&= f_{ij} \sum_{k=0}^{\infty} p_{jj}(k) \leq \sum_{k=0}^{\infty} p_{jj}(k) < \infty
\end{aligned}$$

**Proposición 2.11** *Toda cadena de Markov finita tiene por lo menos un estado recurrente.*

**Desmotración**

Por contradicción, supongamos que todos los estados son transitorios. Entonces para cualquier estado $i$ y $j$, se cumple $\sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Sumando sobre el conjunto finito de todos los posibles estados $j$ se obtiene $\sum_j \sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Por otro lado, intercambiando el orden de las sumas se llega a la afirmación contraria $\sum_{n=1}^{\infty} \sum_j p_{ij}(n) = \sum_{n=1}^{\infty} 1 = \infty$. Por lo tanto es erróneo suponer que todos los estados son transitorios, debe existir por lo menos uno que es recurrente.

En consecuencia, toda cadena finita e irreducible es recurrente. Más adelante demostraremos que en tal caso, con probabilidad uno la cadena visita cada uno de sus estados una infinidad de veces.

## 2.7. Tiempo medio de recurrencia

Hemos visto que si una cadena de Markov inicia en un estado recurrente, entonces regresa a él una infinidad de veces con probabilidad uno. Y hemos definido el tiempo de primera visita a un estado $j$, a partir de cualquier estado $i$, como la posibilidad de que tome el valor infinito. Vamos a definir el tiempo medio de recurrencia como la esperanza de esta variable aleatoria en el caso cuando el estado a visitar es recurrente.

**Definición 2.12** *El tiempo medio de recurrencia de un estado recurrente $j$, a partir del estado $i$, se define como la esperanza de $\tau_{ij}$, y se denota por $\mu_{ij}$, es decir,*

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}(n)$$

Recordemos que cuando el tiempo de primera visita se refiere al mismo estado de inicio y de llegada $j$, se escribe $\tau_j$ en lugar de $\tau_{jj}$. En este caso el tiempo medio de recurrencia se escribe simplemente como $\mu_j$. Esta esperanza puede ser finita o infinita, y representa el número de pasos promedio que a la cadena le toma regresar al estado recurrente $j$.

## 2.8. Clases cerradas

Las clases cerradas son subconjuntos de estados de una cadena de Markov que cumplen la propiedad de que partiendo de cualquier estado de este subconjunto, no se puede pasar a cualquier otro estado fuera del subconjunto. Esta propiedad hace a este subconjunto de estados un subsistema propio de la cadena de Markov.

**Definición 2.13** *Una colección de estados no vacía $\mathscr{C}$ es cerrada si ningún estado fuera de $\mathscr{C}$ es accesible desde algún estado dentro de $\mathscr{C}$, es decir, si para cualquier $i \in \mathscr{C}$ y $j \notin \mathscr{C}$, $i \not\to j$.*

Por ejemplo si $i$ es un estado absorbente, entonces la colección $\{i\}$ es claramente una clase cerrada. Como un ejemplo adicional consideremos cualquier clase de comunicación recurrente. Es claro que esta clase es cerrada pues de lo contrario, si $i \in \mathscr{C}$ y $j \notin \mathscr{C}$ con $\mathscr{C}$ recurrente, e $i \to j$, entonces necesariamente $j \to i$ pues hemos supuesto que $i$ es recurrente. Por lo tanto $i \longleftrightarrow j$, y entonces $j \in \mathscr{C}$, contrario a la hipotesis $j \notin C$. Por lo tanto no es posible salir de una clase de comunicación recurrente. El siguiente resultado es una especie de recíproco de lo que acabamos de mencionar.

**Proposición 2.12** *Toda colección de estados que es cerrada e irreducible es clase de comunicación.*

**Demostración**
Sea $\mathscr{C}$ una colección no vacía de estados que es irreducible y cerrada, y sea $i \in \mathscr{C}$. Entonces $\mathscr{C} \subseteq C(i)$ pues como $\mathscr{C}$ es irreducible, todos sus estados se comunican y por lo tanto deben pertenecer a la misma clase de comunicación.

Como $\mathscr{C}$ es cerrada, no es posible salir de tal colección, de modo que la diferencia $C(i)-\mathscr{C}$ es vacía, si existiera $j \in C(i) - \mathscr{C}$, entonces $i \to j$, lo cual contradice el supuesto de que $\mathscr{C}$ es cerrada. Por lo tanto $\mathscr{C} = C(i)$.

## 2.9. Números de visitas

En esta sección vamos a estudiar la variable aleatoria que registra el número de visitas que una cadena realiza sobre un estado $j$ a partir del estado $i$, es decir, para cualquier tiempo finito $n$ se define la variable aleatoria

$$N_{ij}(n) = \sum_{k=1}^{n} 1_{(X_k=j)}, \quad \textit{cuando} \;\; X_0 = i.$$

Cuando los estados $i$ y $j$ coinciden, se escribe $N_i$ en lugar de $N_{ii}(n)$. Observe que $0 \geq N_{ij}(1) \geq N_{ij}(2) \geq \ldots$, es decir, se trata de una suceción monótona creciente de variables aleatorias no negativas que converge casi seguramente a la variable

$$N_{ij}(n) = \sum_{k=1}^{n} 1_{(X_k=j)}, \quad \textit{cuando } X_0 = i.$$

Los siguientes resultados acerca de estas variables aleatorias permiten distinguir la diferencia cualitativa en el comportamiento de los estados transitorios respecto de los recurrentes.

**Proposición 2.13** *Para cualquier estado $i$ y $j$,*

1. $P(N_{ij} \geq k) = \begin{cases} 1 & \text{si} \quad k = 0, \\ f_{ij}(f_{jj})^{k-1} & \text{si} \quad k \geq 1. \end{cases}$

2. $P(N_{ij} = k) = \begin{cases} 1 - f_{ij} & \text{si} \quad k = 0, \\ f_{ij}(f_{jj})^{k-1}(1 - f_{jj}) & \text{si} \quad k \geq 1. \end{cases}$

3. $E(N_{ij}) = \sum_{n=-1}^{\infty} p_{ij}(n) = \begin{cases} 0 & \text{si} \quad f_{ij} = 0, \\ \frac{f_{ij}}{1-f_{jj}} & \text{si} \quad 0 \leq f_{ij} < 1, \\ \infty & \text{si} \quad f_{ij} \neq 0 \, y \, f_{jj} = 1. \end{cases}$

4. $P(N_{ij} = \infty) = \begin{cases} 0 & \text{si} \quad j \, es \, transitorio, \\ f_{ij} & \text{si} \quad j \, es \, recurrente. \end{cases}$

5. $P(N_{ij} < \infty) = \begin{cases} 1 & \text{si} \quad j \, es \, transitorio, \\ 1 - f_{ij} & \text{si} \quad j \, es \, recurrente. \end{cases}$

**Demostración**

1. La primera parte de esta igualdad es evidente. Para demostrar el caso $k \geq 1$ se usa análisis del primer paso,

$$\begin{aligned} P(N_{ij} \geq k) &= \sum_{n=1}^{\infty} f_{ij}(n) P(N_{jj} \geq k-1) \\ &= f_{ij} P(N_{jj} \geq k-1) \\ &= f_{ij}(f_{jj}^{k}). \end{aligned}$$

2. Este resultado se sigue de la primera fórmula y de la igualdad
   $P(N_{ij} = k) = P(N_{ij} \geq k) - P(N_{ij} \geq k+1)$.

3. Por el teorema de convergencia monótona,

$$\begin{aligned} E(N_{ij}) &= E(\sum_{n=1}^{\infty} 1(_{X_n = j}) \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} E(1(_{X_n=j}) \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} P_{ij}(n). \end{aligned}$$

   Por otro lado, usando la primera fórmula

$$\begin{aligned} E(N_{ij}) &= \sum_{k=1}^{\infty} P(N_{ij} \geq k) \\ &= \sum_{k=1}^{\infty} f_{ij}(f_{jj})^{k-1} \\ &= \begin{cases} 0 & \text{si} \quad f_{ij} = 0, \\ \frac{f_{ij}}{1-f_{jj}} & \text{si} \quad 0 \leq f_{jj} < 1, \\ \infty & \text{si} \quad f_{ij} \neq 0 \, y \, f_{jj} = 1. \end{cases} \end{aligned}$$

4. Por la primera fórmula,

$$\begin{aligned} P(N_{ij} = \infty) &= \lim_{k \to +\infty} P(N_{ij} \geq k) \\ &= \lim_{k \to +\infty} f_{ij}(f_{jj})^{k-1} \\ &= \begin{cases} 0 & \text{si} \quad f \, es \, transitorio, \\ f_{ij} & \text{si} \quad j \, es \, recurrente. \end{cases} \end{aligned}$$

5. Esta probabilidad es el complemento de la anterior.

**Distribución de probabilidad de la variable $N_{ij}$:** La variable aleatoria discreta $N_{ij}$ toma valores en el conjunto $\{0, 1, ..., \infty\}$. La función de probabilidad que hemos encontrado para esta variable incluye los siguientes casos:

**a)** Si $f_{ij} = 0$, entonces no es posible visitar $j$ a partir de $i$, y por lo tanto $P(N_{ij} = 0) = 1$, es decir la probabilidad se concentra completamente en el valor 0.

**b)** Si $f_{ij} > 0$ y $f_{jj} = 1$, es decir, si se puede pasar de $i$ a $j$ y $j$ es recurrente, entonces para cualquier valor de $k \geq 1$, $P(N_{ij} \geq k) = f_{ij}$, y por lo tanto $P(N_{ij} = \infty) = f_{ij}$. Mientras que $P(N_{ij} = 0) = 1 - f_{ij}$. Se trata entonces de una medida de probabilidad concentrada en los valores 0 e $\infty$.

**c)** Si $f_{ij} > 0$ y$f_{jj} < 1$, es decir, si se puede pasar de $i$ a $j$ y $j$ es transitorio, entonces la probabilidad se distribuye sobre los valores finitos $0, 1, \ldots$ como indica la formula 2.

**Recurrencia, transitoriedad y número esperado de visitas.** A partir de estas fórmulas podemos ahora distinguir el comportamiento del número de visitas en los casos cuando el estado $j$ es transitorio o recurrente.

**a)** Si $j$ es transitorio, entonces sin importar el estado incial $i$, con probabilidad uno la cadena realiza sólo un número finito de visitas al estado $j$, esto es lo que dice la fórmula 5 y el número esperado de visitas a tal estado es siempre finito por la fórmula 3 con $f_{jj} < 1$, es decir, $E(N_{ij}) = \sum_{n=1}^{\infty} p_{ij}(n) < \infty$. Por lo tanto, encontramos nuevamente que $\lim_{n \to +\infty} p_{ij}(n) = 0$.

**b)** Si $j$ es recurrente, y si se inicia en $j$, entonces con probabilidad uno la cadena regresa a $j$ una infinidad de veces, esto es lo que dice la fórmula 4 con $f_{ij} = f_{jj} = 1$, y el número esperado de visitas al estado $j$ es infinito. Si la cadena inicia en cualquier otro estado $i$, entonces existe la posibilidad de que la cadena nunca visite $j$ ($f_{ij} = 0$), y el númerado esperado de visitas es naturalmente cero (fórmula 3 con $f_{ij} = 0$). Pero si la cadena visita $j$ alguna vez ($f_{ij} > 0$), entonces regresará a $j$ una infinidad de veces, y el número esperado de visitas al estado $j$ es infinito por la formula 3 con $f_{ij} > 0$ y $f_{jj} = 1$.

Anteriormente habíamos demostrado un criterio para la transitoriedad y la recurrencia del estado $i$ en términos de la convergencia o divergencia de la serie $\sum_{n=1}^{\infty} p_{ij}(n)$. En vista de la fórmula 3 ahora podemos corroborar que un estado $i$ es recurrente si, y sólo si, el número promedio de regresos a él es infinito, y es transitorio si, y sólo si, el número promedio de regresos es finito. También en particular, la fórmula 4 demuestra que toda cadena de Markov irreducible y recurrente, visita cada uno de sus estados una infinidad de veces con probabilidad uno.

**Teorema 2.1 (Ergódico para cadenas de Markov:)** *Para cualesquiera estados $i$ y $j$ de una cadena de Markov irreducible se cumple que*

$$\lim_{n \to \infty} \frac{N_{ij}(n)}{n} = \frac{1}{\mu_j}. \tag{2.4}$$

*siendo este límite cero cuando $\mu_j = \infty$.*

**Demostración**

Si la cadena es transitoria, entonces ambos lados de la igualdad se anulan. Suponga que la cadena es recurrente. El tiempo de primera visita al estado $j$ a partir de $i$ es

$\tau_{ij} = \text{mín}\{n \geq 1 : X_n = j \mid X_0 = i\}$. Dada la recurrencia e irreducibilidad, $P(\tau_{ij} < \infty) = 1$, y entonces para cualquier $n \geq 1$ se cumple la identidad

$$N_{ij}(\tau_{ij} + n) = 1 + N_{jj}(n).$$

Por lo tanto es suficiente demostrar la convergencia para $N_{jj}(n)/n$ pues

$$\begin{aligned}
\lim_{n\to\infty} \frac{N_{ij}(n)}{n} &= \lim_{n\to\infty} \frac{N_{ij}(\tau_{ij}+n)}{\tau_{ij}+n} \\
&= \lim_{n\to\infty} \frac{1+N_{jj}(n)}{\tau_{ij}+n} \\
&= \lim_{n\to\infty} \frac{N_{jj}(n)}{n} \quad \frac{n}{\tau_{ij}+n} \\
&= \lim_{n\to\infty} \frac{N_{jj}(n)}{n}
\end{aligned}$$

Sea $Y(k)$ la variable que registra el número de pasos que transcurren entre la visita $k-1$ y la visita $k$ que la cadena realiza al estado $j$. Sabemos que el tiempo medio de recurrencia es $E(Y(k)) = \mu_j$, para $j = 1, 2, \ldots$, y usando la propiedad fuerte de Markov puede demostrarse que las variables $Y(1), Y(2), \ldots$ son independientes. Se tienen entonces las siguientes estimaciones

$$\frac{Y(1) + \ldots + Y(N_{jj}(n))}{N_{jj}(n)} \leq \frac{n}{N_{jj}(n)} \leq \frac{Y(1) + \ldots + Y(N_{jj}(n)+1)}{N_{jj}(n)}$$

Por la recurrencia, $N_{jj}(n) \to \infty$, cuando $n \to \infty$, de modo que por la ley de los grandes números, los dos extremos de esta desigualdad convergen a $\mu_j$ casi seguramente. Por lo tanto,

$$\lim_{n\to\infty} \frac{N_{jj}(n)}{n} = \frac{1}{\mu_j} \quad c.s.$$

**Interpretación:** Para una cadena de Markov irreducible, el número $\pi_j = 1/\mu_j$ es el tiempo promedio que la cadena permanece en el estado $j$ a largo plazo.

Tomando esperanza en (2.4), por el teorema de convergencia dominada, y para una cadena irreducible, se cumple que

$$\begin{aligned}
E(\lim_{n\to\infty} \frac{N_{ij}(n)}{n}) &= \lim_{n\to\infty} \frac{1}{n} E(N_{ij}(n)) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) \\
&= \frac{1}{\mu_j}
\end{aligned}$$

En particular, cuando el tiempo medio de recurrencia $\mu_j$ es infinito, y aún más particularmente cuando el estado $j$ es transitorio, se tiene que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) = 0.$$

## 2.10. Recurrencia positiva y nula

Hemos visto que si una cadena de Markov inicia en un estado recurrente, entonces regresa a él una infinidad de veces con probabilidad uno. Sin embargo, esta recurrencia puede presentarse de dos formas: cuando el tiempo promedio de retorno es finito o cuando es infinito. Esto lleva a la definición de recurrencia positiva y recurrencia nula respectivamente. Consideremos entonces que $j$ es un estado recurrente. El tiempo de primera visita a este estado, a partir de cualquier otro estado $i$, es la variable aleatoria discreta $\tau_{ij} = \text{mín}\{n \geq 1 : X_n = j \mid X_0 = i\}$. Recordamos que cuando el tiempo de primera visita se refiere al mismo estado recurrente de inicio y de llegada $i$, se escribe simplemente como $\tau_i$ en lugar de $\tau_{ii}$. La esperanza de esta variable aleatoria es naturalmente el tiempo medio de recurrencia.

**Definición 2.14** *El tiempo medio de recurrencia de un estado recurrente $j$, a partir del estado $i$, se define como la esperanza de $\tau_{ij}$, y se denota por $\mu_{ij}$, es decir,*

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}(n).$$

Nuevamente cuando el tiempo medio de recurrencia se refiere al mismo estado recurrente de inicio y de llegada $i$, se escribe simplemente como $\mu_i$. Como hemos mencionado, esta esperanza puede ser finita o infinita, y ello lleva a la siguiente clasificación de estados recurrentes.

**Definición 2.15** *Se dice que un estado recurrente $i$ es*

**a)** *recurrente positivo si $\mu_i < \infty$.*

**b)** *recurrente nulo si $\mu_i = \infty$.*

Demostraremos a continuación que la recurrencia positiva y la recurrencia nula son propiedades de las clases de comunicación. Es decir, dos estados en una misma clase de comunicación recurrente, son ambos recurrentes positivos o recurrentes nulos.

**Proposición 2.14** *La (recurrencia positiva o nula es una propiedad de clase). Sea $i$ un estado recurrente. Entonces*

**a)** *si $i$ es recurrente positivo e $i \leftrightarrow j$, entonces $j$ es recurrente positivo.*

**b)** *si $i$ es recurrente nulo e $i \leftrightarrow j$, entonces $j$ es recurrente nulo.*

**Demostración**

Observe que es suficiente demostrar cualquiera de estas afirmaciones. Demostraremos la primera. Suponga que $i$ es un estado recurrente positivo, es decir, $i$ es recurrente y es tal que $\mu_i < \infty$. Como $i \leftrightarrow j$, se tiene que $j$ es también un estado recurrente. Además existen enteros no negativos $n$ y $m$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Entonces para cualquier entero natural $k$,

$$p_{jj}(n+m+k) \geq p_{ji}(m)p_{ii}(k)p_{ij}(n).$$

Sumando para $k = 1, \ldots, N$, y dividiendo entre $N$,

$$\frac{1}{N}\sum_{k=1}^{N} p_{jj}(n+m+k) \geq p_{ji}(m)\frac{1}{N}\sum_{k=1}^{N} p_{ii}(k)p_{ij}(n).$$

Haciendo N $\longrightarrow \infty$ se obtiene

$$\frac{1}{\mu_j} \geq p_{ji}(m) \frac{1}{\mu_i} p_{ij}(n) > 0.$$

Por lo tanto también $\mu_j$ es finito, es decir, $j$ es recurrente positivo.

**Proposición 2.15** *No existen estados recurrentes nulos en cadenas de Markov finitas.*

**Demostración**
Sea $j$ un estado recurrente y sea $C$ su clase de comunicación. La clase $C$ es cerrada y además es finita pues la cadena completa lo es. Demostraremos que $\mu_j < \infty$. Para cualquier $i \in C$, y $k$ natural,

$$\sum_{j \in C} p_{ij}(k) = 1.$$

Entonces

$$\frac{1}{n} \sum_{k=1}^{n} \sum_{j \in C} p_{ij}(k) = \sum_{j \in C} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) = 1.$$

Haciendo $n \longrightarrow \infty$, por el teorema ergódico aplicado a la clase cerrada $C$ se obtiene

$$\sum_{j \in C} \frac{1}{\mu_j} = 1.$$

Para que esta suma sea uno debe existir por lo menos un valor de $j$ en $C$ tal que $\mu_j < \infty$, pues de lo contrario cada sumando sería cero y la suma total no podría ser uno. Por lo tanto existe un estado $j$ que es recurrente positivo. Dado que la recurrencia positiva es una propiedad de clase, todos los elementos de $C$ son recurrentes positivos.

Observe que en particular, todos los estados de una cadena finita e irreducible son recurrentes positivos.

## 2.11. Evolución de distribuciones

Una matriz estocástica establece una dinámica en el conjunto de las distribuciones de probabilidad definidas sobre el espacio de estados de la correspondiente cadena de Markov. Para explicar la situación de manera simple consideremos un espacio de estados finito $\{0, 1, \ldots, N\}$, y una distribución de probabilidad inicial $\pi(0) = (\pi_0(0), \pi_1(0), \ldots, \pi_N(0))$. Después de transcurrida la primera unidad de tiempo, la cadena se encuentra en cualquiera de sus posibles estados de acuerdo a la distribución $\pi(1) = (\pi_0(1), \pi_1(1), \ldots, \pi_N(1))$ en donde la $j$-ésima entranda de este vector es

$$\begin{aligned}
\pi_j(1) &= P(X_1 = j) \\
&= \sum_{i=0}^{N} P(X_1 = j \mid X_0 = i) P(X_0 = i) \\
&= \pi_0(0) p_{0j}, \pi_1(0) p_{1j}, \ldots, \pi_N(0) p_{Nj}. \\
&= (\pi(0) P)_j.
\end{aligned}$$

Es decir, $\pi(1)$ se obtiene a partir de $\pi(0)$ y de la matriz de probabilidades de transición $P$ a través de la fórmula $\pi(1) = \pi(0)P$,

$$(\pi_0(1), \ldots, \pi_N(1)) = (\pi_0(0), \ldots, \pi_N(0)) \begin{pmatrix} p_{00} & \cdots & p_{0N} \\ \vdots & & \vdots \\ p_{N0} & \cdots & p_{NN} \end{pmatrix}$$

A su vez la distribución $\pi(1)$ se transforma en $\pi(2)$ a través de la ecuación $\pi(2) = \pi(1)P = \pi(0)P^2$, y así sucesivamente. En general, $\pi(n+1) = \pi(n)P = \pi(0)P^{n+1}$. De esta forma se obtiene una sucesión infinita de distribuciones de probabilidad $\pi(0), \pi(1), \pi(2), \ldots,$ en donde cada una de ellas, excepto la primera, es obtenida de la anterior multiplicidad por la derecha por la matriz de probabilidades de transición en un paso.

## 2.12. Distribuciones estacionarias

**Definición 2.16** *Una distribución de probabilidad $\pi = (\pi_0, \pi_1, \ldots)$ es estacionaria o invariante para una cadena de Markov con probabilidades de transición $p_{ij}$ si*

$$\pi_j = \sum_i \pi_i p_{ij}.$$

En términos matriciales $\pi$ es estacionaria si $\pi = \pi P$. Esta identidad tiene como consecuencia el hecho de que para cualquier número natural $n$, se cumpla que $\pi = \pi P^n$, es decir, $\pi$ es también una distribución estacionaria para la matriz $P^n$. Esto significa que si la variable aleatoria inicial $X_0$ tiene esa distribución $\pi$, entonces la distribución de $X_n$ también es $\pi$ pues $P(X_n = j) = \sum_i \pi_i p_{ij}(n) = \pi_j$, es decir, esta distribución no cambia con el paso del tiempo y por ello es que se le llama estacionaria o invariante.

**Proposición 2.16 (Soporte de una distribución estacionaria)** .

*Sea $\pi$ una distribución estacionaria. Si $j$ es un estado transitorio o recurrente nulo, entonces $\pi_j = 0$.*

**Demostración:**

Usaremos el hecho de que si $j$ es un estado transitorio o recurrente nulo, entonces para cualquier estado $i$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k) = 0.$$

Como $\pi$ es una distribución estacionaria,

$$\begin{aligned}
\pi_j &= \sum_i \pi_i p_{ij} \\
&= \sum_i \pi_i p_{ij}(k) \\
&= \frac{1}{n} \sum_{k=1}^{n} \sum_i \pi_i p_{ij}(k) \\
&= \sum_i \pi_i (\frac{1}{n} \sum_{k=1}^{n} p_{ij}(k))
\end{aligned}$$

Tomando el límite cuando $n \to \infty$, por el teorema de convergencia dominada, se obtiene

$$\pi_j = \sum_i \pi_i (\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k)) = 0.$$

En particular, si $\pi_j > 0$, entonces $j$ es un estado recurrente positivo. Esto es una consecuencia inmediata del resultado anterior y para ello puede usarse un argumento por contradicción.

**Proposición 2.17** *(Existencia y unicidad de la distribución estacionaria). Toda cadena de Markov que es irreducible y recurrente positiva tiene una única distribución estacionaria dada por*

$$\pi_j = \frac{1}{\mu_j} > 0.$$

*En particular, toda cadena finita e irreducible tiene una única distribución estacionaria.*

**Demostración**
Sea $\pi_j = 1/\mu_j$. Demostraremos que

1. $\pi_j = \sum_i \pi_i p_{ij}$.
2. $\sum_j \pi_j = 1$.
3. $\pi_j$ es única.

Como la cadena es irreducible y recurrente positiva, se tiene que $\mu_j < \infty$, para cualquier estado $j$. Por lo tanto el cociente $1/\mu_j$ es estrictamente positivo. Por el teorema ergódico para cadenas de Markov se tiene que

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ij}(m) = \frac{1}{\mu_j}.$$

1. Estacionariedad. Para cualquier natural $N$,

$$\begin{aligned}
\sum_{i=0}^{N} \pi_i p_{ij} &= \sum_{i=0}^{N} (\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ki}(m)) p_{ij} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sum_{i=0}^{N} p_{ki}(m) p_{ij} \\
&\leq \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{kj}(m+1) \\
&= \pi_j
\end{aligned}$$

Haciendo $N \longrightarrow \infty$,

$$\sum_i \pi_i p_{ij} \leq \pi_j. \tag{2.5}$$

Suponga que para algún valor de $j$ la desigualdad anterior es estricta. Sumando sobre todos los valores $j$, por el teorema de Fubini,

$$\begin{aligned}
\sum_j \pi_j > \sum_j \sum_i \pi_i p_{ij} &= \sum_i \pi_i \sum_j p_{ij} \\
&= \sum_i \pi_i.
\end{aligned}$$

Lo cual es una contradicción. Por lo tanto (2.5) es una igualdad. Esto demuestra que $\pi$ es estacionaria, y para cualquier $m$ natural,

$$\pi_j = \sum_i \pi_i p_{ij}(m). \tag{2.6}$$

2. Distribución de probabilidad. Para cualquier natural $N$,

$$\begin{aligned}
\sum_{j=0}^{N} \pi_j &= \sum_{j=0}^{N} (\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} p_{ij}(k)) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{j=0}^{N} p_{ij}(k) \\
&\leq \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} 1 \\
&= 1.
\end{aligned}$$

Por otra parte, por (2.6), para cualquier $n$ natural,

$$\pi_j = \sum_i \pi_i (\frac{1}{n} \sum_{m=1}^{n} p_{ij}(m)).$$

Haciendo $n \to \infty$, por el teorema de convergencia dominada de Lebesgue, y usando el hecho de que $\sum_j \pi_j \leq 1$,

$$\begin{aligned}
\pi_j &= \sum_{i=0}^{\infty} \pi_i (\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ij}(m)) \\
&= \sum_{i=0}^{\infty} \pi_i \pi j
\end{aligned}$$

Dado que $\pi_j > 0$, se obtiene $\sum_i \pi_i = 1$.

3. Unicidad. Sera $\pi$ y $\pi^{'}$ dos distribuciones estacionarias para la matriz $P$. Entonces

$$\begin{aligned} \pi_j^{'} &= \sum_i \pi_i^{'} p_{ij}(m) \\ &= \sum_i \pi_i^{'} (\frac{1}{n} \sum_{m=1}^{n} p_{ij}(m)) \\ &\geq \sum_{i=0}^{N} \pi_i^{'} (\frac{1}{n} \sum_{m=1}^{n} p_{ij}(m)). \end{aligned}$$

Haciendo $n \to \infty$,

$$\pi_j^{'} \geq \sum_{i=0}^{N} \pi_i^{'} \pi_j.$$

Ahora hacemos $N \to \infty$ para obtener

$$\pi_j^{'} \geq \pi_j. \tag{2.7}$$

Si esta desigualdad fuera estricta para algún valor de $j$, entonces sumando sobre todos los valores de $j$ se obtiene $1 = \sum_j \pi_j^{'} > \sum_j \pi_j = 1$, lo cual es una contradicción. Por lo tanto (2.7) es una igualdad para cada valor de $j$, y ello demuestra la unicidad.

## 2.13. Distribuciones límite

Como hemos mencionado antes, toda matriz de probabilidades de transición $P$ determina una sucesión de distribuciones de probabilidad $\pi^0, \pi^1, \ldots$ sobre el espacio de estados en donde

$$\pi^{n+1} = \pi^n P = \pi^0 P^{n+1}. \tag{2.8}$$

Bajo ciertas condiciones tal sucesión es convergente a una distribución límite $\pi$. Imaginemos por ahora que tal es el caso y examinaremos algunas propiedades de esta distribución límite. Por (2.8) se tiene que

$$\pi = \lim_{n\to\infty} \pi^n = \pi^0 \lim_{n\to\infty} P^{n+1} = \pi^1 \lim_{n\to\infty} P^n.$$

Esta igualdades revelan varias cosas.

- Primero, la distribución límite no depende de la distribución inicial pues el lado derecho de estas ecuaciones debe ser el mismo.
- Segundo, la distribución límite está dada por el límite de las potencias de $P$ pues si se toma como distribución inicial aquella concentrada en el $i$-ésimo estado, entonces el $j$-ésimo elemento de la distribución límite es $\pi_j = \lim_{n\to\infty} p_{ij}(n)$.
- Tercero, el límite de las potencias de $P$ es una matriz con todos sus renglones idénticos, siendo este reglón la distribución límite.

- Por último, a partir de (2.8) se obtiene que $\lim_{n\to\infty} \pi^{n+1} = \lim_{n\to\infty} \pi^n P$, esto es, $\pi = \pi P$, es decir, la distribución límite es una distribución estacionaria, suponiendo un espacio de estados finito. En esta sección se establecen condiciones para obtener rigurosamente estos resultados.

**Proposición 2.18** *Considere una cadena de Markov con probabilidades de transición $p_{ij}$ tales que los 'limites $\pi_j = \lim_{n\to\infty} p_{ij}(n)$ existen para cada $j$, y no depende del estado $i$. Entonces*

*1.* $\sum_j \pi_j \leq 1$.

*2.* $\pi_j = \sum_i \pi_i p_{ij}(n)$.

*Cuando el espacio de estados es finito se cumple la igualdad en el primer resultado obteniéndose una distribución de probabilidad verdadera.*

**Demostración:**

Supongamos primero el caso cuando el espacio de estados es el conjunto finito $\{0, 1, \ldots, N\}$. Entonces la primera afirmación se cumple con igualdad pues

$$\begin{aligned}\sum_{j=0}^{N} \pi_j &= \sum_{j=0}^{N} \lim_{n\to\infty} p_{ij}(n)\\ &= \lim_{n\to\infty} \sum_{j=0}^{N} p_{ij}(n) = 1.\end{aligned}$$

Para la segunda afirmación se tiene que para cualquier $n \geq 1$,

$$\begin{aligned}\sum_{i=0}^{N} \pi_j p_{ij}(n) &= \sum_{i=0}^{N} \lim_{m\to\infty} p_{ki}(m) p_{ij}(n)\\ &= \lim_{m\to\infty} \sum_{i=0}^{N} p_{ki}(m) p_{ij}(n)\\ &= \lim_{m\to\infty} p_{kj}(m+n)\\ &= \pi_j.\end{aligned}$$

Suponga ahora que el espacio de estados es infinito. En este caso no es posible garantizar la validez del intercambio de límite y suma efectuado antes. Para la primera afirmación se tiene que para cualquier número natural $N \geq 1$,

$$\begin{aligned}\sum_{j=0}^{N} \pi_j &= \sum_{j=0}^{N} \lim_{n\to\infty} p_{ij}(n)\\ &= \lim_{n\to\infty} \sum_{j=0}^{N} p_{ij}(n) \leq \lim_{n\to\infty} 1\\ &= 1\end{aligned}$$

Haciendo $N \longrightarrow \infty$ se obtiene el resultado buscado. Para la segunda afirmación, nuevamente para cualquier número natural $N \geq 1$,

$$\begin{aligned}
\sum_{i=0}^{N} \pi_i p_{ij} &= \sum_{i=0}^{N} \lim_{n\to\infty} p_{ki}(n) p_{ij} \\
&= \lim_{n\to\infty} \sum_{i=0}^{N} p_{ki}(n) p_{ij} \\
&\leq \lim_{n\to\infty} \sum_{i=0}^{N} p_{kj}(n+1) \\
&= \pi_j.
\end{aligned}$$

Si $\pi_j = 0$ para cualquier $j$, entonces estas desigualdades se convierten en identidades como dice el enunciado. Suponga que $\sum_j \pi_j > 0$. Demostraremos que las desigualdades estrictas no son posibles. Suponga que para algún estado $j$, $\sum_i \pi_i p_{ij} < \pi_j$. Entonces

$$\sum_j \pi_j > \sum_{i,j} \pi_i p_{ij} = \sum_i \pi_i \sum_j p_{ij} = \sum_i \pi_i.$$

Esto es una contradicción, por lo tanto la igualdad en el párrafo anterior se cumple.

Ahora se establecen condiciones suficientes para que exista el límite de las probabilidades de transición en $n$ pasos. Este resultado es una especie de recíproco del resultado anterior pues supone la existencia de una distribución estacionaria para concluir que los límites de las probabilidades existen.

**Teorema 2.2 (de convergencia.)** *Considere una cadena de Markov que es*

**a)** *irreducible,*

**b)** *aperódica, y*

**c)** *con distribución estacionaria $\pi$.*

*Entonces para cualquier estados $i$ y $j$, $\lim_{n\to\infty} p_{ij}(n) = \pi_j$.*

**Demostración:**

El método de esta demostración se conoce como técnica de acople.

Sea $\{Y_n\}$ una cadena de Markov independiente de la original $\{X_n\}$, pero con la misma matriz de probabilidades de transición. Entonces el proceso $\{Z_n\}$ definido por $Z_n = (X_n, Y_n)$ es una cadena de Markov con probabilidades de transición.

$$P(Z_{n+1} = (x_{n+1}, y_{n+1}) \mid Z_n = (x_n, y_n)) = p_{x_n, x_{n+1}} p_{y_n, y_{n+1}},$$

y puede fácilmente comprobarse que tiene distribución estacionaria $\pi_{x_n, y_n}$. Puede además verificarse que la cadena $\{Z_n\}$ es recurrente positiva. Además es irreducible pues como $\{X_n\}$ y $\{Y_n\}$ son aperiódicas, existe un número natural $N$ tal que $p_{ij}(n) p_{kl}(n) > 0$, para toda $n > N$. Por lo tanto $p_{(i,k)}(_{j,l})(n) > 0$.

Sea $j$ un estado cualquiera de la cadena original. Defina el primer momento en el que la cadena $\{Z_n\}$ visita al estado $(j,j)$ como $\tau_j = \min\{n \geq 1 : Z_n = (j,j)\}$.

Sea además $\tau_j = \text{mín}\{n \geq 1 : X_n = Y_n\}$. Este es el primer momento de acople de las dos cadenas. Como $\{Z_n\}$ es recurrente, $P(\tau < \infty) = 1$. Además $\tau \geq \tau_j$. Por la propiedad de Markov

$$\begin{aligned}
P(X_n = x, \tau \leq n) &= \sum_j \sum_{r=1}^{n} P(X_n = x, X_r = j, \tau = r) \\
&= \sum_j \sum_{n}^{n} P(X_n = x, \mid X_r = j, \tau = r) P(X_r = j, \tau = r) \\
&= \sum_j \sum_{n}^{n} P(Y_n = x \mid Y_r = j, \tau = r) P(Y_r = j, \tau = r) \\
&= \sum_j \sum_{n}^{n} P(Y_n = x \mid Y_r = j) P(Y_r = j, \tau = r) \\
&= P(Y_n = x, \tau \leq n),
\end{aligned}$$

es decir, sobre el evento $(\tau \leq n)$, las variables $\{X_n\}$ y $\{Y_n\}$ tienen la misma distribución de probabilidad. Por otro lado

$$\begin{aligned}
P(X_n = j) &= P(X_n = j, \tau \leq n) + P(X_n = j, \tau > n) \\
&= P(Y_n = j, \tau \leq n) + P(X_n = j, \tau > n) \\
&\leq P(Y_n = j) + P(\tau > n).
\end{aligned}$$

De manera análoga, $P(Y_n = j) \leq P(X_n = j) + P(\tau > n)$. Por lo tanto,

$$|P(X_n = j) - P(Y_n = j)| \leq P(\tau > n) \longrightarrow 0, \tag{2.9}$$

cuando $n \longrightarrow \infty$. Si ahora se toma $X_0 = i$ con probabilidad uno, y $Y_0$ con la distribución estacionaria $\pi$, entonces $P(X_n = j) = P(X_n = j \mid X_0 = i)P(X_0 = i) = p_{ij}(n)P(X_0 = i) = p_{ij}(n)$, y $P(Y_n = j) = \sum_i P(Y_n = j \mid Y_0 = i)\pi_i = \sum_i \pi_i p_{ij}(n) = \pi_j$.Por lo tanto (2.9) establece que $\mid p_{ij}(n) - \pi_j \mid \longrightarrow 0$.

El siguiente resultado establece condiciones suficientes para la existencia de límite de las probabilidades de transición, asegurando que se trata de una distribución de probabilidad.

**Teorema 2.3 de convergencia** *considere una cadena de Markov que es*

**a)** *irreducible,*

**b)** *recurrente positiva, y*

**c)** *aperiódica.*

*Entonces las probabilidades límite* $\pi_j = \lim_{n\to\infty} p_{ij}(n) = \dfrac{1}{\mu_j}$ *existen, y constituyen la única solución al sistema de ecuaciones*

$$\pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij}, \tag{2.10}$$

*sujeto a las condiciones* $\pi_j \geq 0$, *y* $\displaystyle\sum_{j=0}^{\infty} \pi_j = 1$.

**Demostración**

Como la cadena es irreducible y recurrente positiva, tiene una única distribución estacionaria dada por $\pi_j = 1/\mu_j$. Es decir, es la única solución al sistema de ecuaciones $\pi = \pi P$, con $\pi_j \geq 0$ y $\sum_j \pi_j = 1$. Además, por la aperiódicidad, $p_{ij}(n) \longrightarrow \pi_j$.

## 2.14. Cadenas regulares

Las cadenas de Markov regulares son cadenas finitas que cumplen la propiedad de que a partir de un cierto momento, con probabilidad positiva se puede pasar de un estado a otro cualquiera en un paso. Demostraremos que para este tipo de cadenas existe siempre la distribución límite.

**Definición 2.17** *Se dice que una cadena finita o su matriz de probabilidades de transición es regular si existe un entero natural* $n$ *tal que* $p_{ij}(n) > 0$, *para cualesquiera estados* $i$ *y* $j$.

En palabras, una cadena de Markov finita es regular si alguna potencia de su matriz de probabilidades de transición tiene todas sus entradas estrictamente positivas. Otra forma de definir a una cadena regular es a través del siguiente resultado.

**Proposición 2.19** *Una matriz estocástica es regular si, y sólo si, es finita, irreducible y aperiódica.*

**Demostración**

Si la matriz es regular, entonces claramente es finita, irreducible y aperiódica. Recíprocamente, por la irreducibilidad, para cualesquiera dos estados $i$ y $j$, existe un entero $m$ tal que $p_{ij}(m) > 0$. Entonces existe un entero $N$ tal que $p_{ij}(m + nd(j)) > 0$, para cada $n \geq N$. Como $d(j) = 1$ y siendo la matriz finita, esto implica la regularidad.

Para este tipo particular de cadenas finitas se conocen los siguientes resultados acerca de su comportamiento límite.

**Proposición 2.20** *Toda cadena finita que además es*

1. *regular, tiene como distribución límite la única solución no negativa del sistema (2.10).*
2. *regular y doblemente estocástica, tiene como distribución límite la distribución uniforme.*
3. *irreducible, aperiódica y doblemente estocástica, tiene como distribución límite la distribución uniforme.*

**Demostración**

1. Como la cadena es regular, es irreducible, aperiódica, y es recurrente positiva por ser finita. Por el teorema anterior la distribución límite existe y está dada por el sistema de ecuaciones (2.10).

2. Como la cadena es regular, es aperiódica, irreducible y recurrente positiva por ser finita. Entonces la distribución límite existe. Por la hipótesis de doble estocasticidad y suponiendo que el espacio de estados es $\{0, 1, \ldots, N\}$, se tiene que $\sum_{i=0}^{N} p_{ij}(n) = 1$. Tomando el límite cuando $n$ tiende a infinito se obtiene $\sum_{i=0}^{N} \pi_j = 1$. Por lo tanto $\pi_j = 1/(N+1)$.

3. Este resultado es idéntico al anterior pues hemos demostrado que la regularidad es equivalente a la finitud, irreducibilidad y aperiocidad conjuntamente.

## 2.15. Cadenas reversibles

Sea $\{X_n\}$ una cadena de Markov con probabilidades de transición $p_{ij}$. Sea $m \geq 1$ un entero fijo y defina un nuevo proceso $Y_n = X_{m-n}$ para $n = 0, \ldots, m$, es decir, $Y_n$ es la cadena original pero vista en sentido inverso en el tiempo, ahora del tiempo $m$ al tiempo 0. Este nuevo proceso resulta también ser una cadena Markov pues cumple el criterio de independencia entre pasado y futuro cuando se conoce el presente, las nociones de pasado y futuro se intercambian debido al cambio en el sentido del tiempo. En efecto, para $1 \leq r < n \leq m$, considere la probabilidad condicional

$$P(y_1, \ldots, y_{r-1}, y_{r+1}, \ldots, y_n \mid y_r).$$

en términos del proceso $\{X_n\}$, esta probabilidad es

$$P(X_{m-1} = y_1, \ldots, X_{m-r+1} = y_{r-1},\ \ X_{m-r-1} = y_{r+1}, \ldots, X_{m-n} = y_n \mid X_{m-r} = y_r).$$

Por la propiedad de Markov del proceso $\{X_n\}$, esta probabilidad es el producto

$$P(X_{m-1} = y_1, \ldots, X_{m-r+1} = y_{r-1} \mid X_{m-r} = y_r)$$
$$P(X_{m-r-1} = y_{r+1}, \ldots, X_{m-n} = y_n \mid X_{m-r} = y_r),$$

que en términos del proceso $\{Y_n\}$ es la propiedad de Markov para este proceso

$$P(y_1, \ldots, y_{r-1} \mid y_r) P(y_{r+1}, \ldots, y_n \mid y_r).$$

Sin embargo las probabilidades de transición del nuevo proceso no son homogéneas pues para $0 \leq n < m$,

$$\begin{aligned} P(Y_{n+1} = j \mid Y_n = i) &= \frac{P(X_{m-n} = i, X_{m-n-1} = j)}{P(X_{m-n} = i)} \\ &= P(X_{m-n} = i \mid X_{m-n-1} = j) \frac{P(X_{m-n-1} = j)}{P(X_{m-n} = i)} \\ &= p_{ij} \frac{P(Y_{n+1} = j)}{P(Y_n = i)} \end{aligned}$$

es decir, estas probabilidades dependen de $n$ a través del cociente $P(Y_{n+1} = j)/P(Y_n = i)$. Tal dependencia desaparece cuando se toma como hipótesis la existencia de una distribución estacionaria $\pi$ para $\{X_n\}$, pues en tal caso la igualdad anterior se reduce a

$$P(Y_{n+1} = j \mid Y_n = i) = p_{ij} \frac{\pi_j}{\pi_i}.$$

Bajo tal hipótesis las probabilidades de transición de la nueva cadena son ahora estacionarias. Si adicionalmente se pide que las probabilidades de transición son las mismas para ambas cadenas, entonces de la ecuación anterior se obtiene que debe sastifacerse la ecuación $p_{ij} = p_{ji}\frac{\pi_j}{\pi_i}$, es decir $\pi_i p_{ij} = \pi_j p_{ij}$. Esto lleva a la siguiente definición de reversibilidad la cual añade la hipótesis de irreducibilidad.

**Definición 2.18** *Se dice que una cadena de Markov irreducible con probabilidades de transición $p_{ij}$ y con distribución estacionaria $\pi$ es reversible en el tiempo si para cualquier estado $i$ y $j$,*

$$\pi_i p_{ij} = \pi_j p_{ji}. \tag{2.11}$$

A la ecuación (2.11) se llama ecuación de balance detallado.

**Proposición 2.21** *Considere una cadena irreducible para la cual existe una distribución $\pi$ que cumple (2.11). Entonces $\pi$ es una distribución estacionaria y la cadena es reversible.*

**Demostración**

Si $\pi$ cumple (2.11), entonces es una distribución estacionaria pues $\sum_i \pi_i p_{ij} = \sum_i \pi_j p_{ji} = \pi_j$. Además la cadena es reversible por definición.

# Capítulo 3

# Modelo de Markov Aplicado a un Problema de Epidemiología

El análisis de procesos estocásticos permite predecir el estado en que se encontrará el proceso en el futuro a partir de la información disponible sobre su pasado. La complejidad de estas predicciones depende en gran medida del tipo de proceso; sin embargo, las características de algunos de ellos facilitan su análisis e interpretación. Un caso particular bien estudiado es el proceso de Markov, proceso estocástico cuyo estado futuro dependerá solo del estado en que se encuentre en el presente, pero no de su historia pasada.

Desde que, en 1906, el matemático ruso Andrei Andreyevich Markov definió por primera vez este tipo de procesos, sus aplicaciones en la investigación biomédica han sido múltiples, tanto en la experimentación animal como en los estudios en humanos.

Dependiendo de las condiciones particulares de cada estudio, la metodología para el análisis de datos difiere. Cuando los valores que puede tomar el proceso son discretos suele hablarse de cadenas de Markov, mientras que el término proceso de Markov suele reservarse para procesos con espacio de estados continuos. Aunque la terminología no está estandarizada, la clasificación general de los modelos markovianos puede esquematizarse como se muestra en el cuadro (3.2).

En este capítulo utilizaremos un modelo markoviano para modelar un problema en el campo de la epidemiología y las ciencias de la salud desde un punto de vista eminentemente práctico, prestando especial atención a las principales técnicas de análisis, la interpretación de resultados, la utilidad de los modelos.

TABLA I
CLASIFICACIÓN DE LOS MODELOS DE MARKOV

| | | Espacio de Estados | |
|---|---|---|---|
| | | Discreto | Continuo |
| Tiempo de observación | Discreto | Cadena de Markov discreta | Proceso de Markov discreto |
| | Continuo | Cadena de Markov continua | Proceso de difusión |

Cuadro 3.1: tabla I

Desde comienzos de los años 80, el SIDA se ha convertido en una de las mayores pandemias de nuestra época. En la actualidad, España es uno de los países del oeste de Europa más afectados por el virus, por lo que el conocimiento de la evolución futura de la epidemia es un objetivo prioritario para establecer políticas sanitarias adecuadas. A continuación mostraremos como llevar a cabo un análisis de la epidemia de VIH-SIDA utilizando una cadena de Markov homogénea con tiempos de observación anuales.

**a)** Definición de los estados del proceso y los mecanismos de transición entre ellos. Uno de los modelos más sencillos para el estudio de la epidemia de SIDA establece que cada sujeto de la población puede estar exclusivamente en uno de los siguientes estados:

**Estado 1(S):** susceptible,
**Estado 2(VIH):** infectado por VIH,
**Estado 3(SIDA):** con SIDA,
**Estado 4(M):** Muerto por la enfermedad.

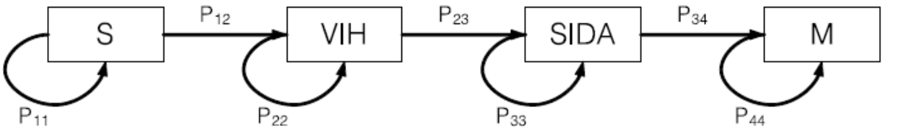


Figura 3.1: Estados en la epidemia de VIH-SIDA

El sujeto susceptible no está infectado (estado 1, -S-). Tras el contacto con un infectado desarrollará la infección y podrá contagiar a otros individuos de la población (estado 2, -VIH-). Cuando el sujeto infectado desarrolla la enfermedad se convertirá en un caso de SIDA (estado 3,-SIDA-), pudiendo morir por causa de la enfermedad (estado 4, -M-). La figura (3.1) muestra las posibles transiciones entre estados junto con las probabilidades de transición teóricas.

Los estados S, VIH y SIDA se denominan transitorios puesto que el individuo nunca volverá a ellos una vez que los abandone para alcanzar otro estado de la enfermedad. El estado M, en cambio, es absorbente ya que los sujetos que lo alcanzan permanecen en él si posibilidad de cambiar a otro estado.

Este modelo epidemiológico es una variante de los modelos SIR Reed-Frost utilizados para estudiar la dinámica de procesos infecciosos en poblaciones.

**b)** Selección del modelo markoviano más apropiado para el estudio del proceso. La observación de los sujetos se realiza de forma anual, por lo que tanto los estados como los instantes de tiempo del proceso son discretos. Además, el estado futuro en que se encuentre un individuo dependerá solo de su estado actual, pero no de su historia pasada. por ello, el proceso estocástico es una cadena de Markov discreta.

Si las politicas de prevención de VIH y las pautas de tratamiento no varían, tanto la incidencia de VIH-SIDA como la tasa de mortalidad, entre los casos de SIDA permanecerán constantes; por lo que las probabilidades de transición no cambiarán en el tiempo. Bajo esta hipótesis, la cadena de Markov discreta será además homogénea, siendo éste el modelo markoviano más apropiado para estudiar la evolución de la epidemia.

**c)** Determinación de las probabilidades de transición entre estados. En epidemias humanas, la matriz de transición suele obtenerse a partir de los datos de población, incidencia, prevalencia y mortalidad publicados en los registros oficiales. Para este trabajo se considerán las cifras registradas sobre la epidemia de VIH-SIDA en España correspondientes al año 2004.

Los datos publicados muestran que el número estimado de nuevos casos fue de 2750 en una población de 43197684 habitantes, lo que supone una tasa de contagio de 64 casos por millón de habitantes (CNE, 2005). Como consecuencia, $p_{12} = 0,000064$. Según el modelo establecido, un sujeto en estado susceptible no puede pasar directamente al estado SIDA ni al estado M, por tanto. La suma de las probabilidades de una fila de la matriz de transición ha de ser 1, por lo que $p_{11} = 1 - p_{12} - p_{13} - p_{14} = 0,000036$, donde $p_{13} = p_{14} = 0$.

La segunda fila de la matriz corresponde a las transiciones que pueden realizarse desde el estado VIH. Así, un individuo infectado no volverá a ser susceptible, por lo que $p_{21} = 0$. La incidencia de SIDA en 2004 fue de 804 nuevos casos (EuroHIV, 2005) en una población prevalente estimada de 125000 individuos con VIH (Cañas et al.., 2003), por lo que $p_{23} = 0,006432$. Siguiendo el modelo propuesto, un sujeto VIH+ no puede pasar directamente al estado M, por tanto $p_{24} = 0$. Además, puesto que los valores de la fila han de sumar 1, se tendrá $p_{22} = 1 - p_{21} - p_{23} - p_{24} = 0,993568$.

Los sujetos que han desarrollado SIDA no retornaran al estado S o VIH, por tanto $p_{31} = p_{32} = 0$. En 2004 se produjeron 542 muertes entre un total de 6609 casos de SIDA (EuroHIV, 2005), lo que supone $p_{34} = 0,0082$. Por último, $p_{33} = 1 - p_{31} - p_{32} - p_{34} = 0,917991$.

Puesto que M es un estado absorvente y los sujetos permanecen en él, se tiene $p_{44} = 1$ y $p_{41} = p_{42} = p_{43} = 0$.

La matriz de transición para la epidemia de VIH-SIDA en España será por tanto

$$p = \begin{pmatrix} 0,000036 & 0,000064 & 0 & 0 \\ 0 & 0,993568 & 0,006432 & 0 \\ 0 & 0 & 0,917991 & 0,0082 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

**d)** Análisis de la evolución de la epidemia en el tiempo. La matriz de transición contiene las probabilidades de paso de un estado a otro en un año, es decir, la probabilidad de que un sujeto susceptible esté infectado al año siguiente es 0,000064. Sin embargo, para conocer la evolución de la epidemia es necesario calcular la probabilidad de que un sujeto susceptible actualmente pueda estar infectado dos años después. Las ecuaciones de Chapman-Kolmogorov permiten calcular estas probabilidades multiplicando la matriz de transición por ella misma y obteniendo la matriz de transición en dos tiempos, es decir

$$p^{(2)} = P \times P = \begin{pmatrix} 0,999872 & 0,000128 & 0,0000004 & 0 \\ 0 & 0,987177 & 0,012295 & 0,000053 \\ 0 & 0 & 0,842708 & 0,015728 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Así, la probabilidad de que un sujeto susceptible en la actualidad esté infectado dos años después es 0,000128 o, de forma equivalente, dentro de dos años habrá 128 nuevos infectados por millón de habitantes. Análogamente, es posible calcular la probabilidad de que un individuo que actualmente se encuentra en un estado esté en otro determinado dentro de tres años, a partir de la matriz de transición en tres tiempos $P^{(3)} = P \times P$. En general, la matriz de transición en $n$ tiempos no es más que un producto sucesivo de matrices dado por $P^{(n)} = P_{(n)} \times P$, con $n \geq 2$ un número entero.

El curso que seguirá la evolución de la epidemia en el tiempo dependerá de las condiciones iniciales de la población estudiada con respecto al virus, es decir, de la distribución actual de los individuos en cada uno de los estados de la enfermedad. En 2004 había en España 43065533 de sujetos susceptibles, 125000 personas vivas infectadas con VIH, 6609 casos de SIDA y 542 muertes causadas por la enfermedad. Estas cifras conforman el vector de casos prevalente iniciales que constituyen el punto de partida para estudiar la evolución futura de la epidemia. Este vector está representado por $v = (43065533, 125000, 6609, 542)$ y la suma de todos sus valores ha de coincidir con la población total estudiada, dada por 43,197.684 habitantes. El número de sujetos que habrá en cada uno de los estados de la enfermedad dentro de $n$ años viene dado por el producto del vector de prevalencias iniciales y la matriz de transición en $n$ tiempos, es decir $v \times P^{(n)}$ [1].

En el estudio de epidemias poblacionales, la manera más habitual de expresar el vector de prevalencias iniciales es como número de casos por millón de habitantes, es decir, $v = (996940, 2894, 153, 13)$. La Figura 2 muestra la evolución de la epidemia tomando este vector como punto de partida y calculando las tasas por millón para cada uno de los estados de la enfermedad hasta 2024, veinte años despues del inicio del estudio, Para cada año, esta tasa viene dada por $v \times P^{(n)}$ con $n = 1, ..., 20$.

Si la incidencia y mortalidad de VIH-SIDA no se modifica los casos de VIH Y SIDA seguirán una tendencia creciente en España. En 2024 el número de infectados por VIH habrá ascendido a 3744 por millón de habitantes, el número de enfermos de SIDA será de 248 por millón de habitantes y las muertes por SIDA llegarán a ser de 46 por millón de habitantes. Estas cifras significan que en 20 años el número de infecciones se habrá incrementado en un 29 por ciento, los casos de SIDA aumentaran un 62 por ciento y las muertes por la enfermedad serán un 154 por ciento superior.

Modificando el vector de prevalencias iniciales se pueden obtener diferentes simulaciones de la evolución hipotética de la epidemia en España. Conocer qué ocurríria en el futuro si la situación actual fuese diferente constituye una importante fuente de información para la toma de decisiones, siendo ésta una de las principales aportaciones de los modelos markovianos a la investigación biomédica.

| PREDICCIONES PARA EL 2024 | | | |
|---|---|---|---|
| AÑO | VIH | SIDA | MUERTE |
| 2024 | 3744 | 248 | 46 |

Cuadro 3.2: Cifras por millón de habitantes

# Conclusión

En este trabajo se analizo la epidemia de VIH-SIDA en España durante el año 2004 utilizando una Cadena de Markov homogénea con tiempos de observación anuales. Además se estudió:

- La teoria de Procesos Estocasticos y sus fundamentos.
- Las principales propiedades de las Cadenas de Markov en tiempo discreto.

Si la incidencia y mortalidad de VIH-SIDA no se modifica los casos de VIH Y SIDA seguirán una tendencia creciente en España. En 2024 el número de infectados por VIH habrá ascendido a 3744 por millón de habitantes, el número de enfermos de SIDA será de 248 por millón de habitantes y las muertes por SIDA llegarán a ser de 46 por millón de habitantes. Estas cifras significan que en 20 años el número de infecciones se habrá incrementado en un 29 por ciento, los casos de SIDA aumentaran un 62 por ciento y las muertes por la enfermedad serán un 154 por ciento superior.

Se recomienda en futuros trabajos solicitar al área de epidemiología del Hospital Central Universitario Antonio María Pineda datos sobre la epidemia de VIH-SIDA para hacer un ánalisis similar con datos pertenecientes a la Región Centro Occidental.

# Bibliografía


[1] Ocaña-Riola R (2009) *Modelos de Markov aplicados a la investigación en ciencias de la salud*. Universidad de Barcelona. España

[2] Athreya K, Lahiri S (2006) *Measure Theory and Probability Theory*. Springer. Iowa State University.

[3] Rincón L (2011) *Introducción a los PROCESOS ESTOCÁSTICOS*. Facultad de Ciencias UNAM. México.