



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**“POLINOMIOS ORTOGONALES Y
PROCESOS ESTOCÁSTICOS”**

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

MATEMÁTICO

P R E S E N T A :

JOSÉ LUIS ARMENTA TREJO



**DIRECTOR DE TESIS:
DR. MANUEL DOMÍNGUEZ DE LA IGLESIA
CIUDAD UNIVERSITARIA, CD. MX. 2017**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

1. Datos del alumno
Armenta
Trejo
José Luis
57940922
Universidad Nacional Autónoma de
México
Facultad de Ciencias
Matemáticas
308556878
2. Datos del tutor
Dr
Manuel
Domínguez
De la Iglesia
3. Datos del sinodal 1
Dra
María Emilia
Caballero
Acosta
4. Datos del sinodal 2
Dr
Sergio Iván
López
Ortega
5. Datos del sinodal 3
Dr
Fernando
Baltazar
Larios
6. Datos del sinodal 4
Dr
Yuri
Salazar
Flores
7. Datos del trabajo escrito
Polinomios ortogonales y
procesos estocásticos
124
2017

Agradecimientos

Quiero agradecer a Dios por acompañarme en cada momento de mi vida, por darme a unos padres que me apoyan y me aman incondicionalmente, por ponerme en el camino correcto y por mostrarme e inspirarme en conocer las matemáticas, que más que una carrera, es mi vida.

Gracias a mis padres por inculcarme desde niño valores como responsabilidad, disciplina, perseverancia y por enseñarme sobre todo a amar, a ser un buen ser humano y por mostrarme que la vida puede ser muy sencilla y que lo que más importa en ésta es ser feliz. Gracias por su esfuerzo y estar pendiente de mis estudios, pero sobre todo, gracias por amarme en la forma en que lo hacen.

Gracias a todas esas personas especiales en mi vida: mis hermanos, mis amigos y a mi novia. Ella por su apoyo y compañía, también mi amiga, por ser mi inspiración para querer mejorar y lograr ser alguien en la vida, por entregarme una parte de tu vida. Ellos por estar conmigo en las buenas y en las malas y querer siempre lo mejor para mí, por creer en mí. Los amo a todos.

Gracias a mis jefes, Pavel Gibrán Esquivel, Juan Guillermo López, y David Gaytán por todo el apoyo que me han brindado, han mostrado ser excelentes personas y me han comprendido para que pueda seguir con mis estudios, pero sobre todo, me han abierto las puertas a un mundo de oportunidades y me han dado la oportunidad de aprender cómo puedo regresar algo de lo que he aprendido a la sociedad, me enseñaron a ver que la ciencia puede cambiar la vida de los demás y que es algo tangible. Los aprecio mucho.

Gracias a todos mis profesores, desde los que conocí en la primaria y secundaria como la maestra Lety y Jorge, los de la preparatoria, Juan Manuel, Claudio Pita y Jorge, hasta los de la universidad, Frank, Manuel, Abelardo y María Emilia. Pues esas personas me han ayudado o han influido de manera significativa a convertirme en lo que hoy soy.

Gracias a mi tutor y profesor, el doctor Manuel Domínguez de la Iglesia, por demostrar ser un excelente ser humano y por representar lo que debe ser un académico: una persona dispuesta a ayudar y a revisar los detalles como símbolo de su entrega a la enseñanza, una persona abierta a la opinión, alguien que se esfuerza por aprender y ayudar a aprender lo que sabe y al mismo tiempo exigente, alguien justo y honesto, pero sobre todo un gran amigo y ser humano. Admiro su trabajo y su dedicación, le agradezco por ayudarme a concretar esta etapa de mi vida y por ayudarme a convertir en un matemático. Infinitas gracias.

Índice general

1. Procesos estocásticos	8
1.1. Repaso de probabilidad	8
1.2. Algunos tipos de procesos estocásticos	14
1.3. Cadenas de Markov a tiempo discreto	16
1.4. Cadenas de Markov a tiempo continuo	23
1.5. Procesos de difusión	30
2. Polinomios Ortogonales	40
2.1. Definición y propiedades generales	43
2.2. Polinomios ortogonales de variable continua	48
2.3. Polinomios ortogonales de variable discreta	62
2.4. Operadores en diferencias	64
3. Polinomios Ortogonales y Procesos estocásticos	80
3.1. Representación espectral para procesos de nacimiento y muerte	80
3.1.1. Representación espectral de Karlin-McGregor	80
3.1.2. Propiedades probabilísticas	93
3.1.3. Métodos para calcular la medida ψ	96
3.1.4. Distribución límite condicional para procesos de nacimiento y muerte	104
3.2. Representación espectral para caminatas aleatorias	108
3.2.1. Distribución límite condicional para caminatas aleatorias	117
3.3. Representación espectral de la densidad de transición para procesos de difusión	119

Introducción

La presente tesis tiene como objetivo mostrar la importancia de los polinomios ortogonales en el área de procesos estocásticos. Para lograrlo se pretende mostrar la relación cercana que hay entre estas dos ramas, que si bien hay un sinnúmero de aplicaciones, hacemos particular énfasis en los teoremas de representación espectral para procesos de nacimiento y muerte, para caminatas aleatorias y para procesos de difusión, todos ellos utilizados para modelar un sinnúmero de fenómenos de interés en nuestra vida diaria.

La historia de los polinomios ortogonales tiene sus orígenes en el siglo XVIII y se relaciona con encontrar solución a problemas de aplicación práctica. Fue gracias a Lagrange que nació la primera familia de polinomios ortogonales de la historia (1784), conocidos como polinomios de Legendre (aunque su denominación es posterior a su descubrimiento). La siguiente familia, en orden de aparición, fue la de los polinomios de Hermite, en honor a Charles Hermite quien los estudió y publicó un ensayo en 1864. La siguiente familia fue la de los polinomios de Laguerre, llamados así debido a que fue Nicolás Laguerre quien los introdujo en 1879. Finalmente, la última clase de polinomios ortogonales clásicos surgen debido a la teoría de las ecuaciones diferenciales de segundo orden, mismas que surgieron en el siglo XVIII como respuesta directa a problemas físicos. Jacobi introduce una familia que generaliza a los polinomios de Legendre a partir de la función hipergeométrica de Gauss. Los polinomios de Jacobi, Laguerre y Hermite constituyen lo que hoy día se conocen como polinomios ortogonales clásicos en variable continua. Posteriormente se clasificaron los polinomios ortogonales clásicos en variable discreta, que son los polinomios de Charlier, Meixner, Krawtchouk y Hahn.

Los operadores infinitesimales asociados a cadenas de nacimientos y muerte (como caminatas aleatorias o procesos de nacimiento y muerte) vienen dados en términos de matrices tridiagonales. Por lo tanto, como veremos, se establece una conexión directa entre estos operadores y polinomios ortogonales mediante la denominada matriz de Jacobi. Para el caso de procesos de difusión, el operador infinitesimal viene dado en términos de un operador diferencial de segundo orden. Si es posible encontrar una familia completa de autofunciones ortogonales que resuelvan el correspondiente problema de Sturm-Liouville, podremos obtener una representación espectral de dicho proceso. Es de resaltar que este tipo de procesos queda totalmente determinado por dicha matriz u operador infinitesimal, y como ésta queda totalmente determinada por su representación, podemos concluir que estos procesos quedan totalmente determinados por estos polinomios y funciones ortogonales.

Los teoremas y proposiciones que veremos concernientes a la representación espectral, son consecuencia de estudiar las propiedades que tienen ciertos operadores autoadjuntos (discretos o continuos) definidos en ciertos espacios de Hilbert \mathcal{H} . En el caso de los polinomios ortogonales clásicos de variable continua, este operador autoadjunto se define a través de un operador diferencial de segundo orden cuyos coeficientes verifican la famosa ecuación

de Pearson, misma que se cumple para el estudio de los polinomios ortogonales clásicos discretos pero definiendo esta vez el operador autoadjunto por medio de un operador en diferencias de segundo orden. En ambos casos, los polinomios ortogonales resultan ser las autofunciones de dichos operadores. A su vez, toda familia de polinomios ortogonales verifica una relación de recurrencia a tres términos (operador de Jacobi), muchos de los cuales están muy relacionados con cadenas de nacimiento y muerte.

Es así, que mediante el teorema de Favard, o teorema espectral, se aplican estos resultados al estudio de caminatas aleatorias, procesos de nacimiento y muerte, y a procesos de difusión. En todos los casos veremos que siempre aparecen ecuaciones en común: una relación de recurrencia a tres términos para definir la medida y los polinomios o funciones ortogonales respecto a esa medida, o un operador diferencial de segundo orden para los procesos de difusión. En ambos casos se relacionan con las famosas ecuaciones de retroceso de Kolmogorov y de evolución o de Folker Planck para demostrar que los operadores definidos son autoadjuntos.

Como primer resultado de la representación espectral, tenemos la fórmula para expresar la distribución invariante, la distribución límite condicional y límite condicional doble por medio de los polinomios ortogonales, distribuciones que nos hablan del comportamiento asintótico de la probabilidad de transición de estos procesos. Así mismo, presentamos ejemplos de la representación para cada tipo de proceso.

De esta forma, el trabajo es un resumen de lo más importante que hay que saber de ambas teorías: definiciones, propiedades y teoremas fundamentales desarrollados desde sus orígenes hasta los últimos años. El pilar y eje de esta síntesis se basa en el trabajo hecho por Wim Schoutens en su libro “Stochastic Processes and Orthogonal Polynomials” [20] y en las notas del curso de “Análisis espectral de procesos estocásticos” de mi tutor Manuel Domínguez de la Iglesia. Como segundo eje y guía me basé también en el libro de Beals y Wong [3]. Así, la tesis está dividida en 3 capítulos.

El primer capítulo es un repaso de la teoría general de procesos estocásticos. Se basa principalmente en el libro de Karlin-Taylor [12]. Usé también resultados del libro de Luis Rincón [16] entre otros para demostrar algunas propiedades y ejemplos importantes, así como definiciones básicas utilizadas a lo largo del trabajo.

En el segundo capítulo enunciamos la definición de polinomios ortogonales continuos y polinomios ortogonales discretos. Se demuestran los resultados principales que conciernen a ambos: relación de recurrencia a tres términos, teoremas de caracterización, de aproximación, relación con ecuaciones diferenciales y en diferencias como eigenfunciones de un operador autoadjunto, fórmula de Rodrigues, relaciones de dualidad, pesos, y finalmente analizamos las propiedades de cada uno de los polinomios ortogonales clásicos. Para este capítulo nos basamos principalmente en el Beals-Wong [3] y en parte del trabajo de Schoutens [20].

La parte central de la tesis se trata en el capítulo tercero, pues es en este momento que discutimos las aplicaciones de los polinomios ortogonales a los procesos estocásticos: los teoremas de representación espectral y sus implicaciones en el cálculo de distribuciones invariantes y condicionales en el infinito. Este capítulo tiene sus bases mayormente en la tesis doctoral de Wim Schoutens [20], en las notas del curso de Manuel Domínguez y en varias publicaciones [10], [8], etc., las más importantes hechas por S. Karlin y J. McGregor.

El campo de aplicación de los polinomios ortogonales no sólo se limita al estudio de la representación espectral y al estudio de probabilidades límite condicionales, sino también

explican cómo se relacionan estos con los sistemas de Sheffer para generar martingalas a través de evaluar procesos en los polinomios ortogonales. Otra de las aplicaciones de los polinomios ortogonales es en la teoría de integración estocástica, dentro de la cuál destaca la propiedad de representación caótica y la propiedad de representación predecible. Este tipo de aplicaciones se pueden ver en [20].

Capítulo 1

Procesos estocásticos

Para poder hablar de procesos estocásticos necesitamos hablar sobre variables aleatorias y sigma álgebras, temas que son vistos a profundidad en un curso de teoría de la medida o probabilidad. Aquí las damos para poder definir un proceso estocástico y pasamos directamente a procesos de Markov y sus propiedades. A continuación damos las definiciones básicas de estos dos conceptos aunque no se explorarán sus propiedades a fondo, pues el propósito es servir como herramienta para desarrollar los preliminares y además el fin de la tesis no es profundizar exclusivamente en el área de probabilidad y procesos estocásticos, sino mostrar la relación de los procesos estocásticos con los polinomios ortogonales.

Dicho lo anterior, suponemos que las definiciones, proposiciones y resultados mostrados en este capítulo son básicas y usualmente bien conocidas, sirve más bien para resaltar la importancia de la matriz de transición de probabilidad de las cadenas de Markov, las aplicaciones que tienen este tipo de procesos y en general los procesos de Markov en el modelaje de incontables fenómenos, modelaje que se apoya también de la interpretación de la integral estocástica. Finalmente, sirve como un resumen de lo que hay que tener presente para poder entender las principales relaciones e implicaciones que se pretenden mostrar como objetivo de la tesis.

1.1. Repaso de probabilidad

Definición 1.1.1. Sea Ω un conjunto. Una clase no vacía $\mathcal{F} \subset \mathcal{P}(\Omega)$ (la clase de todos los subconjuntos de Ω) se llama un σ -anillo de subconjuntos de Ω , si:

1. $E, F \in \mathcal{F} \Rightarrow E - F \in \mathcal{F}$.
2. Si (E_n) es una sucesión de elementos de \mathcal{F} , entonces $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$.

Si además $\Omega \in \mathcal{F}$, entonces \mathcal{F} se llama σ -álgebra de subconjuntos de Ω .

Notemos que

$$\bigcap_{n=1}^{\infty} E_n = E_1 - \bigcup_{n=1}^{\infty} (E_1 - E_n).$$

Es decir, cualquier intersección numerable de elementos de \mathcal{F} es un elemento también de \mathcal{F} . De esta forma tiene sentido hablar de la siguiente definición.

Definición 1.1.2. (σ -álgebra generada). Sea \mathcal{C} una colección no vacía de subconjuntos de Ω . La σ -álgebra generada por \mathcal{C} , denotada por $\sigma(\mathcal{C})$, es la colección

$$\sigma(\mathcal{C}) = \bigcap \{F : F \text{ es } \sigma\text{-álgebra y } \mathcal{C} \subseteq F\}.$$

Es decir, la colección $\sigma(\mathcal{C})$ es la intersección de todas aquellas σ -álgebras que contienen a \mathcal{C} .

Considérese ahora a la colección de todos los intervalos abiertos (a, b) de \mathbb{R} , en donde $a \leq b$. A la mínima σ -álgebra generada por esta colección se le llama σ -álgebra de Borel de \mathbb{R} y se le denota por $\mathcal{B}(\mathbb{R})$.

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a, b) \subseteq \mathbb{R} : a \leq b\}.$$

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama conjuntos de Borel, Borelianos o conjuntos Borel medibles.

Definición 1.1.3. Un espacio medible es una pareja (Ω, \mathcal{F}) en la que Ω es un conjunto no vacío y \mathcal{F} es una σ -álgebra de subconjuntos de Ω . A los elementos de \mathcal{F} se les llama eventos o conjuntos medibles.

Definición 1.1.4. Sean (Ω, \mathcal{F}) y (Ω', \mathcal{F}') dos espacios medibles. Una función $f : \Omega \rightarrow \Omega'$ se llama medible relativa a las σ -álgebras \mathcal{F} y \mathcal{F}' si $f^{-1}(\mathcal{F}') \subset \mathcal{F}$, i.e. $f^{-1}(E') \in \mathcal{F}$ para todo $E' \in \mathcal{F}'$.

Definición 1.1.5. Sea (Ω, \mathcal{F}) un espacio medible. Una medida de probabilidad es una función $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ que satisface

1. $\mathbb{P}(\Omega) = 1$.
2. $\mathbb{P}(A) \geq 0$, para cualquier $A \in \mathcal{F}$.
3. Si $A_1, A_2, \dots \in \mathcal{F}$ son ajenos dos a dos ($A_n \cap A_m = \emptyset$) para $n \neq m$, entonces

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Una definición adicional y de gran importancia es la probabilidad condicional:

Definición 1.1.6. Sean A y B dos eventos ($A, B \in \mathcal{F}$). La probabilidad de ocurrencia del evento A dado B se define como:

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)},$$

siempre que $\mathbb{P}(B) \neq 0$.

Mencionamos ahora algunas propiedades, lemas y teoremas importantes de esta medida de probabilidad de acuerdo con [16], en donde podemos encontrar sus respectivas demostraciones.

Proposición 1.1.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Entonces

1. $\mathbb{P}(\emptyset) = 0$.

2. Si $A_1, \dots, A_n \in \mathcal{F}$ son ajenos dos a dos, entonces

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbb{P}(A_k).$$

3. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$, donde $A^c = \Omega - A$.

4. Si $A \subseteq B$, entonces $\mathbb{P}(B - A) = \mathbb{P}(B) - \mathbb{P}(A)$.

5. Si $A \subset B$, entonces $\mathbb{P}(A) \leq \mathbb{P}(B)$.

6. $0 \leq \mathbb{P}(A) \leq 1$.

7. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

8. $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$.

Proposición 1.1.2. (Desigualdad de Boole). Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión de eventos. Entonces

1.

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

2.

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \geq 1 - \sum_{n=1}^{\infty} \mathbb{P}(A_n^c).$$

Teorema 1.1.1. (Teorema de probabilidad total). Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, y sea $\{A_1, A_2, \dots\}$ una partición de Ω tal que cada elemento de la partición es un evento con probabilidad estrictamente positiva. Entonces

$$\mathbb{P}(B) = \sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n)$$

Teorema 1.1.2. (Teorema de Bayes). Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, y sea $\{A_1, A_2, \dots\}$ una partición de Ω tal que cada elemento de la partición tiene probabilidad estrictamente positiva. Entonces, para cualquier evento B tal que $\mathbb{P}(B) > 0$, y para cualquier $m \geq 1$ fijo tenemos

$$\mathbb{P}(A_m|B) = \frac{\mathbb{P}(B|A_m)\mathbb{P}(A_m)}{\sum_{n=1}^{\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n)}.$$

Proposición 1.1.3. Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión no decreciente de eventos, esto es, $A_1 \subseteq A_2 \subseteq \dots$. Entonces

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

Proposición 1.1.4. (Continuidad de la probabilidad). Sea $\{A_n : n \in \mathbb{N}\}$ una sucesión de eventos convergente al evento A . Entonces

$$\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = \mathbb{P}(A).$$

Definición 1.1.7. Sea (Ω, \mathcal{F}) un espacio medible. Una variable aleatoria real es una función $X : \Omega \rightarrow \mathbb{R}$ tal que para cualquier conjunto Boreliano B , se cumple que el conjunto $X^{-1}(B)$ es un elemento de \mathcal{F} .

Definición 1.1.8. Consideremos los espacios medibles (Ω, \mathcal{F}) y $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Si X es una función de Ω en \mathbb{R} , entonces se definimos la mínima σ -álgebra de subconjuntos de Ω respecto de la cual X es variable aleatoria como

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}.$$

Tenemos de nuevo [16] las siguientes propiedades para variables aleatorias.

Sean X, Y variables aleatorias y $c \in \mathbb{R}$ una constante.

1. La función constante $X = c$ es una variable aleatoria.
2. $X + Y$ es variable aleatoria.
3. XY es variable aleatoria.
4. Si $Y \neq 0$ entonces X/Y es variable aleatoria.
5. $\max\{X, Y\}$ y $\min\{X, Y\}$ son variables aleatorias.
6. $|X|$ es variable aleatoria.

Definición 1.1.9. (Función de distribución). La función de distribución de una variable aleatoria X es la función $F_X(x) : \mathbb{R} \rightarrow [0, 1]$, definida como sigue

$$F_X(x) = \mathbb{P}(X \leq x).$$

Proposición 1.1.5. Sea $F(x)$ la función de distribución de una variable aleatoria. Entonces

1. $\lim_{x \rightarrow \infty} F(x) = 1$.
2. $\lim_{x \rightarrow -\infty} F(x) = 0$.
3. Si $x_1 \leq x_2$, entonces $F(x_1) \leq F(x_2)$.
4. $F(x)$ es continua por la derecha, es decir $\lim_{x \rightarrow a^+} F(x) = F(a)$.

Definición 1.1.10. Una función $F(x) : \mathbb{R} \rightarrow [0, 1]$ es llamada función de distribución si cumple las cuatro propiedades anteriores.

Proposición 1.1.6. Sea X una variable aleatoria con función de distribución $F(x)$ y denotemos por $F(a-) = \lim_{x \rightarrow a^-} F(x)$. Para cualesquiera números reales $a < b$,

1. $\mathbb{P}(X < a) = F(a-)$.
2. $\mathbb{P}(X = a) = F(a) - F(a-)$.
3. $\mathbb{P}(a < X \leq b) = F(b) - F(a)$.
4. $\mathbb{P}(a \leq X \leq b) = F(b) - F(a-)$.
5. $\mathbb{P}(a < X < b) = F(b-) - F(a)$.

$$6. \mathbb{P}(a \leq X < b) = F(b-) - F(a-).$$

Definición 1.1.11. (Variable aleatoria discreta). La variable aleatoria X se llama discreta si su correspondiente función de distribución $F(x)$ es una función constante por pedazos.

Definición 1.1.12. (Variable aleatoria continua). La variable aleatoria X se llama continua si su correspondiente función de distribución es continua.

Definición 1.1.13. (Variable aleatoria absolutamente continua). La variable aleatoria continua X con función de distribución $F(x)$ se llama absolutamente continua, si existe una función no negativa e integrable tal que para cualquier valor de x se cumple

$$F(x) = \int_{-\infty}^x f(u)du.$$

En tal caso a la función $f(x)$ se le llama función de densidad de X .

Definición 1.1.14. (Esperanza). Sea X con función de distribución $F(x)$. La esperanza de X , denotada por $E[X]$, se define como el número

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x dF(x),$$

cuando esta integral sea absolutamente convergente, es decir, cuando $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, y en tal caso se dice que X es integrable, o que tiene esperanza finita.

A la esperanza se le conoce también con el nombre de media, valor esperado, valor promedio o valor medio, y en general se usa la letra griega μ para denotarla.

Definición 1.1.15. (Varianza). La varianza de una variable aleatoria X , se define como la siguiente esperanza, si existe

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

Proposición 1.1.7. (Propiedades de la varianza). Sean X y Y con varianza finita, y sea c una constante. Entonces

1. $\text{Var}(X) \geq 0$.
2. $\text{Var}(c) = 0$.
3. $\text{Var}(cX) = c^2 \text{Var}(X)$.
4. $\text{Var}(X + c) = \text{Var}(X)$.
5. $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}^2[X]$.

Distribuciones importantes

A continuación enlistamos las funciones de densidad de las distribuciones discretas más importantes

1. Bernoulli, $B(p)$

$$(1.1) \quad f(x) = \begin{cases} p, & \text{si } x = 0, \\ 1 - p, & \text{si } x = 1 \end{cases}, \quad \text{sop}(f) = \{0, 1\}, \quad 0 < p < 1.$$

2. Poisson, $P(\lambda)$

$$(1.2) \quad f(x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad \text{sop}(f) = \{0, 1, 2, \dots\}, \quad \lambda > 0.$$

3. Pascal, $Pa(\gamma, \mu)$

$$(1.3) \quad \frac{(1-\mu)^\gamma (\gamma)_x \mu^x}{x!}, \quad \text{sop}(f) = \{0, 1, 2, \dots\}, \quad \gamma > 0, \quad 0 < \mu < 1.$$

4. Binomial, $Bin(N, p)$

$$(1.4) \quad \binom{N}{x} p^x (1-p)^{N-x}, \quad \text{sop}(f) = \{0, 1, 2, \dots, N\}, \quad 0 < p < 1.$$

5. Geométrica, $Geo(p)$

$$(1.5) \quad f(x) = p(1-p)^x, \quad \text{sop}(f) = \{0, 1, 2, \dots\}, \quad 0 < p < 1.$$

6. Hipergeométrica I $Hyp(\alpha, \beta, N)$

$$(1.6) \quad f(x) = \binom{N}{x} \frac{(\alpha+1)_x (\beta+1)_{N-x}}{(\alpha+\beta+2)_N}, \quad \text{sop}(f) = \{0, 1, \dots, N\}, \quad \alpha, \beta > -1, \quad \alpha, \beta < -N.$$

7. Hipergeométrica II $HypII(\alpha, \beta, N)$

$$(1.7) \quad f(x) = \frac{\binom{\alpha}{x} \binom{\beta}{N-x}}{\binom{\alpha+\beta}{N}}, \quad \text{sop}(f) = \{0, 1, \dots, N\}, \quad \alpha, \beta > -1, \quad \alpha, \beta < -N.$$

A continuación enlistamos las funciones de densidad de las distribuciones continuas más importantes

1. Normal, $N(\mu, \sigma^2)$

$$(1.8) \quad f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad \text{sop}(f) = \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \sigma^2 > 0.$$

2. Gamma, $G(\alpha, \beta)$

$$(1.9) \quad f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta}, \quad \text{sop}(f) = (0, \infty), \quad \alpha, \beta > 0.$$

3. Beta, $B(\alpha, \beta)$

$$(1.10) \quad f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad \text{sop}(f) = (0, 1), \quad \alpha, \beta > 0.$$

4. Exponencial $Exp(\lambda)$

$$(1.11) \quad f(x) = \lambda e^{-\lambda x}, \quad \text{sop}(f) = (0, \infty), \quad \lambda > 0.$$

5. Uniforme $U(a, b)$

$$(1.12) \quad f(x) = \frac{1}{b-a}, \quad \text{sop}(f) = (a, b), \quad a < b.$$

1.2. Algunos tipos de procesos estocásticos

Definición 1.2.1. *Un proceso estocástico es una colección de variables aleatorias $\{X_t : t \in \mathcal{T}\}$ parametrizada por un conjunto \mathcal{T} , llamado espacio parametral, en donde las variables toman valores en un conjunto S llamado espacio de estados.*

En los casos más sencillos se toma como espacio parametral el conjunto discreto $\mathcal{T} = \{0, 1, 2, \dots\}$ y estos números se interpretan como tiempos. En este caso se dice que el proceso es a tiempo discreto, y en general este tipo de procesos se denotará por $\{X_n : n = 0, 1, \dots\}$, o explícitamente,

$$X_0, X_1, X_2, \dots$$

Así, para cada n , X_n es el valor del proceso o estado del sistema al tiempo n .

El espacio parametral puede también tomarse como el conjunto continuo $T = [0, \infty)$. Se dice entonces que el proceso es a tiempo continuo, y se denotará por

$$\{X_t : t \geq 0\}.$$

Por lo tanto, seguiremos la convención de que si el subíndice es n , entonces los tiempos son discretos, y si el subíndice es t , el tiempo se mide de manera continua.

A continuación enumeramos los tipos de procesos estocásticos más importantes que hay.

1. *Procesos con Incrementos Independientes Estacionarios*

Sean $t_1 < t_2 < \dots < t_n$, con $t_1, t_2, \dots, t_n \in \mathbb{R}$. Si las variables aleatorias

$$X_{t_2} - X_{t_1}, X_{t_3} - X_{t_2}, \dots, X_{t_n} - X_{t_{n-1}}$$

son independientes para todas las elecciones de t_1, t_2, \dots, t_n decimos entonces que X_t es un proceso con **incrementos independientes**.

Si la distribución de los incrementos $X_{t_1+h} - X_{t_1}$ depende sólo de la longitud del intervalo y no del tiempo t_1 entonces se dice que el proceso tiene **incrementos estacionarios**.

2. *Martingalas*

Sea X_t un proceso estocástico con valores en \mathbb{R} con conjunto indizador continuo o discreto. Decimos que X_t es una **martingala** si

$E[|X_t|] < \infty$ para toda t , y si para cualquier $t_1 < t_2 < \dots < t_{n+1}$,

$$\mathbb{E}[X_{t_{n+1}} | X_{t_1} = a_1, \dots, X_{t_n} = a_n] = a_n$$

para todos los valores de a_1, \dots, a_n .

Para poder hablar de la definición general de martingala necesitamos el concepto de filtración.

Definición 1.2.2. Una filtración es una colección de σ -álgebras $\{\mathcal{F}_n\}_{n \geq 1}$ tal que $\mathcal{F}_n \subseteq \mathcal{F}_m$, cuando $n \leq m$. En particular, la filtración natural o canónica de un proceso $\{X_n : n \geq 1\}$ es aquella sucesión de σ -álgebras definidas por

$$\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}, \quad n \geq 1.$$

Damos ahora la definición de adaptación.

Definición 1.2.3. Se dice que un proceso estocástico $\{X_n : n \geq 1\}$ es adaptado a una filtración $\{\mathcal{F}_n\}_{n \geq 1}$ si la variable X_n es \mathcal{F}_n -medible, para cada $n \geq 1$.

Podemos ahora dar una definición un poco más general de martingala.

Definición 1.2.4. Se dice que un proceso a tiempo discreto $\{X_n : n \geq 1\}$ es una martingala respecto de una filtración $\{\mathcal{F}_n\}_{n \geq 1}$ si:

- Es integrable.
- Es adaptado a la filtración.
- Para cualesquiera $n \leq m$,

$$\mathbb{E}(X_m | \mathcal{F}_n) = X_n.$$

3. *Procesos de Markov*

Sea X_t un proceso estocástico con espacio de estados S , un proceso de Markov es un proceso para el cual

$$\mathbb{P}\{a < X_t \leq b | X_{t_1} = x_1, X_{t_2} = x_2, \dots, X_{t_n} = x_n\} = \mathbb{P}\{a < X_t \leq b | X_{t_n} = x_n\}$$

siempre que $t_1 < t_2 < \dots < t_n < t$ y para toda $x_1, x_2, \dots, x_n \in S$ y $a, b \in S$.

En palabras simples, la probabilidad del comportamiento particular del proceso en el futuro, dado que se conoce su valor presente no se ve alterado por información adicional concerniente a su pasado.

1.3. Cadenas de Markov a tiempo discreto

Una cadena de Markov discreta $\{X_n\}$ es un proceso estocástico de Markov con espacio de estados $E \subset \mathbb{N}$ un conjunto numerable o finito, y para el cual $\mathbb{N} = \{0, 1, 2, \dots\}$. Es conveniente identificar el espacio de estados del proceso con los enteros no negativos $\{0, 1, 2, \dots\}$ y se acostumbra decir que X_n está en el estado i si $X_n = i$.

La probabilidad de que $X_{n+1} = j$ dado que $X_n = i$ se llama **probabilidad de transición de un paso** y se denota por $P_{ij}^{n,n+1}$, es decir,

$$(1.13) \quad P_{ij}^{n,n+1} = \mathbb{P}(X_{n+1} = j | X_n = i).$$

Nótese que la notación hace énfasis en el hecho de que en general las probabilidades de transición son funciones no sólo del estado inicial y el estado final, sino también función del tiempo de transición. Cuando las probabilidades de transición de un paso son independientes del tiempo, es decir del valor de n , decimos que el proceso de Markov tiene probabilidades de transición estacionarias o que la cadena de Markov es **homogénea**.

En ese caso $P_{ij}^{n,n+1}$ es independiente de n y P_{ij} es la probabilidad de que en un intento el valor del estado vaya de i a j . Se acostumbra ordenar estos valores P_{ij} en una matriz de la siguiente forma

$$(1.14) \quad P = \begin{pmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ P_{i0} & P_{i1} & P_{i2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Nos referimos a ella como $P = (P_{ij})$ la **matriz de transición de probabilidades del proceso**.

La fila $i + 1$ de P es la distribución de probabilidad de los valores de X_{n+1} bajo la condición $X_n = i$. Si el número de estados es finito entonces P es una matriz cuadrada finita cuya dimensión es igual al número de estados. Claramente, las cantidades P_{ij} satisfacen las condiciones

$$P_{ij} \geq 0, \quad i, j = 0, 1, 2, \dots$$

$$\sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, 2, \dots$$

El proceso queda totalmente determinado una vez determinada la matriz de transición (1.14) y el valor de X_0 . Se puede probar por inducción y usando la propiedad de Markov que

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, \dots, X_n = i_n) = P_{i_{n-1}, i_n} P_{i_{n-2}, i_{n-1}} \cdots P_{i_0, i_1} p_{i_0}$$

donde

$$p_{i_0} = \mathbb{P}(X_0 = i_0).$$

Definimos ahora la probabilidad de alcanzar el estado j en n transiciones dado que empezamos en el estado i como:

$$P_{ij}^n = \mathbb{P}(X_n = j | X_0 = i).$$

Teorema 1.3.1. *Si la matriz de probabilidad de un paso de una cadena de Markov es (P_{ij}) , entonces*

$$P_{ij}^n = \sum_{k=0}^n P_{ik}^r P_{kj}^s$$

para cualquier par de enteros fijos no negativos r y s que satisfacen $r + s = n$, donde definimos

$$P_{ij}^0 = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Ejemplos

1. El siguiente ejemplo es el modelo genético debido a S. Wright. Asumimos que la población de genes es constante y de tamaño $2N$ compuesta de individuos del tipo a y del tipo A . La siguiente generación está determinada por $2N$ intentos independientes binomiales: si la población ancestral consiste de j genes a y $2N - j$ genes A entonces cada intento resulta en a o en A con probabilidades

$$p_j = \frac{j}{2N}, \quad q_j = 1 - \frac{j}{2N},$$

respectivamente. Se llevan a cabo selecciones repetidas con reemplazo. Con este procedimiento generamos una cadena de Markov $\{X_n\}$ donde X_n es el número de genes a en la n -ésima generación en una población de tamaño $2N$. El espacio de estados contiene los $2N + 1$ valores $\{0, 1, 2, \dots, 2N\}$. La matriz de transición de probabilidad se calcula de acuerdo a la distribución binomial como

$$P_{jk} = \mathbb{P}(X_{n+1} = k | X_n = j) = \binom{2N}{k} p_j^k q_j^{2N-k} \quad (j, k = 0, 1, \dots, 2N).$$

Notemos que una vez que $X_n = 0$ o $X_n = 2N$ entonces $X_{n+k} = 0$ o $X_{n+k} = 2N$ respectivamente para toda $k \geq 0$.

Uno de los problemas de interés es determinar bajo la condición $X_0 = i$ la probabilidad de que la población termine compuesta únicamente de genes del tipo a o del tipo A , de la misma forma es importante saber a que tasa se aproxima a este estado.

2. Uno de los ejemplos más sencillos sobre cadenas de Markov a tiempo discreto son las caminatas aleatorias. Una caminata aleatoria es una cadena de Markov cuyo espacio de espacios es un conjunto finito o infinito de los enteros, en el cual una partícula, si está en el estado i , puede en una sola transición quedarse en el estado i o moverse a los estados adyacentes $i - 1$ ó $i + 1$. Si el estado de espacios se toma como los enteros no negativos, la matriz de transición de la caminata aleatoria tiene la forma

$$(1.15) \quad P = \begin{pmatrix} r_0 & p_0 & 0 & 0 & \cdots & & & & \\ q_1 & r_1 & p_1 & 0 & \cdots & & & & \\ 0 & q_2 & r_2 & p_2 & \cdots & & & & \\ & & & \ddots & \ddots & & & & \\ & & 0 & q_i & r_i & p_i & 0 & \cdots & \\ & & & & & & \ddots & \ddots & \end{pmatrix}$$

Donde $p_i > 0, q_i > 0, r_i \geq 0$ y $q_i + r_i + p_i = 1, i = 1, 2, \dots (i \geq 1), p_0 \geq 0, r_0 \geq 0, r_0 + p_0 \leq 1$. La posibilidad $r_0 + p_0 < 1$ se entiende como que existe una probabilidad $q_0 = 1 - p_0 - r_0$ de que el proceso sea absorbido a cierto estado -1 .

Un ejemplo importante de caminata aleatoria es la caminata aleatoria simétrica, que sirve como versión discreta del movimiento Browniano. La caminata aleatoria simétrica en los enteros es una cadena de Markov con espacio de estados todos los enteros y con

$$P_{ij} = \begin{cases} p & \text{si } j = i + 1 \\ q & \text{si } j = i - 1 \\ 0 & \text{de otra forma.} \end{cases}, \quad i, j \in \mathbb{Z}, \quad p + q = 1$$

El ejemplo más sencillo de caminata aleatoria simétrica corresponde al caso $p = q = \frac{1}{2}$. La caminata aleatoria simétrica en n dimensiones consiste en identificar el espacio de estados con el conjunto de todos los puntos $k = (k_1, k_2, \dots, k_n)$ con entradas en los enteros. La matriz de transición de probabilidad está definida por

$$P_{kl} = \begin{cases} \frac{1}{2n} & \text{si } \sum_{i=1}^n |l_i - k_i| = 1 \\ 0 & \text{de otra forma.} \end{cases}$$

y análogo al caso 1-dimensional, la caminata aleatoria simétrica n -dimensional representa una versión discreta del movimiento Browniano n -dimensional.

Clasificación de los estados de una cadena de Markov

El estado j se llama **accesible** desde el estado i si para algún entero $n \geq 0, P_{ij}^n > 0$, es decir, j es accesible desde el estado i si hay una probabilidad positiva de que en un tiempo finito de transiciones el estado j sea alcanzado comenzando en el estado i . Si dos estados i y j son accesibles entre sí, se dice que **se comunican** y escribimos $i \leftrightarrow j$. El concepto de comunicación es una relación de equivalencia, misma que induce una partición sobre el conjunto de estados.

Decimos que una cadena de Markov es **irreducible** si la relación de equivalencia inducida por la comunicabilidad induce una sola clase. En otras palabras, es irreducible si todos los estados se comunican entre sí.

Definimos ahora el **período** del estado $i, d(i)$, como el máximo común divisor de todos los enteros $n \geq 1$ para los cuales $P_{ii}^n > 0, d(i) = m.c.d\{n \geq 1 : P_{ii}^n > 0\}$. (Si $P_{ii}^n = 0$ para todos los $n \geq 1$ definimos $d(i)$ como $d(i) = 0$). Si en la caminata aleatoria todos los $r_i = 0$,

entonces cada estado tiene período 2. Una cadena de Markov en la cual cada estado tiene período 1 se llama **aperiódica**,

Consideremos ahora un estado arbitrario i fijo. Definimos para cada entero $n \geq 1$

$$f_{ii}^n = \mathbb{P}(X_n = i, X_v \neq i, v = 1, 2, \dots, n-1 | X_0 = i).$$

En otras palabras, f_{ii}^n es la probabilidad de que, empezando en el estado i el primer regreso al estado i ocurra en n transiciones. Tenemos que $f_{ii}^1 = P_{ii}$ y f_{ii}^n puede ser calculado recursivamente de acuerdo a

$$P_{ii}^n = \sum_{k=0}^{n-1} f_{ii}^k P_{ii}^{n-k}, n \geq 1$$

donde definimos f_{ii}^n para toda i .

Definición 1.3.1. La **función generadora** $P_{ij}(s)$ de la sucesión $\{P_{ij}^n\}$ es

$$P_{ij}(s) = \sum_{n=0}^{\infty} P_{ij}^n s^n, \quad |s| < 1.$$

De manera similar definimos la función generadora para la sucesión $\{f_{ij}^n\}$

$$F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^n s^n, \quad |s| < 1.$$

Recordemos además que si

$$A(s) = \sum_{k=0}^{\infty} a_k s^k \quad \text{y} \quad B(s) = \sum_{l=0}^{\infty} b_l s^l$$

entonces

$$A(s)B(s) = C(s) = \sum_{r=0}^{\infty} c_r s^r$$

donde

$$c_r = a_0 b_r + a_1 b_{r-1} + \dots + a_r b_0.$$

Si identificamos a las $a'_k s$ con las $f_{ii}^k s$ y a las $b'_l s$ con las $P_{ii}^l s$ entonces obtenemos

$$F_{ii}(s)P_{ii}(s) = P_{ii}(s) - 1, \quad |s| < 1$$

o

$$P_{ii}(s) = \frac{1}{1 - F_{ii}(s)}, \quad |s| < 1.$$

Obtenemos entonces que

$$P_{ij}^n = \sum_{k=0}^{n-1} f_{ij}^k P_{jj}^{n-k}, \quad i \neq j, \quad n \geq 0$$

donde f_{ij}^k es la probabilidad de que la primera llegada del estado i al estado j ocurra en k transiciones.

De nuevo definimos $f_{ij}^{(0)} = 0$ para todo i y j . Se sigue que

$$P_{ij}(s) = F_{ij}(s)P_{jj}(s), \quad |s| < 1.$$

Definición 1.3.2. Decimos que un estado i es **recurrente** si y sólo si

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

Lo anterior quiere decir que un estado i es recurrente si y sólo si, comenzando en el estado i , la probabilidad de regresar al estado i después de un intervalo finito de tiempo es 1. Un estado no recurrente se llama **transitorio**.

Teorema 1.3.2. Un estado es recurrente si y sólo si

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty.$$

Además, si $i \leftrightarrow j$ y i es recurrente entonces j es recurrente.

Lo anterior demuestra que todos los estados en una clase de equivalencia son recurrentes o transitorios. El número esperado de regresos al estado i dado que $X_0 = i$ es

$$\sum_{n=0}^{\infty} P_{ii}^n.$$

Así, el teorema anterior establece que un estado i es recurrente si y sólo si el número esperado de regresos es infinito.

Enunciamos ahora el teorema básico de las cadenas de Markov.

Teorema 1.3.3. (El teorema límite básico para cadenas de Markov)

a) Consideremos una cadena de Markov recurrente, irreducible y aperiódica. Sea P_{ii}^n la probabilidad de entrar al estado i en la transición n , $n = 0, 1, 2, \dots$, dado $X_0 = i$.

Convenimos $P_{ii}^0 = 1$. Sea f_{ii}^n la probabilidad de regresar por primera vez al estado i en la transición n , $n = 0, 1, 2, \dots$ donde $f_{ii}^0 = 0$. Así

$$P_{ii}^n - \sum_{k=0}^n f_{ii}^{n-k} P_{ii}^k = \begin{cases} 1 & \text{si } n = 0 \\ 0 & \text{si } n > 0 \end{cases}$$

Entonces

$$\lim_{n \rightarrow \infty} P_{ii}^n = \frac{1}{\sum_{n=0}^{\infty} n f_{ii}^n}.$$

b) Bajo las mismas condiciones

$$\lim_{n \rightarrow \infty} P_{ji}^n = \lim_{n \rightarrow \infty} P_{ii}^n.$$

Podemos consultar la demostración en [12].

Hay que notar que si C es una clase recurrente entonces $P_{ij}^n = 0$ para $i \in C, j \notin C$ y cada n . Así, una vez en C , no es posible dejar C . Se sigue que la submatriz $(P_{ij}), i, j \in C$ es una matriz de transición de probabilidad y su cadena de Markov asociada es irreducible y recurrente.

Notemos también que si $a_n \rightarrow a$ conforme $n \rightarrow \infty$, se tiene que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n a_k = a.$$

Así, si i es un miembro de una clase aperiódica recurrente entonces

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n P_{ii}^m = \frac{1}{\sum_{n=0}^{\infty} n f_{ii}^n} = \frac{1}{m_i}$$

donde m_i es el tiempo medio de recurrencia.

Si i es un elemento de una clase recurrente periódica con período d , entonces $P_{ii}^m = 0$ si m no es un múltiplo de d y además

$$\lim_{n \rightarrow \infty} P_{ii}^{nd} = \frac{d}{m_i}.$$

Si $\lim_{n \rightarrow \infty} P_{ii}^n = \pi_i > 0$ para alguna i en una clase recurrente aperiódica, entonces $\pi_j > 0$ para todas las j en la clase de i . En este caso, llamamos la clase **recurrente positiva o fuertemente ergódica**. Si cada $\pi_i = 0$ y la clase es recurrente decimos que la clase es **recurrente nula o débilmente ergódica**.

Teorema 1.3.4. *En una clase recurrente positiva con estados $j = 0, 1, 2, \dots$,*

$$\lim_{n \rightarrow \infty} P_{jj}^n = \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}, \quad \sum_{i=0}^{\infty} \pi_i = 1$$

y las componentes de π quedan únicamente determinados por el conjunto de ecuaciones

$$(1.16) \quad \pi_i \geq 0, \quad \sum_{i=0}^{\infty} \pi_i = 1, \quad \pi_j = \sum_{i=0}^{\infty} \pi_i P_{ij}.$$

Cualquier conjunto $(\pi_i)_{i=0}^{\infty}$ que satisface lo anterior es llamado una **distribución de probabilidad estacionaria** de la cadena de Markov o **distribución invariante**.

Una demostración del teorema anterior se puede encontrar en [12].

Ya establecimos que si j es un estado transitorio, entonces $P_{ij}^n \rightarrow 0$, y que si i, j están en la misma clase recurrente aperiódica, entonces $P_{ij}^n \rightarrow \pi_j \geq 0$. Si i, j están en la misma clase recurrente aperiódica, entonces la misma conclusión es válida si reemplazamos P_{ij}^n por $n^{-1} \sum_{m=1}^n P_{ij}^m$.

Para completar la discusión del comportamiento límite de P_{ij}^n , queda considerar el caso en que i es transitorio y j es recurrente.

Si T es el conjunto de todos los estados transitorios, consideremos entonces

$$x_i^1 = \sum_{j \in T} P_{ij} \leq 1, \quad i \in T,$$

y definamos recursivamente

$$x_i^n = \sum_{j \in T} P_{ij} x_j^{n-1}, \quad n \geq 2 \quad i \in T.$$

Observemos que x_i^n es la probabilidad de que, empezando desde el estado i , el estado del proceso permanezca en T para las siguientes n transiciones. Como $x_i^n \leq 1$ para toda $n \geq 1$, tenemos por inducción que x_i^n es no creciente como función de n . De hecho

$$x_i^2 = \sum_{j \in T} P_{ij} x_j^1 \leq \sum_{j \in T} P_{ij} = x_i^1.$$

Ahora, asumiendo que $x_j^n \leq x_j^{n-1}$ para toda $j \in T$, tenemos que

$$0 \leq x_i^{n+1} = \sum_{j \in T} P_{ij} x_j^n \leq \sum_{j \in T} P_{ij} x_{ij} x_j^{n-1} = x_i^n.$$

Por lo tanto $x_i^n \downarrow x_i$, es decir x_i^n decrece a algún límite x y

$$x_i = \sum_{j \in T} P_{ij} x_j, \quad i \in T.$$

Notemos que si hay sólo un número finito de estados, M , entonces no hay estados recurrentes nulos y no todos los estados pueden ser transitorios. De hecho, dado que $\sum_{j=0}^{M-1} P_{ij}^n = 1$ para toda n , no puede ocurrir que $\lim_{n \rightarrow \infty} P_{ij} = 0$ para toda j .

El mismo caso restringido a las clases recurrentes muestra que no hay estados nulos. Sean C, C_1, C_2, \dots clases recurrentes. Definimos $\pi_i(C)$ como la probabilidad de que el proceso sea absorbido en la clase recurrente C si el estado inicial es el estado transitorio i .

Sea $\pi_i^n(C)$ la probabilidad de que el proceso entre y sea absorbido en C por primera vez en la transición n , dado que el estado inicial es $i \in T$. Entonces

$$\begin{aligned} \pi_i(C) &= \sum_{n=1}^{\infty} \pi_i^n(C) \leq 1 \\ \pi_i^1(C) &= \sum_{j \in C} P_{ij} \\ \pi_i^n(C) &= \sum_{j \in T} P_{ij} \pi_j^{n-1}(C), \quad n \geq 2. \end{aligned}$$

Usando estas ecuaciones obtenemos

$$\begin{aligned} \pi_i(C) &= \pi_i^1(C) + \sum_{n=2}^{\infty} \pi_i^n(C) = \pi_i^1(C) + \sum_{n=2}^{\infty} \sum_{j \in T} P_{ij} \pi_j^{n-1}(C) \\ &= \pi_i^1(C) + \sum_{j \in T} P_{ij} \sum_{n=2}^{\infty} \pi_j^{n-1}(C) \\ \pi_i(C) &= \pi_i^1(C) + \sum_{j \in T} P_{ij} \pi_j(C), \quad i \in T. \end{aligned}$$

Asumiendo que la única solución acotada del conjunto homogéneo de ecuaciones

$$w_i = \sum_{j \in T} P_{ij} w_j, \quad i \in T$$

es el vector cero, entonces $\{\pi_i(C)\}$ queda determinado como la única solución acotada del sistema de ecuaciones anterior. Además, $\pi_i^1(C) > 0$ para alguna $i \in T$ o $\pi_i(C) = 0$ para cada $i \in T$ y así $\pi_i^n(C) = 0$ para toda n .

Teorema 1.3.5. *Sea $j \in C$ (C una clase recurrente aperiódica). Entonces, para $i \in T$, tenemos*

$$\lim_{n \rightarrow \infty} P_{ij}^n = \pi_i(C), \quad \lim_{n \rightarrow \infty} P_{jj}^n = \pi_j(C).$$

Si C es periódico y $j \in C$, una prueba similar puede ser usada para mostrar que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n P_{ij}^m = \pi_i(C) \pi_j.$$

Hacemos énfasis en el hecho de que si i es un estado transitorio y j es un estado recurrente, entonces el límite P_{ij}^n depende tanto de i como de j . Contrastemos contra el caso en que i y j pertenecen a la misma clase recurrente.

1.4. Cadenas de Markov a tiempo continuo

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y consideremos una cadena de Markov a tiempo continuo $\{X_t, t \geq 0\}$ con espacio de estados $\mathbf{E} \subseteq \mathbb{Z}$, que puede ser finito, infinito o doblemente infinito. La propiedad de Markov en este caso puede escribirse como

$$\mathbb{P}(X_{s+t} = j | X_s = i, X_\tau, 0 \leq \tau < s) = \mathbb{P}(X_{s+t} = j | X_s = i) = P_{ij}(s, t), \quad 0 < s < t.$$

Al igual que en el caso de caminatas aleatorias, supondremos que la cadena de Markov es homogénea y por lo tanto las probabilidades anteriores sólo dependen de la diferencia temporal $t - s$, i.e. $P_{ij}(s, t) = P_{ij}(0, t - s)$. En este caso escribiremos siempre $P_{ij}(t)$ y las llamaremos funciones de transición. Estas probabilidades se pueden agrupar en una matriz de probabilidades de transición

$$(1.17) \quad P(t) = \begin{pmatrix} P_{00}(t) & P_{01}(t) & P_{02}(t) & \cdots \\ P_{10}(t) & P_{11}(t) & P_{12}(t) & \cdots \\ P_{20}(t) & P_{21}(t) & P_{22}(t) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Debido a que el tiempo es continuo, para el análisis de $P(t)$ va a ser necesario usar propiedades diferenciales con respecto a t , que nos den información sobre cómo evoluciona infinitesimalmente el proceso a medida que pasa el tiempo. La matriz de transición $P(t)$ cumple con las siguientes propiedades

1. $P_{ij}(t) \geq 0$ para todo $i, j \in \mathbf{E}$ y $P_{ij}(0) = \delta_{ij}$, la delta de Kronecker.
2. $\sum_{i \in \mathbf{E}} P_{ij} \leq 1$ para todo $t \geq 0, i \in \mathbf{E}$. El proceso $P_{ij}(t)$ se dice honesto si cumple $\sum_{j \in \mathbf{E}} P_{ij}(t) = 1$ para todo $t \geq 0, i \in \mathbf{E}$ y deshonesto en caso contrario.

3. Para $i, j \in \mathbf{E}$ y para todo $s, t \geq 0$, se cumple la ecuación de Chapman-Kolmogorov (o la propiedad de semigrupo):

$$(1.18) \quad P_{ij}(s+t) = \sum_{k \in \mathbf{E}} P_{ik}(s)P_{kj}(t).$$

La cuestión contraria también es cierta, es decir, si nos dan una matriz $P(t)$ con las tres propiedades anteriores, siempre es posible construir un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y una cadena de Markov a tiempo continuo $\{X_t, t \geq 0\}$ tal que sus correspondientes transiciones de probabilidad coinciden con las entradas $P_{ij}(t)$ de la matriz $P(t)$. Con lo cual el proceso queda totalmente determinado mediante una matriz de transición que cumple con las propiedades anteriores. Los procesos en los que nos centraremos serán siempre estables, significando que las trayectorias del proceso consistirán en funciones escalonadas que son continuas por la derecha (step functions). Esto es equivalente a que las funciones de transición $P_{ii}(t)$ tienen derivadas finitas de cualquier orden en $t = 0$ para cualquier $i \in \mathbf{E}$.

Aparte de las tres propiedades anteriores, supondremos siempre que las funciones de transición son estándar, es decir, se tiene la propiedad

$$\lim_{t \rightarrow 0} P_{ii}(t) = 1,$$

para todo $i \in \mathbf{E}$ (y por lo tanto, a raíz de la desigualdad $\sum_{j \neq i} P_{ij}(t) \leq 1 - P_{ii}(t)$, se tiene que $P_{ij}(t) \rightarrow \delta_{ij}$ a medida que $t \rightarrow 0$, para todo $i, j \in \mathbf{E}$).

En términos de la matriz de transición $P(t)$ las cuatro propiedades anteriores se pueden reescribir como

1. $P(t) \geq 0$ y $P(0) = I$.
2. $P(t)\mathbf{e} \leq \mathbf{e}$ para todo $t \geq 0$, donde \mathbf{e} es el vector columna con todas las componentes igual a 1.
3. $P(s+t) = P(s)P(t)$, para todo $s, t \geq 0$ (propiedad de semigrupo).
4. $\lim_{t \rightarrow 0} P(t) = I$.

Se puede probar (ver [12]) que $P(t)$ con las cuatro propiedades anteriores es siempre diferenciable en $t = 0$ (y en general se tendrá para cualquier $t \geq 0$). Es decir, se tiene la siguiente

Proposición 1.4.1. *Sea $P_{ij}(t)$ una función de transición. Entonces para $i \in \mathbf{E}$ se tiene que*

1. $-P'_{ii}(0) = \lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} = q_i$ (pero puede ser $+\infty$).
2. Si $q_i < \infty$ entonces $P'_{ij}(0) = \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} = q_{ij} < \infty$, para todo $j \neq i$. La existencia del límite también se cumple cuando $q_j < \infty$.

Una cadena de Markov a tiempo continuo se dice conservativa si

$$\sum_{j \neq i} q_{ij} = q_i < \infty \quad \text{para toda } i.$$

Un estado $i \in \mathbf{E}$ se dice que es estable si $q_i < \infty$ e instantáneo si $q_i = \infty$. El proceso se dirá estable si todos los estados son estables. Todos los procesos que estudiaremos serán estables, i se dice que es absorbente si $q_i = 0$, en cuyo caso $P_{ii}(t) = 1$ para todo $t \geq 0$. Obsérvese que $q_{ij} \geq 0$ mientras que $q_{ii} = -q_i \leq 0$.

Llamaremos operador infinitesimal del proceso a la matriz $\mathcal{A} = (a_{ij})$ con

$$(1.19) \quad a_{ij} = \begin{cases} q_{ij}, & i \neq j, \\ -q_i, & i = j \end{cases}$$

En otras palabras

$$(1.20) \quad \mathcal{A} = \begin{pmatrix} -q_0 & q_{01} & q_{02} & \cdots \\ q_{10} & -q_1 & q_{12} & \cdots \\ q_{20} & q_{21} & -q_2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Las entradas diagonales son no-positivas o posiblemente infinito, mientras que las diagonales no principales tienen entradas finitas y no-negativas. La suma de cada fila es no-positiva. Si todas las componentes diagonales son finitas, entonces \mathcal{A} se dice estable. Si además todas las filas suman 0 (i.e. $\sum_{j \in \mathbf{E}} a_{ij} = 0$ para todo $i \in \mathbf{E}$) entonces \mathcal{A} se dice conservativo.

A raíz de la Proposición 2.4.1 podemos derivar las ecuaciones de Kolmogorov, tanto la ecuación de retroceso como la ecuación de evolución o de Fokker-Planck.

Proposición 1.4.2. *La ecuación de retroceso de Kolmogorov está dada por*

$$(1.21) \quad P'_{ij}(t) = \sum_{k \in \mathbf{E}} q_{ik} P_{kj}(t), \quad t \geq 0, \quad i, j \in \mathbf{E}.$$

Demostración. Por la ecuación de Chapman-Kolmogorov y la Proposición 2.4.1 se tiene que

$$P'_{ij}(t) = \lim_{s \rightarrow 0} \frac{\partial}{\partial s} P_{ij}(s+t) = \lim_{s \rightarrow 0} \frac{\partial}{\partial s} \sum_{k \in \mathbf{E}} P_{ik}(s) P_{kj}(t) = \lim_{s \rightarrow 0} \sum_{k \in \mathbf{E}} P'_{ik}(s) P_{kj}(t) = \sum_{k \in \mathbf{E}} q_{ik} P_{kj}(t).$$

□

De igual manera tenemos la ecuación de evolución:

Proposición 1.4.3. *La ecuación de Kolmogorov o la ecuación de Fokker-Planck está dada por*

$$(1.22) \quad P'_{ij}(t) = \sum_{k \in \mathbf{E}} P_{ik}(t) q_{kj}, \quad t \geq 0, \quad i, j \in \mathbf{E}.$$

Demostración. Por la ecuación de Chapman-Kolmogorov y la Proposición 2.4.1 se tiene que

$$P'_{ij}(t) = \lim_{s \rightarrow 0} \frac{\partial}{\partial s} P_{ij}(s+t) = \lim_{s \rightarrow 0} \frac{\partial}{\partial s} \sum_{k \in \mathbf{E}} P_{ik}(s) P_{kj}(t) = \lim_{s \rightarrow 0} \sum_{k \in \mathbf{E}} P_{ik}(s) P'_{kj}(t) = \sum_{k \in \mathbf{E}} P_{ik}(t) q_{kj}.$$

□

A diferencia de la ecuación de retroceso, la ecuación de evolución puede que no se cumpla, ya que podría suceder que antes de llegar al tiempo t el proceso haya dado infinitos saltos. Esto puede suceder incluso en los casos donde el proceso es conservativo.

Las ecuaciones de Kolmogorov se pueden escribir de manera matricial como

$$(1.23) \quad P'(t) = \mathcal{A}P(t), \quad P'(t) = P(t)\mathcal{A}, \quad P(0) = I,$$

donde \mathcal{A} es el operador infinitesimal (1.20).

Las soluciones de las ecuaciones de Kolmogorov van a depender de cómo sea el operador infinitesimal \mathcal{A} . En la mayor parte de los casos se trabaja con \mathcal{A} como si fuese una matriz con ciertas propiedades. La matriz \mathcal{A} va a servir para describir el proceso para un número finito de saltos, pero no tiene por qué especificar el proceso de manera única. Con lo cual las ecuaciones de Kolmogorov puede que tengan más de una solución, si es que existen soluciones. Esta situación no pasa para el caso de cadenas de Markov a tiempo discreto, donde las soluciones existen y son únicas.

Una función no-negativa $P_{ij}(t)$ que satisface la ecuación de retroceso (o de evolución) de Kolmogorov tal que $\sum_{j \in \mathbf{E}} P_{ij}(t) \leq 1$ para todo $i \in \mathbf{E}$ y $t \geq 0$ se dice que es una pseudosolución de la ecuación de retroceso (o de evolución) de Kolmogorov. Si además $P_{ij}(t)$ satisface las ecuaciones de Chapman-Kolmogorov (1.18), entonces se dice que es una solución propia.

Teorema 1.4.1. *Sea \mathcal{A} estable pero no necesariamente conservativa. Entonces*

1. *Existe una función de transición $f_{ij}(t)$ (posiblemente deshonesto) que satisface ambas ecuaciones de Kolmogorov con la propiedad de que esta solución es minimal, en el sentido de que cualquier otra solución no-negativa de cualquiera de las ecuaciones de Kolmogorov cumple que $f_{ij}(t) \leq P_{ij}(t)$ para todos los estados $i, j \in \mathbf{E}$ y $t \geq 0$. Además $f_{ij}(t)$ es mínima en sentido de que $f'(0) = \mathcal{A}$.*
2. *Si \mathcal{A} es conservativa, la solución mínima $f_{ij}(t)$ de las ecuaciones de Kolmogorov es única si y sólo si es honesta. Además es la única que cumple que $f'(0) = \mathcal{A}$. Si \mathcal{A} no es conservativa entonces $f_{ij}(t)$ no puede ser honesta.*

Proposición 1.4.4. 1. *Si \mathcal{A} es conservativa, la solución minimal $f_{ij}(t)$ de la ecuación de retroceso de Kolmogorov es única si y sólo si la ecuación $\mathcal{A}x = \lambda x, 0 \leq x \leq 1$, es decir*

$$\sum_{j \in \mathbf{E}} q_{ij}x_j = \lambda x_i, \quad 0 \leq x_i \leq 1, \quad i \in \mathbf{E},$$

no tiene solución no-trivial para algún (y por lo tanto para todo) $\lambda > 0$. En este caso se dice que la matriz \mathcal{A} es regular.

2. *Sea \mathcal{A} no necesariamente conservativa pero uniformemente acotada (i.e. $\sup_{i \in \mathbf{E}} q_i < \infty$). Entonces la solución mínima $f_{ij}(t)$ asociada a \mathcal{A} es única.*

En términos de recurrencia, todas las definiciones asociadas a cadenas de Markov a tiempo discreto (comunicación, clases de comunicación, irreducibilidad, recurrencia, transitoriedad, vector invariante, etc.) se pueden extender de manera natural a cadenas de Markov a tiempo continuo. Así, un estado $i \in \mathbf{E}$ se dice recurrente si $\int_0^\infty P_{ii}(t)dt = \infty$, y transitorio en caso

contrario. Un estado $i \in \mathbf{E}$ se dice recurrente positivo si es recurrente y $\lim_{t \rightarrow \infty} P_{ii}(t) > 0$ y recurrente nulo si es recurrente y $\lim_{t \rightarrow \infty} P_{ii}(t) = 0$. Existe otra manera de definir la recurrencia en términos de las distribuciones de primeros tiempos de llegada:

$$(1.24) \quad F_{ij}(t) = \mathbb{P}(X_\tau = j \text{ para algún } \tau, 0 < \tau \leq t | X_0 = i), \quad i \neq j,$$

$$(1.25) \quad F_{ii}(t) = \mathbb{P}(X_{\tau_1} \neq i, X_{\tau_2} = i \text{ para algún } \tau_1, \tau_2, 0 < \tau_1 < \tau_2 \leq t | X_0 = i).$$

Podría ocurrir que alguna de estas distribuciones no sea honesta, i.e. $\int_0^\infty dF_{ij}(t) < 1$. La integral $\int_0^\infty dF_{ii}(t)$ es la probabilidad de que si se comienza en i , salga de i y vuelva de nuevo a i en tiempo finito. El estado i es recurrente si $\int_0^\infty dF_{ii}(t) = 1$ y transitorio en caso contrario. El estado i es recurrente positivo o nulo si es recurrente y si $\int_0^\infty t dF_{ii}(t)$ es finito o infinito, respectivamente.

Las cantidades $P_{ij}(t)$ y $F_{ij}(t)$ están relacionadas mediante las fórmulas

$$P_{ii}(t) = e^{-q_i t} + \int_0^t P_{ii}(t-s) dF_{ii}(s),$$

$$P_{ij}(t) = \int_0^t P_{jj}(t-s) dF_{ij}(s), \quad i \neq j.$$

Un vector $\pi = (\pi_i)_{i \in \mathbf{E}}$ se dice invariante (o estacionario) para la cadena de Markov a tiempo continuo si

$$\pi_j \geq 0 \quad \text{y} \quad \sum_{i \in \mathbf{E}} \pi P_{ij}(t) = \pi_j, \quad t \geq 0, \quad j \in \mathbf{E},$$

o, en versión vectorial, $\pi P(t) = \pi, t \geq 0$. En el caso en el que este vector se pueda normalizar a un vector de probabilidad (i.e. todas son componentes no negativas y $\sum_{i \in \mathbf{E}} \pi_i = 1$), entonces se dice que π es una distribución invariante (o estacionaria) de la cadena de Markov a tiempo continuo. La existencia de esta distribución invariante es equivalente a que la cadena de Markov a tiempo continuo sea irreducible y recurrente positiva.

Para cadenas de Markov a tiempo continuo, una herramienta para estudiar la recurrencia es considerar la transformada de Laplace de las funciones de transición $P_{ij}(t)$. En efecto

$$\hat{P}_{ij}(\lambda) = \int_0^\infty e^{-\lambda t} P_{ij}(t) dt, \quad l > 0, \quad i, j \in \mathbf{E}.$$

Esta función, en términos de λ , también se denomina función resolvente, y cumple las siguientes propiedades

1. $\hat{P}_{ij}(\lambda) \leq 0$ para todo $i, j \in \mathbf{E}$ y $\lambda > 0$.
2. $\lambda \sum_{j \in \mathbf{E}} \hat{P}_{ij}(\lambda) \leq 1$, para todo $i \in \mathbf{E}$ y $\lambda > 0$.
3. $\hat{P}_{ij}(\lambda) - \hat{P}_{ij}(\mu) + (\lambda - \mu) \sum_{k \in \mathbf{E}} \hat{P}_{ik}(\lambda) \hat{P}_{kj}(\mu) = 0$ para todo $i, j \in \mathbf{E}$ y $\lambda, \mu > 0$.
4. $\lim_{\lambda \rightarrow 0} \lambda \hat{P}_{ii}(\lambda) = 1$, para todo $i \in \mathbf{E}$ (y por lo tanto $\lim_{\lambda \rightarrow \infty} \lambda \hat{P}_{ij}(\lambda) = \delta_{ij}$).

La función resolvente $\hat{P}_{ij}(\lambda)$ se llamará honesta si se da la igualdad de la propiedad 2. En términos matriciales, la matriz resolvente $P(\lambda)$ cumple las siguientes propiedades

1. $\hat{P}(\lambda) \geq 0$ para todo $\lambda > 0$.
2. $\lambda \hat{P}(\lambda) \mathbf{e} \leq \mathbf{e}$, para todo $\lambda > 0$.
3. $\hat{P}(\lambda) - \hat{P}(\mu) + (\lambda - \mu) \hat{P}(\lambda) \hat{P}(\mu) = 0$ para todo $\lambda, \mu > 0$
4. $\lim_{\lambda \rightarrow \infty} \lambda \hat{P}(\lambda) = I$.

Procesos de nacimiento y muerte

Los ejemplos más sencillos sobre cadenas de Markov a tiempo continuo son los procesos de nacimiento y muerte. Los procesos de nacimiento y muerte pueden ser vistos como la versión continua de las caminatas aleatorias. Un conjunto de parámetros de nacimiento y muerte es una sucesión $\{(\lambda_n, \mu_n, n \geq 0)\}$ de pares de números tal que $\mu_0 \geq 0, \lambda_0 > 0, \lambda_i, \mu_i > 0, i \geq 1$. Un proceso de nacimiento y muerte X_t es un proceso de Markov con conjunto de parámetros $T = [0, \infty)$ en conjunto de estados $\mathbf{E} = \{-1, 0, 1, 2, \dots\}$ con probabilidades de transición estacionarias:

$$P_{ij}(t) = \mathbb{P}(X_{t+s} = j | X_s = i), \quad i, j \in \mathbf{E} \quad \text{independiente de } s.$$

Además, asumimos que $P_{ij}(t)$ satisface:

1. $P_{i,i+1}(t) = \lambda_i t + o(t)$ conforme $t \downarrow 0, i \in \mathbf{E}$.
2. $P_{i,i-1}(t) = \mu_i t + o(t)$ conforme $t \downarrow 0, i \geq 0$.
3. $P_{i,i}(t) = 1 - (\lambda_i + \mu_i)t + o(t)$ conforme $t \downarrow 0, i \in \mathbf{E}$.
4. $P_{ij}(0) = \delta_{ij}$.
5. $P_{-1,-1}(t) = 1, \quad P_{-1,i}(t) = 0, \quad t \geq 0, \quad i \neq -1$.

$$\mu_0 \geq 0, \quad \lambda_0 > 0, \quad \lambda_i, \mu_i > 0, \quad i \geq 1.$$

Los parámetros λ_i y μ_i son llamados tasas de nacimiento y muerte, respectivamente. Si $\mu_0 > 0$ tenemos un estado absorbente -1 , y una vez que entramos ahí no podemos dejar jamás ese estado. Si $\mu_0 = 0$ tenemos un estado reflejado. Después de entrar al estado 0 siempre regresaremos al estado 1 después de algún tiempo. En este caso el estado -1 jamás es alcanzado.

Normalmente, el espacio de estados en el que se trabaja es $\mathbf{E} = \mathbb{N}$, pero también están espacios de estados infinitos o doblemente infinito (en cuyo caso se llaman procesos de nacimiento y muerte bilaterales). La matriz correspondiente al operador infinitesimal \mathcal{A} viene dada por

$$(1.26) \quad \mathcal{A} = \begin{pmatrix} -(\lambda_0 + \mu_0) & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & 0 & \cdots \\ 0 & 0 & \mu_3 & -(\lambda_3 + \mu_3) & \lambda_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

\mathcal{A} es conservativa si y sólo si $\mu_0 = 0$. La ecuación de retroceso de Kolmogorov está dada por

$$\begin{aligned} P'_{0,j}(t) &= -(\lambda_0 + \mu_0)P_{0,j}(t) + \lambda_0 P_{1,j}(t), \\ P'_{i,j}(t) &= \mu_i P_{i-1,j}(t) - (\lambda_i + \mu_i)P_{i,j}(t) + \lambda_i P_{i+1,j}(t), \quad i \geq 1, \end{aligned}$$

mientras que la ecuación de evolución de Kolmogorov está dada por

$$\begin{aligned} P'_{i,0}(t) &= -(\lambda_0 + \mu_0)P_{i,0}(t) + \mu_1 P_{i,1}(t), \\ P'_{i,j}(t) &= \lambda_{j-1} P_{i,j-1}(t) - (\lambda_j + \mu_j)P_{i,j}(t) + \mu_{j+1} P_{i,j+1}(t), \quad j \geq 1. \end{aligned}$$

Al igual que en el caso de caminatas aleatorias, podemos definir los coeficientes potenciales de la siguiente manera

$$\pi_0 = 1, \quad \pi_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i}, \quad i \geq 1.$$

Estos coeficientes potenciales verifican las ecuaciones de simetría

$$(1.27) \quad \pi_n \lambda_n = \pi_{n+1} \mu_{n+1}, \quad n \geq 0,$$

a raíz de buscar un vector π que haga reversible a la cadena, i.e. $\pi_i P_{ij}(t) = \pi_j P_{ji}(t)$, y usando las ecuaciones de Kolmogorov. \mathcal{A} con esta propiedad se dice que es débilmente simétrica. También se puede comprobar que $\pi \mathcal{A} = 0$, con lo cual π es un vector invariante del proceso de nacimiento y muerte, que se convertirá en distribución siempre y cuando $\sum_{n=0}^{\infty} \pi_n < \infty$.

Un resultado importante es el siguiente

Teorema 1.4.2. *1. Sea \mathcal{A} definido en (1.26) y asúmase que $\lambda_n > 0$ para todo $n \geq 1$. Entonces la solución minimal de $f_{ij}(t)$ asociada a \mathcal{A} de la ecuación de retroceso de Kolmogorov es única si y sólo si $C = \infty$, donde*

$$(1.28) \quad C = \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{i=0}^n \pi_i = \infty.$$

Suponemos las siguientes condiciones sobre los parámetros λ_i y μ_i :

$$(1.29) \quad C = \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{i=0}^n \pi_i = \infty,$$

$$(1.30) \quad D = \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \sum_{i=n+1}^{\infty} \pi_i = \infty,$$

donde los coeficientes potenciales π_n están dados por

$$\pi_i = \frac{\lambda_0 \lambda_1 \dots \lambda_{i-1}}{\mu_1 \mu_2 \dots \mu_i}, \quad i \geq 1$$

la medida invariante del proceso, i.e.

$$\pi\mathcal{A} = 0,$$

y $\pi_0 = 1$. En cierto sentido no permitimos alcanzar el infinito en un tiempo finito. La cantidad C puede ser interpretada como el promedio de la primera ida de 0 a ∞ y D puede ser visto como el tiempo promedio de la primera ida de ∞ a 0.

Notación $A/B/c$ para colas

Para procesos “queueing” o de colas, usamos la notación $A/B/c$, donde c es el número de servidores, A es la distribución del número de llegadas y B es la distribución de servicios. Los símbolos usados en los primeros dos espacios son $G = GI$, independiente general; M , inter llegadas exponenciales o tiempos de servicio; E_k (Erlangian, inter llegada o tiempos de servicio gamma distribuido de orden k , de manera que $E_1 = M$); y D (determinista), un horario fijo de llegadas o de servicios en intervalos fijos.

1.5. Procesos de difusión

Sea τ una variable aleatoria no negativa asociada con un proceso $\{X_t, 0 \leq t < \infty\}$; en otras palabras, asociamos con cada función muestra X_t un número no negativo que denotamos por $\tau(X_t)$

Definición 1.5.1. *La variable aleatoria τ es un tiempo de Markov o tiempo de paro relativo a $\{X_t\}$ si tiene la siguiente propiedad:*

Si X_t y Y_t son dos funciones muestra del proceso tales que $X_t = Y_t$ para $0 \leq t \leq s$ y $\tau(X_t) < s$, entonces $\sigma(X_t) = \sigma(Y_t)$.

Definición 1.5.2. *Si para cada tiempo de paro τ , la distribución de probabilidad de*

$$X_{t_1+\tau}, X_{t_2+\tau}, \dots, X_{t_k+\tau} \quad (t_1 < t_2 < \dots < t_k)$$

dado $X_s, s \leq \tau$ y $X_\tau = x$, es idéntica con la distribución de probabilidad de

$$X_{t_1}, X_{t_2}, \dots, X_{t_k} \quad (t_1 < \dots < t_k)$$

dado $X_0 = x$, entonces se dice que el proceso de Markov tiene la propiedad fuerte de Markov.

Definición 1.5.3. *Un proceso estocástico a tiempo continuo que posee la propiedad fuerte de Markov y para la cual las trayectorias X_t son (casi seguramente) funciones continuas de t se llama **proceso de difusión**.*

Consideremos un proceso de difusión $\{X_t, t \geq 0\}$ cuyo espacio de estados es un intervalo I , con puntos terminales $l < r$ con la posibilidad de que $l = -\infty$ o $r = \infty$. Tal proceso se llama regular si empezando de cualquier punto en el interior de I cualquier otro punto en el interior de I puede ser alcanzado con probabilidad positiva. Más precisamente, para cualquier z en el intervalo I , sea T_z la variable aleatoria igual al primer tiempo que el proceso toma el valor z . En el caso en que el evento z nunca es alcanzado, por convención hacemos $T_z = \infty$.

Definición 1.5.4. *Un proceso de difusión se llama **regular** si*

$$\mathbb{P}(T_z < \infty | X_0 = x) > 0$$

siempre que $l < x, z < r$.

A partir de ahora sólo consideraremos procesos de difusión regulares.

La mayoría de los procesos de Markov, incluyendo el proceso de Poisson (proceso que es un proceso de nacimiento puro, i.e. $\mu_i = 0$ para toda $i \geq 1$ y $\lambda_i = \lambda$ para toda $i \geq 0$), y los procesos de nacimiento y muerte, satisfacen la propiedad

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(|X_h - x| > \epsilon | X_0 = x) = \lambda(x, \epsilon)$$

con $\lambda(x, \epsilon)$ no negativo y posiblemente positivo para ϵ pequeño. De hecho, para el proceso de Poisson tenemos

$$(1.31) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(X_h - i = 1 | X_0 = i) = \lambda, \quad i = 0, 1, \dots$$

donde λ es la tasa media de ocurrencia de los eventos.

En contraste con (1.31) todo proceso de difusión cumple con la siguiente condición:
Para cada $\epsilon > 0$,

$$(1.32) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(|X_{t+h} - x| > \epsilon | X_t = x) = 0, \quad \text{para todo } x \in I.$$

Ésta relación afirma que los desplazamientos largos, que exceden un orden fijo ϵ , son poco probables sobre intervalos suficientemente pequeños de tiempo. Este hecho puede ser visto como una formalización de la propiedad de que las trayectorias del proceso son continuas. Más adelante veremos un teorema que afirma parcialmente el regreso: un proceso de Markov que cumple con (1.32) y con una condición apropiada de uniformidad es un proceso de difusión.

Una gran cantidad de procesos de difusión está caracterizada por dos condiciones básicas además de (1.32) y describen la media y varianza de los desplazamientos infinitesimales. Sea $\Delta_h X(t)$ el incremento en el proceso sobre un intervalo de longitud h . Así $\Delta_h X(t) = X_{t+h} - X_t$. Estas condiciones afirman la existencia de los límites

$$(1.33) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[\Delta_h X_t | X_t = x] = \mu(x, t),$$

y

$$(1.34) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[\{\Delta_h X_t\}^2 | X_t = x] = \sigma^2(x, t),$$

siempre que $l < x < r$. Las funciones $\mu(x, t)$ y $\sigma^2(x, t)$ son llamadas los parámetros infinitesimales del proceso, y, en particular, $\mu(x, t)$ es llamada el parámetro deriva, la media infinitesimal, o el desplazamiento infinitesimal esperado y $\sigma^2(x, t)$ el parámetro de difusión, o varianza infinitesimal.

Generalmente, $\mu(x, t)$ y $\sigma^2(x, t)$ son funciones continuas de x y t y un proceso regular tiene por lo general $\sigma^2(x, t) > 0$ para toda $l < x < r$ y $t > 0$. Se asume que en las relaciones

(1.33) y (1.34) los momentos existen. Una versión más general presenta estas relaciones en términos de los momentos truncados donde $\Delta_h X_t$ es reemplazado por

$$\Delta_{h,\epsilon} X_t = \begin{cases} \Delta_h X_t & \text{si } |\Delta_h X_t| \leq \epsilon, \\ 0 & \text{de lo contrario} \end{cases}$$

Formalmente, las condiciones (1.33) y (1.34) son reemplazadas por

$$(1.35) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[\Delta_{h,\epsilon} X_t | X_t = x] = \mu(x, t)$$

y

$$(1.36) \quad \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[\{\Delta_{h,\epsilon} X_t\}^2 | X_t = x] = \sigma^2(x, t)$$

respectivamente, y de nuevo se debe añadir (1.32) a (1.35) y (1.36) para asegurar que el proceso es una difusión.

En adición a las relaciones (1.33) y (1.34), las siguientes relaciones de momentos de orden mayor se satisfacen:

$$(1.37) \quad \lim_{h \downarrow 0} \frac{\mathbb{E}[|\Delta_h X_t|^r | X_t = x]}{h} = 0, \quad r = 3, 4, \dots$$

Hay otras maneras más generales de definir procesos de difusión, como por ejemplo la siguiente.

Para todo $z > 0$ y a medida que $t \rightarrow 0^+$ se supone que

$$(1.38) \quad \begin{aligned} \mathbb{E}[(X_{s+t} - X_s) \mathcal{X}_{\{|X_{s+t} - X_s| \leq \epsilon\}} | X_s = x] &= t\mu(x) + o(t), \\ \mathbb{E}[(X_{s+t} - X_s)^2 \mathcal{X}_{\{|X_{s+t} - X_s| \leq \epsilon\}} | X_s = x] &= t\sigma^2(x) + o(t), \\ \mathbb{P}(|X_{s+t} - X_s| > \epsilon | X_s = x) &= o(t) \end{aligned}$$

donde $\mathcal{X}_B(x)$ es la función indicadora en el conjunto de Borel B , i.e.

$$\mathcal{X}_B(x) = \begin{cases} 1, & \text{si } x \in B \\ 0, & \text{en otro caso} \end{cases}.$$

Sea ahora $p(t; x, dy) = \mathbb{P}(X_t \in dy | X_0 = x)$.

Si $p(t; x, dy)$ tiene una densidad $p(t; x, y)$, entonces para cualquier subconjunto de Borel $B \subset \mathcal{S}$ se tiene que

$$p(t; x, B) = \int_B p(t; x, y) dy.$$

Para que exista dicha densidad es suficiente que las funciones $\mu(x)$ y $\sigma(x)$ sean continuamente diferenciables con derivadas acotadas en \mathcal{S} , así como que $\sigma''(x)$ exista, sea continua y $\sigma^2(x) > 0$ para todo $x \in \mathcal{S}$.

Consideremos el conjunto $\mathcal{B}(\mathcal{S})$ de todas las funciones medibles Borel, reales y acotadas sobre \mathcal{S} . En este espacio se define el siguiente operador de transición

$$(1.39) \quad (T_t f)(x) = \mathbb{E}[f(X_t)|X_0 = x] = \int_{\mathcal{S}} f(y)p(t; x, dy), \quad t > 0.$$

Entonces $T_t : \mathcal{B}(\mathcal{S}) \rightarrow \mathcal{B}(\mathcal{S})$ es un operador lineal acotado en $\mathcal{B}(f)$ dotado con la norma del supremo $\|f\| = \sup\{|f(y)| : y \in \mathcal{S}\}$ y $(T_0 f)(x) = f(x)$. En efecto, T_t es una contracción, i.e. $\|T_t f\| \leq \|f\|$ para todo $f \in \mathcal{B}(\mathcal{S})$, ya que

$$|(T_t f)(x)| = \left| \int_{\mathcal{S}} f(y)p(t; x, dy) \right| \leq \|f\|.$$

La familia de operadores de transición $\{T_t : t > 0\}$ tiene la propiedad de semigrupo $T_{s+t} = T_s T_t$, pues usando la propiedad $\mathbb{E}[X] = \mathbb{E}\{X|Y\}$, se tiene que

$$(1.40) \quad \begin{aligned} (T_{s+t} f)(x) &= \mathbb{E}[f(X_{s+t})|X_0 = x] = \mathbb{E}[\mathbb{E}[f(X_{s+t})|X_s]|X_0 = x] \\ &= \mathbb{E}[(T_t f)(X_s)|X_0 = x] = T_s(T_t f)(x). \end{aligned}$$

Esto implica que el comportamiento de T_t queda completamente determinado por el comportamiento de T_t cerca de $t = 0$.

Definimos el operador infinitesimal \mathcal{A} de $\{T_t, t > 0\}$, como el operador lineal definido por

$$(1.41) \quad (\mathcal{A}f)(x) = \lim_{t \rightarrow 0^+} \frac{(T_t f)(x) - f(x)}{t},$$

para todas las funciones f en $\mathcal{B}(\mathcal{S})$ tal que la parte derecha converge a alguna función uniformemente en x . La clase de todas estas funciones f componen el dominio $\mathcal{D}_{\mathcal{A}}$ de \mathcal{A} .

Proposición 1.5.1. *Sea f una función acotada y dos veces continuamente diferenciable. Para $x \in \mathcal{S}$ se tiene que*

$$(\mathcal{A}f)(x) = \lim_{t \rightarrow 0^+} \frac{(T_t f)(x) - f(x)}{t} = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x),$$

donde $\mu(x)$ y $\sigma^2(x)$ son los coeficientes deriva y de difusión, respectivamente, definidos en (1.38).

Lo anterior no prueba que cualquier función acotada f y dos veces diferenciable pertenece al dominio $\mathcal{D}_{\mathcal{A}}$, ya que el límite (1.41) no se ha demostrado que sea uniforme en x . Para eliminar esta situación basta con suponer que f es dos veces continuamente diferenciable y se anula fuera de un subintervalo acotado y cerrado de \mathcal{S} . En esta situación se tiene que $f \in \mathcal{D}_{\mathcal{A}}$ y para dicha f , el operador infinitesimal \mathcal{A} es el operador diferencial de segundo orden

$$(1.42) \quad (\mathcal{A}f)(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x).$$

Proposición 1.5.2. *Sea $\{T_t, t > 0\}$ una familia de operadores de transición para un proceso de Markov sobre un espacio de estados \mathcal{S} . Si $f \in \mathcal{D}_{\mathcal{A}}$, entonces se tiene la siguiente ecuación de retroceso de Kolmogorov*

$$(1.43) \quad \frac{\partial}{\partial t}(T_t f)(x) = \mathcal{A}(T_t f)(x), \quad t > 0 \quad x \in \mathcal{S}.$$

Demostración. Si $f \in \mathcal{D}_{\mathcal{A}}$, usando la conmutatividad de T_t con T_s y el hecho de que T_t es una contracción, se tiene, a medida que $s \rightarrow 0^+$, que

$$\left\| \frac{T_s(T_t f) - T_t f}{s} - T_t(\mathcal{A}f) \right\| = \left\| T_t \left(\frac{T_s f - f}{s} - \mathcal{A}f \right) \right\| \leq \left\| \frac{T_s f - f}{s} - \mathcal{A}f \right\| \rightarrow 0.$$

Con lo cual $T_t f \in \mathcal{D}_{\mathcal{A}}$, ya que converge a $T_t(\mathcal{A}f)$, y por lo tanto, por la definición de \mathcal{A} se llega al resultado. \square

Sea $p(t; x, y)$ la densidad de probabilidades de transición de un proceso de difusión $\{X_t, t \geq 0\}$ y consideramos $f = \mathcal{X}_B$ en (1.40). Se tiene entonces que

$$\int_{\mathcal{B}} p(s+t; x, y) dy = \int_{\mathcal{S}} \left(\int_{\mathcal{B}} p(t; z, y) dy \right) p(s; x, z) dz = \int_{\mathcal{B}} \left(\int_{\mathcal{S}} p(t; z, y) p(s; x, z) dz \right) dy,$$

para cualquier conjunto de Borel \mathcal{B} en \mathcal{S} . Esto implica la ecuación de Chapman-Kolmogorov

$$p(s+t; x, y) = \int_{\mathcal{S}} p(t; z, y) p(s; x, z) dz = (T_s f)(x),$$

con $f(z) = p(t; z, y)$. Aplicando (1.41) a esta f se tiene la ecuación de retroceso de Kolmogorov para la densidad de probabilidad de transición $p(t; x, y)$:

$$\frac{\partial p(t; x, y)}{\partial t} = \mu(x) \frac{\partial p(t; x, y)}{\partial x} + \frac{1}{2} \sigma^2(x) \frac{\partial^2 p(t; x, y)}{\partial x^2}.$$

La ecuación de evolución de Kolmogorov o la ecuación de Fokker-Planck gobierna la densidad de probabilidad de X_t cuando X_0 tiene una distribución inicial arbitraria. Supongamos que llamamos a la densidad de esta distribución inicial g . La ecuación de evolución puede ser derivada mediante el adjunto del operador T . La densidad de X_t está dada por

$$(T_t^* g)(y) = \int_{\mathcal{S}} g(x) p(t; x, y) dx.$$

El operador T_t^* transforma una densidad g en otra densidad de probabilidad. Este operador es el adjunto de T_t con respecto al producto interno $(u, v) = \int u(x)v(x)dx$. En efecto,

$$(T_t^* g, f) = \int_{\mathcal{S}} (T_t^* g)(y) f(y) dy = \int_{\mathcal{S}} g(x) (T_t f)(x) dx = (g, T_t f).$$

Si f es dos veces continuamente diferenciable y se anula fuera de un subconjunto compacto de \mathcal{S} , podemos tomar derivada con respecto a t en la relación anterior, y usando (1.43) y que \mathcal{A} conmuta con T_t se llega a

$$(1.44) \quad \left(\frac{\partial}{\partial t} T_t^* g, f \right) = \left(g, \frac{\partial}{\partial t} T_t f \right) = (g, T_t \mathcal{A}f) = (T_t^* g, \mathcal{A}f) = \int_{\mathcal{S}} (T_t^* g)(y) (\mathcal{A}f)(y) dy.$$

Ahora bien, si f, h son ambas dos veces continuamente diferenciables que se anulan fuera de un subintervalo finito se tiene, haciendo integración por partes, que

$$\begin{aligned} (h, \mathcal{A}) &= \int_{\mathcal{S}} h(y) \left[\mu(y) f'(y) + \frac{1}{2} \sigma^2(y) f''(y) \right] dy \\ &= \int_{\mathcal{S}} \left[-\frac{d}{dy} (\mu(y) h(y)) + \frac{1}{2} \frac{d^2}{dy^2} (\sigma^2(y) h(y)) \right] f(y) dy = (\mathcal{A}^* h, f), \end{aligned}$$

donde \mathcal{A}^* es el operador adjunto del operador diferencial \mathcal{A} . Por lo tanto, aplicándoselo a (1.44), se llega a

$$\left(\frac{\partial}{\partial t} T_t^* g, f \right) = (\mathcal{A}^* T_t^* g, f).$$

O equivalentemente,

$$\int_{\mathcal{S}} \frac{\partial p(t; x, y)}{\partial t} g(x) dx = \int_{\mathcal{S}} \left(-\frac{d}{dy} [\mu(y)p(t; x, y)] + \frac{1}{2} \frac{d^2}{dy^2} [\sigma^2(y)p(t; x, y)] \right) g(x) dx.$$

Como esta relación se verifica para suficientes funciones g que son dos veces continuamente diferenciables y que se anulan fuera de algún intervalo cerrado contenido en \mathcal{S} , se tiene la ecuación de Kolmogorov para la densidad de probabilidad de transición $p(t; x, y)$:

$$(1.45) \quad \frac{\partial p(t; x, y)}{\partial t} = -\frac{d}{dy} [\mu(y)p(t; x, y)] + \frac{1}{2} \frac{d^2}{dy^2} [\sigma^2(y)p(t; x, y)], \quad t > 0.$$

Sea $\{X_t, t \geq 0\}$ un proceso de difusión en un espacio de estados $\mathcal{S} = [a, b]$ con $-\infty \leq a < x < b \leq \infty$. Denotamos a τ_y el primer tiempo de llegada a un estado y definido por

$$(1.46) \quad \tau_y = \inf\{t \geq 0; X_t = y\}.$$

También denotaremos para $[c, d] \subset \mathcal{S}$

$$\tau = \tau_{c,d} = \inf\{t \geq 0; X_t \notin (c, d)\} = \tau_c \tau_d$$

Diremos que un punto frontera a es inaccesible o reflectante si no puede ser alcanzado desde ningún punto en el interior. Esto hace que si alguna trayectoria del proceso está cerca de un estado inaccesible, inmediatamente manda el proceso hacia el interior. Diremos que un punto frontera a es absorbente si el proceso de difusión X_t se comporta como \hat{X}_t , donde

$$\hat{X}_t = \begin{cases} X_t, & \text{si } t < \tau_a, \\ a = X_{\tau_a}, & \text{si } t \geq \tau_a. \end{cases}$$

La distribución de probabilidades de transición $p(t; x, dy)$ se divide entonces en dos. Una distribución $p^0(t, x, dy)$ definida en el intervalo $(a, b]$ y la probabilidad de absorción en el extremo a . Por lo tanto (suponiendo que b no es absorbente)

$$p(t; x, B) = \begin{cases} p^0(t; x, B) = \mathbb{P}(X_t \in B \text{ y } \tau_a > t | X_0 = a), & \text{si } B \subset (a, b) \\ \mathbb{P}(\tau_a \leq t | X_0 = x), & \text{si } B = \{a\}. \end{cases}$$

En este caso esperamos una condición de contorno de tipo Dirichlet en $x = a$ en la ecuación de retroceso de Kolmogorov dada por

$$\lim_{x \rightarrow a^+} p^0(t; x, y) = 0,$$

ya que si x comienza desde el estado absorbente hay prácticamente probabilidad 0 de que llegue a un punto del interior. Igual para la ecuación de evolución, donde la condición de contorno en $y = a$ está dada por

$$\lim_{y \rightarrow a^+} p^0(t; x, y) = 0.$$

La distribución de probabilidad sólo aplica al interior del intervalo. Si queremos calcular la probabilidad de absorción $p(t; x, \{a\})$ está dada en términos de $p^0(t; x, y)$ por

$$(1.47) \quad p(t; x, \{a\}) = v(x) - \int_a^b v(y)p^0(t; x, y)dy, \quad t > 0, \quad a < x < b,$$

donde $v(x) = \mathbb{P}(\tau_a < \tau_b | X_0 = x)$.

Para estudiar la recurrencia y transitoriedad de difusiones en términos del operador infinitesimal \mathcal{A} consideramos un proceso de difusión $\{X_t, t \geq 0\}$ con $X_0 = x$ y un intervalo cerrado $[c, d] \subset \mathcal{S}, c < d$. Sea

$$v(x) = v_{c,d}(x) = \mathbb{P}(\tau_c < \tau_d | X_0 = x), \quad c \leq x \leq d,$$

donde τ_y está definido en (1.46). Para un intervalo pequeño de tiempo h , usando la ley de la probabilidad total y el Teorema de Taylor, suponiendo que $v(x)$ es dos veces continuamente diferenciable, se tiene que

$$\begin{aligned} v(x) &= \mathbb{E}[v(X_h) | X_0 = x] + o(h) = \mathbb{E}[v(x + X_h - x) | X_0 = x] \\ &= \mathbb{E} \left[v(x) + (X_h - x)v'(x) + \frac{1}{2}(X_h - x)^2v''(x) | X_0 = x \right] + o(h) \\ &= v(x) + \mu(x)hv'(x) + \frac{1}{2}\sigma^2(x)hv''(x) + o(h). \end{aligned}$$

Se tiene así la siguiente

Proposición 1.5.3. *Las probabilidades $v(x)$ son solución de la siguiente ecuación diferencial con valores iniciales*

$$\begin{aligned} \frac{1}{2}\sigma^2(x)v''(x) + \mu(x)v'(x) &= 0, \quad c < x < d, \\ v(c) &= 1, \quad v(d) = 0. \end{aligned}$$

De manera general se puede definir

$$p_{xy} = \mathbb{P}(\tau_y < \infty | X_0 = x).$$

Un estado u es recurrente si $p_{xy} = 1$ para todo $x \in \mathcal{S}$ tal que $p_{yx} > 0$, y transitorio en caso contrario. Si todos los estados de \mathcal{S} son recurrentes, entonces la difusión se dice recurrente.

Queremos calcular ahora el tiempo medio de escape $E(x)$ de una difusión que comienza en $X_0 = x$, con $x \in (c, d)$, i.e.

$$E(x) = E_{c,d} = \mathbb{E}[\tau_{c,d} | X_0 = x],$$

Al igual que antes, para un intervalo pequeño de tiempo h y asumiendo que $E(x)$ es dos veces continuamente diferenciable, se tiene que

$$E(x) = h + \mathbb{E}[E(X_h) | X_0 = x] + o(h) = h + E(x) + \mu(x)hE'(x) + \frac{1}{2}\sigma^2(x)hE''(x) + o(h).$$

Se tiene entonces la siguiente:

Proposición 1.5.4. *El tiempo medio de escape $E(x)$ es solución de la siguiente ecuación diferencial con valores iniciales*

$$\frac{1}{2}\sigma^2(x)E''(x) + \mu(x)E'(x) = -1, \quad c < x < d,$$

$$E(c) = E(d) = 0.$$

Una difusión recurrente se dice recurrente positiva si $\mathbb{E}[\tau_y|X_0 = x] < \infty$ para todo $x, y \in \mathcal{S}$ y recurrente nula en caso contrario.

Cuando tengamos una difusión recurrente positiva, va a existir una distribución invariante $m(y)$ que necesariamente tiene que satisfacer

$$m(y) = (T_t^* m)(y) = \int_{\mathcal{S}} m(x)p(t; x, y)dx, \quad t > 0.$$

En analogía con los casos de cadenas de Markov, esta distribución estacionaria debe cumplir

$$(1.48) \quad \lim_{t \rightarrow \infty} p(t; x, y) = m(y).$$

Procesos de difusión terminales

Un proceso es llamado un proceso de difusión terminal si las trayectorias se comportan como aquellas de un proceso de difusión regular hasta un tiempo posiblemente aleatorio ζ cuando el proceso acaba.

Escribimos un proceso de difusión terminal en la forma $\{X_t, 0 \leq t < \zeta\}$. Como permitimos el caso $\zeta = \infty$, esto incluye el caso en que el proceso no termina. Si $\zeta = \infty$ desde todos los puntos iniciales $X_0 = x$, el proceso se dice ser conservativo, y escribimos $\{X_t, t \geq 0\}$. Esto es, un proceso de difusión regular $\{X_t\}$ en I es conservativo si

$$\mathbb{P}(X_t \in I | X_0 = x) = \mathbb{P}(\zeta > t | X_0 = x) = 1$$

para toda $t \geq 0$ y $x \in I$.

Para un proceso de difusión terminal, en cada punto x existe una probabilidad $k(x, t)dt + o(dt)$ de que el proceso acabe sobre una duración infinitesimal de tiempo $(t, t + dt)$, y probabilidad $1 - k(x, t)dt + o(dt)$ de que el proceso no acabe. Su formulación rigurosa requiere del concepto de funcionales multiplicativos estocásticos. Notemos que la tasa $k(x, t)$ puede depender de la posición y del tiempo, si permitimos que $k(x, t)$ sea una función tanto del tiempo t como de la posición x . En símbolos,

$$\lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}(t < \zeta < t + h | X_t = x) = k(x, t).$$

Primeros tiempos de llegada

Los tiempos de llegada son puntos y conjuntos que juegan un papel fundamental en el estudio de procesos de difusión de dimensión 1. Formalmente definimos un tiempo de llegada del proceso $\{X_t, 0 \leq t < \zeta\}$ al nivel z por

$$T_z = \begin{cases} \infty & \text{si } X_t \neq z \text{ para } 0 \leq t < \zeta, \\ \inf\{t \geq 0; X_t = z\} & \text{de otra forma.} \end{cases}$$

Por conveniencia escribimos en algunos casos $T(z)$ en lugar de T_z y usamos la notación

$$T^* = T_{a,b} = T(a,b) = \min\{T(a), T(b)\} = T(a) \wedge T(b)$$

para el tiempo de llegada de a o b , el primer tiempo $X_t = a$ o $X_t = b$. Para procesos que empiezan en $X_0 = x$ en (a,b) , esto es lo mismo que el tiempo de salida del intervalo (a,b) :

$$T(a,b) = \inf\{t \geq 0; X_t \notin (a,b)\}, \quad X_0 = x \text{ en } (a,b).$$

Condiciones para que un proceso sea un proceso de difusión

Es de mucha utilidad tener condiciones suficientes sobre las cuales un proceso de Markov es un proceso de difusión. Para esto, introducimos el concepto de proceso estándar. Un proceso de Markov $\{X_t, t \geq 0\}$ es llamado un proceso estándar si las trayectorias poseen las siguientes propiedades de regularidad:

1. X_t es continua por la derecha; i.e., para toda $s \geq 0$

$$\lim_{t \downarrow s} X_t = X_s;$$

2. los límites por la izquierda de X_t existen; i.e.,

$$\lim_{t \uparrow s} X_t \text{ existe para toda } s > 0;$$

y

3. X_t es continua desde la izquierda a través de los tiempos de Markov; i.e., si $T_1 \leq T_2 \leq \dots$ son tiempos de Markov que convergen a $T \leq \infty$, entonces el límite $\lim_{n \rightarrow \infty} X_{T_n} = X_T$ siempre que $T < \infty$.

La tercera condición es llamada algunas veces continuidad quasi-izquierda. No implica que las trayectorias son continuas por la izquierda, sólo que los saltos no pueden ser predecidos. En particular, una trayectoria de un proceso estándar puede en el peor de los casos exhibir discontinuidades del primer tipo (saltos).

Un resultado fuerte es que todo proceso fuerte de Markov $\{X_t, t \geq 0\}$ continuo en probabilidad y sujeto a otras condiciones, posee una versión equivalente $\{\hat{X}_t, t \geq 0\}$ que es un proceso estándar. (Los procesos \hat{X}_t y X_t son equivalentes si \hat{X}_t y X_t tienen mismas distribuciones dimensionalmente finitas). Recordemos que un proceso estocástico es continuo en probabilidad si para toda $\epsilon > 0$ y $s \geq 0$

$$(1.49) \quad \lim_{t \rightarrow s} \mathbb{P}(|X_t - X_s| > \epsilon) = 0.$$

La propiedad (1.49) es bastante débil y se satisface en la mayoría de los modelos estocásticos, así que para propósitos prácticos asumiremos que el proceso de Markov es en primera instancia un proceso estándar.

Una condición suficiente para que un proceso estándar sea un proceso de difusión es la condición de Dynkin:

$$(1.50) \quad \frac{1}{h} \mathbb{P}(|X_{t+h} - X_t| > \epsilon | X_t = x) \rightarrow 0 \quad \epsilon > 0,$$

conforme $h \downarrow 0$, donde la convergencia permanece uniforme para x restringido a cualquier subintervalo compacto de (l, r) y t en cualquier intervalo finito $[0, N]$.
Es así que tenemos el siguiente

Teorema 1.5.1. *Sea $\{X_t, t \geq 0\}$ un proceso estándar y supongamos que la condición de Dynkin se satisface. Entonces $\{X_t, t \geq 0\}$ es un proceso de difusión.*

Capítulo 2

Polinomios Ortogonales

A continuación damos la definición de algunas funciones que estaremos usando con frecuencia durante el desarrollo de la tesis.

La función Gamma

La función gamma está dada por

$$(2.1) \quad \Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt, \quad \operatorname{Re}(z) > 0$$

Propiedades:

1. $\Gamma(1) = 1$,
2. $\Gamma(z+1) = z\Gamma(z)$ con $\operatorname{Re}(z) > 0$. Así, para z entero

$$\Gamma(z+1) = z!$$

3.

$$(2.2) \quad \Gamma(z) = 2 \int_0^{\infty} e^{-t^2} t^{2z-1} dt$$

4.

$$(2.3) \quad \int_0^{\pi/2} \cos^{2x-1} \theta \sin^{2y-1} \theta d\theta = \frac{\Gamma(x)\Gamma(y)}{2\Gamma(x+y)}$$

Función Beta

La función beta se define por

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt, \quad \operatorname{Re}(x), \operatorname{Re}(y) > 0.$$

Sin embargo, haciendo $t = \cos^2 \theta$, $dt = -2 \cos \theta \sin \theta d\theta$ tenemos

$$B(x, y) = \int_{\pi/2}^0 (\cos^2 \theta)^{x-1} (\sin^2 \theta)^{y-1} (-2 \cos \theta \sin \theta d\theta) = 2 \int_0^{\pi/2} \cos^{2x-1} \theta \sin^{2y-1} \theta d\theta$$

y aplicando (2.3) obtenemos que

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Símbolo de Pochhammer

Está definido de la siguiente manera:

$$(a)_n = \begin{cases} 1 & \text{si } n = 0 \\ a(a+1)\dots(a+n-1) & \text{si } n = 1, 2, 3, \dots \end{cases}$$

En términos de la función Gamma, tenemos

$$(a)_n = \frac{\Gamma(a+n)}{\Gamma(a)}, \quad n \geq 0.$$

La serie hipergeométrica generalizada

La serie hipergeométrica generalizada ${}_pF_q$ está definida por

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q; z) = \sum_{j=0}^{\infty} \frac{(a_1)_j \dots (a_p)_j}{(b_1)_j \dots (b_q)_j} \frac{z^j}{j!}$$

donde $b_i \neq 0, -1, -2, \dots, i = 1, \dots, q$. Hay p parámetros en el numerador y q en el denominador.

Los casos ${}_0F_0$ y ${}_1F_0$ son elementales: la serie exponencial y binomial, respectivamente. El caso ${}_2F_1$ es la serie hipergeométrica de Gauss.

Si uno de los parámetros en el numerador $a_i, i = 1, \dots, p$ es un entero negativo, $a_1 = -n$, digamos, la serie es finita y

$${}_pF_q(-n, \dots, a_p; b_1, \dots, b_q; z) = \sum_{j=0}^n \frac{(-n)_j \dots (a_p)_j}{(b_1)_j \dots (b_q)_j} \frac{z^j}{j!}.$$

Cuando la serie es infinita, converge para $|z| < \infty$ si $p \leq q$, converge para $|z| < 1$ si $p = q + 1$, y diverge para toda $z \neq 0$ si $p > q + 1$.

Finalmente introducimos la notación para la N -ésima suma parcial de la serie hipergeométrica generalizada, la cual es de uso si una de las b_j es igual a $-N$. Usamos esta notación en la definición de los polinomios ortogonales discretos. Definimos

$${}_p\hat{F}_q(a_1, \dots, a_p; b_1, \dots, b_q; z) = \sum_{j=0}^N \frac{(a_1)_j \dots (a_p)_j}{(b_1)_j \dots (b_q)_j} \frac{z^j}{j!},$$

donde N denota el entero no negativo que aparece en algunas definiciones de familias de polinomios ortogonales discretos.

La transformada de Stieltjes

La transformada de Stieltjes de una medida ψ definida en un intervalo real I en \mathbb{R} se define como

$$(2.4) \quad B(z; \psi) = \int_I \frac{d\psi(x)}{x - z}, \quad z \in \mathbb{C} - I.$$

Esta transformada puede ser considerada como la función generadora de momentos de la medida ψ , ya que

$$(2.5) \quad B(z; \psi) = -\frac{1}{z} \int_I \frac{1}{1 - x/z} d\psi(x) = -\frac{1}{z} \sum_{n=0}^{\infty} \int_I \frac{x^n}{z^{n+1}} d\psi(x) = -\sum_{n=0}^{\infty} \frac{\mu_n}{z^{n+1}},$$

donde $\mu_n = \int_I x^n d\psi(x)$ son los momentos de la medida. En el caso en el que $\text{sop}(\psi) \subseteq [-A, A]$ entonces $|\mu_n| \leq 2A^n$, implicando que la serie anterior es absolutamente convergente para $|z| > A$. En este caso, la transformada de Stieltjes está completamente determinada en términos de la medida.

Existe una fórmula que permite calcular la medida ψ si se conoce la correspondiente transformada de Stieltjes. A este resultado se le conoce como fórmula de inversión de Perron-Stieltjes, que tiene varias versiones, y la que usamos dice:

Proposición 2.0.5. *Sea ψ una medida de probabilidad con momentos finitos y sea $B(z; \psi)$ su transformada de Stieltjes. Entonces*

$$\int_a^b d\psi(x) + \frac{1}{2}\psi(\{a\}) + \frac{1}{2}\psi(\{b\}) = \frac{1}{\pi} \int_a^b \text{Im}B(x + iz; \psi) dx,$$

donde $\psi(\{a\})$ es el peso de la masa en un punto aislado a . Si la medida es continua en a entonces $\psi(\{a\}) = 0$.

Como consecuencia de la proposición anterior también se tiene la fórmula

$$\int_a^b d\psi(x) = \frac{1}{\pi} \lim_{\epsilon \rightarrow 0^+} \lim_{\eta \rightarrow 0^+} \int_{a+\eta}^{b-\eta} \text{Im}B(x + i\epsilon; \psi) dx.$$

Cuando la medida es continuamente diferenciable con respecto a la medida de Lebesgue en todo su intervalo de definición, i.e. $d\psi(x) = \psi'(x)dx$, se tiene la relación directa

$$\psi'(x) = \lim_{\epsilon \rightarrow 0^+} \frac{B(x + i\epsilon; \psi) - B(x - i\epsilon; \psi)}{2\pi i} dx.$$

Por último, para medidas que tengan una parte continuamente diferenciable y una parte discreta, hay una manera directa de calcular la magnitud del salto. Supongamos que $\psi = \hat{\psi} + \psi(\{a\})\delta_a$. Se tiene entonces que la transformada de Stieltjes cumple

$$B(z; \psi) = B(z; \hat{\psi}) + \frac{\psi(\{a\})}{a - z}.$$

Si $B(z; \hat{\psi})$ está bien definido en $z = a$, entonces

$$(2.6) \quad \psi(\{a\}) = \lim_{z \rightarrow a} (a - z)B(z; \psi).$$

2.1. Definición y propiedades generales

Un sistema de polinomios $\{Q_n(x) : n \in \mathbb{N}\}$ donde $Q_n(x)$ es un polinomio de grado exactamente n es un **sistema de polinomios ortogonales** con respecto a una medida real positiva ϕ , si tenemos las siguientes relaciones de ortogonalidad:

$$(2.7) \quad \int_S Q_n(x)Q_m(x)d\phi(x) = d_n^2\delta_{nm}$$

donde S es el soporte de la medida ϕ y d_n son constantes distintas de cero. Si estas constantes satisfacen $d_n = 1$, decimos que el sistema es **ortonormal**.

La medida ϕ usualmente tiene densidad $w(x)$ diferenciable con respecto a la medida de Lebesgue dx o es una medida discreta con pesos $w(i)$ en los puntos x_i . La relación anterior se convierte entonces en

$$(2.8) \quad \int_S Q_n(x)Q_m(x)w(x)dx = d_n^2\delta_{nm}$$

en el caso continuo y

$$(2.9) \quad \sum_{i=0}^M Q_n(x_i)Q_m(x_i)w_i = d_n^2\delta_{nm}$$

en el caso discreto donde M puede llegar a ser infinito.

Relación de recurrencia a tres términos

Es bien sabido que cualquier familia de polinomios ortogonales $\{Q_n(x)\}$ satisface una **relación de recurrencia a tres términos**

$$xQ_n(x) = a_nQ_{n+1}(x) + b_nQ_n(x) + c_nQ_{n-1}(x),$$

que como veremos más adelante, para el caso de procesos de nacimiento y muerte toma la forma:

$$(2.10) \quad -xQ_n(x) = a_nQ_{n+1}(x) + b_nQ_n(x) + c_nQ_{n-1}(x)$$

donde $a_n, c_n \neq 0$ y $c_n/a_{n-1} > 0$ y donde si además $Q_n(0) = 1$ entonces tenemos que $b_n = -(a_n + c_n)$, pues en ese caso

$$\begin{aligned} -0 \cdot Q_n(0) &= a_nQ_{n+1}(0) + b_nQ_n(0) + c_nQ_{n-1}(0) \\ 0 &= a_n + b_n + c_n \text{ para toda } n \text{ y así } b_n = -(a_n + c_n), \end{aligned}$$

y los polinomios $Q_n(x)$ se pueden definir por la relación de recurrencia

$$(2.11) \quad -xQ_n(x) = a_nQ_{n+1}(x) - (a_n + c_n)Q_n(x) + c_nQ_{n-1}(x)$$

con la condición adicional de que $Q_{-1}(x) = 0$ y $Q_0(x) = 1$. Favard demostró un resultado inverso.

Teorema 2.1.1. (de Favard) Sean A_n, B_n y C_n sucesiones arbitrarias de números reales y sea $\{Q_n(x)\}$ la sucesión de polinomios definida por la relación de recurrencia

$$(2.12) \quad Q_{n+1}(x) = (A_n x + B_n)Q_n(x) - C_n Q_{n-1}(x)$$

con $Q_0(x) = 1$ y $Q_{-1}(x) = 0$

Entonces $\{Q_n(x)\}$ es un sistema de polinomios ortogonales si y sólo si $A_n \neq 0, C_n \neq 0$, y $C_n A_n A_{n-1} > 0$ para toda n . Si todas las $A_n = 1$ entonces tenemos un sistema de polinomios ortogonales **mónicos**.

Comenzamos con algunas propiedades comunes a cualquier sucesión de polinomios ortogonales. Sea $w(x)$ un peso positivo en un intervalo abierto $I = (a, b)$ y asumamos que los momentos

$$(2.13) \quad \mu_n = \int_a^b x^n w(x) dx, n = 0, 1, 2, \dots$$

son finitos.

Sea $\Delta_{-1} = 1$ y sea $\Delta_n, n \geq 0$ el determinante

$$(2.14) \quad \Delta_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}.$$

La **forma cuadrática** asociada

$$\sum_{j,k=0}^n \mu_{j+k} a_j a_k = \int_a^b \left[\sum_{j=0}^n a_j x^j \right]^2 w(x) dx$$

es positiva definida, por lo que el determinante Δ_n es positivo.

Supóngase que $w > 0$ en $I = (a, b)$. El espacio asociado L_w^2 consiste de todas las funciones reales medibles f tales que

$$\int_a^b f^2(x) w(x) dx.$$

El producto interno (f, g) entre dos funciones es

$$(2.15) \quad (f, g) = (f, g)_w = \int_a^b f(x)g(x)w(x)dx.$$

El polinomio

$$Q_n(x) = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} & x \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} & x^n \end{vmatrix}$$

es ortogonal a x^m , $m < n$ mientras que $(Q_n, x^n) = \Delta_n$. Para ver esto, expandimos el determinante anterior a lo largo de la última columna. Calculando el producto interno de Q_n con x^m resulta en un determinante en el cual la última columna del determinante

$$\begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}$$

es remplazada por la columna $m+1$ del mismo determinante. Así, si $m < n$ hay una columna repetida, mientras que si $m = n$ obtenemos Δ_n . Ahora $Q_n(x) = \Delta_{n-1}x^n$ más términos de grado menor, por lo que

$$(Q_n, Q_n) = (Q_n, \Delta_{n-1}x^n) = \Delta_{n-1}(Q_n, x^n) = \Delta_{n-1}\Delta_n.$$

Así, los polinomios

$$P_n(x) = \frac{1}{\sqrt{\Delta_{n-1}\Delta_n}}Q_n(x)$$

son ortonormales. Están únicamente determinados por la condición de que el coeficiente líder sea positivo. El coeficiente líder de P_n es

$$h_n = \sqrt{\Delta_{n-1}/\Delta_n}Q_n(x).$$

Nótese además que $xP_n(x)$ tiene grado $n+1$ y es ortogonal a x^m , $m < n-1$. Se sigue que para algunas constantes a_n, b_n, c_n ocurre que

$$(2.16) \quad xP_n(x) = a_nP_{n+1}(x) + b_nP_n(x) + c_nP_{n-1}(x).$$

Comparando los coeficientes de x^{n+1} , observamos que $a_n = h_n/h_{n+1}$. Por otro lado, tomando el producto interior con P_{n-1} , obtenemos

$$(2.17) \quad c_n = (xP_n, P_{n-1}) = (P_n, xP_{n-1}) = \frac{h_{n-1}}{h_n} = a_{n-1}.$$

Nótese que la existencia de la fórmula de recurrencia a tres términos depende sólo de las propiedades de ortogonalidad de P_n , no en el hecho de que tienen norma uno. Obsérvese además que si $P = (P_0, P_1, \dots)^T$ la relación de recurrencia a tres términos se puede escribir como

$$xP = JP, \quad \text{con} \quad J = \begin{pmatrix} b_0 & a_0 & & & \\ a_1 & b_1 & a_1 & & \\ & a_2 & b_2 & a_2 & \\ & & \ddots & \ddots & \ddots \end{pmatrix}$$

Esta es la llamada matriz de Jacobi, notemos la similitud con la matriz de transición de probabilidad o el operador infinitesimal de una caminata aleatoria o cadena de nacimiento y muerte.

Se sigue de las ecuaciones (2.16) y (2.17) que

$$(x-y)P_n(x)P_n(y) = a_n[P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)] - a_{n-1}[P_n(x)P_{n-1}(y) - P_{n-1}(x)P_n(y)].$$

Iterando, sumando, y dividiendo por $x - y$ obtenemos la fórmula de *Christoffel-Darboux*

$$(2.18) \quad a_n \left[\frac{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)}{x - y} \right] = \sum_{j=0}^n P_j(x)P_j(y).$$

Recordemos que el coeficiente del lado izquierdo es el radio del coeficiente líder de P_n al coeficiente líder de P_{n+1} . Lo anterior tiene consecuencias importantes. Supongamos que q es un polinomio de grado $\leq n$. Entonces q es una combinación lineal de $P_k, k \leq n$, y la ortonormalidad implica que

$$q(x) = \sum_{j=0}^n (q, P_j) P_j(x).$$

Así tenemos la siguiente

Proposición 2.1.1. *Si q es un polinomio de grado $\leq n$, entonces*

$$q(x) = \int_a^b K_n(x, y) q(y) w(y) dy$$

donde

$$K_n(x, y) = a_n \left[\frac{P_{n+1}(x)P_n(y) - P_n(x)P_{n+1}(y)}{x - y} \right]$$

y a_n es el radio del coeficiente líder de P_n al coeficiente líder de P_{n+1} .

La función kernel K_n o núcleo juega el mismo rol con respecto a las expansiones de los polinomios ortogonales como el kernel de Dirichlet juega con respecto a la expansión clásica de Fourier, así que nos referimos a ella como el **Kernel de Dirichlet** para los polinomios $\{P_n\}$. La función kernel puede pensarse como el siguiente determinante

$$(2.19) \quad K_n(x, y) = -\frac{1}{\Delta_n} \begin{vmatrix} 0 & 1 & x & x^2 & \cdots & x^n \\ 1 & A_0 & A_1 & A_2 & \cdots & A_n \\ y & A_1 & A_2 & A_3 & \cdots & A_{n+1} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ y^n & A_n & A_{n+1} & A_{n+2} & \cdots & A_{2n} \end{vmatrix}$$

Tomando el límite cuando $y \rightarrow x$ obtenemos

$$a_n [P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)] = \sum_{j=0}^n P_j^2(x).$$

Una primera consecuencia de la ecuación anterior es que las raíces reales de P_n son simples: si x_0 fuera una raíz doble, el lado izquierdo de la ecuación anterior se anularía en $x = x_0$, pero P_0 es una constante no cero, por lo que el lado derecho de la ecuación es positivo.

Proposición 2.1.2. *P_n tiene n raíces reales, todas en el intervalo I .*

Demostración. El caso $n = 0$ es trivial. Supongamos que $n \geq 1$ y sean x_1, \dots, x_m las raíces reales de P_n que pertenecen a I . Sea

$$q(x) = \Pi(x - x_j).$$

Los cambios de signo de P_n y q en el intervalo I ocurren precisamente en las x_j . Por lo tanto el producto qP_n tiene signo constante en I , así $(P_n, q) \neq 0$. Esto implica que q tiene grado por lo menos n , con lo cual $m = n$, ya que en caso contrario $(p_n, q) = 0$ por la ortogonalidad. \square

Proposición 2.1.3. *Entre cada par de ceros de P_n hay un cero de P_{n-1} .*

Demostración. Para $n < 2$ no hay nada que probar. Para $n \geq 2$,

$$a_n[P'_{n+1}(x)P_n(x) - P_{n+1}(x)P'_n(x)] = \sum_{j=0}^n P_j^2(x)$$

implica que

$$P'_n(x)P_{n-1}(x) - P_n(x)P'_{n-1}(x) > 0.$$

Supongamos que $x_1 < x_2$ son dos ceros consecutivos de P_n . Entonces la desigualdad anterior implica que

$$P'_n(x_j)P_{n-1}(x_j) > 0, j = 1, 2.$$

Como $P'_n(x_1)$ y $P'_n(x_2)$ deben tener diferentes signos, se sigue de la desigualdad anterior que $P_{n-1}(x_j)$ tiene diferentes signos. Así, P_{n-1} tiene un cero en el intervalo (x_1, x_2) por el teorema de Bolzano. \square

Aproximaciones con los polinomios ortogonales

Pasamos ahora a la pregunta de completitud: ¿puede ser cualquier elemento del espacio de Hilbert L_w^2 expresado como una combinación lineal de P_n ? Supongamos que f pertenece a L_w^2 . Considérese el problema de encontrar un elemento en el espacio generado por $\{P_j\}_{j \leq n}$ tal que es el más cercano a f , con respecto a la distancia en L_w^2

$$d(f, g) = \|f - g\| = (f - g, f - g)^{1/2}.$$

Proposición 2.1.4. *Sea*

$$f_n = \sum_{j=0}^n (f, P_j) P_j = \int_a^b K_n(x, y) f(y) w(y) dy.$$

Entonces f_n es la función más próxima a f en el espacio generado por $\{P_0, P_1, \dots, P_n\}$.

Demostración. Escribimos

$$f - g = (f - f_n) + (f_n - g).$$

Calculando los productos interiores vemos que $f - f_n$ es ortogonal a cualquier elemento del generado por $\{P_j\}_{j \leq n}$, así si g está también en el espacio generado, entonces $f - f_n$ y $f_n - g$ son ortogonales, así

$$\|f - g\|^2 = \|f - f_n\|^2 + \|f_n - g\|^2.$$

La parte izquierda es exactamente minimal cuando $g = f_n$. Tomando $g = 0$ en

$$\|f - g\|^2 = \|f - f_n\|^2 + \|f_n - g\|^2$$

y usando la ortonormalidad de P_n obtenemos la **desigualdad de Bessel**

$$(2.20) \quad \|f\|^2 = \sum_{j=0}^n (f, P_j)^2 + \|f - f_n\|^2 \geq \sum_{j=0}^n (f, P_j)^2.$$

\square

Teorema 2.1.2. *Supongamos que w es un peso positivo en el intervalo (a, b) y supongamos que para algún $c > 0$ se tiene que*

$$\int_a^b e^{2c|x|} w(x) dx < \infty.$$

Sean $\{P_n\}$ los polinomios ortogonales con respecto a w . Para cualquier $f \in L_w^2$ se tiene que

$$f = \sum_{n=0}^{\infty} (f, P_n) P_n$$

*en el sentido de que las sumas parciales de la serie en el lado derecho converge a f en norma en L_w^2 . Además, se tiene la **identidad de Parseval***

$$(2.21) \quad \|f\|^2 = \sum_{n=0}^{\infty} (f, P_n)^2.$$

La demostración de este teorema se puede consultar en el Apéndice B de [3].

2.2. Polinomios ortogonales de variable continua

La ecuación diferencial hipergeométrica

Muchos problemas de matemáticas aplicadas y teóricas así como problemas de física conducen a ecuaciones de la forma

$$(2.22) \quad s(x)y'' + \tau(x)y' + \lambda y = 0$$

donde $s(x)$ y $\tau(x)$ son polinomios de grado a lo sumo dos y de grado exactamente uno respectivamente, y λ una constante. Nos referimos a esta ecuación como ecuación diferencial del tipo hipergeométrico, y a sus soluciones como funciones de tipo hipergeométrico. En cierto sentido hay exactamente dos familias de soluciones a ecuaciones diferenciales lineales de segundo grado.

La ecuación de *Kummer* o ecuación *confluyente hipergeométrica*

$$(2.23) \quad xu''(x) + (c-x)u'(x) - au(x) = 0$$

con índices (a, c) y la ecuación *hipergeométrica*

$$(2.24) \quad x(1-x)u''(x) + [c - (a+b+1)x]u'(x) - abu(x) = 0$$

con índices (a, b, c) , donde a, b, c son constantes.

Las demás ecuaciones (Bessel, Whittaker, Hermite, Legendre...) se obtienen de estas por transformaciones estándares o por continuación analítica en la variable independiente.

La ecuación diferencial general lineal de segundo orden en el intervalo I es:

$$(2.25) \quad p(x)u''(x) + q(x)u'(x) + r(x)u(x) = f(x).$$

Con su correspondiente ecuación homogénea:

$$(2.26) \quad p(x)u''(x) + q(x)u'(x) + r(x)u(x) = 0$$

y con operador diferencial asociado:

$$(2.27) \quad L = p(x)\frac{d^2}{dx^2} + q(x)\frac{d}{dx} + r(x).$$

Obsérvese que operadores diferenciables como L son los que aparecen de manera natural en la definición de operadores infinitesimales asociados a procesos de difusión vistos en el capítulo anterior.

Transformaciones y simetría

¿Cómo es que obtenemos los polinomios ortogonales clásicos? Estos pueden obtenerse de preguntarse por la clase de funciones que son autofunciones del operador (2.27), para esto necesitamos entender algunos conceptos importantes: transformaciones gauge y transformación de Liouville.

Una transformación *gauge* de

$$(2.28) \quad p(x)u''(x) + q(x)u'(x) + r(x)u(x) = f(x)$$

es una transformación de la forma

$$(2.29) \quad u(x) = \phi(x)v(x).$$

La función u satisface (2.28) si y sólo si v satisface

$$p(x)v''(x) + \left[2p(x)\frac{\phi'(x)}{\phi(x)} + q(x)\right]v'(x) + \left[p(x)\frac{\phi''(x)}{\phi(x)} + q(x)\frac{\phi'(x)}{\phi(x)} + r(x)\right]v(x) = \frac{f(x)}{\phi(x)}$$

Nótese que el lado izquierdo de la ecuación no cambia si ϕ es remplazado por $C\phi$, con C constante. El operador transformado correspondiente es

$$L_\phi = p(x)\frac{d^2}{dx^2} + \left[2p(x)\frac{\phi'(x)}{\phi(x)} + q(x)\right]\frac{d}{dx} + \left[p(x)\frac{\phi''(x)}{\phi(x)} + q(x)\frac{\phi'(x)}{\phi(x)} + r(x)\right].$$

La utilidad de las transformaciones gauge derivan del hecho de que la ecuación diferencial lineal homogénea de primer orden

$$\phi'(x) = h(x)\phi(x)$$

siempre tiene una solución

$$\exp\left\{\int_{x_0}^x h(y)dy\right\},$$

donde x_0 es cualquier punto del intervalo I . Esta solución no tiene ceros en el intervalo. Además se tiene que si ψ es una segunda solución de

$$\phi'(x) = h(x)\phi(x)$$

entonces el cociente ψ/ϕ tiene derivada 0 y por lo tanto es una constante.

En particular, una transformación gauge puede ser usada para eliminar el término correspondiente al primer orden qu' de (2.28) tomando ϕ tal que $\phi'/\phi = -q/2p$. Un segundo uso es simetrizar el operador

$$L = p(x) \frac{d^2}{dx^2} + q(x) \frac{d}{dx} + r(x).$$

El operador

$$L = p(x) \frac{d^2}{dx^2} + q(x) \frac{d}{dx} + r(x)$$

se llama **simétrico** con respecto al peso w si $(Lu, v) = (u, Lv)$ con el producto interno definido por (2.15) para cada par de funciones u, v que se anulan fuera de un subintervalo cerrado de I .

Proposición 2.2.1. *El operador L es simétrico con respecto al peso w si y sólo si tiene la forma*

$$(2.30) \quad L = p \frac{d^2}{dx^2} + \frac{(pw)'}{w} \frac{d}{dx} + r = \frac{1}{w} \frac{d}{dx} \left(pw \frac{d}{dx} \right) + r.$$

Proposición 2.2.2. *Si L tiene la forma*

$$L = p(x) \frac{d^2}{dx^2} + q(x) \frac{d}{dx} + r(x)$$

entonces existe un peso w , único salvo constantes multiplicativas, tal que L es simétrica con respecto a w .

Proposición 2.2.3. *Dado un operador*

$$L = p(x) \frac{d^2}{dx^2} + q(x) \frac{d}{dx} + r(x)$$

y un peso w en el intervalo I , existe una transformación gauge

$$u(x) = \phi(x)v(x)$$

tal que el correspondiente operador L_ϕ es simétrico respecto a w .

Una transformación invertible T de un espacio $L^2_{w_1}$ a un espacio $L^2_{w_2}$ se llama **unitaria** si

$$(Tf, Tg)_{w_2} = (f, g)_{w_1}$$

para cada par f, g en $L^2_{w_1}$. Los operadores L_1 y L_2 en sus respectivos espacios se dicen ser **unitariamente equivalentes** por T si

$$L_2 = TL_1T^{-1}.$$

Proposición 2.2.4. *Un operador simétrico respecto a un peso w en un intervalo I es unitariamente equivalente, mediante una transformación gauge, a un operador que es simétrico con respecto al peso 1 en el intervalo I .*

Un segundo método para transformar una ecuación diferencial como (2.28) es hacer un cambio de variable. Si

$$y = y(x), u(x) = v(y(x))$$

entonces

$$u'(x) = y'(x)v'(y(x)), u''(x) = [y'(x)]^2v''(y(x)) + y''(x)v'(y(x)).$$

En particular, podemos eliminar el coeficiente p tomando

$$y(x) = \int_{x_0}^x \frac{dt}{\sqrt{p(t)}}.$$

La ecuación (2.28) se vuelve

$$v'' + \left[\frac{q_1(y)}{\sqrt{p_1(y)}} - \frac{p_1'(y)}{2p_1(y)} \right] v' + rv = f.$$

Esto involucra un abuso de notación: las derivadas en v se refieren a derivadas con respecto a y , mientras que las derivadas en p se refieren a derivadas con respecto a x . Para esclarecer esto consideramos p, q, r y f como funciones de $y = y(x)$ tomando $p(x) = p_1(y(x))$. La ecuación se convierte entonces en

$$v''(y) + \left[\frac{q_1(y)}{\sqrt{p_1(y)}} - \frac{p_1'(y)}{2p_1(y)} \right] v'(y) + r_1(y)v(y) = f_1(y).$$

Si eliminamos el término correspondiente al primer orden con una transformación gauge, la transformación compuesta correspondiente se conoce como **transformación de Liouville**.

Polinomios ortogonales clásicos

Es de especial interés extender la condición de simetría a una clase más larga de funciones. En general esto requiere de condiciones de frontera.

Supongamos que $I = (a, b)$ es un intervalo acotado y supongamos que w, w', p, p', q y r se extienden como funciones continuas en el intervalo cerrado $[a, b]$. Supongamos además que u y v son funciones en $C^2(a, b)$ que pertenecen a L_w^2 , y supongamos que u, u', v, v' son continuas en el intervalo cerrado. Supongamos también que L es simétrica. Sabemos que

$$(Lu, v) - (u, Lv) = (pwu'v - pwwv') \Big|_a^b.$$

Si pw se anula en ambos puntos de la frontera entonces no necesitamos condiciones adicionales en la frontera; de otra forma se deben añadir condiciones sobre las funciones u, v . Similarmente, si I es un intervalo semi infinito (a, ∞) , debemos imponer condiciones en $x = a$ a menos que $pw = 0$ en $x = a$. Supongamos que tenemos simetría para una clase maximal de funciones. Una función u que no sea idénticamente cero es una eigenfunción para L con eigenvalor $-\lambda$ si

$$Lu + \lambda u = 0.$$

Si u_1 y u_2 son eigenfunciones con diferentes eigenvalores $-\lambda_1$ y $-\lambda_2$ entonces

$$-\lambda_1(u_1, u_2) = (Lu_1, u_2) = (u_1, Lu_2) = -\lambda_2(u_1, u_2).$$

Así $(u_1, u_2) = 0$: u_1, u_2 son ortogonales. Nos preguntamos bajo qué condiciones se da el caso de que el conjunto de eigenfunciones de L incluye a los polinomios de todos los grados, o por lo menos los grados 0, 1, 2. La condición de simetría implica que L tiene la forma

$$p \frac{d^2}{dx^2} + \frac{(pw)'}{w} \frac{d}{dx} + r.$$

Supongamos que hay polinomios de grados 0, 1, 2 que son eigenfunciones de L . Esto es equivalente a asumir que el espacio de polinomios de grado $\leq k$, $k = 0, 1, 2$ está en el dominio de L y se mapea en sí mismo por L . En particular, las funciones constantes pertenecen a L_w^2 . Así w tiene integral finita

$$\int_a^b w(x) dx < \infty.$$

Aplicando L a la función constante $u_0(x) = 1$ obtenemos $Lu_0 = r$, así que r debe de ser una constante y podemos tomar $r = 0$. De esta forma, aplicando L a $u_1(x) = x$ obtenemos

$$(2.31) \quad Lu_1 = \frac{(pw)'}{w} = p' + p \frac{w'}{w}.$$

Esta expresión debe ser un polinomio de grado a lo más 1. Tomando $u_2(x) = \frac{1}{2}x^2$

$$Lu_2 = p + x \left(p' + p \frac{w'}{w} \right).$$

Como Lu_2 debe ser un polinomio de grado a lo más 2, se sigue que p es un polinomio de grado a lo más 2. La condición de simetría requiere que

$$0 = (Lu, v) - (u, Lv) = \int_a^b [pw(u'v - uv')]'$$

para toda u, v en el dominio. Como se ha visto, una condición necesaria es que $pw \rightarrow 0$ en cada punto final del intervalo. Resumiendo, buscamos soluciones polinomiales $p_n(x)$ que sean autofunciones de un operador diferencial de la forma

$$(2.32) \quad L = p \frac{d^2}{dx^2} + q \frac{d}{dx} = \frac{1}{w} \left(pw \frac{d}{dx} \right)', \quad \deg p \leq 2, \quad \deg q \leq 1,$$

donde $Lp_n = \lambda_n p_n$ y

$$(2.33) \quad (pw)' = qw.$$

A este tipo de operadores diferenciales se les denomina de tipo hipergeométrico, o de tipo Sturm-Liouville. La ecuación (2.33) se denomina la ecuación de Pearson.

Se quiere caracterizar ahora todas las familias de polinomios ortogonales que sean autofunciones de un operador diferencial de tipo Sturm-Liouville (2.32). Como p es un polinomio de grado a lo sumo 2, después de normalizaciones (mapeos afines de la recta, multiplicación del peso, el operador y los polinomios por constantes), se puede reducir a cinco casos, dependiendo del grado y las raíces de p .

1. Caso I: p constante. En este caso, de (2.31), se tiene que w'/w debe ser un polinomio de grado a lo sumo 1. Resolviendo la ecuación diferencial, se tiene que $w(x) = e^h$, donde h es un polinomio real de grado a lo sumo 2. Después de normalizaciones, se puede reducir a dos casos, $w(x) = e^{-x}$ ó $w(x) = e^{\pm x^2}$. En el primer caso, las condiciones de contorno dicen que $pw = w$ se tiene que anular en la frontera de algún intervalo real, pero e^{-x} nunca cumple ese criterio ya que siempre es positiva. Igual para el caso de e^{x^2} . Con lo cual, el único caso que queda es para $w(x) = e^{-x^2}$, en cuyo caso el intervalo de ortogonalidad está forzado a ser $I = (-\infty, \infty)$. El polinomio $q = w'/w = -2x$ es un polinomio de grado exactamente 1. Por lo tanto el operador (2.32) es en este caso

$$L = \frac{d^2}{dx^2} - 2x \frac{d}{dx} \quad \text{en } L_w^2(\mathbb{R}), \quad w(x) = e^{-x^2}.$$

Para cada polinomio de grado n , el operador L lo lleva a otro polinomio de grado también n . El autovalor en este caso, igualando coeficientes es $\lambda_n = -2n$. Con lo cual, los polinomios ortogonales con respecto a la medida e^{-x^2} , que se llaman los polinomios de Hermite y se denotan por $H_n(x)$, cumplen que $LH_n = -2nH_n$. Esta medida, una vez normalizada, corresponde con la distribución normal o gaussiana.

2. Caso II: p lineal. Se normaliza a $p(x) = x$. De (2.31), se tiene que $w'/w = b + \alpha/x$. Resolviendo la ecuación diferencial, se tiene que, salvo constantes, $w(x) = x^\alpha e^{bx}$. Las condiciones de frontera $pw = xw = x^{\alpha+1}e^{bx}$ y la integrabilidad de w obligan a que el intervalo de ortogonalidad sea $I = (0, \infty)$ y que $b < 0$ y $\alpha > -1$. Con lo cual, tras normalizar, $w(x) = x^\alpha e^{-x}$, $\alpha > -1$. El polinomio $q = (xw)'/w = \alpha + 1 - x$ es un polinomio de grado exactamente 1. Por lo tanto el operador (2.32) es en este caso

$$L = x \frac{d^2}{dx^2} + (\alpha + 1 - x) \frac{d}{dx}, \quad \text{en } L_w^2(\mathbb{R}^+), \quad w(x) = x^\alpha e^{-x}, \quad \alpha > -1.$$

Para cada polinomio de grado n , el operador L lo lleva a otro polinomio de grado también n . El autovalor en este caso, igualando coeficientes, es $\lambda_n = -n$. Con lo cual, los polinomios ortogonales con respecto a la medida $x^\alpha e^{-x}$, que se llaman los polinomios de Laguerre y se denotan por $L_n^\alpha(x)$, cumplen que $LL_n^\alpha = -nL_n^\alpha$. Esta medida, una vez normalizada, corresponde con la distribución Gamma y como caso particular, para $\alpha = 0$, la distribución exponencial.

3. Caso III: p cuadrático con raíces reales distintas. Se normaliza a $p(x) = 1 - x^2$. De (2.31) se tiene que $w'/w = \beta/(1+x) + \alpha/(1-x)$. Resolviendo la ecuación diferencial, se tiene que, salvo constantes, $w(x) = (1-x)^\alpha(1+x)^\beta$. Las condiciones de frontera $pw = (1-x^2)w = (1-x)^{\alpha+1}(1+x)^\beta$ y la integrabilidad de w , obligan a que el intervalo de ortogonalidad sea $I = (-1, 1)$ y que $\alpha, \beta > -1$. El polinomio $q = ((1-x^2)w)'/w = \beta - \alpha - x(\alpha + \beta + 2)$ es un polinomio de grado exactamente 1. Por lo tanto el operador (2.32) es en este caso

$$L = (1-x^2) \frac{d^2}{dx^2} + [\beta - \alpha - x(\alpha + \beta + 2)] \frac{d}{dx},$$

$$\text{en } L_w^2(-1, 1) \quad w(x) = (1-x)^\alpha(1+x)^\beta, \quad \alpha, \beta > -1.$$

Para cada polinomio de grado n , el operador L lo lleva a otro polinomio de grado también n . El autovalor en este caso, igualando coeficientes, es cuadrático: $\lambda_n = -n(n + \alpha + \beta + 1)$. Con lo cual, los polinomios ortogonales con respecto a la medida $(1-x)^\alpha(1+x)^\beta$, que se llaman polinomios de Jacobi y se denotan por $P_n^{(\alpha, \beta)}(x)$ cumplen

que $LP_n^{(\alpha+\beta)} = -n(n+\alpha+\beta+1)P_n^{(\alpha,\beta)}$. Los polinomios de Jacobi se pueden construir en cualquier intervalo acotado $[a, b]$, en cuyo caso $p(x) = (b-x)(x-a)$. Una elección bastante usada es para $a = 0$ y $b = 1$, i.e. en $[0, 1]$, en cuyo caso la medida, una vez normalizada, corresponde con la distribución Beta.

4. Caso IV: p cuadrático con raíces complejas distintas. Se normaliza a $p(x) = 1 + x^2$. De (2.31) se tiene que $w'/w = (-2ax + b)/(1 + x^2)$. Resolviendo la ecuación diferencial, se tiene que, salvo constantes,

$$w(x) = (1 + x^2)^{-a} e^{b \arctan x}.$$

Las condiciones de frontera y la positividad e integrabilidad de w , obligan a que el intervalo de ortogonalidad sea $I = (-\infty, \infty)$ y que $a > 1/2, b \in \mathbb{R}$. El polinomio $q = ((1 + x^2)w)'/w = 2x(1 - a) + b$ es un polinomio de grado exactamente 1. Por lo tanto el operador (2.32) es en este caso

$$L(1 + x^2) \frac{d^2}{dx^2} + [2x(1 - a) + b] \frac{d}{dx},$$

en $L_w^2(\mathbb{R})$ $w(x) = (1 + x^2)^{-a} e^{b \arctan x}$, $a > 1/2, b \in \mathbb{R}$.

Para cada polinomio de grado n , el operador L lo lleva a otro polinomio de grado también n . El autovalor en este caso, igualando coeficientes, es de nuevo cuadrático: $\lambda_n = n(n+1-2a)$. Sin embargo el valor de a condiciona la existencia de una sucesión de polinomios ortogonales, ya que para construirlos es necesario que todos los momentos de la medida sean finitos, en este caso sólo hay un número finito de momentos finitos, con lo cual habrá un número finito de polinomios ortogonales $p_n(x; a, b)$, llamados polinomios de Romanovski. El número de polinomios que se pueden construir es para $n < 2a - 1$, que son casualmente los valores para los que el autovalor λ_n es negativo. Esto indica que el operador L tendrá un espectro discreto finito en combinación con un espectro continuo.

Obsérvese que en este caso, para $b = 0$ se tiene que w se reduce a la distribución t de Student.

5. Caso V: p cuadrático con raíz doble. Se normaliza a x^2 . De (2.31), se tiene que $w'/w = -a/x + b/x^2$. Resolviendo la ecuación diferencial, se tiene que, salvo constantes,

$$w(x) = x^{-a} e^{-b/x}.$$

Las condiciones de frontera y la positividad e integrabilidad de w , obligan a que el intervalo de ortogonalidad sea $I = (0, \infty)$ y que $a > -1, b \geq 0$. El polinomio $q = (x^2 w)'/w = x(2 - a) + b$ es un polinomio de grado exactamente 1. Por lo tanto el operador (2.32) es en este caso

$$L = x^2 \frac{d^2}{dx^2} + [x(2 - a) + b] \frac{d}{dx}, \quad \text{en } L_w^2(\mathbb{R}^+), \quad w(x) = x^{-a} e^{-b/x}, \quad a > -1, b \geq 0.$$

El autovalor en este caso, igualando coeficientes, es de nuevo cuadrático: $\lambda_n = n(n + 1 - a)$. Sin embargo, de vuela, sólo hay un número finito de momentos finitos (para $n < a - 1$), con lo cual habrá un número finito de polinomios ortogonales $y_n(x; a, b)$.

También en este caso el operador L tendrá un espectro discreto finito en combinación con un espectro continuo.

Hay dos casos especiales de soluciones polinomiales, pero que no son ortogonales con respecto a ninguna medida soportada en la recta real. Por un lado, cuando $b = 0$, se tiene que $y_n(x; a, 0) = x^n$ es solución de la ecuación diferencial $Ly_n(x; a, 0) = \lambda_n y_n(x; a, 0)$, pero los monomios no son ortogonales con respecto a ninguna medida en \mathbb{R} . Por otro lado, para $b = 1$ y $a = 2 - \alpha$, se tiene que $y_n(x; 2 - \alpha, 1) = B_n(x; \alpha, 1)$ donde $B_n(x; \alpha, \alpha)$ son los denominados polinomios de Bessel. Sin embargo estos polinomios son ortogonales en la circunferencia unitaria, y no en ningún intervalo real.

Con lo anterior, tenemos el siguiente

Teorema 2.2.1. *Salvo normalizaciones, las únicas familias de polinomios ortogonales que son autofunciones de un operador diferencial de segundo orden de la forma (2.30) (i.e. simétrico con respecto a una medida positiva soportada dentro de la recta real) son las familias de Hermite, Laguerre y Jacobi. A estas familias se les suele denominar familias clásicas.*

La fórmula de Rodrigues

Regresamos a los tres casos correspondientes a los polinomios clásicos, con intervalo $I \subset \mathbb{R}$, peso w , y ecuación de eigenvalor de la forma

$$(2.34) \quad p(x)\psi_n''(x) + q(x)\psi_n'(x) + \lambda_n\psi_n(x) = 0, \quad q = \frac{(pw)'}{w}.$$

O equivalentemente

$$(pw\psi_n')' + \lambda_n w\psi_n = 0.$$

La derivada de una solución ψ_n de (2.34) satisface un ecuación similar: diferenciando (2.34) obtenemos

$$p[\psi_n']'' + (q + p')[\psi_n']' + (q' + \lambda_n)\psi_n' = 0.$$

Ahora

$$q + p' = \frac{(pw)'}{w} + p' = \frac{p^2 w' + 2pp'w}{pw} = \frac{[p(pw)]'}{pw}.$$

Así la función pw es también un peso. Como q' es una constante en cada uno de los tres casos, ψ_n es un polinomio ortogonal de grado $n - 1$ para el peso pw .

Continuando ψ_n'' es un polinomio ortogonal de grado $n - 2$ para el peso $p^2 w$, con eigenvalor

$$-\lambda_n - q' - (p' + q)' = -\lambda_n - 2q' - p''.$$

Por inducción, la derivada m -ésima corresponde al peso $p^m w$, con eigenvalor

$$-\lambda_n - mq' - \frac{1}{2}m(m-1)p''.$$

Como $\psi_n^{(n)}$ es constante, el correspondiente eigenvalor es cero y tenemos la fórmula general

$$\lambda_n = -nq' - \frac{1}{2}n(n-1)p''$$

que corresponde a los resultados en los tres casos mencionados

$$\lambda_n = 2n, \lambda_n = n, \lambda_n = n(n + \alpha + \beta + 1)$$

respectivamente.

La ecuación (2.34) puede ser escrita como

$$w\psi_n = -\lambda_n^{-1}(pw\psi_n')'.$$

Como pw es un peso correspondiente a ψ_n' , esto nos lleva a la ecuación

$$w\psi_n = [\lambda_n(\lambda_n + q')]^{-1}(p^2w\psi_n'')''$$

y finalmente a la fórmula de Rodrigues

$$(2.35) \quad w\psi_n = (-1)^n \prod_{m=0}^{n-1} [\lambda_n + mq' + \frac{1}{2}m(m-1)p'']^{-1} (p^n w \psi_n^{(n)})^n.$$

Como $\psi_n^{(n)}$ es constante, podemos normalizar tomando

$$(2.36) \quad \psi_n(x) = w(x)^{-1} \frac{d^n}{dx^n} \{p(x)^n w(x)\} = \frac{1}{w(x)} \left(p^n(x) w(x) \right)^{(n)}.$$

En vista de (2.35), con esta elección de ψ_n , tenemos que

$$\psi_n(x) = a_n x^n + o(x^{n-1})$$

$$(2.37) \quad n!a_n = (-1)^n \prod_{m=0}^{n-1} [\lambda_n + mq' + \frac{1}{2}m(m-1)p''].$$

Como una primera aplicación de la fórmula de Rodrigues, consideramos el cálculo de los productos internos de la forma

$$(f, \psi_n) = (f, \psi_n)_w = \int_a^b f(x) \psi_n(x) w(x) dx.$$

Por (2.36), $\psi_n w = (p^n w)^{(n)}$. La función $p^n w$ se anula bastante rápido en los puntos finales del intervalo I . Por lo tanto, bajo condiciones en la función f , podemos integrar por partes n veces sin obtener términos en la frontera, para obtener

$$\int_a^b \psi_n(x) f(x) w(x) dx = (-1)^n \int_a^b p^n(x) f^{(n)}(x) w(x) dx.$$

Esta idea puede ser utilizada junto con (2.37) para calcular las normas

$$(2.38) \quad \begin{aligned} \int_a^b \psi_n^2(x) w(x) dx &= (-1)^n \int_a^b \psi_n^n p^n(x) w(x) dx \\ &= \prod_{m=0}^{n-1} [\lambda_n + mq' + \frac{1}{2}m(m-1)p''] \int_a^b p^n(x) w(x) dx. \end{aligned}$$

Veamos ahora con más detalle algunas fórmulas asociadas a estas tres familias clásicas.

Los polinomios de Hermite

Los polinomios de Hermite $\{H_n\}$ son polinomios ortogonales asociados con el peso e^{-x^2} en la recta $\mathbb{R} = (-\infty, \infty)$. Es decir

$$I = \mathbb{R} = (-\infty, \infty), w(x) = e^{-x^2}, p(x) = 1, q(x) = -2x;$$

Aquí, el lado derecho de (2.37) es

$$(-2)^n n!,$$

así el coeficiente líder a_n es $(-2)^n$. La integral en (2.38) da

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi},$$

donde Γ y B son las funciones gamma y beta respectivamente. Así el cuadrado de la norma L^2 de ψ_n es respectivamente

$$(2.39) \quad \|\psi_n\|^2 = 2^n n! \sqrt{\pi}.$$

Son eigenfunciones

$$(2.40) \quad H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0,$$

satisfacen la relación

$$H_n'(x) = 2nH_{n-1}(x),$$

y pueden ser definidos por la fórmula de Rodrigues

$$(2.41) \quad H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} \left\{ e^{-x^2} \right\} = \left(2x - \frac{d}{dx} \right)^n \{1\}.$$

Bajo esta normalización, el cálculo previo de las norma L^2 da

$$\|H_n\|^2 = 2^n n! \sqrt{\pi}.$$

Se sigue que el coeficiente líder es 2^n y que

$$(2.42) \quad H_n'(x) - 2xH_n(x) = -H_{n+1}(x).$$

La relación de recurrencia a tres términos se deduce de la siguiente manera: se puede mostrar por inducción que H_n es par si n es par e impar si n es impar:

$$H_n(-x) = (-1)^n H_n(x).$$

Por lo tanto la relación debe de tener la forma

$$(2.43) \quad xH_n(x) = a_n H_{n+1}(x) + b_n H_{n-1}(x).$$

Las identidades (2.42) y (2.43) implican que $H_n' = 2b_n H_{n-1}$. Comparando a los coeficientes líderes, vemos que $a_n = \frac{1}{2}$ y $b_n = n$:

$$xH_n(x) = \frac{1}{2}H_{n+1}(x) + nH_{n-1}(x).$$

Si escribimos $H_n(x) = \sum_{k=0}^n a_k x^k$, entonces la ecuación (2.40) implica la relación

$$(k+2)(k+1)a_{k+2} = 2(k-n)a_k.$$

Como $a_n = 2^n$ y $a_{n-1} = 0$, esta recursión da como resultado

$$H_n(x) = \sum_{2j \leq n} (-1)^j \frac{n!}{j!(n-2j)!} (2x)^{n-2j}.$$

Los primeros seis polinomios son

$$H_0(x) = 1;$$

$$H_1(x) = 2x;$$

$$H_2(x) = 4x^2 - 2;$$

$$H_3(x) = 8x^3 - 12x;$$

$$H_4(x) = 16x^4 - 48x^2 + 12;$$

$$H_5(x) = 32x^5 - 160x^3 + 120x.$$

Su función generadora es

$$\sum_{n=0}^{\infty} \frac{H_n(x)}{n!} s^n = \frac{e^{-(x-s)^2}}{e^{-x^2}} = e^{2xs-s^2}.$$

Los polinomios de Laguerre

Los polinomios de Laguerre $\{L_n^{(\alpha)}\}$ son polinomios ortogonales asociados con el peso $w(x) = x^\alpha e^{-x}$ en la mitad de la recta real $\mathbb{R}_+ = (0, \infty)$. Es decir

$$I = \mathbb{R}_+ = (0, \infty), w(x) = x^\alpha e^{-x}, \alpha > -1, p(x) = x, q(x) = \alpha + 1 - x;$$

Aquí, el lado derecho de (2.37) es

$$(-1)^n n!,$$

así el coeficiente líder a_n es $(-1)^n$. La integral en (2.38) da

$$\int_0^{\infty} x^{n+\alpha} e^{-x} dx = \Gamma(n + \alpha + 1),$$

donde Γ y B son las funciones gamma y beta respectivamente. Así el cuadrado de la norma L^2 de ψ_n es respectivamente

$$\|\psi_n\|^2 = n! \Gamma(n + \alpha + 1).$$

Para $\alpha > -1$ dado, son eigenfunciones

$$(2.44) \quad x[L_n^{(\alpha)}]''(x) + (\alpha + 1 - x)[L_n^{(\alpha)}]'(x) + nL_n^{(\alpha)}(x) = 0.$$

Satisfacen la relación

$$[L_n^{(\alpha)}]'(x) = -L_{n-1}^{(\alpha+1)}(x)$$

y están dados por la fórmula de Rodrigues

$$(2.45) \quad \begin{aligned} L_n^{(\alpha)}(x) &= \frac{1}{n!} x^{-\alpha} e^x \frac{d^n}{dx^n} \{x^\alpha e^{-x} x^n\} \\ &= \frac{1}{n!} \left[\frac{d}{dx} + \frac{\alpha}{x} - 1 \right]^n \{x^n\}, \end{aligned}$$

donde la segunda versión es obtenida usando la transformación gauge $u\phi v$ con $\phi = x^\alpha e^{-x}$. Bajo esta normalización, el cálculo previo de las norma L^2 da

$$\|L_n^\alpha\|^2 = \frac{\Gamma(n + \alpha + 1)}{n!}.$$

Se sigue que el coeficiente líder es $(-1)^n/n!$.

Escribiendo $L_n^{(\alpha)}(x) = \sum_{k=0}^n b_k x^k$, la ecuación (2.44) da la relación de recurrencia

$$(k+1)(k+\alpha+1)b_{k+1} = -(n-k)b_k,$$

de manera que

$$(2.46) \quad L_n^\alpha(x) = \sum_{k=0}^n (-1)^k \frac{(\alpha+1)_n}{k!(n-k)!(\alpha+1)_k} x^k$$

$$(2.47) \quad = \frac{(\alpha+1)_n}{n!} \sum_{k=0}^n \frac{(-n)_k}{(\alpha+1)_k k!} x^k.$$

Los primeros cuatro polinomios de $L_n^{(\alpha)}$ son

$$\begin{aligned} L_0^\alpha(x) &= 1; \\ L_1^\alpha(x) &= \alpha + 1 - x; \\ L_2^\alpha(x) &= \frac{(\alpha+1)(\alpha+2)}{2} - (\alpha+2)x + \frac{1}{2}x^2; \\ L_3^\alpha(x) &= \frac{(\alpha+1)(\alpha+2)(\alpha+3)}{6} - \frac{(\alpha+2)(\alpha+3)}{2}x + \frac{(\alpha+3)}{2}x^2 - \frac{1}{6}x^3; \end{aligned}$$

En particular, para $\alpha = 0$,

$$L_n(x) = \sum_{k=0}^n (-1)^k \frac{n!}{k!(n-k)!k!} x^k.$$

Comparando coeficientes, la relación de recurrencia a tres términos es

$$xL_n^{(\alpha)}(x) = -(n+1)L_{n+1}^{(\alpha)}(x) + (2n+\alpha+1)L_n^{(\alpha)}(x) - (n+\alpha)L_{n-1}^{(\alpha)}(x).$$

Por inducción

$$\frac{d^n}{dx^n}\{xf(x)\} = x\frac{d^n}{dx^n}\{f(x)\} + n\frac{d^{n-1}}{dx^{n-1}}\{f(x)\},$$

de manera que la fórmula de Rodrigues da la relación de recurrencia:

$$(n+1)L_{n+1}^{(\alpha)}(x) = \left[x\frac{d}{dx} + \alpha + n + 1 - x \right] \{L_n^{(\alpha)}(x)\}.$$

Su función generadora es

$$G(x, s) = \sum_{n=0}^{\infty} L_n^{(\alpha)}(x)s^n = \frac{e^{-xs/(1-s)}}{(1-s)^{\alpha+1}}.$$

Polinomios de Jacobi

Los polinomios de Jacobi $\{P_n^{(\alpha, \beta)}\}$ con índices $\alpha, \beta > -1$ son ortogonales con respecto al peso $w(x) = (1-x)^\alpha(1+x)^\beta$ en el intervalo $(-1, 1)$. Es decir

$$I = (-1, 1), w(x) = (1-x)^\alpha(1+x)^\beta, \alpha, \beta > -1, p(x) = 1-x^2, q(x) = \beta - \alpha - (\alpha + \beta + 2)x.$$

Aquí, el lado derecho de (2.37) es

$$(-1)^n n! (\alpha + \beta + n + 1)_n,$$

así el coeficiente líder a_n es $(-1)^n (\alpha + \beta + n + 1)_n$. La integral en (2.38) da

$$\begin{aligned} \int_{-1}^1 (1-x)^{n+\alpha}(1+x)^{n+\beta} dx &= 2^{2n+\alpha+\beta+1} \int_0^1 s^{n+\alpha}(1-s)^{n+\beta} ds \\ &= 2^{2n+\alpha+\beta+1} B(n+\alpha+1, n+\beta+1), \end{aligned}$$

donde Γ y B son las funciones gamma y beta respectivamente. Así el cuadrado de la norma L^2 de ψ_n es respectivamente

$$\|\psi_n\|^2 = 2^{2n+\alpha+\beta+1} n! \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)}.$$

Sin embargo, bajo la normalización definida por la fórmula de Rodrigues, la norma obtenida da

$$\|P_n^{(\alpha, \beta)}\|^2 = \frac{2^{\alpha+\beta+1} \Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{n!(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)}.$$

Así, el coeficiente líder es $\frac{(\alpha + \beta + n + 1)_n}{2^n n!}$.

Cambiando el signo de x ,

$$(2.48) \quad P_n^{(\alpha, \beta)}(-x) = (-1)^n P_n^{(\beta, \alpha)}(x).$$

Las $P_n^{(\alpha, \beta)}$ son eigenfunciones:

$$(2.49) \quad (1-x^2)[P_n^{(\alpha, \beta)}]'' + [\beta - \alpha - (\alpha + \beta + 2)x][P_n^{(\alpha, \beta)}]' + n(n + \alpha + \beta + 1)P_n^{(\alpha, \beta)} = 0.$$

Satisfacen la relación

$$[P_n^{(\alpha,\beta)}]'(x) = \frac{1}{2}(n + \alpha + \beta + 1)P_{n-1}^{(\alpha+1,\beta+1)}(x)$$

y pueden ser definidos por la fórmula de Rodrigues

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{n!2^n}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\{(1-x)^{\alpha+n}(1+x)^{\beta+n}\}.$$

Se sigue de la regla de Leibniz en su forma extendida que

$$(2.50) \quad P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n} \sum_{k=0}^n (-1)^k \frac{1}{k!(n-k)!} \times \frac{(\alpha+1)_n(\beta+1)_n}{(\alpha+1)_{n-k}(\beta+1)_k} (1-x)^{n-k}(1+x)^k,$$

que da los puntos terminales

$$(2.51) \quad P_n^{(\alpha,\beta)}(1) = \frac{(\alpha+1)_n}{n!}, \quad P_n^{(\alpha,\beta)}(-1) = (-1)^n \frac{(\beta+1)_n}{n!}.$$

Su función generadora es

$$\sum_{n=0}^{\infty} P_n^{(\alpha,\beta)}(x)s^n = \frac{2^{\alpha+\beta}}{R(1-s+R)^\alpha(1+s+R)^\beta} \quad R = \sqrt{1-2xs+s^2}.$$

Haciendo uso de (2.51), tenemos que $c_0 = (\alpha+1)_n/n!$, por lo que

$$\begin{aligned} P_n^{(\alpha,\beta)}(x) &= \frac{(\alpha+1)_n}{n!} \sum_{k=0}^n \frac{(\alpha+\beta+1+n)_k(-n)_k}{(\alpha+1)_k k!} y^k \\ &= \frac{(\alpha+1)_n}{n!} {}_2F_1\left(\alpha+\beta+1+n, -n, \alpha+1; \frac{1}{2}(1-x)\right). \end{aligned}$$

Polinomios de Lengendre y Chebychev

Casos especiales de los polinomios de Jacobi son los polinomios de Gegenbauer o polinomios ultrasféricos cuando $\alpha = \beta = \lambda - 1/2$. Se suelen denotar por $C_n^{(\lambda)}(x)$. En el caso especial de $\alpha = \beta = 0$ se denominan los polinomios de Legendre o polinomios esféricos, que son ortogonales en $[-1, 1]$ con peso 1. Se suelen denotar por $P_n(x)$. Finalmente tenemos los polinomios de Chebychev, que hay dos tipos, los polinomios de Chebychev de primera especie (cuando $\alpha = \beta = -1/2$), que se suelen denotar por $T_n(x)$ y los polinomios de Chebychev de segunda especie (cuando $\alpha = \beta = 1/2$), que se suelen denotar por $U_n(x)$. Para todas estas familias las fórmulas se simplifican considerablemente.

Para los polinomios de Gegenbauer, se tiene que

$$(1-x^2)[C_n^{(\lambda)}]'' - (1+2\lambda)x[C_n^{(\lambda)}]' + n(n+2\lambda)C_n^{(\lambda)} = 0, \quad xC_n^{(\lambda)} = \frac{n+1}{2(n+\lambda)}C_{n+1}^{(\lambda)} + \frac{n+2\lambda-1}{2(n+\lambda)}C_{n-1}^{(\lambda)},$$

donde la matriz de Jacobi viene dada por

$$J = \begin{pmatrix} 0 & \frac{1}{2\lambda} & & & \\ \frac{2\lambda}{2(1+\lambda)} & 0 & \frac{2}{2(2+\lambda)} & & \\ & \frac{1+2\lambda}{2(2+\lambda)} & 0 & \frac{3}{2(3+\lambda)} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Para los polinomios de Legendre, se tiene que

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0, \quad xP_n(x) = \frac{n+1}{2n+1}P_{n+1}(x) + \frac{n}{2n+1}P_{n-1}(x),$$

donde la matriz de Jacobi viene dada por

$$J = \begin{pmatrix} 0 & 1 & & & \\ \frac{1}{3} & 0 & \frac{2}{3} & & \\ & \frac{2}{5} & 0 & \frac{3}{5} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Para los polinomios de Chebychev de primera especie se tiene que

$$(2.52) \quad (1-x^2)T_n''(x) - xT_n'(x) + n^2T_n(x) = 0, \quad 2xT_n(x) = T_{n+1}(x) + T_{n-1}(x), \quad T_0 = 1, \quad T_1(x) = x,$$

donde la matriz de Jacobi viene dada por

$$J = \begin{pmatrix} 0 & 1 & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Para los polinomios de Chebychev de segunda especie se tiene que

$$(2.53) \quad (1-x^2)U_n''(x) - 3xU_n'(x) + n(n+2)U_n(x) = 0, \quad 2xU_n(x) = U_{n+1}(x) + U_{n-1}(x),$$

donde la matriz de Jacobi viene dada por

$$J = \begin{pmatrix} 0 & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & \\ & & \ddots & \ddots & \ddots \end{pmatrix}.$$

2.3. Polinomios ortogonales de variable discreta

Introducimos ahora definiciones y resultados análogos para los polinomios ortogonales de variable discreta.

Ecuaciones en diferencias de segundo orden

Supongamos que $w = \{w_n\}_{n=-\infty}^{\infty}$ es una sucesión infinita por ambos lados de números no negativos. Su correspondiente producto interno

$$(f, g) = (f, g)_w = \sum_{m=-\infty}^{\infty} f(m)g(m)w_m$$

está bien definido para todas las funciones reales f y g para las cuales las normas $\|f\|_w$ y $\|g\|_w$ son finitas, donde

$$\|f\|_w^2 = (f, f)_w = \sum_{m=-\infty}^{\infty} f^2(m)w_m.$$

La norma y el producto interno dependen sólo de los valores de la función en los enteros, aunque es conveniente continuar viendo a los polinomios, por ejemplo, como definidos en toda la recta. Los polinomios tienen norma finita si y sólo si los momentos pares

$$\sum_{m=-\infty}^{\infty} m^{2n}w_m$$

son finitos. De ser así, entonces polinomios ortogonales ψ_n pueden ser construidos exactamente como en el caso de pesos continuos. Usualmente normalizamos de manera que w es una distribución de probabilidad

$$\sum_{m=-\infty}^{\infty} w_m = 1.$$

De nuevo, hay un relación de recurrencia a tres términos. Supongamos que

$$\psi_n(x) = a_n x^n + b_n x^{n-1} + \dots$$

El polinomio

$$x\psi_n(x) - \frac{a_n}{a_{n+1}}\psi_{n+1} + \left(\frac{b_{n+1}}{a_{n+1}} - \frac{b_n}{a_n}\right)\psi_n(x)$$

tiene grado $n - 1$ y es ortogonal a los polinomios de grado $< n - 1$, y es por lo tanto un múltiplo de $\gamma_n \psi_{n-1}$. Entonces

$$\gamma_n(\psi_{n-1}, \psi_{n-1})_w = (x\psi_n, \psi_{n-1})_w = (\psi_n, x\psi_{n-1})_w = \frac{a_{n-1}}{a_n}(\psi_n, \psi_n)_w.$$

Así

$$x\psi_n(x) = \alpha_n \psi_{n+1}(x) + \beta_n \psi_n(x) + \gamma_n \psi_{n-1}(x)$$

$$(2.54) \quad \alpha_n = \frac{a_n}{a_{n+1}}, \quad \beta_n = \frac{b_n}{a_n} - \frac{b_{n+1}}{a_{n+1}}, \quad \gamma_n = \frac{a_{n-1}}{a_n} \frac{(\psi_n, \psi_n)_w}{(\psi_{n-1}, \psi_{n-1})_w}.$$

Como antes, la relación de recurrencia a tres términos implica una fórmula de Christoffel Darboux. Si el peso w es positivo sólo en un número finito de números enteros, digamos en $m = 0, 1, \dots, N$ entonces el espacio L^2 tiene dimensión $N + 1$. Los polinomios ortogonales tienen grado $0, 1, \dots, N$ y forman una base. En general hay resultados de completitud análogos al caso continuo.

Teorema 2.3.1. *Supongamos que w es un peso positivo en los enteros y supongamos que para alguna constante $c > 0$*

$$\sum_{m=-\infty}^{\infty} e^{2c|m|} w_m < \infty.$$

Sean $\{P_n\}$ los polinomios ortonormales para w . Entonces $\{P_n\}$ es completo, es decir, para cualquier $f \in L_w^2$

$$\lim_{n \rightarrow \infty} \left\| f - \sum_{k=0}^n (f, P_k) P_k \right\|_w = 0.$$

En el caso discreto, es fácil ver que la convergencia en norma implica convergencia puntual en los puntos m donde $w_m > 0$.

Corolario 2.3.1. *Bajo las hipótesis del teorema (2.3.1), para cualquier $f \in L_w^2$ y cualquier m tal que $w_m > 0$*

$$f(m) = \sum_{n=0}^{\infty} (f, P_n) P_n(m).$$

2.4. Operadores en diferencias

Para funciones definidas en los enteros, la diferenciación puede ser reemplazada por operadores “forward” o “backward” donde introducimos ahora la siguiente notación.

$$(2.55) \quad \Delta_+ f(x) = f(x+1) - f(x) \quad \text{y} \quad \Delta_- f(x) = f(x) - f(x-1).$$

Cada operador mapea polinomios de grado d en polinomios de grado $d-1$, así que el producto

$$[\Delta_+ \Delta_-] f(m) = [\Delta_- \Delta_+] f(m) = f(m+1) + f(m-1) - 2f(m)$$

disminuye el grado del polinomio en dos y juega por lo tanto el papel de segunda derivada. Es conveniente expresar estos operadores en términos de operadores “shift”

$$S_{\pm} f(m) = f(m \pm 1).$$

Así

$$\Delta_+ = S_+ - I, \quad \Delta_- = I - S_-$$

y

$$\Delta_+ \Delta_- = \Delta_- \Delta_+ = S_+ + S_- - 2I = \Delta_+ - \Delta_-.$$

Por lo que el operador general de segundo orden en diferencias puede ser escrito como

$$(2.56) \quad L = p_+ S_+ + p_- S_- + r$$

donde p_+, p_- y r son funciones reales. La condición para que L sea simétrica con respecto a w es que

$$\begin{aligned}
0 &= (Lf, g) - (f, Lg) \\
&= (p_+ S_+ f, g) - (f, p_- S_- g) + (p_- S_- f, g) - (g, p_+ S_+ g) \\
&= \sum_{m=-\infty}^{\infty} [p_+ w - S_+(p_- w)](m) f(m+1) g(m) \\
&\quad + \sum_{m=-\infty}^{\infty} [p_- w - S_-(p_+ w)](m) f(m-1) g(m)
\end{aligned}$$

para cada f y g que se anulan para todos los enteros salvo una cantidad finita. Escogiendo g de manera que se anule salvo en un sólo punto m y f que se anule excepto en $m+1$ o en $m-1$, concluimos que la condición de simetría es equivalente a

$$(2.57) \quad S_-(p_+ w) = p_- w, \quad S_+(p_- w) = p_+ w.$$

Una ecuación en diferencias es del tipo **hipergeométrico** si es de la forma

$$(2.58) \quad s(x) \Delta_+ \Delta_- y(x) + \tau(x) \Delta_+ y(x) + \lambda y(x) = 0$$

donde $s(x)$ y $\tau(x)$ son polinomios de a lo sumo segundo y primero grado respectivamente, y λ es una constante. Usando (2.55) podemos escribir (2.58) como

$$(2.59) \quad (s(x) + \tau(x))y(x+1) - (2s(x) + \tau(x))y(x) + s(x)y(x-1) = -\lambda y(x).$$

Si de nuevo,

$$\lambda = \lambda_n = -n\tau' - \frac{1}{2}n(n-1)s''$$

la ecuación en diferencias del tipo hipergeométrico tiene una solución particular de la forma $y(x) = y_n(x)$ que es un polinomio de grado n .

Nos preguntamos ahora para qué pesos discretos w con momentos finitos y para qué operadores simétricos L con coeficientes $p_{\pm}(x)$ positivos cuando $w_x > 0$ (con excepción de los puntos finales) se puede construir un sistema de autofunciones que sean polinomios. Asumimos que $w_x > 0$ donde esté definido y 0 en el resto. Nos interesaremos en tres casos dependiendo de cuántos puntos estemos considerando en el soporte:

1. $w_x > 0$ si y sólo si $0 \leq x \leq N$.
2. $w_x > 0$ si y sólo si $x \geq 0$.
3. $w_x > 0$ para toda $x \in \mathbb{Z}$.

Veremos que el tercer caso no se puede dar y que tanto el primer caso como el segundo dan dos familias de polinomios.

La condición de simetría (2.56) obliga a que $p_-(0) = 0$ en los casos 1. y 2. y que $p_+(N) = 0$ en el caso 1. Con un cambio de notación en el término de orden 0 ($\Delta = \Delta_+$, $\nabla = \Delta_-$), el operador en diferencias se puede reescribir como

$$(2.60) \quad L = p_+ \Delta - p_- \nabla + r$$

Como L debe llevar polinomios de grado n a polinomios de grado n , en particular tiene que ocurrir para los grados 0, 1 y 2. Como $L(1) = r$, entonces r debe de ser constante y se puede tomar como $r = 0$. Para grado 1, como $\Delta(x) = \nabla(x) = 1$, entonces $L(x) = p_+ - p_-$, con lo cual $p_+ - p_-$ debe ser un polinomio de grado 1. Finalmente, como $\Delta(x^2) = 2x + 1$ y $\nabla(x^2) = 2x - 1$, entonces $L(x^2) = 2x(p_+ - p_-) + p_+ + p_-$, y $p_+ + p_-$ tiene que ser un polinomio de grado a lo sumo 2. Por lo tanto p_+ y p_- son polinomios de grado a lo sumo 2 y al menos uno tiene grado positivo. Si los dos tienen grado 2, entonces deben tener el mismo coeficiente líder. A un operador en diferencias 2.60 con estas condiciones sobre sus coeficientes se le denomina, al igual que en el caso continuo, de tipo hipergeométrico.

La ecuación de Pearson discreta a resolver ahora es

$$w_{x+1} = \frac{p_+(x)}{p_-(x+1)} w_x = \frac{\sigma(x) + \tau(x)}{\sigma(x+1)} = \varphi(x) w_x$$

Veamos, como en el caso continuo, los diferentes casos que pueden aparecer.

Caso I: p_{\pm} tiene grado 0, y el otro grado 1. Entonces $\varphi(x)$ es lineal en una dirección o la otra, i.e. $\varphi(x) = \mathcal{O}(|x|)$, lo que implica que el caso 3. no se puede dar, ya que los momentos no serían finitos. Como $p_-(0) = 0$, esto obliga a que el polinomio constante sea p_+ , con lo cual tampoco puede ocurrir el caso 1. Por lo tanto sólo es posible el caso 2. Normalizando con $p_+(x) = 1$, entonces $p_-(x) = x/a$, $a > 0$ por lo tanto $\varphi(x) = a/(x+1)$. Resolviendo la ecuación de Pearson se tiene que

$$w_x = \frac{a^x}{x!} w_0, \quad x = 0, 1, 2, \dots$$

Como $\sum_{x=0}^{\infty} w_x = w_0 e^a$, se puede escoger $w_0 = e^{-a}$ para normalizar la medida. Esta distribución se conoce como la distribución de Poisson y sus correspondientes polinomios ortogonales son los polinomios de Charlier, que se denotan por $C_n(x; a)$.

Supongamos ahora que p_{\pm} son de grado 1. Si los coeficientes líderes no son iguales, entonces w_x es lineal en una dirección o la otra y de vuelta el caso 3 no es posible, ya que no sería sumable. Si los coeficientes líderes son iguales, entonces asintóticamente $\varphi(x) - 1 \sim b/x$ para alguna constante b , lo que implica que los productos $\prod_{j=0}^x \varphi(j)$, y $\prod_{j=0}^x \varphi(j)^{-1}$ son los dos idénticamente 1 (y $p_+ = p_-$), o uno crece como x y el otro decrece como $1/x$ a medida que $x \rightarrow \infty$. Esto vuelve a excluir el caso 3. Entonces si los dos tienen igual grado tenemos los casos 1. y 2. con $p_-(0) = 0$.

Caso II: p_{\pm} de grado 1, soporte infinito. Normalizemos de tal manera que $p_-(x) = x$ y $p_+(x) = x(x+b)$ con $b, c > 0$ y $c < 1$ para que los momentos sean finitos. Por lo tanto $\varphi(x) = c(x+b)/(x+1)$ y resolviendo la ecuación de Pearson se tiene que

$$w_x = \frac{(b)_x}{x!} c^x w_0, \quad x = 0, 1, 2, \dots$$

Como $\sum_{x=0}^{\infty} w_x = w_0 (1-c)^{-b}$ (serie binomial) se puede escoger $w_0 = (1-c)^b$ para normalizar la medida. Esta distribución se conoce como la distribución de Pascal y sus correspondientes polinomios ortogonales son los polinomios de Meixner, que se denotan por $M_n(x; b, c)$.

Caso III: p_{\pm} de grado 1, soporte finito. Siguiendo el caso anterior, se puede normalizar con $p_{+}(x) = p(N - x)$ y $p_{-}(x) = qx$, donde $p, q > 0$, $p + q = 1$. Por lo tanto, resolviendo la ecuación de Pearson se tiene que

$$w_x = \binom{N}{x} \left(\frac{p}{q}\right)^x w_0, \quad x = 0, 1, 2, \dots, N$$

Como $\sum_{x=0}^N w_x = w_0 q^{-N}$ se puede escoger $w_0 = q^N$ para normalizar la medida. Esta distribución se conoce como la distribución binomial y sus correspondientes polinomios ortogonales son los polinomios de Krawtchouk, que se denotan por $K_n(x; p, N)$.

Caso IV: p_{\pm} de grado 2. Ambos tienen que tener el mismo coeficiente líder, con lo cual

$$\varphi(x) = \frac{p_{+}(x)}{p_{-}(x+1)} = \frac{1 + ax^{-1} + bx^{-2}}{1 + cx^{-1} + dx^{-2}} = 1 + \frac{a-c}{x} + \mathcal{O}\left(\frac{1}{x^2}\right)$$

La condición sobre los momentos obliga a excluir los casos 2 y 3, con lo cual sólo tenemos el caso 1. de soporte finito. Como $p_{-}(0) = 0 = p_{+}(N)$ hay dos casos, que se pueden normalizar mediante

$$p_{-}(x) = x(N + 1 + \beta - x), \quad p_{+}(x) = (N - x)(x + \alpha + 1), \quad \alpha, \beta > 1,$$

o

$$p_{-}(x) = x(x - \beta - N - 1), \quad p_{+}(x) = (N - x)(-x - \alpha - 1), \quad \alpha, \beta < -N.$$

Por lo tanto, resolviendo la ecuación de Pearson, la función peso es de la siguiente forma

$$\begin{aligned} w_x &= \frac{(N - x + 1)_x (\alpha + 1)_x w_0}{x! (N + \beta + 1 - x)_x} = \binom{N}{x} \frac{(\alpha + 1)_x (\beta + 1)_{N-x} w_0}{(\beta + 1)_N} \\ &= \frac{w_0}{(\beta + 1)_N} \binom{\alpha + x}{x} \binom{\beta + N - x}{N - x}. \end{aligned}$$

Como $\sum_{x=0}^N w_x = w_0 \frac{(\alpha + \beta + 2)_N}{N! (\beta + 1)_N}$ se puede escoger $w_0 = \frac{N! (\beta + 1)_N}{(\alpha + \beta + 2)_N}$ para normalizar la medida. Esta distribución se conoce como la distribución hipergeométrica y sus correspondientes polinomios ortogonales son los polinomios de Hahn, que se denotan por $Q_n(x; \alpha, \beta, N)$.

Así, podemos enunciar el siguiente

Teorema 2.4.1. *Salvo normalizaciones, las únicas familias de polinomios ortogonales que son autofunciones de un operador en diferencias de segundo orden simétrico de la forma 2.60 son las familias de Charlier, Meixner, Krawtchouk y Hahn. A estas familias se les suele denominar familias clásicas discretas.*

A continuación vemos con mayor detalle los polinomios clásicos discretos, pero antes, veremos la fórmula de Rodrigues en el caso discreto.

Fórmula de Rodrigues, caso discreto

Supongamos que w es un peso en los enteros y que L es simétrica con respecto a w y que tiene polinomios como eigenfunciones. La ecuación de los eigenvalores para un polinomio ψ_n de grado n es

$$(2.61) \quad p_+ \Delta_+ \psi_n - p_- \Delta_- \psi_n + \lambda_n \psi_n = 0.$$

Aplicando Δ_+ a esta ecuación obtenemos una ecuación para la "derivada" $\psi_n^{(1)}$. Usando la identidad de Leibniz en el caso discreto

$$\Delta_+(fg) = (S_+f)\Delta_+g + g\Delta_+f$$

obtenemos

$$p_+^{(1)} \Delta_+ \psi_n^{(1)} \Delta_+ \psi_n^{(1)} - p_-^{(1)} \Delta_- \psi_n^{(1)} + \lambda_n^{(1)} \psi_n^{(1)}$$

con

$$p_+^{(1)} = S_+p_+, \quad p_-^{(1)} = p_-, \quad \lambda_n^{(1)} = \lambda_n + \Delta_+(p_+ - p_-).$$

El nuevo operador es simétrico con respecto al peso $w^{(1)} = p_+w$:

$$S_+(w^{(1)}p_-^{(1)}) = (S_+p_+)S_+(wp_-) = p_+^{(1)}wp_+ = w^{(1)}p_+^{(1)}.$$

Continuando observamos que las diferencias $\psi_n^{(k)} = (\Delta_+)^k \psi_n$ satisfacen

$$p_+^{(k)} \Delta_+ \psi_n^{(k)} - p_- \Delta_- \psi_n^{(k)} + \lambda_n^{(k)} \psi_n^{(k)} = 0$$

con

$$p_+^{(k)} = S_+^k p_+$$

$$\lambda_n^{(k)} = \lambda_n + (I + S_+ + \dots + S_+^{k-1})\Delta_+p_+ - k\Delta_+p_-$$

y el correspondiente operador es simétrico con respecto al peso

$$w^{(k)} = w \prod_{j=0}^{k-1} S_+^j p_+.$$

Ahora $\psi_n^{(k)}$ tiene grado $n - k$. En particular, $\psi_n^{(n)}$ es constante, por lo que $\lambda_n^{(n)} = 0$, y hemos probado que

$$(2.62) \quad \begin{aligned} \lambda_n &= -(I + S_+ + \dots + S_+^{k-1})\Delta_+p_+ + n\Delta_+p_- \\ &= n\Delta_+(p_- - p_+) + \sum_{j=1}^{n-1} (I - S_+^j)\Delta_+p_+. \end{aligned}$$

La ecuación de eigenvalores puede ser reescrita para obtener $\psi_n^{(k-1)}$ de $\psi_n^{(k)}$, lo que conduce a la fórmula de Rodrigues en el caso discreto para ψ_n . En la primera etapa reescribimos el operador L usando las identidades (2.57) y $S_- \Delta_+ = \Delta_-$

$$wL = wp_+\Delta_+ - S_-(wp_+)\Delta_- = wp_+\Delta_+ - S_-(wp_+\Delta_+)$$

$$\Delta_-(wp_+\Delta_+) = \Delta_-(w^{(1)}\Delta_+).$$

Por lo tanto la ecuación de los eigenvalores puede ser resuelta para obtener

$$\psi_n = -\frac{1}{\lambda_n w} \Delta_-(w^{(1)}\psi_n^{(1)})$$

en los puntos donde w es positiva. Continuamos con este proceso y tomamos la constante $\psi_n^{(n)}$ a ser

$$(2.63) \quad \psi_n^{(n)} = \prod_{k=0}^{n-1} [\lambda_n + (S_+^k - I)p_+ - k\Delta_+p_-] = A_n.$$

El resultado es la fórmula de Rodrigues en el caso discreto, donde $w > 0$

$$(2.64) \quad \psi_n = (-1)^n \frac{1}{w} \Delta_-^n(w^{(n)}), \quad w^{(n)} = w \prod_{k=0}^{n-1} S_+^k p_+.$$

Como $\Delta_- = I - S_-$, podemos expandir $\Delta_-^n = (I - S_-)^n$ y reescribir (2.64) como

$$\psi_n = \frac{1}{w} \sum_{k=1}^n (-1)^{n-k} \binom{n}{k} S_-^k(w^{(n)}).$$

Como una aplicación, obtenemos una fórmula para la norma de ψ_n . Notemos que siempre que las sumas sean convergentes

$$\sum_k \Delta_- f(k)g(k) = -\sum_k f(k)\Delta_+g(k).$$

Por lo tanto

$$\begin{aligned} \lambda_n \|\psi_n\|_w^2 &= -(L\psi_n, \psi_n)_w = -\sum_k (wL\psi_n)(k) \\ &= -\sum_k [\Delta_-(wp_+\psi_n^{(1)})](k)\psi_n(k) = \sum_k [wp_+\psi_n^{(1)}](k)\psi_n^{(1)}(k) \\ &= \|\psi_n^{(1)}\|_{w^{(1)}}^2. \end{aligned}$$

Si siguiendo este proceso, obtenemos

$$A_n \|\psi_n\|_w^2 = \|\psi_n^{(n)}\|_{w^{(n)}}^2 = A_n^2 \|1\|_{w^{(n)}}^2$$

y de esta forma

$$(2.65) \quad \|\psi_n\|_w^2 = A_n \sum_k w^{(n)}(k).$$

Es conveniente escribir el operador (2.64) como

$$\frac{1}{w} \Delta_-^n w^{(n)} = \left(\frac{1}{w} \Delta_- w^{(1)} \right) \left(\frac{1}{w^{(1)}} \Delta_- w^{(2)} \right) \cdots \left(\frac{1}{w^{(n-1)}} \Delta_- w^{(n)} \right).$$

Se sigue de la condición de simetría (2.57) que

$$\frac{1}{w}\Delta_-(w^{(1)}f) = p_+f - \frac{S_-(p_+w)}{w}S_-f = p_+f - p_-S_-f.$$

Entonces

$$\frac{1}{w}\Delta_-w^{(1)} \cdot \frac{1}{w^{(1)}}\Delta_-w^{(2)} = (p_+ - p_-S_-)(p_+^{(1)} - p_-S_-) = p_+p_+^{(1)} - 2p_+p_-S_- + p_-(S_-p_-)S_-^2$$

y por inducción

$$(-1)^n \frac{1}{w}\Delta_-^n w^{(n)} \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \prod_{j=0}^{n-k-1} p_+^{(j)} \prod_{j=0}^{k-1} p_-^{(j)} S_-^k$$

donde $p_-^{(j)} = S_-^j p_-$. Aplicando este operador a la función constante 1 obtenemos

$$(2.66) \quad \psi_n = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \prod_{j=0}^{n-k-1} (S_+^j p_+) \prod_{j=0}^{k-1} (S_-^j p_-).$$

Una segunda aproximación para calcular ψ_n es usar el análogo discreto de las expansiones en series, con los monomios x^k remplazados por los polinomios

$$e_k(x) = (x - k + 1)_k = x(x - 1)(x - 2) \dots (x - k + 1) = (-1)^k (-x)_k.$$

Entonces,

$$\Delta_+ e_k(x) = k e_{k-1}(x), \quad x \Delta_- e_k(x) = k e_k(x), \quad x e_k = e_{k+1} + k e_k.$$

En cada uno de nuestros casos, $p_-(x)$ es divisible por x , así, aplicando el operador $L = p_+ \Delta_+ - p_- \Delta_-$ a la expansión

$$(2.67) \quad \psi_n(x) = \sum_{k=0}^n a_{nk} e_k(x)$$

nos lleva a las relaciones de recurrencia para los coeficientes a_{nk} que identifican a a_{nk} como un múltiplo de $a_{n,k-1}$.

Finalmente, resaltamos que tanto los coeficientes β_n de la relación de recurrencia a tres términos, como los eigenvalores λ_n , pueden ser recuperados de la ecuación (2.61) calculando los coeficientes de x^n y x^{n-1} .

$$(2.68) \quad \begin{aligned} 0 &= L\psi_n(x) + \lambda_n \psi_n(x) \\ &= p_+ \Delta_+(a_n x^n + b_n x^{n-1}) - p_- \Delta_-(a_n x^n + b_n x^{n-1}) + \lambda_n a_n x^n + \lambda_n b_n x^{n-1} + \dots \\ &= (p_+ - p_-)[n a_n x^{n-1} + (n-1) b_n x^{n-2}] + (p_+ + p_-) \left[\binom{n}{2} a_n x^{n-2} + \binom{n-1}{2} b_n x^{n-3} \right] \\ &\quad + \lambda_n a_n x^n + \lambda_n b_n x^{n-1}. \end{aligned}$$

El coeficiente de x^n del lado derecho debe anularse, y esto determina a λ_n . Usando este valor de λ_n en el coeficiente x^{n-1} determinamos el radio b_n/a_n y por lo tanto el término $\beta_n = b_n/a_n - b_{n+1}/a_{n+1}$ en la relación de recurrencia a tres términos.

Los polinomios de Charlier

El intervalo es infinito y

$$p_+(x) = 1, \quad p_-(x) = \frac{x}{a}, \quad w_m = e^{-a} \frac{a^m}{m!}.$$

Notemos que w_m corresponde a la distribución Poisson.

Entonces $p_+^{(k)} = p_+$, $w^{(k)} = w$ y

$$\lambda_n = \frac{n}{a}.$$

Por lo tanto, la constante A_n en (2.63) es $n!/a^n$. De (2.65) obtenemos la norma:

$$\|\psi_n\|_w^2 = \frac{n!}{a^n}.$$

Una normalización estándar es $C_n(x; a) = (-1)^n \psi_n(x)$. La ecuación (2.66) da como resultado

$$\begin{aligned} (2.69) \quad C_n(x; a) &= \sum_{k=0}^n (-1)^k \binom{n}{k} a^{-k} (x - k + 1)_k \\ &= \sum_{k=0}^n \frac{(-n)_k (-x)_k}{k!} \left(-\frac{1}{a}\right)^k \end{aligned}$$

$$(2.70) \quad = {}_2F_0 \left(-n, -x; -\frac{1}{a} \right)$$

donde ${}_2F_0$ es un serie hipergeométrica generalizada. El coeficiente líder es $(-a)^{-n}$. En este caso $\Delta_+ C_n$ es un polinomio ortogonal con respecto al mismo peso, y comparando los coeficientes líderes observamos que

$$(2.71) \quad \Delta_+ C_n(x; a) = -\frac{n}{a} C_{n-1}(x; a).$$

Su ecuación en diferencias asociada es

$$\Delta_+ C_n(x; a) - \frac{x}{a} \Delta_- C_n(x; a) + \frac{n}{a} C_n(x; a) = 0.$$

Los primeros cuatro polinomios son

$$C_0(x; a) = 1;$$

$$C_1(x; a) = -\frac{x}{a} + 1;$$

$$C_2(x; a) = \frac{x(x-1)}{a^2} - \frac{2x}{a} + 1 = \frac{x^2}{a^2} - (1+2a)\frac{x}{a^2} + 1;$$

$$\begin{aligned} C_3(x; a) &= -\frac{x(x-1)(x-2)}{a^3} + \frac{3x(x-1)}{a^2} - \frac{3x}{a} + 1 \\ &= -\frac{x^3}{a^3} + 3(a+1)\frac{x^2}{a^3} - (3a^2 + 3a + 2)\frac{x}{a^3} + 1. \end{aligned}$$

Podemos calcular la función generadora

$$G(x, t; a) = \sum_{n=0}^{\infty} \frac{C_n(x; a)}{n!} t^n.$$

Notemos que el término constante de $C_n(x; a)$ es 1, así

$$G(x, 0; a) = 1, \quad G(0, t; a) = e^t.$$

En general,

$$\begin{aligned} G(x, t; a) &= \sum_{n=0}^{\infty} \sum_{k=0}^n t^n \frac{(-x)_k}{(n-k)!k!} \left(\frac{1}{a}\right)^k \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^m \frac{t^m}{m!} \frac{(-x)_k}{(n-k)!k!} \left(\frac{t}{a}\right)^k \\ &= e^t \left(1 - \frac{t}{a}\right)^x. \end{aligned}$$

La igualdad

$$\sum_{n=0}^{\infty} \frac{C_{n+1}(x; a)}{n!} t^n = \frac{\partial G}{\partial t}(x, t; a)$$

nos lleva a la relación de recurrencia

$$C_{n+1}(x; a) = C_n(x; a) - \frac{x}{a} C_n(x-1; a).$$

La identidad

$$G(x+1, t; a) = \left(1 - \frac{t}{a}\right) G(x, t; a)$$

lleva a (2.71). La relación de recurrencia a tres términos general (2.54) se calcula. Se sigue de (2.70) que

$$(-1)^n C_n(x; a) = \frac{1}{a^n} x^n - \frac{\binom{n}{2} + na}{a^n} x^{n-1}$$

$$(-1)^n C_n(x; a) = \frac{1}{a^n} x^n - \frac{\binom{n}{2} + na}{a^n} x^{n-1}.$$

Así

$$xC_n(x; a) = -aC_{n+1}(x; a) + (n+a)C_n(x; a) - nC_{n-1}(x; a).$$

Los polinomios de Charlier están relacionados con los polinomios de Laguerre

$$C_n(x; a) = (-1)^n \frac{n!}{a^n} L_n^{(x-n)}(a).$$

Los polinomios de Krawtchouk

El intervalo es finito y

$$p_+(x) = p(N-x), \quad p_-(x) = qx, \quad w_m = \binom{N}{m} p^m q^{N-m}$$

donde $p, q > 0, p+q=1$. Notemos aquí que w_m se identifica con la distribución binomial. Entonces $\lambda_n = n$ y la constante A_n en (2.63) es $n!$. Dos normalizaciones estándar son

$$k_n^{(p)}(x, N) = \frac{1}{n!} \psi_n(x)$$

$$K_n(x; p, N) = (-1)^n \frac{(N-n)!}{N! p^n} \psi_n(x) = (-1)^n \binom{N}{n}^{-1} \frac{1}{p^n} k_n^{(p)}(x; N).$$

De (2.65) obtenemos las normas:

$$\|k_n^{(p)}\|_w^2 = \binom{N}{n} (pq)^n$$

$$\|K_n\|_w^2 = \binom{N}{n}^{-1} \left(\frac{q}{p}\right)^n.$$

La mayoría de las identidades tienen una versión más simple en su forma $k_n^{(p)}$. La ecuación (2.66) da

$$(2.72) \quad k_n^{(p)}(x; N) = \sum_{k=0}^n p^{n-k} q^k \frac{(x-N)_{n-k} (x-k+1)_k}{(n-k)! k!}.$$

El coeficiente líder es

$$\frac{1}{n!} \sum_{k=0}^n \binom{n}{k} p^{n-k} q^k = \frac{(p+q)^n}{n!} = \frac{1}{n!}.$$

El polinomio $\Delta_+ k_n^{(p)}$ es una eigenfunción para el peso $w^{(1)}$, el cual, después de normalizar, es el peso asociado a p y $N-1$. Tomando en cuenta los coeficientes líderes, se sigue que

$$\Delta_+ k_n^{(p)}(x; N) = k_{n-1}^{(p)}(x; N-1).$$

La ecuación en diferencias asociada es

$$p(N-x)\Delta_+ k_n^{(p)}(x; N) - qx\Delta_- k_n^{(p)}(x; N) + nk_n^{(p)}(x; N) = 0.$$

Usando esta ecuación y la expansión (2.67), podemos derivar una segunda forma de los polinomios de Krawtchouk:

$$k_n^{(p)}(x; N) = (-p)^n \binom{N}{n} \sum_{k=0}^n \frac{(-n)_k (-x)_k}{(-N)_k k!} p^{-k}$$

$$= (-p)^n \binom{N}{n} {}_2F_1 \left(-n, -x; -N; \frac{1}{p} \right)$$

donde F es la función hipergeométrica. La normalización de la forma alternativa K_n se escoge de tal manera que

$$K_n(x; p, N) = {}_2F_1 \left(-n, -x, -N; \frac{1}{p} \right).$$

Los primeros cuatro polinomios son

$$\begin{aligned} k_0^{(p)}(x; N) &= 1 \\ k_1^{(p)}(x; N) &= x - Np \\ k_2^{(p)}(x; N) &= \frac{1}{2}[x^2 + (2p - 1 - 2Np)x + N(N - 1)p^2] \\ k_3^{(p)}(x; N) &= \frac{1}{6}\{x^3 + (6p - 3 - 3Np)x^2 + [3Np(Np + 1 - 3p) \\ &\quad + 2(3p^2 - 3p + 1)]x - N(N - 1)(N - 2)p^3\}. \end{aligned}$$

Podemos considerar $k_n^{(p)}(x; N)$ como definida por (2.72) para todos los $n = 0, 1, 2, \dots$. Notemos que para $n > N$ estos polinomios se anulan en los puntos $m = 0, 1, 2, \dots, N$. La función generadora es

$$G(x, t; N, p) = \sum_{n=0}^{\infty} k_n^{(p)}(x; N)t^n = (1 + qt)^x (1 - pt)^{N-x}.$$

La identidad

$$\sum_{n=0}^{\infty} (n+1)k_{n+1}^{(p)}(x; N)t^n = \frac{\partial G}{\partial t}(x, t; N, p)$$

lleva a la relación de recurrencia

$$(n+1)k_{n+1}^{(p)}(x; N) = xqk_n^{(p)}(x-1; N-1) - (N-x)pk_n^{(p)}(x; N-1).$$

La identidad

$$(1 - pt)G(x+1, t; N, p) = (1 + qt)G(x, t; N, p)$$

lleva a

$$k_n^{(p)}(x+1; N) - k_n^{(p)}(x; N) = k_{n-1}^{(p)}(x; N).$$

La relación de recurrencia a tres términos

$$xk_n^{(p)}(x; N) = (n+1)k_{n+1}^{(p)}(x; N) + (pN + n - 2pn)k_n^{(p)}(x; N) + pq(N - n + 1)k_{n-1}^{(p)}(x; N)$$

puede ser calculada usando (2.68)

Los polinomios de Meixner

El intervalo es infinito y

$$p_+(x) = c(x + b), \quad p_-(x) = x, \quad w_m = (1 - c)^b \frac{(b)_m}{m!} c^m$$

donde $b > 0$ y $0 < c < 1$. Aquí w_m se identifica con la distribución de Pascal. Por lo tanto

$$\lambda_n = (1 - c)n \quad A_n = (1 - c)^n n!.$$

La ecuación (2.66) nos da como resultado

$$\psi_n(x) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} (x + b)_{n-k} (x - k + 1)_k c^{n-k}$$

con coeficiente líder $(1 - c)^n$. Dos normalizaciones estándar son:

$$m_n(x; b, c) = (-c)^{-n} \psi_n(x) = \sum_{k=0}^n \binom{n}{k} (x + b)_{n-k} (x - k + 1)_k (-c)^{-k}$$

$$M_n(x; b, c) = \frac{(-1)^n}{c^n (b)_n} \psi_n(x) = \frac{1}{(b)_n} m_n(x; b, c).$$

Se sigue que

$$\|m_n\|_w^2 = n! \frac{(b)_n}{c^n}$$

$$\|M_n\|_w^2 = \frac{n!}{(b)_n c^n}.$$

La mayoría de las identidades tienen una forma más simple en la versión m_n . El polinomio $\Delta_+ m_n$ es una eigenfunción para el peso $w^{(1)}$, el cual, después de normalización, es el peso asociado con $b + 1$ y c . El coeficiente líder de m_n es $(1 - 1/c)^n$ por lo que

$$\Delta_+ m_n(x; b, c) = n \left(1 - \frac{1}{c}\right) m_{n-1}(x; b + 1, c).$$

La ecuación en diferencias asociada es

$$c(x + b)\Delta_+ m_n(x; b, c) - x\Delta_- m_n(x; b, c) + (1 - c)nm_n(x; b, c) = 0.$$

Usando esta ecuación y la expansión (2.67) obtenemos una segunda forma:

$$(2.73) \quad m_n(x; b, c) = (b)_n \sum_{k=0}^n \frac{(-n)_k (-x)_k}{(b)_k k!} \left(1 - \frac{1}{c}\right)^k = (b)_n {}_2F_1\left(-n, -x, b; 1 - \frac{1}{c}\right)$$

donde de nuevo F es la función hipergeométrica. La normalización de M_n se escoge de manera que

$$M_n(x; b, c) = {}_2F_1\left(-n, -x, b; 1 - \frac{1}{c}\right).$$

Los primeros cuatro polinomios son:

$$m_0(x; b, c) = 1$$

$$\begin{aligned}
m_1(x; b, c) &= \left(1 - \frac{1}{c}\right)x + b \\
m_2(x; b, c) &= \left(1 - \frac{1}{c}\right)^2 x^2 + \left(2b + 1 - \frac{2b}{c} - \frac{1}{c^2}\right)x + b(b+1) \\
m_3(x; b, c) &= \left(1 - \frac{1}{c}\right)^3 x^3 + \left(3b + 3 - \frac{6b+3}{c} + \frac{3b-3}{c^2} + \frac{3}{c^3}\right)x^2 \\
&\quad + \left(3b^2 + 6b + 2 - \frac{3b^2+3b}{c} - \frac{3b}{c^2} - \frac{2}{c^3}\right)x + b(b+1)(b+2).
\end{aligned}$$

La función generadora es

$$G(x; b, c) = \sum_{n=0}^{\infty} \frac{m_n(x; b, c)}{n!} t^n = (1-t)^{-x-b} \left(1 - \frac{t}{c}\right)^x.$$

La identidad

$$\sum_{n=0}^{\infty} \frac{m_{n+1}(x; b, c)}{n!} t^n = \frac{\partial G}{\partial t}(x; b, c)$$

implica la relación de recurrencia

$$m_{n+1}(x; b, c) = (x+b)m_n(x; b+1, c) - \frac{x}{c}m_n(x-1; b+1, c).$$

La identidad

$$(1-t)G(x+1; b, c) = \left(1 - \frac{t}{c}\right)G(x; b, c)$$

implica

$$m_n(x+1; b, c) - m_n(x; b, c) = nm_{n-1}(x+1; b, c) - \frac{n}{c}m_{n-1}(x; b, c).$$

La relación de recurrencia a tres términos

$$(c-1)xm_n(x; b, c) = cm_{n+1}(x; b, c) - (bc + nc + n)m_n(x; b, c) + n(b+n-1)m_{n-1}(x; b, c)$$

puede ser derivada usando (2.68).

Los polinomios de Hahn

El intervalo es finito y tomamos

$$\begin{aligned}
p_+(x) &= (N-x)(x+\alpha+1), & p_-(x) &= x(N+\beta+1-x) \\
\alpha, \beta &> -1 & \text{o} & \alpha, \beta < -N.
\end{aligned}$$

En el caso $\alpha, \beta < -N$ hemos violado nuestra condición de que p_{\pm} sea positiva en el interior de los puntos $\{1, 2, \dots, N-1\}$. El peso es

$$w_m = C_0 \binom{N}{m} (\alpha+1)_m (\beta+1)_{N-m}$$

donde

$$C_0 = \frac{\Gamma(N + \alpha + \beta + 2)}{\Gamma(\alpha + \beta + 2)}.$$

Aquí w_m se identifica con la distribución hipergeométrica. Con esta elección de C_0 la masa total es

$$\sum_{m=0}^N w_m = 1.$$

De acuerdo con (2.66)

$$(2.74) \quad \psi_n(x) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} \prod_{j=0}^{n-k-1} p_+(x+j) \prod_{j=0}^{k-1} p_-(x-j).$$

Esto parece tener grado $2n$ más que n , pero hay muchos términos que se cancelan. Una forma más útil puede ser obtenida usando (2.67), lo cual conduce a

$$\psi_n(x) = C \sum_{k=0}^n \frac{(-n)_k (-x)_k (n + \alpha + \beta + 1)_k}{(-N)_k (\alpha + 1)_k k!}$$

donde ${}_3F_2$ denota la serie generalizada hipergeométrica. Para determinar la constante C notemos que el término constante en (2.74) proviene del sumando con $k = 0$ y es por lo tanto

$$(-1)^n \prod_{j=0}^{n-1} p_+(j) = (-1)^n (N + 1 - n)_n (\alpha + 1).$$

Se sigue que

$$\psi_n(x) = (-1)^n (N + 1 - n)_n (\alpha + 1)_n \cdot {}_3F_2(-n, -x, n + \alpha + \beta + 1; -N, \alpha + 1; 1).$$

Una normalización es

$$(2.75) \quad \begin{aligned} Q_n(x; \alpha, \beta, N) &= {}_3F_2(-n, -x, n + \alpha + \beta + 1; -N, \alpha + 1; 1) \\ &= \sum_{k=0}^n \frac{(-n)_k (-x)_k (n + \alpha + \beta + 1)_k}{(-N)_k (\alpha + 1)_k k!}. \end{aligned}$$

Una segunda es

$$h_n^{(\alpha, \beta)}(x, N) = (-1)^n \frac{(N - n)_n (\beta + 1)_n}{n!} Q_n(x; \beta, \alpha, N - 1).$$

Se sigue de (2.62) que $\lambda_n = n(n + \alpha + \beta + 1)$, así

$$\|\psi_n\|_w^2 = (n!)^2 \binom{N}{n} \frac{n!(N - n)!(\beta + 1)_n (\alpha + \beta + n + 1)_{N+1}}{(\alpha + \beta + 2)_N (\alpha + \beta + 2n + 1)}.$$

Por lo tanto

$$\|Q_n(x; \alpha, \beta, N)\|^2 = \frac{n!(N - n)!(\beta + 1)_n (\alpha + \beta + n + 1)_{N+1}}{N!(\alpha + 1)_n (\alpha + \beta + 2)_N (\alpha + \beta + 2n + 1)}.$$

El peso $w^{(1)} = p_+ w$ es un múltiplo constante del peso asociado con índices $(\alpha+1, \beta+1, N-1)$. Se sigue de (2.75) que el coeficiente líder de $Q_n(x; \alpha, \beta, N)$ es

$$\frac{(n + \alpha + \beta + 1)_n}{(-N)_n(\alpha + 1)_n}$$

así

$$\Delta_+ Q_n(x; \alpha, \beta, N) = -\frac{n(\alpha + \beta + n + 1)}{N(\alpha + 1)} Q_{n-1}(x; \alpha + 1, \beta + 1, N - 1).$$

La ecuación en diferencias es

$$(N - x)(x + \alpha + 1)\Delta_+ Q_n(x; \alpha, \beta, N) - x(N + \beta + 1 - x)\Delta_- Q_n(x; \alpha, \beta, N) + n(n + \alpha + \beta + 1)Q_n(x; \alpha, \beta, N) = 0.$$

Los primeros tres polinomios son:

$$\begin{aligned} Q_0(x; \alpha, \beta, N) &= 1 \\ Q_1(x; \alpha, \beta, N) &= -\frac{\alpha + \beta + 2}{N(\alpha + 1)}x + 1 \\ Q_2(x; \alpha, \beta, N) &= \frac{(\alpha + \beta + 3)(\alpha + \beta + 4)}{N(N - 1)(\alpha + 1)(\alpha + 2)}x^2 \\ &\quad - \frac{(\alpha + \beta + 3)[\alpha + \beta + 4 + 2(N - 1)(\alpha + 2)]}{N(N - 1)(\alpha + 1)(\alpha + 2)}x + 1. \end{aligned}$$

Una aplicación directa de (2.68) nos da

$$\begin{aligned} Q_n(x; \alpha, \beta, N) &= a_n x^n + b_n x^{n-1} + \dots \\ \frac{b_n}{a_n} &= -\frac{n[2N(\alpha + 1) + (2N + \beta - \alpha)(n - 1)]}{2(\alpha + \beta + 2n)}. \end{aligned}$$

El radio de los coeficientes líderes es

$$\frac{a_n}{a_{n+1}} = -\frac{(n + \alpha + \beta + 1)(\alpha + n + 1)(N - n)}{(2n + \alpha + \beta + 1)(2n + \alpha + \beta + 2)}.$$

Por lo tanto, por (2.54), la relación de recurrencia a tres términos es

$$xQ_n(x; \alpha, \beta, N) = \alpha_n Q_{n+1}(x; \alpha, \beta, N) + \beta_n Q_n(x; \alpha, \beta, N) + \gamma_n Q_{n-1}(x; \alpha, \beta, N)$$

$$\begin{aligned} \alpha_n &= -\frac{(\alpha + \beta + n + 1)(\alpha + n + 1)(N - n)}{(\alpha + \beta + 2n + 1)(\alpha + \beta + 2n + 2)} \\ \gamma_n &= \frac{n(n + \beta)(\alpha + \beta + n + N + 1)}{(\alpha + \beta + 2n)(\alpha + \beta + 2n + 1)} \\ \beta_n &= -(\alpha_n + \gamma_n). \end{aligned}$$

Polinomios de Hahn duales

Sea

$$\lambda(x) = x(x + \gamma + \delta + 1),$$

y considérese el peso

$$w_x(\gamma, \delta, N) = \frac{(2x + \gamma + \delta + 1)(\gamma + 1)_x (-N)_x N!}{(-1)^x (x + \gamma + \delta + 1)_{N+1} (\delta + 1)_x x!}.$$

Los polinomios de Hahn duales $R_n(\lambda(x); \gamma, \delta, N)$ son ortogonales con respecto a la medida anterior y la norma viene dada por

$$\|R_n(\lambda(x); \gamma, \delta, N)\|_w^2 = \binom{\gamma + n}{n}^{-1} \binom{\delta + N - n}{N - n}^{-1}.$$

Se tiene la siguiente relación de dualidad con los polinomios de Hahn:

$$(2.76) \quad R_n(\lambda(x); \gamma, \delta, N) = Q_x(n; \gamma, \delta, N).$$

Por lo tanto, los polinomios duales de Hahn verifican ecuaciones en diferencias y fórmulas de recurrencia a tres términos ‘duales’ de las de Hahn. Específicamente, la ecuación en diferencias de segundo orden es

$$\begin{aligned} B(x)R_n(\lambda(x+1); \gamma, \delta, N) - (B(x) + D(x))R_n(\lambda(x); \gamma, \delta, N) + D(x)R_n(\lambda(x-1); \gamma, \delta, N) \\ = -nR_n(\lambda(x); \gamma, \delta, N), \end{aligned}$$

donde

$$B(x) = \frac{(x + \gamma + \delta + 1)(x + \gamma + 1)(N - x)}{(2x + \gamma + \delta + 1)(2x + \gamma + \delta + 2)}, \quad D(x) = \frac{x(x + \gamma + \delta + N + 1)(x + \gamma)}{(2x + \gamma + \delta)(2x + \gamma + \delta + 1)}.$$

Mientras que la relación de recurrencia a tres términos es

$$\begin{aligned} A_n R_{n+1}(\lambda(x); \gamma, \delta, N) - [A_n + C_n] R_n(\lambda(x); \gamma, \delta, N) + C_n R_{n-1}(\lambda(x); \gamma, \delta, N) \\ = -\lambda(x) R_n(\lambda(x); \gamma, \delta, N), \end{aligned}$$

donde

$$A_n = (N - n)(n + \gamma + 1), \quad C_n = n(\delta + N + 1 - n).$$

Capítulo 3

Polinomios Ortogonales y Procesos estocásticos

3.1. Representación espectral para procesos de nacimiento y muerte

El teorema de representación de Karlin-McGregor de las probabilidades de transición es una herramienta muy útil para entender el comportamiento asintótico de los procesos de nacimiento y muerte. En este teorema aparecen los polinomios ortogonales y además estos también juegan un papel fundamental en el estudio de la ergodicidad exponencial y aparecen también en distribuciones importantes como la distribución límite condicional doble.

Consideremos un proceso de nacimiento y muerte en un espacio de estados $S = \{-1, 0, 1, 2, \dots\}$. El proceso está dado por sus tasas de nacimiento y muerte $\lambda_i, \mu_i, i \in S$ descritas en el capítulo 2. Como antes -1 es un estado absorbente y lo ignoramos si $\mu_0 = 0$ y 0 se convierte entonces en un estado reflejado.

3.1.1. Representación espectral de Karlin-McGregor

En el análisis de los procesos de nacimiento y muerte, son de especial importancia la sucesión de polinomios $\{Q_n(x), n \geq 0\}$, llamados **polinomios de nacimiento y muerte**. Estos están determinados únicamente por la relación de recurrencia

$$(3.1) \quad -xQ_n(x) = \mu_n Q_{n-1}(x) - (\lambda_n + \mu_n)Q_n(x) + \lambda_n Q_{n+1}(x)$$

junto con las condiciones iniciales $Q_{-1}(x) = 0$ y $Q_0(x) = 1$. Karlin y McGregor probaron que la función de transición P puede ser representada como

$$(3.2) \quad P_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i) = \pi_j \int_0^\infty e^{-xt} Q_i(x) Q_j(x) d\phi(x), \quad i, j = 0, 1, \dots, t \geq 0$$

donde ϕ es una medida positiva de Borel con masa total 1 y con soporte en los reales no negativos; ϕ es llamada la **medida espectral** de P . Tomando $t = 0$ en (3.2) se puede ver fácilmente que los polinomios $\{Q_n(x), n \geq 0\}$ son ortogonales con respecto a ϕ .

Es de destacarse que en la representación (3.2) de $P_{ij}(t)$ la dependencia sobre t en el lado derecho de la ecuación está restringida completamente al término exponencial decreciente e^{-tx} y que la dependencia en i y j se factoriza como el producto $Q_i(x)Q_j(x)$.

Para ciertas elecciones de parámetros en los procesos de nacimiento y muerte, las ecuaciones en (3.1) son relaciones de recurrencia reconocibles que definen polinomios ortogonales clásicos para los cuales las medidas de ortogonalización son bien conocidas. Bajo las condiciones (1.28) y (1.30), la medida espectral en (3.2) puede ser solamente una de estas medidas, y la función de transición para nuestros procesos de nacimiento y muerte puede ser calculada directamente de (3.2).

Para el esbozo del método, recordemos que se asume que los coeficientes $\lambda_i, i \geq 0$ y $\mu_i, i \geq 0$ son estrictamente positivos y que $\mu_0 \geq 0$. Si μ_0 es positivo puede ser interpretado como una tasa de absorción del estado cero en un estado con la propiedad de que una vez llegado a ese estado jamás se le puede abandonar. Las relaciones de recurrencia

$$\begin{aligned} -xQ_0(x) &= -(\lambda_0 + \mu_0)Q_0(x) + \lambda_0Q_1(x) \\ (3.3) \quad -xQ_n(x) &= \mu_nQ_{n-1}(x) - (\lambda_n + \mu_n)Q_n(x) + \lambda_nQ_{n+1}(x) \end{aligned}$$

se pueden escribir de forma más compacta como $-xQ = \mathcal{A}Q$. La representación integral puede ser deducida de la siguiente forma: Se escribe la sucesión de funciones

$$f_i(x, t) = \sum_{j=0}^{\infty} P_{ij}(t)Q_j(x)$$

o equivalentemente el vector

$$f(x, t) = P(t)Q(x).$$

Este vector satisface la ecuación

$$\frac{\partial f(x, t)}{\partial t} = P'(t)Q(x) = P(t)\mathcal{A}Q(x) - xf(x, t)$$

y la condición inicial

$$f(x, 0) = Q(x).$$

Así

$$f(x, t) = e^{-xt}Q(x)$$

o

$$f_i(x, t) = e^{-xt}Q_i(x).$$

Ahora, $P_{ij}(t)$ es el j -ésimo **coeficiente de Fourier** de $f_i(x, t)$ y por lo tanto

$$P_{ij}(t) = \pi_i \int_0^{\infty} e^{-xt}Q_i(x)Q_j(x)d\psi(x)$$

donde

$$\int_0^{\infty} Q_j^2(x)d\psi(x) = \frac{1}{\pi_j}.$$

Lo anterior se puede enunciar de forma más rigurosa con el siguiente

Teorema 3.1.1. *Sea \mathcal{A} la matriz del operador infinitesimal asociado al proceso de nacimiento y muerte y $P_{ij}(t)$ una función de transición que es solución de ambas ecuaciones de Kolmogorov (por ejemplo la minimal). Entonces existe una distribución de probabilidad ψ soportada en el intervalo $[0, \infty)$ tal que se tiene la representación integral*

$$(3.4) \quad P_{ij}(t) = \pi_i \int_0^\infty e^{-xt} Q_i(x) Q_j(x) d\psi(x), \quad i, j \in \mathbf{E}, \quad t \geq 0.$$

La representación espectral (3.4) separa claramente las variables i, j, t . Además a medida que $t \rightarrow \infty$ la integral va tendiendo a 0, excepto en el punto $x = 0$, donde puede haber un salto. Para que ocurra este salto es necesario y suficiente que el proceso sea recurrente positivo, en cuyo caso la magnitud del salto viene dada por

$$(3.5) \quad \psi(\{0\}) = \left(\sum_{i=0}^{\infty} \pi_i Q_i^2(0) \right)^{-1}.$$

Obsérvese también que de (3.4) que haciendo $t = 0$ obtenemos la ortogonalidad de los polinomios $Q_n(x)$ definidos en 3.3, i.e.

$$\int_0^\infty Q_i(x) Q_j(x) d\psi(x) = \frac{\delta_{ij}}{\pi_j}, \quad i, j \in \mathbf{E}.$$

Una medida que cumple lo anterior se le llama una solución del problema de momentos de Stieltjes. Mostramos a continuación que dicha medida existe, pero antes, estableceremos la correspondencia entre el conjunto de todas las matrices \mathcal{A} que pertenecen a un proceso de nacimiento y muerte y el conjunto de todos los momentos que son solución de un **problema de momentos de Stieltjes**, que dice que el proceso con μ_0 genera todos los problemas de momentos de Stieltjes, y aquellos con $\mu_0 > 0$ genera todos los que tienen una solución con un momento finito de orden menos uno. Antes haremos una observación importante.

Ya vimos anteriormente que una sucesión de polinomios $\{Q_n(x)\}$ que cumple con la relación de recurrencia a tres términos

$$(3.6) \quad \begin{aligned} Q_0(x) &= 1 \\ -xQ_0(x) &= -(\lambda_0 + \mu_0)Q_0(x) + \lambda_0Q_1(x) \\ -xQ_n(x) &= \mu_nQ_{n-1}(x) - (\lambda_n + \mu_n)Q_n(x) + \lambda_nQ_{n+1}(x), \quad n \geq 1 \end{aligned}$$

se llaman **polinomios que pertenecen a la matriz \mathcal{A}** o también como polinomios de nacimiento y muerte. Como cada λ_n es positivo, $Q_n(x)$ es un polinomio en x de grado exactamente n , el coeficiente de x^n en $Q_n(x)$ es $(-1)^n / \lambda_0 \lambda_1 \cdots \lambda_{n-1}$.

Definimos

$$\pi_0 = 1, \quad \pi_n = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n}.$$

Como $\lambda_n \pi_n = \mu_{n+1} \pi_{n+1}$ la relación de recurrencia a tres términos puede ser escrita de nuevo como

$$-xQ_0(x)\pi_0 = \lambda_0\pi_0[Q_1(x) - Q_0(x)] - \mu_0$$

$$-xQ_n(x)\pi_n = \lambda_n\pi_n[Q_{n+1}(x) - Q_n(x)] - \lambda_{n-1}\pi_{n-1}[Q_n(x) - Q_{n-1}(x)]$$

y como consecuencia

$$(3.7) \quad -x \sum_{j=0}^n Q_j(x)\pi_j = \lambda_n\pi_n[Q_{n+1}(x) - Q_n(x)] - \mu_0 \quad n \geq 0$$

y en consecuencia

$$(3.8) \quad Q_n(x) = 1 + \mu_0 \sum_{m=0}^{n-1} \frac{1}{\lambda_m\pi_m} - x \sum_{m=0}^{n-1} \frac{1}{\lambda_m\pi_m} \sum_{j=0}^m Q_j(x)\pi_j,$$

y se sigue por inducción que para $x < 0$

$$1 = Q_0(x) < Q_1(x) < \cdots < Q_n(x) < Q_{n+1}(x) < \cdots .$$

Si $n \geq 1$

$$Q_n(0) = 1 + \mu_0 \sum_{k=0}^{n-1} \frac{1}{\lambda_k\pi_k}$$

y por lo tanto las desigualdades también son válidas cuando $x = 0$ si $\mu_0 > 0$.
Enunciamos ahora el teorema que es una extensión de un resultado de Favard.

Teorema 3.1.2. *Existe al menos una medida positiva, regular ψ en $0 \leq x < \infty$ de masa total 1 no soportada por un conjunto finito de puntos tal que*

$$\int_0^\infty Q_i(x)Q_j(x)d\psi(x) = \begin{cases} 0 & \text{si } i \neq j \\ \frac{1}{\pi_i} & \text{si } i = j \end{cases}$$

Demostración. El sistema de ecuaciones $\int Q_0 d\psi = 1$, $\int Q_n d\psi = 0$, $n > 0$ puede ser resuelto recursivamente para los momentos $c_n = \int x^n d\psi$. Por ejemplo $\int Q_0 d\psi = 1$ da $c_0 = 1$ y entonces

$$0 = \int Q_1 d\psi = \int \frac{\lambda_0 + \mu_0 - x}{\lambda_0} d\psi$$

da $c_1 = \lambda_0 + \mu_0$. Se puede demostrar [21] que hay una cantidad infinita de medidas en $0 \leq x < \infty$ regulares, pero no necesariamente positivas tales que tienen estos momentos. Sea ψ una de esas medidas.

De la relación de recurrencia se sigue que si $n \geq 1$ entonces

$$\int_0^\infty x^k Q_n(x) d\psi_x = 0, \quad 0 \leq k < n$$

y

$$\int_0^\infty (-x)^n Q_n(x) d\psi(x) = \mu_n \int_0^\infty (-x)^{n-1} Q_{n-1}(x) d\psi(x).$$

Así

$$\int_0^\infty Q_m(x)Q_n(x)d\psi(x) = \frac{\delta_{mn}}{\pi_n}, \quad m, n \geq 0.$$

Para probar el teorema basta probar que los determinantes

$$\Delta_n = \begin{vmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & \cdots & c_{2n} \end{vmatrix} \quad \Delta_n^{(1)} = \begin{vmatrix} c_1 & c_2 & \cdots & c_{n+1} \\ c_2 & c_3 & \cdots & c_{n+2} \\ \vdots & \vdots & & \vdots \\ c_{n+1} & c_{n+2} & \cdots & c_{2n+1} \end{vmatrix}$$

son estrictamente positivos, cosa que ya hicimos. Para $n \geq 1$ el determinante

$$P_n(x) = \begin{vmatrix} c_0 & c_1 & \cdots & c_{n-1} & 1 \\ c_1 & c_2 & \cdots & c_n & x \\ \vdots & \vdots & & \vdots & \\ c_n & c_{n+1} & \cdots & c_{2n-1} & x^n \end{vmatrix}$$

es un polinomio de grado $\leq n$ y

$$\int_0^\infty x^k P_n(x)d\psi(x) = 0, \quad k = 0, 1, 2, \dots, n-1.$$

Por lo tanto $P_n(x)$ es un múltiplo constante de $Q_n(x)$, y comparando coeficientes de x^n

$$P_n(x) = (-1)^n(\lambda_0\lambda_1 \dots \lambda_{n-1})\Delta_{n-1}Q_n(x).$$

Se sigue que

$$\Delta_n = \int_0^\infty x^n P_n(x)d\psi(x) = \frac{(\lambda_0\lambda_1 \dots \lambda_{n-1})^2}{\pi_n} \Delta_{n-1}$$

y como $\Delta_0 = 1$ entonces $\Delta_n > 0, n = 1, 2, \dots$. Además

$$\Delta_n^{(1)} = (-1)^{n+1}P_{n+1}(0) = (\lambda_0\lambda_1 \dots \lambda_n)\Delta_n Q_{n+1}(0) > 0.$$

Lo cual completa la prueba. □

A partir de aquí, una medida positiva con las propiedades requeridas en el teorema anterior se llamará una solución del problema de momentos S. Una medida no necesariamente positiva en $0 \leq x < \infty$ con los mismos momentos se llamará una solución del problema de momentos BVS. Una medida positiva (no necesariamente positiva) en $-\infty < x < \infty$ con los mismos momentos se llamará una solución del problema de momentos H (BVH). S y H son abreviaciones de Stieltjes y Hamburger respectivamente mientras que B.V. vienen del inglés Bounded Variation, es decir, de variación acotada.

El teorema anterior establece una correspondencia entre una matriz de nacimiento y muerte o matriz de Jacobi \mathcal{A} y un problema con solución de Stieltjes. Hay un teorema inverso. Supongamos que $\{Q_n(x)\}$ es una sucesión de polinomios reales, el n -ésimo polinomio de grado exactamente n , ortogonal en $0 \leq x < \infty$ con respecto a una medida ψ . Los ceros del polinomio son entonces interiores al intervalo $0 \leq x < \infty$ y puede por lo tanto asumirse que $Q_n(0) = 1$ para toda n .

Los polinomios satisfacen una relación de recurrencia

$$\begin{aligned}
-xQ_0(x) &= B_0Q_0(x) + C_0Q_1(x) \\
-xQ_n(x) &= A_nQ_{n-1}(x) + B_nQ_n(x) + C_nQ_{n+1}(x)
\end{aligned}$$

donde A_n, B_n, C_n son reales. Tenemos $B_0 + C_0 = 0$ y $A_n + B_n + C_n = 0, n \geq 1$ debido a la normalización. El coeficiente de x^n en $Q_n(x)$ es $(-1)^n/C_0C_1 \dots C_n$ y como $Q_n(0) = 1$ esto es igual a $(-1)^n$ multiplicado por el producto de los recíprocos de las n raíces positivas de $Q_n(x)$. Por lo tanto $C_n > 0$ para $n \geq 1$. Si $n \geq 1$ entonces

$$\int_0^\infty (-x)^n Q_n(x) d\psi(x) = A_n/C_{n-1} \int_0^\infty (-x)^{n-1} Q_{n-1}(x) d\psi(x)$$

así que $A_n > 0$. Por lo tanto la relación de recurrencia de los polinomios $Q_n(x)$ determina una matriz de nacimiento y muerte con $\mu_0 = 0$. Las condiciones bajo las cuales es posible renormalizar los polinomios de tal manera que la relación de recurrencia determine una matriz de nacimiento y muerte con $\mu_0 > 0$ vienen dadas por el siguiente lema.

Lema 3.1.1. *Supongamos que $\mu_0 = 0$ y sea $\{Q_n(x)\}$ una sucesión de polinomios determinada por (3.6). Sea μ un número positivo. Entonces existe una sucesión de constantes positivas $(\alpha_n)_n$ tal que los polinomios $R_n(x) = \alpha_n Q_n(x)$ satisfacen una relación de recurrencia de la forma*

$$\begin{aligned}
R_0(x) &= 1 \\
-xR_0(x) &= -(\lambda'_0 + \mu'_0)R_0(x) + \lambda'_0 R_1(x) \\
-xR_n(x) &= \mu'_n R_{n-1}(x) - (\lambda'_n + \mu'_n)R_n(x) + \lambda'_n R_{n+1}(x) \quad n \geq 1, \mu'_n > 0, \lambda'_n > 0
\end{aligned}$$

con $\mu'_0 = \mu$ si y sólo si la serie

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n}$$

converge y μ satisface

$$\mu \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \leq 1.$$

Demostración. Sea $\{\alpha_n\}$ cualquier sucesión de constantes positivas y $R_n(x) = \alpha_n Q_n(x)$ Entonces

$$\begin{aligned}
R_0(x) &= \alpha_0, \\
-xR_0 &= -\lambda_0 R_0 + \lambda_0 \frac{\alpha_0}{\alpha_1} R_1, \\
-xR_n &= \mu_n \frac{\alpha_n}{\alpha_{n-1}} R_{n-1} - (\lambda_n + \mu_n) R_n + \lambda_n \frac{\alpha_n}{\alpha_{n+1}} R_{n+1}.
\end{aligned}$$

Esta fórmula de recurrencia es de la forma requerida si y sólo si

$$\begin{aligned}
\alpha_0 &= 1, \\
\lambda_0 \frac{\alpha_0}{\alpha_1} + \mu &= \lambda_0, \\
\lambda_n \frac{\alpha_n}{\alpha_{n+1}} + \mu_n \frac{\alpha_n}{\alpha_{n-1}} &= \lambda_n + \mu_n \quad n \geq 1.
\end{aligned}$$

Sea

$$r_0 = \frac{\mu}{\lambda_0}, \quad r_n = \frac{\mu_n}{\lambda_n}, \quad s_n = \frac{\alpha_n}{\alpha_{n-1}}$$

y

$$t_n = r_0 + r_0 r_1 + r_0 r_1 r_2 + \dots + r_0 r_1 \dots r_n = \mu \sum_{i=1}^n \frac{1}{\lambda_i \pi_i}.$$

Entonces las relaciones anteriores pueden ser escritas como

$$1 - \frac{1}{s_1} = r_0,$$

$$(3.9) \quad 1 - \frac{1}{s_{n+1}} = r_n (s_n - 1).$$

Las s_n son positivas y $s_1 = 1/(1 - r_0) > 1$, se sigue por inducción que $s_n > 1$ para toda n . Expresando s_{n+1} en términos de la r_i obtenemos

$$s_{n+1} = 1 + \frac{r_0 r_1 \dots r_n}{1 - t_n}.$$

Por lo tanto (3.9) tiene una solución con $s_n > 1$ para toda n si y sólo si $t_n < 1$ para cada n , esto es, si y sólo si

$$\mu \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} \leq 1.$$

□

Es bien sabido además que Q_n tiene n ceros simples $x_{n1} < x_{n2} < \dots < x_{nn}$ y que los límites

$$\xi_i = \lim_{n \rightarrow \infty} x_{ni}, \quad i \geq 1$$

existen y satisfacen $0 \leq \xi_i \leq \xi_{i+1} < \infty$. Además, debido a las condiciones (1.28) y (1.30) el problema de momento relacionado/correspondiente está determinado y así

$$(3.10) \quad \inf(\text{supp}(\phi)) = \xi_1 \text{ y } \phi(\{\xi\}) = \left(\sum_{k=0}^{\infty} \pi_k Q_k^2(\xi_1) \right)^{-1}$$

donde denotamos por $\text{supp}(\phi)$ al **soporte de la medida** ϕ .

Algunos resultados se siguen directamente de la representación espectral (3.2). Por ejemplo, en el caso $\mu_0 = 0$, tenemos $Q_i(0) = 1, i \geq 0$, y así la distribución condicional límite, de existir, es igual a

$$p_j = \frac{\pi_j}{\sum_{k=0}^{\infty} \pi_k} = \lim_{t \rightarrow \infty} P_{ij}(t) = \pi_j Q_j(0) Q_i(0) \phi(\{0\}) = \pi_j \phi(\{0\}).$$

La siguiente relación de simetría se sigue directamente de (3.2)

$$\pi_i P_{ij}(t) = \pi_j P_{ji}(t).$$

Decimos que $P_{ij}(t)$ es **débilmente simétrica** o **reversible**.

Asociado a un proceso de nacimiento y muerte con $\mu_0 = 0$, existe otro proceso dual de nacimiento y muerte con parámetros $\{(\lambda_n^*, \mu_n^*, n \geq 0)\}$ dado por

$$\lambda_n^* = \mu_{n+1}, \quad \mu_n^* = \lambda_n \quad n \geq 0.$$

El correspondiente operador infinitesimal sería la matriz

$$\mathcal{A}^* = \begin{pmatrix} -(\lambda_1 + \mu_1) & \mu_1 & 0 & 0 & 0 & \cdots \\ \lambda_1 & -(\lambda_1 + \mu_2) & \mu_2 & 0 & 0 & \cdots \\ 0 & \lambda_2 & -(\lambda_2 + \mu_3) & \mu_3 & 0 & \cdots \\ 0 & 0 & \lambda_3 & -(\lambda_3 + \mu_4) & \mu_4 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

cuyos coeficientes potenciales son

$$\pi_0^* = 1, \quad \pi_n^* = \frac{\mu_1 \cdots \mu_n}{\lambda_1 \cdots \lambda_n} = \frac{\lambda_0}{\lambda_n \pi_n}.$$

Los correspondientes polinomios $R_n^*(x)$ asociados a \mathcal{A}^* vienen dados por la relación

$$(3.11) \quad -xR_n^*(x) = \lambda_n \pi_n [Q_{n+1}(x) - Q_n(x)], \quad n \geq 0.$$

Existe también una relación que describe $Q_n(x)$ en términos de $R_n^*(x)$:

$$\pi_{n+1} Q_{n+1}(x) = R_{n+1}^*(x) - R_n^*(x), \quad n \geq 0.$$

La peculiaridad de estos nuevos polinomios es que la ortogonalidad de los polinomios $Q_n(x)$ con respecto a una medida ψ es equivalente a la ortogonalidad de los polinomios $R_n^*(x)$ con respecto a la medida dual ψ^* de ψ , dada por la transformación de Christoffel

$$d\psi^*(x) = \frac{x}{\lambda_0} d\psi(x), \quad x \geq 0.$$

Esta medida dual ψ^* es de probabilidad, ya que

$$0 = \int_0^\infty Q_1(x) d\psi(x) = \int_0^\infty \left(1 - \frac{x}{\lambda_0}\right) d\psi(x) = 1 - \int_0^\infty \frac{x}{\lambda_0} d\psi(x) = 1 - \int_0^\infty d\psi^*(x).$$

Obsérvese que ψ^* no tiene ningún átomo en $x = 0$ y que $\lambda_0 \int_0^\infty d\psi^*(x)/x \leq 1$. La relación con la medida ψ es la siguiente

$$d\psi(x) = \begin{cases} \lambda_0 d\psi^*(x)/x, & \text{si } x > 0 \\ 1 - \lambda_0 \int_0^\infty d\psi^*(t)/t, & \text{si } x = 0. \end{cases}$$

La matriz de transición de esta familia dual, denotada por $P_{ij}^*(t)$ se puede expresar en términos de la matriz de transición original $P_{ij}(t)$ mediante la siguiente

Proposición 3.1.1. Sean $P_{ij}(t)$ y $P_{ij}^*(t)$ las únicas soluciones mínimas de un proceso de nacimiento y muerte y su dual, respectivamente. Entonces

$$P_{ij}^*(t) = \sum_{k=0}^i [P_{jk}(t) - P_{j+1,k}(t)], \quad P_{ij}(t) = \sum_{k=0}^i [P_{jk}^*(t) - P_{j-1,k}^*(t)] \quad i, j \geq 0,$$

con la convención de que $P_{-1,k}^*(t) = 0$ para todo $k \geq 0$. Además, $P_{ij}^*(t)$ satisface las ecuaciones de retroceso de Kolmogorov.

Aparte de la familia dual, existe otra familia que se suele usar bastante, llamado el proceso de nacimiento y muerte asociado. Este proceso tiene como matriz del operador infinitesimal una matriz que se obtiene de \mathcal{A} quitándole la primera fila y primera columna. Por lo tanto será un proceso de nacimiento y muerte con un conjunto de parámetros $\{(\lambda_{n+1}, \mu_{n+1}), n \geq 0\}$. Llamamos $\psi^{(0)}$ a la medida del proceso asociado. Se puede comprobar que los polinomios definidos por

$$Q_n^{(0)} = \int_0^\infty \frac{Q_n(x) - Q_n(y)}{x - y} d\psi(x),$$

verifican la relación de recurrencia generada por el conjunto de parámetros $\{(\lambda_{n+1}, \mu_{n+1}), n \geq 0\}$ con $Q_0^{(0)} = 0$ y $Q_1^{(0)} = \frac{-1}{\lambda_0}$, y son ortogonales con respecto a $\psi^{(0)}$. Se tiene también que

$$(3.12) \quad B(z; \psi) = \frac{1}{\lambda_0 + \mu_0 - z - \lambda_0 \mu_1 B(z; \psi^{(0)})},$$

donde $B(z; \psi)$ es la transformada de Stieltjes, la cual se usa para calcular algunas medidas asociadas a ciertos procesos de nacimiento y muerte.

La función de transición $P_{ij}^{(0)}(t)$ correspondiente al proceso asociado viene también dada por una fórmula del tipo Karlin-McGregor:

$$P_{ij}^{(0)}(t) = \frac{\mu_1}{\lambda_0} \pi_j \int_0^\infty e^{-xt} [-\lambda_0 Q_i^{(0)}(x)] [-\lambda_0 Q_j^{(0)}(x)] d\psi^{(0)}, \quad i, j \geq 1.$$

Se tiene también el siguiente resultado

Proposición 3.1.2. Sean

$$m_{-n} = \lim_{\xi \rightarrow 0^+} \int_0^\infty \frac{d\psi(x)}{(x + \xi)^n}, \quad m_{-n}^a = \lim_{\xi \rightarrow 0^+} \frac{d\psi^{(0)}(x)}{(x + \xi)^n},$$

los momentos negativos asociados a las medidas ψ y $\psi^{(0)}$ respectivamente. Entonces se tiene

$$(3.13) \quad m_{-1}^a = \int_0^\infty \frac{d\psi^{(0)}(x)}{x} \leq \frac{\lambda_0 + \mu_0}{\lambda_0 \mu_1}.$$

Además, las siguientes afirmaciones son ciertas:

1. Si $m_{-1}^a = (\lambda_0 + \mu_0)/\lambda_0 \mu_1$ y $m_{-2}^a < \infty$, entonces $\psi(\{0\}) = 1/(1 + \lambda_0 \mu_1 m_{-2}^a) > 0$.
2. Si $m_{-1}^a = (\lambda_0 + \mu_0)/\lambda_0 \mu_1$ y $m_{-2}^a = \infty$, entonces $\psi(\{0\}) = 0$ y $m_{-1} = \infty$.
3. Si $m_{-1}^a < (\lambda_0 + \mu_0)/\lambda_0 \mu_1$, entonces $\psi(\{0\}) = 0$,

$$m_{-1} = \frac{1}{\lambda_0 + \mu_0 - \lambda_0 \mu_1 m_{-1}^a} < \infty, \quad \frac{m_{-2}}{m_{-1}^2} = 1 + \lambda_0 \mu_1 m_{-2}^a \leq \infty.$$

Demostración. De (3.12), para $z = -\xi$, obtenemos

$$(3.14) \quad \lambda_0 \mu_1 \int_0^\infty \frac{d\psi^{(0)}(x)}{x + \xi} = \lambda_0 + \mu_0 + \xi - \left(\int_0^\infty \frac{d\psi(x)}{x + \xi} \right)^{-1}, \quad \xi > 0.$$

Por lo tanto

$$\int_0^\infty \frac{d\psi^{(0)}(x)}{x + \xi} \leq \frac{\lambda_0 + \mu_0 + \xi}{\lambda_0 \mu_1}, \quad \xi > 0.$$

Haciendo $\xi \rightarrow 0^+$ se consigue la cota (3.13). Por otro lado, haciendo derivada con respecto a ξ en (3.14), se tiene

$$(3.15) \quad -\lambda_0 \mu_1 \int_0^\infty \frac{d\psi^{(0)}(x)}{(x + \xi)^2} = 1 - \frac{\int_0^\infty \frac{d\psi(x)}{(x + \xi)^2}}{\left(\int_0^\infty \frac{d\psi(x)}{x + \xi}\right)^2}, \quad \xi > 0.$$

Haciendo $\xi \rightarrow 0^+$ en (3.14) y (3.15) se obtienen las relaciones

$$(3.16) \quad \lambda_0 \mu_1 m_{-1}^a = \lambda_0 + \mu_0 - \frac{1}{m_{-1}},$$

$$(3.17) \quad \frac{m_{-2}}{m_{-1}^2} = 1 + \lambda_0 \mu_1 m_{-2}^a.$$

También, de (3.12) y (2.6), se llega a que

$$(3.18) \quad \psi(\{0\}) = \lim_{\xi \rightarrow 0^-} \frac{\xi}{\xi - \lambda_0 - \mu_0 + \lambda_0 \mu_1 \int_0^\infty \frac{d\psi^{(0)}(x)}{x + \xi}}$$

La primera afirmación se sigue aplicando L'Hôpital a (3.18) y es positivo por (3.17). La segunda afirmación es consecuencia de la primera afirmación y (3.16). Por último, la tercera afirmación es consecuencia directa de (3.16), (3.17) y (3.18). \square

Veremos ahora una propiedad muy importante de la representación integral, aunque para esto necesitamos el siguiente:

Lema 3.1.2. *Sea ψ una medida definida en $[0, \infty]$ con momentos finitos y sea $P_n(x)$ una familia de polinomios ortogonales con respecto a ψ . Entonces*

$$(3.19) \quad \int_0^\infty e^{-xt} P_n(x) d\psi(x) > 0, \quad t > 0, \quad n \geq 0.$$

Además, si $n \geq 2$ y $p(x)$ es un polinomio de grado r con $0 \leq r < n$ cuyas raíces son todas reales y separadas por algún subconjunto de $r + 1$ raíces de $P_n(x)$, y $p(0) > 0$, entonces

$$(3.20) \quad \int_0^\infty e^{-xt} p(x) P_n(x) d\psi(x) > 0, \quad t > 0, \quad n \geq 0.$$

En particular, para $p(x) = P_m(x)$ (donde se cumplen las condiciones anteriores) se tiene que

$$(3.21) \quad \int_0^\infty e^{-xt} P_m(x) P_n(x) d\psi(x) > 0, \quad t > 0, \quad n, m \geq 0.$$

Demostración. Sea $1 \leq r < n$. Todos los ceros de $P_n(x)$ están dentro del soporte de ortogonalidad y además son simples (véase [3]). Sean $\xi_j, j = 0, \dots, n$, esos ceros ordenados de la siguiente manera $0 < \xi_1 < \dots < \xi_n$. El polinomio $P_n(x)$ se puede escribir como

$$P_n(x) = k_n(\xi_1 - x)(\xi_2 - x) \cdots (\xi_n - x),$$

donde $k_n > 0$ (ya que $P_n(0) > 0$). Para índices $1 \leq i_1 < i_2 < \dots < i_r$ consideramos el r -factor de $P_n(x)$

$$A_{i_1, \dots, i_r}(x) = \prod_{k=1}^r (\xi_{i_k} - x).$$

Probaremos primero que

$$(3.22) \quad f_{i_1, \dots, i_r}(t) = \int_0^\infty e^{-xt} A_{i_1, \dots, i_r}(x) P_n(x) d\psi(x) > 0.$$

Para ello denotamos $1 \leq j_1 < j_2 < \dots < j_{n-r}$ al conjunto complementario de los índices de 1 a n quitándole los índices i_1, \dots, i_r . Definimos, para cualquier función continuamente diferenciable $g(t)$, el operador

$$D_j g(t) = e^{-\xi_j t} \frac{d}{dt} [e^{\xi_j t} g(t)] = g'(t) + \xi_j g(t).$$

Este operador tiene la propiedad siguiente:

$$(3.23) \quad D_j g(t) > 0, \text{ para todo } t > 0 \text{ y } g(0) = 0 \text{ implica que } g(t) > 0 \text{ para todo } t > 0.$$

Esto es debido a que si $g(0) = 0$ entonces $g'(0) > 0$, es decir es estrictamente creciente en 0. Si existiese un $w > 0$ tal que $g(w) \leq 0$, entonces $g'(w) > -g(w)\xi_j \geq 0$, con lo cual g sería estrictamente creciente en w . Esto implica que existe un cero z de g en el intervalo $(0, w)$, i.e. $g(z) = 0$ y además $g'(z) < 0$. Pero eso es contradicción ya que debe ocurrir que $g'(z) + \xi_j g(z) > 0$.

Si j es uno de los índices j_1, \dots, j_{n-r} , entonces

$$D_j f_{i_1, \dots, i_r}(t) = \int_0^\infty e^{-xt} (-x + \xi_j) A_{i_1, \dots, i_r}(x) P_n(x) d\psi(x),$$

e iterando para todos esos índices se tiene que

$$D_{j_1} D_{j_2} \cdots D_{j_{n-r}} f_{i_1, \dots, i_r}(t) = \frac{1}{k_n} \int_0^\infty e^{-xt} [P_n(x)]^2 d\psi(x) > 0.$$

Además, si $k < n - r$, por la propiedad de ortogonalidad de $P_n(x)$, se tiene que

$$D_{j_1} D_{j_2} \cdots D_{j_k} f_{i_1, \dots, i_r}(0) = 0.$$

Aplicando (3.23) $n - r$ veces llegamos a ver que se cumple (3.22).

Veamos ahora (3.20). Sea $p(x)$ con las condiciones del Teorema y elegimos i_1, \dots, i_{r+1} , tal que $\xi_{i_1}, \dots, \xi_{i_{r+1}}$ separan las r raíces de $p(x)$. Entonces, si denotamos $A_k(x) = A_{i_1, \dots, i_{r+1}}(x) / (\xi_{i_k} - x)$, se tiene que

$$\frac{p(x)}{A_{i_1, \dots, i_{r+1}}(x)} = \sum_{k=1}^{r+1} \frac{\alpha_k}{\xi_{i_k} - x}, \quad \alpha_k = \frac{p(\xi_{i_k})}{A_k(\xi_{i_k})} \text{ implica } p(x) = \sum_{k=1}^{r+1} \alpha_k A_k(x),$$

donde los coeficiente α_k son no negativos y no todos 0, ya que los signos de $p(\xi_{i_k})$ y $A_k(\xi_{i_k})$ son siempre los mismos para todo k . Con lo cual, para todo $t > 0$ y usando (3.22), se tiene que

$$\int_0^\infty e^{-xt} p(x) P_n(x) d\psi(x) = \sum_{k=1}^{r+1} \alpha_k \int_0^\infty e^{-xt} A_k(x) P_n(x) d\psi(x) > 0.$$

Veamos ahora (3.19). Para ello definamos

$$F_n(t) = \int_0^\infty e^{-xt} P_n(x) d\psi(x), \quad n \geq 0.$$

Se tiene que $F_0(t) > 0$. Como

$$\frac{d}{dt} [e^{B_0 t} F_1(t)] = e^{-B_0 t} \int_0^\infty e^{-xt} P_1(x) (-x - B_0) d\psi(x) = \frac{C_0}{P_0(x)} e^{-B_0 t} \int_0^\infty e^{-xt} [P_1(x)]^2 d\psi(x) > 0,$$

y $F_1(0) = 0$, podemos aplicar (3.23) para probar que $F_1(t) > 0$. Para $n \geq 2$ se tiene que $F_n(0) = 0$ y que

$$D_j F_n(t) = e^{-\xi_j t} \frac{d}{dt} [\xi_j t F_n(t)] = f_j(t) > 0,$$

con lo cual por la misma razón $F_n(t) > 0$ para todo $t > 0$.

Por último probamos (3.21). Para $m = n$ es trivial. Si $m \neq n$, por ejemplo $m < n$, por la propiedad de separación de los ceros de polinomios ortogonales un subconjunto de ceros de $P_n(x)$ separa a las raíces de $P_m(x)$ y además $P_m(0) > 0$, con lo que podemos aplicar (3.21) para $p(x) = P_m(x)$. □

Estamos ahora en condiciones de enunciar la siguiente:

Proposición 3.1.3. *Sea ψ una medida de probabilidad que es solución de momentos de Sieltjes para el conjunto de parámetros $\{(\lambda_n, \mu_n), n \geq 0\}$ con $\mu_0 = 0$. Entonces la fórmula*

$$(3.24) \quad P_{ij}(t) = \pi_j \int_0^\infty e^{-tx} Q_i(x) Q_j(x) d\psi(x), \quad i, j \in \mathbf{E}, \quad t \geq 0,$$

define una pseudosolución de ambas ecuaciones de retroceso y evolución de Kolmogorov con respecto al operador infinitesimal (1.26).

Demostración. Por definición, hay que probar que $P_{ij}(t)$ es siempre positiva para todo $t > 0$ y que $\sum_{j=0}^\infty P_{ij}(t) \leq 1$ para todo $i \in \mathbf{E}$ y $t \geq 0$. Por un lado, si $i = j$ entonces (3.24) es siempre positiva.

Si $i \neq j$ la positividad es consecuencia de (3.21).

Por otro lado, de (3.7) y (3.11), se tiene que

$$\sum_{j=0}^n P_{ij}(t) = \int_0^\infty e^{-xt} Q_i(x) \sum_{j=0}^n Q_j(x) \pi_j d\psi(x) = \int_0^\infty e^{-xt} Q_i(x) R_n^*(x) d\psi(x).$$

En particular, para $i = 0$, se tiene (usando (3.19)):

$$\frac{d}{dt} \sum_{j=0}^n P_{0,j}(t) = - \int_0^\infty x e^{-xt} R_n^*(x) d\psi(x) = -\lambda_0 \int_0^\infty e^{-xt} R_n^*(x) d\psi^*(x) < 0, \quad t > 0.$$

Como $\sum_{j=0}^n P_{0,j}(0) = 1$ y es decreciente en t , entonces

$$\sum_{j=0}^n P_{0,j}(t) \leq 1, \quad n \geq 0, \quad t > 0.$$

De manera más general, como

$$\sum_{j=0}^n P_{ij}(t) - \sum_{j=0}^n P_{i+1,j}(t) = \frac{\lambda_0}{\lambda_i \pi_i} \int_0^\infty e^{-xt} R_i^*(x) R_n^*(x) d\psi^*(x) > 0,$$

por recurrencia y por (3.21) se puede probar que

$$\sum_{j=0}^n P_{ij}(t) \leq 1, \quad i, n \geq 0, \quad t > 0,$$

y por lo tanto $\sum_{j=0}^\infty P_{ij}(t) \leq 1$. Por último (3.24) satisface las ecuaciones de retroceso y evolución de Kolmogorov usando la simetría (1.27) y la relación de recurrencia (3.6). \square

En cuanto a la unicidad tenemos el siguiente

Teorema 3.1.3. *Sea $\{(\lambda_n, \mu_n, n \geq 0)\}$ un conjunto de parámetros de nacimiento y muerte con coeficientes potenciales π_n . Entonces*

1. Si $\mu_0 = 0$, la solución del problema de momentos de Stieltjes es única si y sólo si

$$(3.25) \quad \sum_{n=0}^{\infty} \left(\pi_n + \frac{1}{\lambda_n \pi_n} \right) = \infty.$$

2. Si $\mu_0 > 0$, la solución del problema de momentos de Stieltjes es única si y sólo si

$$(3.26) \quad \sum_{n=0}^{\infty} \left(1 + \mu_0 \sum_{k=0}^{n-1} \frac{1}{\lambda_k \pi_k} \right)^2 = \infty.$$

3.1.2. Propiedades probabilísticas

Asumiremos a continuación que la solución del problema de Stieltjes es única, con lo cual se tienen las condiciones (3.25) cuando $\mu_0 = 0$ y (3.26) cuando $\mu_0 > 0$.

Sea $\{(\lambda_n, \mu_n, n \geq 0)\}$ un conjunto de parámetros con $\lambda_n, \mu_{n+1} > 0, n \geq 0$, y $\mu_0 \geq 0$. El proceso es entonces irreducible en todos sus estados, excepto quizá en el 0. Con lo cual estudiar cuestiones de recurrencia o absorción se reduce a estudiar el comportamiento en el estado 0. Si $\mu_0 > 0$ se interpreta como la existencia de un estado -1 al cual el proceso puede ser absorbido desde el estado 0.

Sean $F_{ij}(t)$ las distribuciones de los primeros tiempos de llegada definidas en (1.24) y (1.25). Considerando las transformadas de Fourier $\hat{P}_{ij}(\lambda)$ de $P_{ij}(t)$ y $\hat{F}_{ij}(\lambda)$ de $F_{ij}(t)$ obtenemos las relaciones

$$(3.27) \quad \hat{P}_{ii}(\lambda) = \frac{1}{\lambda + \lambda_i + \mu_i} + \hat{P}_{ii}(\lambda)\hat{F}_{ii}(\lambda),$$

$$\hat{P}_{ij}(\lambda) = \hat{P}_{jj}(\lambda)\hat{F}_{ij}(\lambda), \quad i \neq j.$$

En particular para $i = j = 0$, y usando la representación (3.4), se tiene que

$$\begin{aligned} \hat{P}_{00}(\lambda) &= \int_0^\infty e^{-\lambda t} P_{00}(t) dt = \int_0^\infty e^{-\lambda t} \left(\int_0^\infty e^{-xt} d\psi(x) \right) dt \\ &= \int_0^\infty \left(\int_0^\infty e^{-(\lambda+x)t} dt \right) d\psi(x) = \int_0^\infty \frac{d\psi(x)}{\lambda+x}. \end{aligned}$$

Con lo cual se tiene la fórmula

$$(3.28) \quad \hat{F}_{00}(\lambda) = 1 - \frac{1}{(\lambda + \lambda_0 + \mu_0) \int_0^\infty \frac{d\psi(x)}{\lambda+x}}.$$

De manera más general, usando

$$\hat{P}_{ij}(\lambda) = \pi_j \int_0^\infty \frac{Q_i(x)Q_j(x)}{\lambda+x} d\psi(x),$$

se puede llegar a las fórmulas

$$\hat{F}_{ii}(\lambda) = 1 - \frac{1}{(\lambda + \lambda_i + \mu_i) \pi_i \int_0^\infty \frac{Q_i^2(x)}{\lambda+x} d\psi(x)}, \quad \hat{F}_{ij}(\lambda) = \frac{\int_0^\infty \frac{Q_i(x)Q_j(x)}{\lambda+x} d\psi(x)}{\int_0^\infty \frac{Q_j^2(x)}{\lambda+x} d\psi(x)}, \quad i \neq j.$$

Teorema 3.1.4. *Sea $\mu_0 = 0$ (es decir el estado 0 no es absorbente). Son equivalentes:*

1. *El proceso de nacimiento y muerte es recurrente.*

2.

$$\int_0^\infty \frac{d\psi(x)}{x} = \infty.$$

3.

$$\sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n} = \infty, \quad \text{o} \quad \sum_{n=1}^{\infty} \frac{1}{\mu_n \pi_n} = \infty.$$

4.

$$\int_0^{\infty} \frac{d\psi^{(0)}(x)}{x} = 1/\mu_1 \quad \text{donde } \psi^{(0)} \text{ es la medida del proceso asociado.}$$

Y también son equivalentes

1. El proceso de nacimiento y muerte es recurrente positivo.

2. La medida ψ tiene un salto finito y positivo en $x = 0$ de magnitud $\psi(\{0\}) = (\sum_{n=0}^{\infty} \pi_n)^{-1}$.

3.

$$\sum_{n=0}^{\infty} \pi_n < \infty.$$

4.

$$\int_0^{\infty} \frac{d\psi^{(0)}(x)}{x} = 1/\mu_1 \quad \text{y} \quad \int_0^{\infty} \frac{d\psi^{(0)}(x)}{x^2} < \infty, \quad \text{donde } \psi^{(0)} \text{ es la medida del proceso asociado.}$$

Demostración. Como la cadena es irreducible, basta estudiar recurrencia para un estado, por ejemplo el 0. Por definición, el estado 0 es recurrente si y sólo si $\lim_{\lambda \rightarrow 0^+} \hat{F}_{00}(\lambda) = \int_0^{\infty} dF_{00}(t) = 1$. Por (3.28), esto va a ser posible si y sólo si $\lim_{\lambda \rightarrow 0^+} \int_0^{\infty} \frac{d\psi(x)}{\lambda+x} = \infty$, con lo cual tiene que ocurrir que $\int_0^{\infty} \frac{d\psi(x)}{x} = \infty$. Para llegar a la tercera relación hay que demostrar que

$$(3.29) \quad \int_0^{\infty} \frac{d\psi(x)}{x} = \sum_{n=0}^{\infty} \frac{1}{\lambda_n \pi_n}.$$

De manera general, si multiplicamos (3.8) (para $\mu_0 = 0$) por $Q_j(x)/x^k$ e integramos con respecto a la medida ψ , se tiene que

$$(3.30) \quad \int_0^{\infty} Q_n(x) Q_j(x) \frac{d\psi(x)}{x^k} = \int_0^{\infty} Q_j(x) \frac{d\psi(x)}{x^k} - \sum_{m=0}^{n-1} \frac{1}{\lambda_m \pi_m} \sum_{i=0}^m \pi_i \int_0^{\infty} Q_i(x) Q_j(x) \frac{d\psi(x)}{x^{k-1}}.$$

Haciendo tender $n \rightarrow \infty$ en la fórmula anterior y teniendo en cuenta que $\lim_{n \rightarrow \infty} \int_0^{\infty} Q_n(x) Q_j(x) \frac{d\psi(x)}{x^k} = 0$, se tiene que

$$(3.31) \quad \int_0^{\infty} Q_j(x) \frac{d\psi(x)}{x^k} = \sum_{m=0}^{\infty} \frac{1}{\lambda_m \pi_m} \sum_{i=0}^m \pi_i \int_0^{\infty} Q_i(x) Q_j(x) \frac{d\psi(x)}{x^{k-1}}.$$

Combinando (3.30) y (3.31) se tiene finalmente

$$(3.32) \quad \int_0^{\infty} Q_n(x) Q_j(x) \frac{d\psi(x)}{x^k} = \sum_{m=n}^{\infty} \frac{1}{\lambda_m \pi_m} \sum_{i=0}^m \pi_i \int_0^{\infty} Q_i(x) Q_j(x) \frac{d\psi(x)}{x^{k-1}}.$$

Con lo cual se pueden calcular las integrales $\int_0^\infty Q_i(x)Q_j(x)\frac{d\psi(x)}{x^k}$ en términos de $\int_0^\infty Q_i(x)Q_j(x)\frac{d\psi(x)}{x^{k-1}}$. En particular, para $n = j = 0$ y $k = 1$ y usando la ortogonalidad se llega a (3.29). La última parte es consecuencia de (3.16) (ya que $m_{-1} = \infty$ si y sólo si $m_{-1}^a = 1/\mu_1$).

Con respecto a la recurrencia positiva, por definición lo que tiene que ocurrir es que el tiempo esperado de retorno sea finito y eso es equivalente a que $\int_0^\infty t dF_{00}(t) < \infty$. Si $\sum_{n=0}^\infty \pi_n < \infty$ entonces ψ tiene una masa en $x = 0$ de magnitud $\psi(\{0\}) = (\sum_{n=0}^\infty \pi_n)^{-1}$ (véase (3.5)) y $\lim_{t \rightarrow \infty} P_{ij}(t) = \psi(\{0\})\pi_j$. El primer momento de F_{00} viene dado por $-\lim_{\lambda \rightarrow 0^+} \hat{F}'_{00}(\lambda)$. Por lo tanto, usando (3.27), se tiene que

$$-\hat{F}'_{00}(\lambda) = -\frac{\hat{P}'_{00}(\lambda)}{(\lambda + \lambda_0 + \mu_0)\hat{P}_{00}^2(\lambda)} - \frac{1}{(\lambda + \lambda_0 + \mu_0)^2\hat{P}_{00}(\lambda)}.$$

El segundo término tiende a 0 a medida que $\lambda \rightarrow 0^+$, ya que como el proceso es recurrente, entonces $\hat{P}_{00}(\lambda) \rightarrow \infty$. Por otro lado

$$-\frac{\hat{P}'_{00}(\lambda)}{\hat{P}_{00}^2(\lambda)} = \frac{\int_0^\infty \frac{d\psi(x)}{(\lambda+x)^2}}{\left(\int_0^\infty \frac{d\psi(x)}{\lambda+x}\right)^2} = \frac{\frac{\psi(\{0\})}{\lambda^2} + \int_0^\infty \frac{d\psi^+(x)}{(\lambda+x)^2}}{\left(\frac{\psi(\{0\})}{\lambda} + \int_0^\infty \frac{d\psi^+(x)}{\lambda+x}\right)^2},$$

donde ψ^+ denota la medida que se obtiene de ψ quitándole la masa en $x = 0$, con lo cual ψ^+ es continua en $x = 0$. Por lo tanto, multiplicando y dividiendo por λ^2 , se tiene que

$$\int_0^\infty t dF_{00}(t) = -\lim_{\lambda \rightarrow 0^+} \hat{F}'_{00}(\lambda) = \frac{1}{(\lambda_0 + \mu_0)\psi(\{0\})},$$

que es finito. Por otro lado, si el proceso es recurrente positivo, entonces la cadena de Markov $P_{ii}(nh)$ con h fijo es recurrente positivo. Entonces $\lim_{n \rightarrow \infty} P_{ii}(nh) = \psi(\{0\})\pi_j > 0$, y por lo tanto $\psi(\{0\}) > 0$. Como $\psi(\{0\}) = (\sum_{n=0}^\infty \pi_n)^{-1}$ entonces $\sum_{n=0}^\infty \pi_n < \infty$. La última parte es de nuevo consecuencia de la afirmación 1 de la Proposición 3.1.2. Esta proposición también da una expresión alternativa para $\psi(\{0\})$:

$$\psi(\{0\}) = \frac{1}{1 + \lambda_0\mu_1 \int_0^\infty \frac{d\psi^{(0)}(x)}{x^2}}.$$

□

Otro concepto importante es el de α -recurrencia. Para cualquier cadena de Markov a tiempo continuo X_t se define el parámetro de decaimiento de X_t como un número $\alpha \geq 0$ tal que

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log P_{ij}(t) = -\alpha.$$

Si $\alpha > 0$ entonces el proceso es transitorio, ya que $P_{ij}(t)$ se comporta como $e^{-\alpha t}$ a medida que $t \rightarrow \infty$, con lo cual $\int_0^\infty P_{ij}(t) < \infty$. Este parámetro α mide lo rápido que $P_{ij}(t)$ va a tender a cero a medida que $t \rightarrow \infty$.

Se dice entonces que un proceso de nacimiento y muerte es α -recurrente si para algún estado $i \in \mathbf{E}$ se tiene que

$$\int_0^\infty e^{\alpha t} P_{ii}(t) dt = \infty,$$

o α -transitorio en caso contrario. De la misma manera, si el proceso de nacimiento y muerte es α -recurrente, se dice que es α -positivo si para algún estado $i \in \mathbf{E}$ se tiene que

$$\lim_{t \rightarrow \infty} e^{\alpha t} P_{ii}(t) > 0,$$

o α -nulo en caso contrario. Cuando $\alpha = 0$ recuperamos las definiciones convencionales de recurrencia. Es posible probar usando la fórmula de representación de Karlin-McGregor que

$$\alpha = \text{mín } sop(\psi).$$

3.1.3. Métodos para calcular la medida ψ

El principal método para adivinar la medida espectral ψ en términos de la matriz del operador infinitesimal \mathcal{A} es a través de su transformada de Stieltjes y posteriormente aplicar el Teorema de Inversión de Perron-Stieltjes. Para ello una técnica muy común es fijarse en el k -ésimo proceso de nacimiento y muerte, consistente en un nuevo proceso cuya matriz del operador infinitesimal se obtiene de \mathcal{A} quitándole las primera k filas y columnas. La medida correspondiente a este nuevo proceso se suele denotar por $\psi^{(k)}$, $k \geq 0$. Para el caso $k = 0$ tenemos los polinomios asociados y ya se obtuvo una fórmula de relación entre transformadas de Stieltjes. De manera general, se tiene

$$(3.33) \quad B(z; \psi^{(n-1)}) = \frac{1}{\lambda_n + \mu_n - z - \lambda_n \mu_{n+1} B(z; \psi^{(n)})} = \frac{\frac{1}{\lambda_n}}{\frac{\lambda_n + \mu_n - z}{\lambda_n} - \mu_{n+1} B(z; \psi^{(n)})}, \quad n \geq 0.$$

En la anterior fórmula se entiende que $\psi = \psi^{(-1)}$. Iterando la fórmula se tiene que $B(z; \psi)$ se puede escribir com

$$(3.34) \quad B(z; \psi) = \frac{\alpha_n B(z; \psi^{(n-1)}) + \beta_n}{\gamma_n B(z; \psi^{(n-1)}) + \delta_n}, \quad n \geq 0,$$

donde $\alpha_n, \beta_n, \gamma_n, \delta_n$ son funciones de z que no necesariamente tienen que ser únicas. Trate-mos ahora de escribir $B(z; \psi)$ en términos de $B(z; \psi^{(n-1)})$ usando los polinomios $Q_n(x)$ y los asociados $Q_n^{(0)}(x)$. Para $n = 0$, se tiene que $\alpha_0 = \delta_0 = 1, \beta_0 = \gamma_0 = 0$, mientras que para $n = 1$, una solución sería $\alpha_1 = 0, \beta_1 = -1/\lambda_0 = Q_1^{(0)}(z), \gamma_1 = \mu_1, \delta_1 = -(\lambda_0 + \mu_0 - z)/\lambda_0 = -Q_1(z)$. Por lo tanto

$$B(z; \psi) = \frac{Q_1^{(0)}}{\mu_1 B(z; \psi^{(0)}) - Q_1(z)}.$$

Sustituyendo (3.33) en (3.34) y comparando los coeficientes en la fórmula (3.34) para $n = n + 1$, se tienen las siguientes relaciones

$$\begin{aligned} \alpha_{n+1} &= -\mu_{n+1} \beta_n, & \lambda_n \beta_{n+1} &= \alpha_n + (\lambda_n + \mu_n - z) \beta_n, \\ \gamma_{n+1} &= -\mu_{n+1} \delta_n, & \lambda_n \delta_{n+1} &= \gamma_n + (\lambda_n + \mu_n - z) \delta_n. \end{aligned}$$

Sustituyendo la primera en la segunda se tiene que β_n verifica una relación de recurrencia a tres términos

$$\lambda_n \beta_{n+1} = -\mu_n \beta_{n-1} + (\lambda + \mu_n - z) \beta_n, \beta_0 = 0, \quad \beta_1 = -1/\lambda_0 = Q_1^{(0)}(z).$$

La solución para β_n es precisamente los polinomios asociados $Q_n^{(0)}(x)$. Con lo cual $\beta_n = Q_n^{(0)}(z)$. Por lo tanto $\alpha_n = -\mu_n Q_{n-1}^{(0)}(z)$. De igual manera se puede resolver la tercera y cuarta ecuación teniendo en cuenta las condiciones iniciales, resultando $\gamma_n = \mu_n Q_{n-1}(z)$ y $\delta_n = -Q_n(z)$. En consecuencia (3.33) se puede reescribir como

$$(3.35) \quad B(z; \psi) = \frac{-\mu_n Q_{n-1}^{(0)}(z) B(z; \psi^{(n-1)}) + Q_n^{(0)}(z)}{\mu_n Q_{n-1}(z) B(z; \psi^{(n-1)}) - Q_n(z)}, \quad n \geq 0.$$

De esta manera se puede iterar esta fórmula hasta encontrar la transformada de Stieltjes de la medida original.

Como ejemplos de lo anterior tenemos:

1. La cola $M/M/\infty$
2. El proceso lineal de nacimiento y muerte
3. La cola $M/M/K$
4. Modelo cuadrático

A continuación los desarrollamos:

1. **Cola $M/M/\infty$**

La cola $M/M/\infty$ es un proceso de nacimiento y muerte con parámetros de nacimiento y muerte dados por

$$\lambda_n = \lambda, \quad \mu_n = n\mu, \quad n \geq 0,$$

donde $\lambda, \mu > 0$ son las tasas de llegada y de servicio respectivamente. Tenemos

$$\pi_j = \frac{(\lambda/\mu)^j}{j!}, \quad j \geq 0,$$

y los polinomios de nacimiento y muerte satisfacen

$$0 = n\mu Q_{n-1}(x) + (x - \lambda - n\mu) Q_n(x) + \lambda Q_{n+1}(x)$$

donde $Q_0(x) = 1$ y $Q_{-1}(x) = 0$. Si dividimos por μ , obtenemos

$$0 = nQ_{n-1}(x) + \left(\frac{x}{\mu} - \frac{\lambda}{\mu} - n \right) Q_n(x) + \frac{\lambda}{\mu} Q_{n+1}(x), \quad n \geq 0.$$

Comparando esto con las ecuaciones de los polinomios de Charlier $C_n(x; a)$ vistas en el capítulo 3 (2.69),

$$0 = nC_{n-1}(x; a) + (x - a - n)C_n(x; a) + aC_{n+1}(x; a), \quad n \geq 0,$$

donde $C_0(x; a) = 1$ y definiendo $C_{-1}(x; a) = 0$ vemos que

$$Q_n(x) = C_n \left(\frac{x}{\mu}; \frac{\lambda}{\mu} \right), \quad n \geq 0.$$

Los polinomios de Charlier satisfacen una relación de función generatriz

$$\sum_{i=0}^{\infty} C_i(x; a) \frac{z^i}{i!} = e^z \left(1 - \frac{z}{a}\right)^x$$

y son ortogonales con respecto a la distribución Poisson $P(a)$ (1.2) con soporte en los enteros no negativos y dados por

$$(3.36) \quad w_n = \frac{a^n}{n!} e^{-a}, \quad n = 0, 1, 2, \dots$$

Pero esto significa que los $Q_i(x), i \geq 0$ son ortogonales con respecto a la medida de probabilidad que asigna la masa w_n dada en (3.36) a los puntos $n\mu, n \geq 0$, dando como resultado

$$\begin{aligned} P_{ij}(t) &= \pi_j \sum_{n=0}^{\infty} e^{-t\mu n} Q_i(\mu n) Q_j(\mu n) w_n \\ &= \pi_j \sum_{n=0}^{\infty} e^{-t\mu n} C_i\left(n; \frac{\lambda}{\mu}\right) C_j\left(n; \frac{\lambda}{\mu}\right) \frac{(\lambda/\mu)^n}{n!} e^{-(\lambda/\mu)}. \end{aligned}$$

Usando la función generadora y la relación de dualidad

$$\begin{aligned} \sum_{j=0}^{\infty} P_{ij}(t) z^j &= \sum_{n=0}^{\infty} e^{-t\mu n} C_n(i; \lambda/\mu) \frac{(\lambda/\mu)^n}{n!} e^{-\lambda/\mu} \sum_{j=0}^{\infty} \pi_j C_j(n; \lambda/\mu) z^j \\ &= (1 - e^{\mu t} + z e^{-\mu t})^i \exp((\lambda/\mu)(1 - e^{-\mu t})(z - 1)). \end{aligned}$$

La distribución estacionaria está dada por

$$p_i = \frac{\pi_i}{\sum_{k=0}^{\infty} \pi_k} = e^{-(\lambda/\mu)} \frac{(\lambda/\mu)^i}{i!}$$

que no es mas que distribución Poisson $P(\lambda/\mu)$ (1.2).

Este proceso de nacimiento y muerte con $\mu = 1$ tiene un papel importante en la teoría de Stein para las distribuciones Poisson y es llamado algunas veces el proceso de muerte inmigración.

2. El proceso lineal de nacimiento y muerte

Consideramos en este ejemplo procesos de nacimiento y muerte con parámetros lineales de la forma

$$(3.37) \quad \lambda_n = (n + \beta)\lambda, \quad \text{y } \mu_n = n\mu, \quad \text{donde } \lambda, \mu, \beta > 0.$$

Haciendo cuentas obtenemos

$$\pi_j = \frac{(\beta)_j}{j!} \left(\frac{\lambda}{\mu}\right)^j, \quad j = 0, 1, 2, \dots$$

Para continuar, tenemos que considerar tres casos por separado

a) $\lambda < \mu$.

Haciendo cuentas, se puede ver que los polinomios de nacimiento y muerte están dados en términos de los polinomios de Meixner

$$Q_i(x) = M_i\left(\frac{x}{\mu - \lambda}; \beta, \frac{\lambda}{\mu}\right), \quad n = 0, 1, 2, \dots$$

Pero esto significa que los $Q_i(x), i \geq 0$ son ortogonales con respecto a la distribución de probabilidad que asigna masa

$$w_n = \left(1 - \frac{\lambda}{\mu}\right)^\beta \frac{(\beta)_n}{n!} \left(\frac{\lambda}{\mu}\right)^n, \quad n = 0, 1, 2, \dots$$

en los puntos $(\mu - \lambda)n, n = 0, 1, 2, \dots$

b) $\lambda > \mu$.

En este caso los polinomios de nacimiento y muerte están dados de nuevo en términos de los polinomios de Meixner

$$Q_i(x) = \left(\frac{\mu}{\lambda}\right)^i M_i\left(\frac{x}{\lambda - \mu} - \beta; \beta, \frac{\mu}{\lambda}\right), \quad i = 0, 1, 2, \dots$$

y así los $Q_i(x), i \geq 0$ son ortogonales con respecto a la distribución de probabilidad que asigna masa

$$w_n = \left(1 - \frac{\mu}{\lambda}\right)^\beta \frac{(\beta)_n}{n!} \left(\frac{\mu}{\lambda}\right)^n, \quad n = 0, 1, 2, \dots$$

en los puntos $(n + \beta)(\lambda - \mu), n = 0, 1, 2, \dots$

c) $\lambda = \mu$.

Después de algunos cálculos, se puede ver que los polinomios de muerte y nacimiento están dados en términos de los polinomios de Laguerre

$$Q_i(x) = \frac{i!}{(\beta)_i} L_i^{(\beta-1)}(x/\lambda), \quad n \geq 0.$$

Por lo tanto las $Q_i(x), i \geq 0$ son ortogonales en $[0, \infty)$ con respecto a la función de densidad de la distribución Gamma $G(\beta, \lambda)$

$$f(x) = \frac{1}{\lambda^\beta \Gamma(\beta)} x^{\beta-1} e^{-x/\lambda}.$$

3. La cola M/M/K

En este caso tenemos una cola con un número finito de servidores k . Los coeficientes de nacimiento y muerte son los mismos que para la cola $M/M/\infty$, pero limitando el número de servidores, es decir

$$\lambda_n = \lambda, \quad \mu_n = \begin{cases} n\mu, & \text{si } n \leq k \\ k\mu, & \text{si } n > k, \end{cases} \quad \lambda, \mu > 0.$$

Los primeros k polinomios son los mismos que para la cola $M/M/\infty$, es decir, polinomios de Charlier. Entonces

$$Q_n(x) = C_n \left(\frac{x}{\mu}; \frac{\lambda}{\mu} \right), \quad n \leq k.$$

Trataremos de obtener una expresión explícita de la medida ψ . Para ello usamos (3.35) para $n = k$, de tal manera que

$$B(z; \psi) = \frac{-k\mu Q_{k-1}^{(0)}(z)B(z; \psi^{(k-1)}) + Q_k^{(0)}(z)}{k\mu Q_{k-1}(z)B(z; \psi^{(k-1)}) - Q_k(z)}.$$

Pero a partir de quitarle las primeras k filas y columnas, la matriz del operador infinitesimal \mathcal{A} es la misma de ahí en adelante, es decir, $B(z; \psi^{(k-1)}) = B(z; \psi^{(k)})$. Por lo tanto, resolviendo en (3.33) se tiene que

$$B(z; \psi^{(k-1)}) = \frac{\lambda + k\mu - z - \sqrt{(\lambda + k\mu - z)^2 - 4k\lambda\mu}}{2k\lambda\mu},$$

donde la raíz cuadrada se toma positiva para $z < 0$. Sustituyendo esta expresión en la ecuación anterior y racionalizando, con el uso de la fórmula $\lambda_{k-1}\pi_{k-1}[Q_k Q_{k-1}^{(0)} - Q_k^{(0)} Q_{k-1}] = 1$, donde $\lambda_k \pi_k = \lambda(\lambda/\mu)^k/k!$, así como la relación de recurrencia a tres términos, se llega a

$$(3.38) \quad B(z; \psi) = -\frac{L_k(z)}{H_k(z)},$$

donde

$$\begin{aligned} L_k(z) &= 4\lambda^2 Q_k(z)Q_k^{(0)}(z) + 4k\lambda\mu Q_{k-1}(z)Q_{k-1}^{(0)}(z) \\ &\quad - 2\lambda(\lambda + k\mu - z)[Q_k(z)Q_{k-1}^{(0)}(z) + Q_{k-1}(z)Q_k^{(0)}(z)] \\ &= 4\lambda^2 \left[Q_k(z)Q_k^{(0)}(z) - \frac{1}{2} \left(Q_{k+1}(z)Q_{k-1}^{(0)}(z) + Q_{k-1}(z)Q_{k+1}^{(0)}(z) \right) \right] \\ &\quad + 2(k-1)! \left(\frac{\mu}{\lambda} \right)^{k-1} \sqrt{(\lambda + k\mu - z)^2 - 4k\lambda\mu} \end{aligned}$$

y

$$\begin{aligned} H_k(z) &= 4\lambda^2 [Q_k^2(z) - Q_{k-1}(z)Q_{k+1}(z)] \\ &= 4k\lambda\mu [Q_{k-1}^2(z) - Q_{k-2}(z)Q_k(z)] - 4\lambda\mu Q_k(z)[Q_{k-1}(z) - Q_{k-2}(z)]. \end{aligned}$$

Se observa que $H_k(z)$ es un polinomio en z de grado exactamente igual a $2k - 1$ y la parte polinomial de $L_k(z)$ es de grado $2k - 2$.

La fórmula de inversión de Perron-Stieltjes muestra que $\mathcal{T}B(x + i\eta; \psi)$ converge uniformemente a 0 a medida que $\eta \rightarrow 0^+$, siempre que x está en un intervalo que no

contenga a ninguna raíz de $H_k(x)$ y disjuncto al intervalo $|\lambda + k\mu - x| \leq 2\sqrt{k\lambda\mu}$. Por lo tanto, la densidad en su parte continua viene dada por

$$(3.39) \quad \psi'(x) = \frac{(k-1)!}{2\pi\lambda^2} \left(\frac{\mu}{\lambda}\right)^{k-1} \frac{\sqrt{4k\lambda\mu - (\lambda + k\mu - x)^2}}{Q_k^2(x) - Q_{k-1}(x)Q_{k+1}(x)}.$$

Adicionalmente la medida ψ puede tener masas discretas en alguno o todos los ceros de $H_k(x)$.

Determinemos ahora los polinomios $Q_n(x)$ para $n \geq k$. Llamemos $R_n = Q_{n+k}(x)$, $n \geq -1$. Entonces se tiene que

$$(3.40) \quad R_0(x) = Q_k(x), \quad R_{-1}(x) = Q_{k-1}(x)$$

$$-xR_n(x) = k\mu R_{n-1} - (\lambda + k\mu)R_n(x) + \lambda R_{n+1}(x), \quad n \geq 0,$$

donde se observa ahora que los coeficientes de la relación de recurrencia son independientes de n .

Esto indica que los polinomios $R_n(x)$ se podrían expresar como combinación lineal de polinomios de Chebychev $U_n(x), T_n(x)$ vistos en (2.53) y (2.52), respectivamente, que satisfacen

$$T_0(x) = 1, \quad T_{-1}(x) = x, \quad U_0(x) = 1, \quad U_{-1}(x) = 0,$$

$$xP_n(x) = \frac{1}{2}P_{n-1} + \frac{1}{2}P_{n+1}(x), \quad n \geq 1, \quad P_n(x) \in \{T_n(x), U_n(x)\}.$$

En efecto, si llamamos

$$V_n(x) = \left(\frac{k\mu}{\lambda}\right)^{n/2} T_n\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right), \quad W_n(x) = \left(\frac{k\mu}{\lambda}\right)^{n/2} U_n\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right)$$

observamos que $V_n(x)$ y $W_n(x)$ son soluciones de (3.39) donde

$$V_0(x) = 1, \quad V_{-1}(x) = \frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}, \quad W_0(x) = 1, \quad W_{-1}(x) = 0.$$

Por lo tanto $R_n(x) = v(x)V_n(x) + w(x)W_n(x)$ y resolviendo para $n = 0, -1$, se tiene que

$$\begin{aligned} R_n(x) &= \frac{2k\mu}{\lambda + k\mu - x} Q_{k-1}(x)V_n(x) + \left[Q_k(x) - \frac{2k\mu}{\lambda + k\mu - x} Q_{k-1}(x) \right] W_n(x) \\ &= Q_k(x)W_n(x) - \frac{k\mu}{\lambda} Q_{k-1}(x)W_{n-1}(x), \quad n \geq 0. \end{aligned}$$

La última igualdad es consecuencia de la conocida fórmula $T_n(x) - U - n(x) = xU_{n-1}(x)$ aplicada a $V_n(x)$ y $W_n(x)$. Por lo tanto

$$Q_{n+k}(x) = \left(\frac{k\mu}{\lambda}\right)^{n/2} \left[Q_k(x)U_n\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right) - \sqrt{\frac{k\mu}{\lambda}} Q_{k-1}(x)U_{n-1}\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right) \right], \quad n \geq 0.$$

De una manera similar

$$Q_{n+k}^{(0)}(x) = \left(\frac{k\mu}{\lambda}\right)^{n/2} \left[Q_k^{(0)}(x)U_n\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right) - \sqrt{\frac{k\mu}{\lambda}} Q_{k-1}^{(0)}(x)U_{n-1}\left(\frac{\lambda + k\mu - x}{2\sqrt{k\lambda\mu}}\right) \right], \quad n \geq 0.$$

Estudiemos ahora el caso particular de $k = 2$ servidores. Las expresiones explícitas de $Q_n, Q_n^{(0)}$ para $n = 0, 1, 2$, vienen dadas por

$$Q_0(x) = 1, \quad Q_1(x) = \frac{\lambda - x}{\lambda}, \quad Q_2(x) = \frac{(\lambda + \mu - x)(\lambda - x)}{\lambda^2} - \frac{\mu}{\lambda},$$

$$Q_0^{(0)}(x) = 0, \quad Q_1^{(0)}(x) = -\frac{1}{\lambda}, \quad Q_2^{(0)}(x) = -\frac{\lambda + \mu - x}{\lambda^2}.$$

Gracias a esto podemos calcular directamente las funciones $L_2(z)$ y $H_2(z)$ en (3.38), que vienen dadas por

$$L_2(z) = \frac{2\mu}{\lambda^2} \left[(\lambda - z)(2\mu - \lambda - 2z) + \lambda\sqrt{(\lambda + 2\mu - z)^2 - 8\lambda\mu} \right],$$

$$H_2(z) = \frac{4\mu z}{\lambda^2} [z^2 - (2\lambda + \mu)z + \lambda(\lambda + 2\mu)].$$

Por lo tanto

$$B(z; \psi) = -\frac{(\lambda - z)(2\mu - \lambda - 2z) + \lambda\sqrt{(\lambda + 2\mu - z)^2 - 8\lambda\mu}}{2z[z^2 - (2\lambda + \mu)z + \lambda(\lambda + 2\mu)]}.$$

Con lo cual, siguiendo (3.40), la parte continua de la densidad ψ viene dada por

$$\psi'(x) = \frac{\lambda\sqrt{8\lambda\mu - (\lambda + 2\mu - x)^2}}{2\pi x[x^2 - (2\lambda + \mu)x + \lambda(\lambda + 2\mu)]}, \quad |\lambda + 2\mu - x| \leq \sqrt{8\lambda\mu}.$$

Obsérvese que el soporte de esta parte continua viene dado por $[(\sqrt{\lambda} - \sqrt{2\mu})^2, (\sqrt{\lambda} + \sqrt{2\mu})^2]$. La medida puede además tener saltos localizados en los ceros del denominador de la expresión anterior. Estos ceros vienen dados por

$$s_0 = 0, \quad s_1 = \frac{2\lambda + \mu}{2} - \frac{1}{2}\sqrt{\mu(\mu - 4\lambda)}, \quad s_2 = \frac{2\lambda + \mu}{2} + \frac{1}{2}\sqrt{\mu(\mu - 4\lambda)}.$$

La magnitud del salto en $x = s_0 = 0$ viene dado por

$$\lim_{z \rightarrow 0} -zB(z; \psi) = \frac{(2\mu - \lambda) + \sqrt{(2\mu - \lambda)}}{2(2\mu + \lambda)} = \begin{cases} 0, & \text{si } 2\mu \leq \lambda, \\ \frac{2\mu - \lambda}{2\mu + \lambda}, & \text{si } 2\mu > \lambda \end{cases}.$$

Los otros dos ceros s_1, s_2 son reales siempre y cuando $\mu \geq 4\lambda$. Como $B(z; \psi)$ no puede tener polos que no sean reales (por la definición de la transformada), podemos asumir que $\mu \geq 4\lambda$. Cuando $\mu = 4\lambda$ hay un cero doble en $z = 3\lambda$. Sin embargo en este caso es fácil ver que el numerador de $B(z; \psi)$ también tiene un cero doble en $z = 3\lambda$, con lo cual no habría salto.

Si $\mu > 4\lambda$, la magnitud del salto en $x = s_2$ viene dado por

$$\begin{aligned} \lim_{z \rightarrow s_2} (s_2 - z)B(z; \psi) &= - \lim_{z \rightarrow s_2} \frac{(\lambda - z)(2\mu - \lambda - 2z) + \lambda\sqrt{(\lambda + 2\mu - z)^2 - 8\lambda\mu}}{2z(s_1 - z)} \\ &= - \frac{(\lambda - s_2)(2\mu - \lambda - 2s_2) + \lambda\sqrt{(\lambda + 2\mu - s_2)^2 - 8\lambda\mu}}{2s_2(s_1 - s_2)} \\ &= \frac{\lambda}{2} \frac{[\mu - 3\sqrt{\mu(\mu - 4\lambda)} - |\mu - 3\sqrt{\mu(\mu - 4\lambda)}|]}{\sqrt{\mu(\mu - 4\lambda)} [2\lambda + \mu + \sqrt{\mu(\mu - 4\lambda)}]} \\ &= \begin{cases} 0, & \text{si } \mu \geq 3\sqrt{\mu(\mu - 4\lambda)} \\ \frac{\lambda [3\sqrt{\mu(\mu - 4\lambda)} - \mu]}{\sqrt{\mu(\mu - 4\lambda)} [2\lambda + \mu + \sqrt{\mu(\mu - 4\lambda)}]} & \text{si } \mu < 3\sqrt{\mu(\mu - 4\lambda)}. \end{cases} \end{aligned}$$

La condición $\mu < 3\sqrt{\mu(\mu - 4\lambda)}$ es equivalente a $\mu/\lambda > 9/2$. Un cálculo similar demuestra que nunca hay salto en el punto $x = s_1$ debido a que no cumple la condición de que $\mu > 4\lambda$. Si $\mu > 4\lambda$ se tiene que $s_2 < (\sqrt{\lambda} - \sqrt{2\mu})^2$ (es decir fuera del soporte de la parte continua de la medida), y se obtiene igualdad precisamente para $\mu = 4\lambda$, en cuyo caso no hay salto, como ya hemos visto.

4. Modelo de urnas de Laplace Bernoulli (Modelo cuadrático)

Para el caso de un espacio de estados finito $\{0, 1, \dots, N\}$, los polinomios se generan mediante las relaciones

$$\begin{aligned} xQ_n(x) &= \lambda_n Q_{n+1}(x) - (\lambda_n + \mu_n)Q_n(x) + \mu_n Q_{n-1}(x), \quad n = 0, \dots, N-1, \\ xQ_N(x) &= \mu_N Q_{N-1}(x) - \mu_N Q_N(x), \\ Q_0(x) &= 1, \quad Q_{-1}(x) = 0. \end{aligned}$$

Analicemos ahora el proceso de nacimiento y muerte X_t con tasas de nacimiento y muerte cuadráticas

$$\lambda_n = (N - n)(a - n) \quad \text{y} \quad \mu_n = n(b - (N - n)), \quad n = 0, 1, 2, \dots, N$$

respectivamente, con $a, b \geq N$. Así la matriz del operador infinitesimal viene dada por

$$\mathcal{A} = \begin{pmatrix} -Na & Na & & & & \\ b - N + 1 & -\lambda_1 - \mu_1 & (N-1)(a-1) & & & \\ & \ddots & \ddots & \ddots & & \\ & & (N-1)(b+1) & -\lambda_{N-1} - \mu_{N-1} & a - N + 1 & \\ & & & Nb & -Nb & \end{pmatrix}$$

Se puede mostrar que este proceso tiene como distribución estacionaria la distribución hipergeométrica $HypII(a, b, N)$. De hecho la distribución estacionaria está dada por

$$p_i = \lim_{t \rightarrow \infty} \mathbb{P}(X_t = i | X_0 = k) = \frac{\pi_i}{\sum_{j=0}^N \pi_j} = \frac{\binom{a}{i} \binom{b}{N-i}}{\binom{a+b}{N}}.$$

Usando la representación espectral de Karlin y McGregor (3.2), podemos escribir las probabilidades de transición de este proceso X_t como sigue

$$\mathbb{P}(X_t = j | X_0 = i) = \pi_j \sum_{x=0}^N e^{\lambda(x)t} R_i(\lambda(x); a, b, N) R_j(\lambda(x); a, b, N) \rho_x$$

donde

$$\rho_x = \frac{\binom{N-b-1}{N} N! (-N)_x (-a)_x (2x - a - b - 1)}{(-1)^x x! (-b)_x (x - a - b - 1)_{N+1}}$$

y

$$\lambda(x) = x(x - a - b - 1)$$

y los R_i son los polinomios duales de Hahn vistos en (2.76) y definidos por

$$R_i(\lambda(x)) = R_i(\lambda(x); a, b, N) = {}_3\hat{F}_2(-i, -x, x - a - b - 1; -a, -N; 1)$$

para $i = 0, 1, \dots, N$.

3.1.4. Distribución límite condicional para procesos de nacimiento y muerte

Caso Absorbente

Si $\mu_0 > 0$ entonces -1 es un estado absorbente. Podemos entonces considerar la distribución límite condicional $\lim_{t \rightarrow \infty} r_{ij}(t)$ con

$$(3.41) \quad r_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i, t < T < \infty)$$

donde T denota el tiempo de absorción en el estado -1 . Esta es la probabilidad límite de estar en el estado j al tiempo t dado que el proceso no es absorbido al tiempo t , pero esa absorción en el estado -1 ocurre eventualmente.

Distribución límite condicional doble

La distribución invariante doble $\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} r_{ij}(t, s)$ con

$$(3.42) \quad r_{ij}(t, s) = \mathbb{P}(X_t = j | X_0 = i, t + s < T < \infty)$$

y donde T denota de nuevo el tiempo de absorción en el estado -1 , es también de interés. Esta es la probabilidad límite de estar en el estado j dado que el proceso no dejara $\{0, 1, 2, 3, \dots\}$ en el futuro distante pero que el estado de absorción en el estado -1 ocurre eventualmente. Recordando la interpretación de ξ_1 en (3.10) tenemos el siguiente resultado.

Teorema 3.1.5. *Supongamos que las condiciones (1.28) y (1.30) se preservan. La distribución límite condicional (3.41) y la distribución límite condicional doble en (3.42) están dadas por*

$$\lim_{t \rightarrow \infty} r_{ij}(t) = \frac{\pi_j Q_j(\xi_1)}{\sum_{k=0}^{\infty} \pi_k Q_k(\xi_1)}, \quad \lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} r_{ij}(t, s) = \frac{\pi_j Q_j^2(\xi_1)}{\sum_{k=0}^{\infty} \pi_k Q_k^2(\xi_1)}$$

respectivamente, con la interpretación de que el límite es 0 siempre que la suma en el denominador diverge.

Demostración. Para la primera distribución límite condicional, hacemos referencia a [13]. Para la segunda tenemos

$$\begin{aligned} r_{ij}(t, s) &= \mathbb{P}(X_t = j | X_0 = i, t + s < T < \infty) \\ &= \frac{\mathbb{P}(X_t = j, t + s < T < \infty | X_0 = i)}{\mathbb{P}(t + s < T < \infty | X_0 = i)} \\ &= P_{ij}(t) \frac{\mathbb{P}(t + s < T < \infty | X_t = j)}{\mathbb{P}(t + s < T < \infty | X_0 = i)} \\ &= P_{ij}(t) \frac{\mathbb{P}(s < T < \infty | X_0 = j)}{\mathbb{P}(t + s < T < \infty | X_0 = i)} \\ &= P_{ij}(t) = \frac{\int_0^{\infty} e^{-sx} Q_j(x) x^{-1} d\phi(x)}{\int_0^{\infty} e^{-(s+t)x} Q_i(x) x^{-1} d\phi(x)}. \end{aligned}$$

La última igualdad se sigue de la ecuación (3.7) presentada en [13]. Nótese que como $\int_0^{\infty} x^{-1} d\phi(x) < \infty$ ambas integrales en la última ecuación convergen. Karlin y McGregor prueban que para cualquier función continua f

$$\lim_{t \rightarrow \infty} \frac{\int_0^{\infty} e^{-xt} f(x) d\phi(x)}{\int_0^{\infty} e^{-xt} d\phi(x)} = f(\xi_1).$$

Así que tenemos

$$\lim_{s \rightarrow \infty} r_{ij}(t, s) = P_{ij}(t) \frac{Q_j(\xi_1)}{e^{-\xi_1 t} Q_i(\xi_1)}$$

$$\begin{aligned}
&= \frac{Q_j(\xi_1)}{Q_i(\xi_1)} \pi_j \int_0^\infty e^{-(x-\xi_1)t} Q_i(x) Q_j(x) d\phi(x) \\
&= \frac{Q_j(\xi_1)}{Q_i(\xi_1)} \pi_j \left[Q_i(\xi_1) Q_j(\xi_1) \phi(\{\xi_1\}) + \int_{(\xi_1, \infty)} e^{-(x-\xi_1)t} Q_i(x) Q_j(x) d\phi(x) \right].
\end{aligned}$$

Y usando el Teorema de la Convergencia Dominada obtenemos

$$\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} r_{ij}(t, s) = \phi(\{\xi_1\}) \pi_j Q_j^2(\xi_1).$$

Usando (3.10) el teorema se sigue. \square

Ilustramos ahora estos resultados calculando la distribución límite condicional doble para el ejemplo del proceso de nacimiento y muerte lineal con tasas

$$\lambda_n = \lambda(n+1) \quad y \quad \mu_n = \mu(n+1), \quad n = 0, 1, 2, \dots$$

Primero hagamos $\lambda < \mu$, con lo que la absorción tiene que pasar, y estamos calculando

$$\lim_{t \rightarrow \infty} \mathbb{P}(X_t = j | X_0 = i, t < T)$$

y

$$\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} \mathbb{P}(X_t = j | X_0 = i, t + s < T).$$

Después de algunos cálculos se puede ver que los polinomios pueden ser expresados en términos de los polinomios de Meixner y que son ortogonales con respecto a la medida que tiene sus puntos de masa en los puntos $(\mu - \lambda)(n+1)$, $n = 0, 1, 2, \dots$. Así que tenemos $\xi_1 = \mu - \lambda$. Se puede calcular por inducción que

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \frac{1}{n+1} \quad y \quad Q_n(\xi_1) = (n+1).$$

Y haciendo cálculos de nuevo

$$\lim_{t \rightarrow \infty} r_{ij}(t) = \left(\frac{\lambda}{\mu}\right)^j \left(1 - \frac{\lambda}{\mu}\right)^{-1}$$

y

$$\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} r_{ij}(t, s) = \left(1 - \frac{\lambda}{\mu}\right)^2 (j+1) \left(\frac{\lambda}{\mu}\right)^j.$$

Podemos ver que la distribución límite condicional es la distribución geométrica $Geo(\lambda/\mu)$ y que la distribución límite condicional doble se distribuye Pascal $Pa(2, \lambda/\mu)$.

Un razonamiento análogo para $\lambda > \mu$ nos lleva a

$$\lim_{t \rightarrow \infty} r_{ij}(t) = \left(1 - \frac{\mu}{\lambda}\right)^{-1} \left(\frac{\mu}{\lambda}\right)^j$$

y

$$\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} r_{ij}(t, s) = \left(1 - \frac{\mu}{\lambda}\right)^2 (j+1) \left(\frac{\mu}{\lambda}\right)^j.$$

Aquí usamos que $\xi_1 = \lambda - \mu$ y $Q_n(\lambda - \mu) = (n+1)(\mu/\lambda)^n$.

Caso reflejado

Distribución Invariante

Si $\mu_0 = 0$ entonces ignoramos el estado -1 y el estado 0 es un estado reflejado. La distribución límite estacionaria del proceso, si existe ($\sum_k \pi_k < \infty$), está dada por

$$p_j = \lim_{t \rightarrow \infty} \mathbb{P}(X_t = j | X_0 = i) = \frac{\pi_j}{\sum_{k=0}^{\infty} \pi_k} = \pi_j \phi(\{0\}).$$

Así que la distribución invariante existe siempre que ϕ tiene masa positiva en 0 .

Distribución límite condicional

Si $\sum_k \pi_k = \infty$, no tenemos una distribución límite estacionaria. Podemos entonces considerar la distribución límite condicional $\lim_{t \rightarrow \infty} \hat{r}_{ij}(t)$ con

$$(3.43) \quad \bar{r}_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i, t < S)$$

y donde S denota el último tiempo de salida del estado 0 . Está es la probabilidad límite de estar en el estado j dado que el proceso no ha comenzado el “drift” al infinito.

Distribución límite condicional doble

La distribución límite condicional doble $\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} \hat{r}_{ij}(t, s)$ con

$$(3.44) \quad \hat{r}_{ij}(t, s) = \mathbb{P}(X_t = j | X_0 = i, t + s < S)$$

y donde S denota el último tiempo de salida del estado 0 , puede ser también de interés.

Ésta es la probabilidad límite de estar en el estado j dado que el proceso no “drift” al infinito en el futuro distante. La siguiente función juega un papel importante para determinar la distribución límite condicional. Definimos

$$G_i(t) = \mathbb{P}(t < S | X_0 = i), \quad i \geq 0, t \geq 0.$$

En ecuación (3.7) de [13] se muestra que

$$G_i(t) = \frac{\int_0^{\infty} e^{-xt} x^{-1} Q_i(x) d\phi(x)}{\int_0^{\infty} x^{-1} d\phi(x)} < \infty$$

dado que el proceso es **transitorio**. Necesitamos las constantes

$$a_i = G_i(0) = \mathbb{P}(0 < S | X_0 = i), \quad i = 0, 1, \dots$$

Para el caso reflejado tenemos

Teorema 3.1.6. *Supongamos que $\sum_k \pi_k = \infty$, (1.28) y (1.30). La distribución límite condicional de (3.43) y la distribución límite condicional doble de (3.44) están dadas por*

$$\lim_{t \rightarrow \infty} \hat{r}_{ij}(t) = \frac{a_j \pi_j Q_j(\xi_1)}{\sum_{k=0}^{\infty} a_k \pi_k Q_k(\xi_1)}$$

y

$$\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} \hat{r}_{ij}(t, s) = \frac{\pi_j Q_j^2(\xi_1)}{\sum_{k=0}^{\infty} a_k \pi_k Q_k^2(\xi_1)}$$

con la interpretación de que el límite es 0 siempre que la suma en el denominador diverge.

Demostración. La demostración para la distribución límite condicional se puede ver una vez más en [13].

Ahora calcularemos $\lim_{t \rightarrow \infty} \lim_{s \rightarrow \infty} \hat{r}_{ij}(t, s)$. De forma similar al caso absorbente tenemos

$$\begin{aligned} \hat{r}_{ij}(t, s) &= \mathbb{P}(X_t = j | X_0 = i, t + s < S) \\ &= P_{ij}(t) \frac{\mathbb{P}(s < S | X_0 = j)}{\mathbb{P}(t + s < S | X_0 = i)} \\ &= P_{ij}(t) \frac{G_j(s)}{G_i(t + s)} \\ &= P_{ij} \frac{\int_0^\infty e^{-sx} Q_j(x) x^{-1} d\phi(x)}{\int_0^\infty e^{-(s+t)x} Q_i(x) x^{-1} d\phi(x)}. \end{aligned}$$

Y por los mismos argumentos que en el caso absorbente el teorema se sigue. \square

Como ejemplo, de nuevo analizaremos el proceso lineal de nacimiento y muerte con tasas dadas por

$$\lambda_n = (n + 1)\lambda \quad \text{y} \quad \mu_n = \mu n, \quad n \geq 0.$$

Supongamos que $\lambda > \mu$, de manera que tengamos un “drift” al infinito. Como vimos antes en (3.37), los polinomios de nacimiento y muerte para este proceso pueden escribirse en función de los polinomios de Meixner, y la medida espectral tiene sus puntos de masa en $(n + 1)(\lambda - \mu)$, $n = 0, 1, \dots$. Así, $\xi_1 = \lambda - \mu$. Calculando vemos que

$$\pi_n = \left(\frac{\lambda}{\mu}\right)^n \quad \text{y} \quad Q_n(\xi_1) = \left(\frac{\mu}{\lambda}\right)^n$$

y se deduce que

$$\frac{\pi_j Q_j^2(\xi_1)}{\sum_{n=0}^\infty \pi_n Q_n^2(\xi_1)} = \left(1 - \frac{\mu}{\lambda}\right) \left(\frac{\mu}{\lambda}\right)^j, \quad j = 0, 1, \dots$$

y por lo tanto, la distribución límite condicional es la distribución geométrica $Geo(\mu/\lambda)$.

3.2. Representación espectral para caminatas aleatorias

Como ya lo mencionamos, una clase importante de cadenas de Markov discretas son las caminatas aleatorias $X_n, n = 0, 1, 2, \dots$, cuyo espacio de estados es un conjunto de enteros consecutivos y cuyas probabilidades de transición de un paso

$$P_{ij} = \mathbb{P}\{X_{n+1} = j | X_n = i\}$$

forman una matriz de Jacobi; esto es, $P_{ij} = 0$ si $|i - j| > 1$. Como las transiciones sólo pueden ocurrir en los estados vecinos, podemos ver a nuestro proceso como una versión discreta del modelo continuo de difusión.

Usaremos la siguiente notación $P_{i,i-1} = q_i, P_{i,i} = r_i, P_{i,i+1} = p_i$. La probabilidad de transición de m -pasos

$$P_{ij}^{(m)} = \mathbb{P}\{X_{n+m} = j | X_n = i\}$$

forma una matriz $P^{(m)}$ que satisface

$$P^{(0)} = I, \quad P^{(1)} = (P_{ij}) = P,$$

$$P^{(m+1)} = P^{(m)}P = PP^{(m)}$$

de manera que $P^{(m)}$ es simplemente el producto de m copias de P .

Es conveniente distinguir 3 casos de acuerdo a si el espacio de estados es el conjunto finito $0, 1, \dots, N$, o el conjunto semi-finito $0, 1, \dots, n, \dots$, o el conjunto doblemente infinito $\dots, -1, 0, 1, \dots$.

El problema de las caminatas aleatorias puede ser descrito como sigue: La matriz fundamental P está dada y se necesita relacionar propiedades cualitativas del proceso de Markov a las propiedades cualitativas de P y calcular funcionales del proceso en términos de P .

Nuestro tratamiento obtiene una representación integral de la matriz de transición a través de la cual la estructura probabilística del proceso pueda ser analizada. La representación integral involucra un sistema de polinomios ortogonales con respecto a una distribución $\psi(x)$ en el intervalo cerrado $[-1, 1]$.

Una manera de ver la representación espectral es como sigue: Asociemos con el n -ésimo estado un polinomio $Q_n(x)$ de grado n , definido por las relaciones de recurrencia

$$(3.45) \quad xQ_n(x) = p_n Q_{n+1}(x) + r_n Q_n(x) + q_n Q_{n-1}(x), \quad n \geq 0,$$

$$Q_0(x) = 1, \quad Q_{-1}(x) = 0.$$

Notemos que el coeficiente de $Q_m(x)$ del lado derecho es precisamente la probabilidad de moverse del estado n al estado m en una transición. De manera similar vemos que

$$x^k Q_n(x) = \sum_{m=0}^{\infty} P_{nm}^k Q_m(x).$$

Haciendo uso de la teoría de momentos [10], se puede mostrar que $Q_n(x)$ constituye un sistema de polinomios ortogonales con respecto a una distribución $\psi(x)$. Usando esta propiedad de ortogonalidad de Q_n deducimos que

$$(3.46) \quad P_{nm}^k = \int_{-1}^1 x^k Q_n(x) Q_m(x) d\psi(x) / \int_{-1}^1 Q_m^2(x) d\psi(x).$$

De (3.45), se sigue que

$$\int Q_m^2 d\psi(x) = \frac{q_1 q_2 \cdots q_m}{p_0 p_1 \cdots p_{m-1}}.$$

La representación integral (3.46) exhibe la manera más simple posible de la dependencia de P_{nm}^k sobre k, n y m . Utilizando las propiedades de las componentes $Q_n(x)$ y $\psi(x)$ de la fórmula integral podemos analizar las características usuales de recurrencia y absorción del proceso.

Fórmula de representación

Asumimos que $p_i > 0, r_i \geq 0, q_{i+1} > 0$ para $i \geq 0$, y $p_i + q_i + r_i \leq 1$. La desigualdad $p_i + q_i + r_i < 1$ puede ser interpretada en términos de un estado ignorado i^* que es un estado de absorción permanente del proceso, con probabilidad de transición de un paso de i a i^* igual a $1 - (p_i + q_i + r_i)$.

La matriz P determina una transformación lineal en el espacio de todas las sucesiones $f = \{f(i)\}, i \geq 0$, de números complejos, por medias de la fórmula

$$(Pf)(i) = \sum_{j=0}^{\infty} P_{ij}f(j).$$

La serie en el lado derecho de la igualdad tiene a lo más 3 términos no cero. La solución $P\phi = x\phi$, donde x es una constante real o compleja, es única salvo un factor constante. Normalizamos la sucesión $\phi = \{\phi(i)\}$ por la relación $\phi(0) = 1$. El n -ésimo término de la sucesión ϕ es entonces un polinomio $Q_n(x)$ en x de grado exactamente n , y estos polinomios satisfacen las relaciones de recurrencia

$$\begin{aligned} Q_0(x) &= 1, \\ xQ_0(x) &= r_0Q_0(x) + p_0Q_1(x), \\ (3.47) \quad xQ_n(x) &= q_nQ_{n-1}(x) + r_nQ_n(x) + p_nQ_{n+1}(x). \end{aligned}$$

Introducimos ahora las cantidades $\{\pi_i\}$ definidas como las soluciones de las ecuaciones de simetría $P_{ij}\pi_i = P_{ji}\pi_j$ normalizadas por la condición $\pi_0 = 1$. Tenemos entonces que para $n \geq 1, \pi_n = (p_0p_1 \cdots p_{n-1}) / (q_1q_2 \cdots q_n)$. Con esto formamos el espacio de Hilbert $L_2(\pi)$ de todas las sucesiones $f = \{f(i)\}$ de números complejos tales que

$$\|f\|^2 = \sum_{i=0}^{\infty} |f(i)|^2 \pi_i$$

es finito, y notemos el siguiente:

Lema 3.2.1. *La sucesión de transformaciones $f \rightarrow Pf$ induce en $L_2(\pi)$ un operador lineal autoadjunto y acotado T de norma ≤ 1 .*

Demostración. La propiedad de ser autoadjunto se sigue de las ecuaciones de simetría, y la desigualdad de la norma del hecho de que

$$\sum_i \pi_i P_{ij} \leq \pi_j$$

para cada j .

Así $f \in L_2(\pi)$ implica $P(f) = T(f)$, e iterando $P^{(n)}f = T^n f$. Para expresar las componentes de la matriz $P^{(n)}$ en términos del operador T , introducimos las sucesiones $e^{(i)} = \{e_j^{(i)}\}$ definidas por $e_j^{(i)} = \delta_j^i / \pi_i$. Usando las ecuaciones de simetría en la forma $p_i\pi_i = q_{i+1}\pi_{i+1}$, vemos que

$$Te^{(i)} = q_i e^{(i-1)} + r_i e^{(i)} + p_i e^{(i+1)},$$

con una modificación para $i = 0$, y por lo tanto por un argumento inductivo basado en la relación de recurrencia (3.47), $Q_n(T)e^{(0)} = e^{(n)}, n = 0, 1, 2, \dots$. Ahora el producto interno $(T^n e^{(j)}, e^{(i)})$ es igual a $P_{ij}^{(n)} / \pi_j$, y por lo tanto

$$P_{ij}^{(n)} = \pi_j(T^n Q_j(T)e^{(0)}, Q_i(T)e^{(0)}).$$

Los polinomios $Q_n(x)$ tienen coeficientes reales así que los polinomios $Q_n(T)$ son autoadjuntos, y como conmutan con T ,

$$P_{ij}^{(n)} = \pi_j(T^n Q_i(T)Q_j(T)e^{(0)}, e^{(0)}).$$

Como consecuencia, si $\{E_x\}$ es la solución a la igualdad correspondiente al operador autoadjunto T , y si $\psi(x) = (E_x e^{(0)}, e^{(0)})$, entonces

$$(3.48) \quad P_{ij}^{(n)} = \pi_j \int_{-1}^1 x^n Q_i(x) Q_j(x) d\psi(x),$$

donde la integral incluye cualquier salto que pueda estar presente en $x = 1$ o en $x = -1$. Esta fórmula es la representación básica de la matriz de transición de probabilidad de la caminata aleatoria, y ψ es llamada la medida espectral de la caminata aleatoria. \square

Teorema 3.2.1. *Existe una única distribución positiva regular ψ en $-1 \leq x \leq 1$ tal que (3.48) es válido para toda i, j, n .*

Demostración. La existencia de tal ψ ha sido demostrada. Pero (3.48) con $i = j = 0$ determina todos los momentos de ψ y por lo tanto determina a ψ de manera única.

Haciendo $n = 0$ en (3.48) vemos que los polinomios $Q_n(x)$ son polinomios ortogonales que pertenecen a ψ , de hecho,

$$(3.49) \quad \pi_j \int_{-1}^1 Q_i(x) Q_j(x) d\psi(x) = \delta_{ij}.$$

\square

Un caso importante es cuando cada r_n es 0. Si éste es el caso, decimos que la caminata aleatoria es simétrica, el nombre se justifica en parte por que la distribución ψ es simétrica alrededor de $x = 0$. El nombre es además justificado por el siguiente resultado inverso.

Lema 3.2.2. *Cada distribución positiva y simétrica en $-1 \leq x \leq 1$ no soportada por un conjunto finito de puntos, y con masa total igual a 1, es la medida espectral de una caminata aleatoria simétrica. La caminata aleatoria es única si además incluimos la condición de que $p_0 = 1$ y $p_n + q_n = 1$ para $n \geq 1$.*

Demostración. El sistema de polinomios ortogonales con respecto a ψ tiene todos sus ceros en el intervalo abierto $-1 < x < 1$. Sean $Q_n(x)$ estos polinomios normalizados por la condición $Q_n(1) = 1$. La simetría de ψ implica que Q_n es par o impar con n , así que la relación de recurrencia es de la forma

$$xQ_n = q_n Q_{n-1} + p_n Q_{n+1}.$$

Como $Q_n(x) \geq 1$ para $x \geq 1$, se sigue que para toda $p_n > 0$, y haciendo $x = 1$ obtenemos $p_n + q_n = 1$. Multiplicando la fórmula de recurrencia por x^{n-1} y usando la ortogonalidad obtenemos $I_n = q_n I_{n-1}$, donde $I_n = \int_{-1}^1 Q_n(x) d\psi(x)$. Como el coeficiente de x^n en Q_n es más grande que 0, vemos que $I_n > 0$ y por lo tanto $q_n > 0$.

Así la relación de recurrencia determina una única matriz P (construida a partir de las p'_n s y las q'_n s) que es la matriz de transición de probabilidad de una caminata aleatoria simétrica, y está claro que ψ es la medida espectral de esta caminata aleatoria. \square

Recurrencia, absorción y ejemplos

La fórmula de representación (3.48) provee una herramienta para relacionar propiedades de recurrencia del proceso con propiedades de la medida espectral ψ por un lado, y propiedades de la matriz básica P por el otro. Consideremos por ejemplo la distribución del primer tiempo de llegada. Para $i \neq j$, sea f_{ij}^n la probabilidad de que si el estado inicial es i , entonces el estado j sea alcanzado por primera vez en la transición n , y sea f_{ii}^n la probabilidad de que si el estado inicial es i , entonces i se alcanza de nuevo por primera vez en la n -ésima transición. Así las funciones generadoras

$$P_{ij}(s) = \sum_{n=0}^{\infty} P_{ij}^n s^n, \quad F_{ij}(s) = \sum_{n=1}^{\infty} f_{ij}^n s^n$$

están conectadas por las fórmulas

$$(3.50) \quad \begin{aligned} P_{ij}(s) &= F_{ij}(s)P_{jj}(s), \quad i \neq j \\ P_{ii}(s) &= 1 + F_{ii}(s)P_{ii}(s). \end{aligned}$$

Estas relaciones junto con (3.48) hacen posible expresa las $F_{ij}(s)$ en términos de las integrales espectrales. Por ejemplo,

$$(3.51) \quad F_{00}(s) = 1 - 1 / \int_{-1}^1 \frac{d\psi(x)}{1 - xs}.$$

A pesar de que la distribución de los momentos de primera llegada se han determinado por otros métodos [5], es conveniente describir cómo la representación integral conducen a este resultado. El proceso es recurrente si y sólo si $F_{00} \rightarrow 1$ conforme $s \rightarrow 1$, y esto es equivalente a la divergencia de $\int_{-1}^1 (1-x)^{-1} d\psi(x)$. Usando las herramientas analíticas de la teoría de los polinomios ortogonales, integrales de la forma

$$\int_{-1}^1 \frac{Q_i(x)Q_j(x)}{(1-x)^n} d\psi(x),$$

la cual aparece al calcular de las distribuciones de primeras llegadas, pueden ser evaluadas en términos de las constantes p_n, q_n, r_n del proceso. Los cálculos pueden ser hechos de manera que dependan de resultados de cálculos similares llevados a cabo en [8]. Por ejemplo, una condición necesaria para la recurrencia es que $r_0 + p_0 = 1$ y para $n \geq 1, q_n + r_n + p_n = 1$. Si esta condición se satisface, entonces los polinomios $R_n(x) = Q_n(1-x)$ satisfacen una relación de recurrencia

$$-xR_n = q_n R_{n-1} - (p_n + q_n)R_n + p_n R_{n+1}$$

y son ortogonales en $0 \leq x \leq 2$ con respecto a la distribución θ obtenida de ψ haciendo un cambio de variable apropiado. Usando resultados de [8],

$$\int_{-1}^1 \frac{d\psi(x)}{1-x} = \int_0^2 \frac{d\theta(x)}{x} = \sum_{n=0}^{\infty} \frac{1}{p_n \pi_n}.$$

Así la caminata aleatoria es recurrente si y sólo si $\sum 1/p_n \pi_n$ es divergente [5].

Para un proceso recurrente los tiempos de primera llegada esperados son todos finitos si y sólo si $\lim_{n \rightarrow \infty} P_{00}^{2n}$ es positivo, y en este caso el proceso se llama ergódico.

Como $x^{2n} \rightarrow 0$ monótonamente en $-1 < x < 1$ conforme $n \rightarrow \infty$, se puede ver de (3.48) que el proceso es ergódico si y sólo si ψ tiene un salto en $x = 1$ o en $x = -1$. Si ψ no tiene salto en $x = 1$, entonces la cantidad del salto en $x = -1$ es

$$-\lim_{n \rightarrow \infty} \int_{-1}^1 x^{2n+1} d\psi(x) = -\lim_{n \rightarrow \infty} P_{00}^{2n+1} \leq 0,$$

por lo cual no hay salto en $x = -1$. Como consecuencia, el proceso es ergódico si y sólo si ψ tiene un salto en $x = 1$. Ahora un salto en $x = 1$ ocurre si y sólo si $x = 1$ es un eigenvalor del operador autoadjunto T , y éste es el caso si y sólo si la serie

$$\frac{1}{\rho} = \sum_{n=0}^{\infty} Q_n^2(1)\pi_n$$

es convergente, en cuyo caso ρ es la cantidad del salto. Se puede deducir de la fórmula de recurrencia que la serie diverge si $q_n + r_n + p_n < 1$ para alguna n . Está también claro que esta condición, la cual significa probabilidad de absorción positiva en n , hace al proceso transitorio. Si $r_0 + p_0 = 1$ y $q_n + r_n + p_n = 1$ para toda $n \geq 1$, entonces $Q_n(1) = 1$ para toda n y $\rho^{-1} = \sum_{n=0}^{\infty} \pi_n$, y el proceso es ergódico si y sólo si la serie converge. Notamos que si todas las r_n se anulan de manera que ψ es simétrica, entonces cualquier salto en $x = 1$ está ligado con un salto igual en $x = -1$; pero por otro lado si alguna r_n es positiva, entonces $\lim_{n \rightarrow \infty} P_{00}^n$ existe, y esto implica que no hay salto en $x = -1$.

Supongamos ahora que $q_0 = 1 - (r_0 + p_0)$ es positiva. Cada vez que el estado cero es visitado hay una probabilidad q_0 de que la absorción ocurrirá en la siguiente transición. Sea A_i^n la probabilidad de que cuando un estado inicial es i , la absorción del estado cero ocurre en la n -ésima transición. Claramente $A_i^n = q_0 P_{i0}^{n-1}$ y por lo tanto

$$A_i(s) = \sum_{n=1}^{\infty} A_i^n s^n = q_0 s \int_{-1}^1 \frac{Q_i(x)}{1 - xs} d\psi(x).$$

Si $q_n + r_n + p_n = 1$ para $n \geq 1$, el cálculo de los momentos de la distribución de los tiempos de absorción pueden ser reducidos a resultados de [8] por medias de lo descrito anteriormente. La k -ésima caminata aleatoria asociada ($k \geq 0$) que pertenece a la caminata aleatoria dada está definida como el proceso obtenido de empezar la caminata aleatoria dada en un estado $i > k$ y pararla cuando el estado k es alcanzado por primera vez. En particular, la 0-ésima caminata aleatoria asociada tiene al estado cero como un estado ignorado de absorción. La idea de la k -ésima caminata aleatoria no es nueva y se ha usado de manera implícita en [4], [5]. Sea ψ la medida espectral de una caminata aleatoria y α la medida espectral de la 0-ésima caminata aleatoria asociada. El evento de primera llegada del estado 1 al estado 0 para la caminata aleatoria dada puede ser vista como un evento de absorción para el 0-ésimo proceso asociado. Expresando la distribución del tiempo de ocurrencia de este evento por un lado en términos de ψ , y por otro lado en términos de α , obtenemos la identidad

$$\frac{P_{10}(s)}{P_{00}(s)} = F_{10}(s) = q_{1s} \int_{-1}^1 \frac{d\alpha(x)}{1 - xs},$$

que haciendo simplificaciones lleva a

$$(3.52) \quad \int_{-1}^1 \frac{d\psi(x)}{1 - xs} = 1 / \left(1 - r_0 s - q_{1p_0} \int_{-1}^1 \frac{d\alpha(x)}{1 - xs} \right).$$

La transformada de Stieltjes de una medida finita θ en $-1 \leq x \leq 1$ está definida por

$$B(z; \theta) = \int_{-1}^1 \frac{d\theta(x)}{x - z}.$$

La ecuación (3.52) es equivalente a la identidad

$$(3.53) \quad B(x; \psi) = -1/(z - r_0 + p_0 q_1 B(z; \alpha))$$

entre la transformada de Stieltjes de ψ y α . Esta identidad es usada frecuentemente para calcular medidas espectrales.

Ejemplos

1. La más sencilla de las caminatas aleatorias semi-infinitas tiene $p_n = p$ para $n \geq 0$, $q_n = q$ para $n \geq 1$. El proceso es algunas veces llamado como “la apuesta contra el apostador de riqueza infinita” (véase [2] para interpretaciones). La transformada de Stieltjes de la medida espectral $B(z; \psi) = B(z)$ satisface (3.53) la cual se reduce a

$$pqB^2(z) + zB(z) + 1 = 0,$$

y así

$$(3.54) \quad B(z) = (-z + (z^2 - 4pq)^{1/2})/2pq,$$

donde la raíz cuadrada está determinada por la continuación analítica de valores positivos para $z > 1$. La fórmula la cual da ψ en términos de $B(z; \psi)$ es

$$\int_{-1}^{x_0} d\psi(x) = \frac{1}{\pi} \lim_{\eta \rightarrow 0^+} \int_{-1-\epsilon}^{x_0} \text{Im} B(\xi + i\eta) d\xi$$

donde $\epsilon > 0$ y x_0 es cualquier punto donde ψ no tiene saltos. Como consecuencia ψ consiste de la densidad continua

$$\psi'(x) = (4pq - x^2)^{1/2}/2\pi pq$$

sobre el intervalo $-(4pq)^{1/2} < x < (4pq)^{1/2}$. La fórmula de recurrencia

$$xQ_n = qQ_{n-1} + pQ_{n+1}$$

puede ser reducida por transformación de variables a

$$\xi R_n = \frac{1}{2}R_{n-1} + \frac{1}{2}R_{n+1},$$

y uno encuentra que

$$Q_n(x) = (q/p)^{n/2} U_n(x/(4pq)^{1/2})$$

donde las $U_n(\xi)$ son los polinomios de Chebycheff del segundo tipo [1].

2. Para continuar ilustrando el uso de (3.53), consideremos la caminata aleatoria con $q_n, p_n = p$ para toda $n \geq 1$, pero p_0 y r_0 arbitrarios. Si ψ es la medida espectral del proceso, y α la medida espectral del 0-ésimo proceso asociado, entonces α es el mismo que ψ del ejemplo anterior y (3.53) se convierte en

$$B(z; \psi) = \frac{r_0 - (1 - p_0/2p)z + (p_0/2p)(z^2 - 4pq)^{1/2}}{(1 - p_0/p)z^2 - 2r_0(1 - p_0/2p)z + r_0^2 + p_0^2q/p},$$

del cual ψ puede ser calculado.

Ponemos atención a dos casos:

- a) Sea $r_0 = 0$. Entonces

$$\psi'(x) = \begin{cases} \frac{1}{2\pi} \frac{p_0(4pq - x^2)^{1/2}}{x^2(p - p_0) + p_0^2q} & \text{si } -(4pq)^{1/2} < x < (4pq)^{1/2} \\ 0 & \text{en otro caso} \end{cases}$$

También, $\psi(x)$ tiene saltos localizados en $\pm(p_0^2q/(p_0 - p))^{1/2}$ de magnitud

$$\frac{1}{2}(p_0 - 2p)/(p_0 - p)$$

dado $p_0 > 2p$. Cuando $p_0 < 2p$, entonces la distribución ψ es pura densidad como se describió anteriormente.

- b) Sea $p_0 = p$ y $r_0 = q = 1 - p$. Entonces

$$\psi'(x) = \begin{cases} \frac{1}{2\pi(1-p)} \frac{(4pq - x^2)^{1/2}}{1-x} & \text{si } -(4pq)^{1/2} < x < (4pq)^{1/2} \\ 0 & \text{de otra forma} \end{cases}$$

y para $p < 1/2$, $\psi(x)$ posee un salto adicional localizado en 1 de magnitud $(1 - 2p)/(1 - p)$.

Por (3.53), o ψ o α pueden ser calculados si el otro es conocido. Por iteración de esta relación la medida espectral de una caminata aleatoria puede ser calculada dados todos salvo un número finito de las p_n, r_n, q_n que son iguales a aquellos de una caminata aleatoria cuya medida espectral es conocida.

3. Modelo de Ehrenfest

En este modelo la caminata aleatoria se mueve sobre un espacio de estados finito $\{0, 1, 2, \dots, 2N\}$. Aquí $2N$ se interpreta como el número de bolas en total enumeradas y repartidas en dos urnas A y B , donde inicialmente en A hay i bolas y en B hay $2N - i$. El proceso evoluciona acorde a las siguiente estrategia: escogemos un número al azar entre 1 y $2N$. Si esa bola está en la urna A , la pasamos a la urna B en la siguiente tirada. Mientras que si está en la urna B , la pasamos a la urna A . La caminata aleatoria cuenta el número de bolas en A a medida que repetimos este intercambio en tiempo discreto.

Los coeficientes de la caminata aleatoria vienen dados por

$$p_n = \frac{2N - n}{2N}, n = 0, 1, \dots, 2N - 1, \quad q_n = \frac{n}{2N}, n = 1, 2, \dots, 2N, \quad r_n = 0, \quad n = 0, 1, \dots, 2N,$$

4. Modelo de Laplace-Bernoulli

Este modelo también corresponde con un modelo de urnas y se puede describir de la siguiente manera: inicialmente tenemos dos urnas A y B , donde en la urna A hay α bolas blancas y en la urna B hay β bolas negras. En la siguiente transición tomamos una bola de A y otra de B y las intercambiamos. El estado de la caminata aleatoria viene descrita por el número de bolas blancas n en la urna A en cierto momento. Claramente el espacio de estados es discreto $\{0, 1, 2, \dots, \alpha\}$ y las probabilidades de transición vienen dadas por

$$\begin{aligned} p_n &= \frac{\alpha - n}{\alpha} \frac{\alpha - n}{\beta}, \quad n = 0, 1, \dots, \alpha - 1, \\ q_n &= \frac{n}{\alpha} \frac{\beta - \alpha + n}{\beta}, \quad n = 1, 2, \dots, \alpha, \\ r_n &= \frac{n}{\alpha} \frac{\alpha - n}{\beta} + \frac{\alpha - n}{\alpha} \frac{\beta - \alpha + n}{\beta}, \quad n = 0, 1, \dots, \alpha. \end{aligned}$$

En este caso, al igual que el anterior, se pueden describir todos los autovalores de la matriz de transición de probabilidades P (aunque este fenómeno en general no va a ser posible). que vienen dados por

$$1 + \frac{x(x - \alpha - \beta - 1)}{\alpha\beta}, \quad x = 0, 1, \dots, \alpha.$$

Los polinomios generados por P son los conocidos polinomios ortogonales duales de Hahn $R_n(x)$, y que son ortogonales con respecto a la medida $(\gamma = -\alpha - 1, \delta = -\beta - 1, N = \alpha)$:

$$\mu(x) = \sum_{x=0}^{\alpha} \frac{(2x - \alpha - \beta - 1)\alpha!(-\alpha)_x(-\alpha)_x}{(-1)^x(x - \alpha - \beta - 1)_{\alpha+1}(-\beta)_x x!} \delta_x,$$

es decir,

$$\sum_{x=0}^{\alpha} R_i(\lambda(x))R_j(\lambda(x))\mu(x) = \pi_j^{-1}\delta_{ij}, \quad \pi_j = \binom{j - \alpha - 1}{j} \binom{\alpha - \beta - j - 1}{\alpha - j},$$

donde $\lambda(x) = x(x - \alpha - \beta - 1)$. La representación integral de Karlin-McGregor en este caso viene dada por

$$P_{ij}^{(n)} = \pi_j \sum_{x=0}^{\alpha} \left(1 + \frac{x(x - \alpha - \beta - 1)^n}{\alpha\beta}\right)^n R_i(\lambda(x))R_j(\lambda(x))\mu(x).$$

3.2.1. Distribución límite condicional para caminatas aleatorias

Algunas veces la cadena no es recurrente positiva pero tiene un estado absorbente. Podemos considerar la distribución límite condicional $\lim_{n \rightarrow \infty} r_{ij}(n)$ con

$$(3.55) \quad r_{ij}(n) = \mathbb{P}(X_n = j | X_0 = i, n < T < \infty)$$

donde T denota el tiempo de absorción al estado -1 . Esta es la probabilidad límite de estar en el estado j dado que el proceso está al tiempo n en $\{0, 1, 2, \dots\}$ pero el estado de absorción -1 ocurre eventualmente.

Distribución límite condicional doble

La distribución límite condicional doble $\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} r_{ij}(n, m)$ con

$$(3.56) \quad r_{ij}(n, m) = \mathbb{P}(X_n = j | X_0 = i, n + m < T < \infty)$$

y donde T de nuevo denota el tiempo de absorción en el estado -1 , puede ser de nuevo de interés. Esta es la probabilidad límite de estar en el estado j dado que el proceso no dejará $\{0, 1, 2, \dots\}$ en el futuro distante, pero que la absorción en el estado -1 ocurre eventualmente. Tenemos el siguiente teorema, donde hacemos

$$\eta = \sup \text{supp}(\phi)$$

y

$$C_k(\phi) = \frac{\int_{-1}^0 (-x)^k d\phi(x)}{\int_0^1 x^k d\phi(x)}.$$

Teorema 3.2.2. *Si $(X_n)_n$ es aperiódica, entonces la distribución límite condicional de (3.55) y la distribución límite condicional doble de (3.56) existen, si $\eta < 1$ y $C_k(\phi) \rightarrow 0$ conforme $k \rightarrow \infty$, en cuyo caso los límites están dados por*

$$\lim_{n \rightarrow \infty} r_{ij}(n) = \frac{\pi_j Q_j(\eta)}{\sum_{k=0}^{\infty} \pi_k Q_k(\eta)} \quad y \quad \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} r_{ij}(n, m) = \frac{\pi_j Q_j^2(\eta)}{\sum_{k=0}^{\infty} \pi_k Q_k^2(\eta)}.$$

Demostración. Para probar el primer límite acudimos a [18].

Para el segundo procedemos como en los teoremas (3.43) y (3.44) del caso continuo. Tenemos

$$\begin{aligned} r_{ij}(n, m) &= \mathbb{P}(X_n = j | X_0 = i, n + m < T < \infty) \\ &= P_{ij}(n) \frac{\mathbb{P}(m < T < \infty | X_0 = i)}{\mathbb{P}(n + m < T < \infty | X_0 = i)}. \end{aligned}$$

Por la proposición 2 de [18] tenemos que si $(X_n)_n$ es aperiódica entonces

$$\lim_{m \rightarrow \infty} \frac{\mathbb{P}(m < T < \infty | X_0 = j)}{\mathbb{P}(m + n < T < \infty | X_0 = i)}$$

existe si $\eta = 1$ o $\eta < 1$ y $C_k(\phi) \rightarrow \infty$ conforme $k \rightarrow \infty$ y

$$\lim_{m \rightarrow \infty} \frac{\mathbb{P}(m < T < \infty | X_0 = j)}{\mathbb{P}(m + n < T < \infty | X_0 = i)} = \frac{Q_j(\eta)}{\eta^n Q_i(\eta)}.$$

Así que tenemos

$$\lim_{m \rightarrow \infty} r_{ij}(n, m) = P_{ij}(n) \frac{Q_j(\eta)}{Q_i(\eta) \eta^n}.$$

Y así, por el teorema de representación espectral y dado que $(X_n)_n$ es aperiódica

$$\phi(\{-\eta\}) = 0$$

con lo que concluimos que

$$\begin{aligned} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} r_{ij}(n, m) &= \lim_{n \rightarrow \infty} \frac{Q_j(\eta)}{Q_i(\eta)} \int_{-1}^1 \left(\frac{x}{\eta}\right)^n Q_i(x) Q_j(x) d\phi(x) \\ &= Q_j^2(\eta) \pi_j \phi(\{\eta\}) \\ &= \frac{\pi_j Q_j^2(\eta)}{\sum_{k=0}^{\infty} \pi_k Q_k^2(\eta)}. \end{aligned}$$

Y esto completa la demostración. \square

3.3. Representación espectral de la densidad de transición para procesos de difusión

Sea $\{X_t, t > 0\}$ un proceso de difusión con espacio de estados \mathcal{S} y operador infinitesimal \mathcal{A} dado por (1.42), con coeficiente deriva $\mu(x)$ y coeficiente de difusión $\sigma^2(x)$. Para un x_0 arbitrario, definamos la función

$$(3.57) \quad I(x_0, x) = \int_{x_0}^x \frac{2\mu(z)}{\sigma^2(z)} dz.$$

Obsérvese que

$$\frac{d}{dx} I(x_0, x) = \frac{2\mu(x)}{\sigma^2(x)}.$$

Para esta función definimos ahora la medida (salvo constante multiplicativa)

$$(3.58) \quad m(x) = \frac{2}{\sigma^2(x)} e^{I(x_0, x)}.$$

La elección de x_0 no influye en el comportamiento del proceso de difusión. Se tiene entonces que el operador infinitesimal \mathcal{A} se puede reescribir como

$$(3.59) \quad \mathcal{A} = \frac{1}{m(x)} \frac{d}{dx} \left(\frac{\sigma^2(x)}{2} m(x) \frac{d}{dx} \right).$$

En efecto, sea f cualquier función dos veces continuamente diferenciable. Entonces

$$\begin{aligned} \frac{1}{m(x)} \frac{d}{dx} \left(\frac{\sigma^2}{2} m(x) \frac{d}{dx} \right) f &= \frac{1}{m(x)} \frac{d}{dx} \left(e^{I(x_0, x)} \frac{d}{dx} \right) f = \frac{1}{m(x)} \left(\left[\frac{d}{dx} e^{I(x_0, x)} \right] \frac{d}{dx} + e^{I(x_0, x)} \frac{d^2}{dx^2} \right) f \\ &= \frac{\sigma^2(x)}{2} e^{-I(x_0, x)} \left(e^{I(x_0, x)} \left[\frac{2\mu(x)}{\sigma^2(x)} \frac{d}{dx} + \frac{d^2}{dx^2} \right] \right) f = \frac{1}{2} \sigma^2(x) f''(x) + \mu(x) f'(x) = (\mathcal{A}f)(x). \end{aligned}$$

Considérese ahora el espacio de Hilbert $L_m^2(\mathcal{S})$ con el producto interno

$$(f, g)_m = \int_{\mathcal{S}} f(x)g(x)m(x)dx.$$

Debido a (2.30) y el hecho de que \mathcal{A} se puede reescribir de la forma (3.59), se tiene que \mathcal{A} es autoadjunto con respecto al producto interno anterior, es decir $(\mathcal{A}f, g)_m = (f, \mathcal{A}g)_m$ para toda $f, g \in L_m^2(\mathcal{S})$. En particular, la medida $m(x)$ cumple con la ecuación de Pearson

$$\frac{1}{2}(\sigma^2(x)m(x))' = \mu(x)m(x),$$

con lo cual, en el caso de existir la distribución invariante debe coincidir con $m(x)$.

Sea ahora f una función continua y acotada en \mathcal{S} y consideremos la función

$$(3.60) \quad u(t, x) = \mathbb{E}[f(X_t)|X_0 = x].$$

Sabemos que esta función verifica la ecuación de retroceso de Kolmogorov, i.e.,

$$(3.61) \quad \frac{\partial}{\partial t}u(t, x) = \mathcal{A}u(t, x), \quad u(0, x) = f(x).$$

La idea es tratar de calcular esta función $u(t, x)$ por el método de separación de variables. Supongamos que

$$u(t, x) = c(t)\phi(x).$$

Sustituyendo en la ecuación de Kolmogorov se tiene que

$$(3.62) \quad \frac{c'(t)}{c(t)} = \frac{\mathcal{A}\phi(x)}{\phi(x)} = \lambda,$$

El lado izquierdo depende sólo de la variable en t , mientras que el lado derecho involucra sólo a la variable x . Esto es consistente sólo si (3.62) es constante, obteniendo así

$$c'(t) = \lambda c(t)$$

y

$$(3.63) \quad \mathcal{A}\phi(x) = -\lambda\phi(x),$$

donde λ es cierta constante. La ecuación diferencial para $c(t)$ se puede resolver a partir de las condiciones iniciales. Con lo cual $c(t) = e^{\lambda t}$. La segunda ecuación diferencial es de segundo orden y en el caso de existir solución, ésta sería una autofunción del operador \mathcal{A} con autovalor λ . Si podemos encontrar dicha solución, se tendrá que una solución $u(t, x)$ vendrá dada por

$$u(t, x) = e^{\lambda t}\phi(x).$$

Supongamos ahora que el conjunto de autovalores de \mathcal{A} (contando multiplicidades) es numerable, y los llamamos

$$\lambda_0, \lambda_1, \lambda_2, \dots$$

con autofunciones

$$\phi_0(x), \phi_1(x), \phi_2(x), \dots$$

de tal manera que los normalizamos con respecto al producto interno, i.e. $\|\phi_n\|_m = 1, n \geq 0$. Si $\lambda_i \neq \lambda_j$, entonces $\phi_i(x)$ y $\phi_j(x)$ son ortogonales con respecto al producto interno, i.e.

$$(\phi_i, \phi_j)_m = 0, i \neq j.$$

Supongamos que esta sucesión de autofunciones $(\phi_n)_n$ es completa en $L_m^2(\mathcal{S})$. Por lo tanto, para toda $f \in L_m^2(\mathcal{S})$, usando la expansión de Fourier con respecto a la base $(\phi_n)_n$, se tiene que

$$f(x) = \sum_{n=0}^{\infty} (f, \phi_n)_m \phi_n(x).$$

Por el principio de superposición, cada una de estas soluciones es solución de $u(t, x)$. Por lo tanto para una clase amplia de funciones f , se tendrá que $(T_t f)(x) = u(t, x)$. Por lo tanto

$$\begin{aligned} \int_{\mathcal{S}} f(y) p(t; x, dy) &= \sum_{n=0}^{\infty} e^{\lambda_n t} (f, \phi_n)_m \phi_n(x) = \sum_{n=0}^{\infty} e^{\lambda_n t} \left(\int_{\mathcal{S}} f(y) \phi_n(y) m(y) dy \right) \phi_n(x) \\ &= \int_{\mathcal{S}} f(y) \left(\sum_{n=0}^{\infty} e^{\lambda_n t} \phi_n(x) \phi_n(y) m(y) \right) dy. \end{aligned}$$

Para hacer el intercambio de la integral y sumatorio en la última igualdad va a ser suficiente con que $\lambda_n \leq 0$, como veremos a continuación. Como consecuencia, la distribución de transición de probabilidades $p(t; x, dy)$ tendría una densidad $p(t; x, y)$ dada por

$$(3.64) \quad p(t; x, y) = m(y) \sum_{n=0}^{\infty} e^{\lambda_n t} \phi_n(x) \phi_n(y).$$

En muchos casos esta expresión va a ser legítima, pero va a depender de las condiciones iniciales de la ecuación de retroceso de Kolmogorov. Siguiendo los argumentos sobre las condiciones de contorno de soluciones de la ecuación de retroceso de Kolmogorov, en la mayoría de los casos tendremos que resolver una ecuación diferencial de segundo orden para $u(t, x)$ de tipo Sturm-Liouville en $\mathcal{S} = [a, b]$ con valores iniciales de la forma

$$(3.65) \quad \begin{aligned} \frac{\partial u}{\partial t} = \mathcal{A}u &= \frac{1}{m(x)} \frac{\partial}{\partial x} \left(\frac{\sigma^2}{2} m(x) \frac{\partial u}{\partial x} \right), \\ c_{1,1} u(t, a) + c_{1,2} \left(\frac{\sigma^2(x)}{2} m(x) \frac{\partial u}{\partial x} \right) (t, a) &= 0, \\ c_{2,1} u(t, b) + c_{2,2} \left(\frac{\sigma^2(x)}{2} m(x) \frac{\partial u}{\partial x} \right) (t, b) &= 0, \end{aligned}$$

para ciertas constantes $c_{i,j}, i, j = 1, 2$. De la teoría general de problemas de Sturm-Liouville se sabe (por lo menos para el caso en el que los extremos son reflectantes, i.e. $c_{1,1} = c_{2,1} = 0$) que el espectro de \mathcal{A} es discreto si a y b son finitos. En dicho caso también se tiene, usando las condiciones de contorno y la ecuación diferencial, que el autovalor se puede escribir como

$$\begin{aligned}\lambda_n &= (\mathcal{A}\phi_n, \phi_n)_m = \int_a^b \phi_n(x) \frac{d}{dx} \left[\frac{1}{2} \sigma^2(x) m(x) \phi_n'(x) \right] dx \\ &= \frac{1}{2} \phi_n(x) \sigma^2(x) m(x) \phi_n'(x) \Big|_a^b - \frac{1}{2} \int_a^b \sigma^2(x) m(x) (\phi_n'(x))^2 dx = -\frac{1}{2} \int_a^b \sigma^2(x) m(x) (\phi_n'(x))^2 dx.\end{aligned}$$

Con lo cual se tiene que (en particular $\lambda_0 = 0$)

$$\lambda_n < 0, \quad n \geq 0.$$

Si uno o ambos extremos son infinitos, podría ocurrir que el espectro de \mathcal{A} fuese continuo. Llamemos $\Xi \subset \mathbb{R}$ al espectro continuo del operador \mathcal{A} . La ortonormalidad de $\phi(x, \lambda)$ en este caso se entiende como

$$(3.66) \quad \int_{\mathcal{S}} m(x) \phi(x, \lambda) \phi(x, \mu) dx = \frac{\delta(\lambda - \mu)}{\psi(\lambda)},$$

para cierta función $\psi(\lambda)$. Si $f \in L_m^2(\mathcal{S})$, entonces existe una función \hat{f} tal que

$$f(x) = \int_{\Xi} \hat{f}(\lambda) \phi(x, \lambda) \psi(\lambda) d\lambda.$$

Para calcular $\hat{f}(\lambda)$ multiplicamos la expresión anterior por $\phi(x, \mu) m(x)$, integramos sobre \mathcal{S} y usamos la ortogonalidad (3.66). En efecto,

$$\begin{aligned}\int_{\mathcal{S}} f(x) \phi(x, \mu) m(x) dx &= \int_{\mathcal{S}} \left[\int_{\Xi} \hat{f}(\lambda) \phi(x, \lambda) \psi(\lambda) d\lambda \right] \phi(x, \mu) m(x) dx \\ &= \int_{\Xi} \hat{f}(\lambda) \left[\int_{\mathcal{S}} \phi(x, \lambda) \phi(x, \mu) m(x) dx \right] \psi(\lambda) d\lambda = \int_{\Xi} \hat{f}(\lambda) \delta(\lambda - \mu) d\lambda = \hat{f}(\mu).\end{aligned}$$

Las soluciones $u(t, x)$ en (3.60) se pueden escribir como

$$u(t, x) = \int_{\Xi} e^{\lambda t} \hat{f}(\lambda) \phi(x, \lambda) \psi(\lambda) d\lambda.$$

Por la definición de $(T_t f)(x)$ en (1.39) y ya que verifica la ecuación de retroceso de Kolmogorov para una clase amplia de funciones f , se tendrá que $(T_t f)(x) = u(t, x)$. Por lo tanto

$$\begin{aligned}\int_{\mathcal{S}} f(y) p(t; x, dy) &= \int_{\Xi} e^{\lambda t} \hat{f}(\lambda) \phi(x, \lambda) \psi(\lambda) d\lambda = \int_{\Xi} e^{\lambda t} \left[\int_{\mathcal{S}} f(y) \phi(y, \lambda) m(y) dy \right] \phi(x, \lambda) \psi(\lambda) d\lambda \\ &= \int_{\mathcal{S}} f(y) m(y) \left[\int_{\Xi} e^{\lambda t} \phi(x, \lambda) \phi(y, \lambda) \psi(\lambda) d\lambda \right] dy.\end{aligned}$$

Como consecuencia, la distribución de transición de probabilidades $p(t; x, dy)$ tendría una densidad $p(t; x, y)$ dada por

$$(3.67) \quad p(t; x, y) = m(y) \int_{\Xi} e^{\lambda t} \phi(x, \lambda) \phi(y, \lambda) \psi(\lambda) d\lambda.$$

También podría ocurrir que el espectro tenga una parte continua Ξ y otra discreta Θ . En ese caso la densidad de probabilidad $p(t; x, y)$ vendrá dada por

$$(3.68) \quad p(t; x, y) = m(y) \left[\sum_{n \in \Theta} e^{\lambda_n t} \phi_n(x) \phi_n(y) + \int_{\Xi} e^{\lambda t} \phi(x, \lambda) \phi(y, \lambda) \psi(\lambda) d\lambda \right].$$

Ejemplos

Movimiento Browniano en $[0, d]$ con bordes reflectantes

En este caso $\mathcal{S} = [0, d]$ y la ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = \frac{1}{2} \sigma^2 \frac{\partial^2 p}{\partial x^2}, \quad t > 0, \quad 0 < x < d, \quad y \in \mathcal{S},$$

con las condiciones de contorno de tipo Neumann

$$\left. \frac{\partial p}{\partial x} \right|_{x=0, d} = 0, \quad t > 0, y \in \mathcal{S}.$$

Buscamos autovalores λ y autofunciones ϕ del operador infinitesimal \mathcal{A} dado en este caso por

$$\mathcal{A} = \frac{1}{2} \sigma^2 \frac{d^2}{dx^2}.$$

En otras palabras

$$\begin{aligned} \frac{1}{2} \sigma^2 \phi''(x) &= \mathcal{A} \phi(x) \\ \phi'(0) &= \phi'(d) = 0. \end{aligned}$$

La solución general de esta ecuación diferencial viene expresada en términos de exponenciales. En efecto

$$\phi(x) = C_1 e^{\frac{\sqrt{2\lambda}x}{\sigma}} + C_2 e^{-\frac{\sqrt{2\lambda}x}{\sigma}}.$$

La condición de frontera $\phi'(0) = 0$ obliga a que $C_1 = C_2$, mientras que la segunda condición de frontera $\phi'(d) = 0$ obliga a que

$$e^{\frac{\sqrt{2\lambda}d}{\sigma}} - e^{-\frac{\sqrt{2\lambda}d}{\sigma}} = 0.$$

Los valores de λ para los que esta ecuación tiene solución pasan obligatoriamente por que sean negativos. En ese caso la expresión anterior se puede escribir en términos de la función seno, i.e.

$$\sin\left(\frac{\sqrt{-2\lambda}d}{\sigma}\right) = 0.$$

Con lo cual los ceros son múltiplos enteros de π , i.e.

$$\frac{\sqrt{-2\lambda_n}d}{\sigma} = n\pi, \quad n \in \mathbb{Z}.$$

Por lo tanto los autovalores vienen dados por

$$\lambda_n = -\frac{n^2\pi^2\sigma^2}{2d^2}, \quad n \geq 0.$$

Sustituyendo en $\phi(x)$, para cada n , las autofunciones vienen dadas por

$$\phi_n(x) = b_n \cos\left(\frac{n\pi x}{d}\right), \quad n \geq 0.$$

En este caso, y ya que $I(x_0, x) = 0$, la medida $m(x)$ en (3.58) se puede elegir (salvo constante normalizadora) como $m(x) = 1$. Para concluir normalizamos las autofunciones ϕ_n . Para ello, se tiene que

$$\int_0^d dx = d, \quad \int_0^d \cos^2\left(\frac{n\pi x}{d}\right) = \frac{d}{2}.$$

Con lo cual la constante normalizadora b_n , viene dada por

$$b_n = \sqrt{\frac{2}{d}}, \quad n \geq 1, \quad b_0 = \sqrt{\frac{1}{d}}.$$

Es bien sabido (series de Fourier) que el sistema de autofunciones $(\phi_n)_n$ es completo en $L_m^2([0, d])$. Por lo tanto, para $t > 0$ y $0 < x, y < d$ se tiene que

$$(3.69) \quad p(t; x, y) = \frac{1}{d} + \frac{2}{d} \sum_{n=1}^{\infty} \exp\left(-\frac{\sigma^2\pi^2 n^2}{2d^2}t\right) \cos\left(\frac{n\pi x}{d}\right) \cos\left(\frac{n\pi y}{d}\right).$$

Movimiento Browniano con un borde reflectante

En este caso podemos tomar $\mathcal{S} = [0, \infty)$ y la ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2}, \quad t > 0, \quad x > 0, \quad y \geq 0,$$

con las condiciones de contorno de tipo Neumann

$$\left. \frac{\partial p}{\partial x} \right|_{x=0} = 0, \quad t > 0, \quad y \geq 0.$$

Esta difusión es un caso límite de la anterior en (3.69) haciendo $d \rightarrow \infty$. En efecto,

$$\begin{aligned} p(t; x, y) &= \lim_{d \rightarrow \infty} \frac{1}{d} \sum_{n=-\infty}^{\infty} \exp\left(-\frac{\sigma^2\pi^2 n^2}{2d^2}t\right) \cos\left(\frac{n\pi x}{d}\right) \cos\left(\frac{n\pi y}{d}\right) \\ &= \int_{-\infty}^{\infty} \exp\left(-\frac{t\sigma^2\pi^2 u^2}{2}\right) \cos(\pi x u) \cos(\pi y u) du \quad (\text{integral de Riemann}) \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \exp\left(-\frac{t\sigma^2\pi^2 u^2}{2}\right) [\cos(\pi(x+y)u) + \cos(\pi(x-y)u)] du \quad (\text{fórmula trigonométrica}) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp\left(-\frac{t\sigma^2\xi^2}{2}\right) [e^{-i\xi(x+y)} e^{-i\xi(x-y)}] d\xi, \quad (\xi = \pi u \text{ e imparidad}) \end{aligned}$$

$$= \frac{1}{\sqrt{2\pi t\sigma^2}} \left(\exp\left(-\frac{(x+y)^2}{2t\sigma^2}\right) + \exp\left(-\frac{(x-y)^2}{2t\sigma^2}\right) \right),$$

donde la última igualdad se tiene por la fórmula de inversión de Fourier y el hecho de que $\exp(-t\sigma^2\xi^2/2)$ es la función característica de la distribución normal con media 0 y varianza $t\sigma^2$.

Movimiento Browniano con deriva μ en $[0, d]$ con bordes absorbentes

En este caso $\mathcal{S} = [0, d]$ y la ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 p}{\partial x^2} + \mu \frac{\partial p}{\partial x}, \quad t > 0, \quad 0 < x < d, \quad y \in \mathcal{S},$$

con las condiciones de contorno de tipo Dirichlet

$$\lim_{x \rightarrow 0^+} p(t; x, y) = 0, \quad \lim_{x \rightarrow d^-} p(t; x, y) = 0, \quad t > 0, \quad y \in \mathcal{S}.$$

Buscamos autovalores λ y autofunciones ϕ del operador infinitesimal \mathcal{A} dado en este caso por

$$\mathcal{A} = \frac{1}{2}\sigma^2 \frac{d^2}{dx^2} + \mu \frac{d}{dx}.$$

O equivalentemente

$$\begin{aligned} \frac{1}{2}\sigma^2 \phi''(x) + \mu \phi'(x) &= \lambda \phi(x) \\ \lim_{x \rightarrow 0^+} \phi(x) &= \lim_{x \rightarrow d^-} \phi(x) = 0. \end{aligned}$$

La solución general de esta ecuación diferencial viene expresada en términos de exponenciales.

$$\phi(x) = e^{-\frac{\mu x}{\sigma^2}} \left(C_1 e^{\frac{\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2} x} + C_2 e^{-\frac{\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2} x} \right).$$

La condición de frontera $\phi(0)$ obliga a que $C_1 = -C_2$, mientras que la segunda condición de frontera $\phi(d) = 0$ obliga a que

$$e^{\frac{\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2} d} - e^{-\frac{\sqrt{\mu^2 + 2\lambda\sigma^2}}{\sigma^2} d} = 0.$$

Los valores de λ para los que esta ecuación tiene solución pasan obligatoriamente por que el interior del radicando sea negativo. En ese caso la expresión anterior se puede escribir en términos de la función seno, i.e.

$$\sin\left(\frac{\sqrt{-\mu^2 - 2\lambda\sigma^2}}{\sigma^2} d\right) = 0.$$

Con lo cual los ceros son múltiplos enteros de π , i.e.

$$\frac{\sqrt{-\mu^2 - 2\lambda\sigma^2}}{\sigma^2} d = n\pi, \quad n \in \mathbb{Z}.$$

Por lo tanto los autovalores vienen dados por

$$\lambda_n = -\frac{n^2\pi^2\sigma^2}{2d^2} - \frac{\mu^2}{2\sigma^2}, \quad n \geq 1.$$

Sustituyendo en $\phi(x)$, para cada n , las autofunciones vienen dadas por

$$\phi_n(x) = b_n \exp\left(-\frac{\mu x}{\sigma^2}\right) \sin\left(\frac{n\pi x}{d}\right), \quad n \geq 1.$$

En este caso, la medida $m(x)$ en (3.58) se puede elegir (salvo constante normalizadora) como

$$m(x) = \exp\left(\frac{2\mu x}{\sigma^2}\right), \quad 0 \leq x \leq d.$$

Las constantes normalizadoras b_n vienen dadas, al igual que antes, por

$$b_n = \sqrt{\frac{2}{d}}, \quad n \geq 1.$$

Para $t > 0$ y $0 < x, y < d$, la expansión en autofunciones de p viene dada por

$$(3.70) \quad p(t; x, y) = \frac{2}{d} \exp\left(\frac{\mu(y-x)}{\sigma^2}\right) \exp\left(-\frac{\mu^2 t}{2\sigma^2}\right) \sum_{n=1}^{\infty} \exp\left(-\frac{\sigma^2\pi^2 n^2 t}{2d^2}\right) \sin\left(\frac{n\pi x}{d}\right) \sin\left(\frac{n\pi y}{d}\right).$$

Como los bordes son absorbentes faltaría calcular las probabilidades de absorción

$$p^0(t; x, \{0\}) = \mathbb{P}(X_t = 0 | X_0 = x), \quad p(t; x, \{d\}) = \mathbb{P}(X_t = d | X_0 = x), \quad 0 < x < d.$$

Como

$$p(t; x, \{0\}) + p(t; x, \{d\}) = 1 - \int_0^d p^0(t; x, y) dy,$$

se tiene que

$$p(t; x, \{0\}) + p(t; x, \{d\}) = 1 - \frac{2}{d} \exp\left(-\frac{\mu x}{\sigma^2}\right) \sum_{n=1}^{\infty} a_n \exp\left(-\frac{\sigma^2\pi^2 n^2 t}{2d^2}\right) \sin\left(\frac{n\pi x}{d}\right),$$

donde

$$a_n = \int_0^d \exp\left(\frac{\mu y}{\sigma^2}\right) \sin\left(\frac{n\pi y}{d}\right) dy = \frac{nd\pi\sigma^4}{\mu^2 d^2 + n^2\pi^2\sigma^4} \left(1 - (-1)^n \exp\left(\frac{d\mu}{\sigma^2}\right)\right).$$

Con lo cual basta calcular $p(t; x, \{0\})$ y a partir de lo anterior se podrá calcular $p(t; x, \{d\})$. Siguiendo (1.47) se tiene que

$$(3.71) \quad p(t; x, \{0\}) = v(x) - \int_0^d v(y) p^0(t; x, y) dy.$$

Movimiento Browniano con deriva μ en $[0, \infty)$ con borde absorbente

Al igual que antes basta hacer límite en (3.70) a medida que $d \rightarrow \infty$, en cuyo caso se tiene que

$$\begin{aligned}
p^0(t; x, y) &= 2 \exp\left(\frac{\mu(y-x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}\right) \int_0^\infty \exp\left(-\frac{\pi^2 \sigma^2 t u^2}{2}\right) \operatorname{sen}(\pi x u) \operatorname{sen}(\pi y u) du \\
&= \frac{2}{\pi} \exp\left(\frac{\mu(y-x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}\right) \int_0^\infty \exp\left(-\frac{\sigma^2 t \xi^2}{2}\right) \operatorname{sen}(\xi x) \operatorname{sen}(\xi y) d\xi \\
&= \frac{1}{\pi} \exp\left(\frac{\mu(y-x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}\right) \int_0^\infty \exp\left(-\frac{\sigma^2 t \xi^2}{2}\right) [\cos(\xi(x-y)) - \cos(\xi(x+y))] d\xi \\
&= \frac{1}{2\pi} \exp\left(\frac{\mu(y-x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}\right) \int_0^\infty \exp\left(-\frac{\sigma^2 t \xi^2}{2}\right) [e^{-i\xi(x-y)} - e^{-i\xi(x+y)}] d\xi \\
&\quad \exp\left(\frac{\mu(y-x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2 t}} \left[\exp\left(-\frac{(x-y)^2}{2\sigma^2}\right) - \exp\left(-\frac{(x+y)^2}{2\sigma^2 t}\right) \right],
\end{aligned}$$

donde al igual que antes la última igualdad se tiene por la fórmula de inversión de Fourier y el hecho de que $\exp(-t\sigma^2\xi^2/2)$ es la función característica de la distribución normal con media 0 y varianza $t\sigma^2$.

Movimiento Browniano estándar

En este ejemplo el espectro es continuo. Tenemos que buscar autofunciones que cumplan

$$\frac{1}{2}\phi''(x) = \hat{\lambda}\phi(x),$$

con la condición de que cada $\phi(x)$ esté acotada para todo $x \in \mathbb{R}$. La condición de acotada obliga a que el autovalor $\hat{\lambda}$ sea negativo. Llamemos $\hat{\lambda} = -\lambda^2/2$. Entonces buscamos autofunciones $\phi(x)$ tal que

$$\phi''(x) = -\lambda^2\phi(x),$$

con las mismas condiciones iniciales. La solución general de esta ecuación diferencial viene dada por

$$\phi(x, \lambda) = C_1 \operatorname{sen}(\lambda x) + C_2 \operatorname{cos}(\lambda x), \quad \lambda, x \in \mathbb{R},$$

que siempre va a estar acotada. Debido a que $I(x_0, x) = 0$, se puede escoger la medida como $m(x) = 1$. En el espacio $L^2(\mathbb{R})$ (ahora con el producto interno $(f, g) = \int_{\mathbb{R}} f(x)g(\bar{x})dx$) se tiene que para cada $\lambda \in \mathbb{R}$, las autofunciones se pueden escribir como

$$\phi(x, \lambda) = e^{ix\lambda}.$$

La ortogonalidad implica que

$$\int_{\mathbb{R}} \phi(x, \lambda)\phi(\bar{x}, \mu)dx = \int_{\mathbb{R}} e^{ix(\lambda-\mu)}dx = 2\pi\delta(\lambda - \mu).$$

Por lo tanto la representación espectral de la densidad sería

$$p(t; x, y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\lambda^2 t/2} e^{ix\lambda} e^{-iy\lambda} d\lambda,$$

que es la transformada inversa de Fourier del núcleo gaussiano. Por lo tanto

$$p(t; x, y) = \frac{1}{\sqrt{2\pi}} e^{-(x-y)^2/2t}.$$

Proceso de Ornstein-Uhlenbeck

En este caso se tiene que los coeficiente de deriva y difusión vienen dados por

$$\mu(x) = -x, \quad \sigma^2(x) = 1.$$

El espacio de estados del proceso se interpreta como la velocidad de una partícula que sigue un movimiento Browniano, pero existe una fuerza de atracción proporcional a la distancia al origen.

Con $x_0 = 0$, se tiene que $I(0, x) = -x^2$ (3.57). Por lo tanto la medida $m(x)$ se puede elegir (salvo constante normalizadora) como

$$m(x) = e^{-x^2}.$$

Así, el espacio de estados donde se está moviendo este proceso sería $\mathcal{S} = (-\infty, \infty)$. La ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} - x \frac{\partial p}{\partial x}, \quad t > 0.$$

Buscamos entonces autofunciones y autovalores discretos del operador infinitesimal

$$\mathcal{A} = \frac{1}{2} \frac{d^2}{dx^2} - x \frac{d}{dx}.$$

O equivalentemente

$$\frac{1}{2} \phi''(x) - x \phi'(x) = \lambda \phi(x), \quad -\infty < x < \infty,$$

de tal manera que $\int_{-\infty}^{\infty} e^{-x^2} [\phi(x)]^2 dx < \infty$. Esta ecuación diferencial es identificable ya que corresponde a la de los polinomios de Hermite $(H_n)_n$ con autovalor

$$\lambda_n = -n, \quad n \geq 0.$$

Las normas de $H_n(x)$ vienen dadas por (2.39). Por lo tanto, la representación espectral de la densidad $p(t; x, y)$ se pueden escribir como

$$p(t; x, y) = e^{-y^2} \sum_{n=0}^{\infty} e^{-nt} \frac{H_n(x)H_n(y)}{2^n n! \sqrt{\pi}} = \frac{e^{-y^2}}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{H_n(x)H_n(y)}{n!} \left(\frac{e^{-t}}{2}\right)^n$$

usando el hecho de que

$$\sum_{n=0}^{\infty} \frac{H_n(x)H_n(y)}{n!} \left(\frac{u}{2}\right)^n = \frac{1}{\sqrt{1-u^2}} \exp\left(\frac{2uxy}{1+u} - \frac{u^2(x-y)^2}{1-u^2}\right).$$

Así

$$\begin{aligned} p(t; x, y) &= \frac{e^{-y^2}}{\sqrt{\pi}\sqrt{1-e^{-2t}}} \exp\left(\frac{2e^{-t}xy}{1+e^{-t}} - \frac{e^{-2t}(x-y)^2}{1-e^{-2t}}\right) \\ &= \frac{e^{-y^2}}{\sqrt{\pi}\sqrt{1-e^{-2t}}} \exp\left(-\frac{e^{-2t}(x^2+y^2)}{1-e^{-2t}}\right) \exp\left(\frac{2xye^{-t}}{1-e^{-2t}}\right). \end{aligned}$$

Modelo de crecimiento poblacional

En este caso se tiene que los coeficientes de deriva y difusión vienen dados por

$$\mu(x) = bx + c, \quad \sigma^2(x) = 2ax, \quad 0 < x < \infty,$$

donde a, b, c son constantes con $a > 0$. Con $x_0 = 1$, se tiene que

$$I(1, x) = \int_1^x \frac{bz + c}{az} dz = \frac{b}{z}(x-1) + \frac{c}{a} \log x.$$

Por lo tanto, la medida $m(x)$ (3.58) se puede escoger como

$$m(x) = x^\alpha e^{bx/a}, \quad \alpha = \frac{c}{a} - 1, \quad x > 0.$$

La ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = ax \frac{\partial^2 p}{\partial x^2} + (bx + c) \frac{\partial p}{\partial x}, \quad t > 0.$$

Buscamos entonces autofunciones y autovalores discretos del operador infinitesimal

$$\mathcal{A} = ax \frac{d^2}{dx^2} + (bx + c) \frac{d}{dx}.$$

En otras palabras

$$ax\phi''(x) + (bx + c)\phi'(x) = \lambda\phi(x), \quad x > 0,$$

de tal manera que $\int_0^\infty x^\alpha e^{bx/a} [\phi(x)]^2 dx < \infty$. Esta ecuación diferencial de vuelta es identificable ya que corresponde a la de los polinomios de Laguerre $(L_n^{(\alpha)})_n$. En concreto para

$$L_n^{(\alpha)} \left(-\frac{bx}{a} \right).$$

Como $a > 0$, para que estos polinomios sean integrables tiene que ocurrir que $b < 0$ (por la exponencial) y $c > 0$ (por el monomio). El correspondiente autovalor es

$$\lambda_n = bn, \quad n \geq 0.$$

Usando las normas de $L_n^\alpha(x)$ se puede llegar a la representación espectral de la densidad $p(t; x, y)$:

$$\begin{aligned} p(t; x, y) &= \frac{(|b|/a)^{\alpha+1} y^\alpha e^{by/a}}{\Gamma(\alpha+1)} \sum_{n=0}^{\infty} e^{bnt} \frac{n!}{\Gamma(n+\alpha+1)} L_n^{(\alpha)} \left(-\frac{bx}{a} \right) L_n^{(\alpha)} \left(-\frac{by}{a} \right) \\ &= \frac{(|b|/a)^{\alpha+1} y^\alpha e^{by/a}}{\Gamma(\alpha+1)} \frac{\exp \left(\frac{be^{bt}(x+y)}{a(1-e^{bt})} \right) I_\alpha \left(\frac{2|b|\sqrt{xye^{bt}}}{a(1-e^{bt})} \right)}{\left(\frac{b^2}{a^2} xy e^{bt} \right)^{\alpha/2} (1-e^{bt})} \\ &= \frac{|b|e^{by/a} e^{-bt\alpha/2}}{a\Gamma(\alpha+1)(1-e^{bt})} \left(\frac{y}{x} \right)^{\alpha/2} \exp \left(\frac{be^{bt}(x+y)}{a(1-e^{bt})} \right) I_\alpha \left(\frac{2|b|\sqrt{xye^{bt}}}{a(1-e^{bt})} \right). \end{aligned}$$

Donde $I_\alpha(z)$ es la función de Bessel.

Modelo de Wright-Fisher

Este modelo tiene como coeficientes de deriva y difusión los siguientes

$$\mu(x) = \gamma_2(1-x) - \gamma_1x, \quad \sigma^2(x) = x(1-x), \quad 0 < x < 1, \quad \gamma_1, \gamma_2 > 0.$$

El espacio de estados es $\mathcal{S} = [0, 1]$. Los coeficientes γ_1, γ_2 se interpretan como intensidades de mutación de una población A en otra B (y viceversa) que están compitiendo entre ellas.

Este ejemplo se identifica con los polinomios de Jacobi, pero normalizándolos de tal manera que estén definidos en el intervalo $[0, 1]$. En este caso las intensidades de mutación γ_1, γ_2 se identifican con los parámetros $\alpha, \beta > -1$ de la siguiente forma

$$\gamma_1 = \frac{\beta + 1}{2}, \quad \gamma_2 = \frac{\alpha + 1}{2}.$$

La función $I_{x_0, x}$ se puede calcular explícitamente:

$$I(x_0, x) = \int_{x_0}^x \frac{2\gamma_2(1-z) - 2\gamma_1z}{z(1-z)} dz = \int_{x_0}^x \left(\frac{2\gamma_2}{z} - \frac{2\gamma_1}{1-z} \right) dz = \log[x^{2\gamma_2}(1-x)^{2\gamma_1}] + C.$$

Por lo tanto, la medida $m(x)$ (3.58) se puede escoger como

$$m(x) = x^{2\gamma_2-1}(1-x)^{2\gamma_1-1}, \quad 0 < x < 1.$$

La ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = \frac{1}{2}x(1-x)\frac{\partial^2 p}{\partial x^2} + [\gamma_2(1-x) - \gamma_1x]\frac{\partial p}{\partial x}, \quad t > 0.$$

Buscamos entonces autofunciones y autovalores discretos del operador infinitesimal

$$\mathcal{A} = \frac{1}{2}x(1-x)\frac{d^2}{dx^2} + [\gamma_2(1-x) - \gamma_1x]\frac{d}{dx}.$$

En otras palabras

$$\frac{1}{2}x(1-x)\phi''(x) + [\gamma_2(1-x) - \gamma_1x]\phi'(x) = \lambda\phi(x), \quad 0 < x < 1,$$

de tal manera que se cumplan las condiciones iniciales

$$\frac{1}{2}x^{2\gamma_2}\frac{d\phi}{dx}\Big|_{x=0^+} = 0, \quad \frac{1}{2}(1-x)^{2\gamma_1}\frac{d\phi}{dx}\Big|_{x=1^-} = 0.$$

La solución de la ecuación diferencial se puede expresar en términos de los polinomios de Jacobi $(P_n(\alpha, \beta)(y))_n$ en $[-1, 1]$ mediante el cambio de variable

$$x = \frac{1+y}{2}.$$

El correspondiente autovalor es

$$\lambda_n = \frac{1}{2}n(n + \alpha + \beta + 1), \quad n \geq 0.$$

Se tiene entonces que la representación espectral de la densidad $p(t; x, y)$ viene dada por

$$p(t; x, y) = \frac{y^{2\gamma_2-1}(1-y)^{2\gamma_1}}{B(2\gamma_2, 2\gamma_1)} \sum_{n=0}^{\infty} e^{-n(n+2\gamma_1+2\gamma_2-1)t/2} P_n^{(2\gamma_1-1, 2\gamma_2-1)}(2x-1) P_n^{(2\gamma_1-1, 2\gamma_2-1)}(2y-1) \times \\ \times \frac{B(2\gamma_2, 2\gamma_1)(2n+2\gamma_1+2\gamma_2-1)n!\Gamma(n+2\gamma_1+2\gamma_2-1)}{\Gamma(n+2\gamma_2)\Gamma(n+2\gamma_1)},$$

donde $B(x, y)$ es la función Beta. Al contrario de lo que ocurre en el proceso anterior, no se sabe si existe una expresión explícita de esta serie en términos de funciones elementales.

Proceso de Jacobi

Escogemos los coeficientes de deriva y difusión como

$$\mu(x) = \frac{1}{2}[(\beta+1)(1-x) - (\alpha+1)(1+x)], \quad \sigma^2(x) = 1-x^2, \quad -1 < x < 1,$$

donde α, β son constantes. Con $x_0 = 0$, se tiene que

$$I(0, x) = \int_0^x \frac{(\beta+1)(1-z) - (\alpha+1)(1+z)}{(1-z)(1+z)} dz \\ = \int_0^x \left(\frac{\beta+1}{1+z} - \frac{\alpha+1}{1-z} \right) dz = \log[(1+x)^{\beta+1}(1-x)^{\alpha+1}].$$

Por lo tanto la medida $m(x)$ en (3.58) se puede escoger como

$$m(x) = (1-x)^\alpha(1+x)^\beta, \quad -1 < x < 1.$$

Las condiciones de integrabilidad en $-1 < x < 1$ obligan a que $\alpha, \beta > -1$. La ecuación de retroceso de Kolmogorov para la densidad de probabilidades de transición $p(t; x, y)$ viene dada por

$$\frac{\partial p}{\partial t} = (1-x^2) \frac{\partial^2 p}{\partial x^2} + \frac{1}{2}[(\beta+1)(1-x) - (\alpha+1)(1+x)] \frac{\partial p}{\partial x}, \quad t > 0.$$

Buscamos autofunciones y autovalores discretos del operador infinitesimal

$$\mathcal{A} = (1-x^2) \frac{d^2}{dx^2} + \frac{1}{2}[(\beta+1)(1-x) - (\alpha+1)(1+x)] \frac{d}{dx}.$$

En otras palabras

$$(1-x^2)\phi''(x) + \frac{1}{2}[(\beta+1)(1-x) - (\alpha+1)(1+x)]\phi'(x) = \lambda\phi(x), \quad -1 < x < 1,$$

de tal manera que se cumplan las condiciones iniciales

$$\frac{1}{2}(1-x)^{\alpha+1} \frac{d\phi}{dx} \Big|_{x=1-} = 0, \quad \frac{1}{2}(1+x)^{\beta+1} \frac{d\phi}{dx} \Big|_{x=-1+} = 0.$$

La solución de la ecuación diferencial es la de los polinomios de Jacobi $(P_n^{(\alpha, \beta)})_n$. El correspondiente autovalor, es

$$\lambda_n = -\frac{1}{2}n(n+\alpha+\beta+1), \quad n \geq 0.$$

Normalizando los polinomios de tal manera que $P_n^{(\alpha, \beta)}(1) = 1$ se tiene que la representación espectral de la densidad $p(t; x, y)$ viene dada por

$$p(t; x, y) = \frac{(1-y)^\alpha (1+y)^\beta}{2^{\alpha+\beta+1} \Gamma(\alpha+1)^2} \sum_{n=0}^{\infty} e^{-n(n+\alpha+\beta+1)t/2} P_n^{(\alpha, \beta)}(x) P_n^{(\alpha, \beta)}(y) \times \\ \times \frac{(2n + \alpha + \beta + 1) \Gamma(n + \alpha + 1) \Gamma(\alpha + \beta + 1)}{n! \Gamma(n + \beta + 1)}.$$

Una vez más, no se sabe si existe una expresión explícita de esta serie en términos de funciones elementales.

Proceso radial de Movimiento Browniano absorbente en la superficie de la bola unitaria

Sean $\{B_i(t), t \geq 0\}, i = 1, \dots, N$, movimientos Brownianos estándar independientes. Se considera el proceso radial

$$Z_t = B_1^2(t) + \dots + B_N^2(t).$$

Se puede comprobar que Z_t es un proceso de difusión con coeficientes de deriva y difusión dados por

$$\mu_Z(x) = N, \quad \sigma_Z^2(x) = 4x, \quad x > 0.$$

El proceso de Bessel se define como la distancia euclídeana de un movimiento Browniano N -dimensional, i.e.,

$$Y_t = \sqrt{Z_t} = \sqrt{B_1^2(t) + \dots + B_N^2(t)}.$$

Se puede demostrar también que los coeficientes de deriva y de difusión para el proceso de Bessel vienen dados por

$$\mu(x) = \frac{N-1}{2x}, \quad \sigma^2(x) = 1, \quad x > 0.$$

Para $N = 1$ recuperamos el movimiento Browniano reflejado $|B_t|$. Con $x_0 = 1$, se tiene que

$$I(1, x) = \int_1^x \frac{N-1}{z} dz = (N-1) \log x.$$

Por lo tanto la medida $m(x)$ en (3.58) se puede escoger como

$$m(x) = x^{N-1}, \quad x > 0.$$

Consideramos ahora el proceso de Bessel en el espacio de estados $(0, 1]$ donde hay un extremo absorbente en $x = 1$ (en el espacio N -dimensional, el espacio de estados correspondería al interior de la bola de radio unidad incluyendo la superficie esférica).

La ecuación de retroceso de Kolmogorov sería

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} + \frac{N-1}{2x} \frac{\partial p}{\partial x} = \frac{1}{2} \frac{1}{x^{N-1}} \frac{\partial}{\partial x} \left(x^{N-1} \frac{\partial p}{\partial x} \right), \quad t > 0,$$

con condiciones iniciales

$$p(t; 1, y) = 0, \quad x^{N-1} \frac{\partial p}{\partial x} \Big|_{x=0^+} = 0.$$

Por lo tanto buscamos autofunciones de \mathcal{A} tal que

$$\frac{1}{2} \phi''(x) + \frac{N-1}{2x} \phi'(x) = \lambda \phi(x), \quad 0 \leq x \leq 1,$$

condicionado a que

$$\phi(1) = 0, \quad x^{N-1} \frac{d\phi}{dx} \Big|_{x=0^+} = 0.$$

Las autofunciones de esta ecuación diferencial con valores iniciales se identifican con modificaciones de las funciones de Bessel $J_\alpha(x)$:

$$\phi_n(x) = x^{-(N-2)/2} J_{(N-2)/2}(x\sqrt{\lambda_n}), \quad n \geq 0,$$

donde $\sqrt{\lambda_n}$ es la sucesión de ceros positivos de $J_{(N-2)/2}(x)$. Los autovalores vienen dados por $\lambda = -2\lambda_n$.

Así, la representación espectral de $p(t; x, y)$ se puede escribir como

$$\begin{aligned} p(t; x, y) &= y^{N-1} \sum_{n=0}^{\infty} e^{-2\lambda_n t} \phi_n(x) \phi_n(y) \pi_n \\ &= y \left(\frac{y}{x}\right)^{\frac{N-2}{2}} \sum_{n=0}^{\infty} e^{-2\lambda_n t} J_{\frac{N-2}{2}}(x\sqrt{\lambda_n}) J_{\frac{N-2}{2}}(y\sqrt{\lambda_n}) \pi_n, \end{aligned}$$

donde $\pi_n^{-1} = \int_0^1 \phi_n^2(x) x^{N-1} dx = \int_0^1 x J_{(N-2)/2}^2(x\sqrt{\lambda_n}) dx$.

Haciendo tender el borde al infinito nos da una representación espectral de la densidad del movimiento Browniano radial en todo \mathbb{R}^N . En este caso, el espectro será continuo y la representación espectral de la densidad $p(t; x, y)$ vendría dada por

$$p(t; x, y) = \left[\frac{\Gamma((N+1)/2)}{\Gamma(N/2)} \right]^2 y^{N-1} \int_0^\infty e^{-\lambda^2 t/2} J_{\frac{N-2}{2}}(\lambda x) J_{\frac{N-2}{2}}(\lambda y) \lambda^{N-1} d\lambda.$$

Proceso de Orstein-Uhlenbeck radial

Tomemos el proceso de Bessel pero imponiéndole una fuerza atractiva proporcional a la distancia al origen. Los coeficientes de deriva y de difusión están dados por

$$\mu(x) = \frac{N-1}{2x} - x, \quad \sigma^2(x) = 1, \quad x > 0.$$

Con $x_0 = 1$, se tiene que

$$\int_1^x \left(\frac{N-1}{2z} - z \right) dz = (N-1) \log x - \frac{x^2}{2} + \frac{1}{2}.$$

Por lo tanto la medida $m(x)$ se puede escoger como

$$m(x) = x^{N-1} e^{-x^2}, \quad x > 0.$$

La ecuación de retroceso de Kolmogorov sería

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2} + \left(\frac{N-1}{2x} - x \right) \frac{\partial p}{\partial x}, \quad t > 0,$$

con condiciones iniciales

$$x^{N-1} \frac{\partial p}{\partial x} \Big|_{x=0^+} = 0.$$

Por lo tanto buscamos autofunciones de \mathcal{A} tal que

$$\frac{1}{2} \phi''(x) + \left(\frac{N-1}{2x} - x \right) \phi'(x) = \lambda \phi(x), \quad x > 0.$$

Las autofunciones se pueden escribir en términos de polinomios de Laguerre pero evaluados en $y = x^2$ y con parámetro

$$\alpha = \frac{N-2}{2}.$$

Por lo tanto, la representación espectral de $p(t; x, y)$ se puede escribir como

$$\begin{aligned} p(t; x, y) &= 2y^{N-1} e^{-y^2} \sum_{n=0}^{\infty} e^{-2nt} \frac{n!}{\Gamma(n + \alpha + 1)} L_n^{(\alpha)}(x^2) L_n^{(\alpha)}(y^2) \\ &= \frac{2y^{N-1} e^{-y^2}}{(xye^{-t})^\alpha (1 - e^{-2t})} \exp\left(-\frac{e^{-2t}(x^2 + y^2)}{1 - e^{-2t}}\right) I_\alpha\left(\frac{2xye^{-t}}{1 - e^{-2t}}\right), \end{aligned}$$

donde para llegar al resultado se utilizó la fórmula del núcleo de Poisson para los polinomios de Laguerre y que $I_\alpha(z) = i^{-\alpha} J_\alpha(iz)$, donde de nuevo J_α es la función de Bessel estándar.

Bibliografía

- [1] ÁLVAREZ, RENATO (2003) , *Polinomios hipergeométricos y q-polinomios*, Prensas Universitarias de Zaragoza, Zaragoza.
- [2] BAVINCK, H. (1998), *Differential operators having Laguerre type and Sobolev type Laguerre orthogonal polynomials as eigenfunction: A survey in special function and differential equations, proceedings workshop*, Allied Publishers, New Delhi.
- [3] BEALS, RICHARD y WONG, RODERICK (2010) , *Special Functions, A Graduate Text*, Cambridge University Press, New York.
- [4] BERTOIN, J., *Lévy Processes*, Cambridge University Press, Cambridge.
- [5] CHESTER, C., FRIEDMAN, B. y URSELL, F. (1957), *An extension of the method of steepest descents*, Proc. Cambridge Philos. Soc. 53, 599-611.
- [6] CHIHARA, THEODORE S. (1978), *An Introduction to Orthogonal Polynomials*, Gordon and Breach, New York.
- [7] EVANS, LAWRENCE C. (2013), *Introduction to Stochastic Differential Equations*, versión 1.2, Departamento de Matemáticas, UC Berkeley, Library of Congress Cataloging-in-Publication Data.
- [8] GRÜNBAUM, F. ALBERTO (2010), *An Urn Model Associated with Jacobi Polynomials*, Communications in Applied Mathematics and Computational Science, vol. 5, no.1.
- [9] JIN, X. y WONG, R. (1998), *Asymptotic formulas for the Meixner polynomials*, J. Approx. Theory 14, 281-300.
- [10] KARLIN, SAMUEL y MCGREGOR, JAMES L. (1957), *The Differential Equations of Birth-and-Death Processes, and the Stieltjes Moment Problem*, Transactions of the American Mathematical Society, American Mathematical Society, vol. 85, no. 2, 489-546.
- [11] KARLIN, SAMUEL y MCGREGOR, JAMES L. (1959), *Random Walks*, Volumen 3, no.1, Illinois Journal of Mathematics, 66-81.
- [12] KARLIN, SAMUEL y TAYLOR, HOWARD M. (1974), *A First Course in Stochastic Processes*, segunda edición, Academic Press, New York.
- [13] KIJIMA, M., NAIR, M. G., POLLET, P. K. y VAN DOORN, E. A. (1997), *Limiting conditional distributions for birth-death processes*, Adv. Appl. Prob. 29, 185-204.
- [14] KOEKOEK, ROELOF y SWARTTOUW, RENÉ F. (1998), *The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue*, Delft University of Technology, Reporte no. 98-17.

- [15] LITTLEJOHN, LANCE (1986), *An application of a new theorem on orthogonal polynomials and differential equations*, Quaestiones Mathematicae, 10(1), 49-61.
- [16] RINCÓN, LUIS (2011), *Curso intermedio de probabilidad*, primera edición, facultad de ciencias de la UNAM, cd. de México.
- [17] SCHOUTENS, WIM y TEUGELS, JOZEF L. (1998), *Lévy processes, polynomials and martingales*, Communications in Statistics, Stochastic Models 14(1), 335-349.
- [18] SCHRIJNER, P. (1995), *Quasi-Stationarity of Discrete-Time Markov Chains*, tesis doctoral, Universiteit Twente, Enschede.
- [19] SCHRIJNER, P. y VAN DOORN, E. A. (1997), *Weak convergence of conditioned birth-death processes in discrete time*, J. Appl. Prob. 34, 46-53.
- [20] SCHOUTENS, WIM (2000), *Stochastic Processes and Orthogonal Polynomials*, Lecture notes in statistics, v. 146, Springer-Verlag.
- [21] SHOHAT, J. A. y TAMARKIN, J. D. (1943), *The problem of moments*, Mathematical Surveys, vol. 1.
- [22] WONG, R. (2001), *Asymptotic Approximations of Integrals*, Academic Press, Boston, SIAM, Philadelphia.