



UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO

FACULTAD DE CIENCIAS

Distribuciones de momentos tipo fase y
matriz exponencial.

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
MATEMÁTICO

PRESENTA:

JOSE RODRIGO INIESTA MIRANDA

TUTOR:

DR. MOGENS BLADT PETERSEN

2015

CIUDAD UNIVERSITARIA, D.F.





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.

Hoja de Datos del Jurado

1. Datos del alumno
Iniesta
Miranda
Jose Rodrigo
58 49 14 74
Universidad Nacional Autónoma de México
Facultad de Ciencias
Matemáticas
308273360
2. Datos del tutor
Dr.
Bladt
Petersen
Mogens
3. Datos del sinodal 1
Dr.
Baltazar
Larios
Fernando
4. Datos del sinodal 2
Dr.
Mena
Chávez
Ramsés Humberto
5. Datos del sinodal 3
Dr.
Pérez
Garmendia
José Luis Ángel
6. Datos del sinodal 4
Act.
Santoyo
Cano
Alejandro
7. Datos del trabajo escrito
Distribuciones de momentos tipo fase y matriz exponencial.
81 p
2015

Dedicatoria y agradecimientos.

Dedico este esfuerzo con amor a mi familia; a Andrea y Wendy que me motivan a crecer, a mi mamá María de Jesús a mis hermanos Diego y Vanesa que me dan su apoyo incondicional y sabios consejos, a mi abuelito Elfego por sus enseñanzas, a mis tíos y primas Elfego, Marcela, Gabriela, Tania y Cecilia con quienes puedo contar y me dan su apoyo.

También dedico este trabajo a mis amigos por su apoyo y compañía.

Un agradecimiento especial para mi mamá María y para Medarda por su gran y valiosa ayuda.

Agradezco al Dr. Mogens Bladt por las enseñanzas.

Gracias a la Facultad de Ciencias, UNAM.

Contenido

Introducción	1
1 Procesos de Markov	3
1.1 Cadenas de Markov	3
1.2 Procesos de Markov de saltos	14
1.3 Cadenas y procesos de Markov revertidos en el tiempo	25
2 Distribuciones tipo fase y matriz exponencial	29
2.1 Distribuciones tipo fase	29
2.2 Distribuciones matriz exponencial	38
2.2.1 Transformada de Laplace racional	38
2.2.2 Distribución matriz exponencial	46
3 Teoría de renovación	49
3.1 Proceso de renovación puro	49
3.2 Procesos de renovación estacionarios y retrasados	54
3.3 Teorema de renovación de Blackwell	56
3.4 Teorema principal de renovación	59
3.5 Proceso de edad y residual de vida	60
3.6 Teoría de renovación tipo fase	62
4 Distribuciones de momentos tipo fase y matriz exponencial	65
4.1 Distribuciones de momentos matriz exponencial	65
4.2 Distribuciones de momentos tipo fase	68
4.3 Índice de Gini y curva de Lorenz para distribuciones matriz exponencial	73
Apéndice A	75
A.1 Algunas propiedades de matrices y exponencial de una matriz	75
A.2 Notación de Kronecker.	79

Introducción

En este trabajo se consideran distribuciones de momentos de orden natural cuya distribución principal es tipo fase o matriz exponencial. El objetivo principal es demostrar que si la distribución subyacente es tipo fase entonces las correspondientes distribuciones de momentos son tipo fase y si la distribución subyacente es matriz exponencial entonces las correspondientes distribuciones de momentos son matriz exponencial.

La importancia de las distribuciones tipo fase y matriz exponencial se debe a que las distribuciones tipo fase son "densas" en la clase de distribuciones con soporte positivo y las matriz exponencial son un clase más grande de distribuciones que contiene a las distribuciones tipo fase de hecho toda distribución tipo fase es matriz exponencial pero no al revés.

Se busca encontrar representaciones de las distribuciones de momentos que sean accesibles y que además pertenezcan a la misma distribución ya que si se tiene una representación para las distribuciones matriz exponencial entonces esta sería también una representación para las tipo fase sin embargo no podemos asegurar que esta representación pertenezca a las distribuciones tipo fase.

Las distribuciones tipo fase tienen propiedades fáciles de demostrar gracias a que se caracterizan por ser el tiempo de absorción en un proceso de Markov de saltos con algún número natural de estados transitorio y un estado absorbente, de hecho la mayoría de propiedades de esta distribución se pueden demostrar usando su cadena de Markov asociada.

Las distribuciones de momentos son importantes en varios campos aplicados de las matemáticas, en inglés son comúnmente referidas como "*length biased*" o "*size-biased sampling*".

Este trabajo se divide en cuatro capítulos. El primer capítulo trata sobre cadenas y procesos de Markov de saltos estos procesos son necesarios para desarrollar la teoría de distribuciones tipo fase, también se desarrolla lo necesario para considerar a las cadenas y procesos de Markov de saltos con revertidos en el tiempo esto será necesario para encontrar dos representaciones distintas de las distribuciones de momentos para distribuciones tipo fase.

En el segundo capítulo se desarrollan las distribuciones tipo fase y matriz exponencial. Las distribuciones tipo fase son caracterizadas por un proceso de Markov de saltos y las distribuciones matriz exponencial se caracterizan por la forma de su densidad y de su transformada de Laplace, de esta distribución se encontrara una representación de orden minimal que es la representación que se utilizara en el último capítulo.

En el tercer capítulo se desarrolla la teoría de renovación pues para demostrar que la

distribución del primer momento de una distribución tipo fase es tipo fase se utilizaran argumentos pertenecientes a esta teoría.

Finalmente en el cuarto capítulo se demuestra lo comentado anteriormente y se da un pequeño ejemplo teórico en el que se utiliza la distribución del primer momento matriz exponencial para encontrar dos fórmulas de la curva de Lorenz y del índice de Gini.

La curva de Lorenz es una representación gráfica utilizada normalmente para ver que tan concentrados están los ingresos en una población y el índice de Gini para medir la concentración de ingresos o el grado de desigualdad social.

Capítulo 1

Procesos de Markov

Pensemos que en los procesos estocásticos el índice del proceso es tiempo entonces los procesos de Markov son los procesos estocásticos en que el tiempo futuro depende únicamente del tiempo presente, esto es lo que se estudiara en este capítulo que sera fundamental para los siguientes capítulos. Lo referente a cadenas de Markov y a procesos de Markov de saltos fue extraído de [1],[2],[9] y la parte de reversibilidad de [2] y de [9].

1.1 Cadenas de Markov

Las cadenas de Markov son procesos de Markov cuyo tiempo es discreto. Sea $(X_n)_{n \in \mathbb{N}}$ un proceso estocástico que toma valores en un conjunto discreto E ; llamaremos a E espacio de estados.

Definición 1.1.1. (P.M) *El proceso estocástico $X = (X_n)_{n \in \mathbb{N}}$ cumple la propiedad de Markov si*

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \dots, X_0 = i_0) = P(X_{n+1} = j | X_n = i)$$

para todo $n \in \mathbb{N}$ y $j, i, i_{n-1}, \dots, i_0 \in E$.

Si un proceso estocástico X cumple lo anterior es llamado cadena de Markov, además X es llamado cadena de Markov homogénea si no depende de n , es decir, si $P(X_{n+1} = j | X_n = i)$ es igual para toda $n \geq 0$ en cuyo caso escribiremos $p_{i,j} = P(X_n = j | X_{n-1} = i)$ y nos referimos a $p_{i,j}$ como probabilidad de transición en un paso del estado i al estado j . La matriz de transición de una cadena de Markov $(X_n)_{n \geq 0}$ se define como

$$\mathbf{P} = (p_{i,j})_{i,j \in E}$$

Como se verá en el siguiente teorema la distribución de las cadenas de Markov es completamente conocida dada su matriz de transición y su distribución inicial la cual denotamos como π_i , es decir, $\pi_{i_0} = P(X_0 = i_0)$. A partir de ahora estudiaremos cadenas de Markov homogéneas.

Teorema 1.1.1. Sean $i_0, i_1, \dots, i_n \in E$ y $n \in \mathbb{N}$ entonces X es una cadena de Markov si y solo si

$$P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n) = p_{i_n, i_{n-1}} p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0} \pi_{i_0}$$

Demostración. \Rightarrow)

$$\begin{aligned} & P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i_n) \\ &= P(X_n = i_n | X_0 = i_0, \dots, X_{n-1} = i_{n-1}) P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \\ &= P(X_n = i_n | X_{n-1} = i_{n-1}) P(X_0 = i_0, \dots, X_{n-1} = i_{n-1}) \text{(prop. Markov)} \\ &= p_{i_n, i_{n-1}} P(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}) \\ &\vdots \\ &= p_{i_n, i_{n-1}} p_{i_{n-1}, i_{n-2}} \cdots p_{i_1, i_0} \pi_{i_0} \text{(repetiendo los pasos anteriores } n \text{ veces)} \end{aligned}$$

\Leftarrow) es suficiente con escribir la definición de probabilidad condicional y utilizar la hipótesis. \square

Veamos que la propiedad de Markov implica que la probabilidad en el tiempo futuro únicamente depende del último tiempo conocido para lograr esto demostraremos antes un teorema.

Teorema 1.1.2. Sea $n, m \in \mathbb{N}$ y $i_n, i_{n-1}, \dots, i_m, \dots, i_0 \in E$ tales que $0 < m < n$ entonces

$$\begin{aligned} P(X_n = i_n, \dots, X_m = i_m | X_{m-1} = i_{m-1}, \dots, X_0 = i_0) = \\ P(X_n = i_n, \dots, X_m = i_m | X_{m-1} = i_{m-1}) \end{aligned}$$

Demostración. Utilizando el teorema 1.1.1 tenemos lo siguiente

$$\begin{aligned} & P(X_n = i_n, \dots, X_m = i_m | X_{m-1} = i_{m-1}, \dots, X_0 = i_0) \\ &= \frac{P(X_n = i_n, \dots, X_m = i_m, X_{m-1} = i_{m-1}, \dots, X_0 = i_0)}{P(X_{m-1} = i_{m-1}, \dots, X_0 = i_0)} \\ &= \frac{p_{i_n, i_{n-1}} \cdots p_{i_m, i_{m-1}} \cdots p_{i_1, i_0} \pi_{i_0}}{p_{i_m, i_{m-1}} \cdots p_{i_1, i_0} \pi_{i_0}} \\ &= p_{i_n, i_{n-1}} \cdots p_{i_{m+1}, i_m} \\ &= P(X_n = i_n, \dots, X_m = i_m | X_{m-1} = i_{m-1}) \end{aligned}$$

\square

Corolario 1.1.1. Sea $i_{n_k} \in E$ con $k \in 1, 2, \dots, m$ y $n_0, n_1, \dots, n_m \in \mathbb{N}$ tales que $n_0 < n_1 < \dots < n_m$ entonces

$$\begin{aligned} P(X_{n_m} = i_{n_m} | X_{n_{m-1}} = i_{n_{m-1}}, X_{n_{m-2}} = i_{n_{m-2}}, \dots, X_{n_0} = i_{n_0}) \\ = P(X_{n_m} = i_{n_m} | X_{n_{m-1}} = i_{n_{m-1}}) \end{aligned}$$

Demostración. Basta con incluir los índices que hacen falta en la probabilidad utilizar la propiedad de Markov y luego sumar sobre los índices que sobren. \square

Se define la probabilidad de transición en n pasos como $p_{i,j}^{(n)} = P(X_n = j | X_0 = i)$ y la matriz de transición en n pasos como $\mathbf{P}^{(n)} = (p_{i,j}^{(n)})_{i,j \in E}$, en el siguiente teorema se demuestra que en realidad la matriz de transición en n pasos de la cadena es igual a la n -ésima potencia de la matriz de transición en un paso.

Teorema 1.1.3 (Ecuación de Chapman-Kolmogorov). *Para todo $n, m \in \mathbb{N}$ y $i, j \in E$ se cumple*

$$p_{i,j}^{(n+m)} = \sum_{k \in E} p_{i,k}^{(n)} p_{k,j}^{(m)}$$

Demostración. Sean $n, m \in \mathbb{N}$ y $i, j, k \in E$

$$\begin{aligned} & P(X_{n+m} = j | X_0 = i) \\ &= \sum_{k \in E} P(X_{n+m} = j, X_n = k | X_0 = i) \\ &= \sum_{k \in E} P(X_{n+m} = j | X_n = k, X_0 = i) P(X_n = k | X_0 = i) \\ &= \sum_{k \in E} P(X_{n+m} = j | X_n = k) P(X_n = k | X_0 = i) \text{ (Propiedad de Markov)} \end{aligned}$$

\square

Por el teorema anterior tenemos lo siguiente

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{n+m} = \mathbf{P}^n \mathbf{P}^m = \mathbf{P}^{(n)} \mathbf{P}^{(m)}.$$

Sea $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ la sigma álgebra generada por el proceso hasta el tiempo n y $\mathcal{F}_\infty = \sigma(X_0, X_1, X_2, \dots)$ la sigma álgebra generada por todo el proceso.

Definición 1.1.2. *Un tiempo de paro τ para la cadena de Markov $X = (X_n)_{n \in \mathbb{N}}$ es una variable aleatoria que toma valores en $\mathbb{N} \cup \{\infty\}$ tal que $\{\tau = n\} \in \mathcal{F}_n$ para toda n .*

La definición anterior nos dice que los tiempos de paro son variables aleatorias (tiempos aleatorios) sobre las cuales se puede tomar la decisión de si ya han ocurrido o no, sabiendo cómo se ha comportado el proceso.

Definición 1.1.3. *Sea $A \in E$ un subconjunto del espacio de estados, se define*

$$T_A = \inf\{n \geq 0 | X_n \in A\}$$

T_A es el tiempo en que el proceso llega al conjunto A .

Es claro que T_A es un tiempo de paro, si $A = \{i\}$ con $i \in E$ y se escribirá T_i

Teorema 1.1.4 (P.F.M). $X = (X_n)_{n \in \mathbb{N}}$ cumple para toda $n \in \mathbb{N}$ y tiempo de paro $\tau < \infty$

$$P(X_{\tau+n} = j | X_\tau = i) = p_{i,j}^n$$

llamada propiedad fuerte de Markov.

Demostración.

$$\begin{aligned} & P(X_{\tau+n} = j | X_\tau = i) \\ &= \sum_{k \in \mathbb{N}} P(X_{\tau+n} = j | X_\tau = i, \tau = k) P(\tau = k) \\ &= \sum_{k \in \mathbb{N}} P(X_{k+n} = j | X_k = i) P(\tau = k) \\ &= \sum_{k \in \mathbb{N}} P(X_n = j | X_0 = i) P(\tau = k) \text{ (homogeneidad)} \\ &= P(X_n = j | X_0 = i) \sum_{k \in \mathbb{N}} P(\tau = k) \\ &= P(X_n = j | X_0 = i) \end{aligned}$$

□

Lo anterior significa que la propiedad de Markov sigue siendo válida para tiempos aleatorios.

Definición 1.1.4. *Definimos*

$$\rho_{i,j} = P_i(T_j < \infty) = P(T_j < \infty | X_0 = i)$$

la probabilidad de que el proceso inicie en el estado i y llegue al estado j .

En particular $\rho_{j,j}$ es la probabilidad de regresar a j dado que el proceso inicio en j .

Definición 1.1.5. *Decimos que un estado j es absorbente si $\rho_{j,j} = 1$, recurrente si $\rho_{j,j} = 1$ y transitorio si $\rho_{j,j} < 1$*

Esto significa que si j es un estado absorbente y en algún momento el proceso esta en el estado j a partir de ese momento el proceso siempre estará en el estado j , si j es un estado recurrente y el proceso inicia en j en algún momento el proceso volverá a estar en el estado j y si j es un estado transitorio y el proceso inicia en el estado j con probabilidad $1 - \rho_{j,j}$ nunca volverá a estar en ese estado.

Usaremos la siguiente notación para simplificar la escritura

$$P(\cdot | X_0 = i) = P_i(\cdot), E(\cdot | X_0 = i) = E_i(\cdot).$$

Definición 1.1.6. *El número de visitas al estado i es*

$$N_i = \sum_{j=1}^{\infty} 1_{\{X_j=i\}}$$

Teorema 1.1.5. *Sea i algún estado de X cadena de Markov. Entonces las siguientes son equivalentes:*

1. i es recurrente
2. $N_i = \infty$ en casi todo punto
3. $E_i(N_i) = \sum_{m=1}^{\infty} p_{ii}^{(m)} = \infty$

Demostración. Sea

$$T_i^n = \inf\{n > T_i^{n-1} | X_n = i\}, T_i^1 = T_i$$

los tiempo entre visitas al estado i . Entonces

$$\begin{aligned} P_i(T_i^{k+1} < \infty) &= P_i(T_i^{k+1} < \infty, T_i^k < \infty) \\ &= E_i(P_i(T_i^{k+1} < \infty, T_i^k < \infty | \mathcal{F}_{T_i^k})) \text{ (suavizamiento)} \\ &= E_i(I\{T_i^k < \infty\} P_i(T_i^{k+1} < \infty | \mathcal{F}_{T_i^k})) \text{ (medible)} \\ &= E_i(I\{T_i^k < \infty\} P_{X_{T_i^k}}(T_i^{k+1} < \infty)) \text{ (P.F.M)} \\ &= P_i(T_i^{k+1} < \infty) E_i(I\{T_i^k < \infty\} | X_{T_i^k} = i) \\ &= P_i(T_i^{k+1} < \infty) P_i(T_i^k < \infty) \\ &\vdots \\ &= P_i(T_i^1 < \infty)^{k+1} \end{aligned}$$

Si i es recurrente entonces $P_i(T_i^{k+1} < \infty) = 1$ para toda k , es decir $T_i^{k+1} < \infty$ en casi todo punto. Como

$$N_i = \sup\{k \in \mathbb{N} | T_i^k < \infty\}$$

se tiene que $N_i = \infty$ en casi todo punto. Ahora por el teorema de convergencia monótona

$$\begin{aligned} E_i(N_i) &= E_i\left(\sum_{j=1}^{\infty} 1_{\{X_j=i\}}\right) \\ &= \sum_{j=1}^{\infty} E_i(1_{\{X_j=i\}}) \\ &= \sum_{j=1}^{\infty} P_i(X_j = i) \end{aligned}$$

Es claro que $E_i(N_i) = \infty$ por la segunda equivalencia.

Para la implicación que falta supongamos que i es transitorio. Entonces $P_i(T_i < \infty) < 1$

y

$$\begin{aligned}
E_i(N_i) &= \sum_{j=0}^{\infty} P_i(N_i > j) \\
&= \sum_{j=0}^{\infty} P_i(T_i^j < \infty) \\
&= \sum_{j=0}^{\infty} P_i(T_i < \infty)^j \\
&< \infty
\end{aligned}$$

□

Corolario 1.1.2. *Sea i algún estado de X_n . Entonces las siguientes son equivalentes:*

1. i es transitorio
2. $N_i < \infty$ en casi todo punto
3. $E_i(N_i) = \sum_{m=1}^{\infty} p_{ii}^{(m)} < \infty$

Definición 1.1.7. *El estado i comunica con el estado j si existe $m \in \mathbb{N}$ tal que $p_{ij}^{(m)} > 0$ y lo denotamos por $i \rightarrow j$. Dos estados i y j se comunican si $i \rightarrow j$ y $j \rightarrow i$ y lo denotaremos por $i \sim j$.*

La relación \sim es de equivalencia en el espacio de estados, es decir, $i \sim i$, $i \sim j \Leftrightarrow j \sim i$ y si $i \sim j, j \sim k$ entonces $i \sim k$. Por lo anterior el espacio de estados se puede descomponer en clases de equivalencia. Veamos algunas propiedades.

Teorema 1.1.6. *Si i es recurrente y $i \sim j$ entonces j es recurrente.*

Demostración. Sean m_1, m_2 tales que $p_{ij}^{(m_1)} > 0$ y $p_{ji}^{(m_2)} > 0$, entonces

$$E_j(N_j) = \sum_{n=0}^{\infty} p_{jj}^{(n)} \geq \sum_{n=0}^{\infty} p_{ji}^{(m_2)} p_{ii}^{(n)} p_{ij}^{(m_1)} = \infty$$

la desigualdad se da escogiendo un término de la ecuación de Chapman-Kolmogorov, se concluye que j es recurrente por el teorema 1.1.5. □

Por el teorema anterior si una clase tiene algún estado recurrente entonces todos sus estados son recurrentes, análogamente si una clase tiene un estado transitorio todos sus estados son transitorios de lo contrario los estados serian recurrentes.

Entonces podemos dividir al espacio de estados como la unión de las clases de equivalencia inducidas por \sim y el conjunto de todos los estados transitorios.

Definición 1.1.8. *Una cadena de Markov se llama irreducible si todos sus estados se comunican.*

Si una cadena es irreducible entonces todos sus estados son recurrentes o transitorios por lo que las llamaremos cadena irreducible recurrente o transitoria respectivamente.

Definición 1.1.9. Un vector renglón $\nu = (\nu_j)_{j \in E}$ es llamado medida estacionaria o invariante de la cadena de Markov $(X_n)_{n \geq 0}$ si ν es finita, no cero, no negativa y

$$\nu P = \nu$$

La condición anterior implica que si ν es distribución de X_n entonces es distribución de X_{n+1} , es decir, para toda $n \in \mathbb{N}$

$$\nu = \nu P^n$$

Teorema 1.1.7. Si i es recurrente entonces podemos definir una medida estacionaria ν como sigue

$$\nu_j = E_i \left(\sum_{n=0}^{T_i-1} I\{X_n = j\} \right)$$

ν_j es el número de visitas al estado j entre dos visitas al estado i .

Demostración.

$$\begin{aligned} \nu_j &= E_i \left(\sum_{n=0}^{T_i-1} I\{X_n = j\} \right) \\ &= E_i \left(\sum_{n=1}^{T_i} I\{X_n = j\} \right) (X_{T_i} = X_0 = i) \\ &= E_i \left(\sum_{n=1}^{\infty} I\{X_n = j, T_i \geq n\} \right) \\ &= \sum_{n=1}^{\infty} E_i (I\{X_n = j, T_i \geq n\}) \text{ (T.C.M)} \\ &= \sum_{n=1}^{\infty} P_i(X_n = j, T_i > n-1) \\ &= \sum_{n=1}^{\infty} E_i(P_i(X_n = j, T_i > n-1 | \mathcal{F}_{n-1})) \\ &= \sum_{k \in E} p_{kj} \sum_{n=1}^{\infty} P_i(T_i > n-1, X_{n-1} = k) \\ &= \sum_{k \in E} p_{kj} \nu_k \end{aligned}$$

pues

$$\begin{aligned}
& E_i(P_i(X_n = j, T_i > n - 1 | \mathcal{F}_{n-1})) \\
&= E_i(I\{T_i > n - 1\}P_i(X_n = j | \mathcal{F}_{n-1})) \\
&= E_i(I\{T_i > n - 1\}p_{X_{n-1}j}) \\
&= E_i\left(\sum_{k \in E} I\{T_i > n - 1, X_{n-1} = k\}p_{X_{n-1}j}\right) \\
&= \sum_{k \in E} p_{kj}P_i(T_i > n - 1, X_{n-1} = k) \text{ (T.C.M)}
\end{aligned}$$

La primer igualdad se da porque $I\{T_i > n - 1\}$ es medible en \mathcal{F}_{n-1} . Entonces se cumple $\nu P = \nu$, si j no esta en la clase de i entonces $\nu_j = 0$, si j esta en la misma clase que i existe m tal que $p_{ji}^{(m)} > 0$ y $\nu = \nu P = \dots = \nu P^m$ entonces

$$\infty > 1 = \nu_i = \sum_{k \in E} \nu_k p_{ki}^{(m)} \geq \nu_j p_{ji}^{(m)}$$

que implica $\nu_j < \infty$, por lo anterior ν es una medida estacionaria. \square

Corolario 1.1.3. Si una cadena de Markov es irreducible y recurrente entonces para toda medida invariante ν se tiene $\nu_i > 0$ para toda $i \in E$

Demostración. Como $\nu \neq 0$ existe j tal que $\nu_j > 0$ y por ser una cadena recurrente existe m tal que $p_{ji}^{(m)} > 0$ entonces

$$\nu_i = \sum_{k \in E} \nu_k p_{ki}^{(m)} \geq \nu_j p_{ji}^{(m)} > 0$$

\square

Definimos $\nu^{(i)}$ como la medida estacionaria dada por

$$\nu_j^{(i)} = E_i\left(\sum_{n=0}^{T_i-1} I\{X_n = j\}\right)$$

El índice superior (i) indica la dependencia de la elección del estado recurrente i . Si consideramos $j = i$ solo tenemos $I\{X_0 = j\} = 1$ por lo que $\nu_i^{(i)} = 1$. Llamamos medida estacionaria canónica a $\nu^{(i)}$.

Corolario 1.1.4. $\nu_j^{(i)} = \sum_{n=0}^{\infty} P_i(X_n = j, T_i > n)$

Demostración.

$$\begin{aligned}
\nu_j &= E_i \left(\sum_{n=0}^{T_i-1} I\{X_n = j\} \right) \\
&= \sum_{n=1}^{\infty} E_i (I\{X_n = j, T_i \geq n\}) \quad (T.C.M) \\
&= \sum_{n=1}^{\infty} P_i(X_n = j, T_i > n-1)
\end{aligned}$$

□

Lema 1.1.1. *Sea ν una medida estacionaria con $\nu_i = 1$ entonces $\nu = \nu^{(i)}$*

Demostración. Por el corolario anterior

$$\nu_j^{(i)} = \sum_{n=0}^{\infty} P_i(X_n = j, T_i > n)$$

esto implica que si $n \rightarrow \infty$ entonces $P_i(X_n = j, T_i > n) \rightarrow 0$. $P_k(X_n = j, T_i > n)$ es la probabilidad de ir de k a j sin pasar por el estado i . Supongamos $j \neq k$, para $n = 1$ la probabilidad anterior es la probabilidad de transición en un paso. Para $n = 2$

$$P_k(X_n = j, T_i > n) = \sum_{m \neq i} p_{km} p_{mj}.$$

Definimos \hat{P} como la matriz de transición \mathbf{P} de la cadena pero reemplazando la i -ésima columna con ceros. Entonces

$$P_k(X_n = j, T_i > n) = \sum_{m \in E} \hat{p}_{km} \hat{p}_{mj}.$$

Con un proceso inductivo

$$(P_k(X_n = j, T_i > n))_{j,k \in E} = \hat{P}^n$$

es decir $P_k(X_n = j, T_i > n)$ es el kj -ésimo elemento de \hat{P}^n que también se puede escribir como $P_k(X_n = j, T_i > n) = e_j \hat{P}^n e'_k$ donde e_j es un vector renglón con la misma dimensión que el espacio de estados con ceros en las entradas excepto en la j -ésima entrada que llevara un uno, entonces

$$\nu^{(i)} = (\nu_j^{(i)})_{j \in E} = e_i \sum_{n=0}^{\infty} \hat{P}^n$$

Como ν es estacionaria con $\nu_i = 1$ tenemos que

$$\nu_j = \delta_{ij} + e_i \sum_{n=0}^{\infty} \hat{P}^n$$

Como ν es estacionaria con $\nu_i = 1$

$$\nu_j = \delta_{ij} + (\nu\hat{P})_j$$

pues $(\nu\hat{P})_i = 0$ y si $j \neq i$ entonces $(\nu\hat{P})_j = \nu_j$ por ser estacionaria, en consecuencia

$$\begin{aligned} \nu &= e_i + \nu\hat{P} \\ &= e_i + (e_i + \nu\hat{P})\hat{P} \\ &= e_i(I + \hat{P}) + \nu\hat{P}^2 \\ &\vdots \\ &= e_i \sum_{n=0}^m \hat{P}^n + \nu\hat{P}^{m+1} \end{aligned}$$

Como $\hat{P}^{m+1} \rightarrow 0$ si $m \rightarrow \infty$,

$$\lim_{m \rightarrow \infty} \nu = e_i \sum_{n=0}^{\infty} \hat{P}^n = \nu^{(i)}$$

□

Corolario 1.1.5. *Si una cadena de Markov es irreducible y recurrente entonces existe una medida estacionaria. Todas las medidas estacionarias son proporcionales.*

Demostración. Si ν es estacionaria con $\nu_i = c$ entonces $\nu_i^{(i)} = \nu_i/c$ □

Definición 1.1.10. *Un estado recurrente i es recurrente positivo si $E_i(T_i) < \infty$ y recurrente nulo si $E_i(T_i) = \infty$*

Corolario 1.1.6. *Si una cadena de Markov es irreducible y recurrente entonces todos sus estados son recurrentes positivos o recurrentes nulos.*

Demostración. Todas las medidas invariantes son proporcionales a $\nu^{(j)}$, $j \in E$ además

$$\begin{aligned} |\nu^{(j)}| &= \sum_{k \in E} \nu_k^{(j)} \\ &= \sum_{k \in E} \sum_{n=0}^{\infty} P_j(X_n = k, T_j > n) \\ &= \sum_{n=0}^{\infty} \sum_{k \in E} P_j(X_n = k, T_j > n) \\ &= \sum_{n=0}^{\infty} P_j(T_j > n) \\ &= E_j(T_j) \end{aligned}$$

así que $|\nu^{(j)}|$ es finita o infinita y por lo tanto toda la cadena. □

De lo anterior se tiene que la recurrencia nula o positiva es una propiedad de clase. Para cadenas de Markov irreducibles y recurrentes existe una medida estacionaria salvo multiplicación por constantes. Ahora podemos encontrar distribuciones estacionarias en el caso de que se pueda normalizar una medida invariante.

Corolario 1.1.7. *Si una cadena de Markov es irreducible y recurrente positiva existe una única distribución estacionaria dada por*

$$\pi_j = \frac{E_i \left(\sum_{n=0}^{T_i-1} I\{X_n = j\} \right)}{E_i(T_i)}$$

Demostración. Como la cadena es recurrente $|\nu^{(i)}| = E_i(T_i) < \infty$ entonces

$$\pi = \frac{\nu^{(i)}}{|\nu^{(i)}|}$$

es distribución. □

Definición 1.1.11. *El periodo de un estado i es el numero natural más grande $d(i)$ que cumple*

$$P_i(T_i \in L_{d(i)}) = 1$$

donde $L_d = \{d, 2d, 3d, \dots\}$. Si el periodo es uno el estado es llamado aperiódico.

Teorema 1.1.8. *La periodicidad es una propiedad de clase. Si i, j están la misma clase recurrente entonces tienen el mismo periodo.*

Demostración. Sea i un estado recurrente con periodo $d(i)$ y se j otro estado en la misma clase. Como i y j se comunican existen $n, m > 0$ tales que $p_{ij}^n > 0$ y $p_{ji}^m > 0$ así que

$$p_{ii}^{n+m} = \sum_k p_{ik}^n p_{ki}^m \geq p_{ij}^n p_{ji}^m > 0$$

entonces $n + m \in L_{d(i)}$ ahora tomando r tal que $p_{jj}^r > 0$

$$p_{ii}^{n+m+r} \geq p_{ij}^n p_{jj}^r p_{ji}^m > 0$$

entonces $n + m + r \in L_{d(i)}$ que implica $r \in L_{d(i)}$ y $d(i) \leq d(j)$ análogamente se tiene que $d(j) \leq d(i)$. □

Teorema 1.1.9. *Si p_{ij} son las probabilidades de transición de una cadena de Markov aperiódica entonces existe N tal que $p_{ii}^n > 0$ para toda $n \geq N$.*

Demostración. Sea $C = \{n \in \mathbb{N} | p_{ij}^n > 0\}$ entonces existe n tal que $n, n + 1 \in C$ de lo contrario el periodo seria mayor o igual a dos. Todo entero mayor o igual que $n(n + 1)$ puede escribirse como combinación lineal de n y $n + 1$ pues en $n(n + 1) = n + n + \dots + n$ son $n + 1$ sumandos por lo que se puede incrementar de n a $n + 1$ uno por uno entonces si $m \geq n(n + 1)$ se tiene que $m = qn + r(n + 1)$ y

$$p_{ii}^m \geq (p_{ii}^n)^q (p_{ii}^{n+1})^r > 0$$

□

Corolario 1.1.8. Si $\{X_n\}_{n \in \mathbb{N}}$ es irreducible y aperiódica entonces para todo $i, j \in E$ existe N tal que $p_{ij}^n > 0$ para toda $n \geq N$.

Demostración. Elegimos k tal que $p_{ij}^k > 0$ y r tal que $p_{ii}^r > 0$ para todo $m \geq r$ entonces haciendo $n = m + r$ tenemos el resultado. \square

1.2 Procesos de Markov de saltos

Los procesos de Markov son procesos estocásticos en los que el tiempo presente solo depende del tiempo pasado y cuyo tiempo es continuo. Sea $X = (X_t)_{t \geq 0}$ ($t \in \mathbb{R}$) un proceso estocástico que toma valores en un conjunto discreto E llamado espacio de estados.

Definición 1.2.1. (P.M) El proceso $X = (X_t)_{t \geq 0}$ ($t \in \mathbb{R}$) satisface la propiedad de Markov si

$$P(X_{t_m} = j_{t_m} | X_{t_{m-1}} = j_{t_{m-1}}, \dots, X_{t_0} = j_{t_0}) = P(X_{t_m} = j_{t_m} | X_{t_{m-1}} = j_{t_{m-1}})$$

para $t_m > t_{m-1} > \dots > t_0 > 0$ y $j_{t_m}, j_{t_{m-1}}, \dots, j_0 \in E$

Si un proceso X cumple lo anterior es llamado proceso de Markov de saltos además es llamado proceso de Markov de saltos homogéneo si se cumple $P(X_{t+h} = j | X_t = i) = P(X_h = j | X_0 = i)$ y en este caso escribiremos $P(X_h = j | X_0 = i) = p_{i,j}^{(h)}$, es decir, si el proceso es homogéneo entonces no depende del tiempo en que está el proceso solo depende de la diferencia entre el tiempo presente y el último tiempo conocido; a partir de ahora trabajaremos con procesos de Markov de saltos homogéneos. Llamamos a $p_{i,j}^{(h)}$ probabilidades de transición y se define a la matriz de transición como

$$\mathbf{P}^{(h)} = (p_{i,j}^{(h)})_{i,j \in E}$$

Supondremos que $p_{i,j}^{(0)} = \delta_{i,j}$ es la delta de Kronecker (aunque esto no siempre sucede) por lo anterior $\mathbf{P}^{(0)} = \mathbf{I}$ es la matriz identidad.

Teorema 1.2.1. (Ecuación de Chapman-Kolmogorov) Para todo $t, s \geq 0$ el proceso X cumple la ecuación de Chapman-Kolmogorov

$$p_{i,j}^{(t+s)} = \sum_{k \in E} p_{i,k}^{(t)} p_{k,j}^{(s)}$$

Demostración.

$$\begin{aligned}
& P(X_{t+s} = j | X_0 = i) \\
&= \sum_{k \in E} P(X_{t+s} = j, X_t = k | X_0 = i) \\
&= \sum_{k \in E} P(X_{t+s} = j | X_t = k, X_0 = i) P(X_t = k | X_0 = i) \\
&= \sum_{k \in E} P(X_{t+s} = j | X_t = k) P(X_t = k | X_0 = i) \\
&= \sum_{k \in E} P(X_s = j | X_0 = k) P(X_t = k | X_0 = i) \text{ (homogeneidad)}
\end{aligned}$$

□

Lo anterior es equivalente a

$$\mathbf{P}^{(t+s)} = \mathbf{P}^{(t)} \mathbf{P}^{(s)}$$

Sea $\mathcal{F}_t = \sigma(X_s : s \leq t)$ la sigma álgebra generada por el proceso hasta el tiempo t .

Definición 1.2.2. *Un tiempo de paro τ para un proceso de Markov de saltos $X = (X_t)_{t \geq 0}$ ($t \in \mathbb{R}$) es una variable aleatoria no negativa que cumple $\{\tau \leq t\} \in \mathcal{F}_t$*

Un tiempo de paro τ al igual que en el caso discreto es una variables aleatoria sobre la cual se puede verificar la ocurrencia del evento $\{\tau \leq t\}$ a partir del proceso hasta el tiempo t . La sigma álgebra \mathcal{F}_τ generada por τ consiste de los conjuntos

$$A \cap \{\tau \leq t\} \in \mathcal{F}_t$$

tales que A es un conjunto medible en \mathcal{F}_t .

Teorema 1.2.2. *(P.F.M) Los proceso de Markov de saltos $X = (X_t)_{t \geq 0}$ ($t \in \mathbb{R}$) cumplen para todo tiempo de paro $\tau < \infty$ y $s \geq 0$*

$$P(X_{\tau+s} = i_s | \mathcal{F}_\tau) = P(X_s = i_s | X_0 = X_\tau)$$

la anterior es llamada propiedad fuerte de Markov.

Sean $S_0 = 0 < S_1 < S_2 < \dots$ los tiempos en que el proceso cambia de estado. Las diferencias $T_n = S_{n+1} - S_n$ son los tiempos entre saltos también llamados tiempos de interarribo, en el caso de que S_n sea el último estado visitado definimos $T_m = \infty$ para todo $m \geq n$ con $m \in \mathbb{N}$. Veamos que los tiempos de interarribo siguen una distribución exponencial.

Teorema 1.2.3. *Para cualquier $i \in E$ y $t \geq 0$ se tiene que $P(T_n > t | X_{S_n} = i) = e^{-\lambda t}$*

Demostración. Primero denotemos por $f(t) = P(T_n > t | X_{S_n} = i)$ y notemos que por homogeneidad $f(t) = P_i(T_0 > t)$, sean $t, s \geq 0$ entonces:

$$\begin{aligned}
f(t+s) &= P_i(T_0 > t+s) \\
&= P_i(T_0 > t+s, T_0 > t) \\
&= P_i(T_0 > t+s | T_0 > t) P_i(T_0 > t) \\
&= P_i(X_r = i, \forall r \in [0, t+s] | X_r = i, \forall r \in [0, t]) P_i(T_0 > t) \\
&= P_i(X_r = i, \forall r \in [0, t+s] | X_t = i) P_i(T_0 > t) \text{ (P.M)} \\
&= P_i(X_r = i, \forall r \in [0, s] | X_0 = i) P_i(T_0 > t) \text{ (homogeneidad)} \\
&= P_i(T_0 > s) P_i(T_0 > t) \\
&= f(s)f(t)
\end{aligned}$$

Además de lo anterior la función esta entre 0 y 1 por lo que $f(t) = e^{-ct}$ para alguna $c > 0$, denotaremos a c como λ_i pues depende del estado i \square

Definición 1.2.3. Un estado $i \in E$ es llamado absorbente si $\lambda_i = 0$, estable si $0 < \lambda_i < \infty$ e instantáneo si $\lambda_i = \infty$.

Lo anterior significa que si un estado es absorbente el tiempo de interarribo desde la llegada a ese estado es infinito con probabilidad uno en otras palabras si en algún momento el proceso está en un estado absorbente se quedara ahí por siempre, si un estado es estable el proceso se quedara ahí por un tiempo infinito y si el estado es instantáneo inmediatamente cambia de estado una vez que el proceso llegue a este.

Sea $Y_n = X_{S_n}$ la sucesión de estados visitados; en el caso de que S_n sea el último estado visitado definimos $Y_m = X_{S_n}$ para todo $m \geq n$ con $m \in \mathbb{N}$.

Veamos que el proceso de Markov de saltos X se puede reconstruir en términos de $(Y_n, S_n)_{n \in \mathbb{N}}$.

Teorema 1.2.4. Sean $n \in \mathbb{N}, j \in E$ y $t > 0$ entonces

$$P(Y_{n+1} = j, T_n > t | Y_0, \dots, Y_{n-1}, Y_n = i, S_0, \dots, S_n) = q_{i,j} e^{-\lambda_i t}$$

Demostración. S_n es un tiempo de paro pues si el tiempo del n -ésimo cambio de estado es menor a t entonces dado el proceso hasta el tiempo t se conoce el tiempo de este

cambio, es decir, $\{S_n \leq t\} \in \mathcal{F}_t$.

$$\begin{aligned}
& P(Y_{n+1} = j, T_n > t | Y_0, \dots, Y_n = i, S_0, \dots, S_n) \\
&= P(X_{S_{n+1}} = j, X_r = X_{S_n}, \forall r \in [S_n, S_n + t] | X_u, u \leq S_n) \\
&= P(X_{S_{n+1}} = j, X_r = X_{S_n}, \forall r \in [S_n, S_n + t] | X_{S_n} = i) \text{ (P.M)} \\
&= P(X_{S_0+1} = j, X_r = X_{S_0}, \forall r \in [S_0, S_0 + t] | X_{S_0} = i) \text{ (P.F.M)} \\
&= P(Y_1 = j, T_0 > t | X_0 = i) \\
&= P_i(Y_1 = j, T_0 > t) \\
&= \mathbb{E}_i(P_i(Y_1 = j, T_0 > t | \mathcal{F}_{S_1})) \text{ (esperanza iterada, suavizamiento)} \\
&= \mathbb{E}_i(1_{Y_1=j} P_i(T_0 > t | \mathcal{F}_{S_1})) \text{ (medibilidad)} \\
&= \mathbb{E}_i(1_{Y_1=j} P_i(T_0 > t)) \text{ (P.M)} \\
&= \mathbb{E}_i(1_{Y_1=j}) P_i(T_0 > t) \\
&= P_i(Y_1 = j) e^{-\lambda_i t} \text{ (teo 1.2.3)} \\
&= q_{i,j} e^{-\lambda_i t}
\end{aligned}$$

Donde $q_{i,j} = P_i(Y_1 = j)$ □

Teorema 1.2.5. *Existen números $\lambda_i = \lambda(i) \geq 0$ y una matriz de transición \mathbf{Q} tales que*

$$P_i(Y_k = i_k, T_{k-1} > t_k, k = 1, 2, \dots, n) = \prod_{k=1}^n q_{i_{k-1}i_k} e^{\lambda_{i_{k-1}} t_k}$$

Demostración. La prueba será por inducción, la base está en la demostración del teorema (1.2.4) ahora supongamos que es válido para $n - 1$ entonces:

$$\begin{aligned}
& P_i(Y_k = i_k, T_{k-1} > t_k, k = 1, 2, \dots, n) \\
&= P_i(Y_n = i_n, T_{n-1} > t_n, Y_k = i_k, T_{k-1} > t_k, k = 1, 2, \dots, n - 1) \\
&= \mathbb{E}_i(P_i(Y_n = i_n, T_{n-1} > t_n, Y_k = i_k, T_{k-1} > t_k, k = 1, 2, \dots, n - 1 | \mathcal{F}_{S_{n-1}})) \\
&= \mathbb{E}_i(1_{Y_k=i_k, T_{k-1}>t_k, k=1,2,\dots,n-1} P_i(Y_n = i_n, T_{n-1} > t_n | \mathcal{F}_{S_{n-1}})) \text{ (medibilidad)} \\
&= \mathbb{E}_i(1_{Y_k=i_k, T_{k-1}>t_k, k=1,2,\dots,n-1} P_i(Y_n = i_n, T_{n-1} > t_n | Y_{n-1} = i_{n-1})) \text{ (P.M)} \\
&= \mathbb{E}_i(1_{Y_k=i_k, T_{k-1}>t_k, k=1,2,\dots,n-1}) P_{i_{n-1}}(Y_1 = i_n, T_0 > t_n) \\
&= P_i(Y_k = i_k, T_{k-1} > t_k, k = 1, 2, \dots, n - 1) P_{i_{n-1}}(Y_1 = i_n, T_0 > t_n) \\
&= q_{i_{n-1}i_n} e^{\lambda_{i_{n-1}} t_n} \prod_{k=1}^{n-1} q_{i_{k-1}i_k} e^{\lambda_{i_{k-1}} t_k} \text{ (hipótesis)}
\end{aligned}$$

□

De lo anterior se ve que la sucesión $(Y_n)_{n \in \mathbb{N}}$ es una cadena de Markov que llamaremos **cadena asociada**.

Corolario 1.2.1. *La sucesión $(Y_n)_{n \in \mathbb{N}}$ es una cadena de Markov, los tiempos de interarribo son condicionalmente independientes dada la sucesión de estados visitados y tienen distribución exponencial, es decir,*

$$P(T_0 > t_0, \dots, T_{n-1} > t_{n-1} | Y_0, \dots, Y_n) = \prod_{k=1}^{n-1} e^{\lambda_{i_k} t_k}$$

Demostración. De (1.2.4) tomando $t = 0$ y por (1.1.1) $(Y_n)_{n \in \mathbb{N}}$ es una cadena de Markov

$$\begin{aligned} P(T_0 > t_0, \dots, T_{n-1} > t_{n-1} | Y_0, \dots, Y_n) &= \\ &= \frac{P(T_0 > t_0, \dots, T_{n-1} > t_{n-1}, Y_0, \dots, Y_n)}{P(Y_0, \dots, Y_n)} \\ &= \frac{\prod_{k=1}^n q_{i_{k-1}i_k} e^{\lambda_{i_{k-1}} t_{k-1}}}{\prod_{k=1}^n q_{i_{k-1}i_k}} \text{ (cadena de markov y teo 1.2.5)} \\ &= \prod_{k=1}^{n-1} e^{\lambda_{i_k} t_k} \end{aligned}$$

□

Vimos que si X es un proceso de Markov de saltos entonces existen probabilidades $q_{i,j}$ y algunos números $\lambda_i \geq 0$ que cumplen lo que hemos visto; ahora veamos el resultado opuesto para ello supondremos una cadena de Markov Y con matriz de transición \mathbf{Q} y $\lambda_i > 0$; notemos que por la naturaleza del proceso Y se necesita $q_{i,i} = 0$.

De aquí en adelante supondremos que el proceso de Markov de saltos **no tiene estado instantáneos**. Veamos una expresión para las probabilidades de transición $p_{i,j}^{(t)}$

Proposición 1.2.1. *Para cualquier $i, j \in E$ y $t > 0$,*

$$p_{i,j}^{(t)} = e^{-\lambda_i t} \delta_{i,j} + \int_0^t \lambda_i e^{-\lambda_i s} \sum_{k \neq i} q_{i,k} p_{k,j}^{(t-s)} ds$$

Demostración. Es claro que si i es absorbente se sigue cumpliendo pues $\lambda_i = 0$. Supongamos que i es estable, podemos particionar con el tiempo del primer cambio de estado S_1 entonces tenemos lo siguiente

$$p_{i,j}^{(t)} = P_i(X_t = j, S_1 \leq t) + P_i(X_t = j, S_1 > t)$$

Si el tiempo del primer cambio de estado es mayor a t entonces el proceso ha permanecido en el mismo estado desde que inició el proceso hasta el tiempo t entonces

$$P_i(X_t = j, S_1 > t) = e^{-\lambda_i t} \delta_{i,j}$$

pues los tiempos de interarribo se distribuyen exponencial.

Ahora

$$\begin{aligned}
& P_i(X_t = j, S_1 = s) \\
&= \sum_{k \neq i} P_i(X_t = j, S_1 = s, Y_1 = k) \\
&= \sum_{k \neq i} P(X_t = j | S_1 = s, Y_1 = k, X_0 = i) P(Y_1 = k | S_1 = s, X_0 = i) P_i(S_1 = s) \\
&= \sum_{k \neq i} P(X_t = j | S_1 = s, Y_1 = k) P(Y_1 = k | S_1 = s, X_0 = i) \lambda_i e^{-\lambda_i s} \text{(P.M)} \\
&= \sum_{k \neq i} P(X_t = j | X_s = k) P(Y_1 = k | Y_0 = i) \lambda_i e^{-\lambda_i s} \text{(independencia)} \\
&= \sum_{k \neq i} p_{k,j}^{(t-s)} q_{i,k} \lambda_i e^{-\lambda_i s}
\end{aligned}$$

Ahora podemos integrar para encontrar la probabilidad acumulada

$$\begin{aligned}
P_i(X_t = j, S_1 \leq t) &= \int_0^t P_i(X_t = j, S_1 = s) ds \\
&= \int_0^t \lambda_i e^{-\lambda_i s} \sum_{k \neq i} q_{i,k} p_{k,j}^{(t-s)} ds
\end{aligned}$$

□

El siguiente teorema nos dará una expresión más amable para las probabilidades de transición en términos de $q_{i,j}$ y λ_i , pero antes veamos una definición que nos ayudara a simplificar la notación.

Definición 1.2.4. La matriz de intensidades también llamada generador infinitesimal $\mathbf{\Lambda} = (\lambda_{ij})_{i,j \in E}$ del proceso de Markov de saltos X está definida por:

$$\lambda_{ij} = \begin{cases} \lambda_i q_{ij} & \text{si } i \neq j \\ -\sum_{i \neq j} \lambda_{ij} = -\lambda_i & \text{si } i = j \end{cases}$$

Es claro que los renglones de la matriz $\mathbf{\Lambda}$ suman cero.

Teorema 1.2.6. (Ecuaciones diferenciales de Kolmogorov) Para cualquier $i, j \in E$, la función $t \rightarrow p_{ij}^{(t)}$ es diferenciable y su derivada es

$$\frac{d}{dt} p_{ij}^{(t)} = \sum_{k \in E} \lambda_{ik} p_{kj}^{(t)} = \sum_{k \in E} p_{ik}^{(t)} \lambda_{kj}$$

Demostración. Tomando la expresión de la proposición (1.2.1) y haciendo el cambio de variable $u = t - s$ tenemos

$$p_{i,j}^{(t)} = e^{-\lambda_i t} \left(\delta_{i,j} + \int_0^t \lambda_i e^{\lambda_i u} \sum_{k \neq i} q_{i,k} p_{k,j}^{(u)} du \right)$$

Notemos que el integrando es acotado en $[0, t]$ por lo que $p_{i,j}^{(t)}$ es derivable entonces

$$\begin{aligned} \frac{d}{dt} p_{ij}^{(t)} &= -\lambda_i e^{-\lambda_i t} \left(\delta_{ij} + \int_0^t \lambda_i e^{\lambda_i u} \sum_{k \neq i} q_{ik} p_{kj}^{(u)} du \right) + e^{-\lambda_i t} \lambda_i e^{\lambda_i t} \sum_{k \neq i} q_{ik} p_{kj}^{(t)} \\ &= -\lambda_i p_{ij}^{(t)} + \lambda_i \sum_{k \neq i} q_{ik} p_{kj}^{(t)} \\ &= \lambda_{ii} p_{ij}^{(t)} + \sum_{k \neq i} \lambda_{ik} p_{kj}^{(t)} \quad (\text{def. 1.2.4}) \\ &= \sum_k \lambda_{ik} p_{kj}^{(t)} \end{aligned}$$

Que es equivalente a

$$\frac{d}{dt} \mathbf{P}^t = \mathbf{\Lambda} \mathbf{P}^t$$

llamada ecuación backward de Kolmogorov.

Para demostrar que las probabilidades de transición cumplen la ecuación forward de Kolmogorov necesitaremos suponer que las intensidades son acotadas, es decir $\sup \lambda_i < \infty$.

Sabemos que el tiempo que el proceso se mantiene en el estado i se distribuye exponencial con parámetro λ_i entonces la probabilidad de cambiar de estado en el intervalo $[t, t + dt)$ es $\lambda_i dt$, además condicionando a que hay un cambio de estado en este intervalo y suponiendo que $X_{t-} = i$ la probabilidad de que el siguiente estado sea j es $q_{i,j}$ entonces la probabilidad de tener un cambio de estado en el intervalo $[t, t + dt)$ es $\lambda_i dt q_{i,j} = \lambda_{i,j} dt$ entonces

$$p_{ij}^{(h)} = \lambda_{ij} h + o(h)$$

si $h \rightarrow 0$ y si $i = j$

$$p_{ii}^{(h)} = 1 - \lambda_i h + o(h)$$

entonces

$$\frac{p_{ij}^{(h)} - \delta_{ij}}{h} \rightarrow \lambda_{ij}$$

Necesitamos que la expresión anterior sea uniformemente acotada en i, j y h . Primero

para $i \neq j$

$$\begin{aligned}
 0 \leq p_{ij}^s &\leq \text{probabilidad de algun salto en } [0,s] \\
 &= \int_0^s \lambda_i e^{-\lambda_i u} du \\
 &\leq \lambda_i \int_0^s du \\
 &= \lambda_i s
 \end{aligned}$$

que implica

$$0 \leq \frac{p_{ij}^s}{s} \leq \lambda_i \leq \sup \lambda_i$$

análogamente

$$0 \leq 1 - p_{ii}^s \leq \text{probabilidad de algun salto en } [0,s] \leq \lambda_i s$$

entonces por el teorema de convergencia dominada

$$\begin{aligned}
 \frac{p_{ij}^{s+h} - p_{ij}^s}{h} &= \frac{1}{h} \left(\sum_k p_{ik}^s p_{kj}^h - p_{ij}^s \right) \\
 &= \sum_k p_{ik}^s \frac{p_{kj}^h - \delta_{kj}}{h} \\
 &\rightarrow \sum_k p_{ik}^s \lambda_{kj}
 \end{aligned}$$

que es equivalente a

$$\frac{d}{dt} \mathbf{P}^t = \mathbf{P}^t \mathbf{\Lambda}$$

llamada ecuación forward de Kolmogorov. □

A la función $\exp(At) = \sum_{n=0}^{\infty} (At)^n / n!$ donde A es una matriz cuadrada de dimensión finita la llamamos matriz exponencial.

Corolario 1.2.2. *Si E es finito entonces*

$$\mathbf{P}^t = \exp(\mathbf{\Lambda}t) = \sum_{n=0}^{\infty} (\mathbf{\Lambda}t)^n / n!$$

Demostración. Se sigue de las ecuaciones diferenciales de Kolmogorov y de que $\mathbf{P}^{(0)} = \mathbf{I}$. □

Para los procesos de Markov de saltos la clasificación de los estados es análoga a la clasificación en las cadenas de Markov.

Teorema 1.2.7. *Las siguientes afirmaciones son equivalentes:*

1. $\{Y_n\}_{n \in \mathbb{N}}$ es irreducible
2. $\forall i, j \in E, \exists t > 0 : p_{ij}^t > 0$
3. $\forall i, j \in E, \forall t > 0 : p_{ij}^t > 0$

Demostración. Se ve que 3. implica 2. pues si $p_{ij}^t > 0$ pasa para todo tiempo positivo entonces existe uno para el que se cumple $p_{ij}^t > 0$, también se ve que 2. implica 1. pues para todos los estados existe una probabilidad positiva de llegar a ellos para algún tiempo solo falta ver que 1. implica 3.

Sean $i, j \in E$ como i comunica con j en la cadena asociada existe un recorrido de estados de i a j digamos i, i_1, \dots, i_n, j tales que

$$q_{i,i_1} > 0, q_{i_1,i_2} > 0, \dots, q_{i_n,j} > 0$$

entonces p_{ij}^t es mayor o igual que la probabilidad de

$$i \rightarrow i_1 \rightarrow \dots \rightarrow i_n \rightarrow j$$

en un tiempo t positivo y esto último es una suma de exponenciales es decir una distribución gama que tiene densidad positiva en todos los puntos mayores a cero por lo tanto $p_{ij}^t > 0$. \square

Definición 1.2.5. $\{X_t\}$ es irreducible si existe $t > 0$ tal que $p_{ij}^t > 0$ para todo $i, j \in E$

Por el teorema anterior $\{X_t\}$ es irreducible si y solo si $\{Y_n\}$ es irreducible.

Definición 1.2.6. $\{X_t\}$ es recurrente o transitoria si $\{Y_n\}$ es recurrente o transitoria respectivamente.

Definición 1.2.7. Un vector renglón $\nu = (\nu_j)_{j \in E}$ es llamado medida estacionaria o invariante de un proceso de Markov con matriz de probabilidades de transición \mathbf{P}^t si ν es finita, no cero, no negativa y satisface

$$\nu = \nu \mathbf{P}^t$$

para todo $t \geq 0$.

Teorema 1.2.8. Sea $\{X_t\}_{t \geq 0}$ un proceso de Markov de saltos irreducible, recurrente y sea k un estado arbitrario del proceso, definimos $\nu = \{\nu_i\}_{i \in E}$ como

$$\nu_i = E_k \left(\int_0^{U_k} 1_{\{X_s=i\}} ds \right)$$

donde $U_k = \inf\{t > 0 | X_t = k, X_{t-} \neq k\}$ es decir el tiempo del primer retorno al estado k ($X_{t-} = \lim_{s \uparrow t} X_s$ limite por la izquierda). Entonces ν es una medida estacionaria de $\{X_t\}$ y es única salvo multiplicación por escalares.

Demostración.

$$\begin{aligned}
\nu_i &= E_k \left(\int_0^{U_k} 1_{\{X_s=i\}} ds \right) \\
&= E_k \left(\int_0^\infty 1_{\{X_s=i, U_k>s\}} ds \right) \\
&= \int_0^\infty P_k(X_s = i, U_k > s) ds \quad (\text{Fubini})
\end{aligned}$$

$\{U_k > s\}$ es $\sigma(X_u : u \leq s)$ -medible porque podemos decidir si ya ocurrió o no U_k al tiempo s si ya conocemos la evolución de X_t hasta el tiempo s entonces U_k es un tiempo de paro y tenemos lo siguiente

$$p_{ij}^t = P(X_{s+t} = j | X_s = i, U_k > s)$$

entonces

$$\begin{aligned}
\nu_i p_{ij}^t &= \int_0^\infty P_k(X_s = i, U_k > s) P(X_{s+t} = j | X_s = i, U_k > s) ds \\
&= \int_0^\infty P(X_{s+t} = j, X_s = i, U_k > s) ds
\end{aligned}$$

sumando sobre i y luego intercambiando integral con esperanza

$$\begin{aligned}
(\nu \mathbf{P}^t)_j &= \int_0^\infty P(X_{s+t} = j, U_k > s) ds \\
&= \int_0^\infty E_k(1_{\{X_{s+t}=j, U_k>s\}}) ds \\
&= E_k \left(\int_0^{U_k} 1_{\{X_{s+t}=j\}} ds \right) \\
&= E_k \left(\int_t^{t+U_k} 1_{\{X_s=j\}} ds \right) \\
&= E_k \left(\int_t^{U_k} 1_{\{X_s=j\}} ds \right) + E_k \left(\int_{U_k}^{t+U_k} 1_{\{X_s=j\}} ds \right) \\
&= E_k \left(\int_t^{U_k} 1_{\{X_s=j\}} ds \right) + E_k \left(\int_0^t 1_{\{X_s=j\}} ds \right) \\
&= E_k \left(\int_0^{U_k} 1_{\{X_s=j\}} ds \right) \\
&= \nu_j
\end{aligned}$$

por lo anterior $\nu = \nu \mathbf{P}$ además X_0, X_1, \dots es una cadena de Markov irreducible entonces $p_{ij}^t > 0$ para toda $t > 0$ en particular para $t = 1$ y por el teorema (1.1.5) solo falta ver que esta cadena es recurrente y tendremos que la medida ν es única salvo multiplicación

por constantes. Por hipótesis $\{X_t\}$ es recurrente entonces i sera visitado infinitamente en $\{X_t\}$ pero no sabemos si los tiempos de visita al estado i coinciden con los tiempos $1, 2, \dots$ para los cuales tenemos la cadena. Sea T_1^i, T_2^i, \dots los tiempos de espera en el estado i para $\{X_t\}$, estos tiempo tienen distribución exponencial y por lo tanto probabilidad positiva de ser mayores a uno por lo que $X_t = i$ para una infinidad de tiempos naturales entonces la cadena $\{X_n\}_{n \in \mathbb{N}}$ es recurrente \square

Ahora consideremos $\{X_t\}$ con i un estado recurrente y la cadena asociada $\{Y_n\}_{n \in \mathbb{N}}$, veremos cómo están relacionadas las medidas estacionarias definidas en el teorema anterior y en el teorema (1.1.7)

Teorema 1.2.9. *Sea $\{X_t\}_{t \geq 0}$ irreducible, recurrente con matriz de intensidades $\mathbf{\Lambda} = \{\lambda_{ij}\}_{ij \in E}$ y $k \in E$. Entonces la relación entre la medida estacionaria canónica para cadenas de Markov (teorema 1.1.7) y la medida definida en el teorema anterior en términos de k esta dada por $\mu_j = \lambda_j \nu_j$*

Demostración. El proceso de Markov es constante a trozos con tiempos de interarribo T_0, T_1, \dots sea $\tau_k = \inf\{n \geq 0 | Y_n = k\}$

$$\begin{aligned}
\nu_j &= E_k \left(\int_0^{U_k} 1_{\{X_s=j\}} ds \right) \\
&= E_k \left(\sum_{n=0}^{\tau_k-1} T_n 1_{\{Y_n=j\}} \right) \\
&= E_k \left(\sum_{n=0}^{\infty} T_n 1_{\{Y_n=j, \tau_k > n\}} \right) \\
&= \sum_{n=0}^{\infty} E_k (T_n 1_{\{Y_n=j, \tau_k > n\}}) \\
&= \sum_{n=0}^{\infty} E_k (E_k (T_n 1_{\{Y_n=j, \tau_k > n\}} | \{Y_n\}_{n \in \mathbb{N}})) \\
&= \sum_{n=0}^{\infty} E_k (1_{\{Y_n=j, \tau_k > n\}} E_k (T_n | Y_n = j)) \quad (\text{medibilidad}) \\
&= \sum_{n=0}^{\infty} E_k (T_n | Y_n = j) E_k (1_{\{Y_n=j, \tau_k > n\}}) \\
&= \lambda_j^{-1} \mu_j
\end{aligned}$$

\square

Teorema 1.2.10. *Una medida estacionaria ν para un proceso de Markov con matriz de intensidades $\mathbf{\Lambda}$ satisface*

$$\nu \mathbf{\Lambda} = 0$$

Demostración. Derivando $\nu = \nu \mathbf{P}^t$ con respecto de t se tiene $\nu \mathbf{\Lambda} \mathbf{P}^t = 0$ y luego haciendo tender t a cero $\nu \mathbf{\Lambda} = 0$. \square

El teorema anterior nos da un forma de encontrar medidas estacionarias ν como solución de $\nu\Lambda = 0$.

1.3 Cadenas y procesos de Markov revertidos en el tiempo

En las cadenas y procesos de Markov con tiempo revertido se toma el proceso en el sentido contrario del usual es decir el tiempo decrece.

Definición 1.3.1. Consideremos una cadena de Markov homogénea $\{X_n\}_{n \geq 0}$ con espacio de estados discreto E y matriz de transición $\mathbf{P} = \{p_{ij}\}_{i,j \in E}$. Fijamos un numero $N \in \mathbb{N}$ y definimos la cadena de Markov revertida en el tiempo como

$$\tilde{X}_i = X_{N-i}, \quad i = 0, 1, \dots, N$$

Teorema 1.3.1. Sea $\{X_n\}_{n \geq 0}$ una cadena de Markov estacionaria con probabilidades de transición p_{ij} y π su distribución estacionaria tal que para todo i en el espacio de estados $\pi_i > 0$. Entonces $\{\tilde{X}_n\}$ la cadena revertida en tiempo con $n = 0, 1, \dots, N$ es una cadena de Markov homogénea con probabilidades de transición

$$\tilde{p}_{ij} = \frac{p_{ji}\pi_j}{\pi_i}$$

para todo $i, j \in E$.

La matriz de transiciones $\tilde{\mathbf{P}}$ se puede escribir como

$$\tilde{\mathbf{P}} = \Delta^{-1}(\pi)\mathbf{P}'\Delta(\pi)$$

donde $\Delta(\pi)$ es una matriz diagonal con π en la diagonal.

Demostración. Consideremos las probabilidades de transición de la cadena revertida en tiempo y $\pi_i > 0$ para todo $i \in E$ entonces

$$\begin{aligned} P(\tilde{X}_{n+1} = j | \tilde{X}_n = i) &= P(X_{N-n-1} = j | X_{N-n} = i) \\ &= \frac{P(X_{N-n} = i | X_{N-n-1} = j)P(X_{N-n-1} = j)}{P(X_{N-n} = i)} \\ &= \frac{P(X_{N-n} = i | X_{N-n-1} = j)\pi_j}{\pi_i} \end{aligned}$$

y

$$\begin{aligned} P(\tilde{X}_0 = i_0, \tilde{X}_1 = i_1, \dots, \tilde{X}_n = i_n) &= P(X_N = i_0, X_{N-1} = i_1, \dots, X_{N-n} = i_n) \\ &= P(X_n = i_0, X_{n-1} = i_1, \dots, X_0 = i_n) \\ &= p_{i_1 i_0} p_{i_2 i_1} \cdots p_{i_n i_{n-1}} \pi_{i_n} \\ &= \frac{\tilde{p}_{i_0 i_1} \pi_{i_0}}{\pi_{i_1}} \frac{\tilde{p}_{i_1 i_2} \pi_{i_1}}{\pi_{i_2}} \cdots \frac{\tilde{p}_{i_{n-1} i_n} \pi_{i_{n-1}}}{\pi_{i_n}} \pi_{i_n} \\ &= \pi_{i_0} \prod_{r=0}^{n-1} \tilde{p}_{i_r i_{r+1}} \end{aligned}$$

por teorema 1.1.1 esta cadena es homogénea. La forma de escribir la matriz de transición de la cadena revertida en tiempo se sigue de que la inversa de una matriz diagonal es poner en cada elemento de la diagonal el inverso correspondiente de la matriz original y de que $\tilde{p}_{ij} = \frac{p_{ji}\pi_j}{\pi_i}$ \square

Si π cumple las hipótesis del teorema anterior entonces también es distribución estacionaria para la cadena revertida en tiempo ya que

$$(\pi \tilde{\mathbf{P}})_j = \sum_{i \in E} \pi_i \tilde{p}_{ij} = \sum_{i \in E} \pi_i \frac{p_{ji}\pi_j}{\pi_i} = \pi_j$$

Veamos el resultado opuesto del teorema anterior.

Teorema 1.3.2. *Si la cadena revertida en tiempo es homogénea, para toda i y n $P(X_n = i) > 0$ y si $p_{ij} > 0$ para todo i, j entonces la cadena original es estacionaria.*

Demostración. Del teorema anterior y por homogeneidad de la cadena revertida en el tiempo

$$\tilde{p}_{ij} = p_{ji} \frac{P(X_{N-n-1} = j)}{P(X_{N-n} = i)}$$

si definimos $\rho_i = \frac{p_{ii}}{p_{ii}}$ tenemos que

$$\begin{aligned} P(X_n = i) &= \rho_i P(X_{n-1} = i) \\ &= \dots \\ &= \rho_i^n P(X_0 = i) \end{aligned}$$

reemplazando en la expresión anterior y por homogeneidad tenemos

$$\left(\frac{\rho_i}{\rho_j} \right)^{N-n-1} = \rho_i^{-1} \frac{P(X_0 = j)}{P(X_0 = i)}$$

la expresión anterior es válida para toda n entonces $\rho_i = \rho_j = \rho$ para toda i, j entonces

$$P(X_n = i) = \rho^n P(X_0 = i)$$

y sumando sobre i llegamos a que $\rho = 1$ y esto implica

$$P(X_n = i) = P(X_0 = i)$$

\square

Corolario 1.3.1. *Bajo las hipótesis del teorema anterior la cadena revertida en el tiempo y la cadena original tienen la misma distribución si y solo si se cumple $\pi_i p_{ij} = \pi_j p_{ji}$ y en este caso decimos que la cadena es reversible.*

Demostración. Del teorema anterior la cadena tiene distribución estacionaria π , si las cadenas tienen la misma distribución entonces $\pi_i p_{ij} = \pi_j p_{ji}$ se sigue de que

$$\tilde{p}_{ij} = \frac{p_{ji}\pi_j}{\pi_i}$$

Por el teorema (1.1.1) las probabilidades de transición caracterizan a la distribución de la cadena y tenemos

$$\tilde{p}_{ij} = \frac{p_{ji}\pi_j}{\pi_i} = \frac{p_{ij}\pi_i}{\pi_i} = p_{ij}$$

es decir las transiciones son iguales para todo $i, j \in E$. □

Veamos el proceso de saltos de Markov $\{X_t\}_{t \geq 0}$ revertido en el tiempo con espacio de estados discreto E y matriz de intensidades $\mathbf{\Lambda}$.

Definición 1.3.2. Sea $\{X_t\}_{t \in \mathbb{R}}$ un proceso de Markov en tiempo doble infinito. Entonces definimos el proceso revertido en el tiempo $\{\tilde{X}_t\}_{t \in \mathbb{R}}$ como

$$\tilde{X}_t = X_{-t-} = \lim_{s \uparrow -t} X_s$$

La razón de definir así el proceso revertido en el tiempo es para preservar la continuidad por la derecha y limite por la izquierda y el tiempo doble infinito es para evitar el problema de elegir un tiempo a partir del cual iniciar el proceso al revés.

Teorema 1.3.3. Sea $\{X_t\}_{t \in \mathbb{R}}$ un proceso de Markov irreducible, recurrente y estacionario con distribución estacionaria π . Entonces el proceso revertido en el tiempo $\{\tilde{X}_t\}_{t \in \mathbb{R}}$ es un proceso de Markov bien definido con matriz de intensidades

$$\tilde{\mathbf{\Lambda}} = \Delta^{-1} \mathbf{\Lambda}' \Delta$$

más aun $\{\tilde{X}_t\}_{t \in \mathbb{R}}$ tiene la misma distribución estacionaria π .

Demostración. Puesto que $\{\tilde{X}_t\}$ es estacionaria no tiene estados instantáneos y entonces el proceso revertido en el tiempo es un proceso de Markov de saltos.

Consideremos el proceso "discretizado" con pasos de longitud h . Este proceso es irreducible y recurrente esto se ve como en la demostración del teorema (1.2.8). Además es aperiódica debido a que los tiempos de estancia en cada estado son exponenciales esto hace que la cadena discretizada se encuentre en el mismo estado para tiempos consecutivos con probabilidad positiva. Por lo anterior tenemos que $\pi_i > 0$ para toda $i \in E$ entonces por el teorema(1.3.1) la cadena revertida en el tiempo es homogénea y por esto sucede lo mismo con el proceso de Markov de revertido. Aun mas es irreducible recurrente y con distribución estacionaria π debido a que podemos aproximar p_{ij}^x arbitrariamente cerca discretizando el proceso y a que las probabilidades de transición son

continuas (ecuaciones Kolmogorov).

$$\begin{aligned}
\tilde{\lambda}_{ij} &= \lim_{h \downarrow 0} \frac{\tilde{p}_{ij}^h}{h} \\
&= \lim_{h \downarrow 0} \frac{P(\tilde{X}_0 = j, \tilde{X}_{-h} = i)}{P(\tilde{X}_{-h} = i)h} \\
&= \lim_{h \downarrow 0} \frac{P(X_0 = j, X_h = i)}{P(X_h = i)h} \\
&= \lim_{h \downarrow 0} \frac{p_{ji}^h \pi_j}{\pi_i h} \\
&= \frac{\lambda_{ji} \pi_j}{\pi_i}
\end{aligned}$$

□

Teorema 1.3.4. (*Lema de Kelly*) Considere un proceso de Markov estacionario irreducible y recurrente con matriz de intensidades $\mathbf{\Lambda} = \{\lambda_{ij}\}_{i,j \in E}$. Si existen números $\tilde{\lambda}_{ij} > 0, i \neq j$ tales que

$$\sum_{j \neq i} \lambda_{ij} = \sum_{j \neq i} \tilde{\lambda}_{ij}$$

y una distribución π tal que

$$\pi_j \lambda_{ji} = \pi_i \tilde{\lambda}_{ij}, i \neq j$$

entonces $\tilde{\lambda}_{ij}$ con $i \neq j$ son intensidades de transición del proceso revertido en el tiempo y π es una distribución estacionaria para el proceso original y para el proceso revertido.

Demostración.

$$\begin{aligned}
\sum_j \pi_j \lambda_{ji} &= \sum_{j \neq i} \pi_j \lambda_{ji} + \pi_i \lambda_{ii} \\
&= \sum_{j \neq i} \pi_i \tilde{\lambda}_{ij} + \pi_i \lambda_{ii} \\
&= \pi_i \sum_{j \neq i} \tilde{\lambda}_{ij} + \pi_i \lambda_{ii} \\
&= \pi_i \sum_{j \neq i} \lambda_{ij} + \pi_i \lambda_{ii} \\
&= 0
\end{aligned}$$

equivalentemente $\pi \mathbf{\Lambda} = 0$ entonces π es estacionaria y por el teorema anterior también lo es para el proceso revertido en el tiempo. □

Capítulo 2

Distribuciones tipo fase y matriz exponencial

En este capítulo se desarrollaran dos tipos de distribuciones con propiedades interesantes y accesibles. Las distribuciones tipo fase que fueron extraídas de [1] y [2], y una generalización de estas llamadas distribuciones matriz exponencial que fueron extraídas de [5] , [6] y [8].

2.1 Distribuciones tipo fase

Las distribuciones tipo fase son generadas por un proceso de Markov de saltos o una cadena de Markov con un numero finito de estados. Sea X un proceso de Markov de saltos con espacio de estados finito E con p estados transitorios y un estado $p + 1$ absorbente, entonces su matriz de intensidades es de la forma

$$\Lambda = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}$$

donde \mathbf{T} es una submatriz de intensidades de dimensión $p \times p$. Puesto que los renglones de una matriz de intensidades suman cero tenemos que

$$\mathbf{t} = -\mathbf{T}\mathbf{e}$$

donde \mathbf{e} es un vector columna de unos de dimensión p , el vector \mathbf{t} es llamado vector de intensidad de salida, además supongamos que el proceso tiene una distribución inicial(vector fila) π definida en los p estados transitorios es decir $\pi_i = P(X_0 = i)$ con $i \in \{1, 2, \dots, p\}$ y $\pi\mathbf{e} = \sum_{i=1}^p \pi_i = 1$

Definición 2.1.1. (*Distribución tipo fase*) Una distribución tipo fase es la distribución del tiempo de paro $\tau = \inf\{t > 0 | X_t = p + 1\}$ antes de la absorción en un proceso de saltos de Markov con espacio de estados finito $E = \{1, 2, \dots, p, p + 1\}$, un estado absorbente $\{p + 1\}$ y todos los demás transitorios $\{1, 2, \dots, p\}$, si este proceso tiene distribución inicial π y submatriz de intensidades \mathbf{T} se denotara como $\tau \sim PH(\pi, \mathbf{T})$ o $\tau \sim PH_p(\pi, \mathbf{T})$ si es necesario especificar la dimensión del espacio de estados.

Ejemplo 2.1.1. (*Distribución exponencial*) Sea $\tau \sim PH_1(\pi, \mathbf{T})$, entonces existe un solo estado transitorio con tiempo de interarribo distribuido exponencial, es decir, la distribución del tiempo de llegada al estado absorbente es exponencial.

Ejemplo 2.1.2. (*Erlang (Convolución de exponenciales)*) Sea $Y = X_1 + X_2 + \dots + X_p$ donde X_1, X_2, \dots son independientes y exponenciales con media $1/\lambda_i$ respectivamente, para representar esta distribución podemos pensar que un usuario debe pasar p nodos y en cada nodo debe esperar un tiempo distribuido exponencial con media $1/\lambda_i$.

El usuario llega al nodo 1 y permanece en este nodo un tiempo con distribución exponencial y media $1/\lambda_1$, después pasa al nodo 2 y permanece en este nodo un tiempo con distribución exponencial y media $1/\lambda_2$, continuando este proceso p veces el usuario sale del nodo p en un tiempo Y este proceso se puede describir con un proceso de Markov pues los tiempos de interarribo son exponenciales, pensemos que cada nodo es un estado y agregaremos un estado absorbente al que el usuario llega después de haber pasado los p nodos, es decir el tiempo en que ocurre el paro(absorción) será Y ; entonces este proceso de Markov tendrá $p + 1$ estados y el proceso saltara de un estado al siguiente con probabilidad uno y con una tasa de λ_i con $i \in \{1, 2, \dots, p\}$ además el proceso inicia en el primer estado por lo que $P(X_0 = 1) = 1$ y la matriz de intensidades será

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 & 0 \\ 0 & -\lambda_2 & \lambda_2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p & \lambda_p \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Por lo anterior $Y \sim PH(\pi, \mathbf{T})$ donde

$$\pi = (1, 0, \dots, 0)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & \lambda_1 & 0 & \dots & 0 \\ 0 & -\lambda_2 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p \end{pmatrix}$$

Si $\lambda = \lambda_i$ para toda i a la distribución anterior se le llama Erlang y se denota por $Y \sim \text{Erl}(p, \lambda)$

Ejemplo 2.1.3. (*Mezcla de exponenciales*) Sea f_i la densidad de una distribución exponencial con media $1/\lambda_i$ respectivamente y π_i con $i \in \{1, 2, \dots, p\}$ tales que $\pi_i > 0$ y $\sum_{i=1}^p \pi_i = 1$. Entonces

$$f = \sum_{i=1}^p \pi_i f_i$$

es una densidad a la que llamamos mezcla de distribuciones exponenciales. Para encontrar una representación tipo fase de esta distribución pensemos en un usuario que debe pasar por un nodo i de p posibles, llega a este nodo con probabilidad π_i y tarda un

tiempo exponencial con media $1/\lambda_i$ y después entra a un nodo del que no sale; el tiempo que el usuario permanece en alguno de los p nodos tiene una distribución con densidad f . La idea anterior la podemos representar con un proceso de Markov con espacio de estados $E = \{1, 2, \dots, p, p+1\}$ donde los primeros p estados son transitorios y el estado $p+1$ es absorbente, este proceso se inicia con probabilidad π_i en el estado i y después pasa al estado $p+1$ con probabilidad uno entonces el generador infinitesimal de este proceso es

$$\mathbf{\Lambda} = \begin{pmatrix} -\lambda_1 & 0 & 0 & \dots & 0 & \lambda_1 \\ 0 & -\lambda_2 & 0 & \dots & 0 & \lambda_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p & \lambda_p \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

Por lo anterior f se distribuye $PH(\pi, \mathbf{T})$ con

$$\pi = (\pi_1, \pi_2, \dots, \pi_p)$$

$$\mathbf{T} = \begin{pmatrix} -\lambda_1 & 0 & 0 & \dots & 0 \\ 0 & -\lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_p \end{pmatrix}$$

En los ejemplos anteriores se puede ver que la representación no es única pues si permutamos el orden de los estados tendremos la misma distribución. El siguiente lema nos ayudara a encontrar la densidad de una distribución tipo fase.

Lema 2.1.1. Si el proceso de Markov asociado a una distribución tipo fase de dimensión p tiene generador infinitesimal $\mathbf{\Lambda}$ entonces

$$\exp(\mathbf{\Lambda}s) = \begin{pmatrix} \exp(\mathbf{T}s) & \mathbf{e} - \exp(\mathbf{T}s)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}$$

Demostración. Usando que $\mathbf{t} = -\mathbf{T}\mathbf{e}$ se tiene

$$\mathbf{\Lambda}^n = \begin{pmatrix} \mathbf{T}^n & -\mathbf{T}^n\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}$$

$$\begin{aligned}
\exp(\Lambda s) &= \sum_{n=0}^{\infty} \Lambda^n s^n / n! \\
&= \mathbf{I}_{p+1} + \sum_{n=1}^{\infty} \Lambda^n s^n / n! \\
&= \mathbf{I}_{p+1} + \sum_{n=1}^{\infty} \begin{pmatrix} \mathbf{T}^n & -\mathbf{T}^n \mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} s^n / n! \\
&= \mathbf{I}_{p+1} + \begin{pmatrix} \sum_{n=1}^{\infty} \mathbf{T}^n s^n / n! & \sum_{n=1}^{\infty} -\mathbf{T}^n s^n \mathbf{e} / n! \\ \mathbf{0} & 0 \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{I}_p + \sum_{n=1}^{\infty} \mathbf{T}^n s^n / n! & \mathbf{I}_p \mathbf{e} - \mathbf{I}_p \mathbf{e} + \sum_{n=1}^{\infty} -\mathbf{T}^n s^n \mathbf{e} / n! \\ \mathbf{0} & 1 \end{pmatrix} \\
&= \begin{pmatrix} \exp(\mathbf{T}s) & \mathbf{e} - \exp(\mathbf{T}s)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}
\end{aligned}$$

□

Como Λ es una matriz de intensidades asociada a una distribución tipo fase entonces por lo visto en el capítulo anterior

$$\mathbf{P}^s = \exp(\Lambda s)$$

donde \mathbf{P}^s es la matriz de transición del proceso de Markov, en particular si restringimos la matriz al conjunto $E \setminus \{p+1\}$ de los p estados transitorios

$$\mathbf{P}^s|_{E \setminus \{p+1\}} = \exp(\mathbf{T}s)$$

Teorema 2.1.1. Sea $\tau \sim PH_p(\pi, \mathbf{T})$ entonces

1. $P(\tau > x) = \pi e^{\mathbf{T}x} \mathbf{e}$
2. La densidad de τ está dada por $f(x) = \pi e^{(\mathbf{T}x)} \mathbf{t}$

Demostración. 1. La Probabilidad de que la absorción ocurra después de un tiempo x es igual a la probabilidad de que el proceso se encuentre en alguno de los estados transitorios en el tiempo x . Sea E el conjunto de los estados transitorios.

$$\begin{aligned}
P(\tau > x) &= P(X_x \in E) \\
&= \sum_{j=1}^p P(X_x = j) \\
&= \sum_{j=1}^p \sum_{i=1}^p P(X_0 = i) P(X_x = j | X_0 = i) \\
&= \sum_{j=1}^p \sum_{i=1}^p \pi_i \mathbf{P}_{i,j}^x \\
&= \pi e^{\mathbf{T}x} \mathbf{e}
\end{aligned}$$

2. Derivando la expresión anterior y recordando que $-\frac{d}{dx}P(\tau > x) = f_\tau(x)$

$$\begin{aligned} f_\tau(x) &= -\frac{d}{dt}P(\tau > x) = -\pi \mathbf{T} e^{\mathbf{T}x} \mathbf{e} \\ &= -\pi e^{\mathbf{T}x} \mathbf{T} \mathbf{e} \\ &= -\pi e^{\mathbf{T}x} (-\mathbf{t}) \\ &= \pi e^{\mathbf{T}x} \mathbf{t} \end{aligned}$$

□

Corolario 2.1.1. Si $\tau \sim PH_p(\pi, \mathbf{T})$ entonces la función de distribución de τ es

$$F(x) = 1 - \pi e^{\mathbf{T}x} \mathbf{e}$$

Teorema 2.1.2. En un proceso de Markov de saltos con $p + 1$ un estados y matriz de intensidades

$$\mathbf{\Lambda} = \begin{pmatrix} \mathbf{T} & \mathbf{t} \\ \mathbf{0} & 0 \end{pmatrix}$$

la matriz \mathbf{T} es invertible si y sólo si los estados $1, 2, \dots, p$ son transitorios.

Demostración. \Rightarrow) Sea a_i la probabilidad de absorción al estado $p + 1$ iniciando en el estado $i \in \{1, 2, \dots, p\}$ y $q_{i,j}$ la probabilidad de transición del estado i a j de la cadena de Markov (Y_n) asociada al proceso de Markov entonces

$$\begin{aligned} a_i &= P_i(\text{absorción}) \\ &= \sum_{j=1}^{p+1} P_i(\text{absorción} | Y_1 = j) P_i(Y_1 = j) \\ &= q_{i,p+1} + \sum_{j=1, j \neq i}^p P_j(\text{absorción}) P_i(Y_1 = j) \\ &= q_{i,p+1} + \sum_{j=1, j \neq i}^p a_j q_{i,j} \end{aligned}$$

Como $q_{i,j} = -\lambda_{i,j} / \lambda_{i,i}$ tenemos

$$-\lambda_{i,i} a_i = \lambda_{i,p+1} + \sum_{j=1, j \neq i}^p a_j \lambda_{i,j}$$

que es equivalente a

$$0 = \mathbf{t} + \mathbf{T} \mathbf{a}$$

como \mathbf{T} es invertible

$$\mathbf{a} = -\mathbf{T}^{-1} \mathbf{t} = -\mathbf{T}^{-1} \mathbf{T} \mathbf{e} = \mathbf{e}$$

entonces $a_i = 1$ con $i \in \{1, 2, \dots, p\}$ es decir los estados del 1 al p son transitorios.
 \Leftarrow) Consideramos $\mathbf{vT} = 0$ que es equivalente a

$$v_i \lambda_i = \sum_{j \neq i} v_j \lambda_{ji} = \sum_{j \neq i} v_j \lambda_j q_{ji}$$

pues \mathbf{T} es un subgenerador infinitesimal además $\lambda_i > 0$ pues de lo contrario el estado i no es transitorio y q_{ji} son la probabilidades de transición de la cadena asociada al proceso por lo que $q_{ji} = \lambda_{ji}/\lambda_i$. Sea $\mathbf{Q} = (q_{ij})_{ij \in \{1, 2, \dots, p\}}$ con $q_{ii} = 0$. Si hacemos $\mathbf{u} = (\lambda_1 v_1, \lambda_2 v_2, \dots, \lambda_p v_p)$ tenemos $\mathbf{u} = \mathbf{uQ}$ e iterando $\mathbf{u} = \mathbf{uQ}^n$ que implica

$$\mathbf{u} = \lim_{n \rightarrow \infty} \mathbf{uQ}^n = \mathbf{0}$$

pues los estados $1, 2, \dots, p$ son transitorios, entonces $\mathbf{v} = \mathbf{0}$ es decir las columnas de \mathbf{T} son linealmente independientes. \square

Corolario 2.1.2. Si $\tau \sim PH_p(\pi, \mathbf{T})$ entonces \mathbf{T} es invertible.

Demostración. Una distribución tipo Fase está definida como el tiempo de absorción en un proceso de Markov de saltos con $p + 1$ estados con los estados $1, 2, \dots, p$ transitorios y el estado $p + 1$ es absorbente. \square

Lema 2.1.2. Si $\tau \sim PH(\pi, \mathbf{T})$ entonces para todo $s \geq 0$ la matriz $(\mathbf{T} - s\mathbf{I})$ es una matriz de subintensidad y por lo tanto invertible.

Demostración.

$$(\mathbf{T} - s\mathbf{I})_{ij} = (\mathbf{T})_{ij} \geq 0$$

si $i \neq j$ y si $i = j$

$$(\mathbf{T} - s\mathbf{I})_{ij} \leq (\mathbf{T})_{ij} < 0$$

entonces es un subgenerador y por el corolario anterior tenemos la invertibilidad. \square

Lema 2.1.3. Si A es invertible $\int_0^\infty e^{sA} ds = -A^{-1}$

Demostración. Si A es invertible $\int e^{sA} ds = A^{-1}e^{sA}$ por el corolario (A.1.1) entonces

$$\begin{aligned} \int_0^\infty e^{sA} ds &= A^{-1}e^{sA} \Big|_0^\infty \\ &= A^{-1} \left(\lim_{s \rightarrow \infty} e^{sA} - \mathbf{I} \right) \\ &= A^{-1} \left(\lim_{s \rightarrow \infty} \mathbf{P}^s - \mathbf{I} \right) \\ &= A^{-1} (0 - \mathbf{I}) \end{aligned}$$

\square

Teorema 2.1.3. La transformada de Laplace $L_\tau(s)$ de $\tau \sim PH(\pi, \mathbf{T})$ es

$$L_\tau(s) = E(e^{-s\tau}) = \pi(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t}$$

que está bien definida para s mayores al valor propio más grande de \mathbf{T} , en particular para $s > 0$.

Demostración.

$$\begin{aligned} L_\tau(s) &= \int_0^\infty e^{-sx} \pi e^{(\mathbf{T}x)} \mathbf{t} dx \\ &= \int_0^\infty \pi e^{-sx\mathbf{I}} e^{(\mathbf{T}x)} \mathbf{t} dx \\ &= \pi \int_0^\infty e^{(\mathbf{T}x - sx\mathbf{I})} dx \mathbf{t} \\ &= -\pi(\mathbf{T} - s\mathbf{I})^{-1}\mathbf{t} \\ &= \pi(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{t} \end{aligned}$$

□

Ahora podemos encontrar los momentos de las distribuciones tipo fase simplemente con la transformada de Laplace.

Lema 2.1.4. Si \mathbf{T} es una submatriz de intensidad entonces

$$\frac{d^n}{ds^n} (s\mathbf{I} - \mathbf{T})^{-1} = (-1)^n n! (s\mathbf{I} - \mathbf{T})^{-n-1}$$

Demostración. Por inducción

$$\begin{aligned} 0 &= \frac{d}{ds} \mathbf{I} = \frac{d}{ds} (s\mathbf{I} - \mathbf{T})(s\mathbf{I} - \mathbf{T})^{-1} \\ &= (s\mathbf{I} - \mathbf{T})^{-1} + (s\mathbf{I} - \mathbf{T}) \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-1} \end{aligned}$$

que implica

$$\frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-1} = (-1)(s\mathbf{I} - \mathbf{T})^{-2}$$

ahora suponemos la hipótesis cierta para n entonces

$$\begin{aligned} \frac{d^{n+1}}{ds^{n+1}} (s\mathbf{I} - \mathbf{T})^{-1} &= \frac{d}{ds} \frac{d^n}{ds^n} (s\mathbf{I} - \mathbf{T})^{-1} \\ &= (-1)^n n! \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-n-1} \\ &= (-1)^{(n+1)} (n+1)! (s\mathbf{I} - \mathbf{T})^{-n-2} \end{aligned}$$

pues

$$\begin{aligned} 0 &= \frac{d}{ds} \mathbf{I} = \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{n+1} (s\mathbf{I} - \mathbf{T})^{-n-1} \\ &= (n+1)(s\mathbf{I} - \mathbf{T})^{-1} + (s\mathbf{I} - \mathbf{T})^{n+1} \frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-n-1} \end{aligned}$$

que implica

$$\frac{d}{ds} (s\mathbf{I} - \mathbf{T})^{-n-1} = -(n+1)(s\mathbf{I} - \mathbf{T})^{-n-2}$$

□

Teorema 2.1.4. *El n -ésimo momento de $\tau \sim PH(\pi, \mathbf{T})$ es*

$$E(\tau^n) = n! \pi (-\mathbf{T}^{-1})^n \mathbf{e}$$

Demostración.

$$\begin{aligned} E(\tau^n) &= (-1)^n \frac{d^n}{ds^n} L_\tau(s) \Big|_{s=0} \\ &= (-1)^n \pi (-1)^n n! (s\mathbf{I} - \mathbf{T})^{-n-1} \mathbf{t} \Big|_{s=0} \\ &= (-1)^{2n} \pi n! (-\mathbf{T})^{-n-1} \mathbf{t} \\ &= \pi n! (-\mathbf{T})^{-n-1} (-\mathbf{T}\mathbf{e}) \\ &= \pi n! (-\mathbf{T})^{-n} \mathbf{e} \end{aligned}$$

□

Teorema 2.1.5. *(Convolución) Si $X \sim PH_p(\alpha, \mathbf{S})$ y $Y \sim PH_q(\beta, \mathbf{T})$ son independientes entonces $Z = X + Y$ tiene distribución tipo fase con distribución inicial $\pi = (\alpha, \mathbf{0})$ donde $\mathbf{0}$ es un vector de dimensión q de ceros y con submatriz de intensidades \mathbf{U} dada por*

$$\mathbf{U} = \begin{pmatrix} \mathbf{S} & s\beta \\ \mathbf{0} & \mathbf{T} \end{pmatrix}$$

aquí $s = -\mathbf{S}\mathbf{e}$ y $\mathbf{0}$ es una matriz de dimensión $q \times p$.

Demostración. La variable Z es la suma de los tiempos en que se absorben los procesos de Markov asociados a X y Y entonces podemos pensar sin perder generalidad que primero inicia el proceso X e inmediatamente después de que se absorbe inicia el proceso Y , esto lo representamos con un proceso de Markov con espacio de estados $\{1, 2, \dots, p, p+1, \dots, p+q, p+q+1\}$ en donde los estados $\{1, 2, \dots, p\}$ son los estados transitorios de X , $\{p+1, \dots, p+q\}$ son los estados transitorios de Y que conservan sus intensidades entre los estados pertenecientes al mismo proceso y $p+q+1$ es un estado absorbente. Las intensidades con que el proceso pasa de los estados i pertenecientes a X a los estados j pertenecientes a Y son $s_i \beta_j$ es decir el proceso asociado a X llega a la absorción en el estado i y en ese mismo instante el proceso asociado a Y inicia en el estado j , además

las intensidades de absorción de Y son las mismas por lo anterior el proceso asociado a Z tiene matriz de intensidades

$$\Lambda = \begin{pmatrix} \mathbf{S} & s\beta & \mathbf{0} \\ \mathbf{0} & \mathbf{T} & \mathbf{t} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

y Z es el tiempo antes de la absorción de $X + Y$ por lo tanto Z es una distribución tipo fase. \square

Teorema 2.1.6. (Mezcla) Si $X \sim PH_p(\alpha, \mathbf{S})$, $Y \sim PH_q(\beta, \mathbf{T})$ son independientes y

$$U = \begin{cases} X & \text{con probabilidad } p_1 \\ Y & \text{con probabilidad } p_2 \end{cases}$$

donde $p_1 + p_2 = 1$. Entonces U tiene distribución tipo fase con vector inicial $\pi = (p_1\alpha, p_2\beta)$ y submatriz de intensidades

$$\mathbf{W} = \begin{pmatrix} \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{pmatrix}$$

Demostración. Primero notemos que U es igual a X con probabilidad p_1 y a Y con probabilidad p_2 . Es decir el proceso de Markov asociado a U solo estará en el proceso asociado a X o a Y pero no habrá transiciones entre ellos y se absorberá al salir de cualquiera de estos, además iniciara en algún estado i del proceso asociado a X con probabilidad $p_1\alpha_i$ o en algún estado j del proceso asociado a Y con probabilidad $p_2\beta_j$ esto lo representamos con un proceso de Markov con espacio de estados $\{1, 2, \dots, p, p+1, \dots, p+q, p+q+1\}$ en donde los estados $\{1, 2, \dots, p\}$ son los estados transitorios de X , $\{p+1, \dots, p+q\}$ son los estados transitorios de Y que conservan sus intensidades entre los estados pertenecientes al mismo proceso y $p+q+1$ es un estados absorbente que reemplaza al estado absorbente de X y Y , por lo anterior es clara la estructura del generador infinitesimal de U y su distribución inicial. \square

Corolario 2.1.3. Las distribuciones tipo fase son cerradas bajo mezclas finitas de distribuciones tipo fase y también son cerradas bajo convoluciones finitas de distribuciones tipo fase.

Lema 2.1.5. Para toda $a > 0$ la sucesión $X_k \sim \{Erl(k, k/a)\}_{k \in \mathbb{N}}$ converge en probabilidad a la distribución de Dirac ($1_{[a, \infty)}$)

Demostración. Por la desigualdad de Chebyshev

$$P(|X_k - a| \geq \epsilon) \leq \frac{a^2}{k\epsilon^2} \rightarrow 0$$

se tiene la convergencia en probabilidad y por las funciones generadoras de momentos

$$M_{x_k}(t) = \left(1 - \frac{at}{k}\right)^{-k} \rightarrow e^{at} = M_{\delta_a}(t)$$

la convergencia a la distribución delta de Dirac \square

Teorema 2.1.7. *La clase de las distribuciones tipo fase son densas "débiles" en la clase de las distribuciones con soporte en $(0, \infty)$*

Demostración. Sea F una función de distribución con soporte en $(0, \infty)$, como F es monótona entonces es medible además F es positiva y acotada entonces gracias a teoría de la medida existe una sucesión creciente de funciones simples que converge a F puntualmente, es decir para ϵ arbitrario y positivo puedo encontrar g una función simple tal que $|F(t) - g(t)| < \epsilon$ para toda $t \geq 0$. Sean t_n con $n \in \mathbb{N}$ los puntos (ordenados con el índice) en que g salta, g es continua por la derecha por como se construye, $g(-\infty) = 0$ pero no podemos asegurar que $g(\infty) = 1$ pues la sucesión converge por debajo, por esto truncamos g para que a partir de este truncamiento se cumpla $g(\infty) = 1$ y sea una distribución de probabilidad, sea T tal que $1 - g(T) < \epsilon$, existe N tal que $t_n > T$ para toda $n > N$ y definimos

$$G = \sum_{n=1}^{N-1} (g(t_n) - g(t_{n-1}))1_{[t_n, \infty)} + (1 - g(t_N))1_{[t_N, \infty)}$$

notemos que $|F(t) - G(t)| < \epsilon$ para toda $t \geq 0$. Ahora por el lema anterior podemos hacer $|1_{[t_n, \infty)} - E_n| < \epsilon$ con E_n una distribución Erlang y así tendremos que $|G - H| < \epsilon$ donde

$$H = \sum_{n=1}^{N-1} (g(t_n) - g(t_{n-1}))E_n + (1 - g(t_N))E_N$$

además por 2.1.2 y 2.1.3, H es una distribución tipo fase que cumple $|F(t) - H(t)| < \epsilon$ para cada $t \geq 0$ punto de continuidad de F . \square

2.2 Distribuciones matriz exponencial

2.2.1 Transformada de Laplace racional

La transformada de Laplace para una variable aleatoria $X \geq 0$ con densidad f es

$$L_X(s) = E(e^{-sX}) = \int_0^{\infty} e^{-sx} f(x) dx$$

como X es no negativa y f es una densidad L_X está bien definida para $RE(s) \geq 0$. Para s real positivo se tiene la siguiente interpretación probabilista, si Y se distribuye exponencial con parámetro s entonces

$$P(X < Y) = E(P(X < Y|X)) = E(e^{-sX}) = L_X(s)$$

Supongamos que L_X es una función racional es decir

$$L_X(s) = \frac{p(s)}{q(s)}$$

donde p y q son funciones polinomiales

$$p(s) = p_0 s^m + p_1 s^{m-1} + \dots + p_m$$

$$q(s) = q_0 s^n + q_1 s^{n-1} + \dots + q_n$$

Supondremos que el coeficiente de la mayor potencia de p y q , p_0 y q_0 respectivamente son diferentes de cero además supondremos sin perder generalidad que $q_0 = 1$ pues en caso contrario dividimos por q_0 . De la interpretación probabilista de L_X se sigue que $L_X(s) \rightarrow 0$ cuando $s \rightarrow \infty$ esto implica que $m \leq n$, así la forma general para la transformada de Laplace de una variable aleatoria X se escribe como

$$L_X(s) = \frac{p_1 s^{n-1} + \dots + p_n}{s^n + q_1 s^{n-1} + \dots + q_n}$$

donde $\{p_1, \dots, p_{n-1}\}$ pueden ser cero excepto p_n pues si L_X es irreducible, es decir p y q no tienen raíces en común, $L_X(0) = 1$ y también de esto concluimos que $p_n = q_n$. De ahora en adelante asumiremos que L_X es irreducible.

Sea $p(\lambda)$ un polinomio mónico

$$p(\lambda) = \lambda^n + a_1 \lambda^{n-1} + \dots + a_{n-1} \lambda + a_n$$

Se define la matriz compañera de p , $\mathbf{C}(p)$, como

$$\mathbf{C}(p) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -a_n & -a_{n-1} & -a_{n-2} & \dots & -a_1 \end{pmatrix}$$

Teorema 2.2.1. Para un polinomio mónico p se tiene

$$p(\lambda) = \det(\lambda \mathbf{I} - \mathbf{C}(p))$$

Demostración.

$$\det(\lambda \mathbf{I} - \mathbf{C}(p)) = \begin{vmatrix} \lambda & -1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & -1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & -1 \\ a_n & a_{n-1} & a_{n-2} & \dots & a_2 & a_1 + \lambda \end{vmatrix}$$

multiplicando y dividiendo la columna i por λ^{i-1}

$$\det(\lambda \mathbf{I} - \mathbf{C}(p)) = \left(\prod_{i=2}^n \lambda^{i-1} \right)^{-1} \begin{vmatrix} \lambda & -\lambda & 0 & \dots & 0 & 0 \\ 0 & \lambda^2 & -\lambda^2 & \dots & 0 & 0 \\ 0 & 0 & \lambda^3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^{n-1} & -\lambda^{n-1} \\ a_n & a_{n-1} \lambda & a_{n-2} \lambda^2 & \dots & a_2 \lambda^{n-2} & a_1 \lambda^{n-1} + \lambda^n \end{vmatrix}$$

sumando las columnas i desde 2 hasta n a la columna 1 y utilizando el hecho de que $\det(v_1, v_2, \dots, v_n) = \det(v_1 + v_2 + \dots + v_n, v_2, \dots, v_n)$

$$\det(\lambda \mathbf{I} - \mathbf{C}(p)) = \left(\prod_{i=2}^n \lambda^{i-1} \right)^{-1} \begin{vmatrix} 0 & -\lambda & 0 & \dots & 0 & 0 \\ 0 & \lambda^2 & -\lambda^2 & \dots & 0 & 0 \\ 0 & 0 & \lambda^3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda^{n-1} & -\lambda^{n-1} \\ p(\lambda) & a_{n-1}\lambda & a_{n-2}\lambda^2 & \dots & a_2\lambda^{n-2} & a_1\lambda^{n-1} + \lambda^n \end{vmatrix}$$

dividiendo la columna i por λ^{i-1} y desarrollando

$$\begin{aligned} \det(\lambda \mathbf{I} - \mathbf{C}(p)) &= \begin{vmatrix} 0 & -1 & 0 & \dots & 0 & 0 \\ 0 & \lambda & -1 & \dots & 0 & 0 \\ 0 & 0 & \lambda & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda & -1 \\ p(\lambda) & a_{n-1} & a_{n-2} & \dots & a_2 & a_1 + \lambda \end{vmatrix} \\ &= (-1)^{n-1} p(\lambda) (-1)^{n-1} \\ &= p(\lambda) \end{aligned}$$

□

Ahora podemos dar una representación matricial de una fracción de polinomios irreducible.

Teorema 2.2.2. *Sea*

$$f(\lambda) = \frac{p(\lambda)}{q(\lambda)} = \frac{p_1\lambda^{n-1} + p_2\lambda^{n-2} + \dots + p_n}{\lambda^n + q_1\lambda^{n-1} + \dots + q_n}$$

entonces

$$f(\lambda) = \mathbf{p} (\lambda \mathbf{I} - \mathbf{C}(q))^{-1} \mathbf{e}_n$$

donde $\mathbf{p} = (p_n, p_{n-1}, \dots, p_1)$ y $\mathbf{e}_n = (0, 0, \dots, 0, 1)'$.

Demostración. Sabemos que

$$(\lambda \mathbf{I} - \mathbf{C}(q))^{-1} = \frac{\text{adj}(\lambda \mathbf{I} - \mathbf{C}(q))}{\det(\lambda \mathbf{I} - \mathbf{C}(q))}$$

donde $\text{adj}()$ es la matriz adjunta, que es la transpuesta de la matriz de cofactores. La matriz de cofactores K de A se define como

$$K_{i,j} = (-1)^{i+j} \det(A_{i,j})$$

donde $A_{i,j}$ es la matriz que se obtiene de quitar el i -ésimo renglón y la j -ésima columna a la matriz A , así $\text{adj}(A) = K'$. Al hacer el producto de $(\lambda \mathbf{I} - \mathbf{C}(q))^{-1} \mathbf{e}_n$ resulta un

vector cuyas entradas son iguales a la de la n -ésima columna de $\text{adj}(A) = K'$ por lo que solo nos fijaremos en los términos $k_{n,i}$ con $i \in \{1, 2, \dots, n\}$. Consideremos la matriz $(\lambda \mathbf{I} - \mathbf{C}(q))$ sin el n -ésimo renglón y sin la i -ésima columna

$$\left(\begin{array}{cccc|cccc} \lambda & -1 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda & 0 & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \lambda & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & -1 \end{array} \right)$$

cuyo determinante es $\lambda^{i-1}(-1)^{n-i}$. Así

$$\text{adj}(\lambda \mathbf{I} - \mathbf{C}(q))_{i,n} = (-1)^{n+i} \lambda^{i-1} (-1)^{n-i} = \lambda^{i-1}$$

entonces $(\lambda \mathbf{I} - \mathbf{C}(q))^{-1} \mathbf{e}_n = (1, \lambda, \dots, \lambda^{n-1})'$, ahora solo resta multiplicar el vector anterior por \mathbf{p} para obtener el resultado. \square

Corolario 2.2.1. *Los polos de*

$$f(\lambda) = \frac{p(\lambda)}{q(\lambda)} = \frac{p_1 \lambda^{n-1} + p_2 \lambda^{n-2} + \dots + p_n}{\lambda^n + q_1 \lambda^{n-1} + \dots + q_n}$$

son los valores propios de la matriz compañera $\mathbf{C}(q)$.

Definición 2.2.1. *Un polo dominante es un polo que tiene la mayor parte real y la mayor multiplicidad de los polos que tienen la mayor parte real.*

Veremos que si una variable aleatoria tiene transformada de Laplace racional entonces los polos de la transformada tienen parte real negativa, esto es importante para encontrar una forma de la densidad de la variable aleatoria.

Lema 2.2.1. *Si $f(x) = x^{n-1} e^{-ax}$ entonces su transformada de Laplace es*

$$\hat{f}(s) = \frac{(n-1)!}{(s+a)^n}$$

y es válido cuando $\text{Re}(s) > -\text{Re}(a)$

Demostración. Por inducción sobre n .

Si $n = 1$ y $\text{Re}(s+a) > 0$

$$\begin{aligned} \hat{f}(s) &= \int_0^\infty e^{-x(s+a)} dx \\ &= \frac{-e^{-x(s+a)}}{s+a} \Big|_{x=0}^\infty \\ &= \frac{1}{s+a} \end{aligned}$$

supongo que se cumple para $n - 2$ y que $Re(s + a) > 0$ entonces integrando por partes

$$\begin{aligned}\hat{f}(s) &= \int_0^\infty e^{-x(s+a)} x^{n-1} dx \\ &= \frac{-x^{n-1} e^{-x(s+a)}}{s+a} \Big|_{x=0}^\infty + \int_0^\infty \frac{e^{-x(s+a)} (n-1) x^{n-2}}{s+a} dx \\ &= \frac{(n-1)}{s+a} \int_0^\infty e^{-x(s+a)} x^{n-2} dx \\ &= \left(\frac{n-1}{s+a} \right) \frac{(n-2)!}{(s+a)^{n-1}}\end{aligned}$$

además solo se necesitan que $Re(s + a) > 0$ pues la parte imaginaria de $e^{-x(s+a)}$ está acotada. \square

Lema 2.2.2. Las transformadas de Laplace de $f_1(x) = x^{n-1} \cos(ax)$ y de $f_2(x) = x^{n-1} \sin(ax)$ son

$$\hat{f}_1(s) = (n-1)! Re \left(\frac{(s+ia)^n}{(s^2+a^2)^n} \right)$$

y

$$\hat{f}_2(s) = (n-1)! Im \left(\frac{(s+ia)^n}{(s^2+a^2)^n} \right)$$

Demostración. Si $f(x) = x^{n-1} e^{iax}$ por el lema anterior

$$\hat{f} = \frac{(n-1)!}{(s-ia)^n} = (n-1)! \left(\frac{s+ia}{s^2+a^2} \right)^n$$

pues si z es número complejo diferente de cero $z\bar{z} = |z|^2$ entonces el resultado surge de que $Re(f) = f_1$ y $Im(f) = f_2$. \square

Teorema 2.2.3. Sea X una variable aleatoria con transformada de Laplace racional

$$L_X(s) = \frac{p(s)}{q(s)}$$

entonces $L_X(s)$ tiene un polo dominante que es real.

Demostración. Del conjunto de polos de $L_X(s)$ sean $r_j = -\alpha + iw_j$ $j \in 1, 2, \dots, 2k$ los polos que tienen máxima parte real y sea r_0 un polo de máxima parte real. Si un número es polo también su conjugado pues ambos son raíces de q por eso se considera un número par de polos. Buscamos demostrar que r_0 es un polo real igual a $-\alpha$ cuya multiplicidad n_0 es a lo menos la máxima de las multiplicidades n_1, n_2, \dots, n_k de los polos r_1, r_2, \dots, r_{2k} , aquí solo definimos k multiplicidades pues los polos conjugados tienen la misma multiplicidad. Si $n_0 \geq \max\{n_1, n_2, \dots, n_k\}$ entonces tendríamos la existencia de un polo real que además será dominante. Supongamos lo contrario es decir $n_0 < n_i$ para alguna $i \in \{1, 2, \dots, k\}$ sin perder generalidad supongamos que

$i = 1$ y además supongamos que n_1 es el mayor que cumple $n_1 > n_0$. Si la multiplicidad de un polo $-\alpha + i\beta$ es n entonces también es la multiplicidad de su conjugado por lo que en el denominador de $L_X(s)$ aparecerá $(s + \alpha + i\beta)^n (s + \alpha - i\beta)^n = ((s + \alpha)^2 + \beta^2)^n$ por consiguiente podemos escribir

$$L_X(s) = \frac{p(s)}{((s + \alpha)^2 + w_1^2)^{n_1} \dots ((s + \alpha)^2 + w_k^2)^{n_k} (s + \alpha)^{n_0} q_1(s)}$$

donde q_1 es un polinomio cuyas raíces tienen parte real menor a $-\alpha$. Entonces

$$\hat{L}_X(s) = L_X(s - \alpha) = \frac{p(s)}{(s^2 + w_1^2)^{n_1} \dots (s^2 + w_k^2)^{n_k} s^{n_0} q_1(s - \alpha)}$$

ahora $q_i(s - \alpha)$ tiene raíces con parte real negativa. Por los lemas anteriores vemos que la transformada de Laplace inversa \hat{f} de $\hat{L}_X(s)$ puede ser escrita como

$$\hat{f}(s) = \sum_{i=1}^k (P_i(s) \sin(w_i s) + Q_i(s) \cos(w_i s)) + P_0(s) + \phi(s)$$

donde P_i, Q_i son polinomios de grado a lo más $n_i - 1$ y $\phi(s)$ es una función que tiende a cero exponencialmente rápido. Como n_1 es el exponente dominante podemos reescribir

$$\hat{f}(s) = s^{n_1-1} \left(\sum_{i=1}^k (p_i(s) \sin(w_i s) + q_i(s) \cos(w_i s)) + p_0(s) + \phi_1(s) \right)$$

donde las funciones q_i, p_i son acotadas y los límites

$$\bar{p}_i = \lim_{s \rightarrow \infty} \frac{P_i(s)}{s^{n_1-1}}$$

$$\bar{q}_i = \lim_{s \rightarrow \infty} \frac{Q_i(s)}{s^{n_1-1}}$$

existen y al menos uno de \bar{p}_1 o \bar{q}_1 es diferente de cero. Ahora consideramos únicamente la suma

$$g(s) = \sum_{i=1}^k (\bar{p}_i(s) \sin(w_i s) + \bar{q}_i(s) \cos(w_i s)).$$

Esta es una función periódica que cambia de signo (en una infinidad de puntos). Debido a que

$$\frac{\hat{f}(s)}{s^{n_1-1}} = g(s) + o(1)$$

cuando s tiende a infinito, entonces $\hat{f}(s)$ cambia de signo puesto que

$$\hat{L}_X(s) = \int_0^\infty e^{-(s-\alpha)x} f(x) dx = \int_0^\infty e^{-sx} (e^{ax} f(x)) dx$$

tenemos que $\hat{f}(x) = e^{ax} f(x)$, esto implica que f la densidad de X cambia de signo lo que contradice que f sea densidad de probabilidad. \square

Teorema 2.2.4. *Si X tiene transformada de Laplace racional el polo dominante es estrictamente negativo. Esto implica que los valores propios de la matriz compañera del denominador tienen parte real estrictamente negativa.*

Demostración. Sea λ el polo dominante real. Puede haber otros polos, digamos $2N$, que tienen la misma parte real

$$a_i = \lambda + i\sigma_i \quad b_i = \bar{a}_i$$

pues los polos aparecen en parejas conjugadas. Sea I el conjunto de índices de máxima parte real λ . Por la demostración del teorema anterior podemos escribir la densidad f de X como

$$f(x) = C(x)e^{\lambda x} + \sum_{i \in I} (Q_i(x)e^{a_i x} + R_i(x)e^{b_i x}) + \sum_{i \notin I} P_i(x)e^{\lambda_i x}$$

para algunos polinomios C, P_i y Q_i . Si suponemos $\lambda > 0$ entonces $e^{\lambda x} \rightarrow \infty$ cuando $x \rightarrow \infty$ y C no se anula entonces no se satisface $f \rightarrow 0$ cuando $x \rightarrow \infty$. Si $\lambda = 0$ tenemos dos posibilidades $C(x)$ es constante o $C(x)$ tiene orden mayor a una constante entonces no se satisface $f \rightarrow 0$ cuando $x \rightarrow \infty$, lo mismo sucede con P_i y Q_i en caso de existir. Así la única posibilidad es que $\lambda < 0$. La segunda parte se sigue del corolario 2.2.1 y de lo ya demostrado. \square

Corolario 2.2.2. *Si X tiene transformada de Laplace racional entonces está bien definida para números complejos cuya parte real es mayor o igual a cero.*

Demostración. La transformada es racional por lo que no estará bien definida en los polos y por el teorema anterior todos tienen parte real estrictamente negativa. \square

Corolario 2.2.3. *Para la matriz compañera \mathbf{C} correspondiente a una transformada de Laplace racional de una variable aleatoria X se tiene*

$$e^{\mathbf{C}x} \rightarrow \mathbf{0}$$

cuando $x \rightarrow \infty$, donde $\mathbf{0}$ es una matriz de ceros.

Demostración. Se sigue de que los valores propios de \mathbf{C} tienen parte real negativa y de aplicar la función exponencial a la representación canónica de Jordan de la proposición (A.1.4) de \mathbf{C} . \square

Ahora buscamos la forma de la densidad de una variable aleatoria no negativa cuya transformada de Laplace sea racional. Para eso se utilizarán algunas definiciones pertenecientes al cálculo funcional, que nos permiten evaluar cierta clase de operadores en funciones analíticas (holomorfas) este tema es tratado a detalle en [3].

Sea A un operador acotado en un espacio de Banach $(X, \|\cdot\|)$. Llamamos operador resolvente a $(z\mathbf{I} - A)^{-1}$, el conjunto de valores $z \in \mathbb{C}$ para los cuales el resolvente está bien definido (existe es densamente definido y acotado) se llama conjunto resolvente. El espectro de A es el complemento del conjunto resolvente.

Sea E espacio vectorial. Si $A : D(A) \rightarrow E$ con $D(A) \leq E$ y $\dim(E) < \infty$ entonces A se puede representar por una matriz y en este caso el espectro de A es el conjunto de valores propios.

Definición 2.2.2. Sea A un operador acotado en un espacio de Banach $(X, \|\cdot\|)$ y sea γ un camino simple y cerrado que encierra al espectro de A . Si f es analítica en el interior y sobre γ , se define

$$f(A) = \frac{1}{2\pi i} \oint_{\gamma} f(z)(z\mathbf{I} - A)^{-1} dz$$

Ahora encontraremos la forma de la densidad de una variable cuya transformada de Laplace sea racional utilizando la definición anterior.

Teorema 2.2.5. Sea X una variable aleatoria no negativa con transformada de Laplace dada por

$$L_X(\lambda) = \frac{p(\lambda)}{q(\lambda)} = \frac{p_1\lambda^{n-1} + p_2\lambda^{n-2} + \dots + p_n}{\lambda^n + q_1\lambda^{n-1} + \dots + q_n}$$

entonces X tiene densidad de la forma

$$f(x) = \mathbf{p}e^{\mathbf{C}(q)x} \mathbf{e}_n$$

donde $\mathbf{p} = (p_n, p_{n-1}, \dots, p_1)$ y $\mathbf{e}_n = (0, 0, \dots, 0, 1)'$.

Demostración. Todos los polos de L_X tienen parte real estrictamente negativa y son un número finito. Sea γ un camino simple cerrado que encierra a los polos. Como todas las singularidades son los polos, la transformada inversa de Laplace está dada por

$$f(x) = \frac{1}{2\pi i} \oint_{\gamma} e^{zx} L_X(z) dz$$

por el teorema 2.2.2 tenemos lo siguiente

$$f(x) = \frac{1}{2\pi i} \mathbf{p} \oint_{\gamma} e^{zx} (z\mathbf{I} - \mathbf{C}(q))^{-1} dz \mathbf{e}_n$$

ahora por la definición anterior

$$f(x) = \mathbf{p}e^{\mathbf{C}(q)x} \mathbf{e}_n$$

□

A la representación anterior de la densidad se le llama forma canónica.

Teorema 2.2.6. Si X tiene transformada de Laplace racional, entonces la matriz compañera del denominador $\mathbf{C}(q)$ es invertible y el α -ésimo momento de la distribución está dado por

$$\mu_{\alpha} = E(X^{\alpha}) = \Gamma(\alpha + 1) \mathbf{p}(-\mathbf{C}(q))^{-(\alpha+1)} \mathbf{e}_n$$

Demostración. El determinante de $\mathbf{C}(q)$ es $(-1)^{n-1}q_n$ y $L_X(0) = \frac{p_n}{q_n} = 1$ entonces $q_n \neq 0$ por lo tanto $\mathbf{C}(q)$ es invertible. $\mathbf{C}(q)$ tiene valores propios con parte real estrictamente negativa por lo que $-\mathbf{C}(q)$ tiene sus valores propios en el semiplano positivo, entonces se puede encontrar un camino simple cerrado que contenga su espectro. Además la transformada de Laplace es analítica en su dominio de convergencia por lo que podemos hacer $L_{X^\alpha}(-\mathbf{C}(q))$ como en la definición 2.2.2.

$$\begin{aligned}\mu_\alpha &= \int_0^\infty x^\alpha \mathbf{p} e^{\mathbf{C}(\mathbf{q})x} \mathbf{e}_n dx \\ &= \mathbf{p} L_{X^\alpha}(-\mathbf{C}(\mathbf{q})) \mathbf{e}_n \\ &= \Gamma(\alpha + 1) \mathbf{p} (-\mathbf{C}(q))^{-(\alpha+1)} \mathbf{e}_n\end{aligned}$$

pues $L_{X^\alpha}(s) = \frac{\Gamma(\alpha+1)}{s^{\alpha+1}}$ que es convergente para números complejos cuya parte real es positiva. \square

2.2.2 Distribución matriz exponencial

Definición 2.2.3. Sea τ una variable aleatoria con densidad

$$f_\tau(t) = \alpha e^{\mathbf{S}t} s$$

con $t \geq 0$ para algún vector α , un vector columna s y una matriz \mathbf{S} que admiten entradas complejas. Si X es una variable aleatoria con tal distribución diremos que X se distribuye matriz exponencial y se denotara por $\tau \sim ME_n(\alpha, \mathbf{S}, s)$ donde n es la dimensión de los vectores. El triplete (α, \mathbf{S}, s) es llamado representación de la distribución y n el orden.

Notemos que las singularidades de $u \rightarrow (u\mathbf{I} - \mathbf{S})^{-1}$ son los valores propios de \mathbf{S} . El problema aquí es que la dimensión de \mathbf{S} puede ser mayor que el grado del polinomio en el denominador de la siguiente expresión

$$\alpha(u\mathbf{I} - \mathbf{S})^{-1} \mathbf{s} = \frac{p(u)}{q(u)}.$$

Ahora buscamos encontrar una representación minimal que coincida con el grado de el polinomio q . Si la representación es minimal entonces los valores propios de \mathbf{S} coinciden con las raíces de q , es decir, los polos de la transformada de Laplace.

Podemos asumir sin perder generalidad que los valores propios de \mathbf{S} tienen parte real negativa.

Teorema 2.2.7. La variable aleatoria X se distribuye matriz exponencial si y sólo si tiene transformada de Laplace racional.

Demostración. \Leftarrow) Si X tiene transformada de Laplace racional por el teorema 2.2.5 X se distribuye matriz exponencial.

⇒) Si X se distribuye matriz exponencial entonces tiene densidad de la forma

$$\begin{aligned}\alpha e^{\mathbf{S}x} s &= \frac{1}{2\pi i} \alpha \oint_{\gamma} e^{zx} (z\mathbf{I} - \mathbf{S})^{-1} dz s \\ &= \frac{1}{2\pi i} \oint_{\gamma} e^{zx} \alpha (z\mathbf{I} - \mathbf{S})^{-1} s dz\end{aligned}$$

donde γ es un camino simple cerrado que encierra las singularidades de $z \rightarrow (z\mathbf{I} - \mathbf{S})^{-1}$ entonces

$$L_X(z) = \alpha (z\mathbf{I} - \mathbf{S})^{-1} s$$

que es racional. □

A la representación $(\mathbf{p}, \mathbf{C}(q), \mathbf{e}_n)$ como en el teorema 2.2.5 de la transformada de Laplace racional de una variable aleatoria X la llamaremos representación canónica.

Definición 2.2.4. *El grado de una distribución matriz exponencial con transformada de Laplace $L_X(s) = \frac{p(s)}{q(s)}$, donde $L_X(s)$ es irreducible es el grado de $q(s)$.*

Una representación (α, \mathbf{S}, s) de una matriz exponencial no necesariamente es minimal, bastaría con que su transformada de Laplace se pudiera reducir y con el teorema 2.2.5 encontrar otra representación con la transformada ya reducida para tener otra representación de orden menor. Por eso veremos la siguiente definición.

Definición 2.2.5. *Si el orden de (α, \mathbf{S}, s) es igual al grado entonces la representación (α, \mathbf{S}, s) es de orden minimal.*

Corolario 2.2.4. *Una distribución matriz exponencial puede ser expresada por una representación de orden minimal.*

Demostración. Para su transformada de Laplace irreducible, la forma canónica del teorema 2.2.5 es minimal. □

Teorema 2.2.8. *Sea $\tau \sim ME_p(\alpha, \mathbf{S}, s)$ entonces su función de distribución es*

$$F(x) = 1 + \alpha e^{\mathbf{S}x} \mathbf{S}^{-1} s,$$

su función generadora de momentos es

$$M_X(r) = \alpha (-r\mathbf{I} - \mathbf{S})^{-1} s,$$

y su α -ésimo momento es

$$\mu_\beta = E(X^\beta) = \Gamma(\beta) \alpha (-\mathbf{S})^{-(\beta+1)} s$$

Demostración. Por el corolario anterior podemos tomar a \mathbf{S} invertible entonces

$$1 = \int_0^\infty \alpha e^{\mathbf{S}x} s dx = \alpha [\mathbf{S}^{-1} e^{\mathbf{S}x}] \Big|_0^\infty s = \alpha [-\mathbf{S}^{-1}] s$$

y

$$\begin{aligned}
F(t) &= \int_0^t \alpha e^{\mathbf{S}x} s dx \\
&= \alpha [\mathbf{S}^{-1} e^{\mathbf{S}x}] \Big|_0^t s \\
&= \alpha [\mathbf{S}^{-1} e^{\mathbf{S}t}] s - \alpha [\mathbf{S}^{-1}] s \\
&= 1 + \alpha [\mathbf{S}^{-1} e^{\mathbf{S}t}] s
\end{aligned}$$

La función generadora de momentos se encuentra de forma análoga a la distribución y el α -ésimo momento utilizando calculo funcional como en el teorema 2.2.6. \square

Teorema 2.2.9. *Para toda distribución matriz exponencial es posible elegir una representación minimal (α, \mathbf{S}, s) tal que $s = -\mathbf{S}\mathbf{e}$ y $\alpha\mathbf{e} = 1$, además α y \mathbf{S} son real valuadas.*

Demostración. Empezamos con una representación canónica que es minimal de dimensión n . Sea M una matriz no singular. Entonces tenemos que

$$\begin{aligned}
f(x) &= \mathbf{p}e^{\mathbf{C}(q)x}\mathbf{e}_n \\
&= \mathbf{p}M e^{M^{-1}\mathbf{C}(q)Mx} M^{-1}\mathbf{e}_n
\end{aligned}$$

Se requiere que $-M^{-1}\mathbf{C}(q)M\mathbf{e} = M^{-1}\mathbf{e}_n$ o equivalentemente $M\mathbf{e} = -\mathbf{C}(q)^{-1}\mathbf{e}_n$. Ahora

$$\mathbf{C}(q)^{-1} = \begin{pmatrix} -\frac{q_{n-1}}{q_n} & -\frac{q_{n-2}}{q_n} & \cdots & -\frac{q_1}{q_n} & -\frac{1}{q_n} \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

entonces

$$M\mathbf{e} = -\mathbf{C}(q)^{-1}\mathbf{e}_n = \left(\frac{1}{q_n}, 0, \dots, 0 \right)'$$

buscamos una matriz M que sume $1/q_n$ en la primer columna, cero en las demás y sea no singular, por ejemplo

$$M = \begin{pmatrix} \frac{1}{q_n} & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ -1 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$\mathbf{p}M\mathbf{e} = (p_n, p_{n-1}, \dots, p_1) \left(\frac{1}{q_n}, 0, \dots, 0 \right)' = p_n/q_n = 1$$

\square

Capítulo 3

Teoría de renovación

Lo referente a teoría de renovación está basado en [1] y [2], y la parte de teoría de renovación tipo fase en [2].

Los procesos de renovación son procesos puntuales con tiempos entre arribos idénticamente distribuidos excepto el primero en algunos casos. La mayoría de la teoría se desarrollará en términos de integrales de Lebesgue-Stieltjes de la forma

$$\int_0^{\infty} a(x)dF(x)$$

donde a es una función y F es una función de distribución. Si F es absolutamente continua lo anterior es una integral y si F es discreta la expresión será una suma. Para la mayoría de la teoría no es necesario distinguir entre el caso discreto y continuo y por ello se utilizan estas integrales que abarcan ambos casos.

Sean T_1, T_2, \dots i.i.d. $\sim F$ con $F(0) = 0$ y suponemos que las variables T_i son estrictamente positivas. Para $n \geq 1$, $S_n = T_1 + T_2 + \dots + T_n$ son los tiempo de llegada y $\{S_n\}_{n \in \mathbb{N}}$ es llamado proceso de renovación. Los T_i son tiempos de interarribo o distribución entre arribos. Es imposible tener varios arribos al mismo tiempo pues $F(0) = 0$. En algunos casos es conveniente suponer que el tiempo del primer arribo tiene distribución diferente a las demás, cuando estemos en este caso definiremos otro arribo S_0 que tiene función de distribución G que puede ser una distribución degenerada en cero, es decir $S_0 = 0$ casi seguramente. Sea T_0 el tiempo hasta el arribo S_0 aquí ($T_0 = S_0$).

Si $S_0 = 0$ al proceso de renovación se le llama puro o cero-retrasado, mientras que si S_0 es no degenerado en cero se llamará proceso de renovación retrasado.

3.1 Proceso de renovación puro

Definimos el proceso de conteo $\{N(t)\}_{t \geq 0}$ de un proceso de renovación puro como

$$N(t) = \inf\{n | S_n > t\} = \sup\{n | S_n < t\} + 1$$

$N(t)$ es el número de llegadas hasta el tiempo t , incluyendo uno en cero que no es un arribo real. Se define la función de renovación $U(t) = E(N(t))$. Existe la siguiente

relación entre los tiempos de llegada y el proceso de conteo

$$N(t) \leq n \text{ si y solo si } S_n > t$$

aplicando esta equivalencia a la función de renovación

$$\begin{aligned} U(t) &= \sum_{n=0}^{\infty} P(N(t) > n) \\ &= \sum_{n=0}^{\infty} P(S_n \leq t) \\ &= 1 + \sum_{n=1}^{\infty} F^{*n}(t) \\ &= \sum_{n=0}^{\infty} F^{*n}(t) \end{aligned}$$

donde $*$ denota a la convolución, $F^{*0} = 1$, $F^{*2} = F * F$, $F^{*n} = F * F^{n-1}$.

Definición 3.1.1. Si F es absolutamente continua entonces U es diferenciable y $u(t) = U'(t)$ es llamada densidad de renovación del proceso de renovación.

u no es una densidad en el sentido clásico de probabilidad pero tiene la siguiente propiedad $u(t)dt$ es la probabilidad de tener un arribo en el intervalo $[t, t + dt)$ pues $F(0) = 0$ entonces puede ocurrir a lo más un arribo en $[t, t + dt)$ dos o más tienen probabilidad $o(t)$, si denotamos como $1_{\{[t, t+dt)\}}$ a la indicadora de tener un arribo en el intervalo $[t, t + dt)$ entonces

$$u(t)dt = U(t + dt) - U(t) = E(N[t, t + dt]) = E(1_{\{[t, t+dt)\}}) = P([t, t + dt))$$

Teorema 3.1.1. $U(t) < \infty$ para toda $t > 0$

Demostración.

$$\begin{aligned} F^{*(n+m)}(t) &= F^{*n} * F^{*m}(t) \\ &= \int_0^t F^{*n}(t-x) dF^{*m}(x) \\ &\leq F^{*n}(t) F^{*m}(t) \end{aligned}$$

Sea $n = mr + k$ para alguna r fija entonces

$$F^{*n}(t) = F^{*(mr+k)}(t) \leq [F^{*r}(t)]^m F^{*k}(t)$$

Como las variables T_i son positivas podemos elegir $r \in \mathbb{N}$ tal que $P(T_1 + T_2 + \dots + T_r > t) > 0$ y $F^{*r}(t) < 1$ entonces

$$\begin{aligned} U(t) &= \sum_{n=0}^{\infty} F^{*n}(t) \\ &= \sum_{m=0}^{\infty} \sum_{k=0}^{r-1} F^{*(mr+k)}(t) \\ &\leq \sum_{k=0}^{r-1} F^{*k}(t) \sum_{m=0}^{\infty} [F^{*r}(t)]^m < \infty \end{aligned}$$

□

Corolario 3.1.1. $F^{*n}(t) \rightarrow 0$ cuando $n \rightarrow \infty$

Demostración. Por el teorema anterior tenemos que $\sum_{n=0}^{\infty} F^{*n}(t)$ converge y por lo tanto la cola de la serie converge a cero. □

Si condicionamos al tiempo del primer arribo $T_1 = x$ en la función de renovación

$$\begin{aligned} U(t) &= E(N(t)) \\ &= \int_0^{\infty} E(N(t)|T_1 = x) dF(x) \\ &= \int_0^t E(N(t)|T_1 = x) dF(x) + \int_t^{\infty} 1 dF(x) \\ &= 1 - F(t) + \int_0^t 1 + U(t-x) dF(x) \\ &= 1 - F(t) + F(t) + \int_0^t U(t-x) dF(x) \\ &= 1 + U * F(t) \end{aligned}$$

de lo anterior U satisface

$$U = 1 + U * F.$$

que es una ecuación de renovación.

Condicionar al tiempo de primer arribo se llama argumento de renovación. En general las ecuaciones de renovación son de la forma

$$A(t) = a(t) + \int_0^t A(t-x) dF(x)$$

donde a es una función conocida y A es una función desconocida que resuelve la ecuación.

Teorema 3.1.2. Sea $a(t)$ una función que es acotada en intervalos acotados. Entonces existe solución única $A(t)$ a la ecuación

$$A(t) = a(t) + \int_0^t A(t-x)dF(x)$$

que es acotada en intervalos acotados. La solución es

$$A(t) = \int_0^t a(t-x)dU(x) = U * a(t)$$

Demostración. Primero veamos que la solución propuesta cumple con la ecuación

$$\begin{aligned} A(t) &= U * a(t) \\ &= a(t) + \sum_{n=1}^{\infty} F^{*n} * a(t) \\ &= a(t) + F * \left(\sum_{n=0}^{\infty} F^{*n} * a(t) \right) \\ &= a(t) + F * A(t) \end{aligned}$$

Veamos que la solución es acotada en intervalos acotados

$$\begin{aligned} \sup_{0 \leq t \leq T} |A(t)| &= \sup_{0 \leq t \leq T} \left| \int_0^t a(t-x)dU(x) \right| \\ &\leq \int_0^T \sup_{0 \leq s \leq T} |a(s)|dU(x) \\ &= \sup_{0 \leq s \leq T} |a(s)|U(T) \end{aligned}$$

Para la unicidad supongamos otra solución B que satisface $B = a + F * B$

$$\begin{aligned} |A(t) - B(t)| &= |F * A(t) - F * B(t)| \\ &= |F * (A - B)(t)| \\ &= \dots \\ &= |F^{*n} * (A - B)(t)| \\ &= \left| \int_0^t (A - B)(t-x)dF^{*n}(x) \right| \\ &\leq \int_0^t \sup_{0 \leq s \leq t} |(A - B)(s)|dF^{*n}(x) \\ &\leq \sup_{0 \leq s \leq t} |(A - B)(s)|F^{*n}(t) \rightarrow 0 \end{aligned}$$

si $n \rightarrow \infty$ pues A y B son acotados en intervalos acotados. □

Veamos una aplicación del teorema anterior

Teorema 3.1.3. $E(S_{N(t)}) = E(T_1)U(t)$

Demostración. Consideremos el tiempo esperado hasta el primer arribo después de t , $E(S_{N(t)})$. Sea $A(t) = E(S_{N(t)})$ condicionando a que $T_1 = x$ si $x > t$ entonces $S_{N(t)} = x$ en caso contrario $E(S_{N(t)}|T_1 = x) = x + A(t - x)$.

$$\begin{aligned} A(t) &= \int_0^t x + A(t - x)dF(x) + \int_t^\infty x dF(x) \\ &= E(T_1) + \int_0^t A(t - x)dF(x) \end{aligned}$$

lo anterior es una ecuación de renovación con $a(t) = E(T_1)$ constante y por lo tanto acotada, del teorema anterior $A(t) = a*U(t) = E(T_1)U(t)$ y en efecto $A(t)$ está acotada, pues $S_{N(t)}$ es el primer arribo después de t , haciendo un argumento de renovación con el tiempo de interarribo en el que se encuentra t y teniendo en cuenta que $S_{N(t)-1} \leq t < S_{N(t)}$ llegamos a que $P(S_{N(t)} > x) \leq P(T_1 + t > x)$ que es equivalente a $P(S_{N(t)} - t > x) \leq P(T_1 > x)$ ahora integrando tenemos que $A(t)$ es acotada en intervalos acotados. \square

Teorema 3.1.4. (*Teorema elemental de renovación*) Si $\{N(t)\}_{t \geq 0}$ es un procesos de conteo de un proceso de renovación con tiempos de interarribo con media finita y positiva μ , entonces

$$\frac{N(t)}{t} \rightarrow \frac{1}{\mu}$$

y

$$\frac{U(t)}{t} \rightarrow \frac{1}{\mu}$$

casi en todo punto cuando $t \rightarrow \infty$.

Demostración. Como $S_{N(t)-1} \leq t < S_{N(t)}$ entonces

$$\frac{1}{N(t)} \sum_{j=1}^{N(t)-1} T_j \leq \frac{t}{N(t)} < \frac{1}{N(t)} \sum_{j=1}^{N(t)} T_j$$

haciendo $t \rightarrow \infty$ tenemos el primer resultado. Para el segundo $t < S_{N(t)}$ y por el teorema anterior

$$t < E(S_{N(t)}) = \mu U(t)$$

entonces

$$\liminf_{t \rightarrow \infty} \frac{U(t)}{t} \geq \frac{1}{\mu}$$

Para establecer la otra igualdad acotemos las variables T_i por una constante $c > 0$

$$T_i^c = \begin{cases} T_i & \text{si } T_i \leq c \\ c & \text{si } T_i > c \end{cases}$$

Ahora consideremos el proceso de conteo $\{N^c(t)\}_{t \geq 0}$ con tiempos de interarribo T_i^c . Como T_i^c son acotadas por c tenemos que $S_{N^c(t)} \leq t + c$ y entonces

$$t + c \geq E(S_{N^c(t)}) = \mu^c U^c(t)$$

donde μ^c es la media de T_i^c y $U^c(t)$ su correspondiente función de renovación. Por otro lado $T_i^c \leq T_i$ implica que $N^c(t) \geq N(t)$ para todo $t > 0$ y esto implica que $U^c(t) \geq U(t)$ con esto

$$t + c \geq \mu^c U^c(t) \geq \mu^c U(t)$$

que es equivalente a

$$\frac{1 + c/t}{\mu^c} \geq \frac{U(t)}{t}$$

tomando limite

$$\limsup \frac{U(t)}{t} \leq \frac{1}{\mu^c}$$

que es válido para toda c entonces

$$\limsup \frac{U(t)}{t} \leq \lim_{c \rightarrow \infty} \frac{1}{\mu^c}$$

y si $T_i \sim F$

$$\begin{aligned} \lim_{c \rightarrow \infty} \mu^c &= \lim_{c \rightarrow \infty} \int_0^\infty P(T_i^c > x) dx \\ &= \lim_{c \rightarrow \infty} \int_0^c P(T_i^c > x) dx \\ &= \lim_{c \rightarrow \infty} \int_0^c (1 - F(x)) dx \\ &= \int_0^\infty (1 - F(x)) dx = \mu \end{aligned}$$

entonces tenemos

$$\limsup \frac{U(t)}{t} \leq \frac{1}{\mu} \leq \liminf \frac{U(t)}{t}$$

□

3.2 Procesos de renovación estacionarios y retrasados

Los procesos de renovación retrasados son procesos de renovación en los que el primer tiempo no tiene la misma distribución que los demás. Sea $T_0 \sim G$ y $T_1, T_2, \dots \sim F$ *i.i.d.*

Sea $S_0 = T_0$ y $S_n = T_0 + T_1 + \dots + T_n$ llamamos a $\{S_n\}_{\mathbb{N}^0}$ proceso de renovación retrasado. Sea $N(t)$ el número de arribos en el proceso retrasado y

$$U_D(t) = E(N(t))$$

es importante notar la diferencia entre U y U_D que corresponden a un proceso puro y a uno retrasado respectivamente.

Aplicando un argumento de renovación tenemos

$$\begin{aligned} U_D(t) &= E(N(t)) \\ &= \int_0^\infty E(N(t)|T_0 = x)dG(x) \\ &= \int_0^t 0 + E(N(t)|T_0 = x)dG(x) + \int_t^\infty 0dG(x) \\ &= \int_0^t U(t-x)dG(x) \end{aligned}$$

como la convolución es conmutativa

$$U_D = U * G = G * U$$

entonces

$$\begin{aligned} U_D(t) &= G * U(t) \\ &= G * (1 + F * U)(t) \\ &= G(t) + G * F * U(t) \\ &= G(t) + F * U_D(t) \end{aligned}$$

Para los procesos estacionarios supongamos que en el tiempo que un observador arribó el proceso ya había iniciado una cantidad infinita de tiempo atrás. ¿La distribución del tiempo hasta el primer arribo después de que el observador llegó es igual a la de los demás tiempos de interarribo? Para que eso sucediera necesitaríamos que para cualquier intervalo de tiempo con longitud fija el número de llegadas en ese intervalo sea igual sin importar en que momento se considere ese intervalo.

Definición 3.2.1. *Un proceso de renovación es estacionario si*

$$\{N(t+s) - N(t)\} \stackrel{D}{=} \{N(s)\}$$

Por la definición de procesos de renovación estacionarios se tiene que $U_D(t) = ct$ pues

$$U_D(t+s) = U(t)_D + U_D(s)$$

Veamos un teorema que nos dice como debe ser la distribución del primer tiempo de arribo para que el procesos de renovación retrasado sea estacionario.

Teorema 3.2.1. *Un proceso de renovación es estacionario si es retrasado y la distribución del primer tiempo es*

$$G(t) = \frac{1}{\mu} \int_0^t (1 - F(x)) dx$$

Si F es absolutamente continua entonces la densidad de la distribución G del primer tiempo es

$$g(x) = \frac{1}{\mu} (1 - F(x))$$

Demostración.

$$U_D(t) = G(t) + \int_0^t G(t-x) dF(x)$$

y $U_D(t) = ct$ si y solo si

$$G(t) = ct - \int_0^t c(t-x) dF(x)$$

integrando por partes

$$\int_0^t c(t-x) dF(x) = [c(t-x)F(x)]_0^t + \int_0^t cF(x) dx$$

así

$$G(t) = ct - \int_0^t cF(x) dx = c \int_0^t (1 - F(x)) dx$$

Si G es no defectuosa es decir, $G(\infty) = 1$ si $Y \sim F$ entonces $1 = c \int_0^\infty (1 - F(x)) dx = c \int_0^\infty P(Y > x) dx = c\mu$ así tenemos el primer resultado para el segundo solo se deriva el primero. \square

También es sencillo encontrar la densidad de renovación si el proceso de renovación es estacionario

$$E(N(t, t+h)) = U_D(t+h) - U_D(t) = (t+h)/\mu - t/\mu$$

entonces

$$U'_D(t) = 1/\mu$$

3.3 Teorema de renovación de Blackwell

Para el siguiente teorema de renovación distinguiremos dos tipos de distribuciones las lattice y no-lattice.

Definición 3.3.1. *Sea F una función de distribución, x es llamado punto de incremento de F si para todo $\epsilon > 0$*

$$F(x) - F(x - \epsilon) > 0$$

Definición 3.3.2. Una distribución es llamada *lattice* (o aritmética) con generador $d > 0$ si todos los puntos de incremento de F están contenidos del $d\mathbb{N}$ y d es el mayor número para el que lo anterior es verdadero.

Ejemplo: La distribución Poisson tiene sus incrementos en los naturales entonces $d = 1$

Ahora demostraremos que la hipótesis de ser o no distribución *lattice* asegura que si se inician dos procesos retrasados independientes a la larga los tiempos de arribo estarán arbitrariamente cercanos. Si iniciamos un proceso de renovación que tiene distribución *lattice* para sus tiempos de interarribo con generador d y el mismo pero retrasado en $d/2$ los tiempos de arribo no estarán más cerca que $d/2$ entonces la condición de ser *lattice* es necesaria.

Lema 3.3.1. Una distribución F que está concentrada en n números racionales r_1, r_2, \dots, r_n es *lattice*.

Demostración. Sea $r_i = n_i/D_i$ para $i \in 1, 2, \dots, n$ si hacemos $d' = 1/(D_1 D_2 \dots D_n)$ entonces F tendrá sus incrementos en $d'\mathbb{N}$ pero aún podría ser $d > d'$ \square

Lema 3.3.2. Sea F una distribución concentrada en los números racionales $r_i = n_i/D_i$, $i = 1, 2, \dots$ si $D_i \leq D < \infty$ para alguna $D \in \mathbb{N}$ entonces F es *lattice*.

Demostración. Como D_i son acotadas por D en el peor de los casos tomando $d = 1/D! > 0$ aseguramos que los puntos de incremento de F se encuentran en $d\mathbb{N}$ aún podría ser $d > 1/D!$ \square

Lema 3.3.3. Sea F no-*lattice* y z_1 un punto de incremento de F . Sea $y \geq 0$. Entonces para toda $\epsilon > 0$ existen números naturales k_1 y k_2 y otro punto de incremento z_2 tal que la distancia entre $y + k_1 z_1$ y $k_2 z_2$ es menor a ϵ es decir

$$|y + k_1 z_1 - k_2 z_2| < \epsilon$$

Demostración. Supongamos que $\epsilon > 0$ y que todos los puntos de incremento de F son racionales. Entonces para cualquier punto de incremento z_2 tenemos

$$\frac{z_2}{z_1} = \frac{N}{D}$$

que es equivalente a $z_2 = z_1 N/D$, suponemos que N y D no tienen primos en común, por el lema anterior el número de puntos de incremento (racionales) de F no puede ser finito y D no es acotado entonces podemos hacer que $D > z_1/\epsilon$, existen números $k'_1, k'_2 \in \mathbb{Z}$ tales que $k'_2 N - k'_1 D = 1$ y entonces

$$k'_1 z_1 - k'_2 z_2 = k'_1 z_1 - k'_2 z_1 N/D = -z_1/D$$

definimos

$$k_1 = \left[\frac{yD}{z_1} \right] k'_1$$

$$k_2 = \left[\frac{yD}{z_1} \right] k'_2$$

aquí $[x]$ es la función parte entera de x . Entonces

$$\begin{aligned} y + k_1 z_1 - k_2 z_2 &= y + \left[\frac{yD}{z_1} \right] k'_1 z_1 - \left[\frac{yD}{z_1} \right] k'_2 z_2 \\ &= y + \left[\frac{yD}{z_1} \right] k'_1 z_1 - \left[\frac{yD}{z_1} \right] \left(k'_1 z_1 + \frac{z_1}{D} \right) \\ &= \frac{yD}{z_1} \left(\frac{z_1}{D} \right) - \left[\frac{yD}{z_1} \right] \left(\frac{z_1}{D} \right) \end{aligned}$$

como $(yD/z_1) - [yD/z_1] \in [0, 1)$ concluimos que

$$0 \leq y + k_1 z_1 - k_2 z_2 \leq \frac{z_1}{D} < \epsilon$$

En el caso de que el punto de incremento z_2 sea tal que z_2/z_1 es irracional lo aproximamos con números racionales para esto elegimos k'_1 y k'_2 números naturales tales que $0 < y' = k'_2 z_2 - k'_1 z_1 < \epsilon$ entonces con

$$k_1 = \left[\frac{y}{y'} \right] k'_1$$

$$k_2 = \left[\frac{y}{y'} \right] k'_2$$

$$\begin{aligned} 0 &\leq y + k_1 z_1 - k_2 z_2 \\ &= y + \left[\frac{y}{y'} \right] k'_1 z_1 - \left[\frac{y}{y'} \right] (y' + k'_1 z_1) \\ &= y' \left(\frac{y}{y'} \right) - \left[\frac{y}{y'} \right] y' \leq y' < \epsilon \end{aligned}$$

□

Por el lema anterior si iniciamos dos procesos de renovación diferentes uno arbitrario y otro en y , los tiempos de llegada son cercanos después de cierto número k_1 de llegas para un proceso y k_2 para el otro y esto ocurre con probabilidad positiva.

Teorema 3.3.1. (Blackwell) *Para un proceso de renovación con tiempos de interarribo distribuidos F no-lattice con media $0 < \mu < \infty$ la función de renovación $U(t)$ satisface*

$$U(t+a) - U(t) \rightarrow \frac{a}{\mu}$$

cuando $t \rightarrow \infty$

Demostración. Ver la página 155 de [1]

□

3.4 Teorema principal de renovación

Es necesario conocer la definición de integración de Riemann directa pues es una hipótesis del teorema principal de renovación.

Definición 3.4.1. Sea $z : \mathbb{R} \rightarrow \mathbb{R}^+$ una función. Definimos

$$z_h^{\min}(n) = \inf\{z(t) | nh \leq t < (n+1)h\}$$

$$z_h^{\max}(n) = \sup\{z(t) | nh \leq t < (n+1)h\}$$

para toda $n \in \mathbb{N}$ y las "sumas de Riemann"

$$S^{\min}(h) = h \sum_{n \in \mathbb{N}} z_h^{\min}(n)$$

$$S^{\max}(h) = h \sum_{n \in \mathbb{N}} z_h^{\max}(n).$$

Decimos que z es directamente Riemann integrable si las sumas anteriores convergen para toda $h > 0$ y

$$\lim_{h \rightarrow 0} S^{\min}(h) = \lim_{h \rightarrow 0} S^{\max}(h)$$

el límite común es denotado igual que la integral de Riemann como $\int_0^\infty z(x)dx$. Si z toma valores negativos decimos que es directamente Riemann integrable si z_+ y z_- lo son y hacemos $z = z_+ + z_-$

Teorema 3.4.1. (Teorema principal de renovación.) Sea F no-lattice con media μ y no defectuosa la distribución de tiempos de interraibo. Si $z \geq 0$ es directamente Riemann integrable entonces

$$\lim_{t \rightarrow \infty} U * z(t) = \frac{1}{\mu} \int_0^\infty z(s)ds.$$

Sea F no defectuosa, lattice con generador $d > 0$ y con media μ . Si $z \geq 0$ y

$$\sum_{k=0}^{\infty} z(x + kd) < \infty$$

entonces

$$\lim_{n \rightarrow \infty} U * z(x + nd) = \frac{d}{\mu} \sum_{k=0}^{\infty} z(x + kd).$$

Demostración. Ver página 155 de [1] □

3.5 Proceso de edad y residual de vida

Definición 3.5.1. Sea $\{N(t)\}_{t>0}$ el proceso de conteo de un proceso de renovación (posiblemente retrasado) con tiempos de llegadas S_0, S_1, \dots . Se definen los siguientes procesos:

1. Residual de vida $R_t = S_{N(t)} - t$, es decir el tiempo transcurrido desde t hasta el siguiente arribo.
2. Proceso de edad $A_t = t - S_{N(t)-1}$, es el tiempo transcurrido hasta t desde el arribo antes de t .
3. Tiempo total de vida o spread $S_t = A_t + R_t = S_{N(t)} - S_{N(t)-1}$, es el tiempo entre dos arribos consecutivos tal que t está en ese intervalo.

La distribución de los procesos anteriores se puede encontrar usando la función de renovación.

Teorema 3.5.1.

$$P(A_t > x) = \begin{cases} \bar{G}(t) + \int_0^{t-x} \bar{F}(t-z) dU_D(z) & x < t \\ 0 & t \leq x \end{cases}$$

$$P(R_t > y) = \bar{G}(t+y) + \int_0^t \bar{F}(y+t-z) dU_D(z)$$

$$P(A_t > x, R_t > y) = \begin{cases} \bar{G}(t+y) + \int_0^{t-x} \bar{F}(y+t-z) dU_D(z) & x < t \\ 0 & t \leq x \end{cases}$$

$$P(S_t > x) = \begin{cases} \bar{G}(t) + \int_0^{t-x} \bar{F}(t-z) dU_D(z) + \int_{t-x}^t \bar{F}(x) dU_D(z) & x < t \\ \bar{G}(x) + \int_0^t \bar{F}(x) dU_D(z) & t \leq x \end{cases}$$

Demostración. Primero notemos las siguientes equivalencias si $t - x > 0$

$$R_{t-x} > x \Leftrightarrow \text{No hay renovación en } [t-x, t] \Leftrightarrow A_t > x$$

de lo anterior $P(A_t > x)$ vale cero en $[t-x, t]$. Supongamos un proceso de renovación puro, condicionando al tiempo del primer arribo T_1

$$\begin{aligned} P(A_t > x) &= \int_0^\infty P(A_t > x | T_1 = z) dF(z) \\ &= \int_0^{t-x} P(A_{t-z} > x) dF(z) + \int_t^\infty dF(z) \end{aligned}$$

pues

$$P(A_t > x | T_1 = z) = \begin{cases} P(A_{t-z} > x) & z < t-x \\ 0 & t-x \leq z \leq t \\ 1 & z > t \end{cases}$$

si definimos $H_x(t) = P(A_t > x)$ entonces tenemos

$$H_x(t) = \bar{F}(t) + \int_0^{t-x} H_x(t-z) dF(z) = \bar{F} + (H_x 1_{\{x < t\}}) * F(t)$$

cuya solución es

$$H_x(t) = 1_{\{x < t\}} \bar{F} * U(t)$$

por el teorema(3.1.2). Si ahora consideramos la posibilidad de un proceso retrasado tenemos dos casos, si al tiempo t ya ocurrió alguna renovación, sólo faltaría reemplazar en la ecuación que encontramos la función de renovación de un proceso puro por la de uno retrasado y el otro caso es que no han ocurrido renovaciones al tiempo t y solo restaría sumar esta probabilidad, es decir

$$H_x(t) = \bar{G}(t) + \bar{F} * U_D(t)$$

que es la primera expresión.

Por la nota al principio de la demostración tenemos que

$$R_t > x \Leftrightarrow A_{t+x} > x$$

que demuestra la segunda expresión.

Para la tercera expresión

$$R_{t-x} > x + y \Leftrightarrow S_{N(t-x)} - t > y$$

y los eventos $R_{t-x} > x + y$ y $A_t \leq x$ no pueden ocurrir al mismo tiempo entonces tenemos la siguiente igualdad de conjuntos

$$\{R_{t-x} > x + y\} \Leftrightarrow \{R_{t-x} > x + y, A_t > x\}$$

si $A_t > x$ es verdadero entonces $N(t-x) = N(t)$ que da la siguiente equivalencia

$$\{R_{t-x} > x + y\} \Leftrightarrow \{R_t > y, A_t > x\}$$

que demuestra la tercer expresión.

Para la tercera si $x > t$ y no hay arribos en $(0, t)$ entonces el tiempo sin arribos debe ser mayor a x o si el último arribo ocurrió en $z < t$ entonces el intervalo que inicia en z hasta el próximo arribo debe ser mayor a x , es decir

$$P(S_t > x) = \bar{G}(x) + \int_0^t \bar{F}(x) dU_D(z).$$

Si $x < t$ y no han ocurrido arribos el tiempo del primer arribo puede ser a lo menos t , si el último arribo antes de t ocurrió en $z < t - x$ entonces la longitud del intervalo debe ser al menos $t - z$, si el último arribo ocurrió en $(t - x, t)$ entonces la longitud del intervalo es al menos x , es decir

$$P(S_t > x) = \bar{G}(t) + \int_0^{t-x} \bar{F}(t-z) dU_D(z) + \int_{t-x}^t \bar{F}(x) dU_D(z)$$

□

Corolario 3.5.1. Para distribuciones no-lattice F los procesos A_t, R_t y S_t tienen distribución límite cuando $t \rightarrow \infty$. Las correspondientes distribuciones límite A_∞, R_∞ y S_∞ son:

$$\begin{aligned}
 P(A_\infty > x) &= P(R_\infty > x) = \mu^{-1} \int_x^\infty \bar{F}(z) dz \\
 P(A_\infty > x, R_\infty > y) &= \mu^{-1} \int_{x+y}^\infty \bar{F}(z) dz \\
 P(S_\infty > x) &= \mu^{-1} \int_x^\infty \bar{F}(z) dz + \mu^{-1} x \bar{F}(x) = \mu^{-1} \int_0^\infty \bar{F}(\max(x, z)) dz
 \end{aligned}$$

Si F es absolutamente continua entonces la densidad de A_∞ y R_∞ es $\mu^{-1} \bar{F}(x)$. Si F tiene densidad f entonces las densidades de $\{A_\infty, R_\infty\}$ y S_∞ existen y son $\mu^{-1} f(x+y)$ y $\mu^{-1} x f(x)$ respectivamente.

Demostración. Cuando $t \rightarrow \infty$ el término \bar{F} se anula. Para resolver las integrales se utiliza el teorema principal de renovación (3.4.1) que también es válido para procesos retrasados. La función $\bar{F}(z)1_{z>x}$ hace el papel de z en el teorema de renovación para el caso de A_t, R_t y S_t para S_t también se utiliza el teorema de Blackwell, para el caso de la distribución conjunta de $\{A_\infty > x, R_\infty > y\}$ se utiliza $\bar{F}(z) > x+y$ o que $\{R_{t-x} > x+y\} \Leftrightarrow \{R_t > y, A_t > x\}$. Para encontrar las densidades solo hay que derivar. \square

La distribución límite de A_t y R_t es igual a la distribución inicial de un proceso estacionario. Esto está acorde con la interpretación de que un proceso estacionario inicio mucho tiempo atrás cuando inicio su observación. La distribución límite del tiempo de vida es llamada distribución del primer momento.

Teorema 3.5.2. En un proceso de renovación estacionario la distribución del tiempo de vida (spread) tiene la distribución del primer momento con densidad dada por

$$f_{S_t} = \frac{x f(x)}{\mu}$$

De lo anterior podemos notar que S_t no tiene densidad f , a pesar de que los tiempos entre llegadas tengan esa densidad, esto se conoce como "paradoja de inspección". Como el proceso es estacionario podemos considerar cualquier tiempo t aleatorio y elegir el intervalo entre dos arribos que contiene a t este intervalo tendrá densidad $x f(x)/\mu$ y no f .

3.6 Teoría de renovación tipo fase

Veamos algunas propiedades de proceso de renovación con tiempos de interarribo distribuidos tipo fase i.i.d $PH_p(\pi, \mathbf{T})$.

Teorema 3.6.1. *Para un proceso de renovación tipo fase puro con tiempo de interarribo $PH_p(\pi, \mathbf{T})$ la densidad de renovación es*

$$u(x) = \pi e^{(\mathbf{T} + \mathbf{t}\pi)x} \mathbf{t}$$

Demostración. Sean T_1, T_2, \dots los tiempos de interarribo i.i.d $PH_p(\pi, \mathbf{T})$. Concatenando los procesos de Markov asociados a las T_i hasta el tiempo de la absorción obtenemos un proceso de Markov con submatriz de intensidades $\mathbf{T} + \mathbf{t}\pi$, para ver esto pensemos en un pequeño intervalo de tiempo en ese intervalo tenemos dos tipo de cambio de estado, los que ocurren en el mismo T_i y los que ocurren cuando se absorbe un T_i e inicia T_{i+1} , los del primer tipo ocurren con las intensidades de \mathbf{T} y para las del segundo tipo supongamos un cambio de i a j entonces T_i termina con intensidad \mathbf{t}_i y T_{i+1} inicia de con probabilidad π_j , los eventos anterior son excluyentes entonces concluimos que los cambios ocurren con intensidad $\mathbf{T} + \mathbf{t}\pi$.

$u(x)$ es la probabilidad de tener una renovación en $[x, x + dx)$ que es equivalente a que el proceso asociado entre al estado absorbente condicionando al estado actual tenemos

$$\begin{aligned} u(x)dx &= \sum_{i,j=1}^p \pi_i p_{ij}^x \mathbf{t}_j dx \\ &= \sum_{i,j=1}^p \pi_i (e^{(\mathbf{T} + \mathbf{t}\pi)x})_{ij} \mathbf{t}_j dx \\ &= \pi e^{(\mathbf{T} + \mathbf{t}\pi)x} \mathbf{t} \end{aligned}$$

□

Teorema 3.6.2. *Considere un proceso de renovación retrasado con $T_0 \sim PH_p(\alpha, \mathbf{T})$ y T_1, T_2, \dots i.i.d $PH_p(\pi, \mathbf{T})$, entonces la densidad de renovación es*

$$u(x) = \alpha e^{(\mathbf{T} + \mathbf{t}\pi)x} \mathbf{t}$$

Demostración. La demostración es igual a la del teorema anterior pero el proceso inicia con distribución α en vez de π . □

El siguiente teorema muestra que el tiempo transcurrido entre el inicio de una observación y el siguiente arribo es tipo fase.

Teorema 3.6.3. *El tiempo residual de vida R_t de un proceso de renovación tipo fase se distribuye tipo fase. Si el proceso es puro con T_1, T_2, \dots i.i.d $PH_p(\pi, \mathbf{T})$ entonces*

$$R_x \sim PH_p(\pi e^{(\mathbf{T} + \mathbf{t}\pi)x}, \mathbf{T})$$

y si el proceso es retrasado con distribución inicial $T_0 \sim PH_p(\alpha, \mathbf{T})$ entonces

$$R_x \sim PH_p(\alpha e^{(\mathbf{T} + \mathbf{t}\pi)x}, \mathbf{T})$$

Demostración. La distribución del proceso al tiempo x es $\pi e^{(\mathbf{T}+\mathbf{t}\pi)x}$ en el caso puro y $\alpha e^{(\mathbf{T}+\mathbf{t}\pi)x}$ en el caso retrasado y esta distribución sirve como inicial para una tipo fase que inicia al tiempo x . \square

Teorema 3.6.4. *Un proceso de renovación estacionario, con T_1, T_2, \dots i.i.d $PH_p(\pi, \mathbf{T})$ es un proceso de renovación retrasado cuyo tiempo de primer arribo es*

$$T_0 \sim PH_p \left(\frac{\pi(-\mathbf{T})^{-1}}{\pi(-\mathbf{T})^{-1}\mathbf{e}}, \mathbf{T} \right)$$

Demostración. Por el teorema (3.2.1) la densidad del primer arribo debe ser

$$\begin{aligned} f(x) &= \frac{1}{\mu} (1 - F(x)) \\ &= \frac{1}{\pi(-\mathbf{T})^{-1}\mathbf{e}} (1 - (1 - \pi e^{\mathbf{T}x}e)) \\ &= \frac{1}{\pi(-\mathbf{T})^{-1}\mathbf{e}} (\pi e^{\mathbf{T}x}e) \end{aligned}$$

pero la expresión anterior no es la densidad de una distribución tipo fase, para que se cumpla se necesita que el último término de la multiplicación sea $\mathbf{t} = -\mathbf{T}e$ si multiplicamos por $-\mathbf{T}$ y su inversa no afectamos la expresión y además llegamos a lo deseado, ocupando el hecho de que la exponencial de una matriz conmuta con la matriz y con su inversa tenemos que

$$\begin{aligned} \frac{1}{\pi(-\mathbf{T})^{-1}\mathbf{e}} (\pi e^{\mathbf{T}x}e) &= \frac{1}{\pi(-\mathbf{T})^{-1}\mathbf{e}} (\pi(-\mathbf{T})^{-1}e^{\mathbf{T}x}(-\mathbf{T})e) \\ &= \frac{\pi(-\mathbf{T})^{-1}}{\pi(-\mathbf{T})^{-1}\mathbf{e}} (e^{\mathbf{T}x}\mathbf{t}) \end{aligned}$$

\square

Capítulo 4

Distribuciones de momentos tipo fase y matriz exponencial

En este capítulo se demostrará que las distribuciones tipo fase y matriz exponencial son cerradas bajo las distribuciones de momentos de orden natural, es decir si $n \in \mathbb{N}$ entonces la distribución del n -ésimo momento de una distribución tipo fase es tipo fase y análogamente con las distribuciones matriz exponencial, lo referente a este fue extraído de [4] y de [8].

Definición 4.0.1. Sea f la función de densidad de una variable aleatoria no negativa X entonces

$$f_i(x) = \frac{x^i f(x)}{\mu_i} \text{ donde } \mu_i = \int_0^\infty x^i f(x) dx$$

con $i \in \mathbb{N}$ son las densidades de algunas variables aleatorias $X^{(i)}$ cuando μ_i exista. Llamamos a f_i la densidad del i -ésimo momento de f .

4.1 Distribuciones de momentos matriz exponencial

Si f es la densidad de una distribución matriz exponencial y \hat{f} denota la transformada de Laplace de f es fácil ver que los momentos naturales son distribuidos matriz exponencial pues

$$\frac{d^i}{ds^i} \hat{f}(s) = \int_0^\infty (-1)^i x^i e^{-sx} f(x) dx = (-1)^i \mu_i \hat{f}_i(s)$$

y \hat{f} es una función racional entonces también su i -ésima derivada y por lo tanto $\hat{f}_i(s)$ (para llegar a la primera igualdad se hace un intercambio de integral con derivada que es verdadero pues $e^{-sx} f(x)$ es continua en las variables s y x positivas además las derivadas con respecto a s son continuas).

Ahora buscaremos representaciones para estas distribuciones.

De aquí en adelante, si hablamos de una distribución $ME(\alpha, \mathbf{S}, s)$ supondremos que se cumple $s = -\mathbf{S}e$ y $ae = 1$ como en el teorema (2.2.9)

Teorema 4.1.1. Sea (α, \mathbf{S}, s) la representación de una distribución matriz exponencial, entonces el n -ésimo momento es también matriz exponencial con representación $(\alpha_n, \mathbf{S}_n, s_n)$ donde

$$\alpha_n = \left(\frac{\alpha \mathbf{S}^{-n}}{\alpha \mathbf{S}^{-n} e}, 0, \dots, 0 \right) \quad \mathbf{S}_n = \begin{pmatrix} \mathbf{S} & -\mathbf{S} & 0 & \dots & 0 \\ 0 & \mathbf{S} & -\mathbf{S} & \dots & 0 \\ 0 & 0 & \mathbf{S} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{S} \end{pmatrix} \quad s_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ s \end{pmatrix}$$

y \mathbf{S}_n es una matriz de dimensión $(n+1)p \times (n+1)p$.

Demostración. Sea $X \sim ME_p(\alpha, \mathbf{S}, s)$ entonces

$$\mu_n = \Gamma(n) \alpha (-\mathbf{S})^{-(n+1)} s = \Gamma(n) \alpha (-\mathbf{S})^{-(n+1)} (-\mathbf{S}) e = n! \alpha (-\mathbf{S})^{-n} e$$

consideremos

$$f_n(x) = \frac{x^n \alpha e^{\mathbf{S}x} s}{\mu_n}.$$

Sea

$$er_k(x; \lambda) = \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}$$

la densidad de la distribución de Erlang de orden k .

Como se vio en el ejemplo (2.1.2) la distribución de Erlang es una distribución matriz exponencial (tipo fase) que puede ser escrita como

$$er_k(x; \lambda) = (1, 0, \dots, 0) \exp \left(\begin{pmatrix} \begin{pmatrix} -s & s & 0 & \dots & 0 \\ 0 & -s & s & \dots & 0 \\ 0 & 0 & -s & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -s \end{pmatrix} x \right) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ s \end{pmatrix}$$

en lo anterior la dimensión de las matrices es k . La función $z \rightarrow er_k(x; z)$ es analítica en el semiplano positivo y el espectro de $-\mathbf{S}$ esta en el semiplano positivo entonces podemos utilizar la definición (2.2.2) completando la densidad f_n a una densidad de

Erlang como sigue

$$\begin{aligned}
f_n(x) &= \frac{x^n \alpha e^{\mathbf{S}x} s}{\mu_n} \\
&= \frac{\alpha (-\mathbf{S})^{-n-1} (-\mathbf{S})^{n+1} x^n e^{\mathbf{S}x} s}{\alpha (-\mathbf{S})^{-n} e n!} \\
&= \frac{\alpha (-\mathbf{S})^{-n-1}}{\alpha (-\mathbf{S})^{-n} e} er_{n+1}(x, -\mathbf{S}) s \\
&= \frac{\alpha (-\mathbf{S})^{-n-1}}{\alpha (-\mathbf{S})^{-n} e} (\mathbf{I}, 0, \dots, 0) \exp \left(\left(\begin{pmatrix} \mathbf{S} & -\mathbf{S} & 0 & \dots & 0 \\ 0 & \mathbf{S} & -\mathbf{S} & \dots & 0 \\ 0 & 0 & \mathbf{S} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{S} \end{pmatrix} x \right) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\mathbf{S} \end{pmatrix} s \right) \\
&= \left(\frac{\alpha (-\mathbf{S})^{-n}}{\alpha (-\mathbf{S})^{-n} e}, 0, \dots, 0 \right) \exp \left(\left(\begin{pmatrix} \mathbf{S} & -\mathbf{S} & 0 & \dots & 0 \\ 0 & \mathbf{S} & -\mathbf{S} & \dots & 0 \\ 0 & 0 & \mathbf{S} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{S} \end{pmatrix} x \right) - \mathbf{S}^{-1} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\mathbf{S} \end{pmatrix} s \right)
\end{aligned}$$

en lo anterior la dimensión de las matrices es $(n+1)p \times (n+1)p$ ($n+1$ bloques de tamaño p). \square

Corolario 4.1.1.

$$F_n(x) = 1 - \frac{\alpha (-\mathbf{S})^{-n}}{\alpha (-\mathbf{S})^{-n} e} \sum_{i=0}^n \frac{(-\mathbf{S}x)^i}{i!} e^{\mathbf{S}x} e$$

Demostración. Tenemos que

$$F_n(x) = 1 + \alpha_n e^{\mathbf{S}n x} \mathbf{S}^{-1} s = 1 - \alpha_n e^{\mathbf{S}n x} e$$

y con un desarrollo análogo al del corolario (A.1.4)

$$e^{\mathbf{S}n x} = e^{\mathbf{S}x} \begin{pmatrix} \mathbf{I} & \frac{-\mathbf{S}x}{1!} & \frac{(-\mathbf{S}x)^2}{2!} & \dots & \frac{(-\mathbf{S}x)^{n-1}}{(n-1)!} & \frac{(-\mathbf{S}x)^n}{(n)!} \\ 0 & \mathbf{I} & \frac{-\mathbf{S}x}{1!} & \dots & \frac{(-\mathbf{S}x)^{n-2}}{(n-2)!} & \frac{(-\mathbf{S}x)^{n-1}}{(n-1)!} \\ 0 & 0 & \mathbf{I} & \dots & \frac{(-\mathbf{S}x)^{n-3}}{(n-3)!} & \frac{(-\mathbf{S}x)^{n-2}}{(n-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{I} & \frac{-\mathbf{S}x}{1!} \\ 0 & 0 & 0 & \dots & 0 & \mathbf{I} \end{pmatrix}$$

y ahora desarrollando las multiplicaciones se llega al resultado. \square

Si la distribución es tipo fase con representación (α, \mathbf{S}) , entonces la representación para la distribución del n -ésimo momento encontrada en el teorema anterior no es tipo fase pues no sólo tiene elementos negativos en la diagonal.

4.2 Distribuciones de momentos tipo fase

Ahora probaremos que la distribución del n -ésimo momento de una distribución tipo fase es tipo fase, para eso es suficiente probar que la distribución del primer momento es tipo fase y luego obtener la distribución del n -ésimo momento repitiendo el desarrollo. Para el siguiente desarrollo necesitaremos la representación revertida en el tiempo de una distribución tipo fase, sea (α, \mathbf{S}, s) la representación de una matriz exponencial y M una matriz no singular entonces $(\alpha M^{-1}, M\mathbf{S}M^{-1}, Ms)$ es otra representación de la misma distribución, en el caso de las distribuciones tipo fase \mathbf{S} es el generador de un proceso de Markov y como se vio en la sección (1.3) si $M = \Delta(m)$ con m una medida estacionaria ($\Delta(w)$ denota una matriz con el vector w en la diagonal) y $m_i \neq 0$ entonces $M\mathbf{S}M^{-1}$ es el generador del proceso con tiempo al revés.

Para una distribución tipo fase (α, \mathbf{S}) la representación revertida en el tiempo es $(\widehat{\alpha}, \widehat{\mathbf{S}})$, dada por $\widehat{\alpha} = s'M$, $\widehat{\mathbf{S}} = M^{-1}\mathbf{S}M$ y $m = \alpha(-\mathbf{S})^{-1}$. Lo último es debido a que en un proceso de renovación estacionario tipo fase cuyos tiempos de interarribo tienen representación (α, \mathbf{S}) el generador del proceso de Markov asociado al proceso de renovación es $\mathbf{S} + s\alpha$ y se satisface

$$\alpha(-\mathbf{S})^{-1}(\mathbf{S} + s\alpha) = 0$$

entonces por el teorema(1.2.10), $\alpha(-\mathbf{S})^{-1}$ es una medida estacionaria.

Teorema 4.2.1. *Sea (α, \mathbf{S}) la representación de una distribución tipo fase entonces la distribución del primer momento es tipo fase y una representación es $(\widehat{\alpha}_1, \widehat{\mathbf{S}}_1)$ dada por*

$$\widehat{\alpha}_1 = \left(s' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right), 0 \right)$$

$$\widehat{\mathbf{S}}_1 = \begin{pmatrix} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) & \frac{1}{\alpha(-\mathbf{S})^{-1}e} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \Delta (\alpha(-\mathbf{S})^{-1}) \\ 0 & \Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) \end{pmatrix}$$

Demostración. Consideremos un proceso de renovación estacionario con tiempos de interarribo tipo fase (α, \mathbf{S}) . Entonces la distribución de los estados del proceso asociado al tiempo t es $\pi_1 = \frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}$. Por los teoremas (3.5.1) y (3.6.3) los procesos R_t y A_t son distribuidos tipo fase con representación (π_1, \mathbf{S}) o $(\pi_1, \widehat{\mathbf{S}})$ con $\widehat{\mathbf{S}} = \Delta^{-1}(\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1})$. Por el teorema (3.5.2) $A_t + R_t$ es la distribución del primer momento. La representación de $A_t + R_t$ se puede encontrar usando argumentos de distribuciones tipo fase como sigue, el estado al tiempo t del proceso de Markov es el estado antes de la absorción de A_t y el estado inicial de R_t entonces para obtener una distribución tipo fase debemos asegurar que el estado inicial de R_t es el último estado antes de la absorción de A_t o si consideramos el proceso de reversa que el estado inicial de \widehat{A}_t es el último estado antes de la absorción de \widehat{R}_t .

La representación de \widehat{R}_t es decir la representación revertida en el tiempo de (π_1, \mathbf{S}) que tiene vector estacionario $\pi_1(-\mathbf{S})^{-1}$ es

$$\left(s' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right), \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \right)$$

cuyo vector de salida es

$$\begin{aligned}
& \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) e \\
&= \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right)' \\
&= \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \mathbf{S} \right)' \\
&= \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \frac{(\alpha(-\mathbf{S})^{-1})'}{\alpha(-\mathbf{S})^{-1}e}
\end{aligned}$$

Con la intensidad que termina \widehat{R}_t inicia \widehat{A}_t por lo que solo hace falta concatenarlos de acuerdo a la intensidad de salida de \widehat{R}_t esto se logra poniendo este vector en diagonal al concatenar los procesos. \square

La representación anterior se hizo con el tiempo al revés ahora veremos la representación anterior revertida en el tiempo es decir la representación con tiempo normal.

Corolario 4.2.1. *Sea (α, \mathbf{S}) la representación de una distribución tipo fase. La distribución de su primer momento se puede representar por una distribución tipo fase (α_1, \mathbf{S}_1) con*

$$\begin{aligned}
\alpha_1 &= \left(\frac{\alpha \Delta(-\mathbf{S}^{-1}e)}{\alpha(-\mathbf{S})^{-1}e}, 0 \right) \\
\mathbf{S}_1 &= \begin{pmatrix} \Delta^{-1}(-\mathbf{S}^{-1}e) \mathbf{S} \Delta(-\mathbf{S}^{-1}e) & \Delta(-\mathbf{S}^{-1}e) \\ 0 & \mathbf{S} \end{pmatrix}
\end{aligned}$$

Demostración. Encontremos la representación revertida en el tiempo de la representación encontrada en el teorema anterior es decir el vector inicial $(-\widehat{\mathbf{S}}_1 e)' \Delta(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1})$ y el generador $\Delta^{-1}(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1}) \widehat{\mathbf{S}}_1' \Delta(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1})$. Encontremos la inversa de una matriz por bloques triangular similar al generador encontrado en el teorema anterior

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}$$

entonces

$$\begin{aligned}
& \widehat{\mathbf{S}}_1^{-1} \\
&= \begin{pmatrix} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) (\mathbf{S}^{-1})' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) & \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) (\mathbf{S}^{-2})' \Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \\ 0 & \Delta^{-1} (\alpha(-\mathbf{S})^{-1}) (\mathbf{S}^{-1})' \Delta (\alpha(-\mathbf{S})^{-1}) \end{pmatrix}
\end{aligned}$$

entonces

$$\begin{aligned}
\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1} &= \left(s'(\mathbf{S}^{-1})'\Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right), s'(\mathbf{S}^{-2})'\Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \right) \\
&= \left(e'\Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right), (\mathbf{S}^{-1}e)'\Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \right) \\
&= \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}, (\mathbf{S}^{-1}e)'\Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \right)
\end{aligned}$$

con lo anterior

$$\Delta(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1}) = \begin{pmatrix} \Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ 0 & \Delta(\mathbf{S}^{-1}e)\Delta\left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}\right) \end{pmatrix}$$

ahora podemos encontrar el generador

$$\begin{aligned}
&\Delta^{-1}(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1})\widehat{\mathbf{S}}_1'\Delta(\widehat{\alpha}_1(-\widehat{\mathbf{S}}_1)^{-1}) \\
&= \begin{pmatrix} \Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ 0 & \Delta(\mathbf{S}^{-1}e)\Delta\left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}\right) \end{pmatrix}^{-1} \\
&\times \begin{pmatrix} \Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right)\mathbf{S}'\Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & \frac{1}{\alpha(-\mathbf{S})^{-1}e}\Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right)\Delta(\alpha(-\mathbf{S})^{-1}) \\ 0 & \Delta^{-1}(\alpha(-\mathbf{S})^{-1})\mathbf{S}'\Delta(\alpha(-\mathbf{S})^{-1}) \end{pmatrix}' \\
&\times \begin{pmatrix} \Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ 0 & \Delta(\mathbf{S}^{-1}e)\Delta\left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}\right) \end{pmatrix} \\
&= \begin{pmatrix} \Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ 0 & \Delta^{-1}(\mathbf{S}^{-1}e)\Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}\right) \end{pmatrix} \\
&\times \begin{pmatrix} \Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right)\mathbf{S}\Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ \frac{1}{\alpha(-\mathbf{S})^{-1}e}\Delta^{-1}\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right)\Delta(\alpha(-\mathbf{S})^{-1}) & \Delta(\alpha(-\mathbf{S})^{-1})\mathbf{S}\Delta^{-1}(\alpha(-\mathbf{S})^{-1}) \end{pmatrix} \\
&\times \begin{pmatrix} \Delta\left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e}\right) & 0 \\ 0 & \Delta(\mathbf{S}^{-1}e)\Delta\left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e}\right) \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{S} & 0 \\ \Delta^{-1}(\mathbf{S}^{-1}e) & \Delta^{-1}(\mathbf{S}^{-1}e)\mathbf{S}\Delta(\mathbf{S}^{-1}e) \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
& -\widehat{\mathbf{S}}_1 e \\
&= - \begin{pmatrix} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) & \frac{1}{\alpha(-\mathbf{S})^{-1}e} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \Delta (\alpha(-\mathbf{S})^{-1}) \\ 0 & \Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) \end{pmatrix} e \\
&= \begin{pmatrix} -\Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \mathbf{S}' \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) e & -\frac{1}{\alpha(-\mathbf{S})^{-1}e} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \Delta (\alpha(-\mathbf{S})^{-1}) e \\ -\Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) e & \end{pmatrix} \\
&= \begin{pmatrix} -\Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \mathbf{S}' \right)' & -\frac{1}{\alpha(-\mathbf{S})^{-1}e} \Delta^{-1} \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) (\alpha(-\mathbf{S})^{-1})' \\ -\Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) e & \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ -\Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) e \end{pmatrix}
\end{aligned}$$

entonces

$$\begin{aligned}
& (-\widehat{\mathbf{S}}_1 e)' \Delta (\widehat{\alpha}_1 (-\widehat{\mathbf{S}}_1)^{-1}) \\
&= \left(0, -(\Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \mathbf{S}' \Delta (\alpha(-\mathbf{S})^{-1}) e)' \right) \begin{pmatrix} \Delta \left(\frac{\alpha(-\mathbf{S})^{-2}}{\alpha(-\mathbf{S})^{-1}e} \right) & 0 \\ 0 & \Delta (\mathbf{S}^{-1}e) \Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \end{pmatrix} \\
&= \left(0, -e' \Delta (\alpha(-\mathbf{S})^{-1}) \mathbf{S} \Delta^{-1} (\alpha(-\mathbf{S})^{-1}) \Delta (\mathbf{S}^{-1}e) \Delta \left(\frac{\alpha(-\mathbf{S})^{-1}}{\alpha(-\mathbf{S})^{-1}e} \right) \right) \\
&= \left(0, -\alpha(-\mathbf{S})^{-1} \mathbf{S} \Delta (\mathbf{S}^{-1}e) \frac{1}{\alpha(-\mathbf{S})^{-1}e} \right) \\
&= \left(0, \frac{\alpha \Delta (\mathbf{S}^{-1}e)}{\alpha(-\mathbf{S})^{-1}e} \right)
\end{aligned}$$

reacomodando los estados de forma que los de A_t sigan a los de R_t se tiene el resultado. \square

Si ahora suponemos que los tiempos de interarribo tienen la distribución del primer momento tipo fase entonces $A_t + R_t$ tendrá la distribución del segundo momento de esta manera podemos encontrar la distribución del n -ésimo momento sin embargo si la distribución principal tiene orden p entonces la distribución del primer momento tiene orden $2p$ y al seguir obteniendo las distribuciones de orden n llegamos a que estas tienen orden $2^n p$. Veremos una representación alternativa de dimensión $(n+1)p \times (n+1)p$.

Teorema 4.2.2. *Sea (α, \mathbf{S}) la representación de una distribución tipo fase entonces la distribución del n -ésimo momento es tipo fase y una representación es $(\widehat{\alpha}_n, \widehat{\mathbf{S}}_n)$ dada*

por

$$\widehat{\alpha}_n = \left(\frac{\rho_{n+1}}{\rho_n} s' \Delta(\pi_{n+1}), 0, \dots, 0 \right)$$

$$\widehat{\mathbf{S}}_n = \begin{pmatrix} \mathbf{C}_{n+1} & \mathbf{D}_{n+1} & 0 & \dots & 0 & 0 \\ 0 & \mathbf{C}_n & \mathbf{D}_n & \dots & 0 & 0 \\ 0 & 0 & \mathbf{C}_{n-1} & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{C}_2 & \mathbf{D}_2 \\ 0 & 0 & 0 & \dots & 0 & \mathbf{C}_1 \end{pmatrix}$$

con

$$\rho_i = \alpha(-\mathbf{S})^{-i} e \quad \pi_i = \frac{\alpha(-\mathbf{S})^{-i}}{\rho_i}$$

$$\mathbf{C}_i = \Delta^{-1}(\pi_i) \mathbf{S}' \Delta(\pi_i) \quad \text{y} \quad \mathbf{D}_i = \frac{\rho_{i-1}}{\rho_i} \Delta^{-1}(\pi_i) \Delta(\pi_{i-1})$$

Demostración. Probaremos que la transformada de Laplace de $(\widehat{\alpha}_n, \widehat{\mathbf{S}}_n)$ es igual a la del teorema(4.1.1). Notemos que $(u\mathbf{I} - \Delta^{-1}\mathbf{S}'\Delta)^{-1} = \Delta^{-1}(u\mathbf{I} - \mathbf{S}')^{-1}\Delta$ y que $\frac{\rho_{i-1}}{\rho_i} \Delta^{-1}(\pi_i) \pi_{i-1}$ es el vector de salida del generador \mathbf{C}_i entonces el vector de salida de $\widehat{\mathbf{S}}_n$ solo tiene la última entrada (bloque) distinta de cero entonces al desarrollar la transformada de Laplace $\widehat{\alpha}_n (u\mathbf{I} - \widehat{\mathbf{S}}_n)^{-1} \widehat{s}_n$ solo es de importancia el bloque superior derecho de la matriz $(u\mathbf{I} - \widehat{\mathbf{S}}_n)^{-1}$ y utilizando el teorema(A.1.1) tenemos

$$\begin{aligned} L(u) &= (-1)^{n+2} \left(\frac{\rho_{n+1}}{\rho_n} s' \Delta(\pi_{n+1}) \right) \\ &\times \left(\prod_{i=0}^{n-1} (u\mathbf{I} - \Delta^{-1}(\pi_{n+1-i}) \mathbf{S}' \Delta(\pi_{n+1-i}))^{-1} \Delta^{-1}(\pi_{n+1-i}) \Delta(\pi_{n-i}) \frac{\rho_{n-i}}{\rho_{n+1-i}} \right) \\ &\times (u\mathbf{I} - \Delta^{-1}(\pi_1) \mathbf{S}' \Delta(\pi_1))^{-1} \Delta^{-1}(\pi_1) \pi'_0 \frac{1}{\rho_1} \\ &= (-1)^n \rho_n^{-1} s' (u\mathbf{I} - \mathbf{S}')^{-n-1} \alpha' \\ &= (-1)^n \rho_n^{-1} \alpha (u\mathbf{I} - \mathbf{S})^{-n-1} s \\ &= (-1)^n \pi_n [(u\mathbf{I} - \mathbf{S})^{-1} (-\mathbf{S})]^n (u\mathbf{I} - \mathbf{S})^{-1} s \end{aligned}$$

que coincide con la transformada de la representación vista en el teorema(4.1.1). \square

Corolario 4.2.2. Sea (α, \mathbf{S}) la representación de una distribución tipo fase entonces la distribución del n -ésimo momento es tipo fase y una representación es (α_n, \mathbf{S}_n) dada

por

$$\alpha_n = (\phi_n^{-1} \alpha \Delta(\rho_n), 0, \dots, 0)$$

$$\mathbf{S}_n = \begin{pmatrix} \Delta^{-1}(\rho_n) \mathbf{S} \Delta(\rho_n) & \Delta^{-1}(\rho_n) \Delta(\rho_{n-1}) & 0 & \dots & 0 \\ 0 & \Delta^{-1}(\rho_{n-1}) \mathbf{S} \Delta(\rho_{n-1}) & \Delta^{-1}(\rho_{n-1}) \Delta(\rho_{n-2}) & \dots & 0 \\ 0 & 0 & \Delta^{-1}(\rho_{n-2}) \mathbf{S} \Delta(\rho_{n-2}) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \Delta^{-1}(\rho_1) \\ 0 & 0 & 0 & \dots & \mathbf{S} \end{pmatrix}$$

con

$$\rho_i = (-\mathbf{S})^{-i} e \quad \phi_i = \alpha \rho_i$$

Demostración. La demostración se puede hacer encontrando la representación revertida en el tiempo del teorema anterior como en el corolario (4.2.1) o encontrando su transformada de Laplace como en la demostración del teorema anterior. \square

4.3 Índice de Gini y curva de Lorenz para distribuciones matriz exponencial

La curva de Lorenz es una representación gráfica de la función de distribución empírica de la riqueza o ingresos, esta fue desarrollada por Max O. Lorenz en 1905 para representar la desigualdad de la distribución de la riqueza. El coeficiente de Gini es una medida que representa la distribución de ingresos de los residentes de un país y es comúnmente usada como medida de desigualdad, este coeficiente fue desarrollado por Corrado Gini en 1912.

Si F es una función de distribución y F_1 es la función de distribución del primer momento entonces la curva $\gamma : t \rightarrow (F(t), F_1(t))$, $t \in [0, \infty)$ es llamada curva de Lorenz. La curva de Lorenz se utiliza para ilustrar la desigualdad en una sociedad con la interpretación de que los más pobres $x = F(t)$ por ciento de una población poseen $y = F_1(t)$ por ciento de la riqueza total. Una medida que se desprende de esta curva es el índice de Gini, que se define como dos veces el área entre la curva γ y la recta $y = x$. Este índice es proporcional a el área entre la curva γ y la recta $y = x$ con el área del triángulo formado por la recta $y = x$ con $x \in [0, 1]$. Cuanto mayor sea el índice de Gini mayor será la desigualdad de los ingresos y cuanto menor sea mayor será la igualdad de ingresos.

Cuando F es matriz exponencial se pueden encontrar fórmulas de la curva de Lorenz y el índice de Gini.

Teorema 4.3.1. *Sea F una función de distribución matriz exponencial con representación (α, \mathbf{S}, s) con $s = -\mathbf{S}e$. Entonces la curva de Lorenz es*

$$\gamma : t \rightarrow \left(1 - \alpha e^{\mathbf{S}t} e, 1 - \frac{\alpha \mathbf{S}^{-1}}{\alpha \mathbf{S}^{-1} e} (e^{\mathbf{S}t} e + t e^{\mathbf{S}t} s) \right)$$

y el índice de Gini G es

$$G = 2(\alpha \otimes \alpha_1)(-(\mathbf{S} \oplus \mathbf{S}_1))^{-1}(s \otimes e) - 1$$

Demostración. La curva de Lorenz se sigue del teorema (4.1.1). El área A bajo la curva γ es

$$\begin{aligned} A &= \int_0^\infty F_1(t) dF(t) \\ &= \int_0^\infty \alpha e^{\mathbf{S}t} s (1 - \alpha_1 e^{\mathbf{S}_1 t} e) dt \\ &= 1 - \int_0^\infty \alpha e^{\mathbf{S}t} s \alpha_1 e^{\mathbf{S}_1 t} e dt \\ &= 1 + (\alpha \otimes \alpha_1)(\mathbf{S} \oplus \mathbf{S}_1)^{-1}(s \otimes e) \end{aligned}$$

\oplus, \otimes son la suma y producto de Kronecker, la última igualdad se debe a que la función de distribución es un número real, al teorema (A.2.3) y la matriz $\mathbf{S} \oplus \mathbf{S}_1$ es invertible por el teorema (A.2.2) pues cada una lo es. El resultado se sigue de que $G = 2(\frac{1}{2} - A)$. \square

Apéndice A

A.1 Algunas propiedades de matrices y exponencial de una matriz

Proposición A.1.1. La serie $\sum_{n=0}^{\infty} \frac{(sA)^n}{n!}$ donde A es una matriz de dimensión $r \times r$ converge uniformemente para todo $s \in [-h, h]$ con $h > 0$.

Demostración. Dada $h > 0$, $s \in [-h, h]$ y $m \in \mathbb{N}$

$$\begin{aligned} \left\| \sum_{n=0}^{\infty} \frac{(sA)^n}{n!} \right\| &\leq \sum_{n=0}^{\infty} \frac{\|sA\|^n}{n!} \\ &\leq \sum_{n=0}^{\infty} \frac{\|h\|^n \|A\|^n}{n!} < \infty \end{aligned}$$

por lo tanto converge uniforme y absolutamente por el criterio de Weierstrass. \square

Por la proposición anterior tiene sentido hablar de la función exponencial de una matriz A definida como e^{sA} , más aún se tiene la continuidad y convergencia uniforme para todo número real s .

Proposición A.1.2. La matriz A conmuta con $\exp(At)$ además si A es invertible entonces A^{-1} conmuta con $\exp(At)$.

Demostración.

$$\begin{aligned} A \exp(At) &= A \sum_{n=0}^{\infty} (At)^n / n! \\ &= \sum_{n=0}^{\infty} A^{n+1} t^n / n! \\ &= \left(\sum_{n=0}^{\infty} (At)^n / n! \right) A = \exp(At) A \end{aligned}$$

$$\begin{aligned}
A^{-1} \exp(At) &= \sum_{n=0}^{\infty} A^{n-1} I t^n / n! \\
&= \sum_{n=0}^{\infty} A^{n-1} A A^{-1} t^n / n! \\
&= \left(\sum_{n=0}^{\infty} (At)^n / n! \right) A^{-1} = \exp(At) A^{-1}
\end{aligned}$$

□

Ahora veamos que tiene sentido encontrar la derivada y por consiguiente la integral de la función exponencial de una matriz.

Proposición A.1.3. *La función e^{sA} es derivable para todo $s \in \mathbb{R}$ y*

$$\frac{d}{ds} e^{sA} = A e^{sA}$$

Demostración. En la serie $\sum_{n=0}^{\infty} \frac{(sA)^n}{n!}$ cada suma truncada hasta n es derivable, la serie converge uniformemente y la serie de derivadas también converge uniformemente por lo que podemos intercambiar el signo de serie y derivada

$$\begin{aligned}
\frac{d}{ds} e^{sA} &= \frac{d}{ds} \sum_{n=0}^{\infty} \frac{(sA)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{d}{ds} \frac{(sA)^n}{n!} \\
&= \sum_{n=1}^{\infty} A \frac{s^{n-1} A^{n-1}}{(n-1)!} \\
&= A \sum_{n=0}^{\infty} \frac{s^n A^n}{n!}
\end{aligned}$$

□

Corolario A.1.1. *Si A es invertible $\int e^{sA} ds = A^{-1} e^{sA}$*

Sea \mathbf{C} una matriz de dimensión n entonces \mathbf{C} es similar a una matriz diagonal por bloques, es decir ,

$$P^{-1} \mathbf{C} P = \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_m \end{pmatrix} = \mathbf{J}$$

donde J_i es una matriz de dimensión $n_i \times n_i$, $\sum_{i=1}^m n_i = \dim(\mathbf{C}) = n$ y

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda_i \end{pmatrix}$$

es el bloque de Jordan correspondiente al valor propio λ_i de \mathbf{C} con multiplicidad n_i . Las columnas de P son vectores propios generalizados. A lo anterior se le llama forma canónica de Jordan.

Proposición A.1.4. *Sea \mathbf{C} una matriz de dimensión n con forma canónica de Jordan*

$$\mathbf{C} = P \begin{pmatrix} J_1 & 0 & \dots & 0 \\ 0 & J_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & J_m \end{pmatrix} P^{-1} = P\mathbf{J}P^{-1}$$

entonces

$$e^{\mathbf{C}x} = P \begin{pmatrix} e^{J_1 x} & 0 & \dots & 0 \\ 0 & e^{J_2 x} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{J_m x} \end{pmatrix} P^{-1}$$

y

$$e^{J_i x} = \begin{pmatrix} e^{\lambda_i x} & \frac{e^{\lambda_i x} x}{1!} & \frac{e^{\lambda_i x} x^2}{2!} & \dots & \frac{e^{\lambda_i x} x^{n-2}}{(n-2)!} & \frac{e^{\lambda_i x} x^{n-1}}{(n-1)!} \\ 0 & e^{\lambda_i x} & \frac{e^{\lambda_i x} x}{1!} & \dots & \frac{e^{\lambda_i x} x^{n-3}}{(n-3)!} & \frac{e^{\lambda_i x} x^{n-2}}{(n-2)!} \\ 0 & 0 & e^{\lambda_i x} & \dots & \frac{e^{\lambda_i x} x^{n-4}}{(n-4)!} & \frac{e^{\lambda_i x} x^{n-3}}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & e^{\lambda_i x} & \frac{e^{\lambda_i x} x}{1!} \\ 0 & 0 & 0 & \dots & 0 & e^{\lambda_i x} \end{pmatrix}$$

Demostración. Usando que $\mathbf{C} = P\mathbf{J}P^{-1}$ se tiene

$$\begin{aligned}
 e^{\mathbf{C}x} &= e^{P\mathbf{J}xP^{-1}} \\
 &= \sum_{k=0}^{\infty} \frac{(P\mathbf{J}xP^{-1})^k}{k!} \\
 &= P \sum_{k=0}^{\infty} \frac{(\mathbf{J}x)^k}{k!} P^{-1} \\
 &= Pe^{\mathbf{J}x}P^{-1} \\
 &= P \begin{pmatrix} e^{J_1x} & 0 & \dots & 0 \\ 0 & e^{J_2x} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{J_mx} \end{pmatrix} P^{-1}
 \end{aligned}$$

pues la exponencial de una matriz diagonal es igual a aplicar la función exponencial a cada elemento de la diagonal y lo mismo se puede hacer a matrices por bloques. Ahora veamos la forma general de la exponencial aplicada a un bloque de Jordan sea J un bloque de Jordan de dimensión n

$$\begin{aligned}
 J &= \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix} \\
 &= \lambda \mathbf{I} + \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix} \\
 &= \lambda \mathbf{I} + N
 \end{aligned}$$

Notemos que

$$N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}$$

y en general N^k es una matriz con unos en la k -ésima diagonal arriba de la principal si

$k < n$ y $N^k = \mathbf{0}$ si $k \geq n$.

$$\begin{aligned}
e^{Jx} &= \sum_{i=0}^{\infty} \frac{Jx^i}{i!} \\
&= \sum_{i=0}^{\infty} \frac{(\lambda x \mathbf{I} + Nx)^i}{i!} \\
&= \sum_{i=0}^{\infty} \sum_{k=0}^i \binom{i}{k} \frac{(\lambda x \mathbf{I})^{i-k} (Nx)^k}{i!} \\
&= \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \frac{(\lambda x)^{i-k} (Nx)^k}{(i-k)! k!} 1_{\{i \geq k\}} \\
&= \sum_{k=0}^{\infty} \frac{(Nx)^k}{k!} \sum_{i=0}^{\infty} \frac{(\lambda x)^{i-k}}{(i-k)!} 1_{\{i \geq k\}} \\
&= \sum_{k=0}^{\infty} \frac{(Nx)^k}{k!} \sum_{r=0}^{\infty} \frac{(\lambda x)^r}{(r)!} \\
&= e^{\lambda x} \sum_{k=0}^{n-1} \frac{(Nx)^k}{k!} \\
&= e^{\lambda x} \begin{pmatrix} 1 & \frac{x}{1!} & \frac{x^2}{2!} & \cdots & \frac{x^{n-2}}{(n-2)!} & \frac{x^{n-1}}{(n-1)!} \\ 0 & 1 & \frac{x}{1!} & \cdots & \frac{x^{n-3}}{(n-3)!} & \frac{x^{n-2}}{(n-2)!} \\ 0 & 0 & 1 & \cdots & \frac{x^{n-4}}{(n-4)!} & \frac{x^{n-3}}{(n-3)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & \frac{x}{1!} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}
\end{aligned}$$

□

Teorema A.1.1. Sea R una matriz bidiagonal por bloques es decir $R_{ii} \neq 0 \neq R_{i(i+1)}$ y cada uno de los anteriores son bloques de igual dimensión entonces

$$R_{ij}^{-1} = \begin{cases} (-1)^{i+j} \prod_{m=i}^j \prod_{p=m}^{\min(j,m+1)} R_{mp}^{(-1)^{i+j+1}} & i \leq j \\ 0 & i > j \end{cases}$$

Demostración. Esta matriz se encuentra con el método de Gauss o se comprueba multiplicando R por R^{-1} utilizando el hecho de que R es bidiagonal. □

A.2 Notación de Kronecker.

Aquí solo se desarrollan algunas propiedades necesarias del producto y suma de Kronecker, para más detalles se puede consultar [7].

Definición A.2.1. Sea $A = \{a_{ij}\}$ una matriz de $m \times n$ y $B = \{b_{ij}\}$ una matriz de $r \times s$ entonces el producto de Kronecker $A \otimes B$ se define como la matriz de $rm \times sn$ dada por

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}$$

Es fácil ver que

$$\begin{aligned} (A + B) \otimes C &= A \otimes C + B \otimes C \\ A \otimes (B + C) &= A \otimes B + A \otimes C \\ (A \otimes B) \otimes C &= A \otimes (B \otimes C) \\ (\lambda A) \otimes B &= A \otimes (\lambda B) = \lambda(A \otimes B) \end{aligned}$$

Proposición A.2.1.

$$(A \otimes B)(C \otimes D) = AC \otimes BD$$

siempre que las dimensiones permitan que los productos estén bien definidas.

Demostración. Los ij -ésimos bloques coinciden pues

$$\begin{aligned} (a_{i1}B \quad a_{i2}B \quad \dots \quad a_{in}B) \begin{pmatrix} c_{1j}B \\ c_{2j}B \\ \dots \\ c_{nj}B \end{pmatrix} &= \left(\sum_k a_{ik}c_{kj} \right) BD \\ &= (AC)_{ij} BD \end{aligned}$$

□

Corolario A.2.1.

$$\prod_{i=1}^n A_i \otimes B_i = \left(\prod_{i=1}^n A_i \right) \otimes \left(\prod_{i=1}^n B_i \right)$$

siempre que las dimensiones permitan que los productos estén bien definidas.

Definición A.2.2. Si A de tamaño $n \times n$ y otra matriz B de tamaño $m \times m$. La suma de Kronecker de dos matrices $A \oplus B$ se define como

$$A \oplus B = A \otimes I_m + I_n \otimes B$$

donde I_s es la matriz identidad de dimensión s .

Proposición A.2.2. Si $\{\lambda_i\}$ y $\{\mu_j\}$ son los eigenvalores de A y B con eigenvectores x_i, y_j respectivamente entonces $\{\lambda_i + \mu_j\}$ son los eigenvalores de $A \oplus B$ con eigenvectores $x_i \otimes y_j$

Demostración. Sean x, y eigenvectores correspondientes a los eigenvalores λ y μ de A y B respectivamente entonces

$$\begin{aligned}(A \oplus B)(x \otimes y) &= (A \otimes I)(x \otimes y) + (I \otimes B)(x \otimes y) \\ &= (Ax \otimes y) + (x \otimes By) \\ &= \lambda(x \otimes y) + \mu(x \otimes y)\end{aligned}$$

□

Proposición A.2.3.

$$e^{A \oplus B} = e^A \otimes e^B$$

Demostración. Por el corolario anterior $(A \otimes I_m)^p = A^p \otimes I_m$ y por el binomio de Newton tenemos que

$$(A \oplus B)^n = \sum_{p+q=n} \frac{n!(A \otimes I_m)^p (I_n \otimes B)^q}{p!q!} = \sum_{p+q=n} \frac{n!A^p \otimes B^q}{p!q!}$$

con q y p enteros no negativos, notemos que para la primer igualdad se necesita la conmutatividad de las matrices $A \otimes I_m$ y $I_n \otimes B$, esto es cierto y se debe a que ambos productos dan como resultado $A \otimes B$ también por esto la segunda igualdad pues $(A \otimes I_m)^p (I_n \otimes B)^q = (A^p \otimes I_m)(I_n \otimes B^q) = A^p \otimes B^q$ entonces

$$\begin{aligned}e^{A \oplus B} &= \sum_{n=0}^{\infty} \frac{(A \oplus B)^n}{n!} \\ &= \sum_{n=0}^{\infty} \sum_{p+q=n} \frac{A^p \otimes B^q}{p!q!} \\ &= \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{A^p \otimes B^q}{p!q!} \\ &= \sum_{p=0}^{\infty} \frac{A^p}{p!} \otimes \sum_{q=0}^{\infty} \frac{B^q}{q!}\end{aligned}$$

□

Bibliografía

- [1] BREUER, L. and BAUM, D. (2005). *An Introduction to Queueing Theory and Matrix-Analytic Methods*. Dordrecht: Springer.
- [2] ASMUSSEN, S. (2003). *Applied Probability and Queues*. (2nd ed.). New York: Springer.
- [3] HIGHAM, N. J. (2008). *Functions of Matrices: Theory and Computation*. Philadelphia: Society for Industrial and Applied Mathematics.
- [4] BLADT, M. and NIELSEN, B. (2011). *Moment Distributions of Phase Type*. *Stochastic Models*, 27(4), 651-663.
- [5] FACKRELL, M. (2003). *Characterization of Matrix-exponential Distributions*. Doctoral Thesis. University of Adelaide.
- [6] BLADT, M. (2015). *Applied Probability: A matrix-exponential approach*. Lecture notes(unpublished).
- [7] GRAHAM, A. (1981). *Kronecker Products and Matrix Calculus with Applications*. London: Ellis Horwood Series.
- [8] BLADT, M., CAMPILLO, A. and NIELSEN, B (2014). *On the use of functional calculus for phase-type and related distributions*. DTU compute technical report.
- [9] NORRIS, J. (1998). *Cambridge Series in Statistical and Probabilistic Mathematics: Markov Chains*. United States of America: Cambridge university press.