



**UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO**

---

---

**FACULTAD DE CIENCIAS**

**Estimación de la Mortalidad Vía  
Distribuciones de Tipo Fase**

**T E S I S**

**QUE PARA OBTENER EL TÍTULO DE:**

**ACTUARIO**

**P R E S E N T A:**

**OSCAR LUCIO CANO VACA**



**DIRECTOR DE TESIS:  
DR. FERNANDO BALTAZAR LARIOS**

**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.

1. Datos del Alumno

Cano

Vaca

Oscar Lucio

56 85 63 48

Universidad Nacional Autónoma de México

Facultad de Ciencias

Actuaría

305323198

2. Datos del Tutor

Dr.

Fernando

Baltazar

Larios

3. Datos del sinodal 1

Dra.

Ana

Meda

Guardiola

4. Datos del sinodal 2

M. en C.

David Josafat

Santana

Cobian

5. Datos del sinodal 3

Act.

Blanca Rebeca

Ayala

Martínez

6. Datos del sinodal 4

Act.

Jonathan

Jaimes

Pérez

7. Datos de la tesis

Estimación de la Mortalidad Vía Distribuciones de Tipo Fase

78 p.

2015



# Dedicatoria y Agradecimientos

A mi familia y amigos

Les dedico esta tesis ya que han estado a mi lado durante todo este tiempo, agradezco su paciencia y tiempo que me han dedicado, ya que gracias a ustedes, su apoyo y consejos he podido llegar a donde estoy y dar este gran paso dentro de mi vida profesional.

A mi director de tesis Dr. Fernando Baltazar Larios

Le agradezco por su tiempo y enseñanzas indispensables para la elaboración de este trabajo

A la Universidad Nacional Autónoma de México

Por abrirme sus puertas para poder llevar a cabo mis estudios profesionales.

# Índice

## Introducción

## Capítulo 1

Antecedentes del Estudio de la Mortalidad	1
---	---

## Capítulo 2

Proceso de Saltos de Markov	12
-----------------------------	----

### Sección 2.1

Definiciones Elementales	13
--------------------------	----

### Sección 2.2

Propiedades	15
-------------	----

Características de los Tiempos de Estancia	15
--	----

Probabilidades de Transición	18
------------------------------	----

Generador Infinitesimal	20
-------------------------	----

Tiempo de Paro	24
----------------	----

### Sección 2.3

Algoritmo para Simular un Proceso de Saltos de Markov	25
---	----

### Sección 2.4

Ejemplo (Proceso de Nacimiento y Muerte)	27
--	----

## Capítulo 3

Distribuciones Tipo Fase	32
--------------------------	----

### Sección 3.1

Definiciones Elementales	32
--------------------------	----

Sección 3.2	
Propiedades. . . . .	33
Caracterización de las Distribuciones Tipo Fase . . . . .	33
Función Generadora de Momentos . . . . .	35
Sección 3.3	
Ejemplo . . . . .	39
Capítulo 4	
Mortalidad a Través de Distribuciones Tipo Fase	43
Sección 4.1	
Modelo de X. Sheldon Lin y Xiaoming Liu . . . . .	43
Sección 4.2	
Modelo Alterno de Mortalidad . . . . .	49
Desarrollo . . . . .	50
Implementación a Datos Reales . . . . .	53
Implementación a Datos Teóricos. . . . .	57
Conclusiones	63
Apéndices	65
Bibliografía	69

# Introducción

A lo largo de la historia de la humanidad, el estudio y medición del riesgo por mortalidad es y ha sido de suma importancia. En la actualidad es básico contar con una forma de medir ésta en la valuación de seguros y rentas contingentes, así como en la gestión de los sistemas de seguridad social y de pensiones; ya que la solvencia y la estabilidad financiera de las instituciones depende, entre otros aspectos, de la disponibilidad de herramientas apropiadas que reflejen una adecuada medición de la siniestralidad a la que se deberán enfrentar, así como sus posibles desviaciones.

Los actuarios han estado utilizando tablas de mortalidad como herramienta básica para describir la mortalidad a través de una estructura por edades. Sin embargo, una tabla es sólo una alternativa cuando no hay una ley matemática disponible. Las tablas de mortalidad constituyen un elemento técnico de primera importancia para una operación sólida, competitiva y eficiente de algunas entidades financieras. En la producción de estas tablas, la evidencia empírica observada se utiliza con el propósito de estimar la probabilidad de muerte. Diversos factores pueden alterar esta probabilidad; el caso más extensamente considerado es el de la edad pero existen otras características relevantes como sexo, historial clínico, tabaquismo, antigüedad de la póliza, etc. En cualquier caso, una tabla de mortalidad habitualmente es un arreglo de probabilidades de muerte dispuestas de acuerdo con la edad de los individuos de la población.

Las revisiones más recientes de las tablas de mortalidad que utiliza el sector asegurador mexicano han ocurrido con intervalos de cuando menos diez años. Es por ello que contar con una herramienta que nos permita medir de una manera constante y sencilla la mortalidad se ha vuelto de suma importancia.

Por esta razón, dentro de este trabajo se propone un modelo para medir la mortalidad con la ayuda de una distribución tipo fase. Este tipo de distribuciones son utilizadas para modelar el tiempo de arribo al estado absorbente dentro de un proceso de saltos de Markov a tiempo continuo con espacio de estados finito. De tal forma que si el proceso de Markov representa la vida de un individuo a través del tiempo y la muerte es el estado absorbente, la distribución tipo fase mide el tiempo de muerte del individuo.

Ya con la distribución tipo fase y tomando en cuenta los aspectos básicos que intervienen en la mortalidad, se pretende estimar los parámetros necesarios de una forma simple y de esta manera obtener una herramienta que nos permita medir el fenómeno de manera sencilla y constante.

Para ello, el desarrollo del trabajo se lleva a cabo en cuatro capítulos, en el [Capítulo 1](#) se presenta el marco histórico, donde hablaremos de los antecedentes del estudio de la mortalidad así como las herramientas que se han utilizado para medir dicho fenómeno a través de la historia. La bibliografía que se utiliza dentro de este capítulo puede consultarse en [\[ 1 \]](#), [\[ 3 \]](#), [\[ 8 \]](#), [\[ 12 \]](#), [\[ 13 \]](#) y [\[ 18 \]](#).

Dentro del [Capítulo 2](#) se describe un proceso de saltos de Markov a tiempo continuo con un número finito de estados, así como las principales propiedades que cumple y que son necesarias para poder dar paso al [Capítulo 3](#) donde hablaremos de las distribuciones tipo fase y sus características, que representan la base para el modelo que se propone, y así con estos dos capítulos tener el marco teórico que dará sustento a nuestro modelo. En estos capítulos la bibliografía utilizada se encuentra en [\[ 2 \]](#), [\[ 4 \]](#), [\[ 5 \]](#), [\[ 6 \]](#), [\[ 7 \]](#), [\[ 9 \]](#), [\[ 10 \]](#), [\[ 11 \]](#), [\[ 14 \]](#), [\[ 15 \]](#), [\[ 16 \]](#) y [\[ 17 \]](#).

En el [Capítulo 4](#) se contextualizará el problema así como la manera en que se propone darle solución, esto describiendo en primera instancia el modelo propuesto por X. Sheldon Lin y Xiaoming Liu, que es un antecedente importante dentro del uso de distribuciones tipo fase en la medición de la mortalidad. Una vez descrito este modelo y teniendo en cuenta sus fortalezas y debilidades se da paso a la construcción del modelo propio, basándonos en la teoría descrita en los capítulos anteriores así como en supuestos especiales que se tienen que hacer para que el modelo pueda ser aplicado, esto utilizando la bibliografía en [\[ 12 \]](#), [\[ 13 \]](#), [\[ 17 \]](#) y [\[ 18 \]](#). Para poner a prueba nuestro modelo, dentro del mismo [Capítulo 4](#) se llevarán a cabo dos implementaciones, con las cuales se pretende ver la eficiencia del modelo propuesto bajo dos situaciones distintas.

Finalmente se dará paso a la sección de [Conclusiones](#), donde se describe lo desarrollado dentro del trabajo y las conclusiones a las que se llega a partir de los resultados obtenidos.

# Capítulo 1

En este capítulo se dará a conocer cómo se ha estudiado el fenómeno de la mortalidad a través del tiempo y las soluciones que se han dado para medir dicho fenómeno.

## Antecedentes del Estudio de la Mortalidad

Los estudios acerca del comportamiento de la mortalidad en el ser humano se remontan principalmente a principios del siglo XVII, aunque existen antecedentes previos a esta fecha.

Existen evidencias de que en el año 220 d.C., **Domicio Ulpiano**, diseñó un arreglo basado en la edad y la supuesta esperanza de vida asociada a esa edad, aunque existen dudas si ésta realmente representaba el concepto de esperanza de vida o simplemente el valor correspondiente a esa edad para definir una renta unitaria vitalicia. Esto es relevante dado que son conceptos distantes a los antecedentes formales del siglo XVII.

En el año de 1693 el astrónomo inglés **Edmund Halley** publicó “*An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslau; with an attempt to ascertain the prices of annuities upon lives*”, en el que ayudándose de un arreglo matricial resumió la experiencia de las defunciones de 1687 a 1691 en la ciudad de Breslau, Alemania, y a partir de ella se pudo construir la primer tabla de rentas vitalicias en función de la probable duración de la vida.

Este arreglo matricial se presentaba de la siguiente manera:

$$\begin{pmatrix} d_{0,1687} & d_{0,1688} & \cdots & d_{0,1691} \\ d_{1,1687} & d_{1,1688} & \cdots & d_{1,1691} \\ \vdots & \vdots & \ddots & \vdots \\ d_{\omega,1687} & d_{\omega,1688} & \cdots & d_{\omega,1691} \end{pmatrix},$$

donde:

$\omega$  := edad última observada

$d_{i,j}$  := número de personas que fallecieron entre las edades  $i$  e  $i+1$  en el año  $j$ ;

y con ello se calculó la media aritmética  $\bar{d}_i$  de las defunciones por edad, es decir:

$$\bar{d}_i = \frac{\sum_{j=1687}^{1691} d_{i,j}}{5}.$$

Ya con el cálculo de estas medias de la mortalidad por edad para todas las edades, se hace la estimación del grupo inicial de expuestos  $l_0$ , o como es conocido en la actualidad “radix” mediante la suma de todas estas medias:

$$l_0 = \sum_{i=0} \bar{d}_i.$$

En el año de 1725 el matemático francés, **Abraham de Moivre**, en su tratado llamado “*Annuities upon lives*” formuló el modelo relativo de número de sobrevivientes en edades sucesivas desde los 12 hasta los 86 años de edad, basado en el trabajo de **Edmund Halley**, estableciendo como función de aproximación una función decreciente no negativa donde el número de personas vivas a edad  $x$  queda definida de la siguiente forma:

$$l_x = 86 - x, \quad \text{para toda } x \in [12, 86].$$

Esta expresión se puede generalizar de la siguiente manera:

$$l_x = a - bx.$$

Dicha expresión tiene las siguientes restricciones:  $b > 0$ ,  $a = l_0$ ,  $\omega = (l_0 / b)$ ,  $x \leq (l_0 / b)$ , donde  $\omega$  es la última edad teórica de sobrevivencia y  $b$  representa la intensidad con la que decrece la población a través del tiempo.

Con esta expresión, se puede describir la función  $d_x$ , que determina el número de fallecimientos entre las edades  $x$  y  $x+1$ , de esta forma podemos escribir a  $d_x$  de la siguiente manera:

$$d_x = l_x - l_{x+1}.$$

Aunque para la descripción de  $l_x$  propuesta hasta el momento, resulta limitada la obtención de las  $d_x$  ya que éstas siempre serán la constante  $b$ .

Este documento resulta interesante ya que es de los primeros esfuerzos por describir la mortalidad de una forma analítica.

Para el año de 1817 el matemático e ingeniero escocés **Edgard Sang**<sup>1</sup> propone una expresión basada en que el número de sobrevivientes en un grupo debe decrecer de manera geométrica conjuntamente con otro factor diferente a la edad, es decir:

$$l_x = a + kb^x,$$

donde:

$$0 < b < 1; a = \frac{l_0 b^\omega}{1 - b^\omega}; k = \frac{l_0}{1 - b^\omega}$$

$l_0$  := número de personas vivas a edad cero (radix)

$$l_\omega = 0$$

$\omega$  := última edad teórica.

En esa misma época el ingeniero francés **Emile Dormoy**<sup>2</sup> propone una expresión parecida a la de Sang, omitiendo el término constante, lo que da como resultado la siguiente expresión:

$$l_x = ab^x,$$

donde:

$$0 < b < 1; a = l_0 \text{ y } b^x \rightarrow 0 \text{ cuando } x \rightarrow \omega.$$

Que es equivalente a una función exponencial no negativa, decreciente y convexa respecto a la variable edad.

Al aplicar la expresión de Dormoy a la definición de  $q_x$ , que representa la probabilidad de morir entre la edad  $x$  y  $x+1$ , y que se define como el número de muertos a edad  $x$  entre el número de expuestos con esta misma edad, se obtiene:

$$q_x = \frac{d_x}{l_x} = \frac{l_x - l_{x+1}}{l_x} = \frac{ab^x - ab^{x+1}}{ab^x} = \frac{ab^x(1 - b)}{ab^x} = 1 - b,$$

lo que nos dice que la probabilidad de muerte es la misma para todas las edades.

Aunque existieron muchos esfuerzos para establecer una ley que describa matemáticamente el comportamiento de la mortalidad del ser humano en el tiempo, fue hasta 1825 que el matemático inglés **Benjamín Gompertz**, estableció de forma más racional y científica una

---

<sup>1</sup> Alex Craik; Edward Sang (1805-1890): calculator extraordinary, special number of Newsletter in memory of J.G. Fauvel; British Society for the History of Mathematics Newsletter.

<sup>2</sup> Emile Dormoy (1829-1891), Théorie des assurances sur la vie (1878), Journal des actuaires (t. IX, p. 64, 1880).

hipótesis en un documento dirigido a la *Royal Society of London*, citando en su célebre memoria que:

*“Es posible que la muerte sea la consecuencia de dos causas coexistentes: una el azar, sin disposición previa a la muerte o al deterioro; otra, una deterioración o una impotencia creciente para resistir a la destrucción.”*

*“La intensidad de la mortalidad podría tenerse como constante. Si no hubiera otras enfermedades, la vida tendría, en todas las edades, el mismo valor, tanto el número de sobrevivientes como el de muertos, decrecería con la edad en progresión geométrica, mientras que las edades crecerían en progresión aritmética.”*

*“Si el agotamiento del poder del hombre para evitar la muerte fuera tal que, en promedio, y al fin de periodos de tiempo infinitamente pequeños, pero de igual duración, perdiera, también, proporciones iguales del poder de oponerse a la muerte que tenía al principio de dicho intervalo, entonces a la edad  $x$ , la intensidad de la mortalidad ( $\mu_x$ ), podría ser representada por  $aq^x$  siendo  $a$  y  $q$  constantes a determinar.”<sup>3</sup>*

Gompertz consideró el recíproco de la intensidad de la muerte para representar la resistencia a ésta, expresándolo de la siguiente manera:

$$\frac{d}{dx} \left( \frac{1}{\mu_x} \right) = -h \frac{1}{\mu_x},$$

donde  $h$  es la constante de proporcionalidad. Integrando tenemos que

$$\ln \frac{1}{\mu_x} = -hx - \ln(B),$$

entonces:

$$\ln \mu_x = hx + \ln(B),$$

donde  $-\ln(B)$  es la constante de integración, entonces

$$e^{\ln \mu_x} = e^{hx + \ln B},$$

finalmente haciendo  $e^h = c$ , tenemos que

$$\mu_x = Bc^x,$$

---

<sup>3</sup> “On the Nature of the Function of de Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies”. Philosophical Transactions of the Royal Society of London 115: 513-585. 1825

donde la hipótesis de Gompertz es equivalente a suponer que la fuerza de mortalidad se incrementa en progresión geométrica.

A pesar de que Gompertz dentro de sus hipótesis plantea que la mortalidad es consecuencia de dos causas coexistentes: el azar sin disposición previa a la muerte o al deterioro, y la impotencia creciente para resistir a la destrucción, su modelo comprende únicamente el segundo de estos componentes.

En el año de 1860, el matemático inglés **William Matthew Makeham** publicó un trabajo llamado “*On the law of mortality and the construction of annuity tables*”, donde incorpora al trabajo de Gompertz esta primer variable, cuyo valor constante representa a todos aquellos factores al azar sin disposición a la muerte, por ejemplo accidentes, epidemias, etc.

Con ello se obtiene la expresión denominada **Ley de Gompertz-Makeham**<sup>4</sup>, que queda expresada de la siguiente manera:

$$\mu_x = A + Bc^x.$$

Debido a los cambios que ha tenido la mortalidad a través del tiempo y la importancia que ha obtenido para la modelación de seguros de vida, sistemas de seguridad social, planes de pensiones, etc., es deseable disponer de un modelo de mortalidad que se ajuste a los datos de mortalidad observados, pueda asociar sus parámetros al deterioro biológico y fisiológico del envejecimiento y permita el análisis cuantitativo de la muerte. La idea de construir un modelo de mortalidad con estas características ha sido considerada antes y para ello en los últimos años se ha tratado de implementar el uso de los procesos de Markov y distribuciones de tipo fase para su modelación.

Por ejemplo, **Gavrilov y Gavrilova** en 1991 (véase [ 8 ] ) utilizan un proceso de Markov para derivar una fórmula **Gompertz-Makeham** bajo supuestos especiales. **Alen** en 1995 (véase [ 2 ] ) explora el potencial teórico de la utilización de distribuciones de tipo fase al modelar de una forma diferente las tasas de riesgo, lo que sugiere que tal modelo debe encontrar una mayor aplicación en el análisis de supervivencia.

En abril de 2008 **X. Sheldon Lin y Xiaoming Liu**<sup>5</sup> en su publicación titulada “*Markov Aging Process and Phase-Type Law of Mortality*”, proponen un modelo que consta de una

---

<sup>4</sup> En la actualidad las expresiones con 2 y 3 parámetros (Gompertz-Makeham), no presentan un grado razonable de ajuste para las edades inferiores a 12 años, ni superiores a 80, sin embargo, han sido aplicados extensamente para graduar tablas de mortalidad dado el nivel de exactitud con el que se ajustan a las tasas empíricas y por la propiedad que poseen de obedecer a la llamada “*Ley del envejecimiento uniforme*”

<sup>5</sup> X. Sheldon Lin, es Profesor de Actuaría en el departamento de Estadística de la Universidad de Toronto.

cadena de Markov con un número finito de estados donde uno de ellos es un estado absorbente el cual representa la muerte.

Aún en la actualidad, el estudio y la estimación de la mortalidad así como el desarrollo de tablas de mortalidad, no es fácil. En México se han optado por varias metodologías para la construcción de tablas de mortalidad, entre ellas tenemos las siguientes:

La *Comisión Nacional de Seguros y Fianzas* realizó en Septiembre de 1991 la tabla de mortalidad “*Experiencia Mexicana 1982-1989*”, con el objeto de proponer un instrumento que permita medir fielmente las probabilidades de muerte relativas a los asegurados.

Para dicha tabla, tomó como base estadística la experiencia de las compañías mexicanas durante el periodo 1982-1989. Debido a que no se contaba con suficiente información para obtener resultados representativos por sexo, los datos se refieren a la experiencia combinada de hombres y mujeres.

Primero se determinaron el número de expuestos al riesgo a edad alcanzada, mediante la expresión:

$$E_x^t = P_{x+1}^t + d_x^t,$$

donde:

$E_x^t$  := Expuestos al riesgo de edad  $x$ , al principio del año  $t$ .

$P_{x+1}^t$  := Pólizas en vigor de edad  $x + 1$ , reportados al 31 de diciembre del año  $t$ ; en el caso del seguro de grupo, certificados en vigor.

$d_x^t$  := Número de asegurados muertos de edad  $x$ , en el transcurso del año  $t$ .

De esta manera, se obtuvieron los expuestos al riesgo por edad y año de exposición. Se reunió la experiencia de todas las instituciones por cada año del periodo 1982 a 1989, obteniendo el total de expuestos por año de observación.

Las experiencias recogidas fueron tratadas siguiendo el principio fundamental de Ackland el cual establece:

*“Toda persona de edad  $x$ , en observación durante unidades de tiempo, en el intervalo de investigación que se considere, equivale, de acuerdo a los efectos de la formación de las clases de edades, a  $t$  personas de edades  $x, x + 1, \dots, x + t - 1$ , puestas todas en observación durante una sola unidad de tiempo”.*

---

Xiaoming Liu es profesor adjunto en el departamento de Actuaría y Estadística de la Universidad de Western Ontario, London.

Asimismo, tomando en consideración los expuestos al riesgo, al principio del periodo anual, y las muertes ocurridas en el transcurso de cada año, el total de expuestos al riesgo se obtuvo aplicando la siguiente fórmula:

$$E_x = \frac{1}{2}(E_x^{82} + E_x^{89}) + \sum_{t=83}^{89} E_x^t,$$

donde:

$E_x$  := Expuestos al riesgo de edad  $x$ .

$E_x^t$  := Expuestos al riesgo de edad  $x$ , al principio del año  $t$ .

El total de muertes, se determinó como la suma de los muertos de edad  $x$ , durante el período de observación, esto es:

$$d_x = \sum_{t=82}^{89} d_x^t,$$

donde:

$d_x$  := Número de asegurados muertos de edad  $x$ .

$d_x^t$  := Número de asegurados muertos de edad  $x$ , en el transcurso del año  $t$ .

De esta forma para el periodo 1982-1989 se obtuvieron las defunciones  $d_x$  para el seguro de vida individual.

Con dichos resultados se pudo realizar el cálculo de la tasa central de mortalidad, la que se calculó a partir de la expresión:

$$m_x = \frac{d_x}{E_x}.$$

Dada la fluctuación que existía en los datos, se dedujo la existencia de errores que afectan al resultado, estos errores pueden ser causados por la insuficiencia en los datos para alguna edad, inexactitud en las declaraciones o influencia de causas adversas, o excepcionalmente favorables. Tal situación obligó a efectuar metodologías de suavizamiento y graduación para obtener tasas con la suavidad y fidelidad suficientes.

En primer término, se agruparon los expuestos y las muertes ocurridas por intervalos quinquenales, lo que produce un suavizamiento inmediato y proporciona un mayor volumen de información por cada intervalo de edad.

Después se optó por ajustar la curva utilizando el método de ecuaciones diferenciales conocido con el nombre de Whittaken-Henderson, Tipo A. Este método se fundamenta en la formulación de una expresión analítica que mide la combinación suavidad-fidelidad a través de las derivadas parciales de la relación entre ambos factores, de tal manera que conforme ésta se minimiza, la divergencia se incrementa hasta que el suavizamiento es favorable a la fidelidad.

De acuerdo con los principios establecidos en este método, la mejor graduación se obtiene cuando la expresión  $F+hs$ , se hace mínima, lo que se obtiene haciendo cero las derivadas parciales de esta expresión, con respecto a una variable independiente.

Con la aplicación de esta metodología se obtuvo como resultado una curva suave y continua de “*Tasas Brutas de Mortalidad*”.

Estas tasas brutas, se graduaron para obtener tasas básicas, representativas de la mortalidad ocurrida en el sector asegurador, durante el período 1982-1989. Para dicha graduación de la tabla se empleó el método de Makeham, fundamentalmente por la aplicación del principio del envejecimiento uniforme, de modo tal que permite una gran simplicidad de cálculo para funciones de vida conjunta, primas, reservas y anualidades.

Recordemos que la hipótesis de Makeham define la tasa instantánea de mortalidad como:

$$\mu_x = A + Bc^x .$$

Para el cálculo de las constantes de Makeham, se determinaron tasas anuales de mortalidad, a partir de la expresión:

$$q_x = \frac{2m_x}{2 + m_x} .$$

Las constantes de Makeham se calcularon utilizando el método de King y Hardy, partiendo de:

$$l_x = Ks^x + g^{c^x} ,$$

aplicando logaritmo obtenemos

$$\log(l_x) = \log(K) + x \log(s) + c^x \log(g) .$$

Por otro lado

$${}_tP_x = \frac{l_{x+t}}{l_x},$$

aplicando logaritmo obtenemos

$$\log({}_tP_x) = \log(l_{x+t}) - \log(l_x),$$

sustituyendo  $\log(l_x)$

$$\log({}_tP_x) = \log(K) + (x+t)\log(s) + c^{x+t}\log(g) - [\log(K) + x\log(s) + c^x\log(g)],$$

entonces

$$\log({}_tP_x) = t\log(s) + c^x(c^t - 1)\log(g).$$

Partiendo de cuatro grupos de edades equidistantes y tomando primeras diferencias tenemos:

$$\log({}_tP_x) = c^x(c^t - 1)^2\log(g),$$

$$c^t = \frac{\log({}_tP_{x+1})}{\log({}_tP_x)},$$

$$\log(g) = \frac{\log({}_tP_x)}{c^x(c^t - 1)^2},$$

$$\log(s) = \frac{1}{t}[\log({}_tP_x) - c^x(c^t - 1)^2\log(g)],$$

$$A = -\log(s),$$

$$B = -\log(c)\log(g).$$

A partir de dichas constantes se calcularon las tasas de probabilidad anuales, que integran la “Tabla de Mortalidad 1989”.

En Mayo del 2000, la *Comisión Nacional de Seguros y Fianzas*, realizó una actualización de esta tabla utilizando la información de la experiencia en mortalidad de todas las empresas de seguros que operaron el ramo de vida durante los años 1991 a 1998, se utilizó la metodología estadística conocida como Análisis Bayesiano de Modelos de Regresión Logística.

Para cada edad  $x$  y dado el número de asegurados expuestos al principio del año  $E_x$ , se tiene que, bajo los supuestos de independencia y homogeneidad,  $d_x$  puede ser modelado con una distribución Binomial con probabilidad de ocurrencia  $q_x$ , es decir:

$$P(d_x|E_x, q_x) \approx \text{Bin}(E_x, q_x) .$$

El reto era pronosticar la cantidad futura  $d'_x$ , o de manera equivalente  $q'_x = d'_x/E_x$ . Para el caso de los datos del seguro de vida individual se ajustó el modelo de regresión logística dado por:

$$Y_x = \ln \frac{q_x}{1 - q_x} = \alpha + \beta x + \varepsilon .$$

Donde  $x$  representa la edad, siendo la única variable independiente del modelo. Las modas de las distribuciones finales de los parámetros en el modelo son las siguientes:

$$\hat{\alpha} = -9.146130 (0.035469)$$

$$\hat{\beta} = 0.074355 (0.000700)$$

$$\hat{\sigma}^2 = 0.107152$$

$$\hat{\tau} = 9.3325$$

donde  $\hat{\tau} = 1/\hat{\sigma}^2$  y el coeficiente de determinación alcanza un valor de 0.9545.

Para el seguro de vida grupo el modelo de regresión logística que se ajustó fue:

$$Y_x = \ln \frac{q_x}{1 - q_x} = \alpha + \beta x^2 + \varepsilon ,$$

con:

$$\hat{\alpha} = -7.518257 (0.029504)$$

$$\hat{\beta} = 0.000793 (0.0000066)$$

$$\hat{\sigma}^2 = 0.146294$$

$$\hat{\tau} = 6.83556$$

Cuyo coeficiente de determinación tiene un valor de 0.9687.

Las tablas de mortalidad fueron formadas con los percentiles del 90% de confianza para vida individual, y el del 75% de confianza para el caso de vida grupo.

La construcción de tablas de mortalidad con base en los modelos hasta ahora descritos, tiene un alto grado de complejidad en su implementación y actualización, por lo que se busca cubrir la necesidad de contar con un modelo que nos permita actualizarlo constantemente, medir el riesgo por mortalidad ante distintos escenarios y que además, su implementación sea sencilla. Es por ello que dentro de este trabajo exploraremos la opción de medir el riesgo de mortalidad mediante distribuciones tipo fase, para ello partiremos de la teoría con la cual se construye el concepto de distribución tipo fase, que es la de procesos de saltos de Markov a tiempo continuo.

# Capítulo 2

Dentro de este capítulo se define un proceso de saltos de Markov a tiempo continuo y se estudian sus principales propiedades. El objetivo de estudiar este tipo de procesos es que nos sirve como preliminar para definir lo que es una distribución tipo fase y como base para la construcción del modelo de mortalidad.

## Proceso de Saltos de Markov

Consideremos un sistema donde se puede estar en cualquiera de los estados de un conjunto previamente especificado y que evoluciona o cambia de un estado a otro a través del tiempo de acuerdo con una cierta ley de movimiento. Sea  $X_t$  el estado del sistema al tiempo  $t$ . Si consideramos que la forma en la que el sistema evoluciona no es determinista, sino aleatoria, entonces puede considerarse que  $X_t$  es una variable aleatoria para cada valor de  $t$ . La colección de estas variables aleatorias es la definición de un proceso estocástico, donde todas las variables aleatorias están definidas dentro de un espacio de probabilidad  $\{\Omega, \mathcal{F}, P\}$ . Dicha definición puede ser enunciada de la siguiente forma.

**Definición 2.1:** Un proceso estocástico es una colección de variables aleatorias  $\{X_t \mid t \in T\}$ , definidas en un espacio de probabilidad común y parametrizadas por un conjunto  $T$  llamado espacio parametral, en donde las variables toman valores en un conjunto  $E$  llamado espacio de estados.

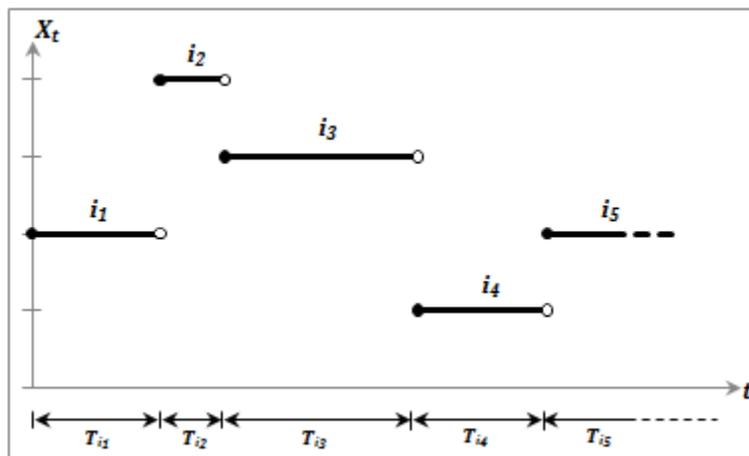
En general, las variables aleatorias que conforman un proceso no son independientes entre si, sino que están relacionadas unas con otras de alguna manera particular. Las distintas formas en que pueden darse estas dependencias es una de las características que distingue a un proceso de los otros.

Una forma de clasificar los diferentes tipos de procesos estocásticos se obtienen de considerar las distintas posibilidades para el espacio parametral, el espacio de estados, las características de las trayectorias, y principalmente las relaciones de dependencia entre las variables aleatorias que conforman el proceso. A continuación nos enfocaremos en lo que es un proceso de saltos de Markov a tiempo continuo, describiendo algunas de las propiedades que este cumple.

## Sección 2.1

### Definiciones Elementales

Consideremos un proceso estocástico a tiempo continuo  $\{X_t\}_{t \geq 0}$  sobre un espacio de estados finito  $E = \{1, 2, \dots, m\}$ . Dicho proceso inicia en un estado  $i_1 \in E$  al tiempo  $T_0$  (por convencionalidad  $T_0 = 0$ ) y permanece allí un tiempo aleatorio  $T_{i_1}$ , después de este tiempo salta a un estado  $i_2$  distinto de  $i_1$ , permaneciendo en este nuevo estado un tiempo  $T_{i_2}$ , al cabo del cual pasa a un estado  $i_3$  distinto del inmediato anterior, y así sucesivamente. Una trayectoria de dicho comportamiento se muestra en la siguiente gráfica.



Gráfica 2.1.1 Trayectoria de un Proceso de Saltos.

El estado  $i_n$ , es el estado en el que se encuentra el proceso antes de realizar el n-ésimo salto. Los tiempos aleatorios  $T_{i_n}$ , son llamados tiempos de estancia y se refieren al tiempo en que el proceso permanece constante en un estado  $i_n$ . Los instantes en los que el proceso cambia de estado son los tiempos  $W_n$ , definidos de la siguiente forma:

$$W_n = T_{i_1} + T_{i_2} + \dots + T_{i_n}, \quad \text{con } n \geq 1, i_n \in E;$$

donde  $W_n$  es el tiempo de ocurrencia del n-ésimo salto.

Un proceso estocástico con el comportamiento descrito, es llamado un proceso de saltos. Si dicho proceso además satisface la siguiente propiedad es llamado proceso de saltos de Markov.

**Propiedad de Markov:** Consideremos un proceso de saltos  $\{X_t\}_{t \geq 0}$  sobre un espacio de estados finito  $E = \{1, 2, \dots, m\}$  y cualesquiera tiempos  $0 \leq t_1 < t_2 < \dots < t_n$ ,  $n \in \mathbb{N}$ , se dice que dicho proceso satisface o tiene la propiedad de Markov si:

$$\mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1) = \mathbb{P}(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1}), \forall i \in E.$$

Un concepto importante en un proceso de saltos es la probabilidad de salto que enunciaremos a continuación.

**Probabilidades de salto:** Denotemos por  $p_{ij}$  a la probabilidad de que el proceso salte del estado  $i$  al estado  $j$  (al efectuar un salto), dichas probabilidades cumplen las siguientes condiciones:

- a)  $p_{ij} \geq 0$  para todo  $i, j \in E$
- b)  $p_{ii} = 0$
- c)  $\sum_{j \in E} p_{ij} = 1$ .

En forma matricial, las probabilidades de saltos constituyen una matriz estocástica de la siguiente forma:

$$P = \begin{pmatrix} 0 & p_{1,2} & \cdots & p_{1,m} \\ p_{2,1} & 0 & \cdots & p_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & 0 \end{pmatrix}.$$

Los procesos de saltos de Markov que nos interesan serán los que cumplan con los siguientes supuestos.

**Reducción:** Puede suceder que los tiempos de estancia  $T_i$  sean cada vez más pequeños, de tal forma que  $\lim_{n \rightarrow \infty} W_n < \infty$ , esto querría decir que el proceso tiene la posibilidad de realizar una cantidad infinita de saltos en un tiempo finito, de tal forma que el proceso no está bien definido para toda  $t \geq 0$  y se dice que el proceso explota en un tiempo finito. Para evitar tal evento se supone que  $\lim_{n \rightarrow \infty} W_n = \infty$ , por lo tanto, para cada  $t \geq 0$  el valor  $X_t$  está bien definido.

**Estados absorbentes o no absorbentes:** Supondremos que existen sólo 2 tipos de estados, absorbentes o no absorbentes, en otras palabras con probabilidad uno el tiempo de estancia en un estado es finito o con probabilidad uno es infinito. Si  $T_i = \infty$  se dice que el proceso deja de saltar y permanece en el estado  $i$  el resto del tiempo, es decir el estado  $i$  es absorbente.

Ya que sabemos lo que es un proceso de Saltos de Markov, procederemos a enunciar algunas de las propiedades principales que este cumple y que nos servirán como sustento en el modelo de mortalidad que se propone.

## Sección 2.2

### Propiedades

Una de las características importantes de un proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$  es que tiene asociada una cadena de Markov a tiempo discreto  $\{Y_n\}_{n \in \mathbb{N}}$ . Dicha cadena está dada por los tiempos donde se efectúan los saltos y donde  $Y_n = i$  se interpreta como que en el  $n$ -ésimo salto, la cadena se encuentra en el estado  $i$ .

Al igual que en las cadenas de Markov a tiempo discreto, el proceso de saltos de Markov tiene asociada una distribución inicial que denotaremos por:

$$\theta = (\theta_1, \dots, \theta_m),$$

que es la distribución de  $X_0$ .

### Características de los Tiempos de Estancia

Una de las grandes consecuencias que tiene la propiedad de Markov en nuestro proceso es que hace que los tiempos de estancia se distribuyan de una forma exponencial con un cierto parámetro  $\lambda_i$ , es decir:

$$T_i \sim \text{Exp}(\lambda_i).$$

Para poder enunciar formalmente dicha propiedad introduciremos la siguiente definición.

**Definición 2.2.1:** Suponemos que  $X_t = i$ , y que el proceso saldrá del estado  $i$  después de un tiempo  $s$ , con  $s > 0$ , es decir  $X_{t+s} \neq i$ . Definamos la variable aleatoria  $\Delta(t)$  de la siguiente manera:

$$\Delta(t) = \inf\{s > 0 | X_{t+s} \neq i\}.$$

Con ayuda de esta variable aleatoria podemos enunciar la siguiente proposición donde queda demostrado que los tiempos de estancia tienen una distribución exponencial.

**Proposición 2.2.1:** Sea  $\{X_t\}_{t \geq 0}$  un proceso de saltos de Markov con espacio de estados  $E = \{1, 2, \dots, m\}$ . Para cada estado  $i \in E$ , existe un parámetro de intensidad denotado por  $\lambda_i$ , independiente del tiempo  $t$ , tal que:

$$\mathbb{P}(\Delta(t) > s | X_t = i) = e^{-\lambda_i s},$$

para toda  $s > 0$ .

**Demostración:** Partiendo de la [Definición 2.2.1](#) tenemos que por la homogeneidad del proceso de saltos de Markov:

$$\mathbb{P}(\Delta(t) > s | X_t = i) = \mathbb{P}(\Delta(0) > s | X_0 = i),$$

y si definimos:

$$h(s) = \mathbb{P}(\Delta(0) > s | X_0 = i),$$

entonces tenemos que:

$$\begin{aligned} h(t + s) &= \mathbb{P}(\Delta(0) > t + s | X_0 = i) \\ &= \mathbb{P}(\Delta(0) > t, \Delta(t) > s | X_0 = i) \\ &= \mathbb{P}(\Delta(0) > t | X_0 = i) \mathbb{P}(\Delta(t) > s | \Delta(0) > t, X_0 = i) \\ &= h(t) \mathbb{P}(\Delta(t) > s | X_t = i) \\ &= h(t) \mathbb{P}(\Delta(0) > s | X_0 = i) \\ &= h(t) h(s). \end{aligned} \tag{2.1}$$

Y dado que las probabilidades de transición son continuas en cero,  $h$  es continua en cero. Y como,  $0 \leq h \leq 1$ , entonces la única solución a la ecuación (2.1) es:

$$h(s) = e^{-\lambda_i s}, \quad \text{con } \lambda_i \in [0, \infty).$$

Por lo tanto

$$\mathbb{P}(\Delta(t) > s | X_t = i) = e^{-\lambda_i s}. \quad \blacksquare$$

**Definición 2.2.2:** Si  $i \in E$ , con una intensidad  $\lambda_i$  asociada. Entonces el estado  $i$  es llamado:

1. Absorbente, si  $\lambda_i = 0$ .
2. No absorbente, si  $0 < \lambda_i < \infty$ .

Obsérvese que un proceso de Markov a tiempo continuo queda completamente especificado por los siguientes tres elementos: una distribución inicial  $\theta$  en el espacio de estados, el conjunto de los parámetros  $\lambda_i$  y las probabilidades de saltos  $p_{ij}$ .

Una propiedad importante en un proceso de saltos de Markov es la independencia que existe entre el tiempo de estancia en el estado  $i$  y el estado al que se salta, dicha propiedad queda demostrada con la siguiente proposición.

**Proposición 2.2.2:** (Independencia condicional). Consideremos un proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$  con espacio de estados  $E = \{1, 2, \dots, m\}$  y suponemos  $W_{n+1} < \infty$ , entonces condicionando con  $X_{W_n} = i$ , las variables aleatorias  $T_i$  y  $X_{W_{n+1}}$  son independientes, es decir:

$$\mathbb{P}(T_i > t, X_{W_{n+1}} = j | X_{W_n} = i) = \mathbb{P}(T_i > t | X_{W_n} = i) \mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i),$$

con  $n \in \mathbb{N}$ ,  $i \neq j$ ,  $i, j \in E$ .

**Demostración:** Utilizando la propiedad de Bayes tenemos que

$$\begin{aligned} \mathbb{P}(T_i > t, X_{W_{n+1}} = j | X_{W_n} = i) &= \frac{\mathbb{P}(T_i > t, X_{W_{n+1}} = j, X_{W_n} = i)}{\mathbb{P}(X_{W_n} = i)} \\ &= \mathbb{P}(T_i > t | X_{W_n} = i) \mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i, T_i > t), \end{aligned}$$

entonces por demostrar que:

$$\mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i, T_i > t) = \mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i).$$

Utilizando la propiedad de Markov tenemos que:

$$\begin{aligned} \mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i, T_i > t) &= \mathbb{P}(X_{W_{n+1}} = j, X_s = i, W_n \leq s < W_{n+1} | X_{W_n+t} = i) \\ &= \mathbb{P}(X_{W_{n+1} + \Delta(W_n+t)} = j, X_s = i, W_n \leq s < W_{n+1} + \Delta(W_n+t) | X_{W_n+t} = i) \\ &= \mathbb{P}(X_{W_{n+1} + \Delta(W_n+t)} = j | X_{W_n+t} = i) \\ &= \mathbb{P}(X_{W_{n+1}} = j | X_{W_n} = i). \quad \blacksquare \end{aligned}$$

## Probabilidades de Transición

Otro concepto importante dentro del proceso de saltos de Markov es el de probabilidad de transición, que se refiere a la probabilidad de estar en un estado  $j$  después de un tiempo  $t$ , dado que se inicia en un estado  $i$ , y su notación es la siguiente:

$$p_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i),$$

en particular si  $t = 0$ , se define a  $p_{ij}(0)$  como la función:

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

conocida como delta de Kronecker.

Para un tiempo  $t \geq 0$ , podemos escribir las probabilidades de transición de la siguiente forma matricial:

$$P_t = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \cdots & p_{1m}(t) \\ p_{21}(t) & p_{22}(t) & \cdots & p_{2m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}(t) & p_{m2}(t) & \cdots & p_{mm}(t) \end{pmatrix}.$$

El conjunto de dichas matrices de transición  $\{P_t \mid t \geq 0\}$  constituye un semigrupo de matrices estocásticas, es decir, cumple las siguientes propiedades:

- $P_0 = I_m$ , donde  $I_m$  es la matriz identidad de dimensión  $m$ .
- $P_t$  es una matriz estocástica.
- $P_{t+s} = P_t P_s$ , para cualesquiera  $t, s \geq 0$ .

La tercera propiedad es la versión continua de la ecuación de Chapman-Kolmogorov (véase [ 16 ] ), conocida como la propiedad de semigrupo, y es demostrada con la siguiente proposición.

**Proposición 2.2.3:** (Ecuación de Chapman-Kolmogorov). Para cualesquiera estados  $i, j \in E$ , y para todo  $t \geq 0$  y  $s \geq 0$ ,

$$p_{ij}(t + s) = \sum_{k=1}^m p_{ik}(t) p_{kj}(s).$$

En notación matricial,  $P_{t+s} = P_t P_s$ .

**Demostración:** Partiendo del lado izquierdo de la igualdad:

$$\begin{aligned} p_{ij}(t+s) &= \mathbb{P}(X_{t+s} = j | X_0 = i) \\ &= \sum_{k=1}^m \mathbb{P}(X_{t+s} = j, X_t = k | X_0 = i) \\ &= \sum_{k=1}^m \mathbb{P}(X_{t+s} = j | X_t = k) \mathbb{P}(X_t = k | X_0 = i), \end{aligned}$$

usando la homogeneidad de proceso de saltos de Markov,

$$\begin{aligned} &= \sum_{k=1}^m \mathbb{P}(X_t = k | X_0 = i) \mathbb{P}(X_s = j | X_0 = k) \\ &= \sum_{k=1}^m p_{ik}(t) p_{kj}(s). \quad \blacksquare \end{aligned}$$

Debemos tomar en cuenta que dentro del proceso las probabilidades de salto  $p_{ij}$  y las probabilidades de transición  $p_{ij}(t)$  se refieren a aspectos distintos.  $p_{ij}$  se refiere a la probabilidad de saltar del estado  $i$  al estado  $j$  dado que existe un salto, mientras que  $p_{ij}(t)$  es la probabilidad de que dado que se está en el estado  $i$ , después de un tiempo  $t$  estar en el estado  $j$ .

Una vez que contamos con el concepto de probabilidad de transición, podemos hablar de otro concepto importante, el de distribuciones estacionarias.

**Definición 2.2.3:** Un vector  $\pi = (\pi_1, \dots, \pi_m)$  es una distribución estacionaria del proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , con el semigrupo de matrices estocásticas  $\{P_t | t \geq 0\}$  asociado si:

1.  $\pi_i \geq 0$  con  $i \in E$
2.  $\sum_{i=1}^m \pi_i = 1$
3.  $\pi = \pi P_t$  para toda  $t \geq 0$ .

Por lo tanto, dada la distribución inicial de  $X_0$  como distribución estacionaria  $\pi$ , entonces para toda  $t > 0$ ,

$$\mathbb{P}(X_t = j) = \sum_{i=1}^m \pi_i p_{ij}(t) = \pi_j,$$

es decir, la variable aleatoria  $X_t$  tiene la misma distribución  $\pi$  para cualquier valor de  $t$ .

En la siguiente proposición obtenemos una forma explícita para las probabilidades de transición de un proceso de saltos de Markov.

**Proposición 2.2.4:** Sean  $i, j \in E$ . Para toda  $t \geq 0$ ,

$$p_{ij}(t) = \delta_{ij}e^{-\lambda_i t} + \int_0^t \sum_{k \neq i} p_{kj}(t-u) p_{ik} \lambda_i e^{-\lambda_i u} du .$$

**Demostración:** Si  $i$  es un estado absorbente, es decir  $\lambda_i = 0$ , entonces la fórmula se reduce a  $p_{ij}(t) = \delta_{ij}$ , lo cual es correcto. Si  $i$  es un estado no absorbente, entonces:

$$\begin{aligned} p_{ij}(t) &= \mathbb{P}(X_t = j | X_0 = i) \\ &= \mathbb{P}(X_t = j, T_i > t | X_0 = i) + \mathbb{P}(X_t = j, T_i \leq t | X_0 = i) \\ &= \delta_{ij}e^{-\lambda_i t} + \int_0^t f_{X_t, T_i | X_0}(j, u | i) du \\ &= \delta_{ij}e^{-\lambda_i t} + \int_0^t \sum_{k \neq i} f_{X_t, X_u, T_i | X_0}(j, k, u | i) du , \end{aligned}$$

en donde por la propiedad de Markov e independencia,

$$\begin{aligned} f_{X_t, X_u, T_i | X_0}(j, k, u | i) &= f_{X_t | X_u, T_i, X_0}(j | k, u, i) f_{X_u | T_i, X_0}(k | u, i) f_{T_i | X_0}(u | i) \\ &= p_{kj}(t-u) p_{ik} \lambda_i e^{-\lambda_i u} . \end{aligned}$$

Por lo tanto,

$$p_{ij}(t) = \delta_{ij}e^{-\lambda_i t} + \int_0^t \sum_{k \neq i} p_{kj}(t-u) p_{ik} \lambda_i e^{-\lambda_i u} du . \quad \blacksquare$$

## Generador Infinitesimal

Una vez que contamos con todas las características anteriores que cumple un proceso de saltos de Markov podemos preguntarnos: ¿qué pasa con las probabilidades de transición en intervalos de tiempo muy pequeños?, en particular en los momentos de saltos.

Supongamos que nuestro proceso se encuentra en el estado  $i$  al tiempo  $t$ , es decir,  $X_t = i$ . Nos interesa saber que pasa en el intervalo de tiempo  $[t, t+h)$  para  $h$  infinitamente pequeña, entonces tenemos los siguientes casos:

1. El proceso se encuentra en el estado  $i$  al tiempo  $t+h$ , este caso ocurre con la probabilidad  $p_{ii}(h) + o(h)$ , donde  $o(h)$  representa la posibilidad de que el proceso salga del estado  $i$ , y regrese a él antes o al tiempo  $t+h$ .

2. El proceso salta a un estado  $j$ , este caso tiene probabilidad  $p_{ij}(h) + o(h)$ , donde  $o(h)$  representa la posibilidad de 2 o más transiciones en el intervalo  $[t, t + h)$ .

En el caso donde  $h$  es infinitamente pequeña se tiene la siguiente proposición.

**Proposición 2.2.5:** Considere el semigrupo estocástico  $\{P_t \mid t \geq 0\}$  asociado al proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , entonces el siguiente límite existe:

$$\lim_{h \rightarrow 0^+} \frac{P_h - I}{h},$$

dicho límite es equivalente a:

$$\lim_{h \rightarrow 0^+} \frac{1 - p_{ii}(h)}{h} = g_i,$$

$$\lim_{h \rightarrow 0^+} \frac{p_{ij}(h)}{h} = g_{ij},$$

con  $0 \leq g_{ij} < \infty$  y  $0 \leq g_i \leq \infty$ .

**Demostración:** Véase [ 9 ] y [ 10 ].

Que dicho límite exista se interpreta como que la matriz de transición converge más rápido a la identidad de lo que el tiempo  $h$  converge a cero, lo cual quiere decir que en intervalos de tiempo muy pequeños el proceso se mantiene en el mismo estado o a lo más salta una vez, y aunque existe la probabilidad de que salte más veces, dicha probabilidad es despreciable.

Recordemos que la matriz  $P_h$  es una matriz estocástica, entonces para todo  $i, j \in E$ :

$$1 = p_{ii}(h) + \sum_{i \neq j} p_{ij}(h),$$

es decir:

$$1 - p_{ii}(h) = \sum_{i \neq j} p_{ij}(h),$$

dividiendo entre  $h$  y tomando el límite cuando  $h$  tiende a cero, tenemos:

$$g_i = \sum_{i \neq j} g_{ij}.$$

**Definición 2.2.4:** Consideremos la matriz  $G$  de  $m \times m$ , que queda definida de la siguiente forma:

$$G = \begin{pmatrix} -g_1 & g_{12} & \cdots & g_{1m} \\ g_{21} & -g_2 & \cdots & g_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \cdots & -g_m \end{pmatrix}$$

La matriz  $G$  es llamada el generador infinitesimal del proceso de saltos de Markov  $\{X_{ij}\}_{t \geq 0}$  y los valores  $g_{ij}$  son llamados tasas de intensidad.

Ahora partamos de la siguiente expresión, que se cumple dado que  $\{P_t | t \geq 0\}$  satisface la ecuación Chapman-Kolmogorov:

$$\frac{P_{t+h} - P_t}{h} = \frac{P_t P_h - P_t}{h} = \frac{P_t [P_h - I]}{h} = P_t \frac{P_h - I}{h}.$$

Tomando el límite cuando  $h$  tiende a cero de ambos lados de la igualdad obtenemos que la derivada de  $P_t$  tiene que ser:

$$P'_t = P_t G = G P_t,$$

donde  $P'_t$  representa la matriz cuyas entradas son los elementos  $p'_{ij}(t)$  para todo  $i, j \in E$ .

Dicha expresión conforma una ecuación diferencial ordinaria, que puede ser resuelta con la condición inicial  $P_0=I$ , y la solución asociada es la siguiente:

$$P_t = \exp(tG) = \sum_{n=0}^{\infty} \frac{t^n G^n}{n!},$$

donde,  $\exp(tG)$  denota la matriz exponencial de  $tG$ .

La matriz  $G$  determina de manera única el comportamiento del proceso de saltos de Markov y cumple con las siguientes propiedades:

1.  $-g_{ii} \leq 0$ , donde  $g_{ii} = g_i$ .
2.  $g_{ij} \geq 0$ , si  $i \neq j$ .
3.  $\sum_j g_{ij} = 0$ .

Una propiedad importante que cumple el generador infinitesimal en relación con la distribución  $\pi$  es la siguiente.

**Proposición 2.2.6:** La distribución  $\pi$  es estacionaria, es decir,  $\pi = \pi P(t)$  si y sólo si  $\pi G = \bar{0}$ , donde  $\bar{0}$  es el vector cuyas entradas son cero.

**Demostración:** Partimos de la siguiente igualdad:

$$\pi G = \bar{0},$$

esto pasa si y sólo si:

$$\pi G^n = \bar{0} \text{ para toda } n \in \mathbb{N},$$

sumando desde  $n$  igual a uno hasta infinito:

$$\sum_{n=1}^{\infty} \frac{t^n}{n!} \pi G^n = \bar{0},$$

sumando  $\pi$  de ambos lados:

$$\pi + \sum_{n=1}^{\infty} \frac{t^n}{n!} \pi G^n = \pi,$$

esto es equivalente a:

$$\pi \sum_{n=0}^{\infty} \frac{t^n}{n!} \pi G^n = \pi,$$

por lo tanto:

$$\pi P(t) = \pi \quad \blacksquare$$

La siguiente definición nos da la relación que existe entre las tasas de intensidad y las probabilidades de salto.

**Definición 2.2.5:** Sea  $G$  el generador infinitesimal asociado al proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , entonces las probabilidades de saltos  $p_{ij}$  quedan definidas de la siguiente manera:

$$p_{ij} = \frac{g_{ij}}{g_i}.$$

Recordemos que en la [Proposición 2.2.1](#) demostramos que los tiempos de estancia en el estado  $i$  se distribuyen de forma exponencial con cierto parámetro  $\lambda_i > 0$ , con el siguiente teorema quedará demostrado que ese parámetro  $\lambda_i$  es igual al valor  $g_i$  del generador infinitesimal.

**Teorema 2.2.1:** Sea  $G$  el generador infinitesimal del proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , entonces condicionando a que el proceso se encuentre en el estado  $i$  al tiempo  $t$ , la variable aleatoria  $\Delta(t)$  se distribuye de forma exponencial con parámetro  $g_i$ , es decir:

$$\mathbb{P}(\Delta(t) > s | X_t = i) = \exp(-g_i s).$$

**Demostración:** Por la [Proposición 2.2.1](#) tenemos que la distribución  $F_{\Delta}(t)$  de  $\Delta(t)$  es exponencial con parámetro  $\lambda_i$  entonces:

$$1 - F_{\Delta}(t) = e^{-\lambda_i t},$$

donde:

$$\lambda_i = F'_{\Delta}(t) = g_i \quad \blacksquare$$

## Tiempo de Paro

Un concepto que definiremos a continuación y que es importante para poder hablar de distribuciones tipo fase es el de tiempo de paro, pero antes tendremos que definir lo que es una filtración.

**Definición 2.2.6: (Filtración).** Una filtración es una colección no numerable de  $\sigma$ -álgebras  $\{\mathcal{F}_t\}_{t \geq 0}$  tal que  $\mathcal{F}_s \subseteq \mathcal{F}_t$ , cuando  $0 \leq s \leq t$ . La filtración natural o canónica de un proceso a tiempo continuo  $\{X_t\}_{t \geq 0}$  es la colección de  $\sigma$ -álgebras  $\{\mathcal{F}_t\}_{t \geq 0}$  dadas por  $\mathcal{F}_t = \sigma(X_s | 0 \leq s \leq t)$ , esto es,  $\mathcal{F}_t$  es la mínima  $\sigma$ -álgebra que hace que cada una de las variables  $X_s$ , para valores de  $s$  en el intervalo  $[0, t]$ , sean medibles. A la  $\sigma$ -álgebra  $\mathcal{F}_t$  se le denomina la historia del proceso al tiempo  $t$ .

Una vez que contamos con el concepto de filtración podemos definir lo que es un tiempo de paro.

**Definición 2.2.7: (Tiempo de paro).** Una variable aleatoria no negativa  $Z$  es un tiempo de paro para el proceso  $\{X_t\}_{t \geq 0}$  si  $\{Z \leq t\} \in \mathcal{F}_t$  para todo  $t \geq 0$ , donde  $\mathcal{F}_t$  es la filtración del proceso  $X$  al tiempo  $t$ .

Un ejemplo de tiempo de paro es la variable aleatoria  $\tau$  que representa el tiempo que el proceso  $\{X_t\}_{t \geq 0}$  tarda en llegar al estado  $i$ , que representa el tiempo del primer arribo al estado  $i$ , con  $i \in E$ , definido de la siguiente manera:

$$\tau = \inf\{t \geq 0 | X_t = i\}. \quad (2.2)$$

Con la siguiente proposición demostraremos que esta variable aleatoria es tiempo de paro.

**Proposición 2.2.7:** Consideremos la variable aleatoria  $\tau$  como en la expresión (2.2) y sea  $\mathcal{F}_t$  la filtración natural asociada al proceso  $\{X_t\}_{t \geq 0}$ , entonces  $\tau$  es tiempo de paro si:

$$\{\tau \leq t\} \in \mathcal{F}_t, \quad \text{con } t \geq 0.$$

**Demostración:** Sea  $\mathcal{F}_t$  la filtración natural asociada al proceso  $\{X_t\}_{t \geq 0}$ , la variable aleatoria  $\tau$  y  $0 \leq t' \leq t$ , entonces:

$$\{\tau = t'\} = \{X_s \neq i | 0 \leq s < t'\} \cap \{X_{t'} = i\} \in \mathcal{F}_{t'} \subset \mathcal{F}_t \quad \blacksquare$$

De manera general se puede demostrar que el  $n$ -ésimo tiempo de arribo al estado  $i$  es un tiempo de paro.

A continuación daremos un algoritmo para poder simular trayectorias de un proceso de saltos de Markov.

## Sección 2.3

### Algoritmo para Simular un Proceso de Saltos de Markov

Dentro de esta sección tenemos un algoritmo para poder simular trayectorias de un proceso de saltos de Markov, para dicho algoritmo partimos de los siguientes supuestos.

Consideremos un proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , con un espacio de estados  $E = \{1, 2, \dots, m\}$ , generador infinitesimal  $G$  y distribución inicial  $\pi$ . Recordemos que todo proceso de saltos de Markov tiene asociada una cadena de Markov a tiempo discreto  $\{Y_n\}_{n \in \mathbb{N}}$ , ya que esta cadena de Markov nos ayudará en la simulación. Partamos de que conocemos el generador infinitesimal  $G$ , a partir del cual construimos la matriz de probabilidades de saltos  $P$  con ayuda de la [Definición 2.2.5](#). Una vez que contamos con estos elementos procedemos a la descripción del algoritmo de simulación de nuestro proceso.

### Algoritmo 2.3.1:

Primero se definen las variables a utilizar:

$X_t$  := Estado en el que se encuentra el proceso al tiempo  $t$ .

$Y_n$  := Estado al que se salta en el  $n$ -ésimo salto.

$W_n$  := Tiempo de ocurrencia del  $n$ -ésimo salto.

$i$  := Auxiliar, que denota un estado,  $i \in E$ .

$T_i$  := Tiempo de estancia en el estado  $i$ .

$n$  := Contador para el número de saltos,  $n \in \mathbb{N}$ .

Empieza el algoritmo.

1. Se inicializan las variables:

$$n = 0$$

$$W_n = 0$$

$Y_n$  = estado tomado aleatoriamente de acuerdo a la distribución inicial  $\pi$ .

2.  $i = Y_n$ .

3. Generamos  $T_i$  igual a un número aleatorio proveniente de una distribución exponencial con parámetro  $g_{ii}$ .

4. Hacemos  $W_{n+1} = W_n + T_i$ .

5.  $X_t = Y_n$  para  $W_n \leq t < W_{n+1}$ .

6.  $Y_{n+1}$  igual a un estado aleatorio proveniente de la distribución del  $i$ -ésimo renglón de la matriz  $P$ .

7.  $n = n+1$ .

8. Regresamos al paso 2.

Dado que este algoritmo es cíclico, es necesaria una condición de paro, que dependiendo del objeto de estudio podría ser, por ejemplo, un número finito de saltos, un tiempo  $t$  finito o la llegada a un estado en especial, en el primer caso se realizará el algoritmo un número de veces especificado, mientras que en el segundo caso se realizará el algoritmo mientras el tiempo de ocurrencia del  $n$ -ésimo salto no sea mayor al tiempo  $t$ , es decir  $W_n \leq t$  y para el tercer caso se realizará el algoritmo mientras el proceso no llegue al estado de interés. Dichas condición puede ser validada al principio o al final del algoritmo.

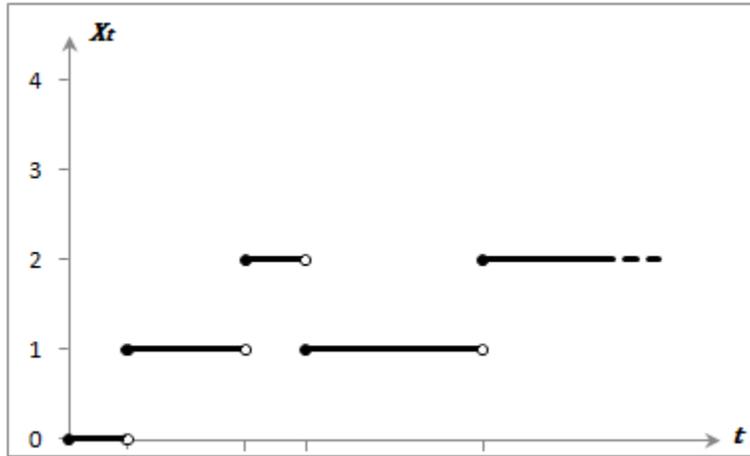
Con esto tenemos los elementos suficientes para poder definir lo que es una distribución tipo fase, pero antes daremos un ejemplo de un proceso de saltos de Markov a tiempo continuo, llamado Proceso de Nacimiento y Muerte.

## Sección 2.4

### Ejemplo (Proceso de Nacimiento y Muerte)

Consideremos un proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , con un espacio de estados  $E = \{0, 1, 2, \dots, m\}$ , donde dicho proceso modela el tamaño de una población a un tiempo  $t$ , es decir  $X_t = i$  quiere decir que al tiempo  $t$  el número de individuos en la población es igual a  $i$  con  $i \in E$ . Dicha población es afectada por entradas y salidas. Este proceso es conocido como Proceso de Nacimiento y Muerte.

La forma de entrar a la población es mediante nacimientos y la forma de salir es la muerte. Este proceso sólo puede realizar saltos una unidad hacia arriba que se interpreta como un nacimiento, lo cual ocurre con una tasa de intensidad asociada  $\lambda_i$  que sólo depende del estado en el que se está antes de dar el salto, o una unidad hacia abajo que se interpreta como una muerte, que ocurre con la tasa de intensidad  $\mu_i$  que también depende solamente del estado en el que se está antes de que ocurra el salto. Se establece que al tiempo cero el tamaño de la población es cero y que el tamaño de esta puede llegar a un máximo de  $m$  individuos. Una posible trayectoria del comportamiento de este tipo de proceso se observa en la siguiente gráfica.



Gráfica 2.4.1 Trayectoria de un Proceso de Nacimiento y Muerte.

Empezaremos construyendo el generador infinitesimal a partir de las tasas de intensidad, obteniendo la siguiente matriz:

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & \dots & 0 & 0 \\ \mu_1 & -(\mu_1 + \lambda_1) & \lambda_1 & \dots & 0 & 0 \\ 0 & \mu_2 & -(\mu_2 + \lambda_2) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \mu_m & -\mu_m \end{pmatrix}.$$

La tasa  $\lambda_i$  son conocidas como tasas de nacimiento y las tasas  $\mu_i$  tasas de muerte. Cuando la población se encuentra en el estado cero, dado que ya no puede sufrir ni una muerte más, se define  $\mu_0=0$ , y cuando la población se encuentra en el estado  $m$  como ha llegado a su máximo y no puede ocurrir un nacimiento más, se define la tasa  $\lambda_m=0$ .

Como consecuencia del generador infinitesimal  $G$  y el [Teorema 2.2.1](#), el tiempo de estancia en el estado  $i$  tiene distribución exponencial con parámetro  $\lambda_i + \mu_i$ , es decir:

$$T_i \sim \text{Exp}(\lambda_i + \mu_i).$$

Por la [Definición 2.2.5](#) las probabilidades de salto en el proceso de nacimiento y muerte quedan definidas de la siguiente manera:

$$p_{ij} = \begin{cases} \frac{\lambda_i}{\lambda_i + \mu_i} & \text{si } j = i + 1 \\ \frac{\mu_i}{\lambda_i + \mu_i} & \text{si } j = i - 1 \\ 0 & \text{e. o. c.} \end{cases} .$$

Una aplicación de este proceso es la siguiente. Supongamos que tenemos una población en peligro de extinción que en la actualidad cuenta con  $n$  ejemplares ( $X_0=n$ ), dado que dicha población está en peligro de extinción, la tasa de muerte es mayor a la de nacimiento para todo estado, es decir  $\lambda_i < \mu_i$  para cualquier estado  $i \in E$ .

Como implicación directa de esta hipótesis en las tasas de intensidad, la probabilidad de que exista un nacimiento dado que se da un salto es menor a la probabilidad de que exista una muerte, es decir, se cumple lo siguiente:

$$\lambda_i < \mu_i ,$$

entonces:

$$\frac{\lambda_i}{\lambda_i + \mu_i} < \frac{\mu_i}{\lambda_i + \mu_i} ,$$

es decir:

$$p_{i,i+1} < p_{i,i-1} .$$

En dicho experimento el estado cero es un estado absorbente y resulta interesante saber el tiempo esperado en que la población se extinguirá, es decir, el tiempo en que nuestro proceso tarda en llegar al estado cero, donde dicho tiempo cumple con la [Definición 2.2.7](#) de tiempo de paro. Para poder dar respuesta a problemas de este tipo es que se introducirá el concepto de distribuciones tipo fase.

A continuación tenemos una aplicación del [Algoritmo 2.3.1](#) para un caso muy particular del Proceso de Nacimiento y Muerte.

Supongamos que tenemos el proceso de nacimiento y muerte descrito hasta el momento, con las siguientes peculiaridades. Solo contamos con 6 estados, es decir,  $E=\{0,1,2,3,4,5\}$ , cuya distribución inicial  $\pi$  es la siguiente:

$$\pi = ( 0, 0.075, 0.25, 0.35, 0.25, 0.075 ),$$

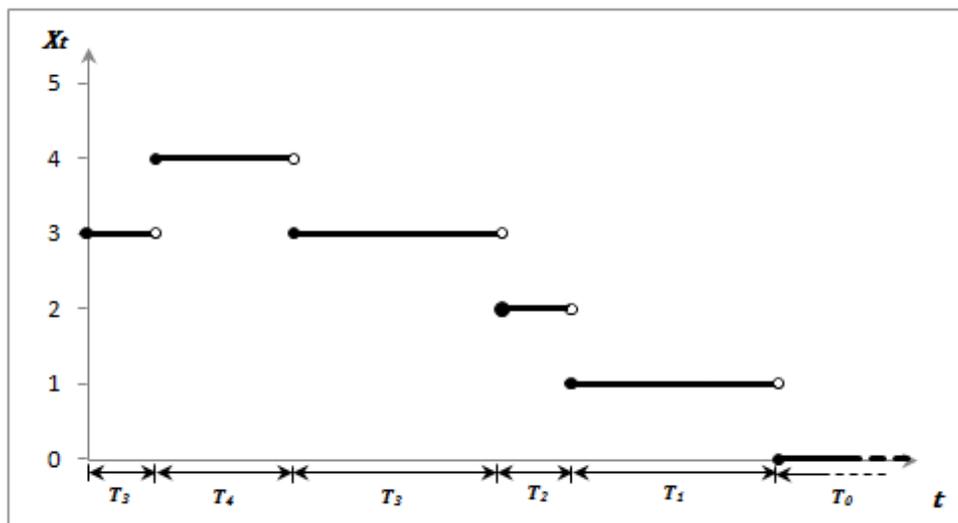
y el generador infinitesimal  $G$  de nuestro proceso es:

$$G = \begin{pmatrix} -0.9 & 0.9 & 0 & 0 & 0 & 0 \\ 1.2 & -1.5 & 0.3 & 0 & 0 & 0 \\ 0 & 0.78 & -1.3 & 0.52 & 0 & 0 \\ 0 & 0 & 0.55 & -1.1 & 0.55 & 0 \\ 0 & 0 & 0 & 0.84 & -1.2 & 0.36 \\ 0 & 0 & 0 & 0 & 1.5 & -1.5 \end{pmatrix}.$$

Con el generador infinitesimal y la [Definición 2.2.5](#) calculamos la matriz de saltos  $P$ , obteniendo:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0.8 & 0 & 0.2 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0.7 & 0 & 0.3 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Con dichas condiciones, el [Algoritmo 2.3.1](#) y tomando como condición de paro de llegar al estado absorbente, es decir, llegar al estado cero; procedemos a simular una posible trayectoria, que se esquematiza con la siguiente gráfica:



Gráfica 2.4.2 Trayectoria simulada de un Proceso de Nacimiento y Muerte.

Esta trayectoria fue simulada con el código del apéndice [A.1](#).

Con la herramienta estudiada hasta el momento sólo se puede obtener una estimación del tiempo esperado de absorción mediante simulaciones Monte Carlo (véase [ 17 ] ), es decir, el tiempo esperado en el que la población llegaría a su extinción.

Por este motivo, en el siguiente capítulo utilizaremos la teoría descrita hasta el momento para introducir el concepto de distribución tipo fase, con lo cual se puede obtener una solución analítica para conocer el valor esperado del tiempo de absorción.

# Capítulo 3

En este capítulo definiremos el concepto de distribución tipo fase, así como sus principales características, ya que el modelo de mortalidad propuesto parte de este concepto.

## Distribuciones Tipo Fase

### Sección 3.1

#### Definiciones Elementales

Consideremos un proceso de saltos de Markov a tiempo continuo  $\{X_{ij}\}_{t \geq 0}$ , con un espacio de estados  $E = \{1, 2, \dots, m, m+1\}$ , donde los estados  $1, 2, \dots, m$  son estados no absorbentes y el estado  $m+1$  es un estado absorbente. Dadas estas condiciones el generador infinitesimal de nuestro proceso tiene la siguiente forma:

$$\Lambda = \begin{pmatrix} G & g \\ \mathbf{0} & 0 \end{pmatrix}, \quad (3.1)$$

donde  $G$  es una matriz cuadrada de dimensión  $m$ ,  $g$  un vector columna de dimensión  $m$  y  $\mathbf{0}$  es un vector fila de dimensión  $m$  donde todas sus entradas son cero. Como  $\Lambda$  es el generador infinitesimal sus filas deben sumar cero, entonces:

$$g = -Ge,$$

donde  $e$  es el vector columna de dimensión  $m$  donde todas las entradas son uno.

$G$  es conocida como la matriz de subintensidades y  $g$  como el vector de intensidades de salida.

Retomemos la distribución inicial  $\pi'$  asociada al proceso  $\{X_{ij}\}_{t \geq 0}$ , es decir:

$$\pi' = (\pi_1, \pi_2, \dots, \pi_m, \pi_{m+1}).$$

Ya que no tiene sentido que el proceso inicie en el estado absorbente  $m+1$ , se define que la probabilidad de iniciar en él es cero, con esto podemos redefinir la distribución inicial de la siguiente manera:

$$\pi = (\pi_1, \dots, \pi_m).$$

Ahora, consideremos una variable aleatoria  $\tau$  como en la expresión ( 2.2 ), pero adaptándola al tiempo que el proceso  $\{X_t\}_{t \geq 0}$  tarda en llegar al estado absorbente  $m+1$ , es decir:

$$\tau = \inf\{t \geq 0 | X_t = m + 1\}, \quad (3.2)$$

y como quedo demostrado en la [Proposición 2.2.7](#), esta variable aleatoria es un tiempo de paro .

Una vez que contamos con estos conceptos, podemos definir lo que es una distribución tipo fase, que hace referencia a la distribución del tiempo de paro  $\tau$  de la expresión ( 3.2 ).

**Definición 3.1.1: (Distribuciones tipo fase).** Una distribución tipo fase con representación  $(\pi, G)$  es la distribución del tiempo de paro  $\tau$  para un proceso de saltos de Markov  $\{X_t\}_{t \geq 0}$ , con espacio de estados  $E = \{1, 2, \dots, m, m+1\}$ , donde los estados  $1, 2, \dots, m$  son estados no absorbentes y el estado  $m+1$  es un estado absorbente, se denota por:

$$\tau \sim PH(\pi, G).$$

Algunas de las propiedades que cumplen las distribuciones tipo fase serán enunciadas a continuación.

## Sección 3.2

### Propiedades

Para hablar de las propiedades que tienen las distribuciones tipo fase empezaremos viendo como es la matriz exponencial del generador infinitesimal.

### Caracterización de las Distribuciones Tipo Fase

**Lema 3.2.1:** Sea  $\Lambda$  el generador infinitesimal como en la expresión ( 3.1 ), entonces:

$$1) \Lambda^n = \begin{pmatrix} G^n & -G^n \mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}.$$

$$2) \exp(\Lambda x) = \begin{pmatrix} \exp(Gx) & \mathbf{e} - \exp(Gx)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}.$$

**Demostración:** 1) La demostración se hará por inducción sobre  $n$ . Para  $n = 1$  tenemos que:

$$\Lambda^1 = \Lambda = \begin{pmatrix} G & g \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} G & -G\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} G^1 & (-G\mathbf{e})^1 \\ \mathbf{0} & 0 \end{pmatrix} = \begin{pmatrix} G^1 & -G^1\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}.$$

Asumimos que se cumple la siguiente hipótesis de inducción:  $\Lambda^n = \begin{pmatrix} G^n & -G^n\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}$ .

Para  $n+1$ ,

$$\begin{aligned} \Lambda^{n+1} &= \Lambda^n \Lambda \\ &= \begin{pmatrix} G^n & -G^n\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} \begin{pmatrix} G & -G\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} \\ &= \begin{pmatrix} G^n G - G^n\mathbf{e}\mathbf{0} & G^n(-G\mathbf{e}) + (-G^n\mathbf{e})0 \\ \mathbf{0}G + \mathbf{0}0 & \mathbf{0}(-G\mathbf{e}) + 0 \end{pmatrix} \\ &= \begin{pmatrix} G^{n+1} & -(G^n G)\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} \\ &= \begin{pmatrix} G^{n+1} & -G^{n+1}\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix}. \end{aligned}$$

2) Si  $\mathbf{I}_n$  es la matriz identidad de dimensión  $n$ ,  $n \in \mathbb{N}$  y  $t \geq 0$ :

$$\begin{aligned} \exp(\Lambda t) &= \sum_{n=0}^{\infty} \frac{\Lambda^n t^n}{n!} \\ &= \mathbf{I}_{m+1} + \sum_{n=1}^{\infty} \frac{\Lambda^n t^n}{n!} \\ &= \mathbf{I}_{m+1} + \sum_{n=1}^{\infty} \begin{pmatrix} (Gt)^n/n! & -(Gt)^n\mathbf{e}/n! \\ \mathbf{0} & 0 \end{pmatrix} \\ &= \mathbf{I}_{m+1} + \begin{pmatrix} \exp(Gt) - \mathbf{I}_m & -(\exp(Gt) - \mathbf{I}_m)\mathbf{e} \\ \mathbf{0} & 0 \end{pmatrix} \\ &= \begin{pmatrix} \exp(Gt) & \mathbf{e} - \exp(Gt)\mathbf{e} \\ \mathbf{0} & 1 \end{pmatrix}. \quad \blacksquare \end{aligned}$$

El siguiente teorema nos ayudará a obtener la función de distribución de la variable aleatoria  $\tau$ .

**Teorema 3.2.1:** Sea  $\tau \sim PH(\pi, G)$ , entonces para todo  $t \geq 0$ :

- 1)  $\mathbb{P}(\tau > t) = \pi \exp(Gt) \mathbf{e}$
- 2)  $\tau$  tiene densidad  $f_\tau$  dada por:

$$f_\tau(t) = \pi \exp(Gt) g.$$

**Demostración:** 1) Observemos que  $\tau > t$  si y sólo si  $X_t \in \{1, 2, \dots, m\}$ . Por la ley de la probabilidad total,

$$\begin{aligned} \mathbb{P}(\tau > t) &= \pi \mathbb{P}(X_t \in \{1, 2, \dots, m\}) \\ &= \sum_{j=1}^m \sum_{i=1}^m \mathbb{P}(X_t = j | X_0 = i) \mathbb{P}(X_0 = i) \\ &= \sum_{i,j \in E \setminus \{m+1\}} \pi_i \exp(Gt)_{ij} \\ &= \pi \exp(Gt) \mathbf{e}. \end{aligned}$$

2)  $f_\tau(t) dt$  es la probabilidad de que  $\tau \in [t, t+dt)$ . Una vez más por la ley de la probabilidad total:

$$\begin{aligned} f_\tau(t) dt &= \mathbb{P}(\tau \in [t, t+dt)) \\ &= \sum_{i=1}^m \pi_i \sum_{j=1}^m p_{ij}(t) g_j dt \\ &= \sum_{i=1}^m \sum_{j=1}^m \pi_i \exp(Gt)_{ij} g_j dt \\ &= \pi \exp(Gt) g dt. \quad \blacksquare \end{aligned}$$

Como consecuencia directa de este teorema la función de distribución es la siguiente.

**Corolario 3.2.1:** La función de distribución  $F_\tau$ , está dada por:

$$F_\tau(t) = 1 - \pi \exp(Gt) \mathbf{e}.$$

## Función Generadora de Momentos

Una propiedad muy importante que cumple la matriz de subintensidades  $G$ , que nos ayuda a simplificar varios cálculos y hace que las distribuciones tipo fase cumpla con otras propiedades, es que  $G$  es invertible, esto queda demostrado con el siguiente teorema.

**Teorema 3.2.2:** La matriz  $G$  es invertible si y sólo si los estados  $1, 2, \dots, m$  son no absorbentes.

**Demostración:** Supongamos  $G$  invertible, entonces sea  $a_i$  la probabilidad de absorción eventual iniciando desde el estado  $i$ .

Condicionando en la primera transición,

$$\begin{aligned} a_i &= p_{i,m+1} + \sum_{j \in E \setminus \{i,m+1\}} p_{i,j} a_j, \\ &= -\frac{g_i}{G_{ii}} - \sum_{j \in E \setminus \{i,m+1\}} \frac{G_{ij}}{G_{ii}} a_j. \end{aligned}$$

Entonces;

$$-a_i G_{ii} = g_i + \sum_{j \in E \setminus \{i,m+1\}} G_{ij} a_j,$$

es decir:

$$\sum_{j \in E \setminus \{m+1\}} G_{ij} a_j + g_i = 0.$$

Si  $\mathbf{a}$  es el vector columna de dimensión  $m$  cuyas entradas son los valores  $a_i$ , en notación vectorial esto es;

$$\mathbf{Ga} + \mathbf{g} = \mathbf{0}.$$

Como  $G$  es invertible, entonces:

$$\mathbf{a} = -G^{-1}\mathbf{g} = G^{-1}\mathbf{Ge} = \mathbf{e}.$$

Entonces  $a_i = 1$  quiere decir que la absorción eventual es segura partiendo desde el estado  $i$  para toda  $i \in E$  distinta de  $m+1$ , por lo tanto los estados  $1, 2, \dots, m$  son no absorbentes.

Ahora para demostrar el regreso, por contradicción, supongamos que los estados  $1, 2, \dots, m$  son no absorbentes y que  $G$  es singular, es decir, existe un vector fila  $\mathbf{y} \neq \mathbf{0}$  de dimensión  $m$  tal que

$$\mathbf{yG} = \mathbf{0}.$$

Entonces para todo  $t \geq 0$ ,

$$\begin{aligned}
 y \exp(Gt) y^T &= y \sum_{n=0}^{\infty} \frac{G^n t^n}{n!} y^T \\
 &= y I_m y^T + y \sum_{n=1}^{\infty} \frac{G^n t^n}{n!} y^T \\
 &= y y^T + y G \sum_{n=0}^{\infty} \frac{G^n t^n}{n!} y^T \\
 &= y y^T .
 \end{aligned}$$

Como  $y \neq 0$ , entonces  $y y^T > 0$ .

Ya que  $\{\exp(Gt)\}_{ij}$  es la probabilidad de, dado que se está en el estado  $i$ , después de un tiempo  $t$  estar en el estado  $j$ , el  $\lim_{t \rightarrow \infty} \exp(Gt)$  existe y es finito, esto debido a que  $\exp(Gt)$  es acotada entrada a entrada, entonces:

$$y \lim_{t \rightarrow \infty} \exp(Gt) y^T = \lim_{t \rightarrow \infty} y \exp(Gt) y^T = y y^T > 0$$

entonces;

$$\lim_{t \rightarrow \infty} \exp(Gt) \neq \mathbf{0}$$

es decir, existen  $i, j \in \{E \setminus m+1\}$  tal que;

$$\lim_{t \rightarrow \infty} \exp(Gt)_{ij} \neq 0$$

entonces;

$$\lim_{t \rightarrow \infty} \mathbb{P}(X_t = j | X_0 = i) = \lim_{t \rightarrow \infty} \exp(Gt)_{ij} \neq 0 .$$

Es decir, si se empieza en el estado  $i$ , la probabilidad de encontrarse en el estado  $j$  cuando  $t \rightarrow \infty$  será estrictamente mayor que  $0$ . Esto querría decir que el estado  $m+1$  no es un estado absorbente, lo cual implica una contradicción, que se genera de suponer  $G$  singular. Por lo tanto si los estados  $1, 2, \dots, m$  son no absorbentes, entonces  $G$  es invertible. ■

Una característica importante de una variable aleatoria son sus momentos, ya que con ellos se pueden inferir muchas características de dicha variable. Para poder obtener una expresión del  $n$ -ésimo momento de una distribución tipo fase, demostraremos el siguiente teorema donde obtenemos la forma que tiene la transformada de Laplace.

**Teorema 3.2.3:** La transformada de Laplace  $L_\tau(s) = \mathbb{E}[\exp(-s\tau)]$  con  $s \geq 0$  de  $\tau \sim PH(\pi, G)$  está dado por:

$$L_\tau(s) = \pi(s\mathbf{I}_m - G)^{-1}g.$$

**Demostración:**

$$\begin{aligned} L_\tau(s) &= \int_0^\infty \exp(-st) f_\tau(t) dt \\ &= \int_0^\infty \exp(-st) \pi \exp(Gt) g dt \\ &= \int_0^\infty \pi \exp(-st\mathbf{I}_m) \exp(Gt) g dt \\ &= \pi \int_0^\infty \exp[t(G - s\mathbf{I}_m)] dt g \end{aligned}$$

como  $(G - s\mathbf{I}_m)$  es de subintensidad para todo  $s \geq 0$ , entonces;

$$\begin{aligned} L_\tau(s) &= -\pi(G - s\mathbf{I}_m)^{-1}g \\ &= \pi(s\mathbf{I}_m - G)^{-1}g \end{aligned} \quad \blacksquare$$

Con la ayuda de la transformada de Laplace podemos demostrar el siguiente teorema, el cual nos indica la forma que tiene en el n-ésimo momento de una distribución tipo fase.

**Teorema 3.2.4:** El n-ésimo momento de  $\tau \sim PH(\pi, G)$  es:

$$\mathbb{E}[\tau^n] = n! (-1)^n \pi (G^{-1})^n \mathbf{e}.$$

**Demostración:** La n-ésima derivada de la transformada de Laplace es:

$$L_\tau^{(n)}(s) = \frac{d^n}{ds^n} \pi(s\mathbf{I} - G)^{-1}g = \pi \left[ \frac{d^n}{ds^n} (s\mathbf{I} - G)^{-1} \right] g = \pi n! (-1)^n (s\mathbf{I} - G)^{-n-1}g.$$

Entonces:

$$\begin{aligned} \mathbb{E}[\tau^n] &= (-1)^n L_\tau^{(n)}(0) \\ &= (-1)^n [\pi n! (-1)^n (-G)^{-n-1}g] \\ &= n! \pi (-G)^{-n} (-G)^{-1} (-G\mathbf{e}) \\ &= (-1)^n n! \pi G^{-n} G^{-1} G\mathbf{e} \\ &= n! (-1)^n \pi (G^{-1})^{-n} \mathbf{e}. \end{aligned} \quad \blacksquare$$

Enseguida daremos un ejemplo para aclarar los conceptos enunciados hasta el momento.

## Sección 3.3

### Ejemplo

Retomando el proceso de nacimiento y muerte como en el ejemplo de la [Sección 2.4](#), con la peculiaridad que el espacio de estados es  $E=\{1, 2, 3, 0\}$  y suponiendo el estado cero como el estado absorbente, es decir, cuando la población llega a la extinción permanece allí el resto del tiempo. De esta forma la matriz  $G$  de subintensidades y la distribución inicial  $\pi$  son las siguientes:

$$G = \begin{pmatrix} -(\mu_1 + \lambda_1) & \lambda_1 & 0 \\ \mu_2 & -(\mu_2 + \lambda_2) & \lambda_2 \\ 0 & \mu_3 & -\mu_3 \end{pmatrix},$$

$$\pi = (\pi_1, \pi_2, \pi_3).$$

Calculando el vector  $g$  como  $-Ge$ :

$$g = -Ge = \begin{pmatrix} -(\mu_1 + \lambda_1) & \lambda_1 & 0 \\ \mu_2 & -(\mu_2 + \lambda_2) & \lambda_2 \\ 0 & \mu_3 & -\mu_3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ 0 \\ 0 \end{pmatrix},$$

el generador infinitesimal es el siguiente:

$$\Lambda = \begin{pmatrix} -(\mu_1 + \lambda_1) & \lambda_1 & 0 & \mu_1 \\ \mu_2 & -(\mu_2 + \lambda_2) & \lambda_2 & 0 \\ 0 & \mu_3 & -\mu_3 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Una vez que contamos con estos elementos, las funciones de densidad y distribución de la variable aleatoria  $\tau$  quedarían expresadas de la siguiente manera:

$$f_\tau(t) = (\pi_1, \pi_2, \pi_3) \exp \begin{pmatrix} -t(\mu_1 + \lambda_1) & t\lambda_1 & 0 \\ t\mu_2 & -t(\mu_2 + \lambda_2) & t\lambda_2 \\ 0 & t\mu_3 & -t\mu_3 \end{pmatrix} \begin{pmatrix} \mu_1 \\ 0 \\ 0 \end{pmatrix},$$

$$F_\tau(t) = 1 - (\pi_1, \pi_2, \pi_3) \exp \begin{pmatrix} -t(\mu_1 + \lambda_1) & t\lambda_1 & 0 \\ t\mu_2 & -t(\mu_2 + \lambda_2) & t\lambda_2 \\ 0 & t\mu_3 & -t\mu_3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

La matriz inversa de  $G$  que nos sirve para calcular el  $n$ -ésimo momento de  $\tau$ , es la siguiente:

$$G^{-1} = \begin{pmatrix} \frac{-1}{\mu_1} & \frac{-\lambda_1}{\mu_1\mu_2} & \frac{-\lambda_1\lambda_2}{\mu_1\mu_2\mu_3} \\ -1 & \frac{-(\mu_1 + \lambda_1)}{\mu_1\mu_2} & \frac{-\lambda_2(\mu_1 + \lambda_1)}{\mu_1\mu_2\mu_3} \\ \frac{-1}{\mu_1} & \frac{-(\mu_1 + \lambda_1)}{\mu_1\mu_2} & \frac{-(\mu_1\lambda_2 + \mu_1\mu_2 + \lambda_1\lambda_2)}{\mu_1\mu_2\mu_3} \end{pmatrix}.$$

De esta manera el  $n$ -ésimo momento de  $\tau$  queda de la siguiente manera:

$$\mathbb{E}[\tau^n] = (-1)^n n! (\pi_1, \pi_2, \pi_3) (G^{-1})^n \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

por ejemplo, el valor esperado de  $\tau$  tiene la siguiente expresión:

$$\begin{aligned} \mathbb{E}[\tau] &= (-1)^1 1! (\pi_1, \pi_2, \pi_3) (G^{-1})^1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= (\pi_1, \pi_2, \pi_3) \begin{pmatrix} \frac{(\lambda_1 + \mu_2)(\lambda_2 + \mu_3) - \mu_2\lambda_2}{\mu_1\mu_2\mu_3} \\ \frac{\mu_2\mu_3 + (\mu_1 + \lambda_1)(\lambda_2 + \mu_3)}{\mu_1\mu_2\mu_3} \\ \frac{\mu_2(\mu_3 - \lambda_1) + (\mu_1 + \lambda_1)(\mu_2 + \lambda_2 + \mu_3)}{\mu_1\mu_2\mu_3} \end{pmatrix}. \end{aligned}$$

Si lo que tenemos es una distribución tipo fase que represente el proceso de nacimiento y muerte con los siguientes parámetros:

$$G = \begin{pmatrix} -1 & 0.5 & 0 \\ 0.7 & -1.1 & 0.4 \\ 0 & 0.8 & -0.8 \end{pmatrix},$$

$$\pi = (0.25, 0.5, 0.25).$$

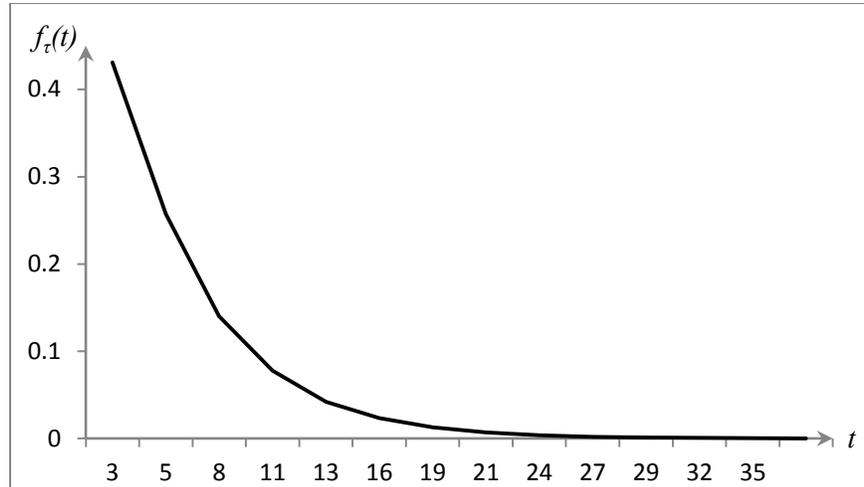
Con esta matriz de sub intensidades el vector  $g$  sería:

$$g = - \begin{pmatrix} -1 & 0.5 & 0 \\ 0.7 & -1.1 & 0.4 \\ 0 & 0.8 & -0.8 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0 \\ 0 \end{pmatrix}.$$

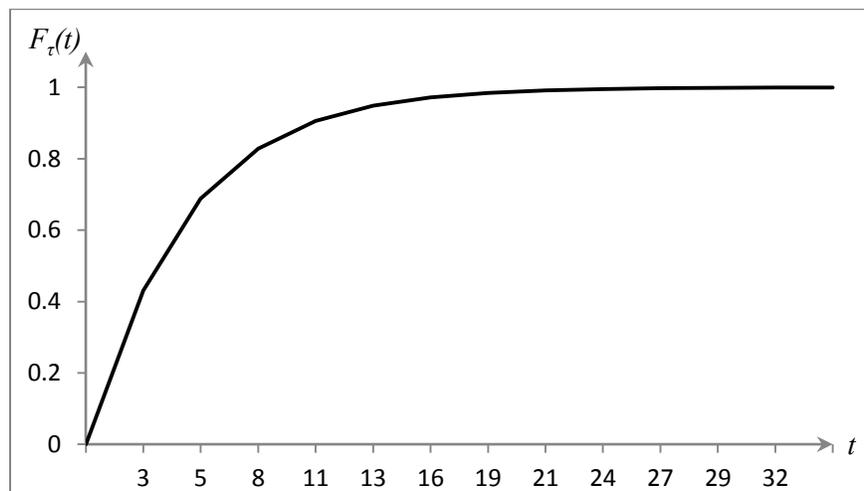
Las funciones de densidad y distribución quedarían de la siguiente manera:

$$f_{\tau}(t) = (0.25, 0.5, 0.25) \exp \begin{pmatrix} -0.9t & 0.9t & 0 \\ 0.7t & -1.1t & 0.4t \\ 0 & 0.8 & -1.5t \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$F_{\tau}(t) = 1 - (0.25, 0.5, 0.25) \exp \begin{pmatrix} -0.9t & 0.9t & 0 \\ 0.7t & -1.1t & 0.4t \\ 0 & 0.8 & -1.5t \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$



Gráfica 3.3.1 Función de densidad



Gráfica 3.3.2 Función de distribución.

La matriz inversa de la matriz de subintensidades es:

$$G^{-1} = \begin{pmatrix} \frac{-1}{0.5} & \frac{-0.5}{0.5(0.7)} & \frac{-0.5(0.4)}{0.5(0.7)(0.8)} \\ \frac{-1}{0.5} & \frac{-1}{0.5(0.7)} & \frac{-1(0.4)}{0.5(0.7)(0.8)} \\ \frac{-1}{0.5} & \frac{-1}{0.5(0.7)} & \frac{-[0.5(0.4) + 0.5(0.7) + 0.5(0.4)]}{0.5(0.7)(0.8)} \end{pmatrix}$$

$$= \begin{pmatrix} -2 & -1.429 & -0.714 \\ -2 & -2.857 & -1.429 \\ -2 & -2.857 & -2.679 \end{pmatrix}.$$

Con ella podemos calcular el valor esperado de la variable aleatoria:

$$\mathbb{E}[\tau] = (0.25, 0.5, 0.25) \begin{pmatrix} \frac{(0.5 + 0.7)(0.4 + 0.8) - 0.7(0.4)}{0.5(0.7)(0.8)} \\ \frac{0.7(0.8) + (0.5 + 0.5)(0.4 + 0.8)}{0.5(0.7)(0.8)} \\ \frac{0.7(0.8 - 0.5) + (0.5 + 0.5)(0.7 + 0.4 + 0.8)}{0.5(0.7)(0.8)} \end{pmatrix} = 6.0625 .$$

Una vez que contamos con el concepto de distribución tipo fase, se dará paso al modelo de mortalidad.

# Capítulo 4

En este capítulo se describirán dos modelos para valorar el riesgo de mortalidad por medio de una distribución tipo fase. El primero, el modelo propuesto por X. Sheldon Lin y Xiaoming Liu, es un antecedente importante, y el segundo, que corresponde al modelo de mortalidad propuesto en este trabajo.

## Mortalidad a Través de Distribuciones Tipo Fase

La medición del riesgo de mortalidad dentro de una población ha adquirido gran importancia debido al alto impacto dentro de valuaciones financieras como lo son los seguros de vida y los sistemas de pensiones. Es por ello que contar con un modelo adecuado para la estimación de ésta resulta de suma importancia.

En la actualidad dichas valuaciones se llevan a cabo utilizando tablas de mortalidad<sup>6</sup>, la desventaja de medir el riesgo de mortalidad con esta herramienta es que se vuelve una valuación determinista incapaz de tomar en cuenta las variaciones que ésta pueda sufrir en distintos escenarios. Es por eso que la necesidad de contar con un modelo que brinde un grado de aleatoriedad a dichas valuaciones adquiere una gran importancia. Se han llevado a cabo muchos esfuerzos por obtener un modelo adecuado, un ejemplo de ello es el propuesto por X. Sheldon Lin y Xiaoming Liu (abril 2008), el cual se describe a continuación.

### Sección 4.1

#### Modelo de X. Sheldon Lin y Xiaoming Liu

En abril de 2008 X. Sheldon Lin y Xiaoming Liu dentro de su documento “*Markov Aging Process and Phase-Type Law of Mortality*” (véase [ 18 ] ); proponen modelar el tiempo de muerte mediante una distribución tipo fase. Con esta propuesta buscan disponer de un modelo de mortalidad que se adapte a los datos observados y pueda vincular sus parámetros con el mecanismo biológico y fisiológico de envejecimiento; lo que permitiría considerar

---

<sup>6</sup> En el mercado asegurador mexicano la Comisión Nacional de Seguros y Fianzas publica las tablas de mortalidad utilizadas por el sector. Las tablas actualmente utilizadas son la tabla CNSF 2000-I para vida individual y CNSF 2000-G para vida grupo.

las opiniones de los expertos en dicho mecanismo con mayor facilidad, además de permitir el análisis cuantitativo de la muerte.

En dicho modelo, suponen que la mortalidad es descrita mediante un proceso de saltos de Markov a tiempo continuo de estados finitos, con un estado absorbente que es la muerte, lo que garantiza que el tiempo de arribo a este estado, es decir, el tiempo de ocurrencia de la muerte sigue una distribución tipo fase.

El lograr modelar la mortalidad mediante una distribución tipo fase resulta bastante ingenioso, ya que aunque a simple vista parece imposible poder segregar la vida del ser humano en un conjunto de estados cuyo tiempo de estancia en cada uno de estos se distribuya de manera exponencial, se logra al introducir el concepto de edad fisiológica. Es por ello que la propuesta de los estados del proceso de saltos de Markov parte de la teoría del envejecimiento del ser humano a través del tiempo, que describe la disminución genéticamente determinada, progresiva e irreversible de la capacidad de adaptarse a su entorno, que se manifiesta como la susceptibilidad a contraer ciertas enfermedades y culminar con la muerte del individuo.

Motivado por esto se introduce el concepto de edad fisiológica, que puede interpretarse como un índice relativo de salud que representa el grado de envejecimiento en el individuo. Al tratarse de un índice relativo, el tiempo de permanencia de un individuo en determinada edad fisiológica puede no ser considerado dentro de un intervalo acotado de tiempo, por lo que se puede suponer que este tiempo tiene un comportamiento exponencial. Así es como cada edad fisiológica pasa a ser un estado del proceso de Markov, por lo que durante el tiempo que el individuo permanece en este estado, sus capacidades para adaptarse a su entorno son las mismas y cuando éste pasa a la siguiente edad fisiológica es porque estas capacidades se vieron afectadas. Una vez contextualizado el problema, lo que sigue es la construcción de los parámetros para la distribución tipo fase y así poder llegar a la cuantificación del fenómeno de estudio.

Para que la distribución tipo fase que modela el fenómeno de mortalidad quede caracterizada es necesario contar con sus dos parámetros, la matriz de subintensidades  $G$  y la distribución inicial  $\pi$ . Para ambos parámetros los autores proponen características y supuestos especiales.

La dimensión de estos parámetros queda determinada a partir del número de estados no absorbentes que se tienen, es decir el número de edades fisiológicas, de esta forma si suponemos de manera general que se cuenta con  $m$  edades fisiológicas, la matriz  $G$  es una matriz cuadrada de dimensión  $m$  y la distribución  $\pi$  un vector fila de dimensión  $m$ .

Hablando de la matriz de subintensidades, se tiene que entre sus características sólo cuenta con valores en la diagonal y las entradas  $(i, i+1)$  con  $i \in \{1, 2, \dots, m-1\}$ , esto ya que si estamos en una edad fisiológica  $i$  sólo se puede pasar a la edad fisiológica  $i+1$  o morir. Los valores con los que se deben de llenar estas entradas son propuestos de manera muy especial. Para la diagonal se propone utilizar dos parámetros, el primero que describe el proceso de envejecimiento con el que un individuo pasa de una edad fisiológica a la siguiente, llamado tasa de envejecimiento y que se denota como  $\lambda_i$ . El segundo es el que describe la susceptibilidad que se tiene a la muerte, llamado tasa de mortalidad y denotado por  $\mu_i$ . La suma de estas tasas es el parámetro de la exponencial que describe el tiempo de estancia en la edad fisiológica  $i$  (la construcción de estas tasas es descrita más adelante). Entonces la matriz de subintensidades asociada a la distribución tipo fase es de la siguiente manera:

$$G = \begin{pmatrix} -(\lambda_1 + \mu_1) & \lambda_1 & \cdots & 0 \\ 0 & -(\lambda_2 + \mu_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\mu_m \end{pmatrix};$$

con lo que obtenemos la siguiente matriz de intensidades:

$$\Lambda = \begin{pmatrix} -(\lambda_1 + \mu_1) & \lambda_1 & \cdots & 0 & \mu_1 \\ 0 & -(\lambda_2 + \mu_2) & \cdots & 0 & \mu_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\mu_m & \mu_m \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

La distribución inicial  $\pi$  depende directamente de la edad fisiológica en la que se encuentre el individuo, empezando con probabilidad uno en ésta y cero en todas las demás. Si suponemos que todo individuo empieza en el estado 1, entonces la distribución inicial sería:

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

De esta forma la mortalidad queda modelada por la distribución tipo fase con representación  $PH(\pi, G)$ .

Las tasas  $\lambda_i$  y  $\mu_i$  son propuestas de una manera muy especiales. Los autores justifican con base al lento descenso que presentan las capacidades fisiológicas observadas en distintos estudios experimentales, que  $\lambda_i$  tenga la siguiente forma:

$$\lambda_i = \lambda \text{ para } i = k + 1, \dots, m - 1;$$

donde  $k$  es un número de estados o edades fisiológicas en los que el individuo tarda en adaptarse a su medio ambiente desde el nacimiento hasta alcanzar su rendimiento fisiológico máximo. Los autores proponen que  $k$  no sea mayor a cuatro.

Para la tasa  $\mu_i$ , se parte del supuesto que existen dos tipos de muertes, las ocasionadas por el deterioro fisiológico y las muertes independientes al envejecimiento, por ejemplo las causadas por accidentes. Por tal motivo la tasa de mortalidad  $\mu_i$  es la siguiente:

$$\mu_i = h_1(i) + h_2(i),$$

donde  $h_1(i)$  representa la tasa asociada a las muertes ocasionadas por motivos independientes al envejecimiento y  $h_2(i)$  la tasa de las muertes asociadas al deterioro fisiológico.

La tasa  $h_1(i)$  se propone como:

$$h_1(i) = \begin{cases} a + b & \text{para } i_1 < i \leq i_2 \\ b & \text{e. o. c} \end{cases},$$

donde la constante  $b$  se considera como una tasa base y la constante  $a$  es una tasa asociada a los accidentes relacionados con el comportamiento de la edad, es por ello que sólo aparece entre los estados  $i_1$  e  $i_2$ . Estos dos estados pueden representar por ejemplo aquellas edades en que el ser humano es más susceptible a morir en un accidente.

Se propone la tasa  $h_2(i)$  como una función creciente con respecto de  $i$ , para representar el aumento en la susceptibilidad a la muerte con respecto al envejecimiento. Se asume  $h_2(i)$  como una función potencia de la forma:

$$h_2(i) = i^r s,$$

donde  $s$  es un parámetro de escala y  $r$  mide la susceptibilidad en relación al envejecimiento. De esta forma la tasa  $\mu_i$  se expresa como:

$$\mu_i = \begin{cases} a + b + i^r s & \text{para } i_1 < i \leq i_2 \\ b + i^r s & \text{e. o. c} \end{cases}.$$

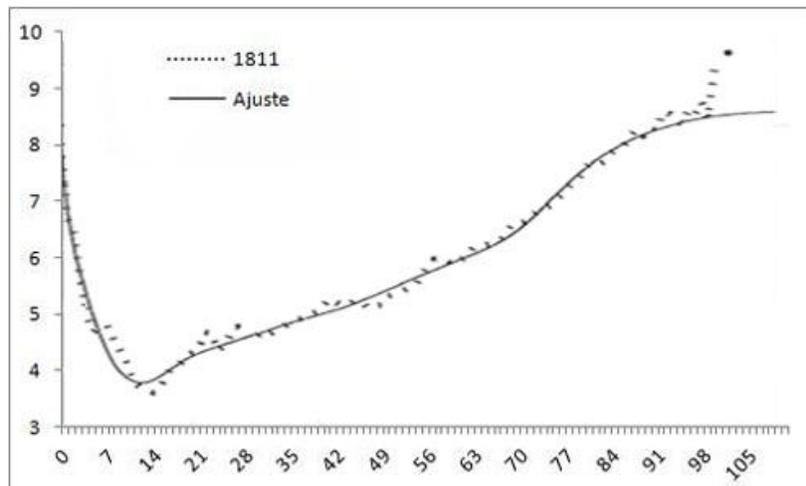
El hecho de que la tasa de mortalidad esté compuesta por varios parámetros hace que la realización de escenarios bajo desviaciones de la mortalidad por causas específicas se pueda hacer con facilidad.

Se realizaron dos aplicaciones de este modelo, la primera de ellas en las cohortes suecas de 1811, 1861 y 1911, utilizando los datos observados de mortalidad de dichas cohortes. Y la segunda a los datos de mortalidad de la seguridad social en E.U. utilizando la tabla que se conoce como “SSA Tabla 1950B” que pertenece al estudio actuarial No. 107.

Para la aplicación en las cohortes suecas fue utilizado en algoritmo simplex “Nelder y Mead 1965”<sup>7</sup> a diferentes combinaciones de valores  $m$ ,  $i_1$  e  $i_2$  y así determinar la máxima edad fisiológica y el rango de edad para alta incidencia de muerte accidental. Con ello se determino  $m=200$  para las tres cohortes y rangos diferentes para cada una de estas; estos rangos así como los parámetros estimados se muestran en la siguiente tabla:

Cohorte	Parametros					
	$\lambda$	$b$	$a$	$[i_1, i_2]$	$s$	$r$
1811	2.5657	3.1504E-03	1.9888E-03	[ 42, 99 ]	9.3157E-09	3
1861	2.4794	4.4825E-03	1.9033E-03	[ 42, 89 ]	2.6351E-13	5
1911	2.3707	9.0987E-04	2.8939E-03	[ 33, 70 ]	1.8872E-15	6

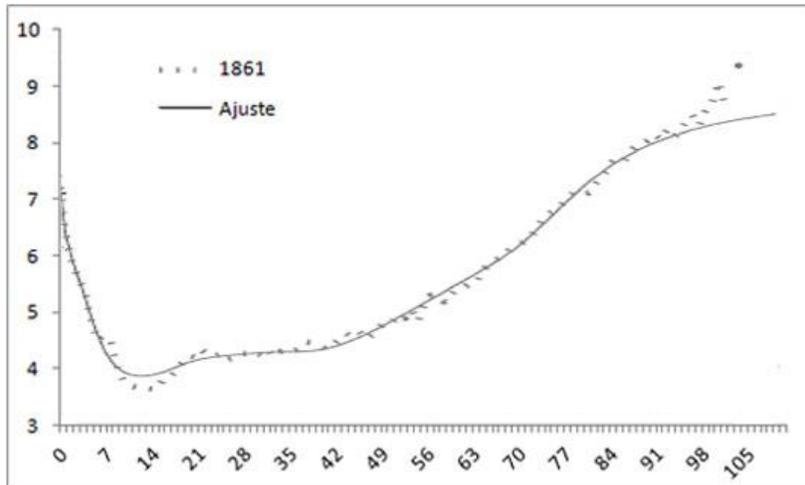
El ajuste obtenido para cada una de las cortes se puede observar en las siguientes graficas<sup>8</sup>.



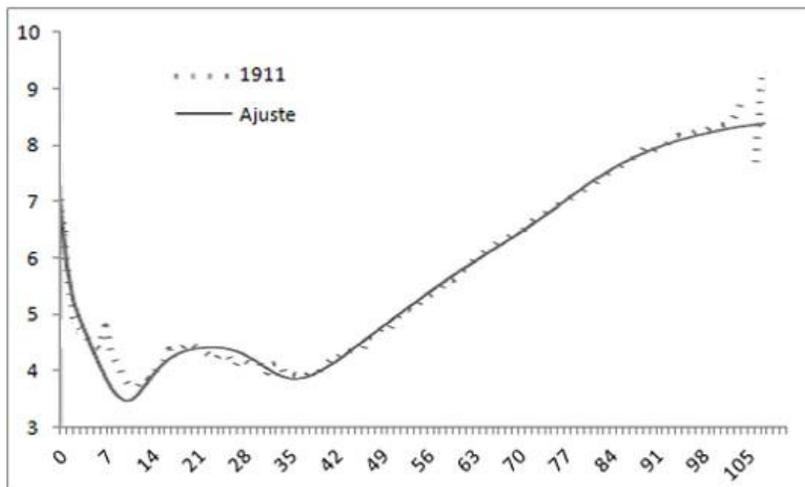
Gráfica 4.1.1 Ajuste a la cohorte de 1811.

<sup>7</sup> El algoritmo simplex Nelder-Mead es un método de búsqueda directa para la minimización sin restricciones de funciones multidimensionales. Desde su primera publicación en 1965, este algoritmo se ha convertido en uno de los métodos más ampliamente usados para la optimización no lineal sin restricciones. A pesar de ser ampliamente usado, apenas se han probado de forma explícita resultados teóricos del algoritmo. El método Nelder-Mead trata de minimizar una función escalar no lineal de  $n$  variables usando sólo valores de la función, sin obtener ninguna información de la derivada (ni implícita ni explícitamente).

<sup>8</sup> Las gráficas 4.1, 4.2, 4.3 y 4.4 están en la escala  $\ln(10000 \cdot q_x)$ .



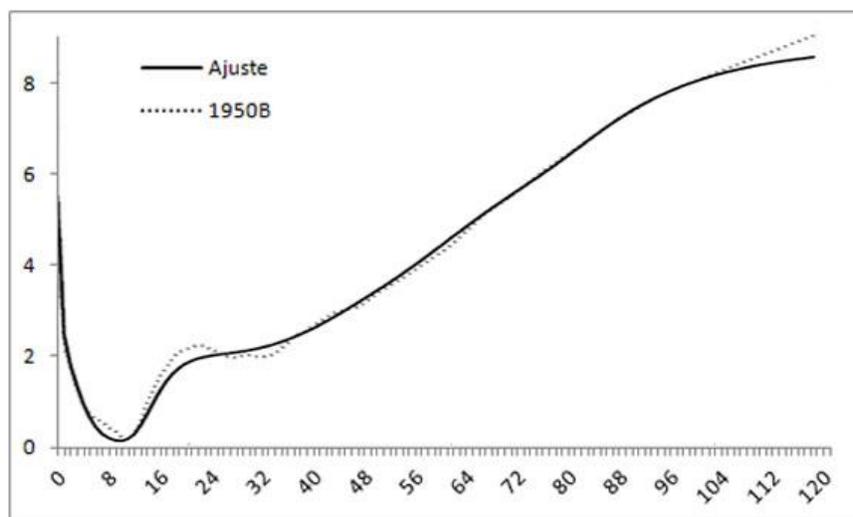
Gráfica 4.1.2 Ajuste a la cohorte de 1861.



Gráfica 4.1.3 Ajuste a la cohorte de 1911.

En la aplicación a los datos de la seguridad social en E.U. también fue utilizado el algoritmo simplex “Nelder y Mead 1965”, pero en este caso fue para reducir al mínimo la suma de errores cuadrados ponderados. A continuación tenemos los parámetros estimados para esta implementación así como la grafica del ajuste.

Parametros					
$\lambda$	$b$	$a$	$[i_1, i_2]$	$s$	$r$
2.2484	2.7671E-04	1.5059E-03	[ 37, 120 ]	2.171E-15	6



Gráfica 4.1.4 Ajuste a la tabla “SSA Tabla 1950B”

Las pruebas de ajuste realizadas a ambas implementaciones fueron Lee-Carter y Heligman-Pollard, así como el cálculo del valor presente actuarial de ciertos seguros y rentas vitalicias. Bajo las tres pruebas los resultados del modelo resultan ser bastante buenos.

## Sección 4.2

### Modelo Alternativo de Mortalidad

El modelo anterior nos abre el panorama acerca de las herramientas para poder medir riesgos como el de mortalidad, ya que la forma en que se utiliza una distribución tipo fase resulta ser bastante ingeniosa y atractiva para su aplicación.

La problemática con la que nos enfrentamos al tratar de implementar dicho modelo es que la estimación de sus parámetros parece ser bastante compleja y demasiada la información que se necesita. Para poder clasificar a un asegurado dentro de una edad fisiológica sería necesario conocer el estado de salud en el que se encuentra, por lo que si pensamos en los productos de venta masiva, por ejemplo los otorgados por el banco al momento de obtener un crédito o los que otorgan algunas empresas a sus empleados; en estos no se realiza un cuestionario de salud a cada asegurado por lo que no se sabría la edad fisiológica en la que se encuentra, por otro lado una vez que el asegurado entra a la cartera de la institución, tampoco se está monitoreando el estado de salud de este durante el periodo de cobertura.

Otro problema es el detalle que se necesita para estimar las tasas de mortalidad, ya que es necesario conocer los detalles de las muertes para saber en cuál de todos los componentes de la tasa va a ser incluido dicho dato. Y aunque estos dos puntos se podrían empezar a

subsana a partir de ahora, el tiempo en el que la institución se tardaría en recolectar la información suficiente para poder estimar los parámetros de una manera adecuada, es considerable, por lo que pensar en la implementación de este modelo se hace prácticamente imposible.

Es por esto que lo que se busca proponer dentro de este trabajo es un modelo que sirva para medir el riesgo por mortalidad a través de una distribución tipo fase, haciendo los supuestos necesarios para que la información que se necesite para su implementación sea la mínima posible y así tratar de garantizar que al momento que se quiera implementar se cuente con ella. Lo cual sería de gran ayuda para empresas que acaban de iniciar operaciones o con poca experiencia como para poder implementar un modelo robusto.

## Desarrollo

Recordemos que una distribución tipo fase surge de un proceso de saltos de Markov a tiempo continuo con espacio de estados finito y un estado absorbente, además ésta queda definida por sus dos parámetros, la matriz de subintensidades y la distribución inicial. Por tanto, para que nuestro modelo quede bien definido es necesario determinar cómo serán el espacio de estados, la matriz de subintensidades y la distribución inicial.

Empezaremos por el espacio de estados. El más fácil de determinar es el estado absorbente, que es el que representa la muerte y estará denotado por  $d$ . Para los estados no absorbentes que representarán el tiempo en que el individuo permanece con vida, su construcción parece no ser fácil y nada intuitiva, ya que si recordamos el tiempo de estancia en cada uno de estos tiene una distribución exponencial, y si pensamos en poder segregar la vida del ser humano en un conjunto finito que cumplan con esta característica pareciera imposible, es por eso que los supuestos utilizados en esta parte del modelo resultan ser bastante especiales.

Para poder determinar el conjunto de estados cuyo tiempo de estancia se distribuye exponencial es necesario no perder el concepto de edad fisiológica, que se interpreta como un índice relativo de salud que representa el grado de envejecimiento en el individuo. Teniendo en cuenta que el tiempo que un individuo permanece en alguna de estas edades no necesariamente es concebido dentro de un intervalo de tiempo acotado, nos permite cumplir con la distribución exponencial y así decir que cada estado no absorbente estaría representado por una edad fisiológica. Sin embargo, que este concepto esté ligado directamente con el estado de salud en el que se encuentra el individuo, nos hace regresar al problema de la falta de información para su determinación. Con esto, el problema al que nos enfrentamos ahora es decir en cuántas edades fisiológicas se segrega la vida del ser

humano, cómo queda determinada cada una de éstas y cómo se clasificará a cada individuo dentro de éstas.

Uno de los supuestos más utilizado dentro del estudio de poblaciones humanas es que un individuo conserva las mismas capacidades fisiológicas durante la unidad de tiempo en el estudio o edad biológica, que generalmente es anual o quinquenal, esto quiere decir que una persona de edad 30 tiene las mismas capacidades mientras tenga 30 años y al cumplir edad 31 estas se ven afectadas, o una persona que se encuentra en el grupo de edad de 21 a 25 años ve afectadas sus capacidades hasta que cumple 26 años y pertenece al siguiente grupo de edades, en base a este supuesto se construyen las tablas de mortalidad con tasas constantes mientras el individuo pertenece a la misma edad biológica.

Si suponemos lo anterior, podemos decir que el valor esperado del tiempo de estancia en una edad fisiológica está representado por la unidad de tiempo que representa una edad biológica. De esta manera no solo tenemos definido como es un estado del proceso, sino también el parámetro de la distribución de cada estado y el número de estados que tendrá nuestro proceso. Y la estimación de estos será fácil para cada implementación que se quiera hacer.

Por ejemplo si pensamos en una compañía que acaba de iniciar operaciones y el regulador le proporciona una tabla de mortalidad que va de edad 0 a edad 100 y más, tenemos 101 grupos de edades, por lo que el número de estados no absorbentes dentro del proceso será 101, y cuyo valor esperado dentro de cada estado sería uno. Otro ejemplo sería, si la información con la que se cuenta, es de manera quinquenal y se tiene la información a partir del quinquenio de seis a diez años, así hasta el quinquenio de 86 a 90, se tendrían 18 estados no absorbentes 17 pertenecientes a los grupos quinquenales y uno más que represente a las personas de 91 y mas, y cuyo valor esperado dentro de cada estado sería 5

Esta manera de definir los estados también hace que cualquier asegurado pueda ser clasificado de una manera fácil y precisa dentro de algún estado. Ya que al saber la edad biológica que tiene este se puede clasificar dentro del estado que le corresponde.

Con esto obtenemos el espacio de estados en el que estará definida nuestra distribución tipo fase. De aquí en adelante podemos suponer de forma general que el número de estados no absorbentes en el que se definirá la distribución será  $m$ , con  $m \in \mathbb{N}$ .

Lo siguiente que resolveremos es cómo será la matriz de subintensidades. Para ésta lo primero que definiremos es la dimensión que tendrá. Si recordamos dentro de ésta están representados los estados no absorbentes, por lo que si contamos con  $m$  estados no absorbentes, la dimensión de ésta será de  $m$  renglones por  $m$  columnas.

Otra de las características que tiene la matriz de subintensidades es que en la diagonal de la misma, se encuentra el valor negativo del parámetro de la distribución exponencial que modela el tiempo de estancia, de esta forma si suponemos que el valor esperado del tiempo de estancia en el  $i$ -ésimo estado no absorbente es  $1/\lambda_i$  con  $i \in \{1, 2, \dots, m\}$ , la entrada  $(i, i)$  de la matriz tendría el valor  $-\lambda_i$ . Así determinaremos los elementos de la diagonal de la matriz.

Veamos que una vez que se sale del estado  $i$  solo existen dos opciones, pasar al estado absorbente  $d$  o al estado  $i+1$ , y como el estado absorbente no está representado dentro de la matriz de subintensidades, los únicos elementos por determinar a parte de la diagonal son las entradas  $(i, i+1)$  con  $i \in \{1, 2, \dots, m-1\}$  y todas las demás entradas de la matriz de subintensidades estarán en cero.

Para las entradas  $(i, i+1)$  nos ayudaremos de la relación que existe entre la matriz de subintensidades y las probabilidades de salto, que nos dice:

$$p_{ij} = \frac{G_{ij}}{\lambda_i} ;$$

donde  $p_{ij}$  es la probabilidad de pasar al estado  $j$  desde el estado  $i$  una vez que se sale de éste,  $\lambda_i$  es el parámetro de la exponencial asociada al estado  $i$  y  $G_{ij}$  la entrada  $(i, j)$  de la matriz. De esta manera el valor de la entrada  $(i, i+1)$  sería el siguiente:

$$G_{ii+1} = \lambda_i p_{ij} .$$

El único valor desconocido hasta el momento sería  $p_{ij}$ . Ya que este valor representa la probabilidad de pasar del estado  $i$  al estado  $j$  una vez que se sale del estado  $i$ , dicha probabilidad se puede interpretar como la probabilidad de sobrevivir en el estado  $i$ , que es el complemento de la probabilidad de morir en el estado  $i$ , es decir:

$$p_{ij} = (1 - q_i) ;$$

con  $q_i$  la probabilidad de morir en el estado  $i$  o dicho de otra forma, la probabilidad de pasar del estado  $i$  al estado absorbente  $d$  una vez que se sale del estado  $i$ . Así el valor de la entrada  $(i, i+1)$  estaría determinado de la siguiente manera:

$$G_{ii+1} = \lambda_i (1 - q_i) .$$

De esta manera la matriz de subintensidades queda de la siguiente manera:

$$G = \begin{pmatrix} -\lambda_1 & \lambda_1(1 - q_1) & 0 & \dots & 0 & 0 \\ 0 & -\lambda_2 & \lambda_2(1 - q_2) & \dots & 0 & 0 \\ 0 & 0 & -\lambda_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -\lambda_{m-1} & \lambda_{m-1}(1 - q_{m-1}) \\ 0 & 0 & 0 & \dots & 0 & -\lambda_m \end{pmatrix},$$

con esto sólo haría falta definir  $q_i$ , para que la matriz de subintensidades quede bien definida.

Para obtener las  $q_i$  o probabilidades de morir en el estado  $i$ , ya que lo que nos interesa es medir el fenómeno de mortalidad sin importar la causa y como fue definido el espacio de estados, la estimación de éstas no resulta complicada, ya que si se cuenta con tasas de mortalidad para las edades biológicas, entonces éstas representan la probabilidad de morir en las edades fisiológicas asociadas a éstas. Estas tasas de mortalidad podrían ser las proporcionadas por el regulador, las observadas dentro de la compañía o algunas otras que sean de interés para la valuación. De esta manera la tasa de mortalidad asociada al  $i$ -ésimo estado será la tasa de mortalidad asociada a la edad biológica que dio pie a la existencia del  $i$ -ésimo estado.

Lo único que haría falta para que nuestra distribución tipo fase y nuestro modelo queden bien definidos es la distribución inicial  $\pi$ . Esta depende directamente del objeto de estudio para el que será utilizado el modelo, ya que dependiendo la edad fisiológica en la que se encuentre el individuo al momento del estudio la distribución inicial asociada sería uno en la entrada correspondiente y cero en todas las demás.

De esta manera queda concluido el desarrollo de nuestro modelo. A continuación se procederá a ponerlo a prueba mediante dos implementaciones, la primera de ellas a datos reales cuyo detalle no es el más adecuado, ya que la información que se cuenta viene agrupada en quinquenios, así se podrá apreciar cómo funciona nuestro modelo en el caso en que la información con la que se cuente no tenga un buen detalle. Para la segunda implementación se tomará como base la tabla de mortalidad utilizada actualmente por el sector asegurador, y ya que ésta se presenta agrupada de forma anual podremos apreciar cómo funciona el modelo cuando el detalle en los datos es mayor.

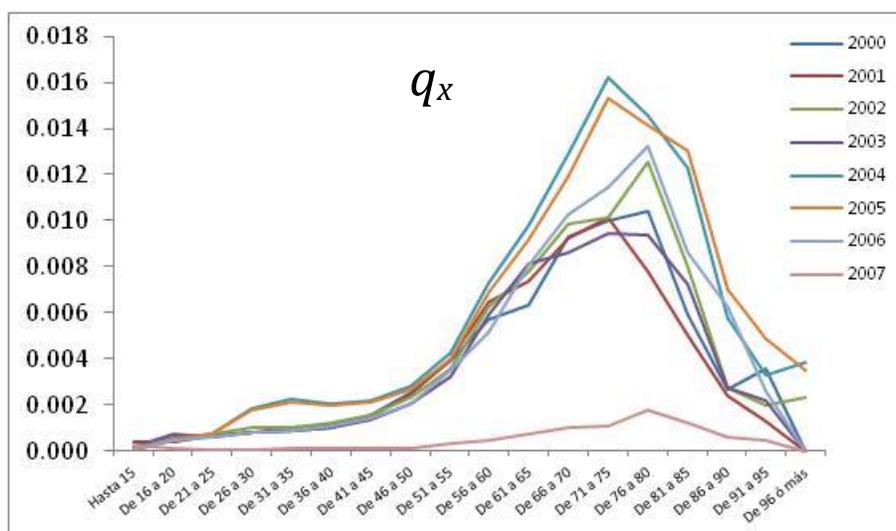
## Implementación a Datos Reales

Para esta implementación se utilizó la información de número de asegurados y defunciones observadas en el sector asegurador mexicano del año 2000 al 2007, estos datos no son

presentados en grupos de edades quinquenales, con excepción del primer y el último grupo; ya que el primer grupo es hasta 15 años y el último es el grupo abierto de 96 años y más.

Años más recientes no fueron considerados ya que la forma de reportar la información cambió y ahora en vez de tener una tabla con defunciones se tiene el dato de siniestros del beneficio básico, que no necesariamente es por defunción<sup>9</sup>.

Con los datos de expuestos y defunciones, se calcularon las  $q_x$ <sup>10</sup> para todos los años, estas probabilidades las tenemos en la siguiente gráfica:



Gráfica 4.2.1 Probabilidad de muerte dado que se está en el grupo  $x$ .

En base a estos datos se decidió excluir la información del año 2007, así como la información del primer grupo que es hasta 15 años y de los grupos mayores a 75 años. Con esto nos quedamos con la información pertenecientes a los 12 grupos de edad restantes y un grupo más que corresponderá a las personas de 76 años o más cuya probabilidad de muerte es definida como uno.

Para estos 13 grupos de edad se calculó la  $q_x$  promedio por grupo, que van a ser las que utilizaremos para estimar los parámetros de la distribución tipo fase.

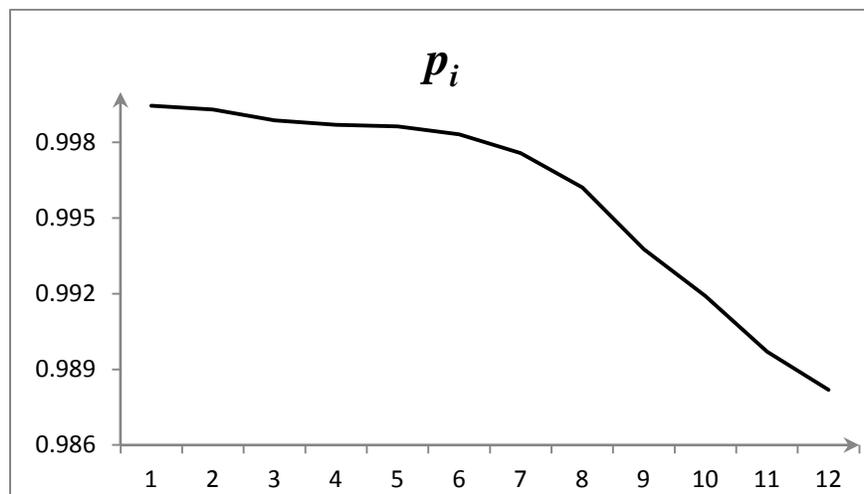
<sup>9</sup> Esta información puede ser descargada de la página de la Comisión Nacional de Seguros y Fianzas en la siguiente liga: <http://www.cnsf.gob.mx/Difusion/Paginas/AAnuarios.aspx>

<sup>10</sup> Dentro de esta implementación cuando se haga referencia a la probabilidad de muerte en algún grupo de edad se denotara como  $q_x$ , con  $x \in \{\text{hasta 15, de 16 a 20, de 21 a 25, ..., de 96 o más}\}$ , y cuando se haga referencia a la probabilidad de muerte en algún estado del proceso de saltos de Markov se denotará como  $q_i$ , con  $i \in \{1, 2, \dots, 12, 13\}$ .

De esta manera el espacio de estados queda definido por 14 estados, 13 de estos no absorbentes que pertenecen a los grupos mencionados y uno más que es el estado absorbente  $d$ , es decir:

$$E = \{1, 2, 3, \dots, 12, 13, d\}.$$

En base a las  $q_x$  promedio, se obtiene la probabilidad de muerte para los estados no absorbentes denotada por  $q_i$  y con las cuales se calculo la  $p_i$ . Con dichas probabilidades quedan definidas las probabilidades de salto, ya que si nos encontramos en un estado  $i \in \{E \setminus d\}$  y se efectúa un salto, solo se puede pasar al estado  $i+1$  con probabilidad  $p_i$  o al estado  $d$  con probabilidad  $q_i$ . En la siguiente gráfica tenemos las probabilidades de sobrevivir en el estado  $i$ .



Gráfica 4.2.2 Probabilidades de sobrevivir en el estado  $i$ <sup>11</sup>.

Por otro lado, como los grupos con los que contamos son quinquenales, asumiremos que el valor esperado del tiempo de estancia en cada uno de los estados no absorbentes es cinco, con esto los parámetros  $\lambda_i$  de las distribuciones exponencial es 0.2, para toda  $i \in \{E \setminus d\}$ .

La construcción de la matriz de subintensidades  $G$  se hará a partir de los valores de  $\lambda_i$  y la probabilidad de sobrevivir en el estado  $i$ . La determinación de los elementos de la matriz  $G$  será de la siguiente manera. La dimensión de la matriz  $G$  es de 13 renglones y 13 columnas, donde cada renglón y columna representa un estado no absorbente, la diagonal estará determinada por los parámetros  $\lambda_i$  con signo negativo, que en nuestro caso es -0.2.

Dado que si nos encontramos en un estado  $i$  solo podemos pasar al estado  $i+1$  o al estado  $d$  y como el estado  $d$  no está representado dentro de la matriz  $G$ , aparte de la diagonal sólo

<sup>11</sup> En esta gráfica se omite el valor de la probabilidad de sobrevivir en el estado 13 ya que es cero, y por cuestiones de escala no se podría apreciar el comportamiento de las  $p_i$ .

hay que determinar las entradas  $(i, i + 1)$ . Estas entradas serán determinadas por los valores  $p_i$  multiplicado por  $\lambda_i$ , esto ya que los renglones de la matriz de intensidades deben sumar cero y que a partir de esta se puede construir la matriz de probabilidades de saltos, y dichas probabilidades corresponde a las  $p_i$  y  $q_i$ . Todas las demás entradas de la matriz  $G$  valdrán cero. Es decir la matriz  $G$  queda de la siguiente manera:

$$G = \begin{pmatrix} -0.2 & 0.199891 & 0 & \dots & 0 & 0 \\ 0 & -0.2 & 0.199860 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -0.2 & 0.197639 \\ 0 & 0 & 0 & \dots & 0 & -0.2 \end{pmatrix}.$$

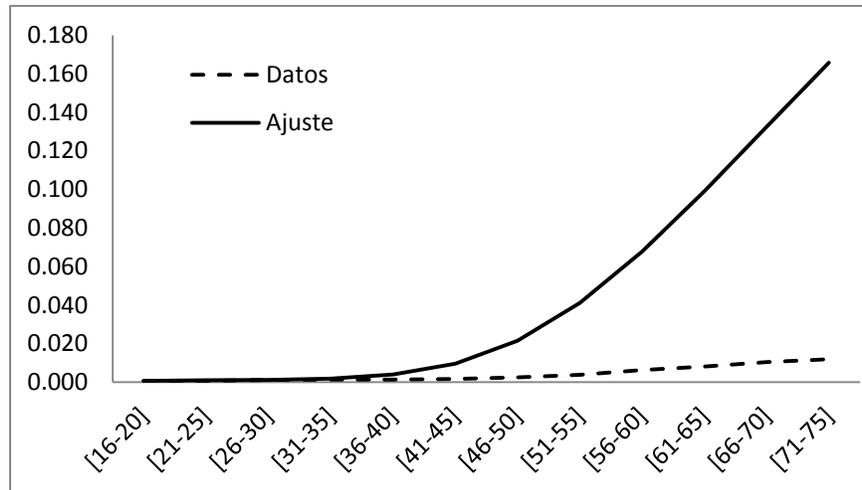
Con esta matriz se puede construir el vector fila  $g$ , que coincide con las probabilidades de muerte multiplicadas por  $\lambda_i$ .

Para tener completamente definida la distribución tipo fase nos hace falta el vector  $\pi$  que representa la distribución inicial. Para este caso se realizó el ejercicio suponiendo que se iniciaba en el estado uno.

Ya con nuestra distribución tipo fase definida, obtenemos que el valor esperado es 64.14, que representaría el tiempo esperado de vida de una persona que se encuentra dentro del estado 1, que no difiere significativamente con el valor de la esperanza de vida que se obtiene a partir de los datos reales que es de 61.64. También se obtuvo el segundo momento de la distribución y con ello la varianza de ésta, que es de 343.71, que equivale a una desviación estándar de 18.54. También se realizó el siguiente ejercicio.

**Ejercicio 4.2.1:** Se simularon 250 tablas de mortalidad aleatorias, cada una suponiendo un radix de 100,000, por lo que se simularon 25,000,000 números aleatorios con distribución tipo fase (100,000 para cada tabla), estos números fueron simulados con el código del apéndice A.2. Con los números simulados se procedió a construir las 250 tablas de mortalidad.

Ya con las 250 tablas, se sacó la  $q_x$  promedio para cada grupo de edad obteniendo el siguiente resultado:



Gráfica 4.2.3 Ajuste a datos empíricos.

Como se puede apreciar en la gráfica, el ajuste no es bueno para casi todos los grupos de edad, esto se puede deber a diversos factores, pero principalmente se le atribuye al detalle en los datos, ya que al no tener un buen detalle en la información los supuestos utilizados son más graves, por ejemplo, al crear grupos más grandes de edad nos implica directamente un mayor error al momento de suponer las edades fisiológicas. Por lo que se puede concluir que el ajuste a este conjunto de datos no fue bueno.

## Implementación a Datos Teóricos

En esta implementación se tomaron los datos de la tabla de mortalidad publicada por la Comisión Nacional de Seguros y Fianzas para las valuaciones de seguros de vida individual, llamada CNSF 2000-I<sup>12</sup>.

$x$	$q_x$								
12	0.000396	30	0.001508	48	0.005725	66	0.021490	84	0.077337
13	0.000427	31	0.001624	49	0.006164	67	0.023111	85	0.082817
14	0.000460	32	0.001749	50	0.006637	68	0.024851	86	0.088649
15	0.000495	33	0.001884	51	0.007145	69	0.026720	87	0.094850
16	0.000533	34	0.002029	52	0.007693	70	0.028724	88	0.101436
17	0.000575	35	0.002186	53	0.008282	71	0.030874	89	0.108424
18	0.000619	36	0.002354	54	0.008915	72	0.033180	90	0.115832
19	0.000667	37	0.002535	55	0.009597	73	0.035651	91	0.123677
20	0.000718	38	0.002730	56	0.010330	74	0.038300	92	0.131973
21	0.000773	39	0.002940	57	0.011119	75	0.041136	93	0.140737

<sup>12</sup> esta tabla puede ser descargada desde el portal de la comisión en la siguiente liga:  
[http://www.cnsf.gob.mx/Normativa/InstReglas/ReglasCNSF/149%203.6.-%20ACUERDO\\_1999-12-31.pdf](http://www.cnsf.gob.mx/Normativa/InstReglas/ReglasCNSF/149%203.6.-%20ACUERDO_1999-12-31.pdf)

22	0.000833	40	0.003166	58	0.011967	76	0.044174	94	0.149983
23	0.000897	41	0.003410	59	0.012879	77	0.047424	95	0.159723
24	0.000966	42	0.003672	60	0.013860	78	0.050902	96	0.169970
25	0.001041	43	0.003954	61	0.014914	79	0.054619	97	0.180733
26	0.001121	44	0.004258	62	0.016048	80	0.058592	98	0.192020
27	0.001207	45	0.004585	63	0.017265	81	0.062834	99	0.203837
28	0.001300	46	0.004938	64	0.018574	82	0.067362	100 <sup>+</sup>	1.000000
29	0.001400	47	0.005317	65	0.019980	83	0.072190		

En dicha tabla contamos con la probabilidad de morir dado que se tiene cierta edad, es decir, las  $q_x$ <sup>13</sup>.

Para determinar la distribución tipo fase, que modela la mortalidad, es necesario determinar el espacio de estados y sus parámetros. Esto es, hay que estimar la matriz de subintensidades  $G$  y la distribución inicial  $\pi$ .

Empezaremos determinando el espacio de estados, como la información con la que se cuenta va de edad 12 a edad 100<sup>+</sup> (100 años y más) en grupos anuales, asumiremos que cada grupo es un estado, con ello contaremos con 89 estados no absorbentes y el estado absorbente denotado por  $d$ , que representa la muerte, es decir:

$$E = \{1, 2, 3, \dots, 88, 89, d\}.$$

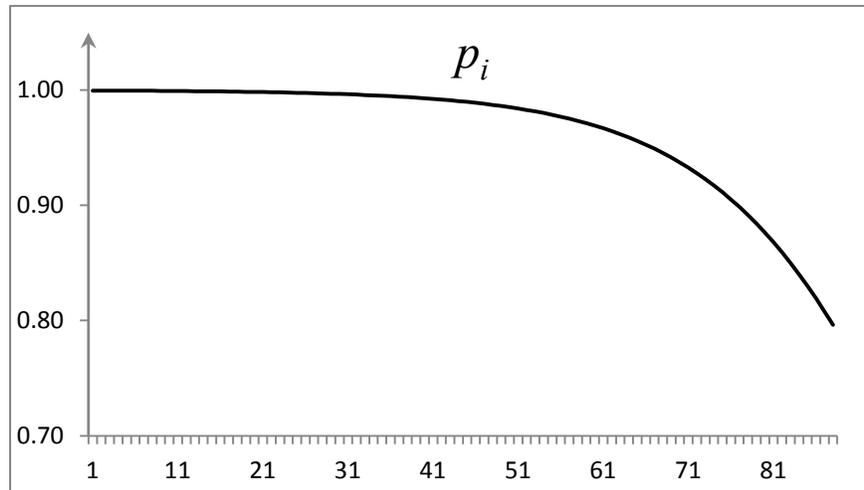
Con esto, para los estados no absorbentes las probabilidades de salto quedan definidas a través de la tabla de mortalidad, ya que si nos encontramos en el estado  $i$  y se efectúa un salto, la probabilidad de pasar al estado absorbente será  $q_i$ , donde:

$$q_i = q_{x-11}$$

y la probabilidad de pasar al siguiente estado será el complemento, es decir  $p_i = 1 - q_i$ . En la siguiente gráfica tenemos las probabilidades de sobrevivencia para cada estado no absorbente.

---

<sup>13</sup> Dentro de esta implementación cuando se haga referencia a la probabilidad de muerte en alguna edad biológica se denotará como  $q_x$ , con  $x \in \{12, 13, \dots, 99, 100^+\}$ , y cuando se haga referencia a la probabilidad de muerte en algún estado del proceso de saltos de Markov se denotará como  $q_i$ , con  $i \in \{1, 2, \dots, 88, 89\}$ .



Gráfica 4.2.4 Probabilidades de sobrevivencia<sup>14</sup>.

Por otro lado, como los grupos con los que contamos son anuales, asumiremos que el valor esperado del tiempo de estancia en cada uno de los estados no absorbentes es uno, con esto los parámetros  $\lambda_i$  de las distribuciones exponencial es uno, para toda  $i \in \{E \setminus d\}$ .

La construcción de la matriz de subintensidades  $G$  se hará a partir de los valores de  $\lambda_i$  y la probabilidad de sobrevivir en el estado  $i$ . Los elementos de la matriz  $G$  serán determinados de la siguiente manera. La dimensión de la matriz  $G$  es de 89 renglones y 89 columnas, donde cada renglón y columna se asocia a un estado no absorbente, los elementos de la diagonal serán los parámetros  $\lambda_i$  con signo negativo, en nuestro caso, estos términos serán -1.

Dado que si nos encontramos en un estado  $i$ , sólo podemos pasar al estado  $i+1$  o al estado  $d$  y como el estado  $d$  no está representado dentro de la matriz  $G$ , únicamente será necesario determinar los valores de las entradas  $(i, i+1)$ . Éstas, quedaran determinadas con los valores  $p_i$ . Para las demás entradas de la matriz  $G$ , el valor será cero. Así la matriz  $G$  queda de la siguiente manera:

$$G = \begin{pmatrix} -1 & 0.999604 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0.999573 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 0.796163 \\ 0 & 0 & 0 & \cdots & 0 & -1 \end{pmatrix}.$$

Con esta matriz se puede construir el vector fila  $g$ , que coincide con las probabilidades de muerte.

<sup>14</sup> En esta gráfica se omite el valor de la probabilidad de sobrevivir en el estado 89 ya que es cero, y por cuestiones de escala no se podría apreciar el comportamiento de las  $p_i$ .

Para tener completamente definida la distribución tipo fase nos hace falta el vector  $\pi$  que representa la distribución inicial. Pero este vector  $\pi$  depende directamente del objeto de estudio, ya que por ejemplo, si se desea obtener el tiempo esperado de vida de un individuo que se encuentra situado en el estado  $i$ , el vector  $\pi$  sería uno en la  $i$ -ésima entrada y cero en todas las demás entradas. O si lo que queremos es construir una tabla de mortalidad aleatoria, se supone que el proceso empieza en el estado uno, por lo tanto el vector  $\pi$  sería uno en la primer entrada y cero en todas las demás.

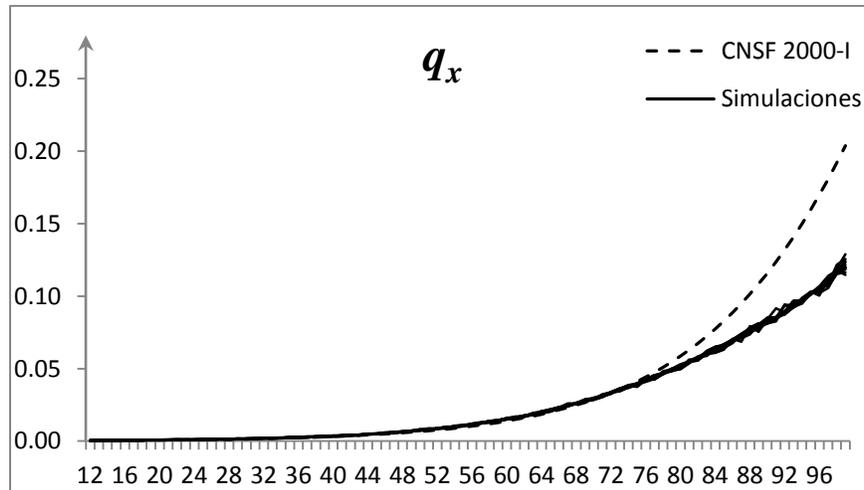
En esta implementación se supuso que se iniciaba en el estado uno, es decir:

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

con ello la distribución tipo fase que modela el tiempo restante de vida para un individuo de edad biológica 12, queda definida con los parámetros  $\pi$  y  $G$ .

Ya con la distribución tipo fase bien definida obtenemos que el valor esperado es de 64.18 años, y representaría el tiempo esperado de vida de una persona con edad 12, que no difiere de manera significativa con el valor de la esperanza de vida que se obtiene a partir de la tabla de mortalidad CNSF 2000-I que es 63.68 años. También se obtuvo el segundo momento de la distribución y con ello la varianza de ésta, que es de 332.64, que equivale a una desviación estándar de 18.24.

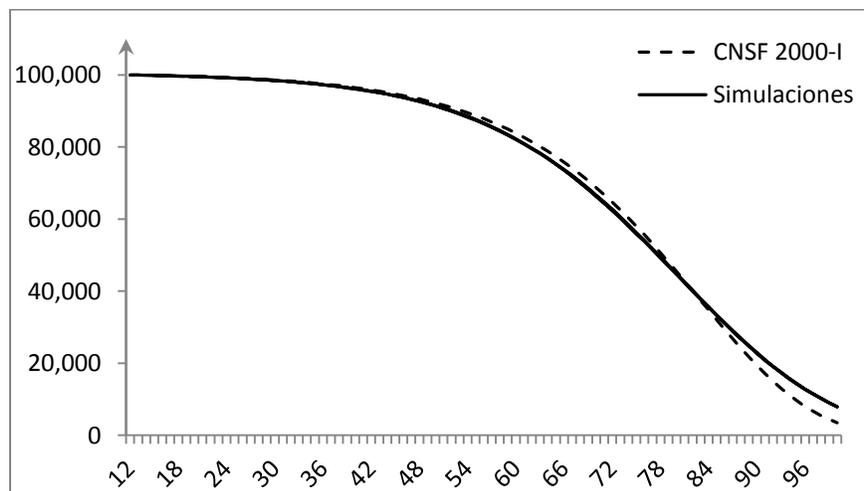
**Ejercicio 4.2.2:** Se simularon diez tablas de mortalidad aleatorias, cada una suponiendo un radix de 100,000, por lo que se simularon 1,000,000 de números aleatorios con distribución tipo fase (100,000 para cada tabla), estos números fueron simulados con el código del apéndice [A.2](#). Con los números simulados se procedió a construir las 10 tablas de mortalidad simuladas, los resultados obteniendo se muestran en la siguiente gráfica:



Gráfica 4.2.5 Tablas de mortalidad<sup>15</sup>.

En la gráfica podemos observar que el ajuste en edades grandes no es bueno, la razón de esto puede ser el supuesto en los parámetros  $\lambda_i$ , ya que para edades grandes la llegada al estado absorbente se está efectuando en un tiempo relativamente grande. Pero para edades pequeñas el ajuste es bueno.

Con la siguiente gráfica podemos ver el comportamiento de la población tanto teórica como la que equivale a cada una de las 10 simulaciones, tenemos que el número de sobrevivientes para edades grandes en las simulaciones queda por arriba del número de sobrevivientes teóricos obtenidos con la tabla de mortalidad, lo que implica que el número de defunciones es menor.



Gráfica 4.2.6 Sobrevivientes a edad  $x$ .

<sup>15</sup> En esta gráfica se omite el valor de la probabilidad de muerte a la edad 100 y más ya que es uno, y por cuestiones de escala no se podría apreciar el comportamiento de las  $q_x$ .

Dentro de esta implementación, al contar con un mayor detalle en los datos y con ello contar con grupos de edades más homogéneos, el ajuste en la mayoría de las edades resulta ser adecuado ya que las curvas simuladas se apegan al comportamiento de la curva teórica, por lo que podríamos decir que la implementación en estos datos es aceptable.

# Conclusiones

Dentro de este trabajo revisó cómo se ha estudiado la mortalidad a través del tiempo y la importancia que ésta tiene dentro de la población humana. Para poder proponer una solución para la medición del riesgo por muerte, se estudiaron los procesos de saltos de Markov así como sus principales propiedades, ya que estos procesos, son la base para definir lo que es una distribución tipo fase, cuya teoría sustenta el modelo de mortalidad propuesto. Antes de desarrollar tal modelo, se describió el modelo propuesto por X. Sheldon Lin y Xiaoming Liu, ya que éste constituye un antecedente muy importante en lo que es la utilización de distribuciones tipo fase para el estudio de la mortalidad.

Una vez que se tenía contextualizado el problema y analizadas las posibles herramientas a utilizar, se desarrolló un modelo de mortalidad basado en una distribución tipo fase. En este modelo se buscó mantener bastante simplicidad para que la información necesaria así como el esfuerzo para estimar sus parámetros fueran los mínimos necesarios, esto pensando en las instituciones que no cuentan con los recursos suficientes para la utilización de un modelo más robusto. Esto se logró tomando como base algunos supuestos y conceptos esenciales, como lo es el concepto de edad fisiológica, ya que sin éste resulta inconcebible el poder utilizar una distribución tipo fase. Otro de los supuestos más importantes fue el de la relación que existe entre la edad fisiológica y la edad biológica, ya que con éste se simplifica bastante la estimación de los estados del proceso de saltos de Markov, así como la estimación de los parámetros que modelan el tiempo de estancia en éstos.

Ya con los supuestos necesarios y las tasas de mortalidad para cada estado, se describió cómo sería la estimación de los dos parámetros que caracterizan a nuestra distribución, la matriz de subintensidades y la distribución inicial, con lo que queda concluido el desarrollo del modelo.

Con el modelo desarrollado se procedió a realizar dos implementaciones, una donde la calidad y suficiencia de información no cuenta con un buen detalle, y la segunda donde el detalle que se tiene en los datos es mayor. De los resultados obtenidos en estas implementaciones se puede concluir lo siguiente:

Una primera conclusión es que, como en la mayoría de los modelos, el que proponemos basa sus resultados en la calidad de los datos con los que se estiman sus parámetros, ya que dentro de las dos implementaciones realizadas los resultados obtenidos son bastantes diferentes. En la [Implementación a Datos Reales](#) donde se cuenta con datos agrupados de manera quinquenal, los resultados que se obtuvieron son inadecuados, ya que la curva obtenida por medio del modelo dista de los datos observados, esto atribuido a que el

supuesto utilizado para la estimación de los parámetros  $\lambda$ , tiene implícito un error proporcional al grupo de edad utilizado, con lo que el resultado obtenido en esta implementación dista bastante de los datos observados.

Ya en la [Implementación a Datos Teóricos](#) donde se cuenta con un mayor detalle en los datos, ya que los grupos que tenemos son anuales, los resultados obtenidos son satisfactorios, esto atribuido a que para este caso los supuestos propuestos se adaptan de mejor manera a los datos con los que se cuenta. Además, al tener grupos de edades más pequeños, el error que existe entre el modelo y los datos reales es menor.

De manera general los resultados obtenidos son satisfactorios ya que como pudimos apreciar en la [Implementación a Datos Teóricos](#), en la que el detalle de los datos es mejor, para las edades donde se encuentra el grueso de la población el modelo describe bastante bien la mortalidad de la población. Donde el ajuste parece no ser adecuado en edades grandes, pero hasta el momento la mayoría de los modelos con los que se cuentan tienen deficiencias en los grupos de edades mayores, esto atribuido a que la mortalidad en estas edades resulta ser bastante errática y los datos con los que se cuenta no son suficientes para poder estimarla de manera precisa.

Ya que pudimos apreciar que el modelo es especialmente sensible a los parámetros  $\lambda$ , un trabajo futuro sería robustecer la estimación de éstas, de manera que se pueda incluir de manera más precisa el tiempo en el que ocurre la muerte del individuo.

Finalmente hay que tomar en cuenta que para poder contar con un modelo más robusto, se tiene un costo directo de que la información necesaria para su implementación, ya que se necesita un mayor detalle y suficiencia. Por lo que se debe de tener cuidado en que el modelo siga siendo aplicable y tratar de no terminar en un modelo como el de X. Sheldon Lin y Xiaoming Liu, cuya implementación resulta difícil.

# Apéndices

## A.1 Proceso de Nacimiento y Muerte

El [Algoritmo 2.3.1](#) para simular trayectorias de un proceso de nacimiento y muerte fue desarrollado en el software R, en particular para el ejemplo de la [Sección 2.4](#) fue desarrollado el siguiente código:

```
m=50          # No. máximo de saltos a simular
n=6           # No.de estados (tomando en cuenta el estado cero)

X<-vector(mode="integer",lengt=m+1) # Vector donde se guarda el estado en el que se está en cada
salto
T<-vector(mode="double",lengt=m)    # Vector donde se guardan los tiempos de estancia
W<-vector(mode="double",lengt=m)    # Vector donde se guardan los tiempos de arribo
I<-vector(mode="double",lengt=n)    # Vector de distribución inicial
G=matrix(0,n,n)                    # Generador infinitesimal
P=matrix(0,n,n)                    # Probabilidades de Salto

#Se llena el vector de distribución inicial
I[1]=0
I[2]=0.075
I[3]=0.25
I[4]=0.35
I[5]=0.25
I[6]=0.075

#Se llena la matriz generadora
G[1,2]=.9
G[2,1]=1.2
G[2,3]=.3
G[3,2]=.78
G[3,4]=.52
G[4,3]=.55
G[4,5]=.55
```

G[5,4]=.84

G[5,6]=.36

G[6,5]=.1.5

G[1,1]=-G[1,2]

G[n,n]=-G[n,n-1]

for(i in 2:(n-1))

{

    G[i,i]=-(G[i,i-1]+G[i,i+1])

}

#Se llena la matriz de probabilidades de salto a partir del generador infinitesimal

P[1,2]=G[1,2]/G[1,2]

P[n,n-1]=G[n,n-1]/G[n,n-1]

for(i in 2:(n-1))

{

    P[i,i-1]=G[i,i-1]/(G[i,i-1]+G[i,i+1])

    P[i,i+1]=G[i,i+1]/(G[i,i-1]+G[i,i+1])

}

#Se inicializa el proceso

X[1]=sample(n,1,replace=FALSE,I)-1

for(i in 1:m)

{

#Se checa la condición de paro

    if (X[i]==0)

    {

        i=m

    }

    else

    {

        # Se simula el tiempo de estancia dependiendo del estado en el que se encuentre

        T[i]=rexp(1,-G[X[i]+1,X[i]+1])

```

# Se acumulan los tiempos de arribo
if(i==1)
{
  W[i]=T[i]
}
else
{
  W[i]=W[i-1]+T[i]
}

# Se simula el estado al que se va a saltar a través de una bernoulli, tomando como éxito la
probabilidad de muerte
aux=rbinom(1,1,P[X[i]+1,X[i]])
if(aux==1)
{
  X[i+1]=X[i]-1
}
else
{
  X[i+1]=X[i]+1
}
}
}

```

## A.2 Distribuciones Tipo Fase

Para simular números aleatorios con una distribución tipo fase fue utilizada la función `rphtype(n,  $\pi$ , G)` en la librería “*actuar*” del software R, dentro de esta función los argumentos que solicita son el tamaño de la muestra que se quiere generar, la distribución inicial  $\pi$  y la matriz de subintensidades  $G$ . Esta función fue utilizada en el [Ejercicio 4.2.1](#) y el [Ejercicio 4.2.2](#). El código del [Ejercicio 4.2.1](#) es el siguiente, la única diferencia con el otro ejercicio es la distribución inicial y la matriz de subintensidades.

```
library(actuar)
```

```

n=13    # numero de estados no absorbentes

I<-vector(mode="double",length=n)    # distribución inicial
G<-matrix(0,n,n)    # matriz de subintensidades

#Se llena el vector de distribución inicial
I[1]=1

for(i in 2:n)
{
  I[i]=0
}

#Se llena la matriz de subintensidades
for(i in 1:n)
{
  G[i,i]=-0.2
}

G[1,2]=0.199890749
G[2,3]=0.199860141
G[3,4]=0.199775636
G[4,5]=0.199739664
G[5,6]=0.199728073
G[6,7]=0.199663184
G[7,8]=0.199515407
G[8,9]=0.199242776
G[9,10]=0.198752187
G[10,11]=0.19838105
G[11,12]=0.197940027
G[12,13]=0.197638608

# Genera los tiempos aleatorios
x <- rphtype(25000000, I, G)

```

# Bibliografía

[ 1 ] Alatorre Salgado Rosa María, *Tabla de Mortalidad Experiencia Mexicana 1982-1989*, Serie Documentos de Trabajo CNSF, Documento de trabajo No. 2, Septiembre 1991.

[ 2 ] Alen O., *On Phase Type Distributions in Survival Analysis*, Scandinavian Journal of Statistics 22: 447–63, 1995.

[ 3 ] Aranda M. Oscar, *Hipótesis relativa a la tasa instantánea de mortalidad (1era parte)*, UNAM F. Ciencias, Junio, 2011.

[ 4 ] Asmussen Soren, Nerman Olle and Olsson Marita, *Fitting Phase-Type Distributions via the EM Algorithm*, Scandinavian Journal of Statistics, Vol. 23, No. 4 (Dec., 1996), pp. 419-441.

[ 5 ] Bladt M., *A Review on Phase – Type Distributions and Their Use in Risk Theory*, Astin Bulletin, Vol. 35, No. 1, pp. 145 – 161.

[ 6 ] Bladt M., *Teoría del Riesgo I*. Enero 2004.

[ 7 ] Baltazar Larios, F., *Estimation of discretely observed Markov Processes*, Tesis para obtener grado de Doctor en Ciencias. UNAM, México D.F 2011.

[ 8 ] Gavrilov L. A. and S. Gavrilova., *The Biology of Life Span: A Quantitative Approach*. Chur, Switzerland: Harwood Academic Publishers, 1991.

[ 9 ] Karlin Samuel and Taylor Howard M., *A First Course in Stochastic Processes*, Academic Press. New York 1975.

[ 10 ] Karlin Samuel and Taylor Howard M., *A Second Course in Stochastic Processes*, Academic Press. New York 1981.

[ 11 ] Martín Braun, *Ecuaciones Diferenciales y sus Aplicaciones*, Grupo Editorial Iberoamérica, 1990.

[ 12 ] Mendoza Ramírez Manuel, Madrigal Gómez Ana María, Martínez Torres Evangelina, *Tablas de Mortalidad CNSF 2000-I y CNSF 2000-G*, Serie Documentos de Trabajo CNSF, Documento de trabajo No. 80, Mayo 2000.

[ 13 ] Newton L. Bowers JR., Hans U. Gerber, James C. Hickman, Donald A. Jones, Cecil J. Nesbitt, *Actuarial Mathematics*, The Society of Actuaries 1997.

[ 14 ] Peralta Gutiérrez Oscar, *Modelos de reserva con reclamaciones tipo fase y matriz exponencial*, Tesis para obtener grado de Actuario. UNAM, México D.F 2013.

[ 15 ] Rincón Luis, *Curso Intermedio de Probabilidad*, Departamento de Matemáticas, Facultad de Ciencias UNAM, Octubre 2007.

[ 16 ] Rincón Luis, *Introducción a los Procesos Estocásticos*, Las Prensas de Ciencias, Facultad de Ciencias UNAM, Enero 2012.

[ 17 ] Rubinstein, R. Y. and Kroese, D. P. *Simulation and Monte Carlo Method*. Second edition. Jhon Wiley & Sons, Inc. 2008.

[ 18 ] X. Sheldon Lin and Xiaoming Liu. *Markov Aging Process and Phase-Type Law of Mortality*, North American Actuarial Journal, Vol. 11, No.4 (Abril 2008), pp. 92 – 109.