



UNIVERSIDAD NACIONAL AUTÓNOMA DE  
MÉXICO

---

---

FACULTAD DE CIENCIAS

"LOS PROCESOS DE BERNOULLI Y POISSON EN EL  
ESTUDIO DE CASOS DE INCUMPLIMIENTO DE LA NORMA  
PARA OZONO"

TESIS

QUE PARA OBTENER EL TÍTULO DE ACTUARIA

PRESENTA:

ROMUALDA HERMELINDA PÉREZ MUÑOZ

DIRECTOR DE TESIS: Dra. ELIANE REGINA RODRÍGUEZ CALONI

2006





Universidad Nacional  
Autónoma de México



**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.

<<Par devoir vis à vis de quelqu'un qui en très peu de  
temps m'appris bien des choses, et une tout  
particulièrement, le goût du bonheur.>> Lévy

## Agradecimientos

Gracias a mis padres porque han sido el pilar de este proyecto de vida desde su inicio.

Este trabajo fue posible gracias a ti, Eliane, porque compartiste conmigo tus valiosos conocimientos y tu excelente forma de trabajar, de la cual tuve la oportunidad de aprender durante el desarrollo del mismo.

Mi reconocimiento especial a la Facultad de Ciencias por la formación que me brindó en sus aulas, a través de profesores comprometidos con su labor docente. De manera particular al profesor Ferrer por el apoyo que siempre recibí de él en sus cursos de cálculo; así como al profesor Minzoni, por haberme infundido otras perspectivas de desarrollo. Pero sobre todo a la UNAM porque nos permite crecer en todos los ámbitos en los que es posible hacerlo.

Asimismo, agradezco al instituto de Matemáticas por brindarme un espacio de trabajo mediante su programa de Becas de Lugar y por el gran apoyo recibido por parte del personal de cómputo, Carlos, Rubén y Federico; del personal de Biblioteca; así como de Leo del departamento de publicaciones.

Mi agradecimiento especial al Dr. Luis Gorostiza por el apoyo que me brindó por medio del SNI y a PROBETEL por haber hecho posible la realización de este proyecto de investigación.

A mis sinodales, Luis, Juan, Angel y Pilar les agradezco por dedicar tiempo a la revisión de mi tesis y por contribuir con su experiencia al enriquecerla con sus comentarios.

Je voudrais aussi remercier mes professeurs au Cele, M. Diaz Barriga par la grande motivation qui nous a donné dans ses bons cours de Français, au même temps que à Rebeca Navarro, Danielle Martineau, Leticia Trujillo et Caroline Caset par des bonnes expériences.

Ich möchte an Heide danken, weil meine beste lehrerin auf Deutsch war. Gleichfalls danke ich an Rocio Madrid von der gute erfahrungen mit ihr, und Lille für die KonversationKurs.

Le hasard, c'est la forme que prend Dieu pour passer incognito.  
Jean Cocteau

# Índice general

<b>Introducción</b>	<b>vii</b>
<b>1. Preliminares para el índice por ozono</b>	<b>1</b>
1.1. Sistema de monitoreo . . . . .	1
1.2. Normas de calidad del aire . . . . .	4
1.3. Cálculo de equivalencias . . . . .	6
1.4. Consecuencias del contaminante ozono . . . . .	7
<b>2. Cadenas de Markov en tiempo continuo</b>	<b>13</b>
2.1. Resultados preliminares . . . . .	13
2.2. Ecuaciones diferenciales de Kolmogorov . . . . .	18
2.3. Probabilidades límite . . . . .	23
2.4. Procesos de nacimiento y muerte . . . . .	27
2.4.1. Probabilidades límite . . . . .	28
2.4.2. Ecuaciones diferenciales de Kolmogorov . . . . .	31
<b>3. Proceso Poisson</b>	<b>35</b>
3.1. Propiedades del Proceso Poisson . . . . .	35
3.2. Distribución de tiempo entre llegadas y de espera . . . . .	43
3.3. Distribución condicional de los tiempos de llegada . . . . .	45
3.4. Proceso Poisson no homogéneo . . . . .	46
<b>4. Proceso Bernoulli y Poisson en el estudio de problemas ambientales</b>	<b>53</b>
4.1. Condiciones para el proceso Bernoulli . . . . .	54
4.1.1. Aplicación a modelos ambientales . . . . .	58
4.1.2. Distribución de probabilidad del número de excedentes . . . . .	59
4.2. Proceso Poisson . . . . .	63

---

4.2.1. Empleo en problemas ambientales . . . . .	64
4.2.2. Distribución de probabilidad para el número de excedentes . . . . .	64
<b>5. Distribución del número de excedentes para ozono</b>	<b>67</b>
5.1. Algunos resultados de la Inferencia Bayesiana . . . . .	67
5.2. Distribuciones del parámetro $\lambda$ . . . . .	70
<b>Conclusiones</b>	<b>79</b>
<b>A. Probabilidad condicional</b>	<b>81</b>
<b>B. Distribución Poisson</b>	<b>83</b>
<b>C. Distribución Binomial</b>	<b>85</b>
<b>D. Distribución Gama</b>	<b>87</b>
<b>E. Funciones <math>o(x)</math></b>	<b>93</b>
<b>F. Gráficas del parámetro <math>\lambda</math></b>	<b>95</b>
<b>Bibliografía</b>	<b>101</b>

# Introducción

Actualmente uno de los principales problemas de carácter internacional lo constituye la contaminación atmosférica. Por lo que se han implementado normas internacionales para los distintos contaminantes, no obstante, lo más importante no lo constituye la implementación de dichas normas, sino el correcto cumplimiento de las mismas. Así, en el presente trabajo se utilizan modelos estocásticos para poder inferir acerca del comportamiento del incumplimiento de la norma de un contaminante en particular, a saber, el ozono.

Dicha aplicación se realizará a través de una modificación del método que utiliza los procesos Bernoulli y Poisson propuestos por Javits (1980).

Por lo que el presente trabajo, está dividido de la siguiente manera:

En el Capítulo 1 se describen las condiciones y métodos de análisis con que cuenta la Zona Metropolitana del Valle de México para hacer frente a los problemas de control ambiental, en particular con el contaminante ozono.

En el Capítulo 2 se enuncian los resultados más sobresalientes que caracterizan a las cadenas de Markov a tiempo continuo.

En el Capítulo 3 se desarrolla un caso particular de las cadenas de Markov: el proceso Poisson.

En el Capítulo 4 se muestra la aplicación de los modelos Binomial y Poisson en problemas de control ambiental, específicamente en la interpretación de la norma internacional para ozono.

En el Capítulo 5 se especifican algunos resultados de la inferencia Bayesiana, así como la propuesta de utilizar ésta teoría para inferir el valor del parámetro  $\lambda$  del modelo Poisson que registra el número de días que la concentración de ozono rebasa los 100, 150, 200 y 240 puntos IMECA para cada año de 1997 hasta 2003.



Finalmente, se presentan las conclusiones y en los apéndices algunos resultados complementarios además de las gráficas de las distribuciones a priori y a posteriori para los diferentes casos considerados.

# Capítulo 1

## Preliminares para el índice por ozono

Uno de los principales problemas de contaminación atmosférica en el Valle de México es generado por las altas concentraciones de ozono. Esto ocasiona que la población de la ciudad se vea expuesta con frecuencia y, por periodos de una o más horas a concentraciones de ozono superiores a la norma ambiental actual. En el trabajo se analizará el comportamiento de dicho contaminante (en términos del incumplimiento de la norma o no) a través de modelos probabilísticos como lo son: los procesos Bernoulli y Poisson.

El presente capítulo está dividido por secciones: en la primera sección se explica cómo funciona el sistema de monitoreo para el control de la calidad del aire en la Zona Metropolitana del Valle de México. Mientras que en la Sección 1.2 se define lo que son las normas de calidad del aire y en la sección subsecuente se muestra cómo se lleva a cabo la equivalencia entre unidades del índice metropolitano de la calidad del aire y las concentraciones máximas diarias y en la última sección se dan a conocer los daños que dicho contaminante ocasiona a la salud humana. Lo anterior se puede revisar con más detalle en: DDF et. al. (1996), INE et. al. (1997) e INE et. al. (2000).

### 1.1. Sistema de monitoreo

El sistema de monitoreo en el Distrito Federal cuenta con dos sistemas de regulación: uno manual en el cual se cuenta con 19 estaciones para el muestreo de partículas suspendidas totales y la Red Automática de Monitoreo

Ambiental (RAMA), ahora conocido como Sistema Integrado de Monitoreo Ambiental (SIMA) que cuenta con 32 estaciones de monitoreo atmosférico entre ellas 10 estaciones micrometeorológicas, una torre meteorológica, un radar acústico y una ecosonda. Es relevante señalar que 21 de éstas se encuentran en el Distrito Federal y 11 en el estado de México. Cada sitio de monitoreo cuenta con el equipamiento necesario para efectuar las mediciones de los contaminantes derivados de la actividad y uso del suelo más representativo de la región. Esta Red opera continuamente durante las 24 horas del día, todos los días del año, por lo que es posible mantener una vigilancia constante del comportamiento espacial y temporal de los contaminantes e informar de manera oportuna a la población la situación prevaleciente de la calidad del aire mediante el Índice Metropolitano de la Calidad del Aire (IMECA), así como, poner en marcha el programa de contingencia ambiental cuando los niveles de contaminación son críticos. Los equipos de medición con que cuenta ésta red son analizadores de gases específicos para los siguientes contaminantes (Tabla 1.1):

Contaminante	Principio de Operación
Ozono( $O_3$ )	Fotometría en el rango de ultravioleta
Óxidos de nitrógeno( $NO, NO_2$ )	Quimiluminiscencia
Bióxido de azufre( $SO_2$ )	Fluorescencia pulsante
Monóxido de carbono( $CO$ )	Espectroscopía no dispersiva por correlación del filtro gaseoso
Acido Sulfhídrico( $H_2S$ )	Ionización de flama
Partículas suspendidas fracción respirable( $PM_{10}$ )	Atenuación de radiación Beta y balanza de oscilación.

Tabla 1.1: Principio de Operación de los equipos de medición del RAMA (Ver INE et. al. (1997)).

Dado que se concentrarán esfuerzos en el estudio del incumplimiento de la norma para ozono, a continuación, se define lo que es el ozono:

**Definición 1.1.** *El ozono es un contaminante que se forma en la atmósfera a partir de reacciones muy complejas, es decir, no es emitido directamente de los escapes o chimeneas como algunos contaminantes.*

Existen dos ciclos generales de reacciones fotoquímicas en la formación del ozono troposférico, en las que participan el oxígeno molecular y dos de los denominados precursores del ozono: los óxidos de nitrógeno ( $\text{NO}_x$ ) y los hidrocarburos (HC).

El primer ciclo se forma a partir de oxígeno ( $\text{O}_2$ ), óxido nítrico (NO), átomos de oxígeno (O),  $\text{O}_2$  atmosférico y bióxido de nitrógeno mientras que en el segundo participan compuestos orgánicos volátiles (COV), bióxido de nitrógeno ( $\text{NO}_2$ ), óxido nítrico (NO), átomos de oxígeno (O) y  $\text{O}_2$  atmosférico. Sin embargo, la complejidad de los procesos reactivos implica que las concentraciones pico de ozono no sean directamente proporcionales a las de sus precursores (óxidos de nitrógeno e hidrocarburos).

Dado que se estará analizando el caso del Distrito Federal, en la Figura 1.1, se presenta la distribución de dichas estaciones

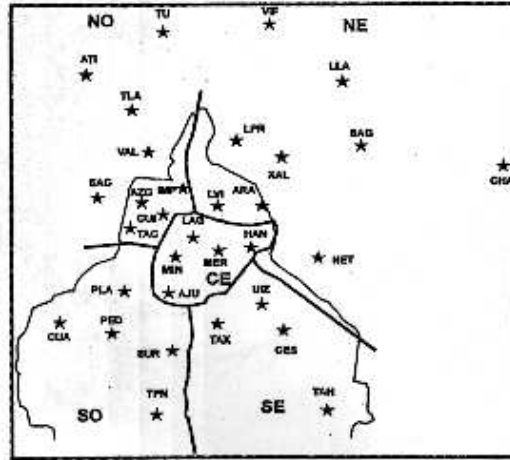


Figura 1.1: Distribución de las estaciones de monitoreo en el D.F. (Ver INE et. al. (1997)). Donde: **TAC** = Tacuba, **AZC** = Azcapotzalco, **IMP** = Instituto Mexicano del Petróleo, **CUI** = Cuitlahuac, **LVI** = La Villa, **ARA** = Aragón, **LAG** = Lagunilla, **MER** = Merced, **HAN** = Hangares, **MIN** = Metro Insurgentes, **BJU** = Benito Juárez, **SUR** = Santa Ursula, **PED** = Pedregal, **PLA** = Plateros, **CUA** = Cuajimalpa, **TPN** = Tlalpan, **CES** = Cerro de la Estrella, **UIZ** = UAM-Iztapalapa, **TAX** = Taxqueña y **TAH** = Tlahuac.

## 1.2. Normas de calidad del aire

En esta sección se explica cómo se lleva a cabo el control de la calidad del aire en México, particularmente, para el caso del ozono.

**Definición 1.2.** *La norma de calidad del aire para determinado contaminante atmosférico es el valor que establece la concentración máxima de contaminantes en el ambiente, la cual no debe sobrepasarse más de una vez por año, para que pueda garantizarse la salud de la población.*

En México se cuenta con una norma o estándar de calidad del aire para los siguientes contaminantes atmosféricos: Bióxido de azufre ( $SO_2$ ); Monóxido de carbono (CO); Bióxido de nitrógeno ( $NO_2$ ); Ozono ( $O_3$ ); partículas suspendidas totales (PST); partículas menores a 10 micrómetros de diámetro ( $PM_{10}$ ) y Plomo (Pb).

**Observación.** Los niveles o concentraciones de los contaminantes en el aire se expresan en unidades como: partes por millón (*ppm*), partes por billón (*ppb*), o microgramos por metro cúbico ( $\mu g/m^3$ ). Sin embargo debido a la dificultad de dicha nomenclatura, se han desarrollado índices de contaminación. De tal manera que a continuación se da una definición para dicho término.

**Definición 1.3.** *Un índice de calidad del aire pondera y transforma las concentraciones de un conjunto de contaminantes a un número adimensional, el cual indica el nivel de contaminación presente en una localidad determinada y que pueda ser fácilmente entendido por el público.*

La forma óptima para manejar las concentraciones de los contaminantes con objeto de obtener un número significativo depende del algoritmo que se utilice para el cálculo del índice. De tal manera que, la implicación más importante es como ponderar los efectos de los contaminantes. En México, se utilizan tanto las normas de calidad del aire como los niveles de daño significativo para ponderar los efectos de los contaminantes y así elaborar reportes diarios de la calidad del aire. De esta forma, se cuenta con el Índice Metropolitano de la Calidad del Aire (IMECA), el cual se basa en la utilización de funciones lineales segmentadas, correspondientes a los estándares internacionales de calidad del aire, en cuanto a los criterios de episodios y los niveles de daño significativo. Dicho índice incluye las variables CO,  $O_3$ ,  $NO_2$ , PST,  $PM_{10}$  y  $SO_2$ .

**Definición 1.4.** *Formalmente la función que define el Índice Metropolitano de la Calidad del Aire (IMECA) se expresa de la siguiente manera:*

$$IMECA = \text{máx}(I_1, I_2, I_3, \dots, I_n),$$

donde  $I_1, I_2, I_3, \dots, I_n$  son los subíndices individuales para cada uno de los contaminantes.

Como ya se mencionó, el cálculo de los subíndices se basa en funciones lineales segmentadas, y para los contaminantes mencionados arriba se utilizan los puntos de quiebre (valores correspondientes a los estándares de calidad del aire a partir de información local) siguientes (Tabla 1.2):

Puntos de quiebre	PST (24hr)	PM10 (24hr)	SO <sub>2</sub> (24hr)	NO <sub>2</sub> (1hr)	CO (8hr)	O <sub>3</sub> (1hr)
$x_1$	260	150	0.13	0.21	11	0.11
$x_2$	546	350	0.35	0.66	22	0.23
$x_3$	627	420	0.56	1.1	31	0.35
$x_4$	864	510	0.78	1.6	41	0.48
$x_5$	1000	600	1.00	2.00	50	0.60

Tabla 1.2: Puntos de quiebre del IMECA (Ver INE et. al. (1997)).

y sus significados son:

Calidad del aire	IMECA
Buena o satisfactoria	0-100
Regular o No satisfactoria	101-150
Mala	151-200
Muy Mala	201 en adelante

Tabla 1.3: IMECA (Ver INE et. al. (2000)).

El primer punto de quiebre  $x_1$  representa la norma de calidad del aire para los distintos contaminantes y equivale a 100 puntos IMECA. Mientras que los puntos de quiebre subsecuentes se diseñaron tomando en cuenta los criterios de salud ambiental. De tal forma que los puntos de quiebre son los valores de la concentración del contaminante que se relaciona con la salud.

Cabe señalar que las normas de calidad del aire se publicaron en el Diario Oficial de la Federación en diciembre de 1994 por la Secretaría de Salud. Dando como resultado para el ozono un valor límite de concentración de 0.11ppm en un tiempo promedio de una hora con una frecuencia máxima aceptable de 1 vez cada 3 años.

En julio de 2000 se publicó el proyecto de modificación de la norma de ozono, la cual mantiene el límite vigente y adiciona uno nuevo de 0.08 ppm en promedio de 8 horas. Por otro lado, las normas como las ya mencionadas se establecen a partir del análisis de estudios epidemiológicos, toxicológicos y de exposición, donde se identifica el nivel más bajo de contaminación que es capaz de causar un impacto negativo en la salud de algún grupo de la población. Sin embargo, en México, dichos estudios todavía no producen resultados, por eso se han tomado como base fundamental la revisión de normas establecidas por la Organización Mundial de la Salud o por Estados Unidos de América, país donde se genera la mayoría del conocimiento sobre dichos tópicos.

### 1.3. Cálculo de equivalencias

Puesto que en México se utiliza el IMECA para dar los respectivos informes de calidad del aire, es necesario mostrar el algoritmo a través del cual se hace la transformación de las concentraciones máximas a IMECA. Éste algoritmo es presentado a continuación:

#### Algoritmo para el cálculo del IMECA

Si  $y$  indica el índice IMECA y  $x$  indica la medida del contaminante, entonces:

- $y = \frac{100}{x_1}x$ ,
- $y = \frac{100}{x_2-x_1}(x - x_2) + 200, \quad x_1 < x \leq x_2$
- $y = \frac{100}{x_3-x_2}(x - x_3) + 300, \quad x_2 < x \leq x_3$
- $y = \frac{100}{x_4-x_3}(x - x_4), \quad x_3 < x \leq x_4$
- $y = \frac{100}{x_5-x_4}(x - x_5) + 500, \quad x_4 < x \leq x_5,$

donde los puntos de quiebre  $x_1, x_2, x_3, x_4$  y  $x_5$  para los contaminantes son dados en la Tabla 1.2. Un ejemplo del funcionamiento de dicho algoritmo se da a continuación.

**Ejemplo:**

Supóngase que para cierto día la concentración máxima diaria de ozono fue de 0.19 ppm. Ahora bien, para obtener el índice *IMECA* asociado a dicha concentración se hace lo siguiente:

De acuerdo con los puntos de quiebre para ozono se tiene que los cortes del algoritmo para ozono son:  $x_1=0.11, x_2=0.23, x_3=0.35, x_4=0.48$  y  $x_5=0.60$ , lo cual da como resultado los siguientes intervalos:

- medición mayor que 0.00 ppm y menor que 0.11 ppm  $(0, x_1]$
- medición mayor que 0.11 ppm y menor que 0.23 ppm  $(x_1, x_2]$
- medición mayor que 0.23 ppm y menor que 0.35 ppm  $(x_2, x_3]$
- medición mayor que 0.35 ppm y menor que 0.48 ppm  $(x_3, x_4]$
- medición mayor que 0.48 ppm y menor que 0.60 ppm  $(x_4, x_5]$ .

Dado que  $x=0.19$ , esto quiere decir que dicho valor cae entre  $x_1=0.11$  y  $x_2=0.23$  de tal manera que se utilizará la ecuación en la que se encuentran ambos cortes ( $x_1$  y  $x_2$ ). De lo cual resulta que el *IMECA* se obtiene con la ecuación

$$y = \frac{100}{x_2 - x_1}(x - x_2) + 200 = \frac{100}{0.23 - 0.11}(0.19 - 0.23) + 200 = 166.67$$

De esta forma, el *IMECA* correspondiente a ese día tiene un valor de 166.67 puntos, el cual por conveniencia, en general, es redondeado al siguiente valor, en éste caso 167 puntos *IMECA*.

Por lo tanto, el *IMECA* correspondiente a ozono en ese día es 167 puntos.

## 1.4. Consecuencias del contaminante ozono

Algunos de los efectos que causa el ozono en humanos bajo ciertas condiciones de exposición (dependiendo de la concentración, duración y actividad física realizada) son (ver, por ejemplo, INE et. al. (2000)):



- i) inflamación pulmonar,
- ii) depresión del sistema inmunológico frente a infecciones pulmonares,
- iii) cambios agudos en la función, estructura y metabolismo pulmonar, y
- iv) efectos sistemáticos en órganos blandos distantes al pulmón, como, por ejemplo, el hígado.

Algunos estudios (DDF et. al. (1996)) han revelado que cuando se expone a niños y adultos jóvenes a concentraciones de 0.12 a 0.16 ppm por periodos de 1 o 2 horas, mientras llevan a cabo diferentes niveles de ejercicio, ellos pueden presentar decrementos en su función pulmonar. Mientras que por periodos de exposición prolongada (7 horas), y, concentraciones de ozono en el intervalo de 0.08 a 0.12 ppm, algunos estudios indican que existe un decremento progresivo de la función pulmonar, así como un incremento en los síntomas respiratorios en situaciones de ejercicio moderado.

Los efectos pulmonares observados en seres humanos saludables expuestos a concentraciones urbanas típicas de ozono consisten en un decremento de la capacidad inspiratoria, una broncoconstricción moderada y síntomas subjetivos de tos y dolor al inspirar prolongadamente. La reducción de la capacidad inspiratoria da como resultado una reducción en la capacidad forzada (es el volumen máximo de aire que se puede soplar a una velocidad máxima después de inhalar profundamente) y en la capacidad pulmonar total (cantidad de aire contenido en los pulmones tras una inspiración máxima), y en combinación con la broncoconstricción (asma) contribuye a una reducción en el volumen expiratorio forzado en un segundo (es el volumen máximo de aire que se puede exhalar durante el primer segundo después de una expiración completa) (ver, INE et. al. (2000)).

La Secretaría de Salud, a través de su sistema de vigilancia epidemiológica, analizó un total de 81 episodios de contingencia ambiental ocurridos entre 1992 y 1994, en situaciones en las que se sobrepasaron los 250 puntos del IMECA. La zona más afectada fue la suroeste con 58 episodios. Los síntomas que se presentan en la salud de la población de la zona tienen una clara correlación positiva con el aumento en el nivel del índice IMECA. Los síntomas comunmente observados son: disnea (dificultad para respirar), cefalea, conjuntivitis, irritación de las mucosas respiratorias y tos productiva.

Información estadística reciente muestra que en los años 1994, 1995, 1996 (de acuerdo con los datos de la Tabla 1.4) los niveles del contaminante ozono tienden a estabilizarse, contrariamente a lo registrado en la mayor parte de la ciudad entre 1986 y 1991. No obstante, las concentraciones de ozono superan frecuentemente la norma de calidad del aire, alcanzando niveles que superan en más de un 100 % el límite establecido.

	Mayor o igual a 100	Mayor o igual a 150	Mayor o igual a 200	Mayor o igual a 250	Mayor o igual a 300
1990	93 %(338)	64 %(235)	26 %(94)	8 %(28)	1 %(4)
1991	96 %(350)	78 %(284)	44 %(162)	15 %(56)	2 %(7)
1992	90 %(329)	72 %(264)	34 %(125)	11 %(39)	3 %(12)
1993	89 %(324)	66 %(240)	23 %(85)	4 %(14)	0.3 %(1)
1994	95 %(346)	72 %(263)	27 %(100)	1 %(4)	0
1995	89 %(325)	72 %(263)	26 %(94)	2 %(7)	0
1996	89 %(327)	64 %(235)	19 %(71)	1 %(5)	0

Tabla 1.4: Porcentaje y número de días por encima de los 100, 150, 200, 250 y 300 puntos IMECA de ozono en la ZMVM de 1990 a 1996 (Ver INE et. al. (1997)).

Por otro lado, la calidad del aire de una cuenca atmosférica también depende, en primer instancia, del volumen de contaminantes emitidos, del comportamiento físico químico de éstos y de la dinámica meteorológica que determina su dispersión, transformación y remoción en la atmósfera. En particular el comportamiento del ozono depende en un 30 % de las fluctuaciones meteorológicas y en un 70 % de los cambios en el volumen de emisiones de contaminantes. Sin embargo, las condiciones fisiográficas y climáticas del Valle de México, contribuyen de manera determinante a los problemas de contaminación de la ciudad. Dichas condiciones son:

- i) Se encuentra a una altura de 2 mil 240 metros, por lo que el contenido de oxígeno del aire es 23 % menor que al nivel del mar, provocando que los procesos de combustión interna sean menos eficientes y produzcan por tanto una mayor cantidad de contaminantes.
- ii) Está rodeado por las montañas de las sierras del Ajusco, Chichinautzin, Nevada, Las Cruces, Guadalupe y Santa Catarina, las que constituyen

una barrera física natural para la circulación del viento, impidiendo el desalojo del aire contaminado fuera del Valle.

- iii) Se localiza dentro de la región central del país, por lo cual está sujeto también a la influencia de sistemas anticiclónicos, generados tanto en el Golfo de México como en el Océano Pacífico. Ocasionalmente una gran estabilidad atmosférica, inhibiendo el mezclado vertical del aire.
- iv) Presenta con frecuencia inversiones térmicas que provocan el estancamiento de los contaminantes. Por las mañanas, la capa de aire que se encuentra en contacto con la superficie del suelo adquiere una temperatura menor que las capas superiores, por lo que se vuelve más densa y pesada. Las capas de aire que se encuentran a mayor altura y que están relativamente más calientes actúan entonces como una cubierta que impide el movimiento ascendente del aire contaminado.
- v) Recibe una abundante radiación solar debido a su latitud de  $19^\circ$  N, lo que hace que su atmósfera sea altamente fotorreactiva. De tal manera que en presencia de la luz solar, los hidrocarburos y los óxidos de nitrógeno reaccionan fácilmente para formar ozono y otros oxidantes.

Existen dos técnicas con las que se puede analizar la evolución histórica de la contaminación ambiental:

- 1) Frecuencia de excedencias sobre los niveles establecidos.
- 2) Valor promedio de los máximos diarios.

A partir de ellas se puede obtener valiosa información, como lo muestra la primera técnica que se relaciona directamente con la salud, mientras que la segunda da una idea cuantitativa del avance alcanzado en el control de la contaminación. Ambas técnicas fueron aplicadas en el caso de ozono por la Secretaría del Medio Ambiente a partir de una serie de 11 años de información (donde se señala que el principal problema al que se enfrenta al estudiar las tendencias es el efecto que las condiciones meteorológicas tienen en la formación de ozono). De acuerdo a los resultados obtenidos por ambas técnicas se pudo concluir que el año en que se registró el mayor número de excedencias sobre 200, 250, y 300 IMECA fue el año de 1991. Sin embargo, cuando se compara el porcentaje de reducción en el nivel promedio anual del contaminante contra el porcentaje de reducción en el número de

excedencias, la última técnica rinde resultados más optimistas. Es decir, al comparar diferentes años se obtiene una mayor reducción en el número de excedencias. Dejando claro que aún un pequeño porcentaje de reducción en el promedio anual del nivel del contaminante se traducirá en un considerable efecto benéfico en la salud.

A partir de la Tabla 1.4 se tiene que el número de días con IMECA superior a la norma fue prácticamente el mismo en 1996 y 1995 (89 % de los días), con una disminución en los valores superiores a los 200 y 250 puntos. En este último caso, en 1996 se tuvieron dos días menos que en 1995, alcanzándose el valor de contingencia en 5 ocasiones.

De acuerdo con DDF et. al. (1996), los análisis estadísticos demuestran que la frecuencia de las concentraciones de ozono tienden a aproximarse a una distribución normal. Este hecho estadístico es fundamental en el diseño de la estrategia para elevar la calidad del aire de la Zona Metropolitana del Valle de México. Con base en ello, el propósito fundamental de la estrategia es reducir la media de la distribución hacia valores más bajos de tal forma que los valores máximos y la frecuencia de incumplimiento de las normas establecidos se reduzca. Consistentemente, se busca incrementar el número de días con bajos niveles de ozono.

# Capítulo 2

## Cadenas de Markov en tiempo continuo

Con este capítulo se hará una presentación de algunos resultados referentes a la teoría de las cadenas de Markov a tiempo continuo, cuya importancia reside en el manejo del concepto de periodos aleatorios de tiempo que hay entre cada cambio de estado. En la Sección 2.1 se dará una breve semblanza de los resultados elementales de la teoría de probabilidades. En la Sección 2.2 se desarrollan las ecuaciones diferenciales que las probabilidades de transición de una cadena de Markov a tiempo continuo satisfacen. En la Sección 2.3 se definen las probabilidades límite de una cadena de Markov a tiempo continuo. En la Sección 2.4 se describen las probabilidades de transición, las ecuaciones diferenciales y las probabilidades límite para el proceso de nacimiento y muerte.

### 2.1. Resultados preliminares

En esta sección se expondrá brevemente algunos resultados básicos de la teoría de la probabilidad, los cuales se encuentran mejor desarrollados en Ash (1972), Grimmett y Stirzaker (1982), Karlin y Taylor (1975) y (1981), y Ross (1996).

**Definición 2.1.** *Un acontecimiento que puede ser reproducido y observado en cualquier momento bajo las mismas circunstancias es llamado **experimento**. Sin embargo, éste se denominará un **experimento aleatorio** si la*

reproducción del experimento en varias ocasiones puede arrojar resultados diferentes.

**Definición 2.2.** El conjunto de todos los posibles resultados de un experimento aleatorio se denotará por  $\Omega$  y se denominará **espacio muestral**, es decir,

$$\Omega = \{\omega \mid \omega \text{ es un resultado del experimento aleatorio}\}.$$

**Definición 2.3.** Cualquier colección de posibles resultados de un experimento aleatorio se llamará un **evento** y se denotará por  $A$ .

**Definición 2.4.** Sea  $\mathcal{A}$  una clase de subconjuntos de  $\Omega$ . Se dice que  $\mathcal{A}$  es una  $\sigma$ -álgebra si y sólo si, se cumple con las siguientes condiciones:

- (i)  $\Omega \in \mathcal{A}$ ,
- (ii) Si  $A \in \mathcal{A}$ , entonces  $A^c \in \mathcal{A}$ ,
- (iii) Si  $A_1, A_2, \dots \in \mathcal{A}$ , entonces  $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ .

**Definición 2.5.** Sea  $\Omega$  un espacio muestral y  $\mathcal{A}$  una  $\sigma$ -álgebra de  $\Omega$ , entonces la pareja  $(\Omega, \mathcal{A})$  es llamada **espacio medible**.

**Definición 2.6.** Sea  $\Omega$  un espacio muestral,  $\mathcal{A}$  una  $\sigma$ -álgebra de subconjuntos de  $\Omega$ . Una **medida** sobre una  $\sigma$ -álgebra es **una función**  $\mu: \mathcal{A} \rightarrow [0, \infty]$  tal que:

- (i)  $\mu(\emptyset) = 0$ ,
- (ii)  $\mu(A) \geq 0$ ,
- (iii) Si  $\{A_i\}$ ,  $i \in T$  es una secuencia de conjuntos disjuntos de  $\Omega$ , es decir,  $A_i \cap A_j = \emptyset$  para toda  $i \neq j$ , y  $T$  un conjunto de índices finito o infinito numerable, se tiene que:

$$\mu \left( \bigcup_{i \in T} A_i \right) = \sum_{i \in T} \mu(A_i).$$

Para el caso en que  $\mu(\Omega) = 1$ ,  $\mu$  se llamará medida de probabilidad y  $\mu: \mathcal{A} \rightarrow [0, 1]$ , será denotada por  $P$ .

**Definición 2.7.** *Un espacio de medida es una terna  $(\Omega, \mathcal{A}, \mu)$ , donde  $\Omega$  es un conjunto,  $\mathcal{A}$  es una  $\sigma$ -álgebra de subconjuntos de  $\Omega$  y  $\mu$  es una medida de  $\mathcal{A}$ . Si  $\mu$  es una medida de probabilidad, entonces  $(\Omega, \mathcal{A}, \mu)$  se denomina **espacio de probabilidad** y se denotará por  $(\Omega, \mathcal{A}, P)$ .*

**Definición 2.8.** *Sea  $(\Omega, \mathcal{A}, \mathcal{P})$  un espacio de probabilidad y defínase  $h$  como la función  $h: \Omega \rightarrow \mathbb{R}$  y a  $B$  como cualquier intervalo de  $\mathbb{R}$ . Si  $h^{-1}(B) \in \mathcal{A}$  para todo  $B \subset \mathbb{R}$ , entonces  $h$  es una **función medible** en  $(\Omega, \mathcal{A})$ .*

**Definición 2.9.** *Una **variable aleatoria**  $X$  en un espacio de probabilidad  $(\Omega, \mathcal{A}, P)$  es una función de  $\Omega$  en  $\mathbb{R}$  que es medible respecto a  $\mathcal{A}$ .*

**Definición 2.10.** *Sea  $(\Omega, \mathcal{A}, P)$  un espacio de probabilidad discreto y  $A$  un evento. Se define la **probabilidad del evento  $A$**  como:*

$$P(A) = \sum_{\omega \in A} P(\omega).$$

*siempre que el  $\sigma$ -álgebra contiene los puntos.*

**Definición 2.11.** *Un **proceso estocástico**  $X = \{X_t : t \in T\}$  es una colección de variables aleatorias definidas en un mismo espacio de probabilidad, indexadas por un conjunto  $T$  y tomando valores en un mismo conjunto  $S$ .  $T$  es llamado un conjunto de índices y  $S$  es el espacio de estados de  $X$ .*

**Observación.** Generalmente se interpreta a  $t$  como el tiempo y se llama  $X_t$  al estado del proceso al tiempo  $t$ . Si el conjunto de índices  $T$  es un conjunto contable, se llamará a  $X$  un proceso estocástico de tiempo discreto, y si  $T$  es un conjunto continuo se llamará un proceso de tiempo continuo.

**Definición 2.12.** *Un proceso estocástico  $X = \{X_t : t \geq 0\}$  con espacio de estados  $S$  finito o infinito numerable, se llamará **proceso de semi-Markov** si, dado que  $X$  está en el estado  $i \in S$ :*

- (a) *el siguiente estado al que se entrará será el estado  $j$  con una probabilidad indicada por  $P_{ij}$ ,  $i, j \in S$  con  $\sum_{j \in S} P_{ij} = 1$ .*
- (b) *dado que el siguiente estado al que se entrará es el estado  $j$ , el tiempo hasta que la transición de  $i$  a  $j$  ocurra tiene una distribución  $F_{ij}$ ,  $i, j \in S$ .*

**Definición 2.13.** Sea  $X = \{X_t : t \geq 0\}$  un proceso de semi-Markov con espacio de estados  $S$ . La sucesión de variables aleatorias  $X^s$  dado por  $\{X_n^s : n = 0, 1, \dots\}$  donde  $X_n^s$  simboliza el estado del  $n$ -ésimo cambio de estado de  $X$  tal que

$$P(X_{n+1}^s = j \mid X_n^s = i) = P_{ij}, \quad i, j \in S, \quad n = 0, 1, 2, \dots$$

es llamada **cadena de saltos asociada a  $X$** .

**Definición 2.14.** Sea  $X = \{X_t : t \geq 0\}$  un proceso semi-Markov con espacio de estados  $S$  finito o infinito numerable. Si  $X$  tiene la propiedad de que cada vez que entra al estado  $i \in S$ .

- (a) La cantidad de tiempo que permanece en dicho estado antes de hacer una transición a un estado diferente se distribuye exponencialmente con tasa  $\nu_i$ , y
- (b) Cuando el proceso deja el estado  $i$ , el siguiente estado al que entrará será el estado  $j \in S$ ,  $i \neq j$  con probabilidad  $P_{ij}$ ,  $\sum_{i \neq j} P_{ij} = 1$ , entonces  $X$  es llamada una **cadena de Markov continua**.

Ver Ross (1996).

### Observación.

1. Por lo tanto, una cadena de Markov continua es un proceso estocástico que se mueve de un estado a otro de forma que la cantidad de tiempo que permanece en un estado antes de proceder a un estado distinto, es distribuido exponencialmente.
2. Observe que, un estado  $i \in S$  para el cual  $\nu_i = \infty$  es llamado un estado instantáneo, dado que cuando entra inmediatamente sale. Mientras que si  $\nu_i = 0$ ,  $i \in S$  entonces  $i$  es llamado absorbente dado que una vez que entra al estado nunca lo deja.

**Definición 2.15.** Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov en tiempo continuo con espacio de estados  $S$ . La **tasa de transición de  $i$  a  $j$** ,  $i, j \in S$ ,  $i \neq j$ , la cual se denotará por  $q_{ij}$ , es definida por:

$$q_{ij} = \nu_i P_{ij},$$

donde  $P_{ij}$  es la probabilidad de transición y  $\nu_i$  es la tasa de la distribución exponencial dada en la Definición 2.14.



**Observación.** Observe que de acuerdo a la definición de tasa de transición de  $i$  a  $j$ ,  $i, j \in S$ ,  $i \neq j$ , se tiene que:

$$\sum_{j \neq i} q_{ij} = \sum_{j \neq i} \nu_i P_{ij} = \nu_i \sum_{j \neq i} P_{ij} = \nu_i.$$

Por tanto  $\nu_i$  es la tasa a la cual el proceso deja el estado  $i$  y  $P_{ij}$  es la probabilidad de ir a  $j$ , dado que se está en  $i$ . Adicionalmente,  $q_{ij}$  es la tasa con que el proceso hace una transición del estado  $i$  al estado  $j$ . Ahora, se presentará otra manera de definir una cadena de Markov de tiempo continuo.

**Definición 2.16.** *Considere un proceso estocástico a tiempo continuo  $X = \{X_t : t \geq 0\}$  que toma valores en el conjunto finito o numerable  $S$ . Se dice que dicho proceso es una **cadena de Markov de tiempo continuo** o **cadena de Markov continua** si para todo  $s, t \geq 0$ ,  $0 \leq t_1 < t_2 < \dots < t_n < s$ ,  $A \subseteq S$ ,  $i_1, i_2, \dots, i_n, i \in S$  vale*

$$\begin{aligned} P\{X_{t+s} \in A \mid X_{t_1} = i_1, X_{t_2} = i_2, \dots, X_{t_n} = i_n, X_s = i\} \\ = P\{X_{t+s} \in A \mid X_s = i\}. \end{aligned}$$

En otras palabras, una cadena de Markov a tiempo continuo es un proceso estocástico que tiene la propiedad de Markov, es decir, la distribución condicional del estado al tiempo  $t + s$ , dado el estado al tiempo  $s$  y todos los estados a tiempos anteriores a  $s$ , depende sólo del estado en el pasado más reciente (es decir, al tiempo  $s$ ) y es independiente del pasado.

**Definición 2.17.** *Considere la cadena de Markov en tiempo continuo  $X = \{X_t : t \geq 0\}$  con espacio de estados  $S$ , y  $t, s \in [0, \infty)$ . La probabilidad  $P_{ij}(s, s + t)$ ,  $i, j \in S$ , dada por*

$$P_{ij}(s, s + t) = P\{X_{t+s} = j \mid X_s = i\}$$

es llamada **probabilidad de transición de la cadena  $X$** . La matriz  $P(s + t) = (P_{ij}(s + t))_{i, j \in S}$  es llamada **matriz de transición de  $X$** . Cuando  $P_{ij}(s, s + t)$ ,  $i, j \in S$ ,  $s, t \geq 0$  no depende de  $s$ , se dice que  $X$  es **homogénea en el tiempo** y la probabilidad y matriz de transición son indicadas por  $P_{ij}(t)$ ,  $i, j \in S$ , y  $P(t)$ , respectivamente. Si la cadena de Markov es tal que para todo  $t, s \in [0, \infty)$  se tiene que:  $X(t + s) - X(t)$  tiene la misma distribución para todo  $t \geq 0$ , entonces  $X$  es llamada **cadena con incrementos estacionarios**.

**Observación.**

Note que

$$\sum_{j \in S} P_{ij}(t) = 1, \quad 0 \leq P_{ij}(t) \leq 1, \quad i, j \in S, t \geq 0.$$

**2.2. Ecuaciones diferenciales de Kolmogorov**

Antes de desarrollar dichas ecuaciones se enunciará un pequeño lema que será de gran utilidad.

**Lema 2.1 (Ecuaciones de Chapman-Kolmogorov).** *Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov continua homogénea en el tiempo con espacio de estados  $S$  y probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S$ ,  $t \geq 0$ . Entonces para  $t, s \geq 0$  se satisface:*

$$P_{ij}(t+s) = \sum_{k \in S} P_{ik}(s)P_{kj}(t), \quad i, j \in S.$$

**Demostración:** Tome  $s, t \geq 0$ , entonces:

$$\begin{aligned} P_{ij}(t+s) &= P\{X_{t+s} = j \mid X_0 = i\} \\ &= \frac{P\{X_{t+s} = j, X_0 = i\}}{P\{X_0 = i\}} \\ &\stackrel{(1)}{=} \sum_{k \in S} \frac{P\{X_{t+s} = j, X_s = k, X_0 = i\}}{P\{X_0 = i\}} \\ &\stackrel{(2)}{=} \sum_{k \in S} P\{X_{t+s} = j \mid X_s = k, X_0 = i\} \frac{P\{X_s = k, X_0 = i\}}{P\{X_0 = i\}} \\ &\stackrel{(2)}{=} \sum_{k \in S} P\{X_{t+s} = j \mid X_s = k, X_0 = i\} P\{X_s = k \mid X_0 = i\} \\ &\stackrel{(3)}{=} \sum_{k \in S} P\{X_{t+s} = j \mid X_s = k\} P\{X_s = k \mid X_0 = i\} \end{aligned}$$

$$\stackrel{(4)}{=} \sum_{k \in S} P_{ik}(s) P_{kj}(t).$$

□

Recuerde que  $P_{ij}(t) = P\{X_{t+s} = j \mid X_s = i\}$ ,  $i, j \in S$  representa la probabilidad de que la cadena  $X = \{X_t : t \geq 0\}$  con espacio de estados  $S$ , estando en el estado  $i$  estará en  $j$  después de un tiempo  $t$ .

Se derivarán dos conjuntos de ecuaciones diferenciales para  $P_{ij}(t)$ ,  $i, j \in S, t \geq 0$ . Sin embargo, antes de hacerlo se enunciará el siguiente lema:

**Lema 2.2.** *Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov en tiempo continuo con espacio de estados  $S$  y probabilidades de transición  $P_{ij}$ ,  $i, j \in S, t \geq 0$*

- (a) *para cada  $i \in S$ , se tiene que  $\lim_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} = \nu_i$  y se le llamará la tasa de salida de  $i$ .*
- (b) *para toda  $i, j \in S, i \neq j$ , se tiene que  $\lim_{h \rightarrow 0} \frac{P_{ij}(h)}{h} = q_{ij}$  y se le llamará la tasa de transición de  $i$  a  $j$ .*

**Demostración:** Ver por ejemplo, Ross (1996). □

Las ecuaciones diferenciales de Kolmogorov muestran como se lleva a cabo el proceso de salto de un estado  $i$  a otro  $j$ . Es decir, en las ecuaciones hacia atrás se supone que se hace un salto muy pequeño desde el estado  $i$ , antes de llegar al estado  $j$ , mientras que en las ecuaciones hacia adelante lo que se considera es que se hace un salto muy grande desde el estado  $i$  antes de llegar al estado  $j$ . Luego entonces lo que proporcionan estas ecuaciones es la posibilidad de estar en el estado  $j$  después de haber realizado ya sea un salto muy pequeño o un salto muy grande entre el estado  $i$  y el estado  $j$ .

**Teorema 2.1. Ecuaciones diferenciales hacia atrás de Kolmogorov.**  
*Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov en tiempo continuo con espacio*

- 
- (1). Ley de probabilidad total.
  - (2). Definición de probabilidad condicional.
  - (3). Propiedad de Markov.
  - (4). Definición de probabilidad de transición.

de estados  $S$ , probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S, t \geq 0$ , tasa de transición  $q_{ij}$ ,  $i \neq j$ ,  $i, j \in S$  y tasa de salida de  $i \in S$  dada por  $\nu_i$ . Entonces la derivada satisface:

$$P'_{ij}(t) = \frac{dP_{ij}(t)}{dt} = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t).$$

**Demostración:** Por las ecuaciones de Chapman-Kolmogorov, se tiene que para todo  $t, h \geq 0$ , e  $i, j \in S$  vale

$$P_{ij}(t+h) = \sum_{k \in S} P_{ik}(h) P_{kj}(t). \quad (2.1)$$

Adicionalmente, por la definición de derivada de una función se tiene que,

$$P'_{ij}(t) = \lim_{h \rightarrow 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h}, \quad t \geq 0, i, j \in S.$$

De esta forma, por (2.1), se tiene que:

$$\begin{aligned} P_{ij}(t+h) - P_{ij}(t) &= \left[ \sum_{k \in S} P_{ik}(h) P_{kj}(t) \right] - P_{ij}(t) \\ &= \sum_{k \in S, k \neq i} P_{ik}(h) P_{kj}(t) - [1 - P_{ii}(h)] P_{ij}(t). \end{aligned}$$

Al dividir por  $h$ , se tiene,

$$\frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) - \frac{[1 - P_{ii}(h)]}{h} P_{ij}(t).$$

Tomando el límite cuando  $h \rightarrow 0$  resulta,

$$\lim_{h \rightarrow 0} \frac{P_{ij}(t+h) - P_{ij}(t)}{h} = \lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) - \lim_{h \rightarrow 0} \frac{[1 - P_{ii}(h)]}{h} P_{ij}(t).$$

pero, se tiene, por el Lema 2.2, que

$$\lim_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} = \nu_i.$$

Así,

$$P'_{ij}(t) = \lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) - \nu_i P_{ij}(t).$$

Para ver que se cumple

$$\lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t), \quad (2.2)$$

se dividirá la demostración en dos partes. En la primera, se tomará  $S$  finito y en la segunda se tomará  $S$  infinito numerable. Adicionalmente, se usará el hecho que si

$$\liminf_{h \rightarrow 0} a_h \leq \lim_{h \rightarrow 0} a_h \leq \limsup_{h \rightarrow 0} a_h \quad (2.3)$$

y si

$$\limsup_{h \rightarrow 0} a_h = \liminf_{h \rightarrow 0} a_h$$

entonces existe  $\lim_{h \rightarrow 0} a_h$  y

$$\liminf_{h \rightarrow 0} a_h = \limsup_{h \rightarrow 0} a_h = \lim_{h \rightarrow 0} a_h.$$

### I) Suponga $S$ finito

La ecuación (2.2) se cumple a partir de que  $S$  es finito y de que el Lema 2.2 garantiza  $\lim_{h \rightarrow 0} \frac{P_{ik}(h)}{h} = q_{ik}$ . Esto se debe a que si para cada elemento de la sucesión  $(a_n^{(k)})_{n \geq 0}$ ,  $k = 1, 2, \dots, N$ , existe  $\lim_{n \rightarrow \infty} a_n^{(k)}$ , entonces  $\lim_{n \rightarrow \infty} \sum_{k=1}^N a_n^{(k)}$  también existe.

### II) Suponga $S$ infinito numerable

Usando el resultado (2.3), sólo queda demostrar que:

- (a)  $\lim_{h \rightarrow 0} \inf \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \geq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t)$   
 (b)  $\lim_{h \rightarrow 0} \sup \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t)$   
 a) Tome  $M$  arbitraria, con  $M < \infty$ ,  $i < M$ , entonces,

$$\begin{aligned} \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) &\geq \liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < M} \frac{P_{ik}(h)}{h} P_{kj}(t) \\ &= \sum_{k \in S, k \neq i, k < M} P_{kj}(t) \liminf_{h \rightarrow 0} \frac{P_{ik}(h)}{h} \\ &= \sum_{k \in S, k \neq i, k < M} q_{ik} P_{kj}(t). \end{aligned}$$

Haciendo  $M$  tender a infinito, se tiene que

$$\liminf_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \geq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t)$$

b) Sea  $M$  arbitraria tal que  $i < M$ , entonces

$$\begin{aligned} \sum_{k \in S, k \neq i} P_{ik}(h) P_{kj}(t) &\stackrel{(1)}{=} \sum_{k \in S, k \neq i, k < M} P_{ik}(h) P_{kj}(t) + \sum_{k \in S, k \geq M} P_{ik}(h) P_{kj}(t) \\ &\stackrel{(2)}{\leq} \sum_{k \in S, k \neq i, k < M} P_{ik}(h) P_{kj}(t) + \sum_{k \in S, k \geq M} P_{ik}(h) \\ &\stackrel{(3)}{=} \sum_{k \in S, k \neq i, k < M} P_{ik}(h) P_{kj}(t) + 1 - \left[ P_{ii}(h) + \sum_{k \in S, k \neq i, k < M} P_{ik}(h) \right]. \end{aligned}$$

Entonces:

$$\begin{aligned} \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) &\stackrel{(4)}{\leq} \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < M} \frac{P_{ik}(h)}{h} P_{kj}(t) \\ &\quad + \limsup_{h \rightarrow 0} \frac{1 - P_{ii}(h)}{h} - \limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i, k < M} \frac{P_{ik}(h)}{h} \\ &\stackrel{(5)}{=} \sum_{k \in S, k \neq i, k < M} q_{ik} P_{kj}(t) + \nu_i - \sum_{k \in S, k \neq i, k < M} q_{ik}. \end{aligned}$$

Es decir, para toda  $M > i$  se cumple

$$\limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \leq \sum_{k \in S, k \neq i, k < M} q_{ik} P_{kj}(t) + \nu_i - \sum_{k \in S, k \neq i, k < M} q_{ik}.$$

Haciendo  $M \rightarrow \infty$  se tiene que

$$\limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \stackrel{(6)}{\leq} \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) + \nu_i - \nu_i = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t).$$

(1). Dividiendo la suma en  $k \geq M$  y  $k < M$ . Dado que  $M > i$  y  $k \geq M$ , se tiene que para  $k > i$  vale  $k \neq i$ .

(2). Para todo  $k, j \in S$  y  $t \geq 0$ ,  $P_{kj}(t) \leq 1$ .

(3).  $\sum_{k \geq M} P_{ik}(h) = 1 - \sum_{k < M} P_{ik}(h)$ .

(4). Por propiedades de límite.

(5). Ya que es una suma finita se puede intercambiar la suma y el límite, y por los resultados del Lema 2.2.

Por lo que

$$\limsup_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) \leq \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t).$$

Así de a), b) y la ecuación (2.3) sigue que

$$\lim_{h \rightarrow 0} \sum_{k \in S, k \neq i} \frac{P_{ik}(h)}{h} P_{kj}(t) = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t).$$

Por lo tanto

$$P'_{ij}(t) = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t),$$

con lo que queda terminada dicha demostración.  $\square$

**Teorema 2.2. Ecuaciones diferenciales hacia adelante de Kolmogorov.** *Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov en tiempo continuo con espacio de estados  $S$  y probabilidades de transición  $P_{ij}$ ,  $i, j \in S$ ,  $t \geq 0$ , y sí, además, para todo  $i \in S$  y todo  $t \geq 0$ ,  $\sum_{k \in S} P_{ik} q_k < \infty$ . Entonces, se cumple que:*

$$P'_{ij}(t) = \sum_{k \in S, k \neq j} P_{ik}(t) q_{kj} - \nu_j P_{ij}(t).$$

**Demostración:** Ver Norris (1997) y Brémaud (1999).  $\square$

## 2.3. Probabilidades límite

En la presente sección se dan a conocer las probabilidades límite para una cadena de Markov a tiempo continuo y que pueden ser encontradas más ampliamente en: Sánchez (2002), Alfaro R. (2003) y Ross (1996).

**Definición 2.18.** *Sea  $X \geq 0$  una variable aleatoria. Se dice que la variable aleatoria  $X$  es **lattice** si existe  $k \geq 0$  tal que*

$$\sum_{n=0}^{\infty} P\{X = nk\} = 1,$$

---

(6).  $\nu_i = \sum_{j \neq i, j \in S} q_{ij}$ .

es decir, la variable aleatoria  $X$  es *lattice* si sólo toma valores múltiplos de un entero no negativo  $k$ . Se definirá el período de  $X$  como el valor más grande de  $k$  que cumpla con dicha propiedad. Además, si  $X$  es *lattice* y  $F$  constituye la función de distribución de  $X$ , entonces se dirá que  $F$  también es *lattice*.

**Definición 2.19.** Sea  $X = \{X_n : n = 0, 1, \dots\}$  una cadena de Markov de tiempo discreto con espacio de estados  $S$  y probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S$ . Se dice que  $X$  es **irreducible** si para toda  $i, j \in S$  existe un entero  $m \geq 0$  tal que:

$$P\{X_{n+m} = j \mid X_n = i\} = P_{ij}^{(m)} > 0,$$

es decir,  $X$  es irreducible si todos los estados están comunicados entre sí.

**Definición 2.20.** Sea  $X = \{X_n : n = 0, 1, \dots\}$  una cadena de Markov a tiempo discreto con espacio de estados  $S$  y probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S$ . Se dice que  $X$  es **recurrente positiva** si para todo  $i \in S$ , dado que  $X_0 = i$  y

$$R_{ii} = \min\{n \geq 1, X_n = i\}$$

es el tiempo de primer retorno a  $i$ , se tiene que:

$$\mu_{ii} = E[R_{ii} \mid X_0 = i] < \infty.$$

**Definición 2.21.** Sea  $X = \{X_n : n = 0, 1, \dots\}$  una cadena de Markov a tiempo discreto, con espacio de estados  $S$  y probabilidades de transición  $P_{ij}$ ,  $i, j \in S$ ,  $i \neq j$ . La **distribución estacionaria** de  $X$  es una colección de números  $\pi(i)$ ,  $i \in S$ , que satisfacen para toda  $i \in S$ :

- (a)  $\pi(i) \geq 0, i \in S$ ,
- (b)  $\sum_{i \in S} \pi(i) = 1$ ,
- (c)  $\pi(i) = \sum_{k \in S} \pi(k)P_{ki}$ .

**Definición 2.22.** Sea  $X = \{X_t : t \geq 0\}$  un proceso de semi-Markov con espacio de estados  $S$ . Sea  $X^s = \{X_n^s : n = 0, 1, 2, \dots\}$  la cadena de saltos asociada a  $X$ . Se dice que el proceso  $X$  es **irreducible** si la cadena de saltos es irreducible.

En lo subsecuente, denótese por  $\mu_i$  la esperanza del tiempo de permanencia por el proceso semi-Markov en el estado  $i$  antes de pasar a un estado  $j$ . Asimismo, defínase  $R_{ii}$  como el tiempo entre transiciones sucesivas al estado  $i$ , y  $\mu_{ii} = E[R_{ii}]$  como el tiempo esperado de regreso al estado  $i$ .



**Teorema 2.3.** Sea  $X = \{X_t : t \geq 0\}$  un proceso de semi-Markov homogéneo en el tiempo irreducible con espacio de estados  $S$  y con  $R_{ii}$ ,  $i \in S$ , no lattice y media finita, entonces

$$P(i) = \lim_{t \rightarrow \infty} P\{X_t = i \mid X_0 = j\},$$

existe y es independiente del estado inicial  $j \in S$ , además para  $i \in S$ ,

$$P(i) = \frac{\mu_i}{\mu_{ii}}.$$

**Demostración:** Ver Ross (1996). □

**Teorema 2.4.** Sea  $X = \{X_t : t \geq 0\}$  un proceso de semi-Markov homogéneo en el tiempo, irreducible y con espacio de estados  $S$ , con  $R_{ii}$ ,  $i \in S$ , no lattice. Sea  $X^s = \{X_n^s : n = 0, 1, \dots\}$  la cadena de saltos asociada a  $X$ , recurrente positiva y con distribución estacionaria  $\pi^s(i)$ ,  $i \in S$ . entonces:

$$P(i) = \lim_{t \rightarrow \infty} P\{X_t = i \mid X_0 = j\} = \frac{\pi^s(i)\mu_i}{\sum_{j \in S} \pi^s(j)\mu_j} > 0.$$

**Demostración:** Ver Ross (1996). □

**Teorema 2.5.** Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov continua homogénea en el tiempo con espacio de estados  $S$  y probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S$ ,  $t \geq 0$ . Sea  $X^s = \{X_n^s : n = 0, 1, \dots\}$  la cadena de saltos de  $X$ , irreducible y recurrente positiva con distribución estacionaria  $\pi^s(i)$ ,  $i \in S$ . Entonces, existe el límite de probabilidades de  $X$ :

$$P(j) = \lim_{t \rightarrow \infty} P_{ij}(t),$$

y está dado por

$$P(j) = \frac{\frac{\pi^s(j)}{\nu_j}}{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i}}, \quad j \in S \quad (2.4)$$

**Demostración:** Es claro que se cumplen las condiciones del Teorema 2.4, y  $\mu_i = \frac{1}{\nu_i}$  ya que una cadena de Markov en tiempo continuo es un proceso de semi-Markov con distribución  $F_{ij}$ ,  $i, j \in S$ , igual a una distribución exponencial de parámetro  $\nu_i > 0$ , por lo que

$$P(i) = \lim_{t \rightarrow \infty} P\{X_t = i \mid X_0 = j\} = \frac{\pi^s(i)\mu_i}{\sum_{j \in S} \pi^s(j)\mu_j} = \frac{\frac{\pi^s(i)}{\nu_i}}{\sum_{i \in S} \frac{\pi^s(j)}{\nu_j}}.$$

□

**Proposición 2.1.** Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov continua homogénea en el tiempo, con espacio de estados  $S$ , que cumple con las condiciones del Teorema 2.5, entonces:

- (a)  $P(j) \geq 0, j \in S$ ,
- (b)  $\sum_{j \in S} P(j) = 1$ ,
- (c)  $\nu_j P(j) = \sum_{i \in S} \nu_i P(i) P_{ij}$ ,
- (d)  $\nu_j P(j) = \sum_{i \in S} P(i) q_{ij}$ .

**Demostración:** Por (2.4) se tiene que (a) y (b) se cumplen, dado que  $\pi^s(j) \geq 0, j \in S$  y también  $\nu_i > 0, i \in S$ , entonces:

$$P(j) = \frac{\frac{\pi^s(j)}{\nu_j}}{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i}} \geq 0,$$

asimismo

$$\sum_{j \in S} P(j) = \sum_{j \in S} \frac{\frac{\pi^s(j)}{\nu_j}}{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i}} = \frac{\sum_{j \in S} \frac{\pi^s(j)}{\nu_j}}{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i}} = 1$$

Dado que se cumplen las condiciones del Teorema 2.5 se tiene que:

$$\begin{aligned} \nu_j P(j) &\stackrel{(1)}{=} \frac{\pi^s(j)}{\sum_{k \in S} \frac{\pi^s(k)}{\nu_k}} \stackrel{(2)}{=} \frac{\sum_{i \in S} \pi^s(i) P_{ij}}{\sum_{k \in S} \frac{\pi^s(k)}{\nu_k}} \\ &= \frac{\sum_{i \in S} \frac{\pi^s(i)}{\nu_i} P_{ij} \nu_i}{\sum_{k \in S} \frac{\pi^s(k)}{\nu_k}} = \sum_{i \in S} \frac{\frac{\pi^s(i)}{\nu_i}}{\sum_{k \in S} \frac{\pi^s(k)}{\nu_k}} P_{ij} \nu_i \\ &\stackrel{(1)}{=} \sum_{i \in S} P(i) P_{ij} \nu_i. \end{aligned}$$

es decir,

$$\nu_j P(j) = \sum_{i \in S} P(i) P_{ij} \nu_i = \sum_{i \in S} P(i) q_{ij}.$$

□

---

(1). Por (2.4) y la definición de  $q_{ij}$  (2.15).

(2). Se cumple ya que  $\pi^s(i), i \in S$ , es la distribución estacionaria de la cadena de saltos asociada a  $X$ .

Por lo que se puede concluir que las tasas a la cual el proceso entra y sale de un estado  $i \in S$  son las mismas. Y se les llama las **ecuaciones de balance total**.

## 2.4. Procesos de nacimiento y muerte

Para apreciar la aplicación de los resultados presentados en las secciones anteriores, se expondrá a continuación el caso más sencillo de las cadenas de Markov a tiempo continuo, el proceso de nacimiento y muerte, dando a conocer aspectos particulares del mismo.

**Definición 2.23.** Sea  $X = \{X_t : t \geq 0\}$  una cadena de Markov en tiempo continuo con espacio de estados  $S = \{0, 1, 2, \dots\}$ , cuya tasa de transición del estado  $i$  al  $j$ ,  $i \neq j$ ,  $i, j \in S$ , es tal que  $q_{ij} = 0$  si  $|i - j| > 1$ . Dicho proceso se llama **proceso de nacimiento y muerte**.

Es decir, un proceso de nacimiento y muerte es una cadena de Markov de tiempo continuo para la cual las transiciones desde un estado  $i$ , sólo pueden hacerse para los estados vecinos  $i - 1$  e  $i + 1$ .

**Definición 2.24.** Se llamará **tasas de nacimiento y tasas de muerte** a los valores  $\lambda_i$  y  $\mu_i$  con  $i \in S = \{0, 1, 2, \dots\}$  respectivamente, definidos por:

$$\begin{aligned}\lambda_i &= q_{ii+1} > 0, & i &= 0, 1, 2, \dots \\ \mu_i &= q_{ii-1} > 0, & i &= 1, 2, \dots \\ \mu_0 &= 0\end{aligned}$$

Cabe señalar que si  $S = 0, 1, \dots, N$ , entonces, se tomará  $\lambda_N = 0$ .

**Ejemplo:** Considérese un sistema de atención al público cuyo estado en un instante dado es representado por el número de personas presentes en el sistema. Supóngase, que hay  $i$  personas en el sistema y que

- (i) un nuevo arribo entra al sistema a una tasa exponencial con parámetro  $\lambda_i > 0$  y
- (ii) una persona deja el sistema a una tasa exponencial con parámetro  $\mu_i > 0$ .

Es decir, si hay  $i$  personas en el sistema, entonces el tiempo hasta el siguiente arribo es distribuido exponencialmente con media  $\frac{1}{\lambda_i}$  y es independiente del tiempo hasta la siguiente partida que es distribuido exponencialmente con media  $\frac{1}{\mu_i}$ . Dicho sistema es un proceso de nacimiento y muerte. Los parámetros  $(\lambda_i)_{i=0}^{\infty}$  y  $(\mu_i)_{i=1}^{\infty}$  son las tasas de llegada (o nacimiento) y partida (o muerte).

**Observación:**

Es sabido que  $\sum_{i \in S, j \neq i} q_{ij} = \nu_i$ ,  $i \in S$ , por lo que se tiene lo siguiente para un proceso de nacimiento y muerte,

$$\nu_i = \lambda_i + \mu_i \quad i \in S,$$

y a partir de que  $q_{ij} = \nu_i P_{ij}$  se tiene que,

$$P_{ij} = \begin{cases} \frac{\lambda_i}{\lambda_i + \mu_i}, & \text{si } j = i + 1 \quad i \in S, \\ \frac{\mu_i}{\lambda_i + \mu_i}, & \text{si } j = i - 1 \quad i \in S, \\ 0, & \text{en otro caso.} \end{cases}$$

### 2.4.1. Probabilidades límite

En esta sección se presentará como obtener las probabilidades límite, cuando éstas existen, en el caso particular de un proceso de nacimiento y muerte

**Teorema 2.6.** *Sea  $X = \{X_t : t \geq 0\}$  un proceso de nacimiento y muerte con espacio de estados  $S = \{0, 1, \dots\}$  con tasas de nacimiento y muerte  $\{\lambda_i : i = 0, 1, 2, \dots\}$  y  $\{\mu_i : i = 1, 2, \dots\}$ ,  $\mu_0 = 0$  respectivamente. Entonces la probabilidad límite de  $X$  está dada por*

$$P(0) = \left[ 1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \dots \mu_2 \mu_1} \right]^{-1}$$

$$P(m) = \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \dots \mu_2 \mu_1 \left[ 1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \dots \mu_2 \mu_1} \right]}, \quad m = 1, 2, \dots$$

siempre que  $\sum_{m=1}^{\infty} \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \dots \mu_2 \mu_1} < \infty$ .

**Demostración:** Denótese por  $P(i)$ ,  $i \in S$  la probabilidad límite de la cadena  $X$ . Note que para el proceso de nacimiento y muerte se tiene que las tasas de transición son dadas por:

$$q_{ij} = \begin{cases} \lambda_i & \text{si } j = i + 1, \quad i = 0, 1, \dots \\ \mu_i & \text{si } j = i - 1, \quad i = 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$$

y además

$$\nu_i = \lambda_i + \mu_i \quad i = 0, 1, \dots$$

Por las ecuaciones de balance total para una cadena de Markov a tiempo continuo se tiene que para el caso particular de un proceso de nacimiento y muerte se cumple:

$$\nu_j P(j) = P(j-1)q_{j-1j} + P(j+1)q_{j+1j}, \quad j \in S$$

o equivalentemente

$$(\lambda_j + \mu_j)P(j) = \lambda_{j-1}P(j-1) + \mu_{j+1}P(j+1), \quad j \in S.$$

Puesto que  $S = \{0, 1, \dots\}$  se tiene que  $\mu_0 = 0$  lo que implica

$$\lambda_0 P(0) = \mu_1 P(1) \tag{2.5}$$

también se tiene que

$$(\lambda_n + \mu_n)P(n) = \lambda_{n-1}P(n-1) + \mu_{n+1}P(n+1), \quad n \geq 1. \tag{2.6}$$

Puesto que la igualdad (2.5) es de gran utilidad para obtener las probabilidades límite del proceso de nacimiento y muerte se demostrará por inducción sobre  $n$  que es válida para todo  $n \geq 0$ .

$$\lambda_n P(n) = \mu_{n+1} P(n+1), \quad n \geq 0.$$

Se cumple para  $n = 0$  por (2.5). Para demostrar para  $n = 1$  se tiene que por (2.6):

$$(\lambda_1 + \mu_1)P(1) = \lambda_0 P(0) + \mu_2 P(2) = \mu_1 P(1) + \mu_2 P(2),$$

y por lo tanto

$$\lambda_1 P(1) = \mu_1 P(1) + \mu_2 P(2) - \mu_1 P(1) = \mu_2 P(2).$$

Por hipótesis de inducción se cumple para  $n = k$ , es decir,

$$\lambda_k P(k) = \mu_{k+1} P(k+1).$$

De tal forma que sólo queda demostrar que se cumple para  $n = k+1$ . Se sabe por (2.6) que

$$(\lambda_{k+1} + \mu_{k+1})P(k+1) = \lambda_k P(k) + \mu_{k+2} P(k+2),$$

entonces

$$\begin{aligned} \lambda_{k+1} P(k+1) &= \mu_{k+2} P(k+2) + \lambda_k P(k) - \mu_{k+1} P(k+1) \\ &\stackrel{(1)}{=} \mu_{k+2} P(k+2) + \mu_{k+1} P(k+1) - \mu_{k+1} P(k+1) \\ &= \mu_{k+2} P(k+2). \end{aligned}$$

Por lo tanto se tiene que:

$$P(n) = \frac{\lambda_{n-1}}{\mu_n} P(n-1), \quad n \geq 1. \quad (2.7)$$

Utilizando (2.7) se demostrará por inducción que se cumple

$$P(m) = \frac{\lambda_{m-1} \lambda_{m-2} \dots \lambda_1 \lambda_0}{\mu_m \mu_{m-1} \dots \mu_2 \mu_1} P(0), \quad m \geq 1. \quad (2.8)$$

Para  $m = 1$ , de (2.7) se tiene que

$$P(1) = \frac{\lambda_0}{\mu_1} P(0).$$

Por hipótesis de inducción se vale la igualdad (2.8) para  $m = k$ , es decir,

$$P(k) = \frac{\lambda_{k-1} \lambda_{k-2} \dots \lambda_1 \lambda_0}{\mu_k \mu_{k-1} \dots \mu_2 \mu_1} P(0),$$

por lo que queda demostrar que se cumple para  $m = k+1$ . A partir de que se cumple (2.7), es decir,

$$P(n) = \frac{\lambda_{n-1}}{\mu_n} P(n-1),$$

---

(1). Por hipótesis.

se tiene que para  $n = k + 1$

$$P(k + 1) = \frac{\lambda_k}{\mu_{k+1}} P(k),$$

y por hipótesis de inducción  $P(k) = \frac{\lambda_{k-1}\lambda_{k-2}\dots\lambda_1\lambda_0}{\mu_k\mu_{k-1}\dots\mu_2\mu_1} P(0)$  lo que implica que

$$P(k + 1) = \frac{\lambda_k\lambda_{k-1}\lambda_{k-2}\dots\lambda_1\lambda_0}{\mu_{k+1}\mu_k\mu_{k-1}\dots\mu_2\mu_1} P(0)$$

Por lo que la igualdad (2.8) queda demostrada.

Dado que  $\sum_{m=1}^{\infty} P(m) = 1$  se tiene que

$$1 = P(0) + P(0) \sum_{m=1}^{\infty} \frac{\lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\dots\mu_2\mu_1}$$

luego,

$$P(0) = \left[ 1 + \sum_{m=1}^{\infty} \frac{\lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\dots\mu_2\mu_1} \right]^{-1}.$$

Note que si

$$\sum_{m=1}^{\infty} \frac{\lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\dots\mu_2\mu_1} = \infty,$$

entonces  $P(0) = 0$  y por tanto  $P(m) = 0$  para todo  $m \in S$ . Por lo que, para que  $P(j) > 0$ ,  $j \in S$ , se debe cumplir que:

$$\sum_{m=1}^{\infty} \frac{\lambda_{m-1}\lambda_{m-2}\dots\lambda_1\lambda_0}{\mu_m\mu_{m-1}\dots\mu_2\mu_1} < \infty$$

por lo que el teorema queda demostrado.  $\square$

## 2.4.2. Ecuaciones diferenciales de Kolmogorov

A continuación, se darán las ecuaciones diferenciales correspondientes al proceso de nacimiento y muerte. Las cuales son como sigue:

**Teorema 2.7.** *Sea  $X = \{X_t : t \geq 0\}$  un proceso de nacimiento y muerte con espacio de estados  $S = \{0, 1, 2, \dots\}$ , probabilidades de transición  $P_{ij}(t)$ ,  $i, j \in S$  y tasas de nacimiento y muerte dadas por  $\lambda_i > 0$ ,  $i = 0, 1, \dots$ ,  $\mu_i > 0$ ,  $i =$*

$1, 2, \dots, \mu_0 = 0$ , respectivamente. Las ecuaciones hacia atrás y hacia adelante de Kolmogorov de  $X$  están dadas, por:

**i) Ecuaciones hacia atrás**

$$P'_{0j}(t) = \lambda_0 P_{1j}(t) - \lambda_0 P_{0j}(t)$$

$$P'_{ij}(t) = \lambda_i P_{i+1j}(t) + \mu_i P_{i-1j}(t) - (\lambda_i + \mu_i) P_{ij}(t), \quad i = 1, 2, \dots$$

**ii) Ecuaciones hacia adelante**

$$P'_{i0}(t) = \mu_1 P_{i1}(t) - \lambda_0 P_{i0}(t)$$

$$P'_{ij}(t) = \lambda_{j-1} P_{ij-1}(t) + \mu_{j+1} P_{ij+1}(t) - (\lambda_j + \mu_j) P_{ij}(t), \quad j = 1, 2, \dots$$

**Demostración:** i) Para una cadena de Markov a tiempo continuo se tiene

$$P'_{ij}(t) = \sum_{k \in S, k \neq i} q_{ik} P_{kj}(t) - \nu_i P_{ij}(t),$$

donde,  $q_{ik}$  es la tasa de transición de  $i$  a  $k$ ,  $i \neq k$ ,  $k \in S$  y  $\nu_i$  es la tasa de salida de  $i$ . Debido a que el proceso de nacimiento y muerte es un caso particular de una cadena de Markov a tiempo continuo, se tiene que:

$$q_{ik} = \begin{cases} \lambda_i & \text{si } k = i + 1, i = 0, 1, \dots \\ \mu_i & \text{si } k = i - 1, i = 1, 2, \dots \\ 0 & \text{en otro caso.} \end{cases}$$

y para  $i \in S$  vale

$$\nu_i = \lambda_i + \mu_i.$$

**a)** Considérese  $i = 0$ . Entonces, por el Teorema 2.1 se tiene que para  $t > 0$

$$P'_{0j}(t) = q_{01} P_{1j}(t) - \nu_0 P_{0j}(t) = \lambda_0 P_{1j}(t) - \lambda_0 P_{0j}(t).$$

**b)** Sea  $i = 1, 2, \dots$ . Entonces, también por el Teorema 2.1 se tiene que para  $t > 0$

$$\begin{aligned} P'_{ij}(t) &= q_{ii+1} P_{i+1j}(t) + q_{ii-1} P_{i-1j}(t) - \nu_i P_{ij}(t) \\ &= \lambda_i P_{i+1j}(t) + \mu_i P_{i-1j}(t) - (\lambda_i + \mu_i) P_{ij}(t). \end{aligned}$$



c) Para  $t = 0$  y para todo  $i, j = 0, 1, 2, \dots$  vale

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

Por lo que de a), b), c) se obtiene que las ecuaciones hacia atrás para un proceso de nacimiento y muerte están dadas por:

$$P'_{0j}(t) = \lambda_0 P_{1j}(t) - \lambda_0 P_{0j}(t)$$

y

$$P'_{ij}(t) = \lambda_i P_{i+1j}(t) + \mu_i P_{i-1j}(t) - (\lambda_i + \mu_i) P_{ij}(t) \quad i = 1, 2, \dots$$

ii) Para una cadena de Markov a tiempo continuo se tiene

$$P'_{ij}(t) = \sum_{k \in S, k \neq j} P_{ik}(t) q_{kj} - \nu_j P_{ij}(t)$$

donde como  $q_{kj}$  es la tasa de transición de  $k$  a  $j$ ,  $k \neq j$ ,  $k \in S$  y  $\nu_j$  es la tasa de salida de  $j$ ,  $j \in S$ . Luego, para un proceso de nacimiento y muerte, se tiene que:

(a) Para el caso de  $j = 0$  por el Teorema 2.2 vale

$$P'_{i0}(t) = P_{i1}(t) q_{10}(t) - \nu_0 P_{i0}(t) = P_{i1}(t) \mu_1 - \lambda_0 P_{i0}(t)$$

(b) Para cualquier  $j = 1, 2, \dots$  también por el Teorema 2.2 vale

$$\begin{aligned} P'_{ij}(t) &= P_{ij-1}(t) q_{j-1j} + P_{ij+1}(t) q_{j+1j} - \nu_j P_{ij}(t) \\ &= P_{ij-1}(t) \lambda_{j-1} + P_{ij+1}(t) \mu_{j+1} - (\lambda_j + \mu_j) P_{ij}(t) \end{aligned}$$

(c) Para  $t = 0$  y para todo  $i, j = 0, 1, 2, \dots$

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

De (a), (b), (c) se sigue que las ecuaciones hacia adelante de un proceso de nacimiento y muerte están dadas por:

$$\begin{aligned} P'_{i0}(t) &= \mu_1 P_{i1}(t) - \lambda_0 P_{i0}(t) \\ P'_{ij}(t) &= \lambda_{j-1} P_{ij-1}(t) + \mu_{j+1} P_{ij+1}(t) - (\lambda_j + \mu_j) P_{ij}(t), \quad j = 1, 2, \dots \end{aligned}$$

□

# Capítulo 3

## Proceso Poisson

El capítulo previo fue dedicado a la elaboración de los conceptos básicos y métodos de las cadenas de Markov de tiempo continuo. Ahora en este capítulo se presentará con más detalle un caso particular de cadenas de Markov a tiempo continuo, a saber, lo que constituye el proceso Poisson.

En la primera sección se define un proceso de conteo, a partir del cual se dará la definición del proceso Poisson. Consecuentemente, se desarrolla la distribución de los tiempos de llegada y de espera. En la Sección 3.2 se muestra la distribución condicional de los tiempos de llegada. Finalmente, en la última sección se presenta el proceso Poisson no homogéneo. Dichos resultados se encuentran desarrollados más ampliamente en Ross (1989).

### 3.1. Propiedades del Proceso Poisson

**Definición 3.1.** *Un proceso estocástico  $N = \{N(t) : t \geq 0\}$  es un proceso de conteo si  $N(t)$  representa el número total de eventos que han ocurrido hasta el tiempo  $t$ . Formalmente, un **proceso de conteo**  $N(t)$  debe satisfacer:*

- (i)  $N(t) \geq 0, \quad t \geq 0$
- (ii)  $N(t)$  es un valor entero,  $t \geq 0$
- (iii) Si  $s < t$ , entonces  $N(s) \leq N(t), \quad s, t \geq 0$
- (iv) Para  $s < t$ ,  $N(t) - N(s)$  es igual al número de eventos que han ocurrido en el intervalo  $(s, t]$ ,  $s, t \geq 0$ .

**Definición 3.2.** Sea  $N = \{N(t) : t \geq 0\}$  un proceso de conteo. Se dice que dicho proceso posee **incrementos independientes** si el número de eventos ocurridos en el intervalo de tiempo  $(s, t]$ , es decir,  $N(t) - N(s)$  es independiente del número de eventos ocurridos en el intervalo de tiempo  $(r, u]$ ,  $r \leq u$ ,  $r, u > t$ , es decir,  $N(t) - N(s)$  es independiente de  $N(u) - N(r)$ .

**Definición 3.3.** Sea  $N = \{N(t); t \geq 0\}$  un proceso de conteo. Se dice que es un **proceso con incrementos estacionarios** si la distribución del número de eventos que ocurren en algún intervalo de tiempo depende sólo de la duración del intervalo. Es decir, si para todo  $t_1 > t_2$  y  $s > 0$  el número de eventos en el intervalo  $(t_1 + s, t_2 + s)$  es decir,  $N(t_2 + s) - N(t_1 + s)$  tiene la misma distribución que el número de eventos en el intervalo  $(t_1, t_2]$ , es decir,  $N(t_2) - N(t_1)$ .

Uno de los más importantes procesos de conteo es el proceso Poisson, el cual es definido como sigue:

**Definición 3.4.** El proceso de conteo  $N = \{N(t) : t \geq 0\}$  será un **proceso de Poisson** con tasa  $\lambda \in \mathbb{R}$ ,  $\lambda > 0$ , si:

- (i)  $N(0) = 0$ ;
- (ii)  $N$  posee incrementos independientes; y
- (iii) el número de eventos en algún intervalo de duración  $t$  tiene distribución Poisson con media  $\lambda t$ . Es decir, para todo  $s, t \geq 0$ ,

$$P\{N(t+s) - N(s) = n\} = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

**Observación.**

1. Nótese que la condición (iii) de la Definición 3.4, muestra que un proceso Poisson tiene incrementos estacionarios. Adicionalmente, se tiene que

$$E[N(t)] = E[N(t) - N(0)] = \lambda t,$$

(Ver Apéndice B) lo cual deja claro porqué  $\lambda$  es llamada la tasa del proceso.

2. Note que la condición (i) de la Definición 3.4 dice que el conteo de eventos empieza al tiempo  $t = 0$ . Dado un proceso  $X = \{X_t : t \geq 0\}$  esta condición es fácilmente verificada. La condición (ii) puede ser verificada directamente a partir del proceso  $X$  considerado. Sin embargo, no es siempre del todo claro cómo determinar que la condición (iii) de dicha definición se cumple. Por esta razón una definición equivalente del proceso Poisson debe ser usada.

**Definición 3.5.** *El proceso de conteo  $N = \{N(t) : t \geq 0\}$  será un proceso Poisson con tasa  $\lambda \in \mathbb{R}$ ,  $\lambda > 0$ , si*

- (i)  $N(0) = 0$ ;
- (ii) *el proceso  $N$  tiene incrementos independientes y estacionarios;*
- (iii)  $P\{N(h) = 1\} = \lambda h + o(h)$ ;
- (iv)  $P\{N(h) \geq 2\} = o(h)$ ,

donde  $o(h)$  está definida en el Apéndice E, dentro de la Definición E.1.

Para ver que se pueden usar ambas definiciones indistintamente, es necesario mostrar que ambas definiciones son equivalentes.

**Lema 3.1.** *Las Definiciones 3.4 y 3.5 son equivalentes.*

Esta demostración se hará en dos partes. En la primera parte se demostrará que la Definición 3.5 implica la Definición 3.4, y en la segunda parte se demostrará que la Definición 3.4 implica la Definición 3.5.

**Demostración:** a) Definición 3.5 implica la Definición 3.4. Note que basta demostrar que las condiciones (iii) y (iv) de la Definición 3.5 implican la condición (iii) de la Definición 3.4. Para ello, sea:

$$P_n(t) = P\{N(t) = n\}$$

que indica la probabilidad de que  $n$  eventos hayan ocurrido hasta el tiempo  $t$ . Se usará inducción matemática sobre  $n$  para demostrar la implicación. En primer lugar se obtendrá una ecuación diferencial para  $P_0(t)$ :

$$\begin{aligned} P_0(t+h) &= P\{N(t+h) = 0\} \\ &= P\{N(t) = 0, N(t+h) - N(t) = 0\} \\ &\stackrel{(1)}{=} P\{N(t) = 0\}P\{N(t+h) - N(t) = 0\} \\ &\stackrel{(2)}{=} P_0(t)[1 - \lambda h + o(h)] \end{aligned}$$

$$\frac{P_0(t+h) - P_0(t)}{h} = -\lambda P_0(t) + \frac{o(h)}{h}.$$

Tomando el límite cuando  $h \rightarrow 0$  se tiene que

$$P_0'(t) = -\lambda P_0(t),$$

o bien

$$\frac{P_0'(t)}{P_0(t)} = -\lambda.$$

Integrando con respecto al tiempo en el intervalo de 0 a  $t$  se tiene

$$\int_0^t \frac{P_0'(s)}{P_0(s)} ds = - \int_0^t \lambda ds,$$

y por lo tanto

$$\log P_0(t) - \log P_0(0) = -\lambda t.$$

Dado que  $P_0(0) = P(N(0) = 0) = 1$ , se tiene que

$$\log P_0(t) = -\lambda t,$$

o bien

$$P_0(t) = e^{-\lambda t}$$

que equivale a

$$\frac{(\lambda t)^0}{0!} e^{-\lambda t} = P\{N(t) - N(0) = 0\}.$$

Por tanto, se tiene que el apartado (iii) de la Definición 3.4 se cumple para  $n = 0$ .

Del mismo modo se obtendrá una ecuación diferencial para  $n \geq 1$ :

- 
- (1). Dado que  $N$  tiene incrementos independientes.  
 (2). Puesto que iii) y iv) de la Definición 3.5 implican que  $P\{N(h) = 0\} = 1 - \lambda h + o(h)$ .

$$\begin{aligned}
P_n(t+h) &= P\{N(t+h) = n\} \\
&= P\{N(t) = n, N(t+h) - N(t) = 0\} \\
&+ P\{N(t) = n-1, N(t+h) - N(t) = 1\} \\
&+ \sum_{k=2}^n P\{N(t+h) = n-k, N(t+h) - N(t) = k\} \\
&= P\{N(t) = n\}P\{N(h) - N(0) = 0\} \\
&+ P\{N(t) = n-1\}P\{N(h) - N(0) = 1\} \\
&+ \sum_{k=2}^n P\{N(t) = n-k\}P\{N(h) - N(0) = k\} \quad (3.1)
\end{aligned}$$

Sin embargo, por el inciso (iv) de la Definición 3.5, el último término de (3.1) es un  $o(h)$ , dado que es una suma finita de funciones que son  $o(h)$ . Usando (iii) y (iv) de la Definición 3.5 y la definición de  $P_n(t)$  se obtiene que (3.1) se puede escribir como:

$$P_n(t+h) = P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + o(h) = (1-\lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h).$$

Por lo tanto,

$$\frac{P_n(t+h) - P_n(t)}{h} = -\lambda P_n(t) + \lambda P_{n-1}(t) + \frac{o(h)}{h}.$$

Tomando el límite cuando  $h \rightarrow 0$ , se tiene que

$$P_n'(t) = -\lambda P_n(t) + \lambda P_{n-1}(t),$$

o equivalentemente,

$$e^{\lambda t}[P_n'(t) + \lambda P_n(t)] = \lambda e^{\lambda t} P_{n-1}(t). \quad (3.2)$$

Note que

$$\frac{d}{dt}(e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_n(t) + e^{\lambda t} P_n'(t),$$

por tanto (3.2) puede escribirse como

$$\frac{d}{dt}(e^{\lambda t} P_n(t)) = \lambda e^{\lambda t} P_{n-1}(t). \quad (3.3)$$

Suponga que (iii) de la Definición 3.4 vale para  $n < k$ , es decir,

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

Ahora, se demostrará que dicho resultado se cumple para  $n = k$ . De la ecuación (3.3) y de la hipótesis de inducción se tiene que

$$\frac{d}{dt}(e^{\lambda t} P_k(t)) = e^{\lambda t} \lambda P_{k-1}(t) = \lambda e^{\lambda t} \frac{(\lambda t)^{k-1}}{(k-1)!} e^{-\lambda t} = \frac{\lambda^k}{(k-1)!} t^{k-1}$$

Integrando con respecto al tiempo en el intervalo de 0 a  $t$  se tiene que

$$\int_0^t \frac{d}{ds}(e^{\lambda s} P_k(s)) ds = \frac{\lambda^k}{(k-1)!} \int_0^t s^{k-1} ds$$

y por tanto

$$e^{\lambda t} P_k(t) - e^{\lambda 0} P_k(0) = \frac{\lambda^k}{k(k-1)!} t^k$$

y dado que  $P_k(0) = P\{N(0) = k\} = 0$  para  $k \neq 0$  se tiene que

$$P_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

Por lo tanto queda demostrado que la Definición 3.5 implica la Definición 3.4.

**b)** Ahora para ver que la Definición 3.4 implica la Definición 3.5, sólo queda demostrar que los incisos (iii) y (iv) de la Definición 3.5 son válidos, puesto que los dos primeros incisos se cumplen en ambas definiciones. De la condición (iii) de la Definición 3.4 se tiene que:

$$P\{N(h) = 0\} = P\{N(h) - N(0) = 0\} = e^{-\lambda h}$$

$$P\{N(h) = 1\} = P\{N(h) - N(0) = 1\} = e^{-\lambda h} \lambda h$$

y

$$\begin{aligned} P\{N(h) \geq 2\} &= P\{N(h) - N(0) \geq 2\} \\ &= 1 - P\{N(h) - N(0) = 0\} - P\{N(h) - N(0) = 1\} \\ &= 1 - e^{-\lambda h} - e^{-\lambda h} \lambda h. \end{aligned}$$

Note que

$$\frac{P\{N(h) = 1\} - \lambda h}{h} = \frac{e^{-\lambda h} \lambda h - \lambda h}{h} = \frac{\lambda h}{h} (e^{-\lambda h} - 1) = \lambda (e^{-\lambda h} - 1).$$

Tomando el límite cuando  $h \rightarrow 0$  se tiene que

$$\lim_{h \rightarrow 0} \frac{P\{N(h) = 1\} - \lambda h}{h} = 0.$$

Por tanto,  $P\{N(h) = 1\} - \lambda h = o(h)$  y (iii) de la Definición 3.5 sigue. Adicionalmente, note que

$$\begin{aligned} \frac{P\{N(h) \geq 2\}}{h} &= \frac{1 - P\{N(h) = 0\} - P\{N(h) = 1\}}{h} \\ &= \frac{1 - e^{-\lambda h} - \lambda h e^{-\lambda h}}{h} \\ &= \frac{1 - e^{-\lambda h}}{h} - \lambda e^{-\lambda h}. \end{aligned}$$

Tomando el límite cuando  $h \rightarrow 0$  se tiene que

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{P\{N(h) \geq 2\}}{h} &= \lim_{h \rightarrow 0} \frac{1 - e^{-\lambda h}}{h} - \lambda \lim_{h \rightarrow 0} e^{-\lambda h} \\ &\stackrel{(1)}{=} \lim_{h \rightarrow 0} \lambda e^{-\lambda h} - \lambda \lim_{h \rightarrow 0} e^{-\lambda h} \\ &= 0. \end{aligned}$$

Por tanto  $P\{N(t) \geq 2\} = o(h)$  y el inciso (iv) de la Definición 3.5 sigue.  $\square$

De esta forma, queda demostrado que la Definición 3.4 implica la Definición 3.5 y por lo tanto ambas definiciones son equivalentes.

Considérese un proceso Poisson  $N = \{N(t), t \geq 0\}$  con tasa  $\lambda$ , y supóngase que cada vez que un evento ocurre es clasificado como del tipo 1 con probabilidad  $p$  o del tipo 2 con probabilidad  $1-p$  y que a su vez es independiente de los demás eventos. Sean  $N_1(t)$  y  $N_2(t)$  el número de eventos del tipo 1 y del tipo 2 que ocurren en  $[0, t]$ , respectivamente. Nótese que  $N(t) = N_1(t) + N_2(t)$ .

---

(1). Por la regla de L'Hôpital.



**Proposición 3.1.** *Los procesos de conteo  $N_1 = \{N_1(t), t \geq 0\}$  y  $N_2 = \{N_2(t), t \geq 0\}$  son procesos Poisson independientes con tasas respectivas  $\lambda p$  y  $\lambda(1-p)$ .*

**Demostración:** En primer lugar se calculará la distribución de probabilidad conjunta

$$P\{N_1(t) = n, N_2(t) = m\}.$$

Para ello, se condicionará sobre  $N(t)$  obteniendo:

$$\begin{aligned} P\{N_1(t) = n, N_2(t) = m\} &= \sum_{k=0}^{\infty} P\{N_1(t) = n, N_2(t) = m \mid N(t) = k\} \\ &\times P\{N(t) = k\}. \end{aligned}$$

Considerando que han sido  $n$  eventos del tipo 1 y  $m$  del tipo 2, se tiene que han ocurrido un total de  $n + m$  eventos en  $[0, t]$ . Es decir,

$$P\{N_1(t) = n, N_2(t) = m \mid N(t) = k\} = 0 \quad \text{cuando } k \neq n + m.$$

Por lo tanto,

$$\begin{aligned} P\{N_1(t) = n, N_2(t) = m\} &= P\{N_1(t) = n, N_2(t) = m \mid N(t) = n + m\} \\ &\times P\{N(t) = n + m\} \\ &= P\{N_1(t) = n, N_2(t) = m \mid N(t) = n + m\} \\ &\times e^{-\lambda t} \frac{(\lambda t)^{n+m}}{(n+m)!}. \end{aligned}$$

Sin embargo, dado que  $n + m$  eventos ocurrieron y que cada evento tiene probabilidad  $p$  de tener un evento del tipo 1 y probabilidad  $1 - p$  de ser un evento del tipo 2, se sigue que la probabilidad que  $n$  de ellos sean tipo 1 y  $m$  de ellos del tipo 2 está dada por la probabilidad Binomial:  $\binom{n+m}{n} p^n (1-p)^m$ . Es decir,

$$\begin{aligned} P\{N_1(t) = n, N_2(t) = m\} &= \left[ \binom{n+m}{n} p^n (1-p)^m \right] e^{-\lambda t} \frac{(\lambda t)^{n+m}}{(n+m)!} \\ &= \left[ \frac{(n+m)!}{m!n!} p^n (1-p)^m \right] e^{-\lambda t} \frac{(\lambda t)^{n+m}}{(n+m)!} \\ &= \frac{p^n (1-p)^m e^{-\lambda t} e^{-\lambda t p + \lambda t p} (\lambda t)^{n+m}}{m!n!} \\ &= e^{-\lambda t p} \frac{(\lambda t p)^n}{(n)!} e^{-\lambda t(1-p)} \frac{(\lambda t(1-p))^m}{(m)!}. \quad (3.4) \end{aligned}$$

Note que,

$$\begin{aligned}
 P\{N_1(t) = n\} &= \sum_{m=0}^{\infty} P\{N_1(t) = n, N_2(t) = m\} \\
 &= e^{-\lambda tp} \frac{(\lambda tp)^n}{(n)!} \sum_{m=0}^{\infty} e^{-\lambda t(1-p)} \frac{(\lambda t(1-p))^m}{(m)!} \\
 &= e^{-\lambda tp} \frac{(\lambda tp)^n}{(n)!}.
 \end{aligned} \tag{3.5}$$

De tal manera que,  $N_1 = \{N_1(t), t \geq 0\}$  es un proceso Poisson con tasa  $\lambda p$ . De forma similar se obtiene que

$$P\{N_2(t) = m\} = e^{-\lambda t(1-p)} \frac{(\lambda t(1-p))^m}{(m)!}, \tag{3.6}$$

es decir,  $N_2 = \{N_2(t), t \geq 0\}$  es un proceso Poisson con tasa  $\lambda(1-p)$ . Por lo tanto de (3.4), (3.5), (3.6) se tiene que  $N_1$  y  $N_2$  son procesos de Poisson independientes con tasas  $\lambda p$  y  $\lambda(1-p)$ , respectivamente.  $\square$

## 3.2. Distribución de tiempo entre llegadas y de espera

**Definición 3.6.** Sea  $N = \{N(t) : t \geq 0\}$  un proceso Poisson con parámetro  $\lambda > 0$ , la secuencia  $X = \{X_n : n \geq 1\}$  es llamada la **secuencia de tiempo entre llegadas** cuando  $X_n$  denota el tiempo entre la ocurrencia del  $(n-1)$ -ésimo y del  $n$ -ésimo evento para  $n \geq 1$ , donde  $X_0 = 0$  y  $X_1$  denota el tiempo de ocurrencia del primer evento.

De tal manera que se puede enunciar la siguiente proposición:

**Proposición 3.2.** Las variables  $X = \{X_n, n = 1, 2, \dots\}$  que integran la secuencia de tiempos entre llegadas se distribuyen independiente e idénticamente como una variable aleatoria con distribución exponencial con parámetro  $\lambda > 0$ .

**Demostración:** Sea  $n = 1$ . En primer lugar, nótese que el evento  $\{X_1 > t\}$  toma lugar si y sólo si, no ocurren eventos del tipo Poisson en el intervalo  $[0, t]$ , por lo tanto

$$P\{X_1 > t\} = P\{N(t) = 0\} = e^{-\lambda t}.$$

De tal manera que,  $X_1$  tiene una distribución exponencial con parámetro  $\lambda$ . Para obtener la distribución de  $X_2$  se condicionarará sobre  $X_1$ . Esto da:

$$\begin{aligned} P\{X_2 > t \mid X_1 = s\} &= P\{0 \text{ eventos en } (s, s+t] \mid X_1 = s\} \\ &= P\{N(s+t) - N(s) = 0\} \\ &\stackrel{(1)}{=} e^{-\lambda t}. \end{aligned}$$

De lo anterior se puede concluir que  $X_2$  también es una variable aleatoria exponencial con parámetro  $\lambda$ , y que  $X_2$  es independiente de  $X_1$ . Suponga que el resultado es válido para  $n = k$ , es decir,

$$P(X_k > t \mid X_{k-1} = s) = e^{-\lambda t}.$$

Queda demostrar que se cumple para  $n = k + 1$ , note que

$$\begin{aligned} P\{X_{k+1} > t \mid X_k = s\} &= P\{0 \text{ eventos en } (s, s+t] \mid X_k = s\} \\ &= P\{0 \text{ eventos en } (s, s+t]\} = P\{N(s+t) - N(s) = 0\} \\ &= e^{-\lambda t}. \end{aligned}$$

Por lo tanto  $X_{k+1}$  es una variable aleatoria exponencial de parámetro  $\lambda$ .  $\square$

**Definición 3.7.** *Considérese un proceso Poisson y  $X = \{X_n : n \geq 1\}$  la respectiva secuencia de tiempo entre llegadas. El tiempo de espera hasta la ocurrencia del n-ésimo evento indicado por  $S_n$ , se define:*

$$S_n = \sum_{i=1}^n X_i, \quad n \geq 1.$$

**Observación:**

La suposición de incrementos independientes y estacionarios, significa que el proceso en un instante dado es independiente de todo lo que ocurrió anteriormente a dicho instante (por tener incrementos independientes) y que conserva la distribución original (por tener incrementos estacionarios). De tal manera, que se puede decir que el proceso no tiene memoria.

---

(1). Por incrementos estacionarios.

Lo anterior muestra como se distribuye el tiempo entre la ocurrencia de dos eventos consecutivos. Sin embargo, también se puede obtener la distribución del tiempo de llegada del  $n$ -ésimo evento o el tiempo de espera hasta que el  $n$ -ésimo evento ocurra.

Note que  $S_n$  tiene una distribución Gamma con parámetros  $n$  y  $\lambda$ . Es decir, su función de densidad de probabilidad es:

$$f(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \quad t \geq 0.$$

Este resultado sigue de la Proposición 3.2 dado que  $X = \{X_n, n = 1, 2, \dots\}$  se distribuye exponencial y del hecho que la suma de  $n$  variables aleatorias independientes con distribución exponencial con un mismo parámetro  $\lambda$  tiene una distribución Gamma con parámetros  $n$  y  $\lambda$  (Ver Apéndice D).

### 3.3. Distribución condicional de los tiempos de llegada

En la sección anterior ya se obtuvo como se distribuye la secuencia de tiempo entre llegadas  $X = \{X_n : n \geq 0\}$  y  $\{S_n : n = 0, 1, \dots\}$  el tiempo de espera hasta la ocurrencia del  $n$ -ésimo evento. Ahora, es necesario saber como se distribuye el tiempo de ocurrencia de los eventos dado que un número específico de éstos ha ocurrido. Para obtener dicha distribución es necesario hacer las siguientes especificaciones.

**Observación:**

Sean  $Y_i, i = 1, 2, \dots, n$ , variables aleatorias y sean  $Y_{(i)}, i = 1, 2, \dots, n$  los estadísticos de orden correspondientes a  $Y_1, Y_2, \dots, Y_n$  es decir,  $Y_{(k)}$  es el  $k$ -ésimo valor más pequeño entre  $Y_1, \dots, Y_n, k = 1, 2, \dots, n$ .

**Lema 3.2.** *Si  $Y_i, i = 1, 2, \dots, n$  son variables aleatorias independientes e idénticamente distribuidas con función de densidad  $f$ , entonces la función de densidad conjunta  $g$  de los estadísticos de orden  $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$  está dado por:*

$$g(y_1, y_2, \dots, y_n) = n! \prod_{i=1}^n f(y_i) \quad y_1 < y_2 < \dots < y_n.$$

**Demostración:** Lo anterior se cumple a partir de que:

- (i)  $(Y_{(1)}, Y_{(2)}, \dots, Y_{(n)})$  es igual a  $(y_1, y_2, \dots, y_n)$  si  $(Y_1, Y_2, \dots, Y_n)$  es igual a alguna de las  $n!$  permutaciones de  $(y_1, y_2, \dots, y_n)$ , y
- (ii) La función de densidad de  $(Y_1, Y_2, \dots, Y_n)$  evaluada en  $y_{i_1}, \dots, y_{i_n}$ , es  $\prod_{j=1}^n f(y_{i_j}) = \prod_{j=1}^n f(y_j)$  cuando  $i_1, i_2, \dots, i_n$  es una permutación de  $1, 2, \dots, n$ .

□

**Lema 3.3.** Sean  $Y_i, i = 1, \dots, n$  variables aleatorias distribuidas uniformemente sobre  $(0, t)$ , entonces la función de densidad de los estadísticos de orden  $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$  estará dada por:

$$g(y_1, y_2, \dots, y_n) = \frac{n!}{t^n} \quad 0 < y_1 < y_2 < \dots < y_n < t.$$

**Demostración:** Inmediata a partir del Lema 3.2.

□

Ahora, se puede enunciar el siguiente teorema:

**Teorema 3.1.** Si se cumple que  $N(t) = n$ , entonces, los  $n$  tiempos de llegadas  $S_1, \dots, S_n$  tienen la misma distribución como los estadísticos de orden correspondientes para  $n$  variables aleatorias independientes uniformemente distribuidos sobre el intervalo  $(0, t)$ .

**Demostración:** Ver, Ross (1996).

□

### 3.4. Proceso Poisson no homogéneo

En esta sección se considerará el proceso Poisson no homogéneo o también llamado no estacionario, el cual es obtenido al dejar que la tasa del proceso sea una función de  $t$ .

**Definición 3.8.** Sea  $N = \{N(t) : t \geq 0\}$  un proceso de conteo. Dicho proceso será un **proceso Poisson no homogéneo** en el tiempo con función de intensidad  $\lambda(t) > 0, t \geq 0$  si:

- (i)  $N(0) = 0$
- (ii)  $N = \{N(t), t \geq 0\}$  tiene incrementos independientes

$$(iii) P\{N(t+h) - N(t) = 1\} = \lambda(t)h + o(h)$$

$$(iv) P\{N(t+h) - N(t) \geq 2\} = o(h)$$

De lo anterior se puede enunciar y demostrar el siguiente teorema:

**Teorema 3.2.** Sea  $N = \{N(t), t \geq 0\}$  un proceso de Poisson no homogéneo y  $m(t) = \int_0^t \lambda(s)ds$ , entonces se cumple que  $N(t+s) - N(t)$  es distribuido Poisson con media  $m(t+s) - m(t)$ , es decir:

$$P\{N(t+s) - N(t) = n\} = e^{-[m(t+s)-m(t)]} \frac{[m(t+s) - m(t)]^n}{n!}, \quad n \geq 0.$$

**Demostración:** Defínase para todo  $n \geq 0, s, t \geq 0$ :

$$P_n(s) = P\{N(t+s) - N(t) = n\}. \quad (3.7)$$

Ahora, se usara inducción sobre  $n$  para demostrar el resultado. Así, tome  $n = 0$ , entonces:

$$\begin{aligned} P_0(s+h) &= P\{N(t+s+h) - N(t) = 0\} \\ &\stackrel{(1)}{=} P\{N(t+s) - N(t) = 0, N(t+s+h) - N(t+s) = 0\} \\ &= P\{N(t+s) - N(t) = 0\}P\{N(t+s+h) - N(t+s) = 0\} \\ &= P_0(s)P\{N(t+s+h) - N(t+s) = 0\} \\ &\stackrel{(2)}{=} P_0(s)[1 - \lambda(t+s)h - o(h)] \end{aligned}$$

Por lo tanto,

$$\frac{P_0(s+h) - P_0(s)}{h} = -\lambda(t+s)P_0(s) + \frac{o(h)}{h}.$$

Tomando el límite cuando  $h \rightarrow 0$  se tiene que

$$P'_0(s) = -\lambda(t+s)P_0(s),$$

es decir,

$$\frac{P'_0(s)}{P_0(s)} = -\lambda(t+s).$$

---

(1). Dado que  $N$  tiene incrementos independientes.

(2). Por (iii) y (iv) de la Definición 3.7.

Lo que implica que

$$\int_0^s \frac{P_0'(u)}{P_0(u)} du = \int_0^s -\lambda(t+u) du$$

y por lo tanto

$$\log P_0(s) - \log P_0(0) = - \int_0^s \lambda(t+u) du$$

Adicionalmente se tiene que

$$\int_0^s \lambda(t+u) du = \int_t^{t+s} \lambda(\mu) d\mu = \int_0^{t+s} \lambda(\mu) d\mu - \int_0^t \lambda(\mu) d\mu = m(t+s) - m(t).$$

Por tanto, dado que  $P_0(0) = P\{N(0) = 0\} = 1$  se tiene que

$$P_0(s) = e^{-[m(t+s)-m(t)]}.$$

Note que para  $n \geq 1$ . De (3.7) se puede escribir

$$\begin{aligned} P_n(s+h) &= P\{N(t+s+h) - N(t) = n\} \\ &= P\{n \text{ eventos en } (t, t+s), 0 \text{ eventos en } [t+s, t+s+h]\} \\ &+ P\{n-1 \text{ eventos en } (t, t+s), 1 \text{ evento en } [t+s, t+s+h]\} \\ &+ \sum_{k=2}^n P\{n-k \text{ eventos en } (t, t+s), k \text{ eventos en } [t+s, t+s+h]\}, \end{aligned}$$

pero a partir de las condiciones ii), iii) y de iv) de la Definición 3.8 se tiene:

$$\begin{aligned} P_n(s+h) &= P\{n \text{ eventos en } (t, t+s)\} P\{0 \text{ eventos en } [t+s, t+s+h]\} \\ &+ P\{n-1 \text{ eventos en } (t, t+s)\} P\{1 \text{ evento en } [t+s, t+s+h]\} \\ &+ o(h) \\ &= P_n(s)[1 - \lambda(t+s)h + o(h)] \\ &+ P_{n-1}(s)[\lambda(t+s)h + o(h)] \\ &+ o(h), \end{aligned}$$

lo cual implica que:

$$\begin{aligned} P_n(s+h) &= P_n(s)[1 - \lambda(t+s)h + o(h)] \\ &+ P_{n-1}(s)[\lambda(t+s)h + o(h)] \\ &+ o(h). \end{aligned}$$

Es decir,

$$\frac{P_n(s+h) - P_n(s)}{h} = \frac{-\lambda(t+s)hP_n(s)}{h} + \frac{\lambda(t+s)hP_{n-1}(s)}{h} + \frac{o(h)}{h}.$$

Tomando el límite cuando  $h \rightarrow 0$  se tiene que

$$P'_n(s) = -\lambda(t+s)P_n(s) + \lambda(t+s)P_{n-1}(s),$$

o equivalentemente

$$e^{\int_t^{t+s} \lambda(s) ds} [P'_n(s) + \lambda(t+s)P_n(s)] = P_{n-1}(s)\lambda(t+s)e^{\int_t^{t+s} \lambda(s) ds}$$

Por lo tanto

$$\frac{d}{ds} [e^{\int_t^{t+s} \lambda(u) du} P_n(s)] = \lambda(t+s)e^{\int_t^{t+s} \lambda(u) du} P_{n-1}(s). \quad (3.8)$$

Sea  $n = 1$ , entonces

$$\begin{aligned} \frac{d}{dr} [e^{\int_t^{t+r} \lambda(u) du} P_1(r)] &= \lambda(t+r)e^{\int_t^{t+r} \lambda(u) du} P_0(r) \\ &\stackrel{(1)}{=} \lambda(t+r)e^{\int_t^{t+r} \lambda(u) du} e^{-\int_t^{t+r} \lambda(u) du} \\ &= \lambda(t+r). \end{aligned}$$

Integrando con respecto a  $r$  en el intervalo  $[0, s]$

$$\begin{aligned} P_1(s) &= \left( \int_t^{t+s} \lambda(u) du + P_1(0) \right) e^{-\int_t^{t+s} \lambda(u) du} \\ &\stackrel{(2)}{=} \left[ \int_t^{t+s} \lambda(u) du \right] e^{-\int_t^{t+s} \lambda(u) du} \\ &= [m(t+s) - m(t)] e^{-[m(t+s) - m(t)]}. \end{aligned}$$

Por lo que queda demostrado para  $n = 1$ . Ahora para mostrar que se cumple para  $k$ , supóngase que se cumple para  $n = k - 1$ , es decir,

$$P\{N(t+s) - N(t) = k - 1\} = e^{-[m(t+s) - m(t)]} \frac{[m(t+s) - m(t)]^{k-1}}{(k-1)!}.$$

---

(1). Dado que  $P_0(r) = e^{-[m(t+r) - m(t)]}$ .

(2). Dado que  $P_1(0) = P\{N(0) = 1\} = 0$ .



De (3.8) se tiene

$$\begin{aligned} \frac{d}{dr} [e^{\int_t^{t+r} \lambda(u) du} P_k(r)] &= \lambda(t+r) e^{\int_t^{t+r} \lambda(u) du} P_{k-1}(r) \\ &\stackrel{(3)}{=} \lambda(t+r) \left( e^{\int_t^{t+r} \lambda(u) du} e^{-\int_t^{t+r} \lambda(u) du} \right) \frac{[\int_t^{t+r} \lambda(u) du]^{k-1}}{(k-1)!} \\ &= \lambda(t+r) \frac{[\int_t^{t+r} \lambda(u) du]^{k-1}}{(k-1)!}. \end{aligned}$$

Integrando con respecto a  $r$  en el intervalo  $[0, s]$

$$e^{\int_t^{t+s} \lambda(u) du} P_k(s) = \frac{[\int_t^{t+s} \lambda(u) du]^k}{(k)!} + P_k(0),$$

dado que  $P_k(0) = P\{N(0) = k\} = 0$  se tiene:

$$P_k(s) = \frac{[\int_t^{t+s} \lambda(u) du]^k}{(k)!} e^{-\int_t^{t+s} \lambda(u) du},$$

es decir,

$$P\{N(t+s) - N(t) = k\} = e^{-[m(t+s) - m(t)]} \frac{[m(t+s) - m(t)]^k}{k!}.$$

Por tanto, el resultado vale para todo  $n \geq 0$  y el teorema queda demostrado.  $\square$

La importancia del proceso Poisson no homogéneo reside en el hecho que no se requiere de la condición de incrementos estacionarios, permitiendo la posibilidad de que algunos eventos puedan ocurrir con mayor o menor intensidad en diferentes periodos de tiempo.

**Observación.** Cuando una función de intensidad  $\lambda(t)$  es acotada, se puede observar que un proceso no homogéneo es una muestra aleatoria de eventos que ocurren de acuerdo a un proceso Poisson homogéneo. Específicamente, sea  $\lambda$  tal que:

$$\lambda(t) \leq \lambda, \quad t \geq 0$$

y considérese un proceso Poisson con tasa  $\lambda$ . Supóngase, que un evento del proceso Poisson que ocurre al tiempo  $t$  es registrado con probabilidad  $\frac{\lambda(t)}{\lambda}$ .

---

(3). Por hipótesis de inducción.

Entonces este proceso de conteo de los eventos es un proceso Poisson no homogéneo con función de intensidad  $\lambda(t)$ . Este hecho sigue a partir de la definición de un proceso no homogéneo. Por ejemplo, i), ii) y iv) se cumplen para un proceso Poisson homogéneo, mientras que el axioma iii) se cumple a partir de lo siguiente:

$$\begin{aligned} P\{\text{un evento contado en } (t, t+h)\} &\stackrel{(1)}{=} P\{\text{un evento en } (t, t+h)\} \frac{\lambda(t)}{\lambda} + o(h) \\ &= \lambda h \frac{\lambda(t)}{\lambda} + o(h) \\ &= \lambda(t)h + o(h). \end{aligned}$$

---

(1). Por la Definición 3.5.

## Capítulo 4

# Proceso Bernoulli y Poisson en el estudio de problemas ambientales

Los capítulos anteriores detallaron el funcionamiento de los procesos Poisson y de las cadenas de Markov. En este capítulo se explicará cómo utilizar algunos procesos, en particular, el proceso de Bernoulli y el proceso de Poisson, en el análisis de la contaminación atmosférica, particularmente, para la norma internacional del ozono. Lo cual, puede consultarse más ampliamente en Ott (1995).

En la primera sección se explica el proceso Bernoulli, así como las ventajas que presenta la utilización de éste modelo en el estudio de contaminación atmosférica. En la segunda sección se analiza el proceso Poisson en comparación con el proceso Bernoulli, presentando ventajas y desventajas, ante la ausencia de algunas suposiciones básicas. La utilización de estos procesos para estudios de problemas de contaminación atmosférica fue propuesto por Javits (1980). En el Capítulo 5 se propondrá y se utilizará el proceso de Poisson como modelo, sin embargo el parámetro del proceso será estimado utilizando una formulación Bayesiana. Al contrario de lo que pasa con la formulación de Javits (1980), el parámetro del proceso de Poisson será una variable aleatoria.

Antes de comenzar a desarrollar el proceso Bernoulli, se observará que hay diversos fenómenos a nuestro alrededor que usualmente son modelados como un proceso Bernoulli, el cual, en general, involucra eventos discretos y la estimación del número total de ocurrencias del evento en estudio. Para

observar más detalladamente lo anterior, se tienen los siguientes ejemplos:

- 1.- Uno de los ejemplos más sencillos, es suponer que una moneda equilibrada (con probabilidad  $\frac{1}{2}$  de salir sol o águila) es lanzada 10 veces independientemente. ¿Cuál es la probabilidad de que cuatro caras aparezcan?
- 2.- Cuando en una área metropolitana se cumple con la norma de calidad del aire, el número esperado de excedentes por año (es decir, el número esperado de incumplimiento de la norma) es 1.0 o menor. Si la calidad del aire en una área determinada es tal que el número esperado de excedentes es exactamente 1.0, ¿cuál es la probabilidad de que:
  - a) más de un excedente ocurra en un año?
  - b) más de 3 excedentes ocurran en 3 años?
  - c) 6 o más excedentes ocurran en 3 años?
- 3.- Un campo es seleccionado para investigar si la contaminación por pesticida está presente en el suelo de cierta área. Cuando la contaminación está presente, experiencias pasadas, indican que muestras recogidas del área darán lecturas positivas con una probabilidad conocida, indicada por  $p \in (0, 1)$ . ¿Cuántas muestras de suelo serían colectadas en el área para obtener un 95 % de confianza de que una lectura positiva no es accidental?

## 4.1. Condiciones para el proceso Bernoulli

Los ejemplos anteriores sirven de base para definir lo que es el proceso Bernoulli, ya que se puede observar la existencia de ciertas condiciones que se cumplen para todos los ejemplos. Así, todos ellos comparten que:

- a) La probabilidad de ocurrencia de un evento de interés es la misma para todos los ensayos.
- b) La cantidad final de interés es el número total de salidas exitosas.
- c) Todos los ensayos son independientes.

Ahora bien se puede dar la siguiente definición para el proceso Bernoulli:

**Definición 4.1.** *Un proceso Bernoulli indica cuando un evento  $A$  ocurre en un ensayo, donde la probabilidad de ocurrencia del evento  $A$  en cada ensayo es la misma y a su vez cada ensayo es realizado de forma independiente.*

De esta forma, es posible definir una variable aleatoria  $K$ , que denote el número total de veces que el evento  $A$  ocurrió en una realización de  $N$  ensayos, cuyo espacio muestral estará dado por  $S = \{0, 1, 2, 3, \dots, N\}$ . Asimismo, la probabilidad de que el evento  $A$  ocurra sobre cada ensayo estará denotado por  $P\{A\} = p \in (0, 1)$  y la correspondiente probabilidad de la no ocurrencia de  $A$ ,  $A^c$ , será denotado por  $P\{A^c\} = 1 - p = q$ .

**Ejemplo:**

Supóngase que, en cierta población, no se sabe quien es un fumador y quien no lo es. Supóngase además, que, en un instante dado, la probabilidad de que una persona seleccionada al azar sea fumadora es  $P\{A\} = \frac{1}{9}$ , donde  $A$  denota el hábito de fumar. Ahora supóngase que en una oficina se encuentran dos de éstas personas, y que la preferencia de una persona por fumar no influye en la otra.

Defínase  $A_1$  como el evento de que la persona 1 sea fumadora y  $A_2$  como el evento de que la persona dos sea fumadora. Por hipótesis se tiene que  $P\{A_1\} = P\{A_2\} = \frac{1}{9}$ . Entonces, en un instante dado pueden darse alguna de las siguientes situaciones:

- (i) Nadie es fumador;
- (ii) Una de las dos personas es fumadora;
- (iii) Las dos personas son fumadoras.

Por tanto la variable aleatoria discreta  $K$  indicando el número total de personas fumadoras en un tiempo dado tiene espacio muestral  $S = \{0, 1, 2\}$ .

Así, se tiene lo siguiente:

- (i) Puesto que en el primer caso, ninguna de las dos personas sea fumadora, tenemos que calcular la probabilidad de que  $K = 0$ , la cual puede ser calculada como la intersección de dos eventos independientes  $A_1^c$  y  $A_2^c$ , es decir, que no sea fumadora ni la persona 1 ni la persona 2:

$$P\{K = 0\} = P\{A_1^c \cap A_2^c\} = qq = \left(\frac{8}{9}\right) \left(\frac{8}{9}\right) = 0.790.$$

(ii) El evento de que una persona sea fumadora, se cumplirá si y solamente si, alguna de las dos personas es fumadora, es decir, tenemos los siguientes dos casos:

a) La persona 1 podría ser fumadora mientras que la persona 2 no. Entonces tenemos lo siguiente:

$$P\{A_1 \cap A_2^c\} = pq = \left(\frac{1}{9}\right) \left(\frac{8}{9}\right) \approx 0.099.$$

b) La persona 2 podría ser fumadora mientras que la persona 1 no. En este caso tenemos lo siguiente:

$$P\{A_1^c \cap A_2\} = qp = \left(\frac{8}{9}\right) \left(\frac{1}{9}\right) \approx 0.099.$$

Ahora bien, el evento de que una persona sea fumadora es la combinación de (a) y (b). Así la probabilidad de que una persona sea fumadora está dado por la propiedad de la unión de dos eventos independientes, es decir,

$$P\{K = 1\} = pq + qp = 2pq = \left(\frac{1}{9}\right) \left(\frac{8}{9}\right) + \left(\frac{8}{9}\right) \left(\frac{1}{9}\right) \approx 0.198.$$

(iii) Finalmente, la probabilidad de que las dos personas sean fumadoras al mismo tiempo es calculado como la intersección de los eventos independientes  $A_1$  y  $A_2$ :

$$P\{A_1 \cap A_2\} = pp = \left(\frac{1}{9}\right)^2 \approx 0.012.$$

De esta forma, se pudo obtener las probabilidades correspondientes a los eventos que hacen parte del espacio muestral del experimento. Se puede observar que efectivamente la suma total de las probabilidades es uno, es decir, es una función de probabilidad:

$$P\{K = 0\} = q^2 = \frac{64}{81} = 0.790$$

$$P\{K = 1\} = 2pq = \frac{16}{81} = 0.198$$

$$P\{K = 2\} = p^2 = \frac{1}{81} = 0.012$$

Asimismo, se deja ver que la función está sesgada hacia la izquierda y que su moda se ubica en el origen ( $K=0$ ).

Note que la variable aleatoria que cuenta el número de sucesos en una sucesión de Bernoulli tiene distribución Binomial.

**Proposición 4.1.** Sean  $X$  e  $Y$  variables aleatorias con distribución Binomial de parámetros  $(n, p)$  y  $(m, p)$ , respectivamente. Entonces la variable  $X + Y$  se distribuye Binomial con parámetros  $n + m$  y  $p$ .

**Demostración:**

$$\begin{aligned} P\{X + Y = k\} &= \sum_{i=0}^k P\{X = i, Y = k - i\} \\ &= \sum_{i=0}^k P\{X = i\}P\{Y = k - i\} \\ &= \sum_{i=0}^k \binom{n}{i} p^i q^{n-i} \binom{m}{k-i} p^{k-i} q^{m-k+i} \end{aligned}$$

donde  $q = 1 - p$  y  $\binom{r}{j} = 0$  cuando  $j > r$ .

Por lo tanto

$$P\{X + Y = k\} = p^k q^{n+m-k} \sum_{i=0}^k \binom{n}{i} \binom{m}{k-i} = p^k q^{n+m-k} \binom{n+m}{k}$$

□

**Lema 4.1.** Sean  $X_1, X_2, \dots, X_n$  la sucesión de Bernoulli donde la probabilidad de suceso es  $p \in (0, 1)$ . Entonces,  $S_n = \sum_{k=1}^n X_k$  tiene distribución Binomial( $n, p$ ).

**Demostración:** Una variable aleatoria  $X$  que se distribuye Bernoulli con parámetro  $p$  es una variable aleatoria con distribución Binomial( $1, p$ ). Por lo que a partir de la Proposición 4.1 se cumple que  $\sum_{i=1}^n X_i$  se distribuye Binomial con parámetro  $n$  y  $p$ . □

### 4.1.1. Aplicación a modelos ambientales

La distribución Binomial también es apta para una variedad de problemas de control ambiental y tiene un uso especial en la interpretación de normas ambientales. El análisis de esta sección se basará en la Norma Internacional, en la cual se especifica, que la concentración máxima de contaminación en el aire no puede excederse más de una vez por año en una localidad monitoreada. Esto implica que el control de la calidad del aire se enfoque en el segundo valor más alto observado durante el año, en una localidad dada. Sin embargo, esta forma presenta algunas desventajas:

- a. No hay ninguna especificación para la ausencia de observaciones. Por ejemplo, si exactamente dos o tres lecturas fueron registradas en una localidad, la comparación de la segunda lectura más alta de éste lugar con la segunda lectura más alta de un sitio en el cual hubo un año lleno de observaciones, tendría poco sentido.
- b. Una segunda desventaja de dicha forma determinística es que no considera casos raros e inusuales. Por ejemplo, condiciones meteorológicas muy inusuales pueden ser causantes de altas concentraciones de contaminación en un año determinado, no obstante ésta localidad podría cumplir perfectamente con la norma en otros años.

Así, la norma para ozono tendrá una transición de un modelo determinístico a uno probabilístico, donde éste último se basa en el número esperado de días que la concentración del contaminante rebasa los 0.12 ppm.

**Definición 4.2.** *Sea  $X$  la variable que denota la concentración máxima de ozono observada durante un día y  $X_s$  como el valor numérico de la norma. Se dirá que hay un **excedente** de la norma si el evento  $\{X > X_s\}$  toma lugar, cabe señalar que en este caso particular  $X_s = 0.12$  ppm.*

Debido a que la norma ambiental de calidad del aire se cumple cuando el número esperado de concentraciones diarias por año calendario que exceden los 0.12 ppm es menor o igual a uno, es importante saber el número total de días durante un año calendario que presentaron una excedencia. Si  $K$  es la variable aleatoria que denota el número de excedentes observados durante algún periodo de tiempo. Entonces, se cumplirá con la norma si  $E[K] \leq 1.0$  y no así, cuando  $E[K] > 1$ .



Sin embargo, los valores esperados resultan ser difíciles de estimar en condiciones reales, por lo que se especifica que el número esperado de excedentes por año en una localidad será calculada promediando el número estimado de excedentes por cada año de observaciones durante los últimos 3 años calendario. No obstante, algunos años presentarán ausencia de observaciones, por lo que para ello se utilizará una tasa de excedencia definida por:

**Definición 4.3.** *Se define la tasa de excedente como la división del número de excedentes observados entre el número de observaciones disponibles para un periodo de tiempo.*

Dicha tasa es necesaria para los años que presentan ausencia de observaciones.

#### 4.1.2. Distribución de probabilidad del número de excedentes

Considérese el caso en el que el monitoreo de una localidad presenta como resultado que se cumple con la norma de calidad del aire en el límite de dicha norma, es decir, se asume que el número esperado de excedentes en ese sitio sea exactamente 1.0 o en su caso que  $E[K]=1.0$  (Nótese que es posible que en algunos años se obtengan 0 excedentes, y en otros años se obtengan más de un excedente, sin embargo, el número esperado de excedentes por año sera 1.0).

Así, es importante saber cual es la distribución de probabilidad de  $K$  que describe el número  $k$  de excedentes observados durante un año particular cuyo valor esperado sea  $E[K]=1.0$ .

A partir de lo expuesto en las secciones anteriores, se ha considerado que la realización de  $K$  sea descrita como el resultado de una sucesión Bernoulli (que registra incumplimiento o no de la norma), y la aplicación de la distribución Binomial en la interpretación de la norma para ozono, de acuerdo con Javits (1980). En este análisis, se consideran los excedentes diarios como eventos independientes, y se considera que los 365 días de cada año calendario son una serie de 365 ensayos independientes Bernoulli.

Sea  $K$  una variable aleatoria que describe el número total de excedentes encontrados en un periodo de un año, y sea  $M$  una variable aleatoria que representa el número total de excedentes contados en un periodo de 3 años. Considérese la situación en la cual la calidad del aire está en el punto crítico,

y que cumple apenas con la forma estadística de la norma de calidad del aire (pero sin exceder). En este punto crítico, el número esperado de excedentes para un periodo de un año y de tres años será como sigue:

$$E[K] = 1.0 \quad \text{y} \quad E[M] = 3E[K] = 3.0.$$

Bajo el supuesto de que el número de excedentes ocurrido durante este periodo de tiempo puede ser representado como una variable aleatoria con distribución Binomial. Se determinarán los parámetros  $n$  y  $p$  de la distribución para dichos periodos de tiempo.

**Periodo de 1 año:** Se sabe que  $n = 365$  días y que  $E[K] = 1$ , y tomando en cuenta que  $E[K] = np$  para una distribución binomial, se tiene que,  $1 = np = 365p$ , lo cual implica que  $p = \frac{1}{365}$ .

**Periodo de 3 años:** En este caso  $n = 1095$  días y  $E[M] = 3$ , lo cual implica que  $1095p = 3$  y por lo tanto  $p = \frac{3}{1095} = \frac{1}{365}$ .

#### Observación.

Lo anterior se explica a partir de que el número esperado de excedentes por año es uno y de que la tasa de excedente  $r = \frac{1}{365}$  es siempre la misma, sin importar, el periodo de tiempo.

Ahora bien, a partir de los parámetros  $n$  y  $p$ , la distribución de probabilidad para el número de excedentes en un año de 365 días estará dado por el modelo de probabilidad Binomial con parámetro  $n = 365$  y  $p = \frac{1}{365}$  o Binomial( $365, \frac{1}{365}$ ), donde  $q = 1 - p = 1 - \frac{1}{365} = 0.36365$ :

$$P\{K = k\} = \binom{365}{k} \left(\frac{1}{365}\right)^k \left(\frac{364}{365}\right)^{365-k} \quad k = 0, 1, 2, \dots, 365.$$

Los términos iniciales de la distribución pueden ser calculados fácilmente. Por ejemplo, la probabilidad de  $K = 0$  puede ser calculada como sigue:

$$P\{K = 0\} = \binom{365}{0} \left(\frac{1}{365}\right)^0 \left(\frac{364}{365}\right)^{365-0} = (1)(1) \left(\frac{364}{365}\right)^{365} = 0.3674.$$

Este primer resultado indica que, si la norma para ozono se cumple de tal manera que la  $E[K]=1.0$ , entonces la probabilidad de que en un año no haya excedentes es 0.3674. En otras palabras, en 100 años de observaciones se esperaría encontrar  $(0.367)(100)=36.7$  años, sin excedentes.

Por otro lado la probabilidad de que uno o más excedentes ocurran en un año determinado estará dado por:

$$P\{K \geq 1\} = 1 - P\{K = 0\} = 1 - 0.3674 = 0.633.$$

y

$$P\{K = 0\} + P\{K = 1\} = 0.3674 + 0.3684 = 0.7358.$$

La probabilidad de que más de un excedente ocurra en un año está dado por:

$$P\{K > 1\} = 1 - P\{K = 0\} - P\{K = 1\} = 1 - 0.3674 - 0.3684 = 1 - 0.7358 = 0.2642.$$

lo cual implica que en promedio el 26.4% de los años experimentarán dos o más excedentes, aunque, la norma para ozono se haya cumplido.

Se puede observar que para  $E[K]=1.0$ , por ejemplo, la probabilidad de más de 4 excedentes por año está dada por:

$$P\{K > 4\} = 1 - 0.9664 = 0.0036,$$

lo que implica que para un área que apenas cumple con la norma, habrá sólo 36 años en 10,000 en los cuales 5 o más excedentes ocurrirán.

Debido a que 5 o más excedentes son muy raros, la ocurrencia de cinco o más excedentes en una estación real de monitoreo durante un año particular sugiere fuertemente que la localidad no está cumpliendo con la parte estadística de la norma. Es decir, es probable que el número esperado de excedentes en dicho lugar sea mayor a uno o que  $E[K] > 1.0$ .

Ahora bien, la norma de calidad del aire para ozono puede verse como la constitución de dos partes:

- a) La parte estadística, la cual requiere que el número esperado de excedentes sea menor o igual a uno, es decir,  $E[K] \leq 1.0$ , y
- b) La parte determinística, que especifica que el número promedio de excedentes observados en algún periodo de 3 años sea menor o igual a tres.

Si  $k_1, k_2, k_3$  denotan el número de excedentes observados en el primer, segundo y tercer años de un periodo de 3 años, entonces la parte determinística de la norma requiere que el promedio de estos tres números, sea tal que

$$\frac{(k_1 + k_2 + k_3)}{3} \leq 1.$$

Sin embargo, esta condición puede cumplirse sólo si el número total de excedentes en un periodo de 3 años,  $m = k_1 + k_2 + k_3$  es menor o igual a tres. En el caso, en que  $M$  tenga un valor de 4 o más, la localidad no estará cumpliendo con la norma. Pero, ¿cuál es la probabilidad de que 4 o más excedentes ocurran en algún periodo de 3 años?. Usando el modelo de Bernoulli, la distribución del número total de excedentes en 3 años, o  $(3)(365)=1095$  días será Binomial  $(1095, \frac{1}{365})$ , y la variable aleatoria  $M$  tendría la siguiente función de densidad:

$$P\{M = k\} = \binom{1095}{k} \left(\frac{1}{365}\right)^k \left(\frac{364}{365}\right)^{1095-k}.$$

A partir de que es de gran interés el evento de que ocurran cuatro o más excedentes, se tienen los siguientes resultados:

$$P\{M = 0\} = \left(\frac{364}{365}\right)^{1095} = 0.04958,$$

$$P\{M = 1\} = \frac{1095}{1} \left(\frac{1}{365}\right)^1 \left(\frac{364}{365}\right)^{1094} = 0.14916,$$

$$P\{M = 2\} = \frac{(1095) \times (1094)}{1 \times 2} \left(\frac{1}{365}\right)^2 \left(\frac{364}{365}\right)^{1093} = 0.22414,$$

$$P\{M = 3\} = \frac{(1095) \times (1094) \times (1093)}{1 \times 2 \times 3} \left(\frac{1}{365}\right)^3 \left(\frac{364}{365}\right)^{1092} = 0.22435.$$

Así, la probabilidad de que ocurra el evento, 4 o más excedentes es calculado de estos resultados como sigue:

$$\begin{aligned} P\{M \leq 3\} &= P\{M = 0\} + P\{M = 1\} + P\{M = 2\} + P\{M = 3\} \\ &= 0.04958 + 0.14916 + 0.22414 + 0.22435 = 0.64723. \end{aligned}$$

$$P\{M \geq 4\} = P\{M > 3\} = 1 - P\{M \leq 3\} = 1 - 0.64723 = 0.35277.$$

Lo cual indica que si la norma es cumplida de tal forma que el número esperado de excedentes es 1.0, entonces 4 o más excedentes ocurrirán en algún periodo de 3 años con probabilidad de aproximadamente  $p=0.353$ . Es decir, a pesar de que la base estadística para la calidad estándar del aire ha sido cumplida, no obstante, niveles actuales no cumplirán con la norma en 35.5 % de los periodos de 3 años.

Para el caso, en el que más de 6 excedentes ocurren, se obtiene que  $F_M(6)^* = 0.96670$ , lo cual da una probabilidad de  $P\{M > 6\} = 1 - F_M(6) = 1 - 0.96670 = 0.0333$ . Es decir, si la parte estadística de la norma (1.0 excedente esperado por año), ha sido cumplida, entonces más de 6 excedencias ocurrirán en 3.3 % de los periodos de 3 años en promedio.

Lo anterior sugiere que la frecuencia con la que cierta localidad estará fuera de la norma puede ser muy grande, aunque se cumpla exactamente con la norma. Consecuentemente, se concluye que los componentes estadísticos y de cumplimiento de la norma no son consistentes.

Puesto que la distribución Poisson es una buena aproximación para la distribución binomial, en la sección subsecuente se analizará el proceso Poisson ya que se está utilizando una  $n = 365$  suficientemente grande y una  $p = \frac{1}{365}$  pequeña para poder hacer una buena aproximación.

## 4.2. Proceso Poisson

Para esta sección se considerará un proceso que evoluciona continuamente en el tiempo, a saber el proceso Poisson (desarrollado en el Capítulo 2 del presente trabajo). De tal manera que la variable aleatoria de interés en este caso, será el número de eventos ocurridos sobre algún periodo de tiempo, donde cada ocurrencia de un evento, se considerará como una llegada. Es decir, nos interesa el número total de llegadas ocurridas durante algún periodo de observación, donde el número de llegadas no se ve afectado por algún otro periodo de observación anterior. Adicionalmente, se considera que la distribución del número de llegadas será independiente del tiempo al cual el periodo de observación empezó. Es claro, que se trata de un proceso Poisson, ya que:

**Definición 4.4.** *Un proceso Poisson describe el número total de eventos independientes que ocurren durante un periodo específico de observación en el cual la tasa de llegada es constante.*

De acuerdo a lo expuesto en el Capítulo 2, se sabe que el modelo probabilístico que caracteriza a un proceso Poisson es la distribución Poisson, en el cual, se considera  $\lambda$  una constante positiva que representa la tasa de llegada

---

\*La **función de distribución** acumulada  $F_X(x)$  de una variable aleatoria  $X$  que toma valores en un conjunto  $S \subset \mathbb{R}$ , se define como  $F_X(x) = P(X \leq x), x \in S$ .

del proceso y  $N(t + s) - N(t)$  (donde  $N$  es definida de la misma manera que en el Capítulo 2) es el número de llegadas en el periodo de observación de interés, es decir:

$$P\{N(t + s) - N(t) = n\} = \frac{e^{-\lambda s}(\lambda s)^n}{n!}, \quad n = 0, 1, \dots$$

#### 4.2.1. Empleo en problemas ambientales

Al igual que antes interesa el cálculo de las probabilidades asociadas con el número de veces que un límite de concentración es excedido, o en su caso, que una norma ambiental no sea cumplida.

Supóngase que se tiene una tasa promedio de excedencias de  $r$  excedencias por unidad de tiempo. Así, para poder aplicar el proceso Poisson, se considerará que la tasa de llegada del proceso es igual a la tasa promedio de excedencias, de tal manera que:  $\lambda = r$ . Si  $K_t$  es la variable aleatoria que denota el número de excedentes en el periodo de tiempo de 0 a  $t$ , entonces se tiene

$$P\{K_t = k\} = \frac{e^{-rt}(rt)^k}{k!} \quad (4.1)$$

la cual es la función de densidad de la distribución Poisson, donde  $t$  expresará la duración del periodo de observación.

**Ejemplo:** Supóngase que agua contaminada es descargada en un manantial continuamente y que el número de violaciones por unidad de tiempo de la norma de calidad del agua puede ser tratada como constante, es decir, se puede considerar análoga a una tasa de llegada constante.

#### 4.2.2. Distribución de probabilidad para el número de excedentes

En la Sección 4.1 el número de excedentes diario de la norma internacional de ozono fue modelado con el proceso Bernoulli. En seguida, se considerará el mismo problema pero tomando como modelo un proceso Poisson.

La variable aleatoria  $K_t$  denotará el número de excedentes ocurridos durante un periodo de un año y la variable aleatoria  $M_t$  denotará el número de excedentes en un periodo de tres años. Así, se asume la situación crítica, en la cual la norma es apenas cumplida, de tal forma que no se rebasa, es decir, que

$$E[K_t] = 1.0 \quad y \quad E[M_t] = 3E[K_t] = 3.0.$$

Así, no tomando en cuenta los años bisiestos, se puede calcular la tasa de excedencia como sigue, con  $t = 365$ .

$$E[K_t] = 1.0 = rt = (r)(365 \text{ días}).$$

Si se resuelve la ecuación para obtener el valor para  $r$ , se obtiene que  $r = \frac{1}{365}$  excedentes por día y al sustituir dicho valor en (4.1), se obtiene

$$P\{K_t = k\} = \frac{e^{-\frac{t}{365}} \left(\frac{t}{365}\right)^k}{k!} \quad k = 0, 1, 2, \dots$$

Así,

- i) La distribución de probabilidad para un periodo de un año se obtiene al sustituir  $t = 365$  días en la relación anterior, lo cual da la distribución Poisson con tasa 1.
- ii) Para calcular la distribución de probabilidad para periodos de 3 años, se tendría que sustituir

$$t = (3 \text{ años}) = (365 \text{ días por año}) = 1095 \text{ días}.$$

en dicha ecuación, obteniendo la distribución Poisson con tasa 3.

Después de haber obtenido el modelo para cada periodo de tiempo, se puede hacer una comparación con el modelo Binomial.

Por ejemplo, con el modelo Poisson, la probabilidad de que haya un excedente o menos está dado por,

$$F_{K_t}(1) = P\{K_t \leq 1\} = \frac{e^{-1}(1)}{0!} + \frac{e^{-1}(1)^k}{1!} = 0.73576.$$

donde  $F_{K_t}(1)$  es la función de distribución acumulada de la variable aleatoria  $K_t$  y cuyo valor es similar al obtenido con el modelo Binomial. Es decir, la probabilidad del evento más de un excedente en un periodo de un año es la misma para ambos modelos:

$$P\{K_t > 1\} = 1 - P\{K_t \leq 1\} = 1 - F_{K_t}(1) = 1 - 0.73567 = 0.26424.$$

Por otro lado, la probabilidad de 0 excedencias en un periodo de 3 años está dado por

$$P\{M_t = 0\} = \frac{e^{-3}(3)^0}{0!} = 0.04979$$

aplicando la distribución Poisson comparado con  $P\{M_t = 0\} = 0.04958$  al utilizar la distribución Binomial. Si estos dos valores son redondeados, ambos resultados serían iguales a 0.050.

Asimismo, si se considera el evento más de tres excedentes en un periodo de 3 años entonces,

$$P\{M_t > 3\} = 1 - P\{M_t \leq 3\} = 1 - F_{M_t}(3) = 1 - 0.64723 = 0.35277$$

donde  $F_{M_t}(3)$  es la función de distribución acumulada de  $M_t$ , por lo que, se obtiene aproximadamente el mismo resultado obtenido por el modelo Binomial.

Finalmente, si se redondean los resultados del modelo Poisson a 3 dígitos, se puede observar que el 26.4 % de los periodos de un año tendrían más de un excedente y 35.3 % de los periodos de 3 años tendrían más de 3 excedentes, lo cual es igual al modelo Binomial bajo el supuesto de que  $E[K]=1.0$ . Los cálculos anteriores dejan claro la buena aproximación del modelo Poisson al Binomial para el caso en que  $p$  es pequeño y  $n$  es suficientemente grande. Dado que  $p = \frac{1}{365}$  es un valor muy pequeño, se está modelando un evento raro, lo cual implica que el modelo Poisson, es un buen modelo para dicho tipo de eventos. Dado que ambos modelos son buenos para resolver dichos problemas, se debe tomar una decisión de cual modelo utilizar a partir de la perspectiva con la que se aborde el problema en cuestión, puesto que para el proceso Poisson el problema es visto como la constitución de eventos independientes (excedentes) en el tiempo, que ocurren con una tasa de llegada constante  $\lambda = r = \frac{1}{365}$  excedentes por día y usualmente contados sobre periodos de observación de 1 o 3 años. Mientras que en el modelo Binomial, el problema es visto como la composición de 365 ensayos independientes por año, con probabilidad  $p = \frac{1}{365}$  de obtener un excedente en un ensayo. Luego, ambos modelos arrojan resultados aproximados para dichos parámetros. Cabe señalar que el modelo Poisson es aproximado por el modelo binomial.



# Capítulo 5

## Distribución del número de excedentes para ozono

En la presente sección se desarrolla la aplicación del trabajo. En lugar de utilizar los modelos Binomial y Poisson con los parámetros dados en el Capítulo 4, se utilizará el punto de vista Bayesiano para estimarlo. Se presentarán las distribuciones a posteriori del parámetro de la distribución Poisson que corresponde a la distribución de la variable aleatoria que registra el número de días en los cuales la concentración  $i$  de ozono sobrepasa la norma. El valor  $i$  representa los puntos IMECA, de tal manera que  $i$  tomará valores en  $\{100, 150, 200, 240\}$ . Estos valores son seleccionados por las siguientes razones: el valor 100 es el valor de la norma internacional, 150 es un valor intermedio entre 100 y 200, el valor 240 es el utilizado por las autoridades ambientales de la Ciudad de México para declarar una emergencia ambiental.

### 5.1. Algunos resultados de la Inferencia Bayesiana

Uno de los objetivos de los análisis estadísticos es encontrar el modelo más adecuado para un conjunto de datos. En dicho modelo se trata de incorporar la información más importante de los datos mostrando los rasgos significativos de los mismos. Sin embargo, existen diversos métodos para llevar a cabo dicho análisis. Uno de ellos lo constituye la inferencia Bayesiana. En esta sección se definirán y utilizarán los conceptos de función de verosimilitud, distribución a priori, a posteriori, la familia conjugada y en particular se

especifican dichos conceptos en el caso de la distribución Gamma y Poisson.

Detalles adicionales a los presentados aquí, pueden ser encontrados en los siguientes textos: Leonard y Hsu (1999) y Lee (1997).

En general, al utilizar un modelo para describir un experimento aleatorio se desea encontrar la distribución de los parámetros para este modelo que describe un conjunto de datos obtenidos a partir de una serie de repeticiones del experimento aleatorio. Así, se indicará por  $\theta$  el vector con los parámetros del modelo y  $\mathbf{X}$  representará el vector con los datos descritos por el modelo. El objetivo es obtener  $P(\theta | \mathbf{X})$ . Sin embargo, del modelo se puede obtener la probabilidad de obtener cierto conjunto de datos, si se conocen los parámetros, es decir,  $P(\mathbf{X} | \theta)$ . Se verá más adelante cómo utilizar dicha información para obtener  $P(\theta | \mathbf{X})$ . Para lograr el objetivo se empezará con algunas definiciones:

**Definición 5.1.** Se llamará **distribución a priori** a  $P_{\text{priori}}(\theta)$ , si ésta constituye la distribución de probabilidad para el parámetro  $\theta \in \Theta$  (espacio de los parámetros), sin tomar en cuenta la información proporcionada por los datos.

**Definición 5.2.** Sea  $f(\mathbf{X} | \theta)$  que denota la función de densidad conjunta o función de masa de probabilidad de una muestra  $\mathbf{X} = (X_1, \dots, X_n)$ , bajo un modelo con parámetro  $\theta$ , es decir, la función de densidad de  $X$  dado  $\theta$ . La función de  $\theta$  definida por

$$\mathcal{L}(\mathbf{X} | \theta) \propto f(\mathbf{X} | \theta).$$

es llamada la **función de verosimilitud**, donde el signo  $\propto$  denota proporcionalidad. La función de verosimilitud normalizada se define como:

$$\frac{\mathcal{L}(X | \theta)}{\int_{\theta} \mathcal{L}(X | \theta)}.$$

**Definición 5.3.** Sea  $P_{\text{priori}}(\theta)$  la distribución a priori del parámetro  $\theta$  y sea  $\mathcal{L}(\mathbf{X} | \theta)$  para  $X \in S$  y  $\theta \in \Theta$  la función de verosimilitud de  $\mathbf{X}$ , entonces la **distribución a posteriori** de  $\theta$  es:

$$P_{\text{posteriori}}(\theta | \mathbf{X}) \propto \mathcal{L}(\mathbf{X} | \theta)P_{\text{priori}}(\theta) \quad (\theta \in \Theta). \quad (5.1)$$

**Observación.** La Definición 5.2 es una consecuencia del Teorema de Bayes (Ver apéndice A), que utilizado en el caso de las distribuciones a priori y a posteriori, produce

$$P_{\text{posteriori}}(\theta | y) = \frac{P_{\text{priori}}(\theta)f(y | \theta)}{f(y)} \quad (\theta \in \Theta)$$

donde

$$f(y) = \int_{\Theta} f(y, \theta) d\theta = \int_{\Theta} P_{\text{priori}}(\theta) f(y | \theta) d\theta.$$

La constante de proporcionalidad puede depender de  $\mathbf{X}$  pero no de  $\theta$ .

La ecuación (5.1) se puede ver como:

$$\text{DENSIDAD A POSTERIORI} \propto \text{FUNCIÓN DE VEROSIMILITUD} \times \text{DENSIDAD A PRIORI}.$$

De tal manera, que la densidad a posteriori resume la información total, después de haberse observado los datos y proveer las bases para la inferencia a posteriori de  $\theta$ .

**Definición 5.4.** Sea  $\mathcal{L}(\mathbf{X} | \theta)$  una función de verosimilitud. Una clase  $\Pi$  de distribuciones a priori se dice que forma **una familia conjugada** con  $\mathcal{L}(\mathbf{X} | \theta)$  si la densidad a posteriori

$$P_{\text{posteriori}}(\theta | \mathbf{X}) \propto \mathcal{L}(\mathbf{X} | \theta) P_{\text{priori}}(\theta)$$

está en la clase  $\Pi$  para todo  $X$ , siempre que la densidad a priori esté en  $\Pi$ . Es decir, *priori* y *posteriori* están en la misma familia de distribuciones.

**Teorema 5.1.** Sea  $\mathbf{X}$  que se distribuye Poisson con parámetro  $\lambda$ . Supóngase que la distribución a priori de  $\theta = \lambda$  es Gamma con parámetros  $\alpha$  y  $\beta$ , (media  $\mu = \frac{\alpha}{\beta}$  y varianza  $\sigma^2 = \frac{\alpha}{\beta^2}$ ). Entonces la distribución a posteriori de  $\theta = \lambda$  dado  $X = k$  es una distribución Gamma con parámetros  $\alpha + k$  y  $\beta + 1$ , y

$$E(\theta | X = k) = \frac{\alpha + k}{\beta + 1},$$

$$\text{Var}(\theta | X = k) = \frac{\alpha + k}{(\beta + 1)^2}.$$

**Demostración:** Por hipótesis se tiene que:  $\mathcal{L}(X | \theta = \lambda)$  es proporcional a una distribución  $Poisson(\lambda)$  y que  $P_{\text{priori}}(\theta)$  es una distribución  $Gamma(\alpha, \beta)$ . Por la Definición 5.3 se tiene que

$$P_{\text{posteriori}}(\theta | X) \propto \mathcal{L}(X | \theta) P_{\text{priori}}(\theta).$$

Entonces,

$$\begin{aligned} P_{\text{posteriori}}(\theta = \lambda | X = k) &\propto e^{-\lambda} \lambda^k \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \\ &\propto \beta^\alpha \lambda^{k+\alpha-1} e^{-\beta\lambda-\lambda} \\ &\propto \frac{\beta^\alpha \lambda^{\alpha+k-1} e^{-(\beta+1)\lambda}}{\Gamma(\alpha + k)}. \end{aligned}$$

Por lo que  $P_{posteriori}(\theta)$  es una distribución Gamma( $\alpha + k, \beta + 1$ ) y por tanto (Ver Apéndice D)  $\mu = \frac{\alpha+k}{\beta+1}$  y  $\sigma^2 = \frac{\alpha+k}{(\beta+1)^2}$ .  $\square$

Se verá en la siguiente sección como utilizar este resultado para predecir el número de violaciones de la norma de ozono en un año dado.

## 5.2. Distribuciones del parámetro $\lambda$

Los datos utilizados aquí pueden ser obtenidos en [www.sma.df.gob/simat](http://www.sma.df.gob/simat) y consisten de los máximos diarios de ozono de la Zona Metropolitana del Valle de México durante los años de 1997 hasta 2003.

Utilizando los registros de concentraciones diarias observadas para los años 1997-2003, se hace un conteo de los días que rebasan los 100, 150, 200 y 240 puntos IMECA, para cada año. En la Tabla 5.1:

Puntos IMECA	PPM	1997	1998	1999	2000	2001	2002	2003
> 100	>.11	318	315	295	317	290	291	273
> 150	>.17	208	209	183	172	135	105	55
> 200	>.23	50	56	30	19	12	9	2
> 240	>.278	7	9	5	2	0	1	0

Tabla 5.1: Número de días que rebasan los 100, 150, 200 y 240 puntos IMECA.

Defínase  $M(i)$  como la variable aleatoria que representa el número de días con medición estrictamente arriba de la concentración  $i$ ,  $i = \{100, 150, 200, 240\}$ . De tal manera que es posible obtener el valor esperado de días arriba de la concentración  $i$ , para toda  $i$ . Se ha considerado dos casos: el primero son las mediciones de los años 1997 hasta 1999 (año en que se implementarán nuevas medidas del control de la contaminación por ozono); el segundo son las mediciones de los años 2000 hasta 2003. Los resultados obtenidos para el periodo 1997-1999 son:

$$\begin{array}{ll}
E[M(100)] = 309.33 & \text{Var}[M(100)] = 156.33 \\
E[M(150)] = 200 & \text{Var}[M(150)] = 217 \\
E[M(200)] = 45.33 & \text{Var}[M(200)] = 185.33 \\
E[M(240)] = 7 & \text{Var}[M(240)] = 4
\end{array}$$

y para el periodo de 2000 a 2003 son:

$$\begin{array}{ll}
E[M(100)] = 292.75 & \text{Var}[M(100)] = 329.58 \\
E[M(150)] = 116.75 & \text{Var}[M(150)] = 2445.58 \\
E[M(200)] = 10.5 & \text{Var}[M(200)] = 49.67 \\
E[M(240)] = 0.75 & \text{Var}[M(240)] = 0.92
\end{array}$$

Se utilizará el modelo Poisson para registrar el número de días que la concentración de ozono rebasa  $i$ . La diferencia con el modelo presentado en el Capítulo 4 es que aquí se estimará el valor  $\lambda$  de la distribución Poisson utilizando los datos reales de la red de monitoreo de la Ciudad de México. De esta forma, suponga que la media de la distribución Poisson tiene distribución a priori una Gamma( $\alpha, \beta$ ). Se sabe que una variable aleatoria Gamma( $\alpha, \beta$ ) tiene una  $\sigma^2 = \frac{\alpha}{\beta^2}$  y  $\mu = \frac{\alpha}{\beta}$ , luego a partir de los valores obtenidos para la media y la varianza de  $M(i)$ ,  $i = \{100, 150, 200, 240\}$  se puede obtener que parámetros  $\alpha$  y  $\beta$  se ajustan mejor a los valores de la media y varianza  $\lambda$ .

De esta forma, si  $E(M(i))_{1997-1999}$  indica el valor de la media de la variable aleatoria  $M(i)$  durante los años 1997 hasta 1999, entonces se tiene que

- para  $E[M(100)]_{1997-1999}$  lo más adecuado es tomar Gamma(612.07, 1.98);
- para  $E[M(150)]_{1997-1999}$  tomar Gamma(184.33, 0.92);
- para  $E[M(200)]_{1997-1999}$  tomar Gamma(11.09, 0.24);
- para  $E[M(240)]_{1997-1999}$  tomar Gamma(12.25, 1.75).

Para los años de 2000 al 2003 se tiene que

- para  $E[M(100)]_{2000-2003}$  lo más adecuado es tomar Gamma(260.03, 0.89);

- para  $E[M(150)]_{2000-2003}$  tomar  $\text{Gamma}(5.57, 0.05)$ ;
- para  $E[M(200)]_{2000-2003}$  tomar  $\text{Gamma}(2.22, 0.21)$ ;
- para  $E[M(240)]_{2000-2003}$  tomar  $\text{Gamma}(0.61, 0.82)$ .

Ahora, se obtiene que las distribuciones a posteriori, se distribuyen  $\text{Gamma}(\alpha + k, \beta + 1)$  a partir del Teorema 5.1, donde  $k$  representa el valor que se rebasó en cada año desde 1997 hasta el 2003. Por lo que se tiene lo siguiente:

1.  $E[M(100)]$

- a) Para el año 1997 se distribuye  $\text{Gamma}(930.07, 2.98)$  con  $\mu=312.24$ ,  $\sigma^2=104.83$  y moda=311.91.
- b) Para el año 1998 se distribuye  $\text{Gamma}(927.07, 2.98)$  con  $\mu=311.24$ ,  $\sigma^2=104.49$  y moda=310.9.
- c) Para el año 1999 se distribuye  $\text{Gamma}(907.07, 2.98)$  con  $\mu=304.52$ ,  $\sigma^2=102.23$  y moda=304.19.
- d) Para el año 2000 se distribuye  $\text{Gamma}(577.03, 1.89)$  con  $\mu=305.59$ ,  $\sigma^2=161.84$  y moda=305.06.
- e) Para el año 2001 se distribuye  $\text{Gamma}(550.03, 1.89)$  con  $\mu=291.29$ ,  $\sigma^2=154.27$  y moda=290.76.
- f) Para el año 2002 se distribuye  $\text{Gamma}(551.03, 1.89)$  con  $\mu=291.82$ ,  $\sigma^2=154.55$  y moda=291.29.
- g) Para el año 2003 se distribuye  $\text{Gamma}(533.03, 1.89)$  con  $\mu=282.29$ ,  $\sigma^2=149.5$  y moda=281.76.

2.  $E[M(150)]$

- a) Para el año 1997 se distribuye  $\text{Gamma}(392.33, 1.92)$  con  $\mu=204.16$ ,  $\sigma^2=106.24$  y moda=203.64.
- b) Para el año 1998 se distribuye  $\text{Gamma}(393.33, 1.92)$  con  $\mu=204.68$ ,  $\sigma^2=106.51$  y moda=204.16.
- c) Para el año 1999 se distribuye  $\text{Gamma}(367.33, 1.92)$  con  $\mu=191.15$ ,  $\sigma^2=99.47$  y moda=190.63.
- d) Para el año 2000 se distribuye  $\text{Gamma}(177.57, 1.05)$  con  $\mu=169.48$ ,  $\sigma^2=161.76$  y moda=168.53.

- e) Para el año 2001 se distribuye Gamma(140.57, 1.05) con  $\mu=134.17$ ,  $\sigma^2=128.06$  y moda=133.21.
- f) Para el año 2002 se distribuye Gamma(110.57, 1.05) con  $\mu=105.54$ ,  $\sigma^2=100.73$  y moda=104.58.
- g) Para el año 2003 se distribuye Gamma(60.57, 1.05) con  $\mu=57.81$ ,  $\sigma^2=55.18$  y moda=56.86.

### 3. E[M(200)]

- a) Para el año 1997 se distribuye Gamma(61.09, 1.24) con  $\mu=49.08$ ,  $\sigma^2=39.44$  y moda=48.28.
- b) Para el año 1998 se distribuye Gamma(67.09, 1.24) con  $\mu=53.9$ ,  $\sigma^2=43.31$  y moda=53.1.
- c) Para el año 1999 se distribuye Gamma(41.09, 1.24) con  $\mu=33.01$ ,  $\sigma^2=26.53$  y moda=32.21.
- d) Para el año 2000 se distribuye Gamma(21.22, 1.21) con  $\mu=17.52$ ,  $\sigma^2=14.46$  y moda=16.69.
- e) Para el año 2001 se distribuye Gamma(14.22, 1.21) con  $\mu=11.74$ ,  $\sigma^2=9.69$  y moda=10.91.
- f) Para el año 2002 se distribuye Gamma(11.22, 1.21) con  $\mu=9.26$ ,  $\sigma^2=7.65$  y moda=8.44.
- g) Para el año 2003 se distribuye Gamma(4.22, 1.21) con  $\mu=3.48$ ,  $\sigma^2=2.88$  y moda=2.66.

### 4. E[M(240)]

- a) Para el año 1997 se distribuye Gamma(19.25, 2.75) con  $\mu=7$ ,  $\sigma^2=2.55$  y moda=6.64.
- b) Para el año 1998 se distribuye Gamma(21.25, 2.75) con  $\mu=7.73$ ,  $\sigma^2=2.81$  y moda=7.36.
- c) Para el año 1999 se distribuye Gamma(17.25, 2.75) con  $\mu=6.27$ ,  $\sigma^2=2.28$  y moda=5.91.
- d) Para el año 2000 se distribuye Gamma(2.61, 1.82) con  $\mu=1.44$ ,  $\sigma^2=0.79$  y moda=0.89.
- e) Para el año 2001 se distribuye Gamma(0.61, 1.82) con  $\mu=0.34$ ,  $\sigma^2=0.19$ .

- f) Para el año 2002 se distribuye Gamma(1.61, 1.82) con  $\mu=0.89$ ,  $\sigma^2=0.49$  y moda=0.34.
- g) Para el año 2003 se distribuye Gamma(0.61, 1.82) con  $\mu=0.34$ ,  $\sigma^2=0.19$ .

**Observación.** Para una variable aleatoria Gamma, la moda está definida para valores  $\alpha \geq 1$ .

Con las distribuciones posteriores obtenidas, se tiene un ajuste más preciso del parámetro  $\lambda$ , como lo muestran las gráficas de dichas distribuciones en el Apéndice F.

Considerando que la moda de la gamma posteriori es un valor representativo de  $\lambda$ , se tomará dicho valor como el parámetro de la distribución Poisson que registra el número de días que la concentración de ozono rebasa  $i$ , en caso que no se pueda tomar la moda se tomará la media de la distribución a posteriori. Para finalmente obtener:

#### 1. M(100)

- a) Para el año 1997 se distribuye Poisson(311.91). Entonces  $P(M(100)>300) = 0.7390$  y  $P(M(100)>320) = 0.3108$ .
- b) Para el año 1998 se distribuye Poisson(310.9). Entonces  $P(M(100)>300) = 0.7202$  y  $P(M(100)>320) = 0.2907$ .
- c) Para el año 1999 se distribuye Poisson(304.19). Entonces  $P(M(100)>300) = 0.5802$  y  $P(M(100)>315) = 0.2565$ .
- d) Para el año 2000 se distribuye Poisson(305.06). Entonces  $P(M(100)>300) = 0.5995$  y  $P(M(100)>325) = 0.1216$ .
- e) Para el año 2001 se distribuye Poisson(290.76). Entonces  $P(M(100)>280) = 0.7242$  y  $P(M(100)>300) = 0.2817$ .
- f) Para el año 2002 se distribuye Poisson(291.29). Entonces  $P(M(100)>280) = 0.7344$  y  $P(M(100)>300) = 0.2923$ .
- g) Para el año 2003 se distribuye Poisson(281.76). Entonces  $P(M(100)>270) = 0.7470$  y  $P(M(100)>300) = 0.1325$ .

#### 2. M(150)

- a) Para el año 1997 se distribuye Poisson(203.64). Entonces  $P(M(150)>195) = 0.7131$  y  $P(M(150)>215) = 0.2019$ .



- b) Para el año 1998 se distribuye Poisson(204.16). Entonces  $P(M(150) > 195) = 0.7253$  y  $P(M(150) > 215) = 0.2124$ .
- c) Para el año 1999 se distribuye Poisson(190.63). Entonces  $P(M(150) > 180) = 0.7667$  y  $P(M(150) > 200) = 0.2355$ .
- d) Para el año 2000 se distribuye Poisson(168.53). Entonces  $P(M(150) > 155) = 0.8423$  y  $P(M(150) > 180) = 0.1777$ .
- e) Para el año 2001 se distribuye Poisson(133.21). Entonces  $P(M(150) > 120) = 0.8653$  y  $P(M(150) > 140) = 0.2609$ .
- f) Para el año 2002 se distribuye Poisson(104.58). Entonces  $P(M(150) > 90) = 0.9181$  y  $P(M(150) > 115) = 0.1432$ .
- g) Para el año 2003 se distribuye Poisson(56.86). Entonces  $P(M(150) > 45) = 0.9380$  y  $P(M(150) > 65) = 0.1271$ .

### 3. M(200)

- a) Para el año 1997 se distribuye Poisson(48.28). Entonces  $P(M(200) > 40) = 0.8701$  y  $P(M(200) > 60) = 0.0432$ .
- b) Para el año 1998 se distribuye Poisson(53.1). Entonces  $P(M(200) > 45) = 0.8522$  y  $P(M(200) > 65) = 0.0481$ .
- c) Para el año 1999 se distribuye Poisson(32.21). Entonces  $P(M(200) > 20) = 0.9853$  y  $P(M(200) > 40) = 0.0760$ .
- d) Para el año 2000 se distribuye Poisson(16.69). Entonces  $P(M(200) > 5) = 0.9981$  y  $P(M(200) > 25) = 0.0208$ .
- e) Para el año 2001 se distribuye Poisson(10.91). Entonces  $P(M(200) > 1) = 0.9997$  y  $P(M(200) > 20) = 0.0042$ .
- f) Para el año 2002 se distribuye Poisson(8.44). Entonces  $P(M(200) > 1) = 0.9979$  y  $P(M(200) > 15) = 0.0130$ .
- g) Para el año 2003 se distribuye Poisson(2.66). Entonces  $P(M(200) > 1) = 0.7439$ ,  $P(M(200) > 5) = 0.0535$  y  $P(M(200) > 10) = 0.0001$ .

### 4. M(240)

- a) Para el año 1997 se distribuye Poisson(6.64). Entonces  $P(M(240) > 1) = 0.9900$ ,  $P(M(240) > 5) = 0.6509$  y  $P(M(240) > 10) = 0.0749$ .
- b) Para el año 1998 se distribuye Poisson(7.36). Entonces  $P(M(240) > 1) = 0.9946$ ,  $P(M(240) > 5) = 0.7428$  y  $P(M(240) > 10) = 0.1260$ .

- c) Para el año 1999 se distribuye Poisson(5.91). Entonces  
 $P(M(240) > 1) = 0.9812$ ,  $P(M(240) > 5) = 0.5397$  y  $P(M(240) > 10) = 0.0390$ .
- d) Para el año 2000 se distribuye Poisson(0.89). Entonces  
 $P(M(240) > 1) = 0.2238$ ,  $P(M(240) > 5) = 0.0003$  y  $P(M(240) > 10) = 0$ .
- e) Para el año 2001 se distribuye Poisson(0.34). Entonces  
 $P(M(240) > 1) = 0.0462$ ,  $P(M(240) > 5) = 0$ . y  $P(M(240) > 10) = 0$ .
- f) Para el año 2002 se distribuye Poisson(0.34). Entonces  
 $P(M(240) > 1) = 0.0462$ ,  $P(M(240) > 5) = 0$ . y  $P(M(240) > 10) = 0$ .
- g) Para el año 2003 se distribuye Poisson(0.34). Entonces  
 $P(M(240) > 1) = 0.0462$ ,  $P(M(240) > 5) = 0$ . y  $P(M(240) > 10) = 0$ .

Usando el programa Mathematica, se pudo obtener las probabilidades asociadas a las distribuciones dadas arriba, de tal manera, que para  $M(100)$ , la cual mantiene una media alrededor de 300 para los años de 1997 a 2003, se puede observar que la probabilidad de que haya más de 300 días que rebasan los 100 puntos IMECA va disminuyendo paulatinamente a la medida que pasan los años. Por ejemplo, para el año 1997 se tiene una probabilidad de 0.7390, para 1999 una probabilidad de 0.5802, mientras que para el año 2003 se tiene una probabilidad de 0.1325. Es decir, con lo anterior se ve reflejado la influencia del cambio del control de calidad del aire, modificado en 1999.

Por otro lado para  $M(150)$ , se observa que la media para los distintos años va disminuyendo comparada con la media del año 1997 (203.64), presentándose un cambio notable a partir de 1999, dando como resultado una media mucho menor (56.86) para el año 2003.

$M(200)$  presenta condiciones similares, al comparar la media del año 2003 con los otros años, respectivamente. Dando como resultado que la probabilidad de que se rebase más de 10 días una concentración de 200, es muy baja para el año 2003.

La observación de  $M(240)$ , muestra el cambio drástico que hay al comparar la probabilidad de que se presentara una contingencia ambiental del año 1999 (0.9812) al año 2000 (0.2238).

Por último se puede observar que después de haber tenido una probabilidad alta (0.9900) de que se declarará, más de una vez contingencia ambiental en el año 1997, actualmente, se puede decir, que dicha posibilidad tiene una probabilidad casi nula. Y una vez más se observa como, a partir de la implementación de un nuevo control ambiental en el año de 1999, para los años

---

subsecuentes a él, la probabilidad es mucho menor comparada con el año de 1997.

# Conclusiones

El análisis de tendencias de la calidad del aire a lo largo de los años permite inferir si existe un problema de deterioro creciente o una mejoría paulatina para determinado contaminante. Así, en el presente trabajo se ha recurrido a la estadística Bayesiana y a los procesos estocásticos para mostrar el comportamiento del contaminante ozono, en el Valle de México, durante el periodo de 1997 al 2003.

Para lo cual se consideró el número de días que se rebasó una concentración determinada, como descrito por un modelo Poisson, donde la aplicación central del trabajo fue la determinación del parámetro ( $\lambda$ ) bajo el cual dicho modelo funcionaba. El parámetro ( $\lambda$ ) fue analizado como una variable aleatoria a la que, antes de la evidencia muestral (rebase cada año), se le asignó una distribución a priori Poisson (con base en el grado de creencia con respecto al comportamiento del parámetro aleatorio). Después de usar la evidencia muestral, la distribución a priori fue modificada para obtener una distribución a posteriori Gamma, la cual fue empleada para formular inferencias con respecto al parámetro (la cual es la estructura del método bayesiano).

Se puede concluir que la estimación de  $\lambda$  facilitó el cálculo debido a que la información inicial fue descrita por un miembro de la familia conjugada Gamma, el modelo Poisson.

El método bayesiano permite que tanto la distribución a priori como la distribución a posteriori proporcionen información necesaria para poder tomar decisiones respecto al parámetro analizado.

Finalmente, de acuerdo a los resultados obtenidos, se puede observar que para el contaminante ozono en el caso de contingencia ambiental ésta tiene una probabilidad casi nula de presentarse (debido a que el parámetro que se utiliza para descartarla es raramente alcanzado). No obstante, se puede observar que casi el 80 % de los días del año se rebasa la norma internacional,

aunque se presente una ligera disminución en comparación con el año de inicio del periodo analizado (1997).

Asimismo, se pudo observar, una clara disminución de los días que rebasaban una concentración determinada. a partir de la implementación de un nuevo régimen ambiental en 1999.

# Apéndice A

## Probabilidad condicional

**Definición A.1.** Sea  $E$  un experimento aleatorio, con un espacio muestral  $\Omega$  de posibles salidas, y sea  $\mathcal{A}$  la  $\sigma$ -álgebra sobre  $\Omega$  de acuerdo a la Definición 2.4. Indique por  $P$  la función de probabilidad definida sobre  $\mathcal{A}$ . Sea  $\{B_1, B_2, \dots, B_r\}$  eventos en  $\mathcal{A}$  tal que  $B_1 \cup B_2 \cup \dots \cup B_r = \Omega$  y  $B_i \cap B_j = \emptyset$ ,  $i \neq j$ . Entonces, para cualesquiera eventos  $A$  y  $B \in \mathcal{A}$  con  $P(A) > 0$ ,

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

denota la **probabilidad condicional** que el evento  $B$  ocurra, dado que el evento  $A$  ya ocurrió.

**Teorema A.1 (Teorema de Bayes).** Si  $\{B_1, B_2, \dots, B_r\}$  constituyen una partición de un espacio muestral  $\Omega$ , entonces para cualquier evento  $A \subset \Omega$ ,

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{P(A)} \quad (i = 1, 2, \dots, r),$$

siempre que  $P(A) > 0$ , donde por la ley de probabilidad total,

$$P(A) = \sum_{j=1}^r P(A | B_j)P(B_j).$$

**Teorema A.2 (Teorema de Bayes para densidades continuas).** Sean  $p_y(y)$  o  $f_y(y)$  la función de probabilidad o de densidad de probabilidad a priori de  $Y$ , respectivamente, y sea  $f(x|y)$  la función de verosimilitud. Entonces la

probabilidad a posteriori o función de densidad de probabilidad a posteriori de  $Y$  dada la evidencia muestral  $X$  es

$$P(y|x) = \frac{f(x|y)P_y(y)}{\sum_y f(x|y)P_y(y)} \quad \text{si } y \text{ es discreta}$$

$$f(y|x) = \frac{f(x|y)f_y(y)}{\int_y f(x|y)f_y(y)dy} \quad \text{si } y \text{ es continua}$$

# Apéndice B

## Distribución Poisson

**Lema B.1.** *Sea  $X$  una variable aleatoria que se distribuye Poisson con parámetro  $\lambda$ , es decir,*

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!},$$

entonces  $E[X] = \lambda = \text{Var}[X]$ .

***Demostración:** Por definición se tiene que*

$$E[X] = \sum_{x=0}^{\infty} x \left( \frac{e^{-\lambda} \lambda^x}{x!} \right) = \sum_{x=1}^{\infty} \frac{e^{-\lambda} \lambda^x}{(x-1)!}.$$

Tomando  $\alpha = x - 1$ , se tiene

$$E[X] = \sum_{\alpha=0}^{\infty} \frac{e^{-\lambda} \lambda^{\alpha+1}}{\alpha!} = \lambda \sum_{\alpha=0}^{\infty} \frac{e^{-\lambda} \lambda^{\alpha}}{\alpha!} = \lambda.$$



También por definición,  $\sigma^2 = \text{Var}[X] = E[X^2] - [E(X)]^2$ , entonces

$$\begin{aligned}
 E[X^2] &= \sum_{x=0}^{\infty} \frac{x^2 e^{-\lambda} \lambda^x}{x!} \\
 &= \sum_{x=1}^{\infty} \frac{x e^{-\lambda} \lambda^x}{(x-1)!} \\
 &= \sum_{\alpha=0}^{\infty} \frac{(\alpha+1) e^{-\lambda} \lambda^{\alpha+1}}{\alpha!} \\
 &= \lambda \sum_{x=0}^{\infty} \frac{(\alpha+1) e^{-\lambda} \lambda^{\alpha}}{\alpha!} \\
 &= \lambda \sum_{\alpha=0}^{\infty} \frac{\alpha e^{-\lambda} \lambda^{\alpha}}{\alpha!} + \lambda \sum_{\alpha=0}^{\infty} \frac{e^{-\lambda} \lambda^{\alpha}}{\alpha!} \\
 &= \lambda E[X] + \lambda \\
 &= \lambda^2 + \lambda,
 \end{aligned}$$

por lo que

$$\begin{aligned}
 \sigma^2 &= \text{Var}(X) \\
 &= E(X^2) - (E[X])^2 = \lambda^2 + \lambda - \lambda^2 \\
 &= \lambda.
 \end{aligned}$$

□

# Apéndice C

## Distribución Binomial

**Definición C.1.** Una variable aleatoria tiene una **distribución de probabilidad Binomial** con parámetros  $n$  y  $p$  si se cumple que:

$$P\{K = k\} = \binom{n}{k} p^k q^{n-k} = \binom{n}{k} p^k (1-p)^{n-k}.$$

**Definición C.2.** En todo anillo conmutativo, cualquiera que sea  $n \in N$ , se cumple

$$(a + b)^n = a^n + \binom{n}{1} a^{n-1} b + \dots + \binom{n}{p} a^{n-p} b^p + \dots + b^n = \sum_{p=0}^n \binom{n}{p} a^{n-p} b^p.$$

lo cual constituye el **Teorema del Binomio**.

**Lema C.1.** Sea  $X$  una variable aleatoria que se distribuye Binomial con parámetro  $n$  y  $p$ , entonces su media y varianza están dadas, respectivamente, por

$$\mu = E[X] = np; \quad \sigma^2 = Var[X] = np(1-p) = npq.$$

**Demostración:** Se sabe que

$$\mu = \sum_{x=0}^n x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = \sum_{x=0}^n \frac{x}{x!} \frac{n!}{(n-x)!} p^x q^{n-x}. \quad (C.1)$$

Dado que  $\frac{x}{x!} = \frac{1}{(1-x)!}$  y que el término  $x = 0$  se define como cero, reagrupando los términos de (C.1)

$$\mu = np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

tomando  $y \geq 0$ ,  $y = x - 1$  se tiene

$$\mu = np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} = np \sum_{y=0}^{n-1} \binom{n-1}{y} p^y q^{n-1-y}$$

y por el Teorema del Binomio

$$\sum_{y=0}^{n-1} \binom{n-1}{y} q^{(n-1)-y} p^y = (q+p)^{n-1} = 1^{n-1} = 1,$$

y por tanto  $\mu = np$ .

Por definición, se tiene que,

$$\sigma^2 = \text{Var}[X] = \sum_{x=0}^n x^2 f(x) - \mu^2 = E[X^2] - \mu^2.$$

Calculando  $E[X^2]$  se obtiene,

$$E[X^2] = \sum_{x=0}^n x^2 \frac{n!}{x!(n-x)!} p^x q^{n-x} = np \sum_{x=1}^n x \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}.$$

Haciendo  $y = x - 1$

$$E[X^2] = np \sum_{y=0}^{n-1} (y+1) \frac{(n-1)!}{y!(n-1-y)!} p^y q^{n-y-1} = np \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^y q^{(n-1)-y}.$$

Note que

$$\begin{aligned} \sum_{y=0}^{n-1} (y+1) \binom{n-1}{y} p^y q^{(n-1)-y} &= \sum_{y=0}^{n-1} y \binom{n-1}{y} p^y q^{(n-1)-y} \\ &+ \sum_{y=0}^{n-1} \binom{n-1}{y} p^y q^{(n-1)-y} \\ &= (n-1)p + 1. \end{aligned}$$

Por consiguiente

$$E[X^2] = np[(n-1)p + 1] = np[np + 1 - p] = np[np + q] = (np)^2 + npq.$$

Así que,

$$\sigma^2 = (np)^2 + npq - (np)^2 = npq.$$

□

# Apéndice D

## Distribución Gama

Debido a que la variable aleatoria continua con distribución Gama es elemental en el desarrollo del Capítulo 5, se enuncian algunos resultados de la misma que pueden ser encontrados más ampliamente en Canavos (1987), Grimmett y Welsh (1986) y Hogg (1977).

**Definición D.1.** La función Gama  $\Gamma(t)$  se define por

$$\Gamma(t) = \int_0^{\infty} e^{-y} y^{t-1} dy$$

**Observación:**

$$\begin{aligned} \Gamma(t) &= -e^{-y} y^{t-1} \Big|_0^{\infty} + \int_0^{\infty} e^{-y} (t-1) y^{t-2} dy \\ &= (t-1) \int_0^{\infty} e^{-y} y^{t-2} dy \\ &= (t-1) \Gamma(t-1). \end{aligned}$$

es decir,

$$\frac{\Gamma(t)}{\Gamma(t-1)} = (t-1). \quad (\text{D.1})$$

De (D.1) y dado que  $\Gamma(1) = \int_0^{\infty} e^{-x} dx = 1$  se cumple que

$$\Gamma(n) = (n-1)! \quad (\text{D.2})$$

**Definición D.2.** Una variable aleatoria  $X$  se **distribuye Gama** con parámetros  $\alpha$  y  $\beta$  si su función de densidad está dada por

$$f(x) = \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad x \geq 0.$$

**Lema D.1.** Sea  $X$  una variable aleatoria que se distribuye  $\text{Gama}(\alpha, \beta)$ , entonces

$$E[X] = \frac{\alpha}{\beta}, \quad \text{Var}[X] = \frac{\alpha}{\beta^2}, \quad \text{Moda}[X] = \frac{\alpha - 1}{\beta}.$$

**Demostración:**

$$\begin{aligned} E[X] &= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} x \beta e^{-\beta x} (\beta x)^{\alpha-1} dx \\ &= \frac{1}{\beta \Gamma(\alpha)} \int_0^{\infty} \beta e^{-\beta x} (\beta x)^{\alpha} dx \\ &= \frac{\Gamma(\alpha + 1)}{\beta \Gamma(\alpha)} \\ &= \frac{\alpha}{\beta}. \end{aligned}$$

Por definición  $\sigma^2 = \text{Var}[X] = E[X^2] - [E[X]]^2$ , entonces

$$\begin{aligned} E[X^2] &= \frac{1}{\Gamma(\alpha)} \int_0^{\infty} x^2 \beta e^{-\beta x} (\beta x)^{\alpha-1} dx \\ &= \frac{1}{\beta \Gamma(\alpha)} \int_0^{\infty} x \beta e^{-\beta x} (\beta x)^{\alpha} dx \\ &= \frac{1}{\beta^2 \Gamma(\alpha)} \int_0^{\infty} \beta (x \beta)^{\alpha+1} e^{-\beta x} dx \\ &= \frac{\Gamma(\alpha + 2)}{\beta^2 \Gamma(\alpha)} = \frac{(\alpha + 1) \Gamma(\alpha + 1)}{\beta^2 \Gamma(\alpha)} = \frac{(\alpha + 1) \alpha \Gamma(\alpha)}{\beta^2 \Gamma(\alpha)} \\ &= \frac{\alpha^2 + \alpha}{\beta^2}. \end{aligned}$$

Por lo que

$$\text{Var}[X] = \frac{\alpha^2 + \alpha}{\beta^2} - \left( \frac{\alpha}{\beta} \right)^2 = \frac{\alpha}{\beta^2}.$$

Puesto que  $X$  se distribuye  $Gama(\alpha, \beta)$ , entonces,

$$f(x) = \frac{\beta(\beta x)^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad x \geq 0.$$

De lo cual se tiene que

$$\begin{aligned} f'(x) &= \frac{\beta^\alpha}{\Gamma(\alpha)} [(\alpha-1)x^{\alpha-2} e^{-\beta x} + e^{-\beta x}(-\beta)x^{\alpha-1}] \\ &= \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)} [(\alpha-1)x^{\alpha-2} - \beta x^{\alpha-1}]. \end{aligned}$$

y

$$\begin{aligned} f''(x) &= \frac{\beta^\alpha}{\Gamma(\alpha)} [-(\alpha-1)e^{-\beta x} \beta x^{\alpha-2} + (\alpha-1)(\alpha-2)x^{\alpha-3} e^{-\beta x} \\ &\quad + e^{-\beta x} \beta^2 x^{\alpha-1} - (\alpha-1)x^{\alpha-2} e^{-\beta x} \beta]. \end{aligned}$$

Tomando  $f'(x)=0$  se tiene

$$(\alpha-1)x^{\alpha-2} = \beta x^{\alpha-1},$$

por lo que

$$x = \frac{\alpha-1}{\beta}.$$

Al sustituir  $x = \frac{\alpha-1}{\beta}$  en  $f''(x)$  se tiene

$$f''\left(\frac{\alpha-1}{\beta}\right) = \frac{\beta^3 e^{-\alpha+1}}{\Gamma(\alpha)} [(\alpha-2)(\alpha-1)^{\alpha-2} - (\alpha-1)(\alpha-1)^{\alpha-2}].$$

Si  $f''(x)$  fuera positiva, implicaría que

$$\begin{aligned} [(\alpha-2)(\alpha-1)^{\alpha-2} - (\alpha-1)(\alpha-1)^{\alpha-2}] &> 0 \\ (\alpha-2) &> (\alpha-1) \\ -1 &> 0. \end{aligned}$$

pero eso es una contradicción, por lo que  $f''\left(\frac{\alpha-1}{\beta}\right) < 0$ .

Luego, a partir de que  $f'\left(\frac{\alpha-1}{\beta}\right) = 0$  y de que  $f''\left(\frac{\alpha-1}{\beta}\right) < 0$  se tiene un máximo local en  $x = \frac{\alpha-1}{\beta}$ . Es decir, la moda de una distribución Gama está dada por

$$\text{Moda}[X] = \frac{\alpha-1}{\beta} \quad \alpha > 1.$$

□

**Teorema D.1.** Si las variables aleatorias  $X$  e  $Y$  tienen funciones de densidad  $f_X(x)^*$  y  $f_Y(y)$  conjuntamente continuas e independientes, entonces la función de densidad de  $X + Y$  esta dada por

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy \quad \text{para todo } a \in \mathbb{R}. \quad (\text{D.3})$$

**Demostración:** Ver Grimmett y Welsh (1986). □

**Proposición D.1.** Sean  $X$  e  $Y$  variables aleatorias que se distribuyen Gama independientemente con parámetros  $(s, \lambda)$  y  $(t, \lambda)$ , respectivamente, entonces,  $X + Y$  es una variable aleatoria Gama con parámetros  $(s + t, \lambda)$ .

**Demostración:** Sean  $f_X(x)$  y  $f_Y(y)$  las funciones de densidad respectivas de las variables  $X$  e  $Y$ .

Por (D.3)

$$f_{X+Y}(a) = \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy.$$

Por lo tanto

$$\begin{aligned} f_{X+Y}(a) &= \frac{1}{\Gamma(s)\Gamma(t)} \int_0^a \lambda e^{-\lambda(a-y)} [\lambda(a-y)]^{s-1} \lambda e^{-\lambda y} (\lambda y)^{t-1} dy \\ &\stackrel{(1)}{=} k e^{-\lambda a} \int_0^a (a-y)^{s-1} y^{t-1} dy \\ &\stackrel{(2)}{=} k e^{-\lambda a} a^{s+t-1} \int_0^1 (1-x)^{s-1} x^{t-1} dx \\ &\stackrel{(3)}{=} C e^{-\lambda a} a^{s+t-1} \quad \text{si } a > 0. \end{aligned}$$

a partir de que es una función de densidad debe sumar uno y por lo tanto

$$f_{x+y}(a) = \frac{\lambda e^{-\lambda a} (\lambda a)^{s+t-1}}{\Gamma(s+t)}. \quad (\text{D.4})$$

□

---

\*La **función de densidad** es una función no negativa, cuya integral, extendida sobre todo el eje X, es igual a la unidad, es decir,  $f_X(x) \geq 0$  y  $\int_{-\infty}^{\infty} f_X(x) = 1$  para todo x.

\* (1). Donde  $k = \frac{\lambda^{s+t}}{\Gamma(t)\Gamma(s)}$ .

\* (2). Al sustituir  $y = \frac{x}{a}$ .

\* (3). Donde  $C = \frac{\lambda^{s+t}}{\Gamma(t)\Gamma(s)} \int_0^1 (1-x)^{s-1} x^{t-1} dx$ .

**Teorema D.2.** Sean  $X_1, X_2, \dots, X_n$  variables aleatorias que se distribuyen exponencialmente e independientemente. Entonces  $X_1 + X_2 + \dots + X_n$  se distribuye Gama( $n, \lambda$ ).

**Demostración:** Una variable aleatoria que se distribuye exponencialmente con parámetro  $\lambda$  es una distribución Gama( $1, \lambda$ ). Por lo que, a partir de la Proposición D.1 se cumple que  $\sum_{i=1}^n X_i$  se distribuye Gama con parámetros  $n$  y  $\lambda$ .  $\square$





# Apéndice E

## Funciones $o(x)$

**Definición E.1.** Una función  $f$  será  $o(x)$  si:

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = 0.$$

**Ejemplos y contra ejemplos:**

1. La función  $f(x) = x^2$  es  $o(x)$  dado que:

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = \lim_{x \rightarrow 0} \frac{x^2}{x} = \lim_{x \rightarrow 0} x = 0.$$

2. La función  $f(x) = x$  no es  $o(x)$  dado que:

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = \lim_{x \rightarrow 0} \frac{x}{x} = \lim_{x \rightarrow 0} 1 = 1.$$

**Observación:**

1. Si  $f(\cdot)$  es  $o(x)$  y  $g(\cdot)$  es  $o(x)$ , entonces  $f(\cdot) + g(\cdot)$  también es  $o(x)$ . Lo cual sigue a partir de que:

$$\lim_{x \rightarrow 0} \frac{f(x) + g(x)}{x} = \lim_{x \rightarrow 0} \frac{f(x)}{x} + \lim_{x \rightarrow 0} \frac{g(x)}{x} = 0 + 0 = 0.$$

2. Si  $f(\cdot)$  es  $o(x)$ , entonces  $g(\cdot) = cf(\cdot)$  también es  $o(x)$ . Lo cual se cumple a partir de que:

$$\lim_{x \rightarrow 0} \frac{cf(x)}{x} = c \lim_{x \rightarrow 0} \frac{f(x)}{x} = c0 = 0.$$

Por lo tanto cualquier combinación lineal de funciones que son  $o(x)$  es también una  $o(x)$ .



# Apéndice F

## Gráficas del parámetro $\lambda$

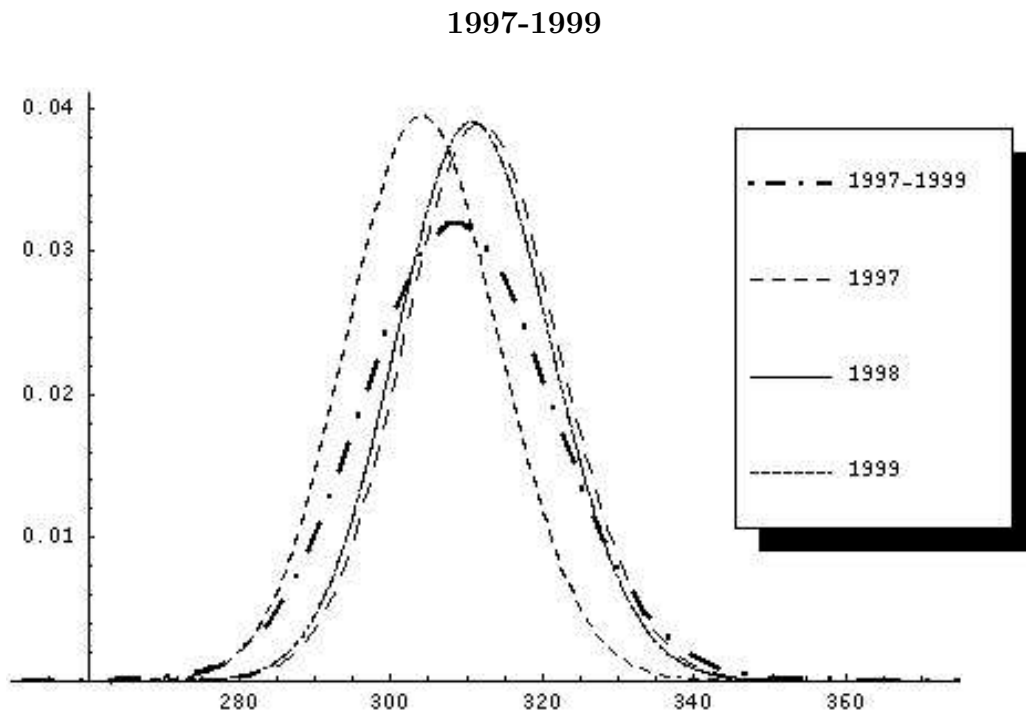


Figura F.1: Distribución a priori de  $\lambda$  para  $M(100)_{1997-1999}$  comparado con las distribuciones posteriores de  $M(100)_i$ ,  $i = 1997, 1998, 1999$ .

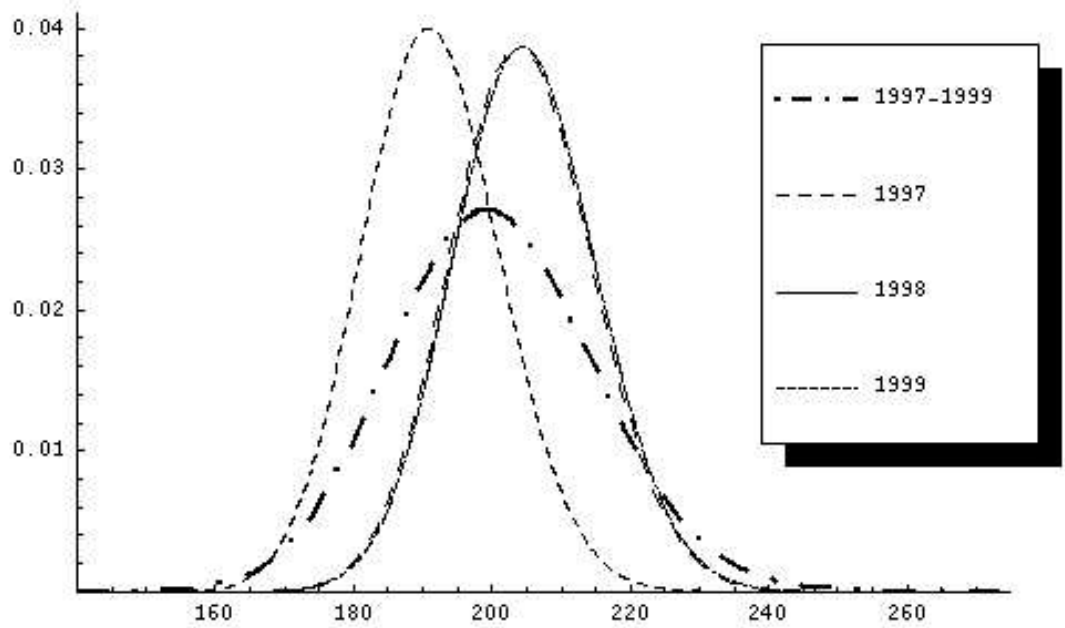


Figura F.2: Distribución a priori de  $\lambda$  para  $M(150)_{1997-1999}$  comparado con las distribuciones posteriores de  $M(150)_i$ ,  $i = 1997, 1998, 1999$ .

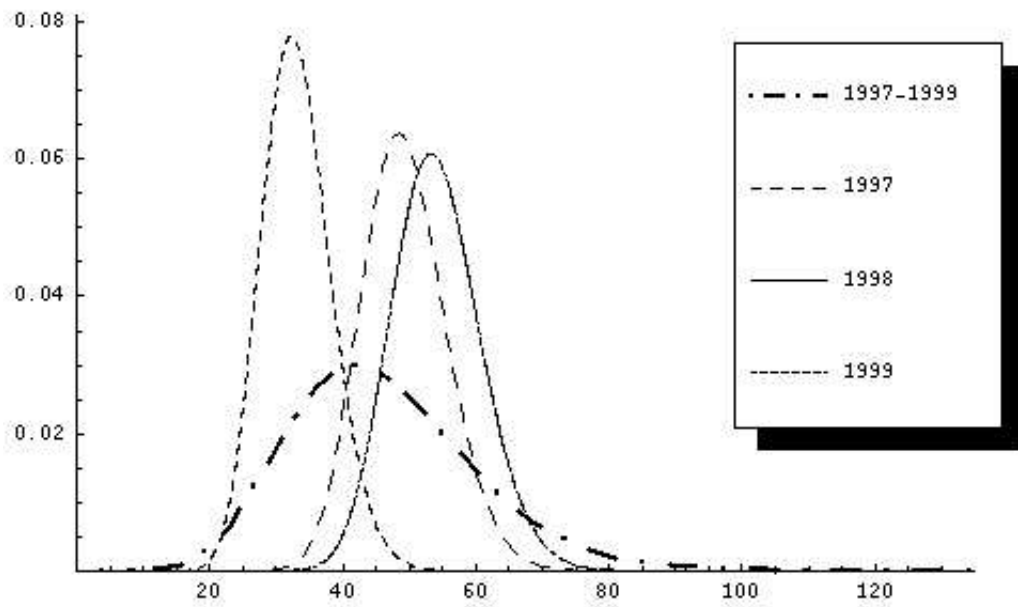


Figura F.3: Distribución a priori de  $\lambda$  para  $M(200)_{1997-1999}$  comparado con las distribuciones posteriores de  $M(200)_i$ ,  $i = 1997, 1998, 1999$ .

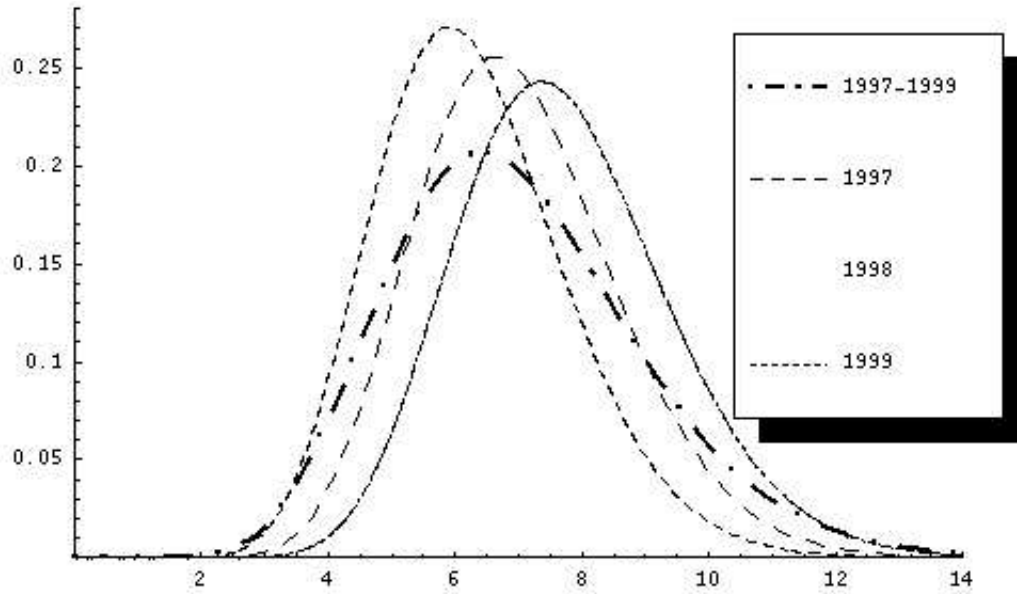


Figura F.4: Distribución a priori de  $\lambda$  para  $M(240)_{1997-1999}$  comparado con las distribuciones posteriores de  $M(240)_i$ ,  $i = 1997, 1998, 1999$ .

**2000-2003**

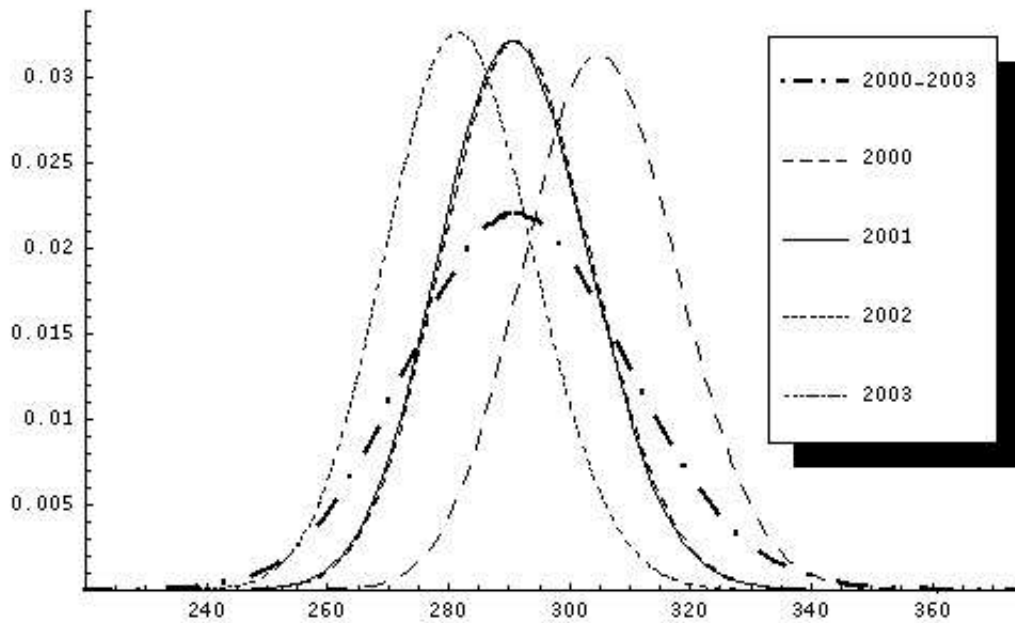


Figura F.5: Distribución a priori de  $\lambda$  para  $M(100)_{2000-2003}$  comparado con las distribuciones posteriores de  $M(100)_i$ ,  $i = 2000, 2001, 2002, 2003$ .

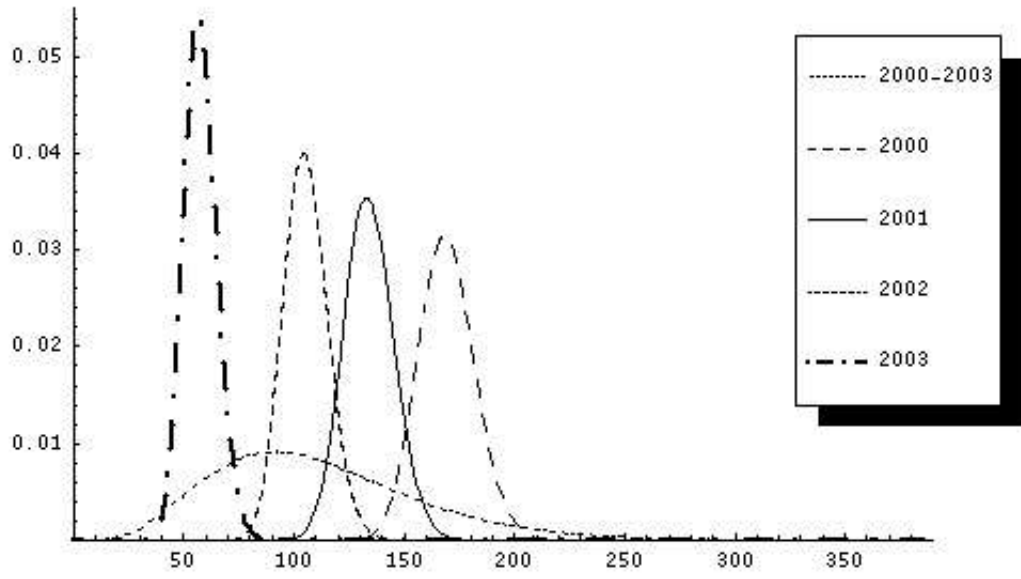


Figura F.6: Distribución a priori de  $\lambda$  para  $M(150)_{2000-2003}$  comparado con las distribuciones posteriores de  $M(150)_i$ ,  $i = 2000, 2001, 2002, 2003$ .

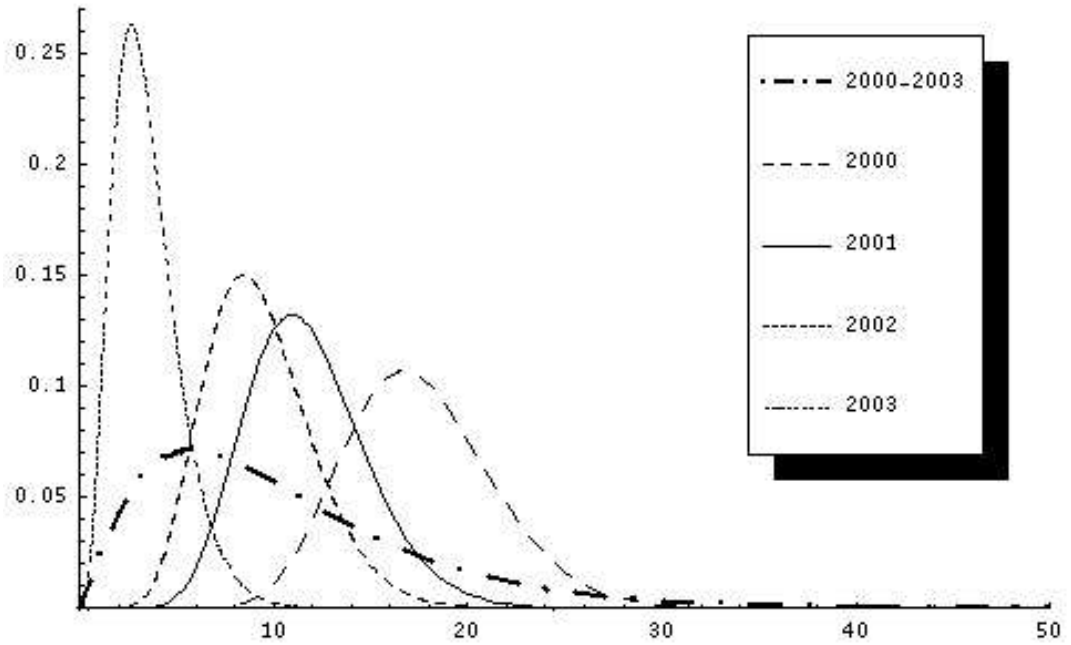


Figura F.7: Distribución a priori de  $\lambda$  para  $M(200)_{2000-2003}$  comparado con las distribuciones posteriores de  $M(200)_i$ ,  $i = 2000, 2001, 2002, 2003$ .

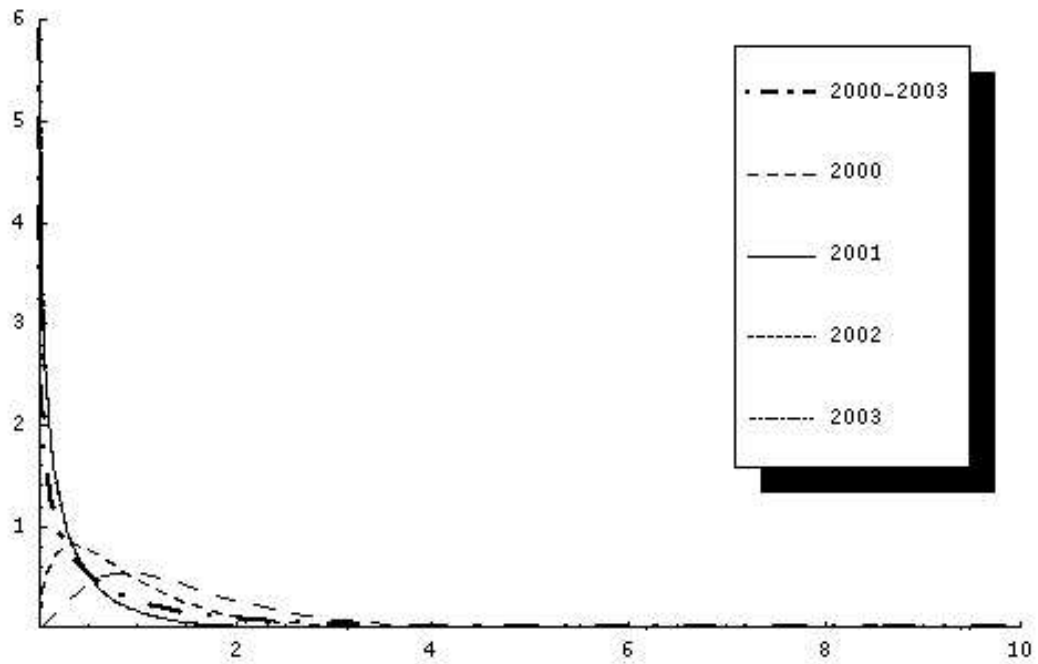


Figura F.8: Distribución a priori de  $\lambda$  para  $M(240)_{2000-2003}$  comparado con las distribuciones posteriores de  $M(240)_i$ ,  $i = 2000, 2001, 2002, 2003$ .



# Bibliografía

- [1] *Y. Alfaro R., 2003, Estimación de líneas ancestrales en genética utilizando el coalescente de Kingman y métodos de Monte Carlo, Tesis de licenciatura, Facultad de Ciencias-UNAM.*
- [2] *R. B Ash, 1972, Real Analysis and Probability, Academic Press, Inc.*
- [3] *P. Brémaud, 1999, Markov Chains, Gibbs fields, Monte Carlo simulation, and queues, Springer.*
- [4] *G. C. Canavos, 1987, Probabilidad y Estadística: Aplicaciones y métodos, Mc Graw-Hill.*
- [5] *DDF. 1996, Programa para Mejorar la Calidad del Aire en el Valle de México 1995-2000, Departamento del Distrito Federal, Gobierno del Estado de México, Secretaría de Medio Ambiente, Recursos Naturales, y Pesca, Secretaría de Salud.*
- [6] *G. R. Grimmett y D. R. Stirzaker, 1982, Probability and Random Processes, Oxford Sciences Publications.*
- [7] *G. Grimmett y D. Welsh, 1986, Probability An Introduction, Oxford Sciences Publications.*
- [8] *R. V. Hogg y E. A. Tanis, 1977, Probability and Statistical Inference, Macmillan Publishin Co., INC.*
- [9] *INE. et al 1997, Primer Informe sobre la Calidad del Aire en Ciudades Mexicanas 1996, Centro Nacional de Investigación y Capacitación Ambiental, Instituto Nacional de Ecología, Dirección General de Gestión e Información Ambiental, Secretaría de Medio Ambiente, Recursos Naturales, y Pesca.*

- [10] *INE et. al. 2000, Gestión de la Calidad del Aire en México, Logros y Retos para el Desarrollo sustentable 1995-2000, Instituto Nacional de Ecología, Dirección General de Gestión e Información Ambiental, Secretaría de Medio Ambiente, Recursos Naturales, y Pesca.*
- [11] *J. S. Javits, 1980, Statistical Interdependence in the Ozone National Ambient Air Quality Standard, J. Air Poll. Control Assoc. 30, 58-59.*
- [12] *S. Karlin, H. M. y Taylor, 1975, A first course in Stochastic Processes, Academic Press, Second Edition.*
- [13] *S. Karlin, H. M. y Taylor, 1981, A second course in Stochastic Processes, Academic Press.*
- [14] *P. K. Lee, 1997, Bayesian Statistics: An Introduction, John Wiley and Sons, Second Edition.*
- [15] *T. Leonard y J. S. Hsu, 1999, Bayesian Methods: An analysis for Statisticians and Interdisciplinary Researchers, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.*
- [16] *J. R. Norris, 1997, Markov chains, Cambridge University Press.*
- [17] *W. R. Ott, 1995, Environmental Statistics and Data Analysis, Lewis Publishers.*
- [18] *S. Ross, 1989, Introduction to Probability Models, Academic Press, Fourth Edition.*
- [19] *S. Ross, 1996, Stochastic Processes, John Wiley and Sons, Second Edition.*
- [20] *P. Sánchez R., 2002, El proceso de líneas de descendencia en el estudio genético de poblaciones, Tesis de licenciatura, Facultad de Ciencias-UNAM.*