



BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

*FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS
POSGRADO EN MATEMÁTICAS*

T E S I S

*Modelación Poisson no homogénea para el caso de
la contaminación atmosférica por ozono en Puebla*

que para obtener el título de:

***MAESTRO EN CIENCIAS
(MATEMÁTICAS)***

presenta:

JUAN ANTONIO CRUZ JUÁREZ

Directores de tesis:

DRA. HORTENSIA REYES CERVANTES

DRA. ELIANE REGINA RODRIGUES

PUEBLA, PUE.

ABRIL 2016

*Dedico esta tesis con todo mi cariño a:
Toda mi familia, amigos y profesores.*

Agradecimientos

Le doy gracias a Dios por darme la oportunidad de llegar a este momento tan importante de mi vida, por darme la dicha de tener una familia maravillosa y por brindarme buenos amigos y profesores.

Le agradezco a mis padres, Antonio Cruz y María Juárez, porque siempre me han apoyado en todas mis decisiones, han estado conmigo en los momentos más felices y tristes de mi vida, día a día me han llenado de su amor, alegría, felicidad y por que sin su apoyo esto no sería posible.

Le agradezco a mis hermanos, Jose Luis, Filiberto, Jesus y Maribel, porque siempre han estado conmigo, por su confianza, por su cariño y su alegría, que han sido parte importante de mi vida.

Le agradezco a la Dra. Hortensia Reyes Cervantes y a la Dra. Eliane Regina Rodrigues por aceptar dirigir esta tesis, por el tiempo que le dedicaron, por sus observaciones, pero muy especialmente por su amistad y sus consejos.

Le agradezco a la Dra. Galdys Linares, Dr. Hugo Cruz, Dr. Francisco Tajonar y al Dr. Humberto Vaquera, por aceptar revisar esté trabajo de tesis y de está forma mejorarlo con sus consejos y correcciones.

Le agradezco a amigos y compañeros: Marisol, Karen, Alicia, Ana Luisa, Rocio, Gelacio, Diego, Aarón, Abel, Leonardo, Enrique, Octavio, Mike, Casas,... que me han apoyado durante el estudio de la maestría, así como los retos de la vida diaria.

Se agradece a Conacyt, por haber tenido la oportunidad de apoyarme económicamente en el transcurso de mis estudios de maestría.

Se agradece a los proyectos, PAPIIT-IN102713 y PAPIIT-IN102416 de la Dirección General de Apoyo al Personal Académico de la Universidad Nacional Autónoma de México (UNAM), por el apoyo económico otorgado para realizar visitas al Instituto de Matemáticas (UNAM). También se agradece al Instituto de Matemáticas (UNAM) por la hospitalidad.

Se agradece a la Vicerrectoría de Investigación y estudios de Posgrado (VIEP-BUAP) por el apoyo para la impresión de la tesis por medio del proyecto llamado La incertidumbre de la estimación en intervalos de proporción binomial (00051).

Se agradece a la Secretaría de Desarrollo Rural, Sustentabilidad Ambiental y Ordenamiento Territorial (SDRSOT) y a Guadalupe Tzintzun Cervantes (Instituto Nacional de Ecología y Cambio Climático), por la información proporcionada y por facilitarnos los datos recabados a través de la Red de monitoreo atmosférico del Estado de Puebla.

Introducción

En las últimas décadas han surgido diversos problemas en las grandes ciudades debido a la expansión poblacional y su fortalecimiento en sus economías mediante crecimiento industrial. Datos de "Population Pyramid" muestran por medio de una pirámide poblacional, el crecimiento de la población en el mundo, que hasta el año pasado fue de 7 349 472 000 habitantes y se estima que para el año 2020 se cuente con una población mundial de 7 758 156 000, ver [33]. En esta misma página se muestra que la población en México el año pasado, fue de 127 017 000 habitantes. Uno de los problemas graves en los países desarrollados o en vía de desarrollo, es la contaminación ambiental. Los 10 países que presentan mayores niveles de contaminación son, de mayor a menor: China, Brasil, Estados Unidos, Indonesia, Japón, México, Rusia, Australia y Perú, ver [47].

Según las últimas estimaciones de la Organización Mundial de la Salud (OMS) sobre la carga mundial de morbilidad, la contaminación del aire exterior e interior provoca unos siete millones de defunciones prematuras, ver [31]. Nuestro país está dentro de los países más contaminados por poseer varias ciudades con problemas graves de contaminación del aire, entre ellas la Ciudad de México, Guadalajara, Monterrey, Puebla, Toluca, entre otras, ver [10]. Las investigaciones en nuestro país respecto a este tipo de contaminación son diversas en varios sentidos, desde conocer las tendencias de los contaminantes químicos mediante modelos matemáticos, ver [5], [2], [19], [20], [42], hasta promover la reducción de dichos contaminantes mediante políticas ambientales, ver [40], [42].

Analizar los datos reales del medio ambiente, como suelen ser: contaminantes químicos, físicos, meteorológicos y otros, resulta ser un trabajo no tan sencillo como pareciera en los modelos matemáticos, pues se deben verificar los supuestos y tener las bases de datos en forma adecuada para la aplicación. En algunos análisis matemáticos y/o estadísticos generalmente se verifican las distribuciones resultantes de las estimaciones, sus parámetros, residuales, intervalos de confianza, entre otras. Muchas veces depende si en la literatura ya se cuentan con pruebas estadísticas fundamentadas. También es importante decir que al elegir un modelo

hay un proceso iterativo que no se dice, pero se realiza, por ejemplo cambiando los modelos para lograr un modelo que prediga o infiera de la mejor forma, teniendo rasgos y propiedades que al investigador y al estadístico le ayuden.

Cuando trabajamos modelos teóricos-matemáticos, el investigador propone condiciones, va creando teoría y por medio de la simulación y programas construidos con esas condiciones verifica si se cumplen o no. Estos modelos también requieren de mucho análisis y conocimiento previo del problema de interés.

Una parte fundamental de los estudios relacionados a la contaminación atmosférica, es resultado de las políticas ambientales que guían a cierto país o región. Entre ellas los límites permisibles al grado de concentración que se encuentran presentes en el ambiente de cierto contaminante (umbrales¹) y los métodos de recabar información (datos), que permitan tener mejores resultados.

La OMS ha realizado recomendaciones sobre los niveles adecuados que deben de mantenerse en los países y actualmente se ha construido un mapa por continentes mostrando los sitios en donde se rebasan. Un estudio dado por el Instituto Mexicano para la Competitividad (IMCO) presenta que el número de muertes por contaminación en México al año es de 5,065 defunciones. Un organismo no gubernamental británico llamado Clean Air Institute (ONG) realizó un estudio sobre los niveles de contaminación de 21 ciudades con más de un millón de habitantes de Latinoamérica, algunas de ellas en México son Guadalajara, Monterrey, Puebla y la Ciudad de México (D.F.), donde la mayor concentración de contaminantes atmosféricos se da en la Ciudad de México (D.F.), ver [14].

La ciudad de Puebla al considerarse como la cuarta ciudad con crecimiento poblacional y desarrollo industrial en el país, empieza a presentar problemas causados por la contaminación atmosférica. Estudios como el Programa de Gestión de Calidad del Aire del Estado de Puebla (PROAIRE) 2012-2020, ver [42], sugieren que si los estándares ambientales fueran más estrictos se reducirían en gran cantidad las enfermedades causadas por el aire que se respira, y esto traería como consecuencia una disminución en el gasto corriente en materia de salud. Este mismo estudio analiza el panorama en el que se encuentra el Estado de Puebla en

¹Concepto teórico para la concentración de una sustancia que representa la transición entre la exposición máxima que no produce efectos adversos y la exposición mínima que produce un efecto adverso bajo condiciones definidas.

cuanto a contaminación ambiental hasta el año 2009 en términos generales.

Tomando en cuenta los trabajos Álvarez et al. y Achcar et al. (2005, 2008, 2011), ver [3],[4], [1], y centrados en el trabajo Achcar (2012), ver [2], donde se realiza una comparación de diferentes funciones de riesgo para un proceso Poisson no homogéneo, aplicado a datos máximos de ozono obtenidos a partir del sistema de monitoreo atmosférico de la Zona Metropolitana del Valle de México (ZMVM). En [2] también se aborda el método bayesiano para hallar la estimación de los hiperparámetros considerando puntos de cambio.

Se hará un análisis semejante utilizando los datos de Ozono de la Ciudad de Puebla de las cuatro estaciones de monitoreo de la Ciudad de Puebla en el periodo de 2001-2009 usando diferentes funciones de riesgo en un proceso Poisson no homogéneo. En el trabajo se plantean los siguientes puntos.

- Utilizar algunas funciones de riesgo para un proceso Poisson no homogéneo que se presentan en el artículo Achcar et al. (2012), ver [2]. Las funciones de riesgo son:
 - Función de riesgo de Weibull
 - Función de riesgo Musa-Okumoto
 - Función de riesgo Goel-Okumoto
 - Función de riesgo Goel-Okumoto generalizada
 - Función de riesgo Weibull-Geométrica
- Utilizar el punto de vista Bayesiano para estimar los parámetros presentes en los modelos considerados.
- Utilizar un programa en R y/o Winbugs para estimar los parámetros del modelo Bayesiano.

La tesis está dividida de la siguiente manera: en los primeros capítulos se da teoría que es necesaria para entender el problema y su modelación matemática y en los subsecuentes se describe método propuesto y la solución, cabe destacar que el trabajo presente solo es una aproximación al fenómeno físico estudiado. En el primer

capítulo se dan conceptos básicos importantes relacionados con cadenas de Markov. En el segundo capítulo se dan definiciones acerca de los procesos Poisson tanto en el sentido homogéneo como el no homogéneo, también se da una pequeña introducción referente a modelos con puntos de cambio. Se presentan algunas funciones de tasa interesantes para nuestra modelación. En el tercer capítulo se dan definiciones importantes de la inferencia bayesiana y se explica el método para calcular las constantes de proporcionalidad que se requieren para identificar la distribución a posteriori de los modelos. Finalmente, el cuarto capítulo se desarrolla la aplicación del método propuesto para la contaminación por ozono en la capital de Puebla, se conoce el estado en el que se encuentra la región de estudio. Posteriormente se dan las conclusiones, algunas consideraciones que se hacen al momento de realizar las corridas en el programa OpenBugs para estimar los parámetros, así como los trabajos futuros a realizar.

Índice general

1. Cadenas de Markov	13
1.1. Cadenas de Markov	13
1.1.1. Probabilidades de transición en n -pasos	16
1.1.2. Tiempos de alcance	17
1.2. Distribuciones estacionarias	20
1.3. Reversibilidad	23
 2. Procesos Poisson	 25
2.1. Procesos Poisson	26
2.1.1. Proceso Poisson no homogéneo	27
2.2. Funciones de Supervivencia y Riesgo (Intensidad, Tasa)	28
2.3. Modelos con puntos de cambio	29
2.4. Funciones de tasa	31
2.4.1. Distribución Weibull (W)	31
2.4.2. Distribución Musa-Okumoto (MO)	33

2.4.3. Distribución Goel-Okumoto (GO)	35
2.4.4. Distribución Goel-Okumoto Generalizada (GOG)	36
2.4.5. Distribución Weibull-Geométrica (WG)	38
3. Inferencia Bayesiana	41
3.1. Distribucion a priori (Inicial)	43
3.2. Función de verosimilitud	44
3.3. Distribución a posteriori (Final)	44
3.4. Criterio Bayesiano para Comparación de Modelos	45
3.5. Método Monte Carlo	45
3.5.1. Muestreador de Gibbs	48
3.5.2. El muestreador de Gibbs como caso particular del algoritmo de Metropolis-Hastings	49
4. Aplicación a Niveles de Ozono en la Ciudad de Puebla	51
4.1. Descripción del modelo	52
4.2. Inferencia Bayesiana sobre el modelo planteado	53
4.3. Modelos utilizados	56
4.3.1. Weibull (W)	57
4.3.2. Musa-Okumoto (MO)	58
4.3.3. Goel-Okumoto (GO)	58
4.3.4. Goel-Okumoto Generalizada (GOG)	59
4.3.5. Weibull-Geométrica (WG)	60

<i>ÍNDICE GENERAL</i>	7
4.4. Caso de estudio	61
4.4.1. Extensión Geográfica del Estado de Puebla	63
4.4.2. Calidad del Aire en Puebla	65
4.5. Resultados	69
A. Teoremas y definiciones.	81

Índice de cuadros

4.1. Contaminates Atmosféricos. Fuente: Proaire 2012-2020	63
4.2. Estadística descriptiva para las 4 estaciones de monitoreo atmosférico .	67
4.3. Resumen estadístico para las funciones de riesgo sin puntos de cambio .	71
4.4. Resumen estadístico con un punto de cambio	74

Índice de figuras

2.1. Función de riesgo de Weibull	33
2.2. Función de riesgo de Musa-Okumoto	34
2.3. Función de riesgo de Goel-Okumoto	36
2.4. Función de riesgo de Goel-Okumoto generalizada ($a, c \geq 0$)	38
2.5. Función de riesgo de Weibull-Geométrica ($b, c \geq 0$)	40
4.1. Vías primarias y secundarias de comunicación en el Estado de Puebla. Fuente: Proaire 2012-2020	64
4.2. Estaciones de Monitoreo Atmosférico de Puebla, 1.- Tecnológico, 2.- Serdán, 3.-Ninfas y 4.- Agua Santa	65
4.3. Series de tiempo para los máximos en las 4 estaciones de monitoreo at- mosférico, del periodo 1/01/2001 al 31/12/2009. La línea horizontal co- rresponde al nivel del umbral (110 ppm)	66
4.4. Comportamiento del ozono en la ciudad de Puebla, umbral $u = 110$ ppb	68
4.5. Gráficos de densidades de la distribución a posteriori para las funcio- nes de riesgo: 1.-Weibull; 2.-Musa-Okumoto; 3.-Goel-Okumoto; 4.-Goel- Okumoto generalizada; 5.-Weibull-Geométrica	70

4.6.	Gráfico de las medias acumuladas, para el proceso Poisson no homogéneo con funciones de riesgo Weibull, Musa-Okumoto, Goel-Okumoto, Goel-Okumoto generalizada y Weibul-Geométrica, respecto a la media de las observaciones y sin puntos de cambio	73
4.7.	Gráficos de densidades de la distribución a posteriori para las funciones de riesgo: 1.-Weibull; 2.-Musa-Okumoto; con un punto de cambio	75
4.8.	Gráfico de las medias acumuladas, para el proceso Poisson no homogéneo con funciones de riesgo Weibull, Musa-Okumoto, respecto a la media de las observaciones y con un punto de cambio	76

Capítulo 1

Cadenas de Markov

En este capítulo se presenta la teoría relacionada con cadenas de Markov, que corresponde al estudio de sucesiones de variables aleatorias, que toman valores en conjuntos finitos o contables, llamado *espacio de estados*, y satisfacen la *propiedad de Markov*. Además, se definen los tipos de estados que presenta una cadena de Markov, que pueden ser recurrentes o transitorios. Así mismo, las características de una cadena que presenta distribución (es) estacionaria (s), ver [36]. En este capítulo se presentan algunas demostraciones, la mayoría se pueden encontrar en las referencias citadas, [7] y [15].

1.1. Cadenas de Markov

Definición 1.1 *Un proceso estocástico es una colección de variables aleatorias $\{X_t : t \in \Theta\}$ indexadas por un conjunto Θ , llamado conjunto de índices, este conjunto puede tomar valores $\Theta = \{0, 1, 2, 3, \dots\}$ o $\Theta = [0, \infty)$. El conjunto S en donde son tomados los valores de las variables aleatorias, es llamado espacio de estados. Este conjunto puede ser discreto o continuo.*

Definición 1.2 *Un proceso estocástico $\{X_t : t = 0, 1, 2, \dots\}$, con espacio de esta-*

dos contable S , es llamado *cadena de Markov*, si cumple la siguiente propiedad:

$$P(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} | X_n = x_n), \quad (1.1)$$

$x_0, x_1, x_2, \dots, x_n, x_{n+1} \in S$, $n \geq 0$. A esta propiedad se le conoce como **propiedad de Markov**.

La propiedad de Markov, dice que la distribución de probabilidad del valor futuro de una variable aleatoria depende únicamente de su estado actual, siendo su pasado irrelevante. En general, una cadena de Markov se puede presentar en términos de una ecuación de la siguiente forma, ver [15]:

$$X_{n+1} = F(X_n, \xi_n),$$

con $F : S \times \mathbb{R} \longrightarrow S$, una función que depende solo de X_n y S el espacio de estados de la cadena, las ξ_n son variables independientes e idénticamente distribuidas.

Definición 1.3 *La probabilidad de cambiar del estado x al estado y , dado que se encuentra en el estado x , es decir,*

$$P(X_{n+1} = y | X_n = x), \quad x, y \in S, \quad (1.2)$$

*se le conoce como **probabilidad de transición**, y es denotado por $p(x, y)$.*

Toda cadena de Markov empieza con un valor inicial. Tal valor inicial puede obtenerse a través de una distribución, dada por:

$$\pi_0(x) = P(X_0 = x), x \in S, \quad (1.3)$$

llamada **distribución inicial de la cadena**.

Definición 1.4 *Una cadena de Markov se dice **homogénea**, si su probabilidad de transición no depende del valor del tiempo n , es decir,*

$$P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x). \quad (1.4)$$

A la matriz $P = (p(x, y))_{x, y \in S}$, se le conoce como matriz de transición y es una matriz estocástica, es decir, cumple que:

$$\sum_{y \in S} p(x, y) = 1. \quad (1.5)$$

En efecto, pues se tiene que:

$$\sum_{y \in S} p(x, y) = \sum (P(X_{n+1} = y | X_n = x)) = P(\bigcup_{y \in S} \{X_{n+1} = y\} | X_n = x) = P(\Omega | X_n = x) = 1.$$

La distribución conjunta de $X_0, X_1, \dots, X_n$ se puede expresar en términos de la distribución de transición $p(x, y)$ y la distribución inicial π_0 .

Teorema 1.1 *Sea $X_0, X_1, \dots, X_n$ una cadena de Markov con distribución inicial π_0 y distribución de transición $p(x, y)$, entonces,*

$$P(X_0 = x_0, \dots, X_n = x_n) = \pi_0(x_0) \cdot p(x_0, x_1) \cdot p(x_1, x_2) \cdot \dots \cdot p(x_{n-1}, x_n).$$

Demostración 1.1 *La demostración se realiza usando inducción sobre m . Para $m=1$, usando la definición de probabilidad condicional, se tiene:*

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1) &= P(X_1 = x_1 | X_0 = x_0) P(X_0 = x_0). \\ &= \pi_0(x_0) \cdot p(x_0, x_1). \end{aligned}$$

Asuma válido para $m=n-1$, es decir:

$$P(X_0 = x_0, \dots, X_{n-1} = x_{n-1}) = \pi_0(x_0) \cdot p(x_0, x_1) \cdot p(x_1, x_2) \cdot \dots \cdot p(x_{n-2}, x_{n-1}).$$

Ahora demostremos para $m=n$. Usando la regla del producto y la definición de cadenas de markov se tiene que,

$$\begin{aligned} P(X_0 = x_0, X_1 = x_1, \dots, X_n = x_n) &= P(X_n = x_n | X_0 = x_0, \dots, X_{n-1} = x_{n-1}) \\ &\quad P(X_0 = x_0, \dots, X_{n-1} = x_{n-1}). \end{aligned}$$

Ahora por ser una cadena de Markov, el lado derecho de la igualdad es:

$$p(x_{n-1}, x_n) \cdot P(X_0 = x_0 \dots X_{n-1} = x_{n-1}),$$

finalmente, usando la hipótesis de inducción, se tiene que:

$$\begin{aligned} P(X_0 = x_0, \dots, X_n = x_n) &= \pi_0(x_0) \cdot p(x_0, x_1) \cdot p(x_1, x_2) \cdot \dots \cdot p(x_{n-2}, x_{n-1}) \cdot \\ &\quad p(x_{n-1}, x_n). \end{aligned}$$

Note que la distribución marginal de un estado $y \in S$, es dada por:

$$\begin{aligned} P(X_n = y) &= P(X_n = y, \bigcup_{x \in S} \{X_{n-1} = x\}) = \sum_{x \in S} P(X_n = y, X_{n-1} = x) \\ &= \sum_{x \in S} P(X_n = y | X_{n-1} = x) P(X_{n-1} = x), \end{aligned}$$

si se usa la siguiente notación, $\pi_n(x) := P(X_n = x)$, se tiene:

$$P(X_n = y) = \pi_n(y) = \sum_{x \in S} p(x, y) \pi_{n-1}(x). \quad (1.6)$$

1.1.1. Probabilidades de transición en n -pasos

La función de transición en n -pasos $p^n(x, y)$, denota la probabilidad de ir al estado y , en n -pasos, dado que se inicia en el estado x , se define por:

$$p^n(x, y) = P(X_n = y | X_0 = x), \quad (1.7)$$

con $x, y \in S$, donde:

1. Si $n = 0$, $x, y \in S$

$$p^0(x, y) = \delta_{(x, y)}, \text{ donde } \delta_{(x, y)} = 1, \text{ si } x = y, \text{ y es cero si } x \neq y.$$

2. Si $n = 1$, $x, y \in S$

$$p^1(x, y) = p(x, y).$$

3. Mientras que para $n \geq 2$, $x, y \in S$

$$p^n(x, y) = \sum_{y_1 \in S} \sum_{y_2 \in S} \cdots \sum_{y_{n-1} \in S} p(x, y_1) p(y_1, y_2) \cdots p(y_{n-1}, y)$$

Ecuación de Chapman-Kolmogorov

Para $n, m \geq 1$, $x, y \in S$, se puede obtener $p^{n+1}(x, y)$, de la siguiente forma:

$$p^{n+m}(x, y) = P(X_{n+m} = y | X_0 = x) = \sum_{z \in S} p^n(x, z) p^m(z, y).$$

La expresión anterior se puede escribir:

$$\begin{aligned}
 p^{n+m}(x, y) &= \frac{P(X_0 = x, X_{n+m} = y)}{P(X_0 = x)} \\
 &= \sum_{z \in S} \frac{P(X_0 = x, X_n = z, X_{n+m} = y)}{P(X_0 = x)} \cdot \frac{P(X_0 = x, X_n = z)}{P(X_0 = x, X_n = z)} \\
 &= \sum_{z \in S} P(X_{n+m} = y | X_0 = x, X_n = z) \cdot P(X_n = z | X_0 = x) \\
 &= \sum_{z \in S} P(X_{n+m} = y | X_n = z) \cdot P(X_n = z | X_0 = x) \\
 &= \sum_{z \in S} p^n(x, z) p^m(z, y).
 \end{aligned}$$

Observación 1.1 Otra forma de obtener $P(X_n = y)$ es:

$$\begin{aligned}
 P(X_n = y) &= P(X_n = y, \bigcup_{x \in S} \{X_0 = x\}) \\
 &= \sum_{x \in S} P(X_n = y, X_0 = x) \\
 &= \sum_{x \in S} P(X_n = y | X_0 = x) P(X_0 = x) \\
 &= \sum_{x \in S} p^n(x, y) \pi_0(x).
 \end{aligned}$$

1.1.2. Tiempos de alcance

Definición 1.5 Sea $x \in S$, definimos el tiempo de alcance como

$$\tau_x = \inf\{n \geq 1 | X_n = x\}, \quad (1.8)$$

así, τ_x es una variable aleatoria.

Notación: Sea $P(\cdot | X_0 = x) = P_x(\cdot)$ y $E(\cdot | X_0 = x) = E_x(\cdot)$.

De esta forma, la distribución de τ_x esta dada por:

$$\begin{aligned}
 P_u(\tau_x = 1) &= p(u, x) \\
 P_u(\tau_x = 2) &= \sum_{z \neq x} p(u, z) p(z, x) = \sum_{z \neq x} p(u, z) P_z(\tau_x = 1) \\
 P_u(\tau_x = 3) &= \sum_{z \neq x} p(u, z) P_z(\tau_x = 2),
 \end{aligned}$$

y en general,

$$P_u(\tau_x = m + 1) = \sum_{z \neq x} p(u, z) P_z(\tau_x = m).$$

Note que,

$$\begin{aligned} P_u(\tau_x = \infty) &= 1 - P_u(\tau_x < \infty) \\ &= 1 - \lim_{n \rightarrow \infty} P_u(\tau_x \leq n) \\ &= 1 - \lim_{n \rightarrow \infty} \sum_{k=1}^n P_z(\tau_x = k) \\ &= 1 - \sum_{k=1}^{\infty} \sum_{z \neq x} p(u, z) P_z(\tau_x = k - 1). \end{aligned}$$

Clasificación de estados

Definición 1.6 Sea $\rho_{xy} := P_x(\tau_y < \infty)$, $x, y \in S$. En el caso que $\rho_{yy} = 1$, es decir, la probabilidad de regresar al estado y , dado que inicialmente se encuentra en el estado y es 1, entonces a y se le conoce como **estado recurrente**.

Cuando la probabilidad de no retorno es positiva, es decir, $0 < 1 - \rho_{yy}$, entonces a y se le conoce como, **estado transitorio**.

Observación 1.2 Si la totalidad de los estados de una cadena son recurrentes, se dirá que la cadena es recurrente. Cuando todos los estados de la cadena son transitorios, se dice que la cadena es transitoria.

Definición 1.7 (Cadenas Irreducibles) Una cadena de Markov se dice irreducible, si todos los estados son alcanzables desde cualquier otro estado en un número finito de pasos, esto es, para cualquiera $x, y \in S$ existen $n \geq 0$ y $m \geq 0$, tales que $p^n(x, y) > 0$ y $p^m(y, x) > 0$.

Sea $N(y)$ = número de visitas al estado $y \in S$, es decir,

$$N(y) = \sum_{n=1}^{\infty} I_y(X_n), \quad (1.9)$$

donde,

$$I_y(X_n) = \begin{cases} 1, & \text{si } X_n = y, \\ 0, & \text{cualquier otro caso.} \end{cases}$$

Note que $N(y)$, es una variable aleatoria. La distribución de $N(y)$ se expresa,

$$\begin{aligned} P_x(N(y) \geq 1) &= P_x\left(\bigcup_{m=1}^{\infty} \{\tau_y = m\}\right) \\ &= P_x(\tau_y < \infty) = \rho_{xy}. \end{aligned}$$

$$\begin{aligned} P_x(N(y) \geq 2) &= \sum_{m_1} P_x(\tau_y = m_1) \sum_{m_2} P_y(\tau_y = m_2) \\ &= \rho_{xy} \rho_{yy}. \end{aligned}$$

En general,

$$P_x(N(y) \geq m) = \rho_{xy} \rho_{yy}^{m-1}, \text{ con } m \geq 1.$$

Entonces,

$$\begin{aligned} P_x(N(y) = m) &= P_x(N(y) \geq m) - P_x(N(y) \geq m+1) \\ &= \rho_{xy} \rho_{yy}^{m-1} - \rho_{xy} \rho_{yy}^m \\ &= \rho_{xy} \rho_{yy}^m (1 - \rho_{yy}). \end{aligned}$$

Finalmente,

$$P_x(N(y) = 0) = 1 - P_x(N(y) \geq 1) = 1 - \rho_{xy}.$$

Ahora, sea $G(x, y)$ que denota la esperanza de la variable aleatoria $N(y)$ dado que la cadena de Markov inicia en el estado x , es decir,

$$G(x, y) = E_x(N(y)) = E_x \left[\sum_{n=1}^{\infty} I_y(X_n) \right].$$

Entonces, se tiene que:

$$\begin{aligned} G(x, y) &= \sum_{n=1}^{\infty} E_x(I_y(X_n)) = \sum_{n=1}^{\infty} P_x(X_n = y) \\ &= \sum_{n=1}^{\infty} p^n(x, y). \end{aligned}$$

A continuación se enuncian dos teoremas que caracterizan a los estados, tanto transitorios como recurrentes.

Teorema 1.2 (Estados transitorios) *Sea y un estado transitorio, entonces:*

1. $P_x(N(y) < \infty) = 1$.

2. $G(x, y) = \frac{\rho_{xy}}{1 - \rho_{yy}}$.

Demostración 1.2 *Ver [15].*

Teorema 1.3 (Estados recurrentes) *Sea y un estado recurrente, entonces,*

1. $P_x(N(y) = \infty) = 1, x \in S$.

2. Si $\rho_{xy} > 0$, entonces, $G(x, y) = \infty$.

- Si $\rho_{xy} = 0$, entonces, $G(x, y) = 0$.

La demostración se puede consultar en [15].

Definición 1.8 *Sean $x, y \in S$, se dirá que x accede a y , denotado por $x \rightarrow y$, si y sólo si, existe $m \geq 0$ tal que $p^m(x, y) > 0$.*

Teorema 1.4 *Si x es un estado recurrente en donde $x \rightarrow y$, entonces y es recurrente y $\rho_{xy} = \rho_{yx} = 1$.*

Demostración 1.3 *Ver [15].*

1.2. Distribuciones estacionarias

Definición 1.9 *Una distribución π es llamada estacionaria si satisface las siguientes propiedades:*

1. $\sum_{x \in S} \pi(x) = 1$, para cada $x \in S$.

2. $\pi(y) = \sum_x \pi(x)p(x, y)$, $y \in S$.

Observación 1.3 *Supóngase que la distribución inicial de una cadena de Markov $\{X_n\}$, es la distribución estacionaria π , entonces,*

$$\pi(y) = \lim_{n \rightarrow \infty} P(X_n = y).$$

Si $N_n(y)$, $n \geq 0$, denota el número de visitas al estado y , en las primeras n -unidades de tiempo, es decir,

$$N_n(y) = \sum_{k=1}^n I_y(X_k). \quad (1.10)$$

Además, si $G_n(x, y) := E_x(N_n(y))$, es el valor esperado de visitas al estado y , en las primeras n -unidades de tiempo, dado que la cadena empieza en el estado x , entonces,

$$\lim_{n \rightarrow \infty} N_n(y) = N(y), \text{ con } x, y \in S, \lim_{n \rightarrow \infty} G_n(x, y) = G(x, y).$$

Ahora si y , es un estado transitorio,

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{N_n(y)}{n} &= 0, \\ \lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} &= 0. \end{aligned}$$

Sea $y \in S$ un estado recurrente, se define

$$m_y := E_y(\tau_y), \quad (1.11)$$

como el tiempo promedio de retorno al estado y .

Teorema 1.5 *Si $y \in S$ es un estado recurrente, entonces,*

- $\lim_{n \rightarrow \infty} \frac{N_n(y)}{n} = \frac{I_{\{\tau_y < \infty\}}}{m_y}$, casi seguramente.
- $\lim_{n \rightarrow \infty} \frac{G_n(x, y)}{n} = \frac{\rho_{xy}}{m_y}$, $x \in S$.

Demostración 1.4 *Ver [15].*

Clasificación de estados recurrentes

Definición 1.10 *Sea $y \in S$ un estado recurrente, entonces:*

1. *y es un estado recurrente nulo si, $m_y = \infty$.*
2. *y es un estado recurrente positivo si, $m_y < \infty$.*

Corolario 1.1 Si $x \in S$, es un estado recurrente positivo y si $x \rightarrow y$, entonces, y es recurrente positivo.

Corolario 1.2 Si $\{X_n\}$ es una cadena de Markov finita, entonces, la cadena tiene al menos un estado recurrente y este es recurrente positivo.

Teorema 1.6 Si $\pi(\cdot)$ es una distribución estacionaria de una cadena de Markov y se tiene que $x \in S$, es un estado transitorio o recurrente nulo, entonces, $\pi(x) = 0$.

Teorema 1.7 Si $\{X_n\}$ es una cadena de Markov, irreducible de estados recurrentes positivos, entonces, la cadena tiene una única distribución estacionaria π , dada por

$$\pi(x) = \frac{1}{m_x}, \quad x \in S. \quad (1.12)$$

Cadenas periódicas

Recordar que para $I \subseteq \mathbb{Z}$ se define el máximo común divisor (denotado por m.c.d.) como: Sí $d := \text{mcd}(I)$ cumple,

1. si $c \in I$, entonces $d|c$, es decir, d divide a c .
2. si existe $l \in I$, tal que $l|c$, para cualquier $c \in I$, entonces $l \leq d$.

De esta forma para $x \in S$, se toma a $d_x = \text{mcd}(I_x)$, donde $I_x := \{n \geq 1 | p^n(x, x) > 0\}$, como el período del estado $x \in S$.

Proposición 1.1 Si $x \leftrightarrow y$, es decir, si x accede desde y y viceversa, con $x, y \in S$ entonces $d_x = d_y$.

Definición 1.11 Sea $\{X_n\}$ una cadena de Markov irreducible y sea d su período, si $d = 1$, se dice que la cadena es **aperiódica**. En el caso de que $d > 1$, la cadena es **periódica**, con período d .

El siguiente teorema caracteriza cuando una cadena de Markov tiene una distribución estacionaria única.

Teorema 1.8 *Sea $\{X_n\}$ una cadena de Markov irreducible, recurrente positiva y aperiódica, y sea π su distribución estacionaria, entonces*

$$\lim_{n \rightarrow \infty} p^n(x, y) = \pi(y), \quad y \in S. \quad (1.13)$$

Demostración 1.5 *Ver [15].*

1.3. Reversibilidad

Sea $M = \{X_n\}$ una cadena de Markov con matriz de transición P y espacio de estados S , y defina $Y_n = X_{-n}$, entonces, si se define a $M' = \{Y_n\}$, M' es también una cadena de Markov y se conoce como cadena reversa en el tiempo, de M .

Definición 1.12 *Una cadena de Markov M es reversible si la probabilidad de transición de la cadena de Markov reversa en el tiempo coincide con las probabilidades de transición de M , es decir, $P(X_{n+1} = y | X_n = x) = P(Y_{n+1} = y | Y_n = x)$.*

Teorema 1.9 *Supóngase que $M = \{X_n\}$ es una cadena de Markov con matriz de transición P , con distribución estacionaria única π . Entonces, M es reversible con respecto a P si y sólo si, para $x, y \in S$*

$$\pi(x)p(x, y) = \pi(y)p(y, x). \quad (1.14)$$

*La ecuación (1.15) es conocida como **ecuación de equilibrio**.*

Capítulo 2

Procesos Poisson

El proceso de Poisson es usado a menudo para modelar el número de fallas en sistemas en los cuales los eventos se presentan en periodos de tiempo con una distribución exponencial. Este proceso es un caso particular de los conocidos procesos de conteo, ver [36].

Definición 2.1 *Se dice que un proceso estocástico $\{N(t); t \geq 0\}$ es un proceso de conteo, si representa el número de eventos ocurridos hasta el tiempo t , esto es, satisface, ver [32]:*

1. $N(0) \geq 0$.
2. $N(t)$ toma valores enteros.
3. Es no decreciente, i.e., si $s < t$, entonces $N(s) \leq N(t)$.
4. Para $s < t$, $N(t) - N(s)$ es igual al número de eventos que ocurren en el intervalo $(s, t]$.

Definición 2.2 *Un proceso de conteo $\{N(t); t \geq 0\}$ se dice de **incrementos independientes**, si el número de eventos que ocurren en intervalos de tiempo disjuntos son independientes, es decir, para $s < t < r < u$, $N(t) - N(s)$ y $N(u) - N(r)$, son variables aleatorias independientes.*

Definición 2.3 *Un proceso de conteo $\{N(t); t \geq 0\}$ se dice de **incrementos estacionarios**, si la distribución de la cantidad de eventos en un intervalo depende únicamente de la longitud del intervalo, es decir, la distribución de $N(t) - N(s)$, $s < t$ depende solamente de $t - s$.*

2.1. Procesos Poisson

Suponga que un mismo evento ocurre repetidas veces de manera aleatoria a lo largo del tiempo. Tal evento puede ser, la llegada de una reclamación a una compañía de seguros, la llegada de un cliente a una ventanilla para solicitar un servicio, la ocurrencia de que cierto contaminante rebase un nivel dado, etcétera. De esta forma se da la siguiente definición.

Definición 2.4 *Sea $T_1, T_2, \dots$ una sucesión de variables aleatorias independientes cada una con distribución exponencial con parámetro λ . El proceso de Poisson de parámetro λ , es el proceso a tiempo continuo $\{N(t) : t \geq 0\}$ definido de la siguiente manera:*

$$N(t) = \max\{n \geq 0 : T_1 + T_2 + \dots + T_n \leq t\}. \quad (2.1)$$

En palabras, la variable $N(t)$ es el entero n máximo tal que $T_1 + T_2 + \dots + T_n$ es menor o igual a t , y ello equivale a contar el número de eventos que ocurren hasta el tiempo t . A continuación se dan definiciones alternativas, ver [36] .

Definición 2.5 *Un proceso de conteo $\{N(t); t \geq 0\}$ es un **proceso de Poisson homogéneo** con tasa λ , (definido sobre un espacio de probabilidad $(\Omega, F, \mathbb{P})$) si satisface:*

1. $N(0) = 0$.
2. Tiene incrementos independientes y estacionarios.
3. El número de eventos en cualquier intervalo de longitud t tiene distribución Poisson con parámetro λ , es decir,

$$P(N(t+s) - N(s) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots$$

Definición 2.6 Se dice que una función f es $o(h)$ si cumple que

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

Bajo la definición anterior se da una definición equivalente a (2.5):

Definición 2.7 Un proceso de conteo $\{N(t); t \geq 0\}$ es un **proceso de Poisson** con tasa λ , si satisface:

1. $N(0) = 0$.
2. Tiene incrementos independientes y estacionarios.
3. $P(N(h) = 1) = \lambda h + o(h)$.
4. $P(N(h) \geq 2) = o(h)$.

2.1.1. Proceso Poisson no homogéneo

Definición 2.8 Un proceso de conteo $\{N(t); t \geq 0\}$ es un **proceso de Poisson no homogéneo o no estacionario** con función de riesgo $\lambda(t) > 0$, si satisface:

1. $N(0) = 0$.
2. Tiene incrementos independientes.
3. El número de eventos en cualquier intervalo de longitud t tiene distribución Poisson con parámetro $m(t+s) - m(t)$, es decir,

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

Para un proceso Poisson $\{N(t); t \geq 0\}$ con intensidad $\lambda > 0$, se define la función de riesgo o tasa de ocurrencia de eventos como, ver [5]:

$$\lambda(t) = \frac{d}{dt}(E(N(t))) = \frac{d}{dt}(m(t)).$$

A $m(t)$ se le llama, la función media del proceso, es decir,

$$m(t) = \int_0^t \lambda(s) ds.$$

Si $m(t) = \lambda t$, implica que $\lambda(t) = \frac{d}{dt}(\lambda t) = \lambda$.

2.2. Funciones de Supervivencia y Riesgo (Intensidad, Tasa)

El análisis de supervivencia se utiliza para describir el estudio de datos que corresponden al tiempo que transcurre desde un origen hasta que se produce un evento de interés.

Se considera que el tiempo de vida de un objeto, persona o grupo (de personas) se estudia a través de una variable aleatoria T . Esta variable aleatoria puede ser continua, discreta o mixta, no negativa, que representa el tiempo de vida del evento de interés que se encuentra en la población de estudio. Por lo que esta variable aleatoria está definida sobre el intervalo de tiempo $[0; \infty)$, y $f(t)$ denota la función de densidad de probabilidad de T .

Definición 2.9 *La probabilidad de que el evento de interés ocurra hasta un tiempo t está definida como*

$$S(t) = P(T \geq t) = \int_t^\infty f(u) du. \quad (2.2)$$

A esta función se le conoce como función de supervivencia.

Algunas propiedades de la función de supervivencia son las siguientes:

1. La función de supervivencia es complementaria con la función distribución.

$$S(t) = 1 - F(t) = 1 - P(T \leq t). \quad (2.3)$$

2. $S(0) = 1$.

3. $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$.

4. Si $t_1 \leq t_2$, entonces $S(t_1) \geq S(t_2)$.

La función de riesgo es fundamental en análisis de supervivencia. Se le conoce también como la tasa de falla condicional en análisis de confiabilidad, tasa de mortalidad en demografía o función de intensidad en procesos estocásticos.

Definición 2.10 *La función de riesgo $h(t)$, se define como:*

$$\begin{aligned}
 h(t) &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{(\Delta t)P(T \geq t)} \\
 &= \lim_{\Delta t \rightarrow 0} \frac{S(t) - S(t + \Delta t)}{(\Delta t)P(T \geq t)} \\
 &= \frac{f(t)}{1 - F(t)} \\
 &= \frac{f(t)}{S(t)}.
 \end{aligned}$$

Para este trabajo la función de riesgo puede interpretarse como la intensidad con la que ocurren los rebases al umbral de cierto contaminante¹ durante el proceso estocástico.

Observación 2.1 *Cuando la función de riesgo es constante, la probabilidad es independiente del tiempo. Sin embargo, en muchos problemas reales la probabilidad suele cambiar con el tiempo.*

Dado que $f(t) = -S(t)'$ se tiene que,

$$h(t) = -\frac{d}{dt} \log S(t).$$

2.3. Modelos con puntos de cambio

Los modelos con puntos de cambio suelen ocurrir en áreas como la medicina, en donde el interés es basado en el seguimiento de la evolución del paciente después de haberle practicado una cirugía o cierto tratamiento médico; problemas de

¹Concepto teórico para la concentración de una sustancia que representa la transición entre la exposición máxima que no produce efectos adversos y la exposición mínima que produce un efecto adverso bajo condiciones definidas.

control de calidad, fiabilidad y econometría, ver [22]. En general, un modelo con punto de cambio se describe de la siguiente manera: Sea $X = (X_1, X_2, \dots)$ un proceso aleatorio, donde lo que interesa es determinar si las observaciones siguen un proceso homogéneo o no, es decir, cuando las observaciones $X_1, X_2, \dots, X_{\tau-1}$ siguen cierta distribución F_0 , mientras que $X_\tau, X_{\tau+1}, \dots$, siguen cierta distribución F_1 , con el punto de cambio τ desconocido y $F_0 \neq F_1$.

Sea $\{N(t)|t \geq 0\}$ un proceso Poisson homogéneo, definimos la función de riesgo del proceso Poisson con punto de cambio en τ como, ver [22]:

$$\lambda(t) = \begin{cases} \lambda(t|\theta_1) & , \quad 0 \leq t \leq \tau, \\ \lambda(t|\theta_2) & , \quad t > \tau, \end{cases} \quad (2.4)$$

donde τ es desconocido. El análisis del proceso es basado sobre la continuidad de las observaciones en el periodo de tiempo $[0, T]$, donde $N(t) = n$ es el número de ocurrencias del evento de interés.

Teorema 2.1 Sea $\theta = (\theta_1, \theta_2)$, el vector de parámetros del proceso, además τ , el punto de cambio del proceso, si la función de riesgo es definida como en la ecuación 2.4, entonces, la función media del proceso es dada por, ver [3]:

$$m(t) = \begin{cases} m(t|\theta_1) & , \quad 0 \leq t < \tau, \\ m(\tau|\theta_1) + m(t|\theta_2) - m(\tau|\theta_2) & , \quad t > \tau. \end{cases}$$

Demostración 2.1 Sea τ el punto de cambio del proceso con función de riesgo dada por 2.4, usando la definición de la función media del proceso Poisson se tiene que,

si $0 \leq t \leq \tau$:

$$m(t) = \int_0^t \lambda(s) ds = \int_0^t \lambda(s|\theta_1) ds = m(t|\theta_1).$$

Si $t > \tau$:

$$\begin{aligned} m(t) &= \int_0^t \lambda(s) ds \\ &= \int_0^\tau \lambda(s|\theta_1) ds + \int_\tau^t \lambda(s|\theta_2) ds \\ &= m(\tau|\theta_1) + m(t|\theta_2) - m(\tau|\theta_2). \end{aligned}$$

De esta forma la función media del proceso con un punto de cambio esta dada por,

$$m(t) = \begin{cases} m(t|\theta_1) & , \quad 0 \leq t < \tau, \\ m(\tau|\theta_1) + m(t|\theta_2) - m(\tau|\theta_2) & , \quad t > \tau. \end{cases}$$

2.4. Funciones de tasa

En la presente sección se muestran las distribuciones que se utilizan en este trabajo. La importancia de ellas es debido a que se han realizado estudios, en donde se muestra que los datos que son modelados mediante un proceso Poisson no homogéneo, se ajustan con funciones de riesgo dadas por estas distribuciones, ver [2].

2.4.1. Distribución Weibull (W)

La distribución Weibull es una de las funciones más utilizadas en infinidad de aplicaciones debido a su gran versatilidad, por ejemplo al tratar problemas con tiempos de vida en fiabilidad industrial, analisis de supervivencia en areas como la medicina, además de ser una función sencilla de integrar. Su función de riesgo esta dada por, ver[17]:

$$h(t) = \frac{\alpha}{\sigma} \left[\frac{t}{\sigma} \right]^{(\alpha-1)} \quad (2.5)$$

donde $\alpha, \sigma > 0$ y $t > 0$ son parámetros reales.

Las funciones de densidad y de supervivencia son:

$$f(t) = \frac{\alpha}{\sigma} \left[\frac{t}{\sigma} \right]^{(\alpha-1)} \exp^{-(t/\sigma)^\alpha}, \quad t > 0,$$

y

$$S(t) = e^{-(t/\sigma)^\alpha},$$

respectivamente.

La función de riesgo adopta las siguientes formas, dependiendo del parámetro α .

1. Si $\alpha < 1$ la función de riesgo es decreciente, es decir, la tasa de fallo disminuye al aumentar el tiempo.
2. Si $\alpha = 1$ la función de riesgo es constante, por lo que no depende del tiempo. En este caso la distribución Weibull coincide con la distribución exponencial.
3. Si $\alpha > 1$ la función de riesgo es creciente. En particular, si $1 < \alpha < 2$ la función de riesgo crece rápidamente en el origen y muy poco a medida que el tiempo crece. Si $\alpha = 2$ el riesgo crece linealmente con el tiempo y si $\alpha > 2$ crece poco, cuando el tiempo es cercano a cero y cuando crece rápido es por que el tiempo se aleja de cero.

Ahora suponiendo que se tiene un proceso Poisson no homogéneo, con función de riesgo Weibull, se obtiene la función media del proceso dada por,

$$\begin{aligned}
 m(t) &= \int_0^t \lambda(s) ds \\
 &= \int_0^t \frac{\alpha}{\sigma} \left(\frac{s}{\sigma}\right)^{(\alpha-1)} ds \\
 &= \frac{\alpha}{\sigma} \left(\frac{1}{\sigma}\right)^{(\alpha-1)} \int_0^t s^{\alpha-1} ds \\
 &= \left(\frac{s}{\sigma}\right)^{\alpha} \Big|_0^t \\
 &= \left(\frac{t}{\sigma}\right)^{\alpha}.
 \end{aligned}$$

En la Figura 2.1 se muestran gráficos de la función de riesgo Weibull con diferentes valores de α y σ .

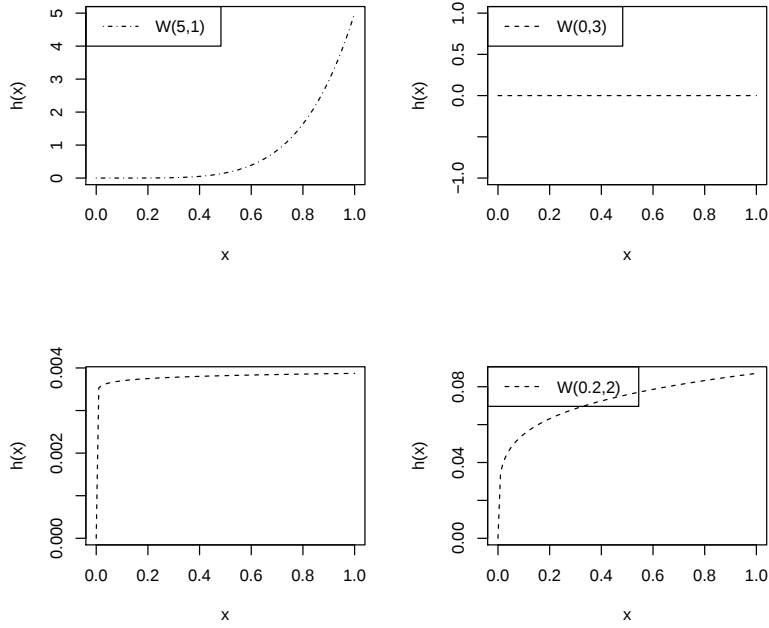


Figura 2.1: Función de riesgo de Weibull

2.4.2. Distribución Musa-Okumoto (MO)

La función de riesgo es dada por, ver [29]

$$h(t) = \frac{\beta}{t + \alpha} \quad (2.6)$$

con $\alpha, \beta > 0$ y $t > 0$.

Su función de densidad es dada por

$$f(t) = \frac{\beta}{(t + \alpha)(t/\alpha + 1)^\beta},$$

y su función de supervivencia,

$$S(t) = \frac{1}{(t/\alpha + 1)^\beta}.$$

Supóngase que se tiene un proceso Poisson no homogéneo, con función de riesgo Musa-Okumoto, la función media del proceso es dada por

$$\begin{aligned}
 m(t) &= \int_0^t \lambda(s) ds \\
 &= \int_0^t \frac{\beta}{s + \alpha} ds \\
 &= \beta \int_0^t \frac{ds}{s + \alpha} \\
 &= \beta \log(s + \alpha) \Big|_0^t \\
 &= \beta \log(t/\alpha + 1).
 \end{aligned}$$

En la Figura 2.2 se observan las formas que tiene la función de riesgo Musa-Okumoto con diferentes valores para α y β .

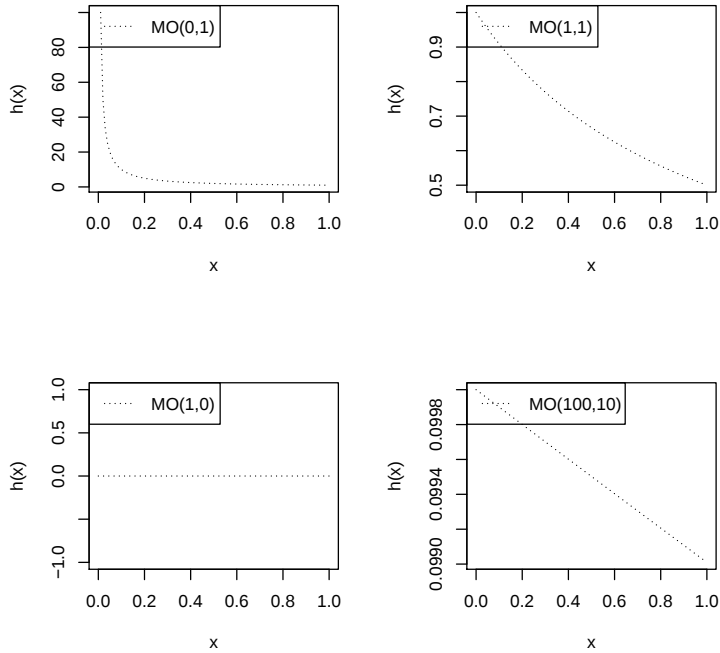


Figura 2.2: Función de riesgo de Musa-Okumoto

2.4.3. Distribución Goel-Okumoto (GO)

La función de riesgo es, ver [27]

$$h(t) = \alpha\beta e^{-\beta t} \quad (2.7)$$

donde, $\alpha, \beta > 0$ y $t > 0$.

Su función de densidad es

$$f(t) = \alpha\beta e^{-\beta t + \alpha(1 - \exp(-\beta t))}.$$

Mientras que la función de supervivencia es:

$$S(t) = e^{\alpha(1 - \exp(-\beta t))}.$$

Cuando se utiliza este modelo, el parámetro α es el número de eventos ocurridos inicialmente, mientras que el parámetro β es la tasa de ocurrencia del evento [27].

Además, si se tiene un proceso Poisson no homogéneo con función de riesgo dada por la distribución Goel-Okumoto, entonces la función media del proceso es

$$\begin{aligned} m(t) &= \int_0^t \lambda(s) ds \\ &= \int_0^t \alpha\beta e^{-\beta s} ds \\ &= \alpha\beta \int_0^t e^{-\beta s} ds \\ &= -\alpha e^{-\beta s} \Big|_0^t \\ &= \alpha(1 - e^{-\beta t}). \end{aligned}$$

En la Figura 2.3 se observan las formas que tiene la función de riesgo Goel-Okumoto con diferentes valores para α y β .

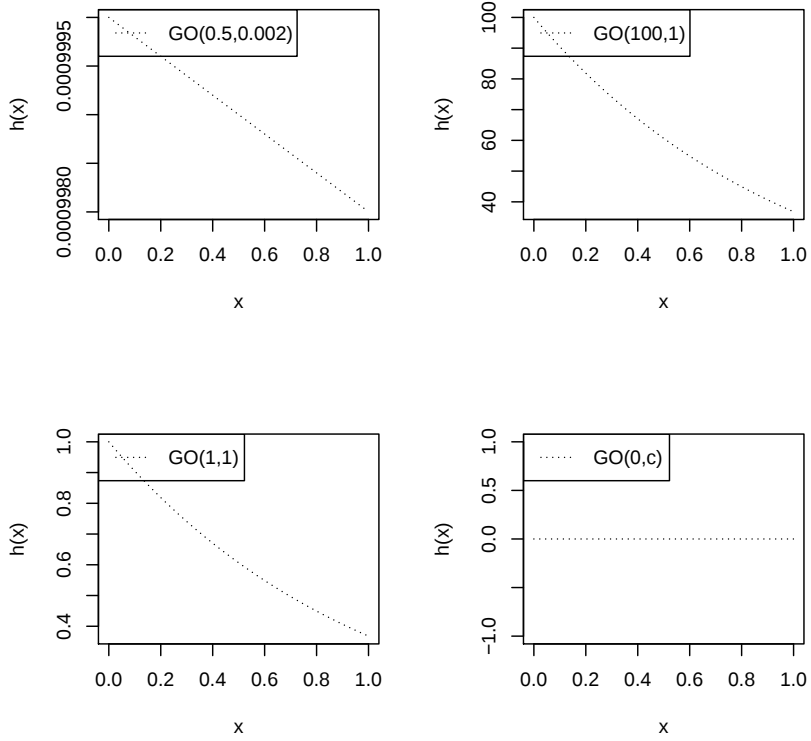


Figura 2.3: Función de riesgo de Goel-Okumoto

2.4.4. Distribución Goel-Okumoto Generalizada (GOG)

Su función de riesgo es dada por, ver [6]

$$h(t) = \alpha\beta\gamma t^{\gamma-1} e^{-\beta t^\gamma}, \quad (2.8)$$

como se puede observar $h(t)$ depende de tres parámetros, $\alpha, \beta, \gamma > 0$ y $t > 0$.

La función de densidad y de supervivencia de esta distribución es dada por

$$f(t) = \alpha\beta\gamma t^{\gamma-1} e^{-\beta t^\gamma + \alpha(1 - \exp(-\beta t^\gamma))},$$

y

$$S(t) = e^{\alpha(1-\exp(-\beta t^\gamma))},$$

respectivamente.

Considérese un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto generalizada, entonces la respectiva función media del proceso es

$$m(t) = \alpha(1 - e^{-\beta t^\gamma}).$$

En efecto, dado que $\lambda(t) = -\frac{d}{dt} \log S(t)$, se tiene que

$$\begin{aligned} m(t) &= \int_0^t \lambda(s) ds \\ &= \int_0^t \left(-\frac{d}{ds} \log(e^{\alpha(1-\exp(-\beta s^\gamma))})\right) ds \\ &= \alpha(1 - e^{-\beta t^\gamma}). \end{aligned}$$

En la Figura 2.4 se observan las formas que tiene la función de riesgo Goel-Okumoto generalizada con diferentes valores para α , β y γ .

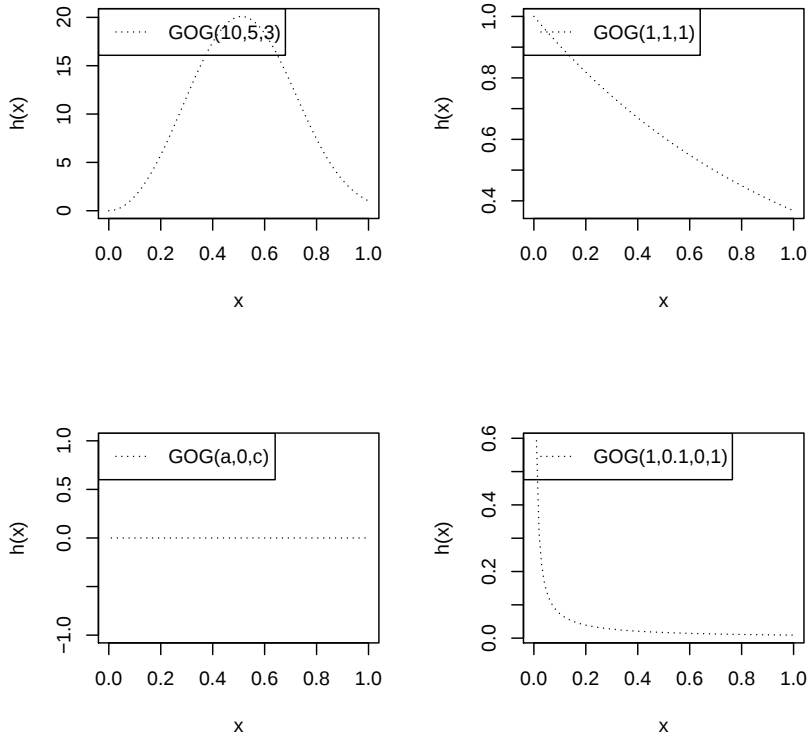


Figura 2.4: Función de riesgo de Goel-Okumoto generalizada ($a, c \geq 0$)

2.4.5. Distribución Weibull-Geométrica (WG)

La función de riesgo es, ver [6],

$$h(t) = (\alpha/\sigma)(t/\alpha)^{\alpha-1}(1 - pe^{-(t/\sigma)^\alpha})^{-1}, \quad (2.9)$$

con $\alpha, \sigma > 0, p \in (0, 1)$ y $t > 0$.

La función de densidad es:

$$f(t) = \frac{\alpha}{\sigma^\alpha}(1-p)t^{\alpha-1}e^{-(t/\sigma)^\alpha}(1 - pe^{-(t/\sigma)^\alpha})^{-2},$$

y la función de supervivencia es dada por

$$S(t) = \frac{(1-p)e^{-(t/\sigma)^\alpha}}{1-pe^{-(t/\sigma)^\alpha}}.$$

La función de riesgo es decreciente para $0 < \alpha < 1$, mientras que, si $\alpha > 1$ la función de riesgo toma diferentes formas.

Suponga que se tiene un proceso Poisson no homogéneo con función de riesgo Weibull-Geométrica, entonces la función media del proceso es dada por

$$m(t) = -\log \left(\frac{(1-p)e^{-(t/\sigma)^\alpha}}{1-pe^{-(t/\sigma)^\alpha}} \right).$$

En efecto, dado que $\lambda(t) = -\frac{d}{dt} \log S(t)$, se tiene,

$$\begin{aligned} m(t) &= \int_0^t \lambda(s) ds \\ &= \int_0^t \left(-\frac{d}{ds} \log \left(\frac{(1-p)e^{-(s/\sigma)^\alpha}}{1-pe^{-(s/\sigma)^\alpha}} \right) \right) ds \\ &= -\log \left(\frac{(1-p)e^{-(t/\sigma)^\alpha}}{1-pe^{-(t/\sigma)^\alpha}} \right). \end{aligned}$$

En la Figura 2.5 se observan las formas que tiene la función de riesgo Weibull-Geométrica con diferentes valores para α , σ y p .

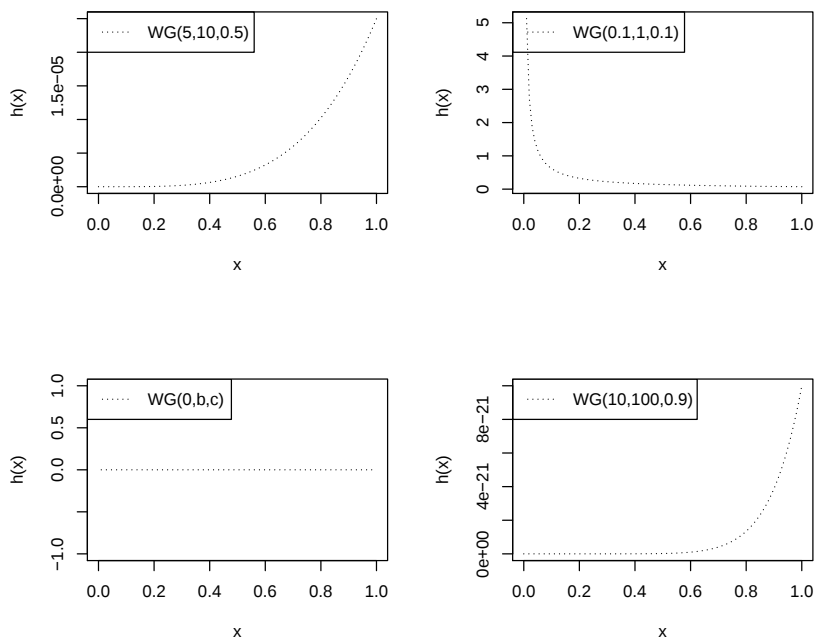


Figura 2.5: Función de riesgo de Weibull-Geométrica ($b, c \geq 0$)

Capítulo 3

Inferencia Bayesiana

En muchos problemas de carácter estadístico, la distribución de probabilidades que generan los datos de cierto experimento es completamente conocida, excepto para los valores de uno o más parámetros. Por ejemplo, supóngase que la distribución de las alturas de cierta cantidad de individuos de alguna población es conocida, esta distribución es una normal con media μ y varianza σ^2 , pero se desconoce la exactitud de los valores de μ y σ^2 .

Como se puede observar, las alturas de los individuos son una muestra aleatoria seleccionada de una población dada, entonces para estas observaciones con alguna información adicional, se realiza inferencia acerca de las alturas de la población, para los valores de μ y σ^2 . En el problema de inferencia estadística, los valores desconocidos como la media y la varianza, son llamados parámetros de la distribución. Al conjunto Θ de todos los posibles valores de un parámetro, digamos θ , o del vector de parámetros $(\theta_1, \theta_2, \dots, \theta_n)$ es llamado *espacio parametral*.

En este capítulo se dan algunos conceptos importantes de la inferencia de parámetros desde el punto Bayesiano, ver [8] y [25]. También se presenta la utilidad del método de Monte Carlo vía Cadenas de Markov (MCMC, por sus siglas en inglés) para obtener información sobre la llamada distribución a posteriori en

un análisis Bayesiano, ver [34], [38].

Sea D los datos observados, y θ los parámetros del modelo y/o los datos faltantes. La inferencia bayesiana se basa en información obtenida de la distribución a posteriori para obtener los parámetros. Esta distribución denotada por $p(\theta|D)$, es obtenida a partir de una distribución a priori sobre los parámetros, indicada por $p(\theta)$ que refleja el conocimiento del investigador acerca de θ , y la función de verosimilitud del modelo es indicada por $l(D|\theta)$.

Así, usando propiedades de probabilidad condicional:

$$p(\theta|D) = \frac{p(\theta)p(D|\theta)}{\int p(\theta)p(D|\theta)d\theta},$$

donde $p(D|\theta)$, es la distribución de probabilidad de los datos observados D , que dependen del parámetro θ .

Cualquier característica de la distribución a posteriori puede ser obtenida a través de inferencia bayesiana. Algunas de estas características pueden ser los momentos centrales o de dispersión, cuantiles, intervalos de credibilidad, entre otros. Las cantidades anteriores pueden expresarse en términos de la esperanza de la distribución a posteriori del parámetro θ .

La esperanza a posteriori de una función $f(\theta)$, es dada por:

$$E[f(\theta)|D] = \frac{\int f(\theta)p(\theta)p(D|\theta)d\theta}{\int p(\theta)p(D|\theta)d\theta}. \quad (3.1)$$

La obtención de esta esperanza es la fuente de la mayor parte de las dificultades prácticas en la inferencia Bayesiana, especialmente en altas dimensiones. En la mayoría de las aplicaciones, la evaluación analítica de $E[f(\theta)|D]$ es imposible, así que han surgido enfoques alternativos, en las que se incluyen las evaluaciones numéricas como la aproximación de Laplace y la integración de Monte Carlo, incluyendo métodos Monte Carlo vía cadenas de Markov (MCMC).

3.1. Distribucion a priori (Inicial)

Definición 3.1 Sea $X = \{x_1, x_2, \dots, x_n\}$ una muestra aleatoria y $\theta \in \Theta$ el parámetro de la distribución que produce la muestra X . Si se cumple que el parámetro θ es una variable aleatoria, entonces a la distribución del parámetro que no toma en cuenta la muestra X , se le conoce como **distribución a priori (inicial)**.

Notación 1 Dada una muestra aleatoria $X = \{x_1, x_2, \dots, x_n\}$, si $\theta \in \Theta$ es el parámetro del modelo que describe el comportamiento de la muestra aleatoria, entonces $p(\theta)$ denota la distribución a priori del parámetro θ .

Definición 3.2 Si la distribución a priori del parámetro θ , no favorece a algún valor en el espacio parametral Θ , a esta distribución se le llama a priori no informativa.

La distribución a priori de un parámetro θ puede ser especificada por intuición del investigador o basada en estudios previos, sin embargo, si no existe información alguna, se suele utilizar una distribución uniforme, como la a priori del parámetro θ , o una a priori no informativa.

Ejemplo 3.1 (Ver [21]), Sea $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$, entonces una a priori no informativa suele ser:

$$p(\theta_i) = \frac{1}{n}, \quad i = 1, \dots, k. \quad (3.2)$$

Si el espacio parametral es continuo y acotado, digamos $\Theta = [a, b]$, suele ser:

$$p(\theta_i) = \frac{1}{b-a}, \quad i = 1, \dots, k. \quad (3.3)$$

Si $\Theta = (-\infty, \infty)$, una distribución a priori podría ser:

$$p(\theta_i) = c, \text{ donde } c > 0, \text{ es una constante e } i = 1, \dots, k. \quad (3.4)$$

En el ejemplo anterior, el espacio parámetro es toda la recta real, la distribución a priori, no es una distribución propia, ya que:

$$\int_{-\infty}^{\infty} p(\theta) d\theta = \int_0^{\infty} c d\theta = \infty.$$

Sin embargo, este tipo de a priori, suelen ser útiles para ciertos casos.

3.2. Función de verosimilitud

Definición 3.3 Sea $X = (x_1, x_2, \dots, x_n)$ una muestra aleatoria con función de densidad conjunta dada por $f(X|\theta)$, donde θ es desconocido y con valores en el espacio parametral Θ . La **función de verosimilitud** es proporcional a $f(X|\theta)$ y se denota por $l(X|\theta)$, ver [8].

Cuando X es una sucesión de variables aleatorias independientes, la función de verosimilitud es dada por

$$l(X|\theta) \propto \prod_{i=1}^n f(x_i|\theta). \quad (3.5)$$

3.3. Distribución a posteriori (Final)

Definición 3.4 Sea $X = (x_1, x_2, \dots, x_n)$ una muestra aleatoria y supongamos $\theta \in \Theta$, el parámetro de la distribución que produjo la muestra X . A la distribución del parámetro dada la muestra, denotada por $p(\theta|X)$, se le conoce como **distribución a posteriori**.

Además, esta distribución es obtenida de la siguiente forma

$$p(\theta|X) = \frac{p(\theta)p(X|\theta)}{p(X)}, \quad (3.6)$$

se observa que, $p(X)$ no depende del parámetro θ . Así que la relación (4.6), se suele simplificar como

$$p(\theta|X) \propto p(\theta)p(X|\theta) \propto l(X|\theta)p(\theta), \quad (3.7)$$

con $\propto$ como el símbolo de proporcionalidad, es decir, la distribución a posteriori es proporcional al producto de la distribución a priori del parámetro θ , con la distribución de la muestra dado el parámetro θ .

3.4. Criterio Bayesiano para Comparación de Modelos

En la literatura existen diversos criterios para comparar modelos, factor de Bayes, método Bayesian Model Averaging (BMA), Akaike, Bayesian Information Criterion (BIC), Arbol de regresión Bayesiana, entre otros, ver [9], [23]. En este trabajo se propone comparar los modelos presentados usando, el criterio de información de desviación (DIC, por sus siglas en inglés). La desviación es definida por $Dev(\theta) = -\log[l(D|\theta)] + c$, donde θ es el vector de parámetros del modelo, D son los datos observados, $l(D|\theta)$ es la función de verosimilitud del modelo y c es una constante que no es necesaria al momento de comparar los modelos. El DIC es dado por, ver [43]:

$$DIC = Dev(\hat{\theta}) + 2n_D,$$

donde $Dev(\hat{\theta})$ es la desviación evaluada en la media a posteriori ($\hat{\theta}$) de θ y $n_D = E[Dev(\theta)] - Dev(\hat{\theta})$, que corresponde al número de parámetros del modelo. Valores pequeños de DIC indican, mejores modelos.

3.5. Método Monte Carlo

El método Monte Carlo es considerado como un método usado para aproximar integrales de funciones que no son tan sencillas de evaluar con métodos comunes, dando soluciones aproximadas. Como un ejemplo tenemos el problema del cálculo de la esperanza que ocurre tanto en distribuciones de carácter frecuentista como bayesiana, veamos el caso general. Sea X un vector de k variables aleatorias (en el caso de inferencia bayesiana, X denota el vector de parámetros y/o los datos perdidos), con distribución $\pi(\cdot)$ (como la distribución a posteriori). Lo que interesa calcular es:

$$E[f(X)] = \int f(x)\pi(x)dx,$$

para una función $f(\cdot)$ de interés. Aquí se podría suponer que la distribución de X es conocida, salvo una constante normalizadora. Esto es,

$$\int \pi(x)dx,$$

es desconocida. Esta situación es común en la práctica, por ejemplo en inferencia bayesiana se sabe que $p(\theta|D) \propto p(\theta)p(D|\theta)$. En este caso puede ser difícil evaluar la constante normalizadora $\int p(\theta)p(D|\theta)d\theta$, es decir, si $\pi \propto \hat{f}$, existe k tal que $\pi(x) = k\hat{f}(x)$ y $\int \hat{f}(x)dx = \frac{1}{k}$ con k desconocido o difícil de obtener.

La integración Monte Carlo evalúa $E[f(X)]$, por medio de muestras $\{x_t : t = 1, 2, \dots, n\}$ de $\pi(\cdot)$ y luego se aproxima mediante el promedio,

$$E[f(X)] \approx \frac{1}{n} \sum_{i=1}^n f(x_i).$$

Así que la media de la población de $f(X)$ se estima por una media de la muestra. Cuando las muestras x_t son independientes, la Ley de los Grandes Números, da garantía de que la aproximación puede hacerse tan precisa como se desea, mediante el aumento del tamaño de la muestra n .

En general, las muestras independientes de $\{x_t\}$ no son factibles, esto se debe a que la distribución $\pi(\cdot)$, puede ser atípica. Sin embargo, la independencia de la muestra $\{x_t\}$ puede no existir. Aún así, la muestra puede ser generada a través de algún proceso, que produzca muestras a partir de $\pi(\cdot)$. Una forma de hacerlo es a través de una cadena de Markov que tiene a $\pi(\cdot)$ como su distribución estacionaria. A este método se conoce como Método Monte Carlo vía Cadenas de Markov (MCMC). En términos generales se da la siguiente definición, ver [38].

Definición 3.5 *El Método Monte Carlo vía Cadenas de Markov para la simulación de una función de distribución f , es algún método que produce una cadena de Markov ergódica $\{X_t\}$ cuya distribución estacionaria es f .*

Uno de los principales algoritmos MCMC conocidos, es el de Metropolis-Hasting. El algoritmo fue presentado inicialmente por Metropolis et. al. en 1953 y extendido en 1970 por Hastings. La dinámica Metropolis-Hasting considera las siguientes condiciones:

1. Sea $S = \{0, 1, 2, \dots\}$ un espacio de estados contable y π una distribución definida sobre S .

2. Tome una función de distribución q sobre todos los estados de S (proposición de cambio) de la cual es relativamente fácil generar valores.
3. Sea $\alpha : S \times S \rightarrow (0, 1)$ la función de aceptación de cambio.

El algoritmo Metropolis-Hasting es el siguiente: Sea x un estado inicial, entonces

1. Se propone el cambio de $x \rightarrow y$, siguiendo la distribución propuesta $q(x, \cdot)$.
2. Se acepta el cambio con probabilidad $\alpha(x, y)$. De lo contrario, no se modifica el valor de x .

$$p(x, y) = \begin{cases} q(x, y)\alpha(x, y), & \text{si } x \neq 0, \\ q(x, y) + \sum_{y: y \neq x} q(x, y)[1 - \alpha(x, y)], & \text{en otro caso.} \end{cases}$$

La dinámica del cambio consiste en que se cumpla la condición de reversibilidad:

$$\text{Para todo } x, y \in S, x \neq y : \pi(x)q(x, y)\alpha(x, y) = \pi(y)q(y, x)\alpha(y, x),$$

equivalentemente,

$$\frac{\alpha(x, y)}{\alpha(y, x)} = \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}.$$

De esta forma se tiene que

$$\alpha(x, y) = \min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\}. \quad (3.8)$$

Para determinar el estado X_{n+1} , a partir del estado $X_n = x$, se toma lo siguiente:

1. Generar un estado y que tenga como estado inicial a x con distribución $q(x, y)$, la cual es una distribución que se escoge libremente, de tal forma que satisfaga las siguientes propiedades:

- $q(x, y) = 0 \Rightarrow q(y, x) = 0$.
- $q(x, y)$ es la matriz de transición de una cadena de Markov en S .

2. Aceptación con probabilidad

$$\alpha(y|x) = \min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\},$$

hacemos $X_{n+1} = y$ (aceptamos y), en otro caso $X_n = x$ (rechazamos y).

Observación 3.1 Como los pasos de aceptación y rechazo son independientes, se tiene que la probabilidad de transición para ir de x a y es:

$$P(X_{n+1} = y | X_n = x) = q(x, y)\alpha(x, y),$$

con $x \neq y$.

Teorema 3.1 Sea π una distribución de probabilidad dada. La cadena de Markov simulada por medio del algoritmo de Metropolis-Hasting es reversible. Si además dicha cadena es irreducible, aperiódica y recurrente positiva, entonces dicho algoritmo determina a una única cadena de Markov ergódica con distribución estacionaria π .

Demostración 3.1 Suponga, sin pérdida de generalidad que

$$\pi(y)q(y, x) > \pi(x)q(x, y).$$

Como

$$p(x, y) = q(x, y)\alpha(x, y) = q(x, y) \min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\} = q(x, y).$$

Por otro lado, se tiene que

$$\begin{aligned} p(y, x) &= q(y, x)\alpha(y, x) \\ &= q(y, x) \min\left\{1, \frac{\pi(x)q(x, y)}{\pi(y)q(y, x)}\right\} = q(y, x) \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)} \\ &= \frac{\pi(x)}{\pi(y)} q(x, y). \end{aligned}$$

Por tanto, se tiene que,

$$\pi(x)p(x, y) = \pi(y)p(y, x).$$

3.5.1. Muestreador de Gibbs

Suponga que se desea generar una muestra aleatoria con distribución $\pi(\theta)$, donde $\theta = (\theta_1, \theta_2, \dots, \theta_n) \in \mathbb{R}^n$ es el vector de parámetros, ver [38]. Además suponga que se conocen todas las distribuciones condicionales,

$$P(\theta_i | \theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n), \quad i \in \{1, 2, \dots, n\}.$$

En diversas ocasiones el generar una muestra de una distribución dada directamente es complicada, es por ello que suele utilizarse el muestreador de Gibbs, ver [38].

El muestreador de Gibbs es una versión del método de Metropolis-Hastings donde se toma a la distribución condicional obtenida de la distribución original de donde se quería simular el parámetro θ_i . En términos generales se describe el algoritmo del muestreador de Gibbs:

- Sea $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_n^{(0)})$, un vector inicial.
- Dado el vector $\theta^{(j-1)}$ en la $(j-1)$ -ésima iteración del algoritmo, genera n muestras aleatorias de la siguiente manera:

$$\begin{aligned}
 \theta_1^{(j)} &\leftarrow \pi(\theta_1 | \theta_2^{(j-1)}, \theta_3^{(j-1)}, \dots, \theta_n^{(j-1)}), \\
 \theta_2^{(j)} &\leftarrow \pi(\theta_2 | \theta_1^{(j)}, \theta_3^{(j-1)}, \dots, \theta_n^{(j-1)}), \\
 &\vdots \\
 \theta_i^{(j)} &\leftarrow \pi(\theta_i | \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{i-1}^{(j)}, \theta_{i+1}^{(j-1)}, \dots, \theta_n^{(j-1)}), \\
 &\vdots \\
 \theta_n^{(j)} &\leftarrow \pi(\theta_n | \theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_{n-1}^{(j)}).
 \end{aligned}$$

- Tomar $\theta^{(j)} = (\theta_1^{(j)}, \theta_2^{(j)}, \dots, \theta_n^{(j)})$ y volver al punto anterior.

Como se observa, el muestreador de Gibbs realiza muestras a partir de las distribuciones condicionales, cambiando el valor de cada parámetro a partir de su distribución condicional dado el valor actual del resto de los parámetros.

3.5.2. El muestreador de Gibbs como caso particular del algoritmo de Metropolis-Hastings

Suponga que se desea obtener una muestra de variables aleatorias con cierta distribución $\pi(\theta)$, donde $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ es el vector de parámetros. Para obtener dicha muestra se utiliza el algoritmo de Metropolis-Hastings, donde se debe

suponer que pueden obtener muestras aleatorias de las distribuciones condicionales del muestreador de Gibbs,

$$\pi_i = \pi(\theta_i | \theta_1, \theta_2, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_n),$$

entonces la opción para generar la matriz de transición q es tomar $q(i, j) = \pi_j$. De forma que la probabilidad de aceptación será dada por:

$$\min\left\{1, \frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}\right\} = \min\left\{1, \frac{\pi(y)\pi(x)}{\pi(x)\pi(y)}\right\} = 1.$$

En este caso, todas las propuesta generadas por el algoritmo mediante q serán aceptadas con probabilidad 1, es decir, el algoritmo generado es el muestreador de Gibbs, ver [38].

Capítulo 4

Aplicación a Niveles de Ozono en la Ciudad de Puebla

En el presente capítulo se presenta un método matemático para estudiar el comportamiento del contaminante ozono en la Ciudad de Puebla. A lo largo de la historia del planeta se han producido diversos cambios climáticos. Sin embargo, en las últimas décadas se gesta un nuevo cambio a consecuencia directa de factores humanos, teniendo como punto de partida la revolución industrial de finales del siglo XIX, donde se observa que el consumo de combustibles fósiles, la degradación de los recursos naturales, la descomposición de residuos orgánicos, los cambios de uso de suelo; vienen contribuyendo y alterando significativamente las propiedades del clima.

Uno de los principales problemas que se tiene como consecuencia de los factores humanos se encuentra en la calidad del aire que se respira. Esto aunado al aumento en el consumo de combustibles fósiles que incrementan la expulsión de sustancias tóxicas, como lo son, dióxido de nitrógeno (NO_2), Dióxido de carbono (CO_2),

Partículas suspendidas (PM y PM2.5), Ozono (O_3), entre otras.

En muchas ciudades se estudia este fenómeno ambiental dada la importancia que tiene en efectos adversos a la salud. En particular, el estudio del ozono se realiza en zonas urbanas por las grandes cantidades de emisiones al ambiente de sus precursores, y los efectos relacionados a muertes prematuras.

A una altura de 20 km sobre la superficie se encuentra la llamada capa de ozono (O_3) u ozono estratosférico. Esta capa de ozono esta de forma natural y actúa de forma benéfica absorbiendo radiación UV proveniente del Sol y evitando así que llegue a la superficie de la Tierra. Sobre la Tropósfera (capa de la atmósfera situada sobre la superficie de la Tierra, hasta una distancia de unos 10 Km) se sitúa el ozono denominado ozono troposférico, que es consecuencia de la acción del hombre. Este ozono se puede encontrar en concentraciones superiores a las naturales, actuando entonces como un contaminante atmosférico por sus efectos nocivos sobre el medio [39].

4.1. Descripción del modelo

El problema a considerar es estudiar el número de veces en el que cierto umbral ambiental es rebasado en un intervalo de tiempo dado. Considere la siguiente formulación matemática: Sea $N(t) \geq 0$, $t \geq 0$ el número de veces que un estándar ambiental de cierto contaminante ha sido rebasado en un intervalo de tiempo $[0, t)$. Considérese que el comportamiento observado es un proceso Poisson no homogéneo, ver [24]¹.

Dado que se tiene un proceso Poisson no homogéneo, se requiere de una función de riesgo $\lambda(t) > 0$, $t > 0$. Cuando se conoce el comportamiento de $\lambda(t)$, entonces la probabilidad de que el estándar ambiental sea violado en el intervalo de tiempo $[t_1, t_2)$, podrá ser calculado a través de:

$$P(N(t_2) - N(t_1) = n) = e^{-[m(t_2) - m(t_1)]} \frac{[m(t_2) - m(t_1)]^n}{n!}, \quad t_2 > t_1,$$

¹El sustento de tomar este conjunto por medio de un proceso Poisson, utiliza conceptos como procesos puntuales y teoría de valores extremos.

donde $m(t) = \int_0^t \lambda(s)ds$, es la función media del proceso.

Observación 4.1 *Obsérvese que la función de riesgo puede depender de un vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_n)$. Por lo tanto, el problema se reduce a estimar el vector de parámetros que mejor describa el comportamiento de los datos observados.*

4.2. Inferencia Bayesiana sobre el modelo planteado

Sea $D = \{d_1, d_2, \dots, d_K\}$ el conjunto de días en los cuales el umbral ambiental asignado a un contaminante de interés ha sido rebazado dentro del intervalo de tiempo $[0, T]$. Sea $\lambda(t)$ la función de riesgo respecto al proceso Poisson no homogéneo que describe el conjunto de observaciones D , que depende de un vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_n)$. Utilizando inferencia bayesiana, existe una relación entre la distribución a priori del vector de parámetros θ con la distribución a posteriori considerando las observaciones dadas D . Es decir,

$$p(\theta|D) \propto l(D|\theta)p(\theta).$$

Teorema 4.1 *Sea $D = \{d_1, d_2, \dots, d_k\}$ el conjunto de días en los cuales un estándar ambiental de cierto contaminante a sido violado durante el intervalo de tiempo $[0, K]$, contados por un proceso Poisson no homogéneo con función de riesgo $\lambda(t) > 0$, y dependiente de un vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_m) \in \mathbb{R}^m$, entonces bajo el modelo Poisson no homogéneo, la función de verosimilitud $l(D|\theta)$ es dada por:*

$$l(D|\theta) = \left[\prod_{i=0}^k \lambda(d_i) \right] e^{-m(K)}. \quad (4.1)$$

Demostración 4.1 *Para la demostración se usa la definición de proceso Poisson no homogéneo, y se puede consultar en, ver [18].*

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

Por la definición se tiene que

$$P(N(d_1) = 0) = e^{-m(d_1)} \frac{(m(d_1))^0}{0!},$$

asi, la función de densidad respecto a la ocurrencia del primer rebase es dada por,

$$\begin{aligned} f(d_1) &= -\frac{d}{dd_1} P(N(d_1) = 0) \\ &= -\frac{d}{dd_1} e^{-m(d_1)} \\ &= \left[\frac{d}{dd_1} m(d_1) \right] e^{-m(d_1)} \\ &= \lambda(d_1) e^{-m(d_1)}. \end{aligned}$$

Para el caso entre dos observaciones de rebase

$$\begin{aligned} P(N(d_2) - N(d_1) = 0) &= e^{-(m(d_2) - m(d_1))} \frac{(m(d_2) - m(d_1))^0}{0!} \\ &= e^{-(m(d_2) - m(d_1))}. \end{aligned}$$

De esta forma, la función de densidad respecto al tiempo entre el primer y segundo rebase, esta dada por,

$$\begin{aligned} f(d_2 - d_1 | d_1) &= -\frac{d}{dd_2} P(N(d_2) - N(d_1) = 0) \\ &= -\frac{d}{dd_2} e^{-(m(d_2) - m(d_1))} \\ &= \left[\frac{d}{dd_2} (m(d_2) - m(d_1)) \right] e^{-(m(d_2) - m(d_1))} \\ &= \lambda(d_2) e^{-(m(d_2) - m(d_1))}. \end{aligned}$$

De manera general se tiene que la función de densidad respecto al tiempo entre la $(i - 1)$ e i -ésima observación de rebase,

$$\begin{aligned} f(d_i - d_{i-1} | d_{i-1} - d_{i-2}) &= -\frac{d}{dd_i} P(N(d_i) - N(d_{i-1}) = 0) \\ &= -\frac{d}{dd_i} e^{-(m(d_i) - m(d_{i-1}))} \\ &= \left[\frac{d}{dd_i} (m(d_i) - m(d_{i-1})) \right] e^{-(m(d_i) - m(d_{i-1}))} \\ &= \lambda(d_i) e^{-(m(d_i) - m(d_{i-1}))}. \end{aligned}$$

Así, la función de verosimilitud tiene la forma

$$\begin{aligned}
 l(d_1, d_2, d_3, \dots, d_k | \lambda) &\propto f(d_1) \prod_{i=2}^k f(d_i - d_{i-1} | d_{i-1} - d_{i_2}) \\
 &= \lambda(d_1) e^{-m(d_1)} \prod_{i=2}^k \left[\lambda(d_i) e^{-(m(d_i) - m(d_{i-1}))} \right] \\
 &= \lambda(d_1) e^{-m(d_1)} \left[\prod_{i=2}^k \lambda(d_i) \right] e^{-\sum_{i=2}^K m(d_i) + \sum_{i=2}^K m(d_{i-1})} \\
 &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-\sum_{i=1}^K m(d_i) + \sum_{i=2}^K m(d_{i-1})} \\
 &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(d_K)}.
 \end{aligned}$$

Cuando existe un punto de cambio se tiene el siguiente resultado, ver [3]:

Corolario 4.1 Sea $D = \{d_1, d_2, \dots, d_K\}$ los días en que el rebase del umbral ambiental ocurrio, asuma que el número de rebases sigue un proceso Poisson no homogéneo con función de riesgo

$$\lambda(t) = \begin{cases} \lambda(t|\theta_1) & , \quad 0 \leq t \leq \tau, \\ \lambda(t|\theta_2) & , \quad t > \tau. \end{cases} ,$$

dependiente de un vector de parámetros $\theta = (\theta_1, \theta_2)$, además de presentar un punto de cambio τ , se tiene que bajo el modelo Poisson no homogéneo, la función de verosimilitud $l(D|\theta)$ es dada por:

$$\begin{aligned}
 l(\omega|D) &= \left(\prod_{i=1}^{N_\tau} \lambda(d_i|\omega_1) \right) e^{-m(\tau|\omega_1)} \\
 &\quad \prod_{i=N_\tau+1}^K \lambda(d_i|\omega_2) e^{-m(K|\omega_2) - m(\tau|\omega_2)}.
 \end{aligned}$$

Donde, $\omega_i = (\theta_i, \tau)$, $i = 1, 2$, corresponde al parámetro antes y después del punto de cambio.

Demostración 4.2 *Para la demostración se sigue un planteamiento similar al teorema anterior, usando la definición de proceso Poisson no homogéneo sobre los intervalos $[0, N_\tau]$ y $[N_\tau + 1, K]$.*

Corolario 4.2 *Sea $D = \{d_1, d_2, \dots, d_k\}$, $k > 0$ el conjunto de días en los cuales un estándar ambiental de cierto contaminante a sido violado. Sea $\lambda(t)$ la función de riesgo del proceso Poisson no homogéneo que registra los días en los cuales el estándar fue sobrepasado, dependiente de un vector de parámetros $\theta = (\theta_1, \theta_2, \dots, \theta_m) \in \mathbb{R}^m$, entonces,*

$$p(\theta|D) \propto \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\theta). \quad (4.2)$$

Demostración 4.3 *Usando inferencia bayesiana y el Teorema 5.1.*

Debido a que las funciones de densidad a posteriori con respecto a los parámetros de la función de riesgo $\lambda(t)$ pueden resultar en distribuciones en las cuales es difícil calcular medias, varianzas u otras características. Se utiliza el método Monte Carlo vía Cadenas de Markov, para simular una cadena de Markov a partir de una función de intensidad dada, tal que su distribución estacionaria sea $p(\theta|D)$. De esta manera una vez que la cadena alcance la convergencia, los estados en los que se sitúe la cadena podrán ser tomados como observaciones de la distribución a posteriori.

La forma en la que se estiman las densidades a posteriori, es mediante el programa OpenBugs, ver[30]. Para cada una de las funciones de intensidad se toma el vector de parámetros a estimar y para cada parámetro se elige una distribución a priori no informativa, con 5 cadenas para verificar la convergencia o no del modelo.

Para el calentamiento se utilizan por lo general 40000 iteraciones, de las cuales se desechan 20000. Los resultados que nos arroja el programa se dan más adelante.

4.3. Modelos utilizados

Los modelos que son utilizados en el presente trabajo, generalmente son usados en los modelos de confiabilidad.

4.3.1. Weibull (W)

La función de riesgo de la distribución Weibull está dada por,

$$h(t) = \frac{\alpha}{\sigma} \left[\frac{t}{\sigma} \right]^{(\alpha-1)}, \quad t > 0.$$

Por lo tanto, el vector de parámetros que se desea estimar es $\theta = (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$. Además, si se tiene un proceso Poisson no homogéneo con función de riesgo Weibull, su función media es dada por

$$m(t) = \left[\frac{t}{\sigma} \right]^\alpha.$$

Considerando lo anterior se tiene el siguiente resultado, ver [3].

Teorema 4.2 *Considere un proceso Poisson no homogéneo con función de riesgo Weibull, además, si se asume que los parámetros de la función de riesgo son a priori independientes, entonces la distribución a posteriori de los parámetros está dada por,*

$$p(\alpha, \sigma | D) = \left(\frac{\alpha}{\sigma^\alpha} \right)^k \left[\prod_{i=1}^k d_i^{\alpha-1} \right] e^{-(T/\sigma)^\alpha} p(\alpha) p(\sigma). \quad (4.3)$$

Demostración 4.4 *Sustituyendo la forma de la función de tasa así como la función de verosimilitud en (5.2), se tiene*

$$\begin{aligned} p(\alpha, \sigma | D) &\propto l(\theta | D) p(\alpha, \sigma) \\ &= l(\theta | D) p(\alpha) p(\sigma) \\ &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\alpha) p(\sigma) \\ &= \left[\prod_{i=1}^k \left(\frac{\alpha}{\sigma} \right) \left(\frac{d_i}{\sigma} \right)^{\alpha-1} \right] e^{-(T/\sigma)^\alpha} p(\alpha) p(\sigma) \\ &= \left(\frac{\alpha}{\sigma^\alpha} \right)^k \left[\prod_{i=1}^k d_i^{\alpha-1} \right] e^{-(T/\sigma)^\alpha} p(\alpha) p(\sigma). \end{aligned}$$

4.3.2. Musa-Okumoto (MO)

La función de riesgo de la distribución Musa-Okumoto está dada por,

$$h(t) = \frac{\beta}{t + \alpha}, \quad t > 0.$$

Por lo tanto, el vector de parámetros que se desea estimar es $\theta = (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$.

Además, si se tiene un proceso Poisson no homogéneo con función de riesgo Musa-Okumoto, su función media es dada por,

$$m(t) = \beta \log(t/\alpha + 1).$$

Considerando lo anterior se tiene el siguiente resultado, ver [21].

Teorema 4.3 *Considere un proceso Poisson no homogéneo con función de riesgo Musa-Okumoto, además, si se asume que los parámetros de la función de riesgo son a priori independientes, entonces la distribución a posteriori de los parámetros está dada por,*

$$p(\alpha, \beta|D) = \beta^k \left[\prod_{i=1}^k (1/d_i + \alpha) \right] e^{-\log(T/\alpha+1)} p(\alpha)p(\beta). \quad (4.4)$$

Demostración 4.5 *Sustituyendo la forma de la función de tasa así mismo la función de verosimilitud en (5.2), se tiene*

$$\begin{aligned} p(\alpha, \beta|D) &\propto l(\theta|D)p(\alpha, \beta) \\ &= l(\theta|D)p(\alpha)p(\beta) \\ &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\alpha)p(\beta) \\ &= \left[\prod_{i=1}^k \left(\frac{\beta}{d_i + \alpha} \right) \right] e^{-\beta \log(T/\alpha+1)} p(\alpha)p(\beta) \\ &= \beta^k \left[\prod_{i=1}^k (1/d_i + \alpha) \right] e^{-\log(T/\alpha+1)} p(\alpha)p(\beta). \end{aligned}$$

4.3.3. Goel-Okumoto (GO)

La función de riesgo de la distribución Goel-Okumoto está dada por,

$$h(t) = \alpha\beta e^{-\beta t}, \quad t > 0.$$

Por lo tanto, el vector de parámetros que se desea estimar es $\theta = (\alpha, \beta) \in \mathbb{R}^+ \times \mathbb{R}^+$. Además, si se tiene un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto, su función media es dada por

$$m(t) = \alpha(1 - e^{-\beta t}).$$

Considerando lo anterior se tiene el siguiente resultado, ver [3].

Teorema 4.4 *Considere un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto, además, si se asume que los parámetros de la función de riesgo son a priori independientes, entonces la distribución a posteriori de los parámetros está dada por,*

$$p(\alpha, \beta | D) = (\alpha\beta)^k e^{\alpha(1-e^{-\beta T})} e^{-\beta \sum_{i=1}^k d_i} p(\alpha)p(\beta). \quad (4.5)$$

Demostración 4.6 *Sustituyendo la forma de la función de tasa así como la función de verosimilitud en (5.2), se tiene*

$$\begin{aligned} p(\alpha, \beta | D) &\propto l(\theta | D) p(\alpha, \beta) \\ &= l(\theta | D) p(\alpha) p(\beta) \\ &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\alpha) p(\beta) \\ &= \left[\prod_{i=1}^k \alpha\beta e^{-\beta d_i} \right] e^{-\alpha(1-e^{-\beta T})} p(\alpha) p(\beta) \\ &= (\alpha\beta)^k e^{-\alpha(1-e^{-\beta T})} e^{-\beta \sum_{i=1}^k d_i} p(\alpha) p(\beta). \end{aligned}$$

4.3.4. Goel-Okumoto Generalizada (GOG)

La función de riesgo de la distribución Goel-Okumoto generalizada está dada por,

$$h(t) = \alpha\beta\gamma t^{\gamma-1} e^{-\beta t^\gamma}, \quad t > 0.$$

Por lo tanto, el vector de parámetros que se desea estimar es $\theta = (\alpha, \beta, \gamma) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$. Además, si se tiene un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto generalizada, su función media es dada por

$$m(t) = \alpha(1 - e^{-\beta t^\gamma}).$$

Considerando lo anterior se tiene el siguiente resultado, ver [21].

Teorema 4.5 *Considere un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto generalizada, además, si se asume que los parámetros de la función de intensidad son a priori independientes, entonces la distribución a posteriori de los parámetros está dada por,*

$$p(\alpha, \beta, \gamma | D) = (\alpha\beta\gamma)^k \left[\prod_{i=1}^k d_i^{\gamma-1} \right] e^{-\alpha(1-e^{-(\beta T^\gamma)})} e^{-\beta \sum_{i=1}^k (d_i^\gamma)} p(\alpha)p(\beta)p(\gamma). \quad (4.6)$$

Demostración 4.7 *Sustituyendo la forma de la función de tasa así como la función de verosimilitud en (5.2), se tiene*

$$\begin{aligned} p(\alpha, \beta, \gamma | D) &\propto l(\theta | D) p(\alpha, \beta, \gamma) \\ &= l(\theta | D) p(\alpha) p(\beta) \\ &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\alpha) p(\beta) p(\gamma) \\ &= \left[\prod_{i=1}^k \alpha \beta \gamma d_i^{\gamma-1} e^{-\beta d_i^\gamma} \right] e^{-\alpha(1-e^{-(\beta T^\gamma)})} p(\alpha) p(\beta) p(\gamma) \\ &= (\alpha\beta\gamma)^k \left[\prod_{i=1}^k d_i^{\gamma-1} \right] e^{-\alpha(1-e^{-(\beta T^\gamma)})} e^{-\beta \sum_{i=1}^k (d_i^\gamma)} p(\alpha) p(\beta) p(\gamma). \end{aligned}$$

4.3.5. Weibull-Geométrica (WG)

La función de riesgo de la distribución Weibull-Geométrica está dada por,

$$h(t) = \frac{(\alpha/\sigma)(t/\sigma)^{\alpha-1}}{1 - pe^{-(t/\sigma)^\alpha}}, \quad t > 0.$$

Por lo tanto, el vector de parámetros que se desea estimar es $\theta = (\alpha, \beta, \mathbf{p}) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$. Además, si se tiene un proceso Poisson no homogéneo con función de riesgo Weibul-Geométrica, su función media es dada por

$$m(t) = -\log \left[\frac{(1-p)e^{-(t/\sigma)^\alpha}}{1 - pe^{-(t/\sigma)^\alpha}} \right].$$

Considerando lo anterior se tiene el siguiente resultado, ver [2].

Teorema 4.6 *Considere un proceso Poisson no homogéneo con función de riesgo Weibull-Geométrica, además, si se asume que los parámetros de la función de intensidad son a priori independientes, entonces la distribución a posteriori de los parámetros está dada por,*

$$p(\alpha, \sigma, \mathbf{p}|D) = (\alpha/\sigma^\alpha)^k \left[\prod_{i=1}^k \frac{d_i^{\alpha-1}}{1 - \mathbf{p}e^{-(d_i/\sigma)^\alpha}} \right] \left[\frac{(1 - \mathbf{p})e^{-(T/\sigma)^\alpha}}{1 - \mathbf{p}e^{-(T/\sigma)^\alpha}} \right] p(\alpha)p(\sigma)p(\mathbf{p}). \quad (4.7)$$

Demostración 4.8 *Sustituyendo la forma de la función de tasa así como la función de verosimilitud en (5.2), se tiene*

$$\begin{aligned} p(\alpha, \sigma, \mathbf{p}|D) &\propto l(\theta|D)p(\alpha, \sigma, \mathbf{p}) \\ &= l(\theta|D)p(\alpha)p(\sigma)p(\mathbf{p}) \\ &= \left[\prod_{i=1}^k \lambda(d_i) \right] e^{-m(T)} p(\alpha)p(\sigma)p(\mathbf{p}) \\ &= \left[\prod_{i=1}^k \frac{(\alpha/\sigma)(d_i/\sigma)^{\alpha-1}}{1 - \mathbf{p}e^{-(d_i/\sigma)^\alpha}} \right] e^{-(-\log \left[\frac{(1-\mathbf{p})e^{-(T/\sigma)^\alpha}}{1 - \mathbf{p}e^{-(T/\sigma)^\alpha}} \right])} p(\alpha)p(\sigma)p(\mathbf{p}) \\ &= (\alpha/\sigma^\alpha)^k \left[\prod_{i=1}^k \frac{d_i^{\alpha-1}}{1 - \mathbf{p}e^{-(d_i/\sigma)^\alpha}} \right] \left[\frac{(1 - \mathbf{p})e^{-(T/\sigma)^\alpha}}{1 - \mathbf{p}e^{-(T/\sigma)^\alpha}} \right] p(\alpha)p(\sigma)p(\mathbf{p}). \end{aligned}$$

4.4. Caso de estudio

Es posible que en ocasiones individuos viviendo en grandes ciudades tengan irritación en los ojos o la garganta. Esto puede ser causado por la contaminación atmosférica que es un problema en las grandes ciudades y en las zonas con actividades industriales, ya que en estos lugares es donde se liberan grandes cantidades de contaminantes a la atmósfera. Precisamente son algunos de estos contaminantes los que causan molestias en los ojos, piel, vías respiratorias y forman esas brumas que en la Ciudad de México y otras ciudades, impiden a sus pobladores ver cerros y volcanes que la rodean, ver [40].

La contaminación del aire es un grave problema que atañe desde hace unas décadas al mundo. Desde los problemas en los años cincuentas y sesentas en los países Europeos, como la famosa niebla tóxica londinense de 1952, el deterioro de los

bosques europeos por la lluvia ácida hasta la grave situación de la calidad del aire en ciudades como, la Ciudad de México, Tokio y Sao Paulo, ver [40]. Este tipo de contaminación, no solo es causado por las actividades humanas, también la misma naturaleza contribuye a la contaminación atmosférica, por ejemplo en las emisiones volcánicas. Afortunadamente, la naturaleza depura en cierta medida los contaminantes del aire por medio de la lluvia, vegetación o viento. Sin embargo, todo tiene un límite, y cuando las emisiones de los contaminantes sobrepasa la capacidad de depurar surgen problemas en el medio que rodea.

Dados los problemas que causa la contaminación del aire, se inició en diferentes países, la toma de decisiones que avalen una reducción en la emisión de contaminantes. Una de ellas es fijar límites permisibles de los contaminantes criterio, así como el establecimiento de redes para recabar información en tiempo (monitoreo atmosférico).

El monitoreo atmosférico proporciona información que sirve para determinar la calidad del aire en una zona determinada, a partir de su comparación contra los niveles permisibles establecidos para proteger la salud y bienestar de la población. La distribución espacial de los contaminantes depende de varios factores ambientales, ya que gracias a ellos se cuenta con datos que nos muestran con que rapidez, dispersión, transporte, reacción y acumulación, se presentan los contaminantes en una cuenca atmosférica.

En los últimos años en México se han realizado estudios para determinar, analizar y dar propuestas para solucionar los problemas que afectan a las ciudades pobladas e industrializadas, ver [2], [3], [10], [20], [28], [41],[42]. Los niveles actuales de contaminación ambiental en las ciudades producen mortalidad en los seres humanos, estos pueden ser a corto o largo plazo. Los resultados de estos estudios han sido publicado en Estados Unidos y Europa, ver [11], [26]. La Organización Mundial de la Salud (OMS) recomienda que los niveles de O_3 sean menores que 100 ug/m³ como media en 8h. Las normas oficiales mexicanas (NOM-020-SSA1-1993) permiten en concentración y tiempo promedio de 0.11ppm (216 ug/m³ en 1 hora).

En el Cuadro 4.1 se presenta los valores límites de exposición recomendados por Normas Oficiales Mexicanas de los contaminantes criterio, es importante destacar que la selección de un límite (umbral), es crucial debido a que si se reducen los límites habría disminución en el número de enfermos que acuden a los centros de salud, pero esto afectaría la infraestructura social, económica y política del país, ver [42].

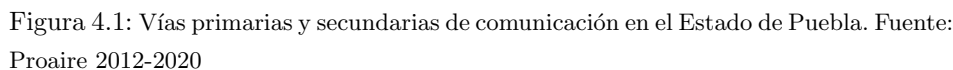
<i>contaminante</i>	<i>símbolo</i>	<i>valor límite</i>
Ozono	O_3	0.11ppm
Partículas menores a 10 micrómetros	PM_{10}	$210 \mu g/m^3$
Bióxido de azufre	SO_2	0.11ppm
Monóxido de carbono	CO	0.11ppm
Bióxido de nitrógeno	NO_2	0.21ppm

Cuadro 4.1: Contaminates Atmosféricos. Fuente: Proaire 2012-2020

Las entidades de la federación que estan comprometidas con la recabación de información de los niveles de los contaminantes del aire son: Durango, Morelos, Hidalgo, Michoacán, Chiapas, Baja California Sur, Chihuahua, Guanajuato, San Luis Potosí, Coahuila, Tabasco, Puebla, Aguascalientes, México, Baja California, Jalisco, Nuevo León, Ciudad de México (Distrito Federal). Los estados que cuentan con un sistema de medición durante un periodo largo son: Ciudad de México (Distrito Federal), Nuevo Leon, Jalisco y Baja California. El Estado de Puebla cuenta con un sistema de monitoreo desde el año 2000 y es considerado con un nivel intermedio de operación con un puntaje de 6 de una escala de 1-10, ver [20]. Esta evaluación considera la calidad de los datos, estudios realizados sobre estos datos, operación de la red, entre otros aspectos.

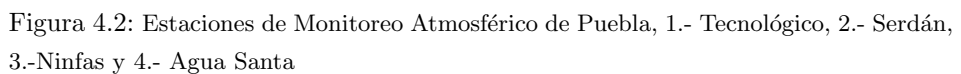
4.4.1. Extensión Geográfica del Estado de Puebla

El Estado de Puebla es una de las 31 entidades que conforman la República Mexicana y esta integrado por 217 municipios. Se ubica en la zona centro-oriental del país, colinda al este con el Estado de Veracruz; al poniente con el Estado de México, Hidalgo, Tlaxcala y Morelos, al sur con los Estados de Oaxaca y Gue-



Una buena calidad del aire puede definirse como la suma de concentraciones de componentes presentes en la atmósfera en un periodo de tiempo dado, que satisfacen la salud, el bienestar de la población, el equilibrio ecológico, y los materiales con valor económico [42].

En la Figura 4.2 se tiene el sitio geográfico y las estaciones existentes hasta el año 2011 de la REMA.



Fuente: http://sinaica.ine.gob.mx/rama_puebla.html

Para el estudio del contaminante ozono, la red de monitoreo atmosférico se toma como una sola estación que registra los máximos diarios que corresponde al municipio de Puebla. Las bases de datos son recabadas por la Secretaría de Desarrollo Rural, Sustentabilidad Ambiental y Ordenamiento Territorial del Estado de Puebla, se puede notar algunas observaciones faltantes, ver Figura 4.3. Dado que se trabajan con modelos Poisson no homogéneos se necesita tener las bases completas por lo cual se utiliza un método para completar la base.

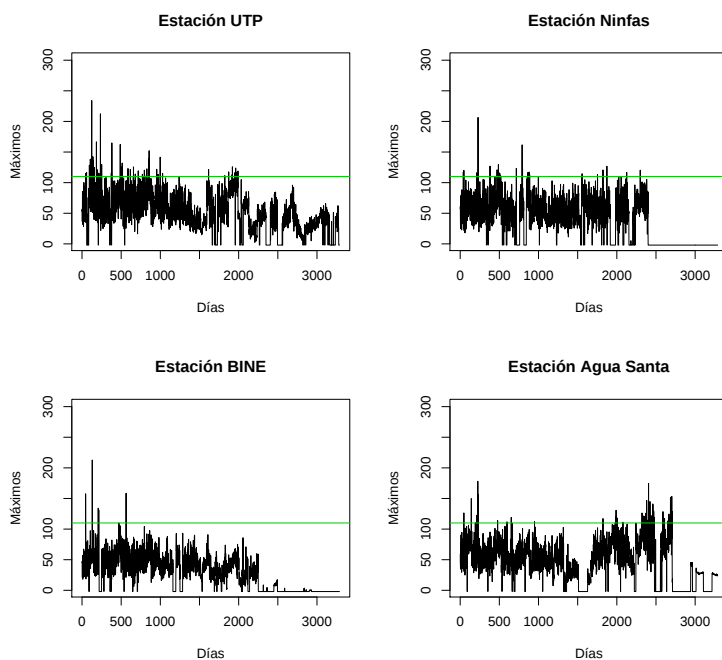


Figura 4.3: Series de tiempo para los máximos en las 4 estaciones de monitoreo atmosférico, del periodo 1/01/2001 al 31/12/2009. La línea horizontal corresponde al nivel del umbral (110 ppm)

Se realiza un resumen estadístico descriptivo sobre los datos para cada una de las estaciones, ver Cuadro 4.2.

Estación	Mín.	1er. cuan.	Mediana	Media	3er. cuan.	Máx.	Datos Faltantes
UTP	1.80	36.00	50.50	54.39	70.10	234.00	10 %
Ninfas	3.00	41.00	55.70	57.18	71.10	206.50	36 %
BINE	0.30	29.10	40.90	42.02	53.30	212.80	35 %
Agua Santa	7.50	38.30	55.20	56.95	72.20	178.10	21 %

Cuadro 4.2: Estadística descriptiva para las 4 estaciones de monitoreo atmosférico

Se observa que en el periodo 2001-2009, las cuatro estaciones de monitoreo en algún momento han presentado observaciones que sobrepasan el umbral (0.11 ppm equivalentes a 110 ppb). De manera similar se observa que en los gráficos de series de tiempo (ver Figura 4.3), la estación que en los últimos años ha presentado más niveles que sobrepasan el umbral es Agua Santa, seguida de la estación Ninfas. Cabe señalar que en los gráficos no se hace presente normalidad en los datos.

Método de llenado de bases de datos

Para el llenado de esta base se sigue el procedimiento que a continuación se describe, para cada estación y por consiguiente para la región de Puebla:

- Estos puntos se realizan para cada estación, del periodo 01-01-2001 al 31-12-2009.
 - Se seleccionan los máximos diarios de cada año.
 - Se compara el año a llenar con sus años más cercanos.
 - El día faltante se obtiene al promediar los días de los años cercanos, es decir, si el dato faltante es el z_i , $i = 1, \dots, 365$ del año Z , tomamos el mismo día pero de los años anterior y posterior al año Z y promediamos estas dos cantidades para obtener el día z_i .
- Después de completar la base de datos para cada estación, se selecciona el máximo diario de las cuatro estaciones.

- La base resultante se usa como los máximos diarios para la región de Puebla.

Observación: El método usado para llenar datos faltantes no es el mejor, dado que existen grandes discrepancias entre las observaciones.

Así mismo, al considerar el número de observaciones que exceden un umbral fijado de antemano de un contaminante, se construyen programas OpenBugs para estimar los parámetros y se obtienen los siguientes resultados. El Gráfico 4.4 muestra el comportamiento del contaminante ozono en la ciudad de Puebla, tomando en cuenta el método de llenado de datos.

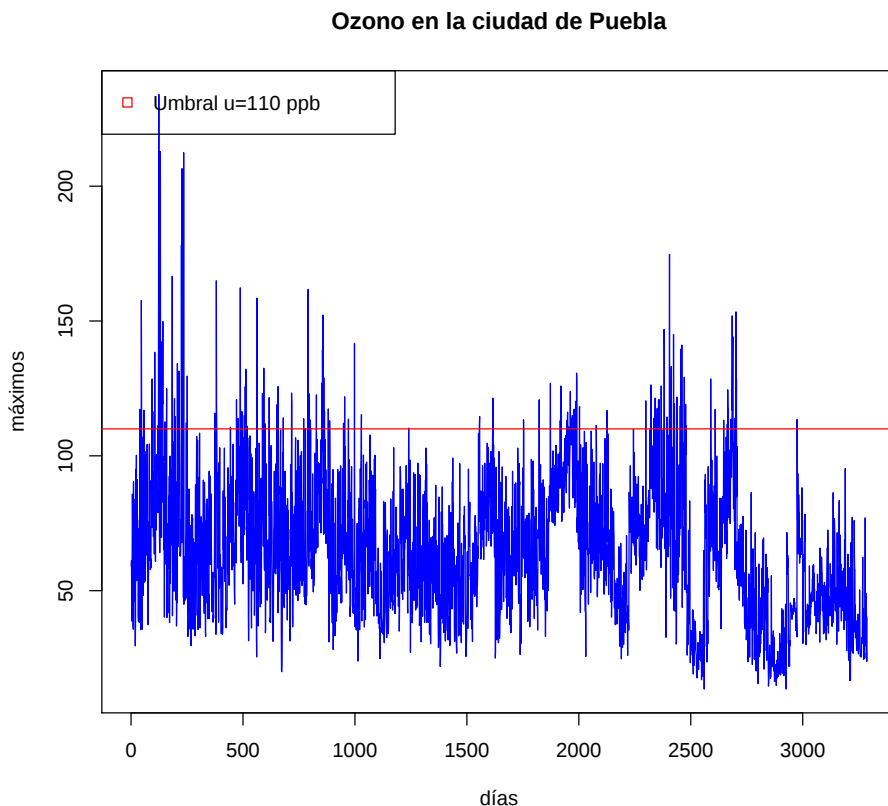


Figura 4.4: Comportamiento del ozono en la ciudad de Puebla, umbral $u = 110$ ppb

4.5. Resultados

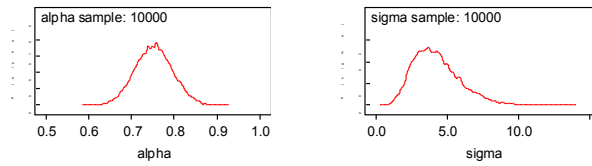
Las cadenas de Markov se simularon en el paquete OpenBugs, para cada función de riesgo se dan distribuciones a prioris uniformes en intervalos apropiados o distribuciones a priori gammas para los parámetros de las funciones de tasa, de la siguiente forma:

- Para la función de riesgo Weibull se utiliza una a priori uniforme en $(0, 1.5)$ para el parámetro α , mientras que para el parámetro σ una a priori gamma $(3.86, 0.76)$.
- Para la función de riesgo Musa-Okumoto se utiliza una a priori gamma $(66.01, 0.57)$ para el parámetro α y una a priori gamma $(68.77, 2.55)$ para el parámetro β .
- En el caso de la función de riesgo Goel-Okumoto se utiliza una a priori uniforme en $(150, 250)$ para el parámetro α , y una a priori uniforme en $(0, 0.003)$ en el caso del parámetro β .
- La función de riesgo Goel-Okumoto generalizada presenta tres parámetros, para cada uno de ellos se utilizan a prioris uniformes en los diferentes intervalos, con $(100, 500)$, $(0.5, 1)$, $(0.00001, 0.001)$, para el parámetro α, γ, β , respectivamente.
- Finalmente, en el caso de la función de riesgo Weibull-Geométrica se tienen a prioris uniformes en intervalos para los parámetros α y p con $(0.3, 0.9)$ y $(0, 1)$, respectivamente. Mientras para el parámetro σ la distribución a priori corresponde a una gamma $(3, 1.1)$.

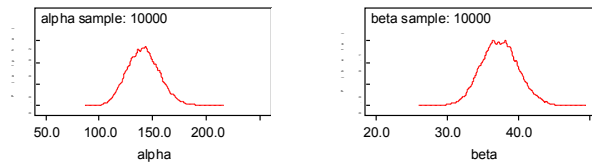
La Figura 4.5 muestra los gráficos de las densidades a posteriori para cada uno de los modelos utilizados.

70CAPÍTULO 4. APLICACIÓN A NIVELES DE OZONO EN LA CIUDAD DE PUEBLA

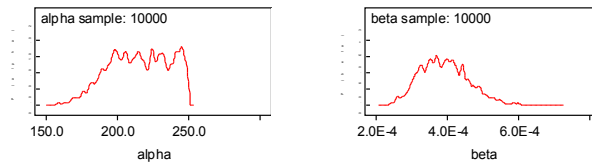
1



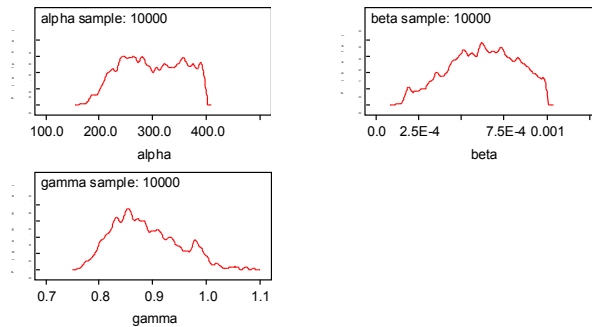
2



3



4



5

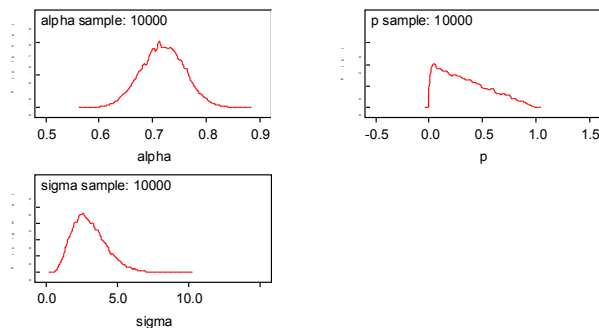


Figura 4.5: Gráficos de densidades de la distribución a posteriori para las funciones de riesgo: 1.-Weibull; 2.-Musa-Okumoto; 3.-Goel-Okumoto; 4.-Goel-Okumoto generalizada; 5.-Weibull-Geométrica

En el siguiente Cuadro 4.3, se establecen las estimaciones de los parámetros para cada una de las funciones de riesgo. Como se mencionó con anterioridad para estimar los parámetros se realizaron simulaciones de 5 cadenas en el programa OpenBugs, esto con el fin de observar por medio de los gráficos de traza, así como los de historial, la convergencia de las cadenas a la distribución a posteriori (La selección de las 5 cadenas se realiza para observar la convergencia, dado que con un número menor no podría ser visible esta convergencia). De esta manera al observar la convergencia a la distribución deseada para cada parámetro, se obtienen los gráficos de la distribución a posteriori para cada uno de los parámetros (ver Figura 4.5), que como se observa en algunos caso no se presenta normalidad. Los resultados presentados se han hecho en ausencia de puntos de cambio.

Modelo	Parámetro	media	sd	MC error	DIC
Weibull	$\hat{\alpha}$	0.7526	0.04364	5.258E-4	
	$\hat{\sigma}$	4.281	1.644	0.01871	1246.0
Musa-Okumoto	$\hat{\alpha}$	141.4	15.21	0.1493	
	$\hat{\beta}$	37.42	2.629	0.02982	1288.0
Goel-Okumoto	$\hat{\alpha}$	215.4	21.07	1.087	
	$\hat{\beta}$	3.937E-4	6.874E-	2.677E-6	1245.0
Goel-Okumoto generalizada	$\hat{\alpha}$	298.8	57.57	3.074	
	$\hat{\beta}$	6.254E-4	2.029E-4	1.009E-5	
	$\hat{\gamma}$	0.887	0.06204	0.7888	1248
Weibull- Geométrica	$\hat{\alpha}$	0.7191	0.04117	5.615E-4	
	$\hat{p}$	0.3366	0.2389	0.0030066	
	$\hat{\sigma}$	3.098	1.222	0.01709	1250

Cuadro 4.3: Resumen estadístico para las funciones de riesgo sin puntos de cambio

Con estos resultados, se calcularon la función media para cada una de las funciones de riesgo analizada

- Para un proceso Poisson no homogéneo con función de riesgo de Weibull, con los parámetros estimados, ver Cuadro 4.3 se tiene la función media esta

dada por

$$m(t) = \left(\frac{t}{\sigma}\right)^\alpha = \left(\frac{t}{4.281}\right)^{0.75}.$$

- De igual forma para un proceso Poisson no homogéneo con función de riesgo Musa-Okumoto, se tiene que la función media del proceso está dada por

$$m(t) = \beta \log\left(\frac{t}{\alpha} + 1\right) = 37.42 \log\left(\frac{t}{141.1} + 1\right).$$

- Para un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto, la función media del proceso es dada por

$$m(t) = \alpha(1 - e^{-\beta t}) = 215.4(1 - e^{-3.94t}).$$

- Cuando se generalizan los modelos anteriores se tiene un proceso Poisson no homogéneo con función de riesgo Goel-Okumoto generalizada, con función media del proceso dada por

$$m(t) = \alpha(1 - e^{-\beta t^\gamma}) = 298.8(1 - e^{-6.25t^{0.88}}).$$

- Finalmente, para un proceso Poisson no homogéneo con función de riesgo Weibull-Geométrica, su correspondiente función media del proceso con los parámetros estimados es dada por

$$m(t) = -\log \left[\frac{(1-p)e^{-(t/\sigma)^\alpha}}{1 - pe^{-(t/\sigma)^\alpha}} \right] = -\log \left[\frac{(0.67)e^{-(t/3.09)^{0.71}}}{1 - 0.33e^{-(t/3.09)^{0.71}}} \right].$$

Para cada una de las funciones de riesgo para un proceso Poisson no homogéneo, la probabilidad de que ocurra el evento de interés esta dado por

$$P(N(t_2) - N(t_1) = n) = e^{-[m(t_2) - m(t_1)]} \frac{[m(t_2) - m(t_1)]^n}{n!}, \quad t_2 > t_1, \quad n \geq 0, \quad (4.8)$$

en cada caso se sustituye el valor de la función media del proceso.

En la Figura 4.6 se presentan los gráficos con las funciones medias del proceso Poisson con las funciones de riesgo Weibull, Musa-Okumoto, Goel-Okumoto, Goel-Okumoto generalizada y Weibull-Geométrica, así como la media de las observaciones. Se observa que modelos se ajustan mejor al promedio de las observaciones.

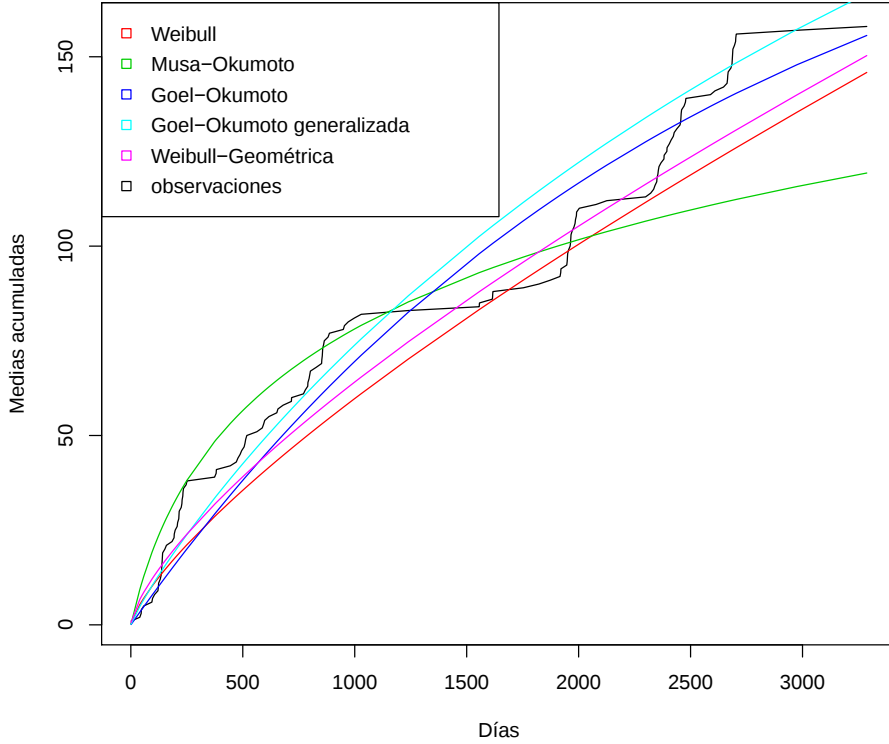


Figura 4.6: Gráfico de las medias acumuladas, para el proceso Poisson no homogéneo con funciones de riesgo Weibull, Musa-Okumoto, Goel-Okumoto, Goel-Okumoto generalizada y Weibull-Geométrica, respecto a la media de las observaciones y sin puntos de cambio

Los siguientes resultados se ha considerado un punto de cambio y se presentan las densidades a posteriori para cada una de la funciones de riesgo respectivas, (sólo para la funciones de riesgo Weibull y Musa-Okumoto).

- Para el proceso Poisson no homogéneo con función de riesgo Weibull y considerando un punto de cambio se tienen las siguientes a priori para cada parámetro antes y después del punto de cambio τ . Distribuciones a priori uniformes con intervalos $(0, 3)$; $(0, 2)$; $(10, 70)$, $(150, 300)$, respecto a los

parámetros $\alpha_1, \alpha_2, \sigma_1$ y el punto de cambio τ , mientras que el parámetro σ_2 se utilizó una a priori gamma (6.56, 0.18).

- Para el proceso Poisson con función de riesgo Musa-Okumoto se tienen a priori uniformes con intervalos en (0, 500000); (0, 2000); (10, 1000); (50, 200); (50, 150), para los parámetros $\alpha_1, \alpha_2, \beta_1, \beta_2$ y al punto de cambio τ , respectivamente.

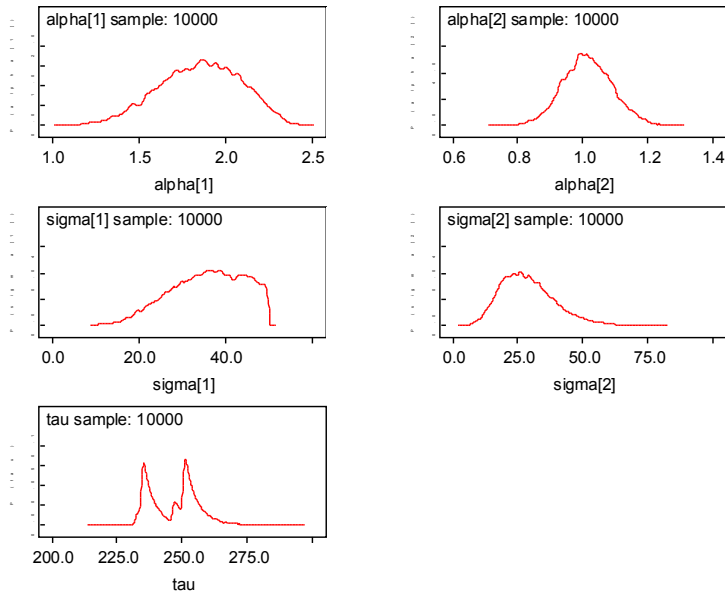
En el siguiente Cuadro 4.4, se establecen las estimaciones de los parámetros para cada una de las funciones de riesgo.

Modelo	Parámetro	media	sd	MC error	DIC
Weibull	$\hat{\alpha}_1$	1.854	0.2364	0.004591	
	$\hat{\alpha}_2$	1.015	0.07628	0.00167	
	$\hat{\sigma}_1$	35.89	8.526	0.1651	
	$\hat{\sigma}_2$	28.91	10.44	0.2346	
	$\hat{\tau}$	247	9.392	0.1645	1226
Musa-Okumoto	$\hat{\alpha}_1$	17750.0	10760.0	191.7	
	$\hat{\alpha}_2$	1076.0	414.6	7.957	
	$\hat{\beta}_1$	652.3	239.4	4.222	
	$\hat{\beta}_2$	114.2	24.42	0.443	
	$\hat{\tau}$	91	20.74	0.4255	1249

Cuadro 4.4: Resumen estadístico con un punto de cambio

En la Figura 4.7 se muestran los gráficos de densidades a posteriori dadas las anteriores distribuciones a priori, de manera similar se relizan corridas de 40000 iteraciones de las cuales se desechan las primeras 20000 sobre 5 cadenas simuladas, y solo se eligen las iteraciones cada 10 observaciones (el desechar las primeras 20000 iteraciones es dado que en ellas no se observan la convergencia de las 5 cadenas).

1



2

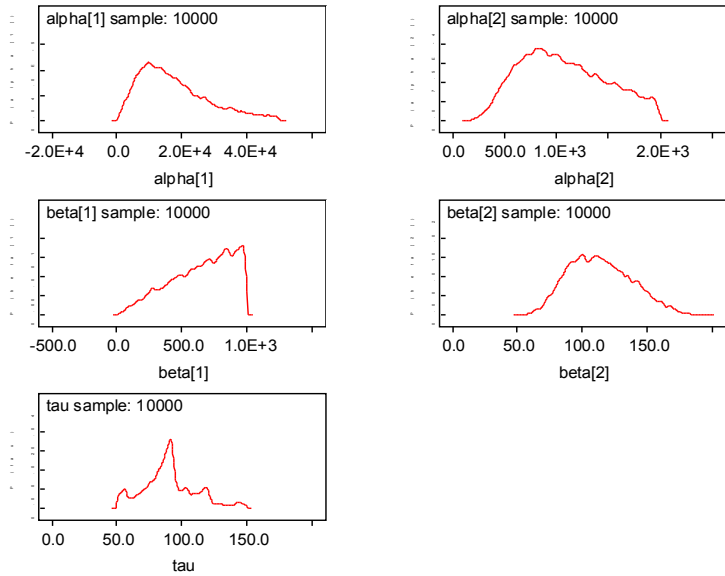


Figura 4.7: Gráficos de densidades de la distribución a posteriori para las funciones de riesgo: 1.-Weibull; 2.-Musa-Okumoto; con un punto de cambio

En la Figura 4.8 se presentan los gráficos con las funciones medias del proceso Poisson, con las funciones de riesgo de Weibull y Musa-Okumoto, considerando un punto de cambio que se da en $\tau = 247$ para el modelo Weibull y $\tau = 91$, para el caso del modelo con función de riesgo Musa-Okumoto (para hallar el punto de cambio τ , se considera como una variable aleatoria sobre el intervalo $[0, T]$ y se estima usando inferencia bayesiana). Se observa que modelo se ajustan mejor al promedio de las observaciones.

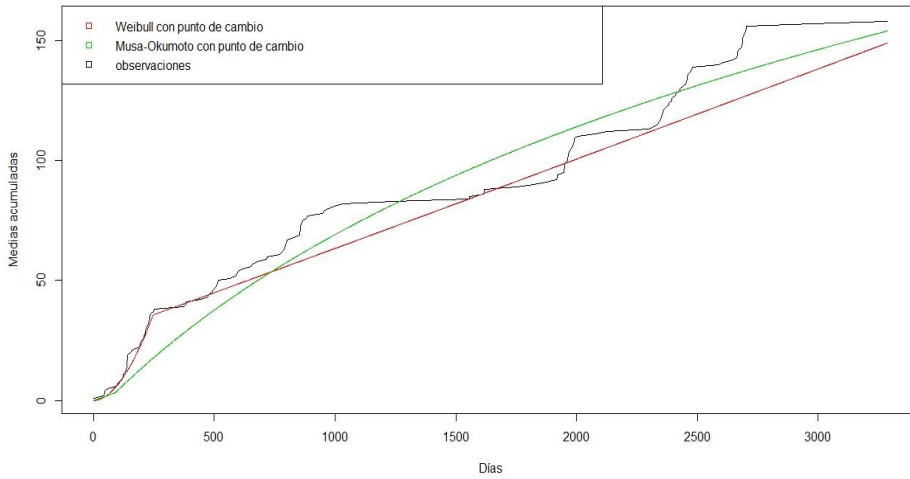


Figura 4.8: Gráfico de las medias acumuladas, para el proceso Poisson no homogéneo con funciones de riesgo Weibull, Musa-Okumoto, respecto a la media de las observaciones y con un punto de cambio

Conclusiones

De acuerdo al modelo presentado se tienen las siguientes conclusiones parciales, tomando en cuenta el objetivo principal del estudio que es elegir el modelo que ajusta mejor las observaciones de la red de monitoreo atmosférico del Estado de Puebla:

- El gráfico de las medias acumuladas con respecto a la media de las observaciones Figura 4.6, se observa que los modelos sin punto de cambio que se describieron, los que mejor se ajustan son: el Weibull, Musa-Okumoto, Goel-Okumoto, Goel-Okumoto generalizada y Weibull-Geométrica y usando el Cuadro 4.3, el modelo que presenta menor DIC es el modelo con función de riesgo Goel-Okumoto.
- Para los modelos con un punto de cambio se observa, ver Figura 4.8 un comportamiento similar en el gráfico de las medias acumuladas, ambos modelos se van ajustando a la media de las observaciones. El modelo que presenta menor DIC es el proceso Poisson con función de riesgo Weibull, ver Cuadro 4.4. En este punto, se considera la misma función de riesgo antes y después del punto de cambio, sin embargo, lo que cambia son los parámetros estimados.
- Al comparar los DIC en cada una de las tablas, con punto de cambio y sin punto de cambio, se observa que el modelo se va ajustando más a las observaciones cuando se consideran modelos con un punto de cambio.

El método que se está usando en este trabajo, es una aplicación de un método ya usado en los datos de Ozono en la zona Metropolitana del Valle de México. Durante el desarrollo del trabajo se observa que las bases originales de nuestro problema son extensas (más de 70 000 observaciones por estación) y que presentan datos faltantes, un problema es que aún no se cuenta con el método más adecuado para su llenado. Al final, usando el método de llenado explicado con anterioridad, se contó con una base de datos (considerando las bases llenas y máximos diarios), con un tamaño de 3287 observaciones.

Por ello urge que se implementen medidas para mejorar la recolección de datos sobre contaminantes de la atmósfera, esto con el fin de tener resultados más acertados a la realidad, así como un informe detallado de lo que está pasando con el aire que se respira en la Ciudad de Puebla².

En el proceso de simulación de las cadenas de Markov en el paquete de Openbugs se consideran los siguientes aspectos:

- Dado que no se cuenta con información acerca del comportamiento de cada uno de los parámetros considerados para cada modelo, se utilizan a priori no informativas (en este caso distribuciones uniformes sobre intervalos).
- Al observar los gráficos de las densidades a posteriori para cada parámetro, si este gráfico tiene una forma conocida, a la distribución a priori del parámetro se le asigna dicha distribución.
- Se consideran la simulación de 5 cadenas para observar la convergencia gráficamente, por medio de los gráficos de traza e historial (se pueden considerar más cadenas, solo que los tiempos de corrida serían más tardados).

²Puebla y su zona conurbada está creciendo desmesuradamente, cada día cuenta con una flotilla vehicular más grande, que trae como consecuencia que los niveles de contaminación sigan creciendo. A pesar de que se han dado recomendaciones y se han considerado nuevas formas de transporte como lo es el transporte articulado RUTA, sigue habiendo problemas con la circulación, como semáforos no sincronizados, topes de reducción de velocidad, entre otros.

Hace falta mucho trabajo para tener resultados, como lo es el mantenimiento a vías primarias y secundarias, la sincronización efectiva de semáforos, eliminación de topes donde no sean necesarios, mejorar el sistema de recolección de datos atmosféricos así como incrementar la red con más estaciones de monitoreo, construir pulmones para la ciudad, así como conservar los pocos que existen.

- Se simulan 40000 iteraciones para que la convergencia se observe de mejor manera y esto se garantice.
- Se desechan las primeras 20000 iteraciones, dado que en los gráficos de historial no se nota la convergencia en estas iteraciones.
- En el gráfico de las medias acumuladas sin considerar puntos de cambio, bajo el modelo Poisson no homogéneo con cada una de las funciones de riesgo consideradas, se hace notorio el comportamiento de la media de las observaciones, es por ello que consideramos puntos de cambio.
- Como no se conoce información acerca de este punto de cambio, se toma como una variable a considerar (parámetro) y de la misma forma se estima mediante inferencia bayesiana.

Trabajos a futuro:

- Un trabajo de importancia es hallar un método efectivo de llenar bases de datos con observaciones faltantes, que garantice resultados apegados a la realidad.
- Siguiendo la línea de investigación, lo ideal es considerar otros modelos y considerar más puntos de cambio con diferentes funciones de riesgo.

Apéndice A

Teoremas y definiciones.

Definición A.1 Probabilidad condicional. Sean A y B dos eventos y supongamos que $P(B) > 0$. La **probabilidad condicional** del evento A dado el evento B se denota con el símbolo $P(A|B)$ y se define como el siguiente cociente:

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

Observe que la probabilidad condicional no esta definida si $P(B) = 0$.

La definición anterior y el teorema que sigue se pueden consultar en, ver [8], [37].

Teorema A.1 Teorema de Bayes. Sea $B_1, B_2, \dots, B_k$ una partición del espacio Ω , tal que cumple, $P(B_j) > 0$, $j = 1, 2, \dots, k$ y sea A un evento tal que $P(A) > 0$. Entonces para $j = 1, 2, \dots, k$ se cumple:

$$P(B_j|A) = \frac{P(B_j)P(A|B_j)}{\sum_{j=1}^k P(B_j)P(A|B_j)} \quad (\text{A.1})$$

Teorema A.2 Ley Fuerte de los Grandes Números. Sea $\{X_n\}$ una sucesion de variables aleatorias independientes e identicamente distribuidas, con media μ . Entonces

$$\frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{\text{c.s.}} \mu$$

Este último teorema se tomó de [37].

Bibliografía

- [1] Achcar A. Jorge, Fernández-Bremauntz Adrián A., Rodrigues Eliane, Tzintzun Guadalupe, *Estimating the number of ozone peaks in Mexico City using a non-homogeneous Poisson model*, Environmetrics, 19: 469-485, 2008.
- [2] Achcar A. Jorge, Barrios M. Juan, Rodrigues Eliane R., *Comparing the Adequacy of Some Non-Homogeneous Poisson Models to Estimate Ozone Exceedances in Mexico City*, Journal of Environmental Protection, 3, 1213-1227, 2012.
- [3] Achcar A. Jorge, Rodrigues Eliane, Tzintzun Guadalupe, *Using non-homogeneous Poisson models with multiple change-points to estimate the number of ozone exceedances in Mexico City*, Environmetrics; 22: 1-12, 2011.
- [4] Álvarez J. Luis, Fernández-Bremauntz A. Adrián, Rodrigues R. Eliane y Tzintzun Guadalupe, *Maximum a Posteriori Estimation of the Daily Ozone Peaks in Mexico City*, 2005.
- [5] Ariza Hernández Francisco J., *Pruebas de Bondad de Ajuste y de Razón de Verosimilitudes Generalizada para Procesos Poisson No Homogéneo*, Tesis de Maestría, COLPOS, 2005.
- [6] Barreto-Souza Wagner, Lemos de Moraes Alice , Cordeiro Gauss M., *The Weibull-Geometric Distribution*, 2008.
- [7] Brzeźniak Zdzislaw, Zastawniak Tomasz, *Basic Stochastic Processes*, 7th edition, Springer, 2005.

- [8] DeGroot Morris H., Schervish Mark J., *Probability and Statistics*, Fourth edition, Addison-Wesley publishing Company, 2012.
- [9] Díaz G.L., Vega V. Sistachs , Covarrubias M. D., Hernández N. I., *Selección de Modelos Bajo el Enfoque Bayesiano: Una Aplicación al Estado Cognitivo de los Adultos Mayores en el Estado de Guerrero*, Revista Investigación Operacional, Vol. 35(02), 180-188, 2014.
- [10] Dirección General de Gestión e Información Ambiental del Instituto Nacional Ecología, Agencia de Cooperación Internacional de Japón (JICA), *Almanaque de datos y tendencias de la calidad del aire en ciudades mexicanas*, <http://www.ine.gob.mx>, 2000.
- [11] Gauderman, W. J. et al. , *Effect of exposure to traffic on lung development from 10 to 18 years of age: a cohort study*, Lancet, 369(9561); p. 571-7, 2007.
- [12] Gilks Walter R., S. Richardson, Spiegelhalter D.J., *Markov Chain Monte Carlo in Practice*, Springer-science business media, B.V., First edition, 1996.
- [13] Goel Amrit L., Okumoto Kazu, *Time-Dependent Error-Detection Rate Model for Software Reliability and Other Performance Measures*, IEEE Transactions on reliability, Vol. R-28, No. 3, 1979.
- [14] Green Joanne, Sánchez Sergio, *La Calidad de Aire en América Latina: Una Visión Panorámica*, Clean Air Institute, 2013.
- [15] Hoel Paul G. , Port Sidney C. , Stone Charles J. , *Introduction to Stochastic Processes*, Waveland Pr Inc, 1986.
- [16] Hogg V. Robert, Craig Allen, McKean W. Joseph, *Introduction to Mathematical statistics*, sixth edition, Pearson Prentice Hall, 2005.
- [17] Horst Rinne, *The Weibull distribution: A Handbook*, Chapman & Hall, 2009.
- [18] Ibarra Carrasco Raymundo, *Modelo de Poisson no homogéneo aplicado a la estimación de la contaminación atmosférica*, tesina, UNAM, 2011.
- [19] Instituto Nacional de Ecología, *Cuarto Almanaque de datos y tendencias de la calidad del aire en 20 ciudades mexicanas (2000-2009)*, primera edición, 2011.

- [20] Instituto Nacional de Ecología y Cambio Climático, *Diagnóstico de la Medición de la Calidad del Aire en México, 60 años monitoreando la calidad del aire*, 2012.
- [21] Jara Ettinger José Julián, *Estimando el número de excedencias de ozono en la Ciudad de México usando procesos de Poisson no homogéneos y el muestreador de Gibbs*, Tesis, <http://web.mit.edu/jjara/www/docs/thesis.pdf>, 2011.
- [22] Jensen Uwe and Lutkebohmert Constanze, *Change-Point Models*, 2009.
- [23] Juárez O., Castells E., *Modelos de arbol de regresión Bayesiano: Un estudio de caso*, Revista Investigación Operacional, Vol. 31(2), 109-125, 2010.
- [24] Leadbetter M. R., Lindgren Georg, Rootzen Holger, *Extremes and related properties of random sequences and processes*, Springer-Verlag, 1983.
- [25] Lee Peter M. , *Bayesian Statistics: An Introduction*, second edition, A John Wiley & Sons, Ltd., Publication, 2012.
- [26] McConnell, R ., et al., *Traffic, susceptibility, and childhood asthma* Environ Health Perspect, 114(5):p. 766-72, 2006.
- [27] Mohan G. Krishna, Rao B. Srinivasa , Prasad R. Satya, *A Comparative of Software Reliability Models Using SPC on Ungrouped Data*, International Journal of Advanced Research in Computer Science and Software engineering, 2012.
- [28] Molina Luisa T., Molina Mario J., *Air Quality in the Mexico Megacity: An Integrated Assessment*, SPRINGER -SCIENCE+BUSINESS MEDIA, B.V., 2002.
- [29] Musa J. D., Okumoto K., *A Logarithmic Poisson Execution Time Model for Software Reliability Measurement*, Proceedings of Seventh International Conference on Software Engineering, Orlando, pp. 230-238, 1984.
- [30] *OpenBUGS*, www.openbugs.net
- [31] Organización Mundial de la Salud (OMS), *Calidad de aire (exterior) y salud*, nota descriptiva No. 313, 2014.

- [32] Pérez Abreu Víctor M., Reynoso Valle David, *Introducción a los Procesos de Poisson*, notas, 2010.
- [33] Population Pyramid, <http://populationpyramid.net/es/mundo/2015>.
- [34] Raftery A. E. , *Hypothesis Testing and Model Selection*, In: W. Gilks, S. Richardson and D. J. Spiegelhalter, Eds., *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London, pp. 163-187, 1996.
- [35] Reiss M. Thomas R.-D , *Statistical analysis of Extreme Values from Insurance, Finance, Hydrology and Other Fields*, Second Edition, Bertelsmann Springer Publishing Group, 2001.
- [36] Rincón Luis, *Introducción a los procesos estocásticos*, Departamento de Matemáticas, Facultad de ciencias UNAM, 2012.
- [37] Rincón Luis, *Introducción a la probabilidad*, Departamento de Matemáticas, Facultad de ciencias UNAM, 2013.
- [38] Robert Christian P., Casella George, *Monte Carlo Statistical Methods*, second edition, Springer, 2010.
- [39] Secretaría de Desarrollo Rural, Sustentabilidad Ambiental y Ordenamiento Territorial (SDRSOT), *Red de Monitoreo Atmosférico de Puebla (REMA)*, www.remapuebla.gob.mx
- [40] SEMARNAT, *Contaminación*, http://www.semarnat.gob.mx/archivosanteriores/informacionambiental/Documents/05_serie/yelmedioambiente/5_contaminacion_v08.pdf.
- [41] Semarnat, SSAOT, Puebla, *Programa de Gestión de calidad de la zona Metropolitana del Valle de Puebla 2006-2011*, 2006.
- [42] Semarnat, SSAOT, Benemérita Universidad Autónoma de Puebla, *Programa de Gestión de Calidad del aire del estado de Puebla 2012-2020*, 2012.
- [43] Spiegelhalter D. J., Best N. G., Carlin B. P. and van der Linde A. , *Bayesian Measures of Model Complexity and Fit*, Journal of the Royal Statistical Society Series B, Vol. 64, No. 4, pp. 583-639, doi:10.1111/1467-9868.00353, 2002.

- [44] Stroock Daniel W., *An Introduction to Markov Processes*, Springer, 2000.
- [45] Villaseñor-Alva José A., Gonzáles-Estrada Elizabeth, Guzmán-Martínez María, *Extreme Value modeling for environmental data analysis*, Boletín de la Sociedad Matemática Mexicana, Third Series, Springer, 2014.
- [46] *La Ciudad de Puebla*, www.laciudaddepuebla.com
- [47] <http://blogecologista.com/los-10-paises-mas-contaminados>.