BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

FACULTAD DE CIENCIAS FÍSICO MATEMÁTICAS

POSGRADO EN CIENCIAS MATEMÁTICAS

EL USO DE CÓPULAS PARA LA OBTENCIÓN DE DISTRIBUCIONES MULTIVARIADAS PARA VALORES EXTREMOS: UNA APLICACIÓN A DATOS DE CONTAMINACIÓN ATMOSFÉRICA

Tesis

PARA OBTENER EL GRADO DE:

DOCTOR EN CIENCIAS MATEMÁTICAS

PRESENTA:

JUAN ANTONIO VAZQUEZ MORALES

DIRECTORES DE TESIS:

DRA. HORTENSIA JOSEFINA REYES CERVANTES
DRA. ELIANE REGINA RODRIGUES

PUEBLA, PUEBLA, DICIEMBRE 2024

DR. SEVERINO MUÑOZ AGUIRRE
SECRETARIO DE INVESTIGACIÓN Y
ESTUDIOS DE POSGRADO, FCFM-BUAP
P R E S E N T E:

Por este medio le informo que el C:

JUAN ANTONIO VÁZQUEZ MORALES

estudiante del Doctorado en Ciencias (Matemáticas), ha cumplido con las indicaciones que el Jurado le señaló en el Coloquio que se realizó el día 21 de noviembre de 2024, con la tesis titulada:

El uso de cópulas para la obtención de distribuciones multivariadas para valores extremos: una aplicación a datos de contaminación atmosférica

Por lo que se le autoriza a proceder con los trámites y realizar el examen de grado en la fecha que se le asigne.

A T E N T A M E N T E.
H. Puebla de Z. a 22 de noviembre de 2024

DR. RAÚL ESCOBEDO CONDE
COORDINADOR DEL POSGRADO
EN MATEMÁTICAS.

Facultad
de Ciencias
Físico Matemáticas

Av. San Claudio y 18 sur, edif. FM1
Ciudad Universitaria, Col. San
Manuel, Puebla, Pue. C.P. 72570
(222) 229 55 00 Ext. 7550 y 7552

En memoria de mi abuela, Concepción Hernández Nicanor.

A la memoria de mi director de tesis de licenciatura, Dr. Guillermo López Mayo.

A mi familia, amigos, alumnos y profesores.

Agradecimientos

A mis padres, Isela y Constantino, quienes me brindaron su comprensión durante estos años. A mi abuela Concepción, que me brindó su cariño y apoyo incondicional. A mis hermanas, Rocio y Lucero, quienes me apoyaron en mis responsabilidades para que pudiera concentrarme en este trabajo. A mi sobrino Julian, quien entendió que no podía jugar con él todos los días.

A mis amigos y compañeros de la facultad: América, David, Gustavo, Julio, Paty, Roque y Yasmín, por su ayuda, comprensión, consejos y los buenos momentos compartidos. Sin embargo, también quiero extender mi gratitud a todas aquellas personas cuya contribución, aunque no mencionada específicamente, ha sido igualmente valiosa.

A mis asesoras de tesis, Dra. Eliane Regina Rodrigues y Dra. Hortensia Josefina Reyes Cervantes, por su invaluable orientación y apoyo a lo largo de este proceso.

A los miembros del jurado: Dr. Bulmaro Juárez Hernández,

Dr. Francisco Solano Tajonar Sanabria, Dra. Gladys Linares Fleites, Dr. Hugo Adán Cruz Suárez y Humberto Vaquera Huerta, por su tiempo y dedicación en la revisión de esta tesis.

Al Consejo Nacional de Humanidades, Ciencias y Tecnologías (Conahcyt) por el apoyo financiero y académico durante el desarrollo de este proyecto. A la Benemérita Universidad Autónoma de Puebla (BUAP) por la formación integral y los recursos necesarios para llevar a cabo esta investigación.

En las matemáticas, el ajedrez y la vida,
no existe el fracaso, solo el aprendizaje.

Introducción

La contaminación ambiental ha sido tema de estudio en los últimos años, en especial sus efectos en la salud y cambios climáticos [2, 51]. El ozono, las partículas menores a 10 y 2.5 micrómetros son de los principales contaminantes analizados en diversos países, y varios artículos se concentran en analizar tales contaminantes [3, 32, 57], y la ciudad de México es una ciudad afectada por tales contaminantes.

Se han aplicado diversas técnicas matemáticas al tema de la contaminación, por ejemplo, mediante modelos en ecuaciones diferenciales parciales [4] o estimaciones de parámetros de funciones que establecen la cantidad de masa de basura en diversas zonas del mar [9]. Las técnicas estadísticas empleadas para analizar los niveles elevados de contaminación son diversas, por ejemplo, usando la teoría de valores extremos para analizar los niveles de dióxido de azufre y el dióxido de nitrógeno en la ciudad de Istanbul [13] y las cópulas para

analizar la relación de partículas menores a 10 micrómetros y otros contaminantes en la ciudad de Kland [32]. En este trabajo se combina la técnica de valores extremos y la teoría de cópulas para analizar la contaminación en la Zona Metropolitana del Valle de México (ZMVM) de 1990 a 2022, con el fin de modelar la relación que existe entre los niveles máximos mensuales de ozono y otros contaminantes, los cuales son registrados por la Red Automática de Monitoreo Atmosférico (RAMA) de la ZMVM.

Dentro de los objetivos del presente trabajo se encuentran:

- Desarrollar y entender la teoría necesaria para el caso de aplicación.
- Encontrar distribuciones marginales de los niveles de máximos mensuales.
- Encontrar distribuciones multivariadas donde se conserve la relación entre los conjuntos de niveles de máximos mensuales.

¿Por qué usar cópulas y valores extremos en la ZMVM? Para responder esta pregunta, se habla un poco de la ciudad de México. La zona de la ciudad está construida en un valle[1], por lo cual las corrientes de aire no son abundantes, y por

[1]Llanura entre montes y montañas.

consiguiente, no propician la renovación del aire en la zona de la ciudad, añadido de que al ser la capital del país, ha provocado que sea una de las ciudades más pobladas del país, provocando movilidad de transporte motorizado, creación de fábricas en la periferia de la ciudad y otras actividades que contribuyen a la emisión de contaminantes.

Como se puede ver, estudiar los niveles altos de contaminación en esta zona es muy importante, en especial los máximos mensuales son los que se deben analizar, ya que se busca que estos sean bajos, o bien, estén inferiores a un umbral. Examinar y predecir el comportamiento aleatorio de estos datos usando teoría de valores extremos es lógico, ya que es una teoría que se adapta a los propósitos planteados.

En la ZMVM se cuenta con la Red Automática de Monitoreo Atmosférico, la cual pone a disposición del público en general datos de contaminación [40]. En tal base se observa que hay datos por día desde 1986, sin embargo, en años recientes, se puede observar que hay horas, incluso días sin datos. Tener datos faltantes es un problema, por suerte, se ha desarrollado la teoría de cópulas, la que permite acoplar varias distribuciones de probabilidad y obtener una función de densidad conjunta donde cada marginal sería una distribución univariada. De acuerdo a Fisher [14]:

"Las cópulas son de interés para los estadísticos por dos razones principales: en primer lugar, como una forma de estudiar medidas de dependencia sin escala; y, en segundo lugar, como punto de partida para la construcción de familias de distribuciones conjuntas, a veces con miras a la simulación".

Por este motivo, usar la teoría de cópulas para analizar el comportamiento conjunto es una idea razonable, ya que incluso teniendo algunos valores faltantes, se puede ajustar y modelar la relación de los contaminantes a partir de las parejas de valores registrados.

Para ajustar las cópulas y las densidades a los datos obtenidos, se usa la estadística bayesiana, en especial los algoritmos para aproximar los parámetros, los llamados métodos Monte Carlo de la cadena de Markov. Estos métodos ya han sido implementados en el paquete OpenBUGS [11, 31, 54], con lo cual se puede aplicar la teoría matemática desarrollada en este trabajo, así como los conceptos en cópulas, valores extremos, estadística bayesiana y métodos Monte Carlo de cadenas de Markov. Para complementar estos temas, se incluye un capítulo de conceptos y resultados de probabilidad, generación de variables aleatorias y estadística, al igual de algunas referencias importantes para una consulta más profunda.

Este trabajo presenta el Capítulo 1, donde se explica qué es una cópula, iniciando con las bivariadas y después las de dimensión mayor a dos. En ambos casos, se formula el Teorema de Sklar, que permite encontrar distribuciones multivariadas dadas las distribuciones marginales. Se maneja el concepto de medida de concordancia a partir de la cópula, cópulas máx-estables y una generalización de las cópulas bivariadas y las cópulas asimétricas.

El Capítulo 2 presenta la teoría de valores extremos, es decir, la distribución de valor extremo generalizada y sus casos particulares. Se resumen los resultados de Smith [50], donde se aclaran las limitantes de las estimaciones de máxima verosimilitud. Por los resultados de Smith, se considera otro camino para encontrar aproximaciones de los parámetros usando la estadística bayesiana, que se desarrolla en el Capítulo 3, que a partir de suposiciones lógicas sobre las distribuciones y parámetros e incluyendo muestras, se puede encontrar una aproximación más exacta de los parámetros.

Aunque el paquete OpenBUGS permite encontrar las estimaciones requeridas con la teoría bayesiana, y no es necesario comprender los algoritmos internos, se incluye el Capítulo 4. Se explica como funciona el llamado "muestrador de Gibbs" que de acuerdo a Lunn et al. [31], es el algoritmo utilizado en

el paquete OpenBUGS. Tal algoritmo entra dentro de los llamados "métodos Monte Carlo de la cadena de Markov", por lo que se hace una breve explicación de lo que es una cadena de Markov y cómo se usa para la inferencia bayesiana. Para introducir el muestreo de Gibbs, se hace considerándolo como una modificación del "algoritmo EM", así como un ejemplo del funcionamiento de ambos algoritmos.

Por último, en el Capítulo 5 se aplica la teoría presentada en los capítulos anteriores a los datos de máximos mensuales en las distintas regiones de la ZMVM. Además, dado que el ozono es un contaminante que se ha observado durante más tiempo en la zona, se analiza la concordancia de este contaminante con otros estudiados por expertos ambientalistas. Asimismo, se utilizan las cópulas asimétricas para examinar la relación en las regiones Centro, Noreste y Suroeste, que forman el llamado "corredor del aire".

Las distribuciones obtenidas en este trabajo, son tales que se conserva la relación entre contaminantes, además de que si por alguna razón fuera del control de los analistas, no se tuvieran algunos datos de los máximos mensuales, podrán utilizar los resultados presentados para simular tales datos faltantes a partir de distribuciones condicionales si observados. También, se podrán obtener los llamados tiempos de retorno.

Al final, se encuentra que las densidades conjuntas y bivariadas se ajustan bien a los datos observados, por lo que se puede utilizar como herramienta de prevención de máximos niveles de contaminación en la ZMVM.

Índice general

Capítulo 1

Cópulas

Las cópulas son una de las herramientas más útiles en estadística para modelar y entender la dependencia entre variables aleatorias. Una cópula es una función que une las funciones de distribución marginal de varias variables aleatorias para formar su función de distribución conjunta.

La idea detrás de las cópulas es que permiten separar la estructura de dependencia de las marginales, es decir, se pueden elegir las distribuciones marginales que mejor se ajusten a los datos sin preocuparse por cómo estas interactúan entre sí. Luego, se puede elegir una cópula que capture la dependencia entre las variables.

Las cópulas son especialmente útiles cuando se trabaja con datos de dimensión alta o cuando las relaciones entre las variables no son lineales.

1.1. Primeros conceptos

Para desarrollar la teoría de este capítulo y posteriores, se introduce la notación siguiente:

- $\mathbb{R}$ denota a los números reales.
- $\overline{\mathbb{R}}$ denota a los reales extendidos $[-\infty, \infty]$.
- $\overline{\mathbb{R}}^2$ denota el plano $\overline{\mathbb{R}} \times \overline{\mathbb{R}}$.

A continuación se mencionan algunos conceptos y resultados para la comprensión del concepto de cópula son sus propiedades.

Definición 1.1 (Rectángulo y sus vértices)**.** *Un rectángulo B en $\overline{\mathbb{R}}^2$ es el producto cartesiano de dos intervalos cerrados, es decir, $B = [x_1, x_2] \times [y_1, y_2]$. Los vértices de un rectángulo son los puntos (x_1, y_1), (x_1, y_2), (x_2, y_1) y (x_2, y_2).*

Por lo anterior, se considera un rectángulo especial, la **unidad cuadrada** I^2 es el producto $I \times I$, donde $I = [0, 1]$.

Definición 1.2 (2-función real)**.** *Una 2-función real es una función con dominio $DomH \subset \overline{\mathbb{R}}^2$, y con rango $RanH \subset \mathbb{R}$.*

Definición 1.3 (H-volumen)**.** *Sean S_1 y S_2 subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una 2-función real tal que $DomH = S_1 \times S_2$. Sea $B = [x_1, x_2] \times [y_1, y_2]$ un rectángulo cuyos vértices están en $DomH$. Entonces el H-volumen de B viene dado por*

$$V_H(B) = H(x_2, y_2) - H(x_2, y_1) - H(x_1, y_2) + H(x_1, y_1). \quad (1.1)$$

Definición 1.4 (2-creciente)**.** *Una 2-función real es 2-creciente si $V_H(B) \geq 0$ para todos los rectángulos B cuyos vértices se encuentran en $DomH$.*

Lema 1.1. *Sean S_1 y S_2 subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una función 2-creciente con dominio $S_1 \times S_2$. Sea $x_1, x_2 \in S_1$ con $x_1 \leq x_2$, y sea $y_1, y_2 \in S_2$ con $y_1 \leq y_2$. Entonces, la función $t \to H(t, y_2) - H(t, y_1)$ es no decreciente en S_1, y la función $t \to H(x_2, t) - H(x_1, t)$ es no decreciente en S_2.*

Demostración. Ver [36]. □

Si se supone que S_1 tiene un elemento mínimo a_1 y que S_2 tiene un elemento mínimo a_2. Se dice que una función H de $S_1 \times S_2$ a $\mathbb{R}$ está **conectada a tierra** si $(x, a_2) = 0 = H(a_1, y)$ para todo (x, y) en $S_1 \times S_2$. Por lo tanto se tiene el siguiente Lema.

Lema 1.2. *Sean S_1 y S_2 subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una función 2-creciente conectada a tierra con dominio $S_1 \times S_2$. Entonces H es no decreciente en cada argumento.*

Demostración. Ver [36]. □

Ahora se supone que S_1 tiene un elemento máximo b_1 y que S_2 tiene un elemento máximo b_2. Entonces se dice que una función H de $S_1 \times S_2$ a $\mathbb{R}$ tiene marginales, y que las marginales de H son las funciones F y G dadas por:

$$DomF = S_1, \text{ y } F(x) = H(x, b_2) \text{ para todo } x \in S_1,$$
$$DomG = S_2, \text{ y } G(y) = H(b_1, y) \text{ para todo } y \in S_2.$$

Se cierra esta sección con un Lema importante sobre las funciones 2-crecientes con marginales.

Lema 1.3. *Sean S_1 y S_2 subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una función 2-creciente y conectada a tierra, con marginales con $DomH = S_1 \times S_2$. Sean (x_1, y_1) y (x_2, y_2) puntos cualesquiera en $S_1 \times S_2$. Entonces*

$$|H(x_2, y_2) - H(x_1, y_1)| \leq |F(x_2) - F(x_1)| + |G(y_2) - G(y_1)|.$$

Demostración. Ver [36]. □

1.2. Cópulas

Para comenzar el tema, primero se define el concepto de subcópulas como una cierta clase de funciones 2-crecientes, conectadas a tierra y con marginales continuas; posteriormente se define el concepto de cópula usando las subcópulas con dominio I^2.

Definición 1.5 (Subcópula)**.** *Una subcópula bidimensional (o 2-subcópula, o brevemente, una subcópula) es una función C' con las propiedades siguientes:*

1. *$DomC' = S_1 \times S_2$, donde S_1 y S_2 son subconjuntos de I que contiene a 0 y 1.*
2. *C' está conectada a tierra y es 2-creciente.*
3. *Para cada u en S_1 y cada v en S_2,*

$$C'(u, 1) = 1 \text{ y } C'(1, v) = v.$$

Notar que para cada (u, v) en $DomC'$, $0 \leq C'(u, v) \leq 1$, de modo que $RanC'$ también es un subconjunto de I.

Definición 1.6 (Cópula)**.** *Una cópula bidimensional (o 2-cópula, o brevemente, cópula) es una subcópula C cuyo dominio es I^2.*

De manera equivalente, una cópula es una función C de I^2 a I con las propiedades siguientes:

1. Para cada u, v en I,

$$C(u, 0) = 0 = C(0, v) \tag{1.2a}$$

y

$$C(u, 1) = u \text{ y } C(1, v) = v. \tag{1.2b}$$

2. Para todo u_1, u_2, v_1, v_2 en I tal que $u_1 \leq u_2$ y $v_1 \leq v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \geq 0. \quad (1.3)$$

El Teorema siguiente se obtiene directamente del Lema 1.3, estableciendo la continuidad de las subcópulas.

Teorema 1.1. *Sea C' una subcópula. Entonces, para cada (u_1, u_2), (v_1, v_2) en $DomC'$,*

$$|C'(u_2, v_2) - C'(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

Por tanto, C' es uniformemente continuo en su dominio.

1.3. Teorema de Sklar

El Teorema de Sklar es fundamental para la teoría de las cópulas y es la base de muchas, si no la mayoría, de las aplicaciones de esa teoría a la estadística. El Teorema de Sklar aclara el papel que juegan las cópulas en la relación entre las funciones de distribución multivariadas y sus marginales univariados. Para la demostración de tal Teorema, se necesita los siguientes Lemas.

Lema 1.4. *Sea H una función de distribución conjunta con marginales F y G. Entonces existe una subcópula única C' tal que*

1. *$DomC' = RanF \times RanG$.*
2. *Para todo x, y en $\overline{\mathbb{R}}$, $H(x, y) = C'(F(x), G(y))$.*

Demostración. Ver [36]. □

Lema 1.5. *Sea C' una subcópula. Entonces existe una cópula C tal que $C(u, v) = C'(u, v)$ para todo (u, v) en $DomC'$; es decir, cualquier subcópula puede extenderse a una cópula. La extensión generalmente no es única.*

Demostración. Prueba en [36]. Sea $DomC' = S_1 \times S_2$. Usando el Teorema 1.1 y el hecho de que C' no es decreciente en cada lugar, se puede extender C' por continuidad a una función C'' con dominio $\bar{S}_1 \times \bar{S}_2$, donde $\bar{S}_1$ es la clausura de S_1 y $\bar{S}_2$ es el cierre de S_2. Se observa que C'' también es una subcópula. A continuación, se extiende C'' a una función C con dominio I^2. Con este fin, sea (a, b) cualquier punto en I^2, sean a_1 y a_2, respectivamente, los elementos mayor y menor de $\bar{S}_1$ tales que $a_1 \leq a \leq a_2$; y sean b_1 y b_2, respectivamente, los elementos mayor y menor de $\bar{S}_2$ tales que $b_1 \leq b \leq b_2$. Tenga en cuenta que si a está en $\bar{S}_1$, entonces $a_1 = a = a_2$; y si b está en $\bar{S}_2$, entonces $b_1 = b = b_2$. Ahora sea

$$\lambda_1 = \begin{cases} (a - a_1)/(a_2 - a_1), & \text{si } a_1 < a_2, \\ 1, & \text{si } a_1 = a_2; \end{cases}$$

$$\mu_1 = \begin{cases} (b - b_1)/(b_2 - b_1), & \text{si } b_1 < b_2, \\ 1, & \text{si } b_1 = b_2; \end{cases}$$

y se define

$$\begin{aligned} C(a, b) = {} & (1 - \lambda_1)(1 - \mu_1)C''(a_1, b_1) \\ & + (1 - \lambda_1)\mu_1 C''(a_1, b_2) \\ & + \lambda_1(1 - \mu_1)C''(a_2, b_1) + \lambda_1\mu_1 C(a_2, b_2). \end{aligned} \tag{1.4}$$

Se observa que la interpolación definida en (1.4) es lineal en cada lugar porque λ_1 y μ_1 son lineales en a y b, respectivamente.

Es claro que $DomC = I^2$, que $C(a, b) = C''(a, b)$ para cualquier (a, b) en $DomC''$; y que C satisface (1.2a) y (1.2b). Por tanto, solo se debe demostrar que C satisface (1.3). Para lograr esto, sea (c, d) otro punto en I^2 tal que $c \geq a$ y $d \geq b$, y sean $c_1, d_1, c_2, d_2, \lambda_2, \mu_2$ relacionados con c y d como $a_1, b_1, a_2, b_2, \lambda_1, \mu_1$ están relacionados con a y b. Al evaluar $V_C(B)$ para el rectángulo $B = [a, c] \times [b, d]$, habrá varios casos a considerar, dependiendo de si hay o no un punto en $\bar{S}_1$

estrictamente entre a y c, y si hay o no un punto en $\bar{S}_2$ estrictamente entre b y d. En el más simple de estos casos, no hay ningún punto en $\bar{S}_1$ estrictamente entre a y c, y ningún punto en $\bar{S}_2$ estrictamente entre b y d, de modo que $c_1 = a_1$, $c_2 = a_2$, $d_1 = b_1$ y $d_2 = b_2$. Sustituyendo (1.4) y los términos correspondientes para $C(a,d), C(c,b)$ y $C(c,d)$ en la expresión dada por (1.1) para $V_C(B)$ y simplificando se tiene

$$\begin{aligned} V_C(B) &= V_C([a,c] \times [b,d]) \\ &= (\lambda_2 - \lambda_1)(\mu_2 - \mu_1) V_C([a_1,a_2] \times [b_1,b_2]), \end{aligned}$$

de lo que se sigue que $V_C(B) \geq 0$ en este caso, ya que $c \geq a$ y $d \geq b$ implican $\lambda_2 \geq \lambda_1$ y $\mu_2 \geq \mu_1$.

Por otro lado, el caso menos simple ocurre cuando hay al menos un punto en $\bar{S}_1$ estrictamente entre a y c, y al menos un punto en $\bar{S}_2$ estrictamente entre b y d, de modo que $a < a_2 \leq c_1 < c$ y $b < b_2 \leq d_1 < d$. En este caso, sustituyendo (1.4) y los términos correspondientes para $C(a,d)$, $C(c,b)$ y $C(c,d)$ en la expresión dada por (1.1) para $V_C(B)$ y reordenando los términos produce

$$\begin{aligned} V_C(B) = {} & (1-\lambda_1)\mu_2 V_C([a_1,a_2] \times [d_1,d_2]) \\ & + \mu_2 V_C([a_2,c_1][d_1,d_2]) \\ & + \lambda_2\mu_2 V_C([c_1,c_2] \times [d_1,d_2]) \\ & + (1-\lambda_1) V_C([a_1,a_2] \times [b_2,d_1]) \\ & + V_C([a_2,c_1] \times [b_2,d_1]) + \lambda_2 V_C([c_1,c_2] \times [b_2,d_1]) \\ & + (1-\lambda_1)(1-\mu_1) V_C([a_1,a_2] \times [b_1,b_2]) \\ & + (1-\mu_1) V_C([a_2,c_1] \times [b_1,b_2]) \\ & + \lambda_2(1-\mu_1) V_C([c_1,c_2] \times [b_1,b_2]). \end{aligned}$$

El lado derecho de la expresión anterior es una combinación de nueve cantidades no negativas (los C-volúmenes de los nueve rectángulos determinados) con coeficientes no negativos y, por lo tanto, no es negativa. Los casos restantes son similares, lo que completa la prueba. □

Teorema 1.2 (Teorema de Sklar)**.** *Sea H una función de distribución conjunta con marginales F y G. Así existe una cópula C tal que para todo x, y en $\overline{\mathbb{R}}$,*

$$H(x, y) = C(F(x), G(y)). \tag{1.5}$$

Si F y G son continuas, entonces C es único; de lo contrario, C se determina únicamente en el $RanF \times RanG$. Por el contrario, si C es una cópula y F y G son funciones de distribución, entonces la función H definida por (1.5) es una función de distribución conjunta con marginales F y G.

Demostración. Prueba en [36]. La existencia de una cópula C tal que la ecuación (1.5) se cumple para todo x, y en $\overline{\mathbb{R}}$ se sigue de los Lemas 1.4 y 1.5. Si F y G son continuas, entonces $RanF = RanG = I$, de modo que la única subcópula en el Lema 1.4 es una cópula. Lo contrario es sencillo usando las definiciones de función de distribución. □

La ecuación (1.5) da una expresión para las funciones de distribución conjunta en términos de una cópula y dos funciones de distribución univariadas. Sin embargo (1.5) se puede invertir para expresar cópulas en términos de una función de distribución conjunta y las "inversas" de las dos marginales. Sin embargo, si una marginal no aumenta estrictamente, entonces no posee una inversa en el sentido habitual. Por lo tanto, primero necesitamos definir "quasi-inversas"de funciones de distribución.

Definición 1.7 (quasi-inversa)**.** *Sea F una función de distribución, entonces una quasi-inversa de F es una función $F^{(-1)}$ con dominio I tal que*

1. *Si t está en el $RanF$, entonces $F^{(-1)}(t)$ es algún número x en $\overline{\mathbb{R}}$ tal que $F(x) = t$, es decir, para todo t en el $RanF$,*

$$F(F^{(-1)}(t)) = t.$$

2. Si t no está en el $RanF$, entonces

$$F^{(-t)} = \text{ínf}\{x|F(x) \geq t\} = \sup\{x|F(x) \leq t\}.$$

Si F es estrictamente creciente, entonces solo tiene una quasi-inversa, que es por supuesto la inversa ordinaria, para el cual se usa la notación habitual F^{-1}.

Usando las quasi-inversas de funciones de distribución, ahora se tiene el Corolario siguiente del Lema 1.4.

Corolario 1.1. *Sean H, F, G y C' como en el Lema 1.4, y sean $F^{(-1)}$ y $G^{(-1)}$ las quasi-inversas de F y G, respectivamente, entonces, para cualquier (u, v) en $DomC'$,*

$$C'(u, v) = H(F^{(-1)}(u), G^{(-1)}(v)).$$

Cuando F y G son continuas, el resultado anterior también es válido para las cópulas y proporciona un método para construir cópulas a partir de funciones de distribución conjunta.

Con una extensión apropiada de su dominio a $\overline{\mathbb{R}}^2$, cada cópula es una función de distribución conjunta con marginales que son uniformes en I. Para ser precisos, sea C una cópula y se define la función H_C en $\overline{\mathbb{R}}^2$ mediante

$$H_c(x, y) = \begin{cases} 0, & x < 0 \text{ o } y < 0, \\ C(x, y), & (x, y) \in I^2, \\ x, & y > 1,\ x \in I, \\ y, & x > 1,\ t \in I, \\ 1, & x > 1,\ y > 1. \end{cases}$$

Así, H_C es una función de distribución cuyas marginales son $Unif(0, 1)$. De hecho, a menudo es muy útil pensar en las cópulas como restricciones a I^2 de las funciones de distribución conjunta cuyas marginales son $Unif(0, 1)$.

1.3.1. Cópulas y variables aleatorias bidimensionales

Ahora se esta en condiciones de interpretar el Teorema de Sklar en términos de variables aleatorias y sus funciones de distribución:

Teorema 1.3. *Sea X e Y variables aleatorias con funciones de distribución F y G, respectivamente, y función de distribución conjunta H. Entonces existe una cópula C tal que (1.5) es cierta. Si F y G son continuas, C es única. De lo contrario, C es únicamente determinada en $RanF \times RanG$.*

La cópula C del Teorema 1.3 se denominará cópula de X e Y, y se denominará C_{XY} cuando su identificación con las variables aleatorias X e Y sea ventajosa.

El Teorema siguiente muestra que la cópula del producto $\Pi(u, v) = uv$ caracteriza las variables aleatorias independientes cuando las funciones de distribución son continuas.

Teorema 1.4. *Sean X e Y variables aleatorias continuas. Luego, X e Y son independientes si y sólo si $C_{XY} = \Pi$.*

Demostración. Ver [36]. □

Gran parte de la utilidad de las cópulas en el estudio de la estadística no paramétrica se deriva del hecho de que para las transformaciones estrictamente monótonas de las variables aleatorias, las cópulas son invariantes o cambian de manera predecible. Se recuerda que si la función de distribución de una variable aleatoria X es continua, y si α es una función estrictamente monótona cuyo dominio contiene $RanX$, entonces la función de distribución de la variable aleatoria $\alpha(X)$ también es continua. Se trata primero el caso de las transformaciones estrictamente crecientes.

Teorema 1.5. *Sean X e Y variables aleatorias continuas con cópula C_{XY}. Si α y β aumentan estrictamente en $RanX$ y $RanY$, respectivamente, entonces $C_{\alpha(X)\beta(Y)} = C_{XY}$. Por tanto, C_{XY} es invariante bajo transformaciones estrictamente crecientes de X e Y.*

Demostración. Prueba en [36]. Sean F_1, G_1, F_2 y G_2 las funciones de distribución de X, Y, $\alpha(X)$ y $\beta(Y)$, respectivamente. Como α y β son estrictamente crecientes,

$$F_2(x) = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x)),$$

y de la misma manera $G_2(y) = G_1(\beta^{-1}(y))$. Por tanto, para cualquier x, y en $\overline{\mathbb{R}}$,

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(Y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(X), G_2(Y)). \end{aligned}$$

Como X e Y son continuos, $RanF_2 = RanG_2 = I$, se sigue que $C_{\alpha(X)\beta(Y)} = C_{XY}$ en I^2. □

Cuando al menos uno de α y β es estrictamente decreciente, se obtiene resultados en los que la cópula de las variables aleatorias $\alpha(X)$ y $\beta(Y)$ es una transformación simple de C_{XY}. Específicamente, se tiene:

Teorema 1.6. *Sean X e Y variables aleatorias continuas con cópula C_{XY}. Se deja que a y b sean estrictamente monótonos en $RanX$ y $RanY$, respectivamente.*

1. *Si α es estrictamente creciente y β es estrictamente decreciente, entonces*

$$C_{\alpha(X),\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v).$$

2. *Si α es estrictamente decreciente y β es estrictamente creciente, entonces*

$$C_{\alpha(X),\beta(Y)}(u,v) = v - C_{XY}(1-u,v).$$

3. *Si α y β son ambas estrictamente decrecientes, entonces*

$$C_{\alpha(X),\beta(Y)}(u,v) = u + v - 1 + C_{XY}(1-u,1-v).$$

Demostración. Sean F_1, G_1, F_2 y G_2 las funciones de distribución de X, Y, $\alpha(X)$ y $\beta(Y)$, respectivamente.

Para el caso 1., α es estrictamente creciente, $F_2(x) = F_1(\alpha^{-1}(x))$, y como β es estrictamente decreciente,

$$G_2(y) = P[\beta(Y) \leq y] = 1 - P[Y \leq \beta^{-1}(y)] = 1 - G_1(\beta^{-1}(y)).$$

Para todo x, y en $\overline{\mathbf{R}}$,

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(Y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= P[X \leq \alpha^{-1}(x)](1 - P[Y \leq \beta^{-1}(y)]) \\ &= F_1(\alpha^{-1}(x)) \\ &\quad - P[F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))] \\ &= F_2(x) - C_{XY}(F_2(x), 1 - G_2(y)). \end{aligned}$$

Como X e Y son continuos, $RanF_2 = RanG_2 = I$, se sigue que

$$C_{\alpha(X)\beta(Y)}(u,v) = u - C_{XY}(u, 1-v)$$

en I^2. El segundo caso es análogo al primero.

Para el tercer caso, como α y β son estrictamente decrecientes, $F_2(x) = 1 - F_1(\alpha^{-1}(x))$ y $G_2(y) = 1 - G_1(\beta^{-1}(y))$,

luego para todo x, y en $\overline{\mathbb{R}}$,

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(F_2(x), G_2(Y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= (1 - P[X \leq \alpha^{-1}(x)]) \\ &\quad (1 - P[Y \leq \beta^{-1}(y)]) \\ &= 1 - F_1(\alpha^{-1}(x)) - G_1(\beta^{-1}(y)) + \\ &\quad + C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= F_2(x) + G_2(y) - 1 + \\ &\quad + C_{XY}(1 - F_2(x), 1 - G_2(x)). \end{aligned}$$

Como X e Y son continuos, $RanF_2 = RanG_2 = I$, se sigue que

$$C_{\alpha(X),\beta(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v).$$

en I^2. □

1.3.2. La cópula de Gumbel-Hougaard en 2 variables

La cópula de Gumbel-Hougaard está definida por

$$C_\theta(u, v) = \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{1/\theta}\right), \qquad (1.6)$$

donde $\theta \in [1, \infty)$. Esta cópula representa la dependencia no negativa, siendo que cuando $\theta = 1$ representa dependencia nula, es decir, representa dos variables independientes, y cuando $\theta \to \infty$ representando dependencia perfecta, es decir, de 1. Por este hecho, parece adecuada para representar la dependencia entre contaminantes, ya que se piensa que si hay contaminación de un contaminante, supongamos ozono, los otros contaminantes también tendrían valores altos, por ejemplo el dióxido de nitrógeno.

El paquete OpenBUGS trabaja con funciones de densidad, por lo que es importante calcular la densidad de una distribución dada por la Ecuación (1.5). Sea f, g y h las densidades

asociada a la distribuciones F, G y H respectivamente, entonces

$$\begin{aligned}
h(x,y) &= \frac{\partial^2 H(x,y)}{\partial x \partial y} = \frac{\partial^2 C(F(x),G(y))}{\partial x \partial y} \\
&= \frac{\partial}{\partial x}\left(\frac{\partial C(F(x),G(y))}{\partial y}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial C(F(x),G(y))}{\partial F(x)} \cdot \frac{\partial F(x)}{\partial y}\right. \\
&\quad \left. + \frac{\partial C(F(x),G(y))}{\partial G(y)} \cdot \frac{\partial G(y)}{\partial y}\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial C(F(x),G(y))}{\partial G(y)} g(y)\right) \\
&= \frac{\partial}{\partial x}\left(\frac{\partial C(F(x),G(y))}{\partial G(y)}\right) g(y) \\
&\quad + \frac{\partial C(F(x),G(y))}{\partial G(y)} \cdot \frac{\partial}{\partial x} g(y) \\
&= g(y) \frac{\partial}{\partial G(y)}\left(\frac{\partial C(F(x),G(y))}{\partial x}\right) \\
&= g(y) \frac{\partial}{\partial G(y)}\left(\frac{\partial C(F(x),G(y))}{\partial F(x)} \cdot \frac{\partial F(x)}{\partial x}\right. \\
&\quad \left. + \frac{\partial C(F(x),G(y))}{\partial G(y)} \cdot \frac{\partial G(y)}{\partial x}\right) \\
&= f(x)g(y) \frac{\partial^2 C(F(x),G(y))}{\partial F(x) \partial G(y)},
\end{aligned}$$

y tomando $u = F(x)$ y $v = G(y)$

$$h(x,y) = f(x)g(y) \frac{\partial^2 C(u,v)}{\partial u \partial v}.$$

Dado que la cópula de Gumbel-Hougaard es absolutamente continua, se puede escribir la densidad bivariada de la cópu-

la como

$$
\begin{aligned}
c_\theta(u,v) = \frac{\partial^2 C_\theta(u,v)}{\partial u \partial v} &= \frac{\partial}{\partial u}\left[\frac{\partial C_\theta(u,v)}{\partial v}\right] \\
&= \frac{\partial}{\partial u}\left[\exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right)(-1)\left(\frac{1}{\theta}\right)\right. \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}-1}\theta(-\log\ v)^{\theta-1}\left(-\frac{1}{v}\right)\right] \\
&= \frac{(-\log\ v)^{\theta-1}}{v}\frac{\partial}{\partial u}\left[\exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right)\right. \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}-1}\right] \\
&= \frac{(-\log\ v)^{\theta-1}}{v}\left[\frac{(-\log\ u)^{\theta-1}}{u}\right. \\
&\times \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right) \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{2}{\theta}-2}\right. \\
&+\exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right)\left(\frac{1}{\theta}-1\right) \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}-2}\theta(-\log\ u)^{\theta-1}\left(-\frac{1}{u}\right)\right] \\
&= \frac{(-\log\ v)^{\theta-1}}{v}\left[\frac{(-\log\ u)^{\theta-1}}{u}\right. \\
&\times \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right) \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{2}{\theta}-2} + \frac{(-\log\ u)^{\theta-1}}{u}\right. \\
&\times \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right)(\theta-1) \\
&\left.\times\ \left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}-2}\right]
\end{aligned}
$$

es decir,

$$\begin{aligned} c_\theta(u,v) = & \frac{(-\log\ u)^{\theta-1}(-\log\ v)^{\theta-1}}{uv} \\ & \times \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}}\right) \\ & \times \left(\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{2}{\theta}-2}\right. \\ & \times \left. +(\theta-1)\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{\frac{1}{\theta}-2}\right). \quad (1.7) \end{aligned}$$

Por tanto, la densidad conjunta para dos variables aleatorias X e Z, se puede escribir como

$$f_{Y,Z,\theta}(y,z) = f_Y(y) f_Z(z) c_\theta(F_Y(y), F_Z(z)), \qquad (1.8)$$

donde c_θ es representada por la ecuación (1.7), F denota funciones de distribución, f funciones de densidad y los subíndices a que variable pertenece.

1.4. Cópulas Arquímedianas

Definición 1.8. *Sea φ una función continua y estrictamente decreciente de I a $[0,\infty]$ tal que $\varphi(1) = 0$. La pseudoinversa de φ es la función $\varphi^{[-1]}$ con $Dom\varphi^{[-1]} = [0,\infty]$ y $Ran\varphi^{[-1]} = I$ dado por*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \le t \le \varphi(0), \\ 0, & \varphi(0) \le t \le \infty. \end{cases}$$

Lema 1.6. *Sea φ una función continua y estrictamente decreciente de I a $[0,\infty]$ tal que $\varphi(1) = 0$, y sea $\varphi^{[-1]}$ la pseudoinversa de φ. Sea C una función de I^2 a I definida por*

$$C(u,v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)), \qquad (1.9)$$

es una cópula.

Demostración. Ver [36]. □

A la familia de cópulas que cumplen la Ecuación (1.9) se le conoce como cópulas Arquímedianas. Esta familia de cópulas son de mucha utilidad en la práctica, debido a que se puede pueden generar cópulas a partir de una función φ como en el Lema 1.6, además de que por la misma función es relativamente sencillo generar variables aleatorias, como se indica en [26].

La cópula de Gumbel-Hougaard forma parte de las llamadas cópulas Arquímedianas, y está generada por $\varphi(t) = (-\log\ t)^{\theta}$.

1.5. Dependencia

Hay una variedad de formas de medir la dependencia y algunas de estas propiedades y medidas son invariantes bajo escala, es decir, permanecen sin cambios bajo transformaciones estrictamente crecientes de las variables aleatorias. Las propiedades de dependencia y las medias de asociación están interrelacionadas, y las más conocidas son las versiones poblacionales de la τ de Kendall y la ρ de Spearman, las cuales miden una forma de dependencia conocida como concordancia. Dentro de las aplicaciones de tales medidas se encuentran la biología [5], la clasificación de cópulas [39], psicología [48] e hidrología [55].

En palabras simples, un par de variables aleatorias son concordantes si los valores "grandes" de una tienden a estar asociados con los valores "grandes" de la otra y los valores "pequeños" de una con los valores "pequeños" de la otra. De forma más precisa, sean (x_i, y_i) y (x_j, y_j) dos observaciones de un vector (X, Y) de variables aleatorias continuas. Se dice que (x_i, y_i) y (x_j, y_j) son concordantes si $x_i < x_j$ e $y_i < y_j$, o

si $x_i > x_j$ e $y_i > y_j$. De manera similar, Se dice que (x_i, y_i) y (x_j, y_j) son discordantes si $x_i < x_j$ e $y_i > y_j$ o si $x_i > x_j$ e $y_i < y_j$. Una formulación alternativa es: (x_i, y_i) y (x_j, y_j) son concordantes si $(x_i - x_j)(y_i - y_j) > 0$ y discordantes si $(x_i - x_j)(y_i - y_j) < 0$.

1.5.1. τ de Kendall

La versión muestral de la medida de asociación τ de Kendall se define en términos de concordancia de la manera siguiente: Sea $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ una muestra aleatoria de n observaciones de un vector (X, Y) de variables aleatorias continuas. Hay $\binom{n}{2}$ pares distintos (x_i, y_i) e (x_j, y_j) de observaciones en la muestra, y cada par es concordante o discordante; sea c el número de pares concordantes y d el número de pares discordantes. Entonces, la τ de Kendall para la muestra se define como

$$t = \frac{c-d}{c+d} = \frac{c-d}{\binom{n}{2}}.$$

De manera equivalente, t es la probabilidad de concordancia menos la probabilidad de discordancia para un par de observaciones (x_i, y_i) y (x_j, y_j) que se eligen aleatoriamente de la muestra. La versión poblacional de la τ de Kendall para un vector (X, Y) de variables aleatorias continuas con función de distribución conjunta H se define de manera similar. Sean (X_1, Y_1) y (X_2, Y_2) vectores aleatorios independientes e idénticamente distribuidos, cada uno con una función de distribución conjunta H. Luego, la versión poblacional de la τ de Kendall se define como la probabilidad de concordancia menos la probabilidad de discordancia:

$$\begin{aligned}\tau &= \tau_{X,Y} \\ &= P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].\end{aligned}$$

Primero se define una función de concordancia Q, que es la diferencia de las probabilidades de concordancia y discordancia entre dos vectores (X_1, Y_1) y (X_2, Y_2) de variables aleatorias continuas con (posiblemente) diferentes distribuciones conjuntas H_1 y H_2 , pero con marginales comunes F y G. Luego se mostrará que esta función depende de las distribuciones de (X_1, Y_1) y (X_2, Y_2) sólo a través de sus cópulas.

Teorema 1.7. *Sean (X_1, Y_1) y (X_2, Y_2) vectores independientes de variables aleatorias continuas con funciones de distribución conjunta H_1 y H_2, respectivamente, con marginales comunes F (de X_1 y X_2) y G (de Y_1 e Y_2). Sean C_1 y C_2 las cópulas de (X_1, Y_1) y (X_2, Y_2), respectivamente, de modo que $H_1(x, y) = C_1(F(x), G(y))$ y $H_2(x, y) = C_2(F(x), G(y))$. Sea Q la diferencia entre las probabilidades de concordancia y discordancia de (X_1, Y_1) y (X_2, Y_2), es decir, sea*

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0],$$

entonces,

$$Q = Q(C_1, C_2) = 4 \iint_{I^2} C_2(u, v) dC_1(u, v) - 1. \qquad (1.10)$$

Demostración. Ver [36]. □

Corolario 1.2. *Para C_1, C_2 y Q dadas como en el Teorema 1.7, se tiene que Q es simétrica en sus argumentos, es decir, $Q(C_1, C_2) = Q(C_2, C_1)$.*

El Teorema anterior permite establecer el valor de τ cuando sólo se tiene un par de variables aleatorias continuas X e Y relacionadas a partir de la con cópula C, tal como se muestra a continuación.

Teorema 1.8. *Sean X e Y variables aleatorias continuas cuya cópula es C. Entonces la versión poblacional de la τ de*

Kendall para X e Y (que se denota por $\tau_{X,Y}$ o τ_C) está dada por

$$\tau_{X,Y} = \tau_C = Q(C,C) = 4\iint_{I^2} C(u,v)dC(u,v) - 1. \quad (1.11)$$

Ejemplo 1.1. *Sea C un miembro de la familia Farlie-Gumbel-Morgenstern (FGM) dada por*

$$C_\theta(u,v) = uv + \theta uv(1-u)(1-v),$$

donde $\theta \in [-1,1]$. Dado que C_θ es absolutamente continua, se tiene que

$$dC_\theta(u,v) = \frac{\partial^2 C_\theta(u,v)}{\partial u \partial v}dudv = [1+\theta(1-2u)(1-2v)]dudv,$$

de modo que

$$\iint_{I^2} C_\theta(u,v)dC_\theta(u,v) = \frac{1}{4} + \frac{\theta}{18},$$

por tanto $\tau_\theta \in [-2/9, 2/9]$.

El Teorema siguiente permite calcular la integral de Q en la ecuación (1.10), y en especial τ_C en la ecuación (1.11).

Teorema 1.9. *Sean C_1 y C_2 cópulas, entonces*

$$\iint_{I^2} C_1(u,v)dC_2(u,v) = \frac{1}{2} - \iint_{I^2} \frac{\partial}{\partial u}C_1(u,v)\frac{\partial}{\partial v}C_2(u,v)dudv.$$

Demostración. Ver [36]. □

1.5.2. ρ de Spearman

La versión poblacional de la medida de asociación conocida como ρ de Spearman[1] se basa en la concordancia y la discordancia. Para obtener la versión poblacional de esta medida,

[1]También se le conoce como coeficiente de rangos de Spearman o coeficiente de rangos ordenados, que mide mejor la relación entre los datos ordenados.

sean (X_1, Y_1), (X_2, Y_2) y (X_3, Y_3) tres vectores aleatorios independientes con una función común de distribución conjunta H (cuyas marginales son F y G) y cópula C. La versión poblacional $\rho_{X,Y}$ de la ρ de Spearman se define como proporcional a la probabilidad de concordancia menos la probabilidad de discordancia para los dos vectores (X_1, Y_2) y (X_2, Y_3), es decir, un par de vectores con las mismas marginales, pero un vector tiene función de distribución H, mientras que los componentes del otro son independientes:

$$\rho_{X,Y} = 3(P[(X_1-X_2)(Y_1-Y_3) > 0]-P[(X_1-X_2)(Y_1-Y_3) < 0]).$$

Mientras la función de distribución conjunta de (X_1, Y_1) es $H(x, y)$, la función de distribución conjunta de (X_2, Y_3) es $F(x)G(y)$, ya que X_2 e Y_3 son independientes. Así, la cópula de X_2 e Y_3 es Π, y usando el Teorema 1.7 y el Corolario 1.2, se tiene el Teorema siguiente.

Teorema 1.10. *Sean X e Y variables aleatorias continuas cuya cópula es C. Entonces la versión poblacional de la ρ de Spearman para X e Y (denotada por $\rho_{X,Y}$ o ρ_C) está dada por*

$$\begin{aligned}\rho_{X,Y} = \rho_C &= 3Q(C, \Pi), \\ &= 12 \iint_{I^2} uvd(C(u, v)) - 3, \\ &= 12 \iint_{I^2} C(u, v)dudv - 3.\end{aligned}$$

Ejemplo 1.2. *Regresando a la cópula FGM en el Ejemplo 1.1,*

$$\iint_{I^2} C_\theta(u, v)dudv = \frac{1}{4} + \frac{\theta}{36},$$

por tanto, $\rho_\theta \in [-1/3, 1/3]$.

1.5.3. Dependencia de las colas

A continuación se muestra los conceptos básicos de dependencia en las colas. En varias ocasiones no solo se necesita dependencia en general, sino asegurar la dependencia para valores grandes o pequeños.

Definición 1.9 (Dependencia en la cola superior e inferior)**.** *Sean X e Y variables aleatorias continuas con funciones de distribución F y G, respectivamente. El parámetro de dependencia de la cola superior λ_U es el límite (si existe) de la probabilidad condicional de que Y sea mayor que el 100-percentil de G dado que X es mayor que el 100-percentil de F cuando t se aproxima a 1, es decir,*

$$\lambda_U = \lim_{t \to 1^-} P[Y > G^{(-1)}(t) | X > F^{(-1)}(t)].$$

De manera similar, el parámetro de dependencia de la cola inferior λ_L es el límite (si existe) de la probabilidad condicional de que Y sea menor o igual al 100-percentil de G dado que X es menor o igual al 100-percentil de F cuando t tiende a 0, es decir,

$$\lambda_L = \lim_{t \to 0^+} P[Y \leq G^{(-1)}(t) | X \leq F^{(-1)}(t)].$$

Teorema 1.11. *Sean X, Y, F, G, λ_U y λ_L como en la Definición 1.9, y sea C la cópula de X e Y, entonces*

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - C(t,t)}{1 - t},$$

y

$$\lambda_L = \lim_{t \to 0^+} \frac{C(t,t)}{t}.$$

Demostración. Ver [36]. □

1.6. Cópulas máx-estables

Una propiedad importante en las cópulas es que sean dependientes para valores extremos. Esta idea esta relacionada con las propiedades de obtener una cópula para valores extremos.

Definición 1.10 (Cópula máx-estable)**.** *Una cópula C es máx-estable si para cada $r > 0$ y todos $u, v \in I$*

$$C(u, v) = C^r(u^{1/r}, v^{1/r}).$$

Teorema 1.12. *Si C es una cópula y n un entero positivo, entonces la función*

$$C_{(n)}(u, v) = C^n(u^{1/n}, v^{1/n}) \tag{1.12}$$

para $u, v \in I$ es una cópula. Además, si (X_i, Y_i), para $i = 1, 2, \ldots, n$ son pares de variables aleatorias independientes e idénticamente distribuidas (iid) con cópula C, entonces $C_{(n)}$ es la cópula de $X_{(n)} = \max\{X_i\}$ y $Y_{(n)} = \max\{Y_i\}$.

Demostración. Ver [36]. □

Ahora se considera el caso en donde n tiende a infinito en la Ecuación (1.12). Se observa un parecido a los conceptos de valores extremos presentados en el Capítulo 2. Como consecuencia del Teorema 1.12 se tiene la definición siguiente.

Definición 1.11 (Cópula de valor extremo)**.** *Una cópula C_* es una cópula de valor extremo si existe una cópula C tal que*

$$C_*(u, v) = \lim_{n \to \infty} C^n(u^{1/n}, v^{1/n}),$$

para $u, v \in I$.

Teorema 1.13. *Una cópula es máx-estable si y solo si es una cópula de valor extremo.*

Demostración. Ver [36]. □

1.7. Cópulas multivariadas

Las Definiciones y Teoremas presentados sobre las cópulas en 2 dimensiones, en su mayoría se pueden formular en más dimensiones, e incluso sus demostraciones son similares. Con esto en mente, se presentan ahora los resultados y Definiciones que son importantes el caso de estudio presentado en esta tesis.

Para cualquier entero positivo n, se denota al espacio vectorial de dimensión n, $\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ como $\mathbb{R}^n$. De manera análoga, se denota al espacio de dimensión n extendido $\overline{\mathbb{R}} \times \overline{\mathbb{R}} \times \cdots \times \overline{\mathbb{R}}$ por $\overline{\mathbb{R}}^n$.

Se denota por letras negritas a los vectores $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ en $\mathbb{R}^2$ (o $\overline{\mathbb{R}}^n$). Se consideran dos vectores son $\mathbf{a} \leq \mathbf{b}$ en $\mathbb{R}^2$ (o $\overline{\mathbb{R}}^n$) si y solo si $a_k \leq b_k$ para todo $k \in \{1, 2, \ldots, n\}$. Para $\mathbf{a} \leq \mathbf{b}$ se puede considerar el rectángulo de n-dimensional $[\mathbf{a}, \mathbf{b}]$ por $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, con lo que se puede definir el cubo unitario por I^n como el producto cartesiano de n intervalos unidad $I = [0, 1]$. Dado $B = [\mathbf{a}, \mathbf{b}]$, los puntos $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ donde c_k es igual a a_k o b_k son los **vértices del rectángulo** B. Por último se considera una n-función real H como es función con dominio $DomH \subset \overline{\mathbb{R}}^n$, y con rango $RanH \subset \mathbb{R}$.

Definición 1.12 (H-volumen)**.** *Sea $S_1, S_2, \ldots, S_n$ subconjuntos no vacíos de $\overline{\mathbb{R}}^n$ y H una n-función real con $DomH = S_1 \times S_2 \times \cdots \times S_n$. Sea $B = [\mathbf{a}, \mathbf{b}] \subset DomH$, entonces el H-volumen de B esta dado por*

$$V_H(B) = \sum sgn(\mathbf{c}) H(\mathbf{c}),$$

donde la suma es tomada sobre todos los vértices $\mathbf{c}$ de B, y

$$sgn(\mathbf{c}) = \begin{cases} 1, & \text{si } c_k = a_k \text{ para un número par de } k\text{'s}, \\ -1, & \text{si } c_k = a_k \text{ para un número impar de } k\text{'s}. \end{cases}$$

Para ejemplificar, se supone que H es una 3-función real con $DomH = \overline{\mathbb{R}}^3$, y un rectángulo $B = [x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$. El H-volumen de B es :

$$\begin{aligned} V_H(B) = & H(x_2, y_2, z_2) - H(x_2, y_2, z_1) - H(x_2, y_1, z_2) \\ & - H(x_1, y_2, z_2) + H(x_2, y_1, z_1) + H(x_1, y_2, z_1) \\ & + H(x_1, y_1, z_2) - H(x_1, y_1, z_1). \end{aligned}$$

Ahora se supone una n-función H con $DomH = S_1 \times S_2 \times \cdots \times S_n$, donde cada S_k tiene un elemento mínimo a_k, se dice que H está **conectada a tierra** si $H(\mathbf{t}) = 0$ para toda $\mathbf{t} \in DomH$ tal que $t_k = a_k$ para al menos algún k. Si $S_k \neq \emptyset$ para cada k y cada S_k tiene un elemento máximo b_k, entonces se dice que H tiene **marginales**, y las marginales unidimensionales de H son dadas por las funciones H_k con $DomH_k = S_k$ y

$$H_k(x) = H(b_1, \ldots, b_{k-1}, x, b_{k+1}, \ldots, b_n), \quad x \in S_k. \tag{1.13}$$

Definición 1.13 (n-creciente)**.** *Una n-función real H es n-creciente si $V_H(B) \geq 0$ para todos los rectángulos B cuyos vértices se encuentran en $DomH$.*

Lema 1.7. *Sean $S_1, S_2, \ldots, S_n$ subconjuntos no vacíos de $\overline{\mathbb{R}}^2$, y sea H una función n-creciente con $DomH = S_1 \times S_2 \times \cdots \times S_n$. Entonces H es no decreciente en cada argumento, es decir, si*

$$(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n)$$

y

$$(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n$$

en el $DomH$ y $x < y$, entonces

$$H(t_1, \ldots, t_{k-1}, x, t_{k+1}, \ldots, t_n) \leq H(t_1, \ldots, t_{k-1}, y, t_{k+1}, \ldots, t_n).$$

Demostración. Ver [47]. □

El resultado siguiente es análogo al Lema 1.3, el cual es necesario para demostrar que las n-cópulas son uniformemente continuas, y la prueba del Teorema de Sklar para dimensiones mayores a 2.

Lema 1.8. *Sean $S_1, S_2, \ldots, S_n$ subconjuntos no vacíos de $\overline{\mathbb{R}}$, y sea H una función n-creciente y conectada a tierra, con marginales con $DomH = S_1 \times S_2 \times \cdots \times S_n$. Sean $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ e $\mathbf{y} = (y_1, y_2, \ldots, y_n)$ puntos cualesquiera en $S_1 \times S_2 \times \cdots \times S_n$. Entonces*

$$|H(\mathbf{x}) - H(\mathbf{y})| \leq \sum_{i=1}^{n} |H_k(x_k) - H_k(y_k)|.$$

Las Definiciones siguientes son análogas a las Definiciones 1.5 y 1.6 respectivamente.

Definición 1.14 (Subcópula)**.** *Una subcópula n-dimensional (o n-subcópula, o brevemente, una subcópula) es una función C' con las propiedades siguientes:*

1. *$DomC' = S_1 \times S_2 \times \cdots \times S_n$, donde cada $S_k \subset I$ que contiene a 0 y 1.*
2. *C' está conectada a tierra y es n-creciente.*
3. *C' tiene marginales (unidimensionales) C'_k, $k = 1, 2, \ldots, n$, que satisfacen*

$$C'(u) = u, \quad u \in S_k. \tag{1.14}$$

Definición 1.15 (Cópula)**.** *Una cópula n-dimensional (o n-cópula, o simplemente, cópula) es una subcópula C cuyo dominio es I^n.*

De manera equivalente, una cópula es una función C de I^n a I con las propiedades siguientes:

1. Para cada $\mathbf{u}$ en I^n,

$$C(\mathbf{u}) = 0$$

si al menos una coordenada de $\mathbf{u}$ es 0, y

$$C(\mathbf{u}) = u_k,$$

si todas las coordenadas de $\mathbf{u}$ son 1 excepto u_k.

2. Para todo $\mathbf{a}, \mathbf{b} \in I^n$ tal que $\mathbf{a} \leq \mathbf{b}$,

$$V_C([\mathbf{a}, \mathbf{b}]) \geq 0.$$

Como consecuencia del Lema 1.8, se tiene que las funciones n-crecientes son uniformemente continuas, y por tanto también las cópulas son uniformemente continuas.

Teorema 1.14. *Sea C' una subcópula. Entonces, para cada $\mathbf{u}, \mathbf{v} \in DomC'$,*

$$|C'(\mathbf{v}) - C'(u)| \leq \sum_{i=1}^{n} |v_k - u_k|.$$

Por tanto, C' es uniformemente continua en su dominio.

Para tener el Teorema de Sklar para dimensiones mayores a 2, antes se debe dar la Definición siguiente:

Definición 1.16. *Una función de distribución (n-dimensional) es una función H con $DomH = \overline{\mathbb{R}}^n$ tal que:*

1. *H es n-creciente,*
2. *$H(\mathbf{t}) = 0$ para todo $\mathbf{t} \in \overline{\mathbb{R}}^n$ tal que $t_k = -\infty$ para al menos un $k \in \mathbb{N}$ entre 1 y n, y $H(\infty, \infty, \ldots, \infty) = 1$.*

Por lo anterior y del Lema 1.7, se deduce que las marginales dadas en la Ecuación (1.13) de una función de distribución (n-dimensional) son funciones de distribución.

Teorema 1.15 (Teorema de Sklar)**.** *Sea H una función de distribución n-dimensional con marginales $F_1, F_2, \ldots, F_n$. Así, existe una cópula C tal que para todo $\mathbf{x} \in \overline{\mathbb{R}}$,*

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n)). \quad (1.15)$$

Si $F_1, F_2, \ldots, F_n$ son continuas, entonces C es única; de lo contrario, C se determina únicamente en $RanF_1 \times RanF_2 \times \cdots \times RanF_n$. Por el contrario, si C es una cópula y $F_1, F_2, \ldots, F_n$ son funciones de distribución, entonces la función H definida por (1.15) es una función de distribución conjunta con marginales $F_1, F_2, \ldots, F_n$.

Demostración. Ver [12, 33, 49] □

1.7.1. La cópula de Gumbel-Hougaard en 3 variables

La cópula de Gumbel-Houggard en tres dimensiones está representada por

$$\begin{aligned} &C_\theta(u_1, u_2, u_3) \\ =&\exp\left(-\left((-\log\ u_1)^\theta + (-\log\ u_2)^\theta + (-\log\ u_3)^\theta\right)^{1/\theta}\right), \end{aligned} \quad (1.16)$$

donde $\theta \in [1, \infty]$.

Al igual que se hizo en el caso de dos dimensiones, la función de densidad en para tres componentes esta dada por

$$\begin{aligned} &f_{X_1,X_2,X_3,\theta}(x_1, x_2, x_3) \\ =&f_{X_1}(x_1) f_{X_2}(x_2) f_{X_3}(x_3) \\ &\times c_\theta(F_{X_1}(x_1), F_{X_2}(x_2), F_{X_3}(x_3)), \end{aligned} \quad (1.17)$$

donde las F's representan las funciones de distribución de las variables aleatorias, las f's las funciones de densidad de

probabilidad y las X's las variables aleatorias, donde

$$c_\theta(u_1, u_2, u_3) = \frac{\partial^3 C_\theta(u_1, u_2, u_3)}{\partial u_1 \partial u_2 \partial u_3}.$$

Así, la derivada de la cópula de Gumbel-Hougaard está dada por

$$\begin{aligned} c_\theta(u_1, u_2, u_3) = \frac{1}{u_1 u_2 u_3} &(-\log u_1)^{\theta-1}(-\log u_2)^{\theta-1} \\ &\times (-\log u_3)^{\theta-1} \exp(-\kappa^{1/\theta})\kappa^{-3+1/\theta} \\ &\times [1 + 2\theta^2 + 3\theta(-1 + \kappa^{1/\theta}) \\ &- 3\kappa^{1/\theta} + \kappa^{2/\theta}], \end{aligned} \tag{1.18}$$

donde

$$\kappa = (-\log u_1)^\theta + (-\log u_2)^\theta + (-\log u_3)^\theta.$$

1.7.2. Cópulas Asimétricas

Una generalización para las cópulas Arquímedianas son las denominadas "cópulas asimétricas", que para el caso de $n = 3$, está dada por

$$C_1(u_3, C_2(u_2, u_1)).$$

Si las cumple que $C(u_1, u_2) = C(u_2, u_1)$, entonces

$$C_1(C_2(u_1, u_2), u_3). \tag{1.19}$$

Sean θ_1, θ_2 los parámetros de C_1 y C_2 las cópulas en la Ecuación (1.19) entonces $\theta_1 < \theta_2$, donde las cópulas pertenecen a la misma familia. Para mayor detalle sobre la construcción y condiciones, revisar [21]. Por consiguiente, la función de distribución con marginales $F(x)$, $G(y)$ y $H(z)$ es

$$C_1(C_2(F(x), G(y)), H(z)). \tag{1.20}$$

Ahora, se calcula la función de densidad asociada a la función de distribución (1.20) Sean f, g y h las densidades asociadas a la distribuciones F, G y H respectivamente,

$$
\begin{aligned}
&\frac{\partial^3 C_1[C_2(F(x),G(y)),H(z)]}{\partial x\partial y\partial z}\\
=&\frac{\partial^2}{\partial x\partial y}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial z}\right)\\
=&\frac{\partial^2}{\partial x\partial y}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\cdot\frac{\partial C_2(F(x),G(y))}{\partial z}\right.\\
&\left.+\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot\frac{\partial H(z)}{\partial z}\right)\\
=&\frac{\partial^2}{\partial x\partial y}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot h(z)\right)\\
=&h(z)\frac{\partial^2}{\partial x\partial y}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\right)\\
=&h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\right)\right]\\
=&h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial y}\right)\right]\\
=&h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\frac{\partial C_2(F(x),G(y))}{\partial y}\\
&\left.\left.+\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot\frac{\partial H(z)}{\partial y}\right)\right]\\
=&h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\left\{\frac{\partial C_2(F(x),G(y))}{\partial F(x)}\cdot\frac{\partial F(x)}{\partial y}\right.\\
&\left.\left.\left.+\frac{\partial C_2(F(x),G(y))}{\partial G(y)}\cdot\frac{\partial G(y)}{\partial y}\right\}\right)\right],
\end{aligned}
$$

y continuando con las operaciones,

$$
\begin{aligned}
&\frac{\partial^3 C_1[C_2(F(x),G(y)),H(z)]}{\partial x\partial y\partial z}\\
=&h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\left.\left.\times\frac{\partial C_2(F(x),G(y))}{\partial G(y)}\cdot g(y)\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\left.\left.\times\frac{\partial C_2(F(x),G(y))}{\partial G(y)}\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\frac{\partial C_2(F(x),G(y))}{\partial G(y)}\\
&\left.\left.+\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot\frac{\partial H(z)}{\partial G(y)}\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial x}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial G(y)}\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial x}\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\frac{\partial C_2(F(x),G(y))}{\partial x}\\
&\left.\left.+\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot\frac{\partial H(z)}{\partial x}\right)\right]\\
=&g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\left\{\frac{\partial C_2(F(x),G(y))}{\partial F(x)}\cdot\frac{\partial F(x)}{\partial x}\right.\\
&\left.\left.\left.+\frac{\partial C_2(F(x),G(y))}{\partial G(y)}\cdot\frac{\partial G(y)}{\partial x}\right\}\right)\right],
\end{aligned}
$$

por consiguiente,

$$
\begin{aligned}
&\frac{\partial^3 C_1[C_2(F(x),G(y)),H(z)]}{\partial x\partial y\partial z}\\
=&g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\left.\left.\times\ \frac{\partial C_2(F(x),G(y))}{\partial F(x)}\cdot f(x)\right)\right],\\
=&f(x)g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\left.\left.\times\ \frac{\partial C_2(F(x),G(y))}{\partial F(x)}\right)\right]\\
=&f(x)g(y)h(z)\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial C_2(F(x),G(y))}\right.\right.\\
&\times\ \frac{\partial C_2(F(x),G(y))}{\partial F(x)}\\
&\left.\left.+\ \frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial H(z)}\cdot\frac{\partial H(z)}{\partial F(x)}\right)\right]\\
=&f(x)g(y)h(z)\\
&\times\frac{\partial}{\partial G(y)}\left[\frac{\partial}{\partial H(z)}\left(\frac{\partial C_1[C_2(F(x),G(y)),H(z)]}{\partial F(x)}\right)\right]\\
=&f(x)g(y)h(z)\frac{\partial^3 C_1[C_2(F(x),G(y)),H(z)]}{\partial F(x)\partial G(y)\partial H(z)}.
\end{aligned}
$$

y tomando $u = F(x)$, $v = G(y)$ y $w = H(Z)$, se tiene que

$$
\begin{aligned}
&\frac{\partial^3 C_1[C_2(F(x),G(y)),H(z)]}{\partial x\partial y\partial z}\\
=&f(x)g(y)h(z)\frac{\partial^3 C_1[C_2(u,v),w]}{\partial u\partial v\partial w}.
\end{aligned}
$$

Ahora se calcula la última derivada, que se denota por $c_{\theta_1,\theta_2}(u,v,w)$ y será la densidad asociada a la distribución

(1.19), la cual está dada por

$$\begin{aligned} c_{\theta_1,\theta_2}(u,v,w) &= \frac{\partial^3 C_1[C_2(u,v),w]}{\partial u \partial v \partial w} \\ &= \frac{\partial^2}{\partial u \partial v}\left(\frac{\partial C_1[C_2(u,v),w]}{\partial w}\right) \\ &= \frac{\partial^2}{\partial u \partial v}\left(\frac{\partial C_1[C_2(u,v),w]}{\partial C_2(u,v)} \cdot \frac{\partial C_2(u,v)}{\partial w}\right. \\ &\quad \left. + \frac{\partial C_1[C_2(u,v),w]}{\partial w} \cdot \frac{\partial w}{\partial w}\right) \\ &= \frac{\partial^2}{\partial u \partial v}\left(\frac{\partial C_1[C_2(u,v),w]}{\partial w}\right). \end{aligned}$$

Sea

$$C_1^{(0,1)}[C_2(u,v),w] := \frac{\partial C_1[C_2(u,v),w]}{\partial w},$$

por lo que

$$\begin{aligned} c_{\theta_1,\theta_2}(u,v,w) &= \frac{\partial^2}{\partial u \partial v}\left[C_1^{(0,1)}[C_2(u,v),w]\right] \\ &= \frac{\partial}{\partial u}\left[\frac{\partial}{\partial v}\left(C_1^{(0,1)}[C_2(u,v),w]\right)\right] \\ &= \frac{\partial}{\partial u}\left[\frac{\partial C_1^{(0,1)}[C_2(u,v),w]}{\partial C_2(u,v)} \cdot \frac{\partial C_2(u,v)}{\partial v}\right. \\ &\quad \left. + \frac{\partial C_1^{(0,1)}[C_2(u,v),w]}{\partial w} \cdot \frac{\partial w}{\partial v}\right] \\ &= \frac{\partial}{\partial u}\left[\frac{\partial C_1^{(0,1)}[C_2(u,v),w]}{\partial C_2(u,v)} \cdot \frac{\partial C_2(u,v)}{\partial v}\right]. \end{aligned}$$

Tomando

$$C_1^{(1,1)}[C_2(u,v),w] := \frac{\partial C_1^{(0,1)}[C_2(u,v),w]}{\partial C_2(u,v)},$$

simplificando a

$$
\begin{aligned}
& c_{\theta_1,\theta_2}(u,v,w) \\
=& \frac{\partial}{\partial u}\left[C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial C_2(u,v)}{\partial v}\right] \\
=& \frac{\partial}{\partial u}\left[C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial C_2(u,v)}{\partial v}\right] \\
=& \frac{\partial C_1^{(1,1)}[C_2(u,v),w]}{\partial u}\cdot\frac{\partial C_2(u,v)}{\partial v} \\
& + C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v} \\
=& \left(\frac{\partial C_1^{(1,1)}[C_2(u,v),w]}{\partial C_2(u,v)}\cdot\frac{\partial C_2(u,v)}{\partial u}\right. \\
& \left. + \frac{\partial C_1^{(1,1)}[C_2(u,v),w]}{\partial w}\cdot\frac{\partial w}{\partial u}\right) \\
& \times\frac{\partial C_2(u,v)}{\partial v} + C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v} \\
=& \frac{\partial C_1^{(1,1)}[C_2(u,v),w]}{\partial C_2(u,v)}\cdot\frac{\partial C_2(u,v)}{\partial u}\cdot\frac{\partial C_2(u,v)}{\partial v} \\
& + C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v} \\
=& \frac{\partial C_1^{(1,1)}[C_2(u,v),w]}{\partial C_2(u,v)}\cdot\frac{\partial C_2(u,v)}{\partial u}\cdot\frac{\partial C_2(u,v)}{\partial v} \\
& + C_1^{(1,1)}[C_2(u,v),w]\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v} \\
=& \frac{\partial^2 C_1^{(0,1)}[C_2(u,v),w]}{\partial(C_2(u,v))^2}\cdot\frac{\partial C_2(u,v)}{\partial u}\cdot\frac{\partial C_2(u,v)}{\partial v} \\
& + \frac{\partial C_1^{(0,1)}[C_2(u,v),w]}{\partial C_2(u,v)}\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v} \\
=& \frac{\partial^3 C_1[C_2(u,v),w]}{\partial(C_2(u,v))^2\partial w}\cdot\frac{\partial C_2(u,v)}{\partial u}\cdot\frac{\partial C_2(u,v)}{\partial v} \\
& + \frac{\partial^2 C_1[C_2(u,v),w]}{\partial C_2(u,v)\partial w}\cdot\frac{\partial^2 C_2(u,v)}{\partial u\partial v},
\end{aligned}
$$

es decir,

$$\begin{aligned} & c_{\theta_1,\theta_2}(u,v,w) \\ =& \frac{\partial^3 C_1[C_2(u,v),w]}{\partial (C_2(u,v))^2 \partial w} \cdot \frac{\partial C_2(u,v)}{\partial u} \cdot \frac{\partial C_2(u,v)}{\partial v} \\ & + \frac{\partial^2 C_1[C_2(u,v),w]}{\partial C_2(u,v) \partial w} \cdot \frac{\partial^2 C_2(u,v)}{\partial u \partial v}. \end{aligned} \tag{1.21}$$

Así, la densidad de las variables X_1, X_2 y X_3 expresada en cópulas asimétricas está dada por

$$\begin{aligned} & f_{X_1,X_2,X_3,\theta_1,\theta_2}(x_1,x_2,x_3) \\ =& f_{X_1}(x) f_{X_2}(x_2) f_{X_3}(x_3) \\ & \times c_{\theta_1,\theta_2}(F_{X_1}(x_1), F_{X_2}(x_2), F_{X_3}(x_3)). \end{aligned} \tag{1.22}$$

1.7.3. La cópula asimétrica de Gumbel - Hougaard en 3 variables

La cópula asimétrica de Gumbel-Houggard en tres dimensiones está representada por

$$\begin{aligned} & C_1(u_3, C_2(u_2,u_1)) \\ =& \exp\left(- \left(\left((-\log\ u_1)^{\theta_2} + (-\log\ u_2)^{\theta_2} \right)^{\frac{\theta_1}{\theta_2}} \right.\right. \\ & \left.\left. + (-\log\ u_3)^{\theta_1} \right)^{\frac{1}{\theta_1}} \right), \end{aligned}$$

donde $\theta_1, \theta_2 \in [1,\infty]$ y $\theta_1 < \theta_2$. Ahora, la derivada dada por (1.21) es necesaria para encontrar la función de densidad, que

para este caso,

$$\begin{aligned}
&\frac{\partial^3 C_1(u_3, C_2(u_2, u_1))}{\partial u_1 \partial u_2 \partial u_3} \\
=&\frac{1}{u_1 u_2 u_3}(-\log\ u_1)^{\theta_2-1}(-\log\ u_2)^{\theta_2-1}(-\log\ u_3)^{\theta_1-1} \\
&\times k_1^{\frac{1}{\theta_1}-3} k_2^{\frac{\theta_1}{\theta_2}-2} \exp\left(-k_1^{\frac{1}{\theta_1}}\right) \\
&\times \left\{ k_2^{\frac{\theta_1}{\theta_2}} \left[1 - 3k_1^{\frac{1}{\theta_1}} + k_1^{\frac{2}{\theta_1}}\right] + \theta_1^2 \left(k_2^{\frac{\theta_1}{\theta_2}} - (-\log\ u_3)^{\theta_1}\right) \right. \\
&+ \theta_1 \left(-1 + k_1^{\frac{1}{\theta_1}}\right) \left(2k_2^{\frac{\theta_1}{\theta_2}} - (-\log\ u_3)^{\theta_1}\right) \\
&\left. +\theta_1\theta_2 k_1 + \theta_2 k_1 \left(-1 + k_1^{\frac{1}{\theta_1}}\right) \right\},
\end{aligned}$$

donde

$$k_1 = \left((-\log\ u_1)^{\theta_2} + (-\log\ u_2)^{\theta_2}\right)^{\frac{\theta_1}{\theta_2}} + (-\log\ u_3)^{\theta_1},$$

y

$$k_2 = (-\log\ u_1)^{\theta_2} + (-\log\ u_2)^{\theta_2}.$$

Capítulo 2

Distribución de Valores Extremos

Las distribuciones de valores extremos son modelos límites para los máximos y mínimos de un conjunto de datos. Estas distribuciones modelan qué tan grandes (o pequeños) serán probablemente los datos. Estas distribuciones son útiles para modelar la dependencia entre variables aleatorias y se utilizan en diversos campos como las finanzas, la biología, las ciencias sociales y cualquier área en la cual se quiera analizar máximos o mínimos de un conjunto de datos en una determinada unidad de tiempo, como se indica en [10, 22, 29, 42].

2.1. Máximos por bloques

Se asume que los datos M_n son máximos, es decir,

$$M_n = \text{máx}\{X_1, \ldots, X_n\}, \tag{2.1}$$

donde las X_i con $i = 1, \ldots, n$ pueden ser no observables.

Si el experto puede observar las X_i en (2.1), entonces elegir los máximos de los bloques observados es otra posibilidad,

además de extraer valores extremos superiores de un conjunto de datos (o tomar excedencias). Este método se llama máximos por bloques o método de Gumbel.

Para variables aleatorias iid $X_1, \ldots, X_n$, M_n tiene función de distribución:

$$P\left[\max_{i\leq n} X_i \leq x\right] = P\left[X_1 \leq x, \ldots, X_n \leq x\right] = F^n(x).$$

Por tanto, las M_n en (2.1) están dependiendo de F^n si las x_i's tienen función de distribución F.

2.2. Teoremas de tipo extremo

El problema surge cuando F es desconocida, por lo que se buscan familias aproximadas de modelos para F^n, que pueden estimarse sobre la base de datos extremos. Esto es similar a la práctica habitual de aproximar la distribución de las medias muestrales mediante la distribución normal, tal como lo justifica el Teorema de Límite Central.

Se procede observando el comportamiento de F^n cuando $n \to \infty$. Pero esto por sí solo no es suficiente: para cualquier $x < x^*$, donde x^* es el punto final superior[1] de F, $F^n(z) \to \infty$ cuando $n \to \infty$, de modo que la distribución de M_n degenera en una masa puntual en x^*.

Se supone que existen $\{a_n > 0\}$ y $\{b_n\}$ sucesiones reales, tales que

$$M_n^* = \frac{M_n - b_n}{a_n},$$

[1] x^* es el más pequeño de los x tal que $F(x) = 1$, es decir,

$$x^* = \text{ínf}\{z : F(x) = 1\}.$$

tiene una distribución límite no degenerada[2] cuando $n \to \infty$, es decir,

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x).$$

Teorema 2.1. *Existen sucesiones de constantes* $\{a_n > 0\}$ *y* $\{b_n\}$ *reales tales que*

$$P\left(\frac{M_b - b_n}{a_n} \leq x\right) \to G(x) \quad \textit{cuando } n \to \infty,$$

para una función de distribución G *no degenerada, entonces* G *es un miembro de la familia de distribuciones de valores extremos generalizada (GEV por sus siglas en inglés),*

$$G_\xi = \exp\left(-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right),\ 1+\xi\left(\frac{x-\mu}{\sigma}\right) > 0, \tag{2.2}$$

donde ξ *es real y donde para* $\xi = 0$ *el lado derecho se puede interpretar como* $\exp(-e^{-x})$ *(caso límite).*

Demostración. Véase [22]. □

Definición 2.1. *El parámetro* ξ *es llamado el* ***índice de valor extremo****.*

El Teorema 2.1 demuestra que las funciones de distribución límite forman una sencilla familia de un solo parámetro explícito aparte de los parámetros de escala y ubicación. La Figura B.1 ilustra esta familia para algunos valores de ξ, con $\mu = 0$ y $\sigma = 1$.

Siguiendo con los valores estándar de $\mu = 0$ y $\sigma = 1$,

a) Para $\xi = 0$, el punto final superior de la distribución es infinito. La distribución, sin embargo, es bastante ligera: $1 - G_0(x) \sim e^x$ cuando $x \to \infty$.

[2]Una distribución degenerada es una distribución de probabilidad en un espacio donde el soporte está necesariamente en un espacio de dimensión más baja.

b) Para $\xi > 0$, $G_\xi(x) < 1$ para todo x, es decir, el punto final superior de la distribución es infinito. Además, cuando $x \to \infty$, $1 - G_\xi(x) \sim \xi^{-1/\xi} x^{-1/\xi}$, es decir, la distribución tiene una cola derecha bastante pesada.

c) Para $\xi < 0$, el punto final superior de la distribución es $-1/\xi$, por lo que tiene una cola corta, verificando $1 - G_\xi(-\xi^{-1} - 1) \sim (-\xi x)^{-1/\xi}$, cuando $x \to 0^+$.

La familia GEV se puede reparametrizar, esto con el fin de tener otras distribuciones más específicas, es decir,

a) El subconjunto de la familia GEV con $\xi = 0$, se obtiene la distribución
$$G_0(x) = \exp(-e^{-x}),$$
para todo x, y es llamada la distribución Gumbel.

b) Para $\xi > 0$, se usa $G_\xi\left(\frac{x-1}{\xi}\right)$ y se toma $\alpha = 1/\xi > 0$,
$$G_\alpha = \begin{cases} 0, & x \leq 0, \\ \exp(-x^{-\alpha}), & x > 0. \end{cases}$$
A esta familia de distribuciones se le conoce como la familia de distribuciones Fréchet.

c) Para $x_i < 0$ se usa $G_\xi\left(-\frac{1+x}{\xi}\right)$ y tomando $\alpha = -1/\xi > 0$,
$$G_\alpha = \begin{cases} \exp(-(x-)^{-\alpha}), & x < 0. \\ 1 & x \geq 0. \end{cases}$$
La cual es la familia de distribuciones Weibull.

Cabe señalar que las tres distribuciones anteriores se encuentran en su forma estándar, es decir, donde el parámetro de ubicación es μ es 0 y el parámetro de escala σ es 1. Además, al parámetro α es el llamado **parámetro de forma**.

Al Teorema 2.1 se puede reformular de acuerdo a Coles [10] de la manera siguiente.

Teorema 2.2. *Si la variable aleatoria X tiene función de distribución F, entonces $\mu+\sigma X$ tiene función de distribución y escala $F_{\mu,\sigma}(x) = F((x-\mu)/\sigma)$, donde μ y $\sigma > 0$ son los parámetros de ubicación y escala respectivamente. Así, la distribución de GEV es una de los siguientes:*

- *Gumbel:*

$$G_0(x) = \exp(-e^{-(x-\mu)/\sigma}),$$

  para toda $x \in \mathbb{R}$.

- *Fréchet, $\alpha > 0$:*

$$G_{1,\alpha}(x) = \exp\left(-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right),$$

  para $x \geq \mu$.

- *Weibull, $\alpha > 0$:*

$$G_{2,\alpha}(x) = \exp\left(-\left(-\left(\frac{x-\mu}{\sigma}\right)\right)^{\alpha}\right),$$

  para $x \leq \mu$.

El Teorema 2.2 muestra las distintas familias de distribuciones GEV que se obtienen, las cuales tienen propiedades y características particulares.

Nota 2.1. *En el Teorema anterior, $\mu+\sigma X$ es miembro de alguna de las familias mencionadas, sin embargo, al desplazamiento o multiplicación por una constante positiva será igualmente miembro de tales familias. Por tanto, la distribución de X es miembro de la familia Gumbell, Frechét o Weibull.*

2.3. Inferencia sobre la función de distribución de GEV

Una dificultad con el uso de métodos de verosimilitud para la distribución GEV, es que las condiciones de regularidad que se requieren para que las propiedades asintóticas asociadas habitualmente con el estimador de máxima verosimilitud sean válidas. La distribución GEV no cumplen estas condiciones porque los puntos finales de la distribución son funciones de los valores de los parámetros: $\mu - \sigma/\xi$ es un punto final superior de la distribución cuando $\xi < 0$, y un punto final inferior cuando $\xi > 0$. Smith en [50] estudió este problema en detalle y obtuvo los resultados siguientes:

- Cuando $\xi > -0.5$, los estimadores de máxima verosimilitud son regulares, en el sentido de tener las propiedades asintóticas habituales.
- Cuando $-1 < \xi < -0.5$, los estimadores de máxima verosimilitud generalmente se pueden obtener, pero no tienen las propiedades asintóticas estándar.
- Cuando $\xi < -1$, es poco probable que se puedan obtener estimadores de máxima verosimilitud.

El caso $\xi \leq -0.5$ corresponde a distribuciones con una cola superior acotada muy corta. Esta situación es rara en aplicaciones de modelos de valores extremos, por lo que las limitaciones teóricas del enfoque de máxima verosimilitud no suelen ser un obstáculo en la práctica.

2.3.1. Estimadores de máxima verosimilitud

Bajo el supuesto de que $X_1, \ldots, X_n$ son variables independientes que tienen la distribución GEV, la log-verosimilitud

de los parámetros de la distribución GEV cuando $\xi \neq 0$ es

$$l(\mu, \sigma, \xi) = -n \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n} \log \left(1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right) - \sum_{i=1}^{n} \left(1 + \left(\frac{x_i - \mu}{\sigma}\right)\right)^{-1/\xi}, \quad (2.3)$$

siempre que

$$1 + \xi \left(\frac{x_i - \mu}{\sigma}\right) > 0, \quad \text{para } i = 1, \ldots, m. \quad (2.4)$$

Para el caso donde $\xi = 0$, donde se tiene una distribución Gumbel da la log-verosimilitud

$$l(\mu, \sigma) = -n \log \sigma - \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma}\right) - \sum_{i=1}^{n} \exp \left(\frac{x_i - \mu}{\sigma}\right). \quad (2.5)$$

Maximizar el par de ecuaciones (2.3) y (2.5) con respecto al vector de parámetros (μ, σ, ξ) conduce a la estimación de máxima verosimilitud. No existe solución analítica, sin embargo, para cualquier conjunto de datos, la maximización es sencilla utilizando algoritmos de optimización numérica. Se necesita tener cuidado para asegurar que los algoritmos no pasen a combinaciones de parámetros que incumplan (2.4), y también eviten las dificultades numéricas que surgirían de la evaluación de (2.3) alrededor de $\xi = 0$.

2.3.2. Inferencia para niveles de retorno

Sea $X_1, X_2, \ldots$ una serie de observaciones independientes. Los datos se bloquean en secuencias de observaciones de longitud n, con n suficientemente grande, generando una serie de máximos de bloque, $M_1, \ldots, M_m$, a los que se puede ajustar

la distribución GEV. Las estimaciones de cuantiles extremos de la distribución máxima se obtienen invirtiendo la ecuación (2.2):

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi}\left(1 - (-\log(1-p))^{-\xi}\right), & \text{para } \xi \neq 0, \\ \mu - \sigma \log(-\log(1-p)), & \text{para } \xi = 0, \end{cases} \tag{2.6}$$

donde $G(z_p) = 1 - p$. El z_p es el nivel de retorno asociado con el borde del período de retorno, ya que con un grado razonable de precisión, se espera que el nivel z_p se exceda en promedio una vez cada $1/p$ unidades de tiempo, es decir, z_p es excedido por el máximo en el periodo de tiempo en cualquier año en particular con probabilidad p.

Por la sustitución de las estimaciones de máxima verosimilitud de los parámetros GEV en (2.6), la estimación de máxima verosimilitud de z_p para $0 < p < 1$, el nivel $1/p$ de retorno, se obtiene como

$$z_p = \begin{cases} \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}\left(1 - (-\log(1-p))^{-\hat{\xi}}\right), & \text{para } \hat{\xi} \neq 0, \\ \hat{\mu} - \hat{\sigma} \log(-\log(1-p)), & \text{para } \hat{\xi} = 0. \end{cases}$$

Para $\hat{\xi} < 0$, es posible hacer inferencias sobre z^* de la distribución, que es el período de retorno de observación infinita, correspondiente a z_p con $p = 0$. La estimación de máxima verosimilitud es

$$\hat{z}_0 = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}}.$$

Cuando $\hat{\xi} \geq 0$, la estimación de máxima verosimilitud de z^* es infinito.

2.4. Las distribuciones

La distribución $Fréchet(\alpha, s, \mu)$ está dada por

$$F(x; \alpha, s) = e^{-((x-\mu)/s)^{-\alpha}}, \quad x > \mu.$$

donde α es el parámetro de forma y $s > 0$ es el parámetro de escala. Su respectiva función de densidad está dada por

$$f(x;\alpha,s,\mu) = \frac{\alpha}{s}\left(\frac{x-\mu}{s}\right)^{-\alpha-1} e^{-((x-\mu)/s)^{-\alpha}}, \quad x > \mu.$$

Para la densidad Weibull presentada en el Teorema 2.2 se tiene que es para valores menores que μ, sin embargo, en aplicaciones se considera la distribución de su transformación negativa, es decir, si Y tiene distribución Weibull donde α es el parámetro de forma y k es el parámetro de escala, se considera la transformación $X = \varphi(Y) = -Y$, la cual es una transformación decreciente, por lo que

$$\begin{aligned} G_X(x) &= P(X \leq x) \\ &= P(\varphi(Y) \leq x) \\ &= P(Y \geq \varphi^{-1}(x)) \\ &= 1 - P(Y \leq \varphi^{-1}(x)) \\ &= 1 - \exp\left(-\left(-\left(\frac{-x-\mu}{\sigma}\right)\right)^{\alpha}\right) \\ &= 1 - \exp\left[-\left(\frac{x-(-\mu)}{k}\right)^{\alpha}\right]. \end{aligned}$$

Así, a partir de ahora, cuando se haga mención de la distribución Weibull, se hace referencia a la ecuación anterior, por lo que la distribución $Weibull(\alpha, k, \mu)$ está dada por[3]

$$F(x;\alpha,k) = 1 - e^{-((x-\mu)/k)^{\alpha}}, \quad x > \mu,$$

donde $\alpha > 0$ es el parámetro de forma y $k > 0$ es el parámetro de escala. La función de densidad asociada es

$$f(x;\alpha,k) = \frac{\alpha}{k}\left(\frac{x-\mu}{k}\right)^{\alpha-1} e^{-((x-\mu)/k)^{\alpha}}, \quad x > \mu,$$

[3]Como $-\mu$ es una constante, se puede reescribir a μ como $-\mu$

Una forma alternativa en la densidad $Weibull(\alpha, k, \mu)$ es tomar $b = k^{-\alpha}$,

$$\begin{aligned} f(x;\alpha,k) &= \frac{\alpha}{k}\left(\frac{x-\mu}{k}\right)^{\alpha-1} e^{-((x-\mu)/k)^{\alpha}} \\ &= \frac{\alpha}{k}\cdot\frac{(x-\mu)^{\alpha-1}}{k^{\alpha-1}} e^{-(x-\mu)^{\alpha}/k^{\alpha}} \\ &= \frac{\alpha}{k^{\alpha}}\cdot(x-\mu)^{\alpha-1} e^{-(x-\mu)^{\alpha}/k^{\alpha}} \\ &= b\alpha\cdot(x-\mu)^{\alpha-1} e^{-b(x-\mu)^{\alpha}}, \end{aligned}$$

para $x > \mu$, la función de distribución es

$$F(x;\alpha,k) = 1 - e^{-b(x-\mu)x^{\alpha}}.$$

Esta reparametrización es la que se encuentra implementada en el paquete OpenBUGS, la cual se denotará por $Weibull_2(\alpha, b, \mu)$.

Capítulo 3

Estadística Bayesiana

En la teoría de probabilidad, se toman a menudo a la Definición A.1 como los axiomas de la probabilidad. De estas propiedades, se establecen varias propiedades y Definiciones, y una de ellas es el Teorema de Bayes, cuya idea es esencial para la estadística bayesiana. La idea clave de la estadística bayesiana es que se puede utilizar la probabilidad para actualizar las "creencias".

3.1. Teorema de Bayes

A menudo, se necesita calcular la probabilidad de un evento A sabiendo que antes ha ocurrido B. A esta probabilidad se le denota por $P(A|B)$, llamada la **probabilidad condicional de A dado B**.

Definición 3.1 (Probabilidad condicional)**.** *Sean A y B eventos de un espacio muestral Ω y se supone que $P(B) > 0$. Así, la probabilidad condicional del evento A dado el evento B se define como*

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

La Definición de probabilidad condicional da una forma de calcular las probabilidades de intersecciones de conjuntos, es decir,

$$P(A \cap B) = P(B)P(A|B).$$

Aunque en la Definición de probabilidad condicional no da un argumento de porque toma esa forma, esté puede consultarse en [23], con un enfoque frecuencial de la probabilidad.

Si E es un evento en un espacio muestral Ω, es posible conocer $P(E)$ en términos de las probabilidades condicionales de una partición de Ω.

Definición 3.2 (Partición)**.** *Se dice que los eventos* A_1, $A_2, \ldots, A_n$ *de* Ω *es una partición de* Ω*, si*

a) $A_i \cap A_j = \emptyset$ *si* $i \neq j$ *(son ajenos dos a dos),*

b) $\Omega = \cup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n$.

Una vez determinada la Definición de partición, se puede llegar al llamado Teorema de la probabilidad total.

Teorema 3.1 (Teorema de la Probabilidad Total)**.** *Si* E *es un evento en* Ω *y* $A_1, A_2, \ldots, A_n$ *una partición de* Ω*, entonces*

$$P(E) = \sum_{i=1}^{n} P(A_i)P(E|A_i).$$

Demostración. Ver [23]. □

Como consecuencia del Teorema de la probabilidad total y la Definición de probabilidad condicional, se obtiene el Teorema base de la estadística bayesiana, el Teorema de Bayes.

Teorema 3.2. *Sea $A_1, A_2, \ldots, A_n$ una partición sobre el espacio muestral Ω y sea E un suceso en Ω, entonces:*

$$P(A_i|E) = \frac{P(A_i)P(E|A_i)}{\sum_{j=1}^{n} P(A_j)P(E|A_j)}, \quad \textit{para } i = 1, 2, \ldots, n. \quad (3.1)$$

Demostración. Ver [23]. □

A la ecuación (3.1) es la llamada **fórmula de Bayes**.

3.2. La idea de la estadística Bayesiana

Sean X e Y variables aleatorias con espacios muestrales Ω_X y Ω_Y respectivamente. Sea $A_1, A_2, \ldots, A_n$ es una partición de Ω_Y, y E un evento de Ω_x. Pensando en la probabilidad de $P(Y \in A_i|X \in E)$, al usar el Teorema de Bayes,

$$P(Y \in A_i|X \in E) = \frac{P(Y \in A_i)P(X \in E|Y \in A_i)}{\sum_{j=1}^{n} P(Y \in A_j)P(X \in E|Y \in A_j)},$$

así, ignorando el denominador de la parte derecha[1], se llega a la expresión siguiente,

$$P(Y \in A_i|X \in E) \propto P(Y \in A_i)P(X \in E|Y \in A_i).^2$$

Condicionar la variable aleatoria Y por el valor de X no cambia los tamaños relativos de las probabilidades de esos

[1] A tal fracción se le conoce por lo regular como "constante de proporcionalidad".

[2] El símbolo $\propto$ se puede interpretar como "proporcional a..."

pares (x, y) que aún pueden ocurrir. Es decir, la probabilidad $P(Y|X)$ es proporcional a $P(X, Y)$ y la constante de proporcionalidad es justo lo que se necesita.

El enfoque que toma la estadística bayesiana, es el de poder obtener funciones de densidades para los parámetros desconocidos, esto con ayuda de información basada en observaciones muestrales sobre las variables de interés y ayudándose de información teórica de la distribución del parámetro desconocido.

En la literatura de la metodología bayesiana se usan letras minúsculas (o griegas). Así, de forma general

$$P(y|x) = \frac{P(x,y)}{P(x)} = \frac{P(y)P(x|y)}{P(x)},$$

por lo que,

$$P(y|x) \propto P(y)P(x|y). \tag{3.2}$$

A (3.2) se le conoce como el Teorema de Bayes para variables aleatorias. A las probabilidades $P(y)$, $P(x|y)$, $P(y|x)$ y $P(x)$ se les conoce como la apriori, la verosimilitud, la posteriori y la predictiva, respectivamente. Las variables x e y pueden ser continuas o discretas. En el caso continuo, la constante de proporcionalidad es

$$\frac{1}{P(x)} = \frac{1}{\int P(y)P(x|y)dy},$$

y en el caso discreto

$$\frac{1}{P(x)} = \frac{1}{\sum_y P(y)P(x|y)}.$$

3.3. Ejemplo sobre la estadística Bayesiana

Sea y el tiempo antes de la primera aparición de una desintegración radiactiva que se mide con un instrumento, pero que, debido a que hay un retraso incorporado en el mecanismo, la desintegración se registra como si hubiera tenido lugar en un tiempo $x > y$. Se tiene el valor de x, pero se quiere decir algo sobre el valor de y, por lo que

$$\begin{aligned} P(y) &= e^{-y} && (0 < y < \infty), \\ P(x|y) &= ke^{-k(x-y)} && (y < x < \infty). \end{aligned}$$

Así,

$$\begin{aligned} P(y|x) &\propto P(y)P(x,y) \\ &\propto e^{(k-1)y} \qquad (0 < y < x). \end{aligned}$$

A menudo basta con obtener un resultado hasta una constante de proporcionalidad. Si se necesita la constante, es sencillo obtenerla, ya que basta con integrar sobre el soporte de la variable de interés. Por lo tanto, en este caso

$$P(y|x) = \frac{(k-1)e^{(k-1)y}}{e^{(k-1)x} - 1} \qquad (0 < y < x).$$

Capítulo 4

Modelos Bayesianos Complejos: Método Monte Carlo de la Cadena de Markov

En la metodología bayesiana que utiliza familia conjugadas[1] es posible realizar una inferencia posteriori exacta. En modelos bayesianos que representan la realidad suelen ser más complejos encontrándose muchos problemas en su implementación a pesar de que existe la teoría. El lograr modelar toda la información es un reto que se aprende en cada situación particular. Se tiene que identificar qué variables usar, identificar sus interacciones, las distribuciones más adecuadas a usar, los posibles cálculos analíticos que sean factibles para lograr el objetivo, entre otros.

Debido a que se está usando el enfoque Bayesiano, siem-

[1]En estadística bayesiana, si las distribuciones posteriori $p(\theta|x)$ y apriori $p(\theta)$ pertenecen a la misma familia de distribuciones de probabilidad, entonces a las distribuciones apriori y posteriori se les denomina distribuciones conjugadas [54].

pre que se pueda escribir la verosimilitud y la(s) apriori(s) en forma matemática, se obtiene una expresión proporcional a la distribución posteriori resultante. Sin embargo, para trazar una densidad, se necesita la constante de normalización, de modo que el área bajo nuestra gráfica de densidad sea 1. Cuando la apriori es no conjugada y la densidad posteriori no es de una familia reconocible, la constante normalizadora debe obtenerse por integración: se debe averiguar a qué valor numérico integra la densidad no normalizada, y luego la constante normalizadora es su inversa.

4.1. Cadenas de Markov

Las cadenas de Markov son variables aleatorias que se generan secuencialmente a lo largo del tiempo. Se dice que una cadena de Markov comienza en el "tiempo 0" con algún valor inicial. En el tiempo 1, la cadena se mueve a un valor aleatorio generado a partir de una distribución de probabilidad cuyos parámetros dependen del valor inicial desde el tiempo 0. En cada punto de tiempo sucesivo, la cadena vuelve a moverse a un nuevo valor aleatorio generado a partir de la misma forma de distribución de probabilidad, pero con parámetros que dependen del valor del punto de tiempo inmediatamente anterior.

La notación común para una cadena de Markov es $\{X_t\}_{t=0}^{\infty}$, donde X_t representa la variable aleatoria en el tiempo t, y una vez alcanzado el tiempo t, x_t denota el valor realizado. El valor x_t se denomina estado de la cadena en el tiempo t. Los puntos de tiempo t en los que una cadena de Markov genera nuevos valores a menudo se denominan iteraciones y los valores generados x_t como iterados. Los valores de una cadena de Markov pueden ser escalares o vectores.

El soporte del que se extraen todas las variables aleatorias

X_t se denomina espacio de estados de la cadena de Markov. La distribución de probabilidad a partir de la cual se extrae el estado en cada momento t, condicionada por el estado del momento anterior, se denomina núcleo de transición de la cadena y se denota por $P(X_t|X_{t-1} = x_{t-1})$.

La característica definitoria de las cadenas de Markov es la propiedad de Markov, que es

$$\begin{aligned} &P(X_t|X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \ldots, X_0 = x_0) \\ =&P(X_t|X_{t-1} = x_{t-1}). \end{aligned}$$

Bajo condiciones de regularidad, las extracciones generadas por una cadena de Markov convergerán en distribución para extraer una distribución de probabilidad objetivo. Así, si se permite que una cadena de Markov trabaje durante el tiempo suficiente, entonces se produce la convergencia y todas las iteraciones posteriori se extraen de esta distribución objetivo.

Incluso después de que una cadena de Markov ha convergido, las iteraciones posteriori siguen siendo dependientes. Esto es porque el núcleo de transición no cambia: cada nuevo valor se genera a partir de una distribución de probabilidad que depende de los valores de la iteración anterior.

4.2. Cadenas de Markov para inferencia bayesiana

Las cadenas de Markov son importantes en la estadística bayesiana porque generalmente es posible construir una cadena de Markov (es decir, definir su núcleo de transición) de tal manera que la distribución objetivo sea la distribución posteriori conjunta de todos los parámetros desconocidos en el modelo bayesiano de interés. Incluso para modelos de dimen-

sión alta en los que no es factible extraer muestras directamente de la posteriori conjunta, a menudo es sencillo definir un núcleo de transición. Por lo tanto, los métodos de Monte Carlo de la cadena de Markov (MCMC por sus siglas en inglés) proporcionan una forma de extraer muestras de la distribución conjunta posteriori en modelos bayesianos realistas de alta dimensión.

El algoritmo de muestreo de Gibbs es una forma particular de construir un núcleo de transición para producir una cadena de Markov con la distribución objetivo deseada. Es el método descrito por Gelfand y Smith en [16], y en él se basan los algoritmos utilizados en WinBUGS y OpenBUGS.

4.3. El algoritmo EM

El algoritmo EM (Expectación Máxima) es una técnica numérica que encuentra el valor en el que $P(\eta|\mathbf{x})$ es un máximo, o equivalentemente, donde $\log P(\eta|\mathbf{x})$ es un máximo, pero donde no se tiene información completa sobre la distribución posteriori.

Un ejemplo del algoritmo EM es el enlace genético desarrollado por Rao [41]. Se tienen las observaciones $\mathbf{x} = (x_1, x_2, x_3, x_4)$ con probabilidades de celda

$$\left(\frac{1}{2} + \frac{1}{4}\eta, \frac{1}{4}(1-\eta), \frac{1}{4}(1-\eta), \frac{1}{4}\eta\right),$$

y se quiere estimar η. La verosimilitud es entonces

$$\left(\frac{1}{2} + \frac{1}{4}\eta\right)^{x_1} \left(\frac{1}{4}(1-\eta)\right)^{x_2} \left(\frac{1}{4}(1-\eta)\right)^{x_3} \left(\frac{1}{4}\eta\right)^{x_4} \propto (2+\eta)^{x_1}(1-\eta)^{x_2+x_3}\eta^{x_4}.$$

La estrategia consiste en aumentar los datos $\mathbf{x}$ agregando más datos de z para producir datos aumentados $\mathbf{y}$. Cabe señalar

que, en gran medida, la distinción entre los parámetros de un modelo y los aumentos de los datos es artificial. En el ejemplo en cuestión, el aumento consiste simplemente en dividir la primera celda en dos celdas con probabilidades $\frac{1}{2}$ y $\frac{1}{4}\eta$. La ventaja de esto es que la expresión obtenida es más sencilla.

$$\left(\frac{1}{2}\right)^{y_0}\left(\frac{1}{4}\eta\right)^{y_1}\left(\frac{1}{4}(1-\eta)\right)^{y_2}\left(\frac{1}{4}(1-\eta)\right)^{y_3}\left(\frac{1}{4}\eta\right)^{y_4} \propto \eta^{y_1+y_4}(1-\eta)^{y_2+y_3},$$

y si se asigna la referencia estándar apriori de $Be(0,0)$, entonces la posteriori tiene una distribución beta

$$P(\eta|\mathbf{y}) \propto \eta^{y_1+y_4-1}(1-\eta)^{y_2+y_3-1}.$$

El algoritmo EM para encontrar la distribución posteriori, es un método iterativo que inicia de alguna conjetura plausible $\eta^{(0)}$ para el valor de η. En la etapa $t > 0$, se supone que la estimación actual es $\eta^{(t)}$. Cada etapa tiene dos E-pasos. En el primero, el E-paso (paso de expectativa), se calcula

$$Q(\eta, \eta^{(t)}) = E_{\eta^{(t)}} \log P(\eta|\mathbf{y}), \qquad t > 0,$$

es decir, la expectativa de la función de log-verosimilitud, calculando la expectativa en $\eta = \eta^{(t)}$, de modo que

$$Q(\eta, \eta^{(t)}) = \int \log(P(\eta|\mathbf{y}))P(\mathbf{y}|\eta^{(t)}, \mathbf{x})d\mathbf{y}.$$

En el segundo paso, el M-paso (el paso de maximización), se encuentra ese valor $\eta^{(t+1)}$ de η que maximiza $Q(\eta, \eta^{(t)})$. En este ejemplo en particular, como $y_i = x_i$ para $i > 1$

$$\begin{aligned} & Q(\eta, \eta^{(t)}) \\ =& E((y_1 + y_4 - 1)\log\eta + (y_2 + y_3 - 1)\log(1-\eta)|\eta^{(t)}, \mathbf{x}) \\ =& (E(y_1|\eta^{(t)}, \mathbf{x}) + x_4 + 1)\log\eta + (x_2 + x_3 - 1)\log(1-\eta). \end{aligned}$$

Para el M-paso, se nota que

$$\frac{\partial Q(\eta, \eta^{(t)})}{\partial \eta} = \frac{E(y_1|\eta^{(t)}, \mathbf{x}) + x_4 - 1}{\eta} - \frac{x_2 + x_3 - 1}{1 - \eta},$$

e igualando a cero,

$$\eta^{(t+1)} = \frac{E(y_1|\eta^{(t)}, \mathbf{x}) + x_4 - 1}{E(y_1|\eta^{(t)}, \mathbf{x}) + x_2 + x_3 + x_4 - 2}.$$

Dado que y_1 tiene una distribución binomial $B(x_1, n)$ con

$$\pi = \frac{\frac{1}{4}\eta^{(t)}}{\frac{1}{2} + \frac{1}{4}\eta^{(t)}} = \frac{\eta^{(t)}}{\eta^{(t)} + 2},$$

así que

$$E(y_1|\eta^{(t)}, \mathbf{y}) = x_1\eta^{(t)}/(\eta^{(t)} + 2)$$

la iteración es dada por

$$\begin{aligned}\eta^{(t+1)} &= \frac{x_1\eta^{(t)}/(\eta^{(t)} + 2) + x_4 - 1}{x_1\eta^{(t)}/(\eta^{(t)} + 2) + x_2 + x_3 + x_4 - 2} \\ &= \frac{\eta^{(t)}(x_1 + x_4 - 1) + 2(x_4 - 1)}{\eta^{(t)}(x_1 + x_2 + x_3 + x_4 - 2) + 2(x_2 + x_3 + x_4 - 2)}.\end{aligned}$$

Los valores realmente observados fueron $x_1 = 125$, $x_2 = 18$, $x_3 = 20$, $x_4 = 34$. Entonces se estima η por iteración comenzando, por ejemplo, $\eta^{(0)} = 0.5$ usando

$$\eta^{(t+1)} = \frac{158\eta^{(t)} + 66}{195\eta^{(t)} + 140}.$$

De hecho, la iteración convergerá a la raíz positiva de $195\eta^2 - 18\eta - 66 = 0$, que es 0.630.

Una demostración de la convergencia del algoritmo EM se encuentra en [30].

4.4. El muestreador de Gibbs

4.4.1. Aumento de datos encadenados

Ahora se restringirá la atención donde los datos aumentados $\mathbf{y}$ consisten en los datos originales $\mathbf{x}$ aumentados por un solo escalar z. Entonces, el algoritmo se puede expresar de la manera siguiente: comenzar con un valor $\eta^{(0)}$ generado a partir de la distribución apriori para η y luego repita de la manera siguiente:

(a_1) Elegir $\eta^{(i+1)}$ de η a partir de la densidad $P(\eta|z^{(i)}, \mathbf{x})$.

(a_2) Elegir $z^{(i+1)}$ de z a partir de la densidad $P(z|\eta^{(i+1)}, \mathbf{x})$.

Existe una simetría entre η y z, y la notación se usa simplemente porque surgió en relación con el primer ejemplo que se consideró en el algoritmo de aumento de datos. Esta versión del algoritmo se denomina como "aumento de datos encadenados", ya que es fácil ver que la distribución del par de valores siguiente (η, z) dados los valores que hasta ahora solo dependen del par actual y, por lo tanto, estos pares se mueven como una cadena de Markov. Es un caso particular de un método numérico al que se refiere como muestreador de Gibbs. Como resultado de las propiedades de las cadenas de Markov, después de un número razonablemente grande de iteraciones T, los valores resultantes de η y z tienen una densidad conjunta cercana a $P(\eta, z|\mathbf{x})$, independientemente de cómo comenzó la cadena.

Las observaciones sucesivas del par (η, z) no serán, en general, independientes, por lo que para obtener un conjunto de observaciones iid, se puede ejecutar el mencionado proceso a través de las T iteraciones sucesivas, reteniendo solo el valor final, sobre las diferentes repeticiones.

Se ilustrará el procedimiento con el ejemplo siguiente de

Casella y George [7]. Se supone que π e y tienen la distribución conjunta siguiente:

$$P(y,\pi) = \binom{n}{y}\pi^{y+\alpha-1}(1-\pi)^{n-y+\beta-1}, \qquad n>0,$$

para $x = 0, 1, \ldots, n$ y $0 \leq y \leq 1$, y que se está interesado en la distribución marginal de y. En lugar de integrar con respecto a π, que mostrará que y tiene una distribución beta-binomial, se procede a encontrar la distribución requerida a partir de las dos distribuciones condicionales:

$$\begin{aligned} &y|\pi \sim B(n,\pi), \\ &\pi|y \sim Be(y+\alpha, n-y+\beta), \qquad \alpha,\beta>0. \end{aligned}$$

Este es un caso simple en el que no hay datos observados $\mathbf{x}$. Se necesita inicializar el proceso en algún lugar, por lo que también se puede comenzar con un valor de π que se elige de una distribución $U(0,1)$. Una implementación en R con $T = 10$ iteraciones replicadas $m = 500$ veces para el caso $n = 16$, $\alpha = 2$ y $\beta = 4$, como señalan Casella y George en [45], dará una muy buena aproximación a la densidad binomial beta.

4.4.2. Un ejemplo con datos observados

En la Tabla B.1 se proporciona un conjunto pequeño de datos que representan fallas de bombas en varios sistemas de la planta nuclear Farley 1 [15] . La aparente variación en las tasas de falla tiene varias fuentes, como se observa en la Tabla B.1. Debería parecer plausible que el número de fallas en cualquier intervalo de tiempo fijo tenga una distribución de Poisson y que la media de esta distribución sea proporcional a la longitud del intervalo, con una constante de proporcionalidad que varíe de una bomba a otra. Parece sensato suponer que estas constantes provienen de la familia conjugada, los

múltiplos de chi-cuadrado, de modo que

$$y_i|\theta_i \sim P(\theta_i t_i), \qquad \theta_i \sim S_0^{-1}\chi_v^2.$$

Se considera que $v = 1.4$. Buscando las distribuciones marginales $P(\theta_i|\mathbf{y})$ se encuentra que no tienen una forma cerrada. Se necesita escribir

$$\begin{aligned}\boldsymbol{\theta} &= (\theta_1, \theta_2, \ldots, \theta_k),\\ \boldsymbol{\theta}_{-i} &= (\theta_1, \ldots, \theta_{i-1}, \theta_{i+1}, \ldots, \theta_k).\end{aligned}$$

Ahora

$$\begin{aligned}P(\boldsymbol{\theta}, \mathbf{y}, S_0) &= \prod P(y_i, \theta_i|S_0)P(S_0)\\ &= \prod P(y_i|\theta_i)P(\theta_i|S_0)P(S_0),\end{aligned}$$

de lo que se sigue

$$P(\theta_i|\boldsymbol{\theta}_{-i}, \mathbf{y}, S_0) \propto \theta_i^{(v+2y_i)/2-1} \exp\left(-\frac{1}{2}\theta_i(S_0 + 2t_i)\right).$$

Entonces se desprende del Apéndice A.2 en [30] que

$$\theta_i|\boldsymbol{\theta}_{-i}, S_0, \mathbf{y} \sim S_1^{-1}\chi_{v'}^2,$$

donde

$$S_1 = S_0 + 2t_i, \qquad v' = v + 2y_i.$$

Se encuentra que si se toma una apriori $U_0^{-1}\chi_\rho^2$ para S_0, entonces

$$P(S_0|\boldsymbol{\theta}, \mathbf{y}) \propto S_0^{(\rho+kv)/2} \exp\left(-\frac{1}{2}\left(U_0 + \sum \theta_i\right) S_0\right).$$

Esta es una distribución chi-cuadrado, de modo que

$$S_0|\boldsymbol{\theta}, \mathbf{y} \sim U_1^{-1}\chi_{\rho'}^2$$

con

$$U_1 = U_0 + \sum \theta_i, \qquad \rho' = \rho + kv.$$

De acuerdo con el principio general del muestreador de Gibbs, ahora se toma un valor de S_0 y después se generan valores de θ_i, luego se usan esos valores para generar un valor de S_0, posteriormente se usa este valor nuevo para generar nuevos valores de θ_i, etcétera.

Los resultados de $N = 10,000$ iteraciones (de los cuales se ignoran los primeros 1000), se muestran en la Tabla B.2. Resulta que S_0 tiene una media de 1.849713 y $s = 0.7906609$.

4.4.3. El muestreador de Gibbs como una extensión del aumento de datos encadenado

Con el algoritmo de aumento de datos encadenados, se tienen dos etapas en las que se estima alternativamente dos parámetros η y z. El muestreador de Gibbs puede considerarse como una extensión multivariante del algoritmo de aumento de datos encadenados en el se estiman r parámetros $\theta_1, \theta_2, \ldots, \theta_r$. Para usarlo, se toma un punto de partida $(\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_r^{(0)})$ y luego iterar de la manera siguiente:

(a_1) Elegir $\theta_1^{(t+1)}$ de θ_1 a partir de la densidad $P(\theta_1|\theta_2^{(t)}, \theta_3^{(t)}, \ldots, \theta_r^{(t)}, \mathbf{x})$.

(a_2) Elegir $\theta_2^{(t+1)}$ de θ_2 a partir de la densidad $P(\theta_2|\theta_1^{(t+1)}, \theta_3^{(t)}, \ldots, \theta_r^{(t)}, \mathbf{x})$.

$\vdots$

(a_r) Elegir $\theta_r^{(t+1)}$ de θ_r a partir de la densidad $P(\theta_r|\theta_1^{(t+1)}, \theta_2^{(t+1)}, \ldots, \theta_{r-1}^{(t+1)}, \mathbf{x})$.

Entonces los valores de $\boldsymbol{\theta} = (\theta_1^{(t)}, \ldots, \theta_r^{(t)})$ se mueven como una cadena de Markov, de modo que una vez que llegan a

una posición particular, al pasar a la iteración siguiente, el procedimiento no depende del estado anterior. En muchos casos esto significará que los valores tomados después de un gran número de iteraciones deben tener una distribución que no depende de su distribución inicial. Hay que tener en cuenta las circunstancias en las que se produce la convergencia a una distribución de equilibrio y, suponiendo que ocurra, la velocidad a la que se alcanza el equilibrio, lo cual se puede consultar en [30] o [45].

4.5. El diagnóstico de Brooks Gelman y Rubin

El paquete OpenBUGS incluye una de las variantes del estadístico Brooks y Gelman [6], es un popular diagnóstico de convergencia propuesto por primera vez por Gelman y Rubin [17]. Este algoritmo se puede usar para decidir cuántas iteraciones de calentamiento usar[2].

La idea intuitiva detrás del diagnóstico es que, si dos o más cadenas de MCMC se han ejecutado desde valores iniciales sobredispersos, se puede evaluar si las cadenas han escapado de sus valores iniciales y han encontrado la distribución objetivo de la cadena de Markov comparando la variabilidad dentro de cada cadena con la variabilidad de las muestras combinadas de todas las cadenas. Una vez que todas las cadenas han convergido (al menos aproximadamente) a la distribución objetivo, la variabilidad dentro de las cadenas debe ser aproximadamente igual a la variabilidad entre cadenas. Antes de la convergencia, mientras que cada cadena se extrae de una

[2]Las iteraciones de calentamiento (o quemado) se usan para dar tiempo que las itereciones presenten cierto comportamiento esperado, en este caso, se usan para observar la convergencia, es decir, se descartan un número de iteraciones iniciales para solo utilizar las que han convergido a la distribución objetivo.

parte diferente del espacio de parámetros, es probable que la variabilidad dentro de la cadena sea menor que la variabilidad de la muestra agrupada. Brooks y Gelman [6] sugirieron varias medidas diferentes de variabilidad que podrían usarse. Este algoritmo implementado en OpenBUGS son los anchos de los conjuntos creíbles del 80 % estimados a partir de las muestras. El diagnóstico numérico, llamado R por relación, es el ancho del conjunto creíble de muestra agrupada dividido por el ancho medio de los conjuntos creíbles dentro de la cadena.

En OpenBUGS se traza el gráfico "Diagnóstico BGR" (BGR diagnostic), siendo la línea verde la evolución de la variabilidad de la muestra agrupada (estadística B), la línea azul la evolución de la variabilidad promedio de las cadenas (estadístico W) y la línea roja la evolución del estadístico $R = B/W$. Hay que tomar en cuenta que B y W están normalizados para que el ancho de intervalo máximo estimado sea uno[3].

Para observar la convergencia en el gráfico, la línea verde y azul se deben estabilizar, lo que hará que la línea roja también se estabilice. De ejemplo se tiene la Figura B.2, donde se observa la convergencia alrededor de las 20,000 iteraciones, las mismas que servirán de calentamiento.

El "truco de ceros" (zeros trick) descrito en [31] se utiliza para codificar esta función de probabilidad en OpenBUGS. Este truco consiste en asignar una distribución de Poisson a un vector de ceros, con parámetro (media de Poisson) igual al negativo del logaritmo de la función de verosimilitud.

[3]La relación $R = B/W$ de los anchos de intervalo agrupados y promedio debe ser mayor que 1, si los valores iniciales están adecuadamente sobredispersados; también tenderá a 1 a medida que se acerque a la convergencia, por lo que en varias ocasiones se supone la convergencia para propósitos prácticos si $R < 1.05$.

Capítulo 5

Análisis de la Contaminación en la ZMVM

La contaminación es un problema ambiental que enfrenta la sociedad actualmente. Esta se produce cuando se introducen sustancias o gases dañinos al ecosistema, causando un desequilibrio. Este fenómeno tiene efectos perjudiciales en recursos naturales, la salud, las actividades humanas y la calidad del agua [28].

En particular, la contaminación del aire se refiere a la presencia en el aire de compuestos que implican riesgo, daño o molestia grave para las personas. Puede ser causada por diversas fuentes, por ejemplo, los gases producidos por vehículos, las emisiones de las fábricas, el polvo, el polen y las esporas de moho. Los efectos de la contaminación del aire pueden variar desde problemas de salud como enfermedades respiratorias y cardiovasculares hasta el calentamiento global [2].

5.1. Los contaminantes y datos

Los eventos de niveles altos de contaminación atmosférica pueden ocurrir de forma que afecten uno o mas contaminantes. Por ejemplo, se puede tener altos niveles de ozono (O_3) y también de monóxido de carbono (CO), por lo que el O_3 es precursor de los altos niveles en otros contaminantes [27]. De esta forma, es interesante estudiar conjuntamente varios contaminantes para saber si altos niveles de O_3 tienen influencia en los altos niveles de dióxido de nitrógeno (NO_2), dióxido de azufre (SO_2), partículas menores a 10 micrómetros (PM_{10}), partículas menores a 2.5 micrómetros ($PM_{2.5}$) y CO. Esto puede ser realizado por medio de las distribuciones conjuntas de las mediciones de estos contaminantes. Sin embargo, en muchos casos obtener una estas distribuciones puede no ser una tarea directa. Una forma de resolver esto es a través de cópulas que permiten obtener estas distribuciones conjuntas, al mismo tiempo que se preserva las propiedades de las distribuciones marginales particulares de cada contaminante, así la información relacionada de la asociación que pueda existir entre las mediciones de los contaminantes de interés.

El paquete OpenBUGS permite usar las técnicas descritas en el Capítulo 4, por lo que simular un gran número de valores para los parámetros de ciertas distribuciones es una tarea factible, con el fin de hacer estimaciones y pronósticos de los comportamientos de las variables de interés. Para ello se considera lo siguiente:

1. La ZMVM se divide de acuerdo a las zonas dadas por la RAMA en las regiones siguientes: Noroeste (NO), Noreste (NE), Centro (CE), Suroeste (SO) y Sureste (SE) [40].

2. Se dividen las estaciones que analizan los componentes en las regiones de la ZMVM. Los contaminantes analizados son el O_3, NO_2, SO_2, PM_{10}, $PM_{2.5}$ y CO.

3. Se obtienen los máximos mensuales de los contaminantes de las estaciones divididas en sus regiones correspondientes, y se analizan las posibles distribuciones y estimaciones de sus parámetros con ayuda de OpenBUGS.

4. Se ajustan las posibles densidades que podrían tener los máximos mensuales descritas en el Capítulo 2. Dado que los máximos mensuales obtenidos son estrictamente mayores que cero, se piensa que los valores deben tener una densidad Fréchet, además de considerar la densidad Weibull como se hace en otras aplicaciones. Por tal motivo se considera el parámetro de ubicación igual a cero.

5. Se hace un análisis conjunto del O_3 con los demás contaminantes con ayuda de OpenBUGS. Esto se hace acuerdo a la cópula de Gumbel-Hougaard que se observa en el Capítulo 1.

5.1.1. Estadísticas de los máximos mensuales

Los datos se consideran a partir de enero de 1990 (año en el que se consideran las bases confiables) a diciembre de 2021, considerando la fecha de inicio de registro de los contaminantes, es decir, para el CO, NO_2, O_3 y SO_2 a partir de enero de 1990; para las PM_{10} a partir de enero de 1995 y a las $PM_{2.5}$ a partir de agosto de 2003.

Los datos están en partes por billón (ppb) para los casos de O_3, NO_2 y SO_2; en microgramos/metro cúbico ($\mu g/m^3$) para PM_{10} y $PM_{2.5}$; y en partes por millón (ppm) para CO.

En la Tabla B.3 se tienen las estadísticas de los máximos mensuales de los contaminante divididos por las zonas de la ZMVM.

Se observa que las zona con las medias más altas son la NO seguida de la NE, lo que tiene sentido, ya que en estas zonas son las que tienen una mayor concentración de fabricas industriales. Las zonas con menor contaminación son las SE y SO, siendo la zona SO la más afectada por el contaminante O_3.

5.2. Análisis univariado de los contaminantes

Primero, se busca determinar cuál modelo probabilista se ajusta mejor a los máximos mensuales de cada contaminante. En OpenBUGS, se calculan los parámetros de forma y ubicación. Los modelos se expresan de la manera siguiente:

$$\begin{aligned} y &\sim Fréchet(\alpha_y, s_y), \\ \alpha_y &\sim U(0,10), \\ s_y &\sim U(0,a). \end{aligned} \tag{5.1}$$

$$\begin{aligned} y &\sim Weibull_2(\alpha_y, b_y), \\ \alpha_y &\sim U(0,10), \\ b_y &\sim U(0,10), \\ k_y &= b^{-1/\alpha_y}. \end{aligned} \tag{5.2}$$

Se recuerda que se quiere hacer inferencia sobre la densidad de los parámetros de forma (α) y los parámetros de escala (s y k para la Fréchet y Weibull respectivamente), por lo que en el caso de suponer caso Weibull, se utiliza la forma $Weibull_2$, debido a que está implementada en OpenBUGS, solo se calcula para cada generación de b y α con su respectivo k. Por otro lado, se usa el subíndice z para indicar que se trata del O_3, y para los otros contaminantes considerados usa el subíndice y.

Como el parámetro de escala para una Fréchet está relacionado con la varianza de la densidad, el parámetro s aumenta si

la varianza lo hace, por lo que para cada caso, el valor cambia de acuerdo a los datos, tomando los valores que se muestran en la Tabla B.4.

Las estimaciones de los parámetros de interés se hace de la manera siguiente:

1. Se crean $100,000$ simulaciones por cadena. Se espera tener 5 cadenas, pero en caso de que las simulaciones tarden en converger, se tomarán 3 cadenas.

2. Se analizan los gráficos BGR para observar la convergencia para decidir el número de iteraciones de calentamiento. Como mínimo se toman $20,000$ iteraciones de calentamiento.

3. La muestra final se toma cada 50 elementos, para tener una muestra aleatoria y evitar correlación entre los datos generados[1].

4. Para saber si la muestra es adecuada, se usa la regla empírica de que el error Monte Carlo (error M. C.) debe ser al menos $1/20$ veces más pequeño que la desviación estándar (d. e.) en cada parámetro del cual se quiere hacer inferencia. Si esto no se cumple se hacen otras $100,000$ iteraciones y se regresa al paso 2.

Para ejemplificar como se analizan los datos, se observa primero como se hace esto para las Región CE y el contaminante O_3. Se supone una densidad Fréchet en el comportamiento del contaminante, como se muestra en la ecuación (5.1). En la Figura B.2 se tienen los gráficos BGR, donde se observa que en general se obtuvo la convergencia de todos los parámetros a partir de las $20,000$ iteraciones, mismas que se

[1]Esto no es del todo necesario, ya que, por las propiedades de Ergodicidad, la media de la muestra converge al valor real de los parámetros. Para más detalles ver [20].

usaran como iteraciones de calentamiento, por lo que se puede hacer inferencia de los parámetros.

Para tener una muestra aleatoria, los elementos de los datos generados se eligen cada 50 elementos después de las iteraciones de calentamiento, y se procede a observar las densidades estimadas de cada parámetro. La Figura B.3 muestra las densidades de los parámetros elegidos, donde todos tienen la forma de una densidad unimodal, lo cual indica una buena muestra.

Esta sería la forma en la que se hace en todas las regiones, contaminantes y modelos mencionados. A continuación solo se mostraran los resultados estadísticos. Si se requiere consultar los gráficos generados en OpenBUGS y códigos se puede consultar en [52].

Con ello se puede estimar los parámetros de ubicación y escala en los modelos (5.1) y (5.2), y para todos los contaminantes y todas las zonas, se tienen las estimaciones de la Tabla B.5.

5.2.1. Elección del mejor modelo

Obteniendo una vez las estimaciones, es necesario elegir que modelo se ajusta mejor, para ello se piensan en dos criterios. Uno de los métodos que se usa es el DIC (Deviance Information Criterion), prefiriendo el que tenga un DIC más pequeño [11]. Para los objetivos establecidos, para comparar el DIC es necesario que los datos estén en la misma escala, recordando que las estimaciones se hicieron en ppb para los casos de O_3, NO_2 y SO_2; en $\mu g/m^3$ para PM_{10} y $PM_{2.5}$ y en ppm para CO. Con ayuda de OpenBUGS se calcula dicha cantidad para los modelos a comparar.

Otra forma de elegir modelo es calcular los MLF (Marginal Likelihood Function) de cada modelo y preferir el modelo con

el MLF más grande [3]. Para ello se utilizará la aproximación

$$V_l = \frac{1}{M}\sum_{i=1}^{M} L(D|\theta^{(i,l)}),$$

donde D son los datos, y $\theta^{(i,l)}$ es el i-ésimo elemento de la muestra generada usando el modelo "l". Aquí se prefiere el MLF (v_l) más grande.

Los DIC, y v_l de cada modelo y cada región se resumen en la Tabla B.6. De la elección del modelo individual con ayuda de los gráficos mostrados en las Figuras B.4 a B.9, apoyan la idea que el mejor criterio es el del MLF, ya se basa en la función de verosimilitud. El criterio del DIC no parece ser adecuado, pues al revisar la literatura, todo indica que este se ve afectado por el método que se utiliza para simular una variable [31], y ya que en el caso Weibull se utiliza la función ya implementada en el paquete OpenBUGS, en cambio para el caso Fréchet se utiliza el "truco de los ceros", lo cual incrementa el DIC.

De acuerdo al criterio del MLF, en todos casos se prefiere un modelo Fréchet en el contaminante O_3, y en la mayoría de los otros contaminantes un modelo Fréchet. En los casos en los cuales se prefiere el modelo Weibull es para el contaminante SO_2 en todas las regiones, y en el contaminante CO en las zonas CE y NO. Esta información se encuentra resumida en la Tabla B.7.

5.3. Elección de la cópula

Para este trabajo se esta interesado en trabajar con cópulas máx-estables, esto debido que los datos son máximos mensuales, por lo que se espera que se tenga relación en la cola derecha. Las cópulas máx-estables encontradas en la literatura son las familias de Galambos, de Gumbel–Hougaard, de

Hüsler–Reiss, de Tawn y t-EV. Jenkin y Clemitshaw hablan de la importancia de estudiar las reacciones químicas en el estudio de la contaminación y muestran las relaciones de los otros contaminantes a través de reacciones químicas [27]. Por tal razón se necesita una cópula que tenga concordancia positiva (usando la ρ de Spearman), hecho comprobado para los máximos mensuales utilizados en el presente escrito, cuyos resultados se muestran en la Tabla B.8.

La cópula de Twan se descarta ya que no mide concordancias cercanas a 1 y la de cópula t-EV se descarta que para cualquiera valores de v y p (parámetros de la cópula) da resultados parecidos a cópula de Gumbel-Hougaard y a la de Galambos [24].

Ahora se mencionaran las propiedades que tienen las cópulas máx-estables, y se observan las propiedades de que tienen para hacer una elección [46].

- Cópula de Gumbel-Hougaard. Esta cópula esta definida por

$$C_\theta(u,v) = \exp\left(-\left((-\log\ u)^\theta + (-\log\ v)^\theta\right)^{1/\theta}\right),$$

para $\theta \in [1,\infty)$, tomando valores de ρ entre $(0,1)$, además de ser la única cópula máx-estable y Arquimediana, donde $\lambda_L = 0$ y $\lambda = 2 - 2^{1/\theta}$.

- Cópula de Galambos. Está cópula esta definida por

$$C_\theta(u,v) = uv\exp\left(-\left((-\log\ u)^{-\theta} + (-\log\ v)^{-\theta}\right)^{-1/\theta}\right),$$

para $\theta \in [0,\infty)$, tomando valores de ρ entre (0,1), donde $\lambda_L = 0$ y $\lambda = 2^{-1/\theta}$.

- Cópula de Hüsler-Reiss. Esta cópula esta definida por

$$C_\theta(u,v) = \exp\left((\log\ u)\Phi\left(\frac{1}{\theta}+\frac{\theta}{2}\log\left(\frac{\log\ u}{\log\ v}\right)\right)\right.$$
$$\left.+(\log\ v)\Phi\left(\frac{1}{\theta}+\frac{\theta}{2}\log\left(\frac{\log\ v}{\log\ u}\right)\right)\right),$$

  donde Φ es la distribución normal estándar univariada y $\theta \in [0,\infty)$, y además $\lambda_L = 0$ y $\lambda = 2 - 2\Phi[1/\theta]$.

La cópula de Galambos se descarta por su parecido a la cópula de Gumbel-Hougaard y además existen varias conexiones profundas entre estas dos familias paramétricas de cópulas en cualquier dimensión, incluso su creación fue hecha para ver tales propiedades [19].

Por último, un análisis de comparación hecho por Genest y Favre [18], se llega a que no hay diferencia significativa entre usar las cópulas de Gumbel-Hougaard, Galambos y Hüsler-Reiss. Además, La cópula de Hüsler-Reiss no es tan sencilla de extender a dimensiones más altas como la de Gumbel-Hougaard y Galambos.

Con todo esto en mente, se llega a la conclusión de que por sus propiedades y sus usos, la mejor cópula para los objetivos establecidos en este trabajo es la cópula de Gumbel-Hougaard.

5.4. Análisis bivariado de los contaminantes

Una vez comprobado que la concordancia entre el O_3 y los otros contaminantes es positiva en las diversas regiones, es posible utilizar la cópula de Gumbel-Hougaard de la ecuación (1.6), que mide concordancia positiva.

Ahora, para estimar los parámetros del comportamiento conjunto tendrá la forma siguiente:

$$
\begin{aligned}
y &\sim Fréchet(\alpha_z, s_y),\\
z &\sim Fréchet(\alpha_y, s_z),\\
(y,z) &\sim f_{y,z,\theta}(y,z),\\
\alpha_y &\sim U(0,10),\\
s_y &\sim U(0,a_y),\\
\alpha_z &\sim U(0,10),\\
s_z &\sim U(0,a_z),\\
\theta &\sim U(1,15),
\end{aligned}
\tag{5.3}
$$

$$
\begin{aligned}
y &\sim Weibull_2(\alpha_y, b_y),\\
z &\sim Weibull_2(\alpha_z, b_z),\\
(y,z) &\sim f_{y,z,\theta},\\
\alpha_y &\sim U(0,10),\\
b_Y &\sim U(0,10),\\
k_y &= b_y^{-1/\alpha_y},\\
\alpha_z &\sim U(0,10),\\
b_z &\sim U(0,10),\\
k_z &= b_z^{-1/\alpha_z},\\
\theta &\sim U(1,15).
\end{aligned}
\tag{5.5}
$$

$$
\begin{aligned}
y &\sim Fréchet(\alpha_y, s_y),\\
z &\sim Weibull_2(\alpha_z, b_z),\\
(y,z) &\sim f_{y,z,\theta},\\
\alpha_y &\sim U(0,10),\\
s_y &\sim U(0,a_y),\\
\alpha_z &\sim U(0,10),\\
b_z &\sim U(0,10),\\
k_z &= b_z^{-1/\alpha_z},\\
\theta &\sim U(1,15).
\end{aligned}
\tag{5.4}
$$

$$
\begin{aligned}
y &\sim Weibull_2(\alpha_y, b_y),\\
z &\sim Fréchet(\alpha_z, s_z),\\
(y,z) &\sim f_{y,z,\theta},\\
\alpha_Y &\sim U(0,10),\\
b_y &\sim U(0,10),\\
k_y &= b_y^{-1/\alpha_y},\\
\alpha_z &\sim U(0,10),\\
s_z &\sim U(0,a_z),\\
\theta &\sim U(1,15).
\end{aligned}
\tag{5.6}
$$

En los modelos, el subíndice z indica que se trata del O_3, mientras que el subíndice y indica cualquier otro contaminante considerado. Los valores de a_y y a_z se toman de acuerdo a la Tabla B.4. Por otro lado, en la expresión $f_{y,z,\theta}$ hace referencia a la densidad conjunta dada por (1.8), donde se considera la

cópula de Gumbel-Hougaard[2]. Así, para el modelo (5.3),

$$\begin{aligned} f_{y,z,\theta}(y,z) = & \left(\frac{\alpha_y \alpha_z}{s_y s_z}\right)\left(\frac{y}{s_y}\right)^{-\alpha_y - 1}\left(\frac{z}{s_z}\right)^{-\alpha_z - 1} \\ & \times \exp\left(-\left[\left(\frac{y}{s_y}\right)^{-\alpha_y \theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{1}{\theta}}\right) \\ & \times \left\{\left[\left(\frac{y}{s_y}\right)^{-\alpha_y \theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{2}{\theta}-2}\right. \\ & \left.+(\theta-1)\left[\left(\frac{y}{s_y}\right)^{-\alpha_y \theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{1}{\theta}-2}\right\}, \end{aligned} \tag{5.7}$$

para (5.4) y (5.6)[3]

$$\begin{aligned} f_{y,z,\theta}(y,z) = & \left(\frac{b_y \alpha_y \alpha_z}{s_z^{-\alpha_z \theta}}\right)\left(y^{\alpha_y - 1} z^{\alpha_z - 1}\right)\exp\left(-b_y y^{\alpha_y}\right) \\ & \times \frac{\left(-\log\left[1 - e^{-b_y y^{\alpha_y}}\right]\right)^{\theta-1}}{\left[1 - e^{-b_y y^{\alpha_y}}\right]} \\ & \times \exp\left(-\left[\left(-\log\left[1 - e^{-b_y y^{\alpha_y}}\right]\right)^{\theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{1}{\theta}}\right) \\ & \left\{\left[\left(-\log\left[1 - e^{-b_y y^{\alpha_y}}\right]\right)^{\theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{2}{\theta}-2}\right. \\ & \left.+(\theta-1)\left[\left(-\log\left[1 - e^{-b_y y^{\alpha_y}}\right]\right)^{\theta} + \left(\frac{z}{s_z}\right)^{-\alpha_z \theta}\right]^{\frac{1}{\theta}-2}\right\}, \end{aligned}$$

[2]En la ecuación (1.8) se tienen subíndices con letras mayúsculas, sin embargo, debido a que en la estadística bayesiana se usan letras minúsculas para las variables aleatorias, se hace el cambio en ese capítulo.

[3]Se presenta la ecuación para el modelo (5.6), para el modelo (5.4) basta con cambiar las y's por z's y viceversa.

y para el modelo (5.5),

$$
\begin{aligned}
f_{y,z,\theta}(y,z) = {} & (\alpha_y\alpha_z)(b_y b_z)(y^{\alpha_y-1}z^{\alpha_z-1})\exp\left(-b_y y^{\alpha_y} - b_z z^{\alpha_z}\right) \\
& \times \frac{\left(-\log\left[1-e^{-b_y y^{\alpha_y}}\right]\right)^{\theta-1}\left(-\log\left[1-e^{-b_z z^{\alpha_z}}\right]\right)^{\theta-1}}{\left[1-e^{-b_y y^{\alpha_y}}\right]\left[1-e^{-b_z z^{\alpha_z}}\right]} \\
& \times \exp\left(-\left[\left(-\log\left[1-e^{-b_y y^{\alpha_y}}\right]\right)^{\theta}\right.\right. \\
& \left.\left.+\left(-\log\left[1-e^{-b_z z^{\alpha_z}}\right]\right)^{\theta}\right]^{\frac{1}{\theta}}\right) \\
& \left\{\left[\left(-\log\left[1-e^{-b_y y^{\alpha_y}}\right]\right)^{\theta}+\left(-\log\left[1-e^{-b_z z^{\alpha_z}}\right]\right)^{\theta}\right]^{\frac{2}{\theta}-2}\right. \\
& +(\theta-1)\left[\left(-\log\left[1-e^{-b_y y^{\alpha_y}}\right]\right)^{\theta}\right. \\
& \left.\left.+\left(-\log\left[1-e^{-b_z z^{\alpha_z}}\right]\right)^{\theta}\right]^{\frac{1}{\theta}-2}\right\}.
\end{aligned}
$$

Se hace el mismo procedimiento del presente Capítulo para estimar los parámetros en el caso univariado, así primero se analiza los gráficos BGR del número de iteraciones de calentamiento. Se encuentra las densidades de los parámetros después de las iteraciones de calentamiento y de tomar la muestra final cada 50 elementos, que tienen la forma de una densidad unimodal. Por último, en la Tablas B.9 a B.13 se encuentran las estimaciones de la media junto con su d. e. y Error M.C. de todos los parámetros en todas las regiones, cumpliendo la regla de que el Error M.C. es al menos 1/20 más pequeño que la d. e., teniendo buenas estimaciones. Los gráficos correspondientes se encuentran en [52].

Para asegurar tener buenas estimaciones, se analizan nuevamente los valores DIC y MLF de cada modelo en cada región. Los resultados se resumen en la Tabla B.14. Nuevamente se prefiere el criterio del MLF por las razones explicadas en la Sección anterior.

De acuerdo al criterio del MLF, en todos casos se prefiere un modelo Fréchet en el contaminante O_3, y en la mayoría de

los otros contaminantes un modelo Fréchet. En los casos en los cuales se prefiere el modelo Weibull es para el contaminante SO_2 en todas las regiones, y en el contaminante CO en las zonas CE y NO. Toda esta información se resume en la Tabla B.15.

En la Tabla B.16 se muestran las $\hat{\rho}$ asociadas a las θ's estimadas en las Tablas B.9-B.13. En estas tablas, se observa que la estimación de la asociación es aproximada a la muestral que se encuentra en la Tabla B.8.

5.5. Análisis trivariado de los contaminantes

El comportamiento de los contaminantes en las regiones CE, NE y SO es de gran importancia, ya que es la zona con mayor contaminación, mismas zonas que corresponden al llamado corredor del aire. Tomando esto en cuenta, se analizará los 3 principales contaminantes de estudio, O_3, PM_{10} y $PM_{2.5}$. Para ello se utilizará las cópulas tridimensionales y las cópulas asimétricas.

Como se observará mas adelante, solo se tomó el caso de todos los contaminantes se comportan como una distribución Fréchet, esto por las distribuciones elegidas del caso univariado y bivariado analizados en las secciones anteriores del Capítulo.

Para el contaminante O_3 se tiene que

$$\begin{aligned}\rho(CE, NE) &= 0.7819816,\\ \rho(CE, SO) &= 0.9415753,\\ \rho(NE, SO) &= 0.7318198,\end{aligned}$$

para PM_{10},

$$\begin{aligned}\rho(CE, NE) &= 0.6758211,\\ \rho(CE, SO) &= 0.7484863,\\ \rho(NE, SO) &= 0.6759925,\end{aligned}$$

y para $PM_{2.5}$,

$$\begin{aligned}\rho(CE, NE) &= 0.6191048,\\ \rho(CE, SO) &= 0.6328578,\\ \rho(NE, SO) &= 0.6460842.\end{aligned}$$

5.5.1. Usando cópulas trivariadas

El modelo usado en OpenBUGS para estimar los parámetros de la función de densidad que gobierna el comportamiento conjunto del O_3, PM_{10} y $PM_{2.5}$ de las variables X_1, X_2 y X_3 correspondiente a las regiones NE, CE y SO respectivamente, tienen la forma siguiente:

$$\begin{aligned}x_1 &\sim Fréchet(\alpha_{x_1}, s_{x_1}),\\ x_2 &\sim Fréchet(\alpha_{x_2}, s_{x_2}),\\ x_3 &\sim Fréchet(\alpha_{x_3}, s_{x_3}),\\ (x_1, x_2, x_3) &\sim f_{x_1,x_2,x_3,\theta}(x_1, x_2, x_3),\\ \alpha_{x_1} &\sim Unif(0, 10),\\ s_{x_1} &\sim Unif(0, a_1),\\ \alpha_{x_2} &\sim Unif(0, 10),\\ s_{x_2} &\sim Unif(0, a_2),\\ \alpha_{x_3} &\sim Unif(0, 10),\\ s_{x_3} &\sim Unif(0, a_3),\\ \theta &\sim Unif(1, 15).\end{aligned} \tag{5.8}$$

Los valores de las a_i's cambian de acuerdo al contaminante, tales valores se encuentran en la Tabla B.17, además la densidad conjunta $f_{x_1,x_2,x_3,\theta}(x_1,x_2,x_3)$ está dada por la Ecuación (1.17) aplicadas a la ecuaciones (1.16) y (1.18). Así, tomando

$$k_3 = \left(\frac{x_1}{s_{x_1}}\right)^{-\alpha_{x_1}\theta} + \left(\frac{x_2}{s_{x_2}}\right)^{-\alpha_{x_2}\theta} + \left(\frac{x_3}{s_{x_3}}\right)^{-\alpha_{x_3}\theta},$$

para el modelo (5.8) se tiene

$$\begin{aligned}
&f_{x_1,x_2,x_3,\theta}(x_1,x_2,x_3))\\
&= \left(\frac{\alpha_{x_1}\alpha_{x_2}\alpha_{x_3}}{x_1x_2x_3}\right)\left(\frac{x_1}{s_{x_1}}\right)^{2\alpha_{x_1}\theta}\left(\frac{x_2}{s_{x_2}}\right)^{2\alpha_{x_2}\theta}\left(\frac{x_3}{s_{x_3}}\right)^{2\alpha_{x_3}\theta}\\
&\quad\times \exp\left(-k_3^{\frac{1}{\theta}}\right)\left(\left(\frac{x_1}{s_{x_1}}\right)^{\alpha_{x_1}\theta}\left(\frac{x_2}{s_{x_2}}\right)^{\alpha_{x_2}\theta}\right.\\
&\quad\left.+\left(\frac{x_1}{s_{x_1}}\right)^{\alpha_{x_1}\theta}\left(\frac{x_3}{s_{x_3}}\right)^{\alpha_{x_3}\theta}+\left(\frac{x_2}{s_{x_2}}\right)^{\alpha_{x_2}\theta}\left(\frac{x_3}{s_{x_3}}\right)^{\alpha_{x_3}\theta}\right)^{-3}\\
&\quad\times k_3^{\frac{1}{\theta}}\left(1+2\theta^2-3k_3^{\frac{1}{\theta}}+k_3^{\frac{2}{\theta}}+3\theta\left(-1+k_3^{\frac{1}{\theta}}\right)\right).
\end{aligned}$$

En la Tabla B.18 se tienen las estadísticas obtenidas. En todos los casos se corrieron 5 cadenas y 100, 000 iteraciones, de las cuales se necesitaron 200, 000 iteraciones de calentamiento y se tomó la muestra final cada 50 elementos. Los gráficos correspondientes se encuentran en [52].

Se observa una buena concordancia, pues los valores de θ son relativamente grandes. Si se considera los promedios de las medidas ρ de Spearman presentadas al inicio de esta Sección, son parecidas a la concordancia asociada al parámetro θ presentado en la Tabla B.18, las cuales son 0.785496, 0.690288 y 0.572556 para los contaminantes O_3, PM_{10} y $PM_{2.5}$ respectivamente.

5.5.2. Usando cópulas asimétricas

El problema de usar cópulas tridimensionales es que no es fácil encontrar la medida de concordancia entre los contaminantes, esto debido que la forma de calcularlo solo es de manera directa para dos dimensiones. Con las cópulas asimétricas se permite ver concordancia entre X_1 y X_2, y (X_1, X_2) y X_3. Teniendo más variables X's, esta idea se puede generalizar.

De acuerdo la Sección 1.7.2, siguiendo la misma notación, si la cópula utilizada es de un parámetro, $\theta_1 < \theta_2$, por lo que se observa primero la medida ρ de Spearman de los contaminantes entre las tres regiones analizadas en esta Sección y cuyas medidas de ρ de Spearman se presentan al inicio de esta Sección.

Con lo observado, para el O_3, la pareja con mayor concordancia es CE y SO, que se consideran como las variables X_1 y X_2 respectivamente, y la región NE como X_3. Para las PM_{10} la pareja con mayor concordancia es CE y SO, que se consideran como las variables X_1 y X_2 respectivamente, y la región NE como X_3. Y por último, para las $PM_{2.5}$ la pareja con mayor concordancia es NE y SO, que se consideran como las variables X_1 y X_2 respectivamente, y la región CE como X_3.

Con lo anterior, se puede considerar la cópula asimétrica

$$C_1(C_2(u_1, u_2), u_3), \quad 1 < \theta_1 < \theta_2,$$

donde C_1 y C_2 son cópulas de Gumbel-Hougaard de parámetros θ_1 y θ_2 respectivamente, así el modelo aplicado en Open-

BUGS es de la forma siguiente:

$$\begin{aligned}
x_1 &\sim Fréchet(\alpha_{x_1}, s_{x_1}),\\
x_2 &\sim Fréchet(\alpha_{x_2}, s_{x_2}),\\
x_3 &\sim Fréchet(\alpha_{x_3}, s_{x_3}),\\
(x_1, x_2) &\sim f_{x_1,x_2,\theta_2}(x_1, x_2),\\
(x_1, x_2, x_3) &\sim f_{x_1,x_2,x_3,\theta_1,\theta_2}(x_1, x_2, x_3),\\
\alpha_{x_1} &\sim Unif(0, 10),\\
\alpha_{x_2} &\sim Unif(0, 10),\\
\alpha_{x_3} &\sim Unif(0, 10),\\
s_{x_1} &\sim Unif(0, a_1),\\
s_{x_2} &\sim Unif(0, a_2),\\
s_{x_3} &\sim Unif(0, a_3),\\
\theta_1 &\sim Unif(1, b_1),\\
\theta_2 &\sim Unif(b_2, 15).
\end{aligned} \tag{5.9}$$

En el modelo (5.9) la densidad conjunta $f_{x_1,x_2,\theta_2}(x_1, x_2)$ esta dada por la Ecuación (1.8) aplicada a la Ecuación (1.7). La densidad conjunta $f_{x_1,x_2,x_3,\theta_1,\theta_2}(x_1, x_2, x_3)$ está dada por la Ecuación (1.22) aplicada a las Ecuaciones (1.6) y (1.21). Los valores de los a_i se encuentran en la Tabla B.4. Así, en el modelo (5.9), $f_{x_1,x_2,\theta_2}(x_1, x_2)$ está dada por la Ecuación (5.7) y tomando

$$k_4 = \left(\left(\frac{x}{s_{x_1}} \right)^{-\alpha_{x_1}\theta_2} + \left(\frac{x_1}{s_{x_2}} \right)^{-\alpha_{x_2}\theta_2} \right)^{\frac{\theta_1}{\theta_2}} + \left(\frac{z_3}{s_{x_3}} \right)^{-\alpha_{x_3}\theta_1},$$

y

$$k_5 = \left(\frac{x}{s_{x_1}} \right)^{-\alpha_{x_1}\theta_2} + \left(\frac{x_1}{s_{x_2}} \right)^{-\alpha_{x_2}\theta_2},$$

se tiene que

$$
\begin{aligned}
& f_{x_1,x_2,x_3}(x_1,x_2,x_3) \\
=& k_4^{\frac{1}{\theta_1}} \left(\frac{\alpha_{x_1}\alpha_{x_2}\alpha_{x_3}}{x_1x_2x_3}\right) \left(\frac{x_1}{s_{x_1}}\right)^{\alpha_{x_1}\theta_2} \left(\frac{x_2}{s_{x_2}}\right)^{\alpha_{x_2}\theta_2} \left(\frac{x_3}{s_{x_3}}\right)^{\alpha_{x_3}\theta_1} \\
& \times \left(\left(\frac{x_1}{s_{x_1}}\right)^{\alpha_{x_1}\theta_2} + \left(\frac{x_2}{s_{x_2}}\right)^{\alpha_{x_2}\theta_2}\right)^{-2} \left(1 + k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_3}}\right)^{\alpha_{x_3}\theta_1}\right)^{-3} \\
& \times \exp\left(-k_4^{\frac{1}{\theta_1}}\right) \left\{\theta_1^2 \left(-1 + k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_2}}\right)^{\alpha_{x_2}\theta_1}\right)\right. \\
& + \theta_2 \left(1 + k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_2}}\right)^{\alpha_{x_2}\theta_1}\right) \left(-1 + k_4^{\frac{1}{\theta_1}}\right) \\
& + k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_2}}\right)^{\alpha_{x_2}\theta_1} \left[1 - 3k_4^{\frac{1}{\theta_1}} + k_4^{\frac{2}{\theta_1}}\right] \\
& + \theta_1 \left[\theta_2 + \theta_2 k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_2}}\right)^{\alpha_{x_2}\theta_1}\right. \\
& \left.\left. + \left(-1 + 2k_5^{\frac{\theta_1}{\theta_2}} \left(\frac{x_3}{s_{x_2}}\right)^{\alpha_{x_2}\theta_1}\right) \left(-1 + k_4^{\frac{1}{\theta_1}}\right)\right]\right\}.
\end{aligned}
$$

Dado que $\theta_1 < \theta_2$ los valores de b_1 y b_2 en el modelo (5.9) varían de acuerdo al contaminante analizado. Para el caso de O_3 se toma $b_1 = 2.5$ y $b_2 = 2.6$, para las PM_{10}, $b_1 = 2$ y $b_2 = 2.1$ y para el caso de las $PM_{2.5}$, $b_1 = 1.67$ y $b_2 = 1.671$.

En la Tabla B.19 se muestran las estadísticas obtenidas a través de OpenBUGS. En todos los casos, se corrieron 5 cadenas con $100,000$ iteraciones, de las cuales se tomaron $20,000$ de calentamiento para el O_3 y las $PM_{2.5}$, y $30,000$ para las PM_{10}. También se observa que el Error M. C. es al menos 20 veces mas pequeño que la d. e., esto para todos los parámetros. Los gráficos correspondientes se encuentran en [52].

5.5.3. Las DIC's para el caso tridimensional

Para seleccionar el mejor enfoque se utiliza el criterio del DIC, seleccionando del modelo con menor DIC. Tales estadísticas se encuentran en la Tabla B.20, en la que se observa que el enfoque tradicional de cópulas trivariadas es mejor para para todos los contaminantes.

Capítulo 6

Conclusiones

La contaminación del aire es un problema ambiental que ha empeorando con el paso del tiempo. Este tipo de contaminación es la que altera los gases suspendidos en la atmósfera. Las sustancias que inundan la capa de la atmósfera van incrementándose a partir de las fuentes contaminantes.

La contaminación del aire tiene consecuencias para la salud humana y el medio ambiente. Entre las consecuencias se encuentra el aumento del riesgo de infecciones respiratorias, enfermedades cardíacas, accidentes cerebrovasculares y cáncer de pulmón. Además, afecta el equilibrio climático del planeta, creando distintos eventos meteorológicos negativos que causan daños terrestres permanentes [2].

La contaminación del aire tiene múltiples causas, entre las que se incluyen en su mayoría los combustibles fósiles. En especial, en la ZMVM, al ser el centro del país de México, y por tanto, el lugar donde se concentra un gran número de la población a diferencia de otras zonas, tiene un grave problema de contaminación, en especial la relacionada con el aire. En la Tabla B.3 se observa que el valor máximo mensual promedio del O_3 es más grande al doble de la norma mexicana de ozono (95 ppb [37]), norma vigente hasta el año de observación de

los datos (2022), por consiguiente es interesante saber como se relacionan otros contaminantes con el ozono.

Se consideraron los máximos mensuales en todas las estaciones correspondientes a la región analizada, esto con la idea de que los máximos sean independientes mes con mes. Como dice la teoría de máximos extremos, las densidades Fréchet y la modificación de la Weibull se ajustaron a los datos de máximo mensuales de los diversos contaminantes, es más, en la mayoría de los casos el modelo Fréchet fue el elegido en el caso univariado, lo cual tiene sentido, ya que desde un principio, la subfamilia Fréchet es la que tiene una distribución sobre los valores positivos. Sin embargo, es buena idea considerar igualmente la distribución $Weibull_2$, ya que permitió un mejor ajuste para el contaminante SO_2 y algunos casos del CO.

Aunque el criterio de visualización de los histogramas y de las densidades obtenidas por medio de OpenBUGS es bueno para elegir un modelo que se ajuste a las densidades univariadas, para evitar dar opiniones personales del mejor ajuste, se aplicaron dos criterios, las estadísticas del DIC y el MLF. El criterio del DIC presenta problemas, ya que este esta relacionado con la variabilidad de la muestra obtenida, la cual a su vez se ve afectado por el método de la simulación de las densidades, como en el caso de las densidades Weibull que esta implementada en el paquete OpenBugs y la densidad Fréchet la cual se implemento "el truco de los ceros", forma que se siguiere en la literatura, sin embargo, Lunn y colaboradores aclaran que esta técnica es la manera de lograr dicho objetivo en OpenBUGS, pero se exhorta en evitarla lo más posible [31]. Por tal motivo se calcula el MLF por medio de su análogo muestral es la mejor opción, ya que este se basa en la función de máxima verosimilitud [3].

Escogiendo como criterio de decisión el estadístico MLF, se toma como mejor modelo el que tenga valor más grande, representando las elecciones en la Tabla B.7. Se observa en

las Figuras B.4 a B.9 que de las densidades con parámetros de la Tabla B.5 en efecto son buenos, otro punto a favor para utilizar las densidades elegidas con el criterio del MLF.

Para evitar sobre ajuste, se compararon las densidades marginales con los máximos mensuales del año 2023, datos que no se utilizaron en la estimación de los parámetros, ya que solo se consideraron para ello los datos hasta el año 2022 [53].

Con las estimaciones del análisis univariado, ya se podría elegir el modelo marginal para hacer los análisis multivariados, sin embargo, esto no se hizo, ya que no se quería omitir información por la concordancia, por lo que solo se utiliza los valores de la Tabla B.5 para elegir puntos de inicio para las diversas cadenas. La concordancia se modelo por medio de la cópula de Gumbel-Hougard, ya que como muestra la Tabla B.8, las concordancia de los máximos mensuales en las distintas regiones con el ozono es positiva, además de que dicha cópula mide dependencias en las colas.

Nuevamente se utilizo el estadístico MLF para seleccionar el mejor modelo bivariado de los contaminantes, tales MLF's se encuentran registrados en la Tabla B.14, observando que para el modelo bivariado, es mejor utilizar el modelo Fréchet en todas las regiones y con todas las combinaciones con los otros contaminantes para el O_3. Las elecciones de los modelos bivariados resumidos en la Tabla B.15, casi siempre con la elección del caso univariado resumidos en la Tabla B.7. No sorprende el cambio de elegir un Weibull a una Fréchet para el caso del contaminante PM_{10} región NO, ya que el MLF univariado de tal contaminante solo difiere de un cero en sus decimales (orden de e^{-959} y e^{-958} para caso Fréchet y Weibull respectivamente). El cambio de un modelo Weibull a Fréchet en la región NE para el contaminante CO no parece ser tan raro, debido que en la Figura B.9 inciso b), ambos ajustes parecen adecuados, esto se debe a que el rango de los máximos mensuales no es tan grande, y en el histograma no se ve un

pico tan alto como para ser una distribución Fréchet.

Tiene sentido que en la mayoría de los casos bivariados la mayoría sean ambas marginales Fréchet, esto debido a que si el O_3 es Fréchet, el cual tiene un efecto de hacer menos variables los datos a diferencia del caso Weibull, al tener la distribución conjunta, el otro contaminante se comporta de manera similar si la concordancia es alta. En el caso del SO_2, se observa que la concordancia es de no más del 42 % con el O_3 en todas las regiones, por lo que se permite conservar su densidad Weibull. Para el caso de la densidad conjunta del CO con el O_3 tiene las concordancia más altas, sin embargo, no en todos se elige que ambas marginales deban ser Fréchet, esto se explica debido a que no hay valores tan grandes, en especial en las regiones CE y SO, y al ser la densidad Weibull de cola más ligera que la Fréchet, se ajusta mejor a las regiones mencionadas. En la Tablas B.9-B.13 se informan los parámetros estimados en todos los modelos.

Observando la Tablas B.9-B.13, se ve que de los parámetros de asociación θ's presentes en los modelos, los valores más altos se dan por el par (CO, O_3) en todas las regiones. El más alto se encuentra en la región NO, seguido de las regiones CE, SO, SE y NE. Recordando que la región NO es una región con un número extremadamente alto de circulación de automóviles, autobuses y camiones, por lo tanto, los niveles de CO están destinados a ser altos y eso tiene efectos en la concentración de ozono. Los valores más pequeños de los parámetros de asociación se encuentran principalmente cuando se toma en cuenta el par $(PM_{2.5}, O_3)$, y el valor más bajo se produce cuando se consideran los datos de la región CE, seguido de cerca por el valor de la región SO. Notar que la región SO es la que tiene la media máxima mensual de $PM_{2.5}$ mas baja. Por lo tanto, parece que la concentración presente en esa región no es lo suficientemente alta como para producir algún efecto en los máximos mensuales de ozono. El segundo promedio máximo mensual más bajo de $PM_{2.5}$ se produce en

la región CE.

Los valores de ρ, para los modelos, obtenidos utilizando el software MATHEMATICA de Wolfram, se dan en la Tabla B.8 para todas las regiones. Se observa una buena estimación de la asociación, en especial para los modelos seleccionados. Por ejemplo, en el caso de (CO, O_3), que tiene parámetros de asociación más altos, también tiene correlaciones de muestra más altas (superiores a 0.7 en todas las regiones) y ρ de Spearman más altas.

También gustaría llamar la atención sobre el hecho de que en el presente caso no siempre el modelo bidimensional más adecuado utilizando cópula corresponde al caso donde las distribuciones marginales son las seleccionadas en el análisis unidimensional. Por ejemplo, en el caso de CO y O_3 y la región NE, se tiene que las distribuciones unidimensionales respectivas son Weibull y Fréchet (Tabla B.7). Sin embargo, cuando se considera el modelo bidimensional tenemos que el modelo que supone la distribución de Fréchet para ambos contaminantes es el modelo seleccionado para esa región. Esto podría explicarse por el hecho de que la distribución de Fréchet también representa bien el comportamiento de los valores máximos mensuales de CO en la región NE con valores mayores para la función de densidad en comparación con los dados por la distribución de Weibull.

Los modelos bidimensionales seleccionados proporcionaron, en la mayoría de los casos, buenas aproximaciones, a través del ρ de Spearman, a las correlaciones muestrales de los pares de contaminantes (véase la Tabla B.16). Incluso en los pocos casos en los que las aproximaciones no son óptimas, sus diferencias no son demasiado grandes. Por lo tanto, proporcionan información sobre algunos precursores del ozono que tienen una mayor influencia en su concentración. Por ejemplo, en el caso del CO, en todas las regiones, tenemos grandes correlaciones con O_3. Por lo tanto, la disminución de su emisión podría resultar en una disminución de los niveles de ozono

en todas las regiones. Por otro lado, la influencia del SO_2, aunque relevante, parece no ser tan grave como la del CO en todas las regiones.

Al final de toda la explicación anterior, hay que señalar que las estimaciones son significativamente apropiadas, en especial del caso bivariado. Estos resultados permiten entender la relación entre los máximos mensuales de O_3 con los otros contaminantes, tal vez excepto para el $PM_{2.5}$, debido a que los dados se empiezan a analizar por la RAMA en agosto de 2003, lo que hace que se tengan menos datos, lo cual afectaría que las estimaciones no sean del todo confiables. Esto también explica que para el análisis bivariado del O_3 y $PM_{2.5}$ si fue posible hacer estimaciones suponiendo el $PM_{2.5}$ Fréchet y O_3 Weibull, pues la cantidad de datos no es lo suficientemente grande para que se observe la convergencia de su densidad[1].

Además de estimar correlaciones futuras entre pares de contaminantes y, por lo tanto, cuánta correlación podrían tener en el comportamiento futuro de los datos si no ocurren cambios, también se puede obtener la probabilidad de que sus mediciones pertenezcan a un conjunto dado. Por ejemplo, tomando el par (CO, O_3) y la región SO. Se Supone que se esta interesado en conocer la probabilidad de tener CO en el intervalo $[11.00, 13.00]$ que corresponde a tener mala calidad del aire en la región de acuerdo con el índice de calidad "Aire y Salud" [38] y de tener O_3 en el intervalo $[154, 204]$ que corresponde al intervalo de la Fase II de la emergencia ambiental en la Ciudad de México [34]. Por lo tanto, se necesita obtener,

$$P((CO, O_3) \in [11, 13] \times [154, 204]) = 1.99E - 2.$$

El enfoque de las cópulas no solo se puede usar para anali-

[1]En los otros análisis no fue posible suponer la primera marginal Fréchet y la segunda marginal correspondiente al O_3 Weibull, al intentar correr el modelo en OpenBUGS, este paraba en no más de 50 iteraciones por cadena, mostrando que los datos no se ajustaban bien a estas suposiciones

zar contaminantes con densidades bivariadas, se puede hacer con más dimensiones. Un problema en la ZMVM es el llamado corredor del aire, comprendida por las zonas CE, NE y SO. Por las características de la cópula elegida, se necesita que las concordancias de las variables aleatorias sean positivas en triadas.

Para comprobar que la relación del O_3 en las regiones NE, CE y SO es positiva, se utiliza la ρ de Spearman, esto para las permutaciones de dos elementos tomados de las tres regiones. Se tiene que todas las concordancias son positivas, por lo que se considera usar la cópula de Gumbel-Hougard, considerando X_1, X_2 y X_3 como los niveles máximos mensuales en las regiones NE, CE y SO. Este mismo enfoque se aplico para los contaminantes PM_{10} $PM_{2.5}$, con el mismo orden de variables y la misma cópula al tener igualmente concordancias positivas.

De acuerdo a los análisis hechos, es factible y lógico considerar que las marginales sean Fréchet. Aunque se considero que las marginales fueran Weibull, al ejecutarlas en OpenBUGS, no se lograron hacer más de 1,000 iteraciones, caso contrario al considerar todas las marginales Fréchet.

La teoría clásica de cópulas multivariadas parece describir la relación entre las tres regiones, sin embargo, al tener solo un parámetro asociado a la concordancia podría no modelar tal medida adecuadamente. Por consiguiente, se considera las cópulas asimétricas, que permiten modelar mejor la concordancia, en especial con la cópula de Gumbel-Hougard, que al considerar 3 variables, se tienen 2 parámetros para modelar la relación entre las regiones del corredor del aire.

Para el caso de las cópulas asimétricas, hay que considerar a X_1 y X_2 como la pareja que tenga mayor concordancia entre las permutaciones de dos elementos tomados de las tres regiones, y como X_3 a la región restante. Para el O_3 y PM_{10} el orden las variables es CE, SO y NE, ya que la ρ de Spearman

muestrales para las parejas (CE, NE), (CE,SO) y (NE,SO) es de 0.7819816, 0.9415753 y 0.7318198 para el O_3 y 0.6758211, 0.7484863 y 0.6759925 para las PM_{10} respectivamente. En el caso de las $PM_{2.5}$ el orden es NE, SO y CE, ya que las concordancias muestrales estimadas para las parejas (CE, NE), (CE,SO) y (NE,SO) son de 0.6191048, 0.6328578 y 0.6460842 respectivamente.

Comparando los resultados de las cópulas en 3 dimensiones y las cópulas asimétricas para 3 variables presentados en las Tablas B.18 y B.19 respectivamente, se observa que las cópulas asimétricas son mejores para modelar la concordancia entre las variables para los contaminantes O_3 y PM_{10}. Esto se concluye por el hecho de que θ_1 y θ_2 presentado en la Tabla B.19 son significativamente distintos, como también se observa en las ρ de Spearman muestrales. Para el caso de las $PM_{2.5}$ se considera mejor usar las cópulas trivariadas, esto es porque las concordancias entre los pares de regiones no es tan distinta, y un solo parámetro es adecuado, junto con que los valores de θ, θ_1 y θ_2 son muy parecido, con lo que concuerda con las ρ de Spearman muestrales observadas.

Utilizando el DIC para seleccionar el modelo que mejor se ajusta a los datos, la cópula habitual de Gumbel-Hougaard es el modelo seleccionado para los tres contaminantes considerados en este estudio (ver Tabla B.20). Al observar la Tablas B.18 y B.19, que el mayor parámetro de asociación en el modelo seleccionado está relacionado con el contaminante O_3, seguido de PM_{10} y $PM_{2.5}$. Este es un resultado razonable, ya que el O_3 viaja largas distancias y sus precursores viajan en la dirección donde se encuentra las regiones que conforman el corredor del aire. Por lo tanto, es más probable que el parámetro de asociación relacionado con este contaminante sea el mayor. En el caso de PM_{10} se tiene la presencia de fuentes de este contaminante en la región noreste debido al gran número de fábricas y también en la región sur y noroeste debido a la gran presencia de camiones. En el caso de $PM_{2.5}$

hay una gran presencia de fábricas en la región noreste, pero también una gran cantidad de vehículos en las regiones centro y suroeste. Por lo tanto, tal vez esta sea la razón para el parámetro de asociación más pequeño producido.

Además de la información que proporcionan los valores de los parámetros de asociación, otra información que se puede obtener está relacionada con la probabilidad de tener un determinado contaminante en intervalos de interés en cada una de las regiones. Por ejemplo, si se considera la regla para declarar la Fase I de una alerta de emergencia en la Ciudad de México [8], entonces es necesario tener O_3 por encima del umbral de 154 ppb en una de sus regiones. En el caso de PM_{10}, es necesario que el umbral correspondiente de 214 $\mu g/m^3$ se exceda en al menos dos de sus regiones. También se puede interesar el caso de que no se declare la Fase II debido a excedencias de O_3 en estas tres regiones del corredor del aire. Por lo tanto, las mediciones de O_3 deberían estar en un intervalo de interés dado en todas las regiones. De manera similar, en los casos de PM_{10} y $PM_{2.5}$.

Se supone que se está interesado en conocer la probabilidad de que se supere el umbral necesario para declarar la Fase I, pero no la Fase II, por el máximo mensual en las tres regiones. Esto significa que en al menos un día en un mes determinado se declararía una alerta de emergencia de Fase I debido a que se superan los umbrales de O_3 o PM_{10} de 154 ppb y 214 $\mu g/m^3$, respectivamente. Sin embargo, los umbrales de 204 ppb y 354 $\mu g/m^3$ para la Fase II no se superan en ninguna de las tres regiones. En este caso, se calcularían las probabilidades siguientes. Si el contaminante de interés es el O_3, entonces

$$P((O_3^{NE}, O_3^{CE}, O_3^{SO}) \in [154, 204) \times [154, 204) \times [154, 204)) \\ = 4.36E-2,$$

y

$$P((PM_{10}^{NE}, PM_{10}^{CE}, PM_{10}^{SO}) \in [214, 354) \times [214, 354) \times [214, 354)) \\ = 1.56E-2,$$

si el contaminante de interés es PM_{10}. Estas probabilidades fueron calculadas usando las cópulas tridimensionales.

Como se puede observar, se logro obtener buenos modelos para analizar la contaminación tanto de forma univariada, bivariada y trivariada. Con la información expuesta, se pueden hacer varios análisis a futuro, como la imputación de datos faltantes o la predicción a futuro, en especial, los niveles de retorno para predecir que en algún momento los niveles de contaminación lleguen a niveles críticos, como exponen Vazquez, Rodrigues y Reyes [53]. La presente tesis contribuye al conocimiento de la matemática, la forma de analizar la contaminación y sobre todo, como idea de aplicación a otras áreas del conocimiento. Otra contribución es que al buscar investigaciones sobre contaminación y cópulas asimétricas, no hay muchos trabajos con tales ideas, ya que se han aplicado principalmente al campo de la hidrología, y para la contaminación con más de 2 variables, se encontró la aplicación de las C-Vine cópulas, idea que podría ser implementada en trabajos futuros y comparar cual enfoque es mejor.

Se propone continuar con el análisis de los umbrales, cuyos primeros avances se encuentran en el Capítulo B. Se busca seleccionar una mejor función de distribución para las marginales y analizar de manera más exhaustiva las propiedades conjuntas de los datos, con el fin de elegir otra cópula que represente mejor la asociación entre las variables.

Apéndice A

Conceptos y Resultados Básicos

A continuación, se presentan algunos conceptos y resultados necesarios para la comprensión de este trabajo. Para este apéndice, es necesario tener conocimientos básicos de teoría de conjuntos, los cuales pueden ser consultados en la sección 3 del capítulo 1 en [43], además de conocimientos de cálculo de límites, disponibles en [1].

A.1. Variables aleatorias

Definición A.1 (Medida de probabilidad)**.** *Sea* $(\Omega, \mathscr{F})$ *un espacio medible. Una medida de probabilidad es una función* $P : \mathscr{F} \to [0, 1]$ *que satisface*

1. $P(\Omega) = 1$.

2. $P(A) \geq 0$, *para cualquier* $A \in \mathscr{F}$.

3. *Si* $A_1, A_2, \ldots \in \mathscr{F}$ *son ajenos dos a dos, esto es,* $A_n \cap$

$$A_m = \emptyset \text{ para } n \neq m, \text{ entonces } P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Si se desea saber propiedades de la medida de probabilidad, se puede consultar [44].

Definición A.2 (Espacio de probabilidad)**.** *Un espacio de probabilidad es una terna* $(\Omega, \mathscr{F}, P)$*, en donde* Ω *es un conjunto arbitrario,* $\mathscr{F}$ *es una* σ*-álgebra de subconjuntos de* Ω*, y* P *es una medida de probabilidad definida sobre* $\mathscr{F}$*.*

Definición A.3 (Variable aleatoria)**.** *Una variable aleatoria es una transformación* X *del espacio de resultados* Ω *al conjunto de números reales, es decir,*

$$X : \Omega \to \mathbb{R},$$

tal que para cualquier número real x*,*

$$\{\omega \in \Omega : X(w) \leq x\} \in \mathscr{F}.$$

Proposición A.1. *Sean* X *y* Y *variables aleatorias y* c *una constante, entonces*

a) cX*,*

b) $X + Y$*,*

c) XY*,*

d) X/Y *donde* $Y \neq 0$*,*

e) máx$\{X, Y\}$ *y*

f) mín$\{X, Y\}$

son variables aleatorias.

Definición A.4 (Función de distribución)**.** *La función de distribución de una variable aleatoria X es la función $F : \mathbb{R} \to [0, 1]$, definida por*

$$F(x) = P(X \leq x), \qquad x \in \mathbb{R}.$$

Proposición A.2. *Sea $F(x)$ la función de distribución de una variable aleatoria. Entonces*

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

Demostración. Véase [44]. □

Definición A.5 (Variable aleatoria discreta)**.** *La variable aleatoria X se llama discreta si su correspondiente función de distribución $F(x)$ es una función constante por pedazos. Sean $x_1, x_2, \ldots$ los puntos de discontinuidad de $F(x)$. En cada uno de estos puntos el tamaño de la discontinuidad es $P(X = x_i) = F(x_i) - F(x_i-) > 0$. A la función $f(x)$ que indica estos incrementos se le llama función de probabilidad de X, y se define como sigue*

$$f(x) = \begin{cases} P(X = x) & \text{si } x = x_1, x_2, \ldots \\ 0 & \text{otro caso.} \end{cases}$$

La función de distribución se reconstruye de la forma siguiente

$$F(x) = \sum_{u\leq x} f(u).$$

Definición A.6 (Variable aleatoria continua)**.** *La variable aleatoria X se llama continua si su correspondiente función de distribución es una función continua.*

Definición A.7 (Variable aleatoria absolutamente continua)**.** *La variable aleatoria continua X con función de distribución $F(x)$ se llama absolutamente continua, si existe una función no negativa e integrable f tal que para cualquier valor de x se cumple*

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

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

Para hablar de una variable aleatoria, es suficiente hacer referencia a su función de distribución o función de densidad, ya que de una se puede obtener la otra.

Definición A.8 (Independencia de variables aleatorias)**.** *Se dice que las variables aleatorias X y Y son independientes si los eventos $(X \leq x)$ y $(Y \leq y)$ son independientes para cualesquiera valores reales de x y y, es decir, si se cumple la igualdad*

$$P[(X \leq x) \cap (Y \leq y)] = P(X \leq x)P(Y \leq y).$$

Definición A.9 (Esperanza)**.** *Sea X una variable aleatoria discreta con función de probabilidad $f(x)$. La esperanza de X se define como el número*

$$E[X] = \sum_{x} xf(x).$$

suponiendo que esta suma es absolutamente convergente, es decir, cuando la suma de los valores absolutos es convergente. Por otro lado, si X es continua con función de densidad $f(x)$, entonces la esperanza es

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

suponiendo que esta integral es absolutamente convergente, es decir, cuando la integral de los valores absolutos es convergente.

Definición A.10 (Varianza)**.** *Sea X una variable aleatoria discreta con función de probabilidad $f(x)$. La varianza de X se define como el número*

$$V[X] = \sum_x (x-\mu)^2 f(x).$$

cuando esta suma es convergente y en donde μ es la esperanza de X. Para una variable aleatoria continua X con función de densidad $f(x)$ se define

$$V[X] = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx,$$

cuando esta integral es convergente.

Proposición A.3. *Sean X y Y dos variables aleatorias con varianza finita y sea c una constante. Entonces*

1. $V[X] \geq 0$.
2. $V[c] = 0$.
3. $V[cX] = c^2 V[X]$.
4. $V[X+c] = V[X]$.
5. $V[X] = E[X^2] - E^2[X]$.
6. *En general,* $V[X+Y] \neq V[X] + V[Y]$.
7. *Si X y Y son variables independientes, entonces*

$$V[X+Y] = V[X] + V[Y].$$

Demostración. Véase [43]. □

A.2. Simulación de variables aleatorias

Actualmente la simulación de una variable aleatoria con cierta distribución esta programada en varios software como R. Por otro lado, existen variables aleatorias no tan comunes que no están programadas o bien, un investigador propone una nueva variable, por ello, se necesita métodos para simular variables aleatorias.

Todos los software o paquete estadístico ya tiene programada la simulación de una variable aleatoria uniforme, hecho que se utiliza para simular otras variables, como se mostrará a continuación.

Definición A.11. *Para una función no decreciente F en $\mathbb{R}$, el inverso generalizado de F, F^-, es la función definida por*

$$F^-(u) = \text{ínf}\{x : F(x) \geq u\}.$$

Lema A.1 (Transformación Integral de Probabilidad)**.** *Si $U \sim Unif[0,1]$ y F una función de distribución, entonces la variable aleatoria $F^-(U)$ tienen la distribución F.*

Demostración. Véase [45]. □

Por lo tanto, de acuerdo al Lema anterior, para generar una variable aleatoria X que tenga función de distribución F, es suficiente generar $U \sim Unif[0,1]$ y luego hacer la transformación $x = F^-(u)$.

Ejemplo A.1. *Si $X \sim Gumbel(\mu, \sigma)$, entonces $F(x) = e^{-e^{-(x-\mu)/\sigma}}$, entonces resolviendo para x en $u = e^{-e^{-(x-\mu)/\sigma}}$ se tiene $x = \mu - \sigma\log(-\log(u))$. Por lo tanto, si $U \sim Unif[a,b]$, ls variable aleatoria $X = \mu - \sigma\log(-\log(U))$ tiene distribución Gumbel.*

Teorema A.1 (Teorema fundamental de simulación)**.** *Simular X con función de densidad $f(x)$ es equivalente a simular*

$$(X, U) \sim Unif\{(x, u) : 0 < u < f(x)\}.^{1}$$

Demostración. Véase [45]. □

Ejemplo A.2. *Sea X una variable aleatoria con función de densidad $f(x)$ y $m > 0$ tal que $f(x) \leq m$, para toda $x \in \mathbb{R}$. Se supone que $f(x) = 0$ para $x \notin [a, b]$, donde $a < b$, entonces se puede simular el par $(Y, U) \sim Unif\{(y, u) : 0 < u < m\}$ y $U|Y \sim Unif(0, m)$, y tomando el par solo si $0 < u < f(y)$ es satisfecho. Esto da como resultado la distribución correcta del valor aceptado de Y, es decir X, porque*

$$\begin{aligned} P(X \leq x) &= P(Y \leq x | U < f(Y)) \\ &= \frac{\int_a^x \int_0^f (y) du dy}{\int_a^b \int_0^f (y) du dy} \\ &= \int_a^x f(y) dy. \end{aligned}$$

A.3. Estadísticas

En esta sección, se analizan los conceptos de estadística que permiten una comprensión del trabajo realizado. Si se desea ver ejemplos de las definiciones, véase [35].

Definición A.12 (Muestra aleatoria)**.** *Una muestra aleatoria de tamaño n es una muestra elegida por un método en el que cada colección de n elementos de la población tiene la misma probabilidad de formar la muestra.*

[1] Vector uniforme en $\Omega = \{(x, u) : 0 < u < f(x)\}$, para mas información véase [45].

Definición A.13 (Media (muestral))**.** *Sea $X_1, \ldots, X_n$ una muestra. La media muestral es*

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Definición A.14 (Varianza (muestral))**.** *Sea $X_1, \ldots, X_n$ una muestra. La varianza muestral es la cantidad*

$$V = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2. \tag{A.1}$$

Lema A.2. *Una fórmula equivalente a (A.1), que puede ser más fácil de calcular, es*

$$V = \frac{1}{n-1} \left(\sum_{i=1}^{n} X_i^2 - n\bar{X} \right).$$

Demostración. Véase [25]. □

Definición A.15 (Mediana (muestral))**.** *Si n números están ordenados del más pequeño al más grande:*

- *Si n es impar, la mediana muestral es el número en la posición $\dfrac{n+1}{2}$.*
- *Si n es par, la mediana muestral representa el promedio de los números en las posiciones $\dfrac{n}{2}$ y $\dfrac{n}{2} + 1$.*

Definición A.16 (Cuartiles)**.** *Los cuartiles dividen la muestra tanto como sea posible en cuartos. Una muestra tiene tres de aquéllos. El método más simple cuando se calcula es el siguiente: Sea n el tamaño de la muestra. Ordene los valores de la muestra del más pequeño al más grande. Para encontrar el primer cuartil, calcule el valor $0.25(n+1)$. Si éste es un entero, entonces el valor de la muestra en esa posición es el*

primer cuartil. Si no, tome entonces el promedio de los valores de la muestra de cualquier lado de este valor. El tercer cuartil se calcula de la misma manera, excepto que se usa el valor $0.75(n+1)$. *El segundo cuartil*[2] *usa el valor* $0.5(n+1)$.

Definición A.17. *El p-ésimo percentil de una muestra, para un número p entre 0 y 100, se divide a la muestra tanto como sea posible, el* $p\,\%$ *de los valores de la muestra es menor que el p-ésimo percentil y el* $(100-p)\,\%$ *son mayores. Se ordena los valores de la muestra del más pequeño al más grande y después se calcula la cantidad* $(p/100)(n+1)$, *donde n es el tamaño de la muestra. Si esta cantidad es un entero, el valor de la muestra en esta posición es el p-ésimo percentil. Por otro lado, se promedia los dos valores de la muestra en cualquier lado.*

[2]El segundo cuartil coincide con la mediana.

Apéndice B

Análisis por Umbrales

Un enfoque que se intento fue el de analizar los máximos diarios a través de los cuantiles, más específicamente aquellos datos mayores al 97.5-percentil. Para ello se considero el mismo conjunto de datos que son descritos en el Capítulo 5.

B.1. Estadísticas y distribuciones para los umbrales de forma marginal

En la Tabla B.21 se presentan algunas estadísticas de los máximos diarios en toda la ZMVM, en especial el 97.5-percentil. Ahora también es importante encontrar la relación entre los máximos diarios de O_3 con los otros contaminantes, para ello se ocupa la ρ de Spearman, cuya version muestral se encuentra en la Tabla B.22.

Con estas estadísticas se procede a encontrar el ajuste a las distribuciones candidatas.

B.2. Análisis univariado

Como se están trabajando con máximos diarios, se piensa que los valores igual tendrían una densidad Fréchet o Weibull como se explica en el Capítulo 2. A diferencia del análisis principal presentado en este trabajo, ahora se quiere solo analizar los datos que son mayores al 97.5-percentil. Así, los modelos usados en OpenBUGS son de la forma:

$$\begin{aligned} x &\sim Fréchet(\alpha_x, s_x, b_x), \\ \alpha &\sim Unif(0, 10), \\ s &\sim Unif(0, a_x). \end{aligned} \qquad \text{(B.1)}$$

$$\begin{aligned} x &\sim Weibull(\alpha_x, k_x, b_x), \\ \alpha &\sim Unif(0, 10), \\ k &\sim Unif(0, b_x). \end{aligned} \qquad \text{(B.2)}$$

En los modelos (B.1) y (B.2) los valores de los limites superiores en los parámetros de escala se encuentran en la Tabla B.23, estos valores cambian de acuerdo a cada contaminante. De igual manera, el valor de μ_x es la parte entera del 97.5-percentil de cada contaminante, los cuales se encuentran en la Tabla B.21.

En OpenBUGS se corrieron 5 cadenas en puntos iniciales distintos, con un total de $100,000$ iteraciones. Se tomaron $20,000$ iteraciones de calentamiento en todos los casos, ademas para tener una muestra aleatoria, se tomo la muestra cada 50 elementos. Los resultados se resumen en la Tabla B.24, donde se observa que se cumple la regla empirica de que el error M. C. sea 1/20 veces más pequeño que la d. e. en cada parámetro del cual se quiere hacer inferencia.

En la Tabla B.25 se registran los DIC y MLF donde se observa que la distribución que mejor se ajusta es la Weibull, esto acompañado de las densidades observadas en la Figura B.10, esto debido a que no hay un pico muy pronunciado al inicio del histograma.

B.3. Análisis conjunto de los umbrales

Ahora se analiza el O_3 con los otros contaminantes considerados. Los vectores de datos considerados son los que en ambos entradas son mayores a su 97.5-percentil correspondiente. Dado que hay dependencias tanto positivas como negativas, se utiliza la cópula de Frank para modelar la dependencia mostrada en la Tabla B.22.

La cópula de Frank esta dada por

$$C_\theta(u,v) = -\frac{1}{\theta}\log\left(1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right),$$

donde $\theta(-\infty,\infty)\setminus\{0\}$, siendo el caso límite de $\theta \to 0$ la dependencia nula. Para hacer la simulaciones en OpenBUGS, se necesita la densidad asociada a la cópula, por lo que su derivada cruzada es dada por

$$c_\theta(u,v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u,v) = \frac{\theta e^{-\theta(u+v)}(1-e^{-\theta})}{(e^{-\theta(u+v)}-e^{-\theta u}-e^{-\theta v}+e^{-\theta})^2}.$$

Para obtener la expresión dada por Zhang y Shigh [56], la expresión anterior se multiplica al numerador y denominador por $(e^{\theta(1+u+v)})^2$, por lo tanto

$$c_\theta(u,v) = \frac{\theta(e^{-\theta}-1)e^{-\theta(1+u+v)}}{(e^{\theta(u+v)}-e^{\theta(1+u)}-e^{\theta(1+v)}+e^{\theta})^2}.$$

Por tanto, la densidad conjunta de un vector aleatorio (x,y) por medio del Teorema de Sklar está dada por

$$f_{X,Y,\theta}(x,y) = f_X(x) f_Y(y) c_\theta(x,y),$$

donde f una función de densidad.

Aunque en el caso univariado se determinó que se debe usar una marginal Weibull, para el análisis bivariado se considera utilizar una marginal Fréchet, con el fin de no perder información.

Los modelos para estimar los parámetros del comportamiento conjunto tendrán la forma siguiente[1]:

$$\begin{aligned} y &\sim Weibull(\alpha_y, k_y, \mu_y), \\ z &\sim Weibull(\alpha_z, k_z, \mu_z), \\ (y,z) &\sim f_{y,z,\theta}, \\ \alpha_y &\sim U(0,10), \\ k_y &\sim U(0,b_y), \\ \alpha_z &\sim U(0,10), \\ k_z &\sim U(0,b_z), \\ \theta &\sim U(-35,35). \end{aligned} \tag{B.3}$$

$$\begin{aligned} y &\sim Fréchet(\alpha_y, s_y, \mu_y), \\ z &\sim Weibull(\alpha_z, k_z, \mu_z), \\ (y,z) &\sim f_{y,z,\theta}, \\ \alpha_y &\sim U(0,10), \\ s_y &\sim U(0,a_y), \\ \alpha_z &\sim U(0,10), \\ k_z &\sim U(0,b_z), \\ \theta &\sim U(-35,35). \end{aligned} \tag{B.4}$$

$$\begin{aligned} y &\sim Weibull_2(\alpha_y, k_y, \mu_y), \\ z &\sim Fréchet(\alpha_z, s_z, \mu_z), \\ (y,z) &\sim f_{y,z,\theta} \\ \alpha_y &\sim U(0,10), \\ k_y &\sim U(0,b_y), \\ \alpha_z &\sim U(0,10), \\ s_z &\sim U(0,a_z), \\ \theta &\sim U(-35,35). \end{aligned} \tag{B.5}$$

Los valores de las a's y b's se encuentran en la Tabla B.26

[1]Para el caso de suponer las dos variables Fréchet no fue posible, ya que, en OpenBUGS el parámetro de escala para el O_3 tendia a crecer demasiado, superando el valor de $2^{31}-1$, el valor más grande posible en OpenBUGS.

y el valor de los parámetros μ's se toman el 0.975-percentil encontrados en la Tabla B.21. El subindice z se refiere al O_3, mientras que el subindice y se usa para los otros contaminantes.

Para todos los casos se corrieron 5 cadenas en puntos iniciales distintos, con un total de $100,000$ iteraciones, de las cuales se tomaron $20,000$ iteraciones de calentamiento. Después, para tener una muestra aleatoria se toma la muestra final cada 50 elementos. Se tiene que hay que hacer tantas simulaciones hasta que el d. e. sea $1/20$ más pequeño que el Error Monte Carlo, como se muestra en la Tabla B.27.

En la Tabla B.28 se muestra que en general los mejores modelos son cuando ambas densidades son Weibull, esto con el criterio del MLF, excepto el caso de las $PM_{2.5}$ y el O_3 que se elige marginal Weibull para las $PM_{2.5}$ y marginal Fréchet para el O_3, aunque solo por poco.

Índice de Figuras

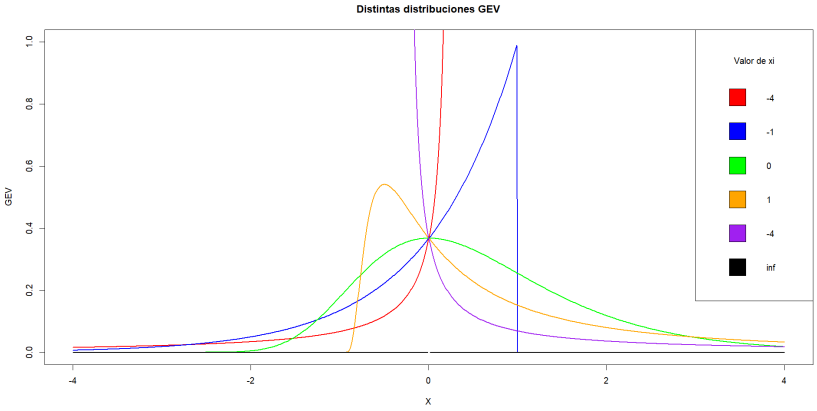
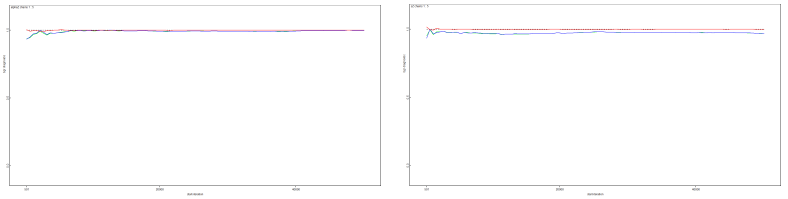
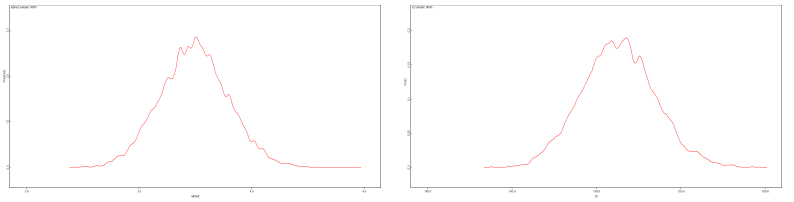


Figura B.1: *Gráfica de la función de distribución GEV.*



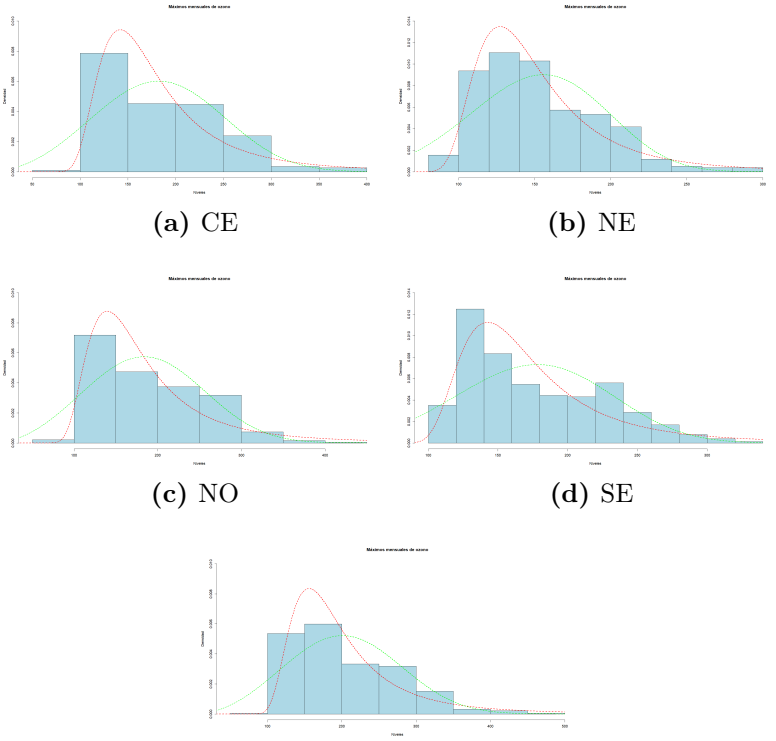
(a) α, CE **(b)** s, CE

Figura B.2: *Gráfico BGR de los parámetros de ubicación(α) y de escala(s) para los máximos mensuales de O_3 en la región CE, suponiendo un modelo Fréchet.*



(a) α, CE **(b)** s, CE

Figura B.3: *Densidad de los parámetros de ubicación(α) y de escala(s) para los máximos mensuales de O_3 en la región CE, suponiendo un modelo Fréchet.*

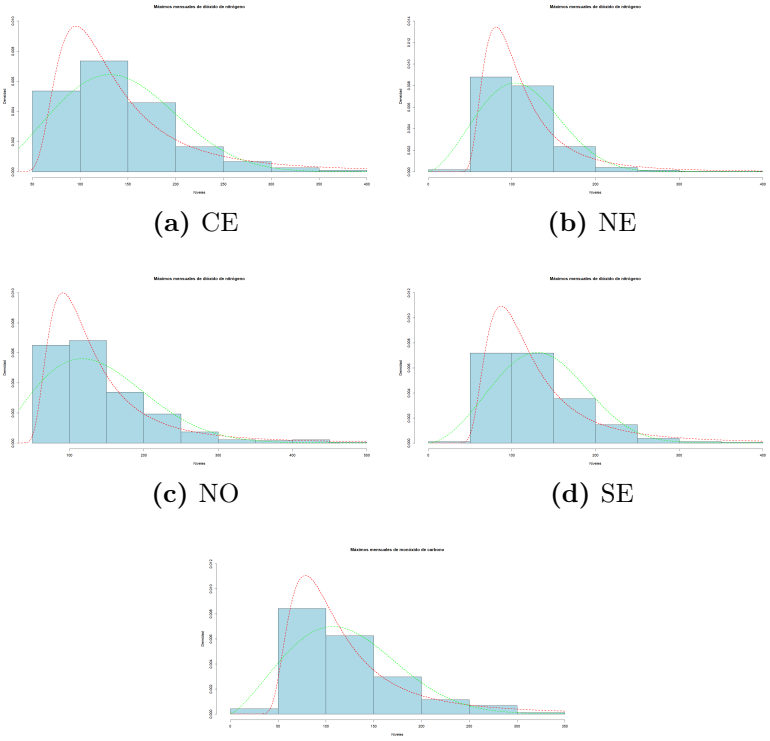


(a) CE **(b)** NE

(c) NO **(d)** SE

(e) SO

Figura B.4: *Comparación de las densidades obtenidas para los máximos mensuales de O_3. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*

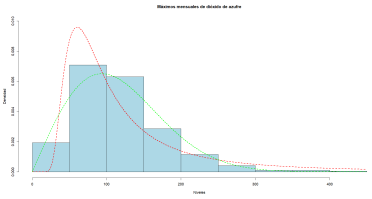


(a) CE **(b)** NE

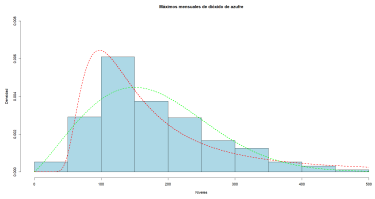
(c) NO **(d)** SE

(e) SO

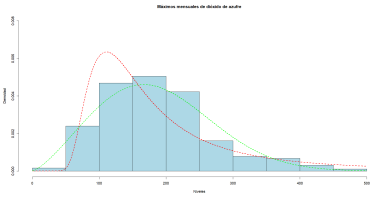
Figura B.5: *Comparación de las densidades obtenidas para los máximos mensuales de NO_2. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*



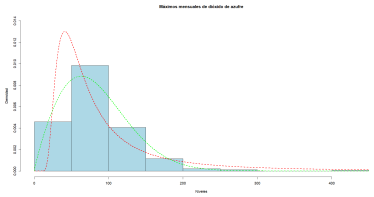
(a) CE



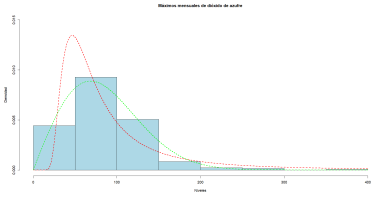
(b) NE



(c) NO

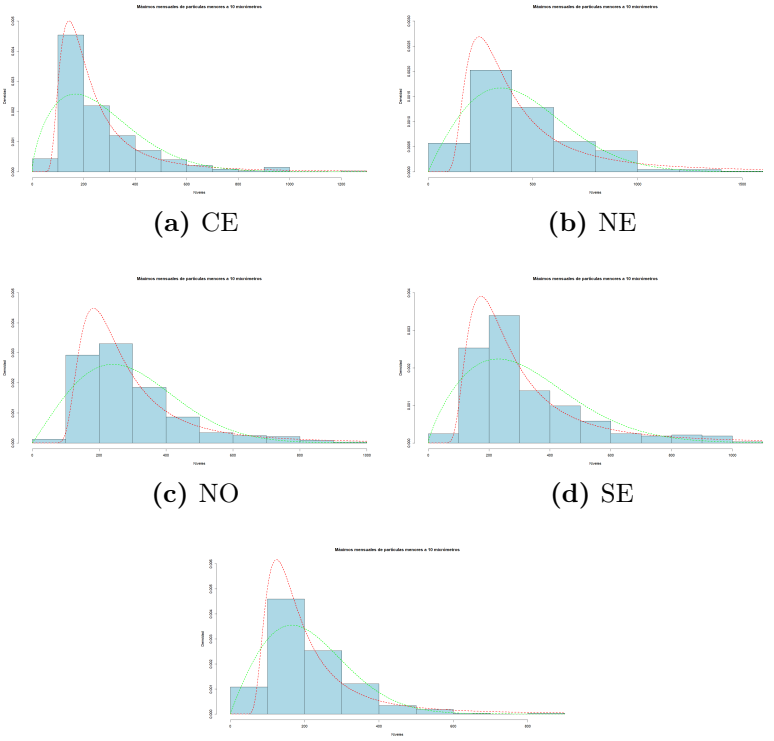


(d) SE



(e) SO

Figura B.6: *Comparación de las densidades obtenidas para los máximos mensuales de SO_2. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*

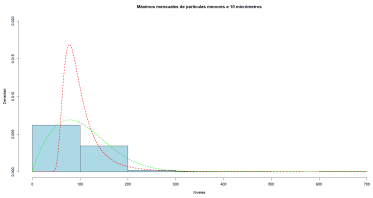


(a) CE **(b)** NE

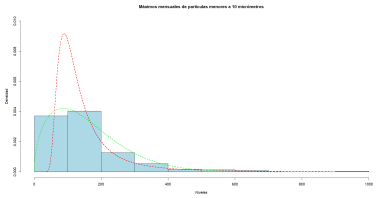
(c) NO **(d)** SE

(e) SO

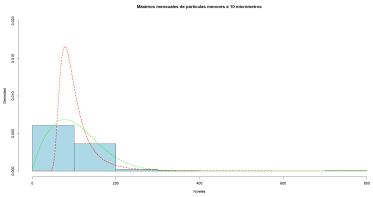
Figura B.7: *Comparación de las densidades obtenidas para los máximos mensuales de PM_{10}. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*



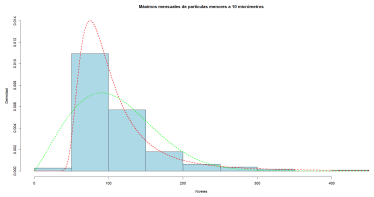
(a) CE



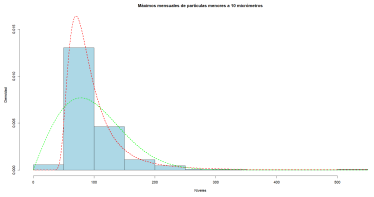
(b) NE



(c) NO

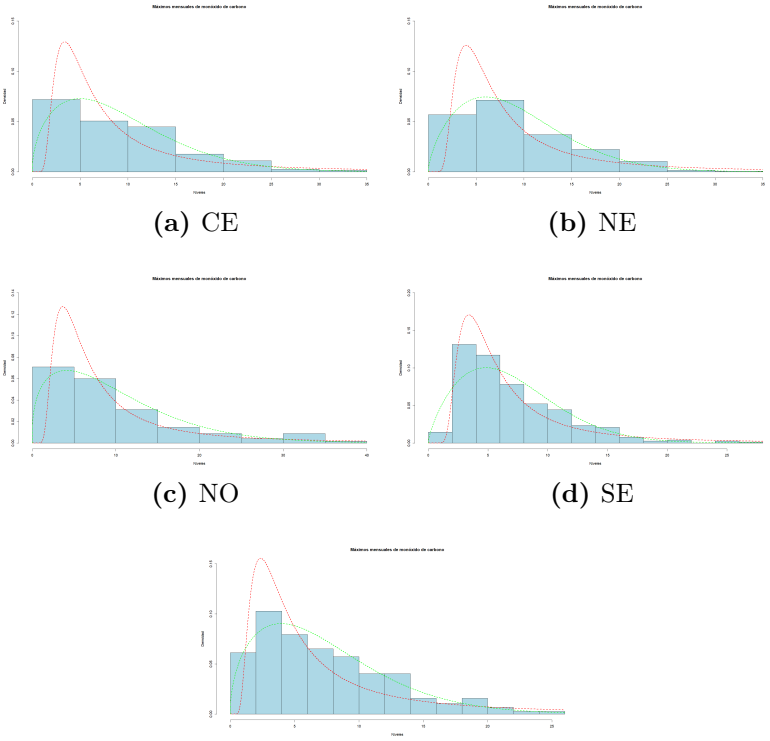


(d) SE



(e) SO

Figura B.8: *Comparación de las densidades obtenidas para los máximos mensuales de $PM_{2.5}$. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*

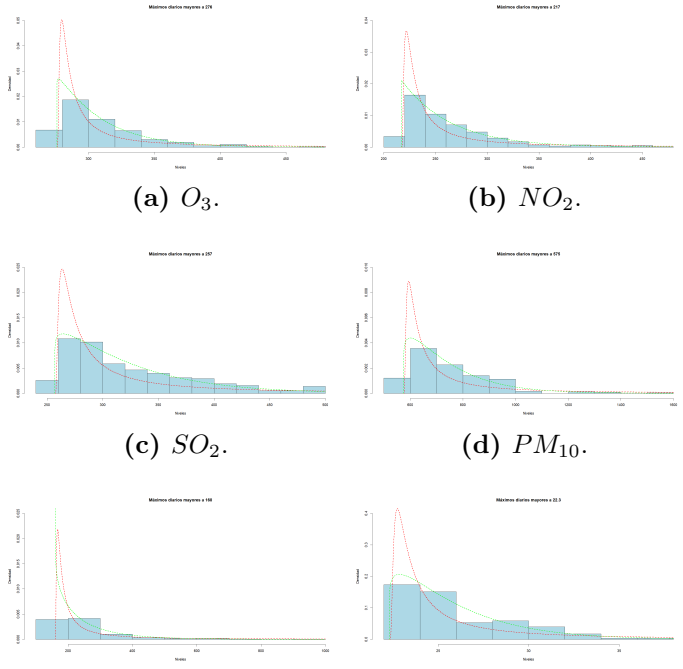


(a) CE **(b)** NE

(c) NO **(d)** SE

(e) SO

Figura B.9: *Comparación de las densidades obtenidas para los máximos mensuales de CO. La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*



(a) O_3. **(b)** NO_2.

(c) SO_2. **(d)** PM_{10}.

(e) $PM_{2.5}$. **(f)** CO.

Figura B.10: *Comparación de las densidades obtenidas para los modelos (B.1) y (B.2). La linea roja representa el ajuste Fréchet y la verde el ajuste Weibull.*

Índice de Tablas

Tabla B.1: *Número y_i de fallas de la bomba en t_i miles de horas.*

Sistema (i)	y_i	t_i	$r_i = y_i t_i$
1	5	94.320	0.05301103
2	1	15.720	0.06361323
3	5	62.880	0.07951654
4	14	125.760	0.11132316
5	3	5.240	0.57251908
6	19	31.440	0.60432570
7	1	1.048	0.95419847
8	1	1.048	0.95419847
9	4	2.096	1.90839695
10	22	10.480	2.09923664

Tabla B.2: *Resultado de usar el muestrador de Gibbs en el ejemplo de la planta nuclear Farley 1.*

Sistema (i)	Media	s
1	0.05989735	0.02507362
2	0.10256774	0.07870071
3	0.08914099	0.03705638
4	0.11561089	0.03005465
5	0.60907691	0.31762446
6	0.60667401	0.13747133
7	0.89859718	0.72278527
8	0.89560013	0.71677873
9	1.58454532	0.75715678
10	1.99107727	0.42022407

Tabla B.3: *Estadísticas de los máximos mensuales. Las estadísticas están dadas en ppb para los casos de O_3, NO_2 y SO_2; en $\mu g/m^3$ para PM_{10} y $PM_{2.5}$ y en ppm para CO.*

	CE		NE		NO		SE		SO	
	Media	d. e.	Media	d. e.	Media	d. e.	Media	d. e.	Media	d. e.
O_3	184.92	59.96	153.06	38.29	187.45	64.06	176.16	48.99	205.34	69.29
NO_2	141.94	57.70	112.10	43.24	140.57	69.63	127.34	51.14	121.22	54.60
SO_2	116.94	61.55	177.97	87.09	186.40	82.39	83.21	46.25	85.68	44.93
PM_{10}	265.35	177.69	447.34	242.23	295.27	150.86	318.99	189.79	211.33	113.75
$PM_{2.5}$	105.79	60.84	159.10	120.74	106.42	61.34	108.73	52.91	98.90	51.34
CO	9.22	6.43	9.24	5.91	9.86	7.95	6.99	4.32	7.43	5.14

Tabla B.4: *Valores de a para las diversas zonas y contaminantes.*

	CE	NE	NO	SE	SO
O_3	200	200	200	200	200
NO_2	150	150	150	150	150
SO_2	100	150	200	100	100
PM_{10}	200	350	250	250	200
$PM_{2.5}$	100	150	100	100	100
CO	10	10	10	10	10

Tabla B.5: *Estimaciones de los parámetros de forma y escala suponiendo distribuciones Fréchet(F) y Weibull(W).*

			CE			NE			NO			SE			SO		
			Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
O_3	F	α	3.746E0	1.511E-1	1.769E-3	4.749E0	1.864E-1	2.011E-3	3.451E0	1.370E-1	1.489E-3	4.454E0	1.813E-1	2.047E-3	3.655E0	1.462E-1	1.766E-3
		s	1.513E2	2.174E0	2.304E-2	1.324E2	1.505E0	1.497E-2	1.504E2	2.361E0	2.735E-2	1.492E2	1.801E0	2.136E-2	1.667E2	2.496E0	2.745E-2
	W	α	3.172E0	1.230E-1	5.992E-3	3.960E0	1.370E-1	5.978E-3	3.071E0	1.265E-1	6.117E-3	3.702E0	1.379E-1	5.917E-3	3.052E0	1.140E-1	5.204E-3
		k	2.055E2	3.517E0	6.105E-2	1.675E2	2.287E0	3.435E-2	2.090E2	3.687E0	6.376E-2	1.941E2	2.832E0	4.104E-2	2.288E2	3.996E0	6.424E-2
NO_2	F	α	2.660E0	9.997E-2	1.078E-3	3.093E0	1.170E-1	1.260E-3	2.625E0	1.021E-1	1.133E-3	2.736E0	1.016E-1	1.240E-3	2.518E0	9.342E-2	1.100E-3
		s	1.080E2	2.202E0	2.407E-2	8.874E1	1.567E0	1.652E-2	1.031E2	2.117E0	2.416E-2	9.769E1	1.933E0	2.195E-2	8.980E1	1.960E0	2.069E-2
	W	α	2.553E0	9.753E-2	3.872E-3	2.570E0	8.824E-2	3.166E-3	2.111E0	7.920E-2	3.822E-3	2.790E0	1.160E-1	4.885E-3	2.326E0	8.273E-2	4.087E-3
		k	1.595E2	3.389E0	5.283E-2	1.254E2	2.648E0	3.449E-2	1.587E2	4.077E0	7.897E-2	1.526E2	3.217E0	5.157E-2	1.368E2	3.137E0	6.411E-2
SO_2	F	α	1.771E0	6.352E-2	6.442E-4	1.893E0	6.809E-2	7.943E-4	2.087E0	7.423E-2	8.304E-4	1.660E0	5.718E-2	7.102E-4	1.888E0	6.760E-2	6.649E-4
		s	7.784E1	2.376E0	2.643E-2	1.219E2	3.485E0	4.040E-2	1.338E2	3.460E0	3.684E-2	5.471E1	1.791E0	1.918E-2	5.829E1	1.663E0	1.834E-2
	W	α	2.006E0	7.150E-2	3.023E-3	2.143E0	8.496E-2	3.174E-3	2.363E0	8.766E-2	3.461E-3	1.893E0	6.523E-2	2.276E-3	2.003E0	7.209E-2	2.939E-3
		k	1.322E2	3.537E0	6.067E-2	2.008E2	5.082E0	7.431E-2	2.100E2	4.757E0	6.845E-2	9.378E1	2.650E0	4.190E-2	9.673E1	2.610E0	4.299E-2
PM_{10}	F	α	2.117E0	9.078E-2	1.020E-3	1.956E0	7.943E-2	8.717E-4	2.390E0	9.782E-2	1.027E-3	2.018E0	8.340E-2	9.016E-4	2.253E0	9.526E-2	1.153E-3
		s	1.717E2	4.792E0	5.284E-2	3.001E2	9.165E0	1.043E-1	2.120E2	5.193E0	6.241E-2	2.114E2	6.200E0	7.091E-2	1.466E2	3.841E0	4.321E-2
	W	α	1.640E0	6.211E-2	2.402E-3	1.937E0	7.868E-2	3.918E-3	2.052E0	7.882E-2	2.795E-3	1.800E0	7.238E-2	3.313E-3	1.958E0	7.682E-2	3.324E-3
		k	2.985E2	1.059E1	1.845E-1	5.043E2	1.527E1	2.937E-1	3.333E2	9.483E0	1.305E-1	3.599E2	1.178E1	2.304E-1	2.385E2	7.102E0	1.284E-1
$PM_{2.5}$	F	α	3.676E0	1.922E-1	1.930E-3	2.582E0	1.291E-1	1.350E-3	3.649E0	1.894E-1	2.135E-3	2.996E0	1.538E-1	1.753E-3	3.261E0	1.684E-1	1.814E-3
		s	8.306E1	1.616E0	1.672E-2	1.034E2	3.072E0	3.708E-2	8.385E1	1.620E0	1.667E-2	8.267E1	1.985E0	2.301E-2	7.639E1	1.664E0	1.798E-2
	W	α	1.852E0	7.413E-2	2.343E-3	1.513E0	6.940E-2	2.433E-3	1.841E0	7.211E-2	2.174E-3	2.119E0	9.340E-2	3.963E-3	1.987E0	8.300E-2	2.884E-3
		k	1.188E2	4.618E0	6.637E-2	1.779E2	8.370E0	1.363E-1	1.192E2	4.648E0	6.605E-2	1.225E2	4.028E0	7.297E-2	1.112E2	4.055E0	6.858E-2
CO	F	α	1.415E0	5.336E-2	6.083E-4	1.566E0	5.857E-2	6.732E-4	1.480E0	5.589E-2	5.339E-4	1.774E0	6.699E-2	7.824E-4	1.239E0	4.569E-2	5.535E-4
		s	4.945E0	1.878E-1	2.138E-3	5.428E0	1.881E-1	2.133E-3	5.177E0	1.908E-1	2.108E-3	4.378E0	1.336E-1	1.434E-3	3.809E0	1.665E-1	1.937E-3
	W	α	1.509E0	5.777E-2	9.065E-4	1.663E0	6.499E-2	1.309E-3	1.363E0	5.103E-2	7.370E-4	1.740E0	6.493E-2	1.061E-3	1.498E0	5.961E-2	9.088E-4
		k	1.026E1	3.646E-1	4.420E-3	1.039E1	3.348E-1	4.291E-3	1.086E1	4.327E-1	4.802E-3	7.888E0	2.478E-1	3.113E-3	8.256E0	2.958E-1	3.561E-3

Tabla B.6: *Comparación de los DIC y MLF en los modelos individuales.*

		CE		NE		NO		SE		SO	
		DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF
O_3	F	7.680E7	4.574E-904	7.680E7	1.301E-833	7.680E7	1.230E-918	7.680E7	5.561E-871	7.680E7	1.619E-925
	W	4.218E3	1.181E-919	3.863E3	2.777E-851	4.240E3	6.191E-929	4.032E3	1.803E-888	4.326E3	6.217E-943
NO_2	F	7.680E7	5.194E-905	7.680E7	8.548E-844	7.680E7	1.428E-906	7.680E7	1.227E-882	7.680E7	5.136E-886
	W	4.176E3	7.331E-907	3.964E3	5.338E-862	4.278E3	1.401E-929	3.435E3	5.191E-890	4.116E3	1.765E-894
SO_2	F	7.680E7	1.501E-924	7.680E7	5.907E-987	7.680E7	8.133E-982	7.680E7	5.333E-875	7.680E7	9.402E-865
	W	4.186E3	1.662E-909	4.466E3	3.585E-971	4.437E3	1.640E-965	3.946E3	2.604E-857	3.939E3	6.793E-856
PM_{10}	F	6.480E7	1.341E-876	6.480E7	2.847E-959	6.480E7	1.031E-998	6.480E7	1.303E-907	6.480E7	1.490E-839
	W	4.142E3	4.869E-900	4.416E3	2.177E-958	4.105E3	2.864E-1012	4.223E3	8.041E-918	3.914E3	8.928E-851
$PM_{2.5}$	F	4.420E7	2.843E-467	4.420E7	2.067E-540	4.420E7	2.132E-468	4.420E7	1.191E-486	4.420E7	8.206E-471
	W	2.358E3	1.046E-512	2.618E3	5.101E-569	2.360E3	7.079E-513	2.338E3	1.011E-508	2.311E3	2.413E-502
CO	F	7.680E7	8.897E-526	7.680E7	1.401E-515	7.680E7	5.095E-527	7.680E7	2.119E-454	7.680E7	1.733E-505
	W	2.385E3	3.962E-518	2.343E3	5.920E-509	2.471E3	9.975E-537	2.101E3	2.712E-456	2.225E3	2.309E-483

Tabla B.7: *Elección del modelo para el análisis univariado.*

	CE	NE	NO	SE	SO
O_3	F	F	F	F	F
NO_2	F	F	F	F	F
SO_2	W	W	W	W	W
PM_{10}	F	W	F	F	F
$PM_{2.5}$	F	F	F	F	F
CO	W	W	F	F	W

Tabla B.8: *Calculo de la ρ de Spearman muestral de los máximos mensuales del O_3 con los otros contaminantes.*

	CE	NE	NO	SE	SO
NO_2 y O_3	0.64041070	0.5411453	0.5605169	0.6414327	0.57207730
SO_2 y O_3	0.41057210	0.3512676	0.2412945	0.3948958	0.38476550
PM_{10} y O_3	0.45670790	0.5199240	0.3104575	0.5381821	0.41344370
$PM_{2.5}$ y O_3	0.06109886	0.4185697	0.3175970	0.2173969	0.05371831
CO y O_3	0.82883330	0.7206002	0.8598176	0.7677488	0.82494090

Tabla B.9: *Estimaciones de los parámetros suponiendo un modelo Fréchet(F) o Weibull(W) en el comportamiento del $NO_2(y)$ y $O_3(z)$ relacionadas a través de la cópula de Gumbel-Hougaard.*

		CE			NE			NO			SE			SO		
		Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
F-F	α_y	2.664E0	6.658E-2	7.412E-4	3.084E0	7.945E-2	9.845E-4	2.605E0	6.885E-2	8.620E-4	2.756E0	6.803E-2	8.991E-4	2.515E0	6.498E-2	7.411E-4
	α_z	3.664E0	9.984E-2	1.262E-3	4.682E0	1.247E-1	1.507E-3	3.451E0	9.351E-2	1.097E-3	4.338E0	1.176E-1	1.480E-3	3.597E0	9.737E-2	1.192E-3
	s_y	1.078E2	1.425E0	1.764E-2	8.882E1	1.044E0	1.301E-2	1.033E2	1.426E0	1.711E-2	9.747E1	1.269E0	1.523E-2	8.967E1	1.299E0	1.448E-2
	s_z	1.514E2	1.470E0	1.855E-2	1.325E2	1.027E0	1.276E-2	1.502E2	1.576E0	1.796E-2	1.495E2	1.234E0	1.567E-2	1.669E2	1.665E0	1.872E-2
	θ	1.884E0	8.319E-2	9.824E-4	1.700E0	7.555E-2	9.864E-4	1.685E0	7.230E-2	8.553E-4	1.849E0	7.930E-2	9.662E-4	1.726E0	7.430E-2	9.564E-4
W-W	α_y	2.491E0	6.538E-2	1.446E-3	2.530E0	6.229E-2	1.651E-3	2.078E0	5.181E-2	1.711E-3	2.529E0	6.513E-2	1.560E-3	2.278E0	6.148E-2	2.127E-3
	α_z	3.114E0	8.450E-2	2.586E-3	3.901E0	1.048E-1	4.679E-3	3.041E0	8.315E-2	3.429E-3	3.658E0	9.549E-2	2.950E-3	3.028E0	8.019E-2	3.564E-3
	k_y	1.609E2	2.355E0	2.894E-2	1.259E2	1.826E0	2.947E-2	1.601E2	2.807E0	5.236E-2	1.441E2	2.113E0	2.714E-2	1.381E2	2.229E0	4.310E-2
	k_z	2.059E2	2.403E0	3.636E-2	1.680E2	1.579E0	3.587E-2	2.096E2	2.512E0	5.026E-2	1.948E2	1.936E0	2.744E-2	2.289E2	2.727E0	6.693E-2
	θ	1.635E0	8.098E-2	1.179E-3	1.412E0	7.299E-2	1.444E-3	1.468E0	7.027E-2	1.548E-3	1.554E0	7.609E-2	1.054E-3	1.496E0	7.324E-2	1.531E-3
W-F	α_y	2.467E0	6.253E-2	2.097E-3	2.502E0	6.268E-2	1.943E-3	2.058E0	5.184E-2	1.260E-3	2.507E0	6.370E-2	2.211E-3	2.253E0	5.901E-2	1.682E-3
	α_z	3.721E0	1.017E-1	1.573E-3	4.689E0	1.299E-1	2.186E-3	3.426E0	9.352E-2	1.464E-3	4.414E0	1.229E-1	1.737E-3	3.652E0	1.003E-1	1.631E-3
	k_y	1.604E2	2.318E0	4.345E-2	1.260E2	1.807E0	3.568E-2	1.600E2	2.830E0	4.852E-2	1.439E2	2.087E0	3.961E-2	1.379E2	2.220E0	3.653E-2
	s_z	1.513E2	1.469E0	2.217E-2	1.324E2	1.055E0	1.647E-2	1.504E2	1.606E0	2.469E-2	1.492E2	1.217E0	1.858E-2	1.665E2	1.640E0	2.427E-2
	θ	1.712E0	7.897E-2	1.271E-3	1.505E0	7.160E-2	1.320E-3	1.531E0	7.088E-2	1.071E-3	1.648E0	7.488E-2	1.183E-3	1.557E0	7.019E-2	1.141E-3

Tabla B.10: *Estimaciones de los parámetros suponiendo un modelo Fréchet(F) o Weibull(W) en el comportamiento del $SO_2(y)$ y $O_3(z)$ relacionadas a través de la cópula de Gumbel-Hougaard.*

		CE			NE			NO			SE			SO		
		Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
F-F	α_y	1.788E0	4.294E-2	5.713E-4	1.895E0	4.659E-2	7.352E-4	2.088E0	5.275E-2	9.291E-4	1.675E0	3.988E-2	4.740E-4	1.903E0	4.801E-2	5.945E-4
	α_z	3.613E0	9.999E-2	1.398E-3	4.621E0	1.267E-1	1.992E-3	3.376E0	9.598E-2	1.440E-3	4.253E0	1.228E-1	1.508E-3	3.530E0	1.001E-1	1.219E-3
	s_y	7.703E1	1.583E0	1.972E-2	1.214E2	2.386E0	3.310E-2	1.335E2	2.416E0	3.485E-2	5.415E1	1.223E0	1.326E-2	5.778E1	1.141E0	1.217E-2
	s_z	1.523E2	1.553E0	1.970E-2	1.328E2	1.052E0	1.598E-2	1.512E2	1.668E0	2.678E-2	1.503E2	1.296E0	1.574E-2	1.678E2	1.770E0	2.000E-2
	θ	1.508E0	6.626E-2	8.554E-4	1.420E0	6.392E-2	9.867E-4	1.297E0	5.768E-2	8.986E-4	1.444E0	6.490E-2	7.415E-4	1.415E0	6.149E-2	6.400E-4
W-W	α_y	1.994E0	5.060E-2	1.222E-3	2.151E0	5.825E-2	1.757E-3	2.367E0	6.211E-2	2.152E-3	1.883E0	4.730E-2	1.167E-3	1.996E0	5.240E-2	1.076E-3
	α_z	3.183E0	8.529E-2	3.532E-3	3.978E0	1.050E-1	4.762E-3	3.095E0	8.371E-2	3.465E-3	3.713E0	1.012E-1	4.916E-3	3.058E0	8.199E-2	2.870E-3
	k_y	1.323E2	2.397E0	4.024E-2	2.016E2	3.528E0	6.717E-2	2.105E2	3.374E0	6.684E-2	9.393E1	1.910E0	3.715E-2	9.657E1	1.835E0	2.997E-2
	k_z	2.064E2	2.479E0	5.385E-2	1.681E2	1.624E0	3.485E-2	2.096E2	2.607E0	4.873E-2	1.952E2	2.005E0	4.972E-2	2.304E2	2.832E0	5.154E-2
	θ	1.302E0	5.640E-2	9.703E-4	1.162E0	4.773E-2	8.561E-4	1.056E0	3.461E-2	5.623E-4	1.177E0	5.151E-2	9.383E-4	1.203E0	5.065E-2	8.475E-4
W-F	α_y	1.967E0	5.528E-2	2.738E-3	2.129E0	5.769E-2	1.673E-3	2.356E0	6.380E-2	1.760E-3	1.871E0	4.518E-2	1.659E-3	1.974E0	5.341E-2	2.193E-3
	α_z	3.738E0	1.042E-1	8.713E-4	4.734E0	1.301E-1	1.955E-3	3.440E0	9.796E-2	1.692E-3	4.392E0	1.267E-1	1.079E-3	3.616E0	1.026E-1	8.011E-4
	k_y	1.327E2	2.526E0	5.655E-2	2.022E2	3.627E0	7.160E-2	2.110E2	3.379E0	5.926E-2	9.387E1	1.871E0	3.098E-2	9.674E1	1.868E0	3.612E-2
	s_z	1.515E2	1.512E0	1.402E-2	1.325E2	1.053E0	1.859E-2	1.506E2	1.675E0	2.878E-2	1.497E2	1.287E0	1.048E-2	1.672E2	1.743E0	1.440E-2
	θ	1.352E0	5.682E-2	7.642E-4	1.257E0	5.200E-2	7.943E-4	1.126E0	4.574E-2	7.073E-4	1.243E0	5.272E-2	5.105E-4	1.260E0	5.336E-2	5.604E-4

Tabla B.11: *Estimaciones de los parámetros suponiendo un modelo Fréchet(F) o Weibull(W) en el comportamiento del $PM_{10}(y)$ y $O_3(z)$ relacionadas a través de la cópula de Gumbel-Hougaard.*

		CE			NE			NO			SE			SO		
		Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
F-F	α_y	2.105E0	6.155E-2	7.710E-4	1.967E0	5.416E-2	6.143E-4	2.388E0	6.745E-2	1.062E-3	2.039E0	5.760E-2	9.369E-4	2.251E0	6.563E-2	7.500E-4
	α_z	4.396E0	1.309E-1	1.566E-3	5.164E0	1.528E-1	1.701E-3	3.994E0	1.210E-1	1.800E-3	5.100E0	1.555E-1	2.538E-3	4.313E0	1.310E-1	1.442E-3
	s_y	1.718E2	3.250E0	4.189E-2	2.973E2	6.086E0	7.329E-2	2.120E2	3.635E0	4.392E-2	2.104E2	4.122E0	5.937E-2	1.464E2	2.624E0	3.034E-2
	s_z	1.432E2	1.302E0	1.581E-2	1.291E2	9.792E-1	1.036E-2	1.419E2	1.445E0	2.189E-2	1.430E2	1.112E0	1.610E-2	1.576E2	1.469E0	1.588E-2
	θ	1.490E0	6.711E-2	8.676E-4	1.608E0	7.677E-2	8.936E-4	1.299E0	5.705E-2	7.847E-4	1.567E0	7.181E-2	9.376E-4	1.413E0	6.268E-2	7.222E-4
W-W	α_y	1.621E0	4.451E-2	1.084E-3	1.930E0	5.725E-2	1.251E-3	2.050E0	5.746E-2	1.588E-3	1.775E0	5.167E-2	1.364E-3	1.952E0	5.508E-2	1.432E-3
	α_z	3.560E0	1.045E-1	5.175E-3	4.070E0	1.105E-1	3.260E-3	3.348E0	9.990E-2	3.794E-3	3.851E0	1.120E-1	5.446E-3	3.592E0	1.049E-1	4.550E-3
	k_y	3.013E2	7.653E0	1.344E-1	5.064E2	1.066E1	1.444E-1	3.356E2	6.721E0	1.191E-1	3.630E2	8.338E0	1.452E-1	2.397E2	4.996E0	8.740E-2
	k_z	1.873E2	2.114E0	4.824E-2	1.601E2	1.608E0	2.269E-2	1.898E2	2.367E0	4.242E-2	1.809E2	1.901E0	4.501E-2	2.060E2	2.336E0	4.794E-2
	θ	1.369E0	6.917E-2	1.326E-3	1.346E0	6.861E-2	8.877E-4	1.193E0	4.986E-2	8.170E-4	1.340E0	6.761E-2	1.273E-3	1.213E0	5.750E-2	9.270E-4
W-F	α_y	1.614E0	4.464E-2	1.986E-3	1.910E0	5.546E-2	1.568E-3	2.041E0	5.786E-2	1.695E-3	1.765E0	5.017E-2	1.178E-3	1.939E0	5.651E-2	1.422E-3
	α_z	4.387E0	1.352E-1	1.157E-3	5.275E0	1.578E-1	2.514E-3	4.044E0	1.245E-1	1.669E-3	5.186E0	1.608E-1	2.170E-3	4.335E0	1.308E-1	2.017E-3
	k_y	3.012E2	7.553E0	1.478E-1	5.062E2	1.047E1	1.796E-1	3.358E2	6.738E0	1.182E-1	3.617E2	8.185E0	1.339E-1	2.395E2	5.049E0	8.480E-2
	s_z	1.433E2	1.325E0	1.190E-2	1.288E2	9.610E-1	1.424E-2	1.418E2	1.447E0	1.942E-2	1.428E2	1.099E0	1.615E-2	1.573E2	1.466E0	1.999E-2
	θ	1.391E0	6.621E-2	7.892E-4	1.434E0	6.918E-2	1.171E-3	1.217E0	5.170E-2	7.197E-4	1.393E0	6.514E-2	9.876E-4	1.290E0	6.044E-2	9.304E-4

Tabla B.12: *Estimaciones de los parámetros suponiendo un modelo Fréchet(F) o Weibull(W) en el comportamiento del $PM_{2.5}(y)$ y $O_3(z)$ relacionadas a través de la cópula de Gumbel-Hougaard.*

		CE			NE			NO			SE			SO		
		Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
F-F	α_y	3.670E0	1.354E-1	2.029E-3	2.322E0	8.922E-2	1.409E-3	3.617E0	1.328E-1	1.621E-3	2.979E0	1.071E-1	1.671E-3	3.250E0	1.213E-1	1.577E-3
	α_z	6.903E0	2.467E-1	4.308E-3	7.123E0	2.444E-1	3.500E-3	6.312E0	2.282E-1	2.922E-3	8.554E0	3.047E-1	3.650E-3	6.781E0	2.392E-1	3.016E-3
	s_y	8.305E1	1.128E0	1.595E-2	1.038E2	2.193E0	2.899E-2	8.395E1	1.152E0	1.502E-2	8.272E1	1.399E0	1.818E-2	7.641E1	1.171E0	1.355E-2
	s_z	1.296E2	9.494E-1	1.274E-2	1.202E2	8.362E-1	1.069E-2	1.274E2	1.002E0	1.249E-2	1.311E2	7.641E-1	8.391E-3	1.421E2	1.058E0	1.258E-2
	θ	1.052E0	3.766E-2	4.916E-4	1.363E0	7.417E-2	9.820E-4	1.228E0	6.573E-2	8.123E-4	1.153E0	6.095E-2	6.975E-4	1.059E0	4.153E-2	4.615E-4
W-W	α_y	1.864E0	5.299E-2	7.138E-4	1.504E0	4.871E-2	6.126E-4	1.845E0	4.839E-2	6.191E-4	2.125E0	6.840E-2	1.219E-3	2.000E0	5.938E-2	6.849E-4
	α_z	6.290E0	2.112E-1	8.769E-3	6.588E0	2.301E-1	9.130E-3	5.618E0	2.073E-1	8.265E-3	6.770E0	2.398E-1	1.163E-2	6.396E0	2.320E-1	9.602E-3
	k_y	1.119E2	3.252E0	3.247E-2	1.788E2	5.895E0	6.070E-2	1.198E2	3.255E0	3.561E-2	1.229E2	2.934E0	3.926E-2	1.114E2	2.807E0	2.760E-2
	k_z	1.506E2	1.194E0	2.057E-2	1.393E2	1.052E0	1.461E-2	1.511E2	1.367E0	2.249E-2	1.489E2	1.106E0	2.294E-2	1.652E2	1.300E0	2.044E-2
	θ	1.024E0	2.207E-2	2.122E-4	1.210E0	6.856E-2	7.479E-4	1.056E0	4.249E-2	4.776E-4	1.041E0	3.193E-2	3.630E-4	1.026E0	2.200E-2	1.956E-4
F-W	α_y	3.676E0	1.363E-1	1.509E-3	—	—	—	3.636E0	1.347E-1	1.439E-3	—	—	—	3.263E0	1.200E-1	1.289E-3
	α_z	6.288E0	2.148E-1	9.174E-3	—	—	—	5.607E0	1.975E-1	8.462E-3	—	—	—	6.396E0	2.209E-1	9.969E-3
	k_z	1.506E2	1.191E0	2.054E-2	—	—	—	1.513E2	1.337E0	2.229E-2	—	—	—	1.652E2	1.289E0	2.127E-2
	s_y	8.302E1	1.118E0	1.185E-2	—	—	—	8.389E1	1.151E0	1.211E-2	—	—	—	7.634E1	1.183E0	1.265E-2
	θ	1.041E0	3.069E-2	3.612E-4	—	—	—	1.150E0	5.885E-2	7.247E-4	—	—	—	1.042E0	3.107E-2	3.629E-4
W-F	α_y	1.860E0	5.271E-2	1.052E-3	1.488E0	4.910E-2	9.393E-4	1.837E0	4.900E-2	9.813E-4	2.123E0	6.904E-2	1.808E-3	1.997E0	6.172E-2	2.488E-3
	α_z	6.899E0	2.460E-1	3.602E-3	7.040E0	2.444E-1	3.639E-3	6.308E0	2.298E-1	3.432E-3	8.577E0	3.080E-1	4.695E-3	6.785E0	2.338E-1	1.941E-3
	k_y	1.192E2	3.242E0	4.753E-2	1.802E2	6.089E0	9.087E-2	1.197E2	3.298E0	5.288E-2	1.230E2	2.982E0	6.060E-2	1.116E2	2.837E0	5.332E-2
	s_z	1.295E2	9.306E-1	1.344E-2	1.203E2	8.394E-1	1.130E-2	1.273E2	1.014E0	1.489E-2	1.310E2	7.864E-1	1.334E-2	1.421E2	1.050E0	8.336E-3
	θ	1.031E0	2.708E-2	3.978E-4	1.258E0	7.708E-2	1.154E-3	1.095E0	6.074E-2	8.715E-4	1.060E0	4.247E-2	7.290E-4	1.032E0	2.816E-2	2.691E-4

Tabla B.13: *Estimaciones de los parámetros suponiendo un modelo Fréchet(F) o Weibull(W) en el comportamiento del $CO(y)$ y $O_3(z)$ relacionadas a través de la cópula de Gumbel-Hougaard.*

		CE			NE			NO			SE			SO		
		Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.	Media	d. e.	Error M.C.
F-F	α_y	1.443E0	3.530E-2	3.913E-4	1.583E0	3.898E-2	4.487E-4	1.489E0	3.651E-2	4.314E-4	1.801E0	4.399E-2	5.195E-4	1.276E0	3.023E-2	3.532E-4
	α_z	3.582E0	9.157E-2	1.080E-3	4.580E0	1.200E-1	1.387E-3	3.372E0	8.587E-2	1.002E-3	4.288E0	1.155E-1	1.402E-3	3.421E0	9.052E-2	1.048E-3
	s_y	4.889E0	1.161E-1	1.406E-3	5.349E0	1.198E-1	1.219E-3	5.145E0	1.163E-1	1.389E-3	4.345E0	8.388E-2	1.018E-3	3.733E0	1.022E-1	1.183E-3
	s_z	1.516E2	1.435E0	1.548E-2	1.329E2	9.954E-1	1.124E-2	1.506E2	5.190E-1	1.719E-2	1.496E2	1.218E0	1.426E-2	1.673E2	1.693E0	1.907E-2
	θ	2.698E0	1.247E-1	1.303E-3	2.142E0	9.707E-2	1.136E-3	2.865E0	1.307E-1	1.459E-3	2.259E0	1.015E-1	1.097E-3	2.548E0	1.187E-1	1.360E-3
W-W	α_y	1.459E0	3.817E-2	1.108E-3	1.608E0	4.386E-2	6.573E-4	1.287E0	3.222E-2	5.794E-4	1.664E0	4.341E-2	8.468E-4	1.450E0	3.849E-2	9.313E-4
	α_z	3.065E0	7.752E-2	3.863E-3	3.880E0	9.691E-2	3.333E-3	2.996E0	7.696E-2	2.555E-3	3.659E0	9.855E-2	4.522E-3	2.976E0	7.274E-2	3.230E-3
	k_y	1.031E1	2.404E-1	4.390E-3	1.039E1	2.255E-1	2.440E-3	1.096E1	2.884E-1	3.051E-3	7.985E0	1.727E-1	2.994E-3	8.323E0	1.966E-1	3.280E-3
	k_z	2.066E2	2.314E0	5.629E-2	1.687E2	1.524E0	2.168E-2	2.103E2	2.382E0	3.289E-2	1.947E2	1.887E0	4.433E-2	2.294E2	2.643E0	5.666E-2
	θ	2.610E0	1.340E-1	2.615E-3	1.874E0	9.563E-2	1.388E-3	2.644E0	1.412E-1	2.237E-3	1.979E0	1.019E-1	1.706E-3	2.396E0	1.216E-1	2.346E-3
W-F	α_y	1.415E0	3.643E-2	1.086E-3	1.586E0	4.191E-2	1.204E-3	1.255E0	3.197E-2	8.575E-4	1.652E0	4.218E-2	1.057E-3	1.420E0	3.743E-2	1.045E-3
	α_z	3.694E0	9.814E-2	1.201E-3	4.697E0	1.223E-1	1.480E-3	3.375E0	9.091E-2	1.052E-3	4.386E0	1.224E-1	1.332E-3	3.633E0	9.551E-2	1.016E-3
	k_y	1.026E1	2.483E-1	4.313E-3	1.037E1	2.311E-1	3.933E-3	1.093E1	2.992E-1	5.012E-3	7.928E0	1.705E-1	2.640E-3	8.267E0	2.000E-1	3.547E-3
	s_z	1.517E2	1.402E0	2.003E-2	1.328E2	1.004E0	1.258E-2	1.507E2	1.514E0	2.176E-2	1.494E2	1.200E0	1.482E-2	1.670E2	1.569E0	2.297E-2
	θ	2.582E0	1.266E-1	1.798E-3	1.952E0	9.143E-2	1.163E-3	2.433E0	1.215E-1	1.647E-3	1.986E0	9.404E-2	1.217E-3	2.474E0	1.183E-1	1.588E-3

Tabla B.14: *Comparación de los DIC y MLF en los modelos bivariados.*

		CE		NE		NO		SE		SO	
		DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF
NO_2	F-F	2.304E8	2.653E-1763	2.304E8	1.789E-1644	2.304E8	1.234E-1789	2.304E8	2.008E-1708	2.304E8	1.036E-1774
y	W-W	7.682E7	3.577E-1799	7.682E7	5.905E-1699	7.682E7	5.715E-1839	7.682E7	1.370E-1751	7.682E7	1.452E-1815
O_3	W-F	1.536E8	9.894E-1777	1.536E8	4.306E-1676	1.536E8	3.250E-1825	1.536E8	5.461E-1727	1.536E8	2.179E-1793
SO_2	F-F	2.304E8	8.297E-1809	2.304E8	6.875E-1807	2.304E8	2.249E-1892	2.304E8	8.526E-1731	2.304E8	9.005E-1776
y	W-W	7.862E7	3.539E-1816	7.862E7	7.566E-1818	7.862E7	2.809E-1893	7.862E7	1.730E-1741	7.862E7	2.216E-1998
O_3	W-F	1.536E8	2.032E-1799	1.536E8	1.529E-1796	1.536E8	1.823E-1881	1.536E8	2.756E-1720	1.536E8	2.256E-1773
PM_{10}	F-F	1.944E8	2.546E-1587	1.944E8	3.549E-1619	1.944E8	4.642E-1613	1.944E8	2.015E-1589	1.944E8	1.371E-1569
y	W-W	6.482E7	2.718E-1635	6.482E7	1.281E-1652	6.482E7	7.073E-1646	6.482E7	6.781E-1639	6.481E7	2.051E-1606
O_3	W-F	1.296E8	8.252E-1617	1.296E8	1.777E-1625	1.296E8	3.212E-1629	1.296E8	2.034E-1608	1.296E8	6.632E-1587
$PM_{2.5}$	F-F	1.326E8	2.630E-902	1.326E8	3.134E-938	1.326E8	2.031E-909	1.326E8	1.571E-900	1.326E8	1.787E-915
y	W-W	4.421E7	1.722E-951	4.421E7	1.322E-992	4.421E7	1.032E-962	4.421E7	8.457E-938	4.421E7	6.282E-949
O_3	F-W	8.841E7	6.438E-906	—	—	8.841E7	1.263E-916	—	—	8.841E7	2.642E-917
	W-F	8.841E7	6.724E-948	8.841E7	5.223E-991	8.841E7	7.947E-957	8.841E7	5.124E-924	8.841E7	2.445E-947
CO	F-F	2.304E8	6.927E-1334	2.304E8	1.632E-1288	2.304E8	2.671E-1337	2.304E8	3.422E-1253	2.304E8	2.977E-1346
y	W-W	7.681E7	6.747E-1350	7.681E7	2.248E-1317	7.681E7	5.068E-1377	7.681E7	1.088E-1293	7.681E7	2.620E-1348
O_3	W-F	1.536E8	1.262E-1330	1.536E8	2.552E-1291	1.536E8	1.472E-1375	1.536E8	1.292E-1276	1.536E8	1.291E-1321

Tabla B.15: *Elección del modelo. Caso bivariado.*

	CE	NE	NO	SE	SO
NO_2 y O_3	F-F	F-F	F-F	F-F	F-F
SO_2 y O_3	W-F	W-F	W-F	W-F	W-F
PM_{10} y O_3	F-F	F-F	F-F	F-F	F-F
$PM_{2.5}$ y O_3	F-F	F-F	F-F	F-F	F-F
CO y O_3	W-F	F-F	F-F	F-F	W-F

Tabla B.16: *Valores de $\hat{\rho}$ asociado a los θ estimados en las Tablas B.9-B.13.*

		CE	NE	NO	SE	SO
NO_2	F-F	0.646752	0.577332	0.570804	0.634908	0.588288
y	W-W	0.547968	0.420996	0.457356	0.507036	0.474300
O_3	W-F	0.582444	0.479592	0.494424	0.554076	0.508644
SO_2	F-F	0.481332	0.426396	0.334284	0.442188	0.423024
y	W-W	0.338424	0.206244	0.079200	0.222156	0.248652
O_3	W-F	0.377952	0.299664	0.166104	0.286932	0.302352
PM_{10}	F-F	0.470736	0.534900	0.335940	0.513960	0.421668
y	W-W	0.390612	0.373392	0.238620	0.368784	0.258492
O_3	W-F	0.406428	0.435684	0.262380	0.407844	0.328404
$PM_{2.5}$	F-F	0.073848	0.386184	0.272916	0.200844	0.083184
y	W-W	0.035088	0.255564	0.079200	0.058896	0.037932
O_3	F-W	0.058896	—	0.193176	—	0.060264
	W-F	0.045000	0.300564	0.129144	0.084504	0.046404
NO_2	F-F	0.815724	0.718956	0.835224	0.744708	0.795156
y	W-W	0.804048	0.643428	0.808692	0.676212	0.770664
O_3	W-F	0.800112	0.668822	0.777012	0.678240	0.783744

Tabla B.17: *Valores para el modelo (5.8) para las diversas zonas y diversos contaminantes.*

	a_1	a_2	a_3
O_3	200	200	200
PM_{10}	350	200	200
$PM_{2.5}$	150	100	100

Tabla B.18: *Estadísticas de la muestra final de los parámetros suponiendo un modelo Fréchet en el comportamiento del O_3, PM_{10} y $PM_{2.5}$ en las regiones NE, CE y SO relacionadas a través de la cópula de Gumbel-Hougaard tridimensional.*

	Parámetro	Media	d. e.	Error M. C.
O_3	α_{X_1}	4.720	0.11210	0.0013900
	α_{X_2}	3.759	0.08819	0.0011400
	α_{X_3}	3.649	0.08724	0.0011120
	s_{X_1}	131.300	0.89630	0.0113300
	s_{X_2}	152.600	1.24200	0.0157700
	s_{X_3}	167.000	1.42100	0.0175000
	θ	2.485	0.08377	0.0010520
PM_{10}	α_{X_1}	1.983	0.05066	0.0005869
	α_{X_2}	2.051	0.05685	0.0006937
	α_{X_3}	2.218	0.05964	0.0007035
	s_{X_1}	296.100	5.49300	0.0709800
	s_{X_2}	173.000	3.00700	0.0370900
	s_{X_3}	147.900	2.38600	0.0293100
	θ	2.029	0.06930	0.0009002
$PM_{2.5}$	α_{X_1}	2.243	0.08292	0.0010150
	α_{X_2}	3.610	0.12480	0.0014720
	α_{X_3}	3.343	0.10990	0.0013120
	s_{X_1}	103.400	1.97800	0.0237400
	s_{X_2}	83.400	1.02900	0.0116000
	s_{X_3}	76.320	1.04900	0.0124100
	θ	1.689	0.06839	0.0008252

Tabla B.19: *Estadísticas de la muestra final de los parámetros suponiendo un modelo Fréchet en el comportamiento del O_3, PM_{10} y $PM_{2.5}$ en las regiones NE, CE y SO relacionadas a través de la cópula de Gumbel-Hougaard asimétrica.*

	Parámetro	Media	d. e.	Error M. C.
O_3	α_{X_1}	3.764	0.07232	0.0016930
	α_{X_2}	3.644	0.06961	0.0015640
	α_{X_3}	4.739	0.11020	0.0027090
	s_{X_1}	151.300	1.04200	0.0206700
	s_{X_2}	166.100	1.17800	0.0222400
	s_{X_3}	132.800	0.90420	0.0174400
	θ_1	2.210	0.09034	0.0025590
	θ_2	4.694	0.16300	0.0037760
PM_{10}	α_{X_1}	2.063	0.04509	0.0007886
	α_{X_2}	2.238	0.04726	0.0007896
	α_{X_3}	1.984	0.04878	0.0008530
	s_{X_1}	172.100	2.45500	0.0384400
	s_{X_2}	147.400	1.96200	0.0317800
	s_{X_3}	298.000	5.33900	0.0835000
	θ_1	1.928	0.04958	0.0008492
	θ_2	2.215	0.06546	0.0011270
$PM_{2.5}$	α_{X_1}	2.326	0.06458	0.0008634
	α_{X_2}	3.340	0.08902	0.0012710
	α_{X_3}	3.623	0.11940	0.0015310
	s_{X_1}	103.700	1.69600	0.0244800
	s_{X_2}	76.290	0.86430	0.0109900
	s_{X_3}	83.470	1.02400	0.0136300
	θ_1	1.636	0.02924	0.0004246
	θ_2	1.704	0.02840	0.0003381

Tabla B.20: *Los DIC's Para el análisis trivariado.*

Contaminante	Cópula tridimensional	Cópula asimétrica
O_3	3.072E8	3.84E8
PM_{10}	2.592E8	3.24E8
$PM_{2.5}$	1.768E8	2.21E8

Tabla B.21: *Estadísticas de los máximos diarios en la ZMVM.*

Componente	Media	d. e.	97.5-percentil
O_3 (ppb)	138.500	60.406	276.00
NO_2 (ppb)	95.080	48.193	217.00
SO_2 (ppb)	76.080	68.298	257.00
PM_{10} ($\mu g/m^3$)	220.200	133.755	575.95
$PM_{2.5}$ ($\mu g/m^3$)	74.120	42.728	160.00
CO (ppm)	7.009	5.604	22.30

Tabla B.22: *Calculo de la ρ de Spearman muestral de los máximos mensuales del O_3 con los otros contaminantes.*

Contaminante	ρ muestral
NO_2 y O_3	-0.32593140
SO_2 y O_3	-0.03623814
PM_{10} y O_3	0.22003870
$PM_{2.5}$ y O_3	0.69456670
CO y O_3	0.02075625

Tabla B.23: *Límites superiores para los parámetros de escala en los modelos (B.1) y (B.2).*

Contaminante	a_x	b_x
O_3	50	50
NO_2	50	100
SO_2	50	100
PM_{10}	100	300
$PM_{2.5}$	50	200
CO	10	10

Tabla B.24: *Estimaciones de los parámetros de los diversos componentes para los modelos (B.1) y (B.2).*

			Media	d. e.	Error M.C.
O_3	F	α	0.7841	0.03261	0.0003769
		s	10.4000	0.84730	0.0094750
	W	α	1.0280	0.04782	0.0005393
		k	33.6700	2.07700	0.0227000
NO_2	F	α	0.7599	0.03090	0.0003455
		s	14.2800	1.16500	0.0124500
	W	α	1.002	0.04519	0.0005066
		k	47.140	2.88100	0.0309900
SO_2	F	α	0.7281	0.02944	0.0003238
		s	21.3800	1.84900	0.0204900
	W	α	1.0930	0.05069	0.0005758
		k	69.0400	3.89800	0.0468400
PM_{10}	F	α	0.7521	0.03216	0.0003601
		s	58.9900	5.28300	0.0551100
	W	α	1.1250	0.05561	0.0006187
		k	179.6000	10.76000	0.1258000
$PM_{2.5}$	F	α	0.7358	0.04097	0.0004816
		s	24.1100	2.75300	0.0318800
	W	α	0.8786	0.05137	0.0005440
		k	88.5400	8.20600	0.0863500
CO	F	α	0.7886	0.03170	0.0003549
		s	1.2610	0.10090	0.0011950
	W	α	1.1270	0.05265	0.0005808
		k	3.8190	0.21090	0.0023520

Tabla B.25: *Estadísticos de comparación para los diversos contaminantes para los modelos (B.1) y (B.2).*

	Modelo Fréchet		Modelo Weibull	
	DIC	MLF	DIC	MLF
O_3	5.62E7	1.638390E-572	5.62E7	8.757836E-550
NO_2	5.84E7	1.468775E-639	5.84E7	1.550686E-615
SO_2	5.80E7	1.635221E-689	5.80E7	1.074329E-654
PM_{10}	4.94E7	1.417759E-688	4.94E7	9.819631E-659
$PM_{2.5}$	3.28E7	8.458427E-402	3.28E7	9.428889E-393
CO	5.82E7	8.635076E-320	5.82E7	3.818431E-289

Tabla B.26: *Valores del parámetro de escala en los modelos (B.3), (B.4) y (B.5).*

	WW		FW		WF	
	b_y	b_z	a_y	b_z	b_y	a_z
NO_2 y O_3	100	50	100	50	100	50
SO_2 y O_3	100	100	50	100	200	50
PM_{10} y O_3	300	50	100	50	300	50
$PM_{2.5}$ y O_3	300	50	200	50	300	50
CO y O_3	10	100	10	100	10	50

Tabla B.27: *Estimaciones de los parámetros suponiendo marginales Fréchet(F) o Weibull(W) en el comportamiento de los umbrales relacionadas a través de la cópula de Frank.*

		NO_2 y O_3		SO_2 y O_3		PM_{10} y O_3		$PM_{2.5}$ y O_3		CO y O_3	
		Media	Error M.C.	Media	Error M.C.	Media	Error M.C.	Media	Error M.C.	Media	Error M.C.
			d. e.		d. e.		d. e.		d. e.		d. e.
W-W	α_y	1.326E0	1.410E-3	1.256E0	1.586E-3	8.935E-1	1.237E-3	1.224E0	2.496E-3	1.089E0	8.936E-4
	α_z	9.319E-1	9.943E-4	1.362E0	1.478E-3	1.137E0	1.381E-3	1.498E0	3.135E-3	9.969E-1	7.743E-4
	k_y	5.587E1	6.013E-2	6.951E1	8.686E-2	1.285E2	4.064E-1	1.572E2	3.643E-1	4.383E0	4.094E-3
	k_z	2.389E1	4.134E-2	4.432E1	5.250E-2	2.532E1	2.464E-2	1.460E1	2.771E-2	3.873E1	4.084E-2
	θ	-2.060E0	1.166E-2	-2.633E-1	1.331E-2	1.440E0	1.261E-2	5.202E0	3.043E-2	9.956E-3	7.218E-3
F-W	α_y	9.105E-1	8.931E-4	1.026E0	1.137E-3	6.507E-1	7.375E-4	8.309E-1	1.606E-3	6.876E-1	4.790E-4
	α_z	9.266E-1	1.088E-3	1.352E0	1.627E-2	1.127E0	1.498E-3	1.403E0	3.217E-3	9.960E-1	8.091E-4
	k_z	2.419E1	3.804E-2	4.595E0	5.100E-2	2.530E1	3.999E-2	1.464E1	2.922E-2	3.484E0	4.069E-2
	s_y	2.215E1	3.378E-2	3.757E0	4.354E-2	3.210E1	8.909E-2	6.122E1	1.850E-1	1.696E-1	1.775E-3
	θ	-2.303E0	1.555E-2	3.160E-1	1.448E-2	1.744E0	1.481E-2	5.731E0	3.496E-2	7.998E-1	1.012E-2
W-F	α_y	1.325E0	1.479E-3	2.175E0	2.611E-3	8.826E-1	1.186E-3	1.186E0	2.490E-3	1.087E0	7.760E-4
	α_z	8.365E-1	9.029E-4	7.534E-1	6.389E-4	7.556E-1	8.487E-4	1.319E0	2.818E-3	7.493E-1	5.611E-4
	k_y	5.669E1	5.995E-2	1.179E2	8.848E-2	1.276E2	2.224E-1	1.522E2	3.979E-1	4.393E0	4.008E-3
	s_z	7.641E0	1.370E-2	1.687E1	3.525E-2	8.332E0	1.760E-2	6.704E0	1.310E-2	1.138E1	1.516E-2
	θ	-2.439E0	1.641E-2	-2.187E-1	1.765E-2	1.675E0	1.476E-2	5.134E0	3.258E-2	-6.381E-2	7.828E-3

Tabla B.28: *Comparación de los DIC y MLF en los modelos bivariados del comportamiento de los umbrales relacionadas a través de la cópula de Frank.*

	NO_2 y O_3		SO_2 y O_3		PM_{10} y O_3		$PM_{2.5}$ y O_3		CO y O_3	
	DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF	DIC	MLF
WW	1.92E7	3.697E-126	1.74E7	1.249E-123	1.5E7	2.381E-109	5.4E6	2.464E-36	4.26E7	2.533E-219
FW	1.92E7	6.229E-130	1.74E7	2.798E-125	1.5E7	1.826E-111	5.4E6	1.987E-37	4.26E7	6.535E-229
WF	1.92E7	4.931E-127	1.74E7	8.604E-131	1.5E7	2.227E-112	5.4E6	1.176E-36	4.26E7	2.748E-225

Referencias

[1] Arizmendi, H., Carrillo, A., and Lara, M. *Cálculo*, 2da ed. Instituto de matemáticas UNAM, 2016.

[2] Ballester, F. Contaminación atmosférica, cambio climático y salud. *Revista Española de salud pública 79* (2005), 159–175.

[3] Barrios, J., and Rodrigues, E. A queueing model to study the occurrence and duration of ozone exceedances in mexico city. *Journal of Applied Statistics 42*, 1 (2015), 214–230.

[4] Bhatter, S., Nishant, and Shyamsunder. Mathematical model on the effects of environmental pollution on biological populations. In *International workshop of Mathematical Modelling, Applied Analysis and Computation* (2022), Springer, pp. 488–496.

[5] Bolboaca, S.-D., and Jäntschi, L. Pearson versus spearman, kendall's tau correlation analysis on structure-activity relationships of biologic active compounds. *Leonardo Journal of Sciences 5*, 9 (2006), 179–200.

[6] Brooks, S., and Gelman, A. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics 7*, 4 (1998), 434–455.

[7] Casella, G., and George, E. Explaining the gibbs sampler. *The American Statistician 46*, 3 (1992), 167–174.

[8] CDMX. Gaceta oficial de la ciudad de méxico no. 230. decima novema Época. administración pública de la ciudad de méxico, 2016. 27 de diciembre.

[9] Chaturvedi, S., Yadav, B. P., Siddiqui, N. A., and Chaturvedi, S. K. Mathematical modelling and analysis of plastic waste pollution and its impact on the ocean surface. *Journal of Ocean Engineering and Science 5*, 2 (2020), 136–163.

[10] Coles, S. *An Introduction to Statistical Modeling of Extreme Values*. Springer, 2007.

[11] Cowles, M. *Applied Bayesian Statistics: with R and OpenBUGS Examples*. Springer Science & Business Media, 2013.

[12] Deheuvels, P. Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. In *Annales de l'ISUP* (1978), vol. 23, pp. 1–36.

[13] Ercelebi, S. G., and Toros, H. Extreme value analysis of istanbul air pollution data. *CLEAN–Soil, Air, Water 37*, 2 (2009), 122–131.

[14] Fisher, N. Copulas. In *Encyclopedia of Statistical Sciences*, 2 ed. Wiley, New York, 2016, pp. 1363–1367.

[15] Gaver, D., and O'Muircheartaigh, I. Robust empirical bayes analyses of event rates. *Technometrics 29*, 1 (1987), 1–15.

[16] Gelfand, A., and Smith, A. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association 85*, 410 (1990), 398–409.

[17] Gelman, A., and Rubin, D. Inference from iterative simulation using multiple sequences. *Statistical science* (1992), 457–472.

[18] Genest, C., and Favre, A.-C. Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of hydrologic engineering 12*, 4 (2007), 347–368.

[19] Genest, C., and Nešlehová, J. G. When gumbel met galambos. In *Copulas and Dependence Models with Applications: Contributions in Honor of Roger B. Nelsen* (2017), Springer, pp. 83–93.

[20] Gray, R. *Probability, Random Processes, and Ergodic Properties*, 2 ed. Springer, 2009.

[21] Grimaldi, S., and Serinaldi, F. Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources 29*, 8 (2006), 1155–1167.

[22] Haan, L. D., and Ferreira, A. *Extreme Value Theory: an Introduction*. Springer Science & Business Media, 2006.

[23] Hernández, A., and Hernández, O. *Elementos de Probabilidad y Estadística*, 2nd ed. Sociedad Matemática Mexicana, 2003.

[24] Hofert, M., Kojadinovic, I., Mächler, M., and Yan, J. *Elements of Copula Modeling With R*. Springer, 2018.

[25] Infante, S., and Zárate, G. *Métodos Estadísticos*, 2da ed. Trillas, 1986.

[26] Jan-FrederikMai, and Scherer, M. *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*, 2nd ed., vol. 6. World Scientific, 2017.

[27] Jenkin, M. E., and Clemitshaw, K. C. Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer. *Atmospheric Environment 34*, 16 (2000), 2499–2527.

[28] Jiménez, B. E. *La Contaminación Ambiental en México*. Editorial Limusa, 2001.

[29] Kotz, S., and Nadarajah, S. *Extreme Value Distributions: Theory and Applications*. World Scientific, 2000.

[30] Lee, P. *Bayesian Statistics: an Introduction*, 4th ed. John Wiley & Sons, 2012.

[31] Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. *The BUGS Book. A Practical Introduction to Bayesian Analysis*. Chapman Hall, 2013.

[32] Masseran, N., and Hussain, S. I. Copula modelling on the dynamic dependence structure of multiple air pollutant variables. *Mathematics 8*, 11 (2020), 1910.

[33] Moore, D. S., and Spruill, M. C. Unified large-sample theory of general chi-squared statistics for tests of fit. *The Annals of Statistics* (1975), 599–616.

[34] N.A.D.F. Norma ambiental para el distrito federal nadf-009-aire-2017. gaceta oficial de la ciudad de méxico, 2018. 14 de noviembre.

[35] Navidi, W. *Estadística Para Ingenieros*, 5ta ed. McGraw Hill Interamericana, 2022.

[36] Nelsen, R. *An Introduction to Copulas*. Springer Science & Business Media, 2007.

[37] N.O.M. Norma oficial mexicana nom-020-ssa1-2014. diario oficial de la federación., 2014. 19 de agosto.

[38] N.O.M. Norma oficial mexicana nom-172-semarnat-2019. diario oficial de la federación., 2019. 20 de noviembre.

[39] QUAN DONG, Y. Value ranges of spearman's rho and kendall's tau of a class of copulas. In *2010 International Conference on Computational and Information Sciences* (2010), IEEE, pp. 182–185.

[40] RAMA. Dirección de monitoreo atmosférico. http://www.aire.cdmx.gob.mx, 2017. Accedido 26-05-2023.

[41] RAO, C. *Linear Statistical Inference and its Applications*, 2nd ed. Wiley New York, 1973.

[42] REISS, R., AND THOMAS, M. *Statistical Analysis of Extreme Values: With Applications to Insurance, Finance, Hydrology, and Other Fields*, 3rd ed. Birkhäuser Verlag, 2001.

[43] RINCÓN, L. *Introducción a la Probabilidad*. Universidad Nacional Autónoma de México, Facultad de Ciencias, 2014.

[44] RINCÓN, L. *Curso Intermedio de Probabilidad*. Universidad Nacional Autónoma de México, Facultad de Ciencias, 2015.

[45] ROBERT, C., AND CASELLA, G. *Monte Carlo Statistical Methods*, 2nd ed. Springer Texts in Statistics. Springer, 2004.

[46] SALVADORI, G., DE MICHELE, C., KOTTEGODA, N. T., AND ROSSO, R. *Extremes in Nature: an Approach Using Copulas*, vol. 56. Springer Science & Business Media, 2007.

[47] SCHWEIZER, B., AND SKLAR, A. *Probabilistic Metric Spaces*. Courier Corporation, 2011.

[48] Shong Chok, N. Pearson's versus spearman's and kendall's correlation coefficients for continuous data. *Master of Science Thesis, University of Pittsburgh, Graduate School of Public Health* (2010).

[49] Sklar, A. Random variables, distribution functions, and copulas: a personal look backward and forward. *Lecture notes-monograph series* (1996), 1–14.

[50] Smith, R. Maximum likelihood estimation in a class of nonregular cases. *Biometrika 72*, 1 (1985), 67–90.

[51] Vallejo, M., Jáuregui-Renaud, K., Hermosillo, A. G., Márquez, M. F., and Cárdenas, M. Efectos de la contaminación atmosférica en la salud y su importancia en la ciudad de méxico. *Gaceta médica de México 139*, 1 (2003), 57–63.

[52] Vazquez, J. A. Información profesional. https://juan0314.github.io/, 2024. Accedido 22-11-2024.

[53] Vazquez, J. A., Rodrigues, E. R., and Reyes, H. J. Bivariate analysis of pollutants monthly maxima in mexico city using extreme value distributions and copula. *Journal of Environmental Protection 15*, 07 (2024), 796–826.

[54] Yau, C. R tutorial with bayesian statistics using openbugs, 2014.

[55] Yue, S., Pilon, P., and Cavadias, G. Power of the mann–kendall and spearman's rho tests for detecting monotonic trends in hydrological series. *Journal of hydrology 259*, 1-4 (2002), 254–271.

[56] Zhang, L., and Singh, V. P. *Copulas and Their Applications in Water Resources Engineering*. Cambridge University Press, 2019.

[57] Zhonghui, J., and Xueqin, L. Comparative analysis of pm2. 5 pollution risk in china using three-dimensional archimedean copula method. *Geomatics, Natural Hazards and Risk 10*, 1 (2019), 2368–2386.