

UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA
SIMULACIÓN DE MODELOS MÁX-ESTABLES
TESIS

PARA OBTENER EL TÍTULO DE:

LICENCIADO EN MATEMÁTICAS APLICADAS

PRESENTA:

CRISTIAN LUCERO GARZÓN

DIRECTOR DE TESIS:

Dr. JOSÉ DEL CARMEN JIMÉNEZ HERNÁNDEZ

HUAJUAPAN DE LEÓN, OAXACA.

JUNIO DE 2022

Agradecimientos

Son muchas las personas que me han ayudado para llevar a buen término esta tesis. Desde luego, el Dr. José del Carmen Jiménez Hernández, mi director debe aparecer en el primer lugar por su generosidad, buenos consejos y una elevada dosis de paciencia. A mis sinodales, Dra. Marisol López Cerino, Dr. Alejandro Iván Aguirre Salado, M.M.M. Armando Alcalá Vallejo y M.C. Tirso Miguel Ángel Ramírez Solano por su contribución al desarrollo de esta tesis. Volviendo la vista atrás, es preciso mencionar también a cada uno de mis profesores de la Universidad que me iniciaron en los arcanos de la simulación y la investigación. A Dolores, mi madre, por su comprensión y ayuda en momentos malos y menos malos. Me ha enseñado a encarar las adversidades sin desfallecer en el intento. Me ha dado todo lo que soy como persona, mis valores, mis principios, mi perseverancia y mi empeño, y todo ello con una gran dosis de amor y sin pedir nunca nada a cambio. A mi hijo Joshua Miguel que me ayudó mucho a nunca darme por vencido, además de ofrecerme siempre su cariño. A Karen Michelle, mi compañera, que me ha apoyado en todo momento. A mis hermanas y hermano quienes me alentaron y apoyaron en cada paso para que no sólo este trabajo pudiese terminarse si no que también en el proceso universitario.

Índice general

Introducción	1
1. Preliminares	7
1.1. Conceptos básicos	7
1.1.1. Espacio de probabilidad	8
1.1.2. Probabilidad de eventos	13
1.1.3. Variables aleatorias y sus distribuciones	15
1.1.4. Esperanza	18
1.1.5. Varianza	20
1.1.6. Variables aleatorias multidimensionales y sus distribuciones	22
1.1.7. Esperanza condicional	24
1.1.8. Covarianza	25
1.1.9. Correlación	25
1.1.10. Distribución normal multivariada	26
1.1.11. Teoremas límite	27
1.2. Procesos estocásticos	30
1.2.1. Procesos con incrementos independientes	32
1.2.2. Procesos estacionarios	32
1.2.3. Procesos con incrementos estacionarios	32
1.2.4. Cadena de Markov a tiempo continuo	32
1.2.5. Proceso de Poisson	34
1.2.6. Procesos gaussianos	36
2. Teoría de valores extremos	39
2.1. Teoría de valores extremos clásica	40
2.1.1. Formulación del modelo	40
2.1.2. Teorema de valores extremos	43

2.1.3.	Distribución generalizada de valores extremos	44
2.1.4.	Demostración del teorema de valores extremos	47
2.1.5.	Ejemplos	48
2.1.6.	Modelos de máximos por bloques	49
2.1.7.	Cuantiles extremos	50
2.2.	Teoría de valores extremos multivariante	51
2.2.1.	Caso bidimensional	52
2.2.2.	Caso multidimensional	54
3.	Procesos máx-estables	61
3.1.	Conceptos iniciales	62
3.2.	Caracterización espectral	64
3.2.1.	Construcción de L. de Haan (1984) de procesos máx-estables	64
3.2.2.	Construcción de Schlather (2002) de procesos máx-estables	65
3.3.	Distribuciones finito dimensionales de los procesos máx-estables .	66
4.	Modelos máx-estables	69
4.1.	Modelo de Smith	69
4.2.	Modelo de Schlather	74
4.3.	Modelo de Brown-Resnick	85
4.4.	Modelo t-extremal	86
5.	Simulación de los modelos máx-estables	89
5.1.	Simulación incondicional	90
5.1.1.	Aplicación con R al modelo de Smith	92
5.1.2.	Aplicación con R al modelo Schlather	95
5.1.3.	Aplicación con R al modelo Brown-Resnick	97
5.1.4.	Aplicación con R al modelo t-extremal	99
5.2.	Simulación condicional	101
5.2.1.	Aplicación con R al modelo Schlather	101
5.2.2.	Aplicación con R al modelo Brown-Resnick	103
5.2.3.	Aplicación con R al modelo t-extremal	105
	Conclusiones	107
A.	Códigos de simulaciones incondicionales	109
A.1.	Modelo de Smith	109

A.1.1. Modelo de Smith unidimensional	109
A.1.2. Modelo de Smith bidimensional	109
A.2. Modelo de Schlather	111
A.2.1. Modelo de Schlather bidimensional	111
A.3. Modelo de Brown-Resnick	112
A.3.1. Modelo de Brown-Resnick bidimensional	112
A.4. Modelo t-extremal	113
A.4.1. Modelo t-extremal bidimensional	113
B. Códigos de simulaciones condicionales	115
B.1. Modelo de Schlather	115
B.1.1. Modelo de Schlather bidimensional	115
B.2. Modelo de Brown-Resnick	117
B.2.1. Modelo de Brown-Resnick bidimensional	117
B.3. Modelo t-extremal	118
B.3.1. Modelo t-extremal bidimensional	118
Bibliografía	120

Índice de figuras

2.1. Simulación de la distribución acumulada exponencial degenerada.	42
5.1. Realizaciones del modelo de Smith con marginales Fréchet unitarias en \mathbb{R} , para diferentes valores del parámetro var . (a) $var = 0.1$; (b) $var = 1$; (c) $var = 4$; (d) $var = 50$; (e) $var = 100$; (f) $var = 1000$.	93
5.2. Realizaciones del modelo de Smith con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros de la matriz de covarianza. (a) $\sigma_{11} = 9/8, \sigma_{12} = 0, \sigma_{22} = 9/8$; (b) $\sigma_{11} = 9/8, \sigma_{12} = 3/4, \sigma_{22} = 9/8$; (c) $\sigma_{11} = 2, \sigma_{12} = 3/4, \sigma_{22} = 9/8$; (d) $\sigma_{11} = 1, \sigma_{12} = 3/4, \sigma_{22} = 9/8$; (e) $\sigma_{11} = 9/8, \sigma_{12} = 1/2, \sigma_{22} = 3/4$; (f) $\sigma_{11} = 9/8, \sigma_{12} = 1/2, \sigma_{22} = 9/8$.	94
5.3. Realizaciones del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2, \nu = 0.5$ y (b) $c_2 = 4, \nu = 1.9$.	96
5.4. Realizaciones del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2, \nu = 0.5$ y (b) $c_2 = 4, \nu = 1.9$.	96
5.5. Realizaciones del modelo de Brown-Resnick con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros. (a) $c_2 = 100000, \nu = 1.9$; (b) $c_2 = 0.00001, \nu = 1.5$; (c) $c_2 = 1, \nu = 1$; y (d) $c_2 = 0.001, \nu = 0.000005$.	98
5.6. Realizaciones del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2, \nu = 0.5$ y (b) $c_2 = 4, \nu = 1.9$.	100

- 5.7. Realizaciones del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$ 100
- 5.8. Simulaciones condicionales del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$ 102
- 5.9. Simulaciones condicionales del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$ 102
- 5.10. Simulaciones condicionales del modelo de Brown-Resnick con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros. (a) $c_2 = 3$, $\nu = 1.9$; (b) $c_2 = 1$, $\nu = 0.1$; (c) $c_2 = 1$, $\nu = 1$; y (d) $c_2 = 0.1$, $\nu = 0.05$ 104
- 5.11. Simulaciones condicionales del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$ 106
- 5.12. Simulaciones condicionales del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$ 106

Introducción

Cuando se tiene un registro de situaciones inusuales, estas se representan bajo el concepto de valores atípicos, esto motivó el desarrollo de la teoría de valores extremos, sustentada matemáticamente por [de Haan y Ferreira \(2006\)](#) y estadísticamente por [Coles *et al.* \(2001\)](#) en la actualidad. Las bases de esta teoría fueron sentadas por Fisher, Tippet y Gnedenko al enunciar y demostrar el primer teorema de la teoría de valores extremos. Un ejemplo donde se aplica esta teoría, es en problemas de precipitación ([Moreno y Sánchez, 2016](#)), es común que los mínimos y máximos de estas se eliminan de primera instancia y no se preocupa por la información que puedan proporcionar acerca de inundaciones o sequías, por ello surge la necesidad de estudiar y analizar este tipo de datos en teoría de valores extremos, la cual durante las últimas décadas, se ha visto desarrollada tanto a nivel teórico como práctico.

Otro ejemplo es el de protección contra el mar por diques ([de Haan y Ferreira, 2006](#)) es de gran interés su estudio, debido a que, los diques tienen que soportar mareas de tormenta que eleva el nivel del agua de mar a lo largo de la costa. En general la teoría de valores extremos se ocupa de modelar fenómenos en hidrología, contaminación, corrosión, medio ambiente, meteorología, seguros, aplicaciones financieras, geología, oceanografía, ingeniería eólica, entre otras ([Coles *et al.*, 2001](#)).

La funcionalidad de estos modelos de valores extremos es la adaptación relativamente simple a fenómenos poco usuales, como lo es cuantificar el riesgo en econometría financiera ([Olmo, 2005](#)), además otras ventajas es incluir la facilidad con la que se puede incorporar toda la información relevante en una inferencia y la facilidad para cuantificar incertidumbres en la estimación ([Coles *et al.*, 2001](#)).

La teoría de valores extremos se originó principalmente de las necesidades de los astrónomos al utilizar o rechazar observaciones periféricas ([Kotz y Nadarajah,](#)

2000). Por lo general, los valores extremos son menos estudiados que los promedios. En algunas aplicaciones, como en los ejemplos anteriores, es necesario extrapolar los datos de los niveles observados a niveles no observados, para dar predicciones del fenómeno, esta teoría proporciona modelos que permiten tal extrapolación (Coles *et al.*, 2001). La teoría de valores extremos proporciona una base teórica sólida y un marco para la extrapolación, fundamentada en teoría asintótica, ya que en muchas aplicaciones es necesario que se haga de esa manera, para poder obtener resultados que sirvan para predecir fenómenos que pueden causar daños irreversibles.

Después de analizar el comportamiento de fenómenos atípicos con la teoría de valores extremos para la modelación de lo inusual en lugar de lo habitual, se busca considerar otras componentes, con ello surgen interrogantes como por ejemplo, ¿qué pasaría si ahora se considera la ubicación espacial de un fenómeno? Es decir, ¿dónde se recolectan, (ubicación espacial) los datos del fenómeno? Dado que en muchas aplicaciones los datos se obtienen en diferentes estaciones meteorológicas ubicadas espacialmente, es necesario considerar la componente espacial en la teoría de valores extremos.

Actualmente, los avances en la teoría de valores extremos se centran en el desarrollo de modelos y métodos para valores extremos de fenómenos espaciales y otras estructuras más complejas, para ello surge la teoría de valores extremos máx-estables, la cual se encarga del estudio de fenómenos espaciales.

La teoría funcional de valores extremos, es decir, los extremos de los procesos estocásticos, fue especialmente activa con la caracterización asintótica y espectral de procesos máx-estables y la construcción de modelos paramétricos máx-estables (Ribatet, 2013). La teoría y la práctica estadística de la teoría de valores extremos univariados está bien desarrollada, sin embargo, se carece de análisis en el modelado de extremos espaciales, esto es problemático, ya que muchos procesos ambientales tienen un dominio espacial.

Los procesos máx-estables son una extensión natural de la teoría de valores extremos al caso dimensional infinito, es decir, los extremos de los procesos estocásticos, y por lo tanto juegan un papel esencial en el modelado estadístico de extremos espaciales (Dombry *et al.*, 2013). Estos procesos fueron caracterizados

completamente por [de Haan \(1984\)](#). Una descripción más precisa de los procesos máx-estables se puede obtener a partir de su representación espectral, esta se determina a partir de dos posibles enfoques. Una primera caracterización es dada por [de Haan \(1984\)](#), proporciona dicha representación considerando funciones de tipo determinístico. Una segunda caracterización de los procesos máx-estables es proporcionada por [Schlather \(2002\)](#) donde considera funciones aleatorias en lugar de funciones determinísticas.

Basado en estos enfoques, se han considerado algunos modelos para procesos máx-estables. Un primer modelo fue desarrollado por [Smith \(1990\)](#) para la modelación de lluvias a través de procesos máx-estables continuos en el tiempo, este modelo se basa en procesos de Gauss y se fundamenta con la teoría de valores extremos, se deduce a partir de la caracterización espectral propuesta por [de Haan \(1984\)](#), conocido como modelo de Smith, también recibe el nombre de proceso de valores extremos gaussianos o modelo perfil de la tormenta. La matriz de varianzas y covarianzas para el proceso gaussiano se desconoce y por tanto el objetivo es estimarla. Este modelo se desarrolla de forma bidimensional, debido a que únicamente se puede obtener de forma explícita la función de distribución acumulada bivariada, es aplicado en general para modelar lluvias del tipo de convección ([Moreno y Ortega, 2013](#)). En este proceso, que Smith interpretó en términos de precipitaciones, depende del tamaño y el tipo de tormenta ([Villanueva e Ibáñez, 2016](#)).

La propuesta de [Schlather \(2002\)](#) es una extensión del modelo de [Smith \(1990\)](#), sustituyendo la función determinística por una forma aleatoria, dicho modelo se desarrolla utilizando réplicas independientes de un proceso estándar gaussiano con función de correlación, partiendo de distribuciones acumuladas finitas dimensionales de un proceso máx-estable, que son expresadas de forma bidimensional. Algunas de las posibles funciones de correlación utilizadas, extraídas de [Gaetan y Guyon \(2010\)](#) son: exponencial, exponencial generalizada, exponencial potencia, esférica, gaussiana, Matérn o bien función de Bessel modificada, Whittle-Matérn y Cauchy, la elección de la función de correlación depende del fenómeno de estudio, el objetivo de este modelo es estimar los parámetros que intervienen en la función de correlación. En general este modelo es aplicado para lluvias ciclónicas, dado que la forma de la tormenta es aleatoria y su estructura está dada por la función

de correlación. En el caso de una tormenta fija, esta mantiene la misma estructura de dependencia y solamente varía la magnitud (Moreno y Ortega, 2013).

El modelo de Brown-Resnick (Brown y Resnick, 1977) desarrollado por Kabluchko *et al.* (2009), analizado y generalizado desde el punto de vista de los procesos máx-estables, surge de elegir la caracterización de Schlather (2002). Nuevamente para este modelo se encuentra presente el proceso de tipo gaussiano con incrementos estacionarios, distribuciones finitas dimensionales, que son expresadas en dimensión dos debido a que únicamente se puede obtener de forma explícita la función de distribución acumulada bivariada. Además dado que es una generalización del modelo de Schlather (2002) con la elección correcta de la función es posible derivar de este, el modelo de Smith (1990).

El modelo t -extremal fue desarrollado por Opitz (2013) y también analizado por Ribatet y Sedki (2013), es una generalización del proceso Brown-Resnick desarrollado por Kabluchko *et al.* (2009), al igual que en los modelos anteriores, este depende del proceso gaussiano propuesto por Schlather (2002), este modelo concibe una construcción espectral que generaliza el proceso gaussiano, además de ser analizado de forma bivariada dado que únicamente se conoce de forma explícita la función de distribución acumulada bivariada.

Como planteamiento del problema se tienen algunas de las interrogantes que se han presentado al exponer los modelos anteriores, las cuales son: ¿Según estos modelos que se proponen acerca de procesos máx-estables, cuál se adapta mejor a un determinado problema? Habiendo estudiado un fenómeno determinado con estos modelos, ¿cuál es el que mejor se ajusta? Si todos los modelos se ajustan a un fenómeno, ¿estos pueden dar predicciones de la misma exactitud? No solamente eso, grandes empresas ignoran el porqué de muchos fenómenos ambientales que dañan la población y medio ambiente. Esta preocupación provoca el estudio del comportamiento de dichos fenómenos por medio de los modelos a describir, para esto es necesario conocer la funcionalidad y el proceso que debe hacer cada modelo para poder llegar a resultados que brinden información de ayuda, es decir, la simulación es esencial para que en un futuro los modelos se apliquen a problemas reales y se tenga un panorama más amplio acerca de la interpretación de los resultados obtenidos por los modelos aplicados al fenómeno.

Como en gran parte de áreas en matemáticas, la teoría desarrollada se encuentra en textos escritos en inglés y de manera dispersa, es por esto que, en esta tesis, se recopila, analiza y redacta de manera más clara, completa y precisa, la información acerca de estos modelos a estudiar, para que en un futuro sirva de referencia para el desarrollo de más teoría acerca de estos modelos y sus aplicaciones. No obstante, esta tesis sirva como guía básica para la comprensión de dichos modelos y su utilidad.

El objetivo de esta tesis es realizar una revisión de los principales enfoques existentes en la teoría de valores extremos espaciales, prestando especial atención a los modelos máx-estables para realizar simulaciones de estos modelos.

La tesis consta de cinco capítulos, al inicio de cada uno de ellos se da el objetivo de dicho capítulo.

Capítulo 1

Preliminares

El objetivo de este capítulo es proporcionar conceptos básicos y resultados de la teoría de probabilidad y estadística que serán esenciales al desarrollo de esta tesis. Los conceptos y teoremas son una recopilación de “Probabilistic methods in combinatorial analysis” de [Sachkov \(1997\)](#), “Probabilidad y estadística, aplicaciones y métodos” de [George \(1988\)](#), “Introducción a la probabilidad” de [Rincón \(2014\)](#), “Estadística básica para estudiantes de Ciencias” de [Gorgas *et al.* \(2011\)](#) y “An introduction to statistical modeling of extreme values” de [Coles *et al.* \(2001\)](#). Los conceptos y teoremas mencionados en las siguientes secciones son de naturaleza auxiliar y se incluyen como referencia, por esta razón, la mayoría de las declaraciones se dan sin pruebas.

1.1. Conceptos básicos

En las ciencias, es frecuente el uso de “variables”. El diccionario de la Real Academia Española (RAE) define una variable como, “magnitud que puede tener un valor cualquiera de los comprendidos en un conjunto”. En estadística, se estudian las variables aleatorias, cuyo valor depende del azar, a estas el diccionario de la RAE las define como, “variable que tiene asociada una determinada ley o distribución de probabilidad, en la que a cada uno de los valores que puede tomar le corresponde una frecuencia relativa o de probabilidad específica”. Por otro lado el diccionario de la RAE define variable estadística, o variable estocástica, como, “función real definida sobre una población finita o una muestra, que toma los valores de cada una de las modalidades de un atributo, y a las que asocia una distribución de frecuencias”. Las variables aleatorias suelen corresponder a dos ca-

tegorías, a saber, discretas y continuas. En esta sección se inicia con el aprendizaje de estas variables aleatorias, y el resto de ella se desarrolla utilizando las mismas.

1.1.1. Espacio de probabilidad

Para el estudio de variables aleatorias, primero que nada se requiere saber que sucede con un fenómeno o experimento en la naturaleza de tipo aleatorio.

Definición 1.1. *Un experimento aleatorio es aquel que cuando se repite bajo las mismas condiciones, el resultado que se observa no siempre es el mismo y tampoco es predecible.*

Dado este tipo de experimentos, no se sabe cual será el resultado, así que por lo menos conviene agrupar en un conjunto a todos los resultados posibles. Parte de las definiciones en probabilidad se fundamenta en teoría de conjuntos y es de ahí de donde se partirá. La colección de todos los posibles resultados de un experimento aleatorio es útil en la definición de probabilidad.

Definición 1.2. *Un espacio muestral es un conjunto arbitrario no vacío Ω que representa todos los posibles resultados de un experimento aleatorio.*

Nota 1.1. *Los elementos de Ω se les llama **eventos elementales, simples o resultados.***

El conjunto de todos los posibles resultados Ω puede ser finito, infinito numerable o infinito no numerable, a continuación se presentan algunos ejemplos que ilustran estas afirmaciones.

Ejemplo 1.1.

Espacio muestral finito. *El número de reservaciones sin cancelar en un restaurante constituye un espacio muestral finito, dado que el número de reservaciones no canceladas puede ser cualquier entero positivo no mayor que la capacidad que tenga dicho lugar.*

Espacio muestral infinito numerable. *El número de llegadas al recibir un servicio constituye un espacio muestral infinito numerable, dado que es posible colocar los resultados en una correspondencia uno a uno con los enteros positivos.*

Espacio muestral infinito no numerable. *La duración de una componente constituye un espacio muestral infinito no numerable, dado que puede ser cualquier número real positivo.*

En este momento, es conveniente dar las formas en que se puede clasificar el espacio muestral Ω en las siguientes definiciones.

Definición 1.3. *Se dice que un espacio muestral Ω es **discreto** si este es finito o infinito numerable y w_i son eventos elementales tal que*

$$\Omega = \{w_1, w_2, \dots, w_n\},$$

o

$$\Omega = \{w_1, w_2, \dots, w_n, \dots\},$$

esto es, si Ω puede ponerse en una correspondencia uno a uno con el conjunto de los enteros positivos.

Definición 1.4. *Se dice que un espacio muestral Ω es **continuo** si sus resultados consisten de un intervalo de números reales.*

Con respecto a Ω , se puede estar particularmente interesado por un subconjunto de este.

Definición 1.5. *Un **evento** de Ω es un grupo de resultados contenidos en este, es decir, cualquier subconjunto $E \subset \Omega$, cuyos miembros tienen una característica en común.*

Por característica común debe entenderse que únicamente un grupo de resultados en particular satisface la característica y los restantes, contenidos en Ω , no. Para poder complementar el conocimiento acerca de eventos en un espacio muestral se da la siguiente nota.

Nota 1.2.

- *Se dice que un evento E_1 es un caso particular de un evento E_2 si se produce la inclusión $E_1 \subset E_2$.*
- *El espacio muestral Ω , es evento en si mismo y puede entenderse como un **evento seguro**.*

- El evento que no contiene a ningún evento simple del espacio muestral recibe el nombre de **evento nulo o vacío** y se representa \emptyset .

En muchos casos, no basta conocer los eventos simplemente, es decir, uno busca más acerca de ellos como por ejemplo: una medida que indique la certeza o incertidumbre que suceda en un cierto evento. Esto lleva a introducir un nuevo concepto que involucre un evento elemental del espacio muestral Ω con un número real que mide la posibilidad de que suceda este, cuando el experimento se lleve a cabo. La probabilidad de un evento es un número real que mide la posibilidad de ocurrencia ya sea colectiva o no, de los resultados del evento cuando se lleve a efecto el experimento. Esto se manifiesta en la definición axiomática de la probabilidad.

Definición 1.6. Sea Ω cualquier espacio muestral y E cualquier evento de este. Se llama **función de probabilidad** sobre el espacio muestral Ω a $P(E)$ si satisface las siguientes axiomas:

1. $P(E) \geq 0$.
2. $P(\Omega) = 1$.
3. Si, para los eventos $E_1, E_2, \dots, E_n, \dots$, con $E_i \cap E_j = \emptyset$ para cada $i \neq j$, entonces

$$P(E_1 \cup E_2 \cup \dots \cup E_n \cup \dots) = P(E_1) + P(E_2) + \dots + P(E_n) + \dots$$

Estos axiomas constituyen la base sobre la que se puede construir toda la teoría del cálculo de probabilidades. En seguida se enuncia algunas de las propiedades consecuencia de los axiomas en la definición (1.6).

Propiedades 1.1.

1. $P(\emptyset) = 0$.
2. Para cualquier evento $E \subset \Omega$, se tiene que $0 \leq P(E) \leq 1$.
3. Para cualesquiera dos eventos $E_1, E_2 \subset \Omega$, se tiene que

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2).$$

4. Para cualesquiera dos eventos $E_1, E_2 \subset \Omega$, si $E_1 \subseteq E_2$, entonces

- $P(E_1) \leq P(E_2)$,
- $P(E_2 - E_1) = P(E_2) - P(E_1)$.

Nota 1.3. Como consecuencia de las propiedades (1.1) inciso 2, se tiene la definición de **probabilidad del complemento**. Para $E \subset \Omega$ cualquier evento se tiene que la probabilidad del complemento de E , denotado por \overline{E} está dada por,

$$P(\overline{E}) = 1 - P(E).$$

A partir de este momento y en lo siguiente se hará uso de la notación que se muestra a continuación

$$\bigcup_{i=1}^n E_i = E_1 \cup E_2 \cup \dots \cup E_n,$$

$$\bigcap_{i=1}^n E_i = E_1 \cap E_2 \cap \dots \cap E_n.$$

Con esto se tiene la siguiente desigualdad, conocida como desigualdad de Boole, para cada número finito de eventos E_1, E_2, \dots, E_n , se cumple que

$$P\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n P(E_i).$$

Nota 1.4. Si E_1, E_2, \dots, E_n son eventos disjuntos a pares, es decir, $E_i \cap E_j = \emptyset$ para cada $i \neq j$, se tiene que

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i).$$

Para lo que sigue será conveniente recordar la definición de σ -álgebra.

Definición 1.7. Sea Ω un espacio muestral y \mathcal{F} una familia de subconjuntos de Ω . Se dice que \mathcal{F} es una **σ -álgebra** si satisface las siguientes condiciones:

1. $\Omega \in \mathcal{F}$.
2. Si $E \in \mathcal{F}$ entonces $\overline{E} \in \mathcal{F}$.

3. Si $\{E_n\}$ es una sucesión de elementos de \mathcal{F} entonces

$$\bigcup_{i=1}^{\infty} E_i \in \mathcal{F} \quad y \quad \bigcap_{i=1}^{\infty} E_i \in \mathcal{F}.$$

A partir de ahora no se le llamará evento a cualquier subconjunto del espacio muestral sino únicamente a aquellos elementos que pertenezcan a una σ -álgebra asociada al espacio muestral.

De esta forma en la σ -álgebra, \mathcal{F} agrupa a todos los subconjuntos de Ω para los que se está interesado en calcular su probabilidad y tal colección constituye el dominio sobre el cual se define una medida de probabilidad.

Así, a cada experimento aleatorio particular se le puede asociar una pareja (Ω, \mathcal{F}) compuesta por el espacio muestral y una σ -álgebra de eventos.

Nota 1.5. Si \mathcal{F} es una σ -álgebra de Ω entonces el par ordenado (Ω, \mathcal{F}) es llamado **espacio medible**. Los elementos de la σ -álgebra \mathcal{F} se les llama eventos correspondientes al espacio medible (Ω, \mathcal{F}) .

Una σ -álgebra de gran estudio y que será de ayuda para dar algunas definiciones en lo que sigue, es la σ -álgebra de Borel. Se tomará como espacio muestral el conjunto de números reales \mathbb{R} y se considerará la colección de intervalos de la forma $(-\infty, x]$ para cualquier número real x , es decir,

$$\mathcal{C} = \{(-\infty, x] : x \in \mathbb{R}\}.$$

La σ -álgebra de Borel de \mathbb{R} se puede construir a partir de esta colección considerando la σ -álgebra más pequeña de subconjuntos de \mathbb{R} que contiene a la colección \mathcal{C} .

Definición 1.8. La σ -álgebra de Borel de \mathbb{R} se denota por $\mathcal{B}(\mathbb{R})$ y se define como la mínima σ -álgebra de subconjuntos de \mathbb{R} que contiene a todos los intervalos de la forma $(-\infty, x]$ esto se escribe de la forma

$$\mathcal{B}(\mathbb{R}) = \sigma\{(-\infty, x] : x \in \mathbb{R}\}.$$

A los elementos de $\mathcal{B}(\mathbb{R})$ se les llama conjuntos de Borel, conjuntos Borel medibles o simplemente Borelianos de \mathbb{R} .

El símbolo σ en la expresión anterior significa que se está tomando la mínima σ -álgebra generada por la colección $\{(-\infty, x] : x \in \mathbb{R}\}$, y el adjetivo mínimo significa

que si \mathcal{F} es una σ -álgebra que contiene a la colección \mathcal{C} , entonces $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}$, es decir, $\mathcal{B}(\mathbb{R})$ es la más pequeña.

A continuación se reescribe la definición (1.6) axiomática de la probabilidad, ahora en términos de σ -álgebra considerando un espacio medible, definiendo así un espacio de probabilidad.

Definición 1.9. Sean (Ω, \mathcal{F}) espacio medible y $P : \mathcal{F} \rightarrow [0, 1]$ una función numérica, llamada **medida de probabilidad** definida sobre \mathcal{F} , tal que satisface los axiomas de la definición (1.6):

1. $P(E) \geq 0$ para cada $E \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. Si, para los eventos $E_1, E_2, \dots, E_n, \dots$, con $E_i \cap E_j = \emptyset$ para cada $i \neq j$, entonces

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

La terna ordenada (Ω, \mathcal{F}, P) se llama **espacio de probabilidad**.

1.1.2. Probabilidad de eventos

Para continuar con el estudio de la probabilidad de eventos en un espacio muestral se examinarán los conceptos de probabilidad condicional, mutuamente independientes, probabilidad total y se desarrolla la ley de multiplicación de probabilidad.

En muchos casos interesa conocer la probabilidad de un evento E_1 en el caso de que haya ocurrido otro suceso E_2 . A esta probabilidad de que se cumpla E_1 bajo la condición de que se cumpla E_2 se le llama probabilidad de E_1 condicionada a E_2 , y se denota por $P(E_1 | E_2)$. La definición matemática de la probabilidad condicionada se presenta a continuación.

Definición 1.10. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y $E_1, E_2 \in \mathcal{F}$. Si $P(E_2) > 0$ entonces el número $P(E_1 | E_2)$, definido por la fórmula

$$P(E_1 | E_2) := \frac{P(E_1 \cap E_2)}{P(E_2)},$$

se llama **probabilidad condicional** del evento E_1 bajo la condición E_2 .

La definición (1.10) de probabilidad condicional permite calcular la probabilidad de la intersección de dos eventos. En este caso, la probabilidad de que E_1 ocurra no está afectada por la ocurrencia o no ocurrencia de E_2 y se dice que los dos sucesos son independientes. En general, se tiene la siguiente definición para un número finitos de eventos que se dan de manera independiente.

Definición 1.11. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y $E_1, E_2, \dots, E_n \in \mathcal{F}$ se dice que son **mutuamente independientes** si, para todas las combinaciones de los subíndices $1 \leq i_1 < \dots < i_k \leq n$, $k = 2, 3, \dots, n$, se tiene

$$P\left(\bigcap_{s=1}^k E_{i_s}\right) = \prod_{s=1}^k P(E_{i_s}).$$

Nota 1.6. Se dice que una colección infinita de eventos es independiente si cualquier subcolección finita de ella lo es.

Luego de poder deducir si una sucesión de eventos son independientes surgen diversas incertidumbres, entre ellas, saber si es posible calcular la probabilidad de un evento, por medio de probabilidades de alguna sucesión de eventos que formen un conjunto completo; esto es referido a que los eventos $E_1, E_2, \dots, E_n, \dots$ constituyen un conjunto completo de eventos si $E_1 \cup E_2 \cup \dots \cup E_n \cup \dots = \Omega$, siendo cada E_i disjuntos a pares, sobre el espacio en el cual esté contenido este evento. Es decir, si $\{E_i\}$ forman un conjunto completo de eventos de Ω la probabilidad de que ocurra $E \subset \Omega$ es la suma de las probabilidades de los eventos E_i por las probabilidades de E condicionadas a cada E_i . Formalmente se describe en el siguiente teorema.

Teorema 1.1. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y $\{E_i\} \subset \mathcal{F}$. Si, para los eventos $E_1, E_2, \dots, E_n, \dots$, con $E_i \cap E_j = \emptyset$ para cada $i \neq j$, $P(E_i) > 0$ y $E \subseteq \bigcup E_i$, entonces se tiene la fórmula de **probabilidad total**

$$P(E) = \sum_{i=1}^{\infty} P(E_i)P(E | E_i).$$

Existe un aspecto adicional que debe mencionarse antes de comenzar con el estudio de las variables aleatorias. Dados los eventos E_1 y E_2 , es posible calcular $P(E_1 \cap E_2)$ si se parte del supuesto de que estos son independientes. Por ende el resultado es simplemente el producto de las probabilidades de cada evento. Para el

caso de dependencia, se tiene la **ley de multiplicación de probabilidad** como consecuencia de la definición (1.10) de probabilidad condicional.

$$P(E_1 \cap E_2) = P(E_1 | E_2)P(E_2).$$

1.1.3. Variables aleatorias y sus distribuciones

Hasta ahora se han examinado conceptos básicos de probabilidad con respecto a eventos que se encuentran en un espacio muestral. Los experimentos se conciben de manera que los resultados del espacio muestral son cualitativos o cuantitativos. En cualquier experimento, hay numerosas características que se pueden observar o medir, pero en la mayor parte de los casos un experimentador se centra en algún aspecto específico o aspecto de una muestra. En general, cada resultado de un experimento se puede asociar con un número si se especifica una regla asociada. Esta regla de asociación se llama variable aleatoria, una variable porque toma diferentes valores numéricos, y aleatoria porque el valor observado depende de que resultado experimental resulte. El concepto de variable aleatoria proporciona un medio para relacionar cualquier resultado con una medida cuantitativa.

Definición 1.12. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Una **variable aleatoria** $\xi : \Omega \rightarrow \mathbb{R}$ es una función numérica tal que

$$\xi := \xi(\omega), \quad \omega \in \Omega.$$

Esto es, un mapeo de Ω en el conjunto de los números reales \mathbb{R} tal que para cada $x \in \mathbb{R}$ el conjunto $\{\omega \in \Omega : \xi(\omega) \leq x\}$ pertenece a la σ -álgebra \mathcal{F} . O bien, en términos de la σ -álgebra de Borel, para cualquier conjunto de Borel $B \in \mathcal{B}(\mathbb{R})$, el subconjunto de Ω dado por,

$$\{\xi \in B\} := \{\omega \in \Omega : \xi(\omega) \in B\}$$

está en la σ -álgebra \mathcal{F} .

Nota 1.7. Una variable aleatoria puede tomar dos tipos de valores:

1. Se dice que una variable aleatoria ξ es **discreta** si el conjunto de sus valores que puede tomar es finito o infinito numerable.

2. Se dice que una variable aleatoria ξ es **continua** si sus valores consisten en uno o más intervalos de la recta real. Es decir, toma todos los valores dentro de un intervalo $(a, b) \subset \mathbb{R}$.

Cabe aclarar que a partir de este momento se centrará el estudio en variables aleatorias continuas únicamente, debido al objetivo de esta tesis.

Para describir completamente una variable aleatoria hay que indicar las probabilidades que tome en algún intervalo.

A partir de este momento se empezará a utilizar el concepto de integral de una función.

Definición 1.13. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y una variable aleatoria continua $\xi : \Omega \rightarrow \mathbb{R}$. Se dice que la función integrable y no negativa $f : \mathbb{R} \rightarrow \mathbb{R}$ es la **función de densidad de probabilidad** de ξ , si para cada intervalo $[a, b]$ de \mathbb{R} se cumple la igualdad,

$$P(a \leq \xi \leq b) = \int_a^b f(x)dx.$$

La probabilidad de que una variable aleatoria continua tome un valor específico es cero.

Nota 1.8. De la definición (1.13) la función de densidad de probabilidad cumple las siguientes dos propiedades:

1. $f(x) \geq 0$ para toda $x \in \mathbb{R}$.
2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

Otra función que puede asociarse a una variable aleatoria es la función de distribución acumulada.

Definición 1.14. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad y una variable aleatoria $\xi : \Omega \rightarrow \mathbb{R}$. La **función de distribución acumulada**, o simplemente, **la distribución de la variable** ξ se define y denota como

$$F_{\xi}(x) := P\{\xi \leq x\}, \text{ para toda } x \in \mathbb{R}.$$

Esto no es más que, la función de densidad de probabilidad evaluada en todos aquellos valores menores o iguales a x . Por simplicidad F_{ξ} se denotará sólo como F .

Ahora, interesa conocer como se calcula dicha función de distribución acumulada, lo cual se hace de la siguiente manera.

Nota 1.9. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad, una variable aleatoria continua ξ , $x \in \mathbb{R}$ y f la función de densidad de probabilidad de ξ , la función de distribución acumulada se calcula como sigue

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

Una vez visto como se define la función de probabilidad y la función de distribución, para el caso de variables aleatorias continuas se desprenden algunas propiedades.

Propiedades 1.2. Si ξ es una variable aleatoria continua:

- La probabilidad de que la variable aleatoria ξ esté en un cierto intervalo (x_i, x_j) se podrá expresar como:

$$P\{x_i < \xi < x_j\} = \int_{x_i}^{x_j} f(x)dx = F(x_j) - F(x_i).$$

- Si F es la función de distribución acumulada, entonces la función de densidad de probabilidad f se obtiene como,

$$f(x) = \frac{dF(x)}{dx}.$$

A continuación se muestra algunas propiedades generales válidas para toda función de distribución.

Propiedades 1.3. Sean (Ω, \mathcal{F}, P) un espacio de probabilidad, $x, x' \in \mathbb{R}$ y $F(x)$ función de distribución acumulada. Así, F satisface las siguientes propiedades:

1. $\lim_{x \rightarrow \infty} F(x) = 1$
2. $\lim_{x \rightarrow -\infty} F(x) = 0$
3. Si $x \leq x'$, entonces $F(x) \leq F(x')$. En este caso se dice que F es no decreciente.

El comportamiento de una variable aleatoria queda, en general, descrito por su distribución de probabilidad, o función de densidad de probabilidad $f(x)$. La cual se desarrolla con base en ciertas hipótesis que se suponen válidas para los fenómenos aleatorios. Es por ello que a continuación se muestra una de las distribuciones de mayor uso de todas las distribuciones continuas de probabilidad denominada **distribución normal o gaussiana**:

Definición 1.15. *Se dice que una variable aleatoria ξ se distribuye normalmente con parámetros (μ, σ) , y se denota por $\xi \sim N(\mu, \sigma)$, si su función de distribución acumulada $\Phi_{\mu, \sigma}(x)$ está dada por*

$$\Phi_{\mu, \sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x \exp\left\{-\frac{(u - \mu)^2}{2\sigma^2}\right\} du$$

con $\sigma > 0$. Si $\mu = 0$ y $\sigma = 1$ se dice que ξ tiene una distribución normal estándar.

Otra de las distribuciones continuas que será de gran interés y apoyo para lo que sigue es la distribución exponencial.

Definición 1.16. *Se dice que una variable aleatoria ξ tiene **distribución exponencial** con parámetro $\lambda > 0$, y se denota por $\xi \sim \exp(\lambda)$, si su función de distribución acumulada $F(x)$ está dada por*

$$F(x) = \int_{-\infty}^x \lambda e^{-\lambda u} du, \text{ para cada } x > 0.$$

1.1.4. Esperanza

Con lo anterior no termina el estudio de medidas para las variables aleatorias, existe medidas características para la distribución de una variable aleatoria, dividiéndose estas en medidas de centralización y medidas de dispersión.

Para poder continuar con el estudio de este tipo de medidas, primero se recordará la definición de la integral de Stieltjes para una función de distribución acumulada $F(x)$ y una función $f(x)$ continua en un intervalo $[a, b]$ de \mathbb{R} . Se divide el intervalo $[a, b]$ en n subintervalos $[x_i, x_{i+1}]$ de modo que $a = x_0 < x_1 < x_2 < \dots < x_n = b$ y calcular una suma

$$S := \sum_{i=1}^n f(\tilde{x}_i)[F(x_i) - F(x_{i-1})]$$

donde \tilde{x}_i es cualquier número del intervalo $[x_i, x_{i+1}]$. Si la suma S tiende a un límite finito como

$$\max_{1 \leq i \leq n} |x_{i+1} - x_i| \rightarrow 0, \quad n \rightarrow \infty,$$

y este límite no depende de la sucesión particular de particiones y en la elección de los puntos x_i , entonces se llama la integral de Stieltjes de $f(x)$ con respecto a la función de distribución $F(x)$ y se denota por

$$\int_a^b f(x) dF(x).$$

En lo que sigue asumimos que la integral de Stieltjes de una función $f(x)$ con respecto a una función $F(x)$ existe si y sólo si la correspondiente integral de la función $|f(x)|$ existe. Por definición,

$$\int_{-\infty}^{\infty} f(x) dF(x) = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x) dF(x).$$

En este sentido, la principal medida de centralización de la distribución de una variable aleatoria es la media o bien **esperanza matemática** representada por $E(\cdot)$ la cual representa el promedio. Es decir, el valor promedio de una variable aleatoria después de un número grande de experimentos, es su valor esperado.

Definición 1.17. Sea ξ una variable aleatoria definida en un espacio de probabilidad (Ω, \mathcal{F}, P) . La **esperanza matemática** (o valor medio o, simplemente, media) de la variable aleatoria continua ξ es el número,

$$\mathbb{E}(\xi) := \int_{-\infty}^{\infty} x dF(x),$$

donde, $F(x)$ es la función de distribución acumulada de ξ .

Nota 1.10. Si ξ es una variable aleatoria continua con función de densidad de probabilidad $f(x)$, la integral en la definición (1.17) de esperanza se reduce a calcular la integral usual de Riemann,

$$\mathbb{E}(\xi) := \int_{-\infty}^{\infty} x f(x) dx.$$

A continuación se presenta una lista de propiedades que cumple la esperanza.

Propiedades 1.4. Sean ξ, ξ_1 y ξ_2 variables aleatorias definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) .

1. $\mathbb{E}(c) = c$, para cada constante c .
2. $\mathbb{E}(c\xi) = c\mathbb{E}(\xi)$, para cada constante c .
3. Si $\xi \geq 0$, entonces, $\mathbb{E}(\xi) \geq 0$.
4. $\mathbb{E}(\xi_1 + \xi_2) = \mathbb{E}(\xi_1) + \mathbb{E}(\xi_2)$, si la esperanza matemática de $\mathbb{E}(\xi_1)$ y $\mathbb{E}(\xi_2)$ existen.
5. Si ξ_1 y ξ_2 son variables aleatorias independientes entonces,

$$\mathbb{E}(\xi_1\xi_2) = \mathbb{E}(\xi_1)\mathbb{E}(\xi_2).$$

Observe que la segunda y cuarta propiedad establecen que la esperanza matemática es lineal, es decir, separa sumas y separa multiplicaciones por constantes. Así como también en la quinta propiedad se observa una aplicación del concepto de independencia en el cálculo de la esperanza.

Nota 1.11. El recíproco de la quinta propiedad de la lista de propiedades (1.4) no es válido en general, es decir, la condición $\mathbb{E}(\xi_1\xi_2) = \mathbb{E}(\xi_1)\mathbb{E}(\xi_2)$ no es suficiente para concluir que ξ_1 y ξ_2 sean independientes.

1.1.5. Varianza

La esperanza matemática por sí sola no proporciona una completa descripción de la distribución de la variable aleatoria. Además de conocer en qué valor se centra esa distribución es importante determinar la dispersión o variación de los valores de la variable aleatoria en torno a la esperanza. Para ello se define la **varianza**, denotada por σ^2 o $\text{Var}(\cdot)$.

Definición 1.18. La **varianza** de una variable aleatoria ξ se define por la fórmula,

$$\text{Var}(\xi) := \mathbb{E}[\xi - \mathbb{E}(\xi)]^2.$$

Al igual que la esperanza, la varianza también cuenta con propiedades básicas las cuales se enumeran a continuación.

Propiedades 1.5. Sean ξ, ξ_1 y ξ_2 variables aleatorias definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) .

1. $\text{Var}(\xi) \geq 0$.
2. $\text{Var}(c) = 0$, para cada constante c .
3. $\text{Var}(c\xi) = c^2\text{Var}(\xi)$, para cada constante c .
4. $\text{Var}(\xi + c) = \text{Var}(\xi)$, para cada constante c .
5. $\text{Var}(\xi) = \mathbb{E}(\xi^2) - \mathbb{E}(\xi)^2$.
6. Si ξ_1 y ξ_2 son variables aleatorias independientes entonces,

$$\text{Var}(\xi_1 + \xi_2) = \text{Var}(\xi_1) + \text{Var}(\xi_2).$$

Nota 1.12. En general, $\text{Var}(\xi_1 + \xi_2) \neq \text{Var}(\xi_1) + \text{Var}(\xi_2)$.

De las propiedades básicas (1.5) se obtiene en particular que la varianza es siempre una cantidad no negativa y que no cumple la propiedad de linealidad, pues en general no separa sumas y cuando aparecen constantes como producto, las constantes se extraen de la varianza elevándolas al cuadrado.

El siguiente resultado haciendo uso de la varianza, proporciona una cota superior para la probabilidad de que una variable aleatoria tome un valor que diste de su media en más de un factor de su desviación estándar.

Propiedades 1.6. Sea ξ una variable aleatoria no negativa y sea ϵ un número arbitrario positivo. Se tiene la siguiente desigualdad,

$$P\{\xi \geq \epsilon\} \leq \frac{\mathbb{E}(\xi)}{\epsilon}.$$

Esta última desigualdad implica, que para cada variable aleatoria ξ , la **desigualdad de Chebyshev** se define como sigue,

$$P\{|\xi - \mathbb{E}(\xi)| \geq \epsilon\} \leq \frac{\text{Var}(\xi)}{\epsilon^2}.$$

Observe que el parámetro ϵ que aparece en la desigualdad de Chebyshev debe ser en realidad estrictamente mayor a la varianza de la variable aleatoria ξ pues

de lo contrario, si $0 < \epsilon \leq \text{Var}(\xi)$, entonces $\text{Var}(\xi)/\epsilon^2 \geq 1$ y tal cantidad no proporciona ninguna información útil como cota superior para una probabilidad. También observar que la desigualdad de Chebyshev es óptima en el sentido de que sin hipótesis adicionales puede alcanzarse la cota superior.

1.1.6. Variables aleatorias multidimensionales y sus distribuciones

Hasta ahora, el interés se ha centrado en una sola variable aleatoria de tipo continuo. Existen problemas en los que deben estudiarse simultáneamente dos o más variables aleatorias. A fin de resolver este tipo de problemas es fundamental el estudio de variables aleatorias multidimensionales o también llamadas vectores aleatorios de los tipos discreto y continuo.

Un vector aleatorio es discreto, si todas las variables aleatorias que lo conforman son discretas. Un vector aleatorio es continuo, si todas las variables aleatorias que lo conforman son continuas. Para el objetivo de esta tesis, se considerará únicamente vectores aleatorios continuos.

Definición 1.19. Sean $\xi_1, \xi_2, \dots, \xi_n$ variables aleatorias definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) . El vector $(\xi_1, \xi_2, \dots, \xi_n)$ se dice que es una **variable aleatoria n -dimensional** o bien un **vector aleatorio**.

Definición 1.20. Sea $(\xi_1, \xi_2, \dots, \xi_n)$ un vector aleatorio definido en el espacio de probabilidad (Ω, \mathcal{F}, P) . La **función de distribución acumulada n -dimensional** está dada por,

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = P\{\xi_1 \leq x_1, \dots, \xi_n \leq x_n\},$$

para $x_i \in \mathbb{R}$, $i = 1, 2, \dots, n$.

Definición 1.21. Sea $(\xi_1, \xi_2, \dots, \xi_n)$ un vector aleatorio n -dimensional continuo. Se dice que la función integrable y no negativa $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow [0, \infty)$, es la **función de densidad conjunta** de la variable aleatoria n -dimensional $(\xi_1, \xi_2, \dots, \xi_n)$ o simplemente **función de densidad** de $(\xi_1, \xi_2, \dots, \xi_n)$ si para todo (x_1, \dots, x_n) en \mathbb{R}^n se cumple la igualdad,

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

La integral múltiple anterior representa el volumen bajo la superficie dada por la función $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ sobre la región que se encuentra a la izquierda y abajo del punto (x_1, \dots, x_n) . Toda función de densidad $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$ de estas características satisface las siguientes dos propiedades.

Propiedades 1.7.

1. $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) \geq 0$.

2.

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) dx_1 \cdots dx_n = 1.$$

Recíprocamente, se dice que una función $f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow [0, \infty)$ es **función de densidad conjunta** de la variable aleatoria n -dimensional (ξ_1, \dots, ξ_n) si cumple con las dos condiciones en propiedades (1.7).

Nota 1.13. Las funciones de densidad de probabilidad de cada una de las ξ_i individuales, denominadas **funciones de densidad marginal**, se obtienen integrando las otras componentes. Por ejemplo,

$$f_{\xi_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, u_2, \dots, u_n) du_n \cdots du_2,$$

es la función de densidad de probabilidad marginal de la componente ξ_1 . Similarmente,

$$f_{\xi_1, \xi_2}(x_1, x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, u_3, \dots, u_n) du_n \cdots du_3,$$

es la función de densidad marginal conjunta de (ξ_1, ξ_2) .

Definición 1.22. Sea $(\xi_1, \xi_2, \dots, \xi_n)$ un vector aleatorio definido en el espacio de probabilidad (Ω, \mathcal{F}, P) , con función de distribución conjunta $F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)$. Supóngase que las respectivas funciones de distribución marginales están dadas por $F_{\xi_1}(x_1), \dots, F_{\xi_n}(x_n)$. Se dice que las variables aleatorias $\xi_1, \xi_2, \dots, \xi_n$ son **mutuamente independientes** si,

$$F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \cdots F_{\xi_n}(x_n),$$

para cada vector $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Alternativamente, puede definirse la independencia en términos de la función de densidad como sigue.

Nota 1.14. Sea $(\xi_1, \xi_2, \dots, \xi_n)$ un vector aleatorio continuo definido en el espacio de probabilidad (Ω, \mathcal{F}, P) y las variables aleatorias $\xi_1, \xi_2, \dots, \xi_n$ mutuamente independientes, entonces las funciones de densidad de las variables están relacionadas por la fórmula,

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = f_{\xi_1}(x_1) \cdots f_{\xi_n}(x_n),$$

para cada vector $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Adicionalmente se tiene la siguiente extensión del concepto de independencia de variables aleatorias.

Nota 1.15. Se dice que las variables aleatorias de una sucesión infinita $\{\xi_k\}$ son independientes si la igualdad anterior de la definición 1.22 se cumple para cualquier n . Es decir, que un conjunto infinito de variables aleatorias es independiente si cualquier subconjunto finito de él lo es.

De manera más general, la influencia de una variable aleatoria en la estructura de probabilidad de otra se caracteriza por la **función de densidad condicional**.

Definición 1.23. Sea (ξ_1, ξ_2) una variable aleatoria bidimensional con función de densidad conjunta $f_{\xi_1, \xi_2}(\cdot, \cdot)$ y funciones de densidad de probabilidad marginales $f_{\xi_1}(\cdot)$ y $f_{\xi_2}(\cdot)$. Entonces,

1. La **densidad condicional** de ξ_1 , dada $\xi_2 = x_2$, si $f_{\xi_2}(x_2) > 0$, se denota como $f_{\xi_1|\xi_2}(\cdot | \cdot)$ y está dada por,

$$f_{\xi_1|\xi_2}(x_1 | \xi_2 = x_2) = \frac{f_{\xi_1, \xi_2}(x_1, x_2)}{f_{\xi_2}(x_2)}.$$

2. La **densidad condicional** de ξ_2 , dada $\xi_1 = x_1$, si $f_{\xi_1}(x_1) > 0$, se denota como $f_{\xi_2|\xi_1}(\cdot | \cdot)$ y está dada por,

$$f_{\xi_2|\xi_1}(x_2 | \xi_1 = x_1) = \frac{f_{\xi_1, \xi_2}(x_1, x_2)}{f_{\xi_1}(x_1)}.$$

1.1.7. Esperanza condicional

Así como fue de interés calcular probabilidades condicionales. A menudo interesa calcular esperanzas cuando hay alguna información parcial disponible, por lo que la esperanza deseada es la **esperanza condicional**.

Definición 1.24. Sea (ξ_1, ξ_2) una variable aleatoria bidimensional continua con función de densidad condicional de ξ_2 dado $\xi_1 = x_1$ definida como $f_{\xi_2|\xi_1}(\cdot | \xi_1 = x_1)$, y función de densidad de probabilidad de ξ_1 , $f_{\xi_1}(\cdot)$. Para todo x_1 tal que $f_{\xi_1}(x_1) > 0$ se define la **esperanza condicional** de ξ_2 dado $\xi_1 = x_1$ por

$$\mathbb{E}(\xi_2 | \xi_1 = x_1) = \int_{-\infty}^{\infty} x_2 f_{\xi_2|\xi_1}(x_2 | \xi_1 = x_1) dx_2.$$

1.1.8. Covarianza

La varianza de una variable aleatoria es una medida de su variabilidad, y la covarianza de dos variables aleatorias es una medida de su variabilidad conjunta, o su grado de asociación. Después de definir la covarianza, se desarrollaran algunas de sus propiedades y discutirá una medida de asociación llamada correlación, que se define en términos de covarianza.

Definición 1.25. Sean ξ_1 y ξ_2 variables aleatorias definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) . La **covarianza** de ξ_1 y ξ_2 , que se denota como $\text{Cov}(\xi_1, \xi_2)$ o σ_{ξ_1, ξ_2} está dada por,

$$\text{Cov}(\xi_1, \xi_2) = \mathbb{E}\{[\xi_1 - \mathbb{E}(\xi_1)][\xi_2 - \mathbb{E}(\xi_2)]\}.$$

Nota 1.16. La covarianza de ξ_1 y ξ_2 , con función de densidad conjunta $f_{\xi_1, \xi_2}(x_1, x_2)$ está dada por,

$$\text{Cov}(\xi_1, \xi_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - \mathbb{E}(\xi_1)][x_2 - \mathbb{E}(\xi_2)] f_{\xi_1, \xi_2}(x_1, x_2) dx_1 dx_2.$$

La covarianza pocas veces se calcula a partir de la definición (1.25). En vez de ello, se aplica la fórmula de cálculo siguiente.

Teorema 1.2. La fórmula para el cálculo de covarianza es,

$$\text{Cov}(\xi_1, \xi_2) = \mathbb{E}(\xi_1 \xi_2) - \mathbb{E}(\xi_1) \mathbb{E}(\xi_2).$$

1.1.9. Correlación

La covarianza entre dos variables aleatorias dígase ξ_1 y ξ_2 sólo brinda una indicación de la relación de esas variables. No se pretende que describa el tipo

o intensidad de tal relación. Es frecuente que interese conocer si dos variables aleatorias guardan relación lineal o no lo hacen. La covarianza a menudo se re-escala para obtener una medida en un intervalo fijo. Esto conduce al coeficiente de correlación.

Definición 1.26. Sean ξ_1 y ξ_2 variables aleatorias definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) . La **correlación** de ξ_1 y ξ_2 , que se denota como $\rho_{\xi_1, \xi_2} = \text{Corr}(\xi_1, \xi_2)$ está dada por,

$$\text{Corr}(\xi_1, \xi_2) = \frac{\text{Cov}(\xi_1, \xi_2)}{\sqrt{\text{Var}(\xi_1)\text{Var}(\xi_2)}}.$$

El coeficiente de correlación tiene una interpretación similar a la covarianza.

Nota 1.17. El coeficiente de correlación $\text{Corr}(\xi_1, \xi_2)$ de dos variables aleatorias ξ_1 y ξ_2 cualesquiera es tal que $-1 \leq \text{Corr}(\xi_1, \xi_2) \leq 1$.

Nota 1.18. En general para un vector aleatorio $\xi = (\xi_1, \dots, \xi_n)$, es posible recopilar en un conjunto toda la información sobre varianzas y covarianzas entre cada par de variables, este conjunto será una matriz, a la cual se le llama **matriz de varianzas y covarianzas**, definida como,

$$\Sigma = \begin{bmatrix} \sigma_{1,1} & \cdots & \sigma_{1,n} \\ \vdots & \ddots & \vdots \\ & \sigma_{j,i} & \ddots \\ \sigma_{n,1} & \cdots & \sigma_{n,n} \end{bmatrix},$$

donde $\sigma_{i,i} = \text{Var}(\xi_i)$ y $\sigma_{i,j} = \text{Cov}(\xi_i, \xi_j)$ para $i \neq j$ con $i, j = 1, \dots, n$.

1.1.10. Distribución normal multivariada

Como en el caso univariante, hay familias estándar de distribuciones de probabilidad para vectores aleatorios. En particular, el análogo multivariante de la distribución normal es la **distribución normal multivariada**

Definición 1.27. Sea $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ un vector aleatorio continuo definido en el espacio de probabilidad (Ω, \mathcal{F}, P) . Con vector de medias $\mu = (\mu_1, \dots, \mu_n)$ y matriz de varianzas y covarianzas Σ . La distribución de la variable aleatoria continua n -dimensional ξ se dice ser **normal multivariada**, denotada $\xi \sim \text{DNM}_n(\mu, \Sigma)$, si esta tiene función de densidad conjunta de la forma,

$$f_{\xi}(x) = (\det(\Sigma))^{-1/2} (2\pi)^{-n/2} \exp \left\{ -\frac{1}{2} (x - \mu)^t \Sigma^{-1} (x - \mu) \right\},$$

donde $x \in \mathbb{R}^n$ y $\det(\Sigma)$ es el determinante de la matriz Σ .

1.1.11. Teoremas límite

Se darán dos de los teoremas límite esenciales en probabilidad: la ley de los grandes números y el teorema central del límite.

Para lo que sigue, será útil definir la convergencia de una variable aleatoria, pero antes de ello se dará la definición de **convergencia de una sucesión**.

Definición 1.28. *Una sucesión $\{x_k\}$ de números reales es convergente a un número x si para cada $\epsilon > 0$ existe un número natural N a partir del cual los elementos de la sucesión se encuentran cercanos al número x , es decir, para $n \geq N$,*

$$|x_n - x| < \epsilon.$$

La variedad de formas en las que puede definirse la convergencia de variables aleatorias estará dada por las formas en las que se decida medir la cercanía de la sucesión con el límite a través de la medida de probabilidad.

El primer tipo de convergencia que se definirá de las variables aleatorias es cuando una sucesión numérica es convergente para cada elemento del espacio muestral fijo.

Definición 1.29. *Sea $\{\xi_k\}$ una sucesión de variables aleatorias definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) . Para cada $\omega \in \Omega$ fijo, la sucesión $\xi_1(\omega), \xi_2(\omega), \dots$ es una sucesión de números reales. En este caso la variable aleatoria límite se define de forma puntual $\xi(\omega) := \lim_{n \rightarrow \infty} \xi_n(\omega)$. A este tipo de convergencia se le llama **convergencia puntual** y se escribe,*

$$\xi_n \rightarrow \xi,$$

para cada $\omega \in \Omega$.

Un tipo de convergencia un tanto menos estricto que el anterior ocurre cuando se permite que la convergencia puntual se observe sobre un conjunto de probabilidad uno.

Definición 1.30 (Convergencia casi segura). Sea $\{\xi_k\}$ una sucesión de variables aleatorias definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) . Se dice que la sucesión $\{\xi_k\}$ converge **casi seguramente**, o **casi dondequiera**, a la variable aleatoria ξ si para casi toda ω , $\xi_n(\omega)$ converge a $\xi(\omega)$, en símbolos,

$$P(\omega \in \Omega : \xi_n(\omega) \rightarrow \xi(\omega)) = 1,$$

y se escribe,

$$\xi_n \xrightarrow{\text{c.s.}} \xi.$$

De este modo se permite que exista un subconjunto de Ω en donde no se verifique la convergencia, pero tal subconjunto debe tener medida de probabilidad cero. Es claro que si una sucesión de variables aleatorias es convergente puntualmente, entonces es también convergente en el sentido casi seguro. El recíproco es falso. Otra forma aún menos restrictiva que la convergencia casi segura es la siguiente.

Definición 1.31. Sea $\{\xi_k\}$ una sucesión de variables aleatorias definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) . Se dice que la sucesión $\{\xi_k\}$ **converge en probabilidad** a la variable aleatoria ξ si para cualquier $\epsilon > 0$,

$$P(\omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| > \epsilon) \rightarrow 0,$$

cuando n tiende a infinito. En este caso se escribe,

$$\xi_n \xrightarrow{p} \xi.$$

Finalmente el último tipo de convergencia que se mencionara hace uso de las funciones de distribución de las variables aleatorias.

A menudo es difícil realizar cálculos exactos con las distribuciones de probabilidad, esto podría ser porque la distribución es desconocida. Es por esto que puede ser posible aproximarse a la verdadera distribución por una distribución más simple. Esto requiere una definición de convergencia de variables aleatorias, es la convergencia en distribución que se define como sigue.

Definición 1.32. Sea $\{\xi_k\}$ una sucesión de variables aleatorias definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) , con funciones de distribución $F_{\xi_1}, F_{\xi_2}, \dots$ respectivamente. Se dice que la sucesión $\{\xi_k\}$ **converge en distribución**, o que

converge débilmente, a la variable aleatoria ξ si para todo punto x en donde $F_\xi(x)$ la función de distribución de ξ es continua se cumple que cuando $n \rightarrow \infty$,

$$F_{\xi_n}(x) \rightarrow F_\xi(x).$$

Es decir, para aquellos valores reales x que cumplan la condición mencionada, debe verificarse que,

$$\lim_{n \rightarrow \infty} P\{\xi_n \leq x\} = P\{\xi \leq x\}.$$

En tal caso se escribe,

$$\xi_n \xrightarrow{d} \xi.$$

Existen otros tipos de convergencia para variables aleatorias pero los que se han mencionado son suficientes para poder enunciar algunos teoremas límite importantes en probabilidad.

El teorema conocido como la **ley de los grandes números** es un resultado que puede observarse en la naturaleza. Constituye uno de los resultados de la teoría de la probabilidad y tiene mucha relevancia en las aplicaciones tanto teóricas como prácticas. Este teorema establece que bajo ciertas condiciones, el promedio aritmético de variables aleatorias converge a una constante cuando el número de sumandos crece a infinito.

Teorema 1.3. *Sea $\{\xi_k\}$ una sucesión de variables aleatorias mutuamente independientes y con la misma distribución, definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) , con media finita μ . Entonces, cuando $n \rightarrow \infty$,*

$$\frac{1}{n} \sum_{i=1}^n \xi_i \rightarrow \mu.$$

La convergencia se verifica en el sentido casi seguro (ley fuerte) y también en probabilidad (ley débil).

La ley de límite más famosa en estadística es el **teorema de límite central**, tiene diversas aplicaciones para simplificar el cálculo de ciertas probabilidades y aproximar algunas distribuciones, declarado aquí en su forma más simple.

Teorema 1.4. *Sea $\{\xi_k\}$ una sucesión de variables aleatorias mutuamente independientes y con la misma distribución de la población (idénticamente distribuidas)*

definidas en un espacio de probabilidad común (Ω, \mathcal{F}, P) , con media μ finita y varianza σ^2 positiva. Entonces,

$$Z_n = \frac{(\xi_1 + \cdots + \xi_n) - n\mu}{\sqrt{n\sigma^2}},$$

tiende en distribución a la función de distribución normal estándar cuando n tiende a infinito.

Observación 1.1. Obsérvese que dividiendo el numerador y denominador por n , y definiendo $\bar{\xi} = (\xi_1 + \cdots + \xi_n)/n$, la variable Z_n puede escribirse de la siguiente forma,

$$Z_n = \frac{\bar{\xi} - \mu}{\sqrt{\sigma^2/n}}. \quad (1.1)$$

Es interesante observar también que la ley de los grandes números asegura que el numerador de la ecuación (1.1) converge a cero conforme n tiende a infinito, sin embargo el denominador de esta expresión también converge a cero y estos límites ocurren de tal manera que este cociente no es constante sino una variable aleatoria con distribución normal estándar.

1.2. Procesos estocásticos

Hasta ahora, se ha abordado el estudio de los conceptos básicos de teoría de la probabilidad, así como de las variables aleatorias enfocado al caso continuo a través de un planteamiento incremental de una a n variables. Lo que sigue es considerar los procesos estocásticos, estos serán una familia de variables aleatorias definidas en un espacio de probabilidad y, por lo tanto, será de gran interés definir una ley de probabilidad en el conjunto de trayectorias del proceso. Más específicamente, los procesos estocásticos generalizan la noción de vectores (finitos dimensionales) de variables aleatorias al caso de cualquier familia de variables aleatorias indexadas en un conjunto general. Típicamente, este último conjunto representa el “tiempo” y es un intervalo de \mathbb{R} (en el caso continuo) o \mathbb{N} (en el caso discreto). Así como en la sección anterior se enfocó el estudio en variables aleatorias continuas del mismo modo para esta sección el tiempo que se estudiara será el caso continuo.

Definición 1.33. Sean (Ω, \mathcal{F}, P) espacio de probabilidad, T un conjunto de índices llamado espacio parametral y $S \subseteq \mathbb{R}$ con $(S, \mathcal{B}(\mathbb{R}))$ un espacio medible. Un

proceso estocástico en (Ω, \mathcal{F}, P) es una familia $(X_t)_{t \in T}$ de variables aleatorias $X_t : (\Omega, \mathcal{F}) \rightarrow (S, \mathcal{B}(\mathbb{R}))$ para $t \in T$.

El espacio (Ω, \mathcal{F}, P) se denomina **espacio de probabilidad subyacente del proceso** $(X_t)_{t \in T}$, mientras que $(S, \mathcal{B}(\mathbb{R}))$ es el **espacio de estado** o el **espacio de fase**.

Nota 1.19.

- Al fijar $t \in T$, la variable aleatoria X_t es el **estado del proceso** en el tiempo t .
- Un proceso estocástico, también llamado **proceso aleatorio**, puede considerarse como una función de dos variables, para todo $\omega \in \Omega$, el mapeo $X(\cdot, \omega) : t \in T \rightarrow X_t(\omega) \in S$ se llama la **trayectoria** o **trayectoria del proceso** correspondiente a ω , es decir, a la pareja (t, ω) se le asocia el valor o estado $X(t, \omega)$, lo cual también puede escribirse como $X_t(\omega)$.
- Cualquier trayectoria $X(\cdot, \omega)$ del proceso pertenece al espacio S^T de funciones definidas en T y con valores en S .

Nota 1.20. Si A es un conjunto de estados, el evento $\{\omega \in \Omega : X_n(\omega) \in A\}$, corresponde a la situación en donde al tiempo n el proceso toma algún valor dentro del conjunto A . En particular, $\{\omega \in \Omega : X_n(\omega) = x\}$ es el evento en donde al tiempo n el proceso se encuentra en el estado x . Considerando distintos tiempos, se estará interesado en eventos de la forma $\{\omega \in \Omega : X_{n_1}(\omega) = x_1, X_{n_2}(\omega) = x_2, \dots, X_{n_k}(\omega) = x_k\}$.

A partir de este momento el espacio parametral se considera como el conjunto continuo $T = [0, \infty)$. Se dice entonces que el proceso es a tiempo continuo, y se denota por $\{X_t : t \geq 0\}$.

Los diferentes tipos de procesos estocásticos se obtienen al considerar las distintas posibilidades para el espacio parametral, el espacio de estados, las características de las trayectorias, y principalmente las relaciones de dependencia entre las variables aleatorias que conforman el proceso. A continuación se darán algunos de los procesos estocásticos que se pueden estudiar. Estos son procesos que cumplen una cierta propiedad particular, no necesariamente excluyentes unas de otras.

1.2.1. Procesos con incrementos independientes

Definición 1.34. Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geq 0\}$ tiene incrementos independientes si para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$, las variables $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes. Esto quiere decir que los desplazamientos que tiene el proceso en estos intervalos disjuntos de tiempo son independientes unos de otros.

1.2.2. Procesos estacionarios

Definición 1.35. Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geq 0\}$ es estacionario en el sentido estricto si para cualesquiera tiempos t_1, \dots, t_n , la distribución del vector $(X_{t_1}, \dots, X_{t_n})$ es la misma que la del vector $(X_{t_1+h}, \dots, X_{t_n+h})$ para cualquier valor de $h > 0$. En particular, la distribución de X_t es la misma que la de X_{t+h} para cualquier $h > 0$.

1.2.3. Procesos con incrementos estacionarios

Definición 1.36. Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geq 0\}$ tiene incrementos estacionarios si para cualesquiera tiempos $s < t$, y para cualquier $h > 0$, las variables $X_{t+h} - X_{s+h}$ y $X_t - X_s$ tienen la misma distribución de probabilidad. Es decir, el incremento que tiene el proceso entre los tiempos s y t sólo depende de estos tiempos a través de la diferencia $t - s$, y no de los valores específicos de s y t .

1.2.4. Cadena de Markov a tiempo continuo

Para definir una cadena de Markov en donde el tiempo es continuo y las variables toman valores enteros, primero se definirá un proceso de saltos.

Definición 1.37. Consideremos un proceso a tiempo continuo $\{X_t : t \geq 0\}$ que inicia en un estado i_1 al tiempo cero. El proceso permanece en ese estado un tiempo aleatorio T_{i_1} después salta a un nuevo estado i_2 distinto del i_1 anterior. El sistema permanece ahora en el estado i_2 un tiempo aleatorio T_{i_2} al cabo del cual brinca a otro estado i_3 distinto del inmediato anterior, y así sucesivamente. Los tiempos aleatorios T son los tiempos en los que el proceso permanece constante en alguno de sus estados, y se llaman tiempos de estancia. Los momentos en donde el

proceso tiene saltos son los tiempos $W_n = T_{i_1} + \cdots + T_{i_n}$, para $n \geq 1$. El **proceso de saltos** se define de la siguiente manera,

$$X_t = \begin{cases} i_1 & \text{si } 0 \leq t < W_1, \\ i_2 & \text{si } W_1 \leq t < W_2, \\ i_3 & \text{si } W_2 \leq t < W_3, \\ \vdots & \end{cases}$$

Sin pérdida de generalidad suponer que el espacio de estados es el conjunto,

$$S = \{0, 1, \dots\}$$

y que el tiempo de estancia asociado el estado i es la variable aleatoria T_i , la cual se supone positiva con función de distribución $F_i(t)$. Se denotará por p_{ij} a la probabilidad de que la cadena pase del estado i al estado j al efectuar un salto. Adicionalmente $p_{ii} = 0$, y con ello se imposibilita que la cadena salte al mismo estado de partida. Las probabilidades de saltos deben satisfacer las siguientes condiciones:

1. $p_{ij} \geq 0$.
2. $p_{ii} = 0$.
3. $\sum_j p_{ij} = 1$.

Además, se supone que los tiempos de estancia T_{i_1}, T_{i_2}, \dots son independientes entre sí, y también son independientes de la elección del estado j al cual la cadena salta después de estar en cualquier otro estado i . Mas aún, suponerse que cada variable T_i es finita con probabilidad uno, o bien, es infinita con probabilidad uno. En el primer caso se dice que el estado i es no absorbente, y en el segundo caso que es absorbente. El hecho de que $T_i = \infty$ se interpreta en el sentido de que el proceso deja de saltar y permanece en el estado i el resto del tiempo, es decir, el estado i es absorbente. sólo hay dos tipos de estados: absorbentes o no absorbentes. En otras palabras, con probabilidad uno el tiempo de estancia es finito o con probabilidad uno es infinito.

Teorema 1.5. *Un proceso de las características arriba especificadas satisface la propiedad de Markov si, y sólo si, los tiempos de estancia en los estados no absorbentes tienen distribución exponencial.*

Suponga que el tiempo de estancia en un estado no absorbente i tiene distribución $\exp(\lambda_i)$, con $\lambda_i > 0$, es decir,

$$F_i(t) = 1 - e^{-\lambda_i t}, \text{ para } t > 0.$$

Observe que puede considerarse que $\lambda_i = 0$ en el caso cuando $T_i = \infty$.

Suponga que las probabilidades de transición son estacionarias en el tiempo, esto significa que para cada $s \geq 0$ y $t \geq 0$, la probabilidad $P(X_{t+s} = j \mid X_s = i)$ es idéntica a $P(X_t = j \mid X_0 = i)$, es decir, no hay dependencia del valor de s . Esta probabilidad se escribe de manera breve mediante la expresión $p_{ij}(t)$, para i y j enteros no negativos. Es decir,

$$p_{ij}(t) = P(X_{t+s} = j \mid X_s = i) = P(X_t = j \mid X_0 = i).$$

Definición 1.38. *A un proceso de saltos con las características y postulados arriba señalados se le llama **cadena de Markov a tiempo continuo**.*

1.2.5. Proceso de Poisson

Proceso de Poisson homogéneo

Existen varias maneras de formular la definición de proceso de Poisson homogéneo, todas ellas equivalentes entre sí. Aquí se enuncia una de estas definiciones. Como consecuencia directa de la cadena de Markov a tiempo continuo, se tiene la definición más usual y funcional de proceso de Poisson homogéneo.

Definición 1.39. *Una sucesión de variables aleatorias $\{N(t) : t \geq 0\}$ definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) se llama **proceso de Poisson (homogéneo)** con intensidad $\lambda > 0$ si satisface las siguientes propiedades:*

- i) $P(N(0) = 0) = 1$, esto es, $N(0)$ es siempre 0.*
- ii) Para cualesquiera $s > 0, t \geq 0$, tales que $0 < s < t$, $N(t) - N(s)$ tiene distribución de Poisson de parámetro $\lambda(t - s)$. Alternativamente la distribución de $N(t + s) - N(t)$ es igual a la de $N(s)$.*
- iii) Para cualesquiera $t_1, \dots, t_n \geq 0$, tales que $0 \leq t_1 < \dots < t_n$, $n \geq 1$, es decir, para todo conjunto finito de tiempos, las variables aleatorias $N(t_n) - N(t_{n-1}), \dots, N(t_2) - N(t_1), N(t_1)$, son variables aleatorias independientes.*

La definición previa establece que el proceso tiene las propiedades de que comienza en cero (i), tiene incrementos estacionarios (ii) y sus incrementos son independientes (iii).

Observese que para cada $t, s > 0$, la variable aleatoria X_t tiene distribución de Poisson de parámetro λt . Por el inciso (ii) se tiene que $X_t - X_0$ tiene distribución de Poisson con parámetro $\lambda(t-0)$ y por el inciso (i) se tiene que X_t tiene distribución de Poisson con parámetro λt . Ahora, si $0 < s < t$ se tiene que por el inciso (ii) que $X_t - X_s$ tiene distribución de Poisson con parámetro $\lambda(t-s)$. Con esto se observa que la distribución del incremento $X_t - X_s$ y la de la variable aleatoria X_{t-s} es la misma. Además esto dice que $P((X_t - X_s) \geq 0) = 1$, pues

$$P((X_t - X_s) \geq 0) = \sum_{k=0}^{\infty} P((X_t - X_s) = k) = 1,$$

esto es, la distribución sólo depende de la longitud del intervalo $t - s$. Esta propiedad se conoce como propiedad de incrementos estacionarios. Nótese además que esta propiedad garantiza que el proceso es no decreciente.

Propiedades 1.8.

- La probabilidad de que ocurra un evento en un intervalo de tiempo pequeño es proporcional al tamaño del intervalo,

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

- La probabilidad de ocurrencia de dos o más eventos en un intervalo muy pequeño es cero,

$$\lim_{h \rightarrow 0} \frac{P(N(h) \geq 2)}{h} = 0.$$

Proposición 1.1. *Supongamos que $N(t)$ es el número de llegadas en el intervalo de tiempo $[0, t]$, que forma un proceso de Poisson de intensidad λ . Entonces, la distribución de cada $N(t)$ es de Poisson con intensidad λt .*

Proceso de Poisson no homogéneo

Definición 1.40. *Una sucesión de variables aleatorias $\{N(t) : t \geq 0\}$ definidas en un espacio de probabilidad (Ω, \mathcal{F}, P) se llama **proceso de Poisson (no homogéneo)** con intensidad $\lambda(t) > 0, t \geq 0$ si satisface las siguientes propiedades:*

- i) $P(N(0) = 0) = 1$, esto es, $N(0)$ es siempre 0.
- ii) Para cualesquiera $t_1, \dots, t_n \geq 0$, tales que $0 \leq t_1 < \dots < t_n$, $n \geq 1$, es decir, para todo conjunto finito de tiempos, las variables aleatorias $N(t_n) - N(t_{n-1}), \dots, N(t_2) - N(t_1), N(t_1)$, son variables aleatorias independientes.

La definición previa garantiza que el proceso comienza en cero (i) y sus incrementos son independientes (ii).

Propiedades 1.9.

- $\lim_{h \rightarrow 0} \frac{P(\text{Exactamente un evento entre } t \text{ y } t+h)}{h} = \lim_{h \rightarrow 0} \frac{P([N(t+h)-N(t)]=1)}{h} = \lambda(t)$.
- $\lim_{h \rightarrow 0} \frac{P(\text{Dos o más eventos entre } t \text{ y } t+h)}{h} = \lim_{h \rightarrow 0} \frac{P([N(t+h)-N(t)] \geq 2)}{h} = 0$.

Definición 1.41. El valor medio del proceso de Poisson no homogéneo se define como,

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

En particular, si $\lambda(t) = \lambda$, constante, entonces $m(t) = \lambda t$.

Propiedades 1.10. Para cada $t \geq 0$ y $s > 0$ se tiene que $N(t+s) - N(t)$ es una variable aleatoria de Poisson con media

$$m(t+s) - m(t) = \int_t^{t+s} \lambda(x) dx.$$

Corolario 1.1. Si $\lambda(t) = \lambda$ (es constante), $N(t+s) - N(t)$ es una variable aleatoria de Poisson con media λt .

1.2.6. Procesos gaussianos

Este tipo de procesos se le pondrá mayor atención debido a que será de gran utilidad para el desarrollo de los siguientes capítulos.

Definición 1.42. Se dice que un proceso estocástico a tiempo continuo $\{X_t : t \geq 0\}$ es un **proceso gaussiano** si para cualesquiera colección finita de tiempos t_1, \dots, t_n , el vector $(X_{t_1}, \dots, X_{t_n})$ tiene distribución normal multivariada o gaussiana multivariada. Equivalentemente el proceso es gaussiano si cada combinación lineal $\alpha_1 X_{t_1} + \dots + \alpha_n X_{t_n}$, para $\alpha_i \in \mathbb{R}$ con $i = 1, \dots, n$, tiene distribución normal univariada.

Cada proceso gaussiano se describe de manera única por sus dos parámetros, las funciones de media y covarianza, dadas respectivamente por,

$$\mu(t) = \mathbb{E}(X_t), \text{ con } t \geq 0,$$

y

$$\Gamma(s, t) = \text{Cov}(X_s, X_t) = \mathbb{E}[(X_s - \mathbb{E}(X_s))(X_t - \mathbb{E}(X_t))] \text{ con } s, t \geq 0.$$

La función de covarianza es definida positiva en el sentido de que por cada $n = 1, 2, \dots$, números reales $\alpha_1, \dots, \alpha_n$ y elementos $t_1, \dots, t_n \geq 0$,

$$\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \Gamma(t_i, t_j) \geq 0.$$

Nota 1.21. *Dada una función de media arbitraria $\mu(t)$ y una función de covarianza $\Gamma(s, t)$, positiva definida, entonces, existe un correspondiente proceso gaussiano.*

Nota 1.22. *El movimiento browniano es el proceso gaussiano único que tiene trayectorias continuas, función de media $\mu(t) = 0$ y covarianza $\Gamma(s, t) = \sigma^2 \min\{s, t\}$, para $s, t \geq 0$.*

Los procesos gaussianos también surgen como límites de sumas normalizadas de funciones aleatorias independientes e idénticamente distribuidas. Sean $\xi_1(t), \xi_2(t), \dots$ variables aleatorias independientes e idénticamente distribuidas, $\mu(t) = E(\xi(t))$ y $\Gamma(s, t) = \text{Cov}(\xi(s), \xi(t))$ funciones de media y covarianza respectivamente. Motivado por el teorema del límite central, se define

$$X_N(t) = \frac{\sum_{i=1}^N (\xi_i(t) - \mu(t))}{\sqrt{N}}.$$

El teorema del límite central establece que la distribución de X_N converge a la distribución normal para cada punto de tiempo fijo t . Una extensión multivariante del teorema de límite central afirma que para cualquier conjunto finito de puntos de tiempo (t_1, \dots, t_n) , el vector aleatorio

$$(X_N(t_1), \dots, X_N(t_n)),$$

tiene, en el límite para N grande, una distribución normal multivariada. No es difícil creer, entonces, que en circunstancias ordinarias, los procesos estocásticos $\{X_N(t) : t \geq 0\}$ convergerían, en un sentido apropiado, a un proceso gaussiano $\{X(t) : t \geq 0\}$ cuya media es cero y cuya función de covarianza es $\Gamma(s, t)$. A esto se le llama el **principio de límite central para funciones aleatorias**.

Capítulo 2

Teoría de valores extremos

El objetivo de este capítulo es proporcionar conceptos y resultados de la teoría de valores extremos, la cual se ocupa del comportamiento estocástico del mínimo o máximo de variables aleatorias independientes e idénticamente distribuidas, y de sus propiedades. Por lo general, estos son menos estudiados que los promedios. En algunas aplicaciones, como en hidrología, contaminación, corrosión, medio ambiente, meteorología, seguros, aplicaciones financieras, geología, oceanografía, ingeniería eólica, entre otras, es necesario extrapolar los datos de los niveles observados a niveles no observados, para dar predicciones del fenómeno. Esta teoría proporciona modelos que permiten tal extrapolación. La teoría de valores extremos proporciona una base teórica sólida y un marco para la extrapolación, fundamentada en teoría asintótica, ya que en muchas aplicaciones es necesario que se haga de esa manera, para poder obtener resultados que sirvan para predecir fenómenos que puedan causar daños irreversibles.

La teoría a desarrollar en este capítulo es extraída de “An introduction to statistical modeling of extreme values” de [Coles *et al.* \(2001\)](#), “Extreme value theory: an introduction” de [de Haan y Ferreira \(2006\)](#), “Análisis de Valores Extremos” de [Villanueva e Ibáñez \(2016\)](#), “Precipitaciones Máximas en el Estado de Guanajuato, México” de [Moreno y Ortega \(2013\)](#), “Análisis de máximos para datos espaciales de lluvias” de [González y Sánchez \(2011\)](#) y “Extreme value and related models with applications in engineering and science” de [Castillo *et al.* \(2005\)](#).

2.1. Teoría de valores extremos clásica

La teoría de valores extremos surge como el desarrollo de un modelo en el cual a partir de una muestra de variables aleatorias es posible derivar nuevas en base a los máximos o mínimos de estas.

Como resultado central en la teoría de valores extremos se encuentra la demostración de Fréchet, quien identificó una distribución límite posible para valores máximos y luego Fisher y Tippett, demostraron que sólo es posible tres familias paramétricas de distribuciones límites para máximos y sus equivalentes para mínimos.

2.1.1. Formulación del modelo

El modelo para el que se desarrolla la teoría de valores extremos se centra en describir el comportamiento estadístico de

$$M_n = \text{máx}\{X_1, \dots, X_n\},$$

donde X_1, \dots, X_n es una sucesión de variables aleatorias independientes que tienen una distribución común F y M_n representa el máximo del proceso de n unidades de tiempo de observación.

La distribución de M_n puede derivarse de manera exacta para todos los valores de n a partir de la distribución de las n variables, teniendo en cuenta las propiedades de independencia:

$$\begin{aligned} P\{M_n \leq z\} &= P\{X_1 \leq z, \dots, X_n \leq z\} \\ &= P\{X_1 \leq z\} \cdots P\{X_n \leq z\} \\ &= \{F(z)\}^n \\ &= F^n(z). \end{aligned}$$

Sin embargo, esto no es inmediatamente útil en la práctica, ya que la función de distribución F es desconocida. Esto lleva a buscar familias de modelos (distribuciones) que sirvan de aproximación para F^n . Antes de proseguir véase el comportamiento de F^n cuando $n \rightarrow \infty$.

Observación 2.1. Defínase $z_+ = \sup\{z : F(z) < 1\}$, de modo que z_+ es el punto

final superior de F . Cuando $n \rightarrow \infty$,

$$M_n \xrightarrow{p} z_+.$$

Es decir, M_n converge en probabilidad a la constante z_+ .

Además se tiene, siempre que $n \rightarrow \infty$, dos casos:

- Para cualquier $z < z_+$, $F^n(z) \rightarrow 0$.
- Para cualquier $z \geq z_+$, $F^n(z) \rightarrow 1$.

En este caso se tiene que F^n siempre converge a una distribución degenerada.

Ejemplo 2.1. Considere la distribución acumulada exponencial $(1 - \exp\{-\frac{x}{\theta}\})$, $x > 0$) con parámetro $\theta = 9$. La Figura (2.1) muestra el comportamiento de F^n para diferentes valores de n . Note que cuando $n \rightarrow \infty$ esta converge a una distribución degenerada.

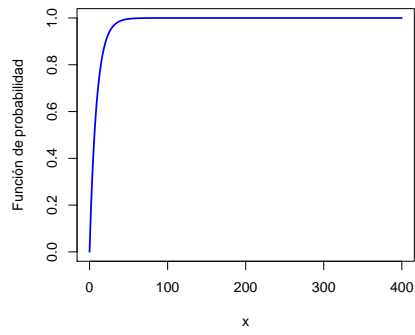
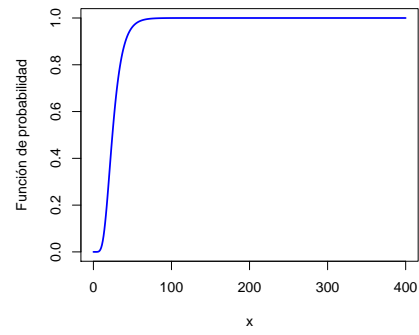
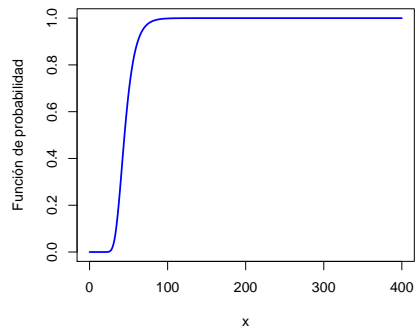
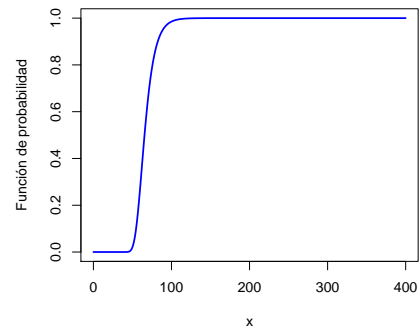
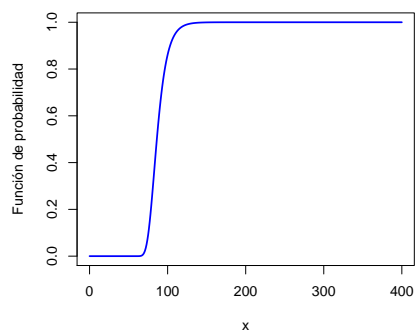
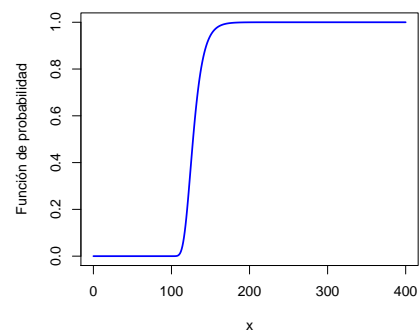
(a) $n=1$ (b) $n=10$ (c) $n=100$ (d) $n=1000$ (e) $n=10000$ (f) $n=100000$

Figura 2.1: Simulación de la distribución acumulada exponencial degenerada.

Esta dificultad se evita permitiendo una normalización lineal de la variable M_n , el cual es uno de los intereses principales de la teoría de valores extremos.

Nota 2.1. *Suponga que existe una secuencia de constantes $\{a_n > 0\}$ y $\{b_n\}$, tal que la normalización para M_n está dada por,*

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

Elecciones apropiadas de $\{a_n\}$ y $\{b_n\}$ estabilizan la ubicación y la escala de M_n^* a medida que aumenta n , evitando las dificultades que surgen con la variable M_n .

Así, la distribución de M_n^* , cuando $n \rightarrow \infty$, siempre converge a una distribución no degenerada, es decir,

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

Por lo tanto, se busca distribuciones límite para M_n^* con elecciones apropiadas de $\{a_n\}$ y $\{b_n\}$ en lugar de M_n , el cual es otro de los problemas que se encuentran en la teoría de valores extremos determinar las posibles distribuciones límite. Una respuesta a este problema se presenta en el siguiente resultado, conocido como el teorema de valores extremos o de Fisher-Tippet (1928).

2.1.2. Teorema de valores extremos

Las únicas posibles distribuciones límite para M_n^* se muestran en el **teorema de valores extremos**.

Teorema 2.1. *Dada una sucesión de variables aleatorias X_1, \dots, X_n independientes que tienen una distribución común F y $M_n = \max\{X_1, \dots, X_n\}$. Si existen sucesiones de constantes $\{a_n > 0\}$ y $\{b_n\}$ tales que*

$$P \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} = F^n(a_n z + b_n) \rightarrow G(z), \text{ cuando } n \rightarrow \infty,$$

donde G es una distribución no degenerada, entonces G pertenece a alguna de las siguientes tres familias de **distribuciones de valores extremos**:

I: Gumbel $G(z) = \exp \left\{ -\exp \left[-\left(\frac{z-b}{a} \right) \right] \right\}$, $-\infty < z < \infty$;

II: Fréchet

$$G(z) = \begin{cases} 0 & \text{si } z \leq b, \\ \exp \left\{ - \left(\frac{z-b}{a} \right)^{-\alpha} \right\} & \text{si } z > b; \end{cases}$$

III: Weibull

$$G(z) = \begin{cases} \exp \left\{ - \left[- \left(\frac{z-b}{a} \right)^\alpha \right] \right\} & \text{si } z < b, \\ 1 & \text{si } z \geq b. \end{cases}$$

Para los parámetros $a > 0$, b y, en el caso de las familias II y III, $\alpha > 0$.

Cada distribución tiene un parámetro de ubicación y escala, b y a respectivamente; adicionalmente, los tipos Fréchet y Weibull tienen un parámetro de forma α .

El teorema (2.1) de valores extremos implica que cuando M_n se puede estabilizar con sucesiones adecuadas $\{a_n\}$ y $\{b_n\}$, la variable normalizada correspondiente M_n^* tiene una distribución límite que debe ser uno de los tres tipos de distribución de valores extremos.

Además, notar que este teorema no garantiza la existencia de las constantes de normalización $\{a_n\}$ y $\{b_n\}$, puesto que, en general dichas constantes podrían no existir.

Nota 2.2. *La característica notable de este resultado es que los tres tipos de distribuciones de valores extremos son los únicos límites posibles para las distribuciones M_n^* , independientemente de la distribución F para la población.*

2.1.3. Distribución generalizada de valores extremos

Los tres tipos de límite que surgen en el teorema 2.1 tienen distintas formas que corresponde a las diferentes formas para la función de distribución F de las variables X_i . Esto puede hacerse preciso considerando el comportamiento de la distribución límite G en z_+ , su punto final superior.

Nota 2.3. *Para la distribución de Weibull, z_+ es finita, mientras que para las distribuciones Fréchet y Gumbel, $z_+ = \infty$. Sin embargo, la densidad de G decae exponencialmente para la distribución de Gumbel y polinomialmente para la distribución de Fréchet.*

Para un mejor análisis de la distribución límite de M_n^* se requiere de una reformulación de los modelos en el teorema 2.1. Las familias Gumbel, Fréchet y

Weibull se pueden combinar en una sola familia de modelos evidenciado en el siguiente resultado.

Teorema 2.2. *La familia de distribuciones de valor extremo generalizada (DVE o DGVE) o distribución de Jenkinson-Von Misses se define como:*

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}, \quad (2.1)$$

definida en el conjunto $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$, donde los parámetros de localización, escala y forma, satisfacen, respectivamente $-\infty < \mu < \infty$, $\sigma > 0$ y $-\infty < \xi < \infty$.

La especificación de ξ determinará el comportamiento de la cola de la distribución, de forma que según el valor que tome este parámetro se tendrá una de las distribuciones del teorema 2.1 de valores extremos.

Nota 2.4. *Las clases de tipo II y tipo III de distribución de valores extremos corresponden respectivamente a los casos $\xi > 0$ y $\xi < 0$ en esta parametrización. El subconjunto de la familia DGVE con $\xi = 0$ se interpreta como el límite de la ecuación (2.1) cuando $\xi \rightarrow 0$, lo que lleva a la familia Gumbel con la función de distribución*

$$G(z) = \exp \left[- \exp \left\{ - \left(\frac{z - \mu}{\sigma} \right) \right\} \right], \text{ con } -\infty < z < \infty.$$

Con esto en mente, se tiene que el teorema 2.1 de valores extremos se replantea de la siguiente manera.

Teorema 2.3. *Si existen sucesiones de constantes $\{a_n > 0\}$ y $\{b_n\}$ tal que*

$$P \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} \rightarrow G(z) \text{ cuando } n \rightarrow \infty.$$

Para una distribución G no degenerada, entonces G pertenece a la familia de distribuciones DGVE,

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}.$$

definida en el conjunto $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$, donde los parámetros de localización, escala y forma, satisfacen, respectivamente $-\infty < \mu < \infty$, $\sigma > 0$ y

$-\infty < \xi < \infty$.

Interpretando el límite en el teorema 2.3 como una aproximación para valores grandes de n sugiere el uso de la familia DGVE para modelar la distribución de máximos de sucesiones largas. La dificultad aparente de que las constantes de normalización serán desconocidas, en la práctica es posible resolverlo de la siguiente forma.

Nota 2.5. *Suponer que existen sucesiones de constantes $\{a_n > 0\}$ y $\{b_n\}$ tal que*

$$P \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} \rightarrow G(z) \text{ cuando } n \rightarrow \infty. \quad (2.2)$$

Para una distribución G no degenerada. De otra manera,

$$P \left\{ \frac{M_n - b_n}{a_n} \leq z \right\} \approx G(z)$$

para n lo suficientemente grande. Equivalentemente,

$$\begin{aligned} P\{M_n \leq z\} &\approx G \left\{ \frac{z - b_n}{a_n} \right\} \\ &= G^*(z), \end{aligned}$$

donde G^ es otro miembro de la familia DGVE.*

En otras palabras, si el teorema 2.3 permite la aproximación de la distribución M_n^* por un miembro de la familia DGVE para n grande, la distribución M_n también se puede aproximar por un miembro diferente de la misma familia, simplemente estimando los parámetros de ubicación y de escala.

El concepto de máx-estabilidad, que se introduce a continuación, está estrechamente relacionado con el teorema 2.3.

Definición 2.1. *Se dice que una distribución G es máx-estable, si para cada $n = 2, 3, \dots$, hay constantes $\alpha_n > 0$ y β_n tales que:*

$$G^n(\alpha_n z + \beta_n) = G(z).$$

Dado que G^n es la función de distribución de $M_n = \max\{X_1, \dots, X_n\}$, donde las X_i son variables independientes cada una con función de distribución G , la

máx-estabilidad es una propiedad que satisfacen las distribuciones para las cuales la operación de tomar máximos muestrales conduce a una distribución idéntica, con distintos parámetros de escala y ubicación. La conexión con las leyes de límite de valor extremo se realiza mediante el siguiente resultado.

Teorema 2.4. *Una distribución es máx-estable si y sólo si es una distribución de valores extremos generalizada.*

2.1.4. Demostración del teorema de valores extremos

Hasta este momento se han dado propiedades y observaciones sobre el teorema (2.1) de valores extremos, con lo enunciado hasta ahora se dará una demostración a este teorema.

El teorema (2.4) se usa directamente en la prueba del teorema de valores extremos, y aquí a continuación la prueba.

Demostración. (del teorema de valores extremos (2.1)) La idea es considerar M_{nk} , la variable aleatoria máxima en una sucesión de $n \times k$ variables para algún valor grande de n . Esto puede considerarse como el máximo de una única sucesión de $n \times k$, o como el máximo de k máximos.

Cada uno de los cuales es el máximo de n observaciones. Más precisamente, suponga que la distribución límite de $(M_n - b_n)/a_n$ es G . Entonces; para n suficientemente grande, por el teorema (2.3),

$$P\{(M_n - b_n)/a_n \leq z\} \approx G(z).$$

Por lo tanto, para cualquier número entero k , dado que nk es grande,

$$P\{(M_{nk} - b_{nk})/a_{nk} \leq z\} \approx G(z). \quad (2.3)$$

Pero dado que M_{nk} es el máximo de k variables que tienen la misma distribución que M_n ,

$$P\{(M_{nk} - b_n)/a_n \leq z\} = [P\{(M_n - b_n)/a_n \leq z\}]^k. \quad (2.4)$$

Por lo tanto, por (2.3) y (2.4) respectivamente,

$$P\{M_{nk} \leq z\} \approx G\left(\frac{z - b_{nk}}{a_{nk}}\right)$$

y

$$P\{M_{nk} \leq z\} \approx G^k \left(\frac{z - b_n}{a_n} \right).$$

Por la unicidad del límite, G y G^k son idénticos aparte de los coeficientes de ubicación y de escala, esto es, $G(z) = G^k(\alpha_k z + \beta_k)$. De ello se deduce que G es máx-estable y, por lo tanto, un miembro de la familia DGVE por teorema 2.4. \square

2.1.5. Ejemplos

Dada una función de distribución F , cómo establecer si realmente se puede lograr la convergencia de la distribución M normalizada. Además de ello se tiene preguntas adicionales:

¿Qué opciones de sucesiones de normalización a_n y b_n son necesarias? y ¿qué miembro de la familia DGVE se obtiene como límite?

La consideración principal es la inferencia estadística de los datos reales para los que se desconoce la distribución subyacente F , se dará sólo unos pocos ejemplos que ilustran cómo la elección cuidadosa de las sucesiones de normalización conduce a una familia DGVE, como lo implica el teorema 2.3.

Ejemplo 2.2. Si X_1, X_2, \dots es una sucesión de variables aleatorias independientes con distribución exponencial estándar $\text{Exp}(1)$, $F(x) = 1 - \exp\{-x\}$ para $x > 0$. En este caso, dejando $a_n = 1$ y $b_n = \log(n)$,

$$\begin{aligned} P\{(M_n - b_n)/a_n \leq z\} &= F^n(z + \log(n)) \\ &= \{1 - \exp[-(z + \log(n))]\}^n \\ &= \{1 - n^{-1} \exp[-z]\}^n \\ &\rightarrow \exp\{-\exp[-z]\}, \end{aligned}$$

como $n \rightarrow \infty$, para $z \in \mathbb{R}$ fija. Por lo tanto, con la elección de a_n y b_n , la distribución límite de M_n cuando $n \rightarrow \infty$ es la distribución Gumbel, correspondiente a $\xi = 0$ en la familia DGVE.

Ejemplo 2.3. Si X_1, X_2, \dots , es una sucesión de variables aleatorias independientes con distribución Fréchet estándar, $F(x) = \exp(-\frac{1}{x})$ para $x > 0$. Dejar $a_n = n$

y $b_n = 0$,

$$\begin{aligned} P\{(M_n - b_n)/a_n \leq z\} &= F^n(nz) \\ &= \left\{ \exp \left[-\frac{1}{nz} \right] \right\}^n \\ &= \exp \left(-\frac{1}{z} \right), \end{aligned}$$

cuando $n \rightarrow \infty$, para cada $z > 0$ fijo. Por lo tanto, el límite en este caso es un resultado exacto para todo n , porque la distribución estándar de Fréchet es la estabilidad máxima de $F: \xi = 1$ en la familia DGVE.

Ejemplo 2.4. Si X_1, X_2, \dots , son variables aleatorias independientes con distribución uniforme $U(0, 1)$, $F(x) = x$ para $0 \leq x \leq 1$. Para $z < 0$ fijo, supongamos $n > -z$ y sea $a_n = \frac{1}{n}$ y $b_n = 1$. Entonces,

$$\begin{aligned} P\{(M_n - b_n)/a_n \leq z\} &= F^n(n^{-1}z + 1) \\ &= \left(1 + \frac{z}{n} \right)^n \\ &\rightarrow \exp(z), \end{aligned}$$

como $n \rightarrow \infty$. Por lo tanto la distribución límite es de tipo Weibull, con $\xi = -1$ en la familia DGVE.

Hay cierta libertad en la elección de $\{a_n\}$ y $\{b_n\}$ en tales ejemplos. Sin embargo, las diferentes opciones que conducen a un límite no degenerado siempre producen una distribución límite en la familia DGVE con el mismo valor de ξ , aunque posiblemente con otros valores de los parámetros de ubicación y de escala.

2.1.6. Modelos de máximos por bloques

La familia de distribuciones DGVE será útil para modelar la distribución de los máximos por bloques. El procedimiento consiste en agrupar los datos (una serie de observaciones independientes X_1, X_2, \dots) en bloques de sucesiones de observaciones de longitud n , para un gran valor de n , generando una serie de bloques máximos, $M_{n,1}, \dots, M_{n,m}$, digamos a la que se puede adaptar la distribución DGVE.

El principal problema que presenta este método reside en la elección del tamaño

de los bloques, para la cual habrá que encontrar un equilibrio entre el sesgo y la varianza.

La elección de bloques muy pequeños conducirá a una pobre aproximación del modelo, con lo que se aumentará el sesgo al estimar y extrapolar. Por el contrario, la elección de bloques demasiado grandes aumentará la varianza de las estimaciones. Por cuestiones prácticas, en sucesiones de datos temporales mensuales, a menudo, los bloques se eligen para que correspondan a un período de tiempo de un año, en cuyo caso n es el número de observaciones en un año, de esta manera los máximos se distribuyen de manera similar en cada uno de los bloques.

2.1.7. Cuantiles extremos

Se denomina niveles de retorno a los cuantiles de la distribución de valores extremos generalizada; se denotan como z_p y se obtendrán como la inversa de la función de distribución DGVE.

Teorema 2.5. *La estimación de los cuantiles extremos de la distribución máxima en un período de tiempo esta dada por*

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - \{-\log(1-p)\}^{-\xi}] & \text{para } \xi \neq 0, \\ \mu - \sigma \log\{-\log(1-p)\} & \text{para } \xi = 0, \end{cases} \quad (2.5)$$

donde $G(z_p) = 1 - p$.

Nota 2.6. *El termino, z_p es el nivel de retorno asociado con el período de retorno $\frac{1}{p}$, ya que con un grado razonable de precisión, se espera que el nivel z_p se exceda en promedio una vez cada $\frac{1}{p}$ años. Más precisamente, z_p se excede por el máximo anual en cualquier año particular con probabilidad p .*

Dado que los cuantiles permiten que los modelos de probabilidad se expresen en la escala de datos, la relación de modelo DGVE a sus parámetros es más fácil de interpretar en términos de las expresiones cuantiles (2.5). En particular, definir $y_p = -\log(1-p)$, así que

$$z_p = \begin{cases} \mu - \frac{\sigma}{\xi} [1 - y_p^{-\xi}] & \text{para } \xi \neq 0, \\ \mu - \sigma \log y_p & \text{para } \xi = 0. \end{cases}$$

Nota 2.7. *Se deduce que, si z_p se representa gráficamente frente a y_p en una escala logarítmica o equivalente, si y_p se representa gráficamente frente a $\log y_p$, la representación gráfica es lineal en el caso $\xi = 0$. Si $\xi < 0$ la representación gráfica es convexa con un límite asintótico como $p \rightarrow 0$ en $\frac{\mu-\sigma}{\xi}$; si $\xi > 0$ la gráfica es cóncava y no tiene límite finito. Este gráfico es conocido como un **gráfico de nivel de retorno**.*

Debido a la simplicidad de la interpretación y a que la elección de la escala comprime la cola de la distribución de modo que se resalta el efecto de la extrapolación, los gráficos de nivel de retorno son prácticamente convenientes tanto para la presentación como para la validación del modelo.

2.2. Teoría de valores extremos multivariante

Anteriormente, se ha hecho el estudio de modelos para valores extremos de un sólo proceso. Ahora conviene poner atención al estudio de valores extremos multivariados. Al estudiar los extremos de dos o más procesos, cada proceso individual puede modelarse utilizando técnicas univariantes, pero existen argumentos sólidos para estudiar también las interrelaciones de valor extremo.

Primero, puede ser que alguna combinación de los procesos sea de mayor interés que los procesos individuales en sí mismos; segundo, en un modelo multivariante, existe la posibilidad de que los datos de cada variable informen inferencias en cada uno de los otros.

Así como en el caso univariado, los valores extremos multivariados también dependen de la justificación asintótica. Además, el uso de modelos en teoría de valores extremos multivariados conduce a problemas que se han discutido anteriormente en el contexto univariado, como lo son los modelos de exceso de umbral máximo y de bloque para valores extremos.

Para dar un gran paso al caso n -dimensional, primeramente se dará una especial atención al caso bidimensional o bivariado. Esto permite resaltar los conceptos y problemas principales sin poderse enredar en la complejidad de la notación que requeriría un tratamiento multivariado completo.

2.2.1. Caso bidimensional

Suponga que $(X_1, Y_1), (X_2, Y_2), \dots$ es una sucesión de vectores que son versiones independientes de un vector aleatorio que tiene la función de distribución $F(x, y)$.

Como en el caso univariante, la teoría clásica para caracterizar el comportamiento extremo de los extremos multivariados se basa en el comportamiento límite de los máximos de bloques. Esto requiere una nueva definición.

Definición 2.2. *Si se define*

$$M_{x,n} = \max_{i=1,\dots,n} \{X_i\} \quad y \quad M_{y,n} = \max_{i=1,\dots,n} \{Y_i\},$$

entonces

$$M_n = (M_{x,n}, M_{y,n}), \quad (2.6)$$

es el vector de las componentes máximas.

El siguiente teorema propone una caracterización asintótica de la distribución de valores extremos, a partir de la función de distribución conjunta límite de $M_n^* = (M_{x,n}/n, M_{y,n}/n)$, cuando $n \rightarrow \infty$, proporcionando un análogo bivariado del teorema 2.1 de valores extremos.

Teorema 2.6. *Sea $M_n^* = (M_{x,n}^*, M_{y,n}^*)$, con (X_i, Y_i) , $i = 1, \dots, n$ vectores independientes con distribución marginal Fréchet. Entonces, si para $n \rightarrow \infty$*

$$P\{M_{x,n}^* \leq x, M_{y,n}^* \leq y\} \xrightarrow{d} G(x, y), \quad (2.7)$$

donde G es una función de distribución no degenerada, se tiene que G es de la forma,

$$G(x, y) = \exp\{-V(x, y)\}, \quad x > 0, \quad y > 0, \quad (2.8)$$

donde,

$$V(x, y) = 2 \int_0^1 \max\left(\frac{w}{x}, \frac{1-w}{y}\right) dH(w), \quad (2.9)$$

y H es una función de distribución en $[0, 1]$, que satisface la condición

$$\int_0^1 w dH(w) = 1/2.$$

La familia de distribuciones que surgen como límites en (2.7) se denomina **clase de distribuciones bivariadas de valores extremos**.

A diferencia del caso univariante, esta distribución engloba a infinitas familias de distribuciones. En seguida se muestran algunos ejemplos.

Ejemplo 2.5. Si se distribuye la masa de la función H en los puntos $w = 0$ y $w = 1$, se obtiene,

$$V(x, y) = x^{-1} + y^{-1},$$

y la correspondiente distribuciones bivariadas de valores extremos generalizada será,

$$G(x, y) = \exp\{-(x^{-1} + y^{-1})\}, \quad x > 0, y > 0.$$

Esta función se puede factorizar en función de x y y , por lo que ambas variables son independientes.

Ejemplo 2.6. Si se concentra toda la masa de la función H en $w = 0.5$, se obtiene,

$$G(x, y) = \exp\{-\max(x^{-1}, y^{-1})\}, \quad x > 0, y > 0,$$

la cual es la función de distribución de variables que son marginalmente estándar de Frechet, pero que son perfectamente dependientes: $X = Y$ con probabilidad 1.

La clase completa de funciones de distribución de valores extremos bivariate se puede obtener mediante una generalización de las distribuciones marginales. Específicamente, tomando

$$\tilde{x} = \left[1 + \xi_x \left(\frac{x - \mu_x}{\sigma_x}\right)\right]^{1/\xi_x} \quad \text{y} \quad \tilde{y} = \left[1 + \xi_y \left(\frac{y - \mu_y}{\sigma_y}\right)\right]^{1/\xi_y},$$

se obtiene

$$G(x, y) = \exp\{-V(\tilde{x}, \tilde{y})\},$$

donde la función V satisface la ecuación (2.9) para cualquier H .

Las distribuciones marginales serán distribuciones generalizadas de valores extremos (DGVE) con parámetros (μ_x, σ_x, ξ_x) y (μ_y, σ_y, ξ_y) respectivamente.

La función V será **homogénea de orden -1**, en el sentido de que para cualquier constante $a > 0$ se tiene

$$\begin{aligned}
V(a^{-1}x, a^{-1}y) &= 2 \int_0^1 \text{máx} \left(\frac{w}{a^{-1}x}, \frac{1-w}{a^{-1}y} \right) \\
&= 2a \int_0^1 \text{máx} \left(\frac{w}{x}, \frac{1-w}{y} \right) \\
&= aV(x, y)
\end{aligned}$$

Usando la propiedad de que V es homogénea de orden -1 , en la ecuación (2.7) se obtiene que

$$G^n(x, y) = G(n^{-1}x, n^{-1}y), \quad (2.10)$$

para $n = 2, 3, \dots$

Nota 2.8. Si (X, Y) tiene la función de distribución G , entonces M_n también tiene la función de distribución G , aparte de un reescalado por n^{-1} .

Por lo tanto, G posee una versión multivariada de la propiedad máx-estable presentada en la definición (2.1). Un argumento similar al de la definición (2.1) implica que las distribuciones límite en la ecuación (2.7) deben tener esta propiedad de máx-estabilidad y, como en el caso univariante aquí, este argumento forma la base de una prueba del teorema (2.6). Es decir, de la ecuación (2.10), distribuciones del tipo ecuación (2.8) tienen la propiedad de máx-estabilidad, y se puede demostrar que son las únicas distribuciones que tienen esta propiedad, sujetas a la especificación marginal. Aunque el teorema (2.6) proporciona una caracterización completa de las distribuciones de límites bivariados, la clase de límites posibles es amplia, estando limitada sólo por la ecuación (2.9).

2.2.2. Caso multidimensional

En la subsección anterior, se dio el estudio de valores extremos bivariado. Ahora, se introduce la generalización al caso multivariado. Afortunadamente, la teoría del valor extremo también está bien desarrollada para el caso multivariante y es posible usar este conocimiento para resolver muchos problemas prácticos. Desafortunadamente, los resultados únicos para el caso univariante de tener una sola familia como el único límite posible se descarta, dado que en el caso multivariante, se tiene un conjunto mucho más amplio de posibilidades.

Formulación del modelo

Considérese $\{\mathbf{X}_n\}_{n \geq 1}$ una sucesión de variables aleatorias k -dimensionales y se denotará como \mathbf{x} al vector $(x_1, \dots, x_k) \in \mathbb{R}^k$. La j -ésima componente de \mathbf{X}_i , $i \in \{1, \dots, n\}$ se denota por $X_{i,j}$, $j \in \{1, \dots, k\}$. Se usará a partir de ahora y en lo sucesivo letras en negrita para denotar vectores de datos multivariados.

Definición 2.3. *Supóngase que se tiene un conjunto de n vectores aleatorios, $\mathbf{X}_1, \dots, \mathbf{X}_n$, independientes e idénticamente distribuidos, de modo que el **vector de valores máximos o máximo multivariado** se define como el vector \mathbf{M}_n tal que cada una de sus componentes son los máximos de las componentes de los vectores $\mathbf{X}_1, \dots, \mathbf{X}_n$, es decir,*

$$\mathbf{M}_n = \left(\max_{1 \leq i \leq n} X_{i,1}, \dots, \max_{1 \leq i \leq n} X_{i,k} \right).$$

Tomando, $M_{n,j} = \max_{1 \leq i \leq n} X_{i,j}$, se tiene que,

$$\mathbf{M}_n = (M_{n,1}, \dots, M_{n,k}).$$

Observación 2.2. *Sea \mathbf{M}_n los vectores de máximos, es decir, los vectores cuyos componentes son los máximos de las componentes. La distribución de \mathbf{M}_n puede derivarse de manera exacta para todos los valores de n a partir de la distribución de las n variables aleatorias, teniendo en cuenta las propiedades de independencia:*

$$\begin{aligned} P\{\mathbf{M}_n \leq \mathbf{x}\} &= P\{\max_{1 \leq i \leq n} X_{i,1} \leq x_1, \dots, \max_{1 \leq i \leq n} X_{i,k} \leq x_k\} \\ &= P\{X_{1,1} \leq x_1, \dots, X_{n,1} \leq x_1, \dots, X_{1,k} \leq x_k, \dots, X_{n,k} \leq x_k\} \\ &= P\{X_{1,1} \leq x_1, \dots, X_{1,k} \leq x_1, \dots, X_{n,1} \leq x_k, \dots, X_{n,k} \leq x_k\} \\ &= F(x_1, \dots, x_k) \cdots F(x_1, \dots, x_k) \\ &= \{F(x_1, \dots, x_k)\}^n \\ &= F^n(x_1, \dots, x_k). \end{aligned}$$

Definición 2.4. *Sean $\mathbf{x} \in \mathbb{R}^k$ y $\mathbf{X}_1, \dots, \mathbf{X}_n$ vectores aleatorios independientes e idénticamente distribuidos con distribución multivariada común $F(\mathbf{x})$. Sea \mathbf{M}_n el vector de máximos de los \mathbf{X}_i . Se dice que G es una **distribución de valores extremos multivariada** si existen sucesiones de vectores de constantes $\{\mathbf{a}_n = (a_{n,1}, \dots, a_{n,k})\}_{n \geq 1}$, $\{\mathbf{b}_n = (b_{n,1}, \dots, b_{n,k})\}_{n \geq 1}$ con $a_{n,i} > 0$, $b_{n,i} \in \mathbb{R}$, para cada*

$i \in \{1, \dots, n\}$, tal que

$$P \left\{ \frac{M_{n,1} - b_{n,1}}{a_{n,1}} \leq x_1, \dots, \frac{M_{n,k} - b_{n,k}}{a_{n,k}} \leq x_k \right\} = F^n(\mathbf{a}_n \mathbf{x} + \mathbf{b}_n) \xrightarrow{d} G(\mathbf{x}) \quad (2.11)$$

y además, G es no degenerada en el sentido de que todas sus marginales univariadas son no degeneradas.

Nota 2.9. Si dos funciones multivariadas F, G satisfacen (2.11) de la definición 2.4 de distribución de valores extremos multivariados, para ciertos valores de \mathbf{a}_n y \mathbf{b}_n se dice que F está en el dominio de atracción de G .

Para derivar la distribución límite de una cierta distribución multivariada será de gran utilidad definir el concepto de función de dependencia de n variables aleatorias.

Definición 2.5. Sea $F(\mathbf{x})$ la distribución de un vector \mathbf{X} k -variado con distribuciones marginales univariadas $F_i(x_i)$, $i \in \{1, \dots, n\}$. Se define la **función de dependencia** asociada a $F(\mathbf{x})$, denotada por $D_F(y_1, \dots, y_k)$, como

$$D_F(F_1(x_1), \dots, F_k(x_k)) = F(x_1, \dots, x_n) = F(\mathbf{x}). \quad (2.12)$$

Nota 2.10. Si las marginales unitarias $F_i(x_i)$ son estrictamente crecientes, la ecuación (2.12) de la función de dependencia en la definición 2.5 será,

$$D_F(y_1, \dots, y_k) = F(F_1^{-1}(y_1), \dots, F_k^{-1}(y_k)) = F(\mathbf{x}).$$

Teorema 2.7. Sean $\mathbf{X}_1, \dots, \mathbf{X}_n$ vectores aleatorios k -dimensionales con función de distribución común F , existen sucesiones de vectores de constantes $\{\mathbf{a}_n = (a_{n,1}, \dots, a_{n,k})\}_{n \geq 1}$, $\{\mathbf{b}_n = (b_{n,1}, \dots, b_{n,k})\}_{n \geq 1}$ con $a_{n,i} > 0$, $b_{n,i} \in \mathbb{R}$, para cada $i \in \{1, \dots, n\}$, tales que

$$P \left\{ \frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n} \leq \mathbf{x} \right\} \xrightarrow{d} G(\mathbf{x}),$$

con G no degenerada, si y sólo si cada marginal pertenece al dominio de atracción de alguna $G_j(x)$, con $j \in \{1, \dots, k\}$ y

$$\lim_{n \rightarrow \infty} D_{F^n} \left(y_1^{1/n}, \dots, y_k^{1/n} \right) = D_G(y_1, \dots, y_k).$$

Teorema 2.8. Una función de distribución d -dimensional G , es una distribución límite maximal (distribución límite de valores extremos) si y sólo si,

1. Sus distribuciones marginales univariadas pertenecen al dominio de atracción de $H_j(x)$.
2. Su función de dependencia $D_G(y_1, \dots, y_d)$ satisface la ecuación funcional

$$D_G^m \left(y_1^{1/m}, \dots, y_k^{1/m} \right) = D_G(y_1, \dots, y_k),$$

para todo $m \geq 1$.

Para obtener la estructura general de distribuciones límite para extremos multivariados y otros resultados, se derivará el proceso de Poisson no homogéneo asociado con los puntos de extremos multivariados. Para ello, primeramente será de gran utilidad la definición de la **representación espectral vía procesos puntuales de Poisson** para caracterizar las familias de distribuciones límites de extremos multivariados.

Teorema 2.9. *Sea $\{\mathbf{X}_n\}_{n \geq 1}$ una sucesión de vectores aleatorios independientes e idénticamente distribuidos k -dimensionales, con entradas no negativas y función de distribución común F , donde F pertenece al dominio de atracción de una distribución multivariada G . Supongamos que las distribuciones marginales de F son idénticamente Fréchet univariadas y considere el proceso puntual $P_n = \{\frac{\mathbf{X}_i}{n} : i = 1, \dots, n\}$, entonces P_n converge en distribución a un proceso Poisson no homogéneo P en $\mathbb{R}_+^k - \{0\}$ cuando $n \rightarrow \infty$, con medida de intensidad*

$$\lambda(dr \times d\mathbf{w}) = m \frac{dr}{r^2} dS(\mathbf{w}), \quad (2.13)$$

donde r_i y $\mathbf{w}_{i,j}$ son las coordenadas pseudo-polar y angular, respectivamente, es decir

$$r_i = \frac{1}{n} \sum_{j=1}^k X_{i,j}, \quad w_{i,j} = \frac{X_{i,j}}{\sum_{j=1}^k X_{i,j}}, \quad i = 1, \dots, n, \quad j = 1, \dots, k, \quad (2.14)$$

y S es una medida de probabilidad en el simplex unitario,

$$S_k = \left\{ (w_1, \dots, w_k) : \sum_{j=1}^k w_j = 1, w_j \geq 0, j = 1, \dots, k \right\},$$

que satisface la condición,

$$\int_{S_k} w_j dS(\mathbf{w}) = \frac{1}{k}, \quad j = 1, \dots, k. \quad (2.15)$$

Nota 2.11. En el teorema 2.9, se puede utilizar la medida de intensidad

$$\lambda^*(dr \times d\mathbf{w}) = m \frac{dr}{r^2} dS^*(\mathbf{w})$$

donde ahora, $\mathbf{w} = (w_1, \dots, w_{k-1})$ y

$$S_k^* = \left\{ (w_1, \dots, w_{k-1}) : \sum_{j=1}^{k-1} w_j \leq 1, w_j \geq 0, j = 1, \dots, k-1 \right\}$$

y S^* satisface la condición,

$$\int_{S_k^*} w_j dS^*(\mathbf{w}) = \frac{1}{k}, \quad j = 1, \dots, k-1.$$

El teorema (2.9) es relevante en la practica dado que afirma que para cualquier sucesión de vectores aleatorios $\{\mathbf{X}_n\}_{n \geq 1}$ existe una medida positiva y finita $S(\mathbf{w})$ que satisface la condición de la integral (2.15) del teorema (2.9) de representación espectral vía procesos puntuales Poisson y es tal que el proceso Poisson no homogéneo asociado tiene como intensidad la ecuación (2.13). Por tanto, dada $S(\mathbf{w})$ es posible calcular las probabilidades de ocurrencia del proceso Poisson en cualquier región dada A . Una de estas aplicaciones consiste en el siguiente corolario, que muestra la caracterización de todas las posibles distribuciones límite multivariadas.

Corolario 2.1. *Cualquier distribución límite de vectores de máximo normalizados, con marginales Fréchet unitarias, tiene la forma*

$$H(\mathbf{x}) = \exp[-v(\mathbf{x})]$$

donde,

$$v(\mathbf{x}) = k \int_{S_k} \max_{1 \leq j \leq k} \left(\frac{w_j}{x_j} \right) dS(\mathbf{w}),$$

para algún $S(\mathbf{w})$ medida de probabilidad en el simplex unitario

$$S_k = \left\{ (w_1, \dots, w_k) : \sum_{j=1}^k w_j = 1, w_j \geq 0, j = 1, \dots, k \right\}$$

que satisface la condición

$$\int_{S_k} w_j dS(\mathbf{w}) = \frac{1}{k}, j = 1, \dots, k.$$

Tomando en cuenta lo expuesto en los dos resultados anteriores se puede apreciar la importancia de contar con datos cuyas distribuciones marginales sean Fréchet unitarias. En la práctica esto no ocurre, por lo que resulta necesario realizar una transformación a los datos de modo que estos tengan las marginales deseadas. Es posible hacer dicha transformación sin perder generalidad debido a la relación entre las tres distribuciones de valores extremos mencionada con anterioridad.

Observación 2.3. Si X es una variable aleatoria que tiene DGVE con parámetros a, b y α , la variable aleatoria $Y = -(\log(F(X))^{-1})$ tiene una distribución Fréchet unitaria, es decir, si X tiene DGVE con parámetros a, b y α , entonces la variable aleatoria Y dada por

$$Y = \begin{cases} [1 - \frac{1}{\alpha} (\frac{X-b}{a})]^{-\alpha} & \text{para } \alpha \neq 0, \\ \exp(\frac{X-b}{a}) & \text{para } \alpha = 0. \end{cases}$$

sigue una distribución Fréchet unitaria.

Capítulo 3

Procesos máx-estables

Los procesos máx-estables son una extensión natural de la teoría de valores extremos multivariante al caso dimensional infinito, es decir, los extremos de los procesos estocásticos, y por lo tanto juegan un papel esencial en el modelado estadístico de extremos espaciales. Estos procesos fueron caracterizados completamente por de Haan (1984). Una descripción más precisa de los procesos máx-estables se puede obtener a partir de su representación espectral, esta se determina a partir de dos posibles enfoques. Una primera caracterización es dada por de Haan (1984), proporciona dicha representación considerando funciones de tipo determinístico. Una segunda caracterización de los procesos máx-estables es proporcionada por Schlather (2002) donde considera funciones aleatorias en lugar de funciones determinísticas.

Este capítulo tiene como objetivo, proporcionar una extensión de la propiedad de máx-estabilidad de procesos, dando lugar a los llamados procesos máx-estables. Los procesos máx-estables constituyen la base para el desarrollo de modelos que se presentaran en el siguiente capítulo.

Este capítulo está basado en “An introduction to statistical modeling of extreme values” de (Coles *et al.* (2001)), “Extreme value theory: an introduction” de de Haan y Ferreira (2006), “Análisis de Valores Extremos” de Villanueva e Ibáñez (2016), “Precipitaciones Máximas en el Estado de Guanajuato, México” de Moreno y Ortega (2013) y “Análisis de máximos para datos espaciales de lluvias” de González y Sánchez (2011).

3.1. Conceptos iniciales

Para el caso de extremos espaciales son fundamentales los procesos máx-estables, análogos infinito-dimensionales a los vectores aleatorios máx-estables, definidos en el capítulo anterior. La principal diferencia con los extremos multivariados es que los extremos ahora se definen continuamente en un dominio espacial $\chi \subset \mathbb{R}^d$, $d \geq 1$. Aunque es posible relajar estos supuestos, asumiendo que χ es un subconjunto compacto de \mathbb{R}^d y que todos los procesos estocásticos a considerar son continuos.

Definición 3.1. Sea $\{Z_i(x) : x \in \chi, i \geq 1\}$ una sucesión de réplicas independientes de un proceso estocástico $\{Z(x) : x \in \chi\}$. Si existen sucesiones de funciones $\{a_n(x) > 0 : x \in \chi, n \geq 1\}$ y $\{b_n(x) : x \in \chi, n \geq 1\}$ tales que para todo $n \geq 1$,

$$\left\{ \frac{\max_{i=1, \dots, n} Z_i(x) - b_n(x)}{a_n(x)} : x \in \chi \right\} \stackrel{d}{=} \{Z(x) : x \in \chi\},$$

entonces, el proceso estocástico $\{Z(x) : x \in \chi\}$ es un **proceso máx-estable**.

Debido a lo supuesto previamente, la igualdad anterior se entiende en sentido de todas las distribuciones dimensionales finitas.

Teorema 3.1. Sea $\{X_i(x) : x \in \chi, i \geq 1\}$ una sucesión de réplicas independientes de un proceso estocástico $\{X(x) : x \in \chi\}$. Si existen sucesiones de funciones continuas $\{c_n(x) > 0 : x \in \chi, n \geq 1\}$ y $\{d_n(x) : x \in \chi, n \geq 1\}$ tales que,

$$\left\{ \frac{\max_{i=1, \dots, n} X_i(x) - d_n(x)}{c_n(x)} : x \in \chi \right\} \rightarrow \{Z(x) : x \in \chi\},$$

cuando $n \rightarrow \infty$, entonces, siempre que no sea degenerado, el proceso estocástico $\{Z(x) : x \in \chi\}$ es un **proceso máx-estable**.

Tener en cuenta que la convergencia previa se refiere a la convergencia débil en el espacio de funciones continuas en χ .

Observación 3.1. Para ser coherente con la teoría de valores extremos univariados, el teorema (3.1) implica que la distribución marginal de $\{Z(x) : x \in \chi\}$ debe ser una distribución de valores extremos generalizada.

La motivación estadística para usar procesos máx-estable para modelar extremos espaciales es la siguiente.

Basado en n réplicas independientes, se asume que el proceso límite $\{Z(x) : x \in \chi\}$ es probable que sea un buen candidato para modelar el proceso máximo parcial,

$$\left\{ \max_{i=1, \dots, n} X_i(x) : x \in \chi \right\},$$

siempre que n sea lo suficientemente grande. La lógica más allá de esto es exactamente la misma que para el análisis univariado de valores extremos donde uno prefiere trabajar directamente con la distribución asintótica de los máximos de bloques, es decir, la distribución generalizada de valores extremos en lugar de estimar la distribución de $\{X(x) : x \in \chi\}$ y elevarlo a la potencia n para estimar la distribución de los máximos parciales.

Hasta ahora, ni la definición (3.1) ni el teorema (3.1) dan una descripción precisa de los procesos máx-estables, y sería bueno tener una mejor imagen de ellos en los casos univariados y multivariantes.

Esta descripción se conoce como la representación espectral de procesos máx-estables. Como es sabido que las distribuciones marginales de $\{Z(x) : x \in \chi\}$ tiene que ser una distribución de valores extremos generalizados es más conveniente establecer las marginales en una distribución dada.

Una opción ampliamente utilizada es tratar con marginales de Fréchet unitarias, es decir, $P\{Z(x) \leq z\} = \exp\{-1/z\}$, para todo $x \in \chi$ y $z > 0$, y con estas marginales específicas $\{Z(x) : x \in \chi\}$ se dice que es un **proceso máx-estable simple**. En este caso se puede observar que las constantes de normalización son $a_n(x) = n$ y $b_n(x) = 0$. Por tanto $\{Z(x) : x \in \chi\}$ es un proceso máx-estable simple si las marginales son Fréchet unitarias y $\max_{i=1, \dots, n} \{X_i(x)/n\} \xrightarrow{d} Z(x)$. Con ello se deriva el siguiente resultado.

Teorema 3.2. *Sea $X(x)$ un proceso aleatorio con marginales Fréchet unitarias en un espacio de índices χ , y sea*

$$Z_n^*(x) = \max_{i=1, \dots, n} \{X_i(x)/n\}.$$

Si,

$$Z_n^*(x) \xrightarrow{d} Z(x),$$

cuando $n \rightarrow \infty$, donde $Z(x)$ es un proceso bien definido, entonces $Z(x)$ es un proceso máx-estable.

La relevancia de los procesos máx-estables en la modelización de extremos espaciales estará basada en argumentos asintóticos.

3.2. Caracterización espectral

Los procesos máx-estables simples presentan una sencilla representación, a través de lo que se llama la representación espectral del proceso. Se exponen dos posibles enfoques.

3.2.1. Construcción de L. de Haan (1984) de procesos máx-estables

Desde un punto de vista inferencial, es probable que el proceso se observe en un número finito de ubicaciones, debido a que el ajuste de procesos máx-estables es similar al ajuste de distribuciones multivariadas de valores extremos. Sin embargo, dado que se trabaja con procesos espaciales, la mayoría de las veces se desea poder obtener predicciones en ubicaciones no observadas que difieren significativamente del análisis de extremos multivariados. Para permitir tales predicciones, es necesario poder tener una extensión del teorema 2.9 (Representación espectral via procesos puntuales Poisson) del capítulo previo, que sea válida sobre todo el conjunto χ . Esta extensión se conoce como la representación espectral de procesos máx-estables (de Haan, 1984; Penrose, 1992).

Teorema 3.3. *Cualquier proceso máx-estable simple no degenerado $\{Z(x) : x \in \chi\}$ definido en un conjunto compacto $\chi \subset \mathbb{R}^d$, $d \geq 1$, con patrones de muestreo continuos satisface,*

$$Z(x) \stackrel{d}{=} \max_{i \geq 1} \zeta_i f_i(x), \quad x \in \chi, \quad (3.1)$$

donde, $\{(\zeta_i, f_i(x)) : i \geq 1\}$ son las coordenadas de un proceso de Poisson en $(0, \infty) \times \mathcal{C}$ con medida de intensidad $\zeta^{-2} d\zeta v(df)$ para alguna medida local finita v definida en el espacio de las funciones continuas no negativas \mathcal{C} en χ , tal que

$$\int f(x) v(df) = 1, \quad x \in \chi.$$

La caracterización espectral 3.1 no es única en el sentido de que diferentes medidas ν pueden conducir al mismo proceso máx-estable $\{Z(x) : x \in \chi\}$. Además la restricción de la función no negativa sólo es necesaria por conveniencia y, dado que la caracterización espectral consiste en tomar el máximo sobre un número infinito de tales funciones, se pueden considerar funciones reales siempre que $\nu\{f(x) > 0\} > 0$ para todo $x \in \chi$.

Por último, un caso de especial atención de la caracterización espectral 3.1 es cuando ν es una medida de probabilidad, ya que en este caso la representación espectral puede escribirse como,

$$Z(x) \stackrel{d}{=} \max_{i \geq 1} \zeta_i Y_i(x), \quad x \in \chi,$$

donde $\{\zeta_i : i \geq 1\}$ son los puntos de un proceso de Poisson en $(0, \infty)$, Y_1, Y_2, \dots una sucesión de réplicas independientes de un proceso estocástico no negativo $\{Y(x) : x \in \chi\}$ con patrones de muestreo continuos y de tal manera que $\mathbb{E}(Y(x)) = 1$ para todo $x \in \chi$.

3.2.2. Construcción de Schlather (2002) de procesos máx-estables

Una segunda caracterización de los procesos máx-estables es proporcionada por Martin Schlather en 2002, donde considera funciones aleatorias en lugar de funciones determinísticas.

Teorema 3.4. *Sea $\{\zeta_i : i \geq 1\}$ un proceso puntual de Poisson en $(0, \infty)$ con medida de intensidad $\zeta^{-2}d\zeta$ y un proceso estocástico no negativo $\{Y(x) : x \in \chi\}$, independiente del anterior, con trayectorias continuas, tal que $\mathbb{E}(Y(x)) = 1$, para todo $x \in \chi$. El proceso,*

$$Z(x) \stackrel{d}{=} \max_{i \geq 1} \zeta_i Y_i(x), \quad x \in \chi, \tag{3.2}$$

donde, Y_1, Y_2, \dots son réplicas independientes de $\{Y(x) : x \in \chi\}$, es un proceso máx-estable.

3.3. Distribuciones finito dimensionales de los procesos máx-estables

Las funciones de distribución acumulada finito dimensionales de un proceso máx-estable se obtiene a través de su representación espectral. Se determina la función de distribución acumulada k -dimensional F . Más precisamente de la representación espectral de de Haan (3.1) se deriva la distribución finito dimensional de $\{Z(x) : x \in \chi\}$, para cualquier $\mathbf{x} = (x_1, \dots, x_k) \in \chi^k$, $k > 1$, y $\mathbf{z} = (z_1, \dots, z_k) \in (0, \infty)^k$, se tiene,

$$\begin{aligned}
F(z_1, z_2, \dots, z_k) &= P \left[\max_i \{\psi_i Y_i(x_j)\} \leq z_j, j = 1, 2, \dots, k \right] \\
&= P \left[\psi \leq \frac{z_j}{Y_i(x_j)}, i \geq 1, j = 1, 2, \dots, k \right] \\
&= \exp \left\{ - \int_{\mathbb{R}^d} \int_{\mathbb{R}} \mathbf{1}_{\{\psi > \min_j \frac{z_j}{y_j}\}} \psi^{-2} d\psi dP^Y(y_1, \dots, y_k) \right\} \\
&= \exp \left\{ - \int_{\mathbb{R}^d} \max_j \frac{y_j}{z_j} dP^Y(y_1, \dots, y_k) \right\} \\
&= \exp \left\{ -\mathbb{E} \left[\max_j \frac{Y(x_j)}{z_j} \right] \right\} \\
&= \exp \{-V_{\mathbf{x}}(z_1, \dots, z_k)\}.
\end{aligned} \tag{3.3}$$

Donde la función,

$$V_{\mathbf{x}}(z_1, \dots, z_k) = \mathbb{E} \left[\max_{j=1, \dots, k} \frac{Y(x_j)}{z_j} \right], \tag{3.4}$$

llamada **función exponente**, caracteriza completamente la distribución conjunta de $Z(\cdot)$. Note que de la ecuación (3.4), poniendo k -veces z se tiene que,

$$V_{\mathbf{x}}(z, \dots, z) = \frac{\theta(\mathbf{x})}{z}, \quad \theta(\mathbf{x}) = \mathbb{E} \left\{ \max_{j=1, \dots, k} Y(x_j) \right\}. \tag{3.5}$$

En la ecuación anterior $\theta(\mathbf{x})$ se llama **coeficiente extremo** (k -dimensional) y es una medida resumida de dependencia a través del elemento del vector aleatorio $Z(\mathbf{x})$. Debido a la propiedad de independencia entre los componentes radial y angular de los extremos multivariados, el coeficiente extremal es independiente del radio, es decir, el nivel z aparece en $V_{\mathbf{x}}(z, \dots, z)$ y se enfoca sólo en la dependencia.

En un contexto espacial, es más conveniente restringir nuestra atención al caso bivariado y definir lo que se conoce como **función de coeficiente extremo**, es decir,

$$\theta : h \mapsto \mathbb{E}[\max\{Y(x), Y(x+h)\}]. \quad (3.6)$$

La función de coeficiente extremo toma valores en $[1, 2]$ donde el límite inferior corresponde a la dependencia perfecta y el superior a la independencia.

De hecho, para estos dos casos límites se tiene respectivamente,

$$\begin{aligned} P\{Z(x+h) \leq z \mid Z(x) \leq z\} &= P\{Z(x+h) \leq z\}^{\theta(h)-1} & (3.7) \\ &= \begin{cases} 1, & \text{perfectamente dependiente,} \\ P\{Z(x+h) \leq z\}, & \text{independiente.} \end{cases} & (3.8) \end{aligned}$$

Debe notarse que la función exponente (3.4) puede ser difícil de evaluar explícitamente cuando $k > 2$ y esta es la razón por la cual la distribución dimensional finita de los procesos máx-estables se restringió primero al caso bivariado solamente.

En otras palabras, cuando la dimensión k es lo suficientemente grande, la densidad conjunta se torna intratable, al menos numéricamente, por la cantidad de términos que ella presenta.

Capítulo 4

Modelos máx-estables

Hasta ahora, se tienen definidos los procesos máx-estables, la forma general de las distribuciones dimensionales finitas y los resultados recientes sobre estos, así como la representación espectral simple y la posibilidad de derivar las distribuciones dimensionales finitas a partir de ella. En este capítulo se presentan los modelos máx-estables y se hará en un orden histórico: Smith, Schlather, Brown-Resnick y t -extremal propuestos respectivamente por [Smith \(1990\)](#), [Schlather \(2002\)](#), [Kablichko *et al.* \(2009\)](#) y [Opitz \(2013\)](#).

Además este capítulo está basado en “An introduction to statistical modeling of extreme values” de [Coles *et al.* \(2001\)](#), “Extreme value theory: an introduction” de [de Haan y Ferreira \(2006\)](#), “Análisis de Valores Extremos” de [Villanueva e Ibáñez \(2016\)](#), “Precipitaciones Máximas en el Estado de Guanajuato, México” de [Moreno y Ortega \(2013\)](#), “Análisis de máximos para datos espaciales de lluvias” de [González y Sánchez \(2011\)](#), “A user’s guide to the SpatialExtremes package” de [Ribatet \(2009\)](#) y “Matemáticas y sus aplicaciones 15” de [Romero y Carrasco \(2020\)](#).

4.1. Modelo de Smith

Un primer modelo fue desarrollado por [Smith \(1990\)](#) para la modelación de lluvias a través de procesos máx-estables continuos en el tiempo, este modelo se basa en procesos gaussianos y se fundamenta con la teoría de valores extremos, se deduce a partir de la caracterización de L. de Haan (1984), conocido como modelo de Smith, también recibe el nombre de proceso de valores extremos gaussianos o modelo perfil de la tormenta. La matriz de varianzas y covarianzas para el pro-

ceso gaussiano se desconoce y por tanto el objetivo es estimarla. Este modelo se desarrolla de forma bidimensional, debido a que únicamente se conoce de forma explícita la función de distribución acumulada bivariada, es aplicado en general para modelar lluvias de convección, esta se producen cuando el aire asciende por diferencias de temperatura a causa de un calentamiento local, se generan tanto por la subida de aire cálido como de aire húmedo. En este proceso, que Smith interpretó en términos de precipitaciones, depende del tamaño y el tipo de tormenta.

L. de Haan fue el primero en proponer una caracterización de procesos máx-estables, la cual se vio en la sección anterior. Luego, Smith usa esta caracterización para proporcionar un modelo paramétrico para extremos espaciales. La construcción es la siguiente.

Teorema 4.1. *Sea $\{(\xi_i, y_i), i \geq 1\}$ que denota los puntos de un proceso de Poisson en $(0, +\infty) \times \mathbb{R}^d$ con medida de intensidad $\xi^{-2}d\xi\nu(dy)$, donde $\nu(dy)$ es una medida positiva en \mathbb{R}^d . Entonces, una caracterización de un proceso máx-estable con marginales Fréchet unitarias es,*

$$Z(x) = \max_i \{\xi_i f(y_i, x)\}, \quad x \in \mathbb{R}^d, \quad (4.1)$$

donde, $\{f(y, x), x, y \in \mathbb{R}^d\}$ es una función no negativa tal que

$$\int_{\mathbb{R}^d} f(x, y)\nu(dy) = 1, \quad \forall x \in \mathbb{R}^d. \quad (4.2)$$

Demostración. Para ver que la ecuación (4.1) define un proceso estacionario máx-estable con marginales Fréchet unitarias, se tiene que verificar que las marginales son de hecho Fréchet unitarias y que $Z(x)$ satisface la propiedad de un proceso máx-estable.

Para encontrar su distribución, considerar el conjunto definido por,

$$E = \{(\xi, y) \in \mathbb{R}^+ \times \mathbb{R}^d : \xi f(y, x) > z\},$$

para una ubicación fija $x \in \mathbb{R}^d$ y $z > 0$. Entonces,

$$\begin{aligned}
P[Z(x) \leq z] &= P[\text{No existan puntos en E}] \\
&= \exp \left\{ - \int_{\mathbb{R}^d} \int_{z/f(y,x)}^{+\infty} \xi^{-2} d\xi \nu(dy) \right\} \\
&= \exp \left\{ - \int_{\mathbb{R}^d} z^{-1} f(x, y) \nu(dy) \right\} \\
&= \exp \left(-\frac{1}{z} \right).
\end{aligned}$$

y las marginales son Fréchet unitarias.

La propiedad máx-estable de $Z(\cdot)$ se debe a que la superposición de n procesos de Poisson idénticos e independientes es un proceso de Poisson con su intensidad multiplicada por n . Más precisamente,

$$\left\{ \max_{i=1}^n Z_i(x_1), \dots, \max_{i=1}^n Z_i(x_k) \right\} \sim n \{Z(x_1), \dots, Z(x_k)\}, \quad k \in \mathbb{N}.$$

□

El proceso definido por la ecuación (4.1) a menudo se conoce como el proceso de lluvia y tormenta. Pensar en y_i como realizaciones de los centros de tormentas de lluvias en \mathbb{R}^d y $\nu(dy)$ como la distribución espacial de estos centros de tormenta sobre \mathbb{R}^d , generalmente $d = 2$. Cada ξ_i representa la intensidad de la i -ésima tormenta y por lo tanto $\xi_i f(y_i, x)$ representa la cantidad de lluvia para este evento específico en la ubicación x . En otras palabras, $f(y_i, \cdot)$ determina como se concentró la i -ésima tormenta en y_i en el espacio.

El teorema (4.1) es básicamente general y Smith consideró un ajuste particular que se deduce a partir de la caracterización de L. de Haan.

Teorema 4.2. Sean $\nu(dy)$ la medida de Lebesgue y

$$f(s, t) = f_0(t - s),$$

donde $f_0(t - s)$ es la función de densidad de una distribución normal multivariada con media cero y matriz de varianzas y covarianzas Σ , es decir,

$$f(s, t) = (2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp \left\{ -\frac{1}{2} (s - t)^t \Sigma^{-1} (s - t) \right\}.$$

De esta manera, para dos elementos del proceso máx-estable en cuestión la

distribución conjunta vía el modelo de Smith queda determinada en forma bidimensional por la ecuación,

$$P[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = \exp \left\{ -\frac{1}{z_1} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1} \right) - \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2} \right) \right\}, \quad (4.3)$$

donde, $\Phi(\cdot)$ denota la función de distribución acumulada de una normal estándar en \mathbb{R} y para las ubicaciones x_1 y x_2 , $a^2 = (x_1 - x_2)^t \Sigma^{-1} (x_1 - x_2)$.

Demostración. (de ecuación 4.3).

Primeramente note que:

$$\begin{aligned} -\log P[Z(x_1) \leq z_1, Z(x_2) \leq z_2] &= \int_{\mathbb{R}^d} \int_{\min\left\{\frac{z_1}{f_0(s-x_1)}, \frac{z_2}{f_0(s-x_2)}\right\}}^{+\infty} \xi^{-2} d\xi ds \\ &= \int_{\mathbb{R}^d} -\xi^{-1} \Big|_{\min\left\{\frac{z_1}{f_0(s-x_1)}, \frac{z_2}{f_0(s-x_2)}\right\}}^{+\infty} ds \\ &= \int_{\mathbb{R}^d} \max \left\{ \frac{f_0(s-x_1)}{z_1}, \frac{f_0(s-x_2)}{z_2} \right\} ds \\ &= \int_{\mathbb{R}^d} \frac{f_0(s-x_1)}{z_1} \mathbf{1}_{\left(\frac{f_0(s-x_1)}{z_1} > \frac{f_0(s-x_2)}{z_2}\right)} \\ &\quad + \int_{\mathbb{R}^d} \frac{f_0(s-x_2)}{z_2} \mathbf{1}_{\left(\frac{f_0(s-x_1)}{z_1} \leq \frac{f_0(s-x_2)}{z_2}\right)} \\ &= \int_{\mathbb{R}^d} \frac{f_0(s)}{z_1} \mathbf{1}_{\left(\frac{f_0(s)}{z_1} > \frac{f_0(s-x_2+x_1)}{z_2}\right)} \\ &\quad + \int_{\mathbb{R}^d} \frac{f_0(s)}{z_2} \mathbf{1}_{\left(\frac{f_0(s-x_1+x_2)}{z_1} \leq \frac{f_0(s)}{z_2}\right)} \\ &= \frac{1}{z_1} \mathbb{E} \left[\mathbf{1}_{\left(\frac{f_0(X)}{z_1} > \frac{f_0(X-x_2+x_1)}{z_2}\right)} \right] \\ &\quad + \frac{1}{z_2} \mathbb{E} \left[\mathbf{1}_{\left(\frac{f_0(X-x_1+x_2)}{z_1} \leq \frac{f_0(X)}{z_2}\right)} \right] \end{aligned}$$

donde $\mathbf{1}_A$ es la función indicadora del conjunto A .

Sea X una variable aleatoria, con densidad normal multivariada f_0 , media 0 y matriz de covarianza Σ . Para obtener la forma cerrada de la distribución bivariada, queda calcular las probabilidades del evento $\{f_0(X)/z_1 > f_0(X-x_2+x_1)/z_2\}$.

Note que:

$$\begin{aligned}
\frac{f_0(X)}{z_1} &> \frac{f_0(X - x_2 + x_1)}{z_2} \\
\iff &\frac{(2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp \left\{ -\frac{1}{2}(X - 0)^t \Sigma^{-1}(X - 0) \right\}}{z_1} \\
&> \frac{(2\pi)^{-d/2} \det(\Sigma)^{-1/2} \exp \left\{ -\frac{1}{2}(X - x_2 + x_1 - 0)^t \Sigma^{-1}(X - x_2 + x_1 - 0) \right\}}{z_2} \\
\iff &\exp \left\{ -\frac{1}{2} X^t \Sigma^{-1} X \right\} > \frac{z_1}{z_2} \exp \left\{ -\frac{1}{2} (X - x_2 + x_1)^t \Sigma^{-1} (X - x_2 + x_1) \right\} \\
\iff &X^t \Sigma^{-1} \left(-\frac{1}{2} X \right) > \log \left(\frac{z_1}{z_2} \right) - \frac{1}{2} (X - x_2 + x_1)^t \Sigma^{-1} (X - x_2 + x_1)
\end{aligned}$$

y esto ocurre si y sólo si,

$$X^t \Sigma^{-1} (x_1 - x_2) > \log \left(\frac{z_1}{z_2} \right) - \frac{1}{2} (x_1 - x_2)^t \Sigma^{-1} (x_1 - x_2).$$

Luego, como X tiene densidad f_0 , $X^t \Sigma^{-1} (x_1 - x_2)$ es normal con media 0 y varianza $a^2 = (x_1 - x_2)^t \Sigma^{-1} (x_1 - x_2)$. Se obtiene

$$\begin{aligned}
\frac{1}{z_1} \mathbb{E} \left[\mathbf{I} \left(\frac{f_0(X)}{z_1} > \frac{f_0(X - x_2 + x_1)}{z_2} \right) \right] &= \frac{1}{z_1} \left\{ 1 - \Phi \left(\frac{\log \left(\frac{z_1}{z_2} \right)}{a} - \frac{a}{2} \right) \right\} \\
&= \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{\log \left(\frac{z_1}{z_2} \right)}{a} \right).
\end{aligned}$$

Finalmente, aplicando lo anterior para los índices 1 y 2 intercambiados.

$$\frac{1}{z_1} \mathbb{E} \left[\mathbf{I} \left(\frac{f_0(X - x_1 + x_2)}{z_1} \leq \frac{f_0(X)}{z_2} \right) \right] = \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{\log \left(\frac{z_1}{z_2} \right)}{a} \right).$$

Con todo esto se tiene el resultado. □

Siguiendo los conceptos vertidos por Smith en 1990, este es un proceso homogéneo en el espacio, es decir, no importa en el modelo la ubicación de los puntos, en este caso las estaciones, sino la distancia entre ellos. Se trabaja en un espacio de dimensión $d = 2$. El proceso $Z(t)$ viene a representar en este caso un máxi-

mo sobre un conjunto de tormentas independientes con la misma forma pero de intensidad variable.

4.2. Modelo de Schlather

La propuesta de [Schlather \(2002\)](#) es una extensión del modelo de [Smith \(1990\)](#), sustituyendo la función determinística por una forma aleatoria, dicho modelo se desarrolla utilizando réplicas independientes de un proceso estándar gaussiano con correlación dada, partiendo de distribuciones acumuladas finito dimensionales de un proceso máx-estable, que son expresadas de forma bidimensional. Algunas de las posibles funciones de correlación utilizadas, son: exponencial, exponencial generalizada, exponencial potencia, esférica, gaussiana, Matérn o bien función de Bessel modificada, Whittle-Matérn y Cauchy. La elección de la función de correlación depende del fenómeno de estudio, el objetivo de este modelo es estimar los parámetros que intervienen en la función de covarianza.

En general este modelo es aplicado para lluvias ciclónicas, se producen precisamente, cuando se crea una especie de remolino debido a diferencias de presiones y no debido a diferencias de temperaturas, cuya característica principal es que ocurren en puntos aislados dispersos en la región de estudio (tal dispersión puede variar considerablemente), dado que la forma de la tormenta es aleatoria y su estructura está dada por la función de correlación. En el caso de una tormenta fija, esta mantiene la misma estructura de dependencia y sólo varía la magnitud.

La propuesta de Schlather es a través de su construcción de procesos máx-estables, esto es considerar la caracterización de L. de Haan sustituyendo la función determinística $f(\cdot)$ por una forma aleatoria $Y(\cdot)$.

El siguiente teorema trata funciones aleatorias arbitrarias $Y(\cdot)$ cuya parte positiva $\max\{0, Y(\cdot)\}$, es integrable.

Teorema 4.3. *Sea Y una función aleatoria medible definida en \mathbb{R}^d y*

$$\mu = \mathbb{E} \left(\int_{\mathbb{R}^d} \max\{0, Y(x)\} dx \right) \in (0, \infty).$$

Sean $\{\xi_i, U_i\}$ puntos de un proceso de Poisson en $(0, \infty) \times \mathbb{R}^d$ con medida de intensidad $d\Lambda(y, \xi) = (\mu^{-1} dy) \times (\xi^{-2} d\xi)$ y $Y_i(\cdot)$ son copias independientes e idénticas

ticamente distribuidas de $Y(\cdot)$; luego

$$Z(x) = \sup_{i \geq 1} \{\xi_i Y_i(x - U_i)\} = \sup_{i \geq 1} \xi_i \max\{0, Y_i(x - U_i)\},$$

es un proceso aleatorio máx-estable estacionario con distribuciones marginales Fréchet unitarias.

Se introduce una clase de procesos máx-estables basada en procesos estocásticos estacionarios con medida finita.

Denote por o el origen del sistema de coordenadas. El siguiente teorema da lugar al modelo denominado modelo de Schlather.

Teorema 4.4. *Sea $Y(\cdot)$ un proceso estacionario en \mathbb{R}^d con $\mu = \mathbb{E}[\max\{0, Y(o)\}] \in (0, \infty)$ y sean $\{\xi_i, i \geq 1\}$ puntos de un proceso de Poisson definido en $(0, \infty)$ con medida de intensidad $d\Lambda(\xi) = \mu^{-1} \xi^{-2} d\xi$. Luego,*

$$Z(\mathbf{x}) = \max_{i \geq 1} \{\xi_i Y_i(\mathbf{x})\} = \max_{i \geq 1} \xi_i \max\{0, Y_i(\mathbf{x})\}, \quad (4.4)$$

es un proceso máx-estable estacionario con distribuciones marginales de Fréchet unitarias, donde las $\{Y_i(\cdot)\}_{i \geq 1}$ son copias independientes e idénticamente distribuidas de $Y(\cdot)$.

Demostración. Como antes, la propiedad máx-estable de $Z(\cdot)$ proviene de la superposición de n procesos de Poisson idénticos e independientes. Más precisamente,

$$\left\{ \max_{i=1}^n Z_i(x_1), \dots, \max_{i=1}^n Z_i(x_k) \right\} \sim n \{Z(x_1), \dots, Z(x_k)\}, \quad k \in \mathbb{N}.$$

Las marginales unitarias de Fréchet se mantienen con el mismo argumento que para el modelo de Smith. De hecho, se considera el siguiente conjunto.

$$E = \{(\xi, y(\mathbf{x})) \in \mathbb{R}^+ \times \mathbb{R}^d : \xi \max(0, y(\mathbf{x})) > z\}.$$

para una ubicación fija $\mathbf{x} \in \mathbb{R}^d$ y $z > 0$. Entonces,

$$\begin{aligned}
P[Z(\mathbf{x}) \leq z] &= P(\text{No existan puntos en } E) \\
&= \exp \left\{ - \int_{\mathbb{R}^d} \int_{z/\text{máx}(0, y(\mathbf{x}))}^{+\infty} \mu^{-1} \xi^{-2} d\nu(dy(\mathbf{x})) \right\} \\
&= \exp \left\{ -\mu^{-1} \int_{\mathbb{R}^d} z^{-1} \text{máx}\{0, y(\mathbf{x})\} \nu(dy(\mathbf{x})) \right\},
\end{aligned}$$

pero como $\mathbb{E}\{Y(\mathbf{x})\} = \mathbb{E}\{0, Y(\mathbf{x})\} = \mu$,

$$P[Z(\mathbf{x}) \leq z] = \exp(-z^{-1}) = \exp\left(-\frac{1}{z}\right),$$

de modo que las marginales son Fréchet unitarias. \square

Al igual que con el modelo de Smith, el proceso definido en la ecuación (4.4) tiene una interpretación práctica. Pensar en $\xi_i Y_i(\cdot)$ como los eventos individuales en el espacio de lluvia diaria, estos tienen todos la misma estructura de dependencia y difieren sólo en la magnitud ξ_i . Este modelo difiere ligeramente del de Smith ya que ahora no se tiene una forma determinista, como una densidad normal multivariada para las tormentas, sino una forma aleatoria impulsada por el proceso $Y(\cdot)$.

Las caracterizaciones de Schlather y Smith tienen fuertes conexiones. Para ver esto, se considera el caso en el que $Y_i(x) = f_0(x - X_i)$ donde f_0 es una función de densidad de probabilidad y $\{X_i\}$ es un proceso de Poisson homogéneo en \mathbb{R}^d . Con esta configuración particular, el modelo (4.4) es idéntico para modelar (4.1).

La ecuación (4.4) es muy general y se necesitan suposiciones adicionales para obtener modelos prácticos.

Teorema 4.5. *Sea $\{Y_i(\cdot)\}_{i \geq 1}$ copias de un proceso gaussiano estándar estacionario $Y(\cdot)$ con función de correlación $\rho(h)$, escalado de modo que $\mathbb{E}[\text{máx}\{0, Y_i(x)\}] = 1$ y Π un proceso de Poisson en $(0, \infty)$, con medida de intensidad $d\Lambda(\xi) = \sqrt{2\pi} \xi^{-2} d\xi$. Con estos nuevos supuestos aplicados al teorema (4.4), se tiene la siguiente expresión para la **distribución conjunta bivariada** del proceso $Z(x) = \text{máx}_{i \geq 1} \xi_i \text{máx}\{0, Y_i(x)\}$.*

$$P[Z(x_1) \leq z_1, Z(x_2) \leq z_2] = \exp \left[-\frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left(1 + \sqrt{1 - 2(\rho(h) + 1) \frac{z_1 z_2}{(z_1 + z_2)^2}} \right) \right], \quad (4.5)$$

donde $h = \|x_1 - x_2\|$, es decir, $h \in \mathbb{R}^+$ es la distancia euclidiana entre la ubicación 1 y la ubicación 2.

Demostración. (de la ecuación 4.5).

Del teorema (4.4) se tiene que el proceso gaussiano extremo se define como

$$Z(\mathbf{x}) = \max_{i \geq 1} \{\xi_i Y_i(\mathbf{x})\},$$

donde los Y_i son copias independientes idénticamente distribuidas del proceso gaussiano estacionario Y , con función de correlación ρ y Π un proceso de Poisson en $(0, \infty)$, con medida de intensidad $d\Lambda(\xi) = \sqrt{2\pi}\xi^{-2}d\xi$. Así, escribiendo ρ_h en lugar de $\rho(h)$ para abreviar, se obtiene

$$\begin{aligned} -\log P[Z(x_1) \leq z_1, Z(x_2) \leq z_2] &= \sqrt{2\pi} \int_{\mathbb{R}^d} \int_{\min\left\{\frac{z_1}{\max\{0, \xi_1(x_1)\}}, \frac{z_2}{\max\{0, \xi_2(x_2)\}}\right\}}^{+\infty} \xi^{-2} d\xi dP \\ &= \sqrt{2\pi} \int_{\mathbb{R}^d} -\xi^{-1} \Big|_{\min\left\{\frac{z_1}{\max\{0, \xi_1(x_1)\}}, \frac{z_2}{\max\{0, \xi_2(x_2)\}}\right\}}^{+\infty} dP \\ &= \sqrt{2\pi} \int_{\mathbb{R}^d} \max\left\{0, \frac{\xi_1(x_1)}{z_1}, \frac{\xi_2(x_2)}{z_2}\right\} dP \\ &= \sqrt{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \max\left\{0, \frac{v}{z_1}, \frac{u}{z_2}\right\} \frac{\exp\left\{-\frac{v^2 - 2uv\rho_h + u^2}{2(1-\rho_h^2)}\right\}}{2\pi(1-\rho_h^2)^{1/2}} dv du. \end{aligned}$$

Para simplificar, defina

$$f(u, v) = \frac{v^2 - 2uv\rho_h + u^2}{2(1-\rho_h^2)} \quad \text{y} \quad V(z_1, z_2) = -\log P[Z(x_1) \leq z_1, Z(x_2) \leq z_2]$$

además, note que

$$\frac{\sqrt{2\pi}}{2\pi(1-\rho_h^2)^{1/2}} = \frac{1}{\sqrt{2\pi(1-\rho_h^2)}} = \frac{1}{k},$$

donde

$$k = \sqrt{2\pi(1-\rho_h^2)}, \quad (4.6)$$

resultando

$$V(z_1, z_2) = \frac{1}{k} \int_0^{\infty} \int_0^{\infty} \max\left\{0, \frac{v}{z_1}, \frac{u}{z_2}\right\} \exp\{-f(u, v)\} dv du.$$

Aquí se tienen tres casos al tomar el máximo, cuando es i) 0, ii) $\frac{v}{z_1}$ y iii) $\frac{u}{z_2}$.

i) De esta forma, cuando el máximo es 0, $V(z_1, z_2)$ resulta,

$$\begin{aligned} A &= \frac{1}{k} \int_0^\infty \int_0^\infty \frac{v}{z_1} \exp\{-f(u, v)\} dv du + \frac{1}{k} \int_0^\infty \int_0^\infty \frac{u}{z_2} \exp\{-f(u, v)\} dv du \\ &= \frac{1}{kz_1} \int_0^\infty \int_0^\infty v \exp\{-f(u, v)\} dv du + \frac{1}{kz_2} \int_0^\infty \int_0^\infty u \exp\{-f(u, v)\} dv du. \end{aligned}$$

defina

$$I_1 = \int_0^\infty \int_0^\infty v \exp\{-f(u, v)\} dv du$$

e

$$I_{11} = \int_0^\infty \int_0^\infty u \exp\{-f(u, v)\} dv du,$$

Observe que el resultado de I_1 debe ser el mismo que el de I_{11} , por lo que se puede escribir lo siguiente,

$$\begin{aligned} I_1 + I_{11} &= \int_0^\infty \int_0^\infty (v + u) \exp\{-f(u, v)\} dv du \\ \iff 2I_1 &= \int_0^\infty \int_0^\infty (v + u) \exp\{-f(u, v)\} dv du \\ \text{Por lo tanto, } I_1 &= \frac{1}{2} \int_0^\infty \int_0^\infty (v + u) \exp\{-f(u, v)\} dv du \end{aligned}$$

Así,

$$\begin{aligned} A &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \frac{1}{2} \int_0^\infty \int_0^\infty (v + u) \exp\{-f(u, v)\} dv du \\ &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) I_1 \end{aligned} \tag{4.7}$$

ii) Si el máximo es $\frac{v}{z_1}$, se debe cumplir que $\frac{v}{z_1} > \frac{u}{z_2}$, alternativamente $v > u \frac{z_1}{z_2}$,

por lo que $V(z_1, z_2)$ es,

$$\begin{aligned} B &= \frac{1}{k} \int_0^\infty \int_0^\infty \frac{v}{z_1} \exp\{-f(u, v)\} dv du \\ &= \frac{1}{kz_1} \int_0^\infty \int_0^\infty v \exp\{-f(u, v)\} dv du \\ &= \frac{1}{kz_1} \int_0^\infty \int_0^\infty \mathbf{1}_{\{v > u \frac{z_1}{z_2}\}} \exp\{-f(u, v)\} dv du, \end{aligned}$$

donde $\mathbf{1}_A$ es la función indicadora del conjunto A . Si se define

$$I_2 = \int_0^\infty \int_0^\infty \mathbf{1}_{\{v > u \frac{z_1}{z_2}\}} \exp\{-f(u, v)\} dv du$$

entonces,

$$B = \frac{1}{kz_1} I_2 \quad (4.8)$$

iii) Análogamente a ii), si el máximo es $\frac{u}{z_2}$, entonces $u > v \frac{z_2}{z_1}$, y $V(z_1, z_2)$ resulta,

$$\begin{aligned} C &= \frac{1}{kz_2} \int_0^\infty \int_0^\infty \mathbf{1}_{\{u > v \frac{z_2}{z_1}\}} \exp\{-f(u, v)\} dv du \\ &= \frac{1}{kz_2} I_3. \end{aligned} \quad (4.9)$$

Una vez analizados los caso previos obteniendo las ecuaciones (4.7), (4.8) y (4.9), $V(z_1, z_2)$ queda determinado como sigue.

$$V(z_1, z_2) = \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) I_1 + \frac{1}{kz_1} I_2 + \frac{1}{kz_2} I_3. \quad (4.10)$$

Se procede a resolver cada una de las integrales anteriores. Para I_1 se usa la transformación

$$\begin{aligned} u &= x + y \\ v &= x - y \end{aligned}$$

El jacobiano es 2, $u + v = 2x$ y sustituyendo en $f(u, v)$,

$$f(x, y) = \frac{2[x^2(1 - \rho_h) + y^2(1 + \rho_h)]}{2(1 - \rho_h^2)} = k_1 x^2 + k_2 y^2.$$

Siendo

$$k_1 = \frac{1 - \rho_h}{1 - \rho_h^2} \quad y \quad k_2 = \frac{1 + \rho_h}{1 - \rho_h^2},$$

De esta forma

$$\begin{aligned} I_1 &= \frac{1}{2} \int_0^\infty \int_{|y|}^\infty 2x \exp(-k_1 x^2 - k_2 y^2) 2dx dy \\ &= \int_0^\infty \left(\int_{|y|}^\infty 2x \exp(-k_1 x^2) dx \right) \exp(-k_2 y^2) dy, \end{aligned}$$

pero

$$\int_{|y|}^\infty 2x \exp(-k_1 x^2) dx = \frac{1}{k_1} \exp(-k_1 y^2).$$

Así

$$I_1 = \frac{1}{k_1} \int_0^\infty \exp(-k_1 y^2) \exp(-k_2 y^2) dy = \frac{1}{k_1} \int_0^\infty \exp\{-(k_1 + k_2)y^2\} dy.$$

Recuerde que $\int_{-\infty}^\infty \exp(-r^2) dr = \sqrt{\pi}$; con esto en mente,

$$\begin{aligned} I_1 &= (1 + \rho_h) \frac{\sqrt{1 - \rho_h^2}}{\sqrt{2}} \sqrt{\pi} \\ &= \frac{(1 + \rho_h) \sqrt{2\pi(1 - \rho_h^2)}}{2} \\ &= \frac{(1 + \rho_h)k}{2}, \end{aligned} \tag{4.11}$$

donde k está dado en (4.6). Para resolver I_2 e I_3 se define,

$$g(z) = \int_0^\infty \int_0^\infty \mathbf{1}_{\{u > vz\}} \exp\{-f(u, v)\} dv du,$$

con esto se observa que

$$I_2 = g(z_1 z_2^{-1}) \quad e \quad I_3 = g(z_2 z_1^{-1}),$$

esto sugiere que sólo basta calcular $g(z)$.

Observe que

$$\begin{aligned}
 g(z) &= \int_0^\infty \int_{uz}^\infty v \exp\{-f(u, v)\} dv du \\
 &= \int_0^\infty \int_{uz}^\infty (v - u\rho_h) \exp\{-f(u, v)\} dv du \\
 &\quad - \rho_h \int_0^\infty \int_0^{uz} u \exp\{-f(u, v)\} dv du \\
 &\quad + \rho_h \int_0^\infty \int_0^\infty u \exp\{-f(u, v)\} dv du.
 \end{aligned}$$

También defina,

$$h(z) = \int_0^\infty \int_{uz}^\infty (v - u\rho_h) \exp\{-f(u, v)\} dv du,$$

y observe que

$$\int_0^\infty \int_0^{uz} u \exp\{-f(u, v)\} dv du = g(z^{-1})$$

y

$$\int_0^\infty \int_0^\infty u \exp\{-f(u, v)\} dv du = I_1,$$

así,

$$g(z) = h(z) - \rho_h g(z^{-1}) + \rho_h I_1.$$

Sustituyendo $g(z^{-1})$ en $g(z)$ y simplificando se obtiene,

$$g(z) = \frac{h(z) - \rho_h h(z^{-1}) + \rho_h(1 - \rho_h)I_1}{1 - \rho_h^2}. \quad (4.12)$$

Ahora se resuelve $h(z)$ por sustitución simple, para ello sea $t = f(u, v)$, entonces

$$\frac{dt}{dv} = \frac{2v - 2u\rho_h}{2(1 - \rho_h^2)} = \frac{v - u\rho_h}{1 - \rho_h^2},$$

de aquí se obtiene que,

$$(1 - \rho_h^2)dt = (v - u\rho_h)dv.$$

Note que si $v = uz$ entonces $t = au^2$, con

$$a = \frac{z^2 - 2Z\rho_h + 1}{2(1 - \rho_h^2)}.$$

Con todo esto en mente procedemos a resolver,

$$\begin{aligned} h(z) &= \int_0^\infty \int_{uz}^\infty (v - u\rho_h) \exp\{-f(u, v)\} dv du \\ &= \int_0^\infty \int_{uz^2}^\infty (1 - \rho_h^2) \exp\{-t\} dv du \\ &= (1 - \rho_h^2) \int_0^\infty \exp\{-au^2\} du \\ &= (1 - \rho_h^2) \frac{\sqrt{\pi}}{2\sqrt{a}}, \text{ sustituyendo } a \\ &= \frac{(1 - \rho_h^2) \sqrt{2\pi(1 - \rho_h^2)}}{2\sqrt{z^2 - 2z\rho_h + 1}} \\ &= \frac{(1 - \rho_h^2)k}{2\sqrt{z^2 - 2z\rho_h + 1}}. \end{aligned} \tag{4.13}$$

Si se sustituye z^{-1} en (4.13) el radical del denominador se puede escribir como

$$\sqrt{z^2 - 2z\rho_h + 1} = \frac{1}{z} \sqrt{z^2 - 2z\rho_h + 1}.$$

Sustituyendo (4.13) y (4.11) en (4.12), se obtiene

$$\begin{aligned} g(z) &= \frac{k}{2\sqrt{z^2 - 2z\rho_h + 1}} - \frac{z\rho_h k}{2\sqrt{z^2 - 2z\rho_h + 1}} + \frac{\rho_h k}{2} \\ &= \frac{k}{2} \left(\frac{1}{\sqrt{z^2 - 2z\rho_h + 1}} - \frac{z\rho_h}{\sqrt{z^2 - 2z\rho_h + 1}} + \rho_h \right) \end{aligned}$$

Si $z = z_1/z_2$,

$$\sqrt{z^2 - 2z\rho_h + 1} = \frac{1}{z_2} \sqrt{(z_1 + z_2)^2 - 2z_1 z_2 (1 + \rho_h)}.$$

Así,

$$\begin{aligned} \frac{2}{k} g\left(\frac{z_1}{z_2}\right) &= \frac{z_2}{\sqrt{(z_1 + z_2)^2 - 2z_1z_2(1 + \rho_h)}} - \frac{z_1z_2\rho_h}{z_2\sqrt{(z_1 + z_2)^2 - 2z_1z_2(1 + \rho_h)}} + \rho_h \\ &= \frac{(z_2 - z_1\rho_h)}{\sqrt{(z_1 + z_2)^2 - 2z_1z_2(1 + \rho_h)}} + \rho_h. \end{aligned}$$

Similarmente, si $z = z_2/z_1$ se tiene,

$$g\left(\frac{z_2}{z_1}\right) = \frac{k}{2} \left[\frac{(z_1 - z_2\rho_h)}{\sqrt{(z_1 + z_2)^2 - 2z_1z_2(1 + \rho_h)}} + \rho_h \right].$$

Simplificando,

$$\begin{aligned} \frac{1}{kz_1} I_2 + \frac{1}{kz_2} I_3 &= \frac{1}{kz_1} g\left(\frac{z_1}{z_2}\right) + \frac{1}{kz_2} g\left(\frac{z_2}{z_1}\right) \\ &= \frac{1}{kz_1} \frac{k}{2} \left[\frac{(z_2 - z_1\rho_h)}{\sqrt{(z_2 + z_1)^2 - 2z_1z_2(1 + \rho_h)}} + \rho_h \right] + \\ &\quad \frac{1}{kz_2} \frac{k}{2} \left[\frac{(z_1 - z_2\rho_h)}{\sqrt{(z_2 + z_1)^2 - 2z_1z_2(1 + \rho_h)}} + \rho_h \right] \\ &= \frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[\sqrt{1 - \frac{2z_1z_2(1 + \rho_h)}{(z_1 + z_2)^2}} + \rho_h \right]. \end{aligned}$$

Finalmente sustituyendo en (4.10),

$$\begin{aligned} V(z_1, z_2) &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) I_1 + \frac{1}{kz_1} I_2 + \frac{1}{kz_2} I_3 \\ &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \frac{(1 - \rho_h)k}{2} + \frac{1}{2} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[\sqrt{1 - \frac{2z_1z_2(1 + \rho_h)}{(z_1 + z_2)^2}} + \rho_h \right] \\ &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[1 - \rho_h + \sqrt{1 - \frac{2z_1z_2(1 + \rho_h)}{(z_1 + z_2)^2}} + \rho_h \right] \\ &= \frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[1 + \sqrt{1 - \frac{2z_1z_2(1 + \rho_h)}{(z_1 + z_2)^2}} \right] \end{aligned}$$

por lo tanto,

$$\begin{aligned}
F(z_1, z_2) &= P \{Z(x_1) \leq z_1, Z(x_2) \leq z_2\} \\
&= \exp \left\{ -\frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[1 + \sqrt{1 - \frac{2z_1z_2(1 + \rho_h)}{(z_1 + z_2)^2}} \right] \right\} \\
&= \exp \left\{ -\frac{1}{k} \left(\frac{1}{z_1} + \frac{1}{z_2} \right) \left[1 + \sqrt{1 - \frac{2z_1z_2(1 + \rho(h))}{(z_1 + z_2)^2}} \right] \right\}, \quad z_1, z_2 > 0.
\end{aligned}$$

□

El proceso máx-estable $Z(\cdot)$ dado en (4.4) se conoce como **proceso de extremos gaussiano**. Por lo general, $\rho(h)$, se elige entre una de las familias paramétricas válidas, como

$$\text{Whittle-Matérn } \rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2} \right)^\nu K_\nu \left(\frac{h}{c_2} \right), \quad c_2 > 0, \quad \nu > 0.$$

$$\text{Cauchy } \rho(h) = \left[1 + \left(\frac{h}{c_2} \right)^2 \right]^{-\nu}, \quad c_2 > 0, \quad \nu > 0.$$

$$\text{Cauchy generalizada } \rho(h) = \left[1 + \left(\frac{h}{c_2} \right)^{\nu_2} \right]^{-\nu_1/\nu_2}, \quad c_2 > 0, \quad \nu_1, \nu_2 > 0.$$

$$\text{Exponencial potencia } \rho(h) = \exp \left[- \left(\frac{h}{c_2} \right)^\nu \right], \quad c_2 > 0, \quad 0 < \nu \leq 2.$$

$$\text{Bessel } \rho(h) = \left(\frac{2c_2}{h} \right)^\nu \Gamma(\nu + 1) J_\nu \left(\frac{h}{c_2} \right), \quad c_2 > 0, \quad \nu \geq \frac{d-2}{2}.$$

Donde c_2 es un parámetro de rango y ν , ν_1 y ν_2 son parámetros de suavidad de la función de correlación. Γ es la función gamma, J_ν y K_ν son las funciones Bessel y Bessel modificada del tercer tipo con orden ν y dimensión del espacio d .

De acuerdo con los procesos gaussianos, a estas funciones de correlación se les puede agregar un parámetro de umbral σ y otro llamado *efecto nugget* denotado por η , el cual hace referencia a una pequeña pero importante desviación del parámetro de umbral con respecto a la unidad, en el caso en que la cantidad h de las funciones de covarianza anteriores es cero. Por tal razón se tiene la relación $\sigma = 1 - \eta$. Las funciones de covarianza agregando los parámetros de umbral y *efecto nugget* están dadas por la ecuación,

$$\rho(h) = \begin{cases} \sigma + \eta & \text{para } h = 0, \\ \sigma \rho(h) & \text{para } h > 0. \end{cases}$$

De este modo, el ajuste del modelo de Schlather en la práctica se reduce a la estimación de los parámetros de rango, suavidad y umbral correspondientes a la función de covarianza utilizada (equivalentemente se pueden estimar los paráme-

tros de rango, suavidad y *efecto nugget*).

Comúnmente, el modelo de Schlather se utiliza para el estudio de lluvias llamadas **lluvias ciclónicas**, cuya característica principal es que ocurren en puntos aislados dispersos en la región de estudio (tal dispersión puede variar considerablemente).

Las funciones de correlación introducidas anteriormente son todas isotrópicas, pero el modelo (4.4) no requiere esta suposición. A partir de una función de correlación válida ρ siempre es posible obtener una función de correlación elíptica ρ_ϵ utilizando la siguiente transformación:

$$\rho_\epsilon(\Delta x) = \rho\left(\sqrt{\Delta x^T A \Delta x}\right), \quad (4.14)$$

donde Δx es el vector de distancia entre dos estaciones, A es cualquier matriz semidefinida positiva que puede involucrar parámetros adicionales. El coeficiente de correlación $\rho(h)$ puede tomar cualquier valor en $[-1, 1]$. La dependencia completa se alcanza cuando $\rho(h) = 1$, mientras que la independencia ocurre cuando $\rho(h) = -1$. Sin embargo la mayoría de las familias de correlación paramétrica no permite valores negativos, por lo que nunca se alcanza la independencia.

4.3. Modelo de Brown-Resnick

El modelo de Brown-Resnick, proceso introducido por [Brown y Resnick \(1977\)](#), pero desarrollado por [Kablichko et al. \(2009\)](#), analizado y generalizado desde el punto de vista de los procesos máx-estables, surge de elegir la caracterización de [Schlather \(2002\)](#). Nuevamente para este modelo se encuentra presente el proceso de tipo gaussiano con incrementos estacionarios, distribuciones finito dimensionales, que son expresadas en dimensión dos debido a que sólo se conoce de forma explícita la función de distribución acumulada bivariada. Además dado que es una generalización del modelo de [Schlather \(2002\)](#) con la elección correcta de la función es posible derivar de este, el modelo de [Smith \(1990\)](#).

Surge de elegir la ecuación (3.2) denominada caracterización de [Schlather](#)

(2002) considerando el proceso introducido por [Brown y Resnick \(1977\)](#),

$$Y_i(x) = \exp \left\{ \epsilon_i(x) - \frac{\sigma^2(x)}{2} \right\}, \quad (4.15)$$

donde ϵ es un proceso gaussiano continuo con incrementos estacionarios y varianza $\sigma^2(x) = \text{Var}(\epsilon(x))$. La **función de distribución conjunta bivariada** está dada por:

$$P(Z(x_1) \leq z_1, Z(x_2) \leq z_2) = \exp \left\{ -\frac{1}{z_1} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1} \right) - \frac{1}{z_2} \Phi \left(\frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2} \right) \right\},$$

donde, Φ denota la función de distribución acumulada de una normal estándar en \mathbb{R} y siendo ahora $a^2 = \text{Var}(Y(x_1 - x_2)) = \sigma^2(x_1 - x_2)$. En este caso si bien Y no es un proceso estacionario el proceso máx-estable asociado si lo es. Si se cumple que $a^2 = \sigma^2(h) \propto h^2$, $h \geq 0$, el proceso coincide con el proceso de Smith.

4.4. Modelo t-extremal

El modelo t -extremal fue desarrollado por [Opitz \(2013\)](#) y también analizado por [Ribatet y Sedki \(2013\)](#), es una generalización del proceso Brown-Resnick desarrollado por [Kabluchko et al. \(2009\)](#), al igual que en los modelos anteriores, este depende del proceso gaussiano propuesto por [Schlather \(2002\)](#), este modelo concibe una construcción espectral que generaliza la del proceso gaussiano, además de ser analizado de forma bivariada dado que sólo se conoce de forma explícita la función de distribución acumulada bivariada.

Surge de elegir,

$$Y_i(x) = \sqrt{\pi} 2^{-(v-2)/2} \Gamma \left(\frac{v+1}{2} \right)^{-1} \max\{0, \epsilon_i(x)\}^v, \quad v \geq 1, \quad (4.16)$$

donde ϵ es un proceso estándar gaussiano con función de correlación $\rho(\cdot)$, Γ es la función gamma y $d\Lambda(t) = vt^{-(v+1)}dt$ medida de intensidad para el proceso de Poisson en $(0, \infty)$. En este caso la función de distribución bivalente obtenida a partir de (3.3)

$$F(z_1, z_2, \dots, z_k) = \exp \{-V(z_1, \dots, z_k)\},$$

siendo $V(z_1, \dots, z_k) = E \left(\max_j \frac{Y(x_j)}{z_j} \right)$ es,

$$P(Z(x_1) \leq z_1, Z(x_2) \leq z_2) = \exp \left\{ -\frac{1}{z_1} T_{v+1} \left[-\frac{\rho(x_1 - x_2)}{b} + \frac{1}{b} \left(\frac{z_2}{z_1} \right)^{1/v} \right] - \frac{1}{z_2} T_{v+1} \left[-\frac{\rho(x_1 - x_2)}{b} + \frac{1}{b} \left(\frac{z_1}{z_2} \right)^{1/v} \right] \right\},$$

donde T_v denota la función de distribución acumulada de una variable aleatoria estándar t-student con $v > 0$ grados de libertad.

Capítulo 5

Simulación de los modelos máx-estables

En este capítulo, se presenta la simulación de los modelos máx-estables discutidos en el capítulo anterior, para ilustrar su comportamiento. Se divide en dos tipos de simulaciones las cuales se darán y explicarán conforme se desglosa el capítulo.

Las simulaciones espaciales permiten la recuperación de información en cualquier localización, en otras palabras, a través de la simulación espacial del proceso puede reconstruir la información en cualquier sitio. Se llaman **simulaciones condicionadas** aquellas que se ajustan a los datos observados; en otro caso, se denominarán **no condicionadas** o **incondicionadas**.

Para estas simulaciones se empleará el Software **R** con el paquete **SpatialExtremes**, este paquete proporciona herramientas para el modelado estadístico de extremos espaciales mediante procesos máx-estable, cópulas o modelos jerárquicos bayesianos. Más precisamente, este paquete permite simulaciones (condicionales) de varios modelos paramétricos máx-estable, análisis de la dependencia espacial extrema, el ajuste de tales procesos usando probabilidades compuestas o mínimos cuadrados (sólo procesos máx-estable simples), verificación y selección de modelos y predicción.

Este capítulo está basado en “Análisis de Valores Extremos” de [Villanueva e Ibáñez \(2016\)](#), “Precipitaciones Máximas en el Estado de Guanajuato, México” de [Moreno y Ortega \(2013\)](#) , “Análisis de máximos para datos espaciales de lluvias” de [González y Sánchez \(2011\)](#), “Spatial extremes: Max-stable processes at

work” de Ribatet (2013), “Spatial Modeling of Extreme Concentrations of Carbon Monoxide Pollution in Urban Regions” de Jiménez Hernández *et al.* (2017), “Simulation of Max-Stable Processes” de Oesting *et al.* (2016), “A User’s Guide to the SpatialExtremes Package” de Ribatet (2009) y “Modelling spatial extremes using max-stable processes” de Ribatet (2017).

5.1. Simulación incondicional

Simulaciones incondicionales son aquellas en las que se toman muestras de una distribución dada sin más restricciones.

La definición de un proceso máx-estable surge de la convergencia del máximo de un número infinito de réplicas normalizado de un proceso aleatorio, es decir, los procesos máx-estables se han definido basándose en infinitas réplicas de un proceso estocástico; sin embargo, en la práctica sólo se podrá obtener un número finito de realizaciones de un proceso, esto es, para fines de simulación, el número de repeticiones es necesariamente finito. A pesar de esto existen resultados teóricos mediante los que será posible obtener simulaciones exactas o aproximadas de algunos modelos máx-estables.

No obstante, Schlather (2002) muestra que es posible obtener simulaciones exactas en una región de muestreo finita.

Teorema 5.1. *Sea Y una función aleatoria medible tal que $\mathbb{E} \left[\int_{\mathbb{R}^d} \max\{0, Y(x)\} dx \right] = 1$, sea Π un proceso de Poisson de $\mathbb{R}^d \times (0, +\infty)$ con medida de intensidad $d\Lambda(y, \xi) = \xi^{-2} dy d\xi$ y $Z(\mathbf{x}) = \max_{i \geq 1} \xi_i \max\{0, Y_i(\mathbf{x})\}$. Suponga que Y está uniformemente acotado por $C \in (0, +\infty)$ y tiene soporte en la bola $b(o, r)$ para algún $r \in (0, +\infty)$. Sea B un conjunto compacto, Y_i son réplicas independientes e idénticamente distribuidas de Y , U_i son variables independientes e idénticamente distribuidas uniformemente en $B_r = \bigcup_{x \in B} b(x, r)$, ξ_i son variables aleatorias independientes e idénticamente distribuidas exponencial estándar y Π, Y_i, ξ_i, U_i mutuamente independientes. Luego, en B ,*

$$Z_*(x) = |B_r| \sup_i \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k}, \quad x \in B, \quad i = 1, 2, \dots \quad (5.1)$$

es igual a los procesos aleatorios máx-estables Z en distribución, y

$$Z_*(x) = |B_r| \sup \left\{ \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k} : i = 1, \dots, m, \text{ } m \text{ es tal que } \frac{C}{\sum_{k=1}^i \xi_k} \leq \max_{1 \leq i \leq m} \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k} \right\} \quad (5.2)$$

casi seguro.

Demostración. Recordar que un proceso de Poisson en \mathbb{R}^d con intensidad 1 puede presentarse como una suma de variables aleatorias exponencial estándar independientes e idénticamente distribuidas es decir $\Pi = \{\sum_{i=1}^n \xi_i : n = 1, 2, \dots\}$. La aplicación del mapeo $\xi \rightarrow |B_r| \xi^{-1}$ a los puntos de Π cede a un nuevo proceso en \mathbb{R}^+ con medida de intensidad $|B_r| \xi^{-2} d\xi$. Como las U_i son distribuciones uniformes, independientes e idénticamente distribuidas en B_r , se tiene que el conjunto aleatorio,

$$\left\{ \left(U_n, \frac{|B_r|}{\sum_{i=1}^n \xi_i} \right) : n = 1, 2, \dots \right\}, \quad (5.3)$$

es un proceso de Poisson en $B_r \times (0, +\infty)$ con medida de intensidad $d\Lambda(y, \xi) = \xi^{-2} dy d\xi$ y el proceso Z_* es igual a Z en distribución.

Para mostrar la segunda afirmación, primero note que m es necesariamente finito como la sucesión $\sum_{k=1}^m \xi_k$ no es decreciente y Y está acotada uniformemente por C de modo que

$$\lim_{m \rightarrow +\infty} \frac{C}{\sum_{k=1}^m \xi_k} = 0.$$

Por tanto existe m finito tal que

$$P \left\{ \max_{1 \leq i \leq m} \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k} \geq \frac{C}{\sum_{k=1}^m \xi_k} \right\} = 1$$

y

$$P \left\{ \left| \max_{1 \leq i \leq m} \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k} - \sup_{i \geq 1} \frac{Y_i(x - U_i)}{\sum_{k=1}^i \xi_k} \right| \neq 0 \right\} = 0$$

□

Si no se cumplen las condiciones del teorema (5.1), es decir, para funciones aleatorias Y cuyo soporte no está incluido en una bola $b(0, r)$ o que no está uniformemente acotada por la constante C , se deben usar aproximaciones para r y

C . Aunque las simulaciones no serán exactas, parece que las aproximaciones para r y C conducen a simulaciones precisas.

5.1.1. Aplicación con R al modelo de Smith

Recuerde que el modelo de Smith está definido por

$$Z(x) = \max_i \{\xi_i f(y_i, x)\}, \quad f(y_i, x) = f_0(x - y_i) \quad (5.4)$$

donde f_0 es la función de densidad normal multivariante con media cero, matriz de covarianza Σ y $\{(\xi_i, y_i)\}_{i \geq 1}$ son los puntos de un proceso de Poisson con medida de intensidad $d\Lambda(y, \xi) = \xi^{-2} d\xi dy$.

El uso del teorema (5.1) para generar realizaciones a partir del modelo de Smith no está teóricamente justificado ya que el soporte de la distribución normal multivariada no está incluido en la bola finita $b(o, r)$, $r < +\infty$. Sin embargo, si r es lo suficientemente grande, la densidad normal multivariante debe estar cerca de o y es poco probable que los puntos $x \in \mathbb{R}^d \setminus b(o, r)$ contribuyan a $Z(x)$.

La sugerencia de Schlather (2002) es usar r tal que $\varphi(r) = 0.001$ es decir $r \approx 3.46$, donde φ denota la densidad normal estándar. Teniendo en cuenta que la varianza en la matriz de covarianza Σ puede ser grande, una opción razonable es dejar

$$r = 3.46 \sqrt{\max\{\sigma_{ii} : i = 1, \dots, d\}} \quad (5.5)$$

donde σ_{ii} son los elemento diagonales de la matriz de covarianza Σ .

Se va a emplear el Software **R** con el paquete *SpatialExtremes*. La función *rmaxstab* permite la simulación del modelo de Smith. Ver código en apéndice A.1.

Las figuras 5.1 y 5.2, a) - f) muestran las simulaciones para el modelo de Smith con diferentes valores del parámetro varianza *var* para el caso univariado y los parámetros de covarianza σ_{11} , σ_{12} y σ_{22} para el caso bivariado.

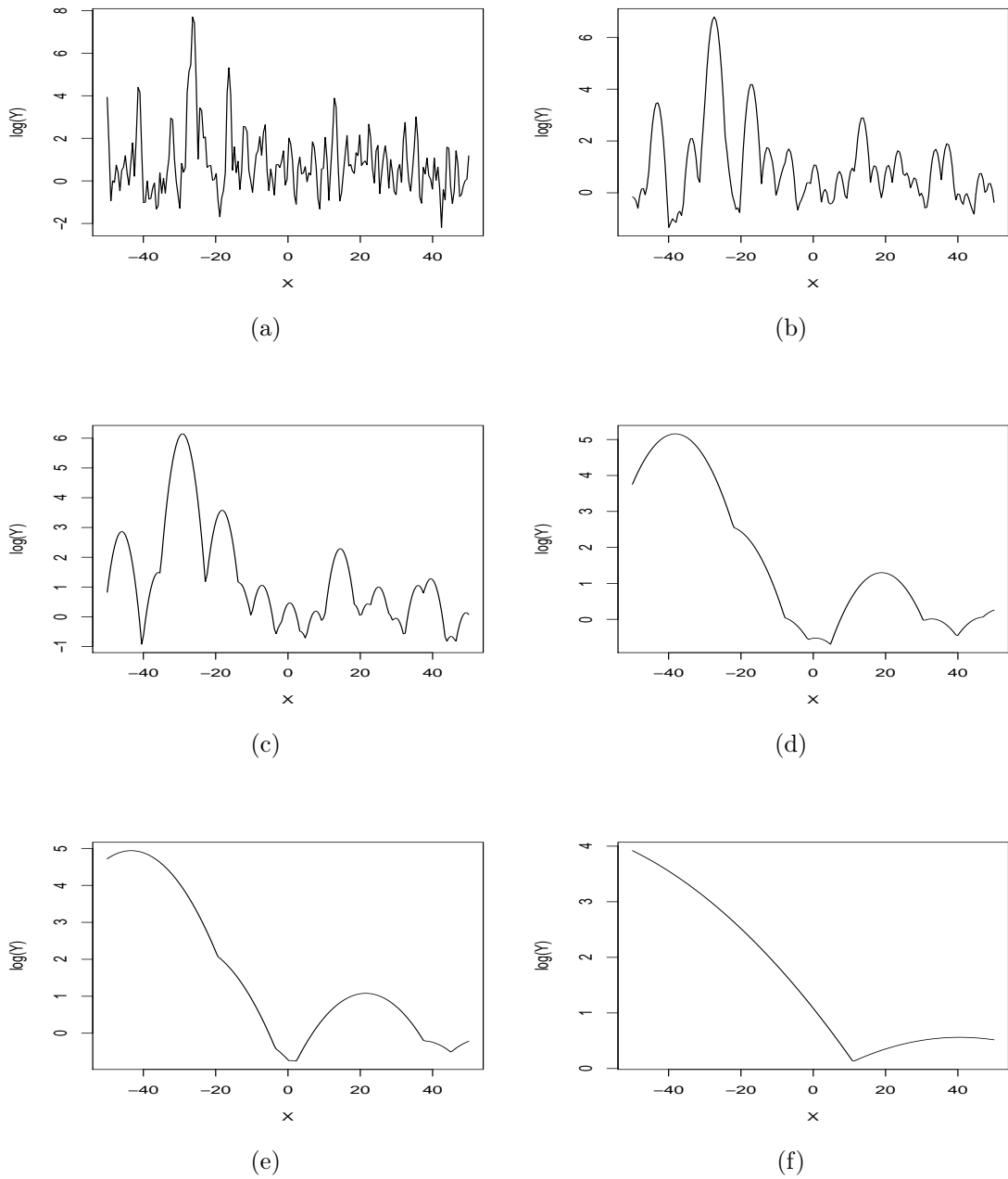


Figura 5.1: Realizaciones del modelo de Smith con marginales Fréchet unitarias en \mathbb{R} , para diferentes valores del parámetro var . (a) $var = 0.1$; (b) $var = 1$; (c) $var = 4$; (d) $var = 50$; (e) $var = 100$; (f) $var = 1000$.

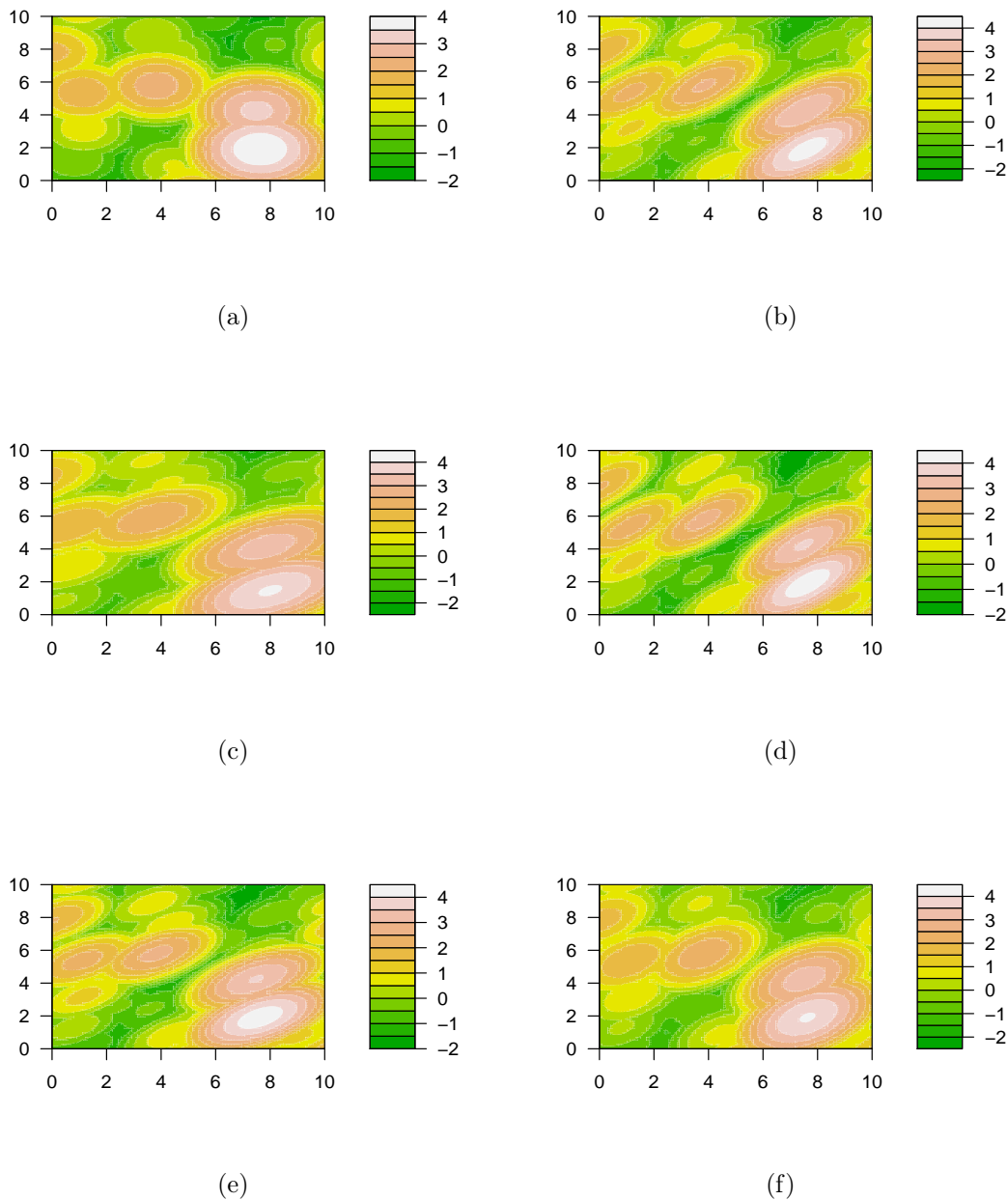


Figura 5.2: Realizaciones del modelo de Smith con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros de la matriz de covarianza. (a) $\sigma_{11} = 9/8$, $\sigma_{12} = 0$, $\sigma_{22} = 9/8$; (b) $\sigma_{11} = 9/8$, $\sigma_{12} = 3/4$, $\sigma_{22} = 9/8$; (c) $\sigma_{11} = 2$, $\sigma_{12} = 3/4$, $\sigma_{22} = 9/8$; (d) $\sigma_{11} = 1$, $\sigma_{12} = 3/4$, $\sigma_{22} = 9/8$; (e) $\sigma_{11} = 9/8$, $\sigma_{12} = 1/2$, $\sigma_{22} = 3/4$; (f) $\sigma_{11} = 9/8$, $\sigma_{12} = 1/2$, $\sigma_{22} = 9/8$.

5.1.2. Aplicación con R al modelo Schlather

Recuerde que el modelo de Schlather está definido por

$$Z(\mathbf{x}) = \max_{i \geq 1} \xi_i \max\{0, Y_i(\mathbf{x})\}, \quad (5.6)$$

donde $Y_i(\cdot)$ son procesos gaussianos estándar estacionarios independientes e idénticamente distribuidos con función de correlación $\rho(h)$, escalados de modo que $\mathbb{E}[\max\{0, Y_i(x)\}] = 1$ y $\{\xi_i\}_{i \geq 1}$ son los puntos de un proceso de Poisson en $(0, \infty)$ con medida de intensidad $d\Lambda(\xi) = \xi^{-2} d\xi$.

Nuevamente, las condiciones requeridas por el teorema (5.1) no se satisfacen ya que los procesos gaussianos no están uniformemente acotados. Sin embargo, Schlather (2002) afirma que al elegir una constante C tal que $P[\max\{0, Y_i(x)\} > C]$ sea pequeño, el procedimiento de simulación sigue siendo preciso y recomienda el uso de $C = 3$.

Como en el modelo anterior se va a emplear el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo de Schlather para algunas funciones de correlación. Ver código en apéndice A.2.

Las figuras 5.3 y 5.4, a) - b) muestran las simulaciones para el modelo de Schlather con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

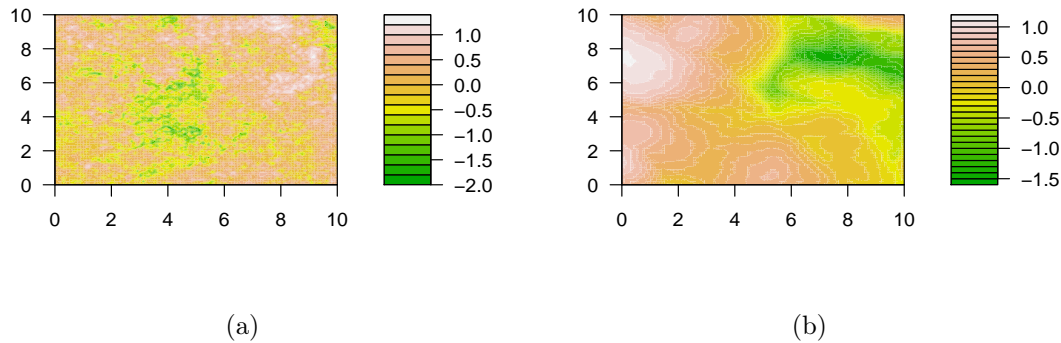


Figura 5.3: Realizaciones del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp \left[- \left(\frac{h}{c_2} \right)^\nu \right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

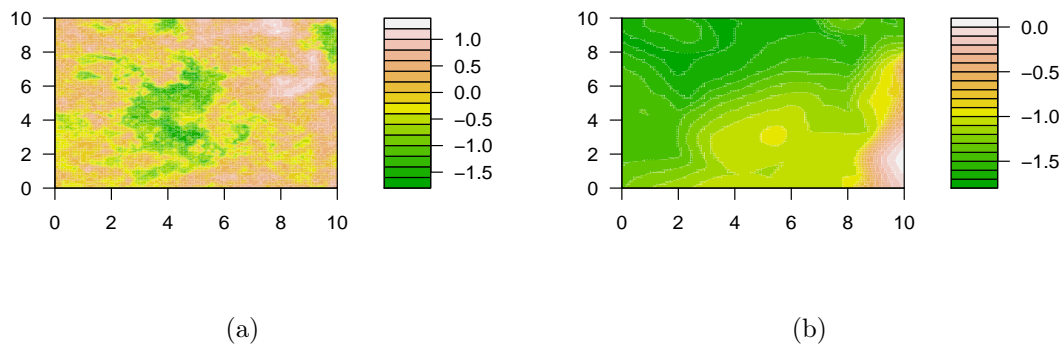


Figura 5.4: Realizaciones del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2} \right)^\nu K_\nu \left(\frac{h}{c_2} \right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

5.1.3. Aplicación con R al modelo Brown-Resnick

Recuerde que el modelo Brown-Resnick surge de elegir la caracterización de la ecuación (3.2) de Schlather (2002),

$$Y_i(x) = \exp \left\{ \epsilon(x) - \frac{\sigma^2(x)}{2} \right\}, \quad (5.7)$$

donde ϵ es un proceso gaussiano continuo con incrementos estacionarios y varianza $\sigma^2(x)$.

De la misma manera que en los modelos previos, las condiciones requeridas por el teorema (5.1) no se satisfacen ya que los procesos gaussianos no están uniformemente acotados. Sin embargo, al elegir una constante $C < \infty$ tal que $P[Y_i(x) \leq C] = 1$ para todo $x \in \chi$, el procedimiento de simulación sigue siendo preciso. Puede ser que la suposición $P[Y_i(x) \leq C] = 1, x \in \chi$, no se cumpla. En tales situaciones, se puede definir un límite pseudo uniforme, esto es, elegir $C < \infty$ tal que $P[Y_i(x) \leq C] = \epsilon$, para algún $\epsilon > 0$ suficientemente pequeño.

Como en los modelos anteriores se va a usar el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo de Brown-Resnick. Ver código en apéndice A.3.

La figura 5.5, a) - d) muestran las simulaciones para el modelo Brown-Resnick con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

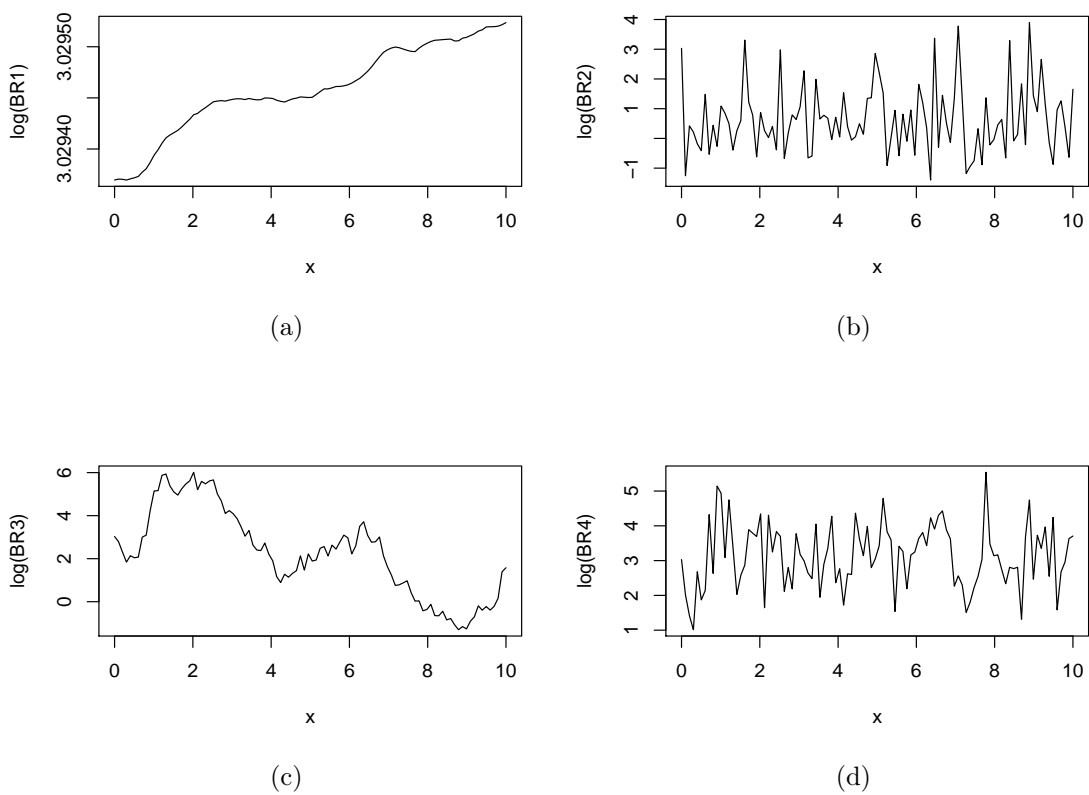


Figura 5.5: Realizaciones del modelo de Brown-Resnick con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros. (a) $c_2 = 100000$, $\nu = 1.9$; (b) $c_2 = 0.00001$, $\nu = 1.5$; (c) $c_2 = 1$, $\nu = 1$; y (d) $c_2 = 0.001$, $\nu = 0.000005$.

5.1.4. Aplicación con **R** al modelo t-extremal

Recuerde que el modelo t-extremal surge de elegir,

$$Y_i(x) = \sqrt{\pi} 2^{-(v-2)/2} \Gamma\left(\frac{v+1}{2}\right)^{-1} \max\{0, \epsilon_i(x)\}^v, \quad v \geq 1, \quad (5.8)$$

donde ϵ es un proceso estándar gaussiano con función de correlación $\rho(\cdot)$, Γ es la función gamma y $d\Lambda(t) = vt^{-(v+1)}dt$ medida de intensidad para el proceso de Poisson en $(0, \infty)$.

Como en todos los modelos anteriores, las condiciones requeridas por el teorema (5.1) no se satisfacen ya que los procesos gaussianos no están uniformemente acotados. Sin embargo, el procedimiento de simulación sigue siendo preciso.

De la misma manera que en los modelos anteriores se va a utilizar el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo t-extremal para algunas funciones de correlación. Ver código en apéndice A.4.

Las figuras 5.6 y 5.7, a) - b) muestran las simulaciones para el modelo t-extremal con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

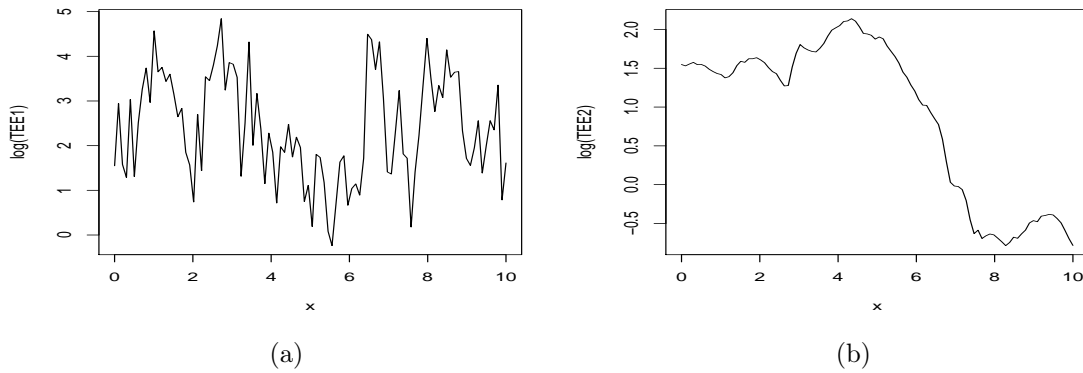


Figura 5.6: Realizaciones del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp \left[- \left(\frac{h}{c_2} \right)^\nu \right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

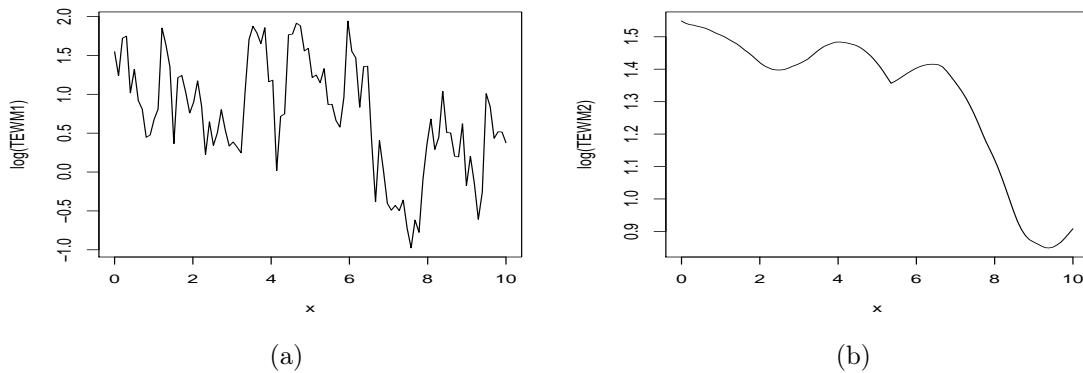


Figura 5.7: Realizaciones del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2} \right)^\nu K_\nu \left(\frac{h}{c_2} \right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

5.2. Simulación condicional

Realizar simulaciones condicionales equivale a obtener realizaciones independientes de una distribución dada con la característica adicional de que cada realización debe satisfacer alguna restricción.

Aunque son posibles otras restricciones, se restringirá la atención a las más naturales, es decir, se realizarán simulaciones a partir de un proceso máx-estable simple sujeto a

$$Z(x_1) = z_1, \dots, Z(x_k) = z_k, \quad x_1, \dots, x_k \in \chi, \quad z_1, \dots, z_k \in (0, \infty), \quad k \geq 1.$$

Este tipo de restricción es vital ya que permite caracterizar el comportamiento del proceso $\{Z(x) : x \in \chi\}$ en algunas ubicaciones nuevas $s \in \chi$ dado que se ha observado algunos valores conocidos en las ubicaciones $x = (x_1, \dots, x_k)$.

5.2.1. Aplicación con **R** al modelo Schlather

Recuerde que el modelo de Schlather está definido por

$$Z(\mathbf{x}) = \max_{i \geq 1} \xi_i \max\{0, Y_i(\mathbf{x})\}, \quad (5.9)$$

donde $Y(\cdot)$ son procesos gaussianos estándar estacionarios independientes e idénticamente distribuidos con función de correlación $\rho(h)$, escalados de modo que $\mathbb{E}[\max\{0, Y_i(x)\}] = 1$ y $\{\xi_i\}_{i \geq 1}$ son los puntos de un proceso de Poisson en \mathbb{R}_*^+ con medida de intensidad $d\Lambda(\xi) = \xi^{-2} d\xi$.

Como en el modelo anterior se va a emplear el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo de Schlather. Ver código en apéndice B.1.

Las figuras 5.8 y 5.9, a) - b) muestran las simulaciones para el modelo de Schlather con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

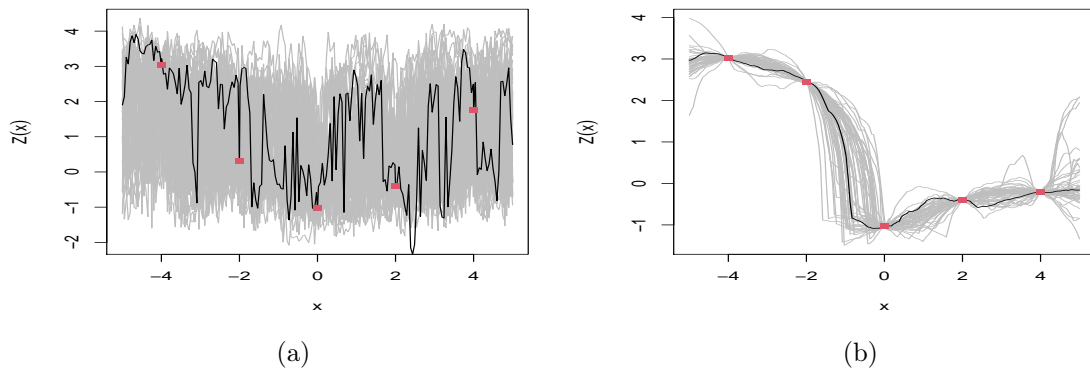


Figura 5.8: Simulaciones condicionales del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

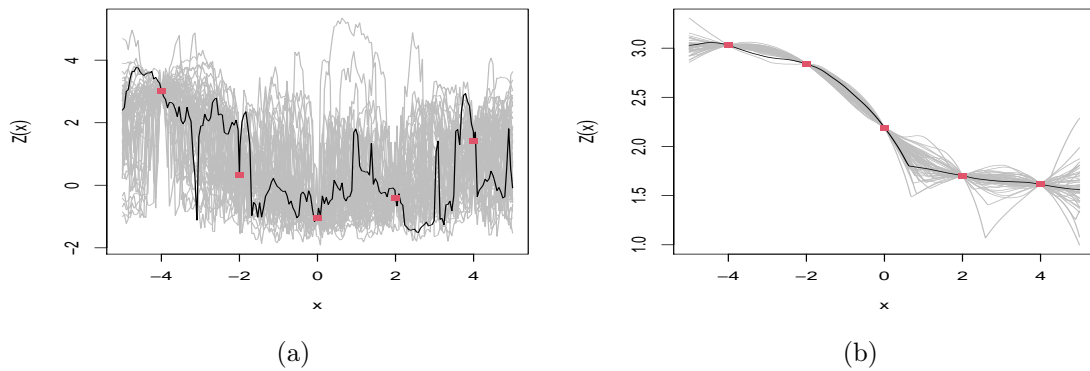


Figura 5.9: Simulaciones condicionales del modelo de Schlather con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

5.2.2. Aplicación con **R** al modelo Brown-Resnick

Recuerde que el modelo Brown-Resnick surge de elegir en la caracterización (3.2) de Schlather (2002),

$$Y_i(x) = \exp \left\{ \epsilon(x) - \frac{\sigma^2(x)}{2} \right\} \quad (5.10)$$

donde ϵ es un proceso gaussiano con incrementos estacionarios y $\sigma^2(x) = V[\epsilon(x)]$.

Como en los modelos anteriores se va a usar el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo de Brown-Resnick. Ver código en apéndice B.2.

La figura 5.10, a) - d) muestran las simulaciones para el modelo Brown-Resnick con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

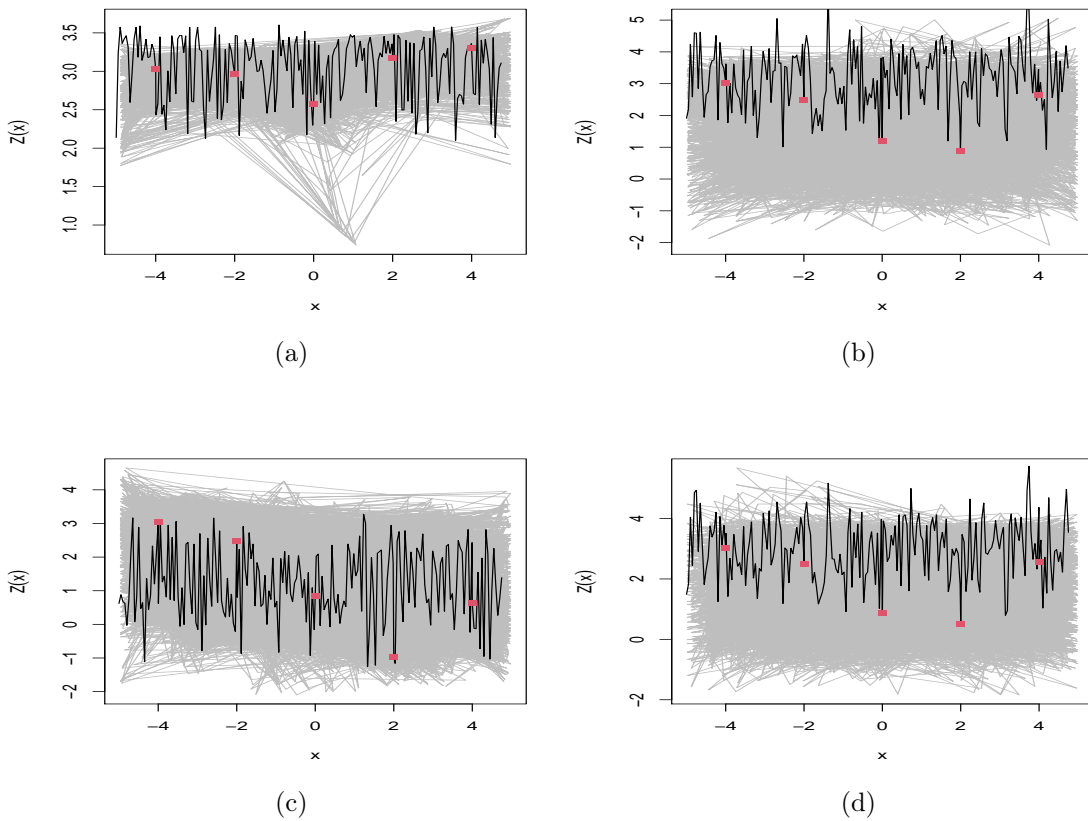


Figura 5.10: Simulaciones condicionales del modelo de Brown-Resnick con marginales Fréchet unitarias en \mathbb{R}^2 , para diferentes valores de los parámetros. (a) $c_2 = 3$, $\nu = 1.9$; (b) $c_2 = 1$, $\nu = 0.1$; (c) $c_2 = 1$, $\nu = 1$; y (d) $c_2 = 0.1$, $\nu = 0.05$.

5.2.3. Aplicación con R al modelo t-extremal

Recuerde que el modelo t-extremal surge de elegir,

$$Y_i(x) = \sqrt{\pi} 2^{-(v-2)/2} \Gamma\left(\frac{v+1}{2}\right)^{-1} \max\{0, \epsilon_i(x)\}^v, \quad v \geq 1, \quad (5.11)$$

donde ϵ es un proceso estándar gaussiano con función de correlación ρ y Γ es la función gamma.

De la misma manera que en los modelos anteriores se va a utilizar el Software **R** con el paquete *SpatialExtremes*. El paquete permite la simulación de procesos máx-estables mediante el uso de la función *rmaxstab*, que se utilizará para la simulación del modelo t-extremal. Ver código en apéndice [B.3](#).

Las figuras [5.11](#) y [5.12](#), a) - b) muestran las simulaciones para el modelo t-extremal con diferentes valores de los parámetros rango (c_2) y suavizado (ν).

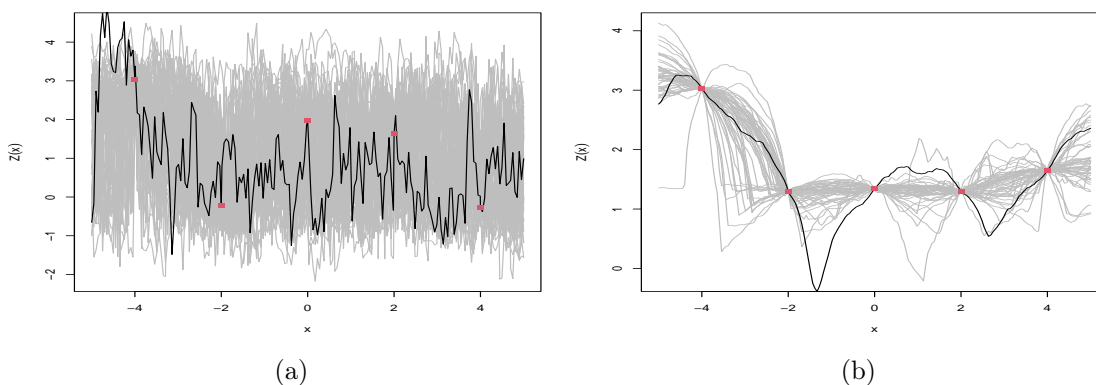


Figura 5.11: Simulaciones condicionales del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación exponencial potencia $\rho(h) = \exp\left[-\left(\frac{h}{c_2}\right)^\nu\right]$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

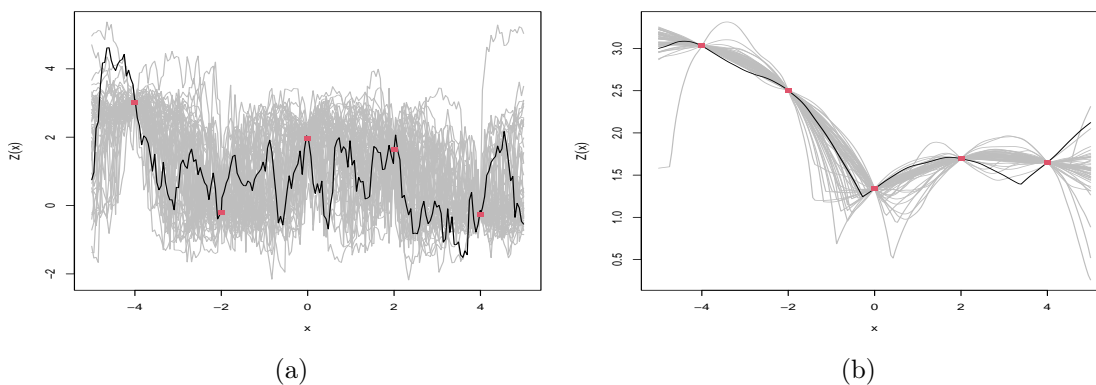


Figura 5.12: Simulaciones condicionales del modelo t-extremal con marginales Fréchet unitarias en \mathbb{R}^2 usando la función de correlación Whittle-Matérn $\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{h}{c_2}\right)^\nu K_\nu\left(\frac{h}{c_2}\right)$, para diferentes valores de los parámetros. (a) $c_2 = 2$, $\nu = 0.5$ y (b) $c_2 = 4$, $\nu = 1.9$.

Conclusiones

El principal objetivo de este trabajo ha sido realizar una revisión y presentación de la teoría de los principales enfoques en el análisis de valores extremos, centrándose principalmente en los procesos máx-estables y su uso para modelar extremos espaciales empleando el paquete **SpatialExtremes**. Se enfoca en modelar máximos puntuales en los casos univariados y multivariados, para después simular los casos bivariados únicamente. En la última década han surgido muchos avances en la teoría de valores extremos, fundamentalmente en el campo de la geoestadística a través del uso de extremos espaciales, siendo los procesos máx-estables una solución elegante para modelizar estos valores.

Debido al tipo de función que presentan algunos modelos no fue del todo posible poder generar las simulaciones. De tal modo que se presenta únicamente aquellas que si fueron simuladas con sus debidos parámetros. Como ejemplo a este problema es la simulación condicional del modelo de Smith.

Al realizar la simulación del modelo de Smith, en las gráficas se observa un patrón de huellas ovaladas que varían ínfimamente en forma y color de acuerdo a la variación numérica de la paleta de colores. Esto quiere decir que el fenómeno como tal al variar los parámetros este mantiene un cierto patrón.

El modelo de Smith presenta dificultades con la cercanía a cero de los valores propios de la matriz de varianzas y covarianzas, lo que tiene por consecuencia problemas de invertibilidad de la matriz.

Para el modelo de Schlather, en estas simulaciones no fue posible observar un patrón en el comportamiento de los datos del fenómeno como ocurriera en el caso de los procesos simulados bajo el modelo de Smith. Pero si es posible notar que al

variar el parámetro rango afecta de manera mínima en los resultados, en cambio al variar el parámetro suavidad afecta considerablemente en los resultados. Las simulaciones arrojan resultados casi idénticos, por lo que no se tiene la seguridad para afirmar que la variación en la elección de la función de correlación modifica de manera significativa los resultados obtenidos.

En general, es bastante difícil decidir qué representación espectral elegir para la simulación de un proceso de Brown-Resnick. Esta elección depende tanto del semivariograma correspondiente como del dominio de simulación.

El modelo t-extremal presenta una alta flexibilidad, lo que lo hace muy útil en la modelación. Tanto en la modelación global como local, variando los grados de libertad de un período al otro, es el modelo que mejor se ajusta en la mayor parte de los casos.

Las respuestas a las preguntas que se formulan como parte del planteamiento del problema abordado en esta tesis, si bien cada uno de los modelos analizados dan predicciones significativas y de mucha utilidad, el modelo t-extremal además de mostrar una alta flexibilidad, muestra que puede dar predicciones de mayor exactitud en comparación a los demás modelos, debido a que es una generalización de ellos. El problema que podría presentar es la alta demanda de procesos que necesita para realizar grandes simulaciones.

Con este trabajo se proporciona una base teórica sólida para aquellos interesados en modelar extremos espaciales. Finalmente, el modelado de extremos espaciales utilizando argumentos de valores extremos es un área de investigación extremadamente activa y se espera que se logren avances en los próximos años. Sin embargo, hasta ahora, la implementación de software de esta metodología es bastante limitada, pero puede haber grandes avances en un futuro cercano.

Apéndice A

Códigos de simulaciones incondicionales

A.1. Modelo de Smith

A.1.1. Modelo de Smith unidimensional

Se considera una sucesión *coord* entre -50, 50 con 200 elementos. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, *coord* un vector de coordenadas de cada ubicación, *cov.mod* = "gauss" el cual corresponde al modelo de Smith. Se especifica el parámetro *var*. Finalmente con *plot* se obtienen los gráficos.

```
> library(SpatialExtremes)
> n.site <- 200
> coord <- seq(-50, 50, length = n.site)
> set.seed(7)
> all.data <- rmaxstab(1, coord, "gauss",
+ var=1000)
> plot(coord, log(all.data), type = "l",
+ xlab = "X", ylab = "log(Y)")
```

A.1.2. Modelo de Smith bidimensional

Se consideran dos sucesiones x y y entre 0 y 10 con 100 elementos. En seguida *cbind*(x, y) toma las sucesiones de datos generadas con anterioridad y los combina

por columnas. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, *coord* un vector de coordenadas de cada ubicación, *cov.mod* = "gauss" el cual corresponde al modelo de Smith. Debido a que *coord* es univariado, se especifican los parámetros de la matriz de covarianza, *cov11*, *cov12* y *cov22*. Como las coordenadas representan puntos de la cuadrícula, *grid = TRUE*. Finalmente con *filled.contour()* se obtienen los gráficos de contorno con las áreas entre las curvas de nivel rellenas en color sólido (Cleveland llama a esto un gráfico de nivel). A la derecha del gráfico se muestra una clave que muestra cómo se asignan los colores a los valores de las realizaciones.

```
> library(SpatzialExtremes)
> x <- y <- seq(0, 10, length = 100)
> coord <- cbind(x, y)
> set.seed(8)
> Sm1 <- rmaxstab(1, coord, "gauss", cov11 = 9/8,
+ cov12 = 0, cov22 = 9/8, grid = TRUE)
> set.seed(8)
> Sm2 <- rmaxstab(1, coord, "gauss", cov11 = 9/8,
+ cov12 = 3/4, cov22 = 9/8, grid = TRUE)
> set.seed(8)
> Sm3 <- rmaxstab(1, coord, "gauss", cov11 = 2,
+ cov12 = 3/4, cov22 = 9/8, grid = TRUE)
> set.seed(8)
> Sm4 <- rmaxstab(1, coord, "gauss", cov11 = 1,
+ cov12 = 3/4, cov22 = 9/8, grid = TRUE)
> set.seed(8)
> Sm5 <- rmaxstab(1, coord, "gauss", cov11 = 9/8,
+ cov12 = 1/2, cov22 = 3/4, grid = TRUE)
> set.seed(8)
> Sm6 <- rmaxstab(1, coord, "gauss", cov11 = 9/8,
+ cov12 = 1/2, cov22 = 9/8, grid = TRUE)
> filled.contour(x, y, log(Sm1), color.palette =
+ terrain.colors)
> filled.contour(x, y, log(Sm2), color.palette =
+ terrain.colors)
```



```

> filled.contour(x, y, log(Sm3), color.palette =
+ terrain.colors)
> filled.contour(x, y, log(Sm4), color.palette =
+ terrain.colors)
> filled.contour(x, y, log(Sm5), color.palette =
+ terrain.colors)
> filled.contour(x, y, log(Sm6), color.palette =
+ terrain.colors)

```

A.2. Modelo de Schlather

A.2.1. Modelo de Schlather bidimensional

Se consideran dos sucesiones x y y entre 0 y 10 con 100 elementos. En seguida $cbind(x, y)$ toma las sucesiones de datos generadas con anterioridad y los combina por columnas. Con la función $rmaxstab$ se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, $coord$ un vector de coordenadas de cada ubicación, $cov.mod$ el cual corresponde a la función de correlación a utilizar del modelo de Schlather. Debido a que las funciones de correlación son todas válidas, el efecto $nugget = 0$, se especifican los parámetros $range$ y $smooth$. Como las coordenadas representan puntos de la cuadrícula, $grid = TRUE$. Finalmente con $filled.contour()$ se obtienen los gráficos de contorno con las áreas entre las curvas de nivel rellenas en color sólido (Cleveland llama a esto un gráfico de nivel). A la derecha del gráfico se muestra una clave que muestra cómo se asignan los colores a los valores de las realizaciones.

```

> library(SpatialExtremes)
> x <- y <- seq(0, 10, length = 100)
> coord <- cbind(x, y)
> #Realizaciones con la función exponencial
> set.seed(5)
> ShE1 <- rmaxstab(1, coord, cov.mod = "powexp",
+ nugget = 0, range = 2, smooth = 0.5, grid = TRUE)
> set.seed(5)
> ShE2 <- rmaxstab(1, coord, cov.mod = "powexp",
+ nugget = 0, range = 4, smooth = 1.9, grid = TRUE)

```

```

> filled.contour(x, y, log(ShE1),color.palette =
+ terrain.colors)
> filled.contour(x, y, log(ShE2),color.palette =
+ terrain.colors)
> #Realizaciones con la función Whittle-Matérn
> set.seed(5)
> ShWM1 <- rmaxstab(1, coord, cov.mod = "whitmat",
+ nugget = 0, range = 2, smooth = 0.5, grid = TRUE)
> set.seed(5)
> ShWM2 <- rmaxstab(1, coord, cov.mod = "whitmat",
+ nugget = 0, range = 4, smooth = 1.9, grid = TRUE)
> filled.contour(x, y, log(ShWM1),color.palette =
+ terrain.colors)
> filled.contour(x, y, log(ShWM2),color.palette =
+ terrain.colors)

```

A.3. Modelo de Brown-Resnick

A.3.1. Modelo de Brown-Resnick bidimensional

Se considera una sucesión x entre 0 y 10 con 100 elementos. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, x un vector de coordenadas de cada ubicación, *cov.mod* = "brown" el cual corresponde al modelo de Brown-Resnick. Se especifican los parámetros *range* y *smooth*. Finalmente con *plot* se obtienen los gráficos.

```

> library(SpatialExtremes)
> x <- seq(0, 10, length = 100)
> set.seed(7)
> BR1 <- rmaxstab(1, x, cov.mod = "brown",
+ range = 100000, smooth = 1.9)
> set.seed(7)
> BR2 <- rmaxstab(1, x, cov.mod = "brown",
+ range = 0.00001, smooth = 1.5)
> set.seed(7)
> BR3 <- rmaxstab(1, x, cov.mod = "brown",

```

```

+ range =1, smooth =1 )
> set.seed(7)
> BR4 <- rmaxstab(1, x, cov.mod = "brown",
+ range =0.001, smooth =0.000005)
> plot(x, log(BR1), type = "l")
> plot(x, log(BR2), type = "l")
> plot(x, log(BR3), type = "l")
> plot(x, log(BR4), type = "l")

```

A.4. Modelo t-extremal

A.4.1. Modelo t-extremal bidimensional

Se considera una sucesión x entre 0 y 10 con 100 elementos. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, *coord* un vector de coordenadas de cada ubicación, *cov.mod* el cual corresponde a la función de correlación a utilizar del modelo de t-extremal. Debido a que las funciones de correlación son todas válidas, el efecto *nugget* = 0, se especifican los parámetros *range*, *smooth* y *DoF*. Finalmente con *plot* se obtienen los gráficos.

```

> library(SpatialExtremes)
> x <-seq(0, 10, length = 100)
> set.seed(6)
> TEE1 <- rmaxstab(1, x,"tpowexp", DoF = 4,
+ nugget = 0, range = 2, smooth = 0.5)
> set.seed(6)
> TEE2 <- rmaxstab(1, x,"tpowexp", DoF = 4,
+ nugget = 0, range = 4, smooth = 1.9)
> plot(x, log(TEE1), type = "l")
> plot(x, log(TEE2), type = "l")
> set.seed(6)
> TEWM1 <- rmaxstab(1, x, "twhitmat", DoF = 4,
+ nugget = 0, range = 2, smooth = 0.5)
> set.seed(6)
> TEWM2 <- rmaxstab(1, x, "twhitmat", DoF = 4,

```

```
+ nugget = 0, range = 4, smooth = 1.9)  
> plot(x, log(TEWM1), type = "l")  
> plot(x, log(TEWM2), type = "l")
```

Apéndice B

Códigos de simulaciones condicionales

B.1. Modelo de Schlather

B.1.1. Modelo de Schlather bidimensional

Se consideran dos sucesiones *coord* y *cond.coord* entre -5, 5 y -4, 4 con 200 y 5 elementos respectivamente. En seguida *c(cond.coord, coord)* toma las sucesiones de datos generadas con anterioridad y los combina por columnas. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, *n = 1* número de observaciones, *all.coord* un vector de coordenadas de cada ubicación, *cov.mod* el cual corresponde a la función de correlación a utilizar del modelo de Schlather. Debido a que las funciones de correlación son todas válidas, el efecto *nugget = 0*, se especifican los parámetros *range* y *smooth*. Como las coordenadas representan puntos de la cuadrícula, *grid = TRUE*. En seguida con la función *condrmaxstab* se genera una simulación condicional de procesos máx-estables con los parámetros, *n.sim = 50* el número de simulaciones condicionales que se generarán, *coord* un vector o matriz que proporciona las coordenadas de cada ubicación, *cond.coord* un vector o matriz que proporciona las coordenadas de cada ubicación condicional, *cond.data* un vector que da los valores condicionales en las ubicaciones de condicionamiento correspondientes, *cov.mod* el cual corresponde a la función de correlación a utilizar del modelo de Schlather. Debido a que las funciones de correlación son todas válidas, el efecto *nugget = 0*, se especifican los parámetros *range* y *smooth*.

Finalmente con `matplot()` se grafica las simulaciones condicionales, con `lines()` se grafica las realizaciones del proceso máx-estable, con `points()` se grafican puntos sobre el proceso máx-estable.

```
> library(SpatialExtremes)

> n.sim <- 50
> n.cond <- 5

> range <- 4
> smooth <- 1.9

> n.site <- 200
> coord <- seq(-5, 5, length = n.site)
> cond.coord <- seq(-4, 4, length = n.cond)
> all.coord <- c(cond.coord, coord)

> #Exponencial potencia
> set.seed(7)
> all.cond.data <- rmaxstab(1, all.coord, "powexp",
+ nugget = 0, range = range, smooth = smooth)
> cond.data <- all.cond.data[1:n.cond]

> ans <- condrmaxstab(n.sim, coord, cond.coord, cond.
+ data, range = range, smooth = smooth, cov.mod = "powexp")

> idx <- order(all.coord)
> matplot(coord, t(log(ans$sim)), type = "l", col = "grey",
+ lty = 1, xlab = expression(x), ylab = expression(Z(x)))
> lines(all.coord[idx], log(all.cond.data)[idx])
> points(cond.coord, log(cond.data), pch = 15, col = 2)

> #Whittle-Matérn
> set.seed(7)
> all.cond.data <- rmaxstab(1, all.coord, "whitmat",
```

```

+ nugget = 0, range = range, smooth = smooth)
> cond.data <- all.cond.data[1:n.cond]

> ans <- condrmaxstab(n.sim, coord, cond.coord, cond.data,
+range = range, smooth = smooth, cov.mod = "whitmat")
> idx <- order(all.coord)
> matplot(coord, t(log(ans$sim)), type = "l", col = "grey",
+ lty = 1, xlab = expression(x), ylab = expression(Z(x)))
> lines(all.coord[idx], log(all.cond.data)[idx])
> points(cond.coord, log(cond.data), pch = 15, col = 2)

```

B.2. Modelo de Brown-Resnick

B.2.1. Modelo de Brown-Resnick bidimensional

Se consideran dos sucesiones *coord* y *cond.coord* entre -5, 5 y -4, 4 con 200 y 5 elementos respectivamente. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, *coord* un vector de coordenadas de cada ubicación, *cov.mod* el cual corresponde al modelo de Brown-Resnick. Se especifican los parámetros *range* y *smooth*. En seguida con la función *condrmaxstab* se genera una simulación condicional de procesos máx-estables con los parámetros, $n.sim = 50$ el número de simulaciones condicionales que se generarán, *coord* un vector o matriz que proporciona las coordenadas de cada ubicación, *cond.coord* un vector o matriz que proporciona las coordenadas de cada ubicación condicional, *cond.data* un vector que da los valores condicionales en las ubicaciones de condicionamiento correspondientes, se especifican los parámetros *range* y *smooth*.

Finalmente con *matplot()* se grafica las simulaciones condicionales, con *lines()* se grafica las realizaciones del proceso máx-estable, con *points()* se grafican puntos sobre el proceso máx-estable.

```

> library(SpatialExtremes)

> n.sim <- 50
> n.cond <- 5

```

```

> range <- 0.1
> smooth <- 0.05

> n.site <- 200
> coord <- runif(n.site,-5, 5)
> cond.coord <- seq(-4, 4, length = n.cond)

> set.seed(7)
> all.cond.data <- rmaxstab(1, coord, "brown",
+ nugget = 0, range = range, smooth = smooth)
> cond.data <- all.cond.data[1:n.cond]

> ans <- condrmaxstab(n.sim, coord, cond.coord,
+ cond.data, range = range, smooth = smooth)

> idx <- order(all.coord)
> matplot(coord, t(log(ans$sim)), type = "l",
> col = "grey", lty = 1, xlab = expression(x),
> ylab = expression(Z(x)))
> lines(all.coord[idx], log(all.cond.data)[idx])
> points(cond.coord, log(cond.data), pch = 15, col = 2)

```

B.3. Modelo t-extremal

B.3.1. Modelo t-extremal bidimensional

Se consideran dos sucesiones *coord* y *cond.coord* entre -5, 5 y -4, 4 con 200 y 5 elementos respectivamente. En seguida $c(\text{cond.coord}, \text{coord})$ toma las sucesiones de datos generadas con anterioridad y los combina por columnas. Con la función *rmaxstab* se genera la realización de un proceso máx-estable con los parámetros, $n = 1$ número de observaciones, *all.coord* un vector de coordenadas de cada ubicación, *cov.mod* el cual corresponde a la función de correlación a utilizar del modelo t-extremal. Debido a que las funciones de correlación son todas válidas,

el efecto $nugget = 0$, se especifican los parámetros $range$ y $smooth$. Como las coordenadas representan puntos de la cuadrícula, $grid = TRUE$. En seguida con la función `condrmaxstab` se genera una simulación condicional de procesos máx-estables con los parámetros, $n.sim = 50$ el número de simulaciones condicionales que se generarán, $coord$ un vector o matriz que proporciona las coordenadas de cada ubicación, $cond.coord$ un vector o matriz que proporciona las coordenadas de cada ubicación condicional, $cond.data$ un vector que da los valores condicionales en las ubicaciones de condicionamiento correspondientes, $cov.mod$ el cual corresponde a la función de correlación a utilizar del modelo de t-extremal. Debido a que las funciones de correlación son todas válidas, el efecto $nugget = 0$, se especifican los parámetros $range$ y $smooth$.

Finalmente con `matplot()` se grafica las simulaciones condicionales, con `lines()` se grafica las realizaciones del proceso máx-estable, con `points()` se grafican puntos sobre el proceso máx-estable.

```
> library(SpatialExtremes)

> n.sim <- 50
> n.cond <- 5

> range <- 4
> smooth <- 1.9

> n.site <- 200
> coord <- seq(-5, 5, length = n.site)
> cond.coord <- seq(-4, 4, length = n.cond)
> all.coord <- c(cond.coord, coord)

> #Exponencial potencia
> set.seed(7)
> all.cond.data <- rmaxstab(1, all.coord, "tpowexp",
+ DoF = 4, nugget = 0, range = range, smooth = smooth)
> cond.data <- all.cond.data[1:n.cond]

> ans <- condrmaxstab(n.sim, coord, cond.coord, cond.data,
+ range = range, smooth = smooth, cov.mod = "whitmat")
```

```
> idx <- order(all.coord)
> matplot(coord, t(log(ans$sim)), type = "l", col = "grey",
+ lty = 1, xlab = expression(x), ylab = expression(Z(x)))
> lines(all.coord[idx], log(all.cond.data)[idx])
> points(cond.coord, log(cond.data), pch = 15, col = 2)

> #Whittle-Matérn
> set.seed(7)
> all.cond.data <- rmaxstab(1, all.coord, "twhitmat",
+ DoF = 4, nugget = 0, range = range, smooth = smooth)
> cond.data <- all.cond.data[1:n.cond]

> ans <- condrmaxstab(n.sim, coord, cond.coord, cond.data,
> range = range, smooth = smooth, cov.mod = "whitmat")
> idx <- order(all.coord)
> matplot(coord, t(log(ans$sim)), type = "l", col = "grey",
+ lty = 1, xlab = expression(x), ylab = expression(Z(x)))
> lines(all.coord[idx], log(all.cond.data)[idx])
> points(cond.coord, log(cond.data), pch = 15, col = 2)
```

Bibliografía

- Brown, B. M. y Resnick, S. I. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14, 4, 732–739.
- Castillo, E., Hadi, A. S., Balakrishnan, N. y Sarabia, J.-M. (2005). Extreme value and related models with applications in engineering and science.
- Coles, S., Bawa, J., Trenner, L. y Dorazio, P. (2001). *An introduction to statistical modeling of extreme values*, tomo 208. Springer.
- de Haan, L. (1984). A spectral representation for max-stable processes. *The annals of probability*, 12, 4, 1194–1204.
- de Haan, L. y Ferreira, A. (2006). *Extreme value theory: an introduction*. Springer Science & Business Media.
- Dombry, C., Éyi Minko, F. y Ribatet, M. (2013). Conditional simulation of max-stable processes. *Biometrika*, 100, 1, 111–124.
- Gaetan, C. y Guyon, X. (2010). *Spatial statistics and modeling*, tomo 90. Springer.
- George, C. C. (1988). Probabilidad y estadística, aplicaciones y métodos. *McGrawHill/Interamericana de México, SA. México*.
- González, E. M. y Sánchez, J. O. (2011). *Análisis de máximos para datos espaciales de lluvias*. Proyecto Fin de Carrera, Maestría en Ciencias con Orientación en Probabilidad y Estadística, Cimat.
- Gorgas, J., Cardiel, N. y Zamorano, J. (2011). Estadística básica para estudiantes de Ciencias. *Universidad Complutense de Madrid, España, Edición de Febrero*.
- Jiménez Hernández, J. d. C., Vaquera Huerta, H. y Pérez Rodríguez, P. (2017). Spatial modeling of extreme concentrations of carbon monoxide pollution in

- urban regions. *American Journal of Environmental Engineering and Science*, 4, 6, 60–70.
- Kabluchko, Z., Schlather, M., De Haan, L. *et al.* (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37, 5, 2042–2065.
- Kotz, S. y Nadarajah, S. (2000). *Extreme value distributions: theory and applications*. World Scientific.
- Moreno, L. y Ortega, J. (2013). *Precipitaciones Máximas en el Estado de Guanajuato, México*. Proyecto Fin de Carrera, Magister en Ingeniería Matemática, Universidad de la República.
- Moreno, L. y Sánchez, J. O. (2016). Bondad del ajuste en predicciones de precipitaciones extremas.
- Oesting, M., Ribatet, M. y Dombry, C. (2016). Simulation of max-stable processes. *Extreme value modeling and risk analysis. Chapman and Hall/CRC, Boca Raton*, 195–214.
- Olmo, J. (2005). Aplicaciones de la teoría de valores extremos a la gestión del riesgo.
- Opitz, T. (2013). Extremal t processes: Elliptical domain of attraction and a spectral representation. *Journal of Multivariate Analysis*, 122, 409–413.
- Ribatet, M. (2009). A user's guide to the SpatialExtremes package. *EPFL, Lausanne, Switzerland*.
- Ribatet, M. (2013). Spatial extremes: Max-stable processes at work. *Journal de la Société Française de Statistique*, 154, 2, 156–177.
- Ribatet, M. (2017). Modelling spatial extremes using max-stable processes.
- Ribatet, M. y Sedki, M. (2013). Extreme value copulas and max-stable processes. *Journal de la Société Française de Statistique*, 154, 1, 138–150.
- Rincón, L. (2014). *Introducción a la probabilidad*. Universidad Nacional Autónoma de México, Facultad de Ciencias.

- Romero, F. M. y Carrasco, D. H. (2020). *Matemáticas y sus aplicaciones 15*. Benemérita Universidad Autónoma de Puebla.
- Sachkov, V. N. (1997). *Probabilistic methods in combinatorial analysis*, tomo 56. Cambridge University Press.
- Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5, 1, 33–44.
- Smith, R. L. (1990). Max-stable processes and spatial extremes. *Unpublished manuscript*, 205.
- Villanueva e Ibáñez (2016). *Análisis de Valores Extremos*. Proyecto Fin de Carrera, Maestría en Estadística Aplicada, Universidad de Granada.