

UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA  
INSTITUTO DE FÍSICA Y MATEMÁTICAS



**CARACTERIZACIÓN DE EVENTOS  
EXTREMOS EN SERIES DE TIEMPO  
PRODUCIDAS POR EL MAPEO DE LÉVY**

TESIS

PARA OBTENER EL TÍTULO DE:  
**INGENIERO EN FÍSICA APLICADA**

PRESENTA

**ARTEMISA VILLALOBOS RAMÍREZ**

DIRECTOR DE TESIS

**Dr. JOSÉ ANTONIO MÉNDEZ BERMÚDEZ**

Co-DIRECTOR DE TESIS

**Dr. HUGO DAVID SÁNCHEZ CHÁVEZ**



*DEDICATORIA*

*Dédico esté trabajo a mi madre,  
quien me ha sabido apoyar.  
Daniela Villalobos Ramírez.*



# Agradecimientos

Quiero expresar mi sincero agradecimiento a mi madre Daniela Villalobos Ramírez quien me ha brindado su apoyo y amor a pesar de los tropiezos que he tenido a lo largo de este camino. Por estar conmigo en los buenos tiempos y también en los malos, además de que siempre ha tratado de ayudarme en todo. De todo corazón te doy las gracias.

Agradezco a mi *brother* Elizabeth Brena Villalobos por todo el apoyo que me ha dado, todo el amor que me ha demostrado y ser la mejor compañera de vida. Gracias por todo Eli.

Agradezco a mi *brother* Lupita porque siempre ha creído en mi y me sabe demostrar su amor. A mis tíos, Alan e Ivonne, porque me han querido como una hija. Y a mi persona favorita Keyla por esas alegrías que siempre lleva a donde vaya.

Agradezco a mi director de tesis, el Dr. José Antonio Méndez Bermudez por darme la oportunidad de realizar este trabajo bajo su asesoramiento y brindarme su tiempo y conocimiento; a mi co-director, el Dr. Hugo David Sánchez Chávez, por el apoyo interno, consejos y confianza que me brindó. También, a mis sinodales, Dr. Ricardo Rosas Rodríguez, Dr. Jorge González García y a M.C. Maxwell Gustavo Jiménez Escamilla, por las revisiones que me proporcionaron en este trabajo.

Me encuentro agradecida con la universidad (UTM) por todo lo que aprendí en esta, los buenos profesores que tuve, los amigos que hice, en particular mis amistades con Claudia y Enrique.



# Resumen

Este trabajo consistió en la caracterización de eventos extremos en series producidos por el mapeo de Lévy. Por lo que, para desarrollar esta tesis se comenzó reproduciendo resultados previamente reportados de distintos mapeos, para sus diferentes casos (integrable, mixto y caótico). Presentando así nuestro modelo, el cual es el mapeo de Lévy.

Partiendo ahora del régimen del caos global se crearon las series de tiempo discretas creadas del mapeo de Lévy, el cual es un generador de números pseudo-aleatorios descrito por dicha distribución por una ley de potencia. Para poder caracterizar los eventos, se crearon umbrales ( $q_n$ ). De estos umbrales se distinguieron los intervalos de ocurrencia, obteniendo una nueva distribución a partir de la diferencia de tiempos entre eventos. Usamos la misma metodología, pero ahora, para una serie de retornos financieros con el fin de observar alguna relación con nuestro modelo y esta serie realista.



# Abstract

This work consisted in the characterization of extreme events in series of Lévy's map. Therefore, to develop this thesis, we started by reproducing previously reported results of the different maps, for the different cases (integrable, mixed and chaotic). Introducing our model, which is Lévy's mapping.

Starting now from the global chaos regime, the discrete time series created from the Lévy mapping were created, which is a pseudo-random generator number described by said distribution by a power law. In order to characterize the events, thresholds were created ( $q_n$ ). From these thresholds, the intervals of occurrence were distinguished, obtaining a new distribution from the time difference between events. We use the same methodology, but now, for a series of financial returns in order to observe some relationship between our model and this realistic series.



# Índice

<b>Agradecimientos</b>	<b>iii</b>
<b>Resumen</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Aspectos preliminares</b>	<b>1</b>
<b>Introducción</b>	<b>3</b>
<b>1. Mapeos</b>	<b>7</b>
1.1. Sistemas Dinámicos . . . . .	7
1.2. Caos . . . . .	8
1.3. El Mapeo estándar . . . . .	9
1.3.1. Rotor pateado . . . . .	9
1.3.2. Mapeo estándar . . . . .	10
1.4. El mapeo de Poincaré . . . . .	11
1.4.1. Billar cosenoidal o billar ondulado . . . . .	11
1.4.2. Mapeo de Poincaré . . . . .	12
1.5. El mapeo de Lévy . . . . .	13
1.5.1. Densidad de distribución tipo Lévy . . . . .	13
1.5.2. Mapeo de Lévy . . . . .	14
<b>2. Las series de tiempo</b>	<b>19</b>
2.1. Series de tiempo . . . . .	19
2.2. Tipos de series . . . . .	20
2.3. Análisis de las series . . . . .	22
2.3.1. Función de distribución . . . . .	22
2.3.2. Gráficas . . . . .	23
2.4. Series producidas por el mapeo de Lévy . . . . .	24
2.5. Series financieras . . . . .	24
2.5.1. Criptomoneda . . . . .	25

2.5.2. Retornos . . . . .	26
<b>3. Eventos extremos</b>	<b>29</b>
3.1. Eventos extremos . . . . .	29
3.2. Teoría de los valores extremos . . . . .	30
3.2.1. Estimador de Hill . . . . .	31
3.3. Desarrollo . . . . .	31
3.3.1. Consideraciones . . . . .	31
3.3.2. Umbral sobre umbral . . . . .	31
3.3.3. Retornos de los precios de las criptomoneda . . . . .	33
<b>4. Resultados y análisis de resultados</b>	<b>35</b>
4.1. Resultados. Parte 1 . . . . .	35
4.2. Resultados. Parte 2 . . . . .	42
<b>5. Conclusiones</b>	<b>47</b>
5.1. Trabajos a futuro . . . . .	48
<b>Apéndice</b>	<b>49</b>
<b>A. Figuras complementarias</b>	<b>51</b>
<b>B. Otras distribuciones</b>	<b>55</b>
<b>Bibliografía</b>	<b>57</b>

# Aspectos preliminares

## Planteamiento del Problema

En esta Tesis se realizará una caracterización de los eventos extremos en secuencias de números pseudo-aleatorios producidas por el mapeo de Lévy con las herramientas estandarizadas tal como la definición de umbrales que se utilizan para catalogar eventos en series de tiempo. También en este trabajo identificamos el papel que juega el parámetro que describe la distribución en las propiedades de los eventos extremos en nuestras series de tiempo que producimos.

## Justificación

Los eventos extremos aparecen tanto en sistemas naturales como artificiales. Entre los ejemplos más representativos de eventos extremos podemos mencionar: sismos de gran magnitud, inundaciones, sequías, caídas en las bolsas de valores, ataques epilépticos, ataques cardíacos, etc. La caracterización, entendimiento y posible predicción de eventos extremos tiene un interés inherente a la protección civil, estabilidad económica y salud pública. En esta Tesis planteamos el estudio y caracterización de eventos extremos en series de tiempo artificiales producidas por el mapeo de Lévy. Esperamos que el entendimiento de nuestras series de tiempo artificiales nos permita ampliar el conocimiento y entendimiento de eventos extremos realistas (en particular en los sistemas financieros como las criptomonedas). En este sentido, las series de tiempo producidas por el mapeo de Lévy pueden servir como un modelo de referencia (modelo nulo) de series de tiempo realistas.

## Hipótesis

Se espera que las propiedades de los eventos extremos (periodicidad, amplitud, correlaciones, etc.) en series de tiempo producidas por el mapeo de Lévy muestren una dependencia importante del valor que caracteriza la distribución.

## **Objetivos**

### **Objetivo general**

El objetivo general de este trabajo consiste en caracterizar eventos extremos en series de tiempo producidas por el mapeo de Lévy.

### **Objetivos particulares**

1. Implementar el mapeo de Lévy y caracterizar los números pseudo-aleatorios generados de acuerdo a su distribución de probabilidad.
2. Construir series de tiempo (discreto) a partir de las secuencias de los números pseudo-aleatorios obtenidas del mapeo de Lévy.
3. Analizar las series de tiempo del mapeo de Lévy con las herramientas que se utilizan para caracterizar eventos extremos en series de tiempo.

# Introducción

En el universo existen fenómenos que observamos durante nuestra vida, los cuales son registrados para tener una comprensión más amplia de lo que sucede en éste. Dentro de los fenómenos hay eventos que se salen de los estándares de los registros, a estos acontecimientos se les conoce como eventos extremos. Los eventos extremos aparecen en diferentes campos tales como en la economía, salud, medio ambiente. El impacto que éstos tienen a menudo causan consecuencias catastróficas, es por eso que es imprescindible tener una comprensión amplia de la ocurrencia y una estadística ya que, con una posible predicción de éstos, podemos tomar ciertas precauciones o decisiones [1].

Como mencionamos al principio los eventos extremos los podemos encontrar en diferentes áreas y estos pueden ser como:

- **Medio ambiente.** En este campo nos encontraremos con eventos extremos tales como las inundaciones, tormentas, sequías, etc. Estos cambios climáticos nos ha llevado a una estadística especial de eventos extremos con un enfoque más preciso [2]. Entre estos eventos naturales encontramos terremotos de gran magnitud.
- **Economía.** Los sistemas en esta área son artificiales ya que fueron generados por el hombre, tal como la compra y venta de acciones, en los mercados volátiles tenemos caída en la bolsa de valores [1, 3, 4].
- **Salud.** Aquí nos encontramos con sistemas muy bien comportados, esto se refiere a que en principio la mayoría de ellos son periódicos, y cuando aparecen eventos extremos surgirán irregularidades, entre los eventos extremos que podemos mencionar están los ataques cardíacos, ataques epilépticos, presiones altas etc. [5].

Una motivación para el estudio de estos eventos radica en averiguar cómo se ve afectada la dinámica de un sistema “complejo”, cuando tal sistema sufre un evento extremo.

Actualmente se sabe que, para muchas series, sus eventos extremos se encuentran en las “colas” de las distribuciones y tienen un decaimiento lento semejante a una ley de potencia, asociando así que estos sistemas tienen una distribución de tipo Lévy.

Este tipo de eventos emergentes que no siguen ninguna regla aparente, tienen a su vez una gran influencia en el comportamiento total de un modelo [6, 1].

Ahora bien, si quisiéramos modelar de mejor manera tales series de tiempo con una distribución tipo Lévy, tendríamos que encontrar el método para explicar la distribución con la cual tales eventos ocurren. Nosotros proponemos basarnos en la distribución tipo Lévy y generar series de tiempo las cuales se distribuirán de esta manera. Aunque otra respuesta tentativa sería la Teoría de Valor Extremo (TVE). La TVE también ha sido una herramienta en problemas financieros, pues los mercados financieros se han caracterizado por una alta volatilidad y una inestabilidad significativa.

Los eventos extremos provocan un gran desconcierto y una de estas razones podría deberse a que tenemos una estadística “pobre”, es decir, tenemos limitada información en algunos sistemas, lo que representa un problema debido a que no se pueden dar buenos pronósticos o proyecciones. Es por eso que en este trabajo generamos las series de tiempo y nuestra base de datos sería muy “rica” debido al modelo que proponemos utilizar.

Para lograr los objetivos de este trabajo, los capítulos se distribuyen de la siguiente manera:

En el capítulo 1 presentamos los sistemas dinámicos lineales y no lineales de forma muy breve, se enfatiza en la dinámica caótica y se particulariza en una forma de atacar este tipo de problemas por medio de los mapeos, los cuales se ejemplificarán con algunos casos conocidos, los cuales son representativos para este trabajo y también presentaremos el *mapeo de Lévy*, el cual utilizamos para esta tesis.

En el capítulo 2 explicamos qué es una serie de tiempo, introducimos algunos conceptos que caracterizan las series de tiempo, además se mostrarán series de tiempo realistas enfatizando en las series dentro del campo de la economía, mostramos las series generadas por el mapeo de Lévy y presentamos nuestros datos realistas que utilizamos.

El capítulo 3 trata sobre los eventos extremos, hablamos de la TVE, describimos el desarrollo de la metodología ocupada para la caracterización de los dichos eventos y por último realizamos un tratamiento a nuestras series realistas y aplicando el mismo método.

El capítulo 4 se destina a presentar los resultados obtenidos al aplicar umbrales a nuestras series producidas. Se obtuvieron las respectivas curvas de ajustes para los diferentes intervalos de ocurrencia. Después se trasladó esta misma metodología a las series de retornos.

El último capítulo finaliza con las conclusiones de la investigación.



# Capítulo 1

## Mapeos

Los mapeos son una forma de solucionar un sistema dinámico caótico por medio de ecuaciones que describen el comportamiento de algunas variables, adicionalmente se visualizan los resultados que se generan al iterar  $n$ -veces el mapeo utilizando un conjunto de valores iniciales. En este capítulo se presentan algunos conceptos generales que sirven para entender el trabajo que se realizó y el mapeo que se utilizó.

### 1.1. Sistemas Dinámicos

Un sistema dinámico es aquel que evoluciona a través del tiempo. Estos sistemas se describen por medio de ecuaciones lineales o ecuaciones no lineales.

Las teorías lineales están conformadas por muchos teoremas útiles, como por ejemplo, la existencia de una solución general, o la superposición de las ecuaciones lineales. Pero hay sistemas que no son lineales, y esto ha sido un foco de interés para muchos científicos, en particular para el matemático francés Henri Poincaré quien fué uno de los que comenzó a estudiar este tipo de dinámica. Una motivación de Poincaré fue proporcionada por el problema de las órbitas de tres cuerpos celestes que experimentan atracción gravitacional mútua.

Para el estudio de los sistemas no lineales se formulan ecuaciones no lineales que no cuentan con la mayoría de las herramientas que proporcionan las teorías lineales tales como la superposición. El caos es una situación que ocurre en muchos sistemas no lineales, en donde su evolución es irregular e impredecible, pero no estocástica.

La forma en la que se pueden analizar estos problemas es por medio de ecuaciones de movimiento no lineales en formas diferentes, entre ellas se encuentran las ecuaciones diferenciales parciales (EDP), las ecuaciones diferenciales ordinarias (EDO), ecuaciones en diferencias, mapeos de iteración, etc. [7, 8].

En particular los mapeos son modelos que pueden capturar la complejidad de los sistemas Hamiltonianos desde la integrabilidad<sup>1</sup> hasta los movimientos caóticos.

---

<sup>1</sup>Se dice que un sistema es integrable si tiene un comportamiento regular o predecible [9].

## 1.2. Caos

En el campo de la ciencia el caos no se debe interpretar como cuestión azarosa o totalmente aleatoria, sino que **el caos es determinista**, lo que implica que al referirnos a la dinámica de un sistema cuyo comportamiento es aperiódico, las reglas que lo definen no contienen componentes aleatorias o ruido. El determinismo comprende que el estado actual del sistema depende de sus estados previos [8, 10, 7].

Aunque el comportamiento dinámico producido por el ruido y las oscilaciones o fluctuaciones que suceden al azar, pueden aparentar ser caóticas éstas no lo son. Recordemos que el caos es de características deterministas porque está confinado a una región de un espacio de fases. En cambio los fenómenos estocásticos aparecen en cualquier región del espacio de fases [11].

En el libro de Richard L. Ingraham [7] se menciona que “el caos cumple con cualquiera o todas las propiedades siguientes: dependencia sensible de las condiciones iniciales, espectros de potencia de banda ancha, correlaciones en descomposición o imprevisibilidad de las órbitas, medida por la complejidad algorítmica positiva o entropías de varios tipos”. Las propiedades que se comentaron anteriormente no son todas independientes, pero algunas de ellas pueden ser difíciles de cuantificar en los sistemas prácticos.

De todas las propiedades anteriores, la de mayor importancia es la dependencia sensible a las condiciones iniciales. La dependencia sensible a las condiciones iniciales quiere decir que aún con condiciones iniciales infinitesimalmente cercanas éstas se comportarán de manera diferente al evolucionar el sistema y la tasa de crecimiento de la distancia entre las trayectorias u órbitas de estos puntos se puede cuantificar mediante el exponente de Lyapunov<sup>2</sup>  $\lambda$  [11].

El exponente de Lyapunov nos da una idea para determinar cómo evolucionan dos órbitas, tomando en cuenta un solo exponente se tiene que para  $\lambda < 0$  las trayectorias tienden a aproximarse (convergen); si  $\lambda = 0$  la distancia entre ellas se mantiene paralelas; y si  $\lambda > 0$  estas distancias divergen. También puede existir la situación de tener más de un exponentes, y ligado a esto se cuenta con la entropía de Kolmogorov, definida como la suma de los exponentes de Lyapunov [12].

Todo lo anterior se puede manifestar en la impredecibilidad de las órbitas, que es cuando cualquier conjunto de condiciones iniciales se extiende más de un diámetro específico que representa la precisión de predicción de interés. Cuando esto sucede, el sistema es impredecible en el sentido de que la predicción basada en cualquier conjunto de condiciones iniciales es tan imprecisa que es imposible determinar el resultado del sistema con la precisión deseada. Esto no quiere decir que si sabemos exactamente el estado de un sistema no podamos predecir donde estará en un momento determinado en el futuro, ya que ese tipo de predicciones se hacen constantemente con modelos por computadora.

---

<sup>2</sup>El exponente lleva el nombre de Aleksandr Lyapunov.

## 1.3. El Mapeo estándar

### 1.3.1. Rotor pateado

Un ejemplo para presentar a los sistemas dinámicos caóticos es el problema del rotor pateado, el cual también usaremos para deducir el “mapeo estándar”<sup>3</sup>. Este tipo de sistema fue estudiado por Chirikov [13].

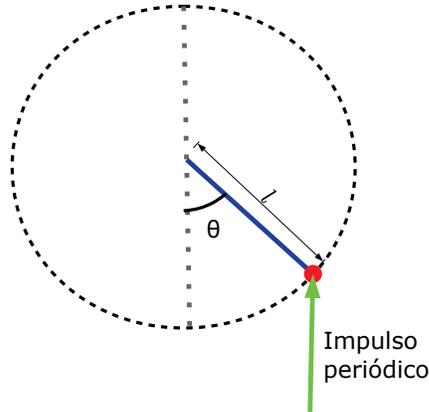


Figura 1.1: Rotor pateado

La idea general de este problema trata de una partícula que está sujeta a un movimiento circular cuyo momento de inercia es  $I$  y cuya longitud es  $l$ , sin fricción y sin que actúen los efectos de la gravedad como se muestra en la figura 1.1. Sobre la partícula actúa un impulso periódico con una fuerza de  $K/l$  que se acciona por un instante de tiempo con la misma dirección e intensidad, y ésta se puede aproximar a una función delta de Dirac (con periodo  $T$ ). Esta fuerza provoca un cambio instantáneo en el valor del impulso angular proporcional a la componente tangencial de dicha fuerza [8, 11, 14].

El Hamiltoniano del sistema se puede escribir de la siguiente forma:

$$H(P_\theta, \theta, t) = \frac{P_\theta^2}{2I} + K \cos \theta \sum \delta(t - nT), \quad (1.1)$$

$$\frac{d\theta}{dt} = \frac{P_\theta}{I}, \quad (1.2a)$$

$$\frac{dP_\theta}{dt} = K \sin \theta \sum \delta(t - nT), \quad (1.2b)$$

donde  $\delta$  es la función delta de Dirac y  $P_\theta$  es el momento angular. Ahora  $P_n$  y  $\theta_n$  denotarán  $P_\theta$  y  $\theta$  para el tiempo  $t = nT$ ,  $\theta_n$  es el ángulo formado con la vertical en el instante que recibe la  $n$ -ésima patada y  $P_n$  el momento generado por esta patada.

<sup>3</sup>También se le conoce como mapeo de Chirikov.

Este sistema Hamiltoniano es unidimensional, pero no es independiente ni integrable; sin embargo, lo podemos resolver porque, entre cada una de las patadas, la velocidad angular permanece constante, lo que involucra que entre los impulsos  $n$  y  $n + 1$  el momento angular es  $P_n$  [9, 8]. Al integrar la ec. (1.2a) entre  $n$  y  $n + 1$  se tiene

$$\theta_{n+1} - \theta_n = \frac{P_n T}{I}, \quad (1.3a)$$

y realizando el mismo procedimiento a la ec.(1.2b) obtenemos

$$P_{n+1} - P_n = K \text{sen } \theta_{n+1}. \quad (1.3b)$$

De las ec. (1.3) obtenemos un mapeo el cual depende de su valor anterior, para poder realizar su análisis por medio de sus iteraciones se deben introducir sus valores iniciales, es decir,  $\theta_0$  y  $P_0$ .

### 1.3.2. Mapeo estándar

De la expresión anterior (1.3), sin perder la generalidad, podemos tomar el período como unidad de tiempo y el momento de inercia también como unidad ( $T = 1$ ,  $I = 1$ ). Lo que reduce al mapeo que dependa de un solo parámetro  $K$  que mide la intensidad de la perturbación aplicada al sistema como se muestra a continuación

$$\begin{aligned} \theta_{n+1} &= \theta_n + P_n, & \text{módulo } 2\pi \\ P_{n+1} &= P_n + K \text{sen } \theta_{n+1}. \end{aligned} \quad (1.4)$$

Por último al mapeo (1.4) le agregamos un módulo  $2\pi$  ya que  $\theta$  es un ángulo, y deseamos limitar su valor para que se encuentre entre cero y  $2\pi$ , este mapeo se conoce como “mapeo estándar”.

Para el mapeo (1.4) podemos observar que con el parámetro  $K = 0$ , el sistema se reduce a un péndulo simple sin gravedad, por lo que es integrable y sus órbitas son periódicas o cuasi-periódicas<sup>4</sup> como se observa en la figura 1.2 (a), y al crecer el valor de  $K$  existen regiones donde el espacio fase contiene islas de movimiento periódico rodeadas por regiones de caos limitadas por curvas invariantes, a este caso se le conoce como espacio fase mixto, tal como se muestra en la figura 1.2 (b) y (c), y por último, el caso totalmente caótico está en la figura 1.2 (d).

---

<sup>4</sup>Que tiene un comportamiento aperiódico pero muestra tener un patrón de recurrencia con una componente de imprevisibilidad a una posición cercana de un estado previo del sistema. “Suele aparecer cuando interactúan dos osciladores de frecuencias diferentes sin que se de acoplamiento ni caos” [12].

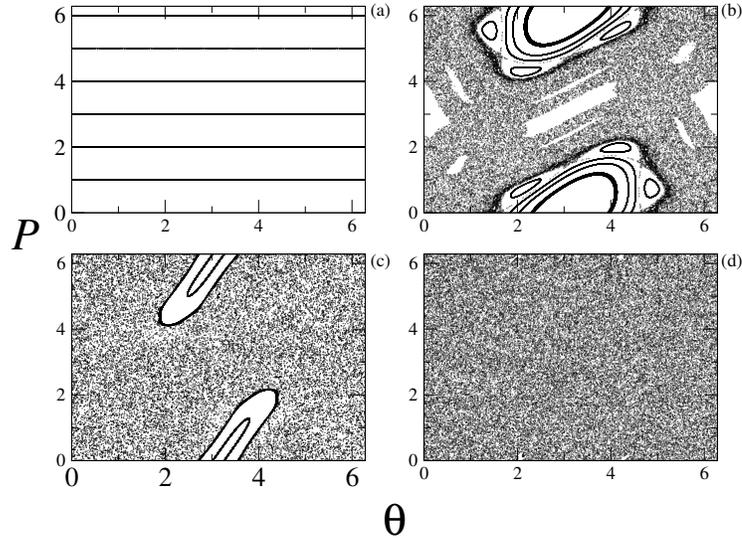


Figura 1.2: Mapeo estándar con (a)  $K = 0$ , (b)  $K = 1.5$ , (c)  $K = 4$  y (d)  $K = 7$ , con el momento angular modulado por  $2\pi$ .

El nombre de mapeo estándar se le dió porque muchos sistemas pueden ser reducidos a este mapeo. Por lo que demuestra ser un modelo muy conveniente para el estudio del comportamiento caótico típico como por ejemplo, la dinámica del cometa Halley [15], y para confinamiento de partículas cargadas en trampas magnéticas de espejo [13, 16, 17]

## 1.4. El mapeo de Poincaré

### 1.4.1. Billar cosenoidal o billar ondulado

El mapeo de Poincaré puede describir la dinámica de un billar cosenoidal. Entonces es conveniente explicar el concepto de un modelo de tipo billar ejemplificándolo por medio del juego de “billar” donde el “taco” golpea a una de las bolas dentro de la mesa, lo que provoca que entren en carambola y éstas a su vez golpean las paredes, e incluso algunas bolas caen en las buchacas. Entonces la representación de éstos sistemas tipo billar sería similar, pero las bolas jugarán el papel de partículas puntuales, sin la existencia de buchacas, y adicionalmente la geometría de la región puede variar. Por lo tanto, un billar describe la dinámica de una partícula moviéndose en una región cuyos choques con los bordes que la contienen se consideran colisiones elásticas. La partícula se mueve sin la influencia de ninguna fuerza externa o interna, en una región abierta o acotada [18]. Los billares han servido como modelos para el estudio de las características de las dinámicas generales de los sistemas deterministas.

El billar cosenoidal es el que ocuparemos para derivar el “*mapeo de Poincaré*”, este sistema dinámico lo describiremos por una partícula puntual moviéndose en una región confinada en un billar cuya geometría es el de dos paredes infinitas o moduladas, una de las paredes sería plana es decir  $y = 0$  y la otra pared tendrá un

perfil de una función coseno es decir  $y = d + w \cos x$ , donde  $d$  es el ancho del billar y  $w$  es la amplitud que tendrá la pared deformada, como se observa en la figura 1.3.

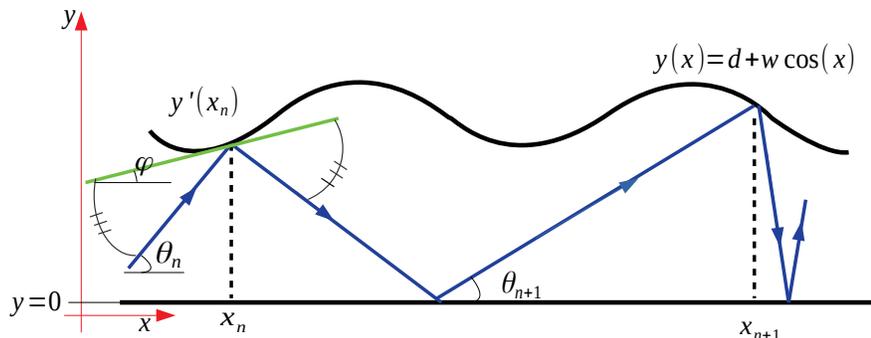


Figura 1.3: Geometría del billar cosenoidal u ondulado

Para describir el sistema nos interesamos en la posición ( $x_n$ ) de la partícula cuando toca con la pared superior, seguida por el ángulo que se forma al tocar la pared inferior ( $\theta_n$ ). Las ecuaciones para las variables de interés son deducidas en [10] y están dadas por

$$\theta_{n+1} = \theta_n - 2\varphi, \quad (1.5a)$$

$$x_{n+1} = x_n + (2d + w(\cos x_{n+1} + \cos x_n)) \cot \theta_{n+1}, \quad (1.5b)$$

donde  $\varphi$  es el ángulo de la función con respecto a la horizontal en el punto que la partícula choca con la pared ondulada, entonces  $\varphi = \arctan(y'(x))$ . Para el análisis anterior no se tomaron en cuenta las colisiones múltiples con la pared ondulada del billar. Con colisiones múltiples queremos decir que una vez que la partícula ha colisionado con la pared ondulada lo haga nuevamente, pero sin chocar con la pared plana [10].

### 1.4.2. Mapeo de Poincaré

Para reducir la expresión anterior realizaremos dos consideraciones, la primera es que  $w \ll d$ , lo que nos lleva a que podemos despreciar  $w(\cos x_{n+1} + \cos x_n)$  dejando a la ecuación (1.5b) de la siguiente manera:

$$x_{n+1} = x_n + 2d \cot(\theta_{n+1}),$$

y como segunda consideración es que el producto de  $w \sin x_n \ll 1$  esto implica que al ser lo suficientemente pequeño  $\varphi = \arctan w \sin x_n \approx w \sin x_n$  obteniendo la siguiente forma para el mapeo

$$M : \begin{cases} \theta_{n+1} = \theta_n - 2w \sin x_n & \text{módulo } \pi \\ x_{n+1} = x_n + 2d \cot(\theta_{n+1}) & \text{módulo } 2\pi \end{cases} \quad (1.6)$$

Observamos que el mapeo (1.6) depende de dos parámetros  $w$  y  $d$ , además agregamos un módulo  $\pi$  al ángulo porque éste se encuentra entre  $0$  y  $\pi$  ya que es la máxima apertura que puede tener con la pared inferior. Se pueden estudiar diferentes casos de la dinámica de este billar, para producir la figura 1.4, como ejemplo dejamos fijo el ancho de las paredes y variamos la amplitud de la función, es habitual que la variable  $x$  se module en un intervalo  $(-\pi, \pi)$

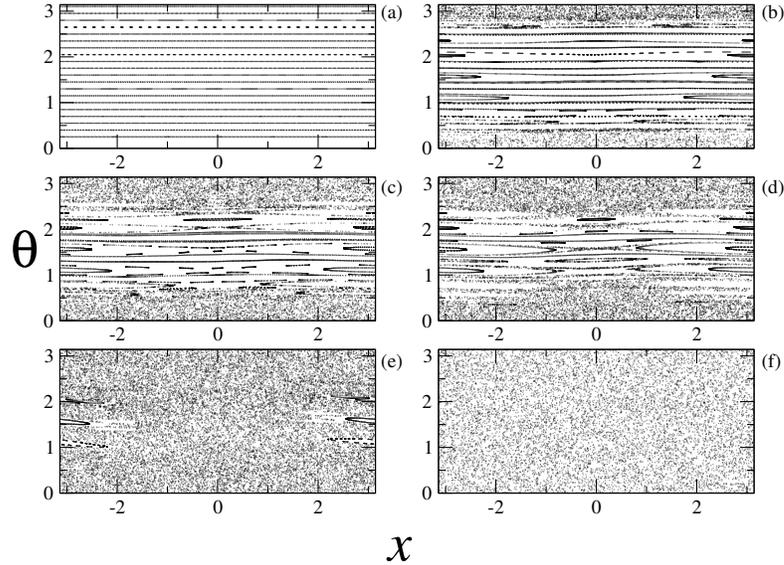


Figura 1.4: Mapeos de Poincaré con el parámetro  $d = 2\pi$ , para  $x_0 = 1$  y (a)-(e)  $\theta_0$  entre  $0.3$  hasta  $3.1$  con  $10^4$  iteraciones, (a)  $w = 0$ , (b)  $w = 0.01$ , (c)  $w = 0.02$ , (d)  $w = 0.03$ , (e)  $w = 0.1$ , (f)  $w = 2\pi/10$  con  $\theta_0 = 0.1$

## 1.5. El mapeo de Lévy

### 1.5.1. Densidad de distribución tipo Lévy

La distribución tipo Lévy pertenece a la familia de las distribuciones  $\alpha$ -estable definidas en base a la transformada de Fourier. La función de distribución tipo Lévy no admite una fórmula analítica cerrada (expresada en términos de funciones elementales), sin embargo, las colas de la distribución, sí poseen un función característica cuya forma es una ley de potencia [19]

$$P(l) \sim \frac{1}{l^{\alpha+1}}. \tag{1.7}$$

Esta densidad de distribución no tiene varianza definida para  $\alpha < 2$  ni tampoco media definida para  $\alpha \leq 1$  ya que estas divergen.

Existen muchos fenómenos que presentan un decaimiento lento en las colas y pueden ser descritos por la ec. (1.7); en el trabajo de Méndez-Bermúdez, et al. [20]

se particulariza en el modelo del billar cosenoidal, el cual explica en la sección 1.4 el mapeo de Poincaré, pero aunque la dinámica de este sistema es descrita bastante bien, hay información que se pierde al aplicar el módulo de  $2\pi$  tal como la longitud entre dos colisiones sucesivas  $l \equiv x_{n+1} - x_n$ .

Considerando la dinámica del mapeo (1.6) en el régimen caótico, entonces una sola trayectoria puede explorar el espacio de fase disponible de forma homogénea. En la figura 1.5 presentamos la distribución de las fases  $P(l)$ , en (b), donde la mayor parte de la concentración de la información se encuentra en las longitudes cercanas, existe una probabilidad no despreciable en las colas para  $l \approx \pm 1000$  y para  $P(\theta)$  se tiene que la distribución es constante e igual a  $1/\pi$ , entonces de la segunda ecuación en el mapeo (1.6) obtenemos  $\theta = \arctan(2d/l)$ . Utilizando  $P(l) = P(\theta)|d\theta/dl|$ , encontramos una ecuación de la distribución para  $P(l)$ :

$$P(l) = \frac{2d}{\pi(1+l^2)}. \quad (1.8)$$

El resultado obtenido es análogo a tener  $\alpha = 1$  en la ec. (1.7). En la figura 1.5 se compara el histograma que se generó del mapeo (1.6) con la ec. (1.8), demostrando su correspondencia y apreciando el decaimiento de las colas en forma de leyes de potencias con  $\alpha = 1$ . Dado que una trayectoria caótica típica llena uniformemente el espacio de fase disponible como se ve en la figura 1.4 (f). Un rasgo interesante de  $\theta$  y  $l$  es que para ángulos cercanos a  $\pi/2$  la distancia recorrida por la partícula es pequeña, mientras que para ángulos cercanos a cero o  $\pi$  se tiene que  $l \rightarrow \infty$ . La distribución de probabilidad tipo Lévy de la ec. (1.8) para las longitudes  $l$  en el billar cosenoidal es la prueba de la existencia de vuelos de Lévy de los cuales se sabe que producen funciones de distribución con colas que decaen lentamente.

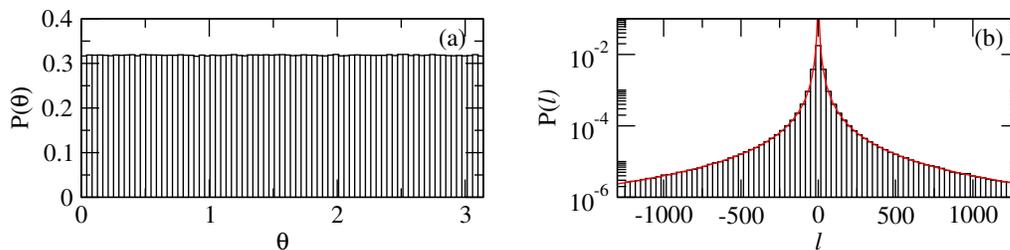


Figura 1.5: Distribuciones de las fases del mapeo de Poincaré (a)  $P(\theta)$ , (b)  $P(l)$ .

### 1.5.2. Mapeo de Lévy

El mapeo de Lévy se deriva de emplear el mapeo (1.6) en el régimen del caos global entonces bajo ese argumento tenemos las siguientes consecuencias

(i)  $P(\theta)$  constante

$$P(\theta) = \frac{1}{\pi}, \quad (1.9)$$

(ii) introducimos la variable  $l$

$$l \equiv x_{n+1} - x_n \quad (1.10)$$

(iii) Los números  $l$  están caracterizados por una distribución tipo Lévy

$$P(l) = \frac{C}{l^{1+\alpha}}, \quad (1.11)$$

donde  $C$  es una constante de normalización. Usando el hecho de que  $P(l) = P(\theta)|d\theta/dl|$ , podemos obtener  $\theta$

$$\theta \equiv \int \frac{P(l)}{P(\theta)} dl = \pi C \int \frac{dl}{l^{\alpha+1}} = -\frac{\pi C}{\alpha l^\alpha}. \quad (1.12)$$

De la ec. (1.12) despejamos el valor de  $l = (-\alpha\theta/\pi C)^{-1/\alpha}$  y se sustituye en la segunda ec. (1.6), de tal manera que se obtiene el mapeo [20]

$$M_\alpha = \begin{cases} \theta_{n+1} = \theta_n - 2w \operatorname{sen} x_n, \\ x_{n+1} = x_n + \gamma |\alpha \theta_{n+1}|^{-1/\alpha}, \end{cases} \quad (1.13)$$

donde  $\gamma = (\pi C)^{1/\alpha}$  y  $w$  es un parámetro, ahora el mapeo depende explícitamente de  $\alpha$ . Por último, se introduce el valor absoluto en la segunda ecuación del mapeo el cual permite el cálculo de potencias fraccionarias de ángulos negativos y hace que todas las longitudes sean positivas.

Recuperamos el mapeo (1.6) de (1.13), si consideramos que  $\alpha = 1$ ,  $\theta_{n+1} \ll 1$ ,  $\cot \theta_{n+1} \approx 1/\theta_{n+1}$ , y  $\gamma = 2d$ . También se observa que tiene una forma similar a los mapeos estudiados en [21-24] en el sentido de que la función  $f(\theta_{n+1})$ , en la segunda línea del mapeo  $M_\alpha$ , es inversamente proporcional a  $\theta_{n+1}$  a una potencia racional. En general, dependiendo de los valores de los parámetros  $(w, \gamma, \alpha)$ , la dinámica del mapeo de Lévy puede ser integrable, mixta o totalmente caótica. Es decir, el espacio fase desarrolla la transición genérica al caos como función de  $(w, \gamma, \alpha)$ . En la figura 1.6 se muestra la transición desde el caso integrable al caos global para  $\alpha = 1$ . En el apéndice A se muestran los mapeos para los diferentes casos y valores de  $\alpha$ .

Pero, para que la ec. (1.9) sea válida, la dinámica debe de ser caótica. Se ha aplicado el criterio del traslape de resonancia de Chirikov para identificar las condiciones que deben cumplir los parámetros y éstos se encuentren en el caos [20], entonces se tiene que

$$w \gtrsim \frac{\gamma^\alpha}{2(2\pi)^{\alpha+1}} = \frac{C}{4(2\pi)^\alpha}. \quad (1.14)$$

Dado que este mapeo está propuesto para usarse en el régimen del caos global las variables que utilizaremos que cumplen este criterio son para los valores de  $w = C = 1$  que servirá para cualquier valor de  $\alpha$  como se muestra en la figura 1.7 de las

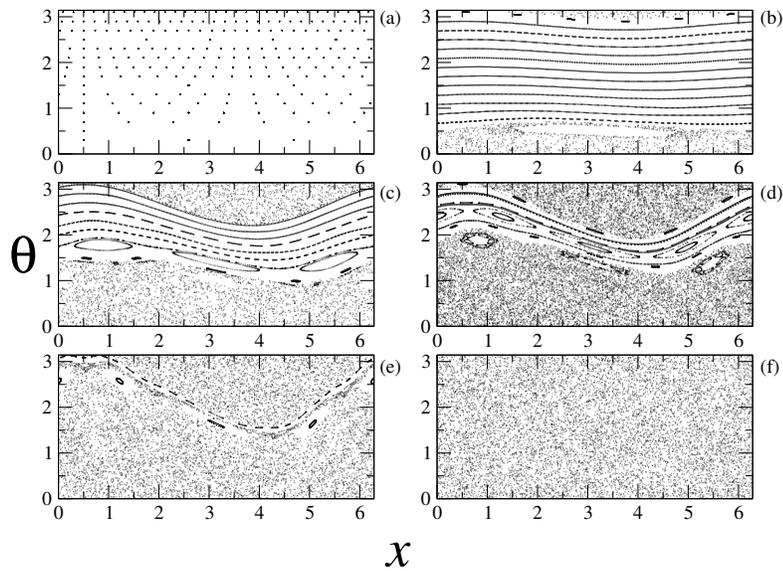


Figura 1.6: Mapeo de Lévy con  $\alpha = 1$ ,  $C = 1$ , iterado  $10^4$  veces y cuyas condiciones iniciales son  $x_0 = 3$  y  $\theta_0 = 0.3 - 3.1$  (a)-(e), (a)  $w = 0$ , (b)  $w = 0.05$ , (c)  $w = 0.25$ , (d)  $w = 0.35$ , (e)  $w = 0.5$  y (f)  $w = 1$  y  $\theta_0 = x_0 = 0.1$ .

distribuciones de fases, y en [20] se encontró que la función que mejor se ajusta es

$$P(\ln l) = l^\alpha. \quad (1.15)$$

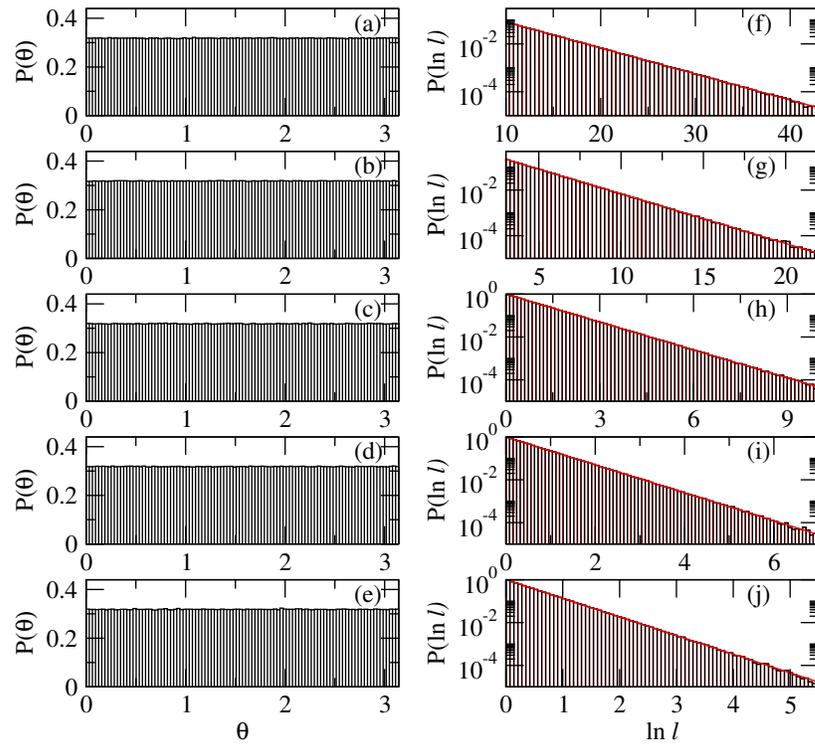


Figura 1.7: (a)-(e) distribución de fase  $P(\theta)$ , (f)-(j) fase  $P(\ln l)$ , (a) y (f)  $\alpha = 1/4$ , (b) y (g)  $\alpha = 1/2$ , (c) y (h)  $\alpha = 1$ , (d) e (i)  $\alpha = 3/2$ , (e) y (j)  $\alpha = 2$ . Con  $x_0 = \theta_0 = 0.1$  y  $w = c = 1$  para  $10^7$  iteraciones y la curva de regresión  $P(\ln l) = l^{-\alpha}$ .



# Capítulo 2

## Las series de tiempo

Las series de tiempo se analizan para comprender el pasado y en principio predecir el futuro. El análisis de series de tiempo cuantifica las características principales de los datos y las variaciones aleatorias. Es por eso que existen modelos de series de tiempo, y en este trabajo enfatizaremos en unas series financieras. A continuación, explicamos lo que son las series de tiempo, algunos conceptos que usamos, y presentamos nuestro modelo serie de tiempo.

### 2.1. Series de tiempo

En el libro de Robert A. Yaffee [25] se cita que una serie de tiempo se describe como “una secuencia de observaciones ordenadas a través del tiempo”, es decir, que si consideramos una serie de tiempo como una secuencia de variables aleatorias,  $x_1, x_2, x_3, \dots$ , donde  $x$  es la variable aleatoria y el subíndice indica el orden en que se registraron los datos en el tiempo  $t$ . La recopilación de esta información suele ser limitada, estos datos pueden ser registrados de manera continua o discreta. Las series de tiempo continuas se graban de forma instantánea y constante, mientras que en las discretas, los datos se registran en intervalos regulares de tiempo [26].

Es interesante notar que muchas series de tiempo conocidas ocurren en el campo de la economía, que se exponen continuamente a las cotizaciones diarias del mercado de valores o las cifras mensuales de desempleo, mientras que en las ciencias sociales, estas series son como las tasas de natalidad o las matrículas escolares, y en medicina, pueden ser mediciones como la presión arterial o registros de actividad eléctrica del corazón que se produce en cada latido cardíaco.

El principal propósito de las series de tiempo es ser analizadas mediante un modelo estadístico que proporcione algún pronóstico de eventos<sup>1</sup> futuros con un grado de incertidumbre. A partir de la información se puede contar con un modelo para una

---

<sup>1</sup>“Un evento es algo que sucede dentro de un espacio y tiempo limitados. Su ocurrencia puede surgir por casualidad o necesidad o por una combinación de ambos; Por causas naturales o provocadas por el hombre o una combinación de ambas [27] ”.

única variable (univariado) o contar con varias series relacionadas, en esta situación es mejor utilizar modelos de varias variables (multivariada).

Con el objetivo de proporcionar alguna estructura estadística para describir los datos que aparentemente fluctúan de manera aleatoria a lo largo del tiempo, mostraremos algunas propiedades y características que presentan las series. Además, existen dos enfoques para el análisis de series de tiempo, comúnmente identificados como el enfoque del dominio del tiempo y el enfoque del dominio de la frecuencia. En este trabajo usaremos el enfoque del dominio del tiempo, porque generalmente se produce por la presunción de que la correlación entre puntos contiguos en el tiempo se explica mejor el valor actual en términos de una dependencia de los valores pasados.

## 2.2. Tipos de series

### Continuas o discretas

En la sección anterior se mencionó que una serie de tiempo puede ser continua o discreta. Una serie continua se podría ejemplificar por medio del registro de una señal de audio (figura 2.1 (a)), mientras que una serie discreta podría ser originada de un muestreo de una señal continua o de un proceso cuyos eventos sean de origen discreto, tal como, el de un indicador económico (ver figura 2.1 (b)).

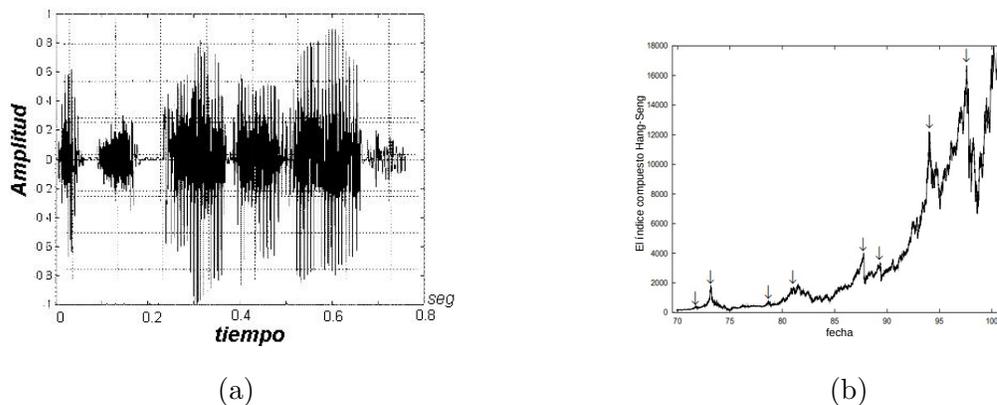


Figura 2.1: Serie de tiempo (a) continua (señal de voz [imagen tomada de Wikipedia]) y (b) discreta (el índice compuesto Hang-Seng de la bolsa de valores de Hong Kong [28]).

El caso particular del análisis de un modelado de series del habla es un problema importante porque está relacionado con la transmisión eficiente de grabaciones de voz. Las características comunes de una serie de tiempo conocida como el espectro de potencia se utilizan para ayudar a las computadoras a reconocer y traducir palabras [26].

Por otra parte, el análisis de la figura 2.1 (b) nos proporciona un criterio de la evolución del mercado de valores de China, en donde podemos observar que hay una tendencia de crecimiento en la serie. Los eventos que interesan predecir en este tipo de series son los aumentos o declives (pronunciados) que tienen.

### Multivariada o univariada

Otra forma en que se nos puede presentar la información es el caso de tener varias series relacionadas entre sí, tal como se muestra en la figura 2.2 (a), o una sola serie aislada, es decir, de una variable figura 2.2 (b).

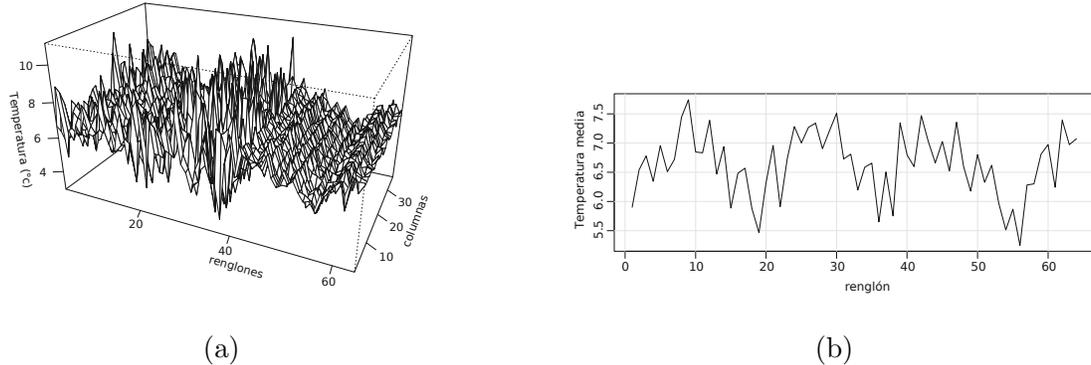


Figura 2.2: Serie (a) multivariada (temperatura de un campo) y (b) univariada (temperatura de un solo renglón del campo) [26].

El estudio de una serie multivariada también puede ser analizado por una sola variable como en la figura 2.2 (b), pero, aunque un modelo particular se ajuste a un par de renglones no significa que se pueda generalizar para todo el campo, y es por eso que se debe estudiar todo el conjunto de los valores y tener un modelo que ajustemos a la colección de datos que corresponden a un solo fenómeno.

### Estacionaria o no estacionaria

Otra característica que puede presentar una serie es que sea estacionaria o no. Cuando una serie es estacionaria<sup>2</sup> conserva sus características estadísticas a lo largo de un tiempo. Esto quiere decir que no es transitoria y tiene (teóricamente) duración y largo infinito, como por ejemplo las series periódicas y el ruido blanco [29]. Por otra parte, las series no estacionarias tienen una tendencia y/o variabilidad que cambian en el tiempo. Los cambios en su primer momento determinan una tendencia a crecer o decrecer, por lo que la serie no oscila alrededor de un valor constante.

<sup>2</sup>Una serie débilmente estacionaria,  $x_t$ , es un proceso de variación finita tal que la función de valor medio,  $\mu t$ , definida en (2.3) es constante y no depende del tiempo  $t$ , y la función de covarianza,  $cov(s, t)$ , definida en (2.4), depende de  $s$  y  $t$  sólo a través de su diferencia  $|s - t|$ .

Una serie estrictamente estacionaria es aquella para la cual el comportamiento probabilístico de una colección de valores  $x_1, x_2, \dots, x_k$ , es idéntico al del conjunto de valores desplazados  $x_{1+h}, x_{2+h}, \dots, x_{k+h}$ , es decir,  $F(x_{1+h}, x_{2+h}, \dots, x_{k+h}) = F(x_1, x_2, \dots, x_k)$  para todos los puntos de tiempo  $t = 1, 2, \dots, k$ , y todo los cambios de tiempo  $h = 0, 1, 2, \dots$

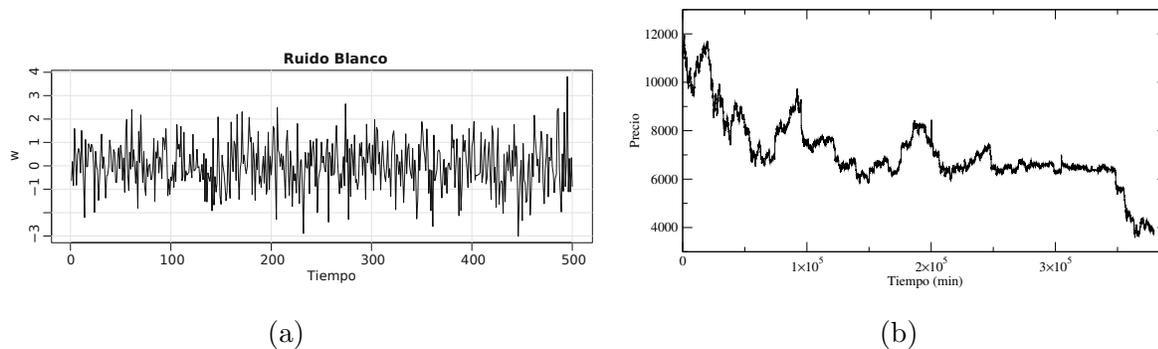


Figura 2.3: Serie de tiempo (a) estacionaria (ruido blanco) [26] y (b) no estacionaria (precio del bitcoin)

El ruido blanco, ver figura 2.3 (a) es un tipo simple de serie, con una colección de variables aleatorias no relacionadas,  $w_t$ , con media 0 y varianza finita  $\sigma_w^2$ . Muchas señales pueden contener ruido y es por eso que varias son tratadas haciendo uso de filtros y suavizadores (medias móviles), desarrollados originalmente para la estimación y el control en aplicaciones de transmisión o registro de información. Mientras que en la figura 2.3 (b) tenemos los precios de la criptomoneda del bitcoin aquí también encontramos fluctuaciones, pero estas no pertenecen a ruido ni errores en los registros.

Las series anteriores son solo ejemplos de bases de datos experimentales que sirven para ilustrar algunas características. En particular, las series que ocupamos serán discretas, de una sola variable y estacionarias.

## 2.3. Análisis de las series

Para usar alguna metodología estadística lo primero es estudiar los datos registrados a lo largo del tiempo. Existe un enfoque muy utilizado en el análisis de las series, consiste en suponer que los datos observados son el resultado de sumas de series cada una con una estructura, pero este enfoque no es indispensable para nosotros, debido a que estamos interesados en estudiar eventos ubicados en las colas de las distribuciones (eventos de mayor amplitud). A continuación, mostramos algunos elementos a considerar.

### 2.3.1. Función de distribución

Una descripción de una serie de tiempo es observada como una colección de  $n$  eventos en puntos de tiempo  $t_1, t_2, \dots, t_n$ , para cualquier entero positivo  $n$ , es proporcionada por la función de distribución que incorpora todas las funciones de distribución univariadas para cualquier subconjunto finito de variables.

$$F_t(x) = P\{x_t \leq x\}, \quad (2.1)$$

o por su correspondiente función de densidad

$$f_t(x) = \frac{\partial F_t(x)}{\partial x}. \quad (2.2)$$

Aunque se pueden identificar características que definen una serie de tiempo solo por su función de distribución o su función de densidad, también podemos usar otra descripción cuantitativa para obtener más información (como sus dos primeros momentos).

**Definición 2.1. Primer momento**

El primer momento alrededor de cero se conoce como *media*:

$$\mu_{x_t} = E(x_t) = \int_{-\infty}^{\infty} x f_t(x) dx, \quad (2.3)$$

donde  $E$  denota el operador de valor esperado habitual. Cuando no exista confusión sobre a qué serie de tiempo nos referimos, escribiremos  $\mu_{x_t}$  como  $\mu_t$ .

**Definición 2.2. Segundo momento**

El segundo momento alrededor de la media para todos  $s$  y  $t$  es la función de *covarianza*:

$$\text{cov}(x_s, x_t) = E[(x_s - \mu_s)(x_t - \mu_t)]. \quad (2.4)$$

La covarianza mide la dependencia lineal entre dos puntos de la misma serie observada en diferentes momentos. Si  $s = t$ , la covarianza se reduce a la *varianza*:

$$E[(x_t - \mu_t)^2] = \text{var}(x_t). \quad (2.5)$$

### 2.3.2. Gráficas

Las gráficas de las series de tiempo nos proporcionan la primera entrada de información, por lo que es importante elegir las escalas, el tamaño del intervalo y la forma en que se trazan los puntos. Para series que muestran grandes variaciones en la amplitud de sus eventos, una escala logarítmica también permite mostrar con más detalles su amplio rango (por ejemplo, al medir el sonido, los ingenieros usan decibelios que se miden en una escala logarítmica). Por ejemplo, “Jenkins y Watts [30] sugieren que las estimaciones del espectro siempre deben representarse en una escala logarítmica”.

El logaritmo natural (base  $e$ ) se puede usar para transformar la amplitud de los eventos  $x_t$  como

$$y_t = \ln(x_t), \quad (2.6)$$

esta transformación se pueden tomar sólo si la serie  $x_t$  toma todos los valores positivos; es decir,  $x_t > 0$  para todos  $t$  [31].

## 2.4. Series producidas por el mapeo de Lévy

En el capítulo anterior presentamos el mapeo de Lévy el cual puede ser usado como un generador de números pseudo-aleatorios, y esta secuencia la interpretamos como una serie de tiempo, donde la amplitud de los eventos serán  $l_t$  y estarán trazadas en el orden de las iteraciones que representan al tiempo  $t$ . Lo estacionario en nuestras series se ve reflejada en la figura 1.7 donde la distribución de los valores tienden a oscilar alrededor de un valor constante, pero nosotros no nos enfocamos en observaciones comunes sino en las observaciones atípicas que se encuentran en la cola de la distribución.

Al estar caracterizado este modelo por una cola que decae lentamente, se tendrán eventos con amplitudes grandes que podemos interpretar como eventos extremos. Debemos enfatizar que entre más pequeño es el valor de  $\alpha$ , más grande es la amplitud de los eventos, como se ve en la figura 2.4.

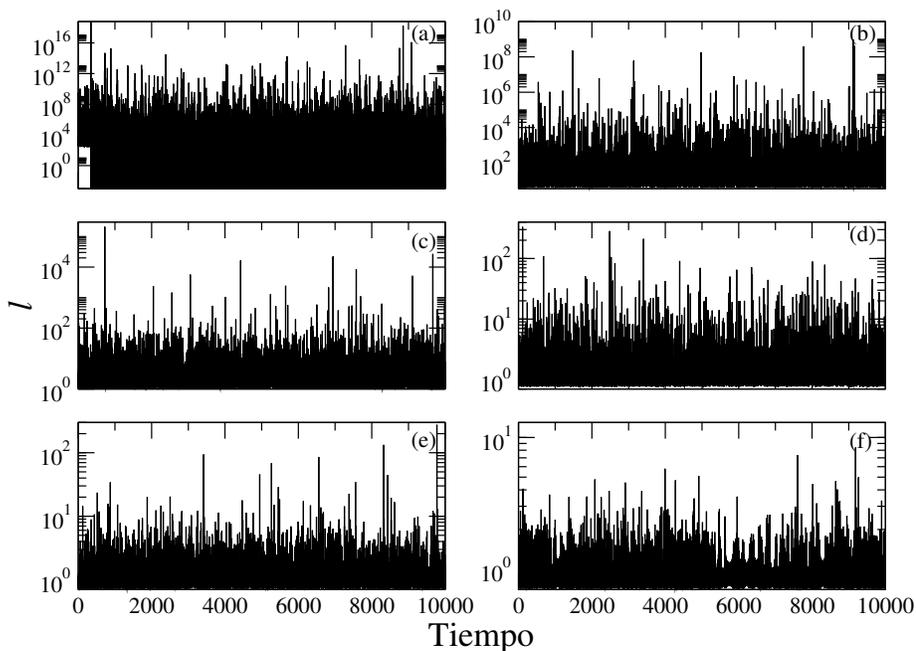


Figura 2.4: Serie de tiempo producida por el mapeo de Lévy, (a)  $\alpha = 1/4$ , (b)  $\alpha = 1/2$ , (c)  $\alpha = 1$ , (d)  $\alpha = 3/2$ , (e)  $\alpha = 2$ , (f)  $\alpha = 4$ .

Debido a las grandes amplitudes que presentan las series realizaremos el análisis a  $\ln l_t$ .

## 2.5. Series financieras

El registro de una secuencia de precios de productos o activos produce una serie de tiempo. El análisis de las series financieras ha sido de gran interés, no solo para los profesionales en esta disciplina, sino también para los teóricos que hacen inferencias

y predicciones. La incertidumbre inherente en las series temporales financieras y teóricas las hacen especialmente interesantes para economistas, estadísticos y físicos [1].

Hoy en día, todas las transacciones en los mercados financieros se registran paso a paso, es decir, cada evento es una acción que se registra con una marca de tiempo definida, lo que lleva a grandes cantidades de datos. Nosotros contamos con una base de datos de la cotización de 10 criptomonedas<sup>3</sup> (la décima moneda se muestra en la figura 2.3 (b)).

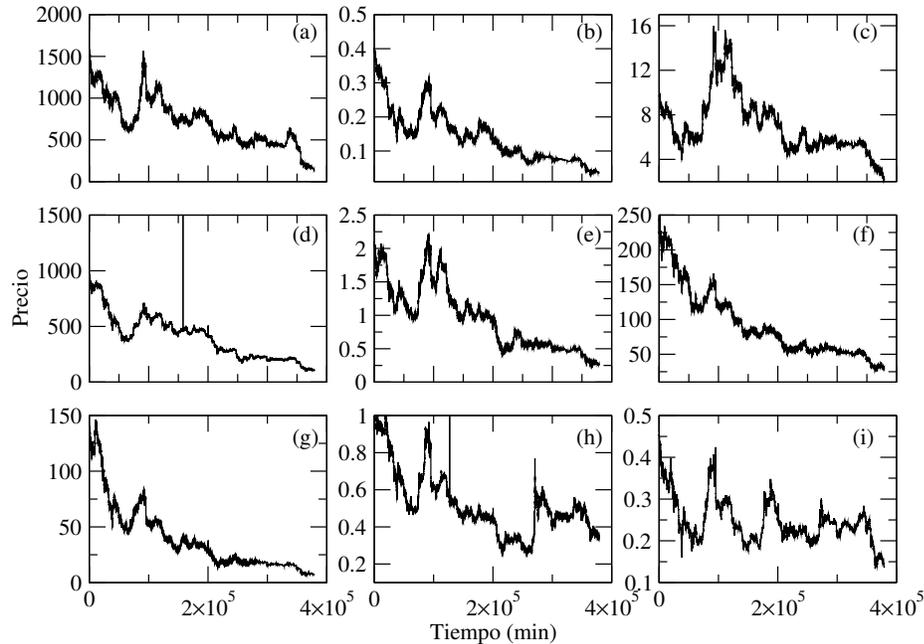


Figura 2.5: Precio de Criptomonedas (a) bitcoin Cash, (b) cardano, (c) eos, (d) ethereum, (e) iota, (f) litecoin, (g) neo, (h) ripple, (i) stellar.

En la figura 2.5 se tiene un conjunto de datos que están muestreados en intervalos de un minuto (aproximadamente) entre el 19 de febrero del 2018 y el 6 de diciembre del 2018, con un total de datos igual a 379417.

### 2.5.1. Criptomoneda

Una criptomoneda es el nombre que recibe una moneda digital que se considera segura debido a la criptografía<sup>4</sup>. Esta divisa alternativa no dependen de ningún banco, y funciona de la misma manera que el dinero convencional tal como para comprar y/o invertir. La primera criptomoneda que empezó a operar fue el bitcoin en 2009 y, desde entonces, han aparecido muchas otras, entre las de mayor crecimiento se encuentran el bitcoin, cardano, eos y ethereum.

<sup>3</sup>La base de datos fue proporcionada por el Dr. Andrés García Medina, investigador del CIMAT de Monterrey.

<sup>4</sup>Es un tipo particular de método de encriptación.

### 2.5.2. Retornos

La mayoría de los datos de las series financieras no son estacionarias, pero podemos obtener una aproximación a la estacionariedad tomando los retornos de las observaciones. Sea el precio de un activo financiero en el momento  $t$ , definimos su retorno logarítmico como [1, 32]

$$r_{t_{ln}} = \frac{p_t - p_{t-1}}{p_t} \approx \ln(p_t) - \ln(p_{t-1}) \quad (2.7)$$

mientras que a su retorno lineal es

$$r_{t_{lineal}} \approx p_t - p_{t-1} \quad (2.8)$$

Estos retornos son las diferencias de precios, cuando  $r_t < 0$  el precio disminuye, cuando  $r_t > 0$  el valor del activo aumenta y si  $r_t = 0$  no hay cambios en su valor. A continuación, son mostradas las graficas de retorno lineal y logarítmico de las criptomonedas:

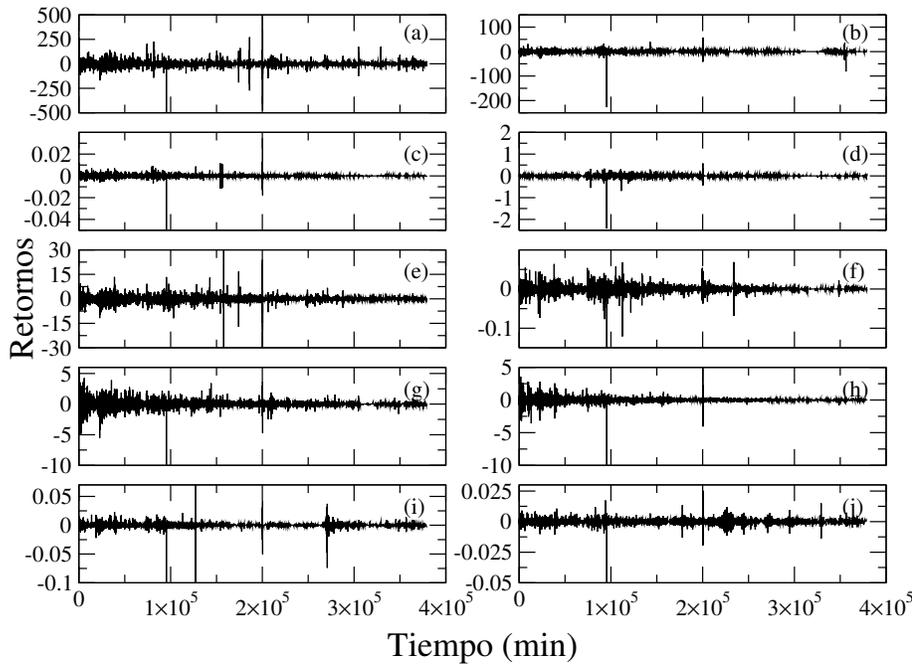


Figura 2.6: Retornos lineales de las criptomonedas (a) bitcoin, (b) bitcoin Cash, (c) cardano, (d) eos, (e) ethereum, (f) iota, (g) litecoin, (h) neo, (i) ripple, (j) stellar.

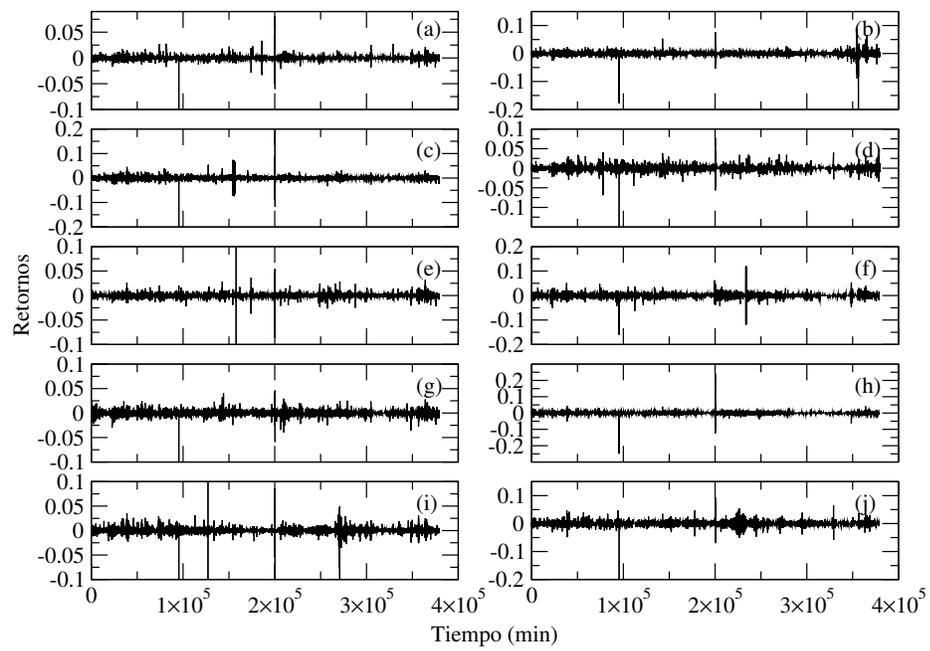


Figura 2.7: Retornos logarítmicos de las criptomonedas (a) bitcoin, (b) bitcoin Cash, (c) cardano, (d) eos, (e) ethereum, (f) iota, (g) litecoin, (h) neo, (i) ripple, (j) stellar.



# Capítulo 3

## Eventos extremos

Para el análisis de eventos extremos existen algunas teorías y métodos que establecen una base de cómo debe realizarse el estudio, en particular hemos usado el método de umbral sobre umbral para analizar las series de tiempo producidas por el mapeo de Lévy. Además, hemos integrado datos reales, con el fin de utilizar la misma metodología tanto en nuestro modelo de series como en el de las series financieras.

### 3.1. Eventos extremos

Para describir qué es un evento extremo recordaremos que el término “evento” se encuentra en el capítulo anterior, mientras que para la interpretación de “extremo” viene a ser más complicado debido a que abarca una colección de cualidades tales como: raros, excepcionales, catastróficos y sorprendentes. Al hablar de raro queremos decir que tiene baja probabilidad y al mencionar catastrófico nos referimos a la dimensión del impacto o cambio que éstas tienen [27].

El interés que se genera al estudiar los eventos extremos es lo impredecible que éstos son, debido a las enormes desviaciones de las mediciones que aparecen en las series de tiempo que poseen una naturaleza explosiva y con una aparente singularidad.

La primera caracterización que se le puede dar a un evento extremo es por los elementos más importantes que lo componen tales como: sus propiedades estadísticas y dinámicas, posibles similitudes y analogías observacionales. Desde una perspectiva estadística los eventos extremos ocurren en las colas de las distribuciones de probabilidad. Estas colas con decaimiento lento, también se conocen como colas “pesadas” o “gruesas”, por ejemplo, las leyes de potencia tales como en la ec. (1.7) para  $\alpha > 0$ . Las leyes de potencia caen mucho más lentamente que las distribuciones exponenciales o gaussianas (ver figuras del apéndice B), lo que indica una mayor probabilidad de ocurrencia.

Aunque la aparición de eventos extremos parece ser aleatorio, la realidad es que son manifestaciones de una dinámica compleja, generalmente lejos del equilibrio, don-

de la variabilidad del sistema y los efectos colectivos son dominantes. Por lo tanto, buscamos mecanismos dinámicos que permitan a un sistema realizar una excursión lejos de su estado normal. Se conocen varios de estos escenarios, entre ellos los conceptos denominados teoría de grandes desviaciones, caos determinista y turbulencia completamente desarrollada, por mencionar sólo algunos.

De todo lo que hemos comentado en este capítulo, hay una amplia literatura donde existen varios conceptos y teoremas que nos ofrecen cierta caracterización de los eventos extremos, pero en ésta se encuentran dificultades tales como la elección de parámetros y la incertidumbre que esto conlleva, como veremos más adelante. Dentro de esta problemática también está que no hay un concepto generalmente aceptado para cuantificar la precisión de un error de predicción [27].

## 3.2. Teoría de los valores extremos

En la sección anterior hemos comentado que los eventos extremos se encuentran en las colas de las distribuciones, en esta sección hablaremos del comportamiento de las colas de las distribuciones estables con diferentes índices, es decir, con diferentes valores de  $\alpha$ . Este tema normalmente se conoce como TVE y nos dice que el valor más grande o más pequeño de un conjunto de valores tomados de la misma distribución original tiende a una distribución asintótica que sólo depende de la cola de la distribución original.

La TVE intenta proporcionar una caracterización del comportamiento de la cola de todos los tipos de distribuciones de probabilidad, argumentando que el comportamiento de la cola sólo puede, en la práctica tener un pequeño número de formas posibles. Tal teoría nos sugiere que podemos identificar las probabilidades de eventos extremos “simplemente” mediante el uso del siguiente procedimiento:

- (i) Identificar el aparente tipo de comportamiento de las colas (mínimo y máximos) o la cola que muestra la variable en cuestión.
- (ii) Estimar los parámetros que luego caracterizan el comportamiento de la cola.
- (iii) Calcular la probabilidad de ocurrencia, pero, muchas veces en la práctica se acostumbra que estén asociadas a algún nivel de cuantiles<sup>1</sup> deseado.

El procedimiento que se acaba de mencionar en la práctica no es tan simple porque lo que realmente se desea hacer es la extrapolación en la cola, pero esto es todo un desafío, no porque sea difícil identificar posibles distribuciones que puedan ajustarse a los datos, sino porque el rango de respuestas, que se puede obtener de forma plausible, puede ser muy amplio, especialmente si queremos extrapolar a puntos más alejados en la cola y que existan pocas observaciones, si es que hay alguna.

---

<sup>1</sup>El cuantil de orden  $p$  de una distribución (con  $0 < p < 1$ ) es el valor de la variable  $x_p$  que marca un corte de modo que una proporción  $p$  de valores de la población es menor o igual que  $x_p$ . Por ejemplo, los cuantiles serían 0.25, 0.5 y 0.75; los deciles 0.1, 0.2, ...; los percentiles 0.01, 0.02 hasta el 1. También ésto se puede presentar en porcentaje [33].

### 3.2.1. Estimador de Hill

En la TVE se tiene el siguiente estimador para el parámetro  $\alpha$  de la cola (llamado el estimador de Hill)

$$\hat{\alpha}^{-1} = \frac{1}{n} \sum_{i=1}^n \log \left( \frac{x_i}{x_{i_{cola}+1}} \right), \quad (3.1)$$

donde  $n$  es el número de observaciones en la cola e  $i_{cola}+1$  es el índice de la observación donde se supone que comienza la cola. Sin embargo, como comentamos al principio, encontramos que tampoco indica que hay una regla definitiva sobre cómo determinar la cola, es decir, sobre dónde se considera que comenzó la cola. Existen sugerencias de como trazar el índice de la cola contra la elección de la cola  $i$ , y de aquí elegir una cola y, por lo tanto, el índice de la cola coincide con el lugar donde dicha trama parece ser más o menos horizontal.

El estimador de Hill (y otros enfoques equivalentes) asume que sabemos cuándo comienza la cola y también asume implícitamente que conocemos el valor distributivo exacto en ese momento [4, 34].

## 3.3. Desarrollo

### 3.3.1. Consideraciones

De la sección anterior sólo se comenta una forma de metodología y teoría, pero existen otros métodos (las cópulas, POT<sup>2</sup>, etc.).

Debido a que no hay una estructura única para determinar el inicio de las colas, se propone una metodología que considera la separación de eventos en varios intervalos de acuerdo a sus amplitudes (vea por ejemplo [4]).

El trazado de la cola de la distribución la daremos en escala log-log; éste ha sido durante mucho tiempo un método popular para la inspección visual del comportamiento de las leyes de potencia.

### 3.3.2. Umbral sobre umbral

La metodología que proponemos es semejante a la de cuantil sobre cuantil [4], y éstos son los umbrales que determinan un rango para los eventos de la series.

Los umbrales que nosotros proponemos son equidistantes, los cuales establecemos por medio de un límite superior y uno inferior y dividimos en  $n$  partes. Para la sección del límite inferior usamos el valor mínimo de  $\ln l$ , mientras que para el límite superior usamos la distribución  $P(\ln l)$ , y fijando el “mínimo” de la distribución que corresponde al valor máximo de la serie. Entonces, se decidió que el mínimo de la distribución fuera 5 ordenes de magnitud más pequeño que el máximo de la distribución y de ahí despejamos el valor de  $l$ , es decir,

<sup>2</sup>Por sus siglas en ingles (peaks over a threshold) picos sobre un umbral.

$$l_{max} \Rightarrow P(\ln l)_{min} = P(\ln l)_{max} \times 10^{-5}. \quad (3.2)$$

El último umbral, el cual corresponde al límite superior es el único umbral que queda abierto, en la Figura 3.1, a modo de ejemplo, se muestra una serie con 3 umbrales.

Una vez definidos los umbrales ( $q_n$ ), se generaron las series de tiempo que están producidas por el mapeo de Lévy cuya componente estocástica para estas series es una semilla aleatoria (se produce por el reloj del sistema cuando ejecutamos el inicio de las series) que nos proporciona las condiciones iniciales ( $\theta_0$  y  $x_0$ ).

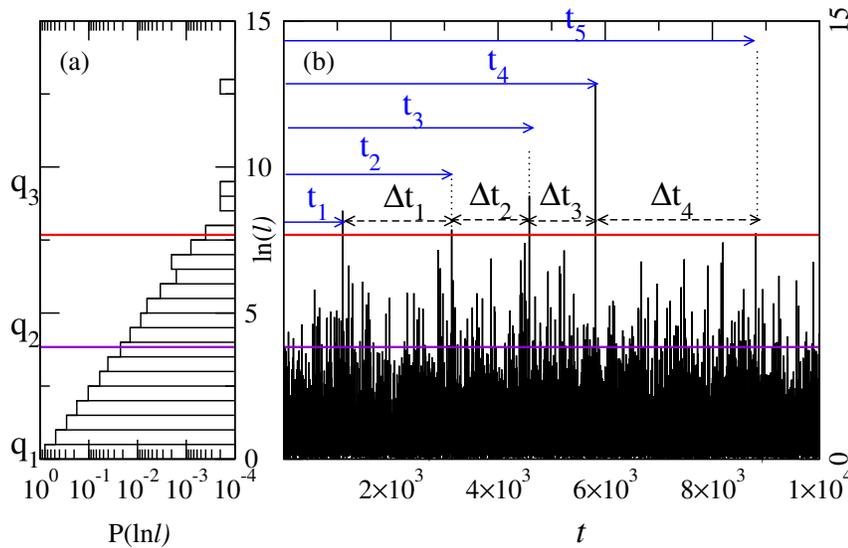


Figura 3.1: Ejemplo de serie de tiempo para el caso  $n = 3$  con  $10^4$  iteraciones para  $\alpha = 1$ . (a) Gráfica que muestra la distribución  $P(\ln l)$  obtenida de la serie de tiempo que se encuentra en (b).  $q_1$  corresponde al intervalo entre 0.0 y 3.8376 mientras que  $q_2$  se encuentran entre 3.8376 (línea violeta) y 7.6752 (línea roja) y  $q_3$  al intervalo abierto de 7.6752 en adelante.

Para realizar este trabajo se uso  $n = 10$  para diferentes valores de  $\alpha$ . Una vez separados los eventos en sus umbrales, la información que deseamos obtener es la estadística de ocurrencia y ver si existe una relación con respecto a su umbral correspondiente.

La ocurrencia entre eventos consecutivos para cada umbral se definió de la siguiente manera

$$\Delta t_n = t_{n+1} - t_n. \quad (3.3)$$

Para cada uno de los diferentes  $q_i$  se creó una serie de datos  $\Delta t$  y se obtuvo su distribución. En la Figura 3.1 se puede observar que para el último umbral ( $q_3$ ) se encuentran 5 eventos y existen 4  $\Delta t$ . La cantidad de información para cada uno de los umbrales fue de  $10^6$  datos.

### 3.3.3. Retornos de los precios de las criptomoneda

En el capítulo anterior en la sección 2.5 mostramos nuestra base de datos de series realistas. Además, obtuvimos una aproximación a la estacionariedad por medio de los retornos. En contraste con nuestras series artificiales, las series financieras tienen eventos extremos negativos (es decir, pérdida de su valor). Con la intención de usar la misma metodología que nuestras series de tiempo se aplicó un valor absoluto a los retornos, y el análisis se realizó a cuatro criptomonedas que son de mayor popularidad, como vemos en las siguientes figuras:

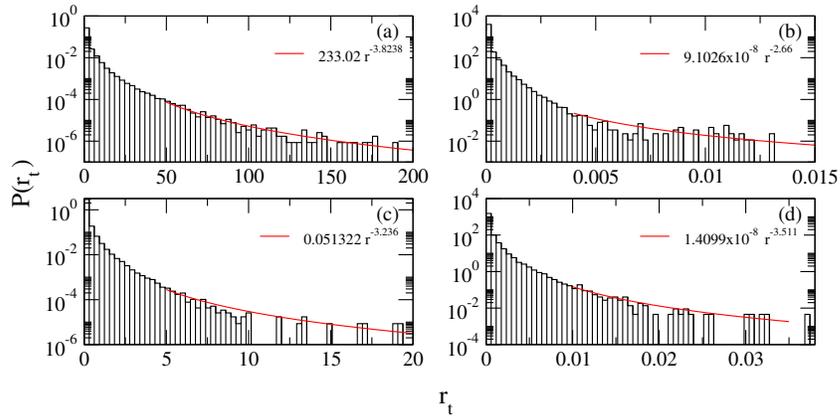


Figura 3.2: Distribución de los retornos lineales (a) bitcoin, (b) cardano, (c) ethereum y (d) ripple.

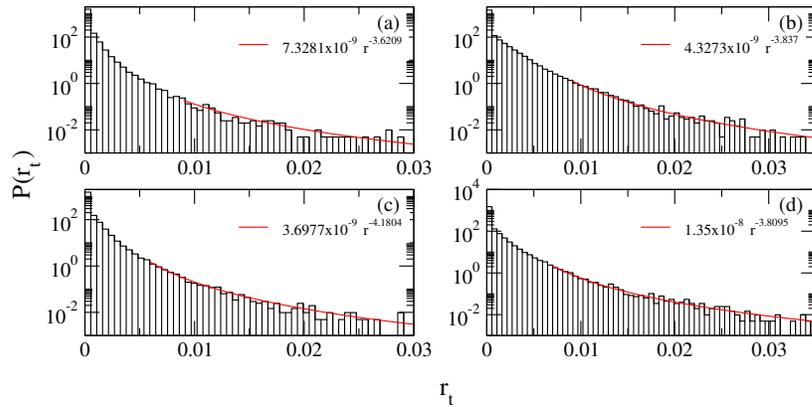


Figura 3.3: Distribución de los retornos logarítmicos (a) bitcoin, (b) cardano, (c) ethereum y (d) ripple.

Como podemos apreciar los retornos lineales decaen más lentamente que los logarítmicos, en la Figura 3.2 también se observa que existen espacios sin información dentro de la distribución y esto se debe a la amplitud de los eventos. Es por esta razón que se utilizaron los decaimientos logarítmicos, y con valores de  $\alpha$  entre 3 y 4 [4].

Para tener una forma semejante a la distribución de nuestras series, le sumamos una unidad a los retornos lineales después de obtener su valor absoluto y tomamos el logaritmo natural de éste, esperando encontrar alguna similitud. Podemos ver (figura 3.4) que sólo ethereum tiene una distribución similar a la nuestra.

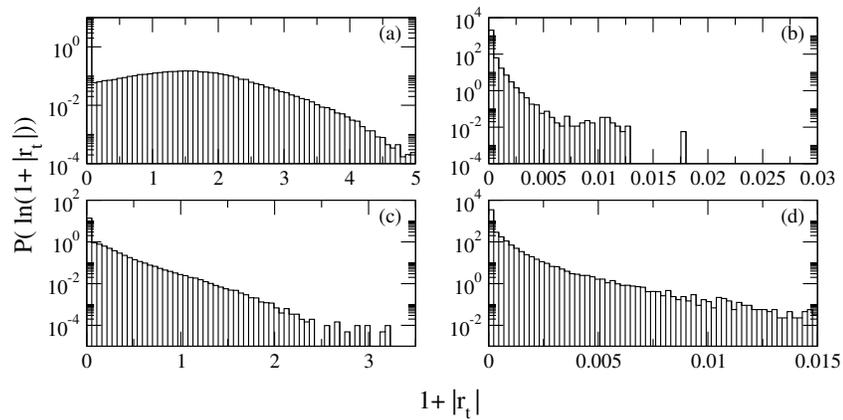


Figura 3.4: Distribución de los datos con el nuevo criterio (a) bitcoin, (b) cardano, (c) ethereum, (d) ripple.

# Capítulo 4

## Resultados y análisis de resultados

A continuación, presentamos los resultados de este trabajo, los cuales podemos dividir en dos partes, la primera parte corresponde al análisis de las series producidas por el mapeo de Lévy, mientras que la segunda parte está formada por las series de los retornos.

### 4.1. Resultados. Parte 1

En la figura 4.1 (a-e) se muestra la distribución de la fracción de datos por umbral, los cuales tienen la siguiente forma

$$\text{fracción de datos por umbral} = \gamma' e^{-\beta' q_i}, \quad (4.1)$$

donde  $q_i$  es el punto medio de cada umbral al que pertenece. También en esta misma figura 4.1 mostramos otros dos tipos de distribuciones (g) y (h), las cuales podemos comparar con las de tipo Lévy.

En particular se establecieron 10 umbrales para cada una de los diferentes  $\alpha$  (1/4, 1/2, 1, 3/2, 2, 4), para la distribución normal ( $\mu = 0, \sigma = 1$ ) y la exponencial ( $\lambda = 1$ ) y se obtuvieron las distribuciones de  $P(\Delta t)$  para los últimos 5 umbrales. Estos umbrales pertenecen al percentil .99 de su distribución.

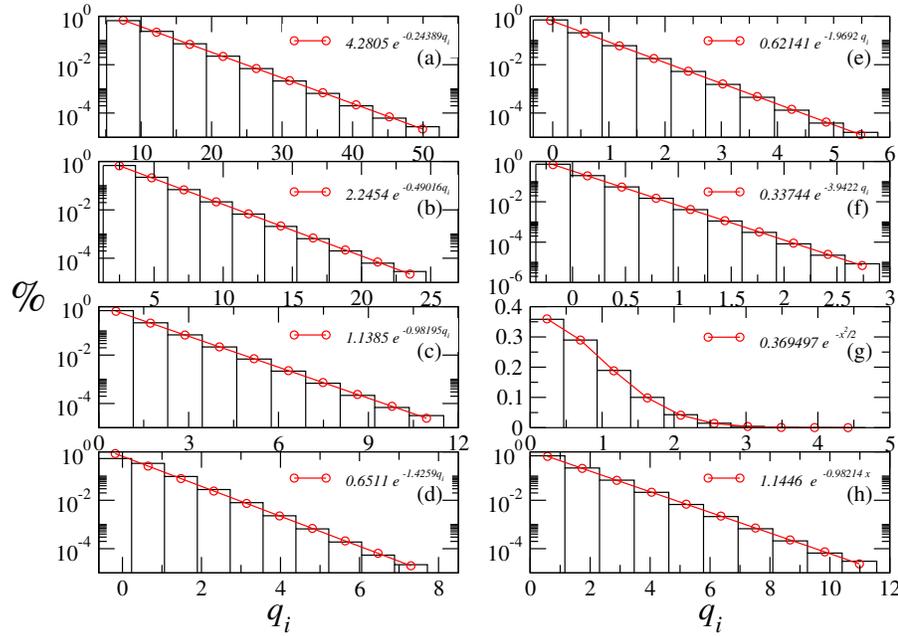


Figura 4.1: Porcentaje de los datos que aparecen en cada uno de los umbrales con (a)  $\alpha = 1/4$ , (b)  $\alpha = 1/2$ , (c)  $\alpha = 1$ , (d)  $\alpha = 3/2$  (e)  $\alpha = 2$ , (f)  $\alpha = 4$ , (g) distribución normal con media cero y desviación estándar uno y (h) distribución exponencial con  $\lambda = 1$ .

En las tablas 4.1-4.6 se encuentra toda la información de los umbrales, donde la primera columna pertenece al intervalo que tiene cada umbral, la segunda columna indica la cantidad de datos por umbral y la última columna en la que se tiene “%” representa la fracción por umbral.

Tabla 4.1: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.		
$\alpha = 1/4$		
Umbral	# de datos	%
166.08 - $1.85 \times 10^4$	24034322902	$6.5746 \times 10^{-1}$
$1.85 \times 10^4$ - $2.08 \times 10^6$	8672405085	$2.3723 \times 10^{-1}$
$2.08 \times 10^6$ - $2.33 \times 10^8$	2666041772	$7.2930 \times 10^{-2}$
$2.33 \times 10^8$ - $2.61 \times 10^{10}$	819565197	$2.2419 \times 10^{-2}$
$2.61 \times 10^{10}$ - $2.92 \times 10^{12}$	251972411	$6.8927 \times 10^{-3}$
$2.92 \times 10^{12}$ - $3.27 \times 10^{14}$	77440010	$2.1183 \times 10^{-3}$
$3.27 \times 10^{14}$ - $3.66 \times 10^{16}$	23814723	$6.5145 \times 10^{-4}$
$3.66 \times 10^{16}$ - $4.10 \times 10^{18}$	7321118	$2.0027 \times 10^{-4}$
$4.10 \times 10^{16}$ - $4.59 \times 10^{20}$	2250756	$6.1569 \times 10^{-5}$
$>4.59 \times 10^{20}$	1000001	$2.7355 \times 10^{-5}$

Tabla 4.2: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

$\alpha = 1/2$		
Umbral	# de datos	%
3.74 - 38.71	23991110529	$6.7855 \times 10^{-1}$
38.71 - 399.72	7828063486	$2.2140 \times 10^{-1}$
399.72 - 4127.44	2436245073	$6.8905 \times 10^{-2}$
4127.44 - 42618.46	758203705	$2.1444 \times 10^{-2}$
42618.46 - 440062.04	235931275	$6.6729 \times 10^{-3}$
440062.04 - 4543913.57	73435889	$2.0770 \times 10^{-3}$
4543913.57 - 46918725.76	22855191	$6.4642 \times 10^{-4}$
46918725.76 - 484464941.86	7108803	$2.0106 \times 10^{-4}$
484464941.86 - 5002400982.74	2214767	$6.2641 \times 10^{-5}$
>5002400982.74	1000001	$2.8283 \times 10^{-5}$

Tabla 4.3: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

$\alpha = 1$		
Umbral	# de datos	%
1 - 3.16	21601591925	$6.8377 \times 10^{-1}$
3.16 - 9.99	6831023193	$2.1622 \times 10^{-1}$
9.99 - 31.62	2160151508	$6.8376 \times 10^{-2}$
31.62 - 99.99	683129412	$2.1623 \times 10^{-2}$
99.99 - 316.22	216013722	$6.8376 \times 10^{-3}$
316.22 - 999.99	68324158	$2.1627 \times 10^{-3}$
999.99 - 3162.27	21604099	$6.8385 \times 10^{-4}$
3162.27 - 9999.98	6829008	$2.1616 \times 10^{-4}$
9999.98 - 31622.77	2159860	$6.8367 \times 10^{-5}$
>31622.77	1000001	$3.1653 \times 10^{-5}$

Tabla 4.4: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

$\alpha = 3/2$		
Umbral	# de datos	%
0.54 - 1.26	24271013107	$5.2967 \times 10^{-1}$
1.26 - 2.89	15360555459	$3.3521 \times 10^{-1}$
2.89 - 6.65	4412738538	$9.6299 \times 10^{-2}$
6.65 - 15.29	1267690832	$2.7665 \times 10^{-2}$
15.29 - 35.12	364128390	$7.9464 \times 10^{-3}$
35.12 - 80.66	104599974	$2.2827 \times 10^{-3}$
80.66 - 185.28	30046911	$6.5571 \times 10^{-4}$
185.28 - 425.57	8636571	$1.8847 \times 10^{-4}$

425.57 - 977.48	2480457	$5.4131 \times 10^{-5}$
> 977.48	1000001	$2.1823 \times 10^{-5}$

Tabla 4.5: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

$\alpha = 2$		
Umbral	# de datos	%
0.70 - 1.30	44177051842	$7.0693 \times 10^{-1}$
1.30 - 2.41	12937572447	$2.0703 \times 10^{-1}$
2.41 - 4.45	3799715753	$6.0804 \times 10^{-2}$
4.45 - 8.22	1114176837	$1.7829 \times 10^{-2}$
8.22 - 15.18	326713207	$5.2281 \times 10^{-3}$
15.18 - 28.04	95807985	$1.5331 \times 10^{-3}$
28.04 - 51.79	28093267	$4.4955 \times 10^{-4}$
51.79 - 95.64	8237108	$1.3181 \times 10^{-4}$
95.64 - 176.62	2415943	$3.8660 \times 10^{-5}$
>176.62	1000001	$1.6002 \times 10^{-5}$

Tabla 4.6: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

$\alpha = 4$		
Umbral	# de datos	%
0.70 - 0.97	85723901637	$7.2592 \times 10^{-1}$
0.97 - 1.35	23468073139	$1.9873 \times 10^{-1}$
1.35 - 1.87	6471174536	$5.4798 \times 10^{-2}$
1.87 - 2.59	1764692242	$1.4943 \times 10^{-2}$
2.59 - 3.58	481257453	$4.0753 \times 10^{-3}$
3.58 - 4.96	131265828	$1.1115 \times 10^{-3}$
4.96 - 6.87	35790939	$3.0308 \times 10^{-4}$
6.87 - 9.50	9757065	$8.2624 \times 10^{-5}$
9.50 - 13.15	9757065	$2.2554 \times 10^{-5}$
>13.15	1000001	$8.4681 \times 10^{-6}$

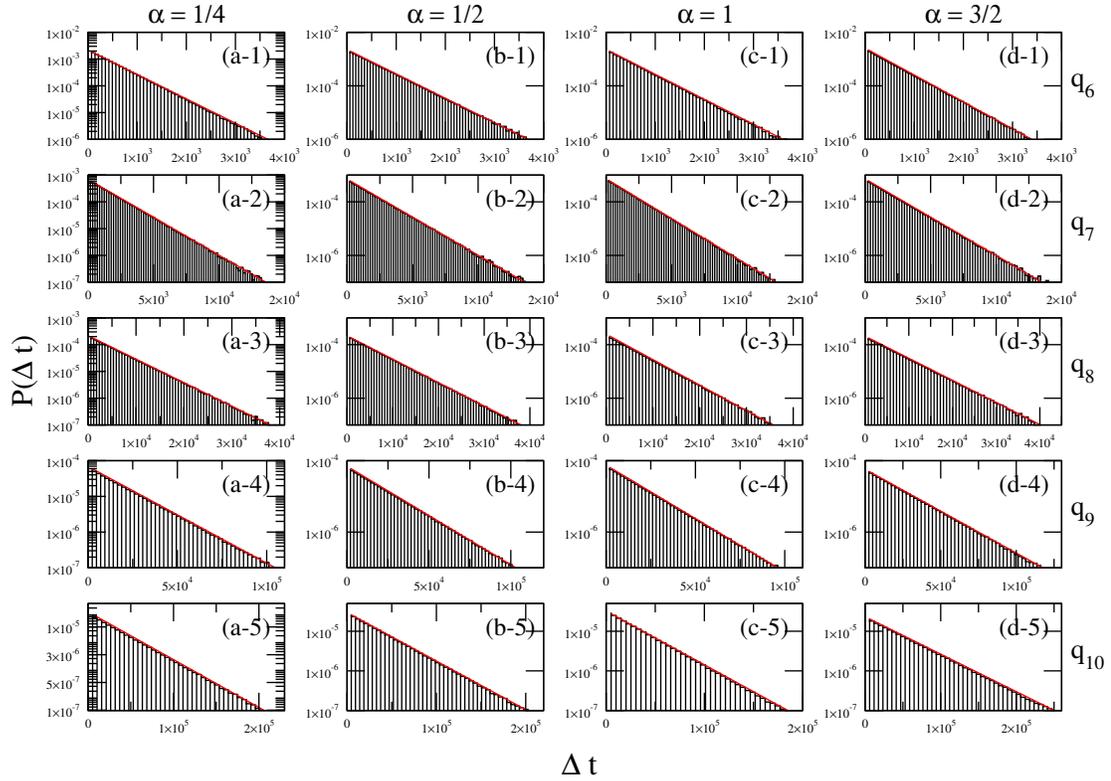


Figura 4.2: Distribución  $P(\Delta t)$  obtenida de la serie de tiempo para los umbrales  $q_6$ ,  $q_7$ ,  $q_8$ ,  $q_9$ , y  $q_{10}$  en (a-1 - a-5) con  $\alpha = 1/4$ , (b-1 - b-5)  $\alpha = 1/2$  (c-1 - c-5)  $\alpha = 1$  y (d-1 - d-5)  $\alpha = 3/2$ .

Tabla 4.7: Valores de  $\beta$  y  $\gamma$  correspondiente a la ecuación  $P(\Delta t) = \gamma e^{\beta \Delta t}$  para los diferentes umbrales.

$\alpha$	1/4		1/2	
$q_i$	$\beta$	$\gamma$	$\beta$	$\gamma$
$q_6$	0.0020989	0.002175	0.0020924	0.0022497
$q_7$	0.00065191	0.00069075	0.0006543	0.00070734
$q_8$	0.00020223	0.00021731	0.00020143	0.00021454
$q_9$	$6.2327 \times 10^{-5}$	$6.8759 \times 10^{-5}$	$6.35446 \times 10^{-5}$	$6.975 \times 10^{-5}$
$q_{10}$	$2.7726 \times 10^{-5}$	$3.0791 \times 10^{-5}$	$2.8197 \times 10^{-5}$	$2.9731 \times 10^{-5}$

Tabla 4.8: Valores de  $\beta$  y  $\gamma$  correspondiente a la ecuación  $P(\Delta t) = \gamma e^{\beta \Delta t}$  para los diferentes umbrales.

$\alpha$	1		3/2	
$q_i$	$\beta$	$\gamma$	$\beta$	$\gamma$
$q_6$	0.0021494	0.0023074	0.0023126	0.0025501
$q_7$	0.0006855	0.00072647	0.00065831	0.00071854
$q_8$	0.00021958	0.00024859	0.00019112	0.00020712
$q_9$	$6.8922 \times 10^{-5}$	$7.5058 \times 10^{-5}$	$5.4031 \times 10^{-5}$	$5.752 \times 10^{-5}$
$q_{10}$	$3.1203 \times 10^{-5}$	$3.2323 \times 10^{-5}$	$2.174 \times 10^{-5}$	$2.3218 \times 10^{-5}$

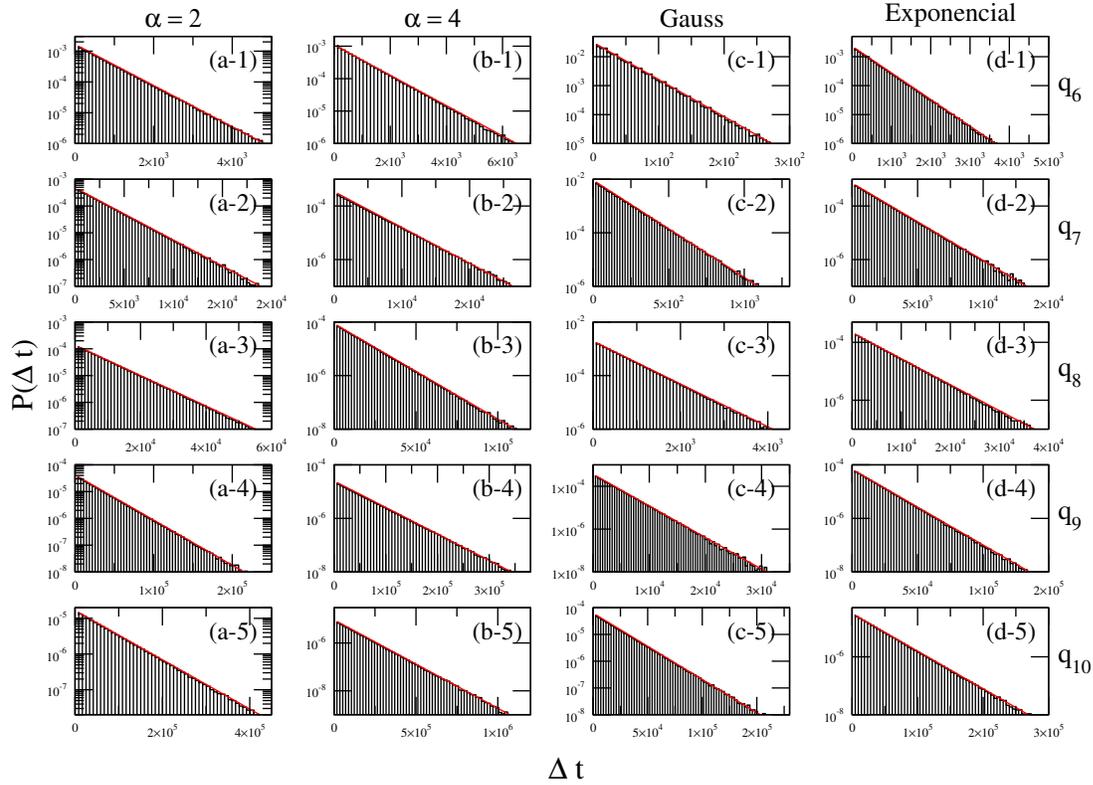


Figura 4.3: Distribución  $P(\Delta t)$  obtenida de la serie de tiempo para los umbrales  $q_6$ ,  $q_7$ ,  $q_8$ ,  $q_9$ , y  $q_{10}$  en (a-1 - a-5) con  $\alpha = 2$ , (b-1 - b-5)  $\alpha = 4$  (c-1 - c-5) normal y (d-1 - d-5) exponencial.

Tabla 4.9: Valores de  $\beta$  y  $\gamma$  correspondiente a la ecuación  $P(\Delta t) = \gamma e^{\beta \Delta t}$  para los diferentes umbrales.

$\alpha$	2		4	
$q_i$	$\beta$	$\gamma$	$\beta$	$\gamma$
$q_6$	0.00154	0.00166	0.0010777	0.0010983
$q_7$	0.0004483	0.00047796	0.00030842	0.00030591
$q_8$	0.00013092	0.0001375	$8.2122 \times 10^{-5}$	$8.7366 \times 10^{-5}$
$q_9$	$3.8819 \times 10^{-5}$	$4.2048 \times 10^{-5}$	$2.2783 \times 10^{-5}$	$2.4984 \times 10^{-5}$
$q_{10}$	$1.593 \times 10^{-5}$	$1.6824 \times 10^{-5}$	$8.5116 \times 10^{-6}$	$9.2802 \times 10^{-6}$

Tabla 4.10: Valores de  $\beta$  y  $\gamma$  correspondiente a la ecuación  $P(\Delta t) = \gamma e^{\beta \Delta t}$  para los diferentes umbrales.

	normal		exponencial	
$q_i$	$\beta$	$\gamma$	$\beta$	$\gamma$
$q_6$	0.029534	0.031314	0.0021178	0.0022882

$q_7$	0.0083592	0.0092346	0.0006651	0.00071494
$q_8$	0.0018396	0.0019288	0.0002067	0.00021395
$q_9$	0.00033944	0.00036101	$6.5128 \times 10^{-5}$	$6.832 \times 10^{-5}$
$q_{10}$	$5.706 \times 10^{-5}$	$6.2418 \times 10^{-5}$	$2.9883 \times 10^{-5}$	$3.1227 \times 10^{-5}$

En la figura 4.2 y 4.3 se tiene la distribución de ocurrencia  $\Delta t$ , y en esta podemos observar que cada una de las distribuciones tiene un decaimiento exponencial de la forma

$$P(\Delta t) = \gamma e^{-\beta \Delta t}. \quad (4.2)$$

Los valores de  $\gamma$  y  $\beta$  se muestran en las tablas 4.7-4.10. En la columna de los exponentes  $\beta$  encontramos que los valores son semejantes a los mostrados en las tablas 4.1-4.6 con respecto a la columna %.

En la figura 4.4 comparamos los valores que tiene el exponente  $\beta$  y el  $\ln \beta$ . También en (c) y (d) de esta misma figura, encontramos que el último umbral está cerrado, refiriéndose a que este está acotado por la misma longitud que los otros umbrales, es decir, descartando todos los eventos que se salen de este rango.

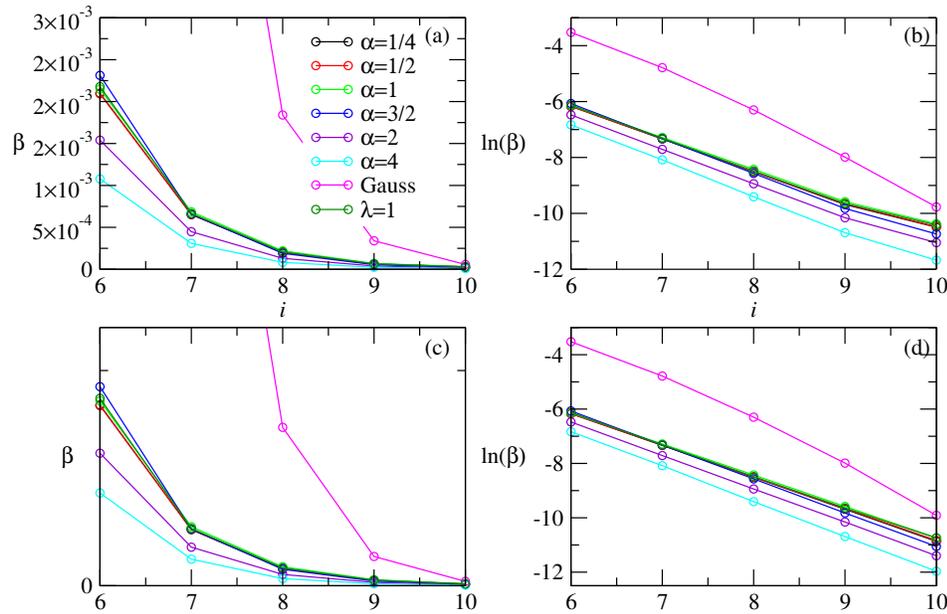


Figura 4.4: (a) y (c) exponente  $\beta$ , (b) y (d)  $\ln \beta$  para cada uno de los umbrales, (a) y (b) tienen el último umbral abierto, (c) y (d) el último umbral cerrado.

En la figura 4.4 (d) se aprecia que el  $\ln \beta$  es una línea recta para los diferentes valores  $\alpha$ , y estas son líneas prácticamente paralelas con una pendiente similar, mientras que para el caso del umbral abierto (b) existe una desviación de la línea para el último umbral que hacen que no sean líneas rectas, esto se debe a que entran otros datos que pueden tener valores de evento con amplitudes mayores al del intervalo establecido.

## 4.2. Resultados. Parte 2

A continuación, presentamos los umbrales para las series de los retornos logarítmicos.

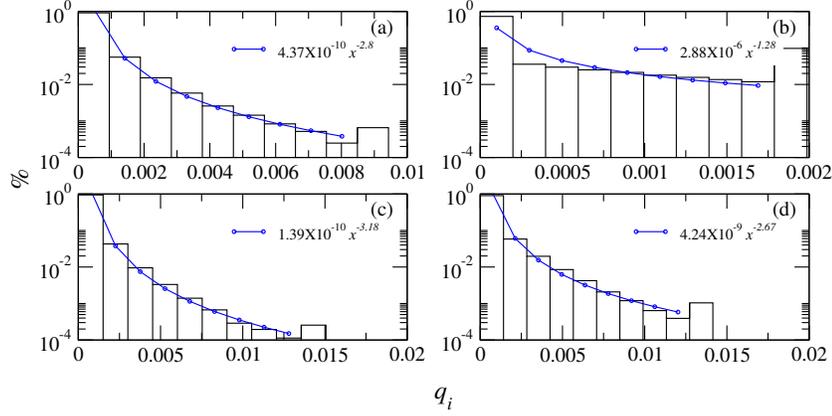


Figura 4.5: Distribución de los datos para los 10 umbrales (a) bitcoin, (b) cardano, (c) ethereum, (d) ripple.

Tabla 4.11: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.		
Bitcoin		
Umbral	# de datos	%
0 - $9.43 \times 10^{-4}$	347871	$9.1685 \times 10^{-1}$
$9.43 \times 10^{-4}$ - $1.88 \times 10^{-3}$	21200	$5.5875 \times 10^{-2}$
$1.88 \times 10^{-3}$ - $2.82 \times 10^{-3}$	5766	$1.5197 \times 10^{-2}$
$2.82 \times 10^{-3}$ - $3.77 \times 10^{-3}$	2204	$5.8089 \times 10^{-3}$
$3.77 \times 10^{-3}$ - $4.71 \times 10^{-3}$	977	$2.5750 \times 10^{-3}$
$4.71 \times 10^{-3}$ - $5.65 \times 10^{-3}$	545	$1.4364 \times 10^{-3}$
$5.65 \times 10^{-3}$ - $6.60 \times 10^{-3}$	315	$8.3022 \times 10^{-3}$
$6.60 \times 10^{-3}$ - $7.54 \times 10^{-3}$	196	$5.1658 \times 10^{-4}$
$7.54 \times 10^{-3}$ - $8.48 \times 10^{-3}$	94	$2.4774 \times 10^{-4}$
$> 8.48 \times 10^{-3}$	248	$6.5363 \times 10^{-4}$

Tabla 4.12: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.		
Cardano		
Umbral	# de datos	%
0 - $1.98 \times 10^{-4}$	277471	$7.3131 \times 10^{-1}$
$1.98 \times 10^{-4}$ - $3.97 \times 10^{-4}$	13614	$3.5881 \times 10^{-2}$
$3.97 \times 10^{-4}$ - $5.96 \times 10^{-4}$	11314	$2.9819 \times 10^{-2}$

$5.96 \times 10^{-4} - 7.94 \times 10^{-4}$	9524	$2.5101 \times 10^{-2}$
$7.94 \times 10^{-4} - 9.93 \times 10^{-4}$	8112	$2.1380 \times 10^{-2}$
$9.93 \times 10^{-4} - 1.19 \times 10^{-3}$	6802	$1.7927 \times 10^{-2}$
$1.19 \times 10^{-3} - 1.39 \times 10^{-3}$	5919	$1.5600 \times 10^{-2}$
$1.39 \times 10^{-3} - 1.58 \times 10^{-3}$	5168	$1.1860 \times 10^{-2}$
$1.58 \times 10^{-3} - 1.78 \times 10^{-3}$	4500	$5.7720 \times 10^{-2}$
$> 1.78 \times 10^{-3}$	36992	$9.7497 \times 10^{-2}$

Tabla 4.13: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

Ethereum		
Umbral	# de datos	%
$0 - 1.50 \times 10^{-3}$	357174	$9.4137 \times 10^{-1}$
$1.50 \times 10^{-3} - 3.01 \times 10^{-3}$	16257	$4.2847 \times 10^{-2}$
$3.01 \times 10^{-3} - 4.51 \times 10^{-3}$	3612	$9.5198 \times 10^{-3}$
$4.51 \times 10^{-3} - 6.02 \times 10^{-3}$	1265	$3.3340 \times 10^{-3}$
$6.02 \times 10^{-3} - 7.53 \times 10^{-3}$	530	$1.3968 \times 10^{-3}$
$7.53 \times 10^{-3} - 9.03 \times 10^{-3}$	254	$6.6944 \times 10^{-4}$
$9.03 \times 10^{-3} - 1.05 \times 10^{-2}$	110	$2.8991 \times 10^{-4}$
$1.05 \times 10^{-2} - 1.20 \times 10^{-2}$	74	$1.9503 \times 10^{-4}$
$1.20 \times 10^{-3} - 1.35 \times 10^{-2}$	43	$1.1333 \times 10^{-4}$
$> 1.35 \times 10^{-2}$	97	$2.5565 \times 10^{-4}$

Tabla 4.14: Valores del intervalo para cada uno de los umbrales, número de eventos encontrados por umbral y su fracción correspondiente.

Ripple		
Umbral	# de datos	%
$0 - 1.41 \times 10^{-3}$	342975	$9.0395 \times 10^{-1}$
$1.41 \times 10^{-3} - 2.82 \times 10^{-3}$	22109	$5.8271 \times 10^{-2}$
$2.82 \times 10^{-3} - 4.24 \times 10^{-3}$	7462	$1.9667 \times 10^{-2}$
$4.24 \times 10^{-3} - 5.65 \times 10^{-3}$	3233	$8.5209 \times 10^{-3}$
$5.65 \times 10^{-3} - 7.07 \times 10^{-3}$	1605	$4.2301 \times 10^{-3}$
$7.07 \times 10^{-3} - 8.48 \times 10^{-3}$	790	$2.0821 \times 10^{-3}$
$8.48 \times 10^{-3} - 9.90 \times 10^{-3}$	454	$1.1965 \times 10^{-3}$
$9.90 \times 10^{-3} - 1.13 \times 10^{-2}$	243	$6.4045 \times 10^{-4}$
$1.13 \times 10^{-3} - 1.27 \times 10^{-2}$	149	$3.9270 \times 10^{-4}$
$> 1.27 \times 10^{-2}$	396	$1.0437 \times 10^{-3}$

Tanto en la figura 4.5 como en las tablas 4.11-4.14 se observa que para el último umbral, los eventos más extremos no son tan escasos, pero esto se debe a que este

umbral queda abierto con la finalidad de no omitir eventos que podrían tener una magnitud mayor al intervalo dado. También nos encontramos que que estos tiene un comportamiento en forma de ley de potencia.

En la figura 4.6 encontramos que la mayoría de los  $P(\Delta t)$  se comportan casi constante a partir de cierto  $\Delta t$ . Sin embargo en los retornos del cardano y ripple podemos apreciar una decaimiento lento en las colas, semejante al de una ley de potencia. Dado a las fluctuaciones presentes en las distribuciones no se realizó ninguna curva de ajuste.

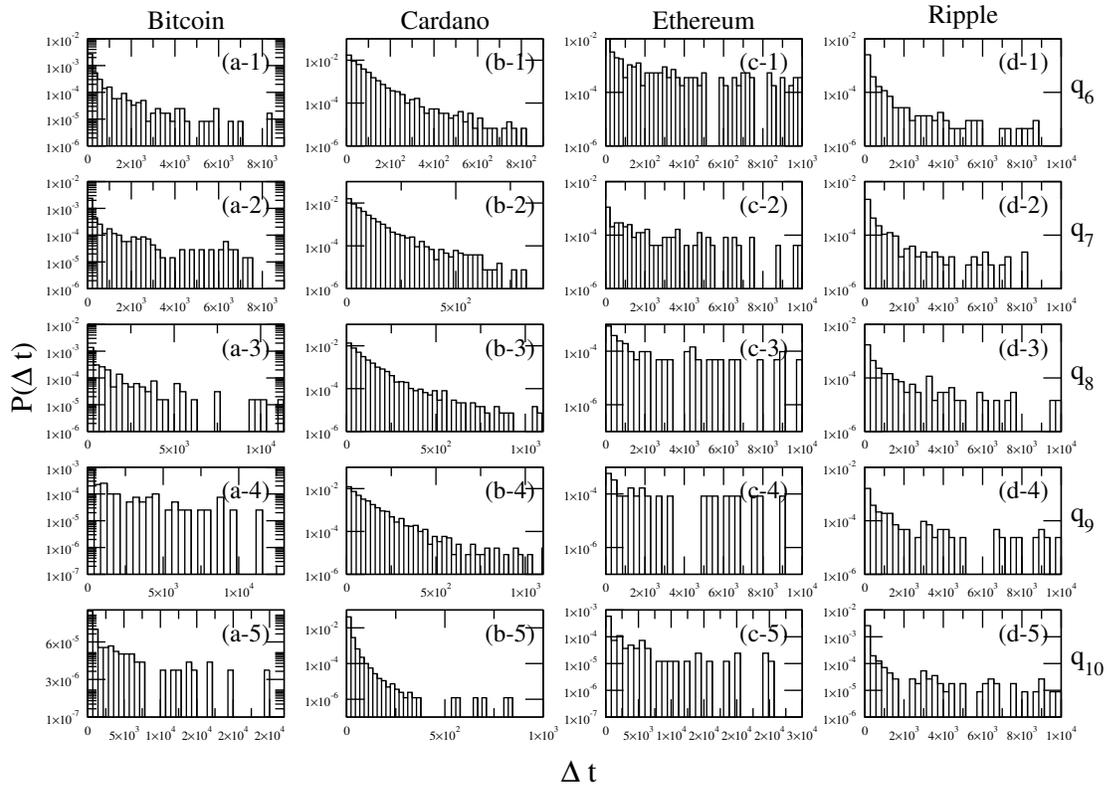


Figura 4.6: Distribución  $P(\Delta t)$  obtenida de de los retorno, para los umbrales  $q_6$ ,  $q_7$ ,  $q_8$ ,  $q_9$ , y  $q_{10}$  en (a-1 - a-5) con bitcoin, (b-1 - b-5) cardano (c-1 - c-5) ethereum y (d-1 - d-5) ripple.

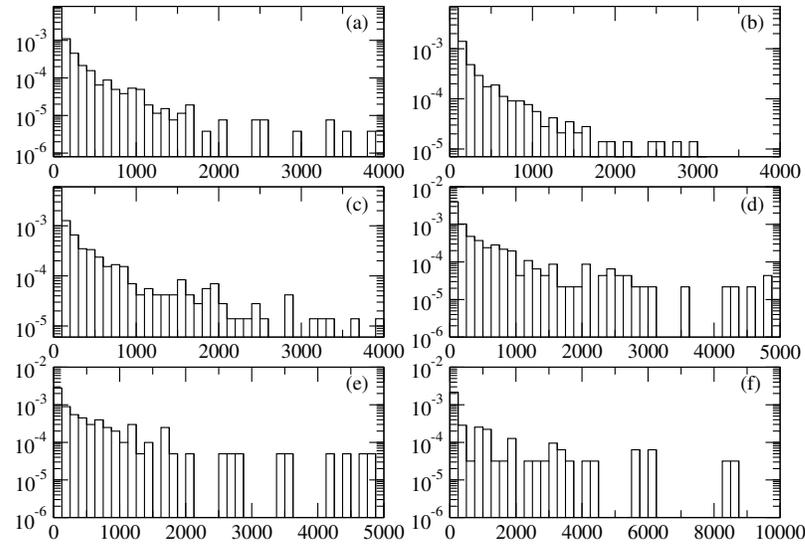


Figura 4.7: Distribución de intervalos e tiempo para cada los 6 últimos umbrales para la criptomoneda del ethereum.

La figura 4.7 se creo a partir del retorno propuesto de la figura 3.4 ya que esta distribución se asemejaba mejor a la nuestra. Sin embargo observamos que esta aparenta seguir teniendo un decaimiento en forma de ley de potencia en la distribución  $\Delta t$ .



# Capítulo 5

## Conclusiones

En este trabajo se realizó un estudio de las características de las series producidas por el mapeo de Lévy, encontrando el papel que juega el parámetro  $\alpha$ , estas observaciones las encontramos en las tablas 4.1-4.6 en la columna que indica el tamaño del intervalo por umbral, en donde podemos apreciar que entre más pequeño es  $\alpha$  produce eventos de mayor magnitud, y cuando  $\alpha$  aumenta de valor la magnitud de los eventos disminuye. Además vemos que para la ec. (4.1)  $\beta' \approx \alpha$ .

Por otra parte, encontramos que la distribución  $P(\Delta t)$  se comporta como una ley exponencial ec. (4.2). El coeficiente y el exponente muestran tener un valor con el mismo orden de magnitud. Y vemos que el valor  $\beta$  es aproximadamente igual al valor que encontramos en la columna de los % de las tablas 4.1-4.6. Lo que nos lleva a pensar que el valor del exponente es  $\beta \approx \gamma' e^{-\alpha q_i}$ , concluyendo que el parámetro  $\alpha$  también se encuentra presente en la distribución de  $P(\Delta t)$ . Aunado a esto, en la figura 4.4 vemos que conforme el umbral aumenta el valor  $\beta$  disminuye, además encontramos una relación lineal para  $\ln \beta$ .

Con respecto a las series financieras el análisis se realizó mejor con sus retornos logarítmicos, ya que la serie originalmente muestra una tendencia no estacionaria (porque la distribución no fluctuá alrededor de algún valor) y con los retornos logarítmicos tenemos una distribución casi simétrica, en forma de leyes de potencia. Al aplicar los mismos criterios a la serie de retornos lineales nos encontramos que muchos retornos eran cero, por lo que aplicar un logaritmo directamente a esta nueva serie nos genera una indeterminación . Por eso el análisis se realizó a los retornos logarítmicos donde encontramos que la distribución de los umbrales era semejante a su distribución original (excepto el cardano).

Por otro lado, tenemos  $P(\Delta t)$  tiene un comportamiento más o menos constante o en algunos casos similar a una ley de potencia, lo que nos lleva pensar que si la base de datos fuera más grande, tal vez podríamos apreciar una ley de potencia en todas las  $P(\Delta t)$ .

La importancia de haber creado estas series de tiempo es que están caracterizadas por la distribución tipo Lévy la cual se encuentra presente en muchos fenómenos y sistemas dinámicos que se comportan como variables aleatorias que se distribuyen

de esta forma, tales como: alimentación animal silvestre [35], movilidad humana [36], estadística de recurrencia de terremotos [37], vuelo de mosquitos [38], series financieras, etc.

Aparte de servir como un generador de números pseudo-aleatorios, una aplicación de este modelo es que describe la dinámica del billar cosenoidal, también puede ser usado como un tipo de señal, tal como es el caso de ruido blanco.

## 5.1. Trabajos a futuro

En esta tesis se estudiaron, series producidas por el mapeo de Lévy esperando encontrar un comportamiento similar en las series de los retornos de las criptomonedas. Las series financieras estudiadas en este trabajo provienen de sistemas complejos tales que no se ajustaron a nuestro modelo propuestos, pero esto queda abierto para el estudio de otras series.

# Apéndice



# Apéndice A

## Figuras complementarias

Mapeos de Lévy para diferentes valores de  $\alpha$ .

En este apartado podemos ver la transición del caso integrable al mixto para los valores de  $\alpha$ . Observamos que para  $\alpha$  pequeños su transición a las situaciones caóticas es más apresurada, mientras que para  $\alpha$  más grandes, esta transición es más lenta.

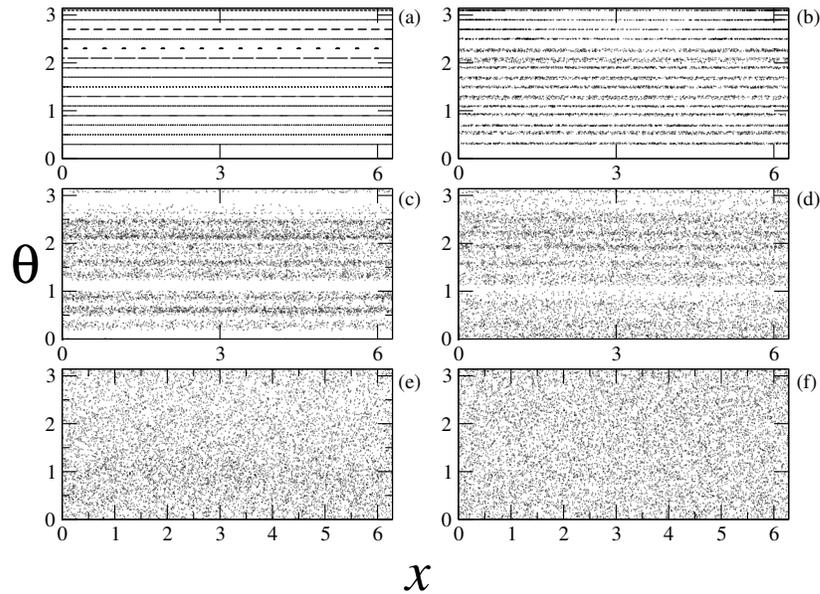


Figura A.1: Mapeo de Lévy con  $\alpha = 1/4$ ,  $c = 1$  y  $x_0 = 0.1$  (a)-(e) iterado 665 veces para cada uno de los diferentes  $\theta_0 = 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.5, 2.7, 2.9, 3.1$  (a)  $w = 0$ , (b)  $w = 0.001$ , (c)  $w = 0.005$ , (d)  $w = 0.01$ , (e)  $w = 0.1$ , (f) con  $10^4$  iteraciones,  $w = 1$  y  $\theta_0 = 0.1$ .

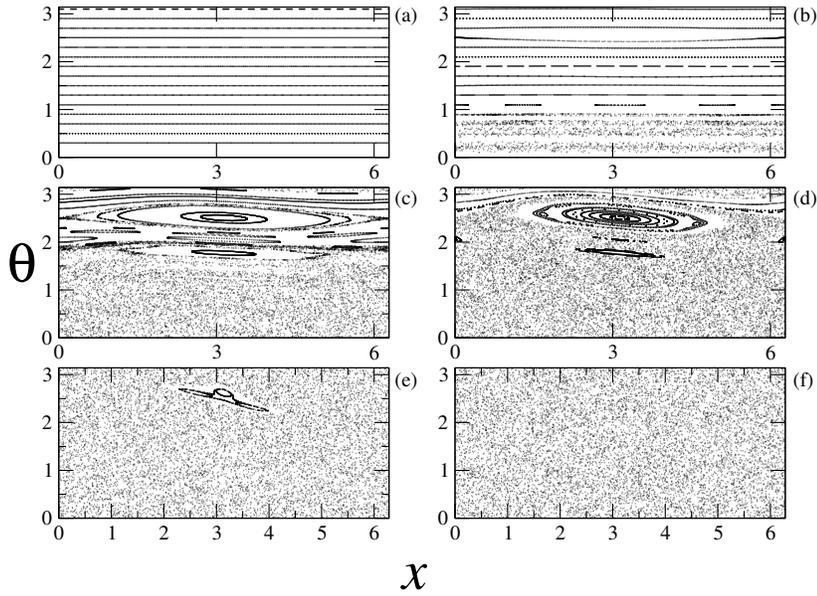


Figura A.2: Mapeo de Levy con  $\alpha = 1/2$ ,  $c = 1$  y  $x_0 = 0.1$  (a)-(e) iterado 665 veces para cada uno de los diferentes  $\theta_0 = 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.5, 2.7, 2.9, 3.1$ , (a)  $w = 0$ , (b)  $w = 0.005$ , (c)  $w = 0.05$ , (d)  $w = 0.1$ , (e)  $w = 0.3$ , (f) con  $10^4$  iteraciones,  $w = 1$  y  $\theta_0 = 0.1$ .

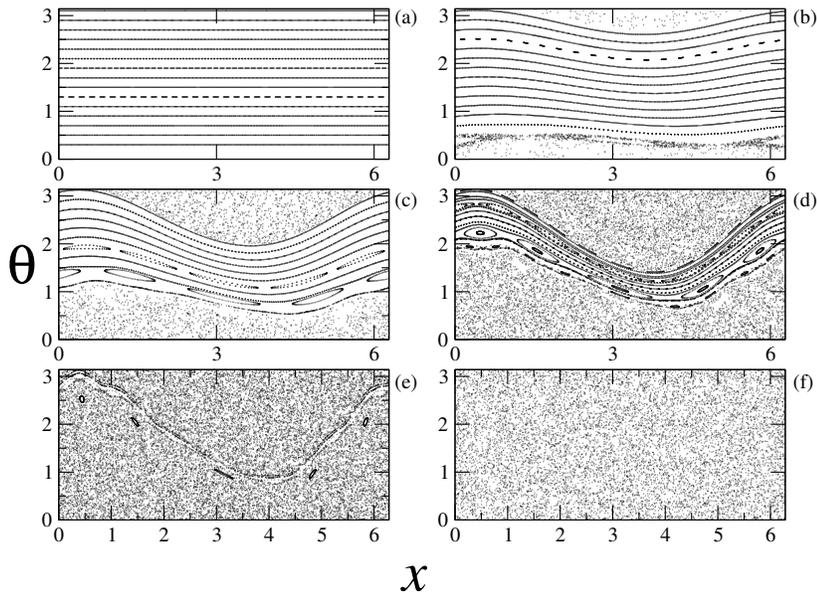


Figura A.3: Mapeo de Levy con  $\alpha = 3/2$ ,  $c = 1$   $x_0 = 0.1$  (a)-(e) iterado 665 veces para cada uno de los diferentes  $\theta_0 = 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.5, 2.7, 2.9, 3.1$ , (a)  $w = 0$ , (b)  $w = 0.1$ , (c)  $w = 0.25$ , (d)  $w = 0.4$ , (e)  $w = 0.55$ , (f)  $w = 1$ .

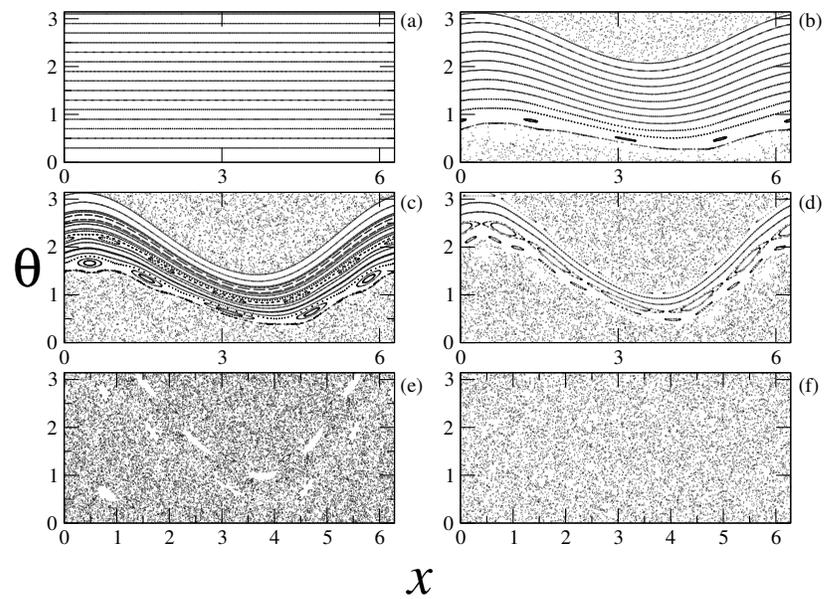


Figura A.4: Mapeo de Levy para  $\alpha = 2$ ,  $c = 1$  y  $x_0 = 0.1$  (a)-(e) iterado 665 veces para cada uno de los diferentes  $\theta_0 = 0.3, 0.5, 0.7, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.5, 2.7, 2.9, 3.1$ , (a)  $w = 0$ , (b)  $w = 0.2$ , (c)  $w = 0.35$ , (d)  $w = 0.45$ , (e)  $w = 0.6$ , (f)  $w = 1$ .



# Apéndice B

## Otras distribuciones

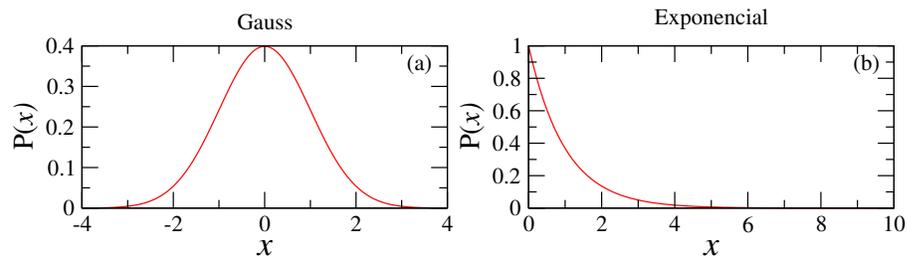


Figura B.1: distribuciones

### Exponencial

La distribución exponencial es un proceso en el que los eventos ocurren de manera continua e independiente a una tasa promedio constante

#### Función de densidad de probabilidad

$$f(x) = \lambda \exp(-\lambda x) \quad x \in (0, \infty] \quad (\text{B.1})$$

y  $\lambda > 0$

### Gausiana

Las distribuciones normales son importantes en las estadísticas y a menudo se usan en las ciencias naturales y sociales para representar variables aleatorias de valor real cuyas distribuciones no se conocen.

### Función de densidad de probabilidad

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) \quad (\text{B.2})$$

donde  $\mu$  es su media,  $\sigma^2$  es la varianza y en principio  $x \in (-\infty, \infty)$

# Bibliografía

- [1] A. Chakraborti, I. M. Toke, M. Patriarca, and F. Abergel, “Econophysics review: I. empirical facts,” *Quantitative Finance*, vol. 11, no. 7, pp. 991–1012, 2011.
- [2] M. Chavez, M. Ghil, and J. Urrutia-Fucugauchi, *Extreme events: Observations, modeling, and economics*. John Wiley & Sons, 2016.
- [3] R. Chicheportiche and A. Chakraborti, “Copulas and time series with long-ranged dependencies,” *Physical Review E*, vol. 89, no. 4, p. 042117, 2014.
- [4] S. Begušić, Z. Kostanjčar, H. E. Stanley, and B. Podobnik, “Scaling properties of extreme price fluctuations in bitcoin markets,” *Physica A: Statistical Mechanics and its Applications*, vol. 510, pp. 400–406, 2018.
- [5] R. Aggarwal and H. Prabhakar, “Electrocardiogram abnormalities,” in *Complications in Neuroanesthesia*. Elsevier, 2016, pp. 111–115.
- [6] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling extremal events: for insurance and finance*. Springer Science & Business Media, 2013, vol. 33.
- [7] R. L. Ingraham, *A Survey of Nonlinear Dynamics: chaos Theory*. World scientific, 1992, vol. 91, no. 1.
- [8] E. Ott, *Chaos in dynamical systems*. Cambridge university press, 2002.
- [9] C. Gignoux and B. Silvestre-Brac, *Solved problems in Lagrangian and Hamiltonian mechanics*. Springer Science & Business Media, 2009.
- [10] J. A. Méndez-Bermúdez, “Estudio de la correspondencia clásica-cuántica en billares caóticos,” Ph.D. dissertation, IFUAP, 2003.
- [11] S. N. Rasband, *Chaotic dynamics of nonlinear systems*. Courier Dover Publications, 2015.
- [12] J. J. B. Nieto, *Caos, orden y desorden: en el sistema monetario y financiero internacional, el caso de México*. Plaza y Valdés, 2002.
- [13] B. V. Chirikov, “A universal instability of many-dimensional oscillator systems,” *Physics reports*, vol. 52, no. 5, pp. 263–379, 1979.

- 
- [14] C. M. Bender, J. Feinberg, D. W. Hook, and D. J. Weir, “Chaotic systems in complex phase space,” *Pramana*, vol. 73, no. 3, p. 453, 2009.
- [15] R. Chirikov and V. Vecheslavov, “Chaotic dynamics of comet halley,” *Astronomy and Astrophysics*, vol. 221, pp. 146–154, 1989.
- [16] B. V. Chirikov, “Resonance processes in magnetic traps,” *The Soviet Journal of Atomic Energy*, vol. 6, no. 6, pp. 464–470, 1960.
- [17] B. Chirikov, “Particle confinement and adiabatic invariance,” *Proc. R. Soc. Lond. A*, vol. 413, no. 1844, pp. 145–156, 1987.
- [18] N. Chernov and R. Markarian, *Introducción al estudio de billares caóticos*. Courier Dover Publications, 2005.
- [19] D. Heymann, R. Perazzo, and M. Zimmermann, *Economía de fronteras abiertas: exploraciones en sistemas sociales complejos*. Teseo, 2013.
- [20] J. Méndez-Bermúdez, J. A. de Oliveira, and E. D. Leonel, “Two-dimensional nonlinear map characterized by tunable lévy flights,” *Physical Review E*, vol. 90, no. 4, p. 042138, 2014.
- [21] J. A. de Oliveira, R. Bizão, and E. D. Leonel, “Finding critical exponents for two-dimensional hamiltonian maps,” *Physical Review E*, vol. 81, no. 4, p. 046212, 2010.
- [22] E. D. Leonel, J. A. De Oliveira, and F. Saif, “Critical exponents for a transition from integrability to non-integrability via localization of invariant tori in the hamiltonian system,” *Journal of Physics A: Mathematical and Theoretical*, vol. 44, no. 30, p. 302001, 2011.
- [23] J. A. De Oliveira and E. D. Leonel, “Locating invariant tori for a family of two-dimensional hamiltonian mappings,” *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 21-22, pp. 3727–3731, 2011.
- [24] J. A. De Oliveira, C. P. Dettmann, D. R. Da Costa, and E. D. Leonel, “Scaling invariance of the diffusion coefficient in a family of two-dimensional hamiltonian mappings,” *Physical Review E*, vol. 87, no. 6, p. 062904, 2013.
- [25] R. Yaffee and M. McGee, *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS*. Elsevier Science, 2000.
- [26] R. H. Shumway and D. S. Stoffer, *Time series analysis and its applications: with R examples*. Springer, 2017.
- [27] S. Albeverio, V. Jentsch, and H. Kantz, *Extreme events in nature and society*. Springer Science & Business Media, 2006.

- [28] D. Sornette, A. Johansen *et al.*, “Significance of log-periodic precursors to financial crashes,” *Quantitative Finance*, vol. 1, no. 4, pp. 452–471, 2001.
- [29] F. A. Nava, *Procesamiento de series de tiempo*. Fondo de Cultura Económica, 2015.
- [30] C. Chatfield, *The analysis of time series: an introduction*. Chapman and Hall/CRC, 2016.
- [31] A. V. Metcalfe and P. S. Cowpertwait, *Introductory time series with R*. Springer, 2009.
- [32] F. Wang, K. Yamasaki, S. Havlin, and H. E. Stanley, “Scaling and memory of intraday volatility return intervals in stock markets,” *Physical Review E*, vol. 73, no. 2, p. 026117, 2006.
- [33] S. F. Fernández, J. M. C. Sánchez, A. Córdoba, J. M. Cordero, and A. C. Largo, *Estadística descriptiva*. Esic Editorial, 2002.
- [34] M. Kemp, *Extreme Events-robust portfolio construction in the presence of fat tails*. John Wiley & Sons, 2011.
- [35] G. Viswanathan, E. Raposo, and M. Da Luz, “Lévy flights and superdiffusion in the context of biological encounters and random searches,” *Physics of Life Reviews*, vol. 5, no. 3, pp. 133–150, 2008.
- [36] D. Brockmann, L. Hufnagel, and T. Geisel, “The scaling laws of human travel,” *Nature*, vol. 439, no. 7075, p. 462, 2006.
- [37] O. Sotolongo-Costa, J. Antoranz, A. Posadas, F. Vidal, and A. Vazquez, “Lévy flights and earthquakes,” *Geophysical research letters*, vol. 27, no. 13, pp. 1965–1968, 2000.
- [38] S. y. L. E. D. Botari, Tiago y Alves, “Explaining the high number of infected people by dengue in rio de janeiro in 2008 using a susceptible-infective-recovered model,” *Physical Review E*, no. 3, p. 037101, 2011.
- [39] O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, *Lévy processes: theory and applications*. Springer Science & Business Media, 2012.
- [40] T. Petrosky, “Chaos and cometary clouds in the solar system,” *Physics Letters A*, vol. 117, no. 7, pp. 328–332, 1986.
- [41] D. L. Shepelyansky and A. D. Stone, “Chaotic landau level mixing in classical and quantum wells,” *Physical review letters*, vol. 74, no. 11, p. 2098, 1995.