

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS

INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

CENTRO DE INVESTIGACIÓN EN CIENCIAS

ANÁLISIS DE IRREVERSIBILIDAD EN SISTEMAS  
FINANCIEROS

TESIS PROFESIONAL PARA OBTENER EL GRADO DE:

DOCTORADO EN CIENCIAS

PRESENTA:

JESSICA MORALES HERRERA

DIRECTOR: DR. RAÚL SALGADO GARCÍA

CUERNAVACA, MORELOS

MAYO, 2024



## Resumen

La eficiencia del mercado es un pilar central en la teoría económica. Propuesta en 1970 por Eugene Fama, la Hipótesis de Mercados Eficientes (HME) sostiene que los precios de los activos financieros reflejan toda la información disponible en un momento dado. Esencialmente, sostiene que es altamente improbable obtener rendimientos superiores al promedio, independientemente de la información que se posea. Aunque esta hipótesis ha sido cuestionada en varias ocasiones por economistas que argumentan que el mercado es ineficiente, citando comportamientos irracionales de los inversores y crisis financieras como evidencia.

Algunas investigaciones apuntan a que la ineficiencia del mercado puede estar vinculada a la irreversibilidad de las series de tiempo del mercado. La lógica es que si una serie de tiempo es irreversible, se puede establecer una dirección temporal, lo que nos permite determinar qué sucedió antes y qué después. Esto se traduce en una especie de “predicción” del comportamiento del mercado, lo cual contradice la idea de que los mercados son eficientes.

En este estudio, nos enfocamos en buscar una métrica que sea sensible a la irreversibilidad en las series de tiempo financieras. Además, tratamos de determinar si un mercado ineficiente es también irreversible. Para hacer esto, comparamos un índice de ineficiencia descrito en la literatura con los hallazgos de nuestro análisis. Utilizamos series temporales de criptomonedas, como el Bitcoin, en nuestro estudio, con el fin de explorar la eficiencia de un mercado con alta volatilidad.

Empleamos técnicas como los tiempos de retorno, el índice de asimetría desfasada y los patrones de tendencias para examinar la irreversibilidad de las series de tiempo. Partimos de la hipótesis de que el comportamiento del índice de irreversibilidad sería estadísticamente similar al índice de ineficiencia. Aunque los resultados obtenidos no responden directamente a la pregunta de si la irreversibilidad es también una medida de ineficiencia, son sumamente interesantes y abren la puerta a investigaciones más profundas sobre la relación entre la irreversibilidad y la ineficiencia en una serie de tiempo. En futuros estudios, nos gustaría explorar más a fondo esta relación.

**Abstract**

Market efficiency is a central pillar in economic theory. Proposed in 1970 by Eugene Fama, the Efficient Market Hypothesis (EMH) holds that the prices of financial assets reflect all the available information at a given moment. Essentially, it argues that it is highly unlikely to achieve above-average returns, regardless of the information one possesses. Although this hypothesis has been challenged on several occasions by economists who argue that the market is inefficient, citing irrational investor behavior and financial crises as evidence.

Some research points to the market's inefficiency being linked to the irreversibility of market time series. The logic is that if a time series is irreversible, a temporal direction can be established, which allows us to determine what happened before and what after. This translates into a sort of "prediction" of market behavior, which contradicts the idea that markets are efficient.

In this study, we focus on looking for a metric that is sensitive to irreversibility in financial time series. In addition, we try to determine if an inefficient market is also irreversible. To do this, we compare an inefficiency index described in the literature with the findings of our analysis. We use cryptocurrency time series such as Bitcoin in our study, as we wanted to explore the efficiency of a market with high volatility.

We employ techniques such as return times, lagged skewness index, and trend patterns to examine the irreversibility of time series. We start from the hypothesis that the behavior of the irreversibility index would be statistically similar to the inefficiency index. Although the results obtained do not directly answer the question of whether irreversibility is also a measure of inefficiency, they are extremely interesting and open the door to deeper investigations into the relationship between irreversibility and inefficiency in a time series. In future studies, we would like to explore this relationship more thoroughly.





# Índice general

<b>1. Introducción.</b>	<b>1</b>
<b>2. Marco teórico.</b>	<b>7</b>
2.1. Procesos estocásticos. . . . .	7
2.2. Cadenas de Markov. . . . .	8
2.3. Teoría de la información. . . . .	11
2.4. Series de tiempo financieras. . . . .	15
2.5. Objetivos. . . . .	17
<b>3. Metodología.</b>	<b>19</b>
3.1. Tiempos de recurrencia. . . . .	19
3.1.1. Tiempo de retorno. . . . .	21
3.1.2. Estimación de parámetros. . . . .	22
3.2. Protocolos de muestreo. . . . .	25
3.2.1. Prueba de control: tiempo de retorno. . . . .	26
3.2.2. Resultados experimento controlado para los distintos valores de $p$ . . . . .	28
3.3. <i>Matching times</i> . . . . .	33
3.3.1. Experimento controlado: cadena de Markov. . . . .	35
3.3.2. Experimento controlado: datos aleatorios con distribución normal. . . . .	38
3.4. Datos sustitutos y prueba de hipótesis. . . . .	39
3.5. Índice de asimetría desfasada. . . . .	40
3.6. Tratamiento de datos. . . . .	42
3.7. Patrones de tendencias. . . . .	44
3.7.1. Obtención de muestras. . . . .	45
3.8. Análisis de eficiencia. . . . .	46
<b>4. Resultados.</b>	<b>59</b>
4.1. Tiempos de recurrencia. . . . .	59
4.1.1. Tiempo de retorno. . . . .	59

---

4.1.2. <i>Matching times</i> . . . . .	60
4.2. Índice de asimetría desfasada. . . . .	66
4.3. Análisis de rachas para irreversibilidad e ineficiencia. . . . .	71
4.4. Comparación de resultados: índice de asimetría desfasada vs rachas. . . . .	77
4.5. Discusión de resultados. . . . .	78
<b>5. Conclusiones.</b>	<b>81</b>
<b>A. Apéndice.</b>	<b>83</b>
<b>Bibliografía</b>	<b>87</b>

# Capítulo 1

## Introducción.

El término de irreversibilidad, que proviene de la termodinámica, se refiere a los procesos que no pueden revertirse con el tiempo. En el siglo XIX, fue Rudolf Clausius quien introdujo la entropía, una medida que nos dice si un sistema está o no en equilibrio termodinámico [1]. Según [2], la irreversibilidad puede interpretarse también como el grado en que las propiedades estadísticas de las señales permanecen inalteradas incluso cuando se invierte la dirección del tiempo. La irreversibilidad es un concepto que suele surgir en los sistemas que se alejan del equilibrio termodinámico, es decir, aquellos sistemas en los que ya no es posible regresar al estado inicial. Este fenómeno puede observarse en una variedad de situaciones y procesos comunes. Algunos ejemplos claros de este fenómeno incluyen la combinación de diversas sustancias para formar una mezcla homogénea, en la cual los componentes originales ya no pueden ser separados fácilmente. Otro ejemplo típico es la transferencia de calor, que ocurre naturalmente desde una región de mayor temperatura hacia una región de menor temperatura, proceso que no puede revertirse en condiciones normales. Claramente, la irreversibilidad temporal nos da detalles microscópicos sobre qué tan lejos del equilibrio termodinámico está el proceso en juego [3]. Se ha aplicado recientemente en varios estudios, incluyendo secuencias de ADN [4], lo cual es relevante en el estudio del genoma humano. También se ha usado en el análisis de irreversibilidad en composiciones musicales aplicando física estadística, dinámica no lineal y teoría de grafos [2]. Además, se han llevado a cabo estudios en sistemas financieros. En [5], se utilizan algoritmos de visibilidad para medir la irreversibilidad de índices financieros. Este estudio también sugiere evaluar la relación entre la irreversibilidad y la predictibilidad de una forma cuantificable.

Generalmente, se considera que los precios pueden modelarse a través de una distribución lognormal y también, que siguen un proceso de Markov [6]. Del mismo modo, los rendimientos logarítmicos se ajustan a una distribución normal y se consideran indepen-

---

dientes e idénticamente distribuidos, lo que podría parecerse a una caminata aleatoria. De acuerdo con la teoría de la termodinámica, este comportamiento es similar a un proceso reversible y en el que no podemos determinar una dirección temporal para cada evento, ni predecir eventos futuros. En vista de lo anterior, dado que los precios son markovianos [6] y los rendimientos son independientes, ambos deberían cumplir con la hipótesis de los mercados eficientes. Eugene Fama presentó esta hipótesis en 1970 [7], proponiendo que los precios actuales reflejan el valor intrínseco de un activo. Esta teoría tiene tres versiones: débil, semi-fuerte y fuerte. La versión débil sugiere que el análisis técnico de los precios (análisis histórico) no ayuda a obtener rendimientos superiores a la media. La versión semi-fuerte, además de incluir la versión débil, indica que cualquier información pública sobre un activo se reflejará casi instantáneamente en su precio. Finalmente, la versión fuerte incluye las dos primeras y además sugiere que no es posible obtener rendimientos extraordinarios, incluso con información privada [8]. A lo largo de los años, diversos estudios han cuestionado y examinado esta teoría, demostrando que no siempre es válida [5, 9, 10]. Recientemente, estos estudios se han extendido al ámbito de las criptomonedas, obteniendo resultados similares [11–15].

La investigación de sistemas financieros en relación a la irreversibilidad temporal se fundamenta en la noción de que si un sistema es irreversible, se puede determinar una dirección temporal entre eventos, y por lo tanto, podría ser en cierto modo “predecible”. Esto contradice la hipótesis de mercados eficientes que se mencionó previamente.

Los sistemas financieros son una serie de instituciones responsables de impulsar y mantener la estabilidad económica de un país. Dentro de estos sistemas, encontramos los mercados financieros. Estos mercados, que pueden ser físicos o virtuales, son espacios donde se realizan intercambios de activos financieros y se establece su precio. Hay diversos tipos de mercados financieros algunos de ellos son el mercado de valores, mercado de bonos, mercado de derivados, mercado de divisas y el mercado de criptomonedas.

En este trabajo, nuestro enfoque principal será el mercado de criptomonedas que ha experimentado un crecimiento significativo en los últimos años. Este mercado gestiona el intercambio de dinero fiduciario, como pesos, dólares, euros, por activos digitales basados en la tecnología blockchain. Este tipo de monedas digitales no ha sido bien recibido por los gobiernos o instituciones bancarias debido a su naturaleza descentralizada. Esto significa que no existe ninguna entidad central o intermediario que regule o controle su uso. Las transacciones se realizan de manera *peer-to-peer*. Algunas de las principales criptomonedas (no estables) según su capitalización de mercado, como se cita en coinmarket [16], son Bitcoin (BTC), Ether (ETH), XRP Ledger (XRP), Dogecoin (DODGE), entre otros. Bitcoin, la primera y más reconocida criptomoneda, fue introducida en 2008 y posteriormente implementada en 2009 por una entidad conocida como Satoshi Nakamoto. Para entender

cómo funciona la red *blockchain*, se puede consultar el artículo publicado bajo el mismo nombre [17]. La creación de esta criptomoneda ha sido el catalizador para el surgimiento de innumerables otras. Con base en la información proporcionada por CoinMarketCap, un sitio web reconocido a nivel mundial por su seguimiento de precios de criptomonedas, en agosto de 2023, el universo criptográfico ha crecido hasta albergar a 9337 criptomonedas diferentes [16]. Este número refleja claramente el aumento de popularidad y aceptación de las criptomonedas a nivel global. Desde su creación, el precio de Bitcoin ha seguido un patrón de crecimiento que se podría describir prácticamente como exponencial, a pesar de haber experimentado caídas significativas en ciertos momentos. Esta criptomoneda, al igual que la mayoría de las criptomonedas, es conocida por su alta volatilidad, un factor que puede ser tanto un atractivo como un desafío para los inversores. Sin embargo, existen criptomonedas estables como DAI, USD Coin y USD Tether, que se caracterizan por mantener su precio en un nivel estable, basado en activos de referencia tales como el dólar estadounidense. Esto proporciona una opción para aquellos que buscan los beneficios de la criptomoneda sin la volatilidad asociada. Para ilustrar la evolución del mercado de criptomonedas, a continuación, se presenta un gráfico que muestra el comportamiento del precio de Bitcoin (Figura 1.1) desde el año 2011 y de Ether (Figura 1.2) desde el año 2016.



Figura 1.1: Precios del Bitcoin desde 2011. Figura tomada de [16].

En este estudio, nos enfocaremos en analizar las series de tiempo de algunas de las criptomonedas más destacadas por su capitalización de mercado, como Bitcoin, Ether, XRP Ledger, Litecoin y Bitcoin Cash. Luego, compararemos los resultados obtenidos. Como ya hemos señalado, Bitcoin [18] fue la primera criptomoneda en ser lanzada y sigue siendo la más relevante con una capitalización de mercado de \$529,640,380,704<sup>1</sup>.

Por otro lado, el Ether, lanzado en 2015, se ubica en la segunda posición con una capitalización de mercado de \$204,483,131,235 [19], y se distingue por no tener una cantidad

<sup>1</sup>Las capitalizaciones de mercado de todas las criptomonedas citadas se consultaron en agosto de 2023.



Figura 1.2: Precios del Ether desde 2016. Figura tomada de [16].

fija de moneda, a diferencia de Bitcoin. Además, según [ethereum.org](https://ethereum.org) [20], Ethereum no solo es programable, sino que también permite la construcción de aplicaciones descentralizadas. Esto significa que Ethereum no se limita a ser una simple red de pagos como Bitcoin. Se ha transformado en un mercado de servicios financieros completo, brindando apoyo a una amplia gama de usuarios, incluyendo inversores, desarrolladores, empresas, artistas, músicos y escritores, entre otros [20].

XRP Ledger [21], con una capitalización de \$27,868,057,510 y ocupando el quinto lugar, es una creación de David Schwartz, Jed McCaleb y Arthur Britto en 2011. Su objetivo era desarrollar un activo digital más sostenible que Bitcoin, optimizado específicamente para pagos. Lanzado en junio de 2012, era conocido inicialmente como Ripple. Se distingue por ser más rápido, escalable, sostenible y de coste más bajo que Bitcoin. Sin embargo, comparte una característica con Bitcoin: ambas monedas tienen una cantidad finita, siendo 100 mil millones para XRP y 21 millones para BTC [17, 21].

Litecoin [22] y Bitcoin Cash [23], aunque no tan prominentes como Bitcoin, ofrecen ciertas mejoras. Litecoin, lanzada en 2011, es un proyecto de software libre con un límite de 84 millones de monedas [22]. Se destaca por su velocidad, manejando más transacciones en el mismo tiempo que Bitcoin, además de permitir pagos instantáneos con costos cercanos a cero [22]. A menudo se la compara con la plata frente al oro de Bitcoin, según Andrew Torba [24] y Tara Mandjee [25].

Finalmente Bitcoin Cash, un proyecto de código abierto que se creó en agosto de 2017, es en realidad una bifurcación del Bitcoin original. Sus seguidores argumentaban que creían en la visión original de Bitcoin: un sistema de dinero electrónico peer-to-peer que cualquiera podría usar, con bajas comisiones [23].

La naturaleza volátil de estos activos ha llevado a su uso como un elemento de diversificación en la gestión de carteras, de acuerdo con [26], o como un activo puramente especulativo, según [27]. Esto nos lleva a cuestionarnos: ¿Son las series de tiempo de las criptomonedas predecibles y, en caso afirmativo, presentan un cierto grado de irreversi-

bilidad? ¿Cumplen estos activos financieros con la hipótesis de mercados eficientes? Y, finalmente, ¿es posible relacionar de algún modo la irreversibilidad con la ineficiencia del mercado? La investigación sobre la irreversibilidad en series temporales financieras es un campo emergente, lo que nos motiva a dedicar nuestro estudio a este tema.

En este análisis, exploraremos la irreversibilidad y la eficiencia en series temporales de las criptomonedas seleccionadas. Nuestro objetivo es descubrir posibles relaciones entre estos aspectos y responder a la pregunta: ¿si un mercado es irreversible, también es ineficiente?





# Capítulo 2

## Marco teórico.

En este capítulo, presentamos las herramientas que utilizamos en nuestra investigación. Para una comprensión más profunda, por favor, consulte [6, 28–32].

### 2.1. Procesos estocásticos.

Iniciaremos esta sección con la definición de proceso estocástico, un concepto clave en este trabajo.

**Definición 2.1.1.** *Un proceso estocástico es una colección de variables aleatorias  $\{X_t : t \in T\}$  parametrizada por un conjunto  $T$ , llamado espacio parametral, en donde las variables toman valores en un conjunto  $S$  llamado espacio de estados.*

En la definición anterior si el espacio parametral  $T$  es el conjunto discreto  $T = \{0, 1, 2, \dots\}$  entonces el proceso se considera un proceso a tiempo discreto. En otro caso el espacio parametral será el conjunto  $T = [0, \infty)$  y entonces se tendrá un proceso a tiempo continuo y se denota  $\{X_t : t \geq 0\}$ .

**Definición 2.1.2.** *Un proceso estocástico a tiempo continuo  $X_t : t \geq 0$  es estacionario estrictamente hablando si para cualesquiera tiempos  $t_1, \dots, t_n$  la distribución del vector  $(X_{t_1}, \dots, X_{t_n})$  es la misma que la del vector  $(X_{t_1+h}, \dots, X_{t_n+h})$  para cualquier valor de  $h > 0$ .*

Según la teoría financiera del paseo aleatorio, que Eugene Fama aborda en su artículo [33], esta teoría tiene una conexión directa con la hipótesis de los mercados eficientes. La teoría sugiere que los precios evolucionan siguiendo un camino aleatorio, mismo que definiremos a continuación.

**Definición 2.1.3.** Una caminata aleatoria simple sobre el conjunto de los enteros  $\mathbb{Z}$  es un proceso estocástico a tiempo discreto  $\{X_n : n = 0, 1, \dots\}$  tal que  $X_0 = 0$  y para cada  $n \in \mathbb{N}$  se tiene que

$$\mathbb{P}(X_{n+1} = j | X_n = i) = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = i - 1, \\ 0 & \text{otro caso.} \end{cases} \quad (2.1)$$

De la definición anterior podemos observar que si un proceso estocástico sigue una caminata aleatoria, entonces el valor que toma el proceso al tiempo  $n$  depende únicamente del valor tomado por dicho proceso en el tiempo inmediato anterior, es decir al tiempo  $n - 1$ , con lo cual se cumple la siguiente propiedad.

## 2.2. Cadenas de Markov.

**Propiedad 2.2.1** (de Markov). Sea  $\{X_n : n = 0, 1, \dots\}$  un proceso estocástico a tiempo discreto. Se dice que este proceso satisface la propiedad de Markov si para cada  $n \in \mathbb{N}$  y para cualesquiera enteros  $x_0, x_1, \dots, x_{n+1}$  se tiene que

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n). \quad (2.2)$$

**Definición 2.2.2.** Una cadena de Markov es un proceso estocástico a tiempo discreto  $\{X_n : n = 0, 1, \dots\}$  con espacio de estados discreto, y que satisface la propiedad de Markov.

De aquí podemos concluir que cualquier caminata simple sobre  $\mathbb{Z}$  es una cadena de Markov a tiempo discreto. También podemos observar que cualquier proceso que cumpla la propiedad de Markov, es consistente con la hipótesis de mercados eficientes propuesta por Eugene Fama en 1970 [7].

**Definición 2.2.3.** Sean  $i$  y  $j$  dos estados de una cadena de Markov. A la probabilidad de pasar del estado  $i$  al estado  $j$  al tiempo  $n$  y  $n + 1$  se le denota  $p_{ij}$  y se define como

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i), \quad (2.3)$$

**Definición 2.2.4.** Se define la matriz de transición como

$$P = \begin{pmatrix} p_{00} & p_{01} & \cdots & p_{0n} \\ p_{10} & p_{11} & \cdots & p_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n0} & p_{n1} & \cdots & p_{nn} \end{pmatrix}, \quad (2.4)$$

donde  $p_{ij}$  es la probabilidad de pasar del estado  $i$  al estado  $j$  en un paso.

**Proposición 2.2.5.** La matriz de probabilidades de transición  $P = (p_{ij})$  cumple las siguientes propiedades.

a)  $p_{ij} \geq 0$ .

b)  $\sum_j p_{ij} = 1$ .

**Definición 2.2.6.** Toda matriz cuadrada que cumpla las propiedades dadas en la proposición 2.2.5 se dice que es una matriz estocástica. Si además la matriz satisface la condición  $\sum_i p_{ij} = 1$  entonces se dice que la matriz es doblemente estocástica.

**Proposición 2.2.7.** La probabilidad de transición en  $n$  pasos,  $p_{ij}(n)$ , está dada por la entrada  $(i, j)$  de la  $n$ -ésima potencia de la matriz  $P$ , esto es,

$$p_{ij} = (P^n)_{ij}. \quad (2.5)$$

A continuación daremos algunas definiciones que serán necesarias en resultados más adelante. Notemos que una cadena de Markov establece la dinámica en el conjunto de las distribuciones de probabilidad definidas sobre el espacio de estados de dicha cadena [28].

Sea  $S = \{0, 1, 2, \dots, n\}$  el espacio finito de estados de una cadena de Markov y sea  $\pi^0$  la distribución de probabilidad inicial, definida como

$$\pi^0 = (\pi_0^0, \pi_1^0, \dots, \pi_n^0), \quad (2.6)$$

donde  $\pi_i^0$  es la probabilidad de inicializar la cadena en el estado  $i$  al tiempo cero. Entonces después de transcurrida la primera unidad de tiempo, la cadena se encuentra en alguno de sus  $n$  posibles estados, siguiendo el vector de distribución

$$\pi^1 = (\pi_0^1, \pi_1^1, \dots, \pi_n^1), \quad (2.7)$$

en donde

$$\pi_j^1 = \mathbb{P}(X_1 = j) = \sum_{i=0}^n P(X_1 = j | X_0 = i) \mathbb{P}(X_0 = i) = \sum_{i=0}^n \pi_i^0 p_{ij} = \sum_{i=0}^n \pi_i^0 p_{ij} = \pi^0 P_j, \quad (2.8)$$

en donde  $P_j$  es la columna  $j$  de la matriz  $P$ . En general, para  $m \geq 1$ , obtenemos una sucesión de distribuciones  $\pi = (\pi^0, \pi^1, \dots)$ , en donde

$$\pi^m = \pi^{m-1} P = \pi^0 P^m. \quad (2.9)$$

**Definición 2.2.8.** Una distribución de probabilidad  $\pi = (\pi^0, \pi^1, \dots)$  es estacionaria o invariante si para una cadena de Markov con matriz de probabilidades de transición  $P = (p_{ij})$  si

$$\pi^m = \sum_i \pi^{m-1} p_{ij}. \quad (2.10)$$

En términos matriciales, la distribución de probabilidad  $\pi$ , es estacionaria si  $\pi = \pi P$ .

En 1828, Robert Brown notó movimientos irregulares en las partículas de polen al observarlas bajo el microscopio. Sin embargo, no fue hasta 1923 que Norbert Wiener logró demostrar la existencia y unicidad de un proceso que se alineaba con lo que Brown había observado. Hoy en día, este proceso es conocido como movimiento browniano o proceso de Wiener. En general, se prefiere usar el primer término para describir el fenómeno físico, mientras que el segundo se usa más para referirse al modelo matemático [28]. Este proceso también tiene aplicaciones en el campo de las finanzas, particularmente en la valoración de opciones. Esto se realiza a través de la ecuación de Black-Scholes-Merton, que asume que los precios de los activos siguen un proceso generalizado de Wiener. Según este modelo, se asume que los precios de los activos financieros siguen una distribución lognormal, mientras que su tasa de rendimiento seguirá una distribución normal [6]. A continuación, definiremos formalmente los procesos que hemos mencionado anteriormente.

**Definición 2.2.9.** Un movimiento browniano unidimensional de parámetro  $\sigma^2$  es un proceso estocástico  $\{B_t : t \geq 0\}$  con valores en  $\mathbb{R}$  que cumple las siguientes propiedades.

1.  $B_0 = 0$  casi seguramente (c.s).
2. Las trayectorias  $t \mapsto B_t$  son continuas.
3. El proceso tiene incrementos  $B_t - B_s$  independientes.

4. Para cualesquiera tiempos  $0 \leq s < t$ , la variable incremento  $B_t - B_s$  tiene distribución  $N(0, \sigma^2(t - s))$ .

Se dice que el movimiento browniano es estándar cuando  $\sigma^2 = 1$ .

El proceso anterior nos ayudará a definir el siguiente concepto.

**Definición 2.2.10.** Sean  $\mu$  y  $\sigma > 0$  constantes, se define el movimiento browniano geométrico (o proceso de Wiener generalizado), como el proceso  $\{X_t : t \geq 0\}$  solución de la ecuación estocástica

$$dX_t = \mu X_t dt + \sigma X_t dB_t. \quad (2.11)$$

Obsérvese que si  $X_0 = x_0$  y  $x_0 > 0$ , entonces

$$X_t = x_0 e^{[(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t]}, \quad (2.12)$$

es solución de la ecuación 2.11.

### 2.3. Teoría de la información.

Como ya mencionamos, es de suma importancia estudiar la irreversibilidad de un sistema ya que si una serie temporal irreversible, existe la posibilidad de realizar predicciones sobre los precios futuros. Es importante recordar que cuando una serie temporal es irreversible, se puede establecer una dirección temporal clara, lo que nos permite discernir entre los eventos pasados y los eventos futuros. Esta distinción es fundamental para el análisis técnico, que implica el estudio de los precios históricos para realizar proyecciones sobre el comportamiento futuro del mercado. Esto contradice la hipótesis de los mercados eficientes. Sin embargo, queremos destacar la teoría propuesta por Andrew Lo, la hipótesis de los mercados adaptativos, que sugiere que los mercados no son siempre eficientes, por lo que hay momentos en los que es posible hacer predicciones en el mercado [34]. Pueden consultar un estudio sobre la predictibilidad en los rendimientos en [35].

A continuación, vamos a definir lo que es entropía. Esencialmente, es una forma de medir la incertidumbre de una variable aleatoria. Para mayor detalle de todos los resultados presentados en esta sección véase [30]. La demostración de los resultados presentados en esta sección pueden consultarse en el capítulo A.

**Definición 2.3.1.** Sea  $X$  una variable aleatoria discreta con espacio de estados  $S$  y función masa de probabilidad  $p(x) = P\{X = x\}$ ,  $x \in S$ . Se define la entropía de Shannon  $H(X)$  como

$$H(X) = -\mathbb{E}[\ln(p(X))]. \quad (2.13)$$

En particular cuando  $X$  es discreta, la entropía de Shannon será

$$H(X) = -\sum_x p(x) \ln(p(x)), \quad (2.14)$$

mientras que, para el caso continuo se define como

$$H(X) = -\int_{-\infty}^{\infty} p(x) \ln(p(x)) dx. \quad (2.15)$$

**Definición 2.3.2.** Se define la entropía conjunta  $H(X_1, X_2, \dots, X_n)$  de un vector aleatorio de variables discretas  $(X_1, X_2, \dots, X_n)$  con distribución de probabilidad conjunta  $p(x_1, x_2, \dots, x_n)$  como

$$H(X_1, X_2, \dots, X_n) = -\sum_{x_1 \in S_1} \sum_{x_2 \in S_2} \cdots \sum_{x_n \in S_n} p(x_1, x_2, \dots, x_n) \ln p(x_1, x_2, \dots, x_n), \quad (2.16)$$

o lo que es lo mismo,

$$H(X_1, X_2, \dots, X_n) = -\mathbb{E}[\ln p(X_1, X_2, \dots, X_n)]. \quad (2.17)$$

**Definición 2.3.3.** Si  $(X, Y) \sim p(x, y)$ , la entropía condicional  $H(Y|X)$  se define como

$$\begin{aligned} H(Y|X) &= \sum_{x \in S_X} p(x) H(Y|X = x) \\ &= -\sum_{x \in S_X} p(x) \sum_{y \in S_Y} p(y|x) \ln p(y|x) \\ &= -\sum_{x \in S_X} \sum_{y \in S_Y} p(x, y) \ln p(y|x) \\ &= -\mathbb{E}[\ln p(Y|X)]. \end{aligned} \quad (2.18)$$

Definimos la entropía relativa, la cual nos ayuda para medir la cercanía entre dos distribuciones, asimismo mide la ineficiencia de asumir que la distribución de una variable aleatoria es  $q$  cuando la verdadera distribución es  $p$ .

**Definición 2.3.4.** La entropía relativa o divergencia de Kullback-Leibler entre dos funciones de probabilidad  $p(x)$  y  $q(x)$  se define como,

$$D(p||q) = \sum_{x \in S} p(x) \ln \frac{p(x)}{q(x)} = \mathbb{E}_p \left[ \ln \frac{p(X)}{q(X)} \right]. \quad (2.19)$$

Cabe señalar que la divergencia Kullback-Leibler no puede entenderse como la distancia entre dos distribuciones, pues ésta no cumple la propiedad de simetría, es decir

$$D(p||q) \neq D(q||p). \quad (2.20)$$

**Definición 2.3.5.** *Un proceso estocástico es reversible si se cumple que,*

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_n, X_2 = x_{n-1}, \dots, X_n = x_1). \quad (2.21)$$

*Si no se cumple la ecuación anterior, se dice que el proceso es irreversible.*

A continuación definimos la tasa de producción de entropía para un proceso estocástico, la cual nos ayudará a calcular de manera teórica la irreversibilidad de una serie de tiempo.

**Definición 2.3.6.** *Sea  $\{X_j : j \in \mathbb{N}\}$  un proceso estocástico. La tasa de entropía  $h$  se define como,*

$$h = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n), \quad (2.22)$$

*mientras que la tasa de entropía reversa  $h_R$  está dada por*

$$h_R = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_n, X_{n-1}, \dots, X_1). \quad (2.23)$$

**Teorema 2.3.7.** *Sea  $\{X_j : j \in \mathbb{N}\}$  un proceso estocástico estacionario, entonces*

$$\lim_{n \rightarrow \infty} \frac{H(X_1, X_2, \dots, X_n)}{n} = \lim_{n \rightarrow \infty} H(X_n | X_{n-1}, \dots, X_1). \quad (2.24)$$

**Proposición 2.3.8.** *Sea  $\{X_j : j \in \mathbb{N}\}$  una cadena de Markov con distribución estacionaria  $\pi$  y matriz de transición  $P$ . Suponiendo que el proceso inicia en la distribución estacionaria, entonces la tasa de entropía  $h$  está dada por*

$$h = - \sum_{i,j} \pi_i p_{ij} \ln p_{ij}, \quad (2.25)$$

*mientras que la tasa de entropía reversa  $h_R$ , correspondiente a la tasa de entropía obtenida cuando el proceso ocurre en sentido inverso, es decir  $X_0 = x_n \dots X_n = x_0$  está dada por*

$$h_R = - \sum_{i,j} \pi_i p_{ij} \ln p_{ji}. \quad (2.26)$$

*Demostración.* La tasa de entropía  $h$  para una cadena de Markov estacionaria se define como,

$$\begin{aligned} h &= \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) \\ &= \lim_{n \rightarrow \infty} H(X_n | X_{n-1}, \dots, X_1) \\ &= \lim_{n \rightarrow \infty} H(X_n | X_{n-1}) \\ &= H(X_2 | X_1), \end{aligned}$$

donde  $H(X_2 | X_1) = \sum_i \pi_i \left( - \sum_j P_{ij} \ln P_{ij} \right)$ . Análogamente

$$h_R = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_n, X_{n-1}, \dots, X_1) \quad (2.27)$$

$$= \lim_{n \rightarrow \infty} H(X_1 | X_2, \dots, X_n) \quad (2.28)$$

$$= \lim_{n \rightarrow \infty} H(X_1 | X_2) \quad (2.29)$$

$$= H(X_1 | X_2), \quad (2.30)$$

donde  $H(X_1 | X_2) = \sum_i \pi_i \left( - \sum_j P_{ji} \ln P_{ji} \right)$ .  $\square$

**Definición 2.3.9.** Definimos la producción de entropía  $e_p$  para un proceso estocástico estacionario  $\{X_n : n \in \mathbb{N}\}$  como

$$e_p = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[ \ln \frac{\mathbb{P}(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)}{\mathbb{P}(X_1 = x_n, X_2 = x_{n-1}, \dots, X_n = x_1)} \right]. \quad (2.31)$$

Como podemos observar de la ecuación 2.31, si  $e_p = 0$  el proceso es reversible, mientras que si  $e_p \neq 0$  el proceso es irreversible.

A continuación se define la tasa de producción de entropía para el caso de cadenas de Markov.

**Teorema 2.3.10.** Sea  $X = \{X_n : n = 0, 1, \dots\}$  una cadena de Markov, estacionaria, con espacio de estados  $S$  numerable. Sea  $P = (p_{ij})_{i,j \in S}$  la matriz de transición y una distribución única e invariante  $\pi = (\pi_i)_{i \in S}$  entonces la tasa de producción de entropía para una cadena de Markov puede escribirse como

$$e_p = \frac{1}{2} \sum_{i,j \in S} (\pi_i p_{ij} - \pi_j p_{ji}) \log \frac{\pi_i p_{ij}}{\pi_j p_{ji}}. \quad (2.32)$$

Del resultado anterior se puede obtener lo siguiente.



**Corolario 2.3.11.** *Sea  $h$  la tasa de entropía directa para una cadena de Markov estacionaria y sea  $h_R$  la tasa de entropía reversa, correspondiente a la tasa de entropía obtenida cuando el proceso ocurre en sentido inverso entonces*

$$e_p = h_R - h. \quad (2.33)$$

A continuación daremos un ejemplo, a manera de ilustrar el cálculo de la entropía para una cadena de Markov, mismo que puede encontrarse en [36].

Sea  $\{X_t \in S : t \in \mathbb{N}\}$  un proceso estocástico discreto con espacio de estados  $S = \{1, 2, 3\}$ . Se define el proceso como una cadena de Markov de tres estados (véase la Figura 2.1) tal que  $M : S \times S \rightarrow [0, 1] \subset \mathbb{R}$  definida como

$$M = \begin{pmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 1-p & 0 \end{pmatrix}, \quad (2.34)$$

donde  $p$  es un parámetro tal que  $p \in [0, 1]$ . Podemos observar que la matriz es doblemente estocástica, con lo cual el vector de distribución invariante  $\pi = \pi M$ , está dado por  $\pi = (1/3, 1/3, 1/3)$ . Calcularemos la tasa de entropía de dicha cadena.

$$e_p = \frac{1}{6} \left[ ((p - (1-p)) \log \frac{p}{(1-p)} + ((1-p) - p) \log \frac{(1-p)}{p} + \right. \quad (2.35)$$

$$\left. ((1-p) - p) \log \frac{(1-p)}{p} + (p - (1-p)) \log \frac{p}{(1-p)} + \right. \quad (2.36)$$

$$\left. (p - (1-p)) \log \frac{p}{(1-p)} + ((1-p) - p) \log \frac{(1-p)}{p} \right]. \quad (2.37)$$

Es fácil ver que la tasa de entropía después de un poco de álgebra, está dada por

$$e_p = (2p - 1) \log \left( \frac{p}{1-p} \right). \quad (2.38)$$

## 2.4. Series de tiempo financieras.

También definimos el concepto de una serie temporal, un elemento clave de este trabajo. Una serie temporal se puede describir como un conjunto de observaciones, cada una obtenida en un momento específico  $t$  [37].

Empezaremos por definir diferentes tipos de rendimientos, el cual es la ganancia obtenida al realizar una inversión. Antes de proceder a los análisis, convertimos los precios

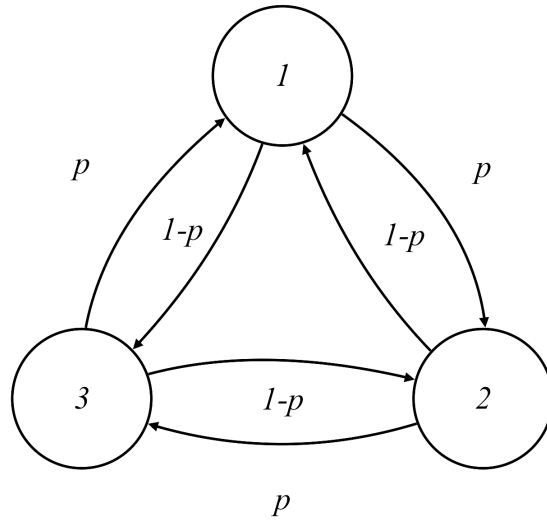


Figura 2.1: Cadena de Markov a tres estados.

a rendimientos logarítmicos. Observaremos que los rendimientos de precios se expresan como un producto, mientras que los rendimientos logarítmicos se pueden expresar como una suma. Estos tienen características estadísticas favorables. Si asumimos que los precios siguen una distribución lognormal, entonces los rendimientos seguirán una distribución normal. Además, si asumimos que estos rendimientos son independientes, la aplicación del teorema del límite central se simplifica. En las siguientes definiciones, haremos uso de la palabra activo. Un activo es un título que otorga al comprador el derecho de recibir un ingreso futuro. Para las siguientes definiciones sea  $P_t$  el precio de un activo al tiempo  $t$ .

**Definición 2.4.1.** *Rendimiento simple (periodo 1).*

*Si se mantiene la posición de un activo por un periodo desde el tiempo  $t-1$  al tiempo  $t$  se tendrá el rendimiento bruto simple  $R_t$  como sigue:*

$$R_t = \frac{P_t}{P_{t-1}} - 1, \quad (2.39)$$

o bien

$$P_t = P_{t-1}(1 + R_t). \quad (2.40)$$

A continuación generalizamos la definición anterior.

**Definición 2.4.2.** *Rendimiento simple multiperiodo.*

*Si se mantiene la posición de un activo por  $k$  periodos entre el tiempo  $t-k$  y el tiempo  $t$  se define el rendimiento bruto simple a  $k$ -periodos como sigue:*

$$\begin{aligned}
1 + R_t[k] &= \frac{P_t}{P_{t-k}} = \frac{P_t}{P_{t-1}} \times \frac{P_{t-1}}{P_{t-2}} \times \dots \times \frac{P_{t-k+1}}{P_{t-k}} \\
&= (1 + R_t)(1 + R_{t-1}) \dots (1 + R_{t-k+1}) \\
&= \prod_{j=0}^{k-1} (1 + R_{t-j}). \tag{2.41}
\end{aligned}$$

Daremos un ejemplo intuitivo para introducir el siguiente concepto. Suponga que se depositan 10 pesos en un banco y se gana una tasa de interés del 10 % anual. Entonces después de un año, se tendrían  $10 + 10(0.1) = 10(1 + 0.1) = 11$  pesos. En otro caso, si la tasa de interés se paga semestralmente, después de un semestre se tendrían  $10 + 10(0.05) = 10(1 + 0.05) = 10.5$  pesos y finalmente después del primer año se tendrían  $10.5 + 10.5(0.05) = 10(1.05)(1.05) = 10(1 + \frac{0.1}{2})^2 = 11.025$  pesos. Siguiendo el mismo razonamiento, después de  $m$  periodos de tiempo, la cantidad obtenida sería de  $10(1 + \frac{0.1}{m})^m$  pesos. Con esto podemos observar que, si hacemos los periodos de tiempo cada vez más pequeños, la cantidad inicial crece más rápido, es decir obtendremos un rendimiento continuo.

**Definición 2.4.3.** *Rendimiento compuesto continuamente.*

*El rendimiento logarítmico  $\ell$  de un activo es llamado rendimiento compuesto continuamente y se define como*

$$\ell = \ln(1 + R_t) = \ln \frac{P_t}{P_{t-1}}.$$

## 2.5. Objetivos.

Este estudio se ha llevado a cabo con el objetivo de confirmar o refutar nuestra hipótesis inicial. Esta hipótesis plantea la posibilidad de que exista una relación entre la ineficiencia del mercado y la irreversibilidad de la serie temporal de dicho mercado. Esta relación podría ser crucial para entender cómo funcionan los mercados y cómo se pueden prever sus movimientos. En el capítulo 3, detallamos los diferentes análisis que hemos llevado a cabo para probar esta hipótesis. Estos análisis se han basado en los resultados que se presentan en este capítulo. Posteriormente, en el capítulo 4, presentamos los resultados que hemos obtenido de los análisis realizados en las series de tiempo de las criptomonedas.



# Capítulo 3

## Metodología.

En este capítulo, abordaremos los procedimientos que seguimos para cada análisis realizado durante nuestra investigación. Estos análisis se llevaron a cabo con el objetivo de buscar un índice de irreversibilidad y, por lo tanto, detectar el grado de irreversibilidad en las series de tiempo financieras, específicamente en las series de tiempo de BTC [18], ETH [20], XRP Ledger [21], LTC [22] y BCH [23]. Los datos se obtuvieron de *CryptoDataDownload* [38], que proporciona datos históricos de algunos de los mercados digitales (*exchanges*) más populares, con diferentes frecuencias de muestreo. Las bases de datos tienen nueve columnas: tiempo unix<sup>1</sup>, fecha, símbolo, precios de apertura, más alto, más bajo, cierre, volumen de venta de la criptomoneda y volumen (cantidad) en dólares. Consulte la Figura 3.1.

En este proyecto, abordamos métodos como los tiempos de recurrencia, el índice de asimetría desfasada, y el análisis de patrones de tendencia. Además, describimos la estrategia implementada para manejar los datos faltantes, un problema frecuente en la mayoría de las bases de datos con las que trabajamos. Como último paso, aplicamos el método propuesto por Brouty y Garcin para determinar la ineficiencia de una serie de tiempo [39].

### 3.1. Tiempos de recurrencia.

Recordemos que utilizamos la tasa de producción de entropía  $e_p$  para medir el grado de irreversibilidad de un proceso estocástico. De la ecuación 2.31 podemos observar que teóricamente, la irreversibilidad se obtiene mediante un límite haciendo  $n$  tender a infinito. Por lo general, en las series de tiempo reales, es imposible tener una gran cantidad de datos, o bien, tener varias muestras del mismo suceso. Lo anterior imposibilita la tarea de

---

<sup>1</sup>El tiempo unix es utilizado en muchos sistemas computacionales y se define como el número de segundos transcurridos desde el 1 de enero de 1970 a la fecha en cuestión.

unix	date	symbol	open	high	low	close	Volume BTC	Volume USD
1420070460	01/01/2015 00:01	BTC/USD	321	321	321	321	0	0
1420070520	01/01/2015 00:02	BTC/USD	321	321	321	321	1.41333333	453.679999
1420070580	01/01/2015 00:03	BTC/USD	321	321	321	321	2	642
1420070640	01/01/2015 00:04	BTC/USD	320.98	320.98	320.43	320.43	2.84388	911.264468
1420070700	01/01/2015 00:05	BTC/USD	320.5	320.51	320.49	320.5	8.255	2645.7275
1420070760	01/01/2015 00:06	BTC/USD	321	321	320.52	320.52	1.44731	463.891801
1420070820	01/01/2015 00:07	BTC/USD	320.54	320.54	320.54	320.54	0.312	100.00848
1420070880	01/01/2015 00:08	BTC/USD	320.54	320.54	320.54	320.54	0	0
1420070940	01/01/2015 00:09	BTC/USD	320.34	320.34	319.94	319.94	0.73234	234.30486
1420071000	01/01/2015 00:10	BTC/USD	319.94	319.94	319.94	319.94	0	0
1420071060	01/01/2015 00:11	BTC/USD	319.94	319.94	319.94	319.94	0	0
1420071120	01/01/2015 00:12	BTC/USD	319.67	319.67	319.67	319.67	0.0165	5.274555
1420071180	01/01/2015 00:13	BTC/USD	319.67	319.67	319.67	319.67	0	0
1420071240	01/01/2015 00:14	BTC/USD	319.7	319.7	319.7	319.7	0.0165	5.27505
1420071300	01/01/2015 00:15	BTC/USD	319.7	319.7	319.7	319.7	0	0
1420071360	01/01/2015 00:16	BTC/USD	319.7	319.7	319.7	319.7	0	0
1420071420	01/01/2015 00:17	BTC/USD	319.54	319.54	319.54	319.54	0.0165	5.27241
1420071480	01/01/2015 00:18	BTC/USD	319.54	319.54	319.54	319.54	0	0
1420071540	01/01/2015 00:19	BTC/USD	319.57	319.57	319.57	319.57	0.0638	20.388566
1420071600	01/01/2015 00:20	BTC/USD	319.57	319.57	319.57	319.57	0	0
1420071660	01/01/2015 00:21	BTC/USD	319.6	320	319.6	320	0.45898142	146.874054
1420071720	01/01/2015 00:22	BTC/USD	320.08	320.08	320.08	320.08	0.2504	80.148032
1420071780	01/01/2015 00:23	BTC/USD	320.08	320.08	320.08	320.08	0	0

Figura 3.1: Base de datos de precios minuto a minuto del Bitcoin desde enero de 2015, extraída de [38]. Se tienen nueve columnas en las que encontramos el tiempo unix, la fecha con horas y minutos, el símbolo de la criptomoneda, los precios de apertura, más alto, más bajo, cierre y los volúmenes de la criptomoneda y en dólares.

obtener el grado de irreversibilidad de dichas series temporales. conocidos como tiempos de recurrencia. Estos métodos nos permiten obtener estimaciones para la tasa de producción de entropía. Entre estos tiempos de recurrencia, encontramos el tiempo de retorno y el *matching time*, que se pueden consultar en las referencias [3] y [36] respectivamente.

Como mencionamos anteriormente, ya que las series temporales disponibles son finitas, podemos considerar que estas series están truncadas hasta cierto punto. Esto sugiere la posibilidad de obtener una estimación empírica no óptima de la tasa de entropía. Los tiempos de retorno ofrecen una solución a este problema, ajustando las estimaciones para obtener un valor más preciso. A continuación, presentamos la metodología para calcular la tasa de producción de entropía de una serie temporal utilizando los tiempos de retorno.

**Definición 3.1.1.** Consideremos un conjunto finito de símbolos  $A$  al que llamamos alfabeto. Sea  $X := \{X_n : n \in \mathbb{N}\}$  un proceso estacionario, cuyas realizaciones son sucesiones infinitas de símbolos tomados del conjunto  $A$ . Una realización infinita del proceso se denota como  $x = x_1x_2x_3 \dots$  con  $x_i \in A \forall i$  y  $x \in A^{\mathbb{N}}$ . Denotamos como  $x_k^{k+\ell-1}$  a la secuencia de símbolos de tamaño  $\ell$ , con  $\ell$  un entero positivo, que comienza en el  $k$ -ésimo lugar de la realización  $x$ .

**Definición 3.1.2.** Una secuencia finita de  $\ell$  símbolos tomados del alfabeto  $A$ ,  $a = a_1a_2a_3 \dots a_\ell$ , se llama una palabra o bloque de tamaño  $\ell$ . Definimos también la secuencia finita inversa  $\bar{a}$ , de  $\ell$  símbolos tomados del alfabeto  $A$  como  $\bar{a} = a_\ell a_{\ell-1} \dots a_2 a_1$

La ocurrencia de la palabra de tamaño  $\ell$ , puede ser observada a lo largo de la realización del proceso en el  $k$ -ésimo lugar, para ello se utilizará la siguiente notación:

$$x_k^{k+\ell-1} = a, \quad (3.1)$$

mientras que  $\bar{a}$  denota la ocurrencia de la palabra de tamaño  $\ell$ , en sentido inverso, es decir:

$$x_k^{k+\ell-1} = \bar{a}. \quad (3.2)$$

### 3.1.1. Tiempo de retorno.

**Definición 3.1.3.** *Definimos el tiempo de retorno  $\rho_\ell$  como:*

$$\rho_\ell(x) := \inf\{k > 1 : x_k^{k+\ell-1} = x_1^\ell\}. \quad (3.3)$$

Lo anterior significa que realizaremos una búsqueda de la palabra de longitud  $\ell$  a lo largo de la serie de tiempo y el tiempo de retorno  $\rho$  nos proporciona el lugar  $k$  más pequeño en dónde dicha palabra es encontrada.

De acuerdo con [40], Ornstein y Weiss se tiene el siguiente resultado

$$h = \lim_{\ell \rightarrow \infty} \frac{1}{\ell} \ln(\rho_\ell), \quad \mathbb{P} - \text{c.s.}, \quad (3.4)$$

el cual se basa en el teorema de Shannon-McMillan-Breiman, que dice que dado el conjunto  $[x_1^\ell] = \{z \in A^{\mathbb{N}} : z_1^\ell = x_1^\ell\}$  entonces

$$h = \lim_{\ell \rightarrow \infty} -\frac{1}{\ell} \log \mathbb{P}([x_1^\ell]), \quad \mathbb{P} - \text{c.s.} \quad (3.5)$$

Por lo anterior, podemos inferir que la entropía puede ser estimada mediante el límite del tiempo de retorno de la definición 3.1.3, cuando  $\ell$  tiende a infinito. Con lo anterior podemos notar que, si deseamos calcular la irreversibilidad de una serie de tiempo debido a que únicamente contamos con una sola trayectoria finita, estaremos realizando un truncamiento hasta un tiempo  $N$ , es decir

$$x = x_1, x_2, \dots, x_N, \quad (3.6)$$

será la serie de tiempo con la que contamos, de tal manera que, el tiempo de recurrencia máximo que podremos establecer corresponde a  $N - \ell + 1$ , ya que si esta cota es superada, el bloque no es encontrado totalmente dentro de la serie de tiempo y dicha muestra quedará indeterminada (véase la Figura 3.2).

Kontoyiannis en la referencia [41] obtiene una aproximación fuerte para el tiempo de recurrencia y el tiempo de espera, lo que le permitió obtener la convergencia para tiempo

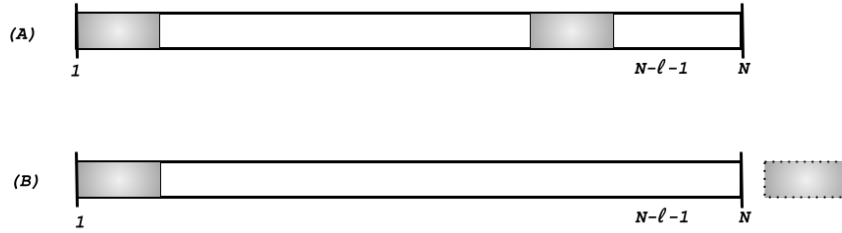


Figura 3.2: Búsqueda del tiempo de retorno en la serie de tiempo. En el caso (A) el bloque de tamaño  $\ell$  es encontrado dentro de la serie de tiempo, por lo tanto el tiempo de retorno está bien definido. En el caso (B) bloque no es encontrado dentro de la serie de tiempo por lo tanto el tiempo de retorno queda numéricamente indeterminado.

de espera en procesos  $\psi$ -mixing, finalmente extendiendo estos resultados a cadenas de Markov. También obtuvo un principio de invarianza de convergencia casi segura para  $\log \rho_\ell$ , con lo cual se satisface un teorema de límite central y una ley de logaritmos iterados para procesos estacionarios de valores finitos, incluyendo las cadenas de Markov estacionarias irreducibles y aperiódicas. Por lo anterior podemos suponer  $\ln \rho_\ell$  satisface un teorema central del límite, por lo que para  $\ell$  suficientemente grande, la distribución de  $\ln(\rho_\ell)/\ell$  puede ser considerada normal y cuya media y varianza deberán ser estimadas. A continuación se detalla la estimación de los parámetros antes mencionados, misma que puede ser consultada en [3].

### 3.1.2. Estimación de parámetros.

Sea  $\Theta_\ell$  la variable aleatoria cuyas realizaciones son estimaciones de la tasa de entropía de un bloque de tamaño  $\ell$  obtenidas por medio del tiempo de retorno, es decir:

$$\Theta_\ell = \frac{1}{\ell} \ln(\rho_\ell), \quad (3.7)$$

dónde  $\Theta_\ell$  se considera normal, con media  $h_\ell$  y varianza  $\sigma_\ell^2$ . Dichos parámetros deberán ser estimados. Definimos ahora el conjunto

$$\mathcal{H} = \{h_i : 1 \leq i \leq m\}, \quad (3.8)$$

como el conjunto de  $m$  muestras de realizaciones de la variable  $\Theta_\ell$ . Obsérvese que dicho conjunto contiene las estimaciones de la tasa de entropía utilizando el tiempo de retorno



de los bloques de tamaño  $\ell$ . Recordemos que el tiempo de retorno podía quedar numéricamente indeterminado si el bloque no se encuentra dentro de la serie de tiempo. Dado lo anterior, es claro ver que la estimación de la tasa de producción de entropía para dicho bloque quedará indeterminada también. De lo anterior tendemos la siguiente definición.

**Definición 3.1.4.** *Se dice que la variable  $h_i$  es una muestra censurada si el valor de  $h_i$  es numéricamente indeterminado. De lo contrario, se dice que  $h_i$  es no censurada.*

Dado que el tiempo de retorno siempre está acotado por arriba, de aquí en adelante a dicha cota la llamaremos tiempo de censura y la denotaremos como  $T_c$ . Asimismo la estimación de la tasa de entropía queda acotada por arriba por

$$h_c = \frac{\ln T_c}{\ell}. \quad (3.9)$$

Conviene establecer un  $T_c$  que permita tener uniformidad en la obtención de las muestras de tiempos de retorno; más adelante ahondaremos más en este detalle. En la Figura 3.3, podemos observar que en la búsqueda del tiempo de retorno se puede presentar cualquiera de los siguientes escenarios: el tiempo de retorno esté definido (muestra no censurada), es decir  $\rho_\ell = t$  (Caso (A)), o bien, el tiempo de retorno nunca es encontrado (muestra censurada), con lo que se tiene que  $\rho_\ell > k + n$  o bien,  $\rho_\ell > 2n$  (casos (B) y (C) respectivamente).

Supongamos que se tienen  $k$  muestras no censuradas, con  $k < m$ . Por lo tanto el número de muestras censuradas es  $m - k$ . Asumiendo independencia en las realizaciones, entonces la variable aleatoria  $k$  que cuenta el número de muestras no censuradas tiene distribución binomial. De esta manera,  $\hat{p} = k/m$  es el estimador para la tasa de muestras no censuradas.

Sea  $\mathcal{M}_\ell^\rho \{a_i : a_i = x(i\Delta + 1, i\Delta + \ell), 0 \leq i < m\}$  el conjunto de palabras de longitud  $\ell$ , espaciadas por una cantidad  $\Delta$  y de cardinalidad  $m$  con  $m = \lfloor \frac{n}{\Delta} \rfloor$  y  $n = N/2$ . Luego los siguientes conjuntos

$$\mathcal{R}_\ell = \{t \in \mathbb{N} : \rho_\ell^n(a) = t, a \in \mathcal{M}_\ell^\rho\}, \quad (3.10)$$

y

$$\bar{\mathcal{R}}_\ell = \{t \in \mathbb{N} : \rho_\ell^n(\bar{a}) = t, \bar{a} \in \mathcal{M}_\ell^\rho\}, \quad (3.11)$$

denotan los conjuntos de tiempos de retorno directo e inverso respectivamente. Por otro lado los conjuntos

$$\mathcal{H}_\ell^\rho = \{h = \ln(t)/\ell : t \in \mathcal{R}_\ell\}, \quad (3.12)$$

y

$$\bar{\mathcal{H}}_\ell^\rho = \{h = \ln(t)/\ell : t \in \bar{\mathcal{R}}_\ell\}. \quad (3.13)$$

denotan los conjuntos de entropía por bloques directo y reverso respectivamente.

Para este punto, podemos ya estimar los parámetros  $\hat{h}_\ell$  y  $\hat{\sigma}_\ell^2$  antes mencionados como sigue.

$$\hat{h}_\ell = \bar{h} + \hat{\zeta}(h_c - \bar{h}), \quad (3.14)$$

$$\hat{\sigma}_\ell^2 = s^2 + \hat{\zeta}(h_c - \bar{h})^2. \quad (3.15)$$

Cabe señalar que la barra en  $\bar{h}$  es usada aquí para representar la media aritmética y no hace referencia a alguna palabra en sentido reverso, dada en la ecuación 3.2. Recordemos que el detalle de esta estimación puede consultarse en [3]. Los parámetros en las ecuaciones anteriores se obtienen de la siguiente manera

$$\bar{h} = \frac{1}{k} \sum_{i=1}^k h_i, \quad (3.16)$$

$$s^2 = \frac{1}{k} \sum_{i=1}^k (h_i - \bar{h})^2, \quad (3.17)$$

$$\hat{\zeta} = \frac{\phi(\hat{\xi})}{\hat{p}\hat{\xi} + \phi(\hat{\xi})}, \quad (3.18)$$

$$\hat{\xi} = \Phi^{-1}(\hat{p}), \quad (3.19)$$

donde  $\hat{p} = k/m$  como se mencionó anteriormente, es el estimador para la tasa de muestras no censuradas,  $\phi(x)$  es función de densidad de una distribución normal estándar, misma que queda determinada por

$$\phi(x) = e^{-x^2/2}/\sqrt{2\pi}, \quad (3.20)$$

y  $\Phi$  es la función de distribución acumulada. La corrección anterior, que puede consultarse en [3], puede verse como un ajuste estadístico, ya que dicho valor queda censurado conforme aumentamos el valor de  $\ell$ .

Repitiendo los pasos anteriores para el conjunto  $\bar{\mathcal{R}}_\ell$  con el fin de estimar la tasa de entropía reversa por bloques, finalmente obtenemos la tasa de producción de entropía por medio de la siguiente ecuación

$$\hat{e}_p = \hat{h}_\ell^R - \hat{h}_\ell. \quad (3.21)$$

## 3.2. Protocolos de muestreo.

Es vital establecer un límite uniforme para todas las muestras de tiempos de retorno, tal como se mencionó anteriormente, dado que las realizaciones del conjunto de la ecuación 3.8 están censuradas por un  $h_c$ . Esto asegura la homogeneidad en las muestras censuradas. Por ejemplo, una muestra tomada al inicio de la serie de tiempo tendrá un tiempo de búsqueda mayor que una muestra tomada a mitad de la serie de tiempo. Al establecer este límite, garantizamos que todos los bloques tengan el mismo tiempo de búsqueda. Consideraremos a  $T_c = N/2$  como el tiempo máximo de búsqueda, lo que significa que los tiempos de retorno tendrán un rango de búsqueda de  $k + N - \ell$  (consulte la Figura 3.3).

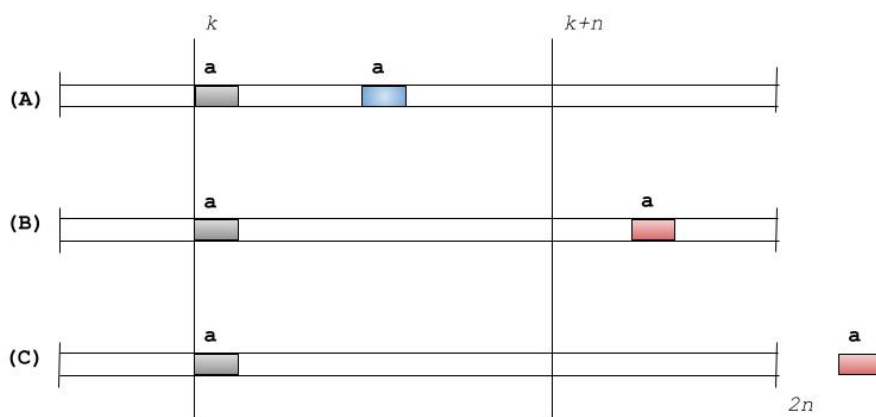


Figura 3.3: Recolección de muestras para el tiempo de retorno. En el caso (A) la muestra queda no censurada, pues el tiempo de retorno está bien definido, sin embargo en los casos (B) y (C) la muestra queda censurada, ya sea porque supera el tiempo de censura  $T_c$  (caso (B)) o porque el tiempo de retorno jamás se encuentra (caso (C)).

Si establecemos el tiempo de censura como  $T_c = N/2$ , la entropía censurada de la ecuación 3.9 tendrá solo un parámetro variable, que es el tamaño  $\ell$  del bloque. Es importante considerar cómo calcular un  $\ell$  óptimo, que llamaremos  $\ell^*$ . Como se puede ver en la Figura 3.4, si aumentamos el tamaño  $\ell$  de los bloques, la precisión mejora gracias al teorema del límite central (TLC), pero también perdemos datos significativos porque es menos probable encontrar ese bloque en la serie de tiempo.

Por eso, elegiremos  $\ell^*$  de manera que el número de muestras no censuradas  $k$  sea aproximadamente la mitad del total de muestras, es decir  $k \approx m/2$ . Según los cálculos en [3], si elegimos  $\ell^*$  de esta manera, los errores de estimación serán menores a 0.06 (según pruebas realizadas en una cadena de Markov con diferentes parámetros).

A continuación, se proporciona una descripción detallada de cómo se llevó a cabo el

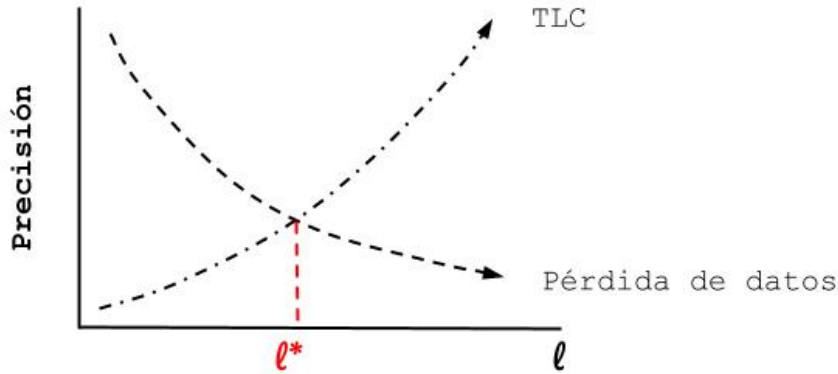


Figura 3.4: Cómo afecta aumentar el tamaño de  $\ell$ . Al aumentar el tamaño de  $\ell$  estaremos ganando precisión con respecto al teorema del límite central, pero al mismo tiempo, perdemos precisión pues conforme el bloque de tamaño  $\ell$  aumenta existe una probabilidad menor de encontrarlo en la realización  $x$ .

muestreo de datos para el análisis anterior.

1. Se partió de una única trayectoria de tamaño  $N$ .
2. Se tomó un número  $m$  de muestras. Dado el espaciamiento  $\Delta$  entre cada muestra, el máximo número de muestras  $m$  debe cumplir que  $m \leq \lfloor \frac{n}{\Delta} \rfloor$ , con  $n = N/2$ .
3. El valor umbral de censura se tomó como  $T_c = n$ .
4. El muestreo se obtuvo seleccionando como bloque inicial, el que está localizado en la posición  $(k\Delta + 1, k\Delta + \ell)$ , donde  $0 \leq k \leq m$ .

### 3.2.1. Prueba de control: tiempo de retorno.

Realizamos este análisis para examinar la eficacia del método de tiempos de retorno. Para ello generamos secuencias usando una cadena de Markov con un espacio de estados  $S = \{0, 1, 2\}$ , un vector de distribución inicial  $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ , y una matriz estocástica  $P : S \times S \rightarrow [0, 1]$  definida como sigue:

$$P = \begin{pmatrix} 0 & p & 1-p \\ 1-p & 0 & p \\ p & 1-p & 0 \end{pmatrix}, \quad (3.22)$$

con  $p \in [0, 1]$ .

Para la prueba de control creamos un programa en Python con las siguientes propiedades:

1. Se tomó una serie de tiempo de tamaño  $N = 100,000$ .
2. El espaciamiento entre cada muestra se consideró como  $\Delta = 20$ .
3. El número de muestras se tomó como  $m = \lfloor \frac{n}{\Delta} \rfloor = 2,500$
4. Se tomaron los siguientes parámetros para la cadena de Markov:  $p = 0.5, 0.6$  y  $0.7$ .
5. Se calculó las tasas de entropía, utilizando los tiempos de retorno.
6. Se realizó la búsqueda de  $\ell^*$ .
7. Se graficaron los resultados.

Esperábamos que este método que hemos implementado demostrara sensibilidad a la hora de detectar la irreversibilidad de un proceso. Para corroborar esta suposición, generamos una serie de tiempo utilizando una cadena de Markov con los parámetros que habíamos mencionado anteriormente. En el escenario donde  $p = 0.5$ , anticipábamos que la serie de tiempo resultante sería reversible. Esto significa que  $e_p \approx 0$ . Por otro lado, para  $p = 0.6$  y  $p = 0.7$ , suponíamos que se presentaría un cierto grado de irreversibilidad, es decir,  $e_p > 0$ .

Para determinar el valor óptimo de  $\ell^*$ , desarrollamos un código en Python que utilizó una base de datos similar a la de la tabla 3.2. Este código fue diseñado específicamente para analizar los datos y proporcionar un resultado más óptimo para  $\ell^*$ .

A continuación, adjuntamos el seudo código de este programa, para proporcionar una visión clara de cómo se realizó el análisis y cómo se obtuvieron los resultados.

```
#Definimos una variable K_l como la mitad del número de muestras
K_L = int(m/2)
minimo = K_L

for i in range(1,longitud_columna L): #longitud de la columna L
#medimos la distancia entre el k_l y cada uno de los ks en K
    dist = abs(K_L-K[i])
    # elegimos la distancia mínima entre K_l y los k's disponibles
    if dist<minimo:
        minimo = dist
        l_ast = L[i]
```

```

l_ast_place = i
K_ast = K[i]

```

```
# l* = l_ast
```

Por lo tanto, seleccionamos el  $\ell^*$  de forma que la distancia entre el  $k_\ell = m/2$  seleccionado sea lo más pequeña posible en comparación con todos los  $k$  disponibles (consulte la columna  $k$  en la tabla 3.2). El código se probó utilizando datos de una cadena de Markov definida en la ecuación 3.22. Es importante recordar que las tasas de entropía directa y reversa para una cadena de Markov se pueden calcular utilizando las ecuaciones 2.25 y 2.26. Al considerar que es igualmente probable iniciar la cadena en cualquiera de los tres estados, la evaluación de las tasas de entropía se simplifica a las siguientes ecuaciones:

$$h(p) = -p \ln(p) - (1 - p) \ln(1 - p), \quad (3.23)$$

y

$$h_R(p) = -(1 - p) \ln(p) - p \ln(1 - p), \quad (3.24)$$

Con lo cual, después de un poco de álgebra podemos concluir que

$$e_p(p) = (2p - 1) \log \left( \frac{p}{1 - p} \right). \quad (3.25)$$

La tabla 3.1 muestra los valores obtenidos<sup>2</sup>. Como se puede apreciar, para la cadena de Markov con parámetro  $p = 0.5$ , las tasas de entropía directa y reversa son iguales, lo que significa que la tasa de producción de entropía es cero. Esto nos lleva a concluir que el proceso es reversible, tal como se anticipó. En los otros dos casos,  $p = 0.6$  y  $p = 0.7$ , la tasa de entropía reversa supera a la tasa de entropía directa, resultando en una tasa de producción de entropía superior a cero y, por ende, un proceso irreversible. Esto también coincide con nuestras expectativas. Este control experimental nos permite conocer el valor teórico para cada caso (proceso reversible e irreversible).

### 3.2.2. Resultados experimento controlado para los distintos valores de $p$ .

La tabla 3.2 presenta los resultados para  $h$  y  $h_R$  en una cadena de Markov con  $p = 0.5$ , considerando diferentes tamaños de bloque  $\ell$ . La tabla 3.1 muestra que los valores teóricos

---

<sup>2</sup>Los valores para  $h$  y  $\hat{h}_R$  se redondean a cuatro dígitos

$p$	$\hat{h}$	$\hat{h}_R$
0.5	$\ln(2) \approx 0.6931$	$\ln(2) \approx 0.6931$
0.6	$\frac{1}{5}(5 \ln(5) - 3 \ln(3) - 2 \ln(2)) \approx 0.6730$	$\frac{1}{5}(5 \ln(5) - 2 \ln(3) - 3 \ln(2)) \approx 0.7541$
0.7	$\frac{1}{10}(10 \ln(10) - 7 \ln(7) - 3 \ln(3)) \approx 0.6108$	$\frac{1}{10}(10 \ln(10) - 3 \ln(7) - 7 \ln(3)) \approx 0.9497$

Tabla 3.1: Cálculo de las tasas de entropía para una cadena de Markov. La primera columna muestra los diferentes valores de  $p$ . La segunda y tercera columna representan la entropía directa y reversa respectivamente, calculadas para cada valor de  $p$ . Como se puede ver, tal como se anticipaba, para el caso  $p = 0.5$ ,  $h = h_R$ . Sin embargo, para  $p = 0.6$  y  $p = 0.7$ ,  $h_R > h$ .

de las tasas de entropía directa y reversa son  $h \approx 0.6931$  y  $h_R \approx 0.6931$  respectivamente<sup>3</sup>. Estos se alcanzan aproximadamente cuando  $15 \leq \ell \leq 16$ . Es crucial considerar el tamaño óptimo de bloque ( $\ell^*$ ) para tener alrededor de la mitad de las muestras no censuradas. Observamos que justamente para  $15 \leq \ell \leq 16$  obtenemos el número de muestras deseado. Seleccionamos  $\ell = 15$ , ya que con este tamaño obtenemos un número de muestras  $k$  donde  $k > n$  no son censuradas, mientras que con  $\ell = 16$  tendríamos  $k < n$ .

En la Figura 3.5, se presentan las distribuciones de las estimaciones de  $h$  y  $h_R$  relacionadas con los diferentes tamaños de bloque  $\ell$  utilizados en la tabla 3.2. Notamos que a medida que el tamaño del bloque  $\ell$  aumenta, las distribuciones pierden su forma acampanada. Tanto para la entropía directa (panel (a)), como para la entropía reversa (panel (b)), con  $\ell^* = 15$ , el histograma se divide aproximadamente a la mitad, es decir, únicamente se mantienen, como se esperaba, cerca de  $k = m/2$  de las muestras censuradas. Para  $\ell > 15$ , el histograma muestra una disminución en el número de muestras.

En el panel (c) de la Figura 3.5, observamos cómo las tasas de entropía directa  $\hat{h}$  y reversa  $\hat{h}_R$  (línea verde y azul, respectivamente) coinciden para valores de  $\ell$  cercanos a 15. Sin embargo, para valores menores, las tasas de entropía difieren en aproximadamente un orden de magnitud cercano a  $10^{-2}$ . Además, a partir de  $\ell = 16$ , ambas tasas de entropía parecen diverger. Esto sugeriría que  $\ell = 15$  es el valor óptimo para realizar el experimento con valores reales, ya que es en este punto donde se alcanzan valores cercanos a las tasas de entropía teóricas  $h$  y  $h_R$ .

La tabla 3.3 resume los resultados del cálculo para  $\hat{h}$  y  $h_R$  para una cadena de Markov con  $p = 0.6$  considerando diferentes tamaños de bloque  $\ell$ . En la tabla 3.1, vemos que las tasas de entropía directa y reversa son  $\hat{h} \approx 0.6730$  y  $\hat{h}_R \approx 0.7541$ , respectivamente<sup>4</sup>. Estos

<sup>3</sup>Los valores para  $\hat{h}$  y  $h_R$  están truncados a cuatro dígitos.

<sup>4</sup>Los valores para  $\hat{h}$  y  $\hat{h}_R$  son truncados a cuatro dígitos.

$\ell$	$k$	$k_R$	$\hat{h}^*$	$\hat{h}_R^*$
5	2500	2500	0.5958	0.4679
6	2500	2500	0.6719	0.5666
7	2500	2500	0.6642	0.6321
8	2500	2500	0.6694	0.6165
9	2500	2500	0.6706	0.6458
10	2500	2500	0.6749	0.6403
11	2500	2500	0.6773	0.6597
12	2499	2498	0.6743	0.6578
13	2453	2453	0.6798	0.6721
14	2145	2163	0.6853	0.6746
<b>15</b>	1582	1607	<b>0.6901</b>	<b>0.6857</b>
<b>16</b>	1012	1027	<b>0.7002</b>	<b>0.7024</b>
17	594	562	0.7159	0.7371
18	331	323	0.7217	0.7834
19	158	156	0.7405	0.8026

Tabla 3.2: Estimación para  $h$  y  $h_R$  para  $p = 0.5$ . Los valores teóricos esperados para las tasas de entropía directa y reversa son  $\hat{h} \approx 0.6931$  y  $\hat{h}_R \approx 0.6931$  respectivamente. El valor  $\ell = 15$  es tomado como  $\ell^*$ , pues con este valor se tiene un número de muestras no censuradas mayor a la 1250. \*Los valores de  $\hat{h}$  y  $\hat{h}_R$  son truncados a cuatro dígitos.

valores se alcanzan aproximadamente para  $\ell = 17$  para  $\hat{h}$  y en  $\ell = 14$  para  $\hat{h}_R$ . Aquí,  $\ell^* = 16$  y  $\ell^* = 14$  son el tamaño de bloque óptimo para estimar  $h$  y  $h_R$ . Tenga en cuenta que al usar este  $\ell$ , la estimación de  $e_p$  puede no ser tan precisa, pero aún es evidente que el proceso es irreversible ya que  $\hat{h}_R > \hat{h}$ .

La Figura 3.6 ilustra la distribución de la estimación de  $h$  (panel (a)) y  $h_R$  (panel (b)) para cada valor de  $\ell$  que corresponde a los datos en la tabla 3.3 para  $p = 0.6$ . Similar al caso anterior, se puede notar que a medida que aumenta el tamaño del bloque, el histograma se reduce. Es fundamental señalar que para  $\ell = 16$  y  $\ell = 14$  en la distribución de  $\hat{h}$  y  $\hat{h}_R$  respectivamente, los histogramas se reducen aproximadamente a la mitad, garantizando así que el número de muestras censuradas sea  $k \approx m/2$ . En el panel (c) se puede ver cómo las tasas de entropía directa  $\hat{h}$  y reversa  $\hat{h}_R$  (representadas por las líneas verde y azul, respectivamente) coinciden para valores de  $\ell$  cercanos a 5 y 7. Sin embargo, cuando los valores son mayores, la diferencia entre ambas curvas puede llegar hasta casi un orden de magnitud de  $2 \times 10^{-2}$ . Esto nos indica que los valores óptimos para llevar



$\ell$	$k$	$k_R$	$\hat{h}^*$	$\hat{h}_R^*$
5	2500	2500	0.6657	0.66956
6	2500	2500	0.6566	0.62552
7	2500	2500	0.6560	0.67880
8	2500	2500	0.6566	0.67107
9	2500	2500	0.6556	0.70284
10	2500	2500	0.6561	0.69966
11	2500	2491	0.6616	0.71524
12	2481	2391	0.6629	0.71884
13	2410	2080	0.6631	0.72690
<b>14</b>	2162	1523	0.6654	<b>0.7380</b>
15	1752	973	0.6687	0.7581
<b>16</b>	1267	579	<b>0.6745</b>	0.7860
17	812	313	0.6820	0.8140
18	488	177	0.6877	0.8796
19	273	85	0.7002	0.8998

Tabla 3.3: Estimación para  $h$  y  $h_R$  con  $p = 0.6$ . Los valores teóricos esperados para las tasas de entropía directa y reversa son  $\hat{h} \approx 0.6730$  y  $\hat{h}_R \approx 0.7541$  respectivamente. El tamaño de bloque óptimo en este caso es  $\ell^* = 16$ . \*Los valores de  $\hat{h}$  y  $\hat{h}_R$  son truncados a cuatro dígitos.

a cabo el experimento en un proceso irreversible son aquellos en los que  $14 \leq \ell \leq 16$ . Para concluir, la tabla 3.7 resume los resultados para  $\hat{h}$  y  $\hat{h}_R$  de una cadena de Markov con  $p = 0.7$ , considerando diferentes tamaños de bloque  $\ell$ . La tabla 3.1 muestra que las tasas de entropía directa y reversa son aproximadamente  $\hat{h} \approx 0.6108$  y  $\hat{h}_R \approx 0.9497$ , respectivamente<sup>5</sup>. Estos valores se alcanzan, más o menos, para  $\ell = 18$  en  $\hat{h}$  y  $\ell = 13$  en  $\hat{h}_R$ , los cuales se consideraron como  $\ell^*$ . Es importante notar que para  $\ell = 13$ , el número de muestras censuradas es casi el doble del número de muestras requeridas, pero elegimos este valor porque nos ofrece una mejor aproximación a la  $\hat{h}_R$  teórica. La Figura 3.7 muestra las distribuciones para el cálculo de las tasas de entropía directa y reversa, que corresponden a los datos de la tabla 3.7. Al igual que en casos anteriores, la distribución de los  $\ell^*$  para ambas entropías se corta aproximadamente a la mitad, con  $\ell = 18$  para la entropía directa y  $\ell = 13$  para la reversa. En el panel (c) de la Figura 3.7, podemos ver claramente que las tasas de entropía directa  $\hat{h}$  y reversa  $\hat{h}_R$  (representadas por las líneas verde y azul

<sup>5</sup>Los valores para  $\hat{h}$  y  $\hat{h}_R$  se han truncado a cuatro dígitos.

respectivamente) no coinciden en ningún punto de  $\ell$ . Esto indica claramente un proceso irreversible.

Las tablas 3.4, 3.5 y 3.6 presentan un resumen de los valores obtenidos en los análisis de esta sección. Hemos definido  $\Delta\hat{h} : |h - \hat{h}|$  y  $\Delta\hat{h}_R : |h_R - \hat{h}_R|$  para calcular el error relativo de cada estimación. Es importante mencionar que en la tabla 3.6 solo se calcula el error absoluto  $\Delta\hat{e}_p : |e_p - \hat{e}_p|$  al estimar  $e_p$ , ya que no es posible realizar el cálculo numérico de  $\hat{e}_p$  cuando  $p = 0.5$ . Mantenemos este cálculo para  $p = 0.6$  y  $p = 0.7$  para mantener la consistencia en los cálculos.

$p$	$\ell$	$h^*$	$\hat{h}^*$	$\Delta\hat{h}/h$	Porcentaje
0.5	15	0.6931	0.6901	0.0432	0.4328 %
0.6	16	0.6730	0.6745	0.0022	0.2229 %
0.7	18	0.6108	0.6048	0.0098	0.9823 %

Tabla 3.4: Tabla comparativa entre  $\hat{h}$  y los valores obtenidos para  $h$ . En la última columna podemos observar que en todos los casos el porcentaje de error de estimación es menor al 1%. \*Los valores fueron truncados a cuatro dígitos.

$p$	$\ell$	$h_R^*$	$\hat{h}_R^*$	$\Delta\hat{h}_R/h$	Porcentaje
0.5	15	0.6931	0.6857	0.0106	1.0677 %
0.6	14	0.7541	0.7380	0.0213	2.1349 %
0.7	13	0.9497	0.9389	0.0113	1.1372 %

Tabla 3.5: Tabla comparativa entre  $h_R$  y los valores obtenidos para  $\hat{h}_R$ . En la última columna podemos observar que el porcentaje de error de estimación es ligeramente mayor al 1% con  $p = 0.5$  y  $p = 0,7$  mientras que para  $p = 0.6$  se alcanza un porcentaje mayor a 2%. \*Los valores fueron truncados a cuatro dígitos.

$p$	$e_p^*$	$\hat{e}_p^*$	$\Delta\hat{e}_p$	Porcentaje
0.5	0	-0.0044	0.0044	0.44 %
0.6	0.0811	0.0635	0.0176	1.76 %
0.7	0.3389	0.3341	0.0048	0.48 %

Tabla 3.6: Tabla comparativa entre  $e_p$  y los valores obtenidos para  $\hat{e}_p$ . Nuevamente la última columna muestra el porcentaje de error de estimación es menor al 2%. \*Los valores fueron truncados a cuatro dígitos.

Los resultados que hemos obtenido, como se muestra en las tablas comparativas (3.4, 3.5 y 3.6), indican que el método de tiempos de retorno puede ser una herramienta efectiva para calcular la tasa de producción de entropía en series de tiempo reales. Durante nuestras pruebas, este método ha sido útil para determinar si un proceso es reversible o irreversible. En la mayoría de los casos, el error fue inferior al 1.5 %, lo que muestra una alta precisión. Sin embargo, es crucial recordar que estos resultados son preliminares y solo se aplican a los casos específicos que hemos estudiado. En el capítulo 4, profundizaremos en los resultados que hemos obtenido al analizar la irreversibilidad de un periodo específico de la serie de tiempo de Bitcoin con este método. Este análisis nos ayudará a entender mejor la efectividad de este método en situaciones reales.

$\ell$	$k$	$k_R$	$\hat{h}^*$	$\hat{h}_R^*$
5	2500	2500	0.6144	0.8385
6	2500	2500	0.6081	0.8172
7	2500	2500	0.6034	0.8664
8	2500	2500	0.5983	0.8635
9	2500	2450	0.5979	0.8902
10	2499	2194	0.6007	0.8902
11	2487	1726	0.5995	0.9016
12	2454	1145	0.6004	0.9200
<b>13</b>	2383	667	0.6027	<b>0.9389</b>
14	2240	354	0.6047	0.9958
15	2056	182	0.6043	0.9738
16	1794	89	0.6048	0.9956
17	1506	37	0.6055	1.0223
<b>18</b>	1217	17	<b>0.6048</b>	1.0338
19	916	9	0.6062	0.9660

Tabla 3.7: Estimación para  $h$  y  $h_R$  para  $p = 0.7$ . Los valores teóricos esperados para las tasas de entropía directa y reversa son  $h \approx 0.6108$  y  $h_R \approx 0.9497$  respectivamente. \*Los valores de  $\hat{h}$  y  $\hat{h}_R$  son truncados a cuatro dígitos.

### 3.3. *Matching times.*

En esta sección, proporcionamos una explicación clara del procedimiento para calcular las tasas de entropía utilizando los *matching times*. Tenga en cuenta que puede encontrar

más detalles sobre este método en [36].

Sea  $X = \{X_n : n \in \mathbb{N}_0\}$  un proceso estacionario ergódico generado por una ley  $\mathbb{P}$ , cuyas realizaciones son sucesiones de símbolos infinitas tomadas de un conjunto  $A$ , por lo tanto el conjunto de todas las posibles realizaciones es un subconjunto de  $A^{\mathbb{N}}$ . Se denota  $x = (x_0x_1x_2x_3\dots)$  a una realización infinita del proceso  $X$ . Sea  $\ell$  un entero positivo que denota el tamaño una realización del proceso  $X$ . A la cadena de los primeros  $\ell$  símbolos se denota como  $X_0^{\ell-1}$ .

**Definición 3.3.1.** *A una cadena finita compuesta de  $\ell$  símbolos se le llama palabra o bloque de tamaño  $\ell$ . Más aún, se dice que  $a$  ocurre en el lugar  $k$  de la trayectoria  $x$  si*

$$x_k^{k+\ell-1} = a. \quad (3.26)$$

Se definen a continuación los *matching times*.

**Definición 3.3.2.** *Dada una realización  $x = (x_0x_1x_2x_3\dots, x_t)$  hasta un tiempo  $t$  del proceso  $X$ , el matching time  $L_t^+$  se define como el  $\ell$  más pequeño tal que la  $\ell$ -palabra  $x_0^{\ell-1} = (x_0, x_1, \dots, x_{\ell-1}, x_\ell)$  no reaparece en  $x$ , es decir*

$$L_t^+(x) = \min\{\ell : x_0^{\ell-1} \neq x_j^{j+\ell-1}, \forall j = 1, 2, \dots, t - \ell + 1\}, \quad (3.27)$$

análogamente, se define matching time reverso  $L_t^-$  como la  $\ell$  más pequeña, tal que la  $\ell$ -palabra temporalmente invertida  $x_{\ell-1}^0 = (x_{\ell-1}, x_{\ell-2}, \dots, x_1, x_0)$  no reaparece en  $x$ , es decir

$$L_t^-(x) = \min\{\ell : x_{\ell-1}^0 \neq x_j^{j+\ell-1}, \forall j = 1, 2, \dots, t - \ell + 1\}. \quad (3.28)$$

De acuerdo con Kontoyannis [41], tanto  $L_t^+$  como  $L_t^-$ , satisfacen la ley de los grandes números. Asimismo, dichas variables aleatorias satisfacen el teorema central del límite, ya que las siguientes expresiones

$$\sqrt{\frac{h^3 \log(t)}{\sigma^2}} \left( \frac{L_t^+}{\log(t)} - \frac{1}{h} \right) \rightarrow \mathcal{N}(0, 1), \quad (3.29)$$

y

$$\sqrt{\frac{h_R^3 \log(t)}{\sigma_R^2}} \left( \frac{L_t^-}{\log(t)} - \frac{1}{h_R} \right) \rightarrow \mathcal{N}(0, 1), \quad (3.30)$$

convergen en distribución, al incrementar  $t$ . De esta manera

$$X_t^+ = \frac{L_t^+}{\log(t)} \xrightarrow{cs} \frac{1}{h}, \quad t \rightarrow \infty, \quad (3.31)$$

y

$$X_t^- = \frac{L_t^-}{\log(t)} \xrightarrow{cs} \frac{1}{h_R}, \quad t \rightarrow \infty, \quad (3.32)$$

convergen casi seguramente  $\frac{1}{h}$  y  $\frac{1}{h_R}$  respectivamente, cuando  $t \rightarrow \infty$ . Para estimar la tasa de entropía se realiza lo siguiente:

**Definición 3.3.3.** Sea  $\mathcal{W} = \{x_i : |x_i| = t_i, 1 \leq i \leq m\}$  el conjunto de  $m$  muestras de secuencias (observaciones del proceso  $X$ ). Sea  $t$  el tamaño de las muestras. La colección de sucesiones de tamaño  $t$  se denota como:

$$\mathcal{T} = \{t_i : 1 \leq i \leq m\}.$$

Aplicando las ecuaciones 3.27 y 3.28 a cada secuencia en  $\mathcal{W}$  se obtienen los siguientes conjuntos

$$\mathcal{L}^+ = \{\ell^+ = L_t^+(x_i) : x_i \in \mathcal{W}, 1 \leq i \leq m\}, \quad (3.33)$$

y

$$\mathcal{L}^- = \{\ell^- = L_t^-(x_i) : x_i \in \mathcal{W}, 1 \leq i \leq m\}, \quad (3.34)$$

los cuales contienen los *matching times* directos y reversos, respectivamente de cada secuencia en  $\mathcal{W}$ . Podemos definir mediante los conjuntos anteriores, los estimadores  $\hat{h}$  y  $\hat{h}_R$  para la tasa de entropía directa y reversa, utilizando los *matching times*, de la siguiente manera:

$$\hat{h} = \frac{\ln(t)}{\left(\frac{1}{m}\right) \sum_{j=1}^m \ell_j^+}, \quad (3.35)$$

y

$$\hat{h}_R = \frac{\ln(t)}{\left(\frac{1}{m}\right) \sum_{j=1}^m \ell_j^-}. \quad (3.36)$$

Con lo anterior, nuevamente podemos obtener la tasa de producción de entropía [42] de la siguiente manera

$$e_p = h_R - h. \quad (3.37)$$

### 3.3.1. Experimento controlado: cadena de Markov.

Para evaluar la eficacia de este método en la identificación de la irreversibilidad en un proceso, realizamos un experimento de control, similar a los ejercicios que llevamos a cabo con los tiempos de retorno. Hicimos uso de Python para implementar el experimento, que implicaba la generación de secuencias de datos a partir de una cadena de Markov. En este

contexto, utilizamos una cadena de Markov con ciertas características para evaluar la eficacia del método en la detección de la irreversibilidad. A continuación, proporcionaremos más detalles sobre estas características.

- Se tomaron  $m$  muestras, con  $m = 500$ .
- Se utilizaron series de distintos tamaños  $t = 10^5$  y  $10^6$ .
- Se realizó el experimento para los siguientes valores de  $p = 0.5, 0.6, 0.7$ . Nuevamente, se esperaba que los resultados arrojarán que el proceso es reversible para  $p = 0.5$  e irreversible para  $p = 0.6$  y  $p = 0.7$ .
- Se realizó el cálculo de las tasas de entropía, utilizando *matching times*.
- Se graficó la estimación de las tasas de entropía directa y reversa.

La tabla 3.8 muestra los valores de  $\hat{h}$  y  $\hat{h}_R$  para diferentes valores de  $p$  ( $p = 0.5, 0.6, 0.7$ ) y para  $t = 10^5$  y  $10^6$ . La última columna da el valor aproximado para  $e_p$  teóricamente<sup>6</sup>. Comparándolo con los valores en la quinta columna, que corresponde a la estimación  $\hat{e}_p$ , vemos que excepto para el caso  $p = 0.5$  la aproximación es mejor cuando  $t = 10^5$ . También notamos que aunque el método no hace una aproximación exacta de la tasa de producción de entropía, es coherente con lo que se esperaba: un proceso reversible ( $e_p$  cercano a cero) cuando  $p = 0.5$  y un proceso irreversible ( $e_p > 0$ ) cuando  $p > 0.5$ .

$p$	$\ell$	$\hat{h}^*$	$\hat{h}_R^*$	$\hat{e}_p^*$	$h$ teórica*	$h_R$ teórica*	$e_p$ teórica*
0.5	$10^5$	0.6646	0.6683	0.0037	0.6931	0.6931	0
	$10^6$	0.5562	0.5590	0.0028			
0.6	$10^5$	0.6386	0.7157	0.0771	0.6730	0.7510	0.078
	$10^6$	0.5406	0.6005	0.0599			
0.7	$10^5$	0.5833	0.8762	0.2929	0.6108	0.9497	0.3389
	$10^6$	0.4911	0.7410	0.2498			

Tabla 3.8: Estimación para  $h$  y  $h_R$  con distintos valores de  $p$ . Los valores teóricos se muestran en las columnas  $h$ ,  $h_R$  y  $e_p$  respectivamente. \*Los valores son aproximados a cuatro dígitos.

La Figura 3.8 presenta los resultados de las estimaciones de las tasas de producción de entropía directa y reversa (líneas azul y verde, respectivamente) para  $p = 0.5$  y para los

<sup>6</sup>Los valores para  $\hat{h}$ ,  $\hat{h}_R$  y  $e_p$ , tanto teóricos como estimados, se redondean a cuatro dígitos.

diferentes tamaños de  $t$ . Se puede ver que el comportamiento de ambas tasas de entropía es estadísticamente similar, lo que indica un proceso reversible para cualquier tamaño de bloque, tal como se anticipó.

En la Figura 3.9, se exhiben los resultados de las estimaciones de las tasas de producción de entropía directa y reversa (líneas azul y verde, respectivamente) para  $p = 0.6$  y para los diferentes tamaños de  $t$ . En este escenario, las tasas de entropía parecen desviarse ligeramente una de la otra. Esto concuerda con nuestras expectativas, ya que se observa un proceso con cierto grado de irreversibilidad.

Finalmente, en la Figura 3.10, se muestran los resultados de las estimaciones de las tasas de producción de entropía directa y reversa (líneas azul y verde, respectivamente) para  $p = 0.7$  y para los diferentes tamaños de  $t$ . Aquí, podemos ver una mayor diferencia entre las curvas correspondientes a las tasas de producción de entropía directa  $\hat{h}$  y la tasa de entropía reversa  $\hat{h}_R$ , lo que indica que el proceso es irreversible, como se esperaba.

La tabla 3.9 muestra el porcentaje de error en la estimación de  $e_p$  para diferentes valores de  $p$  y  $t$ . Al examinar la tabla, notamos que la mejor estimación se obtiene con  $t = 10^5$ , donde los errores generalmente no superan el 1%, excepto para  $p = 0.7$ . Estos resultados son acordes con nuestras expectativas, lo que sugiere que los *matching times* podrían ser útiles para estimar la tasa de producción de entropía. Para tener más seguridad en la capacidad del método para detectar la irreversibilidad en un proceso, en la siguiente sección presentamos los resultados obtenidos al aplicar este método a una serie de tiempo que sigue una distribución normal.

$t$	$p$	$\Delta\hat{e}_p$	porcentaje
$10^5$	0.5	0.0037	0.37 %
	0.6	0.0009	0.09 %
	0.7	0.04591	4.59 %
$10^6$	0.5	0.0028	0.28 %
	0.6	0.0181	1.81 %
	0.7	0.0891	8.91 %

Tabla 3.9: Tabla comparativa para las estimaciones de  $e_p$  con los valores de  $t$  utilizados. Podemos observar que con  $t = 10^5$  los errores de estimación para  $e_p$ , en la mayoría de los casos se mantienen por debajo del 1% salvo cuando  $p = 0.7$ .

### 3.3.2. Experimento controlado: datos aleatorios con distribución normal.

En esta sección, presentamos un experimento adicional que se llevó a cabo con el objetivo de proporcionar una evaluación más completa del método de los *matching times*. Este experimento fue diseñado como una prueba de control que nos permitiera examinar el rendimiento del método en un entorno controlado y predecible.

La realización de este experimento se llevó a cabo utilizando el lenguaje de programación Python. El experimento se basó en el uso de series de datos aleatorios generadas para este propósito. Estos datos fueron generados para tener características específicas que nos permitieran probar de manera efectiva las capacidades del método de los *matching times*. Es importante destacar que los datos utilizados en este experimento fueron generados para seguir una distribución normal. A continuación, vamos a desglosar las características del experimento y los datos que consideramos.

- Se tomó una cantidad de muestras  $m = 500$ .
- Se elaboraron series de tiempo con los siguientes tamaños  $t = 10^5, 10^6$ .
- Las series anteriores se crearon como datos normales con media cero y varianza 0.005 ( $N(0, 0.005)$ ).
- Se transformaron los datos a logaritmos de retorno.
- Se codificaron los datos de manera binaria, dependiendo si el dato obtenido en el paso anterior era mayor o menor a cero.
- Se calculó de tasas de entropía, utilizando *matching times*.
- Se graficó la estimación de las tasas de entropía directa y reversa.

Dada la naturaleza aleatoria, independiente y normalmente distribuida de estos datos generados artificialmente, esperábamos que fueran reversibles, indicando un valor de  $e_p$  cercano a cero. La tabla 3.10 muestra los resultados obtenidos para la estimación de la tasa de entropía directa y reversa, así como la estimación de  $e_p$  con los parámetros mencionados<sup>7</sup>. Nos encontramos con un resultado inesperado: el valor obtenido para  $e_p$  es negativo, lo que probablemente se deba a algún error estadístico, ya que sabemos que  $e_p \geq 0$ . Sin embargo, este valor sigue estando muy cerca de cero.

---

<sup>7</sup>Los valores para  $\hat{h}$ ,  $\hat{h}_R$  y  $e_p$  son truncados a cuatro dígitos.



La Figura 3.11 presenta los resultados obtenidos al calcular las tasas de entropía  $\hat{h}$  y  $\hat{h}_R$  (líneas azul y verde, respectivamente) utilizando series de tiempo de tamaños  $t = 10^5$  y  $t = 10^6$ . Como se puede observar, no hay una diferencia significativa entre las curvas de ambos estimadores, lo que está en línea con nuestras expectativas de que el proceso era reversible.

Basándonos en los resultados que hemos obtenido, podemos llegar a la conclusión de que la metodología de los *matching times*, que se ha implementado en nuestro estudio, puede llegar a ser una herramienta útil y efectiva para medir la irreversibilidad. Esta afirmación se ve respaldada por los resultados satisfactorios que hemos obtenido en las pruebas de control que se llevaron a cabo.

$m$	$t$	$\hat{h}^*$	$\hat{h}_R^*$	$\hat{e}_p^*$
500	100000	0.6398	0.6371	-0.0028
	1000000	0.6486	0.6428	-0.0057

Tabla 3.10: Estimación de tasas de entropía datos aleatorios con distribución normal ( $\mathcal{N}(0, 0.005)$ ). La primera columna indica el número de muestras utilizadas, que en todos los casos fueron 500. La segunda columna muestra los tamaños de las series de tiempo utilizadas, mientras que las siguientes columnas muestran los resultados de la estimación de las tasas de entropía directa  $\hat{h}$  y reversa  $\hat{h}_R$ . Finalmente la última columna muestra la tasa de producción de entropía  $e_p$  estimada. Podemos observar que a pesar de incrementar el tamaño de la serie de tiempo, la estimación de  $e_p$  permanece en el mismo orden de magnitud. \*Los valores son aproximados a cuatro dígitos.

### 3.4. Datos sustitutos y prueba de hipótesis.

En los análisis que vamos a presentar a continuación, llevamos a cabo el mismo estudio, pero también utilizando datos sustitutos. Nuestra intención con este enfoque es obtener significancia estadística que permita afirmar con mayor seguridad las conclusiones derivadas de nuestra investigación. Los datos sustitutos son útiles en el análisis estadístico y se crean al permutar los datos originales. Esta nueva serie, preserva la media, la varianza, la distribución de la serie original y tiene una característica estadística importante: todos sus elementos son independientes. Además al permutar los datos, cualquier estructura temporal existente en la serie de tiempo original, es destruida en la nueva serie [43], lo que les da la propiedad de ser reversibles. Por lo tanto, podemos hacer una prueba de hipótesis para ver si las series de tiempo reales son o no irreversibles, con un cierto nivel

de confianza. Aquí se describe cómo se realizó la prueba de hipótesis. Sea  $\mu_s = 0$  la media de algún estimador para la tasa de producción de entropía  $\hat{e}_p$  de los datos sustitutos y sea  $\mu$  la media de dicho estimador, pero para los datos reales. Definimos la hipótesis nula  $H_0$  y la hipótesis alternativa  $H_1$  como

$$\begin{aligned} H_0 : \mu_s &= \mu, \\ H_1 : \mu &> \mu_s. \end{aligned} \tag{3.38}$$

Podemos observar en la ecuación 3.4 que la hipótesis nula sugiere que los datos reales son reversibles. Por otro lado, la hipótesis alternativa indica la existencia de un cierto grado de irreversibilidad en los datos. Sea  $\alpha = 0.0495$  el nivel de significancia. A continuación obtenemos la desviación estándar para los datos sustitutos definimos el valor crítico  $c$  como

$$c = \mu_s + \gamma_\alpha \sigma, \tag{3.39}$$

con  $\gamma_\alpha = 1.65$ , la cual será de ayuda para determinar si se rechaza o no la hipótesis nula. Si el estimador  $\hat{e}_p > c$ , entonces la hipótesis alternativa  $H_1$  se cumple, por lo que rechazamos la hipótesis nula y concluimos que los datos son irreversibles con un nivel  $1 - \alpha$  de confiabilidad. Si no, mantenemos la hipótesis nula y asumimos que los datos son reversibles.

### 3.5. Índice de asimetría desfasada.

La técnica que aplicamos se basa en la referencia [44], que destaca sus beneficios, especialmente en cuanto a tiempos de recurrencia. La ventaja principal es que los parámetros a estimar son menos en comparación con otros métodos. Para nuestro análisis, recurrimos a los rendimientos logarítmicos  $\ell$  de Bitcoin desde 2017 hasta 2022, por trimestre, dando como resultado 4 subseries de tiempo anualmente. Luego, calculamos la media y la varianza de los datos para codificarlos. La codificación se realizó en función de la posición de cada dato, respecto a la media y la varianza, tal como se detalla a continuación.

$$l_t = \begin{cases} 3, & \text{si } \mu + \sigma < \ell < \infty, \\ 2, & \text{si } 0 < \ell < \mu + \sigma, \\ 1, & \text{si } \mu - \sigma < \ell < 0, \\ 0, & \text{si } -\infty < \ell < \mu - \sigma. \end{cases}$$

Una vez teniendo la serie codificada  $S_\tau$  de longitud  $n$ , la obtención de muestras se realizó de la siguiente manera: nos situamos en el primer elemento de la serie y tomamos

una muestra mediante una ventana de tamaño  $\tau$ , donde  $3 \leq \tau \leq 22$ , realizando un deslizamiento de tamaño 1. Al estar en el segundo elemento tomamos nuevamente una muestra de tamaño  $\tau$  continuando así hasta llegar al elemento  $n - 3$  (véase la fig. 3.14). Posteriormente calculamos la probabilidad condicional de empezar en el estado  $i$  y llegar al estado  $j$  mediante la siguiente ecuación

$$P(S_{j+\tau} = j | S_j = i), \quad (3.40)$$

para  $0 \leq i, j \leq 3$ . Posteriormente, estimamos la matriz de transición  $Q^\tau \in M_{n \times n}$ , donde

$$Q_{ij}^\tau = P(S_{j+\tau} = j | S_j = i), \quad (3.41)$$

y calculamos el vector estacionario  $\pi_s$ , asociado al *eigenvalor*  $\lambda = 1$ . Calculamos la tasas de entropía dependientes de  $\tau$ ,  $\hat{h}(\tau)$  y  $\hat{h}_R(\tau)$  haciendo uso de las ecuaciones 2.25 y 2.26, donde

$$h(\tau) = \sum_i \pi_{s_i} Q_{ij}^\tau \ln Q_{ij}^\tau, \quad (3.42)$$

y

$$h_R(\tau) = \sum_i \pi_{s_i} Q_{ij}^\tau \ln Q_{ji}^\tau. \quad (3.43)$$

Finalmente calculamos el índice de asimetría desfasada,  $e_p(\tau)$ , mediante la siguiente ecuación

$$e_p(\tau) = h_R(\tau) - h(\tau). \quad (3.44)$$

El procedimiento que acabamos de detallar también se ha llevado a cabo con datos sustitutos para poder comparar los resultados y obtener una confiabilidad estadística. Es importante recordar que cuando se realiza el análisis con datos sustitutos, esperamos que los resultados indiquen que el proceso es reversible, es decir, que  $e_p(\tau) \approx 0$ . En el Capítulo 4, presentaremos los resultados obtenidos para varios trimestres representativos. La Figura 3.13 muestra los resultados para el segundo y tercer trimestre de 2017, como un ejemplo para explicar el comportamiento del índice de asimetría desfasada  $e_p(\tau)$ . En las figuras se muestra el promedio de  $e_p(\tau)$  para los datos sustitutos, destacado con una línea roja. En el panel (b), se puede notar el comportamiento esperado de  $e_p(\tau)$  (línea azul), que disminuye rápidamente hacia cero y el área debajo de esta curva es pequeña. Sin embargo, en el panel (a) se puede notar un comportamiento diferente. La curva correspondiente a  $e_p(\tau)$  disminuye a cero, pero lo hace de forma más gradual, resultando en un área bajo la curva mayor que la observada en el panel (b). En referencia a [44], los autores realizaron diversas pruebas para una cadena de Markov con diferentes parámetros:  $p = 0.5$ ,  $p =$

0.6,  $p = 0.7$ ,  $p = 0.8$  y  $p = 0.9$  y  $0 \leq \tau \leq 100$ . Los resultados muestran que la curva correspondiente a  $e_p(\tau)$  para la cadena de Markov con  $p = 0.5$  tiene un área bajo la curva menor en comparación a cuando  $p = 0.9$ . Por esta razón, en nuestra sección de resultados incluimos también las curvas que representan el área bajo la curva de los trimestres en conjunto. La determinación del área bajo la curva se llevó a cabo del siguiente modo.

1. Calculamos la media  $\mu_s(\tau)$  de  $e_p(\tau)$  para los datos sustitutos.
2. Calculamos la diferencia  $d$  entre el valor obtenido de  $e_p(\tau)$  para el trimestre  $i$  y  $\mu_s(\tau)$ , con  $1 \leq i \leq 24$ .
3. Si  $d > 0$  preservamos dicho valor; en otro caso  $d = 0$ . Lo anterior con la finalidad de sólo quedarnos con el valor positivo del área bajo la curva.
4. Graficamos los valores obtenidos en el paso anterior.

Las curvas que se presentan en el análisis ofrecen una visión mucho más detallada y precisa de la irreversibilidad durante cada uno de los periodos que se han analizado. Estas curvas son herramientas esenciales en nuestro estudio, ya que nos permiten cuantificar y visualizar la irreversibilidad de una manera comprensible y significativa. Esencialmente, la interpretación de estas curvas es bastante directa: un área más grande bajo la curva se traduce en un mayor grado de irreversibilidad durante ese periodo específico. Esto significa que, si observamos un incremento en el área bajo la curva, podemos interpretar esto como un aumento en la irreversibilidad durante ese intervalo de tiempo.

### 3.6. Tratamiento de datos.

Como señalamos antes, descubrimos que las bases de datos que obtuvimos de [38] a menudo carecían de datos o tenían periodos de tiempo ausentes. A pesar de estos vacíos, pudimos realizar nuestro análisis con los datos disponibles. Sin embargo, estos huecos en los datos podrían provocar desfases en nuestras curvas y dificultar el análisis preciso de periodos de irreversibilidad o eficiencia. A continuación, detallamos el proceso que seguimos para manejar estas bases de datos. La tabla 3.11 ofrece un resumen de la cantidad de datos ausentes y el periodo analizado para cada criptomoneda que estudiamos. Podemos ver que la serie temporal con más datos faltantes es la de Litecoin, seguida de Ether, y Bitcoin en tercer lugar. A pesar de que las bases de datos de XRP Ledger y Bitcoin Cash estaban casi completas, aún realizamos el siguiente procedimiento para rellenar los huecos. Utilizamos Python 3 y la paquetería Pandas para realizar el relleno en la base de datos de los rendimientos logarítmicos.

Criptomoneda	Periodo analizado	Cantidad de datos faltantes
BTC	2015-2022	0.9802 %
ETH	2017-2022	3.4315 %
XRP	2017-2022	0.0582 %
BCH	2018-2022	0.0002 %
LTH	2019-2022	3.5554 %

Tabla 3.11: Información de datos faltantes en las bases de datos utilizadas para cada criptomoneda. Podemos observar una mayor cantidad de datos faltantes para Litecoin, seguido de Ether; la base de datos con menor pérdida de datos es Bitcoin Cash con sólo un 0.0002 % de datos faltantes.

1. Se buscaron las fechas faltantes, comparando las fechas disponibles en la base de datos con una lista de fechas que contenía horas y minutos. Se tuvo en cuenta que los años 2016 y 2020 son años bisiestos, por lo cual se creó una lista específica para años bisiestos que contenía desde el minuto 0:01 hasta las 23:59 del 29 de febrero.
2. Se tenían tres casos en los que debimos modificar o añadir algún dato faltante.
  - a) Los precios al tiempo  $t$  y  $t + 1$  eran iguales.
  - b) Los precios al tiempo  $t$  y  $t + 1$  no existían.
  - c) El precio al tiempo  $t$ , para algún  $t$  era cero.
3. Sabemos que los rendimientos logarítmicos de los precios se distribuyen normalmente. Por lo tanto, calculamos la media ( $\mu$ ) y la desviación estándar ( $s$ ) de los datos disponibles. Con estos parámetros, generamos datos aleatorios con distribución normal para completar los espacios de los casos anteriores.
4. En otro caso, se calculó el rendimiento logarítmico como se describió en la definición 2.4.3.
5. Se verificó que se tuvieran 525,600 datos por año y se realizaron los análisis correspondientes.

Gracias al procedimiento detallado anteriormente, podemos asegurar que los resultados presentados se alineen de manera perfecta con los periodos de tiempo analizados, lo que nos permite tener una coherencia en los datos. Esto es crucial porque nos asegura que tenemos un número constante de datos por cada año, lo que a su vez garantiza que nuestras

conclusiones sean representativas de cada periodo de tiempo en cuestión. De esta manera, podemos presentar resultados concretos y precisos, aumentando así la confiabilidad de nuestros hallazgos.

### 3.7. Patrones de tendencias.

En esta sección presentamos un análisis que hemos denominado patrones de tendencias de subida y bajada de precios. En la mayoría de los casos, nos referiremos a este proceso como el análisis de rachas. Nuevamente utilizamos las series de tiempo de rendimientos logarítmicos, previamente tratados como se describió en la sección anterior.

Sea  $(\ell_0, \ell_1, \dots, \ell_N)$  una muestra de rendimientos logarítmicos de tamaño  $N$ . Una racha de subida se define como una subsucesión  $(\ell_t, \ell_{t+1}, \dots, \ell_{t+n})$  que ocurre en la muestra y que cumple la siguiente condición

$$\ell_t < \ell_{t+1} < \dots < \ell_{t+n}, \quad (3.45)$$

donde  $t \in \mathbb{N}_0$  y  $n \in \mathbb{N}$ . Similarmente se define una racha de bajada como una subsucesión  $(\ell_t, \ell_{t+1}, \dots, \ell_{t+n})$  que ocurre en la muestra y que cumple la siguiente condición

$$\ell_t > \ell_{t+1} > \dots > \ell_{t+n}.$$

También es posible definir las rachas constantes, como aquellas que no son ni de subida ni de bajada. Estas rachas no serán consideradas en este análisis ya que en la realidad, la probabilidad de que una racha de este tipo ocurra es muy baja, salvo errores en la base de datos como mencionamos anteriormente.

El análisis consiste en contar la duración  $t$ , en el que los rendimientos logarítmicos, tuvieron una racha de subida, o bien una racha de bajada, por lo tanto definimos lo siguiente.

**Definición 3.7.1.** *Se define la duración de tiempo  $T_{\uparrow}(t)$  en la que una racha es monótona creciente, comenzando en el tiempo  $t$  como:*

$$T_{\uparrow}(t) = \text{máx}\{n : \ell_t < \ell_{t+1} < \ell_{t+2} < \dots < \ell_{t+n}\}, \quad (3.46)$$

*análogamente, la duración de tiempo  $T_{\downarrow}(t)$  en la que una racha es monótona decreciente, comenzando en el tiempo  $t$  como:*

$$T_{\downarrow}(t) = \text{máx}\{n : \ell_t > \ell_{t+1} > \ell_{t+2} > \dots > \ell_{t+n}\}. \quad (3.47)$$

Definimos a continuación las distribuciones correspondientes a cada una de las variables aleatorias dadas en la definición anterior.

**Definición 3.7.2.** Sea  $X := \{X_t : t \in \mathbb{N}_0\}$  un proceso estocástico, y sea  $T_{\uparrow}(0)$  la duración de una racha de subida que comienza en  $t = 0$ . Denotamos por  $P_{\uparrow}(n)$  a la probabilidad de que  $T_{\uparrow}(0)$  tome un valor  $n \in \mathbb{N}$ , i.e.,

$$P_{\uparrow}(n) := \mathbb{P}(T_{\uparrow}(0) = n). \quad (3.48)$$

Denotamos como  $P_{\downarrow}(n)$  a la probabilidad de que  $T_{\downarrow}(0)$  tome un valor  $n \in \mathbb{N}$ , i.e.,

$$P_{\downarrow}(n) := \mathbb{P}(T_{\downarrow}(0) = n). \quad (3.49)$$

En lo subsecuente, nos referimos a  $P_{\uparrow}$  y a  $P_{\downarrow}$  como la distribución de subida y distribución de bajada del proceso  $X$ , respectivamente.

Podemos observar que las distribuciones de probabilidad de bajada y subida coinciden cuando el proceso estocástico  $X = \{X_t : t \in T\}$  es una sucesión de variables aleatorias independientes e idénticamente distribuidas (i.i.d) y no existe una correlación entre ellas. Debido a la simetría del proceso, la probabilidad de encontrar una trayectoria  $a \in S$  en la muestra, debería ser igual a la probabilidad de encontrar la trayectoria opuesta  $\bar{a} \in S$ . De aquí que la probabilidad de encontrar una racha de subida, sea igual a encontrar una racha de bajada. Obsérvese que lo anterior se cumple para un proceso estocástico reversible. Bajo la hipótesis de que la serie de tiempo de los rendimientos logarítmicos es una serie de tiempo reversible, entonces se espera que la distribución de las rachas de subida, sea estadísticamente igual a la distribución de rachas de bajada. Pero veamos que sucede en una serie de tiempo real: la Figura 3.15 muestra la distribución de la duración en escala semilogarítmica de las rachas de subida y bajada para el tercer trimestre de 2019 de la serie de tiempo del Bitcoin. Podemos observar que las distribución de rachas de subida (barras verdes) es muy similar a la distribución de las rachas de bajada (barras rojas), pero observemos el caso de las rachas de 6 minutos de duración, ambas distribuciones parecen desfasarse e incluso en las rachas de 7 minutos de duración ni siquiera observamos rachas de subida. Dada la observación anterior, nos preguntamos, ¿esta diferencia es significativa y al ser ambas distribuciones distintas, en realidad se tiene un proceso reversible? Para medir la diferencia entre la distribución de subida  $P_{\uparrow}$  y de bajada  $P_{\downarrow}$ , utilizamos la divergencia de Kullback-Leibler dada en la definición 2.3.4.

### 3.7.1. Obtención de muestras.

Para poder llevar a cabo el análisis de manera efectiva, decidimos efectuar una extracción de datos en intervalos trimestrales. Este proceso se realizó utilizando una estrategia

de ventana de deslizamiento de diez días. Este enfoque particular se adoptó deliberadamente con la intención de poder recopilar una cantidad significativamente grande de muestras para el análisis, lo que a su vez nos permitiría obtener conclusiones más precisas y detalladas. Consulte la Figura 3.16 para obtener una representación visual de cómo se llevó a cabo este proceso. Para el análisis en cuestión, optamos por recurrir a las bases de datos que se indican en la tabla 3.11. Sin embargo, antes de hacer uso de estos datos, aplicamos el tratamiento previamente descrito.

Realizamos el mismo análisis con un conjunto de 500 series de tiempo de datos sustitutos, cada una de longitud  $t = 284$  (vea la fig. 3.17), para reforzar la confiabilidad estadística de nuestros resultados. En el siguiente capítulo, presentamos las figuras con nuestros hallazgos. Las curvas que representan las estimaciones de la medida de irreversibilidad para los datos sustitutos en los resultados son en realidad el promedio  $\mu_s$ , que, como recordará, detallamos en la sección 3.4 y se muestra en la ecuación 3.39.

### 3.8. Análisis de eficiencia.

Para concluir, y con el objetivo de contrastar y comparar los resultados obtenidos previamente, hemos llevado a cabo un análisis de eficiencia aplicado a las criptomonedas que se mencionaron anteriormente. Este análisis de eficiencia se basa en un método que ha sido propuesto y descrito por Brouty y Garcin en la referencia [39]. A continuación una descripción detallada del procedimiento que seguimos para llevar a cabo este análisis de eficiencia.

Comenzamos un un conjunto de  $n + 1$  rendimientos logarítmicos consecutivos denotados por  $\{\ell_0, \ell_1, \dots, \ell_n\}$ . La serie se codifica de manera binaria como sigue:

$$a_i = \begin{cases} 1 & \text{if } \ell_i > 0, \\ 0 & \text{if } \ell_i \leq 0. \end{cases} \quad (3.50)$$

suponiendo que  $\{a_i : 1 \leq i \leq n\}$  representa una realización finita del proceso estocástico  $\{X_t : t \in \mathbb{N}_0\}$ , la probabilidad de obtener una trayectoria  $\mathbf{a} = (a_1, a_1, \dots, a_L)$  de longitud  $L$  y espacio de estados ( $a_i \in \{0, 1\}$ ) se puede expresar de la siguiente manera,

$$p(\mathbf{a}) := \mathbb{P}((X_1, X_2, \dots, X_L) = \mathbf{a}). \quad (3.51)$$

En nuestro caso tomamos un bloque de tamaño  $L = 3$  y calculamos la probabilidad de ocurrencia de cada bloque. Calculamos la entropía conjunta para un vector aleatorio  $(X_1, X_2, \dots, X_L)$  como se menciona en [39].

$$H^L = H(X_1, X_2, \dots, X_L) - \sum_a p(a) \log_2(p(a)). \quad (3.52)$$



De acuerdo con Brouty y Garcin [39], al tomar un bloque de tamaño  $L$ ,  $a = (X_1, X_2, \dots, X_L)$  si el mercado es ineficiente podemos predecir el siguiente elemento en la sucesión  $X_{L+1}$  con la información del bloque de tamaño  $L$ . Se define  $X_{L+1}$  como una variable aleatoria tipo Bernoulli, como un parámetro  $\pi_i^L \in [0, 1]$ . Entonces la sucesión completa  $(X_1, X_2, \dots, X_L, X_{L+1})$  tomará los valores  $(a, 1)$  con probabilidad  $p_i^L \pi_i^L$  o bien tomará los valores  $(a, 0)$  con probabilidad  $p_i^L(1 - \pi_i^L)$ . La entropía conjunta para el bloque anterior puede calcularse de la siguiente manera

$$H^{L+1} = -\sum_{i=1}^{2^L} p_i^L \pi_i^L \log_2(p_i^L \pi_i^L) + p_i^L(1 - \pi_i^L) \log_2(p_i^L(1 - \pi_i^L)). \quad (3.53)$$

Si el mercado es eficiente, el elemento  $X_{L+1}$  es independiente de la historia pasada y la probabilidad de ocurrencia de  $X_{L+1}$  es entonces  $\pi_i^L = \frac{1}{2}$  y entonces se tiene:

$$H_*^{L+1} = -\sum_{i=1}^{2^L} p_i^L \log_2\left(\frac{p_i^L}{2}\right) = 1 + H^L. \quad (3.54)$$

Brouty y Garcin proponen el índice de ineficiencia siguiente

$$I_*^{L+1} = H_*^{L+1} - H^{L+1}. \quad (3.55)$$

De lo anterior se tiene el siguiente resultado, propuesto en la referencia [39].

**Teorema 3.8.1.** *Para  $L \in \mathbb{N}^*$ , se tiene que  $I_*^{L+1} \geq 0$ . Más aún,*

$$I_*^{L+1} = 0 \iff \forall i \in [1, 2^L] \setminus \{j | p_j^L = 0\}, \pi_i^L = \frac{1}{2}. \quad (3.56)$$

El teorema anterior nos asegura que dicho estimador solamente será igual a cero siempre y cuando el mercado sea eficiente. Utilizando este estimador en datos reales, es altamente probable que no obtengamos un cero exacto, pero si, al aplicar dicho estimador a los datos podemos concluir que es por lo menos estadísticamente cero, podríamos deducir que dicho mercado podría ser eficiente.

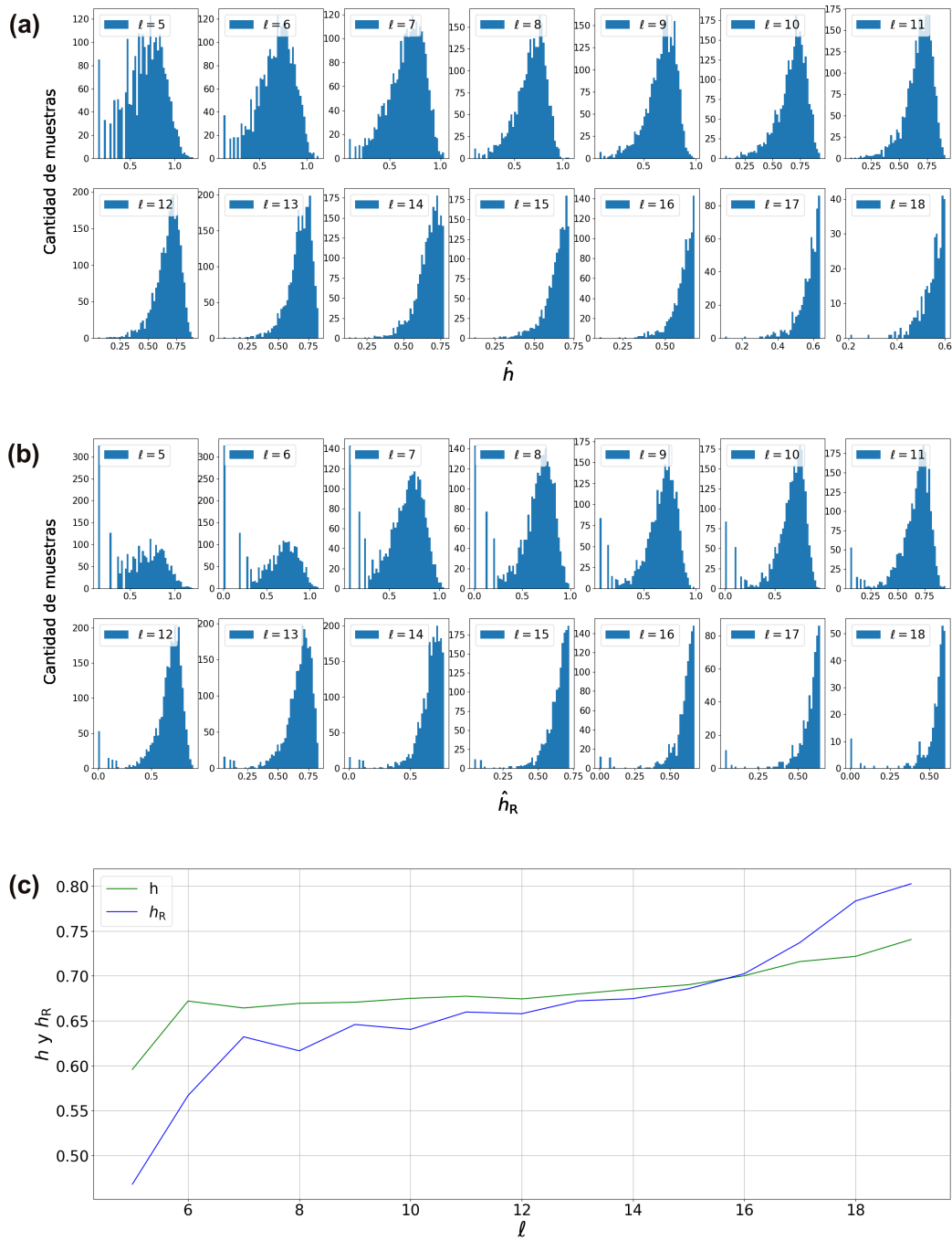


Figura 3.5: Tasas de entropía directa y reversa para  $p = 0.5$ . El panel (a) muestra la distribución de la tasa de entropía directa y el panel (b) la tasa de entropía reversa, ambas calculadas mediante el tiempo de retorno, para diferentes valores de  $\ell$ . En el panel (c) se presentan la tasa de entropía directa y reversa calculadas usando tiempos de retorno. Las líneas verde y azul representan la tasa de entropía directa y reversa respectivamente.

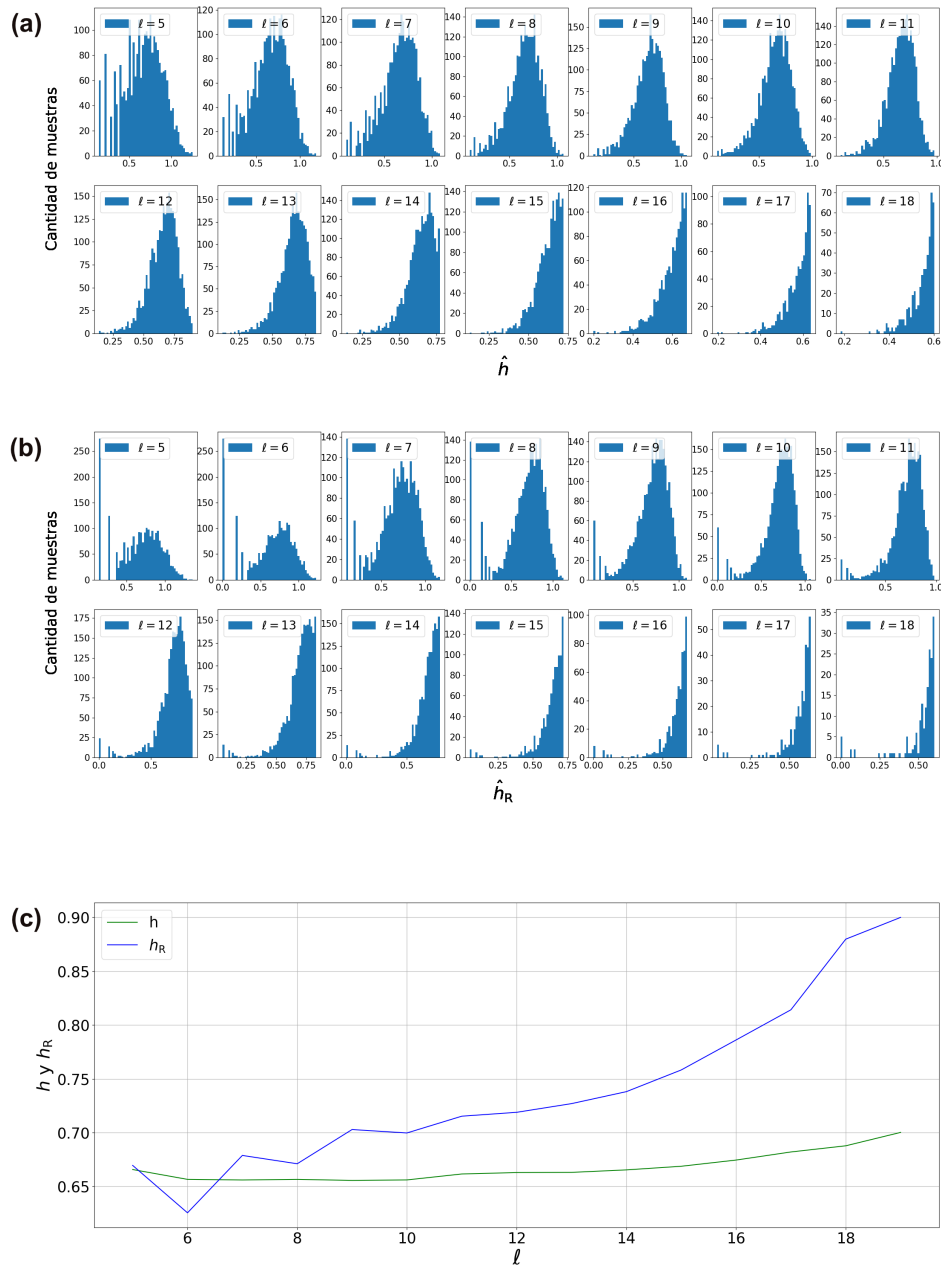


Figura 3.6: Tasas de entropía directa y reversa para  $p = 0.6$ . En el panel (a) observamos la distribución de la tasa de entropía directa mientras que la distribución para la tasa de entropía reversa se muestra en el panel b, ambas calculadas por medio del tiempo de retorno, para distintos valores de  $\ell$ . En el panel (c) se presentan la tasa de entropía directa y reversa calculadas usando tiempos de retorno. Las líneas verde y azul representan la tasa de entropía directa y reversa respectivamente.

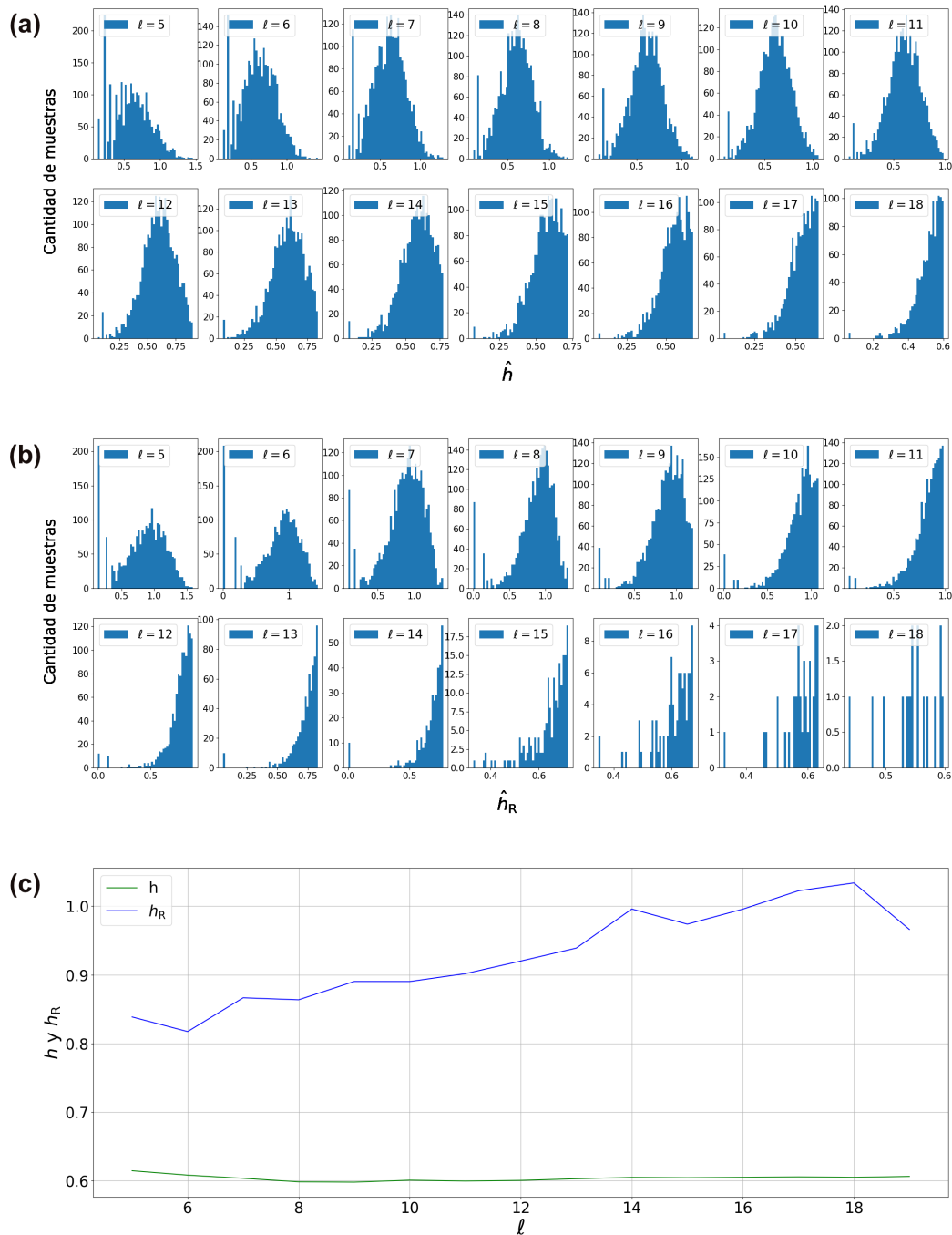


Figura 3.7: Tasas de entropía directa y reversa para  $p = 0.7$  calculada por medio del tiempo de retorno. En el panel (a) y (b) observamos la tasa de entropía directa y reversa respectivamente, para distintos valores de  $\ell$ . En el panel (c) se muestra Tasa de entropía directa y reversa. Las líneas verde y azul representan la tasa de entropía directa y reversa respectivamente. En este caso para ningún valor de  $\ell$  los estimadores coinciden. Este es un claro ejemplo de proceso irreversible.

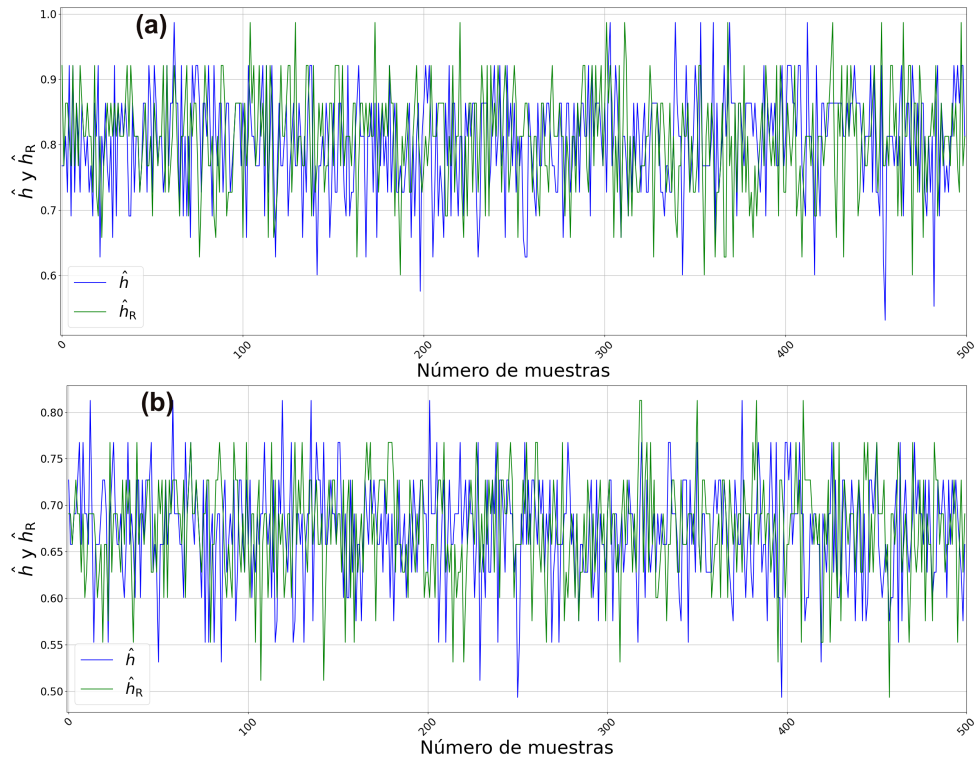


Figura 3.8: Estimación de la tasa de entropía directa (línea azul) y reversa (línea verde) con  $p = 0.5$ , y de tamaño  $t = 100,000$  y  $t = 1,000,000$  (panel a) (panel b). Podemos observar que no es posible distinguir entre ambos procesos, por lo que se puede suponer que se trata de un proceso reversible.

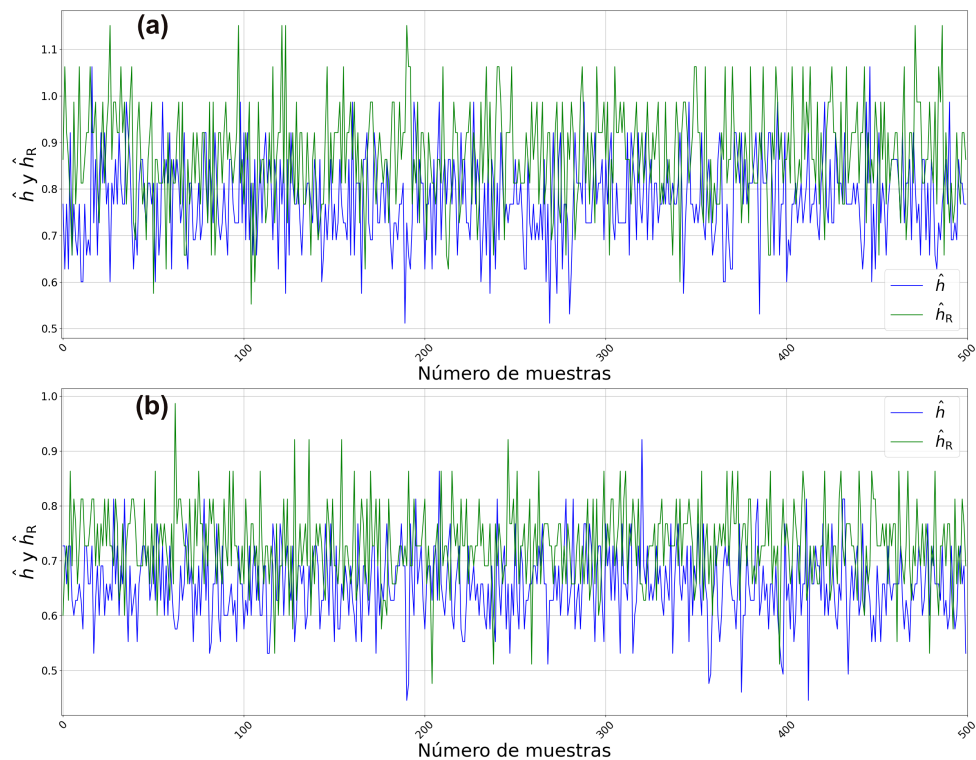


Figura 3.9: Estimación de tasas de entropía directa (línea azul) y reversa (línea verde) con  $p = 0.6$ , y una serie de tiempo de tamaño  $t = 100,000$  (panel a) y  $t = 1,000,000$  (panel b). En estos resultados podemos observar un ligero desfase entre ambos procesos con lo cual se puede suponer que se tiene un proceso irreversible.

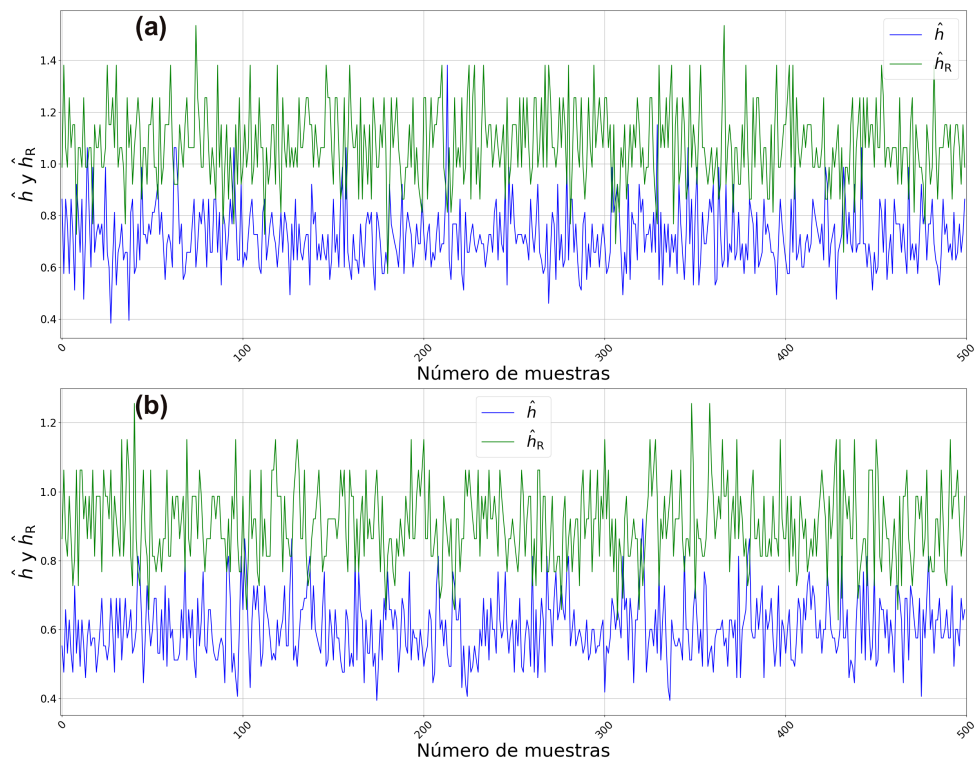


Figura 3.10: Estimación de tasas de entropía directa (línea azul) y reversa (línea verde) con  $p = 0.7$ , y una serie de tiempo de tamaño  $t = 100,000$  (panel a) y  $t = 1,000,000$  (panel b). En este caso sí podemos distinguir entre ambos procesos con lo cual se puede suponer que se tiene un proceso irreversible.

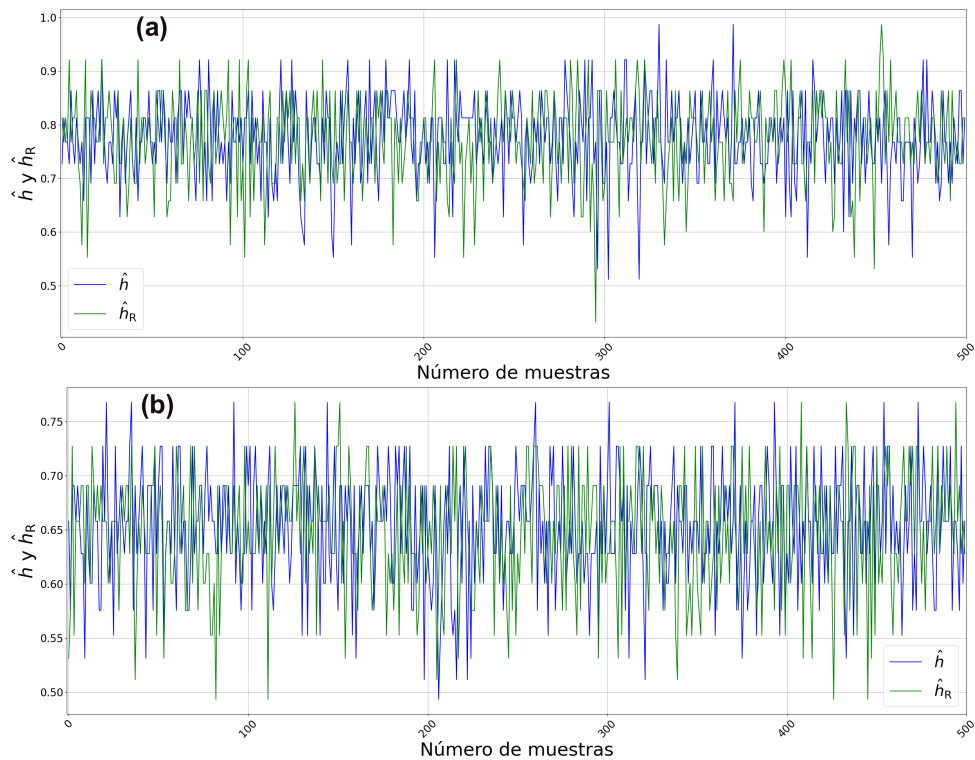


Figura 3.11: Estimación de tasas de entropía datos aleatorios con  $t = 100,000$  (panel a) y  $t = 1,000,000$  (panel b). Los resultados muestran que la serie de tiempo analizada, corresponde a un proceso reversible.

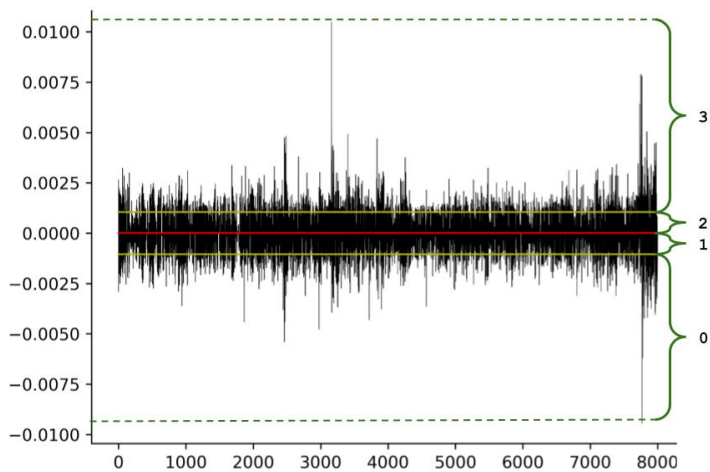


Figura 3.12: Esquema que muestra el número asignado de acuerdo a la posición del rendimiento logarítmico. La línea roja representa la media  $\mu$  y las líneas amarillas la varianza negativa  $-\sigma$  y positiva  $\sigma$  respectivamente.



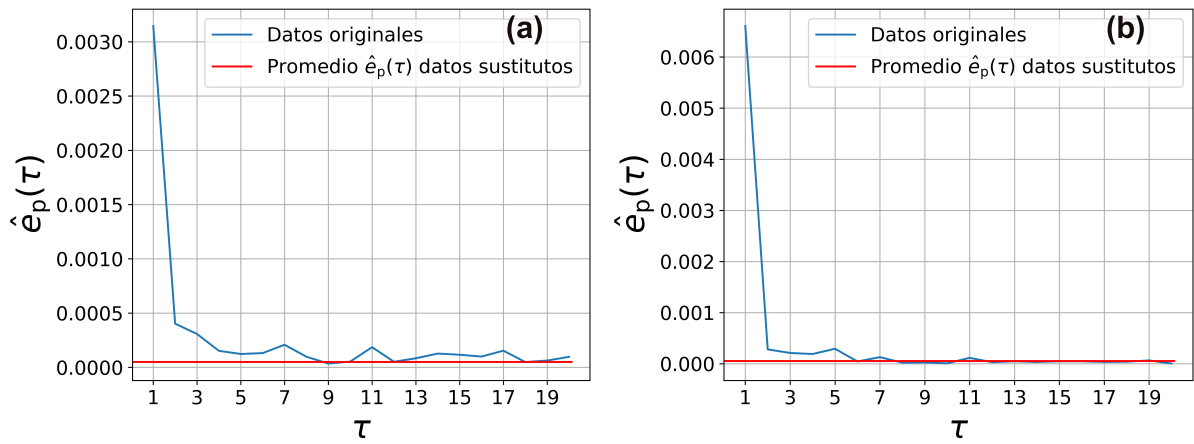


Figura 3.13: Índice de asimetría desfasada para (a) el segundo trimestre de 2017 y (b) el tercer trimestre de 2017, ambos procedentes de la serie de tiempo de Bitcoin. En las figuras podemos observar que aunque el comportamiento de el índice de asimetría desfasado es como se esperaba, en cada trimestre se observa un área bajo la curva de dicho índice distinta. Esta área nos ayudará a determinar el grado de irreversibilidad en cada periodo.



Figura 3.14: Esquema de desplazamiento de ventana para  $\tau = 3$ . En el panel (a) observamos que para tomar la primera muestra, nos situamos en el primer elemento de la serie el cual es 1 y después de avanzar con una ventana de tamaño 3 el dato encontrado es 2. Para la siguiente muestra, (panel (b)) comenzamos en el segundo elemento de la serie de tiempo, el número 3, y después de tomar una ventana de desplazamiento de tamaño 3, el dato encontrado es 3.

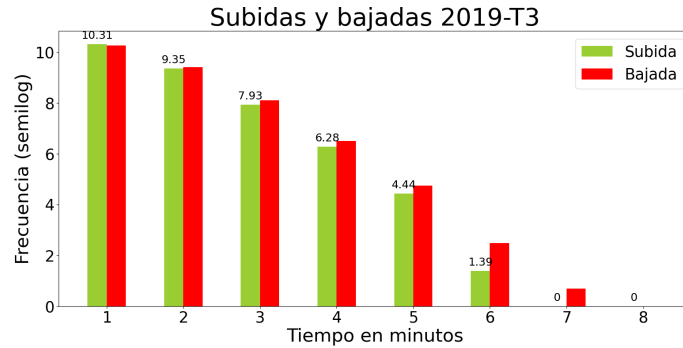


Figura 3.15: Distribución de la duración en minutos, de rachas de subida y bajada (en escala semilogarítmica). Los datos mostrados corresponden al tercer trimestre del 2019 para la serie de tiempo de Bitcoin. Las barras verdes y rojas representan la duración de las rachas de subida, y bajada respectivamente. Si el proceso fuera reversible, esperaríamos un comportamiento estadísticamente similar entre ambas distribuciones. En la figura podemos observar que esto no se cumple en todos los casos; véase el comportamiento en las barras de duración 3 minutos en adelante.

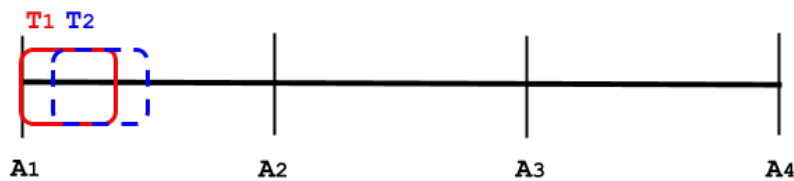


Figura 3.16: Obtención de muestras. En la figura se esquematiza la manera en la que se realizó el corrimiento. Cada cuadro representa un trimestre y la separación simboliza 10 días.

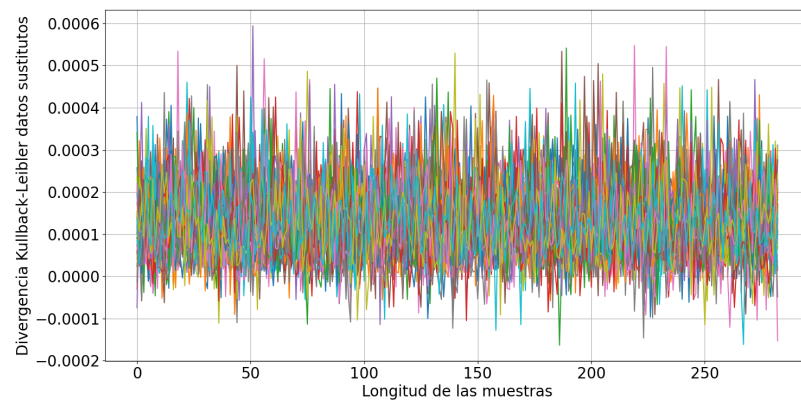


Figura 3.17: Ensemble de datos sustitutos de la serie de tiempo de Bitcoin. Cada curva representa el cálculo de la divergencia Kullback-Leibler para cada realización de datos sustitutos de longitud 284. En las curvas de la sección 4 se utilizó el promedio sobre las 500 realizaciones y se muestra con una línea horizontal.



# Capítulo 4

## Resultados.

En el presente capítulo, vamos a profundizar en los resultados que se obtuvieron de cada uno de los análisis que se mencionaron previamente en el capítulo anterior. Dirigiremos nuestra atención principalmente hacia la serie de tiempo correspondiente al Bitcoin en el contexto de los análisis que se llevaron a cabo utilizando los tiempos de retorno y los *matching times*. Esto se debe a que, en su momento, no pudimos llegar a conclusiones definitivas con respecto a estos análisis. No obstante, hemos decidido incluirlos aquí para poder realizar una comparación efectiva con los otros métodos que hemos utilizado.

Además, vamos a revisar los datos obtenidos del índice de asimetría desfasada, enfocándonos en el análisis de Bitcoin y Ether. También veremos los resultados del análisis de irreversibilidad, que se hizo usando el enfoque de rachas, ya que resultó ser más eficaz en comparación con otros métodos probados. Para este enfoque, usamos series de tiempo de Bitcoin, Ether, XRP, Litecoin y Bitcoin Cash. Para el análisis de eficiencia, usamos las mismas series de tiempo que en el análisis de rachas.

Los datos para los análisis posteriores se tomaron de la página web *CryptoDownload* [38], y usamos el *exchange Bitstamp*. Es importante mencionar que, en todos los casos, las bases de datos originales se muestrearon en intervalos de un minuto.

### 4.1. Tiempos de recurrencia.

#### 4.1.1. Tiempo de retorno.

En nuestro análisis, hemos empleado los rendimientos logarítmicos del precio de apertura de Bitcoin, minuto a minuto. El conjunto de datos consiste en 1 millón de datos recogidos entre septiembre de 2019 y agosto de 2021. Como se mencionó en la sección 3.2, es esencial encontrar un  $\ell^*$  óptimo para el análisis. Nuestra prueba de control para el

proceso reversible indicó que un  $\ell^* = 15$  proporciona una buena estimación para las tasas de entropía  $h$  y  $h_R$ . Observamos que al aumentar el tamaño del bloque, los histogramas perdían parte de la información. En  $\ell = 15$ , se conservaba aproximadamente la mitad de las muestras no censuradas. La Figura 4.1 presenta los resultados obtenidos de la estimación de la tasa de entropía directa (mostrada en el panel (a)) y la tasa de entropía reversa (mostrada en el panel (b)), ambas aplicadas a nuestra serie de tiempo. En esta figura, podemos observar de manera clara y detallada los resultados obtenidos. Además, en el panel (a) podemos observar que cuando  $\ell = 11$ , el histograma conserva aproximadamente la mitad de las muestras censuradas. Sin embargo, cuando  $\ell = 14$ , el histograma pierde su forma característica acampanada, lo que es una observación importante a tomar en cuenta. A pesar de esto, logramos obtener una buena estimación de la tasa de entropía directa, que se encuentra alrededor de 0.9, como se puede observar en el subpanel donde  $\ell = 11$ . Por otro lado, en el panel (b) observamos un escenario diferente. En este caso, las distribuciones de la tasa de entropía reversa no adoptan una forma acampanada, lo que nos impide obtener una estimación estadísticamente significativa para esta tasa. Esta limitación nos impide estimar la tasa de producción de entropía  $e_p$  y, por lo tanto, no podemos determinar el grado de irreversibilidad de la serie de tiempo en cuestión. Basándonos en estas observaciones, concluimos que el uso de los tiempos de retorno no nos proporcionaría una buena estimación para las tasas de entropía directa y reversa. Debido a esto, tomamos la decisión de descartar este método en particular para nuestro análisis.

#### 4.1.2. *Matching times.*

En esta sección, vamos a compartir los resultados que hemos obtenido al aplicar la técnica de *matching times* a la serie de tiempo del Bitcoin. Esta serie de tiempo cubre un periodo de cuatro años, desde el año 2017 hasta el 2021. Cabe mencionar que este análisis se llevó a cabo durante el año 2021, por lo que no incluye datos completos de este mismo año. Más adelante, entraremos en detalles sobre esta particularidad. Para aumentar la confiabilidad estadística de nuestros resultados, empleamos dos estrategias de muestreo diferentes. El primer método de muestreo comprendió periodos de diez días, con una ventana móvil de un día para la siguiente muestra. En el segundo método de muestreo, consideramos periodos más largos, de un trimestre, y utilizamos una ventana móvil de una semana. Esto resultó en una menor superposición entre las muestras sucesivas, lo cual es un aspecto importante a tener en cuenta. Para analizar los rendimientos logarítmicos, utilizamos dos tipos de codificaciones, la binaria y la ternaria. Sin embargo, los resultados que presentamos en este informe corresponden a la codificación binaria. Más adelante,

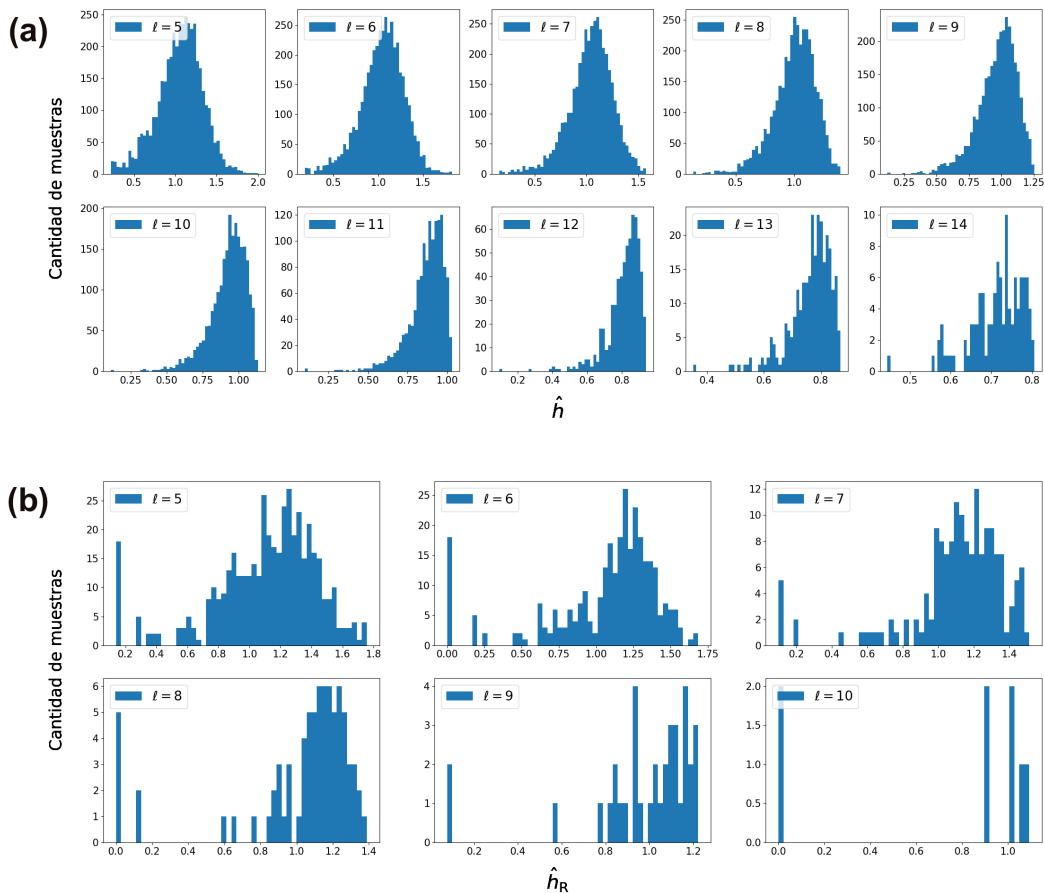


Figura 4.1: Estimación de  $h$  y  $h_R$  (panel (a) y (b) respectivamente) para los distintos valores de  $\ell$  utilizando 1 millón de datos. En el panel a observamos que para  $\ell = 12$  el histograma empieza a perder forma, mientras que en el panel b, observamos que en ningún momento se presenta una distribución semejante a una normal como se esperaba.

vamos a explicar las razones detrás de esta elección. A continuación, vamos a describir con más detalle el procedimiento que seguimos para llevar a cabo este análisis, proporcionando una visión más profunda de nuestra metodología.

- Para el primer muestreo, se utilizaron y extrajeron los precios de apertura por cada diez días. Se tomó una ventana de deslizamiento de un día y nuevamente se extrajeron los datos de los diez días subsecuentes y se repitió el proceso. Con lo cual se obtuvieron  $t$  subsecuencias de longitud  $m$  con  $t = 14,400$  y  $m = 355$ .
- Para el segundo muestreo, se utilizaron y extrajeron los precios de apertura por trimestre. Se tomó una ventana de deslizamiento de una semana, se tomó el trimestre

siguiente y se repitió el proceso. De esta forma se obtuvieron  $t$  subsecuencias de tamaño  $m$ , con  $t = 120,960$  y  $m = 41$ .

- Se realizó la conversión a rendimientos logarítmicos.
- Se realizaron dos tipos de codificación: binaria y ternaria, obteniendo mejores resultados con la primera.
- Se calcularon las tasas de entropía utilizando la metodología descrita anteriormente para los *matching times*.
- Se graficaron los resultados para las tasas de entropía directa y reversa.

### **Análisis por cada diez días.**

Comenzamos presentando los resultados del análisis que se llevó a cabo en periodos de diez días, utilizando una ventana de deslizamiento de un día. En la tabla 4.1 se muestran los valores obtenidos para  $h$  y  $h_R$  (ambos redondeados a 4 cifras) en el análisis de los rendimientos logarítmicos de los precios de apertura de Bitcoin durante el periodo de cinco años que va desde 2017 hasta 2021. Es importante señalar que los valores de  $m$ , que corresponden al número de muestras obtenidas de las series de tiempo, varían para los años 2020 y 2021. Esta variación se debe a que la base de datos de 2020 contenía ciertos datos faltantes y, dado que este estudio se realizó durante el año 2021, la base de datos para ese año aún no estaba completa al momento de la investigación. En la columna  $t$  se indica la cantidad total de datos utilizados cada año, ascendiendo a un total de 14,400 datos. Mientras tanto, la columna  $e_p$  muestra los resultados de la estimación de la tasa de producción de entropía para cada año. Aunque en teoría esta tasa debería ser estrictamente positiva, en la práctica se obtuvo un valor negativo debido a un posible error de estimación. Sin embargo, es importante destacar que estos valores son prácticamente cero, por lo que no representan una gran desviación de lo que se esperaría. Esto nos lleva a concluir que la serie de rendimientos logarítmicos de Bitcoin de 2017 a 2021 puede considerarse un proceso reversible.

En la Figura 4.2 se muestran la tasa de producción de entropía directa y reversa estimada (línea azul y verde respectivamente) para los años 2017 a 2021. Podemos observar que ambas gráficas se mantienen muy cerca la una de la otra. Estas curvas pueden compararse con las curvas expuestas en el capítulo 3 para el caso en donde se tenía una cadena de Markov con  $p = 0.5$  el cual era un proceso reversible.



Año	m	t	$h^*$	$h_R^*$	$e_p^*$
2017	355	14400	0.6039	0.5946	-0.0093
2018	355	14400	0.6282	0.6276	-0.0006
2019	355	14400	0.6216	0.6232	0.0016
2020	329	14400	0.6303	0.6244	-0.0059
2021	310	14400	0.6367	0.6267	-0.0099

Tabla 4.1: Estimación para  $h$  y  $h_R$  para las series de tiempo del Bitcoin de 2017 a 2021, con muestreo a diez días y ventana de deslizamiento de un día. Valores teóricos  $h \approx 0.6931$ ,  $h_R \approx 0.6931$ . De acuerdo con la estimación de  $e_p$ , podría decirse que las series de tiempo son reversibles. \*Los valores son aproximados a cuatro dígitos.

### Análisis trimestral.

La tabla 4.2 muestra los valores obtenidos para  $h$  y  $h_R$  al tomar periodos trimestrales con una ventana de deslizamiento semanal. Como recordatorio, se utilizaron los rendimientos logarítmicos del precio de apertura de Bitcoin, con una frecuencia de un minuto. Observamos una variación en la columna  $m$ , que representa la cantidad de trimestres obtenidos en el año, en 2020 y 2021, por las mismas razones ya mencionadas. Cada serie temporal trimestral, de longitud  $t$ , contiene 120,960 datos (rendimientos minuto a minuto). La columna  $h$  muestra los valores estimados para la tasa de entropía de cada año, mientras que  $h_R$  muestra los valores estimados para la tasa de entropía reversa. La columna  $e_p$  indica la estimación para la tasa de producción de entropía, donde, en algunos casos, se puede observar una variación mayor que en el análisis realizado con periodos de diez días. Por ejemplo, en 2020, la estimación de  $\hat{e}_p = -0.0321$  y en 2021, la estimación de  $\hat{e}_p = -0.0357$ . Aunque las tasas de producción de entropía no son tan pequeñas como en la estimación anterior, siguen siendo cercanas a cero, lo que nos indica de nuevo que podríamos estar ante un proceso reversible. Sin embargo, al examinar los resultados gráficamente, la Figura 4.3, que representa la estimación de la tasa de entropía directa  $h$  (línea azul) y reversa  $h_R$  (línea verde), nos sugiere algo diferente. En estas gráficas, se aprecian periodos en los que ambas líneas se separan, lo que podría indicar un cierto grado de irreversibilidad. Un ejemplo de esto se puede observar en la gráfica del año 2020, en el trimestre número 21 (marcado como  $T21$  en el eje  $X$ ), donde las gráficas parecen ir en sentidos contrarios. Si comparamos los resultados, vemos que mientras que las estimaciones realizadas con el primer muestreo parecen indicar que el proceso no presenta irreversibilidad, con el segundo muestreo sí se observa un grado de irreversibilidad. Por

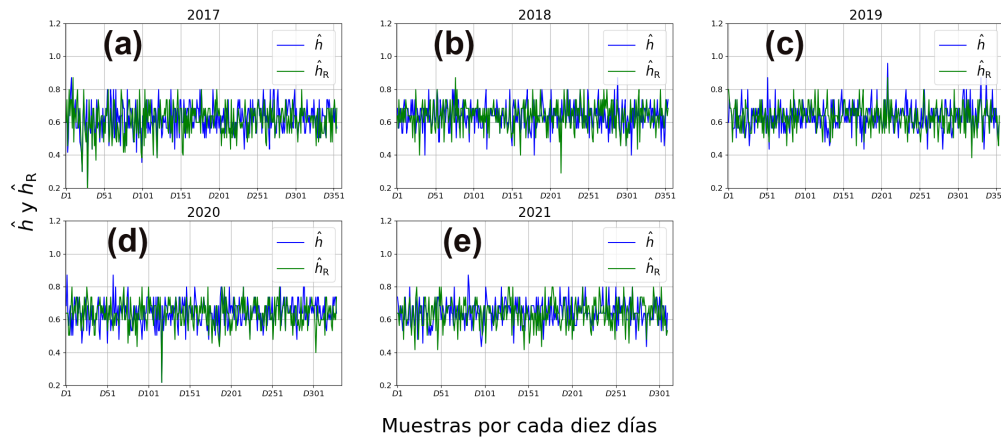


Figura 4.2: Análisis periodo diez días con ventana de deslizamiento de un día, utilizando los rendimientos logarítmicos del precio de apertura de los datos. Las gráficas representan la estimación de la tasa de entropía directa  $\hat{h}$  y reversa  $\hat{h}_R$  (línea azul y verde) respectivamente. Se observa que los datos no presentan un comportamiento de irreversibilidad.

lo tanto, concluimos que el segundo muestreo parece ser más sensible para detectar la irreversibilidad de la serie de tiempo.

Año	m	t	$h^*$	$h_R^*$	$e_p^*$
2017	41	120960	0.5931	0.5895	-0.0036
2018	41	120960	0.6475	0.6528	-0.0053
2019	41	120960	0.6355	0.6339	0.0017
2020	37	120960	0.6894	0.7216	-0.0321
2021	20	120960	0.5723	0.6080	-0.0357

Tabla 4.2: Estimación para  $h$  y  $h_R$  para las series de tiempo del Bitcoin de 2017 a 2021, con muestreo trimestral con ventana de deslizamiento de una semana. Valores teóricos  $h \approx 0.6931$ ,  $h_R \approx 0.6931$ . En este caso, a pesar de que los valores obtenidos son cercanos a cero, no son tan pequeños como en la tabla 4.1. Véase el valor de  $e_p$  en 2020 y 2021.

\*Los valores son aproximados a cuatro dígitos.

En la descripción del análisis, señalamos que aplicamos dos tipos de codificación a los datos de rendimientos logarítmicos: binaria y ternaria. En la codificación binaria, sólo consideramos si el rendimiento era negativo, en cuyo caso lo reemplazamos por  $a$ , o positivo, reemplazado por  $b$ . Para la codificación ternaria, agregamos una tercera categoría para los rendimientos que eran cero, reemplazados por  $c$ . Tomamos en cuenta este último

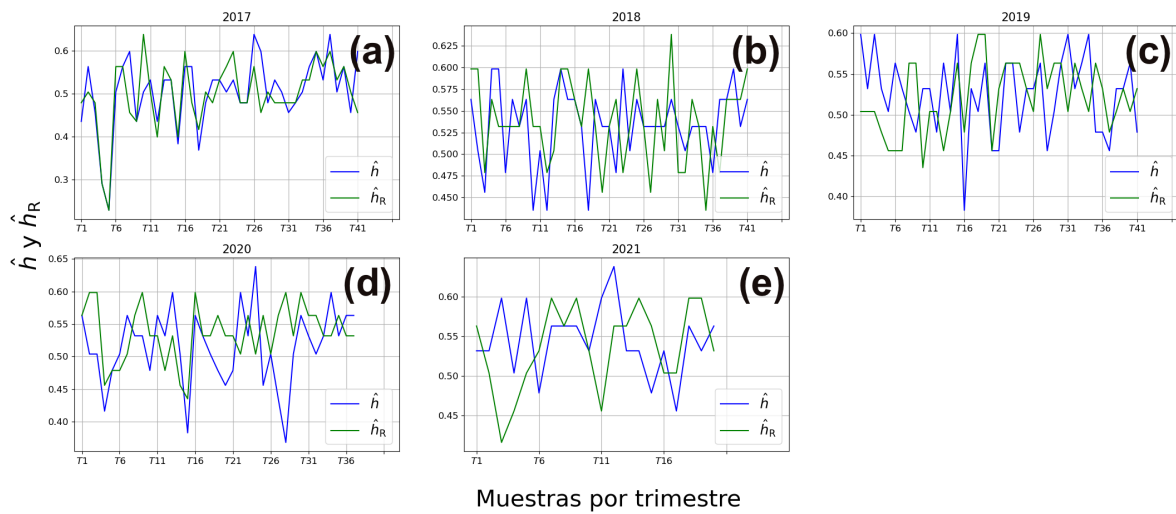


Figura 4.3: Análisis trimestral con ventana de deslizamiento de una semana utilizando la codificación binaria. Las curvas representan la estimación de la tasa de entropía directa  $h$  y reversa  $h_R$  (línea azul y verde) respectivamente. En este caso sí podemos ver un cierto grado de irreversibilidad en las curvas, ya que podemos ver periodos en donde dichas curvas no coinciden o incluso se mueven en sentidos contrarios; véase el panel (e) correspondiente al año 2021 en trimestre 11 ( $T_{11}$ ).

caso debido a que hubo periodos en la base de datos en los que el precio no cambió, resultando en un rendimiento de cero. Es importante notar que esto podría ser debido a un error durante la recolección de datos, ya que es improbable que un precio permanezca constante durante un periodo prolongado.

Las Figura 4.4 se incluye para comparar los resultados de cada codificación y cada año analizado, utilizando un muestreo trimestral. El código de colores es el mismo que en las figuras anteriores. El panel (a) de la Figura 4.4 corresponde al año 2017. Se puede observar que la codificación ternaria muestra un superposición más notable de ambas gráficas, lo que nos llevaría a la conclusión de que el proceso puede ser reversible. Sin embargo, al observar el panel (a) de la figura 4.3, con la codificación binaria, es evidente que hay periodos de superposición, pero también momentos donde las gráficas se comportan de manera diferente. Esto se hará aún más claro en las figuras correspondientes a los años 2018 a 2021.

En el panel (b) de la Figura 4.4 correspondiente al año 2018, se puede ver con claridad lo que se mencionó antes. Al analizar la figura de la codificación ternaria, parece que el proceso tiene cierta irreversibilidad, y en algunos momentos, parece reversible. Sin embargo, en la figura de la codificación binaria, se puede observar que hay un mayor grado

de irreversibilidad durante la mayoría de los periodos. La tendencia anterior se mantiene para el año 2019, como se puede apreciar en el panel (c). El panel (d) correspondiente a 2020, la observación inicial no es tan evidente. Al examinar las gráficas, en ambos casos es aparente que no hay superposiciones como en las gráficas de 2017, pero sí se puede apreciar en la gráfica de codificación ternaria que ambas podrían tener un comportamiento correlacionado, aunque con cierto desfase. Sin embargo, al examinar la gráfica de codificación binaria, se puede notar que en ciertos periodos (como los trimestres 18, 19 y 20) ambas gráficas parecen tener comportamientos contrastantes. Esto podría indicar que en este año hay un mayor grado de irreversibilidad que puede ser detectado por ambas codificaciones. Por último, las curvas correspondientes a 2021, mostradas en el panel (c), conservan en cierta medida lo observado en los años anteriores. Como en 2020, la diferencia tal vez no es obvia y podríamos argumentar que faltan aún más datos que en el año anterior, pero incluso con esta falta, también se puede apreciar en algunos periodos (como del trimestre 3 al 5) que la contraposición de comportamientos entre ambas gráficas se mantiene. Esto podría sugerir que existe cierto grado de irreversibilidad en 2021. Hemos decidido mantener los resultados obtenidos a partir de la codificación binaria. Esta decisión se basa en que este tipo de codificación se alinea de manera más natural con la esencia de los datos que estamos manejando. Para ser más claros, los rendimientos en el mundo de las finanzas suelen ser, o bien positivos, o bien negativos, siendo muy raro que estos sean cero. Por lo tanto, la codificación binaria captura de una forma más adecuada esta característica intrínseca de los datos financieros. Al analizar estos resultados, se puede deducir que durante los años 2020 y 2021 se produjo una mayor irreversibilidad en comparación con el periodo comprendido entre 2017 y 2019. Podemos llegar a esta conclusión debido a que ambas codificaciones, tanto la binaria como la alternativa, fueron capaces de detectar dicha irreversibilidad. Esto sugiere que los movimientos financieros durante los últimos años han sido menos reversibles que en años anteriores.

## 4.2. Índice de asimetría desfasada.

Dentro de esta sección, nos proponemos presentar de manera detallada los resultados que surgieron del análisis que se realizó utilizando el índice de asimetría desfasada. Comenzaremos nuestra discusión con un análisis en profundidad de la serie temporal correspondiente al Bitcoin, seguido de una revisión de los resultados obtenidos para Ether. Como hemos mencionado anteriormente en el capítulo 3, la tarea de calcular el área debajo de la curva para estimar  $\hat{e}_p(\tau)$  es crucial. Esta operación nos proporciona una información mucho más precisa y detallada sobre el grado de irreversibilidad de la serie temporal que

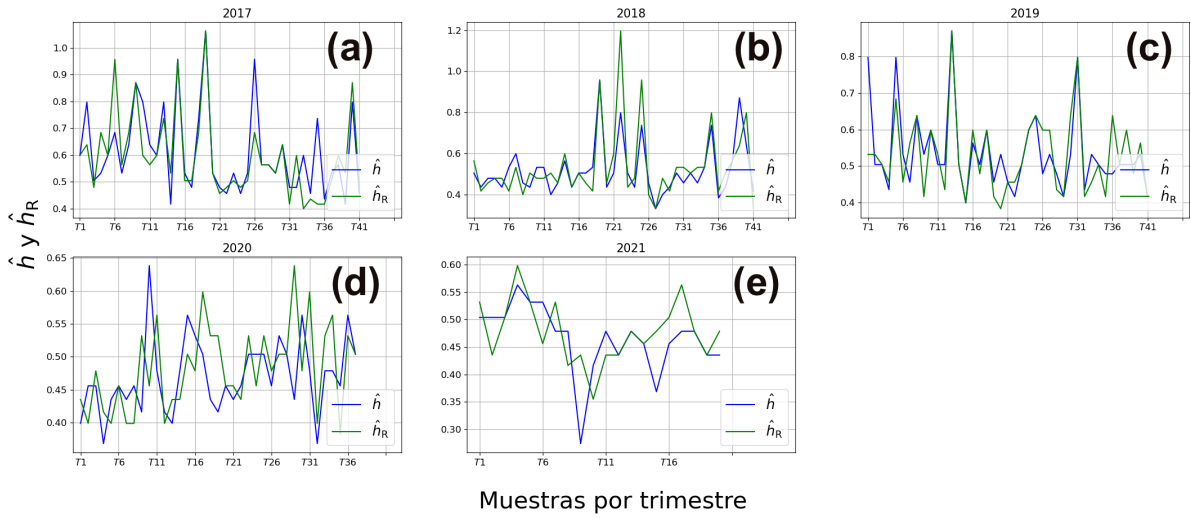


Figura 4.4: Análisis trimestral con ventana de deslizamiento de una semana y codificación ternaria. Al comparar los resultados aquí expuestos con los de la Figura 4.3, podemos notar que con la codificación binaria se obtiene mayor sensibilidad a la irreversibilidad. La excepción a lo anterior es el año 2021, mostrado en el panel (e), el cual, con ambas codificaciones presenta signos de irreversibilidad.

estamos analizando, en comparación con los análisis anteriores. Por lo tanto, también nos enfocaremos en mostrar los resultados que emergieron de este análisis específico.

La Figura 4.5 presenta los resultados para algunos trimestres que son representativos del comportamiento general. Es importante recordar que analizamos desde el año 2015 hasta el segundo trimestre del 2022, que es cuando se finalizó este análisis. La línea azul representa la estimación de  $e_p(\tau)$  para los datos de la criptomoneda en particular, mientras que la línea roja muestra el promedio de la estimación de  $\hat{e}_p(\tau)$  para los datos sustitutos. El desplazamiento de la ventana se consideró para valores  $\tau \in [1, 9]$ . Según los resultados de [44], si el área bajo la curva de  $\hat{e}_p(\tau)$  es mayor, entonces el grado de irreversibilidad también lo es. Es importante recordar que la línea roja nos proporciona una base para la significancia estadística de nuestros resultados. Por lo tanto, cualquier punto en la línea azul que esté por encima de la línea roja se puede considerar estadísticamente mayor que cero. Teniendo esto en cuenta, podemos clasificar los trimestres en la Figura 4.5 de mayor a menor irreversibilidad. Así, el panel (b) (tercer trimestre de 2017) muestra el mayor grado de irreversibilidad, seguido por el panel (c) (primer trimestre de 2021), luego el panel (a) (primer trimestre de 2015), y finalmente, el panel (d) (tercer trimestre de 2022), que sería la menos irreversible.

Para facilitar una comparación más precisa, hemos determinado el área bajo la curva entre la línea azul y la línea roja para cada trimestre, como se muestra en el panel (a) de

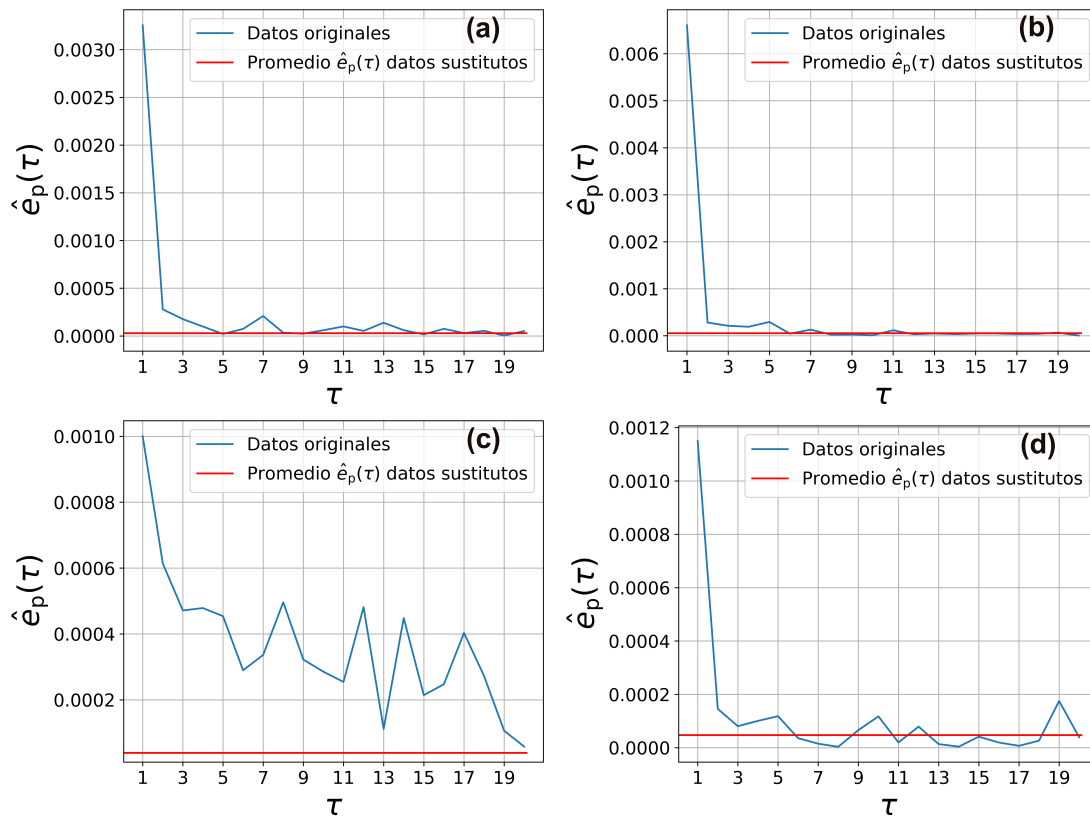


Figura 4.5: Índice de asimetría desfasada para Bitcoin en 2015, 2017, 2021 y 2022. En el panel (a) observamos los resultados del análisis para el primer trimestre del 2015, en el panel (b) el tercer trimestre de 2017, en el panel (c) el primer trimestre de 2021 y finalmente en el panel (d) el tercer trimestre de 2022. El grado de irreversibilidad puede ser estimado observando el área bajo la curva en cada gráfica.

la Figura 4.7. Adicionalmente, se observa una tendencia notable hacia la reversibilidad a medida que nos acercamos a 2022. Es posible identificar periodos de alta irreversibilidad, por ejemplo, el lapso entre 2015 y 2016, o el pico evidente en el tercer trimestre de 2017.

Hemos realizado el mismo análisis para Ether, manteniendo la misma clave de colores. En este caso, se utilizaron datos desde 2018 hasta el segundo trimestre de 2022, similar al estudio de Bitcoin. Utilizamos el mismo rango de valores para  $\tau$  que el empleado en el análisis de Bitcoin.

La Figura 4.6 muestra trimestres representativos del comportamiento general durante este periodo. Según el grado de irreversibilidad, el primer trimestre de 2021 (panel (c)) tiene el mayor grado, seguido por el primer trimestre de 2018 (panel (a)), el tercer trimestre de 2019 (panel (b)), y finalmente el cuarto trimestre de 2022 (panel (d)) con el menor grado.

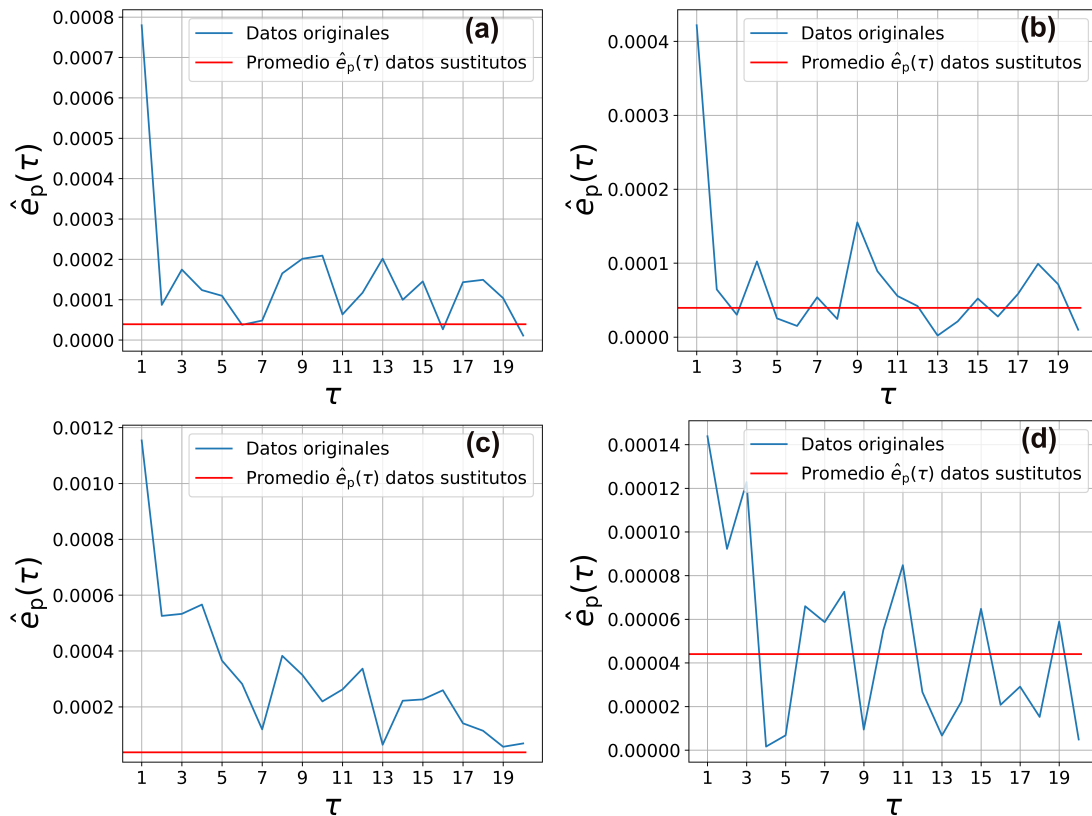


Figura 4.6: Índice de asimetría desfasada para Ether en 2018, 2019, 2021 y 2022. En el panel (a) observamos los resultados del análisis para el primer trimestre del 2018, en el panel (b) el tercer trimestre de 2019, en el panel (c) el primer trimestre de 2021 y finalmente en el panel (d) el cuarto trimestre de 2022. El grado de irreversibilidad puede ser estimado observando el área bajo la curva en cada gráfica.

Para obtener una comparación más precisa de la irreversibilidad entre cada trimestre, hemos llevado a cabo también, el cálculo del área bajo la curva de la función  $\hat{e}_p(\tau)$ . Los resultados de este cálculo se pueden apreciar claramente en la Figura 4.7 panel (b). Al examinar estos resultados, se puede observar un incremento notable en la irreversibilidad de la serie de tiempo entre el segundo trimestre del año 2020 y el cuarto trimestre del año 2021, una tendencia que parece ir disminuyendo de manera gradual conforme nos acercamos al final del año 2022.

Los resultados que se presentan en esta sección parecen ser distintos a los que mencionamos en la sección anterior. En dicha sección, utilizamos los *matching times* para llevar a cabo nuestro análisis, que se centró exclusivamente en la serie de tiempo de Bitcoin. De acuerdo con los *matching times* que se muestran en la Figura 4.3, los años 2020 y 2021 fueron los que presentaron un mayor grado de irreversibilidad.

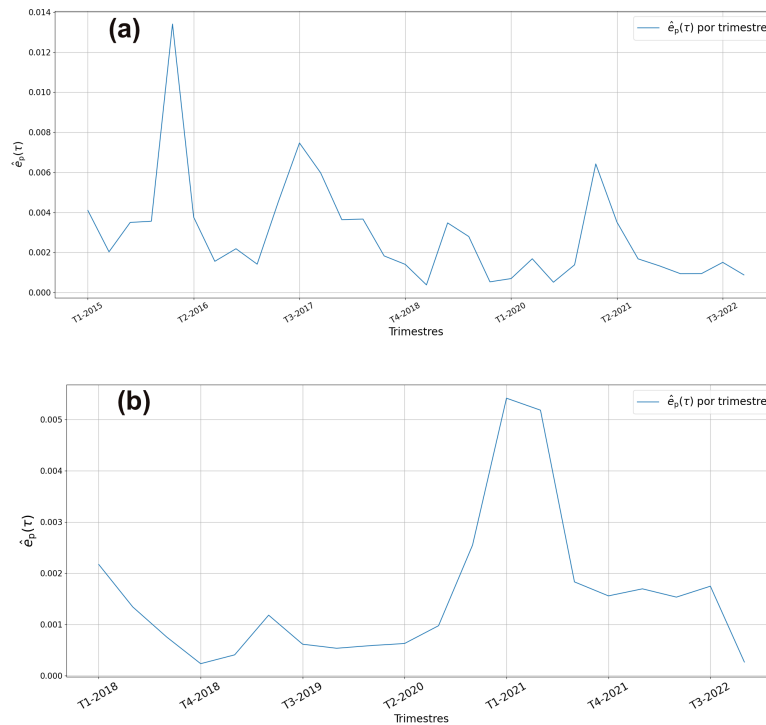


Figura 4.7: Estimación de áreas bajo la curva de  $e_p(\tau)$  para Bitcoin y Ether. En el panel (a) observamos los resultados correspondientes a los trimestres de 2015 a 2022 para Bitcoin. De acuerdo con esta gráfica podríamos concluir que la serie de tiempo del Bitcoin tiende a la reversibilidad. En el panel (b) se presentan los resultados de los trimestres de 2018 a 2022 para Ether. Podemos observar un mayor grado de irreversibilidad en el periodo comprendido entre 2020 y 2021 y, así como en el caso de Bitcoin parece tender a la eficiencia cuando nos acercamos al último trimestre de 2022.

Sin embargo, si nos enfocamos en el análisis realizado con el índice de asimetría desfasada (ilustrado en la Figura 4.7), observamos un panorama diferente. Aunque el año 2021 muestra cierta irreversibilidad, esta se ve superada por la irreversibilidad estimada para el año 2015. En el caso de 2020, la irreversibilidad es notablemente menor que la estimada para 2021.

Hasta este punto, podemos concluir que el método de los *matching times* nos brindó la posibilidad de identificar algunos periodos de irreversibilidad, pero no nos permitió cuantificarla con precisión. Por otro lado, el índice de asimetría desfasada nos proporciona un medio para realizar una comparación más clara entre los grados de irreversibilidad de diferentes periodos.

Es importante destacar que los resultados obtenidos a través de los dos métodos pueden variar, dado que el protocolo de muestreo utilizado fue diferente en cada caso. En secciones



posteriores, pondremos en contraste los resultados de este análisis con los obtenidos a través del análisis de rachas, lo que nos permitirá tener una visión más completa de la situación.

### 4.3. Análisis de rachas para irreversibilidad e ineficiencia.

Como señalamos previamente, este análisis resultó ser el más exitoso que hemos llevado a cabo hasta ahora. Las conclusiones obtenidas de este estudio nos proporcionaron una comprensión más profunda de los periodos de irreversibilidad en las series de tiempo que analizamos. Este análisis, que utilizó el análisis de rachas, nos permitió identificar y entender mejor estos periodos de irreversibilidad. Además de esto, también pudimos examinar la ineficiencia de las series de tiempo, según el índice propuesto Brouty y Garcin, en [39]. Este análisis se llevó a cabo utilizando cinco criptomonedas: Bitcoin, Ether, XRP, Litecoin y Bitcoin Cash. El protocolo de recolección de muestras para este análisis se describió en el capítulo anterior, específicamente en la sección 3.7. Además, para estos análisis, también se realizó un tratamiento previo de los datos. Este proceso de preparación de los datos se describió en la sección 3.6. Este tratamiento previo aseguró que los datos estuvieran completos, lo que contribuyó a la eficacia y al éxito de nuestro análisis.

Comenzamos esta sección al introducir los resultados obtenidos para Bitcoin y Ether, como se muestra claramente en la Figura 4.3. El eje  $x$  ofrece una representación cronológica de fechas representativas, presentadas trimestralmente y en formato  $mm/dd/yy$ , mientras que el eje  $y$  proporciona una visualización de los índices de irreversibilidad  $I_T$  y de eficiencia  $I_*$ . Para Bitcoin, tal como se puede apreciar en la Figura 4.3 panel (a), se identifica una clara señal de irreversibilidad e ineficiencia en el periodo comprendido entre el año 2015 y junio de 2020. No obstante, también se pueden observar periodos donde la irreversibilidad es reducida, como es el caso del intervalo de tiempo que va desde febrero de 2016 hasta marzo de 2017. Es importante destacar que el año 2020 fue un año particularmente distintivo debido al inicio de la pandemia de COVID-19. Esta situación de salud mundial puede haber sido un factor determinante en el notable incremento del índice de eficiencia durante este periodo. Al observar la Figura 4.3, se puede apreciar que esta tendencia se evidencia claramente en los picos que experimenta la curva de color azul, que se extienden desde mayo de 2019 hasta junio de 2020. Este fenómeno puede ser indicativo de cómo situaciones extraordinarias, como una pandemia, pueden tener un impacto significativo en la eficiencia de ciertos sistemas y procesos. Hacia a finales del año 2020, ambos índices parecen converger hacia el cero, un fenómeno que proporciona datos

interesantes acerca de la eficiencia del mercado. De acuerdo a nuestros resultados, parece que para Bitcoin, el mercado tiende hacia la reversibilidad y la eficiencia, coincidiendo con los hallazgos de estudios previos realizados en [45–47]. Posteriormente, observamos un repunte en la irreversibilidad hacia finales de 2022, un dato que no puede pasar desapercibido. Es importante subrayar que estos índices de irreversibilidad son estadísticamente mayores a cero con un nivel de confiabilidad del 95 %, (ver la sección sobre pruebas de hipótesis 3.4).

Podemos apreciar que durante ciertas etapas o periodos específicos, los dos índices presentan un patrón de comportamiento similar, como es el caso del periodo comprendido entre 2015 y 2016. Durante este periodo, es posible identificar dos picos que se asemejan en su comportamiento: uno corresponde a la irreversibilidad y el otro a la ineficiencia. Estos picos son seguidos por un patrón de comportamiento que, aunque similar, presenta un desfase en el tiempo, comenzando cerca del mes de diciembre del año 2015. Otra observación interesante es el pico que se forma el 15 de abril de 2018, donde se puede notar un comportamiento bastante similar entre ambas curvas. No obstante, este comportamiento correlacionado, o en otras palabras, esta tendencia de comportamiento similar entre los dos índices, no resulta ser evidente o fácilmente identificable después de esta fecha específica.

En el caso de Ether, se observa un patrón distinto (Figura 4.3 panel (b)). Durante el inicio del periodo que se analiza, se puede apreciar un alto grado de ineficiencia en su comportamiento, mientras que la irreversibilidad es menos evidente. Ambos índices, tanto la ineficiencia como la irreversibilidad, alcanzan su punto máximo durante el periodo comprendido entre 2017 y 2018. Este fenómeno tiene sentido si consideramos que Ether era un mercado emergente en 2017 y, según estudios como el de [9], estos mercados pueden presentar niveles de ineficiencia superiores a los de los mercados ya desarrollados. A partir de ese momento, se observa una disminución en ambos índices, lo que sugiere que Ether ha estado evolucionando progresivamente hacia una mayor eficiencia y reversibilidad. En general, el comportamiento de ambos índices presenta similitudes a lo largo del tiempo, lo que podría indicar una correlación entre la eficiencia y la reversibilidad en el mercado de Ether. Es importante notar que al examinar los índices de ineficiencia de Ether y Bitcoin, encontramos una diferencia significativa en su evolución desde 2019 hasta 2020. Con Bitcoin, vemos un grado de irreversibilidad menor en comparación con su índice de ineficiencia. Por otro lado, Ether mantuvo un índice de ineficiencia bastante constante durante ese mismo periodo, mientras que el índice de irreversibilidad parece haber aumentado más. Este punto es importante, ya que nos ofrece una perspectiva más detallada de cómo un mismo evento puede impactar de manera diferente a los índices de estas dos criptomonedas, lo que podría ser muy útil para comprender mejor estos conceptos.

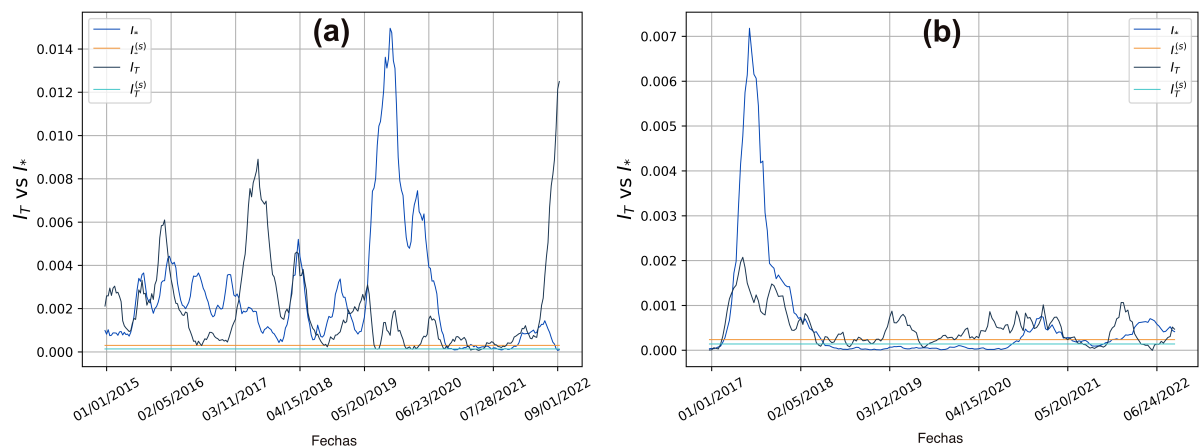


Figura 4.8: Irreversibilidad vs. eficiencia: (a) Bitcoin y (b) Ether. La línea en azul señala el índice de irreversibilidad  $I_T$ , la línea negra representa el índice de ineficiencia  $I_*$ , mientras que las líneas amarilla y azul representan los mismos índices pero para los promedios de los datos sustitutos. Ambas criptomonedas parecen tender a la reversibilidad, pero sólo Ether parece tender a la eficiencia ya que en el caso de Bitcoin, podemos observar un incremento en el índice de la ineficiencia.

En la Figura 4.9, se presentan las curvas de  $I_T$  e  $I_*$ , las cuales corresponden a las criptomonedas XRP y Litecoin. Al observar la Figura 4.9 panel (a), se puede notar un comportamiento similar al de Bitcoin en ambos índices. Este comportamiento consiste en claras señales de irreversibilidad y de ineficiencia para XRP, que parecen disminuir alrededor de agosto de 2020. Es notable que durante ciertos periodos, ambos índices muestran una gran similitud, como entre junio de 2019 y agosto de 2020. Durante este periodo, los picos de las curvas muestran un comportamiento muy parecido, pero con un ligero desfase. Este patrón podría indicar que la irreversibilidad en algunos casos parece adelantarse a la ineficiencia, un hallazgo que puede tener implicaciones importantes para la comprensión del comportamiento del mercado de criptomonedas. Al igual que las criptomonedas mencionadas anteriormente, XRP parece tener una tendencia hacia la reversibilidad y la eficiencia, lo que sugiere que esta criptomoneda puede estar avanzando hacia un estado de mayor estabilidad. En el caso de Litecoin, como se puede apreciar en el panel (b) de la Figura 4.9, la irreversibilidad y la ineficiencia son evidentes, pero no parecen disminuir al menos durante el periodo analizado (2019-2022). Este patrón podría sugerir que dicha criptomoneda aún no ha alcanzado su estabilidad en términos de eficiencia de mercado. Aunque también se puede observar una similitud menos evidente entre ambos estimadores, en el periodo comprendido entre junio de 2020 y finales de 2022, esta similitud es menos marcada que en el caso de XRP. Este hecho puede indicar que

Litecoin se comporta de manera diferente a XRP y Bitcoin en ciertos aspectos clave.

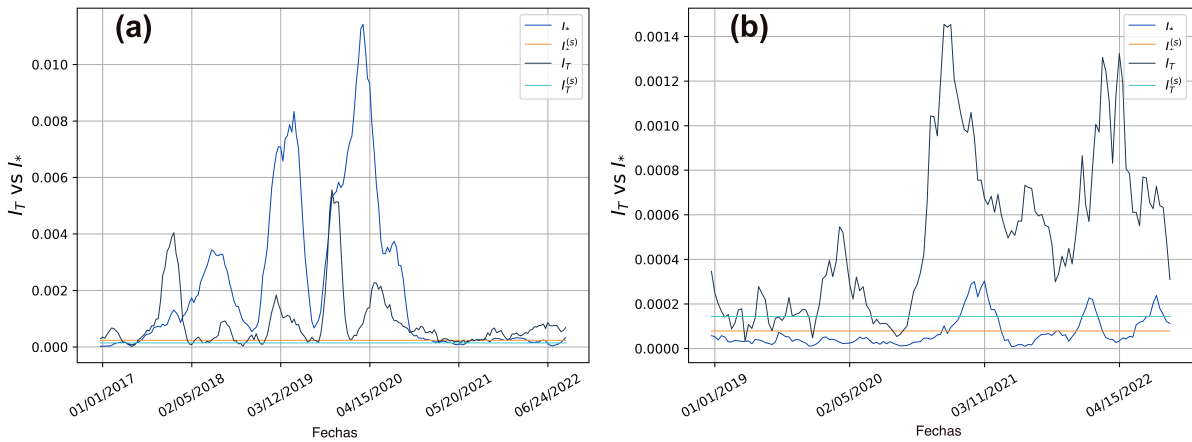


Figura 4.9: Irreversibilidad vs eficiencia: panel (a) XRP y panel (b) Litecoin. La línea azul señala el índice de irreversibilidad  $I_T$ , la línea negra representa el índice de ineficiencia  $I_*$ , mientras que las líneas amarilla y azul representan los mismos índices pero para los promedios de los datos sustitutos. En cuanto a la ineficiencia e irreversibilidad XRP, parece tender, tanto a la reversibilidad como a la eficiencia, mientras que Litecoin esta tendencia no es clara, y más bien permanece irreversible e ineficiente.

Finalmente, en la Figura 4.10, observamos los resultados de los índices de irreversibilidad e ineficiencia correspondientes a Bitcoin Cash. Se puede apreciar un patrón de comportamiento bastante similar al observado en Ether, caracterizado por un alto grado de ineficiencia que se extendió desde enero hasta septiembre de 2018. Este fenómeno podría explicarse, en parte, por el hecho de que Bitcoin Cash fue lanzado en el año 2017. A lo largo del periodo analizado, se observan incrementos en el índice de irreversibilidad que guardan cierta similitud con las fluctuaciones experimentadas por Ethereum, aunque es interesante notar que, el índice de ineficiencia no aumenta en la misma proporción, a excepción de algunos incrementos notables que se manifiestan a principios de los años 2021 y 2022. Este patrón de comportamiento guarda una gran similitud con el observado en el caso de Ether, como se puede constatar en el panel (b) de la Figura 4.3. No obstante, una característica que distingue a Bitcoin Cash de las otras cuatro criptomonedas analizadas es que parece ser la única que no muestra una posible correlación entre los comportamientos de los índices de irreversibilidad e ineficiencia.

También nos dimos a la tarea de calcular el coeficiente de correlación entre ambos estimadores, para cada una de las criptomonedas. Recordemos que el coeficiente de correlación entre dos variables aleatorias  $X$  e  $Y$ , es una medida que nos permite conocer la dependencia lineal entre dos variables y se calcula de la siguiente manera [48]

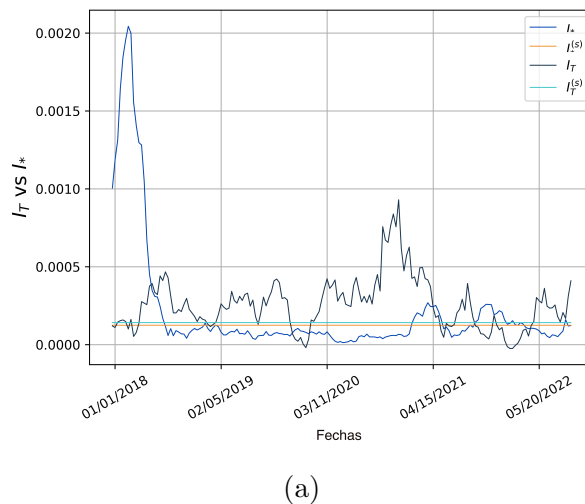


Figura 4.10: Irreversibilidad vs eficiencia Bitcoin Cash. La línea azul señala el índice de irreversibilidad  $I_T$ , la línea negra representa el índice de ineficiencia  $I_*$ , mientras que las líneas amarilla y azul representan los mismos índices pero para los promedios de los datos sustitutos. Bitcoin Cash es la única criptomoneda en la que la correlación entre ambas gráficas no es clara. Pese a lo anterior podemos observar que se mantiene ineficiente la mayor parte del tiempo y parece tender a la eficiencia, mientras que sólo es irreversible durante el 2018.

$$r_{xy} = \frac{\sum_i^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^n (x_i - \bar{x})^2 \sum_i^n (y_i - \bar{y})^2}} \quad (4.1)$$

donde  $n$  es el tamaño de la muestra,  $x_i$  y  $y_i$  son las observaciones de las variables  $X$  e  $Y$  y finalmente,  $\bar{x}$  y  $\bar{y}$  son las medias muestrales.

Los resultados obtenidos al calcular este coeficiente se muestra en la tabla 4.3. Podemos observar que la criptomoneda que presenta mayor correlación entre ambos índices en Ether, lo cual es consistente con los resultados obtenidos en la gráfica. Las siguientes criptomonedas con un índice alto de correlación serían Bitcoin Cash y XRP. Sin embargo para Bitcoin y Litecoin, los índices de correlación son bajos e incluso negativos. Lo anterior nos dice que si bien en algunos casos o periodos ambos índices parecen comportarse de la misma manera, no es así para todas las series de tiempo analizadas y mucho menos para todos los periodos, lo que nos lleva a la siguiente pregunta: ¿Por qué para algunas criptomonedas y para algunos periodos ambos índices parecen tener comportamientos similares y para otros no?

Destaca claramente que Ether tiene la mayor correlación entre los dos índices. Esto concuerda con los resultados que se observan en el panel (b) la Figura , lo que indica

un alto grado de coherencia entre los dos conjuntos de datos para esta criptomoneda. Además, Litecoin y XRP también presentan una correlación significativa, aunque no tan fuerte como Ether. Esto sugiere que hay una relación observable entre los dos índices para estas criptomonedas, aunque no es tan consistente como en el caso de Ether. Sin embargo, es importante notar que Bitcoin y Bitcoin Cash muestran correlaciones bajas e incluso negativas. Este resultado es congruente con lo que se puede observar en las figuras correspondientes, indicando que a pesar de que en algunos periodos los dos índices pueden comportarse de manera similar, no es el caso para todas las series de tiempo y periodos analizados. Estos hallazgos plantean una cuestión intrigante: ¿Por qué para algunas criptomonedas y en algunos periodos ambos índices muestran comportamientos similares y en otros no? Esto podría ser un área de investigación futura interesante, ya que entender estas diferencias podría arrojar luz sobre la dinámica subyacente de estos mercados de criptomonedas.

Nuestra expectativa inicial era descubrir una correlación entre los dos índices que nos permitiera identificar una relación entre la irreversibilidad y la ineficiencia a lo largo del tiempo. A pesar de esto, no siempre se cumplió esta expectativa. Nos encontramos con casos particulares, como Bitcoin y Bitcoin Cash, en los que no pudimos identificar ninguna correlación evidente. Esto contrasta con otros casos como Ether, Litecoin o XRP, en los cuales observamos algunas similitudes durante ciertos periodos de tiempo. Ether resultó ser particularmente interesante para nuestro estudio, ya que su comportamiento demostró ser bastante similar en casi la totalidad del periodo que analizamos. En la siguiente sección de nuestro trabajo, nos proponemos comparar estos hallazgos con los resultados que obtuvimos a través del uso del índice de asimetría desfasada.

Criptomoneda	Coefficiente de correlación
BTC	-0.1111
ETH	0.6733
XRP	0.3253
LTH	0.4611
BCH	-0.2087

Tabla 4.3: Coeficiente de Pearson entre  $I_T$  e  $I_*$  para todas las criptomonedas analizadas. Podemos observar que Ether es la criptomoneda que presenta una mayor correlación entre ambos estimadores, seguido de Litecoin y XRP. Bitcoin y Bitcoin Cash parecen no estar correlacionadas (véase fig 4.3, fig 4.9 y fig 4.10).

#### 4.4. Comparación de resultados: índice de asimetría desfasada vs rachas.

En esta sección, nos dedicaremos a hacer una comparación entre los resultados obtenidos a través del análisis de rachas y los resultados derivados del índice de asimetría desfasada y el análisis de ineficiencia. Nuestro objetivo principal es determinar si estos estimadores, nos proporcionan la misma información esencial. Anteriormente, en la sección 3.5, discutimos la importancia de calcular el área bajo la curva para  $\hat{e}_p(\tau)$ . En línea con esta idea, en la Figura 4.11, hemos presentado el área bajo la curva. Para hacerlo, hemos seleccionado un representante para cada trimestre analizado. Este representante se muestra mediante una línea roja punteada y horizontal en cada periodo correspondiente. Además, hemos incluido las curvas de la sección anterior, las cuales muestran el estimador para la irreversibilidad  $I_{\mathcal{T}}$  y el estimador para la ineficiencia  $I_*$ . Para facilitar la comparación y la coherencia, hemos mantenido los mismos colores que se usaron en la sección anterior. En el panel (a) de la Figura , podemos observar la comparación para la serie de tiempo de la Bitcoin. Resulta interesante cómo la trayectoria de la línea roja, en ciertas ocasiones, parece imitar a la línea azul que representa el estimador  $I_*$  de ineficiencia. Sin embargo, en otras ocasiones parece tener más similitud con la línea negra que denota el estimador de irreversibilidad  $I_{\mathcal{T}}$ . De manera similar, cuando nos trasladamos a la Figura panel (b) correspondiente a Ether, notamos que la interrelación entre las dos curvas es aún más notable, alternando entre similitudes con  $I_{\mathcal{T}}$  y  $I_*$ . Es importante mencionar que, a pesar de estas similitudes, la escala entre los dos estimadores es diferente, como se evidencia claramente en el eje derecho de ambas figuras.

A partir de estas observaciones, podemos concluir que, aunque no fue posible responder de manera definitiva a la pregunta de si medir la irreversibilidad es lo mismo que medir la ineficiencia en una serie temporal, sí podemos afirmar que ambos conceptos están relacionados de alguna manera, como se evidencia en la Figura 4.11. En esta figura, podemos observar que ambos estimadores para la irreversibilidad de la serie de tiempo muestran comportamientos similares con el estimador de la ineficiencia propuesto en [39]. En el futuro, sería útil y explorar y analizar la aplicación de otros estimadores para medir la ineficiencia y la irreversibilidad de una serie temporal. Estos nuevos enfoques podrían proporcionar una visión más clara del comportamiento de la serie temporal. Luego, podríamos comparar estos nuevos resultados con los que hemos obtenido en esta sección. Esta comparación podría revelar nuevas ideas o confirmar nuestras conclusiones actuales. Esto ciertamente nos ayudaría a entender mejor la relación y el comportamiento entre la irreversibilidad e ineficiencia en las series de tiempo financieras.

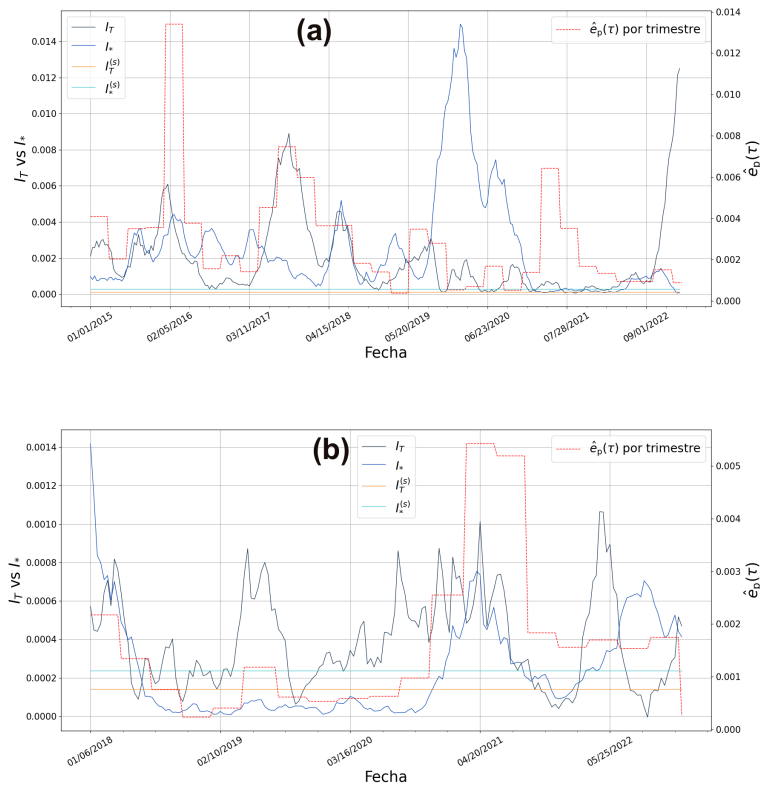


Figura 4.11: Áreas bajo la curva de  $e_p(\tau)$  para trimestres de 2015 a 2022 para Bitcoin (panel (a)) y de 2018 a 2022 para Ether (panel (b)), vs índices de irreversibilidad e ineficiencia. El código de colores se mantiene como en las gráficas anteriores. La línea roja punteada representa la estimación de  $e_p(\tau)$  por trimestre. En el lado izquierdo vemos la escala para los índices de irreversibilidad e ineficiencia, mientras que del lado derecho vemos la escala de  $e_p(\tau)$ .

## 4.5. Discusión de resultados.

En este estudio, nos encontramos con resultados bastante sorprendentes. Anticipábamos que los índices de eficiencia e irreversibilidad mostrarían un comportamiento estadísticamente similar, tal como se sugiere en [5], y esperábamos establecer una correlación entre ellos. Al concluir, quedamos con algunas preguntas no resueltas que nos gustaría explorar en el futuro. En esta sección final, ofrecemos una visión de los posibles estudios que podrían llevarse a cabo.

Al analizar los resultados del índice de asimetría desfasada, que se muestran en la Figura 4.11, se puede destacar que este índice es un estimador más eficaz para la irreversibilidad que los tiempos de retorno y los matching times. Esto se debe a que nos



brinda la capacidad de cuantificar el nivel de irreversibilidad en la serie de tiempo y nos proporciona más detalles sobre los periodos donde dicha irreversibilidad es evidente. Lo anterior también es válido para el análisis realizado a través de los patrones de tendencia mostrados en las Figuras 4.3-4.10. Sería útil profundizar en la correlación entre ambos índices de irreversibilidad y entender por qué no se comportan de manera estadísticamente similar en todos los periodos. También sería interesante investigar si existe alguna correlación entre el índice de asimetría desfasada y el estimador para ineficiencia  $I_*$ . Nos preguntamos por qué los índices para la irreversibilidad y la ineficiencia sólo muestran correlación en ciertos periodos.

Otro aspecto a tener en cuenta es investigar de manera teórica si existe una correlación entre la irreversibilidad y la ineficiencia. Considerar variables económicas significativas que quizás hemos pasado por alto en este estudio, nos permitiría obtener una visión más completa de los periodos en los que esta correlación se manifiesta. La irreversibilidad puede ser una herramienta útil para obtener información sobre la ineficiencia del mercado durante periodos de alta inestabilidad. Durante estos tiempos, los picos en los índices de ineficiencia e irreversibilidad pueden ser más notables, lo que nos lleva a pensar que la correlación entre estos dos índices puede ser más fuerte. Específicamente, estos periodos pueden coincidir con momentos de turbulencia financiera, como el inicio de la pandemia de COVID-19.

Una pregunta que podría surgir al examinar el comportamiento de las criptomonedas es: ¿por qué todas las criptomonedas no presentan la misma conducta en términos de irreversibilidad e ineficiencia? Como hemos señalado anteriormente, cada moneda ha surgido en diferentes momentos y bajo circunstancias distintas. Esto podría significar que algunas de las criptomonedas más recientes aún no han alcanzado la estabilidad en términos de eficiencia. Además, es relevante recordar que cada criptomoneda posee un modelo de operación y una economía propios, factores que pueden influir en su eficiencia y estabilidad. Por consiguiente, es de esperar que existan variaciones en el comportamiento de la irreversibilidad e ineficiencia entre las diferentes criptomonedas.

Una hipótesis que podría explorarse en futuros estudios es la de los Mercados Adaptativos [34,49]. Esta teoría propone que los mercados se comportan de manera eficiente en el largo plazo, pero pueden presentar ineficiencias en el corto plazo debido a las conductas y decisiones de los agentes que participan en ellos. Los agentes se adaptan constantemente a las nuevas informaciones y cambian sus estrategias de inversión, lo que puede generar fluctuaciones en los precios de los activos y, por ende, ineficiencias temporales en el mercado. Explorar esta hipótesis podría ayudar a entender las correlaciones observadas en ciertos periodos entre los índices de irreversibilidad e ineficiencia y cómo estos se relacionan con los momentos de alta inestabilidad en los mercados.

Finalmente, nos gustaría buscar o proponer otro índice, tanto para la irreversibilidad como para la ineficiencia, que permita contrastar los resultados obtenidos en este estudio. Esta línea de investigación podría arrojar luz sobre las discrepancias observadas en este estudio y proporcionar un marco más completo para entender la eficiencia del mercado.

# Capítulo 5

## Conclusiones.

En el desarrollo de este trabajo, hemos ejecutado una serie de análisis con el objetivo de medir la irreversibilidad en series de tiempo financieras, centrándonos especialmente en el Bitcoin y otras criptomonedas relevantes. A lo largo de este proceso, hemos descubierto que algunos de los métodos empleados no proporcionan la precisión suficiente para determinar con fiabilidad el grado de irreversibilidad y sus periodos correspondientes. Sin embargo, hemos logrado resultados satisfactorios con los análisis realizados a través del índice de asimetría desfasada y mediante el uso de patrones de tendencia, también llamado análisis de rachas. El método de rachas no solo es sencillo de aplicar y no paramétrico, sino que también nos permite entender mejor cuán irreversible es una serie de tiempo y los periodos donde esta irreversibilidad se manifiesta.

En cuanto a la conexión entre la irreversibilidad de la serie de tiempo y su ineficiencia, aún no hemos llegado a una conclusión definitiva. Aunque para la criptomoneda Ether parece haber una alta correlación, en el caso de Bitcoin, XRP, Litecoin y Bitcoin Cash, la correlación sólo se manifiesta en ciertos periodos y resulta negativa o inexistente en otros. Los resultados arrojados por el análisis del índice de asimetría desfasada nos proporcionan una visión más profunda. En algunos periodos, este estimador se comportó de manera similar al estimador de irreversibilidad, mientras que en otros, seguía al estimador de ineficiencia. Los resultados obtenidos con el análisis del índice de asimetría desfasada nos brindan mayor información sobre el punto anterior. Este estimador, en ciertos periodos se comportó como el estimador de irreversibilidad, mientras que en otros, su comportamiento seguía al estimador de ineficiencia. Aunque no hemos identificado una correlación directa entre los dos indicadores (irreversibilidad e ineficiencia), este estudio sugiere que la irreversibilidad podría ser útil para analizar la ineficiencia en ciertos mercados o periodos. Esto es evidente en el surgimiento de nuevos mercados o en periodos de inestabilidad económica, como el inicio de la pandemia de COVID-19. Dado que estos

dos indicadores no están correlacionados la mayor parte del tiempo, no podemos afirmar que medir la irreversibilidad de una serie temporal sea lo mismo que medir la ineficiencia. No obstante, al observar un comportamiento similar entre los dos indicadores, podemos inferir que existe una relación, aunque puede haber factores que aún no hemos tenido en cuenta.

En vista de los resultados obtenidos, planeamos realizar un estudio posterior para analizar esta correlación desde la perspectiva de la ineficiencia semi-fuerte. Es decir, queremos explorar cómo las noticias, ya sean positivas o negativas, influenciaron el movimiento de los precios en los periodos con algún grado de irreversibilidad. Esto podría llevarse a cabo con la ayuda del análisis de sentimientos, una técnica utilizada en el campo del *machine learning*.

Además, consideramos que es importante buscar otro índice de ineficiencia e irreversibilidad que nos permita contrastar con los resultados de este estudio y obtener más información sobre el comportamiento mutuo de los índices utilizados. Queremos entender por qué en algunos periodos ambos índices parecen comportarse de manera similar y por qué en otros no, o por qué en algunos periodos el comportamiento es similar pero con cierto desfase. Esperamos responder a estas y otras preguntas en futuros estudios.

# Apéndice A

## Apéndice.

En esta sección daremos la demostración de algunos de los resultados dados en el capítulo 2. Comenzamos con el teorema 2.3.10.

**Teorema.** Sea  $X = \{X_n\}_{n \in \mathbb{Z}}$  una cadena de Markov, estacionaria, con espacio de estados  $S$  numerable. Sea  $P = (p_{ij})_{i,j \in S}$  la matriz de transición y una distribución única e invariante  $\Pi = (\pi_i)_{i \in S}$  entonces la tasa de producción de entropía para una cadena de Markov puede escribirse como

$$e_p = \frac{1}{2} \sum_{i,j \in S} (\pi_i p_{ij} - \pi_j p_{ji}) \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}}.$$

**Demostración.** Sabemos que  $e_p$  puede obtenerse mediante la ecuación

$$e_p = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left( \ln \frac{P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)}{P(X_n = x_1, X_{n-1} = x_2, \dots, X_1 = x_n)} \right), \quad (\text{A.1})$$

para la cadena una cadena de Markov se tiene que

$$e_p = \lim_{n \rightarrow \infty} \frac{1}{n} \left( \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} \pi_0 p_{01, p_{12}, \dots, p_{(n-1)n}} \ln \frac{\pi_0 p_{12}, \dots, p_{(n-1)n}}{\pi_n p_{n(n-1), \dots, p_{10}} \right).$$

Sabemos que  $P$  es una matriz estocástica, entonces

$$\sum_j p_{ij} = 1,$$

y además, ya que la cadena es estacionaria

$$\sum_j \pi_i p_{ij} = \pi(j).$$

Entonces se tiene que, la siguiente expresión

$$\sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} \pi_0 p_{01, p_{12}, \dots, p_{(n-1)n}} \ln \frac{\pi_0 p_{12}, \dots, p_{(n-1)n}}{\pi_n p_{n(n-1), \dots, p_{10}},$$

puede simplificarse para cada  $j \in [1, n]$ , a lo siguiente

$$= \sum_{j-1} \sum_j \pi_{j-1} p_{(j-1)j} \ln \frac{\pi_{j-1} p_{(j-1)j}}{\pi_j p_{j(j-1)}},$$

tomando la sumando sobre  $n$  y reetiquetando los índices se tiene

$$\begin{aligned} &= \sum_{k=1}^n \left( \sum_{j-1} \sum_j \pi_{j-1} p_{(j-1)j} \ln \frac{\pi_{j-1} p_{(j-1)j}}{\pi_j p_{j(j-1)}} \right) \\ &= n \sum_i \sum_j \pi_i p_{ij} \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}}. \end{aligned} \tag{A.2}$$

Entonces

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{1}{n} \left( \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} \pi_0 p_{01, p_{12}, \dots, p_{(n-1)n}} \ln \frac{\pi(x_0) p_{12}, \dots, p_{(n-1)n}}{\pi(x_n) p_{n(n-1), \dots, p_{10}} \right) \\ &= \sum_i \sum_j \pi_i p_{ij} \ln \frac{\pi(x_i) p_{ij}}{\pi(x_j) p_{ji}} \\ &= \frac{1}{2} \sum_{i,j \in S} (\pi_i p_{ij} - \pi_j p_{ji}) \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}}. \end{aligned}$$

**Corolario.** Sea  $h$  la tasa de entropía directa para una cadena de Markov estacionaria y sea  $h_R$  la tasa de entropía reversa, correspondiente a la tasa de entropía obtenida cuando el proceso ocurre en sentido inverso entonces

$$e_p = h_R - h. \tag{A.3}$$

**Demostración.** Sabemos que

$$\begin{aligned}
e_p &= \frac{1}{2} \sum_{i,j \in S} (\pi_i p_{ij} - \pi_j p_{ji}) \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}} \\
&= \frac{1}{2} \sum_{i,j \in S} \pi_i p_{ij} \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}} - \frac{1}{2} \sum_{i,j \in S} \pi_j p_{ji} \ln \frac{\pi_i p_{ij}}{\pi_j p_{ji}} \\
&= \frac{1}{2} \sum_{i,j \in S} \pi_i p_{ij} \ln \pi_i p_{ij} - \frac{1}{2} \sum_{i,j \in S} \pi_i p_{ij} \ln \pi_j p_{ji} - \frac{1}{2} \sum_{i,j \in S} \pi_j p_{ji} \ln \pi_i p_{ij} + \frac{1}{2} \sum_{i,j \in S} \pi_j p_{ji} \ln \pi_j p_{ji} \\
&= \frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} \ln \pi_i - \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ij} - \sum_{i,j \in S} \pi_i p_{ij} \ln \pi_j + \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ji} \right. \\
&\quad \left. - \sum_{i,j \in S} \pi_j p_{ji} \ln \pi_i + \sum_{i,j \in S} \pi_j p_{ji} \ln p_{ij} + \sum_{i,j \in S} \pi_j p_{ji} \ln \pi_j - \sum_{i,j \in S} \pi_j p_{ji} \ln p_{ji} \right] \\
&= \frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} (\ln \pi_i - \ln \pi_j) \right] + \frac{1}{2} \left[ \sum_{i,j \in S} \pi_j p_{ji} (\ln \pi_j - \ln \pi_i) \right] \\
&\quad + \frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} (\ln p_{ji} - \ln p_{ij}) \right] + \frac{1}{2} \left[ \sum_{i,j \in S} \pi_j p_{ji} (\ln p_{ij} - \ln p_{ji}) \right].
\end{aligned}$$

Obsérvese que, por estacionariedad de la cadena de Markov

$$\sum_i \pi_i p_{ij} = \pi_j,$$

y además, ya que la matriz es estocástica

$$\sum_j p_{ij} = 1.$$

Entonces la ecuación anterior se reduce a

$$\begin{aligned}
&\frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ji} - \pi_i p_{ij} \ln p_{ij} + \pi_j p_{ji} \ln p_{ij} - \pi_j p_{ji} \ln p_{ji} \right] \\
&= \frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ji} + \pi_j p_{ji} \ln p_{ij} \right] - \frac{1}{2} \left[ \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ij} + \pi_j p_{ji} \ln p_{ji} \right].
\end{aligned}$$

Reetiquetando se tiene

$$\begin{aligned}
&\sum_{i,j \in S} \pi_i p_{ij} \ln p_{ji} - \sum_{i,j \in S} \pi_i p_{ij} \ln p_{ij} \\
&= h_R - h.
\end{aligned}$$





# Bibliografía

- [1] Rudolf Clausius. *Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie: vorgetragen in der naturforsch. Gesellschaft den 24. April 1865.* Éditeur inconnu, 1865.
- [2] Alfredo González-Espinoza, Gustavo Martínez-Mekler, and Lucas Lacasa. Arrow of time across five centuries of classical music. *Physical Review Research*, 2(3):033166, 2020.
- [3] Raul Salgado-Garcia and Cesar Maldonado. Estimating entropy rate from censored symbolic time series: A test for time-irreversibility. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(1), 2021.
- [4] Astero Provata, C Nicolis, and G Nicolis. Dna viewed as an out-of-equilibrium structure. *Physical Review E*, 89(5):052105, 2014.
- [5] Ryan Flanagan and Lucas Lacasa. Irreversibility of financial time series: a graph-theoretical approach. *Physics Letters A*, 380(20):1689–1697, 2016.
- [6] John C Hull. *Options, Futures, and Other Derivatives, th ed.* Prentice Hall, New Jersey, 2000.
- [7] Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2):383–417, 1970.
- [8] Arjun Singh. Are cryptocurrency markets, efficient markets? *Efficient Markets*, 2020.
- [9] Luciano Zunino, Massimiliano Zanin, Benjamin M Tabak, Darío G Pérez, and Osvaldo A Rosso. Forbidden patterns, permutation entropy and stock market inefficiency. *Physica A: Statistical Mechanics and its Applications*, 388(14):2854–2864, 2009.
- [10] Andrii O Bielskiy, Serhii V Hushko, Andriy V Matviychuk, Oleksandr A Serdyuk, Serhiy O Semerikov, and Vladimir N Soloviev. Irreversibility of financial time series: a case of crisis. 2021.

- 
- [11] Aurelio F Bariviera. The inefficiency of bitcoin revisited: A dynamic approach. *Economics Letters*, 161:1–4, 2017.
- [12] Eng-Tuck Cheah, Tapas Mishra, Mamata Parhi, and Zhuang Zhang. Long memory interdependency and inefficiency in bitcoin markets. *Economics Letters*, 167:18–25, 2018.
- [13] Saralees Nadarajah and Jeffrey Chu. On the inefficiency of bitcoin. *Economics Letters*, 150:6–9, 2017.
- [14] Ahmet Sensoy. The inefficiency of bitcoin revisited: A high-frequency analysis with alternative currencies. *Finance Research Letters*, 28:68–73, 2019.
- [15] Andrew Urquhart. The inefficiency of bitcoin. *Economics Letters*, 148:80–82, 2016.
- [16] Coinmarketcap. <https://coinmarketcap.com/>. Accessed: Febrero, 2023).
- [17] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Decentralized Business Review*, page 21260, 2008.
- [18] Bitcoin.org. <https://bitcoin.org/es/>. Accessed: Agosto, 2023).
- [19] Coinmarketcap. <https://coinmarketcap.com/>. Accessed: Agosto, 2023).
- [20] ethereum.org. <https://ethereum.org/es/what-is-ethereum/>. Accessed: Agosto, 2023).
- [21] xrpl.org. <https://xrpl.org/xrp-overview.html>. Accessed: Agosto, 2023).
- [22] litecoin.org. <https://litecoin.org/es/>. Accessed: Agosto, 2023).
- [23] bitcoincash.org. <https://bitcoincash.org/>. Accessed: Agosto, 2023).
- [24] Is litecoin the silver to bitcoin’s gold? <https://www.coindesk.com/markets/2013/11/27/is-litecoin-the-silver-to-bitcoins-gold/>. Accessed: Agosto, 2023).
- [25] Tara Mandjee. Bitcoin, its legal classification and its regulatory framework. *J. Bus. & Sec. L.*, 15:1, 2014.
- [26] Anne Haubo Dyhrberg. Hedging capabilities of bitcoin. is it the virtual gold? *Finance Research Letters*, 16:139–144, 2016.
- [27] Eng-Tuck Cheah and John Fry. Speculative bubbles in bitcoin markets? an empirical investigation into the fundamental value of bitcoin. *Economics letters*, 130:32–36, 2015.

- 
- [28] Luis Rincón. *Introducción a los procesos estocásticos*. UNAM, Facultad de Ciencias, 2012.
- [29] Luis Rincón. *Introducción a la teoría del riesgo*. 2012.
- [30] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.
- [31] Ruey S Tsay. *Analysis of financial time series*. John wiley & sons, 2005.
- [32] Da-Quan Jiang and Donghua Jiang. *Mathematical theory of nonequilibrium steady states: on the frontier of probability and dynamical systems*. Springer Science & Business Media, 2004.
- [33] Eugene F Fama. Random walks in stock market prices. *Financial analysts journal*, 51(1):75–80, 1995.
- [34] Andrew W Lo. Adaptive markets and the new world order (corrected may 2012). *Financial Analysts Journal*, 68(2):18–29, 2012.
- [35] Katherine Julieth Sierra Suárez, Juan Benjamín Duarte Duarte, and Victor Alfonso Rueda Ortíz. Predictability of returns in the colombian stock market and the adaptive market hypothesis. *Estudios Gerenciales*, 31(137):411–418, 2015.
- [36] Raul Salgado-García. Time-irreversibility test for random-length time series: The matching-time approach applied to dna. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(12), 2021.
- [37] Peter J Brockwell and Richard A Davis. *Introduction to time series and forecasting*. Springer, 2002.
- [38] Cryptodatadownload. <https://www.cryptodatadownload.com/>. Accessed: February, 2023).
- [39] Xavier Brouty and Matthieu Garcin. A statistical test of market efficiency based on information theory. *Quantitative finance*, 23(6):1003–1018, 2023.
- [40] Donald S Ornstein and Benjamin Weiss. How sampling reveals a process. *The Annals of Probability*, 18(3):905–930, 1990.
- [41] Ioannis Kontoyiannis. Asymptotic recurrence and waiting times for stationary processes. *Journal of Theoretical Probability*, 11:795–811, 1998.

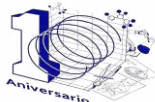
- 
- [42] Pierre Gaspard. Time-reversed dynamical entropy and irreversibility in markovian random processes. *Journal of statistical physics*, 117:599–615, 2004.
- [43] Gemma Lancaster, Dmytro Iatsenko, Aleksandra Pidde, Valentina Ticcinelli, and Aneta Stefanovska. Surrogate data for hypothesis testing of physical systems. *Physics Reports*, 748:1–60, 2018.
- [44] Nazul Merino-Negrete, Cesar Maldonado, and Raúl Salgado-García. Sorting eggs by lag irreversibility. *Physica D: Nonlinear Phenomena*, 459:134022, 2024.
- [45] Andrew Urquhart. The inefficiency of bitcoin. *Economics Letters*, 148:80–82, 2016.
- [46] Aurelio F. Bariviera. The inefficiency of bitcoin revisited: A dynamic approach. *Economics Letters*, 161:1–4, 2017.
- [47] Aviral Kumar Tiwari, R.K. Jana, Debojyoti Das, and David Roubaud. Informational efficiency of bitcoin—an extension. *Economics Letters*, 163:106–109, 2018.
- [48] Luis Rincón. *Estadística descriptiva*. UNAM, Facultad de Ciencias, 2017.
- [49] Andrew Lo. *Adaptive markets: Financial evolution at the speed of thought*. Princeton University Press, 2017.



# INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

POSGRADO EN CIENCIAS

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS



Instituto de Investigación en Ciencias Básicas y Aplicadas



Cuernavaca, Mor., a 23 de mayo de 2024

**DRA. LINA ANDREA RIVILLAS ACEVEDO  
COORDINADORA DEL POSGRADO EN CIENCIAS**

**PRESENTE**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la tesis titulada: ANÁLISIS DE IRREVERSIBILIDAD EN SISTEMAS FINANCIEROS, que presenta la M. en C. Jessica Morales Herrera (10009554) para obtener el título de Doctora en Ciencias.

Director de tesis: Dr. Raúl Salgado García

Unidad Académica: Instituto de Investigación en Ciencias Básicas y Aplicadas (IICBA)

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Markus Franziskus Müller CInC - UAEM	APROBADO	
Dr. Antonio Daniel Rivera López CInC - UAEM	APROBADO	
Dr. Cesar Octavio Maldonado Ahumada Inst. Potosino de Investigación Científica y Tecnológica AC	APROBADO	
Dra. Paola Vanessa Olguín Rodríguez CInC - UAEM	APROBADO	
Dr. José Daniel Arzate Mena CInC - UAEM	APROBADO	
Dra. Brenda Ivette García Maya Tecnológico de Monterrey- Campus Guadalajara	APROBADO	
Dr. Raúl Salgado García CInC - UAEM	APROBADO	





UNIVERSIDAD AUTÓNOMA DEL  
ESTADO DE MORELOS

Se expide el presente documento firmado electrónicamente de conformidad con el ACUERDO GENERAL PARA LA CONTINUIDAD DEL FUNCIONAMIENTO DE LA UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS DURANTE LA EMERGENCIA SANITARIA PROVOCADA POR EL VIRUS SARS-COV2 (COVID-19) emitido el 27 de abril del 2020.

El presente documento cuenta con la firma electrónica UAEM del funcionario universitario competente, amparada por un certificado vigente a la fecha de su elaboración y es válido de conformidad con los LINEAMIENTOS EN MATERIA DE FIRMA ELECTRÓNICA PARA LA UNIVERSIDAD AUTÓNOMA DE ESTADO DE MORELOS emitidos el 13 de noviembre del 2019 mediante circular No. 32.

### Sello electrónico

**RAUL SALGADO GARCIA | Fecha:2024-05-23 08:51:09 | FIRMANTE**

brbt0BxphpblyqDfhsEDZCld7T8LHTjfyekh1CFH9Rn81z7jz5ennV8iLpY60SsC5J/sQG8u7REfzmtVWLNXpZrGr7kfzUc/8UigBUXgYg+iypr7sM3QsVZLxo2l+a8HpXv9cT3iNnBQ  
BvxsRV+fEAB8EXdteCoaJfxE5TYxWPFsFLKDIkMcZQMY97ZPUK3oap1x53SQD8DzakYp6IC+cjRwpJqLqzdw3aVdTQibZGNWCxP/thJGORdsw8BwGxh35wL8LasQlsQACmITP  
5ZelCEsIVAJ0zUCVKENUBz8eUZEb7qrZx5J7nOnRctz/T3e/vS4BZNN08lyfjTKdTX5g==

**CÉSAR OCTAVIO MALDONADO AHUMADA | Fecha:2024-05-23 09:25:26 | FIRMANTE**

soHqJ73p7fR032JvC5o2X2kUwKjZuO1wq6THTupgRM6GVfymppJcdeWlcr1TpsunHDsX4eyde2vlzr5mlsfomU4pU6xClupc4TxWhsMGlVXNA/Z6a66x6F6mA8bUBVRng2Dmd  
9njsh7r+wPpG4KEoG4Uej/tKElJLFnpliSn3IKxYqSK4o8qrr6/6Z0pO1UBp1t4B+5nwtivEZ7HuWtIgleaeCxV1Klqq/LMQ/UzicVK72uWHf0f6pWofJf3WkWehN9OOmzAUbRGe  
lgR6PVHUro7k3WA8OgSXAMm+IMFRooNYCibOL1aWzNIRmtzyq8Mw0MoSfmUo3c/SCg==

**PAOLA VANESSA OLGUÍN RODRÍGUEZ | Fecha:2024-05-23 09:46:33 | FIRMANTE**

SYaAjko6ASFmHGcx1q4Gw66/KECms2pt9htD9SsCg1lwnnHef1Wo+ETab6K5gQf9os8RckSWMSJGjaOoUQTyIXalTYKAUhoOoDQ53OdYw9YaYx8j4xkRE6YtC5h5Z+Gx+t2  
5wRgvi3LKANu3xmsVyOmQuD0HeS+jNjz2JYSJOIZY+PCM8bBP4ZA5zQDAG5QUFDZjxWrc+Didd4P4hXFoWWhgsmfjJF9Orxy1H99b28arglVa37L40IVPefaxAFonFdTP8gPIIG  
TpXO/BFik9xdi4AbkzlpndYqHBDotf3hj8CZ9go60HHpwV/9BEYJmEsF/WiWxtr+WWovlv5+fw==

**MARKUS FRANZISKUS MULLER | Fecha:2024-05-23 10:45:47 | FIRMANTE**

a85ybtioqg4XYUueiw9OxfcgZRupf0KsLQAU8UKUYkmrqys+ZiOMTmU2+oSg6wbT58D32AwMclhYgz91Rfk3GFzSy8mwNPLKtdRFRKfjYsQvc5w2iibKXDMkJtITWHV/waQBpBa  
TcFJR28fv7QyFPXbFaraE3l32Rit4o6bi8b/DD1hZ43oXuO+5Ux//knmAP/78dO769ar2YjTUzUkQBBornL9HOA/BEJ7J02tOogim1zhD3PXF9aneYJuOmMimvT7jyaZNa8oUMqb  
8BoQ/CpkdGOEijRoh8At5oIMo+7t8/IUcmSPOR6MtsHivRnSbPKdfpJJE/vZRDHv9zf2A==

**JOSE DANIEL ARZATE MENA | Fecha:2024-05-23 12:09:22 | FIRMANTE**

PuP8aEz3brj11vpDBcQDej8CHWkqAKOSVdqxoCKOPXDzPT6PP13EoVGD5VkwObPxyvJl+5t3IM/3J8EAq80npQ4KqKqGwSvnpVpL/+oxQNIg9elsm0CeOk2DH8aDME5A4FB  
TbvAsdveybl27b9N+pZVHwmivESfMp//IAv2TcdyuWtSEMb1RQjKeP6LHZayHsDuf3cUnqFrmJ8jtQwhZSuu2z5DBC6VeseulXekjvxcqrEZwzP0j3kvN+QvvCsmJUxmpavfZ4FCDr  
TA8xyPKMxyi2wEByqigU1OtD0o+NN0+XFwlfzI9S5/0y/H8v6fcoy8eWUzdo1b6aMLI+1Q==

**BRENDA IVETTE GARCÍA MAYA | Fecha:2024-05-24 10:23:05 | FIRMANTE**

gmOzUo0zKF19Ja8HhgeQSa7N1AHnd19YjUD3IXOLHYV3jXsTMknRAfp4RGmCnLznxjYydl6ihJO1ecNgUqtmLihRmDOCINx/t35dneu6qfc6ijmnuGg8vKY9YXYIQctzMUvfmW7  
43cfnqxj4wOMI+f+IODS6dVxyzwEKmPQFvprna9ISs+IXO2SvaprEMMOBpfxmpnfh1PWZd2ZcGwKZ0ck0plxwbOFELDMAgJEYhlzqvN+xxUWdRdHLP/aQ5PVEt7XHdjY3nzINK  
KTBPiW7fA7PNtIBNrszmFcvqFhJ0y4Th1lxFSxsZcxBDulCXTE/NcbLLOyFwT0Eue25CeYkRQ==

**ANTONIO DANIEL RIVERA LOPEZ | Fecha:2024-05-27 13:55:41 | FIRMANTE**

rNdOKi+M+ai0JA86ZRC9ISWxamwK01CfMLmSrijQFGC24XilsLMcBMHUNFeE/1Ps/nBOU9W4XKCGrmzzLIDCGp/tO22Ahx30RuLyLueiK9x87k7xsPYw5w0Jmm4EEIQkukt7  
a1gq90KfmOOZ2TRNBf/OZqaGLCOIBaObXV6wmsOfmp3Ouf+XDpn6lnScPnE6ltycr87AVSVJpuF4pvsoHPFqjL498C9VVJDionOtsfv7h9OcYbkdsuEum6t2/UQ+cClunC90DO  
gDJYB0e4TPvVeRwxwezPxSNo/z9VJi11/Cklqhsz100/qznVul5SwimKpftME0j8ytj+7VA==

Puede verificar la autenticidad del documento en la siguiente dirección electrónica o  
escaneando el código QR ingresando la siguiente clave:



YqShJACLI

<https://efirma.uaem.mx/noRepudio/gDfffN0aMF8NtptA7YLSjxVINo7Q6tF0>



UAEM  
RECTORÍA  
2023-2029