



Facultad de Ciencias
Instituto de Estadística
Ingeniería en Estadística

Entropía cruzada muestral de perfiles para series de tiempo financieras.

Alejandro Gabriel Brito Martínez
14 de diciembre de 2021

Profesor Guía

Dr. Javier Contreras

Instituto de Estadística, Universidad de Valparaíso

Proyecto de titulación para optar al:

grado académico de: *Licenciado en Estadística*

título profesional de: *Ingeniero en Estadística*

minor en: *Estadística Financiera*

Resumen

En este trabajo se presenta una estimación eficiente y robusta de la sincronización, concretamente la medida de entropía cruzada de perfiles (CEP), para analizar series temporales complejas. A diferencia de la entropía cruzada muestral (CSE) original que se basa en un parámetro de tolerancia fijo r , la CEP propuesta se basa en el método de histograma acumulativo (CHM) en conjunto con la regla de Freedman-Diaconis para obtener un rango de valores de entropía con diferentes selecciones de r en un rango determinado. Además, se redefine la medida de disimilitud en la CEP, en lugar de la función de distancia utilizada en la CSE. Se realizan ensayos con datos simulados y series de tasas de cambio para hacer un estudio comparativo. La CSE original se introduce como comparación para ilustrar que la CEP propuesta es capaz de extraer relaciones más específicas de las series temporales, y también es un método más robusto para describir la sincronización entre ellas. Los resultados muestran que el nuevo método es capaz de distinguir diferentes tipos de series temporales, como fue demostrado en la aplicación de tipo de cambio de divisas. Se sugiere que la CEP propuesta puede convertirse potencialmente en un nuevo método fiable para la estimación de la sincronización de series temporales complejas y no necesariamente estacionarias.

Agradecimientos

Quiero agradecer a mi familia y amigos que han sido mis principales apoyo en momentos de flaqueza y debilidad, ya que con sus consejos, enseñanzas y tolerancia me han brindado ese respaldo necesario para culminar esta etapa de mi vida.

Gracias a mi pareja, por brindarme su apoyo y amor incondicional durante todo el desarrollo de la realización de este proyecto.

Agradecer también de manera especial a mi profesor guía, por compartir sus conocimientos y guiarme en el proceso de la presente tesis.

Índice general

Resumen	2
Agradecimientos	3
1. Introducción	5
1.1. Estado del Arte	6
1.2. Revisión de Literatura	7
1.2.1. Cuantificadores de información para variables discretas	7
1.2.2. Cuantificadores de información para variables continuas	8
2. Objetivos	11
2.1. Objetivo General	11
2.2. Objetivos Específicos	11
3. Metodología	12
3.1. Entropía cruzada muestral	12
3.1.1. Elección de los parámetros m y r	13
3.2. Entropía de perfiles	15
3.2.1. Regla de Freedman-Diaconis	16
3.3. Entropía cruzada de perfiles	18
4. Resultados	22
4.1. Simulación	22
4.1.1. Mapa de Hénon	22
4.2. Tasas de cambio de divisas	25
4.2.1. Visualización de las tasas de cambio	25
4.2.2. Prueba de Dickey-Fuller Aumentada	28
4.2.3. Selección del modelo	28
4.2.4. Estimación de parámetros	28
4.2.5. Prueba de causalidad de Granger	29
4.2.6. Aplicación de la entropía cruzada de perfiles	32
5. Conclusiones	35
A. Código	36
Referencias	43

Capítulo 1

Introducción

La entropía es un término científico, así como una propiedad física medible que se asocia comúnmente con un estado de desorden, aleatoriedad o incertidumbre. La entropía se utiliza en diversos campos, desde la termodinámica clásica, donde se reconoció por primera vez, hasta la descripción microscópica de la naturaleza en la física estadística y los principios de la teoría de la información ([Cover T. y Thomas J., 2012](#)).

Cuando se desarrolla la teoría de información, se adapta la idea de entropía de partículas a entropía de información y transferencia de bits, a través de la entropía de Shannon ([Shannon, C. E., 1948](#)). Gracias a esta conexión con la estadística matemática, se establecen conceptos como la entropía de distribuciones, entropía cruzada y divergencia de Kullback-Leibler ([Kullback, S. y Leibler, R. A., 1951](#)) entre otras como medidas complementarias de información en el análisis de datos.

La econofísica es una rama de la economía heterodoxa que aplica las leyes de la termodinámica a la teoría económica, más específicamente se puede aplicar análisis estadístico de datos que provienen de indicadores micro y macroeconómicos utilizando estas herramientas de la termodinámica y teoría de la información ([Mantegna R.N. y Stanley H.E., 1999](#)). En esta rama, se han introducido recientemente un indicador de asociación, el cual permite comparar el grado de sincronización de dos conjuntos de datos (series de tiempo) extraídos en un intervalo de tiempo específico, este se denomina entropía cruzada muestral ([Lake, Richman, Griffin, y Moorman, 2002](#)).

El presente documento se enfoca en analizar los principales atributos de la entropía cruzada muestral a partir del contenido establecido en la literatura reciente, y buscar generalizar estos atributos desde otro enfoque. Específicamente, el interés está en reducir la arbitrariedad de los parámetros de este indicador.

1.1. Estado del Arte

Dentro del área de teoría de la información, se puede analizar desde un punto de vista estadístico el concepto de entropía (Shannon, C. E., 1948), el cual es un índice que mide el nivel de incertidumbre de una función de distribución. Gracias a este concepto se puede definir la divergencia de Kullback-Leibler (Kullback, S. y Leibler, R. A., 1951) que permite comparar dos distribuciones, como por ejemplo dos distribuciones Poisson con distintos parámetros, o dos distribuciones gaussianas con distintos parámetros de escala, en otras. Desarrollando esta divergencia nace la entropía cruzada que representa una especie de “diferencia” de entropías (Cover T. y Thomas J., 2012).

La entropía cruzada ha sido utilizada como base para poder desarrollar un índice que mide la asincronía en series de tiempo financieras. Este índice recibe el nombre de entropía cruzada muestral (CSE), la cual fue usada primero para comparar la variabilidad entre señales del corazón en bebés (Lake y cols., 2002), y posteriormente para medir la asincronía de dos series de retorno de tasas de cambio (DKK/USD, NOK/USD, CAD/USD, JPY/USD, KRW/USD, SGD/USD, THB/USD y TWD/USD)¹ para el periodo de 1995 a 1998 (antes de la crisis monetaria asiática) y de 1999 a 2003 (después de la crisis monetaria asiática), respectivamente (Li-Zhi L., Xi-Yuan Q., y Heng-Yao L., 2010). Se encontró que la CSE de cada una de estas dos series de retorno de tasas de cambio es mayor después de la crisis monetaria asiática, lo que indica una mayor asincronía entre las tasas de cambio. En comparación con el coeficiente de correlación, la CSE es superior para describir la correlación entre las series temporales, lo que proporciona un método alternativo y prometedor para la demostración de las correlaciones entre las series temporales. También se espera que la CSE tenga importantes aplicaciones para las estrategias de diversificación y selección de portafolio (Xia y Shang, 2012).

A pesar de la efectividad de la CSE, esta carece de eficiencia debido a que se basa en parámetros que son considerados arbitrarios por miembros de la comunidad científica, por lo que naturalmente se intentan descubrir nuevas maneras de mejorar dicho método. Recientemente se han desarrollado procedimientos como la entropía de perfil (EP) (Udhayakumar, Karmakar, y Palaniswami, 2017), que es una extensión de la entropía muestral (Richman y Moorman, 2000) y que mejora la arbitrariedad y eficiencia del parámetro de tolerancia r (Karmakar C., Udhayakumar R., y Palaniswami M., 2020). Además, se ha corregido la arbitrariedad del parámetro de tolerancia de la CSE utilizando métodos no paramétricos, como el de bootstrap (Ramírez-Parietti, I., Contreras-Reyes, J.E., y Idrovo-Aguirre, B.J., 2021). Sin embargo, el trabajo más reciente y el que motiva esta tesis, ataca esta problemática considerando la entropía cruzada muestral refinada² (Shang D., Shang P., y Zhang Z., 2021).

¹En orden: corona danesa, corona coruega, dólar canadiense, yen japonés, won sur coreano, dólar de singapur, bat tailandés, dólar taiwanés

²En esta tesis, y de forma de ser consistente con la entropía de perfiles, la llamaremos entropía cruzada muestral de perfiles).

1.2. Revisión de Literatura

En esta sección, se revisan brevemente los principales cuantificadores de información para variables aleatorias discretas y continuas, que nos permitirán comprender de mejor forma el índice de entropía cruzada muestral.

1.2.1. Cuantificadores de información para variables discretas

En el área de teoría de la información, la entropía representa la información promedio entorno a la observación de una variable aleatoria X . Se considera X como una variable aleatoria discreta con una función de distribución de probabilidad $\mathbb{P}(X = x)$, tal que para un evento específico x_i se asigna la probabilidad:

$$\mathbb{P}(X = x_i) = p_i \quad , \forall i \mid 1 \leq i \leq n.$$

Por otro lado, la esperanza de X está dada por:

$$\mu = \mathbb{E}[X] = \sum_{i=1}^n x_i p_i.$$

Esta esperanza se puede extender a funciones entorno a variables aleatorias, siempre y cuando dichas funciones estén bien definidas, de la forma:

$$\mu = \mathbb{E}[f(X)] = \sum_{i=1}^n f(x_i) p_i.$$

Relacionando lo expuesto con el concepto de entropía se puede definir lo siguiente:

Definición 1 (*Entropía Discreta*). Definimos $H(X)$ como la entropía de una variable aleatoria discreta X como:

$$H(X) = \mathbb{E}[g(x)] = \sum_{x \in X} p(x) \log \frac{1}{p(x)} = - \sum_{x \in X} p(x) \log p(x),$$

donde $p(x)$ es la función de distribución de probabilidad de X y $g(x) = -\log p(x)$.

Propiedad 1 (*Entropía Discreta*). De la Definición 1 se obtiene lo siguiente:

i) $H(x) \geq 0$, ya que $0 < p(x) < 1$.

ii) Sean a y b dos bases distintas de las funciones logarítmicas correspondientes a las entropías $H_a(X)$ y $H_b(X)$, por propiedad de aritmética de logaritmos tenemos la siguiente relación:

$$H_b(X) = (\log_b a) H_a(X).$$

Definición 2 (*Entropía Relativa o divergencia de Kullback–Leibler*). Se define la divergencia de Kullback–Leibler como la medida de la distancia entre dos funciones de distribución de probabilidades (*Kullback, S. y Leibler, R. A., 1951*) y se denota como:

$$D(p \parallel q) = \sum_{x \in X} p(x) \ln \frac{p(x)}{q(x)}.$$

Propiedad 2 La entropía de Kullback Leibler es siempre positiva y es cero sólo cuando $p(x) = q(x)$.

Observación 1 La entropía relativa no es una distancia, dado que no cumple la desigualdad triangular. En general, ocurre que $D(p \parallel q) \neq D(q \parallel p)$.

1.2.2. Cuantificadores de información para variables continuas

De manera similar a la sección anterior se introducen las definiciones de cuantificadores pero esta vez respecto a una variable continua. Se define X como una variable aleatoria continua en un soporte S con función de densidad $f(x)$ tal que su función de distribución acumulada es:

$$F(x) = \int_S f(x)dx.$$

Su esperanza se denota como:

$$\mu = E[X] = \int_S xf(x)dx.$$

Se denota la esperanza para funciones en torno a las probabilidades asociadas de X como:

$$\mu = E[g(X)] = \int_S g(x)f(x)dx.$$

Definición 3 (*Entropía Diferencial*). Definimos $H(X)$ como la entropía de una variable aleatoria continua X con función de densidad f como:

$$H(X) = E[g(x)] = \int_S f(x) \ln \frac{1}{f(x)} dx = - \int_S f(x) \log f(x) dx,$$

donde S es el conjunto soporte de la variable aleatoria y $g(x) = -\log p(x)$.

Ejemplo 1 (*Distribución Normal*). Se calcula la entropía diferencial para una variable aleatoria normal $X \sim \mathcal{N}(\mu, \sigma^2)$, con parámetro de locación $\mu \in \mathbb{R}$ y de escala $\sigma^2 > 0$. La función de densidad es dada por

$$\phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

La entropía de X se calcula como:

$$\begin{aligned} H(X) &= - \int_{-\infty}^{\infty} \phi(x) \log \phi(x) dx \\ &= - \int_{-\infty}^{\infty} \phi(x) \log \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx. \end{aligned}$$

Luego, se separa en sumas de logaritmos y se distribuye la integral con $\phi(x)$:

$$\begin{aligned}
H(X) &= - \int_{-\infty}^{\infty} \phi(x) \left(\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{-(x-\mu)^2}{2\sigma^2} \right) dx \\
&= - \int_{-\infty}^{\infty} \phi(x) \log(2\pi\sigma^2)^{-\frac{1}{2}} + \int_{-\infty}^{\infty} \phi(x) \frac{(x-\mu)^2}{2\sigma^2} dx \\
&= \frac{1}{2} \ln(2\pi\sigma^2) \int_{-\infty}^{\infty} \phi(x) dx + \frac{1}{2\sigma^2} \int_{-\infty}^{\infty} \phi(x)(x-\mu)^2 dx.
\end{aligned}$$

Por propiedad de función de distribución para variables aleatorias continuas se sabe que:

$$\begin{aligned}
\int_{-\infty}^{\infty} \phi(x)(x-\mu)^2 dx &= \sigma^2 \\
\int_{-\infty}^{\infty} \phi(x) dx &= 1.
\end{aligned}$$

Entonces:

$$\begin{aligned}
H(X) &= \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2} \\
&= \frac{1}{2} \log(2\pi\sigma^2) + \frac{1}{2} \log(e) \\
&= \frac{1}{2} \log(2\pi e\sigma^2).
\end{aligned}$$

Observación 2 La entropía diferencial de una variable aleatoria normal no depende del parámetro de locación μ , sólo depende del parámetro de escala σ^2 .

Definición 4 (Entropía Relativa o Divergencia de Kullback-Leibler). Se define $D(f \parallel g)$ a la entropía relativa entre dos funciones de densidad $f(x)$ y $g(x)$ de una variable aleatoria X como:

$$D(f \parallel g) = \int_{S_x} f(x) \log \frac{f(x)}{g(x)} dx.$$

Ejemplo 2 (Distribución Normal; Divergencia K-L). Para dos variables aleatorias continuas con distribuciones normales $X \sim \mathbb{N}(\mu_1, \sigma_1^2)$ e $Y \sim \mathbb{N}(\mu_2, \sigma_2^2)$, y sus respectivas densidades

$$\phi_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \quad \text{y} \quad \phi_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}},$$

la divergencia de Kullback-Leibler entre $\phi_1(x)$ y $\phi_2(x)$ (Cover T. y Thomas J., 2012) está dada por:

$$\begin{aligned}
D(\phi_1 \parallel \phi_2) &= \int_{-\infty}^{\infty} \phi_1 \log \frac{\phi_1}{\phi_2} dx \\
&= \int_{-\infty}^{\infty} \phi_1 \log \frac{\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\mu)^2}{2\sigma_1^2}}}{\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\mu)^2}{2\sigma_2^2}}} dx \\
&= \int_{-\infty}^{\infty} \phi_1 \log \left(\frac{\sigma_2}{\sigma_1} \right)^2 e^{\frac{1}{2} \left[\left(\frac{x-\mu_2}{\sigma_2} \right)^2 - \left(\frac{x-\mu_1}{\sigma_1} \right)^2 \right]} dx \\
&= \int_{-\infty}^{\infty} \phi_1 \log \left(\frac{\sigma_2}{\sigma_1} \right) + \frac{1}{2} \left[\left(\frac{x-\mu_2}{\sigma_2} \right)^2 - \left(\frac{x-\mu_1}{\sigma_1} \right)^2 \right] dx \\
&= \log \left(\frac{\sigma_2}{\sigma_1} \right) \int_{-\infty}^{\infty} \phi_1 dx + \frac{1}{2\sigma_2} \int_{-\infty}^{\infty} \phi_1 (x-\mu_2)^2 dx - \frac{1}{2\sigma_1^2} \int_{-\infty}^{\infty} \phi_1 (x-\mu_1)^2 dx \\
&= \log \left(\frac{\sigma_2}{\sigma_1} \right) + \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \phi_1 (x-\mu_1 + \mu_1 - \mu_2)^2 dx - \frac{1}{2} \\
&= \log \left(\frac{\sigma_2}{\sigma_1} \right) + \frac{1}{2\sigma_2^2} \int_{-\infty}^{\infty} \phi_1 [(x-\mu_1)^2 - 2(x-\mu_1)(\mu_1 - \mu_2) + (\mu_1 - \mu_2)^2] dx - \frac{1}{2} \\
&= \log \left(\frac{\sigma_2}{\sigma_1} \right) + \frac{1}{2\sigma_2^2} \left(\sigma_1^2 - 2(\mu_1 - \mu_2) \left(\int_{-\infty}^{\infty} \phi_1 x dx - \mu_1 \int_{-\infty}^{\infty} \phi_1 dx \right) + (\mu_1 - \mu_2)^2 \right) - \frac{1}{2} \\
&= \\
&= \log \left(\frac{\sigma_2}{\sigma_1} \right) + \frac{\sigma_1^2 + (\mu_1 - \mu_2)^2}{2\sigma_2^2} - \frac{1}{2}.
\end{aligned}$$

Observación 3 *Vemos que en el Ejemplo 2 la divergencia depende tanto de las varianzas como las medias, pero al asumir que $\mu_1 = \mu_2$, la divergencia depende sólo de las varianzas.*

Propiedad 3 *Cabe destacar que $D(f \parallel g)$ tiene un resultado finito, sí y sólo sí, el soporte de $f(x)$ está contenido en el soporte de $g(x)$.*

La teoría antes expuesta nos da a entender que la incertidumbre es una propiedad de las probabilidades asociadas a eventos aleatorios, la cual es desarrollada a partir de la definición de entropía y originada por la medición de observaciones de una variable aleatoria para el conteo de combinatorias que llevan un evento a ser probable, esto permite hacer de la incertidumbre una medida cuantificable en unidades de información. Asumiendo que los eventos poseen probabilidades de ocurrencia más bajas, conllevan a un conteo de información promedio más alto y que, inversamente, los eventos con probabilidades de ocurrencia altas se interpretan como eventos poco informativos. En resumidas cuentas, la probabilidad de un evento depende del número de combinaciones que hacen este estado probable y ese conteo de combinaciones es una cuantificación de incertidumbre medida en unidades de información.

Capítulo 2

Objetivos

2.1. Objetivo General

Estudiar el índice de entropía cruzada muestral (CSE) y una posible extensión dada por la entropía cruzada de perfiles (CEP), para comparar y analizar el nivel de sincronía/asincronía entre dos series temporales.

2.2. Objetivos Específicos

Para poder alcanzar el objetivo general se plantean los siguientes objetivos específicos:

- Estudiar el índice de entropía de perfiles (EP) y su implementación en el software R.
- Extender el índice de entropía de perfiles (EP) al índice de entropía cruzada de perfiles (CEP).
- Implementar la CSE y CEP en el software R.
- Aplicar la CEP en datos simulados por el mapa bidimensional de Hénon.
- Aplicar la CEP en series temporales de tasas de cambio y comparar los resultados con la CSE .

Capítulo 3

Metodología

3.1. Entropía cruzada muestral

Sean $u = (u(1), u(2), \dots, u(N))$ y $v = (v(1), v(2), \dots, v(N))$ dos series de tiempo no necesariamente estacionarias y de largo N . Se fijan los parámetros de entrada m y r , donde m es la longitud del vector a comparar, y r es la tolerancia para aceptar coincidencias. Se forman las siguientes secuencias de vectores:

$$\begin{aligned}x_m(i) &= (u(i), u(i+1), \dots, u(i+m-1)), & 1 \leq i \leq N-m \\y_m(j) &= (v(j), v(j+1), \dots, v(j+m-1)), & 1 \leq j \leq N-m\end{aligned}$$

Luego para cada $i \leq N-m$, se fija:

$$B_i^m(r)(v \parallel u) = \frac{\text{número de } 1 \leq j \leq N-m \text{ tal que } d[x_m(i), y_m(j)] \leq r}{N-m},$$

donde

$$d[x_m(i), y_m(j)] = \text{máx}\{|u(i+k) - v(j+k)| : 0 \leq k \leq m-1\},$$

es decir, la máxima diferencia de sus componentes escalares. Luego se define:

$$B^m(r)(v \parallel u) = \frac{1}{N-m} \sum_{i=0}^{N-m} B_i^m(r)(v \parallel u),$$

que corresponde al valor promedio de $B_i^m(r)(v \parallel u)$.

Similarmente se define:

$$A_i^m(r)(v \parallel u) = \frac{\text{número de } 1 \leq j \leq N-m \text{ tal que } d[x_{m+1}(i), y_{m+1}(j)] \leq r}{N-m}$$

y su promedio:

$$A^m(r)(v \parallel u) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r)(v \parallel u).$$

Finalmente, la entropía cruzada muestral (Lake y cols., 2002) se define como:

$$CSE(m, r, N) = -\log \left\{ \frac{A^m(r)(v \parallel u)}{B^m(r)(v \parallel u)} \right\}.$$

Sea $B = (N - m)^2 B^m(r)(v \parallel u)$ y $A = (N - m)^2 A^m(r)(v \parallel u)$ de modo que B es el número total de pares de vectores de longitud m de las dos series que coinciden dentro de r y A es el número de pares de coincidencias desfasadas de longitud $m + 1$.

De lo anterior, tenemos que la entropía cruzada muestral determina cuántos vectores conformados por datos de series temporales ocurren con un rango estadístico significativo que puede ser entendido como similitud. Cuando un par de series de tiempo están muy sincronizadas, debe ocurrir que este conjunto presenta menores valores de entropía cruzada muestral que el otro. Mientras que en el caso contrario, la entropía cruzada muestral debe ser mayor para un par de series menos sincronizadas. Por lo tanto, la entropía cruzada muestral puede ser comprendida como una medida de correlación cruzada entre series temporales (Li-Zhi L. y cols., 2010).

3.1.1. Elección de los parámetros m y r

Las experiencias generales llevan a utilizar valores de r entre 0.1 y 0.25 y valores de m de 1 o 2 para registros de datos de longitud N que van de 100 a 5000. Una forma de determinar m es mediante una función del orden estimado del proceso AR de cada serie temporal de retorno, donde el orden del proceso AR se estima utilizando el método de máxima verosimilitud y el criterio AIC (Li-Zhi L. y cols., 2010). Se utiliza el criterio para seleccionar el r que minimiza (minmax):

$$\text{máx} \left(\frac{\sigma_{CP}}{CP}, \frac{\sigma_{CP}}{-\log(CP)CP} \right),$$

el cual es el máximo del error relativo de la entropía muestral y del estimador $CP = A/B$, respectivamente; donde CP estima la probabilidad condicional de una coincidencia entre u y v para $m + 1$ puntos dada la existencia de una coincidencia entre u y v para m puntos.

El número de coincidencias entre u y v de largo $m + 1$ se puede expresar como:

$$A = \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} U_{ij}, \quad \text{donde} \quad U_{ij} = \begin{cases} 1 & \text{si } d[x_{m+1}(i), y_{m+1}(j)] \leq r \\ 0 & \text{e.o.c.} \end{cases}.$$

La suma puede restringirse a los B pares (i, j) de coincidencias de longitud m , donde $d[x_m(i), y_m(j)] \leq r$. Así, la varianza de CP es:

$$\sigma_{CP}^2 = \frac{\text{Var}(A)}{B^2} = \frac{1}{B^2} \sum_{i=1}^{N-m} \sum_{j=1}^{N-m} \sum_{k=1}^{N-m} \sum_{l=1}^{N-m} \text{Cov}(U_{ij}, U_{kl}).$$

Para los B pares donde $i = k$ y $j = l$, se tiene que

$$\text{Cov}(U_{ij}, U_{kl}) = \text{Var}(U_{ij}) = CP(1 - CP).$$

Si los vectores plantilla para U_{ij} y U_{kl} no tienen puntos en común, significa que son independientes y por lo tanto no están correlacionados, es decir, $\text{Cov}(U_{ij}, U_{kl}) = 0$. Ahora si los vectores se solapan, es decir si $\text{mín}\{|i - k|, |j - l|\} \leq m$, se tiene que:

$$Cov(U_{ij}, U_{kl}) = U_{ij}U_{kl} - CP^2 = \begin{cases} 1 - (CP)^2, & \text{si coinciden vectores en } m + 1 \text{ puntos} \\ -(CP)^2, & \text{e.o.c.} \end{cases}$$

Finalmente, se estima la varianza de CP como:

$$\sigma_{CP}^2 = \frac{CP(1 - CP)}{B} + \frac{1}{B^2}[K_A - K_B(CP)^2],$$

donde K_A es el número de pares de vectores plantilla coincidentes de longitud $m + 1$ que se solapan y K_B es el número de pares de vectores plantilla coincidentes de longitud m que se solapan.

3.2. Entropía de perfiles

Para la entropía de perfiles (Udhayakumar y cols., 2017), se utiliza el método del histograma acumulativo (CHM) para generar un perfil de entropía muestral. El método utiliza un enfoque basado en los datos para obtener un potencial conjunto de valores de r . Este conjunto de valores se utiliza para construir el perfil de entropía. Los pasos se enumeran a continuación.

Sea $\{X(t)\}$ una serie temporal no necesariamente estacionaria y de longitud N , $1 \leq n \leq N$. Para un valor dado de la dimensión de incrustación m :

- (1) Formar $N - m$ vectores, cada uno de largo m , dado por

$$\{X_i^m : 1 \leq i \leq (N - m)\}$$

donde

$$X_i^m = \{x(i + k) : 0 \leq k \leq m - 1\}$$

- (2) Similarmente, formar $N - m$ vectores, cada uno de largo $m + 1$, dado por

$$\{X_i^{m+1} : 1 \leq i \leq (N - m)\}$$

con

$$X_i^{m+1} = \{x(i + k) : 0 \leq k \leq m\}$$

- (3) Tomar cada vector X_i^m del primer paso como un vector plantilla y encontrar su distancia por cada vector X_j^m , donde la distancia está dada por

$$d_{ij}^m = \{\text{máx} |X_i^m - X_j^m| : 1 \leq j \leq (N - m), j \neq i\}.$$

De esta forma, para un i -ésimo vector plantilla, el vector de distancia para una dimensión de incrustación m será

$$d_i^m = \{d_{ij}^m : 1 \leq j \leq (N - m), j \neq i\}.$$

Similarmente para una dimensión de incrustación $m + 1$, el vector de distancia puede ser obtenido como

$$d_i^{m+1} = \{d_{ij}^{m+1} : 1 \leq j \leq (N - m), j \neq i\}.$$

- (4) Sea D la matriz que contiene todos los elementos de d^m y d^{m+1} , se define un *rango* como el conjunto de todos los elementos únicos de D , ordenados de forma ascendente. Adicionalmente, n_{bin} es el número de elementos en este *rango*.
- (5) Desde el vector de distancias d_i^m , la función de distribución acumulada cdf_i^m es calculada como:

$$cdf_i^m(q) = p(d_i^m \leq rango(q)), \text{ para } 1 \leq q \leq n_{bin},$$

donde p es la probabilidad.

- (6) Repetir el paso 5 para todo $d_i^m : 1 \leq i \leq N - m$, para luego calcular la matriz de distribución acumulada completa cdf^m para la dimensión de incrustación m como se muestra:

$$cdf^m = \begin{bmatrix} cdf_1^m(1) & \dots & cdf_1^m(n_{bin}) \\ cdf_2^m(1) & \dots & cdf_2^m(n_{bin}) \\ \vdots & \ddots & \vdots \\ cdf_{N-m}^m(1) & \dots & cdf_{N-m}^m(n_{bin}) \end{bmatrix}$$

- (7) Similarmente, para la dimensión de incrustación $m + 1$ la matriz de distribución acumulada cdf^{m+1} puede ser obtenida repitiendo el paso 4 para todo d_i^{m+1} :

$$cdf^{m+1} = \begin{bmatrix} cdf_1^{m+1}(1) & \dots & cdf_1^{m+1}(n_{bin}) \\ cdf_2^{m+1}(1) & \dots & cdf_2^{m+1}(n_{bin}) \\ \vdots & \ddots & \vdots \\ cdf_{N-m}^{m+1}(1) & \dots & cdf_{N-m}^{m+1}(n_{bin}) \end{bmatrix}$$

- (8) Ahora, para cada columna de cdf^m y cdf^{m+1} , se calcula el promedio de las probabilidades, los cuales son representados por $\theta^m(q)$ y $\theta^{m+1}(q)$ respectivamente.

$$\theta^m(q) = \frac{1}{N-m} \sum_{i=1}^{N-m} cdf_{iq}^m, \quad 1 \leq q \leq n_{bin}$$

$$\theta^{m+1}(q) = \frac{1}{N-m} \sum_{i=1}^{N-m} cdf_{iq}^{m+1}, \quad 1 \leq q \leq n_{bin}.$$

- (9) La entropía muestral, la cual es una aproximación de la probabilidad condicional de dos segmentos que coinciden en una longitud $m + 1$ si coinciden en m , puede ser definida como

$$SampEn = \log \left(\frac{\theta^m}{\theta^{m+1}} \right).$$

En este caso, se tiene n_{bin} número de valores $SampEn$ ya que θ^m y θ^{m+1} son calculados entre $1 \leq q \leq n_{bin}$. Esto ultimo puede ser expresado como

$$SampEn(q) = \log \left(\frac{\theta^m(q)}{\theta^{m+1}(q)} \right), \quad 1 \leq q \leq n_{bin}.$$

Hay que tener en cuenta que cada valor de q es de hecho un valor único de r . Por lo tanto, habrá n_{bin} número de valores únicos de r utilizados para el perfil de entropía. El algoritmo selecciona automáticamente el valor de n_{bin} , derivado de la dinámica de la señal. En este caso, n_{bin} es un valor basado en los datos (que abarca todos los posibles valores de r para la señal dada) y, por tanto, la mejor elección para generar un perfil de entropía completo.

Cabe señalar que Karmakar C. y cols. (2020) no presentan un método para la selección de n_{bin} . Para ello es posible considerar la regla de Freedman-Diaconis (Freedman y Diaconis, 1981) para el número de segmentos n_{bin} del estimador de histograma (Contreras-Reyes, J.E., 2021), tal como se presenta a continuación.

3.2.1. Regla de Freedman-Diaconis

Sea $\{y_i\}$, $i = 1, \dots, N$, cualquier vector de observaciones de longitud N . La frecuencia relativa de la estimación del histograma se considera para $\{y_i\}$, que se divide en n_{bin} segmentos. Sea

$$IQR(\{y_i\}) = Q_3(\{y_i\}) - Q_1(\{y_i\})$$

el rango intercuartil de $\{y_i\}$, donde $Q_1(\{y_i\})$ y $Q_3(\{y_i\})$ son el primer y tercer cuartil del conjunto de observaciones $\{y_i\}$. Entonces, el número de segmentos se calcula a partir de la regla de Freedman–Diaconis como

$$n_{bin}^{FD} = \frac{2IQR(\{y_i\})}{N^{1/3}}. \quad (3.1)$$

Para un conjunto de observaciones $\{y_i\}$, $i = 1, \dots, N$, muestreadas a partir de alguna distribución de probabilidad, la regla de Freedman-Diaconis (3.1) minimiza la integral teórica de la diferencia al cuadrado entre la densidad de frecuencia relativa y la teórica.

3.3. Entropía cruzada de perfiles

A diferencia de la entropía cruzada muestral (CSE) original que se basa en una tolerancia fija r , la CEP propuesta se basa en un concepto llamado método de histograma acumulativo (CHM) para obtener un rango de valores de entropía con diferentes selecciones de r en un rango determinado (Shang D. y cols., 2021). Además, se redefine la medida de disimilitud en el CEP, en lugar de la función de distancia utilizada en el CSE. El método consiste en:

- (1) Sean $u = [u(1), u(2), \dots, u(N)]$ y $v = [v(1), v(2), \dots, v(N)]$ dos series de tiempo no necesariamente estacionarias y de longitud N .
- (2) Se fija un parámetro de entrada m , donde m es el largo de los vectores extraídos de u y v . Se extraen $N - m$ vectores de largo m :

$$\begin{aligned} X_m(i) &= [u(i), u(i+1), \dots, u(i+m-1)], & 1 \leq i \leq N-m, \\ Y_m(j) &= [v(j), v(j+1), \dots, v(j+m-1)], & 1 \leq j \leq N-m. \end{aligned}$$

Para u y v separadamente.

- (3) Permutar los dos vectores $X_m(i)$ e $Y_m(i)$ por el rango de sus elementos, y sustituir los elementos de los vectores por enteros que van desde 0 hasta $m-1$. Entonces $X_m(i)$ e $Y_m(i)$ son permutados como $X_m^p(i)$ e $Y_m^p(i)$, respectivamente, donde

$$\begin{aligned} X_m^p(i) &= [\tau_1, \tau_2, \dots, \tau_m], \\ Y_m^p(i) &= [\omega_1, \omega_2, \dots, \omega_m]. \end{aligned}$$

Aquí, τ_j y ω_j son los enteros permutados relacionados con los elementos originales.

- (4) Para definir la disimilitud o distancia entre dos vectores, tenemos que introducir la función delta de Kronecker. La función delta de Kronecker es una función entre dos variables, si las dos variables son iguales, la función devuelve uno, en caso contrario devuelve cero. Se define como:

$$\delta_{\tau_k \omega_k} = \begin{cases} 0 & \text{si } \tau_k \neq \omega_k \\ 1 & \text{si } \tau_k = \omega_k \end{cases}.$$

Además, se define

$$\theta_{ij}(m) = \frac{1}{m} \sum_{i=1}^m \delta_{\tau_i \omega_i}.$$

Luego, la disimilitud entre dos vectores extraídos de dos series temporales es:

$$d_{ij}^m = e^{-\theta_{ij}(m)} d[X_m(i), Y_m(j)],$$

donde

$$d[X_m(i), Y_m(j)] = \max\{|u(i+k) - v(j+k)| : 0 \leq k \leq m-1\}.$$

(5) Se construye el vector de distancia d_i^m :

$$d_i^m = [d_{i1}^m d_{i2}^m d_{i3}^m \dots d_{i(N-m+1)}^m].$$

De manera similar para el vector de largo $m + 1$, el vector de distancia se construye como:

$$d_i^{m+1} = [d_{i1}^{m+1} d_{i2}^{m+1} d_{i3}^{m+1} \dots d_{i(N-m)}^{m+1}].$$

(6) Para todos los vectores de distancia, se reordenan según su orden de menor a mayor y se colocan en el conjunto $\{D_i\}, i = 1, 2, \dots, n_e$, donde $n_e = (N - m + 1)(N - m)$. Sea $\delta = \min\{|D_i - D_{i+1}| : D_i \neq D_{i+1}, 1 \leq i \leq n_e - 1\}$, entonces el número total de todos los segmentos que se dividen por la longitud mínima δ es

$$n_{bin} = \frac{\max(\{D_i\}) - \min(\{D_i\})}{\delta}. \quad (3.2)$$

Aquí se propone la regla de Freedman-Diaconis dada en (3.1) para obtener

$$n_{bin}^{FD} = \frac{2IQR(\{D_i\})}{(N - m + 1)^{1/3}(N - m)^{1/3}}, \quad (3.3)$$

y, haciendo coincidir (3.2) y (3.3), el parámetro δ_{FD} basado en la regla de Freedman-Diaconis es

$$\delta_{FD} = \frac{(N - m + 1)^{1/3}(N - m)^{1/3}}{IQR(\{D_i\})} \left(\frac{\max(\{D_i\}) - \min(\{D_i\})}{2} \right).$$

(7) Para un vector de distancia d_i^m , se calcula la función de distribución acumulativa (cdf i m):

$$cdf_i^m(q\delta_{FD}) = p(d_i^m \leq q\delta_{FD}), \quad 1 \leq q \leq n_{bin}^{FD}.$$

(8) Se crea la matriz de distribución acumulativa completa cdf^m repitiendo el paso (7) para todas las $d_i^m, i = 1, 2, \dots, (N - m + 1)$:

$$cdf^m = \begin{bmatrix} cdf_1^m(1\delta_{FD}) & cdf_1^m(2\delta_{FD}) & \dots & cdf_1^m(n_{bin}^{FD}\delta_{FD}) \\ cdf_2^m(1\delta_{FD}) & cdf_2^m(2\delta_{FD}) & \dots & cdf_2^m(n_{bin}^{FD}\delta_{FD}) \\ \vdots & \vdots & \vdots & \vdots \\ cdf_{N-m+1}^m(1\delta_{FD}) & cdf_{N-m+1}^m(2\delta_{FD}) & \dots & cdf_{N-m+1}^m(n_{bin}^{FD}\delta_{FD}) \end{bmatrix}.$$

Del mismo modo, la matriz de distribución acumulativa completa para la dimensión $m + 1$ es:

$$cdf^{m+1} = \begin{bmatrix} cdf_1^{m+1}(1\delta_{FD}) & cdf_1^{m+1}(2\delta_{FD}) & \dots & cdf_1^{m+1}(n_{bin}^{FD}\delta_{FD}) \\ cdf_2^{m+1}(1\delta_{FD}) & cdf_2^{m+1}(2\delta_{FD}) & \dots & cdf_2^{m+1}(n_{bin}^{FD}\delta_{FD}) \\ \vdots & \vdots & \vdots & \vdots \\ cdf_{N-m}^{m+1}(1\delta_{FD}) & cdf_{N-m}^{m+1}(2\delta_{FD}) & \dots & cdf_{N-m}^{m+1}(n_{bin}^{FD}\delta_{FD}) \end{bmatrix}.$$

(9) A continuación, se define la media del logaritmo natural de cdf_i^m :

$$\Phi^m(q\delta_{FD})(v \parallel u) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \log cdf_i^m(q\delta_{FD}), \quad 1 \leq q \leq n_{bin}^{FD}.$$

Análogamente la media del logaritmo natural de cdf_i^{m+1} es:

$$\Phi^{m+1}(q\delta_{FD})(v \parallel u) = \frac{1}{N - m} \sum_{i=1}^{N-m} \log cdf_i^{m+1}(q\delta_{FD}), \quad 1 \leq q \leq n_{bin}^{FD}.$$

(10) Se observa que para todo lo anterior, $X_m(i)$ es la plantilla y $Y_m(j)$ es la coincidencia de la plantilla (u objetivo). Ahora, para aclarar bien la dependencia de la dirección, se tiene que definir $\Phi^m(q\delta_{FD})(u \parallel v)$ y $\Phi^{m+1}(q\delta_{FD})(u \parallel v)$ cambiando $X_m(i)$ e $Y_m(i)$, y luego repetir los pasos (2)-(9). Así, se puede definir:

$$B = \frac{\Phi^m(q\delta_{FD})(v \parallel u) + (\Phi^m(q\delta_{FD})(u \parallel v))}{2},$$

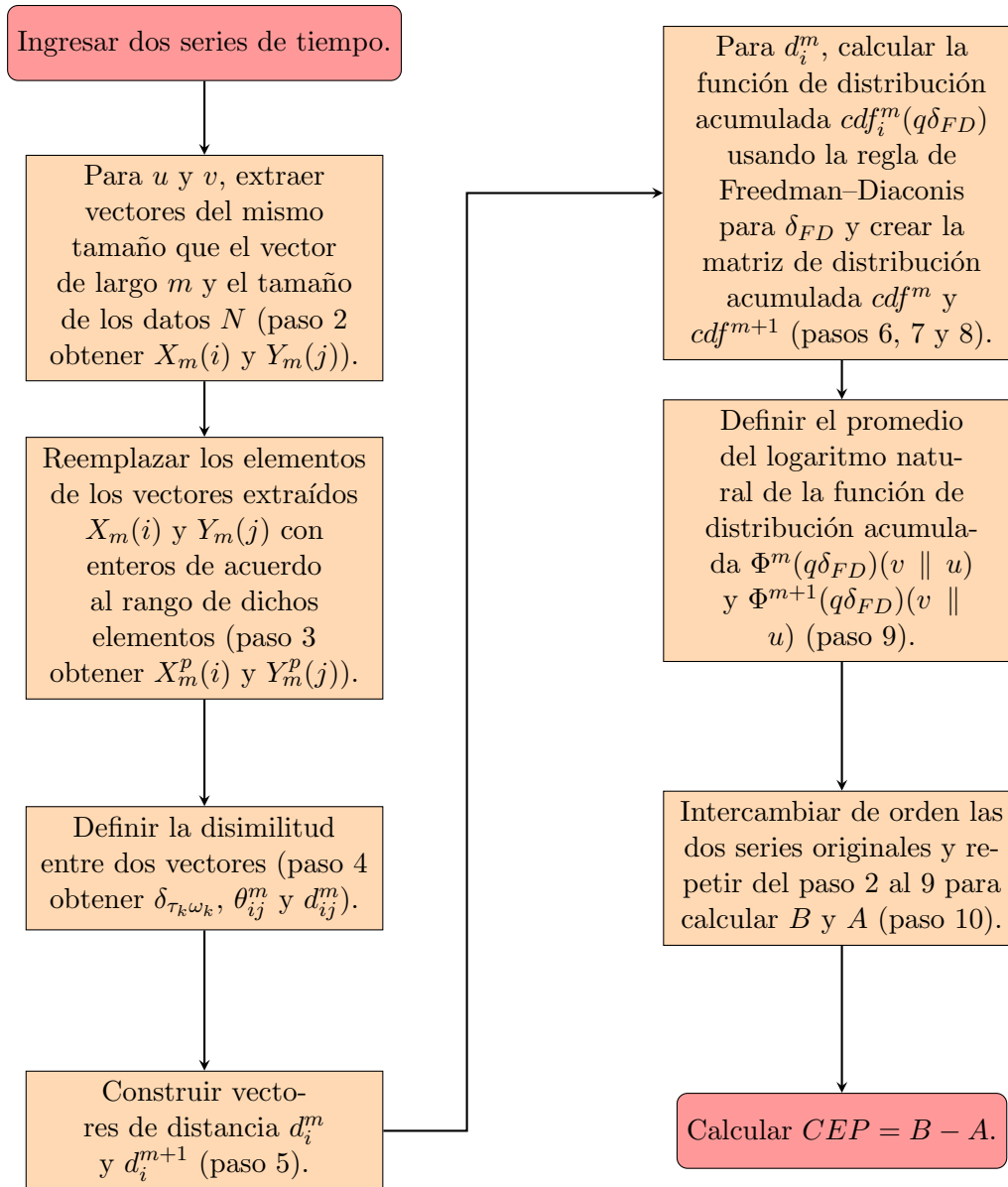
$$A = \frac{\Phi^{m+1}(q\delta_{FD})(v \parallel u) + \Phi^{m+1}(q\delta_{FD})(u \parallel v)}{2}.$$

La CEP finalmente se obtiene como:

$$CEP(N, m) = B - A.$$

Según la definición de CEP (Shang D. y cols., 2021), ésta se basa en el método del histograma acumulativo (CHM) considerando todos los valores posibles de r , lo que permite observar la sincronización entre series temporales de forma más exhaustiva con la adición de una “dimensión”. En realidad, la mayoría de las mediciones de disimilitud se basan en las distancias espaciales, lo que puede conducir fácilmente a juicios erróneos de secuencias con tendencias similares pero con distancias espaciales distantes. Por lo tanto, se propone la nueva medida de disimilitud basada en la función delta de Kronecker, la secuencia de símbolos, la distancia de Chebyshev y la regla de Freedman–Diaconis. Esto proporciona una medida alternativa más razonable y lógica para resolver los problemas de sincronización. La CEP puede encargarse de estimar la sincronización entre dos series temporales diferentes. Cuando hay poca sincronización entre ellas, el valor de la CEP es alto. Por el contrario, una fuerte sincronización entre ellas da lugar a un valor de CEP bajo.

A continuación se ilustra un diagrama de flujo (pasos 1-10) que se utilizará en la implementación computacional para las funciones en el software R (ver Apéndice A).



Capítulo 4

Resultados

4.1. Simulación

4.1.1. Mapa de Hénon

El mapa bidimensional de Hénon es uno de los ejemplos de sistemas dinámicos más estudiado ya que muestran un comportamiento caótico. El mapa de Hénon toma un punto (x_t, y_t) en el plano y lo mapea a un nuevo punto. Este se define como:

$$\begin{cases} x_{t+1} = 1 - ax_t^2 + y_t \\ y_{t+1} = bx_t \end{cases} .$$

Los valores de los dos parámetros a y b varían para generar tres conjuntos de series: $a = 1.4$ y $b = 0.3$ para Hénon 1, $a = 1.3$ y $b = 0.35$ para Hénon 2, por último $a = 1.1$ y $b = 0.35$ para Hénon 3 como se ve en la Figura 4.1.

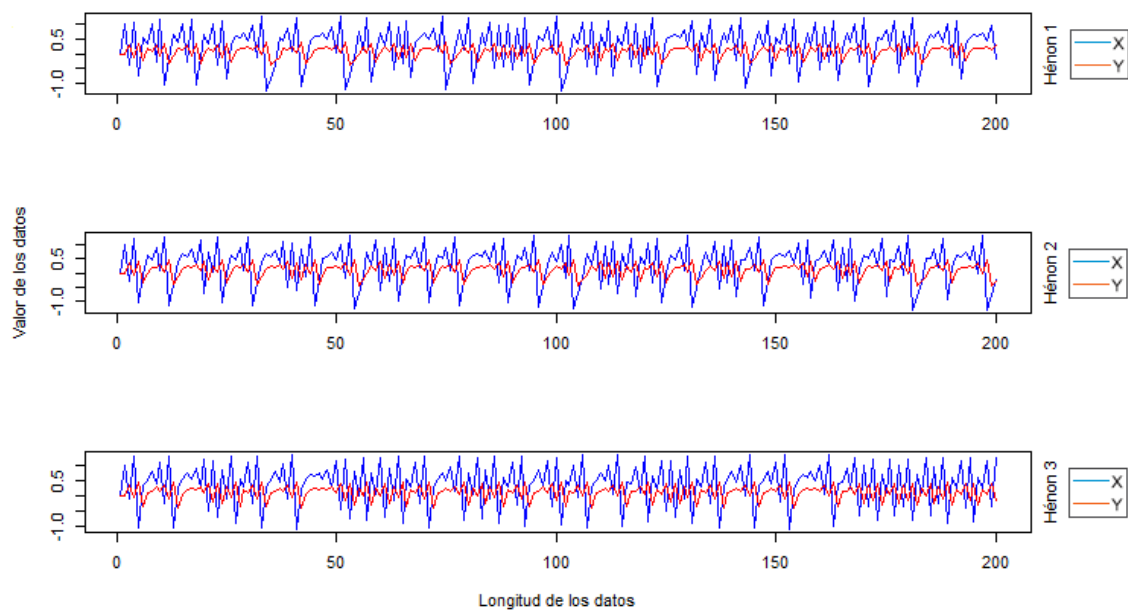


Figura 4.1: Diagramas esquemáticos de las tres series de Hénon

Fuente: Elaboración propia.

Los resultados obtenidos se basan en aplicar la CEP a estos datos simulados, tomando como series de entrada Hénon 1 con Hénon 3 y luego Hénon 1 con Hénon 2 para un $m = 3$, $m = 4$ y $m = 5$.

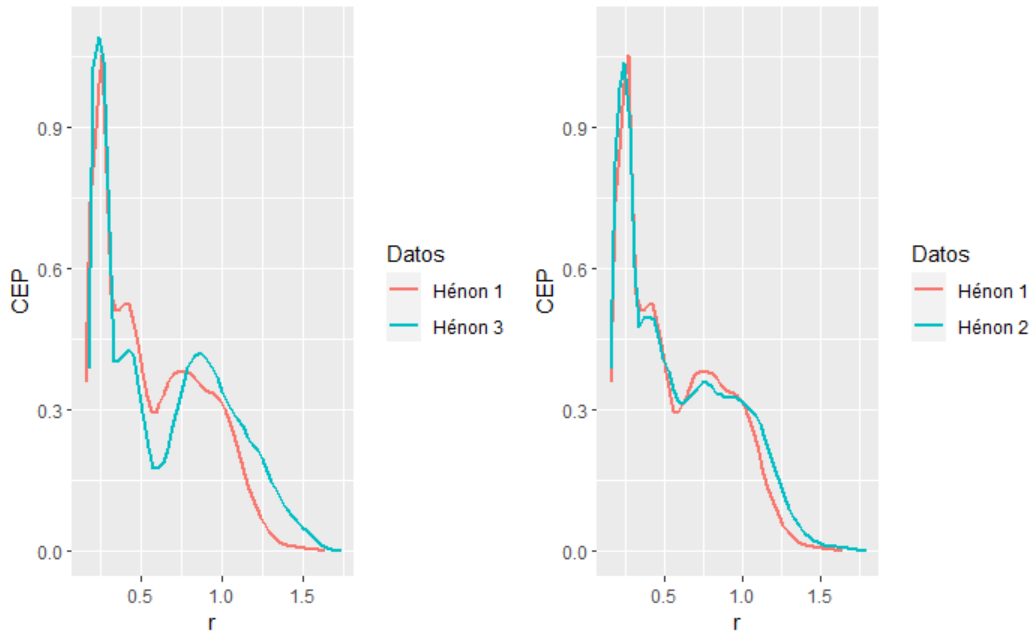


Figura 4.2: Relación entre la CEP y r con $m = 3$ para las series de mapas de Hénon
Fuente: Elaboración propia.

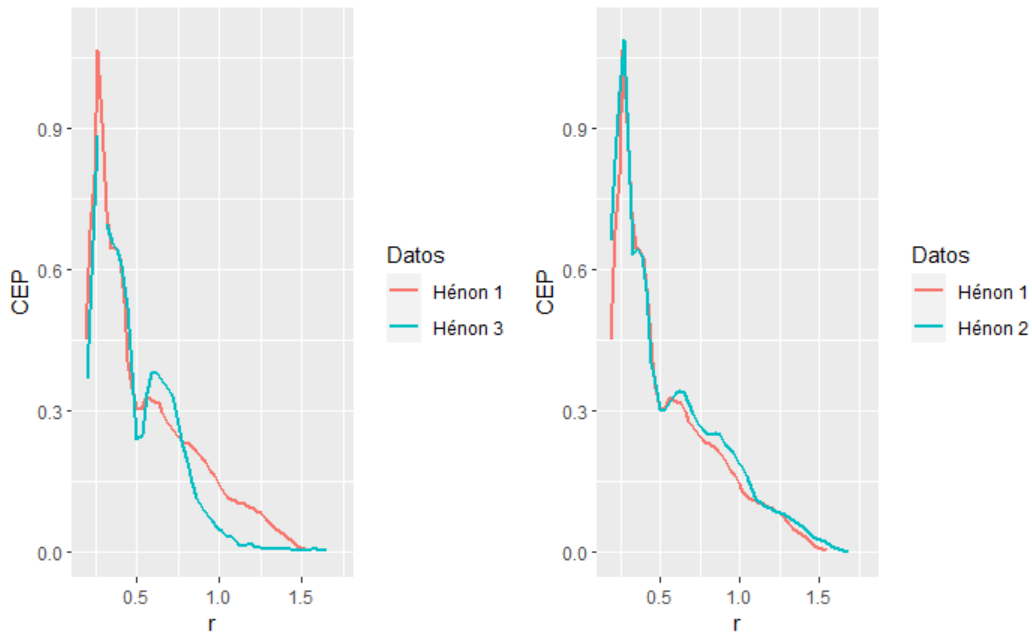


Figura 4.3: Relación entre la CEP y r con $m = 4$ para las series de mapas de Hénon
Fuente: Elaboración propia.

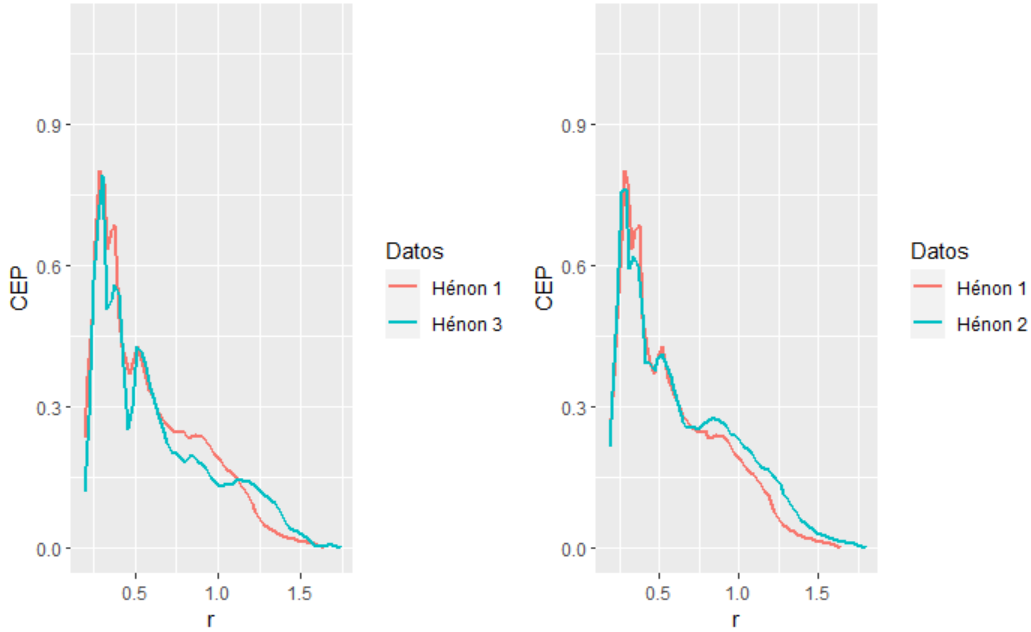


Figura 4.4: Relación entre la CEP y r con $m = 5$ para las series de mapas de Hénon
Fuente: Elaboración propia.

En cuanto a las series producidas por el mapa bidimensional de Hénon, se calcula el valor de la CEP de x e y por los 3 casos. Sabemos que Hénon 1, 2 y 3 son tres conjuntos de series que tienen condiciones similares debido a las selecciones de a y b . Las series de Hénon 1 son un tanto caóticas, y para comparar se utiliza éste mapa como referencia en las figuras 4.2, 4.3 y 4.4. Así podemos ver que el valor de la CEP de Hénon 3 es más fluctuante y difiere más que la CEP de Hénon 2 durante casi todo el proceso cuando r varía. Esto indica que el nivel de sincronía de las variables x e y de Hénon 1 y 2 son bastante similares, no así las variables x e y de Hénon 1 y 3 como se ve en la Figura 4.4, donde la sincronía de esta última es más fuerte (mientras menor sea la CEP, más fuerte es la sincronía) que Hénon 1 para distintos valores de r .

Por otra parte, como los tres casos pueden considerarse como dos series de vibraciones irregulares, el impacto en sus valores de la CEP provocado por la variación de r es casi imperceptible durante el último periodo ($r > 1.5$). Cabe señalar que al aumentar el parámetro m , en general la CEP tiende a disminuir en magnitud para Hénon 1, 2 y 3.

4.2. Tasas de cambio de divisas

Para probar este método se consideran las series de tiempo de tasas de cambio de dólar canadiense a americano (CAD/USD) y de dólar de Singapur a dólar americano (SGD/USD). Ambas series de tiempo son diarias y corresponden a las fluctuaciones que ocurrieron tras la crisis asiática. Se realizan los siguientes análisis previos utilizando los métodos más convencionales:

- Verificar si existe raíz unitaria mediante el test de Dickey–Fuller aumentado (Dickey, D. A. y Fuller, W. A., 1979) a las series originales y las de log-retorno.
- Realizar un ajuste ARMA-GARCH (Ramírez-Parietti, I. y cols., 2021) a cada serie, escogiendo el mejor modelo de cada serie usando el criterio AIC.
- Mostrar los parámetros estimados de cada modelo escogido, y chequear los supuestos del residuo.
- Realizar un test de causalidad de Granger (Granger, C. W. J., 1969) para ambas series originales y sus log-retornos.

4.2.1. Visualización de las tasas de cambio

En primera instancia se visualizan las series en las figuras 4.5 y 4.6. Principalmente se puede apreciar una alta volatilidad en ambas series, la cual comienza a aumentar posteriormente a la crisis asiática (comienzos de 1998).

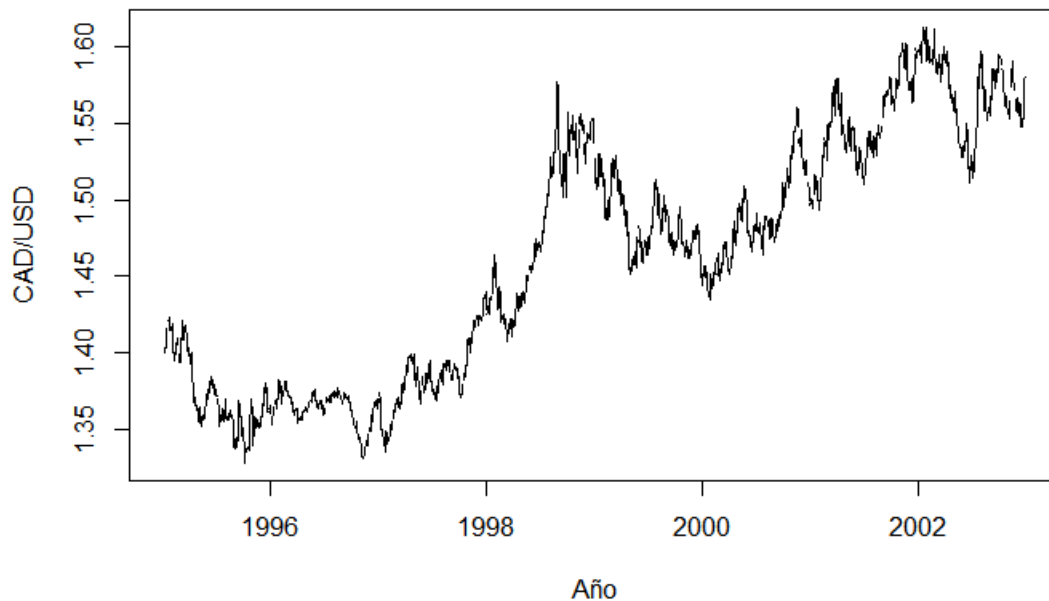


Figura 4.5: Dólar canadiense con respecto al dólar desde 1995 hasta 2002

Fuente: Elaboración propia.

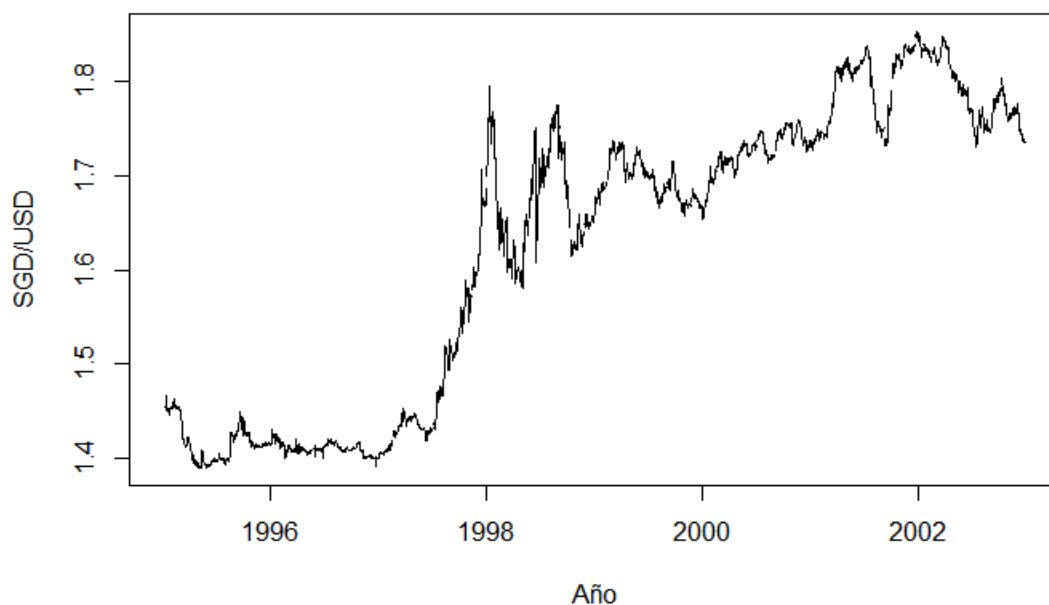


Figura 4.6: Dólar de Singapur con respecto al dólar desde 1995 hasta 2002
Fuente: Elaboración propia.

Se definen los log-retornos como $r_t = \log y_t - \log y_{t-1}$, donde y_t es la tasa de cambio. Para realizar un análisis comparativo se aplican los log-retornos a ambas series tal como se ilustra en las figuras 4.7 y 4.8. En la Figura 4.8 se puede observar que a mediados de 1997 el dólar de Singapur comenzó a tener una gran volatilidad. Ésta volatilidad también se presenta en la Figura 4.7 correspondiente al dólar canadiense, pero que sin embargo, la alta volatilidad comienza a fines de 1998.

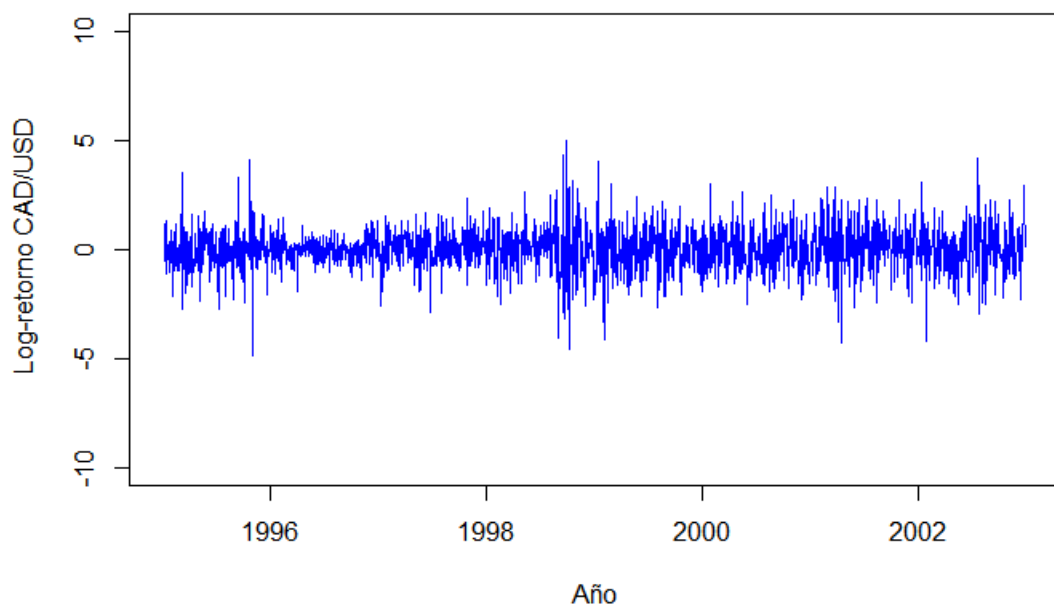


Figura 4.7: Log-retornos para CAD/USD desde 1995 hasta 2002
Fuente: Elaboración propia.

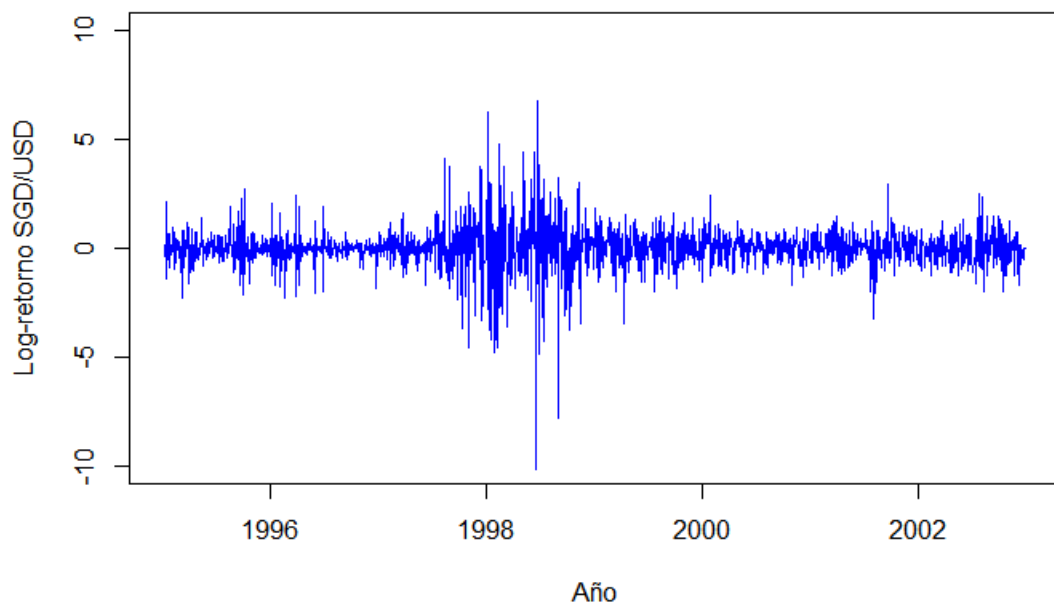


Figura 4.8: Log-retornos para SGD/USD desde 1995 hasta 2002
Fuente: Elaboración propia.

4.2.2. Prueba de Dickey-Fuller Aumentada

Para verificar si existe raíz unitaria es necesario aplicar la prueba ADF (Dickey, D. A. y Fuller, W. A., 1979), como se muestra en la Tabla 4.1. La estadística de la prueba ADF resulta en un número negativo. Cuanto más negativo es, más fuerte es el rechazo de la hipótesis nula de que existe una raíz unitaria para un cierto nivel de confianza. Al fijarnos en el valor p, en este caso para el dólar canadiense, no existe raíz unitaria lo que nos da indicios de que es un proceso estacionario, no así el dólar de Singapur en el cual si existe raíz unitaria, es decir es un proceso no estacionario. Con respecto a sus respectivos log-retornos y tomando en cuenta la diferenciación, se tiene que en ambos casos no existe raíz unitaria, por lo que son procesos estacionarios. De esta forma continuamos trabajando con los log-retornos de las series originales.

Serie	Dickey-Fuller	valor p
CAD/USD	-3.5278	0.0396
SGD/USD	-2.1805	0.5019
Log-retorno CAD/USD	-32.198	0.01
Log-retorno SGD/USD	-31.308	0.01

Cuadro 4.1: Resumen de prueba de Dickey-Fuller Aumentada para las series originales y para los log-retornos de las series.

4.2.3. Selección del modelo

Se procede a analizar las series de tiempo de tasas de cambio de divisas en dos periodos separados (Li-Zhi L. y cols., 2010; Ramírez-Parietti, I. y cols., 2021). Estos periodos se dividen en los años 1995-1998 y 1999-2003. Entre varios modelos, el modelo que mejor se ajusta a las series de tasas de cambio, fue el modelo ARMA(1,1)-GARCH(1,1) (basado en el AIC), el cual se define como:

$$\begin{aligned}
 \Phi(L)r_t &= \Theta(L)\epsilon_t \\
 \Phi(L) &= 1 - \phi_1 L \\
 \Theta(L) &= 1 - \theta_1 L \\
 \epsilon_t &= \sigma_t \nu_t \\
 \sigma_t^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,
 \end{aligned}$$

donde L es el operador de rezago ($Lr_t = r_t - r_{t-1}$), se asume que ν_t sigue una distribución t de Student con media cero, varianza unitaria y k grados de libertad, y $\Phi(L)$ y $\Theta(L)$ son los polinomios autorregresivos y de medias móviles, respectivamente

El modelo ARMA(1,1)-GARCH(1,1) es una composición del modelo ARMA para los log-retornos, y sus innovaciones se modelan a través de un modelo GARCH, el cual considera innovaciones presentes y pasadas modeladas por parámetros autorregresivos ($\alpha_i, i = 0, 1$), mientras que la varianza en el tiempo actual depende de la varianza en el tiempo pasada. Así, la varianza está en función del tiempo observado. Los resultados del ajuste se presentan en las tablas 4.2 y 4.3.

4.2.4. Estimación de parámetros

Los parámetros estimados para ambos periodos son:

Periodo	Serie	Modelo	AIC
1995-1998	CAD/USD	ARMA(1,1)-GARCH(1,1)	2.6368
1995-1998	SGD/USD	ARMA(1,1)-GARCH(1,1)	1.8387
1999-2003	CAD/USD	ARMA(1,1)-GARCH(1,1)	2.8317
1999-2003	SGD/USD	ARMA(1,1)-GARCH(1,1)	2.8032

Cuadro 4.2: Resumen de modelos ARMA-GARCH para log-retornos por periodos.

Periodo	Parámetro	CAN/USD		SGD/USD	
		Estimador	D.E.	Estimador	D.E.
1995-1998	ϕ_1	-0.518	0.176	0.413	0.175
	θ_1	0.583	0.164	-0.525	0.162
	α_0	0.012	0.007	0.005	0.002
	α_1	0.074	0.017	0.145	0.029
	β_1	0.916	0.018	0.853	0.030
1999-2002	ϕ_1	-0.405	0.224	0.260	0.417
	θ_1	0.464	0.219	-0.290	0.412
	α_0	0.034	0.007	0.031	0.022
	α_1	0.016	0.009	0.051	0.020
	β_1	0.948	0.007	0.915	0.038

Cuadro 4.3: Parámetros del modelo ARMA(1,1)-GARCH(1,1) para log-retornos de las tasas de cambio CAN/USD y SGD/USD

Con respecto al análisis residual de los modelos antes ajustados, se obtuvo que los residuos son ruido blanco utilizando la prueba de Ljung-Box (Fisher y Gallagher, 2012).

4.2.5. Prueba de causalidad de Granger

La causalidad de Granger (Granger, C. W. J., 1969) es una forma de investigar la causalidad entre dos variables en una serie temporal. El método es una explicación probabilística de la causalidad; utiliza conjuntos de datos empíricos para encontrar patrones de correlación. La causalidad está estrechamente relacionada con la idea de causa y efecto, aunque no es exactamente lo mismo. Una variable X es causal para la variable Y si X es la causa de Y o Y es la causa de X . Sin embargo, con la causalidad de Granger, no se está probando una verdadera relación de causa y efecto; lo que se quiere saber es si una variable concreta es anterior a otra en la serie temporal. En otras palabras, si encuentra causalidad de Granger en sus datos, no hay una relación causal en el verdadero sentido de la palabra. La hipótesis nula de la prueba es que los valores de x retardados no explican la variación de y . En otras palabras, se supone que $x(t)$ no es Granger causado por $y(t)$.

Se muestran las pruebas de causalidad de Granger hasta 50 rezagos, primero para cada serie original (Figura 4.9) y luego para sus log-retornos (Figura 4.10). Finalmente, se realiza una comparación entre el período previo y posterior a la crisis asiática (Figura 4.11). Para las series originales se tiene que el dólar canadiense es Granger-causado por el dólar de Singapur, pero esto no ocurre de forma inversa, lo cual tiene sentido ya que los efectos de la crisis ocurrieron en Asia primero y esto afectó a occidente un tiempo después. En cambio para los log-retornos se observa que el dólar canadiense no es Granger-causado por el dólar de Singapur, esto ocurre de forma inversa.

Podemos decir que al considerar los log-retornos, desaparece la causalidad de Granger, indicando que la prueba es afectada por la volatilidad. En este caso, se ve que solo en el periodo previo a la crisis el dólar canadiense es Granger-causado por el dólar de Singapur, en cualquier otro caso no ocurre, esto puede tener distintas explicaciones, principalmente que la prueba es susceptible al hecho de segmentar la serie por periodos.

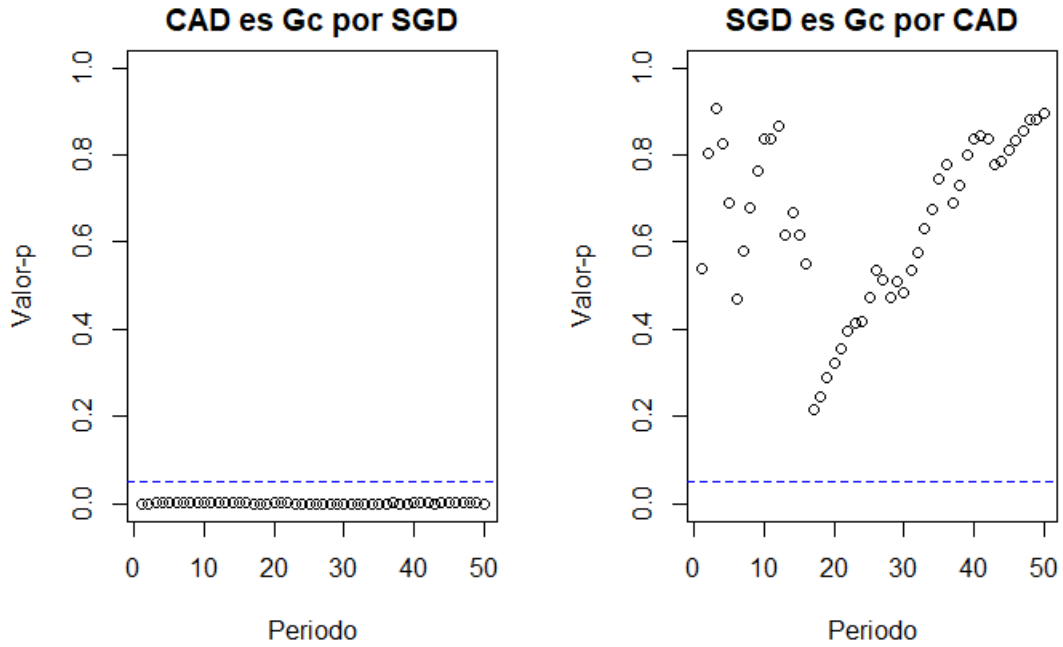


Figura 4.9: Prueba de causalidad de Granger para CAD/USD y SGD/USD
Fuente: Elaboración propia.

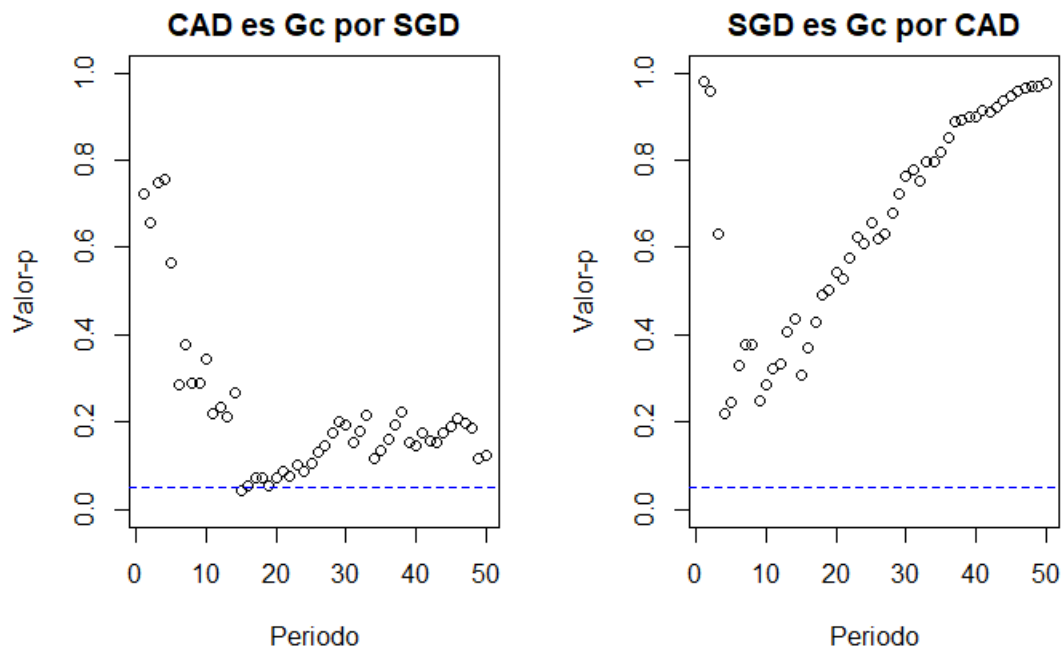


Figura 4.10: Prueba de causalidad de Granger para log-retornos de CAD/USD y SGD/USD
Fuente: Elaboración propia.

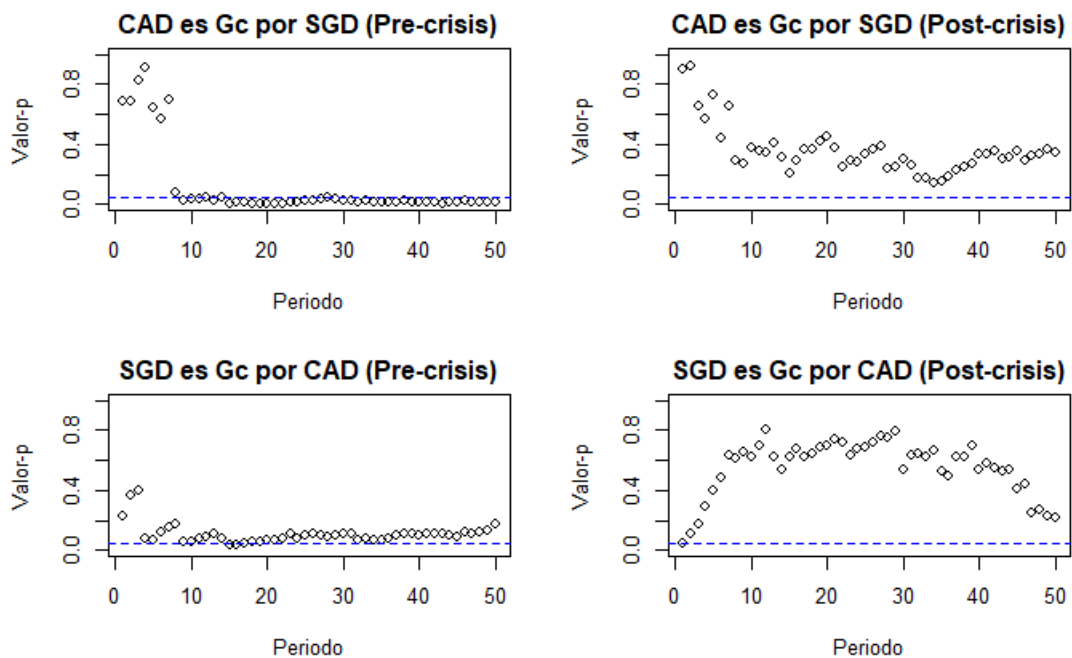


Figura 4.11: Prueba de causalidad de Granger para log-retornos de CAD/USD y SGD/USD separadas en 1995-1998 (Pre-crisis) y 1999-2002 (Post-crisis)
Fuente: Elaboración propia.

4.2.6. Aplicación de la entropía cruzada de perfiles

Tras la validación del método propuesto con los datos simulados, se desea probar la viabilidad de la CEP con estas tasas de cambio. Con ambas tasas divididas en un periodo pre-crisis (1995-1998) y en otro periodo post-crisis (1999-2002), se investiga la relación entre el CEP y la r calculada con valores de los log-retornos de las mismas tasas, antes y después de la crisis asiática.

Consideraremos las dimensiones $m = 2, 3, 4$ y 5 . Los resultados se muestran en las figuras 4.12, 4.13, 4.14 y 4.15 donde las líneas azules son la CEP que varía con r , generadas con los valores de los log-retornos de las tasas de cambio después de la crisis, y las líneas rojas se generan con valores de los mismos datos pero en un periodo previo a la crisis.

Se puede apreciar que las líneas azules se sitúan principalmente por encima de las líneas rojas, lo que implica que los valores de la CEP de los log-retornos de las tasas de cambio después de la crisis son mayores que los log-retornos de las tasas de cambio antes de la crisis. Los resultados revelan que los log-retornos de las tasas de cambio antes de la crisis tienen una mayor sincronización que después de la crisis, lo que también se ajusta al sentido común, ya que en el periodo posterior a la crisis la volatilidad de las tasas de cambio aumentó. Esto es consistente con los cálculos de la CSE (Li-Zhi L. y cols., 2010; Ramírez-Parietti, I. y cols., 2021).

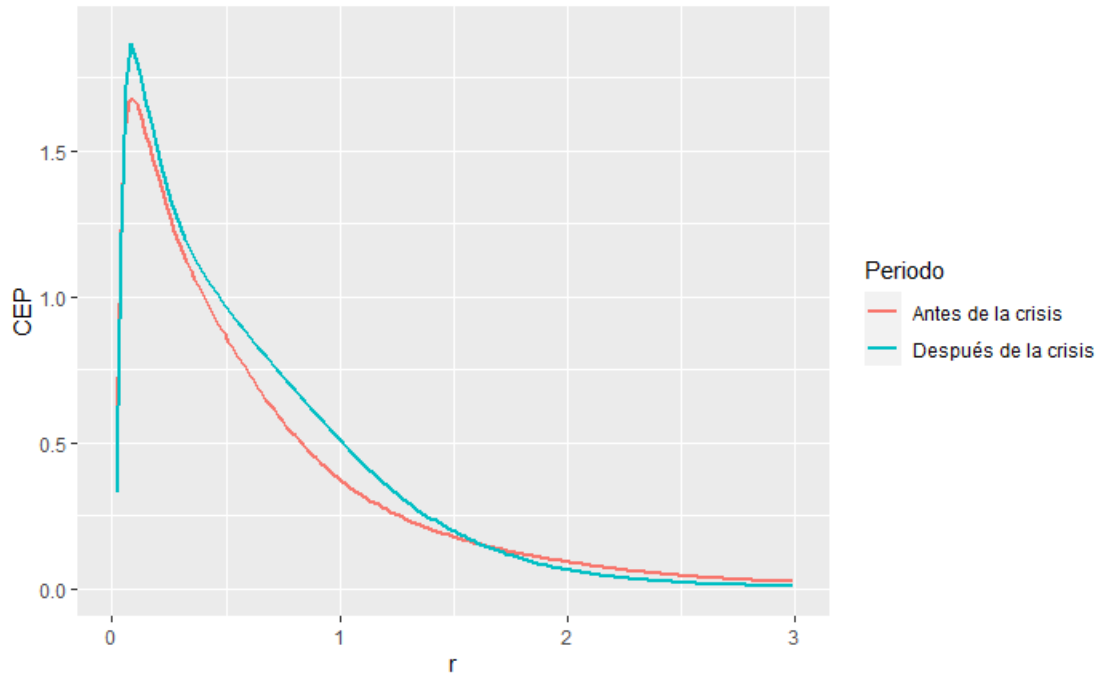


Figura 4.12: CEP para log-retornos de CAD/USD y SGD/USD separadas en 1995-1998 (pre-crisis) y 1999-2002 (post-crisis) para $m = 2$

Fuente: Elaboración propia.

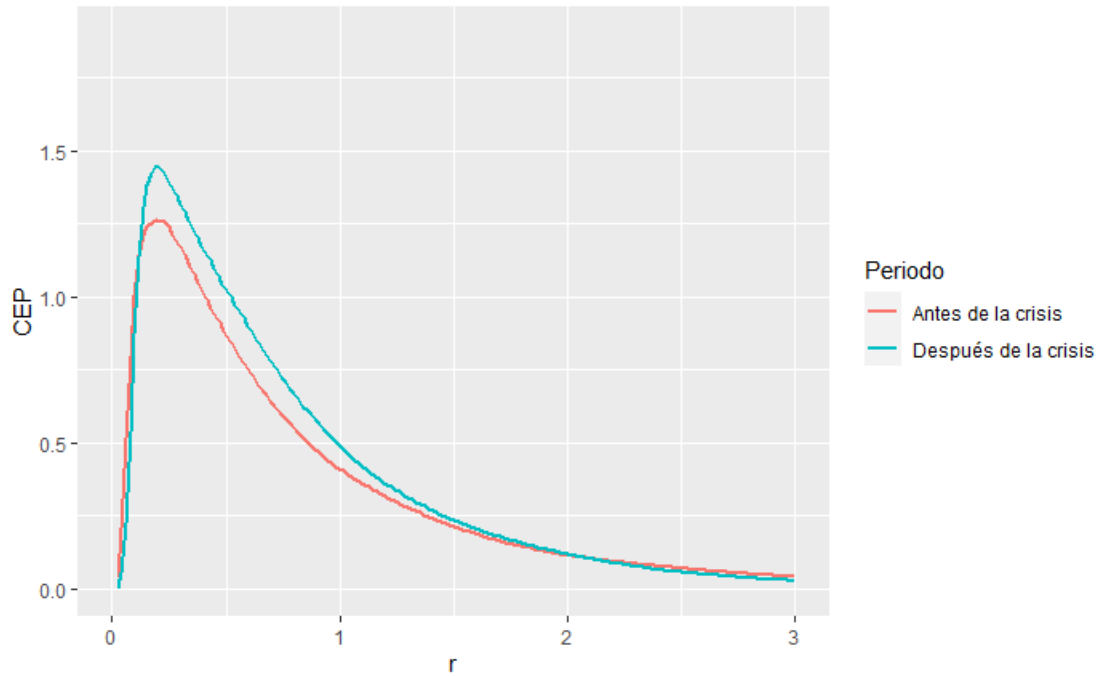


Figura 4.13: CEP para log-retornos de CAD/USD y SGD/USD separadas en 1995-1998 (pre-crisis) y 1999-2002 (post-crisis) para $m = 3$
Fuente: Elaboración propia.

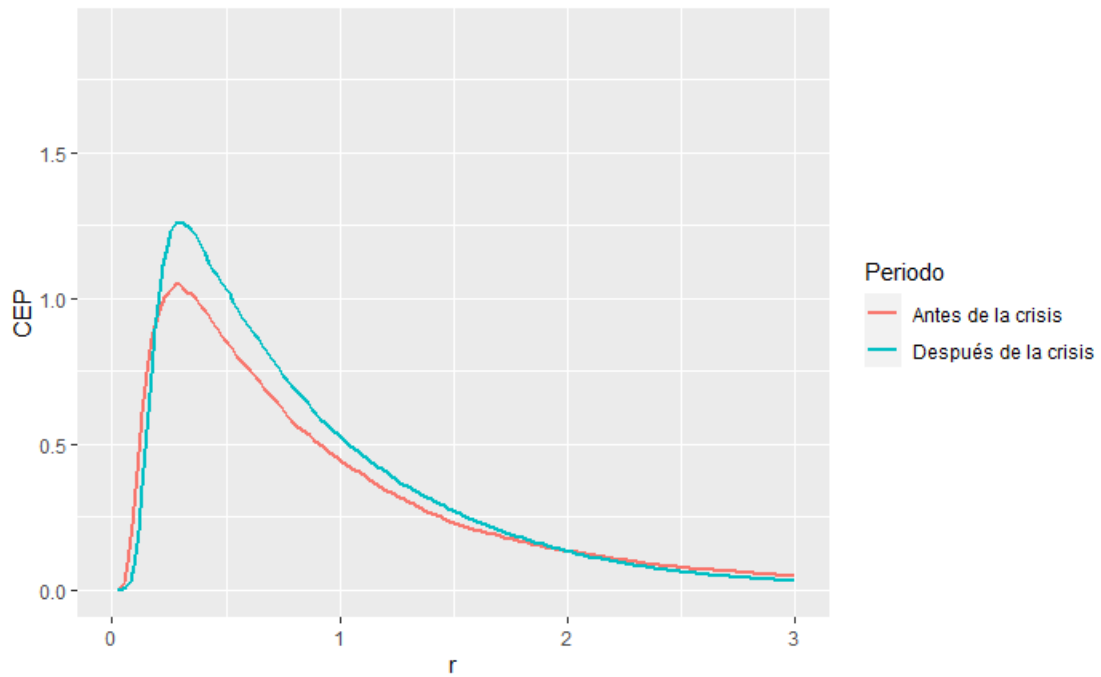


Figura 4.14: CEP para log-retornos de CAD/USD y SGD/USD separadas en 1995-1998 (pre-crisis) y 1999-2002 (post-crisis) para $m = 4$
Fuente: Elaboración propia.

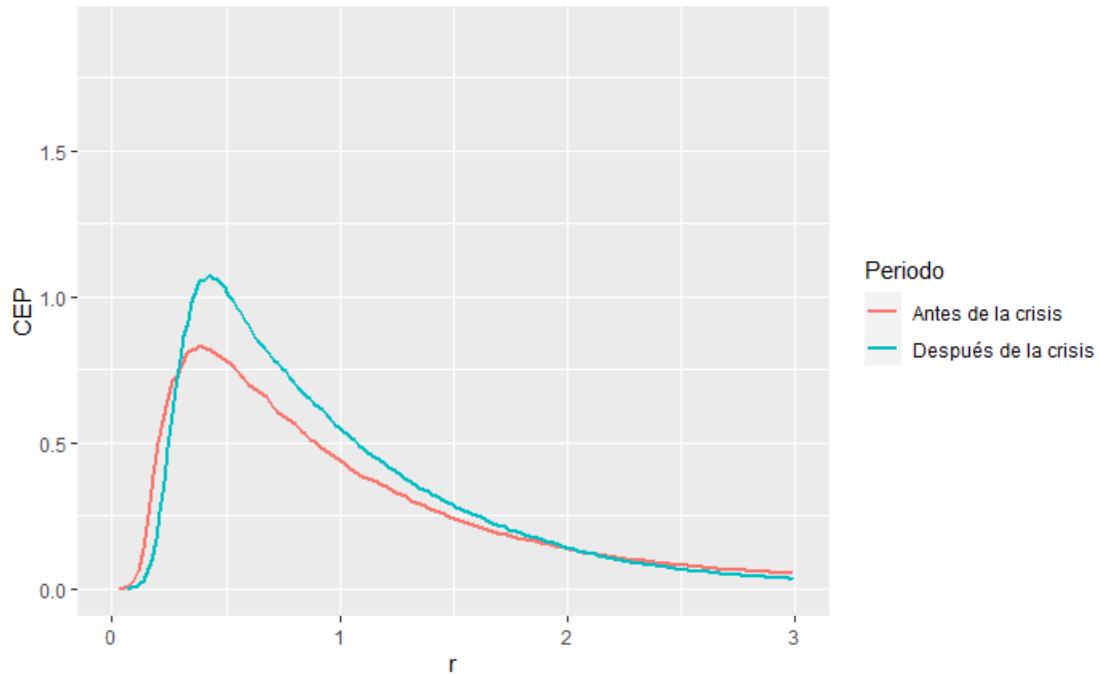


Figura 4.15: CEP para log-retornos de CAD/USD y SGD/USD separadas en 1995-1998 (pre-crisis) y 1999-2002 (post-crisis) para $m = 5$
Fuente: Elaboración propia.

Para validar la eficacia de este método cuando se aplica en la práctica, es útil generar ruido blanco como perturbación y añadirlo a la serie de tasa de cambio original, en este caso con el uso de sus log-retornos. El efecto del ruido se presenta en la Figura 4.7 y Figura 4.8. Los resultados anteriores muestran que la CEP no es sensible al ruido blanco, al menos hasta cierto punto.

Capítulo 5

Conclusiones

A lo largo de este trabajo se ha revisado la base teórica para la entropía cruzada muestral (CSE), se ha implementado el algoritmo para el índice de entropía de perfiles (EP), y se ha construido de una forma más robusta el índice de entropía cruzada de perfiles (CEP). Esta versión de la CEP se aplicó a series simuladas por mapas binarios de Hénon y a datos de las series de tasas de cambio estudiadas en el periodo previo y posterior a la crisis asiática. En concreto la CEP fue redefinida de tal manera que el número de segmentos del estimador del histograma fue obtenido usando la regla de Freedman-Diaconis (paso 6 de la sección 3.3), esto modificó el cálculo de las matrices de funciones de distribuciones acumuladas producidas por el CHM y el parámetro δ para el cálculo del rango (pasos 7 a 10 sección 3.3).

Además, el costo computacional de la CEP basada en la regla de Freedman-Diaconis es considerablemente menor al que utilizada la CSE. Ya que la CSE realiza una búsqueda de coincidencia de las distancias según el parámetro de tolerancia r , mientras que la CEP estima directamente el número de segmentos usando la regla de Freedman-Diaconis.

De esta forma, podemos mencionar que el aporte de esta tesis se resume en tres aspectos:

- 1) Se ha implementado en el software R el cálculo de la CEP permitiendo a otros investigadores utilizar esta herramienta.
- 2) Se ha obtenido el número de segmentos del estimador del histograma usando la regla de Freedman-Diaconis.
- 3) Se aplicó la CEP a series temporales de tasas de cambio (CAD/USD y SGD/USD) para analizar la sincronización antes y después de la crisis asiática.

Cabe señalar que existen otros métodos distintos al de Freedman-Diaconis para estimar el número de segmentos del estimador del histograma. Como, por ejemplo, las reglas de Sturges y Doane, las cuales fueron previamente consideradas en esta tesis, pero que no fueron presentadas dado que arrojaban valores similares. De igual forma se propone utilizar la CEP como un índice de sincronización para cualquier tipo de series de tiempo no necesariamente estacionarias, tales como: meteorológicas, oceanográficas, epidemiológicas (COVID 19), u otros tipos de series financieras.

Apéndice A

Código

```
library(readxl)
library(zoo)
library(ggplot2)

setwd("C:\\Directorio de trabajo")

CADDATA <- read_excel("DEXCAUS.xls", range = cell_rows(12:2097))
SGDDATA <- read_excel("DEXSIUS.xls", range = cell_rows(12:2097))

caddates <- CADDATA$`34702`
CADDATA$`1.4035`[CADDATA$`1.4035` == 0] <- NA
cadts <- zoo(CADDATA$`1.4035`, caddates)

sgddates <- SGDDATA$`34702`
SGDDATA$`1.454`[SGDDATA$`1.454` == 0] <- NA
sgdts <- zoo(SGDDATA$`1.454`, sgddates)

sgdts<-sgdts[!is.na(sgdts)]
cadts<-cadts[!is.na(cadts)]

n <- length(sgdts)
lrcad <- diff(log(cadts), lag=1)
lrsgd <- diff(log(sgdts), lag=1)

lrcad <- na.omit(lrcad)
lrsgd <- na.omit(lrsgd)

lrcad1 <- scale(lrcad[1:967])
lrcad2 <- scale(lrcad[968:1934])
lrsgd1 <- scale(lrsgd[1:967])
lrsgd2 <- scale(lrsgd[968:1934])

SCE <- function(u, v, m, tol){
  require(Trading)
```

```

A_m <- function(u, v, m, tol){
  A <- rep(0, length(u)-(m))
  x <- matrix(0, ncol = length(u)-(m), nrow = m+1)
  y <- matrix(0, ncol = length(u)-(m), nrow = m+1)
  for (i in 1:(length(u)-(m))) {
    x[,i] <- u[i:(m+i)]
    y[,i] <- v[i:(m+i)]
  }
  x.cols <- split(x, col(x))
  y.cols <- split(y, col(y))
  distsmat <- outer(x.cols, y.cols, FUN = Vectorize(Chebyshev_distance))
  for(i in 1:(length(u)-(m))){
    A[i] <- length(which(distsmat[i,] < tol))
  }
  output <- list(A = sum(A), dists = distsmat)
  return(output)
}

B_m <- function(u, v, m, tol){
  B <- rep(0, length(u)-m)
  distsmat <- matrix(0, ncol = length(u)-m, nrow = length(u)-m)
  x <- matrix(0, ncol = length(u)-m, nrow = m)
  y <- matrix(0, ncol = length(u)-m, nrow = m)
  for (i in 1:dim(x)[2]) {
    x[,i] <- u[i:(m+i)-1]
    y[,i] <- v[i:(m+i)-1] ### SE CONSTRUYEN LAS SERIES X Y RELLENANDO POR COLUMNAS
  }
  x.cols <- split(x, col(x))
  y.cols <- split(y, col(y))
  distsmat <- outer(x.cols, y.cols, FUN = Vectorize(Chebyshev_distance))
  for(i in 1:(length(u)-m)){
    B[i] <- length(which(distsmat[i,] < tol))
  }
  output <- list(B = sum(B), dists = distsmat)
  return(output)
}

idxA <- which( A_m(u,v,m,tol)$dists < tol )
matchmatA <- matrix(0, nrow = nrow(A_m(u,v,m,tol)$dists),
ncol = ncol(A_m(u,v,m,tol)$dists))
matchmatA[idxA] <- 1
KA <- 0

for (i in 1:(nrow(matchmatA))) {
  for (j in 1:(ncol(matchmatA))) {
    for (k in 1:(nrow(matchmatA))) {
      for (l in 1:(ncol(matchmatA))) {
        if (min(abs(i-k),abs(j-l)) <= m ) {

```

```

        KA <- KA + 1
    }
}
}
}

idxB <- which( B_m(u,v,m,tol)$dists < tol )
matchmatB <- matrix(0, nrow = nrow(B_m(u,v,m,tol)$dists),
ncol = ncol(B_m(u,v,m,tol)$dists))
matchmatB[idxB] <- 1
KB <- 0

for (i in 1:(nrow(matchmatB))) {
  for (j in 1:(ncol(matchmatB))) {
    for (k in 1:(nrow(matchmatB))) {
      for (l in 1:(ncol(matchmatB))) {
        if (min(abs(i-k),abs(j-l)) <= m ) {
          KB <- KB + 1
        }
      }
    }
  }
}

cse <- log((B_m(u,v,m,tol)$B)/(A_m(u,v,m,tol)$A+0.000000001))
cp <- exp(-cse)
estvar <- (cp-cp^2)/(sum(matchmatB))+(1/(sum(matchmatB))^2)*(KA-KB*(cp)^2)

final <- list(crossSampEn = cse, CP = cp, sterror = sqrt(estvar))
return(final)
}

D_matrix = function(u,v, m){

# Paso 1
N<-length(u)

# Paso 2 definimos los vectores
tmpltMatu<-matrix(0,m,(N-m+1))
tmpltMatv<-matrix(0,m,(N-m+1))
for (i in 1:(N-m+1)){
  tmpltMatu[,i]<-u[0:(m-1)+i]
  tmpltMatv[,i]<-v[0:(m-1)+i]
}

# Paso 3 calcular los vectores permutados
Pu<-matrix(0,m,(N-m+1))

```

```

Pv<-matrix(0,m,(N-m+1))
for (i in 1:(N-m)){

  # Vector auxiliar
  ui<-tmpltMatu[,i]

  # Nombra cada elemento con un numero entero
  names(ui)<-0:(m-1)

  # Ordenar segun el rango
  ui_s<-sort(ui)

  # Se guarda la permutación
  Pu[,i]<-as.numeric(names(ui_s))

  # Se repite lo mismo pa vi
  vi<-tmpltMatv[,i]
  names(vi)<-0:(m-1)
  vi_s<-sort(vi)
  Pv[,i]<-as.numeric(names(vi_s))
}

# Paso 4 Calcular la matriz de distancias
DisMatm<-matrix(NA,(N-m+1),(N-m+1))
for (i in 1:(N-m+1)){
  #Calculamos la disimilitud

  # Tomar el vector i-ésimo de las permutaciones
  tmpltVecM_p<-Pu[,i]

  # Genera una matriz que contiene este vector N-1 veces
  tmpltMatMa_p<-matrix(tmpltVecM_p,m,N-m+1)

  # Guarda las sumas de las coincidencias para cada vector
  di0<-apply(tmpltMatMa_p==Pv,2,sum)

  # Redondeamos y dividimos por m
  di0<-round(di0,3)/m

  #Calculamos un vector con la medida de distancia

  # Tomar el vector i-ésimo
  tmpltVecM<-tmpltMatu[,i]

  # Genera una matriz que contiene el vector i-ésimo repetido matLenM-1 veces
  tmpltMatMa<-matrix(tmpltVecM,m,N-m+1)

  # Guarda las distancias calculadas de la resta de las matrices anteriores

```

```

    di1<-apply(abs(tmplMatMa-tmplMatv),2,max)
    di1<-round(di1,3)

    # Calculamos la disimilitud y guardamos
    di=exp(-di0)*di1
    DisMatm[,i]<-di
  }
  return(DisMatm)
}

phi<-function(v,u,m){
  N<-length(u)
  # Calculamos Dm y Dm+1

  # Generamos la matriz de distancias para m
  Dm<-D_matrix(u,v,m)

  # Generamos la matriz de distancias para m+1
  Dm1<-D_matrix(u,v,m+1)

  # Las unimos en un conjunto
  D<-c(Dm,Dm1)
  A<- diff(range(D, na.rm=TRUE))
  B<- IQR(D, na.rm=TRUE)

  # Número de barras del histograma
  nbins<-round(A/(2*B/length(D)^(1/3)),0)

  # Calculaos delta y el rango de los elemento
  delta<-(max(D)-min(D))/nbins
  rango<-c(1:nbins)*delta+min(D)

  # Matriz vacia para las probabilidades acumuladas de tamaño N-m+1 x nbins
  allTmplVectHistM<-matrix(NA,N-m+1,nbins)

  #Matriz vacia para las probabilidades acumuladas de tamaño N-m x nbins
  allTmplVectHistM1<-matrix(NA,N-m,nbins)

  for (i in 1:(N-m+1)){
    # Calcula la probabilidad acumulada de las distacioas del i-esimo vector para m
    tplveci<-ecdf(Dm[,i])(rango)
    tplveci[tplveci==0]<-0.1^3
    allTmplVectHistM[i,<-tplveci
  }

  for (i in 1:(N-m)){
    # Calcula la probabilidad acumulada de las distacioas del i-esimo vector para m+1
    tplveci<-ecdf(Dm1[,i])(rango)
  }
}

```

```

    tplveci[tplveci==0]<-0.1^3
    allTpltVectHistM1[i,]<-tplveci
  }
  # Calcula la media del logaritmo de la matriz de probabilidades acumuladas para m
  phi_vu_m<-apply(log(allTpltVectHistM),2,mean)

  # Calcula la media del logaritmo de la matriz de probabilidades acumuladas para m+1
  phi_vu_m1<-apply(log(allTpltVectHistM1),2,mean)

  names(phi_vu_m)<-round(rango,5)
  names(phi_vu_m1)<-round(rango,5)

  # Se juntan en una lista para retornar la funcion
  sol<-list(m=phi_vu_m,m1=phi_vu_m1)
  return(sol)
}

# Calcular la CEP

CEP<- function(u,v,m){

  # Función para eliminar valores NA e inf.
  busq <- function(x){
    n=length(x)
    indice=NULL
    i=1
    while (is.na(x[i]) || is.infinite(x[i])) {
      i=i+1
    }
    indice = i
    x2 <- x[indice:n]
    return(x2)
  }

  sol_vu<-phi(v,u,m)
  sol_uv<-phi(u,v,m)
  A<-(sol_vu$m1+sol_uv$m1)/2
  B<-(sol_vu$m+sol_uv$m)/2
  sol0<-B-A
  names(sol0)<-names(sol_vu$m)
  sol<-busq(sol0)
  return(sol)
}

# Probamos para m=2
sol1<-CEP(lrcad1,lrsgd1,2)
sol2<-CEP(lrcad2,lrsgd2,2)

```

```
sol0<-c(sol2,sol1)
col<-c(rep('Después de la crisis',length(sol2)),rep('Antes de la crisis',length(sol1)))
dataplot<-data.frame(CEP=sol0,r=as.numeric(c(names(sol2),names(sol1))),Periodo=col)

ggplot(dataplot,aes(x=r,y=CEP))
+geom_line(aes(color=Periodo),size=1)
+xlim(c(0,3))+ylim(c(0,1.9))
```

Referencias

- Fisher, T., y Gallagher, C. (2012). New weighted portmanteau statistics for time series goodness of fit testing. *Journal of the American Statistical Society*, 107, 777-787.
- Freedman, D., y Diaconis, P. (1981). On the histogram as a density estimator: L2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 57, 453-476.
- Lake, D., Richman, J., Griffin, M., y Moorman, J. (2002). Sample entropy analysis of neonatal heart rate variability. *American journal of physiology. Regulatory, integrative and comparative physiology*, 283, R789-97. doi: 10.1152/ajpregu.00069.2002
- Richman, J., y Moorman, J. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American journal of physiology. Heart and circulatory physiology*, 278, H2039-49. doi: 10.1152/ajpheart.2000.278.6.H2039
- Contreras-Reyes, J.E. (2021). Mutual information matrix based on asymmetric shannon entropy for nonlinear interactions of time series. *Nonlinear Dynamics, in press.*, 104(4), 3913-3924. Descargado de <https://link.springer.com/article/10.1007/s11071-021-06498-w> doi: <https://doi.org/10.1007/s11071-021-06498-w>
- Cover T., y Thomas J. (2012). *Elements of information theory*. Somerset: 2nd ed. Wiley.
- Dickey, D. A., y Fuller, W. A. (1979). Distribution of the estimates for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427-431. Descargado de <https://www.jstor.org/stable/2286348> doi: <https://doi.org/10.2307/2286348>
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424-438. Descargado de <https://www.jstor.org/stable/1912791> doi: <https://doi.org/10.2307/1912791>
- Karmakar C., Udhayakumar R., y Palaniswami M. (2020). Entropy profiling: A reduced-parametric measure of kolmogorov-sinai entropy from short-term hrv signal. *Entropy*, 22, 1396. Descargado de <https://www.mdpi.com/1099-4300/22/12/1396> doi: <https://doi.org/10.3390/e22121396>
- Kullback, S., y Leibler, R. A. (1951). On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79-86. Descargado de <https://www.projecteuclid.org/journals/annals-of-mathematical-statistics/volume-22/issue-1/On-Information-and-Sufficiency/10.1214/aoms/1177729694.full?tab=ArticleLink> doi: <https://doi.org/10.1214/aoms/1177729694>
- Li-Zhi L., Xi-Yuan Q., y Heng-Yao L. (2010). Cross-sample entropy of foreign exchangetime series. *Physica A*, 389, 4785 - 4792. Descargado de <https://www.sciencedirect.com/science/article/abs/pii/S0378437110005030?via%3Dihub> doi: <https://doi.org/10.1016/j.physa.2010.06.013>
- Mantegna R.N., y Stanley H.E. (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge University press.
- Ramírez-Parietti, I., Contreras-Reyes, J.E., y Idrovo-Aguirre, B.J. (2021). Cross-sample entropy estimation for time series analysis: a nonparametric approach. *Nonlinear Dynamics*, 105.

- Descargado de <https://link.springer.com/article/10.1007/s11071-021-06759-8> doi: <https://doi.org/10.1007/s11071-021-06759-8>
- Shang D., Shang P., y Zhang Z. (2021). Efficient synchronization estimation for complex time series using refined cross-sample entropy measure. *Commun Nonlinear Sci Numer Simulat*, 94, 105556. Descargado de <https://www.sciencedirect.com/science/article/abs/pii/S1007570420303865?via%3DiHub> doi: <https://doi.org/10.1016/j.cnsns.2020.105556>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379-423. Descargado de <https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1948.tb01338.x> doi: <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Udhayakumar, R., Karmakar, C., y Palaniswami, M. (2017). Approximate entropy profile: a novel approach to comprehend irregularity of short-term hrv signal. *Nonlinear Dynamics*, 88, 1-15. doi: 10.1007/s11071-016-3278-z
- Xia, J., y Shang, P. (2012). Multiscale entropy analysis of financial time series. *Fluctuation and Noise Letters*, 11. doi: 10.1142/S0219477512500332