**UNIVERSIDAD POLITÉCNICA DE CARTAGENA**

*DEPARTAMENTO DE MÉTODOS CUANTITATIVOS E INFORMÁTICOS*

# *ENTROPÍAS TOPOLÓGICA Y DE PERMUTACIÓN. APLICACIÓN A LA DINÁMICA ECONÓMICA*

*José Miguel Rodriguez Gómez*

*2009*

***UNIVERSIDAD POLITÉCNICA DE CARTAGENA***

*DEPARTAMENTO DE MÉTODOS CUANTITATIVOS E INFORMÁTICOS*

# *ENTROPÍAS TOPOLÓGICA Y DE PERMUTACIÓN. APLICACIÓN A LA DINÁMICA ECONÓMICA*

**Doctorando:**

*José Miguel Rodriguez Gómez*

**Directores:**

*Dr. D. José Salvador Cánovas Peña*

*Dr. D. Manuel Ruiz Marín*

La realización de esta memoria no habría sido posible sin el apoyo y la ayuda de varias personas. Vaya para ellos todo mi agradecimiento.

Como no podia ser de otro modo mi primer pensamiento va dirigido a mi mujer, quien con sus comentarios y paciencia ha hecho posible que me enfrente a esta memoria.

En segundo lugar todo mi agradecimiento y admiración a mis dos directores, el Dr. D. Jose Salvador Cánovas y el Dr. D. Manuel Ruiz, no sólo por su ayuda que ha sido toda, sino porque he tenido el placer de poder trabajar con dos personas, que no dan la apariencia de lo que conocen, pero que son dos auténticos sabios en la materia, que la dominan como nadie y de ahí el reconocimiento de la comunidad científica, que refiere en muchas ocasiones sus trabajos.

A todos vosotros muchas gracias.

# RESUMEN

La entropía es la noción conductora de esta memoria, tanto en el contexto de los sistemas dinámicos discretos como en su aplicación al análisis de series temporales.

El Capítulo 1 lo hemos dedicado a una pequeña introducción donde hablamos del origen de la palabra entropía, de la relación de la misma con la Física, La Economía, La Ecología, etc. Definimos no formalmente lo que se entiende por un sistema dinámico y damos algún ejemplo.

Nuestro Capítulo 2 reúne una serie de conceptos y propiedades que serán de gran utilidad al lector, para la mejor compresión de esta memoria.

Será en nuestro Capítulo 3 donde introduciremos nuestra nueva definición de entropía topológica $ent(f)$ y comprobaremos como de buena puede llegar a ser, cuando tratamos con funciones definidas en conjuntos no compactos, que modelan algunos fenómenos de las ciencias experimentales y en particular de la economía.

A $ent(f)$ la denominaremos entropía *Cánovas-Rodriguez* de la función $f$, que se inspira en cierto sentido, en la definición establecida por *Gurevič* para el caso de los homeomorfismos shift. Una vez establecida dicha definición, comprobaremos que cumple la mayoría de las propiedades clásicas de la entropía topológica introducidas para compactos, con anterioridad por otros autores. También es cierto que esta definición sí verifica la filosofía de que si una función presenta entropía *Cánovas-Rodriguez* positiva la función tiene una dinámica compleja.

Es el Capítulo 4 el que dedicaremos a poner de manifiesto el uso de la entropía como medida de la dependencia entre series temporales. Más concretamente utilizaremos el concepto de entropía de permutación para construir un test consistente de independencia entre series temporales que se distribuye asintóticamente como una Chi-cuadrado.

En el Capítulo 5 estableceremos algunas relaciones entre la entropía topológica y la de permutación para funciones definidas en la recta real. También demostraremos algunas propiedades que satisface la entropía de permutación.

Después pasaremos a comprobar como hay ocasiones en que ambas definiciones coinciden como por ejemplo, para funciones monótonas con un numero finito trozos en el intervalo. También veremos como para una función $f$ definida en la recta real tal que $f$ no tiene un número finito de trozos, entonces $ent(f)$ es distinta de su entropía de permutación, tal como ocurrie para las funciones definidas en el intervalo.

# ABSTRACT

Entropy is the driving concept of this thesis, both in the context of discrete dynamical systems and in its application to time series analysis.

Chapter 1 is devoted to a brief introduction where we discuss the origin of the word entropy. Its relationship with physics, economy, ecology, etc. Then we not formally introduce what is meant by a dynamical system and we give some examples.

Our Chapter 2 brings together a number of concepts and properties that will be useful to the reader, to better understand the following chapters.

In Chapter 3 we introduce our new definition $ent(f)$ of entropy for functions defined in non-compact sets, which is very common when modeling some phenomena of the experimental sciences and especially economics.

This new concept of entropy is what we cal *Canovas-Rodriguez* entropy which is inspired in a sense, in the definition given by *Gurevič* for the case of homeomorphism shift. It is also checked that this definiton of entropy meets most of the classical properties of topological entropy. This definition verifies that if a function presents positive *Canovas-Rodriguez* entropy then the function has complex dynamic.

Chapter 4 is devoted to highlighting the use of entropy as a measure of dependence among time series. Moreover we use the concept of permutation entropy to construct a consist test for independence that asymptotically distributes as a Chi-Square distribution.

Chapter 5 establishes some relations between topological entropy and permutation entropy for functions defined in the real line. We also show some properties that are satisfied by permutation entropy. Then we will check when both definitions coincide, for example for piecewise monotonous functions in the interval with a finite number of pieces. We will also see that if $f$ is a function defined in the real line with an infinite number of pieces, then

$ent(f)$ is not equal to permutation entropy of $f$.

# Índice general

# Capítulo 1

# Introducción

Originalmente, “entropía” surgió como palabra acuñada del griego, de em (en - en, sobre, cerca de...) y sqopg (tropêe - mudanza, giro, alternativa, cambio, evolución...). El término fue usado primeramente en 1850 por el físico alemán Rudolf Julius Emmanuel Clausius 1822-1888 que intentó definir una magnitud que midiera el desorden presente en el estado físico de un sistema, e introdujo el concepto de entropía en el marco de la Termodinámica. Enunciando el segundo principio de la Termodinámica en términos de entropía como:

**En todo proceso físico, la entropía permanece constante o aumenta, y si aumenta, el proceso es irreversible.**

El término Entropía (tendencia natural de la pérdida del orden) puede referirse:

1. <u>En física y química</u> a:

   **Entropía termodinámica**, una magnitud que mide la parte de la energía que no puede utilizarse para producir un trabajo; es el grado de desorden que poseen las moléculas que integran un cuerpo, o también el grado de irreversibilidad alcanzada después de un proceso que implique transformación de energía.

**Entropía de Kolmogorov** objeto o dimensión estudiado en la física y matemáticas a partir de las homotecias.

2. <u>En teoría de la información</u> a:

   El grado de incertidumbre que existe sobre un conjunto de datos.

3. <u>En matemáticas</u> a:

   **Entropía topológica**, la correspondiente a la cantidad real asociada a todo sistema topológicamente dinámico.

   **Entropía métrica**, la correspondiente a la cantidad real asociada a todo sistema dinámico medible.

4. <u>En ecología</u> a:

   Es una medida asociada a la biodiversidad.

Vamos a imaginar que tenemos una caja con tres divisiones; dentro de la caja y en cada división se encuentran tres tipos diferentes de canicas: azules, amarillas y rojas, respectivamente. Las divisiones son movibles así que me decido a quitar la primera de ellas, la que separa a las canicas azules de las amarillas. Lo que estamos haciendo dentro del punto de vista de la entropía es quitar un grado o índice de restricción al sistema; antes de que quitáramos la primera división, las canicas se encontraban separadas y ordenadas en colores: en la primera división las azules, en la segunda las amarillas y en la tercera las rojas, estaban restringidas a un cierto orden. Al quitar la segunda división, estamos quitando también otro grado de restricción. Las canicas se han mezclado unas con otras de tal manera que ahora no las podemos tener ordenadas pues las barreras que les restringían han sido quitadas. La entropía de este sistema ha aumentado al ir quitando las restricciones, pues inicialmente, había un orden establecido y al final del proceso (el proceso es en este caso el quitar las divisiones de la caja) no existe orden alguno dentro de la caja. La entropía es en este caso una medida del orden (o desorden) de un sistema o de la falta de grados de restricción.

Teniendo en cuenta nuestra caja ya sin las separaciones, tenemos a las canicas revueltas unas con otras, es decir, sin un orden. Si el proceso que efectuamos de quitar las divisiones fuera reversible, las canicas tendrían que ordenarse espontáneamente en azules, amarillas y rojas, según el orden de las divisiones. Esto no ocurrirá. El proceso que efectuamos con nuestra caja de canicas fue un proceso irreversible, en donde una vez terminado, el orden que había en las condiciones iniciales del sistema ya nunca volverá a establecerse. El estudio de este tipo de procesos es importante porque en la naturaleza todos los procesos son irreversibles.

¿Para qué sirve la entropía?

La entropía, como medida del grado de restricción o como medida del desorden de un sistema, o bien en ingeniería, como concepto auxiliar en los problemas del rendimiento energético de las máquinas, es una de las variables termodinámicas más importantes. Su relación con la teoría del caos le abre un nuevo campo de estudio e investigación.

Para referirse a esta tendencia natural al desorden, el Físico Alemán Rudolf Julius Emmanuel Clausius, fue el primero en utilizar el término Entropía como ya hemos advertido al inicio de este Capítulo, ¿pero entonces por qué pensar, en la entropía, cuando se piensa en la economía?.

Entre las muchas definiciones que pueden encontrarse de economía, algunas la establecen como una ciencia social, y otras mas exactamente como la manera en que la sociedad administra sus recursos escasos. La economía se refiere a la sociedad y ésta a su vez puede ser considerada como un sistema complejo, por la forma como cada uno de sus individuos se interrelacionan, porque la acción de cada uno de ellos incide sobre las decisiones de los demás.

Esta consideración de la sociedad como un sistema complejo, es lo que nos permite trasladar al análisis de la economía las herramientas desarrolladas en la física.

En economía resulta muy importante el concepto de equidad, al igual que el concepto de eficiencia. Estos conceptos sobresalen constantemente en cualquier situación, bien sea en el desarrollo o explicación de una teoría o bien en las discusiones relacionadas con la definición de la política económica.

La eficiencia hace referencia a la maximización de la producción con los recursos escasos de que dispone la sociedad, y la equidad sugiere la distribución equitativa de los beneficios obtenidos con dichos recursos, entre los miembros de la sociedad.

Si se tratara de ordenar la sociedad como se hace con una baraja de cartas, teniendo como elementos de orden la eficiencia y la equidad, nos encontraríamos con que al igual que en la baraja, existe un infinito de posibilidades en que puede organizarse una sociedad, pero a aquella que conjuga la eficiencia y la equidad le corresponde una bajísima probabilidad. De hecho, en los textos de economía se menciona el hecho de que la sociedad se enfrenta a una disyuntiva entre estos dos principios.

De esta manera, al considerar la sociedad como un sistema complejo y teniendo en cuenta que los principios de eficiencia y equidad, serían los referentes básicos de lo que podría considerarse el orden de la sociedad desde el punto de vista económico, es posible concluir que en la sociedad existe un alto nivel de entropía. Conclusión que resulta obvia si consideramos el principio o ley que determina: “Un sistema formado de un número muy grande de componentes tiende a evolucionar espontáneamente hacia las situaciones de máxima entropía". Si la sociedad como un sistema complejo que es, tiende hacia la máxima entropía, se requiere de una fuerza ordenadora que le permita dejar de ser un sistema aislado, de tal forma que en su interior se produzcan intercambios positivos para propia la sociedad. Su equivalente en la física sería el intercambio de energía que se genera entre los sistemas no aislados.

En términos económicos, mantener la sociedad como un "sistema aislado" equivaldría a una economía en la que su funcionamiento está determinado por las fuerzas de los mercados, es decir, un sistema al que no se le adiciona un componente ordenador, por considerar que aquellos, los mercados, son la opción que garantiza la mayor eficiencia para asignar los recursos.

En los sistemas económicos también resulta pues indispensable utilizar el principio de entropía, puede ayudarnos a entender mejor el paper ordenador que constitucionalmente le corresponde al estado y que permite o garantiza la mayor eficiencia en la asignación de recursos.

Un sistema dinámico es un modo de describir el recorrido a lo largo del tiempo de todos los puntos de un espacio dado $\mathbb{E}$. El espacio $\mathbb{E}$ puede imaginarse como el espacio de estados de cierto sistema físico, económico, etc. Matemáticamente, $\mathbb{E}$ puede ser un espacio topológico o un subconjunto abierto de un espacio topológico. En nuestro caso y a lo largo de esta memoria consideraremos como espacio topológico a $\mathbb{R}$, es decir, la recta real y mas concretamente siempre que podamos nos restringiremos al intervalo $I = [0, 1]$.

Cuando un sistema dinámico, donde la evolución sigue un estado inicial típico, es muy sensible a perturbaciones de este estado inicial, en general, tales perturbaciones crecen exponencialmente.

Para finalizar esta introducción explicaremos como se ha estructurado la memoria que aquí presentamos.

Hemos introducido un Capítulo 2 al cual hemos titulado preliminares, y lo hemos dedicado a recoger conceptos básicos, que serán necesarios y que se han utilizado durante el desarrollo de la presente memoria, la mayoría de los cuales nos son familiares y nos permiten entender mejor el contenido de la misma.

En el Capítulo 3 hemos introducido una nueva definición de entropía topológica para funciones continuas tales que, al menos las funciones continuas reales con esta nueva definición mantengan la filosofía general, es decir, entropía topológica positiva implica que la función tiene un comportamiento dinámico complicado. Además, hemos perseguido que nuestra definición mantenga algunas propiedades que se cumplen en la definición clásica de entropía topológica introducida para conjuntos compactos.

Estudiaremos por tanto sus propiedades fundamentales y probaremos algunos resultados para esta nueva definición, que son equivalentes a resultados análogos para la entropía topológica de funciones continuas sobre intervalos compactos. Todos estos resultados se han recogido en una publicación, ver en Cánovas J., Rodriguez J. (2005) que ha sido referenciada por Dai, Xiongping y Jiang, Yunping; (2008), K. Brucks y H. Bruin, (2004) y Liu, L., Wang, Y., Wei, G. (2009).

Es en el Capítulo 4 donde nos enfrentamos con un problema crucial en economía, la identificación de una posible relación dinámica entre dos series temporales. El hecho de que algunos fenómenos económicos puedan ser descritos por series temporales, indica que es útil y relevante un test que determine la existencia de una relación, de dependencia entre series, para luego investigar la naturaleza dinámica de la misma. Aclarando ese tipo de dependencia (y, si es posible, su causa) tendremos una información importante que nos permita luego llevar a cabo las predicciones.

Han sido escasos los intentos en la literatura para contrastar la dependencia entre series temporales. Fue Haug (1976) quien propuso un test de independencia entre dos series temporales basado en la función residual de correlación cruzada. Pierce (1977) encontró que, cuando Haug utiliza su test, la hipótesis de independencia puede llegar a ser rechazada para el caso de algunas series de medidas de la actividad económica que tienen una relación que parece estar bien establecida. En relación con esto, Geweke (1981) encontró que el test de Haugh a menudo presenta una baja potencia en series reales, lo que podría explicar los resultados de Pierce (1977).

El procedimiento de Haug (1976) ha sido extendido y mejorado por diferentes autores. Koch, et al. (1986) lo extendió con una modificación que permite introducir un patrón mas potente para los coeficientes de correlación cruzada. En un influyente artículo, Hong (1989) propuso un criterio, basado en técnicas de estimación por núcleo. Existen otras extensiones del contraste de Haug disponibles, (véase, por ejemplo, Pham et al. 2003 y las referencias allí escritas).

En todos estos artículos es necesario asumir que los procesos generadores de datos de las series $\{X_t\}$ e $\{Y_t\}$ son lineales, básicamente autoregresivos. En general, estos procedimientos siguen la estructura principal de Haug (1976). El contenido de este Capítulo ha sido recogido en Matilla-García, et al (2009).

En el Capítulo 5 lo que hacemos es comprobar algunas relaciones entre las entropías introducidas en los Capítulos 2 y 3 y la entropía de permutación utilizada en el Capítulo 4.

Queremos remarcar que existe margen para trabajos futuros. En todos los capítulos se presenta una colección de problemas abiertos que pensamos son de interés y que trazan posibles líneas futuras de evolución para los conceptos desarrollados en esta memoria.

# Capítulo 2

# Preliminares

En este capítulo presentamos algunos conceptos y resultados generales de la teoría de sistemas dinámicos y series temporales. La naturaleza de los problemas estudiados en esta memoria nos obliga a incluir una serie de nociones y resultados. Antes de sumergirnos, es conveniente que nos familiaricemos con algunas definiciones y propiedades de los elementos con los que vamos ha trabajar.

Así que este capítulo es sólo una lista de *nociones para consultar*, elaborada de forma totalmente subjetiva, pensando que dichos conceptos hallan podido ser olvidados. De antemano pedimos disculpas a aquellos lectores que no se encuentren afectados por nuestro criterio.

Es importante advertir al lector que de la lista de nociones que introducimos en este capítulo, están enfocadas hacia el uso que tiene en los resultados y técnicas empleadas en nuestro trabajo; así que algunos conceptos pueden que no aparezcan expresados en su forma más clásica ó conocida, por ejemplo, la definición de *entropía topológica de Ader, Konheim y McAndrew* (ver R. L. Adler, et al. 1965).

## 2.1. Sistemas dinámicos

¿Qué entendemos por sistemas dinámicos?. En el presente contexto: serán modelos matemáticos donde la ecuación que define la dinámica del sistema dependerá del tiempo. Estos sistemas se sustentan en dos pilares básicos. El primero es el espacio donde se desarrolla el sistema: conjunto cuyos elementos son los estados posibles del sistema. El segundo pilar será una ley de evolución, que describe cómo cambian los estados del sistema como función del tiempo, una vez que el estado inicial es dado. Generalmente asumiremos que cada vez que el estado de un sistema es conocido para el tiempo $t$, la ley de evolución determina completamente los estados en cualquier tiempo. Esto define un sistema determinístico, opuesto al sistema estocástico donde la evolución esta dada en términos de probabilidad. También supondremos que no podemos influenciar la dinámica, a excepción del estado inicial.

Cuando el comportamiento puede ser sólo predicho en periodos cortos; después de los cuales la incertidumbre del estado inicial, posiblemente combinada con los errores de precisión en los cálculos, hace imposible futuras predicciones, los sistemas se denominaran sistemas dinámicos caóticos y hablaremos entonces de dinámica caótica o simplemente caos.

El siguiente paso será la exposición de una serie de conceptos que nos serán de gran utilidad en el momento de realizar el estudio o análisis de un sistema dinámicos cualquiera.

Recordemos pues algunos conceptos como el de espacio topológico $T_2$ o de Hausdorff. Así pues dado un espacio topológico $(X, \tau)$ diremos que este es de Hausdorff si para cada par de puntos $x_1 \neq x_2$ de $X$, existen entornos $U_1$ y $U_2$ de $x_1, x_2$, respectivamente, que son disjuntos (ver Munkres J.R., 2002).

Análogamente recordaremos que un espacio topológico $(X, \tau)$ se dice que es $T_1$ si los conjuntos con un número finito de puntos son cerrados (ver Munkres J.R., 2002).

Y por ultimo diremos que un espacio topológico $(X, \tau)$ se dice que es normal si se verifica que: Dados $F_1$ y $F_2$ subconjuntos cerrados y disjuntos de $X$, existen conjuntos abiertos y disjuntos $G_1$ y $G_2$ tales que $F_1 \subset G_1$ y $F_2 \subset G_2$ (ver Seymour L., 1970).

Nos disponemos ahora a definir con un poco más de precisión lo que se entiende por un sistema dinamíco.

**Definición 2.1.1.** *Sea $X$ un espacio topológico (normalmente $T_2$ o Hausdorff) y $G = \mathbb{R}$ o $\mathbb{Z}$ (o cualquier otro semigrupo aditivo $G \subset \mathbb{R}$). Un **sistema dinámico** es una terna $(X, G, \Phi)$, con $\Phi : X \times G \to X$ una aplicación continua tal que:*

$$\begin{cases} \Phi(x, 0) = x, & \forall x \in X \\ \Phi(\Phi(x, s), t) = \Phi(x, s + t), & \forall x \in X, \forall s, t \in G \end{cases}$$

*Si notamos $\Phi(x, t)$ por $\Phi_t(x), t \in G, x \in X$ y consideramos $\Phi_t : X \to X$, en términos de estas aplicaciones continuas las propiedades anteriores se reformulan como sigue:*

$$\begin{cases} \Phi_0 = Identidad \\ \Phi_t \circ \Phi_s = \Phi_{s+t}, \quad \forall x \in X, \forall s, t \in G \end{cases}$$

*$X$ es el **espacio de fases o de estados**, $G$ es el conjunto de **tiempos**, y $\Phi$ nos indica la evolución del espacio de fases $X$ a lo largo del tiempo, ya que $\Phi(x, t)$ significa el estado del sistema en el instante $t$, partiendo de $x$ en el instante inicial $t = 0$. Si $G = \mathbb{Z}$ o $G = \mathbb{N} \cup \{0\}$, el sistema se dice **discreto.***

Así pues, dada una función continua $f$ en un espacio topológico $X$, al par $(X, f)$ le llamaremos sistema dinámico discreto, con $G = \mathbb{N} \cup \{0\}$, y

$$\Phi(n, x) = f^n(x),$$

donde $f^n$ representa la iterada $n$-ésima de $f$, de manera que

$$f^n := f \circ f^{n-1} \quad f^1 := f \text{ y} f^0$$

es la identidad en $X$.

La forma mas habitual de representar dichos sistemas dinámicos sería mediante una ecuación en diferencias de la forma:

$$x_{n+1} = f(x_n), \quad n = 0, 1, 2, \ldots \tag{2.1}$$

Sea un sistema dinámico $(X, f)$, llamaremos *trayectoria u órbita* de un punto $x \in X$ al conjunto

$$\{x, f(x), f^2(x), \ldots\} = \bigcup_{n=0}^{\infty} \quad \{f^n(x)\}. \tag{2.2}$$

este conjunto así definido lo representaremos por $Orb(x)$ ó por $Orb_f(x)$. Estudiar la evolución de un sistema dinámico discreto consiste, básicamente, en analizar la evolución asintótica de las órbitas $Orb_f(x) = \{f^n(x)\}_{n \geq 0}$, para $x \in X$.

**Definición 2.1.2.** *Sea $X$ un espacio métrico y sea $f : X \to X$ una función continua. Llamaremos* ***punto fijo*** *de $f$ a todo punto $x \in X$ que verifique*

$$f(x) = x. \tag{2.3}$$

*Al conjunto de todos los puntos fijos de $f$ lo representaremos por* $\mathrm{Fix}(f)$ *(ver Holmgren, Richard 1996).*

**Definición 2.1.3.** *Sea $X$ un espacio métrico y sea $f : X \to X$ una función continua, entonces $x$ es un **punto periódico** de $f$ si*

$$f^n(x) = x \textit{ para algún } n \geq 1. \tag{2.4}$$

*Si $n = 1$ tendremos que $x$ es un punto de los que hemos definido como puntos fijos (ver Holmgren, Richard 1996).*

**Definición 2.1.4.** *Sea $X$ un espacio métrico y sea $f : X \to X$ una transformación continua. Un subconjunto $K \subseteq X$ diremos que es invariante (por $f$) si $f(K) \subseteq K$ y diremos que es estrictamente invariante si $f(K) = K$(ver Holmgren, Richard 1996).*

**Definición 2.1.5.** *Sea $X$ un espacio métrico y sea $f : X \to X$ una transformación continua, $x \in X$. Llamaremos **conjunto** $\omega-$limite de $x$ al conjunto de todos los puntos limite de $\{f^n(x) \mid n \geq 0\}$ y lo representaremos por $\omega(x)$. Esto es*

$$\omega(x) = \{y \in X \mid \textit{ existe una sucesión } n_i \textit{ tal que } \lim_{i\to\infty} f^{n_i}(x) = y\}. \tag{2.5}$$

*Los puntos $x \in \omega(x)$ se denominarán puntos **recurrentes** (ver Alsedà, et al. 2000.)*

**Definición 2.1.6.** *Sea $X$ un espacio métrico y sea $f : X \to X$ continua. Dado un punto $x \in X$ le denominaremos **wandering (errante)** de $f$ si existe un entorno abierto $U$ de $x$ tal que los conjuntos $f^{-n}(U)$, $n \geq 0$ son disjuntos ($U \cap f^n(U) = \emptyset$). El conjunto **non-wandering** de $f$, que representaremos por $\Omega(f)$, estará formado por todos los puntos que no son wandering de $f$, esto es:*

$$\Omega(f) = \{x \in X \mid \textit{ para cualquier } U \textit{ de } x \quad \exists\, n \geq 1 \textit{ con } U \cap f^n(U) \neq \emptyset\}. \tag{2.6}$$

*Resumiendo, diremos que un punto es non-wandering, si cualquier entorno abierto de $x$ contiene al menos dos puntos de alguna órbita de $x$ (ver Alsedà, et al. 2000.)*

Dentro de las funciones con dinámica compleja podemos encontrar las llamadas topológicamente transitivas.

**Definición 2.1.7.** *Sea $X$ un espacio métrico y sea $f : X \to X$ una aplicación continua. Diremos que $f$ es* ***topológicamente transitiva*** *si para cualquier par de subconjuntos abiertos no vacíos $U, V \subset X$, existe un entero positivo $k$ tal que $f^k(U) \cap V \neq \emptyset$ (ver Block, et al. 1992).*

Intuitivamente, una función topológicamente transitiva tiene puntos los cuales, eventualmente, bajo las iteraciones se mueven desde un entorno arbitrariamente pequeño hacia cualquier otro.

**Definición 2.1.8.** *Decimos que un subconjunto $M$ de $X$ es minimal para una función $f$, si este es no vacío, cerrado e invariante y si no existe otro subconjunto de $M$ teniendo las mismas propiedades (ver Block, et al. 1992).*

**Lema 1.** *Si $f : X \to X$ es una aplicación continua, entonces las siguientes propiedades son equivalentes:*

*(i) $f$ es topológicamente transitiva.*

*(ii) Para cualquier conjunto abierto no vacío $W$ en $X$, $\bigcup_{n=1}^{\infty} f^n(W)$ es denso en $X$.*

*(iii) Para cualquier par de conjuntos abiertos no vacíos $U$ y $V$ en $X$, existe un entero positivo $k$ tal que $f^{-k}(U) \cap V \neq \emptyset$.*

*(iv) Para cualquier conjunto abierto no vacío $W$ en $X$, $\bigcup_{n=1}^{\infty} f^{-n}(W)$ es denso en $X$.*

(v) *Cualquier subconjunto cerrado invariante adecuado de $X$ tiene interior vacío*

(ver Block, et al. 1992).

## 2.2. Definiciones de entropía topológica

Antes de introducir las distintas definiciones de entropía topológica, vamos a describir unas serie de pasos que nos conducirán a la introducción de la definición de entropía, sobre un espacio métrico cualquiera, para después de forma cronológica continuar introduciendo las distintas definiciones de entropía topológica que van surgiendo, hasta llegar a la introducida por Cánovas-Rodriguez que es la que origina el siguiente capítulo de esta memoria.

A lo largo de esta sección nos referiremos con asiduidad al concepto función o transformación que conserva la medida, por lo que creemos razonable definir dicho concepto.

Sea $(X, \mathfrak{B}, P)$ un espacio probabilístico, donde $X$ un espacio métrico, por $\mathfrak{B}$ representaremos la $\sigma$-algebra de los subconjuntos de Borel de $X$ y por $P$ una medida de probabilidad en $\mathfrak{B}$.

**Definición 2.2.1.** *Supongamos que $(X_1, \mathfrak{B}_1, P_1), (X_2, \mathfrak{B}_2, P_2)$ son dos espacios probabilísticos. Diremos entonces que:*

1. *Una aplicación $f : X_1 \to X_2$ es medible si $f^{-1}(\mathfrak{B}_2) \subset \mathfrak{B}_1$ (i.e. $B_2 \in \mathfrak{B}_2 \Rightarrow f^{-1}(B_2) \in \mathfrak{B}_1$).*

2. *Una aplicación $f : X_1 \to X_2$ diremos que conserva la medida si $f$ es medible y $P_1(f^{-1}(B_2)) = P_2(B_2)$ para todo $B_2 \in \mathfrak{B}_2$.*

*(ver P. Walters, 1982)*

Estamos ya pues en condiciones de introducir el concepto de función ergódica.

**Definición 2.2.2.** *Sea $(X, \mathfrak{B}, P)$ un espacio probabilístico, y sea $f : X \to X$, una aplicación que conserva la medida. Diremos que $f$ es ergódica si el único elemento $B$ de $\mathfrak{B}$ verificando $f^{-1}(B) = B$ satisface $P(B) = 0$ o $P(B) = 1$ (ver P. Walters, 1982).*

Hay varias formas de presentar las condiciones de ergoicidad aquí hemos expuesto una de ellas.

**Teorema 2.2.3.** *Si $f : X \to X$ es una aplicación que conserva la medida del espacio de probabilidad $(X, \mathfrak{B}, P)$, entonces los siguientes resultados son equivalentes:*

1. *$f$ es ergódica.*
2. *Para cualquier $A \in \mathfrak{B}$ con $P(A) > 0$ tenemos que $P(\bigcup_{n=1}^{\infty} f^{-n}A) = 1$.*
3. *Para cualquier $A, B \in \mathfrak{B}$ con $P(A) > 0$, $P(B) > 0$ existe un $n > 0$ con $P(f^{-n}A \cap B) > 0$.*

El primer gran resultado en la teoría de ergoicidad fue debido a **Birkhoff** y lo expuso en el siguiente teorema.

**Teorema 2.2.4.** *Teorema de ergoicidad de Birkoff*

*Supongamos que $f : (X, \mathfrak{B}, P) \to (X, \mathfrak{B}, P)$ es una aplicación, de importancia relevante a la ahora de introducir el concepto de entropía, que conserva la medida de $(X, \mathfrak{B}, P)$ y sea $g$ una función medible Lebesgue en el espacio $(X, \mathfrak{B}, P)$. Entonces:*

$$\frac{1}{n}\sum_{i=0}^{n-1} g(f^i(x)) \tag{2.7}$$

*converge hacia una función $g^*$ medible Lebesgue en $(X, \mathfrak{B}, P)$. También se verifica que: $g^* \circ f = g^*$ y si $m(X) < \infty$, entonces $\int g^* dm = \int g dm$.*

(ver Walters, P., 1982). Donde $m : \mathfrak{B} \to \mathbb{R} \bigcup \{\infty\}$ es una aplicación tal que $m(\emptyset) = 0$ y $m(\bigcup\limits_{n=1}^{\infty})B_n) = \sum\limits_{n=1}^{\infty} m(B_n)$ siendo $\{B_n\}$ una sucesión de elementos de $\mathfrak{B}$ disjuntos dos a dos.

El siguiente paso será introducir el concepto de partición de un espacio probabilístico $(X, \mathfrak{B}, P)$.

**Definición 2.2.5.** *Llamaremos partición de $(X, \mathfrak{B}, P)$ a una colección disjunta de elementos de $\mathfrak{B}$ cuya unión es $X$ y que será representada por la expresión*

$$\xi = \{A_1, \ldots, A_k\}$$

*(ver P. Walters, 1982).*

Sea $\xi$ partición de $(X, \mathfrak{B}, P)$ que representa una lista de los posibles resultados de un experimento, donde la probabilidad de los resultados $A_i$ es $P(A_i)$. Asociaremos a este experimento el número $H(\xi)$ que medirá la cantidad de incertidumbre sobre el resultado del experimento representado por $\xi$.

**Definición 2.2.6.** *Sea $\xi = \{A_1, \ldots, A_k\}$ una partición finita de $X$. Llamaremos entropía métrica de $\xi$ al número*

$$H(\xi) = -\sum_{i=1}^{k} P(A_i) \log P(A_i) \tag{2.8}$$

*(ver P. Walters, 1982).*

De la anterior definición deduciremos los siguientes resultados:

1. Si $\xi = \{X, \emptyset\}$, entonces $H(\xi) = 0$. Aquí $\xi$ representa los resultados de un cierto experimento por lo que no hay incertidumbre sobre el resultado.

2. Si $\xi = \{A_1, \ldots, A_k\}$ donde $P(A_i) = \frac{1}{k}$, entonces

$$H(\xi) = -\sum_{i=1}^{k} \frac{1}{k} \log \frac{1}{k}. \tag{2.9}$$

3. $H(\xi) \geq 0$.

4. Si $f : X \to X$ es una aplicación que conserva la medida, entonces $H(f^{-1}\xi) = H(\xi)$.

Estamos pues en condiciones de dar el ultimo paso, que nos llevara hasta la definición de entropía métrica de una función continua.

**Definición 2.2.7.** *Supongamos que $f : X \to X$ es una aplicación que conserva la medida del espacio probabilístico $(X, \mathfrak{B}, P)$. Si $\xi$ representa una partición finita de $X$, entonces llamaremos*

$$h(f, \xi) = \lim_{n\to\infty} \frac{1}{n} H(\bigvee_{i=0}^{n-1} f^{-i}(\xi)), \tag{2.10}$$

*y la denominaremos **entropía de la función f** con respecto a $\xi$ (ver P. Walters, 1982).*

Obsérvese:

1. Que $h(f, \xi) \geq 0$.
2. Que $\bigvee_{i=0}^{n-1} f^{-i}(\xi)$ son todos los conjuntos de la forma $\bigcap_{i=0}^{n-1} f^{-i}(A_{ji})$.
3. Si nosotros suponemos que la aplicación $f$ representa el paso de un día de tiempo, entonces $H(\bigvee_{i=0}^{n-1} f^{-i}(\xi))$ representará el experimento combinado de realizar el experimento original, representado por $\xi$, en $n$ dias consecutivos.

4. Que $h(f, \xi)$ es el promedio de información por día que uno obtiene al realizar el experimento original diariamente.

**Definición 2.2.8.** *Si $f : X \to X$ es una aplicación que conserva la medida del espacio probabilístico $(X, \mathfrak{B}, P)$, entonces a*

$$h_P(f) = \sup_{\xi}\{h(f, \xi)\} \tag{2.11}$$

*donde el supremo esta considerado sobre todas las particiones finitas $\xi$ de $X$, la denominaremos* ***entropía métrica de f*** *y dicha entropía dependerá de la probabilidad $P$ (ver P. Walters 1982).*

Si, como hicimos en la definición anterior, tomamos $f$ como una aplicación que representa el paso de un día de tiempo, entonces $h_P(f)$ es el promedio máximo de información por día obtenido mediante la realización del mismo experimento diariamente.

Una vez introducido el concepto y alguna de las propiedades de la entropía de una función, pasamos ha definir la entropía topológica de una función que es en realidad uno de los ejes de esta memoria.

La definición original fue introducida por *Adler, Konheim y McAndrew* (ver R. L. Adler, et al. 1965). Su idea de asignar un número a un recubrimiento por abiertos para medir su tamaño fue inspirada por *Kolmogorov.* Para espacios métricos una definición diferente fue introducida por *Bowen.*

**Definición 2.2.9.** *Sea $\mathcal{X}$ un conjunto compacto, $f : \mathcal{X} \to \mathcal{X}$ una aplicación continua y sea $\mathcal{U}$ un cubrimiento por abiertos de $\mathcal{X}$. Por $\mathcal{N}(\mathcal{U})$ representaremos el cardinal mas pequeño de un subrecubrimiento de $\mathcal{U}$ (esto es, una subfamilia de $\mathcal{U}$ cuya unión es estrictamente igual a $X$). Por compacidad $\mathcal{N}(\mathcal{U})$ es siempre finito. Si $\mathcal{U}$ y $\mathcal{V}$ son recubrimientos por abiertos de $X$ entonces a*

$$\mathcal{U} \bigvee \mathcal{V} = \{\mathrm{U} \bigcap \mathrm{V} : \mathrm{U} \in \mathcal{U}, \mathrm{V} \in \mathcal{V}\}$$

*le denominaremos su* refinamiento común.*(ver R. L. Adler, et al. 1965.)*

**Definición 2.2.10.** *Entropía topológica de Alder,Konheim y McAndrew.*

*Sea $\mathcal{U}$ un recubrimiento por abiertos de X*

$$\mathcal{U}^n = \mathcal{U} \bigvee f^{-1}\mathcal{U} \bigvee \ldots \bigvee f^{-n+1}\mathcal{U},$$

*donde $f^{-k}(\mathcal{U}) = \{f^{-k}(\mathrm{U}) : \mathrm{U} \in \mathcal{U}\}$. Utilizando argumentos de subaditividad, se prueba que*

$$h(f, U) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{N}(U^n) \tag{2.12}$$

*existe para cualquier recubrimiento abierto $\mathcal{U}$ (y es igual a $\inf_{n\in N}\{\frac{1}{n}\log\mathcal{N}(U^n)\}$). La **entropía topológica** de $(X, f)$ se define pues como*

$$h(f) = \sup h(f, U) \tag{2.13}$$

*para todos los cubrimientos por abiertos $\mathcal{U}$ de $\mathcal{X}$ (ver R. L. Adler, et al. 1965.)*

La siguiente definición fue introducida por *Rufus Bowen* (ver R. Bowen, 1973) la novedad que presenta esta definición es que: la función $f$ es uniformemente continua y el espacio $X$ es métrico pero no necesariamente compacto

**Definición 2.2.11.** *Sea $(X, d)$ un espacio métrico, $f : X \to X$ uniformemente continua, sea $n \in \mathbb{N}$, $\varepsilon > 0$ y sea $K$ un subconjunto compacto de $X$. Un subconjunto $F$ de $X$ se denomina $(n, \varepsilon)$-span $K$ con respecto a $f$ si para todo $x \in K$ existe $y \in F$ con $d_n(x, y) \leq \varepsilon$ tal que:*

$$K \subset \bigcup_{y\in F} \bigcap_{i=0}^{n-1} f^{-i}(\overline{B}) f^i(y) \quad (\varepsilon > 0) \tag{2.14}$$

*notaremos por $r_n(\varepsilon, K)$ el mínimo cardinal de un conjunto $(n, \varepsilon) - spanning$ de $K$ respecto de $f$ (o $r_n(\varepsilon, K, f)$). Sea*

$$h(f; K) = \lim_{\epsilon\to 0} \limsup_{n\to\infty} \frac{1}{n} \log r_n(\varepsilon, K, f).$$

*Llamaremos* <u>*entropía topológica de Bowen*</u> *de f y la representaremos por h(f) a la siguiente expresión:*

$$h(f) = \sup_K \{ \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log r_n(\varepsilon, K, f) \ \ con\ K \subset X\ compacto \} \tag{2.15}$$

*estando el supremo tomado sobre el conjunto de todos los subconjuntos compactos de X. En ocasiones denotaremos dicha entropía de las forma $h_d(f)$ para de esta manera hacer constar la dependencia de la distancia d.*

A continuación introduciremos una definición equivalente pero "dual", en el sentido de que los conjuntos spanning serán sustituidos por sus duales, los conjuntos separados.

**Definición 2.2.12.** <u>*Entropía topológica de Bowen.*</u> *Sea $(X, d)$ un espacio métrico, $f : X \to X$ uniformemente continua y sean $K \subset X$ un conjunto compacto, $\epsilon > 0$ y $n \in \mathbb{N}$. Un subconjunto $E \subset K$ se dice que es un conjunto $(n, \epsilon, K, f)$– separado si para todo $x, y \in E$, $x \neq y$, existe $i \in \{0, 1, ..., n-1\}$ tal que $d(f^i(x), f^i(y)) > \epsilon$. Representaremos por $s_n(\epsilon, K, f)$ el máximo cardinal de un conjunto $(n, \epsilon, K, f)$-separado de $K$. Llamaremos* entropía topológica *de $f$ y la notaremos por $h(f)$ a la siguiente expresión:*

$$h(f) := \sup_K \left\{ \lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{1}{n} \log s_n(\epsilon, K, f) \ \ con\ K \subset X\ \ compacto \right\}. \tag{2.16}$$

*(ver P. Walters, 1982.)*

Nuestro siguiente paso será mostrar como la entropía métrica introducida por *Alder, Konheim y McAndrew* (ver R. L. Adler, et al. 1965) que representaremos de momento por $h^*(f)$ y la entropía introducida por *Bowen* (ver P. Walters, 1982) y que representaremos por $h(f)$ coinciden, cuando $X$ es un espacio métrico compacto.

Empezaremos considerando el espacio métrico $(X, d)$ donde por $diam(\alpha)$ representaremos el diámetro de un recubrimiento $\alpha$ y lo definiremos como

$diam(\alpha) = \sup_{A \in \alpha} diam(A)$ siendo, $diam(A) = \sup_{x,y \in A} d(x,y)$, el diámetro del conjunto A.

**Teorema 2.2.13.** *Sea $(X,d)$ un espacio métrico compacto. Si $\{\alpha_n\}_1^\infty$ es una sucesión de recubrimientos de $X$ con $diam(\alpha_n) \to 0$, entonces si $h^*(f) < \infty$ el $\lim_{n\to\infty} h^*(f,\alpha_n)$ existe y además es igual a $h^*(f)$, y si $h^*(f) = \infty$ entonces $\lim_{n\to\infty} h^*(f,\alpha_n) = \infty$.*

El siguiente resultado establece las relaciones básicas entre ambas definiciones de entropía topológica, es decir, la introducida por medio de conjuntos spanning y la introducida por medio de recubrimientos.

**Teorema 2.2.14.** *Sea $f : X \to X$ una aplicación continua del espacio métrico compacto $(X,d)$.*

1. *Si $\alpha$ es un recubrimiento por abiertos de $X$ con numero de Lebesgue $\delta$ entonces*
$$\mathcal{N}(\bigvee_{i=0}^{n-1} f^{-i}\alpha) \leq r_n(\delta/2, X) \leq s_n(\delta/2, X). \tag{2.17}$$

2. *Si $\varepsilon > 0$ y $\gamma$ es un recubrimiento por abiertos con $diam(\gamma) \leq \varepsilon$ entonces*
$$r_n(\varepsilon, X) \leq s_n(\varepsilon, X) \leq \mathcal{N}(\bigvee_{i=0}^{n-1} f^{-i}\gamma). \tag{2.18}$$

**Colorario 1.** *Sea $f : X \to X$ una aplicación continua del espacio métrico compacto $(X,d)$. Sea $\varepsilon > 0$. Sea $\alpha_\varepsilon$ el recubrimientos de $X$ por bolas abiertas de radio $2\varepsilon$ y sea $\gamma_\varepsilon$ un recubrimiento de $X$ por bolas abiertas de radio $\varepsilon/2$. Entonces*
$$\mathcal{N}(\bigvee_{i=0}^{n-1} f^{-i}\alpha_\varepsilon) \leq r_n(\varepsilon, X) \leq s_n(\varepsilon, X) \leq \mathcal{N}(\bigvee_{i=0}^{n-1} f^{-i}\gamma_\varepsilon). \tag{2.19}$$

Consecuencia inmediata de los resultados expuestos con anterioridad será el siguiente teorema.

**Teorema 2.2.15.** *Si $f : X \to X$ una aplicación continua del espacio métrico compacto $(X, d)$ entonces $h(f) = h^*(f)$, es decir, las dos definiciones de entropía topológica coinciden.*

Nuestro siguiente paso, es enunciar las propiedades que se cumplen en la definición clásica de entropía topológica introducida para conjuntos compactos, y que vamos a resumir en el siguiente resultado (ver P. Walters 1982, L. Alsedá, et al. 1993, R. L. Adler, et al. 1965, R. Bowen 1971 y R. A. Dana, et al. 1986).

**Teorema 2.2.16.** *Sean $X$ y $Y$ dos conjuntos compactos (métricos) y sea $f : X \to X$ y $g : Y \to Y$ dos aplicaciones continuas. Entonces las siguientes propiedades son ciertas:*

*(a) Sea $\varphi : X \to Y$ continua tal que $g \circ \varphi = \varphi \circ f$. Entonces:*

*(a1) Si la función $\varphi$ es inyectiva, entonces*

$$h(f) \leq h(g).$$

*(a2) Si la función $\varphi$ es sobreyectiva, entonces*

$$h(f) \geq h(g).$$

*(a3) Si la función $\varphi$ es biyectiva, entonces*

$$h(f) = h(g).$$

*(b) Supongamos que $X = \bigcup_{i=1}^{n} X_i$, donde $X_i$ son compactos e invariantes por $f$. Entonces*

$$h(f) = \max \{h(f|_{X_i}) : i = 1, 2, ...., n\}.$$

*(c) Para un entero* $n \geq 0$ *se cumple*

$$h(f^n) = nh(f).$$

*(d) Sea* $f \times g : X \times Y \to X \times Y$ *definida por* $(f \times g)(x, y) = (f(x), g(y))$ *para todo* $(x, y) \in X \times Y$*. Entonces*

$$h(f \times g) = h(f) + h(g).$$

*(e) Si* $f$ *es un homeomorfismo, entonces*

$$h(f) = h(f^{-1}).$$

*(f) Sea* $\varphi : X \to Y$ *una aplicación sobreyectiva continua tal que* $\varphi \circ f = g \circ \varphi$*. Entonces*

$$\begin{aligned} &\text{máx}\{h(g), \sup\{h(f, \varphi^{-1}(y)) : y \in Y\}\} \\ &\leq h(f) \\ &\leq h(g) + \sup\{h(f, \varphi^{-1}(y)) : y \in Y\}. \end{aligned} \tag{2.20}$$

*(g) Si* $f : X \to Y$ *y* $g : Y \to X$ *son continuas, entonces*

$$h(f \circ g) = h(g \circ f).$$

*(h) Sean* $f : X \to Y$*,* $g : Y \to X$ *continuas y sea* $F : X \times Y \to X \times Y$ *definida por* $F(x, y) = (g(y), f(x))$ *para todo* $(x, y) \in X \times Y$*. Entonces*

$$h(F) = h(f \circ g) = h(g \circ f).$$

*(i) Si* $X_\infty = \bigcap\limits_{n \geq 0} f^n(X)$ *entonces*

$$h(f) = h(f|_{X_\infty}).$$

*(j)* $h(f) = h(f|_{\Omega(f)})$ *donde* $x \in \Omega(f)$ *si para todo entorno* $U$ *de* $x$ *existe* $n > 0$ *tal que*

$$f^n(U) \cap U \neq \emptyset.$$

Para el apartado siguiente, supondremos que $(X, d)$ es un espacio métrico compacto y $f : X \to X$ continua.

## 2.3. Relación entre entropía métrica y topológica

Sea $\mathfrak{B}(X)$ la $\sigma$-algebra de Borel de los subconjuntos de $X$ y representemos por $\mathcal{M}(X, f)$ el espacio de todas las medidas de probabilidad, en el espacio medible $(X, \mathfrak{B}(X))$, las cuales son conservadas por f. Se sabe que $\mathcal{M}(X, f)$ es un conjunto convexo no vacío el cual es compacto con la topología discreta. Si $\mu \in \mathcal{M}(X, f)$ entonces f es una aplicación que conserva la medida del espacio $(X, \mathfrak{B}(X), \mu)$ y por lo tanto, tiene una *entropía métrica* también conocida como *entropía de Kolmogorov* que representaremos por $h_\mu$(f).

**Definición 2.3.1.** *La función* $h : \mathcal{M}(X, f) \to [0, \infty)$ *que a todo* $\mu \to h_\mu(f)$ *la denominaremos entropía de la aplicación* $f$*.(ver P. Walters, 1982).*

El siguiente paso será ver la relación existente entre la entropía topológica y la entropía métrica de Kolmogorov, para lo cual introduciremos el principio variacional.

Sea $f : X \to X$ una aplicación continua de $X$ espacio métrico compacto, entonces $h(f) = \sup\{h_\mu(f)$ tal que $\mu \in \mathcal{M}(X, f)\}$. La desigualdad

$$\sup\{h_\mu(f) \text{ tal que } \mu \in \mathcal{M}(X, f)\} \leq h(f)$$

fue probada por *Goodwyn* en 1968. En 1970 fue *Dinaburg* quien prueba la igualdad para el caso en que $X$ tiene recubrimientos de dimensión finita y fue también en 1970 cuando *Goodman* lo prueba para el caso general.

**Teorema 2.3.2.** *Sea $f : X \to X$ una aplicación continua de un espacio métrico y compacto $X$ en si mismo. Entonces*

$$h(f) = \sup\{h_\mu(f) \colon \mu \in \mathcal{M}(X, f)\}. \tag{2.21}$$

## 2.4. Estudio de funciones del intervalo

Esta sección la dedicaremos al estudio de propiedades de la entropía de funciones continuas en el intervalo $I = [0, 1]$.

**Definición 2.4.1.** *Funciones monótonas a trozos*

*Sea $f : I \to I$ continua, diremos que es* ***monótona a trozos****, si existe una partición finita de $I$ en intervalos, tal que, en cada uno de los elementos de la partición, $f$ es monótona. Diremos que $f$ es $\mathtt{P}$-monótona para algún conjunto finito $\mathtt{P} \subset I$, donde $\mathtt{P}$ es el conjunto de extremos de los intervalos de la partición. Al conjunto de todas las funciones continuas monótonas a trozos $f : I \to I$ lo representaremos por $\mathcal{C}$ (ver Alsedà, et al. 2000).*

De la anterior definición podemos deducir que:

1. Si $f \in \mathcal{C}$ entonces $h(f) = \sup_{\mathcal{A}} h(f, \mathcal{A})$, también se cumple si el supremo es tomado sobre cualquier recubrimiento finito $\mathcal{A}$ de $I$ por intervalos.

2. Si $f \in \mathcal{C}$, diremos que un recubrimiento $\mathcal{A}$ es *f-mono* si este es finito, esta formado por intervalos y $f$ es monótona en cada uno de sus elementos.

3. Si $f \in \mathcal{C}$ y $\mathcal{A}$ es un recubrimiento *f-mono* de $I$ entonces $h(f, \mathcal{A}) = h(f)$.

**Teorema 2.4.2.** *Si $f \in \mathcal{C}$ entonces*

$$\lim_{n\to\infty} \frac{1}{n} \log c_n = h(f) \tag{2.22}$$

*y $\frac{1}{n}\log c_n \geq h(f)$ para algún $n$, donde $c_n$ el mínimo cardinal de todos los $f^n -$ mono recubrimientos de I. Se puede comprobar que $c_n$ es el mínimo de trozos monótonos de $f^n$.*

En la teoría de sistemas dinámicos, uno de los mas famosos es la *herradura* de Smale. Su construcción es de la forma siguiente. Tomemos un intervalo cerrado I y dos subintervalos cerrados disjuntos $I_1, I_2$. Entonces aplicamos los rectángulos $I \times I_1$ e $I \times I_2$ mediante una aplicación afín en $\mathbb{R}^2$ de tal manera que los lados de los rectángulos y sus imágenes son paralelos, las imágenes de los dos rectángulos son disjuntas y atraviesan el cuadrado $I \times I$ de abajo hacia arriba (o de arriba hacia abajo). Esto puede hacerse de tal manera que podamos extender esta aplicación f a un difeomorfismo de $I \times I$ en su imagen y obtener $f(I \times I)$ con forma de herradura (ver Alsedà, et al. 2000.)

Haciendo abuso del lenguaje, llamaremos habitualmente *herradura* a cualquier función que se parece a la herradura de Smale y tenga propiedades similares.

En un sistema unidimensional podemos detectar objetos similares. Si se aplica un intervalo en sí mismo varias veces, obtendremos una *herradura* multiple. Dicho de otra forma, si $f \in \mathcal{C}$ y $s \geq 2$, entonces una *s-herradura* de $f$ es un intervalo $J \subset I$ y una partición $\mathcal{D}$ de $J$ en $s$ subintervalos tal que la clausura de cada elemento de $\mathcal{D}$ $f-$recubre $J$. Representaremos por $(J, \mathcal{D})$ a este tipo de herraduras.

A continuación estableceremos una conexión entre la entropía de $f$ y la existencia de herraduras de $f$ y sus iteradas.

**Lema 2.** *Si $(J, \mathcal{D})$ es una s-herradura de f entonces existe una partición $\mathcal{D}_n$*

*de $J$ tal que $(J, \mathcal{D}_n)$ es una $s^n$-herradura de $f^n$.*

**Lema 3.** *Si $f$ tiene una $s$-herradura entonces*

$$h(f) \geq \log s. \tag{2.23}$$

**Teorema 2.4.3.** *Si $f \in \mathcal{C}$ tiene entropía positiva entonces existen sucesiones $\{k_n\}_{n=1}^{\infty}$ y $\{s_n\}_{n=1}^{\infty}$ de enteros positivos tales que, para cada $n$ la función $f^{k_n}$ tiene una $s_n$-herradura y además*

$$\lim_{n\to\infty} \frac{1}{k_n} \log s_n = h(f). \tag{2.24}$$

## 2.5. Compactificación

Como es sabido, la recta real $\mathbb{R}$ junto con la topología usual $\tau_u$ ,es decir, el espacio topológico $(\mathbb{R}, \tau_u)$ es homeomorfo al conjunto abierto $(0,1)$. Bastará con considerar la aplicación $\varphi : \mathbb{R} \longrightarrow (0,1)$ definida por:

$$\varphi(x) = \tfrac{x}{1+|x|}.$$

Pero si añadimos al intervalo $(0,1)$ los puntos 0 y 1 convertimos el abierto en un cerrado $[0,1]$ que es compacto y tiene a $(0,1)$ como subespacio. Esto sugiere un procedimiento para obtener a partir de $(\mathbb{R}, \tau_u)$ un espacio compacto que contenga a $(\mathbb{R}, \tau_u)$ como subespacio. Este proceso se conoce como *compactificación* por dos puntos. Consiste en añadir dos puntos $p$ y $q$ a $\mathbb{R}$, y extender la aplicación $\varphi$ a una nueva aplicación $\overline{\varphi} : \overline{\mathbb{R}} = \mathbb{R} \bigcup \{p\} \bigcup \{q\} \longrightarrow [0,1]$ definiendo $\overline{\varphi}(p) = 0, \overline{\varphi}(q) = 1, \overline{\varphi} \mid \mathbb{R} = \varphi$, y dotar de una nueva topología $\overline{\tau}$ al nuevo conjunto $\overline{\mathbb{R}}$ de modo que $\overline{\varphi}$ sea un homeomorfismo. Con esta topología el espacio $(\overline{\mathbb{R}}, \overline{\tau})$ es compacto por ser homeomorfo al compacto $[0, 1]$. Si como puntos $p$ y $q$ tomamos $+\infty$ y $-\infty$ entonces $\overline{\mathbb{R}}$ se le denomina recta real ampliada (ver Arregui, 1988).

**Definición 2.5.1.** *Sea* $(X, \tau)$ *un espacio topológico y de Hausdorff. Llamaremos* compactificación *de* $X$ *a un par* $(Y, \varphi)$ *donde* $Y$ *es un espacio topológico compacto y de Hausdorff y* $\varphi : X \to Y$ *es un homeomorfismo sobre la imagen* $\varphi(X)$ *tal que, esta es, un abierto denso de* $Y$ *( ver Seymour L., 1970 ).*

Si $X$ es localmente compacto vamos a construir una compactificación de $X$ llamada de (Alexandroff). Definimos un conjunto $X^* := X \cup \{\omega\}$ siendo $\omega$ un elemento que no está en $X$. Consideremos la siguiente familia del conjunto $\mathcal{P}(X^*)$:

$$\tau^* := \tau \cup \{(X \backslash \mathrm{K})^* \text{ con } \mathrm{K} \subset X \textit{ compacto}\},$$

donde si $A \subset X$, $A^* := A \cup \{\omega\}$. Donde $\tau^*$ es la topología cuyos abiertos son los abiertos de $\tau$ de la forma $A^* := A \cup \{\omega\}$ tales que $X^* - A^*$ es un cerrado y compacto de $(X, \tau)$. El homeomorfismo $\varphi$ sera la inclusión $\varphi : \mathbb{X} \to \mathbb{X}^*$. Entonces se verifican las siguientes **propiedades**:

1. $\tau^*$ es una topología de forma que el par $(X^*, \varphi)$ es una compactificación (llamada **compactificación de Alexandroff**). Se cumple que $(X^*, \varphi^*)$ es de Hausdorff y compacto si y solo si $(X, \varphi)$ es de Hausdorff localmente compacto.

2. Si $X$ e $Y$ son dos espacios topológicos no compactos, Hausdorff, localmente compactos y homeomorfos, entonces sus compactificaciones de Alexandroff son homeomorfas. Es decir, que las compactificaciones de Alexandroff están bien definidas salvo homeomorfismos.

3. Sea $Y$ un espacio topológico compacto y Hausdorff y sea $y \in Y$ no aislado. Denotaremos $X := Y \backslash \{y\}$. Entonces, $Y$ es homeomorfo a la compactificación de Alexandroff de $X$.

4. Si $f : X \to X$ es un homeomorfismo, entonces existe una única extensión $f^* : X^* \to X^*$ tal que $f^*(w) = w$, de manera que $f^*$ es un homeomorfismo.

En general, el resultado 4 no es cierto si $f : X \to X$ es simplemente continua. Esta propiedad se logra gracias a la compactificación de **Stone-Čech**, que pasamos a introducir. La compactificación de Alexandroff de un espacio topológico $X$ es considerarda, en algún sentido, como la compactificación minimal de $X$ ya que dado $X$ le añadíamos un punto que no era de $X$ y partiendo de estas situación se realizaba todo el proceso de obtención del espacio compacto y de Hausdorff. En este mismo sentido la de Stone-Čech puede ser considerada como la compactificación maximal de $X$.

La compactificación de Stone-Čech es pues una técnica para extender una aplicación de un espacio topológico $X$ en un espacio compacto de Hausdorff $\beta(X)$. La compactificación de Stone-Čech $\beta(X)$ de un espacio topológico $X$ es, por tanto, el mayor espacio compacto de Hausdorff "generado"por $X$, en el sentido de que cualquier aplicación de $X$ en un espacio compacto de Hausdorff factoriza a través de $\beta(X)$(de una manera única).

**Definición 2.5.2.** *Sea $X$ es un espacio completamente regular y separable, entonces la función de $X$ en su imagen en $\beta(X)$ es una homeomorfismo, así de esta manera $X$ puede ser visto como un subespacio denso de $\beta(X)$. En general para cualquier espacio topológico $X$, la función de $X$ en $\beta(X)$ no tiene por qué ser inyectiva ( ver Munkres, J. R., 1975 ).*

Como consecuencia inmediata de la anterior definición tendremos que:

1. Sea $X$ un espacio topológico cualquiera y sea $\beta(X)$ un espacio compacto de Hausdorff, consideremos una aplicación continua $h$ de $X$ en $\beta(X)$. Entonces se verifica la siguiente propiedad universal: para cualquier aplicación continua $f : X \to K$, donde $K$ es un espacio compacto de Hausdorff, existe una aplicación continua $\beta(f) : \beta(X) \to K$ tal que $\beta(f) \circ h = f$ y $h(x) = x$ para todo $x \in X$. Como es habitual en las propiedades universales, esta propiedad universal, junto con el hecho de que $\beta(X)$ es un espacio compacto de Hausdorff que contiene $X$, caracteriza a $\beta(X)$ salvo homeomorfismo.

2. La compactificación de Stone-Čech se caracterizará por el hecho de que, cualquier aplicación continua $f : X \to C$ siendo $C$ un espacio compacto de Hausdorff, se extiende de forma única a una aplicación continua $g : \beta(X) \to C$ de forma que $f(x) = g(x)$ en $X$ .

## 2.6. Series temporales

Es frecuente, no solo en el ámbito económico y en general en las ciencias sociales, que las observaciones de los caracteres de una población se realicen ligadas al tiempo o a fechas en instantes determinados del tiempo. Así, por ejemplo, podemos estar interesados en estudiar el comportamiento y la evolución temporal de uno de los caracteres de una empresa susceptible de ser observado. En este caso esa observación se realizará de forma repetida durante una serie de momentos del tiempo. Esa observación repetida en el tiempo da lugar a una serie temporal.

La teoría de las series temporales estudia, esencialmente, modelos matemáticos para describir magnitudes que evolucionan a lo largo de tiempo (medido de manera discreta) y para predecir su evolución. Para alcanzar este doble objetivo se utiliza una metodología bastante consolidada, según la cual se admite que la serie temporal es una función del tiempo:

$$X_t = f(t).$$

Desde este punto de vista, cualquier serie temporal se supone que es el resultado de cuatro componentes: tendencia, variaciones estacionales, variaciones cíclicas y variaciones residuales o accidentales. Desde la aparición del libro de Box y Jenkins en los setenta, (ver Box, et al. 1970) su aplicabilidad real a través de programas informáticos, ha progresado considerablemente. Una definición más formal de serie temporal sería:

**Definición 2.6.1.** *Una serie temporal se puede definir como una sucesión de observaciones cuantitativas cronológicamente ordenadas, o como una distribución bidimensional en la cual una de las componentes es el tiempo y la otra la propia variable en estudio. La representaremos como* $\{X_t\}_{t\in I}$ *donde* $X_t$ *es la observación correspondiente al instante* $t$ *de la variable que estamos estudiando e* $I = \{1, 2, \ldots, t, \ldots\}$ *(ver Box, et al. 1970).*

Con el fin de estudiar las series temporales desde un punto de vista no descriptivo, sino inferencial, y poder estudiar propiedades de independencia de la series y entre series temporales, necesitamos construir modelos matemáticos que nos permitan detectar estas características. Para ello una herramienta esencial es la variable aleatoria.

**Definición 2.6.2.** *Sea* $(\Omega, \mathcal{S}, P)$ *un espacio probabilístico, donde* $\Omega$ *es un espacio muestral,* $\mathcal{S}$ *una* $\sigma$*-álgebra y* $P$ *una función de probabilidad. Diremos que una función*

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

*es una variable aleatoria si la antiimagen mediante $X$ de cualquier intervalo de la forma $(-\infty, a]$ es un suceso en la $\sigma$-álgebra $\mathcal{S}$ para todo $a \in \mathbb{R}$, es decir,*

$$X^{-1}((-\infty, a]) \in \mathcal{S} \quad \textit{para todo } a \in \mathbb{R}$$

*( ver Durá J.M., et al. 1988. )*

A continuación pasaremos a distinguir entre dos tipos de variables aleatorias, las variables aleatorias discretas y las variables aleatorias continuas. Ésto nos facilitara la lectura del Capítulo 4 de esta memoria.

En general una variable aleatoria se denomina *discreta* cuando entre dos valores de la variable es sólo posible encontrar una cantidad finita de valores distintos intermedios. Por el contrario, diremos que una variable es *continua* si entre dos valores cualesquiera que toma dicha variable es siempre posible encontrar infinitos valores intermedios distintos.

Como un primer ejemplo de variable aleatoria discreta consideraremos la denominada de **Bernoulli**.

Supongamos que realizamos un experimento aleatorio donde el resultado del mismo sólo puede ser éxito o fracaso. Este tipo de experimentos tiene asociada la variable aleatoria de Bernoulli. Esta variable aleatoria $X$ está definida en un espacio muestral $\Omega = \{A, A^c\}$ con dos únicos sucesos que denominaremos $A =$"éxito" y $A^c =$"fracaso" y que toma los valores:

$$\begin{aligned} X: \quad & \Omega \longrightarrow \mathbb{R} \\ & A \mapsto X(A) = 1 \\ & A^c \mapsto X(A^c) = 0. \end{aligned}$$

Su función puntual de probabilidad viene dada por

$$P(X = 1) = p \quad \text{y} \quad P(X = 0) = 1 - p,$$

donde $p$ es la probabilidad de obtener éxito en el experimento aleatorio. En este caso denotaremos a la variable aleatoria de Bernoulli por $B(p)$.

A continuación consideremos el siguiente conjunto

$$\{X_1, X_2, \ldots, X_n\}$$

de $n$ variables aleatorias independientes identicamente distribuidas $B(p)$. Entonces definimos la variable aleatoria Binomial de parámetros $n$ y $p$, $B(n, p)$ como

$$B(n,p) = \sum_{i=1}^{n} X_i.$$

Obsérvese que esta variable calcula el número de éxitos en $n$ realizaciones de un experimento de Bernoulli. Su función puntual de probabilidad es

$$P(B(n,p) = r) = \binom{n}{r} p^r (1-p)^{(n-r)}.$$

Para finalizar la presentación de algunas variables aleatorias discretas introduciremos la **Multinomial**.

Sean

$$\{X_1, X_2, \ldots, X_k\}$$

$k$ variables aleatorias binomiales independientes, es decir, $X_1 = B(n, p_1)$, $X_2 = B(n, p_2), \ldots, X_k = B(n, p_k)$. Llamaremos variable aleatoria **Multinomial** a la variable multivariante $Y = (X_1, X_2, \ldots, X_k)$ con función puntual de probabilidad

$$P[X_1 = n_1, X_2 = n_2 \ldots, X_k = n_k] = \frac{n!}{n_1! n_2! \cdots n_k!} p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$$

siendo $n = n_1 + n_2 + \cdots + n_k$.

El siguiente paso que daremos será presentar unos ejemplos de variables aleatorias continuas. Dentro de este tipo de variables una de las que tiene mayor relevancia es la que denominaremos como variable aleatoria **Normal**.

La distribución Normal, también llamada distribución de Gauss o distribución gaussiana, es la distribución de probabilidad que con más frecuencia aparece en estadística y teoría de probabilidades. Esto se debe a dos razones fundamentalmente:

1. Su función de densidad es simétrica y con forma de campana, lo que favorece su aplicación como modelo a gran número de variables estadísticas.

2. Es, además, límite de otras distribuciones y aparece relacionada con multitud de resultados ligados a la teoría de las probabilidades gracias a sus propiedades matemáticas.

Su función de densidad (el equivalente a la función puntual de probabilidad en el caso discreto) viene definida por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

donde $\mu$ representa a la media (valor esperado o esperanza matemática) y $\sigma$ su desviación típica. Esta variable aleatoria se representa por $N(\mu, \sigma)$

Esta función de densidad presenta una serie de peculiaridades:

1. La función $f(x)$ tiene forma de campana de Gauss.

2. Es simétrica respecto a su media.

3. Su media coincide con su mediana.

4. Alcanza el máximo en $x = \mu$.

5. Tiene los puntos de inflexión en $\mu - \sigma$ y $\mu + \sigma$.

6. La distribución de una variable aleatoria normal es única respecto de $\mu$, $\sigma$.

Para finalizar con la presentacion de las variables aleatorias que vamos a utilizar en esta memoria, vamos a definir la variable **Chi-cuadrado**.

Consideremos ahora $\{X_1, X_2, \ldots, X_n\}$ un conjunto de n variables independientes normales $N(0,1)$. Llamaremos variable aleatoria **Chi-cuadrado** con $n$ grados de libertad y la representaremos como $\chi_n^2$ a la variable aleatoria definida como

$$\chi_n^2 = X_1^2 + X_2^2 + \ldots + X_n^2.$$

Su función de distribución, al igual que para la distribución normal está tabulada para los diferentes valores de $n$ (grados de libertad).

Otro concepto de crucial importancia en el análisis de series temporales es el de independencia serial. Por ello pasamos a introducir el concepto de independencia entre variables aleatorias.

**Definición 2.6.3.** *Sean $X$ e $Y$ dos variables aleatorias con funciones de densidad $f_X(x)$ y $f_Y(y)$ respectivamente. Consideremos la variable aleatoria bidimensional $(X,Y)$ con función de densidad conjunta $f(x,y)$. Diremos que las variables $X$ e $Y$ son independientes si*

$$f(x,y) = f_X(x) f_Y(y).$$

*(si la v. a. es discreta hablaríamos de función puntual de probabilidad)( ver Durá J.M., et al. 1988. )*

Uno de los grandes problemas en economía es determinar la dependencia o independencia de una serie temporal. Para ello existen numerosos métodos en la literatura, tanto paramétricos como no paramétricos.

Uno de los problemas más complejos y necesarios de contrastar a la hora de realizar predicciones es el que se aborda en el Capítulo 4 de esta memoria. El problema consiste en, dadas dos series temporales $\{X_t\}_{t\in I}$ e

$\{Y_t\}_{t \in I}$, construir un estadístico con distribución asintótica estandar que nos permita aceptar o rechazar la hipótesis de dependencia entre las series. El test construido para ello es no paramétrico y necesita de la hipótesis de estacionariedad de las series a contrastar.

Desde un punto de vista intuitivo, un proceso estacionario se describe por una secuencia de datos o valores que no presentan ningún cambio sistemático en la media, ni en varianza. Así, se dice, que un proceso es estacionario cuando, en cada uno de los puntos del tiempo, la observación registrada puede ser considerada una variable aleatoria a la que está asociada una función de densidad de probabilidad. Función que está usualmente descrita por sus momentos, valor esperado y varianza.

Una serie temporal $\{X_t\}_{t \in I}$ se dice que es *estacionaria* si su distribución de probabilidad no cambia en el tiempo. Consideremos la probabilidad conjunta de un conjunto de variables aleatorias

$$F(w_{t_1}, w_{t_2}, \cdots w_{t_n}) = P(X_{t_1} \leq w_{t_1}, X_{t_2} \leq w_{t_2}, \cdots X_{t_n} \leq w_{t_n})$$

diremos entonces que es *estacionaria de orden* $n$ si

$$F(w_{t_1}, \cdots w_{t_n}) = F(w_{t_1+k}, \cdots w_{t_{n+k}},)$$

para todo $t_1, \cdots t_n, k$.

Un proceso se dirá *estrictamente estacionario* si es estacionario de orden $n$ para cada $n$.

## 2.7. Estadísticos

La estadística nos proporciona herramientas que formalizan y uniforman los procedimientos para sacar conclusiones siempre que las muestras seleccionadas sean representativas de la población que han sido extraídas. Esta representatividad permite extender los valores que describen a las muestras, es decir, los estadísticos, a la población correspondiente.

A continuación definimos dos estadísticos que utilizaremos con posterioridad en el Capítulo 4 que servirán para contrastar independencia entre series temporales.

El estadístico de Haugh

$$S_M = n \sum_{j=-M}^{M} r^2_{\hat{u}\hat{v}}(j) \tag{2.25}$$

donde

$$r_{\hat{u}\hat{v}}(j) = \sum_{t=j+1}^{n} \hat{u}_t \hat{v}_{t-j} / \left( \sum_{t=1}^{n} \hat{u}_t^2 \sum_{t=1}^{n} \hat{v}_t^2 \right)^{1/2}$$

son los residuos de las correlaciones cruzadas para $0 \leq j \leq n-1$, $r_{\hat{u}\hat{v}}(j) = r_{\hat{u}\hat{v}}(-j)$ para $1-n \leq j < 0$ y $\hat{u}_t, \hat{v}_t$ , $t = 1, ..., n$ son las dos series residuales de longitud $n$, obtenidas por ajuste univariante para cada uno de los modelos de la serie. La constante $M \leq n-1$ es un entero fijo y debe ser elegido a priori. La distribución asintótica de $S_M$ es chi-cuadrado bajo la hipótesis nula de independencia y la hipótesis es rechaza para valores grandes del estadístico.

El estadístico Hong generaliza los estadísticos de Haugh. De hecho, el test de Hong es una suma ponderada de los residuos de las correlaciones cruzadas de la forma

$$Q_n = \frac{n \sum_{j=1-n}^{n-1} k^2(j/d) r^2_{\hat{u}\hat{v}}(j) - M_n(k)}{\left[2V_n(k)\right]^{1/2}} \tag{2.26}$$

donde

$$M_n(k) = \sum_{j=1-n}^{n-1} \left(1 - |j|/n\right) k^2(j/d)$$

y

$$V_n(k) = \sum_{j=2-n}^{n-2} \left(1 - |j|/n\right) \left(1 - (|j|+1)/n\right) k^4(j/d).$$

La ponderación depende de una función núcleo $k$ y un parámetro $d$ (ambos tiene que estar seleccionados a priori). Bajo la hipótesis nula, el test estadístico $Q_n$ es asintóticamente una $N(0,1)$ y se rechaza la hipótesis nula para valores grandes de $Q_n$. En general, algunos autores (véase Fan, Y., et al. 1999) han indicado que el comportamiento asintótico del test basado en el núcleo para muestras finitas tiene poco valor y que sus resultados dependen en gran medida del ancho de banda (véase Skaug, H. J., et al. 1993).

## 2.8. Permutaciones

**Definición 2.8.1.** *Dado un conjunto finito con todos sus elementos diferentes, llamamos permutación a cada una de las posibles ordenaciones de los elementos de dicho conjunto. Así, formalmente, una permutación de un conjunto X es una biyección, $\sigma$, de X en sí mismo. La primera forma de representar una permutación $\sigma$ consiste en escribirla en forma de matriz, situando en primera línea los elementos del dominio $0, 1, 2, ..., n-1$, y en segunda línea las imágenes correspondientes a los elementos reordenados por la biyección, $\sigma(0), \sigma(1), \sigma(2), ..., \sigma(n-1)$:*

$$\begin{pmatrix} 0 & 1 & \dots & n-1 \\ \sigma(0) & \sigma(1) & \dots & \sigma(n-1) \end{pmatrix}$$

*Denotaremos por $S_n$ el grupo formado por todas las permutaciones de un conjunto de n elementos. Obsérvese que el cardinal de este conjunto $S_n$ es $n! = n(n-1)(n-2)\cdots 2$.*

Ejemplo: Dado un conjunto ordenado $\{0, 1, 2, 3, 4, 5, 6, 7\}$ podemos expresar una permutación $\sigma$ sobre éste mediante una matriz de correspondencias de la forma:

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 4 & 5 & 6 & 7 & 0 & 1 \end{pmatrix}$$

Otra forma de notación de las permutaciones es la llamada notación de ciclos. Un ciclo de longitud $s$ es una permutación que intercambia cíclicamente $s$ elementos y fija los restantes. Esta notación revela mejor la estructura interna de la permutación. Para ello:

1 Empezamos con cualquier elemento. Lo escribimos, a su derecha escribimos su imagen, a la derecha de esta, la imagen de su imagen, y seguimos así hasta que se complete un ciclo.

2 Luego cogemos cualquier elemento no contenido en el primer ciclo, volvemos a escribir su imagen a su derecha, y continuamos hasta completar el segundo ciclo.

3 El proceso continúa hasta que la permutación entera ha quedado descrita como producto de ciclos disjuntos.

Utilizando el ejemplo inicial, vamos a expresar la permutación $\sigma$ como composición de ciclos disjuntos:

$$\sigma = (0245)(1367) = (6713)(5024)$$

así pues, $\sigma(0) = 2, \sigma(2) = 4, \sigma(4) = 5, \sigma(5) = 0$ de esta manera hemos cerrado el primer ciclo, a continuación describimos el segundo $\sigma(1) = 3, \sigma(3) = 6, \sigma(6) = 7, \sigma(7) = 1$, quedando de esta forma descompuesta la permutación inicial $\sigma$ en dos ciclos $(0245)(1367)$ disjuntos. De forma análoga se puede comprobar la otra descomposición propuesta.

**Definición 2.8.2.** *(Permutación cíclica) Un tipo especial de permutación es aquella donde no se fija ningún elemento, en cuyo caso la denominaríamos permutación cíclica. Más concretamente, si un ciclo afecta a un elemento $x$ cualquiera del conjunto, al aplicar dicho ciclo reiteradamente todos los elementos afectados por el reordenamiento pasarán por la posición de $x$ en algún momento. Y de forma recíproca, el elemento $x$ pasará por todas las posiciones de todos los elementos afectados por la permutación.*

Sea $n \geq 2$ y $2 \leq r \leq n$. Un *ciclo de longitud* $r$ o $r$*-ciclo* de $S_n$ es una permutación $\sigma$ tal que del conjunto $\{1, 2, ..., n\}$ hay $r$ elementos diferentes secuenciados, $\{a_1, a_2, ..., a_r\}$, para los cuales se cumple que:

1. $\sigma(x) = x$ si $x \notin \{a_1, a_2, ..., a_r\}$ y

2. $\sigma(a_1) = a_2, \sigma(a_2) = a_3, ..., \sigma(a_{r-1}) = a_r$ y $\sigma(a_r) = a_1$.

# Capítulo 3

# Entropía topológica de funciones en la recta real

## 3.1. Introducción

Sea $X$ un espacio topológico compacto (habitualmente se considera métrico), y sea $f : X \to X$ una aplicación continua. Reiterando esta aplicación nosotros obtendremos un sistema dinámico. La entropía topológica surge con la idea de ser una herramienta importante para poder medir cuán complicada es su dinámica, esto es, el número de órbitas muy diferentes que tienen, la rapidez con que se mezcla diversos conjuntos, etc. La entropía topológica de un sistema dinámico será pues un numero no negativo que determinará la complejidad del sistema.

Introducida como ya hemos advertido por *Adler, Konheim y McAndrew* (ver R. L. Adler, et al 1965). Dicha definición se inspira en la dada por *Kolmogorov-Sinaí* ( ver Kolmogorov, A. N., et al,. 1961.) Más tarde, *Dinaburg* (ver Dinaburg, E. I. 1970) por un lado y *Rufus Bowen* (ver R. Bowen, 1971) por otro, dieron una diferente, que recuerda a una definición equivalente de la dimensión de Hausdorff, ellos utilizaron para ello la noción de puntos $\varepsilon$-separados. Esta segunda definición aclara el significado de la entropía topológica: así pues, para un sistema dado por la iterada de una función, la entropía topológica en dicho sistema representará el crecimiento exponencial del número de órbitas de las iteradas.

El objetivo de este capitulo es la introducción de una definición de entropía topológica para funciones continuas tales que, al menos las funciones continuas reales, mantienen la filosofía general siguiente: entropía topológica positiva implica que la función tiene un comportamiento dinámico complicado. Además, nosotros perseguimos que nuestra definición mantenga algunas de las propiedades que se verifican para la definición clásica introducida para compactos.

En el siguiente paso será resaltar de forma cronológica varias definiciones de entropía topológica para espacios topológicos no compactos, anteriores a la introducida por nosotros en el presente capítulo y que justifican en cierto modo la forma de definir la nuestra.

## 3.2. Entropías topológicas sobre no compactos

Aparte de la definición de **Bowen** introducida en el capítulo anterior en la que el espacio no era necesariamente compacto, existen otras definiciones de entropía que a continuación presentamos.

1. Entropía topológica de Bowen

   Sea $X$ un espacio compacto, sea $f : X \to X$ una función continua, sea $Y \subset X$, y $\mathcal{U}$ un recubrimiento finito por abiertos de X. Denotaremos

$E \prec \mathcal{U}$ si E es un abierto que esta contenido en algún elemento de $\mathcal{U}$ y denotaremos $\{E_i\} \prec \mathcal{U}$ una colección de abiertos tales que $E_i \prec \mathcal{U}$ para todo i.

*a*) Llamemos $n_{f,\mathcal{U}}(E)$ al mayor número entero no negativo tal que

$$f^k(E) \prec \mathcal{U} \text{ para todo } k \in [0, n_{f,\mathcal{U}}(E));$$

$n_{f,\mathcal{U}}(E) = 0$ si $E \not\prec \mathcal{U}$ y $n_{f,\mathcal{U}}(E) = +\infty$ si todo $f^k(E) \prec \mathcal{U}$.

*b*) Sea

$$D_{\mathcal{U}}(E) = \exp(-n_{f,\mathcal{U}}(E)) \text{ y } D_{\mathcal{U}}(\mathfrak{E}, \lambda) = \sum_{i=1}^{\infty} D_{\mathcal{U}}(E_i)^{\lambda}$$

siendo $\mathfrak{E} = \{E_i\}_{i=1}^{\infty}$ y $\lambda \in \mathbb{R}$.

*c*) Llamemos entonces $m_{\mathcal{U},\lambda}$ a una medida que definimos de la forma:

$$m_{\mathcal{U},\lambda}(Y) = \lim_{\epsilon \to 0} \inf\{D_{\mathcal{U}}(\mathfrak{E}, \lambda) \text{ tal que } \bigcup E_i \supset Y \text{ y } D_{\mathcal{U}}(E_i) < \epsilon\}.$$

Notemos que $m_{\mathcal{U},\lambda}(Y) \leq m_{\mathcal{U},\lambda'}(Y)$ para $\lambda > \lambda'$ y $m_{\mathcal{U},\lambda}(Y) \notin \{0, \infty\}$ para al menos un $\lambda \in \mathbb{R}$.

Definiremos entonces:

$$h_{\mathcal{U}}(f, Y) = \inf\{\lambda \text{ tal que } m_{\mathcal{U},\lambda}(Y) = 0\}.$$

Y finalmente definiremos entropía topológica de f para Y como:

$$h(f, Y) = \sup_{\mathcal{U}} h_{\mathcal{U}}(f, Y)$$

si hacemos $Y = X$, obtendremos la entropía de f, $\widetilde{h}(f) = h(f, X)$.

2. Entropía topológica de Pesin

En Pesin et al. (1984) se da una manera alternativa de construir la definición de **Bowen** mediante el uso de la noción de presión topológica. Sea $Y$ un espacio métrico arbitrario no compacto, $f : Y \to Y$ una aplicación continua, y $\varphi : Y \to \mathbb{C}$ ó $\mathbb{R}$ una función continua en $Y$. Sea $\mathcal{C}(X)$ el conjunto de la funciones continuas. Consideremos $\mathcal{U}$ un recubrimiento arbitrario finito por abiertos de $Y$. Denotemos por $\mathcal{W}_m(\mathcal{U})$ el conjunto de colecciones de longitud $m$ de elementos del recubrimiento $\mathcal{U}$,

$$\mathcal{W}_m(\mathcal{U}) = \{m(\underline{U}) = \{U_{i_1}, U_{i_2}, \ldots, U_{i_{m-1}}\} \text{ con } U_{ij} \in \mathcal{U}\}.$$

Para una función $\varphi : X \to X$ continua llamaremos

$$S_m\varphi(\underline{U}) = \sup\{\sum_{k=0}^{m-1} \varphi(f^k(x)) : \ x \in Z(\underline{U})\}$$

donde $Z(\underline{U}) = \{x \in Z : \ \ f^k(x) \in U_{i_k}, \ \ k = 0, 1, \ldots, m-1\}$. Si $Z(\underline{U}) = \emptyset$, consideraremos que $S_m\varphi(\underline{U}) = -\infty$.

Sea $\mathcal{W}(\mathcal{U}) = \bigcup\limits_{m \geq 0} \mathcal{W}_m(\mathcal{U})$ diremos que $\Gamma \subset \mathcal{W}(\mathcal{U})$ es un recubrimiento de $Z$ si $Z \subset \bigcup\limits_{\underline{U} \in \Gamma} Z(\underline{U})$. El numero de elementos de la colección $\underline{U}$ lo representaremos por $m(\underline{U})$. Sea

$$\begin{aligned} M(\mathcal{U}, \lambda, Z, \varphi, N) = \inf_{\Gamma \subset \mathcal{W}_m(\mathcal{U})} \{ & \sum_{\underline{\mathcal{U}} \in \Gamma} \exp(-\lambda m(\underline{U}) + S_{m(\underline{U})}\varphi(\underline{U})) \\ & \text{siendo } \Gamma \text{ recubrimiento de } Z \\ & \text{y para cualquier } \underline{U} \in \Gamma,\ m(\underline{U}) \geq N\} \end{aligned} \tag{3.1}$$

*a*) Definimos en $Y$ una medida $m(\mathcal{U}, \lambda, Z, \varphi)$, con $Z \subset Y$ mediante la formula:

$$m(\mathcal{U}, \lambda, Z, \varphi) = \lim_{N \to \infty} M(\mathcal{U}, \lambda, Z, \varphi, N)$$

b) Definimos ahora la presión topológica en el conjunto $Z$ correspondiente a la función $\varphi$, mediante la igualdad:

$$P_Z^*(\varphi) = \sup P_Z(\mathcal{U}, \varphi)$$

siendo $P_Z(\mathcal{U}, \varphi) = \text{ínf}\{\lambda; m(\mathcal{U}, \lambda, Z, \varphi)\}$.

c) Si $Y$ es un subconjunto de un espacio métrico compacto $X$ y la métrica de $Y$ es la inducida por la métrica de $X$ se demuestra que $P_Z^*(\varphi) = P_Z(\varphi)$ para todo $\varphi \in C(X)$. En lo que sigue que limitaremos nuestra discusión al caso $\varphi = 0$.

d) Llamaremos entonces entropía topológica de la aplicación $f$ del conjunto no compacto $Y$ a la cantidad

$$h^*(Y, f) = P_Y^*(0).$$

e) Representaremos la definición de entropía topológica de Bowen para conjuntos no compactos por la expresión $\widetilde{h}$. Entonces podemos asegurar que se verifica:

$$h^*(Y, f) = \widetilde{h}(f)$$

(ver Pesin, et al. 1984).

3. Entropía topológica de Hofer

Sea X un espacio Hausdorff no compacto, y sea $f : X \to X$ una aplicación continua. En la definición de la entropía topológica para los espacios no compactos, por lo menos dos enfoques parecen naturales: uno es la compactificación del espacio y considerar la extensión $f^*$ de $f$ de la compactificación $X^*$ de X; el segundo enfoque sería considerar solo recubrimientos finitos de X. Hofer define la entropía topológica de f para espacios no compactos de dos formas diferentes: $h^2(f)$ y $h^3(f)$.

*a*) La entropía $h^2(f)$ se define por la expresión $h^2(f) = h(f^*)$, donde $f^*$ es la única extensión continua de f en la compactificación de Stone-Čech $X^*$ de X siendo h la entropía topológica introducida por Adler-Konheim-McAndrew.

*b*) La entropía $h^3(f)$ se define como $h^3(f) = \sup_{\alpha\in\mathcal{A}_g} h(f, \alpha,)$. Entendiendo por $\mathcal{A}_g$ la clase de todos los recubrimientos finitos por abiertos de X.

*c*) Si $X$ es un espacio compacto pero no Hausdorff entonces $h^2(f) = h^3(f)$.

*d*) Si $X$ es un espacio no compacto pero sí es Hausdorff entonces $h^2(f) \leq h^3(f)$.

*e*) Sea $X$ es un espacio normal $T_1$. Entonces $h^2(f) = h^3(f)$.

4. Entropía de *Gurevič*

Sea $\sum = \mathbb{N}^{\mathbb{Z}}$ y M es una matriz infinita cuyas entradas son 0 y 1. Construimos un subconjunto $X \subset \sum$ de forma que $(x_n) \in X$ si y solo si el término de M dado por $m_{x_n x_{n+1}} = 1$. La métrica $d((x_n), (y_n)) = \frac{1}{2^i}$, donde $i \geq 0$ es el primer natural tal que $x_i \neq y_i$, hace de X un espacio métrico. Si el número de unos en M es finito, entonces el número de diferentes símbolos en sucesiones de X es finito y X es un compacto. En este caso, la aplicación shift $G : X \to X$ dada por $G(x_n) = (x_n + 1)$ es continua y su entropía fue estudiada por *Parry, W.(1964)* siendo

$$h(G) = \log \lambda M$$

donde $\lambda M$ es el radio espectral de M ó el máximo del módulo de los valores propios de M.

Otros resultados prueban que existe una única medida $\mu_0$, invariante relativa a $T_\Gamma$ (la topología de las cadenas de Marcov), para la cual la entropía métrica $h_{\mu_0}(T_\Gamma)$ coincide con la entropía topológica de $T_\Gamma$, $h_{top}(T_\Gamma)$ (para el resto de las medidas invariantes se verifica que $h_\mu(T_\Gamma) < h_{top}(T_\Gamma)$).

Cuando el número de unos en M es infinito, X no es compacto aunque G es un homeomorfismo y $\Gamma$ es un grafo conexo, siendo $\Gamma'$ un subgrafo conexo de $\Gamma$. Nos encontramos con la dificultad que supone la no compacidad del espacio $X$, dificultad que puede superarse mediante la definición de la entropía topológica como

$$h(T_\Gamma) = \sup_{\Gamma'} \log \lambda(\Gamma') = \sup_{\Gamma'} h_{top}(\sigma'_\Gamma)$$

donde el supremo esta tomado sobre todas las submatrices de M con cantidad finita de unos. Esta entropía topológica $h(T_{\Gamma)}$ está convenientemente calculada, y no resulta complicado poder construir una sucesión $\{\mu_n\}$ de medidas, invariantes relativas a $h(T_\Gamma)$, tal que

$$h_{\mu_n}(T_\Gamma) \to h(T_\Gamma) \quad para \quad n \to \infty.$$

Otra forma natural de resolver el problema que supone la no compacidad de $X$, es compactificando dicho espacio, y aplicando entonces la definición ordinaria de entropía topológica esto es

$$\widetilde{h}_G = h(G^*)$$

donde $G^*$ es la extensión de G a la compactificación por un punto de X. Se prueba en *Gurevič* (1969) que

$$\widetilde{h}_G(T_\Gamma) = h_G(G)$$

A continuación presentamos nuestra definición de entropía topológica para espacios topológicos no compactos y sus propiedades generales. En segundo lugar, estudiaremos una variación de esta noción. Finalmente, estudiaremos el caso particular de entropía topológica para funciones definidas en la recta real.

## 3.3. Definición y propiedades generales

Sea $(X, d)$ un espacio métrico y sea $f$ una aplicación continua de $X$ en si mismo. Nuestra definición de entropía topológica para $f$ es la siguiente:

$$\text{ent}(f) := \sup\{h(f|_K) \text{ con } K \subseteq X, \text{ compacto e invariante por } f\}.$$

Por el Teorema 2.2.16 (i), tenemos que

$$\text{ent}(f) = \sup\{h(f|_K) \text{ con } K \in \mathcal{K}(X, f)\},$$

donde $\mathcal{K}(X, f)$ es la familia de todos los subconjuntos compactos de $X$ los cuales son estrictamente invariantes por $f$. Nótese que esta definición tiene sentido cuando es $X$ métrico o simplemente es un espacio topológico.

Nuestra definición está, en cierto sentido, inspirada en la definición establecida por *Gurevič* para el caso de los homeomorfismos shift. Aunque este echo no es una garantía para nuestra definición (ver Sección 3.4).

Comenzamos con el estudio de algunas propiedades de la nueva definición de entropía topológica, las cuales resumimos en el siguiente resultado.

**Teorema 3.3.1.** *Sea $X$ e $Y$ dos espacios métricos, y sean $f : X \to X$ y $g : Y \to Y$ dos aplicaciones continuas. Entonces las siguientes propiedades son ciertas:*

*(a) Sea $\varphi : X \to Y$ continua tal que*

$$g \circ \varphi = \varphi \circ f.$$

*Entonces:*

*(a1) Si $\varphi : X \to Y$ es inyectiva, entonces*

$$\mathrm{ent}(f) \leq \mathrm{ent}(g).$$

*(a2) Si $\varphi : X \to Y$ es sobreyectiva y tal que el conjunto $\varphi^{-1}(K)$ es compacto para todo compacto $K \subseteq X$, entonces*

$$\mathrm{ent}(f) \geq \mathrm{ent}(g).$$

*(a3) Si $\varphi : X \to Y$ es un homeomorfismo, entonces*

$$\mathrm{ent}(f) = \mathrm{ent}(g).$$

*(b) La formula*

$$\mathrm{ent}(f) = \mathrm{máx}\{h(f|_{X_i}) : i = 1, 2, ...., n\}$$

*donde $X = \bigcup_{i=1}^{n} X_i$, y $X_i$ son subconjuntos invariantes por $f$ para $i = 1, 2 \ldots n$, en general no es cierta.*

*(c) Para todo $n \in \mathbb{N}$,*

$$\mathrm{ent}(f^n) = n \cdot \mathrm{ent}(f).$$

*(d) Sea $f \times g : X \times Y \to X \times Y$ tal que $(f \times g)(x, y) = (f(x), g(y))$ para todo $(x, y) \in X \times Y$. Entonces*

$$\mathrm{ent}(f \times g) = \mathrm{ent}(f) + \mathrm{ent}(g).$$

*(e) Si $f$ es un homeomorfismo, entonces*

$$\mathrm{ent}(f) = \mathrm{ent}(f^{-1}).$$

*(f)* *Sea $\varphi : X \to Y$ continua, sobreyectiva, verificando que $\varphi \circ f = g \circ \varphi$ y tal que si $K$ es compacto en $Y$ entonces $\varphi^{-1}(K)$ es compacto en $X$. Entonces*

$$\begin{aligned}
&\max\{\mathrm{ent}(g), \sup\{h(f, \varphi^{-1}(y)) : y \in \varphi(K), K \in \mathcal{K}(X, f)\}\} \\
&\le \mathrm{ent}(f) \\
&\le \mathrm{ent}(g) + \sup\{h(f, \varphi^{-1}(y)) : y \in Y\}.
\end{aligned} \tag{3.2}$$

*(g)* *Sean $f : X \to Y$ y $g : Y \to X$ continuas. Entonces*

$$\mathrm{ent}(g \circ f) = \mathrm{ent}(f \circ g).$$

*(h)* *Sean $f : X \to Y$ y $g : Y \to X$ continuas, y $F : X \times Y \to X \times Y$ se define por $F(x, y) = (g(y), f(x))$, $(x, y) \in X \times Y$. Entonces*

$$\mathrm{ent}(F) = \mathrm{ent}(f \circ g) = \mathrm{ent}(g \circ f).$$

*(i)* *Sea $X_\infty = \bigcap_{n \ge 0} f^n(X)$. Entonces*

$$\mathrm{ent}(f) = \mathrm{ent}(f|_{X_\infty}).$$

*(j)* *Sea $\Omega(f)$ el conjunto non-wandering de $f$. Entonces*

$$\mathrm{ent}(f) = \mathrm{ent}(f|_{\Omega(f)}).$$

*Demostración.* (a1) Sea $K \in \mathcal{K}(X, f)$. Entonces $\varphi(K) \in \mathcal{K}(Y, g)$. Por el teorema 2.2.16 (a1), $(f|_K) \le h(g_{\varphi(K)})$. Entonces

$$\begin{aligned}
\mathrm{ent}(f) &= \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\} \\
&\le \sup\{h(g|_{\varphi(K)}) : K \in \mathcal{K}(X, f)\} \\
&\le \sup\{h(g|_L) : L \in \mathcal{K}(Y, g)\} \\
&= \mathrm{ent}(g).
\end{aligned} \tag{3.3}$$

(a2) Sea $K \in \mathcal{K}(Y, g)$. Entonces $\varphi^{-1}(K) \in \mathcal{K}(X, f)$. Como $\varphi$ es sobreyectiva, por el Teorema 2.2.16 (a2), $h(f|_{\varphi^{-1}(K)}) \geq h(g|_K)$. Entonces

$$\begin{aligned}
\operatorname{ent}(f) &= \sup\{h(f|_L) : L \in \mathcal{K}(X, f)\} \\
&\geq \sup\{h(f|_{\varphi^{-1}(K)}) : K \in \mathcal{K}(Y, g)\} \\
&\geq \sup\{h(g|_K) : K \in \mathcal{K}(X, f)\} \\
&= \operatorname{ent}(g).
\end{aligned} \tag{3.4}$$

(a3) Esta es consecuencia de (a1) y (a2).

(b) Sea $f : X \to X$ homeomorfismo minimal definido en un conjunto compacto con entropía topológica positiva (ver M. Rees 1981). Sea $x \in X$ y definimos $X_1 := \operatorname{FullOrb}_f(x) = \{f^n(x) : n \in \mathbb{Z}\}$ y $X_2 := X \setminus X_1$. Es evidente que ambos conjuntos son invariantes por $f$ y que no tienen subconjuntos compactos invariantes (en el caso de que ellos tengan subconjuntos compactos invariantes, la función $f$ no es minimal). Por lo tanto $\operatorname{ent}(f|_{X_i}) = 0$, $i = 1, 2$, mientras que $\operatorname{ent}(f) = h(f) > 0$.

(c) Sea $K \in \mathcal{K}(X, f)$. Entonces $K \in \mathcal{K}(X, f^n)$. Por el Teorema 2.2.16 (c), $h(f^n|_K) = h((f|_K)^n) = n \cdot h(f|_K)$. Entonces

$$\begin{aligned}
\operatorname{ent}(f^n) &= \sup\{h(f^n|_L) : L \in \mathcal{K}(X, f^n)\} \\
&\geq \sup\{h(f^n|_K) : K \in \mathcal{K}(X, f)\} \\
&= n \cdot \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\} \\
&= n \cdot \operatorname{ent}(f).
\end{aligned} \tag{3.5}$$

Ahora demostraremos la desigualdad en el sentido contrario. Sea $K \in \mathcal{K}(X, f^n)$, $n \in \mathbb{N}$. El conjunto $\widehat{K} = \bigcup_{i=0}^{n-1} f^i(K)$ es compacto en $X$ y tal que

$$f(\widehat{K}) = f\left(\bigcup_{i=0}^{n-1} f^i(K)\right) = f(K) \cup f^2(K) \cup \cdots \cup f^{n-1}(k) \cup K = \widehat{K}.$$

Por lo tanto $\widehat{K} \in \mathcal{K}(X, f)$. Entonces, por el Teorema 2.2.16 (b) y (c),

$$\begin{aligned} n \cdot h(f|_{\widehat{K}}) &= h((f|_{\widehat{K}})^n) = h(f^n|_{\widehat{K}}) \\ &= \sup\{h(f^n|_{f^i(K)}) : i = 0, 1, \ldots, n-1\} \\ &\geq h(f^n|_K). \end{aligned} \tag{3.6}$$

entonces

$$\begin{aligned} n \cdot \mathrm{ent}(f) &\geq n \cdot \sup\{h(f|_{\widehat{K}}) : K \in \mathcal{K}(X, f^n)\} \\ &= \sup\{h(f^n|_K) : K \in \mathcal{K}(X, f^n)\} \\ &= \mathrm{ent}(f^n). \end{aligned} \tag{3.7}$$

(d) Sea $K \in \mathcal{K}(X \times Y, f \times g)$. Sean $\Pi_1(x, y) = x$ y $\Pi_2(x, y) = y$. Por el Teorema 2.2.16 (d), $h(f|_K) \leq h(f \times g|_{K_1 \times K_2}) = h(f|_{K_1}) + h(g|_{K_2})$, donde $K_i = \Pi_i(K)$, $i = 1, 2$ (recordemos que $K \subseteq K_1 \times K_2$). Entonces:

$$\begin{aligned} \mathrm{ent}(f \times g) &= \sup\{h(f \times g|_K) : K \in \mathcal{K}(X \times Y, f \times g)\} \\ &\leq \sup\{h(f \times g|_{K_1 \times K_2}) : K_1 \times K_2 \in \mathcal{K}(X \times Y, f \times g)\} \\ &\leq \sup\{h(f|_{K_1}) : K_1 \in \mathcal{K}(X, f)\} + \sup\{h(g|_{K_2}) : K_2 \in \mathcal{K}(Y, g)\} \\ &= \mathrm{ent}(f) + \mathrm{ent}(g). \end{aligned}$$

Ahora provaremos la desigualdad en sentido contrario. Sea $K_1 \in \mathcal{K}(X, f)$ y $K_2 \in \mathcal{K}(Y, g)$. Por el Teorema 2.2.16 (d), $h(f \times g|_{K_1 \times K_2}) = h(f|_{K_1}) + h(g|_{K_2})$. Entonces:

$$\begin{aligned} \mathrm{ent}(f \times g) &= \sup\{h(f \times g|_K) : K \in \mathcal{K}(X \times Y, f \times g)\} \\ &\geq \sup\{h(f \times g|_{K_1 \times K_2}) : K_1 \in \mathcal{K}(X, f) \text{ y } K_2 \in \mathcal{K}(Y, g)\} \\ &= \sup\{h(f|_{K_1}) : K_1 \in \mathcal{K}(X, f)\} + \sup\{h(g|_{K_2}) : K_2 \in \mathcal{K}(Y, g)\} \\ &= \mathrm{ent}(f) + \mathrm{ent}(g). \end{aligned}$$

(e) Sea $K \in \mathcal{K}(X, f)$. Está claro que $K \in \mathcal{K}(X, f^{-1})$. Por el Teorema

2.2.16 (d), $h(f|_K) = h(f^{-1}|_K)$. Por lo tanto

$$\begin{aligned} \text{ent}(f) &= \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\} \\ &= \sup\{h(f^{-1}|_K) : K \in \mathcal{K}(X, f)\} \\ &= \text{ent}(f^{-1}). \end{aligned}$$

(f) Sea $K \in \mathcal{K}(X, f)$. Entonces $\varphi(K) \in \mathcal{K}(Y, g)$. Por el Teorema 2.2.16 (f),

$$h(f|_K) \leq h(g|_{\varphi(K)}) + \sup\{h(g, \varphi^{-1}(y)) : y \in \varphi(K)\}$$

y entonces

$$\begin{aligned} \text{ent}(f) &= \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\} \\ &\leq \sup\{h(g|_{\varphi(K)}) + \sup\{h(g, \varphi^{-1}(y)) : y \in \varphi(K)\}\} \\ &\leq \sup\{h(g|_L) : L \in \mathcal{K}(Y, g)\} + \sup\{h(g, \varphi^{-1}(y)) : y \in Y\} \\ &= \text{ent}(g) + \sup\{h(g, \varphi^{-1}(y)) : y \in Y\}. \end{aligned}$$

Como la otra desigualdad es inmediata, finalizamos la demostración.

(g) Sea $K \in \mathcal{K}(X, g \circ f)$. Entonces $f(K) \in \mathcal{K}(Y, f \circ g)$. Por el Teorema 2.2.16 (g), $h(g \circ f|_K) = h(f \circ g|_{f(K)})$. Entonces

$$\begin{aligned} \text{ent}(g \circ f) &= \sup\{h(g \circ f|_K) : K \in \mathcal{K}(X, g \circ f\} \\ &= \sup\{h(f \circ g|_{f(K)}) : K \in \mathcal{K}(X, g \circ f)\} \\ &\leq \sup\{h(f \circ g|_L) : L \in \mathcal{K}(Y, f \circ g)\} \\ &= \text{ent}(f \circ g). \end{aligned}$$

Mediante un razonamiento semejante, probamos que $\text{ent}(f \circ g) \leq \text{ent}(g \circ f)$.

(h) Como $F^2(x,y) = ((g\circ f)(x), (f\circ g)(y))$, entonces

$$\begin{aligned} 2\cdot \mathrm{ent}(F) &= \mathrm{ent}(F^2) = \mathrm{ent}((g\circ f)\times (f\circ g)) \\ &= \mathrm{ent}(g\circ f) + \mathrm{ent}(f\circ g) \\ &= 2\cdot \mathrm{ent}(g\circ f) \\ &= 2\cdot \mathrm{ent}(f\circ g). \end{aligned}$$

(i)Esta igualdad es inmediata porque cualquier $K \in \mathcal{K}(X,f)$ esta contenido en $X_\infty$.

(j) Sea $i : \Omega(f) \to X$ definida por $i(x) = x$. Recordemos que $(f\circ i)(x) = (i\circ f|_{\Omega(f)})(x)$ para cualquier $x \in \Omega(f)$. Como $i$ es inyectiva, por el apartado (a1), $\mathrm{ent}(f) \geq \mathrm{ent}(f|_{\Omega(f)})$.

Sea $K \in \mathcal{K}(X,f)$ y sea $\widehat{K} = K\cap\Omega(f) \in \mathcal{K}(X, f|_{\Omega(f)})$. Además, $(f|_K)|_{\Omega(f)} = f|_{\widehat{K}} = (f|_{\Omega(f)})|_K$. Por el Teorema 2.2.16 (j), $h(f|_K) = h(f|_{K\cap\Omega(f)})$. Entonces:

$$\begin{aligned} \mathrm{ent}(f) &= \sup\{h(f|_K) : K \in \mathcal{K}(X,f)\} \\ &= \sup\{h(f|_{\Omega(f)\cap K}) : K \in \mathcal{K}(X,f)\} \\ &\leq \sup\{h(f|_L) : L \in \mathcal{K}(X, f|_{\Omega(f)})\} \\ &= \mathrm{ent}(f|_{\Omega(f)}). \end{aligned}$$

$\square$

**Ejemplo 1.** *El siguiente ejemplo debido a Henk Bruin muestra que el Teorema 3.3.1 (a2) no es cierto en general si $\varphi$ no cumple la condición $\varphi^{-1}(K)$ es compacto para todo $K$ compacto. Sea $f : \mathbb{R} \to \mathbb{R}$ definida por $f(x) = 2x$ para todo $x \in \mathbb{R}$ y sea $g : S^1 \to S^1$ definido también por $g(x) = 2x$ para todo $x \in S^1$. Entonces $g$ es un factor de $f$ y $\mathrm{ent}(f) = 0$ mientras que $\mathrm{ent}(g) = h(g) = \log 2$ (ver R. L. Adler, et al. 1965). Algunos otros ejemplos de este hecho se pueden ver en ( D. Fiebig, et al. 1986).*

**Ejemplo 2.** *Cuando $X$ es un espacio topológico compacto (no métrico) y $f : X \to X$ es continua, existe una definición de entropía topológica de $f$ que utiliza recubrimientos por abiertos de $X$ (ver R. L. Adler, et al. 1965). Si $X$ no es compacto, podemos definir de nuevo* $\mathrm{ent}(f) = \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\}$. *Entonces podemos probar que las propiedades (a), (c)–(e) y (g)–(j) del Teorema 3.3.1 son también ciertas en este contexto, que es, para espacios topológicos no metrizables.*

Una vez introducida nuestra definición, debemos precisar que no es equivalente a ninguna de las definiciones mencionadas anteriormente en esta memoria. Este hecho es claro en el caso de la 1ª definición de *Bowen*. Para la segunda, puede verse en (Pesin et al. 1984) que en general no tienen porqué coincidir las entropías de $f$ y la $f|\Omega(f)$, lo que muestra que nuestra definición no puede ser equivalente. Igualmente, en el caso de la definición de *Hofer*, vemos que la entropía para la función $f : \mathbb{Z} \to \mathbb{Z}$ dada por $f(n) = n+1$ cumple que $h^2(f) = \infty$. Al ser $\mathbb{Z}$ normal concluimos que también $h^3(f) = \infty$, mientras que $ent(f) = 0$.

## 3.4. Un principio variacional restringido y la relación con compactificación

Sea $f : X \to X$ continua con $X$ métrico. Sea $\mathcal{M}(X, f)$ el conjunto de todas las medidas de probabilidad $\mu$ definida en la $\sigma$–algebra de Borel de $X$, $\beta$, de tal manera que se tiene $\mu(A) = \mu(f^{-1}(A))$ para todo $A \in \beta$. Sea $\mathcal{E}(X, f)$ el conjunto de las medidas ergódicas de f, esto es, $\mu \in \mathcal{M}(X, f)$ para los que la condición $f^{-1}(A) = A$ implica que $\mu(A) = 0$ or $\mu(A) = 1$. El principio variacional para la entropía topológica (ver P. Walters 1982) establece que si $X$ es compacto, entonces $h(f) = \sup\{h_\mu(f) : \mu \in \mathcal{E}(X, f)\}$, donde $h_\mu(f)$ es la entropía métrica de $f$. Si $X$ es no compacto, podemos definir

$$h_V(f) = \sup\{h_\mu(f) : \mu \in \mathcal{E}(X, f)\}$$

(ver M. Handel, et al. 1995). Entonces es inmediato que $\mathrm{ent}(f) \leq h_V(f)$. El siguiente ejemplo prueba que la desigualdad puede ser estricta.

**Ejemplo 3.** *Consideremos el ejemplo del Teorema 3.3.1 (b) y la función* $f|_{X_2} : X_2 \to X_2$. *Ya que* $X_1$ *es una orbita no periódica,* $\mathcal{M}(X_1, f) = \emptyset$*, y por lo tanto* $\mathcal{E}(X_1, f) = \emptyset$*. Entonces* $\sup\{h_\mu(f|_{X_2}) : \mu \in \mathcal{E}(X_2, f|_{X_2})\} > 0$*. Por otra parte* $\mathrm{ent}(f|_{X_2}) = 0$ *por que* $\mathcal{K}(X_2, f_{X_2}) = \emptyset$*.*

Representaremos por $\mathcal{B}(X, f)$ el conjunto de medidas ergódicas invariantes $\mu$ de $f$ tal que $\mathrm{supp}\mu \subseteq K$, para algún $K \in \mathcal{K}(X, f)$, donde $\mathrm{supp}\mu$ representa el subconjunto compacto más pequeño de medida uno. Entonces

$$\begin{aligned}
\mathrm{ent}(f) &= \sup\{h(f|_K) : K \in \mathcal{K}(X, f)\} \\
&= \sup_{K \in \mathcal{K}(X,f)} \sup\{h_\mu(f) : \mu \in \mathcal{M}(K, f|_K)\} \\
&= \sup\{h_\mu(f) : \mu \in \mathcal{B}(X, f)\},
\end{aligned}$$

lo que constituye un principio variacional para la entropía topológica de conjuntos no compactos.

El principio variacional está conectado con la siguiente cuestión. Sea $f : X \to X$ una aplicación continua donde $X$ es no compacto. Supongamos que $X^*$ es una compactificación de $X$ tal que existe una extensión continua de $f$ en $X^*$, representada por $f^*$. ¿Cual es la relación entre $\mathrm{ent}(f)$ y $h(f^*)$? En general está claro que $h(f^*) \geq \mathrm{ent}(f)$. Nosotros estudiaremos bajo que condiciones se verifica la igualdad. El siguiente ejemplo de (J. E. Hofer. 1974) muestra que esto no es posible en general.

**Ejemplo 4.** *Sea $\mathbb{Z}$ el conjunto de los numeros enteros y sea $f : \mathbb{Z} \to \mathbb{Z}$ la función definida por $f(n) = n + 1$ para todo $n \in \mathbb{Z}$. Sean $\mathbb{Z}^*$ y $f^*$ la compactificación de Stone-Cech de $\mathbb{Z}$. Entonces podemos ver en (J. E. Hofer 1974) que $h(f^*) = \infty$. Ya que $\mathcal{K}(\mathbb{Z}, f) = \emptyset$, nosotros obtendremos que* $\mathrm{ent}(f) = 0$.

*Es importante señalar que la función del ejemplo anterior tiene una dinámica muy simple. En efecto, la función $F : \mathbb{R} \to \mathbb{R}$ definida por $F(x) = x+1$, tiene una dinámica simple, baste con resaltar que no tiene puntos fijos, no tiene puntos recurrentes, no tiene ciclos, no hay compactos invariantes y cualquier orbita que consideremos acabará siempre en el $\infty$. Es claro por otro lado que $f = F|_{\mathbb{Z}}$.*

**Teorema 3.4.1.** *Sea $X$ métrico y sea $f : X \to X$ continua tal que $\mathcal{B}(X, f) = \mathcal{E}(X, f)$. Supongamos que existe una compactificación $X^*$ la cual es métrica y tal que $X^* \setminus X$ es numerable. Supongamos que existe una extensión continua $f^* : X^* \to X^*$. Entonces* $\mathrm{ent}(f) = h(f^*)$.

*Demostración.* El argumento para la demostración es similar a la del ( M. Handel, et al. 1995). Como alguna medida ergódica de $f$ pertenece a $\mathcal{B}(X, f)$ y $X^* \setminus X$ es a lo sumo numerable, entonces el conjunto $\mathcal{E}(X^*, f^*) \setminus \mathcal{E}(X, f)$ contiene como mucho medidas asociadas a puntos periódicos. Así, $h_\mu(f^*) = 0$ para todo $\mu \in \mathcal{E}(X^*, f^*) \setminus \mathcal{E}(X, f)$. Entonces $h(f^*) = \sup\{h_\mu(f^*) : \mu \in \mathcal{E}(X^*, f^*)\} = \sup\{h_\mu(f^*) : \mu \in \mathcal{E}(X^*, f^*) \setminus \mathcal{M}(X, f)\}$, y por lo tanto $h(f^*) = \mathrm{ent}(f)$. $\square$

La condición $\mathcal{B}(X, f) = \mathcal{E}(X, f)$ no se pueden evitar, como muestra el siguiente ejemplo.

**Ejemplo 5.** *Consideremos el ejemplo definido en la demostración del Teorema 3.3.1 (b). En este ejemplo la compactificación del conjunto $X_2$ es $X$ y la extensión de $f_2$ es $f$. Entonces $h(f) > 0$ mientras que* $\mathrm{ent}(f_2) = 0$.

## 3.5. Entropía topológica de funciones en la recta real

En esta sección estudiaremos la entropía topológica de funciones continuas definidas en la recta real. Dado un intervalo compacto $[a,b]$, sea $f_{[a,b]} : [a,b] \to [a,b]$ la función continua definida por

$$f_{[a,b]}(x) := \begin{cases} f(x) & \text{si } f(x) \in [a,b], \\ b & \text{si } f(x) > b, \\ a & \text{si } f(x) < a. \end{cases}$$

Recordemos que si $x \in [a,b]$, el conjunto $\omega_{f_{[a,b]}}(x)$ es el conjunto de los puntos limites de la orbita de $x$ bajo $f_{[a,b]}$. Entonces $\omega(f_{[a,b]}) = \bigcup_{x\in[a,b]} \omega_{f_{[a,b]}}(x)$.

**Lema 4.** *Sea $f : \mathbb{R} \to \mathbb{R}$ continua y supongamos que $\mathrm{Orb}_f(a)$ y $\mathrm{Orb}_f(b)$ están en $[a,b]$. Entonces existe $K \in \mathcal{K}(\mathbb{R}, f)$, $K \subseteq [a,b]$, tal que $h(f_{[a,b]}) = h(f|_K)$.*

*Demostración.* Si $[a,b]$ es invariante por $f$ no hay nada que probar. Supongamos que $[a,b]$ no es invariante y sea $\mathcal{A} := \{x \in [a,b] : f(x) \notin [a,b]\}$ y $\mathcal{A}_\infty := \bigcup_{n\geq 0} f^{-n}(\mathcal{A})$. Está claro que $\mathrm{Orb}_f(a)$ y $\mathrm{Orb}_f(b)$ están contenidos en $[a,b] \setminus \mathcal{A}_\infty$. Nosotros probamos que para todo $x \in [a,b]$, $\omega_{f_{[a,b]}}(x) \cap \mathcal{A}_\infty = \emptyset$. Sea $y \in \mathcal{A}_\infty$. Sea $(c,d) \subset \mathcal{A}_\infty$ tal que $y \in (c,d)$. Sea $x \in [a,b]$ tal que $y \in \omega_{f_{[a,b]}}(x)$. Entonces existe $n \geq 0$ tal que $f^n_{[a,b]}(x) \in (c,d)$ y entonces existe $m \geq 0$ tal que $f^{n+m}_{[a,b]}(x) \in \mathcal{A}$. Por lo tanto $f^{n+m+1}_{[a,b]}(x) = a$ o $f^{n+m+1}_{[a,b]}(x) = b$. Entonces $\{f^k_{[a,b]}(x) : k \geq n+m+1\} \cap (c,d) = \emptyset$, y entonces $y \notin \omega_{f_{[a,b]}}(x)$. Como $h(f_{[a,b]}) = h(f_{[a,b]}|_{\omega(f_{[a,b]})})$ y $\omega(f_{[a,b]})$ es compacto (ver L. Block, et al. 1992), la que demuestra el lema. □

**Teorema 3.5.1.** *Sea $f : \mathbb{R} \to \mathbb{R}$ continua. Entonces*

$$\mathrm{ent}(f) = \sup\{h(f_I) : I \textit{ intervalo compacto conteniendo } K \in \mathcal{K}(\mathbb{R}, f)\}.$$

*Demostración.* Sea $\alpha = \sup\{h(f_I) : I$ intervalo compacto conteniendo $K \in \mathcal{K}(\mathbb{R}, f)\}$. Sea $K \in \mathcal{K}(\mathbb{R}, f)$ y se $J$ el intervalo más pequeño compacto que contiene a $K$. Entonces está claro que $h(f|_K) \leq h(f_J)$. Entonces $\mathrm{ent}(f) \leq \alpha$. Por el Lema 4, obtendremos la desigualdad en sentido contrario. □

### 3.5.1. Funciones monótonas a trozos

Una aplicación continua $f : \mathbb{R} \to \mathbb{R}$ se dice que es monótona a trozos si existe $-\infty = x_1 < x_2 < \cdots < x_n = +\infty$ tal que $f|_{(x_i, x_{i+1})}$ es creciente o decreciente para $i = 1, 2, \ldots n$.

**Teorema 3.5.2.** *Sea $f : \mathbb{R} \to \mathbb{R}$ continua y monótona a trozos. Entonces existe un intervalo compacto $[a, b]$ tal que* $\mathrm{ent}(f) = h(f_{[a,b]})$.

*Demostración.* Por el Teorema 3.3.1 (i), podemos considerar que $f$ es sobreyectiva trabajando con $f : \mathbb{R}_\infty \to \mathbb{R}_\infty$. Si $\mathbb{R}_\infty$ esta acotado, no hay nada que probar. Por lo tanto, supondremos que es no acotado de la forma $\mathbb{R}_\infty = [\alpha, +\infty)$, $\alpha \in \mathbb{R}$ (los demas casos son similares). Sea $L$ el conjunto definido por la unión de todos los conjuntos compactos de $\mathcal{K}(X_\infty, f)$. Consideraremos dos casos. Si $L$ es acotado, sea $a := \text{mín} L$ y $b := \text{máx} L$. Si $L$ es no acotado, sea $\mathcal{F} := L \setminus \mathrm{Fix}(f)$, donde $\mathrm{Fix}(f)$ representa el conjunto de puntos fijos de $f$. Ya que $\displaystyle\lim_{x \to +\infty} f(x) = +\infty$ y $f$ es monótona a trozo, existe $b \in \mathrm{Fix}(f)$ tal que $y < b$ para todo $y \in \mathcal{F}$. Sea $a := \text{mín} L$. En ambos casos se considera la función $f_{[a,b]}$. Está claro que $\mathrm{Orb}_f(a)$ y $\mathrm{Orb}_f(b)$ están contenidas en $[a, b]$. Por el Lema 4, existe $K \in \mathcal{K}(\mathbb{R}_\infty, f)$ tal que $h(f_{[a,b]}) = h(f|_K)$. Puesto que no existen elementos de $\mathcal{K}(\mathbb{R}_\infty, f)$ fuera de $[a, b]$, excepto los puntos fijos de $f$, por el Teorema 3.5.1, concluiremos que $\mathrm{ent}(f) = h(f_{[a,b]})$. □

**Nota 1.** *El resultado anterior no es cierto cuando la función no es monótona a trozos. Por ejemplo consideremos que la función $f : \mathbb{R} \to \mathbb{R}$ tal que* $[n, n+1]$

*es estrictamente invariante por $f$, $f(n) = n$ para todo $n \in \mathbb{N}$, y $f(x)$=0 para todo $x < 0$. Definimos $f|_{[n,n+1]}$ tal que tiene $2n+1$ trozos lineales cubriendo $[n, n+1]$. Entonces $h(f|_{[n,n+1]}) = \log(2n+1)$ para todo $n \geq 1$. Entonces* $\mathrm{ent}(f) = \infty$*, mientras que $h(f|_K)$ es finito para todo $K \in \mathcal{K}(\mathbb{R}, f)$.*

Es evidente que el ejemplo anterior puede ser construido tal que $f$ será de clase $C^\infty$. El mismo hecho no puede ocurrir cuando la función se define en un intervalo compacto $[a, b]$, donde la entropía esta acotada por $\log k$, donde $k = \text{máx}\{|f'(x)| : x \in [a, b]\}$ ver P. Walters. (1982).

Para una función monótona a trozos $f$ representamos por $c_n(f)$ el numero de trozos monótonos de $f^n$, $n \geq 1$. En el caso de funciones continuas en un intervalo compacto, se sabe que

$$h(f) = \lim_{n\to\infty} \frac{1}{n} \log c_n(f)$$

ver L. Alsedá, et al. (1993). Ahora probaremos un resultado analogo.

**Teorema 3.5.3.** *Supongamos que $f : \mathbb{R} \to \mathbb{R}$ es continua y monótona a trozos. Entonces*

$$\mathrm{ent}(f) = \lim_{n\to\infty} \frac{1}{n} \log c_n(f)$$

*Demostración.* Puesto que $\lim_{x\to+\infty} f(x) = \beta$ y $\lim_{x\to-\infty} f(x) = \alpha$, consideremos $[-\infty, +\infty]$, los dos puntos de la compactificación de la recta real y definimos una extensión continua de $f$ por $f^*(x) = f(x)$ si $x \in \mathbb{R}$, $f^*(-\infty) = \alpha$ y $f^*(+\infty) = \beta$. Por el Teorema 3.5.2, existe un intervalo compacto $[a, b]$ tal que $\mathrm{ent}(f) = h(f_{[a,b]})$. Además, siguiendo la demostración del Teorema 3.5.2, podemos ver que $[a, b]$ puede ser elegido de tal manera que $\mathbb{R} \setminus [a, b]$ no contiene subconjuntos compactos invariantes excepto para puntos fijos o puntos de periodo dos (donde $f^*(-\infty) = +\infty$ y $f^*(+\infty) = -\infty$). Entonces $\mathcal{B}(\mathbb{R}, f) = \mathcal{E}(\mathbb{R}, f)$ por el Teorema 3.4.1 $h(f^*) = \mathrm{ent}(f)$. Por otra parte, ya que $f^*$ es monótona a trozos, $h(f^*) = \lim_{n\to\infty} \frac{1}{n} \log c_n(f^*)$. Puesto que $c_n(f^*) = c_n(f)$, concluimos la demostración. □

### 3.5.2. Entropía y herraduras

Recordemos que una aplicación continua $f : [a,b] \to [a,b]$ tiene una $k$–herradura, $k \in \mathbb{N}$, si existe $k$ subintervalo compacto $J_i$, $1 \leq i \leq k$ tal que $\bigcup_{i=1}^{k} J_i \subseteq f(J_i)$, $1 \leq i \leq k$. Puede verse en L. Alsedá, et al. (1993) que si $f$ tiene un $k$–herradura, entonces $h(f) \geq \log k$. Por lo tanto, si $h(f) > 0$, existe una sucesión de enteros positivos $s_n$ y $k_n$ tal que $f^{k_n}$ tiene un $s_n$–herradura y

$$h(f) = \lim_{n\to\infty} \frac{1}{k_n} \log s_n.$$

Nosotros probaremos un resultado similar para funciones continuas en la recta real.

**Teorema 3.5.4.** *Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua. Entonces*

*(a) Si $f$ tiene un $s$–herradura, entonces* $\mathrm{ent}(f) > 0$.

*(b) Si* $\mathrm{ent}(f) > 0$, *entonces existe una sucesión de enteros positivos $s_n$ y $k_n$ tal que $f^{k_n}$ tiene un $s_n$–herradura y*

$$\mathrm{ent}(f) = \lim_{n\to\infty} \frac{1}{k_n} \log s_n.$$

*Demostración.* (a) Sea $J_1, ..., J_s$ un $s$–herradura de $f$ y sea $J$ el intervalo más pequeño conteniendo $J_1, ..., J_s$. Entonces está claro que $h(f_J) > \log s$ y por el Teorema 3.5.1, $\mathrm{ent}(f) > \log s$.

(b) Primero, supondremos que $\mathrm{ent}(f) = \alpha < \infty$. Sea $\epsilon_n$ una sucesión de numeros reales positivos tal que $\lim_{n\to\infty} \epsilon_n = 0$. Por el Teorema 3.5.1, para algún $m \in \mathbb{N}$ existe un intervalo compacto $I_m$ y $\delta_m > 0$ tal que $h(f_{I_m}) > 0$

y $\alpha - \epsilon_m \leq h(f_{I_m}) - \delta_m < h(f_{I_m}) < h(f_{I_m}) + \delta_m \leq \alpha + \epsilon_m$. Puesto que $h(f_{I_m}) > 0$, existe $n_0(m) \in \mathbb{N}$ tal que si $n \geq n_0(m)$, entonces

$$h(f_{I_m}) - \delta_m \leq \frac{1}{k_n(m)} \log s_n(m) \leq h(f_{I_m}) + \delta_m,$$

donde $f_{I_m}^{k_n(m)}$ tiene un $s_n(m)$–herradura. Esto significa que $f^{k_n(m)}$ tiene un $s_n(m)$–herradura (por que si $I_m \subseteq f_{I_m}^{k_n(m)}(J)$ para algún intervalo $J \subseteq I_m$, entonces $I_m \subseteq f_{I_m}^{k_n(m)}(J) \subseteq f^{k_n(m)}(J)$). Ahora, fijamos $n_m \geq n_0(m)$ tal que

$$h(f_{I_m}) - \delta_m \leq \frac{1}{k_{n_m}} \log s_{n_m} \leq h(f_{I_m}) + \delta_m,$$

para algún $m \in \mathbb{N}$. Entonces, existe una sucesión de enteros positivos $n_m$ que converge hacia el infinito tal que $\alpha - \epsilon_m \leq \frac{1}{k_{n_m}} \log s_{n_m} \leq \alpha + \epsilon_m$. Por lo tanto

$$\lim_{m\to\infty} \frac{1}{k_{n_m}} \log s_{n_m} = \text{ent}(f).$$

Ahora suponemos que $\text{ent}(f) = +\infty$. Para algún $m \in \mathbb{N}$, por el Teorema 3.5.1, existe un subintervalo compacto $I_m$ tal que $h(f_{I_m}) > m$. Entonces existe $n_0(m) \in \mathbb{N}$ tal que

$$\frac{1}{k_n(m)} \log s_n(m) > m$$

para todo $n \geq n_0(m)$ y tal que $f_{I_m}^{k_n(m)}$ tiene un $s_n(m)$–herradura, el cual es una herradura para $f^{k_n(m)}$. Fijamos $n_m \geq n_0(m)$ y entonces

$$\frac{1}{k_{n_m}} \log s_{n_m} > m$$

para todo $m \in \mathbb{N}$. Entonces

$$\lim_{m\to\infty} \frac{1}{k_{n_m}} \log s_{n_m} = +\infty$$

y concluimos la demostración. □

### 3.5.3. Entropía de funciones topológicamente transitivas

La funciones transitivas en espacios compactos han sido estudiadas ver S. F. Kolyada, et al. (1997). En este estudio, podemos ver que si $f : [a, b] \to [a, b]$ es transitiva, entonces $h(f) > 0$. Nos preguntamos si es verdad el mismo resultado para funciones transitivas de la recta real.

Ahora estudiamos el caso de las funciones topológicamente transitivas en la recta real. Una aplicación continua $f : \mathbb{R} \to \mathbb{R}$ se dice que es *topológicamente transitiva* si para cualquier subconjunto abierto no vacío $V$ y $W$ de $\mathbb{R}$, existe $n \in \mathbb{N}$ tal que $f^n(V) \bigcap W \neq \emptyset$. Está claro que si $f$ es una función transitiva y $K$ es invariante, entonces $\mathrm{Int}(K) = \emptyset$. Las funciones transitivas en la recta real han sido estudiadas en N. Anima, et al. (2001) donde se prueba que si $\mathcal{P}$ es el conjunto de puntos críticos $f$, entonces la transitividad de $f$ implica que $\mathcal{P}$ es no acotado (nosotros decimos que $x \in \mathbb{R}$ es un *punto critico* de $f$ si para cualquier entorno $V$ de $x$ existe distintos $y, z \in V$ y tal que $f(y) = f(z)$).

**Teorema 3.5.5.** *Sea $f : \mathbb{R} \to \mathbb{R}$ continua y transitiva, entonces* $\mathrm{ent}(f) > 0$. *Por tanto, para cualquier real positivo $\alpha$ existe una función $f : \mathbb{R} \to \mathbb{R}$ transitiva tal que* $\mathrm{ent}(f) \geq \alpha$.

*Demostración.* Está claro que $\mathrm{Fix}(f) \neq \emptyset$. Dividimos la demostración en dos casos: (1) existe $a \in \mathrm{Fix}(f)$ tal que $f$ es creciente en $[a, \epsilon)$ o decreciente en $(-\epsilon, a]$ para algún $\epsilon > a$ y (2) para todo $a \in \mathrm{Fix}(f)$ la función $f$ es decreciente en $[a, \epsilon)$ y creciente en $(-\epsilon, a]$ para algún $\epsilon > a$.

(1) Supongamos por ejemplo que $f$ es creciente en $[a, \epsilon)$ para algún $\epsilon > a$ (el otro caso es análogo). Primeramente, notemos que $f(x) > x$ para todo $x \in (a, \epsilon)$, por que de otra forma el intervalo $[a, \epsilon)$ sería invariante por $f$, y

esto contradice que $f$ es transitiva. Sea $c \in \mathrm{Fix}(f)$ el menor punto fijo de $f$ tal que $a < c$. Es evidente que el punto $c$ existe por que de otra forma $f([a, +\infty)) \subset [a, +\infty)$ y entonces $f$ no seria transitiva. Entonces existe $d > c$ tal que $f(d) < a$, por que de otra forma $[a, \infty)$ sería de nuevo invariante y con interior no vacío, una contradicción. Sea $b$ el menor numero en $(c, d)$ tal que $f(b) = a$. Entonces debe existir $e \in (a, b)$ tal que $f(e) > b$, por que de otra forma $[a, b]$ sería invariante. Sea $J_1 = [a, e]$ y $J_2 = [e, b]$. Está claro que $f_{[a,b]}(J_1) = f_{[a,b]}(J_2) = [a, b]$, entonces $f_{[a,b]}$ tiene un 2–herradura y por lo tanto $h(f_{[a,b]}) \geq \log 2$. Por el Teorema 3.5.1 (ver también el Teorema 3.5.4), nosotros concluimos que $\mathrm{ent}(f) \geq \log 2$.

(2) Sea $a \in \mathrm{Fix}(f)$ y consideremos la función $f^2$. Ahora $f^2$ es creciente en $[a, \epsilon)$ para algún $\epsilon > a$. Por N. Anima, et al. (2001), sostenemos que $f^2$ es transitiva para todo $\mathbb{R}$ ó en cualquiera de los conjuntos $[a, \infty)$ y $(-\infty, a]$. Razonando como en el caso anterior (supongamos $f^2|_{[a,\infty)}$ es transitiva, si esta no lo es en toda la recta real) existen $b > e > a$ tal que $f^2(b) = a$, $f^2(e) > b$ y tal que $h(f^2_{[a,b]}) \geq \log 2$. Entonces existe $K \in \mathcal{K}(\mathbb{R}, f^2)$ tal que $h(f^2|_K) = h(f^2_{[a,b]}) \geq \log 2$ y, por el Teorema 3.3.1 (c), $\mathrm{ent}(f) > 0$.

Ahora probaremos la segunda parte del resultado. Dado $\alpha \in \mathbb{R}$, sea $m \in \mathbb{Z}^+$ tal que $\log m \geq \text{máx}\{\alpha, 4\}$. Ahora nosotros construimos una función continua $f : \mathbb{R} \to \mathbb{R}$ como sigue. Consideremos la sucesión de pares $(n, a_n)$, $n \geq 0$ tal que $a_n = -nm/2$ si $n$ es par y $a_n = (n + 1)m/2$ si $n$ es impar. Entonces $f$ en $[n, n + 1]$ es lineal y continua. Finalmente, se define $x < 0$ sea $f(x) = -f(-x)$. Claramente, para $n = 0, ..., m - 1$ es cierto que $[0, m] \subset f([n, n + 1])$ y entonces $f_{[0,m]}$ es un $m$–herradura. Entonces $h(f_{[0,m]}) \geq \log m$ y por lo tanto $\mathrm{ent}(f) \geq \log m \geq \alpha$.

Para concluir la demostración nosotros necesitamos mostrar que $f$ es transitiva. Con este fin, notemos que $\lim\limits_{i\to\infty} f^i(n, n + 1) = \mathbb{R}$ para todo $n \in \mathbb{Z}$.

Esto significa que para cualquier subconjunto abierto $U$ de la recta real y cualquier entero $n$ existe $k \in \mathbb{N}$ tal que $f^k(n, n+1) \cap U \neq \emptyset$. Ahora sea $(a, b) \subset \mathbb{R}$. Dado que la pendiente de cada trozo de pieza lineal de $f$ tiene pendiente con módulo superior a 4, es inmediato comprobar que existe $k \in \mathbb{N}$ tal que $f^k(a, b)$ contiene un intervalo $(n, n+1)$ para algún $n \in \mathbb{N}$. Esto concluye la demostración. $\square$

## 3.6. Cuestiones abiertas

Acabamos este capítulo con una serie de cuestiones abiertas que pensamos, son de interés:

1. Sean $X$ un conjunto localmente compacto y $f : X \to X$ un homeomorfismo. ¿Es cierta la igualdad $ent(f) = h(f^*)$, siendo $f^* : X^* \to X^*$ su extensión a la compactificación por un punto de X.? Nótese que trivialmente es cierto para homeomorfismos de la recta real ya que en este caso ambas son 0.

2. ¿Es posible encontrar funciones transitivas reales cuya entropía sea finita.? Nótese que estas funciones tienen una cantidad infinita de trozos monótonos.

3. Encontrar qué tipo de espacio topológico hace que diferentes definiciones de entropía coincidan. Por ejemplo si $X$ es compacto y métrico, todas las definiciones de esta memoria son equivalentes. Estudiar la posibilidad de encontrar equivalencias con hipótesis mas débiles.

86

# Capítulo 4

# Un test de independencia entre series temporales basado en dinámica simbólica

## 4.1. Introducción

En este capítulo, dados dos procesos temporales $\{X_t\}_{t\in I}$ e $\{Y_t\}_{t\in I}$, estudiaremos la potencial dependencia existente entre ellos, traduciendo el problema (y el conjunto de información) en símbolos a través de una sencilla técnica de simbolización y, a continuación, utilizaremos la entropía asociada a estos símbolos para construir un contraste de independencia entre ambas series. Como resultado, se obtiene un nuevo test no paramétrico de independencia entre series temporales, consistente, que evita restricciones tales como la linealidad o normalidad en los datos y que no necesita la estimación del proceso generador de datos que, si no están bien especificados o no están consistentemente estimados, podría invalidar los resultados asintóticos de los contrastes tipo Haug (1976).

Para probar la dependencia entre $\{X_t\}$ e $\{Y_t\}$ utilizaremos la dinámica simbólica y la entropía de permutación como una medida de la dependencia.

El uso de la entropía ha desempeñado un papel de primer orden como medida de la dependencia entre series temporales en las últimas dos décadas véase, por ejemplo, Harry (1989). Recientemente se ha utilizado el concepto de entropía y dinámica simbólica para aplicarlo al estudio de la dependencia de series temporales véase Matilla y Ruiz (2008). Más concretamente, Matilla y Ruiz (2008), simbolizan las observaciones de una serie temporal con elementos del grupo simétrico introduciendo así un patrón de orden dentro de la serie. Posteriormente deriva la distribución asintótica de una trasformación afín de la entropía de los símbolos (permutaciones) bajo la hipótesis nula de independencia, construyendo de esta manera un test consistente que se distribuye asintóticamente como una Chi-cuadrado. Nosotros utilizaremos la misma idea de simbolización con permutaciones de Matilla y Ruiz (2008) para contruir nuestro test.

A continuación presentamos la notación y varias definiciones de gran utilidad para describir la metodología usada en la construcción del estadístico.

## 4.2. Definiciones y notación

Sea $\{\mathbf{W}_t\}_{t\in I}$ una serie temporal 2-dimensional con

$$\mathbf{W}_t = (X_t, Y_t),$$

donde $X_t$ e $Y_t$ son series temporales de datos reales. Para un entero positivo $m \geq 2$ representamos por

$$S_m = \{\pi_1, \pi_2, \dots, \pi_{m!}\}$$

el grupo simétrico de orden $m!$, grupo formado por todas las permutaciones de longitud $m$. Sea

$$\pi = (i_1, i_2, \ldots, i_m) \in S_m.$$

Llamaremos al elemento $\pi$ del grupo simétrico $S_m$ un símbolo. Consideremos el producto directo de los grupos simétricos $\mathcal{S}_m = S_m \times S_m$. Ahora definiremos un patrón de orden para un símbolo $\eta_{ij} = (\pi_i^x, \pi_j^y) \in \mathcal{S}_m$, para un tiempo $t \in I$. Con este fin consideramos que la serie está (embebida) sumergida en un espacio $2m$-dimensional como sigue:

$$\mathbf{W}_m(t) = (X_{t+1}, X_{t+2}, \ldots, X_{t+m}, Y_{t+1}, Y_{t+2}, \ldots, Y_{t+m})$$

para cada $t \in I$ y denominaremos $2m$-historia al vector $\mathbf{W}_m(t)$.

Al parámetro $m$ lo llamaremos *dimensión de embebimiento*. Sea

$$\pi_i^x = (i_1, i_2, \ldots, i_m) \in S_m$$

un símbolo. Entonces decimos que $t$ es de $\pi_i^x$-tipo para la serie $\{X_t\}_{t\in I}$ si y solo si el simbolo $\pi_i^x = (i_1, i_2, \ldots, i_m)$ es el único elemento en el grupo simétrico $S_m$ satisfaciendo las dos condiciones siguientes:

$$(a) \quad X_{t+i_1} \leq X_{t+i_2} \leq \cdots \leq X_{t+i_m},$$

$$(b) \quad i_{s-1} < i_s \text{ si } X_{t+i_{s-1}} = X_{t+i_s}$$

donde la condición $(b)$ garantiza la unicidad en la asignación del símbolo $\pi_i^x$.

Esto se justifica si los valores de $X_t$ tienen una distribución continua de modo que la igualdad de valores sea muy poco frecuente, con una probabilidad teórica de que se produzca de 0. Análogamente, simbolizaremos la serie $\{Y_t\}_{t\in I}$ con los símbolos $\pi_j^y \in S_m$.

Ahora, sea $\eta_{ij} = (\pi_i^x, \pi_j^y) \in \mathcal{S}_m$. Diremos que $t$ es de $\eta_{ij}$-tipo para $\mathbf{W} = (X, Y)$ o simplemente que $t$ es de $\eta_{ij}$-tipo, si y solo si $t$ es de $\pi_i^x$-tipo para $X$

y de $\pi_j^y$-tipo para $Y$. Notemos que para todo $t$ tal que $t$ es de $\eta_{ij}$-tipo para $\mathbf{W}$ la $2m$-historia $\mathbf{W}_m(t)$ se convierte en un símbolo único $\eta_{ij}$. Con el fin de ver esto, el siguiente ejemplo ayudará al lector.

**Ejemplo 6.** *Tomaremos como dimensión de embebimiento $m = 3$. Así, el grupo simétrico es*

$$S_3 = \{(0,1,2),(0,2,1),(1,0,2),(1,2,0),(2,0,1),(2,1,0)\}.$$

*Consideremos la serie temporal finita de siete valores*

$$\{(X_1,Y_1) = (8,2),(X_2,Y_2) = (7,9),(X_3,Y_3) = (6,5),(X_4,Y_4) = (9,7),$$
$$(X_5,Y_5) = (4,3),(X_6,Y_6) = (10,12),(X_7,Y_7) = (2,5)\}.$$

*Entonces para $t = 2$ tenemos que $X_{t+1} = 7 < X_{t+0} = 6 < X_{t+2} = 9$ y $Y_{t+1} = 5 < Y_{t+2} = 7 < Y_{t+0} = 9$ por lo tanto, tenemos que el periodo $t = 2$ es de $((1,0,2),(1,2,0))$-tipo.*

Una vez estudiado y comprendido el ejemplo anterior, consideremos la serie temporal bidimensional $\{\mathbf{W}_t\}_{t\in I}$ y fijemos $m$ la dimensión de embebimiento. Ahora uno puede fácilmente calcular la frecuencia relativa de los símbolos $\pi_i^x \in S_m$ y $\pi_j^y \in S_m$ sin mas que utilizar la teoría de la probabilidad, teniendo presente que la frecuencia absoluta del simbolo $\pi_i^x$ es

$$n_{\pi_i^x} = \sharp\ \{t \in I \mid t \text{ es } \pi_i^x - \text{tipo para } X\}.$$

Este valor representa el numero de veces que $t$ es de $\pi_i^x -$ tipo para la serie $\{X_t\}_{t\in I}$. Analogamente para la serie $\{Y_t\}_{t\in I}$ se define la frecuencia absoluta del símbolo $\pi_j^y$ como

$$n_{\pi_j^y} = \sharp\ \{t \in I \mid t \text{ es } \pi_j^y - \text{tipo para } Y\}.$$

Una vez calculadas las frecuencias absolutas se pueden calcular las frecuencias relativas facilmente:

$$p(\pi_i^x) := p_{\pi_i^x} = \frac{\sharp\ \{t \in I \mid t \text{ es } \pi_i^x - \text{ tipo para } X\}}{|I| - m + 1} = \frac{n_{\pi_i^x}}{|I| - m + 1} \tag{4.1}$$

para el simbolo $pi_i^x$ en la serie $\{X_t\}_{t\in I}$ y

$$p(\pi_j^y) := p_{\pi_j^y} = \frac{\sharp\,\{t \in I \mid t \text{ es } \pi_j^y - \text{ tipo para } Y\}}{|I| - m + 1} = \frac{n_{\pi_j^y}}{|I| - m + 1} \qquad (4.2)$$

para el simbolo $pi_j^y$ en la serie $\{Y_t\}_{t\in I}$, donde $|I|$ representa el cardinal del conjunto $I$.

De manera similar, para $\eta_{ij} \in \mathcal{S}_m$ obtenemos que esta frecuencia relativa es

$$p(\eta_{ij}) := p_{\eta_{ij}} = \frac{\sharp\,\{t \in I \mid t \text{ es } \eta_{ij} - \text{tipo}\}}{|I| - m + 1} = \frac{n_{\eta_{ij}}}{|I| - m + 1} \qquad (4.3)$$

Una vez introducida la notación, los símbolos y los conceptos necesarios nos encontramos en condiciones de introducir la definición de *entropía de permutación* para una serie temporal 2-dimensional $\{\mathbf{W}_t\}_{t\in I}$ para una dimensión de embebimiento $m \geq 2$. Esta entropía está definida como la entropía de *Shannon* para los $m!^2$ símbolos distintos, de la forma siguiente:

$$h_{\mathbf{W}}(m) = -\sum_{\eta\in\mathcal{S}_m} p_\eta \log(p_\eta). \qquad (4.4)$$

De manera similar se pueden definir las entropías marginales de permutación como

$$h_X(m) = -\sum_{\pi_i^x\in S_m} p_{\pi_i^x} \log(p_{\pi_i^x}) \quad \text{y} \quad h_Y(m) = -\sum_{\pi_j^y\in S_m} p_{\pi_j^y} \log(p_{\pi_j^y})$$

para $\{X_t\}$ y $\{Y_t\}$ respectivamente. La entropía de permutación, es pues un indicador de la información contenida en la serie temporal, obtenida comparando $m$ valores consecutivos de dicha serie temporal.

El siguiente paso en el presente capitulo es la construcción propiamente dicha del test de independencia entre series basado en la entropía de permutación y que se distribuye asintóticamente como una $\chi^2$. Además demostraremos que el test es consistente.

## 4.3. Construcción y propiedades del test de independencia

En este sección construiremos el test de independencia para series temporales, para lo que utilizaremos todas las herramientas definidas en la Sección 4.2.

Para ello consideremos $\{\mathbf{W}_t\}_{t\in I}$ una serie temporal 2-dimensional y $m$ una dimensión de embebimiento fija. Con el fin de construir el test de independencia entre las series $\{X_t\}_{t\in I}$ e $\{Y_t\}_{t\in I}$, comenzaremos considerando como hipótesis nula la siguiente:

$$H_0: \quad \{X_t\}_{t\in I} \quad \text{e} \quad \{Y_t\}_{t\in I} \quad \text{son independientes} \tag{4.5}$$

frente a cualquier otra alternativa.

Ahora para un símbolo $\eta \in \mathcal{S}_m$ definimos la variable aleatoria $Z_{\eta t}$ como sigue:

$$Z_{\eta t} = \begin{cases} 1 & \text{si } t \text{ es de } \eta\text{-tipo} \\ \\ 0 & \text{en caso contrario.} \end{cases} \tag{4.6}$$

Entonces $Z_{\eta t}$ es una variable Bernoulli con probabilidad de "exito" $p_\eta$, donde "exito" significa que $t$ es de $\eta$-tipo. Es sencillo ver que

$$\sum_{\eta \in \mathcal{S}_m} p_\eta = 1 \tag{4.7}$$

Ahora supondremos que el conjunto $I$ es finito y de orden $T$. Entonces estamos interesados en saber cuántos $t$'s son $\eta$-tipo para todo símbolo $\eta \in \mathcal{S}_m$. Llamemos $K = T - m + 1$. Con el fin de responder a la pregunta planteada anteriormente construimos la siguiente variable

$$Y_\eta = \sum_{t=1}^{K} Z_{\eta t} \tag{4.8}$$

La variable $Y_\eta$ así definida puede tomar los valores $\{0, 1, 2, \ldots, K\}$. Luego de ello se deduce que $Y_\eta$ puede ser aproximada a una variable aleatoria Binomial

$$Y_\eta \approx B(K, p_\eta). \tag{4.9}$$

Luego, bajo la hipótesis nula $H_0$, la función puntual de probabilidad conjunta de las $m!^2$ variables $(Y_{\eta_1}, Y_{\eta_2}, \ldots, Y_{\eta_{m!^2}})$ es:

$$P(Y_{\eta_{11}} = a_1, \ldots, Y_{\eta_{m!m!}} = a_{m!^2}) = \frac{(a_1 + a_2 + \cdots + a_{m!^2})!}{a_1!a_2! \cdot \cdots \cdot a_{m!^2}!} p_{\eta_{11}}^{a_1} p_{\eta_{12}}^{a_2} \cdots p_{\eta_{m!m!}}^{a_{m!^2}} \tag{4.10}$$

donde $a_1 + a_2 + \cdots + a_{m!^2} = K$. Como consecuencia la distribución conjunta de las $m!^2$ variables $(Y_{\eta_{11}}, Y_{\eta_{12}}, \ldots, Y_{\eta_{m!m!}})$ es una distribución multinomial.

La función de verosimilitud de la distribución (4.10) es:

$$L(p_{\eta_{11}}, p_{\eta_{12}}, \ldots, p_{\eta_{m!^2}}) = \frac{K!}{n_{\eta_{11}}!n_{\eta_{12}}! \cdot \cdots \cdot n_{\eta_{m!m!}}!} p_{\eta_{11}}^{n_{\eta_{11}}} p_{\eta_{12}}^{n_{\eta_{12}}} \cdots p_{\eta_{m!m!}}^{n_{\eta_{m!m!}}} \tag{4.11}$$

y puesto que $\sum_{i=1}^{m!^2} p_{\eta_i} = 1$ se sigue que $L(p_{\eta_{11}}, p_{\eta_{12}}, \ldots, p_{\eta_{m!m!}})$ tiene la siguiente expresión

$$\frac{K!}{n_{\eta_{11}}!n_{\eta_{12}}! \cdot \cdots \cdot n_{\eta_{m!m!}}!} p_{\eta_{11}}^{n_{\eta_{11}}} p_{\eta_{12}}^{n_{\eta_{12}}} \cdots (1 - p_{\eta_{11}} - p_{\eta_{12}} \cdots - p_{\eta_{m!^2-1}})^{n_{\eta_{m!m!}}} \tag{4.12}$$

y el logaritmo de esta función de verosimilitud $\log(L(p_{\eta_{11}}, p_{\eta_{12}}, \ldots, p_{\eta_{m!m!}}))$ es

$$\begin{aligned} &\log\left(\frac{K!}{n_{\eta_{11}}!n_{\eta_{12}}! \cdot \cdots \cdot n_{\eta_{m!m!}}!}\right) + \sum_{i=1}^{m!} \sum_{j=1}^{m!-1} n_{\eta_{ij}} \log(p_{\eta_{ij}}) \\ &\quad + n_{\eta_{m!m!}} Ln(1 - p_{\eta_{11}} - p_{\eta_{12}} - \cdots - p_{\eta_{m!(m!-1)}}). \end{aligned} \tag{4.13}$$

Con el fin de obtener los estimadores máximo verosímil $\widehat{p}_{\eta_{ij}}$ de $p_{\eta_{ij}}$ para todo $i, j = 1, 2, \ldots, m!$, resolvemos la siguiente ecuación

$$\frac{\partial \log(L(p_{\eta_{11}}, p_{\eta_{12}}, \ldots, p_{\eta_{m!m!}}))}{p_{\eta_{ij}}} = 0 \tag{4.14}$$

obteniendo que

$$\widehat{p}_{\eta_{ij}} = \frac{n_{\eta_{ij}}}{K}. \tag{4.15}$$

Entonces el estadístico ratio de verosimilitud es ver Lehmann (1986):

$$\begin{aligned}
\lambda(Y) &= \frac{\frac{K!}{n_{\eta_{11}}!n_{\eta_{12}}!\cdots\cdots n_{\eta_{m!m!}}!} p_{\eta_{11}}^{n_{\eta_{11}}} p_{\eta_{12}}^{n_{\eta_{12}}} \cdots p_{\eta_{m!m!}}^{n_{\eta_{m!n}}}}{\frac{K}{n_{\eta_{11}}!n_{\eta_{12}}!\cdots\cdots n_{\eta_{m!n}}!} \widehat{p}_{\eta_{11}}^{n_{\eta_{11}}} \widehat{p}_{\eta_{12}}^{n_{\eta_{12}}} \cdots \widehat{p}_{\eta_{m!m!}}^{n_{\eta_{m!m!}}}} = \frac{\prod_{i=1}^{m!} \prod_{j=1}^{m!} p_{\eta_{ij}}^{n_{\eta_{ij}}}}{\prod_{i=1}^{m!} \prod_{j=1}^{m!} \left(\frac{n_{\eta_{ij}}}{K}\right)^{n_{\eta_{ij}}}} = \\
&= K^{\sum_{i=1}^{m!} \sum_{j=1}^{m!} n_{\eta_{ij}}} \prod_{i=1}^{m!} \prod_{j=1}^{m!} \left(\frac{p_{\eta_{ij}}}{n_{\eta_{ij}}}\right)^{n_{\eta_{ij}}} = K^K \prod_{i=1}^{m!} \prod_{j=1}^{m!} \left(\frac{p_{\eta_{ij}}}{n_{\eta_{ij}}}\right)^{n_{\eta_{ij}}}.
\end{aligned} \tag{4.16}$$

Por otra parte, $\Lambda(m) = -2\log(\lambda(Y))$ sigue asintóticamente una distribución Chi-cuadrado con $(m! - 1)^2$ grados de libertad ver Lehmann (1986). Por lo tanto

$$\Lambda(m) = -2\log(\lambda(Y)) = -2[K\log(K) + \sum_{i=1}^{m!} \sum_{j=1}^{m!} n_{\eta_{ij}} \log\left(\frac{p_{\eta_{ij}}}{n_{\eta_{ij}}}\right)] \sim \chi^2_{(m!-1)^2} \tag{4.17}$$

Recordemos que $\sum_{i=1}^{m!} \sum_{j=1}^{m!} \frac{n_{\eta_{ij}}}{K} = 1$. Sea $\eta_{ij} = (\pi_i^x, \pi_j^y) \in \mathcal{S}_m$. Entonces, y bajo la hipótesis nula $H_0$ se satisface que

$$\frac{n_{\eta_{ij}}}{K} = p_{\eta_{ij}} = p_{\pi_i^x} p_{\pi_j^y} = \frac{n_{\pi_i^x}}{K} \frac{n_{\pi_j^y}}{K}.$$

Por lo tanto se deduce que

$$
\begin{aligned}
\Lambda(m) &= -2K\left[\log(K)+\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{p_{\eta_{ij}}}{n_{\eta_{ij}}}\right)\right] \\
&= -2K\left[\log(K)+\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{\frac{n_{\pi_i^x}}{K}\frac{n_{\pi_j^y}}{K}}{n_{\eta_{ij}}}\right)\right] \\
&= -2K\left[\log(K)+\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\left(\log\left(\frac{n_{\pi_i^x}}{K}\right)+\log\left(\frac{n_{\pi_j^y}}{K}\right)\right)\right. \\
&\quad \left.-\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{n_{\eta_{ij}}}{K}\right)-\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log(K)\right] \\
&= -2K\left[\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\left(\log\left(\frac{n_{\pi_i^x}}{K}\right)+\log\left(\frac{n_{\pi_j^y}}{K}\right)\right)-\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{n_{\eta_{ij}}}{K}\right)\right] \\
&= -2K\left[\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\pi_i^x}}{K}\log\left(\frac{n_{\pi_i^x}}{K}\right)\frac{n_{\pi_j^y}}{K}+\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\pi_j^y}}{K}\log\left(\frac{n_{\pi_j^y}}{K}\right)\frac{n_{\pi_i^x}}{K}\right. \\
&\quad \left.-\sum_{i=1}^{m!}\sum_{j=1}^{m!}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{n_{\eta_{ij}}}{K}\right)\right]
\end{aligned}
\tag{4.18}
$$

Recordemos que $\sum_{i=1}^{m!}\frac{n_{\pi_i^x}}{K}=\sum_{i=j}^{m!}\frac{n_{\pi_j^y}}{K}=1$. Entonces el estadistico, $\Lambda(m)$ permanece como:

$$
\begin{aligned}
\Lambda(m) &= -2K\left[\frac{n_{\pi_1^x}}{K}\log\left(\frac{n_{\pi_1^x}}{K}\right)\sum_{j=1}^{m!}\frac{n_{\pi_j^y}}{K}+\frac{n_{\pi_2^x}}{K}\log\left(\frac{n_{\pi_2^x}}{K}\right)\sum_{j=1}^{m!}\frac{n_{\pi_j^y}}{K}+\dots\right. \\
&\quad +\frac{n_{\pi_{m!}^x}}{K}\log\left(\frac{n_{\pi_{m!}^x}}{K}\right)\sum_{j=1}^{m!}\frac{n_{\pi_j^y}}{K}+\frac{n_{\pi_1^y}}{K}\log\left(\frac{n_{\pi_1^y}}{K}\right)\sum_{i=1}^{m!}\frac{n_{\pi_i^x}}{K} \\
&\quad +\frac{n_{\pi_2^y}}{K}\log\left(\frac{n_{\pi_2^y}}{K}\right)\sum_{i=1}^{m!}\frac{n_{\pi_i^x}}{K}+\dots+\frac{n_{\pi_{m!}^y}}{K}\log\left(\frac{n_{\pi_{m!}^y}}{K}\right)\sum_{i=1}^{m!}\frac{n_{\pi_i^x}}{K} \\
&\quad \left.-\sum_{j=1}^{m!^2}\frac{n_{\eta_{ij}}}{K}\log\left(\frac{n_{\eta_{ij}}}{K}\right)\right] = -2K\left[h_{\mathbf{W}}(m)-h_X(m)-h_Y(m)\right]
\end{aligned}
$$

Por tanto hemos demostrado el siguiente teorema.

**Teorema 4.3.1.** *Sea* $\{\mathbf{W}_t = (X_t, Y_t)\}_{t\in I}$ *una serie temporal* 2*-dimensional con* $|I| = T$. *Sean* $h_{\mathbf{W}}(m)$, $h_X(m)$ *y* $h_Y(m)$ *las entropías de permutación definidas en (4.4) para una dimensión de embebimiento fija* $m > 2$, *con* $m \in \mathbb{N}$. *Si las series temporales* $\{X_t\}_{t\in I}$ *y* $\{Y_t\}_{t\in I}$ *son independientes entonces el estadístico*

$$\Lambda(m) = 2(T - m + 1)\left[h_X(m) + h_Y(m) - h_{\mathbf{W}}(m)\right] \tag{4.19}$$

*sigue asintóticamente una distribución* $\chi^2_{(m!-1)^2}$.

Sea $\alpha$ un numero real con $0 \leq \alpha \leq 1$. Sea $\chi^2_\alpha$ tal que

$$P(\chi^2_{(m!-1)^2} > \chi^2_\alpha) = \alpha.$$

Entonces para contrastar

$$H_0 : \{X_t\}_{t\in I} \quad \text{e} \quad \{Y_t\}_{t\in I} \quad \text{son independientes}$$

la regla de decisión en la aplicación del test $\Lambda(m)$ con un nivel de confianza del $100(1-\alpha)\,\%$ es:

$$\begin{array}{ll} \textit{Si } 0 \leq \Lambda(m) \leq \chi^2_\alpha & \textit{Aceptamos} \quad H_0 \\ \text{en caso contrario} & \textit{Rechazamos} \quad H_0 \end{array} \tag{4.20}$$

El $\Lambda(m)$-test puede ser generalizado de forma natural para el caso de $n$ variables, $\mathbf{W} = (X_1, X_2, \ldots, X_n)$, como se puede ver en el siguiente corolario. Para su demostración bastará con seguir los pasos de la demostración del Teorema 4.3.1.

**Colorario 2.** *Sea* $\{\mathbf{W}_t = (X_{1t}, \ldots, X_{nt})\}_{t\in I}$ *una serie temporal* $n$*-dimensional con* $|I| = T$. *Sean* $h_{\mathbf{W}}(m)$ *y* $h_{X_i}(m)$ *las entropías de permutación definida en (4.4) para una dimensión de embebimiento fija* $m \geq 2$ *con* $m \in \mathbb{N}$. *Si las* *series temporales* $\{X_{it}\}_{t\in I}$ *son independientes para* $i = 1, 2, \ldots, n$ *entonces el estadístico*

$$\Lambda(m) = 2(T - m + 1)\left[\sum_{i=1}^{n} h_{X_i}(m) - h_{\mathbf{W}}(m)\right] \tag{4.21}$$

*sigue asintóticamente una distribución* $\chi^2_{m!^n-1-n(m!-1)}$.

Una vez construido el test, el siguiente paso lógico será comprobar la utilidad del mismo. Para ello tratamos de medir la consistencia del $\Lambda(m)$-test, de gran importancia, pues determina lo valioso que puede ser el uso de dicho test cuando el tamaño muestral aumenta.

### 4.3.1. Consistencia

A continuación probaremos que el test $\Lambda(m)$ es consistente. Se trata de una valiosa propiedad ya que el test rechazará asintóticamente la independencia entre procesos siempre que exista dependencia entre $\{X_t\}_{t\in I}$ y $\{Y_t\}_{t\in I}$. Representemos por $\widehat{\Lambda}(m)$ el estimador de $\Lambda(m)$.

**Teorema 4.3.2.** *Sea* $\{X_t\}_{t\in I}$ *e* $\{Y_t\}_{t\in I}$ *dos procesos estrictamente estacionarios, y* $m > 2$ *con* $m \in \mathbb{N}$. *Entonces bajo dependencia de orden* $\leq m$,

$$\lim_{T\to\infty} \Pr(\widehat{\Lambda}(m) > C) = 1$$

*para todo* $0 < C < \infty, C \in \mathbb{R}$.

*Demostración.* Primero notemos que el estimador $\widehat{h}(m) = -\sum_{\pi\in S_m} \widehat{p}_\pi \log(\widehat{p}_\pi)$ de $h(m)$, donde $\widehat{p}_\pi = \frac{n_\pi}{T-m+1}$, es consistente por que $p\lim_{T\to\infty} \widehat{p}_\pi = p_\pi$ para cualquier proceso estacionario ver Bandt, et al. (2002), y por tanto tenemos que

$$p\lim_{T\to\infty} \widehat{h}(m) = h(m). \tag{4.22}$$

Recordemos que $\widehat{\Lambda}(m) = 2(T-m+1)[\widehat{h}_X(m) + \widehat{h}_Y(m) - \widehat{h}_{\mathbf{W}}(m)]$. Obsérvese que

$$
\begin{aligned}
&-\sum_{i=1}^{m!} \tfrac{n_{\pi_i^x}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\right) - \sum_{j=1}^{m!} \tfrac{n_{\pi_j^y}}{K} \log\left(\tfrac{n_{\pi_j^y}}{K}\right) = \\
&-\sum_{i=1}^{m!} \left(\tfrac{n_{\eta_{i1}}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\right) + \cdots + \tfrac{n_{\eta_{im!}}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\right)\right) \\
&-\sum_{j=1}^{m!} \left(\tfrac{n_{\eta_{1j}}}{K} \log\left(\tfrac{n_{\pi_j^y}}{K}\right) + \cdots + \tfrac{n_{\eta_{m!j}}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\right)\right) \\
&= -\sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{ij}}}{K} \left(\log\left(\tfrac{n_{\pi_i^x}}{K}\right) + \log\left(\tfrac{n_{\pi_j^y}}{K}\right)\right) \\
&= -\sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{1j}}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\tfrac{n_{\pi_j^y}}{K}\right)
\end{aligned}
$$

y puesto que

$$
\widehat{h}_X(m) + \widehat{h}_Y(m) = -\sum_{i=1}^{m!}\sum_{j=1}^{m!} \frac{n_{\eta_{1j}}}{K} \log\left(\frac{n_{\pi_i^x}}{K}\frac{n_{\pi_j^y}}{K}\right) \tag{4.23}
$$

Por otro lado, para cualquier numero real positivo $x \in \mathbb{R}$ tenemos que $-\log(x) \geq 1 - x$. De hecho, para $x \neq 1$ tenemos que $-\log(x) > 1 - x$. Entonces, utilizando este hecho, que los procesos son estacionarios y (4.23) tenemos que bajo dependencia entre las series de orden $\leq m$

$$
\begin{aligned}
&\widehat{h}_X(m) + \widehat{h}_Y(m) \widehat{-h}_{\mathbf{W}}(m) = -\sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{1j}}}{K} \log\left(\tfrac{n_{\pi_i^x}}{K}\tfrac{n_{\pi_j^y}}{K}\right) \\
&+\sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{ij}}}{K} \log\left(\tfrac{n_{\eta_{ij}}}{K}\right) = \\
&= -\sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{ij}}}{K} \log\left(\frac{\frac{n_{\pi_i^x}}{K}\frac{n_{\pi_j^y}}{K}}{\frac{n_{\eta_{ij}}}{K}}\right) > \sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{ij}}}{K} (1 - \frac{\frac{n_{\pi_i^x}}{K}\frac{n_{\pi_j^y}}{K}}{\frac{n_{\eta_{ij}}}{K}}) \\
&= \sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\eta_{ij}}}{K} - \sum_{i=1}^{m!}\sum_{j=1}^{m!} \tfrac{n_{\pi_i^x}}{K}\tfrac{n_{\pi_j^y}}{K} = 1 - 1 = 0
\end{aligned}
$$

Por lo tanto hemos obtenido que bajo dependencia de orden $\leq m$

$$
\widehat{h}_X(m) + \widehat{h}_Y(m) - \widehat{h}_{\mathbf{W}}(m) > 0 \tag{4.24}
$$

Sea $0 < C < \infty$ con $C \in \mathbb{R}$ y tomemos $T$ suficientemente grande tal que

$$
\frac{C}{2(T-m+1)} < \widehat{h}_X(m) + \widehat{h}_Y(m) - \widehat{h}_{\mathbf{W}}(m). \tag{4.25}
$$

Entonces

$$\Pr[\widehat{\Lambda}(m) > C] = \Pr[2(T-m+1)\widehat{h}_X(m) + \widehat{h}_Y(m) - \widehat{h}_{\mathbf{W}}(m) > C] =$$
$$\Pr[\widehat{h}_X(m) + \widehat{h}_Y(m) - \widehat{h}_{\mathbf{W}}(m) > \tfrac{C}{2(T-m+1)}]$$

Por lo tanto, por(4.25) tendremos que

$$\lim_{T\to\infty} \Pr(\widehat{\Lambda}(m) > C) = 1$$

como queriamos demostrar. □

Como $\Lambda(m) \to +\infty$ bajo la hipótesis alternativa de dependencia de orden $\leq m$ entre los procesos, los valores críticos de la $\chi^2$ para el rechazo de la hipótesis nula se pueden usar. La propiedad de consistencia del test puede ser generalizada para el caso en que se consideren $n$ series temporales dependientes. Este hecho queda recogido en el siguiente corolario, cuya demostración es inmediata a partir de la del Teorema 4.3.2.

**Colorario 3.** *Sea* $\{\mathbf{W}_t = (X_{1t}, X_{2t}, \ldots, X_{nt})\}_{t\in I}$ *una serie $n$-dimensional estrictamente estacionaria con* $|I| = T$, *y* $m \geq 2$, $m \in \mathbb{N}$. *Entonces, bajo dependencia de orden* $\leq m$ *entre las seires* $\{X_{it}\}_{t\in I}$ *para* $i = 1, 2, \ldots, n$,

$$\lim_{T\to\infty} \Pr(\widehat{\Lambda}(m) > C) = 1$$

*para todo* $0 < C < \infty, C \in \mathbb{R}$.

Desde un punto de vista mas amplio, la utilización del nuevo test requiere seleccionar el parámetro $m$. El número de símbolos (eventos) aumenta con el valor de $m$ y, a continuación, la selección tiene que estar en relación con el tamaño de la muestra disponible. Para $m = 3$, el número de símbolos utilizados para el cálculo de $\Lambda(3)$ es 36, mientras que para $m = 4$ la prueba evalúa más de 576 posibles símbolos. Siguiendo Rohatgi (1976), para el test Chi-Cuadrado, el tamaño de la serie bivariante debe estar muy por encima del número de posibles eventos (símbolos), y por tanto, en general la recomendación es trabajar con muestras que contienen cerca de 10 veces el número de símbolos. Huelga decir que los requisitos del tamaño de la muestra son superiores si se aplica la versión multivariante del test ($n \geq 3$). En general, en las ciencias sociales son relativamente escasos los datos, en comparación con las ciencias naturales. Para las simulaciones realizadas, hemos considerado un conjunto de datos relativamente pequeño y por lo tanto, hemos fijado $m = 3$.

En la siguiente sección, estudiaremos el tamaño y la potencia del nuevo test realizando un ejercicio de simulación de MonteCarlo.

## 4.4. Comportamiento del $\Lambda(m)$-test en muestras finitas

Desde un punto de vista práctico, es natural investigar el comportamiento de las propiedades de nuevo test para muestras finitas y comparar los resultados con los de otros test disponibles. Además del estadístico $\Lambda(m)$, utilizaremos los estadísticos $S_M$ de Haugh (2.25) y $Q_n$ de Hong (2.26) en el proceso de simulación con el fin de comparar los resultados de potencia y de tamaño.

Con el fin de estudiar el comportamiento del tamaño y la potencia de los tests hemos utilizado los siguientes procesos de generación de datos (PGD):

$$\begin{aligned}
PGD\ 1 &: X_t = 0{,}5X_{t-1} + u_t & Y_t &= 0{,}5Y_{t-1} + v_t \\
PGD\ 2 &: X_t = u_t & Y_t &= X_{t-2} \\
PGD\ 3 &: X_t = 0{,}8X_{t-1} + u_t & Y_t &= 0{,}8Y_{t-1} + X_t + v_t \\
PGD\ 4 &: X_t = u_t & Y_t &= 0{,}8X_{t-1}^2 + v_t \\
PGD\ 5 &: X_t = 0{,}6X_{t-1} + u_t & Y_t &= 0{,}6u_{t-1}^2 + v_t
\end{aligned}$$

donde ambos $u_t$ y $v_t$ son i.i.d. $N(0,1)$ e independientes entre si.

Usamos PGD 1, para especificar dos procesos Gaussianos independientes AR (1), para determinar el tamaño de los tres estadísticos en muestras finitas. PGD 2 y PGD 3 son procesos lineales dependientes que han sido ya utilizados en alguna ocasión en condiciones ideales (normalidad y linealidad) para la aplicación de los estadísticos de Haugh (1976) y de Hong (1996). Finalmente consideramos el caso de no linealidad con los procesos PGD 4 y PGD 5.

Hemos contemplado tres tamaños muestrales $T = 500, 1000$ y 3000. Por cada $T$, generamos $T + 200$ observaciones y entonces descartamos las 200 primeras para reducir los efectos de los valores iniciales. En esta simulación hemos utilizado el núcleo de Daniell, es decir, $k(z) = \frac{\sin(\pi z)}{\pi z}, z \in \mathbb{R}$, para calcular el estadístico de Hong (1996). Además, también es necesario para calcular los test de Haugh (1976)[1] y el de Hong (1996) elegir el ratio $d$. A este respecto, consideramos $d = \lfloor \log n \rfloor$ ($\lfloor a \rfloor$ denota la parte entera de $a$). Como ya se ha comentado, hemos fijado m = 3 cuando utilizamos el nuevo test.

[1]Es decir M = $d$.

Tabla 1. Tasas de rechazo de 2000 repeticiones en virtud de PGD 1.

| | | $T = 500$ | $T = 1000$ | $T = 3000$ |
|---|---|---|---|---|
| 5 % | $\Lambda(m)$ | 5,3 | 5,1 | 4,9 |
| | $S_M$ | 4,5 | 4,8 | 5,3 |
| | $Q_n$ | 6,3 | 5,4 | 5,4 |
| 10 % | $\Lambda(m)$ | 9,4 | 10,1 | 9,7 |
| | $S_M$ | 9,1 | 9,9 | 10,0 |
| | $Q_n$ | 11,1 | 9,8 | 10,1 |

PGD 1 $X_t$=0,5$X_{t-1}$+$u_t$;$Y_t$=0,5$Y_{t-1}$+$v_t$.

Tabla 2. Tasas de rechazo de 2000 repeticiones al nivel del 5 % bajo cuatro hipótesis alternativas

| | PGD 2 | | | PGD 3 | | |
|---|---|---|---|---|---|---|
| | $T = 500$ | $T = 1000$ | $T = 3000$ | $T = 500$ | $T = 1000$ | $T = 3000$ |
| $S_M$ | 100 | 100 | 100 | 100 | 100 | 100 |
| $Q_m$ | 100 | 100 | 100 | 100 | 100 | 100 |
| $\Lambda(m)$ | 100 | 100 | 100 | 100 | 100 | 100 |

| | PGD 4 | | | PGD 5 | | |
|---|---|---|---|---|---|---|
| | $T = 500$ | $T = 1000$ | $T = 3000$ | $T = 500$ | $T = 1000$ | $T = 3000$ |
| $S_M$ | 10 | 11 | 37 | 14 | 15 | 17 |
| $Q_m$ | 23 | 28 | 27 | 31 | 34 | 33 |
| $\Lambda(m)$ | 39 | 80 | 100 | 50 | 76 | 100 |

Donde $u_t, v_t \sim N(0, 1)$

La Tabla 1 informa del tamaño de los tres test bajo la hipótesis nula (PGD 1). Para ello hemos realizado 2000 repeticiones. La principal observación es que para estos tamaños muestrales los tres test tiene un buen comportamiento, ajustandose los valores empiricos a los teóricos. Por otra parte, la Tabla 2 muestra el comportamiento empírico del test bajo varias hipótesis alternativas (ya sea lineal o no lineal). En relación con los procesos lineales (PGD 2 y 3), el estadístico de Haugh (1976) y el $\Lambda(3)$ se comportan muy bien con independencia del tamaño de la muestra. Sin embargo, el $\Lambda(3)-$test supera a los test de Haugh (1976) y de Hong (1996) cuando el proceso generador de datos es no lineal. De hecho, cuando el tamaño de la muestra aumenta, el test $\Lambda(3)$ aumenta rápidamente en potencia, mientras que los otros dos estadísticos no mejoran.

En resumen, hemos presentado un nuevo test para la dependencia entre dos series temporales estacionarias, que, en contraste con el test de Haugh (1976), evita pre-especificaciones tipo ARMA y la selección del núcleo estocástico. Además, es adecuado para cualquiera proceso estacionario lineal o no lineal y es consistente frente a cualquier alternativa de dependencia de orden $\leq m$.

El test y sus propiedades teóricas son fácilmente generalizables para detectar la dependencia entre $n-$variables distintas ($n \geq 3$). Simulaciones empíricas indican y ponen de relieve la utilidad general del test.

## 4.5. Programa en mathematica 5.2

En esta sección incluimos el programa utilizado para las simulaciones de MoteCarlo. El lenguaje de programación utilizado es Mathematica 5.2.

A modo de ejemplo ponemos la programación del proceso generador de datos PGD1:

**For[T=3000;**

**i=1;**

**X=0;**

```
Y=0;
aleatorioX=RandomArray[NormalDistribution[0,1],T+210];
aleatorioY=RandomArray[NormalDistribution[0,1],T+210];
DatosX=List[];
DatosY=List[];
i≤T+210,i++,
Xnew=0.5*X+Extract[aleatorioX,i];
Ynew=0.5*Y+Extract[aleatorioY,i];
Append[DatosX,Xnew];
Append[DatosY,Ynew];
X=Xnew;
Y=Ynew;
];
```

Una vez generadas las dos series DatosX y DatosY, presentamos la programación para determinar el valor del estadístico Chi-Cuadrado para una dimension de embebimiento m=3:

```
m=3;
T=3000;
Simbolos=Permutations[Table[i,{i,1,m}]];
cero=Table[0,{i,1,m}];
CCX=Table[0,{i,1,m!}];
CCY=Table[0,{i,1,m!}];
CC=Table[0,{i,1,m!^2}];
Array[CCpiX,m!];
Array[CCpiY,m!];
Array[CCpiXY,{m!,m!}];
For[j=1,j≤m!,j++,
```

```
CCpiY[j]=List[]];
For[
j=1,j<=m!,j++,
For[
i=201,i≤200+T-m+1,i++,
piY=Take[DatosY,{i,i+m-1}];
If[Ordering[piY]-Extract[Simbolos,j]==cero,
AppendTo[CCpiY[j],i]]
];
];
For[j=1,j≤m!,j++,
CCpiX[j]=List[]];
For[
j=1,j<=m!,j++,
For[
i=201,i≤200+T-m+1,i++,
piX=Take[DatosX,{i,i+m-1}];
If[Ordering[piX]-Extract[Simbolos,j]==cero,
AppendTo[CCpiX[j],i]]
];
];
For[l=1,l≤m!,l++,
For[j=1,j≤m!,j++,
CCpiXY[l,j]=List[]]
];
For[l=1,l≤m!,l++,
For[
```

```
j=1,j<=m!,j++,
For[
i=201,i≤200+T-m+1,i++,
piXX=Take[MMX,{i,i+m-1}];
piYY=Take[MMY,{i,i+m-1}];
If[{Ordering[piXX],Ordering[piYY]}-
{Extract[Simbolos,l],Extract[Simbolos,j]}== {cero,cero},
AppendTo[CCpiXY[l,j],i]]
]; ]; ]
For[ProbX=;ProbY=;ProbXY=;j=1,j≤m!,j++,
AppendTo[ProbX,1/(T-m+1)*Length[Extract[
Array[CCpiX,m!],j]]];
AppendTo[ProbY,1/(T-m+1)*Length[Extract[
Array[CCpiY,m!],j]]];
For[l=1,l≤m!,l++,
AppendTo[ProbXY,1/(T-m+1)*Length[Extract[
Array[CCpiXY,{m!,m!}],{j,l}]]]
] ];
LogProbX=Table[Piecewise[{{Log[Extract[ProbX,n]],
Extract[ProbX,n]≠0}},0],{n,1,m!}];
productoX=ProbX*LogProbX;
LogProbY=Table[Piecewise[{{Log[Extract[ProbY,n]],
Extract[ProbY,n]≠0}},0],{n,1,m!}];
productoY=ProbY*LogProbY;
EntropiaX=-N[m!*Mean[productoX]];
EntropiaY=-N[m!*Mean[productoY]];
LogProbXY=Table[Piecewise[{{Log[Extract[ProbXY,n]],
```

```
Extract[ProbXY,n]≠0}},0],{n,1,m!^2}];
    productoXY=ProbXY*LogProbXY;
    EntropiaXY=-N[m!^2*Mean[productoXY]];
    Chi2=-2(T-m+1)*(EntropiaXY-EntropiaX-EntropiaY)
```

## 4.6. Cuestiones abiertas

Acabamos este capítulo con una serie de cuestiones abiertas que creemos pueden ser estudiadas con las técnicas desarrolladas en el mismo.

1. Estudio de la Causalidad en procesos estocásticos, ya sean temporales o espaciales. Con el $\Lambda$-test somos capaces de determinar dependencia entre dos series temporales, el siguiente paso es determinar cual de las series es la causa de la dependencia.

2. Desarrollar nuevas simbolizaciones que sean capaces de detectar cambios estructurales y procesos heterocedásticos (como pueden ser los ARCH y GARCH) muy comunes en el estudio de las series financieras.

# Capítulo 5

# Relación entre entropía topológica y de permutación

## 5.1. Introducción

En este Capítulo estableceremos algunas relaciones entre la entropía topológica y la de permutación para funciones definidas en la recta real. También demostraremos algunas propiedades que satisface la entropía de permutación.

Comenzaremos recordando algunos conceptos básicos y las definiciones de *entropía Cánovas-Rodriguez* (2005) y de *entropia de permutación*.

**Definición 5.1.1.** *Sea $f : I \to I$ continua y sea $\pi$ una permutación del conjunto $S_n = \{0, 1, 2 \cdots , n-1\}$. Diremos que $\pi$ es $f$-**admisible** (o simplemente admisible si la relación con la aplicación $f$ esta clara dentro del contexto) si:*

$$P_\pi = \{x \in \mathbb{R} \mid f^{\pi(0)}(x) < f^{\pi(1)}(x) < \cdots f^{\pi(n-1)}(x)\} \neq \emptyset.$$

*Es decir, el conjunto de puntos cuyo segmento de orbita de longitud n está or-*

*denado según $\pi$.*

Llamaremos $P_n^*$ a la familia de todos los conjuntos no vacíos $P_\pi$ para $\pi$ una permutación del conjunto $S_n = \{0, 1, 2, \ldots, n-1\}$ $f$-admisible. Representaremos el cardinal de un conjunto por $|\cdot|$, y llamaremos a

$$h^*(f, n) = \frac{1}{n-1} \log | P_n^* |$$

entropía de permutación de f

El siguiente resultado pone de manifiesto la estrecha relación entropía topológica y de permutación. Mas concretamente el siguiente teorema, para el caso de funciones en el intervalo monótonas a trozos demuestra la igualdad entre las entropías.

**Teorema 5.1.2.** *(C. Bandt, G. Keller y B. Pompe 2002)*

*Si $I$ es un intervalo cerrado y $f : I \to I$ una función monótona a trozos entonces el limite $\lim_{n\to\infty} h^*(f, n)$ existe y es igual a la entropía topológica $h(f)$ de $f$.*

Posteriormente Misiurewicz (2003) construye un ejemplo de una función $f$ definida en un compacto invariante con un número infinito de trozos para la que la igualdad entre las entropías topológica y de permutación no se verifica.

También en Misiurewicz (2003) se introduce la definición de entropía $h^\sharp(f, n)$ como sigue. Sea $R_n(\pi) = \{x \in P_\pi |\ x = f^n(x)\}$ y denotemos por $\mathcal{R}_n^*$ la familia de todos los conjuntos $R_\pi$ no vacíos. Entonces se define

$$h^\sharp(f, n) = \frac{1}{n} \log |\mathcal{R}_n^*|$$

y por tanto tenemos que

$$h^\sharp(f) = \limsup_{n\to\infty} h^\sharp(n).$$

Además Misiurewicz (2003) demostró el siguiente resultado.

**Teorema 5.1.3.** *(Misiurewicz)(2003) Sea $f : I \to I$ una función monotona a trozos con un número finito de trozos, en un intervalo compacto $I$. Entonces*

$$\limsup_{n\to\infty} h^{\sharp}(n) = h(f)$$

En este contexto podemos extender la definición de entropía de *Cánovas-Rodriguez* $\text{ent}(f)$ al de entropía de permutación $\text{ent}^*(f)$ de funciones sobre la recta real como:

$$\text{ent}(f) = \sup\{h(f|_K) : K \in \mathcal{K}(\mathbb{R}, f)\}, \tag{5.1}$$

$$\text{ent}^*(f) = \lim_{n\to\infty} \frac{1}{n-1} \log \mid P_n^* \mid, \tag{5.2}$$

A continuación vamos a estudiar la relación existente entre ambas definiciones.

## 5.2. Relación entre entropía topológica y entropía de permutación

En esta sección daremos condiciones necesarias y suficientes en las que se verifica que $ent(f) = ent^*(f)$, para $f$ una función en la recta real.

### 5.2.1. f tiene un numero finito de trozos.

Por el Teorema 3.5.2 tenemos que $ent(f) = h(f_I)$ para $I = [a, b]$ un intervalo compacto tal que fuera de $I$ no existen compactos invariantes por $f$ que no sean puntos fijos, es decir,

$$(\mathbb{R} \setminus I) \cap (\mathcal{K}(\mathbb{R}, f) \setminus \mathcal{F}(f)) = \emptyset.$$

Más aún, también se verifica que $Orb_f(a), Orb_f(b) \subseteq I$. Por otro lado si $K \in \mathcal{K}(\mathbb{R}, f)$ y $\pi$ es una permutación admisible para $f_{|K}$ entonces también tenemos que $\pi$ es admisible para $f_I$. Luego $h^*(f_{|_K}) \leq h^*(f_I)$. Ahora, como $f_I$ es una función en un intervalo compacto con un número finito de trozos monótonos, tenemos que por Bandt, et al. (2002) se verifica que $h(f_I) = \lim_{n\to\infty} h^*(f_I, n)$ y por tanto $ent^*(f) \leq ent(f)$. Veamos ahora la desigualdad en sentido contrario.

Queremos demostrar que $ent(f) \leq ent^*(f)$. Por un lado sabemos que:

$$\limsup_{n\to\infty} h^\sharp(n, f_I) = h(f_I) = ent(f) = h^*(f_I)$$

donde $I = [a, b]$ y $Orb_f(a), Orb_f(b) \subseteq I$. Ahora vamos a demostrar que se verifica la siguiente desigualdad

$$\limsup_{n\to\infty} h^\sharp(n, f_I) \leq \limsup_{n\to\infty} h^\sharp(n, f) \tag{5.3}$$

Sea $x \in R_n(\pi, f)$. Entonces existen dos posibilidades:

1. $Orb_f(x) \subseteq I$, entonces $x \in R_n(\pi, f)$ y por tanto la desigualdad (5.3) se satisface.

2. $Orb_f(x) \nsubseteq I$ y por tanto existe $0 \leq k \leq n-1$ tal que $f^k(x) > b$ (ó $f^k(x) < a$) luego concluimos que $f_I^k(x) = b$ (ó $f_I^k(x) = a$). Entonces tenemos que $Orb_{f_I}(x) = Orb_{f_I}(b) = Orb_f(b) \subseteq I$, y por tanto $b \in R_n(\pi, f)$. Luego hemos demostrado que $|P_n^*(f_I)| \leq |P_n^*(f)|$ y por tanto

   $$ent(f) = h(f_I) = h^*(f_I) = \limsup_{n\to\infty} h^\sharp(n, f_I) \leq \limsup_{n\to\infty} h^\sharp(n, f) \leq ent^*(f)$$

   tal y como queriamos demostrar.

### 5.2.2. f no tiene un numero finito de trozos.

A continuación vamos a construir un ejemplo que pone de manifiesto que si $f$ no tiene un número finito de trozos $ent(f) \neq ent^*(f)$ tal como ocurrió para las funciones en el intervalo. Sea $f : I \to I$ la función definida en Misiurewicz, (2003), que cumple las siguientes condiciones:

P1) $f(0) = 0$ , $f(1) = 1$, y $f(0,1) \subseteq (0,1)$.

P2) $h(f) = 0$ y $h^*(f) > 0$.

P3) Para cada intervalo de la forma $[a,1]$, $a \in (0,1)$, la función $f|_{[a,1]}$ es monótona a trozos.

A partir de ella definimos una nueva función $\widetilde{f} : (0,1) \to (0,1)$ definida como $\widetilde{f}(x) = f(x)\ \forall x \in (0,1)$. Tomamos $\varphi : (0,1) \to \mathbb{R}$ un homeomorfismo y definimos $g : \mathbb{R} \to \mathbb{R}$ por $g(x) = (\varphi \circ \widetilde{f} \circ \varphi^{-1})(x)$. Aplicando la propiedad $P3)$ vemos que $g$ es monótona a trozos. Aplicando el teorema 3.3.1(a3), se tiene que $ent(g) = ent(\widetilde{f}) \leq h(f) = 0$, por lo que deducimos que $ent(g) = 0$.

Por otro lado, $\varphi$ preserva las permutaciones, tenemos que $ent^*(g) > 0$. Hemos pues construido un ejemplo que muestra como $ent(g) \lneq ent^*(g)$.

## 5.3. Funciones totalmente admisibles

Diremos que una función continua $f$ es totalmente admisible si cualquier permutación $\pi$ es $f$-admisible.

**Teorema 5.3.1.** *Existe una función $f : [0,1] \to [0,1]$ continua y totalmente admisible*

*Demostración.* Ahora vamos a construir una función $f$ en el intervalo $[0,1]$ que sea continua y totalmente admisible.

Sean $I_i = [\frac{1}{2i}, \frac{1}{2i-1}]$ y $J_i = [\frac{1}{2i+1}, \frac{1}{2i}]$. Para cada $i = 1, 2, \ldots$ sea $\pi_i = (i_0, i_1, i_2, \ldots, i_{n-1})$ una permutación en el grupo simétrico $S_n$. Consideremos ahora $n$ puntos $P_{i_0}, P_{i_1}, \ldots, P_{i_{n-1}} \in J_i$ satisfaciendo

$$\frac{1}{2i+1} = P_{i_0} < P_{i_1} < \ldots < P_{i_{n-1}} = \frac{1}{2i}.$$

En lo que resta de capítulo entenderemos que si para un cierto subíndice $j$ se verifica que $i_j = n-1$, entonces $i_j + 1 = 0$. Entonces definimos la aplicación $g_{ij}$ como la recta que une los puntos $(P_{i_j}, P_{i_j+1})$ y $(P_{i_{j+1}}, P_{i_{j+1}+1})$ es decir:

$$g_{ij}(x) = (\frac{P_{i_{j+1}+1} - P_{i_j+1}}{P_{i_{j+1}} - P_{i_j}})x + (\frac{P_{i_j+1}P_{i_{j+1}} - P_{i_j}P_{i_{j+1}+1}}{P_{i_{j+1}} - P_{i_j}}) \tag{5.4}$$

para $j = 0, 1, \ldots, n-2$.

Por otro lado, en los extremos de los intervalos $I_i$, la aplicación está definida por $g_{i(n-1)}$ y por $g_{(i-1)0}$ para $i > 1$. Entonces para $i > 1$, definimos la aplicación $h_i : I_i \to I_i$ como la aplicación afín uniendo $(\frac{1}{2i}, g_{i(n-1)}(\frac{1}{2i}))$ y $(\frac{1}{2i-1}, g_{(i-1)0}(\frac{1}{2i-1}))$, es decir

$$h_i(x) = g_{i(n-1)}(\frac{1}{2i}) + 2i(2i-1)(g_{(i-1)0}(\frac{1}{2i-1}) - g_{i(n-1)}(\frac{1}{2i}))(x - \frac{1}{2i}) \tag{5.5}$$

y definimos $h_1 : [\frac{1}{2}, 1] \to [\frac{1}{2}, 1]$ como la recta uniendo los puntos $(\frac{1}{2}, g_{1(n-1)}(\frac{1}{2}))$ y $(1, 1)$

$$h_1(x) = g_{1(n-1)}(\frac{1}{2}) + 2(1 - g_{1(n-1)}(\frac{1}{2}))(x - \frac{1}{2}) \tag{5.6}$$

Entonces definimos la función $f : [0,1] \to [0,1]$ como sigue

$$f(x) = \begin{cases} 0 & \text{si } x = 0 \\ g_{ij}(x) & \text{si } x \in [P_{ij}, P_{i(j+1)}] \\ h_i(x) & \text{si } x \in I_i \end{cases} \tag{5.7}$$

Obsérvese que para la aplicación $f$ todas las permutaciones son admisibles, y por tanto $h^*(f) = \infty$. $\square$

Construyamos ahora en el intervalo $[0,1]$ una función continua de entropía de permutación infinita y no totalmente admisible. Nótese en primer lugar que la función $\widetilde{f}:[0,\frac{1}{2}]\to[0,\frac{1}{2}]$ definida por $\widetilde{f}(x)=\frac{1}{2}f(2x)$ con $f$ como en (5.7) es totalmente admisible y por tanto $h^*(\widetilde{f})=\infty$.

Consideramos la siguiente función

$$\overline{f}(x)=\begin{cases}\widetilde{f}(x)+\frac{1}{2} & \text{si } 0\leq x\leq \frac{1}{2}\\ 1-x & \text{si } \frac{1}{2}\leq x\leq 1\end{cases}\tag{5.8}$$

Veamos si le falta alguna permutación. Para ello consideremos $x\in(0,\frac{1}{2})$. Entonces es claro que $\overline{f}(x)\in(\frac{1}{2},1)$ y que $\overline{f}^2(x)\in(0,\frac{1}{2})$ y por tanto tendremos que se presentan las siguientes posibilidades $x\leq\overline{f}^2(x)\leq\overline{f}(x)$ ó $\overline{f}^2(x)\leq$ $\leq x\leq\overline{f}(x)$ lo que origina que sólo se pueden dar las permutaciones $(0,2,1)$ ó $(2,0,1)$, respectivamente. Consideremos ahora el caso $x\in(\frac{1}{2},1)$. Entonces $\overline{f}(x)\in(0,\frac{1}{2})$ y $\overline{f}^2(x)\in(\frac{1}{2},1)$ por lo que $\overline{f}(x)\leq x\leq\overline{f}^2(x)$ ó $\overline{f}(x)\leq$ $\overline{f}^2(x)\leq x$ dando lugar a las siguientes permutaciones $(1,0,2)$ ó $(1,2,0)$, respectivamente. Resulta pues evidente que le faltan dos permutaciones a $\overline{f}$ para ser totalmente admisible.

Veamos a continuación que la entropía de permutación de $\overline{f}$ es infinita. Para ello veamos primero que $h^*(\overline{f}^2)=\infty$. En efecto,

$$\overline{f}^2_{|_{[0,\frac{1}{2}]}}(x)=\frac{1}{2}-\frac{1}{2}f(2x)=\frac{1}{2}-\widetilde{f}(x)$$

y por tanto $h^*(\overline{f}^2)\geq h^*(\overline{f}^2_{|_{[0,\frac{1}{2}]}})=h^*(\widetilde{f})=h^*(f)=\infty$ tal y como queriamos demostrar.

Por otro lado, sea $\pi$ una permutación de longitud $n$ de $\overline{f}^2$. Dicha permutación tiene asociada al menos una permutación $\overline{\pi}$ de longitud $2n-1$ de $\overline{f}$. Ahora como

$$h^*(\overline{f}^2)\geq h^*(\overline{f}^2_{|_{[0,\frac{1}{2}]}})=\lim_{n\to\infty}\frac{1}{n}\log\mid P_n^*\mid$$

tendremos que

$$\mid P_n^*(\overline{f}^2)\mid\leq\mid P_{2n-1}^*(\overline{f})\mid$$

luego

$$h^*(\overline{f}^2) = \lim_{n\to\infty} \frac{1}{n} \log \mid P_n^*(\overline{f}^2) \mid \leq 2 \lim_{n\to\infty} \frac{1}{2n-1} \log \mid P_{2n-1}^*(\overline{f}) \mid = 2h^*(\overline{f}).$$

Ahora como $h^*(\overline{f}^2) = \infty$ entonces tenemos que $h^*(\overline{f}) = \infty$. Por tanto hemos demostrado que $\overline{f}$ es una función no totalmente admisible y sin embargo tiene entropía de permutación infinita.

### 5.3.1. Entropía de permutación en aplicaciones bidimensionales

Otra propiedad que podemos resaltar de la entropía de permutación es la siguiente. Consideremos el cuadrado $[0,1] \times [0,1]$ ordenado según el orden lexicográfico. Entonces, de manera natural, para esta relación de orden se puede definir la entropía de permutación para aplicaciones definidas en $[0,1] \times [0,1]$.

Recordemos qué entendemos por orden lexicográfico. Consideremos ahora dos conjuntos $A$ y $B$ parcialmente ordenados por las relaciones $\leq_A$ y $\leq_B$, respectivamente. Entonces definimos un orden lexicográfico en el producto cartesiano $A \times B$ como una relación binaria $\leq$ de orden total definida como sigue:

Para todo $(a,b), (c,d) \in A \times B$ tenemos que

$$(a,b) \leq (c,d) \Leftrightarrow \begin{cases} a \leq_A c & \text{si } a \neq c \\ b \leq_B d & \text{si } a = c \end{cases}$$

Ahora consideremos una aplicación

$$F = (g \times s) : [0,1] \times [0,1] \to [0,1] \times [0,1]$$

definida por $F(x,y) = (g(x), s(y))$ donde $s$ y $g$ son funciones continuas en el intervalo $[0,1]$. Por tanto existen $x_0$ e $y_0$ puntos fijos de $g$ y $s$ respectivamente.

Por tanto si $x \in P_\pi(g)$ e $y \in P_\sigma(s)$ tenemos que $(x, y_0) \in P_\pi(F)$ y $(x_0, y) \in P_\sigma(F)$. Así que se verifica

$$h^*(F) \geq \text{máx}\{h^*(s), h^*(g)\}.$$

Sin embargo esta ultima desigualdad no es cierta si consideramos la aplicación

$$F = (g \times s) : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \times \mathbb{R}.$$

En efecto, consideremos como función $g$ aquella definida por $g(x) = 1 + x$. Obsérvese que esta función no tiene puntos fijos y tiene como única permutación admisible a $(0, 1, 2, \ldots, n-1)$. Tomemos como aplicación $s$ la aplicación $f$ definida en (5.7) extendida a todo $\mathbb{R}$ por una aplicación afín. Entonces tenemos que $ent^*(g) = 0$ y $ent^*(s) = \infty$. Sea ahora $\pi = (i_0, i_1, \ldots, i_{n-1}) \in S_n \setminus \{(0, 1, 2, \ldots, n-1)\}$. Supongamos que existe un $(x, y) \in P_\pi(F)$. Entonces tenemos que

$$(g^{i_0}(x), s^{i_0}(y)) < (g^{i_1}(x), s^{i_1}(y)) < \ldots < (g^{i_{n-1}}(x), s^{i_{n-1}}(y))$$

pero al ser $g$ estrictamente creciente y sin puntos fijos, tenemos que $\pi$ tiene que ser admisible para $g$ y por tanto $\pi = (0, 1, 2, \ldots, n-1)$, obteniendo una contradicción. Por tanto la única permutación admisible para $F$ es $(0, 1, 2, \ldots, n-1)$, luego

$$ent^*(F) = 0 \ngeq \text{máx}\{ent^*(g), ent^*(s)\} = \infty$$

tal como queríamos demostrar. Nótese que si cambiamos el orden de las funciones en el producto, es decir, tomamos $G = s \times g$, entonces $ent^*(G) = \infty = \text{máx}\{ent^*(g), ent^*(s)\}$, lo que prueba que no necesariamente $ent^*(g \times s) = ent^*(s \times g)$.

## 5.4. Cuestiones abiertas

1. Estudiar condiciones en $f$ para que la igualdad

$$ent(f) = ent^*(f)$$

   sea cierta para funciones $f : \mathbb{R} \to \mathbb{R}$ monótonas a trozos, en el sentido de que $f|_{[a,b]}$ es monótona a trozos para cada $[a, b] \subset \mathbb{R}$.

2. Estudiar las propiedades generales de $ent^*(f)$.

3. Encontrar funciones $f$ con $h(f) = \infty$ que sean totalmente admisibles.

# Bibliografía


[1] Adler, R. L., Konheim, A. G. and McAndrew, M. H., *Topological entropy,* Trans. Amer. Math. Soc. **114**, pp.309–319, 1965.

[2] Alsedà, Ll.,Llibre, J., and Misiurewicz, M., *Combinatorial dynamics and entropy in dimension one.* Second edition. Advanced Series in Nonlinear Dynamics, **5**, 2000.

[3] Anima, N., Kannan V. and Sesha, S.P., *Properties of topologically transitive maps on the real line,* Real Analysis Exchange. **27**, pp. 325–334, 2001.

[4] Anima, N., Kannan V. and Sesha, S.P., *Real Analysis Exchange*, Volume **27** (1), pp. 325-334, 2001.

[5] Arregui F., J., *Topológia.* Código UNED: 0108304UD01A02, 1988.

[6] Bandt, C. and Pompe, B., *Permutation entropy - a natural complexity measure for time series.* Physics Review Letters, **88**, 174102, 2002.

[7] Bandt, C., Keller, G. and Pompe, B., *Entropy of interval maps via permutations.* Nonlinearity **15** (5), pp. 1595–1602, 2002.

[8] Block, L., and Coppel, W., A., *Dynamics in one dimension* **1513**, Springer, Berlin, 1992.

[9] Bowen, R., *Topological entropy for noncompact sets*, Trans. Amer. Math. Soc. **184**, pp. 125–136, 1973.

[10] Box, G., and Jenkins, G., *Times series analysis. Forecasting and control.* Holden-Day, San Francisco, Calif.-London-Amsterdam 1970.

[11] Brucks K. and Bruin, H., *Topics from one-dimensional dynamics*, London Math. Soc. Student Texts **62**, 2004.

[12] Cánovas, J.S. and Rodríguez, J.M., *Topological entropy of maps on the real line Topology and its Applications*, **153**, (5-6), 1, pp. 735–746, 2005.

[13] Çinlar, E., *Introduction to stochastic processes.* Prentice-Hall, Inc. Englewood Cliffs, New Jersey. 1975.

[14] Dai, X., and Jiang, Y., *Distance entropy of dynamical systems on noncompact-phase spaces.* Discrete Contin. Dyn. Syst. **20** (2), pp. 313–333, 2008.

[15] Durá, J., and López, J., *Fundamentos de Estadística* Ariel, S.A. Barcelona 1988.

[16] Fan, Y., and Linton, O., *Some higher order theory for a consistent non-parametric model specification test.* The Journal of Statistical Planning and Inference **109** (1-2), pp. 125-154, 1999.

[17] Geweke, J., *A comparison of tests of independence of two covariance stationary time series.* J. Am. Statist. Assoc. **76**, pp. 363–73, 1981.

[18] Hao, B. and Zheng, W., *Applied Symbolic Dynamics and Chaos.* World Scientific, Singapore. 1998.

[19] Harry, J., *Estimation of entropy and other functionals of a multivariate density.* Ann. Inst. Statist. Math. **41** (4), pp. 683–697, 1989.

[20] Haugh, L.D., *Checking the independence of two covariance-stationary time series: a univariate residual cross correlation approach.* J. Am. Statist. Assoc. **71**, pp. 378–85, 1976.

[21] Gurevic, B. M. *Topological entropy of a countable Markov chain.* (Russian) Dokl. Akad. Nauk SSSR **187**, pp. 715–718, 1969 .

[22] Holmgren, R., *A first course in discrete dynamical systems.* Second edition. Universitext. Springer-Verlag, New York, 1996.

[23] Hong, Y., *Testing for Independence Between Two Covariance Stationary Time Series.* Biometrika, Vol. **83**, pp. 615-625, 1996.

[24] Koch, P. D. and Yang, S. S., *A method for testing the independence of two time series that accounts for a potential pattern in the cross-correlation function.* J. Amer. Statist. Assoc. **81**, pp. 533–544, 1986.

[25] Kolmogorov, A. N., and Tihomirov, V. M., *$\varepsilon$-entropy and $\varepsilon$-capacity of sets in functional space.* Amer. Math. Soc. Transl. **17** (2), pp. 277–364, 1961.

[26] Lehmann, E. L., *Testing statistical hypotheses.* Second edition. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986.

[27] Liu, L., and Wang, Y., Wei, G., *Topological entropy of continuous functions on topological spaces* Chaos, Solitons and Fractals,**39** (1), pp. 417 - 427, 2009.

[28] Matilla-García, M. and Ruiz, M., *A nonparametric independence test using permutation entropy.* Journal of Econometrics. **144**(1), pp. 139-155, 2008.

[29] Matilla-García, M., Rodríguez, J.M. and Ruiz M., *A Symbolic Test for Testing Independence Between Time Series"* en segunda revisión en Journal of Time Series Analysis (2009)

[30] Misiurewicz, M., *Permutations and topological entropy for interval maps.* Nonlinearity **16** (3), pp.971–976, 2003.

[31] Munkres, J., *Topology: a first course.* Prentice-Hall, Inc., Englewood Cliffs, N.J., **275**, 1975.

[32] Parry, W., *Intrinsic Markov Chains*, Trans. Amer. Math. Soc. **112**, pp. 55–66, 1964.

[33] Pham, D. T., Roy, R. and Cedras, L., *Tests for non-correlation of two cointegrated ARMA time series.* Journal of Time Series Analysis, **24**, 553-577, 2003.

[34] Pierce, A., *Lack of dependence among economic variables.* J. Am. Statist. Assoc. **72**, pp. 11–22, 1977.

[35] Rohatgi, V.K., *An Introduction to Probability Theory and Mathematical Statistics.* John Wiley & Sons, New York. 1976.

[36] Skaug, H. J. and Tjostheim D., *Nonparametric Tests of Serial Independence.* Developments in Time Series Analysis, pp. 207–229, Chapman & Hall, London, 1993.

[37] Walters, P., *An introduction to ergodic theory,* Springer Verlag 1982.