



UNIVERSIDAD VERACRUZANA

FACULTAD DE FÍSICA E INTELIGENCIA ARTIFICIAL

**Aquí va el título del trabajo, hay títulos que
tienen más de un renglón a veces dos e
incluso tres renglones**

Trabajo recepcional en la modalidad de:

TESIS, TESINA, MONOGRAFÍA, ETC

que como requisito parcial para obtener el título de:

Licenciado(a) en Física

P R E S E N T A

NOMBRE DEL ALUMNO

ASESOR(ES):

NOMBRE DEL ASESOR O ASESORES

Xalapa Enríquez, Veracruz

Mes año

Filtro de entropía aplicado a series de tiempo de mercados financieros

Versión a corrección

12 de abril de 2020

Version: 1

Versión a corrección

Filtro de entropía aplicado a series de tiempo de mercados financieros

Tesis, 12 de abril de 2020

Universidad Veracruzana

Facultad de Física

Gonzalo Aguirre Beltrán, Isleta

91060 and Xalapa

Resumen

El presente trabajo de tesis presenta la identificación de distribuciones estadísticas en datos de mercados financieros.

En el presente trabajo de tesis se presentan y discuten los resultados obtenidos tras aplicar un filtro de entropía a series de tiempo de mercados financieros.

Las técnicas de procesamiento de datos empleadas son básicamente dos, medias móviles y entropía. Las medias móviles son empleadas para poder interpretar de una mejor manera los datos tras aplicar un filtro de entropía ya que permiten ajustar una ventana temporal de los mercados. Los indicadores de precios al mercado utilizados fueron Dow Jones, DAX Performance, BMV IPC, y Nikkei 225.

Las medias móviles son mejor conocidas por los analistas de datos como Moving Average o MAV, la cual obtiene un promedio de un determinado conjunto de datos y esto a su vez permite aminorar el ruido de la serie de tiempo, para que al aplicar el filtro de entropía el resultado obtenido no manifieste un mayor número de máximos o un comportamiento mayormente uniforme.

La selección de las ventanas de tiempo se hace manualmente por el autor, las cuales se pueden realizar para cualquier ventana de tiempo bajo dos regímenes, el filtrado de entropía en los precios del mercado a través de un Moving Average, o el filtrado de entropía en los precios del mercado sin un Moving Average.

Se utilizó software que permitiera hacer los cálculos y un análisis más preciso en menor tiempo, esto permitió identificar que es posible ajustar distribuciones estadísticas a los resultados obtenidos para la distribuciones después de haber aplicado un filtro de entropía con o sin medias móviles.

Motivación

La búsqueda de avances en la ciencia y el progreso económico de las diferentes naciones viene muchas impulsado por intereses particulares, a ellos les resulta de suma importancia estar a la vanguardia implementando los avances científicos a la tecnología, a la industria y a la economía. Aunque esta última se ve directamente beneficiada por el progreso de las dos anteriores, existe una gran cantidad de personas interesadas en el estricto funcionamiento de la economía y su progreso, que pese a estar muchas veces sujeto a intereses que van desde gobiernos, hasta sencillamente individuos estudiosos de la materia, que de igual manera buscan implementar el conocimiento de Física, Estadística o Matemáticas como herramientas para alcanzar sus diferentes objetivos, lo cual resulta de suma importancia en estos tiempos, dado que ahora se cuenta con el poder computacional necesario como para probar teoremas matemáticos, o incluso diseñar algoritmos que automaticen una predicción en tiempo real sobre los mercados.

Objetivos generales

- Aplicar un primer filtro de entropía en diferentes mercados para identificar la distribución estadística que mejor se ajuste.
- Aplicar un segundo filtro de entropía en diferentes mercados tras utilizar medias móviles para identificar la distribución estadística.
- Obtener las fechas en la serie de tiempo que manifestaron menor entropía para realizar estudios posteriores de análisis predictivo..

Índice general

1	Introducción	1
2	Contexto teorico	3
2.1	La similitud entre economía y física	3
2.1.1	Economía neoclásica	4
2.1.2	Retornos	4
2.1.3	Propiedades de los retornos	5
2.2	Analogías Física-Economía	6
2.2.1	Ley de Boltzmann-Gibbs y distribución de la probabilidad del dinero (en equilibrio)	6
2.2.2	Segunda Ley de la termodinámica y la Segunda Ley de la econofísica	8
2.2.3	Entropía en la termodinámica y analogía en la econofísica	9
2.2.4	Entropía en la mecánica estadística	9
2.2.5	La ecuación de la entropía	10
2.3	Aplicaciones de la ecuación de Shannon para la entropía.	10
2.3.1	La entropía aplicada a la econofísica	11
3	Metodologia	13
3.1	Algoritmo para el calculo de entropia simple	13
3.1.1	Procesamiento de datos de precios de mercados	13
3.1.2	Calculo de retornos de precios de mercados	14
3.1.3	Calculo de la entropia	15
3.2	Calculo de entropía con retornos de medias moviles	15
3.3	Simulacion de un mercado eficiente	16
3.4	Ajuste de distribuciones estadisticas	16
3.5	Busqueda de minimos de entropia	18
4	Resultados	21
4.1	Resultado del cálculo de entropías simple	21
4.2	Resultado del cálculo de entropías con medias moviles	23
4.3	Simulacion de Mercado eficiente	23
4.4	Pruebas estadisticas	24
5	Conclusion	27

6	Conceptos	29
6.0.1	1.1.1. Econofísica	29
6.0.2	2.1.1. Economía	29
6.0.3	1.2. Agentes	29
6.0.4	1.3. Modelo de agentes	30
6.0.5	1.3.1. Modelo de agentes en econofísica	30
6.0.6	1.4. Sistema complejo	30
6.0.7	2.2. Precio	30
6.0.8	1.5. Precio en econofísica	30
6.0.9	2.3. Dinero	31
6.0.10	2.4. Economía financiera	31
6.0.11	2.5. Hechos estilizados	31
6.0.12	2.6. Mercado	31
6.0.13	2.7. Mercado financiero	32
6.0.14	1.6. Mercado financiero en econofísica	32
6.0.15	2.1.2. Economía en el contexto de econofísica	32
6.0.16	2.8. Microeconomía	32
6.0.17	2.9. Macroeconomía	32
6.0.18	1.7. Circuitos económicos	33
6.0.19	1.8. Media de una distribución	33
6.0.20	1.9. Mediana	34
6.0.21	1.10. Moda	34
6.0.22	1.11. Variable aleatoria	34
6.0.23	1.12. Varianza	34
6.0.24	1.13. Distribución normal	35
6.0.25	1.14. Estandarización	35
6.0.26	1.12. Proceso estacionario	35
6.0.27	1.9. Ruido blanco	36
6.0.28	1.13. Proceso estocástico	36
6.0.29	1.14. Serie de tiempo	37
6.0.30	2.10. Dos tipos de agentes	37
6.0.31	1.16. Tendencias o trends	38
6.0.32	1.17. Medias móviles	38
6.0.33	2.10.1. Agentes técnicos	38
6.0.34	1.16. Modelos fundamentales en econofísica	38
6.0.35	1.16.2. Ley de Boltzmann-Gibbs	38
6.0.36	1.16.3. Distribución de la probabilidad del dinero (en equilibrio)	39
6.0.37	2.11.1. La mano invisible	40
6.0.38	2.11.2. Teoría de juegos	40
6.0.39	2.11.3. Dilema del prisionero	40
6.0.40	2.11.4. Frecuencia	41
6.0.41	2.11.5. Fluctuación	41

6.0.42 2.11.6. Volatilidad	41
6.0.43 Mercado eficiente	41

Introducción

Se puede asumir que para tener un mayor desarrollo humano hay que poseer avances en la ciencia, con base en un entendimiento de las limitaciones que existen hoy en día, para así poder revolucionar la tecnología, de manera que al tener un nuevo avance tecnológico se dé pie a nuevos avances científicos, grosso modo es así como se genera un círculo virtuoso, no obstante el trasfondo es más meritorio que codicioso.

Si bien la economía mundial se mueve a través de intereses particulares de las diferentes naciones, empresas, casas de bolsa, mercados o incluso de los intereses propios de cada individuo, la Física no es precisamente motivada por los mismos intereses, ya que al tratarse de una ciencia natural, las leyes de la naturaleza y en particular las leyes físicas que describen la naturaleza del universo son asunto interés común, y es ahí donde al unir conocimientos de propiedad universal a la economía puede parecer incluso utópico.

Es bien sabido que el estudio de la naturaleza y todo lo que le conforma ha partido de los principios básicos del razonamiento humano, hasta llegar a comprender la dinámica de los objetos que hay alrededor de cada individuo, conforme se concibe dicha descripción física de un fenómeno se puede profundizar más en el comportamiento del fenómeno, gracias a un entero entendimiento del movimiento de los cuerpos en el universo se ha podido comprender la naturaleza del movimiento en cuerpos de mayor tamaño, como lo son planetas, también se ha podido ir un paso más allá y comenzar a interpretar fenómenos tales como el modelado de un sistema de partículas con carga, e incluso adentrarse en el mundo microscópico y modelar fenómenos a escalas que resultaban difíciles de imaginar hace un siglo, y lo que es más meritorio es que hoy en día es posible estudiar grandes cantidades de información y comprender su comportamiento.

En diferentes ámbitos, desde los microscópico-macroscópico, hasta la complejidad en el estudio de los modelos de agentes, la Física se ha encargado de explicar los fenómenos que competen a la naturaleza del universo, la Economía no es la excepción, puesto que es un fenómeno más en el que la humanidad se ve envuelta día a día, tal como lo es la termodinámica del medio en que los seres humanos habitamos el planeta, la economía es también algo de lo que nadie está exento de su sustancial

carácter en el desarrollo de nuestra especie.

Desde la teoría económica de Adam Smith en 1767 (Nordhaus, Samuelson, 2004), hecho que revolucionó a la economía, y es por ello que se le conoce como el padre de la economía moderna, puesto que al día de hoy mucho de su legado se sigue aplicando de una u otra manera, evidentemente con variantes pero existe, hasta el más reciente concepto de entropía (entropía de la información o de Shannon) se encuentra particular interés en la unión de dos ciencias que describen el funcionamiento de los sistemas en que nuestra especie se desenvuelve, ya que con las herramientas de una ciencia se puede explicar cualitativa y cuantitativamente un fenómeno que pareciera ser exclusivo de una sola ciencia.

Aunque no es estrictamente necesario entender los conceptos del área de economía, es fundamental entender que los mercados son sistemas formados por un número muy grande de componentes, o participantes, cada uno viendo por su propio interés, sin saber que al hacerlo están promoviendo un concepto conocido como *la mano invisible*. Al tratarse de un fenómeno donde se experimenta alta competitividad por parte de los otros internautas, y aún conociendo el funcionamiento de los indicadores de los mercados, o estando informado sobre los acontecimientos que influyen en las decisiones de los "traders", existe una gran incertidumbre sobre si es posible predecir una tendencia que sea favorable, esto debido al gran volumen de información que muestra una serie de tiempo financiera.

Con esto en mente esta tesis esta organizada de la siguiente manera:
En el capítulo 1 se introduce al tema de la econofísica donde se tratarán los fundamentos que conforman a esta nueva rama de estudio, y se mencionan conceptos básicos de economía.

En el capítulo 2 se presentan las características de los mercados estudiados, y se describe como funcionan los mercados con base en la perspectiva de la econofísica, planteando el objetivo general de este trabajo, así como los objetivos particulares.

En el capítulo 3 se explica el método empleado para la obtención de entropía de los mercados, se describe brevemente la manera de obtener la mínima entropía de las series de tiempo financieras.

Y finalmente en el capítulo 4 se presentan los resultados y conclusiones de este trabajo.

” *Un sistema físico macroscópico no puede ser descrito solo en términos de variables mecánicas, la llamada entropía era necesaria*

— P. Richmond, J. Mimkes, S. Hutzler
(2013)

La mecánica estadística y la economía son consideradas como dos campos de investigación diferentes, el primero que pertenece a la rama de la física y ciencias naturales, y la segunda a las ciencias sociales. Sin embargo, a partir del siglo XIX la evolución en la investigación de la economía permitió establecer analogías entre las dos ciencias. De este modo la economía puede ser estudiada desde el punto de vista de la física estadística.

En este capítulo se presenta al lector los conceptos esenciales para comprender las analogías Física-Economía aplicadas en esta Tesis. Se introducirán ciertos conceptos de Economía claves para la comprensión de esta Tesis. Del mismo modo, vamos a presentar los conceptos de mecánica estadística que son aplicados. En este Capítulo se introducirá al lector al estado del arte de la econofísica y sobretodo, la aplicación del concepto de entropía a la economía. Finalmente, se presentaran aplicaciones de la econofísica a las finanzas y mercados actuales.

2.1 La similitud entre economía y física

En *Física*, un sistema cerrado conserva la energía, en *Economía*, un sistema económico cerrado conserva el *dinero*. De la anterior analogía se entiende que no hay un flujo externo de dinero, por lo que la cantidad total de dinero se conserva.

La *mecánica estadística* estudiada por los físicos, y la economía, tienen en común que ambas estudian grandes ensambles; colecciones de átomos, y agentes económicos, respectivamente. (Cottrell et al., 2009, pagina 149).

De la anterior analogía, en una colección de átomos se asume que cada uno de ellos posee una energía cinética $\varepsilon_i \geq 0$, mientras que en los mercados financieros cada uno de los agentes participantes tiene un recurso llamado dinero el cual no puede ser negativo, es decir $d_i \geq 0$ (Cottrell et al., 2009, pagina 149).

En el contexto de la mecánica estadística se define un ensamble microcanónico como un conjunto en el que las características y el comportamiento de los compo-

nentes son similares, posee entropía máxima y conservación de energía. Gracias a su naturaleza física, es posible estudiarle como un todo.

Adicionalmente, en *Física*, a menudo se dice que un sistema se encuentra en equilibrio cuando la energía con que interactúan los componentes es conservada.

2.1.1 Economía neoclásica

En los comienzos de la *Econofísica* se encuentran emblemáticos personajes como Adolphe Quétlet (1796-1874), Léon Walras (1834-1910), Vilfredo Pareto (1848-1923), y Adam Smith (1723-1790), quienes ofrecieron un conjunto de ideas que permitieron ampliar la manera en que se estudia la *Economía*. Una de sus mayores contribuciones es considerar una transacción como una consecuencia determinista del intercambio de bienes entre dos agentes. Esta consideración es innovadora con respecto a las corrientes económicas y sociales del siglo XIX (e.g. aspectos sociales, factores políticos, a distribución de la riqueza).

Esta innovación permite a la *Econofísica* estudiar los sistemas complejos intrínsecos de la *Economía*. La econofísica entonces considera que un sistema complejo está compuesto de un gran número de *grados de libertad* en diferentes escalas de tiempo (Richmond et al., 2013, página 17).

La *Economía neoclásica* puede ser considerada como la corriente de pensamiento que da lugar a la *Econofísica*. Del mismo modo que Pareto estudió la distribución de la riqueza en el siglo XIX, Adam Smith propuso la descripción de un fenómeno llamado *la mano invisible*. Este y otros conceptos de economía pueden ser consultados en el capítulo ??.

Función de utilidad

Se utiliza para encontrar el número óptimo de diferentes profesionales en una compañía, la mejor elección de acciones de un mercado (Richmond et al., 2013).

2.1.2 Retornos

Los retornos son usualmente variables de análisis más rentables que el precio por sí, por varias razones. La distribución de los retornos es más simétrica y estable en el tiempo que la distribución de los precios en sí. La estructura de los retornos es cercana a estacionaria, mientras que la estructura de los precios observados en el tiempo no lo es (Gençay et al., 2001, página 17).

El logaritmo de los precios muestra que están normalmente distribuidos, al menos aproximadamente (Osborne, 1959). En esta Tesis se trabaja con los retornos de los registros diarios individuales de cada serie de tiempo financiera, en otras palabras,

la ventana temporal entre cada «tick» es constante y entonces el *lag* que se denota como δt también tiene la constante de un día.

Los retornos son expresados como:

$$R_t = P_{t+1} - P_t .$$

Es decir, el retorno del precio en cuestión (R_t) es equivalente al precio individual de su sucesor temporal (P_{t+1}) menos el precio del objeto temporal en turno (P_t), sin embargo, siguiendo los ideales de Osborne, hay que calcular el logaritmo de los retornos a sabiendas de que los precios se etiquetan como Z_t y son las variables de registro de las mediciones individuales diarias al cierre del mercado, donde $t \in \mathcal{T} = \{T_1, T_2, \dots, T_M\}$. El logaritmo del precio es:

$$Y_t = \ln Z_t .$$

Entonces los *retornos logarítmicos* son definidos como:

$$X_t = Y_t - Y_{t-1} .$$

Cuya interpretación es la diferencia sucesiva de los logaritmos de los precios registrados en el tiempo, en otras palabras, la diferencia del logaritmo del precio en turno y su antecesor temporal. Aunque en términos computacionales es mucho más sencillo trabajar con la siguiente expresión que respeta la proposición de Osborne (Slanina, 2013):

$$\ln R_t = \ln P_{t+1} - \ln P_t .$$

2.1.3 Propiedades de los retornos

Podemos citar algunas propiedades de los retornos propuestas por Cont (2001):

- Ausencia de auto correlacion: Las autocorrelaciones lineales no son significativas en la mayoría de los casos
- Colas largas: La distribución de retornos parece seguir una ley de potencias.
- Asimetría Ganancia/Perdida: Se observan bajas mas claramente que altas en los precios.
- Gaussioidad agregada (aggregational gaussianity): cuando se aumenta el intervalo de tiempo en el que se calculan los retorno se observa que el valor de los retornos sigue una distribución gaussiana.
- Intermitencia: los retornos muestran un gran grado de variabilidad.
- Asimetría en escalas de tiempo: a mayor intervalo de tiempo los retornos muestran volatilidad mas facilmente que cuando el intervalo de tiempo es menor.

Ademas de las propiedades listadas arriba, algunas definiciones de mercado eficiente implican que los retornos no son predecibles. Por ejemplo, Arthur Stalla-Bourdillon de la Banca de Francia en su artículo *Stock Return Predictability: comparing Macro- and Micro-Approaches* (Stalla-Bourdillon, 2022) da el siguiente ejemplo:

Dado que toda la informacion esta contenida en los precios del mercado, cambios en el mercado puede solo ser causados por la llegada de nueva informacion, que es por definicion, impredecible. En otras palabras, los precios deben seguir una caminata aleatoria; una regresion de retornos basada en informacion pasada no debe contener informacion predictiva.

La predictividad de los retornos asi como sus propiedades estadisticas son discutidas en profundidad por Pesaran (2010). En su articulo, Pesaran menciona que la predictibilidad de los retornos puede ser asociada a la volatilidad del mercado sugiriendo que la predictividad se incrementa en tiempos de crisis.

2.2 Analogías Física-Economía

2.2.1 Ley de Boltzmann-Gibbs y distribución de la probabilidad del dinero (en equilibrio)

La ley de la distribución de la probabilidad de la energía en la mecánica estadística:

$$P(\varepsilon) = C e^{-\varepsilon/T},$$

donde:

ε es la energía

C es una constante de normalización

T es la temperatura (Wannier, 1966)

Dada la similitud de un sistema descrito por la Ley de Boltzmann-Gibbs y las condiciones de un sistema económico cerrado, se puede esta ley a una descripción de la *distribución de la probabilidad del dinero* (Cottrell et al., 2009).

Distribución de la probabilidad del dinero (en equilibrio)

De acuerdo con Cottrel, Michelson y Wright (Cottrell et al., 2009) la distribución de la probabilidad del dinero sigue la forma de la *Ley de Boltzmann-Gibbs*

$$P(d) = C e^{-d/T},$$

Donde:

d es el dinero

C es una constante de normalización

T es el promedio de la cantidad de dinero por agente en el sistema económico.

Con base en el artículo *Price variations in a stock market with many agents* por Shubik, Pakzuski y Bak en 1997, la ley de la conservación del dinero establece que el dinero no puede ser manufacturado por agentes del sistema económico pero si puede ser transferido entre ellos. Esta descripción es análoga a la que un sistema conservativo presenta en Física, por ejemplo, cuando en un sistema cerrado sin intercambio de energía con el exterior, los átomos colisionan entre sí (Bak et al., 1997).

Ley de la conservación del dinero.

En relación con lo establecido por Shubik, Pakzuski y Bak en 1997 en el artículo ya mencionado, lo que compete en este trabajo de tesis es el resultado de la interacción entre los agentes i y j .

Sean i y j dos agentes que interactúan en el mercado, correspondiéndole a cada uno de ellos una cantidad finita de dinero d_i y d_j respectivamente, se puede denotar de la siguiente manera: $[d_i, d_j]$.

Si al llevarse a cabo la interacción entre dos agentes y el intercambio de dinero que llevan a cabo es constante, dicho intercambio se etiqueta como Δd . El resultado de la interacción de los agentes en cuestión se expresa como: $[d_i, d_j] \rightarrow [d'_i, d'_j] = [d_i - \Delta d, d_j + \Delta d]$. Dicho lo anterior puede notarse que $d_i + d_j = d'_i + d'_j$ lo cual significa que la cantidad total de dinero en la transacción es conservada si no existe un flujo externo de dinero que modifique la cantidad total d .

En tales condiciones se asume que la distribución de probabilidad del dinero en equilibrio es invariante pese a fuertes fluctuaciones Δd entre los agentes (Cottrell et al., 2009, pagina 149).

como descrito en (Huang & Raney-Yan, 2021).

2.2.2 Segunda Ley de la termodinámica y la Segunda Ley de la econofísica

Existe una formulación conocida por Kelvin-Planck que dice : *No hay proceso termodinámico en estado constante para el cual el calor es completamente convertido en trabajo* (Struchtrup, 2014, pagina 86).

Sin embargo, la primera derivación de la segunda ley fue dada por Rudolf Clausius (1822-1888) basada en el argumento de que la dirección en que se transfiere el calor esta restringida y depende estrictamente de las declaraciones en ciclos termodinámicos (Struchtrup, 2014, pagina 55). Del desarrollo de la segunda ley por Clausius se tiene que *el calor irá del calor al frío por sí mismo, y no al revés* (Struchtrup, 2014, pagina 64).

La segunda ley explica la restricción en la eficiencia y la dirección de los procesos, *el calor no puede ser completamente convertido en trabajo* (Struchtrup, 2014, pagina 5).

Lo anterior se ha definido en la literatura mediante la siguiente ecuación:

$$dS \geq \frac{\delta Q}{T}, \quad (2.1)$$

donde δ es la diferencial inexacta del calor $Q[J]$, entre la temperatura $T[K]$, mientras que dS es el diferencial de la entropía $S \frac{[J]}{[K]}$. Esta relación indica que la entropía es incrementada en procesos espontáneos pero no cambia en equilibrio (Keszei, 2011, pagina 340).

Segunda Ley de la econofísica

Por analogía con la sección anterior, se llama a S como S_e que es la entropía económica (Georgescu-Roegen, 1971), aunque S_e será una cantidad adimensional, mientras que la constante λ es un factor integrante cuya unidad de medida es la moneda según sea el caso (Richmond et al., 2013, pagina 166), y en vez de Q se ocupa M que en este caso es el dinero.

$$dS_e \geq \frac{\delta M}{\lambda} \quad (2.2)$$

Es preciso mencionar que en Economía S_e se le llama función de producción. En el presente trabajo de tesis no se determina el factor integrante, aunque existen métodos para obtenerlo, en termodinámica se utiliza la constante de Boltzmann debido a enlace entre temperatura y energía (Richmond et al., 2013, pagina 166).

2.2.3 Entropía en la termodinámica y analogía en la econofísica

Para Clausius en 1865 la entropía es una energía inutilizable que puede provenir de por ejemplo, de un motor de vapor que utiliza combustible, irremediamente una cantidad de energía no será aprovechada. Adicionalmente sostiene que la energía inutilizable en el universo o cualquier sistema cerrado tiende a incrementar (Cottrell et al., 2009, pagina 21).

La entropía es una cantidad que surge cuando se construye una desigualdad que describe la tendencia al equilibrio (Keszei, 2011, pagina 70).

En el contexto de enofísica, la entropie es la medida de un número total de microestados económicos accesibles o disponibles que pertenecen a un macroestado (mercado-estado fase .económico"). Aquí la analogía se encuentra en que, la energía mide la probabilidad de que un estado en particular en un espacio fase .económico" sea alcanzado. (pg.200, Richmond, Econophysics) (Cottrell et al., 2009).

2.2.4 Entropía en la mecánica estadística

Es la medida de un número total de microestados económicos accesibles o disponibles que pertenecen a un macroestado (mercado-estado fase .económico"), por

analogía, la energía mide la probabilidad de que un estado en particular en un espacio fase "económico" sea alcanzado (Richmond et al., 2013).

2.2.5 La ecuación de la entropía

La ecuación de Boltzmann para la entropía está definida como:

$$S = -k_B \int p(\Gamma, t) \ln p(\Gamma, t) d\Gamma. \quad (2.3)$$

En esta ecuación la entropía S es definida para moléculas en un espacio fase determinado en un macroestado termodinámico. En la ecuación 2.3 k_B es la constante de Boltzmann, $p(\Gamma, t)$ es la función de densidad de probabilidad en el tiempo de Γ . Γ engloba a las variables de posición y momento del espacio fase (Richmond et al., 2013, página 13). Cabe mencionar que un macroestado está relacionado con las variables temperatura, presión, volumen.

Por otro lado, en el estudio de microestados discretos contables y accesibles, la ecuación de la entropía es:

$$S = -k_B \sum p_r \ln p_r. \quad (2.4)$$

Donde p_r es la probabilidad de un microestado.

Otra definición de la entropía es dada por la ecuación de Shannon 2.5 En este trabajo de Tesis se define como la cantidad de pérdida y ganancia de información en un conjunto de datos.

$$H = - \sum_{i=1}^n p_i \log(p_i). \quad (2.5)$$

Donde p_i es la probabilidad del dato en cuestión.

2.3 Aplicaciones de la ecuación de Shannon para la entropía.

Aunque este trabajo de Tesis no está orientado a la inteligencia artificial (AI por sus siglas en inglés), resulta interesante que la AI aplica el concepto de entropía de la información. Por ejemplo (Dos Santos et al., 2012) utilizan el concepto de entropía para analizar la impureza de una base de datos. Ellos están interesados en analizar las edades de las personas y los alimentos de su preferencia. Su método aplica una reducción en la entropía antes de procesar la base de datos con un árbol de decisiones. Esto les permitió tener una ganancia de información (Dos Santos et al.,

2012). Al hacer una clasificación y etiquetar, obtuvieron los resultados que habían planteado en su hipótesis.

En este punto resulta interesante comprender el alcance que tiene la ecuación de Shannon, no solamente en la termodinámica. Por ejemplo, en la publicación *Entropy Calculation, Information Gain and Decision Tree Learning* del sitio web medium (Pranto, 2020) aplican la entropía para mejorar el modelo de árboles de decisiones. En inteligencia artificial un árbol de decisiones es frecuentemente utilizado como clasificador de bases de datos. Calcular la entropía permite tomar los atributos más importantes de la base de datos para la creación de un árbol de decisiones así como medir la homogeneidad de los datos. Por consecuencia, puede ayudar a determinar la calidad del árbol de decisiones.

En otra publicación hecha en el portal tds (towards data science) (Jeremiah, 2020), se utiliza el mismo concepto de un árbol de decisiones. En este caso se tiene como objetivo que el algoritmo de inteligencia artificial determine el impacto de un atributo con respecto a otro basado en la cantidad de información que presentan/

La ecuación de la entropía de Shannon no solamente se aplica para árboles de decisión, también se ha utilizado para mostrar que algunos textos poseen más riqueza en su contenido con respecto a un tema. Además se menciona que en comparación con las técnicas tradicionales del aprendizaje-máquina. La aplicación de la entropía de Shannon disminuye la tasa de falsos positivos en relación con las palabras clave que se refieren a un tema en particular (Chan, 2022).

2.3.1 La entropía aplicada a la econofísica

Una primera aproximación es considerar la entropía como un cálculo de estructuras complejas. Las estructuras con entropía alta tienen poca información que aportar, y las estructuras con baja entropía pueden ser consideradas determinísticas (Pincus & Kalman, 2004).

Un *mercado eficiente* es una estructura con máxima entropía, y entre más alta es la entropía del mercado más eficiente es el mercado.

Una segunda aproximación está dada a partir de la definición de Boltzmann, su aplicación en el contexto de econofísica es la medida del número total de estados económicos accesibles (Richmond et al., 2013).

Una tercera manera es la que se emplea en macroeconomía. La función de utilidad es conocida como función de producción S_e , en referencia a que se ha etiquetado a S como la entropía. La primera y segunda leyes de la econofísica se aplican directamente en este contexto, ya que la entropía captura información crucial. Por ejemplo, la función de producción S_e se obtiene de la cantidad de personas contratadas en diferentes ocupaciones y cuantas unidades producen (Richmond et al., 2013, página 170).

Una cuarta manera de acercarse a este concepto de la manera en que se aborda en esta Tesis es a través del uso que se le da la entropía como función de utilidad en econofísica. El concepto de función de utilidad se ha definido en la economía neoclásica. Entonces la entropía se utiliza considerando el número total de elementos de la característica 1, el número total de elementos de la característica 2, y así sucesivamente. En el contexto matemático el estudio de este concepto lleva a estudiar la maximización de la función de producción, ello implica estudiar la entropía a través de un Lagrangiano. En el contexto de la vida diaria, la maximización de la entropía, o bien, la maximización de la función de utilidad es el número óptimo de empleados para desarrollar una óptima producción.

Una quinta aplicación de la entropía en la econofísica es en el estudio del precio específico que se denota como ψ , el cual es el valor por hora del contenido laboral (Cottrell et al., 2009). La distribución de ψ permite que se pueda calcular la ecuación entropía de Shannon. De lo anterior se encuentra que con una desviación estándar pequeña hay una baja entropía, y cuando la desviación estándar es grande se tiene una entropía grande (Cottrell et al., 2009, página 189). **En esta parte hay que insertar la figura de esa página 189.**

Metodologia

En este capitulo se presenta el metodo y algoritmos desarrollados para calculo de la entropia en mercados financieros. La metodologia presentada tiene como objetivo ayudar a detectar si un mercado es *estable* en el tiempo a partir de la entropia calculada. Adicionalmente, el metodo propuesto puede ser utilizado para detectar momentos (fechas) en los que la entropia es minima. Esto se traduce como un intervalo de tiempo en el que el precio no cambia significativamente.

3.1 Algoritmo para el calculo de entropia simple

La metodologia implementada para el calculo de la entropia asi como el procesamiento de datos se presenta en el diagrama 3.1.

En este diagrama se identifican cuatro etapas principales:

- Carga de datos
- Pre-procesamiento de datos
- Calculo de retornos y cuartiles
- Calculo de entropia

Los algoritmos presentados en esta tesis fueron desarrollados en el language de programacion Wolfram Mathematica y Matlab.

3.1.1 Procesamiento de datos de precios de mercados

Los precios de los mercados utilizados en esta Tesis fueron obtenidos del portal Yahoo Finance. Cabe mencionar que los precios son registrados cada 24h excepto los dias sabado y domingo. Se descargaron las datos de precios entre DD/MM/YYYY y DD/MM/YYYY de los mercados AAA, BBB, CCC, y DDD.

Un ejemplo de la base de datos de precios descargada se presenta en la tabla ???. Las columnas de la base de datos corresponden a la fecha en la que se realizo el registro de precio respectivo. Las bases de datos de cada uno de los mercados fueron depuradas y ordenadas. Esta depuracion corresponde a identificar y eliminar las fechas en las que no hay precio registrado. De este modo se evitan errores de calculo.

Calculo de la entropía de Shannon para mercados

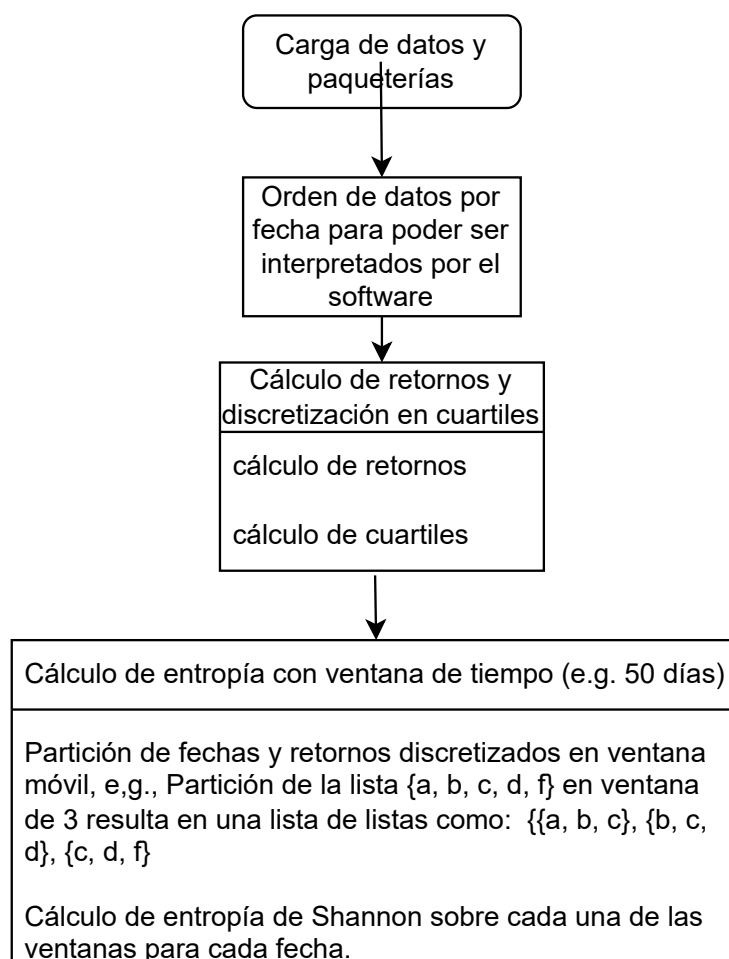


Figure 3.1: Diagrama del algoritmo utilizado para el cálculo de entropía de Shannon en mercados financieros.

3.1.2 Calculo de retornos de precios de mercados

Luego que se han cargado los datos, y se han ordenado es necesario calcular los retornos de los precios. Los retornos permiten analizar fácilmente la tasa de cambio en el tiempo. En esta Tesis se eligió trabajar con los retornos de los precios y no con los precios directamente debido a las propiedades de los retornos presentadas en la subsección 2.1.3. Aunque los retornos son una mejor representación de los precios, aún presentan picos que sobresalen de la media. Posteriormente se procede a discretizar los retornos mediante un proceso que divide en Q cuantiles a los retornos. Por ejemplo, la Tabla ?? muestra un ejemplo de discretización de los retornos en cuatro cuartiles. En este ejemplo a cada cuartil se le asigna la etiqueta 1, 2, 3, 4 dependiendo del intervalo al que corresponde el valor de retorno. De este modo, la

Fecha	precio
DD/MM/YYYY	
DD/MM/YYYY	
DD/MM/YYYY	

Tabla 3.1: Ejemplo de base de datos de mercados financieros.

Cuartiles	Valor de retorno	Etiqueta
Primer cuartil	$(-\infty, Q_2)$	1
Segundo cuartil	$[Q_2, Q_3)$	2
Tercer cuartil	$[Q_3, Q_4)$	3
Cuarto cuartil	$[Q_4, \infty)$	4

Tabla 3.2: Intervalos de cuartiles y etiquetas aplicadas a los retornos.

Tabla ?? muestra un ejemplo de etiquetado para los retornos calculados sobre una base de datos.

3.1.3 Calculo de la entropia

El calculo de la entropia se realizo de acuerdo a la ecuacion de Shannon 2.5. Para poder calcular la entropia se calculo la probabilidad de cada etiqueta sobre intervalos de tiempo glisantes de Δt_{ent}). Los retornos y etiquetas fueron calculados y asignados como se explico en la seccion anterior. Con los intervalos dados por los cuartiles y la discretización de los retornos se obtiene todo lo necesario para calcular la entropía. La ventana de tiempo glisante para el calculo de la entropía se corresponde a subconjuntos de datos agrupados en listas de N elementos. En la Tabla ?? se muestra un ejemplo de la entropia calculada para un conjunto de datos. La ventana glisante aplicada para el calculo de entropia en este ejemplo es de 50 dias.

3.2 Calculo de entropía con retornos de medias moviles

El metodo explicado en la Seccion anterior fue estudiado con una modificacion en el calculo de retornos. Dicha modificacion consiste en el calculo de la media del precio sobre una ventana glisante de tiempo Δt_{MAV}) (i.e., media movil) (ver Figura 3.2). El intervalo de la ventana glisante impacta directamente la curva de precios.

Fecha	retorno	cuartil
DD/MM/YYYY		1
DD/MM/YYYY		2
DD/MM/YYYY		4

Tabla 3.3: Ejemplo de etiquetado para una base de datos de mercados financieros.

Fecha	retorno	entropia
DD/MM/YYYY		N entropia
DD/MM/YYYY		N entropia
DD/MM/YYYY		N entropia

Tabla 3.4: Ejemplo de entropia calculada para una base de datos de mercados financieros.

De este modo se observa que una ventana glisante grande (e.g. 50 dias) tiene un efecto de suavizado de la curva mas claro que una ventana corta (e.g., 10 dias) ver Figura ??.

Una vez que las medias moviles de precio son calculadas, se calculan los retornos como dado por la Ecuacion 2.1.3. Despues se aplica el metodo descrito en la Seccion 3.1.3 para calcular la entropia. Las observaciones asi como conclusiones respecto a la aplicacion de medias moviles seran discutidas en profundidad en la seccion ??.

3.3 Simulacion de un mercado eficiente

Para fines de comparacion y validacion de la metodologia propuesta en las Secciones previas a mercados financieros reales se realizo una simulación de un mercado eficiente. Esta simulacion utiliza los datos del mercado real para calcular la media y desviación estándar de los retornos (no estandarizados) de los precios reales. Posteriormente se obtienen retornos simulados a los cuales se les asigna una fecha.

A partir de este punto se realizaron dos simulaciones que se pueden apreciar en el diagrama 3.3. La primera es que a dichos retornos simulados se les aplica el proceso para el cálculo de entropía de la figura 3.1. En la segunda se aplica el proceso de medias móviles como se presenta en 3.2.

La simulacion de mercado eficiente nos proporciona puede ser utilizada para calcular un umbral que permita seleccionar los valores de entropía mínima en los mercados reales.

3.4 Ajuste de distribuciones estadísticas

Para poder seleccionar los valores de minima entropia se aplico un ajuste de distribuciones estadísticas a los valores de entropia. Este ajuste es realizado utilizando el paquete *Distribution Fitter* de Matlab. Este paquete esta compuesto de una aplicacion que permite de ajustar de manera interactiva funciones de probabilidad. Las funciones de ajuste pueden tambien ser utilizadas fuera del ambiente interactivo.

Las distribuciones que seran ajustadas son:

- Normal
- Poisson

Calculo de la entropia de Shannon para Mercados con medias móviles

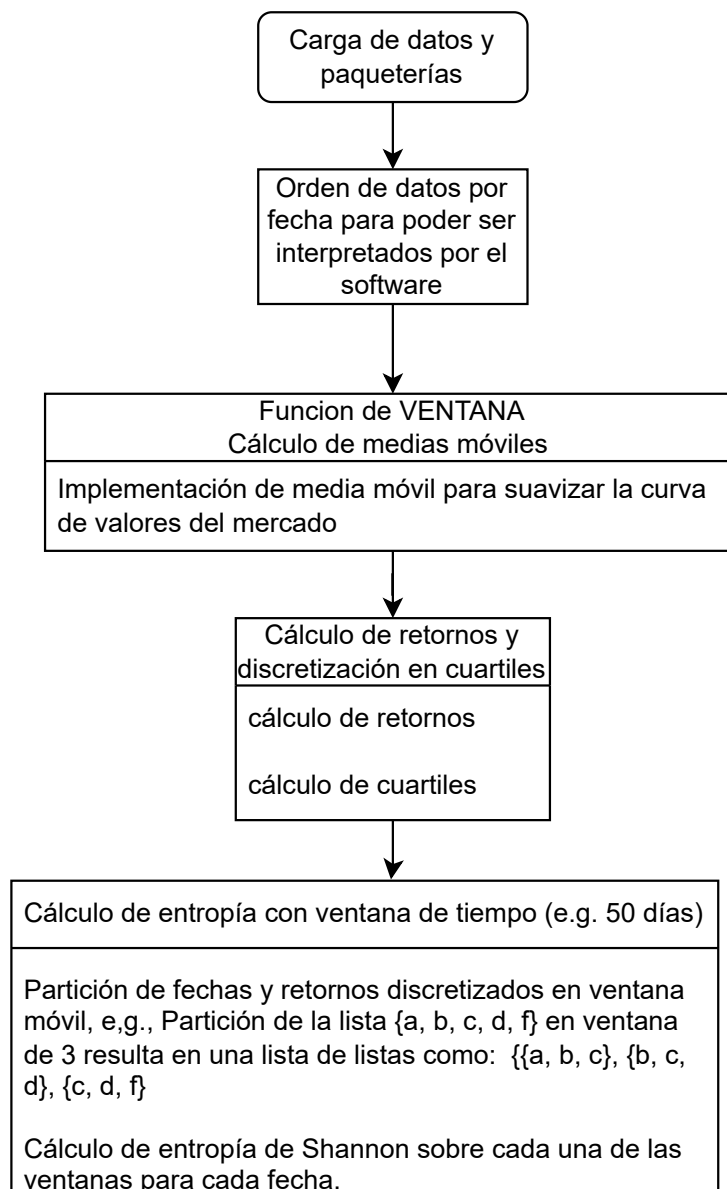


Figure 3.2: Diagrama del algoritmo utilizado para el cálculo de entropía de Shanon en mercados financieros con medias móviles.

Cálculo de la entropía de Shannon para Mercado eficiente en ventana de tiempo

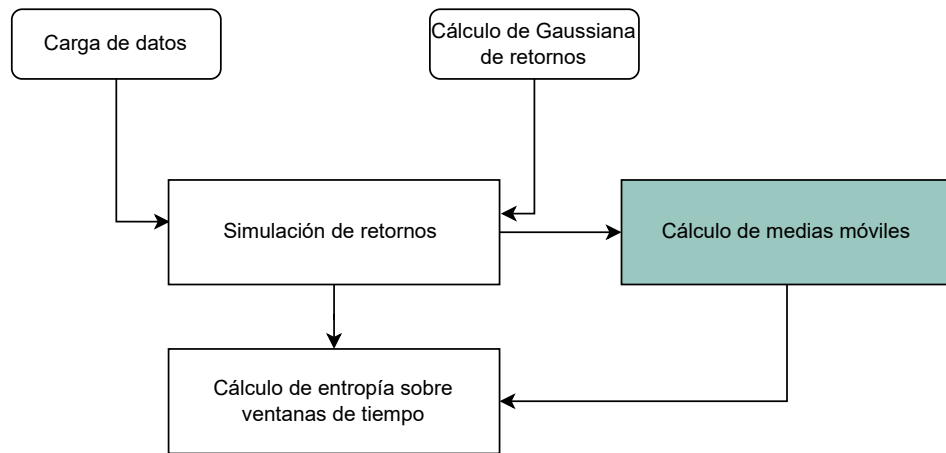


Figure 3.3: Diagrama del algoritmo de calculo de entropia para la simulacion de mercado eficiente.

- ttt
- ttt
- ttt

Por otro lado se aplicaron diferentes test para ajustar una distribucion normal a los valores de entropia calculados.

- Cramér-von Mises
- Kolmogorov-Smirnov
- Pearson X^2
- Anderson-Darling
- Watson U^2

3.5 Búsqueda de minimos de entropia

Uno de los objetivos de esta Tesis es buscar un metodo que permita la identificacion de minimos de entropia en los mercados financieros. Sin embargo, en la metodologia descrita en la Seccion 3.2 se observa que existen tres parametros dados por el usuario:

- la talla de ventana glisante Δt_{MAV} para suavizado de la curva de precios,

- la talla de ventana glisante Δt_{ent} para el calculo de la entropia y
- el numero de cuantiles Q sobre los cuales se calcula la entropia de Shannon.

Estos parametros actuan directamente sobre el calculo de la entropia de la siguiente manera: el valor de Δt_{MAV} tiene un impacto directo sobre el suavizado de la curva de precios, y como consecuencia, sobre la determinacion de los valores de entropia. Δt_{ent} tiene un impacto en el numero de muestras utilizadas en el calculo de la entropia, este valor impacta directamente el valor de la probabilidad de los valores de Q . Por otro lado, el valor de Q es el numero de categorias utilizadas para el calculo de la entropia. Un valor muy grande de Q puede diluir los minimos de entropia mientras que un valor muy pequeno puede resultar en una perdida de la informacion.

Para mejorar el calculo de la entropia se realizo una evaluacion de los valores de Δt_{MAV} , Δt_{ent} y Q . Esta evaluacion busca maximizar el numero de minimas entropias detectadas. Esta maximizacion se basa en el ajuste de las distribuciones de probabilidad listadas en la Seccion 3.4.

Resultados

En este capítulo se presentan los resultados de la implementación de la metodología presentada en el Capítulo 3. Los algoritmos fueron implementados en Mathematica y Matlab y fueron aplicados sobre las bases de datos de mercados financieros como descrito en la Sección 3.1.1. En las siguientes secciones presentaremos los resultados de la estimación de mínimos de entropía así como la optimización de los parámetros libres del algoritmo.

4.1 Resultado del cálculo de entropías simple

La primera etapa de este estudio consistió en aplicar el algoritmo descrito en la sección 3.1. En este caso se calculó la entropía sobre los cuatro mercados analizados.

En la Figura 4.1 se tienen los precios en el mercado y su evolución en el tiempo, tal como se muestran dichos precios no se les ha aplicado ningún filtro de media móvil ni ninguna suavización de la curva. El único proceso aplicado a los datos en esta etapa corresponde a la depuración de datos corruptos.

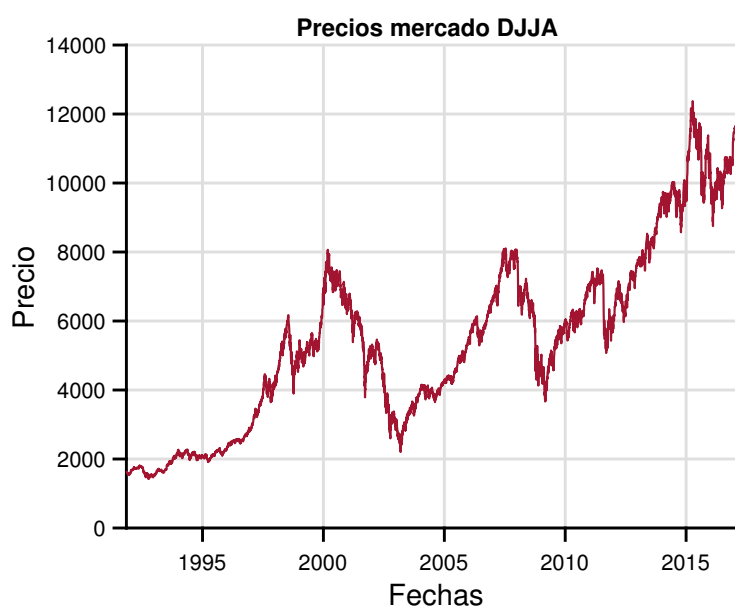


Figure 4.1: . Evolución temporal de los precios en el mercado DJJA.

A partir de los datos depurados y ordenados se calcularon los retornos como se explica en la sección 3. Se aplicó el algoritmo de la Figura 3.1. En la figura 4.2 se aprecia que los retornos fluctúan entorno a una media de valor cero. Sin embargo se aprecia que existen mínimos en el valor del retorno, como es el caso del segundo mínimo, que además es el mínimo global de todos los retornos.

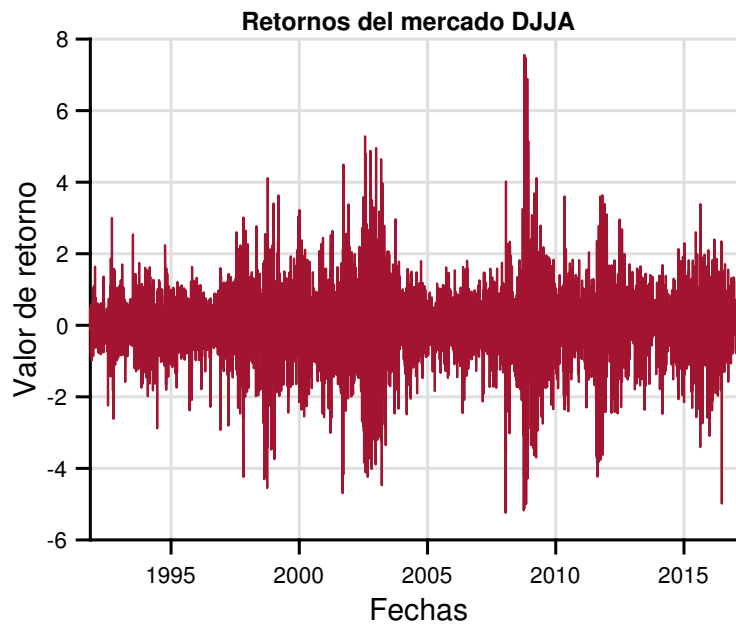


Figure 4.2: Retornos para el mercado DJJA.

Posteriormente se procedio a calcular la entropia de los mercados financieros aplicando el algoritmo descrito en la Seccion 3.1.3. En la figura 4.3 se presentan los resultados del calculo de entropia. Se observa que la entropía posee valores mínimos en la misma fecha (dd/mm/yyyy) del segundo mínimo que muestra el gráfico de los precios (Ver Figura 4.1). Un mínimo es observado cuando varias fechas consecutivas el precio es más bajo, resultando el valor de retorno negativo.

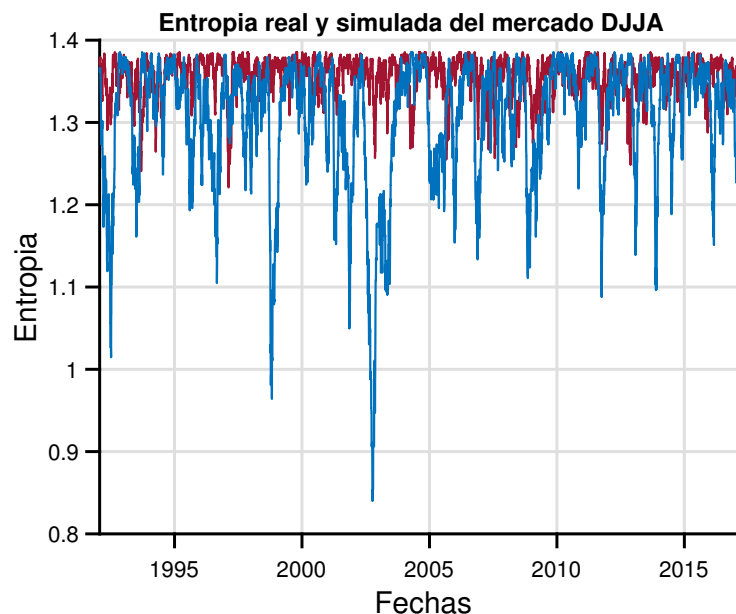


Figure 4.3: Entropía para un mercado simulado y un mercado eficiente. Mercado real DJJA (línea azul) y mercado eficiente DJJA (línea roja).

4.2 Resultado del cálculo de entropías con medias móviles

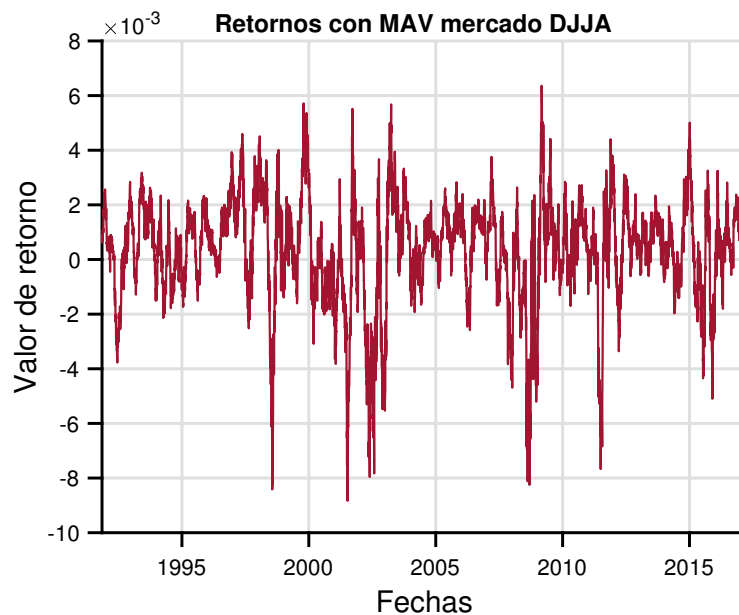


Figure 4.4: Retornos con una media móvil de 50 días aplicada para el mercado DJJA.

En la gráfica 4.4 se puede apreciar que la representación de los datos es menos conglomerada debido a que se aplicó una ventana de 50 días. Este filtrado de datos permite que la gráfica conserve su comportamiento pero que sea fácil identificar puntos de interés.

En la gráfica de comparación (Ver figura 4.5) entre la entropía del mercado real y la entropía del mercado simulado se tiene que cuando se simulan los retornos y se les aplica un filtro de media móvil, el cálculo de la entropía muestra que hay mínimos de la entropía con valor cero. Los retornos de datos reales muestran un comportamiento similar ya que hay más de un punto en que la entropía mínima también es cero.

4.3 Simulación de Mercado eficiente

Se observan valores mínimos de entropía gracias a la manera en que se presentan los resultados, con ello comparar el comportamiento de la entropía del mercado eficiente y del mercado real con fechas, con o sin media móvil requiere únicamente del cálculo de un umbral que diferencie los valores de entropía que se comportan de manera gaussiana de los que no.

Mediante la definición de un umbral que permite seleccionar con un intervalo de confianza de 95 por ciento los valores de entropía mínima en los mercados.

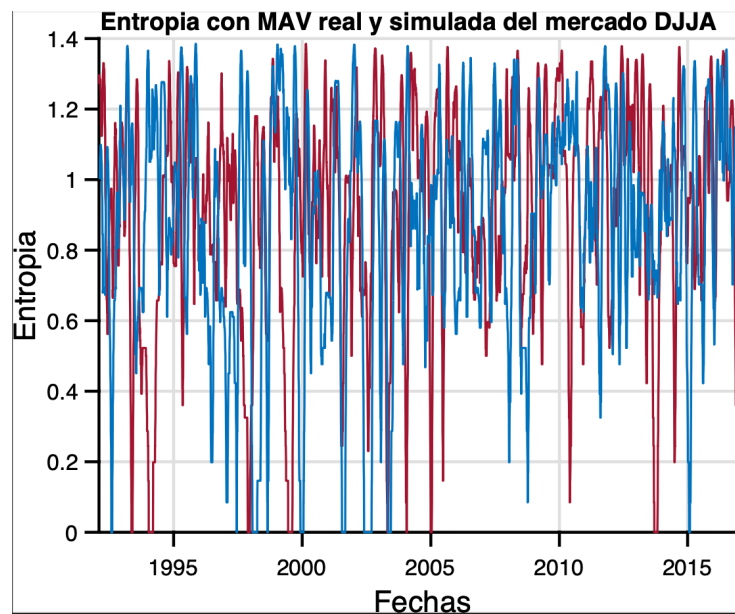


Figure 4.5: Entropía del mercado real DJJA (línea azul) y una simulación de un mercado ideal DJJA (línea roja) con aplicación de medias móviles.

4.4 Pruebas estadísticas

Finalmente, se aplica una prueba estadística de distribuciones a los valores con mínima entropía, tanto de aquellos que fueron tratados con un filtro de media móvil como aquellos que no.

Recordemos que de la ecuación de la entropía de Shannon, se realiza la suma de la probabilidad del estado, multiplicado por la probabilidad del estado. Gracias a que se trabajan con probabilidades es posible graficar un histograma de las entropías, y además se puede estudiar la entropía mediante cuantiles de manera análoga con los retornos de los precios.

Calcular el intervalo de confianza para la distribución de estas entropías sale de este trabajo de tesis, ya que no se asemeja ni a una distribución normal ni a una distribución de Student, ello implicaría tener que realizar interpolaciones entre los cuantiles obtenidos tal como se muestra en el artículo *A Method for Confidence Intervals of High Quantiles* de Mei Ling Huang, y Xiang Raney-Yan publicado en 2021.

Por lo anterior, se propone utilizar como umbral para hallar los valores mínimos de entropía el valor del percentil 95 en vez de un intervalo de confianza. Entre el año 2000 y 2005, se aprecia un valor mínimo en el precio, mismo que tiene impacto al estudiar la entropía.

Entonces, al realizar un gráfico que muestra la distribución de la entropía, y dado que la entropía es obtenida a partir de probabilidades, es fácil obtener cualquier cuantil (percentil 95 en el caso del intervalo de confianza), además que dicho

percentil permite diferenciar aquellos valores de la entropía que son ruido de aquellos que no.

Conclusion

5

” *Un sistema físico macroscópico no puede ser descrito solo en términos de variables mecánicas, la llamada entropía era necesaria*

— P. Richmond, J. Mimkes, S. Hutzler
(2013)

Los conceptos que se presentan a continuación son manifestados en el orden en que se tenga el mayor entendimiento de este trabajo de tesis, no debe confundirse la nomenclatura, ya que 1. se ocupa para todo aquello relacionado directamente con *econofísica* y 2. se ocupa para todo aquello relacionado con *Economía*.

6.0.1 1.1.1. Econofísica

Término acuñado por Eugene Stanley en 1997 para referirse a una rama de la Física (Roehner, 2002, página 3) que a partir de encontrar relaciones entre leyes físicas y variables macroscópicas propias de la Economía, busca un entendimiento más profundo de los fenómenos que en esta se presentan en los mercados.

6.0.2 2.1.1. Economía

Es el estudio de cómo las sociedades utilizan recursos escasos para producir bienes valiosos y distribuirlos entre diferentes personas (Samuelson & Nordhaus, 2009).

6.0.3 1.2. Agentes

Son los individuos, compañías, países, casas de bolsa, familias, comerciantes, inversionistas, capaces de interactuar entre sí de una manera no trivial, que además se encuentran inmersos en un mecanismo auto-organizado pese a poseer intereses particulares cada uno.

6.0.4 1.3. Modelo de agentes

Estructuras presentes en cualquier ciencia, mismas que pese a manifestar fenómenos diametralmente diferentes poseen datos que exhiben características similares tales como; un número N de agentes, la estadística no Gaussiana en cortos periodos de tiempo, adaptabilidad, autocorrelaciones, autointeracción, y la inaccesibilidad al análisis de un solo agente en particular.

6.0.5 1.3.1. Modelo de agentes en econofísica

Sistemas propuestos con base en una estructura Física, y en algunas ocasiones son propuestos únicamente de manera teórica, en los que se plantea que los N agentes (partículas) se rijan bajo una distribución que les permite la interacción entre ellos mismos y a su vez la conservación de la energía, si así fuese el caso.

6.0.6 1.4. Sistema complejo

Compuesto por un número muy grande de agentes, los elementos que le conforman poseen dinámica en sus interacciones entre sí, esto con base en los grados de libertad que el modelo en cuestión les permita, de tal manera que sus interacciones no son triviales, sin embargo, aunque en ocasiones pueda manifestarse como un sistema caótico, también contiene regularidades como lo es la auto-organización.

6.0.7 2.2. Precio

Cuota requerida para el proceso de adquisición, o para que se lleve a cabo un trato de concesión de un bien. Dependiendo del contexto en el modelo de agentes y del sistema complejo que se trate, puede ser una variable dinámica en la que haya intermediarios que le modifican directa e indirectamente.

6.0.8 1.5. Precio en econofísica

En este contexto encontramos que es la unidad en que se «mide» una variable dinámica conformada por N agentes de un respectivo modelo, mismos que disponen de una gama de posibilidades con respecto a los grados de libertad de dicho entorno en que se lleven a cabo los tratos de adquisición, o concesión de un bien (interacciones). Los N agentes causantes de dicha variabilidad pueden ser pertenecientes o no a dicho modelo, estas interacciones entre agentes internos o externos tienen una retroalimentación cuyo efecto es la principal causa de que su comportamiento sea el de un sistema complejo.

6.0.9 2.3. Dinero

Cada agente i posee una simbólica suma de recurso que utiliza para adquirir bienes o comodidades de consumo, su unidad de medida es la «moneda» aunque esta es una unidad arbitrariamente divisible. La moneda *per se* no puede ser producida ni consumida por los actores (Cockshitt W. P, Cottrell A. F., Michaelson G. J., Et Al, 2009). Los agente intercambian *dinero* y obtienen un incremento cuando llevan a cabo tratos de venta o cesión, y obtienen un detrimento cuando llevan a cabo un trato de adquisición o compra de un bien.

6.0.10 2.4. Economía financiera

Rama de la economía que analiza cómo los inversionistas racionales deben invertir sus bienes para alcanzar sus objetivos de la mejor manera posible (Samuelson & Nordhaus, 2005).

6.0.11 2.5. Hechos estilizados

Aunque no se profundiza en este concepto, vale la pena mencionar que este se refiere a que la experiencia recabada en econofísica relaciona una gran cantidad de datos de varias actividades que realiza la economía. Existen ciertas características universales intrínsecas (por no llamarles leyes) a los diferentes sistemas complejos que se traten, independientemente de las consecuencias del intercambio de bienes, o de la región en el planeta que uno se encuentre, o incluso dentro de una razonable brecha de observación hay independencia del tiempo también (Slanina, F. 2014).

Ejemplos de lo anteriormente descrito puede ser la cantidad de votos que tiene un candidato a la presidencia de un país, o la cantidad de gazapos que tiene una hembra conejo. Hay modelos que tienen por objetivo es explicar y predecir el comportamiento de fenómenos, en este caso, económicos, cuya tasa de acierto al ser elevada les permite ser relacionados con los llamados *hechos estilizados*, lo cual a su vez incrementa su grado de confianza.

6.0.12 2.6. Mercado

Un mercado es un mecanismo a través del cual compradores y vendedores interactúan para determinar precios e intercambiar bienes y servicios (Samuelson & Nordhaus, 2005).

6.0.13 2.7. Mercado financiero

Sistemas organizados para la compra-venta de acciones, divisas, opciones, bonos, derivados, etc. (Hernández Montoya, 2018).

Mercados cuyos productos o servicios son instrumentos financieros tales como acciones y bonos (Samuelson & Nordhaus, 2005).

6.0.14 1.6. Mercado financiero en econofísica

Sistemas conformados por un número muy grande de agentes que interactúan unos con otros, mismos que reaccionan a información tanto interna como externa a su entorno para determinar el mejor precio según sea el caso.

6.0.15 2.1.2. Economía en el contexto de econofísica

Le concierne el comportamiento de las personas así como las decisiones que tienen que tomar bajo debidas restricciones. (Hutzler, Mimkes & Richmond, 2013).

Trata las cuestiones que surgen generalmente como resultado intrínseco de estos mecanismos, es decir, la toma de decisiones de un sistema conformado por agentes que llevan a cabo tratos de inversión, compra o venta, dichas cuestiones son denominadas *endógenas*, mientras que las *exógenas* son todas aquellas que si bien no están inmersas en el mecanismo en cuestión, siguen formando parte del sistema complejo, las cuales incluyen desde noticias de alto impacto como lo son revoluciones, el inicio de una guerra, desastres naturales, o especulaciones en los mercados financieros.

Si bien, este trabajo de tesis no se centra en economía, si resulta de suma importancia saber a qué rama de la economía se está estudiando, por esa razón se define *grosso modo* la **macroeconomía** y la **microeconomía** a continuación.

6.0.16 2.8. Microeconomía

Estudia la naturaleza y el comportamiento dinámico de los elementos individuales, como consumidores, dueños de compañías, consejos de administración, quienes interactúan para formar, por ejemplo, precios (Hutzler, Mimkes & Richmond, 2013).

Análisis que explica el comportamiento de elementos individuales de una economía, tales como la determinación del precio de un solo producto o el comportamiento de un solo consumidor o empresa (Samuelson & Nordhaus, 2005).

6.0.17 2.9. Macroeconomía

Considera el comportamiento agregado de consumidores, dueños de compañías, consejos de administración, para describir cantidades como el ingreso doméstico bruto, y el nivel de empleo (Hutzler, Mimkes & Richmond, 2013).

Análisis que trata el comportamiento de la economía en su totalidad con respecto al producto, el ingreso, el nivel de precios, el comercio internacional, el desempleo, y otras variable económicas agregadas. (Samuelson & Nordhaus, 2005).

6.0.18 1.7. Circuitos económicos

No abundaremos mucho en este tema ya que escapa del objetivo principal de este trabajo de tesis, sin embargo es importante concebir a los bienes, derivados o acciones, y más concretamente al *dinero* como la **energía** de un sistema complejo en un contexto meramente financiero, a este tipo de estructuras se les conoce como *sistemas económicos*¹.

Un **circuito económico** es un sistema que puede ser modelado de diferentes maneras, el cual está compuesto por ciclos que pueden ser tan largos como sea permitida la entrada de energía (Hutzler, Mimkes & Richmond, 2013), tomando en cuenta que el balance de energía al final de cada ciclo nunca debe ser negativo. En este trabajo de tesis, se trata con *circuitos monetarios* cuya **energía** es el *precio* de cierre diario de los mercados.

Cabe destacar que conceptos como el de **esperanza** o **media** no han de sorprender, ya que la esperanza $E(x)$ es un concepto mejor conocido en Física por el nombre de *valor esperado*, de manera análoga y bajo este mismo contexto la media poblacional es la mejor conocida como μ . En aras de no escasear la información que solidifica este trabajo de tesis, se definen a continuación.

6.0.19 1.8. Media de una distribución

La media μ de una distribución de probabilidad $P(x)$ para un conjunto finito de valores numéricos x , también llamada *esperanza*, *valor esperado* o coloquialmente, *media*, es el promedio de los valores x multiplicados por las probabilidades de cada uno:

$$\mu = \sum_{all\ x} x P(x)$$

(Probability, Jim Pitman 1993, pg. 162, 163)

¹Véase Econophysics & physical economics, Hutzler, Mimkes & Richmond, 2013

Si bien el concepto anterior puede entenderse perfectamente para obtener la media de una distribución de valores discretos, también existe la definición de mediana para valores continuos. A partir de la distribución X , sea pues $f(x)$ una función de densidad para una variable, el valor esperado $E(X)$ a menudo denotado solo como μ o μ_x es:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

(Time Series Analysis, Cryer Jonathan, Chan Kung-Sik, 2008)

6.0.20 1.9. Mediana

Sea m un número que forma parte de la distribución de X , y sea x un valor numérico que forma parte de la distribución, la *mediana* es un número m tal que tanto la probabilidad de obtener un número x menor o igual que m y la probabilidad de obtener un número x mayor o igual que m es igual a $1/2$. (Probability, Jim Pitman 1993, pg. 165)

6.0.21 1.10. Moda

Sea X una distribución, entonces la *moda* es el valor más probable de X , y puede haber más de uno valor. (Probability, Jim Pitman 1993, pg. 165)

6.0.22 1.11. Variable aleatoria

Una variable aleatoria y es una función real evaluada definida en el espacio muestra Ω tal que para cada número real c , $A_c = \{ \omega \in \Omega \mid y(\omega) \leq c \} \in \mathcal{F}$, donde \mathcal{F} es una suma algebraica de eventos o subconjuntos de Ω . En otras palabras, A_c es un evento para el cual la probabilidad es definida en términos de P_r que es una medida de probabilidad definida en \mathcal{F} . Por completez, la función $F : \mathbb{R} \rightarrow [0,1]$, definida por $F(c) = P_r(A_c)$ es la función distribución de y . (New Introduction to Multiple Time Series ,Lütkepohl, H. 2005, pg. 2)

6.0.23 1.12. Varianza

Sea X una variable aleatoria. La varianza de X denotada como $Var(X)$ es el cociente de la división entre la raíz cuadrada de la diferencia entre el valor de la variable aleatoria X y la media μ del valor X al cuadrado, dividida entre el número de observables. Esto es:

$$\mu = \mu_x = E(X)$$

$$E[(X - \mu)^2]$$

(Probability, Jim Pitman 1993, pg. 185)

6.0.24 1.13. Distribución normal

Hasta ahora se ha tratado el concepto de distribución de manera intuitiva y se ha etiquetado como X , si bien, puede trabajarse solo con uno de los elementos que le componen, es decir, un elemento x o x_i , así como también puede ser tratada como un todo, o sea, cuando $X = x$, en cualquier caso, en el presente trabajo de tesis se trabaja con una *distribución normal*.

La **distribución normal** es una curva normal que se representa por encima del eje x , el área bajo dicha curva normal se rige por los parámetros de la media μ y desviación estándar σ .

$$y = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2}}$$

Donde $z = \frac{x-\mu}{\sigma}$ mide el número de desviaciones estándar desde la media μ hasta el número x . Esta distribución posee media cero y desviación estándar uno. (Probability, Jim Pitman 1993, pg. 94)

6.0.25 1.14. Estandarización

Proceso aplicado a cada uno de los elementos que pertenecen a la distribución cuyo resultado es la variable aleatoria $z = \frac{x-\mu}{\sigma}$ y como consecuencia se obtiene media cero y desviación estándar uno. (Time Series Analysis, Cryer Jonathan, Chan Kung-Sik, 2008)

6.0.26 1.12. Proceso estacionario

Se manifiesta cuando las configuraciones de un sistema permanecen inalteradas en el tiempo, y en *Estadística* el concepto es análogo, ya que lo define como «aquel en que las leyes de la probabilidad que gobiernan el comportamiento del proceso no cambian en el tiempo» (Chan & Crier, 2008).

6.0.27 1.9. Ruido blanco

6.0.28 1.13. Proceso estocástico

Se define por familias de variables aleatorias $X(t; \omega) \mid t \in \mathbb{T}$, para un conjunto indicado \mathbb{T} (Hassler, U. 2016) :

$$X : \mathbb{T} \times \Omega \rightarrow \mathbb{R} \quad (t; \omega) \mapsto X(t; \omega)$$

Cabe mencionar que $t \in \mathbb{T}$ en el contexto de este trabajo de tesis debe ser interpretado como el «tiempo», de tal manera que cuando uno ocupa específicamente un punto o dato en el tiempo t_o el proceso estocástico devuelve simplemente una variable aleatoria,

$$X : \Omega \rightarrow \mathbb{R} \quad \omega \mapsto X(t_o; \omega)$$

Es decir que se puede referir a un resultado ω_o en un conjunto, o una «trayectoria» realizada por el proceso (Hassler, U. 2016).

Por su parte, un proceso estocástico **estacionario** $x(t)$ es aquel proceso estocástico cuya función de densidad de probabilidad (FDP) $P[x(t)]$ es invariante bajo un cambio de tiempo (Mantegna R. & Stanley H., 2004).

En síntesis, es también un objeto *complejo*. Para caracterizarlo matemáticamente, se dispone de vectores aleatorios de longitud finita n con puntos arbitrarios en el tiempo $t_1 < \dots < t_n$, así, matemáticamente se considera²:

$$X(t_i) \equiv (X(t_1; \omega), \dots, X(t_n; \omega))' \quad , \quad t_1 < \dots < t_n$$

²(Hassler, U. 2016).

6.0.29 1.14. Serie de tiempo

Conjunto de variables medidas secuencialmente en el tiempo (Cowpertwait P. & Metcalfe A, 2009), registradas mediante intervalos de separación entre medición y medición. Su esencia es de un proceso estocástico ya que es fácil reconocer que las observaciones realizadas son impredecibles en el tiempo.

Las series de tiempo financieras conllevan una gran cantidad de información que no es redundante. La cantidad de información es tan grande que es difícil extraer un subconjunto de información económica asociada a algún aspecto (Mantegna R. & Stanley H., 2004).

Para entender o modelar mecanismos que puedan manifestar comportamientos estocásticos para predecir futuros valores con base en el historial que manifiesta, o con otras series de tiempo cuyos factores son relacionados (Chan & D. Cryer, 2008).

En otras palabras, es una secuencia de observaciones a una cantidad o cantidades. (Ruppert D. & Matteson D, 2015).

Al utilizar series de tiempo el objetivo es obtener lo que se conoce como «*ticks*» que en un sentido muy general representa a cualesquiera variables cotizadas de cualquier origen de cualquier instrumento financiero (Dacoragna M.M., Et Al, 2001), vale la pena mencionar que gracias al poder computacional es posible registrar cada evento (transacción o trato) minuto a minuto, de aquí que en el lenguaje coloquial se utiliza la expresión «tick a tick» o el también conocido en inglés *data-tick* (Hutzler, Mimkes & Richmond, 2013). La secuencia temporal de estos *ticks* es *inhomogénea* en caso de que los intervalos varíen en tamaño, naturalmente si la serie de tiempo cuyos *ticks* registrados son de manera que sus intervalos permanecen constantes, se le conoce como *homogénea*.

6.0.30 2.10. Dos tipos de agentes

Los mercados que atañen a este trabajo de tesis son conformados por dos tipos de agentes o inversores, los inversionistas *fundamentales*, y los inversionistas *técnicos* (Hutzler, Mimkes, Richmond, Econophysics & Physical economics, 2013), en obsequio a la brevedad, se define únicamente a los del segundo grupo, mismos que con la finalidad de no alejarse de la naturaleza de un *modelo de agentes*, se les presenta como *agentes técnicos*.

6.0.31 1.16. Tendencias o trends

Cambio en las series de tiempo que puede no ser periódico (Introductory Time Series with R, Cowpertwait, Paul S.P., Metcalfe, Andrew V, 2009, pg. 5)

6.0.32 1.17. Medias móviles

Sirven para indicar tendencias o "*trends*" en corto o mediano plazo dependiendo de la ventana de tiempo elegida. Mejor conocida como MAV, por su significado en inglés, "*Moving Average*" o simplemente "*MAV*".

Asumiendo una longitud de n días para la ventana de tiempo, el periodo en cuestión se etiqueta como t y el precio de cierre que registra cada uno de los días es C_t , la media móvil simple se define como sigue:

$$MAV = \left(\frac{1}{n}\right)(C_t + C_{t-2} + \dots C_{t-n+1})$$

Aunque suelen utilizarse con los precios de cierre, pueden ser utilizados con los precios en lo que abre un mercado, los precios más altos o los más bajos. (Mohammadali Mehralizadeh, 2017 pg.55)

6.0.33 2.10.1. Agentes técnicos

En Economía son mejor conocidos como "*traders*" técnicos o *inversionistas técnicos*. Son agentes que ignoran por completo los fundamentos en que se basan los "*traders*" fundamentalistas, cuyo criterio de decisión se basa en la apreciación numérica, matemática y estadística de los datos de un mercado.

El análisis que realizan estos agentes es, como su nombre lo dice, técnico, por lo que el precio, la duración de ese precio, el volumen de interacciones realizadas en un intervalo de tiempo, y las tendencias que se pueden observar en la serie de tiempo del mercado son solo unos de los muchos elementos en los que un agente técnico se basa para tomar decisiones, e incluso, predicciones.

6.0.34 1.16. Modelos fundamentales en econofísica

Debido a la relación que tiene la *Econofísica* con la *mecánica estadística*, es momento de sentar bases respecto a modelos como el de Boltzmann y Gibbs, dado que será de suma utilidad relacionar estas bases con el tema medular de este trabajo de tesis.

6.0.35 1.16.2. Ley de Boltzmann-Gibbs

La ley de la distribución de la probabilidad de la energía en la *mecánica estadística*:

$$P(\varepsilon) = C e^{-\varepsilon/T}$$

Donde:

ε es la energía

C es una constante de normalización

T es la temperatura (Wannier, 1966) (Cottrell, Michelson, Wright, Classical Econophysics, pg.149)

Dada la similitud de un sistema descrito por la Ley de Boltzmann-Gibbs y las condiciones de un sistema económico cerrado, se puede esta ley a una descripción de la *distribución de la probabilidad del dinero*. (Cottrell, Michelson, Wright, Classical Econophysics, pg.149)

6.0.36 1.16.3. Distribución de la probabilidad del dinero (en equilibrio)

De acuerdo con Cottrell, Michelson y Wright (Cottrell, Michelson, Wright, Classical Econophysics, pg.149) la distribución de la probabilidad del dinero sigue la forma de la *Ley de Boltzmann-Gibbs*

$$P(d) = C e^{-d/T}$$

Donde:

d es el dinero

C es una constante de normalización

T es el promedio de la cantidad de dinero por agente en el sistema económico.

Con base en el artículo *Price variations in a stock market with many agents* por Shubik, Pakzuski y Bak en 1997, la ley de la conservación del dinero establece que el dinero no puede ser manufacturado por agentes del sistema económico pero si puede ser transferido entre ellos. Esta descripción es análoga a la que un sistema conservativo presenta en Física, por ejemplo, cuando en un sistema cerrado sin intercambio de energía con el exterior, los átomos colisionan entre sí.

6.0.37 2.11.1. La mano invisible

Idea propuesta por Adam Smith en 1776, guiada por los ideales del *laissez-faire*, que significa "dejar hacer", la cual se refiere a que cuando cada participante busca solo su propio beneficio, a menudo promueve el beneficio de la sociedad de manera más eficaz que si realmente pretendiera promoverlo, como si una mano invisible benevolente estuviera dirigiendo todo el proceso. (Economía, Samuelson Nordhauss, pg.26)

6.0.38 2.11.2. Teoría de juegos

Estudio que puede ser descrito con el uso de las Matemáticas para comprender la competencia y cooperación entre varias partes involucradas . El rango de aplicación va desde estrategias de guerra hasta el entendimiento de la competición económica, desde problemas económicos o sociales hasta el comportamiento de los animales en situaciones de competencia(Peters, Hans. 2015, Springer, Game Theory).

Análisis de situaciones que comprende a dos o más tomadores de decisiones con intereses al menos parcialmente en conflicto. Se puede aplicar a la interacción de dos o más agentes en situaciones de negociación como las huelgas, o conflictos como los juegos y la guerra(Samuelson, Nordhauss, Economía).

6.0.39 2.11.3. Dilema del prisionero

Existen diferentes problemas que estudia la disciplina de la *teoría de juegos*, el que se presenta a continuación es probablemente el más conocido y el que mejor puede extrapolarse a este trabajo de tesis.

Sean dos prisioneros P1 y P2 que han cometido juntos un delito. El fiscal entrevista por separado a cada uno de ellos y les dice : "*Tengo suficientes pruebas sobre los dos para mandarlos un año a la cárcel. Pero haré un trato contigo: si solo confiesas tu, se te condenar a tres meses de cárcel, mientras que tu socio será condenado a 10 años. Si confiesan los dos, ambos serán condenados a cinco años*".

Lo importante en este caso es que cuando ambos prisioneros actúan interesadamente y confiesan, ambos serán condenados a mayores penas de cárcel. Solo son condenados a menos años de cárcel cuando ambos actúan de manera colusiva y altruísta. La situación en cuestión concluye que el equilibrio no cooperativo es ineficiente.

6.0.40 2.11.4. Frecuencia

Si bien, en *Física* la frecuencia es medida en **Hz** debido a la ecuación:

$$f = \frac{1}{T}$$

Donde:

T es el periodo y se mide en segundos.

En el contexto de este trabajo de tesis se entiende la *frecuencia* como una proporción cuantificable de registros de observables, datos, o precios en el tiempo (Jim Pitman, Probability, Springer, 1993).

6.0.41 2.11.5. Fluctuación

Observable registrada en el tiempo tan grande o tan pequeña conforme con el total de observables y la frecuencia de cada una.

6.0.42 2.11.6. Volatilidad

En *Economía* a menudo es calculada como la desviación estándar del cambio en el precio en una apropiada ventana de tiempo. (Mantegna, Econophysics).

Aunque se conocen la *volatilidad completa o histórica*, la *volatilidad implícita*, la *volatilidad instantánea o actual*, la *volatilidad de frontera* y la *volatilidad progresiva*, siendo esta última una de las utilizadas en predicciones. Esta tesis se centra específicamente en la primera.

Es una medida de aleatoriedad de un conjunto de registros para dos puntos cualesquiera en el tiempo (Wilmott, FAQ, Quantitative Finance).

Asumiendo que existe una ventana de tiempo de interés t_i La siguiente ecuación describe la volatilidad en cuestión:

$$v(t_i) = v(\Delta t, n, p; t_i) = \left[\frac{1}{n} \sum_{j=1}^n |r(\Delta t; t_{i-n+j})|^p \right]^{\frac{1}{p}}$$

Donde r son los retornos que se definen en la sección **retornos**, n es el número de la cantidad de observables utilizada para calcular los retornos, y p suele tomar el valor de 2 ya que v^2 es la varianza de los retornos. (An introduction to high-frequency finance, Dacoragna, Gencay, Muller, Olsen, Pictet 2001)

6.0.43 Mercado eficiente

Pesaran definition

At the core of the EMH lies the following three basic premises: 23 1. Investor rationality: It is assumed that investors are rational, in the sense that they correctly update their beliefs when new information is available. 2. Arbitrage: Individual investment

decisions satisfy the arbitrage condition, and trade decisions are made guided by the calculus of the subjective expected utility theory a la Savage. 3. Collective rationality: Differences in beliefs across investors cancel out in the market.

In his 1970 review, Fama distinguishes between three different forms of the EMH:

1. The weak form asserts that all price information is fully reacted in asset prices, in the sense that current price changes can not be predicted from past prices. This weak form was also introduced in an unpublished paper by Roberts (1967).
2. The semi-strong form that requires asset price changes to fully react all publicly available information and not only past prices.
3. The strong form that postulates that prices fully react information even if some investor or group of investors have monopolistic access to some information.

Índice de figuras

3.1	Diagrama del algoritmo utilizado para el cálculo de entropía de Shanon en mercados financieros.	14
3.2	Diagrama del algoritmo utilizado para el cálculo de entropía de Shanon en mercados financieros con medias móviles.	17
3.3	Diagrama del algoritmo de calculo de entropia para la simulacion de mercado eficiente.	18
4.1	. Evolución temporal de los precios en el mercado DJJA.	21
4.2	Retornos para el mercado DJJA.	22
4.3	Entropía para un mercado simulado y un mercado eficiente. Mercado real DJJA (linea azul) y mercado eficiente DJJA (linea roja).	22
4.4	Retornos con una media móvil de 50 días aplicada para el mercado DJJA.	23
4.5	Entropia del mercado real DJJA (linea azul) y un la simulacion de un mercado ideal DJJA (linea roja) con aplicación de medias móviles.	24

Índice de cuadros

3.1	Ejemplo de base de datos de mercados financieros.	15
3.2	Intervalos de cuartiles y etiquetas aplicadas a los retornos.	15
3.3	Ejemplo de etiquetado para una base de datos de mercados financieros.	15
3.4	Ejemplo de entropía calculada para una base de datos de mercados financieros.	16

Bibliografía

- Bak, P., Paczuski, M., & Shubik, M. 1997, *Physica A: Statistical Mechanics and its Applications*, 246, 430
- Chan, A. P. 2022, arXiv preprint arXiv:2206.02363
- Cont, R. 2001, *Quantitative Finance*, 1, 223
- Cottrell, A., Cockshott, P., Michaelson, G., Wright, I., & Yakovenko, V. 2009, *Classical Econophysics*, Routledge Advances in Experimental and Computable Economics (Taylor & Francis)
- Dos Santos, C. N., Milidiú, R. L., dos Santos, C. N., & Milidiú, R. L. 2012, *Entropy guided transformation learning* (Springer)
- Gençay, R., Dacorogna, M., Muller, U., Pictet, O., & Olsen, R. 2001, *An Introduction to High-Frequency Finance* (Elsevier Science)
- Huang, M. L. & Raney-Yan, X. 2021, *Entropy*, 23
- Jeremiah, L. 2020, *Entropy and Information Gain in Decision Trees*
- Keszei, E. 2011, *Chemical Thermodynamics: An Introduction* (Springer Berlin Heidelberg)
- Osborne, M. F. M. 1959, *Operations Research*, 7, 145
- Pesaran, H. 2010, Institute for the Study of Labor (IZA), IZA Discussion Papers
- Pincus, S. & Kalman, R. E. 2004, *Proceedings of the National Academy of Sciences*, 101, 13709
- Pranto, B. 2020, *Entropy Calculation, Information Gain and Decision Tree Learning*
- Richmond, P., Mimkes, J., & Hutzler, S. 2013, *Econophysics and Physical Economics*, Titolo collana (OUP Oxford)
- Roehner, B. 2002, *Patterns of Speculation: A Study in Observational Econophysics* (Cambridge University Press)
- Samuelson, P. & Nordhaus, W. 2009, *Economics* (McGraw-Hill Education)
- Slanina, F. 2013, *Essentials of Econophysics Modelling*, Titolo collana (OUP Oxford)

Stalla-Bourdillon, A. 2022, SSRN Electronic Journal

Struchtrup, H. 2014, Thermodynamics and Energy Conversion (Springer Berlin Heidelberg)