



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

**Vuelos Truncados de Lévy aplicados al estudio de índices de
mercado**

Julia Cassetti

Director: Dr. Pablo Amster

Diciembre de 2011

Índice general

1. Introducción	7
2. Modelización de activos	11
2.1. Procesos estocásticos	12
2.2. Clasificación de los procesos estocásticos	12
2.3. Función de distribución	14
2.4. Momentos de un proceso estocástico	14
2.5. Independencia y estacionaridad	15
2.6. Escala de tiempo o memoria del proceso	16
2.7. Movimiento Browniano	17
2.7.1. Movimiento browniano como límite de paseos al azar	17
2.7.2. Movimiento browniano geométrico	20
2.8. Movimiento Browniano Geométrico	21
3. Distribuciones de Lévy	25
3.1. Distribuciones estables	26
3.2. Vuelos Truncados de Lévy (TLF)	37
4. Estimación no paramétrica	41
5. Estimación de parámetros	51
5.1. Autosimilaridad	51
5.2. Estimación de los parámetros α y γ	52
5.3. Caso Núcleo Gaussiano y $h_{opt} = 1,06\sigma n^{-\frac{1}{5}}$	54
5.4. Una heurística para la elección del h_{opt}	56
6. Conclusiones	67
A. Apéndice estadístico	69
A.1. Algunas definiciones previas	69
A.2. Variables Aleatorias	71
A.2.1. Variables aleatorias discretas	71
A.2.2. Variables aleatorias continuas	72
A.3. Independencia	73
A.4. Variables aleatorias independientes y convolución	74

A.5. Esperanza matemática	75
A.6. Función característica	76
A.6.1. Estabilidad y función característica	77
A.7. Teorema Central del límite	79
A.8. Aproximación gaussiana a una variable aleatoria binomial	79

Resumen

La idea del trabajo es mostrar que el movimiento browniano no ajusta adecuadamente a algunos índices de mercado y presentar a las distribuciones de Lévy como aquellas que modelan la parte central de la distribución de dichos índices. Dado que estas distribuciones tienen varianza infinita para todos los casos salvo para la distribución normal, se proponen los Vuelos Truncados de Lévy para resolver este problema. Estas distribuciones coinciden con las distribuciones de Lévy en la parte central de los datos y luego se anulan o tienen decaimiento exponencial.

El primer capítulo es una introducción donde se presenta el problema y se cuenta la historia del mismo. En el segundo capítulo se presentan los conceptos de procesos estocásticos, movimiento browniano y movimiento browniano geométrico. Se explica el motivo por el cual se propuso este último como modelo para describir la evolución de los índices de mercado y se muestra que este modelo no ajusta bien a los índices de mercado que se estudian en este trabajo. En el tercer capítulo se presentan las distribuciones estables y las distribuciones de Lévy. Se identifican los parámetros de éstas distribuciones que serán necesario estimar y se presentan las propiedades que serán útiles para el objetivo de este trabajo: estimar los parámetros de las distribución de Lévy que mejor ajusta a los índices de mercado presentados. También se presentan los Vuelos Truncados de Lévy que son distribuciones cercanas a las de Lévy pero que difieren en las colas de la distribución. En el capítulo cuatro se dan conceptos de estimación no paramétrica de la densidad ya que se hará uso de estos conceptos al momento de la estimación de los parámetros. Finalmente en el capítulo cinco se estiman los parámetros de la distribución de Lévy que mejor ajusta a la parte central de la distribución de los datos y se presenta una heurística que mejora la estimación de dichos parámetros. En el capítulo seis presentamos nuestras conclusiones.

Capítulo 1

Introducción

Los mercados financieros son sistemas complejos con muchos elementos internos que interactúan entre sí. La dinámica de estos mercados es difícil de entender no sólo por la complejidad de dichas interacciones sino por la cantidad de factores externos que actúan sobre dichos mercados. Debido a que la naturaleza de estas interacciones es desconocida como así también la forma en que los factores externos actúan sobre ellos, se puede tratar de comprender las leyes que gobiernan los mercados financieros a través de estudios empíricos de ciertos observables (precios de transacción, el volumen negociado, la frecuencia de negociación, los valores de los índices de mercado) que se generan a partir de dichas interacciones. Estos estudios pueden tomarse como punto de partida para la comprensión de las leyes empíricas que gobiernan los mismos. El primer modelo para estudiar el cambio de precios fue presentado por Bachelier en 1900 en su tesis doctoral titulada "Teoría de la especulación" donde tiene como director a Poincaré. En este trabajo Bachelier propuso modelar el cambio de precios de ciertos activos financieros por un proceso estocástico Gaussiano (un movimiento Browniano). Este modelo es natural si se considera que el cambio de precios es el resultado de muchos shocks independientes, los cuales conducen, por el teorema central del límite, a una distribución gaussiana de los mismos. Sin embargo existen estudios empíricos de las series de algunos de los índices de mercado más importantes que muestran que en intervalos cortos de tiempo la función de densidad asociada tiene una curtosis más grande que la dada por la distribución gaussiana [3]. A pesar de este hecho empírico, en teoría de finanzas se supone que los retornos tienen distribución Gaussiana debido a la simplificación que se obtiene en los cálculos analíticos y a que es uno de los supuestos de la clásica fórmula de Black and Scholes para el precio de opciones. Para dar una idea del aporte realizado por Bachelier observemos que el modelo de Black and Scholes, considerado como un hito en la teoría de precios de opciones, fue publicado en 1973, es decir, 70 años después del trabajo de Bachelier. La propuesta inicial de Bachelier pronto fue reemplazada por un modelo en el cual el stock de precios sigue una distribución log normal (un movimiento Browniano geométrico)[2], esto es, la diferencia del logaritmo de los precios entre un tiempo t y $t - 1$ (proceso conocido como retorno) sigue

una distribución Gaussiana. Este modelo sólo da una primera aproximación a lo que se observa en los datos reales, por esta razón varios estudios se realizaron en función de lo observado en estudios empíricos. Entre los modelos propuestos el más revolucionario en la teoría de precios especulativos desde el trabajo inicial de Bachelier fue el de Mandelbrot[4]. Él estudió variable estocástica conocida como retorno $r_{\Delta t, t}$, definida como la diferencia del logaritmo de dos índices de precios consecutivos. Es decir, definimos $r_{\Delta t, t}$ como

$$\begin{aligned} r_{\Delta t, t} &= \log(I_{t+\Delta t}) - \log(I_t) \\ &= \log\left(\frac{I_{t+\Delta t}}{I_t}\right) \end{aligned}$$

donde Δt es la diferencia, en minutos (también la llamaremos escala temporal), entre dos mediciones del índice. Él observó que no solo el proceso estocástico de los retornos no es Gaussiano, sino que este proceso muestra otra propiedad interesante: el escalamiento del tiempo. Esto es, que la distribución de los retornos para varias elecciones de Δt tienen formas funcionales similares. Motivado por

- i) las colas pronunciadas
- ii) la forma funcional estable para diferentes escalas de tiempo

Mandelbrot propuso que el logaritmo del cambio de precios sigue una distribución estable de Lévy[5]. Estas distribuciones surgen de la generalización del Teorema Central del Límite y tienen propiedades interesantes. Veremos que estas distribuciones son estables: si sumamos variables aleatorias independientes e idénticamente distribuidas con distribución de Levy de parámetro α , obtenemos una variable aleatoria con distribución de Levy con el mismo parámetro α .

Mantegna y Stanley [3] estudiaron el comportamiento de ciertas series económicas. Ellos estudiaron el índice S&P 500 analizando entre 10^6 y 10^7 datos (en contraste con el trabajo de Mandelbrot que estudió aproximadamente 2000 puntos). Ellos observaron que esta serie podía ser descrita por un proceso no Gaussiano donde la parte central de la distribución corresponde a un proceso estable de Lévy no Gaussiano pero el comportamiento asintótico de la distribución muestra un decaimiento más rápido que el que presenta una distribución de Lévy. Veremos que las distribuciones estables de Lévy están caracterizadas por una función de densidad que tiene un comportamiento asintótico que obedece a una ley potencial. Más aun, la función de densidad puede aproximarse por

$$f(x) \approx \frac{\gamma}{\pi x^{\alpha+1}} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right)$$

Esto nos está diciendo que los procesos estables de Lévy tienen varianza infinita para valores de $\alpha < 2$. Por este motivo otros modelos se han propuesto para resolver este problema. Mostramos que este problema puede evitarse considerando que la evolución temporal de los mercados financieros están descritos por los Vuelos Truncados de Lévy (TLF)[6], procesos cuya parte central coincide

con una distribución de Lévy y sus colas son nulas o, en algunos casos, tienen decaimiento exponencial. El objetivo de este trabajo es mostrar que la distribución de alguno de los índices de mercado no es compatible con la dada por un movimiento browniano. Mostraremos que la distribución de Lévy ajusta mejor a la parte central de esta distribución y presentaremos a los TLF como modelo para resolver al problema de la infinitud de los segundos momentos. Para estudiar lo anterior trabajaremos con índices de mercado correspondientes a cinco países:

- Merval
- MEXBOL
- BOVESPA
- IPSA
- S&P 500

Analizaremos la variable estocástica conocida como retorno $r_{\Delta t, t}$, definida como la diferencia del logaritmo de dos índices de precios consecutivos. Es decir, definiremos $r_{\Delta t, t}$ como

$$\begin{aligned} r_{\Delta t, t} &= \log(I_{t+\Delta t}) - \log(I_t) \\ &= \log\left(\frac{I_{t+\Delta t}}{I_t}\right) \end{aligned}$$

donde Δt es la diferencia, en minutos entre dichos índices (también la llamaremos escala temporal) y estimaremos los parámetros de la distribución de Lévy que proponemos para ajustar a la parte central de los datos (el exponente característico α y γ). La organización del trabajo será la siguiente: en el primer capítulo daremos algunos conceptos de medida y de procesos estocásticos. En el segundo capítulo presentaremos las distribuciones de Levy, conceptos de estabilidad y mostraremos en forma empírica que las distribuciones de Lévy ajustan mejor a la parte central de las distribución de los datos que un movimiento browniano. Presentaremos además los TLF. En el tercer capítulo daremos algunos conceptos de estimación de funciones de densidad ya que lo necesitaremos para estimar el parámetro que caracteriza a las distribuciones de Levy. En el cuarto capítulo analizaremos varios índices de mercado estimando los parámetros de la distribución de Levy que mejor aproximan a los mismos. Por último presentaremos una apéndice con conceptos básicos de probabilidades y estadística que se necesitarán para la comprensión de los temas planteados y un anexo con algunas demostraciones.

Capítulo 2

Modelización del precio de los activos financieros

Cuando se analiza una variable que modela la evolución temporal del precio, del volumen o número de transacciones de un activo financiero, se observa que el comportamiento de la misma corresponde al de un proceso estocástico ya que su valor cambia en el tiempo en forma incierta. En este capítulo daremos algunos conceptos estadísticos que nos servirán para modelar el precio de los activos financieros.

Casi todos los modelos del precio de los activos financieros asumen como válidas las hipótesis de mercado eficiente. Estas afirman que la mejor predicción para el valor de un activo mañana es utilizar el valor que tuvo hoy. El único factor que afecta a los precios es la llegada a los mercados de nueva información, es decir, los mercados responden inmediatamente a cualquier información adicional sobre el activo. Como se supone que esta información llega aleatoriamente, el cambio de los precios también debe ser aleatorio. Como medida del cambio en el precio de los activos consideraremos, en un principio para una mejor comprensión del problema, lo que llamaremos retorno

$$r(t) = \frac{S(t + \Delta t) - S(t)}{S(t)}$$

donde $S(t)$ es el precio del activo al tiempo t y Δt es el incremento en el tiempo. Si consideramos un pequeño incremento en el tiempo dt , el precio del activo cambiará de $S(t)$ a $S(t) + dS(t)$ (que lo llamaremos dS). ¿Cómo cambiará el retorno $\frac{dS}{S}$? El modelo comunmente utilizado surge de descomponer este cambio en dos términos. Uno determinístico que indique el retorno obtenido por la inversión realizada (como si fuese una inversión sin riesgo), y uno aleatorio que modele el cambio del precio del activo debido a factores externos.

La contribución del primer término al retorno está dada por μt , donde μ es una medida de la tasa de crecimiento promedio del precio del activo (μ puede ser una constante o, en modelos más complejos puede ser función de S y de t). El segundo término está representado por una contribución aleatoria, σdW , donde

dW es un movimiento browniano (que veremos más adelante) y σ es un número llamado volatilidad, que mide el desvío estandar del retorno.

De esta forma el retorno queda modelado por la siguiente ecuación estocástica:

$$\frac{dS}{S} = \mu t + \sigma dW \quad \text{o} \quad dS = \mu S t + \sigma S dW$$

Notemos que aún no hemos aclarado el rol de dW . Para ello necesitamos definir procesos estocásticos y algunos conceptos relacionados a ellos.

2.1. Procesos estocásticos

Definición 2.1. Un proceso estocástico es una familia de variables aleatorias $X(t, w)$ que evolucionan en el tiempo, donde w pertenece a un espacio muestral Ω y $t \in T \subset \mathbb{R}$. Este conjunto se denotará

$$\{X(t, w), t \in T, w \in \Omega\}$$

Observación 2.1. Para un $t \in T$ dado, $X(t, w) = X_t(w)$ es una variable aleatoria. Para un $w \in \Omega$, $X(t, w)$ es una función de t , ya no es una variable aleatoria. A esta función la llamaremos una realización, trayectoria o función muestral.

De ahora en más no escribiremos explícitamente el argumento w en la función $X(t, w)$. El proceso estocástico se denotará

$$\{X(t), t \in T\} \text{ o simplemente } X(t)$$

2.2. Clasificación de los procesos estocásticos

Existen cuatro tipos diferentes de procesos estocásticos. Esta clasificación depende del tipo de conjunto que sean tanto T como Ω

Definición 2.2. Un proceso estocástico se dice que es de tiempo discreto si el conjunto T es una sucesión finita o infinita de valores. Es de tiempo continuo si el conjunto T es un intervalo.

Definición 2.3. Hemos dicho que para un t dado, $X(t)$ es una variable aleatoria. Al conjunto de valores que puede tomar esta variable aleatoria lo llamaremos espacio de estado del proceso estocástico y lo denotaremos como $\Omega_{X(t)}$. Si este conjunto es finito o infinito numerable (respectivamente, infinito no numerable) el proceso estocástico se dirá que es de estado discreto (respectivamente, de estado continuo).

Ejemplo 2.1. Un ejemplo clásico de proceso estocástico son los paseos al azar. Supongamos que una partícula a tiempo cero está en el origen. En cada unidad de tiempo t se lanza una moneda. Si sale ceca (respectivamente, cara) la partícula se mueve una unidad a la derecha (respectivamente, izquierda). Consideramos la

variable aleatoria X_n que indica la posición de la partícula después de n tiradas de la moneda. Este es un ejemplo de proceso estocástico de tiempo discreto y de estado discreto ya que $T = \mathbb{N}$ y $\Omega_{X(t)} = \{\pm 1, \pm 2, \dots\}$. Una forma de esquematizar este proceso lo muestra la figura 2.1.

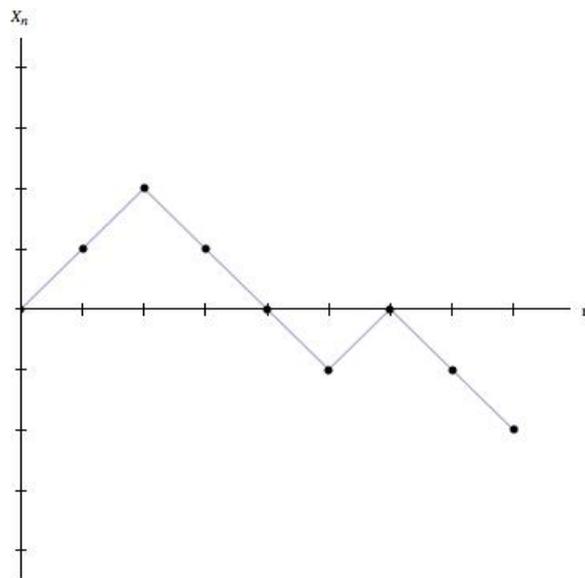


Figura 2.1

Ejemplo 2.2. Un modelo simple del precio de un activo financiero. Éste es otro ejemplo de paseos al azar para tiempos discretos y espacio de estado continuo. En este ejemplo el precio inicial de un activo financiero es X_0 y en cada período el precio del activo puede cambiar a uno de dos valores posibles d y u , con la condición que $0 < d < 1 < u$, de manera que dX_0 represente una disminución en el precio del stock cuando pasamos al próximo período y uX_0 represente un aumento en el precio del mismo. Es decir podemos pensar que el activo tomará valor u con probabilidad p y tomará valor d con probabilidad $1 - p$. Para $t = 2$ el precio del activo será

$$X_2(u, u) = u^2 X_0$$

$$X_2(u, d) = udX_0$$

$$X_2(d, u) = udX_0$$

$$X_2(d, d) = d^2 X_0$$

En este caso la variable aleatoria $X_t(w)$ es el valor del activo a tiempo t .

Ejemplo 2.3. Otro ejemplo podría ser observar el precio de un activo financiero en tiempos t_0, t_1, \dots, t_n que sólo puede tomar valores entre X_d y X_u . En este ejemplo $\Omega = [X_d, X_u]$ y $X_t(w) = w_t$ donde w_t es el valor del activo en tiempo t .

Ejemplo 2.4. Proceso estocástico de tiempo continuo. Un ejemplo sencillo de este tipo de proceso es el definido por

$$X(t) = Y t \quad \text{para } t \geq 0$$

donde Y es una variable aleatoria que tiene una distribución arbitraria.

2.3. Función de distribución

Definición 2.4. La función de distribución de orden k de un proceso estocástico $X(t)$ es la función de distribución conjunta del vector aleatorio $(X(t_1), \dots, X(t_k))$. Es decir

$$F(x_1, \dots, x_k; t_1, \dots, t_k) = P[X(t_1) \leq x_1, \dots, X(t_k) \leq x_k]$$

Definimos, para el caso discreto, la función de masa de probabilidad como

$$p(x_1, \dots, x_k; n_1, \dots, n_k) = P[X_{n_1} = x_1, \dots, X_{n_k} = x_k]$$

y, para el caso continuo, la función de densidad de orden k como

$$f(x_1, \dots, x_k; t_1, \dots, t_k) = \frac{\partial^k}{\partial x_1 \dots \partial x_k} F(x_1, \dots, x_k; t_1, \dots, t_k)$$

cuando la derivada existe.

Ejemplo 2.5. En el ejemplo 2.2 la función de distribución para el caso $t = 2$ es

$$P(X_2 = x) = \begin{cases} p^2 & \text{si } x = u^2 X_0 \\ 2p(1-p) & \text{si } x = udX_0 \\ (1-p)^2 & \text{si } x = d_0^2 X_0 \end{cases}$$

2.4. Momentos de un proceso estocástico

De la misma forma que la media y la varianza nos permiten caracterizar, por lo menos parcialmente, variables aleatorias, los momentos de un proceso estocástico nos permitirán caracterizar a los mismos. Para ello definimos

- Media del proceso

$$\mu_t = E(X_t)$$

- Autocorrelación del proceso en el punto (t_1, t_2)

$$R_X(t_1, t_2) = E(X(t_1)X(t_2))$$

- Autocovarianza del proceso en el punto (t_1, t_2)

$$C_X(t_1, t_2) = E(X(t_1)X(t_2)) - \mu_{t_1}\mu_{t_2}$$

$$C_X(t_1, t_2) = R_X(t_1, t_2) - \mu_{t_1}\mu_{t_2}$$

- Coeficiente de correlación del proceso en el punto (t_1, t_2)

$$\rho_X(t_1, t_2) = \frac{C_X(t_1, t_2)}{\sqrt{C_X(t_1, t_1)C_X(t_2, t_2)}}$$

Observación 2.2. En el contexto de procesos estocásticos usamos el prefijo *auto* para indicar que la función es calculada para dos variables del mismo proceso estocástico.

Observación 2.3. La varianza del proceso se define como

$$\text{Var}(X(t)) = C_X(t, t)$$

Dos propiedades importantes que se necesitarán para definir un movimiento Browniano son las siguientes:

2.5. Independencia y estacionaridad

Definición 2.5. Un proceso estocástico $\{X(t), t \in T\}$ se dice que tiene incrementos independientes si las variables aleatorias

$$X(t_4) - X(t_3) \quad y \quad X(t_2) - X(t_1)$$

son independientes $\forall t_1 < t_2 \leq t_3 < t_4$

Definición 2.6. Un proceso estocástico $\{X(t), t \in T\}$ se dice que tiene incrementos estacionarios si las variables aleatorias

$$X(t_2 + s) - X(t_1 + s) \quad y \quad X(t_2) - X(t_1)$$

tienen la misma función de distribución para todo s

Definición 2.7. Un proceso estocástico se dice que es estacionario de primer orden si su función de distribución es invariante en el tiempo. Es decir, si cumple que

$$F(X(t)) = F(X(t + k)) \quad \forall k$$

Definición 2.8. Un proceso estocástico se dice que es estrictamente estacionario si la función de distribución conjunta de un conjunto de r variables aleatorias es invariante bajo cambios del origen, es decir

$$F(X(t_1), X(t_2), \dots, X(t_r)) = F(X(t_1 + k), X(t_2 + k), \dots, X(t_r + k))$$

para todo r y k .

Algunas veces esta definición es considerada muy estricta. Por eso, daremos una definición menos restrictiva.

Definición 2.9. Un proceso estocástico se dirá *débilmente estacionario* si cumple las siguientes condiciones:

- $E\{X(t)\} = \mu$, es decir, la esperanza del proceso no depende del tiempo.
- La función de autocorrelación $R_x(t_1, t_2) = E\{X_1(t)X_2(t)\}$ cumple que $R_x(t_1, t_2) = R(\tau)$ es una función de $\tau = t_2 - t_1$ (es decir, depende de la longitud del intervalo) donde

$$E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, x_2; t_1 t_2) dx_1 dx_2$$

- $E\{X^2(t)\} = R(0)$, es decir la varianza del proceso $Var(X(t)) = R(0) - \mu^2$ no depende del tiempo.

2.6. Escala de tiempo o memoria del proceso

Centraremos nuestra atención en la clase de memoria temporal de un proceso estocástico. Si el proceso estocástico es estacionario vamos a considerar como medida de la memoria del tiempo a la integral de la función de autocorrelación. Esta integral podrá tomar los siguientes valores:

$$\int_0^{\infty} R(\tau) d\tau = \begin{cases} \text{finita} \\ \infty \\ \text{indeterminada} \end{cases}$$

- Si la integral es finita, existe una memoria temporal τ_c llamada la correlación temporal del proceso. En este caso podemos decir que la correlación está presente hasta τ^* y que no hay correlación para $\tau > \tau^*$, donde τ^* es el área bajo la función de autocorrelación
- Si la integral es infinita nos dice que es imposible elegir una escala temporal que nos permita separar un régimen donde las variables están correlacionadas, de otro donde las variables son independientes de a pares.

Ejemplos:

Caso a) $R(\tau) = e^{-\frac{\tau}{\tau^*}}$

$$\int_0^{\infty} e^{-\frac{\tau}{\tau^*}} d\tau = \tau^*$$

Caso b) $R(\tau) = e^{-\frac{\tau^v}{\tau_0^v}}$

$$\int_0^{\infty} e^{-\frac{\tau^v}{\tau_0^v}} d\tau = \frac{\tau_0^{\frac{1}{v}}}{v} \Gamma\left(\frac{1}{v}\right)$$

Caso c) $R(\tau) \sim \tau^{\eta-1}$. Si $0 < \eta \leq 1$

$$\int_{t_1}^{\infty} \tau^{\eta-1} d\tau = \infty$$

Definición 2.10. Los procesos estocásticos caracterizados por una función de autocorrelación como las del caso c), es decir aquellas donde la integral de la función de autocorrelación es ∞ , se dice que tienen *correlación de rango largo* o que *son fuertemente dependientes*.

Definición 2.11. Los procesos estocásticos caracterizados por una función de autocorrelación como las del caso a), es decir donde la integral de la función de autocorrelación es $< \infty$, se dice que tienen *correlación de rango corto*.

2.7. Movimiento Browniano

El movimiento browniano es un movimiento irregular que se observa en algunas partículas microscópicas que se hallan en un medio fluido (por ejemplo polen en una gota de agua). Recibe su nombre en honor al botánico Robert Brown quien lo describe en 1827. En realidad en 1785, el mismo fenómeno había sido descrito por Jan Ingenhousz sobre partículas de carbón en alcohol, sin embargo el descubrimiento del movimiento browniano se atribuye tradicionalmente al botánico Robert Brown.

El primero en describir matemáticamente el movimiento browniano fue Thorvald N. Thiele, un astrónomo danés en 1880, en un documento sobre el método de los mínimos cuadrados. Fue seguido por Louis Bachelier en 1900 en su tesis doctoral "La teoría de la especulación", en la que se presenta un análisis estocástico del mercado de opciones. Sin embargo, fue el estudio independiente de Albert Einstein en su artículo de 1905 (Über die von der molekularen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen/ Sobre el movimiento requerido por la teoría cinética molecular del calor de pequeñas partículas suspendidas en un líquido estacionario) el que mostró la solución a los físicos, como una forma indirecta de confirmar la existencia de átomos y moléculas. Muchas propiedades de este tipo de procesos fueron dadas por el matemático americano Norbert Wiener en a partir de 1918. A continuación daremos un acercamiento al tema.

2.7.1. Movimiento browniano como límite de paseos al azar

Supongamos que tenemos un paseo al azar como el descrito en el ejemplo 2.1. En este ejemplo, considerábamos un proceso estocástico de tiempo discreto y de espacio de estado discreto. Nosotros podemos "acelerar el proceso" en el siguiente sentido: podemos pensar que los desplazamientos de la partícula se hacen cada δ unidades de tiempo y que la distancia recorrida por la partícula es de ε unidades

ya sea a la derecha o a la izquierda según corresponda (de acuerdo a si salió cara o ceca en el lanzamiento de la moneda). La idea es tomar límite cuando δ y ε tienden a cero, de esta forma la trayectoria de la partícula convergerá a una trayectoria continua (en las dos variables). Sin embargo hay que tener cuidado ya que ε y δ no pueden tender a cero en forma independiente uno del otro, porque si así fuera, la varianza del proceso límite podría ser igual a cero o a infinito lo que carecería de interés.

Supongamos que $X(t)$ denota la posición de la partícula a tiempo t y en $t = 0$ la partícula se encuentra en el origen, es decir $X(0) = 0$. Sea N el número de desplazamientos a la derecha que la partícula hizo luego de las primeras n tiradas de la moneda. De esta forma N es una variable aleatoria con distribución Binomial de parámetros n y $p = \frac{1}{2}$.

Podemos pensar que la posición de la partícula después de $n\delta$ unidades de tiempo está dada por

$$X(n\delta) = (2N - n)\varepsilon$$

Observemos que si los desplazamientos han sido todos hacia la derecha, de manera tal que $N = n$, entonces tenemos que $X(n\delta) = n\varepsilon$. De la misma forma, si todos los desplazamientos fueron hacia la izquierda ($N = 0$) entonces obtenemos que

$$X(n\delta) = -n\varepsilon$$

Dado que la partícula se mueve únicamente para $t = \delta, 2\delta, \dots$ podemos pensar que la posición de la misma a tiempo t está dada por

$$X(t) = X\left(\left[\frac{t}{\delta}\right]\delta\right) \quad \text{para todo } t \geq 0$$

donde $[]$ denota la parte entera. Dado que $N \sim Bi\left(n, \frac{1}{2}\right)$ tenemos que

$$\begin{aligned} E(X(n\delta)) &= \left(2n\frac{1}{2} - n\right)\varepsilon = 0 \\ \text{Var}(X(n\delta)) &= 4\varepsilon^2 \text{Var}(N) = 4\varepsilon^2 n \frac{1}{4} \\ &= n\varepsilon^2 \end{aligned}$$

Por lo tanto,

$$\text{Var}(t)|_{t=n\delta} = \frac{t}{\delta}\varepsilon^2$$

vemos que ε^2 debe tender a cero a la misma velocidad que δ , porque de lo contrario, la varianza sería infinita. Podemos suponer que existe una constante $\sigma > 0$ que cumple que

$$\varepsilon = \sigma\sqrt{\delta}, \text{ es decir, } \varepsilon^2 = \sigma^2\delta$$

Luego, cuando $\delta \rightarrow 0$ tenemos que

$$\begin{aligned} E(X(t)) &= 0 \\ \text{Var}(X(t)) &\rightarrow \sigma^2 t \quad \text{cuando } \delta \rightarrow 0 \text{ y } \forall t \geq 0 \end{aligned}$$

Vimos que $X(n\delta)$ es una variable aleatoria con distribución binomial. Por lo tanto podemos aproximarla por una variable aleatoria normal (ver sección 7 del apéndice estadístico). Luego tenemos que

$$P(X(n\delta) \leq x) \simeq P(N(0, n\epsilon^2) \leq x)$$

y, tomando límite cuando $\delta \rightarrow 0$ tenemos que

$$P(W(t) \leq x) = P(N(0, \sigma^2 t) \leq x)$$

donde

$$W(t) = \lim_{\delta \rightarrow 0} X(t)$$

Definición 2.12. Un movimiento Browniano o un proceso de Wiener $\{W(t), t \geq 0\}$ es un proceso estocástico definido sobre un espacio de probabilidad (Ω, \mathbb{F}, P) que cumple que:

- i) $W(0) = 0$
- ii) $\{W(t), t \geq 0\}$ tiene incrementos estacionarios e independientes
- iii) $W(t) \sim N(0, \sigma^2 t)$

Observaciones:

- Si definimos

$$B(t) = \frac{W(t)}{\sigma}$$

obtenemos un proceso estocástico cuya varianza es

$$\text{Var}(B(t)) = t$$

Este proceso estocástico se llama *Movimiento Browniano Standard*

- Podemos reemplazar la condición iii) por

$$W(t+s) - W(t) \sim N(0, \sigma^2 s) \quad \forall s, t \geq 0$$

De esta forma no es necesario pedir que el proceso tenga incrementos estacionarios ya que sale de esta nueva condición.

- Función de autocovarianza de un proceso de Wiener

$$\begin{aligned} C_W(t, t+s) &= C(W(t), W(t+s)) \\ &= C(W(t), W(t) + W(t+s) - W(t)) \\ &= C(W(t), W(t)) + C(W(t), W(t+s) - W(t)) \\ &= C(W(t), W(t)) \\ &= \sigma^2 t \end{aligned}$$

ya que las variables aleatorias $W(t)$ y $W(t+s) - W(t)$ son independientes $\forall t, s \geq 0$. Dado que la función de autocovarianza depende de t y no de la longitud del intervalo, el proceso de Wiener $W(t)$ no es estacionario, por lo menos en sentido débil. Lo que es estacionario son los incrementos.

Definición 2.13. Movimiento browniano con parámetros (también conocido como movimiento browniano con desplazamiento)

Llamaremos proceso estocástico X con parámetros (μ, σ^2) a aquel que está descrito por la ecuación

$$X(t) = X(0) + \mu t + \sigma W(t) \quad \text{para todo } t$$

donde $W(t)$ es un proceso de Weiner, μ y $\sigma \neq 0$ son constantes reales. Notemos que, dado que $E[W(t)] = 0$ y $Var[W(t)] = t$

$$\begin{aligned} E[X(t)] &= X(0) + \mu t \\ Var[X(t)] &= \sigma^2 t \quad \text{para todo } t \end{aligned}$$

Esto muestra que la esperanza y la varianza del proceso crecen linealmente con t .

2.7.2. Movimiento browniano geométrico

Observemos que un movimiento browniano puede tomar valores negativos. Entonces es discutible utilizar este modelo para describir la evolución del precio de una acción en el tiempo. Por eso utilizaremos el llamado Movimiento Browniano Geométrico en donde el logaritmo del precio de una acción sigue un proceso de Wiener.

Sea $\{X(t), t \geq 0\}$ un proceso de Wiener con parámetros μ y σ (supongamos que $X(0) = 0$), y sea

$$S(t) = e^{X(t)} \quad \text{para } t \geq 0$$

Tomando logaritmos tenemos que

$$\ln(S(t)) = X(t)$$

y de esta forma, para un t fijo, $\ln(S(t)) \sim N(\mu t, \sigma^2 t)$.

Diremos entonces que $S(t)$ sigue un movimiento Browniano geométrico. Notemos que $S(t)$, para un t fijo, tiene una distribución lognormal con parámetros μt y $\sigma^2 t$. Esto significa que la función de densidad de $S(t)$ será:

$$f_{S(t)}(y) = \frac{1}{\sqrt{2\pi\sigma^2 t y}} e^{-\frac{(\ln(y) - \mu t)^2}{2\sigma^2 t}} \quad \text{para } y > 0$$

Observaciones:

- El espacio de estado de un movimiento browniano geométrico es el intervalo $(0, \infty)$. Esto sale de la definición de $S(t)$ y por eso es usado en finanzas para modelar el precio de las acciones.
- Se puede generalizar la definición dada anteriormente, considerando

$$S(t) = S(0)e^{X(t)} \quad \text{para } t \geq 0$$

donde $S(0)$ es una constante positiva que indica el valor del proceso en $t = 0$. El valor de $S(0)$ podría ser aleatorio.

2.8. ¿Por qué el Movimiento Browniano Geométrico no modela adecuadamente a la variable retorno?

El modelo más comunmente utilizado para modelar índices de precios es el movimiento browniano geométrico. Sin embargo, cuando se trabaja con los datos, se observa que las distribuciones empíricas son leptocúrticas¹ (con curtosis mayor que la distribución Gaussiana). Es decir, el comportamiento empírico no corresponde al caso planteado teóricamente. Mostraremos esto estudiando varios índices de mercado.

¹El coeficiente de curtosis mide cuan 'puntiaguda' es una distribución respecto de la distribución normal. Si este coeficiente es mayor que 3 (corresponde al caso Gaussiano) la distribución se llama leptocúrtica.

En este trabajo analizaremos los siguientes índices :

- Argentina (MERVAL)
- México (MEXBOL)
- Brasil (BOVESPA)
- Chile (IPSA)
- Estados Unidos (S&P 500)

Cada serie de datos comprende el período 21 de Diciembre de 2009 hasta el 5 de Julio de 2010. La variable estocástica que vamos a analizar es el retorno definido como la diferencia de los logaritmos de dos índices de precios consecutivos. Es decir

$$r_{\Delta t, t} = \ln(I_{t+\Delta t}) - \ln(I_t) = \ln\left(\frac{I_{t+\Delta t}}{I_t}\right)$$

donde Δt es la escala temporal, la diferencia en el tiempo entre dos valores del índice. Para comparar el retorno para diferentes valores de Δt , vamos a definir el retorno normalizado

$$g(t) = \frac{G(t) - \langle G(t) \rangle_{\Delta t}}{\sigma}$$

donde

$$\sigma^2 = \langle G^2(t) \rangle_{\Delta t} - \langle G(t) \rangle_{\Delta t}^2$$

y $\langle \dots \rangle$ es la media de la serie con escala temporal Δt o la media de la serie de los incrementos.

Queremos mostrar que el movimiento browniano no ajusta bien a los datos propuestos. Dado que las distribuciones estudiadas son simétricas estudiaremos $|g(t)|$ para aprovechar toda la muestra. Compararemos la función de distribución acumulada, como así también las funciones de densidad de los retornos para cada uno de los índices estudiados con los correspondientes al dado por un movimiento browniano generado con media cero y varianza 1. En realidad estudiaremos $H(x) = 1 - F(x)$ siendo $F(x)$ la función de distribución de $|g(t)|$ para observar el comportamiento en las colas de las distribuciones estudiadas. Cabe señalar que $F(0) = P(|g(t)| \leq 0) = 0$ y, por lo tanto, $H(x) = 1$. En la figura 2.2 mostramos esta comparación.

Como se puede observar en cada uno de estos gráficos, el movimiento browniano no describe adecuadamente la evolución de los índices estudiados, ya sea en su parte central como en las colas de la distribución. El propósito de este trabajo es encontrar una distribución perteneciente a la familia de las distribuciones de Lévy que ajuste mejor a la parte central de la distribución de los índices estudiados.

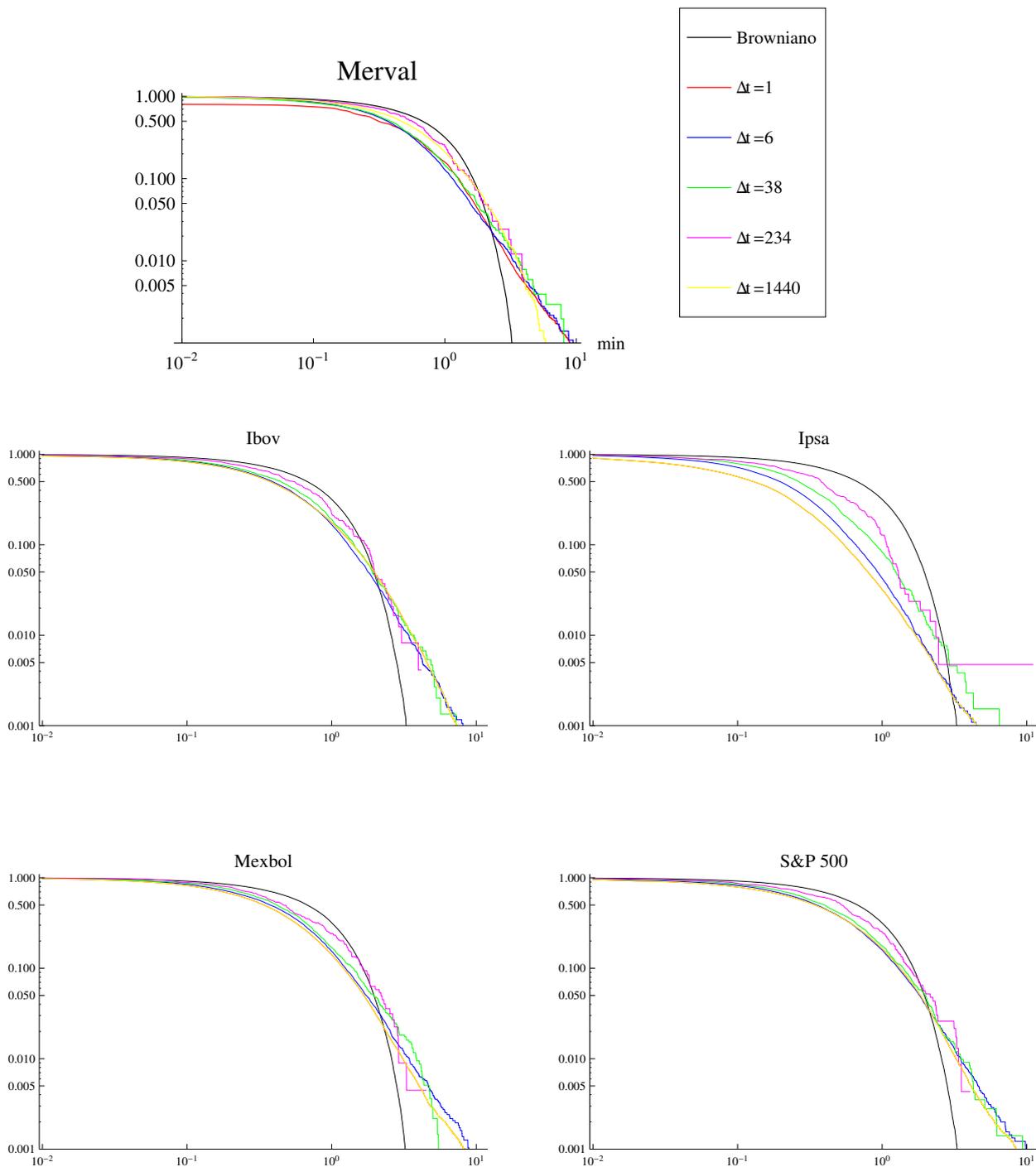


Figura 2.2

Capítulo 3

Distribuciones de Lévy

En el capítulo anterior vimos que el movimiento browniano geométrico no es un buen modelo que sirva para explicar el comportamiento de los índices de mercado debido a la leptocurtosidad que se observó en forma empírica en la función de densidad de probabilidad. Debido a esto se han propuesto modelos alternativos al movimiento Browniano geométrico. Estos modelos difieren entre ellos no solamente con respecto a la forma y leptocurtosidad de la función de densidad de probabilidad, sino también con respecto a ciertas propiedades como la finitud o infinitud de los momentos de orden superior de la distribución ($n \geq 2$).

El primer modelo que tomó en cuenta la leptocurtosidad de la función de distribución fue propuesto por Mandelbrot en 1963 [4], cuando modeló el $\ln S(t)$ para los precios del algodón como un proceso estocástico con incrementos que siguen una distribución de Lévy no Gaussiana. Estas distribuciones si bien se caracterizan por tener segundos momentos infinitos tienen propiedades interesantes como la estabilidad (es decir son distribuciones autosimilares).

Mantegna y Stanley estudiaron el comportamiento de ciertas series económicas. Ellos estudiaron el índice S&P 500 [3] analizando entre 10^6 y 10^7 datos (en contraste con el trabajo de Mandelbrot que estudió aproximadamente 2000 puntos). Ellos observaron que esta serie podía ser descrita por un proceso no Gaussiano donde la parte central de la distribución corresponde a un proceso estable de Lévy no Gaussiano, pero el comportamiento asintótico de la distribución muestra un decaimiento más rápido que el que presenta una distribución de Lévy. Veremos que las distribuciones estables de Lévy están caracterizadas por una función de densidad que tiene un comportamiento asintótico que obedece a una ley potencial. Más aun, la función de densidad puede aproximarse por

$$f(x) \approx \frac{\gamma}{\pi x^{\alpha+1}} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right).$$

Esto nos está diciendo que los procesos estables de Lévy tienen varianza infinita para valores de $\alpha < 2$.

Para resolver el problema de la infinitud de los segundos momentos presentaremos otro modelo conocido como Vuelos Truncados de Lévy (TLF) donde, como su nombre lo indica, la función de densidad es la dada por una distribución de

Lévy para valores de $x \in [-l, l]$ multiplicada por una constante de normalización y cero fuera de este intervalo. Los TLF no son estables pero tienen segundo momento finito.

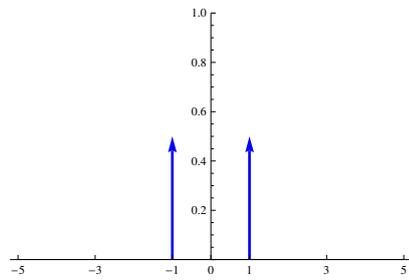
En este capítulo presentaremos a las distribuciones estables de Lévy, sus propiedades y su caracterización a través del exponente característico de las mismas. Estudiaremos el comportamiento asintótico de la función de densidad y presentaremos a los TLF como una propuesta de solución al problema de la varianza infinita que se encuentra en las distribuciones de Lévy no Gaussianas.

3.1. Distribuciones estables

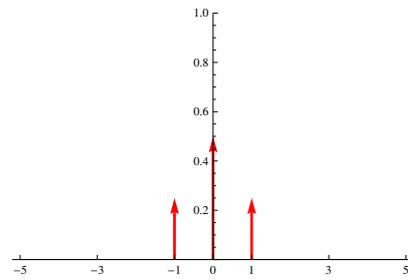
Supongamos un proceso estocástico caracterizado por un paseo al azar unidimensional, con saltos X_i $1 \leq i \leq n$ independientes e idénticamente distribuidos con una función de densidad de probabilidad $f(x)$ y función de distribución $F(X)$. Sea

$$S_n = X_1 + \dots + X_n \quad \text{para } 1 \leq i \leq n$$

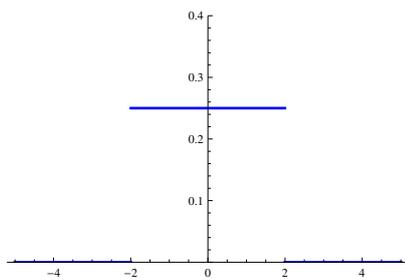
Una pregunta que podemos formularnos es cuándo la probabilidad de que el paseo al azar haya arribado a la posición $x = x_1 + \dots + x_n$ ($F_{S_n}(x)$) luego de n pasos es la misma que $F(x)$ salvo un factor de escala. En otras palabras, ¿puede existir una $f(x)$ que produzca un proceso estocástico con una trayectoria autosimilar?. A continuación vemos algunos ejemplos del comportamiento de las funciones de densidad para S_n para $n=1,2$.



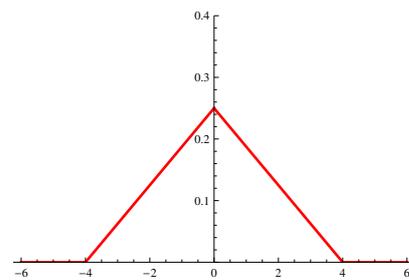
(a) $S_1 = Be(\frac{1}{2})$



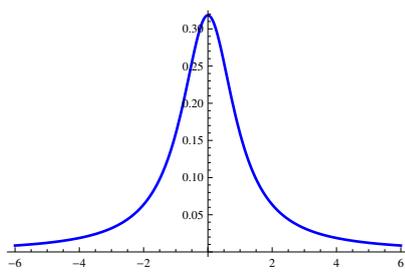
(b) S_2



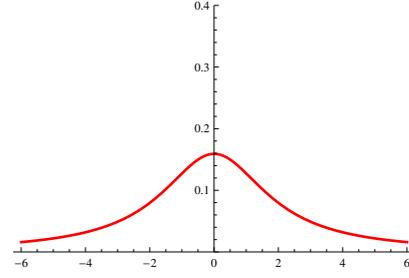
(c) $S_1 = U(-2, 2)$



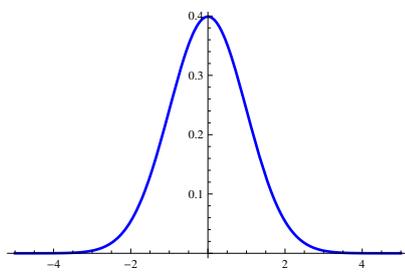
(d) S_2



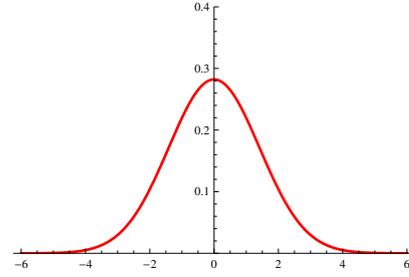
(e) $S_1 = N(0, 1)$



(f) S_2



(g) $S_1 = Cauchy(0, 1)$



(h) S_2

Como vemos, tanto la función de densidad $f_{S_2}(x)$ para la Bernoulli como para la uniforme cambian en escala y en la forma funcional. En cambio la $f_{S_2}(x)$ correspondiente a la distribución normal y Cauchy no cambian en forma pero si en escala. Cuando la forma funcional de $f_{S_n}(x)$ es la misma que la $f(x_i)$ el proceso se dice que es estable.

Definición 3.1. Supongamos que X, X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas con distribución común F , y sea

$$S_n = X_1 + \dots + X_n \quad \text{para } 1 \leq i \leq n$$

- Diremos que la distribución F es estable en sentido amplio si, para cada n , existen constantes $c_n > 0$ y γ_n tales que

$$S_n \stackrel{d}{=} c_n X + \gamma_n$$

y F no está concentrada en un punto (es decir, no es una delta). En otras palabras, la distribución de S_n es la misma que la de X salvo por un parámetro de posición y de escala.

- Diremos que la distribución F es estable en sentido estricto si $\gamma_n = 0 \forall n$. Se puede probar [10] que las constantes de normalización son de la forma

$$c_n = n^{\frac{1}{\alpha}} \quad \text{con } 0 < \alpha \leq 2$$

La constante α será llamada el exponente característico de la distribución.

Las funciones de distribución de Cauchy y Gaussiana son ejemplos de distribuciones estables. Mas aún, la función característica de ambos procesos tiene la misma forma funcional

$$\varphi_n(q) = e^{-n\gamma|q|^\alpha}$$

donde $\alpha = 1$ cuando la distribución es de Cauchy y $\alpha = 2$ para el caso Gaussiano. La demostración de esta afirmación se presenta en A.6.1.

La pregunta que podemos hacernos ahora es si éstas son las únicas distribuciones estables. Lévy y Khintchine resolvieron este problema determinando la clase entera de distribuciones estables. Ellos encontraron que la forma general de la función característica de un proceso estable es:

$$(3.1) \quad \ln \varphi(q) = \begin{cases} i\mu q - \gamma |q|^\alpha \left[1 - i\beta \frac{q}{|q|} \tan\left(\frac{\pi}{2}\alpha\right) \right] & \text{para } \alpha \neq 1 \\ i\mu q - \gamma |q|^\alpha \left[1 + i\beta \frac{q}{|q|} \frac{2}{\pi} \ln(|q|) \right] & \text{para } \alpha = 1 \end{cases}$$

donde α asume valores $0 < \alpha \leq 2$, γ es un factor de escala positivo, μ es un número real y β es un parámetro de asimetría con rango entre -1 y 1 (el valor 0 es para las simétricas).

La forma analítica de la distribución estable de Lévy es conocida sólo para algunos valores de α y β , a saber:

$$\begin{aligned} \alpha = \frac{1}{2} \quad \beta = 1 & \text{ (Lévy - Smirnov)} \\ \alpha = 1 \quad \beta = 0 & \text{ (Cauchy)} \\ \alpha = 2 & \text{ (Gaussiana)} \end{aligned}$$

El parámetro α se conoce como el índice de estabilidad o exponente característico, γ es un parámetro de escala, β un parámetro de sesgo y μ un parámetro de posición. Estos parámetros pueden interpretarse de la siguiente forma:

- α es el exponente característico. Este parámetro toma valores en el intervalo $(0, 2]$. En particular cuando $\alpha = 2$ obtenemos la distribución normal. Luego podemos pensar que dicho parámetro indica apartamientos de la normalidad. La distribución se alejará de la distribución normal a medida que α tome valores lejanos a 2. También nos da una idea del decaimiento de la distribución cuando x toma valores muy grandes.
- γ es un parámetro de escala. No puede interpretarse como el desvío standard del proceso dado que éste existe sólo para el caso $\alpha = 2$. Sin embargo cuanto mayor es el valor de γ , más ancha es la función de densidad.
- β es un parámetro que indica cuán sesgada es la función de densidad. Este parámetro toma valores que pertenecen al intervalo $[-1, 1]$. Cuando $\beta = 1$ la densidad está totalmente sesgada a la derecha y, simétricamente, cuando $\beta = -1$ la densidad está sesgada a la izquierda. Cuando $\beta = 0$ corresponde al caso de densidad simétrica.
- μ es el parámetro de posición. $\mu = E(X)$ en el caso que los primeros momentos existan.

Un resultado conocido es el que enuncia la siguiente proposición.

Proposición 3.1. Sean X_1, \dots, X_n variables aleatorias independientes e idénticamente distribuidas, con función de distribución estable, simétrica y continua. Entonces la función característica de la suma S_n está dada por

$$\varphi_n(q) = e^{-n\gamma|q|^\alpha}$$

Propiedades:

1. Si X es una variable aleatoria con distribución estable de Lévy con exponente característico $0 < \alpha < 2$. Entonces si:
 - a) $0 < \alpha \leq 1$, X no tiene momentos enteros.
 - b) $1 < \alpha < 2$ sólo existe el primer momento entero.

Demostración. Dado que

$$E(X^q) = -i^q \frac{\partial^q \varphi(k)}{\partial^q k} \Big|_{k=0} \quad q = 1, 2, \dots$$

Tenemos que para $q = 1$

$$\begin{aligned} E(X) &= -i \frac{\partial \varphi(q)}{\partial q} \Big|_{q=0} \\ &= -i\gamma \alpha |q|^{\alpha-1} e^{-\gamma|q|^\alpha} \Big|_{q=0} \end{aligned}$$

Entonces para que exista primer momento $\alpha - 1 \geq 0$, es decir $1 \leq \alpha$. Respecto del segundo momento tenemos que

$$\begin{aligned} E(X^2) &= -i^2 \frac{\partial^2 \varphi(q)}{\partial^2 q} \Big|_{q=0} \\ &= \alpha\gamma |q|^{\alpha-2} e^{-\gamma|q|^\alpha} (-\alpha + \alpha\gamma |q|^\alpha + 1) \Big|_{q=0} \end{aligned}$$

Luego el segundo momento existirá si $\alpha - 2 \geq 0$, es decir si $2 \leq \alpha$. Como $0 < \alpha \leq 2$, el segundo momento será finito si $\alpha = 2$. Es decir, este tipo de distribuciones tendrán varianza finita únicamente para $\alpha = 2$ que corresponde al caso normal. \square

2. Si una distribución es estable, para valores grandes de x la función de densidad puede aproximarse por

$$(3.2) \quad f(x) \approx -\frac{1}{\pi} \sum_{k=1}^m \frac{(-1)^k \gamma^k \Gamma(\alpha k + 1)}{k! |x|^{k\alpha+1}} \sin\left(\frac{k\pi\alpha}{2}\right)$$

La demostración de este resultado se puede ver en [13]

3. Si consideramos el primer término de la aproximación anterior tenemos que

$$(3.3) \quad f(x) \approx \frac{\gamma}{\pi x^{\alpha+1}} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right)$$

Una idea de cómo calcular el coeficiente principal de la expansión es la siguiente:

Dado que la transformada de Fourier de la función característica de una variable aleatoria X nos da su función de densidad, usando (3.1) tenemos que la distribución estable simétrica de índice α y factor de escala γ está dada por:

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\gamma|q|^\alpha} e^{-iqx} dq \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\gamma|q|^\alpha} \cos(qx) dq && \text{por paridad de la función coseno} \\ &= \frac{1}{\pi} \int_0^{\infty} e^{-\gamma|q|^\alpha} \cos(q|x|) dq \end{aligned}$$

Haciendo el cambio de variables $v = q|x|$ tenemos que

$$\begin{aligned} f(x) &= \frac{1}{\pi|x|} \int_0^{\infty} e^{-\gamma|\frac{v}{x}|^{\alpha}} \cos(v) dv \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\pi|x|} \int_0^{\infty} e^{-\gamma|\frac{v}{x}|^{\alpha}} e^{-\varepsilon v} \cos(v) dv \end{aligned}$$

donde la última igualdad se cumple por el teorema de convergencia mayorada. Si aproximamos por un desarrollo de Taylor de orden 1 a

$$(3.4) \quad h(v) = e^{-\gamma|\frac{v}{x}|^{\alpha}}$$

tenemos que

$$\begin{aligned} f(x) &\sim \frac{1}{\pi|x|} \lim_{\varepsilon \rightarrow 0^+} \int_0^{\infty} \left(1 - \gamma \left|\frac{v}{x}\right|^{\alpha}\right) e^{-\varepsilon v} \cos(v) dv \\ &= \frac{1}{\pi|x|} \lim_{\varepsilon \rightarrow 0^+} \int_0^{\infty} e^{-\varepsilon v} \cos(v) dv - \frac{\gamma}{\pi|x|^{\alpha+1}} \lim_{\varepsilon \rightarrow 0^+} \int_0^{\infty} e^{-\varepsilon v} |v|^{\alpha} \cos(v) dv \end{aligned}$$

a) Veamos que el primer término de la expresión anterior se anula. Observemos que en este caso no se puede aplicar el teorema de convergencia mayorada dado que, la sucesión de funciones $e^{-\varepsilon v} \cos(v)$ converge puntualmente a $\cos(v)$ cuando $\varepsilon \rightarrow 0$ que no está acotada por ninguna función integrable. Un cálculo directo muestra que

$$\int_0^{\infty} e^{-\varepsilon v} \cos(v) dv = \frac{\varepsilon}{1 + \varepsilon^2}$$

por lo tanto

$$\lim_{\varepsilon \rightarrow 0^+} \int_0^{\infty} e^{-\varepsilon v} \cos(v) dv = \lim_{\varepsilon \rightarrow 0^+} \frac{\varepsilon}{1 + \varepsilon^2} = 0$$

b) Trabajamos entonces con

$$\int_0^{\infty} e^{-\varepsilon v} |v|^{\alpha} \cos(v) dv$$

Tendríamos que probar que

$$\int_0^{\infty} e^{-\varepsilon v} |v|^{\alpha} \cos(v) dv = -\Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right)$$

Para $\varepsilon > 0$ y $\alpha > -1$ definimos la siguiente función

$$J(\varepsilon, \alpha) = \int_0^{\infty} e^{-(\varepsilon-i)v} v^{\alpha} dv$$

Observemos que

$$\int_0^{\infty} e^{-\varepsilon v} |v|^{\alpha} \cos(v) dv = \operatorname{Re}(J(\varepsilon, \alpha))$$

Es decir, bastaría calcular $\operatorname{Re}(J(\varepsilon, \alpha))$ para obtener el resultado deseado. Para esto vamos a considerar cuatro casos.

- Caso 1: $-1 < \alpha < 0$
- Caso 2: $\alpha = 0$
- Caso 3: $\alpha > 0$ y $\alpha \in \mathbb{N}$
- Caso 4: $\alpha > 0$ y $\alpha \notin \mathbb{N}$. Consideramos $n - 1 < \alpha < n$ con $n \in \mathbb{N}$
- Caso 1: $-1 < \alpha < 0$

Consideramos los caminos C_1, \dots, C_4 definidos por:

- $C_1 : z = t, t \in [\delta, R]$,
- $C_2 : z = R e^{i\theta}, \theta \in [0, \theta_0], R \in \mathbb{R}$
- $C_3 : z = \delta e^{i\theta}, \theta \in [0, \theta_0]$,
- $C_4 : z = e^{i\theta_0} t, t \in [\delta, R]$

(ver fig. 3), donde $\theta_0 = -\text{Arg}(\varepsilon - i) < \pi/2$.

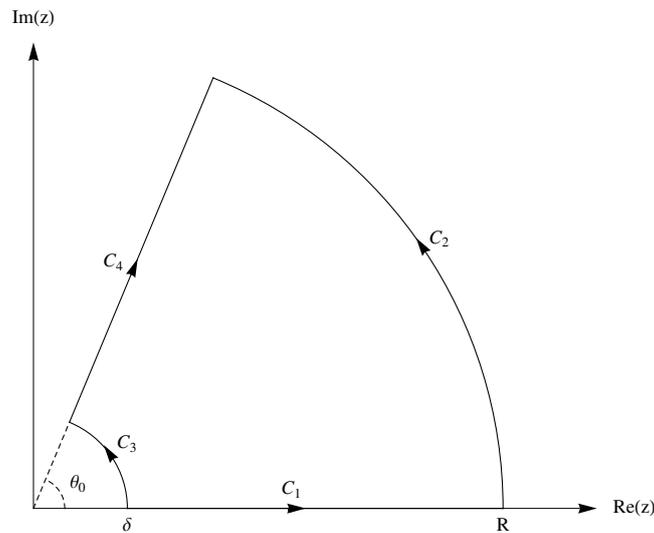


Figura 3.1

Sea $f(z) = e^{-(\varepsilon-i)z} z^\alpha$ con $-1 < \alpha < 0$, entonces

$$J(\varepsilon, \alpha) = \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_1} f(z) dz$$

Notemos que $f(z)$ es una función holomorfa en $\text{Re}(z) > 0$. De esta forma

$$\int_{C_1+C_2} f(z) dz = \int_{C_3+C_4} f(z) dz$$

Veamos que vale la siguiente afirmación

$$\lim_{R \rightarrow \infty} \int_{C_2} f(z) dz = 0 \text{ y } \lim_{\delta \rightarrow 0} \int_{C_3} f(z) dz = 0$$

Si $z = r e^{i\theta}$

$$\begin{aligned} \left| \int_C f(z) dz \right| &= \left| \int_0^{\theta_0} f(r e^{i\theta}) i r e^{i\theta} d\theta \right| \\ &\leq r \int_0^{\theta_0} |f(r e^{i\theta})| |i e^{i\theta}| d\theta \\ &\leq r \int_0^{\frac{\pi}{2}} |e^{-(\varepsilon-i)r e^{i\theta}} (r e^{i\theta})^\alpha| d\theta \\ &= r^{1+\alpha} \int_0^{\frac{\pi}{2}} |e^{-(\varepsilon-i)r e^{i\theta}}| d\theta \\ &= r^{1+\alpha} \int_0^{\frac{\pi}{2}} |e^{-\varepsilon r e^{i\theta}} e^{i r e^{i\theta}}| d\theta \end{aligned}$$

Ahora

$$\begin{aligned} |e^{-\varepsilon r e^{i\theta}}| &= e^{\operatorname{Re}(-\varepsilon r e^{i\theta})} = e^{-\varepsilon r \cos \theta} < 1 \text{ ya que } \theta \in \left[0, \frac{\pi}{2}\right] \text{ y} \\ |e^{i r e^{i\theta}}| &= e^{\operatorname{Re}(i r e^{i\theta})} = e^{-r \sin \theta} \end{aligned}$$

Entonces

$$\begin{aligned} \left| \int_C f(z) dz \right| &\leq r^{1+\alpha} \int_0^{\frac{\pi}{2}} e^{-r \sin \theta} d\theta \\ &\leq r^{1+\alpha} \int_0^{\frac{\pi}{2}} e^{-r \frac{2\theta}{\pi}} d\theta \end{aligned}$$

Esta última desigualdad sale del hecho que $\frac{2\theta}{\pi} \leq \sin \theta$ para valores de $\theta \in \left[\theta, \frac{\pi}{2}\right]$. Resolviendo la integral obtenemos lo siguiente

$$\left| \int_C f(z) dz \right| \leq \frac{\pi}{2} r^\alpha (1 - e^{-r})$$

Dado que $r \in \mathbb{R}$, podemos concluir entonces que valen las siguientes desigualdades:

$$\begin{aligned} \left| \int_{C_2} f(z) dz \right| &\leq \frac{\pi}{2} R^\alpha (1 - e^{-R}) \\ \left| \int_{C_3} f(z) dz \right| &\leq \frac{\pi}{2} \delta^\alpha (1 - e^{-\delta}) \end{aligned}$$

Es fácil ver que $\lim_{R \rightarrow \infty} R^\alpha (1 - e^{-R}) = 0$ y que $\lim_{\delta \rightarrow 0} \delta^\alpha (1 - e^{-\delta}) = 0$, de donde obtenemos la afirmación. Por lo tanto

$$\lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_2} f(z) dz = \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_3} f(z) dz = 0$$

y de esta forma

$$J(\varepsilon, \alpha) = \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_4} f(z) dz$$

Luego basta calcular $\lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_4} f(z) dz$.

Recordemos que $\theta_0 = -\text{Arg}(\varepsilon - i) < \pi/2$. Si llamamos $w = \varepsilon - i$ entonces $|w| = \sqrt{\varepsilon^2 + 1}$. Por lo tanto

$$\varepsilon - i = \sqrt{1 + \varepsilon^2} e^{-i\theta_0}$$

y de esta forma

$$e^{i\theta_0} (\varepsilon - i) = \sqrt{1 + \varepsilon^2}$$

Entonces

$$\begin{aligned} \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_4} f(z) dz &= \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{\delta}^R f(e^{i\theta_0} t) e^{i\theta_0} dt \\ &= \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{\delta}^R e^{-(\varepsilon - i)e^{i\theta_0} t} (e^{i\theta_0} t)^\alpha e^{i\theta_0} dt \\ &= \lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} e^{i(1+\alpha)\theta_0} \int_{\delta}^R e^{-\sqrt{1+\varepsilon^2} t} t^\alpha dt \end{aligned}$$

Haciendo un cambio de variables tenemos que

$$\lim_{\substack{R \rightarrow \infty \\ \delta \rightarrow 0}} \int_{C_4} f(z) dz = e^{i(1+\alpha)\theta_0} (\sqrt{1 + \varepsilon^2})^{-(1+\alpha)} \Gamma[1 + \alpha]$$

Por lo tanto

$$J(\varepsilon, \alpha) = e^{i(1+\alpha)\theta_0} (\sqrt{1 + \varepsilon^2})^{-(1+\alpha)} \Gamma[1 + \alpha]$$

Entonces

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \text{Re}(J(\varepsilon, \alpha)) &= \text{Re}\left(\lim_{\varepsilon \rightarrow 0} J(\varepsilon, \alpha)\right) \\ &= \text{Re}\left(i e^{i\alpha \frac{\pi}{2}}\right) \Gamma[1 + \alpha] \\ &= -\sin\left(\alpha \frac{\pi}{2}\right) \Gamma[1 + \alpha] \end{aligned}$$

Luego

$$f(x) \sim -\frac{\gamma}{\pi |x|^{\alpha+1}} \left(-\sin\left(\alpha \frac{\pi}{2}\right) \Gamma[1 + \alpha]\right) = \frac{\gamma}{\pi |x|^{\alpha+1}} \sin\left(\alpha \frac{\pi}{2}\right) \Gamma[1 + \alpha]$$

- Caso 2: $\alpha = 0$

$$\operatorname{Re}(J(\varepsilon, 0)) = \int_0^{\infty} e^{-\varepsilon v} \cos(v) dv = \frac{\varepsilon}{1 + \varepsilon^2}$$

Por lo tanto, $\lim_{\varepsilon \rightarrow 0} \operatorname{Re}(J(\varepsilon, 0)) = 0$

- Caso 3: $\alpha > 0$ y $\alpha \in \mathbb{N}$

Si $\alpha > 0$, integrando por partes tenemos

$$J(\varepsilon, \alpha) = \alpha(\varepsilon - i)^{-1} J(\varepsilon, \alpha - 1)$$

Si además $\alpha \in \mathbb{N}$, inductivamente tenemos que

$$J(\varepsilon, \alpha) = \alpha! (\varepsilon - i)^{-\alpha} J(\varepsilon, 0)$$

dado que $J(\varepsilon, 0) = (\varepsilon - i)^{-1}$

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \operatorname{Re}(J(\varepsilon, \alpha)) &= \operatorname{Re} \left[\alpha! (-i)^{-\alpha-1} \right] \\ &= \alpha! \operatorname{Re} \left[\left(-\frac{1}{i} \right)^{\alpha+1} \right] \\ &= \alpha! \operatorname{Re} \left[i^{\alpha+1} \right] \\ &= \operatorname{Re} \left[i \left(\cos \left(\alpha \frac{\pi}{2} \right) + i \sin \left(\alpha \frac{\pi}{2} \right) \right) \right] \Gamma(\alpha + 1) \\ &= -\sin \left(\alpha \frac{\pi}{2} \right) \Gamma(\alpha + 1) \end{aligned}$$

Recordamos que si $n \in \mathbb{N}$ se cumple que $\Gamma(n + 1) = n!$

- Caso 4: $\alpha > 0$ y $\alpha \notin \mathbb{N}$.

Consideramos $n - 1 < \alpha < n$ con $n \in \mathbb{N}$. Entonces

$$J(\varepsilon, \alpha) = \alpha(\alpha - 1) \dots (\alpha - n + 1) (\varepsilon - i)^{-n} J(\varepsilon, \alpha - n)$$

con $-1 < \alpha - n < 0$. Si aplicamos el caso 1 considerando $\beta = \alpha - n$ tenemos que

$$J(\varepsilon, \alpha - n) = e^{i(1+\alpha-n)\theta_0} \left(\sqrt{1 + \varepsilon^2} \right)^{-(1+\alpha-n)} \Gamma[1 + \alpha - n]$$

Entonces

$$\begin{aligned} J(\varepsilon, \alpha) &= \alpha(\alpha - 1) \dots (\alpha - (n - 1)) * \\ &\quad \Gamma[\alpha - (n - 1)] \left(\sqrt{1 + \varepsilon^2} \right)^{-(1+\alpha-n)} (\varepsilon - i)^{-n} e^{i(1+\alpha-n)\theta_0} \\ &= \Gamma[1 + \alpha] \left(\sqrt{1 + \varepsilon^2} \right)^{-(1+\alpha-n)} (\varepsilon - i)^{-n} e^{i(1+\alpha-n)\theta_0} \end{aligned}$$

De esta forma

$$\begin{aligned}
 \lim_{\varepsilon \rightarrow 0} \operatorname{Re}(J(\varepsilon, \alpha)) &= \lim_{\varepsilon \rightarrow 0} \operatorname{Re} \left(\Gamma[1 + \alpha] \left(\sqrt{1 + \varepsilon^2} \right)^{-(1+\alpha-n)} (\varepsilon - i)^{-n} e^{i(1+\alpha-n)\theta_0} \right) \\
 &= \Gamma[1 + \alpha] \operatorname{Re} \left[e^{ni\frac{\pi}{2}} e^{i(1+\alpha-n)\theta_0} \right] \\
 &= \Gamma[1 + \alpha] \operatorname{Re} \left[e^{i(1+\alpha)\frac{\pi}{2}} \right] \\
 &= -\sin\left(\alpha\frac{\pi}{2}\right) \Gamma(\alpha + 1)
 \end{aligned}$$

Observación: En la propiedad 1 vimos que las distribuciones de Lévy con exponente característico $0 < \alpha < 2$ no tiene segundo momento finito. Esto también se puede ver como consecuencia de este resultado donde el decaimiento de la función de densidad es $f(x) \sim \frac{1}{x^{\alpha+1}}$.

4. El índice α no solo controla las colas de la distribución sino también afecta al valor de la función de densidad en el origen pues

$$f(x = 0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\gamma^{\frac{1}{\alpha}} \pi \alpha}$$

Esto se puede ver dado que la transformada de Fourier de la función característica de una variable aleatoria X nos da su función de densidad, usando (3.1) tenemos que la distribución estable simétrica de índice α y factor de escala γ está dada por:

$$\begin{aligned}
 f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\gamma|q|^\alpha} e^{-iqx} dq \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\gamma|q|^\alpha} \cos(qx) dq \quad \text{por paridad de la función coseno} \\
 &= \frac{1}{\pi} \int_0^{\infty} e^{-\gamma|q|^\alpha} \cos(q|x|) dq
 \end{aligned}$$

Por lo tanto

$$f(x = 0) = \frac{1}{\pi} \int_0^{\infty} e^{-\gamma|q|^\alpha} dq$$

Si hacemos el cambio de variables

$$t = \gamma q^\alpha$$

tenemos el resultado

Con un razonamiento similar y usando (3.1) tenemos que

$$(3.5) \quad f(S_n = 0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{(\gamma n)^{\frac{1}{\alpha}} \pi \alpha}$$

Esta última igualdad nos va a ser útil al momento de estimar el parámetro α .

3.2. Vuelos Truncados de Lévy (TLF)

Hemos visto que el comportamiento asintótico para valores grandes de X es un comportamiento de ley de potencia. Como consecuencia de esto es que estos procesos tienen varianzas infinitas. Para evitar este problema Mantegna y Stanley [6] consideraron el siguiente proceso estocástico

$$f(x) = \begin{cases} 0 & x > l \\ cf_i(x) & -l < x < l \\ 0 & x < -l \end{cases}$$

donde

$$f_i(x) = \frac{1}{\pi} \int_0^{+\infty} e^{-\gamma q^\alpha} \cos(qx) dq$$

es la función de densidad de una distribución de Lévy simétrica, con exponente característico α ($0 < \alpha \leq 2$) y factor de escala γ , c es una constante de normalización y l es el valor de corte.

Los TLF no son estables porque sabemos que las únicas distribuciones estables son las de Lévy, pero tienen varianzas finitas. Por lo tanto convergen a una distribución Gaussiana, aunque su convergencia es muy lenta. Ellos mostraron que la convergencia de la suma de n variables aleatorias independientes, idénticamente distribuidas con distribución TLF a un proceso normal puede ser tan lenta que se necesita un gran número de eventos independientes ($n \geq 10^4$) para asegurar la convergencia a un proceso Gaussiano. Ellos consideraron la cantidad

$$S_n = \sum_{i=1}^n x_i$$

donde x_i son variables aleatorias independientes, idénticamente distribuidas con distribución TLF.

La función de densidad $f(S_n)$ está bien aproximada por la densidad de Lévy ($f_L(S_n)$) cuando n es pequeño, mientras que $f(S_n)$ convergerá a un proceso Gaussiano en el límite cuando $n \rightarrow \infty$. De esta forma existe un valor de n , que llamaremos n_x , que determina el cambio entre un régimen de Lévy y uno Gaussiano. Es decir existirá un valor n_x tal que

$$f_{S_n}(x) = \begin{cases} f_{S_n}^L(x) & n \ll n_x \\ f_{S_n}^G(x) & n \gg n_x \end{cases}$$

donde $f_{S_n}^G(x)$ es una densidad Gaussiana. Para ver este de cambio de régimen estudiaron la probabilidad del retorno al origen para S_n ($P(S_n = 0)$) en función de n, l y α . Para valores pequeños de n , $P(S_n = 0)$ toma valores muy cercanos al esperado para procesos estables de Lévy. Es decir

$$(3.6) \quad P(S_n = 0) \simeq L(S_n = 0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{\gamma^{\frac{1}{\alpha}} \pi \alpha}$$

Para valores grande de n , $P(S_n = 0)$ se puede aproximar por los valores dados por un proceso normal, es decir

$$(3.7) \quad P(S_n = 0) \simeq N(S_n = 0) = \frac{1}{\sqrt{2\pi} \sigma_0(\alpha, l) n^{\frac{1}{2}}}$$

donde $\sigma_0(\alpha, l)$ es el desvío standard de un proceso dado por un TLF. Si usamos el resultado 2 podemos encontrar una aproximación al desvío standard del TLF y esto nos permitirá encontrar el cambio de régimen. Usando (3.3) tenemos que

$$\begin{aligned} \sigma_0^2(\alpha, l) &\simeq \int_{-l}^l x^2 f(x) dx \\ &= \frac{2\gamma}{\pi} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right) \int_0^l x^{1-\alpha} dx = \\ &\quad \frac{2\gamma}{\pi(2-\alpha)} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right) l^{(2-\alpha)} \end{aligned}$$

por lo tanto

$$(3.8) \quad \sigma_0(\alpha, l) = \left(\frac{2\gamma}{\pi(2-\alpha)} \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right) \right)^{\frac{1}{2}} l^{\frac{(2-\alpha)}{2}}$$

si igualamos (3.6) con (3.7) y usamos (3.8) nos queda que $n_x \simeq A l^\alpha$ donde

$$A = \left[\frac{\pi \alpha}{2\Gamma\left(\frac{1}{\alpha}\right) \left(\gamma \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right) / (2 - \alpha) \right)^{\frac{1}{2}}} \right]^{\frac{2\alpha}{(\alpha-2)}}$$

En ese mismo artículo ellos investigaron numéricamente, a través de simulaciones, la convergencia del proceso cuando n aumenta. Generaron un proceso estable de Lévy de índice $\alpha = 1,5$ y factor de escala $\gamma = 1$ usando un algoritmo [9] propuesto por Mantegna. Estudiaron la probabilidad del retorno al origen, $f_{S_n}(0)$, simulando el proceso S_n para distintos valores de n , considerando el mismo α y

$l = 100$. Generaron 5×10^4 realizaciones y compararon con el régimen de Lévy y el Gaussiano. Mostraron el cruce entre los dos regímenes observando que, para estos parámetros de control $n_x \approx 260$.

Para estos mismos parámetros de control investigaron $f_{S_n}(x)$ para distintos valores de n , simulando TLF para $n = 1$, $n = 10$, $n = 100$, y $n = 1000$. Para poder comparar las formas de las densidades estudiaron las variables escaladas considerando $\tilde{S}_n = \frac{S_n}{n^{\frac{1}{\alpha}}}$. En una escala semilogarítmica mostraron que para valores bajos de n ($n = 1$ y $n = 10$) la parte central de la distribución está bien descrita por una densidad de Lévy con $\alpha = 1,5$ y $\gamma = 1$. En cambio, para valores grandes de n ($n = 1000$) el proceso TLF se aproxima a un régimen Gaussiano.

Es decir, investigando la probabilidad del retorno al origen de un proceso estocástico no normal, 'casi estable' con varianza finita, se observa un claro cruce entre un régimen de Lévy y uno Gaussiano. De esta forma densidades cercanas a una Lévy (que difieren en sus colas) pueden observarse empíricamente para un gran intervalo de tiempo (n grande), aún en presencia de procesos con varianza finita.

Capítulo 4

Estimación no paramétrica

Es conocido el papel importantísimo que juega el conocimiento de la función de densidad en la comprensión del comportamiento estadístico de una variable aleatoria. Conocer la función de densidad de una variable aleatoria es tener una completa descripción de la misma. Por este motivo es de sumo interés poder estimar dicha función a partir de la información proporcionada por una muestra aleatoria de la variable en estudio.

Una forma de abordar este problema es a través de métodos paramétricos. En este caso se considera que la función de densidad que se desea estimar pertenece a una familia de distribuciones conocidas. Por ejemplo: normal, exponencial, Poisson, etc. En este caso, una vez identificada la familia de pertenencia, se estiman los parámetros de dicha distribución. Este tipo de estimación se llama *estimación paramétrica*.

Una segunda posibilidad es no suponer a priori ningún modelo para la distribución de probabilidad y dejar que "los datos hablen por si solos". Es decir, no asumir ninguna hipótesis para la función de densidad sobre la distribución de la muestra. Este enfoque se llama *estimación no paramétrica* de la densidad y es el que utilizaremos en este trabajo.

El más sencillo y más conocido estimador no paramétrico de la densidad es el histograma, que tiene que estar "normalizado" para que sea considerado un estimador de la densidad (es decir, integral 1). Este estimador de la densidad, que lo llamaremos \hat{f}_h se define como

$$\hat{f}_h(x) = \sum_{i=1}^k \frac{N_i}{n(t_{i+1} - t_{i-1})} I_{B_i}(x)$$

donde N_i es la cantidad de datos que caen dentro de $B_i = [t_i, t_{i+1})$. En esta definición asumimos que f tiene soporte en $[a, b]$ y que $a \leq t_1 < t_2 < \dots < t_k = b$ es una partición de dicho intervalo. Si consideramos que la longitud de los intervalos es

constante, es decir que $t_{i+1} - t_{i-1} = h_n$, $\hat{f}_h(x)$ queda definido como

$$\begin{aligned}\hat{f}_h(x) &= \frac{1}{nh_n} \sum_{i=1}^k N_i I_{B_i}(x) \\ &= \frac{1}{nh_n} \sum_{i=1}^n I_{B_i}(x_i) \\ &= \frac{N_i}{nh_n} \text{ si } x \in B_i\end{aligned}$$

Uno de los problemas que plantea la estimación por histogramas es que el estimador resultante es una función discontinua, cosa que podría no suceder con la verdadera densidad. Esto ha motivado la aparición de métodos alternativos como los que presentamos en este trabajo que es la estimación por núcleos.

La idea de este método surge de la definición de función de densidad para variables aleatorias continuas, esto es que si X es una variable aleatoria con distribución F y densidad f , en cada punto de continuidad de f se tiene que

$$f(x) = \lim_{h \rightarrow 0} \frac{1}{2h} F(x+h) - F(x-h)$$

Ahora, dada una muestra X_1, \dots, X_n de variables aleatorias independientes e idénticamente distribuidas con distribución F , para cada h fijo $F(x+h) - F(x-h)$ se puede estimar como la proporción de observaciones que caen dentro del intervalo $(x-h, x+h)$. Por lo tanto, tomando h chico, un estimador natural de la densidad es

$$\hat{f}(x) = \frac{\#\{X_i : X_i \in (x-h, x+h)\}}{2hn} = \frac{F_n(x+h) - F_n(x-h)}{2h}$$

siendo F_n la función de distribución empírica. Este estimador fue propuesto por Rosenblatt (1956) y resulta asintóticamente insesgado dado que

$$\begin{aligned}E(\hat{f}(x)) &= \frac{1}{2h} E(F_n(x+h)) - E(F_n(x-h)) \\ &= \frac{1}{2h} F(x+h) - F(x-h)\end{aligned}$$

por lo tanto

$$\lim_{h \rightarrow 0} E(\hat{f}(x)) = f(x)$$

Notemos que este estimador es discontinuo en los puntos $X_i \pm h$ y es constante entre estos valores. Es de interés encontrar estimadores continuos de la función de densidad. Por eso vamos a generalizar a este último de la siguiente forma:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h} w\left(\frac{x - X_i}{h}\right)$$

donde

$$w(y) = \begin{cases} \frac{1}{2} & y \in [-1, 1) \\ 0 & \text{en otro caso} \end{cases}$$

ya que $\hat{f}(x)$ puede expresarse como

$$\begin{aligned} \hat{f}(x) &= \frac{1}{2hn} \sum_{i=1}^n I(x-h < X_i \leq x+h) \\ &= \frac{1}{2hn} \sum_{i=1}^n I(X_i - h \leq x < X_i + h) \end{aligned}$$

La función w puede ser reemplazada por una función más general, $K(z)$, de esta forma surgen los estimadores tipo núcleo. Éstos se definen como

$$(4.1) \quad \hat{f}(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)$$

La función $K(z)$ se llamará función núcleo, función Kernel o función peso y $\{h_n\}$ es una sucesión de constantes positivas conocidas como ancho de banda, parámetro de suavización. La función $K(z)$ satisface ciertas condiciones de regularidad:

- $\int_{-\infty}^{\infty} K(z) dz = 1$
- $\int_{-\infty}^{\infty} |K(z)| dz < \infty$
- $\int_{-\infty}^{\infty} |z K(z)| dz < \infty$

Usualmente $K(z) \geq 0$ y simétrico, luego cualquier función de densidad simétrica, como por ejemplo la densidad normal, puede usarse como núcleo. Alguno de los núcleos más utilizados son:

a) Núcleo rectangular o unifome

$$K(z) = \begin{cases} \frac{1}{2} & z \in (-1, 1) \\ 0 & \text{en otro caso} \end{cases}$$

En este caso cualquier punto en el intervalo $(x-h, x+h)$ contribuye $\frac{1}{2nh}$ al estimado de $f(x)$ en el punto x , y cualquier punto fuera de ese intervalo no contribuye en nada.

b) Núcleo Gaussiano

$$K(z) = \frac{1}{\sqrt{2\pi}} e^{(-\frac{1}{2}z^2)}$$

c) Núcleo Epanecnikov

$$K(z) = \begin{cases} \frac{3}{4\sqrt{5}} \left(1 - \frac{z^2}{5}\right) & |z| < \sqrt{5} \\ 0 & \text{en otro caso} \end{cases}$$

d) Núcleo triangular

$$K(z) = \begin{cases} 1 - |z| & |z| < 1 \\ 0 & \text{en otro caso} \end{cases}$$

d) Núcleo Biweight

$$K(z) = \begin{cases} \frac{15}{16} (1 - z^2)^2 & |z| < 1 \\ 0 & \text{en otro caso} \end{cases}$$

El parámetro h_n determina las propiedades estadísticas del estimador en cuanto a la consistencia, sesgo y varianza. Parzen [11], basándose en un teorema que se puede encontrar en Bochner [12], estudia el sesgo, la varianza y la consistencia en un punto x para este tipo de estimadores donde la función núcleo $K(x)$ es una función simétrica y acotada que cumple, además de las condiciones anteriores, que $\lim_{x \rightarrow \infty} |x K(x)| = 0$. Estas condiciones son satisfechas por cualquiera de los núcleos presentados. Parzen [11] presenta estos resultados e ilustra el problema básico de la estimación no paramétrica: una rápida convergencia a cero del parámetro h_n provoca una disminución del sesgo y un aumento de la varianza. Esta conclusión surge de los siguientes resultados:

Teorema 4.1. *Sea $K(y)$ una función Borel acotada que satisface las condiciones*

- $\int_{-\infty}^{\infty} K(z) dz = 1$
- $\int_{-\infty}^{\infty} |K(z)| dz < \infty$
- $\int_{-\infty}^{\infty} |z K(z)| dz < \infty$

Sea $g \in L^1$ y sea

$$g_n(x) = \frac{1}{h_n} \int_{-\infty}^{\infty} K\left(\frac{y}{h_n}\right) g(x-y) dy$$

donde $\{h_n\}$ es una sucesión de constantes positivas que satisfacen que $\lim_{h \rightarrow 0} h_n = 0$. Entonces, si x es un punto de continuidad de g , se cumple que

$$\lim_{n \rightarrow \infty} g_n(x) = g(x) \int_{-\infty}^{\infty} K(y) dy$$

Demostración. Ver [12] □

Corolario 4.1. Si $\{h_n\}$ es una sucesión que cumple que $\lim_{n \rightarrow \infty} h_n = 0$, entonces el estimador \hat{f}_n es asintóticamente insesgado en todos los puntos de continuidad de la función de densidad $f(x)$.

Demostración.

$$\begin{aligned} (4.2) \quad E[\hat{f}_n(x)] &= \frac{1}{n} \sum_{i=1}^n E\left[\frac{1}{h_n} K\left(\frac{x-X_i}{h_n}\right)\right] \\ &= \int_{-\infty}^{\infty} \frac{1}{h_n} K\left(\frac{x-y}{h_n}\right) f(y) dy \end{aligned}$$

Usando que $K_{h_n}(y) = \frac{1}{h_n} K\left(\frac{y}{h_n}\right)$ es una aproximación de la identidad cuando $h \rightarrow 0$, es conocido que en todo punto de continuidad de $f(x)$ vale

$$\lim_{n \rightarrow \infty} E[\hat{f}_n(x)] = f(x) \int_{-\infty}^{\infty} K(y) dy$$

y recordando que la función núcleo cumple que $\int_{-\infty}^{\infty} K(z) dz = 1$ obtenemos el resultado. □

Teorema 4.2. Supongamos que

$$\begin{aligned} \lim_{n \rightarrow \infty} n h_n &= \infty \\ \lim_{n \rightarrow \infty} h_n &= 0 \end{aligned}$$

Entonces el estimador $\hat{f}_n(x)$ es consistente en el sentido que

$$MSE[\hat{f}_n(x)] \xrightarrow{n \rightarrow \infty} 0 \quad \forall x \in \mathbb{R}$$

siendo $MSE[\hat{f}_n(x)]$ el error cuadrático medio definido por

$$MSE[\hat{f}_n(x)] = Var[\hat{f}_n(x)] + Sesgo^2[\hat{f}_n(x)]$$

Demostación.

$$\begin{aligned} Var[\hat{f}_n(x)] &= \frac{1}{n} Var\left[\frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right)\right] \\ &\leq \frac{1}{n} E\left[\left(\frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right)\right)^2\right] \\ &= \frac{1}{n h_n} \left[\frac{1}{h_n} \int_{-\infty}^{\infty} \left(K\left(\frac{x - y}{h_n}\right)\right)^2 f(y) dy \right] \end{aligned}$$

y por el teorema 4.1 aplicado a $K^2(y)$ tenemos que

$$\frac{1}{h_n} \int_{-\infty}^{\infty} \left(K\left(\frac{x - y}{h_n}\right)\right)^2 f(y) dy \xrightarrow{n \rightarrow \infty} f(x) \int_{-\infty}^{\infty} K^2(y) dy$$

ya que $\int_{-\infty}^{\infty} K^2(y) dy < \infty$ y además la convolución es conmutativa. Entonces si $\lim_{n \rightarrow \infty} n h_n = \infty$ tenemos que $Var[\hat{f}_n(x)] \rightarrow 0$. Luego, por el corolario 4.1 tenemos que $MSE[\hat{f}_n(x)]$ es asintóticamente insesgado. \square

El ancho de banda más conveniente debe converger a cero a un ritmo más lento que n^{-1} . Si esto no sucediera, el sesgo tendería a cero pero la varianza aumentaría. Por lo tanto la elección del ancho de banda es un punto importante en la estimación de la función de densidad. Esta elección se realiza de modo que se minimice alguna medida de error. En general se utiliza como medida de error el error cuadrático medio integrado definido por

$$MISE(\hat{f}(x)) = \int_{-\infty}^{\infty} Var[\hat{f}_n(x)] dx + \int_{-\infty}^{\infty} Sesgo^2[\hat{f}_n(x)] dx$$

y se minimiza una aproximación de él.

Antes de entrar en el problema planteado mencionemos que Naradaya (1989) establece las siguientes condiciones para la función núcleo.

Definición 4.1. Diremos que una función núcleo $K(x)$ pertenece a la clase H_s (con $s \geq 2$ par) o diremos que $K(x)$ es un núcleo de orden s si satisface las siguientes

condiciones de regularidad

$$\begin{aligned}
 K(x) &= K(-x) \\
 \int_{-\infty}^{\infty} K(z) dz &= 1 \\
 \sup_{-\infty < x < \infty} |K(x)| &< \infty \\
 \int_{-\infty}^{\infty} x^i K(x) dx &= 0 \quad i = 1, \dots, s-1 \\
 \int_{-\infty}^{\infty} x^s K(x) dx &= k_s \neq 0 \\
 \int_{-\infty}^{\infty} x^s |K(x)| dx &< \infty
 \end{aligned}$$

Notemos que si $K(x)$ es una función de densidad debe verificarse que

$$\int_{-\infty}^{\infty} x^2 K(x) dx > 0$$

Si utilizamos como núcleo una función de densidad simétrica, estamos considerando núcleos de orden 2. En particular las funciones núcleo propuestas como ejemplo son núcleos de orden 2. Estas condiciones nos serán de utilidad para encontrar el ancho de banda óptimo. Si en (4.2) hacemos el siguiente cambio de variable $y = x - h_n t$ tenemos que

$$(4.3) \quad E[\hat{f}_n(x)] = \int_{-\infty}^{\infty} \frac{1}{h_n} K(y) f(x - h_n t) dt$$

y haciendo un desarrollo de Taylor de $f(x - h_n t)$ alrededor de $t = 0$ tenemos que

$$(4.4) \quad f(x - h_n t) = f(x) - f'(x)h_n t + f''(x)\frac{h_n^2 t^2}{2} + \dots$$

reemplazando en (4.2) y usando las condiciones de Nadaraya tenemos que

$$E[\hat{f}_n(x)] = f(x) + \frac{f''(x)h_n^2 k_2}{2} + O(h^4)$$

Por lo tanto el sesgo es

$$\text{sesgo}[\hat{f}_n(x)] = \frac{f''(x)h_n^2 k_2}{2} + O(h^4)$$

Además

$$\begin{aligned} \text{Var}[\hat{f}_n(x)] &= \frac{1}{n} \sum_{i=1}^n \text{Var}\left[\frac{1}{h_n} K\left(\frac{x - X_i}{h_n}\right)\right] \\ &= \frac{1}{n} \left[\frac{1}{h_n^2} \int_{-\infty}^{\infty} \left(K\left(\frac{x-y}{h_n}\right)\right)^2 f(y) dy - \left\{f(x) + \text{sesgo}[\hat{f}_n(x)]\right\}^2 \right] \end{aligned}$$

si hacemos el cambio de variable $y = x - h_n t$ tenemos que

$$\text{Var}[\hat{f}_n(x)] = \frac{1}{n h_n} \int_{-\infty}^{\infty} K^2(t) f(x - h_n t) dt - \frac{1}{n} \left\{f(x) + O(h^2)\right\}^2$$

y usando las condiciones de Naradaya junto con (4.4) tenemos que

$$\text{Var}[\hat{f}_n(x)] \simeq \frac{1}{n h_n} f(x) \int_{-\infty}^{\infty} K^2(t) dt$$

Por lo tanto el *MISE* queda

$$\begin{aligned} (4.5) \quad \text{MISE}(\hat{f}(x)) &\simeq \int_{-\infty}^{\infty} \left[\frac{1}{n h_n} f(x) \int_{-\infty}^{\infty} K^2(t) dt \right] dx + \frac{h_n^4 k_2^2}{4} \int_{-\infty}^{\infty} f''(x)^2 dx \\ &= \frac{1}{n h_n} \int_{-\infty}^{\infty} K^2(t) dt + \frac{h_n^4 k_2^2}{4} \int_{-\infty}^{\infty} f''(x)^2 dx \end{aligned}$$

Buscamos el valor de h_n que minimice esta expresión, y obtenemos

$$(4.6) \quad h_{opt} = \left[\frac{\int_{-\infty}^{\infty} K^2(t) dt}{\int_{-\infty}^{\infty} f''(x)^2 dx} \right]^{\frac{1}{5}} n^{-\frac{1}{5}} k_2^{-\frac{2}{5}}$$

Observemos que $h_n \rightarrow 0$ con una velocidad del orden de $n^{-\frac{1}{5}}$, que es menor que n^{-1} (hecho que se había exigido anteriormente). Además el valor del ancho de banda depende del conocimiento de $f(x)$ que es desconocida para nosotros. Una posibilidad es suponer que $f(x)$ es una distribución normal (con media cero y varianza σ^2) y calcular $\int_{-\infty}^{\infty} f''(x)^2 dx$ con esta suposición. En este caso tenemos que

$$\int_{-\infty}^{\infty} f''(x)^2 dx = \frac{3}{8} \pi^{-\frac{1}{2}} \sigma^{-5} \simeq 0,212 \sigma^{-5}$$

De esta forma, si utilizamos como núcleo la distribución Gaussiana, el h_{opt} será

$$(4.7) \quad h_{opt} = 1,06 \sigma n^{-\frac{1}{5}}$$

donde podemos reemplazar el valor de σ por una estimación de la varianza a partir de los datos. En el caso que la densidad $f(x)$ no se asemeje a una densidad normal, este valor de h_{opt} no será adecuado (Bowman (1985) y Silverman (1986)). Una posible modificación es considerar

$$(4.8) \quad h_{opt} = 1,06 A n^{-\frac{1}{5}}$$

donde $A = \min(\sigma, \frac{Q}{1,349})$ siendo Q el rango intercuartil. Este valor de h_{opt} se comporta bien para densidades unimodales o moderadamente bimodales. Silverman también sugiere otra modificación y propone como valor del ancho de banda a

$$h_{opt} = 0,9 A n^{-\frac{1}{5}}$$

En este trabajo hemos estimado la función de densidad usando los valores de h_{opt} dados por (4.7) y (4.8). Hemos comparado los errores usando (4.5) y nos hemos quedado con el h_{opt} dado por (4.7). En el capítulo 5 explicaremos con más detalle el procedimiento empleado.

Capítulo 5

Estimación de parámetros

En el capítulo 2 vimos que el movimiento browniano no modela adecuadamente a la serie de datos propuestos. Siendo que el movimiento browniano es un caso particular de una familia mas general, la familia de distribuciones de Lévy, nuestra propuesta es determinar qué miembro de esta familia ajusta mejor a la parte central de la serie de índices de mercado propuestos para estudiar. Para ello estimaremos los parámetros α y γ que determinan la distribución y presentaremos los Vuelos Truncados de Lévy como aquella distribución que mejor ajusta a estos datos. La variable estocástica analizada es el retorno $r_{\Delta t, t}$, definida como la diferencia del logaritmo de dos índices de precios consecutivos. Es decir, definimos $r_{\Delta t, t}$ como

$$\begin{aligned} r_{\Delta t, t} &= \log(I_{t+\Delta t}) - \log(I_t) \\ &= \log\left(\frac{I_{t+\Delta t}}{I_t}\right) \end{aligned}$$

donde Δt es la diferencia, en minutos (también la llamaremos escala temporal), entre dos mediciones del índice. Para estimar el exponente característico α y γ usaremos dos propiedades que tienen los vuelos de Lévy:

- i) su autosimilaridad
- ii) la probabilidad del retorno al origen.

5.1. Autosimilaridad

Hemos visto que la función característica de la suma de n variables aleatorias, independientes e idénticamente distribuidas con distribución de Lévy de parámetros α, γ con $\beta = 0$, está dada por

$$\varphi_n(q) = e^{-n\gamma|q|^\alpha}$$

Si llamamos $Y(t) = \log(I_t)$, entonces para $\Delta t = 1$ tenemos que

$$r_{1,t} = Y(t+1) - Y(t)$$

Supongamos que hemos probado que $r_{1,t}$ es una Lévy(α, γ). Entonces, por ejemplo, si $\Delta t = 3$ podemos pensar que

$$\begin{aligned} r_{3,t} &= Y(t+3) - Y(t) \\ &= Y(t+3) - Y(t+2) + Y(t+2) - Y(t+1) + Y(t+1) - Y(t) \\ &= X_3 + X_2 + X_1 \end{aligned}$$

donde las $X_i(t) = Y(t+i) - Y(t+(i-1)) \sim \text{Lévy}(\alpha, \gamma)$, independientes para $i = 1, 2, 3$. Por lo tanto

$$r_{3,t} = S_3 = X_1 + X_2 + X_3$$

De esta forma la función característica de r_{3t} está dada por (ver apéndice estadístico)

$$\varphi_{r_{3t}}(q) = e^{-3\gamma|q|^\alpha}$$

donde el parámetro de escala para r_{3t} es igual a 3γ . En general,

$$r_{\Delta t,t} = S_{\Delta t} = \sum_{i=1}^{\Delta t} X_i$$

donde $X_i(t) = Y(t+i) - Y(t+(i-1)) \sim \text{Lévy}(\alpha, \gamma)$, independientes para $i = 1, \dots, \Delta t$. La función característica de $r_{\Delta t,t}$ será

$$(5.1) \quad \varphi_{r_{\Delta t,t}}(q) = e^{-\Delta t \gamma |q|^\alpha}$$

5.2. Estimación de los parámetros α y γ

Estimar los parámetros que caracterizan la distribución es un punto difícil de abordar si se utilizan métodos que investigan las colas de la distribución, especialmente porque para valores grandes de Δt se reduce el número de datos. Es decir, consideramos que el conjunto de datos es insuficiente para estudiar el comportamiento asintótico de la distribución. Por este motivo haremos un abordaje diferente. Estudiaremos la probabilidad del retorno al origen como función de Δt . Hemos elegido este método para estimar los parámetros de la distribución porque en este punto la función de densidad está menos afectada por el ruido introducido por la finitud de los datos experimentales.

De acuerdo a (3.5) tenemos una relación entre la probabilidad del retorno al origen y los parámetros α y γ dada por

$$f(S_n = 0) = f_{S_n}(0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{(\gamma n)^{\frac{1}{\alpha}} \pi \alpha}$$

y de acuerdo a (5.1) de este capítulo tenemos que

$$(5.2) \quad f_{r_{\Delta t,t}}(0) = f_{S_{\Delta t}}(0) = \frac{\Gamma\left(\frac{1}{\alpha}\right)}{(\gamma \Delta t)^{\frac{1}{\alpha}} \pi \alpha}$$

Esto nos servirá para encontrar los estimadores de los parámetros ya que, si tomamos logaritmos a ambos miembros, tenemos una relación lineal entre el logaritmo de la probabilidad del retorno al origen y el logaritmo de Δt , donde la pendiente es una función de α

$$(5.3) \quad \ln(f_{r_{\Delta t,t}}(0)) = \beta_1 + \beta_2 \ln \Delta t$$

con $\beta_2 = -\frac{1}{\alpha}$ y $\beta_1 = \ln\left(\frac{\Gamma(\frac{1}{\alpha})}{\pi\alpha\gamma^{\frac{1}{\alpha}}}\right)$. Por lo tanto $\hat{\alpha} = -\frac{1}{\hat{\beta}_2}$ y

$$(5.4) \quad \hat{\gamma} = \frac{\Gamma(\frac{1}{\alpha})}{\pi \alpha e^{\beta_1}}$$

Necesitamos conocer la probabilidad del retorno al origen $f_{r_{\Delta t,t}}(0)$. Esto lo haremos estimando la función de densidad $f_{r_{\Delta t,t}}(x)$ según lo visto en el capítulo 4, utilizando métodos no paramétricos a través de núcleos. Un punto importante en la estimación es la elección del ancho de banda. Hemos visto que el ancho de banda óptimo está dado por

$$h_{opt} = \left[\frac{\int_{-\infty}^{\infty} K^2(t) dt}{\int_{-\infty}^{\infty} f''(x)^2 dx} \right]^{\frac{1}{5}} n^{-\frac{1}{5}} k_2^{-\frac{2}{5}}$$

que depende de la función de densidad desconocida para nosotros. Recordemos que si se supone $f \sim N(0, \sigma^2)$, $h_{opt} = 1,06 \sigma n^{-\frac{1}{5}}$. Si f no se asemeja a una distribución normal, Bowman (1985) y Silverman (1986) propusieron considerar $h_{opt} = 1,06 A n^{-\frac{1}{5}}$, con $A = \min(\sigma, \frac{Q}{1,349})$ siendo Q el rango intercuartil. En ambos casos se puede estimar σ con el desvío standard de los datos. En este trabajo consideramos estos dos valores de h_{opt} para estimar $f_{r_{\Delta t,t}}(0)$ y comparamos las estimaciones considerando como error a

$$(5.5) \quad MISE(\hat{f}(x)) = \frac{1}{n h_n} \int_{-\infty}^{\infty} K^2(t) dt + \frac{h_n^4 k_2^2}{4} \int_{-\infty}^{\infty} f''(x)^2 dx$$

tomando núcleo Gaussiano y Epanechnikov

Estudiaremos varios índices de mercado correspondientes a cinco países. Los datos disponibles para estos índices van desde el 21 de Diciembre de 2009 hasta el 5 de Julio de 2010. Estas series están dadas por minuto y se consideraron los datos de cierre. La cantidad de datos utilizados para modelar cada índice es la siguiente

- Merval - 38730 datos

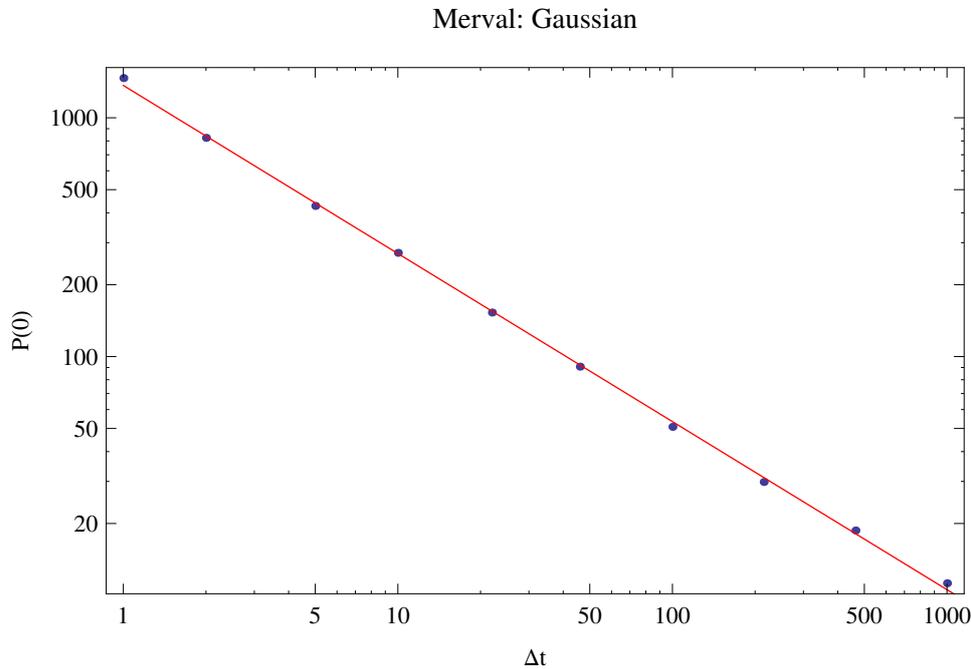


Figura 5.1

- MEXBOL - 52243 datos
- BOVESPA - 56656 datos
- IPSA - 49579 datos
- S&P 500 - 54082 datos

Consideraremos primero el caso de núcleo Gaussiano y luego el caso de núcleo Epanechnikov. En cada caso encontraremos los estimadores de los parámetros y de la función de densidad de los datos analizando cada uno de los h_{opt} propuestos.

5.3. Caso Núcleo Gaussiano y $h_{opt} = 1,06\sigma n^{-\frac{1}{5}}$

Trabajaremos primero con el índice Merval para mostrar la metodología empleada. Luego lo extenderemos a los otros conjuntos de datos. En la figura 5.1 mostramos el ajuste lineal, en un log log plot, de $f_{r_{\Delta t,t}}(0)$ versus Δt . Hemos considerado valores de Δt (medido en minutos) que estén logarítmicamente equiespaciados variando entre 1 y 1000 mín.

A partir del ajuste lineal obtenemos un estimador para α y γ (5.4): $\hat{\alpha} = 1,42$ y $\hat{\gamma} = 6 \times 10^{-6}$ con un error relativo del orden del 1.6% y 15% respectivamente. Vemos experimentalmente que la distribución de los retornos del índice Merval no muestra un comportamiento Gaussiano sino que es compatible con un régimen de Lévy. En la figura 5.2 comparamos la densidad de probabilidad empírica para

$\Delta t = 1$ con la densidad de Lévy de parámetros $\hat{\alpha}$ y $\hat{\gamma}$ donde en la abscisa se reescaló la variable utilizando la dispersión muestral. En la figura 5.3 mostramos la comparación entre las distribuciones respectivas.

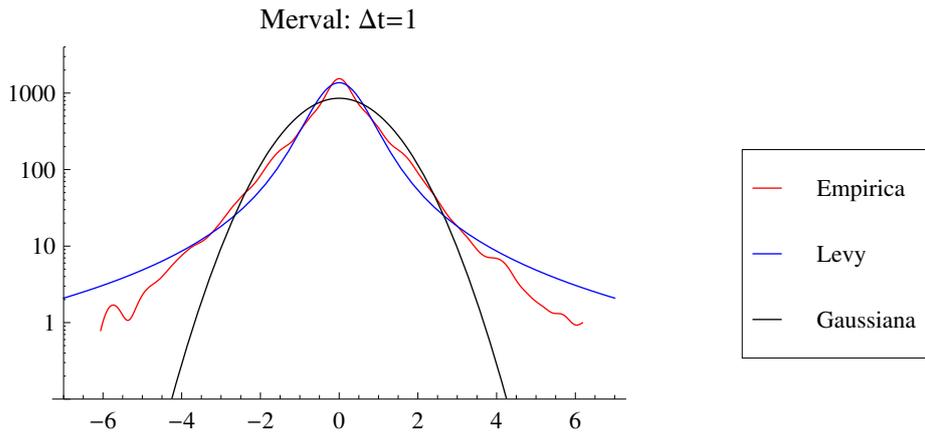


Figura 5.2

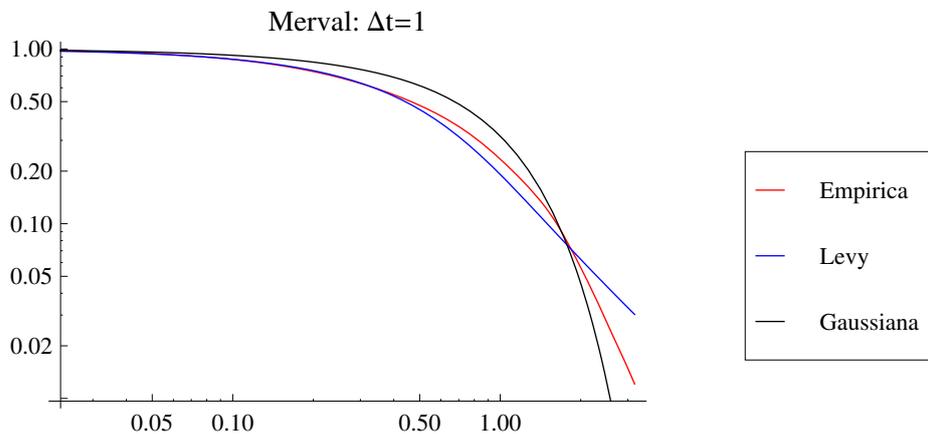


Figura 5.3

Se observa que la distribución de Lévy ajusta bien a los datos para $|r_{\Delta t,t}| \leq 4\sigma$ y también se observa un apartamiento respecto del régimen de Lévy en las colas de la distribución. Este apartamiento se debe a que la distribución de Lévy tiene segundos momentos infinitos si $\alpha < 2$. El comportamiento asintótico de los datos experimentales nos indica que la distribución tiene varianza finita. Por este motivo surge la idea de modelar la serie de los retornos con los Vuelos Truncados de Lévy. Una distribución que responde a una Lévy en la parte central de los datos.

5.4. Una heurística para la elección del h_{opt}

Hemos visto que la elección del ancho de banda óptimo h_{opt} depende de la función de densidad que es desconocida. Nosotros ya tenemos una estimación de la función de densidad y nos preguntamos qué pasaría si utilizáramos esta información para mejorar la estimación. Recordemos que nuestra hipótesis es que nuestros datos se distribuyen con un Vuelo Truncado de Lévy, distribución que coincide con una Lévy en la parte central de los datos. De esta forma la derivada segunda de la distribución de Lévy y de los Vuelos de Lévy no difieren significativamente. Por este motivo

$$f''(x) = \frac{\Gamma\left(\frac{5}{\alpha}\right)}{\alpha\pi(2\gamma)^{\frac{5}{\alpha}}}$$

y de esta forma el h_{opt} queda

$$h_{opt} = \left[\int_{-\infty}^{\infty} K^2(t) dt \right]^{\frac{1}{5}} n^{-\frac{1}{5}} k_2^{-\frac{2}{5}} \left[\frac{\alpha\pi}{\Gamma\left(\frac{5}{\alpha}\right)} \right]^{\frac{1}{5}} (2\gamma)^{\frac{1}{\alpha}}$$

Por otro lado la distribución Gaussiana pertenece a la familia de las distribuciones de Lévy, esto es cuando $\alpha = 2$. En este caso $\gamma = \frac{\sigma^2}{2}$ y, de esta forma, $(2\gamma)^{\frac{1}{\alpha}} = \sigma$. Extrapolando esto, se nos ocurrió investigar cómo daría la estimación de \hat{f} si cambiáramos la elección del h_{opt} . Propusimos un nuevo h_{opt} , que lo llamaremos h^* , dado por

$$h^* = \left[\int_{-\infty}^{\infty} K^2(t) dt \right]^{\frac{1}{5}} n^{-\frac{1}{5}} k_2^{-\frac{2}{5}} \left[\frac{\alpha\pi}{\Gamma\left(\frac{5}{\alpha}\right)} \right]^{\frac{1}{5}} \sigma$$

y tomamos $\alpha = \hat{\alpha}$ y $(2\gamma)^{\frac{1}{\alpha}} = \hat{\sigma}$. Entonces volvimos a estimar $\hat{\alpha}$ con este nuevo valor de ancho de banda y repetimos este proceso hasta que $|\hat{\alpha}_{n+1} - \hat{\alpha}_n| < \varepsilon$. Además se puede elegir entre varios núcleos. Nosotros trabajamos con el núcleo Gaussiano y el Epanechnikov y comparamos las estimaciones obtenidas midiendo (5.5) y nos quedamos con la estimación que nos dio menor error. Es decir, proponemos estimar $\hat{\alpha}$ pensando que tenemos un problema de punto fijo que lo resolvemos de la siguiente forma:

- fijamos el núcleo a usar
- damos valores iniciales para α y γ , tomamos $\alpha_0 = 2$ y $\gamma_0 = \frac{\hat{\sigma}^2}{2}$ (siendo $\hat{\sigma}$ el desvío standard estimado), que corresponde al caso de densidad normal

- para estos valores iniciales obtenemos un valor de $h_K^{\alpha_0, \gamma_0}$. Para el caso de núcleo Gaussiano el ancho de banda es $h_G^{2, \sigma} = 1,06 \hat{\sigma} n^{-1/5}$, mientras que para el caso de núcleo Epanechnikov el ancho de banda es $h_E^{2, \sigma} = 1,049 \hat{\sigma} n^{-1/5}$
- con este valor de $h_K^{\alpha_0, \gamma_0}$ estimamos $f_{r_{\Delta t, t}}(0)$, que lo llamaremos $\hat{f}_{r_{\Delta t, t}}(0)$
- con este valor de $\hat{f}_{r_{\Delta t, t}}(0)$ estimamos un nuevo α que será $\hat{\alpha}_1$ y tomamos $\hat{\gamma}_1 = \gamma_0$
- repetimos el proceso hasta que $|\hat{\alpha}_{n+1} - \hat{\alpha}_n| < \varepsilon$ con ε chico
- con este valor de $\hat{\alpha} = \hat{\alpha}_{n+1}$ estimamos $\hat{\gamma}$ usando (5.4)
- comparamos la estimación de $\hat{f}_{r_{\Delta t, t}}$ con la obtenida en el apartado anterior y estudiamos el error relativo entre ambas estimaciones.
- cambiamos el núcleo y repetimos el proceso.

Trabajamos con el núcleo Gaussiano y Epanechnikov. En ambos casos hicimos 4 iteraciones para lograr que $|\hat{\alpha}_{n+1} - \hat{\alpha}_n| < 10^{-5}$. Luego, para cada Δt , comparamos el error obtenido al estimar $f_{r_{\Delta t, t}}$ usando el α que surge de la cuarta iteración (error 4), con el obtenido utilizando la estimación inicial (error 0). Esta comparación la hicimos estudiando el cociente entre ambos, es decir, analizamos la relación $\frac{\text{Error 4}}{\text{Error 0}}$. Dado que estos cocientes casi no difieren para los distintos valores de Δt , damos un solo valor para cada conjunto de datos estudiado y para cada núcleo utilizado. Esto lo mostramos en la tabla 1:

	Núcleo Gaussiano	Núcleo Epanechnikov
Ibov	0.80	0.80
Ipsa	0.85	0.87
Merval	0.78	0.78
Mexbol	0.79	0.79
Spx	0.86	0.86

Tabla 1

Como se observa obtenemos una mejora entre el 14% y 22% si empleamos esta metodología de estimación, independientemente del núcleo utilizado. También comparamos los errores 4 al utilizar los diferentes núcleos estudiando la relación $\frac{\text{Error 4 Epanechnikov}}{\text{Error 4 Gaussiano}}$. Esto lo mostramos en la tabla 2:

Ibov	0,92
Ipsa	0,76
Merval	0,86
Mexbol	0,94
S&P 500	0,87

Tabla 2

Se puede ver que en todos los casos obtenemos una mejor estimación utilizando el núcleo Epanechnikov, obteniendo una reducción en el error que varía entre el 8 % y el 24 % según el conjunto de datos estudiado. Por lo anteriormente expuesto, para nosotros la mejor estimación se obtiene eligiendo el núcleo Epanechnikov y empleando la heurística propuesta. Con esta elección los valores de los parámetros estimados se muestran en la Tabla 3.

	$\hat{\alpha}$	$\hat{\gamma}$
Ibov	1,54	$2,39 \times 10^{-6}$
Ipsa	1,44	$2,04 \times 10^{-6}$
Merval	1,42	$5,58 \times 10^{-6}$
Mexbol	1,53	$1,45 \times 10^{-6}$
S&P 500	1,62	$1,10 \times 10^{-6}$

Tabla 3

En las figuras 5.4 a 5.8 comparamos, para cada uno de los índices estudiados y para $\Delta t = 1, 2, 5, 10$, los ajustes obtenidos utilizando la heurística propuesta. En las figuras 5.9 y 5.10 mostramos la comparación de la estimación de la función de densidad y de distribución de los diferentes conjuntos de datos utilizando los núcleos Gaussiano y Epanechnikov para $\Delta t = 1, 2, 5, 10$. En la abscisa se reescaló la variable utilizando la dispersión muestral como en los gráficos 5.2 y 5.3

Ibov usando Núcleo Gaussiano

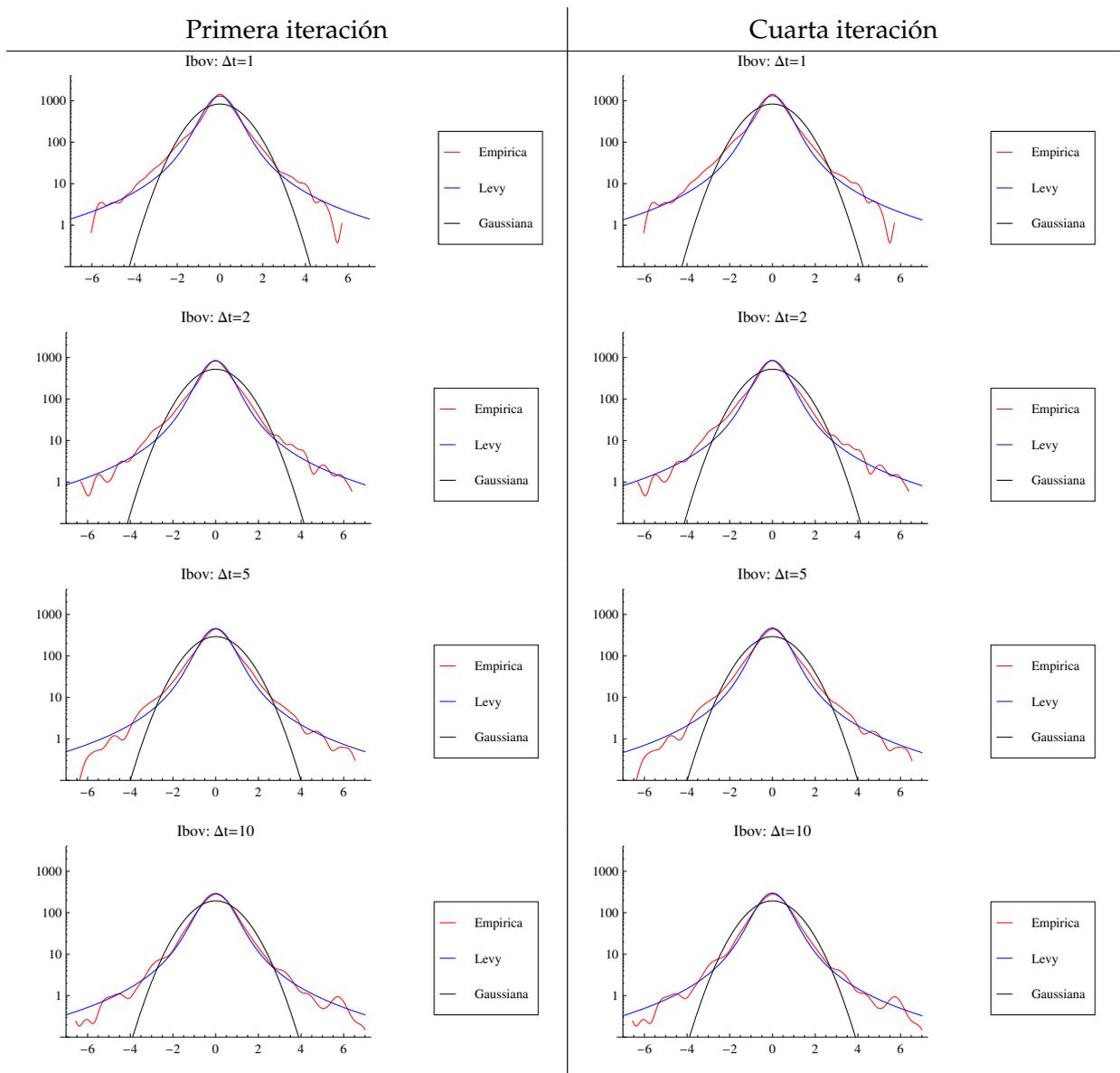


Figura 5.4

Ipsa usando Núcleo Gaussiano

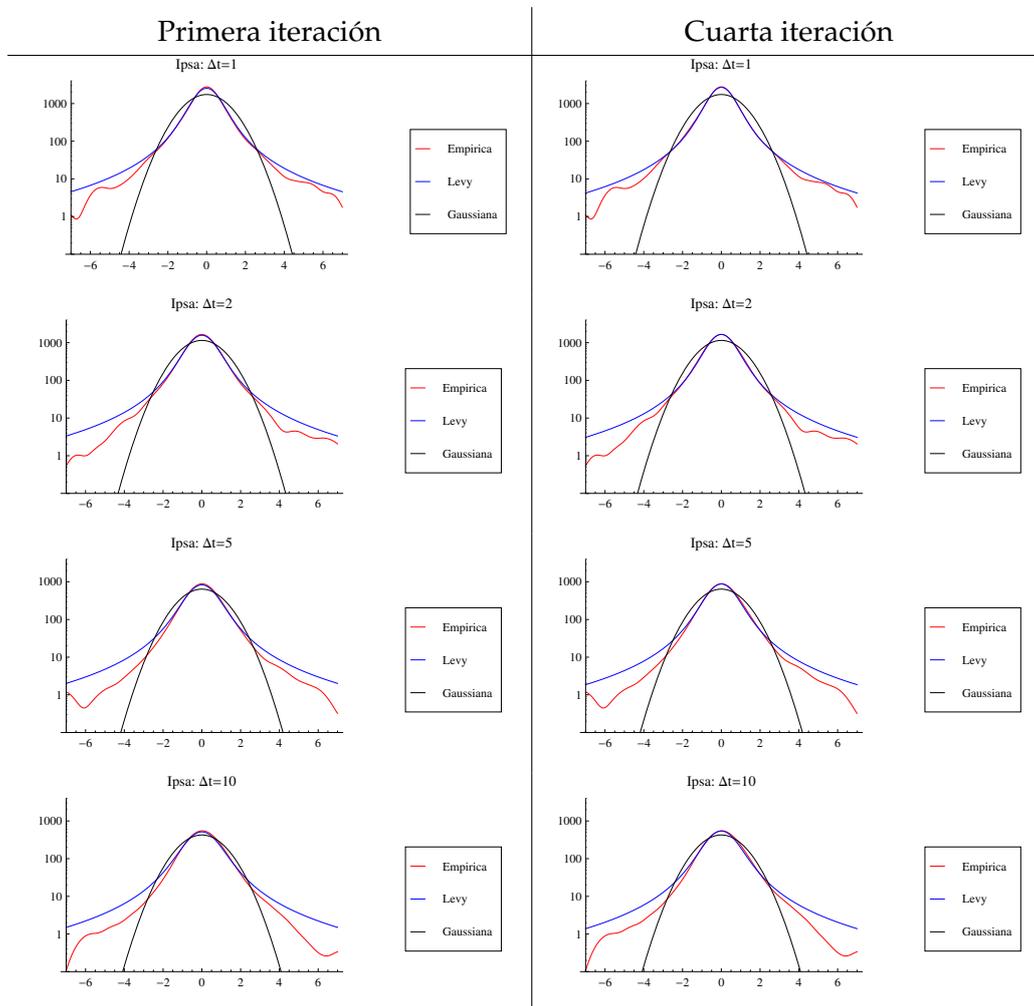


Figura 5.5

Merval usando Núcleo Gaussiano

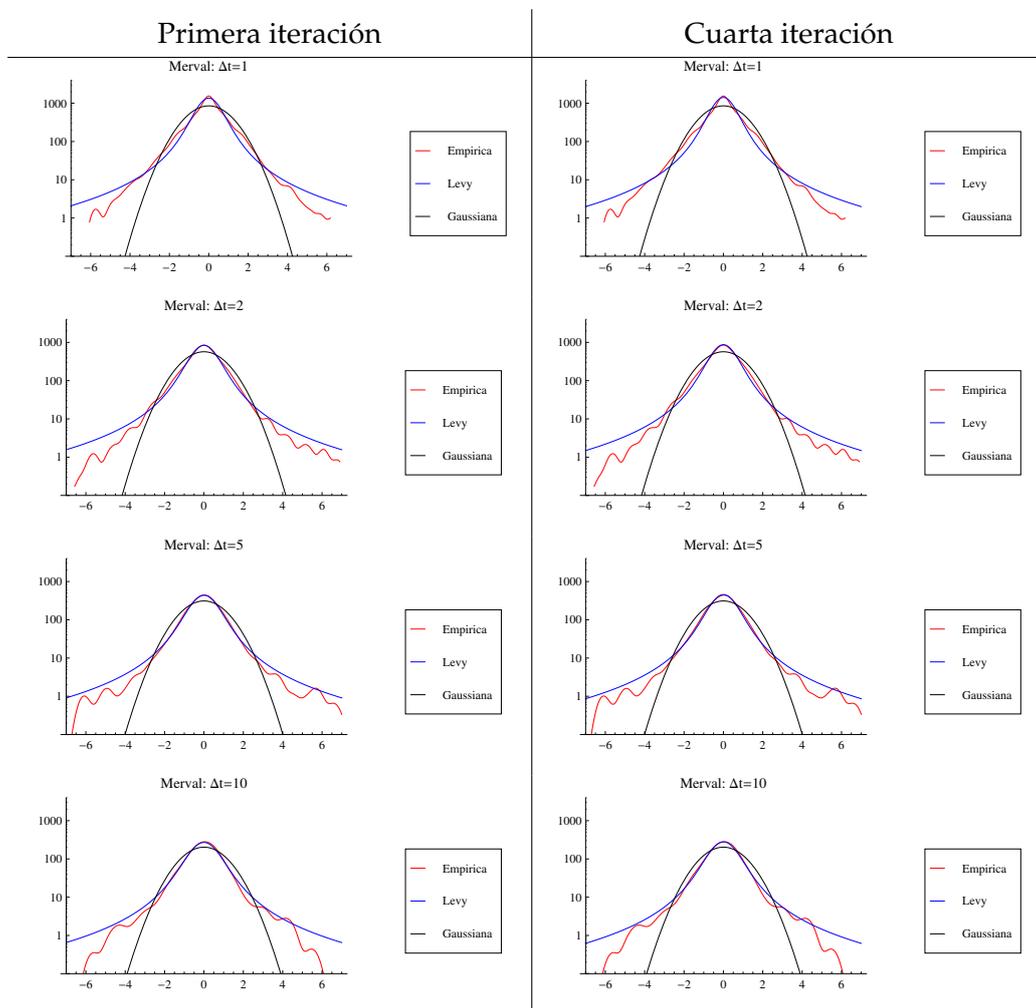


Figura 5.6

Mexbol usando Núcleo Gaussiano

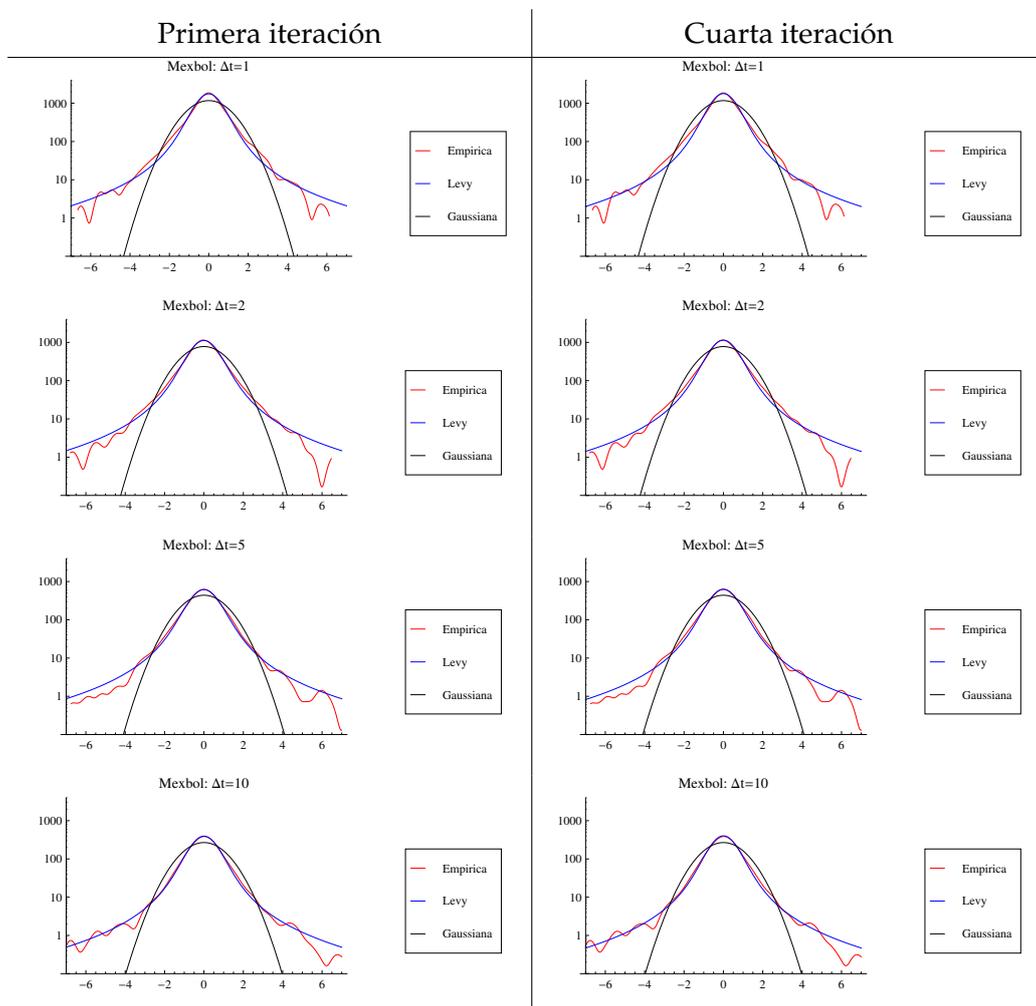


Figura 5.7

Spx usando Núcleo Gaussiano

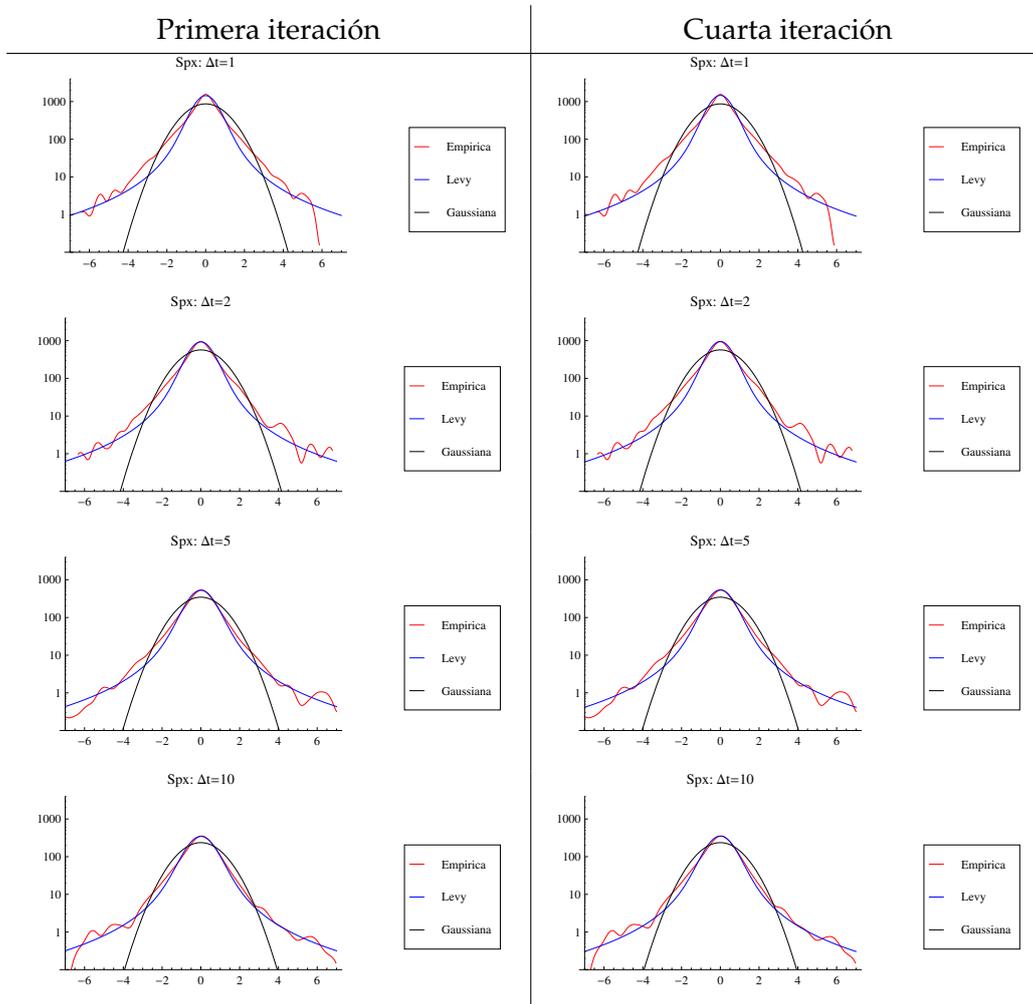


Figura 5.8

Comparación de las densidades para $\Delta t = 1$
con la cuarta iteración para todos los índices

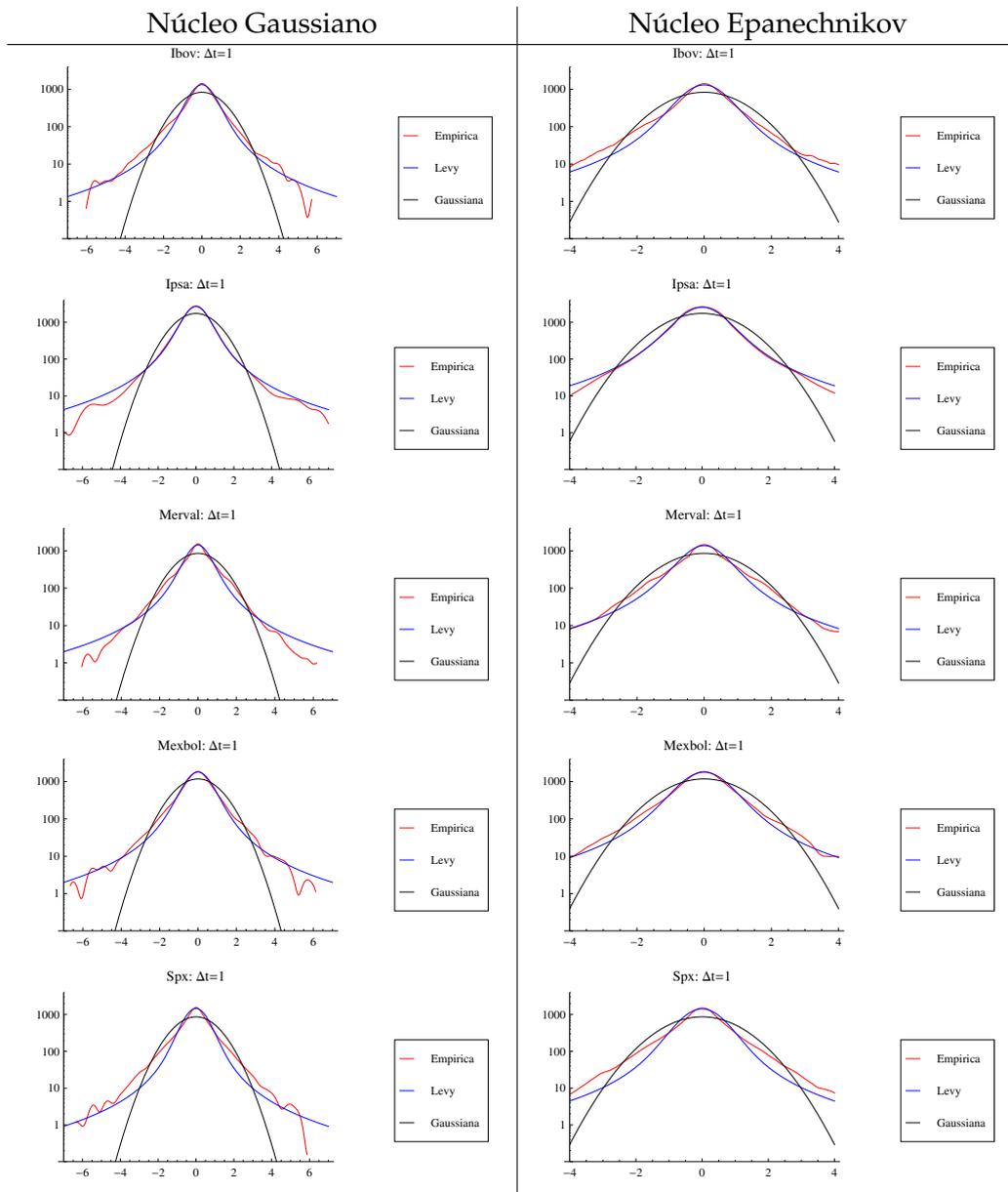


Figura 5.9

Comparación de las distribuciones para $\Delta t = 1$
con la cuarta iteración para todos los índices

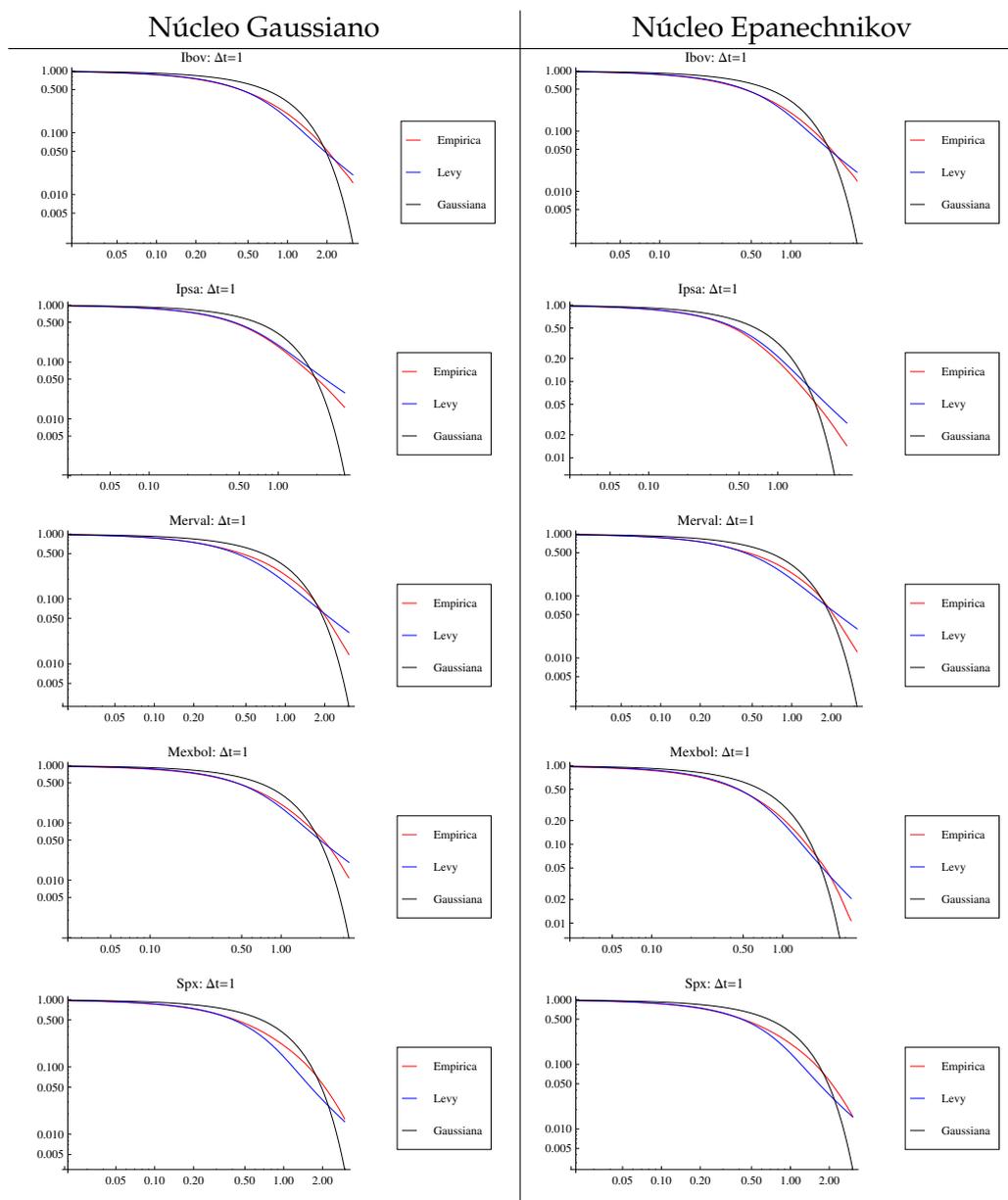


Figura 5.10

Capítulo 6

Conclusiones

Hemos visto que el movimiento browniano no ajusta adecuadamente a la evolución de precios de algunos índices de mercado. Hemos presentado a las distribuciones de Lévy como modelo para ajustar a la parte central de los datos y obtuvimos un mejor ajuste que el dado por el proceso Gaussiano. El problema es que estas distribuciones tiene varianza infinita para valores de $\alpha < 2$ y, por lo tanto, no ajustan bien las colas de la distribución. Para resolver esto presentamos a los Vuelos Truncados de Lévy, distribuciones que coinciden en la parte central de los datos con una distribución de Lévy y cero en las colas. La forma de estas ditribuciones es:

$$f(x) = \begin{cases} 0 & x > l \\ cf_l(x) & -l < x < l \\ 0 & x < -l \end{cases}$$

donde $f_l(x)$ es una distribución de Lévy simétrica y c es una constante de normalización. Estas distribuciones no son estables pero tienen varianza finita. Koponen [8] consideró Vuelos Truncados de Lévy donde el corte de la función de densidad tiene un decaimiento exponencial y está caracterizada por un parámetro l . La función característica de esta distribución está definida por:

$$\varphi(q) = \exp \left\{ c_0 - c_1 \frac{(q^2 + 1/l^2)^{\alpha/2}}{\cos(\pi\alpha/2)} \cos [\alpha \arctan (l|q|)] \right\}$$

con c_0 y c_1 constantes definidas apropiadamente. Koponen propone calcular numéricamente la transformada de Fourier de $\hat{f}(x)$ usando cuadraturas adaptativas diseñadas para integrales oscilantes. Es de interés estudiar la aplicación de este modelo a los datos estudiados.

Apéndice A

Apéndice estadístico

A.1. Algunas definiciones previas

Definición A.1. Un experimento aleatorio es un experimento que se puede repetir bajo las mismas condiciones y su resultado no puede ser predicho con certeza. Uno de los ejemplos más simples de experimento aleatorio es tirar una moneda y por ejemplo, registrar el número de caras que se obtienen cuando se lanza la misma dos veces.

Definición A.2. Un espacio muestral Ω de un experimento aleatorio es el conjunto de todos los resultados posibles de ese experimento. En nuestro ejemplo Ω será:

$$\Omega = \{(C, C); (C, S); (S, C); (S, S); \}$$

Definición A.3. Un evento es un subconjunto de un espacio muestral. En particular cada resultado posible de un experimento aleatorio se llama evento elemental.

Definición A.4. Sea Ω un conjunto no vacío. Una σ -álgebra es una colección G de conjuntos de Ω que cumple con las siguientes propiedades:

- (i) $\emptyset \in G$
- (ii) Si $A \in G$, entonces $A^c \in G$
- (iii) Si A_1, A_2, \dots es una sucesión de conjuntos en G , entonces $\bigcup_{k=1}^{\infty} A_k \in G$

Ejemplos de σ -álgebras. Consideremos el siguiente ejemplo. Supongamos que el precio inicial de un activo financiero es X_0 y supongamos que en cada período (en cada valor de $t = 1, 2, \dots$) el precio del activo puede cambiar a uno de dos valores posibles d y u , con la condición que $0 < d < u$. En el próximo período, el precio del activo será dX_0 o uX_0 . En realidad consideramos $0 < d < 1 < u$ de manera que dX_0 represente una disminución en el precio del activo cuando pasamos al próximo período y uX_0 represente un aumento en el precio del mismo. Es decir podemos pensar que estamos tirando una moneda, si sale cara el valor

del activo será u y si sale seca será d . Luego $X_1(c) = uX_0$ y $X_1(s) = dX_0$. En el segundo período (es decir, después de haber tirado nuevamente la moneda) el valor del activo tomará alguno de estos valores:

$$X_2(CC) = uX_1(C) = u^2X_0 \quad X_2(SC) = dX_1(C) = duX_0$$

$$X_2(CS) = uX_1(C) = udX_0 \quad X_2(SS) = dX_1(C) = d^2X_0$$

Para el tercer período el conjunto de todos los eventos posibles será

$$\Omega = \{CCC, CCS, CSC, CSS, SCC, SCS, SSC, SSS\}$$

Algunas σ -álgebras definidas en este conjunto Ω son: $F_0 = \{\emptyset, \Omega\}$

$$F_1 = \{\emptyset, \Omega, \{CCC, CCS, CSC, CSS\}, \{SCC, SCS, SSC, SSS\}\}$$

$$F_2 = \{\emptyset, \Omega, \{CCC, CCS\}, \{CSC, CSS\}, \{SCC, SCS\}, \{SSC, SSS\},$$

y todas las uniones de estos conjuntos

F_3 es el conjunto de partes de Ω . Podemos interpretar a una σ -álgebra como la información disponible hasta un momento dado. Por ejemplo, F_1 es la σ -álgebra que tiene toda la información hasta $t = 1$, es decir, sabemos qué ocurrió hasta ese momento y no tenemos información que sucede después. Por eso F_1 contiene a los conjuntos \emptyset, Ω , un conjunto formado por la n-uplas cuya primer componente es el resultado de uno de los eventos al tirar una moneda (en este caso cara) y otro conjunto formado por la n-uplas cuya primer componente es el resultado del segundo evento posible (en este caso seca). Luego podemos pensar que la σ -álgebra F_1 contiene la información de la primera tirada de la moneda. Similarmente, F_2 contiene la información hasta el tiempo 2, F_0 no contiene información y F_3 tiene toda la información disponible.

A la menor σ -álgebra que contiene a todos los conjuntos abiertos de \mathbb{R} la llamaremos σ -álgebra de Borel \mathfrak{B} y a un elemento de \mathfrak{B} lo llamaremos boreliano y lo denotaremos con B .

Definición A.5. Sea Ω un conjunto no vacío, F una σ -álgebra en Ω . Una medida de probabilidad P es una función definida sobre la σ -álgebra F $P : F \rightarrow [0, 1]$ que cumple:

- $P(\Omega) = 1$
- $P(\emptyset) = 0$
- Si A_1, A_2, \dots es una sucesión de conjuntos disjuntos en F , entonces

$$P(\cup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k)$$

Podemos dar la siguiente interpretación de una medida de probabilidad. Supongamos que Ω es el conjunto de todos los resultados posibles de algún experimento aleatorio, F una σ -álgebra definida en Ω y sea A un elemento de F . Hay una cierta probabilidad, entre 0 y 1, que cuando el experimento se realice el resultado del experimento pertenezca a A . Luego $P(A)$ será esa probabilidad.

Definición A.6. Llamaremos un espacio de probabilidad a la terna formada por (Ω, F, P) , con Ω un conjunto no vacío, F una σ -álgebra definida sobre Ω y P una medida de probabilidad sobre F .

A.2. Variables Aleatorias

Definición A.7. Sea (Ω, F, P) un espacio de probabilidad. Una variable aleatoria X es una función $X : \Omega \rightarrow \mathbb{R}$ medible. Es decir es una función que cumple que $X^{-1}(B) \in F$ para todo boreliano $B \in \mathfrak{B}$.

Definición A.8. Definimos a la función de distribución de X por

$$F_X(a) = P(\omega \in \Omega : X(\omega) \leq a) = P(X \leq a), \quad a \in \mathbb{R}$$

Definición A.9. Sea X_1, \dots, X_n variables aleatorias. Definimos la función de distribución conjunta de dichas variables a

$$F_{X_1, \dots, X_n}(a_1, \dots, a_n) = P(X_1 \leq a_1, \dots, X_n \leq a_n)$$

Definición A.10. Sea X una variable aleatoria y sea A un evento, es decir, un subconjunto de Ω . La función de distribución de X condicional a la realización del evento A para el cual $P(A) > 0$ se define como

$$F_X(x/A) = \frac{P[\{X \leq x\} \cap A]}{P(A)}$$

A.2.1. Variables aleatorias discretas

Definición A.11. Diremos que una variable aleatoria X es discreta si el conjunto de posibles valores de la misma es finito o numerable. Para este caso la función de distribución queda definida como

$$F_X(a) = \sum_{x_i \leq a} P(X = x_i)$$

Ejemplo A.1. Una variable aleatoria X se dice que tiene una distribución Bernoulli con parámetro p (donde p es la probabilidad de éxito) si su función de probabilidad es

$$p(x) = P(X = x) = p^x(1 - p)^{1-x} \quad \Omega = \{0, 1\}$$

Ejemplo A.2. Una variable aleatoria X se dice que tiene una distribución Binomial con parámetros n y p , y se denota $X \sim Bi(n, p)$ si su función de distribución viene dada por

$$p(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{1-x}$$

con $\Omega = \{0, 1, \dots, n\}$ y p es la probabilidad de éxito.

Ejemplo A.3. Una variable aleatoria X que toma valores $0, 1, 2, \dots$ se dice que es de Poisson con parámetro λ , si $\forall \lambda > 0$ la función de probabilidad puntual está dada por

$$p(x) = P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!} \quad \Omega = \{0, 1, 2, \dots\}$$

La notación en este caso es $X \sim P(\lambda)$.

A.2.2. Variables aleatorias continuas

Definición A.12. Una variable aleatoria X es continua si existe una función $f(x) \geq 0$ definida $\forall x \in \mathbb{R}$ que verifica que

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

En este caso diremos que $f(x)$ es la función de densidad de X .

Ejemplo A.4. Una variable aleatoria se dice que tiene una distribución uniforme (o que está uniformemente distribuida) sobre el intervalo $(0, 1)$ si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{en otro caso} \end{cases}$$

La notación en esta caso es $X \sim U[a, b]$

Ejemplo A.5. Diremos que la variable aleatoria X tiene una distribución exponencial con parámetro $\lambda > 0$ si su función de densidad está dada por

$$f(x) = \lambda e^{-\lambda x} \quad \text{para } x \geq 0$$

La notación en esta caso es $X \sim Exp(\lambda)$

Ejemplo A.6. Diremos que la variable aleatoria X está normalmente distribuida con parámetros μ y σ^2 (notación $X \sim N(\mu, \sigma)$) si su función de densidad está dada por

$$f(x) = \frac{1}{2\pi\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

La notación en este caso es $X \sim N(\mu, \sigma^2)$. Muchas veces se usa la notación

$$\Phi(x) = \frac{1}{2\pi} \int_{-\infty}^x e^{-\frac{x^2}{2}}$$

para indicar que $X \sim N(0, 1)$.

Ejemplo A.7. Una variable aleatoria se dice que tiene una distribución de Cauchy con parámetro θ , $-\infty < \theta < \infty$, si su función de densidad está dada por

$$f(x) = \frac{1}{\pi} \frac{1}{1 + (x - \theta)^2} \quad -\infty < x < \infty$$

Ejemplo A.8. Una variable aleatoria se dice que tiene una distribución gamma con parámetros (α, λ) , $\lambda > 0$ y $\alpha > 0$ si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

con $\Gamma(\alpha) = \int_0^\infty e^{-y} y^{\alpha-1} dy$

A.3. Independencia

Definición A.13. Sea (Ω, F, P) un espacio de probabilidad. Dos eventos $A, B \in F$ son independientes si

$$P(A \cap B) = P(A)P(B)$$

Definición A.14. Dos familias de conjuntos medibles \mathcal{U}, \mathcal{G} son independientes si

$$P(A \cap B) = P(A)P(B) \text{ para todo } A \in \mathcal{U}, B \in \mathcal{G}.$$

Definición A.15. La colección finita de variables aleatorias X_1, X_2, \dots, X_n sobre (Ω, F, P) son mutuamente independientes si para todo conjunto Borel C_1, C_2, \dots, C_n ,

$$P(X_i \in C_i, i = 1, \dots, n) = \prod_{i=1}^n P(X_i \in C_i)$$

donde

$$P(X_i \in C_i) \equiv P(\omega \in \Omega : X_i(\omega) \in C_i)$$

En particular si $C_i = (-\infty, a_i]$, $a_i \in \mathbb{R} \quad \forall i = 1, \dots, n$, entonces X_1, X_2, \dots, X_n serán independientes si

$$P(X_1(\omega) \leq a_1, \dots, X_n(\omega) \leq a_n) = \prod_{i=1}^n P(X_i(\omega) \leq a_i)$$

es decir la función de distribución conjunta cumple que

$$F_{X_1, \dots, X_n}(a_1, \dots, a_n) = \prod_{i=1}^n F_{X_i}(a_i)$$

Definición A.16. La colección infinita de variables aleatorias X_1, X_2, \dots son mutuamente independientes si cualquier colección finita de ellas es mutuamente independiente.

A.4. Variables aleatorias independientes y convolución

Teorema A.1. Sean X_1 y X_2 dos variables aleatorias i.i.d. Sea $S_2 = X_1 + X_2$. Entonces la función de densidad $f(S_2)$ está dada por la convolución de ellas. Es decir

$$f(S_2) = f(X_1) \otimes f(X_2)$$

Demostración. la función de distribución de la suma de dos variables aleatorias está dada por

$$F_{X_1+X_2}(z) = \Pr(X_1 + X_2 \leq z) = \Pr(X_1 + X_2 \in B_z), \text{ con } B_z = \{(x, y)/x + y \leq z\}$$

Luego

$$F_{X_1+X_2}(z) = \int \int_{B_z} f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{z-y} f(x_1, x_2) dx_1 dx_2$$

siendo $f(x_1, x_2)$ la función de densidad conjunta de (X_1, X_2) . Si hacemos el cambio de variables

$$\begin{aligned} s &= x_1 + x_2 \\ t &= x_2 \end{aligned}$$

usando el teorema de cambio de variables en una integral doble y considerando que el jacobiano de la transformación es 1, tenemos que

$$F_{X_1+X_2}(z) = \int_{-\infty}^{\infty} \int_{-\infty}^z f(s-t, t) ds dt = \int_{-\infty}^z \int_{-\infty}^{\infty} f(s-t, t) dt ds$$

Luego la función de densidad de $S_2 = X_1 + X_2$ es

$$f(S_2) = \int_{-\infty}^{\infty} f(s-t, t) dt$$

Pero como X_1 y X_2 son independientes, tenemos

$$f(S_2) = \int_{-\infty}^{\infty} f_{x_1}(s-t) f_{x_2}(t) dt = f_{x_1} \otimes f_{x_2} \quad \square$$

Teorema A.2. La transformada de Fourier de la convolución de dos funciones de L^1 es el producto de las transformadas de Fourier de cada una de ellas. Es decir,

$$F(X_1 \otimes X_2) = F(X_1)F(X_2)$$

siendo F la transformada de Fourier de una función.

Demostración. Sean $f(x)$ y $g(x)$ dos funciones absolutamente integrables, entonces

$$\begin{aligned} F(f(x) \otimes g(x)) &= \int_{-\infty}^{\infty} (f(x) \otimes g(x)) e^{iqx} dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x-t) g(t) dt e^{iqx} dx = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(z) g(t) e^{iqz} e^{iqt} dt dz = \\ &= \int_{-\infty}^{\infty} f(z) e^{iqz} dz \int_{-\infty}^{\infty} g(t) e^{iqt} dt = \\ &= F(f(x)) F(g(x)) = F(q) G(q) \quad \square \end{aligned}$$

A.5. Esperanza matemática

Definición A.17. La esperanza matemática de una variable aleatoria X con función de densidad dada por $f(x)$ se define como

$$E(X) = \sum_{k=1}^{\infty} x_k p_X(x_k) \quad \text{para el caso discreto}$$

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) \quad \text{para el caso continuo}$$

Notemos que el operador matemático E es lineal.

Definición A.18. El momento de orden k de una variable aleatoria X con función de densidad $f(x)$ está dado por

$$E(X^k) \quad k = 0, 1, 2, \dots$$

Definición A.19. La varianza de una variable aleatoria X se define como

$$V(X) = E[(X - E(X))^2]$$

Notemos que si $E(X) = 0$ la varianza de X es el momento de orden 2 de X .

A.6. Función característica

Definición A.20. Sea X una variable aleatoria con función de densidad dada por $P(X)$. La función característica de X es una función $\varphi : \mathbb{R} \rightarrow \mathbb{C}$ definida por

$$\varphi_X(q) = E(e^{iqX})$$

siendo E el operador esperanza de una variable aleatoria y $e^{itX} = \cos(tX) + i \sin(tX)$ con $q \in \mathbb{R}$

Observación A.1. Si X es una variable aleatoria continua la función característica es la transformada de Fourier de su función de densidad

$$\begin{aligned} \varphi_X(q) &= E(e^{iqX}) \\ &= \int_{-\infty}^{\infty} e^{iqx} f(x) dx \\ &= \int_{-\infty}^{\infty} f(x) e^{iqx} dx = F(f(X)) \end{aligned}$$

Por otro lado si antitransformamos obtenemos

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi(q) e^{-iqx} dq$$

que es la función de densidad de X .

Dado que la transformada de Fourier es única, la función $\varphi_X(q)$ caracteriza a la variable aleatoria X . Por ejemplo, hay una única distribución Gaussiana standard cuya función característica es $\varphi_X(q) = e^{-\frac{q^2}{2}}$.

Propiedades

a) $E(X^q) = i^{-q} \frac{\partial^q \varphi(k)}{\partial^q k} \Big|_{k=0} \quad q = 1, 2, \dots$

b) $|\varphi(k)| \leq \varphi(k=0) = 1$

Demostración.

$$\begin{aligned} |\varphi(k)| &= \left| E(e^{ikX}) \right| \\ &\leq \int_{-\infty}^{\infty} |e^{ikx}| f(x) dx \\ &= \int_{-\infty}^{\infty} f(x) dx = 1 \\ &= \varphi(k=0) \quad \square \end{aligned}$$

- c) Sean X_1, \dots, X_n variables aleatorias i.i.d. Sea $S_n = X_1 + \dots + X_n$. Entonces la función característica $\varphi(S_n)$ está dada por el producto de las funciones características.

Demostración.

$$(A.1) \quad \begin{aligned} \varphi_n(q) &= F(f(X_1 + \dots + X_n)) = F(f(X_1) \otimes \dots \otimes f(X_n)) \\ &= F(f(X_1)) \dots F(f(X_n)) = \varphi(q) \dots \varphi(q) = \varphi(q)^n \end{aligned} \quad \square$$

A.6.1. Estabilidad y función característica

En esta sección mostraremos que tanto la distribución de Cauchy como la distribución Gaussianas son estables

- La distribución de Cauchy es estable. Los haremos para el caso $n = 2$.

Sean X_1 y X_2 dos variables aleatorias i.i.d. con distribución de Cauchy. Definimos $S_2 = X_1 + X_2$. Queremos ver que $f(S_2)$ también es de Cauchy. Sea φ_2 la función característica de $X_1 + X_2$. Luego por (A.1)

$$\varphi_2(q) = \varphi(q)^2$$

Recordemos que la función de distribución de Cauchy es

$$f(x) = \frac{\gamma}{\pi} \frac{1}{\gamma^2 + x^2}$$

De esta forma la transformada de Fourier (F) para la función de densidad es

$$\varphi(q) = \int_{-\infty}^{\infty} \frac{\gamma}{\pi} \frac{1}{\gamma^2 + x^2} e^{iqx} dx$$

Resolviendo la integral (ver) queda:

$$\varphi(q) = e^{-\gamma|q|}$$

Luego

$$\varphi_2(q) = (\varphi(q))^2 = e^{-2\gamma|q|}$$

Dado que $\varphi(q) \in L^1$ podemos antitransformar obteniendo la función de densidad correspondiente a esa función característica. Así en este caso la

función de densidad de S_2 será:

$$\begin{aligned}
 f(S_2(x)) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_2(q) e^{-iqx} dq \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-2|q|\gamma} e^{-iqx} dq \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-2|q|\gamma} (\cos qx - i \sin qx) dq \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-2|q|\gamma} \cos qx dq - \frac{i}{2\pi} \int_{-\infty}^{\infty} e^{-2|q|\gamma} \sin qx dq \\
 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-2|q|\gamma} \cos qx dq \\
 &= \frac{1}{\pi} \int_0^{\infty} e^{-2q\gamma} \cos qx dq \\
 &= \frac{1}{\pi} \frac{2\gamma}{4\gamma^2 + x^2}
 \end{aligned}$$

Entonces vemos que la forma funcional de $f(S_2)$ es de Cauchy. Esto nos dice que esta distribución es **estable**.

- Aplicando un razonamiento similar se puede ver que la distribución Gaussiana es estable. En efecto, si X es una variable aleatoria con distribución gaussiana, la función de densidad es

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

Su función característica es

$$\varphi(q) = e^{-\frac{\sigma^2}{2}q^2} = e^{-\gamma q^2}$$

donde $\gamma = \frac{\sigma^2}{2}$. Luego la función característica de S_2 es

$$\varphi_2(q) = e^{-2\gamma q^2}$$

y antitransformando, obtenemos

$$f(S_2) = \frac{1}{\sqrt{8\pi\gamma}} e^{-\frac{x^2}{8\gamma}}$$

Por lo tanto la distribución gaussiana es también una distribución estable, este último término lo reescribimos

$$f(S_2) = \frac{1}{\sqrt{2\pi} \sqrt{2}\sigma} e^{-\frac{x^2}{2(\sqrt{2}\sigma)^2}}$$

y definimos

$$\sigma_2 = \sqrt{2}\sigma$$

Así hemos probado que los procesos estocásticos de Cauchy y Gaussiano son estables. Mas aún, la función característica de ambos procesos tiene la misma forma funcional:

$$\varphi_2(q) = e^{-2\gamma|q|^\alpha}$$

donde $\alpha = 1$ (Cauchy) y $\alpha = 2$ para el proceso Gaussiano.

A.7. Teorema Central del límite

Sean X_1, X_2, \dots una sucesión infinita de variables aleatorias independientes, idénticamente distribuidas (i.i.d), sea $S_n = X_1 + X_2 + \dots + X_n$. Si $E(X_1) = \mu \in \mathbb{R}$ y $Var(X_1) = \sigma^2 \in (0, \infty)$ se verifica que

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq z\right) = P(Y \leq z)$$

siendo Y una variable aleatoria con distribución normal de media 0 y varianza 1 ($Y \sim N(0, 1)$). Esto quiere decir lo siguiente:

$$S_n \sim N(n\mu, n\sigma^2)$$

$$\frac{S_n}{n} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

En general, para $n \geq 30$ la aproximación por la distribución Gaussiana es bastante buena.

A.8. Aproximación gaussiana a una variable aleatoria binomial

Sean z_1 y z_2 dos números fijos y sea $X \sim \text{Bi}(n, p)$. Entonces si $Z \sim N(np, np(1-p))$ se cumple que

$$P(X \leq k) \approx P(Z \leq k)$$

De la misma forma, si llamamos $X^* = \frac{X - np}{\sqrt{np(1-p)}}$ se cumple que

$$P(X^* \leq k) \rightarrow \Phi(k) \text{ cuando } n \rightarrow \infty$$

Hay que tener en cuenta que al aproximar una distribución binomial por una normal estamos aproximando la distribución de una variable aleatoria discreta

con la distribución de una variable aleatoria continua. Por lo tanto hay que tener algún cuidado con los puntos extremos del intervalo considerado. Por ejemplo, para una variable aleatoria continua $P(X = 3) = 0$, mientras que para una variable aleatoria discreta esta probabilidad puede ser positiva. Hay una corrección que se puede hacer que mejora esta aproximación

$$P(X = k) \simeq P\left(k - \frac{1}{2} \leq X \leq k + \frac{1}{2}\right)$$
$$P(a \leq X \leq b) \simeq P\left(a - \frac{1}{2} \leq X \leq b + \frac{1}{2}\right)$$

Bibliografía

- [1] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Polit. Econ. 81, 637-654 (1973).
- [2] P. Willmott, J. N. Dewynne, S. D. Howison, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, Oxford 1993.
- [3] R. N. Mantegna, H. E. Stanley, *Scaling behaviour in the dynamics of an economic index*, Nature 376 (1995) 46.
- [4] B. B. Mandelbrot, *The variation of certain speculatives prices*, J. Business 36, 394 (1963).
- [5] P. Lévy, *Théorie de l'Addition des Variables Aléatoires*, Gauthier-Villars, París, 1937.
- [6] R. N. Mantegna, H. E. Stanley, *Stochastic process with ultraslow convergence to a gaussian: The Truncated Lévy Flight*, Phys. Rev. Lett. 73 (1994) 2946.
- [7] S. V. Buldyrev, A. L. Goldberger et al, *Long range correlation properties of coding an noncoding DNA sequences: GenBankanalysis*, Phys. Rev. E, 51, 5084-5091 (1995).
- [8] I. Koponen, *Analytic approach to the problem of convergence of truncated Lévy flights towards the Gaussian stochastic process*, Phys. Rev. E, 52, 1197-1199 (1995).
- [9] R. N. Mantegna, *Fast, accurate algorithm for numerical simulation of Lévy stable stochastic processes*, Phys. Rev. E 49, 4677-4683 (1994)
- [10] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol I y II*, John Wiley and Sons, New York, 1970
- [11] E. Parzen, *On estimation of a probability density function and mode*, Annals of Mathematical Statistics 33(3), 1065–1076, 1962
- [12] S. Bochner, *Harmonic Analysis and the Theory of Probability*, Univ. of California Press, 1995.
- [13] H. Bergström, *On Some Expansions of Stable Distributions*, Ark. Mathematicae II 18, 375-378, 1952