



CIMAT

Inferencia no paramétrica sobre procesos de Lévy vía la función característica empírica

TESINA

Que para obtener el grado de
Maestro en Ciencias con Especialidad en Probabilidad y
Estadística

Presenta

Guillermo Basulto Elías

Directores de Tesina

Dr. Miguel Nakamura Savoy
Dr. Víctor Manuel Pérez Abreu Carrión

Guanajuato, Guanajuato. 15 de Julio de 2011.

Integrantes del jurado

Presidente: Dr. Víctor Manuel Rivero Mercado

Secretario: Dra. Angélica Hernández Quintero

Vocal: Dr. Miguel Nakamura Savoy

Lector especial: Dr. Víctor Manuel Pérez Abreu
Carrión

Asesores:

Dr. Miguel Nakamura Savoy

Dr. Víctor Manuel Pérez Abreu Carrión

Sustentante:

Guillermo Basulto Elías

*A mi mamá y hermanas,
Lety, Diana y Viry*

Agradecimientos

A mis asesores, los Dres. Miguel Nakamura Savoy y Víctor Manuel Pérez Abreu, por su paciencia, su gran dedicación y sus sabios consejos.

A mis sinodales, los Dres. Víctor Manuel Rivero Mercado y Angélica Hernández Quintero por sus valiosas observaciones.

A mi mamá y mis hermanas, por estar siempre conmigo brindándome su cariño y apoyo incondicional.

A mis profesores, por su entusiasmo y disposición para llevar a cabo su labor de la enseñanza de la mejor manera.

A mis amigos y compañeros, por compartir sus conocimientos y experiencias, haciendo más amena y provechosa esta etapa. Un agradecimiento especial a mi amigo Martín, por sus consejos para la parte computacional de este trabajo.

Al CONACyT, por el apoyo económico proporcionado, y al CIMAT, por todas las facilidades para asistir a congresos y conferencias que enriquecieron mi formación.

Contenido

Introducción	1
1 Preliminares	7
1.1 Procesos de Lévy y fórmula de Lévy-Khintchine	7
1.2 Función característica empírica	10
1.2.1 Estimador de mínima distancia	12
1.2.2 Procedimiento k-L	13
1.2.3 Estimador de momentos	14
2 Estimación de la medida canónica de Kolmogorov	17
2.1 Con inversión de Fourier	18
2.2 Con mínima distancia a la función característica empírica	21
2.3 Con mínima distancia a la función característica empírica considerando la frecuencia de observación	26
3 Estimación de la densidad de Lévy	31
3.1 Con inversión de Fourier en términos de la deriva, la intensidad y la difusión	32
3.2 Con inversión de Fourier en términos de la difusión	36
4 Discusión	41
5 Estimador tipo <i>sieves</i> basado en método de momentos	45
5.1 Método <i>sieves</i>	46

5.2	Método de momentos	49
5.2.1	Normalidad asintótica	50
6	Implementación, nueva propuesta y bandas de confianza	53
6.1	Comentarios sobre cálculo e implementación de estimadores	54
6.2	Implementación del estimador tipo sieves basado en momentos	56
6.3	Otra propuesta de <i>sieves</i>	58
6.4	Bandas <i>bootstrap</i> de confianza	62
A	Procesos de Lévy	65
A.1	Divisibilidad infinita y procesos de Lévy	65
A.2	Descomposición de Lévy-Ito	66
A.3	Ejemplos de procesos de Lévy	69
A.4	Funciones de centro	71
A.5	Momentos de un proceso de Lévy	73
A.6	Actividad finita o infinita	74
A.6.1	Ejemplos	75
B	Otros resultados	77
B.1	Propiedades de las funciones características	77
B.2	Base de ondículas de Haar	79
	Notación	81
	Bibliografía	83

Introducción

Los procesos de Lévy son una familia de procesos estocásticos de gran importancia con muchas aplicaciones, siendo el movimiento browniano, los procesos de Poisson compuestos y los procesos α -estables los ejemplos más conocidos. Actualmente, los procesos de Lévy se usan para modelación en diferentes áreas, tales como finanzas, teoría de riesgo, teoría de colas, física, control óptimo, etcétera. En Cont y Tankov [8] y Barndorff-Nielsen, Mikosch y Resnick [2] se pueden encontrar modelos aplicables a estas áreas que ilustran la pertinencia de los procesos de Levy en este tipo de aplicaciones.

Este trabajo versa sobre la estimación no paramétrica en procesos de Lévy usando como herramienta principal a la función característica empírica. La función característica está en correspondencia biyectiva con la función de distribución. La función característica empírica tiene varias propiedades que la hacen un buen estimador de la función característica teórica; por ejemplo, es insesgada, asintóticamente normal y fuertemente consistente, tanto de manera puntual, como uniformemente, entre otras. Una ventaja adicional de la función característica empírica para hacer inferencia sobre procesos de Lévy es que la función característica, a través de la fórmula de Lévy-Khintchine, descompone de manera muy explícita cada proceso de Lévy como la suma de dos procesos de Lévy independientes: un movimiento browniano con deriva y un proceso de puros saltos. Por otro lado, los procesos de Lévy incluyen algunos ejemplos importantes cuya función de distribución no puede escribirse analíticamente de una manera sencilla, como es el caso de los procesos de Lévy α -estables.

En muchas de las aplicaciones mencionadas, es usual que componentes específicas del

proceso de Lévy no sean enteramente conocidas. Cuando esto ocurre, son observaciones del fenómeno quienes proporcionan información al respecto, haciendo evidente la relevancia de la inferencia estadística en procesos estocásticos; en particular, de la inferencia no paramétrica, como se verá más adelante. La fórmula de Lévy-Khintchine indica que todo proceso de Lévy está caracterizado por dos parámetros reales (de deriva y difusión) y una medida, la medida de Lévy. En algunos procesos de Lévy, la fórmula de Lévy-Khintchine puede caracterizarse por un número finito de parámetros, como es el caso del movimiento browniano, los procesos de Poisson, los procesos gama, los procesos gaussianos inversos y los procesos α -estables, entre otros.

Es deseable hacer inferencias sobre procesos de Lévy a través de la fórmula de Lévy-Khintchine sin asumir que esta puede caracterizarse por un número finito de parámetros. Lo anterior ocurre cuando quien desea aplicar un modelo de procesos de Lévy, no cuenta con conocimiento contextual suficiente para especificar un modelo paramétrico particular; o bien, cuando se desea hacer un análisis exploratorio de datos para fines de discernir un modelo finito paramétrico. A raíz de esto, se motiva la pertinencia o necesidad de estadística no paramétrica para estimar esta medida.

La medida de Lévy puede estimarse a través de la medida canónica de Kolmogorov o mediante la densidad de Lévy, por las ventajas que estas ofrecen. La medida canónica de Kolmogorov combina la medida de Lévy con el parámetro de difusión; esta medida reacomoda la fórmula de Lévy-Khintchine de manera que pueden obtenerse con facilidad el primer y el segundo momento, además de que es una medida finita, a diferencia de la medida de Lévy, que no necesariamente lo es. La densidad de Lévy se obtiene al asumir que la medida de Lévy es absolutamente continua con respecto a la medida de Lebesgue. Cuando existe la densidad de Lévy, pueden usarse métodos de estimación tipo kernel. Por otro lado, la forma de la densidad de Lévy proporciona más información acerca de la distribución de los saltos que la medida de Lévy, pues a través de esta puede observarse dónde se concentran más los saltos de un proceso de Lévy.

Cuando se observa un proceso de Lévy discreta y equidistantemente, la estimación no

paramétrica es fundamentalmente tratada en la literatura con dos enfoques diferentes de estimación, dependiendo de la disponibilidad de datos. Uno de ellos es la llamada frecuencia alta de observación, que es cuando se considera que los saltos y la parte de difusión pueden observarse directamente de la muestra. Por otro lado, se considera que se tiene una frecuencia baja de observaciones cuando la distancia de observaciones no tiende a cero e incluso asintóticamente, los saltos y la parte de difusión no pueden observarse. Estudios de estimación con frecuencia alta de datos pueden encontrarse en Figueroa y Houdré [16], Figueroa [15] y Comte y Genon-Catalot [7].

Hay dos objetivos principales en este trabajo. El primero de ellos es hacer una reseña y estudio cualitativo de los trabajos en los que se estima la medida de Lévy bajo ciertas condiciones, y proponer un método de construcción de bandas de confianza para la medida de Lévy. Estas condiciones son: que la estimación sea no paramétrica, a través de la función característica empírica, con muestreo discreto equidistante asumiendo baja frecuencia de datos; no restringir a que el proceso sea únicamente de saltos o difusión y pidiendo que el proceso sea de actividad finita, y tenga al menos segundo momento finito. En concreto, se hace una revisión de los trabajos de los últimos 8 años que realizan la estimación de la medida de Lévy con las condiciones de modelación mencionadas arriba; estos son de Watteel y Kulperger [29], Neumann y Reiss [26] y Kappus y Reiss [24], quienes estiman la medida de Lévy a través de la medida canónica de Kolmogorov, y Gugushvili [19] y [20], donde se estima la medida de Lévy mediante la densidad de Lévy. La justificación de proponer bandas de confianza es que en la literatura se ha encontrado que se da preferencia a estimadores puntuales y sería muy interesante complementar las inferencias acerca de la medida de Lévy, que no es un parámetro real, sino una función; con ello se obtendría una mejor descripción de los saltos del proceso de Lévy. La discusión que se realiza tiene que ver con las diferencias en los supuestos, que implican diferencias importantes en la modelación; por ejemplo, en [19] y [20] se asume que el número esperado de saltos por unidad es finito y, en cambio, [29], [26] y [24] hacen condiciones fuertes sobre los momentos del proceso, que implican que las colas no pueden ser muy pesadas.

El segundo objetivo es dar una propuesta de estimador para la medida de Lévy bajo los mismos supuestos que los trabajos de estudio y que sea fácil de calcular e implementar. La razón de esto es que el cálculo de los estimadores propuestos en los trabajos considerados es computacionalmente intrincado y lento, por lo que es relevante saber si un estimador que es más simple es capaz de tener un buen desempeño. Específicamente, en este trabajo se hace una propuesta de un estimador de momentos para la medida de Lévy, a través de la medida canónica de Kolmogorov; para conseguir esto, se “parametriza” la medida canónica de Kolmogorov usando la descomposición en ondículas de Haar, luego se igualan la función característica y su versión empírica en diferentes puntos. Puede entenderse este estimador haciendo analogía con uno tipo *sieves*, el cual fue propuesto por Grenander [18] en 1981 y usado por Figueroa [15] para estimar la densidad de Lévy. *Grosso modo*, este método consiste en aproximar una función por el mejor elemento de alguna familia de modelos paramétricos lineales de dimensión finita, que pueden volverse más y más complejos en tanto aumenta la cantidad de datos. Al usar la descomposición de Haar, lo que se obtiene al igualar la función característica con su versión empírica, es un sistema de ecuaciones del cual no es difícil calcular una solución. Por su similitud al concepto de igualar versiones teóricas paramétricas con versiones empíricas para establecer un sistema de ecuaciones a resolver, este método será referido en este trabajo (así como se hace en parte de la literatura, por ejemplo, en Feuerverger y McDunnough [12] y [13]) como el “método de momentos”.

Debido a inconvenientes en la implementación, en este trabajo se plantea sustituir la base de ondículas de Haar por otro modelo lineal de dimensión finita que consiste en aproximar la medida canónica de Kolmogorov por una medida compuesta únicamente de saltos en un retícula especificada. Los resultados obtenidos establecen el potencial de este modelo lineal como otra alternativa de estimación con ventajas de implementación y plantean algunos trabajos futuros.

La estructura de este trabajo es la siguiente: En el Capítulo 1 se introducen los procesos de Lévy y la fórmula de Lévy-Khintchine, además de la función característica empírica y sus propiedades como estimador de la función característica, con algunos ejemplos de estimación

para el caso de un proceso de Lévy caracterizado por un número finito de parámetros. Los Capítulos 2 y 3 reseñan los trabajos que estiman la medida de Lévy a través de la medida canónica de Kolmogorov y a través de la densidad de Lévy, respectivamente. En el Capítulo 4 se estudian las diferencias de supuestos y modelación entre los trabajos considerados en los Capítulos 2 y 3. En el Capítulo 5 se hace una propuesta para estimar la medida de Lévy con un método *sieves* basado en momentos. En el Capítulo 6 se discute la implementación de los estimadores propuestos en los trabajos que se estudiaron en los capítulos previos; se hacen comentarios sobre la implementación del estimador que se propone en el Capítulo 5; se implementa el estimador tipo *sieves* basado en momentos, sustituyendo la base de ondículas de Haar por otro modelo lineal, y se sugieren bandas de confianza para los estimadores de la medida de Lévy, usando el método bootstrap. Con el fin de que la lectura de este trabajo sea autocontenida y fluida, se agregan dos apéndices. En el primero de ellos se enuncian las propiedades más relevantes de los procesos de Lévy. En el segundo apéndice se hace un compendio de resultados varios que serán utilizados a lo largo del trabajo como son la base de ondículas de Haar y algunos resultados de funciones características.

Capítulo 1

Preliminares

La fórmula de Lévy-Khintchine muestra, a través de la función característica, cómo un proceso de Lévy está descrito por dos parámetros reales y una medida. Es por lo anterior que la función característica empírica emerge como herramienta natural en las inferencias sobre procesos de Lévy. En este capítulo se introducen los procesos de Lévy y la fórmula de Lévy-Khintchine. También se enuncian algunas de las propiedades de la función característica empírica como estimador de la función característica, por ejemplo que es insesgada, asintóticamente normal y fuertemente consistente, tanto puntual, como uniformemente, entre otras; además se muestran algunos donde la función característica empírica puede usarse para estimar una cantidad finita de parámetros.

1.1 Procesos de Lévy y fórmula de Lévy-Khintchine

En esta sección se introducen los procesos de Lévy y la fórmula de Lévy-Khintchine, que serán usados a lo largo de este trabajo. Para una descripción más completa y con ejemplos de procesos de Lévy, se sugiere al lector ir al Apéndice A, así como a las referencias bibliográficas allí descritas.

Definición 1.1 *Un proceso de Lévy $\{X_t, t \geq 0\}$ es aquel que satisface las siguientes condiciones:*

- (i) *Tiene incrementos independientes: Para todos $0 \leq t_0 < t_1 < \dots < t_n$, las v.a. $X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes.*
- (ii) *Tiene incrementos estacionarios: Para todos $t_0 < t_1$, se cumple que $X_{t_1} - X_{t_0} \stackrel{\mathcal{L}}{=} X_{t_1 - t_0}$.*
- (iii) *Tiene trayectorias continuas a la derecha con límites a la izquierda (càdlàg), casi seguramente.*
- (iv) *$X_0 = 0$, casi seguramente.*

La representación de Lévy-Khintchine es una manera clásica de escribir el exponente característico, es decir, el logaritmo de la función característica de un proceso de Lévy $\{X_t, t \geq 0\}$ en algún t fijo. La variable t puede interpretarse como el tiempo en que el proceso de Lévy es observado. Existen otras maneras de escribir el exponente característico, usadas en algunos de los trabajos que aquí se estudian, y que son descritas en la Sección A.4.

Es bien conocido que la **función característica** de un proceso de Lévy $\{X_t, t \geq 0\}$ puede escribirse de la forma

$$\varphi_t(u) := \mathbb{E}e^{iuX_t} = e^{t\Psi(u)},$$

donde Ψ se define como

$$\Psi(u) = iub - \frac{\sigma^2}{2}u^2 + \int [e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}}] \nu(dx), \quad \forall u \in \mathbb{R}, \quad (1.1)$$

y ν es una medida en \mathbb{R} que no carga cero (i.e., $\nu(\{0\}) = 0$) y cumple

$$\int (x^2 \wedge 1) \nu(dx) < \infty. \quad (1.2)$$

La función Ψ se conoce como **exponente característico** o **función cumulante** y la expresión en (1.1) se conoce como **fórmula de Lévy-Khintchine**. La medida ν recibe el nombre de **medida de Lévy o medida de saltos**, la constante b se conoce como **término de deriva** y la constante no negativa σ^2 como **término browniano, de difusión o de volatilidad**. Cuando existe una función $\rho : \mathbb{R} \rightarrow \mathbb{R}_+$, tal que $\nu(dx) = \rho(x)dx$ (es decir, ν es absolutamente continua, con respecto a la medida de Lebesgue), esta es llamada **densidad**

de Lévy. Note que la función no negativa ρ , no necesariamente es densidad de probabilidad, pues su integral podría ser infinita o no integrar a uno; esto constituye un abuso de nomenclatura al llamar “densidad” a ρ .

Como puede constatarse en el Apéndice A, cada proceso de Lévy se descompone como un movimiento browniano con deriva más un proceso de puros saltos. Los términos b y σ^2 son la deriva y la varianza del movimiento browniano, mientras que la medida de Lévy, ν , describe el comportamiento de los saltos.

Por otro lado, cada terna (b, σ^2, ν) con $b \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}_+$ y ν medida que no carga cero y satisface (1.2), determina un único proceso de Lévy cuyo exponente característico está dado por (1.1). Observe que b y σ^2 son parámetros reales, mientras que la medida de Lévy, ν , es de dimensión infinita.

En la siguiente sección se explican razones por las cuales la función característica empírica es un estimador trascendental de la función característica. En la literatura que aquí se considera, usando la función característica empírica se estima la medida de Lévy a través de la llamada medida canónica de Kolmogorov o a través de la densidad de Lévy.

Cuando existe el segundo momento del proceso de Lévy, la medida de Lévy puede combinarse con el término de difusión a través de la llamada forma canónica de Kolmogorov. A través de esta pueden recuperarse la medida de Lévy y el término de difusión. La medida canónica de Kolmogorov tiene la ventaja de ser una medida finita, a diferencia de ν , que no necesariamente lo es. Al trabajar con esta medida, el exponente característico toma una forma muy especial con cierta interpretación. En el Capítulo 2 se describe más detalladamente la medida de Kolmogorov, así como los trabajos que hacen uso de ella para estimar ν .

La existencia de la densidad de Lévy es un supuesto deseable cuando se quiere modelar usando un proceso de Lévy. En efecto, al estimar ν a través de la densidad de Lévy, se tiene la posibilidad usar técnicas similares a la del kernel de estimación, que es exactamente a lo que se recurre en Gugushvili [19] y [20]. El Capítulo 3 versa sobre la estimación de la medida de Lévy a través de su densidad y de la descripción de los trabajos [19] y [20].

1.2 Función característica empírica

La función característica empírica es una herramienta que puede usarse para fines de estimación de parámetros; por ejemplo, cuando no se conoce una forma cerrada para la densidad de probabilidad (y por ende, tampoco para la función de verosimilitud). La función característica tiene varias propiedades que la hace conveniente; por ejemplo, se comporta bien ante cambios de localización y escala, sumas de variables aleatorias independientes y además está en correspondencia biyectiva con la función de distribución.

En esta sección se dan algunas propiedades asintóticas de la función característica empírica que motivan su uso para la estimación de la medida de Lévy. Los resultados que aquí se presentan, han sido tomados de Feuerverger y Mureika [11], Feuerverger y McDunnough [12] y [13] y Prakasa Rao [27] (Cap. 8), quien a su vez menciona los resultados más relevantes de funciones características empíricas.

Considere una sucesión de variables aleatorias independientes e idénticamente distribuidas $\{Y_n, n \geq 1\}$, con función de distribución común F y función característica

$$\varphi(t) = \int e^{itx} F(dx).$$

La **función característica empírica** de las primeras n variables aleatorias es

$$\widehat{\varphi}_n(t) = \frac{1}{n} \sum_{j=1}^n e^{itY_j}.$$

Como estimador para φ , la función característica empírica tiene varias propiedades provechosas. De la definición de función característica se sigue que $\widehat{\varphi}_n$ es insesgado para φ , es decir, para cada $t \geq 0$,

$$\mathbb{E}\widehat{\varphi}_n(t) = \varphi(t).$$

Además de esta propiedad de insesgamiento, la función característica empírica es consistente de diversas maneras, como a continuación se explica.

Se sabe, por la Ley Fuerte de Grandes Números, que para cada $t \geq 0$,

$$\widehat{\varphi}_n(t) \rightarrow \varphi(t), \text{ c.s.,}$$

es decir, $\widehat{\varphi}_n$ es un estimador fuertemente consistente para φ , en el sentido puntual. Sin embargo, en [27] se resumen resultados más generales que muestran que esta propiedad puntual no sólo es válida localmente, sino también en forma funcional. La siguiente proposición resume estos resultados. La parte (a) indica que la convergencia es uniforme sobre intervalos finitos; la parte (b), que la convergencia es uniforme sobre toda la recta real, siempre y cuando la función de distribución sea discreta; la parte (c), que la convergencia es uniforme sobre intervalos que van creciendo a cierta tasa, y la parte (d), una manera en que, bajo una noción de norma L^p , la función característica empírica puede converger a la función característica.

Proposition 1.1 *Para la función característica empírica, $\widehat{\varphi}_n$, son válidas las siguientes afirmaciones:*

(a) *Para $0 \leq T < \infty$, se tiene que*

$$\lim_{n \rightarrow \infty} \sup_{|t| \leq T} |\widehat{\varphi}_n(t) - \varphi(t)| = 0, \text{ c.s.}$$

(b) *Si F es discreta, entonces*

$$\lim_{n \rightarrow \infty} \sup_{t \in \mathbb{R}} |\widehat{\varphi}_n(t) - \varphi(t)| = 0, \text{ c.s.}$$

(c) *Si $T_n = o\left(\sqrt{n/\log \log n}\right)$, entonces*

$$\lim_{n \rightarrow \infty} \sup_{|t| \leq T_n} |\widehat{\varphi}_n(t) - \varphi(t)| = 0, \text{ c.s.}$$

(d) *Si $T_n = o(n^{p/2})$, para $0 < p \leq 2$, entonces se cumple la convergencia en probabilidad,*

$$\int_0^{T_n} |\widehat{\varphi}_n(t) - \varphi(t)|^p dt \xrightarrow[n \rightarrow \infty]{} 0.$$

Defina ahora el **proceso característico** como

$$\zeta_n(t) = \sqrt{n} [\widehat{\varphi}_n(t) - \varphi(t)], \quad t \in \mathbb{R}.$$

Observe que

$$\mathbb{E}\zeta_n(t) = 0 \text{ y } \mathbb{E}[\zeta_n(t_1)\zeta_n(t_2)] = \varphi(t_1 + t_2) - \varphi(t_1)\varphi(t_2),$$

para todos $t, t_1, t_2 \in \mathbb{R}$. Sea además ζ un proceso complejo gaussiano, con media cero y la misma estructura de covarianza que ζ_n .

Por el Teorema de Límite Central Multivariado se tiene que para cada $t \in \mathbb{R}$, $\zeta_n(t) \xrightarrow{\mathcal{L}} \zeta(t)$. Más aún, se cuenta con el siguiente resultado acerca de la convergencia débil de ζ_n al proceso gaussiano complejo ζ .

Proposition 1.2 *El proceso $\zeta_n(t)$ converge débilmente a $\zeta(t)$, sobre intervalos finitos.*

A continuación se describen algunos métodos de estimación de parámetros usando la función característica empírica; existen otros que no se mencionan (ver Feuerverger y McDunnough [13]). Aunque aquí no se discute, la función característica empírica puede ser usada también para construir pruebas de bondad de ajuste (ver Rao [27]) y pruebas de simetría (ver Feuerverger y McDunnough [11]), ya sea alrededor del origen o en algún punto no especificado. Los tres métodos que aquí se describen, son aplicables a familias caracterizadas por un número finito de parámetros en un conjunto $\Theta \subseteq \mathbb{R}^p$. Así pues, en el resto de esta sección se considera una sucesión de variables aleatorias independientes e idénticamente distribuidas, $\{Y_n, n \geq 1\}$, con función característica $\varphi(t, \theta)$, para $\theta \in \Theta \subseteq \mathbb{R}^p$.

1.2.1 Estimador de mínima distancia

El estimador que se describe aquí está basado en el error cuadrático integrado, como el título lo sugiere. Bajo ciertas condiciones, tal estimador es asintóticamente normal; su demostración pueden encontrarse en Rao [27].

Definición 1.2 *Un estadístico $\hat{\theta}_n = \hat{\theta}_n(Y_1, \dots, Y_n)$ se dice un **estimador de error cuadrático integrado** con función de peso $G(\cdot)$ si minimiza*

$$I_n(\theta) = \int |\hat{\varphi}_n(t) - \varphi(t; \theta)|^2 G(dt).$$

La función G se asume creciente y de variación total uno.

A continuación se enuncian supuestos para demostrar la asintoticidad normal de $\widehat{\theta}_n$, los cuales coinciden con los supuestos usuales que los llamados M -estimadores para ser asintóticamente normales (ver Huber [22]).

Supuestos de regularidad:

1. (Identificabilidad) $\theta_1 = \theta_2$ si y sólo si $P_{\theta_1} = P_{\theta_2}$, donde P_{θ_j} es la medida de probabilidad correspondiente a la función característica $\varphi(t, \theta)$.
2. $I_n(\theta)$ es dos veces diferenciable con respecto a θ .
3. Si se definen $U(t, \theta) = \text{Re } \varphi(t, \theta)$ y $V(t, \theta) = \text{Im } \varphi(t, \theta)$, entonces para $l, m = 1, \dots, p$, $\partial^2 U(t, \theta) / \partial \theta_l \partial \theta_m$ y $\partial^2 V(t, \theta) / \partial \theta_l \partial \theta_m$, son simultáneamente continuas en t y θ y uniformemente acotadas por una función G -integrable.
4. $(\mathcal{I}(\theta))_{l,m} := -\mathbb{E} \left[\frac{\partial^2 \log f(Y_1, \theta)}{\partial \theta_l \partial \theta_m} \right] < \infty$, donde $f(x, \theta)$ es la densidad de Y_1 .
5. Si $\lambda(\theta) = \int |\varphi'(t, \theta)|^2 G(dt)$, entonces $0 < \lambda(\theta) < \infty$.

Teorema 1.3 *Asuma las condiciones 1-5 y define*

$$K(x, \theta) = \int [\cos tx - U(t, \theta)] \frac{\partial U(t, \theta)}{\partial \theta} G(dt) + \int [\sin tx - V(t, \theta)] \frac{\partial V(t, \theta)}{\partial \theta} G(dt).$$

Entonces existe una raíz $\widehat{\theta}_n$ de la ecuación $I_n(\theta) = 0$, la cual es fuertemente consistente; además, si θ_0 es el verdadero valor del parámetro, entonces

$$\sqrt{n} (\widehat{\theta}_n - \theta_0) \xrightarrow{\mathcal{L}} N_k \left(0, \frac{\text{var } K(Y_1, \theta_0)}{\lambda^2(\theta_0)} \right).$$

1.2.2 Procedimiento k-L

Este es un método asintótico de máxima verosimilitud, basado en observaciones de la función característica empírica en k puntos fijos t_1, \dots, t_k , que no varían con n , el tamaño de la muestra. Su nombre hace alusión a estos k puntos y a la palabra “verosimilitud” en inglés (*likelihood*). Este método es propuesto en Feuerverger y McDunnough [12] y consiste en

aproximar normalmente la logverosimilitud de las partes real e imaginaria de la función característica empírica.

Defínanse los vectores z_θ y z_n , de dimensión $2k$, tales que

$$z_\theta = \begin{pmatrix} \operatorname{Re} \varphi(t_1) \\ \vdots \\ \operatorname{Re} \varphi(t_k) \\ \operatorname{Im} \varphi(t_1) \\ \vdots \\ \operatorname{Im} \varphi(t_k) \end{pmatrix} \quad \text{y} \quad z_n = \begin{pmatrix} \operatorname{Re} \widehat{\varphi}_n(t_1) \\ \vdots \\ \operatorname{Re} \widehat{\varphi}_n(t_k) \\ \operatorname{Im} \widehat{\varphi}_n(t_1) \\ \vdots \\ \operatorname{Im} \widehat{\varphi}_n(t_k) \end{pmatrix}. \quad (1.3)$$

Además defina $\Sigma^{(n)} = \operatorname{var}(z_n)$. Se asume que θ_0 es el verdadero valor del parámetro y se encuentra en algún conjunto abierto y convexo. También se asume que la función de distribución de las observaciones, tiene densidad.

Por la Proposición 1.2, se obtiene que

$$-k \log \det \Sigma^{(n)} - n (z_n - z_\theta)^T (\Sigma^{(n)})^{-1} (z_n - z_\theta),$$

es una aproximación a la logverosimilitud. Esto sugiere una estimación minimizando sobre la función

$$l_n(\theta) := n (z_n - z_\theta)^T (\Sigma^{(n)})^{-1} (z_n - z_\theta).$$

Se tiene que existe $\widehat{\theta}_n$ que es una raíz asintótica de la función $\partial l_n(\theta) / \partial \theta$ y además es consistente para θ_0 . Además, $\widehat{\theta}_n$ es asintóticamente normal con media θ y varianza asintótica

$$\left(\frac{\partial z_\theta^T}{\partial \theta} \right) (\Sigma^{(n)})^{-1} \left(\frac{\partial z_\theta^T}{\partial \theta} \right), \quad (1.4)$$

evaluado en θ_0 .

1.2.3 Estimador de momentos

El método que aquí se describe, se propone en [12]. La idea es igualar la parte real y la parte imaginaria de la función característica con su contraparte empírica, evaluándolas en k diferentes puntos, t_1, \dots, t_k , que no varían con n , el tamaño de la muestra.

1.2. Función característica empírica

Sean z y z_n como en (1.3) y $\Sigma^{(n)} = \text{var}(z_n)$. Considere la ecuación

$$D^T (z_n - z_\theta) = 0, \quad (1.5)$$

donde D es una matriz de tamaño $2k \times p$ definida como

$$D = (\Sigma^{(n)})^{-1} \frac{\partial z_\theta}{\partial \theta^T}.$$

Si $\hat{\theta}_n$ es una raíz consistente de (1.5), entonces se demuestra que $\hat{\theta}_n$ es asintóticamente normal con media θ y varianza asintótica

$$\left(\frac{\partial z_\theta^T}{\partial \theta} \right) (\Sigma^{(n)})^{-1} \left(\frac{\partial z_\theta^T}{\partial \theta} \right),$$

evaluado en θ_0 , el verdadero valor de θ . Por lo tanto, este método y el estimador del procedimiento k-L son asintóticamente equivalentes.

Capítulo 2

Estimación de la medida canónica de Kolmogorov

La medida de Lévy puede estimarse a través de la medida canónica de Kolmogorov, por las ventajas que esta última ofrece. La medida canónica de Kolmogorov se obtiene combinando la medida de Lévy con el término de difusión; para definirse, se requiere la existencia del segundo momento del proceso de Lévy. Con esta medida se reacomoda la fórmula de Lévy-Khintchine de manera que pueden obtenerse con facilidad el primer y el segundo momento, además de que es una medida finita, a diferencia de la medida de Lévy, que no necesariamente lo es.

En este capítulo se introduce la medida canónica de Kolmogorov y se hace una descripción de los trabajos que la estiman de manera no paramétrica, a través de la función característica empírica, con muestreo discreto equidistante asumiendo baja frecuencia de datos, sin restringir a que el proceso sea únicamente de saltos o difusión. Estos trabajos son de Watteel y Kulperger [29], Neumann y Reiss [26] y Kappus y Reiss [24]; el primero usa inversión de Fourier y los últimos dos, un criterio de mínima distancia a la función característica empírica.

El esquema de observación es discreto y uniforme, es decir, un proceso de Lévy $\{X_t, t \geq 0\}$ se observa en los puntos $\Delta, 2\Delta, \dots, n\Delta$. La medida canónica de Kolmogorov, ν_σ , se define

como

$$\nu_\sigma(dx) = \sigma^2 \delta_0(dx) + x^2 \nu(dx),$$

donde σ^2 y ν son el parámetro de difusión y la medida de Lévy del proceso $\{X_t, t \geq 0\}$, respectivamente. Dado que el proceso tiene segundo momento (y por ende ν), se tiene que la medida ν_σ es finita.

Con esta medida, el exponente característico (1.1), se escribe para cada $u \in \mathbb{R}$ como

$$\Psi(u) = iu\mu + \int \frac{e^{iux} - 1 - iux}{x^2} \nu_\sigma(dx), \quad (2.1)$$

donde $\mu = b + \int_{|x| \geq 1} x \nu(dx)$ y el integrando en (2.1) se define continuamente en $x = 0$, es decir,

$$\left. \frac{e^{iux} - 1 - iux}{x^2} \right|_{x=0} = -\frac{u^2}{2}.$$

Además se cumple que μ es la media del proceso y $\mathbb{E}X_1^2 = \nu_\sigma(\mathbb{R})$. El término $\int_{|x| \geq 1} x \nu(dx)$ es finito debido a la existencia del segundo momento.

La transformada de Fourier de la medida canónica de Komogorov, $\mathcal{F}\nu_\sigma$, puede expresarse en términos de la función característica o en términos de la segunda derivada del exponente característico, como a continuación se describe. Si φ es la función característica al tiempo 1 y Ψ el exponente característico, entonces se cumple que

$$-\Psi''(u) = \mathcal{F}\nu_\sigma(u) = -\left\{ \frac{[\varphi'(u)]^2}{\varphi(u)^2} - \frac{\varphi''(u)}{\varphi(u)} \right\}. \quad (2.2)$$

Esta fórmula es usada en Watteel y Kulperger [29], Neumann y Reiss [26], Kappus y Reiss [24] y Gugushvili [19] para formular sus propuestas de estimación.

2.1 Con inversión de Fourier

El trabajo que aquí se describe es de Watteel y Kulperger [29], publicado en 2003. A grandes rasgos, estima para variables aleatorias infinitamente divisibles en L^2 , la medida canónica de Kolmogorov, ν_σ . El método de estimación que usa es expresar la transformada de Fourier de ν_σ en términos de la función característica y luego sustituir esta por la función característica

empírica en la fórmula de inversión de Fourier. Posteriormente se muestra la asintoticidad normal de este estimador sobre intervalos; para ello se asume la existencia del momento $5 + \delta$, para alguna $\delta > 0$. Finalmente se consideran algunos aspectos de implementación y se calculan bandas de confianza; esta última parte se discutirá en el Capítulo 6. Aunque en este artículo no se hace referencia a los procesos de Lévy, es claro que el artículo puede entenderse como estimación de las características de un proceso de Lévy con esquema de observación discreto y equidistante, es decir, $\Delta = 1$.

Usando el teorema de inversión B.2 y la fórmula (2.2), se tiene que para $-\infty < a < b < \infty$,

$$\nu_\sigma([a, b]) = \lim_{S \rightarrow \infty} \frac{1}{2\pi} \int_{-S}^S \frac{e^{-iau} - e^{-ibu}}{iu} \left(\frac{[\varphi'(u)]^2}{\varphi(u)^2} - \frac{\varphi''(u)}{\varphi(u)} \right) du. \quad (2.3)$$

El estimador que se propone es el obtenido al sustituir en (2.3) la función característica, φ , por la función característica empírica, $\widehat{\varphi}_n$. Esto es,

$$\widehat{\nu}_{\sigma,n,S}([a, b]) = \frac{1}{2\pi} \int_{-S}^S \frac{e^{-iau} - e^{-ibu}}{iu} h(\widehat{\varphi}_n(u), \widehat{\varphi}'_n(u), \widehat{\varphi}''_n(u)) du, \quad (2.4)$$

donde $h(v_1, v_2, v_3) = \frac{v_2^2}{v_1^2} - \frac{v_3}{v_1}$.

Siguiendo [29], se muestra que el estimador $\widehat{\nu}_{\sigma,n,S}([a, b])$ de $\nu_\sigma([a, b])$ es asintóticamente normal. Para esto se asume

$$\mathbb{E} |X_1|^{5+\delta} < \infty,$$

para algún $\delta > 0$.

Expandiendo el término $h(\widehat{\varphi}_n(u), \widehat{\varphi}'_n(u), \widehat{\varphi}''_n(u))$ en (2.4) a través de su serie de Taylor alrededor de $h(\varphi(u), \varphi'(u), \varphi''(u))$, se obtiene que

$$h(\widehat{\varphi}_n(u), \widehat{\varphi}'_n(u), \widehat{\varphi}''_n(u)) = h(\varphi(u), \varphi'(u), \varphi''(u)) + \left(\frac{\partial h}{\partial \varphi} \right) \begin{pmatrix} \widehat{\varphi}_n(u) - \varphi(u) \\ \widehat{\varphi}'_n(u) - \varphi'(u) \\ \widehat{\varphi}''_n(u) - \varphi''(u) \end{pmatrix} + o_P\left(\frac{1}{\sqrt{n}}\right),$$

donde $o_P\left(\frac{1}{\sqrt{n}}\right)$ es el residuo de la expansión de Taylor.

Para $j = 0, 1, 2$, defina $\zeta_n^{(j)}(u) = \sqrt{n} \left(\widehat{\varphi}_n^{(j)}(u) - \varphi^{(j)}(u) \right)$. Se obtiene que $\mathbb{E} \zeta_n^{(j)}(u) = 0$ y

$$\text{Cov} \left[\zeta_n^{(j)}(u_1), \zeta_n^{(k)}(u_2) \right] = \varphi^{(j+k)}(u_1 + u_2) - \varphi^{(j)}(u_1) \varphi^{(k)}(u_2),$$

para todas $j, k \in \{0, 1, 2\}$ y $u_1, u_2 \in \mathbb{R}$.

El siguiente resultado afirma que el estimador de la medida ν_σ , $\widehat{\nu}_{\sigma,n,S}$, es asintóticamente normal con una varianza engorrosa. Puede verse también la relación de este teorema con la Proposición (1.2).

Teorema 2.1 *Sea $\Gamma = \{u : |u| < S\}$, $S > 0$, una vecindad donde $\widehat{\varphi}_n \neq 0$. Asuma que $\mathbb{E}|X_j|^{5+\delta} < \infty$, para algún $\delta > 0$. Entonces si se define*

$$\nu_{\sigma,S}([a, b]) := \frac{1}{2\pi} \int_{-S}^S \frac{e^{-iau} - e^{-ibu}}{iu} \left(\frac{[\varphi'(u)]^2}{\varphi(u)^2} - \frac{\varphi''(u)}{\varphi(u)} \right) du,$$

se tiene que

$$\sqrt{n} \{ \widehat{\nu}_{\sigma,n,S}([a, b]) - \nu_{\sigma,S}([a, b]) \} = \frac{1}{2\pi} \int_{\Gamma} \frac{e^{iau} - e^{ibu}}{iu} \sqrt{n} \{ \widehat{\varphi}_n''(u) - \varphi''(u) \} du,$$

es asintóticamente normal con media cero y varianza $\int_{|s|<S} \int_{|u|<s} G(s, u) dw_s dw_u$, donde $dw_u = (2\pi i u)^{-1} (e^{iau} - e^{ibu}) dt$ y G se define como

$$G(u_1, u_2) = \lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[\widehat{\varphi}_n(u_1)^3 \overline{\widehat{\varphi}_n(u_2)^3} g_n(u_1, u_2) \right]}{\varphi(u_1)^3 \overline{\varphi(u_2)^3}},$$

con

$$g_n(u_1, u_2) = \left\{ \left(\frac{\partial h}{\partial \varphi} \right) \begin{pmatrix} \zeta_n(u_1) \\ \zeta_n'(u_1) \\ \zeta_n''(u_1) \end{pmatrix} \right\} \overline{\left\{ \left(\frac{\partial h}{\partial \varphi} \right) \begin{pmatrix} \zeta_n(u_2) \\ \zeta_n'(u_2) \\ \zeta_n''(u_2) \end{pmatrix} \right\}}.$$

Teorema 2.2 *Si una variable aleatoria infinitamente divisible tiene varianza finita, entonces*

$$\lim_{R \rightarrow \infty} \frac{-1}{2\pi} \int_{S < |t| < R} \frac{e^{-iat} - e^{-ibt}}{it} \Psi''(u) du \xrightarrow{S \rightarrow \infty} 0.$$

El Teorema 2.2 junto con el Teorema 2.1, indican que $\widehat{\nu}_{\sigma,n,S}$ aproxima bien ν_σ si S y n son suficientemente grandes.

2.2 Con mínima distancia a la función característica empírica

En esta sección se describe el estimador para la medida canónica de Kolmogorov propuesto en Neumann y Reiss ([26]), publicado en 2009. Se observa discreta y uniformemente un proceso de Lévy en los puntos $1, 2, \dots, n$ (es decir, $\Delta = 1$). Se estiman μ y la medida canónica de Kolmogorov mediante un criterio de mínima distancia a la función característica empírica. Posteriormente se calculan tasas óptimas de convergencia cuando $n \rightarrow \infty$ y se muestra un ejemplo para el cual implementan y aplican su método, mas esta última parte se discute en el Capítulo 6, que aborda el cómputo de los estimadores.

Primero se proponen estimadores fuertemente consistentes para cualquier métrica que satisfaga un par de condiciones generales, sin asumir la existencia siquiera de momentos. Para esto, los autores del trabajo usan una forma equivalente del exponente característico (1.1), la cual surge de usar la función de centro $c(x) = 1/(1+x^2)$. También es importante la equivalencia

$$\int (1 \wedge x^2) \nu(dx) < \infty \text{ si y sólo si } \int \frac{x^2}{1+x^2} \nu(dx) < \infty.$$

Entonces el exponente característico se escribe como

$$\Psi(u) = iu\bar{b} - \frac{1}{2}\sigma^2 u^2 + \int \left[e^{iux} - 1 - \frac{iux}{1+x^2} \right] \nu(dx),$$

con ν una medida en \mathbb{R} que no carga cero y cumple

$$\int \frac{x^2}{1+x^2} \nu(dx) < \infty. \tag{2.5}$$

Se introduce la medida finita $\bar{\nu}_\sigma$ como

$$\bar{\nu}_\sigma(dx) = \sigma^2 \delta_0(dx) + \frac{x^2}{1+x^2} \nu(dx).$$

Esto permite reescribir el exponente característico como

$$\Psi(u) = iu\bar{b} + \int \frac{(e^{iux} - 1)(1 + x^2) - iux}{x^2} \bar{\nu}_\sigma(dx),$$

extendiendo continuamente en $x = 0$ en el integrando, por lo que

$$\left. \frac{(e^{iux} - 1)(1 + x^2) - iux}{x^2} \right|_{x=0} = -\frac{u^2}{2}.$$

Así pues, para una métrica apropiada d , se tratan de definir $\widehat{b}_n, \widehat{\nu}_{\sigma,n}$ de manera que

$$d\left(\widehat{\varphi}_n, \varphi\left(\cdot; \widehat{b}, \widehat{\nu}_\sigma\right)\right) = \inf_{\widetilde{b} \in \mathbb{R}, \widetilde{\nu}_\sigma \in \mathcal{M}(\mathbb{R})} d\left(\widehat{\varphi}_n, \varphi\left(\cdot; \widetilde{b}, \widetilde{\nu}_\sigma\right)\right),$$

donde $\mathcal{M}(\mathbb{R})$ es la clase de medidas finitas en \mathbb{R} . Como no se puede garantizar que el ínfimo se alcanza en la fórmula anterior, se toma una sucesión $(\delta_n)_n$ que converja a cero y \widehat{b}_n y $\widehat{\nu}_{\sigma,n}$ tales que

$$d\left(\widehat{\varphi}_n, \varphi\left(\cdot; \widehat{b}_n, \widehat{\nu}_{\sigma,n}\right)\right) \leq \inf_{\widetilde{b} \in \mathbb{R}, \widetilde{\nu}_\sigma \in \mathcal{M}(\mathbb{R})} d\left(\widehat{\varphi}_n, \varphi\left(\cdot; \widetilde{b}, \widetilde{\nu}_\sigma\right)\right) + \delta_n. \quad (2.6)$$

Es importante señalar que los estimadores \widehat{b}_n y $\widehat{\nu}_{\sigma,n}$, obtenidos en (2.6) no necesariamente son únicos.

Para la métrica d se asumen dos condiciones que permitirán demostrar la consistencia de los estimadores que satisfagan (2.6). Se supone que

$$\lim_{n \rightarrow \infty} d(\widehat{\varphi}_n, \varphi) = 0, \quad P_{\widehat{b}_n, \widehat{\nu}_{\sigma,n}} - \text{casi seguramente} \quad (2.7)$$

y además

$$\lim_{n \rightarrow \infty} d(\varphi(\cdot; \bar{b}_n, \bar{\nu}_{\sigma,n}), \varphi) = 0 \quad (2.8)$$

$$\Rightarrow \lim_{n \rightarrow \infty} \int_s^t \varphi(u; \bar{b}_n, \bar{\nu}_{\sigma,n}) du = \int_s^t \varphi(u) du, \quad \forall s, t \in \mathbb{R}.$$

Un ejemplo de tal métrica, es la norma $L^2(\mathbb{R})$ ponderada por una función integrable. La condición (2.7) se obtiene a partir del Teorema de Convergencia Dominada y la condición (2.8) es inmediata.

Los estimadores en (2.6) son fuertemente consistentes. Esto se sigue del Teorema A.3 y de las condiciones sobre la distancia.

Teorema 2.3 *Si la distancia d satisface las propiedades (2.7) y (2.8), entonces los estimadores de mínima distancia $(\widehat{b}_n, \widehat{\nu}_{\sigma,n})$ son fuertemente consistentes, es decir, con $P_{\widehat{b}, \widehat{\nu}_\sigma}$ -probabilidad 1 se tiene que*

$$\widehat{b}_n \rightarrow \bar{b} \quad \text{y} \quad \widehat{\nu}_{\sigma,n} \Rightarrow \bar{\nu}_\sigma.$$

Finalmente, se demuestra que para una métrica en particular se pueden obtener tasas de convergencia óptimas, definiendo una distancia para medidas finitas, que a continuación se especifica. A partir de aquí se supone la existencia del momento $4 + \delta$, para algún $\delta > 0$. Además, en lugar de estimar $\bar{\nu}_\sigma$ y \bar{b} , se estiman μ y la medida canónica de Kolmogorov.

La forma de medir la desviación entre ν_σ y algún estimador $\tilde{\nu}_{\sigma,n}$ en este artículo es a través de la siguiente norma, que es similar a la norma supremo,

$$l_s(\tilde{\nu}_{\sigma,n}, \nu_\sigma) = \sup \left\{ \left| \int f d\tilde{\nu}_{\sigma,n} - \int f d\nu_\sigma \right| : f \in F_s \right\},$$

donde

$$F_s := \left\{ f : \int (1 + |u|)^s |\mathcal{F}f(u)| du \leq 1 \right\}.$$

La función $\mathcal{F}f$ corresponde a la transformada de Fourier de la función f . Puede verse que la clase F_s está conformada por funciones uniformemente acotadas y equicontinuas. La consistencia con respecto a la convergencia débil implica $l(\tilde{\nu}_{\sigma,n}, \nu_\sigma) \rightarrow 0$. Se menciona que de acuerdo al análisis de Fourier, pedir que $f \in F_s$ es un poco más fuerte que pedir que $f \in C^s$, con $\|f\|_{C^s} \leq 1$, para una norma adecuada de C^s .

La métrica que se usa para calcular tasas de convergencia y que además satisface las condiciones (2.7) y (2.8) es

$$d^{(2)}(\psi, \phi) = \sum_{k=0}^2 \left\| \psi^{(k)} - \phi^{(k)} \right\|_{L^\infty_w},$$

donde $\|\cdot\|_{L^\infty_w}$ es la norma L^∞ ponderada por la función de peso w , es decir,

$$\|\phi\|_{L^\infty_w} = \sup_{u \in \mathbb{R}} |\phi(u) w(u)|.$$

De esta forma los estimadores de ν_σ y μ , $\widehat{\nu}_{\sigma,n}$ y $\widehat{\mu}_n$, respectivamente se definen como aquellos que cumplen

$$d^{(2)}(\widehat{\varphi}_n, \varphi(\cdot; \widehat{\mu}_n, \widehat{\nu}_{\sigma,n})) \leq \inf_{\tilde{\mu} \in \mathbb{R}, \tilde{\nu}_\sigma \in \mathcal{M}(\mathbb{R})} d^{(2)}(\widehat{\varphi}_n, \varphi(\cdot; \tilde{\mu}, \tilde{\nu}_\sigma)) + \delta_n,$$

con $\delta_n = O(n^{-1/2})$. Pese a que se estiman la medida canónica y μ , a diferencia de $\bar{\nu}_\sigma$ y \bar{b} (como en (2.6)), puede demostrarse que el Teorema 2.3 es válido para $\widehat{\mu}$ y $\widehat{\nu}_\sigma$, es decir, son fuertemente consistentes. El uso de una norma que involucra las segundas derivadas se debe a la relación entre la primera y la segunda derivada de la función característica con la medida de Lévy, como a continuación se menciona.

Heurísticamente, mientras más rápido converja a cero $|\varphi|$ en los extremos, más complicado será estimar ν_σ , debido a la fórmula (2.2). Se consideran entonces tres casos particulares de decaimiento en la función característica para calcular las tasas de convergencia:

- (a) Con parte gaussiana. Cuando tienen término de difusión, es decir, $\sigma^2 > 0$.
- (b) Con decaimiento exponencial. La función característica decae a lo más exponencialmente, es decir, existe $\alpha > 0$ tal que

$$\operatorname{Re} \log \varphi(u) \geq -\alpha |u|, \quad \text{para todo } u \in \mathbb{R}.$$

- (c) Decaimiento polinomial de orden β . Para algunos $\beta \geq 0$ y $C > 0$, se tiene que

$$|\varphi(u)| \geq C(1 + |u|)^{-\beta}, \quad \text{para toda } u \in \mathbb{R}.$$

Los siguientes dos teoremas muestran que los estimadores $\widehat{\nu}_{\sigma,n}$ y $\widehat{\mu}_n$ son óptimos; el primer teorema da cotas superiores de convergencia y el segundo muestra que dichas cotas también son cotas inferiores en algún sentido, es decir, las tasas de convergencia son óptimas.

Teorema 2.4 *Sea $w : \mathbb{R} \rightarrow \mathbb{R}$ definida como $w(u) = (\log(e + |u|))^{-1/2-\gamma}$, donde γ es cualquier número positivo. Los estimadores $\widehat{\mu}_n$ y $\widehat{\nu}_{\sigma,n}$ satisfacen las siguientes condiciones*

$$\mathbb{E}_{\mu, \nu_\sigma} |\mu - \widehat{\mu}_n| = O(n^{-1/2}).$$

Además, para cualquier $s > 0$,

$$l_s(\widehat{\nu}_{\sigma,n}, \nu_\sigma) = O_{P_{b,\nu_\sigma}} \left(\frac{1}{\sqrt{n}} \sup_{u \in [0, U_n]} \left\{ \frac{(1+u)^{2-s}}{w(u) |\varphi(u)|} \right\} \right),$$

donde

$$U_n = \inf \left\{ u > 0 : \frac{(1+u)^2 n^{-1/2}}{w(u) |\varphi(u; b, \nu_\sigma)|} \geq 1 \right\}.$$

Para los casos especiales a considerar, se tiene que las tasas respectivas de convergencia en P_{μ, ν_σ} -probabilidad son:

(a) Parte exponencial, $(\log n)^{-s/2}$.

(b) Decaimiento exponencial, $(\log n)^{-s}$.

(c) Decaimiento polinomial de orden $\beta \geq 0$, $\left[(\log n)^{1/2+2\gamma} n^{-1/2} \right]^{s/\beta} \vee n^{-1/2}$.

El siguiente teorema demuestra que las tasas de convergencia calculadas en el Teorema (2.4) para los casos especiales (a), (b) y (c); son óptimas.

Teorema 2.5 Para K y $\bar{K} > 0$ suficientemente grandes y para cualesquiera $\alpha > 0$ y $\beta \geq 0$ se definen las siguientes clases de medidas ν_σ :

$$\mathcal{A}_{(a)}(K, \sigma) = \{ \nu_\sigma \in \mathcal{M}(\mathbb{R}) : \nu_\sigma(\mathbb{R}) \leq K \} \quad (\sigma > 0),$$

$$\mathcal{A}_{(b)}(K, \alpha) = \{ \nu_\sigma \in \mathcal{M}(\mathbb{R}) : \nu_\sigma(\mathbb{R}) \leq K, \operatorname{Re}(\log \varphi(u)) \geq -\alpha |u| \} \quad (\sigma = 0),$$

$$\mathcal{A}_{(c)}(K, \bar{K}, \beta) = \left\{ \nu_\sigma \in \mathcal{M}(\mathbb{R}) : \nu_\sigma(\mathbb{R}) \leq K, |\varphi(u)| \geq \bar{K}^{-1} (1 + |u|)^{-\beta} \right\} \quad (\sigma = 0).$$

Entonces para algún $b \in \mathbb{R}$ y para cualquier $s > 0$ se obtienen las siguientes tasas inferiores de convergencia, donde $\tilde{\nu}_{\sigma,n}$ denota cualquier estimador de ν_σ basado en n observaciones.

(a) Existe $\varepsilon > 0$ tal que

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{\nu}_{\sigma,n}} \sup_{\nu_\sigma \in \mathcal{A}_{(a)}(C, \sigma)} P_{\mu, \nu_\sigma} \left(\frac{l_s(\tilde{\nu}_{\sigma,n}, \nu_\sigma)}{(\log n)^{-s/2}} > \varepsilon \right) > 0.$$

(b) Existe $\varepsilon > 0$ tal que

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{\nu}_{\sigma,n}} \sup_{\nu_{\sigma} \in \mathcal{A}_{(b)}(C, \alpha)} P_{\mu, \nu_{\sigma}} \left(\frac{l_s(\tilde{\nu}_{\sigma,n}, \nu_{\sigma})}{(\log n)^{-s}} > \varepsilon \right) > 0.$$

(c) Existe $\varepsilon > 0$ tal que

$$\liminf_{n \rightarrow \infty} \inf_{\tilde{\nu}_{\sigma,n}} \sup_{\nu_{\sigma} \in \mathcal{A}_{(c)}(C, \bar{C}, \alpha)} P_{\mu, \nu_{\sigma}} \left(\frac{l_s(\tilde{\nu}_{\sigma,n}, \nu_{\sigma})}{n^{-(s/2\beta) \wedge (1/2)}} > \varepsilon \right) > 0.$$

2.3 Con mínima distancia a la función característica empírica considerando la frecuencia de observación

Se describe el estimador para la medida canónica de Kolmogorov propuesto en Kappus y Reiss [24], publicado en 2010. Se observa un proceso de Lévy cada Δ unidades de tiempo hasta un tiempo T . De manera similar a Neumann y Reiss [26], se construye un estimador de mínima distancia para el término de deriva y la medida canónica de Kolmogorov. Se construyen además tasas óptimas de convergencia para estos, simultáneamente en Δ y T , conforme $\Delta \rightarrow 0$ y $T \rightarrow \infty$.

Un proceso de Lévy $(X_t, t \geq 0)$ se observa en n puntos equidistantes $\Delta, \dots, n\Delta = T$. La función característica de X_{Δ} tiene la forma

$$\varphi_{\Delta}(u) = \mathbb{E}e^{iuX_{\Delta}} = e^{\Delta\Psi(u)},$$

donde Ψ es el exponente característico Ψ . Se asumirá a lo largo de la sección que X_1 tiene momentos de orden $4 + \delta$, para alguna $\delta > 0$. La función característica empírica se escribe como

$$\widehat{\varphi}_{\Delta, T}(u) := \frac{1}{n} \sum_{k=1}^n e^{iu(X_{k\Delta} - X_{(k-1)\Delta})}.$$

Para la construcción de la métrica se consideran no sólo la función característica, sino también sus derivadas. Esto debido a la fórmula (2.2), que señala la relación entre la transformada de Fourier y la segunda derivada de la función característica empírica. Involucrando

2.3. Con mínima distancia a la función característica empírica considerando la frecuencia de observación

la frecuencia de observación, se obtiene

$$\mathcal{F}\nu_\sigma(u) = \frac{1}{\Delta} \left(\frac{\varphi'_\Delta(u)^2}{\varphi_\Delta(u)^2} - \frac{\varphi''_\Delta(u)}{\varphi_\Delta(u)} \right). \quad (2.9)$$

Para obtener un estimador con tasa óptima para $T \rightarrow \infty$ con distancia arbitraria de observación Δ , la elección apropiada de una métrica tendrá que depender de Δ y T . Entonces se define la métrica

$$d_\Delta(\varphi, \psi) := \sum_{k=0}^2 \Delta^{-1 \wedge k/2} \left\| \varphi^{(k)} - \psi^{(k)} \right\|_{L^\infty(w)}, \quad (2.10)$$

donde

$$\|f\|_{L^\infty(w)} := \sup_{u \in \mathbb{R}} |f(u)| w(u),$$

para la función de ponderación $w : \mathbb{R} \rightarrow \mathbb{R}^+$, definida como

$$w(u) = (\log(e + |u|))^{-1/2-\delta}, \quad \text{para toda } u \in \mathbb{R},$$

como se definió en la Sección 2.2.

De manera análoga a (2.6), los estimadores $\hat{\mu}_{\Delta,T}$ y $\hat{\nu}_{\sigma_{\Delta,T}}$ son elegidos de manera que

$$d_\Delta(\hat{\varphi}_{\Delta,T}, \varphi_\Delta(\cdot; \hat{\mu}_{\Delta,T}, \hat{\nu}_{\sigma_{\Delta,T}})) \leq \inf_{(\mu, \nu_\sigma)} d_\Delta(\hat{\varphi}_{\Delta,T}, \varphi_\Delta(\cdot; \mu, \nu_\sigma)) + \varepsilon_T, \quad (2.11)$$

con $\varepsilon_T = o(\Delta^{1/2}T^{-1/2})$. Se muestra en Neumann y Reiss [26] que para observaciones equidistantes con Δ fijo, los estimadores de μ y ν_σ definidos de acuerdo a la ecuación (2.11) son fuertemente consistentes bajo condiciones generales en la elección de la métrica d_Δ .

Finalmente, se calculan tasas óptimas de convergencia en probabilidad para dos tipos de decaimiento de la función característica, que abarcan varios de los procesos de Lévy más conocidos, que es lo que muestran los siguientes dos teoremas. Las dos clases de procesos a considerar son las siguientes.

(a) La función característica de X_Δ satisface

$$|\varphi_\Delta(u)| \geq Ce^{-\Delta c|u|^\alpha}, \quad (2.12)$$

para algunas $0 \leq \alpha \leq 2$ y $C, c > 0$.

(b) Se tiene a lo más decrecimiento polinomial de la función característica:

$$|\varphi_{\Delta}(u)| \geq C(1 + |u|)^{-\Delta\beta}, \quad (2.13)$$

para $C > 0$ y $\beta \geq 0$.

La forma de medir la desviación entre ν_{σ} y algún estimador $\tilde{\nu}_{\sigma,n}$ en este artículo es a través de

$$l_s(\tilde{\nu}_{\sigma,n}, \nu_{\sigma}) = \sup \left\{ \left| \int f d\tilde{\nu}_{\sigma,n} - \int f d\nu_{\sigma} \right| : f \in F_s \right\},$$

donde

$$F_s := \left\{ f : \int (1 + |u|)^s |\mathcal{F}f(u)| du \leq 1 \right\}.$$

La función $\mathcal{F}f$ corresponde a la transformada de Fourier de la función f .

A continuación se enuncian los dos resultados del artículo en donde se dan las tasas de convergencia de los estimadores propuestos para las dos clases de procesos mencionadas.

Teorema 2.6 *El estimador de μ , $\hat{\mu}_{\Delta,T}$ satisface*

$$\mathbb{E} |\hat{\mu}_{\Delta,T} - \mu| = O(T^{-1/2}).$$

Para $\hat{\nu}_{\sigma_{\Delta},T}$, se tienen las siguientes tasas de convergencia en probabilidad.

(a) Si la función característica satisface (2.12), entonces

$$l_s(\hat{\nu}_{\sigma_{\Delta},T}, \nu_{\sigma}) = O_P \left(\left(\frac{\log T}{\Delta} \right)^{-s/\alpha} \vee T^{-1/2} \right).$$

(b) Si la función característica satisface (2.13), entonces

$$l_s(\hat{\nu}_{\sigma_{\Delta},T}, \nu_{\sigma}) = O_P \left(T^{-s/(2\Delta\beta)} (\log(e + T))^{\frac{s(1/2+\delta)}{\Delta\beta}} \vee T^{-1/2} \right).$$

El siguiente resultado indica que las tasas de convergencia son óptimas, en el sentido que las cotas superiores son, de alguna forma, cotas inferiores.

2.3. Con mínima distancia a la función característica empírica considerando la frecuencia de observación

Teorema 2.7 Defínanse las siguientes clase de medidas ν_σ :

$$\begin{aligned}\mathcal{A}_{(a)}(C, c, \alpha) &= \left\{ \nu_\sigma : |\varphi_\Delta(u)| \geq C e^{-\Delta c |u|^\alpha} \right\}, \\ \mathcal{A}_{(b)}(K, \alpha) &= \left\{ \nu_\sigma : |\varphi_\Delta(u)| \geq C (1 + |u|)^{-\Delta\beta} \right\}.\end{aligned}$$

Entonces se obtienen las siguientes cotas inferiores uniformemente para $|b| \leq B$, donde B es alguna constante positiva:

(a) Existe $\varepsilon > 0$ tal que

$$\liminf_{T \rightarrow \infty, \Delta \in (0, 1]} \inf_{\bar{\nu}_{\sigma_\Delta, T}} \sup_{\nu_\sigma \in \mathcal{A}_{(a)}(C, c, \alpha)} P_{\mu, \nu_\sigma} \left[\left(\left(\frac{\log T}{\Delta} \right)^{s/\alpha} \wedge T^{1/2} \right) l_s(\bar{\nu}_{\sigma_\Delta, T}, \nu_\sigma) > \varepsilon \right] > 0.$$

(b) Existe $\varepsilon > 0$ tal que

$$\liminf_{T \rightarrow \infty, \Delta \in (0, 1]} \inf_{\bar{\nu}_{\sigma_\Delta, T}} \sup_{\nu_\sigma \in \mathcal{A}_{(b)}(C, \beta)} P_{\mu, \nu_\sigma} [T^{s/(2\Delta\beta) \wedge 1/2} l_s(\bar{\nu}_{\sigma_\Delta, T}, \nu_\sigma) > \varepsilon] > 0,$$

donde el ínfimo se toma sobre todos los posibles estimadores $\bar{\nu}_{\sigma_\Delta, T}$ de ν_σ basados en observaciones del proceso de Lévy, con distancia Δ hasta el tiempo T .

Capítulo 3

Estimación de la densidad de Lévy

La medida de Lévy puede ser indirectamente estimada a través de la densidad de Lévy, siempre que esta exista; esto es, que la medida de Lévy se pueda expresar como $\nu(dx) = \rho(x)dx$, para alguna función $\rho : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}_+$. Cuando existe la densidad de Lévy, pueden usarse métodos tipo estimación kernel. Por otro lado, la forma de la densidad de Lévy dice un poco más acerca de la distribución de los saltos que la medida de Lévy.

En este capítulo se describen dos trabajos que estiman la densidad de Lévy de un proceso de Lévy $\{X_t, t \geq 0\}$ que se observa discreta y uniformemente en los puntos $1, \dots, n$; estos son Gugushvili [19] y [19]. En ambos se asume la existencia del segundo momento del proceso y además que el proceso de Lévy es de actividad finita, es decir, que el número esperado de saltos de $\{X_t, t \geq 0\}$ por unidad de tiempo, es finita.

Bajo las condiciones anteriores, el exponente característico (1.1) se puede expresar como

$$\Psi(u) = iu\tilde{b} - \frac{\sigma^2}{2}u^2 + \int (e^{iux} - 1) \rho(x) dx, \quad (3.1)$$

donde $\tilde{b} = b - \int_{|x|<1} x \rho(x)dx$. La intensidad del proceso se denotará por λ , es decir, $\lambda = \int \rho(x)dx$. La fórmula (3.1) implica que el proceso $\{X_t, t \geq 0\}$ es la suma de un movimiento browniano con deriva \tilde{b} y difusión σ , y un proceso de Poisson compuesto con intensidad λ y distribución de saltos $f(x) := \rho(x)/\lambda$.

3.1 Con inversión de Fourier en términos de la deriva, la intensidad y la difusión

A continuación se hace la descripción de Gugushvili ([19]), quien estima el término de difusión, la intensidad de los saltos y la media del proceso, para los que calcula tasas de convergencia. Posteriormente estima la densidad de Lévy usando el teorema de inversión de Fourier, donde se usan argumentos de sustitución y de estimación kernel. Finalmente, calcula la tasa de convergencia para este estimador, usando el error cuadrático medio integrado. La idea general para calcular estos parámetros, que enseguida se definen, es usar funciones similares a las funciones kernel para estimación de densidades de probabilidad y posteriormente calcula tasas superiores de convergencia para estos estimadores midiendo la discrepancia con el error cuadrático medio.

Se asume que ρ pertenece a la clase de funciones

$$W(\beta, L, \Lambda, K) = \left\{ \rho : \rho(x) = \lambda f(x), f \text{ es densidad, } \int u^2 f(u) du \leq K, \int |u|^\beta |\mathcal{F}f(u)| du \leq L, \lambda \in (0, \Lambda] \right\}, \quad (3.2)$$

donde β, L, Λ y K son números positivos y $\mathcal{F}f$ es la transformada de Fourier de f . Además se supone que $\sigma \in (0, \Sigma]$ y $|b| \leq \Gamma$, donde Σ y Γ son positivos.

Se describirá la forma en que se estima σ^2 , pues las ideas son similares para los otros parámetros que son reales. La idea es usar funciones tipo kernel.

Sea $v : \mathbb{R} \rightarrow \mathbb{R}$ con soporte $[-1, 1]$ y que satisface

$$\int_{-1}^1 v(u) du = 0, \quad \int_{-1}^1 \left(-\frac{u^2}{2} \right) v(u) du = 1 \quad \text{y} \quad v(u) = O(u^\beta) \text{ cuando } u \rightarrow 0, \quad (3.3)$$

donde β es el mismo que en (3.2). El ancho de banda h , para este y los demás estimadores dependerá de n de la forma

$$h_n = (\alpha \log n)^{-1/2}, \quad 0 < \alpha < \Sigma^{-2}.$$

Se escribirá h en lugar de h_n , teniendo en cuenta que h depende de n . Se define la función tipo kernel, v^h , como

$$v^h(u) = h^3 v(hu),$$

para toda $u \in \mathbb{R}$.

Ahora, puede verse que

$$\log |\varphi(u)| = -\lambda + \lambda \operatorname{Re}(\mathcal{F}_f(u)) - \frac{\sigma^2}{2}u^2,$$

de donde se obtiene que, por (3.3),

$$\int_{-1/h}^{1/h} \log |\varphi(u)| v^h(u) du = \lambda \int_{-1/h}^{1/h} \operatorname{Re}(\mathcal{F}_f(u)) v^h(u) du + \sigma^2. \quad (3.4)$$

Se tiene que $\int_{-1/h}^{1/h} \operatorname{Re}(\mathcal{F}_f(u)) v^h(u) du \rightarrow 0$ conforme $h \rightarrow 0$. De esta forma se tiene una expresión para σ^2 en términos de la función característica, por lo que un estimador podría obtenerse al sustituirse la función característica en (3.4), por la función característica empírica. De este modo, un estimador para σ^2 está dado por

$$\tilde{\sigma}_n^2 = \int_{-1/h}^{1/h} \{[\log |\hat{\varphi}_n(u)| \wedge M_n] \vee -M_n\} v^h(u) du, \quad (3.5)$$

donde $h_n^2 M_n = O(\log n)$ y $h_n M_n \rightarrow \infty$. Es necesario truncar con la sucesión $\{M_n\}_n$ para demostrar la consistencia en el sentido que a continuación se enuncia.

Teorema 3.1 *El estimador de σ^2 propuesto en (3.5) satisface la desigualdad*

$$\sup_{|\tilde{b}| \leq \Gamma} \sup_{\sigma \in (0, \Sigma]} \sup_{\rho \in W(\beta, L, \Lambda, K)} \mathbb{E}(\sigma^2 - \tilde{\sigma}_n^2)^2 \leq A (\log n)^{-\beta-3},$$

para alguna constante A .

Análogamente se dan estimadores para λ y b , para los cuales también se demuestra consistencia. El Teorema 3.1, indica que el error cuadrático medio de $\tilde{\sigma}^2$ más grande de entre todas las posibles características del proceso de Lévy, está acotado superiormente por una función que decrece logarítmicamente a cero. Resultados análogos serán formulados para el resto de los parámetros estimados.

Tomando una función $w : \mathbb{R} \rightarrow \mathbb{R}$ con soporte $[-1, 1]$ y que satisface

$$\int_{-1}^1 w(u) du = -1, \quad \int_{-1}^1 u^2 w(u) du \quad \text{y} \quad w(u) = O(u^\beta) \quad \text{cuando } u \rightarrow 0,$$

donde β es el mismo que en (3.2). También sea w^h tal que $w^h(u) = hw(hu)$.

Teorema 3.2 *El estimador de λ dado por*

$$\tilde{\lambda}_n = \int_{-1/h}^{1/h} \{[\log |\widehat{\varphi}_n(u)| \wedge M_n] \vee -M_n\} w^h(u) du,$$

es consistente en el sentido siguiente:

$$\sup_{|\tilde{b}| \leq \Gamma} \sup_{\sigma \in (0, \Sigma]} \sup_{\rho \in W(\beta, L, \Lambda, K)} \mathbb{E} \left(\lambda - \tilde{\lambda}_n \right)^2 \leq B (\log n)^{-\beta-1},$$

para alguna constante B .

Defina la función tipo kernel $y^h(u) = h^2 y(hu)$, donde la función y es real y continua, tiene soporte en $[-1, 1]$ y satisface

$$\int_{-1}^1 u y(u) du = 1, \quad y \quad y(u) = O(u^\beta) \text{ cuando } u \rightarrow 0,$$

donde β es el mismo que en (3.2).

Finalmente, el estimador de \tilde{b} que se propone es

$$\tilde{b}_n = \int_{-1/h}^{1/h} \{[\text{Im}(\text{Log} \widehat{\varphi}_n(u)) \wedge M_n] \vee -M_n\} y^h(u) du,$$

donde Log representa el logaritmo complejo, para el cual, de acuerdo al Teorema B.5, está definido para los $\omega \in \Omega$ tales que $\widehat{\varphi}_n(\cdot)(\omega)$ no se anula en $[-1/h, 1/h]$.

Teorema 3.3 *El estimador de \tilde{b} , \tilde{b}_n , satisface la desigualdad*

$$\mathbb{E} \left(\tilde{b} - \tilde{b}_n \right)^2 \leq C (\log n)^{-\beta-2},$$

para alguna constante B .

Ahora se estima la densidad de Lévy y se da una cota superior usando el error cuadrático medio integrado. De (3.1) se obtiene que la transformada de Fourier de ρ se expresa como

$$\mathcal{F}\rho(u) = \text{Log} \left(\frac{\varphi(u)}{e^{iu\tilde{b}} e^{-\lambda} e^{-\frac{\sigma^2}{2} u^2}} \right).$$

Luego, usando el Teorema B.3 se obtiene que

$$\rho(x) = \frac{1}{2\pi} \int e^{-iux} \text{Log} \left(\frac{\varphi(u)}{e^{iu\tilde{b}} e^{-\lambda} e^{-\frac{\sigma^2}{2} u^2}} \right) du. \quad (3.6)$$

Un estimador para ρ podría obtenerse reemplazando en (3.6) los parámetros \tilde{b} , λ y σ^2 por sus estimadores calculados en la sección anterior. Restaría dar un estimador de φ , para ello note que como φ es integrable (como consecuencia de que $\rho \in W(\beta, L, \Lambda, K)$), entonces X_1 tiene densidad continua, la cual puede aproximarse por un estimador kernel para densidades,

$$q_n(x) = \frac{1}{nh} \sum_{j=1}^n k\left(\frac{x - \Delta X_j}{h}\right),$$

donde k es un kernel. Ahora, como la transformada de Fourier de q_n es $\hat{\varphi}_n(u) \mathcal{F}q_n(ht)$, donde $\mathcal{F}q_n$ es la transformada de Fourier de q_n ; entonces φ en (3.6) se sustituirá por $\hat{\varphi}_n(u) \mathcal{F}q_n(ht)$. Como en el caso de los estimadores $\tilde{\sigma}_n^2$, $\tilde{\lambda}_n$ y \tilde{b}_n , se requiere truncar para demostrar consistencia. De este modo, el estimador propuesto para ρ es

$$\begin{aligned} \tilde{\rho}_n(x) &= -i\tilde{b}_n \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-iux} u du + \tilde{\lambda}_n \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-iux} du + \frac{\tilde{\sigma}_n^2}{2} \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-iux} u^2 du \\ &\quad + \frac{1}{2\pi} \int_{-1/h}^{1/h} e^{-iux} \{[\log |\hat{\varphi}_n(u) \mathcal{F}k(uh)| \wedge M_n] \vee -M_n\} du \\ &\quad + \frac{i}{2\pi} \int_{-1/h}^{1/h} e^{-iux} \{[\arg(\hat{\varphi}_n(u) \mathcal{F}k(uh)) \wedge M_n] \vee -M_n\} du. \end{aligned}$$

Para estudiar las propiedades asintóticas del estimador $\tilde{\rho}_n$, se tomará una función kernel en particular. Sea $k(x) = \sin x / (\pi x) / (\pi x)$. Además se medirá el desempeño del estimador usando el error cuadrático medio integrado (MISE, por sus siglas en inglés),

$$\text{MISE}(\tilde{\rho}_n) = \mathbb{E} \int |\tilde{\rho}_n(x) - \rho(x)|^2 dx.$$

Se tiene que $\tilde{\rho}_n$ es consistente en el sentido que el error cuadrático medio integrado converge.

Teorema 3.4 *Además de las condiciones sobre λ , σ^2 , b y ρ , suponga que*

$$\int |u|^{2\beta} |\mathcal{F}f(u)|^2 dt < \infty,$$

entonces se cumple la desigualdad

$$\text{MISE}(\tilde{\rho}_n) \leq D (\log n)^{-\beta},$$

para alguna constante D .

En la sección anterior, además de calcular un estimador para ρ , se calcularon tasas superiores de convergencia. Aquí se calcula una tasa inferior de convergencia para la densidad de Lévy.

Teorema 3.5 *Sea \mathcal{T} el conjunto de todas las ternas de Lévy tales que $|b| \leq \Gamma$, $\sigma \in (0, \Sigma]$, $\lambda \in (0, \Lambda]$ y además*

$$\int |u|^{2\beta} |\mathcal{F}f(u)|^2 du \leq E,$$

para $\beta \geq 1/2$ y $E > 0$. Entonces

$$\inf_{\hat{\rho}_n} \sup_{\mathcal{T}} \kappa \text{MISE}(\tilde{\rho}_n) \geq (\log n)^{-\beta},$$

para alguna constante $\kappa > 0$. El ínfimo está tomado sobre todos los estimadores $\hat{\rho}_n$ basado en observaciones $\Delta X_1, \dots, \Delta X_n$.

3.2 Con inversión de Fourier en términos de la difusión

El estimador que se describe aquí, propuesto en Gugushvili [20], es para la densidad de Levy. Se expresa la transformada de Fourier de la densidad de Lévy en términos del coeficiente browniano y la función característica; se aplica un teorema de inversión; y se sustituyen σ^2 y la función característica por un estimador de σ^2 y por la función característica empírica. Además se calculan tasas de convergencia para la densidad de Lévy, usando el error cuadrático medio integrado.

De la fórmula 3.1 se tiene que

$$\varphi'(u) = \varphi(u) \left(i\tilde{b} - \sigma^2 u + i \int e^{iux} x \rho(x) dx \right).$$

Despejando y derivando nuevamente se obtiene que

$$\int e^{iux} x^2 \rho(x) dx = \frac{(\varphi'(u))^2 - \varphi(u) \varphi''(u)}{\varphi^2(u)} - \sigma^2.$$

Para $x \neq 0$, se aplica el Teorema B.3 se llega a

$$\rho(x) = \frac{1}{2\pi x^2} \int e^{-iux} \left\{ \frac{(\varphi'(u))^2}{\varphi^2(u)} - \frac{\varphi''(u)}{\varphi(u)} - \sigma^2 \right\} du. \quad (3.7)$$

El estimador que se propone en este artículo, se obtiene al sustituir en (3.7) φ por la función característica empírica y σ^2 por un estimador que se obtiene con un método tipo estimación kernel para densidades.

El integrando de (3.7) es integrable por el supuesto de que la segunda derivada de la transformada de Fourier de ρ , $\mathcal{F}\rho''$, es integrable. Sin embargo, al hacer la sustitución, no necesariamente va a ser integrable, además de que para valores grandes del valor absoluto de u , causan problemas computacionales, por lo que se propone una modificación que trunca la función característica y usa un kernel de estimación. Esto es,

$$\widehat{\rho}(x) = \frac{1}{2\pi x^2} \int e^{-iux} \left\{ \frac{[\widehat{\varphi}'_n(u)]^2}{\widehat{\varphi}_n^2(u)} \mathbf{1}_{G_u} - \frac{\widehat{\varphi}''_n(u)}{\widehat{\varphi}_n(u)} \mathbf{1}_{G_u} - \widehat{\sigma}^2 \right\} \mathcal{F}k(ht) du, \quad (3.8)$$

donde $\mathcal{F}k$ denota la transformada de Fourier de una función kernel k , con ancho de banda h . La integral en (3.8) es finita si $\mathcal{F}k$ tiene soporte compacto. El conjunto G_u que trunca la función se define como

$$G_u = \left\{ |\widehat{\varphi}_n(u)| \geq \kappa_n e^{-\Sigma^2/(2h)} \right\}. \quad (3.9)$$

Observe que G_t depende de h , Σ y una sucesión de constantes κ_n ; todas ellas se describirán en la siguiente sección. Finalmente, $\widehat{\sigma}$ se define como en [19], es decir,

$$\widehat{\sigma}^2 = \int_{-1/h}^{1/h} \{[\log |\widehat{\varphi}_n(u)| \wedge M_n] \vee -M_n\} v^h(u) du, \quad (3.10)$$

donde v^h también es una función tipo kernel que se describe en la siguiente sección.

Para justificar el estimador propuesto y dar sus propiedades asintóticas, primero se enuncian todas las condiciones requeridas. Se asume que la densidad de Lévy pertenece a la clase de funciones

$$\begin{aligned} W(\beta, L, K, \Lambda) = \{ & \rho : \rho(x) = \nu(\mathbb{R}) f(x), f \text{ es una densidad de probabilidad,} \\ & \int |u|^\beta |\mathcal{F}f(u)| du \leq L, \int x^2 f(x) dx \leq K, \\ & \mathcal{F}f'' \text{ es integrable, } \nu(\mathbb{R}) \in (0, \Lambda] \}, \end{aligned}$$

donde β , L , K y Λ son números positivos. También se asume que $\sigma \in (0, \Sigma]$ y $\left| \widetilde{b} \right| \leq \Gamma$, para algunas $\Sigma, \Gamma > 0$.

Por otro lado, también se asume que el ancho de banda para w y v se define como

$$h = h_n = (\eta \log n)^{-1/2}, \quad 0 < \eta < \Sigma^{-2}/2.$$

Para calcular tasas de convergencia se toma el kernel $k(x) = \sin x / (\pi x)$. Observe que la transformada de Fourier de k tiene soporte compacto, $\mathcal{F}k(u) = \mathbf{1}_{[-1,1]}(u)$. La sucesión $\{\kappa_n\}_n$ en (3.9) se toma como $\kappa_n = \kappa (\log(3 \log(3n)))^{-1}$, para una constante $\kappa > 0$. Esta es una condición técnica que se requiere para las demostraciones.

En [19] se mostró que bajo las condiciones que a continuación se describen, $\tilde{\sigma}^2$ es un estimador consistente en el sentido de error cuadrático medio. Sea $v^h(u) = h^3 v(ht)$ la función tipo kernel, donde v es continua, real, con soporte $[-1, 1]$ y tal que

$$\int_{-1}^1 v(u) du = 0, \quad \int_{-1}^1 \left(-\frac{u^2}{2}\right) v(u) du = 1 \quad \text{y} \quad v(u) = O(u^\beta) \quad \text{cuando } u \rightarrow 0.$$

Las cotas de truncamiento en (3.10) son tales que $h_n^2 M_n = \log \log(3n)$. El resultado de consistencia para $\tilde{\sigma}$ es el siguiente.

Teorema 3.6 *El estimador de σ^2 propuesto en (3.10) satisface la desigualdad*

$$\sup_{|\tilde{b}| \leq \Gamma} \sup_{\sigma \in (0, \Sigma]} \sup_{\rho \in W(\beta, L, \Lambda, K)} \mathbb{E} (\sigma^2 - \tilde{\sigma}_n^2)^2 \leq A (\log n)^{-\beta-3},$$

para alguna constante A .

Como en la sección anterior, el Teorema 3.6 indica que el máximo error cuadrático medio, tomado sobre todos las posibles características del proceso, está acotado por una función que decrece lento a cero. Ahora, las tasas de convergencia para la densidad de Lévy se dan en los siguientes resultados.

Teorema 3.7 *Sea \mathcal{T} la colección de las ternas $(\tilde{b}, \sigma^2, \rho)$ tales que $|\tilde{b}| \leq \Gamma$, $\sigma \in (0, \Sigma]$ y $\rho \in W(\beta, L, K, \Lambda)$. Entonces se tiene que*

$$\sup_{\mathcal{T}} \mathbb{E} [\hat{\rho}(x) - \rho(x)]^2 \leq B (\log n)^{-\beta},$$

para una constante $B > 0$ y cada $x \neq 0$.

Teorema 3.8 Sea \mathcal{T} la colección de las ternas $(\tilde{b}, \sigma^2, \rho)$ tales que $|\tilde{b}| \leq \Gamma$, $\sigma \in (0, \Sigma]$ y $\nu(\mathbb{R}) < \infty$ y además

$$\int |u|^\beta |\mathcal{F}_f(u)| du \leq L.$$

Entonces para cada x se satisface la desigualdad

$$\inf_{\tilde{\rho}_n} \sup_{\mathcal{T}} \mathbb{E} [\tilde{\rho}_n(x) - \rho(x)]^2 \geq C (\log n)^{-\beta},$$

para una constante $C > 0$. El ínfimo está tomado sobre todos los estimadores $\tilde{\rho}_n$ basados en los incrementos del proceso.

Capítulo 4

Discusión

En esta sección se hace una comparación cualitativa entre los trabajos descritos en los Capítulos 2 y 3, es decir, de Neumann y Reiss [26], Kappus y Reiss [24], Gugushvili [19], Gugushvili [20] y Watteel y Kulperger [29]. Se distinguen básicamente dos enfoques de muestreo, de alta y de baja frecuencia. Se considera que se tiene una alta frecuencia de observaciones cuando se asume que los saltos de un proceso de Lévy pueden detectarse. Por otro lado, se considera baja frecuencia de observaciones cuando no se puede saber si un incremento grande es debido a un salto (o una infinidad) o solo a la parte de difusión. Todos los trabajos mencionados en los Capítulos 3 y 4 trabajan bajo un esquema de baja frecuencia de observaciones, salvo [24], que pretende conseguir un equilibrio entre alta y baja frecuencia.

Los supuestos que hacen sobre el proceso de Lévy para calcular las características del proceso y sus tasas de convergencia son muy diferentes. Tienen todos en común que la clase de procesos que consideran tienen difusión y saltos, además de que hacen supuestos fuertes sobre los momentos.

En Neumann y Reiss [26] y Kappus y Reiss [24] se usa un criterio de mínima distancia entre la función característica y la función característica generada con las características sobre las cuales se efectúa la minimización. En ambos se asume la existencia del momento $4 + \delta$, para algún $\delta > 0$, con el fin de mostrar que sus estimadores son tasa óptima; sin embargo muestran la consistencia fuerte de sus estimadores sin asumir la existencia de momentos,

mas que algunas condiciones generales sobre una distancia que usan para minimizar. La diferencia entre ellos es que en [24] los estimadores y las tasas de convergencia involucran tanto el tiempo hasta el cual se observa el proceso, como la distancia entre las observaciones (que ya se mencionó que es constante); a diferencia de [26], donde se considera únicamente el tiempo hasta el cual se observa el proceso.

En Gugushvili [19] y [20] se asume la existencia de la densidad de Lévy, del segundo momento del proceso y la integrabilidad de la función característica. Una de las restricciones más fuertes es que se consideran procesos de Lévy con actividad finita, es decir, considera procesos de Lévy que a lo sumo tienen una cantidad finita de saltos por unidad de tiempo. El método usado es expresar la densidad de Lévy en términos de la función característica, usando inversión de Fourier; después aplicar un argumento de sustitución, es decir, sustituir la función característica por su versión empírica, haciendo algunos ajustes. Usa ideas del método de estimación de densidades usando funciones kernel. La diferencia entre estos trabajos es que en [19] se estiman algunos parámetros para poder estimar después la densidad de Lévy (sustituyendo estos por sus estimadores), calculando tasas de convergencia para cada uno de ellos; en [20] la finalidad es estimar la densidad de Lévy, por lo que se hace esto de una manera más directa; de hecho necesita estimar únicamente el coeficiente de difusión y no estima el coeficiente de deriva.

El método de sustitución es también el que se usa en Watteel y Kulperger [29]. De hecho la idea general es la misma que en [19] y [20], expresa la medida canónica de Kolmogorov en términos de la función característica usando inversión de Fourier y después aplica la sustitución. El problema con el estimador que se obtiene para la medida canónica de Kolmogorov, es que podría obtenerse algo que no es medida y se debe tener control sobre ciertos parámetros en la estimación; es por ello que este trabajo cuenta con una sección que considera correcciones para que el estimador que propone sea, efectivamente, una medida y también considera aspectos numéricos y de implementación. Demuestra que la medida estimada de cada intervalo es asintóticamente normal; para ello supone la existencia del momento $5 + \delta$, para alguna $\delta > 0$.

4. Discusión

La siguientes tablas resume las diferencias entre los procesos de Lévy que se consideran en los trabajos estudiados. A continuación se representan los trabajos de Neumann y Reiss [26], Kappus y Reiss [24], Gugushvili [19], Gugushvili [20] y Watteel y Kulperger [29] por NR, KR, Gu1, Gu2 y WK, respectivamente.

	Forma de estimar ν	¿Estima σ^2 ?
WK	Combinándola con σ^2	Sí, indirectamente
Gu1	A través de la densidad de Lévy	Sí
Gu2	A través de la densidad de Lévy	Sí
NR	Combinándola con σ^2	Sí, indirectamente
NK	Combinándola con σ^2	Sí, indirectamente

	¿Estima la deriva?	Actividad infinita	Sin densidad de Lévy
WK	✗	✓	✓
Gu1	✓	✗	✗
Gu2	✗	✗	✗
NR	✓	✓	✓
KR	✓	✓	✓

	Criterio de convergencia	Condición de momentos
WK	Normalidad Asintótica	$5 + \delta$
Gu1	Error cuadrático medio integrado	2
Gu2	Error cuadrático medio	2
NR	Norma tipo supremo	$4 + \delta$
KR	Norma tipo supremo	$4 + \delta$

Capítulo 5

Estimador tipo *sieves* basado en método de momentos

En este capítulo se da una propuesta de estimador para la medida canónica de Kolmogorov, la cual es fácil de calcular e implementar. La idea general es “parametrizar” la medida canónica de Kolmogorov con la base de ondículas de Haar, igualar la función característica empírica con la teórica en distintos puntos y resolver el sistema resultante para obtener los estimadores de los coeficientes de la base de Haar. El objetivo es dar una propuesta de estimador para la medida de Lévy que sea más sencillo de implementar, comparado con los estimadores que se estudian en los Capítulos 2 y 3. Para el método de momentos que se propone, no es necesario que se tenga actividad finita o densidad de Lévy; basta con que el segundo momento del proceso sea finito. La idea de usar la base de Haar surgió después de estudiar la forma en que se implementa el estimador para la medida canónica de Kolmogorov en Neumann y Reiss [26]. Sin embargo es importante mencionar que en el estimador que se propone podría sustituirse la base de ondículas de Haar por alguna otra familia de modelos lineales de dimensión finita, lo cual se reconsidera en el Capítulo 6.

Se ha elegido estimar la medida canónica de Kolmogorov en lugar de la medida de Lévy; el motivo es que, como ya se mencionó en el Capítulo 2, con la medida canónica de Kolmogorov, se estiman con facilidad el primer y el segundo momento, además de que es una medida

finita, a diferencia de la medida de Lévy, que no necesariamente lo es. Al ser finita la medida canónica de Kolmogorov, puede aproximarse bien en un intervalo finito, a diferencia de una medida que no es finita, como puede ser la medida de Lévy.

Sea $\{X_t, t \geq 0\}$ un proceso de Lévy con deriva μ y medida canónica de Kolmogorov ν_σ , el cual se ha observa en los puntos $1, \dots, n$, tomando $X_0 = 0$. Este estimador es válido para procesos de Lévy con segundo momento finito. En las siguientes dos secciones se desarrolla el método de momentos y en la tercera sección, se realiza una analogía con el método *sieves*.

5.1 Método *sieves*

Se explica en esta sección cómo se aproxima la medida canónica de Kolmogorov en términos de la base de ondículas de Haar. La base de ondículas de Haar ha sido seleccionada como un estudio inicial debido a que esta posee buenas propiedades aproximantes. Al ser aproximada con la base de ondículas de Haar, la medida canónica de Kolmogorov se expresa de una manera muy sencilla (como una suma finita de funciones indicadoras, multiplicadas por los coeficientes de la base de Haar a estimar); consecuentemente, las integrales con esta, se calculan con facilidad, además de que quedan lineales en los coeficientes a estimar. Por otro lado, al ser una aproximación con funciones escalonadas, la base de ondículas de Haar permite detectar posibles saltos en la medida canónica de Kolmogorov.

Podría parecer contradictorio que se hable de estimación no paramétrica, cuando se está parametrizando mediante los coeficientes de la base de ondículas de Haar. Los métodos de inferencia no paramétrica encuentran su justificación cuando se tiene un tamaño de muestra grande; a su vez, en el método de momentos que aquí se propone, mientras más grande sea el tamaño de muestra, crece más el número de parámetros, y es esto lo que convierte a este método en no paramétrico. Para ilustrar esta afirmación, considere los histogramas. Los histogramas se consideran como un instrumento estadístico no paramétrico. Sin embargo, hay una base finita de funciones escalonadas que se usa para representar una densidad arbitraria; cuando aumenta el número de observaciones, se puede aumentar el número de

funciones escalonadas, es decir, se puede aumentar el número de clases que tiene el intervalo, y obtener un estimador consistente de la densidad de manera no paramétrica.

Esta propuesta de aproximar con la base de ondículas de Haar la medida canónica de Kolmogorov puede verse como un estimador tipo *sieves*, el cual fue propuesto por Grenander [18] en 1981 y usado por Figueroa [15] en 2009 para estimar la densidad de Lévy (asumiendo una frecuencia alta de observaciones). El método de *sieves* consiste en aproximar una función por el mejor elemento de alguna familia de modelos lineales de dimensión finita, que pueden volverse más y más complejos en tanto aumenta la cantidad de datos. En Grenander [18] se describe el método *sieves* para encontrar este mejor elemento.

Figueroa [15], sigue el método de *sieves*, con la diferencia que él no recurre a máxima verosimilitud, sino otro criterio de minimización. Asimismo, identifica en este método tres componentes principales. Se elige primero una familia de dimensión finita de funciones lineales, que recibe el nombre de *sieves*, con buenas propiedades aproximantes. Ejemplos comunes de este tipo de familias son los *splines*, los polinomios trigonométricos, o las ondículas. El segundo paso es especificar una distancia entre funciones para determinar el mejor aproximador del objetivo, en el caso de [15], la densidad de Lévy. Finalmente, se propone un estimador para la mejor aproximación del objetivo en el modelo lineal, que involucra un problema de optimización. Para el estimador que aquí se propone, la familia de *sieves* que se elige es la base de ondículas de Haar y el criterio de optimización es el de encontrar el valor del parámetro que solucione un sistema de ecuaciones que resulta de igualar la función característica con su parte empírica; en la siguiente sección se describe a detalle este sistema de ecuaciones. Dado que se considera una igualdad para después resolver un sistema, no es necesario definir explícitamente una distancia entre funciones.

Por facilidad de cálculo, se estima la medida canónica en un intervalo con extremos que son potencia de dos; sin embargo, no necesariamente tiene que hacerse así. Se va a estimar la medida canónica de Komogorov en el intervalo $[-2^M, 2^M]$, para algún entero M , donde el refinamiento más fino es el que se obtiene al dividir $[-2^M, 2^M]$ en subintervalos ajenos de longitud 2^{-J} , para el entero $J > -M$. Se tiene entonces que para cada $x \in \mathbb{R}$, que el

estimador propuesto para ν_σ es

$$\widehat{\nu}_{\sigma,\theta}(x) = \sum_{j=-J}^M \sum_{k=-2^{M-j}}^{2^{M-j}-1} \alpha_{j,k} \phi_{j,k}(x), \quad (5.1)$$

donde $\phi_{j,k}(x) = 2^j \phi(2^{-j}x - k)$ y

$$\phi(x) = \begin{cases} -1 & \text{si } x \in (0, 1/2], \\ 1 & \text{si } x \in (1/2, 1], \\ 0 & \text{si } x \notin (0, 1]. \end{cases}$$

Defina el vector θ como aquel cuya primera entrada es μ y el resto son todos los $\alpha_{j,k}$ de la ecuación (5.1). La dimensión de θ es

$$\dim \theta = 1 + \sum_{j=-M}^J 2^{M+j+2} = 2^{J+M+2} - 1. \quad (5.2)$$

Observe que se está abusando de la notación al tomar la medida de Lévy como una función. En realidad se está estimando la función definida por

$$F_{\nu_\sigma}(x) = \nu_\sigma((-\infty, x]),$$

para cada $x \in \mathbb{R}$, la cual determina a la medida ν_σ .

Una observación importante, que será crítica en la implementación descrita en el siguiente capítulo, es el hecho que la fórmula (5.1) podría no cumplir con ser no negativa o creciente. Esto se debe a que los parámetros $\alpha_{j,k}$ están libres, es decir, no tienen restricciones que impongan estas propiedades. Por esta razón, el método de estimación propuesto en la siguiente sección y su implementación imponen las restricciones requeridas. En una primera exploración de la viabilidad de este método, se optó por considerar un número de parámetros y un tamaño de muestra grandes. Esto arrojó que obviar las restricciones no es posible en la práctica, no obstante que la teoría asintótica garantiza una buena aproximación.

Es importante mencionar que en el estimador que en este capítulo se propone podría sustituirse la base de ondículas de Haar por alguna otra familia de modelos lineales de dimensión finita. Esta noción será retomada en el Capítulo 6.

5.2 Método de momentos

Se propone como estimador de θ el vector que resulta de igualar la función característica con su versión empírica, evaluadas en l puntos de la recta real, donde $l = 2^{J+M+1} - 1$ si se utiliza la base de Haar; note que $\dim \theta = 2l + 1$. Puede compararse este método con el propuesto en la Sección 1.2.3.

Sean $t_1, \dots, t_l \in \mathbb{R}$. Defínanse los vectores z_θ y z_n , tales que

$$z_\theta = \begin{pmatrix} \mu \\ \operatorname{Re} \varphi(t_1; \theta) \\ \vdots \\ \operatorname{Re} \varphi(t_l; \theta) \\ \operatorname{Im} \varphi(t_1; \theta) \\ \vdots \\ \operatorname{Im} \varphi(t_l; \theta) \end{pmatrix} \quad \text{y} \quad z_n = \begin{pmatrix} X_n/n \\ \operatorname{Re} \widehat{\varphi}_n(t_1) \\ \vdots \\ \operatorname{Re} \widehat{\varphi}_n(t_l) \\ \operatorname{Im} \widehat{\varphi}_n(t_1) \\ \vdots \\ \operatorname{Im} \widehat{\varphi}_n(t_l) \end{pmatrix},$$

donde

$$\varphi(t_m; \mu, \theta) := \exp \left\{ it_m \mu + \int \frac{e^{it_m x} - 1 - it_m x}{x^2} \widehat{\nu}_{\sigma, \theta}(dx) \right\}, \quad (5.3)$$

para $m = 1, \dots, l$. Hay $2l + 1$ parámetros, y se va a resolver el sistema de ecuaciones resultante de igualar z_θ con z_n , como a continuación se describe; es por ello que se ha definido así la primera coordenada de estos vectores, para tener un sistema de $2l + 1$ ecuaciones con $2l + 1$ incógnitas. Observe que X_n/n es la media muestral de los incrementos de las observaciones del proceso de Lévy (que son variables aleatorias independientes e idénticamente distribuidas), pues

$$\frac{X_n}{n} = \frac{1}{n} \sum_{j=1}^n (X_j - X_{j-1}).$$

Al desarrollar la integral de la fórmula (5.3), puede demostrarse que la expresión dentro de la exponencial es una combinación lineal de los parámetros, pues $\widehat{\nu}_{\sigma, \theta}$ tiene la forma (5.1).

Sea $\tilde{\theta}$ la solución a la ecuación

$$z_\theta = z_n. \quad (5.4)$$

El estimador de la medida canónica de Kolmogorov que se propone es $\widehat{\nu}_{\sigma, \tilde{\theta}}$. Se tiene que este es un estimador consistente en el sentido que

$$\lim_{M \rightarrow \infty} \lim_{J \rightarrow \infty} \lim_{n \rightarrow \infty} \widehat{\nu}_{\sigma, \tilde{\theta}} = \nu_{\sigma},$$

para casi todo punto en \mathbb{R} . Dado que las partes real e imaginaria de una función característica son par e impar, respectivamente, entonces los puntos t_m , $m = 1, \dots, l$ pueden tomarse en los números reales no negativos, sin pérdida de generalidad.

5.2.1 Normalidad asintótica

El estimador de θ , $\tilde{\theta} \in \mathbb{R}^{2l+1}$, resulta ser un M -estimador, como a continuación se muestra. Los M -estimadores son fuertemente consistentes y asintóticamente normales; en esta sección se enuncia el resultado para el estimador que se propone.

Defina la función $\psi : \mathbb{R} \times \mathbb{R}^{2l+1} \rightarrow \mathbb{R}^{2l+1}$ como

$$\psi(y, \theta) = \begin{pmatrix} \mu - y \\ \operatorname{Re} \varphi(t_1; \theta) - \cos(y t_1) \\ \vdots \\ \operatorname{Re} \varphi(t_l; \theta) - \cos(y t_l) \\ \operatorname{Im} \varphi(t_1; \theta) - \operatorname{sen}(y t_1) \\ \vdots \\ \operatorname{Im} \varphi(t_l; \theta) - \operatorname{sen}(y t_l) \end{pmatrix}.$$

De esta forma se ve que $\tilde{\theta}$ es el mismo que el M -estimador resultante de resolver la ecuación

$$\sum_{j=1}^n \psi(X_j - X_{j-1}, \theta) = 0. \quad (5.5)$$

Si existe $\tilde{\theta}$ que resuelva la ecuación (5.5) y $\mathbb{E}\psi(X_1, \theta)$ tiene una matriz de derivadas no singular, dígame Λ , en el verdadero valor del parámetro, θ_0 ; entonces, de acuerdo a (Huber [22], Corolario 3.2, pág. 133), $\sqrt{n}(\tilde{\theta} - \theta_0)$ es asintóticamente normal con media cero y matriz de covarianzas

$$\Lambda^{-1} \operatorname{var}(\psi(X_1, \theta_0)) (\Lambda^T)^{-1}.$$

Dado que para $a < b$, se puede expresar $\widehat{\nu}_{\sigma, \tilde{\theta}}([a, b])$ como combinación lineal $\tilde{\theta}$, también se tiene que $\widehat{\nu}_{\sigma, \tilde{\theta}}([a, b])$ es a su vez asintóticamente normal. De manera más general, para cualquier función g derivable en θ se obtiene que $g(\tilde{\theta})$ es asintóticamente normal, aplicando el método delta.

Capítulo 6

Implementación, nueva propuesta y bandas de confianza

Este capítulo se compone de cuatro secciones. La primera de estas contiene comentarios sobre cálculo e implementación de estimadores de los trabajos de estimación de la medida de Lévy discutidos en los Capítulos 2 y 3. Llevar a la práctica los estimadores que se estudiaron en los Capítulos 2 y 3 acarrea varios problemas, y esto dio motivación a proponer un estimador más sencillo de calcular, que es el que se describió en el Capítulo 5.

Al implementar el estimador tipo *sieves* basado en momentos, usando la base de ondículas de Haar, que se propone en el Capítulo 5, se otro tipo de dificultades, las cuales son descritas en la segunda sección. En la tercera sección se sustituye la base de ondículas de Haar por otro modelo lineal de dimensión finita que consiste en aproximar la medida canónica de Kolmogorov por una medida compuesta únicamente de saltos en un retícula especificada y se presentan los resultados que ilustran el potencial de este modelo lineal como otra alternativa de estimación con ventajas de implementación. Detalles finos de eficiencia computacional y propiedades teóricas se plantean como trabajo futuro.

En la última sección se discute cómo podrían construirse bandas de confianza *bootstrap* para los diferentes estimadores de la medida de Lévy que en este trabajo se han tratado. Es de interés contar con bandas de confianza porque así se puede complementar un estimador

puntual (donde “puntual” es, en este caso, una función). En efecto, el ancho de la banda de confianza muestra qué tanta incertidumbre inferencial existe sobre la forma del estimador de la medida de Lévy.

6.1 Comentarios sobre cálculo e implementación de estimadores

En esta sección se describen los procedimientos para implementar los estimadores propuestos en los Capítulos 2 y 3. La implementación del estimador de la medida canónica de Kolmogorov que se propone en el Capítulo 5, consiste en resolver un sistema de ecuaciones. Fuera de este, las otras propuestas involucran aplicar fórmulas de inversión de Fourier para cada punto donde la medida de Lévy es estimada, lo cual ya hace ineficiente su cómputo. Además del tiempo de cómputo, se pueden producir errores muy grandes debido a los denominadores en la fórmula (2.2), que es usada (o alguna similar, como en [20]) en los trabajos que se están considerando.

En Watteel y Kulperger [29] se hace una propuesta de estimador que consiste únicamente en aplicar la fórmula de inversión de Fourier para obtener la medida canónica de Kolmogorov, mediante el uso de la función característica empírica. Este estimador podría ni siquiera ser una medida; puede ocurrir que no sea monótono creciente. Por esta razón, se proponen en tal trabajo algunas técnicas de suavizamiento, con la finalidad de que el estimador de la medida sea monótono; cada una de estas técnicas es un problema de optimización. Por otro lado, este estimador es sensible a la región donde se integra para aproximar la integral en la fórmula de inversión de Fourier, y se propone entonces un criterio para determinar esta región en ese trabajo.

En Gugushvili [19] y [20], nuevamente se tiene que estimar la fórmula de inversión para cada punto donde se calcula la densidad de Lévy. Además de ello, en ambos trabajos se estiman varios parámetros (los cuales se sustituyen en la fórmula de inversión de Fourier) de forma similar a la estimación de densidades tipo kernel, lo que involucra el cálculo numérico

de más integrales. Por otro lado, se tienen que usar funciones y parámetros que cumplan ciertas condiciones generales (por ejemplo, las funciones v y w en la Sección 3.1). Dependiendo de cuáles se tomen, los resultados podrían variar considerablemente.

Finalmente, recuérdese que en los trabajos de Neumann y Reiss [26] y Kappus y Reiss [24] se usa un criterio de mínima distancia a la función característica empírica. En [26] se da una propuesta de implementación que es computacionalmente muy costosa; esta propuesta es válida también para implementar el estimador en [24]. Además de la poca eficiencia computacional, lo que se hace es aproximar el método propuesto por otro; es decir, no se implementa al pie de la letra el método que el artículo profesa, pues para ello, habría que minimizar sobre el espacio de medidas finitas.

La aproximación mencionada en [26] consiste en tres pasos:

1. Proponer estimadores piloto.
2. Parametrizar estos estimadores.
3. Minimizar la distancia mencionada alrededor de una vecindad de los parámetros.

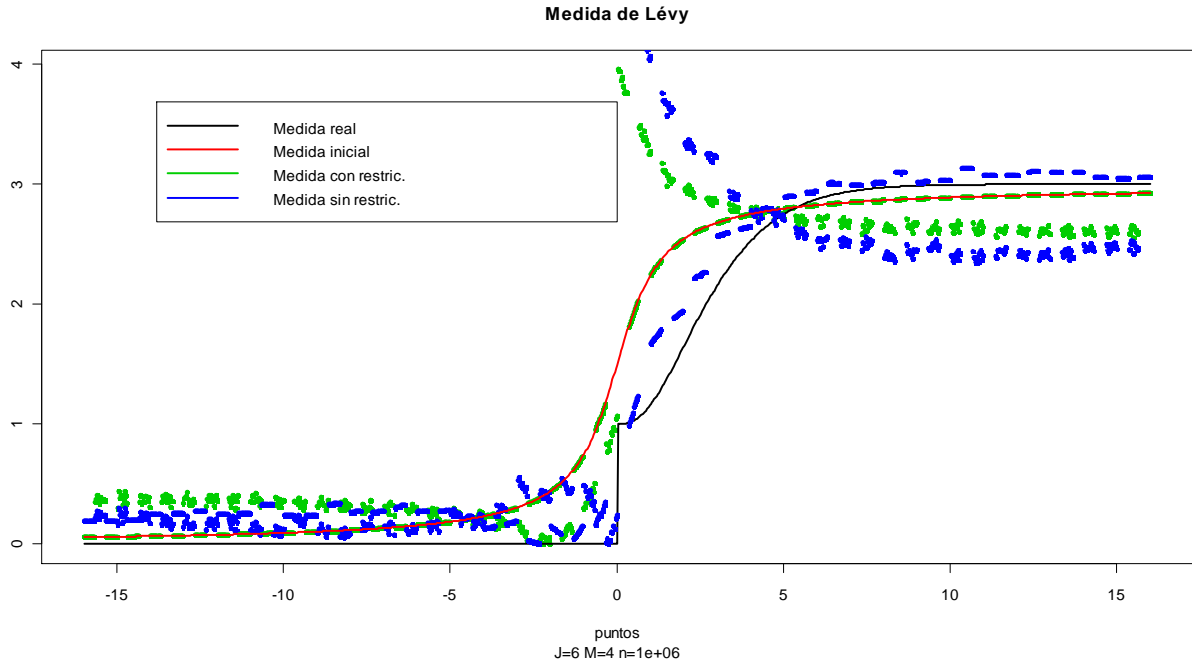
Recuerde que en [26] y [24] se estima la medida canónica de Kolmogorov y la media del proceso. Así, para el primer paso, el estimador piloto de la media es la media muestral y el estimador de la medida canónica es el que se obtiene al aplicar la fórmula de inversión a la fórmula (2.2) (algo similar a como se hace en [29], pero sin los procedimientos de suavizamiento). Luego, el segundo paso se consigue parametrizando la medida canónica mediante el uso de la base de ondículas de Haar para un número finito de coeficientes. De [26] es donde surge la idea del estimador que se propone en el Capítulo 5. Finalmente, alrededor de una vecindad de cada uno de los coeficientes calculados de la base de ondículas de Haar y de la media muestral, se hace la minimización de la distancia que se define en [26] (o en [24]).

6.2 Implementación del estimador tipo sieves basado en momentos

Como ya se mencionó en el Capítulo 5, la base de Haar tiene algunas ventajas para ser usada en la estimación de la medida canónica de Kolmogorov. Entre estas, puede detectar saltos de la medida y además permite dar una expresión cerrada y sencilla de las integrales de la fórmula (5.3), necesarias para resolver el sistema de ecuaciones.

Por otro lado, la base de Haar tiene algunas desventajas. La primera de ellas es que no se tiene control sobre el crecimiento, esto es, en la fórmula (5.2) se observa que si se desea aproximar por una malla más fina de la base de Haar, el crecimiento de los parámetros es exponencial. Otra debilidad es que no existe restricción sobre los parámetros para que la medida estimada sea, efectivamente, una función no negativa y creciente, como ya se mencionó en la Sección 5.1; aunque asintóticamente el método funciona. Pese a lo anterior, se observó que el módulo de la función característica sí se aproxima muy bien, sin embargo no ocurre lo mismo con las partes real e imaginaria por separado.

Se ha observado a través de simulaciones que lo que debería ser la medida estimada, resulta un función que no es creciente. Una solución a este problema es imponer restricciones sobre algunos puntos para que en ellos, la función estimada obtenida sea creciente. La siguiente figura muestra un caso donde tanto el tamaño de muestra como el número de parámetros son muy grandes (1 000 000 y 4 095, respectivamente); como se indica, se grafica la medida canónica de Kolmogorov real, la estimada y una medida estimada con las restricciones mencionadas, tomando una restricción por cada parámetro. Como puede apreciarse, incluso con un tamaño de muestra y un número de parámetros muy grande e incluso poniendo restricciones sobre la forma de la medida de Lévy, el resultado no es satisfactorio. De hecho para este número de parámetros, no hay diferencia significativa al restringir o no la medida canónica en algunos puntos. Aunque no se ilustra, para el caso $J = 5$ (es decir, 2 047 parámetros), la medida sin restricciones queda totalmente fuera del rango de la medida real.



El ejemplo de proceso de Lévy que se usó para generar la gráfica anterior es un proceso de Lévy que se compone de la suma de un movimiento browniano de varianza $\sigma^2 = 1$ y deriva $b = 1$, más un proceso de Poisson compuesto de intensidad 1 y distribución de saltos exponencial de parámetro 1. La implementación se llevó a cabo en el lenguaje de programación R.

Las propiedades mencionadas para el estimador tipo *sieves* basado en momentos, presentado en el Capítulo 5, serían aún válidas si se cambia la base de ondículas de Haar por algún otro modelo lineal de dimensión finita que sea más conveniente. Otro modelo lineal podría permitir un mejor control sobre la dimensión de los parámetros o evitar poner restricciones sobre la medida estimada; aunque también podría generar el problema de integrar numéricamente si no se pudiera dar una expresión cerrada y manejable para las integrales de la fórmula (5.3).

6.3 Otra propuesta de *sieves*

Es esta sección se considera la implementación de una *sieves* en particular, diferente de la base de ondículas de Haar. En general se observa que esta *sieves* proporciona resultados considerablemente mejores a la base de Haar; esta propuesta se basa en aproximar la medida canónica de Kolmogorov por una medida escalonada, con saltos positivos, es decir,

$$\widehat{\nu}_\sigma(x) = \sum_{j=1}^{2J} \theta_j \mathbf{1}_{\{x \leq x_j\}} \quad (6.1)$$

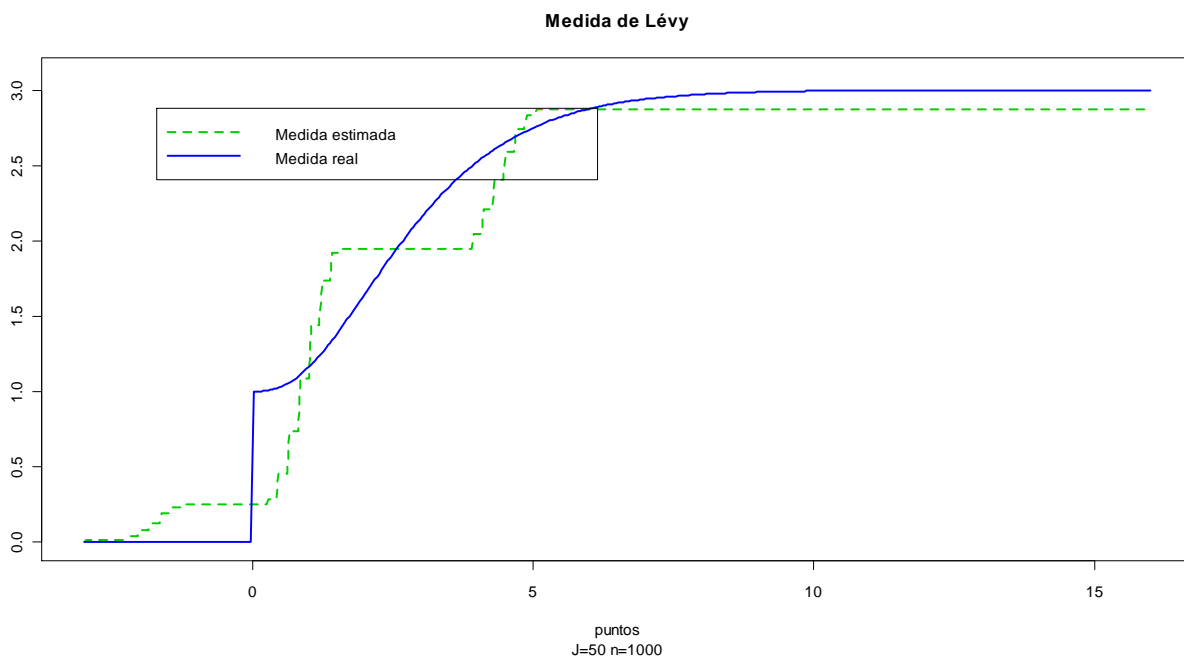
es la aproximación para la medida canónica de Kolmogorov en el intervalo $[a, b]$, donde $\theta_j \geq 0$, $j = 1, \dots, J$ y $a = x_1 < x_2 < \dots < x_J = b$. Observe que el número de parámetros es $2J+1$, incluyendo el parámetro de deriva.

La aproximación definida por la fórmula (6.1) resulta en expresiones cerradas y manejables para las integrales de la fórmula (5.3), que además permite detectar los saltos de la medida canónica de Kolmogorov y su algoritmo es más rápido que el de la base de ondículas de Haar. A continuación se presentan algunas gráficas que ilustran el potencial de este modelo lineal de dimensión finita para el método tipo *sieves* basado en momentos. Los puntos x_j se tomaron equidistantes y, salvo la última gráfica, todas están generadas a partir del mismo proceso de Lévy que se usó en la sección anterior para la base de ondículas de Haar.

La siguiente gráfica muestra que incluso con un número pequeño tanto de parámetros (101) como de muestra (1 000), se obtiene algo superior al mejor caso donde se usó la base

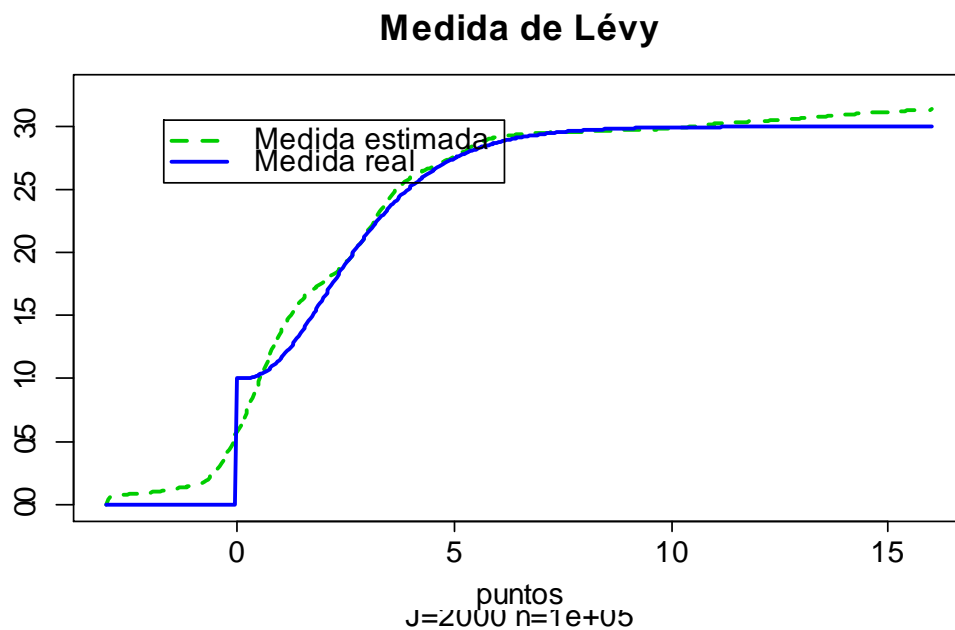
6.3. Otra propuesta de sieves

de ondículas de Haar.

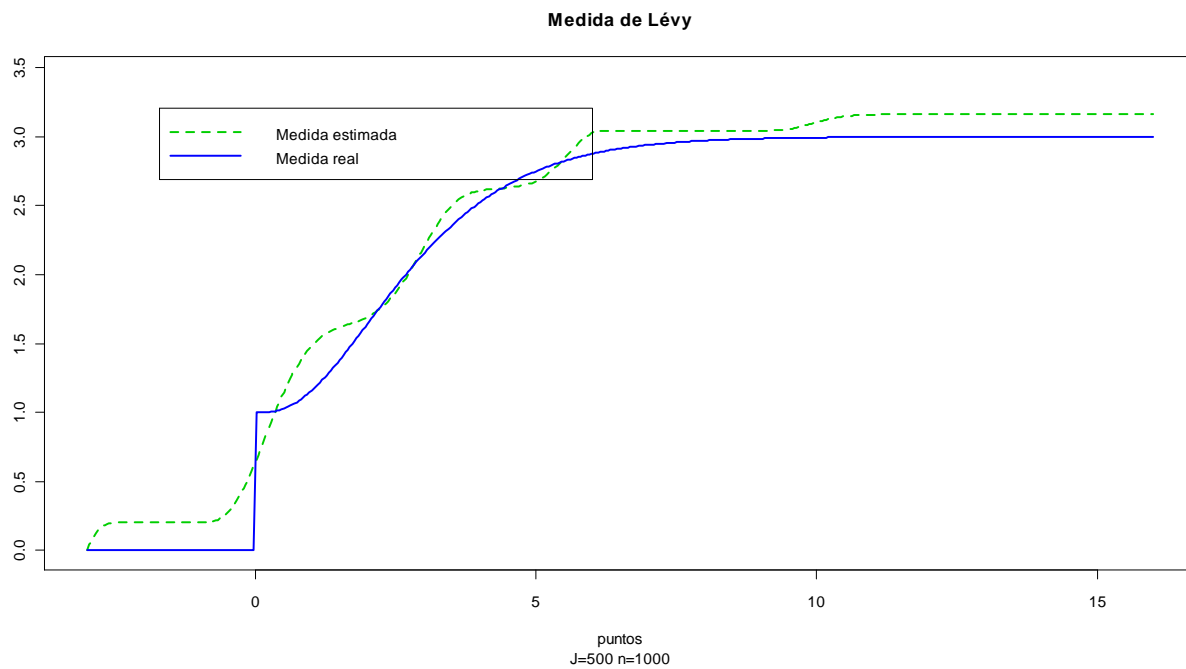


(6.2)

Con el fin de comparar este estimador con el se la sección anterior, se presenta esta gráfica, generada con la misma cantidad de parámetros (4095) y muestra (1 000 000). Se obtiene que la aproximación a la medida canónica de Kolmogorov real es muy buena.



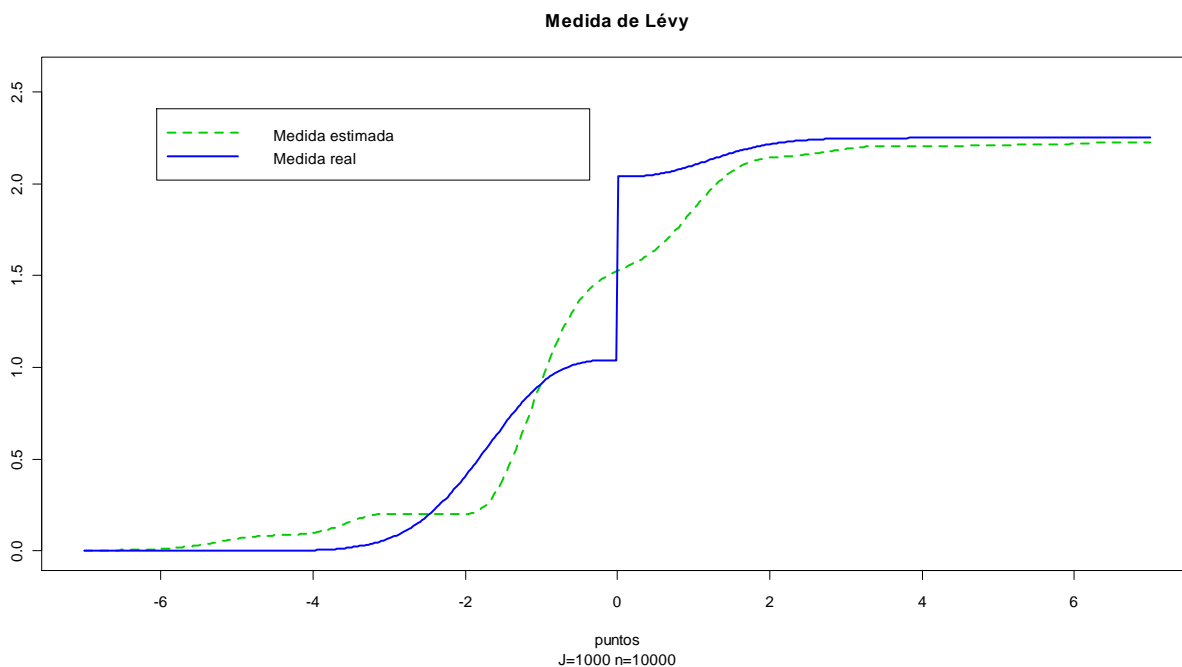
Aumentando el número de parámetros (a 1 001) en (6.2) sin aumentar el tamaño de muestra (1 000), se obtiene un buen estimador para la medida canónica de Kolmogorov, como se observa en la siguiente gráfica.



La siguiente gráfica muestra la medida canónica de Kolmogorov estimada con el mismo método, aplicado ahora a proceso de Lévy que se conforma de la suma de un movimiento browniano con varianza $\sigma^2 = 1$ y deriva -1 y un proceso de Poisson compuesto de intensidad 1 y distribución se saltos normal con media $-1/2$ y varianza 1. Se tomó una muestra de tamaño 10 000 y se aproximó la medida canónica de Kolmogorov con 2 001 parámetros. El

6.3. Otra propuesta de *sieves*

ajuste es razonablemente bueno, dado que el tamaño de muestra no es muy grande.



Pueden explorarse varias opciones para descubrir el verdadero potencial del estimador tipo *sieves* basado en momentos usando este modelo lineal, entre estas acciones se encuentran:

1. Encontrar un número adecuado de parámetros para un tamaño fijo de muestra.
2. En los ejemplos anteriores, los puntos donde salta la medida de Lévy fueron seleccionados de manera equidistante en el intervalo de estimación. Sin embargo, podría trabajarse un sistema adaptativo de puntos de saltos donde la medida de Lévy estimada pueda contener más información (i.e., donde sea menos plana). Una manera de implementar esto es una estimación en dos fases: la primera de ellas es como se ha descrito, y la segunda se efectúa redefiniendo puntos y parámetros con mayor densidad donde la medida de Lévy tiene derivada mayor.
3. Hasta el momento, los puntos donde se evalúan la función característica aproximada por el modelo lineal y la función característica empírica se han tomado equidistantes. Otra forma de seleccionarlos sería donde pudiera contenerse más información acerca de la función característica empírica.

6.4 Bandas *bootstrap* de confianza

En esta sección se plantea de cómo podrían construirse bandas de confianza *bootstrap* para los diferentes estimadores de la medida de Lévy que en este trabajo se han tratado. Las bandas de confianza proporcionan más información acerca de qué tan bueno es el estimador puntual, que en este caso, se trata de una función. Dado que la medida de Lévy describe los saltos del proceso, es importante conocer la forma de esta; las bandas de confianza muestran qué tan diferente puede ser la medida de Lévy estimada de la real. Hay varios ejemplos de cálculo de bandas de confianza para densidades de probabilidad (Hall [21]) o para funciones de distribución de probabilidad. Uno de ellos es el artículo de Bickel y Krieger [4], en el cual se calculan bandas asintóticas de confianza usando *bootstrap* para funciones de distribución, usando el estadístico de Kolmogorov Smirnov.

Una manera de construir bandas de confianza para la medida de Lévy es la siguiente. Para cada punto de la medida de Lévy estimada puede construirse un intervalo de confianza, ya sea a partir de *bootstrap*, normalidad asintótica o distribución exacta (si se está trabajando bajo el supuesto de un modelo finito paramétrico). A partir de estos intervalos de confianza puede construirse una banda de confianza para la densidad de Lévy, siempre y cuando se modifique de manera conveniente la forma de la banda resultante, pues tal cual, subestimaría la probabilidad de que la medida de Lévy se encuentre dentro de esta banda. Bandas de este tipo se calculan en Watteel y Kulperger [29], donde sólo se muestran los gráficos resultantes y se menciona que se usó la normalidad asintótica de su estimador la medida canónica de Kolmogorov sobre intervalos finitos.

Es posible también construir bandas de confianza para estimadores de la medida canónica de Kolmogorov, usando el método *bootstrap*. La idea es estudiar la distribución del supremo de la distancia entre la medida canónica de Kolmogorov y un estimador de ella, de manera similar al estadístico de Kolmogorov-Smirnov, que fue usado para construir bandas de confianza por Bickel y Krieger [4], como ya se mencionó. Considere

$$D(\nu_\sigma, \tilde{\nu}_\sigma) := \sup_{x \in I} |\nu_\sigma(x) - \tilde{\nu}_\sigma(x)|,$$

donde I es el intervalo (suficientemente grande) donde se efectúa la estimación, ν_σ es la medida canónica de Kolmogorov del proceso original y $\tilde{\nu}_\sigma$ es la medida canónica de Kolmogorov estimada, que se obtiene a través de la muestra X_1, \dots, X_n . Se calcula $\tilde{\nu}_\sigma$, el estimador de ν_σ . Luego, se generan p muestras aleatorias del proceso del tamaño n ,

$$\begin{array}{ccc} X_1^{(1)} & \cdots & X_n^{(1)} \\ X_1^{(2)} & \cdots & X_n^{(2)} \\ \vdots & \ddots & \vdots \\ X_1^{(p)} & \cdots & X_n^{(p)} \end{array},$$

de manera que la muestra $X_1^{(j)}, \dots, X_n^{(j)}$, viene del proceso de Lévy cuya medida canónica de Kolmogorov es $\tilde{\nu}_\sigma$ y cuya media es la media muestral, X_n/n . Ahora, se obtienen p estimadores de la medida canónica de Kolmogorov $\tilde{\nu}_\sigma$: $\tilde{\nu}_\sigma^{(1)}, \dots, \tilde{\nu}_\sigma^{(p)}$. Se esta forma, se obtienen $D(\tilde{\nu}_\sigma, \tilde{\nu}_\sigma^{(1)}), \dots, D(\tilde{\nu}_\sigma, \tilde{\nu}_\sigma^{(p)})$. A partir de esta muestra, se estiman el cuantil $Q_{1-\alpha}$. Luego, la banda de confianza al $(1 - \alpha) \times 100\%$ está dada por

$$\mathcal{B} = \left\{ (x, y) : x \in I, y \in \left[\tilde{\nu}(x) - \tilde{Q}_{1-\alpha}, \tilde{\nu}(x) + \tilde{Q}_{1-\alpha} \right] \right\},$$

donde $\tilde{Q}_{1-\alpha}$ es el cuantil estimado. La idea es afín a metodología estándar para construir intervalos de confianza usando bootstrap, particularmente el llamado método de cuantiles (ver Davison y Hinkley [10]).

Es importante notar que, a diferencia del estadístico de Kolmogorov-Smirnov, la distancia $D(\nu_\sigma, \tilde{\nu}_\sigma)$ no necesariamente es libre de distribución. Por esta razón, los cuantiles se estiman vía *bootstrap*. Por otro lado, el desempeño de cada estimador de la medida de Lévy en los estimadores estudiados en los Capítulos 2 y 3, se mide de manera diferente; por ejemplo, a través del error cuadrático medio ([20]), del error cuadrático medio integrado ([19]), o de alguna otra forma, como una norma tipo supremo ([26] y [24]). La distancia $D(\nu_\sigma, \tilde{\nu}_\sigma)$, también permite medir el desempeño de un estimador, ofreciendo la ventaja de que es más sencilla e interpretable (pues mide la máxima diferencia entre una medida y otra).

Apéndice A

Procesos de Lévy

En este apéndice se exponen las características más relevantes de los procesos de Lévy para este trabajo; es de interés especial cómo pueden ser las trayectorias, pues ello permite comprender aspectos de modelación usando procesos de Lévy. La mayor parte de los resultados han sido extraídos de Sato [28] y Cont y Tankov [8].

A.1 Divisibilidad infinita y procesos de Lévy

Es importante mencionar la relación entre variables aleatorias infinitamente divisibles y procesos de Lévy.

Definición A.1 *Se dice que una variable aleatoria Y es **infinitamente divisible** si para cada $n \in \mathbb{N}$, existen Y_1, \dots, Y_n v.a.i.i.d. tales que*

$$Y \stackrel{\mathcal{L}}{=} Y_1 + \dots + Y_n,$$

es decir, Y y $Y_1 + \dots + Y_n$ tienen la misma distribución. En términos de funciones características, la variable aleatoria Y es infinitamente divisible si y solo si para cada $n \in \mathbb{N}$, existe una función característica φ_n tal que la función característica de Y , φ_Y , se escribe como

$$\varphi_Y(u) = [\varphi_n(u)]^n,$$

para toda $u \in \mathbb{R}$.

Los procesos de Lévy están en correspondencia uno a uno con las variables aleatorias infinitamente divisibles. De hecho, Kolmogorov demostró para casos especiales de variables aleatorias infinitamente divisibles la fórmula de Lévy-Khintchine, que a continuación se enuncia, para procesos de Lévy.

A.2 Descomposición de Lévy-Ito

La descomposición de Lévy-Ito describe las trayectorias de un proceso de Lévy. Esta indica que todo proceso de Lévy puede descomponerse como la suma de tres procesos de Lévy independientes; esto es, un movimiento browniano con deriva, un proceso de Poisson compuesto con saltos de longitud mayor a uno y otro proceso límite que puede verse como una suma compensada de procesos de Poisson compuestos. Las definiciones y resultados de esta sección han sido extraídos de Sato [28], donde pueden encontrarse las demostraciones.

Para establecer la descomposición de Lévy-Ito, es necesario primero definir las llamadas medidas aleatorias de Poisson.

Definición A.2 Sean (E, \mathcal{E}, η) y $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida σ -finito y un espacio de probabilidad, respectivamente. Una familia de v.a. en $(\Omega, \mathcal{F}, \mathbb{P})$ con valores en $\mathbb{Z}_+ \cup \{\infty\}$, $\{N(B) : B \in \mathcal{E}\}$, se llama medida aleatoria de Poisson con intensidad η si se cumplen las siguientes condiciones.

- (i) Para cada $B \in \mathcal{E}$, $N(B)$ tiene distribución Poisson de parámetro $\eta(B)$;
- (ii) Si $B_1, \dots, B_n \in \mathcal{E}$ son ajenos, entonces $N(B_1), \dots, N(B_n)$ son v.a. independientes;
- (iii) Para casi todo $\omega \in \Omega$, $N(\cdot, \omega)$ es una medida en E .

El siguiente resultado y su demostración se encuentran en ([28], Teorema 19.2).

Teorema A.1 (Lévy-Ito) Sea $\{X_t; t \geq 0\}$ un proceso de Lévy en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con terna característica (b, σ^2, ν) y defina la medida $\tilde{\nu}$ en $\mathbb{R}_+ \otimes \mathbb{R}$ como

$\tilde{\nu}(dt \times dx) = dt \nu(dx)$. Para cada $B \in \mathbb{R}_+ \otimes \mathbb{R}$ define

$$M(B, \omega) = \begin{cases} \#\{s : (s, X_s(\omega) - X_{s-}(\omega)) \in B\} & \text{si } X(\omega) \text{ es càdlàg,} \\ 0 & \text{en otro caso.} \end{cases}$$

Entonces se cumple lo siguiente.

(i) $\{M(B) : B \in \mathbb{R}_+ \otimes \mathbb{R}\}$ es una medida aleatoria de Poisson en $\mathbb{R}_+ \otimes \mathbb{R}$ con intensidad $\tilde{\nu}$.

(ii) Con probabilidad 1, el proceso

$$X_t^1(\omega) : = \lim_{\varepsilon \downarrow 0} \int_{(0,t] \times \{\varepsilon < |x| < 1\}} \{xM(d(s,x), \omega) - x\tilde{\nu}(d(s,x))\} \quad (\text{A.1}) \\ + \int_{(0,t] \times \{|x| \geq 1\}} xM(d(s,x), \omega)$$

está definido para toda $t \in \mathbb{R}_+$ y la convergencia es uniforme en t en cualquier intervalo acotado. El proceso $\{X_t^1, t \geq 0\}$ es un proceso de Lévy con terna generadora $(0, 0, \nu)$.

(iii) Define $X_t^2 = X_t - X_t^1$. Con probabilidad 1, el proceso X_t^2 es c.s. continuo en t y además tiene terna generadora $(b, \sigma^2, 0)$.

(iv) Los procesos $\{X_t^1, t \geq 0\}$ y $\{X_t^2, t \geq 0\}$ son independientes.

Observe que en el Teorema A.1, $M([0, t] \times A, \omega)$ cuenta el número de saltos de $X(\omega)$ entre el tiempo cero y el tiempo t , cuyo tamaño está contenido en A . El primer sumando en (A.1) es el proceso límite del que se habló al inicio de esta sección (y que resulta ser una martingala cuadrado integrable, pero ese aspecto no será tocado en este trabajo), y el segundo sumando es el proceso de Poisson compuesto con saltos de longitud mayor a uno. Dado que M es una medida aleatoria de Poisson, se tiene que estos dos procesos son independientes entre sí y a su vez son independientes del proceso $\{X_t^2, t \geq 0\}$.

La Fórmula de Lévy Khintchine (1.1) se obtiene del Teorema A.1; basta con observar las funciones características correspondientes de los procesos allí definidos. Un reacomodo en (1.1) es suficiente para obtener lo que podría llamarse versión débil de la descomposición de Lévy-Ito.

Corolario A.1 Sea $\{X_t, t \geq 0\}$ un proceso de Lévy con terna característica (b, σ^2, ν) . Entonces su exponente característico puede expresarse como

$$\Psi(u) = \Psi^{(1)}(u) + \Psi^{(2)}(u) + \Psi^{(3)}(u),$$

donde

$$\begin{aligned} \Psi^{(1)}(u) &= iub - \frac{1}{2}\sigma^2, \\ \Psi^{(2)}(u) &= \nu(\mathbb{R} \setminus (-1, 1)) \int (e^{iux} - 1) \frac{\nu(dx)}{\nu(\mathbb{R} \setminus (-1, 1))} y \\ \Psi^{(3)}(u) &= \sum_{n=0}^{\infty} \left\{ \lambda_n \int_{D_n} (e^{iux} - 1) \frac{\nu(dx)}{\lambda_n} - iu \int_{D_n} x \nu(dx) \right\}, \end{aligned}$$

tomando $D_n = \{x : \frac{1}{2^{n+1}} \leq |x| < \frac{1}{2^n}\}$ y $\lambda_n = \nu(D_n)$.

Del Corolario A.1 se puede concluir que los exponentes característicos $\Psi^{(1)}$, $\Psi^{(2)}$ y $\Psi^{(3)}$ corresponden a un movimiento browniano con deriva, un proceso de Poisson compuesto con saltos de longitud mayor a uno y una suma infinita de procesos de Poisson independientes a los que se les ha restado su media, respectivamente.

El resultado que a continuación se enuncia puede encontrarse en ([28], Teorema 19.3). Afirma que bajo una condición adicional en el Teorema A.1, el proceso de Lévy $\{X_t, t \geq 0\}$ puede expresarse como la suma de dos procesos de Lévy independientes: un movimiento browniano y un proceso de Poisson compuesto.

Teorema A.2 Asuma que el proceso X_t en el Teorema (A.1) satisface $\int_{|x| \leq 1} |x| \nu(dx) < \infty$. Sea b_0 la deriva de X_t . Entonces, con probabilidad 1, el proceso

$$X_t^3(\omega) := \int_{(0,t] \times \{0 < |x|\}} x M(d(s, x), \omega)$$

está definido para toda $t \geq 0$ y satisface la igualdad

$$\mathbb{E}e^{iuX_t^3} = \exp \left\{ t \int (e^{iux} - 1) \nu(dx) \right\}.$$

Defina además $X_t^4 = X_t - X_t^3$. Entonces $\{X_t, t \geq 0\}$ es un proceso de Lévy c.s. continuo tal que

$$\mathbb{E}e^{iuX_t^4} = \exp \left\{ t \left(iub - \frac{\sigma^2}{2}u^2 \right) \right\} \int (e^{iux} - 1) \nu(dx).$$

Los procesos $\{X_t^3, t \geq 0\}$ y $\{X_t^4, t \geq 0\}$ son independientes.

A.3 Ejemplos de procesos de Lévy

Ejemplo A.1 El movimiento browniano es el proceso de Lévy más conocido. Un **movimiento browniano con deriva** b y **desviación estándar** $\sigma > 0$ $\{X_t, t \geq 0\}$ se define como un proceso de Lévy con trayectorias casi seguramente continuas; que además cumple que X_t tiene distribución normal con media bt y varianza $t\sigma^2$. Si $b = 0$ y $\sigma = 1$, se dice que $\{X_t, t \geq 0\}$ es un **movimiento browniano estándar**.

Observe que, por las propiedades de la distribución normal, si $\{Y_t, t \geq 0\}$ es un movimiento browniano estándar, $b \in \mathbb{R}$ y $\sigma^2 > 0$; entonces el proceso definido para cada $t \geq 0$ como $X_t := \sigma Y_t + bt$ es un movimiento browniano con deriva b y desviación estándar $\sigma > 0$. El exponente característico de un movimiento browniano $\{X_t, t \geq 0\}$ con deriva b y desviación estándar $\sigma > 0$ es

$$\Psi(u) = iub - \frac{\sigma^2}{2}u^2.$$

Note que ν es la medida cero.

Ejemplo A.2 Un **proceso de Poisson** con intensidad $\lambda > 0$ es un proceso de Lévy $\{X_t, t \geq 0\}$, tal que para cada $t \geq 0$, X_t tiene distribución Poisson de parámetro λt . Puede demostrarse que $\{X_t, t \geq 0\}$ es un procesos de puros saltos. El exponente característico de este proceso es

$$\Psi(u) = \lambda (e^{iu} - 1).$$

Puede demostrarse que en este caso la medida de Lévy está dada por

$$\nu(dx) = \delta_1(dx).$$

Ejemplo A.3 Los **procesos de Poisson compuestos** son una generalización de los procesos de Poisson. La diferencia reside en que estos últimos son procesos de puros saltos positivos unitarios, mientras que un proceso de Poisson compuesto es un proceso de puros saltos que siguen una ley F . Formalmente, si $\{\xi_n\}_{n \geq 1}$ es una sucesión de v.a.i.i.d. y $\{N_t, t \geq 0\}$ es un proceso de Poisson de parámetro $\lambda > 0$, entonces el proceso $\{X_t, t \geq 0\}$ definido para cada $t \geq 0$ como

$$X_t = \sum_{n=1}^{N_t} \xi_n,$$

es un proceso de Poisson compuesto de intensidad $\lambda \geq 0$ y ley F , donde F es la función de distribución de ξ_1 . El exponente característico en este caso es

$$\Psi(u) = \lambda \int (e^{iux} - 1) F(dx). \quad (\text{A.2})$$

De (A.2) se ve que

$$\nu(dx) = \lambda F(dx).$$

Observe además que un proceso de Poisson es un proceso de Poisson compuesto donde las v.a. $\{\xi_n\}_n$ son degeneradas en uno.

Ejemplo A.4 Los **procesos α -estables** son aquellos donde X_1 es una variable aleatoria α -estable. Estas tienen la propiedad de ser cerradas bajo suma de copias independientes, salvo parámetros de localización y escala, esto es, si n es un número natural cualquiera y Y_1, \dots, Y_n son copias independientes de la variable α -estable Y , entonces existen $a_n > 0$ y $b_n \in \mathbb{R}$ tales que

$$Y_1 + \dots + Y_n = a_n Y + b_n. \quad (\text{A.3})$$

Se ha demostrado que a_n en (A.3) es de la forma $n^{1/\alpha}$, para algún $\alpha \in (0, 2]$, que es precisamente lo que da el nombre a esta clase de variables aleatorias. Si $\alpha = 2$, Y tiene distribución normal y si $\alpha = 1$, tiene distribución de Cauchy. Estos son los únicos casos donde hay una expresión cerrada para la densidad de probabilidad, en los demás casos se puede expresar como una serie. Una propiedad interesante de las variables α -estables es

son de colas pesadas, específicamente, si $\alpha \in (0, 2)$, entonces únicamente se existe hasta el momento α . Para $\alpha \in (0, 1) \cup (1, 2)$, el exponente característico de Y , puede expresarse como

$$\Psi(u) = c|u|^\alpha \left(1 - i\beta \tan \frac{\pi\alpha}{2} \operatorname{sgn} u\right) + iu\eta,$$

para algunos $\beta \in [-1, 1]$, $\eta \in \mathbb{R}$ y $c > 0$; además la medida de Lévy se expresa como

$$\nu(dx) = \begin{cases} c_1 x^{-1-\alpha} dx & \text{si } x > 0 \\ c_2 |x|^{-1-\alpha} dx & \text{si } x < 0 \end{cases},$$

donde $c = c_1 + c_2$ y $\beta = (c_1 - c_2) / (c_1 + c_2)$.

A.4 Funciones de centro

Otras fórmulas similares a (1.1) son usadas y a continuación se describen en detalle. Es importante tener presentes dos propiedades que son fáciles de demostrar. La primera que es si A es boreliano de \mathbb{R} , tal que su cerradura no contiene al cero, entonces $\nu(A) < \infty$. También se tiene que la suma de dos procesos de Lévy independientes es proceso de Lévy. Esto se puede demostrar viendo que la suma de sus exponentes característicos tiene también la forma (1.1).

Sea $c : \mathbb{R} \rightarrow \mathbb{R}$ una función medible y acotada tal que

$$c(x) = 1 + o(x), \quad x \rightarrow 0 \quad \text{y} \quad c(x) = O(1/|x|), \quad |x| \rightarrow \infty. \quad (\text{A.4})$$

Entonces (1.1) puede escribirse como

$$\Psi(u) = iub_c - \frac{\sigma^2}{2}u^2 + \int [e^{iux} - 1 - iux c(x)] \nu(dx), \quad (\text{A.5})$$

donde $b_c = b + \int x [c(x) - \mathbf{1}_{\{|x|<1\}}] \nu(dx)$.

Si además c es medible y para cada $u \in \mathbb{R}$, la función $e^{iux} - 1 - iux c(x)$ es integrable con respecto a ν , entonces la terna $(b, \sigma^2, \nu)_c := (b_c, \sigma^2, \nu)$ caracteriza al proceso de Lévy $\{X_t, t \geq 0\}$. Se hará referencia a esta clase de funciones c como **funciones de centro**. Algunos ejemplos de estas funciones son:

(i) $c(x) = \mathbf{1}_{\{|x| < \varepsilon\}}$, $\varepsilon > 0$.

(ii) $c(x) = \frac{1}{1+x^2}$.

(iii) $c(x) = \mathbf{1}_{\{|x| \leq 1\}} + (2 - |x|) \mathbf{1}_{\{1 < |x| \leq 2\}}$.

(iv) $c(x) = \mathbf{1}_{[-1,1]}(x) + \frac{1}{x} \mathbf{1}_{(1,\infty)}(x) - \frac{1}{x} \mathbf{1}_{(-\infty,-1)}(x)$.

(v) $c(x) = \frac{\text{sen } x}{x}$.

Observe que al sustituir $c(x) = \mathbf{1}_{\{|x| \leq \varepsilon\}}$ en (A.5) se obtiene

$$\Psi(u) = iu\tilde{b} - \frac{1}{2}\sigma^2 + \int [e^{iux} - 1 - iux\mathbf{1}_{\{|x| < \varepsilon\}}] \nu(dx),$$

donde $\tilde{b} = b - \int x \mathbf{1}_{\{\varepsilon \leq |x| < 1\}} \nu(dx)$. De aquí se ve que el uno dentro de la indicadora $\mathbf{1}_{\{|x| < 1\}}$ en la fórmula (1.1) puede ser sustituido por cualquier otra constante positiva. La función de centro $c(x) = (1+x^2)^{-1}$ es usada en Neumann y Reiss [26].

El teorema siguiente se encuentra en ([28], Teorema 8.7). Asevera que bajo algunas condiciones, una sucesión converge si y sólo si su terna característica satisface ciertas condiciones.

Teorema A.3 *Sea $c : \mathbb{R} \rightarrow \mathbb{R}$ continua, acotada y que satisface (A.4). Suponga que para cada $n \in \mathbb{N}$, P_n es una distribución de probabilidad infinitamente divisible con función característica φ_n y con terna característica $(b_n, \sigma^2, \nu)_c$. Sea P una probabilidad. Entonces $P_n \xrightarrow{\mathcal{L}} P$ si y solo si P es infinitamente divisible y su función característica, φ , tiene representación de Lévy-Khintchine $(b, \sigma^2, \nu)_c$ con b, σ^2, ν que satisfacen las siguientes condiciones:*

(i) *Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es continua, acotada y se desvanece en una vecindad del cero, entonces*

$$\lim_{n \rightarrow \infty} \int f(x) \nu_n(dx) = \int f(x) \nu(dx).$$

(ii) *Definiendo*

$$\sigma_{n,\varepsilon}^2 = \sigma_n^2 + \int_{|x| \leq \varepsilon} x^2 \nu_n(dx),$$

se tiene que

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \rightarrow \infty} |\sigma_{n,\varepsilon}^2 - \sigma^2| = 0.$$

(iii) $b_n \rightarrow b$.

A.5 Momentos de un proceso de Lévy

Hay una estrecha relación entre los momentos de un proceso de Lévy y los momentos de su medida de Lévy. Aquí se explica dicha relación, la cual es usada por varios de los artículos que se estudian en este trabajo.

Definición A.3 Sea $g : \mathbb{R} \rightarrow \mathbb{R}$ una función medible y no negativa. El ***g-momento*** de una medida μ se define como $\int g(x) \mu(dx)$.

Definición A.4 Una función $g : \mathbb{R} \rightarrow \mathbb{R}$ se llama ***submultiplicativa*** si es no negativa y existe una constante $a > 0$ tal que

$$g(x + y) \leq a g(x) g(y), \text{ para } x, y \in \mathbb{R}.$$

Los siguientes dos resultados se encuentran en ([28], Teorema 25.3 y Proposición 25.4, respectivamente).

Teorema A.4 Sea g una función submultiplicativa, localmente acotada y medible. Entonces la existencia del g momento de un proceso de Lévy X no es una propiedad distribucional que dependa del tiempo. Además, X_t tiene g -momento finito para cada $t > 0$ si y sólo si $\nu|_{\{|x|>1\}}$ tiene g -momento finito.

Proposition A.5 Se satisfacen las siguientes propiedades.

1. El producto de dos funciones submultiplicativas es submultiplicativo.
2. Si $g(\cdot)$ es submultiplicativa, entonces también lo es $g(a \cdot + b)^c$, para $a, b \in \mathbb{R}$ y $c > 0$.
3. Sea $0 < a \leq 1$, entonces las siguientes funciones son submultiplicativas: $|x| \vee 1$, $x \vee 1$, $\exp |x|^a$, $\exp(x \vee 0)^a$, $\log(|x| \vee e)$, $\log(x \vee e)$, $\log \log(|x| \vee e^e)$ y $\log \log(x \vee e^e)$.

El Teorema A.4 expresa de manera general la relación mencionada. Sin embargo, vale la pena escribirla de manera más explícita de la forma en que es usada en los artículos que se estudian aquí, tal como se hace en ([8], Proposición 3.13).

Proposición A.6 *Para cada $n \in \mathbb{N}$, se tiene que $\mathbb{E}X_t < \infty$ para algún $t > 0$, equivalentemente, para todo $t > 0$, si y solo si $\int_{|x| \geq 1} |x|^n \nu(dx) < \infty$.*

Observe que, debido a (1.2), si $n \geq 2$, la condición $\int_{|x| \geq 1} |x|^n \nu(dx) < \infty$ es equivalente a $\int |x|^n \nu(dx) < \infty$.

Otro resultado interesante concerniente a momentos es el que se enuncia en ([8], Proposición 3.14), que afirma que una relación similar ocurre con los momentos exponenciales del proceso y de la medida de Lévy.

Teorema A.7 *Para cada $u \in \mathbb{R}$, se tiene que $\int e^{ux} \nu(dx) < \infty$ si y solo si $\mathbb{E}e^{uX_t} < \infty$, para algún $t \geq 0$, si y solo si $\mathbb{E}e^{uX_t} < \infty$, para todo $t \geq 0$.*

A.6 Actividad finita o infinita

Los términos “actividad finita” y “actividad infinita” tienen que ver con el número esperado de saltos de un proceso de Lévy, por unidad de tiempo. Enseguida se explica esto formalmente.

Definición A.5 *Se dice que un proceso de Lévy tiene **actividad finita** si $\int_{|x| \leq 1} \nu(dx) < \infty$. En otro caso se dice que el proceso tiene actividad infinita.*

Observe que en la Definición (A.5), la condición $\int_{|x| \leq 1} \nu(dx) < \infty$ es equivalente a la condición $\int \nu(dx) < \infty$, pues por (1.2), la integral es finita fuera el disco unitario. Una forma de definir la medida de Lévy es la siguiente. Para cada $A \in \mathcal{B}(\mathbb{R})$

$$\nu(A) = \mathbb{E}[\#\{t \in [0, 1] : \Delta X_t \neq 0, \Delta X_t \in A\}]. \quad (\text{A.6})$$

Esto es, $\nu(A)$ es el número esperado de saltos por unidad de tiempo cuya longitud pertenezca a A . Por ejemplo, $\nu((1, 2))$ es el número esperado de saltos por unidad de tamaño mayor que 1 y menor que 2. Recuerde que la condición (1.2) garantiza que todo $B \in \mathcal{B}(\mathbb{R})$ tal que $0 \notin \bar{B}$ cumple $\nu(B) < \infty$, por lo que un proceso de Lévy podría únicamente tener una cantidad infinita de saltos de longitud en una vecindad del cero.

Al definir la medida de Lévy como en (A.6), se ve que un proceso de Lévy tiene actividad finita si y solo si en cualquier intervalo de tiempo, el número esperado de saltos es finito. Quizá sea útil mencionar que en [3] se hace alusión a procesos de Lévy con “intensidad finita”, pero se trata precisamente de procesos de Lévy con actividad finita. Lo que se asume en tal artículo es que la medida de Lévy puede escribirse como $\nu(dx) = \rho(x)dx$ y mencionan que el proceso tiene intensidad finita si $\rho \in L^1(\mathbb{R})$, lo que significa que

$$\infty > \int \rho(x) dx = \int \nu(dx),$$

es decir, que el proceso sea de actividad finita.

A.6.1 Ejemplos

Existen procesos de Lévy con actividad infinita, pero que tienen momentos, así como procesos de Lévy con actividad finita y que no tienen momentos. Aquí se describen algunos ejemplos de tales procesos.

Los procesos gaussianos inversos son una clase de procesos de Lévy cuyas trayectorias son crecientes. Se definen como el tiempo que un movimiento browniano estándar con deriva tarda en cruzar una barrera. Formalmente, si $\{B_t, t \geq 0\}$ es un movimiento browniano estándar y $\beta \in \mathbb{R}_+$, un proceso gaussiano inverso $\{X_t, t \geq 0\}$ se define como

$$X_t = \inf \{s > 0 : B_s + \beta s > t\}.$$

En (Kyprianou [25], pág. 8) se demuestra que, efectivamente, $\{X_t, t \geq 0\}$ es proceso de Lévy y que su medida de Lévy es

$$\nu(dx) = \frac{\mathbf{1}_{\{x \geq 0\}}}{\sqrt{2\pi x^3}} e^{-\frac{\beta^2 x}{2}} dx.$$

Tomando $\beta = 1$, se tiene que

$$\begin{aligned} \nu([-1, 1]) &= \int_{-1}^1 \nu(dx) \\ &= \frac{1}{\sqrt{2\pi}} \int_0^1 \frac{1}{x^{3/2}} e^{-\frac{x}{2}} dx \\ &\geq \frac{e^{-1/2}}{\sqrt{2\pi}} \int_0^1 \frac{1}{x^{3/2}} dx \\ &= \infty. \end{aligned}$$

Por otro lado, para $n \in \mathbb{N}$,

$$\int_{|x| \geq 1} |x|^n \nu(dx) = \frac{1}{\sqrt{2\pi}} \int_1^\infty x^{n-3/2} e^{-\frac{x}{2}} dx < \infty.$$

Entonces este es un ejemplo de un proceso de Lévy que tiene todos sus momentos y además tiene actividad infinita, es decir, una infinidad de saltos esperados por unidad de tiempo.

Considérese ahora un proceso de Lévy con terna característica $(0, 0, \nu)$, donde ν se define como

$$\nu(dx) = \frac{1}{\pi} \frac{1}{x^2 + 1} dx.$$

La medida ν es precisamente la medida de una v.a. con distribución Cauchy, por lo que

$$\int 1 \wedge x^2 \nu(dx) \leq \int \nu(dx) < \infty,$$

lo cual indica que se trata de una medida de Lévy. Se tiene además que

$$\nu([-1, 1]) \leq \int \nu(dx) < \infty$$

y

$$\int_{|x| \geq 1} |x| \nu(dx) = \frac{2}{\pi} \int_1^\infty \frac{x}{x^2 + 1} dx = \infty.$$

Un proceso de Lévy así definido tiene actividad finita, pero carece de momentos.

Apéndice B

Otros resultados

Con el objetivo de que la lectura de este trabajo sea lo más autocontenida posible, se exponen en este apéndice varios temas que mejoran la comprensión del mismo. Particularmente, se presentan varias propiedades de funciones características, como lo son las diferentes versiones de la fórmula de inversión de Fourier. También se introduce la base de ondículas de Haar.

B.1 Propiedades de las funciones características

Son usados en este trabajo varios resultados de funciones características; la mayor parte de los que aquí se exponen, han sido extraídos de Chung [5], donde pueden verse también sus demostraciones. El siguiente teorema se encuentra en ([5], Teorema 6.3.3) y da otro criterio para demostrar convergencia en distribución.

Teorema B.1 *Una sucesión de medidas de subprobabilidad $\{P_n\}_{n \geq 1}$ converge vagamente a P_∞ si y solo si la sucesión de sus correspondientes funciones características $\{\varphi_n\}_{n \geq 1}$ satisface*

$$\int_0^t \varphi_n(u) du \xrightarrow{n \rightarrow \infty} \int_0^t \varphi_\infty(u) du,$$

donde φ_∞ es la función característica de P_∞ .

A continuación se enuncian algunas versiones del Teorema de Inversión de Fourier, que tienen la finalidad de recuperar una medida a partir de su transformada de Fourier. Una

demostración de estos resultados puede encontrarse en ([5], Teoremas 6.2.1, 6.2.3 y 6.2.4, respectivamente), donde se demuestran para medidas de probabilidad, mas siguen siendo válidos para medidas finitas.

Teorema B.2 Si μ es una medida finita con transformada de Fourier $\mathcal{F}\mu$ y $x_1 < x_2$, entonces

$$\mu((x_1, x_2)) + \mu(\{x_1\}) + \mu(\{x_2\}) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-iux_1} - e^{-iux_2}}{it} \varphi(u) du.$$

Teorema B.3 Si μ es una medida finita con transformada de Fourier $\mathcal{F}\mu$ integrable, entonces $F(x) := \mu((-\infty, x))$, para $x \in \mathbb{R}$, es continuamente diferenciable y además

$$F'(x) = \frac{1}{2\pi} \int e^{-ixu} \mathcal{F}\mu(u) dt.$$

Teorema B.4 Sea μ una medida finita con transformada de Fourier $\mathcal{F}\mu$. Para cada $x \in \mathbb{R}$ se tiene que

$$\mu(\{x\}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-iux} \mathcal{F}\mu(u) du.$$

El siguiente resultado se encuentra en ([5], Teorema 7.6.2). Da condiciones suficientes de cuándo una función $f : \mathbb{R} \rightarrow \mathbb{C}$, puede expresarse como $f(t) = e^{g(t)}$, donde g es única.

Teorema B.5 Sea $f : \mathbb{R} \rightarrow \mathbb{C}$ tal que $f(0) = 0$ y para algún $T > 0$, f es continua en $[-T, T]$ y no se anula en ese intervalo. Entonces existe una única función $g : [-T, T] \rightarrow \mathbb{C}$ con $g(0) = 0$ que es continua y cumple que

$$f(t) = e^{g(t)}, \quad -T \leq t \leq T.$$

El resultado continúa siendo válido cuando $[-T, T]$ es reemplazado por $(-\infty, \infty)$.

Remark 1 Las condiciones

$$\int (1 \wedge x^2) \nu(dx) < \infty \text{ y } \int \frac{x^2}{1+x^2} \nu(dx) < \infty$$

son equivalentes, pues

$$\frac{x^2}{1+x^2} \leq x^2 \wedge 1 \leq 2 \frac{x^2}{1+x^2}.$$

La primera desigualdad es clara, mientras que la segunda se cumple porque si $|x| < 1$, entonces $x^2 = 2\frac{x^2}{2} \leq 2\frac{x^2}{1+x^2}$. Por otro lado, la función $\frac{x^2}{1+x^2}$ es creciente en $[1, \infty)$ y en uno vale $1/2$, por lo que $1 = 2\frac{1}{2} \leq 2\frac{x^2}{1+x^2}$, para $|x| \geq 1$.

Para la función característica empírica se usa la notación

$$\widehat{\varphi}_n(u) = \frac{1}{n} \sum_{t=1}^n e^{iu(X_t - X_{t-1})}, \quad u \in \mathbb{R}.$$

B.2 Base de ondículas de Haar

Aquí se describe brevemente la base de ondículas de Haar, para entender con mayor facilidad la parte de implementación que se hace en [26]. Las definiciones de esta parte han sido extraídas de [30].

Recuerde que si la familia de funciones $\{f_n\}_n$ es una base ortonormal en $L^2(\mathbb{R})$, entonces toda $f \in L^2(\mathbb{R})$ puede escribirse como

$$f = \sum_n a_n f_n,$$

donde $a_n := \langle f_n, f \rangle = \int f_n(x) f(x) dx$.

Definición B.1 Una **ondícula** es una función $\phi \in L^2(\mathbb{R})$ tal que la familia de funciones

$$\phi_{j,k}(t) := 2^{j/2} \phi(2^j t - k),$$

$k, j \in \mathbb{Z}, t \in \mathbb{R}$, es una base ortonormal en $L^2(\mathbb{R})$.

Observe que el soporte de la función $\phi_{j,k}$ es $\text{sop}(\phi_{j,k}) = [2^{-j}k, 2^{-j}(k+1))$. De esta manera, fijando j , se tiene que $\phi_{j,k}$ son funciones con soporte de longitud 2^{-j} y que son ajenos entre sí.

Definición B.2 La **base de ondículas de Haar** se define como

$$\phi(t) := \begin{cases} 1 & \text{si } t \in [0, 1/2), \\ -1 & \text{si } t \in [1/2, 1), \\ 0 & \text{en otro caso.} \end{cases}$$

No es difícil ver que, efectivamente, ϕ es una ondícula.

Notación

1. \mathbb{R} es el conjunto de números reales.
2. \mathbb{R}_+ es el conjunto de los números no negativos.
3. $\stackrel{\mathcal{L}}{=}$ significa “igualdad en distribución”.
4. $\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}$.
5. $a \wedge b = \text{mín} \{a, b\}$.
6. c.s. significa “casi seguramente”.
7. $\xrightarrow{\mathcal{L}}$ representa la convergencia en distribución.
8. La función $g(n)$ es $o(f(n))$ significa $g(n)/f(n) \xrightarrow{n \rightarrow \infty} 0$.
9. La función $g(n)$ es $O(f(n))$ significa que $\limsup_{n \rightarrow \infty} |g(n)/f(n)| < \infty$.
10. El proceso estocástico Y_n es $O_{\mathbb{P}}(\alpha_n)$ si para todo $\varepsilon > 0$, existe $K > 0$ tal que para toda $n \geq 1$ se cumple
$$\mathbb{P}(|Y_n| > K\alpha_n) < \varepsilon.$$
11. f', f'', \dots son las derivadas primera, segunda, etc., de la función f .
12. $f^{(k)}$ es la k -ésima derivada de la función f .
13. $\delta_y(x) = \begin{cases} 1 & \text{si } x = y \\ 0 & \text{si } x \neq y \end{cases}$.
14. \bar{z} es el conjugado del número complejo z , es decir, si $z = x + iy$, entonces $\bar{z} = x - iy$.
15. C^s es la clase de funciones reales con s -ésima derivada continua.
16. $\mathcal{M}(\mathbb{R})$ es la clase de medidas finitas en \mathbb{R} .

17. $\mathcal{B}(\mathbb{R})$ son los conjuntos de Borel en \mathbb{R} .
18. $\mathcal{F}f$ es la transformada de Fourier de la función f .
19. $\mathcal{F}\nu$ es la transformada de Fourier de la medida ν .

Bibliografía

- [1] Akritas, M. (1982), Asymptotic theory for estimating the parameters of a Lévy process, *Annals of the Institute of Statistical Mathematics*, **Vol. 34**, págs. 259–280.
- [2] Barndorff-Nielsen, O.E., Mikosch, T. y Resnick, S. (eds.) (2001), *Lévy Processes: Theory and Applications*, Birkhäuser, Boston.
- [3] Belomestny, D. y M. Reiss (2006), Spectral calibration of exponential Lévy models, *Finance and Stochastics*, **Vol. 10**, págs. 449–474.
- [4] Bickel P. y Krieger A. (1988), Confidence Bands for a Distribution Function Using the Bootstrap, *Journal of American Statistical Association*, **Vol. 84**, núm. 405, págs. 95–100.
- [5] Chung, K. (1974), *A Course in Probability Theory*, New York: Academic Press.
- [6] Comte, F. y V. Genon-Catalot (2010), Nonparametric adaptive estimation for pure jump Lévy processes, *Annales de l' Institut Henri Poincaré Probability and Statistics*, **Vol. 46**, núm. 3, págs. 595-617.
- [7] Comte, F. y V. Genon-Catalot (2009), Nonparametric estimation for pure jump Lévy processes based on high frequency data, *Stochastic Processes and their Applications*, **Vol. 12**, págs. 4088–4123.
- [8] Cont, R. y Tankov, P. (2004), *Financial Modelling with Jump Processes*, Boca Raton, Florida, USA: Chapman & Hall/CRC.

- [9] DasGupta, A.(2008), *Asymptotic Theory of Statistics and Probability*, 1ra. ed. Springer Text in Statistics.
- [10] Davison, A. y Hinkley, D. (1997) *Bootstrap Methods and their Application*. Cambridge University Press.
- [11] Feuerverger, A. y Mureika, R. (1977), The empirical characteristic function and its applications, *The Annals of Statistics*, **Vol. 55**, págs. 88-97.
- [12] Feuerverger, A. y McDunnough, P. (1981a), On the efficiency of empirical characteristic function procedures, *Journal of the Royal Statistical Society, Serie B* 43, págs. 20-27.
- [13] Feuerverger, A. y McDunnough, P. (1981b), On some Fourier methods for inference, *Journal of the American Statistical Association*, **Vol. 76**, núm. 374, págs. 279-387.
- [14] Figueroa-López, J. E. (2008), Small-time moment asymptotics for Lévy processes, *Statistics & Probability Letters*, **Vol. 78**, págs. 3355–3365.
- [15] Figueroa-López, J. (2009), Nonparametric estimation for Lévy models based on discrete sampling, *IMS Lecture Notes of the 3rd E.L. Lehmann Symposium*, **Vol. 57**, págs. 117–146.
- [16] Figueroa-López, J. E. y Houdré, C. (2006), Risk bounds for the non-parametric estimation of Lévy processes. *IMS Lecture Notes, High Dimensional Probability*, **Vol. 51**, págs. 96-116.
- [17] Goldstein, L. and K. Messer (1992), Optimal plug-in estimators for nonparametric functional estimation, *The Annals of Statistics*, **Vol. 20**, págs. 1306–1328.
- [18] Grenander, U. (1981). *Abstract Inference*. Wiley, New York.
- [19] Gugushvili, S. (2009a), Nonparametric estimation of the characteristic triplet of a discretely observed Lévy process, *Journal of Nonparametric Statistics*, **Vol. 21**, págs. 321–343.

- [20] Gugushvili, S. (2009b), Nonparametric inference for discretely sampled Lévy processes, arXiv:0908.3121v2.
- [21] Hall, P. (1993), On Edgeworth expansion and bootstrap confidence bands in nonparametric curve estimation, *Journal of the Royal Statistical Society*, **Vol. 55**, núm. 1, págs. 291-304.
- [22] Huber, P.(1981), *Robust Statistics*. Wiley, New York.
- [23] Jongbloed, G., Van Der Meulen, F. H. y Van Der Vaart, A.W. (2005), Nonparametric inference for Lévy-driven Ornstein–Uhlenbeck processes, *Bernoulli*, **Vol. 11**, págs. 759–791.
- [24] Kappus, J. y Reiss, M. (2010), Estimation of the characteristics of a Lévy process observed at arbitrary frequency, *Statistica Neerlandica*, **Vol. 64**, núm. 3, págs 314-328.
- [25] Kyprianou, A. (2006), *Introductory Lectures on fluctuations of Lévy process with applications*, Universitext, Springer.
- [26] Neumann, M. y Reiss, M. (2009), Nonparametric estimation for Lévy Processes from low frequency observations, *Bernoulli*, **Vol. 15**, págs. 223-248.
- [27] Rao, B. (1987), *Asymptotic Theory of Statistical Inference*, John Wiley and Sons, New York.
- [28] Sato, K. (1999), *Lévy Processes and Infinitely Divisible Distributions*, Cambridge: Cambridge University Press.
- [29] Watteel, R. N. y Kulperger, R. J. (2003), Nonparametric Estimation of the Canonical Measure for Infinitely Divisible Distributions, *Journal of Statistical Computation and Simulation*, **Vol.73**, No. 7, págs. 525–542.
- [30] Wojtaszczyk, P. (1997), *A Mathematical Introduction to Wavelets*, Cambridge Univ. Press.