PUBLICACIONES MATEMATICAS DEL URUGUAY

VOLUMEN 11

VOLUMEN 11

PUBLICACIONES MATEMÁTICAS DEL URUGUAY
Editadas por el Centro de Matemática de la Facultad de Ciencias, Universidad de la República y el Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" Facultad de Ingeniería, Universidad de la República; con el apoyo del Programa de Desarrollo de las Ciencias Básicas.

Montevideo, Abril de 2006.

ISSN
0797-1443

PUBLICACIONES MATEMÁTICAS DEL URUGUAY

Publicaciones Matemáticas del Uruguay(PMU) fue fundada in 1987 y publicada inicialmente por el Centro de Matemática de la Universidad de la República y auspiciada por el Programa de desarrollo de las ciencias básicas (PEDECIBA). Actualmente se publica en conjunto por el Centro de Matemática de la Facultad de Ciencias y el Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" de la Facultad de Ingeniería –ambas de la Universidad de la República– y cuenta con el auspicio del PEDECIBA.

Su objetivo es el de publicar artículos de investigación en todas las áreas de la matemática. Serán considerados para su publicación, artículos originales, adelantos de resultados de investigación, así como resúmenes panorámicos en áreas importantes de la matemática. Todos los artículos serán cuidadosamente arbitrados.

Las Publicaciones Matemáticas del Uruguay aparecerán con una frecuencia semestral.

Presentación de manuscritos

El Consejo Editor de las PMU recomienda que los trabajos presentados lo sean en formato LaTeX; sin embargo trabajos presentados en cualquiera de los otros formatos TeX standardt serán también considerados. Los archivos electrónicos deben ser enviados a la siguiente dirección: `pmu@cmat.edu.uy`.

Si el autor prefiere enviar una versión del manuscrito en papel, se requieren tres copias. Si el artículo es aceptado, el autor deberá enviar la versión electrónica con las características mencionadas.

La dirección de recepción es la siguiente: Publicaciones Matemáticas del Uruguay, Centro de Matemática, Facultad de Ciencias, IGUA 4225, CEP 11400, Montevideo, Uruguay. Al enviar un artículo el autor está asegurando que los derechos autorales no han sido cedidos y que el artículo no ha sido publicado o enviado a publicación en otra revista. Las Publicaciones Matemáticas del Uruguay –ISSN 0797 1443– están indexadas y resumidas en Mathematical Reviews Database y Zentralblatt MATH.

Intercambios bibliográficos

Propuestas de intercambio de suscripciones de las PMU por cualquier tipo de material bibliográfico son bienvenidas y deben ser enviadas a las direcciones mencionadas.

Publicaciones Matemáticas del Uruguay was founded in 1987 and was intially published by "Centro de Matemática de la Universidad de la República" and sponsored by " Programa de desarrollo de las ciencias básicas" (PEDECIBA). Currently is published jointly by "Centro de Matemática de la Facultad de Ciencias de la Universidad de la República" and "Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" de la Facultad de Ingeniería" –both from "Universidad de la República"– with the sponsorship of PEDECIBA.

It endeavors to publish research articles in all areas of mathematics. Survey articles of significant areas of mathematics as well as research announcements will also be considered for publication. All articles will be carefully reviewed.

It will appear with a frequency of two issues a year.

Submission of manuscripts

Publicaciones Matemáticas del Uruguay recommends for reasons of expediency the submission of papers in LaTeX; however papers submitted in any of the other standardt TeX formats will also be considered. Electronic files should be sent to the following address: `pmu@cmat.edu.uy`.

In the case an author prefers to send hard copies, three are required and if the submitted manuscript is eventually accepted, the author will be asked to send an electronic version as described above.

The receiving address is the following: Publicaciones Matemáticas del Uruguay, Centro de Matemática, Facultad de Ciencias, IGUA 4225, CEP 11400, Montevideo, Uruguay.

The submission of a paper implies the author's assurance that it has not been copyrighted, published or submitted for publication elsewhere.

Publicaciones Matemáticas del Uruguay –ISSN 0797 1443– is indexed/abstracted in Mathematical Reviews Database and Zentralblatt MATH.

Library exchanges

Proposals for exchange of library materials in return for subscriptions to the Publicaciones Matemáticas del Uruguay are most welcome. The proposals should be sent to the above address.

Índice

Prefacio

Los artículos que aparecen en este volumen son contribuciones realizadas por participantes del IX Congreso Latinoamericano de Probabilidad y Estadística Matemática (CLAPEM), realizado en Punta del Este, Uruguay, del 22 al 26 de Marzo del 2004. Realizado por segunda vez en el país, contó con 230 participantes de 25 países. El próximo CLAPEM tendrá lugar en Perú, en Marzo de 2007.

Foreword

The articles appearing in this volume are contributions by participants of the IX Latin American Congress on Probability and Mathematical Statistics held in Punta del Este, Uruguay, 22-26 March 2004. Next CLAPEM will take place in Perú, in March 2007.

Reseña de la Ley del Logaritmo Iterado para Martingalas Auto-Normalizadas

Victor H. de la Peña, M. J. Klass y Tze-Leung Lai

ABSTRACT

In this article we present a review of the law of the iterated logarithm for self normalized processes. In the case of continuous martingales the normalization is a function of the quadratic variation. Starting with Kolmogorov's (non-normalized law) we consider results by Stout (1970; 1973) and finally regard the time martingales.We include also –with the corresponding proof– the upper bound of the law of the iterated logarithm for continuous martingales. The proof is an example of the method of pseudo maximization by integration of the density that is derived from the method of mixtures. This method provides very strong tools in order to attack the highly non linear problems originated y the development of a general theory of self normalization. It also has important consequences in the study of crossing problems.

RESUMEN

En este artículo presentamos una reseña de la ley del logaritmo iterado para procesos auto-normalizados. En el caso de martingalas continuas, la normalizacíon es una funcion de la variacíon cuadrática. Comenzamos citando la ley (no-normalizada) de Kolmogorov pasando por resultados de Stout (1970; 1973) llegando hasta el caso de martingalas en tiempo discreto. Además incluimos (con prueba) la cota superior de la ley del logaritmo iterado para martingalas continuas. La prueba provee un ejemplo del método de seudo-maximización por integración de densidades que se deriva del método de mezclas. Este método provee herramientas muy potentes para atacar los problemas (altamente no lineales) que surgen en el desarrollo de la teoría general de auto-normalización y además tiene implicaciones importantes en el estudio de problemas de cruce de fronteras.

Introducción

Los procesos autonormalizados surgen naturalmente en la práctica estadística. En forma estandar (como cuando estan ligados a el Teorema del Límite Central) no tienen unidades y frecuentemente permiten debilitar o aún eliminar supuestos de momentos. El ejemplo típico de auto-normalización es la estadística t de Student. Esta estadística reemplaza la desviación estandard de la población por la desviación estandard de la muestra.

Desde que William Gosset introdujo la estadística Student-t ha habido mucha actividad en el area de auto-normalización. Stout (1970; 1973) introdujo una ley del logaritmo iterado para auto-normalización de super martingalas, en donde la auto-normalización era hecha con el uso de la varianza condicional. Este resultado extendió directamente la ley del logaritmo iterado de Kolmogorov a el caso de martingalas. Desgraciadamente como la varianza condicional es una constante en el caso de variables independientes el resultado no representó un avance a el caso de variables independientes. Mas aun, el problema con variables (no-simétricas) de varianza infinita continuo abierto por muchos años. En los noventas hubo un resurgimiento en el interés en el área dado en parte por el trabajo de Griffin y Kuelbs (1989; 1991) que provee una versión de la ley del logaritmo iterado para variables independientes e idénticamente distribuidas. En fechas recientes ha aumentado el interés en teoremas del límite central y cotas a los momentos de sumas autonormalizadas de variables independientes e idénticamente distribuidas X_i. En particular, Giné, Götze y Mason (1997), muestran que la estadística t de Student tiene una distribución límite normal si y solo si X_1 se encuentra en el dominio de atracción de la distribución normal. La prueba usa cotas exponenciales y de momentos L_p para las sumas autonormalizadas $U_n = A_n/B_n$, donde $A_n = \sum_{i=1}^n X_i$ y $B_n^2 = \sum_{i=1}^n X_i^2$. Este resultado de Giné et al. provee más evidencia de el efecto robustificante que la auto-normalización tiene sobre resultados límites. En este caso el teorema del limite central no se puede aplicar si la varianza es infinita aunque las variables pueden estar en el dominio de atracción de la ley normal. En Fan, Tam, Vande Woude y Ren (2004) esta propiedad robustificante es usada en su trabajo sobre el análisis de microarreglos. En ese artículo proponen una estadística t ponderada como un método para robustificar que permite tratar los efectos heterosedásticos presentes en comparaciones simultánas múltiples. Estas comparaciones múltiples son hechas a un gran número (en los miles) de genes. En su artículo proveen una extensión de la estadística t para protegerse contra el efecto altamente negativo que la especificación errónea de el modelo tendría sobre miles de pruebas de hipótesis. Una fuente importante sobre resultados de auto-normalización en variables independientes es el articulo de Shao (1997).

1. ¿ Qué tan grande puede ser una suma ?

Para poner nuestro estudio en perspectiva a continuación presentamos la ley del logaritmo iterado de Kolmogorov (Chow y Teicher (1988)).

Teorema 1.1. (Kolmogorov). Sea $\{X_i\}$ una sucesión de variables independientes con $EX_i = 0$. Denote $S_n = \sum_{i=1}^n X_i$ y $s_n^2 = \sum_{i=1}^n EX_i^2$. Supongamos que $s_n^2 \to \infty$ para $n \to \infty$ y que

$$|X_n| \leq \epsilon_n s_n / \sqrt{\log\log(s_n^2 \vee e^e)},$$

para $\epsilon_n \to 0$ cuando $n \to \infty$. Entonces la convergencia

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2s_n^2 \log\log s_n^2}} = 1,$$

se cumple con probabilidad 1.

•

Observe que este resultado es válido cuando las variables están acotadas por funciones que dependen de la varianza de las variables. Este resultado nos da una idea de que tan grande puede ser la suma cuando su tamaño se mide con respecto a una función específica de s_n. Vease Marcinkiewicz y Zygmund (1937) en relación a estos comentarios. El estudio de el caso de variables independientes no acotadas fue iniciado por Hartman y Wintner (1941) bajo condiciones de dominacion de momentos que se cumplen en el caso de variables $\{X_i\}$, idénticamente distribuidas para las cuales $EX_1 = 0$ y $EX^2 < \infty$. (vease tambien de Acosta (1983)).

El objetivo principal de este artículo es presenar un resultado análogo de auto-normalización natural para martingalas en tiempo discreto y para martingalas continuas.

Una extensión del resultado de Kolmogorov a el caso de martingalas en tiempo discreto se debe a de Stout (1970; 1973) a el caso de martingalas autonormalizadas por la variación cuadrática (Vease tambien Einmahl y Mason (1989)). Esta normalizacion se reduce a la normalización no aleatoria en el caso de variables independientes y por lo tanto el resultado extiende directamente la ley de Kolmogorov pero no provee auto-normalización.

Teorema 1.2 Sea $\{d_n, \mathcal{F}_n\}$, $n = 1, ...$ una sucesión adaptada con $E(d_n|\mathcal{F}_{n-1}) \leq 0$. Denotemos $M_n = \sum_{i=1}^n d_i$ y la varianza condicional $\sigma_n^2 = \sum_{i=1}^n E(d_i^2|\mathcal{F}_{i-1})$. Supongamos que

A: $d_n \leq m_n$ para algunas $\mathcal{F}_{n-1}$-medible $m_n \geq 0$.
B: $\sigma_n^2 < \infty$ con probabilidad uno para cada n.
C: $\lim_{n\to\infty} \sigma_n^2 = \infty$ con probabilidad uno.
D: $\limsup_{n\to\infty} \frac{m_n\sqrt{\log\log(\sigma_n^2)}}{\sigma_n} = 0$ con probabilidad uno.
Entonces,

$$\limsup_{n\to\infty} \frac{M_n}{\sqrt{2\sigma_n^2 \log\log \sigma_n}} = 1,$$

con probabilidad uno.

El resultado de Kolmogorov (Teorema 1.1) fue extendido por Marcinkiewicz a el caso de auto-normalización en el caso de variables independientes simétricas (no necesariámente con la misma distribución) probando en la notación del Teorema 1.1.

Teorema 1.3 (Marcinkiewicz).

$$\lim\sup_{n\to\infty} \frac{S_n}{\sqrt{2V_n^2 \log\log V_n^2}} \leq 1,$$

cuando $V_n^2 \to \infty$ donde $V_n^2 = \sum_{i=1}^n X_i^2$.

Este resultado fue extendido por Griffin y Kuelbs (1989; 1991).

A continuación presentamos el resultado análogo al de Kolmogorov para la auto-normalización natural en el caso de martingalas en tiempo discreto y finalmente con el resultado especial para el caso de martingalas en tiempo continuo.

Teorema 1.4. (de la Peña, Klass y Lai (2004)). Sea $\{M_n, \mathcal{F}_n\}$, $n = 1, ...$ una martingala con $\mathcal{F}_n$ una sucesión creciente de σ-algebras y diferencia de martingalas $\{d_i\}$. Denótese $M_n = \sum_{i=1}^n d_i$ y $V_n^2 = \sum_{i=1}^n d_i^2$, $v_n = \frac{V_n}{\log\log(V_n \vee e^e)}$. Supongamos que $v_n \to \infty$, $|d_n| \leq m_n$ con probabilidad uno para algunas variables aleatorias m_n-$\mathcal{F}_{n-1}$-medibles. Con $\{m_n\}$ una sucesión creciente $m_n \to \infty$. Además, supongamos que $V_n \to \infty$ y $\frac{m_n}{V_n} \to 0$ en forma casi segura. Entonces,

$$\lim\sup_{n\to\infty} \frac{M_n}{\sqrt{2V_n^2 \log\log V_n^2}} = 1,$$

en forma casi segura.

La prueba de este resultado requiere del desarrollo de nuevas martingalas exponenciales y un argumento delicado de tiempos de parada para "acotar" a la suma de cuadrados usando una sucesión de constantes crecientes.

Para martingalas en tiempo continuo el resultado análogo es presentado a continuación.

Teorema 1.5 Sea $\{M_t, \mathcal{F}_t, \quad t \geq 0\}$ una martingala continua con variacion cuadrática $< M >_t < \infty$ para toda t finita. Entonces,

$$\lim\sup_{t\to\infty} \frac{M_t}{\sqrt{2 < M >_t \log\log < M >_t}} = 1,$$

con probabilidad uno en el conjunto en el que $< M >_t \rightarrow \infty$ en forma casi segura.

La prueba de la cota superior de este resultado es una aplicación interesante del método de seudo-maximización por integración de densidades. En la siguiente sección presentamos este metodo y la prueba. El método se basa en integrar una martingala exponencial con respecto a una densidad apropiada (dependiendo del problema). En lo que sigue utilizaremos la siguiente condicion de integrabilidad.

Condición de Integrabilidad (CI).

Para un par arbitrario de variables aleatorias A, B, con $B > 0$ diremos que satisfacen la condición CI si

$$E \exp\{\lambda A - \lambda^2 B^2/2\} \leq 1, \tag{1,1}$$

para toda $\lambda, \quad -\infty < \lambda < \infty$.

2. Seudo-Maximización: Método de Mezclas

En esta sección estamos interesados en desarrollar desigualdades exponenciales y la ley del logaritmo iterado en base únicamente de la condición CI introducida en (1.1).

Notese que si el integrando $\exp\{\lambda A - \lambda^2 B^2/2\}$ pudiese ser maximizado para λ dentro del valor esperado (como puede ser hecho si el cociente A/B^2 no es aleatorio), entonces tomando $\lambda = A/B^2$ uno obtendria $E \exp\{\frac{A^2}{2B^2}\} \leq 1$. Esto a su vez nos daria la desigualdad optima del tipo Chebyshev $P(\frac{A}{B} > x) \leq \exp\{\frac{-x^2}{2}\}$. Dado que A/B^2 en general no puede ser tomada como no aleatoria necesitamos encontrar un método alterno para manejar esta maximización.

Una forma de obtener un efecto similar consiste en integrar con respecto a una medida de probabilidad F, y usando el teorema de Fubini intercambiar el orden de integración con respecto a P y F. Para que esto sea efectivo para todos los pares posibles de (A, B), la F adecuada necesitará ser tan uniforme como sea posible para que asi "tome" el valor máximo de $T(\lambda) = \exp\{\lambda A - \lambda^2 B^2/2\}$ independientemente de donde ocurra este valor. Por lo tanto alguna masa será asignada en forma segura cerca del valor $\lambda = A/B^2$ que maximiza $\lambda A - \lambda^2 B^2/2$. Como todas las medidas uniformes son multiplos de la medida de Lebesgue (que es infinita), construimos una medida finita (o una sucesión de medidas) que decrece a cero tan lentamente como es posible.

Este método permite obtener desigualdades exponenciales y desigualdades de momentos para A/B, $\frac{A}{\sqrt{B^2+(EB)^2}}$, $\frac{A}{B\sqrt{\log\log(B\vee e^2)}}$ así como leyes del logaritmo iterado (ver de la Peña, Klass and Lai (2000; 2004)).

El método de mezclas que re-bautizamos como seudo-maximización en base a la discusión en el inicio de esta sección tiene sus orígenes en el trabajo de Robbins y Siegmund (1970).

Teorema 2.1.1. (de la Peña, Klass and Lai (2004)) Sean A, B con $B > 0$ variables aleatorias que satisfacen la condición de integrabilidad (1.1) para toda $\lambda \in \mathbf{R}$. Entonces

$$P(\frac{|A|}{\sqrt{B^2+(EB)^2}} > x) \leq \sqrt{2}\exp(-x^2/4) \qquad (2,1)$$

para toda $x > 0$.

La prueba de este resultado se basa en el Lema 2.1.2 que se usa para "seudo-maximizar". La prueba de este resultado se obtiene multiplicando las dos cotas de (1.1) por la densidad $(2\pi)^{-1/2}EB\exp(-\lambda^2(EB)^2/2)$ (con $EB > 0$) e integrando con respecto a λ. Finalmente se usa la desigualdad de Hölder.

Lema 2.1.2. Sean A, B con $B > 0$ variables aleatorias que satisfacen la condicion de integrabilidad CI para toda $\lambda \in \mathbf{R}$. Entonces,

$$(EB)E\frac{1}{\sqrt{B^2+(EB)^2}}\exp\{\frac{A^2}{2(B^2+(EB)^2)}\} \leq 1.$$

2.2 Interpretación Intuitiva

Cuando $EB \to \infty$ estamos (aproximadamente) colocando una masa puntual alrededor de cero y como para toda $\epsilon > 0$, $P(A/B^2 > \epsilon, B^2 > z) \leq \exp\{-\epsilon^2 z^2/2\} \to 0$ cuando $z \to \infty$, el valor máximo de $T(\lambda, A, B)$ tambien se localiza alrededor de $\lambda = 0$ con gran probabilidad. Un argumento similar nos da una interpretación intuitiva en otras circunstancias.

En la siguiente sección presentamos la prueba de la cota superior para la ley del logaritmo iterado para martingalas continuas. La prueba se basa en el uso de la seudo-maximización y el cruce de fronteras.

3. Cruce de Fronteras y la Ley del Logaritmo Iterado.

A continuación estudiamos (en forma general) el cruce de fronteras utilizando el método de mezclas. Iniciamos presentando las llamadas fronteras de "Robbins-Siegmund" (R-S) que son discutidas en forma extensiva en Lai (1976). Sea F una medida finita y positiva sobre $(0, \lambda_0)$ y supongamos que $F(0, \lambda_0) > 0$. Sea $\Psi(u, v^2) = \int \exp(\lambda u - \lambda^2 v^2/2\} dF(\lambda)$.

Si tomamos $u = M_t$ y $v^2 =< M >_t$ la variable aleatoria $\exp\{\lambda M_t - \lambda^2 < M >_t /2\}$ es una martingala que satisface la condicion de (1.1) (con igualdad). Integrando con respecto a $dF(\lambda)$ obtenemos la super-martingala $\Psi(M_t, < M >_t)$. La siguiente sección muestra como usar esta super-martingala para estudiar cruce de fronteras.

3.1. Cruce de Fronteras

Para $c > 0$ y $v^2 > 0$, la ecuacion

$$\Psi(u, v^2) = c,$$

tiene una solución única

$$u = \beta_F(v^2, c).$$

Además, $\beta_F(v^2, c)$ es una función cóncava de v^2 y

$$\lim_{v^2 \to \infty} \frac{\beta_F(v^2, c)}{v^2} = b/2,$$

donde

$$b = \{\sup y > 0 : \int_0^b dF(\lambda) = 0\},$$

con sup sobre el conjunto vacio igual a infinito.

Para poder usar las fronteras R-S $\beta_F(v^2, c)$ en el cruce de fronteras procedemos de la siguiente forma.

Consideremos el problema de estimar la probabilidad

$$P(A_t \geq g(B_t) \text{ para algún } t \geq 0).$$

Si g es una frontera R-S, entonces $g(B_t) = \beta_F(B_t^2, c)$ para alguna medida F y $c > 0$. Esta probabilidad es igual a

$$P(A_t \geq \beta_F(B_t^2, c) \text{ para algún } 0 \leq t \leq t_0\ ) =$$

$$P(\Psi(A_t, B_t^2) \geq c \text{ para algún } 0 \leq t \leq t_0\ ) \leq$$

$$E\Psi(A_{t_0}, B_{t_0}^2)/c \leq F(0, \lambda_0)/c,$$

usando la desigualdad de Doob con la super-martingala $\Psi(A_t, B_t), 0 \leq t \leq t_0$.

3.2. La Ley del Logaritmo Iterado Para Martingalas Auto-normalizadas

A continuación presentamos una aplicación de la seudo-maximización en el caso de martingalas continuas. En este caso tomamos

$$dF_\delta(\lambda) = \frac{1}{\lambda(\log(1/\lambda))(\log\log(1/\lambda))^{1+\delta}},$$

para cada $\delta > 0$ y $0 < \lambda < e^{-e}$.

Denótese $\log_2(x) = \log\log(x)$ y $\log_3(x) = \log\log_2(x)$. Como se muestra en el Ejemplo 4 de Robbins y Siegmund (1970), en este caso

$$\beta_{F_\delta}(v^2, c) =$$

$$\sqrt{2v^2[\log_2 v^2 + (3/2+\delta)\log_3 v^2 + \log(c/2\sqrt{\pi}) + o(1)]},$$

cuando $v^2 \to \infty$.

Usando lo anterior, tenemos que la probabilidad de interés está acotada por
$F_\delta(0, e^{-e})/c$ para cada $c > 0$. Para $\epsilon > 0$ dado, tómese δ lo suficientemente pequeña y $c = c(\delta)$ lo suficientemente grande para que $F_\delta(0, e^{-e})/c(\delta) < \epsilon$. Como ϵ puede ser arbitrariamente pequeño y para cada $c = c(\delta)$ fija, $\beta_{F_\delta}(v^2, c) \to \sqrt{2v^2 \log\log v}$ conforme $v^2 \to \infty$,

$$\lim\sup_{t\to\infty} \frac{M_t}{\sqrt{2 < M >_t \log\log < M >_t}} \leq 1,$$

en el conjunto en que $< M >_t \to \infty$.

La explicación intuitiva de este resultado es que cuando $\delta \to 0$ tenemos una cantidad creciente de masa de F_0 concentrada cerca $\lambda = 0$ como estamos integrando sobre $dF_\delta(\lambda)$, que fue escogida porque decrece muy lentamente. Mas aun, para $\delta = 0$, esta función no es integrable. Integrando con respecto a esta densidad y después dejando $\delta \to 0$ nos permite el recuperar el crecimiento máximo (casi seguro) de A_t, medido en función de B_t que en este caso resulta ser $\sqrt{2 < M >_t \log\log < M >_t}$.

Para probar la desigualdad inversa se requiere el utilizar un método delicado de cotas.

En el apéndice presentamos una serie de martingalas exponenciales que satisfacen la condición de integrabilidad CI necesaria para el uso de el método de mezclas que presentamos en el artículo. Para aplicar estos resultados, es importante el tener en cuenta los valores de λ para los cuales CI es satisfecha. La lista es muy variada e incluye martingalas en tiempo discreto y sumas de variables condicionalmente simétricas sin condiciones de momentos.

3.3. Ejemplos: Martingalas y Super-Martingalas

En el Apéndice proveemos varios ejemplos de variables aleatorias que satisfacen la suposición canónica. Los Lemas A.1 y A.2 son considerados resultados clásicos en la teoría de martingalas. El Lema A.3 se refiere al caso de martingalas con brincos bajo una condición local de integrabilidad cuadrática. El Lema A.4 trata el caso de variables aleatorias condicionalmente simétricas sin ninguna condición de dependencia o en los momentos. El Lema A.5, introducido por Stout (1973), permite normalización por la raíz cuadrada de la suma de varianzas condicionales. El Lema A.6 es la super-martingala exponencial asociada a la desigualdad de Bernstein.

Apéndice

Lema A. 1 Sea W_t un proceso Browniano estandar. Supóngase que T es un tiempo de paro, dado que $T < \infty$ a.s. Entonces

$$E\exp\{\lambda W_T - \lambda^2 T/2\} \leq 1,$$

para toda λ, $-\infty < \lambda < \infty$.

Lema A. 2 Sea M_t una martingala, continua, cuadráticamente integrable, con $M_0 = 0$. Entonces

$$E\exp\{\lambda M_T - \lambda^2 < M >_t /2\} \leq 1,$$

para toda $t \geq 0$. Además, $\exp\{\lambda M_t - \lambda^2 < M >_t /2\}$ es una supermartingala para toda $-\infty < \lambda < \infty$. (Si asumimos que M_t es una martingala local y continua, entonces la desigualdad también es válida por el Lema de Fatou.)

Lema A. 3 Sea $\{M_t; t \geq 0\}$ una martingala local, cuadráticamente integrable, y $M_0 = 0$. Sea $\{V_t\}$ un proceso creciente, adaptado, puramente discontinuo y localmente integrable; y sea $V^{(p)}$ su proyección dual predecible. Deje $X_t := M_t + V_t$, $C_t := \sum_{s \leq t}((\Delta X_s)^+)^2$, $D_t := \{\sum_{s \leq t}((\Delta X_s)^-)^2\}_t^{(p)}$, $H_t := \langle M \rangle_t^c + C_t + D_t$. Entonces $\exp\{X_t - V_t^{(p)} - 1/2 H_t\}$ es una supermartingala por lo tanto

$$E\exp\{\lambda(X_t - V_t^{(p)}) - \lambda^2 H_t/2\} \leq 1$$

para toda $-\infty < \lambda < \infty$.

El Lema A. 3 , aparece como la Proposición 4.2.1 en Barlow, Jacksa y Yor (1986). El siguiente lema se satisface sin ninguna condición de integrabilidad de la variable en cuestión. Es una generalización del hecho de que si X es una variable aleatoria simétrica, entonces $A = X$ y $B = X^2$ satisfacen la condición canónica (1.3.1). Posee una larga historia, incluyendo a Wang (1989) y Hitczenko (1990). En su presente versión se puede encontrar en de la Peña (1999).

Lema A. 4 Sea $\{d_i\}$ una sucesión de variables adaptadas a un sucesión creciente de σ-algebras $\{\mathcal{F}_i\}$. Asúmase que las d_i's son condicionalmente simétricas (esto es, $\mathcal{L}(d_i|\mathcal{F}_{i-1}) = \mathcal{L}(-d_i|\mathcal{F}_{i-1})$. Entonces, $\exp\{\lambda \sum_{i=1}^{n} d_i - \lambda^2 \sum_{i=1}^{n} d_i^2/2\}$, $n \geq 1$ es una supermartingala con media ≤ 1, para toda $-\infty < \lambda < \infty$.

Observe que la sucesión de variables aleatorias X_i puede ser "simetrizada" para producir una supermartingala exponencial, introduciendo variables aleatorias X_i' tales que

$$\mathcal{L}(X_n'|X_1, X_1', \ldots, X_{n-1}, X_{n-1}', X_n) = \mathcal{L}(X_n|X_1, \ldots, X_{n-1})$$

dado que $d_n = X_n - X_n'$; Vea la Sección 6.1 de de la Peña y Giné (1999).

Lema A. 5 (Stout) Sea $\{d_n\}$ una sucesión de variables aleatorias adaptadas a una sucesión creciente de σ-algebras $\{\mathcal{F}_n\}$ tal que $E(d_n|\mathcal{F}_{n-1}) \leq 0$ y $d_n \leq M$ a.s. para toda n y algunas constante positivas M, no-aleatorias. Sea $0 < \lambda \leq M^{-1}$, $A_n = \sum_{i=1}^{n} d_i$, $B_n^2 = (1 + 1/2\lambda_0 M) \sum_{i=1}^{n} E(d_i^2|\mathcal{F}_{i-1})$, $A_0 = B_0 = 0$. Entonces $\{\exp(\lambda A_n - 1/2\lambda^2 B_n^2), \mathcal{F}_n, n \geq 0\}$ es una supermartingala para toda $0 \leq \lambda \leq \lambda_0$.

El Lema A. 6 extiende las desigualdades de Bernstein y Bennett, a el caso de martingalas. Vease de la Peña y Giné (1999) para una reseña historica.

Lema A. 6 Sea $\{d_n\}$ una sucesion de variables aleatorias adaptadas a una sucesión de σ-algebras $\{\mathcal{F}_n\}$ tal que con probabilidad uno, $E(d_n|\mathcal{F}_{n-1}) = 0$ y $\sigma_n^2 = E(d_n^2|\mathcal{F}_{n-1}) < \infty$. Supóngase que existe una constante positiva M tal que $E(|d_n|^k|\mathcal{F}_{n-1}) \leq (k!/2)\sigma_n^2 M^{k-2}$ o $P(|d_n| \leq M|\mathcal{F}_{n-1}) = 1$ con probabilidad uno para toda $n \geq 1$, $k > 2$. Sea $A_n = \sum_{i=1}^{n} d_i$, $V_n^2 = \sum_{i=1}^{n} E(d_i^2|\mathcal{F}_{i-1})$, $A_0 = V_0 = 0$. Entonces $\{\exp(\lambda A_n - \frac{1}{2(1-M\lambda)}\lambda^2 V_n^2), \mathcal{F}_n, n \geq 0\}$ es una supermartingala para toda $0 \leq \lambda \leq 1/M$.

Nota

Este trabajo es un resumen traducido de una reseña mas extensa que someteremos a la revista Probability Surveys.

Agradecimientos

Le agradecemos al árbitro varias sugerencias que mejoraron el texto y a Gerardo Hernández-del Valle por su ayuda en la traducción de parte del texto.

Referencias

[1] Barlow, M. T., Jacka, S. D., and Yor, M. (1986). Inequalities for a pair of processes stopped at a random time. *Proc. Lond. Math. Soc. Ser. III* **53** 142-172.

[2] Chow, Y. S. and Teicher, H. (1988). Probability Theory: Independence, interchangeability, martingales. Springer-Verlag, New York. Second Edition.

[3] de Acosta, A. (1983). A new proof of the Hartman-Wintner law of the iterated logarithm. *Ann. Probab.* **11** (2), 270-276.

[4] de la Peña, V. H. (1999). A general class of exponential inequalities for martingales and ratios. *Ann. Probab.* **27** 537-564.

[5] de la Peña, V. H. and Giné, E. (1999). *Decoupling: From Dependence to Independence.* Springer, New York.

[6] de la Peña, V. H., Klass, M. J. and Lai, T. L. (2000). Moment bounds for self-normalized processes. In *Progress in Probability, Proceedings of Conference on High Dimensional Probability II.* (E. Giné, D. M. Mason and J. A. Wellner, eds.), 3-12. Birkhäuser, Boston.

[7] de la Peña, V. H., Klass, M. J. and Lai, T. L. (2004). Self-normalized processes:exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.* **32** (3A), 1902-1933.

[8] Einmahl, U. and Mason, D. (1989). Some results on the almost sure behavior of martingales. In *Colloquia Math. Soc. Janós Bolyai: Limit*

Theorems in Probability and Statistics (I. Berkes, E. Csaki, P. Revesz, eds.), 185-195. North Holland, Amsterdam.

[9] Fan, J. Tam, P. Vande Woude, G. and Ren, Y. (2004). Normalization and analysis of cDNA microarrays using within-array replications applied to neuroblastoma cell response to a cytokine. *Proc. Natl. Acad. Sci. USA*, **101** (5), 1135-1140.

[10] Giné, E, Goetze, F. and Mason D. (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25** 1514-1531.

[11] Giné, E. and Mason, D.M. (1998). On the LIL for self-normalized sums of i.i.d. random variables. *J. Theoret. Probab.* **11** 351-370.

[12] Griffin, P. and Kuelbs, J. (1989). Self-normalized laws of the iterated logarithm. *Ann. Probab.* **17** 1571-1601.

[13] Griffin, P. and Kuelbs, J. (1991). Some extensions of the LIL via self-ormalizations. *Ann. Probab.* **19** 380-395.

[14] Hartman, P. and Wintner, A. (1941). On the law of the iterated logarithm. *American J. Math.* **63** 169-176.

[15] Hitczenko, P. (1990). Upper bounds for the L_p-norms of martingales. *Probab. Theory Related Fields* **86** 225-238.

[16] Lai, T. L. (1976). Boundary crossing probabilities for sample sums and confidence sequences. *Ann. Probab.* **4** 299-312.

[17] Marcinkiewicz, J. and Zygmund, A. (1937). Remarque sur la loi du logarithme itéré. *Fund. Math.* **29** 215-222.

[18] Robbins, H. and Siegmund, D. (1970). Boundary crossing probabilities for the Wiener process and sample sums. *Ann. Math. Statist.* **41** 1410-1429.

[19] Shao, Q. (1997). Self-normalized large deviations. *Ann. Probab.* **25** 285-328.

[20] Stout, W. F. (1970). A martingale analogue of Kolmogorov's law of the iterated logarithm. *Z. Wahrsch. verw. Geb.* **15** 279-290.

[21] Stout, W. F. (1973). Maximal inequalities and the law of the iterated logarithm. *Ann. Probab.* **1** 322-328.

[22] Wang, G. (1989). Some sharp inequalities for conditionally symmetric martingales. Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

V. H. DE LA PEÑA
vp@stat.columbia.edu
Department of Statistics, Columbia University

M.J. KLASS
klass@stat.berkeley.edu
Departments of Mathematics and Statistics,
University of California, Berkeley

TZE-LEUNG LAI
lait@stat.stanford.edu
Department of Statistics, Stanford University

Rates of Convergence for the BIC Estimates of Markov Chain Order

C.C.Y. Dorea, C.R. Gonçalves and J.S. Lopes[1]

ABSTRACT

The Bayesian Information Criterion (BIC) estimates the order of a Markov chain with finite state space from a sample path $X_1, X_2, \ldots, X_n$. It is known that the BIC criterion provides a class of strongly consistent estimates for the true order. In this note, we derive rates of convergence for the BIC estimates.

RESUMEN

El criterio de información bayesiano (BIC) estima, a partir de un camino muestral $X_1, X_2, \ldots, X_n$ el orden de un proceso de Markov con espacio de estados finito. Es sabido que el criterio BIC produce una clase de estimadores fuertemente consistentes para el orden. En esta nota estudiamos las velocidades de convergencia para los estimadores BIC.

Mathematics Subject Classification (2000): 62F12, 62M05; Secondary 60J10.

Introduction

Let $X = \{X_n\}_{n \geq 1}$ be a multiple Markov chain with finite state space E and unknown but finite order $0 \leq r \leq K < \infty$. For $0 \leq k \leq K$ let $\hat{L}(k)$ be the maximum log-likelihood function based on the sanple $X_1, X_2, \ldots, X_n$ when the chain is assumed to be of order k. The approximation of Kullback-Leibler information measure by Neyman-Pearson statistics along with the asymptotic χ^2-distribution of the maximum log-likelihood ratio $2 \log \dfrac{\hat{L}(k)}{\hat{L}(r)}$ form the basis to derive the Akaike information criterion (AIC, Akaike (1974)),

$$\text{AIC}(k) = -2 \log \hat{L}(k) + 2|E|^k(|E| - 1)$$

where $|E|$ denotes the cardinality. Schwarz (1978) proposed the BIC estimator that replaces in the penalty term 2 by $\log n$ and this corrects the inconsistency of AIC. Define

$$\hat{r}_n = \arg\min_{0 \leq k \leq K} \text{BIC}(k)$$

[1]Partially supported by CNPq, CAPES-PROCAD and FAPDF-PRONEX.

where

$$\mathrm{BIC}(k) = -2\log \hat{L}(k) + |E|^k(|E|-1)\log n, \tag{1}$$

$$\log \hat{L}(k) = \sum_{i_1\ldots i_{k+1}} n_{i_1\cdots i_{k+1}} \log \frac{n_{i_1\cdots i_{k+1}}}{n_{i_1\cdots i_k}} \tag{2}$$

and

$$n_{i_1\cdots i_k} = \sum_{j=1}^{n-k} 1(X_j = i_1, \ldots, X_{j+k-1} = i_k).$$

The term $n_{i_1\cdots i_k}$ represents the transition counts, that is, the number of occurrences of $(i_1, \ldots, i_k)$ in the sample $(X_1, \ldots, X_n)$ and for $k = 0$ we interpret $n_{i_1\cdots i_k} = n$. The term $\log \hat{L}(k)$ is the k-th order maximum log-likelihood and the sum is over $i_1, \ldots, i_{k+1}$ for which $n_{i_1\cdots i_{k+1}} > 0$.

Several authors have addressed the consistency problem for BIC, see, for example, Katz (1981), Finesso (1992) or Barron, Rissanen and Yu (1998). More recently, Csiszar and Shields (2000) established the strong consistency for BIC with no boundness assumption on the order r. Some works on optimal error exponents for the probability of errors can also be found in the literature : Merjav, Gutman and Ziv (1989), Finesso, Liu and Narayan (1996) and Gassiat and Boucheron (2003). In this note, we are concerned with the rate of convergence of BIC estimator $\hat{r}_n$ to the true value r.

For the process X denote the transition probabilities by

$$p_{i_{n-r}\cdots i_{n-1};i_n} = Pr(X_n = i_n \,|\, X_k = i_k,\ k < n), \quad i_j \in E. \tag{3}$$

If $r = 0$ interpret (3) as $\pi_{i_n} = p_{i_{n-1};i_n}$.

For $k \geq r$ let $Y_n^{(k)} = (X_n, \ldots, X_{n+k-1})$. Then $Y^{(k)} = \{Y_n^{(k)}\}_{n\geq 1}$ is a first order Markov chain on E^k with transition probabilities

$$Pr(Y_{n+1}^{(k)} = (j_1, \ldots, j_k) \mid Y_n^{(k)} = (i_1, \ldots, i_k)) = p_{i_{k-r+1}\cdots i_k;j_k} \tag{4}$$

if $(i_2, \ldots, i_k) = (j_1, \ldots, j_{k-1})$.

Theorem 1. Assume that the Markov chain $Y^{(r)}$ is ergodic and that $c_n > 0$ is a sequence satisfying

$$\limsup_{n\to\infty} \frac{c_n}{\log\log n} < \infty.$$

Then

$$\lim_{n\to\infty} P_\nu(\mathrm{BIC}(k) - \mathrm{BIC}(r) \geq c_n\,, k \neq r) = 1, \quad \forall\, \nu,$$

where P_ν stands for the probability with initial distribution ν.

1. Auxiliary Results

We will assume that the derived first order Markov chain $Y^{(r)}$ is ergodic with stationary (equilibrium) distribution $(\pi_{i_1\cdots i_r})$. Note that

$$\begin{aligned}\pi_{i_1\cdots i_r} &= \sum_{j_1\cdots j_r} \pi_{j_1\cdots j_r} p_{j_1\cdots j_r;i_r} \\ &= \sum_{j} \pi_{ji_1\cdots i_{r-1}} p_{ji_1\cdots i_{r-1};i_r}. \end{aligned} \tag{5}$$

For $k \geq r$ define

$$\pi_{i_1\cdots i_k} = \pi_{i_1\cdots i_r} p_{i_1\cdots i_r;i_{r+1}} \cdots p_{i_{k-r}\cdots i_{k-1};i_k} \tag{6}$$

then by (5),

$$\begin{aligned}\pi_{i_1\cdots i_k} &= \sum_{j} \pi_{ji_1\cdots i_{r-1}} p_{ji_1\cdots i_{r-1};i_r} p_{i_1\cdots i_r;i_{r+1}} \cdots p_{i_{k-r}\cdots i_{k-1};i_k} \\ &= \sum_{j} \pi_{ji_1\cdots i_{k-1}} p_{i_{k-r}\cdots i_{k-1};i_k}. \end{aligned}$$

From (4) we conclude that (6) is a stationary distribution for $Y^{(k)}$.

Proposition 1. For $k \geq r$ the first order Markov chain $Y^{(k)}$ is ergodic with stationary distribution $(\pi_{i_1\cdots i_k})$ given by (6). Moreover,

$$\frac{n_{i_1\cdots i_k}}{n} \longrightarrow \pi_{i_1\cdots i_k} \text{ and } \frac{n_{i_1\cdots i_{k+1}}}{n_{i_1\cdots i_k}} \longrightarrow p_{i_{k-r+1}\cdots i_k;i_{k+1}} \quad \text{a.s. } [P_\nu]. \tag{7}$$

From the ergodicity of $Y^{(r)}$ it is easy to verify that $Y^{(k)}$ is also ergodic. The almost sure (a.s.) convergences (7) follows immediately from the Law of Large Numbers for Markov chains (see, for example, Dacunha-Castelle and Duflo (1986)). Next, we gather theorems 4.7 and 4.8 (Hall and Heyde (1980)) adapted to our needs.

Lemma 1. Let $\{Z_n\}$ and $\{W_n\}$ be sequences of random variables and let $\mathcal{F}_n = \sigma(Z_1, Z_2, \ldots, Z_n)$. For $S_n = \sum_{j=1}^{n} Z_j$ assume that $\{S_n, \mathcal{F}_n\}$ is a zero-mean and square-integrable martingale and that W_n is $\mathcal{F}_{n-1}$-measurable and satisfies

$$W_n \overset{a.s.}{\to} \infty \quad \text{and} \quad \frac{W_n}{W_{n+1}} \overset{a.s.}{\to} 1. \tag{8}$$

Moreover, for $T_n = \sqrt{2W_n^2 \log\log W_n^2}$ the following conditions are satisfied

$$\frac{1}{T_n}\sum_{j=1}^{n}[Z_j 1(|Z_j|>1) - E\{Z_j 1(|Z_j|>1) \mid \mathcal{F}_{j-1}\}] \overset{a.s.}{\to} 0, \tag{9}$$

$$\frac{1}{W_n^2}\sum_{j=1}^{n}[E\{Z_j^2 1(|Z_j|\le 1) \mid \mathcal{F}_{j-1}\} - (E\{Z_j 1(|Z_j|\le 1) \mid \mathcal{F}_{j-1}\})^2] \overset{a.s.}{\to} 1 \tag{10}$$

and

$$\sum_{j=1}^{\infty}\frac{E\{Z_j^4 1(|Z_j|\le 1) \mid \mathcal{F}_{j-1}\}}{W_j^4} < \infty \quad a.s. \tag{11}$$

Then

$$\limsup_{n\to\infty}\frac{S_n}{T_n} = 1 \ \ a.s.$$

For fixed $i_1, \ldots, i_{k+1}$ such that $0 < p_{i_{k-r+1}\cdots i_k;i_{k+1}} < 1$ define

$$Z_j = 1(Y_j^{(k)} = (i_1, \ldots, i_k))[1(Y_{j+1}^{(k)} = (i_2, \ldots, i_{k+1})) - p_{i_{k-r+1}\cdots i_k;i_{k+1}}]$$

and

$$W_n^2 = [p_{i_{k-r+1}\cdots i_k;i_{k+1}}(1 - p_{i_{k-r+1}\cdots i_k;i_{k+1}})]\sum_{j=1}^{n-k}1(Y_j^{(k)} = (i_1, \ldots, i_k)).$$

Let $\mathcal{F}_n = \sigma(Y_1^{(k)}, \ldots, Y_{n+1}^{(k)})$ and $S_n = \sum_{j=1}^{n} Z_j$. Then Z_n is $\mathcal{F}_n$-measurable, W_n is $\mathcal{F}_{n-1}$-measurable, (8) is satisfied,

$$E_\nu(Z_n) = E_\nu(Z_n \mid \mathcal{F}_{n-1}) = 0$$

and $\{S_n, \mathcal{F}_n\}$ is a zero-mean and square-integrable martingale. Since $|Z_j| \le 1$ we have (9). From the equalities

$$E_\nu(Z_j^2 \mid \mathcal{F}_{j-1}) = 1(Y_j^{(k)} = (i_1, \ldots, i_k))p_{i_{k-r+1}\cdots i_k;i_{k+1}}(1 - p_{i_{k-r+1}\cdots i_k;i_{k+1}})$$

and $\sum_{j=1}^{n} E_\nu\{Z_j^2 1(|Z_j|\le 1) \mid \mathcal{F}_j\} = W_{n+k}^2$ we obtain (10). Similarly one verifies (11).

Note that we can write

$$S_n = n_{i_1\cdots i_{k+1}} - p_{i_{k-r+1}\cdots i_k;i_{k+1}} n_{i_1\cdots i_k} + O(1)$$

and

$$W_n^2 = p_{i_{k-r+1}\cdots i_k;i_{k+1}}(1 - p_{i_{k-r+1}\cdots i_k;i_{k+1}}) n_{i_1\cdots i_k}.$$

And from Lemma 1 we can conclude,

Corollary 1. For $k \geq r$ and $n_{i_1\ldots i_{k+1}} > 0$ we have a.s. $[P_\nu]$

$$\limsup_{n\to\infty} \frac{(n_{i_1\cdots i_{k+1}} - p_{i_{k-r+1}\cdots i_k;i_{k+1}} n_{i_1\cdots i_k})^2}{2n_{i_1\cdots i_k} \log\log n} = p_{i_{k-r+1}\cdots i_k;i_{k+1}}(1-p_{i_{k-r+1}\cdots i_k;i_{k+1}}).$$

And if $k = r = 0$

$$\limsup_{n\to\infty} \frac{(n_i - \pi_i n)^2}{2n \log\log n} = \pi_i(1 - \pi_i).$$

2. Proof of Theorem 1

Let r be the true order and ν be any initial distribution.

(a) Assume $k > r$ and write k-th order log-likelihodd function as

$$\log L(k) = \sum_{i_1\ldots i_{k+1}} n_{i_1\cdots i_{k+1}} \log p_{i_{k-r+1}\cdots i_k;i_{k+1}}.$$

Using the inequality $-\log z \leq (1-z) + (1-z)^2 \;,\; z \geq \frac{1}{2}$, and the fact that

$$\sum_{i_{k+1}} (n_{i_1\cdots i_{k+1}} - n_{i_1\cdots i_k} \log p_{i_{k-r+1}\cdots i_k;i_{k+1}}) = 0$$

we have by (2)

$$\log \hat{L}(k) - \log L(k) = -\sum_{i_1\cdots i_{k+1}} n_{i_1\cdots i_{k+1}} \log\left(\frac{n_{i_1\cdots i_k} p_{i_{k-r+1}\cdots i_k;i_{k+1}}}{n_{i_1\cdots i_{k+1}}}\right) \leq$$

$$\sum_{i_1\ldots i_{k+1}} n_{i_1\cdots i_{k+1}} \left(\frac{n_{i_1\cdots i_{k+1}} - n_{i_1\cdots i_k} p_{i_{k-r+1}\cdots i_k;i_{k+1}}}{n_{i_1\cdots i_{k+1}}}\right) +$$

$$\sum_{i_1\ldots i_{k+1}} n_{i_1\cdots i_{k+1}} \left(\frac{n_{i_1\cdots i_{k+1}} - n_{i_1\cdots i_k} p_{i_{k-r+1}\cdots i_k;i_{k+1}}}{n_{i_1\cdots i_{k+1}}}\right)^2 \leq$$

$$\sum_{i_1\ldots i_{k+1}} n_{i_1\cdots i_k} n_{i_1\cdots i_{k+1}} \left(\frac{n_{i_1\cdots i_{k+1}} - n_{i_1\cdots i_k} p_{i_{k-r+1}\cdots i_k;i_{k+1}}}{n_{i_1\cdots i_{k+1}}}\right)^2 \qquad (12)$$

From (7) and Corollary 1 we conclude

$$\limsup_{n\to\infty} \frac{2(\log \hat{L}(k) - \log L(k))}{\log\log n} \le \sum_{i_1\ldots i_{k+1}} (1 - p_{i_{k-r+1}\cdots i_k;i_{k+1}}) \le |E|^{k+1} - |E|^k = |E|^k(|E|-1). \quad (13)$$

(b) Assume $k < r$. Note that

$$\log \hat{L}(0) = \sum_i n_i \log \frac{n_i}{n} = \sum_{i_1\ldots i_{k+1}} \log \frac{n_{i_{k+1}}}{n} + o(n).$$

Using Jensen's inequality we have for $k \ge 1$,

$$\frac{1}{n}[\log \hat{L}(0) - \log \hat{L}(k)] = \sum_{i_1\ldots i_k} \frac{n_{i_1\cdots i_k}}{n} \sum_{i_{k+1}} \frac{n_{i_1\cdots i_{k+1}}}{n_{i_1\cdots i_k}} \log\left(\frac{n_{i_{k+1}}}{n}\frac{n_{i_1\cdots i_k}}{n_{i_1\cdots i_{k+1}}}\right) + o(1)$$

$$\le \sum_{i_1\ldots i_k} \frac{n_{i_1\cdots i_k}}{n} \log 1 + o(1).$$

Similarly we show that for $k \le r-1$

$$\log \hat{L}(r) - \log \hat{L}(k) \le \log \hat{L}(r) - \log \hat{L}(r-1) \le o(1)n.$$

Next, we show that for some $d_1 > 0$ we have

$$\lim_{n\to\infty} \frac{1}{n}(\log \hat{L}(r-1) - \log \hat{L}(r)) \le -d_1. \quad (14)$$

From Proposition 1 we have

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{L}(r) = \sum_{i_1\ldots i_{r+1}} \pi_{i_1\cdots i_r} p_{i_1\cdots i_r;i_{r+1}} \log p_{i_1\cdots i_r;i_{r+1}}.$$

Write

$$\log \hat{L}(r-1) = \sum_{i_2\ldots i_{r+1}} n_{i_2\cdots i_{r+1}} \log \frac{n_{i_2,\cdots,i_{r+1}}}{n_{i_2\cdots i_r}}.$$

From Proposition 1

$$\frac{n_{i_2\cdots i_{r+1}}}{n} \xrightarrow{a.s.} \pi_{i_2\cdots i_{r+1}} \quad , \quad \frac{n_{i_2\cdots i_r}}{n} \xrightarrow{a.s.} \sum_{i_{r+1}} \pi_{i_2\cdots i_{r+1}}$$

and by (5)

$$\pi_{i_2\cdots i_{r+1}} = \sum_{i_1} \pi_{i_1\cdots i_r;i_{r+1}}.$$

It follows that

$$\lim_{n\to\infty} \frac{1}{n} \log \hat{L}(r-1) = \sum_{i_1 \ldots i_{r+1}} \pi_{i_1 \cdots i_r} p_{i_1 \cdots i_r; i_{r+1}} \log \frac{\pi_{i_2 \cdots i_{r+1}}}{\sum_{i_{r+1}} \pi_{i_2 \cdots i_{r+1}}}.$$

Applying Jensen's inequality

$$\sum_{i_1 \ldots i_r} \pi_{i_1 \cdots i_r} \sum_{i_{r+1}} p_{i_1 \cdots i_r; i_{r+1}} \log \frac{\pi_{i_2 \cdots i_{r+1}}}{p_{i_1 \cdots i_r; i_{r+1}} \sum_{i_{r+1}} \pi_{i_2 \cdots i_{r+1}}} \leq 0$$

and equality holds if and only if $p_{i_1 \cdots i_r; i_{r+1}}$ does not depend on i_1, which is a contradiction. Thus we have (13).

(c) Let $\gamma(k) = |E|^k(|E| - 1)$. From (1) and (13) we have for $k < r$

$$\begin{aligned} \text{BIC}(k) - \text{BIC}(r) &= 2 \log \hat{L}(r) - 2 \log \hat{L}(k) + (\gamma(k) - \gamma(r)) \log n \\ &\geq d_1 n + o(1) \geq d_1 \log \log n. \end{aligned}$$

And for $k > r$ write

$$\begin{aligned} \text{BIC}(k) - \text{BIC}(r) = 2(\log \hat{L}(r) &- \log L(r)) - 2(\log \hat{L}(k) - \log L(k)) \\ &+ (\gamma(k) - \gamma(r)) \log n. \end{aligned}$$

Let $d_2 = \gamma(k) - \gamma(r) > 0$. From (12) we have for same constant d_3

$$\frac{\text{BIC}(k) - \text{BIC}(r)}{\log \log n} \geq d_3 + d_2 \frac{\log n}{\log \log n} \geq d_4 > 0.$$

Thus,

$$\text{BIC}(k) - \text{BIC}(r) \geq d_4 \log \log n.$$

And this together with (13) completes the proof.

References

[1] Akaike, H. (1974) - *A new look at the statistical model identification*, IEEE Trans. Automation and Control, vol. 19, 716-723.

[2] Barron, A. Rissanen, J. and Yu, B. (1998) - *The minimum description length principle in coding and modeling*, IEEE Trans. Information Theory, vol. 44, 2743-2760.

[3] Csiszar, I. and Shields, P.C. (2000) - *The consistency of the BIC Markov order estimator*, Annals of Statistics, vol. 28, 1601-1619.

[4] Dacunha-Castelle, D. and Duflo, M. (1986) - *Probability and Statistics*, vol. 2, Springer-Verlag, New York.

[5] Finesso, L. (1992) - *Estimation of the order of a finite Markov chain*, In: Recent Advances in Mathematical Theory of Systems, Control and Network Signals (eds. H. Kimura and S. Kodama), Mita Press, 643-645.

[6] Finesso, L., Liu, C.C. and Narayan, P. (1996) - *The optimal error exponent for Markov order estimation*, IEEE Trans. Information Theory, vol. 42, 1488-1497.

[7] Gassiat, G. and Boucheron, S. (2003) - *Optimal error exponents in hidden Markov model order estimation*, IEEE Trans. Information Theory, vol. 48, 964-978.

[8] Hall, P. and Heyde, C.C. (1980) - *Martingale Limit Theory and its Application*, Academic Press, New York.

[9] Katz, R.W. (1981) - *On some criteria for estimating the order of a Markov chain*, Technometrics, vol. 23, 243-249.

[10] Merjav, N., Gutman, M. and Ziv, J. (1989) - *On the estimation of the order of a Markov chain and universal data compression*, IEEE Trans. Information Theory, vol. 35, 1014-1019.

[11] Schwarz, G. (1978) - *Estimating the dimension of a model*, Annals of Statistics, vol. 6, 461-464.

C. C. Y. Dorea
cdorea@mat.unb.br

C.R. Gonçalves
J.S. Lopes

Departamento de Matemática
Universidade de Brasilia
70910-900 - Brasilia-DF, Brasil

Smoothing of paths and weak approximation of the occupation measure of Lévy processes

Ernesto Mordecki and Mario Wschebor

ABSTRACT

Consider a real-valued Lévy process with non-zero Brownian component and jumps with locally finite variation. We obtain an invariance principle for the speed of approximation of its occupation measure by means of functionals defined on regularizations of the paths.

RESUMEN

Consideramos un proceso de Lévy con valores reales, componente Browniana no nula y saltos con variación localmente finita. Obtenemos un principio de invariancia para la velocidad de aproximación de la medida de ocupación del proceso, mediante funcionales definidos en regularizaciones de las trayectorias.

Mathematics Subject Classification (2000): 60G51, 60J75. Secondary: 60J55.

1. Lévy processes

Let $X = \{X_t \colon t \geq 0\}$ be a real-valued Lévy process, defined on a probability space $(\Omega, \mathcal{F}, P)$, that we represent by

$$X_t = \sigma W_t + S_t + mt. \tag{1}$$

Here $W = \{W_t \colon t \geq 0\}$ is a standard Wiener process, $S = \{S_t \colon t \geq 0\}$ a pure jump process with càdlàg paths, m and σ are real constants, and we assume that the Gaussian part does not vanish, i.e. $\sigma > 0$. Denote by $\mathcal{F} = \{\mathcal{F}_t \colon t \geq 0\}$ the minimal filtration generated by X, that satisfy the usual assumptions (see Jacod and Shiryaev (1987)).

Furthermore, assume that

(FV) the jump part of the process has locally finite variation, i.e. for each positive t, $\sum_{0<r\leq t} |\Delta S_r|$ is almost surely finite,

where, as usual, we denote $f(r-)$ the left limit of a càdlàg function f on the point r, and $\Delta f_r = f(r) - f(r-)$ is the magnitude of its jump at this point.

In view of (FV), the random variable S_t satisfies

$$S_t = \sum_{0<s\leq t} \Delta X_s$$

Given a positive constant a, it will be useful to define the processes $S^a = \{S_t^a : t \geq 0\}$ and $X^a = \{X_t^a : t \geq 0\}$ by

$$S_t^a = \sum_{0<s\leq t} \Delta X_s \mathbf{1}_{\{|\Delta X_s| \geq a\}}, \tag{2}$$

that is the ($a.s.$ finite) sum of jumps of the process greater or equal than a, and

$$X_t^a = mt + \sigma W_t + S_t^a, \tag{3}$$

respectively.

The characteristic function of the random variable X_t has the standard form $E\left(e^{zX_t}\right) = e^{t\kappa(z)}$, where the function $\kappa(z)$ (defined for the complex values of z such that this expectation is finite) has the form

$$\kappa(z) = mz + \frac{1}{2}\sigma^2 z^2 + \int_{\mathbb{R}} (e^{zy} - 1)\Pi(dy). \tag{4}$$

Here $\Pi(dy)$, the Lévy-Khinchine measure of the process, is a non-negative measure defined on $\mathbb{R} \setminus \{0\}$ that, in accordance with condition (FV) above, satisfies $\int (1 \wedge |y|)\Pi(dy) < \infty$. We denote by $\nu_t(dy)$ the Poisson jump measure of the process on the interval $[0,t]$. Note that for each $t > 0$, we have $a.s.$ $\nu_t(\{|x| \geq \delta\}) < \infty$ for every $\delta > 0$. For general references on Lévy processes see Skorokhod (1991), Bertoin (1996) or Sato (1999).

2. Regularized Lévy processes

We now describe the regularization of the trajectories, that, in our context, is interpreted as a partial observation of the process through a physical device. Let $\psi : \mathbb{R} \to \mathbb{R}^+$ be a C^1 function with compact support, say $\text{supp}(\psi) \subset [-1,1]$, such that $\int_{-1}^{1} \psi(t)dt = 1$ and, for $\varepsilon > 0$, define the approximation of unity

$$\psi_\varepsilon(t) = \frac{1}{\varepsilon}\psi\Big(\frac{t}{\varepsilon}\Big).$$

We denote by $\|\psi\| = \left(\int_{-1}^{1} \psi^2(t)dt\right)^{1/2}$ the norm of ψ in $L^2(\mathbb{R}, dt)$. The regularization $X^\varepsilon = \{X_t^\varepsilon : t \geq 0\}$ of the process is obtained by convolution with

ψ_ε in the following way:

$$X_t^\varepsilon = \big(\psi_\varepsilon * X\big)_t = \int_{\mathbb{R}} \psi_\varepsilon(t-s) X_s ds = \int_{-1}^{1} \psi(-w) X_{t+w\varepsilon} dw, \tag{5}$$

where we set $X_s = W_s = S_s = 0$ if $s < 0$. In the same way, we define $W^\varepsilon = \{W_t^\varepsilon : t \geq 0\}$ and $S^\varepsilon = \{S_t^\varepsilon : t \geq 0\}$, and obtain that $X_t^\varepsilon = mt - \varepsilon m\alpha + \sigma W_t^\varepsilon + S_t^\varepsilon$ where $\alpha = \int_{\mathbb{R}} w\psi(w)dw$.

Observe that the regularized processes inherits the regularity properties of ψ, so that X^ε has C^1 paths. For further reference, we compute the time-derivative (denoted with a dot) of the regularized process, that can be written as a stochastic integral:

$$\begin{aligned} \dot{X}_t^\varepsilon &= \int_{\mathbb{R}} \frac{\partial}{\partial t}\big(\psi_\varepsilon(t-s)\big) X_s ds = \frac{1}{\varepsilon}\int_{-1}^{1} \dot{\psi}(-w)\left(X_{t+\varepsilon w} - X_{t-\varepsilon}\right) dw \\ &= \int_{\mathbb{R}} \psi_\varepsilon(t-s) dX_s = \frac{1}{\varepsilon}\int_{-1}^{1} \psi(-w) d^w\big(X_{t+\varepsilon w}\big). \end{aligned} \tag{6}$$

Similar formulae hold for W^ε, S^ε and $S^{a,\varepsilon}$. In particular,

$$\dot{W}_t^\varepsilon = \frac{1}{\varepsilon}\int_{-1}^{1} \dot{\psi}(-w)(W_{t+\varepsilon w} - W_{t-\varepsilon}) dw,$$

which implies, for $t \in [0,T]$ and $0 < \varepsilon < 1$:

$$|\varepsilon \dot{W}_t^\varepsilon| \leq 2\|\dot{\psi}\|_\infty \sup_{|h|<\varepsilon, t\in[0,T]} |W_{t+h} - W_{t-\varepsilon}| \leq C_\eta(\omega)\varepsilon^{1/2-\eta}, \tag{7}$$

with $\eta \in (0, 1/2)$ arbitrary, and $C_\eta(\omega)$ a random constant independent of ε; we also have that $\sqrt{\varepsilon}\dot{W}_t^\varepsilon$ has centered Gaussian distribution, with variance $\|\dot{\psi}\|^2$.

If $F \colon \mathbb{R}^+ \to \mathbb{R}$ is a C^1 function, we denote the *number of crossings* of the level u by the function F on an interval $I = [s,t]$, by

$$N_u^F[s,t] = \sharp\{r \colon F_r = u, r \in I\}, \tag{8}$$

that is, the number of roots belonging to I of the equation $F_t = u$. It is easy to verify, that, for a given continuous function $f \colon \mathbb{R} \to \mathbb{R}$, we have

$$\int_{-\infty}^{\infty} f(u) N_u^F[0,T] du = \int_0^T f(F_t)|\dot{F}_t| dt. \tag{9}$$

3. Main Result

The aim of Theorem 1 below is to approximate the occupation measure of the process X on the interval $[0, T]$ by a re-normalization of the number of crossings of the process $X^\varepsilon = \{X_t^\varepsilon : t \geq 0\}$ with horizontal levels on the same time interval.

THEOREM 1 *Consider a Lévy process $X = \{X_t : t \geq 0\}$ with characteristic exponent given in (4), $\sigma > 0$, finite variation jump component, and the regularization $X^\varepsilon = \{X_t^\varepsilon : t \geq 0\}$ defined in (5). Then, for each C^2-function $f : \mathbb{R} \to \mathbb{R}$ with bounded second derivative, we have*

$$\frac{1}{\sqrt{\varepsilon}} \left[\int_{\mathbb{R}} f(u) C_\varepsilon \sqrt{\varepsilon} N_u^{X^\varepsilon}[0, t] du - \sigma \int_0^t f(X_s) ds \right] - C_\varepsilon \int_0^t f(X_s^\varepsilon) |\dot{S}_s^\varepsilon| ds \Rightarrow D \int_0^t f(X_s) dB_s \quad (10)$$

as $\varepsilon \to 0$, where:

- *$B = \{B_t : t \geq 0\}$ is a Wiener process independent of X;*
- *The first constant is*

$$C_\varepsilon = \frac{\sigma}{E(\sqrt{\varepsilon}|\sigma \dot{W}_1^\varepsilon + m|)} \to \frac{1}{\|\psi\|} \sqrt{\frac{\pi}{2}} = C_0 \quad (\varepsilon \to 0). \quad (11)$$

- *The second constant is*

$$D^2 = 2\sigma^2 \int_0^2 \left(r(t) \operatorname{Arsin} r(t) + \sqrt{1 - r^2(t)} - 1 \right) dt, \quad (12)$$

  where $r(t)$ is a covariance function defined by

$$r(t) = \frac{1}{\|\psi\|^2} \int \psi(t - u) \psi(-u) du.$$

- *$\Rightarrow$ denotes weak convergence in the space $\mathcal{C} = \mathcal{C}([0, +\infty), \mathbb{R})$ of continous functions.*

Before proving the Theorem we make some remarks on the statement.

Remarks.

1.- A simple consequence of Theorem 1 is that for each $t > 0$, one has

$$\int_{\mathbb{R}} f(u) C_\varepsilon \sqrt{\varepsilon} N_u^{X^\varepsilon}[0,t] du \to \sigma \int_0^t f(X_s) ds \text{ in probability} \tag{13}$$

as $\varepsilon \to 0$. This result can be used to estimate σ from the observation of the smoothed path X^ε. Results of type (13) are well-known for semimartingales having continuous paths (Azaïs & Wschebor, 1997) and also other classes of processes (Azaïs & Wschebor, 1996), where almost sure convergence is proved.

2.- Theorem 1 contains the speed of convergence in (13). This allows to make inference on σ from the observation of X^ε.

Analogous results for processes with continuous paths are in Berzin & León (1994) for Brownian motion and in Perera & Wschebor (1998, 2002) for certain classes of continuous semi-martingales having Itô-integrals as martingale part. Even if X is a Brownian motion, the proof below seems to be simpler and more direct than previously published ones.

There exist also some related results for Brownian motion and general diffusions, where the approximation X^ε of the actual path X is replaced by polygonal approximation and the smooth function f by a Dirac-delta function, or considering functionals defined on random walks. See for example, Dacunha-Castelle and Florens (1986), Florens (1993), Génon-Catalot and Jacod (1993), Borodin and Ibragimov (1994), and Jacod (1998, 2000). In this context, if ε is the size of the discretization in time, then the speed of convergence turns out to be of the order $\varepsilon^{1/4}$.

3.- We shall prove (see Proposition 3 in next section) that for each $t > 0$ the bias term

$$\mathcal{L}_\varepsilon(f,t) = C_\varepsilon \int_0^t f(X_s^\varepsilon) \big| \dot{S}_s^\varepsilon \big| ds \tag{14}$$

in (10), almost surely converges, as $\varepsilon \to 0$, to

$$\mathcal{L}_0(f,t) = C_0 \sum_{0<s\leq t} L(f,s) \big| \Delta X_s \big| \tag{15}$$

where

$$L(f,t) = \int_{-1}^{1} \psi(z) f\Big(X_{t-} \int_z^1 \psi(w) dw + X_t \int_{-1}^{z} \psi(w) dw \Big) dz. \tag{16}$$

It follows that one can replace Theorem 1 by the statement

$$\frac{1}{\sqrt{\varepsilon}}\left[\int_{\mathbb{R}} f(u)C_\varepsilon\sqrt{\varepsilon}N_u^{X^\varepsilon}[0,t]du - \sigma\int_0^t f(X_s)ds\right] - \mathcal{L}_0(f,t) \tag{17}$$

converges cilindrically to the law of $D\int_0^t f(X_s)dB_s$.

With this statement, the bias term does not depend on ε, but excepting the case of trivial f, we lose weak convergence when the jump part of X does not vanish.

4. Proofs

Proof of Theorem 1

In order to prove the Theorem we first observe, in view of (9), that

$$\int_{\mathbb{R}} f(u)N_u^{X^\varepsilon}[0,t]du = \int_0^t f(X_s^\varepsilon)|\dot{X}_s^\varepsilon|ds, \qquad a.s.$$

Write our expression as the sum of three terms:

$$\begin{aligned}
&\frac{1}{\sqrt{\varepsilon}}\int_0^t f(X_s^\varepsilon)C_\varepsilon\sqrt{\varepsilon}|\dot{X}_s^\varepsilon|ds - \frac{\sigma}{\sqrt{\varepsilon}}\int_0^t f(X_s)ds - C_\varepsilon\int_0^t f(X_s^\varepsilon)|\dot{S}_s^\varepsilon|ds\\
&\quad = C_\varepsilon\int_0^t f(X_s^\varepsilon)\big(|\dot{X}_s^\varepsilon| - |\sigma\dot{W}_s^\varepsilon + m| - |\dot{S}_s^\varepsilon|\big)ds\\
&\quad + C_\varepsilon\int_0^t f(X_s^\varepsilon)|\sigma\dot{W}_s^\varepsilon + m|ds - C_\varepsilon\int_0^t f(X_s)|\sigma\dot{W}_{s-\varepsilon}^\varepsilon + m|ds\\
&\quad + \frac{1}{\sqrt{\varepsilon}}\int_0^t \big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}_{s-\varepsilon}^\varepsilon + m| - \sigma\big)f(X_s)ds.
\end{aligned}$$

We now introduce the following simplification, that will be useful for the proof. Given an arbitrary $\delta\in(0,1)$ there exists $b>0$ such that there is no jump with absolute value greater than b, with probability greater that $1-\delta$. The given Lévy process can be written as $X_t = (X_t - S_t^b) + S_t^b$, where the random processes $\{S_t^b : t\geq 0\}$ (defined in (2)) and $\{X_t - S_t^b : t\geq 0\}$ are independent. A standard argument shows that it is enough to prove the result for the process $\{X_t - S_t^b : 0\leq t\leq T\}$, so, in what follows, we assume that the support of N is contained in the interval $[-b,b]$. Under this additional hypothesis, it is easy to see that for each $t\geq 0$ the random variable X_t has finite moments of all orders.

In what follows, the parameter of the various processes we will consider vary in a fixed interval $[0,T]$.

We divide the proof into three steps:

1. Proof of

$$Z_t^{1,\varepsilon} = \int_0^t f(X_s^\varepsilon)\big(|\dot{X}_s^\varepsilon| - |\sigma\dot{W}_s^\varepsilon + m| - |\dot{S}_s^\varepsilon|\big)ds \Rightarrow 0. \tag{18}$$

2. Proof of

$$Z_t^{2,\varepsilon} = \int_0^t f(X_s^\varepsilon)|\sigma\dot{W}_s^\varepsilon + m|ds - \int_0^t f(X_s)|\sigma\dot{W}_{s-\varepsilon}^\varepsilon + m|ds \Rightarrow 0 \tag{19}$$

3. Proof of

$$Z_t^{3,\varepsilon} = \frac{1}{\sqrt{\varepsilon}}\int_0^t \big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}_{s-\varepsilon}^\varepsilon + m| - \sigma\big)f(X_s)ds \Rightarrow D\int_0^t f(X_t)dB_s. \tag{20}$$

Proof of Step 1. We prove

$$\sup_{0\le t\le T}\Big|\int_0^t f(X_s^\varepsilon)\big(|\dot{X}_s^\varepsilon| - |\sigma\dot{W}_s^\varepsilon + m| - |\dot{S}_s^\varepsilon|\big)ds\Big| \to 0 \quad a.s. \quad (\varepsilon\to 0).$$

Since there exists an almost surely finite random variable $M(\omega)$ such that

$$\sup_{0\le s\le T}|f(X_s^\varepsilon)| \le M(\omega), \qquad \sup_{0\le s\le T}|f(X_s)| \le M(\omega), \tag{21}$$

it is enough to prove that

$$\int_0^T \Big||\dot{X}_s^\varepsilon| - |\sigma\dot{W}_s^\varepsilon + m| - |\dot{S}_s^\varepsilon|\Big|ds \to 0 \quad a.s. \quad (\varepsilon\to 0).$$

We have

$$\begin{aligned}\Big||\dot{X}_s^\varepsilon| - |\sigma\dot{W}_s^\varepsilon + m| - |\dot{S}_s^\varepsilon|\Big| &\le 2\min\big(|\sigma\dot{W}_s^\varepsilon + m|, |\dot{S}_s^\varepsilon|\big)\\ &\le 2\min\big(|\sigma\dot{W}_s^\varepsilon + m|, |\dot{S}_s^{a,\varepsilon}|\big) + 2|\dot{S}_s^{a,\varepsilon} - \dot{S}_s^\varepsilon|.\end{aligned}$$

We claim that

$$\sup_{0<\varepsilon\le 1}\int_0^T |\dot{S}_s^\varepsilon - \dot{S}_s^{a,\varepsilon}|ds \to 0 \quad a.s. \quad (a\to 0). \tag{22}$$

In fact, denoting $g(x) = |x|\mathbf{1}_{\{|x|<a\}}$, in view of (6), we have

$$\begin{aligned}|\dot{S}_t^\varepsilon - \dot{S}_t^{a,\varepsilon}| &\le \Big|\frac{1}{\varepsilon}\int_{-1}^1 |\dot{\psi}(-w)| \sum_{t-\varepsilon<v\le t+\varepsilon w} g(\Delta X_v)dw\Big|\\ &\le \frac{2}{\varepsilon}\|\dot{\psi}\|_\infty \sum_{t-\varepsilon<v\le t+\varepsilon} g(\Delta X_v).\end{aligned}$$

Furthermore, if $G(t) = \sum_{0<v\leq t} g(\Delta X_v)$ and $0 < \varepsilon \leq 1$, we obtain

$$\begin{aligned}\int_0^T |\dot S_s^\varepsilon - \dot S_s^{a,\varepsilon}|ds &\leq \frac{2}{\varepsilon}\|\dot\psi\|_\infty \int_0^T \big(G(t+\varepsilon) - G(t-\varepsilon)\big)dt \\ &= \frac{2}{\varepsilon}\|\dot\psi\|_\infty \int_0^T \Big(\int_{t-\varepsilon}^{t+\varepsilon} G(dv)\Big)dt \\ &\leq \frac{2}{\varepsilon}\|\dot\psi\|_\infty \int_0^{T+1} \Big(\int_{v-\varepsilon}^{v+\varepsilon} dt\Big)G(dv) \\ &= 4\|\dot\psi\|_\infty G(T+1) \\ &= 4\|\dot\psi\|_\infty \sum_{0<t\leq T+1} |\Delta X_t| \mathbf{1}_{\{|\Delta X_t|<a\}},\end{aligned}$$

and (22) follows.

Consider now a fixed. Put $\tau_0 = 0$, and denote the succesive epochs of jump with absolute value not smaller that a by

$$\tau_n = \inf\{t > \tau_{n-1} \colon |\Delta X_t| \geq a\} \quad (n = 1, 2, \dots).$$

Denote also by $N_t = \max\{n \colon \tau_n \leq t\}$ $(t \geq 0)$ the number of these jumps up to time t. Fix $\omega \in \Omega$, and choose $\varepsilon > 0$ such that the intervals $(\tau_n - \varepsilon, \tau_n + \varepsilon)$ $(n = 1, \dots, N_T)$ are disjoint. Taking into account that $\dot S^{a,\varepsilon} = 0$ outside these intervals, and applying (7) with $\eta = 1/4$, we obtain for $s \in [0, T]$:

$$\min\big(|\sigma \dot W_s^\varepsilon + m|, |\dot S_s^{a,\varepsilon}|\big) \leq \hat C_{1/4}(\omega)\varepsilon^{-3/4} \sum_{n=1}^{N_T} \mathbf{1}_{(\tau_n-\varepsilon,\tau_n+\varepsilon)}(s).$$

where $\hat C_{1/4}$ is a new constant depending on $\omega \in \Omega$, and on the parameters m, σ. We then obtain

$$\begin{aligned}\int_0^T \min\big(|\sigma \dot W_s^\varepsilon + m|, |\dot S_s^{a,\varepsilon}|\big)ds &\leq \sum_{n=1}^{N_T} \int_{\tau_n-\varepsilon}^{\tau_n+\varepsilon} \hat C_{1/4}(\omega)\varepsilon^{-3/4} ds \\ &= 2\hat C_{1/4}(\omega) N_T(\omega)\varepsilon^{1/4}.\end{aligned}$$

From this inequality and (22) the statement of Step 1 follows.

Proof. of Step 2. First observe that

$$\int_0^t f(X_s)|\sigma \dot W_{s-\varepsilon}^\varepsilon + m|ds = \int_{-\varepsilon}^{t-\varepsilon} f(X_{s+\varepsilon})|\sigma \dot W_s^\varepsilon + m|ds$$

which implies

$$Z_t^{2,\varepsilon} = \int_0^t \big(f(X_s^\varepsilon) - f(X_{s+\varepsilon})\big)|\sigma\dot{W}_s^\varepsilon + m|ds$$
$$- \int_{-\varepsilon}^0 f(X_{s+\varepsilon})|\sigma\dot{W}_s^\varepsilon + m|ds + \int_{t-\varepsilon}^t f(X_{s+\varepsilon})|\sigma\dot{W}_s^\varepsilon + m|ds.$$

For the second integral, we have

$$\Big|\int_{-\varepsilon}^0 f(X_{s+\varepsilon})|\sigma\dot{W}_s^\varepsilon + m|ds\Big| \le M(\omega)\hat{C}_{1/4}(\omega)\varepsilon^{1/4},$$

where $M(\omega)$ is given in (21). (Remember that $W_s = 0$ if $s < 0$, but W_s^ε does not necessarily vanishes for $s < 0$.) A similar bound holds for the third integral.

So, in order to obtain (19), we must prove that

$$\hat{Z}_t^{2,\varepsilon} = \int_0^t \big(f(X_s^\varepsilon) - f(X_{s+\varepsilon})\big)|\sigma\dot{W}_s^\varepsilon + m|ds \Rightarrow 0 \tag{23}$$

Denote

$$f_t^\varepsilon = f(X_t^\varepsilon) - f(X_{t+\varepsilon}), \qquad g_t^\varepsilon = |\sigma\dot{W}_t^\varepsilon + m|.$$

Given $S < T$ we compute the second moment:

$$E\big(\hat{Z}_T^{2,\varepsilon} - \hat{Z}_S^{2,\varepsilon}\big)^2 = 2\iint_{S\le s+2\varepsilon\le t\le T} E(f_s^\varepsilon f_t^\varepsilon g_s^\varepsilon g_t^\varepsilon)dsdt \tag{24}$$
$$+ 2\iint_{S\le s\le t\le s+2\varepsilon\le T} E(f_s^\varepsilon f_t^\varepsilon g_s^\varepsilon g_t^\varepsilon)dsdt = 2(L_1 + L_2).$$

We denote by Δ_t^ε the increment

$$\Delta_t^\varepsilon = X_t^\varepsilon - X_{t+\varepsilon} = \int_{-1}^1 \psi(-w)(X_{t+w\varepsilon} - X_{t+\varepsilon})dw$$
$$= \sigma\Delta_t^{\varepsilon,W} + \Delta_t^{\varepsilon,S} - \varepsilon m(\alpha+1)$$

(with the obvious notation). Furthermore

$$E\big(\Delta_t^{\varepsilon,W}\big)^2 = \iint \psi(-u)\psi(-v)E\big((W_{t+\varepsilon u} - W_{t+\varepsilon})(W_{t+\varepsilon v} - W_{t+\varepsilon})\big)dudv$$
$$= \varepsilon\iint \psi(-u)\psi(-v)(1 - u\vee v)dudv.$$

Now we apply Taylor's expansion:

$$
\begin{aligned}
f_t^\varepsilon &= f(X_{t+\varepsilon}+\Delta_t^\varepsilon)-f(X_{t+\varepsilon}) = f'(X_{t+\varepsilon})\Delta_t^\varepsilon + \frac{1}{2}f''(X_{t+\varepsilon}+\theta\Delta_t^\varepsilon)(\Delta_t^\varepsilon)^2/2 \\
&= f'(X_{t-\varepsilon})\Delta_t^\varepsilon + \big(f'(X_{t+\varepsilon})-f'(X_{t-\varepsilon})\big)\Delta_t^\varepsilon \qquad (25)\\
&\quad + \frac{1}{2}f''(X_{t+\varepsilon}+\theta\Delta_t^\varepsilon)(\Delta_t^\varepsilon)^2/2 \qquad (26)
\end{aligned}
$$

where $0<\theta<1$.

Take now conditional expectations in the integrand corresponding to L_1 in (28):

$$E(f_s^\varepsilon f_t^\varepsilon g_s^\varepsilon g_t^\varepsilon) = E\big(f_s^\varepsilon g_s^\varepsilon E(f_t^\varepsilon g_t^\varepsilon/\mathcal{F}_{t-\varepsilon})\big).$$

Plug the Taylor expansion for f_t^ε into the last expectation and consider each term. First, as $\Delta_t^\varepsilon g_t^\varepsilon$ is independent of $\mathcal{F}_{t-\varepsilon}$,

$$
\begin{aligned}
E\big(f'(X_{t-\varepsilon})\Delta_t^\varepsilon g_t^\varepsilon/\mathcal{F}_{t-\varepsilon}\big) &= f'(X_{t-\varepsilon})E(\Delta_t^\varepsilon g_t^\varepsilon)\\
&= f'(X_{t-\varepsilon})\Big[\sigma E(\Delta_t^{\varepsilon,W}g_t^\varepsilon)+E(\Delta_t^{\varepsilon,S})E(g_t^\varepsilon)-\varepsilon m(\alpha+1)E(g_t^\varepsilon)\Big]
\end{aligned}
$$

since S and W are independent processes.

For the first term in brackets, substracting $E\big(\Delta_t^{\varepsilon,W}|\sigma\dot W_t^\varepsilon|\big)=0$, we have

$$
\begin{aligned}
\big|E(\Delta_t^{\varepsilon,W}g_t^\varepsilon)\big| &= \Big|E\big(\Delta_t^{\varepsilon,W}(|\sigma\dot W_t^\varepsilon+m|-|\sigma\dot W_t^\varepsilon|)\big)\Big|\\
&\le |m|E\big|\Delta_t^{\varepsilon,W}\big| = (const)\varepsilon^{1/2}.
\end{aligned}
$$

In what concerns the second term in brackets,

$$E\big|\Delta_t^{\varepsilon,S}\big| \le \|\psi\|_\infty E\Big(\sum_{t-\varepsilon<s\le t+\varepsilon}|\Delta X_s|\Big) = \|\psi\|_\infty 2\varepsilon\int|x|\Pi(dx),$$

so that

$$\big|E(\Delta_t^{\varepsilon,S})E(g_t^\varepsilon)\big| \le (const)\varepsilon^{1/2},$$

and we obtain:

$$
\begin{aligned}
\big|E\big(f'(X_{t-\varepsilon})\Delta_t^\varepsilon g_t^\varepsilon \mid \mathcal{F}_{t-\varepsilon}\big)\big| &\le \big|f'(X_{t-\varepsilon})\big|\ (const)\ \varepsilon^{1/2}\\
&\le \big(\|f''\|_\infty|X_{t-\varepsilon}|+1\big)(const)\varepsilon^{1/2}.
\end{aligned}
$$

Furthermore

$$
\begin{aligned}
&\Big|E\Big(\big(f'(X_{t+\varepsilon})-f'(X_{t-\varepsilon})\big)\Delta_t^\varepsilon g_t^\varepsilon \mid \mathcal{F}_{t-\varepsilon}\Big)\Big|\\
&\quad = \Big|E\Big(\big(f''(X_{t-\varepsilon}+\theta'(X_{t+\varepsilon}-X_{t-\varepsilon})\big)(X_{t+\varepsilon}-X_{t-\varepsilon})\Delta_t^\varepsilon g_t^\varepsilon \mid \mathcal{F}_{t-\varepsilon}\Big)\Big|\\
&\qquad \le \|f''\|_\infty E\Big(\big|\sigma(W_{t+\varepsilon}-W_{t-\varepsilon})+S_{t+\varepsilon}-S_{t-\varepsilon}+2m\varepsilon\big|\times|\Delta_t^\varepsilon|g_t^\varepsilon\Big),
\end{aligned}
$$

where $0 < \theta' < 1$. A standard computation with normal distributions shows that:

$$E\big(|\Delta_t^{\varepsilon,W}|^2\,[g_t^\varepsilon]^2\,\big) \le (const)$$

So, by Cauchy-Schwarz's inequality we obtain

$$E\Big(\big|\sigma(W_{t+\varepsilon} - W_{t-\varepsilon}) + S_{t+\varepsilon} - S_{t-\varepsilon} + 2m\varepsilon\big| \times |\Delta_t^{\varepsilon,W}|g_t^\varepsilon\Big) \le (const)\varepsilon^{1/2}.$$

Also

$$E\Big(\big|W_{t+\varepsilon} - W_{t-\varepsilon}\big| \times |\Delta_t^{\varepsilon,S}|g_t^\varepsilon\Big) = E\Big(\big|W_{t+\varepsilon} - W_{t-\varepsilon}\big|g_t^\varepsilon\Big)E|\Delta_t^{\varepsilon,S}| \le (const)\varepsilon.$$

As for the other term

$$E\Big(\big|S_{t+\varepsilon} - S_{t-\varepsilon}\big| \times |\Delta_t^{\varepsilon,S}|g_t^\varepsilon\Big)$$
$$= E\Big(\big|(S_{t+\varepsilon} - S_{t-\varepsilon})\Delta_t^{\varepsilon,S}\big|\Big)E|g_t^\varepsilon| \le (const)\varepsilon^{1/2},$$

because $E|g_t^\varepsilon| \le (const)\varepsilon^{-1/2}$ and

$$E\Big(\big|(S_{t+\varepsilon} - S_{t-\varepsilon})\Delta_t^{\varepsilon,S}\big|\Big)$$
$$= E\Big|\int_{-1}^{1}\psi(-w)(S_{t+\varepsilon} - S_{t-\varepsilon})(S_{t+w\varepsilon} - S_{t+\varepsilon})dw\Big| \le (const)\varepsilon.$$

Let us now consider the result of plugging the last term of (26) into the conditional expectation. We have:

$$\Big|E\big(f''(X_{t+\varepsilon} + \theta\Delta_t^\varepsilon)(\Delta_t^\varepsilon)^2 g_t^\varepsilon/\mathcal{F}_{t-\varepsilon}\big)\Big| \le$$
$$(const)\|f''\|_\infty E\big(\sigma^2(\Delta_t^{\varepsilon,W})^2 g_t^\varepsilon + (\Delta_t^{\varepsilon,S})^2 g_t^\varepsilon + \varepsilon^2 g_t^\varepsilon\big) \le (const)\varepsilon^{1/2}, \quad (27)$$

based on similar computations. Summing up, we obtain (in the integral L_1):

$$E(f_s^\varepsilon g_s^\varepsilon/\mathcal{F}_{t-\varepsilon}) \le (const)\varepsilon^{1/2}.$$

This also shows that

$$E(f_s^\varepsilon g_s^\varepsilon) \le (const)\varepsilon^{1/2},$$

so that

$$L_1 \le (const)(T-S)^2\varepsilon. \quad (28)$$

On the other hand, let us show that for $s, t \in [0, T]$ and $0 < \varepsilon \le 1$ the expectation $E\big(f_s^\varepsilon f_t^\varepsilon g_s^\varepsilon g_t^\varepsilon\big)$ is bounded.

Applying Cauchy-Schwarz's inequality, it suffices to prove the boundedness of

$$E\left\{(f_t^\varepsilon g_t^\varepsilon)^2\right\}$$

for $t \in [0,T]$ and $0 < \varepsilon < 1$. Check that

$$E\left\{(f_t^\varepsilon g_t^\varepsilon)^2\right\} \leq (const)\left(\|f''\|_\infty + 1\right)^2 \times$$
$$\left[E\left\{X_{t-\varepsilon}^2 (\Delta_t^\varepsilon g_t^\varepsilon)^2 + E\left\{(\Delta_t^\varepsilon)^4 (g_t^\varepsilon)^2\right\}\right\} + E\left\{(X_{t+\varepsilon} - X_{t-\varepsilon})^2 (\Delta_t^\varepsilon)^2 (g_t^\varepsilon)^2\right\}\right],$$

and the proof of the boundedness of this expression follows in much a similar way as the one of L_1.

This implies, first, that

$$E\left((\hat{Z}_T^{2,\varepsilon} - \hat{Z}_S^{2,\varepsilon})^2\right) \leq (const)(T-S)^2$$

for $0 \leq S, T \leq T_0$, hence that $\{\hat{Z}_T^{2,\varepsilon} : 0 \leq T \leq T_0\}$ is tight in $\mathcal{C}([0,T_0],\mathbb{R})$ and, second, that

$$E\left((\hat{Z}_T^{2,\varepsilon})^2\right) \leq (const) T^2 \varepsilon,$$

so that, for fixed T, $\hat{Z}_T^{2,\varepsilon} \to 0$ $(\varepsilon \to 0)$ in L^2. This proves (23).

Proof. of Step 3. Introduce the processes $y^\varepsilon = \{y_t^\varepsilon : t \geq 0\}$ and $Y^\varepsilon = \{Y_t^\varepsilon : t \geq 0\}$ defined by

$$y_t^\varepsilon = C_\varepsilon \sqrt{\varepsilon} |\sigma \dot{W}_{t-\varepsilon}^\varepsilon + m| - \sigma, \qquad Y_t^\varepsilon = \frac{1}{\sqrt{\varepsilon}} \int_0^t y_s^\varepsilon ds, \qquad t \geq 0.$$

Let us prove that

$$Y^\varepsilon \Rightarrow DB, \tag{29}$$

where $B = \{B_t : t \geq 0\}$ is a Wiener process independent of X, and D the constant in (12).

In order to see this, first observe that, since y_t^ε depends on the increments of the process W on the interval $[t-\varepsilon, t+\varepsilon]$, the process y^ε is 2ε-dependent, and as a consequence, the process Y^ε has asymptotically independent increments as $\varepsilon \to 0$. This means that any cluster point in the weak topology for the family of processes $\{Y^\varepsilon : t \geq 0\}$ as $\varepsilon \to 0$ is a process with independent increments. Analogous arguments give that any cluster point has stationary increments. In order to complete the proof of (29), we prove the tightness in the space of continuous functions.

$$
\begin{aligned}
E(Y^\varepsilon_t - Y^\varepsilon_s)^4 &= \frac{1}{\varepsilon^2} E\Big(\int_s^t y^\varepsilon_u du\Big)^4 \\
&= \frac{4!}{\varepsilon^2}\int_s^t du_1 \int_{u_1}^{u_1+2\varepsilon} du_2 \int_{u_2}^t du_3 \int_{u_3}^{u_3+2\varepsilon} du_4 E(y^\varepsilon_{u_1} y^\varepsilon_{u_2} y^\varepsilon_{u_3} y^\varepsilon_{u_4}) \\
&\qquad\qquad \le 4 \times 4!(s-t)^2 E\big(y^\varepsilon_u\big)^4 \le (const)(t-s)^2,
\end{aligned}
$$

where we have used (i) the 2ε-dependence of $\{\dot W^\varepsilon_t : t \ge 0\}$, (ii) the fact that due to the choice of C_ε we have $E(y^\varepsilon_t) = 0$, and (iii) the fact that $E\big(y^\varepsilon_u\big)^4$ converges to a finite limit, as $\varepsilon \to 0$. This proves the tightness property (see 12.51 in Billingsley (1968)). As Y^ε is a centered process, in order to conclude (29) it remains to compute the constant D. This constant can be obtained as

$$
D^2 = \lim_{\varepsilon\to 0} E\big(Y^\varepsilon_1\big)^2.
$$

Now,

$$
\begin{aligned}
E\big(Y^\varepsilon_1\big)^2 &= \frac{1}{\varepsilon}\int_0^1\int_0^1 E\big(y^\varepsilon_s y^\varepsilon_t\big) ds dt = \frac{2}{\varepsilon}\int_0^1 dt \int_t^{(t+\varepsilon)\wedge 1} E\big(y^\varepsilon_s y^\varepsilon_t\big) \\
&\sim \frac{2}{\varepsilon}\int_0^{2\varepsilon} E\big(y^\varepsilon_1 y^\varepsilon_{1+t}\big) dt = 2\int_0^2 E\big(y^\varepsilon_1 y^\varepsilon_{1+\varepsilon u}\big) dt \to 2\sigma^2 \int_0^2 E\big(g(U_0) g(U_u)\big) du.
\end{aligned}
$$

with U defined in (30), and $g(x)$ defined in (32). The rest of the computation of the constant D is presented in the following result.

Given $\varepsilon > 0$ define the process $U^\varepsilon = \{U^\varepsilon_t : t \ge 0\}$ by

$$
U^\varepsilon_t = \frac{\sqrt{\varepsilon}}{\|\psi\|} \dot W^\varepsilon_{\varepsilon t - \varepsilon} \tag{30}
$$

For $t \ge 2$, U^ε is a centered Gaussian stationary process with covariance function

$$
\begin{aligned}
r(t) = E(U^\varepsilon_2 U^\varepsilon_{2+t}) &= \frac{\varepsilon}{\|\psi\|^2} E(\dot W^\varepsilon_\varepsilon \dot W^\varepsilon_{\varepsilon(1+t)}) \\
&= \frac{1}{\|\psi\|^2}\int \psi(t-u)\psi(-u) du, \qquad (31)
\end{aligned}
$$

(where we used (6)). We conclude that the distribution of U^ε does not depend on ε (excluding the interval $[0, 2]$), and introduce the process U as a centered Gaussian stationary process with covariance given by (31), that can be put in place of U^ε for our purposes. Observe that $E(U_t)^2 = 1$.

LEMMA 2 *Define*

$$g(x) = \sqrt{\frac{\pi}{2}}|x| - 1, \ (x \in \mathbb{R}). \tag{32}$$

Then

(1) $E\big(g(U_t)\big) = 0.$

(2) $E\big(g(U_0)g(U_t)\big) = r(t) \operatorname{Arsin} r(t) + \sqrt{1 - r^2(t)} - 1.$

Proof. As U_0 is a standard Gaussian random variable (1) is direct. In order to see (2), denote by

$$p(x, y, r) = \frac{1}{2\pi\sqrt{1 - r^2}} \exp\Big\{\frac{-1}{2(1 - r^2)}\Big[x^2 + y^2 - 2rxy\Big]\Big\},$$

the density of the Gaussian bidimensional vector (U_0, U_t) with $r = r(t)$. If we denote $f(r) = E\big(g(U_0)g(U_t)\big)$, it is not difficult to verify the following formal calculations:

$$\begin{aligned} f''(r) &= \frac{\partial}{\partial r} \iint_{\mathbb{R}^2} g(x)g(y) \frac{\partial}{\partial r} p(x, y, r) dx dy \\ &= \frac{\partial}{\partial r} \iint_{\mathbb{R}^2} g(x)g(y) \frac{\partial^2}{\partial x \partial y} p(x, y, r) dx dy \\ &= \frac{\partial}{\partial r} \iint_{\mathbb{R}^2} g'(x)g'(y)p(x, y, r) dx dy = \frac{\partial}{\partial r} f'(r) \qquad (33) \\ &= \iint_{\mathbb{R}^2} g''(x)g''(y)p(x, y, r) dx dy = 2\pi p(0, 0, r) = \frac{1}{\sqrt{1 - r^2}}. \end{aligned}$$

Here $g''(x) = 2\sqrt{\pi/2}\,\delta_0$ where δ_0 denotes a Dirac delta function at the origin, we twice use $\frac{\partial}{\partial r} p(x, y, r) = \frac{\partial^2}{\partial x \partial y} p(x, y, r)$, and twice integrate by parts. When $r = 0$ the random variables U_0 and U_t are independent. This gives $f(0) = E\big(g(U_0)g(U_t)\big) = E\big(g(U_0)\big)^2 = 0$, by (1); and, by the intermediate step (33) we also have $f'(0) = E\big(g'(U_0)g'(U_t)\big) = E\big(g'(U_0)\big)^2 = 0$. Finally, integrating twice we get

$$E\big(g(U_0)g(U_t)\big) = r(t) \operatorname{Arsin} r(t) + \sqrt{1 - r^2(t)} - 1,$$

concluding the proof of the Lemma.

We now claim that

$$(Y^\varepsilon, W) \Rightarrow (DB, W) \tag{34}$$

where (B, W) is a pair of standard independent Wiener processes.

For this, see first that

$$E\Big(\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_{t-\varepsilon}+m|-\sigma\big)W_s\Big)$$
$$=E\Big(\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_{t-\varepsilon}+m|-\sigma\big)\big(W_{t\wedge s}-W_{(t-\varepsilon)\wedge s}\big)\Big).$$

Now

$$\begin{aligned}\big|E(Y^\varepsilon_t W_s)\big| &= \Big|\frac{1}{D\sqrt{\varepsilon}}\int_0^t E\Big(\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_{r-\varepsilon}+m|-\sigma\big)W_s\Big)dr\Big| \\ &= \Big|\frac{1}{D\sqrt{\varepsilon}}\int_{(s-\varepsilon)\wedge t}^{s\wedge t} E\Big(\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_{r}+m|-\sigma\big)W_s\Big)dr\Big| \\ &\le \frac{\varepsilon}{D\sqrt{\varepsilon}}\Big(E\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_{t-\varepsilon}+m|-\sigma\big)^2E(W_t-W_{t-\varepsilon})^2\Big)^{1/2} \\ &= (const)\varepsilon.\end{aligned}$$

This means that $E(Y^\varepsilon_t W_s)\to 0$ $(\varepsilon\to 0)$, and, as it is direct to obtain that $\{Y^\varepsilon_t W_s\}_{\varepsilon>0}$ is uniformly integrable, we obtain (34). As a consequence, since the jump part is independent from the continous part in our Lévy process, we have the weak convergence

$$(Y^\varepsilon, W, S)\Rightarrow(DB, W, S)$$

where B is independent of X.

Let us finally see (20). Observe that for each $\varepsilon>0$, *a.s.* the process Y^ε has locally finite variation. Applying Ito's formula:

$$\begin{aligned}\frac{1}{\sqrt{\varepsilon}}\int_0^T\big(C_\varepsilon\sqrt{\varepsilon}|\sigma\dot{W}^\varepsilon_t+m|-\sigma\big)f(X_t)dt &= \int_0^T f(X_t)dY^\varepsilon_t \\ &= f(X_T)Y^\varepsilon_T-\int_0^T Y^\varepsilon_t df(X_t), \qquad (35)\end{aligned}$$

where using the hypothesis that f is C^2 it follows that $\{f(X_t)\}$ is a semimartingale. The process (Y^ε, X) is adapted, and weakly converges to (B, X).

As the integrator in the right hand member of (35) is fixed one can verify that the hypotheses of Theorem 2.2 in Kurtz and Protter (1991, see Remark 2.5) hold true, thus obtaining

$$f(X_T)Y^\varepsilon_T-\int_0^T Y^\varepsilon_t df(X_t)\Rightarrow f(X_T)B_T-\int_0^T B_t df(X_t)$$

Now, we apply Ito's formula, taking into account that the quadratic covariation $[X, B] = 0$ and we get

$$f(X_T)B_T - \int_0^T B_t df(X_t) = \int_0^T f(X_t)dB_t,$$

completing the proof of (20).

To finish, we state and prove the proposition announced in Remark 3 after the statement of Theorem 1.

PROPOSITION 3 *Assume that* $X = \{X_t \colon t \geq 0\}$ *and* f *satisfy the hypothesis of Theorem 1. Then, for the processes defined in* (15) *and* (14), *for each* $t \geq 0$, *almost surely*

$$\mathcal{L}_\varepsilon(f, t) \to \mathcal{L}_0(f, t),$$

as $\varepsilon \to 0$.

Proof. On account of (22) it suffices to show that for fixed $a > 0$, almost surely

$$C_0 \int_0^t f(X_s^{a,\varepsilon})|\dot{S}_s^{a,\varepsilon}|ds \to \mathcal{L}_0^a(f, t), \tag{36}$$

as $\varepsilon \to 0$, where $\mathcal{L}_0^a(f, t)$ is obtained from (15) when the process X is replaced by X^a.

Using the same notations as in the last part of Proof of Step 1 in Theorem 1, we can write $a.s.$, for ε sufficiently small

$$\int_0^t f(X_s^{a,\varepsilon})|\dot{S}_s^{a,\varepsilon}|ds = \sum_{n=1}^{N_t} \int_{\tau_n-\varepsilon}^{\tau_n+\varepsilon} f(X_s^{a,\varepsilon})|\dot{S}_s^{a,\varepsilon}|ds. \tag{37}$$

Observe that for $\tau_n - \varepsilon < s < \tau_n + \varepsilon$ one has: $\dot{S}_s^{a,\varepsilon} = \frac{1}{\varepsilon}\psi\left(\frac{s-\tau_n}{\varepsilon}\right)\Delta X_{\tau_n}$, so that

$$\int_{\tau_n-\varepsilon}^{\tau_n+\varepsilon} f(X_s^{a,\varepsilon})|\dot{S}_s^{a,\varepsilon}|ds = \frac{1}{\varepsilon}\int_{\tau_n-\varepsilon}^{\tau_n+\varepsilon} f(X_s^{a,\varepsilon})\psi\left(\frac{s-\tau_n}{\varepsilon}\right)|\Delta X_{\tau_n}|\ ds.$$

Making the change of variables $z = (s - \tau_n)/\varepsilon$ in each integral, we obtain

$$\int_0^t f(X_s^{a,\varepsilon})|\dot{S}_s^{a,\varepsilon}|ds = \sum_{n=1}^{N_t} |\Delta X_{\tau_n}| \int_{-1}^1 f(X_{\tau_n+\varepsilon z}^{a,\varepsilon})\psi(z)dz.$$

To compute the limit as $\varepsilon \to 0$ in the right hand member of the last equality, use that

$$X_{\tau_n+\varepsilon z}^{a,\varepsilon} \to X_{\tau_n}^a \int_{-1}^z \psi(w)dw + X_{\tau_n-}^a \int_z^1 \psi(w)dw.$$

This proves the statement.

References

[1] Azaïs, J-M. and Wschebor, M. Almost sure oscillation of certain random processes, *Bernoulli* **2** (3), 1996, 257–270.

[2] Azaïs, J-M. and Wschebor, M. Oscillation presque sûre de martingales continues, *Séminaire de Probabilités XXXI, Lecture Notes in Mathematics 1655,* Springer, 1997, 69–76.

[3] Bertoin, J. *Lévy Processes.* Cambridge University Press, Cambridge, 1966.

[4] Berzin, C. and León, J.R. Weak convergence of the integrated number of level crossings to the local time of the Wiener process. *Comptes R. Acad. Sc. Paris*, Sér. I **319**, 1994, 1311–1316.

[5] Billingsley, P. *Convergence of probability measures.* Wiley, New York, London, Sydney, 1968.

[6] Borodin, A.N. and Ibragimov, I.A. Limit theorems for functionals of random walks. (Predel'nye teoremy dlya funktsionalov ot sluchajnykh bluzhdanij.) (Russian) Trudy Matematicheskogo Instituta Imeni V. A. Steklova. 195. Nauka, Sankt-Peterburg, 1994. (See also: Borodin, A.N. and Ibragimov, I.A. Limit theorems for functionals of random walks. (English) Proceedings of the Steklov Institute of Mathematics. 195. American Mathematical Society (AMS), Providence, RI, 1995.)

[7] Dacunha-Castelle, D. and Florens-Zmirou, D. Estimation of the coefficient of a diffusion from discrete observations, *Stochastics* **19**, 1986, 263–284.

[8] Florens-Zmirou, D. On estimating the difussion coefficient from discrete observations, *J. Appl. Prob.* **30**, 1993, 790–804.

[9] Génon-Catalot, V. and Jacod, J. On the estimation of the diffusion coefficient for multidimensional diffusion processes, *Ann.Inst. H.Poincaré, Probab. Stat.* **29**, 1993, 119–151.

[10] Jacod, J. Rates of convergence to the local time of a diffusion. *Ann.Inst. H.Poincaré, Probab. Stat.* **34**, 1998, 505–544.

[11] Jacod, J. Non-parametric kernel estimation of the diffusion coefficient of a diffusion. *Scand. J. Statist.* **27** (1), 2000, 83–96.

[12] Jacod, J. and Shiryaev, A.N. *Limit Theorems for Stochastic Processes.* Springer, Berlin, Heidelberg, 1987.

[13] Kurtz, T. G. and Protter, P. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.* **19** (3), 1991, 1035–1070.

[14] Perera, G. and Wschebor, M. Crossings and occupation measures for a class of semimartingales, *Ann. Probab.* **26** (1), 1998, 253–266.

[15] Perera, G. and Wschebor, M. Inference on the Variance and Smoothing of the Paths of Diffusions, *Ann.Inst. H.Poincaré, Probab. Stat.* **38** (6), 2002, 1009–1022.

[16] Sato, Ken-iti. *Lévy processes and infinitely divisible distributions.* Cambridge Studies in Advanced Mathematics, 68. Cambridge University Press, Cambridge, 1999.

[17] Skorokhod, A. V. *Random processes with independent increments.* Kluwer Academic Publishers, Dordrecht, 1991.

Ernesto Mordecki
mordecki@cmat.edu.uy

Mario Wschebor
wschebor@cmat.edu.uy

Centro de Matemática. Facultad de Ciencias.
Universidad de la República.
Calle Iguá 4225. 11400 Montevideo. Uruguay.

On the Lévy-Khintchine Representation of Lévy Processes in Cones of Banach Spaces

Víctor Pérez-Abreu[1] and Alfonso Rocha-Arteaga[2]

ABSTRACT

Subordinators in Banach spaces are considered. The existence of the special Lévy-Khintchine representation is related to the geometry of the space and a Pettis integral with respecto to the underlying Lévy measure.

RESUMEN

Se consideran subordinadores en espacios de Banach. Se relaciona la existencia de la representación especial de Lévy-Khintchine con la geometría del espacio y con una integral de Pettis con respecto a la medida de Lévy subyacente.

Key Words: Pettis integral, Subordinator, geometry of Banach spaces
Mathematics Subject Classification(2000): 60G51, 60B05.

Introduction

One dimensional increasing Lévy processes (also called subordinators) have been widely studied. They have important properties and are useful in building other Lévy processes; see [3], [4] and [17]. A one dimensional subordinator $\{\sigma_t : t \geq 0\}$ is a nonnegative Lévy process characterized by the special form of the Lévy-Khintchine representation of its Fourier transform

$$Ee^{iu\sigma_t} = \exp\left\{ t \int_{(0,\infty)} \left(e^{iux} - 1\right) \nu\left(dx\right) + itu\gamma_0 \right\} \quad u \in \mathbb{R}, \tag{1}$$

where there is no Gaussian part, the *drift* γ_0 is nonnegative and the Lévy measure ν is concentrated in the cone $[0,\infty)$ with the order of singularity

$$\int_{(0,1]} x\nu(dx) < \infty. \tag{2}$$

[1]This paper was written while this author was visiting the School of Mathematics, Georgia Institute of Technology. He acknowledges the support of this Institute.

[2]The work of this author was supported by PROMEP Grant 103.5/04/2765 UASIN-EXB-236.

The study of subordinators in higher dimensional spaces leads to consider Lévy processes with values in cones. When a cone is proper, a Lévy process is cone-valued if and only if it is cone-increasing. Then it is natural to call these processes cone-valued subordinators or simply subordinators. In the finite dimensional case, cone-valued Lévy processes are already discussed in the classical books by Bochner [5] and Skorohod [18] and have been recently studied in [2], [13] and [14].

The purpose of this paper is to study the structure of Lévy processes with values in cones of infinite dimensional Banach spaces and investigate whether there is an intrinsic relation between probability and functional analytic aspects. As a basic point of interest one may ask whether every subordinator has a special form of the Lévy-Khintchine representation (SLKR) similar to the one dimensional case (1). We call these subordinators regular. A restricted class of Banach spaces where the answer is affirmative was considered in [9] and [15]. In the present work we conduct a systematic study of cone-valued Lévy processes in more general Banach spaces and on the convergence of their non-compensated jumps. We find that the existence of the SLKR for every subordinator is related to the geometry of the Banach space. Specifically, it is pointed out that a cone-valued subordinator is always regular for proper normal cones containing no copy of c_0^+, the cone of nonnegative scalar sequences converging to zero. This fact is also related to the existence of a Pettis integral with respect to the underlying Lévy measure.

The paper is organized as follows. In section 2 we review basic facts about convergence in cones of Banach spaces. In Section 3 a systematic and detailed analysis of cone-valued Lévy processes is done. We first review some facts about Banach space-valued Lévy processes and give sufficient conditions to have the SLKR and being cone-valued. These conditions include the existence of a Pettis integral with respect to the Lévy measure. Then we consider the converse problem, by studying the convergence of the non-compensated sum of jumps of cone-valued subordinators and conditions on the geometry of the space for the existence of the SLKR and the associated Pettis integral. In section 4 we concentrate on regular subordinators, including the class of LK-cones in which every subordinator is regular. We prove that the cones considered in [9] are particular cases of LK-cones.

1. Preliminaries on Cones in Banach Spaces

Recall that a nonempty closed convex set K of B is said to be a *cone* if $\lambda \geq 0$ and $x \in K$ imply $\lambda x \in K$. Note that a cone is closed under finite sums and contains the zero element. A cone K is *generating* if $B = K - K$, that is, every $x \in B$ can be written as $x = x_1 - x_2$ for $x_1 \in K$ and $x_2 \in K$. It is *proper* if $x = 0$ whenever x and $-x$ are in K. The *dual cone* K^* of K is defined as $K^* = \{f \in B^* : f(s) \geq 0 \text{ for every } s \in K\}$. A proper cone K of a Banach space B induces a partial order on B by defining $x_1 \leq_K x_2$ whenever $x_2 - x_1 \in K$ for $x_1 \in B$ and $x_2 \in B$. Given a sequence $(x_n) = (x_n)_{n=1}^{\infty}$ in B, if $x_n \leq_K x_{n+1}$ (respectively $x_{n+1} \leq_K x_n$) for each $n \geq 1$, the sequence is called *K-increasing* (respectively *K-decreasing*). Likewise, a function $f : [0, \infty) \to B$ is called *K-increasing* (respectively *K-decreasing*) if $f(t_1) \leq_K f(t_2)$ (respectively $f(t_2) \leq_K f(t_1)$) for $t_1 \leq t_2$.

A sequence (x_n) in K is said to be *K-majorized* if there exists $x \in K$ with $x_n \leq_K x$, for $n \geq 1$. A cone K is said to be *regular* if every K-increasing and K-majorized sequence in K is norm convergent. A cone K is said to have *generating dual* if K^* is generating for B^*. A useful characterization in Banach spaces is that K^* is generating for B^* if and only if K is normal, *i.e.*, $0 \leq_K x \leq_K y$ for $y \in K$, implies $\|x\| \leq \lambda \|y\|$, where $\lambda > 0$ ([10, Th. 1.5.4 and Prop. 1.5.7]).

Given two cones K_1 and K_2 of the Banach spaces B_1 and B_2, it is said that K_1 is *isomorphic to* K_2 if there is an isomorphism φ between $\overline{Span}(K_1)$ and $\overline{Span}(K_2)$ such that $\varphi(K_1) = K_2$.

Let c_0 denote the Banach space of real sequences $a = (a_n)$ converging to zero with norm $\|a\| = \sup_{n\geq 1} |a_n|$ and let c_0^+ denote the cone of c_0 consisting of all sequences with nonnegative terms. The cone c_0^+ plays an important role in the study of convergence of cone valued series and sequences. The following Bessaga-Pelczynski type result for convergence of series in Banach spaces follows straightforward from Theorem 5.8 in [7], assuming that the summands are cone-valued. Recall that a series $\sum_{k=1}^{\infty} x_k$ in B is *weakly unconditionally Cauchy (w.u.C.)* if for all $f \in B^*$, $\sum_{k=1}^{\infty} |f(x_k)|$ is a real convergent series.

PROPOSITION 1 *Let K be a cone of a Banach space B. In order that any w.u.C. series $\sum_{k=1}^{\infty} x_k$, with $x_n \in K$, $n \geq 1$, be (norm) unconditionally convergent, it is necessary and sufficient that K contain no subcone isomorphic to c_0^+.*

The following result is an easy consequence of Proposition 1.

PROPOSITION 2 *A normal cone that contains no subcone isomorphic c_0^+ is regular.*

In view of the main results of this work we introduce the following terminology.

DEFINITION 3 *A proper cone is said to be an LK-cone if it is normal and contains no subcone isomorphic to* c_0^+ .

2. Cone-valued Lévy Processes

Throughout this section B will denote a separable Banach space with norm $\|\cdot\|$, topological dual space B^* and dual norm $\|\cdot\|^*$.

2.1. The special Lévy-Khintchine representation

Recall that a Banach space valued (i.e. a B-valued) *Lévy process* $X = \{X_t : t \geq 0\}$ is a stochastic process with values in B defined on a probability space $(\Omega, \mathcal{F}, P)$ such that i) $X_0 = 0$ a.s., ii) it has independent and stationary increments, iii) it is stochastically continuous with respect to the norm $\|\cdot\|$ (for every $\varepsilon > 0$, $P(\|X_t - X_s\| > \varepsilon) \to 0$ as $s \to t$) and iv) almost surely the paths are right-continuous in $t \geq 0$ and have left-limits in $t > 0$ (*càdlàg*) with respect to the norm.

Let $D_0 = \{0 < \|x\| \leq 1\}$. The following Lévy-Khintchine representation for B-valued Lévy processes is presented in [9].

THEOREM 4 *Let* $\{X_t : t \geq 0\}$ *be a Lévy process in a separable Banach space* B. *Then, its characteristic functional is such that*

$$Ee^{if(X_t)} = \exp\left\{t\left(-\frac{1}{2}(Af, f) + if(\gamma) + \psi(f, \nu)\right)\right\} \qquad f \in B^*, \qquad (3)$$

where

$$\psi(f, \nu) = \int \left[e^{if(x)} - 1 - if(x)1_{D_0}(x)\right] \nu(dx), \qquad (4)$$

$\gamma \in B$, *A is a nonnegative selfadjoint operator from B^* to B, the Lévy measure ν on $\mathcal{B}(B\backslash\{0\})$ the Borel σ-algebra of $B\backslash\{0\}$, is such that for any $f \in B^*$*

$$\int \left(1 \wedge |f(x)|^2\right) \nu(dx) < \infty. \qquad (5)$$

The triplet of parameters (A, ν, γ) in Theorem 4 is called the *generating triplet* of the Lévy process X and it is unique.

REMARK 5 *a) For infinite dimensional Banach spaces, Lévy measures are not characterized by the condition* $\int \left(1 \wedge \|x\|^2\right) \nu(dx) < \infty$; *see [1] and [11]. They are rather identified by the fact that the mapping*

$$f \longmapsto \exp\{\psi(f, \nu)\} \quad f \in B^* \qquad (6)$$

where $\nu(\{0\}) = 0$ is the characteristic functional of some probability measure on B. As a consequence of this definition $\nu(\{x \in B : ||x|| > \varepsilon\}) < \infty$ for every $\varepsilon > 0$.

b) Denote $X_{s-} := \lim_{s\uparrow t} X_s$ and let $C \in \mathcal{B}_0$ the ring of Borel sets of B with positive distance from 0, then the process

$$X_t^C = \sum_{s<t} (X_s - X_{s-}) 1_C (X_s - X_{s-}) \tag{7}$$

is well defined and represents the sum of jumps of the process X occurred up to time t and took place in C. Its characteristic functional has the form

$$Ee^{if(X_t^C)} = exp\left\{t \int_C \left(e^{if(x)} - 1\right) \nu(dx)\right\} \qquad f \in B^*. \tag{8}$$

As in the finite dimensional case, there is a one-to-one relation between Lévy processes and infinitely divisible laws in Banach spaces.

PROPOSITION 6 *Let μ be an infinitely divisible probability measure with generating triplet (A, ν, γ). Then there exists a Lévy process $\{X_t : t \geq 0\}$ with generating triplet (A, ν, γ) such that X_1 has the law μ and viceversa.*

The Lévy-Khintchine representation for Lévy processes of bounded variation is also derived in [9].

PROPOSITION 7 *Let $\{Z_t : t \geq 0\}$ be a B-valued Lévy process. Then, $\{Z_t\}$ has bounded variation on each interval $[0, t]$, almost surely, if and only if, it has characteristic functional given by*

$$Ee^{if(Z_t)} = exp\left\{t \int_B \left(e^{if(x)} - 1\right) \nu(dx) + itf(\gamma)\right\} \qquad f \in B^*,$$

where the Lévy measure ν satisfies

$$\int (1 \wedge \|x\|) \, \nu\,(dx) < \infty. \tag{9}$$

A straightforward but key observation is that a Lévy process is cone-increasing if and only if it is cone-valued. The proof is as in the finite dimensional case in Theorem 83 of [16].

PROPOSITION 8 *Let K be a proper cone of B and let $\{Z_t : t \geq 0\}$ be a B-valued Lévy process. Then the following are equivalent:*

a) For any fixed $t \geq 0$, Z_t is concentrated on K almost surely.

b) Almost surely, $Z_t(\omega)$ is K-increasing in t, i. e. for s,t with $s \leq t$, we have $Z_s \leq_K Z_t$.

In view of the above result and similar to the one dimensional case, we shall say that a K-increasing Lévy process in B is a *K-subordinator* or *subordinator* if the underlying cone is well understood.

Given a Lévy measure ν on $\mathcal{B}(B\backslash\{0\})$, we shall say that an element $I_\nu \in B$ is a *ν-Pettis centering* if

$$\int_{D_0} |f(x)|\, \nu(dx) < \infty \qquad \text{for every } f \in B^* \tag{10}$$

and $f(I_\nu) = \int_{D_0} f(x)\nu(dx)$ for every $f \in B^*$. Sometimes we shall write $I_\nu = \int_{D_0} x\nu(dx)$.

Sufficient conditions on the generating triplet (A, ν, γ) of a Lévy process $\{Z_t : t \geq 0\}$ to be cone-subordinator are now presented. The only assumption on the cone K is that K is proper.

THEOREM 9 *Let K be a proper cone of a separable Banach space B. Let $\{Z_t : t \geq 0\}$ be a Lévy process in B with generating triplet (A, ν, γ). Assume the following three conditions*

a) $A = 0$,

b) $\nu(B\backslash K) = 0$, i.e., ν is concentrated on K and

c) there exists a ν-Pettis centering $I_\nu = \int_{D_0} x\nu(dx)$ such that $\gamma_0 := \gamma - I_\nu \in K$.

Then the process Z is a subordinator.

Observe that assumptions (a)-(c) above give the Fourier transform

$$Ee^{if(Z_t)} = exp\left\{ t \int_K \left(e^{if(x)} - 1 \right) \nu(dx) + itf(\gamma_0) \right\}, \tag{11}$$

since for all $f \in B^*$, $f(\gamma_0) = f(\gamma) - \int_{D_0} f(x)\nu(dx)$. We refer to (11) together conditions (a)-(c) as the *special Lévy-Khintchine representation.* Before proving the theorem we point out the following facts.

REMARK 10 *a) The existence of a ν-Pettis centering is related to the ν-integrability of the function $h(x) = x1_{D_0}(x)$ in the sense of Pettis. Recall that the B-valued function h is Pettis integrable w.r.t. ν if (10) is satisfied and for any $C \in \mathcal{B}(B\backslash\{0\})$ there exists and element I^C in B such that*

$f(I^C) = \int_{D_0} f(x)1_C(x)\nu(dx)$ for every $f \in B^*$. The element I^C is called the Pettis integral and we have that $I_\nu = I^U$, where $U = \{x \in B : x \in D_0\}$. We refer to [12] for a recent survey on the Pettis integral.

b) There are subordinators in Banach spaces whose Lévy measure satisfies the stronger integrability condition (9) for the norm (see [15]). This yields the existence of $I_\nu = \int_{D_0} x\nu(dx)$ as a Bochner integral.

The Laplace transform of a subordinator with Fourier transform (11) is obtained by standard analytic continuation.

PROPOSITION 11 *Let K be a proper cone of a separable Banach space B and let $\{Z_t : t \geq 0\}$ be a Lévy process with the special representation (11) where (10) is satisfied. Then the Laplace transform of Z_t is given by*

$$Ee^{-f(Z_t)} = exp\{-t\Phi(f)\} \qquad f \in K^*. \tag{12}$$

with Laplace exponent

$$\Phi(f) = \int_K \left(1 - e^{-f(x)}\right)\nu(dx) + f(\gamma_0). \tag{13}$$

Proof. [Proof of Theorem 9] We have to show that Z_t takes values in K almost surely. For each $\varepsilon > 0$, consider the jumps sum process $Z_t^{\triangle_\varepsilon}$ defined by (7) where $\triangle_\varepsilon = \{x : \|x\| > \varepsilon\}$. We observe that $Z_t^{\triangle_\varepsilon} \in K$ almost surely. Indeed, suppose that there exists $C \in \mathcal{B}_0$ contained in $B \backslash K$ such that the process in (7) satisfies $Z_t^C \neq 0$ with positive probability, then $0 < P\left(Z_t^C \neq 0\right) \leq 1 - e^{-t\nu(C)}$ which is not possible since $\nu(C) = 0$. Thus $Z_t^{\triangle_\varepsilon} \in K$ a.s.

Similar to the one-dimensional case, since $\gamma_0 \in K$ it is enough to prove that $J_t = Z_t - t\gamma_0 \in K$ almost surely. Notice that J_t and Z_t have the same jumps and therefore $J_t^{\triangle_\varepsilon} = Z_t^{\triangle_\varepsilon}$ a.s. Hence $\left\{J_t^{\triangle_\varepsilon}\right\}$ is a K-valued process and its characteristic functional is given by (8) on the Borel set $K \cap \{x : \|x\| > \varepsilon\}$. Since K is convex and closed then $K = \cap_{k=1}^{\infty}\{x : g_k(x) \geq 0\}$ for a sequence of continuous linear functionals g_k. Then, for $u \geq 0$,

$$Ee^{-ug_k(J_t^{\triangle_\varepsilon})} = exp\left\{t\int_{K\cap\{x:\|x\|>\varepsilon\}}\left(e^{-ug_k(x)} - 1\right)\nu(dx)\right\}.$$

Letting $\varepsilon \downarrow 0$ we get

$$Ee^{-u\lim_{\varepsilon\downarrow 0} g_k(J_t^{\triangle_\varepsilon})} = exp\left\{t\int_K\left(e^{-ug_k(x)} - 1\right)\nu(dx)\right\},$$

where the right hand side is finite by (10) and tends to 1 as u decreases to zero. Therefore $\lim_{\varepsilon\downarrow 0} g_k(J_t^{\triangle_\varepsilon})$ exists a.s. and it is nonnegative for each k. From (11) and the fact that $Z_t^{\triangle_\varepsilon}$ and $Z_t - Z_t^{\triangle_\varepsilon}$ are independent

$$Ee^{ig_k(J_t - J_t^{\triangle_\varepsilon})} = exp\left\{t\int_{K\cap\{x:\|x\|\leq\varepsilon\}}\left(e^{ig_k(x)}-1\right)\nu(dx)\right\}$$

where the right hand side tends to 1 as $\varepsilon\downarrow 0$ and therefore almost surely $g_k(J_t) = \lim_{\varepsilon\downarrow 0} g_k(J_t^{\triangle_\varepsilon})$. Since $g_k\left(J_t^{\triangle_\varepsilon}\right)$ is nonnegative and increasing as $\varepsilon\downarrow 0$ for all k, then $g_k(J_t)\geq 0$ and therefore $J_t\in K$. □

Proposition 7 and Theorem 9 yield a stronger result.

COROLLARY 12 *Let K be a proper cone of B. Let $\{Z_t : t\geq 0\}$ be a B-valued Lévy process with generating triplet (A,ν,γ) satisfying (a), (b), (c) in Theorem 9 as well as the additional condition (9). Then $\int_{D_0} x\nu(dx)$ is a Bochner integral and the process Z is a subordinator of bounded variation.*

While in finite dimensions every subordinator is of bounded variation, for infinite dimensional Banach spaces there are subordinators of unbounded variation. This is the case when the ν-Pettis centering I_ν is not Bochner integrable. For example, let $\sigma_n, n\geq 1$, be a sequence of one-dimensional independent subordinators, where for each $n\geq 1$, σ_n has generating triplet $(0,\nu_n,n^{-1})$ with Lévy measure $\nu_n = n^{-2}\delta_{\{n\}}$. Then $Z_t = (\sigma_1(t),\sigma_2(t),...)$ is a subordinator in c_0^+ with drift $\gamma_0 = 0$ and Lévy measure $\nu(C) = \sum_{n=1}^\infty n^{-2}1_C(ne_n)$, where $\{e_n\}_{n\geq 1}$ is the sequence of unit vectors in c_0^+. Then

$$\int_{D_0}\|x\|\,\nu(dx) = \sum_{n=1}^\infty n^{-1} = \infty,$$

but $I_\nu = \sum_{n=1}^\infty n^{-1}e_n \in c_0^+$ and for any $f\in l_1 = c_0^*$, $f = (f_1,, f_2,...)$,

$$\int_{D_0}|f(x)|\,\nu(dx) = \sum_{n=1}^\infty n^{-1}|f_n| < \infty.$$

From Theorem 9 we have that Z has the special Lévy-Khintchine representation and therefore from (9) in Proposition 7, Z cannot have bounded variation.

2.2. Convergence of non-compensated jumps

Whether the converse of Theorem 9 is true or not, relies in the analysis of the sums of the non-compensated jumps of the subordinator falling into the cone. While in the finite dimensional case these sums are always convergent ([18, Th. 3.21]), for infinite dimensional Banach spaces a more detailed analysis is needed.

Fix $t > 0$ and consider for each $\varepsilon > 0$, the sum of non-compensated jumps of size bigger than ε, see (7),

$$Z_t^{\triangle_\varepsilon} = \sum_{s<t} \Delta Z_s 1_{\triangle_\varepsilon}(\Delta Z_s) \tag{14}$$

where $\triangle_\varepsilon = \{x : \|x\| > \varepsilon\}$ and $\Delta Z_s = Z_s - Z_{s-}$. Observe that for any sequence $\varepsilon_n \downarrow 0$, one has the alternative representation of (14) as sums of independent random elements in K

$$Z_t^{\triangle_{\varepsilon_n}} = \sum_{k=1}^{n} \xi_k \tag{15}$$

where $\xi_1 = Z_t^{\triangle_{\varepsilon_1}}, \xi_k = \sum_{s<t} \Delta Z_s 1_{\triangle_{\varepsilon_k,\varepsilon_{k-1}}}(\Delta Z_s), k \geq 2$, with $\triangle_{\varepsilon_k,\varepsilon_{k-1}} = \{x : \varepsilon_k < \|x\| \leq \varepsilon_{k-1}\}$.

We first prepare a technical lemma on the one-dimensional processes $f(Z_t)$, $f \in B^*$.

LEMMA 13 *Let K be a proper cone of B and let $\{Z_t : t \geq 0\}$ be a K-valued Lévy process. For any $f \in K^*$, the one dimensional family $f(Z_t^{\triangle_\varepsilon})$ is nonnegative, increasing as $\varepsilon \downarrow 0$ and bounded by $f(Z_t)$, almost surely.*

Proof. The process $\{f(Z_t) : t \geq 0\}$ is a one dimensional subordinator since $\{Z_t\}$ is a K-valued Lévy process and $f \in K^*$. Its corresponding sum of non-compensated jumps

$$[f(Z_t)]^{\triangle_\varepsilon^f} := \sum_{s<t} (f(Z_s) - f(Z_{s-})) 1_{\{f(x):|f(x)|>\varepsilon\}}(f(Z_s) - f(Z_{s-}))$$

is nonnegative, increasing as $\varepsilon \downarrow 0$ and bounded by $f(Z_t)$. Hence the limit $lim_{\varepsilon\downarrow 0} [f(Z_t)]^{\triangle_\varepsilon^f}$ exists.

Let ε_n be any decreasing sequence to 0 and let $\varepsilon_n(f) = \varepsilon_n \|f\|^*$. Assume, without loss of generality, that $\|f\|^* > 0$. If $f(\Delta Z_s) > \varepsilon_n(f)$ then $\|\Delta Z_s\| > \varepsilon_n$. Therefore $[f(Z_t)]^{\triangle_{\varepsilon_n(f)}^f} \leq f(Z_t^{\triangle_{\varepsilon_n}})$ almost surely for every $n \geq 1$ and hence $f(Z_t^{\triangle_{\varepsilon_n}})$ is nonnegative almost surely. Next, when $\varepsilon_2 < \varepsilon_1$ we have that

$\triangle_{\varepsilon_1} \subset \triangle_{\varepsilon_2}$ and hence $f(Z_t^{\triangle_{\varepsilon_2}} - Z_t^{\triangle_{\varepsilon_1}}) = \sum_{s<t} f(\Delta Z_s) 1_{\{x:\varepsilon_2<\|x\|\leq\varepsilon_1\}}(\Delta Z_s) \geq 0$, proving the increasingness. Finally, let $\varepsilon > 0$ and note that $f(Z_t^{\triangle_\varepsilon}) = \sum_{s<t} f(\Delta Z_s) 1_{\{x:\|x\|>\varepsilon\}}(\Delta Z_s)$ is bounded by $f(Z_t)$ almost surely since it represents a finite number of jumps of the one-dimensional subordinator $\{f(Z_t)\}$.

□

For each $t > 0$ the jumps sum $Z_t^{\triangle_\varepsilon}$ is K-increasing as function of ε and it is K-majorized by Z_t.

LEMMA 14 *$Z_t^{\triangle_{\varepsilon_2}} - Z_t^{\triangle_{\varepsilon_1}} \in K$ for $\varepsilon_1 < \varepsilon_2$ and $Z_t - Z_t^{\triangle_\varepsilon} \in K$ for $\varepsilon > 0$.*

Proof. If $\varepsilon_1 < \varepsilon_2$ then $Z_t^{\triangle_{\varepsilon_2}} - Z_t^{\triangle_{\varepsilon_1}} = \sum_{s<t} \Delta Z_s 1_{\{x:\varepsilon_1<\|x\|\leq\varepsilon_2\}}(\Delta Z_s) \in K$. This proves the first assertion. If $0 \leq s < t$ then $Z_{t-} - Z_s = \lim_{\varepsilon\downarrow 0} Z_{t-\varepsilon} - Z_s \in K$. Hence, if $0 < s_1 < s_2 < ... < s_n \leq t$

$$Z_t - \sum_{k=1}^{n} (Z_{s_k} - Z_{s_k-}) = Z_t + \sum_{k=1}^{n} (Z_{s_k-} - Z_{s_k})$$
$$= Z_t - Z_{s_n} + (Z_{s_2-} - Z_{s_1}) + ... + (Z_{s_n-} - Z_{s_{n-1}}) + Z_{s_1-}$$

which belongs to K. Thus $Z_t - Z_t^{\triangle_\varepsilon} \in K$ for all $\varepsilon > 0$. □

In order to get more insight into the structure of cone-valued Lévy processes on infinite dimensional Banach spaces, we require additional assumptions on the cone. The following *weak* result always holds for subordinators with values in normal cones.

THEOREM 15 *Let K be a proper normal cone of B and let $\{Z_t : t \geq 0\}$ be a K-valued Lévy process. Then the jumps sum $Z_t^{\triangle_{\varepsilon_n}}$ is w.u.C. almost surely for any sequence $\varepsilon_n \downarrow 0$.*

Proof. Because K is normal, it is generating dual, i.e. any $f \in B^*$ can be decomposed into $f = f^+ - f^-$ where $f^+ \in K^*$ and $f^- \in K^*$. It is enough to prove the assertion for any positive (with respect to K) linear functional since $|f| = f^+ + f^-$. Let $f \in K^*$. From (15) $\sum_{k=1}^n f(\xi_k) = f(Z_t^{\triangle_{\varepsilon_n}})$ which is increasing as function of ε_n and bounded, by Lemma 13. Therefore $\sum_{k=1}^n f(\xi_k)$ converges a.s. □

Under the additional condition on the norm convergence of the jumps sum, the converse of Theorem 9 is true for proper normal cones.

THEOREM 16 *Let K be a proper normal cone of a separable Banach space B and let $\{Z_t : t \geq 0\}$ be a K-subordinator. If the jumps sum $Z_t^{\triangle_\varepsilon}$ converges a.s. in norm as $\varepsilon \downarrow 0$, then the characteristic functional of Z has the special Lévy-Khintchine representation (11) with generating triplet satisfying conditions (a)-(c) in Theorem 9. In particular, there exists a ν-Pettis centering I_ν such that $\gamma_0 = \gamma - I_\nu$.*

Proof. Step 1. We first prove that ν is concentrated on K. For some sequence of continuous linear functionals g_k we have $K = \cap_{k=1}^{\infty} \{x : g_k(x) \geq 0\}$. Then for each $k \geq 1$, the one dimensional subordinator $\{g_k(Z_t)\}$ has only nonnegative jumps since it is nonnegative. So if C is contained in the union $\cup_{k=1}^{\infty} \{x : g_k(x) < 0\}$ then $\nu(C) = 0$. Thus ν is concentrated on K.
Step 2. Next we show that the Gaussian part is zero. Let $\triangle_\varepsilon = \{x : \|x\| > \varepsilon\} \cap K$. By assumption $Z_t^{\triangle_\varepsilon}$ converges strongly to some Z_t^0 as $\varepsilon \downarrow 0$ almost surely. Therefore the process $\{Z_t - Z_t^0\}$ is continuous almost surely and

$$Ee^{if(Z_t - Z_t^0)} = exp\left\{-\frac{1}{2}(Af, f) + if(\gamma_0)\right\}. \tag{16}$$

Since $Z_t - Z_t^{\triangle_\varepsilon} \in K$ for all $\varepsilon > 0$ (Lemma 14) then $Z_t - Z_t^0 \in K$ by closedness of K. Hence for every $f^+ \in K^*$ the process $\{f^+(Z_t - Z_t^0)\}$ is nonnegative, continuous and Gaussian, therefore $var\left(f^+(Z_t - Z_t^0)\right) = (Af^+, f^+) = 0$. This gives $var\left(f(Z_t - Z_t^0)\right) = 0$ for any $f \in B^*$, which shows that the covariance operator A is null.

Step 3. We check the required form of the drift. The fact that $\gamma_0 \in K$ follows since $f(Z_t - Z_t^0) \geq 0$ for every $f \in K^*$ and from (16) we get $\gamma_0 \in K$. Moreover, using $A = 0$, (16) and the fact that (see (8))

$$Ee^{if(Z_t^0)} = \lim_{\varepsilon \downarrow 0} Ee^{if(Z_t^{\triangle_\varepsilon})} = exp\left\{t \int_K \left(e^{if(x)} - 1\right) \nu(dx)\right\}, \tag{17}$$

we obtain (11).

Next, let $f^+ \in K^*$. From (3), (17) and

$$Ee^{if^+(Z_t^0)} = \lim_{\varepsilon \downarrow 0} exp\left\{t \int_{\{x \in K : \varepsilon < \|x\| \leq 1\}} \left[e^{if^+(x)} - 1 - if^+(x)\right] \nu(dx)\right.$$
$$\left. + t \int_{\{x \in K : \|x\| > 1\}} \left(e^{if^+(x)} - 1\right) \nu(dx) + it \int_{\{x \in K : \varepsilon < \|x\| \leq 1\}} f^+(x) \nu(dx)\right\}$$

we have that $exp\left\{it \int_{\varepsilon < \|x\| \leq 1} f^+(x) \nu(dx)\right\}$ converges as $\varepsilon \downarrow 0$. This implies the convergence of the degenerate distribution at point $t \int_{\varepsilon < \|x\| \leq 1} f^+(x) \nu(dx)$

and consequently
$\int_{\varepsilon<\|x\|\leq 1} f^{+}(x)\nu(dx) \to \int_{D_0} f^{+}(x)\nu(dx)$ as $\varepsilon \downarrow 0$. Since ν is concentrated on K and K^* is a generating cone of B^*, (10) holds for every $f \in B^*$.

Finally, $I_\nu = \gamma - \gamma_0 \in B$ is a well defined ν-Pettis centering, since (10) holds and $\int_{D_0} f(x)\nu(dx) = f(\gamma - \gamma_0)$ for any $f \in B^*$ by the uniqueness of the generating triplet of the process Z. □

Lemma 14 and Theorem 16 yield the following result for subordinators with values in regular cones.

THEOREM 17 *Let K be a proper regular cone of a separable Banach space B and let $\{Z_t : t \geq 0\}$ be a K-subordinator. Then*

a) The jump process $Z_t^{\triangle_\varepsilon}$ in (14) is always norm convergent a.s.

b) If K is normal, Z has the special Lévy-Khintchine representation.

3. Regular subordinators

Given a proper cone K of a separable Banach space B, a K-subordinator $\{Z_t : t \geq 0\}$ is called a *regular subordinator* in K or a *K-valued regular subordinator* if it has the special Lévy-Khintchine representation. In the case of LK-cones every subordinator is regular.

THEOREM 18 *Let K be an LK-cone of a separable Banach space B. A B-valued Lévy process is a K-subordinator if and only if it is a regular subordinator.*

Proof. Assume the special Lévy-Khintchine representation, then by Theorem 9 the process Z is K-valued. Conversely, if Z is a K-valued Lévy process then the assertion follows from Proposition 2 and Theorem 17 (b). □

As a special case of Theorem 18 we recover a result formulated in [9, Cor. p. 278], that considered a restricted class of cones. Here we give a rigorous proof of this result by proving that the underlying cone is an LK-cone.

PROPOSITION 19 *Let B be a separable Banach space with a proper cone K such that there is a continuous linear functional f_0 with the property that $k_0 = \inf_{x\in K, \|x\|=1} f_0(x) > 0$.*

A B-valued Lévy process has the special Lévy-Khintchine representation if and only if it is a K-valued process.

Proof. We first observe that for $0 \neq x \in K$, $0 < k_0 \leq f_0(x/\|x\|)$ and therefore $\|x\| \leq k_0^{-1} f_0(x), \qquad x \in K$.

For any nonzero continuous linear functional f on B, define $f_1(x) = f(x) + \|f\|^* k_0^{-1} f_0(x)$ and $f_2(x) = \|f\|^* k_0^{-1} f_0(x)$ for all $x \in B$. Then f_1, $f_2 \in K^*$ and $f_1 - f_2 = f$. Indeed, observe that f_2 is nonnegative on K since $f_0(x) > 0$ for $x \in K$ and using the above inequality we deduce that f_1 is also nonnegative on K. Thus K^* is a generating cone for B^* and hence K is normal.

Let $\sum_{k=1}^{\infty} x_k$ be any w.u.C. series of elements in K and let $s_n = \sum_{k=1}^{n} x_k$. Then $\sum_{k=1}^{\infty} |f_0(x_k)|$ is finite and for $n > m$, $s_n - s_m \in K$ and therefore $\|s_n - s_m\| \leq k_0^{-1} f_0(s_n - s_m)$. Then (s_n) is norm convergent and hence from Proposition 1, K does not contain any subcone isomorphic to c_0^+. Then, K is an LK-cone and the result now follows from Theorem 18. □

As a consequence of Propositions 6 and 11, infinitely divisible random elements in LK-cones are characterized by a special form of their Laplace transform. The next result may be thought as the Banach space analogue of the well known characterization of the Laplace transform of a real nonnegative infinitely divisible random variable in [8, Th. 13.7.2]. It is proved in [6] for normal and regular cones of general ordered vector spaces.

COROLLARY 20 *Let K be an LK-cone of a separable Banach space B. In order that a K-valued random variable S have an infinitely divisible law it is necessary and sufficient that S has the Laplace transform*

$$Ee^{-f(S)} = exp\left\{\int_K \left(e^{-f(x)} - 1\right)\nu(dx) - f(\gamma_0)\right\} \qquad f \in K^*$$

where $\gamma_0 \in K$ and the Lévy measure ν satisfies (10).

Proof. Suppose S is an infinitely divisible random variable in K and let $\{Z_t : t \geq 0\}$ be the associated K-subordinator such that S and Z_1 have the same law. The process Z is a K-regular subordinator by Theorem 18. Then from Proposition 11 Z has the Laplace transform (12). The assertion follows by taking $t = 1$ in (12). The converse is immediate. □

Acknowledgements

We thank Ken-iti Sato for many useful suggestions.

References

[1] Araujo, A. and Giné, E. (1980) *The Central Limit Theorem for Real and Banach Valued Random Variables*, Wiley, New York.

[2] Barndorff-Nielsen, O. E., Pedersen, J. and Sato, K. (2001) Multivariate subordination, self-decomposability and stability, *Adv. Appl. Probab.* **33**, 160–187.

[3] Bertoin, J. (1996) *Lévy Processes,* Cambridge University Press, Cambridge.

[4] Bertoin, J. (1999) Subordinators: examples and applications, *Lectures Notes in Mathematics* **1717**, Springer, Berlin, 1–91.

[5] Bochner, S. (1955) *Harmonic Analysis and the Theory of Probability,* University of California Press, Berkeley and Los Angeles.

[6] Dettweiler, E. (1976) Infinitely divisible measures on the cone of an ordered locally convex vector space, *Ann. Sci. Univ. Clermont* **14**, 61, 11-17.

[7] Diestel, J. (1984) *Sequences and Series in Banach Spaces*, Springer, New York.

[8] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications*, Vol. II, Second Edition, Wiley, New York.

[9] Gihman, I. I. and Skorohod, A. V. (1975) *The Theory of Stochastic Processes*, Vol. II, Springer, Berlin.

[10] Kamthan, P. K. and Gupta, M. (1985) *Theory of Bases and Cones*, Pitman Advances Publishing Proceedings, Research Notes in Mathematics **117**.

[11] Linde, W. (1986) *Probability in Banach Spaces -Stable and Infinitely Divisible Distributions-*, John Wiley & Sons, Berlin.

[12] Musial K. (2002) Pettis integral, *Handbook of Measure Theory, Vol I*, E. Pap (Ed.). North-Holland, Amsterdam. pp531-586.

[13] Pedersen, J. and Sato, K. (2003) Cone-parameter convolution semigroups and their subordination, *Tokyo J. Math.* **26**, 503-525.

[14] Pedersen, J. and Sato, K. (2004) Relations between cone-parameter Lévy processes and convolution semigroups, *J. Math. Soc. Japan* ***56***, 541-559.

[15] Pérez-Abreu, V. and Rocha-Arteaga, A. (2005) Covariance-parameter Lévy processes in the space of trace-class operators, *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **8**, 1-22.

[16] Rocha-Arteaga, A. and Sato, K. (2003) *Topics in Infinitely Divisible Distributions and Lévy Processes*, Aportaciones Matemáticas, Sociedad Matemática Mexicana, Investigación **17**.

[17] Sato, K. (1999) *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, Cambridge.

[18] Skorohod, A. V. (1991) *Random Processes with Independent Increments*, Kluwer Academic Publisher, Dordrecht, Netherlands (Russian original 1986).

V. Pérez-Abreu
pabreu@cimat.mx
Centro de Investigación en Matemáticas
Apdo. Postal 402, Guanajuato 36000, Mexico.

A. Rocha-Arteaga
arteaga@uas.uasnet.mx
Escuela de Ciencias Físico-Matemáticas
Universidad Autónoma de Sinaloa
Culiacan, Mexico.