UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

PROCESOS DE LÉVY, PROCESOS DE MARKOV Y CONVERGENCIA DÉBIL EN ESPACIOS DE SKOROKHOD

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
ALEJANDRO ROSALES ORTIZ

DIRECTORA:
DRA. MARÍA EMILIA CABALLERO ACOSTA
INSTITUTO DE MATEMÁTICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

CIUDAD DE MÉXICO, JUNIO 2019.

Procesos de Lévy, procesos de Markov y convergencia débil en espacios de Skorokhod

Alejandro Rosales Ortiz
Directora: María Emilia Caballero Acosta.

A mis papás, Laura y Ernesto, a mi hermana Laura y a María Emilia. Una vez más, gracias por el cariño y el apoyo. No podría estar más agradecido.

Este trabajo no habría sido posible sin el curso de procesos de Lévy que impartió la profesora María Emilia en el posgrado. Uno de los objetivos fue recopilar parte de la teoría que se estudió (y que a mi parecer es sumamente interesante) para que pueda ser accesible a otros estudiantes interesados por el tema. Este material sentó las bases que me permitieron profundizar en la teoría de procesos de Markov, convergencia débil de procesos así como otros aspectos de la teoría de los procesos de Lévy que se pueden encontrar en esta tesis. Es por esto que el trabajo resultó extenso: se trata en esencia de un trabajo en equipo.

Quisiera agradecer al posgrado de la UNAM y en particular a Silvia, María Inés, María Teresa y Lucía por el apoyo y comprensión a la hora de realizar trámites así como por su cálido trato. Para mi fue un gusto formar parte del posgrado en matemáticas. A los profesores Ana Meda, Sandra Palau, Adrián González y Ramsés Mena por haber aceptado ser mis sinodales. Agradezco mucho sus correcciones así como el tiempo que le dedicaron a revisar el trabajo.

Finalmente, agradezco a CONACYT la beca que recibí a lo largo de estos dos años.

Prefacio

El objetivo de este trabajo consiste en presentar, de forma accesible y lo más auto-contenida posible, una breve introducción a tres temas fundamentales en el estudio de la teoría de los procesos estocásticos: los procesos de Lévy, los procesos de Markov y la convergencia débil de sucesiones de procesos estocásticos. Como se hará notar con frecuencia, estos tres se encuentran fuertemente ligados. En el segundo semestre de mis estudios de maestría, cursé una materia impartida por la profesora María Emilia Caballero en la cual estudiamos la convergencia débil de medidas, las distribuciones infinitamente divisibles y los procesos de Lévy. Bajo supervisión de la profesora Caballero, al finalizar el curso inicié con la redacción del material estudiado durante el semestre, en vista de que la poca literatura disponible suele ser poco accesible. Algunas inquietudes que surgieron conforme fuimos avanzando me llevaron al estudio de la teoría de los procesos de Markov y de Feller, que es el contenido de los capítulos 5 y 6. El objetivo se convirtió en realizar un estudio a profundidad de los procesos de Markov, abordar problemas de existencia, completación de filtraciones y existencia de buenas versiones en espacios adecuados y presentar a los procesos de Lévy como un ejemplo fundamental de esta rica clase de procesos. De esto último trata el capítulo 7: en este se encuentran por un lado, la teoría de los procesos de Lévy estudiada a lo largo de los primeros cuatro capítulos con, por otro lado, la teoría general de los procesos de Markov y Feller. Cerramos con el capítulo 8, en el cual se estudia la convergencia débil de sucesiones de procesos estocásticos al dotar al espacio de las trayectorias cadlag, al cual denotaremos por D, con la topología J1 de Skorokhod. Una vez introducidas las herramientas adecuadas, se establecerán criterios para determinar convergencia de medidas en D y se probarán dos teoremas propios a las familias de procesos de Feller y de procesos de Lévy respectivamente. La tesis se puede dividir en tres grandes bloques:

I- El primer tema de estudio son los procesos de Lévy y las distribuciones infinitamente divisibles: de esto tratan en esencia los primeros cuatro capítulos. Los preliminares se encuentran distribuidos en los dos primeros: el capítulo 1 aborda cuestiones técnicas de la teoría de procesos que se usarán constantemente a lo largo del trabajo mientras que en las secciones 2.1 y 2.3 se hace un repaso breve de la convergencia débil de sucesiones de medidas en $\mathbb{R}$ que será necesario para el estudio de las distribuciones infinitamente divisibles de la sección 2.2, tema centrar del capítulo 2. La convergencia débil en espacios métricos se introduce en la sección 2.3 y únicamente se dan como introducción a lo que haremos en el capítulo 8. Los tres resultados clave del capítulo 2 son el teorema 2.2.7, el teorema 2.2.8, y el teorema 2.2.11. Estos caracterizan a la familia de distribuciones infinitamente divisibles a través de su función característica. El capítulo 3 está dedicado al estudio de las medidas aleaotrias de Poisson (MAP). Trataremos el problema de su existencia para luego estudiar el conecpto de integral de una función con respecto a un MAP. Cerramos estableciendo criterios de integrabilidad con respecto a MAP en espacios métricos y estudiamos la naturaleza del objeto que se obtiene al integrar una función con respecto a una medida aleatoria de Poisson. A grandes rasgos, en esto consiste el teorema de Campbell. Finalmente, con estas herramientas podremos abordar el capítulo 4, en el cual se estudia a los procesos de Lévy. Los resultados del capítulo 3 sobre distribuciones infinitamente divisibles permiten probar la descomposición de Lévy-Kintchine al inicio de este capítulo. Esta, junto con la teoría de medidas aleatorias de Poisson y la teoría de las martingalas cuadrado integrables (la cual se puede hallar en el apéndice), nos llevan a la descomposición de Lévy-Itô en la sección 4.2. No se puede pasar por alto lo rica que resulta la teoría que rodea a los procesos de Lévy. Se finaliza este capítulo con algunas aplicaciones a la variación total de las trayectorias y una breve introducción a los subordinadores.

II- Los procesos de Lévy forman parte de una gran familia de procesos, conocidos como los procesos de Markov. Nos dedicamos al estudio de estos a lo largo de los capítulos 5 y 6. En el capítulo 5 abordamos cuestiones de existencia, introducimos el concepto de un sistema de Markov y probamos la propiedad de Markov para estos en espacios canónicos. En el capítulo 5 nos restringimos a una familia de procesos de Markov, conocidos como los procesos de Feller. Una vez introducidas sus propiedades básicas, probaremos la existencia de versiones con trayectorias cadlag. Este resultado se conoce como el teorema de regularidad de procesos de Feller y una vez más, nos apoyaremos en la teoría de martingalas incluida en el apéndice. Veremos que al trabajar con esta versión, al completar la filtración natural obtenemos una filtración continua por la derecha, y por lo tanto podemos trabajar bajo las llamadas condiciones habituales. Finalizamos el capítulo 6 con la prueba de la propiedad fuerte de Markov y algunas aplicaciones como la casi-continuidad por la izquierda. La propiedad fuerte de Markov será una herramienta fundamental a lo largo del capítulo 7 así como en la prueba de los teoremas de la sección 8.6, mientras que la casi-continuidad por la izquierda nos permitirá extender algunos resultados de convergencia a familias de procesos de Feller en la sección 8.4. Se concluye el estudio de los procesos de Markov con el capítulo 7, al combinar lo estudiado en los dos capítulos previos, con la teoría desarrollada para los procesos de Lévy en la primera parte. Estudiaremos a los procesos de Lévy como una familia fundamental dentro de los procesos de Feller y veremos resultados que le son propios. Empezaremos construyendo el sistema de Markov asociado a un proceso de Lévy en el espacio canónico de las trayectorias cadlag con el cual trabajaremos de ahí en adelante. Daremos dos construcciones posibles, una en la cual se utiliza la teoría de los procesos de Markov y Feller, y otra que no requiere de esta aunque es un tanto más laboriosa. Este es esencialmente el contenido de la sección 7.1. Luego, bajo este marco, probaremos resultados propios a los procesos de Lévy en $\mathbb{R}^d$ que combinan su carácter markoviano con la independencia y estacionariedad de sus incrementos para finalmente darnos a la tarea de estudiar propiedades asintóticas: veremos que podemos clasificar a los procesos de Lévy como recurrentes o transitorios y que bajo hipótesis adecuadas, podemos caracterizar su comportamiento asintótico. Para probar estos resultados, haremos uso de la herramienta desarrollada en el capítulo 6 para los procesos de Feller. Esta es una de las secciones fundamentales de este trabajo ya que una demostración detallada como la que aquí se presenta no se encuentra en la literatura y además nos habla del comportamiento de los procesos de Lévy. Cabe notar que este tipo de resultados se pueden obtener fácilmente para el movimiento browniano usando herramientas de cálculo estocástico.

III- Como este trabajo es fundamentalmente, en parte, un trabajo sobre convergencia débil, se quiso introducir la convergencia débil de procesos, que es una generalización por un lado natural y por otro sorprendente y poderosa del concepto de convergencia débil de medidas en $\mathbb{R}$. Por esto, dedicamos el capítulo 8 a este tema. Iniciamos con un primer ejemplo de convergencia débil: al reescalar una caminata aleatoria se obtiene una sucesión de procesos que convergen al movimiento browniano. Par demostrar este resultado, conocido como el principio de invariancia de Donsker, se hace uso únicamente de resultados de convergencia débil probados en el segundo capítulo. La prueba es larga y compleja lo que nos lleva a desarrollar herramientas más sofisticadas para abordar este tipo de problemas. A esto dedicaremos el resto del capítulo: empezamos con un estudio a profundidad de la convergencia débil de medidas en espacios métricos, lo que generaliza lo hecho en el capítulo 2. Dotaremos al conjunto de medidas en un espacio métrico E con una métrica, conocida como la métrica de Prohorov y probaremos la equivalencia entre tensión de sucesiones de medidas y compacidad relativa cuando E sea un espacio métrico completo y separable. Este resultado es llamado el teorema de Prohorov y será fundamental en lo que resta del capítulo. A continuación, veremos que es posible metrizar al espacio D a modo de que sea un espacio métrico completo y separable. Así, será posible aplicar el teorema de Prohorov a sucesiones de medidas en D, que en nuestro caso corresponderán a las leyes de procesos con trayectorias cadlag. La topología que esta métrica induce en D es llamada la topología J1 de Skorokhod y estudiaremos algunas de sus propiedades fundamentales. Finalmente, regresamos al problema de convergencia débil de medidas en D:
(i) Estableceremos algunos criterios de convergencia para sucesiones de medidas en D. Veremos que tensión + convergencia de distribuciones finito dimensionales permite concluir convergencia débil de la sucesión.
(ii) Será evidente la importancia de establecer criterios que permitan comprobar tensión de colecciones de medidas en D y por esto dedicamos la sección 8.5 a desarrollar dichos criterios. En particular se prueba el criterio de tensión de Aldous.
(iii) Para concluir, aplicamos todo lo estudiado al probar dos teoremas de convergencia: primero uno que caracteriza la convergencia de procesos de Feller a partir de sus semigrupos y un segundo propio a los procesos de Lévy, que lo logra a partir de su exponente característico.

Contenido

Capítulo 1

Preliminares de procesos estocásticos

Sea $(\Omega, \mathscr{F}, \mathbb{P})$ es un espacio de probabilidad y $(E, \mathscr{E})$ un espacio medible. Denotaremos por $E^T = \{x\,;\, x(t) \in E \text{ para cada } t \in T\}$ al espacio de las funciones con valores en E indexadas por $t \in T$, donde T será $\mathbb{R}^+$ o un subconjunto de este. Heurísticamente, un proceso estocástico X definido en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ es una función aleatoria, esto es, para cada ω, $X(\omega)$ es un elemento de E^T. Formalmente, es entonces natural definir un proceso estocástico en $(\Omega, \mathscr{F}, \mathbb{P})$ con valores en E como una función medible $X : \Omega \to E^T$. Esto es, para cada ω, $X_t(\omega) = x(t)$ para toda t para alguna función $x \in E^T$, también llamada la trayectoria de X asociada a esa ω. Sin embargo, obsérvese que para poder hablar de medibilidad de X, es necesario antes definir una σ-álgebra en E^T; resolvemos esto a continuación:

1.1 La σ-álgebra de los cilindros en E^T

Referencia: Spieksma [3] sección 2 y Kallenberg [12] capítulos 1 y 2.

Recordemos la definición de $\sigma-$álgebra $\mathscr{E}^n = \otimes_{i\leq n}\mathscr{E}_i$ en el espacio producto E^n con $n = 1, 2, \ldots, \infty$.

Definición 1.1.1 *Si n es finito, la σ-álgebra en E^n es la generada por los rectángulos finito dimensionales:*

$$\mathscr{E}^n = \sigma\left(A_1 \times \cdots \times A_n\,;\, A_1 \ldots A_n \in \mathscr{E}\right)$$

mientras que en E^∞ se define como

$$\mathscr{E}^\infty = \sigma\left(A_1 \times \cdots \times A_n \times E \times E \times \ldots\,;\, A_1, \ldots, A_n \in \mathscr{E} \text{ y } n \in \mathbb{N}\right).$$

Denotaremos por $\mathcal{R}_n^\infty$ a la colección de conjuntos de la forma $A_1 \times \cdots \times A_n \times E \times E \times \ldots$ y por $\mathcal{R}_n$ a los de la forma $A_1 \times \cdots \times A_n$ con $A_1, \ldots, A_n \in \mathscr{E}$ de modo que $\mathscr{E}^n = \sigma(\mathcal{R}_n)$ y $\mathscr{E}^\infty = \sigma(\mathcal{R}_n^\infty\,;\, n \in \mathbb{N})$.

Definición 1.1.2 *Para $t \in T$, definimos a la proyección $\pi_t : E^T \to E$ como $\pi_t(x) = x_t$, también conocida como la función valuación al tiempo t.*

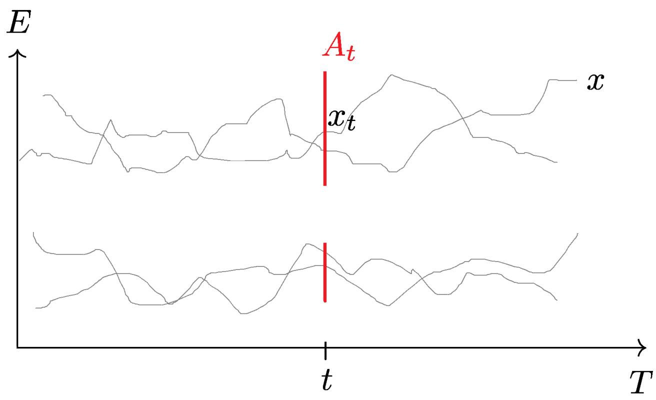


Figure 1.1: Representación del conjunto $\{x \in E^T\,;\, x_t \in A_t\}$

La σ-álgebra en E^T que nos servirá no es más que la generada por la familia de proyecciones $(\pi_t)_{t\in T}$:

Definición 1.1.3 *Un rectángulo 1-dimensional en E^T es un subconjunto de E^T de la forma:*

$$S = \{x \in E^T\,;\, x_t \in A_t\} = \pi_t^{-1}(A_t) \qquad \textit{para } t \in T \textit{ fijo y } A_t \in \mathscr{E}$$

Definimos la $\sigma-$álgebra $\mathscr{E}^T$ en E^T como la generada por los rectángulos 1-dimensionales:

$$\mathscr{E}^T = \sigma\,(S\,;\, S \textit{ rectángulo 1-dimensional }) = \sigma\,(\pi_t\,;\, 0 \le t \le T)$$

Por construcción, es la mínima $\sigma-$álgebra que hace a las proyecciones π_t medibles para todo t y por lo tanto también se puede definir simplemente como la σ-álgebra generada por las proyecciones $\pi_t : x \to x(t)$. El objetivo será caracterizar a los conjuntos en $\mathscr{E}^T$ y ver algunas propiedades sobre su estructura que serán de utilidad.

Definición 1.1.4 *Un rectángulo finito dimensional en E^T es un subconjunto de E^T de la forma:*

$$R = \{x \in E^T\,;\, (x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n\} = \bigcap_{i=1}^{n} \pi_{t_i}^{-1}(A_i) \qquad t_1, \ldots, t_n \in [0, T]$$
$$A_1, \ldots, A_n \in \mathscr{E},$$

para $n \in \mathbb{N}$. Denotaremos por $\mathcal{C}$ a la colección de todos los rectángulos finito dimensionales R.

Está claro que los rectángulos finito dimensionales se obtienen como intersecciones finitas de rectángulos 1-dimensionales y por lo tanto no solo están en $\mathscr{E}^T$ sino que la generan pues contienen a rectángulos 1-dimensionales. A diferencia de estos, la colección $\mathcal{C}$ es un $\pi-$sistema al ser estable ante intersecciones finitas; esta observación será de gran utilidad para extender propiedades a toda la σ-álgebra $\mathscr{E}^T$ usando el lema de clases monótonas.

La definición 1.1.2 se extiende naturalmente a colecciones a lo más numerables de tiempos: Si $t_1, t_2, \ldots t_n \in T$, la proyección $\pi_{(t_1,\ldots,t_n)} : E^T \to E^n$ se define como

$$\pi_{(t_1,\ldots,t_n)}(x) = (x_{t_1}, \ldots, x_{t_n}) \tag{1.1.1}$$

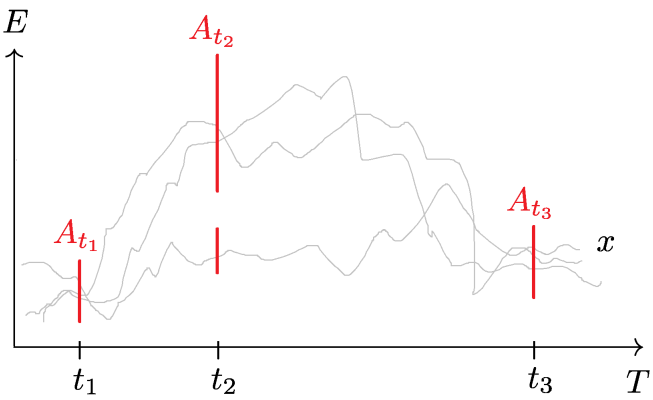


Figure 1.2: Representación del conjunto $\{x \in E^T\,;\,(x_{t_1},\ldots,\,x_{t_n}) \in A_1 \times \ldots \times A_n\}$

Aquí, $\pi_{(t_1,\ldots,t_n)}$ es $\mathscr{E}^T/\mathscr{E}^n$-medible pues $\sigma(\mathcal{R}_n) = \mathscr{E}^n$ y si $R_n \in \mathcal{R}_n$, $\pi^{-1}_{(t_1,\ldots,t_n)}(R_n)$ es justamente un rectángulo finito dimensional (de dimensión n). En el caso numerable, la definición es la misma para el vector $(t_1, t_2, \ldots)$ y la $\mathscr{E}^T/\mathscr{E}^\infty$medibilidad de $\pi_{(t_1,t_2,\ldots)} : E^T \to E^\infty$ también se tiene pues $\mathscr{E}^\infty = \sigma(\mathcal{R}^\infty_n\,; n \in \mathbb{N})$ y si $R \in \mathcal{R}^\infty_n$ para alguna n, $\pi^{-1}_{(t_1,\ldots,t_n)}(R)$ es también un rectángulo finito dimensional (de dimensión n).

Definición 1.1.5 *Diremos que $D \in \mathscr{E}^T$ es un cilindro finito-dimensional si existe $n \in \mathbb{N}$ y una colección finita $t_1, \ldots, t_n \in T$ y $B \in \mathscr{E}^n$ tal que*

$$D = \{x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in B\} = \pi^{-1}_{(t_1,\ldots,t_n)}(B).$$

Si la colección $t_1, t_2, \ldots$ es a lo más numerable y $B \in \mathscr{E}^\infty$, entonces llamaremos a

$$D = \{x \in E^T\,;\,(x_{t_1}, x_{t_2}, \ldots) \in B\} = \pi^{-1}_{(t_1,t_2,\ldots)}(B)$$

un $\sigma-$cilindro.

Nótese que los rectángulos finito dimensionales son en particular cilindros finito-dimensionales con $B = A_1 \times \cdots \times A_n$. Resulta que podemos caracterizar exactamente a la σ-álgebra $\mathscr{E}^T$:

Proposición 1.1.6 *$\mathscr{E}^T$ consiste justamente en la colección de $\sigma-$cilindros.*

Demostración
Primero, probaremos la siguiente contención e igualdad:

$$\{D\,; D \text{ es } \sigma\text{-cilindro}\} \subset \sigma\,(D\,;\, D \text{ es cilindro f.d}) \quad \text{y} \quad \sigma\,(D\,;\, D \text{ es cilindro f.d}) = \mathscr{E}^T. \quad (1.1.2)$$

La igualdad se sigue de lo siguiente:
Sea $D = \{x\,;\,(x_{t_1}, \ldots, x_{t_n}) \in B\} = \pi^{-1}_{(t_1,\ldots,t_n)}(B)$ cilindro finito dimensional con $B \in \mathscr{E}^n$. Como $B \in \sigma\,(\mathcal{R}_n)$ entonces

$$\begin{aligned}
\pi^{-1}_{(t_1,\ldots,t_n)}(B) \in \pi^{-1}_{(t_1,\ldots,t_n)}\{\sigma(\mathcal{R}_n)\} &= \sigma\left(\pi^{-1}_{(t_1,\ldots,t_n)}\{\mathcal{R}_n\}\right) \\
&= \sigma\left(x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \cdots \times A_n\,;\, A_1, \ldots A_n \in \mathscr{E}\right) \\
&= \mathscr{E}^T
\end{aligned}$$

y entonces $D \in \mathscr{E}^T$. Se sigue que $\sigma(D\,;\,D \text{ cilindro f.d.}) \subset \mathscr{E}^T$; como los rectángulos 1-dimensionales son cilindros f.d., entonces se tiene la igualdad.
Probaremos ahora la contención. Sea $D = \{x\,;\,(x_{t_1}, x_{t_2}, \dots) \in B\} = \pi^{-1}_{(t_1,t_2,\dots)}(B)$ un $\sigma-$cilindro para algún $B \in \mathscr{E}^\infty$ y $t_1, t_2, \dots$ colección a lo más numerable. Para probar la contención en (7.3.3), obsérvese que si $B \in \mathscr{E}^\infty$, entonces por definición $B \in \sigma(\mathcal{R}^\infty_n\,;\,n \in \mathbb{N})$ y similarmente a lo que se hizo antes:

$$\begin{aligned}
\pi^{-1}_{(t_1,t_2,\dots)}(B) \in \pi^{-1}_{(t_1,t_2,\dots)}\{\sigma(\mathcal{R}^\infty_n\,;\,n \in \mathbb{N})\} &= \sigma\left(\pi^{-1}_{(t_1,t_2,\dots)}\{\mathcal{R}^\infty_n\,;\,n \in \mathbb{N}\}\right) \\
&\subset \sigma\left(\pi^{-1}_{(s_1,s_2,\dots)}\{\mathcal{R}^\infty_n\,;\,n \in \mathbb{N}\}\, s_1, s_2, \dots \in [0,T]\right) \\
&= \sigma(R\,;\,R \text{ rectángulo f.d.}) \\
&= \mathscr{E}^T
\end{aligned}$$

por lo que se tiene (7.3.3). Como los rectángulos 1-dimensionales son $\sigma-$cilindros, entonces $\sigma(D\,;\,D \text{ es } \sigma-\text{cilindro}) = \mathscr{E}^T$. Bastaría entonces probar que el conjunto de $\sigma-$cilindros es una σ-álgebra.
E^T es un un $\sigma-$cilindro pues $E^T = \{x\,;\,x_T \in E\}$. Además, si $D = \pi^{-1}_{(t_1,t_2,\dots)}(B)$ para algún $B \in \mathscr{E}^\infty$, $D^c = \pi^{-1}_{(t_1,t_2,\dots)}(B^c)$ que también es un $\sigma-$cilindro por lo que el conjunto es cerrado bajo complementos; faltaría ver que lo es bajo intersecciones numerables. Sean

$$D_n = \{x \in E^T\,;\,(x_{t^n_1}, x_{t^n_2}, \dots) \in B_n\} \qquad \begin{array}{l} t^n_1, t^n_2 \dots \in [0,T] \\ D_n \in \mathscr{E}^\infty \end{array}$$

$\sigma-$cilindros para $n \in \mathbb{N}$. Si $S = \cup_{n=1}^\infty \{t^n_1, t^n_2, \dots\}$, sea $f : S \to \mathbb{N}$ biyectiva, de modo que $S = \{f^{-1}(1), f^{-1}(2), \dots\}$. La posición de un elemento t^n_k dentro de S está dictada por el valor $f(t^n_k)$. Reescribiremos a cada D_n en términos de f; para ello, definimos

$$\tilde{B}_n = \{(a_1, a_2, \dots) \in E^\infty\,;\,(a_{f(t^n_1)}, a_{f(t^n_2)}, \dots) \in B_n\}$$

de modo que

$$D_n = \{x \in E^T\,;\,(x_{f^{-1}(1)}, x_{f^{-1}(2)}, \dots) \in \tilde{B}_n\}$$

y por lo tanto:

$$\bigcap_{n=1}^\infty D_n = \{x \in E^T\,;\,(x_{f^{-1}(1)}, x_{f^{-1}(2)}, \dots) \in \cap_n \tilde{B}_n\}$$

que es un $\sigma-$cilindro. ■

1.2 Procesos Estocásticos

Referencias: Kallenberg [12], capítulos 1, 2 y 3.

Por el resto de la sección, $(\Omega, \mathscr{F}, \mathbb{P})$ es un espacio de probabilidad.

1.2.1 Medibilidad

Ahora que $(E^T, \mathscr{E}^T)$ tiene estructura de espacio medible podemos hablar de funciones medibles de $(\Omega, \mathscr{F}, \mathbb{P})$ en $(E^T, \mathscr{E}^T)$. La colección de estos son los procesos estocásticos E-valuados con trayectorias en E^T.

De saber más sobre las propiedades de las trayectorias de X, nos podemos restringir a un subespacio de E^T con más estructura:

Sea U un subconjunto de E^T, medible o no. Si $X : \Omega \to U \subset E^T$, equipamos al espacio U con la σ-álgebra de los cilindros restringida a elementos de U: $U \cap \mathscr{E}^T = \{U \cap B\,;\, B \in \mathscr{E}^T\}$ y que ocasionalmente denotaremos $\mathscr{E}^T_U$. En este caso, la medibilidad de X como función en $(U, U \cap \mathscr{E}^T)$ es equivalente a medibilidad como función en $(E^T, \mathscr{E}^T)$ y $U \cap \mathcal{C}$ es un π-sistema que genera a $U \cap \mathscr{E}^T$ pues $\sigma(U \cap \mathcal{C}) = U \cap \sigma(\mathcal{C}) = U \cap \mathscr{E}^T$. Similarmente, $U \cap \mathscr{E}^T$ es generada por los rectángulos 1-dimensionales restringidos a elementos en U o equivalentemente, por los mapeos de evaluación $\pi_t : U \to E$, $t \in T$. Nos seguiremos refiriendo a ellos como los rectángulos finito/uno-dimensionales respectivamente sin especificar la restricción a U ya que debería seguirse del contexto. Cuando E sea un espacio métrico, (en nuestro caso se tratará de $\mathbb{R}^d$ o un espacio polaco) será de especial interés considerar U el espacio de las funciones continuas en E denotado $C(\mathbb{R}^+, E)$ o el de las continuas por la derecha con límites por la izquierda $D(\mathbb{R}^+, E)$ también llamadas cadlag.

Uno de los principales problemas de trabajar en $(E^T, \mathscr{E}^T)$ es que es un espacio con poca estructura y la σ-álgebra es *muy pobre*: por ejemplo, al utilizar la caracterización que se acaba de probar para los conjuntos en $\mathscr{E}^T$, veamos que $C(\mathbb{R}^+, E)$ y $D(\mathbb{R}^+, E)$ no son subconjuntos medibles de E^T. Para ver esto, bastaría notar que en un sigma-cilindro $D \in \mathscr{E}^T$ siempre habrá funciones que no son continuas por la derecha. Sea

$$D = \{x \in E^T\,;\, (x_{t_1}, x_{t_2}, \dots) \in B\} = \pi^{-1}_{(t_1, t_2, \dots)}(B)$$

arbitrario. Como $\{t_1, t_2 \dots\}$ es a lo más numerable, considerar s_0 y $\{s_1, s_2, \dots\} \subset \mathbb{R}^+ \cap \{t_1, t_2 \dots\}^c$ sucesión tal que $s_n \downarrow s_0$. Sea $b = (b_1, b_2, \dots) \in B$ y definimos a $x \in E^T$ como $x(t) = b_i$ si $t = t_i$ y $x(t) = (-1)^i$ si $t = s_i$. El resto de sus valores se puede definir arbitrariamente. Es claro que $x \in D$ y que por otro lado x no es continua en s_0. Parte del problema está en que en E^T están *todas* las trayectorias de T en E, incluyendo las no Lebesgue-medibles. Es razonable querer trabajar en un espacio con más estructura siempre que sea posible.

Es entonces natural definir un proceso estocástico continuo (resp. cadlag) con valores en un espacio métrico E como una función medible de un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ en $C(\mathbb{R}^+, E)$ (resp. $D(\mathbb{R}^+, E)$) equipado con la σ-álgebra $\mathscr{E}^T_C$ (resp.$\mathscr{E}^T_D$). La definición que dimos de proceso U-valuado es equivalente a la siguiente condición:

Proposición 1.2.1 *Una función $X : \Omega \to U \subseteq E^T$ es $U \cap \mathscr{E}^T$-medible si y solo si $X_t : \Omega \to E$ es $\mathscr{E}$-medible para cada $t \in T$.*

Demostración

Obsérvese que $X_t = \pi_t \circ X$.

Al tomar valores en U, la medibilidad de X como función con valores en $(U, U \cap \mathscr{E}^T)$ es equivalente a medibilidad en $(E^T, \mathscr{E}^T)$ así que consideramos esta última. Primero, si X es $\mathscr{F}/\mathscr{E}^T$-

medible, como π_t es $\mathscr{E}^T/\mathscr{E}$ la composición es $\mathscr{E}$ medible para cada t y se sigue que X_t es $\mathscr{E}$ medible para cada t.

Ahora, supongamos que para todo $t \in T$, X_t es $\mathscr{E}$-medible: entonces, para todo $t \in T$ y $A \in \mathscr{E}$, $(X_t)^{-1}(A) \in \mathscr{F}$ es decir, $X^{-1}(\pi_t^{-1}(A)) \in \mathscr{F}$ para todo $t \in T$ y $A \in \mathscr{E}$. Como $\{\pi_t^{-1}(A); t \in T$ y $A \in \mathscr{E}\}$ es el conjunto de los rectángulos 1-dimensionales, entonces tenemos que $X^{-1}(S) \in \mathscr{F}$ para todo S rectángulo 1-dimensional. Pero como $\sigma(S\,;\, S$ rectángulo 1-dimensional$) = \mathscr{E}^T$, entonces X es $\mathscr{E}^T$-medible. ■

Esto nos lleva a la siguiente definición de proceso estocástico, más manejable ya que en principio no se necesita mencionar a la σ-álgebra de los cilindros y que por la proposición anterior es equivalente a la primera que se mencionó

Definición 1.2.2 *Un proceso $X = (X_t)_{t\in T}$ con trayectorias en $U \subseteq \mathscr{E}^T$ es una colección de variables aleatorias con valores en E es decir, X_t es $\mathscr{E}$-medible para cada $t \in T$ y tal que $(X_t(\omega))_{t\in T} \in U$ para cada $\omega \in \Omega$.*

Recordemos que la σ-álgebra generada por X denotada $\mathscr{F}^0_\infty = \sigma(X)$ es la mínima que hace a X medible. Además, está dada por:

$$\mathscr{F}_0^\infty = X^{-1}(U \cap \mathscr{E}^T) = \{\{X \in B\} : B \in U \cap \mathscr{E}^T\}. \tag{1.2.1}$$

Nótese que no depende de si pensamos a X como función en $(U, U \cap \mathscr{E}^T)$ o en $(E^T, \mathscr{E}^T)$ ya que

$$\mathscr{F}^0_\infty = X^{-1}(U \cap \mathscr{E}^T) = X^{-1}(U) \cap X^{-1}(\mathscr{E}^T) = \Omega \cap X^{-1}(\mathscr{E}^T) = X^{-1}(\mathscr{E}^T)$$

Por lo tanto, sin ambiguedad, los elementos en $\mathscr{F}^0_\infty$ son de la forma $\{X \in D\}$ con D un σ-cilindro. Como el conjunto de los rectángulos 1-dimensionales $\mathcal{R}$ genera a $\mathscr{E}^T$, entonces

$$\begin{aligned}\mathscr{F}^0_\infty = X^{-1}(\sigma(\mathcal{R})) = \sigma(X^{-1}(\mathcal{R})) &= \sigma(\omega\,;\, X(\omega) \in \{x \in E^T\,;\, x_t \in A_t\},\, A_t \in \mathscr{E} \text{ y } t \in T) \\ &= \sigma(\{X_t \in A_t\}\,;\, A_t \in \mathscr{E} \text{ y } t \in T) \\ &= \sigma(X_t\,;\, t \in T).\end{aligned}$$

Por lo tanto, la σ-álgebra generada por un proceso X se puede escribir equivalentemente como

$$\mathscr{F}^0_\infty = \sigma(X_t\,;\, t \in T).$$

A lo largo de este texto trabajaremos con procesos a tiempo continuo, así que T será $\mathbb{R}^+ := [0, \infty)$ y por lo tanto, $\mathscr{F}^0_\infty = \sigma(X_t\,;\, t \in \mathbb{R}^+)$.

1.2.2 Distribución

Ahora que los espacios $(E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$, $(C(\mathbb{R}^+, \mathbb{R}), \mathscr{E}_C^{\mathbb{R}^+})$ y $(D(\mathbb{R}^+, \mathbb{R}), \mathscr{E}_D^{\mathbb{R}^+})$ tienen estructura de espacio de medida, podemos considerar medidas definidas en ellos:

Similarmente a las variables aleatorias, la *ley inducida* por X en $(E^T, \mathscr{E}^T)$ es una medida de probabilidad en el espacio $(E^T, \mathscr{E}^T)$ denotada $\mathbb{P}_X$ y dada por:

$$\mathbb{P}_X(D) = \mathbb{P}\left(X^{-1}(D)\right) = \mathbb{P}\left(\omega\,;\, X(\omega) \in D\right), \qquad D \in \mathscr{E}^T.$$

Decimos que dos procesos X y Y tienen la misma ley si $\mathbb{P}_X = \mathbb{P}_Y$. Como el conjunto de los rectángulos finito dimensionales $\mathcal{C}$ es un $\pi-$sistema que genera a $\mathscr{E}^T$, cualquier medida en $(E^T, \mathscr{E}^T)$ queda completamente determinada por sus valores en $\mathcal{C}$. Por lo tanto, para que X y Y tengan la misma ley basta que su distribución coincida en los elementos de $\mathcal{C}$, esto es que

$$\mathbb{P}_X(x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n) = \mathbb{P}_Y(x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n)$$

o equivalentemente

$$\begin{aligned}&\mathbb{P}(\omega : X(\omega) \in \{x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n\}) \\ &\qquad = \mathbb{P}(\omega : Y(\omega) \in \{x \in E^T\,;\,(x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n\})\end{aligned}$$

para cualquier colección $A_1, \ldots, A_n \in \mathscr{E}$, $t_1, \ldots t_n \in T$ y $n \in \mathbb{N}$. Esto es justamente pedir que se satisfaga la

Definición 1.2.3 *Decimos que dos procesos X y Y tienen la misma distribución si*

$$\mathbb{P}\left(X_{t_1} \in A_1, \ldots X_{t_n} \in A_n\right) = \mathbb{P}\left(Y_{t_1} \in A_1, \ldots Y_{t_n} \in A_n\right).$$

para cualquier $n \in \mathbb{N}$ y cualquier colección $A_1, \ldots, A_n \in \mathscr{E}$, $t_1, \ldots t_n \in T$.

Es entonces natural definir la siguiente colección de medidas en $(E^n, \mathscr{E}^n)$ para $n \in \mathbb{N}$ asociadas a X:

$$\mu_{t_1,t_2,\ldots,t_n}(A_1 \times A_2 \times \cdots \times A_n) = \mathbb{P}\left(X_{t_1} \in A_1, \ldots X_{t_n} \in A_n\right) \quad A_i \in \mathscr{E},\, t_i \in T \tag{1.2.2}$$

en vista de que la ley de X queda completamente determinada por estas. A la colección de medidas definidas en (1.2.2) se les llama la familia de *distribuciones finito dimensionales* de X y no son más que la familia de distribuciones de los vectores $(X_{t_1}, \ldots, X_{t_n})$ en $(E^n, \mathscr{E}^n)$ para $t_i \in T$, $n \in \mathbb{N}$. Así, por la discusión que se tuvo, la ley de un proceso queda determinada por sus distribuciones finito-dimensionales.

1.2.3 Independencia

Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad.

Una *colección de eventos* en $\mathscr{F}$, $\mathcal{A} = (A_i)_{i \in I}$ arbitraria con índices en I es (mutualmente) independiente si para cualquier $n \in \mathbb{N}$ y cualquier colección de índices distintos $i_1, \ldots i_n \in I$ se satisface

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} \mathbb{P}\left(A_{i_k}\right) \tag{1.2.3}$$

Una *colección de familias de eventos* $(\mathcal{A}_i)_{i \in I}$ es independiente si se satisface (1.2.3) para índices $t_1, \cdots, t_n \in I$ distintos, arbitrarios y $A_{t_k} \in \mathcal{A}_{t_k}$ arbitrarios. Es decir,

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_{i_k}\right) = \prod_{k=1}^{n} \mathbb{P}\left(A_{i_k}\right) \quad \text{para todo } A_{t_1} \in \mathcal{A}_{t_1}, \ldots, A_{t_1} \in \mathcal{A}_{t_1}$$

$$\text{y cualesquiera } \{t_1, \ldots, t_n\} \subset I.$$

Una *colección de funciones medibles* $(X_i)_{i\in I}$ definidas en Ω es independiente si la colección de las σ-álgebras generadas por cada una de las funciones $(\sigma(X_i))_{i\in I}$ es independiente. Nótese que no se especifica la naturaleza de cada X_i: pueden tener imágenes en espacios distintos.

Para verificar la independencia entre σ-álgebras, basta con comprobar la independencia en π-sistemas generadores:

Lema 1.2.4 *Si una colección de π-sistemas $\mathcal{C}_i$, $i \in I$ es independiente, entonces la colección de σ-álgebras generadas $\mathscr{G}_i = \sigma(\mathcal{C}_i)$, $i \in I$ es independiente.*

Demostración
Fijamos $i_1, \ldots, i_n$ índices distintos y arbitrarios. Necesitamos probar que se satisface (1.2.3) para $A_{i_k} \in \mathscr{G}_{i_k}$ cualquiera y obsérvese que (1.2.3) ya se satisface para $A_{i_k} \in \mathcal{C}_{i_k}$.
Sea $D \subset \mathscr{G}_{i_1}$ la clase de los $A_{i_1} \in \mathscr{G}_{t_1}$ tales que se cumple (1.2.3) para $A_{i_2} \in \mathcal{C}_{i_2}, \ldots, A_{i_n} \in \mathcal{C}_{i_n}$ cualesquiera. D es una λ-sistema y $\mathcal{C}_{i_1} \subset D$. Por el lema de clases monótonas, $\mathscr{G}_{i_1} = \sigma(\mathcal{C}_{i_1}) \subset D$.
Entonces, (1.2.3) se satisface para $A_{i_1} \in \mathscr{G}_{i_1}$ y $A_{i_2} \in \mathcal{C}_{i_2}, \ldots, A_{i_n} \in \mathcal{C}_{i_n}$ arbitrarios.
Procedemos similarmente para $i_2, i_3 \ldots$ hasta llegar a i_n. ■

Escribiremos $\mathscr{G}_1 \bigvee \mathscr{G}_2 = \sigma\{\mathscr{G}_1, \mathscr{G}_2\}$ y $\bigvee_{i\in I} \mathscr{G}_i = \sigma(\mathscr{G}_i\,;\, i \in I)$.

Lema 1.2.5 *Sea $(\mathscr{G}_i)_{i\in I}$ colección de σ-álgebras independientes y considérese una partición $\mathcal{I}$ de I. Si definimos para cada $S \in \mathcal{I}$*

$$\mathscr{G}_S := \bigvee_{i\in S} \mathscr{G}_i,$$

entonces $\{\mathscr{G}_S\}_{S\in\mathcal{I}}$ vuelve a ser una familia independiente.

Demostración
Para cada $S \in \mathcal{I}$, considérese la familia $\mathcal{C}_S$ de intersecciones finitas $\bigcap_{k=1}^n A_k$ de elementos de $\cup_{i\in S}\mathscr{G}_i$. Es un π-sistema y genera a $\mathscr{G}_S$ pues en particular cada elemento de $\mathscr{G}_i$, $i \in S$ está en ella. Veamos que $(\mathcal{C}_S)_{S\in\mathcal{I}}$ forma una familia independiente: Podemos suponer s.p.g que si $C \in \mathcal{C}_S$ con $C = \bigcap_{k=1}^n A_k$ entonces los elementos A_i pertenecen a σ-álgebras distintas (si A_i, $A_j \in \mathscr{G}_k$ para alguna k con $i \neq j$, renombrar $A_i \cap A_j = \tilde{A}_i$). Sean $C_{s_1}, \ldots C_{s_n}$ con $C_{s_i} \in \mathcal{C}_S$ para alguna $S \in \mathcal{I}$ de modo que si $s_i \neq s_j$, $C_{s_i} \in \mathcal{C}_S$ y $C_{s_i} \in \mathcal{C}_{\tilde{S}}$, entonces $S \neq \tilde{S}$. Esto es, cada C_{s_i} se escribe como $\bigcap_{k=1}^n A_k$ con $A_1, \cdots, A_n$ elementos pertenecientes a sigma-álgebras distintas de la colección $\mathscr{G}_i, i \in S$. Al usar la independencia de $(\mathscr{G}_i)_{i\in I}$ y luego agrupar adecuadamente, obtenemos:

$$\mathbb{P}\left(\bigcap_{i=1}^n C_{s_i}\right) = \prod_{i=1}^n \mathbb{P}\left(C_{s_i}\right)$$

por lo que los π-sistemas $(C_S)_{S\in\mathcal{I}}$ son independientes. Se sigue del lema 1.2.4 que la colección de σ-álgebras $\mathscr{G}_S = \sigma(\mathcal{C}_S)$, $S \in \mathcal{I}$ es independiente. ■

Es interesante notar que las definiciones de ley e independencia de funciones medibles $X : \Omega \to H$ donde $(H, \mathscr{H})$ es un espacio medible, son las mismas sin importar la naturaleza del espacio de estados. Si $H = \mathbb{R}$, X es variable aleatoria , si $H = \mathbb{R}^n$, X es vector aleatorio y si $H = \mathscr{E}^T$, X es proceso estocástico.

Consecuencias al probar independencia de procesos:
Sean $X, Y : \Omega \to U$ procesos con trayectorias en $U \subseteq E^T$. Como la colección de los rectángulos finito dimensionales $\mathcal{C}$ forma un π-sistema en $U \cap \mathscr{E}^T$ y $\sigma(\mathcal{C}) = U \cap \mathscr{E}^T$, entonces $X^{-1}(\mathcal{C})$ y $Y^{-1}(\mathcal{C})$ son π-sistemas en $\mathscr{F}$ y generan a $\mathscr{F}_\infty^{0,X}$ y $\mathscr{F}_\infty^{0,Y}$ respectivamente pues $\mathscr{F}_\infty^{0,X} = X^{-1}(\sigma(\mathcal{C})) = \sigma(X^{-1}(\mathcal{C}))$. Nótese que los elementos de $X^{-1}(\mathcal{C})$ son de la forma

$$\begin{aligned} A &= \{\omega\,;\, X(\omega) \in \{x \in U \cap E^T\,;\, (x_{t_1}, \ldots, x_{t_n}) \in A_1 \times \ldots \times A_n\}\} \\ &= \{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\} \end{aligned}$$

con $t_1, \ldots, t_n \in T$ y $A_1, \ldots, A_n \in \mathscr{E}$ (y por lo tanto $X^{-1}(\mathcal{C})$ es un π-sistema ya que intersecciones finitas vuelve a escribirse de esta forma) mientras que los de $Y^{-1}(\mathcal{C})$ se escriben como

$$B = \{Y_{s_1} \in B_1, \ldots, Y_{s_m} \in B_m\}$$

para $s_1, \ldots, s_m \in [0, T]$ y $B_1, \ldots, B_n \in \mathscr{E}$. Por lo tanto, por el lema 1.2.4, para probar que X y Y son independientes bastaría ver que

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B) \qquad \text{para } A \in X^{-1}(\mathcal{C}), B \in Y^{-1}(\mathcal{C})$$

que es probar que se satisface

$$\begin{aligned} &\mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n, Y_{s_1} \in B_1, \ldots, Y_{s_m} \in B_m) = \\ &\qquad\qquad \mathbb{P}(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)\,\mathbb{P}(Y_{s_1} \in B_1, \ldots, Y_{s_m} \in B_m) \end{aligned} \tag{1.2.4}$$

para toda colección finita $t_1, \ldots, t_n, s_1, \ldots, s_m \in T$ y $A_1, \ldots, A_n, B_1, \ldots, B_n \in \mathscr{E}$ para cualquier $n, m \in \mathbb{N}$. Por lo tanto, definimos

Definición 1.2.6 *Dos procesos X, Y con valores en E son independientes si se satisface (1.2.4) para cualquier colección finita $t_1, \ldots, t_n \in T$ y $A_1, \ldots, A_n \in \mathscr{E}$.*

Nos referiremos al π-sistema en $(\Omega, \mathscr{F})$ de las colecciones de la forma $\{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\}$ como el π-sistema usual que genera a $\mathscr{F}_\infty^0 = \sigma(X_t\,;\, t \in T)$.

1.3 Teorema de existencia de Kolmogorov y espacios canónicos

Definición 1.3.1 *Para todo $n \in \mathbb{N}$ y $t_1 < \cdots < t_n$ con $t_i \in T$, sean $\mu_{t_1, \ldots t_n}$ medidas de probabilidad en $(E^n, \mathscr{E}^n)$. Decimos que la colección de medidas es consistente si*

$$\begin{aligned} &\mu_{t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_n}(A_1 \times \cdots \times A_{i-1} \times A_{i+1} \times \cdots \times A_n) = \\ &\qquad\qquad \mu_{t_1, \ldots, t_n}(A_1 \times \cdots \times A_{i-1} \times E \times A_{i+1} \times \cdots \times A_n) \end{aligned}$$

para todo $A_i \in \mathscr{E}$, $i \in 1, \ldots, n$.

Dada una colección $\mathscr{M}$ de medidas consistentes, el teorema de Kolmogorov garantiza la existencia de un proceso definido en un espacio de probabilidad con $\Omega = E^T$ cuya familia de distribuciones finito-dimensionales es justamente $\mathscr{M}$. En particular, como es claro que las distribuciones finito dimensionales de un proceso forman un sistema de medidas consistente, este resultado nos permite hablar de la versión *canónica* de un proceso Y, que no es más que una versión de Y definida en el espacio de trayectorias $\Omega = E^{\mathbb{R}^+}$.

Teorema 1.3.2 Teorema de consistencia de Kolmogorov
Supongamos que E es un espacio Polaco (espacio métrico completo y separable) y $\mathscr{E}$ su $\sigma-$álgebra. Consideremos una familia de medidas

$$\mathscr{M} = \{\mu_{t_1,\dots,t_n} \in \mathcal{P}(E^n) : n \in \mathbb{N},\, t_1, \dots, t_n \in T\}.$$

donde el espacio E^n siempre está equipado con su sigma-álgebra producto $(E^n, \mathscr{E}^n)$. Si $\mathscr{M}$ forma un sistema de medidas consistente, entonces existe una medida de probabilidad $\mathbb{P}$ en $(E^T, \mathscr{E}^T)$ tal que el proceso canónico $X = (X_t)_{t\in T}$ en $(\Omega = E^T, \mathscr{F} = \mathscr{E}^T, \mathbb{P})$ definido como la identidad de $E^T \to E^T$

$$X(\omega) = \omega \qquad X_t(\omega) = \omega_t$$

tiene por familia de distribuciones finito dimensionales a $\mathscr{M}$.

La prueba de este teorema se puede consultar en Billingsley [7] capítulo 7, página 513. A X se le llama el *proceso coordenado* en el espacio canónico E^T o simplemente *proceso canónico.* Los espacios canónicos $\Omega = E^{\mathbb{R}^+}$, $C(\mathbb{R}^+, E)$ y $D(\mathbb{R}^+, E)$ son los espacios naturales en los que definir a un proceso, proceso continuo y proceso cadlag respectivamente pues sus elementos ω representan la evolución del proceso en el tiempo. En cualquier caso, a $X : \Omega \to \Omega$ definido como la identidad se le llama proceso canónico. Por este teorema, dado un proceso Y, podemos siempre que sea necesario recurrir a su versión canónica X definida en $\Omega = E^{\mathbb{R}^+}$. Nótese la limitante: si el proceso original tenía trayectorias continuas o cadlag, la versión dada por este teorema no tiene por qué tenerlas, solo sabemos que tienen la misma distribución. Más adelante, lidiaremos con este problema al construir una versión canónica de un proceso de Lévy (y que por lo tanto debe tener trayectorias en $D(\mathbb{R}^+, \mathbb{R})$).

Nótese que $X_t(\omega)$ es simplemente la proyección $\pi_t(\omega)$ para toda t, y habría sido equivalente definir a X como $X = (\pi_t)_{t\geq 0}$ pero adaptamos la notación.[1]
Al ser X la identidad de $E^{\mathbb{R}^+}$ en $E^{\mathbb{R}^+}$, obtenemos entonces que si

$$\mathscr{F}^0_\infty = \sigma(X_t\,;\, t \geq 0) \quad \text{entonces} \quad \mathscr{F}^0_\infty = \mathscr{E}^{\mathbb{R}^+}.$$

Como el π-sistema de los rectángulos

$$R = \{x \in E^T\,;\, (x_{t_1}, \dots, x_{t_n}) \in A_1 \times \dots \times A_n\} \qquad t_1 < \cdots < t_n \in \mathbb{R}^+\, A_i \in \mathscr{E}$$

genera a $\mathscr{E}^{\mathbb{R}^+}$, entonces genera a $\mathscr{F}^0_\infty$. Obsérvese que al ser X_t la valuación en t,

$$\begin{aligned} R &= \{x \in E^T\,;\, (\pi_{t_1}(x), \dots, \pi_{t_1}(x)) \in A_1 \times \dots \times A_n\} \\ &= \{\omega \in E^T\,;\, (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in A_1 \times \cdots \times A_n)\} \\ &= \{X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\} \end{aligned}$$

Una de las ventajas de trabajar con versiones canónicas, es que se puede definir el siguiente operador en $\Omega = E^{\mathbb{R}^+}, C(\mathbb{R}^+, E)$ ó $D(\mathbb{R}^+, E)$:

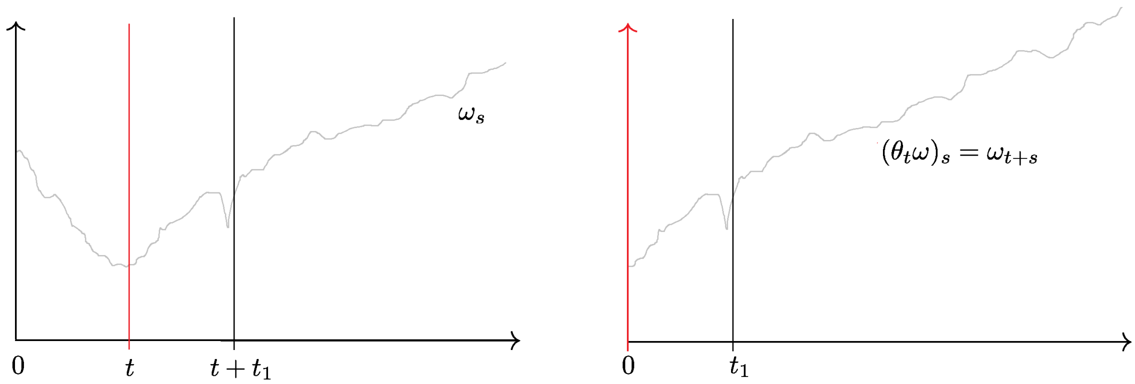
[1]Cuando trabajemos en un espacio canónico, denotamos por $\omega = (\omega_t)_t$ a los elementos de $E^{\mathbb{R}^+}$ cuando se trata de un elemento del espacio de probabilidad mientras que denotaremos por x cuando queramos enfatizar el hecho de que se trata de una trayectoria. Sin embargo representan lo mismo en este contexto.

Definición 1.3.3 *Para cualquier $t \geq 0$, definimos el operador de traslación*

$$\theta_t : E^{\mathbb{R}^+} \to E^{\mathbb{R}^+} \qquad\qquad (\theta_t \omega)_s = \omega_{t+s}$$

para $s \geq 0$ y $\omega \in E^{\mathbb{R}^+}$.

El operador θ_t corta la parte de la trayectoria ω que está antes de t y este se vuelve el nuevo punto de partida:



Claramente $\theta_s \circ \theta_t = \theta_{s+t}$ y es $\mathscr{F}^0_\infty/\mathscr{F}^0_\infty$ medible para cada t: basta probar que $(\theta_t)^{-1}(A) \in \mathscr{F}^0_\infty$ para A en el π-sistema usual:

$$\begin{aligned}
&(\theta_t)^{-1}\left(\omega \in E^T ; (\omega_{t_1}, \ldots, \omega_{t_n}) \in A_1 \times \cdots \times A_n\right) \\
&\qquad = \{\omega \in E^T ; (\omega_{t+t_1}, \ldots, \omega_{t+t_n}) \in A_1 \times \cdots \times A_n ; \} \\
&\qquad = \{X_{t+t_1} \in A_1, \ldots, X_{t+t_n} \in A_n\} \in \sigma(X_s ; s \geq t) \subset \mathscr{F}^\infty_0
\end{aligned}$$

por lo que de hecho, es $\sigma(X_s ; s \geq t)/\mathscr{F}^\infty_0$-medible. Para $\Omega = C(\mathbb{R}^+, E)$ ó $D(\mathbb{R}^+, E)$ consideramos la restricción de θ_t a estos espacios y la prueba de medibilidad es similar.

Capítulo 2

Distribuciones infinitamente divisibles

2.1 Resultados preliminares de convergencia débil en $\mathbb{R}$ y Teorema del Límite Central de Lindberg para arreglos triangulares

Referencias: Billingsley [7] capítulo 5 y Ash [1] sección 7.3

Iniciaremos con un breve repaso no exhaustivo de resultados de convergencia en ley en $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ ya que se usarán a lo largo del capítulo. Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad, consideraremos variables aleatorias definidas en este espacio a lo largo del capítulo. Los resultados que probaremos en esta sección se restringen al caso $E = \mathbb{R}$ aunque algunas definiciones se darán para espacios métricos pues se usarán más adelante.

Definición 2.1.1 *Una sucesión de medidas de probabilidad $\{\mu_n\}$ definidas en un espacio métrico (E, d) equipado con la σ-álgebra de sus borelianos $\mathscr{B}(E)$ converge débilmente, en ley o en distribución a otra medida de probabilidad μ si para toda $f : E \to \mathbb{R}$ continua y acotada, se tiene que*

$$\lim_{n\to\infty} \int_{\mathbb{E}} f d\mu_n = \int_{\mathbb{E}} f d\mu$$

En el caso $E = \mathbb{R}$, si para cada n, F_n es la función de distribución asociada a μ_n y F la de μ, esta definición es equivalente a pedir que $\lim_{n\to\infty} F_n(x) = F(x)$ para todo x en el conjunto de puntos de continuidad de F, que denotaremos por C_F. Diremos que una sucesión de variables aleatorias con valores en un espacio métrico $(E, \mathscr{B}(E))$ converge en ley a una variable aleatoria X si sus leyes $\mathbb{P}_{X_n} = \mu_n$ en $(E, \mathscr{B}(E))$ convergen débilmente a $\mathbb{P}_X = \mu$. En el caso real, por el teorema de cambio de variable esto es equivalente a pedir que

$$\lim_{n\to\infty} \mathbb{E}\left[f(X_n)\right] = \mathbb{E}\left[f(X)\right] \text{ para toda } f \text{ continua y acotada.}$$

Se denotará a menudo tanto $\mu_n \xrightarrow{\mathscr{L}} \mu$ como $X_n \xrightarrow{\mathscr{L}} X$.

Teorema 2.1.2 Selección de Helly
Para cada sucesión $\{F_n\}$ de funciones de distribución en $\mathbb{R}$, existe una subsucesión $\{F_{n_k}\}$ y una función F no decreciente y continua por la derecha tal que

$$\lim_{k\to\infty} F_{n_k}(x) = F(x)$$

para todo $x \in C_F$.

Demostración
Al aplicar el método diagonal, no es difícil probar que existe una subsucesión $\{F_{n_k}\}$ tal que $\lim_{k\to\infty} F_{n_k}(r)$ existe para todo racional r. Entonces para cada racional, sea $G(r) := \lim_{k\to\infty} F_{n_k}(r)$. Obsérvese que G es no decreciente al ser límite de funciones no decrecientes. Definimos para toda x en $\mathbb{R}$,

$$F(x) := \inf\big\{G(r) \,|\, x < r, r \in \mathbb{Q}\big\}.$$

F es no decreciente y si $y < r$ entonces $F(y) \leq G(r)$. Ahora, veamos que F es continua por la derecha:
Sea $\epsilon > 0$ y $x \in \mathbb{R}$. Por construcción de F, existe un racional r tal que $x < r$ y:

$$G(r) < F(x) + \epsilon.$$

Entonces, para toda y tal que $x < y < r$ se cumple que $F(x) \leq F(y) \leq G(r) < F(x) + \epsilon$ por lo que F es continua por la derecha.

Falta probar que $F_{n_k}(x) \to F(x)$ para toda $x \in C_F$:
Sean $\epsilon > 0$ y $x \in C_F$, por continuidad por la izquierda en x, existe $y < x$ tal que $F(x)-\epsilon < F(y)$. Por el mismo argumento que se usó para la continuidad por la derecha en x, existe un racional r tal que $x < r$ y $G(r) < F(x) + \epsilon$. Si además, s es un racional tal que $y < s < x < r$, entonces tenemos que:

$$F(x) - \epsilon < F(y) \leq G(s) \leq G(r) \leq F(x) + \epsilon.$$

Como para toda k, $F_{n_k}(s) \leq F_{n_k}(x) \leq F_{n_k}(r)$ y $F_{n_k}(s) \to G(s)$, $F_{n_k}(r) \to G(r)$ entonces:

$$F(x) - \epsilon < \liminf F_{n_k}(x) \leq \limsup F_{n_k}(x) \leq F(x) + \epsilon$$

por lo que concluimos que

$$\lim_{n\to\infty} F_{n_k}(x) = F(x).$$

■

Nótese que F no tiene por qué ser una función de distribución: por ejemplo, $F_n(x) = \mathbb{1}_{[n,\infty]}(x)$ es función de distribución para cada $n \in \mathbb{N}$ pero $\lim_{n\to\infty} F_n(x) \equiv 0$ que no es una función de distribución. Para garantizar que el límite vuelva a ser una función de distribución, necesitamos garantizar que no se escapa la "masa" y para ello, necesitamos el concepto de *tensión:*

Definición 2.1.3 *Sea $\mathscr{M} = (\mu_i)_{i\in I}$ un familia de medidas de probabilidad en un espacio métrico E con la sigma-álgebra de sus borelianos $\mathscr{B}(E)$.*
1) Decimos que la familia $\mathscr{M}$ es tensa si para cada $\epsilon > 0$, existe un compacto K tal que $\mu_i(K^c) < \epsilon$ para toda $i \in I$. Si el espacio es $\mathbb{R}^d$, podemos remplazar a K por un rectángulo

cerrado en la definición.
2) Decimos que la familia $\mathscr{M}$ es relatívamente compacta si para cualquier sucesión $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathscr{M}$ existe una subsucesión que converge débilmente a una medida de probabilidad ν en $(E, \mathscr{B}(E))$.

En el caso $E = \mathbb{R}$, tenemos el

Teorema 2.1.4 Prokhorov en $\mathbb{R}$
Una sucesión de medidas $\{\mu_n\}$ en $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ es tensa si y solo si es relativamente compacta, es decir, si de toda subsucesión $\{\mu_{n_k}\}$ de $\{\mu_n\}$, se puede extraer otra sub-sucesión $\{\mu_{n_{k_j}}\}$ tal que $\mu_{n_{k_j}} \xrightarrow{\mathscr{L}} \mu$ para alguna medida de probabilidad μ.

Demostración
Primero, supongamos que la sucesión $\{\mu_n\}$ es tensa y sea $\{\mu_{n_k}\}$ subsucesión cualquiera de esta. Por el teorema de selección de Helly 2.1.2, la sucesión de distribuciones asociadas $\{F_{n_k}\}$ tiene una subsucesión $\{F_{n_{k_j}}\}$ que converge a una función no decreciente y continua por la derecha F en sus puntos de continuidad; sea μ la medida asociada a F, probaremos que es de probabilidad. Como la sucesión es tensa, dado $\epsilon > 0$ existen a y b tales que $\mu_{n_{k_j}}(a, b] > 1 - \epsilon$ para toda j y podemos suponer que a y b son puntos de continuidad de F. Entonces,

$$\lim_{j\to\infty} \mu_{n_{k_j}}(a, b] = \lim_{j\to\infty} F_{n_{k_j}}(b) - F_{n_{k_j}}(a) = F(b) - F(a) = \mu(a, b].$$

Por lo tanto, $\mu(a, b] > 1 - \epsilon$, y como esto es para cualquier ϵ, entonces μ debe ser medida de probabilidad.

Si suponemos que la sucesión no es tensa, entonces debe existir $\epsilon > 0$ tal que para cada $k \in \mathbb{N}$

$$\inf_n \mu_n(-k, k] < 1 - \epsilon.$$

Por lo tanto, para cada k, existe $n_k > n_{k-1}$ tal que $\mu_{n_k}(-k, k] < 1 - \epsilon$. Consideramos la sub-sucesión $\{\mu_{n_k}\}$ y supongamos que existe una subsucesión $\{\mu_{n_{k_j}}\}$ de $\{\mu_{n_k}\}$ que converge a una medida de probabilidad μ. Sean a y b puntos de continuidad de μ tales que $\mu(a, b] > 1 - \epsilon$. Para k suficientemente grande, $(a, b] \subset (-k_j, k_j]$, por lo que:

$$1 - \epsilon \geq \lim_{j\to\infty} \mu_{n_{k_j}}(-k_j, k_j] \geq \lim_{n\to\infty} \mu_{n_{k_j}}(a, b] = \mu(a, b]$$

lo cual lleva a una contradicción. ∎

Esto nos permite probar la siguiente caracterización de convergencia débil:

Corolario 2.1.5 *Sea $\{\mu_n\}$ sucesión de medidas tensa en $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. Si toda subsucesión $\{\mu_{n_k}\}$ que converge débilmente, converge a la misma medida μ, entonces $\mu_n \xrightarrow{\mathscr{L}} \mu$.*

Demostración
Supongamos que toda subsucesión convergente de $\{\mu_n\}$ converge a μ pero que $\mu_n \overset{\mathscr{L}}{\nrightarrow} \mu$. Entonces existe $\epsilon > 0$ y x punto de continuidad de F tal que

$$\limsup_n |F_n(x) - F(x)| > \epsilon$$

y por lo tanto, existe una subsucesión $\{F_{n_k}\}$ tal que $|F_{n_k}(x) - F(x)| > \epsilon$ para toda k. Como la sucesión de medidas $\{\mu_{n_k}\}$ es subsucesión de $\{\mu_n\}$ que es tensa, entonces debe existir una subsucesión $\{\mu_{n_{k_j}}\}$ de $\{\mu_{n_k}\}$ que converge a una medida ν, y por hipótesis, $\nu = \mu$. Pero entonces $F_{n_{k_j}}(x) \to F(x)$ al ser x un punto de continuidad, lo cual lleva a una contradicción. ■

Este corolario permite probar el fundamental teorema decontinuidad de Lévy en $\mathbb{R}$, del cual omitiremos la prueba:

Teorema 2.1.6 Continuidad de Lévy en $\mathbb{R}$
Sean X_n, X variables aleatorias reales con funciones características φ_{X_n} y φ_X respectivamente.
1) Si $X_n \overset{\mathscr{L}}{\to} X$, entonces $\varphi_{X_n}(\lambda) \to \varphi_X(\lambda)$ para todo $\lambda \in \mathbb{R}$.
2) Si $\varphi_{X_n}(\lambda)$ converge a una función $f(\lambda)$ para todo $\lambda \in \mathbb{R}$ y f es continua en 0, entonces existe una variable aleatoria X con valores en $\mathbb{R}$ tal que $\varphi_X(\lambda) = f(\lambda)$ y $X_n \overset{\mathscr{L}}{\to} X$.
Además, en cualquiera de los dos casos, la sucesión de medidas $\mu_n = \mathbb{P}_{X_n}$ es tensa.

La prueba se puede consultar en [7] capítulo 5. Con esto, llegamos al teorema central de esta primer sección:

Teorema 2.1.7 Teorema del límite central de Lindberg para arreglos triangulares
Supongamos que para cada n, $\{X_{n1}, \cdots, X_{nr_n}\}$ es una colección de variables aleatorias independientes:

$$\begin{bmatrix} X_{1r_1} & & & & \\ X_{21} & X_{1r_2} & & & \\ X_{31} & X_{32} & X_{3r_3} & & \\ \cdots & & & & \\ X_{n1} & X_{n2} & X_{n3} & \cdots & X_{nr_n} \\ \vdots & & & & \end{bmatrix}$$

donde $r_n \leq r_{n+1}$ y $r_n \to \infty$. A una colección de este tipo se le llama arreglo triangular de variables aleatorias.
Sea $S_n = X_{n1} + \cdots + X_{nr_n} = \sum_{k=1}^{r_n} X_{nk}$ y supongamos que:

$$\mathbb{E}[X_{kl}] = 0, \qquad \mathbb{E}\left[X_{kl}^2\right] = \sigma_{nk}^2 < \infty \qquad s_n^2 := \sum_{k=1}^{r_n} \sigma_{nk}^2.$$

Si además, para todo $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{\{|X_{nk}| > \epsilon \cdot s_n\}} X_{nk}^2 \, d\mathbb{P} = 0 \tag{2.1.1}$$

entonces $\lim_{n\to\infty} \frac{S_n}{s_n} = N$ en ley. A la condición (2.1.1) se le llama condición de Lindberg.

Antes de probar el teorema, veamos que las versiones usuales del teorema del límite central quedan como caso particular del teorema anterior:

1− *Teorema del límite central para variables aleatorias idénticamente distribuidas*

Si $\{X_n\}_{n\in\mathbb{N}}$ son variables aleatorias idénticamente distribuidas con media 0 y varianza finita σ^2, entonces tenemos el arreglo triangular compuesto por $\{X_{nk} = X_k\}$, es decir, para cada n, el n-ésimo renglón del arreglo triangular está compuesto por $\{X_1, \cdots, X_n\}$.
Aquí:

$$s_n^2 = n\sigma^2 \qquad s_n = \sqrt{n\cdot\sigma^2} = \sigma\sqrt{n}, \qquad r_n = n.$$

Entonces la condición de Lindberg se reduce a:

$$\lim_{n\to\infty}\sum_{k=1}^{r_n}\frac{1}{s_n^2}\int_{\{|X_{nk}|>\epsilon\cdot s_n\}} X_{nk}^2\, d\mathbb{P} = \lim_{n\to\infty}\frac{n}{n\sigma^2}\int_{\{|X_1|>\epsilon\cdot\sigma\sqrt{n}\}} X_1^2 d\mathbb{P} = 0$$

por el teorema de convergencia dominada. Así, por el Teorema 2.1.7

$$\lim_{n\to\infty}\frac{S_n}{s_n} = \frac{S_n}{\sigma\sqrt{n}} = N$$

2$-$*El caso uniformemente acotado*
Si suponemos que $\{X_k\}_{k\in\mathbb{N}}$ es una colección con media 0 s.p.g. y varianza finita tal que $|X_k|\leq M$ para toda k y $\lim_{n\to\infty} s_n = \infty$, entonces:

$$\begin{aligned}\int_{\{|X_k|>\epsilon\cdot s_n\}} X_k^2\, d\mathbb{P} &= \mathbb{E}\left[|X_k|^2 \mathbb{1}_{\{|X_k|>\epsilon\cdot s_n\}}\right]\\ &\leq M^2\mathbb{P}\left(|X_k|>\epsilon\cdot s_n\right)\\ &\leq M^2\frac{\mathbb{E}\left[|X_k|^2\right]}{\epsilon^2 s_n^2}\\ &= M^2\frac{\sigma_k^2}{\epsilon^2 s_n^2}\end{aligned}$$

por lo que:

$$\begin{aligned}\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{\{|X_k|>\epsilon\cdot s_n\}} X_k^2\, d\mathbb{P} &\leq \frac{1}{s_n^2}\sum_{k=1}^{n} M^2\frac{\sigma_k^2}{\epsilon^2 s_n^2}\\ &\leq \frac{M^2}{\epsilon^2 s_n^2}\to 0\end{aligned}$$

Como se satisface la condición de Lindberg, $\frac{S_n}{s_n}\to N$.

3$-$*Condiciones Lyapunov*
Bajo las mismas hipótesis del teorema de Linberg, reemplazamos la condición de Lindberg por la siguiente:
supongamos que existe $\delta > 0$ tal que:

$$\lim_{n\to\infty}\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{r_n}\mathbb{E}\left[|X_{nk}|^{2+\delta}\right] = 0.$$

Veamos que si la condición de Lyapunov se cumple, entonces la de Lindberg también:

$$\begin{aligned}\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{\{|X_k|>\epsilon\cdot s_n\}} X_k^2\, d\mathbb{P} &= \frac{1}{s_n^2}\sum_{k=1}^{n}\int_{\{|X_k|>\epsilon\cdot s_n\}} X_k^\delta\frac{X_k^2}{X_k^\delta}\, d\mathbb{P}\\ &\leq \frac{1}{s_n^2 s_n^\delta\epsilon^\delta}\sum_{k=1}^{n}\int_{\{|X_k|>\epsilon\cdot s_n\}}|X_k|^{2+\delta}\, d\mathbb{P}\\ &\leq \frac{1}{s_n^{2+\epsilon}\epsilon^\delta}\sum_{k=1}^{n}\mathbb{E}\left[|X_k|^{2+\delta}\right]\to 0.\end{aligned}$$

Por lo tanto, se satisfacen las hipótesis del teorema 2.1.7 y $\frac{S_n}{s_n} \to N$.

Los siguientes lemas técnicos no se probaran, se pueden consultar sus pruebas en Billinsgley [7] p367 y serán necesarios en la prueba del teorema:

Lema 2.1.8 *Para todo $u \in \mathbb{R}$ se satisface:*

$$\left| e^{iux} - \sum_{k=0}^{n} \frac{(iu)^k}{n!} \right| \leq \min \left\{ \frac{|u|^{n+1}}{(n+1)!}, 2\frac{|u|^n}{n!} \right\}$$

Lema 2.1.9 *Si $z_i, w_i \in \mathbb{C}$ son complejos unitarios para todo $i \in \{1, \cdots, m\}$ entonces:*

$$\left| \prod_{i=1}^{m} w_i - \prod_{i=1}^{m} z_i \right| \leq \sum_{k=1}^{m} |z_k - w_k|$$

Lema 2.1.10 *Para $z \in \mathbb{C}$, se satisface que:*

$$|e^z - 1 - z| \leq |z|^2 \sum_{k=2}^{\infty} \frac{|z|^{k-z}}{k!} \leq |z|^2 e^{|z|}$$

Ahora si, probaremos el Teorema del límite central de Lindberg Feller:

Demostración

Denotaremos por φ_X a la función característica de una variable aleatoria X, es decir $\varphi_X(t) = \mathbb{E}\left[e^{itX}\right]$. Es bien sabido que si N es una variable aleatoria normal $(0,1)$, entonces $\varphi_N = e^{-\frac{t^2}{2}}$. Sin pérdida de generalidad, supondremos que $s_n^2 = 1$ ya que de no ser el caso, se definen nuevas variables aleatorias normalizadas $X'_{nk} = \frac{X_{nk}}{s_n}$. Estas si satisfacen que $(s'_n)^2 = 1$ y nótese que cumplen la condición de Lindberg pues:

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \int_{\{|X'_{nk}|>\epsilon\}} |X'_{nk}|^2 \, d\mathbb{P} = \lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{r_n} \int_{\{|X_{nk}|>\epsilon\cdot s_n\}} X_{nk}^2 \, d\mathbb{P} = 0.$$

Entonces al trabajar en el caso $s_n^2 = 1$, la condición de Lindberg se reduce a:

$$\lim_{n\to\infty} \sum_{k=1}^{r_n} \int_{\{|X_{nk}|>\epsilon\}} |X_{nk}|^2 \, d\mathbb{P} = 0$$

y hay que probar que $S_n \to N$ en ley. Nótese que en este caso, (2.1.1) tiene por consecuencia que

$$\lim_{n\to\infty} \max_{k=1,\cdots,r_n} \sigma_{nk}^2 \to 0.$$

En efecto, como

$$\sigma_{nk} = \int_{|X_{nk}|>\epsilon} X_{nk}^2 d\mathbb{P} + \int_{|X_{nk}|\leq\epsilon} X_{nk}^2 d\mathbb{P} \leq \int_{|X_{nk}|>\epsilon} X_{nk}^2 d\mathbb{P} + \epsilon^2.$$

entonces

$$\max_{k=1,\cdots,r_n} \sigma_{nk}^2 \leq \max_{k=1,\cdots,r_n} \int_{|X_{nk}|>\epsilon} X_{nk}^2 d\mathbb{P} + \epsilon^2 \leq \sum_{k=1}^{r_n} \int_{|X_{nk}|>\epsilon} X_{nk}^2 d\mathbb{P} + \epsilon^2 \longrightarrow 0 + \epsilon^2.$$

La idea de la prueba es la siguiente:

Por el teorema de Lévy en $\mathbb{R}$, bastaría probar que $|\varphi_N - \varphi_{S_n}| \to 0$. Como $\sum_k \sigma_{nk}^2 = 1$, entonces podemos reescribir

$$\varphi_N = e^{\frac{-t^2}{2}} = \prod_{k=1}^{r_n} e^{\frac{-t^2}{2}\sigma_{nk}^2}$$

y por otro lado, por independencia:

$$\varphi_{S_n} = \prod_{k=1}^{r_n} \varphi_{nk} \qquad \varphi_{nk} \text{ función característica de } X_{nk}.$$

Como,

$$\begin{aligned} |\varphi_N - \varphi_{S_n}| &= \left| \prod_{k=1}^{r_n} e^{\frac{-t^2}{2}\sigma_{nk}^2} - \prod_{k=1}^{r_n} \varphi_{nk} \right| \\ &\leq \underbrace{\left| \prod_{k=1}^{r_n} e^{\frac{-t^2}{2}\sigma_{nk}^2} - \prod_{k=1}^{r_n} \left(1 - \frac{t^2}{2}\sigma_{nk}^2\right) \right|}_{A_n} + \underbrace{\left| \prod_{k=1}^{r_n} \left(1 - \frac{t^2}{2}\sigma_{nk}^2\right) + \prod_{k=1}^{r_n} \varphi_{nk} \right|}_{B_n} \end{aligned}$$

bastaría probar que tanto A_n como B_n convergen a 0.
Empezaremos probando que A_n tiende a 0. Por el lema 2.1.8,

$$\left| e^{itx} - \left(1 + itx - t^2 \frac{x^2}{2}\right) \right| \leq \min\{|tx|^2, |tx|^3\}$$

así que al evaluar en $x = X_{nk}$ y tomar esperanza de ambos lados obtenemos que:

$$\left| \mathbb{E}\left[e^{itX_{nk}}\right] - \mathbb{E}\left[1 + itX_{nk} - \frac{t^2}{2} X_{nk}\right] \right| \leq \mathbb{E}\left[\min\{|tX_{nk}|^2, |tX_{nk}|^3\}\right].$$

Si denotamos M_{nk} a $\min\{|tX_{nk}|^2, |tX_{nk}|^3\}$, la expresión anterior se reduce a:

$$\left|\varphi_{nk}(t) - \left(1 + t^2\mathbb{E}\left[X_{nk}^2\right]\right)\right| \leq \mathbb{E}\left[M_{nk}\right]. \tag{2.1.2}$$

Pero:

$$\begin{aligned} \mathbb{E}\left[M_{nk}\right] &= \int_{|X_{nk}|\leq\epsilon} M_{nk} d\mathbb{P} + \int_{|X_{nk}|>\epsilon} M_{nk} d\mathbb{P} \\ &\leq \int_{|X_{nk}|\leq\epsilon} |tX_{nk}|^3 d\mathbb{P} + \int_{|X_{nk}|>\epsilon} |tX_{nk}|^2 d\mathbb{P} \\ &\leq |t|^3\epsilon \int_{|X_{nk}|\leq\epsilon} |X_{nk}|^2 d\mathbb{P} + |t|^2 \int_{|X_{nk}|>\epsilon} |X_{nk}|^2 d\mathbb{P} \end{aligned}$$

Por (2.1.2) y la desigualdad anterior,

$$\left|\varphi_{nk}(t) - \left(1 + t^2\mathbb{E}\left[X_{nk}^2\right]\right)\right| \leq |t|^3\epsilon \int_{|X_{nk}|\leq\epsilon} |X_{nk}|^2 d\mathbb{P} + |t|^2 \int_{|X_{nk}|>\epsilon} |X_{nk}|^2 d\mathbb{P}$$

y entonces al considerar las sumas hasta r_n:

$$\begin{aligned} \sum_{k=1}^{r_n} \left|\varphi_{nk}(t) - \left(1 + t^2\mathbb{E}\left[X_{nk}^2\right]\right)\right| &\leq |t|^3\epsilon \sum_{k=1}^{r_n} \sigma_{nk}^2 + t^2 \sum_{k=1}^{r_n} \int_{|X_{nk}|>\epsilon} |X_{nk}|^2 d\mathbb{P} \\ &\leq |t|^3\epsilon + t^2 \sum_{k=1}^{r_n} \int_{|X_{nk}|>\epsilon} |X_{nk}|^2 d\mathbb{P} \end{aligned}$$

Por la condición de Lindberg, este segundo término tiende a 0, por lo que podemos concluir que:

$$\lim_{n\to\infty}\sum_{k=1}^{r_n}\left|\varphi_{nk}(t)-\left(1+t^2\mathbb{E}\left[X_{nk}^2\right]\right)\right|=0.$$

De poder aplicar el lema 2.1.9 al considerar $z_i=\varphi_{ni}$ y $w_j=1-\frac{t^2}{2}\sigma_{nk}^2$ obtendríamos que:

$$\begin{aligned}\lim_{n\to\infty}A_n=\lim_{n\to\infty}&\left|\prod_{k=1}^{r_n}e^{\frac{-t^2}{2}\sigma_{nk}^2}-\prod_{k=1}^{r_n}\left(1-\frac{t^2}{2}\sigma_{nk}^2\right)\right|\\&\leq\lim_{n\to\infty}\sum_{k=1}^{r_n}\left|\varphi_{nk}(t)-\left(1+t^2\mathbb{E}\left[X_{nk}^2\right]\right)\right|=0.\end{aligned}$$

Sin embargo, para poder aplicar 2.1.9 con $z_i=\varphi_{ni}$, $w_i=1-\frac{t^2}{2}\sigma_{nk}^2$ hay que verificar que en efecto $|z_i|,|w_j|\leq 1$. Claramente $|\varphi_{ni}|\leq 1$ por ser función característica. Para comprobar que $|1-\frac{t^2}{2}\sigma_{nk}^2|\leq 1$ basta notar que como $\max_{k\in 1,\cdots,r_n}\{\sigma_{nk}^2\}\to 0$, entonces a partir de una n suficientemente grande, se va a satisfacer siempre que $|1-\frac{t^2}{2}\sigma_{nk}^2|\leq 1$. Con esto, concluimos que $A_n\to 0$.
Para concluir la prueba, solamente falta probar que

$$\lim_{n\to\infty}B_n=0$$

Para estudiar la convergencia a 0 de B_n, se aplica el lema 2.1.10 tomando $z=\frac{t^2}{2}\sigma_{nk}^2$:

$$\left|e^{\frac{t^2}{2}\sigma_{nk}^2}-\left(1-\frac{t^2}{2}\sigma_{nk}^2\right)\right|\leq|t|^4\frac{\sigma_{nk}^4}{4}e^{\frac{t^2}{2}\sigma_{nk}^2}.$$

Como $\lim_{n\to\infty}\max_{k\in 1,\cdots,r_n}\sigma_{nk}\to 0$, entonces para n suficientemente grande $e^{\frac{t^2}{2}\sigma_{nk}^2}<L$ para alguna $L\in\mathbb{R}^+$ y entonces:

$$\begin{aligned}\sum_{k=1}^{r_n}\left|e^{\frac{t^2}{2}\sigma_{nk}^2}-\left(1-\frac{t^2}{2}\sigma_{nk}^2\right)\right|&\leq|t|^4\sum_{k=1}^{r_n}\frac{\sigma_{nk}^4}{4}e^{\frac{t^2}{2}\sigma_{nk}^2}\\&\leq\frac{|t|^4}{4}\max_{k\in 1,\cdots,r_n}\{\sigma_{nk}^2\}L\sum_{k=1}^{r_n}\sigma_{nk}^2\\&=\frac{|t|^4L}{4}\max_{k\in 1,\cdots,r_n}\{\sigma_{nk}^2\}\to 0.\end{aligned}$$

Así que una vez más, al aplicar el lema 2.1.9, concluimos:

$$\lim_{n\to\infty}B_n=\lim_{n\to\infty}\left|\prod_{k=1}^{r_n}\left(1-\frac{t^2}{2}\sigma_{nk}^2\right)+\prod_{k=1}^{r_n}\varphi_{nk}\right|\leq\lim_{n\to\infty}\sum_{k=1}^{r_n}\left|e^{\frac{t^2}{2}\sigma_{nk}^2}-\left(1-\frac{t^2}{2}\sigma_{nk}^2\right)\right|=0.$$

■

2.2 Distribuciones infinitamente divisibles

Ya probamos los resultados de convergencia débil que necesitaremos en $E=\mathbb{R}$. En esta sección, nos seguiremos interesando en límites de arreglos triangulares similares a los considerados en el teorema de Lindberg. Bajo las condiciones del teorema de Lindberg, sabemos que el arreglo converge a una variable aleatoria normal, pero ¿qué pasa si relajamos las condiciones? ¿cuando podemos asegurar convergencia a una variable aleatoria? Lo que veremos es que bajo hipótesis menos estrictas podremos en muchos casos seguir garantizando la existencia de un límite en ley aunque ya no tendrá por qué ser una variable aleatoria normal. La familia de límites posibles forma una colección de particular interés y será pieza clave en el estudio de los procesos de Lévy.

Definición 2.2.1 Convergencia vaga
Sean μ_n, μ *medidas finitas en* $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ *tales que, para todo* $a, b \in \mathbb{R}$ *que satisfagan* $\mu(\{a\}) = \mu(\{b\}) = 0$ *(también llamados puntos de continuidad de* μ*) se cumple que :*

$$\lim_{n\to\infty} \mu_n\left((a,b]\right) = \mu\left((a,b]\right).$$

Entonces, diremos que μ_n *converge vagamente a* μ *y lo denotaremos* $\mu_n \xrightarrow{v} \mu$*.*

Si μ_n y μ son medidas de probabilidad, entonces $\mu_n \xrightarrow{\mathscr{L}} \mu$ es equivalente a $\mu_n \xrightarrow{v} \mu$. Sin embargo, esto no es cierto en general:
Si para toda $n \in \mathbb{N}$, $\mu_n = \delta_n$, donde $\delta_n(\mathbb{R}) = \delta_n(\{n\}) = 1$ y $\mu \equiv 0$, entonces $\mu_n[a,b) \to 0$ para todo $a, b \in \mathbb{R}$, por lo que $\mu_n \xrightarrow{v} \mu$ pero no se puede hablar de convergencia débil puesto que μ no es medida de probabilidad. Denotaremos por C_0 al conjunto de las funciones f continuas tales que $\lim_{|x|\to\infty} f(x) = 0$. Tenemos la siguiente caracterización de convergencia vaga:

Lema 2.2.2 *Si* $\mu_n \xrightarrow{v} \mu$ *y* $\sup_{n\in\mathbb{N}} \mu_n(\mathbb{R}) < M < \infty$*, entonces, para toda* $f \in C_0$ *se cumple que*

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu.$$

Demostración
Como $\sup_n \mu_n(\mathbb{R}) < M$, entonces $\mu(\mathbb{R}) < M$. Sea $f \in C_0$ y sean a y b puntos de continuidad de μ, tales que $|f(x)| \leq \frac{\epsilon}{M}$ si $x \notin A = (a,b]$. Podemos garantizar que existen pues $\lim_{|x|\to\infty} f(x) = 0$.
Por un lado:

$$\left|\int_{A^c} f d\mu_n\right| \leq \int_{A^c} |f| d\mu_n \leq \mu_n(A^c)\frac{\epsilon}{M} \leq \epsilon$$

por lo que, por el mismo razonamiento para μ, obtenemos que:

$$\left|\int_{A^c} f d\mu_n\right| \leq \epsilon \qquad \text{y} \qquad \left|\int_{A^c} f d\mu\right| \leq \epsilon \tag{2.2.1}$$

Por otro lado, si $\mu(A) > 0$, definimos las siguientes medidas de probabilidad en $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$:

$$\gamma(B) = \frac{\mu(A \bigcap B)}{\mu(A)} \quad y \quad \gamma_n = \frac{\mu_n(A \bigcap B)}{\mu_n(A)}.$$

Veamos que $\gamma_n \xrightarrow{v} \gamma$:
Si $B = (a_0, b_0]$ con a_0, b_0 dos puntos de continuidad de γ, entonces también son puntos de continuidad de μ y entonces:

$$\lim_{n\to\infty} \gamma_n(a_0, b_0] = \lim_{n\to\infty} \frac{\mu_n\left((a,b] \bigcap (a_0, b_0]\right)}{\mu_n(A)} = \frac{\mu\left((a,b] \bigcap (a_0, b_0]\right)}{\mu(A)} = \gamma(a_0, b_0]$$

La segunda igualdad se da pues por un lado, $\mu_n(A) \to \mu(A)$, y $(a,b] \bigcap (a_0, b_0]$ se puede ver como unión de 4 intervalos disjuntos de la forma $(x_i, y_i]$ con $x_i, y_i \in \{a, b, a_0, b_0\}$ y para cada uno de estos, $\mu_n(x_i, y_i] \to \mu(x_i, y_i]$. Se sigue de esto que $\gamma_n \xrightarrow{v} \gamma$ y por lo tanto $\gamma_n \xrightarrow{\mathscr{L}} \gamma$. Por el teorema (8.2.3):

$$\int f d\gamma_n \to \int f d\gamma \qquad \forall f \in C_b \text{ (y en particular, para toda } f \in C_0).$$

Además, como $d\gamma_n = \frac{\mathbb{1}_A}{\mu_n(A)} d\mu_n$ y $d\gamma = \frac{\mathbb{1}_A}{\mu(A)} d\mu$, se sigue que

$$\int f d\gamma_n = \frac{1}{\mu_n(A)} \int f \cdot \mathbb{1}_A d\mu_n \qquad \text{y} \qquad \int f d\gamma = \frac{1}{\mu(A)} \int f \cdot \mathbb{1}_A d\mu$$

y entonces

$$\lim_{n\to\infty} \int_A f d\mu_n = \int_A f d\mu \qquad \text{pues } \mu_n(A) \to \mu(A). \tag{2.2.2}$$

De (5) y (6) concluimos:

$$\left| \int f d\mu_n - \int f d\mu \right| \leq \left| \int_A f d\mu_n - \int_A f d\mu \right| + \left| \int_{A^c} f d\mu_n \right| + \left| \int_{A^c} f d\mu \right| < \epsilon.$$

Ahora, si $\mu(A) = 0$, se tiene que:

$$\lim_{n\to\infty} \left| \int_A f d\mu_n \right| \leq \lim_{n\to\infty} c\mu_n(A) = 0$$

por lo que también se puede concluir. ■

Lema 2.2.3 *Si $\{\mu_n\}$ es una sucesión de medidas finitas tales que* $\sup \mu_n(\mathbb{R}) < \infty$, *entonces existe una subsucesión $\{\mu_{n_k}\}$ y una medida μ tal que $\mu_n \xrightarrow{v} \mu$.*

Demostración
Sea $F_n(x) = \mu_n(-\infty, x]$. Como para cada x, la sucesión $\{F_n(x)\}$ está acotada, por el mismo argumento que en la prueba del teorema de Helly, existe una función continua por la derecha y no decreciente F y una subsucesión tal que $F_{n_k}(x) \to F(x)$ para toda $x \in \mathbb{R}$. Como F es continua por la derecha y no decreciente, tiene asociada una medida $\nu(a,b] = F(b) - F(a)$ y se sigue que $\mu_{n_k} \xrightarrow{\nu} \nu$.

Recordemos que si Z_1 y Z_2 son variables aleatorias reales independientes con función de distribución F_1 y F_2 respectivamente, la distribución F_3 de $Z_1 + Z_2$ está dada por

$$F_3(x) = F_2 * F_1(x) = \int_{-\infty}^{\infty} F_1(x-y) dF_2(dy).$$

Esto se puede consultar en Ash, [1] p328.

Definición 2.2.4 Infinitamente Divisible
Decimos que una variable aleatoria X con distribución Fes infinitamente divisible si para cada n, existen variables aleatorias independientes idénticamente distribuidas $\{X_k^{(n)}\}_{k=1}^n$ tales que

$$X \stackrel{\mathscr{L}}{=} X_1^{(n)} + \cdots + X_n^{(n)}.$$

Las siguientes formulaciones son entonces equivalentes:
·Para todo n, existe F_n función de distribución tal que

$$F = \underbrace{F_n * \cdots * F_n}_{n}.$$

·Para todo n, $\varphi_X = (\varphi_n)^n$ para φ_n función característica de una variable aleatoria.

Por esto, se habla tanto de una función característica φ, variable aleatoria X como de una distribución F infinitamente divisible.

Ejemplos:
1) Si X tiene distribución $\mathcal{N}(0,1)$,

$$\varphi_X(\lambda) = e^{-\frac{\lambda^2}{2}} = \prod_{k=1}^{n} e^{-\frac{\lambda^2}{2n}}.$$

Como $e^{-\frac{\lambda^2}{2n}}$ es la función característica de una variable aleatoria $\mathcal{N}(0,n)$, entonces $X = X_1^{(n)} + \cdots + X_n^{(n)}$ donde $X_k^{(n)}$ tiene distribución $\mathcal{N}(0,n)$ para toda k, por lo que X es infinitamente divisible.

2) Si X tiene distribución Poisson(c),

$$\varphi_X(\lambda) = e^{c(e^{i\lambda}-1)} = \prod_{k=1}^{n} e^{\frac{c}{n}(e^{i\lambda}-1)} = \prod_{k=1}^{n} \varphi_Y(\lambda)$$

donde Y es una variable aleatoria con distribución Poisson$(\frac{c}{n})$, por lo que X es infinitamente divisible. Obsérvese como en ambos casos, la función característica de la variable infinitamente divisible X se escribe a partir de una exponencial.

La clase de variables aleatorias infinitamente divisibles es robusta en muchos sentidos: si φ, φ_1 y φ_2 son funciones características de variables aleatorias infinitamente divisibles, entonces $\varphi_1\varphi_2$, $\bar{\varphi}_1$ y $|\varphi|^2$ vuelven a ser funciones características de variables infinitamente divisibles (donde por $\bar{\varphi}_1$ nos referimos a su conjugado complejo). La prueba es sencilla y se puede leer en Ash [1], sección 7.5. Un hecho más sorprendente es que es estable ante límites débiles:

Teorema 2.2.5 *Para cada n, sea X_n una variable aleatoria infinitamente divisible y supongamos que $X_n \xrightarrow{\mathscr{L}} X$ para alguna variable aleatoria X. Entonces, X es infinitamente divisible.*

Demostración
Fijamos $r \in \mathbb{N}$; como para cada n la variable aleatoria X_n es infinitamente divisible, existen $\{X_n^{(i,r)}\}_{i=1}^r$ variables aleatorias independientes e idénticamente distribuidas tales que

$$X_n \stackrel{\mathscr{L}}{=} X_n^{(1,r)} + X_n^{(2,r)} + \cdots + X_n^{(r,r)}.$$

Como $X_n \xrightarrow{\mathscr{L}} X$, por el teorema de continuidad de Lévy en $\mathbb{R}$ la sucesión $\{X_n\}$ es tensa; veamos que esto implica que la sucesión, $\{X_n^{(1,r)}\}_{n\in\mathbb{N}}$ también es tensa. Al ser $X_n^{(i,r)}$ $i \in \{1 \cdots r\}$ independientes e idénticamente distribuidas,

$$\left(\mathbb{P}\left(X_n^{(1,r)} > M\right)\right)^r = \mathbb{P}\left(X_n^{(1,r)} > M, \cdots, X_n^{(r,r)} > M\right) \leq \mathbb{P}\left(X_n > rM\right)$$

y similarmente

$$\left(\mathbb{P}\left(X_n^{(1,r)} < -M\right)\right)^r \leq \mathbb{P}\left(X_n < -rM\right).$$

Como la sucesión $\{X_n\}_n$ es tensa, podemos hacer ambas cotas arbitrariamente pequeñas y se sigue que la sucesión $\{X_n^{(1,r)}\}_n$ también es tensa. Por Prokhorov, es relatívamente compacta y debe existir una subsucesión que converge débilmente a una variable Y con distribución ν. Como este argumento es válido para cada $\{X_n^{(i,r)}\}_n$ con $i \in 1, \cdots, r$, construimos una subsucesión tal que para cada i, $X_{n_k}^{(i,r)} \xrightarrow{\mathscr{L}} Y$ donde Y tiene distribución ν. Por independencia, al aplicar el teorema de Lévy se sigue que

$$X_{n_k} \overset{\mathscr{L}}{=} X_{n_k}^{(1,r)} + X_{n_k}^{(2,r)} + \cdots + X_{n_k}^{(r,r)} \xrightarrow{\mathscr{L}} Y^{(1)} + \cdots + Y^{(r)}$$

donde $\{Y^{(i)}\}$ son iid con distribución ν. Pero recordemos que $X_n \xrightarrow{\mathscr{L}} X$, por lo que $X \overset{\mathscr{L}}{=} Y^{(1)} + \cdots + Y^{(r)}$ y al ser r arbitrario, concluimos que Y es infinitamente divisible. ■

Antes de pasar a los tres teoremas centrales de esta sección, probamos el siguiente resultado que usaremos más adelante, al definir el exponente característico de un proceso de Lévy:

Proposición 2.2.6 *Si φ es una función característica infinitamente divisible, entonces φ nunca se anula.*

Demostración
Al ser φ infinitamente divisible, $\varphi = (\varphi_n)^n$ para cada n. Primero, supongamos que para cada n, φ y φ_n son reales no-negativas. Entonces, podemos escribir $\varphi_n = \varphi^{\frac{1}{n}}$ y φ_n converge puntualmente a una función g que solamente puede tomar los valores 0 y 1:

$$\lim_{n\to\infty} \varphi_n(\lambda) = \lim_{n\to\infty} \varphi^{\frac{1}{n}}(\lambda) = \begin{cases} 1 & \text{si } \varphi(\lambda) > 0 \\ 0 & \text{si } \varphi(\lambda) = 0. \end{cases}$$

Como $\varphi(0) = 1$, por continuidad $\varphi(\lambda) > 0$ en una vecindad abierta B de 0 y por lo tanto, $g(\lambda) = 1$ en B. Al ser g continua en 0 y límite puntual de una sucesión de funciones características (φ_n), por el teorema deLévy g es una función característica y en particular debe ser continua. Como g solo puede tomar los valores $\{0, 1\}$ y $g(0) = 1$, se sigue que $g \equiv 1$. Obtenemos que para cualquier u, $\varphi_n(\lambda) \to 1$ y en particular $\varphi_n(\lambda) > 0$ a partir de n suficientemente grande. Como $\varphi(\lambda) = (\varphi_n(\lambda))^n$, concluimos que $\varphi(\lambda) > 0$.

Ahora, sean φ y φ_n funciones características arbitrarias. Si $\varphi = (\varphi_n)^n$, entonces $|\varphi|^2 = |\varphi_n|^{2n} = (|\varphi_n|^2)^n$ donde tanto $|\varphi|^2$ como $|\varphi_n|^2$ son funciones características infinitamente divisibles por la discusión que se tuvo previo al teorema 2.2.5, pero además son reales no negativas. Al aplicar el razonamiento anterior a las características $\psi = |\varphi|^2$ y $\psi_n = |\varphi_n|^2$, obtenemos el resultado en el caso general.

Teorema 2.2.7 *Supongamos que*

$$\varphi(\lambda) = \exp\left\{ \int_{\mathbb{R}\setminus 0} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} \mu(dx) - \mu(\{0\}) \frac{\lambda^2}{2} \right\} \tag{2.2.3}$$

para alguna medida μ finita. Entonces φ es la función característica de una variable aleatoria infinitamente divisible.

Observaciones: Por el lema 2.1.8 al considerar $n = 1$, tenemos que

$$|e^{i\lambda x} - 1 - i\lambda x| \leq \frac{1}{2}(\lambda x)^2$$

por lo que al dividir entre $1/x^2$, obtenemos que el integrando esta acotado para toda x por $\lambda^2/2$ y por lo tanto es integrable. De hecho, esto nos permite aplicar teorema deconvergencia dominada y deducir que $\varphi(\lambda)$ definida en (2.2.3) es continua en 0. Esto se usará más adelante cuando probemos que es una función característica.

Una vez más por 2.1.8 tomando $n = 2$, tenemos

$$|\left(e^{i\lambda x} - 1 - i\lambda x\right) + \frac{(tx)^2}{2}| \leq \frac{1}{6}|x|^3$$

por lo que al dividir entre x^2, concluimos que $\lim_{x\to 0} \frac{e^{i\lambda x}-1-i\lambda x}{x^2} = -\frac{\lambda^2}{2}$. Entonces la expresión (2.2.3) también se puede escribir de forma mas compacta:

$$\varphi(\lambda) = \exp\left\{\int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2}\mu(dx)\right\}.$$

Se usará cualquiera de las dos expresiones según sea conveniente. Antes de pasar a la prueba, regresemos a los dos ejemplos que estudiamos antes de enunciar el teorema y hagamos explícita a la medida μ:

1) Si $\mu = \delta_0\sigma^2$ la medida puntual en 0 con valor σ^2, entonces $\varphi(\lambda) = e^{-\frac{t^2\sigma^2}{2}}$ es la característica de una variable aleatoria con ley $\mathcal{N}(0, \sigma^2)$ que en efecto, vimos ya que era infinitamente divisible.

2) Si $\mu = \delta_a ca^2$ la masa puntual en $a \neq 0$ con valor $c \cdot a^2$, entonces

$$\varphi(\lambda) = \exp\left\{\frac{e^{i\lambda a} - 1 - i\lambda a}{a^2} \cdot ca^2\right\} = e^{c\left(e^{i\lambda a}-1\right)}e^{-ic\lambda a}$$

que es la función característica de $Z = a(Z^c - c)$ donde Z^c es una variable aleatoria con distribución Poisson(c). Es infinitamente divisible ya que para cada n,

$$\varphi_Z(\lambda) = e^{c\left(e^{i\lambda a}-1\right)}e^{-ic\lambda a} = \prod_{k=1}^{n} e^{\frac{c}{n}\left(e^{i\lambda a}-1\right)}e^{-i\frac{c}{n}\lambda a} \tag{2.2.4}$$

donde $e^{\frac{c}{n}\left(e^{i\lambda a}-1\right)}e^{-i\frac{c}{n}\lambda a}$ es la característica de $a\left(Z^{\frac{c}{n}} - \frac{c}{n}\right)$.

Demostración

Empezaremos probando que en efecto φ es una función característica. Aproximaremos a la medida μ por una sucesión de medidas atómicas y probaremos que convergen a μ vagamente. A partir de la medida μ, definimos la siguiente sucesión de medidas: para cada k, μ_k es la medida con masa puntual $\mu(\frac{j}{2^k}, \frac{j+1}{2^k}]$ en el punto $\frac{j}{2^k}$ para $j \in I^k = \{0, \pm 1, \pm 2, \cdots, \pm 2^{2^k}\}$, es decir:

$$\mu_k = \sum_{j\in I^k} \delta_{\frac{j}{2^k}}\mu\left(\frac{j}{2^k}, \frac{j+1}{2^k}\right]$$

Primero, probaremos que $\mu_k \xrightarrow{v} \mu$. Para esto, basta probar que $\mu_n(a, b] \to \mu(a, b]$ para todo a, b en un denso (esto no es difícil de verificar). Como los diádicos $\{\frac{j}{2^k} \,|\, j \in \mathbb{Z},,\ k \in \mathbb{N}\}$ son un denso, probaremos que:

$$\lim_{n\to\infty} \mu_n\left(\frac{j}{2^k}, \frac{l}{2^m}\right] = \mu\left(\frac{j}{2^k}, \frac{l}{2^m}\right]$$

Nótese que si $m < k$, entonces cualquier intervalo de la forma $\left(\frac{j}{2^k}, \frac{\tilde{l}}{2^m}\right]$ se puede reescribir al multiplicar numerador y denominador por una potencia de 2 adecuada como $\left(\frac{j}{2^k}, \frac{l}{2^k}\right]$ por lo que podemos suponer que los intervalos están compuestos por diádicos con mismo denumerador 2^k. Ahora, al considerar enteros n suficientemente grandes de modo que $n > k$,

$$\begin{aligned}\mu_n\left(\frac{j}{2^k}, \frac{l}{2^k}\right] = \mu_n\left(\frac{\tilde{j}}{2^n}, \frac{\tilde{l}}{2^n}\right] &= \sum_{\tilde{j}<i\leq\tilde{l}} \mu_n\left(\left\{\frac{i}{2^i}\right\}\right)\\ &= \sum_{\tilde{j}<i\leq\tilde{l}} \mu\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right] = \mu\left(\frac{\tilde{j}}{2^n}, \frac{\tilde{l}}{2^n}\right] = \mu\left(\frac{j}{2^k}, \frac{l}{2^k}\right].\end{aligned}$$

Es decir, si $n > k$, la sucesión $\{\mu_n\left(\frac{j}{2^k}, \frac{l}{2^k}\right]\}_n$ se vuelve constante $\mu_n\left(\frac{j}{2^k}, \frac{l}{2^k}\right] = \mu\left(\frac{j}{2^k}, \frac{l}{2^k}\right]$ para toda n, por lo que

$$\lim_{n\to\infty} \mu_n\left(\frac{j}{2^k}, \frac{l}{2^k}\right] = \mu\left(\frac{j}{2^k}, \frac{l}{2^k}\right].$$

Con esto, concluimos que $\mu_n \xrightarrow{v} \mu$.
Ahora, veamos que φ_{μ_n} definida como sigue es función característica:

$$\varphi_{\mu_n}(\lambda) := \exp\left\{\int_{\mathbb{R}\setminus 0} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2}\mu_k(dx) - \mu_k(\{0\})\frac{\lambda^2}{2}\right\}$$

Obsérvese que φ_{μ_n} NO es la función característica de la medida μ_n, así que hay un ligero abuso de notación al nombrarla de este modo. Se trata de (2.2.3) al considerar a $\mu = \mu_k$.
Como μ_k solo tiene masa puntual $\mu\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right]$ en los diádicos $\frac{i}{2^n}$,

$$\begin{aligned}\varphi_{\mu_n}(\lambda) &= \exp\left\{\sum_{i\in I^n\setminus 0} \frac{e^{i\lambda\frac{i}{2^n}} - 1 - i\lambda\frac{i}{2^n}}{\left(\frac{i}{2^n}\right)^2}\mu\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right] - \mu\left(0, \frac{1}{2^n}\right]\frac{\lambda^2}{2}\right\}\\ &= \prod_{i\in I^n\setminus 0} \exp\left\{\frac{e^{i\lambda\frac{i}{2^n}} - 1 - i\lambda\frac{i}{2^n}}{\left(\frac{i}{2^n}\right)^2}\mu\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right]\right\}\exp\left\{-\mu\left(0, \frac{1}{2^n}\right]\frac{\lambda^2}{2}\right\}\end{aligned}$$

Aunque la expresión es intimidante, nótese que en realidad está compuesta por términos conocidos: por los dos ejemplos previos, cada término de la forma

$$\exp\left\{\frac{e^{i\lambda\frac{i}{2^n}} - 1 - i\lambda\frac{i}{2^n}}{\left(\frac{i}{2^n}\right)^2}\mu\left(\frac{i}{2^n}, \frac{i+1}{2^n}\right]\right\}$$

es la característica de una variable aleatoria de la forma $a(Z^c - c)$, con Z^c variable Poisson(c) para constantes apropiadas a y c, mientras que

$$\exp\left\{-\mu\left(0, \frac{1}{2^n}\right]\frac{\lambda^2}{2}\right\}$$

es la característica de una variable aleatoria normal. Entonces φ_{μ_n} es en efecto una función característica de una variable aleatoria que se descompone como suma de variables aleatorias independientes.

Con esto, concluimos que así definida, φ_{μ_n} es función característica. Bastaría probar que $\varphi_{\mu_k}(\lambda) \to \varphi(\lambda)$ para poder concluir por el teorema de Lévy que φ también lo es.
Para λ fijo, $\lim_{x\to 0} \frac{e^{i\lambda x}-1-i\lambda x}{x^2} = -\frac{\lambda^2}{2} < \infty$, por lo que, si:

$$f(x) = \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} = Re(f) + iIm(f)$$

entonces $Re(f)$, $Im(f) \in C_0$. Por el lema (2.2.2), $\int f d\mu_n \to \int f d\mu$. Finalmente, podemos concluir: como

$$\lim_{n\to\infty} \varphi_{\mu_k}(\lambda) = \varphi(\lambda) \quad \text{para todo } \lambda$$

y al ser φ continua en 0, es función característica. Es claramente infinitamente divisible pues, si definimos $\mu_i^{(n)} := \frac{1}{n}\mu$, entonces $\varphi = \prod_{k=1}^n \varphi_{\mu_1^{(n)}}$ donde $\varphi_{\mu_1^{(n)}}$ es la característica (2.2.3) con medida $\mu = \mu_1$. ■

Sea $\{X_{n1}, \cdots X_{nr_n}\}_{n\in\mathbb{N}}$ un arreglo triangular, cada renglón del arreglo está compuesto por variables aleatorias independientes tales que,

$$\sigma_{nk}^2 = \mathbb{E}\left[X_{nk}^2\right] < \infty \qquad \mathbb{E}\left[X_{nk}\right] = 0 \qquad s_n^2 = \sum_{k=1}^{r_n} \sigma_{nk}^2 > 0 \tag{2.2.5}$$

y sea $S_n = \sum_{k=1}^{r_n} X_{nk}$.

Teorema 2.2.8 *Toda distribución infinitamente divisible con media 0 y varianza finita es límite de S_n para algún arreglo triangular que satisface (2.2.5) y que además cumple las siguientes propiedades:*
a) $\sup_n\{s_n^2\} < \infty$
b) $\lim_{n\to\infty} \max_{k\leq r_n} \sigma_{nk}^2 = 0$.

La idea detrás de *b*) es asegurarse que cada elemento X_{nk} del arreglo aporta poco a la suma S_n conforme crece n: al tener media 0, por Markov

$$\max_{k\leq r_n} \mathbb{P}\left(|X_{nk}| \geq \epsilon\right) \leq \max_{k\leq r_n} \sigma_{nk}^2/\epsilon^2 \to 0.$$

Para probar 2.2.8 necesitaremos el siguiente resultado:

Proposición 2.2.9 *Si X y Y son variables aleatorias independientes tales que $X+Y$ tiene segundo momento finito, entonces X y Y también tienen segundos momentos finitos.*

Demostración
Como $X^2 + Y^2 \leq (X+Y)^2 + 2|XY|$, entonces basta probar que $|XY|$ es integrable. Pero al ser independientes, $\mathbb{E}\left[|X||Y|\right] = \mathbb{E}\left[|X|\right]\mathbb{E}\left[|Y|\right]$ por lo que solo necesitamos probar que X y Y son integrables. Como $|Y| = |Y + x - x| \leq |x| + |x+Y|$, entonces $\mathbb{E}\left[|Y|\right] = \infty$ implicaría que $\mathbb{E}\left[|x+Y|\right] = \infty$ para cada x. Supongamos entonces que $\mathbb{E}\left[|Y|\right] = \infty$. Por independencia y Fubini:

$$\begin{aligned}
\mathbb{E}\left[|X+Y|\right] &= \int_{\mathbb{R}^d\times\mathbb{R}^d} (x+y)\, d\mathbb{P}_{(X,Y)}(dx\times dy) \\
&= \int_{\mathbb{R}^d\times\mathbb{R}^d} (x+y)\, d\mathbb{P}_X(dx)\otimes\mathbb{P}_Y(dy) \\
&= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} (x+y)\, d\mathbb{P}_Y(dx) d\mathbb{P}_X(dy) = \int_{\mathbb{R}^d} \mathbb{E}\left[|x+Y|\right] d\mathbb{P}_X = \infty
\end{aligned}$$

lo cual es una contradicción. Concluimos que $\mathbb{E}[|X|]$, $\mathbb{E}[|Y|] < \infty$. No será relevante pero nótese que también estamos probando que si $X + Y \in L_1$ con X,Y independientes entonces X,$Y \in L_1$. ■

Corolario 2.2.10 *Si $\{X_k\}_{k=1}^n$ son variables aleatorias independientes y $(X_1 + \cdots + X_n) \in L_2$, entonces $X_i \in L_2$ para toda i.*

Con esto, ya podemos probar el teorema (2.2.8).

Demostración
Como X es infinitamente divisible, entonces para cada n, existen variables aleatorias independientes idénticamente distribuidas $\{X_{nk}\}_{k=1}^n$ tales que:

$$X \overset{\mathscr{L}}{=} \sum_{k=1}^{n} X_{nk} =: S_n$$

Consideremos entonces el arreglo triangular formado por estas $\{X_{n1}, \cdots, X_{nn}\}_{n\in\mathbb{N}}$. Para toda n, $X \overset{\mathscr{L}}{=} S_n$ por lo que $S_n \overset{\mathscr{L}}{\to} X$. Falta probar que este arreglo satisface las condiciones (2.2.5), (a) y (b).
Satisface (2.2.5):
$\mathbb{E}[X_{nk}] = 0$ pues $0 = \mathbb{E}[X] = \mathbb{E}[\sum^n X_{nk}] = n\mathbb{E}[X_{nk}]$ para cualquier k por ser idénticamente distribuidas. Similarmente, $\sigma_{nk}^2 < \infty$ puesto que $nVar(X_{nk}) = Var(X) < \infty$ para cualquier k.
Satisface (a) y (b):
Para toda n, $s_n^2 = \sum^n \sigma_{nk}^2 = Var(\sum^n X_{nk}) = Var(X) < \infty$ y entonces $\sup_n s_n^2 < \infty$ por lo que se cumple (a). Además, si $\sigma^2 = Var(X)$, entonces para toda n, $\sigma^2 = \sum^n \sigma_{nk}^2 = n\sigma_{n_1}$ al ser idénticamente distribuidas. Pero entonces,

$$\lim_{n\to\infty} n\sigma_{n1}^2 = \sigma^2 < \infty$$

por lo que $\sigma_{n1}^2 \to 0$ y se sigue que $\max_{k\leq n} \sigma_{nk}^2 \to 0$ una vez más al ser idénticamente distribuidas.
■

Teorema 2.2.11 *Si X es el límite de un arreglo triangular $\{X_{n1} \cdots X_{nr_n}\}_{n\in\mathbb{N}}$ que satisface la propiedad (2.2.5):*

$$\sigma_{nk}^2 = \mathbb{E}\left[X_{nk}^2\right] < \infty \qquad \mathbb{E}[X_{nk}] = 0 \qquad 0 < s_n^2 = \sum_{k=1}^{r_n} \sigma_{nk}^2$$

y las propiedades (a) y (b):

$$\sup\{s_n^2\} < \infty, \qquad \lim_{n\to\infty} \max_{k\leq r_n} \sigma_{nk}^2 = 0$$

entonces la función característica de X es de la forma:

$$\varphi(\lambda) = \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} \mu(dx) \right\}$$

para alguna medida finita μ y por lo tanto es infinitamente divisible.

Nótese que los límites S_n/s_n que consideramos bajo las hipótesis del teorema de Lindberg 2.1.7 a partir de arreglos triangulares $\{X_{n1}, \cdots, X_{nr_n}\}$ caen dentro de estas. Esto se vuelve más claro al normalizar el arreglo: del inicio de la prueba del teorema de Lindberg, el arreglo normalizado $X'_{nk} = \frac{X_{nk}}{s_n}$, $S'_n = \frac{S_n}{s_n}$ satisface las condiciones (2.2.5), cumple (b) (recordemos que esta era una de las consecuencias de la condición de Lindberg al normalizar) y (a) pues $s'_n = 1$. Entonces, el límite en ley de S_n/s_n es el mismo que el de S'_n, y en efecto su límite es infinitamente divisible pues se trata de una variable aleatoria normal. Lo que estamos obteniendo es que al relajar las hipótesis del teorema de Lindberg, los límites posibles de S_n son la clase de las variables aleatorias infinitamente divisibles.

Demostración

Denotaremos por F_{nk} y $\varphi_{X_{nk}}$ a la distribución y función característica respectivamente de X_{nk}. Puesto que $S_n \overset{\mathscr{L}}{\to} X$, $\varphi_{S_n} \to \varphi_X$ y por independencia de los renglones del arreglo triangular, $\varphi_{S_n} = \prod_{k=1}^{r_n} \varphi_{X_{nk}}$. Además, se satisface

$$\lim_{n\to\infty} \left| \prod_{k=1}^{r_n} \varphi_{X_{nk}}(\lambda) - \exp\left\{ \sum_{k=1}^{r_n} (\varphi_{X_{nk}} - 1) \right\} \right| = 0. \tag{2.2.6}$$

Como la prueba de este teorema es laboriosa, la justificación de (2.2.6) se deja al final. Nótese que:

$$\sum_{k=1}^{r_n} (\varphi_{X_{nk}} - 1) = \sum_{k=1}^{r_n} \int_{\mathbb{R}} \left(e^{itx} - 1\right) dF_{nk}(dx) = \sum_{k=1}^{r_n} \int_{\mathbb{R}} \left(e^{itx} - 1 - itx\right) dF_{nk}(dx) \tag{2.2.7}$$

pues $\int_{\mathbb{R}} itx dF_{nk}(dx) = it\mathbb{E}[X_{nk}] = 0$.

Definimos la siguiente medida:

$$\mu_n(-\infty, x] = F_n(x) := \sum_{k=1}^{r_n} \int_{-\infty}^{x} y^2 dF_{nk}(dy). \tag{2.2.8}$$

Es finita pues $\mu_n(\mathbb{R}) = \sum_{k=1}^{r_n} \int_{\mathbb{R}} y^2 dF_{nk} = \sum_{k=1}^{r_n} \sigma_{nk}^2 = s_n^2 < \infty$. Por el teorema 2.2.7, la siguiente función es función característica de una variable aleatoria infinitamente divisible:

$$\begin{aligned}
\varphi_n &= \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - itx}{x^2} \mu_n(dx) \right\} \\
&= \exp\left\{ \sum_{k=1}^{n} \int_{\mathbb{R}} \left(e^{i\lambda x} - 1 - itx\right) dF_{nk}(dx) \right\} \\
&= \exp\left\{ \sum_{k=1}^{r_n} (\varphi_{nk} - 1) \right\}
\end{aligned}$$

donde la última igualdad es consecuencia de (2.2.7) y entonces el límite (2.2.6) se reescribe como sigue:

$$\lim_{n\to\infty} \left| \prod_{k=1}^{r_n} \varphi_{X_{nk}}(\lambda) - \varphi_n(\lambda) \right| = 0. \tag{2.2.9}$$

Como $\mu_n(\mathbb{R}) = s_n^2$, entonces $\sup_n \mu_n(\mathbb{R}) = \sup_n s_n^2 < \infty$ por hipótesis. Por el lema 2.2.3, existe una subsucesión $\{\mu_{n_j}\}$ de $\{\mu_n\}$ que converge vagamente a una medida finita ν. Por el teorema 2.2.2, para toda función $f \in C_0$, $\int f d\mu_{n_j} \to \int f d\nu$, así que:

$$\lim_{j\to\infty} \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} d\mu_{n_j}(dx) = \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} d\nu(dx)$$

y se sigue que:

$$\begin{aligned} \lim_{j\to\infty} \varphi_{n_j} &= \lim_{j\to\infty} \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} d\mu_{n_j}(dx) \right\} \\ &= \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} d\nu(dx) \right\}. \end{aligned}$$

Si definimos:

$$\psi(\lambda) = \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} d\nu(dx) \right\}$$

entonces lo que acabamos de probar es que $|\varphi_{n_j} - \psi| \to 0$ y por lo tanto,

$$\left|\varphi_{S_{n_j}} - \psi\right| = \left|\prod_{k=1}^{r_{n_j}} \varphi_{X_{n_j k}} - \psi\right| \le \left|\prod_{k=1}^{r_{n_j}} \varphi_{X_{n_j k}} - \varphi_{n_j k}\right| + \left|\varphi_{n_j k} - \psi\right|$$

lo cual tiende a 0 por (2.2.9) y lo último. Concluimos entonces que $\varphi_{S_{n_j}} \to \psi$; como por hitpótesis $\varphi_{S_n} \to \varphi_X$, entonces $\varphi_X = \psi$ y con esto concluye la prueba. Solamente falta probar el límite (2.2.9); nótese que

$$\begin{aligned} \left|\prod_{k=1}^{r_n} \varphi_{X_{nk}}(\lambda) - \exp\left\{ \sum_{k=1}^{r_n} (\varphi_{X_{nk}} - 1) \right\}\right| &= \left|\prod_{k=1}^{r_n} \varphi_{X_{nk}}(\lambda) - \prod_{k=1}^{r_n} \exp\left\{ (\varphi_{X_{nk}} - 1) \right\}\right| \\ &\le \sum_{k=1}^{r_n} \left|\varphi_{X_{nk}} - \exp\{\varphi_{X_{nk}} - 1\}\right|. \end{aligned}$$

La última desigualdad es consecuencia del lema (2.1.9): recordemos que si $|z_i|, |w_i| \le 1$ entonces $|\prod w_i - \prod z_i| \le \sum |w_i - z_i|$. Aquí, claramente $|\varphi_{X_{nk}}| \le 1$. Además, si $|z| \le 1$, $|Re(z)| < |z| < 1$ y entonces $|e^{z-1}| = e^{Re(z-1)} = e^{Re(z)-1} \le 1$ ya que $Re(z) - 1 \in [-1, 0]$. Una vez más, como $z = \varphi_{X_{nk}}$ satisface $|z| < 1$, podemos usar el lema 2.1.9. Veamos que este último término tiende a 0. Si definimos $\Theta_{nk}(t) := \varphi_{X_{n_k}}(t) - 1$, entonces:

$$\begin{aligned} |\Theta_{nk}(t)| = |\varphi_{X_{nk}}(t) - 1| = \left|\mathbb{E}\left[e^{itX_{nk}} - 1\right]\right| &\le \left|\mathbb{E}\left[e^{itX_{nk}} - 1 - itX_{nk}\right]\right| \\ &\le \mathbb{E}\left[\min\left\{\frac{|tX_{nk}|^2}{2}, ct|X_{nk}|\right\}\right] \le \frac{t^2}{2}\sigma_{nk}^2 \end{aligned}$$

por lo que, para cada t,

$$\max_{k\le r_n} |\Theta_{nk}(t)| = \max_{k\le r_n} |\varphi_{nk}(t) - 1| \le \frac{t^2}{2} \max_{k\le r_n} \sigma_{nk}^2 \to 0. \tag{2.2.10}$$

Además,

$$\begin{aligned}\sum_{k=1}^{r_n} |\Theta_{nk}(t)| &= \sum_{k=1}^{r_n} |\varphi_{X_{nk}}(t) - 1| \\ &\leq \frac{t^2}{2} \sum_{k=1}^{r_n} \sigma_{nk}^2 \\ &\leq \frac{t^2}{2} \sup_n s_n^2 < \infty \quad \text{para toda } n \end{aligned} \tag{2.2.11}$$

Por (2.2.11), (2.2.10) y Lema 2.1.10 podemos concluir, para n suficientemente grande, que:

$$\begin{aligned}\sum_{k=1}^{r_n} |\varphi_{X_{nk}} - \exp\{\varphi_{X_{nk}} - 1\}| &= \sum_{k=1}^{r_n} \left|1 + \Theta_{nk} - e^{\Theta_{nk}}\right| \\ &\leq e^2 \sum_{k=1}^{r_n} |\Theta_{nk}|^2 \\ &\leq e^2 \max_{k \leq r_n} |\Theta_{nk}| \cdot \sum_{k=1}^{r_n} |\Theta_{nk}| \longrightarrow 0. \end{aligned}$$

■

2.3 Elementos de convergencia débil en $\mathbb{R}^d$ y espacios métricos

Referencias: Bilingsley [7] capítulo 5 y Bilinsgley [8] capítulo 1

A lo largo de esta sección, generalizaremos algunos resultados que se vieron en la sección anterior y probaremos el teorema del Portemanteau, el cual necesitaremos más adelante cuando abordemos la convergencia de procesos. Iniciamos generalizando a $\mathbb{R}^d$ el concepto de función de distribución:

Si X es una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ con valores en $\mathbb{R}^d$ o en un espacio métrico E, la ley de X denotada indistintamente por $\mathbb{P}_X$ ó μ_X es la medida de probabilidad inducida por X en $\mathbb{R}^d$ (resp. E), $\mathbb{P}_X(B) = \mathbb{P}(X \in B)$. Al restringirnos a $\mathbb{R}^d$, podemos generalizar el concepto de función de distribución; si $u = (u_1, \cdots, u_d) \in \mathbb{R}^d$, definimos la función de distribución de X como sigue:

$$F_X(u) = \mathbb{P}\left(X_1 \leq u_1, \cdots, X_d \leq u_d\right) = \mathbb{P}_X\left((-\infty, u_1] \times \cdots, \times(-\infty, u_d]\right)$$

Diremos que $x < y$ para x, $y \in \mathbb{R}^d$ si para toda i, $x_i < y_i$. Con esta notación, $\{x \in \mathbb{R}^d \,|\, a < x \leq b\} =: (a, b]$ es el rectángulo encerrado entre los vértices a y b. Denotaremos $x \uparrow a$ (resp $x \downarrow a$) siempre que $x_i \uparrow a_i$ (resp $x_i \downarrow a_i$) para toda $i \in \{1 \cdots d\}$.
Así definida, F es no decreciente y por continuidad por arriba de la medida μ_X, F es continua por la derecha en el siguiente sentido:

$$\lim_{y \downarrow x} F(y) = F(x).$$

Proposición 2.3.1 *Sea μ una medida de probabilidad en con valores en $\mathbb{R}^d$ y F su distribución $F(x) = \mu(-\infty, x]$. Entonces, F es continua en x si y solo si $\mu\left(\partial(-\infty, x]\right) = 0$.*

Demostración
Primero, si F es continua en x:

$$\mu(-\infty, x) = \lim_{y\uparrow x} F(y) = F(x) = \mu(-\infty, x]$$

por lo que $\mu(\partial(-\infty, x]) = 0$.
Ahora, supongamos $\mu(\partial(-\infty, x]) = 0$. Obsérvese que

$$\lim_{y\uparrow x} F(y) = \lim_{y\uparrow x} \mu(-\infty, y] = \mu(-\infty, x) = \mu(-\infty, x] = F(x) = \lim_{y\downarrow x} F(y)$$

por continuidad por la derecha de F. Entonces existen $x_1 < x < x_2$ tales que $|F(x) - F(x_1)| < \epsilon$ y $|F(x_2) - F(x)| < \epsilon$. Si $y \in (x_1, x_2]$, como F es creciente, concluimos que $|F(x) - F(y)| < \epsilon$, por lo que F es continua en x. ■

Teorema 2.3.2 Portemanteau en espacios métricos (parcial)
Si $\{X_n\}_{n\in\mathbb{N}}$ y X son variables aleatorias con valores en un espacio métrico $(E, \mathscr{B}(E), d)$, entonces las siguientes condiciones son equivalentes, y por lo tanto son equivalentes a que $X_n \xrightarrow{\mathscr{L}} X$:
1-$\lim_{n\to\infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)]$ *para toda función* $f : E \to \mathbb{R}$ *continua y acotada.*

2- $\limsup_{n\to\infty} \mu_{X_n}(K) \leq \mu_X(K)$ *para todo K cerrado.*

3-$\liminf_{n\to\infty} \mu_{X_n}(G) \geq \mu_X(G)$ *para todo G abierto.*

4-$\lim_{n\to\infty} \mu_{X_n}(A) = \mu_X(A)$ *si* $\mu_X(\partial A) = 0$.

Si $E = \mathbb{R}^d$ y F es la distribución de μ_X se tiene además:
5-$\lim_{n\to\infty} F_{X_n}(u) = F_X(u)$ *para todo punto u de continuidad de F_X.*

Demostración
$1 \Rightarrow 2$. La idea consiste en aproximar a $\mathbb{1}_K$ por una sucesión de funciones continuas f_ϵ y pasar al límite. Para $\epsilon > 0$, definimos a $f_\epsilon(x) = (1 - d(x, K))^+$. Es acotada y nótese que $f_\epsilon(x) = 0$ si $x \notin K^\epsilon$ mientras que $f_\epsilon(x) = 1$ si $x \in K$. Al proceder por casos, no es difícil comprobar que es continua y de hecho que $|f_\epsilon(x) - f_\epsilon(y)| \leq d(x, y)$. Se sigue entonces que

$$\mathbb{1}_K(x) \leq f_\epsilon(x) = (1 - d(x, K))^+ \leq \mathbb{1}_{K^\epsilon}(x).$$

Como K es un cerrado, $K^\epsilon \downarrow K$ y por lo tanto, $f_\epsilon(x) \downarrow \mathbb{1}_K(x)$ conforme $\epsilon \downarrow 0$. Al integrar con respecto a μ_{X_n} y pasar al límite, se sigue entonces que

$$\limsup_{n\to\infty} \mu_{X_n}(K) = \limsup_{n\to\infty} \mathbb{P}(X_n \in K) \leq \lim_{n\to\infty} \mathbb{E}[f_\epsilon(X_n)] = \mathbb{E}[f_\epsilon(X)].$$

Ahora, por el teorema de convergencia dominada,

$$\limsup_{n\to\infty} \mu_{X_n}(K) \leq \lim_{\epsilon\downarrow 0} \mathbb{E}[f_\epsilon(X)] = \mathbb{P}(X \in K) = \mu_X(K).$$

$2 \Leftrightarrow 3$. Se prueba al tomar complementos.
2 y $3 \Rightarrow 4$. Sea A en $\mathscr{B}(E)$ tal que $\mu_X(\partial A) = 0$. Obsérvese que en este caso $\mu_X(\overline{A}) = \mu_X(A) = \mu_X(\mathring{A})$. Por un lado,

$$\limsup \mu_{X_n}(A) \leq \limsup \mu_{X_n}(\overline{A}) \overset{2}{\leq} \mu_X(\overline{A}) = \mu_X(A)$$

y por otro lado,

$$\liminf \mu_{X_m}(A) \geq \liminf \mu_{X_n}(\mathring{A}) \overset{3}{\geq} \mu_X(\mathring{A}) = \mu_X(A).$$

Entonces,

$$\mu_X(A) \leq \liminf \mu_{X_n}(A) \leq \limsup \mu_{X_n}(A) \leq \mu_X(A)$$

por lo que $\lim_{n\to\infty} \mu_{X_n} = \mu_X(A)$.

$4 \Rightarrow 1$. Sea f continua y acotada, tal que $|f| \leq K$. Existen $\alpha_0 < \cdots < \alpha_l$ reales tales que para toda $i \in \{1 \cdots l\}$:
(i) $\alpha_1 < -K < K < \alpha_l$
(ii) $\alpha_i - \alpha_{i-1} < \epsilon$
(iii) $\mathbb{P}\left(f(X) = \alpha_i\right) = \mu_X(x \in E \,:\, f(x) = \alpha_i) = 0$
Esta última condición se puede lograr pues los conjuntos $\{x \in E \,:\, f(x) = \alpha\}_{\alpha\in\mathbb{R}}$ son ajenos y por lo tanto hay a lo más una cantidad numerable de reales α para los cuales $\mu_X(\{x \in E \,:\, f(x) = \alpha\}) > 0$. Partimos el dominio de f como sigue: sea $A_i := \{x \,|\, \alpha_{i-1} < f(x) < \alpha_i\}$. Entonces, $\partial A_i \subset \{x \in E \,:\, f(x) = \alpha_{i-1}\} \bigcup \{x \in E \,:\, f(x) = \alpha_i\}$ por lo que $\mu_X(\partial A_i) = 0$. Así, $\{A_i\}$ es una partición del dominio de f. Por lo tanto:

$$\begin{aligned}
\left|\sum_{i=1}^{l} \alpha_i \mathbb{P}\left(X_n \in A_i\right) - \int_\Omega f(X_n) d\mathbb{P}\right| &= \left|\sum_{i=1}^{l} \int_{A_i} \alpha_i \mu_{X_n}(dx) - \sum_{i=1}^{l} \int_{A_i} f(x) \mu_{X_n}(dx)\right| \\
&\leq \sum_{i=1}^{l} \int_{A_i} |\alpha_i - f(x)| \mu_{X_n}(dx) \\
&\leq \epsilon \mu_{X_n}\left(\bigcup_i A_i\right) = \epsilon.
\end{aligned}$$

Como lo anterior es también válido al remplazar X_n por X, concluimos:

$$\begin{aligned}
\left|\int_\Omega f(X_n) d\mathbb{P} - \int_\Omega f(X) d\mathbb{P}\right| \leq& \left|\int_\Omega f(X_n) d\mathbb{P} - \sum_{i=1}^{l} \alpha_i \mathbb{P}\left(X_n \in A_i\right)\right| \\
&+ \left|\sum_{i=1}^{l} \alpha_i \mathbb{P}\left(X_n \in A_i\right) - \sum_{i=1}^{l} \alpha_i \mathbb{P}\left(X \in A_i\right)\right| \\
&+ \left|\sum_{i=1}^{l} \alpha_i \mathbb{P}\left(X \in A_i\right) - \int_\Omega f(X) d\mathbb{P}\right| \\
\leq& 3\epsilon.
\end{aligned}$$

donde segundo término se puede hacer arbitrariamente pequeño pues $\mathbb{P}\left(X_n \in A_i\right) \to \mathbb{P}\left(X \in A_i\right)$ al cumplirse por construcción que $\mu_X\left(\partial A_i\right) = 0$.

$4 \Rightarrow 5$. Si x es un punto de continuidad de F, entonces por la proposición 2.3.1, $\mu_X(\partial(-\infty, x]) = 0$ así que por 5, se sigue que $\mu_{X_n}(-\infty, x] \to \mu_{X_n}(-\infty, x]$ que es justamente $F_n(x) \to F(x)$.

$5 \Rightarrow 3$. Denotaremos por $H^i_c = \left\{x \in \mathbb{R}^d \,|\, x_i = c\right\}$ a los hiperplanos de $\mathbb{R}^d$ y sea $D^i = \left\{c \in \mathbb{R} \,|\, \mu\left(x \in H^i_c\right) > 0\right\}$. Los hiperplanos de medida positiva son el equivalente a los átomos de una distribución en una dimensión. Para cada $i \in \{1, \cdots, d\}$ fija, H^i_c y $H^i_{\tilde{c}}$ son ajenos si $c \neq \tilde{c}$ y por lo tanto, la cardinalidad de D^i es a lo más numerable pues la medida μ_X de la unión de estos está acotada por 1. Además, recordemos que por la proposición anterior $F_X(u)$ es continua en u si y solo si $\mu_X(\partial(-\infty, u]) = 0$.

Si $a = (a_1, \cdots, a_d)$ y $b = (b_1, \cdots, b_d)$, diremos que un rectángulo $(a, b]$ es *bueno* si a_i, $b_i \notin D_i$ para toda i. Entonces, si un rectángulo $(a, b]$ es *bueno*, F es continua en todo $\delta \in \{(c_1, c_2, \cdots, c_n) : c_i \in \{a_i, b_i\}\}$, los vértices de $(a, b]$ pues $\mu_X((-\infty, \delta]) = 0$.
Si $A = (a, b]$ es un rectángulo bueno, al utilizar la fórmula de inclusión-exclusión,

$$\lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \mathbb{P}\left(X_n \in A\right) = \lim_{n\to\infty} \sum_{k=0}^{d} (-1)^k \sum_{\delta \in \Delta_{k,d}} F_n(\delta)$$

$$= \sum_{k=0}^{d} (-1)^k \sum_{\delta \in \Delta_{k,d}} F(\delta) = \mathbb{P}\left(X \in A\right).$$

Por lo tanto, $\lim_{n\to\infty} \mu_{X_n}(A) = \mu_X(A)$ para todo rectángulo A o unión ajena de rectángulos *buenos*. Como D^i tiene cardinalidad numerable, cualquier abierto U de $\mathbb{R}^n$ se puede ver como unión creciente de uniones de rectángulos *buenos* A_j, es decir:

$$U = \lim_{j\to\infty} A_j \qquad A_j \text{ unión creciente de rectángulos buenos.}$$

Entonces, como $A_j \subset U$para toda j,

$$\liminf_n \mu_n(U) \geq \liminf_n \mu_n(A_j) = \lim_{n\to\infty} \mu_n(A_j) = \mu(A_j)$$

por lo que para toda j, $\liminf_n \mu_n(U) \geq \mu(A_j)$. Como $A_j \uparrow U$,

$$\liminf_n \mu_n(U) \geq \lim_{j\to\infty} \mu(A_j) = \mu(U).$$

Con esto, concluimos la prueba. ∎

Tanto el teorema de Helly como el teorema 2.1.4 de Prokhorov admiten una generalización en $\mathbb{R}^d$, ambas pruebas son similares al caso uni-dimensional y se pueden consultar en Ash [1], capítulo 7. Solo enunciaremos el segundo y omitiremos su prueba ya que se probará en un contexto más general cuando abordemos convergencia de procesos.

Teorema 2.3.3 Prokhorov en $\mathbb{R}^d$
Una sucesión de medidas $\{\mu_n\}$ en $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ es tensa si y solo si es relativamente compacta, es decir, si de toda subsucesión $\{\mu_{n_k}\}$ de $\{\mu_n\}$, se puede extraer una sub-sucesión $\{\mu_{n_{k_j}}\}$ tal que $\mu_{n_{k_j}} \xrightarrow{\mathscr{L}} \mu$ para alguna medida de probabilidad μ.

Otro teorema que se extiende a $\mathbb{R}^d$ es el teorema de continuidad de Lévy:

Teorema 2.3.4 Teorema de continuidad de Lévy en $\mathbb{R}^d$
Sean X_n, X variables aleatorias con valores en $\mathbb{R}^d$. Entonces

$$X_n \xrightarrow{\mathscr{L}} X \quad \textit{si y solo si} \quad \varphi_{X_n}(u) \to \varphi(u)$$

para cada $u \in \mathbb{R}^d$.

Demostración
Si $X_n \xrightarrow{\mathscr{L}} X$, entonces para toda función f continua y acotada, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$. Para todo $x \in \mathbb{R}^d$, la función $f : x \to e^{i<u,x>}$ es continua, de módulo 1 y se puede descomponer en parte real e imaginaria. Ambas son funciones reales continuas y acotadas pues están dominadas por el módulo de f. De esto, se sigue que:

$$\lim_{n\to\infty} \varphi_{X_n}(u) = \lim_{n\to\infty} \mathbb{E}\left[e^{i<u,X_n>}\right] = \mathbb{E}\left[e^{i<u,X>}\right] = \varphi_X(u).$$

Ahora, supongamos que para todo $u \in \mathbb{R}^d$, $\varphi_{X_n}(u) \to \varphi_X(u)$. La idea de la prueba es la siguiente: probar que la sucesión de medidas $\mu_{X_n} = \mathbb{P}_{X_n}$ en $\mathbb{R}^d$ es tensa y a partir de esto, probar que para todo x punto de continuidad de F, se cumple que $F_n(x) \to F(x)$.
Sea $u = (u_1, \cdots, u_d) \in \mathbb{R}$ y $s \in \mathbb{R}$. Definimos la sucesión de variables aleatorias reales $Y_n =< u, X_n >$ y $Y =< u, X >$. Entonces, para toda u:

$$\lim_{n\to\infty} \varphi_{Y_n}(s) = \lim_{n\to\infty} \mathbb{E}\left[e^{isY_n}\right] = \lim_{n\to\infty} \varphi_{X_n}(su) = \varphi_X(su) = \varphi_Y(s). \tag{2.3.1}$$

En particular, al considerar $u = e_i$ (el i-ésimo vector canónico), se tiene que $Y_n = X_n^i$ y por (2.3.1),

$$\lim_{n\to\infty} \varphi_{X_n^i}(s) = \varphi_{X^i}(s).$$

Por el teorema de continuidad de Lévy en $\mathbb{R}$, $X_n^i \xrightarrow{\mathscr{L}} X^i$ para toda i y la sucesión $\{\mu_{X_n^i}\}_{n\in\mathbb{N}}$ es tensa, lo cual permite probar sin demasiadas dificultades que la sucesión $\{\mu_{X_n}\}_{n\in\mathbb{N}}$ también lo es. Para facilitar la escritura denotaremos por μ_n a μ_{X_n} y por μ a μ_X.

Por el teorema de Prokhorov 2.3.3, como la sucesión $\{\mu_n\}$ es tensa, existe una subsucesión $\{\mu_{n_k}\}$ de esta, tal que $\mu_{n_k} \xrightarrow{\mathscr{L}} \nu$ para alguna medida de probabilidad ν. Ahora, sean F_n y F las distribuciones de X_n y X respectivamente. Bastaría probar que $F_n(x) \to F(x)$ para todo x punto de continuidad de F.
Supongamos existe x_0 punto de continuidad de F tal que

$$\lim_{n\to\infty} F_n(x_0) \neq F(x_0).$$

Como la sucesión $\{F_n(x_0)\}$ toma valores en $[0,1]$, entonces es una sucesión acotada y por lo tanto existe una subsucesión $\{F_{n_l}(x_0)\}$ convergente, $F_{n_l}(x_0) \to \beta \in [0,1]$, con $\beta \neq F(x_0)$. La subsucesión $\{\mu_{n_l}\}$ es tensa pues es subsucesión de $\{\mu_n\}$ y una vez más, por el teorema 2.3.3 existe una subsubsucesión $\{\mu_{n_{l_k}}\}$ que converge a una medida de probabilidad $\tilde{\nu}$. Pero entonces, como $\mu_{n_{l_k}} \xrightarrow{\mathscr{L}} \tilde{\nu}$, por la primera implicación de este teorema, las características convergen puntualmente:

$$\lim_{k\to\infty} \varphi_{n_{l_k}}(s) = \psi(s) \qquad \psi \text{ función característica de } \tilde{\nu}.$$

Por hipótesis, $\varphi_n \to \varphi_X$ por lo que $\psi = \varphi_X$ y por ende, las leyes asociadas deben ser las mismas: $\tilde{\nu} = \mu_X$. Por lo tanto, $\mu_{n_{l_k}} \xrightarrow{\mathscr{L}} \mu_X$ y entonces para todo x punto de continuidad de F, $F_{n_{l_k}}(x) \to F(x)$. En particular, se debe cumplir que $F_{n_{l_k}}(x_0) \to F(x_0)$, lo cual es una contradicción, pues $\{F_{n_{l_k}}\}$ es una subsucesión de $\{F_{n_l}\}$ y $F_{n_l}(x_0) \to \beta$ donde $\beta \neq F(x_0)$. ■

Recordemos que si X_k, X son variables en $\mathbb{R}^d$, $X_k \to X$ en ley si y solo si, para toda $f : \mathbb{R}^d \to \mathbb{R}$ continua y acotada, $\lim_{k\to\infty} \mathbb{E}[f(X_k)] = \mathbb{E}[f(X)]$

Proposición 2.3.5 *Sean* $X_k = (X_k^1, \cdots, X_k^d)$ *y* $Y = (Y^1, \cdots, Y^d)$ *variables aleatorias con valores en* $(\mathbb{R}, \mathscr{B}(\mathbb{R}^d)$. *Entonces,* $X_k \to Y$ *en Ley si y solo si*

$$\sum_{u=1}^{d} t_n X_n^u \to \sum_{u=1}^{d} t_u Y^u \quad \textit{en Ley}$$

para todo vector $(t_1, \cdots, t_d) \in \mathbb{R}^d$. *Nótese que* $\sum_{u=1}^{d} t_n X_n^u$ *es una combinación lineal del vector* X_k*, y por lo tanto es una variable aleatoria real.*

Demostración
La función $g(x_1, \cdots, x_n) := \sum_{u=1}^{d} t_u x_u$ es una función continua de $\mathbb{R}^d$ en $\mathbb{R}$. Entonces, si $f : \mathbb{R} \to \mathbb{R}$ es una función continua, la composición $g \circ f : \mathbb{R}^d \to \mathbb{R}$ es continua y acotada. Como $X_k \xrightarrow{\mathscr{L}} Y$,

$$\lim_{n\to\infty} \mathbb{E}\left[f \circ g(X_n)\right] = \mathbb{E}\left[f \circ g(Y)\right]$$

por lo que

$$\lim_{n\to\infty} \mathbb{E}\left[f\left(\sum_{u=1}^{d} t_n X_n^u\right)\right] = \mathbb{E}\left[f\left(\sum_{u=1}^{d} t_u Y^u\right)\right]$$

para toda f continua y acotada, y entonces $\sum_{u=1}^{d} t_n X_n^u \xrightarrow{\mathscr{L}} \sum_{u=1}^{d} t_u Y^u$.
Ahora, supongamos para cualquier vector $t = (t_1, \cdots, t_n)$ las combinaciones lineales convergen en ley. Denotamos por φ_k^t, φ_Y^t a las funciones característica de $\sum_{u=1}^{d} t_u X_k^u$ y $\sum_{u=1}^{d} t_u Y^u$ respectivamente. Entonces, por el teorema de continuidad de Lévy, para toda $s \in \mathbb{R}$,

$$\lim_{k\to\infty} \varphi_k^t(s) = \lim_{k\to\infty} \mathbb{E}\left[e^{is\sum_{u=1}^{d} t_u X_k^u}\right] = \mathbb{E}\left[e^{is\sum_{u=1}^{d} t_u Y^u}\right] = \varphi_Y(s).$$

En particular, si $s = 1$, se satisface para todo t que:

$$\lim_{k\to\infty} \mathbb{E}\left[e^{i<t,X_k>}\right] = \mathbb{E}\left[e^{i<t,Y>}\right]$$

que son justamente las características de X_k y Y. Una vez más, por el teorema de continuidad de Lévy en $\mathbb{R}^d$, $X_k \xrightarrow{\mathscr{L}} Y$. ■

Una consecuencia del resultado anterior es una generalización del teorema del límite central para $\mathbb{R}^d$:

Teorema 2.3.6 Teorema del límite central multivariado
Sean $X_n = (X_n^1, \cdots, X_n^d)$ *con* $n \in \mathbb{N}$ *una colección de vectores aleatorios independientes idénticamente distribuidos con vector de medias c y*

$$Var(X_n^i) = \sigma_i^2 < \infty \qquad \mathbb{E}\left[X_n^i\right] = c^i \qquad d_n^{i,j} = \mathbb{E}\left[(X_n^i - c^i)(X_n^j - c^j)\right]$$

y $\Sigma = \{d_n^{i,j}\}$ *la matriz de covarianzas.*
Si

$$S_n = \sum_{k=1}^{n} X_k = \left(\sum_{k=1}^{n} X_k^1, \cdots, \sum_{k=1}^{n} X_k^d\right)$$

entonces:

$$\lim_{n\to\infty} \frac{S_n - nc}{\sqrt{n}} = N \qquad \textit{en Ley}$$

donde $N = (Y_1, \cdots, Y_d)$ *es una variable aleatoria normal(0,Σ).*

Obsérvese que si X_n y X_m son idénticamente distribuidas, para todo $A = A_1 \times \cdots \times A_d \in \mathscr{B}(\mathbb{R}^d)$:

$$\begin{aligned}\mathbb{P}\left(X_n \in A\right) &= \mathbb{P}\left(X_n^1 \in A_1, \cdots, X_n^d \in A_d\right) \\ &= \mathbb{P}\left(X_m^1 \in A_1, \cdots, X_m^d \in A_d\right) = \mathbb{P}\left(X_m \in A\right)\end{aligned}$$

por lo que sus entradas X_n^i y X_m^i también lo son.

Demostración
Al reescribir, tenemos que

$$\frac{S_n - nc}{\sqrt{n}} \xrightarrow{\mathscr{L}} N(0, \Sigma) \qquad \text{si y solo si}$$

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_i^1 - c_1\right), \cdots, \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_i^d - c_d\right)\right) \xrightarrow{\mathscr{L}} N(0, \Sigma)$$

y por la proposición (2.3.5), bastaría probar que para toda $t = (t_1, \cdots, t_n)$ se cumple que

$$\lim_{n\to\infty}\sum_{k=1}^{d}\left(\frac{t_k}{\sqrt{n}}\sum_{i=1}^{n}\left(X_i^k - c_k\right)\right) = \sum_{k=1}^{d} t_k Y_k \quad \text{en Ley.}$$

Como N se distribuye normal multivariada $(0, \Sigma)$, entonces $\sum_{k=1}^{d} t_k Y_k$ se distribuye normal con media 0 y varianza $t^t\,\Sigma\,t$.
Pero:

$$\lim_{n\to\infty}\sum_{k=1}^{d}\frac{t_k}{\sqrt{n}}\left(\sum_{i=1}^{n}\left(X_i^k - c_i\right)\right) = \lim_{n\to\infty}\frac{\sum_{i=1}^{n}\left(\sum_{k=1}^{d} t_k\left(X_i^k - c_k\right)\right)}{\sqrt{n}}$$

donde las variables aleatorias $\{\sum_{k=1}^{d} t_k\left(X_i^k - c_k\right)\}_{i\in\mathbb{N}}$ son independientes, idénticamente distribuidas, de media 0 y no es difícil comprobar que su varianza es $t^t\Sigma t$. Por el Teorema del límite central real, convergen a una variable aleatoria normal con media 0 y varianza $t^t\Sigma t$. ∎

Capítulo 3

Medidas aleatorias de Poisson

Referencias: Sato [15], capítulo 4 y algunos elementos de Kingman [16] capítulos 1-3.

Este capítulo podría parecer un tanto desconectado de los anteriores así que explicaremos brevemente su papel. En el capítulo 4 nos daremos a la tarea de probar un teorema de existencia de procesos de Lévy, conocido como la descomposición de Lévy-Itô. En la prueba de este resultado convergen tres grandes temas: los resultados del capítulo 2 de variables aleatorias infinitamente divisibles, teoría de las martingalas cuadrado integrables y las medidas aleatorias de Poisson. El primer tema ya se cubrió con detalle y la teoría de martingalas es conocida, así que la omitimos de esta presentación. Se pueden consultar en el apéndice las pruebas de los principales resultados. Así, el último paso será estudiar a las medidas aleatorias de Poisson y a esto dedicamos este capítulo. La prueba que incluimos del teorema de Campbell consiste en una generalización para medidas σ-finitas de la prueba de Sato [15], capítulo 4. Se pueden leer otras pruebas en Kyprianou [13] capítulo 2 así como Kingman [16] capítulo 3 aunque a mi parecer, la que proponemos permite comprender mejor las propiedades básicas de las medias aleatorias de Poisson.

3.1 Primeras definiciones y existencia

Sea $(S, \mathscr{S}, \nu)$ un espacio de medida con ν medida σ-finita y $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad.

Definición 3.1.1 *Una medida aleatoria de Poisson con medida de intensidad ν es una colección de variables aleatorias*

$$N = \{N(A) \,:\, A \in \mathscr{S}\}$$

tales que:

i) Para cada ω, $N(\,\cdot\,, \omega)$ es una medida de conteo (valores en $\mathbb{N}$) en $(S, \mathscr{S})$.
ii) Para cada $A \in \mathscr{S}$, $N(A, \cdot\,)$ es una variable aleatoria con distribución Poisson de parámetro $\nu(A)$:

$$\mathbb{P}\left(N(A) = k\right) = e^{-\nu(A)} \frac{(\nu(A))^k}{k!} \tag{3.1.1}$$

iii) Si $A_1, \cdots, A_n \in \mathscr{S}$ son ajenos, entonces las variables aleatorias $N(A_1), \cdots, N(A_n)$ son independientes.

Por convención, una variable aleatoria Poisson de parámetro infinito es una variable aleatoria que toma el valor infinito casi seguramente y una Poisson de parámetro 0 es una variable aleatoria que toma el valor 0 casi seguramente. Empezamos probando la existencia de una colección con estas características. De la construcción debería esclarecerse la naturaleza de N.

Teorema 3.1.2 Existencia de medidas aleatorias de Poisson
Dado un espacio $(S, \mathscr{S}, \nu)$ con ν medida sigma finita, existe un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ y una medida aleatoria de Poisson $N : S \times \Omega \to \mathbb{R}$ con intensidad ν.

Demostración
Caso $\nu(S) < \infty$:
Sean $X_1, \cdots, X_n, \cdots$ variables i.i.d., $X_i : \Omega \to S$ con distribución $\mathbb{P}(X_i \in A) = \frac{\nu(A)}{\nu(S)}$ para $A \in \mathscr{S}$ (se puede pensar como la distribución uniforme en S) y sea Y variable aleatoria Poisson de parámetro $c = \nu(S)$ en el mismo espacio Ω independiente de las anteriores. Para cada $A \in \mathscr{S}$, definimos a la siguiente variable aleatoria:

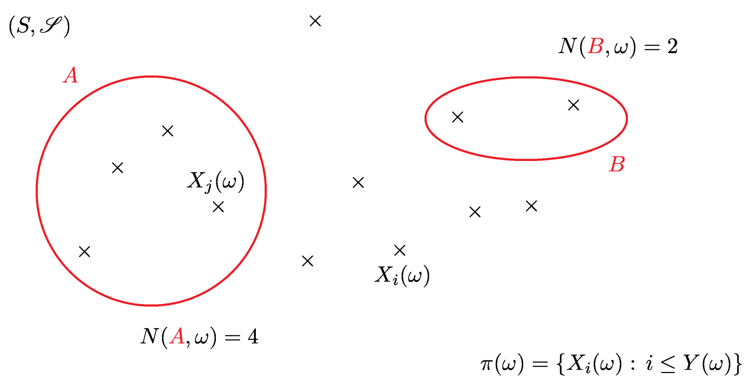
$$N(A, \omega) = \sum_{i=1}^{Y(\omega)} \mathbb{1}_{\{X_i(\omega) \in A\}}. \tag{3.1.2}$$

Nos detenemos brevemente para explicar este nuevo objeto: para cada ω fija, $Y(\omega)$ toma un valor en los naturales. Restringimos nuestra colección de variables aleatorias $\{X_i(\omega)\}_{i \in \mathbb{N}}$ a las primeras $Y(\omega)$ y obtenemos el siguiente subconjunto de elementos de S:

$$\pi(\omega) = \{X_i(\omega) \,:\, i \in \{1 \cdots Y(\omega)\}\}.$$

Entonces $\pi(\omega)$ es una nube de puntos en S de cardinalidad $Y(\omega)$, y la medida $N(A, \omega)$ de un conjunto A se define como el número de variables aleatorias $X_i(\omega)$ con $X_i(\omega) \in \pi(\omega)$ que cayeron dentro de A:

$$N(A, \omega) = \#\pi(\omega) \cap A.$$



Al conjunto aleatorio π se le conoce como el *proceso Poisson puntual* asociado a N y está claro que, así definida, $N(\cdot, \omega)$ es una medida atómica con soporte $\pi(\omega)$. Otra forma análoga de escribirlo es

$$N(\cdot, \omega) = \sum_{X_i(\omega) \in \pi(\omega)} \delta_{X_i(\omega)}.$$

Regresamos a la prueba: para cada A fijo, la variable aleatoria $\mathbb{1}_{\{X_1(\omega)\in A\}}$ se distribuye Bernoulli de parámetro $p = \frac{\nu(A)}{\nu(S)}$ y entonces, por el ejemplo 4.1, $N(A,\omega)$ se distribuye Poisson de parámetro $cp = \nu(A)$. Falta probar que si $A_1, \cdots, A_k$ son ajenos, entonces las variables aleatorias $N(A_1), \cdots, N(A_k)$ son independientes. Consideremos entonces $A_1, \cdots, A_k$ ajenos tales que $\cup_{i=1}^k A_i = S$, $n_1, \cdots, n_k$ naturales y sea $n := n_1 + \cdots + n_k$.

$$\begin{aligned}
\mathbb{P}\left(N(A_1) = n_1, \cdots, N(A_k) = n_k\right) &= \mathbb{P}\left(N(A_1) = n_1, \cdots, N(A_k) = n_k, Y = n\right) \\
&= \mathbb{P}\left(N(A_1) = n_1, \cdots, N(A_k) = n_k \,|\, Y = n\right)\mathbb{P}\left(Y = n\right) \\
&= \mathbb{P}\left(\sum_{j=1}^n \mathbb{1}_{A_1}(X_j) = n_1, \cdots, \sum_{j=1}^n \mathbb{1}_{A_k}(X_j) = n_k\right)\mathbb{P}\left(Y = n\right)
\end{aligned}$$

El primer término corresponde a una distribución multinomial mientras que el segundo a una Poisson:

$$\begin{aligned}
&= \frac{n!}{n_1!\cdots n_k!}\left(\frac{\nu(A_1)}{\nu(S)}\right)^{n_1}\cdot\cdots\cdot\left(\frac{\nu(A_k)}{\nu(S)}\right)^{n_k} e^{-\nu(S)}\frac{\nu(S)^n}{n!} \\
&= \frac{1}{n_1!\cdots n_k!}\nu(A_1)^{n_1}\cdot\cdots\cdot\nu(A_2)^{n_2}e^{-(\nu(A_1)+\cdots+\nu(A_k))} \\
&= \prod_{i=1}^k \frac{\nu(A_i)^{n_i}}{n_i!}e^{-\nu(A_i)} \\
&= \prod_{i=1}^k \mathbb{P}\left(N(A_i) = n_i\right)
\end{aligned}$$

con lo que concluimos el caso $\nu(S) < \infty$.

Caso $\nu(S) = \infty$:
Como el espacio S es sigma finito, existen conjuntos $\Theta_1, \ldots, \Theta_k, \ldots$ ajenos de ν-medida finita tales que $\cup_{k=1}^\infty \Theta_k = S$. Para cada k, definimos la medida finita $\nu_k(B) := \nu(B \cap \Theta_k)$; por el caso anterior, existe un espacio de probabilidad y una medida aleatoria de Poisson $N_k = \{N_k(B) : B \in \mathscr{S}\}$ con medida de intensidad ν_k. Al pasar a un espacio producto adecuando, las podemos construir independientes en un espacio común Ω. Definimos la medida aleatoria:

$$N(B) = \sum_{k=1}^\infty N_k(B) \qquad B \in \mathscr{S}$$

Veamos que en efecto es una variable aleatoria Poisson de parámetro $\nu(B)$. Si $\nu(B) < \infty$, así definido, $N(B)$ es la suma de variables Poisson independientes de parámetro $\nu_k(B)$ por lo que también se distribuye Poisson. Además,

$$\mathbb{E}\left[N(B)\right] = \sum_{k=1}^\infty \mathbb{E}\left[N_k(B)\right] = \sum_{k=1}^\infty \nu_k(B) = \nu(B)$$

así que su parámetro es $\nu(B)$. Falta probar el caso $\nu(B) = \infty$; necesitamos probar que $N(B)$ se distribuye Poisson de parámetro $\nu(B) = \infty$ que, por convención, es que la variable aleatoria

$N(B)$ tome el valor infinito casi seguramente. Para esto, obsérvese que existe un real positivo a' tal que $1 - e^{-x} \geq \frac{x}{2}$ si $0 \leq x \leq a'$ (esto es claro al hacer un dibujo) mientras que si $a' \leq x$, $1 - e^{-x} \geq 1 - e^{-a'}$ ya que $1 - e^{-x}$ es creciente. Si definimos $a = 1 - e^{-a'}$, se tiene que $1 - e^{-x} \geq \frac{x}{2} \wedge a$ y entonces:

$$\sum_{k=1}^{\infty} \mathbb{P}\left(N_k(B) \geq 1\right) = \sum_{k=1}^{\infty} \left(1 - e^{-\nu_k(B)}\right) \geq \sum_{k=1}^{\infty} \left(\frac{1}{2}\nu_k(B) \wedge a\right) = \infty$$

donde la última igualdad es consecuencia de que $\sum_k \nu_k(B) = \nu(B) = \infty$. Por Borel-Cantelli, $\mathbb{P}\left(\limsup\{N_k(B) \geq 1\}\right) = 1$, es decir casi seguramente, para cada ω, el evento $N_k(B) \geq 1$ ocurre una infinidad de veces por lo que $N(B) = \infty$ casi seguramente. Para probar la propiedad de independencia, consideramos conjuntos ajenos $A_1, \cdots, A_r$. Al descomponer cada evento

$$\{N(A_i) = n_i\} = \bigcup_{\sum_{k=1}^{\infty} n_i^k = n_i} \{N_k(A_i) = n_i^k\}$$

usar independencia y agrupar adecuadamente se sigue que

$$\mathbb{P}\left(N(A_1) = n_1, \cdots, N(A_r) = n_r\right) = \prod_{i=1}^{r} \mathbb{P}\left(N(A_i) = n_i\right).$$

■

Obsérvese que si $\nu(S) < \infty$, el soporte de la medida N es finito casi seguramente pues $\mathbb{E}\left[N(S)\right] = \nu(S)$ mientras que si $\nu(S) = \infty$, la variable aleatoria $N(S)$ es Poisson de parámetro infinito por lo que toma el valor infinito casi seguramente y entonces, el soporte de N es numerable no finito casi seguramente: existe una cantidad infinita numerable de puntos con masa positiva. De esto último se sigue que en el caso $\nu(S) < \infty$, N es una medida finita c.s. y en particular es σ-finita. Veamos que cuando la medida soporte cumple $\nu(S) = \infty$, la medida aleatoria N sigue siendo σ-finita: como ν es σ-finita, existen conjuntos A_i con $S = \nu(A_i) < \infty$ y $\cup_i A_i = S$. Entonces, $\mathbb{E}\left[N(S)\right] = \mathbb{E}\left[N(\cup_i A_i)\right]$ donde $\mathbb{E}\left[N(A_i)\right] = \nu(A_i) < \infty$. Así, tenemos que $N(S) = N(\cup_i A_i)$ con $N(A_i) < \infty$ c.s. Esto es, N es casi seguramente una medida σ-finita.

Otra cosa a notar es que si ν tiene algún átomo, es decir, $\nu(\{x\}) > 0$ para algún $x \in S$, entonces $N(\{x\})$ es Poisson con parámetro $\nu(\{x\}) > 0$ por lo que $\mathbb{P}\left(N(\{x\}) > 0\right) > 0$ (es decir, si x es átomo de ν, entonces tiene probabilidad positiva de tener medida de Poisson no nula). Si ν no tiene átomos, $\nu(\{x\}) = 0$ para todo $x \in S$, $N(\{x\}) = 0$ casi seguramente para todo $x \in S$ por ser Poisson de parámetro 0. En el contexto de los procesos de Lévy como se verá más adelante, haremos uso solamente de medidas ν no atómicas.

3.2 Integración con respecto a medidas aleatorias y fórmula de Campbell

Como para cada ω, $N(\cdot, \omega)$ es una medida, es natural buscar definir la integral de una función $g : S \to \mathbb{R}$ con respecto a $N(dx, \omega)$. Como $N(\cdot, \omega)$ es atómica, si denotamos por π al soporte de N y la medida de intensidad ν no tiene átomos, las integrales deberían reducirse a sumas:

$$\int_S g(x) N(dx, \omega) = \sum_{x \in \pi(\omega)} g(x) \tag{3.2.1}$$

Nótese que al ser N σ-finita c.s., podemos aplicar el teorema de convergencia monótona c.s. a integrales con respecto a $N(dx,\omega)$ y una vez hayamos establecido criterios de integrabilidad podremos aplicar el teorema de convergencia dominada. El problema es que aún no es claro cuando expresiones de la forma (3.2.1) convergen, así que establecemos a continuación condiciones bajo las cuales integrar está bien definido:

Teorema 3.2.1 Fórmula de Campbell
Supongamos que $(S,\mathscr{S},\nu)$ es espacio de medida con ν σ-finita y N medida aleatoria de Poisson con intensidad ν. Sea $g:S\to\mathbb{R}$ medible, definimos:

$$Y(\omega)=\int_S g(x)N(dx,\omega).$$

Se satisface lo siguiente:

i)

$$\int_S |g(x)|N(dx)<\infty \quad c.s. \qquad \text{si y solo si} \qquad \int_S (1\wedge|g(x)|)\,\nu(dx)<\infty. \tag{3.2.2}$$

Esto nos da una condición necesaria y suficiente para que g sea integrable c.s. con respecto a $N(dx,\omega)$.

ii) Si $\int_S (1\wedge|g(x)|)\,\nu(dx)<\infty$, entonces

$$\mathbb{E}\left[e^{i\lambda Y}\right]=\exp\left\{-\int_S\left(1-e^{i\lambda g(x)}\right)\nu(dx)\right\} \tag{3.2.3}$$

para todo $\lambda\in\mathbb{R}$ por lo que si ν es finita, Y tiene distribución Poisson compuesto.

iii) Si $\int_S|g(x)|\nu(dx)<\infty$,

$$\mathbb{E}\left[Y\right]=\int_S g(x)\nu(dx) \tag{3.2.4}$$

y si además, $\int_S|g(x)|^2\nu(dx)<\infty$,

$$\mathbb{E}\left[Y^2\right]=\int_S g^2(x)\nu(dx)+\left(\int_S g(x)\nu(dx)\right)^2 \tag{3.2.5}$$

Demostración

En $ii)$ probaremos que, así definida, Y es en efecto una variable aleatoria así que lo asumiremos para la prueba de $i)$. Supondremos que $\nu(S)=\infty$ (la prueba del caso finito es un caso particular de esta) y por lo tanto, la medida aleatoria N tiene soporte en S infinito casi seguramente pues $\nu(S)=\infty$. Iniciamos entonces con la prueba de $i)$:

Supongamos que $\int_S (1\wedge|g(x)|)\,\nu(dx)<\infty$. Como,

$$\int_S 1\wedge|g|d\nu=\nu(x\,;|g(x)|\geq 1)+\int_{|g|<1}|g|\,d\nu \tag{3.2.6}$$

se sigue que

$$\nu\left(x\,;\,|g(x)|\geq 1\right)<\infty \quad \text{y} \quad \int_{|g|<1}|g(x)|\nu(dx)<\infty. \tag{3.2.7}$$

Como $\int_S|g(x)|N(dx)=\int_{|g|\geq 1}|g(x)|N(dx)+\int_{|g|<1}|g(x)|N(dx)$, basta probar que ambas integrales convergen casi seguramente.

Por (3.2.7), N tiene soporte finito casi seguramente en $\left\{|g(x)|\geq 1\right\}$ pues la medida Poisson de este se distribuye Poisson con parámetro $\nu\,(x\,;\,|g(x)|\geq 1)<\infty$. Por lo tanto, la primera integral es una suma finita de la forma $\sum_{x_i\in B} g(x_i)$ donde $B(\omega)=\pi(\omega)\cap\big\{x\,;\,|g(x)|\geq 1\big\}$.

Para probar que $\int_{|g|<1}|g(x)|N(dx)$ es finito c.s., bastaría ver que:

$$\mathbb{E}\left[\int_{|g|<1}|g(x)|N(dx)\right]<\infty \tag{3.2.8}$$

Es sabido que si f es una función medible positiva y acotada en un espacio sigma finito, entonces

$$\int f d\nu<\infty \quad \text{si y solo si} \quad \sum_{n=0}^{\infty}\frac{1}{2^n}\nu\left(x\,:\,f(x)\geq\frac{1}{2^n}\right)<\infty \tag{3.2.9}$$

La prueba se deja como ejercicio. Tomando $f=|g|\mathbb{1}_{|g|<1}$ y aplicando (3.2.7) tenemos que $\int f d\nu<\infty$ y entonces la segunda condición de (3.2.9) se satisface también. Si $A_n=\Big\{x\,;\,|g(x)|\in\big[2^{-(n+1)},2^{-n}\big)\Big\}$, descomponemos como unión ajena

$$\{x\,;\,|g(x)|<1\}=\{x\,;\,g(x)=0\}\cup\bigcup_{n=0}^{\infty}A_n$$

por lo que:

$$\begin{aligned}
\mathbb{E}\left[\int_{|g|<1}|g(x)|N(dx)\right]&=\mathbb{E}\left[\sum_{n=0}^{\infty}\int_{A_n}|g(x)|N(dx)\right]\\
&=\sum_{n=0}^{\infty}\mathbb{E}\left[\int_{A_n}|g(x)|N(dx)\right]\\
&\leq\sum_{n=0}^{\infty}\mathbb{E}\left[N\left(A_n\right)2^{-n}\right]\\
&=\sum_{n=0}^{\infty}\nu(A_n)2^{-n}\\
&\leq\sum_{n=0}^{\infty}2^{-n}\nu\left(x\,;f(x)\geq\ 2^{-(n+1)}\right)\\
&<\infty
\end{aligned}$$

Concluimos que $\int_S|g(x)|N(dx)<\infty$ casi seguramente.

Probemos el recíproco por contradicción; supongamos que $\int_S|g(x)|N(dx)<\infty$ pero que $\int_S 1\wedge|g|\nu(dx)=\infty$. Por (3.2.6),

$$\nu(x\,;\,|g|\geq 1)=\infty \quad \text{o} \quad \int_{|g|<1}|g|\,d\nu=\infty.$$

Veamos que ambas posibilidades nos llevan a una contradicción.

Caso 1: si $\nu(x\,;\,|g|\geq 1)=\infty$, entonces $N(x\,;\,|g|\geq 1)$ es infinito casi seguramente y entonces:

$$\int_S|g(x)|N(dx)\geq\int_{|g|\geq 1}|g(x)|N(dx)\geq\int_{|g|\geq 1}1\,N(dx)=N(x\,;\,|g(x)|\geq 1)=\infty.$$

Caso 2: si $\int_{|g|<1} |g| N(dx) = \infty$, por un lado:

$$\mathbb{E}\left[\int_{|g|<1} |g| N(dx)\right] = \sum_{n=0}^{\infty} \mathbb{E}\left[\int_{A_n} |g| N(dx)\right] \geq \sum_{n=0}^{\infty} 2^{-(n+1)} \nu(A_n). \tag{3.2.10}$$

Por otro lado si definimos $B_n = \{x\,;\, |g(x)| \geq 2^{-n}\}$, como $\int_{|g|<1} |g| N(dx) = \infty$, por (3.2.9) tenemos que

$$\sum_{n=0}^{\infty} 2^{-n} \nu\left(x\,;\, |g(x)| \geq \frac{1}{2^n}\right) = \sum_{n=0}^{\infty} 2^{-n} \nu\left(B_n\right) = \infty.$$

Además, al descomponer $B_n = \cup_{i=0}^{n-1} A_i$ podemos escribir $\nu(B_n) = \sum_{i=1}^{n-1} \nu(A_i)$ al ser ajenos. Entonces:

$$\begin{aligned}
\sum_{n=1}^{\infty} 2^{-n} \nu(B_n) &= \sum_{n=1}^{\infty} 2^{-n} \left(\sum_{i=0}^{n-1} \nu(A_i)\right) \\
&= 2^{-1}\left(\nu(A_0)\right) + 2^{-2}\left(\nu(A_0) + \nu(A_1)\right) + \dots \\
&\qquad \dots + 2^{-n}\left(\nu(A_0) + \dots + \nu(A_{n-1})\right) + \dots \\
&= \nu(A_0) \sum_{i=1}^{\infty} 2^{-i} + \nu(A_1) \sum_{i=2}^{\infty} 2^{-i} + \dots + \nu(A_n) \sum_{i=n+1}^{\infty} 2^{-i} + \dots \\
&= \nu(A_0) \cdot 2^{-1} + \nu(A_1) \cdot 2^{-2} + \dots + \nu(A_n) 2^{-(n+1)} + \dots \\
&= \sum_{n=0}^{\infty} 2^{-(n+1)} \nu(A_n).
\end{aligned}$$

Así que por (3.2.10) y esta última igualdad,

$$\mathbb{E}\left[\int_{|g|<1} |g| N(dx)\right] \geq \sum_{n=1}^{\infty} 2^{-n} \nu(B_n) = \infty$$

lo cual es también una contradicción. Con esto, concluimos la prueba de i).

Prueba de ii) :
Seguiremos usando la notación definida en ii). Sea $g_n = \sum_{k=1}^{n} x_k^n \mathbb{1}_{C_k^n}$ sucesión de funciones simples tales que $|g_n| \uparrow |g|$. Obsérvese que entonces $N(C_k^n) < \infty$ por lo que $N(C_k^n)$ es una variable Poisson en el sentido usual. Si

$$Y_n = \int_S g_n(x) N(dx) = \sum_{k=1}^{n} x_k^n N(C_k^n)$$

entonces Y_n es variable aleatoria para cada n por ser combinación lineal de variables aleatorias Poisson $N(C_k^n)$ y además

$$\lim_{n\to\infty} |Y_n(\omega) - Y(\omega)| \leq \lim_{n\to\infty} \int_S |g_n - g| N(dx, \omega) = 0 \tag{3.2.11}$$

por el teorema deconvergencia dominada, pues N es medida sigma finita casi seguramente (dominamos por $2|g|$ y recordemos que $\int_S |g| N(dx) < \infty$ por hipótesis). Concluimos que $Y_n \to Y$ puntualmente por lo que Y es variable aleatoria. En particular, $Y_n \xrightarrow{\mathscr{L}} Y$ y por el teorema deLévy,

$\varphi_{Y_n}(\lambda) \to \varphi_Y(\lambda)$. Al ser suma de variables Poisson independientes, la característica de Y_n se calcula fácilmente:

$$\begin{aligned}
\varphi_{Y_n} = \mathbb{E}\left[e^{i\lambda Y_n}\right] &= \prod_{k=1}^{n} \mathbb{E}\left[e^{i\lambda x_k^n N(C_k^n)}\right] \\
&= \prod_{k=1}^{n} \exp\left\{\nu(C_k^n)\left(e^{i\lambda x_k^n} - 1\right)\right\} \\
&= \exp\left\{\sum_{k=1}^{n} \nu(C_k^n)\left(e^{i\lambda x_k^n} - 1\right)\right\} \\
&= \exp\left\{\int_S \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx)\right\}
\end{aligned}$$

y puesto que $g_n \to g$, bastaría argumentar por que podemos usar el teorema de convergencia dominada para concluir (3.2.3).
Como

$$\int_S \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx) = \int_{|g|<1} \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx) + \int_{|g|\geq 1} \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx)$$

entonces dominamos cada término por separado:
Para toda n, $\int_{|g|\geq 1} \left|e^{i\lambda g_n(x)} - 1\right| \nu(dx) \leq 2\nu(x\,;\,|g| \geq 1) < \infty$ por (3.2.7). Para dominar el otro término, usamos el lema 2.1.8:

$$\begin{aligned}
\int_{|g|<1} \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx) &\leq \int_{|g|<1} \min\left\{\frac{|\lambda g_n|^2}{2}, 2|\lambda g_n|\right\}\nu(dx) \\
&\leq \int_{|g|<1} \min\left\{\frac{|\lambda g|^2}{2}, 2|\lambda g|\right\}\nu(dx) \\
&\leq 2|\lambda| \int_{|g|<1} |g|\nu(dx) \\
&< \infty \qquad\qquad \text{por (3.2.7).}
\end{aligned}$$

Así, para cada $\lambda \in \mathbb{R}$, $\int_S \left(e^{i\lambda g_n(x)} - 1\right)\nu(dx) \longrightarrow \int_S \left(e^{i\lambda g(x)} - 1\right)\nu(dx)$ y por continuidad se tiene (3.2.3).

Prueba de *iii)*
Sean $\tilde{Y_n} = \int_S |g_n| N(dx)$ y $\tilde{Y} = \int_S |g| N(dx)$. Son finitas casi seguramente pues $0 \leq \tilde{Y_n} \leq \tilde{Y} < \infty$ por hipótesis. Además, $|\tilde{Y_n} - \tilde{Y}| \leq \int_S |g_n - g|\nu(dx) \to 0$ por (3.2.11) por lo que $\tilde{Y_n} \uparrow \tilde{Y}$ puntualmente (recordemos que $|g_n| \uparrow |g|$). Veamos que $\tilde{Y}$ es integrable:

$$\mathbb{E}\left[\tilde{Y_n}\right] = \mathbb{E}\left[\sum_{i=1}^{n} |x_k^n| N(C_k^n)\right] = \sum_{i=1}^{n} |x_k^n|\nu(C_k^n) = \int_S |g_n|\nu(dx) \leq \int_S |g|\nu(dx) < \infty$$

y esto pasa para toda n. Ahora, como $\tilde{Y_n} \uparrow \tilde{Y}$, por el teorema de convergencia monótona $\mathbb{E}\left[\tilde{Y_n}\right] \to \mathbb{E}\left[\tilde{Y}\right] \leq \int_S |g|\nu(dx) < \infty$, por lo que $\tilde{Y}$ es integrable.

Como para cada n se satisface que $\mathbb{E}[Y_n] = \sum_{k=1}^{n} x_k^n \nu(C_k^n) = \int_S g_n \nu(dx)$, entonces aplicando dos veces el teorema de convergencia dominada:

$$\mathbb{E}[Y] = \lim_{n\to\infty} \mathbb{E}[Y_n] = \lim_{n\to\infty} \int_S g_n\,\nu(dx) = \int_S g\,\nu(dx) \tag{3.2.12}$$

donde para la primera aplicación, usamos que $|Y_n|$ está dominada por $\tilde{Y}$ que es integrable y para la segunda que $|g_n|$ está dominada por $|g|$. Con esto, concluimos (3.2.4). La prueba de (3.2.5) es similar; solo es necesario recordar que si X se distribuye Poisson de parámetro c, entonces $\mathbb{E}\left[X^2\right] = c + c^2$. Primero, lo probamos para Y_n:

$$\begin{aligned}
\mathbb{E}\left[Y_n^2\right] = \mathbb{E}\left[\left(\sum_{k=1}^{n} x_k^n N(C_k^n)\right)^2\right] &= \sum_{k \neq q} x_k^n x_q^n \mathbb{E}\left[N(A_k^n)\right] \mathbb{E}\left[N(A_q^n)\right] + \sum_k (x_k^n)^2\, \mathbb{E}\left[N(A_k^n)^2\right] \\
&= \sum_{k \neq q} x_k^n x_q^n \nu(A_k^n)\nu(A_q^n) + \sum_k (x_k^n)^2 \left(\nu(A_k^n) + \nu(A_k^n)^2\right) \\
&= \left(\int_S g_n\, \nu(dx)\right)^2 + \int_S g_n^2 \nu(dx)
\end{aligned}$$

Un argumento similar al que se usó en (3.2.12) nos permite usar el teorema de convergencia dominada de ambos lados de la igualdad para concluir que $\mathbb{E}\left[Y^2\right] = \left(\int_S g\, \nu(dx)\right)^2 + \int_S g^2 \nu(dx)$. ■

Observación 3.2.2

Obtuvimos que la función característica de $Y = \int_S gN(dx)$ *es*

$$\mathbb{E}\left[e^{i\lambda Y}\right] = \exp\left\{-\int_S \left(1 - e^{i\lambda g(x)}\right) \nu(dx)\right\}. \tag{3.2.13}$$

Si ν_g*, la medida inducida por* g *en* $\mathbb{R}$*, es finita (es decir* $\nu_g(\mathbb{R}) = \nu(S) < \infty$*) entonces* $\kappa = \frac{1}{\nu_g(\mathbb{R})}\nu_g$ *es medida de probabilidad en* $\mathbb{R}$ *y por el teorema de cambio de variable:*

$$\begin{aligned}
\exp\left\{\int_S \left(e^{i\lambda g(x)} - 1\right) \nu(dx)\right\} &= \exp\left\{\nu_g(\mathbb{R}) \int_R \left(e^{i\lambda x} - 1\right) \kappa(dx)\right\} \\
&= \exp\left\{\nu(S)\left(\varphi_\kappa(\lambda) - 1\right)\right\}
\end{aligned}$$

por lo que Y *es una variable aleatoria Poisson compuesto de parámetro* $\nu(S)$ *y la ley de sus saltos está dada por* κ.

Corolario 3.2.3 *Usando la misma notación, si* $A, B \in \mathscr{S}$ *son dos conjuntos ajenos y las integrales* $\int_A |g|\nu(dx)$, $\int_B |g|\nu(dx)$ *son finitas, entonces las variables aleatorias* $Y_1 = \int_A |g|N(dx)$ *y* $Y_2 = \int_B |g|N(dx)$ *son independientes.*

Demostración

Se aproxima a $g\mathbb{1}_A$ y $g\mathbb{1}_B$ por una sucesión creciente (en valor absoluto) de simples $\{g_n^1\}$ y $\{g_n^2\}$ con soportes ajenos y se observa que para cada n, $Y_1^n = \int_S g_n^1 N(dx)$ y $Y_2^n = \int_S g_n^2 N(dx)$ son independientes por la condición (iii) de 3.1.1 y el lema 1.2.5. Al ser N sigma finita c.s., por el teorema de convergencia monótona tenemos que $Y_1^n(\omega) \to Y_1(\omega)$ y $Y_2^n(\omega) \to Y_2(\omega)$ c.s. Así, al usar el teorema de continuidad de Lévy es fácil comprobar que Y_1 y Y_2 son independientes. ■

Este resultado se extiende fácilmente a colecciones de la forma $(\int_{A_i} gN(dx))_{i\in\mathbb{N}}$ para $A_i \cap A_j = \emptyset$.

Capítulo 4

Procesos de Lévy

4.1 Fórmula de Lévy-Khintchine

Referencias: Bertoin [9] capítulo 1 y Kyprianou [13] capítulo 2

Sea $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ un espacio de probabilidad.

Definición 4.1.1 Proceso de Lévy
Un proceso $X = (X_t)_{t\geq 0}$ *definido en un espacio de probabilidad* $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ *con valores en* $\mathbb{R}$ *es llamado un proceso de Lévy si satisface las siguientes propiedades:*
i) Las trayectorias de X *son cadlag c.s.*
ii) $X_0 = 0$ *c.s.*
iii) Para $0 \leq s \leq t$, $X_t - X_s$ *es independiente de* $\{X_u : u \leq s\}$.
iv) Para $0 \leq s \leq t$, $X_t - X_s$ *tiene la misma distribución que* X_{t-s}.

Como consecuencia inmediata de la definición, para todo t la variable aleatoria X_t es infinitamente divisible: para cualquier n, podemos escribir a X como una suma de n variables aleatorias independientes idénticamente distribuidas

$$X_t = \sum_{k=1}^{n} X_{k\frac{t}{n}} - X_{(k-1)\frac{t}{n}} \qquad \text{pues} \qquad \left(X_{k\frac{t}{n}} - X_{(k-1)\frac{t}{n}}\right) \sim X_{\frac{t}{n}}.$$

Entonces, por lo que se probó en el capítulo 1, X_t y en particular X_1 debe tener una función característica de la forma:

$$\varphi(\lambda) = \exp\left\{ \int_{\mathbb{R}} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2} \mu(dx) \right\}$$

para alguna medida finita μ. Por la proposición 2.2.6 φ nunca se anula, lo cual nos lleva a la siguiente definición:

Definición 4.1.2 *A* Ψ *definido como*

$$\Psi(\lambda) = -\log\left(\varphi_{X_1}(\lambda)\right)$$

se le llama el exponente característico de X *y por lo tanto,*

$$\varphi_{X_1}(\lambda) = e^{-\Psi(\lambda)}.$$

Basta con conocer a Ψ para poder recuperar a la característica de X_t para cualquier t:

Proposición 4.1.3 *Si Ψ es el exponente característico de un proceso de Lévy X, entonces la función característica de la variable aleatoria X_t está dada por:*

$$\varphi_{X_t}(\lambda) = e^{-t\Psi(\lambda)}.$$

Demostración

Lo probaremos primero para $t = n \in \mathbb{N}$, luego para $t = \frac{p}{q} \in \mathbb{Q}$ y finalmente para t real. Denotaremos por φ_t a φ_{X_t} para aligerar la notación.

Al escribir $X_2 = X_1 + (X_2 - X_1)$ y usar el hecho de que X_1 y $(X_2 - X_1)$ son independientes y con misma ley, se sigue que

$$\varphi_2(\lambda) = \varphi_{X_1+(X_2-X_1)}(\lambda) = \varphi_1(\lambda) \cdot \varphi_1(\lambda) = (\varphi_1(\lambda))^2 .$$

Por lo tanto, $\varphi_2(\lambda) = \varphi_1^2(\lambda) = e^{-2\Psi(\lambda)}$. Lo anterior se puede hacer para cualquier natural, y entonces $\varphi_n(\lambda) = e^{-n\Psi(\lambda)}$ para todo $n \in \mathbb{N}$.

Ahora si $t = \frac{1}{n}$, escribimos $X_1 = \sum_{k=1}^{n} X_{k\frac{t}{n}} - X_{(k-1)\frac{1}{n}}$ donde las variables aleatorias $\left\{X_{k\frac{t}{n}} - X_{(k-1)\frac{1}{n}}\right\}_{k\leq n}$ son independientes y con distribución $X_{\frac{1}{n}}$. Similarmente al caso anterior:

$$\varphi_1(\lambda) = \varphi_{\left\{\sum_k X_{k\frac{t}{n}} - X_{(k-1)\frac{1}{n}}\right\}}(\lambda) = \left(\varphi_{\frac{1}{n}}(\lambda)\right)^n$$

por lo que

$$\varphi_{\frac{1}{n}}(\lambda) = (\varphi_1(\lambda))^{\frac{1}{n}} = e^{-\frac{1}{n}\psi(\lambda)}.$$

Si $t = \frac{p}{q}$ con p, q enteros, entonces $X_{\frac{p}{q}} = \sum_{k=1}^{p} X_{\frac{k}{q}} - X_{\frac{k-1}{q}}$ donde las variables aleatorias $\left\{X_{\frac{k}{q}} - X_{\frac{k-1}{q}}\right\}_{k\leq p}$ son independientes con distribución $X_{\frac{1}{q}}$. Se sigue que

$$\varphi_{\frac{p}{q}}(\lambda) = \left(\varphi_{\frac{1}{q}}(\lambda)\right)^p = \left(e^{-\frac{1}{q}\Psi(\lambda)}\right)^p = e^{-\frac{p}{q}\Psi(\lambda)}.$$

Finalmente, si $t \in \mathbb{R}$, consideramos (q_n) sucesión decreciente de racionales tal que $q_n \downarrow t$. Por continuidad por la derecha de las trayectorias de X,

$$\varphi_t(\lambda) = \mathbb{E}\left[e^{i\lambda X_t}\right] = \lim_{n\to\infty} \mathbb{E}\left[e^{i\lambda X_{q_n}}\right] = \lim_{n\to\infty} \varphi_{q_n}(\lambda) = \lim_{n\to\infty} e^{-q_n\Psi(\lambda)} = e^{-t\Psi(\lambda)}$$

con lo que concluimos. ■

A continuación, probamos que el exponente característico Ψ de un proceso de Lévy caracteriza por completo sus distribuciones finito-dimensionales, y por lo tanto su ley en $D(\mathbb{R}^+, \mathbb{R})$:

Proposición 4.1.4 *Sean X y Y procesos de Lévy en espacios de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ y $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ con el mismo exponente Ψ. Entonces para $A_i \in \mathscr{B}(\mathbb{R})\, i = 1, \ldots n$ y $0 \leq t_1 < t_2 < \cdots < t_n$ cualesquiera se satisface que*

$$\mathbb{P}\left(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right) = \tilde{\mathbb{P}}\left(Y_{t_1} \in A_1, \ldots, Y_{t_n} \in A_n\right).$$

Demostración
Como X y Y tienen el mismo exponente característico Ψ, por la proposición anterior $\varphi_{X_t} = e^{-t\Psi} = \varphi_{Y_t}$ y entonces las variables aleatorias X_t y Y_t tienen la misma ley para cada $t \in \mathbb{R}^+$. Como por estacionariedad de los incrementos

$$\mathbb{P}(X_t - X_s \in A) = \mathbb{P}(X_{t-s} \in A) = \mathbb{P}(Y_{t-s} \in A) = \mathbb{P}(Y_t - Y_s \in A), \tag{4.1.1}$$

se sigue que

$$\begin{aligned}\mathbb{P}\left(X_{t_1} - X_{t_2} \in A\right) \cdot \mathbb{P}\left(X_{t_2} - X_{t_3} \in A\right) \cdot \ldots \cdot \mathbb{P}\left(X_{t_n} - X_{n-1} \in A\right)\\ = \mathbb{P}\left(Y_{t_1} - Y_{t_2} \in A\right) \cdot \mathbb{P}\left(Y_{t_2} - Y_{t_3} \in A\right) \cdot \ldots \cdot \mathbb{P}\left(Y_{t_n} - Y_{n-1} \in A\right).\end{aligned}$$

Ahora, por independencia de los incrementos,

$$\begin{aligned}\mathbb{P}\left(X_{t_1} \in A_1, X_{t_2} - X_{t_1} \in A_2 \ldots, X_{t_n} - X_{t_{n-1}} \in A_n\right)\\ = \tilde{\mathbb{P}}\left(Y_{t_1} \in A_1, Y_{t_2} - Y_{t_1} \in A_2 \ldots, Y_{t_n} - Y_{t_{n-1}} \in A_n\right)\end{aligned} \tag{4.1.2}$$

por lo que las leyes inducidas en $\mathscr{B}(\mathbb{R}^d)$ por los vectores

$$W = \left(X_{t_1}, X_{t_2} - X_{t_1} \ldots, X_{t_n} - X_{t_{n-1}}\right) \quad \text{y} \quad Z = \left(Y_{t_1}, Y_{t_2} - Y_{t_1} \ldots, Y_{t_n} - Y_{t_{n-1}}\right)$$

coinciden en la colección de conjuntos de la forma $\{A_1 \times \cdots \times A_n\,;\, A_i \in \mathscr{B}(\mathbb{R})\}$, que es un π-sistema generador de $\mathscr{B}(\mathbb{R}^d)$. Así, X y Y inducen la misma ley en $\mathscr{B}(\mathbb{R}^d)$ y entonces para cualquier $D \in \mathscr{B}(\mathbb{R}^d)$ y $f = \mathbb{1}_D$, se tiene que

$$\mathbb{E}\left[f(W)\right] = \tilde{\mathbb{E}}\left[f(Z)\right]. \tag{4.1.3}$$

Por linealidad, (4.1.3) también se satisface para f simple y por el teorema de convergencia dominada, se extiende para f medible y acotada. En particular, si $f = \mathbb{1}_D \circ g$ con $g(x_1, \ldots, x_n) = (x_1, x_2 + x_1, \ldots, x_n + x_{n-1})$ y $D = A_1 \times \cdots \times A_n$, la igualdad se convierte en

$$\mathbb{P}\left(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right) = \tilde{\mathbb{P}}\left(Y_{t_1} \in A_1, \ldots, Y_{t_n} \in A_n\right).$$

■

La teoría desarrollada en el capítulo 1 sobre variables aleatorias infinitamente divisibles, nos permite probar

Teorema 4.1.5 Fórmula de Lévy-Khintchine
Si X es un proceso de Lévy en $\mathbb{R}$ con exponente característico Ψ, entonces existe una tripleta (a, q^2, Π) donde a, q son números reales, Π es una medida que satisface $\Pi(\{0\}) = 0$, $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$ que permite escribir a Ψ como:

$$\Psi(\lambda) = ia\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}}\left(1 - e^{i\lambda x}\right)\mathbb{1}_{\{|x| \geq 1\}}\Pi(dx) + \int_{\mathbb{R}}\left(1 - e^{i\lambda x} + i\lambda x\right)\mathbb{1}_{\{|x|<1\}}\Pi(dx) \tag{4.1.4}$$

Una medida Π que satisface $\int 1 \wedge x^2\Pi(dx) < \infty$ se le llama medida de Lévy y a la tripleta (a, q^2, Π) se le conoce como la tripleta característica de X.

Demostración
Por los teoremas 2.2.11, 2.2.7, 2.2.8, si X es infinitamente divisible y de media 0, su exponente característico se escribe:

$$-\Psi(\lambda) = -\mu(\{0\})\frac{\lambda^2}{2} + \int_{\mathbb{R}\setminus 0} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2}\mu(dx)$$

para alguna medida finita μ. Si $\mathbb{E}[X_1] = a' \neq 0$, entonces la variable aleatoria $X_1 - a$ es infinitamente divisible de media 0, por lo que X tiene por exponente característico:

$$-\Psi(\lambda) = ia'\lambda - \mu(\{0\})\frac{\lambda^2}{2} + \int_{\mathbb{R}\setminus 0} \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2}\mu(dx). \tag{4.1.5}$$

Definimos $\Pi(dx) = \mathbb{1}_{x\neq 0}\frac{1}{x^2}\mu(dx)$. Esta nueva medida satisface que $\int_{\mathbb{R}} 1 \wedge x^2\Pi(dx) < \infty$ puesto que

$$\begin{aligned}
\int_{\mathbb{R}} 1 \wedge x^2\Pi(dx) &= \int_{\{|x|<1\}} x^2\Pi(dx) + \int_{\{|x|\geq 1\}} 1\Pi(dx) \\
&= \int_{\{0<|x|<1\}} 1\,\mu(dx) + \int_{\{|x|\geq 1\}} \frac{1}{x^2}\mu(dx) \leq \mu(\mathbb{R}\setminus 0) < \infty
\end{aligned}$$

y $\Pi(\{0\}) = 0$ (μ es una medida finita pero nótese que Π ya no tiene por que serlo). Notemos que

$$\int_{\{|x|\geq 1\}} x\Pi(dx) = \int_{\{|x|\geq 1\}} \frac{1}{x}\mu(dx) \leq \mu(\{|x| \geq 1\}) < \infty,$$

lo cual nos permite reescribir la expresión (4.1.5):

$$\begin{aligned}
\Psi(\lambda) &= -ia'\lambda + \mu(\{0\})\frac{\lambda^2}{2} + \int_{\mathbb{R}\setminus 0} \frac{1 - e^{i\lambda x} + i\lambda x}{x^2}\mu(dx) \\
&= -ia'\lambda + \lambda\int_{\mathbb{R}} \left(ix\mathbb{1}_{|x|\geq 1}\right)\Pi(dx) + \mu(\{0\})\frac{\lambda^2}{2} && (4.1.6) \\
&\quad + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right)\Pi(dx) - \int_{\mathbb{R}} \left(ix\lambda\mathbb{1}_{|x|\geq 1}\right)\Pi(dx) \\
&= ia\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\mathbb{1}_{\{|x|<1\}}\right)\Pi(dx) && (4.1.7) \\
&= ia\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right)\mathbb{1}_{\{|x|\geq 1\}}\Pi(dx) + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right)\mathbb{1}_{\{|x|<1\}}\Pi(dx)
\end{aligned}$$

donde $a = -ia' + \int_{\mathbb{R}} \left(ix\mathbb{1}_{|x|\geq 1}\right)\Pi(dx)$ y $q^2 = \mu(\{0\})$. ■

A la expresión (4.1.7) también se le conoce como la Fórmula de Lévy-Khintchine y usaremos ambas expresiones a lo largo del trabajo. La elección de sumar y restar $\lambda\int_{\mathbb{R}} \left(ix\mathbb{1}_{|x|\geq 1}\right)\Pi(dx)$ en (4.1.6) no es (del todo) arbitraria: por desarrollo de Taylor

$$1 - e^{i\lambda x} + i\lambda x = 1 - \left(1 + i\lambda x + \frac{(i\lambda x)^2}{2} + O(x^3)\right) + i\lambda x$$

que es del orden de $O(x^2)$. Así, en $\{|x| \leq 0\}$ la integral con respecto a Π es finita al cumplirse que $\int 1 \wedge x^2\Pi(dx) < \infty$ La fórmula de Lévy-Kintchine será indispensable para estudiar el comportamiento trayectorial de un proceso de Lévy y nos permitirá probar el teorema central de este capítulo, que es la descomposición de Lévy-Itô. La medida Π que aparece en la expresión podría resultar un tanto misteriosa; a lo largo del capítulo iremos esclareciendo el papel que juega en el comportamiento trayectorial del proceso.

Así, buscaremos identificar la naturaleza de los términos que aparecen en (4.1.4) y explicar como influyen en comportamiento de X. En concreto el objetivo es el siguiente: dado un proceso de Lévy X con terna (a, q^2, Π) y exponente Ψ, construir tres procesos de Lévy independientes X_1, X_2, X_3 tales que $X = X_1 + X_2 + X_3$ en distribución, de modo que

$$\Psi(\lambda) = \underbrace{ia\lambda + q^2\frac{\lambda^2}{2}}_{(1)} + \underbrace{\int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \mathbb{1}_{\{|x|\geq 1\}}\Pi(dx)}_{(2)} + \underbrace{\int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right) \mathbb{1}_{\{|x|<1\}}\Pi(dx)}_{(3)}$$

donde el (i)-esimo término de esta expresión corresponda al exponente Ψ^i del proceso X^i.

Dos ejemplos canónicos:

Asumiremos la existencia y las propiedades básicas del movimiento browniano y del proceso Poisson. Las definiciones de estos procesos se pueden hallar en el apéndice.

1- El movimiento browniano

Si $B = (B_t)$ es un movimiento browniano, es claro que se trata de un proceso de Lévy. Como B_1 tiene distribución $\mathcal{N}(0,1)$, $\varphi_{B_1} = e^{-\frac{\lambda^2}{2}}$ y el exponente característico Ψ de B es entonces $\Psi(\lambda) = \frac{\lambda^2}{2}$. La tripleta característica de B está entonces dada por $a = 0$, $\Pi = 0$ y $q^2 = 1$ y por lo tanto, la medida finita μ es la delta de dirac en 0 (pues recordemos de la prueba de Lévy-Khintchine que $q^2 = \mu(\{0\}) = 1$).

2- *El proceso Poisson compuesto y el proceso Poisson*

Sea $N = (N_t)$ un proceso Poisson de parémtro c. Recordemos que $N_t \sim Poisson(ct)$ y por lo tanto $\mathbb{E}[N_t] = ct$. Como el valor N_t coincide con el número de saltos de N hasta tiempo t, el parámetro c gobierna la frecuencia con la que N salta. Esta es una observación que usaremos más adelante.

Definición 4.1.6 *Si N es un proceso Poisson de parámetro c y $\{Y_i\}_{i\in\mathbb{N}}$ es una colección de variables aletorias independientes idénticamente distribuidas e independientes de N, definimos al proceso Poisson compuesto:*

$$C_t = \sum_{i=1}^{N_t} Y_i.$$

donde por convención, $\sum_{i=k+1}^{k} Y_i = 0$. Este es un proceso similar al proceso Poisson, solo que ahora los saltos no son todos de tamaño uno, son de tamaño aleatorio y tienen por distribución la ley de Y_1. El número de saltos que ha dado C hasta el tiempo t viene dado por el valor de N_t por lo que que la frecuencia con la que C salta coincide con la de N (e insistimos en que esta viene dada por el parámetro c). Si $Y_i \equiv 1$, C es el proceso Poisson N.

Proposición 4.1.7 *El proceso Poisson compuesto C es un proceso de Lévy y su exponente característico está dado por:*

$$\Psi(\lambda) = c\int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \mathbb{P}_{Y_1}(dx) = c\left(1 - \varphi_Y(\lambda)\right). \tag{4.1.8}$$

A c se le llama el parámetro del Poisson compuesto y a $\mathbb{P}_{Y_1}$ la ley de sus saltos.

Demostración
Empezamos calculando la función característica de C_t:

$$\begin{aligned}
\mathbb{E}\left[\exp\{i\lambda C_t\}\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[\exp\left\{i\lambda \sum_{i=1}^{n} Y_i\right\}\middle| N_t = n\right] \mathbb{P}\left(N_t = n\right) \\
&= \sum_{n=0}^{\infty} \prod_{i=1}^{n} \mathbb{E}\left[\exp\left\{i\lambda Y_i\right\}\right] \mathbb{P}\left(N_t = n\right) \\
&= e^{-ct} \sum_{n=0}^{\infty} \varphi_{Y_1}(\lambda)^n \frac{(ct)^n}{n!} \\
&= e^{ct\left(\varphi_{Y_1}(\lambda)-1\right)}.
\end{aligned}$$

De esto, se sigue (4.1.8) al evaluar en $t = 1$. Ahora veamos que si $t > s$, entonces $C_t - C_s$ y C_{t-s} tienen la misma ley. Obsérvese que $C_t = C_s + \sum_{i=N_s+1}^{N_t} Y_i$ y entonces, similarmente a lo que se hizo antes:

$$\begin{aligned}
\varphi_{C_t - C_s}(\lambda) &= \mathbb{E}\left[\exp\left\{i\lambda \sum_{i=N_s+1}^{N_t - N_s + N_s} Y_i\right\}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\exp\left\{i\lambda \sum_{i=n+1}^{N_t - N_s + n} Y_i\right\}\middle| N_s = n\right] \mathbb{P}\left(N_s = n\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[\exp\left\{i\lambda \sum_{i=n+1}^{N_{t-s} + n} Y_i\right\}\right] \mathbb{P}\left(N_s = n\right) \\
&= \mathbb{E}\left[\exp\left\{i\lambda \sum_{i=1}^{N_{t-s}} Y_i\right\}\right] \cdot \sum_{n=0}^{\infty} \mathbb{P}\left(N_s = n\right) \\
&= \varphi_{C_{t-s}}(\lambda)
\end{aligned}$$

donde en la segunda igualdad se utilizó que si dos variables X y Y son independientes, entonces $\mathbb{E}\left[H(X,Y)|Y\right] = \int_{\mathbb{R}} H(x,Y) d\mathbb{P}_X(dx)$ (pues N tiene incrementos independientes) y para probar la cuarta igualdad se sigue un procedimiento idéntico al que se usó para calcular la función característica de C_t. Falta probar que $C_t - C_s$ es independiente de $\sigma\left(C_u\,;\ u \leq s\right)$ es decir, que para cualquier colección finita de tiempos $0 = t_1 < t_2 < \cdots < t_n = s$, $C_t - C_s$ es independiente de $\sigma\left(X_{t_i}\,; i = 1, \cdots, n\right)$, la cual coincide con a la σ-álgebra generada por los incrementos $\sigma\left(X_{t_1},\, X_{t_{i+1}} - X_{t_i}\,; i = 1, \cdots, n-1\right)$ (ver lema 9.1.1 del Apéndice). Lo probaremos para $t_1 = s$ pero un procedimiento análogo permite probarlo para colecciones finitas de incrementos. Veamos entonces que las variables aleatorias $C_t - C_s$ y C_s son independientes:

$$\begin{aligned}
\varphi_{(C_t-C_s,C_s)}(\lambda_1,\lambda_2) &= \mathbb{E}\left[\exp\left\{i\lambda_1\sum_{i=1}^{N_s}Y_i + i\lambda_2\sum_{i=N_s+1}^{N_t-N_s+N_s}Y_i\right\}\right]\\
&= \sum_{n=0}^{\infty}\mathbb{E}\left[\exp\left\{i\lambda_1\sum_{i=1}^{n}Y_i + i\lambda_2\sum_{i=n+1}^{N_t-N_s+n}Y_i\right\}\middle| N_s=n\right]\mathbb{P}\left(N_s=n\right)\\
&= \sum_{n=0}^{\infty}\mathbb{E}\left[\exp\left\{i\lambda_1\sum_{i=1}^{n}Y_i\right\}\right]\cdot\mathbb{E}\left[\left\{i\lambda_2\sum_{i=n+1}^{N_t-N_s+n}Y_i\right\}\right]\mathbb{P}\left(N_s=n\right)\\
&= \mathbb{E}\left[\left\{i\lambda_2\sum_{i=1}^{N_{t-s}}Y_i\right\}\right]\cdot\sum_{n=0}^{\infty}\mathbb{E}\left[\exp\left\{i\lambda_1\sum_{i=1}^{n}Y_i\right\}\middle| N_s=n\right]\mathbb{P}\left(N_s=n\right)\\
&= \varphi_{C_{t-s}}(\lambda_2)\cdot\mathbb{E}\left[\exp\left\{i\lambda_1\sum_{i=1}^{N_s}Y_i\right\}\right]\\
&= \varphi_{C_{t-s}}(\lambda_2)\cdot\varphi_{C_s}(\lambda_1)
\end{aligned}$$

Como la continuidad por la derecha se sigue de la continuidad por la derecha de N, concluimos entonces que C es un proceso de Lévy cuyo exponente característico Ψ está dado por $\Psi(\lambda) = c(1-\varphi_{Y_1}(\lambda))$. ■

Ejemplo Consideremos el siguiente caso particular: si las variables aleatorias $\{Y_i\}_{i\in\mathbb{N}}$ tienen distribución Bernoulli con parámetro p, entonces $\varphi_{Y_1}(\lambda) = pe^{i\lambda}+1-p$ por lo que el exponente característico de C es $\Psi(\lambda) = cp(1-e^{i\lambda})$. Es decir, C resulta ser un proceso Poisson con parámetro cp.

Ya tenemos entonces que C tiene por exponente

$$\Psi(\lambda) = c\int_{\mathbb{R}}\left(1-e^{i\lambda x}\right)\mathbb{P}_{Y_1}(dx)$$

Como ejercicio, nos preguntamos por la tripleta característica que nos permite escribir a Ψ como en (4.1.7), su "forma canónica" de variable aleatoria infinitamente divisible. Si en (4.1.7) consideramos $q^2 = 0$ y medida de Lévy $\Pi(dx) = c\mathbb{P}_{Y_1}(dx)$, al reemplazar obtenemos:

$$ia\lambda + \int_{\mathbb{R}}\left(1-e^{i\lambda x}+i\lambda x\mathbb{1}_{\{|x|<1\}}\right)c\mathbb{P}_{Y_1}(dx)$$

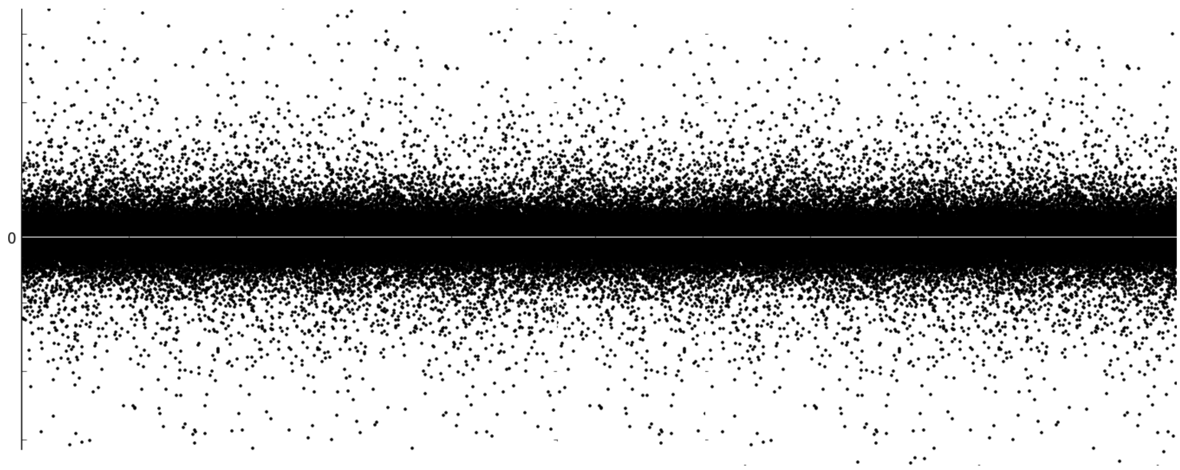
así que bastaría definir $a = -\int_{\mathbb{R}} x\mathbb{1}_{\{|x|<1\}}c\mathbb{P}_{Y_1}(dx)$ para recuperar Ψ.

De esto, se puede deducir fácilmente la tripleta del proceso Poisson N, que es un caso particular de lo anterior ya que recordemos que $N_t = \sum_{i=1}^{N_t}1$ (es decir, $Y_1 \equiv 1$). En ese caso, la medida de Lévy está dada por $\Pi = c\mathbb{P}_{Y_1} = c\delta_1$ y entonces $a = 0$. Obsérvese que la medida Π en los tres ejemplos lleva la información del tamaño y frecuencia de los saltos (es decir, la distribución de Y_i y el parámetro de C). En el caso del browniano, al ser continuo, no tiene saltos y $\Pi = 0$. Esta observación será más evidente una vez hayamos probado la descomposición de Lévy Itô. Para esto, fue necesario introducir a las medidas aleatorias de Poisson. Con esto finalizamos el ejemplo y seguimos nuestro camino.

De aquí en adelante, trabajaremos con medidas aleatorias de Poisson N definidas en el espacio

$$(S, \mathscr{S}, \nu) = ([0,\infty) \times \mathbb{R}, \mathscr{B}([0,\infty)) \otimes \mathscr{B}(\mathbb{R}), dt \times \Pi(dx)) \tag{4.1.9}$$

donde dt es la medida de Lebesgue λ en $\mathbb{R}^+$ y Π es una medida de Lévy en $\mathbb{R}$ y por lo tanto $\Pi(\{0\}) = 0$. La siguiente simulación corresponde a una medida Poisson con intensidad $dt \times \Pi$ al utilizar una medida de Lévy Π estable $1/2$, es decir, $\Pi(dx) = \frac{1}{|x|^{1/2}}dx$. Cada punto corresponde a un átomo de N para esa ocurrencia.



Entonces, $S = [0,\infty) \times \mathbb{R}$ y la medida soporte de N es $\nu = dt \times \Pi(dx)$. El intervalo $[0,\infty)$ se debe pensar como un parámetro temporal mientras que $\mathbb{R}$ como espacial. Nótese que ν no es atómica pues la medida de Lebesgue dt no lo es. Así, $\mathbb{P}\left(N(\{(x,t)\}) > 0\right) = 0$ y nótese que aunque Π no tiene por que ser finita, solo puede ser infinita en vecindades del 0. Si es infinita, cualquier rectángulo abierto $R_1 \times R_2$ que contenga algún punto de $\mathbb{R} \times 0$ tendrá medida $N(R_1 \times R_2) = \infty$ casi seguramente pues $\nu(R_1 \times R_2) = \lambda(R_1) \times \Pi(R_2) = \infty$ como se puede apreciar en la figura anterior.

Solo haremos uso de la fórmula de Campbell 3.2.1 para funciones $g : S \to \mathbb{R}$ de la forma

$$g(x,s) = x\mathbb{1}_{[0,t]}(s)\mathbb{1}_B(x) \tag{4.1.10}$$

para $B \in \mathscr{B}(\mathbb{R})$. Por lo tanto, las integrales y condiciones del teorema de Campbell se simplifican: por ejemplo, las integrales se reducen a expresiones de la forma $\int_S g(x,s)\nu(ds \times dx) = t\int_B x\,\Pi(dx)$. Es conveniente reescribir el teorema 3.2.1 para g como en (4.1.10) en vista de que se usará una y otra vez:

Teorema 4.1.8 *Sea N medida aleatoria de Poisson en (4.1.9). Para $B \in \mathscr{B}(\mathbb{R})$ y t fijo, definimos*

$$Y_t = \int_{[0,t]\times B} xN(ds \times dx). \tag{4.1.11}$$

Entonces:

i)

$$\int_{[0,t]\times B} |x|N(ds \times dx) < \infty \quad c.s. \qquad si\ y\ solo\ si \qquad \int_B (1 \wedge |x|)\,\Pi(dx) < \infty. \tag{4.1.12}$$

ii) Si $\int_B (1 \wedge |x|)\, \Pi(dx) < \infty$*, entonces*

$$\mathbb{E}\left[e^{i\lambda Y_t}\right] = \exp\Big\{ -t \int_B \left(1 - e^{i\lambda x}\right) \Pi(dx)\Big\} \tag{4.1.13}$$

para todo $\lambda \in \mathbb{R}$*. Si además* $\Pi(B)$ *es finita, la variable aleatoria* Y_t *tiene distribución Poisson compuesto (estos puntos se verán con detalle en la prueba del lema 4.1.9).*

iii) Si $\int_B |x|\, \Pi(dx) < \infty$*,*

$$\mathbb{E}\left[Y_t\right] = t \int_B x\, \Pi(dx) \tag{4.1.14}$$

y si además, $\int_B |x|^2\, \Pi(dx) < \infty$*,*

$$\mathbb{E}\left[Y_t^2\right] = t \int_B x^2\, \Pi(dx) + \left(t \int_B x\, \Pi(dx)\right)^2 \tag{4.1.15}$$

A partir de una medida de Poisson, podemos construir procesos Poisson compuestos al hacer variar el parámetro t en (4.1.11):

Lema 4.1.9 *Sea* $B \in \mathscr{B}(\mathbb{R})$ *tal que* $0 < \Pi(B) < \infty$*. Entonces, el proceso*

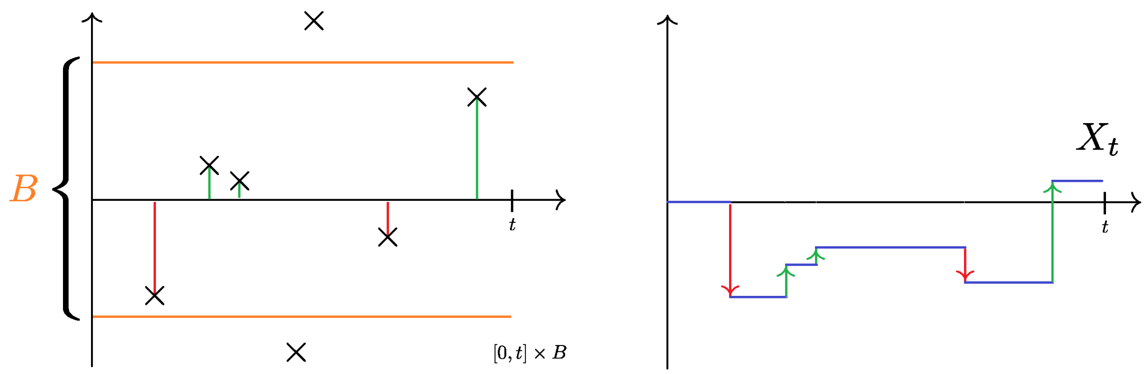
$$X_t = \int_{[0,t]} \int_B x\, N(ds \times dx), \qquad t \geq 0 \tag{4.1.16}$$

es un proceso Poisson compuesto con parámetro $\Pi(B)$ *y distribución de saltos* $\frac{1}{\Pi(B)}\Pi|_B$ *(*Π *restringida a* B*). Además, su exponente característico es*

$$\Psi(\lambda) = \int_B \left(1 - e^{i\lambda x}\right) \Pi(dx). \tag{4.1.17}$$

Si $\pi(\omega)$ es el soporte de $N(\cdot,\omega)$, así definido X_t no es más que la suma acumulada de las coordenadas espaciales de $(x,s) \in \pi(\omega)$ con $x \in B$ y con coordenada temporal menor a t:

$$\int_{[0,t]} \int_B x\, N(ds \times dx) = \sum_{(x,t)\in\pi\cap[0,t]\times B} x \tag{4.1.18}$$



Demostración

Como $\Pi(B) < \infty$, $\nu([0,t) \times B) < \infty$ y entonces el soporte de la medida N en $[0,t) \times B$ es finito casi seguramente. Se sigue que para cada t, (4.1.18) es una suma finita y por lo tanto X_t es cadlag. Primero, veamos que X_t tiene distribución Poisson. Escribimos

$$X_t = \int_S x \mathbb{1}_{[0,t]} \mathbb{1}_B \, N(ds \times dx)$$

y la hipótesis $\Pi(B) < \infty$ garantiza que $\int_B (1 \wedge |x|) \Pi(dx) < \infty$ y por (3.2.3), se tiene que

$$\begin{aligned}
\varphi_{X_t}(\lambda) &= \exp\left\{ \int_S \left(e^{i\lambda x \mathbb{1}_{[0,t]} \mathbb{1}_B} - 1 \right) ds \times \Pi(dx) \right\} \\
&= \exp\left\{ \int_S \mathbb{1}_{[0,t]} \mathbb{1}_B \left(e^{i\lambda x} - 1 \right) ds \times \Pi(dx) \right\} \\
&= \exp\left\{ \int_B \int_0^t \left(e^{i\lambda x} - 1 \right) ds \, \Pi(dx) \right\} \\
&= \exp\left\{ t \int_B \left(e^{i\lambda x} - 1 \right) \Pi(dx) \right\} \\
&= \exp\left\{ t\Pi(B) \int_{\mathbb{R}} \left(e^{i\lambda x} - 1 \right) \frac{\Pi|_B(dx)}{\Pi(B)} \right\}
\end{aligned}$$

que en efecto es la función característica de una variable Poisson compuesto con parámetro $t\Pi(B)$ y ley de saltos $\frac{\Pi|_B(dx)}{\Pi(B)}$. Esto quiere decir que los saltos solo pueden tomar magnitudes $\Delta X_t \in B$. Ahora, veamos que los incrementos son independientes y estacionarios. Si $t > s$,

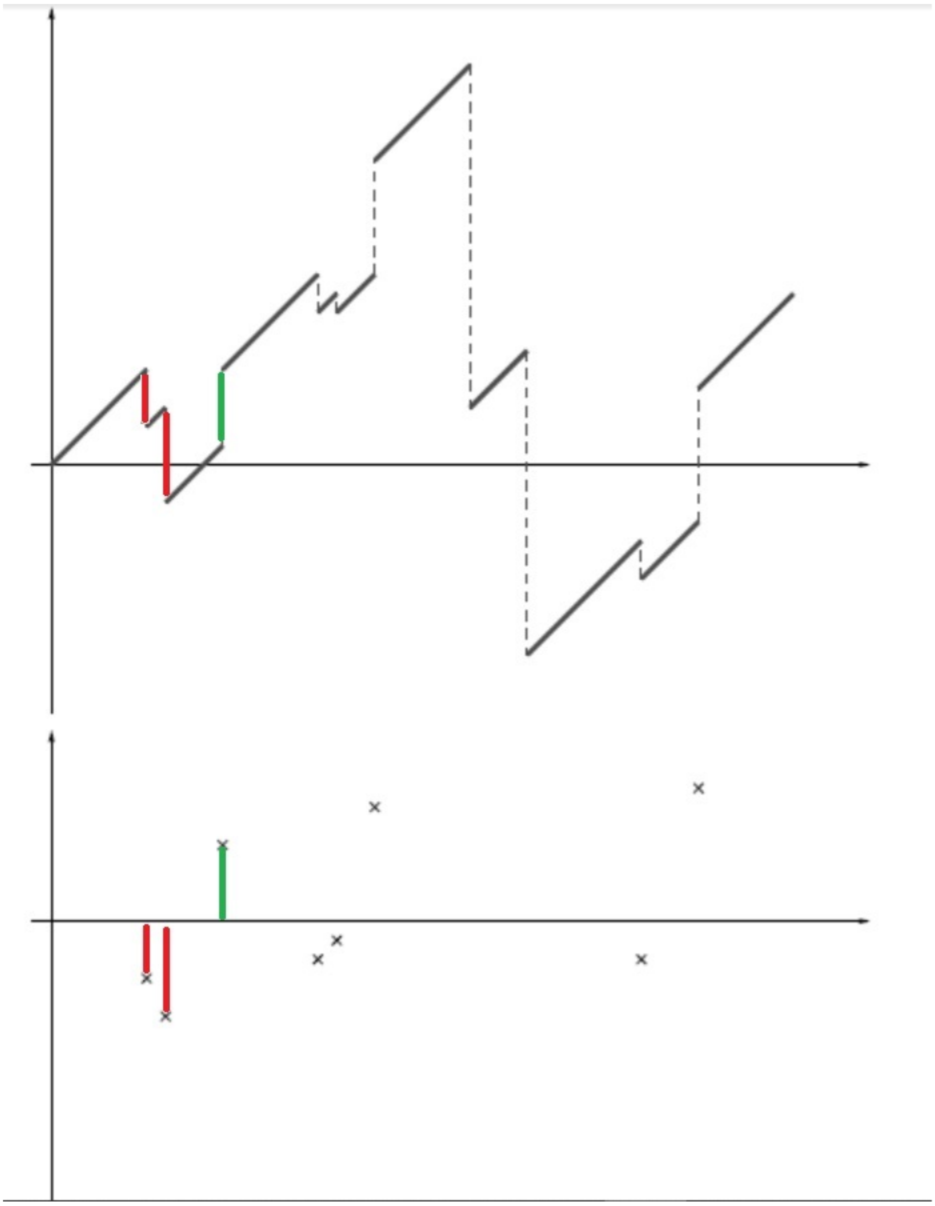
$$X_t - X_s = \int_{[0,t) \times B} x \, N(ds \times dx) - \int_{[0,s] \times B} x \, N(ds \times dx) = \int_{(s,t] \times B} x \, N(ds \times dx) \qquad (4.1.19)$$

así que al aplicar (4.1.13), la característica de $X_t - X_s$ y la de X_{t-s} está dada por

$$\exp\left\{ (t-s) \int_B \left(e^{i\lambda x} - 1 \right) \Pi(dx) \right\}$$

y por lo tanto los incrementos son estacionarios. Si consideramos incrementos $X_{t_i} - X_{t_{i-1}}$ con $t_0 < \cdots < t_n$ al usar (4.1.19) tenemos integrales con respecto a $N(ds \times dx)$ sobre conjuntos ajenos $(t_{i-1}, t_i] \times B$. Por el corolario 3.2.3, la colección formada por los incrementos está compuesta por variables aleatorias independiente. ■

Nos detendremos a analizar la naturaleza de los saltos de un proceso X definido como en (4.1.16) y su relación con la medida de Lévy. De la discusión que se tuvo, es claro que el número de saltos que tiene el proceso X hasta un tiempo t, es el número de átomos de la medida N en $[0,t] \times B$ que es justamente $\#\pi(\omega) \cap ([0,t] \times B)$. $X(\omega)$ permanece constante en un intervalo de la forma $[t_i, t_{i+1})$ siempre y cuando la medida $N(\,\cdot\,, \omega)$ no tenga ningún átomo de la forma (x,t) con $x \in B$ y $t \in [t_i, t_{i+1})$. Si (x,t) con $x \in B$ es átomo de $N(\,\cdot\,, \omega)$, $X(\omega)$ tendrá un salto en t de magnitud x.



La medida de Lévy Π caracteriza la frecuencia y magnitud de estos: como por el lema 4.1.9 el soporte de la distribución de los saltos es B, solo podrán ser de *magnitud* r con $r \in B$; por ejemplo, si $B = [1, \infty)$, los saltos de X serán positivos y de magnitud mayor o igual que 1. En resumen, ***el soporte de*** Π ***caracteriza el tamaño que pueden tener los saltos***.

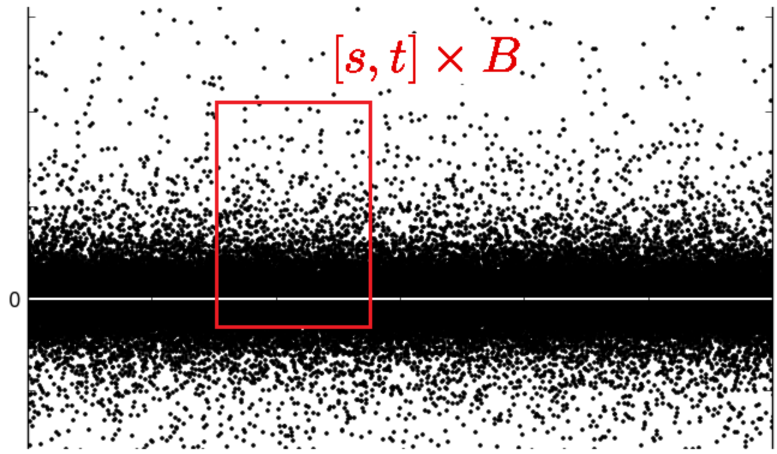
Recordemos que la *frecuencia* con la que ocurren los saltos está determinada por el parámetro del proceso Poisson compuesto X. En este caso, es justamente $\Pi(B)$ por lo que, mientras más grande sea este valor, más probable es que haya una gran cantidad de saltos en un intervalo de tiempo determinado. Sin embargo, el hecho de que $\Pi(B)$ sea finita garantiza que habrá una cantidad finita casi seguramente de saltos en cualquier intervalo finito de tiempo $[0, t]$ ya que recordemos que el número de átomos de N en $[0, t] \times B$ se distribuye Poisson N parámetro $\Pi \times dt\,(B \times [0, t]) < \infty$. En resumen, ***la magnitud de*** $\Pi(B)$ ***caracteriza la frecuencia con la que ocurren los saltos***.

Estas observaciones nos permiten comprender mejor la definición de una medida de Lévy así como las condiciones que aparecen en la fórmula de Campbell. Recuérdese que para que una medida $\Pi(dx)$ sea llamada medida de Lévy, debe cumplir que $\int_{\mathbb{R}}(1 \wedge x^2)\Pi(dx) < \infty$. Esto tiene varias consecuencias: primero, cualquier boreliano B que se quede contenido en el complemento de una vecindad abierta del origen $B(0,\epsilon)$, tiene Π-medida finita y por lo tanto el proceso definido en (4.1.16) será Poisson compuesto de parámetro $\Pi(B)$, tendrá un cantidad finita de saltos en intervalos de tiempo finito y estos serán de tamaño al menos ϵ. En particular, se cumple para abiertos de la forma $B=(-\infty,-\epsilon)\cup(\epsilon,+\infty)$. Como ya se había mencionado Π solo puede ser infinita en vecindades que contengan al origen.

Ya tenemos una buena interpretación del caso $\Pi(B)<\infty$. Ahora, sigamos considerando expresiones de la forma

$$X_t=\int_{[0,t]}\int_B x\,N(ds\times dx),\qquad t\geq 0 \tag{4.1.20}$$

y veamos que puede ocurrir cuando $\Pi(B)=\infty$. Sea entonces B un abierto que contiene al 0 y Π una medida de Lévy no finita. Como, $dt\times\Pi(dx)([s,t]\times B)=\infty$ entonces $N([s,t]\times B)=\infty$ c.s. por lo que el soporte de $N(\cdot,\omega)$ se ve como en la siguiente figura:



Distinguiremos dos casos, cuando

$$\int_{\mathbb{R}} 1\wedge|x|\Pi(dx)<\infty \quad \text{y cuando únicamente} \quad \int_{\mathbb{R}} 1\wedge x^2\Pi(dx)<\infty.$$

La segunda condición se tiene siempre por definición de la medida de Lévy mientras que la primera es una condición más fuerte.

Caso 1: Una primer posibilidad es entonces que la medida de Lévy satisfaga la condición adicional $\int_{\mathbb{R}}(1\wedge|x|)\Pi(dx)<\infty$ con $\Pi(B)=\infty$. El proceso X como se definió en (4.1.20) ya no es un proceso Poisson compuesto pero se puede escribir como una superposición de procesos Poisson independientes: por ejemplo si $B=\mathbb{R}$, denotamos $A_0=[-1,1]^c$ y $A_n=[-\frac{1}{n},-\frac{1}{n+1})\cup(\frac{1}{n+1},\frac{1}{n}]$ y descomponemos:

$$X_t=\int_{[0,t]}\int_{\mathbb{R}} x\,N(ds\times dx)=\sum_{n=0}^{\infty}\int_{[0,t]\times A_n} x\,N(ds\times dx). \tag{4.1.21}$$

Cada $\int_{[0,t]\times A_n} xN(ds\times dx)$ es un proceso Poisson compuesto con saltos cuya magnitud cae en A_n al tener A_n Π-medida finita. La independencia de los sumandos de (4.1.21) se sigue de 3.2.3. La función característica de X viene dada por (4.1.13) pues estamos dentro de las hipótesis de Campbell y la expresión (4.1.21) está bien definida por (4.1.12): a pesar de tener X una infinidad de saltos en cualquier intervalo de tiempo, la suma de estos es finita para cada t pues

$$\sum_{s\leq t}|\Delta X_s| = \int_{[0,t]}\int_{\mathbb{R}}|x|N(ds\times dx) < \infty.$$

La misma prueba que se siguió en el lema 4.1.9 nos permite probar que es un proceso de Lévy, diferente a los que nos hemos encontrado hasta ahora. Tenemos entonces un buen entendimiento de este primer caso.

Caso 2: La segunda posibilidad, es que $\int_{\mathbb{R}} 1\wedge|x|\Pi(dx) = \infty$ con $\Pi(B) = \infty$. En ese caso, solo tenemos que $\int_{\mathbb{R}} 1\wedge x^2\Pi(dx) < \infty$ al tratarse de una medida de Lévy y entonces la expresión (4.1.16) no está bien definida: la suma de los saltos no converge al no satisfacerse la segunda condición de 4.1.12. Esta dificultad se hará evidente a continuación. Un ejemplo de una medida que cae dentro de este caso es la medida α-estable $\Pi(dx) = \frac{1}{|x|^{1+\alpha}}\mathbb{1}_{\mathbb{R}\setminus 0}dx$. Es medida de Lévy, no es finita en ninguna vecindad del 0 y tampoco satisface $\int_B(1\wedge|x|)\Pi(dx) < \infty$.

Recordemos el problema inicial para hacer evidente la dificultad que se presenta: dado un proceso de Lévy X con terna (a, q^2, Π), buscamos construir tres procesos independientes X_1, X_2, X_3 tales que $X = X_1 + X_2 + X_3$. Por la fórmula de Lévy-Kintchine:

$$\Psi(\lambda) = \underbrace{ia\lambda + q^2\frac{\lambda^2}{2}}_{(1)} + \underbrace{\int_{\mathbb{R}}\left(1-e^{i\lambda x}\right)\mathbb{1}_{\{|x|\geq 1\}}\Pi(dx)}_{(2)} + \underbrace{\int_{\mathbb{R}}\left(1-e^{i\lambda x}+i\lambda x\right)\mathbb{1}_{\{|x|<1\}}\Pi(dx)}_{(3)}$$

Los términos deberían ahora resultar familiares: (1) es el exponente característico de un movimiento browniano con deriva $B_t - at$, mientras que (2), por el lema 4.1.9, es el exponente de $X_t^{(2)} = \int_{[0,t]}\int_{(-\infty,-1]\cup[1,\infty)} x\,N(ds\times dx)$ es un proceso Poisson compuesto con saltos de tamaño mayor que uno. Ingenuamente, podríamos descomponer (3) en

$$\int_{\mathbb{R}}\left(1-e^{i\lambda x}\right)\mathbb{1}_{\{|x|<1\}}\Pi(dx) + \int_{\mathbb{R}}(i\lambda x)\,\mathbb{1}_{\{|x|<1\}}\Pi(dx)$$

aplicar Campbell similarmente a (2) y llegar a la conclusión de que (3) es la resta de un proceso Poisson compuesto (o una superposición de estos) con saltos de magnitud menor que 1 y un término de deriva:

$$\int_{[0,t]}\int_{(-1,0)\cup(0,1)} x\,N(ds\times dx) - t\int_{\mathbb{R}}(\lambda x)\mathbb{1}_{\{|x|<1\}}\Pi(dx).$$

Este razonamiento sería correcto de cumplirse $\int_{[-1,1]}(1\wedge|x|)\Pi(dx) < \infty$ pero como ya se vio, esto no tiene por qué suceder y la primer integral de hecho no converge (no podemos aplicar Campbell (4.1.13)). Tendremos que proceder con más cuidado; para superar esta dificultad e identificar la naturaleza de este misterioso tercer término, será necesario hacer uso del espacio de las martingalas cuadrado integrables. En particular, usaremos que se trata de un espacio de Hilbert. En el Apéndice se puede leer la prueba de esta conocida propiedad.

4.2 La descomposición de Lévy-Itô

Referencia: Kyprianou [13] capítulo 2.

Lema 4.2.1 *Sea $B \in \mathscr{B}(\mathbb{R})$ tal que $\int_B |x|\Pi(dx) < \infty$ con Π medida de Lévy.*
i) El proceso Poisson compuesto compensado M definido como

$$M_t = \int_{[0,t]} \int_B x\, N(ds \times dx) - t \int_B x\, \Pi(dx), \qquad t \geq 0 \tag{4.2.1}$$

es una martingala con respecto a $(\mathscr{F}_t)_{t\geq 0}$ la filtración generada por X en (4.1.16).
ii) Además, si $\int_B x^2\Pi(dx) < \infty$ entonces es una martingala cuadrado integrable.

Nótese que por Campbell $M_t = X_t - \mathbb{E}[X_t]$ y que la condición $\int_B x^2\Pi(dx) < \infty$ se satisface para cualquier B acotado (de hecho independientemente de que sea vecindad del 0). Esto es consecuencia de que $\int 1 \wedge x^2\Pi(dx) < \infty$. Aquí es donde la definición de una medida de Lévy juega un papel fundamental, es lo que nos permitirá concluir que M es cuadrado integrable y que por lo tanto que está en $\mathcal{M}_T$.

Demostración
Como por el lema 4.1.9, X tiene incrementos independientes y estacionarios, se sigue que M también. Veamos que para $t > s$, $\mathbb{E}[M_t|\mathscr{F}_s] = M_s$:

$$\begin{aligned}
\mathbb{E}[M_t - M_s|\mathscr{F}_s] &= \mathbb{E}[M_t - M_s] \\
&= \mathbb{E}\left[\int_{(s,t]} \int_B x\, N(du \times dx)\right] - (t-s)\int_B x\, \Pi(dx) \\
&= 0
\end{aligned}$$

donde la última igualdad es consecuencia del teorema de Campbell, (3.2.4). Ahora, supongamos que además, $\int_B x^2\Pi(dx) < \infty$. Entonces, una vez más por el teorema de Campbell, (3.2.5) se tiene que:

$$\begin{aligned}
\mathbb{E}\left[\left(M_t + t\int_B x\, \Pi(dx)\right)^2\right] &= \mathbb{E}\left[\left(\int_{[0,t]}\int_B x\, N(ds \times dx)\right)^2\right] \\
&= t\int_B x^2\, \Pi(dx) + t^2\left(\int_B x\, \Pi(dx)\right)^2
\end{aligned}$$

Al desarrollar el lado izquierdo de la igualdad usando que $\mathbb{E}[M_t] = 0$ por ser martingala nula en 0 obtenemos que

$$\mathbb{E}\left[\left(M_t + t\int_B x\, \Pi(dx)\right)^2\right] = \mathbb{E}\left[M_t^2\right] + t^2\left(\int_B x\, \Pi(dx)\right)^2 \tag{4.2.2}$$

por lo que $\mathbb{E}\left[M_t^2\right] = t\int_B x^2\, \Pi(dx) < \infty$. ■

Teorema 4.2.2 *Supongamos que Π es medida de Lévy y para cada $\epsilon \in (0,1)$, sea $D_\epsilon = (-1,-\epsilon) \cup (\epsilon, 1)$. Definimos la martingala*

$$M_t^\epsilon = \int_{[0,t]} \int_{D_\epsilon} xN(ds \times dx) - t\int_{D_\epsilon} x\, \Pi(dx) \qquad t \geq 0. \tag{4.2.3}$$

Entonces, existe una martingala M con las siguientes propiedades: para cada $T > 0$, existe una sucesión determinista $\{\epsilon_k^T\}_{k\in\mathbb{N}}$ que decrece a 0 tal que

$$\lim_{n\to\infty} \sup_{0\le s\le T} \left(M_s^{\epsilon_n^T} - M_s\right)^2 = 0 \quad c.s. \tag{4.2.4}$$

y por lo tanto, $M^{\epsilon_n^T} \to M$ uniformemente en compactos $\mathbb{P}$-casi seguramente. M es adaptado a $(\mathscr{F}_t)_{t\ge 0}$, tiene $\mathbb{P}$-casi seguramente trayectorias cadlag con una cantidad a lo más numerable de discontinuidades en $[0,T]$ y sus incrementos son independientes y estacionarios. Es decir, es un proceso de Lévy.

Demostración

Primero, probemos (4.2.4). Sean η, $\epsilon \in (0,1)$ con $\epsilon > \eta$. Entonces,

$$\begin{aligned}
\|M^\eta - M^\epsilon\|_{\mathcal{M}_T}^2 &= \mathbb{E}\left[\left(M_T^\eta - M_T^\epsilon\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{[0,T]}\int_{\eta<|x|\le\epsilon} xN(ds\times dx) - T\int_{\eta<|x|\le\epsilon} x\,\Pi(dx)\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_{[0,T]}\int_{\eta<|x|\le\epsilon} xN(ds\times dx)\right)^2\right] + T^2\left(\int_{\eta<|x|\le\epsilon} x\,\Pi(dx)\right)^2 \\
&\qquad - 2T\mathbb{E}\left[\int_{[0,T]}\int_{\eta<|x|\le\epsilon} xN(ds\times dx)\right]\int_{\eta<|x|\le\epsilon} x\,\Pi(dx) \\
&= T\int_{\eta<|x|\le\epsilon} x^2\,\Pi(dx)
\end{aligned}$$

donde en la última igualdad se están usando los incisos (4.1.14) y (4.1.14) de la fórmula de Campbell. Nótese que $\int_{D_\epsilon} x^2\Pi(dx) < \infty$ al ser D_ϵ acotado. Como $\int_{(-1,1)} x^2\,\Pi(dx) < \infty$ por ser medida de Lévy, entonces por el teorema de convergencia dominada

$$\lim_{\epsilon,\eta\to 0} \|M^\epsilon - M^\eta\|_{\mathcal{M}_T}^2 = \lim_{\epsilon,\eta\to 0} T\int_{\eta<|x|\le\epsilon} x^2\,\Pi(dx) = 0$$

por lo que la sucesión $(M^\epsilon)_{\epsilon>0}$ es de Cauchy en $\mathcal{M}_T$. Por el teorema 9.3.12, $\mathcal{M}_T$ es completo y por lo tanto existe una martingala $M = (M_t)_{t\le T}$ en $\mathcal{M}_T$ adaptada a $(\mathscr{F}_t)_{t\le T}$ con trayectorias continuas por la derecha $\mathbb{P}$-casi seguramente tal que $\|M^\epsilon - M\|_{\mathcal{M}_T} \to 0$. Esto se hizo para un T arbitrario; veamos que este límite no depende de T: es decir, si tomamos $0 < T' < T$, las trayectorias de las martingalas límite coinciden en $[0,T']$ $\mathbb{P}$-casi seguramente.
Por la desigualdad de martingalas de Doob 9.3.9, tomando límite cuando $\epsilon \to 0$

$$\lim_{\epsilon\to 0}\mathbb{E}\left[\sup_{0\le s\le T}\left(M_s^\epsilon - M_s\right)^2\right] \le 4\lim_{\epsilon\to 0}\|M^\epsilon - M\|_{\mathcal{M}_T} = 0. \tag{4.2.5}$$

Ajustamos la notación: sean $0 < T' < T$ y

$$M_{\cdot,T'} = \{M_{s,T'} : s\in[0,T']\} \qquad M_{\cdot,T} = \{M_{s,T} : s\in[0,T]\}$$

las martingalas límite de M^ϵ en $\mathcal{M}_{T'}$ y $\mathcal{M}_T$ respectivamente, obtenidas a partir del procedimiento que se siguió previamente (se sobre entiende que M^ϵ se está restringiendo a $[0,T']$ en

$\mathcal{M}_{T'}$). Queremos probar que, salvo por un conjunto $\mathbb{P}$-nulo, $M_{s,T'}(\omega) = M_{s,T}(\omega)$ para toda $s \in [0, T']$. Tenemos que

$$\lim_{\epsilon\to 0} \left\| M^\epsilon - M_{\cdot,T'} \right\|_{\mathcal{M}_{T'}} = 0 \quad \text{y} \quad \lim_{\epsilon\to 0} \| M^\epsilon - M_{\cdot,T} \|_{\mathcal{M}_T} = 0 \tag{4.2.6}$$

pero además, por (4.2.5):

$$\lim_{\epsilon\to 0} \mathbb{E}\left[\sup_{0\le s\le T} (M_s^\epsilon - M_{s,T})^2 \right] \le 4 \lim_{\epsilon\to 0} \| M^\epsilon - M_{\cdot,T} \|_{\mathcal{M}_T} = 0$$

así que en particular, para $s = T'$

$$0 = \lim_{\epsilon\to 0} \mathbb{E}\left[\left(M_{T'}^\epsilon - M_{T',T} \right)^2 \right] = \lim_{\epsilon\to 0} \| M^\epsilon - M_{\cdot,T} \|_{\mathcal{M}_{T'}} \tag{4.2.7}$$

donde una vez más, se sobre entiende que se esta considerando la restricción al intervalo $[0, T']$ de las martingalas M^ϵ, $M_{\cdot,T}$. Por lo tanto, se sigue de (4.2.6), (4.2.7) y la desigualdad del triángulo que:

$$\left\| M_{\cdot,T'} - M_{\cdot,T} \right\|_{\mathcal{M}_{T'}} \le \lim_{\epsilon\to\infty} \left\| M_{\cdot,T'} - M^\epsilon \right\|_{\mathcal{M}_{T'}} + \lim_{\epsilon\to\infty} \| M^\epsilon - M_{\cdot,T} \|_{\mathcal{M}_{T'}} = 0.$$

pues el lado izquierdo de la desigualdad no depende de ϵ. Por la observación que se hizo previo al teorema 9.3.11, la martingala $M_{\cdot,T'}$ es indistinguible de la restricción de $M_{\cdot,T}$ al intervalo $[0, T']$. Como T y T' se pueden tomar arbitrariamente grandes, podemos hablar de la martingala límite $M = (M_t)_{t\ge 0}$ sin restringirnos a un horizonte finito.

Por (4.2.5) y usando que $\sup_n a_n^2 = (\sup_n |a_n|)^2$, se tiene que $\sup_{0\le s\le T} |M_s^\epsilon - M_s| \overset{L^2(\Omega)}{\to} 0$ y por lo tanto existe una subsucesión $\{\epsilon_n^T\}_n$ tal que

$$\lim_{n\to\infty} \sup_{0\le s\le T} |M_s^{\epsilon_n^T} - M_s|(\omega) = 0 \tag{4.2.8}$$

puntualmente, $\mathbb{P}$-c.s. Así, salvo en un conjunto $\mathbb{P}$-nulo $M^{\epsilon_n^T}(\omega)$ converge uniformemente a $M(\omega)$ para $s \in [0, T]$, por lo que se sigue (4.2.4). Las trayectorias de M tienen límite por la izquierda pues las trayectorias de $M^{\epsilon_n^T}$ son cadlag para cada n y la convergencia es uniforme $\mathbb{P}$- c.s. Una función cadlag tiene a lo más una cantidad numerable de discontinuidades por lo que las trayectorias de M tienen a lo más una cantidad numerable de discontinuidades $\mathbb{P}$-c.s. Para mayor detalle de estas afirmaciones, se puede consultar Kyprianou [13] página 64. Otra posibilidad habría sido apelar a la versión cadlag de M dada por el teorema 9.3.6 pero optamos por este camino ya que el resultado no se probó.

Para ver que los incrementos son independientes y estacionaros, sean $0 \le s_1 \le t_1 \le \cdots \le s_n \le t_n \le T < \infty$ y $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Por (4.2.8), para cualquier $t \in [0, T]$ se tiene que $\lim_{n\to\infty} |M_t^{\epsilon_n^T} - M_t|(\omega) = 0$ $\mathbb{P}$- casi seguramente. Por lo tanto, por el teorema de convergencia dominada:

$$\begin{aligned}
\mathbb{E}\left[\prod_{j=1}^{n} e^{i\lambda_j (M_{t_j}-M_{s_j})}\right] &= \lim_{n\to\infty} \mathbb{E}\left[\prod_{j=1}^{n} e^{i\lambda_j (M_{t_j}^{\epsilon_n^T}-M_{s_j}^{\epsilon_n^T})}\right] \\
&= \lim_{n\to\infty} \prod_{j=1}^{n} \mathbb{E}\left[e^{i\lambda_j (M_{t_j}^{\epsilon_n^T}-M_{s_j}^{\epsilon_n^T})}\right] \\
&= \lim_{n\to\infty} \prod_{j=1}^{n} \mathbb{E}\left[e^{i\lambda_j (M_{t_j-s_j}^{\epsilon_n^T})}\right] \\
&= \prod_{j=1}^{n} \mathbb{E}\left[e^{i\lambda_j (M_{t_j-s_j})}\right] \qquad (4.2.9)
\end{aligned}$$

Al evaluar en $(\lambda_1, \cdots, \lambda_n) = (1, 0, \cdots, 0)$ obtenemos que los incrementos son estacionarios y al usar esto en (4.2.9) se sigue que son independientes. En la tercera y cuarta igualdad se está usando que para cada ϵ_n^T, el proceso $M^{\epsilon_n^T}$ (que es el mismo que el que se definió en Lema 4.2.1) es Poisson compuesto y por lo tanto tiene incrementos independientes y estacionarios. ■

Teorema 4.2.3 Descomposición de Lévy-Itô
Dada una tripleta (a, q^2, Π) con Π medida de Lévy, existen tres procesos X_1, X_2 y X_3 tales que el proceso $X = X_1 + X_2 + X_3$ es un proceso de Lévy con exponente característico

$$\Psi(\lambda) = \underbrace{ia\lambda + q^2\frac{\lambda^2}{2}}_{(1)} + \underbrace{\int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \mathbb{1}_{\{|x|\geq 1\}}\Pi(dx)}_{(2)} + \underbrace{\int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right) \mathbb{1}_{\{|x|<1\}}\Pi(dx)}_{(3)}$$

Además el proceso

$$X_t^{(1)} = B_t - at$$

es un movimiento browniano con deriva y tiene por exponente característico (1),

$$X_t^{(2)} = \int_{[0,t]} \int_{(-\infty,-1]\cup[1,\infty)} x\, N(ds \times dx) \qquad (4.2.10)$$

de exponente (2) es un proceso Poisson compuesto con magnitud de saltos mayor o igual que uno y X_3 es una martingala cadlag cuadrado integrable con saltos de magnitud menor que uno y con incrementos independientes y estacionarios (por lo que es proceso de Lévy) cuyo exponente está dado por (3).

Demostración
Sea N una medida de Poisson en $([0,\infty) \times \mathbb{R}, \mathscr{B}([0,\infty)) \otimes \mathscr{B}(\mathbb{R}), dt \times \Pi(dx))$ definida en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ y $X_1 = (B_t - at)_{t\geq 0}$ un browniano con deriva definido en otro espacio de probabilidad $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Está claro que el exponente Ψ_1, Ψ_2 de X_1, X_2 está dado respectivamente por (1) y (2) por la discusión que se tuvo previo a la sección 2.3. Así, en el espacio producto, X_1 y X_2 (definido como en (4.2.10)) son independientes y $X_1 + X_2$ tiene por exponente $\Psi_1 + \Psi_2$. Ahora, construimos al proceso X_3: para cada $\epsilon \in (0, 1)$, consideramos a la martingala

$$M_t^{\epsilon} = \int_{[0,t]} \int_{\epsilon\leq|x|<1} xN(ds \times dx) - t \int_{\epsilon\leq|x|<1} x\, \Pi(dx) \qquad t \geq 0.$$

Por Campbell-(4.1.13), su exponente característico es

$$\Psi_{\epsilon,3}(\lambda) = \int_{\epsilon\leq|x|<1} \left(1 - e^{i\lambda x} + i\lambda x\right) \Pi(dx)$$

y es independiente de X_2 pues estamos integrando con respecto a la medida de Poisson en conjuntos ajenos (Corolario 3.2.3). Por el teorema 4.2.2, existe una martingala cuadrado integrable M (que además es proceso de Lévy) y una sucesión $\{\epsilon_n^T\}_{n\in\mathbb{N}}$ a la cual $M^{\epsilon_n^T}$ converge uniformemente en $[0,T]$ $\mathbb{P}$-c.s. Como en particular $M_1^{\epsilon_n^T}(\omega) \to M_1(\omega)$ $\mathbb{P}$-c.s., por el teorema de Lévy y el teorema de convergencia dominada, el exponente ψ_3 de M es

$$\Psi_3(\lambda) = \int_{|x|<1} \left(1 - e^{i\lambda x} + i\lambda x\right) \Pi(dx).$$

En esta afirmación se está usando que $\int_{\mathbb{R}}(1\wedge x^2)\Pi(dx) < \infty$ y que por el lema 2.1.10, $|1 - e^{i\lambda x} + i\lambda x| \leq |\lambda x|^2 |e^{i\lambda x}| \leq |\lambda x|^2$. Esto permite dominar por $x^2 \mathbb{1}_{|x|<1}$ pues $\lambda^2 \int_{|x|<1} x^2 \Pi(x) < \infty$. Que X_3 sea independiente de X_2 se sigue de que para cada n, $X_3^{\epsilon_n^T}$ es independiente de X_2 y se aplica un procedimiento similar al que se utilizó para probar (4.2.9). Concluimos que el proceso X definido como

$$X_t = X_t^{(1)} + X_t^{(2)} + X_t^{(3)}$$

tiene por exponente característico

$$\begin{aligned}\Psi(\lambda) &= \Psi_1(\lambda) + \Psi_2(\lambda) + \Psi_3(\lambda)\\ &= ia\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \mathbb{1}_{\{|x|\geq 1\}}\Pi(dx) + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right) \mathbb{1}_{\{|x|<1\}}\Pi(dx)\end{aligned}$$

■

Obsérvese que X_3 se definió gracias a que el espacio $\mathcal{M}_T$ es de Hilbert, apelamos a su completitud y lo definimos como el límite de una sucesión de Cauchy de procesos Poisson compuestos compensados $X^{\epsilon_n^T}$ con saltos cada vez más pequeños (con magnitud de saltos $|\Delta X^{\epsilon_n^t}| \geq \epsilon_n^T$). El problema es que al apelar a un resultado de existencia, perdimos la intuición sobre el comportamiento del proceso límite X_3 que era justamente uno de los aspectos que buscábamos esclarecer. Nótese que al pasar a una subsucesión el límite se puede pensar puntual, por lo que, al menos intuitivamente, X_3 debería ser un proceso de "tipo Poisson compensado" con saltos arbitrariamente pequeños. Veamos que esta intuición está bien justificada al argumentar como en (4.1.21). Salvo por un conjunto $\mathbb{P}$-nulo, para cada $t \in [0,T]$ y definiendo $\epsilon_0^T = 1$ se tiene (al considerar una subsucesión adecuada de ser necesario) que

$$\begin{aligned}X_t^{(3)} = \lim_{n\to\infty} X_t^{\epsilon_n^T} &= \lim_{n\to\infty} \int_{[0,t]} \int_{\epsilon_n^T\leq|x|} x\, N(ds\times dx) - t\int_{\epsilon_n^T\leq|x|} x\,\Pi(dx)\\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \left\{ \int_{[0,t]} \int_{\epsilon_k^T\leq|x|<\epsilon_{k-1}^T} x\, N(ds\times dx) - t\int_{\epsilon_k^T\leq|x|<\epsilon_{k-1}^T} x\,\Pi(dx) \right\}\\ &= \sum_{k=1}^{\infty} \left\{ \int_{[0,t]} \int_{\epsilon_k^T\leq|x|<\epsilon_{k-1}^T} x\, N(ds\times dx) - t\int_{\epsilon_k^T\leq|x|<\epsilon_{k-1}^T} x\,\Pi(dx) \right\} \qquad (4.2.11)\end{aligned}$$

así que en efecto, X_3 es superposición de procesos Poisson compuestos independientes y compensados con magnitud de saltos en $[\epsilon_k^T, \epsilon_{k-1}^T)$ para cada k y entonces la intuición que dimos está bien justificada. La existencia de $X_t^{(3)}$ como límite puntual para cada t nos permite pasar al límite en la última igualdad. En principio, (4.2.11) no habría tenido por qué converger ya que cada término de la forma $\int_{\epsilon_k^T \leq |x| < \epsilon_{k-1}^T} x\, N(ds \times dx)$ es una suma finita de términos tanto positivos como negativos con valores en $(\epsilon_{k-1}^T, \epsilon_k^T] \cup [\epsilon_k^T, \epsilon_{k-1}^T)$.

Cabe notar que si Y es un proceso de Lévy con exponente Ψ, entonces por la descomposición de Lévy-Itô sabemos que existen X_1, X_2, X_3 procesos en un espacio de probabilidad tales que $X = X_1 + X_2 + X_3$ es proceso de Lévy con exponente Ψ y por lo tanto X y Y tienen las mismas distribuciones finito dimensionales. Así, X es solo una versión de Y. Hay que ser cuidadosos al probar propiedades relativas a X y quererlas aplicar al proceso original Y. Por ejemplo las propiedades trayectoriales que podamos probar para X no tienen por qué cumplirse para Y ya que la igualdad en ley podría no ser suficiente.

4.3 Variación total y aplicaciones

Si X es un proceso de Lévy con terna (a, q^2, Π), la medida Π debe satisfacer $\int_{\mathbb{R}} 1 \wedge x^2 \Pi(dx) < \infty$. Sin embargo, si además cumple que $\int_{\mathbb{R}} 1 \wedge |x|\, \Pi(dx) < \infty$, la descomposición de Itô es mucho más sencilla: por la fórmula de Lévy Kintchine

$$\Psi(\lambda) = ia\lambda + q^2 \frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x \mathbb{1}_{\{|x|<1\}}\right) \Pi(dx) \tag{4.3.1}$$

La condición $\int_{\mathbb{R}} 1 \wedge |x| \Pi(dx)$ implica que $\int_{\mathbb{R}} x \mathbb{1}_{\{|x|<1\}} \Pi(dx) < \infty$ (a diferencia del caso general) por lo que si definimos $\delta = a + \int_{\mathbb{R}} x \mathbb{1}_{\{|x|<1\}} \Pi(dx)$ podemos reordenar (4.3.1) y obtenemos

$$\Psi(\lambda) = i\delta\lambda + q^2 \frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \Pi(dx). \tag{4.3.2}$$

El primer término es el exponente de un browniano con deriva mientras que el segundo, por Campbell y recordando la discusión que se tuvo en (4.1.21), proviene de

$$X_t^{(2)} = \int_{[0,t]} \int_{\mathbb{R}} x\, N(ds \times dx) \qquad t \geq 0 \tag{4.3.3}$$

pues $\int_{\mathbb{R}} 1 \wedge |x| \Pi(dx) < \infty$; en este caso, $X_3 = 0$. Además, es un proceso Poisson compuesto siempre y cuando $\Pi(\mathbb{R}) < \infty$. Por lo tanto X es suma de un browniano con deriva y un proceso Poisson compuesto si y solo si tiene exponente de la forma (4.3.2) y $\Pi(\mathbb{R}) < \infty$. En el caso $\Pi(\mathbb{R}) = \infty$, se trata de una superposición de procesos Poisson compuestos, habrá una cantidad infinita de saltos Δ_t en $[0, T]$ pero la suma de estos $\sum_{s \leq T} |\Delta_s| = \int_{[0,T]} \int_{\mathbb{R}} |x|\, N(ds \times dx) < \infty$ una vez más por Campbell-(4.1.12).

Definición 4.3.1 *Decimos que un proceso X es de variación acotada si las trayectorias $X_t(\omega)$ $t \in [0, T]$ son de variación acotada $\mathbb{P}$-c.s. para todo $T \geq 0$.*

Nótese que, en general, si dos procesos X, Y tienen la mismas distribuciones finito dimensionales, no necesariamente sus trayectorias satisfacen las mismas propiedades. Por ejemplo, cuando se estudió el espacio $\mathcal{M}_T$, para toda martingala se construyó una versión con trayectorias continuas por la derecha. Por esto, si Y es un procesos de Lévy y X es el proceso construido a partir de la descomposición de Lévy-Itô, hay que asegurarse que las propiedades que se prueben para las trayectorias de X sigan siendo validas para las de Y.

Proposición 4.3.2 *Si X y Y son procesos de Lévy con la misma terna (a, q^2, Π), entonces X es de variación acotada si y solo si Y lo es.*

Demostración
Supongamos que X no es de variación acotada. Por (4.1.2), para cualquier familia finita de índices en $[0, T]$, $0 \leq t_1 < \cdots < t_n = T$, los vectores $W = \left(X_{t_1}, X_{t_2} - X_{t_1} \ldots, X_{t_n} - X_{t_{n-1}}\right)$ y $Z = \left(Y_{t_1}, Y_{t_2} - Y_{t_1} \ldots, Y_{t_n} - Y_{t_{n-1}}\right)$ tienen la misma ley. Sea $\{I_j\}_{j \in \mathbb{N}}$ familia creciente de particiones, I_j de cardinalidad finita para cada j tal que $\cup_j I_j = ([0, T] \cap \mathbb{Q}) \cup \{T\}$. Por lo tanto, para cada k se satisface

$$\mathbb{P}\left(\sum_{t_i \in I_k} |X_{t_i} - X_{t_{i-1}}| > M\right) = \tilde{\mathbb{P}}\left(\sum_{t_i \in I_k} |Y_{t_i} - Y_{t_{i-1}}| > M\right) \tag{4.3.4}$$

Como I_{k+1} es refinamiento de I_k, entonces

$$\left\{\sum_{t_i \in I_k} |X_{t_i} - X_{t_{i-1}}| > M\right\} \subset \left\{\sum_{t_i \in I_{k+1}} |X_{t_i} - X_{t_{i-1}}| > M\right\}$$

y tomando límite cuando $k \to \infty$ en (4.3.4) se tiene que

$$\mathbb{P}\left(\sup_{I_k}\left\{\sum_{t_i \in I_k} |X_{t_i} - X_{t_{i-1}}|\right\} > M\right) = \tilde{\mathbb{P}}\left(\sup_{I_k}\left\{\sum_{t_i \in I_k} |Y_{t_i} - Y_{t_{i-1}}|\right\} > M\right). \tag{4.3.5}$$

Como X no es de variación acotada, existe $A \in \mathscr{F}$ con $\mathbb{P}(A) > 0$ tal que

$$\sup_{I_k} \sum_{t_i \in I_k} |X_{t_i} - X_{t_{i-1}}|(\omega) = \infty \qquad \text{para todo } \omega \in A.$$

Basta con restringirse a las particiones I_k pues X tienen trayectorias cadlag y $\mathbb{Q}$ es un denso. Como (4.3.5) es mayor que $\mathbb{P}(A)$ para cualquier M, se tiene el resultado tomando una sucesión $M_j \uparrow \infty$ y el conjunto

$$\tilde{A} = \bigcap_j \left\{\sup_{I_k}\left\{\sum_{t_i \in I_k} |Y_{t_i} - Y_{t_{i-1}}|\right\} > M_j\right\} = \left\{\sup_{I_k}\left\{\sum_{t_i \in I_k} |Y_{t_i} - Y_{t_{i-1}}|\right\} = \infty\right\}$$

pues $0 < \mathbb{P}(A) \leq \tilde{\mathbb{P}}(\tilde{A})$. ■

¿Cuándo un proceso de Lévy Y con terna (a, q^2, Π) es de variación acotada? Por el resultado anterior, basta establecer cuando el proceso obtenido a partir de la descomposición de Lévy-Itô $X = X_1 + X_2 + X_3$ es de variación acotada. Como el primer término es un movimiento browniano que no es de variación acotada, una condición necesaria es que q sea 0. El segundo término X_2, siempre sera de variación acotada pues se trata de un proceso Poisson compuesto, así que la pregunta se reduce a estudiar cuando X_3 es de variación acotada.

Si X es un proceso con límites por la izquierda, denotamos por $\Delta X_t = X_t - X_{t-}$ a su proceso de saltos.

Lema 4.3.3 *Si X es un proceso de Lévy, entonces $\Delta X_t = 0$ c.s. o dicho de otra forma:*

$$\mathbb{P}\left(X_t = X_{t-}\right) = 1 \qquad \textit{para todo } t \geq 0.$$

Demostración
Denotemos por Ψ al exponente de X. Por existencia de los límites por la izquierda y estacionariedad de los incrementos,

$$\varphi_{\Delta X_t}(\lambda) = \lim_{s\uparrow t} \mathbb{E}\left[e^{i\lambda(X_t - X_s)}\right] = \lim_{s\uparrow t} \varphi_{X_{t-s}}(\lambda) = \lim_{s\uparrow t} e^{(t-s)\Psi(\lambda)} = e^0$$

por lo que ΔX_t tienen la misma función característica que la constante 0 y por lo tanto $\Delta X_t = 0$ c.s. ■

Otra forma de decir esto es:

Corolario 4.3.4 *Si X es un proceso de Lévy, entonces para cada t*

$$\lim_{s\uparrow t} X_s = X_t \quad c.s.$$

a pesar de no ser continuo por la izquierda, lo cual es sorprendente. Este hecho es un caso particular de una propiedad más fuerte, la cual probaremos en el capítulo 4 en 6.3.2, conocida como casi-continuidad por la izquierda. Sin embargo no requeriremos de tanto por ahora y la prueba de este caso particular es significativamente más simple.

Proposición 4.3.5 *Sean C^1 y C^2 dos procesos Poisson compuestos independientes con respecto a la misma filtración $(\mathscr{F}_t)$. Entonces,*

$$\sum_{s\geq 0} |\Delta C_s^1| \cdot |\Delta C_s^2| = 0 \qquad c.s.$$

es decir, casi seguramente, los dos procesos no saltan al mismo tiempo para ninguna t.

Demostración
Sean (T_n) los tiempos de saltos sucesivos de C^1 y nótese que

$$\sum_{s\geq 0} |\Delta C_s^1| \cdot |\Delta C_s^2| = \sum_{n\geq 0} |\Delta C_{T_n}^1| \cdot |\Delta C_{T_n}^2|$$

pues los otros términos son 0. Bastaría entonces probar que $\Delta C_{T_n}^2 = 0$. Como T_n es $\sigma(C_s^1,\, s \geq 0)$-medible, por independencia:

$$\mathbb{E}\left[\Delta C_{T_n}^2\right] = \mathbb{E}\left[\mathbb{E}\left[\Delta C_{T_n}^2 | T_n\right]\right] = \int_{R^+} \mathbb{E}\left[\Delta C_t^2 | T_n = t\right] \mathbb{P}_{T_n}(dt) = \mathbb{E}\left[\Delta C_t^2\right] = 0$$

pues por el lema 4.3.3 $C_t - C_{t-} = 0$ c.s. con lo que concluimos. ■

Lema 4.3.6 *Un proceso de Lévy X con terna característica (a, q^2, Π) es de variación acotada si y solo si $q = 0$ y $\int_{\mathbb{R}}(1 \wedge |x|)\Pi(dx) < \infty$.*

Demostración
Si $\int_{\mathbb{R}}(1 \wedge |x|)\Pi(dx) < \infty$, recordando la discusión que se tuvo al inicio de la sección, el exponente de X se reescribe como

$$\Psi(\lambda) = i\delta\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \Pi(dx). \tag{4.3.6}$$

y por lo tanto es suma de un browniano con deriva X_1 y una superposición de procesos Poisson compuestos

$$X_t^{(2)} = \int_{[0,t]} \int_{\mathbb{R}} x\, N(ds \times dx) \qquad t \geq 0.$$

Como además $q = 0$, X_1 se reduce al término determinista de deriva $X_1 = -\delta t$ y por lo tanto es de variación acotada. El proceso X_2 es un proceso de saltos puro continuo por la derecha y entonces su variación en $[0, t]$ está dada por la suma de la magnitud de estos $\sum_{s \leq t} |\Delta_s X_s^{(2)}|$. Por Campbell, la condición $\int 1 \wedge |x| \Pi(dx) < \infty$ es necesaria y suficiente para que esta cantidad sea finita:

$$\sum_{s \leq t} |\Delta_s X_s^{(2)}| = \int_{[0,t]} \int_{\mathbb{R}} |x|\, N(ds \times dx) < \infty.$$

Ahora, supongamos que X es de variación acotada y fijamos $t \in [0, T]$ arbitrario. Está claro que entonces q debe ser cero al no ser el browniano de variación acotada en ningún intervalo. Entonces el exponente de X se ve como

$$\Psi(\lambda) = \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \mathbb{1}_{\{|x| \geq 1\}} \Pi(dx) + \int_{\mathbb{R}} \left(1 - e^{i\lambda x} + i\lambda x\right) \mathbb{1}_{\{|x| < 1\}} \Pi(dx)$$

Omitimos el término de deriva al no afectar el resultado. Por la discusión que se tuvo en (4.2.11) y la descomposición de Lévy Itô, este es exponente de

$$X_t = \int_{[0,t]} \int_{x \in (-1,1)^c} x\, N(ds \times dx) \\ + \lim_{n \to \infty} \sum_{k=1}^{n} \left\{ \int_{[0,t]} \int_{\epsilon_k^T \leq |x| < \epsilon_{k-1}^T} x\, N(ds \times dx) - t \int_{\epsilon_k^T \leq |x| < \epsilon_{k-1}^T} x\, \Pi(dx) \right\}$$

al considerar una subsucesión adecuada. Obsérvese que las integrales con respecto a $N(ds \times dx)$ que aparecen en esta expresión son sobre conjuntos ajenos de $\mathbb{R}^+ \times \mathbb{R}$ y por tanto son procesos Poisson compuestos independientes. Denotemos por K^k a la constante $\int_{\epsilon_k^T \leq |x| < \epsilon_{k-1}^T} x\, \Pi(dx)$ y por C^k a los procesos Poisson compuestos en la expresión anterior. Reescribimos a X como

$$X_t = C_t^0 + \lim_{n \to \infty} \sum_{k=1}^{n} \left\{ C_t^k - tK^k \right\}$$

Se sigue de la proposición 4.3.5 que los saltos de los procesos Poisson ΔC_t^k ocurren en tiempos distintos. Entonces si para algún t, $\Delta C_t^k = r > 0$ entonces $\Delta X_t = r$. Nótese que de no haber probado que los saltos no pueden ocurrir al mismo tiempo esto no tendría por qué ser cierto: podría compensarse un salto negativo con uno positivo al ocurrir en el mismo instante. Como la variación total de X en $[0, t]$ que denotaremos por $BV(X)_t$ es mayor que la suma de sus saltos $\sum_{s \leq t} \Delta X_t$,

$$BV(X)_t \geq \sum_{s \leq t} \Delta X_t \geq \sum_{s \leq t} \Delta C_s^0 + \sum_{s \leq t} \left(\sum_{k=1}^{n} \Delta C_s^k \right) = \int_{[0,t] \times \mathbb{R} \setminus (-\epsilon_n^T, \epsilon_n^T)} |x| N(ds \times dx).$$

Al tomar el límite $n \to \infty$ obtenemos

$$BV(X)_t \geq \int_{[0,t] \times \mathbb{R}} |x| N(ds \times dx) \tag{4.3.7}$$

que es finito al ser X de variación acotada. Por Campbell (4.1.12), se sigue que

$$\int_{\mathbb{R}} 1 \wedge |x|\Pi(dx) < \infty.$$

■

Una consecuencia directa de este resultado y (4.3.2) es lo siguiente:

Corolario 4.3.7 *Si un proceso de Lévy X con terna (a, q^2, Π) y exponente Ψ es de variación acotada, entonces*

$$\Psi(\lambda) = i\delta\lambda + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \Pi(dx) \tag{4.3.8}$$

donde $\delta = a + \int_{\mathbb{R}} x\mathbb{1}_{\{|x|<1\}}\Pi(dx)$. Además, $\Pi(\mathbb{R}) < \infty$ si y solo si X es un proceso Poisson compuesto con deriva.

En este caso, la descomposición de X queda completamente determinada por la pareja (δ, Π), por lo que al trabajar en el caso $\int_{\mathbb{R}}(1 \wedge |x|)\Pi(dx) < \infty$ se habla de pareja característica en lugar de terna.

Proposición 4.3.8 *Sea X un proceso de Lévy con terna característica (a, q^2, Π) y exponente Ψ. Entonces:*

1.
$$\lim_{|\lambda|\to\infty} \frac{\Psi(\lambda)}{\lambda^2} = \frac{q^2}{2}.$$

2. *Si X es de variación acotada y tiene por pareja característica (δ, Π),*

$$\lim_{|\lambda|\to\infty} \frac{\Psi(\lambda)}{\lambda} = -i\delta.$$

3. *X es un proceso Poisson compuesto si y solo si Ψ es acotada.*

Demostración
Se usarán las siguientes desigualdades:

$$|1 - \cos(a)| \leq 2(1 \wedge a^2) \qquad |1 - \cos(a)| \leq 2(1 \wedge |a|)$$

$$|a - \sin(a)| \leq 2(|a| \wedge a^2) \qquad |\sin(a)| \leq 2(1 \wedge |a|).$$

1. Por la fórmula de Lévy-Kintchine,

$$\frac{\Psi(\lambda^2)}{\lambda^2} = \frac{ia}{\lambda} + \frac{q^2}{2} + \int_{\mathbb{R}} \frac{\left(1 - e^{i\lambda x} + i\lambda x\mathbb{1}_{\{|x|<1\}}\right)}{\lambda^2}\Pi(dx) \tag{4.3.9}$$

y como $\lim_{|\lambda|\to\infty} \frac{\left(1-e^{i\lambda x}+i\lambda x\mathbb{1}_{\{|x|<1\}}\right)}{\lambda^2} = 0$, bastaría argumentar porqué podemos meter el límite dentro de la integral para concluir la prueba. Probaremos que podemos dominar este término por una función integrable; para $|\lambda|$ suficientemente grande,

$$\begin{aligned}
\left|\frac{e^{i\lambda x}-1-i\lambda x\mathbb{1}_{\{|x|\le 1\}}}{\lambda^2}\right| &= \left|\frac{\cos(\lambda x)+i\sin(\lambda x)-1-i\lambda x\mathbb{1}_{\{|x|\le 1\}}}{\lambda^2}\right| \\
&\le \frac{|1-\cos(\lambda x)|}{\lambda^2}+\frac{|(\sin(\lambda x)-\lambda x)\mathbb{1}_{\{|x|\le 1\}}+\sin(\lambda x)\mathbb{1}_{\{|x|>1\}}|}{\lambda^2} \\
&\le \frac{|1-\cos(\lambda x)|}{\lambda^2}+\frac{|\sin(\lambda x)-\lambda x|}{\lambda^2}\mathbb{1}_{\{|x|\le 1\}}+\frac{|\sin(\lambda x)|}{\lambda^2}\mathbb{1}_{\{|x|>1\}} \\
&\le 2\frac{1\wedge(\lambda x)^2}{\lambda^2}+2\frac{|\lambda x|\wedge(\lambda x)^2}{\lambda^2}\mathbb{1}_{\{|x|\le 1\}}+2\frac{1\wedge|\lambda x|}{\lambda^2}\mathbb{1}_{\{|x|>1\}} \\
&\le 2(1\wedge x^2)+2(1\wedge x^2)\mathbb{1}_{\{|x|\le 1\}}+2(1\wedge x^2)\mathbb{1}_{\{|x|>1\}} \\
&\le 4(1\wedge x^2)
\end{aligned}$$

y entonces,

$$\int_{\mathbb{R}}\frac{\left(1-e^{i\lambda x}+i\lambda x\mathbb{1}_{\{|x|<1\}}\right)}{\lambda^2}\Pi(dx)\le 4\int_{\mathbb{R}}(1\wedge x^2)\Pi(dx)<\infty$$

por ser medida de Lévy. Así, tomando el limite $|\lambda|\to\infty$ en (4.3.9) y aplicando el teorema de convergencia dominada, tenemos 1.

2. Al ser X de variación acotada, por el corolario 4.3.7:

$$\frac{\Psi(\lambda)}{\lambda}=i\delta+\int_{\mathbb{R}}\frac{\left(1-e^{i\lambda x}\right)}{\lambda}\Pi(dx) \tag{4.3.10}$$

y similarmente a lo que se hizo en 1, como $\lim_{|\lambda|\to\infty}\frac{\left(1-e^{i\lambda x}\right)}{\lambda}=0$, bastaría con justificar porqué se puede meter el límite dentro de la integral; para $|\lambda|$ suficientemente grande,

$$\begin{aligned}
\frac{|1-e^{i\lambda x}|}{|\lambda|} &\le \frac{|1-\cos(\lambda x)|}{|\lambda|}+\frac{|\sin(\lambda x)|}{|\lambda|} \\
&\le 2\frac{1\wedge|\lambda x|}{|\lambda|}+2\frac{1\wedge|\lambda x|}{|\lambda|} \\
&\le 4(1\wedge|x|).
\end{aligned}$$

Como X es de variación acotada, por el lema 4.3.6 $\int_{\mathbb{R}}(1\wedge|x|)\Pi(dx)<\infty$ y por lo tanto tomando el límite cuando $|\lambda|\to\infty$ y aplicando el teorema de convergencia dominada en (4.3.10) se concluye.

3. Si X es un proceso Poisson compuesto (sin deriva), por el corolario 4.3.7, $\Pi(\mathbb{R})<\infty$ y

$$\Psi(\lambda)=\int_{\mathbb{R}}\left(1-e^{i\lambda x}\right)\Pi(dx).$$

Se sigue que $|\Psi|\le 2\Pi(\mathbb{R})<\infty$.

Ahora, si Ψ es acotada, entonces $\lim_{|\lambda|\to\infty}\frac{\Psi(\lambda)}{\lambda^2}=0$ y por la proposición 4.3.8-2 se tiene que $q^2=0$. Por lo tanto,

$$\Psi(\lambda)=ia\lambda+\int_{\mathbb{R}}\left(1-e^{i\lambda x}+i\lambda x\mathbb{1}_{\{|x|<1\}}\right)\Pi(dx)$$

y

$$\mathrm{Re}\left(\Psi(\lambda)\right)=\int_{\mathbb{R}}\left(1-\cos(\lambda x)\right)\Pi(dx).$$

Como para cada $t > 0$ y $x \in \mathbb{R}$,

$$1 - e^{-\frac{t^2x^2}{2}} = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} (1 - \cos(ux)) e^{-\frac{u^2}{2t}} du \tag{4.3.11}$$

entonces, como el integrando es positivo, por Fubini-Tonelli:

$$\begin{aligned} \int_{\mathbb{R}} \left(1 - e^{-\frac{t^2x^2}{2}}\right) \Pi(dx) &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-\frac{u^2}{2t}} \int_{\mathbb{R}} (1 - \cos(ux)) \Pi(dx)\, du \\ &= \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} e^{-\frac{u^2}{2t}} \operatorname{Re}(\Psi(u))\, du \\ &\leq \sup\{\operatorname{Re}(\Psi(u))\,;\, u \in \mathbb{R}\} \end{aligned} \tag{4.3.12}$$

que es finito por hipótesis, por lo que $\int_{\mathbb{R}} \left(1 - e^{\frac{-t^2x^2}{2}}\right) \Pi(dx)$ es acotada para todo $t \in \mathbb{R}$. Esto nos permite usar el teorema de convergencia dominada y concluir que

$$\Pi(\mathbb{R}) = \lim_{t \to \infty} \int_{\mathbb{R}} \left(1 - e^{\frac{-t^2x^2}{2}}\right) \Pi(dx) < \infty \tag{4.3.13}$$

Como $\Pi(\mathbb{R}) < \infty$, entonces $\int_{\mathbb{R}} 1 \wedge |x| \Pi(dx) < \infty$ y como además $q^2 = 0$, por el lema 4.3.6 X es de variación acotada. Si (δ, Π) es su pareja característica, Ψ se reescribe como

$$\Psi(\lambda) = i\delta\lambda + \int_{\mathbb{R}} \left(1 - e^{i\lambda x}\right) \Pi(dx).$$

Pero entonces, por el inciso anterior,

$$-i\delta = \lim_{|\lambda| \to \infty} \frac{\Psi(\lambda)}{\lambda} = 0$$

pues Ψ es acotada, por lo que $\delta = 0$. Al ser Π finita, se sigue que X es un proceso Poisson compuesto. ■

4.4 Subordinadores

Por falta de tiempo, no fue posible profundizar más en este tema durante la realización del trabajo así que solo daremos un breve recorrido por sus propiedades elementales y presentaremos algunos ejemplos. Se pueden consultar por ejemplo el texto de Bertoin sobre subordinadores [10] así como Kyprianou [13] capítulo 5.

4.4.1 Definiciones y exponente de Laplace

Sea $(\Omega, \mathscr{F}, \mathbb{P})$ un espacio de probabilidad.

Definición 4.4.1 *Un proceso de Lévy $\sigma = (\sigma_t)$ en $\mathbb{R}$ creciente es llamado un subordinador.*

Al ser creciente, σ es positivo, de variación acotada y por el lema 4.3.6, σ no tiene componente browniana y su exponente característico Ψ queda determinado por una pareja (δ, Π) donde Π satisface $\int_{\mathbb{R}} 1 \wedge |x| \Pi(dx) < \infty$. Como no tiene saltos negativos, Π tiene su soporte en $(0, \infty)$ y por el Corolario 4.3.7 Ψ se puede escribir como

$$\Psi(\lambda) = i\delta\lambda + \int_{\mathbb{R}^+} \left(1 - e^{i\lambda x}\right) \Pi(dx)$$

Como $\int_{\mathbb{R}}(1 \wedge |x|)\Pi(dx) < \infty$, por la fórmula de Campbell,

$$\sigma_t = \delta t + \int_{[0,t]\times\mathbb{R}^+} xN(ds \times dx).$$

donde N es medida aleatoria de Poisson en $\mathbb{R}^+ \times \mathbb{R}$ con intensidad $ds \otimes \Pi$. Obsérvese que $\int_{[0,t]\times\mathbb{R}^+} xN(ds \times dx) < \infty$ pues $\int_{\mathbb{R}}(1 \wedge |x|)\Pi(dx) < \infty$ y por lo tanto, aunque σ brinca una infinidad de veces en $[0,t)$ si Π no es finita, la suma de estos es finita. En resumen, tenemos

Teorema 4.4.2 *Si σ es un subordinador, existe una pareja (δ, Π) donde Π es una medida de Lévy con soporte en $(0,\infty)$ que además satisface $\int_{\mathbb{R}^+}(1 \wedge x)\Pi(dx) < \infty$ tal que para todo $\lambda \geq 0$,*

$$\Psi(\lambda) = i\delta\lambda + \int_{\mathbb{R}^+} \left(1 - e^{i\lambda x}\right) \Pi(dx). \tag{4.4.1}$$

Recíprocamente, toda función de la forma (4.4.1) con Π medida de Lévy en $\mathbb{R}^+$ y tal que $\int_{\mathbb{R}^+}(1 \wedge x)\,\Pi(dx) < \infty$ es el exponente de un subordinador.

Al tratar con subordinadores, se suele trabajar con su transformada y exponente de Laplace que denotaremos Φ en lugar del de Fourier Ψ. Obsérvese que como σ_t es positivo, su transfromada de Laplace siempre existe y está dada por:

$$\mathscr{L}_{\sigma_t}(\lambda) := \mathbb{E}\left[e^{-\lambda\sigma_t}\right] = \exp\left\{ - t\delta\lambda - \int_{\mathbb{R}} \left(1 - e^{-\lambda x}\right) \Pi(dx)\right\}.$$

es decir, $\mathscr{L}_{\sigma_t}(\lambda) = \varphi_{\sigma_1}(i\lambda)$. Similarmente al exponente característico, definimos el exponente de Laplace de σ como

$$\Phi(\lambda) = -\log\left\{\mathscr{L}_{\sigma_1}(\lambda)\right\} = \delta\lambda + \int_{\mathbb{R}} \left(1 - e^{-\lambda x}\right) \Pi(dx)$$

por lo que similarmente $\Phi(\lambda) = \Psi(i\lambda)$. Así, el teorema 4.4.2 sigue siendo válido si reemplazamos Ψ por la expresión de Φ.

4.4.2 Ejemplos de subordinadores

Procesos Poisson y Poisson compuesto

Sea $N = (N_t)$ un proceso Poisson de parámetro c. Su exponente de Laplace está dado por

$$\Phi(\lambda) = c(1 - e^{-\lambda}) = c\int_{\mathbb{R}^+}(1 - e^{-\lambda x})\delta_1(dx) \tag{4.4.2}$$

Por lo tanto, N es un subordinador con medida $\Pi(dx) = c\delta_1(dx)$ y drift $\delta = 0$.
Sean (Y_i) una colección de variables i.i.d. con valores en $\mathbb{R}^+$. Entonces

$$C_t = \sum_{i=1}^{N_t} Y_i$$

es un proceso Poisson compuesto de parámetro c con medida de saltos $\mathbb{P}_{Y_1}$. Como $Y_i \in \mathbb{R}^+$, el soporte de $\mathbb{P}_{Y_1}$ está en $\mathbb{R}^+$ y C solo puede tener saltos positivos. Su exponente de Laplace está dado por:

$$\Phi(\lambda) = c\int_{\mathbb{R}^+}(1 - e^{-\lambda x})\mathbb{P}_{Y_1}(dx)$$

Como $\mathbb{P}_{Y_1}(\mathbb{R}^+) = 1$ al ser medida de probabilidad, la medida $\Pi(dx) = c\mathbb{P}_{Y_1}(dx)$ satisface $\int_{\mathbb{R}^+}(1 \wedge x)\,\Pi(dx) < \infty$ y entonces C es un subordinador con pareja característica $(0, \Pi)$.

Subordinadores estables de índice $\alpha \in (0,1)$.

Para cada $\alpha \in (0,1)$, definimos la siguiente familia de exponentes de Laplace:

$$\Phi(\lambda) = \frac{\alpha}{\Gamma(1-\alpha)} \int_{\mathbb{R}^+} (1 - e^{-\lambda x}) \frac{1}{x^{1+\alpha}} \, dx \tag{4.4.3}$$

Si $\Pi(dx) = \frac{\alpha}{\Gamma(\alpha)} \frac{\mathbb{1}_{\mathbb{R}^+}(x)}{x^{1+\alpha}} \, dx$, Π tiene soporte en $\mathbb{R}^+$ y satisface

$$\int_{\mathbb{R}^+} (1 \wedge x) \frac{1}{x^{1+\alpha}} \, dx = \int_0^1 x \frac{1}{x^{1+\alpha}} \, dx + \int_1^\infty \frac{1}{x^{1+\alpha}} \, dx < \infty$$

por lo que $\int_{\mathbb{R}^+} (1 \wedge x) \, \Pi(dx) < \infty$.

Definición 4.4.3 *Un proceso de Lévy con exponente de Laplace*

$$\Phi(\lambda) = \frac{\alpha}{\Gamma(1-\alpha)} \int_{\mathbb{R}^+} (1 - e^{-\lambda x}) \frac{1}{x^{1+\alpha}} \, dx \qquad \alpha \in (0,1)$$

es llamado subordinador estable de índice α.

Obsérvese que $\Pi(\mathbb{R}^+) = \infty$ para cualquier α y por lo tanto, un subordinador estable σ tiene una infinidad de saltos (positivos) en cualquier intervalo de tiempo. La condición $\int_{\mathbb{R}^+} 1 \wedge x \, \Pi(dx) < \infty$ garantiza por Campbell que la suma de estos hasta $t \geq 0$ es finita para cualquier $t \geq 0$. La restricción $\alpha \in (0,1)$ es para garantizar $\int_{\mathbb{R}^+} (1 \wedge x) \, \Pi(dx) < \infty$.

Lema 4.4.4 *Si σ es un subordinador estable de índice $\alpha \in (0,1)$, entonces*

$$\Phi(\lambda) = \lambda^\alpha$$

Demostración
Recordemos que la densidad gamma de parámetros $\lambda, \beta \in (0,\infty)$ tiene soporte en $\mathbb{R}^+$ y está dada por $f(x) = \frac{\lambda^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\lambda x}$ Como $\alpha \in (0,1)$, entonces consideramos a $\beta := 1 - \alpha \in (0,\infty)$ y

$$\frac{\lambda^{1-\alpha}}{\Gamma(1-\alpha)} \int_0^\infty x^{(1-\alpha)-1} e^{-\lambda x} dx = 1$$

al ser una densidad con soporte en $\mathbb{R}^+$. Por lo tanto

$$\lambda^\alpha = \frac{1}{\Gamma(1-\alpha)} \int_0^\infty \underbrace{x^{-\alpha}}_{f} \underbrace{\lambda e^{-\lambda x}}_{g'} dx = \frac{\alpha}{\Gamma(\alpha)} \int_{\mathbb{R}^+} (1 - e^{-\lambda x}) \frac{1}{x^{1+\alpha}} \, dx = \Phi(\lambda)$$

por integración por partes. ■

Similarmente al browniano, los subordinadores estables satisfacen una propiedad de escalamiento:

Proposición 4.4.5 *Sea σ un subordinador estable de índice α. Entonces, para todo $t \geq 0$ el proceso*

$$\tilde{\sigma}_t = k^{-1/\alpha} \sigma_{kt} \quad t \geq 0$$

tiene la misma distribución que σ.

Demostración
Basta probar que $\tilde{\sigma}$ tiene el mismo exponente de Laplace que σ. Al ser σ subordinador estable, tiene exponente $\Phi(\lambda) = \lambda^\alpha$. Calculando la transformada de Laplace de $\tilde{\sigma}_1$ obtenemos que

$$\mathscr{L}(\tilde{\sigma}_1)(\lambda) = \mathbb{E}\left[\exp\{-\lambda\tilde{\sigma}_1\}\right] = \mathbb{E}\left[\exp\{-\lambda k^{-1/\alpha}\sigma_k\}\right] = \mathscr{L}(\sigma_k)(\lambda k^{-1/\alpha})$$

y como $\mathscr{L}(\sigma_1)(\lambda) = \exp\{-k\Phi(\lambda)\}$, al remplazar obtenemos que

$$\begin{aligned}\mathscr{L}(\tilde{\sigma}_1)(\lambda) &= \exp\{-k\Phi(\lambda k^{-1/\alpha})\}\\ &= \exp\{-k\left(\lambda k^{-1/\alpha}\right)^\alpha\}\\ &= \exp\{-\lambda^\alpha\}.\end{aligned}$$

Es decir $\tilde{\sigma}$, tiene exponente de Laplace $\tilde{\Phi} = \lambda^\alpha$ y entonces coincide con el de σ. ■

Proceso de llegadas del movimiento browniano.

Referencia: Bertoin [10], capítulo 1.

Sea B un movimiento browniano y $(\mathscr{F}_t)$ su filtración natural (completada) en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$. Para cada t, definimos al siguiente tiempo de paro:

$$\sigma_t = \inf\{u \geq 0 : B_u > t\}$$

y consideremos al proceso $(\sigma_t : t \geq 0)$; es claramente creciente y al ser σ_t finito casi seguramente, está bien definido. Como $\mathscr{F}_{\sigma_s} \subset \mathscr{F}_{\sigma_t}$ si $s \leq t$, la colección $\tilde{\mathscr{F}}_t = \mathscr{F}_{\sigma_t}$ para $t \geq 0$ es una filtración. Además, al ser σ_t una variable $\mathscr{F}_{\sigma_t}$-medible, el proceso (σ_t) es $\tilde{\mathscr{F}}_t$-adaptado.

Lema 4.4.6 *El proceso (σ_t) es un subordinador adaptado a la filtración $(\tilde{\mathscr{F}}_t)$.*

Demostración
Falta solamente probar que los incrementos son independientes y estacionarios. Sean $s, t > 0$, nótese que por continuidad $B_{\sigma_t} = t$ y como $\sigma_{t+s} > \sigma_t$,

$$\begin{aligned}\sigma_{t+s} - \sigma_t &= \inf\{u \geq \sigma_t : B_u > t + s\} - \sigma_t\\ &= \inf\{u \geq \sigma_t : B_u - B_{\sigma_t} > s\} - \sigma_t\\ &= \left(\inf\{u \geq 0 : B_{u+\sigma_t} - B_{\sigma_t} > s\} + \sigma_t\right) - \sigma_t\end{aligned}$$

Por Propiedad fuerte de Markov, $B' = (B_{u+\sigma_t} - B_{\sigma_t} : u \geq 0)$ es un browniano independiente de $\tilde{\mathscr{F}}_t$ y por lo tanto, $\sigma_{t+s} - \sigma_t \sim \sigma_s$. La independencia de incrementos también se sigue de la Propiedad fuerte de Markov:

$$\sigma_{t+s} - \sigma_t = \inf\{u \geq 0 : B_{u+\sigma_t} - B_{\sigma_t} > s\} \text{ que es independiente de } \sigma_r, \text{ para } r \leq t.$$

al tratarse de una variable $\tilde{\mathscr{F}}_t$-medible. Inductivamente, esto se generaliza para cualquier colección de incrementos y concluimos que los incrementos de σ son independientes y estacionarios. Al ser las trayectorias continuas cadlag, concluimos que σ es un $(\tilde{\mathscr{F}}_t)$-subordinador.
■

4.5 Generalizaciones a $\mathbb{R}^d$

Existen generalizaciones a mayor dimensión de los resultados probados a lo largo de este capítulo. Enunciaremos sin probar algunos de estos. Similaremente, decimos que el vector aleatorio X con valores en $\mathbb{R}^d$ es infinitamente divisible si para todo n, existen vectores aleatorios independientes e idénticamente distribuidos $(X_j)_{j=1}^n$ tales que

$$\varphi_X(\lambda) = (\varphi_{X_n}(\lambda))^n$$

La definición de un proceso de Lévy es la misma que la enunciada para el caso real al considerar que ahora X toma valores en $\mathbb{R}^d$. Si

$$\varphi_{X_1}(\lambda) = \mathbb{E}\left[e^{i\langle\lambda, X_1\rangle}\right] \quad \lambda \in \mathbb{R}^d$$

es la característica de X_1, al ser infinitamente divisible se escribe como

$$\varphi_{X_1}(\lambda) = e^{-\Psi(\lambda)}$$

donde $\Psi(\lambda) : \mathbb{R}^d \to \mathbb{C}$ se define como el exponente característico de X. En este caso, también se tiene la fórmula de Lévy Khintchine

Teorema 4.5.1 *Si X es un proceso de Lévy en $\mathbb{R}^d$ con exponente característico Ψ, entonces existe una tripleta (a, Q, Π) donde $a \in \mathbb{R}^d$ y Q es una forma cuadrática positiva semi-definida en $\mathbb{R}^d$, Π es una medida en $\mathbb{R}^d$ que satisface $\Pi(\{0\}) = 0$ y $\int_{\mathbb{R}^d}(1 \wedge x^2)\Pi(dx) < \infty$ que permite escribir a Ψ como:*

$$\Psi(\lambda) = i\langle a, \lambda\rangle + \frac{Q(\lambda)}{2} + \int_{\mathbb{R}^d}\left(1 - e^{i\langle x, \lambda\rangle}\right)\mathbb{1}_{\{|x|\geq 1\}}\Pi(dx) + \int_{\mathbb{R}^d}\left(1 - e^{i\langle x, \lambda\rangle} + i\langle x, \lambda\rangle\right)\mathbb{1}_{\{|x|<1\}}\Pi(dx)$$

para $\lambda \in \mathbb{R}^d$.

Los resultados estudiados en este capítulo tienen generalizaciones para procesos de Lévy en $\mathbb{R}^d$ aunque los omitiremos. Se puede consultar por ejemplo Bertoin [9] y Sato [15] pues en ambos se trabaja con procesos en $\mathbb{R}^d$. No trabajaremos con procesos de Lévy con valores en $\mathbb{R}^d$ hasta que abordemos cuestiones de recurrencia y transitividad en el capítulo 6.

Capítulo 5

Procesos de Markov

5.1 Filtraciones y tiempos de paro

Referencia: Seguimos la presentación de Le Gall [11], Secciones 3.3 y 3.2 y completamos con algunos ejercicios y definiciones de Revuz & Yor [2] capítulo 3 y Çinlar [5] capítulo 5.

En esta sección se hace un breve repaso de propiedades conocidas de medibilidad de procesos y tiempos de paro. Este material se puede encontrar en cualquier libro de cálculo estocástico así que nos restringimos a los resultados esenciales y dejamos en el Apéndice pruebas de propiedades relativas a tiempos de paro que se usarán con frecuencia. Esta es únicamente una sección de referencia.

5.1.1 Definiciones

Definición 5.1.1 *Una filtración en* $(\Omega, \mathscr{F}, \mathbb{P})$ *es una colección de sub-*σ*-álgebras de* $\mathscr{F}$ $(\mathscr{F}_t)_{0\leq t\leq\infty}$ *tales que* $\mathscr{F}_s \subseteq \mathscr{F}_t$ *si* $s \leq t \leq \infty$*. Entonces, para cada* $0 \leq s < t$

$$\mathscr{F}_0 \subseteq \mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \mathscr{F}_\infty \subseteq \mathscr{F}.$$

A $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ *se le llama un espacio de probabilidad filtrado.*

Si $X = (X_t)$ es un proceso en este espacio, la mínima filtración con respecto a la cual X_t es $\mathscr{F}_t$-medible para cada t es la definida por $\mathscr{F}_t^0 = \sigma(\,X_s\,; 0 \leq\, s \leq t)$ para cada t y $\mathscr{F}_\infty^0 = \sigma(\,X_s\,;\, s \geq 0)$. Se le llama la filtración canónica o natural de X. Heurísticamente, $\mathscr{F}_t$ contiene los eventos que podrían haber ocurrido hasta o en el tiempo t. A partir de $(\mathscr{F}_t)$ podemos construir otra filtración, definiendo para cada t

$$\mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s$$

Está claro que $\mathscr{F}_t \subseteq \mathscr{F}_{t+}$. La filtración $\mathscr{F}_{t+}$ contiene información de los eventos que se pueden observar en un futuro inmediato. Por ejemplo, si Ω es el espacio de las trayectorias continuas por la izquierda $G(\mathbb{R}^+, \mathbb{R})$ y X el proceso coordenado, el evento $\{X \text{ discontinuo en } t\}$ pertenece a $\mathscr{F}_{t+}$ pues basta con conocer información de la trayectoria para un $\{t + \epsilon : \epsilon > 0\}$ para determinar si es discontinua en t o no. Veremos un ejemplo en concreto más adelante.

Definición 5.1.2 *Una filtración es continua por la derecha si* $\mathscr{F}_t = \mathscr{F}_{t+}$ *para cada* t*.*

Como se verá en un ejemplo posterior, esto no tiene por que ocurrir. No es difícil comprobar que la filtración $(\mathscr{F}_{t+})$ es continua por la derecha.

Definición 5.1.3 *Un proceso X es adaptado a una filtración $(\mathscr{F}_t)$ si X_t es $\mathscr{F}_t$ medible para cada t.*

En particular, es obvio que X es adaptado a su filtración canónica. A menudo es necesario trabajar con procesos que satisfacen propiedades más fuertes de medibilidad:

Definición 5.1.4 *Decimos que un proceso X es medible si la aplicación*

$$(t,\omega) \longmapsto X_t(\omega)$$

de $\mathbb{R}^+ \times \Omega$ en $(E,\mathscr{E})$ es una función $\mathscr{B}(\mathbb{R}^+)\otimes\mathscr{F}$-medible.
Un proceso X es progresivamente medible con respecto a una filtración $(\mathscr{F}_t)$ si para cada t, la aplicación $(s,\omega)\to X_s(\omega)$ de $[0,t]\times\Omega$ en E es $\mathscr{B}([0,t])\otimes\mathscr{F}_t$-medible.

Para probar cuestiones de medibilidad, la medibilidad conjunta del proceso es una condición a menudo indispensable. Tenemos la siguiente sencilla caracterización:

Teorema 5.1.5 *Un proceso adaptado a $(\mathscr{F}_t)$ continuo por la derecha es progresivo.*

En particular, todo proceso de Lévy $\mathscr{F}_t$-adaptado es $\mathscr{F}_t$-progresivo. Se puede leer la prueba en LeGall [11] páginas 42 y 43.

Definición 5.1.6 *Una variable aleatoria en $(\Omega,\mathscr{F})$, $T:\Omega\to\mathbb{R}^+\cup\infty$ se le llama un tiempo de paro con respecto a una filtración $(\mathscr{F}_t)$ si $\{T\le t\}\in\mathscr{F}_t$ para todo $t\ge 0$.*

Si no se especifica lo contrario, por el resto de la sección T es tiempo de paro es con respecto a $(\mathscr{F}_t)$. Obsérvese que el evento $\{T<t\}\in\mathscr{F}_t$ para toda $t\ge 0$ pues $\{T<t\}=\cup_{s<t}\{T\le s\}$ pero no es suficiente pedir que $\{T<t\}\in\mathscr{F}_t$ para todo $t\ge 0$ para que T sea un tiempo de paro. Que T sea tiempo de paro es equivalente a que el proceso

$$Z_t = \mathbb{1}_{\{T\le t\}}, \quad t\ge 0$$

sea adaptado a $(\mathscr{F}_t)$. Si interpretamos a T como el momento en que ocurre un evento de interés, Z nos permite detectar para cada tiempo t si el evento ya ocurrió, en vista de que mientras este no ocurra, Z tomará el valor 0. Pensando a $(\mathscr{F}_t)$ como el flujo de información acumulada hasta tiempo t, que Z sea adaptado a $(\mathscr{F}_t)$ implica que $\{Z_t=1\}\in\mathscr{F}_t$ y entonces se puede interpretar como que la información acumulada hasta cada tiempo t es suficiente para determinar si ya ocurrió el evento de interés.

Definición 5.1.7 *Si T es un tiempo de paro, la σ-álgebra del pasado hasta T ó σ-álgebra parada en T se define como*

$$\mathscr{F}_T = \left\{A\in\mathscr{F}_\infty\,;\,\forall\, t\ge 0,\, A\cap\{T\le t\}\in\mathscr{F}_t\,\right\}$$

En efecto, así definida $\mathscr{F}_T$ es una σ-álgebra:

Demostración
Sea $A \in \mathscr{F}_T$, probemos que $A^c \in \mathscr{F}_T$.
Como $A \cap \{T \leq t\} \in \mathscr{F}_t$, entonces $A^c \cup \{T > t\} \in \mathscr{F}_t$ y $\{T \leq t\} \in \mathscr{F}_t$ por ser T un tiempo de paro. De lo anterior, se deduce que

$$A^c \cap \{T \leq t\} = (A^c \cup \{T > t\}) \cap \{T \leq t\} \in \mathscr{F}_t.$$

Como Ω y $\emptyset$ están en $\mathscr{F}_T$, falta ver que es cerrada bajo uniones numerables. Sea $(A_n)_{n\in\mathbb{N}}$ sucesión con $A_n \in \mathscr{F}_T$ para todo n. Veamos que $\cup_n A_i \in \mathscr{F}_T$.

$$\left(\bigcup_{n=1}^{\infty} A_n\right) \cap \{T \leq t\} = \bigcup_{n=1}^{\infty} \underbrace{(A_n \cap \{T \leq t\})}_{\in \mathscr{F}_t} \in \mathscr{F}_t$$

por lo que $\mathscr{F}_T$ es σ-álgebra. ∎

Si T es un tiempo de paro con respecto a $(\mathscr{F}_t)$, también lo es con respecto a $(\mathscr{F}_{t+})$ pero el recíproco no es válido en general.

Definición 5.1.8 *A un tiempo de paro con respecto a $(\mathscr{F}_{t+})$ se le llama un tiempo de paro $\mathscr{F}_t$-opcional.*

Intuitivamente, un tiempo de paro opcional requiere una pizca de información del futuro $t + \epsilon$ para cualquier $\epsilon > 0$ para determinar si el evento de interés ocurrió en el tiempo t. Justamente la σ-álgebra $\mathscr{F}_{t+}$ contiene esa información adicional. El siguiente ejemplo ilustra este hecho:

Ejemplo: Sea X una variable aleatoria con valores en $\{-1, 1\}$. Definimos el proceso

$$Z_t = t \cdot X \qquad\qquad t \geq 0$$

Z es una recta que pasa por el origen y X determina su pendiente. Si $\mathscr{F}_t = \sigma(Z_s\,;\, s \leq t)$ su filtración natural, como $Z_0 \equiv 0$, $\mathscr{F}_0$ es la σ-álgebra trivial mientras que para cualquier $t > 0$,

$$\mathscr{F}_t = \sigma(X) = \big\{\{X = 1\}, \{X = -1\}\big\} \cup \mathscr{F}_0$$

ya que al ser la función $f(x) = t \cdot x$ invertible si $t > 0$, $\sigma(X) = \sigma(f(X)) = \sigma(Z_t)$ (Apéndice 1, 9.1.1). Es claro entonces, que la filtración $(\mathscr{F}_t)$ no es continua por la derecha. La variable aleatoria

$$T = \begin{cases} 0 & \text{si } X = 1 \\ 2 & \text{si } X = -1 \end{cases}$$

no es un tiempo de paro con respecto a $(\mathscr{F}_t)$ pues $\{T = 0\} = \{X = 1\} \notin \mathscr{F}_0$. Con la información que contamos al tiempo $t = 0$, no podemos determinar si X tomó el valor 1 o -1 pues en ambos casos $Z_0 = 0$. Sin embargo, si lo es con respecto a la filtración $(\mathscr{G}_t) = (\mathscr{F}_{t+})$ que en este caso resulta ser la filtración constante:

$$\mathscr{F}_{0^+} = \bigcap_{\epsilon>0} \mathscr{F}_\epsilon = \big\{\{X = 1\}, \{X = -1\}\big\} \cup \mathscr{F}_0$$

Así, T es tiempo de paro $\mathscr{F}_t$-opcional pero no es un $\mathscr{F}_t$ tiempo de paro.

Proposición 5.1.9 (Aproximación de tiempos de paro)
Para cualquier tiempo de paro T, existe una sucesión decreciente (T_n) de tiempos de paro tal que $\lim_{n\to\infty} T_n(\omega) = T(\omega)$ para cada ω.

Demostración
Definimos para cada n

$$T_n(\omega) = \begin{cases} \frac{k}{2^n} & \text{si } T(\omega) \in [\frac{k-1}{2^n}, \frac{k}{2^n}),\ k \in \mathbb{N} \\ \infty & \text{si } T(\omega) = \infty \end{cases}$$

Como $|T(\omega) - T_n(\omega)| \leq \frac{1}{2^n}$ tenemos que $T_n \downarrow T$. Veamos que T_n es tiempo de paro.

Como $T_n \geq T$, por el corolario 9.2.5 del Apéndice, es suficiente probar que T_n es $\mathscr{F}_T$ medible y al ser T_n discreto, basta con probar que $T_n \in \frac{k}{2^n} \in \mathscr{F}_T$ para cada k. Pero

$$\left\{T_n = \frac{k}{2^n}\right\} = \left\{T \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right)\right\} \in \mathscr{F}_T$$

y

$$\{T_n = \infty\} = \{T = \infty\} \in \mathscr{F}_T$$

pues T es $\mathscr{F}_T$-medible.

Teorema 5.1.10 *Si X es un proceso progresivo y si T es un tiempo de paro, X_T es $\mathscr{F}_T$-medible en $\{T < \infty\}$.*

La prueba se puede leer en Yor, proposición 4.8 y 4.9. Finalmente, introducimos una familia de tiempos de paro que usaremos con frecuencia:

Proposición 5.1.11 *Sea (X_t) un proceso adaptado a una filtración $(\mathscr{F}_t)$ con valores en un espacio métrico (E, d). Supongamos que las trayectorias de X son continuas por la derecha y sea O un abierto de E. Entonces*

$$T_O = \inf\{t \geq 0 \,:\, X_t \in O\}$$

es un tiempo de paro con respecto a la filtración $(\mathscr{F}_{t+})$.

Se puede consultar la prueba en LeGall [11].

5.2 La Propiedad de Markov

Referencias: Revuz & Yor [2] capítulo III y Spieksma [3] capítulo 3.

A continuación, daremos las ideas heurísticas de lo que se definirá más adelante como la propiedad de Markov y el semigrupo de un proceso de Markov. $(E, \mathscr{E})$ será R^d o un espacio polaco con la sigma-álgebra de sus borelianos. Un proceso X con filtración natural $\mathscr{F}_t^0 = \sigma(X_u\,; u \leq t)$, $t \geq 0$ en $(E, \mathscr{E})$ para el cual el futuro del proceso X_t condicionado a la σ-álgebra $\mathscr{F}_s^0$ con $s < t$ solo depende de $\sigma(X_s)$ se dice que satisface la propiedad de Markov. Es decir, para cualquier $A \in \mathscr{E}$ y $s < t$,

$$\mathbb{P}\left(X_t \in A | \sigma(X_u\,; u \leq s)\right) = \mathbb{P}\left(X_t \in A | \sigma(X_s)\right) \tag{5.2.1}$$

por lo que, para cada A, la probabilidad condicional con respecto a $\mathscr{F}_s$ es una variable $\sigma(X_s)$ medible. Como $\mathbb{P}(X_t \in A|\sigma(X_s))$ es $\sigma(X_s)$ medible, debe existir una función borel medible $h : E \to \mathbb{R}$ tal que $\mathbb{P}(X_t \in A|\sigma(X_s)) = h(X_s)$. Como además (5.2.1) depende tanto de A como de t y s, denotaremos a esta $h(x)$ por $P_{s,t}(x, A)$. Esto es,

$$\mathbb{P}(X_t \in A|\sigma(X_s)) = P_{s,t}(X_s, A) \tag{5.2.2}$$

donde $x \to P_{s,t}(x, A)$ es $\mathscr{E}$-medible para cada $A \in \mathscr{E}$ mientras que $A \to P_{s,t}(x, A)$ es una medida de probabilidad en E para cada x. Nótese que la definición que dimos de P es independiente de que se satisfaga (5.2.1).

Definición 5.2.1 *Sea $(E, \mathscr{E})$ espacio medible. Un kernel o probabilidad de transición P en E es una función $P : E \times \mathscr{E} \to [0, 1]$ tal que*
i) Para cada $x \in E$, $B \to P(x, B)$ es una medida de probabilidad en $(E, \mathscr{E})$.
ii) Para cada $B \in \mathscr{E}$, $x \to P(x, B)$ es $\mathscr{E}/\mathscr{B}([0, 1])$-medible .

Definición 5.2.2 *Una familia de probabilidades de transición $(P_{s,t})$ es llamada homogénea si para $t > s$ cualesquiera*

$$P_{s,t} = P_{0,t-s}. \tag{5.2.3}$$

En ese caso, denotaremos por P_t a $P_{0,t}$ y trabajaremos con la colección $(P_t)_{t\geq 0}$. De aquí en adelante solo consideraremos familias de este tipo.

Seguimos motivando las definiciones: supóngase que X es un proceso que satisface (5.2.1) y que P es un kernel homogéneo que satisface (5.2.2). Sabemos que $\mathbb{P}(X_{t+s} \in A|\sigma(X_u; u \leq s)) = P_t(X_s, A)$ lo que es equivalente a:

$$\mathbb{E}\left[\mathbb{1}_A(X_{t+s})|\mathscr{F}_s^0\right] = \int_E \mathbb{1}_A(y) P_t(X_s, dy)$$

así que por linealidad y los métodos usuales de aproximación se tiene que

$$\mathbb{E}\left[f(X_{t+s})|\mathscr{F}_s^0\right] = \int_E f(y) P_t(X_s, dy).$$

para cualquier f positiva o acotada. Entonces si $t, s \in \mathbb{R}^+$

$$\begin{aligned} P_{t+s}(X_h, A) &= \mathbb{P}\left(X_{h+t+s} \in A|\mathscr{F}_h^0\right) \\ &= \mathbb{E}\left[\mathbb{P}\left(X_{s+h+t} \in A|\mathscr{F}_{s+h}^0\right)|\mathscr{F}_h^0\right] \\ &= \mathbb{E}\left[P_t(X_{s+h}, A)|\mathscr{F}_h^0\right] \\ &= \int_E P_t(y, A) P_s(X_h, dy) \end{aligned}$$

lo cual motiva la siguiente definición:

Definición 5.2.3 *Una familia de probabilidades de transición (P_t) en $(E, \mathscr{E})$ es llamada función de transición (homogénea) si para todo $s, t \geq 0$, $x \in E$ y $B \in \mathscr{E}$ se satisface*

$$P_{t+s}(x, B) = \int_E P_t(x, dy) P_s(y, B). \tag{5.2.4}$$

A esta relación se le conoce como la ecuación de Chapman-Kolmogorov.

Para cualquier t, P_t actúa sobre las funciones $\mathscr{E}/\mathscr{B}(\mathbb{R})$-medibles y acotadas $b\mathscr{E}$(o positivas $\mathscr{E}^+$) $f : E \to \mathbb{R}$ como sigue:

$$(P_t f)(x) = \int_E P_t(x, dy) f(y) = \int_E f(y) P_t(x, dy).$$

Usaremos ambas notaciones a lo largo del texto. $P_t f$ es a su vez una función $\mathscr{E}/\mathscr{B}(\mathbb{R})$-medible y acotada (o positiva) de E en $\mathbb{R}$ pues se puede ver como límite de medibles al aproximar a f por simples, y es acotada (resp. positiva). Con esta notación, la ecuación de Chapman-Kolmogorov se ve como

$$P_{t+s} \mathbb{1}_B = P_t P_s \mathbb{1}_B$$

y por linealidad y Teorema de convergencia dominada, se extiende para cualquier $f \in b\mathscr{E}$:

$$P_{t+s} f = P_t P_s f \tag{5.2.5}$$

que es una forma abreviada de escribir:

$$\int_E P_{t+s}(x, dy) f(y) = \int_E P_t(x, dy) \int_E P_s(y, dz) f(z).$$

A una familia de operadores $(P_t)_{t \geq 0}$ en $b\mathscr{E}$ que cumple (5.2.5) se le llama *semigrupo de operadores* en $b\mathscr{E}$.

Definición 5.2.4 *Sea $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ un espacio de probabilidad filtrado. Un proceso X adaptado a la filtración $(\mathscr{F}_t)$ satisface la propiedad de Markov con respecto a la filtración $(\mathscr{F}_t)$ y función de transición (P_t) si para cualquier $s, t \in \mathbb{R}^+$ con $s < t$*

$$\mathbb{E}\left[f(X_{t+s}) | \mathscr{F}_s \right] = P_t f(X_s) \tag{5.2.6}$$

para cualquier $f \in \mathscr{E}^+$. A la distribución $\mathbb{P}_{X_0}$ de X_0, se le llama la distribución inicial de X y se le suele denotar por ν.

Cando X satisface la propiedad de Markov con respecto a una filtración $(\mathscr{F}_t)$ que no es necesariamente la natural, se dice que satisface la $\mathscr{F}_t$-propiedad de Markov mientras que si se trata de su filtración natural, solo se dice que satisface la propiedad de Markov. Como $\mathscr{F}_t^0 \subset \mathscr{F}_t$, si X satisface la propiedad de Markov con respecto a $(\mathscr{F}_t)$ entonces lo hace con respecto a $(\mathscr{F}_t^0)$

$$\mathbb{E}\left[f(X_{t+s}) | \mathscr{F}_s^0 \right] = \mathbb{E}\left[\mathbb{E}\left[f(X_{t+s}) | \mathscr{F}_s \right] | \mathscr{F}_s^0 \right] = \mathbb{E}\left[P_t f(X_s) | \mathscr{F}_s^0 \right] = P_t f(X_s).$$

Es importante notar que si modificamos la medida de probabilidad $\mathbb{P}$, X no tiene por qué seguir satisfaciendo la propiedad de Markov. Si X no tiene asociada una función de transición, simplemente decimos que satisface la propiedad de Markov con respecto a $(\mathscr{F}_t)$ si

$$\mathbb{E}\left[f(X_{t+s}) | \mathscr{F}_s \right] = \mathbb{E}\left[f(X_{t+s}) | X_s \right]. \tag{5.2.7}$$

Obsérvese que para probar (5.2.6), basta con que se cumpla

$$\mathbb{P}\left(X_{t+s} \in B | \mathscr{F}_s \right) = P_t(X_s, B) \tag{5.2.8}$$

para $s < t$ y $B \in \mathscr{E}$ arbitrarios: esta igualdad equivale a que se satisfaga

$$\mathbb{E}\left[\mathbb{1}_B(X_{t+s}) | \mathscr{F}_s \right] = P_t \mathbb{1}_B(X_s) \tag{5.2.9}$$

que es (5.2.6) para $f = \mathbb{1}_B$ y por los argumentos habituales se tiene la igualdad para cualquier $f \in b\mathscr{E}$ o $\mathscr{E}^+$. La propiedad de Markov se puede expresar de muchas formas equivalentes, se recomienda consultar Tudor [6], capitulo 5. El siguiente lema nos da un criterio para establecer cuando un proceso satisface la propiedad de Markov con respecto a una función de transición (P_t):

Lema 5.2.5 *Un proceso X satisface la propiedad de Markov con respecto a su filtración natural $(\mathscr{F}_t^0) = (\sigma(X_s\,;\, s \leq t))$ con funciones de transición $(P_t)_{t\geq 0}$ y distribución inicial ν si y solo si para cualquier colección $0 = t_0 < t_1 < \cdots < t_n \quad n \in \mathbb{N}$ y $f_1, \ldots, f_n \in \mathscr{E}^+$*

$$\mathbb{E}\left[\prod_{i=0}^{n} f_i(X_{t_i})\right] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{t_1-t_0}(x_0, dx_1) f_1(x_1) \ldots \int_E P_{t_n-t_{n-1}}(x_{n-1}, dx_n) f_n(x_n) \tag{5.2.10}$$

Obsérvese que entonces si X satisface la propiedad de Markov, sus distribuciones finito-dimensionales quedan determinadas por sus funciones de transición y su distribución inicial ν (y por lo tanto su ley $\mathbb{P}_X$ en $E^{\mathbb{R}^+}$). Esto se sigue que, tomando $f_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{E}$

$$\mathbb{P}\left(X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots X_{t_n} \in A_n\right) = \int_{A_0} \nu(dx_0) \int_{A_1} P_{t_1-t_0}(x_0, dx_1) \ldots \int_{A_n} P_{t_n-t_{n-1}}(x_{n-1}, dx_n).$$

Demostración

Supongamos que X satisface Markov. Entonces, por propiedad torre de la esperanza y propiedad de Markov,

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=0}^{n} f_i(X_{t_i})\right] &= \mathbb{E}\left[\prod_{i=0}^{n-1} f_i(X_{t_i}) \mathbb{E}\left[f_n(X_{t_n}) | \mathscr{F}_{t_{n-1}}^0\right]\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{n-1} f_i(X_{t_i}) P_{t_n-t_{n-1}} f_n(X_{t_{n-1}})\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{n-2} f_i(X_{t_i}) \mathbb{E}\left[f(X_{t_{n-1}}) P_{t_n-t_{n-1}} f_n(X_{t_{n-1}}) | \mathscr{F}_{t_{n-2}}^0\right]\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{n-2} f_i(X_{t_i}) P_{t_{n-1}-t_{n-2}} f_{n-1} P_{t_n-t_{n-1}} f_n(X_{t_{n-2}})\right]
\end{aligned}$$

y se repite el proceso hasta llegar a

$$\begin{aligned}
&= \mathbb{E}\left[f_0(X_{t_0})(P_{t_1-t_0} f_1 \ldots P_{t_n-t_{n-1}} f_n)(X_{t_0})\right] \\
&= \nu f_0 P_{t_1-t_0} f_1 \ldots P_{t_n-t_{n-1}} f_n
\end{aligned}$$

por lo que se tiene la igualdad. Probemos el recíproco: para ver que X satisface

$$\mathbb{E}\left[f(X_{t+s}) | \mathscr{F}_s\right] = P_t f(X_s) \tag{5.2.11}$$

bastaría verificar que para todo $A \in \mathscr{F}_s^0$

$$\int_A f(X_{t+s}) d\mathbb{P} = \int_A P_t f(X_s) d\mathbb{P}. \tag{5.2.12}$$

Lo probaremos primero para el $\pi-$sistema usual que genera a $\mathscr{F}_s^0$:

$$\pi = \{A = \{X_{t_0} \in A_0,\, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\}\,;\, 0 = t_0 < t_1 < \cdots < t_n \leq s,\, A_1 \ldots A_n \in \mathscr{E}\}.$$

Como $\mathbb{1}_A = \prod^n \mathbb{1}_{A_i}(X_{t_i})$, al aplicar (5.2.10) con $f_i = \mathbb{1}_{A_i}$ y f obtenemos

$$\begin{aligned}
\int_A f(X_{t+s})d\mathbb{P} &= \mathbb{E}\left[\prod_{i=1}^n \mathbb{1}_{A_i}(X_{t_i}) f(X_{t+s})\right] \\
&= \nu(dx_0)\mathbb{1}_{A_0}(x_0)P_{t_1-t_0}(x_0, dx_1)\mathbb{1}_{A_1}(x_1)\ldots P_{t+s-t_n}(x_{n-1}, dx_n)f(x_n)
\end{aligned}$$

que por Chapman-Kolmogorov es igual a

$$\begin{aligned}
&= \nu(dx_0)\mathbb{1}_{A_0}(x_0)P_{t_1-t_0}(x_0, dx_1)\mathbb{1}_{A_1}(x_1)\ldots P_{s-t_n}(x_{n-1}, dx_n)P_t(x_n, dy)f(y) \\
&= \nu(dx_0)\mathbb{1}_{A_0}(x_0)P_{t_1-t_0}(x_0, dx_1)\mathbb{1}_{A_1}(x_1)\ldots P_{s-t_n}(x_{n-1}, dx_n)(P_t f)(x_n) \\
&= \mathbb{E}\left[\prod_{i=1}^n \mathbb{1}_{A_i}(X_{t_i})(P_t f)(X_s)\right] \\
&= \int_A P_t f(X_s) d\mathbb{P} \qquad (5.2.13)
\end{aligned}$$

Como (5.2.12) se vale para un π-sistema que genera a $\mathscr{F}_s^0$ y la colección de conjuntos para los cuales se cumple (5.2.12) es un lambda sistema, el lema de clases monótonas nos permite concluir (5.2.12). ■

Para probar el regreso del lema anterior, se pueden debilitar las hipótesis: por la discusión que se tuvo en la sección 5.1, para probar que X satisface Markov con respecto a (P_t) basta ver que para s, t positivos y $A \in \mathscr{E}$ arbitrarios, $\mathbb{P}\left(X_{t+s} \in A | \mathscr{F}_t^0\right) = P_t(X_s, A)$ por lo que basta probar (5.2.11) para $f = \mathbb{1}_A$, $A \in \mathscr{E}$. Pero entonces, haciendo este remplazo en el regreso, se sigue que para probar que X cumple Markov bastaría probar que (5.2.10) se cumple para $f_i = \mathbb{1}_{A_i}$:

Corolario 5.2.6 *Si un proceso X en $(\Omega, \mathscr{F}, \mathbb{P})$ satisface*

$$\mathbb{E}\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \int_{A_0} \nu(dx_0) \int_{A_1} P_{t_1-t_0}(x_0, dx_1) \ldots \int_{A_n} P_{t_n-t_{n-1}}(x_{n-1}, dx_n) \qquad (5.2.14)$$

para cualquier colección finita $f_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{E}$, $0 = t_0 < t_1 < \cdots < t_n \quad n \in \mathbb{N}$ entonces cumple la propiedad de Markov con respecto a su filtración natural $(\mathscr{F}_t^0) = (\sigma(X_s\,;\, s \leq t))$, con funciones de transición (P_t) y distribución inicial ν.

Lo que obtuvimos es que el hecho de que un proceso X con ley inicial ν satisfaga la propiedad de Markov con respecto a (P_t), se reduce a una relación entre sus distribuciones finito dimensionales y el semigrupo de transición (P_t) junto con la medida inicial ν.

5.3 Procesos de Markov en espacios canónicos

Nos dedicamos ahora a probar la existencia de un proceso de Markov con distribución inicial ν y semigrupo (P_t) al conocer únicamente a estos últimos. El corolario 5.2.6 y la observación subsecuente nos dan una idea del camino a seguir: si X es un proceso de Markov con distribución inicial ν y semigrupo (P_t) entonces sus distribuciones finito dimensionales deben satisfacen el corolario 5.2.6. Como dada una familia de medidas de probabilidad consistentes, el teorema de Kolmogorov garantiza la existencia de una medida de probabilidad $\mathbb{P}^\nu$ en $(\Omega, \mathscr{F}^0_\infty) = (E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$ bajo la cual el proceso coordenado X tiene por distribuciones f.d. a esta familia de medidas, el problema se reduce a definir su familia de distribuciones finito-dimensionales. El corolario 5.2.6 nos indica justamente como construirlas a partir de (P_t) y ν de modo que el proceso que obtengamos al usar el teorema de Kolmogorov satisfaga la propiedad de Markov.

Así, a partir del semigrupo (P_t) y una distribución inicial $\nu \in \mathcal{P}(E)$, se construirá usando el corolario 5.2.6 una familia de medidas de probabilidad consistentes y aplicaremos el teorema de Kolmogorov para probar la existencia de una medida de probabilidad $\mathbb{P}^\nu$ en el espacio canónico $(\Omega, \mathscr{F}^0_\infty) = (E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$ bajo la cual el proceso coordenado X tenga a esa familia de medidas como distribuciones finito dimensionales y sea un proceso de Markov con semigrupo (P_t). En particular, tendrá distribución inicial $X_0 \sim \nu$. Al hacer esto para cada $\nu \in \mathcal{P}(E)$, lo que obtenemos es la existencia de una *familia* de medidas de probabilidad $(\mathbb{P}^\nu)_{\nu \in \mathcal{P}(E)}$ en el espacio canónico $(\Omega, \mathscr{F}^0_\infty) = (E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$, de modo que el proceso coordenado X bajo cada medida $\mathbb{P}^\nu$ tiene distribución inicial $\mathbb{P}^\nu(X_0 \in A) = \nu(A)$ y satisface la propiedad de Markov en $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$ con respecto a (P_t). La propiedad de Markov y Markov fuerte serán el puente entre las distribuciones de X en cada uno de estos espacios $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)_{\nu \in \mathcal{P}(E)}$ que en un principio podrían parecer ajenos.

5.3.1 Construcción de un proceso de Markov en $E^{\mathbb{R}^+}$

Sea $(E, \mathscr{E})$ un espacio polaco con la σ-álgebra de sus borelianos. Trabajaremos en el espacio canónico $(E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+}) = (\Omega, \mathscr{F}^0_\infty)$, (θ_t) el operador traslación y $X : \Omega \to \Omega$ la identidad en $E^{\mathbb{R}^+}$

$$X(\omega) = \omega \qquad X(\omega)(s) = \omega_s \quad s \geq 0.$$

Recordemos que entonces

$$\Omega = E^{\mathbb{R}^+} \quad \mathscr{F}^0_t = \sigma(X_u \,;\, u \leq t) \quad \mathscr{F}^0_\infty = \mathscr{E}^{\mathbb{R}^+} = \sigma(X_u \,;\, u \in \mathbb{R}^+).$$

Nótese como aún no hemos definido una medida de probabilidad en $(\Omega, \mathscr{F}^0_\infty)$.

Teorema 5.3.1 *Sean (P_t) funciones de transición en $(E, \mathscr{E})$. Para cualquier distribución ν en $(E, \mathscr{E})$ existe una única medida de probabilidad $\mathbb{P}_\nu$ en $(\Omega, \mathscr{F}^0_\infty)$ tal que el proceso coordenado X en $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$ es un proceso de Markov con respecto a su filtración natural $(\mathscr{F}^0_t)$ con funciones de transición (P_t) y medida inicial ν.*

Demostración
Sea ν medida de probabilidad en $(E, \mathscr{E})$. Para cualquier $n \in \mathbb{N}$ y $0 = t_0 < t_1 < \cdots < t_n$, definimos la siguiente familia de medidas de probabilidad en $(E^{n+1}, \mathscr{E}^{n+1})$ a partir de las probabilidades de transición:

$$P^\nu_{t_0,\ldots,t_n}(A_0 \times A_1 \times \cdots \times A_n) = \int_{A_0} \nu(dx_0) \int_{A_1} P_{t_1 - t_0}(x_0, dx_1) \ldots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n)$$

y en $(E^n, \mathscr{E}^n)$

$$P^\nu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \int_E \nu(dx_0) \int_{A_1} P_{t_1 - t_0}(x_0, dx_1) \ldots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n).$$

Por Chapman-Kolmogorov, forman un sistema consistente:
si $t_j > s > t_{j-1}$

$$\begin{aligned}
&P^{\nu}_{t_0,\dots,t_j,\dots,t_n}\left(A_0 \times A_1 \times \cdots \times A_j \times \cdots \times A_n\right) = \\
&= \int_{A_0} \nu(dx_0) \dots \int_{A_j} P_{t_j - t_{j-1}}(x_{j-1}, dx_j) \dots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n) \\
&= \int_{A_0} \nu(dx_0) \dots \int_{E} P_{t_j - s}(x_{j-1}, dy) \int_{A_j} P_{t_j - s}(y, dx_j) \dots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n) \\
&= P^{\nu}_{t_0,\dots,s,t_j,\dots,t_n}\left(A_0 \times A_1 \times \cdots \times E \times A_j \times \cdots \times A_n\right)
\end{aligned}$$

Por el teorema de consistencia de Kolmogorov 1.3.2 existe una medida de probabilidad que denotaremos $\mathbb{P}^{\nu}$ en $(\Omega, \mathscr{F}^0_{\infty}) = (E^{\mathbb{R}^+}, \mathscr{E}^+)$ tal que el proceso coordenado X tiene por distribuciones f.d. a la colección $\mathbb{P}^{\nu}_{t_0,\dots,t_j,\dots,t_n}$.
Entonces por construcción, se satisface que

$$\begin{aligned}
&\mathbb{P}^{\nu}\left(X_{t_0} \in A_0, X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\right) \\
&\qquad = \int_{A_0} \nu(dx_0) \int_{A_1} P_{t_1 - t_0}(x_0, dx_1) \dots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n) \qquad (5.3.1)
\end{aligned}$$

es decir, se tiene la igualdad del lema 5.2.10 para funciones indicadoras $f_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{E}$. Por el corolario 5.2.6, X en $(\Omega, \mathscr{F}^0_{\infty}, \mathbb{P}^{\nu})$ es un proceso de Markov con respecto a (P_t) y su filtración natural $(\mathscr{F}^0_t)$ ■

Si Z es $\mathscr{F}^0_{\infty}$ medible, denotamos

$$\mathbb{E}^{\nu}[Z] = \int_{\Omega} Z \, d\mathbb{P}^{\nu}.$$

Para cada distribución inicial $\nu \in \mathcal{P}(E)$, X en $(\Omega, \mathscr{F}^0_{\infty}, \mathbb{P}^{\nu})$ satisface la propiedad de Markov

$$\mathbb{E}^{\nu}\left[f(X_{t+s}) | \mathscr{F}^0_s\right] = P_t f(X_s)$$

Como el proceso X es siempre el proceso coordenado, lo que se tiene es una propiedad sobre la familia de medidas $(\mathbb{P}^{\nu})$ en el espacio canónico. Por ello, se habla más bien de un *sistema* $\mathcal{X} = (\Omega, \mathscr{F}^0_{\infty}, (\mathscr{F}^0_t), (\theta_t), (X_t), (\mathbb{P}^{\nu}))$; así, la definición con la que iniciamos la siguiente sección, en un contexto un poco más general, debería resultar más natural.

5.3.2 Sistemas y procesos de Markov

Referencia: Revuz & Yor [2] y Jacod [26] capítulo 3

Consideraremos Ω un espacio canónico, ya sea en el que hemos estado trabajando $(E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$, el espacio de las funciones cadlag $(D(\mathbb{R}^+, E), \mathscr{E}^+_D)$ o continuas $(C(\mathbb{R}^+, E), \mathscr{E}^+_C)$ con su operador traslación (θ_t) y $X : \Omega \to \Omega$ la identidad, proceso canónico usual. Definimos $\mathscr{F}^0_{\infty} = \sigma(X_s\,;\, s \geq 0)$ y $\mathscr{F}^0_t$ su filtración natural. Finalmente, sea $(\mathbb{P}^{\nu})_{\nu \in \mathcal{P}(E)}$ familia de medidas de probabilidad en $(\Omega, \mathscr{F}^0_{\infty})$ tales que $\mathbb{P}^{\nu}(X_0 \in A) = \nu(A)$.

Definición 5.3.2 *Un sistema $\mathcal{X} = (\Omega, \mathscr{F}^0_{\infty}, (\mathscr{F}^0_t), (\theta_t), (X_t), (\mathbb{P}^{\nu}))$ es un proceso de Markov si existe una función de transición (P_t) tal que para todo $\nu \in \mathcal{P}(E)$ y para toda $f \in \mathscr{E}^+$ o $b\mathscr{E}$,*

$$\mathbb{E}^{\nu}\left[f(X_{t+s}) | \mathscr{F}_s\right] = P_t(X_s)$$

Ya tenemos un primer ejemplo de un sistema con estas propiedades:

Es claro que la familia de probabilidades $(\mathbb{P}^\nu)$ construidas en el teorema 5.3.4 en el espacio $(E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$ a partir del semigrupo de transición (P_t) nos da el proceso de Markov $\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), (\theta_t), (X_t), (\mathbb{P}^\nu))$ con respecto a (P_t). A menudo diremos simplemente que X es un proceso de Markov con respecto a (P_t) y no se mencionará al sistema de no ser necesario. Nótese como a partir de un proceso Y en un espacio de probabilidad cualquiera que satisface la propiedad de Markov con respecto a (P_t), se puede construir un sistema $\mathcal{X}$ (en el espacio canónico) que es proceso de Markov con respecto al mismo semigrupo (P_t). Si $\nu_o \sim Y_0$, entonces en particular X bajo $\mathbb{P}_{\nu_0}$ es una versión de Y en el espacio canónico, en vista de que las distribuciones f.d. de ambos cumplirán (5.2.14) y por lo tanto coinciden. Se dice que $\mathcal{X}$ es el sistema de Markov asociado al proceso de Markov Y. Más adelante, tendremos otro ejemplo pues construiremos un sistema $\mathcal{X}$ en $D(\mathbb{R}^+, \mathbb{R})$ asociado a un proceso de Lévy.

A continuación y por el resto del capítulo, trabajaremos con un sistema de Markov

$$\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), (\theta_t), (X_t), (\mathbb{P}^\nu))$$

que satisface la propiedad de Markov con respecto a (P_t). Nos referiremos a X simplemente como proceso de Markov con función de transición (P_t). Para cada $x \in E$, si $\nu = \delta_x$ es la masa puntual en x, tenemos la medida de probabilidad $\mathbb{P}^{\delta_x}$ que denotaremos simplemente $\mathbb{P}^x$. Bajo esta medida, $\mathbb{P}^x(X_0 = x) = \delta_x(\{x\}) = 1$ y por lo tanto, bajo $\mathbb{P}^x$, X es un proceso que inicia en $X_0 = x$, $\mathbb{P}^x$ c.s.

Similarmente, si Z es $\mathscr{F}^0_\infty$ medible, denotamos

$$\mathbb{E}^x[Z] = \int_\Omega Z\, d\mathbb{P}^x.$$

Como en cada $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$, X satisface la propiedad de Markov con respecto a $(\mathscr{F}^0_t)$ y (P_t) con ley inicial ν, por el lema 5.2.5 (con $t_0 = 0$) se tiene lo siguiente:

Corolario 5.3.3 *Para cada medida inicial ν, $f_i \in \mathscr{E}^+$ y $0 = t_0 < t_1 < \cdots < t_n$ se satisface*

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{t_1-t_0}(x_0, dx_1) f_1(x_1) \dots \int_E P_{t_n-t_{n-1}}(x_{n-1}, dx_n) f_n(x_n). \tag{5.3.2}$$

En particular, si $\nu = \delta_x$:

$$\begin{aligned}\mathbb{E}^x\left[\prod_{i=0}^n f_i(X_{t_i})\right] &= \int_E \delta_x(dx_0) f_0(x_0) \int_E P_{t_1-t_0}(x_0, dx_1) f_1(x_1) \dots \int_E P_{t_n-t_{n-1}}(x_{n-1}, dx_n) f_n(x_n) \\ &= f_0(x) \int_E P_{t_1-t_0}(x, dx_1) f_1(x_1) \dots \int_E P_{t_n-t_{n-1}}(x_{n-1}, dx_n) f_n(x_n). \end{aligned} \tag{5.3.3}$$

por lo que $x \to \mathbb{E}^x\left[\prod_{i=0}^n f_i(X_{t_i})\right]$ es $\mathscr{E}$-medible.

Una consecuencia útil de esta última igualdad es que

$$P_t f(x) = \mathbb{E}^x[f(X_t)] \tag{5.3.4}$$

y por lo tanto, cambiar de x en $P_t f$ corresponde a cambiar la posición inicial del proceso X. Obsérvese que en particular, para cada $x \in E$:

$$\mathbb{P}^x\left(X_t \in A\right) = \mathbb{P}^x\left(X_0 \in E,\, X_t \in A\right) = \int_E \delta_x(dx_0) \int_A P_t(x_0, dx_1) = P_t(x, A)$$

por lo que $x \to \mathbb{P}^x(X_t \in A)$ es $\mathscr{E}$-medible y $P_t(x, A)$ se lee como la probabilidad de $X_t \in A$ habiendo X empezado en x. De forma más general, si $0 < t_1 < \cdots < t_n$

$$\begin{aligned}\mathbb{P}^x\left(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right) &= \mathbb{P}^x\left(X_{t_0} \in E, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right)\\ &= \int_E \delta_x(dy) P_{t_1-t_0}(y, dx_1) \mathbb{1}_{A_1}(x_1) \ldots P_{t_n-t_{n-1}}(x_{n-1}, dx_n) \mathbb{1}_{A_n}(x_n)\\ &= P_{t_1-t_0}(x, dx_1) \mathbb{1}_{A_1}(x_1) \ldots P_{t_n-t_{n-1}}(x_{n-1}, dx_n) \mathbb{1}_{A_n}(x_n) \quad (5.3.5)\end{aligned}$$

mientras que si incluimos a $t_0 = 0$ en la colección, obtenemos similarmente

$$\begin{aligned}&\mathbb{P}^x\left(X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right)\\ &\qquad = \mathbb{1}_{A_0}(x) P_{t_1-t_0}(x, dx_1) \mathbb{1}_{A_1}(x_1) \ldots P_{t_n-t_{n-1}}(x_{n-1}, dx_n) \mathbb{1}_{A_n}(x_n)\end{aligned}$$

por lo que también, la aplicación $x \to \mathbb{P}^x\left(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right)$ es medible. Si reemplazamos a δ_x por una distribución inicial ν, por (5.3.2) y el caso anterior

$$\begin{aligned}&\mathbb{P}^\nu\left(X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right)\\ &\qquad = \int \nu(dy) \mathbb{1}_{A_0}(y) P_{t_1-t_0}(y, dx_1) \mathbb{1}_{A_1}(x_1) \ldots P_{t_n-t_{n-1}}(x_{n-1}, dx_n) \mathbb{1}_{A_n}(x_n)\\ &\qquad = \int_E \nu(dy) \mathbb{P}^y\left(X_{t_0} \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\right)\\ &\qquad = \int_E \mathbb{E}^y\left[\prod_{i=0}^n \mathbb{1}_{A_i}(X_{t_i})\right] \nu(dy)\end{aligned}$$

es decir,

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \int_E \mathbb{E}^y\left[\prod_{i=0}^n f_i(X_{t_i})\right] \nu(dy) \qquad f_i = \mathbb{1}_{A_i}. \tag{5.3.6}$$

Esta igualdad es un caso particular de la siguiente proposición, que nos permitirá probar la versión más general de la propiedad de Markov:

Proposición 5.3.4 *Sea Z una variable aleatoria $\mathscr{F}^0_\infty$-medible positiva o acotada. Entonces, $x \to \mathbb{E}^x[Z]$ es $\mathscr{E}$-medible y para cada distribución inicial ν*

$$\mathbb{E}^\nu[Z] = \int_E \mathbb{E}^x[Z]\, \nu(dx). \tag{5.3.7}$$

Demostración
Por (5.3.6), ya se tiene el resultado para $Z = \prod_{i=1}^n f_i(X_{t_i})$ con $f_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{E}$:

$$\mathbb{E}^\nu\left[\prod_{i=1}^n \mathbb{1}_{A_i}(X_{t_i})\right] = \int_E \mathbb{E}^y\left[\prod_{i=1}^n \mathbb{1}_{A_i}(X_{t_i})\right] \nu(dy) \qquad A_i \in \mathscr{E} \tag{5.3.8}$$

Considérese la siguiente clase de conjuntos $\Gamma \in \mathscr{F}^0_\infty$:

$$\mathcal{D} = \left\{\Gamma \in \mathscr{F}^0_\infty;\, x \to \mathbb{E}^x[\mathbb{1}_\Gamma] \text{ es medible y } \mathbb{E}^\nu[\mathbb{1}_\Gamma] = \int_E \mathbb{E}^x[\mathbb{1}_\Gamma]\, \nu(dx)\right\}$$

no es difícil comprobar que es un λ-sistema.
Sea $\mathcal{G}$ el π-sistema usual que genera a $\mathscr{F}^0_\infty$

$$\mathcal{G} = \left\{ A = \{X_{t_0} \in A_0,\, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\}\,;\, 0 \leq t_0 < \cdots < t_n,\, A_1 \ldots A_n \in \mathscr{E} \right\}.$$

Como la igualdad (5.3.8) no es más que:

$$\mathbb{E}^\nu \left[\mathbb{1}_{\{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}} \right] = \int_E \mathbb{E}^y \left[\mathbb{1}_{\{X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\}} \right] \nu(dy) \tag{5.3.9}$$

se tiene que $\mathcal{G} \subset \mathcal{D}$. Por el lema de clases monótonas, $\mathbb{E}^\nu [Z] = \int_E \mathbb{E}^x [Z]\, \nu(dx)$ para $Z = \mathbb{1}_D$ con $D \in \mathscr{F}^0_\infty$, por linealidad se tiene para simples y finalmente por los teoremas de convergencia se tiene para toda Z variable aleatoria $\mathscr{F}^0_\infty$-medible acotada o positiva. ■

Observación: Una consecuencia de esto es que para que un evento $A \in \mathscr{F}^0_\infty$ satisfaga $\mathbb{P}^\nu (A) = 0$ para toda $\nu \in \mathcal{P}(E)$, basta con que $\mathbb{P}^x (A) = 0$ para todo $x \in E$. Esto se sigue de remplazar a Z por $\mathbb{1}_A$ en (5.3.7).

Si Z es $\mathscr{F}^0_\infty$-medible, para cada ω, escribimos $Z \circ \theta_t(\omega) = F(\theta_t \omega)$. Nótese como $Z \circ \theta_t$ solo depende de la trayectoria ω_s para valores $s \geq t$. Formalmente, se tiene lo siguiente:

Lema 5.3.5 *Si Z es $\mathscr{F}^0_\infty$-medible, entonces la variable aleatoria $Z \circ \theta_t$ es $\sigma(X_s\,;\, s \geq t)$-medible.*

Demostración
Como θ_t es $\sigma(X_s\,;\, s \geq t)/\mathscr{F}^0_\infty$-medible y Z es $\mathscr{F}^0_\infty/\mathscr{E}$-medible, entonces $F \circ \theta_t$ es $\sigma(X_s\,;\, s \geq t)/\mathscr{E}$-medible. ■

Usando el operador θ_t, podemos formular la propiedad de Markov en su forma más general:

Teorema 5.3.6 Propiedad de Markov (filtración natural).
Si Z es $\mathscr{F}^0_\infty$-medible y positiva (o acotada), para cada $t \geq 0$ y distribución inicial ν,

$$\mathbb{E}^\nu \left[Z \circ \theta_t | \mathscr{F}^0_t \right] = \mathbb{E}^{X_t} [Z] \qquad \mathbb{P}^\nu \ c.s. \tag{5.3.10}$$

El término de la derecha se debe entender como la variable aleatoria obtenida por la composición de los mapeos $\omega \to X_t(\omega)$ y $x \to \mathbb{E}^x [Z]$. Por lo tanto, para cada ω, $\mathbb{E}^{X_t(\omega)} [Z] = \mathbb{E}^{\delta_{X_t(\omega)}} [Z]$. Como ya se mencionó antes, $Z \circ \theta_t$ es función de los estados futuros X_s, $s \geq t$ y su esperanza condicional con respecto al pasado $\mathscr{F}_t$ es una función del estado presente X_t. Esto generaliza las formulaciones anteriores: la igualdad (5.3.10) implica en particular que $\mathbb{E}^\nu \left[Z \circ \theta_t | \mathscr{F}^0_t \right]$ es igual a una variable aleatoria $\sigma(X_t)$-medible y por lo tanto es inmediato que $\mathbb{E}^\nu \left[Z \circ \theta_t | \mathscr{F}^0_t \right] = \mathbb{E}^\nu [Z \circ \theta_t | X_t]$ que es (5.2.7). Si tomamos $Z = \mathbb{1}_{\{X_s \in A\}}$, (5.3.10) es simplemente

$$\mathbb{P}^\nu (X_{s+t} \in A | \mathscr{F}_t) = \mathbb{P}^{X_t} (X_s \in A) = P_s(X_t, A).$$

mientras que si $Z = f(X_s)$, al recordar que $P_s f(x) = \mathbb{E}^x [f(X_s)]$, (5.3.10) es la propiedad de Markov que formulamos inicialmente:

$$\mathbb{E}^\nu \left[f(X_{t+s}) | \mathscr{F}^0_t \right] = \mathbb{E}^{X_t} [f(X_s)] = P_s f(X_t).$$

Demostración
Bastaría probar que para todo $A \in \mathscr{F}_t$,

$$\int_A Z \circ \theta_t \, d\mathbb{P}^\nu = \int_A \mathbb{E}^{X_t}[Z] \, d\mathbb{P}^\nu. \tag{5.3.11}$$

Se sigue el procedimiento habitual: empezamos probándolo para A en el π-sistema

$$\mathcal{G} = \left\{ A = \{X_{t_1} \in A_1,\, X_{t_2} \in A_2, \ldots, X_{t_n} \in A_n\} \,;\, 0 \le t_1 < \cdots < t_n \le t,\, A_1 \ldots A_n \in \mathscr{E} \right\}$$

que genera a $\mathscr{F}_t$ y $Z = \prod_{i=1}^m f_i(X_{s_i})$, $0 \le s_1 < \cdots < s_m < \infty$. Denotaremos por $G(x) = \mathbb{E}^x\left[\prod_{i=1}^m f(X_{s_i})\right]$. En ese caso,

$$\begin{aligned}
\int_A Z \circ \theta_t \, d\mathbb{P}^\nu &= \int_A \prod_{i=1}^m f_i(X_{s_i}) \circ \theta_t \, d\mathbb{P}^\nu \\
&= \mathbb{E}^\nu\left[\prod_{j=1}^n \mathbb{1}_{A_j}(X_{t_j}) \prod_{i=1}^m f_i(X_{s_i+t})\right] \\
&= \nu \mathbb{1}_E P_{t_1} \mathbb{1}_{A_1} \ldots P_{t_n - t_{n-1}} \mathbb{1}_{A_n} P_{s_1+t-t_n} f_1 P_{s_2+t-(s_1+t)} f_2 \ldots P_{s_m+t-(s_{m-1}+t)} f_m \\
&= \nu \mathbb{1}_E P_{t_1} \mathbb{1}_{A_1} \ldots P_{t_n - t_{n-1}} \mathbb{1}_{A_n} P_{s_1+t-t_n} f_1 P_{s_2-s_1} f_2 \ldots P_{s_m-s_{m-1}} f_m \\
&\qquad\qquad\qquad\qquad \text{y por Chapman-Kolmogorov} \\
&= \nu \mathbb{1}_E P_{t_1} \mathbb{1}_{A_1} \ldots P_{t_n - t_{n-1}} \mathbb{1}_{A_n} P_{t-t_n} P_{s_1} f_1 P_{s_2-s_1} f_2 \ldots P_{s_m-s_{m-1}} f_m \\
&= \nu \mathbb{1}_E P_{t_1} \mathbb{1}_{A_1} \ldots P_{t_n - t_{n-1}} \mathbb{1}_{A_n} P_{t-t_n} G \\
&= \mathbb{E}^\nu\left[\prod_{j=1}^n \mathbb{1}_{A_j}(X_{t_j}) \mathbb{E}^{X_t}\left[\prod_{i=1}^m f(X_{s_i})\right]\right] \\
&= \int_A \mathbb{E}^{X_t}[Z] \, d\mathbb{P}^\nu
\end{aligned}$$

Se supuso que $t_n < t$ y por eso se usó Chapman-Kolmogorov; si $t_n = t$ no hace falta. Como se tiene la igualdad (5.3.11) para $A \in \mathcal{G}$, por el lema de clases monótonas se tiene para todo $A \in \mathscr{F}_t = \sigma(\mathcal{G})$.
Tomando $f_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{E}$ la igualdad se vale para $Z = \mathbb{1}_B$ con B en el π-sistema usual que genera a $\mathscr{F}^0_\infty$. Por lema de clases monótonas y los teoremas de convergencia se extiende para toda variable aleatoria $Z \in \mathscr{F}^0_\infty$ positiva o acotada. ■

Hasta este punto, solo hemos logrado construir procesos de Markov con valores en el espacio $E^{\mathbb{R}^+}$ y con respecto a sus filtraciones naturales $(\mathscr{F}^0_t)$ que no tienen por que ser ni continuas por la derecha ni completas. Para garantizar la existencia de versiones con mejores propiedades trayectoriales y poder trabajar bajo las hipótesis habituales, nos restringiremos a una clase más pequeña de procesos de Markov.

Capítulo 6

Procesos de Feller

En este capítulo, nos restringimos a una subclase de procesos de Markov con propiedades adicionales, a saber los procesos de Feller. Empezaremos probando que para todo proceso de Feller, podemos garantizar la existencia de una versión continua por la derecha. Este resultado se conoce como el teorema de regularización de procesos de Feller. Trabajar con esta versión, nos permitirá probar:
1- Para que una filtración generada por un proceso de Feller sea continua por la derecha y completa, basta con agregar los conjuntos nulos. Será necesario probar que la propiedad de Markov y algunos de los resultados probados en el capítulo anterior siguen siendo válidos para esta nuevo filtración.
2- Probaremos una versión más fuerte de la propiedad de Markov, conocida como la propiedad Fuerte de Markov así como algunas de sus consecuencias.

6.1 Procesos de Feller y Teorema de regularización

Referencias: Le Gall [11] sección 6.3, Revuz & Yor [2] capítulo III sección 2 y Zanten [4] sección 3.3.

A lo largo de este capítulo, trabajaremos en E un espacio LCCB [1] aunque una vez más, por simplicidad podemos pensarlo como $\mathbb{R}^d$. Mencionamos a los espacios LCCB pues al probar el teorema de regularización de procesos de Feller habrá que hacer uso de propiedades topológicas de estos. Algunas propiedades de estos espacios se pueden leer en el apéndice así como en el capítulo 0 de Blumenthal & Getoor [17] y de Revuz & Yor [2].

En este capítulo, consideraremos funciones de transición con propiedades adicionales. Denotaremos por C_0 o $C_0(E)$ al espacio de funciones continuas de E en $\mathbb{R}$ tales que

$$\lim_{|x|\to\infty} f(x) = 0.$$

Al ser acotadas, en este espacio se trabajará con la norma uniforme $\|f\|_\infty = \sup_{x\in E} |f(x)|$ y así, $C_0(E)$ es un espacio de Banach separable. Nos referimos al apéndice para más detalle. Nótese que si (P_t) es función de transición, entonces $\|P_t f\|_\infty \le \|f\|_\infty$ si $f \in C_0$ para cualquier t.

[1] Espacio topológico Hausdorff con base numerable y localmente compacto. Todo espacio LCCB es metrizable con una métrica compatible con la topología y se puede probar que bajo esta métrica es Polaco. Estamos entonces considerando espacios con propiedades adicionales a los usados hasta ahora.

Definición 6.1.1 *Una función de transición (P_t) es función de transición de Feller si para cada $f \in C_0$*

$$\lim_{t \downarrow 0} \|P_t f - f\|_\infty = 0 \qquad y \qquad P_t f \in C_0$$

Si un proceso satisface la propiedad de Markov con respecto a una función de transición de Feller, se dice que es un proceso de Feller.

Esta definición es equivalente a otra, que en principio podría parecer más débil, en la cuál se remplaza la convergencia uniforme por la convergencia puntual $P_t f \to f$. No la usaremos así que solo lo mencionamos. Se puede consultar esto en Revuz, Yor [2] capítulo III proposición 2.4. A lo largo de este capítulo, trabajaremos con un sistema de Markov $\mathcal{X}$ con semigrupo de transición de Feller.

Proposición 6.1.2 *Si (P_t) es función de transición de Feller, entonces es fuertemente continua en el siguiente sentido:*

$$\lim_{t \to t'} \|P_t f - P_{t'} f\|_\infty = 0 \tag{6.1.1}$$

para t, t' cualesquiera en $\mathbb{R}^+$.

Demostración
Si $t > t'$:

$$\begin{aligned}\lim_{t \downarrow t'} \|P_t f - P_{t'} f\|_\infty = \lim_{t \downarrow t'} \|P_{t'} P_{t-t'} f - P_{t'} f\|_\infty &= \lim_{t \downarrow t'} \|P_{t'}(P_{t-t'} f - f)\|_\infty \\ &\leq \lim_{t \downarrow t'} \|P_{t-t'} f - f\|_\infty = 0.\end{aligned}$$

Si $t' > t$:

$$\begin{aligned}\lim_{t \uparrow t'} \|P_{t'} f - P_t f\|_\infty = \lim_{t \uparrow t'} \|P_t P_{t'-t} f - P_t f\|_\infty &= \lim_{t \uparrow t'} \|P_t(P_{t'-t} f - f)\|_\infty \\ &\leq \lim_{t \uparrow t'} \|P_{t'-t} f - f\|_\infty = 0.\end{aligned}$$

■

Si $f \in C_0$, entonces $P_t f \in C_0$ por definición de semigrupo de Feller y entonces la composición $P_t P_s f$ está bien definida. Como $P_t f \to f$ uniformemente cuando $t \downarrow 0$, en particular tenemos la convergencia puntual:

$$\lim_{h \downarrow 0} P_{t+h} f(x) = \lim_{h \downarrow 0} P_h(P_t f)(x) = P_t f(x)$$

es decir, para cada $f \in C_0$ y $x \in E$, la aplicación $t \to P_t f(x)$ es continua por la derecha y por lo tanto medible con respecto a t. Lo cual nos permite definir al siguiente operador:

Definición 6.1.3 *Para $\lambda > 0$ definimos al resolvente de orden λ del semigrupo P_t como*

$$R_\lambda f(x) = \int_0^\infty e^{-\lambda t} P_t f(x) dt.$$

Observación: R_λ es un operador lineal sobre C_0 y $R_\lambda(C_0) \subset C_0$: para t fija la aplicación $x \to P_t f(x)$ es acotada al estar en C_0, así que si $x_n \to x$ por el teorema de convergencia dominada

$$\lim_{n\to\infty} R_\lambda f(x_n) = \lim_{n\to\infty} \int_0^\infty e^{-\lambda t} P_t f(x_n) dt = \int_0^\infty e^{-\lambda t} P_t f(x) dt = R_\lambda f(x)$$

y por lo tanto $R_\lambda f$ es continua. Ahora si (z_n) es una sucesión tal que $\|z_n\| \to \infty$, una vez más por el teorema de convergencia dominada y porque $P_t f \in C_0$, tenemos

$$\lim_{n\to\infty} R_\lambda f(z_n) = \lim_{n\to\infty} \int_0^\infty e^{-\lambda t} P_t f(z_n) dt = \int_0^\infty e^{-\lambda t} \lim_{n\to\infty} P_t f(z_n) dt 0$$

por lo que en efecto, $R_\lambda(C_0) \subset C_0$ para cualquier $\lambda > 0$.

Lema 6.1.4 *Si $f \in C_0$, entonces*

$$\lim_{\lambda\to\infty} \|\lambda R_\lambda f - f\|_\infty = 0 \tag{6.1.2}$$

y por lo tanto, si $\mathcal{H}$ es un denso numerable de C_0, la colección

$$\mathcal{H}' = \{nR_n f \,:\, f \in \mathcal{H},\, n \in \mathbb{N}\}$$

vuelve a ser un denso numerable de C_0.

Demostración

Al aplicar el teorema de convergencia dominada y cambio de variable se sigue que

$$\lim_{\lambda\to\infty} \lambda R_\lambda f = \lim_{\lambda\to\infty} \lambda \int_0^\infty e^{-\lambda t} P_t f dt = \lim_{\lambda\to\infty} \int_0^\infty e^{-t} P_{t/\lambda} f dt \to f.$$

Veamos que la convergencia a f se tiene en C_0. Como $\int_0^\infty e^{-t} dt = 1$ y la convergencia del semigrupo $P_t f \to f$ cuando $t \downarrow 0$ es uniforme,

$$|\lambda \int_0^\infty e^{-t} P_{t/\lambda} f(x) dt - f(x)| \leq \int_0^\infty e^{-t} |P_{t/\lambda} f(x) - f(x)| dt \overset{\|\cdot\|_\infty}{\leq} \int_0^\infty e^{-t} \|P_{t/\lambda} f - f\|_\infty dt \to 0$$

donde este último término converge uniformemente en x a 0 por convergencia dominada. La segunda afirmación se debe a que por la primera parte, a todo elemento f del denso $\mathcal{H}$ se le puede aproximar en C_0 por elementos de $\mathcal{H}'$. ■

La resolvente por si misma satisface propiedades de interés, entre ellas:

Proposición 6.1.5 *R_λ satisface la ecuación resolvente:*

$$R_\mu - R_\lambda + (\mu - \lambda) R_\mu R_\lambda = 0$$

para toda μ y $\lambda > 0$. Además su imagen $R_\lambda(C_0)$ no depende de λ y es densa en C_0.

Como nos estamos restringiendo estrictamente a los resultados necesarios para probar el teorema de regularización y no ocuparemos ninguna de estas afirmaciones, omitimos sus pruebas. Nos referimos a LeGall [11] y Yor [2] para un estudio más detallado de la resolvente. Pasaremos a los últimos resultados necesarios para probar la existencia de una versión cadlag de un proceso de Feller. Empecemos notando que los operadores R_λ y P_s conmutan; esto es consecuencia de Fubini y la propiedad de semigrupo:

$$\begin{aligned}
R_\lambda P_s f(x) &= \int_0^\infty e^{-\lambda t} P_t P_s f(x) dt = \int_0^\infty e^{-\lambda t} P_{t+s} f(x) dt = \int_0^\infty e^{-\lambda t} P_s P_t f(x) dt \\
&= \int_0^\infty e^{-\lambda t} \int_E P_t f(y) P_s(x, dy) dt = \int_E \int_0^\infty e^{-\lambda t} P_t f(y) dt P_s(x, dy) = P_s R_\lambda f(x).
\end{aligned}$$

Un elemento clave será el siguiente resultado:

Lema 6.1.6 *Sea X proceso de Feller con semigrupo (P_t) y resolvente (R_λ). Para $\lambda > 0$ y $f \in C_0^+(E)$ cualesquiera, el proceso*

$$e^{-\lambda t} R_\lambda f(X_t), \quad t \geq 0$$

es una super-martingala con respecto a $(\mathscr{F}_t^0)$, la filtración natural de X, para cualquier distribución inicial $\mathbb{P}_\nu$.

Demostración
Sea $s < t$, por propiedad de Markov aplicado a $R_\lambda f \in C_0(E) \subset b\mathscr{E}$,

$$\mathbb{E}^\nu \left[e^{-\lambda t} R_\lambda f(X_t) | \mathscr{F}_t^0 \right] = e^{-\lambda t} \mathbb{E}^\nu \left[R_\lambda f(X_{t-s} \circ \theta_s) | \mathscr{F}_s^0 \right] = e^{-\lambda t} P_{t-s} R_\lambda f(X_s) \tag{6.1.3}$$

así que si logramos probar la desigualdad

$$e^{-\lambda(t-s)} P_{t-s} R_\lambda f(x) \leq R_\lambda f(x), \tag{6.1.4}$$

al multiplicar ambos lados de (6.1.4) por $e^{-\lambda s}$ y remplazar en (6.1.3) concluiríamos la propiedad de supermartingala:

$$\mathbb{E}^\nu \left[e^{-\lambda t} R_\lambda f(X_t) | \mathscr{F}_t^0 \right] = e^{-\lambda t} P_{t-s} R_\lambda f(X_s) \leq e^{-\lambda s} R_\lambda f(X_s).$$

A continuación probamos (6.1.4). Como los operadores P_{t-s} y R_λ conmutan, al aplicar la propiedad de semigrupo y cambio de variable en la integral:

$$\begin{aligned}
e^{-\lambda(t-s)} P_{t-s} R_\lambda f(x) &= e^{-\lambda(t-s)} R_\lambda P_{t-s} f(x) \\
&= e^{-\lambda(t-s)} \int_0^\infty e^{-\lambda u} P_{t-s+u} f(x) du \\
&= e^{-\lambda(t-s)} \int_{t-s}^\infty e^{-\lambda(u+s-t)} P_u f(x) du = \int_{t-s}^\infty e^{-\lambda u} P_u f(x) du
\end{aligned}$$

y por lo tanto, obtenemos

$$\begin{aligned}
e^{-\lambda(t-s)} P_{t-s} R_\lambda f(x) &= \int_{t-s}^\infty e^{-\lambda u} P_u f(x) du \\
&\leq \int_0^\infty e^{-\lambda u} P_u f(x) du = R_\lambda f(x)
\end{aligned}$$

con lo que concluimos la prueba. ∎

Lema 6.1.7 *Sean X y Y dos variables aleatorias definidas en un mismo espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ con valores en un espacio polaco E (métrico completo y separable) con la σ-álgebra de sus borelianos $\mathscr{B}(E)$. Si para toda pareja f, $g : E \to \mathbb{R}$ de funciones continuas y acotadas se satisface que*

$$\mathbb{E}\left[f(X)g(Y)\right] = \mathbb{E}\left[f(X)g(X)\right]$$

entonces $X = Y$ $\mathbb{P}$-c.s.

Demostración

Sean A, $B \in \mathscr{B}(E)$ borelianos cualesquiera. Al aproximar a $\mathbb{1}_A(x)\mathbb{1}_B(y)$ por una sucesión creciente de funciones continuas $f_n(x)g_n(y)$ y aplicar el teorema de convergencia monótona, se tiene que

$$\mathbb{E}\left[\mathbb{1}_A(X)\mathbb{1}_B(Y)\right] = \mathbb{E}\left[\mathbb{1}_A(X)\mathbb{1}_B(X)\right]$$

para todo A, $B \in \mathscr{B}(E)$. Si consideramos entonces al λ-sistema $\mathcal{D} \subset \mathscr{B}(E^2)$

$$\mathcal{D} = \{D \in \mathscr{B}(E^2)\,:\, \mathbb{E}\left[\mathbb{1}_D(X,Y)\right] = \mathbb{E}\left[\mathbb{1}_D(X,X)\right]\}$$

es claro por lo primero que se hizo que el π-sistema de los rectángulos

$$\mathcal{C} = \{A \times B \,:\, A,\, B \in \mathscr{B}(E)\}$$

se queda contenida en $\mathcal{C}$ y genera a $\mathscr{B}(E^2)$. Por el teorema de clases monótonas $\mathcal{D} = \mathscr{B}(E^2)$ y por lo tanto $\mathbb{E}\left[\mathbb{1}_D(X,Y)\right] = \mathbb{E}\left[\mathbb{1}_D(X,X)\right]$ para todo $D \in \mathscr{B}(E^2)$. Si $F(x,y)$ es acotada o positiva y $\mathscr{B}(E)$-medible, al aproximarla por simples y aplicar los teorema de convergencia habituales se tiene que $\mathbb{E}\left[F(X,Y)\right] = \mathbb{E}\left[F(X,X)\right]$. En particular, si consideramos $F(x,y) = d(x,y)$, tenemos que $d(X,Y) = 0$ c.s. y por lo tanto $X = Y$ c.s. ∎

Los detalles de lo que se enuncia a continuación se pueden encontrar en el Apéndice, solo daremos un breve repaso de las nociones que necesitaremos. Consideremos la compactación de Alexandroff del espacio E: el plano extendido al añadir el punto al infinito $E_\Delta = E \cup \{\Delta\}$ con la topología obtenida al agregar a la topología usual de E las vecindades de Δ. Estas se definen como los complementos de compactos en E. Este nuevo espacio topológico resulta ser nuevamente LCCB y compacto. Como todo espacio LCCB tiene una métrica d compatible con la topología, bajo esta E_Δ es un espacio métrico compacto. Obsérvese que toda función $f \in C_0(E)$ se extiende naturalmente a una continua en $C(E_\Delta)$ al definir $f(\Delta) = 0$. La denotaremos igual y teniendo en mente esta convención, consideraremos a continuación funciones en $C_0(E)$ y $C_0^+(E)$ como elementos de $C(E_\Delta)$. Del apéndice, recordemos el

Corolario 6.1.8

(i) Si $\mathcal{H}$ es un subconjunto denso numerable de $C_0^+(E)$, entonces sus extensiones separan puntos de E_Δ, es decir: para todo x, $y \in E_\Delta$ existe $f \in \mathcal{H}$ tal que $f(x) \neq f(y)$.
(ii) Además, existe $x \in E_\Delta$ tal que $x_n \to x$ en E_Δ si y solo si $\lim_{n\to\infty} f(x_n)$ existe para toda $f \in \mathcal{H}$.

Nótese que si X es un proceso con valores en E, podemos considerarlo como un proceso con valore en E_Δ sin mayor problema. Finalmente, llegamos al

Teorema 6.1.9 Regularización de procesos de Feller
Un proceso de Feller X con distribución inicial ν definido en el espacio de probabilidad canónico $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu) = (E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+}, \mathbb{P}^\nu)$ admite una modificación cadlag.

Si queremos trabajar con la versión cadlag del sistema $\mathcal{X}$ asociado a X, como tenemos una modificación (en principio distinta) *para cada* $\nu \in \mathcal{P}(E)$, se necesita un paso adicional. Lo dejamos como observación al final de la prueba.

La idea de la prueba consiste en verificar que, al componer al proceso de Feller X con elementos h de una familia de funciones adecuada $\mathcal{H}$ (a saber, las estudiadas en el lema 6.1.6), podemos probar la existencia de límites por la izquierda y derecha de $h(X)$ (para lograrlo usaremos el teorema de regularización de martingalas). Al ser la clase $\mathcal{H}$ lo suficientemente rica (separa puntos de E_Δ), podremos probar la existencia de los límites laterales de las trayectorias de X en E_Δ y esto nos permitirá definir una versión cadlag.

Demostración
Por claridad, dividimos la demostración en 4 partes.

Paso 1: existe una modificación continua por la derecha $\tilde{X}$ de X con trayectorias en E_Δ
Sea $\mathcal{H}'$ subconjunto denso numerable de $C_0^+(E)$. Por el lema 6.1.4,

$$\mathcal{H} = \{nR_n f \,:\, f \in \mathcal{H}',\, n \in \mathbb{N}\}$$

vuelve a ser denso y numerable en $C_0^+(E)$ y entonces por el corolario 6.1.8, $\mathcal{H}$ separa puntos de E_Δ. Por otro lado, al aplicar el lema 6.1.6 tenemos que para cada $nR_nf \in \mathcal{H}$, el proceso

$$e^{-nt} R_n f(X_t) \quad t \geq 0$$

es una supermartingala real bajo $\mathbb{P}^\nu$ y por lo tanto, por la proposición 9.3.4 existe un $\Omega(n,f)$ con $\mathbb{P}^\nu(\Omega(n,f)) = 1$ para el cual el límite

$$\lim_{\substack{r\downarrow t\\ r\in\mathbb{Q}}} nR_n f(X_t) \tag{6.1.5}$$

existe para toda t (omitir e^{-at} y multiplicar por n no surte ningún efecto). Como $\mathcal{H}'$ es numerable, entonces

$$\Omega' = \bigcap_{n\in\mathbb{N}, f\in\mathcal{H}'} \Omega(n,f)$$

satisface $\mathbb{P}^\nu(\Omega') = 1$ y en Ω', estos límites existen para todo $t \in \mathbb{R}^+$ y $nR_nf \in \mathcal{H}$. Entonces, tenemos que al restringirnos a Ω', el límite (6.1.5) existe para toda $nR_nf \in \mathcal{H}$ que recordemos es un conjunto que separa puntos de E_Δ. Así que una vez más por el corolario 6.1.8, el límite

$$\lim_{\substack{r\downarrow t\\ r\in\mathbb{Q}}} X_t$$

existe en E_Δ para toda $t \in \mathbb{R}^+$ en Ω'. Nótese que este límite podría ser Δ. Fijamos $x_0 \in E$ y definimos entonces al proceso $\tilde{X}$ como

$$\tilde{X}_t(\omega) = \lim_{\substack{r\downarrow t\\ r\in\mathbb{Q}}} X_r(\omega) \quad \text{si } \omega \in \Omega'$$

y $\tilde{X}_t(\omega) = x_0$ si $\omega \in \Omega \setminus \Omega'$. Veamos ahora que $\tilde{X}$ es versión de X; en efecto, si f y g son dos funciones continuas y acotadas en $C(E_\Delta)$, tenemos al aplicar la propiedad de Markov que

$$\begin{aligned}\mathbb{E}^\nu\left[g(X_t)f(\tilde{X}_t)\right] &= \lim_{\substack{r\downarrow t\\ r\in\mathbb{Q}}} \mathbb{E}^\nu\left[g(X_t)f(X_r)\right]\\ &= \lim_{\substack{r\downarrow t\\ r\in\mathbb{Q}}} \mathbb{E}^\nu\left[g(X_t)P_{r-t}f(X_t)\right] = \mathbb{E}^\nu\left[g(X_t)f(X_t)\right]\end{aligned}$$

Por convergencia uniforme de $P_{r-t}f$ a f cuando $r \downarrow t$. Al recordar que todo espacio $LCCB$ es polaco, por el lema 6.1.7 concluimos que $\mathbb{P}^\nu\left(X_t = \tilde{X}_t\right) = 1$ por lo que $\tilde{X}$ es una versión de X. Obsérvese que P_t se extiende naturalmente como semigrupo con valores en E_Δ al definir $P_t(x, \Delta) = 0$.

Paso 2: $\tilde{X}$ es Markov con respecto a (P_t):

Al recordar el lema 5.2.5, para probar que X es Markov con semigrupo (P_t) con respecto a *su filtración* natural que denotaremos $(\tilde{\mathscr{F}}_t^0)$ bastaría probar que para $f_i \in \mathscr{E}_\Delta^+$ se satisface

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(\tilde{X}_{t_i})\right] = \int_{E_\Delta} \nu(dx_0)f_0(x_0)\int_{E_\Delta} P_{t_1-t_0}(x_0, dx_1)f_1(x_1) \\ \ldots\int_{E_\Delta} P_{t_n-t_{n-1}}(x_{n-1}, dx_n)f_n(x_n). \quad (6.1.6)$$

Toda función $f_i \in \mathscr{E}_\Delta^+$ al restringirla a E es una función en $\mathscr{E}^+$, así que al igual que con el semigrupo no distinguiremos entre ambas. Como X es Markov con respecto a (P_t), satisface

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \int_{E} \nu(dx_0)f_0(x_0)\int_{E} P_{t_1-t_0}(x_0, dx_1)f_1(x_1) \\ \ldots\int_{E} P_{t_n-t_{n-1}}(x_{n-1}, dx_n)f_n(x_n). \quad (6.1.7)$$

Al ser $\tilde{X}$ versión de X, el lado izquierdo de las ecuaciones (6.1.6) y (6.1.7) coincide. Por otro lado, como $P_t(x, \Delta) = 0$, las integrales en (6.1.6) se pueden restringir a E y se sigue que también coinciden los términos a la derecha. Con esto concluimos que X es $(\tilde{\mathscr{F}}_t^0)$ Markov con semigrupo (P_t) y valores en E_Δ.

Paso 3: las trayectorias de $\tilde{X}$ son cadlag en E_Δ c.s.

Como $\tilde{X}$ es Markov con semigrupo (P_t), por el lema 6.1.6 para cada $h \in C_0(E_\Delta)$,

$$Y_t = e^{-\lambda t}R_\lambda f(\tilde{X}_t),\ t \geq 0$$

es una $(\tilde{\mathscr{F}}_t^0)$ supermartingala real, y como $\tilde{X}$ es continuo por la derecha, Y es martingala continua por la derecha. Por lo tanto, por el teorema 9.3.5 Y tiene trayectorias cadlag c.s. y en particular esto es válido para toda $f \in \mathcal{H}'$. Tenemos entonces que las trayectorias de $nR_nf(\tilde{X})$ son cadlag para toda $nR_nf \in \mathcal{H}$, clase que recordemos separa puntos en E_Δ. Como los límites

$$\lim_{s\uparrow t} nR_nf(\tilde{X}_s)$$

existen para toda t, por el corolario 6.1.8, el $\lim_{s\uparrow t}\tilde{X}_s$ existe en E_Δ para toda t y lo denotamos por $\tilde{X}_{t-}$. Nótese que este podría tomar el valor Δ. Obsérvese que como $\mathbb{P}\left(\tilde{X}_t = X_t\right) = 1$, entonces si sabemos que para cada t, $\tilde{X}_t \in E$ c.s. pero esto no garantiza que casi seguramente, toda trayectoria permanezca en E.

Nota: para ilustrar esta última observación, considérese el siguiente ejemplo. Si B es un browniano real, $\mathbb{P}^x(B_t \in \mathbb{R} \setminus 0) = 1$ para todo $t \geq 0$ pero las trayectorias de $(B_t)_{t\geq 0}$ pasan por el origen casi seguramente.

Paso 4: $\tilde{X}$ es proceso con trayectorias cadlag en E

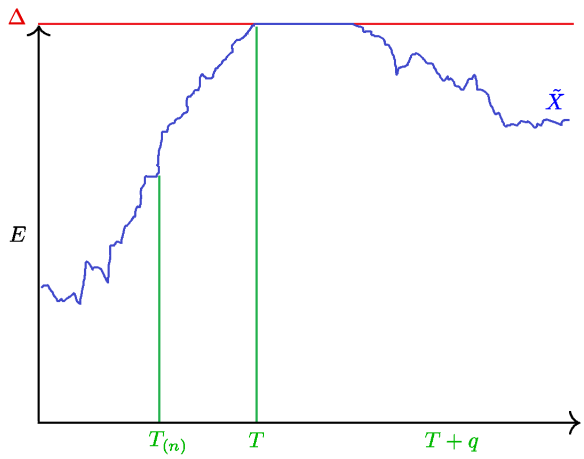
Hasta este punto, tenemos que $t \to \tilde{X}_t$ es cadlag como proceso en E_Δ; solo nos falta verificar que las trayectorias de $\tilde{X}$ permanecen en E y por lo tanto $\tilde{X}$ es cadlag como proceso en E. Consideremos $f \in C_0(E)$ con $f(x) > 0$ para todo $x \in E$ (f solamente se anula en $f(\Delta) = 0$) y a la supermartingala

$$Y_t = e^{-\lambda t} R_\lambda f(\tilde{X}_t)$$

Nótese que $Y = 0$ únicamente cunado $\tilde{X} = \Delta$. Por el teorema 9.3.5, Y vuelve a ser martingala con respecto a la filtración continua por la derecha $(\tilde{\mathscr{F}}^0_+)$ y $\tilde{Y}$ sigue siendo adaptado a esta filtración. Denotemos $(\mathscr{G}_t)$ a la filtración $(\tilde{\mathscr{F}}^0_{t+})$. Entonces, por la proposición 5.1.11

$$T_{(n)} = \inf\{t \geq 0 \,:\, Y_t < 1/n\}$$

es un $(\mathscr{G}_t)$ tiempo de paro, al igual que $\lim_{n\to\infty} T_{(n)} \uparrow T$ por 9.2.6. Terminaríamos la prueba de verificar que $\mathbb{P}^\nu(T < \infty) = 0$; nótese que en $\{T < \infty\}$ $\tilde{X}$ explota en tiempo finito.



Es claro que si $t \in [0, T_{(n)})$, entonces tanto $\tilde{X}_t$ como $\tilde{X}_{t-}$ están en E. Para cada $q \in \mathbb{Q}^+$, tenemos que $T_{(n)} \leq T + q$ y como $(\mathscr{G}_t)$ es continua por la derecha, por el teorema 9.3.13 del apéndice:

$$\mathbb{E}^\nu\left[Y_{T+q} | \mathscr{G}_{T_{(n)}}\right] \leq Y_{T_{(n)}}$$

obtenemos que

$$\mathbb{E}^\nu\left[Y_{T+q}\mathbb{1}_{\{T<\infty\}}\right] \leq \mathbb{E}^\nu\left[Y_{T+q}\mathbb{1}_{\{T_{(n)}<\infty\}}\right] \leq \mathbb{E}^\nu\left[Y_{T_{(n)}}\mathbb{1}_{\{T_{(n)}<\infty\}}\right] \leq \frac{1}{n}$$

pues $\{T_{(n)} < \infty\} \subset \{T < \infty\}$ y $\{T_{(n)} < \infty\} \in \mathscr{G}_{T_{(n)}}$. Como el término de la izquierda no depende de n y Y es no negativa, $\mathbb{E}^\nu\left[Y_{T+q}\mathbb{1}_{\{T<\infty\}}\right] = 0$ y entonces $Y_{T+q} = 0$ c.s. para todo $q \in \mathbb{Q}^+$. Por continuidad por la derecha, obtenemos que $Y = 0$ para todo $t \in [T, \infty)$ en $\{T < \infty\}$. Esto nos dice que si $\tilde{X}$ explota, entonces no vuelve a regresar a E (y la trayectoria en la imagen anterior no puede ocurrir). Por otro lado, como $\tilde{X}_t \in E$ c.s. al ser versión de X, para cada $n \in \mathbb{N}$ tenemos que

$$Y_n = e^{-\lambda n} R_\lambda f(\tilde{X}_n) > 0 \quad \text{c.s.}$$

con lo que concluimos que $\mathbb{P}^\nu (T < \infty) = 0$ y las trayectorias de $\tilde{X}$ se quedan en E c.s.

Con esto, finalizamos la prueba: $\tilde{X}$ es modificación de X con trayectorias cadlag en E y es proceso de Feller con respecto al semigrupo (P_t). ■

Si el proceso inicial Y corresponde a un sistema de Markov $\tilde{X} = (\tilde{\Omega}, \tilde{\mathscr{F}}^0_\infty, (\tilde{\mathscr{F}}^0_t), Y, (\theta_t), (\tilde{\mathbb{P}}^\nu))$, hay que realizar un paso adicional: al aplicar este procedimiento a Y en cada espacio $(\tilde{\Omega}, \tilde{\mathscr{F}}^0_\infty, \tilde{\mathbb{P}}^\nu)$, obtenemos una versión $\tilde{Y}^\nu$ de Y, que es proceso de Markov con semigrupo (P_t), distribución inicial $\mathbb{P}^\nu (Y_0 \in A) = \nu(A)$ y trayectorias cadlag. Si $\tilde{\mathbb{P}}^\nu_{\tilde{X}^\nu}$ es la ley de $\tilde{Y}$ en $(D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+})$, al considerar un nuevo espacio de probabilidad en el espacio canónico,

$$(\Omega, \mathscr{F}, \mathbb{P}^\nu) = (D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+}, \mathbb{P}^\nu_{\tilde{X}^\nu})$$

no es difícil verificar que el proceso coordenado $Y(\omega) = \omega$ es proceso de Markov con distribución inicial ν y semigrupo (P_t). Como algo similar se probará en el capítulo 7, omitimos los detalles. Así, $\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, X, (\mathscr{F}^0_t), (\theta_t), (\mathbb{P}^\nu))$ es la versión cadlag del sistema inicial. Trabajaremos con esta por el resto del capítulo.

6.2 Propiedad de Markov bajo las hipótesis habituales

Referencia: Revuz & Yor [2] capítulo III sección 2 y Zanten [4] sección 3.3.

De aquí en adelante, X es la versión cadlag de un proceso de Feller con semigrupo de transición (P_t). Al aplicar un procedimiento similar al que se seguirá para los procesos de Lévy en el próximo capítulo, podemos pensar que este está definido en el espacio canónico $\Omega = D(\mathbb{R}^+, E)$ con la σ-álgebra de los cilindros $\mathscr{F}^0_\infty = \mathscr{E}_D^{\mathbb{R}^+}$, y que X es el proceso coordenado usual. Seguiremos denotando por $(\mathscr{F}^0_t)$ a su filtración natural. Por ahora, hemos trabajado con la filtración natural $(\mathscr{F}^0_t)$ de X. Esta filtración no es ni continua por la derecha ni completa así que en esta sección abordamos este problema. Construiremos una nueva filtración a partir de la original que sea completa y continua por la derecha y bajo la cual X siga siendo proceso de Markov.

Para cada medida de probabilidad ν en E, tenemos un espacio de probabilidad filtrado $(\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t)_{t\geq 0}, \mathbb{P}^\nu)$; si N^ν son los subconjuntos de $\mathbb{P}^\nu$-nulos de $\mathscr{F}^0_\infty$, denotamos por $\mathscr{F}^\nu_\infty$ a su $\mathbb{P}^\nu$ compleción:

$$\mathscr{F}^\nu_\infty = \sigma(\mathscr{F}^0_\infty \cup N^\nu).$$

Recordemos que $\mathbb{P}^\nu$ tiene una única extensión a esta nueva σ-álgebra así que la denotamos igual. Ahora completamos la filtración: denotaremos por $(\mathscr{F}^\nu_t)_{t\geq 0}$ a la filtración generada al agregar los conjuntos de N^ν a la filtración original:

$$\mathscr{F}^\nu_t = \sigma(\mathscr{F}^0_t \cup N^\nu).$$

De solo querer trabajar con una distribución inicial ν (lo cual sucede con frecuencia), se trabaja en el espacio $(\Omega, \mathscr{F}^\nu_\infty, \mathbb{P}^\nu)$. Sin embargo, en nuestro caso nos interesa trabajar con cualquier

$\nu \in \mathcal{P}(E)$ y el problema es que tenemos una compleción $\mathscr{F}_\infty^\nu$ diferente para cada $\nu \in \mathcal{P}(E)$. Si Z es $\mathscr{F}_\infty^\nu$-medible, no tiene por que ser $\mathscr{F}_\infty^\rho$-medible si $\nu \neq \rho$. Queremos restringirnos a una clase de variables aleatorias medibles "universales", para poder considerar su integral bajo cualquier $\mathbb{P}^\nu$. Por eso, definimos

$$\mathscr{F}_t = \bigcap_{\nu \in \mathcal{P}(E)} \mathscr{F}_t^\nu \quad \text{y} \quad \mathscr{F}_\infty = \bigcap_{\nu \in \mathcal{P}(E)} \mathscr{F}_\infty^\nu.$$

Obsérvese que $\mathscr{F}_\infty^0 \subset \mathscr{F}_\infty$ y que $\mathscr{F}_t^0 \subset \mathscr{F}_t$ por lo que estas intersecciones no son triviales. Además, si $Z \in \mathscr{F}_\infty$, entonces $Z \in \mathscr{F}_\infty^\nu$ para cualquier $\nu \in \mathcal{P}(E)$ y por lo tanto podemos considerar su integral bajo cualquier $\mathbb{P}^\nu$.

Observación: como estamos trabajando con una familia de medidas de probabilidad $(\mathbb{P}^\nu)_\nu$ en $(\Omega, \mathscr{F}_\infty^0)$ y no con una sola, se establecen las siguientes convenciones (Revuz,Yor capítulo 1, definición 4.13 y capítulo 3 página 83):
1) Diremos que un evento $A \in \mathscr{F}_\infty$ es nulo si $\mathbb{P}^\nu(A) = 0$ para toda $\nu \in \mathcal{P}(E)$ y diremos que se vale casi seguramente si $\mathbb{P}^\nu(A) = 1$ para cada $\nu \in \mathcal{P}(E)$.
2) Una filtración $(\mathscr{G}_t)$ en $(\Omega, \mathscr{F}_\infty^0, (\mathbb{P}^\nu))$ es completa si $\mathscr{G}_0$ y por lo tanto cada $\mathscr{G}_t$ contiene a los conjuntos nulos de $\cap_{\nu \in \mathcal{P}(E)} \mathscr{F}_\infty^\nu$.

Es claro que la filtración $(\mathscr{F}_t)$ que construimos satisface esta segunda propiedad. Hasta ahora solo hemos completado las filtraciones $(\mathscr{F}_t^\nu)$ y $(\mathscr{F}_t)$ pero sorprendentemente, esto vuelve tanto a $(\mathscr{F}_t^\nu)$ como a $(\mathscr{F}_t)$ continuas por la derecha:

Teorema 6.2.1 *Las filtraciones $(\mathscr{F}_t^\nu)$ y $(\mathscr{F}_t)$ son continuas por la derecha.*

Demostración
Como

$$\mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s = \bigcap_{s>t} \bigcap_{\nu \in \mathcal{P}(E)} \mathscr{F}_s^\nu = \bigcap_{\nu \in \mathcal{P}(E)} \bigcap_{s>t} \mathscr{F}_s^\nu = \bigcap_{\nu \in \mathcal{P}(E)} \mathscr{F}_{t+}^\nu$$

entonces bastaría probar que $(\mathscr{F}_t^\nu)$ es cad. Trabajaremos en el espacio $(\Omega, \mathscr{F}_\infty^\nu, \mathbb{P}^\nu)$. Primero probaremos que para toda variable aleatoria Z positiva $\mathscr{F}_\infty^0$-medible,

$$\mathbb{E}^\nu\left[Z|\mathscr{F}_t^\nu\right] = \mathbb{E}^\nu\left[Z|\mathscr{F}_{t+}^\nu\right] \qquad \mathbb{P}^\nu \text{ c.s.} \tag{6.2.1}$$

Por los argumentos habituales, basta probarlo para $Z = \prod_{i=1}^n f_i(X_{t_i})$ con $f_i \in C_0$ y $t_1 < \cdots < t_n$. Como $\mathscr{F}_t^\nu$ y $\mathscr{F}_t^0$ difieren solo en conjuntos $\mathbb{P}^\nu$nulos,

$$\mathbb{E}^\nu\left[Z|\mathscr{F}_t^0\right] = \mathbb{E}^\nu\left[Z|\mathscr{F}_t^\nu\right] \qquad \mathbb{P}^\nu \text{ c.s.} \tag{6.2.2}$$

para cualquier $t \geq 0$ (dejamos la explicación al final de la prueba). Ahora tomando h suficientemente pequeño de modo que $t_{k-1} \leq t \leq t+h < t_k$ y usando propiedad de Markov:

$$\begin{aligned}
\mathbb{E}^\nu\left[Z|\mathscr{F}^\nu_{t+h}\right] &= \mathbb{E}^\nu\left[\prod_{i=1}^n f_i(X_{t_i})|\mathscr{F}^0_{t+h}\right] \\
&= \prod_{i=1}^{k-1} f_i(X_{t_i})\mathbb{E}^\nu\left[\prod_{i=k}^n f_i(X_{t_i})\middle|\mathscr{F}^0_{t+h}\right] \\
&= \prod_{i=1}^{k-1} f_i(X_{t_i})\mathbb{E}^\nu\left[\prod_{i=k}^{n-1} f_i(X_{t_i})\mathbb{E}^\nu\left[f_n(X_{t_n})|\mathscr{F}^0_{t_{n-1}}\right]\middle|\mathscr{F}^0_{t+h}\right] \\
&= \prod_{i=1}^{k-1} f_i(X_{t_i})\mathbb{E}^\nu\left[\prod_{i=k}^{n-1} f_i(X_{t_i})P_{t_n-t_{n-1}}f_n(X_{t_{n-1}})\middle|\mathscr{F}^0_{t+h}\right] \\
&\qquad\qquad\qquad\qquad\text{y repitiendo el procedimiento llegamos a} \\
&= \prod_{i=1}^{k-1} f_i(X_{t_i})\mathbb{E}^\nu\left[f_k(X_{t_k})P_{t_{k+1}-t_k}\dots P_{t_n-t_{n-1}}f_n(X_{t_k})\middle|\mathscr{F}^0_{t+h}\right] \\
&= \prod_{i=1}^{k-1} f_i(X_{t_i})\cdot P_{t_k-(t+h)}f_kP_{t_{k+1}-t_k}\dots P_{t_n-t_{n-1}}f_n(X_{t+h}) \\
&= g_h(X_{t+h})\prod_{i=1}^{k-1} f_i(X_{t_i})
\end{aligned}$$

donde $g_h(x) = P_{t_k-(t+h)}f_kP_{t_{k+1}-t_k}\dots P_{t_n-t_{n-1}}f_n(x)$. Por continuidad por la derecha, $\lim_{h\downarrow 0} X_{t+h} \to X_t$ mientras que por continuidad fuerte del semigrupo, $\lim_{h\downarrow 0} g_h(x) \to g(x)$ con:

$$g(x) = P_{t_k-t}f_kP_{t_{k+1}-t_k}\dots P_{t_n-t_{n-1}}f_n(x).$$

Así, por el teorema de convergencia de Lévy hacia arriba, concluimos que

$$\begin{aligned}
\mathbb{E}^\nu\left[Z|\mathscr{F}^\nu_{t+}\right] = \lim_{h\downarrow 0}\mathbb{E}^\nu\left[Z|\mathscr{F}^\nu_{t+h}\right] = \lim_{h\downarrow 0} g_h(X_{t+h})\prod_{i=1}^{k-1} f_i(X_{t_i}) &= g(X_t)\prod_{i=1}^{k-1} f_i(X_{t_i}) \\
&= \mathbb{E}^\nu\left[Z|\mathscr{F}^\nu_t\right] \quad \mathbb{P}^\nu \text{ c.s.}
\end{aligned}$$

por lo que se tiene (6.2.1).

Ahora, veamos que en esto implica que $\mathscr{F}^\nu_{t+} = \mathscr{F}^\nu_t$. Sea $B \in \mathscr{F}^\nu_{t+}$; como $B \in \mathscr{F}^\nu_\infty$, existen $A \in \mathscr{F}^0_\infty$ y $N \in N^\nu$ tales que $B = A \cup N$. Por lo tanto $\mathbb{1}_B = \mathbb{1}_A$, $\mathbb{P}^\nu$ c.s. donde $\mathbb{1}_A$ es $\mathscr{F}^0_\infty$-medible. Por lo anterior, $\mathbb{P}^\nu$ c.s. se tiene que:

$$\mathbb{1}_B = \mathbb{1}_A = \mathbb{E}^\nu\left[\mathbb{1}_A|\mathscr{F}^\nu_{t+}\right] = \mathbb{E}^\nu\left[\mathbb{1}_A|\mathscr{F}^\nu_t\right] = \mathbb{E}^\nu\left[\mathbb{1}_B|\mathscr{F}^\nu_t\right] \quad \mathbb{P}^\nu \text{ c.s.}$$

Como $\mathscr{F}^\nu_t$ es $\mathbb{P}^\nu$ completo, $\mathbb{1}_B$ también es $\mathscr{F}^\nu_t$ medible y en particular $B \in \mathscr{F}^\nu_t$ (recuérdese que si $(\Omega, \mathscr{G}, \mu)$ es espacio de medida completo, f es función $\mathscr{G}$-medible y $f = g$ c.s., entonces g también es $\mathscr{G}$-medible) con lo que concluye la prueba.

Para ver (6.2.2), nótese que $\mathscr{F}^\nu_t = \{A \cup N\,;\, A \in \mathscr{F}^0_t \text{ y } N \in N^\nu\}$. Al integrar ambos términos sobre $A \cup N \in \mathscr{F}^\nu_t$, se tiene por un lado que:

$$\int_{A\cup N}\mathbb{E}^\nu\left[Z|\mathscr{F}^0_t\right]d\mathbb{P}^\nu = \int_A \mathbb{E}^\nu\left[Z|\mathscr{F}^0_t\right]d\mathbb{P}^\nu + \int_{N\setminus A}\mathbb{E}^\nu\left[Z|\mathscr{F}^0_t\right]d\mathbb{P}^\nu = \int_A Z d\mathbb{P}^\nu$$

mientras que

$$\int_{A\cup N} \mathbb{E}^{\nu}\left[Z|\mathscr{F}_t^{\nu}\right] d\mathbb{P}^{\nu} = \int_{A\cup N} Z d\mathbb{P}^{\nu} = \int_A Z d\mathbb{P}^{\nu}.$$

Como tanto $\mathbb{E}^{\nu}\left[Z|\mathscr{F}_t^{0}\right]$ como $\mathbb{E}^{\nu}\left[Z|\mathscr{F}_t^{\nu}\right]$ son $\mathscr{F}_t^{\nu}$-medible, se sigue (6.2.2). ■

A partir de ahora trabajaremos con las σ-álgebras $\mathscr{F}_\infty$, $\mathscr{F}_\infty^{\nu}$ y sus filtraciones completas y continuas por la derecha $(\mathscr{F}_t)$ y $(\mathscr{F}_t^{\nu})$. Entonces, si $Z \in \mathscr{F}_\infty$, denotamos por $\mathbb{E}^{\nu}[Z]$ a la integral de Z con respecto a $\mathbb{P}^{\nu}$, en el espacio de probabilidad $(\Omega, \mathscr{F}_\infty, (\mathscr{F}_t), \mathbb{P}^{\nu})$ para $\nu \in \mathcal{P}(E)$. Como ahora estamos considerando una familia más grande de funciones medibles Z, es importante aclarar si los resultados probados en la sección anterior siguen siendo ciertos y bajo que condiciones. Por ello, será necesario enriquecer la σ-álgebra $\mathscr{E}$. Recordemos que E es un espacio polaco y $\mathscr{E}$ la σ-álgebra de sus borelianos. Para cada $\nu \in \mathcal{P}(E)$, denotamos por $\mathscr{E}^{\nu}$ la ν-compleción de $\mathscr{E}$. Definimos a

$$\mathscr{E}^* = \bigcap_{\nu\in\mathcal{P}(E)} \mathscr{E}^{\nu},$$

llamada la σ-álgebra de los conjuntos universalmente medibles. Similarmente, nótese como $\mathscr{E} \subset \mathscr{E}^*$ por lo que si $f : E \to \mathbb{R}$ es $\mathscr{E}$-medible, entonces también lo es con respecto a $\mathscr{E}^*$. Se sigue valiendo un resultado análogo al teorema 5.3.4:

Proposición 6.2.2 *Si Z es $\mathscr{F}_\infty$-medible y acotada, entonces $x \to \mathbb{E}^x[Z]$ es $\mathscr{E}^*$-medible y*

$$\mathbb{E}^{\nu}[Z] = \int_E \mathbb{E}^x[Z]\,\nu(dx)$$

para cada distribución inicial ν.

Demostración

Sea $\nu \in \mathcal{P}(E)$, como $\mathscr{F}_\infty \subset \mathscr{F}_\infty^{\nu}$, entonces Z es $\mathscr{F}_\infty^{\nu}$-medible. Ahora, por definición de $\mathscr{F}_\infty^{\nu}$, existen Z_1 y Z_2 variables aleatorias $\mathscr{F}_\infty^{0}$-medibles tales que $Z_1 \le Z \le Z_2$ y $\mathbb{E}^{\nu}[Z_2 - Z_1] = 0$, es decir $Z_2 = Z_1$ $\mathbb{P}^{\nu}$ c.s. (si $Z = \mathbb{1}_{AUN}$ es claro tomando $Z_1 = \mathbb{1}_A$ y $Z_2 = \mathbb{1}_{A\cup U}$, $U \in \mathscr{F}_\infty^0$ conjunto $\mathbb{P}^{\nu}$ nulo tal que $N \subset U$. Se sigue para simples y tomando límite para acotadas o positivas). Por monotonía,

$$\mathbb{E}^x[Z_1] \le \mathbb{E}^x[Z] \le \mathbb{E}^x[Z_2].$$

Por el teorema 5.3.4, $x \to \mathbb{E}^x[Z_1]$, $x \to \mathbb{E}^x[Z_2]$ son $\mathscr{E}$-medibles y

$$0 = \mathbb{E}^{\nu}[Z_2 - Z_1] = \int_E \underbrace{\mathbb{E}^x[Z_2] - \mathbb{E}^x[Z_1]}_{\ge 0}\,\nu(dx)$$

por lo que $\mathbb{E}^x[Z_2] = \mathbb{E}^x[Z] = \mathbb{E}^x[Z_1]$ ν c.s. Como $\mathscr{E}^{\nu}$ es ν-completa, $\mathbb{E}^x[Z]$ es $\mathscr{E}^{\nu}$-medible y como esto es válido para cada $\mathscr{E}^{\nu}$, $\mathbb{E}^x[Z]$ es medible con respecto a $\mathscr{E}^*$. Para probar la igualdad, obsérvese que

$$\int_E \mathbb{E}^x[Z_2]\,\nu(dx) = \int_E \mathbb{E}^x[Z]\,\nu(dx) = \int_E \mathbb{E}^x[Z_1]\,\nu(dx)$$

y por el teorema 5.3.4

$$\int_E \mathbb{E}^x[Z_2]\,\nu(dx) = \mathbb{E}^{\nu}[Z_1] \le \mathbb{E}^{\nu}[Z] \le \mathbb{E}^{\nu}[Z_2] = \int_E \mathbb{E}^x[Z_1]\,\nu(dx)$$

por lo que $\mathbb{E}^{\nu}[Z] = \int_E \mathbb{E}^x[Z]\,\nu(dx)$. ■

Lema 6.2.3 *Para toda $t \geq 0$, X_t es una variable aleatoria $\mathscr{F}_t/\mathscr{E}^*$-medible.*

Demostración
Sea $A \in \mathscr{E}^*$; veamos que $X_t^{-1}(A) \in \mathscr{F}_t$.
Fijamos una medida inicial $\nu \in \mathcal{P}(E)$ y consideremos el espacio $(\Omega, \mathscr{F}_t^\nu, \mathbb{P}^\nu)$; denotamos por μ a $\mathbb{P}^\nu_{X_t}$, la ley de X_t en E bajo $\mathbb{P}^\nu$. Como $\mathscr{E}^* = \bigcap_{\rho \in \mathcal{P}(E)} \mathscr{E}^\rho$, en particular $\mathscr{E}^* \subseteq \mathscr{E}^\mu$. Por definición de μ-compleción de $\mathscr{E}$, existen A_1, $A_2 \in \mathscr{E}$ tales que

$$A_1 \subseteq A \subseteq A_2 \quad \text{y} \quad \mu(A_2 \setminus A_1) = 0$$

y por lo tanto, $X_t^{-1}(A_1) \subseteq X_t^{-1}(A) \subseteq X_t^{-1}(A_2)$. Como A_1 y $A_2 \in \mathscr{E}$, entonces $X_t^{-1}(A_1), X_t^{-1}(A_1) \in \mathscr{F}_t^0$. Pero

$$\mathbb{P}^\nu(X_t^{-1}(A_2) \setminus X_t^{-1}(A_1)) = \mathbb{P}^\nu(X_t^{-1}(A_2 \setminus A_1)) = \mu(A_2 \setminus A_1) = 0$$

por lo que, por definición de la $\mathbb{P}^\nu$-compleción $\mathscr{F}_t^\nu$ de $\mathscr{F}_t^0$, se debe cumplir que $X_t^{-1}(A) \in \mathscr{F}_t^\nu$. Como ν es arbitraria, $X_t^{-1}(A) \in \cap_{\rho \in \mathcal{P}(E)} \mathscr{F}_t^\rho = \mathscr{F}_t$. ■

Con esto, podemos probar que X sigue cumpliendo la propiedad de Markov con respecto a la filtración completa y continua por la derecha $(\mathscr{F}_t)$.

Teorema 6.2.4 Propiedad de Markov (filtración aumentada).
Sea Z una variable aleatoria $\mathscr{F}_\infty$-medible no negativa o acotada. Entonces, para cada $t \geq 0$ y distribución inicial ν

$$\mathbb{E}^\nu [Z \circ \theta_t | \mathscr{F}_t] = \mathbb{E}^{X_t} [Z] \quad \mathbb{P}^\nu \ \textit{c.s.}$$

En particular, X sigue siendo proceso de Markov con respecto a $(\mathscr{F}_t)$.

Demostración
Por los dos lemas previos, X_t es $\mathscr{F}_t/\mathscr{E}^*$-medible y $x \to \mathbb{E}^x [Z]$ es $\mathscr{E}^*$-medible por lo que $\mathbb{E}^{X_t} [Z]$ es $\mathscr{F}_t$-medible. Bastaría entonces probar que para todo $A \in \mathscr{F}_t$

$$\int_A Z \circ \theta_t \, d\mathbb{P}^\nu = \int_A \mathbb{E}^{X_t} [Z] \, d\mathbb{P}^\nu \tag{6.2.3}$$

Como en la prueba anterior, denotamos por μ a $\mathbb{P}^\nu_{X_t}$, la ley de X_t en E bajo $\mathbb{P}^\nu$. Como $\mathscr{F}_\infty = \cap_{\nu \in \mathcal{P}(E)} \mathscr{F}_\infty^\nu$, en particular $\mathscr{F}_\infty \subset \mathscr{F}_\infty^\mu$ (la $\mathbb{P}^\mu$-compleción de $\mathscr{F}_\infty^0$). Entonces, existe una variable aleatoria Z', $\mathscr{F}_\infty^0$-medible tal que $\{Z' \neq Z\} \subseteq \Gamma$, con $\Gamma \in \mathscr{F}_\infty^0$ y $\mathbb{P}^\mu(\Gamma) = 0$; en particular $Z = Z' \, \mathbb{P}^\mu$ c.s.

El plan es el siguiente: probar que $Z \circ \theta_t = Z' \circ \theta_t \, \mathbb{P}^\nu$ c.s. y que $\mathbb{E}^{X_t} [Z] = \mathbb{E}^{X_t} [Z'] \, \mathbb{P}^\nu$ c.s.; luego, usando la propiedad de Markov para la filtración natural $(\mathscr{F}_t^0)$ probar (6.2.3) al remplazar Z por Z'.

Nótese que

$$\{Z \circ \theta_t \neq Z' \circ \theta_t\} = \{\omega \, ; \, \theta_t \omega \in \{Z \neq Z'\}\} = \theta_t^{-1}(\{Z \neq Z'\})$$

por lo que $\{Z \circ \theta_t \neq Z' \circ \theta_t\} \subseteq \theta_t^{-1}(\Gamma)$; veamos que este último tiene $\mathbb{P}^\nu$-medida 0.

$$\mathbb{P}^\nu \left(\theta_t^{-1}(\Gamma)\right) = \mathbb{P}^\nu \left(\omega : \theta_t \omega \in \Gamma\right) = \mathbb{E}^\nu \left[\mathbb{1}_\Gamma \circ \theta_t\right] = \mathbb{E}^\nu \left[\mathbb{E}^\nu \left[\mathbb{1}_\Gamma \circ \theta_t | \mathscr{F}_t^0\right]\right]$$

y como $\Gamma \in \mathscr{F}_\infty^0$, por la propiedad de Markov 5.3.6

$$\begin{aligned}\mathbb{E}^\nu\left[\mathbb{E}^\nu\left[\mathbb{1}_\Gamma\circ\theta_t|\mathscr{F}_t^0\right]\right] = \mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[\mathbb{1}_\Gamma\right]\right] &= \int_\Omega \mathbb{E}^x\left[\mathbb{1}_\Gamma\right]\circ X_t\, d\mathbb{P}^\nu \\ &= \int_E \mathbb{P}^x\left(\Gamma\right)\mu(dx) \\ &= \mathbb{P}^\mu\left(\Gamma\right) \\ &= 0\end{aligned}$$

donde se usó el teorema 5.3.4 en la cuarta igualdad. Entonces, $\mathbb{P}^\nu\left(Z\circ\theta_t \neq Z'\circ\theta_t\right) = 0$ por lo que $Z\circ\theta_t = Z'\circ\theta_t$ $\mathbb{P}^\nu$ c.s.

Nótese que para cualquier $\Gamma \in \mathscr{F}_\infty^0$ al usar una vez más 5.3.4, las igualdades se siguen valiendo

$$\mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[\mathbb{1}_\Gamma\right]\right] = \int_\Omega \mathbb{E}^x\left[\mathbb{1}_\Gamma\right]\circ X_t\, d\mathbb{P}^\nu = \int_E \mathbb{P}^x\left(\Gamma\right)\mu(dx) = \mathbb{P}^\mu\left(\Gamma\right)$$

y por lo tanto las medidas en $\mathscr{F}_\infty^0$

$$\Gamma \to \mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[\mathbb{1}_\Gamma\right]\right] \quad \text{y} \quad \Gamma\to\mathbb{E}^\mu\left[\mathbb{1}_\Gamma\right]$$

coinciden. Entonces, la compleción de $\mathscr{F}_\infty^0$ bajo cualquiera de las dos es $\mathscr{F}_\infty^\mu$ y sus extensiones a esta σ-álgebra son la misma, por unicidad de la extensión. Se sigue que si $F \in \mathscr{F}_\infty$, entonces $F \in \mathscr{F}_\infty^\mu$ y por lo tanto su integral con respecto a las dos es la misma, es decir:

$$\mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[F\right]\right] = \mathbb{E}^\mu\left[F\right].$$

De esto, se sigue que $\mathbb{E}^{X_t}\left[Z\right] = \mathbb{E}^{X_t}\left[Z'\right]$ $\mathbb{P}^\nu$ c.s. pues como $\{Z'\neq Z\}\subset\Gamma$ con $\mathbb{P}^\mu\left(\Gamma\right)=0$,

$$\int_\Omega |\mathbb{E}^{X_t}\left[Z\right] - \mathbb{E}^{X_t}\left[Z'\right]|\, d\mathbb{P}^\nu \leq \mathbb{E}^\nu\left[\mathbb{E}^{X_t}\left[|Z-Z'|\right]\right] = \mathbb{E}^\mu\left[|Z-Z'|\right] = 0.$$

Entonces, como Z' es $\mathscr{F}_\infty^0$-medible, por la propiedad de Markov 5.3.6 se tiene la igualdad buscada para todo $A\in\mathscr{F}_\infty^0$:

$$\int_{A'} Z\circ\theta_t\, d\mathbb{P}^\nu = \int_{A'} Z'\circ\theta_t\, d\mathbb{P}^\nu = \int_{A'}\mathbb{E}^{X_t}\left[Z'\right] d\mathbb{P}^\nu = \int_{A'}\mathbb{E}^{X_t}\left[Z\right] d\mathbb{P}^\nu.$$

Finalmente, si $A\in\mathscr{F}_\infty\subset\mathscr{F}_\infty^\nu$, entonces $A = A'\cup N$ con $A'\in\mathscr{F}_\infty^0$ y $N\subset U\in\mathscr{F}_\infty^0, \mathbb{P}^\nu(U)=0$ por lo que también se tiene que

$$\int_A Z\circ\theta_t\, d\mathbb{P}^\nu = \int_A \mathbb{E}^{X_t}\left[Z\right] d\mathbb{P}^\nu.$$

■

6.3 Propiedad fuerte de Markov

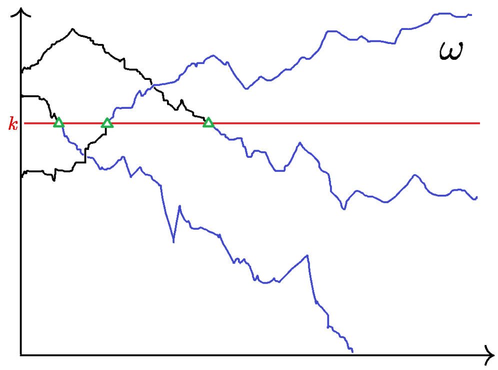
Una versión más fuerte de la propiedad de Markov se sigue valiendo para los procesos de Feller al considerar tiempos de paro además de tiempos deterministas. Seguiremos en el mismo marco que en la sección anterior: $(\Omega,\mathscr{F}_\infty)$ el espacio canónico de las trayectorias cadlag, X la versión canónica de un proceso de Feller, que es un proceso de Markov con respecto a la filtración natural aumentada $(\mathscr{F}_t)$ con semigrupo (P_t). Extendemos la noción del operador traslación a tiempos de paro: sea τ tiempo de paro con respecto a la filtración $(\mathscr{F}_t)$. Para cada ω, θ_τ mapea la trayectoria $s \to \omega_s$ en la trayectoria $s \to \omega_{\tau(\omega)+s}$ por lo que sigue siendo un operador del espacio de trayectorias canónico Ω en si mismo:

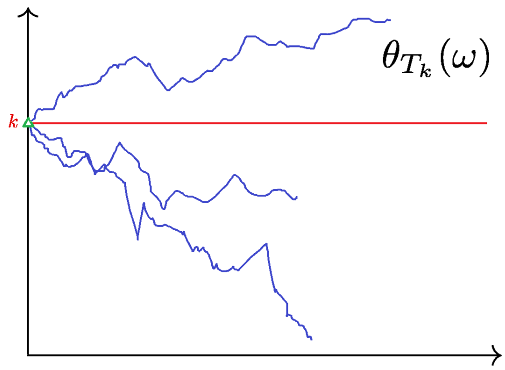
$$(\theta_\tau)(\omega) = \theta_{\tau(\omega)}\omega = \omega_{\tau(\omega)+\bullet}$$

Si τ es determinista $\tau \equiv t$, la definición coincide con la de θ_t. Como X es la identidad en Ω y $X_t(\omega) = \omega_t$ la valuación al tiempo t,

$$(X_t \circ \theta_\tau)(\omega) = X_t \circ \theta_{\tau(\omega)}(\omega) = X_t \circ \omega_{\tau(\omega)+\bullet} = \omega_{\tau(\omega)+t} = X_{\tau(\omega)+t}(\omega).$$

Por ejemplo si $T_k(\omega) = \inf\{t \geq 0 : X_t(\omega) = k\} = \inf\{t \geq 0 : \omega_t = k\}$, la transformación $\theta_{T_k} : \omega \to \theta_{T_k(\omega)}(\omega)$ de los caminos ω se ve como sigue:





Nótese como en este caso, $\theta_{T_k}(\Omega)$ consta únicamente de trayectorias que inician en k.

La medibilidad de θ_τ con respecto a $\mathscr{F}_\infty$ se incluye en la prueba del siguiente teorema:

Teorema 6.3.1 Propiedad Fuerte de Markov
Sea Z una variable aleatoria $\mathscr{F}_\infty$ medible no negativa o acotada. Entonces, para cada tiempo de paro τ con respecto a la filtración $(\mathscr{F}_t)$ y distribución inicial ν, se tiene

$$\mathbb{E}^\nu\left[Z \circ \theta_\tau | \mathscr{F}_\tau\right] = \mathbb{E}^{X_\tau}\left[Z\right] \qquad \mathbb{P}^\nu \text{ c.s. en } \{\tau < \infty\}. \tag{6.3.1}$$

Demostración
Primero, supongamos que τ toma una cantidad numerable de valores en un subconjunto $D \subset \mathbb{R}^+$. Por la propiedad de Markov:

$$\begin{aligned}
\mathbb{1}_{\{\tau<\infty\}}\mathbb{E}^\nu\left[Z \circ \theta_\tau | \mathscr{F}_\tau\right] = \sum_{d\in D} \mathbb{1}_{\{\tau=d\}}\mathbb{E}^\nu\left[Z \circ \theta_\tau | \mathscr{F}_\tau\right] &= \sum_{d\in D} \mathbb{1}_{\{\tau=d\}}\mathbb{E}^\nu\left[Z \circ \theta_d | \mathscr{F}_d\right] \\
&= \sum_{d\in D} \mathbb{1}_{\{\tau=d\}}\mathbb{E}^{X_d}\left[Z\right] \\
&= \mathbb{1}_{\{\tau<\infty\}}\mathbb{E}^{X_\tau}\left[Z\right]
\end{aligned}$$

por lo que se tiene la igualdad. Se dará al final una justificación de la segunda igualdad. Ahora, lo probaremos para $Z \in \mathscr{F}^0_\infty$ y τ arbitrario. Recuérdese que la sucesión de tiempo de paro

$$\tau_k = \frac{[2^k T] + 1}{2^k}$$

decrece a τ, cada uno toma una cantidad a lo más numerable de valores y $\{\tau_k < \infty\} = \{\tau < \infty\}$. Si $Z = \prod_{i=1}^n f(X_{t_i})$, $t_1 < \cdots < t_n$ $f_i \in C_0$ se reduce a probar:

$$\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu \left[\prod_{i=1}^n f(X_{t_i}) \circ \theta_\tau | \mathscr{F}_\tau \right] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\tau} \left[\prod_{i=1}^n f(X_{t_i}) \right]. \tag{6.3.2}$$

Por el caso anterior y (5.3.3), para cada k

$$\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu \left[\prod_{i=1}^n f(X_{t_i}) \circ \theta_{\tau_k} | \mathscr{F}_{\tau_k} \right] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_{\tau_k}} \left[\prod_{i=1}^n f(X_{t_i}) \right] = g(X_{\tau_k})$$

donde $g(x) = P_{t_1} f_1 P_{t_2 - t_1} f_2 \dots P_{t_n - t_{n-1}} f_n(x) = \mathbb{E}^x \left[\prod_i f(X_{t_i})\right]$. Como (P_t) es semigrupo de Feller, $g \in C_0$ (en particular, es continua) y entonces por continuidad por la derecha de X, $g(X_{\tau_k}) \to g(X_\tau)$. Esto es,

$$\lim_{k\to\infty} \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_{\tau_k}} \left[\prod_{i=1}^n f(X_{t_i}) \right] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\tau} \left[\prod_{i=1}^n f(X_{t_i}) \right].$$

Por otro lado, al ser la filtración y X continuos por la derecha:

$$\lim_{k\to\infty} \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu \left[\prod_{i=1}^n f(X_{\tau_k + t_i}) | \mathscr{F}_{\tau_k} \right] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu \left[\prod_{i=1}^n f(X_{\tau + t_i}) | \mathscr{F}_\tau \right]$$

por lo que se tiene (6.3.2). Por el argumento usual de clases monótonas, se vale la igualdad para $Z \in \mathscr{F}^0_\infty$ positiva o acotada:

$$\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu \left[Z \circ \theta_\tau | \mathscr{F}_\tau \right] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\tau} [Z] \qquad \mathbb{P}^\nu \text{ c.s.}$$

Falta entonces probarlo cuando $Z \in \mathscr{F}_\infty$. Obsérvese que en ese caso aún no sabemos si $Z \circ \theta_\tau$ es $\mathscr{F}_\infty$-medible, por lo que no podemos hablar de $\mathbb{E}^\nu \left[Z \circ \theta_\tau | \mathscr{F}_\tau \right]$.

Sea $Z \in \mathscr{F}_\infty$ y supongamos que $\tau < \infty$ c.s.. Fijemos una distribución inicial ν y denotemos por μ a $\mathbb{P}^\nu_{X_\tau}$. Como $\mathscr{F}_\infty \subset \mathscr{F}^\mu_\infty$, en particular $Z \in \mathscr{F}^\mu_\infty$ y por construcción de la compleción existen Z',$Z'' \in \mathscr{F}^0_\infty$ tales que $Z' \le Z \le Z''$ y $Z' = Z''$ $\mathbb{P}^\mu$ c.s. Ahora, como $Z'' - Z' \in \mathscr{F}^0_\infty$, por el caso anterior:

$$\begin{aligned}
\mathbb{E}^\nu \left[Z'' \circ \theta_\tau - Z' \circ \theta_\tau \right] &= \mathbb{E}^\nu \left[\mathbb{E}^\nu \left[Z'' \circ \theta_\tau - Z' \theta_\tau | \mathscr{F}_\tau \right] \right] \\
&= \mathbb{E}^\nu \left[\mathbb{E}^{X_\tau} \left[Z'' - Z' \right] \right] \qquad (6.3.3) \\
&= \int_E \mathbb{E}^x \left[Z'' - Z' \right] \mu(dx) \\
&= \mathbb{E}^\mu \left[Z'' - Z' \right] \\
&= 0 \qquad (6.3.4)
\end{aligned}$$

Entonces $Z' \circ \theta_\tau = Z'' \circ \theta_\tau \, \mathbb{P}^\nu$ c.s. y como $Z' \circ \theta_\tau \leq Z \circ \theta_\tau \leq Z'' \circ \theta_\tau$, se sigue que $Z \circ \theta_\tau$ es $\mathscr{F}_\infty^\nu$ medible por ser la σ-álgebra $\mathbb{P}^\nu$ completa. Como ν es arbitraria, $Z \circ \theta_\tau \in \mathscr{F}_\infty$ por lo que podemos hablar de $\mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau]$. Nótese que entonces $\theta_\tau \in \mathscr{F}_\infty$ ya que podemos considerar en particular a $Z = \mathbb{1}_A$ para cualquier $A \in \mathscr{F}_\infty$. También se deduce de (6.3.3) que $\mathbb{E}^{X_\tau}[Z'] = \mathbb{E}^{X_\tau}[Z] = \mathbb{E}^{X_\tau}[Z''] \; \mathbb{P}^\nu$ c.s.. Pero entonces, como por Markov

$$\underbrace{\mathbb{E}^\nu \left[Z' \circ \theta_\tau | \mathscr{F}_\tau \right]}_{\mathbb{E}^{X_\tau}[Z']} \leq \mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau] \leq \underbrace{\mathbb{E}^\nu \left[Z'' \circ \theta_\tau | \mathscr{F}_\tau \right]}_{\mathbb{E}^{X_\tau}[Z'']}$$

obtenemos que

$$\mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau] = \mathbb{E}^{X_\tau}[Z] \qquad \mathbb{P}^\nu \text{ c.s.}$$

Finalmente, si τ no es finito, definimos el tiempo de paro $\sigma = \tau \mathbb{1}_{\{\tau < \infty\}}$. Como es finito, por el caso anterior $\mathbb{E}^\nu [Z \circ \theta_\sigma | \mathscr{F}_\sigma] = \mathbb{E}^{X_\sigma}[Z]$ y en particular se sigue valiendo en $\{\tau < \infty\}$:

$$\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\sigma | \mathscr{F}_\sigma] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\sigma}[Z] \qquad \mathbb{P}^\nu \text{c.s.}$$

En vista de que $\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\sigma}[Z] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^{X_\tau}[Z]$, bastaría probar que

$$\mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\sigma | \mathscr{F}_\sigma] = \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau].$$

Como $\sigma \leq \tau$, $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$ y entonces la primer esperanza condicional es una v.a. $\mathscr{F}_\tau$-medible y bastaría con probar que integran lo mismo en cualquier conjunto de $\mathscr{F}_\tau$.

Sea $A \in \mathscr{F}_\tau$; por definición de σ-álgebra parada, $\{\tau < \infty\} \in \mathscr{F}_\tau$ y por lo tanto $\{\tau < \infty\} \cap A \in \mathscr{F}_\tau$. Como para $t \in \mathbb{R}^+$ arbitrario, $\{\sigma \leq t\} \cap (\{\tau < \infty\} \cap A) = \{\tau \leq t\} \cap A \in \mathscr{F}_t$ pues $A \in \mathscr{F}_\tau$, se sigue que $\{\tau < \infty\} \cap A \in \mathscr{F}_\sigma$ también. Por lo tanto, como $\sigma = \tau$ en $\{\tau < \infty\}$,

$$\begin{aligned}
\int_A \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau] \, d\mathbb{P}^\nu = \int_{A \cap \{\tau<\infty\}} Z \circ \theta_\tau d\mathbb{P}^\nu &= \int_{A \cap \{\tau<\infty\}} Z \circ \theta_\sigma d\mathbb{P}^\nu \\
&= \int_{A \cap \{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\sigma | \mathscr{F}_\sigma] \, d\mathbb{P}^\nu \\
&= \int_A \mathbb{1}_{\{\tau<\infty\}} \mathbb{E}^\nu [Z \circ \theta_\sigma | \mathscr{F}_\sigma] \, d\mathbb{P}^\nu
\end{aligned}$$

por lo que podemos concluir la prueba. ■

Observación: por el lema 6.2.2 y la Propiedad fuerte de Markov se tiene la siguiente igualdad:

$$\mathbb{E}^\nu [Z \circ \theta_\tau] = \mathbb{E}^\nu [\mathbb{E}^\nu [Z \circ \theta_\tau | \mathscr{F}_\tau]] = \mathbb{E}^\nu \left[\mathbb{E}^{X_\tau}[Z] \right] = \int_E \mathbb{E}^x [Z] \, \mathbb{P}^\nu_{X_\tau}(dx) = \mathbb{E}^\mu [Z]$$

para $\mu = \mathbb{P}^\nu_{X_\tau}$. Es decir, considerar la esperanza de $Z \circ \theta_\tau$ con distribución inicial $X_0 \sim \nu$ es equivalente a recorrer el tiempo τ unidades y calcular la esperanza bajo la distribución inicial $X_0 \sim \mathbb{P}^\nu_{X_\tau}$.

Una consecuencia interesante de la Propiedad fuerte de Markov es la llamada casi-continuidad por la izquierda de los procesos de Feller, la cual probamos a continuación. Denotamos por X_{t-} al límite por la izquierda $\lim_{s \uparrow t} X_s$ y nótese que este existe por ser X cadlag.

Teorema 6.3.2 Casi-continuidad por la izquierda
Sea X proceso de Feller y (τ_n) sucesión de tiempos de paro que crece a τ. Entonces, para cada $x \in E$

$$\lim_{n\to\infty} X_{\tau_n} = X_\tau \quad \mathbb{P}^x \ c.s.$$

Observaciones:
1) Este resultado es obvio si X es un proceso continuo pero nada evidente si solo es cadlag, ya que en ese caso lo único claro es que $X_{\tau_n} \to X_{\tau-}$. Una consecuencia interesante es que si τ es un tiempo de paro tal que $X_\tau \neq X_{\tau-}$, entonces no pude existir una sucesión de tiempos de paro (τ_n) estrictamente creciente $\tau_n < \tau_{n+1}$ tal que $\tau_n \uparrow \tau$. Se dice en ese caso que τ no es *predecible*. En particular, los tiempos de salto de un proceso de Feller no son predecibles (pensar por ejemplo en el proceso Poisson). Para un tiempo de paro de este tipo, si $\tau_n \uparrow \tau$, para cada ω debe existir $n(\omega)$ tal que $\tau_n(\omega) = \tau(\omega)$ para $n > n(\omega)$.
2) Otra consecuencia es que, al considerar tiempos de paro deterministas, obtenemos que $\lim_{s\uparrow t} X_s = X_t$ $\mathbb{P}^x$, c.s. y por lo tanto

$$D(X) := \{t \geq 0 : \mathbb{P}^x (X_t = X_{t-}) = 1\} = \mathbb{R}^+.$$

Esto se usará cuando abordemos convergencia débil de procesos.

Demostración
Podemos suponer que $\tau_n < \tau$ para toda n ya que de no ser el caso, para $n \geq n(\omega)$, $\tau_n(\omega) = \tau(\omega)$ y trivialmente se tiene que $X_{\tau_n}(\omega) \to X_\tau(\omega)$. Es claro entonces que

$$\lim_{n\to\infty} X_{\tau_n} = X_{\tau-}. \tag{6.3.5}$$

por lo que bastaría probar que $X_\tau = X_{\tau-}$ $\mathbb{P}^x$ c.s.
Fijemos $t > 0$ y sean f, g en C_0.
Por un lado, al tomar primero límite cuando $n \uparrow \infty$ y luego cuando $t \downarrow 0$ y aplicar Teorema de convergencia dominada:

$$\lim_{n\to\infty} \mathbb{E}^x \left[f(X_{\tau_n})g(X_{\tau_n+t})\right] = \mathbb{E}^x \left[f(X_{\tau-})g(X_{(\tau+t)-})\right]$$

$$\lim_{t\downarrow 0} \mathbb{E}^x \left[f(X_{\tau-})g(X_{(\tau+t)-})\right] = \mathbb{E}^x \left[f(X_{\tau-})g(X_\tau)\right]$$

La justificación de que $g(X_{(\tau+t)-}) \to g(X_\tau)$ se deja al final. Entonces obtuvimos que

$$\lim_{t\downarrow 0} \lim_{n\to\infty} \mathbb{E}^x \left[f(X_{\tau_n})g(X_{\tau_n+t})\right] = \mathbb{E}^x \left[f(X_{\tau-})g(X_\tau)\right]. \tag{6.3.6}$$

Ahora, por otro lado, por Propiedad fuerte de Markov:

$$\begin{aligned}
\mathbb{E}^x \left[f(X_{\tau_n})g(X_{\tau_n+t})\right] &= \mathbb{E}^x \left[f(X_{\tau_n})\mathbb{E}^x \left[g(X_{\tau_n+t})|\mathscr{F}_{\tau_n}\right]\right] \\
&= \mathbb{E}^x \left[f(X_{\tau_n})\mathbb{E}^{X_{\tau_n}} \left[g(X_t)\right]\right] \\
&= \mathbb{E}^x \left[f(X_{\tau_n})P_t g(X_{\tau_n})\right]
\end{aligned}$$

Notemos que al ser (P_t) semigrupo de Feller, $P_t g(x)$ es continua y entonces al tomar el límite cuando $n \uparrow \infty$ obtenemos

$$\lim_{n\to\infty} \mathbb{E}^x \left[f(X_{\tau_n})g(X_{\tau_n+t})\right] = \mathbb{E}^x \left[f(X_{\tau-})P_t g(X_{\tau-})\right]$$

Recordemos que por ser Feller, $P_t g(x) \to g(x)$ para cualquier $x \in E$. Entonces, para cada ω, $\lim_{t\downarrow 0} P_t g(X_{\tau-}(\omega)) = g(X_{\tau-}(\omega))$ y por lo tanto

$$\lim_{t\downarrow 0} \mathbb{E}^x \left[f(X_{\tau-}) P_t g(X_{\tau-}) \right] = \mathbb{E}^x \left[f(X_{\tau-}) g(X_{\tau-}) \right].$$

Así que por otro lado, obtuvimos

$$\lim_{t\downarrow 0} \lim_{n\to\infty} \mathbb{E}^x \left[f(X_{\tau_n}) g(X_{\tau_n+t}) \right] = \mathbb{E}^x \left[f(X_{\tau-}) g(X_{\tau-}) \right]. \tag{6.3.7}$$

Al juntar lo obtenido en (6.3.6) y (6.3.7) se sigue para f,$g \in C_0$ cualesquiera

$$\mathbb{E}^x \left[f(X_{\tau-}) g(X_\tau) \right] = \mathbb{E}^x \left[f(X_{\tau-}) g(X_{\tau-}) \right].$$

Por el lema 6.1.7 $X_\tau = X_{\tau-}$ $\mathbb{P}^x$ c.s. con lo que concluimos. Finalizamos la prueba argumentando por qué se tiene $g(X_{(\tau+t)-}) \to g(X_\tau)$ puntualmente. Sea (t_n) sucesión de reales tal que $t_n \downarrow 0$, bastaría probar que

$$\lim_{n\to\infty} X_{(\tau+t_n)-} = X_\tau$$

Como X tiene límites por la izquierda, para cada n existe $\tilde{t}_n(\omega) < t_n$ tal que $|X_{\tau+\tilde{t}_n} - X_{(\tau+t_n)-}| < \epsilon$. Obsérvese entonces que $\tilde{t}_n \downarrow 0$.
Por otro lado, por ser X continuo por la derecha para n suficientemente grande, $|X_\tau - X_{\tau+\tilde{t_n}}| < \epsilon$. Con esto concluimos que

$$|X_\tau - X_{(\tau+t_n)-}| \leq |X_\tau - X_{(\tau+\tilde{t_n})}| + |X_{(\tau+\tilde{t_n})} - X_{(\tau+t_n)-}|.$$

■

Referencia: Revuz, Yor [2] p 103-104.
En el siguiente capítulo, trabajaremos con tiempos de paro trasladados: si τ y σ son tiempo de paro, definimos

$$\tau^\sigma := \sigma + \tau \circ \theta_\sigma. \tag{6.3.8}$$

Estos junto con la Propiedad fuerte de Markov forman una herramienta poderosa que permite probar propiedades trayectoriales. Si pensamos a un tiempo de paro como el momento en que ocurre un evento de interés, τ^σ representa la *primera vez que ocurre* τ *después de que ocurrió* σ . Un caso particular que usaremos frecuentemente es cuando T_A y T_B son respectivamente $\inf\{t : X_t \in B\}$ y $\inf\{t : X_t \in A\}$, los tiempo de llegada a dos conjuntos $A, B \in \mathscr{E}$. En ese caso,

$$(T_A)^{T_B} = T_B + T_A \circ \theta_{T_B} = \inf\{t \geq T_B : X_t \in A\}.$$

Antes de probar que esto vuelve a ser un tiempo de paro, haremos algunas observaciones. De la definición (Yor [2] página 104) se puede ver que

$$f(X_\tau) \circ \theta_\sigma = f(X_{\tau^\sigma}) \quad \text{ y } \quad \{\tau^\sigma < \infty\} = \{\tau \circ \sigma < \infty\} \tag{6.3.9}$$

a pesar de que $\tau^\sigma \neq \tau \circ \sigma$. Si $\sigma \equiv t$ es determinista y S es otro tiempo de paro,

$$(\tau^t)^S = S + t \circ \theta_S + \tau \circ \theta_{S+t} = \tau^{S+t} \tag{6.3.10}$$

Si σ no es determinista, lo anterior no es cierto. Para probar que $\sigma + \tau \circ \theta_\sigma$ es en efecto un $\mathscr{F}_t$-tiempo de paro, necesitaremos lo siguiente:

Lema 6.3.3 *Para cualquier $\mathscr{F}_t$-tiempo de paro τ y $t > 0$ se tiene que $\theta_\tau^{-1}(\mathscr{F}_t) \in \mathscr{F}_{\tau+t}$.*

Demostración
Como $\omega \in \theta_\tau^{-1}\{X_t \in A\}$ si y solo si $\theta_\tau\omega \in \{X_t \in A\}$ empecemos observando que

$$\mathbb{1}_{\theta_\tau^{-1}\{X_t\in A\}} = \mathbb{1}_{\{X_t\in A\}} \circ \theta_\tau = \mathbb{1}_A(X_t) \circ \theta_\tau = \mathbb{1}_A(X_{\tau+t}) = \mathbb{1}_{\{X_{\tau+t}\in A\}},$$

es decir, $\theta_\tau^{-1}\{X_t \in A\} = \{X_{\tau+t} \in A\}$.
Ahora, sea $A = \{X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n\}$ con $t_i \leq t$, $A_i \in \mathscr{E}$. Para estos, al usar esta última observación, se cumple que

$$\theta_\tau^{-1}(\{X_{t_1} \in A_1, \cdots, X_{t_n} \in A_n\}) = \{X_{\tau+t_1} \in A_1, \cdots, X_{\tau+t_n} \in A_n\} \in \mathscr{F}_{\tau+t}.$$

Como los conjuntos que se escriben como A forman un π-sistema que genera a $\mathscr{F}_t^0$, se sigue que $\theta_\tau^{-1}(\mathscr{F}_t^0) \in \mathscr{F}_{\tau+t}$.

Ahora sea $\nu \in \mathcal{P}(E)$, $A \in \mathscr{F}_t$ y veamos que $\theta_\tau^{-1}(A) \in \mathscr{F}_{\tau+t}$. Si denotamos $\mu = \mathbb{P}^\nu_{X_\tau}$, en particular, $A \in \mathscr{F}_t^\mu$ y entonces existen A', $A'' \in \mathscr{F}_t^0$ con $A' \subset A \subset A''$ tales que $\mathbb{P}^\mu(A'' \setminus A') = 0$. Como

$$\mathbb{1}_A \circ \theta_\tau = \mathbb{1}_{\theta_\tau^{-1}(A)} \tag{6.3.11}$$

probar $\theta_\tau^{-1}(A) \in \mathscr{F}_{\tau+t}$ equivale a probar que $\mathbb{1}_A \circ \theta_\tau$ es $\mathscr{F}_{\tau+t}$-medible. Al condicionar con respecto a $\mathscr{F}_\tau$ y aplicar Propiedad fuerte de Markov,

$$\mathbb{E}^\nu\left[\mathbb{1}_{A''} \circ \theta_\tau - \mathbb{1}_{A'} \circ \theta_\tau\right] = \mathbb{E}^\nu\left[\mathbb{E}^{X_\tau}\left[\mathbb{1}_{A''\setminus A'}\right]\right] = \int_E \mathbb{E}^x\left[\mathbb{1}_{A''\setminus A'}\right]\mu(dx) = \mathbb{P}^\mu\left(A'' \setminus A'\right) = 0.$$

Entonces, tenemos que

$$\mathbb{1}_{A'} \circ \theta_\tau \leq \mathbb{1}_A \circ \theta_\tau \leq \mathbb{1}_{A''} \circ \theta_\tau$$

con $\mathbb{1}_{A''} \circ \theta_\tau = \mathbb{1}_{A'} \circ \theta_\tau \, \mathbb{P}^\nu$ c.s., ambas $\mathscr{F}^\nu_{\tau+t}$-medibles al estar A',$A'' \in \mathscr{F}_t^0$ por lo que se hizo antes. Al ser completa la sigma-álgebra $\mathscr{F}^\nu_{\tau+t}$, se sigue que $\mathbb{1}_A \circ \theta_\tau$ es $\mathscr{F}^\nu_{\tau+t}$-medible. Como esto se vale para cualquier $\nu \in \mathcal{P}(E)$, no es difícil ver que $\mathbb{1}_A \circ \theta_\tau \in \mathscr{F}_{\tau+t}$. ■

Proposición 6.3.4 *Si σ y τ son tiempos de paro con respecto a $(\mathscr{F}_t)$, entonces $\tau + \sigma \circ \theta_\tau$ también es un tiempo de paro con respecto a la misma filtración.*

Demostración

Por continuidad por la derecha de la filtración, basta probar que $\{\tau + \sigma \circ \theta_\tau < t\} \in \mathscr{F}_t$ para cada t. Descomponemos

$$\{\tau + \sigma \circ \theta_\tau < t\} = \bigcup_{q\in\mathbb{Q}} \{\tau < t - q\} \cap \{\sigma \circ \theta_\tau < q\}$$

y como $\{\sigma \circ \theta_\tau < q\} = \theta_\tau^{-1}(\{\sigma < q\})$ del lema anterior se sigue que $\{\sigma \circ \theta_\tau < q\} \in \mathscr{F}_{\tau+q}$. Pero entonces, por definición de los eventos en la sigma-álgebra parada $\mathscr{F}_{\tau+q}$,

$$\{\tau < t - q\} \cap \{\sigma \circ \theta_\tau < q\} = \{\tau + q < t\} \cap \{\sigma \circ \theta_\tau < q\} \in \mathscr{F}_t$$

con lo que concluimos. ■

Capítulo 7

Procesos de Lévy como procesos de Markov

7.1 Sistema de Markov asociado en $D(\mathbb{R}^+, E)$

En esta sección, nos centraremos en las propiedades elementales de los procesos de Lévy como proceso de Markov y en la construcción de su sistema de Markov asociado $\mathcal{X}$ en el espacio canónico $D(\mathbb{R}^+, E)$, que es el espacio natural en el cual trabajar con un proceso cadlag. Será necesario trabajar con este sistema $\mathcal{X}$ para probar resultados sobre el comportamiento asintótico de los procesos de Lévy, tema que abordamos al final del capítulo. La referencia para este último tema es el texto de Bertoin [9], aunque las herramientas usadas difieren ligeramente: las pruebas que daremos se apoyan fuertemente en lo estudiado en el capítulo anterior y se pueden contrastar con las dadas por Bertoin, que hace uso de la propiedad de Markov de forma más heurística. Por esto nos propusimos darle la mayor formalidad posible a estos argumentos. A lo largo de esta sección, E será el espacio $\mathbb{R}^d$.

El primer paso será probar que cualquier proceso de Lévy Y satisface la propiedad de Markov con respecto a un semigrupo (P_t) que definiremos y que además, resulta ser semigrupo de de Feller. Así, todo proceso de Lévy es un proceso de Markov en el sentido de la definición 5.2.4 con respecto a un semigrupo (P_t) y distribución inicial $\nu = \delta_0$. Después, nos dedicamos a construir un proceso de Lévy y a su proceso de Markov asociado $\mathcal{X}$ en el espacio canónico $\Omega = D(\mathbb{R}^+, E)$. Presentamos dos soluciones a este problema.

Solución 1: esta es directa al aplicar la teoría desarrollada en los capítulos 5 y 6: como Y es un proceso de Markov con semigrupo (P_t), recordemos que, por la discusión que se tuvo en la Sección 5.3 del capítulo 5, el teorema de Kolmogorov nos permite probar la existencia de una colección de medidas $(\mathbb{P}^\nu)$ en $(\tilde{\Omega}, \tilde{\mathscr{F}}^0_\infty) = (E^{\mathbb{R}^+}, \mathscr{E}^{\mathbb{R}^+})$ tales que el sistema $\tilde{\mathcal{X}} = (\tilde{\Omega}, \tilde{\mathscr{F}}^0_\infty, (\tilde{\mathscr{F}}^0_t), \tilde{X}, (\theta_t), (\tilde{\mathbb{P}}^\nu))$ con $\tilde{X}$ el proceso coordenado, es proceso de Markov con respecto al mismo semigrupo (P_t). En particular, $\tilde{X}$ bajo $\mathbb{P}^0$ tiene las distribuciones finito-dimensionales del proceso de Lévy Y, ya que recordemos que estas quedan determinadas por el semigrupo y la distribución inicial $\nu = \delta_0$. Sin embargo, obsérvese que $\tilde{X}$ bajo $\mathbb{P}^0$ no es un proceso de Lévy puesto que sus trayectorias no son cadlag. Al ser (P_t) semigrupo de Feller, el teorema de regularización 6.1.9 (y en particular por la observación que le sigue) nos permite construir un sistema de Markov $\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), X, (\mathbb{P}^\nu), (\theta_t))$ asociado al mismo semigrupo

(P_t) en el espacio D. Así, X tiene trayectorias cadlag bajo cada $\mathbb{P}^\nu$ y entonces $\mathcal{X}$ es *el proceso de Markov asociado* al proceso de Lévy Y. Nótese como bajo $\mathbb{P}^0$, X es un proceso de Lévy y es una versión de Y en el espacio canónico D.

Solución 2: este segundo acercamiento lo proponemos nosotros y no requiere de ninguno de los resultados mencionados en la primer solución. [1] Es más pedestre pero permite familiarizarse con los conceptos, pues construiremos la familia de probabilidades $(\mathbb{P}^\nu)$ explícitamente en el espacio D. Por Lévy-Ito, sabemos ya de la existencia del proceso de Lévy Y, definido en un espacio de probabilidad $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$, así que usaremos esto a nuestro favor. Nuestro objetivo, es construir el proceso de Markov en $\Omega = D(\mathbb{R}^+, E)$

$$\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), X, (\mathbb{P}^\nu), (\theta_t), (P_t))$$

asociado al proceso de Lévy Y. Nótese que para construir $\mathcal{X}$ solo falta definir a la familia $(\mathbb{P}^\nu)$ pues (P_t) es el semigrupo de Y y X será el proceso coordenado. Empezaremos construyendo explícitamente una familia de medidas de probabilidad $(\mathbb{P}^\nu)$ en el espacio canónico $(\Omega, \mathscr{F}^0_\infty) = (D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+})$ donde recordemos del primer capítulo, $\mathscr{E}_D^{\mathbb{R}^+}$ es la sigma-álgebra de los cilindros restringida a D. Cada $\mathbb{P}^\nu$ se construye como la medida inducida en el espacio medible $(D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+})$ por el proceso Y al sumarle una variable aleatoria independiente con distribución ν. Este nuevo proceso se denotará $Z^\nu := Y + \Delta_0$ con $\Delta_0 \sim \nu$. Así, para cada ν tenemos un espacio de probabilidad:

$$(D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+}, \tilde{\mathbb{P}}_{Z^\nu}) = (\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu).$$

Veremos que si X es el mapeo coordenado en este nuevo espacio, el sistema $\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), X, (\theta_t), (\mathbb{P}^\nu))$ es un proceso de Markov con respecto a (P_t). Observemos que en particular, si $\nu = \delta_0$, X bajo $\mathbb{P}^0$ es proceso de Lévy y es una versión de Y. En este caso tenemos obviamente que las trayectorias de X son cadlag.

Ambas soluciones requieren de lo que se probará en la sección 7.1.1. Como la *solución 1* no requiere de mayor explicación, en la primer parte de la sección 7.1.2 nos dedicamos a detallar la *solución 2.* Ambas soluciones se reencuentran en el corolario 7.1.8 y finalizamos la sección 7.1.2. con algunas propiedades del sistema $\mathcal{X}$ construido.

7.1.1 Construcción de la familia de transición de un proceso de Lévy

En todo lo subsecuente, Y será un proceso de Lévy en un espacio de probabilidad $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ con valores en $E = \mathbb{R}^d$. Denotaremos $(\tilde{\mathscr{F}}^0_t)$ a la filtración natural de Y y por μ_t a la ley de Y_t es decir, $\mu_t = \mathbb{P}_{Y_t}$. Si X,Y son dos variables aleatorias independientes con valores en $\mathbb{R}^d$, con ley $\tilde{\mathbb{P}}_X$ y $\tilde{\mathbb{P}}_Y$ respectivamente, la ley de (X+Y) está dada por la convolución [2] $\tilde{\mathbb{P}}_X * \tilde{\mathbb{P}}_Y$ pues

[1] Por esto decidimos incluirlo. Además, fue la primer solución que concebí para resolver el problema al no conocer el teorema de regularización de Feller.

[2] Recordemos que dadas dos medidas de probabilidad μ y ν en $E = \mathbb{R}^d$, se define una nueva medida $\mu * \nu$ llamada convolución de μ y ν como $\mu * \nu(A) = \int_E \int_E \mathbb{1}_A(x+y)\nu(dx)\mu(dy)$.

por propiedades elementales de la esperanza condicional, para toda f acotada

$$\begin{aligned}\tilde{\mathbb{E}}\left[f(X+Y)\right] &= \tilde{\mathbb{E}}\left[\tilde{\mathbb{E}}\left[f(X+Y)|Y\right]\right]\\ &= \int_E \mathbb{E}\left[f(x+Y)\right]\tilde{\mathbb{P}}_X(dx)\\ &= \int_E\int_E f(x+y)\tilde{\mathbb{P}}_X(dx)\tilde{\mathbb{P}}_Y(dy) = \int_E f(z)\tilde{\mathbb{P}}_X * \tilde{\mathbb{P}}_Y(dz).\end{aligned}$$

La segunda igualdad es por el teorema de cambio de variable condicional. Este se usará a menudo y se puede leer en el Apéndice, lema 9.1.2.

Lema 7.1.1 *La familia de distribuciones (μ_t) forma un semigrupo de convolución, es decir:*

$$\mu_{t+s} = \mu_t * \mu_s \qquad\qquad \forall\, s,t > 0.$$

Demostración
Sea $f \in \mathscr{E}$ integrable. Como $Y_{t+s} - Y_t$ tiene distribución μ_s y es independiente de $\sigma(Y_u\,;\, u \leq t)$, entonces $Y_{t+s} = (Y_{t+s} - Y_t) + Y_t$ es una suma de dos variables aleatorias independientes con ley μ_t y μ_s respectivamente. Así, la suma se distribuye $\mu_s * \mu_t$ y se tiene que

$$\int_E f(y)\mu_{t+s}(dy) = \tilde{\mathbb{E}}\left[f(Y_{t+s})\right] = \tilde{\mathbb{E}}\left[f((Y_{t+s} - Y_t) + Y_t)\right] = \int_E f(y)\,\mu_s * \mu_t(dy).$$

■

Para toda $f \in b\mathscr{E}$ o en $\mathscr{E}^+$, definimos la siguiente familia (P_t) de operadores:

$$\begin{aligned}P_t(x,A) &:= \int_E \mathbb{1}_A(x+y)\mu_t(dy) &\qquad (P_t f)(x) &:= \int_E f(x+y)\mu_t(dy)\\ &= \tilde{\mathbb{P}}\left(Y_t + x \in A\right) & &= \tilde{\mathbb{E}}\left[f(x+Y_t)\right].\end{aligned} \tag{7.1.1}$$

Lema 7.1.2 *La colección (P_t) definida como en (7.1.1) satisface la ecuación de Chapman-Kolmogorov y por lo tanto es función de transición (ó semigrupo).*

Demostración
Sean $s, t > 0$

$$\begin{aligned}P_t P_s f = P_t\left(\int_E f(u+y)\mu_s(dy)\right) &= \int_E\int_E f(x+u+y)\mu_t(du)\mu_s(dy)\\ &= \int_E f(x+z)\mu_t * \mu_s(dz)\\ &= \int_E f(x+z)\mu_{t+s}(dz)\\ &= P_{t+s}f.\end{aligned}$$

■

Ahora, veamos que la familia de transición de un proceso de Lévy forma un semigrupo de Feller:

Lema 7.1.3 *La función de transición (P_t) es función de transición de Feller.*

Demostración
Sea $f \in C_0$ y $\epsilon > 0$. Como f es uniformemente continua, existe $\delta > 0$ tal que $|f(y) - f(x)| < \epsilon$ si $|y - x| < \delta$ para todo $x, y \in E$. Como Y es continuo por la derecha, existe $t > 0$ (que depende de ω) tal que $|Y_s(\omega)| < \delta$ si $s \in [0, t]$. Por lo tanto

$$\begin{aligned}
\lim_{t\downarrow 0} |(P_t f)(x) - f(x)| &= \lim_{t\downarrow 0} |\int_E f(x+y)\mu_t(dy) - f(x)| \\
&\leq \lim_{t\downarrow 0} \int_E |f(x+y) - f(x)|\mu_t(dy) \\
&= \lim_{t\downarrow 0} \tilde{\mathbb{E}}\left[|f(x+Y_t) - f(x)|\right] \\
&\leq \lim_{t\downarrow 0} \tilde{\mathbb{E}}\left[\sup_{x\in E} |f(x+Y_t) - f(x)|\right] \\
&< \epsilon
\end{aligned}$$

pues por el teorema de convergencia dominada (f está acotada), y continuidad uniforme:

$$\lim_{t\downarrow 0} \tilde{\mathbb{E}}\left[\sup_{x\in E} |f(x+Y_t) - f(x)|\right] = \tilde{\mathbb{E}}\left[\lim_{t\downarrow 0} \sup_{x\in E} |f(x+Y_t(\omega)) - f(x)|\right] < \epsilon.$$

Así, concluimos que $\|P_t f - f\|_\infty \to 0$. Además,

$$\lim_{u\to x} |P_t f(x) - P_t f(u)| \leq \lim_{u\to x} \int_E |f(x+y) - f(u+y)|\mu_t(dy) = 0$$

por lo que $P_t f$ es continua. Finalmente y una vez más por el teorema de convergencia dominada

$$\lim_{|x|\to\infty} P_t f(x) = \lim_{|x|\to\infty} \int_E f(x+y)\mu_t(dy) = 0$$

con lo que concluimos que $P_t(C_0) \subset C_0$. Así, (P_t) es familia de transición de Feller. ∎

Proposición 7.1.4 *El proceso de Lévy Y en $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ satisface la propiedad de Markov con respecto a su filtración natural $(\tilde{\mathscr{F}}_t)$ y la función de transición (P_t):*

$$\tilde{\mathbb{E}}\left[f(Y_{t+s})|\tilde{\mathscr{F}}_s\right] = P_t f(Y_s) \qquad \tilde{\mathbb{P}}\ c.s. \tag{7.1.2}$$

para todo $t, s > 0$.

Demostración
Sea $f \in b\mathscr{E}$ y $s, t > 0$. Como $Y_{t+s} - Y_s$ es independiente de $\tilde{\mathscr{F}}_s$ y Y_s es $\tilde{\mathscr{F}}_s$-medible,

$$\begin{aligned}
\tilde{\mathbb{E}}\left[f(Y_{t+s})|\tilde{\mathscr{G}}_s\right] = \tilde{\mathbb{E}}\left[f(Y_{t+s} - Y_s + Y_s)|\tilde{\mathscr{F}}_s\right] &= \int_E f(y+Y_s)\mu_t(dy) \\
&= \int_E f(y+x)\mu_t(dy) \circ X_s \\
&= P_t f(Y_s).
\end{aligned}$$

∎

Con esto, concluimos que un proceso de Lévy Y es un proceso de Feller con semigrupo (P_t).

7.1.2 Construcción de la versión canónica del proceso de Markov asociado a un proceso de Lévy en $D(\mathbb{R}^+, E)$

Sea $\nu \in \mathcal{P}(E)$ una distribución arbitraria en E y Δ_0 variable aleatoria independiente de Y con distribución ν. Definimos el siguiente proceso en $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$:

$$Z^\nu = Y + \Delta_0 \qquad\qquad Z_t^\nu = Y_t + \Delta_0 \quad t \geq 0 \tag{7.1.3}$$

Sea $(\mathscr{H}_t^0)$ su filtración natural. Salvo en el caso $\nu = \delta_0$, este proceso no es de Lévy ya que la distribución inicial de Z^ν es $\tilde{\mathbb{P}}\left(Z_0^\nu \in A\right) = \tilde{\mathbb{P}}\left(\Delta_0 \in A\right) = \nu(A)$ (pues $Y_0 = 0$) y recordemos que un proceso de Lévy inicia en 0.

Proposición 7.1.5 *El proceso Z^ν satisface la propiedad de Markov con respecto a $(\mathscr{H}_t^0)$, con la misma función de transición (P_t) que Y y con distribución inicial ν:*

$$\tilde{\mathbb{E}}\left[Z_{t+s}^\nu | \mathscr{H}_s^0\right] = (P_t f)(Z_s^\nu) \quad \tilde{\mathbb{P}} \ c.s.$$

Demostración
Nótese que al ser $Y_{t+s} - Y_s$ independiente de $\tilde{\mathscr{F}}_s^0$ y de $\sigma(\Delta_0)$, por el lemma 1.2.5, $Y_{t+s} - Y_s$ es independiente de

$$\sigma(\Delta_0) \bigvee \tilde{\mathscr{F}}_s^0 = \sigma(\Delta_0\,,\, Y_u\,; u \leq s) \supseteq \sigma(Y_u + \Delta_0\,;\; u \leq s) = \mathscr{H}_s^0 \tag{7.1.4}$$

y como $Y_s + \Delta_0$ es una variable aleatoria $\mathscr{H}_s^0$-medible:

$$\begin{aligned}
\tilde{\mathbb{E}}\left[f(Z_{t+s}^\nu) | \mathscr{H}_s^0\right] = \tilde{\mathbb{E}}\left[f(Y_{t+s} + \Delta_0) | \mathscr{H}_s^0\right] &= \tilde{\mathbb{E}}\left[f(Y_{t+s} + Y_s - Y_s + \Delta_0) | \mathscr{H}_s^0\right] \\
&= \int_E f(y + Z_s^\nu) \mu_t(dy) \qquad (7.1.5) \\
&= (P_t f)(Z_s^\nu)
\end{aligned}$$

pues (7.1.5) es la composición de los mapeos $\omega \to Z_s^\nu(\omega)$ y $x \to \int_E f(y+x)\mu_t(dy)$. En la tercer igualdad se usó el lema 9.1.2. ■

Observación: para cada $t, s > 0$, como

$$Z_{t+s}^\nu - Z_s^\nu = Y_{t+s} + \Delta_0 - (Y_s + \Delta_0) = Y_{t+s} - Y_s \stackrel{\mathscr{L}}{=} Y_t,$$

el proceso (Z_t^ν) satisface que el incremento $Z_{t+s}^\nu - Z_s^\nu$ es independiente de $\mathscr{H}_s^0$ y $Z_{t+s}^\nu - Z_s^\nu$ tiene la misma distribución que Y_s, *no* que Z_s^ν (a diferencia de lo que ocurre con un proceso de Lévy).

Al aplicar el lema 5.2.5, obtenemos:

Corolario 7.1.6 *Para toda colección $0 = t_0 < t_1 < \cdots < t_n \in \mathbb{R}^+$ y $f_i \in b\mathscr{E}$ o $\mathscr{E}^+$*

$$\tilde{\mathbb{E}}\left[\prod_{i=0}^n f_i(Z_{t_i}^\nu)\right] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{t_1 - t_0}(x_0, dx_1) f_1(x_1) \ldots \int_E P_{t_n - t_{n-1}}(x_{n-1}, dx_n) f_n(x_n)$$

Por lo tanto, hemos construido un proceso Z^ν que satisface la propiedad de Markov, continuo por la derecha y con las mismas distribuciones f.d. que Y pero con distribución inicial ν.

Recordemos del primer capítulo que al ser cadlag, Z^ν induce una medida de probabilidad $\tilde{\mathbb{P}}_{Z^\nu}$ en $(D(\mathbb{R}^+, E), \mathscr{E}_D^{\mathbb{R}^+})$: si R es un elemento del π-sistema usual de $\mathscr{E}_D^{\mathbb{R}^+}$

$$R = \{x \in D\,;\,(x_{t_1}, \dots, x_{t_n}) \in A_1 \times \cdots \times A_n\} \quad 0 \le t_1 < \dots t_n \quad A_i \in \mathscr{E}$$

entonces

$$\tilde{\mathbb{P}}_{Z^\nu}(R) = \tilde{\mathbb{P}}\,(Z^\nu \in R) = \tilde{\mathbb{P}}(Z^\nu_{t_1} \in A_1, \dots, Z^\nu_{t_n} \in A_n).$$

Denotemos $\mathbb{P}^\nu = \tilde{\mathbb{P}}_{Z^\nu}$. Para cada $\nu \in \mathcal{P}(E)$, consideremos el espacio de probabilidad $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$ con

$$\Omega = D(\mathbb{R}^+, E) \qquad\qquad \mathscr{F}^0_\infty = \mathscr{E}_D^{\mathbb{R}^+} \qquad\qquad \mathbb{P}^\nu = \tilde{\mathbb{P}}_{Z^\nu}$$

y X el proceso coordenado usual en D, $X(\omega) = \omega$, $X(\omega)(t) = \omega_t$ para cada $\omega \in D$. Sea $\mathscr{F}^0_t = \sigma(X_s\,;\,s \le t)$ su filtración natural y $\mathscr{F}^0_\infty = \sigma(X_s;\, s \ge 0)$.

$$(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}}) \xrightarrow{\;Z^\nu\;} (D, \mathscr{E}_D^{\mathbb{R}^+}, \tilde{\mathbb{P}}_{Z^\nu})$$
$$(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu) \xrightarrow{\;X\;} (D, \mathscr{E}_D^{\mathbb{R}^+})$$

Para cada t, recordemos que μ_t es la distribución de Y_t bajo $\tilde{\mathbb{P}}$ y que P_t es el operador en C_0 definido por

$$P_t f = \int_E f(x+y)\mu_t(dy).$$

Ya probamos que (P_t) es una función de transición.

Proposición 7.1.7 *Para cada ν, el proceso coordenado X en el espacio canónico $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$ es proceso de Markov con respecto a $(\mathscr{F}^0_t)$, función de transición (P_t) y ley inicial ν:*

$$\mathbb{E}^\nu\left[f(X_{t+s})|\mathscr{F}^0_s\right] = P_t f(X_s) \quad \mathbb{P}^\nu c.s.$$

para toda $f \in b\mathscr{E}$ o $\mathscr{E}^+$.

Demostración
Por el lema 5.2.5, basta probar que para todo $0 = t_0 < t_1 < \cdots < t_n \in \mathbb{R}^+$ y $f_i \in b\mathscr{E}$ o $\mathscr{E}^+$

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \int_E \nu(dx_0) f_0(x_0) \int_E P_{t_1 - t_0}(x_0, dx_1) f_1(x_1) \dots \int_E P_{t_n - t_{n-1}}(x_{n-1}, dx_n) f_n(x_n)$$

y por le corolario 7.1.6 es suficiente ver que se da la siguiente igualdad:

$$\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{t_i})\right] = \tilde{\mathbb{E}}\left[\prod_{i=0}^n f_i(Z^\nu_{t_i})\right]. \tag{7.1.6}$$

Veamos que para cada ν, X en $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^\nu)$ y Z^ν en $(\tilde{\Omega}, \mathscr{H}^0_\infty, \tilde{\mathbb{P}})$ tienen las mismas distribuciones finito dimensionales:

$$\begin{aligned}
&\mathbb{P}^\nu\left(X_{t_0} \in A_0, \ldots, X_{t_n} \in A_n\right) \\
&\qquad = \mathbb{P}^\nu(\omega \in D\,;\, X_{t_0}(\omega) \in A_0, \ldots, X_{t_n}(\omega) \in A_n) \\
&\qquad = \mathbb{P}^\nu(\omega \in D\,;\, (\omega_{t_0}, \ldots, \omega_{t_n}) \in A_0 \times \cdots \times A_n) \\
&\qquad = \tilde{\mathbb{P}}_{Z^\nu}(x \in D\,;\, (x_{t_0}, \ldots, x_{t_n}) \in A_0 \times \cdots \times A_n)) \\
&\qquad = \tilde{\mathbb{P}}\left(Z^\nu_{t_0} \in A_0, \ldots Z^\nu_{t_n} \in A_n\right)
\end{aligned}$$

por lo que los vectores $(X_{t_0}, \ldots X_{t_n})$ y $(Z^\nu_{t_0}, \ldots, Z^\nu_{t_n})$ tienen la misma distribución y se tiene (7.1.6). ■

Es claro entonces que el sistema $\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), X, (\mathbb{P}^\nu), (\theta_t))$ es un proceso de Markov en $D(\mathbb{R}^+, E)$ con función de transición (P_t). Hemos entonces construido un ejemplo concreto de un sistema de Markov. Por (7.1.6) obtuvimos que X bajo $\mathbb{P}^\nu$ tiene la misma distribución que Z^ν bajo $\tilde{\mathbb{P}}$. En particular X bajo $\mathbb{P}^0$ tiene la misma ley que Z^{δ_0} bajo $\tilde{\mathbb{P}}$, y este último es justamente Y. Entonces X en $(\Omega, \mathscr{F}^0_\infty, \mathbb{P}^0)$ es una versión de Y en el espacio canónico. [3] Así, estamos dentro del marco de la teoría desarrollada en el capítulo 5 a partir de la definición 5.3.2. En particular, X satisface la propiedad de Markov como se enunció en (5.3.6):

Corolario 7.1.8 *Si Z es $\mathscr{F}^0_\infty$-medible y positiva (o acotada), para cada $t \geq 0$ y distribución inicial ν,*

$$\mathbb{E}^\nu\left[Z \circ \theta_t | \mathscr{F}^0_t\right] = \mathbb{E}^{X_t}[Z] \qquad \mathbb{P}^\nu \text{ c.s.} \tag{7.1.7}$$

Entonces, si $Z = f(X_s)$, tenemos la expresión explícita del término de la derecha en (7.1.7): por (5.3.4) y construcción del semigrupo,

$$\mathbb{E}^x[f(X_t)] = P_t f(x) = \tilde{\mathbb{E}}[f(Y_t + x)] = \mathbb{E}^0[f(X_t + x)] \tag{7.1.8}$$

por lo que $\mathbb{E}^{X_t}[f(X_s)]$ es la composición de $\omega \to X_s(\omega)$ con $x \to \mathbb{E}^0[f(X_t + x)]$.

Para finalizar, haremos algunas observaciones sobre el sistema $\mathcal{X}$ que se obtuvo. Ya notamos que cuando $\nu = \delta_0$, X bajo $\mathbb{P}^0$ es una versión de Y en $D(\mathbb{R}^+, E)$ y por lo tanto es un proceso con incrementos independientes y estacionarios. Si $\nu \neq \delta_0$, X bajo $\mathbb{P}^\nu$ ya no es un proceso de Lévy pues no inicia en 0. Sin embargo, tenemos

Proposición 7.1.9
1) La distribución de $X_{s+t} - X_s$ bajo cada $\mathbb{P}^\nu$ es la misma que la de X_t bajo $\mathbb{P}^0$ y es independiente de $\mathscr{F}^0_s$ para cada $s, t \geq 0$.
2) El proceso X tiene incrementos independientes y estacionarios bajo cada $\mathbb{P}^\nu$. Esto último es que $X_{s+t} - X_s \stackrel{\mathscr{L}}{=} X_t - X_0$ bajo cada $\mathbb{P}^\nu$.

Nótese como debemos restar la condición inicial ahora que X no empieza en 0. Cuando un proceso de Markov satisface 2), se dice que $\mathcal{X}$ es un *PAIS* (processus à acroissements indépendant et stationaires, ver Jacod [26]).

[3] Cabe observar que la independencia y estacionariedad de los incrementos son propiedades que se determinan a partir de las distribuciones finito dimensionales, por lo que en efecto X es un proceso de Lévy.

Demostración
1) Por propiedad de Markov, para cualquier $f \in b\mathscr{E}$

$$\mathbb{E}^x \left[f(X_{s+t} - X_s) | \mathscr{F}_s^0 \right] = \mathbb{E}^x \left[f(X_t - X_0) \circ \theta_s | \mathscr{F}_s^0 \right] = \mathbb{E}^{X_s} \left[f(X_t - X_0) \right].$$

El lado derecho es la composición de los mapeos $\omega \to X_s(\omega)$ con $x \to \mathbb{E}^x [f(X_t - X_0)]$. Pero como por construcción del semigrupo $\mathbb{E}^x [f(X_t - X_0)] = \mathbb{E}^x [f(X_t - x)] = \mathbb{E}^0 [f(X_t)]$, se sigue que la aplicación $x \to \mathbb{E}^x [f(X_t - X_0)]$ es constante igual a $\mathbb{E}^0 [f(X_t)] = P_t f(0)$. Así, para toda $x \in E$

$$\mathbb{E}^x \left[f(X_{s+t} - X_s) | \mathscr{F}_s^0 \right] = \mathbb{E}^0 \left[f(X_t) \right] = P_t f(0). \tag{7.1.9}$$

Al tomar esperanza con respecto a $\mathbb{P}^x$ de ambos lados se tiene que $\mathbb{E}^x [f(X_{s+t} - X_s)] = \mathbb{E}^0 [f(X_t)]$ y por lo tanto la distribución de $X_{s+t} - X_s$ bajo $\mathbb{P}^x$ es la misma que la de X_t bajo $\mathbb{P}^0$. De esto último y (7.1.9) también se concluye que $\mathbb{E}^x \left[f(X_{s+t} - X_s) | \mathscr{F}_s^0 \right] = \mathbb{E}^x [f(X_{s+t} - X_s)]$ y por lo tanto $X_{s+t} - X_s$ es independiente de $\mathscr{F}_s^0$.

2) La independencia de incrementos se sigue de la independencia de $X_{t+s} - X_s$ con respecto a $\mathscr{F}_s^0$ para cada $s, t \geq 0$. Para ver que $X_{t+s} - X_s \stackrel{d}{=} X_t - X_0$, basta notar que por el primer inciso,

$$\left\{ X_{s+t} - X_s \text{ bajo } \mathbb{P}^x \right\} \stackrel{\mathscr{L}}{=} \left\{ X_t \text{ bajo } \mathbb{P}^0 \right\} \stackrel{\mathscr{L}}{=} \left\{ X_t - X_0 \text{ bajo } \mathbb{P}^x \right\}.$$

■

Por esto, llamaremos a X bajo $\mathbb{P}^\nu$ proceso de Lévy con ley inicial ν y en el caso particular en que $\nu = \delta_x$, es un proceso de Lévy iniciado en $X_0 = x$. Es claro entonces que el sistema $\mathcal{X} = (\Omega, \mathscr{F}_\infty^0, (\mathscr{F}_t^0), X, (\mathbb{P}^\nu), (\theta_t), (P_t))$ que se construyó, es el proceso de Markov asociado al proceso de Lévy original Y de semigrupo (P_t). Como ya se mencionó, a menudo no se hace la distinción entre ambos en la literatura. Esto se debe a que dado un proceso que satisface la propiedad de Markov, siempre le podemos asociar un sistema $\mathcal{X}$ que satisface la propiedad de Markov con respecto al mismo semigrupo.

7.2 Propiedad fuerte de Markov de los procesos de Lévy y medida potencial

Referencia: J.Bertoin [9], capítulo 1.

Dedicamos esta sección a resultados propios a los procesos de Lévy. Usaremos la teoría de procesos de Markov desarrollada en los capítulos 5 y 6 y el sistema $\mathcal{X}$ asociado que construimos la sección anterior para estudiar el comportamiento asintótico de los procesos de Lévy. Similarmente a cadenas de Markov, se catalogaran los estados según si son recurrentes o transitorios. Esto se logrará a través de la *medida potencial* y probaremos que para una gran familia de procesos de Lévy, X visita cualquier vecindad para tiempos arbitrariamente grandes o $\|X\|$ tiende a infinito.

Por el resto del capítulo, X es el proceso coordenado usual en el espacio canónico $\Omega = D(\mathbb{R}^+, E)$ tal que bajo $\mathbb{P}^0$ es un proceso de Lévy. Tiene asociado el sistema

$$\mathcal{X} = (\Omega, \mathscr{F}_\infty, (\mathscr{F}_t), X, (\mathbb{P}^\nu), (\theta_t))$$

que satisface la propiedad de Markov con respecto al semigrupo

$$P_t f(x) = \mathbb{E}^0 [f(x + X_t)] = \mathbb{E}^x [f(X_t)] .$$

Así bajo $\mathbb{P}^x$, X es el proceso de Lévy iniciado en x. A menudo, escribiremos simplemente $\mathbb{P}$ en lugar de $\mathbb{P}^0$. Los siguientes resultado son propios a los procesos de Lévy y se les conoce como la propiedad de Markov y la propiedad de Markov fuerte para procesos de Lévy. Empezamos probando la versión más débil, que es una clara generalización de la proposición 7.1.9.

Teorema 7.2.1 Propiedad de Markov para procesos de Lévy
Para cada $a \geq 0$, el proceso $(X_{a+t} - X_a)_{t\geq 0}$ bajo cada $\mathbb{P}^\nu$ es independiente de $\mathscr{F}_a$ y tiene la misma ley que X bajo $\mathbb{P}^0$.

Demostración
Paso 1: Empezamos probando que $(X_{a+t} - X_a)_{t\geq 0}$ bajo $\mathbb{P}^x$ tiene la misma distribución que $(X_t)_{t\geq 0}$ bajo $\mathbb{P}^0$.

Para ver esto, bastaría probar que para $f_1, f_2 \dots f_n \in b\mathscr{E}$ cualesquiera y $t_1 < t_2 < \dots < t_n$ se satisface

$$\mathbb{E}^x \left[\prod_{i=0}^{n} f_i(X_{a+t_i} - X_a) \right] = \mathbb{E}^0 \left[\prod_{i=0}^{n} f_i(X_{t_i}) \right] = P_{t_1} f_1 P_{t_2 - t_1} \dots P_{t_n - t_{n-1}} f(0) \tag{7.2.1}$$

donde la segunda igualdad es por el corolario 5.3.3. Esta es una generalización de lo que se vio en (7.1.9) pues obtendríamos que la aplicación $x \to \mathbb{E}^x \left[\prod_{i=0}^{n} f_i(X_{a+t_i} - X_a)\right]$ es constante. Procedemos por inducción: el caso $n = 1$ sería ver que

$$\mathbb{E}^x [f(X_{a+t_1} - X_a)] = \mathbb{E}^0 [f(X_{t_1})] = P_{t_1} f(0)$$

pero esto es consecuencia de 7.1.9. Ahora, al condicionar y por la independencia de $X_{a+t_n} -$

$X_{a+t_{n-1}}$ bajo $\mathbb{P}^x$ con $\mathscr{F}_{a+t_{n-1}}$ probada en la proposición 7.1.9 obtenemos que

$$\mathbb{E}^x\left[\prod_{i=1}^{n} f_i(X_{a_k+t_i}-X_{a_k})\right]$$
$$=\mathbb{E}^x\left[\prod_{i=1}^{n-1} f_i(X_{a_k+t_i}-X_{a_k})\mathbb{E}^x\left[f_n(\underbrace{X_{a_k+t_n}-X_{a_k+t_{n-1}}}_{\perp\mathscr{F}_{a_k+t_{n-1}}}+X_{a_k+t_{n-1}}-X_{a_k})|\mathscr{F}_{a_k+t_{n-1}}\right]\right]$$

y como $X_{a_k+t_n}-X_{a_k+t_{n-1}}$ bajo $\mathbb{P}^x \overset{\mathscr{L}}{=} X_{t_n-t_{n-1}}$ bajo $\mathbb{P}^0$, ley que hemos denotado $\mu_{t_n-t_{n-1}}$, por independencia y propiedades de la esperanza condicional tenemos que:

$$=\mathbb{E}^x\left[\prod_{i=1}^{n-1} f_i(X_{a_k+t_i}-X_{a_k})\int_E f_n(y+(X_{a_k+t_{n-1}}-X_{a_k}))\mathbb{P}^x_{X_{t_n}-X_{t_{n-1}}}(dy)\right]$$
$$=\mathbb{E}^x\left[\prod_{i=1}^{n-1} f_i(X_{a_k+t_i}-X_{a_k})\int_E f_n(y+(X_{a_k+t_{n-1}}-X_{a_k}))\mu_{t_n-t_{n-1}}(dy)\right]$$

que por construcción del semigrupo de un Lévy es justamente

$$=\mathbb{E}^x\left[\prod_{i=1}^{n-1} f_i(X_{a_k+t_i}-X_{a_k})P_{t_n-t_{n-1}}f_n(X_{a_k+t_{n-1}}-X_{a_k})\right]$$
$$=\mathbb{E}^x\left[\prod_{i=1}^{n-2} f_i(X_{a_k+t_i}-X_{a_k})\cdot\{f_{n-1}P_{t_n-t_{n-1}}f_n\}(X_{a_k+t_{n-1}}-X_{a_k})\right]$$

y por hipótesis de inducción aplicada a las $n-1$ funciones $f_1, f_2, \dots, (f_{n-1}P_{t_n-t_{n-1}}f)$, esto es:

$$=P_{t_1}f_1P_{t_2-t_1}f_2P_{t_3-t_2}\cdots P_{t_{n-2}-t_{n-1}}f_{n-1}P_{t_n-t_{n-1}}f_n(0)$$
$$=\mathbb{E}^0\left[\prod_{i=0}^{n} f_i(X_{t_i})\right].$$

Para probar que esto sigue siendo válido al remplazar x por una medida inicial ν, basta notar que por (5.3.7) se tiene que $\mathbb{E}^\nu\left[\prod_{i=0}^n f_i(X_{a+t_i}-X_a)\right]=\nu(dx)\mathbb{E}^x\left[\prod_{i=0}^n f_i(X_{a+t_i}-X_a)\right]=\mathbb{E}^0\left[\prod_{i=0}^n f_i(X_{t_i})\right]$ o simplemente observar que el argumento sigue siendo válido al considerar una medida ν cualquiera.

Paso 2: Ahora, probemos que $(X_{a+t}-X_a)_{t\geq 0}$ bajo cada $\mathbb{P}^x$ es independiente de $\mathscr{F}_a$. Para ver esto, recordemos que por lo que se probó en la primera parte, la aplicación

$$x\to\mathbb{E}^x\left[\prod_{i=0}^{n} f_i(X_{a+t_i}-X_a)\right] \tag{7.2.2}$$

es constante e igual a $\mathbb{E}^0\left[\prod_{i=0}^n f_i(X_{t_i})\right]$ para cualquier $a\geq 0$. Por propiedad de Markov:

$$\mathbb{E}^x\left[\prod_{i=1}^{n} f(X_{a+t_i}-X_a)|\mathscr{F}_a\right]=\mathbb{E}^x\left[\prod_{i=1}^{n} f(X_{t_i}-X_0)\circ\theta_a|\mathscr{F}_a\right]$$
$$=\mathbb{E}^{X_a}\left[\prod_{i=1}^{n} f(X_{t_i}-X_0)\right]=\mathbb{E}^0\left[\prod_{i=0}^{n} f_i(X_{t_i})\right].$$

Una vez más, este último término es igual a $\mathbb{E}^x\left[\prod_{i=0}^n f_i(X_{a+t_i}-X_a)\right]$ así que con esto concluimos la independencia y la prueba. ■

La propiedad de Markov para procesos de Lévys permite probar una versión aún mas general:

Teorema 7.2.2 Propiedad fuerte de Markov para procesos de Lévy
Si τ es un tiempo de paro finito, el proceso $(X_{\tau+t} - X_\tau)_{t\geq 0}$ bajo cada $\mathbb{P}^\nu$ es independiente de $\mathscr{F}_\tau$ y tiene la misma ley que X bajo $\mathbb{P}^0$.

Demostración
Consideremos $f_1, \cdots, f_n$ continuas y acotadas, $A \in \mathscr{F}_\tau$ cualquiera y tiempos $0 \leq t_1 < \cdots t_n < \infty$. Empezamos probando que

$$\mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau+t_i} - X_\tau)\mathbb{1}_A\right] = \mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau+t_i} - X_\tau)\right]\mathbb{P}^x(A) \tag{7.2.3}$$

cuando τ toma un número finito de valores $\tau \in \{a_1, \cdots, a_n\}$. Por definición de la sigma-álgebra parada, $\{\tau = a_k\} \cap A \in \mathscr{F}_{a_k}$ y entonces al aplicar la propiedad de Markov para procesos de Lévy que se acaba de probar:

$$\begin{aligned}
\mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau+t_i} - X_\tau)\mathbb{1}_A\right] &= \sum_{k=1}^m \mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{a_k+t_i} - X_{a_k})\mathbb{1}_{\{\tau=a_k\}}\mathbb{1}_A\right] \\
&= \sum_{k=1}^m \mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{a_k+t_i} - X_{a_k})\right]\mathbb{P}^x(\tau = a_k, A) \\
&= \mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right]\sum_{k=1}^m \mathbb{P}^x(\tau = a_k, A) \\
&= \mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right]\mathbb{P}^x(A)
\end{aligned}$$

donde en la tercer igualdad se usó que para cada k, el proceso $(X_{a_k+t} - X_{a_k})_{t\geq 0}$ bajo $\mathbb{P}^x$ tiene la misma distribución que X bajo $\mathbb{P}^0$. Como $\mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right] = \mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{a+t_i} - X_a)\right]$, por un lado esto nos da la independencia con respecto a $\mathscr{F}_\tau$ pero por otro lado, al tomar $A = \Omega$, obtenemos que

$$\mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau+t_i} - X_\tau)\right] = \mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right]. \tag{7.2.4}$$

por lo que también se tiene la segunda afirmación. El resultado se tiene entonces para τ tiempo de paro discreto. Si consideramos ahora un tiempo de paro τ finito cualquiera, existe una sucesión (τ_k) de tiempos de paro discretos que decrece a τ. Por lo que acabamos de probar, tenemos que para cada k:

$$\mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau_k+t_i} - X_{\tau_k})\mathbb{1}_A\right] = \mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right]\mathbb{P}^x(A).$$

Por continuidad de cada f_i, continuidad por la derecha de X y convergencia dominada, al tomar el límite cuando $k \to \infty$ obtenemos

$$\mathbb{E}^x\left[\prod_{i=1}^n f_i(X_{\tau+t_i} - X_\tau)\mathbb{1}_A\right] = \mathbb{E}^0\left[\prod_{i=1}^n f_i(X_{t_i})\right]\mathbb{P}^x(A).$$

Esto prueba ambas afirmaciones para τ finito cualquiera por el mismo argumento que se dio en el caso discreto. ■

Definición 7.2.3 *Sea X un proceso de Lévy. Su medida potencial, denotada $U(x,\cdot)$ se define como*

$$U(x,B)=\int_0^\infty \mathbb{P}^x\left(X_t\in B\right)dt=\mathbb{E}^x\left[\int_0^\infty \mathbb{1}_{\{X_t\in B\}}dt\right]$$

para cualquier B boreliano.

$U(x,B)$ es el tiempo esperado que pasa el proceso en un conjunto B, habiendo iniciado en $X_0=x$. No es difícil comprobar que en efecto se trata de una medida en $(\mathbb{R}^d,\mathscr{B}(\mathbb{R}^d))$. La propiedad de Markov fuerte nos permite reescribir a $U(x,\cdot)$ de la siguiente conveniente forma:

Lema 7.2.4 *Para todo $x\in E$ y A abierto o cerrado,*

$$U(x,A)=\int_{\overline{A}} U(y,A)\mathbb{P}^x\left(X_{T_A}\in dy\right) \tag{7.2.5}$$

donde $\overline{A}$ denota la cerradura del conjunto A.

Demostración
Primero, obsérvese que si $t<T_A$, entonces $\mathbb{1}_{\{X_t\in A\}}=0$ y entonces

$$\int_0^\infty \mathbb{1}_{\{X_t\in A\}}dt=\int_{T_A}^\infty \mathbb{1}_{\{X_t\in A\}}dt=\int_0^\infty \mathbb{1}_{\{X_{T_A+t}\in B\}}dt=\int_0^\infty \mathbb{1}_{\{A\}}(X_t)\circ\theta_{T_A}dt.$$

La tercer igualdad se usará frecuentemente. Así, al notar que $X_{T_A}\in\overline{A}$, por Propiedad fuerte de Markov 6.3.1

$$\begin{aligned}
U(x,A)=\mathbb{E}^x\left[\int_0^\infty \mathbb{1}_{\{X_t\in A\}}dt\right]&=\int_0^\infty \mathbb{E}^x\left[\mathbb{1}_{\{A\}}(X_t)\circ\theta_{T_A}\right]dt\\
&=\int_0^\infty \mathbb{E}^x\left[\mathbb{E}^x\left[\mathbb{1}_{\{A\}}(X_t)\circ\theta_{T_A}|\mathscr{F}_{T_A}\right]\right]dt\\
&=\int_0^\infty \mathbb{E}^x\left[\mathbb{E}^{X_{T_A}}\left[\mathbb{1}_{\{X_t\in A\}}\right]\right]dt\\
&=\int_0^\infty \int_{\overline{A}}\mathbb{E}^y\left[\mathbb{1}_{\{X_t\in A\}}\right]\mathbb{P}^x\left(X_{T_A}\in dy\right)dt\\
&=\int_{\overline{A}} U(y,A)\mathbb{P}^x\left(X_{T_A}\in dy\right)
\end{aligned}$$

■

Observación: Esta no es más que una variante de una propiedad que ya estudiamos. Al recordar la observación que se hizo al concluir el teorema 6.3.1, si denotamos por μ_{T_A} a la distribución de X_{T_A} bajo $\mathbb{P}^x$, por la igualdad (7.2.5)

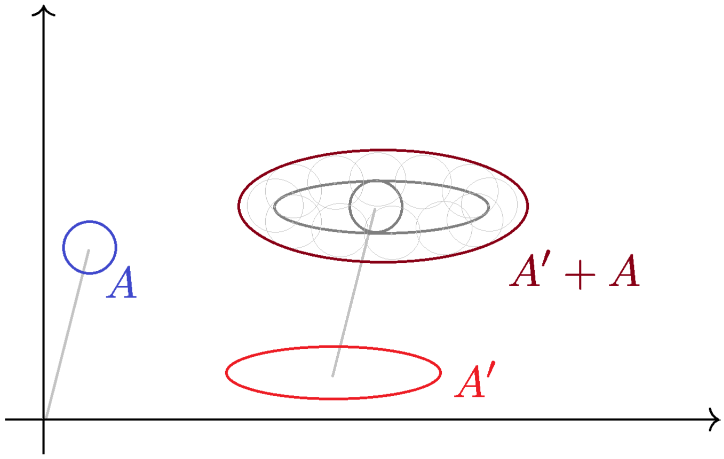
$$\begin{aligned}
\mathbb{E}^x\left[\int_0^\infty \mathbb{1}_{\{X_t\in A\}}dt\right]&=U(x,A)\\
&=\int_{\overline{A}} U(y,A)\mathbb{P}^x\left(X_{T_A}\in dy\right)\\
&=\int_E \mathbb{E}^y\left[\int_0^\infty \mathbb{1}_{\{X_t\in A\}}dt\right]\mathbb{P}^x(X_{T_A}\in dy)=\mathbb{E}^{\mu_{T_A}}\left[\int_0^\infty \mathbb{1}_{\{X_t\in A\}}dt\right].
\end{aligned}$$

7.3 Comportamiento asintótico: recurrencia y transitividad

Introducimos la siguiente notación: si A y A' son subconjuntos de $\mathbb{R}^d$, entonces

$$A + A' = \{x + x',\, x \in A \text{ y } x' \in A'\} \qquad A - A' = \{x - x',\, x \in A \text{ y } x' \in A'\}$$

y denotaremos por $A - x$ a $A - \{x\}$ para $x \in \mathbb{R}^d$.



En la figura anterior se está pensando en conjuntos con interior. Estos no se colorearon para mayor claridad.

La idea es caracterizar el comportamiento de las trayectorias de X según los valores que tome su medida potencial. Heurísticamente, si $U(x, B)$ es finita, X no regresa a B a partir de algún t suficientemente grande, mientras que si es no es finita, X debe pasar mucho tiempo en B o regresar una infinidad de veces. De ahí, las siguientes definiciones:

Definición 7.3.1
1. Decimos que X es transitorio si para todo K compacto, para todo $x \in \mathbb{R}^d$ se satisface $U(x, K) < \infty$.
2. Decimos que X es recurrente si para toda bola $B(0, r) = B$ centrada en 0 y de radio r arbitrario, se satisface que $U(0, B) = \infty$.

Nótese que si se cumple 1 para $x = 0$, entonces se tiene para todo $x \in \mathbb{R}^d$:

$$\begin{aligned} U(x, K) = \int_0^\infty \mathbb{P}^x (X_t \in K)\, dt = \int_0^\infty \mathbb{P}(X_t + x \in K)\, dt &= \int_0^\infty \mathbb{P}(X_t \in K - x)\, dt \\ &= U(0, K - x). \end{aligned} \tag{7.3.1}$$

Esto se cumple en general, y como lo usaremos constantemente lo dejamos en forma de lema:

Lema 7.3.2 *Para todo $D \in \mathscr{B}(\mathbb{R}^d)$ se satisface que*

$$U(x, D) = U(0, D - x). \tag{7.3.2}$$

Así, la definición de transitoriedad es equivalente a

Definición 7.3.3 *Decimos que X es transitorio si para todo K compacto, se satisface que $U(0, K) < \infty$.*

Nótese que X no puede ser recurrente y transitorio a la vez. El siguiente lema, nos permite concluir que X siempre debe ser recurrente o transitorio.

Proposición 7.3.4 *Si existe $\epsilon > 0$ tal que $U(0, B(0,\epsilon)) < \infty$, entonces X es transitorio.*

Demostración
Sea K un compacto cualquiera, hay que probar que

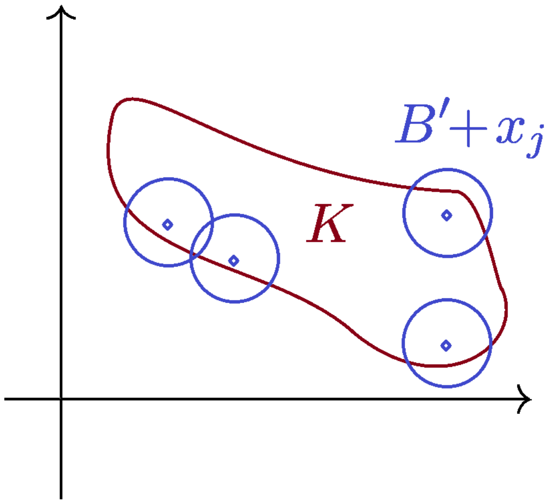
$$U(0, K) < \infty.$$

Primero probaremos que al considerar una bola B' centrada en el origen suficientemente pequeña, se satisface que

$$U(0, B' + x) < \infty$$

para todo $x \in \mathbb{R}^d$. Nótese que por el lema 7.3.2, esto equivale a probar que

$$U(x, B') < \infty,$$



para todo $x \in \mathbb{R}^d$. Al cubrir al compacto K por una cantidad finita de bolas de la forma $x_i + B'$ que por construcción tienen medida potencial finita, finalizaremos la prueba.

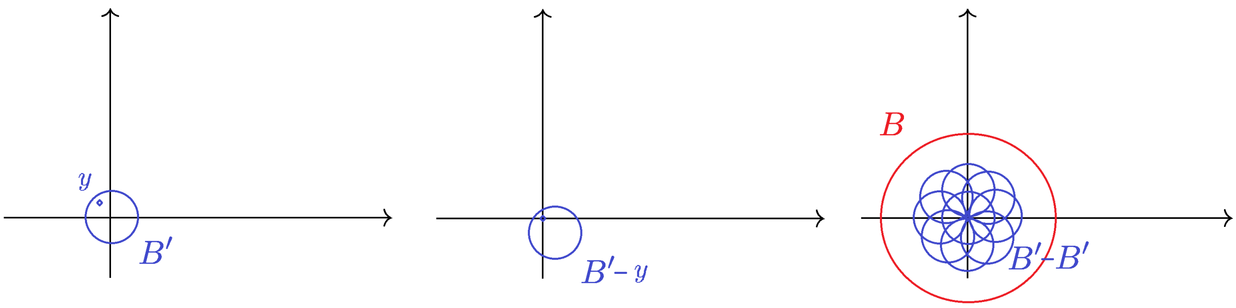
Sea $B = B(0,\epsilon)$ y $B' = \overline{B(0, \frac{\epsilon}{3})}$. Se sigue que $B' - B' = \overline{B(0, \frac{2\epsilon}{3})} \subset B$. Si $T_{B'} = \inf\{t \geq 0\,;\, X_t \in B'\}$,

$$U(x, B') = \int_{B'} U(y, B') d\mathbb{P}^x \left(X_{T_{B'}} \in dy\right)$$

pero

$$\sup_{y \in B'} U(y, B') = \sup_{y \in B'} U(0, B' - y) \leq U(0, B' - B') \leq U(0, B) < \infty$$

ya que por hipótesis, $U(0, B)$ es finito.



Así, obtenemos que $U(x, B') < \infty$ para cualquier $x \in \mathbb{R}^d$. Entonces, para x_j arbitrario

$$U(-x_j, B') < \infty \qquad \text{por lo que} \qquad U(0, B' + x_j) < \infty.$$

Como K se puede cubrir por una cantidad finita de bolas $B' + x_j$ para $j = 1, \ldots n$, concluimos:

$$U(0, K) \leq U(0, \cup^n B' + x_j) \leq \sum_{j=1}^{n} U(0, B' + x_j) < \infty.$$

■

Definición 7.3.5 *Denotaremos por Σ al soporte de la medida $U(0,\cdot)$, es decir*

$$\Sigma = \{x \in \mathbb{R}^d\,;\, U(0,N_x) > 0 \ \ \textit{para toda } N_x \textit{ vecindad abierta de } x\}.$$

Diremos que x es punto posible si $x \in \Sigma$.

Proposición 7.3.6 *Son equivalentes:*

(i) x es punto posible.
(ii) Para todo $\epsilon > 0$, existe $t > 0$ tal que

$$\mathbb{P}\left(X_t \in B(x,\epsilon)\right) > 0.$$

(iii) Para toda bola B con centro en x,

$$\mathbb{P}\left(T_B < \infty\right) > 0.$$

Demostración
$(ii) \Rightarrow (i)$
Sea $\epsilon > 0$ y $B(x,\epsilon)$ bola abierta con centro en x arbitraria, bastaría probar que $U(0,B(x,\epsilon)) > 0$. Denotamos $B = B(x,\epsilon)$ y $B' = B(x,\frac{\epsilon}{2})$. Sabemos que existe $t > 0$ tal que $\mathbb{P}\left(X_t \in B'\right) > 0$. Como

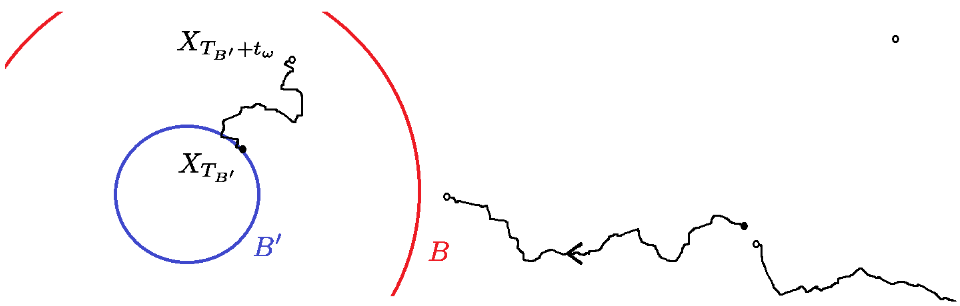
$$\left\{X_t \in B'\right\} \subset \left\{T_{B'} < \infty\right\}$$

entonces

$$0 < \mathbb{P}\left(X_t \in B'\right) < \mathbb{P}\left(T_{B'} < \infty\right). \tag{7.3.3}$$

Sea $\omega \in \left\{T_{B'} < \infty\right\}$, por continuidad por la derecha $X_{T_{B'}}(\omega) \in \overline{B'}$. Una vez más por continuidad por la derecha de X, existe $t_\omega > 0$ tal que

$$X_t(\omega) \in B \qquad \text{para todo } t \in \left[T_{B'}(\omega), T_{B'}(\omega) + t_\omega\right].$$



Entonces, para cada $\omega \in \left\{T_{B'} < \infty\right\}$

$$\int_0^\infty \mathbb{1}_{\{X_t \in B\}}(\omega)dt > \int_{T_{B'}(\omega)}^{T_{B'}(\omega)+t_\omega} \mathbb{1}_{\{X_t \in B\}}(\omega)dt = t_\omega > 0$$

Esta es una desigualdad que se tiene trayectoria por trayectoria. Se sigue que

$$U(0,B) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{X_t \in B\}}dt\right] \geq \mathbb{E}\left[\mathbb{1}_{\{T_{B'}<\infty\}}\int_0^\infty \mathbb{1}_{\{X_t \in B\}}dt\right] > 0$$

pues en el conjunto $\{T_{B'} < \infty\}$ de medida $\mathbb{P}(T_{B'} < \infty) > 0$, tenemos la desigualdad estricta $\int_0^\infty \mathbb{1}_{\{X_t \in B\}} dt > 0$. Así, x es punto posible.

$(i) \Rightarrow (ii)$

Supongamos que x es punto posible. Como $x \in \Sigma$, para toda $B(x,\epsilon)$ se tiene que $U(0, B(x,\epsilon)) > 0$ y entonces

$$0 < \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{X_t \in B(x,\epsilon)\}} dt\right] = \int_0^\infty \mathbb{P}(X_t \in B(x,\epsilon))\, dt$$

por lo que debe existir $t > 0$ tal que $\mathbb{P}(X_t \in B(x,\epsilon)) > 0$.

$(iii) \Rightarrow (i)$

Sea $\epsilon > 0$ y supongamos $\mathbb{P}(T_B < \infty) > 0$ para cualquier bola B centrada en x. Como se satisface (7.3.3), por el mismo argumento que en $(ii) \Rightarrow (i)$ se tiene que $x \in \Sigma$.

$(i) \Rightarrow (iii)$

Por $(i) \Rightarrow (ii)$, para todo $\epsilon > 0$ existe $t > 0$ tal que $\mathbb{P}(X_t \in B(x,\epsilon)) > 0$ y como $\{X_t \in B(x,\epsilon)\} \subset \{T_{B(x,\epsilon)} < \infty\}$, se sigue (iii). ■

Proposición 7.3.7 *Si $x, y \in \Sigma$, entonces $x + y \in \Sigma$.*

Demostración

Sea B_{x+y} la bola de radio ϵ centro $x+y$, bastaría probar que $U(0, B_{x+y}) > 0$. Para ver esto, consideremos B_x, B_y dos bolas cerradas con centro x, y respectivamente de radio suficientemente pequeño de modo que $B_x + B_y \subset B_{x+y}$. Como tanto x como y están en Σ, ambas vecindades tienen medida potencial positiva $U(0, B_x)$, $U(0, B_y) > 0$. Primero haremos pasar a X por B_x, sabemos que esto tiene probabilidad positiva de ocurrir pues $x \in \Sigma$. Luego, al utilizar la homogeneidad temporal que nos da la propiedad de Markov, reiniciaremos al proceso en B_x y probaremos que con probabilidad positiva podemos llegar a B_{x+y}. Aquí se usará la homogeneidad espacial del semigrupo de un proceso de Lévy pues ir de x a B_{x+y} tiene la misma distribución que partir del origen y viajar a B_y. De este último evento, si sabemos que tiene probabilidad positiva puesto que $y \in \Sigma$.

Primero, como $x \in \Sigma$, por 7.3.6 $\mathbb{P}(T_{B_x} < \infty) > 0$ y podemos condicionar:

$$\begin{aligned} U(0, B_{x+y}) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{X_t \in B_{x+y}\}} dt\right] &\geq \mathbb{E}\left[\mathbb{1}_{\{T_{B_x} < \infty\}} \int_0^\infty \mathbb{1}_{B_{x+y}}(X_{T_{B_x}+t}) dt\right] \\ &= \int_0^\infty \mathbb{E}\left[\mathbb{1}_{\{T_{B_x} < \infty\}} \mathbb{E}\left[\mathbb{1}_{B_{x+y}}(X_t) \circ \theta_{T_{B_x}} | \mathscr{F}_{T_{B_x}}\right]\right] dt \\ &= \int_0^\infty \mathbb{E}\left[\mathbb{1}_{\{T_{B_x} < \infty\}} \mathbb{E}^{X_{T_{B_x}}}\left[\mathbb{1}_{B_{x+y}}(X_t)\right]\right] dt \end{aligned}$$

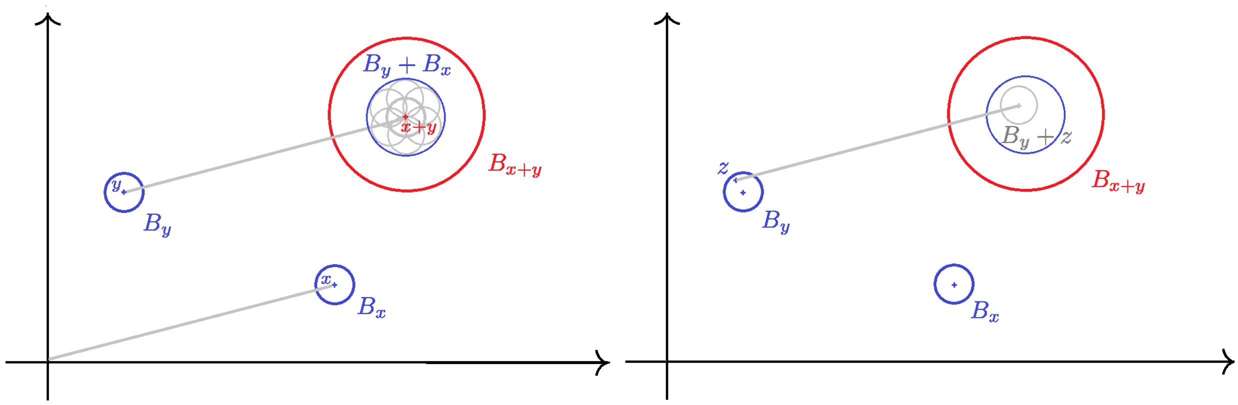
y como $\mathbb{P}(A \,| T_{B_x} < \infty) = \int_A \frac{\mathbb{1}_{\{T_{B_x} < \infty\}}}{\mathbb{P}(T_{B_x} < \infty)} d\mathbb{P}, \quad A \in \mathscr{F}$

$$\begin{aligned} &= \mathbb{P}(T_{B_x} < \infty) \int_0^\infty \mathbb{E}\left[\mathbb{E}^{X_{T_{B_x}}}\left[\mathbb{1}_{B_{x+y}}(X_t)\right] \Big| T_{B_x} < \infty\right] dt \\ &= \mathbb{P}(T_{B_x} < \infty) \int_0^\infty \int_{\overline{B_x}} \mathbb{E}^z\left[\mathbb{1}_{B_{x+y}}(X_t)\right] d\mathbb{P}\left(X_{T_{B_x}} \in dz \,| T_{B_x} < \infty\right) dt \\ &= \mathbb{P}(T_{B_x} < \infty) \int_{\overline{B_x}} U(z, B_{x+y}) \mathbb{P}\left(X_{T_{B_x}} \in dz \,| T_{B_x} < \infty\right). \end{aligned}$$

Bastaría probar que $U(z, B_{x+y}) > 0$ para todo $z \in B_x$ para concluir la prueba. Aquí es donde usamos la homogeneidad espacial: como $y \in \Sigma$, entonces $U(0, B_y) > 0$ y por lo tanto

$$U(z, B_y + z) > 0 \tag{7.3.4}$$

Como por construcción $B_y + B_x \subset B_{x+y}$, para todo $z \in B_x$ se satisface que $B_y + z \subset B_{x+y}$ y por lo tanto $U(z, B_{x+y}) > 0$. ■



Nos centraremos en estudiar procesos de Lévy para los cuales el grupo generado por Σ es $\mathbb{R}^d$ es decir, X no vive en un subconjunto (estricto) de $\mathbb{R}^d$. Llamaremos a este supuesto

$$\text{la hipótesis } (H).$$

Si B_0 es una bola con centro en 0, denotamos por $B_x = x + B_0$ a la bola B_0 trasladada con centro en x.

Introducimos el conjunto de puntos que son visitados para tiempos arbitrariamente grandes:

$$\mathscr{R} = \left\{ x \in \mathbb{R}^d ;\ \text{para cualesquiera} B_0 \ \text{ y } t \geq 0,\, \mathbb{P}\left(\exists\, s \geq t \ \text{ tal que } X_s \in x + B_0\right) = 1 \right\}$$

donde B_0 denota una bola con centro en 0. Nótese que $\mathscr{R} \subset \Sigma$ y obsérvese que $\mathscr{R}$ se puede reescribir usando la notación introducida en (6.3.8) como sigue:

$$\mathscr{R} = \left\{ x \in \mathbb{R}^d ;\, \forall\, B_0 \ \text{ y } t \geq 0,\, \mathbb{P}\left(T^t_{x+B_0} < \infty\right) = 1 \text{ para todo } t \geq 0 \right\}.$$

Empecemos probando que si $x \in \mathscr{R}$, entonces podemos reemplazar a t por un tiempo de paro finito:

Proposición 7.3.8 *Si $x \in \mathscr{R}$ y τ es un tiempo de paro finito, entonces:*

$$\mathbb{P}\left(T^\tau_{x+B_0} < \infty\right) = 1.$$

Si τ no es finito pero $\mathbb{P}(\tau < \infty) > 0$, al condicionar sobre $\{\tau < \infty\}$ obtenemos la propiedad análoga:

$$\mathbb{P}\left(T^\tau_{x+B_0} < \infty \,|\, \tau < \infty\right) = 1.$$

Esta última proposición será clave para lo que probaremos a continuación y nos da una mejor intuición de las propiedades de los puntos en $\mathscr{R}$: si A es un boreliano cualquiera y $\tau = T_A$ el tiempo de arribo de X a subconjunto, la proposición anterior nos garantiza que después de pasar por A, regresaremos eventualmente a cualquier vecindad de x (al restringirnos al evento $\{T_A < \infty\}$).

Demostración
Para cada $N \in \mathbb{N}$, $\tau \wedge N$ es un tiempo de paro finito y como $T^{\tau\wedge N}_{x+B_0} \leq T^{N}_{x+B_0}$,

$$\mathbb{P}\left(T^{\tau\wedge N}_{x+B_0} < \infty\right) \geq \mathbb{P}\left(T^{N}_{x+B_0} < \infty\right) = 1.$$

Como $T^{\tau\wedge N}_{x+B_0} \uparrow T^{\tau}_{x+B_0}$, la sucesión de conjuntos $\left\{T^{\tau\wedge N}_{x+B_0} < \infty\right\}_N$ decrece a $\left\{T^{\tau}_{x+B_0} < \infty\right\}$. Así, al tomar el límite $N \to \infty$ se tiene el resultado:

$$\mathbb{P}\left(T^{\tau}_{x+B_0} < \infty\right) = \lim_{n\to\infty} \mathbb{P}\left(T^{\tau\wedge N}_{x+B_0} < \infty\right) = 1.$$

Si τ no es finito pero $\mathbb{P}(\tau < \infty) > 0$, la sucesión de conjuntos $\left\{T^{\tau\wedge N}_{x+B_0} < \infty, \tau < \infty\right\}_N$ decrece a $\left\{T^{\tau}_{x+B_0} < \infty, \tau < \infty\right\}$ y similarmente

$$\mathbb{P}\left(T^{\tau}_{x+B_0} < \infty, \tau < \infty\right) = \lim_{n\to\infty} \mathbb{P}\left(T^{\tau\wedge N}_{x+B_0} < \infty, \tau < \infty\right). \tag{7.3.5}$$

Además, por lo primero que probamos, sabemos que

$$\mathbb{P}\left(T^{\tau\wedge N}_{x+B_0} < \infty\right) = 1$$

pues $\tau \wedge N$ es un tiempo de paro finito. Se sigue al dividir ambos lados de (7.3.5) por $\mathbb{P}(\tau < \infty)$ que

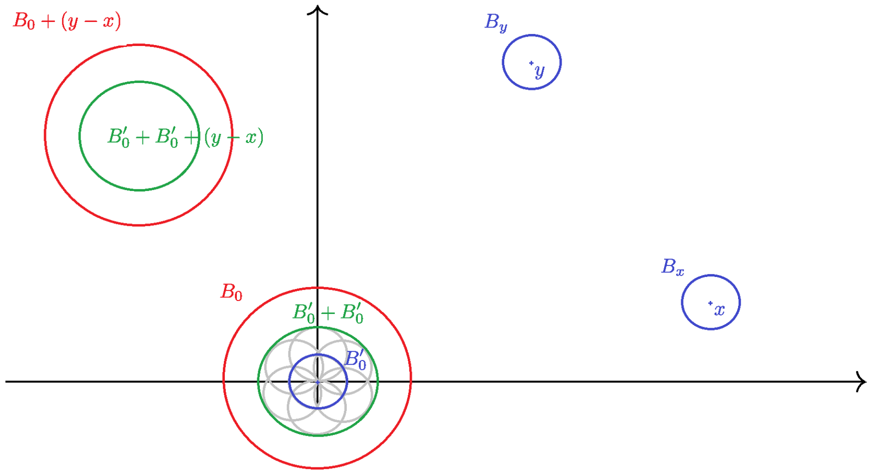
$$\mathbb{P}\left(T^{\tau}_{x+B_0} < \infty | \tau < \infty\right) = \lim_{n\to\infty} \mathbb{P}\left(T^{\tau\wedge N}_{x+B_0} < \infty | \tau < \infty\right) = 1.$$

■

Ahora, probaremos que para procesos de Lévy que satisfacen (H) solo pueden ocurrir dos cosas: todo punto se visita para tiempos arbitrariamente grandes (esto es, la trayectoria $(X_t)_{t\geq T}$ es densa en $\mathbb{R}^d$ para toda $T \in \mathbb{N}$) o en caso contrario, todo punto se deja de visitar eventualmente:

Proposición 7.3.9 *Bajo (H), $\mathscr{R} = \emptyset$ o $\mathscr{R} = \mathbb{R}^d$.*

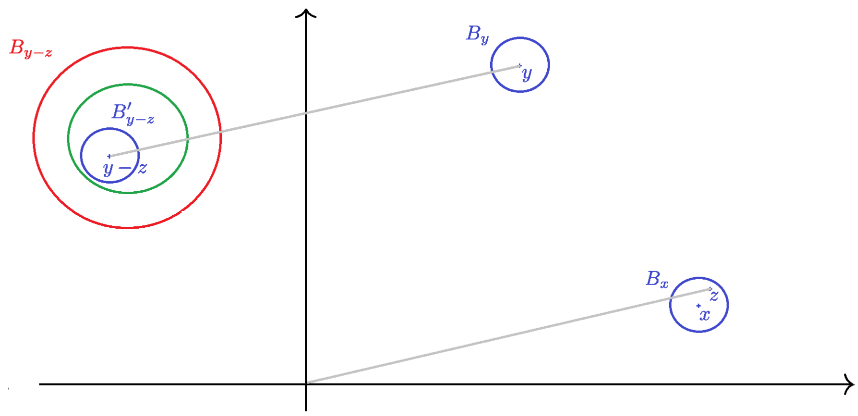
Demostración
Primero, supongamos que $\mathscr{R} \neq \emptyset$ por lo que existe al menos un elemento $y \in \mathscr{R}$. Por (H), $\Sigma = \mathbb{R}^d$. Así, consideramos $x \in \Sigma = \mathbb{R}^d$ cualquiera y B_0 una bola con centro en 0 de radio arbitrario. Sea B_0' otra bola con centro en 0 tal que $\overline{B_0'} + \overline{B_0'} \subset B_0$.



Probaremos que $y - x \in \mathscr{R}$ es decir, que para cualquier t

$$\mathbb{P}\left(T^t_{B_{y-x}} < \infty\right) = 1.$$

La idea es la siguiente: como $x \in \Sigma$, con probabilidad positiva podemos llegar a B'_x. Usando la propiedad de Markov, trasladamos el origen a x y reiniciamos al proceso en z dentro de B'_x. Como $y \in \mathscr{R}$, por la proposición anterior con probabilidad uno eventualmente el proceso ingresa a la vecindad B'_y. Así restringidos al evento $\{T_{B'_x} < \infty\}$, podemos ir de B'_x a B'_y con probabilidad uno. Por la homogeneidad espacial, ir de un punto $z \in B'_x$ a B'_y tiene la misma probabilidad que iniciar en 0 y llegar a B'_{y-z}, bola que se queda dentro B_{y-x} por construcción. Al trasladarnos nuevamente al origen la condición $\{T_{B'_x} < \infty\}$ perderá su efecto pues el proceso inicia en 0. Así concluiremos que con probabilidad 1 ingresamos a la bola B_{y-x} después de t unidades de tiempo. A continuación, procedemos con detalle.



Como $x \in \Sigma$ por la proposición 7.3.6 $\mathbb{P}\left(T_{B'_x} < \infty\right) > 0$ y como $y \in \mathscr{R}$, por la proposición anterior

$$\mathbb{P}\left(\left(T_{B'_y}\right)^{t+T_{B'_x}} < \infty \middle| T_{B'_x} < \infty\right) = 1.$$

Juntando las observaciones (6.3.9), (6.3.10) y la propiedad de Markov

$$\begin{aligned}
\mathbb{P}\left(\left(T_{B'_y}\right)^{t+T_{B'_x}} < \infty\,,\, T_{B'_x} < \infty\right) &= \mathbb{E}\left[\mathbb{1}_{\{T_{B'_x}<\infty\}}\mathbb{1}_{\mathbb{R}^+}\left(T_{B'_y}\right)^t \circ \theta_{T_{B'_x}}\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{T_{B'_x}<\infty\}}\mathbb{E}\left[\mathbb{1}_{\mathbb{R}^+}\left(T_{B'_y}\right)^t \circ \theta_{T_{B'_x}} | \mathscr{F}_{T_{B'_x}}\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{T_{B'_x}<\infty\}}\mathbb{E}^{X_{T_{B'_x}}}\left[\mathbb{1}_{\mathbb{R}^+}\left(T_{B'_y}\right)^t\right]\right] \\
&= \int_\Omega \mathbb{E}^{X_{T_{B'_x}}}\left[\mathbb{1}_{\mathbb{R}^+}\left(T_{B'_y}\right)^t\right]\mathbb{1}_{\{T_{B'_x}<\infty\}} d\mathbb{P},
\end{aligned}$$

por lo que dividiendo ambos lados entre $\mathbb{P}\left(T_{B'_x} < \infty\right)$ y cambio de variable, se tiene que

$$1 = \mathbb{P}\left(\left(T_{B'_y}\right)^{t+T_{B'_x}} < \infty \middle| T_{B'_x} < \infty\right) = \int_{\overline{B'_x}} \mathbb{P}^z\left(\left(T_{B'_y}\right)^t < \infty\right) d\mathbb{P}\left(X_{T_{B'_x}} \in dz \,|\, T_{B'_x} < \infty\right). \tag{7.3.6}$$

Ahora,

$$\begin{aligned}
\mathbb{P}^z\left(\left(T_{B'_y}\right)^t < \infty\right) = \mathbb{P}^z\left(\exists\, s \geq t\,;\, X_s \in B'_y\right) &= \mathbb{P}^0\left(\exists\, s \geq t\,;\, X_s + z \in B'_y\right) \\
&= \mathbb{P}\left(\exists\, s \geq t\,;\, X_s \in\, y - z + B'_0\right) \\
&= \mathbb{P}\left(\exists\, s \geq t\,;\, X_s \in\, (y-x) + (x-z) + B'_0\right)
\end{aligned}$$

y como $z \in \overline{B'_x}$, entonces $x - z \in \overline{B'_0}$ al ser x el centro de B'_x. Pero entonces, por construcción, para todo $z \in \overline{B'_x}$

$$(y-x) + (x-z) + B'_0 \subset (y-x) + \overline{B'_0} + \overline{B'_0} \subset (y-x) + B_0.$$

Así,

$$\begin{aligned}
\mathbb{P}^z\left(\left(T_{B'_y}\right)^t < \infty\right) &= \mathbb{P}\left(\exists\, s \geq t\,;\, X_s \in\, (y-x) + (x-z) + B'_0\right) \\
&\leq \mathbb{P}\left(\exists\, s \geq t\,;\, X_s \in\, (y-x) + B_0\right) \\
&= \mathbb{P}\left(T^t_{B_{y-x}} < \infty\right)
\end{aligned}$$

que no depende de z, por lo que remplazando en (7.3.6) obtenemos:

$$1 = \mathbb{P}\left(\left(T_{B'_y}\right)^{t+T_{B'_x}} < \infty \middle| T_{B'_x} < \infty\right) \leq \mathbb{P}\left(T^t_{B_{y-x}} < \infty\right).$$

Al ser t arbitrario, se sigue que $y - x \in \mathscr{R}$. Como $\Sigma = \mathbb{R}^d$ y cualquier punto $z \in \mathbb{R}^d$ se puede ver como $y - x$ para algún $x \in \Sigma$, concluimos que $\mathscr{R} = \mathbb{R}^d$. ∎

Teorema 7.3.10 *Bajo* (H)*, son equivalentes:*

(1) $\mathscr{R} = \emptyset$

(2) X *es transitorio*

(3) $\lim\limits_{t\to\infty} |X_t| = \infty$

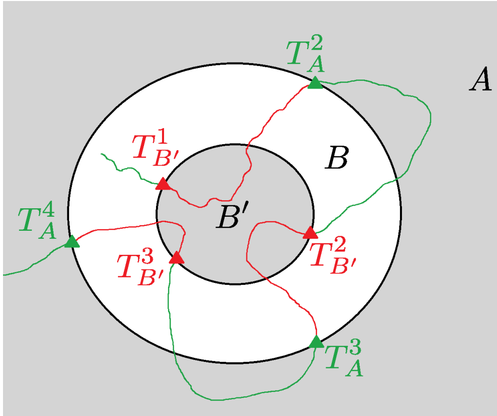
Demostración

(2) $\Rightarrow$ (1)

Supongamos X es transitorio. Sean B' y B dos bolas con centro en 0, B' cerrada, tales $B'+B' \subset B$. Sea $A = B^c$ y $A' = (B')^c$. Supongamos que $\mathscr{R} \neq \emptyset$, así por la proposición anterior, $\mathscr{R} = \mathbb{R}^d$.

Definimos la siguiente sucesión de tiempos de paro:

$$\begin{aligned}
T_A^1 = T_A \quad & = \inf\{t \geq 0\,;\, X_t \in A\} \\
T_{B'}^1 = (T_{B'})^{T_A^1} \quad & = \inf\{t > T_A^1\,;\, X_t \in B'\} \\
T_A^2 = (T_A)^{T_{B'}^1} \quad & = \inf\{t > T_{B'}^1\,;\, X_t \in A\} \\
T_{B'}^2 = (T_{B'})^{T_A^2} \quad & = \inf\{t > T_A^2\,;\, X_t \in B'\} \\
\vdots \quad & \quad \vdots
\end{aligned}$$



Como $\mathscr{R} = \mathbb{R}^d$, por la proposición 7.3.8 son todos finitos y la sucesión es creciente. Entre $T_{B'}^{n-1}$ y T_A^n, el proceso X ya regresó a B' por $n-1$-ésima vez pero aún no regresa a A por n-ésima vez. Así, $X_t \in B$ para $t \in [T_{B'}^{n-1}, T_A^n]$ y entonces X pasa $T_A^n - T_{B'}^{n-1}$ unidades de tiempo en B entre esas dos llegadas. Por lo tanto:

$$U(0,B) = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{X_t \in B\}} dt\right] \geq \mathbb{E}\left[\sum_{n=1}^\infty \left(T_A^n - T_{B'}^{n-1}\right)\right]$$

Recordemos que $T_A^n = (T_A)^{T_{B'}^{n-1}} = T_{B'}^{n-1} + T_A \circ \theta_{T_{B'}^{n-1}}$ y entonces $T_A^n - T_{B'}^{n-1} = T_A \circ \theta_{T_{B'}^{n-1}}$. Por propiedad de Markov:

$$\begin{aligned}
\sum_{n=1}^\infty \mathbb{E}\left[T_A^n - T_{B'}^{n-1}\right] &= \sum_{n=1}^\infty \mathbb{E}\left[\mathbb{E}\left[T_A \circ \theta_{T_{B'}^{n-1}} | \mathscr{F}_{T_{B'}^{n-1}}\right]\right] \\
&= \sum_{n=1}^\infty \mathbb{E}\left[\mathbb{E}^{X_{T_{B'}^{n-1}}}\left[T_A\right]\right] \\
&\geq \sum_{n=1}^\infty \inf_{y \in B'} \mathbb{E}^y\left[T_A\right]
\end{aligned}$$

pues $X_{T_{B'}^{n-1}} \in B'$ al ser una bola cerrada. Para llegar a una contradicción, bastaría probar que esta serie diverge, pues $U(0,B) < \infty$ al ser X transitorio. Notemos que

$$\begin{aligned}
T_{A'} = \inf\{t \geq 0\,;\, X_t \notin B'\} &= \inf\{t \geq 0\,;\, X_t + y \notin \underbrace{B' + y}_{\subset B}\} \\
&\leq \inf\{t \geq 0\,;\, X_t + y \notin B\} = T_A
\end{aligned}$$

pues $y \in B'$ y $B' + B' \subset B$. Entonces, $\mathbb{E}^0\left[T'_A\right] \leq \mathbb{E}^y\left[T_A\right]$ para todo $y \in B'$ y por continuidad por la derecha, $0 < \mathbb{E}^0\left[T'_A\right]$ pues X debe permanecer cerca del 0 en una vecindad de $t = 0$ bajo $\mathbb{P}^0$ y entonces $0 < T'_A$. De esto, se sigue que $0 < \inf_{y \in B'} \mathbb{E}^y\left[T_A\right]$. Al sumar sobre n, el término principal no depende de n y entonces la serie diverge, con lo que concluimos.

(1) $\Rightarrow$ (2)

Supongamos $\mathscr{R} = \emptyset$. Como en particular $0 \notin \mathscr{R}$, debe existir $\epsilon > 0$, una bola B con centro en 0 radio ϵ y un tiempo $t_0 > 0$ tales que

$$\mathbb{P}\left(\exists s \geq t_0\,;\, X_s \in B\right) = \mathbb{P}\left((T_B)^{t_0} < \infty\right) < 1-\epsilon.$$

Probaremos que la medida potencial $U(0,B') < \infty$ y por lo tanto X es transitorio. Construimos una bola cerrada B' con centro en 0 tal que $B'+B' \subset B$ y consideremos la siguiente sucesión de tiempos de paro:

$$\begin{aligned}
T^1_{B'} &= (T_{B'})^{t_0} && = \inf\{t > t_0\,;\, X_t \in B'\} \\
T^2_{B'} &= (T_{B'})^{T^1_{B'}+t_0} && = \inf\{t > T^1_{B'} + t_0\,;\, X_t \in B'\} \quad \text{en } \{T^1_{B'} < \infty\} \\
&\vdots \quad \vdots && \quad \vdots \\
T^n_{B'} &= (T_{B'})^{T^{n-1}_{B'}+t_0} && = \inf\{t > T^{n-1}_{B'} + t_0\,;\, X_t \in B'\} \quad \text{en } \{T^{n-1}_{B'} < \infty\} \\
&\vdots \quad \vdots && \quad \vdots
\end{aligned}$$

La colección de intervalos $[T^n_{B'}, T^n_{B'} + t_0]_{n\in\mathbb{N}}$ cubre los tiempos de visita de X a B' pues en $(T^n_{B'} + t_0\,,\, T^{n+1}_{B'})$, $X \notin B'$ por definición de $T^{n+1}_{B'}$. Ahora:

$$\begin{aligned}
U(0,B') = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\{X_t \in B'\}} dt\right] &= \mathbb{E}\left[\sum_{n=1}^\infty \mathbb{1}_{\{T^n_{B'} < \infty\}} \left(\int_{T^n_{B'}}^{T^n_{B'}+t_0} \mathbb{1}_{\{X_t\in B'\}} dt\right)\right] \\
&\leq \sum_{n=1}^\infty t_0 \mathbb{P}\left(T^n_{B'} < \infty\right)
\end{aligned} \tag{7.3.7}$$

Por lo tanto, bastaría probar que la serie converge para concluir. Probaremos por inducción que $\mathbb{P}\left(T^n_{B'} < \infty\right) < (1-\epsilon)^n$.
Para $n = 1$ como $B' \subset B$,

$$\mathbb{P}\left(T^1_{B'} < \infty\right) \leq \mathbb{P}\left((T_B)^{t_0} < \infty\right) < (1-\epsilon)$$

así que supongámoslo válido para n. Como:

$$T^{n+1}_{B'} = (T_{B'})^{T^n_{B'}+t_0} = T^n_{B'} + t_0 + T_{B'} \circ \theta_{T^n_{B'}+t_0} \quad \text{en } \{T^n_{B'} < \infty\} \tag{7.3.8}$$

y como $\{T^{n+1}_{B'} < \infty\} = \{T^{n+1}_{B'} < \infty\,,\, T^n_{B'} < \infty\} = \{T_{B'} \circ \theta_{T^n_{B'}+t_0} < \infty\,,\, T^n_{B'} < \infty\}$, por propiedad de Markov

$$\begin{aligned}
\mathbb{P}\left(T^{n+1}_{B'} < \infty\right) &= \mathbb{E}\left[\mathbb{1}_{\{T^n_{B'} < \infty\}} \mathbb{1}_{\mathbb{R}^+}(T_{B'}) \circ \theta_{T^n_{B'}+t_0}\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{T^n_{B'} < \infty\}} \mathbb{E}\left[\mathbb{1}_{\mathbb{R}^+}(T_{B'}) \circ \theta_{T^n_{B'}+t_0} | \mathscr{F}_{T^n_{B'}}\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_{\{T^n_{B'} < \infty\}} \mathbb{E}^{X_{T^n_{B'}}}\left[\mathbb{1}_{\mathbb{R}^+}(T_{B'}) \circ \theta_{t_0}\right]\right] \\
&= \int_\Omega \mathbb{P}^{X_{T^n_{B'}}}\left((T_{B'})^{t_0} < \infty\right) \mathbb{1}_{\{T^n_{B'} < \infty\}} d\mathbb{P}
\end{aligned}$$

Dividiendo ambos lados por $\mathbb{P}\left(T^n_{B'} < \infty\right)$ y cambio de variable

$$\mathbb{P}\left(T^{n+1}_{B'} < \infty | T^n_{B'} < \infty\right) = \int_{B'} \mathbb{P}^z\left((T_{B'})^{t_0} < \infty\right) d\mathbb{P}\left(X_{T^n_{B'}} \in dz \middle| T^n_{B'} < \infty\right). \tag{7.3.9}$$

Ahora, $\mathbb{P}^z\left((T_{B'})^{t_0} < \infty\right) = \mathbb{P}\left(\exists\, s \geq t_0\,;\, X_s + z \in B'\right) \leq \mathbb{P}\left(\exists s \geq t_0\,;\, X_s \in B\right)$ ya que $B' - z \subset B$ puesto que $z \in B'$. Por lo tanto, acotamos $\mathbb{P}^z\left((T_{B'})^{t_0} < \infty\right)$ por $\mathbb{P}\left((T_B)^{t_0} < \infty\right)$ para toda $z \in B'$ en (7.3.9). Al no depender este último término de z, obtenemos que:

$$\begin{aligned}\mathbb{P}\left(T_{B'}^{n+1} < \infty | T_{B'}^{n} < \infty\right) &\leq \mathbb{P}\left((T_B)^{t_0} < \infty\right)\\ &= (1-\epsilon).\end{aligned}$$

Multiplicando ambos lados por $\mathbb{P}\left(T_{B'}^{n} < \infty\right)$ que por hipótesis de inducción es igual a $(1-\epsilon)^n$, concluimos que:

$$\mathbb{P}\left(T_{B'}^{n+1} < \infty\right) \leq (1-\epsilon)^{n+1}. \tag{7.3.10}$$

Remplazando en (7.3.7), llegamos a que la serie converge y que por lo tanto $U(0, B') < \infty$.

(1) $\Rightarrow$ (3)
Por lo que se hizo en la implicación anterior la serie $\sum_{n=1}^{\infty} t_0 \mathbb{P}\left(T_{B'}^{n} < \infty\right)$ converge, y por lo tanto $\lim_{n\to\infty} \mathbb{P}^0\left(T_{B'}^{n} < \infty\right) = 0$. Esto es:

$$\lim_{n\to\infty} \mathbb{P}^0\left(\exists\, s > T_{B'}^{n-1} + t_0\,;\, X_s \in B'\right) = \lim_{n\to\infty} \mathbb{P}^0\left(T_{B'}^{n} < \infty\right) = 0.$$

Así,

$$\lim_{t\to\infty} \mathbb{P}^0\left(\exists\, s > t\,;\, X_s \in B'\right) = \lim_{t\to\infty} \mathbb{P}^0\left((T_{B'})^{t} < \infty\right) = 0 \tag{7.3.11}$$

y entonces para t_0 suficientemente grande, $\mathbb{P}^0\left((T_{B'})^{t_0} < \infty\right) \ll \delta$. Usando este hecho, probaremos que ocurre algo similar a (7.3.11) para cualquier posición inicial $x \in \mathbb{R}^d$ al considerar una bola más pequeña:
Sea B'' bola cerrada con centro en 0 tal que $B'' + B'' \subseteq B'$. Probaremos que

$$\lim_{t\to\infty} \mathbb{P}^x\left(\exists\, s > t\,;\, X_s \in B''\right) = \lim_{t\to\infty} \mathbb{P}^x\left((T_{B''})^{t} < \infty\right) = 0 \tag{7.3.12}$$

para cualquier $x \in E$. La idea es iniciar a X en x y hacerlo pasar por B''. Al condicionar en el primer tiempo de llegada a B'' y aplicar propiedad de Markov, reiniciamos al proceso en $z \in B''$. Regresar a B'' después de t_0 unidades de tiempo partiendo de un $z \in B''$ nos deja en una situación similar a la que teníamos en la implicación anterior. [4]

$$\begin{aligned}\mathbb{P}^x\left((T_{B''})^{T_{B''}+t_0} < \infty\right) &= \mathbb{P}^x\left((T_{B''})^{T_{B''}+t_0} < \infty\,,\, T_{B''} < \infty\right)\\ &= \mathbb{E}^x\left[\mathbb{1}_{\{T_{B''}<\infty\}} \mathbb{E}^x\left[\mathbb{1}_{\mathbb{R}^+}\left(T_{B''}\right) \circ \theta_{T_{B''}+t_0} | \mathscr{F}_{T_{B''}}\right]\right]\\ &= \mathbb{E}^x\left[\mathbb{1}_{\{T_{B''}<\infty\}} \mathbb{E}^{X_{T_{B''}}}\left[\mathbb{1}_{\mathbb{R}^+}\left(T_{B''}\right) \circ \theta_{t_0}\right]\right]\\ &= \int_{\Omega} \mathbb{P}^{X_{T_{B''}}}\left((T_{B''})^{t_0} < \infty\right) \mathbb{1}_{\{T_{B''}<\infty\}} d\mathbb{P}^x\end{aligned}$$

por lo cual

$$\mathbb{P}^x\left((T_{B''})^{T_{B''}+t_0} < \infty \,|\, T_{B''} < \infty\right) = \int_{B''} \mathbb{P}^z\left((T_{B''})^{t_0} < \infty\right) \mathbb{P}^x\left(X_{T_{B''}} \in dz \,|\, T_{B''} < \infty\right). \tag{7.3.13}$$

[4] Partíamos de 0 y regresábamos a B'. Nótese como para u, v cualesquiera en B'', $u + v \subseteq B'$.

Ahora, como $B'' - z \subseteq B'$ para todo $z \in B''$,

$$\begin{aligned}\mathbb{P}^z\left((T_{B''})^{t_0} < \infty\right) &= \mathbb{P}\left(\exists\, s \geq t_0\,;\, X_s + z \in B''\right)\\ &\leq \mathbb{P}\left(\exists\, s \geq t_0\,;\, X_s \in B'\right) = \mathbb{P}^0\left((T_{B'})^{t_0} < \infty\right)\end{aligned}$$

que no depende de z. Por esto, al acotar en (7.3.13) por $\mathbb{P}^0\left(T_{B''} < \infty\right)$ y al multiplicar ambos lados por $\mathbb{P}^x\left(T_{B''} < \infty\right)$ obtenemos:

$$\mathbb{P}^x\left((T_{B''})^{T_{B''}+t_0} < \infty, T_{B''} < \infty\right) \leq \mathbb{P}^0\left((T_{B'})^{t_0} < \infty\right)\mathbb{P}^x\left(T_{B''} < \infty\right) \ll \delta$$

y como $\{(T_{B''})^{T_{B''}+t_0} < \infty, T_{B''} < \infty\} = \{(T_{B''})^{T_{B''}+t_0} < \infty\}$, obtenemos que

$$\lim_{t\to\infty} \mathbb{P}^x\left((T_{B''})^{T_{B''}+t} < \infty\right) = 0.$$

Entonces, $\lim_{t\to\infty} \mathbb{P}^x\left((T_{B''})^{t} < \infty\right) = 0$ y por lo tanto,

$$\lim_{t\to\infty} \mathbb{P}^x\left(\exists\, s \geq t\,;\, X_s \in B''\right) = \lim_{t\to\infty} \mathbb{P}^x\left((T_{B''})^{t} < \infty\right) = 0.$$

Como x es cualquiera, obtenemos que para todo $x, y \in \mathbb{R}^d$,

$$\lim_{t\to\infty} \mathbb{P}^x\left(\exists\, s \geq t\,;\, X_s \in B'' + y\right) = 0. \tag{7.3.14}$$

Si K es un compacto, lo cubrimos con un número finito de bolas de la forma $B'' + y$, por lo que para cada $x \in \mathbb{R}^d$

$$\lim_{t\to\infty} \mathbb{P}^x\left(\exists\, s \geq t\,;\, X_s \in K\right) = 0.$$

Tomamos una sucesión de tiempos (t_n) tales que $\mathbb{P}^x\left(\exists\, s \geq t_n\,;\, X_s \in K\right) < \frac{1}{2^n}$. Como

$$\sum_n \mathbb{P}^x\left(\exists\, s \geq t_n\,;\, X_s \in K\right) < \infty,$$

por Borel-Cantelli X no regresa a K a partir de una t suficientemente grande $\mathbb{P}^x$-c.s. Como el compacto es arbitrario, $|X_s| \to \infty$ $\mathbb{P}^x$-c.s. Como esto es válido para todo $x \in \mathbb{R}^d$, $|X_s| \to \infty$ c.s.

Como (3) $\Rightarrow$ (1) es directo, concluimos. ■

Teorema 7.3.11 *Bajo* (H), *son equivalentes:*

(1) $\mathscr{R} = \mathbb{R}^d$

(2) X *es recurrente*

(3) *Para todo boreliano* K *con medida de Lebesgue positiva,*

$$U(x, K) = \infty \quad \text{para casi todo } x \in \mathbb{R}^d.$$

Demostración

Por los dos resultados previos, solo es necesario probar (2) $\Leftrightarrow$ (3) y como el recíproco es claro, probaremos (2) $\Rightarrow$ (3). Primero, probaremos que $U(x, B) = \infty$ para bolas B con centro en 0 (podemos asumir esto sin pérdida de generalidad por (7.3.1)) para cada $x \in \mathbb{R}^d$.

Sea B bola con centro en 0 y B' bola abierta con centro en 0 tal que $\overline{B'}+\overline{B'}\subseteq B$. Como X es recurrente $U(0,B')=\infty$. Por propiedad de Markov fuerte aplicada al primer tiempo de entrada a B'

$$\begin{aligned}
U(x,B)=\mathbb{E}^x\left[\int_0^\infty \mathbb{1}_{\{X_t\in B\}}dt\right] &\geq \mathbb{E}^x\left[\mathbb{1}_{\{T_{B'}<\infty\}}\int_{T_{B'}}^\infty \mathbb{1}_{\{X_t\in B\}}dt\right]\\
&= \mathbb{E}^x\left[\mathbb{1}_{\{T_{B'}<\infty\}}\int_0^\infty \mathbb{1}_B(X_t)\circ\theta_{T_{B'}}dt\right]\\
&= \mathbb{E}^x\left[\mathbb{1}_{\{T_{B'}<\infty\}}\mathbb{E}^x\left[\int_0^\infty \mathbb{1}_B(X_t)\circ\theta_{T_{B'}}dt|\mathscr{F}_{T_{B'}}\right]\right]\\
&= \mathbb{E}^x\left[\mathbb{1}_{\{T_{B'}<\infty\}}\mathbb{E}^{X_{T_{B'}}}\left[\int_0^\infty \mathbb{1}_B(X_t)dt\right]\right]\\
&= \int_\Omega \mathbb{E}^{X_{T_{B'}}}\left[\int_0^\infty \mathbb{1}_B(X_t)dt\right]\mathbb{1}_{\{T_{B'}<\infty\}}d\mathbb{P}^x
\end{aligned}$$

y por lo tanto

$$U(x,B)\geq \mathbb{P}^x\left(T_B'<\infty\right)\int_{B'}\mathbb{E}^z\left[\int_0^\infty \mathbb{1}_B(X_t)dt\right]\mathbb{P}^x\left(X_{T_{B'}}\in dz\,|\,T_{B'}<\infty\right).$$

Por construcción, como $z\in B'$, $B'+z\subset B$ y entonces $B'\subset B-z$. Al ser esto válido para todo $z\in B'$, acotamos inferiormente el integrando:

$$\mathbb{E}^z\left[\int_0^\infty \mathbb{1}_B(X_t)dt\right]=\mathbb{E}^0\left[\int_0^\infty \mathbb{1}_{B-z}(X_t)dt\right]\geq \mathbb{E}^0\left[\int_0^\infty \mathbb{1}_{B'}(X_t)dt\right]=U(0,B')$$

que no es finito. Así,

$$U(x,B)\geq \mathbb{P}^x\left(T_{B'}<\infty\right)U(0,B').$$

Por (H), $\mathbb{P}^x\left(T_{B'}<\infty\right)>0$ y concluimos que $U(x,B)=\infty$.

Ahora lo probaremos para compactos. Sean C y K compactos con medida de Lebesgue positiva. Por Fubini:

$$\begin{aligned}
\int_C U(x,K)\,dx=\int_C U(0,K-x)\,dx &= \int_C\int_{\mathbb{R}^d}\mathbb{1}_{K-x}(y)\,U(0,dy)\,dx\\
&= \int_C\int_{\mathbb{R}^d}\mathbb{1}_K(x+y)\,U(0,dy)\,dx\\
&= \int_{\mathbb{R}^d}\left(\int_C \mathbb{1}_K(x+y)\,dx\right)U(0,dy). \qquad (7.3.15)
\end{aligned}$$

Afirmamos que la función $y\mapsto\int_C \mathbb{1}_K(x+y)\,dx$ es continua y no es idénticamente 0.
La continuidad de

$$y\mapsto\lambda\left(C\cap(K-y)\right)=\int_C \mathbb{1}_K(x+y)\,dx$$

es consecuencia de la regularidad de la medida de Lebesgue: si B es una bola abierta y C un compacto, la aplicación

$$y\mapsto\int \mathbb{1}_C(x)\mathbb{1}_{B-y}(x)dx$$

es continua ya que, si $y_n\to y$, $\mathbb{1}_C(x)\mathbb{1}_{B-y_n}(x)\to\mathbb{1}_C(x)\mathbb{1}_{B-y}(x)$ tanto para todo $x\in(B-y)$ como para todo $x\in\mathbb{R}^d\setminus(\bar{B}-y)$. Como la frontera de $B-y$ tiene medida de Lebesgue 0, se sigue por Teorema de convergencia dominada que

$$\lambda(C\cap(B-y_n))=\int \mathbb{1}_C(x)\mathbb{1}_{B-y_n}(x)dx\to\int \mathbb{1}_C(x)\mathbb{1}_{B-y}(x)dx=\lambda(C\cap(B-y)) \qquad (7.3.16)$$

Esto vuelve a ser cierto si B es un abierto con frontera de medida nula por el mismo argumento. Ahora, si K es un compacto, por regularidad existe una unión finita de bolas $B \subset K$ tal que $\lambda(K \setminus B) < \epsilon$. Como $\lambda(\partial B) = 0$, al aplicar lo anterior y desigualdad triangular obtenemos que, para toda sucesión $y_n \to y$,

$$\begin{aligned}
&|\lambda(C \cap (K - y_n)) - \lambda(C \cap (K - y))| \\
&\quad \leq |\lambda(C \cap (K - y_n)) - \lambda(C \cap (B - y_n))| \\
&\qquad + |\lambda(C \cap (B - y_n)) - \lambda(C \cap (B - y))| \\
&\qquad\quad + |\lambda(C \cap (B - y)) - \lambda(C \cap (K - y))|
\end{aligned}$$

el primer y tercer término son menores que ϵ mientras que el segundo es menor que ϵ al tomar n suficientemente grande por (7.3.16). Probemos ahora lo segundo: al ser positiva bastaría con probar que su integral es estrictamente positiva, es decir que:

$$\int_{\mathbb{R}^d} \left(\int_C \mathbb{1}_K(x + y) dx \right) dy = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_C(x) \mathbb{1}_{K-y}(x) dx dy > 0.$$

Por Fubini e invariancia de la medida de Lebesgue λ bajo traslación:

$$\begin{aligned}
\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_C(x) \mathbb{1}_{K-y}(x) dx dy &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \mathbb{1}_C(x) \mathbb{1}_{K-x}(y) dy dx \\
&= \int_{\mathbb{R}^d} \mathbb{1}_C(x) \lambda(K - x) dx \\
&= \lambda(C) \lambda(K)
\end{aligned}$$

que es estrictamente mayor que 0.

Pero entonces, por continuidad existe una bola abierta B_y con centro en y tal que $u \to \int_C \mathbb{1}_K(x + u) dx > 0$ para todo $u \in B_y$. Ya probamos que $U(x, B) = \infty$ así que al usar este hecho en (7.3.15) obtenemos que

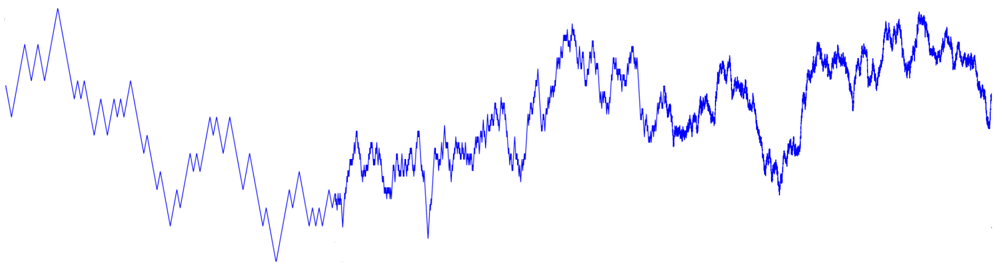
$$\int_C U(x, K)\, dx = \infty.$$

Como esto se cumple para cualquier compacto C con medida de Lebesgue positiva, por el lema 9.1.3 del Apéndice obtenemos que $U(x, K) = \infty$ para casi cualquier $x \in \mathbb{R}^d$. Si D es un boreliano con medida positiva, una vez más usando la regularidad de la medida de Lebesgue podemos hallar un compacto de medida positiva $K \subset D$ y por lo tanto, $U(x, D) = \infty$ para casi cualquier $x \in \mathbb{R}^d$. ■

Capítulo 8

Convergencia débil de procesos

En este capítulo, nos dedicamos al estudio de resultados de convergencia débil de procesos. Para esto, serán necesarios los resultados y definiciones de los capítulos 1 y 2; en particular todo lo relacionado con convergencia débil en espacios métricos. Recordemos que un proceso con trayectorias en $U \subset E^T$ se definió como una función medible de un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ en el espacio medible $(U, \mathscr{E}_U^T)$. Los casos que nos van a interesar serán cuando U sea metrizable y tal que bajo esta métrica, sus borelianos $\mathscr{B}(U)$ coincidan con la σ-álgebra de los cilindros $\mathscr{E}_U^T$ para que las definiciones que involucran medibilidad sean consistentes. Como se verá más delante, este será el caso de los espacios $C([0,T],E)$ y $D(\mathbb{R}^+,E)$, en los cuales trabajaremos.



Al ser U espacio métrico, tendrá sentido definir la convergencia débil de procesos como un caso particular de convergencia de elementos aleatorios con valores en un espacio métrico. Así,

Definición 8.0.12 *Sea $(X^n = (X_t^n : t \in [0,T]))$ con $T \in \mathbb{R}^+ \cup \{\infty\}$ una sucesión de procesos estocásticos con trayectorias en $U \subset E^T$, donde U es un espacio métrico con métrica d bajo la cual $\mathscr{B}(U) = \mathscr{E}_U^T$. Diremos que X^n converge débilmente a un proceso X con trayectorias en U si X_n converge débilmente a X en el espacio métrico (U,d) como se definió en 2.1.1. Esto es, la sucesión $(\mathbb{P}_{X_n})$ de medidas en U converge débilmente*

$$\mathbb{P}_{X_n} \xrightarrow{\mathscr{L}(U)} \mathbb{P}_X$$

en el espacio métrico (U,d). Esto también se denota como

$$X_n \xrightarrow{\mathscr{L}(U)} X.$$

Como se verá en la sección 8.1 al trabajar en $U = C([0,1],\mathbb{R})$, probar convergencia débil de procesos directamente resulta mucho más laborioso que probar convergencia débil de variables y vectores aleatorios. Esto motiva el desarrollo de herramientas más elaboradas para obtener criterios más eficientes que nos permitan probar convergencia de procesos. Para empezar, requeriremos de un estudio a mayor profundidad de la convergencia débil en espacios métricos, este es el contenido de la sección 8.2. Como los procesos con los que hemos trabajado tienen trayectorias en D, en la sección 8.3 construimos una métrica adecuada para este espacio, llamada la métrica de Skorokhod y cuya topología inducida en D se conoce como la topología J1 de Skorokhod. En la sección 8.4 probamos algunos criterios de convergencia de sucesiones de procesos (X_n) con trayectorias en D. Estos requieren a menudo verificar tensión de la sucesión de medidas $(\mathbb{P}_{X_n})$ y es por esto que en 8.5 veremos algunos criterios de tensión para sucesiones de medidas en D. Finalmente, en la sección 8.6 aplicamos todo esto al probar un resultado que caracteriza la convergencia de sucesiones de procesos de Feller.

8.1 El caso del movimiento browniano

Referencia: Morters and Peres [14] capítulo 5.

Empezamos con un primer resultado de convergencia débil de procesos estocásticos: probaremos la convergencia en ley de la caminata aleatoria interpolada al movimiento browniano. Este resultado es conocido como el teorema de invariancia de Donsker. La prueba que daremos no requiere de la teoría que desarrollaremos en las siguientes secciones y servirá para motivar lo que vendrá a continuación. Cabe notar que se pueden dar pruebas substancialmente más cortas al usar los resultados de las secciones 8.2 a 8.7. Aquí, $U = C[0,1]$ y $d = \|\cdot\|_\infty$ es la métrica uniforme. Requeriremos del

Teorema 8.1.1 Encaje de Skorokhod
Sea (B_t) un movimiento browniano, $(\mathscr{F}_t)$ su filtración natural y X una variable aleatoria con $\mathbb{E}[X] = 0$ y $\mathbb{E}[X^2] < \infty$. Entonces, existe un tiempo de paro T con respecto a $(\mathscr{F}_t)$ para el cual B_T tiene la misma ley que X.

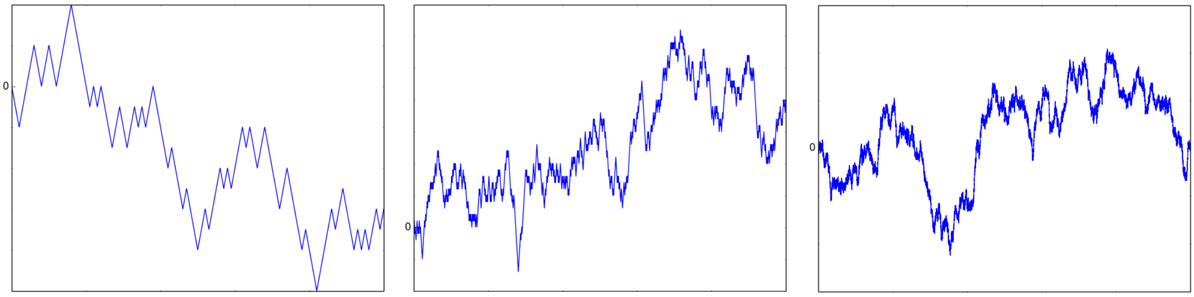
Corresponde al teorema 5.15 en [14]. En [14] se presentan dos pruebas una por Dubin y otra por Azéma y Yor. Para no desviarnos, asumiremos este resultado para no desviarnos. Pasaremos a la prueba de un primer resultado de convergencia de procesos conocido como el principio de invariancia de Donsker. Recordemos que el espacio $C[0,1]$ con la métrica uniforme $\|\cdot\|_\infty$ es un espacio métrico (completo y separable) y como ya se mencionó, un proceso estocástico continuo $X = (X_t : t \in [0,1])$ es una función medible de un espacio de probabilidad en el espacio $C[0,1]$ equipado con la σ-álgebra de los cilindros.

Sea (ξ_n) una sucesión de variables aleatorias idénticamente distribuidas y supongamos que $\mathbb{E}[\xi_n] = 0$ y $Var(\xi_n) = 1$. Esto se puede asumir sin pérdida de generalidad ya que sino podemos considerar la sucesión normalizada y centrada $(\xi_n - \mathbb{E}[\xi_n])/\sqrt{Var(\xi_n)}$. Consideremos a $S_n = \sum_{i=1}^n \xi_k$, la caminata aleatoria generada por la sucesión y a la interpolación de sus puntos:

$$S(t) = S_{\lfloor t\rfloor} + (t - \lfloor t\rfloor)(S_{\lfloor t\rfloor+1} - S_{\lfloor t\rfloor}) \quad t \in [0,\infty). \tag{8.1.1}$$

Es un proceso estocástico continuo y en base a este, definimos para cada $n \in \mathbb{N}$ al proceso reescalado $S_n^*(t) = \frac{S(nt)}{\sqrt{n}}$ para $t \in [0,1]$. Nótese que para cada n, $S_n^* : \Omega \to C[0,1]$ pues es un proceso continuo.

Teorema 8.1.2 Principio de Invariancia de Donsker
En el espacio métrico $(C[0,1], \|\cdot\|_\infty)$*, la sucesión* $(S_n^*)_n$ *converge en ley a un movimiento browniano* $B = (B_t)_{t\in[0,1]}$.



La prueba se apoya fuertemente en el siguiente lema:

Lema 8.1.3 *Supongamos que B es un movimiento browniano. Entonces, para cualquier variable aleatoria X con media 0 y varianza 1 existe una sucesión de tiempos de paro $0 = T_0 \leq T_1 \leq T_2 \leq \cdots$ con respecto a la filtración de B tales que*
a) La sucesión (B_{T_n}) tiene la misma distribución que la caminada aleatoria $S_n = \sum_{k=1}^n \xi_k$ con incrementos independientes ξ_k distribuidos como X.
b) La sucesión $(S_n^*)_n$ construida a partir de esta caminata aleatoria en particular satisface que

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{0\leq t\leq 1}\left|\frac{B_{nt}}{\sqrt{n}} - S_n^*(t)\right| > \epsilon\right) = 0 \tag{8.1.2}$$

Así, la idea es probar que podemos encajar cualquier caminata aleatoria centrada en B y luego probar la convergencia a B para la caminata encajada que construimos.

Demostración
Prueba de a): Sea $T_0 = 0$, por el Teorema del encaje de Skorokhod, existe un tiempo de paro T_1 tal que $\mathbb{E}[T_1] = 1$ y $B_{T_1} \sim X$ y por lo tanto, se tiene que

$$B_{T_1} \sim \sum_{k=1}^{1} \xi_k.$$

Por propiedad de Markov Fuerte, el browniano

$$B_t^{(2)} = B_{T_1+t} - B_{T_1} \quad t \geq 0$$

es independiente de $\mathscr{F}_{T_1}$ y en particular de T_1 y B_{T_1} al ser ambas $\mathscr{F}_{T_1}$-medibles. Definimos similarmente a T_2', tiempo de paro con respecto a la filtración de $B^{(2)}$ tal que $\mathbb{E}[T_2'] = 1$ y $B_{T_2'}^{(2)} \sim X$. Entonces, $T_2 := T_1 + T_2'$ es un tiempo de paro para el browniano original B (dejamos este detalle para el final de la prueba), $\mathbb{E}[T_2] = 2$ y además,

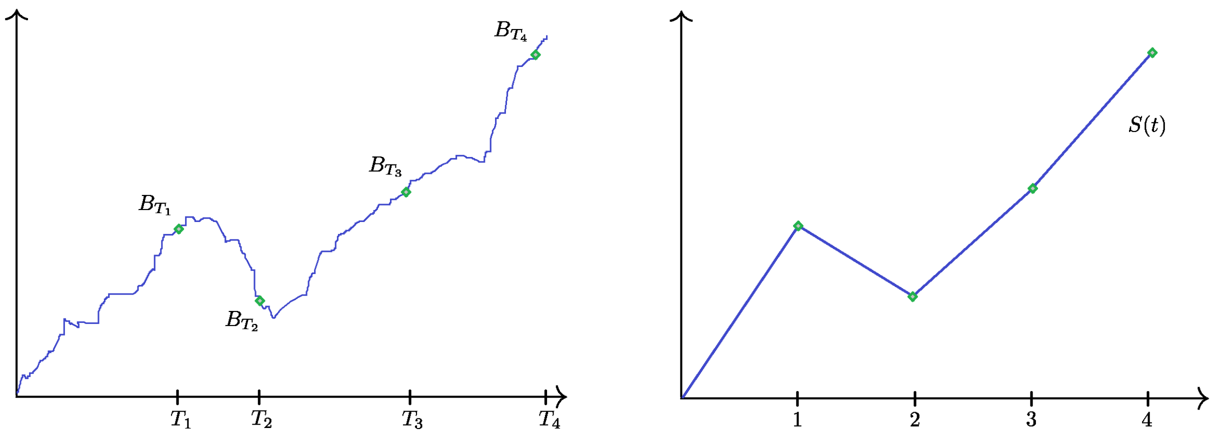
$$B_{T_2} \sim \sum_{k=1}^{2} \xi_k.$$

Esto se debe a que

$$B_{T_2} = B_{T_1+T_2'} = B_{T_1} + \left(B_{T_1+T_2'} - B_{T_1}\right) = B_{T_1} + B^{(2)}_{T_2'}$$

y al ser B_{T_1}, $B^{(2)}_{T_2'}$ independientes con misma distribución que X, se sigue que $B_{T_2} \sim \sum_{k=1}^{2} \xi_k$. Nótese que $T_2 - T_1 = T_2'$ que es independiente de $T_1 - T_0$ por ser tiempo de paro para $B^{(2)}$ y ambos incrementos tienen la misma distribución ya que T_2' y T_1 se construyeron exactamente de la misma forma para brownianos distintos. Podemos ahora considerar el browniano $(B^{(3)}_t = B_{T_2+t} - B_{T_2} : t \geq 0)$ y repetir el procedimiento.

Inductivamente, construimos una sucesión creciente de tiempos de paro con respecto a B, $0 = T_0 \leq T_1 \leq T_2 \leq \cdots$ tales que $\mathbb{E}[T_n] = n$ y $B_{T_n} \sim \sum_{k=1}^{n} \xi_n$ con lo que finalizamos con la prueba de (a). Se dice que la caminata aleatoria S_n está encajada en el movimiento browniano B. De aquí en adelante, trabajaremos con la interpolación $S(t)$ y la sucesión de reescalamientos $S_n^*(t)$ construidas a partir de esta versión en particular de la caminata aleatoria S_n.



Nótese que $T_n - T_{n-1} = T_n'$, que es un tiempo de paro con la misma distribución que $T_1 - T_0$ con respecto a

$$B^{(n)}_t = B^{(n-1)}_{T_{n-1}+t} - B_{T_{n-1}} \quad t \geq 0.$$

Este es un browniano independiente de $\mathscr{F}_{T_{n-1}}$ y por lo tanto de $\mathscr{F}_{T_k}$ para $k \leq n-1$ pues $\mathscr{F}_{T_k} \subset \mathscr{F}_{T_{n-1}}$ por monotonía de la sucesión de tiempos de paro. Por lo tanto, para cada n tenemos que $T_n - T_{n-1}$ es independiente de $\{T_k - T_{k-1}\}_{k \in 1 \cdots n-1}$ de lo cual deducimos que la sucesión de incrementos $\{T_k - T_{k-1}\}_{k \in \mathbb{N}}$ es independiente e idénticamente distribuida; esto se usará más adelante.

Antes de pasar a la prueba de (b), veamos que $T_1 + T_2'$ es tiempo de paro con respecto a la filtración de B. Si denotamos por $(\mathscr{F}_t)$ y $(\mathscr{F}^{(2)}_t)$ a las filtraciones naturales de B y $B^{(2)}$ respectivamente, notemos que

$$\mathscr{F}^{(2)}_t = \sigma\left(B_{T_1+s} - B_{T_1} : s \leq t\right) \subset \mathscr{F}_{T_1+t}.$$

Entonces, recordando que T_2' es un $\mathscr{F}^{(2)}_t$-tiempo de paro, para cada $t \in \mathbb{R}^+$ escribimos:

$$\{T_1 + T_2' > t\} = \bigcup_{r \in \mathbb{Q} \cap [0,t]} \{T_1 > r\} \cap \underbrace{\{T_2' > t - r\}}_{\in \mathscr{F}^{(2)}_{t-r}}.$$

Por la observación que se hizo, para toda $r \in \mathbb{Q} \cap [0, t]$, $\{T_2' > t - r\} \in \mathscr{F}_{T_1+t-r}$. Pero entonces, por definición de lo eventos pertenecientes a la σ-álgebra parada, se sigue que $\{T_1 > r\} \cap \{T_2' > t - r\} \in \mathscr{F}_t$ ya que

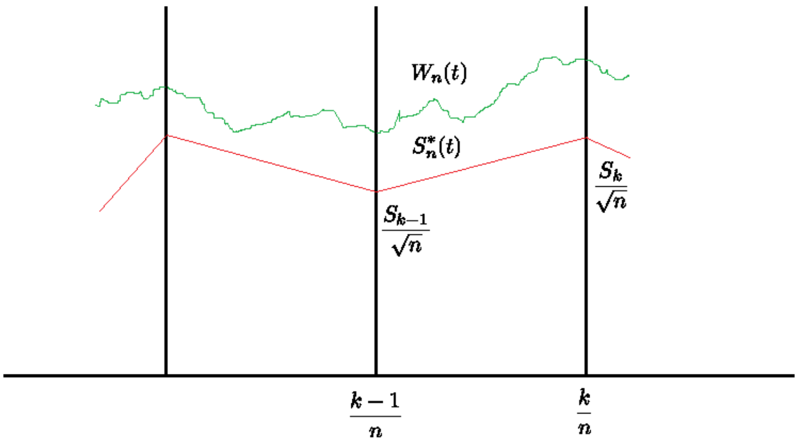
$$\{T_1 > r\} \cap \{T_2' > t - r\} = \{T_1 + t - r > t\} \cap \underbrace{\{T_2' > t - r\}}_{\in \mathscr{F}_{T_1+t-r}} \in \mathscr{F}_t$$

con lo que concluimos que $T_1 + T_2'$ es un $\mathscr{F}_t$-tiempo de paro.

Prueba de (b): Denotemos $W_n(t) = B_{nt}/\sqrt{n}$ y sea

$$A_n = \Big\{\exists\, t \in [0, 1] \text{ tal que } |W_n(t) - S_n^*(t)| > \epsilon\Big\}.$$

Probaremos que $\mathbb{P}(A_n) \to 0$. Sea $k = k(t)$ el único entero tal que $\frac{(k-1)}{n} \leq t \leq \frac{k}{n}$ y obsérvese que por construcción, $S_n^*(t)$ al restringirla a intervalos de la forma $[\frac{(k-1)}{n}, \frac{k}{n}]$ es lineal



con extremos puntos de la caminata aleatoria inicial escalados:

$$S_n^*\left(\frac{k}{n}\right) = \frac{S(k)}{\sqrt{n}} = \frac{S_k}{\sqrt{n}} \quad \text{y} \quad S_n^*\left(\frac{k-1}{n}\right) = \frac{S(k-1)}{\sqrt{n}} = \frac{S_{k-1}}{\sqrt{n}}.$$

Consideremos los dos eventos siguientes:

$$B_n = \Big\{\exists\, t \in [0, 1) \text{ tal que } \Big|W_n(t) - \frac{S_{k(t)}}{\sqrt{n}}\Big| > \epsilon\Big\}$$
$$C_n = \Big\{\exists\, t \in [0, 1) \text{ tal que } \Big|W_n(t) - \frac{S_{k(t)-1}}{\sqrt{n}}\Big| > \epsilon\Big\}$$

y veamos que $A_n \subset B_n \cup C_n$. Supongamos que existe $t \in [0, 1]$ tal que $|W_n(t) - S_n^*(t)| > \epsilon$ y sea $k = k(t)$ el único entero tal que $\frac{k-1}{n} \leq t \leq \frac{k}{n}$. Para cualquier t dentro de este intervalo, se cumple que $\lfloor nt \rfloor = k - 1$ pues $k - 1 \leq nt \leq k$ y por lo tanto:

$$\begin{aligned} S_n^*(t) = \frac{1}{\sqrt{n}} S(nt) &= \frac{1}{\sqrt{n}}\Big(S_{k-1} + (nt - (k-1))\,(S_k - S_{k-1})\Big) \\ &= \frac{1}{\sqrt{n}}\Big((k - nt)S_{k-1} + (nt - (k-1))S_k\Big) \\ &= \frac{1}{\sqrt{n}}\Big(aS_{k-1} + bS_k\Big) \end{aligned}$$

donde $a + b = 1$ y $a, b \geq 0$. Pero entonces, para el t en cuestión tenemos la siguiente cadena de desigualdades:

$$\begin{aligned}\epsilon < |S_n^*(t) - W_n(t)| &= \Big|\frac{S(nt)}{\sqrt{n}} - \frac{B_{nt}}{\sqrt{n}}\Big| \\ &= \Big|\frac{1}{\sqrt{n}}\Big(aS_{k-1} + bS_k\Big) - \frac{1}{\sqrt{n}}\Big(aB_{nt} + bB_{nt}\Big)\Big| \\ &\leq a\Big|\frac{S_{k-1}}{\sqrt{n}} - \frac{B_{nt}}{\sqrt{n}}\Big| + b\Big|\frac{S_k}{\sqrt{n}} - \frac{B_{nt}}{\sqrt{n}}\Big|\end{aligned}$$

por lo que $\Big|\frac{S_{k-1}}{\sqrt{n}} - \frac{B_{nt}}{\sqrt{n}}\Big| > \epsilon$ ó $\Big|\frac{S_k}{\sqrt{n}} - \frac{B_{nt}}{\sqrt{n}}\Big| > \epsilon$ pues $a + b = 1$. Con esto, obtenemos que $A_n \subset B_n \cup C_n$. Como $S_k = B_{T_k} = \sqrt{n}W_n(T_k/n)$, reescribimos

$$\begin{aligned}B_n &= \Big\{\exists\, t \in [0,1) \text{ tal que } \Big|W_n(t) - W_n(T_k/n)\Big| > \epsilon\Big\} \\ C_n &= \Big\{\exists\, t \in [0,1) \text{ tal que } \Big|W_n(t) - W_n(T_{k-1}/n)\Big| > \epsilon\Big\}\end{aligned}$$

y nótese que entonces

$$\begin{aligned}B_n \cup C_n \subset &\Big\{\exists\, s, t \in [0,1] \text{ tal que } |s-t| < \delta \text{ y } \Big|W_n(s) - W_n(t)\Big| > \epsilon\Big\} \\ &\cup \Big\{\exists\, t \in [0,1) \text{ tal que } \Big|\frac{T_k}{n} - t\Big| \vee \Big|\frac{T_{k-1}}{n} - t\Big| \geq \delta\Big\}\end{aligned}$$

pues tanto B_n como C_n están contenidos en esta unión. El primer conjunto, que denotaremos por U, no depende de n pues para cada n, W_n es un movimiento browniano por propiedad de escalamiento. Por continuidad uniforme en el $[0,1]$, podemos hacer $\mathbb{P}(U)$ arbitrariamente pequeño al considerar una δ adecuada. Si denotamos al segundo conjunto V_n, queda entonces probar que para cualquier δ fija, $\mathbb{P}(V_n) \to 0$. Al aplicar ley fuerte de los grandes números a la sucesión de incrementos independientes e idénticamente distribuidos con media uno $\{T_k - T_{k-1}\}_{k\in\mathbb{N}}$, obtenemos que

$$\lim_{n\to\infty} \frac{T_n}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n}(T_k - T_{k-1}) = 1 \text{ c.s.}$$

No es difícil comprobar que si una sucesión de números reales $\{a_n\}$ satisface

$$\lim_{n\to\infty}\frac{a_n}{n} = 1 \quad \text{entonces} \quad \lim_{n\to\infty}\sup_{0\leq k\leq n}\frac{|a_k - k|}{n} = 0$$

así que al aplicarlo a nuestro caso, obtenemos que $\lim_{n\to\infty}\sup_{0\leq k\leq n}\frac{|T_k-k|}{n} = 0$ c.s. y por lo tanto, tenemos convergencia en probabilidad a 0:

$$\lim_{n\to\infty}\mathbb{P}\left(\sup_{0\leq k\leq n}\frac{|T_k - k|}{n} \geq \delta\right) = 0$$

Usando esto, probemos que $\mathbb{P}(V_n) \to 0$; sea $n > \frac{2}{\delta}$. Recordemos que $\frac{k-1}{n} \leq t \leq \frac{k}{n}$ así que si para una t fija ocurre que $|T_k/n - t| > \delta$, necesariamente $|T_k/n - (k-1)/n| > \delta$ o $|T_k/n - k/n| > \delta$

y entonces

$$\begin{aligned}
\mathbb{P}\left(V_n\right) &= \mathbb{P}\left(\exists\, t \in [0,1) \text{ tal que } \left|\frac{T_k}{n} - t\right| \vee \left|\frac{T_{k-1}}{n} - t\right| \geq \delta\right) \\
&\leq \mathbb{P}\left(\sup_{1\leq k\leq n} \left(\frac{T_k}{n} - \frac{k-1}{n}\right) \vee \left(\frac{k}{n} - \frac{T_{k-1}}{n}\right) \geq \delta\right) \\
&\leq \mathbb{P}\left(\sup_{1\leq k\leq n} \left(\frac{T_k-(k-1)}{n}\right) \geq \delta\right) + \mathbb{P}\left(\sup_{1\leq k\leq n} \left(\frac{k-T_{k-1}}{n}\right) \geq \delta\right) \\
&\leq \mathbb{P}\left(\sup_{1\leq k\leq n} \left(\frac{T_k-k}{n}\right) \geq \frac{\delta}{2}\right) + \mathbb{P}\left(\sup_{1\leq k\leq n} \left(\frac{(k-1)-T_{k-1}}{n}\right) \geq \frac{\delta}{2}\right)
\end{aligned}$$

y ambos términos tienden a 0 por lo que se acaba de probar. Esto nos permite concluir que $\mathbb{P}\left(A_n\right) \to 0$ con lo que concluimos el lema. ■

Ahora, probaremos el Principio de Invariancia de Donsker:

Demostración

Manteniendo la notación del lema anterior, sea $(S_n^*)_n$ la sucesión construida a partir del encaje de Skorokhod en B y recordemos que para cualquier n, W_n es un movimiento browniano por propiedad de escalamiento. Sea $K \subset C[0,1]$ un cerrado y sea K_ϵ la ϵ-vecindad de K:

$$K_\epsilon = \{f \in C[0,1] : \|f-g\|_\infty \leq \epsilon \text{ para alguna } g \in K\}.$$

Entonces,

$$\begin{aligned}
\mathbb{P}\left(S_n^* \in K\right) &\leq \mathbb{P}\left(S_n^* \in K, \|S_n^* - W_n\|_\infty \leq \epsilon\right) + \mathbb{P}\left(S_n^* \in K, \|S_n^* - W_n\|_\infty > \epsilon\right) \\
&\leq \mathbb{P}\left(W_n \in K_\epsilon\right) + \mathbb{P}\left(\|S_n^* - W_n\|_\infty > \epsilon\right)
\end{aligned}$$

pero $\mathbb{P}\left(W_n \in K_\epsilon\right) = \mathbb{P}\left(B \in K_\epsilon\right)$ y $\mathbb{P}\left(\|S_n^* - W_n\|_\infty > \epsilon\right) \to 0$ por el lema anterior. Además, al ser K un cerrado:

$$\lim_{\epsilon \downarrow 0} \mathbb{P}\left(B \in K_\epsilon\right) = \mathbb{P}\left(B \in \bigcap_{\epsilon>0} K_\epsilon\right) = \mathbb{P}\left(B \in K\right)$$

por lo que $\limsup_{n\to\infty} \mathbb{P}\left(S_n^* \in K\right) \leq \mathbb{P}\left(B \in K\right)$ para cualquier cerrado K. Por el teorema del Portemanteau, $S_n \xrightarrow{\mathscr{L}} B$. ■

8.2 Convergencia débil en espacios métricos

8.2.1 La métrica de Prohorov en $\mathcal{P}(E)$

Referencia: Ethier, Kurtz [20] capítulo III, sección 1.

Sea (E,d) un espacio métrico, $\mathscr{B}(E)$ la σ-álgebra de sus borelianos y sea $\mathscr{C}$ la colección de cerrados de E. Para $F \subset E$, denotaremos por F^ϵ a su ϵ-vecindad abierta, es decir, al conjunto $F^\epsilon = \{x \in E : d(x,F) < \epsilon\}$. El objetivo será metrizar al espacio $\mathcal{P}(E)$ de modo que convergencia de medidas $\mu_n, \mu \in \mathcal{P}(E)$ bajo la métrica ρ

$$\rho(\mu_n, \mu) \to 0 \quad \text{ en } (\mathcal{P}(E), \rho)$$

sea equivalente a la convergencia débil de medidas usual:

$$\mu_n \overset{\mathscr{L}(E)}{\to} \mu.$$

Para cada μ, $\nu \in \mathcal{P}(E)$, definimos la métrica de Prohorov $\rho(\mu,\nu)$ como sigue:

$$\rho(\mu,\nu) = \inf\{\epsilon > 0 : \mu(F) \leq \nu(F^\epsilon) + \epsilon \text{ para todo } F \in \mathscr{C}\} \tag{8.2.1}$$

Observación: Nótese que de haber considerado $\inf\{\epsilon > 0 : \mu(F) \leq \nu(F^\epsilon)$ para todo $F \in \mathscr{C}\}$ en lugar de (8.2.1) podríamos obtener valores infinitos: por ejemplo, si $\mu, \nu \in \mathcal{P}(\mathbb{R})$, μ con densidad gausiana y $\nu = \delta_0$ la masa puntual en 0 y K es un compacto que contiene al 0, $\mu(K^\epsilon) < 1 = \nu(K)$ para todo $\epsilon > 0$. Sumar ϵ como en (8.2.1) garantiza que $\rho \leq 1$.

Para probar que ρ es una métrica, necesitaremos el siguiente lema:

Lema 8.2.1 *Sean $\mu, \nu \in \mathcal{P}(E)$ y $\alpha, \beta > 0$. Si*

$$\mu(F) \leq \nu(F^\alpha) + \beta \tag{8.2.2}$$

para todo $F \in \mathscr{C}$, entonces

$$\nu(F) \leq \mu(F^\alpha) + \beta \tag{8.2.3}$$

para todo $F \in \mathscr{C}$.

Demostración

Supongamos que se satisface (8.2.2) para todo $F \in \mathscr{C}$. Para probar que se cumple (8.2.3), debemos verificar que para $F_1 \in \mathscr{C}$ cualquiera, se cumple

$$\nu(F_1) \leq \mu(F_1^\alpha) + \beta. \tag{8.2.4}$$

Sea $F_2 = S - F_1^\alpha$. Nótese que $F_2 \in \mathscr{C}$ y por construcción, $d(F_1, F_2) \geq \alpha$. Pero entonces, $F_1 \subset S - F_2^\alpha$ y por (8.2.2) para $F = F_2$, se sigue que

$$\mu(F_1^\alpha) = 1 - \mu(F_2) \geq 1 - \nu(F_2^\alpha) - \beta \geq \nu(F_1) - \beta$$

con lo que obtenemos (8.2.4). ■

Lema 8.2.2 *Así definida, ρ es una métrica de $\mathcal{P}(E)$.*

Demostración

Al aplicar el lema anterior para $\alpha = \beta = \epsilon$ obtenemos que $\rho(\mu,\nu) = \rho(\nu,\mu)$ para toda $\mu, \nu \in \mathcal{P}(E)$. Ahora, si $\rho(\mu,\nu) = 0$, para cualquier ϵ se satisface $\mu(F) \leq \nu(F^\epsilon) + \epsilon$ y como $F^\epsilon \downarrow F$ por ser F un cerrado,

$$\mu(F) \leq \lim_{\epsilon \downarrow 0} \nu(F^\epsilon) + \epsilon$$

por lo que $\mu(F) \leq \nu(F)$. Análogamente se obtiene que $\mu(F) \geq \nu(F)$ y por lo tanto las medidas coinciden para todo $F \in \mathscr{C}$. Como la clase de los cerrados $\mathscr{C}$ es estable ante intersecciones finitas y genera a $\mathscr{B}(E)$, al aplicar el lema de clases monótonas obtenemos que $\mu = \nu$ y nótese que si $\mu = \nu$, $\rho(\mu,\nu) = 0$. Para concluir que ρ es métrica, solo falta probar que cumple la desigualdad del triángulo. Sean μ, ν y $\lambda \in \mathcal{P}(E)$ y supongamos que $\rho(\mu,\nu) < \delta$ y $\rho(\nu,\lambda) < \epsilon$. Entonces, para cualquier $F \in \mathscr{C}$

$$\begin{aligned}\mu(F) &\leq \nu(F^\delta) + \delta \leq \nu(\overline{F^\delta}) + \delta \\ &\leq \nu((\overline{F^\delta})^\epsilon) + \delta + \epsilon \leq \nu(F^{\delta+\epsilon}) + \delta + \epsilon\end{aligned}$$

donde la última desigualdad se sigue de notar que, como $\overline{F^\delta} \subset \{x : d(x,F) \leq \delta\}$, entonces

$$(\overline{F^\delta})^\epsilon \subset \{x : d(x,F) \leq \delta\}^\epsilon \subset \{x : d(x,F) < \delta + \epsilon\} = F^{\delta+\epsilon}.$$

Con esto concluimos que $\rho(\mu,\nu) \leq \delta + \epsilon$ y por lo tanto ρ satisface la desigualdad triangular. ■

Observación: La distancia se basa en comparar las medidas en cerrados: si $\rho(\mu,\nu) < \epsilon$, entonces para todo $F \in \mathscr{C}$ se satisface $\mu(F) \leq \nu(F^\epsilon) + \epsilon$ así como su expresión simétrica y por lo tanto

$$|\mu(F) - \nu(F)| \leq \epsilon + \nu(F^\epsilon \setminus F) \vee \mu(F^\epsilon \setminus F)$$

donde el último término de la derecha tiende a 0 conforme $\epsilon \downarrow 0$ puesto que $F^\epsilon \downarrow F$ por ser F un cerrado. También obsérvese que si F es un compacto tal que $1 - \epsilon \leq \mu(F)$ y $\rho(\mu,\nu) < \epsilon$, entonces $1 - \epsilon \leq \mu(F) \leq \nu(F^\epsilon) + \epsilon$ y por lo tanto $1 - 2\epsilon \leq \nu(F^\epsilon)$. Es decir, a grandes rasgos aislar la masa de μ permite aislar la de ν. Esto se usará por ejemplo al probar el teorema de Prohorov, ya que como ocurrirá a menudo en las próximas pruebas, para probar tensión de una colección $\mathscr{M}$, se prueba tensión para un subconjunto que dista "poco" con respecto a ρ de $\mathscr{M}$ y se aplica lo anterior.

Ahora, probamos que convergencia bajo la métrica de Prohorov es equivalente a convergencia débil de medidas. Con esto completamos el teorema del Portemanteau estudiado en el capítulo 1 para espacios métricos separables:

Teorema 8.2.3 Portemanteau en métricos separables
Si μ_n, μ son medidas de probabilidad en un espacio métrico separable E, entonces las siguientes afirmaciones son equivalentes a la convergencia débil $\mu_n \xrightarrow{\mathscr{L}} \mu$:
0- $\lim_{n\to\infty} \rho(\mu_n,\mu) = 0$.

1- $\lim_{n\to\infty} \int_E f d\mu_n = \int_E f d\mu$ para toda función $f : E \to \mathbb{R}$ continua y acotada.

2- $\limsup_{n\to\infty} \mu_n(K) \leq \mu(K)$ para todo K cerrado.

3- $\liminf_{n\to\infty} \mu_n(G) \geq \mu(G)$ para todo G abierto.

4- $\lim_{n\to\infty} \mu_n(A) = \mu_X(A)$ si $\mu(\partial A) = 0$.

Si $E = \mathbb{R}^d$ y F es la distribución de μ se tiene además:
5- $\lim_{n\to\infty} F_{\mu_n}(u) = F_\mu(u)$ para todo punto u de continuidad de F_μ.

Demostración
Por lo que se hizo en el capítulo 1, 1-4 son equivalentes a la convergencia débil y solo falta probar que la convergencia en la métrica de Prohorov $\rho(\mu_n,\mu) \to 0$ es equivalente a $\mu \xrightarrow{\mathscr{L}} \mu$.

Probemos que $0 \Rightarrow 1$. Para cada n, sea $\epsilon_n = \rho(\mu_n, \mu) + 1/n$. Nótese que, por definición de ρ,

$$\mu_n(F) \leq \mu(F^{\epsilon_n}) + \epsilon_n = \mu(F^{\epsilon_n}) + \rho(\mu_n, \mu) + 1/n \quad \text{para todo } F \in \mathscr{C}.$$

Consideremos primero $f : E \to \mathbb{R}$ acotada y positiva. Por el teorema de cambio de variable y Fubini:

$$\begin{aligned}\int_E f d\mu_n = \int_0^\infty x\, \mu_n(f \in dx) &= \int_0^\infty \int_0^x ds\, \mu_n(f \in dx)\\ &= \int_0^{\|f\|_\infty} \mu_n(f \geq s) ds \leq \int_0^{\|f\|_\infty} \mu(\{f \geq s\}^{\epsilon_n}) ds + \epsilon_n \|f\|_\infty\end{aligned}$$

pues $F = \{f \geq s\}$ es un cerrado. Por esto mismo, $\{\{f \geq s\}^{\epsilon_n}\} \downarrow \{f \geq s\}$ y entonces por el teorema de convergencia dominada: $\int_0^\infty \mu(\{f \geq s\}^{\epsilon_n}) ds \to \int_0^\infty \mu(\{f \geq s\}) ds$. Así, si f es positiva y acotada obtenemos que

$$\limsup_n \int_E f d\mu_n \leq \int_0^{\|f\|_\infty} \mu(f \geq s) ds = \int_E f d\mu.$$

Ahora, si f es solo acotada, al aplicar lo anterior a las funciones positivas $\|f\|_\infty + f$ y $\|f\|_\infty - f$ se sigue que

$$\limsup_n \int_E \|f\|_\infty + f d\mu_n \leq \int_E \|f\|_\infty + f d\mu \quad \text{y} \quad \limsup_n \int_E \|f\|_\infty - f d\mu_n \leq \int_E \|f\|_\infty - f d\mu$$

con lo que concluimos:

$$\int_E f d\mu \leq \liminf_n f\mu_n \leq \limsup_n f\mu_n \leq \int_E f\mu.$$

Ahora, probemos que $3 \Rightarrow 0$. Sea $\epsilon > 0$. Como E es separable, existe $E_1, E_2, \dots \in \mathscr{B}(E)$ una partición de S con diámetro$(E_i) < \epsilon/2$ para $i = 1, 2, \dots$. Sea N el menor entero positivo tal que $\mu(\bigcup_{i=1}^N E_i) \geq 1 - \epsilon/2$. Consideremos la clase $\mathscr{G}$ compuesta por abiertos de la forma $\left(\bigcup_{i \in I} E_i\right)^{\epsilon/2}$ para $I \subset \{1, 2, \dots, N\}$; obsérvese que $\mathscr{G}$ es de cardinalidad finita. Como se satisface 3, para cada $G \in \mathscr{G}$ se cumple que $\mu(G) - \epsilon/2 \leq \mu_n(G)$ a partir de n suficientemente grande y por lo tanto, existe n_0 tal que $\mu(G) \leq \mu_n(G) + \epsilon/2$ si $n \geq n_0$ para todo $G \in \mathscr{G}$. Si F es un cerrado cualquiera, consideremos a su aproximación por elementos de la partición:

$$F_0 = \bigcup_{E_i \cap F \neq \emptyset} E_i.$$

Nótese que $\mu(F) = \mu(F_0) + \mu(F \cap (S \setminus \bigcup_{i=1}^N E_i)) \leq \mu(F_0^{\epsilon/2}) + \epsilon/2$ y además, por construcción, $F_0^{\epsilon/2}$ es un elemento de $\mathscr{G}$. Se sigue entonces que

$$\begin{aligned}\mu(F) &\leq \mu(F_0^{\epsilon/2}) + \epsilon/2\\ &\leq \mu_n(F_0^{\epsilon/2}) + \epsilon \leq \mu_n(F^\epsilon) + \epsilon\end{aligned}$$

donde en la última desigualdad se está utilizando que si $x \in F_0^{\epsilon/2}$,

$$d(x, F) \leq d(x, F_0) + d(F_0, F) < \epsilon/2 + \epsilon/2$$

pues el diámetro$(E_i) < \epsilon/2$. Con esto, concluimos que $\mu(F) \leq \mu_n(F^\epsilon) + \epsilon$ y por lo tanto $\rho(\mu_n, \mu) \to 0$. ■

Observación: Nótese que solo usamos la separabilidad para $3 \Rightarrow 0$, el resto de las implicaciones se siguen valiendo para E un espacio métrico arbitrario.

Obtuvimos que $(\mathcal{P}(E), \rho)$ es entonces un espacio métrico y cuando E es separable, la convergencia débil $\mu_n \overset{\mathscr{L}}{\to} \mu$ es equivalente a la convergencia $\mu_n \to \mu$ en $\mathcal{P}(E)$ con la topología de la métrica de Prokhorov, también conocida como la topología débil. Recordemos que en un espacio métrico S, un subconjunto $K \subset S$ es relativamente compacto si su cerradura es compacta y esto es equivalente a que de toda sucesión $\{a_n\} \in K$ se pueda extraer una subsucesión $\{a_{n_k}\}$ convergente. Así, la definición que dimos de compacidad relativa de una colección de medidas de probabilidad $\mathscr{M} \subset \mathcal{P}(E)$ en 2.1.3 es equivalente a que sea relativamente compacta como subconjunto de $(\mathcal{P}(E), \rho)$, lo que justifica el nombre. Se puede decir más sobre las propiedades de este espacio al imponer condiciones al espacio E:

Teorema 8.2.4 *Sea E un espacio métrico.*
1. Si E es separable, entonces $(\mathcal{P}(E), \rho)$ es separable.
2. Si E es separable y completo, entonces $(\mathcal{P}(E), \rho)$ es separable y completo.

La prueba se puede leer por ejemplo en Billinglsey [8] capítulo 1.

8.2.2 El Teorema de Prohorov

Referencia: Kurtz [20] capítulo III sección 2.

Recordemos que en un espacio métrico completo,
1. $K \subset E$ es compacto si y solo si es cerrado y totalmente acotado (esto es, para todo $\epsilon > 0$, K se puede cubrir por un número finito de bolas abiertas de radio ϵ) y
2. $K \subset E$ es relativamente compacto si y solo si es totalmente acotado.

Lema 8.2.5 *Si (E, d) es completo y separable, entonces para cada $\mu \in \mathcal{P}(E)$, la colección que consta de una única medida $\mathscr{M} = \{\mu\}$, es tensa.*

Demostración
Como E es separable, para todo $1/n$ podemos cubrir a E con una cantidad numerable de bolas de radio $1/n$, es decir $E = \bigcup_{i=1}^{\infty} B(x_i, 1/n)$. Para cada $n \geq 1$, existe entonces $R_n \in \mathbb{N}$ tal que

$$\mu\left(\left\{\bigcup_{i=1}^{R_n} B(x_i, 1/n)\right\}^c\right) \leq \frac{\epsilon}{2^n}.$$

Consideremos al conjunto K, definido como $\bigcap_{n\geq 1} \bigcup_{i=1}^{R_n} B(x_i, 1/n)$. Por construcción, para todo $\epsilon > 0$, K se puede cubrir por una cantidad finita de bolas de radio ϵ y al ser E completo, K es relativamente compacto. Además, satisface que

$$\mu(K^c) = \mu\left(\bigcup_{n\geq 1}\left\{\bigcup_{i=1}^{R_n} B(x_i, 1/n)\right\}^c\right) \leq \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon$$

con lo que concluimos que $\{\mu\}$ es tensa. ∎

Con esto, nos pasamos a uno de los resultados más importantes del capítulo:

Teorema 8.2.6 Prohorov en espacios polacos
Sea (E,d) un espacio métrico completo y separable y sea $\mathscr{M} \subset \mathcal{P}(E)$. Entonces, son equivalentes:
(i) $\mathscr{M}$ es tensa.
(ii) $\mathscr{M}$ es relativamente compacta con respecto a la topología de Prohorov en $\mathcal{P}(E)$.

Demostración
$(i) \Rightarrow (ii)$ Supongamos que $\mathscr{M}$ es tensa. Como por el teorema 8.2.4 el espacio $(\mathcal{P}(E),\rho)$ es completo, para probar que $\mathscr{M} \subset (\mathcal{P}(E),\rho)$ es relativamente compacto bastaría probar que $\mathscr{M}$ es totalmente acotado. Nos damos a la tarea de construir una colección finita de centros y bolas en $\mathcal{P}(E)$ de modo que $\mathscr{M}$ se quede contenido en la unión de estas: dado $\delta > 0$, necesitamos construir un subconjunto finito $\Lambda \subset \mathcal{P}(E)$ tal que

$$\mathscr{M} \subset \bigcup_{\mu\in\Lambda} \{\nu \in \mathcal{P}(E) : \rho(\nu,\mu) < \delta\}, \tag{8.2.5}$$

conjunto formado por la unión de las bolas con centro $\mu \in \Lambda$ y radio δ.

Sea $0 < \epsilon < \delta/2$ y sea $K \subset E$ un compacto tal que $\inf_{\nu\in\mathscr{M}} \nu(K) \geq 1-\epsilon$. Como E es completo y K compacto, existen $\{x_0,x_1,\ldots,x_n\} \in E$ tales que $K \subset \bigcup_{i=1}^{n} B(x_i,2\epsilon)$. Fijamos un punto $x_0 \in E$ y un natural $m \geq n/\epsilon$. Definimos a Λ como la colección finita de medidas de probabilidad atómicas de la forma:

$$\Lambda = \Big\{ \sum_{i=0}^{n} \left(\frac{k_i}{m}\right) \delta_{x_i} : k_i \in \mathbb{N} \text{ y } k_0 + \cdots + k_n = m \Big\}.$$

Nótese como la cardinalidad de Λ crece conforme m es grande. Sea $\nu \in \mathscr{M}$ cualquiera, hay que probar que se queda contenida en una bola de radio δ y centro $\mu \in \Lambda$; exhibimos a μ a continuación al definir sus coeficientes k_i. Sea $k_i = \lfloor m\nu(E_i)\rfloor$ para $i = 1,\ldots,n$ donde $E_i = B(x_i,2\epsilon) - \bigcup_{j=1}^{i-1} B(x_j,2\epsilon)$, $k_0 = m - (k_1+\cdots+k_n)$ y consideremos entonces al elemento de Λ siguiente:

$$\mu = \sum_{i=1}^{n} \frac{\lfloor m\nu(E_i)\rfloor}{m}\delta_{x_i} + \frac{k_0}{m}\delta_{x_0}.$$

Se puede pensar como una aproximación discreta de ν construida a partir de la partición $\{E_i\}$. Como $m\nu(E_i) - \lfloor m\nu(E_i)\rfloor \leq 1$, entonces

$$\nu(E_i) \leq \lfloor m\nu(E_i)\rfloor/m + 1/m \leq \nu(E_i) + 1/m$$

por lo que m dicta que tan buena es la aproximación, y nótese que $\cup_i E_i = \cup_i B(x_i,2\epsilon)$. Así, si

$F \in \mathscr{C}$,

$$\begin{aligned}
\nu(F) &\leq \nu\left(F \cap \bigcup_{i=1}^{n} B(x_i, 2\epsilon)\right) + \epsilon \\
&\leq \sum_{F\cap E_i \neq \emptyset} \nu(E_i) + \epsilon \\
&\leq \sum_{F\cap E_i \neq \emptyset} \left(\frac{\lfloor m\nu(E_i)\rfloor}{m} + \frac{1}{m}\right) + \epsilon \\
&\leq \sum_{F\cap E_i \neq \emptyset} \frac{\lfloor m\nu(E_i)\rfloor}{m} + \frac{n}{m} + \epsilon \\
&\leq \sum_{F\cap E_i \neq \emptyset} \frac{\lfloor m\nu(E_i)\rfloor}{m} \delta_{x_i}(F^{2\epsilon}) + 2\epsilon
\end{aligned}$$

donde en la última igualdad se está usando que si $F \cap B(x_i, 2\epsilon) \neq \emptyset$, entonces $B(x_i, 2\epsilon) \subset F^{2\epsilon}$ y por lo tanto $\delta_{x_i}(F^{2\epsilon}) = 1$. Esto nos lleva a que $\nu(F) \leq \mu(F^{2\epsilon}) + 2\epsilon$ para todo $F \in \mathscr{C}$ con lo cual concluimos que $\rho(\nu, \mu) \leq 2\epsilon < \delta$ y obtenemos (8.2.5).

(ii) $\Rightarrow$ *(i)* Sea $\epsilon > 0$. Como $\mathscr{M}$ es relativamente compacto en $(\mathcal{P}(E), \rho)$ espacio métrico completo, entonces es totalmente acotado. Así, para cada $n \in \mathbb{N}$ existe un subconjunto finito $\Lambda^n \subset \mathcal{P}(E)$ tal que $\mathscr{M}$ se queda totalmente contenido en la unión de las bolas de radio $\epsilon/2^{n+1}$ y centro $\mu \in \Lambda^n$, es decir

$$\mathscr{M} \subset \bigcup_{\mu \in \Lambda^n} \{\nu \in \mathcal{P}(E) : \rho(\nu, \mu) < \epsilon/2^{n+1}\}. \tag{8.2.6}$$

Por el lema 8.2.5, cada $\{\mu\}$ es tensa y como Λ^n tiene cardinalidad finita, para cada n existe un compacto K_n tal que

$$\inf_{\mu \in \Lambda^n} \mu(K_n) \geq 1 - \epsilon/2^{n+1}. \tag{8.2.7}$$

Como se satisface (8.2.6), para cada $\nu \in \mathscr{M}$ existe $\mu_n \in \Lambda^n$ tal que $\rho(\nu, \mu_n) \leq \epsilon/2^{n+1}$ y entonces por definición de ρ y (8.2.7) se satisface que

$$\nu(K_n^{\epsilon/2^{n+1}}) \geq \mu_n(K_n) - \epsilon/2^{n+1} \geq 1 - \epsilon/2^n.$$

Si definimos a K como la cerradura de $\cap_{n\geq 1} K_n^{\epsilon/2^{n+1}}$, obtenemos que es un compacto al ser cerrado y estar contenido en el compacto $K_1^{\epsilon/4}$ y satisface que

$$\nu(K) \geq \nu\left(\cap_{n\geq 1} K_n^{\epsilon/2^{n+1}}\right) \geq 1 - \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = 1 - \epsilon.$$

■

8.2.3 El Teorema de representación de Skorokhod y del mapeo continuo

Referencia: Ethier, Kurtz [20] capítulo III secciones 1 y 7, Zhan Shi [23] chapitre 2.

En esta sección, nos limitamos a enunciar el teorema de Representación de Skorokhod, resultado probado inicialmente por Skorokhod en 1956 (para el caso S un espacio métrico completo y separable) y en 1968 por Dudley para cuando S es solamente completo. Probaremos algunas consecuencias que se deducen fácilmente de este. El Teorema de representación permite a menudo simplificar significativamente algunas pruebas así como debilitar hipótesis de algunos teoremas. Por ejemplo, permitió probar en 1964 una versión más fuerte del principio de invariancia de Donsker, conocido como el principio de invariancia fuerte, en el cual se prueba la convergencia casi segura de la sucesión al considerar un espacio de probabilidad adecuando (Strassen [21], 1964).

Teorema 8.2.7 Representación de Skorokhod
Sea (S,d) un espacio métrico separable. Supongamos que μ_n, $n = 1, 2, \ldots$ y $\mu \in \mathcal{P}(S)$ satisfacen $\mu_n \overset{\mathscr{L}(S)}{\to} \mu$. Entonces, existe un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ en el cual están definidas variables aleatorias X_n, $n = 1, 2, \ldots$ y X con distribución μ_n, $n = 1, 2, \ldots$ y μ respectivamente tales que $\lim_{n\to\infty} X_n \to X$ $\mathbb{P}$-c.s.

Demostración
La prueba se puede leer en Ethier, Kurtz [20] capítulo III, teorema 1.8. ■

Cuando S es separable y completo, podemos además usar como espacio de probabilidad a $(\Omega, \mathscr{F}, \mathbb{P}) = ([0,1], \mathscr{B}[0,1], \lambda)$ donde λ es la medida de Lebesgue (Ikeda, Watanabe [22] capítulo 1 teorema 2.7) y en el caso $E = \mathbb{R}$ la prueba no es tan compleja, como se puede ver en [1], capítulo 7. Aplicaciones sencillas del teorema de representación son versiones del lema de Fatou y el teorema de convergencia dominada para convergencia débil:

Corolario 8.2.8 *Consideremos variables aleatorias X_n, X con valores en $E = \mathbb{R}$. Si $X_n \overset{\mathscr{L}}{\to} X$ entonces $\mathbb{E}[|X|] \leq \liminf \mathbb{E}[|X_n|]$. Si además la colección (X_n) es uniformemente integrable, entonces X es integrable y $\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$.*

Ambas pruebas consisten en pasarse al espacio descrito en 8.2.7 y aplicar los resultados usuales para convergencia casi segura (Ash [1] capítulo 7, ejercicios 2 y 3). Nosotros usarmeos Skorokhod para probar el *Teorema del mapeo continuo*, el cual necesitaremos más adelante. Recordemos que si $(E, \mathscr{E}, \mu)$ es un espacio de medida, $(E', \mathscr{E}')$ un espacio medible y $h : E \to E'$ es medible, entonces

$$\mu \circ h^{-1} : \mathscr{B}(E') \to \mathbb{R}, \qquad \mu \circ h^{-1}(B') = \mu(h^{-1}(B')) = \mu(x \in E : h(x) \in B')$$

es una medida en E'. En el caso de elementos aleatorios definidos en $(\Omega, \mathscr{F}, \mathbb{P})$ con valores en E', hemos estado denotado a esta medida por $\mathbb{P} \circ X^{-1} = \mathbb{P}_X$, la medida inducida por X o ley de X en E'.

Teorema 8.2.9 Teorema del mapeo continuo
Sean (E,d) y (E',d') espacios métricos separables y $h : E \to E'$ boreliana. Supongamos que μ_n, $\mu \in \mathcal{P}(E)$ satisfacen que $\mu_n \overset{\mathscr{L}(E)}{\to} \mu$ y consideremos las medidas $\mu_n \circ h^{-1}$, $\mu \circ h^{-1}$ en E'. Sea C_h el conjunto de puntos de continuidad de h. Si $\mu(C_h) = 1$ (es decir, h es continua μ-c.s.), entonces $\mu_n \circ h^{-1} \overset{\mathscr{L}(E')}{\to} \mu \circ h^{-1}$.

Si h es continua, la prueba es inmediata pues si $f : E' \to \mathbb{R}$ es continua y acotada, $f \circ h : E \to \mathbb{R}$ también lo es y se concluye al usar el teorema del Portemanteau y el teorema de cambio de variable:

$$\int_{E'} f\, \mu_n \circ h^{-1}(dx) = \int_E (f \circ h)\mu_n(dy) \to \int_E (f \circ h)\mu_n(dy) = \int_{E'} f\, \mu \circ h^{-1}(dx).$$

El Teorema del mapeo continuo es una generalización de esto.

Demostración
Como $\mu_n \overset{\mathscr{L}(E)}{\to} \mu$ y los espacios E y E' son separables, por el teorema 8.2.7 existe un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ y variables aleatorias X_n, X con valores en E definidas en este, con distribución μ_n, μ respectivamente y tales que $X_n(\omega) \to X(\omega)$ c.s. Como $X(\omega) \in C_h$ c.s., entonces $h(X_n(\omega)) \to h(X(\omega))$ c.s. Como convergencia casi segura implica convergencia débil, $h(X_n) \overset{\mathscr{L}(E')}{\to} h(X)$. Al notar que la distribución de $h(X_n)$, $h(X)$ es $\mu_n \circ h^{-1}$, $\mu_n \circ h^{-1}$ se sigue que $\mu_n \circ h^{-1} \overset{\mathscr{L}(E')}{\to} \mu \circ h^{-1}$. ■

Detallemos un poco más esta última afirmación para que el siguiente corolario sea claro: si $B' \in E'$,

$$\begin{aligned} \mu \circ h^{-1}(B') = \mathbb{P}_X \circ h^{-1}(B') &= \mathbb{P}_X\left(x : h(x) \in B'\right) \\ &= \mathbb{P}\left(X \in \{x : h(x) \in B'\}\right) = \mathbb{P}\left(h(X) \in B'\right). \end{aligned}$$

Así, si X_n, X son elementos aleatorios con valores en E y $h : E \to E'$ boreliana, probar la convergencia $h(X_n) \overset{\mathscr{L}(E')}{\to} h(X)$ se reduce a probar que $\mathbb{P}_{X_n} \circ h^{-1} \overset{\mathscr{L}(E')}{\to} \mathbb{P}_X \circ h^{-1}$. El teorema anterior sigue siendo válido en espacios métricos arbitrarios pero solo necesitaremos esta versión. El caso general se puede consultar en Kallenberg [12] pero nótese como la prueba es más compleja al no poder usar el teorema de Skorokhod. En el siguiente corolario, $\mathbb{P}_{X_n}, \mathbb{P}_X$ juegan el papel de μ_n, μ y no es más que una reformulación del teorema del mapeo continuo adaptada al uso que se le dará:

Corolario 8.2.10 *Sean (E, d) y (E', d') espacios métricos separables y $h : E \to E'$ boreliana. Si X_n, X son elementos aleatorios con valores en E con $X_n \overset{\mathscr{L}(E)}{\to} X$ y h es continua $\mathbb{P}_X$-c.s. (o dicho de otra forma, h es continua en X c.s.), entonces $h(X_n) \overset{\mathscr{L}(E')}{\to} h(X)$.*

Observaciones: (1) Si $X \in C_h$ c.s., entonces $\mathbb{P}_X(C_h) = \mathbb{P}(X \in C_h) = 1$ y podemos aplicar el corolario anterior.
(2) Nótese como si E y E' son métricos y μ, ν medidas en E, la continuidad de una función $f : E \to E'$ en un punto depende únicamente de las métricas de E, E'. Sin embargo, que f sea continua c.s. depende de la medida considerada. Por ejemplo, $\mathbb{1}_{\mathbb{R}^+}(x) : \mathbb{R} \to \mathbb{R}$ es continua λ c.s. pero discontinua δ_0 c.s.

8.3 El espacio D y topología J1 de Skorokhod

Referencias: Jacod, Shiryaev [18] capítulo VI sección 1, Polland [24] capítulo VI, Shi [23] capítulo 3, Billingsley [8] capítulo 3 sección 12.

Sea $(S, \mathscr{B}(S), d)$ espacio métrico con $\mathscr{B}(S)$ la σ-álgebra de sus borelianos y E un espacio métrico polaco. Recordemos que el concepto de convergencia débil de variables aleatorias con valores en $(S, \mathscr{B}(S), d)$ se reduce, por definición, a la convergencia débil de medias definidas en $(S, \mathscr{B}(S), d)$. Como un proceso estocástico con trayectorias cadlag es una función medible con valores en $(D(\mathbb{R}^+, E), \mathscr{E}_D^T)$, para poder definir convergencia débil de procesos estocásticos con valores en $D(\mathbb{R}^+, E)$ necesitamos poder hablar de convergencia débil de medidas en el espacio $D(\mathbb{R}^+, E)$. Esto desde el inicio presenta algunas dificultades: para empezar necesitamos que D sea un espacio métrico, y metrizarlo como veremos a continuación no es una tarea sencilla ya que de entrada, no hay una métrica "canónica" a diferencia de, por ejemplo, el espacio $C[0,1]$. Además, necesitamos que los borelianos en D obtenidos al construir esta métrica coincidan con $\mathscr{E}_D^T$. Estos serían entonces los requisitos mínimos que la métrica deberá satisfacer para poder hablar de convergencia débil. Sin embargo, recordemos que los resultados sobre convergencia débil que probamos en la sección anterior requieren que S tenga propiedades adicionales: el teorema de representación de Skorokhod y el Teorema del mapeo continuo asumen que S es separable mientras que el teorema de Prohorov, requiere que S sea completo y separable. Como este último caracteriza compacidad relativa a través de tensión, propiedad básica al probar resultados de convergencia débil, será indispensable que D bajo la métrica que consideremos satisfaga estas características.

Para comprender mejor la importancia de estos resultados al probar convergencia de procesos con trayectorias en D, supongamos que logramos nuestro objetivo: es decir, probar la existencia de una métrica d del espacio D bajo la cual (D, d) es un espacio métrico completo y separable. Una vez logrado esto, denotaremos por $D_E[0, \infty)$ al espacio D con esta estructura de espacio métrico. Sean $\mathscr{M} = \{\mu_n\}_{n \geq 0}$, con $\mathscr{M} \subset \mathcal{P}(E)$ y $\mu \in \mathcal{P}(D)$. Podemos pensar que corresponden a las leyes $\mu_n = \mathbb{P}_{X_n}$ de una sucesión de procesos (X_n) con trayectorias en D. Buscamos un criterio que nos permita verificar la convergencia $\mu_n \overset{\mathscr{L}(D)}{\to} \mu$. Recordemos:

Lema 8.3.1 *Sea (S, ρ) un espacio métrico. Una sucesión $\{s_n\}$ converge a un elemento $s_\infty \in S$ si y solo si toda subsucesión de $\{s_n\}$ tiene una sub-subsucesión $\{s_{n_k}\}$ que converge a s_∞.*

Como D es separable, por Portemanteau 8.2.3 son equivalentes

$$\mu_n \xrightarrow{\mathscr{L}(D)} \mu \qquad \text{y} \qquad \mu_n \to \mu \ \text{ en } (\mathcal{P}(D), \rho).$$

Al aplicar el lema 8.3.1 a nuestro caso, para probar que $\mu_n \to \mu$ para alguna $\mu \in \mathcal{P}(D)$, bastaría probar que toda subsucesión de $\{\mu_n\}$ tiene una sub-subsucesión $\{\mu_{n_k}\}$ que converge a μ. Obsérvese que si logramos probar que $\mathscr{M} \subset (\mathcal{P}(D), \rho)$ es relativamente compacto, obtenemos que toda subsucesión de $\mathscr{M}$ tiene una sub-subsucesión convergente. Así, una vez más por el lema 8.3.1 solo quedaría verificar que este límite siempre debe ser μ para poder concluir que

$$\mu_n \xrightarrow{\mathscr{L}(D)} \mu.$$

Entonces, una estrategia viable para probar la convergencia débil de la sucesión medidas $\mathscr{M}$ a la medida μ consiste en probar que :

1) La colección $\mathscr{M}$ es relativamente compacta en $\mathcal{P}(D)$.
2) El único límite posible es μ.

A menudo, se prueba la compacidad relativa de $\mathscr{M}$ a través del teorema de Prohorov: al ser $D_E[0,\infty)$ completo y separable, por el teorema 8.2.4 el espacio $(\mathcal{P}(D),\rho)$ también es completo y separable y entonces, por Prohorov, $\mathscr{M}\subset\mathcal{P}(D)$ es relativamente compacta si y solo si $\mathscr{M}$ es tensa. Como esto último es más fácil de verificar, en la práctica se prueba que

1) La colección $\mathscr{M}$ es tensa.
2) El único límite posible es μ.

La caracterización del límite del segundo punto se puede lograr al probar convergencia de las distribuciones finito dimensionales. Esto se verá en el capítulo 8.4.

El contenido de esta sección es puramente de análisis y en palabras de Jacod y Shiryaev, construir la topología y probar resultados de compacidad en esta puede resultar tedioso. Así, al igual que ellos, *"we would not like to inflict the whole story on our reader"* por lo que en lugar de enfocarnos en las pruebas daremos un recorrido por las ideas principales intentando motivar lo más posible las definiciones y resultados que se van obteniendo. Daremos con detalle la bibliografía usada conforme se vaya avanzando para que se pueda consultar la fuente fácilmente de ser necesario.

8.3.1 El espacio C

Referencia: usamos la notación y seguimos Jacod, Shiryaev [18] capítulo VI así como Kallenberg [12] Apéndice A.2.

Para poner en contexto al espacio D, empezamos con un espacio conocido, el de las funciones continuas definidas en $[0,T]$ y valores en E . Cuando T es finito, $C([0,T],\mathbb{R}^+)$ con la métrica uniforme $\|x-y\|_\infty^T=\sup_{s\leq T}\|x(t)-y(t)\|$ es completo y es separable por Stone Weirestrass. Denotemos a C con su estructura de espacio métrico por $C_E[0,T]$. Cuando T no es finito, se puede equipar a C con la métrica asociada con la convergencia uniforme en compactos:

$$d(x,y)=\sum_{k\in\mathbb{N}}\frac{1}{2^n}\left(1\wedge\|x-y\|_\infty^k\right)$$

y sigue siendo completo y separable, aunque por simplicidad seguiremos considerando el caso $T<\infty$. Esta métrica induce una topología en $C_E[0,T]$ y se prueba que la sigma-álgebra de sus borelianos (la generada por sus bolas abiertas) es justamente la σ-álgebra de los cilindros finito dimensionales $\mathscr{E}_C^T$. Así, el espacio $C_E[0,T]$ ya cuenta con las características que quisiéramos dar a D y por lo tanto podemos aplicar los resultados del capítulo anterior a colecciones de medidas en $\mathcal{P}(C_E[0,\infty))$.

Ya discutimos la importancia que juega la tensión a la hora de probar la convergencia de medidas. Para verificar tensión de colecciones $\mathscr{M}\subset\mathcal{P}(C_E[0,T])$, requerimos de una buena caracterización de los conjuntos compactos (o relativamente compactos) de $C_E[0,T]$. Tenemos ya un resultado clásico de análisis que se hace cargo de esto: para cada $x \in C_E[0,T]$, sea

$$w(x,I) = \sup_{s,t\in I} |x(s) - x(t)|, \quad I \text{ intervalo de } \mathbb{R}^+.$$

A partir de este, se define el *módulo de continuidad* para funciones continuas

$$\begin{aligned} w_T(x,h) &= \sup \big\{ |x(s) - x(t)| : |s-t| \leq h,\ s,t \in [0,T] \big\} \\ &= \sup \big\{ w(x,[t,t+h]) : [t,t+h] \subset [0,T] \big\} \quad \text{para } h \geq 0. \end{aligned}$$

Entre otras cosas, satisface que si $x \in C_E[0,\infty)$, entonces $\lim_{h\downarrow 0} w_T(x,h) = 0$ por continuidad uniforme, por lo que se dice que w_T está bien adaptado al espacio $C_E[0,\infty)$. Tenemos la siguiente caracterización de los conjuntos relativamente compactos:

Teorema 8.3.2 Arzela-Ascoli
Un subconjunto $A \subset C_E[0,\infty)$ es relativamente compacto si y solo si existe un denso $S \subset [0,T]$ tal que

$$\begin{cases} (i)\ \pi_t(A) \text{ es relativamente compacto en } E \text{ para cada } t \in S \\ (ii)\ \lim_{h\downarrow 0} \sup_{x\in A} w_T(x,h) = 0 \end{cases} \tag{8.3.1}$$

donde la condición (ii) es la equicontinuidad de la familia A formulada empleando el módulo de continuidad. Así, en $C_E[0,T]$ ya tenemos todas las herramientas para empezar a probar resultados de convergencia de medidas. Sin embargo, grandes clases de procesos tienen trayectorias en el espacio $D_E[0,\infty)$ (Semimartingalas, procesos de Feller entre otros) y por lo tanto el objetivo ahora es obtener resultados similares a estos para el espacio D.

8.3.2 El espacio D

Referencia: Jacod, Shiryaev [18] capítulo VI sección 1

Empezamos definiendo un módulo de continuidad que esté adaptado a las funciones en $D = D(\mathbb{R}^+, E)$. Se define el *módulo de continuidad modificado* para cada $x \in D$ y $T \in \mathbb{R}^+$ como

$$\tilde{w}_T(x,h) = \inf \big\{ \max_{i\leq r} w(x,[t_{i-1},t_t)) : \{t_i\}_{i=1}^r \text{ partición de } [0,T] \text{ tal que } \inf_{i<r} |t_i - t_{i-1}| \geq h \big\}.$$

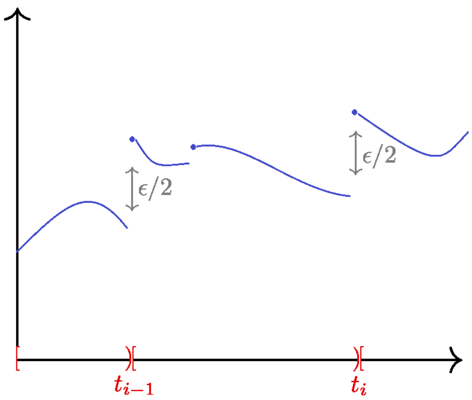
Nótese como el valor de x en el extremo T no influye en nada el valor, y que estamos tomando $i < r$ en la última restricción. Este último detalle puede diferir según el autor y según se trabaje en $[0,T]$ con $T < \infty$ o $\mathbb{R}^+$. Aquí no estamos poniendo restricción sobre la longitud del último intervalo de la partición. Nótese como si x tiene un salto en un punto t_i de la partición π de $[0,T]$, la magnitud de este $\Delta x(t_i)$ deja de ser considerada en $\{\max_{i\leq r} w(x,[t_{i-1},t_t)) : \pi = \{t_i\}_{i=1}^r\}$. Este módulo de continuidad está bien adaptado al espacio D y nos permitirá caracterizar a los conjuntos relativamente compactos en $D_E[0,\infty)$ a partir de un resultado similar a Arzela-Ascoli, como se verá más adelante.

Lema 8.3.3 *Si $x \in D_E[0,\infty)$ entonces,*
(i) Para cada T y $\epsilon > 0$, x tiene una cantidad finita de discontinuidades $\Delta x(t) \geq \epsilon$. Por lo tanto, x tiene una cantidad a lo más numerable de discontinuidades en $\mathbb{R}^+$.
(ii) Para cada T, x es acotado en $[0,T]$.
(iii) $\lim_{h\downarrow 0} w_N(x,h) = 0$ para cada $N \in \mathbb{N}$.

Demostración

(i) Si para algún $T > 0$ y $\epsilon > 0$, x tuviera una cantidad infinita de discontinuidades de tamaño $\Delta x(t) > \epsilon$ en $[0, T]$, entonces debe haber un punto de acumulación de estas lo cual nos lleva a una contradicción. Como (ii) es clara, probemos (iii).

(iii) Fijemos $N \in \mathbb{N}$. Sabemos que existe una cantidad finita de $t \in [0, T]$ tal que $\Delta x(t) > \epsilon/2$. Así, al considerar una partición de $[0, T)$ que contenga a estos puntos de salto, podemos construir una cubierta finita de intervalos disjuntos $\{[s_{i-1}, s_i)\}_{i \in J}$ de $[0, T)$ para la cual $\Delta x(t) \leq \epsilon/2$ para todo $t \in (s_{i-1}, s_i)$. A cada intervalo $[s_{i-1}, s_i)$ de esta cubierta le aplicamos el siguiente procedimiento:



Como en (s_{i-1}, s_i) se satisface $\Delta x(t) \leq \epsilon/2$, por continuidad por la derecha, para cada $t \in [s_{i-1}, s_i)$ existe una vecindad abierta $V_t \subset [s_{i-1}, s_i)$ relativa a la topología de subespacio para la cual $|x(s) - x(t)| < \epsilon$ si $s \in V_t \cap [s_{i-1}, s_i)$. Al existir el límite por la izquierda en $t = s_i$, existe una vecindad abierta $V_t \subset [s_{i-1}, s_i]$ relativa a la topología de subespacio tal que $|x(s) - x(s_i-)| \leq \epsilon$ si $s \in V_t$. La colección de vecindades $\{V_t : t \in [s_{t-1}, s_i]\}$ es una cubierta abierta del intervalo $[s_{t-1}, s_i]$ con la topología de subespacio. Así, existe una subcubierta finita $\{V_t\}_{t \in I_i}$. Al proceder similarmente para cada intervalo $[s_{i-1}, s_i)$ $i \in J$ de la cubierta de $[0, T)$, obtenemos la colección

$$\mathscr{P} = \big\{V_t : V_t \in \{V_t\}_{t \in I_i} \text{para algún } i \in J\big\}.$$

Consideremos la partición $\pi = \{t_i\}$ formada por los extremos de cada $V_t \in \mathscr{P}$, afirmamos que

$$\max_i \{w(x, [t_{i-1}, t_i))\} = \max_i \{|x(s) - x(u)| : s, u \in [t_{i-1}, t_i)\} \leq 2\epsilon.$$

En efecto, si $s, u \in [t_{i-1}, t_i)$, entonces $s, u \in V_t$ para algún $V_t \in \mathscr{P}$ y por construcción $|x(s) - x(u)| \leq |x(s) - x(t)| + |x(t) - x(u)| \leq 2\epsilon$. Si $h = \inf |t_i - t_{i-1}|$, $t_i \in \pi$, tenemos que

$$\tilde{w}_T(x, h) \leq \max_i \{w(x, [t_{i-1}, t_i))\} \leq 2\epsilon$$

al tratarse $\pi = \{t_i\}$ de una partición particular que satisface $|t_i - t_{i-1}| \geq h$. ■

Ahora, nos damos a la tarea de construir una métrica en $D_E[0, \infty)$ que satisfaga las características enunciadas al inicio de la sección.

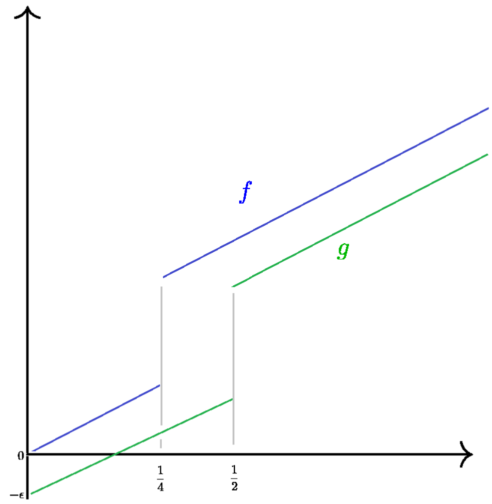
8.3.3 La distancia de Skorokhod y compacidad en $D_E[0, \infty)$

Para ganar algo de intuición antes de enunciar el teorema que caracterizará por completo la topología de Skorokhod, trabajaremos en el intervalo $[0, T]$ para $T < \infty$ y en esta sección nos referimos por D al espacio $D([0, T], E)$ con $E = \mathbb{R}^d$. Empecemos entonces hallando una métrica adecuada. Un primer intento podría ser utilizar la métrica uniforme pues de hecho, el espacio D bajo $\|\cdot\|_\infty$ es completo. Esta es en efecto una estrategia viable sobre todo cuando los procesos límite son continuos. En Pollard [24] el capítulo V está dedicado al estudio de este caso.

El problema de considerar esta métrica es que D bajo esta no es separable [1] y además no es adecuada para estudiar límites que no son continuos. Esto último se debe a que no es una forma adecuada para comparar funciones discontinuas. Para hacer esto más evidente, consideremos el siguiente ejemplo: para algún ϵ pequeño, sean $f, g : [0,T] \to \mathbb{R}^+$ las funciones en D siguientes

$$f(t) = \begin{cases} t & \text{si } t \in [0, \frac{1}{4}) \\ t+1 & \text{si } t \in [\frac{1}{4}, T] \end{cases} \qquad g(t) = \begin{cases} t-\epsilon & \text{si } t \in [0, \frac{1}{2}) \\ t+1-\epsilon & \text{si } t \in [\frac{1}{2}, T] \end{cases}$$

El salto de g ocurre un poco después al de f y esto hace que $\|f-g\|_\infty = 1$ a pesar de ser visualmente muy cercanas, como se puede apreciar en la primer figura. El problema está en que dos funciones discontinuas no pueden estar cerca bajo la norma $\|\cdot\|_\infty$ a no ser que tengan saltos en los mismos instantes. Necesitamos más flexibilidad al tratar con funciones en D. Por esto, introducimos al conjunto de reparametrizaciones o cambios de tiempo del intervalo $[0,T]$:

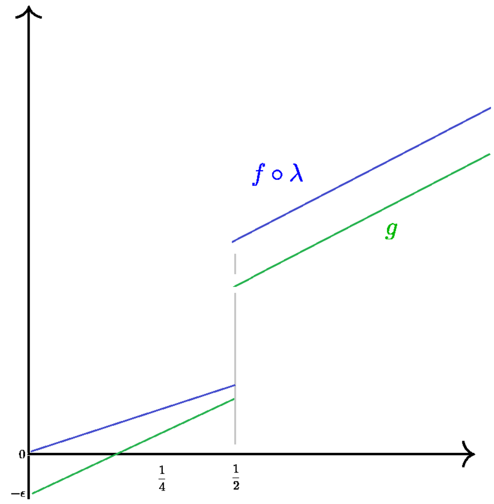


$$\Lambda = \{\lambda : [0,T] \to [0,T] : \lambda \text{ es biyección continua y monótona creciente}\}$$

Al considerar un cambio de tiempo λ adecuado, podemos lograr que los saltos de $f \circ \lambda$ y g coincidan: por ejemplo considerar la función lineal por pedazos $\lambda \in \Lambda$ tal que

$$\lambda[0,1/4) = [0,1/2) \text{ y } \lambda[1/4,T] = [1/2,T].$$

Así, obtenemos que $\|f\circ\lambda - g\|_\infty \sim \epsilon$ y entonces módulo un cambio de tiempo, f y g son en efecto cercanas bajo este nuevo enfoque. Existen muchos cambios de tiempo posibles que logran esto así que buscaremos el más "eficiente", en el sentido que definimos a continuación.



El cambio de tiempo más elemental es la identidad I (pues no estamos modificando a f), y entonces es razonable medir la eficiencia de un cambio de tiempo al considerar la magnitud $\|Id - \lambda\|_\infty$. Se puede pensar como el precio a pagar por componer con λ. Como buscamos minimizar tanto $\|f\circ\lambda\|_\infty$ como $\|\lambda - I\|_\infty$, definimos la métrica de Skorokhod en $D([0,T], E)$ como sigue:

$$d_T(f,g) = \inf_{\lambda\in\Lambda}\{\|f\circ\lambda - g\|_\infty \vee \|\lambda - I\|_\infty\} \tag{8.3.2}$$

[1] Si $x^{(u)}(t) := \mathbb{1}_{[0,u)}(t)$, $t \in [0,T]$, entonces $A = \{x^{(u)}, u \in [0,T]\}$ es no numerable en D y $\|x^{(u)} - x^{(v)}\|_\infty = 1$ si $u \neq v$.

En algunos textos se remplaza $\vee$ por $+$ pero es posible verificar que las dos métricas resultan uniformemente equivalentes y por lo tanto, son equivalentes y sus sucesiones de Cauchy coinciden. No es difícil probar que es una métrica (ver por ejemplo, Billingsley [8] capítulo 3 sección 16 o Shi [23] capítulo 3) y entonces induce una topología en D, llamada la *topología J1 de Skorokhod*. Existen otras topologías como la J2, M1 y M2 pero como solo trabajaremos con la J1, la llamaremos simplemente la topología de Skorokhod. Decimos que $x_n, x \in D$ convergen en la topología de Skorokhod si $d_T(x_n, x) \to 0$.

Lema 8.3.4 *Sean $x_n, x \in D$. Entonces, $x_n \to x$ en D si y solo si existe una sucesión $\lambda_n \in \Lambda$ tal que $\|\lambda_n - I\|_\infty \to 0$ y $\|x_n \circ \lambda_n - x\|_\infty \to 0$.*

Demostración
La ida es clara mientras que el regreso se sigue de que

$$d_T(x_n, x) = \inf_{\lambda \in \Lambda} \{\|x_n \circ \lambda - x\|_\infty \vee \|Id - \lambda\|_\infty\} \leq \|x_n \circ \lambda_n - x\|_\infty \vee \|\lambda_n - I\|_\infty \to 0.$$

■

Observación: Si $\lambda \in \Lambda$, entonces $\lambda^{-1} \in \Lambda$, y además, como se puede leer en Shi [23] Definición 1.3, no es difícil verificar que

$$\begin{aligned} \|f \circ \lambda - g\|_\infty &= \|f - g \circ \lambda^{-1}\|_\infty \\ \|\lambda - I\|_\infty &= \|\lambda^{-1} - I\|_\infty. \end{aligned}$$

Así, este lema es también equivalente a pedir que $\|\lambda_n - I\|_\infty \to 0$ y $\|x_n - x \circ \lambda_n\|_\infty \to 0$ para una sucesión $\lambda_n \in \Lambda$.

En principio, la métrica d_T en $D([0, T], E)$ se puede extender a todo $D(\mathbb{R}^+, E)$ sin demasiadas dificultades: solo se modifica ligeramente la definición para volver a x, y continuas en T al multiplicarlas por una función $g_T(t)$ lineal por pedazos adecuada que solo altera sus valores al final del intervalo $[0, T]$. Así, se redefine

$$d'_T(x, y) = \inf_{\lambda \in \Lambda} \{\|(g_T \cdot x) \circ \lambda - (g_T \cdot y)\|_\infty \vee \|\lambda - I\|_\infty\} \tag{8.3.3}$$

y a partir de esta, se considera la extensión a $D(\mathbb{R}^+)$

$$d(x, y) = \sum_{k=1}^{\infty} \frac{1}{2^k} 1 \wedge d'_k(x, y) \quad x, y \in D(\mathbb{R}^+, E). \tag{8.3.4}$$

Los detalles se pueden leer en Billinglsey [8] p168. Similarmente, se prueba que es una métrica y que bajo esta D es separable[2]. De no estar buscando completez, este camino es viable y de hecho es el seguido por Pollard [24] en el capítulo 6.

En nuestro caso, surge un problema desde la definición de $d_T(x, y)$ para $T < \infty$. La topología de Skorokhod, que es la inducida por la métrica d, es *separable* [3] pero como espacio métrico (D, d) no es *completo* [4]. Sin embargo no todo está perdido: aquí es importante notar que

[2]Pollard [24], capítulo 6 definición 1 y observación p124 así como teorema 6.
[3]Billingsley [8] capítulo 3 teorema 12.2
[4]Billingsley [8] capítulo 3 ejemplo 12.2

mientras que la separabilidad es una propiedad topológica, la completez depende de la métrica usada. Así, de lograr encontrar una métrica d^0 equivalente a d bajo la cual D si sea completo, la topología bajo esta sería la misma pero el espacio métrico D sería *completo y separable.* Afortunadamente:

Teorema 8.3.5 *Existe una métrica d_T^0 en D equivalente a d_T bajo la cual $(D([0,T]), d_T^0)$ es completo y separable. Además, está dada por:*

$$d_T^0(x,y) = \inf_{\lambda\in\Lambda}\{\|x\circ\lambda - y\|_\infty \vee \|\lambda\|^0\} \tag{8.3.5}$$

donde

$$\|\lambda\|^0 = \sup_{\substack{s<t\\ s,t\in[0,T]}} \left|\log\frac{\lambda(t)-\lambda(s)}{t-s}\right|. \tag{8.3.6}$$

La prueba se puede leer en Bilingsley [8] teoremas 12.2 y 12.1. La idea detrás de remplazar $\|\lambda - I\|_\infty$ por (8.3.6) es penalizar aún más los cambios de tiempo con pendiente diferente de 1, que es la pendiente de I. Nótese como si $\lambda = I$, $\|\lambda\|^0 = 0$. Ahora si, modificamos como en (8.3.3):

$$\tilde{d}_T^0(x,y) = \inf_{\lambda\in\Lambda}\{\|(g_T\cdot x)\circ\lambda - (g_T\cdot y)\|_\infty \vee \|\lambda\|^0\}$$

y extendemos a $D(\mathbb{R}^+)$:

$$\tilde{d}^0(x,y) = \sum_{k=1}^{\infty}\frac{1}{2^k}\,1\wedge \tilde{d}_k^0(x,y) \quad x,y\in D(\mathbb{R}^+,E). \tag{8.3.7}$$

Esta métrica es equivalente a la definida en (8.3.4) y por lo tanto generan la misma topología ; se le llama también la topología de Skorokhod en $D(\mathbb{R}^+)$. Sin embargo, bajo la métrica $\tilde{d}^0$, el espacio $D(\mathbb{R}^+)$ es completo y separable. La métrica (8.3.2) fue la que consideró originalmente Skorokhod mientras que $\tilde{d}^0$ fue introducida más adelante por Prokhorov. Esto nos lleva a enunciar el siguiente teorema, el cual caracteriza por completo las propiedades de la topología J1 de Skorokhod:

Teorema 8.3.6 Topología J1 de Skorokhod
1) Existe una topología metrizable en $D(\mathbb{R}^+,E)$ llamada la topología de Skorokhod, y una métrica bajo la cual el espacio es completo y separable. Denotaremos $D_E[0,\infty)$ a D con su estructura de espacio métrico.

2) Se satisface la siguiente caracterización: una sucesión x_n converge a x si y solo si existe una sucesión $\lambda_n\in\Lambda$ tal que

$$\begin{cases} (i)\sup\limits_s |\lambda_n(s) - s| \to 0 \\ (ii)\sup_{s\le N}|x_n\circ\lambda_n(s) - x(s)| \to 0 \qquad \textit{para todo } N\in\mathbb{N}\end{cases} \tag{8.3.8}$$

donde $|\cdot|$ se debe interpretar en (ii) como la métrica del espacio E cuando no se trate de $\mathbb{R}^d$. Esto es, que se satisfaga el lema 8.3.4 para cada N.

3) La σ-álgebra de los borelianos de $D_E[0,\infty)$ coincide con la sigma-álgebra generada por las proyecciones (π_t) para $t\in\mathbb{R}^+$. Es decir, $\mathscr{E}_D^T = \mathscr{B}(D_E[0,\infty))$.

Ahora que podemos usar el teorema de Prohorov para caracterizar tensión de sucesiones de medidas en $D_E[0,\infty)$, requerimos de una caracterización de los conjuntos relativamente compactos (la tensión se puede definir en términos de conjuntos relativamente compactos ya que siempre podemos considerar su cerradura y obtener un compacto).

Teorema 8.3.7 Compacidad en $D_E[0,\infty)$
Un subconjunto $A \subset D_E[0,\infty)$ es relativamente compacto para la topología de Skorokhod si y solo si,

$$\begin{cases} (i)\, \pi_t(A) \text{ es relativamente compacto en } E \text{ para cada } t \in \mathbb{Q}^+ \\ \\ (ii)\, \lim_{h\downarrow 0} \sup_{x\in A} \tilde{w}_N(x,h) = 0 \qquad \text{para cada } N \in \mathbb{N} \end{cases} \tag{8.3.9}$$

Otra forma de decir (i) es que para cada t, existe un compacto K_t tal que $\pi_t(A) \subset K_t$. (pues al ser relativamente compacto, su cerradura es un compacto). Cuando E es localmente compacto, podemos además garantizar que existe un compacto $K_{[0,T]} \subset E$ común: es decir, tal que $\pi_t(A) \subset K_{[0,T]}$ para cada $t \in [0,T]$. En particular en el caso $E = \mathbb{R}^d$, (i) se ve como:

$$(i') \sup_{x\in A} \sup_{s\le T} |x(s)| < \infty \qquad \text{para cada } T \ge 0 \tag{8.3.10}$$

que es como se enuncia en [18] capítulo VI, teorema 1.14. El compacto sería $K_{[0,T]} = \overline{B}(0,M)^d$ para $M \ge \sup_{x\in A} \sup_{s\le T} |x(s)|$.

En Jacod, Shiryaev[18] se recomienda fuertemente al lector no estudiar la prueba de estos resultados en vista de que esta no aporta ninguna intuición adicional a las características de la topología. Seguimos su consejo religiosamente. Los dos teoremas anteriores se pueden consultar en Jacod, Shiryaev [18] (teorema 1.14 capítulo VI) para el caso $E = \mathbb{R}^d$ y en Ethier, Kurtz [20] (capítulo III sección 5) y Kallenberg [12] (Apéndice A2) para espacios métricos. Algunas observaciones son de Kolokoltsov [19] (capítulo I, teorema 1.7.2).

8.3.4 Continuidad de funciones en $D_E[0,\infty)$

La convergencia débil de medidas $\mu_n \overset{\mathscr{L}(D)}{\to} \mu$ en $D_E[0,\infty)$ nos permite, al aplicar el teorema del Portemanteau o el teorema del mapeo continuo, calcular límites interesantes (reales y en ley) al componer con funciones continuas con respecto a la topología de Skorokhod. Aunque resulta complicado hallar funciones $f : D_E[0,\infty) \to \mathbb{R}$ continuas en todo $D_E[0,\infty)$, para aplicar el Teorema del mapeo continuo es más que suficiente con verificar que f es continua en un subconjunto $U \subset C_f \subset D_E[0,\infty)$, para el cual $\mu(U) = 1$ pues entonces, $1 = \mu(U) \le \mu(C_f)$. Finalizamos probando la continuidad de algunas funciones f que aparecen frecuentemente, en subconjuntos U de la forma

$$\{x \in D_E[0,\infty) : x(t) = x(t-)\} = \{x \in D_E[0,\infty) : \Delta x(t) = 0\}.$$

Proposición 8.3.8 Continuidad de funciones en $D_E[0,\infty)$
(1) Para cada $t \in \mathbb{R}^+$, las funciones $\pi_t : x \to x(t)$ y $\pi_{t-} : x \to x(t-)$ son continuas en $\{x \in D_E[0,\infty) : x(t) = x(t-)\} \subset D_E[0,\infty)$.

(2) Para cada $t \in \mathbb{R}^+$, las funciones $S_t : x \to \sup_{s \leq t} |x(s)|$ y $S_t^{\Delta} : x \to \sup_{s \leq t} |\Delta x(s)|$ son continuas en $\{x \in D_E[0, \infty) : x(t) = x(t-)\} \subset D_E[0, \infty)$.

Así, en particular, son continuas en $C(E) \subset D_E[0, \infty)$.

Se pueden hallar otros ejemplos de funciones continuas en Jacod, Shiryaev [18] o por ejemplo Whitt [25].

Demostración
(1) Para probar esto, bastaría ver que para toda $x \in \{x \in D_E[0, \infty) : x(t) = x(t-)\}$, si $x_n \to x$ en $D_E[0, \infty)$, entonces $\pi_t(x_n) \to \pi_t(x)$ en E. Así, supongamos que $x_n \to x$ en $D_E[0, \infty)$ y sea $T > t$. Por la caracterización (2) del teorema 8.3.6 y la observación que sigue al lema 8.3.4, existe una sucesión de cambios de tiempo $\lambda_n \in \Lambda$ tal que $\|x_n - x \circ \lambda_n\|_\infty^T \to 0$ y $\|\lambda_n - I\|_\infty \to 0$. Así,

$$\begin{aligned} |x_n(t) - x(t)| &\leq |x_n(t) - x \circ \lambda_n(t)| + |x \circ \lambda_n(t) - x(t)| \\ &\leq \|x_n - x \circ \lambda_n\|_\infty^T + |x \circ \lambda_n(t) - x(t)|. \end{aligned}$$

Por continuidad en t, el segundo término tiende a 0 mientas que el primero lo hace por elección de los cambios de tiempo λ_n. Las demás pruebas se omiten, se pueden consultar en Jacod, Shiryaev capítulo VI, sección 1 proposición 2.4. ■

Primeras aplicaciones:

Para finalizar esta sección, mostramos como la continuidad de estas funciones junto con el Teorema del mapeo continuo 8.2.10 permite obtener límites débiles que podrían ser de interés. En Whitt [25] se aborda con mucha más profundidad este tipo de métodos, conocidos como *the continuous mapping approach*, para hallar límites débiles.

Sea X un proceso de Feller y supongamos que (X_n) es una sucesión de procesos con valores en D tales que $X_n \overset{\mathscr{L}(D)}{\to} X$. Por la Casi-continuidad por la izquierda 6.3.2, sabemos que $\mathbb{P}(X_t \neq X_{t-}) = 1$ para cada t. Dicho de otro forma, esto es

$$\begin{aligned} &\mathbb{P}_X(x \in D_E[0, \infty) : x(t) = x(t-)) \\ &\qquad = \mathbb{P}(X \in \{x \in D_E[0, \infty) : x(t) = x(t-)\}) = \mathbb{P}(X_t = X_{t-}) = 1. \end{aligned}$$

Así, $X \in \{x \in D_E[0, \infty) : x(t) = x(t-)\}$ c.s. y las funciones $h : D_E[0, \infty) \to E$ definidas en la proposición 8.3.8 son continuas en $\{x \in D_E[0, \infty) : x(t) = x(t-)\}$. Por el corolario 8.2.10 tenemos la convergencia débil de $h(X_n) \overset{\mathscr{L}(E)}{\to} h(X)$. Esto es, para cada t tenemos la convergencia débil de las siguientes variables aleatorias:

$$X_t^n \xrightarrow{\mathscr{L}(E)} X_t \qquad\qquad X_{t-}^n \xrightarrow{\mathscr{L}(E)} X_{t-}$$

$$\sup_{s \leq t} X_s^n \xrightarrow{\mathscr{L}(E)} \sup_{s \leq t} X_s^n \qquad\qquad \sup_{s \leq t} |\Delta X_s^n| \xrightarrow{\mathscr{L}(E)} \sup_{s \leq t} |\Delta X_s|.$$

Por ejemplo, ya probamos la convergencia débil de la caminata interpolada $S_n^*(t)$ al movimiento browniano B por lo que estos límites son válidos al considerar $S_n^*(t) = \frac{S(nt)}{\sqrt{n}} = X^n$ y $X = B$. Así, tenemos el siguiente corolario:

Corolario 8.3.9 Teorema del Límite central
Sean (ξ_n) variables aleatorias independientes e idénticamente distribuidas con valores en $\mathbb{R}$, media $\mathbb{E}[\xi_n] = 0$ y $var(\xi_n) = 0$. Si $S_n = \sum_{k=1}^n \xi_k$, entonces $S_n/\sqrt{n} \overset{\mathscr{L}(\mathbb{R})}{\to} N(0,1)$.

Demostración
Consideremos a la la caminata aleatoria interpolada $S(t)$ como se definió en (8.1.1) y sea $S_n^*(t) = \frac{S(nt)}{\sqrt{n}}$. Por el teorema de Donsker, $S_n^* \overset{\mathscr{L}(D)}{\to} B$. Como B es continuo, para cada t tenemos que $S_n^*(t) \overset{\mathscr{L}(E)}{\longrightarrow} B_t$. En particular, si consideramos $t = 1$, obtenemos que $S_n/\sqrt{n} \overset{\mathscr{L}(E)}{\longrightarrow} B_1$ que tiene distribución $N(0,1)$. ∎

8.4 Convergencia de procesos en $D_E[0,\infty)$

Referencia: Ethier, Kurtz [20] capítulo III sección 7

Empezamos probando que la distribución de un proceso con valores en $D_E[0,\infty)$ queda determinada por sus distribuciones finito dimensionales para una colección de tiempos en un subconjunto denso de $[0,\infty)$:

Lema 8.4.1 *Sean X y Y procesos en $D_E[0,\infty)$, y supongamos que para un denso $\mathcal{D} \subset [0,\infty)$ se satisface que*

$$(X_{t_1}, \dots X_{t_k}) \overset{\mathscr{L}}{=} (Y_{t_1}, \dots, Y_{t_k})$$

para todo subconjunto finito $\{t_1, \dots t_n\} \subset \mathcal{D}$. Entonces, $Y \overset{\mathscr{L}}{=} X$.

Demostración
Para ver esto, bastaría probar que X y Y tienen las mismas distribuciones finito dimensionales. Sean entonces $t_1, \dots t_k \in [0,\infty)$ arbitrarios y consideremos sucesiones $\{t_1^n\} \subset \mathcal{D} \cap [t_1,\infty)$, ..., $\{t_k^n\} \subset \mathcal{D} \cap [t_k,\infty)$ tales que $t_1^n \downarrow t_1$, ..., $t_k^n \downarrow t_k$. Por continuidad por la derecha de X y Y:

$$\lim_{n\to\infty} \pi_{t_1^n,\dots,t_k^n}(X) = \pi_{t_1,\dots,t_k}(X) \quad \text{y} \quad \lim_{n\to\infty} \pi_{t_1^n,\dots,t_k^n}(Y) = \pi_{t_1,\dots,t_k}(Y)$$

puntualmente y por lo tanto en distribución. Pero para cada n, $\pi_{t_1^n,\dots,t_k^n}(X) \overset{\mathscr{L}}{=} \pi_{t_1^n,\dots,t_k^n}(Y)$ por lo cual, por unicidad (débil) del límite en distribución, X y Y deben tener las mismas distribuciones finito dimensionales. ∎

Lema 8.4.2 *Sea E un espacio métrico cualquiera. Si X es un proceso con trayectorias en $D_E[0,\infty)$, entonces el complemento en $[0,\infty)$ de*

$$\begin{aligned} D(X) &= \{t \geq 0 : \mathbb{P}(X_t = X_{t-}) = 1\} \\ &= \{t \geq 0 : \mathbb{P}(X_t \neq X_{t-}) = 0\} \end{aligned} \tag{8.4.1}$$

es a lo más numerable. Así, $D(X)$ es denso en $\mathbb{R}^+$.

Demostración
Supongamos que el complemento de (8.4.1) es no numerable. Como $D(X)^c = \{t \geq 0 :$

$\mathbb{P}(X_t \neq X_{t-}) > 0\}$, debe existir un $T > 0$ para el cual la restricción al intervalo $[0,T]$, $D_T(X)^c = [0,T] \cap D(X)^c$ es no numerable. Descomponemos a $D_T(X)^c$ como sigue:

$$D_T(X)^c = \bigcup_{n,m\in\mathbb{N}} \{t \in [0,T] : \mathbb{P}\big(d(X_t, X_{t-}) > 1/n\big) > 1/m\}$$

y notamos que entonces deben existir $N, M \in \mathbb{N}$ tales que la cardinalidad de

$$U_{M,N} = \{t \in [0,T] : \mathbb{P}\big(d(X_t, X_{t-}) > 1/N\big) > 1/M\}$$

es no numerable. En particular, existe una sucesión $\{t_j\} \subset U_{M,N} \subset [0,T]$ para la cual,

$$\mathbb{P}\left(d(X_{t_j}, X_{t_j-}) > 1/N\right) > 1/M \quad \text{para todo } j.$$

Pero entonces,

$$\begin{aligned}\mathbb{P}\left(d(X_{t_j}, X_{t_j-}) > 1/N \quad i.o.\right) &= \mathbb{P}\left(\cap_n \cup_{j\geq n} \{d(X_{t_j}, X_{t_j-}) > 1/N\}\right) \\ &= \lim_{n\to\infty} \mathbb{P}\left(\cup_{j\geq n}\{d(X_{t_j}, X_{t_j-}) > 1/N\}\right) \geq 1/M.\end{aligned}$$

pues cada uno de los términos $\mathbb{P}\left(d(X_{t_j}, X_{t_j-}) > 1/N\right)$ es mayor que $1/M$. Así,

$$\mathbb{P}\left(d(X_{t_j}, X_{t_j-}) > 1/N \quad i.o.\right) > 0 \tag{8.4.2}$$

lo cual nos lleva a una contradicción, ya que como X tiene trayectorias en $D_E[0,\infty)$, solo puede tener una cantidad finita de discontinuidades de magnitud superior a $1/N$ en el intervalo $[0,T]$ por el lema 8.3.3. ■

Teorema 8.4.3 Convergencia en $D_E[0,\infty)$
Sea E un espacio métrico separable y $\{X^n\}$ procesos con valores en $D_E[0,\infty)$.
(a) Si $X_n \overset{\mathscr{L}(D)}{\to} X$, entonces

$$\big(X^n(t_1), \ldots, X^n(t_n)\big) \overset{\mathscr{L}(E^n)}{\to} \big(X(t_1), \ldots, X(t_n)\big)$$

para cada subconjunto finito $\{t_1, \ldots, t_n\} \subset D(X)$. Además, para cualquier subconjunto finito $\{t_1, \ldots, t_k\} \subset [0,\infty)$, existen sucesiones $\{t_1^n\} \subset [t_1,\infty)$, $\ldots$, $\{t_k^n\} \subset [t_k,\infty)$ que convergen a $t_1, \ldots, t_k$ respectivamente y tales que $(X^n_{t_1^n}, \ldots, X^n_{t_k^n}) \overset{\mathscr{L}(E^n)}{\to} (X_{t_1}, \ldots, X_{t_k})$.

(b) Si la sucesión $\{X^n\}$ es relativamente compacta y existe un subconjunto denso $\mathcal{D} \subset [0,\infty)$ tal que

$$\big(X^n(t_1), \ldots, X^n(t_k)\big) \overset{\mathscr{L}(E^n)}{\to} \big(X(t_1), \ldots, X(t_k)\big)$$

para cada $\{t_1, \ldots, t_k\} \subset \mathcal{D}$, entonces $X_n \overset{\mathscr{L}(D)}{\to} X$.

Observación: (1) Recordemos que en el caso de los procesos de Feller, una consecuencia de la Casi-continuidad por la izquierda 6.3.2 era que $D(X) = \mathbb{R}^+$ y por lo tanto, (a) se reduce a $X_n \overset{\mathscr{L}(D)}{\to} X \Rightarrow X_n \overset{f.d.}{\to} X$.
(2) Cuando en (b) el denso es $\mathcal{D} = D(X)$, la prueba se vuelve particularmente simple y por esto la incluimos al final como observación.

Demostración

(a) Recordemos que $\pi_t : D_E[0,\infty) \to E$ es continua en $C_{\pi_t} = \{x \in D_E[0,\infty) : x(t) = x(t-)\}$. Sea $t \in D(X)$, entonces

$$\begin{aligned}&\mathbb{P}_X(\{x \in D_E[0,\infty) : x(t) = x(t-)\})\\ &\qquad = \mathbb{P}(X \in \{x \in D_E[0,\infty) : x(t) = x(t-)\}) = \mathbb{P}\left(X_t = X_{t-}\right) = 1,\end{aligned}$$

es decir, π_t es continua $\mathbb{P}_X$ c.s. Ahora, si $t_1, \ldots, t_n \in D(X)$, entonces $\pi_{t_1,\ldots,t_n} : D_E[0,\infty) \to E^n$ definida como

$$\pi_{t_1,\ldots,t_n}(x) = (\pi_{t_1}(x), \ldots, \pi_{t_n}(x)) = (x(t_1), \ldots, x(t_n))$$

es continua $\mathbb{P}_X$ c.s. ya que restringida a $C_{\pi_{t_1}} \cap \cdots \cap C_{\pi_{t_n}}$ cada entrada es continua y $\mathbb{P}_X(C_{\pi_{t_1}} \cap \cdots \cap C_{\pi_{t_n}}) = 1$. Por el Teorema del mapeo continuo 8.2.10, $\pi_{t_1,\ldots,t_n}(X^n) \overset{\mathscr{L}(E^k)}{\to} \pi_{t_1,\ldots,t_n}(X)$, es decir

$$(X^n_{t_1}, \ldots, X^n_{t_k}) \overset{\mathscr{L}(E^k)}{\to} (X_{t_1}, \ldots, X_{t_k}).$$

Probemos ahora la segunda afirmación. Nos restringimos al caso en el cual el vector consta de una sola entrada, es decir, probaremos que para cualquier $t \in [0,\infty)$ existe una sucesión $\{t^n\} \subset [t,\infty)$ tal que $t^n \downarrow t$ y

$$X^n_{t^n} \overset{\mathscr{L}(E)}{\longrightarrow} X_t. \tag{8.4.3}$$

Por el teorema 8.2.7 existe un espacio de probabilidad $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{\mathbb{P}})$ y procesos Y_n, Y definidos en este, con valores en $D_E[0,\infty)$ y con la misma distribución que X_n, X tales que $Y_n(\omega) \to Y(\omega)$ $\tilde{\mathbb{P}}$ c.s. Nótese que $D(X) = D(Y)$ ya que si $t \in D(X)$,

$$1 = \mathbb{P}_X(x \in D_E[0,\infty) : x(t) = x(t-)) = \mathbb{P}_Y(x \in D_E[0,\infty) : x(t) = x(t-))$$

y por lo tanto, $Y(\omega) \in C_{\pi_{t^i}}$ c.s. para cada i. Así, por continuidad

$$\pi_{t^i}(Y^n(\omega)) \overset{\tilde{\mathbb{P}}c.s.}{\longrightarrow} \pi_{t^i}(Y(\omega)) \qquad \text{y} \qquad \lim_{t^i \downarrow t} \pi_{t^i}(Y) = \pi_t(Y)$$

para cada i, donde la segunda igualdad se tiene simplemente por continuidad por la derecha. Se sigue entonces por desigualdad del triángulo en (E,d) que $\pi_{t^n}(Y^n) \to \pi_t(Y)$ $\tilde{\mathbb{P}}$ c.s. y por lo tanto la convergencia se tiene también en distribución. Con esto concluimos que

$$\pi_{t^n}(X^n) \overset{\mathscr{L}(E)}{\to} \pi_t(X)$$

$$\begin{tikzcd}
X^n \arrow[r, "\mathscr{L}(D)", "n\to\infty"'] \arrow[d, phantom, "\mathscr{L}\ \|"] & X \arrow[d, phantom, "\|\ \mathscr{L}"] \\
Y^n \arrow[r, "c.s.", "n\to\infty"'] \arrow[d, "\pi_{t^i}"'] & Y \arrow[d, "\pi_{t^i}"] \\
\pi_{t^i}(Y^n) \arrow[r, "c.s.", "n\to\infty"'] \arrow[dr] & \pi_{t^i}(Y) \arrow[d, "t^i \downarrow t", "c.s."'] \\
 & \pi_t(Y)
\end{tikzcd}$$

que es (8.4.3). Para probar el caso general, se razona de la misma manera considerando en lugar de π_t y π_{t^n} a $\pi_{t_1,\ldots,t_k}$ y a $\pi_{t_1^n,\ldots,t_k^n}$. Como $Y \in C_{\pi_{t_1}} \cap \cdots \cap C_{\pi_{t_k}}$ c.s., el argumento es similar.

(b) Como la sucesión (X^n) es relativamente compacta, bastaría probar que cualquier sucesión que converge débilmente, tiene por límite débil a X. Supongamos entonces (al renombrar de ser necesario) que $X_n \stackrel{\mathscr{L}(D)}{\to} Y$ donde Y es un proceso con valores en $D_E[0,\infty)$. Por el lema 8.4.1 para verificar que $X \stackrel{\mathscr{L}}{=} Y$ bastaría probar para $\{t_1,\dots,t_k\}$ cualesquiera en el denso $D(Y)$,

$$(X_{t_1},\dots X_{t_k}) \stackrel{\mathscr{L}(E^k)}{=} (Y_{t_1},\dots,Y_{t_k}).$$

Sean $f_1,\dots,f_k : E\to\mathbb{R}$ continuas y acotadas. Entonces $F(x_1,\dots,x_k)=\prod_i f_i(x_i) : E^n\to\mathbb{R}$ es continua y acotada.

Por la parte (a), sabemos que para todo $\{t_1,\dots,t_k\}\in D(Y)$ se cumple que

$$\left(X^n_{t_1},\dots,X^n_{t_k}\right) \stackrel{\mathscr{L}(E^k)}{\to} (Y_{t_1},\dots,Y_{t_k})$$

y por lo tanto, por Portemanteau

$$\lim_{n\to\infty}\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t_i})\right]=\mathbb{E}\left[\prod_{i=1}^k f_i(Y_{t_i})\right]. \tag{8.4.4}$$

Consideremos sucesiones $\{t^n_1\}\subset\mathcal{D}\cap[0,\infty)$, $\dots$, $\{t^n_k\}\subset\mathcal{D}\cap[0,\infty)$ que decrecen a $t_1,\dots,t_k$ respectivamente. Por hipótesis, para cada m fija tenemos que

$$(X^n_{t^m_1},\dots,X^n_{t^m_k}) \stackrel{\mathscr{L}(E^k)}{\to} (X_{t^m_1},\dots,X_{t^m_k})$$

y entonces una vez más por Portemanteau

$$\lim_{n\to\infty}\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t^m_i})\right]=\mathbb{E}\left[\prod_{i=1}^k f_i(X_{t^m_i})\right]. \tag{8.4.5}$$

Entonces, obtenemos que

$$\begin{aligned}
&\left|\mathbb{E}\left[\prod_{i=1}^k f_i(X_{t_i})\right]-\mathbb{E}\left[\prod_{i=1}^k f_i(Y_{t_i})\right]\right|\\
&\quad\le\left|\mathbb{E}\left[\prod_{i=1}^k f_i(X_{t_i})\right]-\mathbb{E}\left[\prod_{i=1}^k f_i(X_{t^m_i})\right]\right|\\
&\qquad+\left|\mathbb{E}\left[\prod_{i=1}^k f_i(X_{t^m_i})\right]-\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t^m_i})\right]\right|\\
&\qquad\quad+\left|\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t^m_i})\right]-\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t_i})\right]\right|\\
&\qquad\qquad+\left|\mathbb{E}\left[\prod_{i=1}^k f_i(X^n_{t_i})\right]-\mathbb{E}\left[\prod_{i=1}^k f_i(Y_{t_i})\right]\right|
\end{aligned}$$

El primer y tercer término se pueden hacer arbitrariamente pequeños por continuidad por la derecha mientras que el segundo y el cuarto por (8.4.4) y (8.4.5) respectivamente. Al aproximar indicadoras, tenemos que las distribuciones finito dimensionales de X y Y coinciden para todo $\{t_1,\dots t_k\}\in D(Y)$, que es un denso. Concluimos al aplicar el lema 8.4.1 que $X\stackrel{\mathscr{L}}{=}Y$.

Observación: Si en (b) pedimos que el denso $\mathcal{D}$ sea $D(X)$, la prueba se simplifica considerablemente: manteniendo la notación, tenemos que $D(X)\cap D(Y)$ es un denso de $[0,\infty)$ (ya que el complemento de cada uno de estos es a lo más numerable) y por un lado por (a), para todo $t_1,\ldots,t_k \in D(X)\cap D(Y)$ tenemos que $(X^n_{t_1},\ldots,X^n_{t_k}) \overset{\mathscr{L}(E^k)}{\to} (Y_{t_1},\ldots,Y_{t_k})$ mientras que por hipótesis, $(X^n_{t_1},\ldots,X^n_{t_k}) \overset{\mathscr{L}(E^k)}{\to} (X_{t_1},\ldots,X_{t_k})$ por lo que las distribuciones finito dimensionales de X y Y coinciden en el denso $D(X)\cap D(Y)$. Concluimos que tienen la misma distribución al aplicar el lema 8.4.1. ■

8.5 Criterios de tensión para medidas en $D_E[0,\infty)$

Referencias: Kallenberg [12] capítulo 14 y Kolokoltsov [19] capítulo 4 sección 4.8. Algunos elementos de Kurtz [20] capítulo 7 teorema 7.2.

Sea (E,d) un espacio métrico completo y separable. Dimos en el teorema 8.4.3 un criterio de convergencia para sucesiones de medidas en $D_E[0,\infty)$. Este requiere que la sucesión de medidas sea relativamente compacta en $\mathcal{P}(D)$ y como $D_E[0,\infty)$ es completo y separable, sabemos por Prohorov que esto es equivalente a que sea tensa. Así, para aplicar el teorema 8.4.3 necesitamos criterios que nos permitan determinar cuando una sucesión de medidas en $D_E[0,\infty)$ es tensa. Empezamos probando un primer criterio y luego a partir de este obtenemos un segundo resultado, conocido como el criterio de tensión de Aldous. Este es el que usaremos en la última sección para caracterizar la convergencia débil de procesos de Feller.

Teorema 8.5.1 *Sea $\{X^\alpha\}$ familia de procesos con trayectorias en $D_E[0,\infty)$. Entonces $\{X^\alpha\}$ es tensa (y por lo tanto relativamente compacta) si y solo si se satisfacen las dos condiciones siguientes:*

(i) Para cada $t\in\mathbb{Q}$, la familia de variables aleatorias $\{\pi_t(X^\alpha)\}$ es tensa en E, es decir, para cada $\eta>0$ y racional $t\geq 0$ existe un compacto $K_{\eta,t}\subset E$ tal que

$$\inf_\alpha \mathbb{P}\left(X^\alpha_t \in K_{\eta,t}\right) \geq 1-\eta$$

(ii) Para todo $\epsilon>0$ y $T>0$

$$\lim_{h\downarrow 0}\sup_\alpha \mathbb{P}\left(\tilde{w}_T(X^\alpha,h)>\epsilon\right)=0$$

Observaciones: 1) Si $\{X^\alpha\}$ es relativamente compacta y E es locamente compacto, se satisface una condición más fuerte que (i) llamada la *compact containment condition.* Esto es, para cada $\eta\geq 0$ y $T>0$, existe un compacto "común" $K_{\eta,[0,T]}\subset E$ tal que

$$\inf_\alpha \mathbb{P}\left(X^\alpha_t \in K_{\eta,[0,T]} \text{ para } 0\leq t\leq T\right)\geq 1-\eta$$

En el caso $E=\mathbb{R}$, la *compact containment condition* se puede escribir como

$$n\geq n_0 \Rightarrow \mathbb{P}\left(\sup_{t\leq T}|X^n_t|>K\right)\leq\epsilon$$

para algún $K\in\mathbb{R}$ y n_0 suficientemente grande, que es como viene enunciado este resultado en Jacod, Shiryaev [18] capítulo VI teorema 3.21.

2) Nótese que si E es compacto, la condición (i) se tiene trivialmente puesto que siempre podemos considerar el compacto $K=E$.

Como las condiciones de este teorema son difíciles de verificar, probaremos más adelante un criterio más manejable que permite garantizar (ii). Se le conoce como el criterio de Aldous.

Demostración
Para probar que la sucesión de medidas asociadas a $\{X^\alpha\}$ es relativamente compacta en $D_E[0,\infty)$, por el teorema de Prohorov debemos exhibir un conjunto $B \subset D_E[0,\infty)$ relativamente compacto que satisfaga:

$$\sup_\alpha \mathbb{P}_{X^\alpha}(B^c) = \sup_\alpha \mathbb{P}(X^\alpha \in B^c) < \epsilon. \tag{8.5.1}$$

Aquí es donde se usará la caracterización de compacidad probada en el teorema 8.3.7. Sea $T > 0$ cualquiera.

Por (ii), existe una sucesión $h_n \downarrow 0$ tal que

$$\sup_\alpha \mathbb{P}\left(\tilde{w}_T(X^\alpha, h_k) > 2^{-k}\right) \leq 2^{-(k+1)}\epsilon.$$

Por (i) si $\mathbb{Q} = \{t_1, t_2, \ldots\}$, para cada k la familia $\{\pi_{t_k}(X^\alpha)\}_\alpha$ es relativamente compacta. Así, existen compactos C_1, C_2, ... en E tales que

$$\sup_\alpha \mathbb{P}\left(X^\alpha_{t_k} \in E \setminus C_k\right) \leq 2^{-(k+1)}\epsilon.$$

Definimos el subconjunto de $D_E[0,\infty)$ siguiente:

$$B = \bigcap_k \{x : x_{t_k} \in C_k\} \cap \{x : \tilde{w}_T(x, h_k) \leq 2^{-k}\}.$$

B satisface que

$$\lim_{h\downarrow 0} \sup_{x\in B} \tilde{w}_T(x, h) = 0$$

y ademas como $\pi_{t_k}(B) \subset C_k$ (que es un compacto) se sigue que $\pi_{t_k}(B)$ es relativamente compacto en E. Así, como esto se tiene para todo $T > 0$, por el teorema 8.3.7 B es un subconjunto relativamente compacto de $D_E[0,\infty)$. Además, para cada α tenemos que

$$\begin{aligned}
\mathbb{P}_{X^\alpha}(B^c) &\leq \sum_{k=1}^\infty \mathbb{P}_{X^\alpha}(x : x_{t_k} \in E \setminus C_k) + \mathbb{P}_{X^\alpha}\left(x : \tilde{w}_T(x, h_k) > 2^{-k}\right) \\
&= \sum_{k=1}^\infty \mathbb{P}\left(X^\alpha_{t_k} \in E \setminus C_k\right) + \mathbb{P}\left(\tilde{w}_T(X^\alpha, h_k) > 2^{-k}\right) \\
&\leq \sum_{k=1}^\infty \sup_\alpha \mathbb{P}\left(X^\alpha_{t_k} \in E \setminus C_k\right) + \sup_\alpha \mathbb{P}\left(\tilde{w}_T(X^\alpha, h_k) > 2^{-k}\right) \\
&\leq \epsilon
\end{aligned}$$

por lo que se satisface (8.5.1). Con esto concluimos que $\{X^\alpha\}$ es relativamente compacta en $D_E[0,\infty)$. ■

Consideremos procesos X^n con valores en $D_E[0,\infty)$ cada uno definido en su propio espacio de probabilidad filtrado $(\Omega^n, \mathscr{F}^n, (\mathscr{F}^n_t), \mathbb{P}_n)$; las filtraciones se asumen càd.

Definición 8.5.2 *Diremos que la sucesión $\{X^n\}$ satisface la condición de Aldous si para cualquier sucesión de reales positivos $h_n \downarrow 0$ y cualquier colección acotada $\{\tau_n\}$ tal que τ_n es un $\mathscr{F}^n_t$-tiempo de paro se satisface que*

$[A]$
$$\mathbb{P}\text{-}\lim_{n\to\infty} d(X^n_{\tau_n}, X^n_{\tau_n + h_n}) = 0.$$

La condición [A] se puede escribir de varias formas equivalentes, entre ellas que:

$$[A1] \qquad \lim_{h\downarrow 0}\limsup_{n}\sup_{\substack{\tau,\sigma\in\mathscr{F}^n tdp\\ \tau\le\sigma\le\tau+h\le T}} \mathbb{E}\left[d(X^n_\tau, X^n_\sigma)\wedge 1\right] = 0.$$

para cada $T\in\mathbb{R}^+$.

Distintas formulaciones de la condición de Aldous así como las pruebas de sus equivalencias se pueden leer en Kolokolstov [19] capítulo IV proposición 4.8.1. La prueba del siguiente teorema usará $[A1]$ mientras que trabajaremos con $[A]$ en la siguiente sección. Antes de continuar, nos detenemos brevemente a aclarar el significado de la condición $[A]$, pues estamos haciendo un abuso de notación. Como los procesos no viven en el mismo espacio, en principio hablar de límite en probabilidad

$$\lim_{n\to 0}\mathbb{P}\left(d(X^n_{\tau_n}, X^n_{\tau_n+h_n}) > \epsilon\right) = 0$$

no tiene sentido. Sin embargo, nótese que como cada X^n está definido en su espacio de probabilidad $(\Omega^n, \mathscr{F}^n, (\mathscr{F}^n_t), \mathbb{P}_n)$, podemos hablar de

$$\lim_{n\to 0}\mathbb{P}_n\left(d(X^n_{\tau_n}, X^n_{\tau_n+h_n}) > \epsilon\right) = 0. \tag{8.5.2}$$

Una vez más, no se trata de un límite en probabilidad ya que para cada n, estamos considerando una medida $\mathbb{P}_n$ distinta en un espacio Ω^n distinto, pero nótese que (8.5.2) si es equivalente a que

$$d(X^n_{\tau_n}, X^n_{\tau_n+h_n}) \xrightarrow{\mathscr{L}(\mathbb{R})} 0.$$

Así, la condición $[A]$ se refiere a que $d(X^n_{\tau_n}, X^n_{\tau_n+h_n})$ converge débilmente a 0, aunque se suela denotar como un límite en probabilidad a 0.

Teorema 8.5.3 Criterio de tensión de Aldous
La condición de Aldous [A] implica la condición (ii) del teorema 8.5.1. Por lo tanto, la sucesión de procesos $\{X^n\}$ es tensa si para cada $t\in\mathbb{Q}$, la familia $\{\pi_t(X^\alpha)\}$ es tensa en E y se satisface [A].

Demostración
Trabajaremos en E con la métrica $d\wedge 1$, esta es equivalente a d y acotada por 1; la renombramos d. Sea $T>0$, para cada n, sea $\sigma^n_0 = 0$ y definimos recursivamente la siguiente sucesión creciente de tiempos de paro:

$$\sigma^n_{k+1} = \inf\{s\ge\sigma^n_k : d(X_{\sigma^n_k}, X_s) > \epsilon\}$$

Empecemos notando lo siguiente: por desigualdad del triángulo

$$\sup_{s,t\in[\sigma^n_k,\sigma^n_{k+1}]} d(X^n_s, X^n_t) \le 2\epsilon \tag{8.5.3}$$

y por lo tanto, si definimos para cualquier m al conjunto:

$$A^n_m = \bigcap_{k=1}^{m-1}\left(\{\sigma^n_{k+1}-\sigma^n_k > h\}\cup\{\sigma^n_k > T\}\right)\cap\{\sigma^n_m > T\}$$

entonces

$$\tilde{w}_T(X^n,h)\mathbb{1}_{A^n_m} \le 2\epsilon.$$

Esto ocurre pues al restringirnos a A^n_m, la colección $\{\tau^n_1, \tau^n_2, \ldots, \tau^n_m\} \cap [0,T] \cup \{T\}$ es una partición particular de $[0,T]$ para la cual se cumple (8.5.3) para toda $k \leq m$ siempre y cuando $[\sigma^n_k, \sigma^n_{k+1}] \subset [0,T]$. También observemos que como $d \leq 1$, siempre se satisface la desigualdad $\tilde{w}_T(X^n, h) \leq 1$. Así, para todo $m \in \mathbb{N}$, la siguiente cota es trivial si para alguna k ocurre alguno de los eventos $\{\sigma^n_{k+1} - \sigma^n_k \leq h,\, \sigma^n_k \leq T\}$, $\{\sigma^n_m \leq T\}$

$$\begin{aligned}\tilde{w}_T(X^n, h) &\leq 2\epsilon \mathbb{1}_{A^n_m} + \tilde{w}_T(X^n, h)\mathbb{1}_{(A^n_m)^c}\\ &\leq 2\epsilon + \sum_{k=1}^{m-1} \mathbb{1}_{\{\sigma^n_{k+1}-\sigma^n_k \leq h,\, \sigma^n_k \leq T\}} + \mathbb{1}_{\{\sigma^n_m \leq T\}}\end{aligned}$$

por lo que al tomar esperanza de ambos lados obtenemos

$$\mathbb{E}\left[\tilde{w}_T(X^n, h)\right] \leq 2\epsilon + \sum_{k=1}^{m-1} \mathbb{P}\left(\sigma^n_{k+1} - \sigma^n_k \leq h,\, \sigma^n_k \leq T\right) + \mathbb{P}\left(\sigma^n_m \leq T\right) \tag{8.5.4}$$

Recordemos que buscamos probar que se satisface la condición (ii) del teorema 8.5.1, es decir que

$$\lim_{h\downarrow 0} \limsup_n \mathbb{E}\left[\tilde{w}_T(X^n, h)\right] = 0.$$

Probaremos entonces que los términos de la derecha en (8.5.4) tienden a zero. Empecemos con los términos de la serie: por desigualdad de Markov,

$$\begin{aligned}\mathbb{P}\left(\sigma^n_{k+1} - \sigma^n_k \leq h,\, \sigma^n_k \leq T\right) &= \mathbb{P}\left(d(X_{\sigma^n_{k+1}}, X_{\sigma^n_k})\mathbb{1}_{\{\sigma^n_{k+1}-\sigma^n_k \leq h,\, \sigma^n_k \leq T\}} \geq \epsilon\right)\\ &\leq \mathbb{E}\left[d(X_{\sigma^n_{k+1}}, X_{\sigma^n_k})\mathbb{1}_{\{\sigma^n_{k+1}\leq \sigma^n_k + h,\, \sigma^n_k \leq T\}}\right]\epsilon^{-1}\\ &\leq \mathbb{E}\left[d(X_{\sigma^n_{k+1}}, X_{\sigma^n_k})\mathbb{1}_{\{\sigma^n_{k+1}\leq \sigma^n_k + h \leq T+h\}}\right]\epsilon^{-1}.\end{aligned} \tag{8.5.5}$$

Así, al restringirnos a $h < T$, tenemos que $T + h \leq 2T$ y por lo tanto (8.5.5) está acotado por

$$\nu_n(2T, h)\epsilon^{-1} = \sup_{\substack{\tau,\sigma \in \mathscr{F}^n \text{tdp}\\ \tau \leq \sigma \leq \tau + h \leq 2T}} \mathbb{E}\left[d(X^n_\tau, X^n_\sigma)\right]\epsilon^{-1}$$

que por hipótesis satisface $\lim_{h\downarrow 0}\limsup_n \nu_n(2T,h) \to 0$ por $[A1]$. Con esto obtenemos

$$\lim_{h\downarrow 0} \limsup_n \mathbb{E}\left[\tilde{w}_T(X^n, h)\right] \leq 2\epsilon + \limsup_n \mathbb{P}\left(\sigma^n_m \leq T\right).$$

Solamente falta verificar que $\lim_{n\to\infty} \mathbb{P}\left(\sigma^n_m \leq T\right) \to 0$ pues este término no depende de h. Como

$$\mathbb{P}\left(\sigma^n_m \leq T\right) = \mathbb{E}\left[1 \cdot \mathbb{1}_{\{\sigma^n_m \leq T\}}\right] \leq \mathbb{E}\left[e^{T-\sigma^n_m}\mathbb{1}_{\{\sigma^n_m \leq T\}}\right] \leq e^T \mathbb{E}\left[e^{-\sum_{k=0}^{m-1}\sigma^n_{k+1}-\sigma^n_k}\mathbb{1}_{\{\sigma^n_k \leq T\}}\right]$$

al aplicar el lema 9.1.4 y (8.5.5) obtenemos que para todo $c > 0$

$$\begin{aligned}\mathbb{P}\left(\sigma^n_m \leq T\right) &\leq e^T\left(e^{-mc} + \sup_{k<m}\mathbb{P}\left(\sigma^n_{k+1} - \sigma^n_k \leq c, \sigma^n_k \leq T\right)\right)\\ &\leq e^T\left(e^{-mc} + \nu_n(2T,h)\right)\epsilon^{-1}.\end{aligned}$$

Podemos hacer al primer término arbitrariamente pequeño al considerar c suficientemente grande y una vez más como $\lim_{h\downarrow 0}\limsup_n \nu(2T,h) = 0$ por $[A1]$ concluimos que

$$\lim_{n\to\infty} \mathbb{P}\left(\sigma^n_m \leq T\right).$$

■

8.6 Convergencia débil de procesos de Feller

Referencia: Kallenberg [12] capítulo 17.

A continuación, se asume que los espacios de probabilidad son canónicos y satisfacen las hipótesis habituales.

Teorema 8.6.1 Convergencia de Procesos de Feller (Kurtz, Mackevicius)
Sea X proceso de Feller en E definido en $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ con semigrupo (P_t) y X^1, X^2,... procesos de Feller en E con semigrupos (P_t^1), (P_t^2),... cada uno definido en su espacio de probabilidad filtrado $(\Omega^n, \mathscr{F}^n, (\mathscr{F}_t^n), \mathbb{P}_n)$. Entonces, son equivalentes:

(i) $\lim\limits_{n\to\infty} P_t^n f = P_t f$ uniformemente para cada $t \geq 0$ y $f \in C_0$.

(ii) Si $X_0^n \overset{\mathscr{L}}{\to} X_0$ en E, entonces $X^n \overset{\mathscr{L}(D)}{\to} X$ en $D_E[0,\infty)$.

Demostración
Supongamos (i) y que $X_0^n \overset{\mathscr{L}}{\to} X_0$. Por el teorema 8.4.3, esto se tendría de lograr demostrar la convergencia en distribuciones f.d. y tensión de la sucesión (X^n).

Empezamos probando la primer condición: para ver que $X^n \overset{f.d.}{\to} X$ basta con verificar que si $f_0, \ldots, f_m \in C_0(E)$ y $0 = t_0 < t_1 < \cdots < t_m$ entonces:

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{k\leq m} f_k(X_{t_k}^n)\right] = \mathbb{E}\left[\prod_{k\leq m} f_k(X_{t_k})\right]. \tag{8.6.1}$$

Procedemos por inducción: si $m = 0$, $\mathbb{E}[f(X_0^n)] \to \mathbb{E}[f(X_0)]$ puesto que por hipótesis, $X_0^n \overset{\mathscr{L}}{\to} X_0$. Ahora, supongamos válido (8.6.1) para $k \leq m-1$ y $m-1$ funciones en C_0 arbitrarias. Si $t_0 < t_1 < \cdots < t_m$ y $h_i = t_i - t_{i-1}$, al aplicar la propiedad de Markov reescribimos ambos lados de (8.6.1):

$$\mathbb{E}\left[\prod_{k\leq m} f_k(X_{t_k}^n)\right] = \mathbb{E}\left[\prod_{k\leq m-1} f_k(X_{t_k}^n)\mathbb{E}\left[f_m(X_{t_m})|\mathscr{F}_{t_{m-1}}\right]\right] = \mathbb{E}\left[\prod_{k<m} f_k(X_{t_k}^n)P_{h_m}^n f_m(X_{t_{m-1}}^n)\right]$$

y al hacer lo análogo para el término de la derecha en (8.6.1), se obtiene

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{k<m} f_k(X_{t_k}^n)P_{h_m}^n f_m(X_{t_{m-1}}^n)\right] = \mathbb{E}\left[\prod_{k<m} f_k(X_{t_k})P_{h_m} f_m(X_{t_{m-1}})\right]. \tag{8.6.2}$$

Por (i), $P_{h_m}^n f_m \to P_{h_m} f_m$ en C_0. Entonces, verificar (8.6.2) se reduce a ver que

$$\lim_{n\to\infty} \mathbb{E}\left[\prod_{k<m} f_k(X_{t_k}^n)P_{h_m} f_m(X_{t_{m-1}}^n)\right] = \mathbb{E}\left[\prod_{k<m} f_k(X_{t_k})P_{h_m} f_m(X_{t_{m-1}})\right]. \tag{8.6.3}$$

En efecto ya que, por desigualdad del triángulo

$$\begin{aligned}
&\mathbb{E}\left[\left|\prod_{k<m} f_k(X_{t_k}^n)P_{h_m}^n f_m(X_{t_{m-1}}^n) - \prod_{k<m} f_k(X_{t_k})P_{h_m} f_m(X_{t_{m-1}})\right|\right] \\
&\qquad \leq \mathbb{E}\left[\left|\prod_{k<m} f_k(X_{t_k}^n)P_{h_m}^n f_m(X_{t_{m-1}}^n) - \prod_{k<m} f_k(X_{t_k}^n)P_{h_m} f_m(X_{t_{m-1}}^n)\right|\right] \\
&\qquad\quad + \mathbb{E}\left[\left|\prod_{k<m} f_k(X_{t_k}^n)P_{h_m} f_m(X_{t_{m-1}}^n) - \prod_{k<m} f_k(X_{t_k})P_{h_m} f_m(X_{t_{m-1}})\right|\right]
\end{aligned}$$

donde el primer término tiende a 0 al estar mayorado por $\prod_{k<m} \|f_k\|_\infty \|P^n_{h_m} f_m - P_{h_m} f_m\|_\infty$ el cual tiende a 0 al tomar $n \to \infty$ mientras que probar que el segundo término converge a 0 es justamente (8.6.3). Pero (8.6.3) se tienen por hipótesis de inducción, ya que observemos que $f_{m-1} \cdot P_{h_m} f_m \in C_0$, con lo que concluimos.

Con esto, ya probamos la convergencia en distribuciones finito-dimensionales

$$X_n \overset{f.d.}{\to} X.$$

Por el teorema 8.4.3, concluiremos al probar que la sucesión (X^n) es tensa, y lo haremos por medio del criterio de tensión de Aldous 8.5.3. Obsérvese que por compacidad de E_Δ, la primer condición es inmediata (siempre podemos considerar el compacto $K = E_\Delta$) y solo nos hace falta probar que se satisface la condición de Aldous. Para cada n, denotemos por $(\mathscr{F}^n_t)$ y $(\mathscr{F}_t)$ a la filtración càd de X^n y X respectivamente. Consideremos una colección acotada (τ_n) tal que para cada n, τ_n es un $\mathscr{F}^n_t$-tiempo de paro así como una sucesión de reales positivos $h_n \downarrow 0$. La condición de Aldous consiste en probar que

$$\mathbb{P}\text{-}\lim_{n\to\infty} d(X^n_{\tau_n}, X^n_{\tau_n + h_n}) = 0. \tag{8.6.4}$$

Obsérvese que por la Propiedad fuerte de Markov 6.3.1 y la observación subsecuente,

$$\begin{aligned}
\mathbb{P}\left(d(X^n_{\tau_n}, X^n_{\tau_n+h_n}) > \epsilon\right) &= \mathbb{E}\left[\mathbb{1}_{(\epsilon,\infty)}(d(X^n_0, X^n_{h_n})) \circ \theta_{\tau_n}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{(\epsilon,\infty)}(d(X^n_0, X^n_{h_n})) \circ \theta_{\tau_n} | \mathscr{F}^n_{\tau_n}\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}^{X_{\tau_n}}\left[\mathbb{1}_{(\epsilon,\infty)}(d(X^n_0, X^n_{h_n}))\right]\right] \\
&= \mathbb{P}^{\nu_n}\left(d(X^n_0, X^n_{h_n}) > \epsilon\right)
\end{aligned}$$

para $\nu_n = \mathbb{P}_{X^n_{\tau_n}}$, medida con soporte en E. Entonces probaríamos (8.6.4) si para cualquier sucesión de distribuciones iniciales $\nu_n \in \mathcal{P}(E_\Delta)$ con $\nu_n(E) = 1$,

$$\lim_{n\to\infty} \mathbb{P}^{\nu_n}\left(d(X^n_0, X^n_{h_n}) > \epsilon\right) = 0. \tag{8.6.5}$$

Recordemos la observación que se hizo en (8.5.2): esto es probar que la sucesión $\{d(X^n_0, X^n_{h_n})\}_n$, cada una definida en el espacio de probabilidad $(\Omega^n, \mathscr{F}^n, (\mathscr{F}^n_t), \mathbb{P}^{\nu_n}_n)$, converge débilmente a 0. Bastaría entonces probar que para cualquier subsucesión, existe una subsucesión de esta que converge débilmente a 0. Como E_Δ es compacto, la sucesión de medidas ν_n es tensa y por lo tanto podemos suponer, al pasar a una subsucesión de ser necesario, que converge $\nu_n \to \nu$ débilmente a una medida $\nu \in \mathcal{P}(E_\Delta)$. Fijemos $f, g \in C_0$. Como el semigrupo es de Feller, $\lim_{t\downarrow 0} P^n_t g = g$ y como $P^n_t g \to P_t g$ por hipótesis, obtenemos que $P^n_{h_n} g \to g$ por la desigualdad del triángulo. Así, una vez más por propiedad de Markov,

$$\begin{aligned}
\mathbb{E}^{\nu_n}\left[f(X^n_0) g(X^n_{h_n})\right] &= \mathbb{E}^{\nu_n}\left[f(X^n_0)\mathbb{E}^{\nu_n}\left[g(X^n_{h_n}) | \mathscr{F}_0\right]\right] \\
&= \mathbb{E}^{\nu_n}\left[f(X^n_0) P^n_{h_n} g(X^n_0)\right] \\
&= \int_E f(x) P^n_{h_n} g(x) \nu_n(dx)
\end{aligned}$$

pues recordemos que bajo $\mathbb{P}^{\nu_n}$, $X_0^n \sim \nu_n$. Como $\nu_n \to \nu$ débilmente y $P_{h_n}^n g \to g$,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}^{\nu_n}\left[f(X_0^n)g(X_{h_n}^n)\right] &= \lim_{n\to\infty}\int_E f(x)P_{h_n}^n g(x)\nu_n(dx)\\ &= \int_E f(x)g(x)\nu(dx)\\ &= \mathbb{E}^{\nu}\left[f(X_0)g(X_0)\right].\end{aligned}$$

Entonces, como ya se probó el resultado para productos de funciones $f, g \in C_0$, por el lema de clases monótonas funcional obtenemos que $F(X_0^n, X_{h_n}^n) \to F(X_0, X_0)$ para toda $F : E\times E \to \mathbb{R}$ continua y acotada. Así, $(X_0^n, X_{h_n}^n) \xrightarrow{\mathscr{L}} (X_0, X_0)$ y en particular,

$$d(X_0^n, X_{h_n}^n) \xrightarrow{\mathscr{L}} d(X_0^n, X_0^n) = 0.$$

con lo que concluimos la prueba. ■

Teorema 8.6.2 Convergencia de Procesos de Lévy
Sea X un proceso de Lévy en $E = \mathbb{R}$ definido en $(\Omega, \mathscr{F}, \mathbb{P})$ con semigrupo (P_t) y exponente característico Ψ. Sean $X^1, X^2, \ldots$ procesos de Lévy en E con semigrupos (P_t^1), $(P_t^2), \ldots$ y exponentes característicos Ψ_1, Ψ_2, ... cada uno definido en su espacio de probabilidad $(\Omega^n, \mathscr{F}^n, \mathbb{P}_n)$. Entonces, son equivalentes:
(i) $\lim_{n\to\infty} \Psi_n(\lambda) = \Psi(\lambda)$ para todo $\lambda \in \mathbb{R}$.
(ii) $X^n \xrightarrow{\mathscr{L}(D)} X$ en $D_E[0,\infty)$.

Demostración
Empezamos probando la primer implicación. Sea $t > 0$ fija y denotemos por μ_t^n, μ_t a la distribución de X_t^n, X_t respectivamente. Por el Teorema 8.6.1, bastaría probar que para toda $f \in C_0(E)$, los semigrupos convergen $P_t^n f \to P_t f$ uniformemente. Esto es, que

$$\lim_{n\to\infty}\int_E f(y+x)\mu_t^n(dy) = \int_E f(y+x)\mu_t(dy) \quad \text{uniformemente.}$$

Por la proposición 4.1.3, $\varphi_{X_t^n}(\lambda) = e^{-t\Psi_n(\lambda)}$, $\varphi_{X_t}(\lambda) = e^{-t\Psi(\lambda)}$ y como $\Psi_n(\lambda) \to \Psi(\lambda)$, entonces por le Teorema de continuidad de Lévy $X_t^n \xrightarrow{\mathscr{L}(\mathbb{R})} X_t$. Por el teorema de representación de Skorokhod, existe un espacio de probabilidad y variables aleatorias que denotaremos Y_t^n, Y_t con la misma distribución que X_t^n, X_t respectivamente para las cuales la convergencia es puntual: $Y_t^n(\omega) \to Y_t(\omega)$. Así,

$$\begin{aligned}\left|\int_E f(y+x)\mu_t^n(dy) - \int_E f(y+x)\mu_t(dy)\right| &\le \mathbb{E}\left[\left|f(Y_t^n(\omega)+x) - f(Y_t(\omega)+x)\right|\right]\\ &\le \mathbb{E}\left[\sup_{x\in E}\left|f(Y_t^n(\omega)+x) - f(Y_t(\omega)+x)\right|\right].\end{aligned}$$

Como $f \in C_0(E)$, entonces es uniformemente continua. Así, para cada $\epsilon > 0$ existe $\delta > 0$ tal que para todo $u, v \in E$, si $|u-v| < \delta$ entonces $|f(u) - f(v)| < \epsilon$. Al considerar n (que depende de ω) suficientemente grande tal que $|Y_t^n(\omega) - Y_t(\omega)| < \delta$, obtenemos al aplicar el teorema de convergencia dominada que

$$\lim_{n\to\infty}\mathbb{E}\left[\sup_{x\in E}\left|f(Y_t^n(\omega)+x) - f(Y_t(\omega)+x)\right|\right] \le \epsilon$$

pues $|Y_t^n(\omega)+x-(Y_t^n(\omega)+x)|<\delta$. Así, $P_t^n f\to P_t f$ uniformemente para toda $f\in C_0$ y para todo $t>0$, y como $X_0^n \stackrel{\mathscr{L}(E)}{\to} X_0$ concluimos al aplicar el teorema 8.6.1.

Ahora, supongamos que $X^n \stackrel{\mathscr{L}(D)}{\to} X$. Para ver que $\Psi_n(\lambda)\to\Psi(\lambda)$, bastaría probar que $X_1^n \stackrel{\mathscr{L}(E)}{\to} X_1$ pues concluiríamos al aplicar el teorema de continuidad de Lévy. Pero esto es consecuencia del teorema 8.4.3 y en particular de la observación que le sigue, pues todo proceso de Lévy es Feller y por lo tanto es casi-continuo por la izquierda. [5] Al tener que $X^n \stackrel{f.d.}{\to} X$, en particular $X_1^n \stackrel{\mathscr{L}(E)}{\to} X_1$. ■

[5]Notemos que la casi continuidad por la izquierda de los procesos de Lévy se probó por primera vez en el lema 4.3.3 de una forma muy sencilla.

Capítulo 9

Apéndice

9.1 Resultados de teoría de la medida

Lema 9.1.1 *Sea $f = (f_1, \cdots, f_n) : \mathbb{R}^n \to \mathbb{R}^n$ medible e invertible y $X = (X_1, \cdots, X_n)$ vector aleatorio. Entonces $\sigma(X) = \sigma(f(X))$.*

Por un lado, $f(X)$ es $\sigma(X)$-medible y por lo tanto $\sigma(f(X)) \subset \sigma(X)$. Similarmente, $f^{-1} \circ f(X)$ es $\sigma(f(X))$-medible, y entonces X es $\sigma(f(X))$-medible con lo que $\sigma(X) \subset \sigma(f(X))$. En particular, si consideramos la función invertible $f(x_1, \ldots, x_n) = (x_1, x_2 - x_1, \ldots, x_n - x_{n-1})$ y $X = (X_{t_1}, \cdots, X_{t_n})$ se sigue que

$$\sigma\left(X_{t_1}, X_{t_{i+1}} - X_{t_i}; i = 1, \cdots, n-1\right) = \sigma\left(X_{t_i}; i = 1, \cdots, n\right)$$

Esto respeta la heurística de considerar las σ-álgebras como cuerpos de información: al ser f invertible, no se debería perder información al considerar la composición.

Lema 9.1.2 *Si X, Y son dos variables aleatorias independientes en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathbb{P})$ con valores en $E = \mathbb{R}^d$ y $f : E \times E \to \mathbb{R}$ es acotada o positiva, entonces*

$$\mathbb{E}\left[f(X,Y)|Y\right] = \int_E f(x, Y)\mathbb{P}_X(dx).$$

La prueba se puede consultar en cualquier libro de teoría de probabilidad.

Lema 9.1.3 *Sea $f : \mathbb{R}^d \to \mathbb{R}$ una función medible y supongamos que para todo compacto C se satisface que*

$$\int_C f dx = \infty.$$

Entonces, $f(x) = \infty$ casi seguramente.

Demostración
Supongamos que existe un subconjunto A de medida positiva tal que $f(x) < \infty$ para todo $x \in A$. Entonces,

$$0 < \lambda(A) = \lambda\left(\bigcup_{n \in \mathbb{N}} \{x \in A : |f(x)| < n\}\right)$$

por lo que para algún n, $\lambda(x \in A : |f(x)| < n) > 0$. Por regularidad de la medida de Lebesgue, para todo B medible

$$\lambda(B) = \sup\{\lambda(C) : C \subset B,\, C \text{ compacto }\}$$

así que existe un compacto C de medida positiva tal que $C \subset \{x \in A : |f(x)| < n\}$. Por hipótesis,

$$\infty = \int_C f \leq \int_{\{x \in A : |f(x)|<n\}} |f| dx < \infty$$

lo cual es una contradicción. ■

Lema 9.1.4 *Sean $X_1,\ X_2,\ \ldots, X_n \geq 0$ variables aleatorias en $(\Omega, \mathscr{F}, \mathbb{P})$ y $S_n = \sum_{k=1}^n X_k$. Entonces, para toda $c > 0$*

$$\mathbb{E}\left[e^{-S_n}\right] \leq e^{-nc} + \max_{k\leq n} \mathbb{P}\left(X_n \leq c\right).$$

Además, si $A_1, \ldots, A_k \in \mathscr{F}$, se vale la siguiente variante:

$$\mathbb{E}\left[e^{-S_n}\mathbb{1}_{\cap_k A_k}\right] \leq e^{-nc} + \max_{k\leq n} \mathbb{P}\left(X_k \leq c, A_k\right)$$

Referencia: Kallenberg [12] capítulo 14, Proposición 14.13.

Demostración

Por la desigualdad de Hölder:

$$\begin{aligned}
\mathbb{E}\left[e^{-S_n}\right] = \mathbb{E}\left[\prod_{k=1}^n e^{-X_k}\right] &\leq \prod_{k=1}^n \mathbb{E}\left[e^{-nX_k}\right]^{\frac{1}{n}} \\
&\leq \prod_{k=1}^n \mathbb{E}\left[e^{-nc}\mathbb{1}_{\{X_k>c\}} + e^0 \mathbb{1}_{\{X_k \leq c\}}\right]^{\frac{1}{n}} \\
&\leq \prod_{k=1}^n \left(e^{-nc} + \max_{k\leq n} \mathbb{P}\left(X_k \leq c\right)\right)^{\frac{1}{n}} = \left(e^{-nc} + \max_{k\leq n} \mathbb{P}\left(X_k \leq c\right)\right)^{n\cdot\frac{1}{n}}
\end{aligned}$$

Para probar la segunda afirmación, sean $A_1, \ldots, A_k \in \mathscr{F}$, similarmente

$$\begin{aligned}
\mathbb{E}\left[e^{-S_n}\mathbb{1}_{\cap_k A_k}\right] = \mathbb{E}\left[\prod_{k=1}^n \mathbb{1}_{A_k} e^{-X_k}\right] &\leq \prod_{k=1}^n \mathbb{E}\left[\mathbb{1}_{A_k} e^{-nX_k}\right]^{\frac{1}{n}} \\
&\leq \prod_{k=1}^n \mathbb{E}\left[e^{-nc}\mathbb{1}_{\{X_k>c\}} + e^0 \mathbb{1}_{\{X_k\leq c,\}\cap A_k}\right]^{\frac{1}{n}} \\
&\leq \left(e^{-nc} + \max_{k\leq n} \mathbb{P}\left(X_k \leq c, A_k\right)\right)^{n\cdot\frac{1}{n}}.
\end{aligned}$$

■

Teorema 9.1.5 Lema de clases monótonas

Sea Ω un conjunto y λ un λ-sistema en Ω, es decir, una colección de subconjuntos de Ω tales que:

1) $\Omega \in \lambda$.

2) Si $A, B \in \lambda$ y $A \subset B$, entonces $B \setminus A \in \lambda$.

3) Si $\{A_n\}$ es una colección creciente de elementos en λ, entonces $\cup A_n \in \lambda$.

Si π es un π-sistema en Ω (es decir, es estable ante intersecciones finitas) y $\pi \subset \lambda$, entonces $\sigma(\pi) \subset \lambda$.

La prueba se puede consultar en LeGall [11], página 261.

9.2 Teoría de procesos, tiempos de paro

9.2.1 Movimiento browniano y proceso Poisson

En esta sección incluimos resultados y definiciones conocidas de teoría de procesos.

Definición 9.2.1 Movimiento Browniano
Un proceso $B = (B)_{t\geq 0}$ *en* $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ *es un movimiento browniano si satisface:*
i) Las trayectorias de B *son continuas c.s.*
ii) $B_0 = 0$ *c.s.*
iii) Para $0 \leq s \leq t$, $B_t - B_s$ *es independiente de* $\{B_u : u \leq s\}$.
iv) Para $0 \leq s \leq t$, $B_t - B_s$ *tiene la misma distribución que* B_{t-s}.
v) Para cada $t > 0$, B_t *se distribuye* $\mathcal{N}(0, t)$.

La existencia y propiedades de B se pueden consultar en cualquier texto de procesos estocásticos, por ejemplo en Revuz, Yor [2], LeGall [11] y Peres, Morters [14].

Definición 9.2.2 Proceso Poisson
Un proceso $N = (N_t)_{t\geq 0}$ *en* $(\Omega, \mathscr{F}, (\mathscr{F}_t), \mathbb{P})$ *es un proceso Poisson con intensidad* $c > 0$ *si satisface las siguientes propiedades:*
i) Las trayectorias de N *son cadlag c.s.*
ii) $N_0 = 0$ *c.s.*
iii) Para $0 \leq s \leq t$, $N_t - N_s$ *es independiente de* $\{N_u : u \leq s\}$.
iv) Para $0 \leq s \leq t$, $N_t - N_s$ *tiene la misma distribución que* N_{t-s}.
v) Para cada $t > 0$, N_t *se distribuye Poisson*(ct).

9.2.2 Tiempos de paro

En esta sección se pueden hallar las pruebas de propiedades sobre tiempos de paro que se usan continuamente en el texto. Seguimos a LeGall [11], capítulo 3.

Caracterizaremos cuando T es un tiempo de paro con respecto a $(\mathscr{G}_t)$ así como la σ-álgebra parada $\mathscr{G}_T$ en términos de $(\mathscr{F}_t)$.

Proposición 9.2.3 *Escribimos* $\mathscr{G}_t = \mathscr{F}_{t+}$ *para todo* $t \geq 0$.
1) *Son equivalentes:*

(*i*) T *es tiempo de paro c.r.a.* $(\mathscr{G}_t)$

(*ii*) $\{T < t\} \in \mathscr{F}_t$ *para cada* $t \geq 0$.

(*iii*) $T \wedge t$ *es* $\mathscr{F}_t$ *medible para cada* $t \geq 0$.

2) *Sea* T *tiempo de paro c.r.a.* $(\mathscr{G}_t)$. *Entonces,*

$$\mathscr{G}_T = \big\{A \in \mathscr{F}_\infty\,;\, \forall\, t \geq 0,\ A \cap \{T < t\} \in \mathscr{F}_t\big\}.$$

Denotamos $\mathscr{F}_{T+} = \mathscr{G}_T$.

Demostración

Prueba de $a)$

$(i) \Rightarrow (ii)$ Supongamos T es tiempo de paro c.r.a. $(\mathscr{G}_t)$.

Entonces si $s < t$, $\{T \leq s\} \in \mathscr{G}_s = \mathscr{F}_{s+} \subset \mathscr{F}_t$. Por lo tanto,

$$\{T < t\} = \bigcup_{\substack{s<t \\ s\in\mathbb{Q}}} \{T \leq s\} \in \mathscr{F}_t.$$

$(ii) \Rightarrow (i)$ Supongamos $\{T < t\} \in \mathscr{F}_t$ para cada $t \geq 0$. Para cada $s > t$,

$$\{T \leq t\} = \bigcap_{\substack{t<q<s \\ q\in\mathbb{Q}}} \{T < q\} \in \mathscr{F}_s.$$

Como esto es para cada $s > 0$, entonces $\{T \leq t\} \in \cap_{t<s}\mathscr{F}_s = \mathscr{F}_{t+}$ por lo que T es tiempo de paro c.r.a. $(\mathscr{G}_t)$.

$(ii) \Leftrightarrow (iii)$. Sea $t \geq 0$ arbitrario. Para $s \in \mathbb{R}$,

$$\{T \wedge t < s\} = \begin{cases} \Omega & \text{si } s > t \\ \{T < s\} & \text{si } s \leq t \end{cases}$$

Si el proceso $T \wedge t$ es $\mathscr{F}_t$-medible, entonces en particular, tomando $s = t$ se tiene que $\{T < t\} \in \mathscr{F}_t$. Al ser t cualquiera, se tiene (ii). Ahora, si $\{T < t\} \in \mathscr{F}_t$ para todo $t \geq 0$, entonces si $s \leq t$, $\{T < s\} \in \mathscr{F}_s \subseteq \mathscr{F}_t$ y como $\Omega \in \mathscr{F}_t$, se sigue que $T \wedge t$ es adaptado.

Prueba de $b)$

Si T es tiempo de paro c.r.a. $(\mathscr{G}_t)$, por definición

$$\mathscr{G}_T = \big\{A \in \mathscr{F}_\infty\,;\, \forall\, t \geq 0,\ A \cap \{T \leq t\} \in \mathscr{F}_{t+}\big\}.$$

Veamos que es equivalente a

$$\big\{A \in \mathscr{F}_\infty\,;\, \forall\, t \geq 0,\ A \cap \{T < t\} \in \mathscr{F}_t\big\} \tag{9.2.1}$$

Sea $A \in (9.2.1)$. Para cada $t > 0$ y $s > t$

$$A \cap \{T \leq t\} = \bigcap_{\substack{t<q<s \\ q\in\mathbb{Q}}} A \cap \{T < q\} \in \mathscr{F}_s$$

Entonces, $A \cap \{T \leq t\} \in \mathscr{F}_s$ para toda $s > t$ y por lo tanto está en $\cap_{t<s}\mathscr{F}_s = \mathscr{F}_{t^+}$. Concluimos entonces que $A \in \mathscr{G}_T$. Ahora, si $A \in \mathscr{G}_T$, escribimos

$$A \cap \{T < t\} = \bigcup_{\substack{s<t \\ s\in\mathbb{Q}}} A \cap \{T \leq s\}$$

y como $\{T \leq s\} \in \mathscr{F}_{s+} \subset \mathscr{F}_t$ para toda $s < t$, entonces $A \cap \{T < t\} \in \mathscr{F}_t$ por lo que $A \in (9.2.1)$. ■

A continuación, probaremos una colección de propiedades de los tiempos de paro y sus respectivas σ-álgebras que serán de utilidad.

Proposición 9.2.4 Primera lista de propiedades

a) Para cada tiempo de paro T, $\mathscr{F}_T \subseteq \mathscr{F}_{T+}$. Si la filtración es continua por la derecha, $\mathscr{F}_T = \mathscr{F}_{T+}$.
b) Si $T \equiv t$, entonces $\mathscr{F}_T = \mathscr{F}_t$ y $\mathscr{F}_{T+} = \mathscr{F}_{t+}$
c) Sea T es un tiempo de paro, entonces es $\mathscr{F}_T$-medible.
d) Sea T un tiempo de paro y $A \in \mathscr{F}_\infty$. Definimos

$$T^A(\omega) = \begin{cases} T(\omega) & \text{si } \omega \in A \\ \infty & \text{si } \omega \notin A \end{cases}$$

Entonces, $A \in \mathscr{F}_T$ si y solo si T^A es tiempo de paro.
e) S, T tiempos de paro con $S \leq T$. Entonces,

$$\mathscr{F}_S \subseteq \mathscr{F}_T \quad y \quad \mathscr{F}_{S+} \subseteq \mathscr{F}_{T+}$$

Demostración
a) Como $A \in \mathscr{F}_T$, $A \cap \{T \leq t\} \in \mathscr{F}_t$ para todo t.

$$A \cap \{T < t\} = \bigcup_{\substack{s<t \\ s\in\mathbb{Q}}} \underbrace{(A \cap \{T \leq s\})}_{\in \mathscr{F}_s \subseteq \mathscr{F}_t} \in \mathscr{F}_t$$

así que por la caracterización de $\mathscr{F}_{T+}$ que se dio en (ii), $A \in \mathscr{F}_{T+}$.
Si $(\mathscr{F}_t)$ es continua por la derecha, $\mathscr{F}_t = \mathscr{F}_{t+}$ para todo t y por lo tanto,

$$\begin{aligned} \mathscr{F}_{T+} &= \left\{A \in \mathscr{F}_\infty\, ; \forall\, t > 0,\, A \cap \{T \leq t\} \in \mathscr{F}_{t+}\right\} \\ &= \left\{A \in \mathscr{F}_\infty\, ; \forall\, t > 0,\, A \cap \{T \leq t\} \in \mathscr{F}_{t}\right\} = \mathscr{F}_T \end{aligned}$$

b) Es directo pues $A \cap \{T \leq t\} = A$, de modo que $A \in \mathscr{F}_T$ si y solo si $A \in \mathscr{F}_t$.
c) Para probar que T es $\mathscr{F}_T$-medible, basta ver que para todo $s \in \mathbb{R}^+ \cup \infty$, $\{T \leq s\} \in \mathscr{F}_T$. Pero

$$\{T \leq s\} \cap \{T \leq t\} = \{T \leq s \wedge t\} \in \mathscr{F}_{s\wedge t} \subseteq \mathscr{F}_t$$

por lo que $\{T \leq s\} \in \mathscr{F}_T$ para toda s y entonces $\{T = \infty\} = \cap_n \{T > n\} \in \mathscr{F}_T$ y concluimos que T es $\mathscr{F}_T$-medible.
d) Como

$$\{T^A \leq t\} = \{T \leq t\} \cap A \in \mathscr{F}_t$$

entonces es claro que si T^A es tiempo de paro, $\{T \leq t\} \cap A \in \mathscr{F}_t$ por lo que $A \in \mathscr{F}_T$. Recíprocamente, si $A \in \mathscr{F}_T$ entonces $\{T \leq t\} \cap A \in \mathscr{F}_t$ por lo que T^A es tiempo de paro.
e) Supongamos S, T son tiempo de paro con $S \leq T$ y sea $A \in \mathscr{F}_S$. Como $A \cap \{S \leq t\} \in \mathscr{F}_t$ para todo $t \geq 0$ entonces,

$$A \cap \{T \leq t\} = \underbrace{(A \cap \{S \leq t\})}_{\in \mathscr{F}_t} \cap \underbrace{\{T \leq t\}}_{\in \mathscr{F}_t}$$

ya que $\{T \leq t\} \subset \{S \leq t\}$. Entonces, $A \in \mathscr{F}_T$. La prueba para $\mathscr{F}_{S+} \subseteq \mathscr{F}_{T+}$ es similar intercambiando $\leq$ por $<$.

El siguiente corolario se usa al probar que se puede aproximar cualquier tiempo de paro por una sucesión de tiempos de paro discretos decreciente:

Corolario 9.2.5 *Si T es un tiempo de paro y S es $\mathscr{F}_T$-medible con $S \geq T$, entonces S es tiempo de paro.*

Demostración
Para cada $t \geq 0$: $\{S \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t$ pues $\{S \leq t\} \in \mathscr{F}_T$ por ser S una variable aleatoria $\mathscr{F}_T$-medible. Por lo tanto, S es un tiempo de paro.

Proposición 9.2.6 Segunda lista de propiedades
Sean S, T tiempos de paro y y $(T_n)_{n\in\mathbb{N}}$ sucesión de tiempos de paro.
Entonces,

$$S \vee T, \quad S \wedge T, \quad \sup_n \{T_n\}$$

son tiempo de paro. En particular, si (T_n) es sucesión creciente, $\lim_{n\to\infty} T_n = T$ existe y es un tiempo de paro. Además

$$\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T \quad y \quad \{S \leq T\}, \{S = T\} \in \mathscr{F}_{S\wedge T}.$$

Demostración

$$\{S \vee T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \in \mathscr{F}_t \text{ por lo que } S \vee T \text{ es tiempo de paro.}$$
$$\{S \wedge T \leq t\} = \{S \leq t\} \cup \{T \leq t\} \in \mathscr{F}_t \text{ por lo que } S \wedge T \text{ es tiempo de paro.}$$
$$\Big\{\sup_n T_n \leq t\Big\} = \bigcap_n \{T_n \leq t\} \in \mathscr{F}_t \text{ por lo que } \sup_n T_n \text{ es tiempo de paro.}$$

Ahora, probemos que $\mathscr{F}_{S\wedge T} = \mathscr{F}_S \cap \mathscr{F}_T$. Como $S \wedge T$ es tiempo de paro y $S \wedge T \leq T, S$ y entonces por e) $\mathscr{F}_{S\wedge T} \subseteq \mathscr{F}_S, \mathscr{F}_T$ por lo que tenemos la primer contención. Por otra parte, si tomamos $A \in \mathscr{F}_S \cap \mathscr{F}_T$, entonces $A \cap \{S \leq t\} \in \mathscr{F}_t$ y $A \cap \{T \leq t\} \in \mathscr{F}_t$. Pero

$$A \cap \{S \wedge T \leq t\} = A \cap (\{S \leq t\} \cap \{T \leq t\}) = (A \cap \{S \leq t\}) \cap (A \cap \{T \leq t\}) \in \mathscr{F}_t$$

por lo que $A \in \mathscr{F}_{S\wedge T}$.
Finalmente, probemos que $\{S \leq T\}$ y $\{S = T\} \in \mathscr{F}_{S\wedge T}$.

$$\{S \leq T\} \cap \{T \leq t\} = \{S \leq t\} \cap \{T \leq t\} \cap \{S \wedge t \leq T \wedge t\} \in \mathscr{F}_t$$
$$\{S \leq T\} \cap \{S \leq t\} = \{S \wedge t \leq T \wedge t\} \cap \{S \leq t\} \in \mathscr{F}_t$$

ya que por la proposición 9.2.3 $T \wedge t$ es $\mathscr{F}_t$-medible. Como $\mathscr{F}_T \cap \mathscr{F}_S = \mathscr{F}_{S\wedge T}$, $\{S \leq T\} \in \mathscr{F}_{S\wedge T}$.
Esto nos permite concluir que $\{S = T\} = \{S \leq T\} \cap \{S \geq T\} \in \mathscr{F}_{S\wedge T}$. ■

Proposición 9.2.7 Tercer lista de propiedades (Tiempos opcionales y filtraciones cad)
Sea $(T_n)_{n\in\mathbb{N}}$ sucesión de tiempos de paro.
f) $\inf_n\{T_n\}$ es tiempo de paro c.r.a $(\mathscr{F}_{t+})$. En particular, si (T_n) es decreciente, el límite

$\lim_{n\to\infty} T_n = T$ *existe y es un tiempo de paro c.r.a.* $(\mathscr{F}_{t+})$.
g) Si $T_n \downarrow T$

$$\mathscr{F}_{T+} = \bigcap_n \mathscr{F}_{T_n+} \quad \text{y entonces si la filtración es cád:} \quad \mathscr{F}_T = \bigcap_n \mathscr{F}_{T_n}.$$

h) Si la filtración es continua por la derecha, entonces

$$\liminf_n T_n \quad y \quad \limsup_n T_n$$

son tiempos de paro. [1]

Demostración
f) Para cada $t \geq 0$

$$\big\{ \inf_n \{T_n\} < t \big\} = \bigcup_n \{T_n < t\} \in \mathscr{F}_t$$

y por la proposición 9.2.3 $\inf_n\{T_n\}$ es tiempo de paro c.r.a. $(\mathscr{F}_{t+})$. Nótese que la prueba sigue siendo válida si T_n son tiempos de paro c.r.a. $(\mathscr{F}_{t+})$.
g) Para cada n, $\mathscr{F}_{T+} \subseteq \mathscr{F}_{T_n+}$ así que la primer contención es directa. Ahora, sea $A \in \mathscr{F}_{T_n+}$ para toda n. Como $\{T < t\} = \liminf_n \{T_n < t\}$,

$$\begin{aligned} A \cap \{T < t\} = A \cap \liminf_n \{T_n < t\} &= A \cap \left(\bigcap_n \bigcup_{m\geq n} \{T_n < t\} \right) \\ &= \bigcap_n \bigcup_{m\geq n} \underbrace{(A \cap \{T_n < t\})}_{\in \mathscr{F}_t} \in \mathscr{F}_t \end{aligned}$$

así que $A \in \mathscr{F}_{T+}$ y por lo tanto tenemos la segunda contención.
h) Probemos que $\liminf_n T_n$ es tiempo de paro si $(\mathscr{F}_t)$ es cad (en este caso, todo tiempo de paro opcional es tiempo de paro).

$$\big\{ \liminf_n T_n > t \big\} = \big\{ \lim_{n\to\infty} \inf_{m>n} \{T_m\} \geq t \big\} = \bigcup_n \left\{ \inf_{m\geq n} \{T_m\} > t \right\} \in \mathscr{F}_t$$

pues para cada n, $\inf_{m\geq n}\{T_m\}$ es tiempo de paro c.r.a. $(\mathscr{F}_{t+})$ pero al ser $(\mathscr{F}_t)$ cád, es tiempo de paro.
Otra forma de verlo es la siguiente: la sucesión $(\inf_{m\geq n}\{T_m\})_n$ es decreciente y por lo que acabamos de probar, para cada n, $\inf_{m\geq n}\{T_m\}$ es un tiempo de paro (opcional). Como ínfimo de tiempos de paro vuelve a ser tiempo de paro (opcional), la siguiente expresión es tiempo de paro:

$$\inf_n \left\{ \inf_{m\geq n} \{T_m\} \right\} = \lim_{n\to\infty} \inf_{m\geq n} \{T_m\} = \liminf_n T_n$$

y por lo tanto $\liminf_n T_n$ es tiempo de paro. Para probar que $\limsup_n T_n$ es tiempo de paro, hacemos algo similar:

$$\big\{ \limsup_n T_n < t \big\} = \big\{ \lim_{n\to\infty} \sup_{m\geq n} \{T_m\} < t \big\} = \bigcup_n \left\{ \sup_{m\geq n} \{T_m\} < t \right\} \in \mathscr{F}_t.$$

■

[1] Si no lo es, no tienen por que ser ni siquiera tiempos de paro opcionales.

9.3 El espacio de las martingalas cuadrado-integrables $\mathcal{M}_T$

Referencia: Ash & Doléans-Dade [1] capítulo 9.

Definición 9.3.1 *Sea* $(\Omega, \mathscr{F}, \mathbb{P})$ *un espacio de probabilidad y* $(\mathscr{F}_t)_{t\geq 0}$ *una familia no decreciente de sub-*σ*-álgebras de* $\mathscr{F}$*. Una familia de variables aleatorias* $X = (X_t)_{t\geq 0}$ *es martingala con respecto a la filtración* $(\mathscr{F}_t)_{t\geq 0}$ *si:*

a) Para cada t*,* X_t *es* $\mathscr{F}_t$ *medible e integrable.*
b) Si $s < t$*,* $\mathbb{E}[X_t|\mathscr{F}_s] = X_s$ *c.s.*

Definición 9.3.2 *Si* X *es un proceso estocástico, decimos que* Y *es una versión de* X *si, para todo* t*,* $\mathbb{P}(X_t = Y_t) = 1$*.*

Teorema 9.3.3 *Sea* $(\Omega, \mathscr{F}, \mathbb{P})$ *un espacio de probabilidad y* $(\mathscr{F}_t)_{t\geq 0}$ *filtración continua por la derecha. Asumiremos que* $\mathscr{F}$ *es completa (es decir, si* $A \in \mathscr{F}$ *y* $\mathbb{P}(A) = 0$*, entonces para todo* $B \subset A$*,* $B \in \mathscr{F}$*) y para todo* t*,* $\mathscr{F}_t$ *contiene al conjunto* $\mathcal{N} = \{A \in \mathscr{F}\,;\, \mathbb{P}(A) = 0\}$*. Entonces, cualquier martingala* X *con respecto a* $(\mathscr{F}_t)_{t\geq 0}$ *admite una versión continua por la derecha.*

Demostración
Sea S un denso numerable de $\mathbb{R}^+$, a y b dos reales con $a < b$ y $S_n = S \cup [0, n]$. Nótese que si X es martingala continua, la restricción a una cantidad numerable de índices I es una martingala discreta con respecto a $(\mathscr{F}_t)_{t\in I}$. Sea $I \subset S_n$ finito; si $U_{ab}(I)$ es el número de *upcrossings* de $\{X_t\}_{t\in I}$ del intervalo (a, b), por el teorema de *upcrossings* de Doob:

$$\mathbb{E}[U_{ab}] \leq \frac{1}{b-a}\mathbb{E}\left[(X_n - a)^+\right] \qquad \text{ya que } t \leq n \text{ para todo } t \in I$$

y entonces, por Chevychev

$$\mathbb{P}(U_{ab}(I) \geq \lambda) \leq \frac{1}{\lambda}\mathbb{E}[U_{ab}(I)] \leq \frac{1}{\lambda(b-a)}\mathbb{E}\left[(X_n - a)^+\right] \tag{9.3.1}$$

Sea I_k $k = 1, 2, \ldots$ sucesión creciente de subconjuntos de S_n tal que $\cup_k I_k = S_n$. Veamos que X restringido a S_n tiene un número finito de cruces c.s., es decir: $\mathbb{P}(\lim_k U_{ab}(I_k) = \infty) = 0$. Supongamos que existe $\alpha > 0$ tal que $\mathbb{P}(\lim_k U_{ab}(I_k) = \infty) = \alpha$. Como $I_k \subset I_{k+1}$, entonces para cualquier $\lambda \in \mathbb{R}^+$, $\{U_{ab}(I_k) \geq \alpha\} \subset \{U_{ab}(I_{k+1}) \geq \alpha\}$ por lo que la sucesión de conjuntos es creciente. Entonces, para toda λ

$$\begin{aligned}
\lim_{k\to\infty} \mathbb{P}(U_{ab}(I_k) \geq \lambda) &= \mathbb{P}\left(\cup_{k=1}^{\infty}\{U_{ab}(I_k) \geq \lambda\}\right) \\
&= \mathbb{P}\left(\liminf_k \{U_{ab}(I_k) \geq \lambda\}\right) \\
&= \mathbb{P}\left(\exists\, N_\omega \text{ tal que } U_{ab}(I_k)(\omega) \geq \lambda \quad \forall\, n \geq N_\omega\right) \\
&= \mathbb{P}\left(\lim_k U_{ab}(I_k) \geq \lambda\right) \\
&> \frac{\alpha}{2}
\end{aligned}$$

pues $\mathbb{P}(\lim_k U_{ab}(I_k) = \infty) \leq \mathbb{P}(\lim_k U_{ab}(I_k) \geq \alpha)$. Entonces, por (9.3.1) tenemos que para toda $\lambda \geq 0$ se satisface

$$\frac{\alpha}{2} < \mathbb{P}\left(\lim_k U_{ab}(I_k) \geq \lambda\right) \leq \frac{1}{b-a}\mathbb{E}\left[(X_n - a)^+\right]$$

lo cual lleva a una contradicción para λ suficientemente grande. Por lo tanto, se tiene que $\mathbb{P}\left(\lim_k U_{ab}(I_k) = \infty\right) = 0$. Ahora, si

$$A_{n,\,ab} = \left\{\omega\,;\, \exists\, t \in [0,n) \text{ tal que } \liminf_{s\downarrow t\ s\in S} X_s < a < b < \limsup_{s\downarrow t\ s\in S} X_s\right\}$$

entonces $A_{n,\,ab} \subset \{\lim_k U_{ab}(I_k) = \infty\}$ pues para que omega esté en $A_{n,\,ab}$, X debe tener una infinidad de saltos en una vecindad de t.
Por ser $\mathscr{F}$ completa, $A_{n,\,ab}$ es medible y $\mathbb{P}\left(A_{n,\,ab}\right) = 0$. Definimos al conjunto A en el cual X no tiene límite por la derecha en algún t restringiéndonos a sucesiones en el denso S:

$$A = \left\{\omega\,;\, \exists\, t \in \mathbb{R}^+ \text{ tal que } \liminf_{s\downarrow t\ s\in S} X_s < \limsup_{s\downarrow t\ s\in S} X_s\right\}.$$

Claramente

$$A \subset \bigcup_{n\in\mathbb{N}} \bigcup_{\substack{a<b \\ a,\,b\in\mathbb{Q}}} A_{n,\,ab} \qquad \text{y por lo tanto} \quad \mathbb{P}\left(A\right) = 0$$

Definimos al proceso $Y = \{Y_t\}_{t\geq 0}$ como sigue:

$$Y_t(\omega) = \begin{cases} \lim_{s\downarrow t\ s\in S} X_s(\omega) & \omega \in A^c \\ 0 & \omega \in A. \end{cases}$$

Veamos que es adaptado:

$$\{Y_t \geq \alpha\} = \begin{cases} \bigcup_{n\in\mathbb{N}} \left\{X_s \geq \alpha \quad s \in \left(t, t+n^{-1}\right) \cap S\right\} = \bigcup_{n\in\mathbb{N}} \bigcap_{s\in S\cap(t,t+\frac{1}{n})} \{X_s \geq \alpha\} & \text{si } \alpha \geq 0 \\ A \cup \bigcup_{n\in\mathbb{N}} \bigcap_{s\in S\cap(t,t+\frac{1}{n})} \{X_s \geq \alpha\} \quad \text{si } \alpha < 0 \end{cases}$$

donde $\bigcap_{s\in S\cap(t,t+\frac{1}{n})} \{X_s \geq \alpha\} \in \cup_{s>t}\mathscr{F}_s = \mathscr{F}_t^+ = \mathscr{F}_t$ por ser la filtración continua por la derecha y $A \in \mathscr{F}_t$ para toda t por ser de medida 0 (aquí se están usando las dos hipótesis que se pidieron para la filtración) por lo que Y es adaptado. Para cada $\omega \in A^c$ y $t \in \mathbb{R}^+$, existe $\delta > 0$ tal que $|Y_t(\omega) - X_s(\omega)| < \epsilon$ si $t - s < \delta$ con $s \in S$ y $s > t$. Por lo tanto para cualquier $u \in (t, t+\delta)$:

$$\epsilon > \lim_{s\downarrow u\ s\in S} |Y_t(\omega) - X_s(\omega)| = |Y_t(\omega) - Y_u(\omega)|$$

por lo que Y es continuo por la derecha. Falta probar que Y es versión de X y que es martingala. Como X es martingala, $X_s = \mathbb{E}\left[X_u|\mathscr{F}_s\right]$ si $s < u$. Pero por un lado, $\lim_{s\downarrow t\ s\in S} X_s = Y_t$ c.s. y por otro lado, $\lim_{s\downarrow t\ s\in S} \mathbb{E}\left[X_u|\mathscr{F}_s\right] = \mathbb{E}\left[X_u|\mathscr{F}_t^+\right] = \mathbb{E}\left[X_u|\mathscr{F}_t\right] = X_t$ de donde $Y_t = X_t$ c.s. es decir, $\mathbb{P}\left(X_t = Y_t\right) = 1$ por lo que Y es versión de X. Para ver que Y es martingala, obsérvese lo siguiente: para todo $B \in \mathscr{F}_s$ se satisface:

$$\int_B Y_s\, d\mathbb{P} = \int_B Y_s\, d\mathbb{P} = \int_B \mathbb{E}\left[X_t|\mathscr{F}_s\right] d\mathbb{P} = \int_B X_t\, d\mathbb{P} = \int_B Y_t\, d\mathbb{P} = \int_B \mathbb{E}\left[Y_t|\mathscr{F}_s\right] d\mathbb{P}$$

por lo que $Y_s = \mathbb{E}\left[Y_t|\mathscr{F}_s\right]$ c.s. ya que ambas son variables aleatorias $\mathscr{F}_s$ medibles ■

Obsérvese como una consecuencia de la prueba anterior es la

Proposición 9.3.4 *Si (X_t) es una martingala, entonces para casi todo ω, el límite*

$$\lim_{r\downarrow t,\, r\in\mathbb{Q}} X_r(\omega)$$

existe para cada $t \in [0,\infty)$. Si se remplaza a X por una supermartingala, el resultado sigue siendo cierto.

Solo estamos omitiendo la prueba cuando X es supermartingala. Esto permite probar que las trayectorias de X son de hecho cadlag casi seguramente:

Teorema 9.3.5 *Si X es un supermartingala continua por la derecha, entonces*
1) X es supermartigala con respecto a $(\mathscr{F}_{t+})$ y con respecto a su compleción.
2) Casi cada trayectoria de X es cadlag.

Este hecho no se usará hasta el capítulo 5 y solo mencionamos que permite probar una versión más fuerte del teorema de regularización:

Teorema 9.3.6 *Sea $(\mathscr{F}_t)$ filtración continua por la derecha y completa. Sea $X = (X_t)_{t\geq 0}$ supermartingala tal que $t \to \mathbb{E}[X_t]$ es continua por la derecha. Entonces, X admite una modificación cadlag que vuelve a ser supermartingala.*

Las pruebas se pueden consultar en Revuz&Yor capítulo II sección 2.

Regresemos al estudio de las martingalas cuadrado integrables. Los siguientes dos resultados nos permitirán probar que el espacio de las martingalas cuadrado integrables es un espacio de Hilbert, resultado que necesitamos para la descomposición de Lévy-Itô.

Lema 9.3.7 *Sea X submartingala con respecto a $(\mathscr{F}_t)_{t\geq 0}$. Entonces, para todo $\lambda > 0$*

$$\mathbb{P}\left(\sup_{s\leq t} X_s > \lambda\right) \leq \int_{\sup_{s\leq t} X_s > \lambda} X_t \, d\mathbb{P}$$

Demostración
Primero, lo probamos para un número finito de índices $t_1 < \cdots < t_k$. Sea T el siguiente tiempo de paro:

$$T(\omega) = \begin{cases} \min\{t_i\,;\, X_{t_i}(\omega) > \lambda\} & \text{si } \omega \in \{\max_i X_{t_i} > \lambda\} \\ t_k & \text{si } \omega \in \{\max_i X_{t_i} \leq \lambda\} \end{cases}$$

y como $\{\max_i X_{t_i} > \lambda\} = \cup_{i=1}^{k-1}\{T = t_i\} \cup \{T = t_k\,,\ \max_i X_{t_i} > \lambda\}$, se sigue que:

$$\begin{aligned} \lambda\mathbb{P}\left(\max_i X_{t_i} > \lambda\right) &= \mathbb{E}\left[\lambda \mathbb{1}_{\{\max_i X_{t_i} > \lambda\}}\right] \\ &= \mathbb{E}\left[\lambda \sum_{i=1}^{k-1} \mathbb{1}_{\{T=t_i\}} + \lambda \mathbb{1}_{\{T=t_k\,,\max_{t_i} X_{t_i} > \lambda\}}\right] \\ &\leq \sum_{i=1}^{k-1} \mathbb{E}\left[X_{t_i} \mathbb{1}_{\{T=t_i\}}\right] + \mathbb{E}\left[X_{t_k} \mathbb{1}_{\{T=t_k\,,\max_i X_{t_i} > \lambda\}}\right] \\ &\leq \sum_{i=1}^{k-1} \mathbb{E}\left[X_{t_k} \mathbb{1}_{\{T=t_i\}}\right] + \mathbb{E}\left[X_{t_k} \mathbb{1}_{\{T=t_k\,,\max_i X_{t_i} > \lambda\}}\right] \\ &= \int_{\max_i X_{ti} > \lambda} X_{t_k}. \end{aligned}$$

En la cuarta desigualdad, se está usando que, al ser T un tiempo de paro y X submartingala, $\{T = t_i\} \in \mathscr{F}_{t_i}$ y entonces, $\int_{T=t_i} X_{t_k}\, d\mathbb{P} = \int_{T=t_i} \mathbb{E}\left[X_{t_k}|\mathscr{F}_{t_i}\right] d\mathbb{P} \geq \int_{T=t_i} X_{t_i}\, d\mathbb{P}$. Ahora, consideremos una sucesión creciente de conjuntos finitos I_k con $I_0 = \{t\}$ tales que $\cup_k I_k = \mathbb{Q}_t$ donde $\mathbb{Q}_t = \mathbb{Q} \cap [0,t] \cup \{t\}$. Para $k \in \mathbb{N}$, la sucesión de conjuntos $\{\max_{t_i \in I_k} X_{t_i} > \lambda\}$ crece a $\{\sup_{s\in\mathbb{Q}_t} X_s > \lambda\}$ y para cada k, se satisface

$$\lambda\mathbb{P}\left(\max_{t_i\in I_k} X_{t_i} > \lambda\right) = \int_{\max\limits_{t_i\in I_k} X_{t_i} > \lambda} X_t\, d\mathbb{P}$$

por lo que, al tomar el límite cuando k tiende a infinito de ambos lados obtenemos, por el teorema de convergencia dominada

$$\lambda\mathbb{P}\left(\sup_{s\in\mathbb{Q}_t} X_s > \lambda\right) = \int_{\sup\limits_{s\in\mathbb{Q}_t} X_s > \lambda} X_t\, d\mathbb{P}$$

pues X es integrable. Como X es continua por la derecha, $\sup_{s\le t} X_s = \sup_{s\in\mathbb{Q}_t} X_s$ por lo que concluimos

$$\lambda\mathbb{P}\left(\sup_{s\le t} X_s > \lambda\right) = \int_{\sup_{s\le t} X_s > \lambda} X_t\, d\mathbb{P}$$

■

Observación 9.3.8 *Para probar* $\sup_{s\le t} X_s = \sup_{s\in\mathbb{Q}_t} X_s$, *sea* $\{t_k\}$ *sucesión en* $[0,t]$ *tal que* $X_{t_k} \to \sup_{s\le t} X_s$. *Como la sucesión* t_k *es acotada, existe una subsucesión* $\{t_{k_j}\}$*convergente, y converge a algún* $r \in [0,t]$*; además, se sigue cumpliendo que* $X_{t_{k_j}} \to \sup_{s\le t} X_s$. *Ahora, como* X *es continuo por la derecha, para cada* t_k *existe* $s_k \in S$ *con* $t_k \le s_k$ *tal que* $|X_{t_k} - X_{s_k}| < \frac{1}{k}$. *No es difícil ver que* $X_{s_k} \to \sup_{s\le t} X_s$.

Teorema 9.3.9 Desigualdad de martingalas de Doob
Si X *es una martingala continua por la derecha con respecto a una filtración* $(\mathscr{F}_t)_{t\ge 0}$ *y* $X_t \in L^2$ *para cada* t*, entonces:*

$$\left\|\sup_{s\le t} |X_s|\right\|_2 \le 2\,\|X_t\|_2 \tag{9.3.2}$$

Sea $Y = \sup_{s\le t} |X_s|$. Como $|x|$ es convexa, $|X|$ es sub-martingala y por el lema 9.3.7 $\lambda\mathbb{P}(Y > \lambda) \le \int_{Y>\lambda} |X_t| d\mathbb{P}$. Como aún no sabemos si $\|Y\|_2 < \infty$, definimos $Y_n = Y \wedge n$. Como $\{Y_n > \lambda\}$ es vacío si $\lambda \ge n$ y coincide con $\{Y > \lambda\}$ si $\lambda < n$, entonces

$$\lambda\mathbb{P}(Y_n > \lambda) \le \int_{Y_n>\lambda} |X_t| d\mathbb{P} \tag{9.3.3}$$

y como $s^2 = \int_0^s 2\lambda d\lambda$:

$$
\begin{aligned}
\mathbb{E}\left[Y_n^2\right] = \int_0^\infty s^2\, \mathbb{P}_{Y_n}(ds) &= \int_0^\infty \int_0^s (2\lambda)\, d\lambda\, \mathbb{P}_{Y_n}(ds) \\
&= \int_0^\infty 2\lambda \mathbb{P}\left(Y_n > \lambda\right) d\lambda \\
&= 2\int_0^\infty \int_\Omega |X_t| \mathbb{1}_{\{Y_n>\lambda\}} d\mathbb{P} d\lambda \qquad \text{por (9.3.3)} \\
&= 2\int_\Omega |X_t| \int_0^\infty \mathbb{1}_{\{Y_n>\lambda\}} d\lambda d\mathbb{P} \\
&= 2\int_\Omega |X_t| \int_0^{Y_n} 1 d\lambda d\mathbb{P} \\
&= 2\mathbb{E}\left[|X_t| Y_n\right] \\
&\leq 2\|X_t\|_2 \|Y_n\|_2
\end{aligned}
$$

por lo que $\|Y_n\|_2 \leq 2\|X_t\|_2$ y al tomar el límite cuando $n \to \infty$ se obtiene el resultado.

Definición 9.3.10 *Sea $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ espacio de probabilidad completo con $(\mathscr{F}_t)_{t\geq 0}$ filtración continua por la derecha y tal que para cada t, $\mathscr{F}_t$ contiene al conjunto $\mathcal{N} = \{A \in \mathscr{F}\,;\, \mathbb{P}(A) = 0\}$. Denotaremos por $\mathcal{M}_T$ al espacio vectorial de los procesos $X = \{X_t\}_{0\leq t\leq T}$ en $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geq 0}, \mathbb{P})$ tal que X es una $(\mathscr{F}_t)-$martingala continua por la derecha y $X_T \in L^2$.*

Si Y es versión de X en $\mathcal{M}_T$, entonces para todo J denso numerable de $[0,T]$,

$$\mathbb{P}\left(X_t = Y_t\,;\, \forall\, t \in J\right) = 1$$

y por continuidad por la derecha, se tiene que $\mathbb{P}\left(X_t = Y_t\,;\, \forall\, t \in [0,t]\right) = 1$ por lo que X y Y son indistinguibles (si dos funciones continuas por la derecha coinciden en un denso, entonces coinciden en todo su dominio). En $\mathcal{M}_T$ trabajaremos bajo la clase de equivalencia $X \sim Y$ si Y es versión de X, por lo que $X = Y$ si son indistinguibles. Consideremos en $\mathcal{M}_T$ el producto escalar $\langle X, Y\rangle = \mathbb{E}\left[X_T Y_T\right]$. Como consecuencia inmediata del teorema 9.3.9, se tiene lo siguiente:

Teorema 9.3.11
i) Si $\langle X, X\rangle = 0$, entonces $X = 0$ por lo que $\|X\|_{\mathcal{M}^T} = \langle X, X\rangle$ es una norma en $\mathcal{M}_T$.
ii) Si una sucesión $\{X_n\}_{n\in\mathbb{N}}$ converge a X en $\mathcal{M}_T$, entonces

$$\sup_{t\leq T} |X_t^{(n)} - X_t| \xrightarrow{L^2} 0 \tag{9.3.4}$$

y en particular, $X_t^{(n)} \xrightarrow{L^2} X_t$ para cualquier t.

Teorema 9.3.12 *El espacio $\mathcal{M}_T$ con el producto escalar $\langle X, Y\rangle = \mathbb{E}\left[X_T Y_T\right]$ es un espacio de Hilbert y el subespacio de martingalas continuas es cerrado en $\mathcal{M}_T$.*

Demostración
Sea $\{X_n\}_{n\in\mathbb{N}}$ sucesión de Cauchy en $\mathcal{M}_T$. Por (9.3.4), $\{X_T^{(n)}\}_{n\in\mathbb{N}}$ es sucesión de Cauchy en L^2 y por ser completo, converge; sea Z_T su límite en L^2. Para toda $t \in [0,T]$, definimos

$Z_t = \mathbb{E}[Z_T|\mathscr{F}_t]$. Como $Z = \{Z_t\}_{t\leq T}$ es martingala, admite una versión continua por la derecha $Y = \{Y_t\}_{t\leq T}$ y como $Y_T = Z_T$ casi seguramente, entonces

$$\|X_n - Y\|_{\mathcal{M}_T} = \|X_T^{(n)} - Y_T\|_2 = \|X_T^{(n)} - Z_T\|_2 \longrightarrow 0 \tag{9.3.5}$$

por lo que $\mathcal{M}_T$ es completo y por lo tanto es espacio de Hilbert.
Ahora, supongamos que $X_n \overset{\mathcal{M}_T}{\to} X$ y que X_n es martingala continua para cada n. Como

$$\|X_T^{(n)} - X_T\|_{\mathcal{M}_T}^2 = \mathbb{E}\left[|X_T^{(n)} - X_T|^2\right] \to 0$$

podemos suponer sin pérdida de generalidad que $\mathbb{E}\left[|X_T^{(n)} - X_T|^2\right] < 2^{-n}$ (sino, construir una subsucesión que lo cumpla) por lo que $\sum_n \mathbb{E}\left[|X_T^{(n)} - X_T|^2\right] < \infty$. Por el teorema 9.3.9

$$\mathbb{E}\left[\sup_{t\leq T}|X_t^{(n)} - X_t|^2\right] \leq \mathbb{E}\left[|X_T^{(n)} - X_T|^2\right]$$

y entonces

$$\sum_{n=1}^{\infty}\mathbb{E}\left[\sup_{t\leq T}|X_t^{(n)} - X_t|^2\right] \leq \sum_{n=1}^{\infty}\mathbb{E}\left[|X_T^{(n)} - X_T|^2\right] < \infty. \tag{9.3.6}$$

Por otra parte, por la desigualdad de Chevychev:

$$\mathbb{P}\left(\sup_{t\leq T}|X_t^{(n)} - X_t| \geq \epsilon\right) \leq \frac{1}{\epsilon^2}\mathbb{E}\left[\sup_{t\leq T}|X_t^{(n)} - X_t|^2\right]$$

por lo que, usando (9.3.7) concluimos que

$$\sum_{n=1}^{\infty}\mathbb{P}\left(\sup_{t\leq T}|X_t^{(n)} - X_t| \geq \epsilon\right) < \infty.$$

Por Borel-Cantelli, $\mathbb{P}\left(\limsup_n\{\sup_{t\leq T}|X_t^{(n)} - X_t| \geq \epsilon\}\right) = 0$ así que tomando una sucesión $\epsilon_k \downarrow 0$,

$$\mathbb{P}\left(\left\{\sup_{t\leq T}|X_t^{(n)} - X_t| \nrightarrow 0\right\}\right) \leq \sum_{k=1}^{\infty}\mathbb{P}\left(\limsup_n\left\{\sup_{t\leq T}|X_T^{(n)} - X_T| \geq \epsilon_k\right\}\right) = 0 \tag{9.3.7}$$

por lo que $\sup_{t\leq T}|X_t^{(n)}(\omega) - X_t(\omega)| = \|X_n(\omega) - X(\omega)\|_\infty \to 0$ $\mathbb{P}$-casi seguramente. Como hay convergencia uniforma de funciones continuas salvo en un conjunto $\mathbb{P}$-nulo, entonces fuera de este su límite $X(\omega)$ es continua, y por lo tanto X es indistinguible de una martingala continua en $[0,T]$.

Nótese que la desigualdad (9.3.7) se sigue de que

$$\left\{\sup_{t\leq T}|X_t^{(n)} - X_t| \nrightarrow 0\right\} = \bigcup_k \limsup_n\left\{\sup_{t\leq T}|X_T^{(n)} - X_T| \geq \epsilon_k\right\}.$$

■

Teorema 9.3.13 Paro opcional para supermartingalas no negativas
Sea X es una supermartingala no negativa con trayectorias continuas por la derecha y sean S, T tiempos de paro con $S \leq T$,

$$X_S \geq \mathbb{E}[X_T|\mathscr{F}_S]$$

La prueba se puede consultar en Le Gall [11] página 64.

9.4 Espacios LCCB y compactación de Alexandroff

A continuación, los espacios topológicos considerados serán siempre Hausdorff. En esta sección veremos propiedades relativas a los espacios LCCD así como a su compactación de Alexandroff, que serán necesarias únicamente al probar el teorema 6.1.9. En algunos casos no daremos las pruebas y nos referimos a Blumenthal & Getoor [17].

Si E es localmente compacto, podemos considerar la clase de funciones que tienden a 0 en "infinito" $C_0(E)$ en el siguiente sentido: para todo $\epsilon > 0$ existe un compacto K tal que $|f(x)| < \epsilon$ si $x \in K^c$.

Definición 9.4.1 *Diremos que un espacio topológico E es LCCB si es localmente compacto y tiene una base numerable. Esto es:*
1) Para cada $x \in E$, existe una vecindad abierta y un compacto K tal que $x \in U \subset K$.
2) Existe una colección numerable de abiertos $\mathcal{U}$ tal que todo abierto se puede expresar como unión de una sub-familia de estos.

Si E es un espacio LCCB, entonces es metrizable y podemos considerar una métrica d compatible con la topología, de modo que (E, d) sea un espacio métrico separable y todo conjunto cerrado acotado con respecto a la métrica d sea compacto (en particular, E es Polaco). Además, bajo la norma $\|\cdot\|_\infty$, los espacios $C_0^+(E)$ y $C_0(E)$ son espacios de Banach separables.

Definición 9.4.2 *Sea E un $LCCB$ y consideremos al siguiente espacio $E_\Delta = E \cup \{\Delta\}$. Definimos las vecindades de Δ como los conjuntos de la forma $K^c \cap \{\Delta\}$ y al añadir el conjunto de abiertos de E obtenemos una topología en E_Δ. Al espacio E_Δ se le conoce como la compactación en un punto de Alexandroff.*

E_Δ es un espacio topológico compacto y no es difícil comprobar que vuelve a ser Hausdorff. Será importante probar que E_Δ sigue teniendo una base numerable pues al ser compacto, vuelve a ser un espacio LCCD. La clave estará en notar que al utilizar en E la métrica d compatible con la topología de E tal los compactos son cerrados y acotadas con respecto a d, podemos construir una base de vecindades numerable del punto al infinito Δ al considerar complementos de bolas cerradas:

Lema 9.4.3 *Si E es LCCB, la compactación en un punto E_Δ de E tiene una base numerable.*

Demostración
Sea d la metrica en E compatible con la topología tal que todo conjunto acotado (relativo a d) y cerrado es compacto y sea $\mathcal{U}$ base numerable de la topología de E. Fijemos $x_0 \in E$ y consideremos la familia numerable $\mathcal{O} = \{\overline{B(x_0, R)}^c \cup \{\Delta\} : R \in \mathbb{Q}\}$ de vecindades de Δ. Entonces, $\mathcal{B} \cup \mathcal{O}$ es base de la topología de E_Δ. En efecto, si B es un abierto que no contiene a Δ, B se escribe como unión de elementos de $\mathcal{U}$ por ser base de la topología de E. Si $\Delta \in B$, debe existir un compacto K tal que $\Delta \in K^c \subset U$. Entonces, al considerar una bola suficientemente grande, $\Delta \subset \overline{B(x_0, R)}^c \subset K^c$ y escribimos entonces a B como $\overline{B(x_0, R)}^c \cup U \setminus \overline{B(x_0, R)}^c$. Como $U \setminus \overline{B(x_0, R)}^c$ es un abierto que no contiene a Δ, por el caso anterior se escribe como unión numerable de elementos de $\mathcal{U}$ mientras que $\overline{B(x_0, R)}^c \in \mathcal{O}$, con lo que concluimos que B se escribe como unión de elementos de $\mathcal{O} \cup \mathcal{U}$.

Recordemos del capítulo 5 que toda función $f \in C_0(E)$ se puede extender naturalmente a una continua en $C(E_\Delta)$ al definir $f(\Delta) = 0$.

Lema 9.4.4 *Sea E un LCCB y $\mathcal{H}$ subconjunto denso numerable de $C_0^+(E)$. Entonce $\mathcal{H}$ (al considerar las extensiones a E_Δ de cada $f \in \mathcal{H}$) separa puntos de E_Δ, es decir si $x \neq y$ son puntos en E_Δ, existe $f \in \mathcal{H}$ tal que $f(x) \neq f(y)$.*

Demostración
Nos restringimos al caso $x, y \in E$. Primero nótese que claramente existe una función $h \in C_0^+(E)$ tal que $h(x) \neq h(y)$;

$$h(u) = \begin{cases} \frac{d(x,y)-d(x,u)}{d(x,y)} & \text{si } d(x,u) \leq d(x,y) \\ 0 & \text{si } d(x,u) > d(x,y) \end{cases}$$

tiene soporte en $B(x, d(x,y))$ y $h(x) = 1$, $h(y) = 0$. Por densidad, existe $f \in \mathcal{H}$ tal que $\|f - h\|_\infty < \frac{1}{3}$ y entonces en particular $h(x) \geq 2/3 > 1/3 \geq h(y)$. Si ahora $x \in E$ y $y = \Delta$, bastaría hallar $f \in \mathcal{H}$ tal que $f(x) > 0$ (pues por convención $f(\Delta) = 0$), pero esto es claro por un argumento similar. ■

Como E_Δ tiene base numerable y es compacto, es un LCCB. Al considerar una métrica compatible con su topología, es entonces un espacio métrico compacto. Finalmente, llegamos al siguiente lema:

Lema 9.4.5 *Sea M un espacio métrico compacto y $\mathcal{H}$ un subconjunto de $C^+(M)$ tal que $\mathcal{H}$ separa puntos de M. Entonces,*

$$x_n \to x \textit{ si y solo si, para toda } f \in \mathcal{H},\ f(x_n) \to f(x).$$

Demostración
Como la ida se tiene por continuidad, probamos el regreso. Supongamos que $d(x_n, x) \not\to 0$. Entonces para alguna vecindad abierta U_x de x, existe una subsucesión (x_{n_k}) tal que $x_{n_k} \cap U_x = \emptyset$ para toda $k \in \mathbb{N}$. Como M es secuencialmente compacto, la sucesión (x_{n_k}) tiene una subsucesión $(x_{n_{k_j}})$ convergente y converge a un punto $y \neq x$. Sea $f \in \mathcal{H}$ tal que $f(x) \neq f(y)$. Como $(x_{n_{k_j}})$ es subsucesión de la sucesión original (x_n), $f(x_{n_{k_j}}) \to f(x)$. Por otro lado, por continuidad de f, $f(x_{n_{k_j}}) \to f(y)$ lo cual nos lleva a una contradicción. ■

Bajo las hipótesis del lema anterior, se puede refinar el resultado:

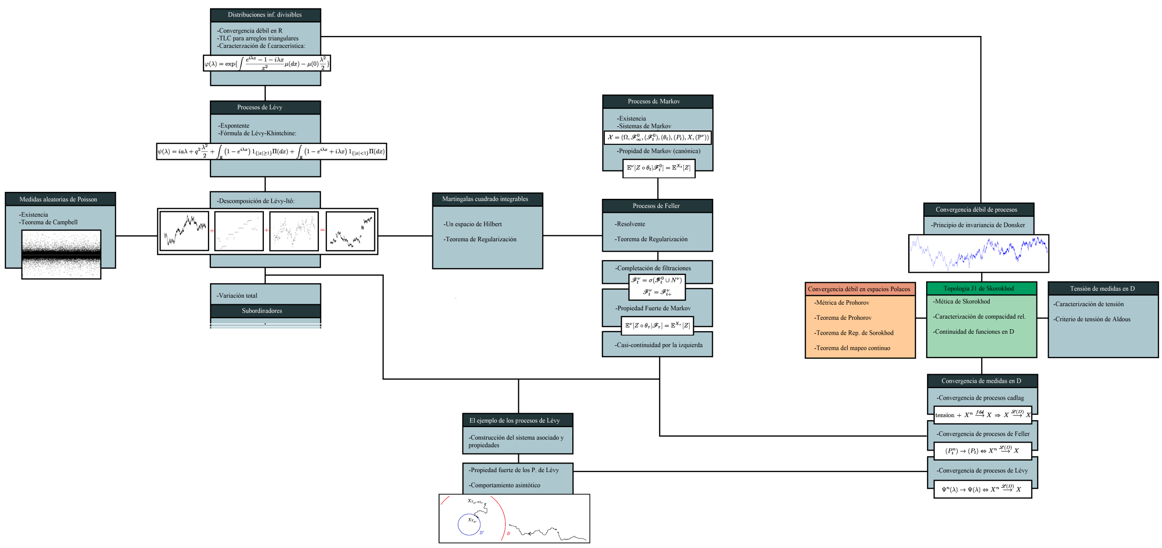
Lema 9.4.6 *Existe $x \in M$ tal que $x_n \to x$ si y solo si $\lim_{n\to\infty} f(x_n)$ existe para toda $f \in \mathcal{H}$.*

Demostración
La idea es directa al aplicar el lema anterior. Para el regreso, obsérvese que la sucesión (x_n) tiene al menos un punto de acumulación x al ser M un espacio métrico compacto. Supongamos que este no es único, esto es, existe $y \in E_\Delta$ otro punto de acumulación de (x_n). Sea $f \in \mathcal{H}$ que separe x y y, es decir, $f(x) \neq f(y)$. Entonces, para esta f no se satisface que el $\lim_{n\to\infty} f(x_n)$ existe lo cual nos lleva a una contradicción. ■

Al combinar el lema 9.4.4 y este último resultado obtenemos:

Corolario 9.4.7 *Si $\mathcal{H}$ es un subconjunto denso numerable de $C_0^+(E)$, entonces sus extensiones separan puntos de E_Δ. Además, existe $x \in E_\Delta$ tal que $x_n \to x$ en E_Δ si y solo si $\lim_{n\to\infty} f(x_n)$ existe para toda $f \in \mathcal{H}$.*

Esto es lo que se usará en la prueba del teorema de regularización de procesos de Feller 6.1.9.


Distribuciones inf. divisibles
-Convergencia débil en R
-TLC para arreglos triangulares
-Caracterzación de f.caracerística:
$\varphi(\lambda) = \exp\{\int \frac{e^{i\lambda x} - 1 - i\lambda x}{x^2}\mu(dx) - \mu(0)\frac{\lambda^2}{2}\}$
Procesos de Lévy
-Expontente
-Fórmula de Lévy-Khintchine:
$\psi(\lambda) = ia\lambda + q^2\frac{\lambda^2}{2} + \int_{\mathbb{R}} (1 - e^{i\lambda x}) 1_{\{|x|\geq 1\}}\Pi(dx) + \int_{\mathbb{R}} (1 - e^{i\lambda x} + i\lambda x) 1_{\{|x|<1\}}\Pi(dx)$
Medidas aleatorias de Poisson
-Existencia
-Teorema de Campbell
-Descomposición de Lévy-Itô:
+
+
=
-Variación total
Subordinadores
Martingalas cuadrado integrables
-Un espacio de Hilbert
-Teorema de Regularización
Procesos de Markov
-Existencia
-Sistemas de Markov
$\mathcal{X} = (\Omega, \mathscr{F}^0_\infty, (\mathscr{F}^0_t), (\theta_t), (P_t), X, (\mathbb{P}^\nu))$
-Propidad de Markov (canónica)
$\mathbb{E}^\nu[Z \circ \theta_t | \mathscr{F}^0_t] = \mathbb{E}^{X_t}[Z]$
Procesos de Feller
-Resolvente
-Teorema de Regularización
-Completación de filtraciones
$\mathscr{F}^\nu_t = \sigma(\mathscr{F}^0_t \cup N^\nu)$
$\mathscr{F}^\nu_t = \mathscr{F}^\nu_{t+}$
-Propiedad Fuerte de Markov
$\mathbb{E}^\nu[Z \circ \theta_\tau | \mathscr{F}_\tau] = \mathbb{E}^{X_\tau}[Z]$
-Casi-continuidad por la izquierda
El ejemplo de los procesos de Lévy
-Construcción del sistema asociado y propiedades
-Propiedad fuerte de los P. de Lévy
-Comportamiento asintótico
$X_{T_{B'}}$
B'
B
Convergencia débil de procesos
-Principio de invariancia de Donsker
Convergencia débil en espacios Polacos
-Métrica de Prohorov
-Teorema de Prohorov
-Teorema de Rep. de Sorokhod
-Teorema del mapeo continuo
Topología J1 de Skorokhod
-Mética de Skorokhod
-Caracterización de compacidad rel.
-Continuidad de funciones en D
Tensión de medidas en D
-Caracterización de tensión
-Criterio de tensión de Aldous
Convergencia de medidas en D
-Convergencia de procesos cadlag
tension $+ X^n \xrightarrow{fdd} X \Rightarrow X \xrightarrow{\mathscr{L}(D)} X$
-Convergencia de procesos de Feller
$(P^n_t) \to (P_t) \Leftrightarrow X^n \xrightarrow{\mathscr{L}(D)} X$
-Convergencia de procesos de Lévy
$\Psi^n(\lambda) \to \Psi(\lambda) \Leftrightarrow X^n \xrightarrow{\mathscr{L}(D)} X$

Bibliography

[1] B. Ash, C. Doléans-Dade *Probability & Measure Theory*, Second Edition 2000.

[2] D. Revuz, M. Yor *Continuous Martingales and Brownian Motion*, First Edition 1990.

[3] F. Spieksma *An introduction to Stochastic Processes in Continuous Time*, Detailed lecture notes from [4] 2016 .

[4] H. van Zanten *An introduction to Stochastic Processes in Continuous Time*, Lecture Notes 2004.

[5] E. Çinlar *Probability and Stochastics*, First Edition 2011.

[6] C. Tudor *Procesos Estocásticos*, Tercera Edición 2002.

[7] P. Billingsley *Probability and Measure*, Third Edition 1995.

[8] P. Billingsley *Converegence of probability measures*, First Edition 1999.

[9] J. Bertoin *Lévy Processes*, First Edition 1996.

[10] J. Bertoin, F.Martinelli, Y. Peres *Lectures on Probability Theory and Statistics. Ecole d'Eté de Probabilités de Saint-Flour XXVII-1997*, First Edition 1999.

[11] J-F. LeGall *Brownian Motion, Martingales, and Stochastic Calculus*, First Edition 2013.

[12] O. Kallenberg *Foundations of Modern Probability*, Second Edition 2001.

[13] A. Kyprianou *Fluctuations of Lévy Processes with Applications*, Second Edition 2014.

[14] P. Morters, Y. Peres *Brownian Motion*, First Edition 2010.

[15] K. Sato *Lévy processes and infinitely divisible distributions*, First Edition 1999.

[16] J. Kingman *Poisson Processes*, First Edition 1993.

[17] M. Blumenthal, K. Getoor *Markov Processes and Potential Theory*, First Edition 1963.

[18] J. Jacod, A. Shiryaev *Limit Theorems for Stochastic Processes*, First Edition 1987.

[19] V. Kolokoltsov *Markov Processes, Semigroups and Generators*, First Edition 2011.

[20] S. ETHIER, T. KURTZ *Markov Processes Characterization and Convergence*, First Edition 2005.

[21] STRASSEN V. *An invariance principle for the law of the iterated logarithm* 1964

[22] IKEDA, WATANABE *Stochastic Diferential Equations and Diffusion Processes* Second Edition 1988

[23] SHI Z. *Support de cours: Théorèmes limites pour des processus stochastiques*

[24] POLLARD D. *Convergence of Stochastic Processes* First Edition 1984

[25] WHITT W. *Stochastic-Processes Limits, An introduction to Stochastic-Process Limits and Their Application to Queues* First Edition 2002

[26] JACOD J. *Support de cours: Processus de Markov, application à la dynamique des populations* 2004