CIMAT

Centro de Investigación
en Matemáticas, A.C.

Procesos de Markov

Notas del curso

Autor:

inm

Profesor:

jcpm

CIMAT

Guanajuato

3 de junio de 2022

Índice general

I. De las clases

1. Preliminares: Clases monótonas

18 de enero

Estamos interesados en diversas versiones del *teorema de clases monótonas*. Estas versiones serán de gran utilidad durante el curso.

1. Definición:
Sean Ω un conjunto y $\mathcal{S}$ una colección de subconjuntos de Ω; entonces

(i) $\mathcal{S}$ es un *π-sistema* (en Ω) si $\mathcal{S}$ es cerrada bajo intersecciones finitas;

(ii) $\mathcal{S}$ es un *δ-sistema* (en Ω) si

(a) $\Omega \in \mathcal{S}$,

(b) si $A, B \in \mathcal{S}$ son tales que $A \subset B$ entonces $B \setminus A \in \mathcal{S}$,
(c) si $\{A_n : n \in \mathbb{N}\} \subset \mathcal{S}$ es una sucesión creciente entonces $\bigcup_n A_n \in \mathcal{S}$.

Claramente una σ-álgebra es tanto un δ-sistema como un π-sistema. Uno puede verificar fácilmente que si $\mathcal{S}$ es un δ-sistema y un π-sistema entonces $\mathcal{S}$ es una σ-álgebra.

Sea $\mathcal{M}$ una colección de subconjuntos de Ω. Vamos a definir a $\delta(\mathcal{M})$ como el δ-sistema más pequeño en Ω que contiene a $\mathcal{M}$. La existencia de $\delta(\mathcal{M})$ es clara, ya que la intersección de un número arbitrario de δ-sistemas es nuevamente δ-sistema y el conjunto potencia es un δ-sistema. A $\delta(\mathcal{M})$ se le conoce como *δ-sistema generado por $\mathcal{M}$*.

1. Teorema:
Sea $\mathcal{S}$ un π-sistema en Ω. Entonces $\delta(\mathcal{S}) = \sigma(\mathcal{S})$.

Demostración: Como $\delta(\mathcal{S}) \subset \sigma(\mathcal{S})$ trivialmente, basta probar que $\delta(\mathcal{S})$ es una σ-álgebra y para ello es necesario ver que $\delta(\mathcal{S})$ es un π-sistema. Pensando en esto definamos

$$D_1 = \left\{B \in \delta(\mathcal{S}) : B \cap A \in \delta(\mathcal{S}) \text{ para toda } A \in \mathcal{S}\right\}.$$

Puesto que $\mathcal{S}$ es un π-sistema es claro que $\mathcal{S} \in D_1$. Por otra parte, es claro que $\Omega \in D_1$ y que para $\{A_n : n \in \mathbb{N}\} \subset D_1$ creciente, $\{A_n \cap A : n \in \mathbb{N}\} \subset \delta(\mathcal{S})$ también lo es y por consiguiente

$$\left(\bigcup_n A_n\right) \cap A = \bigcup_n (A_n \cap A) \in \delta(\mathcal{S});$$

es decir que $\bigcup_n A_n \in D_1$. Por último notemos que dados $B \subset C \in D_1$ con $B \subset C$, se sigue que $B \cap A, C \cap A \in \delta(\mathcal{S})$, $B \cap A \subset C \cap A$ y $(C \cap A) \setminus (B \cap A) = (C \setminus B) \cap A$. Por lo tanto $C \setminus B \in D_1$. Por consiguiente $D_1 = \delta(\mathcal{S})$. Para concluir, consideremos a

$$D_2 = \left\{B \in \delta(\mathcal{S}) : B \cap A \in \delta(\mathcal{S}) \text{ para toda } A \in \delta(\mathcal{S})\right\}.$$

Usando los mismos argumentos de antes se concluye que $D_2 = \delta(\mathcal{S})$ y en consecuencia $\delta(\mathcal{S})$ es un π-sistema. ■

Ahora veremos una versión funcional del teorema 1.

2. Teorema:
Sean Ω un conjunto y $\mathcal{S}$ un π-sistema en Ω. Además consideremos $\mathcal{H}$ un espacio vectorial de funciones $f : \Omega \to \mathbb{R}$ que satisfaga

(i) $\mathbb{1} \in \mathcal{H}$ *y* $\mathbb{1}_A \in \mathcal{H}$ *para toda* $A \in \mathcal{S}$;
(ii) *si* $\{f_n : n \in \mathbb{N}\} \subset \mathcal{H}$ *es una sucesión creciente de funciones no negativas tales que* $f = \sup_n f_n$ *es finita (o acotada) entonces* $f \in \mathcal{H}$.

Bajo estos supuestos $\mathcal{H}$ contiene a todas las funciones $f : \Omega \to \mathbb{R}$ que son $\sigma(\mathcal{S})$-medibles.

Demostración: Consideremos a $\mathcal{D} = \{A : \mathbb{1}_A \in \mathcal{H}\}$. De (i) vemos que $\Omega \in \mathcal{D}$ y $\mathcal{S} \subset \mathcal{D}$. Sean $A_1, A_2 \in \mathcal{D}$ tales que $A_1 \subset A_2$. Puesto que $\mathbb{1}_{A_2 \setminus A_1} = \mathbb{1}_{A_2} - \mathbb{1}_{A_1}$, en vista de que $\mathcal{H}$ es un espacio vectorial, se sigue que $A_2 \setminus A_1 \in \mathcal{D}$. Finalmente es claro que $\mathcal{D}$ es cerrado bajo uniones crecientes en vista de (ii). Efectivamente, si $\{A_n : n \in \mathbb{N}\} \subset \mathcal{D}$ es creciente, entonces

$$\mathbb{1}_{\bigcup_n A_n} = \lim_n \mathbb{1}_{A_n} = \sup_n \mathbb{1}_{A_n}.$$

Por ende $\mathcal{D}$ es un δ-sistema y $\sigma(\mathcal{S}) \subset \mathcal{S}$ debido al teorema 1.

Si $f \in \sigma(\mathcal{S})$ es una función que toma valores reales, entonces $f = f^+ - f^-$, donde f^+ y f^- son funciones no negativas y $\sigma(\mathcal{S})$-medibles. Por otro lado, si $f \in \sigma(\mathcal{S})$ es no negativa, entonces f es el límite creciente de las funciones simples $f_n = (2^{-n}\lfloor 2^n f \rfloor) \wedge 2^n$. Como $f_n \in \mathcal{H}$, obtenemos que $f \in \mathcal{H}$ gracias a (ii) y como consecuencia el resultado. ■

Los teoremas 1 y 2 serán usados de la siguiente manera: Sean Ω un conjunto y $\{(E_i, \mathcal{E}_i) : i \in I\}$ una colección de espacios medibles indexados por un conjunto arbitrario I. Para cada $i \in I$ supondremos que $\mathcal{S}_i$ es un π-sistema tal que $\sigma(\mathcal{S}_i) = \mathcal{E}_i$ y que $f_i : \Omega \to E_i$ es una función. Bajo este escenario obtenemos las proposiciones 1 y 2.

1. Proposición:
Consideremos a $\tilde{\mathcal{S}} := \{\bigcap_{i \in J} f_i^{-1}(A_i) : A_i \in \mathcal{S}_i, J \subset I$ es finito$\}$. Entonces $\tilde{\mathcal{S}}$ es un π-sistema en Ω y $\sigma(\tilde{\mathcal{S}}) = \sigma(f_i, i \in I)$.

2. Proposición:
Sea $\mathcal{H}$ un espacio vectorial de funciones $f : \Omega \to \mathbb{R}$ que satisfacen:

(i) *$1 \in \mathcal{H}$ y $\mathbb{1}_A \in \mathcal{H}$ para toda $A \in \tilde{\mathcal{S}}$.*
(ii) *Si $\{f_n\} \subset \mathcal{H}$ es una sucesión creciente de funciones no negativas tales que $f := \sup f_n$ es acotada, entonces $f \in \mathcal{H}$.*

Bajo estos supuestos $\mathcal{H}$ contiene a las funciones $\sigma(\tilde{\mathcal{S}})$-medibles y acotadas.

La proposición 1 es inmediata y la proposición 2 es consecuencia directa de la proposición 1 y del teorema 2. De ahora en adelante nos vamos a referir a cualquiera de los resultados presentados como el *teorema de clases monótonas*.

2. Construcción de procesos aleatorios

20 de enero

Sean $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, T un espacio temporal (v.g. $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}_+$, $\mathbb{R}$, $\mathbb{R}^d$) y $(E, \mathcal{E})$ un espacio medible.

2. Definición:
Un **proceso aleatorio** con valores en el espacio medible $(E, \mathcal{E})$ es una colección de

variables aleatorias $X = \{X_t : t \in T\}$, donde $X_t : \Omega \to E$ es medible. El espacio E es llamado **espacio de estados**.

Para cada $\omega \in \Omega$, la función $t \mapsto X_t(\omega)$ es una "curva" en E a la cual se le conoce como **trayectoria**. Vamos a pensar en una trayectoria como un punto de que es tomado de forma aleatoria en el espacio de funciones que van de T en E, denotado por E^T.

Por ejemplo, consideremos $T = [0,1]$, $E = \mathbb{R}^d$ y $C([0,1],\mathbb{R}^d)$. Un proceso continuo definido en el intervalo $[0,1]$ y con valores en $\mathbb{R}^d$, X es una variable aleatoria con valores en $C([0,1],\mathbb{R}^d)$. Otro ejemplo es $D([0,1],\mathbb{R}^d)$, el espacio de las funciones càdlàg, funciones continuas por la derecha y con límites por la izquierda.

De ahora en adelante vamos a considerar a T como un conjunto arbitrario y a E como un espacio métrico compacto.[1] Sea $E^T := \{f : T \to E\}$, llamaremos *cilindro de* E^T a todo conjunto de la forma $A^F \times E^{T\setminus F}$, donde $F \subset T$ es finito y A^F es un conjunto medible de $(E^F, \mathcal{E}^{\otimes F})$.

1. Lema:
El conjunto de cilindros $\mathcal{C}$ forma un álgebra.

Demostración: Veamos que $\mathcal{C}$ es estable bajo complementos. Si $A^F \in \mathcal{E}^{\otimes F}$ entonces $(A^F)^c \in \mathcal{E}^{\otimes F}$ y como $(A^F \times E^{T\setminus F})^c = (E^F \setminus A^F) \times E^{T\setminus F}$, y entonces $\mathcal{C}$ es estable bajo complementos.

Ahora consideremos $F, F' \subset T$ finitos y además $A^F \in \mathcal{E}^{\otimes F}$, $B^{F'} \in \mathcal{E}^{\otimes F'}$. Notemos que

$$(A^F \times E^{T\setminus F}) \cap (B^{F'} \times E^{T\setminus F'}) = (C^{F''} \cap D^{F''}) \times E^{T\setminus F''},$$

donde $F'' := F \cup F'$, $C^{F''} := A^F \times E^{F''\setminus F}$ y $D^{F''} := B^{F'} \times E^{F''\setminus F'}$. Aquí pueden usarse las proyecciones que se definen más adelante, *i.e.* $C^{F''} = p_F(A^F)^{-1}$ y $D^{F''} = p_{F'}^{-1}(B^{F'})$, considerando a F'' como el dominio de las proyecciones. Por ende, $\mathcal{C}$ es estable bajo intersecciones finitas. ∎

Vamos a denotar por $\mathcal{E}^{\otimes T}$ a la σ-álgebra generada por los cilíndros y vamos a trabajar con el espacio medible $(E^T, \mathcal{E}^{\otimes T})$. Un ejemplo de un proceso continuo es la *difusión de Feller*, dada por la solución a la ecuación diferencial estocástica

$$Y_t = y + \sigma \int_0^t \sqrt{Y_s}\, \mathrm{d}B_s.$$

En este caso el espacio canónico sería $C(\mathbb{R}_+, \mathbb{R})$. La medida $\mathbb{Q}$ que hace que la difusión de Feller sea el proceso canónico debe satisfacer

$$\mathbb{Q}(e^{-\lambda Y_t}) = e^{-y u_\lambda(t)},$$

donde u_λ satisface la ecuación diferencial parcial, ver la proposición 4.4 de Dawson (2017, pág. 42),

$$\begin{cases} \dfrac{\partial}{\partial t} u_\lambda = -\dfrac{\sigma^2}{2} u_\lambda^2, \\ u_\lambda(0) = \lambda. \end{cases}$$

[1]Este supuesto se puede retirar.

El mismo espacio se puede dotar de otra medida bajo la cual

$$\mathbb{E}[e^{-\lambda W_t}] = e^{\lambda^2 t/2}.$$

Tal medida es la medida de Wiener y da al movimiento browniano como proceso canónico. Moraleja: En un espacio se pueden definir distintos procesos canónicos dependiendo de la medida de probabilidad que se asigne.

Estamos interesados en construir medidas de probabilidad en el espacio medible $(E^T, \mathcal{E}^{\otimes T})$. Supongamos que existe una medida de probabilidad μ_T en dicho espacio y sean $F \subset T$ finito y $\pi_F : E^T \to E^F$ la proyección canónica. Vamos a denotar por μ_F a la medida imagen de μ_T bajo dicha proyección; esto es $\mu_F = \mu_T \circ \pi_F^{-1}$. Ahora consideremos a $F' \subset F$ y a la proyección natural $p_{F'} : E^F \to E^{F'}$ y notemos que $p_{F'} \circ \pi_F = \pi_{F'}$. Esta última relación nos induce una propiedad conocida como *consistencia*: Sea $F' \subset F$, ambos subconjuntos finitos de T, entonces $\mu_{F'} = \mu_F \circ p_{F'}^{-1} = \mu_T \circ \pi_{F'}^{-1}$.

De manera recíproca, nos gustaría encontrar una medida de probabilidad en el espacio $(E^T, \mathcal{E}^{\otimes T})$ de tal manera que se conozcan *a priori* las proyecciones finito-dimensionales. En otras palabras, deseamos encontrar une medida de probabilidad $(E^T, \mathcal{E}^{\otimes T})$ bajo la cual para todo subconjunto finito F de T, $\{X_t : t \in F\}$ tenga una ley dada, donde $X_t : E^T \to E$.

Como ejemplo, supongamos que $\{X_t : t \in T\}$ son variables aleatorias independientes con ley común γ. Para F finito basta considerar $\gamma^{\otimes F}$.

2. Lema:
Si $\mathcal{A}$ es un álgebra y ρ, ρ' son dos medidas de probabilidad en $\sigma(\mathcal{A})$ tales que $\rho(A) = \rho'(A)$ para toda $A \in \mathcal{A}$, entonces $\rho = \rho'$.

3. Teorema (Carathéodory):
Sea $\mu_0 : \mathcal{A} \to [0,1]$, donde $\mathcal{A}$ es un álgebra. Vamos a suponer que para A y B disjuntos, $\mu_0(A \cap B) = \mu_0(A) + \mu_0(B)$. Si para toda sucesión $\{A_n : n \geqslant 1\} \subset \mathcal{A}$ tal que $A_n \searrow \emptyset$, tenemos que $\lim_n \mu_0(A_n) = 0$, entonces existe una medida finita μ sobre $\sigma(\mathcal{A})$ tal que $\mu(A) = \mu_0(A)$ para toda $A \in \mathcal{A}$.

3. Lema:
Sea (E, d) un espacio métrico compacto y μ una medida finita en E. Entonces para todo boreliano B se tiene que

$$\sup\{\mu(K) : K \subset B \text{ es compacto}\} = \mu(B) = \inf\{\mu(U) : U \supset B \text{ es abierto}\}.$$

4. Teorema (Daniell–Kolmogorov):
Sean T un conjunto arbitrario, $(E, \mathcal{E})$ un espacio métrico compacto y para todo subconjunto finito $F \subset T$, μ_F una medida de probabilidad en $(E^F, E^{\otimes F})$. Supongamos que la propiedad de consistencia se satisface, i.e. *si $F' \subset F$ y $p_{F'}$ es la proyección natural de E^F a $E^{F'}$ entonces $\mu_{F'} = \mu_F \circ p_F^{-1}$. Entonces existe una única medida de probabilidad μ en $(E^T, \mathcal{E}^{\otimes T})$ tal que para todo subconjunto finito $F \subset T$, $\mu_F = \mu \circ \pi_F^{-1}$.*

En Tao (2011, págs. 196–198) se da este mismo resultado requiriendo espacios topológicos y medidas de probabilidad regulares por dentro. Alternativamente puede consultarse el teorema 12.1.2 de Dudley (2002, págs. 441–443) para una demostración en la que se consideran espacios universalmente medibles.

25 de enero

Demostración del teorema 4: Primero se demostrará la unicidad. Consideremos dos medidas μ y μ' de probabilidad que satisfacen las condiciones del teorema. Entonces ambas medidas coinciden sobre el conjunto de cilindros $\mathcal{C}$. En efecto, para $A \in \mathcal{E}^{\otimes F}$,

$$\mu(A^F \times E^{T\setminus F}) = \mu_F(A) = \mu'(A^F \times E^{T\setminus F}).$$

Por el lema 2 obtenemos la unicidad.

Pasamos a la existencia, para lo cual usaremos el teorema de Carathéodory con el álgebra $\mathcal{C}$ y una medida $\mu_0 : \mathcal{C} \to [0,1]$ que cumple

$$\mu_0(A^F \times E^{T\setminus F}) = \mu_F(A^F).$$

La propiedad de consistencia garantiza que la medida μ_0 esté bien definida, es decir que no importa la representación que se tome de un conjunto dado. Efectivamente, si $F'' \subset T$ es finito y $F \cap F'' = \emptyset$, entonces

$$A^F \times E^{T\setminus F} = A^F \times E^{F''} \times E^{T\setminus(F\cup F'')},$$

de donde se sigue la afirmación al notar que $\mu_{F\cup F''} = \mu_F \circ p_F^{-1}$, donde $p_F : E^{F\cup F''} \to E^F$ es la proyección natural. Además μ_0 es finitamente aditiva (Ejercicio) . Finalmente hay que verificar que si $\{C_n : n \geqslant 1\} \subset \mathcal{C}$ es tal que $C_n \searrow \emptyset$, entonces $\mu_0(C_n) \to 0$. Consideremos $C_n = A_n \times E^{T\setminus F_n}$ con $F_n \subset T$ finito y $A_n \in \mathcal{E}^{\otimes F_n}$. Sin pérdida de generalidad podemos suponer que $F_n \subset F_{n+1}$. Como E^{F_n} es un espacio métrico compacto, al aplicar el lema 3 deducimos que, dada $\varepsilon > 0$ existe un compacto $K_n \in E^{F_n}$ tal que $K_n \subset A_n$ y $\mu_{F_n}(A_n \setminus K_n) < 2^{-n}\varepsilon$.

Procedemos por contradicción, suponiendo que $\mu_0(C_n) \not\to 0$, de modo que existe $\varepsilon > 0$ tal que $\mu_0(C_n) > 2\varepsilon$. Sea $H_n = K_n \times E^{T\setminus F_n}$, el cual es compacto por el teorema de Tychonov.[2] Como $H_n \subset C_n$, por definición de μ_0 tenemos que $\mu_0(C_n \setminus H_n) \leqslant 2^{-n}\varepsilon$. Entonces

$$\begin{aligned}
\mu_0(H_1 \cap H_2) &= \mu_0(H_1) + \mu_0(H_2) - \mu_0(H_1 \cup H_2) \\
&\geqslant \mu_0(C_1) + \mu_0(C_2) - 2^{-1}\varepsilon - 2^{-2}\varepsilon - \mu_0(C_1) \\
&\geqslant \mu_0(C_2) - 2^{-1}\varepsilon - 2^{-2}\varepsilon.
\end{aligned}$$

Por lo tanto, para toda $n \geqslant 2$ se deduce, recursivamente, que

$$\mu_0\left(\bigcap_{i=1}^n H_i\right) \geqslant \mu_0(C_n) - \sum_{i=1}^n 2^{-i}\varepsilon \geqslant \varepsilon. \tag{$\star$}$$

[2]Ver el teorema 37.3 en Munkres, 2014, pág. 234.

Puesto que $\bigcap_{i=1}^n H_i \subset C_n$, $\bigcap_{n\geqslant 1} H_n = \varnothing$ o bien $\bigcup_n H_n = E^T$. Como E es compacto, existe $N \in \mathbb{N}$ tal que $\bigcap_{i=1}^N H_i = \varnothing$, contradiciendo ($\star$). Por lo tanto $\mu_0(C_n) \to 0$. ∎

Es importante señalar que el teorema de Daniell–Kolmogorov es válido en el caso donde E es un espacio de Lusin, el cual es un *espacio homeomorfo a un subconjunto de Borel en un espacio métrico compacto*. Por ejemplo, $\mathbb{R}^d$ es homeomorfo a un subconjunto de Borel (abierto) de un espacio métrico compacto; en efecto, si consideramos la proyección estereográfica de $\mathbb{S}^d \to \mathbb{R}^{d+1}$ obtenemos el homeomorfismo deseado.

Ejemplo 1: Cadenas de Markov

Sea $T = \mathbb{N}_0$ y E un espacio de estados numerable. Además vamos a considerar

$$\Omega := E^{\mathbb{N}_0} = \{\text{sucesiones con valores en } E\}$$

dotado con la σ-álgebra $\mathcal{F}$, generada por los cilindros. Vamos a definir a la aplicación $X_n : \Omega \to E$ como la proyección canónica (o el mapeo de evaluación) $\omega \mapsto \omega_n$. Sea $\theta_k : \Omega \to \Omega$ el *operador traslación* (*right shift* o *shift*), definido por $\theta_k(\omega) := \{\omega_{k+i} : i \in \mathbb{N}_0\}$, de modo que $X_k = X_0 \circ \theta_k$.

Vamos a definir a $\mathcal{F}_n := \sigma(X_0, X_1, \ldots, X_n)$ y $\mathcal{F} := \bigvee_n \mathcal{F}_n$. En este caso $\{\mathcal{F}_n\}$ es una filtración pues $\mathcal{F}_n \subset \mathcal{F}_{n+1}$.

Llamaremos *matriz de transición* a una matriz $P : E \times E \to [0,1]$ tal que para toda $x \in E$,

$$\sum_{y\in E} P(x,y) = 1.$$

A $P(x,\cdot)$ la podemos ver como una medida de probabilidad sobre E al considerar $P(x,A) = \sum_{y\in A} P(x,y)$ para $A \in \mathcal{E}$.

3. Definición:

Sea $\{\mathbb{P}_x : x \in E\}$ una familia de medidas de probabilidad en el espacio $(\Omega, \mathcal{F})$. Vamos a llamar a esta familia *markoviana* con matriz de transición P si

(i) $\mathbb{P}_x(X_0 = x) = 1$ para toda $x \in E$;
(ii) $\mathbb{P}_{x_0}(X_1 = x_1, \ldots, X_n = x_n) = P(x_0, x_1) \cdots P(x_{n-1}, x_n)$.

Si tomamos $n_1 < n_2 < \cdots < n_k$ enteros positivos se puede verificar que

$$\mathbb{P}_{x_0}(X_{n_1} = x_1, \ldots, X_{n_k} = X_k) = P^{n_1}(x_0, x_1) \cdots P^{n_k - n_{k-1}}(x_{k-1}, x_k), \tag{$\star$}$$

donde $P^2(x,y) = \sum_z P(x,z)P(z,y)$ y $P^{n+1} = P^n P$. Gracias a ($\star$), podemos verificar la propiedad de consistencia, es decir que si $n_1 < n_2 < \cdots < n_k$ y $m_1 < m_2 < \cdots < m_l$ son enteros positivos tales que $\{m_1, \cdots, m_l\} \subset \{n_1, \ldots, n_k\}$, la ley de $(X_{m_1}, \ldots, X_{m_l})$ bajo $\mathbb{P}_{x_0}$ se obtiene bajo la proyección de la ley de $(X_{n_1}, \ldots, X_{n_k})$ bajo la misma probabilidad.

En conclusión del teorema 4 se asume la existencia de una medida de probabilidad $\mathbb{P}_{x_0}$ en Ω para todo $x_0 \in E$. Además, la familia $\{\mathbb{P}_x : x \in E\}$ tiene que ser markoviana con matriz de transición P.

3. Proposición (Propiedad de Markov):
Sea $\mathbb{P}_x$ una medida de probabilidad en Ω. Bajo el evento $\{X_n = y\}$, la cadena trasladada $X \circ \theta_n = \{X_{n+i} : i \in \mathbb{N}_0\}$ es independiente de $\mathcal{F}_n$ y tiene ley $\mathbb{P}_y$.

Demostración: Sea $f : E^k \to \mathbb{R}$ acotada. Notemos que

$$\begin{aligned}
&\mathbb{E}_x[\mathbb{1}_{\{X_1=x_1,\dots,X_n=y\}} f(X_0 \circ \theta_n, \dots, X_{k-1} \circ \theta_n)] \\
&\quad = \sum_{y_0,\dots,y_{k-1}\in E} f(y_0,\dots,y_{k-1}) \mathbb{P}_x(X_1 = x_1, \dots, X_n = y, \dots, X_{n+k-1} = y_{k-1}) \\
&\quad = \sum_{y_0,\dots,y_{k-1}\in E} f(y_0,\dots,y_{k-1}) P(x,x_1)\cdots P(x_{n-1},y)\mathbb{1}_{\{y=y_0\}}\cdots P(y_{k-2},y_{k-1}) \\
&\quad = \mathbb{P}_x(X_1 = x_1, \dots, X_n = y)\mathbb{E}_y[f(X_0,\dots,X_{k-1})].
\end{aligned}$$

Para obtener el mismo cálculo para cualquier conjunto $A \in \mathcal{F}_n$, basta notar que A se puede obtener como una unión disjunta de elementos de la forma $\{X_1 = x_1, \dots, X_n = x_n\}$. Consecuentemente

$$\mathbb{E}_x[\mathbb{1}_{A\cap\{X_n=y\}} f(X_0 \circ \theta_n, \dots, X_{k-1} \circ \theta_n)] = \mathbb{P}_x(A \cap \{X_n = y\})\mathbb{E}_y[f(X_0,\dots,X_{k-1})].$$

Finalmente, por un argumento de clases monótonas (Ejercicio) , el cálculo anterior se puede extender a cualquier funcional $f : E^{\mathbb{N}} \to \mathbb{R}$ medible y acotado, *i.e.*

$$\mathbb{E}_x[\mathbb{1}_{A\cap\{X_n=y\}} f(X \circ \theta_n)] = \mathbb{P}_x(A \cap \{X_n = y\})\mathbb{E}_y[f(X)].$$ ■

27 de enero

Vamos a llamar *tiempo de paro* a una variable aleatoria τ con valores en $\overline{\mathbb{N}}$ tal que para toda n, $\{\tau = n\} \in \mathcal{F}_n$.

(Ejercicio) Probar que $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau = n\} \in \mathcal{F}_n \text{ para toda } n\}$ es una σ-álgebra. Además, ver que $X_\tau \mathbb{1}_{\{\tau<\infty\}}$ es $\mathcal{F}_\tau$-medible.

(Ejercicio) (Propiedad fuerte de Markov) Sea τ un tiempo de paro finito casi seguramente, *i.e.* $\mathbb{P}_x(\tau < \infty) = 1$ para toda $x \in E$. Entonces bajo el evento $\{X_\tau = y\}$, la cadena trasladada $X \circ \theta_\tau$ es independiente de $\mathcal{F}_\tau$ y tiene por ley a $\mathbb{P}_y$.

Ejemplo 2: Movimiento browniano

4. Definición:
Un *movimiento browniano d-dimensional* $B = \{B_t : t \geqslant 0\}$ es un proceso aleatorio en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con valores en $\mathbb{R}^d$ tal que

(i) $\mathbb{P}(B_0 = 0) = 1$,
(ii) tiene incrementos independientes, *i.e.* si $t_0 < t_1 < \cdots < t_n$, entonces

$$B_{t_0}, B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}}$$

son independientes,

(iii) si $t, s \geqslant 0$, entonces

$$\mathbb{P}(B_{t+s} - B_s \in A) = \int_A \frac{1}{(2\pi t)^{d/2}} \exp\left\{-\frac{\|x\|^2}{2t}\right\} \mathrm{d}x$$

para todo $A \in \mathcal{B}(\mathbb{R}^d)$,

(iv) con probabilidad uno, el mapeo $t \mapsto B_t$ es continuo.

Sean $T = \mathbb{R}_+$, $E = \mathbb{R}^d$ y recordemos que $\mathbb{R}^d$ es un espacio de Lusin. Además vamos a considerar el conjunto $\Omega := E^T$, donde la σ-álgebra $\mathcal{F}$ es la generada por los cilindros. Nuestro objetivo es construir una medida de probabilidad $\mathbb{P}$ en el espacio $(\Omega, \mathcal{F})$ de tal manera que el proceso canónico $B = \{B_t : t \geqslant 0\}$ definido por $B_t(\omega) = \omega_t$ es un *casi*[3] movimiento browniano d-dimensional bajo $\mathbb{P}$. Para construir a $\mathbb{P}$ usaremos el teorema de Daniell–Kolmogorov.

Para cada $0 < t_1 < \cdots < t_n$ vamos a definir una medida $\mu_{t_1,\ldots,t_n}$ sobre $(\mathbb{R}^d)^n$ como sigue: para $A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}^d)$ consideremos

$$\mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n) = \int_{A_1} \mathrm{d}x_1 \cdots \int_{A_n} \mathrm{d}x_n \prod_{i=1}^n P_{t_i - t_{i-1}}(x_{i-1}, x),$$

con $x_0 = 0$, $t_0 = 0$ y

$$P_t(x,y) = \frac{1}{(2\pi t)^{d/2}} \exp\left\{-\frac{\|x-y\|^2}{2t}\right\}.$$

Con esta notación es claro que $\mu_{t,t+h}(\mathbb{R}^d \times A) = \mu_{t+h}(A)$. Verifiquemos que esta medida sobre los cilindros satisface la propiedad de consistencia. Para ello vamos a considerar $\{s_1, s_2, \ldots, s_{n-1}\} \subset \{t_1, t_2, \cdots, t_n\}$ con $t_j \notin \{s_1, s_2, \ldots, s_{n-1}\}$. Entonces

$$\mu_{s_1,\ldots,s_{n-1}}(A_1 \times \cdots \times A_{n-1}) = \mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_{j-1} \times \mathbb{R}^d \times A_j \times \cdots \times A_{n-1}).$$

Este procedimiento nos permite construir procesos con incrementos independientes y estacionarios con un kernel $p_t(x,y)$ dado. Por lo tanto del teorema 4 se asume la existencia de una medida de probabilidad en $(\Omega, \mathcal{F})$ tal que

$$\mathbb{P}(\{\omega : B_t(\omega) \in A_i, i = 1, \ldots, n\}) = \mu_{t_1,\ldots,t_n}(A_1 \times \cdots \times A_n).$$

4. Proposición:

Existe una medida $\mathbb{P}$ en $(\Omega, \mathcal{F})$ bajo la cual el proceso canónico dado por $B_t(\omega) = \omega_t$ para $\omega \in \Omega$ y $t \in \mathbb{R}_+$ tiene incrementos independientes y estacionarios. En adición, un incremento $B_t - B_s$, donde $t > s$, se distribuye como una variable aleatoria normal de media 0 y varianza $t - s$

Nuestra construcción estaría completa si probamos que la $\mathbb{P}(C(\mathbb{R}_+, \mathbb{R}^d)) = 1$.

Se puede verificar que el conjunto $C(\mathbb{R}_+, \mathbb{R}^d)$ no es $\mathcal{F}$-medible, de modo que la cantidad $\mathbb{P}(C(\mathbb{R}_+, \mathbb{R}^d))$ no está bien definida. Este problema pone en manifiesto que la

[3] Por la continuidad :'(

σ-álgebra $\mathcal{F}$ no es lo suficientemente grande para nuestros propósitos. En sí ningún conjunto en $\mathcal{F}$ tendrá restricciones en una cantidad no numerable de coordenadas.[4]

(Ejercicio) Probar que el único conjunto $\mathcal{F}$-medible contenido en $C(\mathbb{R}_+, \mathbb{R}^d)$ es el vacío y deducir que $C(\mathbb{R}_+, \mathbb{R}^d)$ no es $\mathcal{F}$-medible.

5. Definición:
Sean $X = \{X_t : t \geqslant 0\}$ y $\widetilde{X} = \{\widetilde{X}_t : t \geqslant 0\}$ dos procesos aleatorios. Decimos que $\widetilde{X}$ es una *modificación* de X si para toda $t \geqslant 0$,

$$\mathbb{P}(X_t = \widetilde{X}_t) = 1.$$

Es importante señalar que X y $\widetilde{X}$ *pueden* tener trayectorias completamente distintas.

6. Definición:
Sean X y $\widetilde{X}$ como en la definición 5. Decimos que X y $\widetilde{X}$ son *indistinguibles* $X_t = \widetilde{X}_t$ para toda $t \geqslant 0$, con probabilidad uno.

Si X y $\widetilde{X}$ son indistinguibles, entonces $\widetilde{X}$ es una modificación de X y además los dos procesos tienen las mismas trayectorias con probabilidad casi seguramente.

(Ejercicio) Sean X y $\widetilde{X}$ dos procesos continuos casi seguramente. Probar que $\widetilde{X}$ es una modificación de X, entonces X y $\widetilde{X}$ son indistinguibles.

5. Teorema (Criterio de continuidad de Kolmogorov)**:**
Sea $X = \{X_t : t \in I\}$ un proceso aleatorio indexado por un intervalo $I \subset \mathbb{R}$ con valores en (E, d) un espacio métrico completo. Supongamos que existen $p, \varepsilon, c > 0$ tal que

$$\mathbb{E}[d(X_t, X_s)^p] \leqslant c|t-s|^{1+\varepsilon}, \qquad s, t \in I.$$

Entonces existe una modificación $\widetilde{X}$ de X cuyas trayectorias son localmente Hölder de índice $\alpha \in (0, \varepsilon/p)$; esto es, para toda $T \geqslant 0$,

$$d(\widetilde{X}_t(\omega), \widetilde{X}_s(\omega)) \leqslant K(\omega, T)|t-s|^{\alpha}, \qquad s, t \in I \cap [-T, T].$$

En particular, X admite una modificación continua que es única salvo indistinguibilidad.

Paul Lévy consideró variables aleatorias gaussianas estándar y usó interpolación para construir, en el límite, un proceso aleatorio continuo que satisface las demás propiedades dadas en la definición 4. Esta construcción se detalla en la sección 3.2 de Schilling y Partzsch (2012, págs. 28–33).

1 de febrero

Demostración del teorema 5: La unicidad es evidente de que si dos modificaciones son casi seguramente continuas entonces son indistinguibles. Probaremos la existencia,

[4] Es imposible determinar a una función en Ω al especificar sus valores en una cantidad numerable de puntos.

suponiendo sin pérdida de generalidad que $I = [0,1]$. Por la desigualdad de Chebyshev vemos que para $a > 0$ y cualesquiera $s,t \in [0,1]$ se tiene que

$$\mathbb{P}(d(X_t,X_s) \geqslant a) \leqslant a^{-p}\mathbb{E}[d(X_t,X_s)^p] \leqslant ca^{-p}|t-s|^{1+\varepsilon}.$$

Al tomar $s = 2^{-n}(i-1)$, $t = 2^{-n}i$ y $a = 2^{-n\alpha}$ se sigue que

$$\mathbb{P}(d(X_{2^{-n}(i-1)},X_{2^{-n}i}) \geqslant 2^{-n\alpha}) \leqslant c2^{-n(1+\varepsilon)+n\alpha p},$$

y en consecuencia

$$\mathbb{P}\left(\max_{i\in[2^n]} d(X_{2^{-n}(i-1)},X_{2^{-n}i}) \geqslant 2^{-n\alpha}\right) \leqslant c2^{-n(\alpha p-\varepsilon)}.$$

De este modo, si $\alpha p - \varepsilon < 0$, como $\sum_n c2^{n(\alpha p-\varepsilon)} < \infty$, por Borel–Cantelli se deduce la existencia de $\widetilde{\Omega} \in \mathcal{F}$ con $\mathbb{P}(\widetilde{\Omega}) = 1$ y tal que para toda $\omega \in \widetilde{\Omega}$ existe $n_0 := n_0(\omega)$ para el cual

$$d(X_{2^{-n}(i-1)},X_{2^{-n}i}) < 2^{-n\alpha}$$

para toda $i \in [2n]$ siempre que $n \geqslant n_0$. Mediante un argumento usual se deduce la existencia de una constante $K(\omega) > 0$ tal que

$$\sup_{n\geqslant 1}\sup_{i\in[2n]} \frac{d(X_{2^{-n}(i-1)},X_{2^{-n}i})}{2^{-n\alpha}} \leqslant K(\omega).$$

Sea ahora D el conjunto de diádicos en $[0,1]$; esto es el conjunto de números $t \in [0,1]$ para los cuales existen N y $\zeta_1,\dots,\zeta_N \in \{0,1\}$ con la propiedad de que

$$t = \sum_{k=1}^{N} 2^{-k}\zeta_k.$$

Consideremos $s,t \in D$ con $s < t$ y sea $p > 1$ el mayor entero tal que $t - s < 2^{-p}$. En adición tomemos $\kappa := \lfloor 2^p \rfloor$ y notemos que $\kappa \leqslant \lfloor 2^p t\rfloor \leqslant \lfloor 2^p s + 1\rfloor = \kappa + 1$. Por último consideremos que

$$s = \frac{\kappa}{2^p} + \frac{\zeta^s_{p+1}}{2^{p+1}} + \cdots + \frac{\zeta^s_{p+l}}{2^{p+l}}$$

y

$$t = \frac{\kappa}{2^p} + \frac{\zeta^t_{p}}{2^{p}} + \frac{\zeta^t_{p+1}}{2^{p+1}} + \cdots + \frac{\zeta^t_{p+m}}{2^{p+m}}.$$

Con base en lo anterior definamos, para $i = 1,\dots,l$,

$$s_i = \frac{\kappa}{2^p} + \frac{\zeta^s_{p+1}}{2^{p+1}} + \cdots + \frac{\zeta^s_{p+i}}{2^{p+i}}$$

y $s_0 = 2^{-p}\kappa$. De manera similar definimos a t_j pero con $t_0 = 2^{-p}\kappa + 2^{-p}\zeta_p^t$. Con todas estas consideraciones, al usar la desigualdad del triángulo y argumentos elementales, deducimos que

$$\begin{aligned}
d(X_t, X_s) = d(X_{t_m}, X_{s_j}) &\leqslant d(X_{t_0}, X_{s_0}) + \sum_{i=1}^{l} d(X_{s_i}, X_{s_{i-1}}) + \sum_{j=1}^{m} d(X_{t_j}, X_{t_{j-1}}) \\
&\leqslant K(\omega)2^{-p\alpha} + \sum_{i=1}^{l} K(\omega)2^{-(p+i)\alpha} + \sum_{j=1}^{m} K(\omega)2^{-(p+j)\alpha} \\
&\leqslant 2K(\omega)2^{-p\alpha}\sum_{i\geqslant 0} 2^{-i\alpha} = \frac{2K(\omega)2^{-p\alpha}}{1-2^{-\alpha}} \\
&\leqslant \frac{2^{\alpha+1}K(\omega)}{1-2^{-\alpha}}(t-s)^{\alpha},
\end{aligned}$$

con la última desigualdad dándose debido a que $t - s \geqslant 2^{-(p+1)}$.

Hemos demostrado que para cada $\omega \in \widetilde{\Omega}$, el mapeo $t \in D \mapsto X_t(\omega)$ es Hölder de índice α, y por ende uniformemente continuo en D. Así, para cada $t \in I$ podemos definir

$$\widetilde{X}_t(\omega) := \begin{cases} \lim\limits_{s\to t, s\in D} X_s(\omega) & \text{si } \omega \in \widetilde{\Omega}, \\ x_0 \in E & \text{si } \omega \notin \widetilde{\Omega}, \end{cases}$$

pues el límite está bien definido gracias a la continuidad uniforme y el criterio de Cauchy. Por definición $\widetilde{X}$ es Hölder de índice α. Resta verificar que es una modificación de X, para lo cual basta notar que $X_t = \lim_{s\to t} X_s$ en probabilidad para cada $t \in I$. ■

1. Corolario:
Sea $B = \{B_t : t \geqslant 0\}$ un movimiento browniano unidimensional. El proceso B admite una modificación continua cuyas trayectorias son localmente Hölder de índice α, con $\alpha = 1/2 - \varepsilon$ y $\varepsilon \in (0, 1/2)$. En particular B admite una modificación continua.

Demostración: Sean $\varepsilon \in (0, 1/2)$ y $p > 0$. Si Z es una variable aleatoria con distribución normal estándar y $c(p) := \mathbb{E}[|Z|^p]$, es claro que

$$\mathbb{E}[|B_t - B_s|^p] = c(p)|t-s|^{1+(p/2-1)}.$$

Al tomar $p > 2$ de tal forma que $1/2 - \varepsilon < (p/2-1)/p$ concluimos gracias al Teorema 5. ■

3. Teoría de renovación

3 de febrero

Sea $\{\xi_i\}_i$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, no negativas. Además definimos el n-ésimo tiempo de renovación mediante

$$T(n) = \sum_{i=1}^{n} \xi_i, \qquad T(0) = 0.$$

Ejemplo: Si X es una cadena de Markov con valores en un espacio numerable. Vamos a suponer que $X_0 = x_0$ y vamos a denotar por $T(n) := \inf\{m > T(n-1) : X_m = x_0\}$ y $T(0) = 0$. Para ver que $T(n)$ es un n-ésimo tiempo de renovación usamos la propiedad de Markov.

3.1. Teorema de renovación

7. Definición:
Un proceso de renovación $N = \{N_t : t \geqslant 0\}$ es un proceso aleatorio tal que

$$N_t = \max\{n \in \mathbb{N} : T(n) \leqslant t\}.$$

Vamos a definir a la función de renovación como $U_t := \mathbb{E}[N_t] + 1$.

Vamos a considerar a F como la función de distribución de ξ_1. Para $k \in \mathbb{N}$, F_k es la función de distribución del k-ésimo tiempo de renovación, T_k, y su ley satisface $\mathrm{d}F_k = (\mathrm{d}F)^{\otimes k}$. Por último, $F_0 := 1$.

4. Lema:
Tenemos que $\mathbb{P}(N_t = k) = F_k(t) - F_{k+1}(t)$ *y* $U_t = \sum_{n\geqslant 0} F_n(t)$.

Demostración: De la igualdad $\{N_t = k\} = \{T(k) \leqslant t < T(k+1)\}$ obtenemos la primera afirmación. Por otra parte,

$$\mathbb{E}[N_t] = \sum_{k=1}^{\infty} k\mathbb{P}(N_t = k) = \sum_{k=1}^{\infty} k(F_k(t) - F_{k+1}(t)) = \sum_{k=1}^{\infty} F_k(t). \qquad \blacksquare$$

2. Corolario (Ecuación de renovación)**:**
La función de renovación U satisface la ecuación

$$U_t = 1 + \int_{[0,t]} U_{t-s}\, \mathrm{d}F(s).$$

Demostración: Notemos que

$$\sum_{k=0}^{\infty} F_k(t) = F_0(t) + \sum_{k=1}^{\infty} F_k(t) = F_0(t) + \sum_{k=1}^{\infty} F_n * F(t). \qquad \blacksquare$$

(Ejercicio) Sea H una función positiva y acotada. Considere la siguiente ecuación:

$$\mu(t) = H(t) + \int_{[0,t]} \mu(t-s)\,\mathrm{d}F(s). \tag{$\star$}$$

Probar que existe una única solución a la ecuación ($\star$) dada por

$$\mu(t) = \int_{[0,t]} H(t-d)\,\mathrm{d}U_s.$$

8. Definición:
Decimos que F (o ξ o el proceso de renovación) es *aritmética* si existe $r > 0$ tal que $\mathbb{P}(\xi_1 \in r\mathbb{N}) = 1$. El máximo entero positivo r que cumple con esta propiedad es llamado el *paso* de F. En caso contrario diremos que F no es aritmética.

Ejemplo: Consideremos a la caminata aleatoria simétrica con valores en $\mathbb{Z}$ y que comienza en 0. Aquí $T(1), T(2), \ldots$ son los tiempos de retorno al 0 de la caminata.

6. Teorema (Blackwell)**:**
Supongamos que $\mathbb{E}[\xi_1] = m < \infty$*.*

(i) *Si F no es aritmética, entonces para toda $h > 0$,*

$$\lim_{x\to\infty} U_{x+h} - U_x = \frac{h}{m}.$$

(ii) *Si F es aritmética de paso $r > 0$, entonces*

$$\lim_{n\to\infty} U_{n+r} - U_n = \frac{r}{m}.$$

Solo probaremos el caso aritmético, ya que el caso no aritmético resulta muy técnico. Definamos al *proceso exceso de vida*, $R = \{R_n : n \geqslant 0\}$, donde $R_n = T(N_n + 1) - n$.

5. Lema:
El proceso R es una cadena de Markov con matriz de transición

$$Q(a,b) = \begin{cases} 1 & si\ a \geqslant 2, b = a-1, \\ 0 & si\ a \geqslant 2, b \neq a-1, \\ \mathrm{d}F(b) & si\ a = 1. \end{cases}$$

Demostración: Sea $\mathcal{G}_n = \sigma(T(0), \ldots, T(n))$. La familia $\mathcal{G}_n$ es una filtración y la variable aleatoria $N_n + 1$ es un tiempo de paro. En efecto, basta notar que

$$\{N_n = k-1\} = \{T(k-1) \leqslant n < T(k)\} \in \mathcal{G}_k.$$

Para toda $n \in \mathbb{N}$, vamos a denotar por $\mathcal{F}_n$ a la σ-álgebra $\mathcal{G}_{N_n+1}$. Ahora probemos que R es una cadena de Markov con respecto a $\{\mathcal{F}_n : n \geqslant 1\}$. Para ello basta calcular la matriz de transición via

$$\mathbb{E}[\mathbb{1}_{\{\Delta, N_n+1=k, R_n=a\}} f(R_{n+1})] \tag{$\star\star$}$$

con $f : \mathbb{N} \to \mathbb{R}$ acotada y $\Delta \in \mathcal{F}_n$. Si $a > 1$ entonces los objetos dentro de ($\star\star$) son iguales a $R_{n+1} = T(k) - (n+1)$ y $R_n = T(k) - a$, de modo que $R_{n+1} = a - 1$. En consecuencia

$$\mathbb{E}[\mathbb{1}_{\{\Delta, N_n+1=k, R_n=a\}} f(R_{n+1})] = f(b)\mathbb{1}_{\{b=a-1\}}\mathbb{P}(\Delta, N_n + 1 = k, R_n = 1).$$

Finalmente, si $a = 1$, $R_{n+1} = \xi_{k+1} \perp\!\!\!\perp \{\xi_1, \ldots, \xi_k\}$ y por lo tanto

$$\mathbb{E}[\mathbb{1}_{\{\Delta, N_n+1=k, R_n=a\}} f(R_{n+1})] = \mathbb{E}[f(\xi_{k+1})]\mathbb{P}(\Delta, N_n + 1 = k, R_n = 1). \qquad \blacksquare$$

Demostración del Teorema 6, parte (ii): Vamos a suponer que F es aritmética de paso 1. Vimos que R es una cadena de Markov que si parte de 1, el primer tiempo de retorno al nivel 1 tiene la misma ley que ξ_1 y por tanto tiene media finita. La cadena además es positiva-recurrente y aperiódica, gracias a la hipótesis de aritmeticidad de ξ_1. En conclusión, R_n tiene una distribución estacionaria γ, la cual es también límite. Entonces $\gamma(k) = \sum_{i=1}^{\infty} \gamma(i)Q(i,k)$. A partir de la fórmula de Q, obtenemos la recursión

$$\gamma(k) = \gamma(k+1) + \gamma(1)\mathbb{P}(\xi_1 = k).$$

Por lo tanto, $\gamma(k) - \gamma(k+1) = \gamma(1)\mathbb{P}(\xi_1 = k)$. Al sumar, obtenemos que

$$\gamma(k) = \gamma(1)\mathbb{P}(\xi_1 \geqslant k) = \gamma(1)\overline{F}(k-1).$$

Sumando sobre toda k obtenemos que

$$1 = \gamma(1) \sum_{k=1}^{\infty} \overline{F}(k-1) = \gamma(1)m,$$

de modo que $\gamma(1) = m^{-1}$.

Finalmente, vemos que

$$U_{n+1} - U_n = \mathbb{E}\left[\sum_{k\geqslant 1} \mathbb{1}_{\{T(k)=n+1\}}\right] = \mathbb{P}(R_n = 1). \qquad N_n = \sum_{k\geqslant 1} \mathbb{1}_{\{T(k)\leqslant n\}}.$$

Por la convergencia deducida previamente concluimos la prueba. $\blacksquare$

8 de febrero

3. Corolario (Teorema de renovación clave):

Supongamos que F no es aritmética y sea $g : \mathbb{R}_+ \to \mathbb{R}_+$ una función continua con soporte compacto. Entonces si $m < \infty$,

$$\lim_{x\to\infty} \int_{[0,x]} g(x-s)\,\mathrm{d}U_s = \frac{1}{m}\int_0^{\infty} g(t)\,\mathrm{d}t.$$

Demostración: Primero observemos que

$$\int_{[0,x]} g(x-s)\,\mathrm{d}U_s = \int_{[0,x]} g(s)\,\mathrm{d}(U_x - U_{x-s}).$$

Del teorema de renovación se tiene que sobre todo conjunto compacto de la forma $[0,a]$, $U_x - U_{x-s} \to s/m$ para $s \in [0,a]$. De la convergencia vaga vemos que

$$\int_{[0,x]} g(s)\,\mathrm{d}(U_x - U_{x-s}) \to \frac{1}{m}\int_0^\infty g(s)\,\mathrm{d}s. \qquad \blacksquare$$

7. Teorema (Teorema de renovación elemental)**:**
Supongamos que $m < \infty$. Entonces

$$\frac{U_t}{t} \to \frac{1}{m} \qquad \textit{conforme } t \to \infty.$$

Demostración: Para toda $t > 0$ se cumple que $T(N_t) \leqslant t < T(N_t + 1)$. Por el lema de Wald[1] tenemos que

$$t \leqslant \mathbb{E}[T(N_T + 1)] = \mathbb{E}\left[\sum_{i=1}^{N_t+1} \xi_i\right] = \mathbb{E}[\xi_1]\mathbb{E}[N_t + 1] = mU_t.$$

En consecuencia $U_t/t \geqslant 1/m$. Por otro lado, como $T(N_t) \leqslant t$, tenemos que

$$t \geqslant \mathbb{E}[T(N_t)] = \mathbb{E}[T(N_t + 1) - \xi_{N_t+1}] = mU_t - \mathbb{E}[\xi_{N_t+1}].$$

Definamos $\xi_j^c = \xi_i \wedge c$, donde $c > 0$. Consideremos al proceso de renovación N^c asociado a $\{\xi_j^c\}_j$. En este caso $m^c = \mathbb{E}[\xi_1^c]$ y por un razonamiento similar al ya expuesto, $t \geqslant m^c\mathbb{E}[N_t^c + 1] - c$.

Como $\xi_j^c \leqslant \xi_j$ para toda $j \geqslant 1$, tendremos que $N_t^c \geqslant N_t$ para toda $t \geqslant 0$, de modo que $\mathbb{E}[N_t^c + 1] \geqslant U_t$. Por lo tanto

$$\frac{1}{m^c} + \frac{c}{tm^c} \geqslant \frac{U_t}{t}.$$

Al aplicar límite superior con respecto a t obtenemos que

$$\limsup_{t\to\infty} \frac{U_t}{t} \leqslant \frac{1}{m^c},$$

y cuando hacemos tender $c \to \infty$ obtenemos $m^c \to m$ por convergencia monótona, y por consiguiente,

$$\limsup_{t\to\infty} \frac{U_t}{t} \leqslant \frac{1}{m}. \qquad \blacksquare$$

[1] Ver la sección 1.8.1 en Resnick, 1992, págs. 47–48.

(Ejercicio) Supongamos $m<\infty$, pruebe que

$$\frac{1}{t}N_t \to \frac{1}{m},$$

casi seguramente, conforme $t\to\infty$.

Sea $f:\mathbb{R}_+\to\mathbb{R}_+$ arbitraria, vamos a definir

$$I^\delta := \sum_{k=0}^{\infty}\delta \sup\Big\{f(x): x\in[k\delta,(k+1)\delta)\Big\} \quad \text{y} \quad I_\delta := \sum_{k=0}^{\infty}\delta \operatorname{ínf}\Big\{f(x): x\in[k\delta,(k+1)\delta)\Big\}.$$

Estas sumas se conocen como *sumas superior e inferior de Riemann que aproximan a la integral de f en $\mathbb{R}_+$*.

9. Definición:
Decimos que una función $f:\mathbb{R}_+\to\mathbb{R}_+$ es *directamente Riemann integrable* si I^δ e I_δ tienen el mismo límite I conforme $\delta\to 0$.

(Ejercicio) Sea $f:\mathbb{R}_+\to\mathbb{R}_+$ decreciente tal que $f(0)<\infty$ y $\int_0^\infty f(x)\,\mathrm{d}x<\infty$. Probar que f es directamente Riemann integrable.

(Ejercicio) Sea $f:\mathbb{R}_+\to\mathbb{R}_+$ continua. Probar que f es directamente Riemann integrable si y solamente si $I^\delta<\infty$ para alguna $\delta>0$.

(Ejercicio) Supongamos que F no es aritmética y sea $g:\mathbb{R}_+\to\mathbb{R}_+$ dRI. Entonces si $m<\infty$,

$$\lim_{x\to\infty}\int_{[0,x]} g(x-s)\,\mathrm{d}U_s = \frac{1}{m}\int_0^\infty g(t)\,\mathrm{d}t.$$

3.2. Edad y exceso de vida

El *proceso de edad* se define como

$$A_t := t - T(N_t)$$

y el *proceso exceso de vida* como

$$R_t := T(N_t+1) - t.$$

Recordemos que $T(k)$ es el k-ésimo tiempo de renovación, $N_t = \max\{k : T(k)\leqslant t\}$ y $\mathcal{G}_k=\sigma(\xi_1,\dots,\xi_k)$.

10. Definición:
Sea $(\Omega,\mathcal{F},\mathbb{P})$ un espacio de probabilidad. Una filtración es una familia $\{\mathcal{F}_t\}_{t\geqslant 0}$ de sub-σ-álgebras de $\mathcal{F}$ tales que $\mathcal{F}_s\subset\mathcal{F}_t$ siempre y cuando $s\leqslant t$. Al sistema

$$(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\geqslant 0},\mathbb{P})$$

se le llama *espacio de probabilidad filtrado*.

Sea $\{X_t : t \geqslant 0\}$ un proceso aleatorio definido en $(\Omega, \mathcal{F}, \mathbb{P})$. La *filtración canónica* asociada a dicho proceso está dada por $\mathcal{F}_t = \sigma(X_s : s \leqslant t)$. Vamos a denotar por

$$\mathcal{F}_{t-} := \bigvee_{s<t} \mathcal{F}_s \qquad \text{y} \qquad \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$$

para todo $t \geqslant 0$. Es claro que $\mathcal{F}_{t-} \subset \mathcal{F}_t \subset \mathcal{F}_{t+}$.

11. Definición:
Sean $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ un espacio de proba filtrado y $\{X_t : t \geqslant 0\}$ un proceso aleatorio definido en $(\Omega, \mathcal{F}, \mathbb{P})$. Se dice que el proceso es *adaptado* a la filtración $\{\mathcal{F}_t\}$ si para toda $t \geqslant 0$ se tiene que la variable aleatoria X_t es $\mathcal{F}_t$-medible.

12. Definición:
Una filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$ se dice *continua por la derecha* si $\mathcal{F}_t = \mathcal{F}_{t+}$ para toda $t \geqslant 0$.

De esta última definición es fácil ver que $\{\mathcal{F}_{t+}\}_{t\geqslant 0}$ es continua por la derecha.

13. Definición:
Sea $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ un espacio de probabilidad filtrado. Una función $\tau : \Omega \to [0, \infty]$ se llama *tiempo de paro* (con respecto a $\{\mathcal{F}_t\}$) si

(i) τ es medible con respecto a $\mathcal{F}_\infty := \bigvee_{t\geqslant 0} \mathcal{F}_t$;
(ii) el conjunto $\{\tau \leqslant t\} \in \mathcal{F}_t$ para toda $t \geqslant 0$;

(Ejercicio) Sea $\{\mathcal{F}_t\}_{t\geqslant 0}$ una filtración continua por la derecha. Probar que τ es un tiempo de paro si, sólo si, el conjunto $\{\tau < t\} \in \mathcal{F}_t$ para toda $t \geqslant 0$.
(Ejercicio) Sea τ un tiempo de paro con respecto a $\{\mathcal{F}_t\}_{t\geqslant 0}$ y $X_t := \mathbb{1}_{[0,\tau]}(t)$. Probemos que el proceso $\{X_t : t \geqslant 0\}$ es adaptado a dicha filtración.

14. Definición:
Definamos al *tiempo de entrada en* A o la primera vez que el proceso $\{X_t : t \geqslant 0\}$ entra en el conjunto A como $T_A = \inf\{t \geqslant 0 : X_t \in A\}$, con $\inf \emptyset = \infty$.

(Ejercicio) Sea $\{X_t : t \geqslant 0\}$ un proceso estocástico adaptado a $\{\mathcal{F}_t\}_{t\geqslant 0}$ y con espacio de estados $(E, \mathcal{E})$, donde E es un espacio métrico. Además consideremos $A \in \mathcal{E}$. Probar que

(i) si X y $\{\mathcal{F}_t\}_{t\geqslant 0}$ con continuos por la derecha y A es abierto entonces T_A es un tiempo de paro con respecto a $\{\mathcal{F}_t\}_{t\geqslant 0}$;
(ii) si X es continuo y A es cerrado entonces T_A es un tiempo de paro con respecto a $\{\mathcal{F}_t\}_{t\geqslant 0}$.

10 de febrero

15. Definición:
Sea τ un tiempo de paro con respecto a $\{\mathcal{F}_t\}$. Denotamos por $\mathcal{F}_\tau$ a la colección de

eventos $A \in \mathcal{F}$ tales que

$$A \cap \{\tau \leqslant t\} \in \mathcal{F}_t \qquad \text{para toda } t \geqslant 0.$$

Llamaremos a $\mathcal{F}_\tau$ la *σ-álgebra parada en τ*.

5. Proposición:
Sea τ un tiempo de paro con respecto a $\{\mathcal{F}_t\}$. Entonces $\mathcal{F}_\tau$ es una σ-álgebra.

6. Proposición:
Sean $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ un espacio de probabilidad filtrado y $\{X_t : t \geqslant 0\}$ un proceso aleatorio càdlàg adaptado y con valores en un espacio métrico E. Además consideremos a τ un tiempo de paro con respecto a $\{\mathcal{F}_t\}$. Entonces $X_\tau \mathbb{1}_{\{\tau<\infty\}}$ es $\mathcal{F}_\tau$-medible. Además el proceso parado $X^\tau := \{X_\tau \wedge t : t \geqslant 0\}$ es adaptado.

Demostración: Primero notemos que Z es $\mathcal{F}_\tau$-medible si y sólo si $Z\mathbb{1}\{\tau \leqslant t\}$ es $\mathcal{F}_t$-medible para toda $t \geqslant 0$, ya que Z es aproximable por funciones de la forma $\sum_i a_i \mathbb{1}_{A_i}$ con $A_i \in \mathcal{F}_\tau$. Como τ es tiempo de paro, la variable aleatoria

$$\tau_n := 2^{-n} \lceil 2^n \tau \rceil$$

es también un tiempo y satisface que $\tau_n \geqslant \tau$. En efecto, que sea tiempo de paro se deduce de

$$\{\tau_n \leqslant t\} = \left\{\tau \leqslant 2^{-n} \lfloor 2^n \rfloor\right\} \in \mathcal{F}_t. \qquad (\lfloor x \rfloor \leqslant y \iff x \leqslant \lceil y \rceil)$$

Por construcción es claro que τ_n toma valores en

$$D_n = \{k2^{-n} : k \in \mathbb{Z}_+\} \cup \{\infty\}$$

y que la sucesión $\{\tau_n\}$ decrece a τ.

Consideremos

$$X_\tau \mathbb{1}_{\{\tau<\infty\}} \mathbb{1}_{\{\tau \leqslant t\}} = X_\tau \mathbb{1}_{\{\tau<t\}} + X_t \mathbb{1}_{\{\tau=t\}} = X_t \mathbb{1}_{\{\tau=t\}} + \lim_{n\to\infty} X_{\tau_n \wedge t} \mathbb{1}_{\{\tau<t\}}.$$

Notemos que $X_t \mathbb{1}_{\{\tau=t\}}$ y $X_{\tau_n \wedge t} \mathbb{1}_{\{\tau<t\}}$ son $\mathcal{F}_t$-medibles. Para la segunda parte de la suma veamos que

$$X_{\tau_n \wedge t} = \sum_{d \in D_n \cap [0,t]} X_d \mathbb{1}_{\{\tau_n = d\}} + X_t \mathbb{1}_{\{t<\tau_n\}}$$

es $\mathcal{F}_t$-medible. Esto nos da la $\mathcal{F}_t$-medibilidad de $X_\tau \mathbb{1}_{\{\tau<\infty\}} \mathbb{1}_{\{\tau \leqslant t\}}$.

Con esto demostrado, basta notar que $\tau \wedge t$ es tiempo de paro y $X_{\tau \wedge t} \in \mathcal{F}_{\tau\wedge t} \subset \mathcal{F}_t$. ■

Recordemos que $N_t + 1$ es tiempo de paro con respecto a $\{\mathcal{G}_k\}$. Vamos a definir las filtraciones $\{\mathcal{F}_t\}$ y $\{\mathcal{H}_t\}$ mediante

$$\mathcal{F}_t := \mathcal{G}_{N_t+1} \qquad \text{y} \qquad \mathcal{H}_t := \sigma(N_s, s \leqslant t).$$

Con ello obtenemos el siguiente resultado.

6. Lema:

(i) *Los procesos de edad y exceso de vida, A y R, son continuos por la derecha, lineales por pedazos con pendientes $+1$ y -1 y con valores en $\mathbb{R}_+$ y $(0,\infty)$ respectivamente.*

(ii) *Las filtraciones $\{\mathcal{F}_t\}$ y $\{\mathcal{H}_t\}$ satisfacen*

$$\mathcal{H}_t \subset \mathcal{F}_t = \mathcal{H}_t \vee \sigma(T(N_t+1)).$$

Demostración: La parte (i) es clara por la definición de los procesos.

Para (ii) probemos primero que $\mathcal{H}_t \subset \mathcal{F}_t$. Como

$$N_t = \sum_{i=1}^{\infty} \mathbb{1}_{\{T(i) \leqslant t\}},$$

bastará probar que $T(n)$ es tiempo de paro con respecto a $\mathcal{F}_t$, lo cual se hará mediante inducción. El caso base $n = 0$ es trivial. Supongamos que $T(n)$ es tiempo de paro para alguna $n \geqslant 0$. Como el proceso R es continuo por la derecha y $T(n)$ es tiempo de paro, por la proposición 6, vemos que $R_{T(n)}$ es $\mathcal{F}_{T(n)}$-medible. Por otro lado se tiene que $T(n+1) = T(n) = R_{T(n)}$, de modo que

$$\begin{aligned}\{T(n+1) > t\} &= \{0 < T(n) < t, T(n) + R_{T(n)} > t\} \cup \{T(n) \geqslant t, R_{T(n)} \geqslant 0\} \\ &= \bigcup_{r \in [0,t] \cap \mathbb{Q}} \{r < T(n) < t, R_{T(n)} > t - r\} \cup \{T(n) \geqslant t, R_{T(n)} \geqslant 0\}.\end{aligned}$$

De la última expresión se deduce que $\{T(n+1) > t\} \in \mathcal{F}_t$, probando la primera parte del resultado.

Finalmente demostremos que

$$\mathcal{F}_t = \mathcal{H}_t \vee \sigma(T(N_t+1)),$$

notando que $\mathcal{F}_t \supset \mathcal{H}_t \vee \sigma(T(N_t+1))$ en vista de que $\sigma(T(N_t+1)) \subset \mathcal{F}_t$. En cuanto a la contención inversa, primero notemos que $T(n)$ es tiempo de paro con respecto a $\{\mathcal{H}_t\}$. Eso es claro pues $\{T(n) \leqslant t\} = \{N_t \geqslant n\}$. Hecho esto, tomemos $s \leqslant t$ y $x \in (0, t-s]$. Tenemos que

$$\begin{aligned}\{R_s \leqslant x\} = \{T(N_s+1) - s \leqslant x\} &= \bigcup_{n \geqslant 0} \{T(n+1) \leqslant x + s, N_s = n\} \\ &= \bigcup_{n \geqslant 0} \{T(n) \leqslant s < T(n+1) \leqslant x + s\} \in \mathcal{H}_{s+x} \subset \mathcal{H}_t.\end{aligned}$$

Si, en cambio, $x > t - s$, vemos que

$$\{t - s < R_s \leqslant x\} = \{R_s \leqslant t - s\}^c \cap \{T(N_s+1) \leqslant x + s\} \in \mathcal{H}_t \vee \sigma(T(N_t+1)).$$

Esto prueba que R_s es $\mathcal{H}_t \vee \sigma(T(N_t+1))$-medible para toda $s \leqslant t$, lo cual nos da el resultado deseado. ∎

16. Definición (Propiedad de Markov):
Un proceso aleatorio $X = \{X_t : t \geqslant 0\}$ con valores en un espacio polaco E adaptado a la filtración $\{\mathcal{F}_t\}$ y que inicia en $x \in E$ es un *proceso de Markov (con respecto a $\{\mathcal{F}_t\}$)* si para cualesquiera función $f : E \to \mathbb{R}$ medible y acotada y $s, t \geqslant 0$,

$$\mathbb{E}_x[f(X_{t+s}) \mid \mathcal{F}_s] = \mathbb{E}_{X_s}[f(X_{t+s})] \qquad \text{casi seguramente.}$$

8. Teorema:
El proceso R (exceso de vida) satisface la propiedad de Markov con respecto a $\{\mathcal{F}_t\}$; i.e. *para f boreliana, positiva y acotada y $s, t \geqslant 0$,*

$$\mathbb{E}[f(R_{t+s}) \mid \mathcal{F}_t] = \begin{cases} f(x-s) & \text{si } R_t = x > s, \qquad (\star) \\ \int_{[0,s-x]} \mathrm{d}U_v \int_{(s-x-v,\infty)} \mathrm{d}F(y) f(y-s+x+v) & \text{si } R_t = x < s. \quad (\star\star) \end{cases}$$

Demostración de la ecuación ($\star$) del teorema 8: Sea $\Delta \in \mathcal{F}_t$. Como $N_t + 1$ es tiempo de paro con respecto a $\{\mathcal{G}_n\}$, el evento $\Delta \cap \{N_t + 1 = k+1\} \in \mathcal{G}_{k+1}$. Supongamos que $s + t < T(k+1)$, de modo que

$$R_{t+s} = T(N_{t+s} + 1) - (t+s) = R_t - s$$

y por ende

$$\mathbb{E}\Big[f(R_{t+s}) \mathbb{1}_{\Delta \cap \{N_t = k\} \cap \{s+t < T(K+1)\}}\Big] = \mathbb{E}\Big[f(R_t - s) \mathbb{1}_{\Delta \cap \{N_t = k\} \cap \{s+t < T(K+1)\}}\Big].$$

Consecuentemente,

$$\mathbb{E}[f(R_{t+s}) \mid \mathcal{F}_t] = f(R_t - s) \qquad \text{si } R_t > s.$$

■

15 de febrero

Demostración de la ecuación ($\star\star$) del teorema 8: Consideremos $\Lambda \in \mathcal{F}_t$. Recordemos que es importante que $N_t + 1$ es $\mathcal{G}_k$-tiempo de paro,

$$\Lambda \cap \{N_t + 1 = k+1\} \in \mathcal{G}_{k+1} \quad \text{y} \quad R_t = T(N_t + 1) - t.$$

Pediremos que $t + s > T(k+1)$. Tomemos $T'(j) = \xi_{k+2} + \ldots + \xi_{k+1+j}$ para $j \geqslant 1$. Asociado a estos tiempos consideramos el proceso de renovación N' y el proceso de exceso de vida R'. Notemos que $G' \perp\!\!\!\perp \mathcal{G}_{k+1}$ y tiene la misma ley que R por construcción.

Buscaremos representar a R_{t+s} en términos de R_t y R', lo cual estará dado por

$$R_{t+s} = R'_{s+t-T(k+1)} R'_{s-R_t}.$$

Al realizar algunos cálculos obtenemos que

$$\begin{aligned} \mathbb{E}\Big[f(R_{t+s}) \mathbb{1}_{\Lambda \cap \{N_t = k\} \cap \{s+t > T(k+1)\}}\Big] &= \mathbb{E}\Big[f(R'_{s-R_t}) \mathbb{1}_{\Lambda \cap \{N_t = k\} \cap \{s+t > T(k+1)\}}\Big] \\ &= \mathbb{E}\Big[\mathbb{1}_{\Lambda \cap \{N_t = k\} \cap \{s+t > T(k+1)\}} g(R_t)\Big], \end{aligned}$$

donde

$$g(x)=\mathbb{E}[f(R_{s-x})],\qquad s>x.$$

Para concluir, calcularemos de forma explícita a g. Mediante cálculos usuales,

$$\begin{aligned}
\mathbb{E}[f(R_v)] &= \sum_{k\geqslant 0}\mathbb{E}\big[f(T(N_v+1)-v)\mathbb{1}_{\{N_v=k\}}\big]\\
&= \sum_{k\geqslant 0}\mathbb{E}\Big[f(\xi_{k+1}+T(k)-v)\mathbb{1}_{\{T(k)\leqslant v\}\cap\{\xi_{k+1}>v-T(k)\}}\Big]\\
&= \sum_{k\geqslant 0}\mathbb{E}\Big[\mathbb{1}_{\{T(k)\leqslant v\}}\int_{(v-T(k),\infty)}\mathrm{d}F(y)f(T(k)-v+y)\Big]\\
&= \int_{[0,v]}\mathrm{d}U_r\int_{(v-r,\infty)}\mathrm{d}F(y)f(r-v+y).
\end{aligned}$$

Al tomar $v=s-x$ concluimos. ■

Para el proceso de edad, la propiedad de Markov es un poco distinta. El proceso sigue siendo de Markov pero no bajo $\{\mathcal{F}_t\}$ sino bajo $\{\mathcal{H}_t\}$. Para $s\geqslant x$ y f boreliana, positiva y acotada, definimos al operador

$$G_sf(x):=\mathbb{E}[f(A_{s-x})].$$

Notemos que

$$\begin{aligned}
G_sf(x) &= \mathbb{E}[f(s-x-T(N_{s-x}))]\\
&= \sum_{k=0}^{\infty}\mathbb{E}\Big[f(s-x-T(k))\mathbb{1}_{\{T(k)\leqslant s-x\}\cap\{\xi_{k+1}>s-x-T(k)\}}\Big]\\
&= \sum_{k\geqslant 0}\mathbb{E}\Big[f(s-x-T(k))\mathbb{1}_{\{T(k)\leqslant s-x\}}\overline{F}(s-x-T(k))\Big]\\
&= \int_{[0,s-x]}\mathrm{d}U_yf(s-y-x)\overline{F}(s-x-y).
\end{aligned}$$

9. Teorema:
El proceso de edad A verifica la propiedad de Markov con respecto a la filtración $\{\mathcal{H}_t\}$; i.e. *para f boreliana, positiva y acotada y $s,t\geqslant 0$,*

$$\mathbb{E}[f(A_{t+s})\mid\mathcal{H}_t]=\begin{cases} f(A_t+s) & \text{si } \{N_t=N_{t+s}\},\\ \dfrac{1}{\overline{F}(A_t)}\displaystyle\int\mathbb{1}_{\{A_t<y\leqslant s+A_t\}}G_sf(y-A_t)\,\mathrm{d}F(y) & \text{en otro caso.}\end{cases}$$

Demostración: Sea $\Lambda\in\mathcal{H}_t$. Como todo conjunto $\mathcal{H}_t$-medible restringido al evento

$$\{N_t=n\}$$

es $\mathcal{G}_n$-medible, tenemos que $\Lambda\cap\{N_t=n\}\in\mathcal{G}_n$.

Primero supondremos que $s+t<T(n+1)$. En este caso se tiene la relación

$$A_{t+s}=A_t+s$$

y por ende

$$\mathbb{E}\Big[f(A_{t+s})\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t<T(n+1)\}}\Big]=\mathbb{E}\Big[f(A_t+s)\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t<T(n+1)\}}\Big],$$

de modo que

$$\mathbb{E}\big[f(A_{t_s})\,\big|\,\mathcal{H}_t\big]=f(A_t+s)\qquad\text{si }\{N_t=N_{t+s}\}.$$

Supondremos ahora que $s+t\geqslant T(n+1)$. Definimos a $T'(j)=\xi_{n+2}+\cdots+\xi_{n+1+j}$ para $j\geqslant 1$. Consideremos además a A' y N' asociados a $\{T'(j)\}$. Nuevamente $A'\perp\!\!\!\perp\mathcal{G}_{n+1}$, con lo cual podemos realizar cálculos. Notando que

$$A_{t+s}=A'_{s+A_t-\xi_{n+1}}.$$

Luego

$$\begin{aligned}
&\mathbb{E}\Big[f(A_{t+s})\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t\geqslant T(n+1)\}}\Big]\\
&\quad=\mathbb{E}\Big[f(A'_{s+A_t-\xi_{n+1}})\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t\geqslant T(n+1)\}}\Big]\\
&\quad=\mathbb{E}\Big[\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t\geqslant T(n)+\xi_{n+1}\}}G_sf(\xi_{n+1}-A_t)\Big]\\
&\quad=\mathbb{E}\Big[\mathbb{1}_{\Lambda\cap\{T(n)\leqslant t\}}\mathbb{1}_{\{A_t<\xi_{n+1}\leqslant s+A_t\}}G_sf(\xi_{n+1}-t+T(n))\Big]\\
&\quad=\mathbb{E}\left[\mathbb{1}_{\Lambda\cap\{T(n)\leqslant t\}}\frac{\overline{F}(A_t)}{\overline{F}(A_t)}\int_{(A_t,s+A_t]}G_sf(y-A_t)\,\mathrm{d}F(y)\right]\\
&\quad=\mathbb{E}\Big[\mathbb{1}_{\Lambda\cap\{A_t\geqslant 0\}}\overline{F}(A_t)H_sf(A_t)\Big],
\end{aligned}$$

donde

$$H_sf(x)=\frac{1}{\overline{F}(x)}\int_{(x,x+s]}G_sf(y-x)\,\mathrm{d}F(y).$$

De este modo,

$$\mathbb{E}\Big[f(A_{t+s})\mathbb{1}_{\Lambda\cap\{N_t=n\}\cap\{s+t\geqslant T(n+1)\}}\Big]=\mathbb{E}\big[\mathbb{1}_{\Lambda\cap\{N_t=n\}}H_sf(A_t)\big].\qquad\blacksquare$$

Ahora estudiaremos el comportamiento asintótico del par (A_t,R_t). Para ello sean U y Z dos variables aleatorias independientes, donde U es uniforme en $[0,1]$ y la ley de Z satisface que

$$\mathbb{P}(Z\in\mathrm{d}z)=\frac{z}{m}\,\mathrm{d}F(z),\qquad \mathrm{d}F(z)=\mathbb{P}(\xi_1\in\mathrm{d}z),\ m=\mathbb{E}[\xi_1]<\infty.$$

10. Teorema:
Supongamos que el proceso de renovación es no aritmético y que $m < \infty$. Entonces

$$(A_t, R_t) \Rightarrow_t (UZ, (1-U)Z).$$

Demostración: Consideremos $a, r \geqslant 0$. Veamos que

$$\begin{aligned}
\mathbb{P}(A_t \leqslant a, R_t > r) &= \sum_{k \geqslant 0} \mathbb{E}\left[\mathbb{1}_{\{T(k) \leqslant t < T(k)+\xi_{k+1}\}} \mathbb{1}_{\{t-T(k) \leqslant a\}} \mathbb{1}_{\{\xi_{k+1}+T(k)-t>r\}}\right] \\
&= \sum_{k \geqslant 0} \mathbb{E}\left[\mathbb{1}_{\{t-a \leqslant T(k) \leqslant t\}} \mathbb{1}_{\{\xi_{k+1}+T(k)>r+t\}}\right] \\
&= \sum_{k \geqslant 0} \mathbb{E}\left[\mathbb{1}_{\{t-a \leqslant T(k) \leqslant t\}} \overline{F}(r+t-T(k))\right] \\
&= \int_{\{t-x \leqslant a\}} \overline{F}(r+t-x)\, \mathrm{d}U_x.
\end{aligned}$$

Por último veamos que U es medida de renovación y notemos que

$$x \longmapsto \mathbb{1}_{\{t-x \leqslant a\}} \overline{F}(t-x+r)$$

es dRI, ya que es positiva continua y decreciente, lo cual implica que

$$\mathbb{P}(A_t \leqslant a, R_t > r) \to \frac{1}{m} \int_0^{\infty} \mathbb{1}_{\{x<a\}} \overline{F}(x+t)\, \mathrm{d}x = \mathbb{P}(UZ \leqslant a, (1-U)Z > r). \qquad \blacksquare$$

3.3. Variación regular y teorema de Dynkin–Lamperti

22 de febrero

El objetivo ahora es estudiar el comportamiento asintótico de los procesos de edad y de exceso de vida en el caso en que el proceso de renovación tiene media infinita; *i.e.* $\mathbb{E}[\xi_1] = \infty$.

17. Definición:
Sean $\alpha \in \mathbb{R}$ y $f : (0, \infty) \to (0, \infty)$ una función medible. Decimos que f *varía regularmente* en $+\infty$ (respectivamente en 0^+) con índice α si para todo $\lambda > 0$,

$$\lim_{x \to \infty / 0^+} \frac{f(\lambda x)}{f(x)} = \lambda^{\alpha}.$$

Si $\alpha = 0$, decimos que f *varía lentamente*.

1. *Observación:*
Si f varía regularmente con índice α, entonces $x \longmapsto x^{-\alpha} f(x)$ varía lentamente. Recíprocamente si $\ell(x)$ varía lentamente entonces $x \longmapsto x^{\alpha}\ell(x)$ varía regularmente con índice α.

(Ejercicio) Probar que si f tiene límite finito y positivo entonces f varía lentamente. Además probar que

$$f(x) = \log x \qquad \text{y} \qquad g(x) = \frac{1}{\log x}$$

varían lentamente.

(Ejercicio) Sean c y ϵ funciones medibles tales que $\lim_{x\to+\infty} c(x)$ existe en $(0,\infty)$ y $\lim_{x\to+\infty} \epsilon(x) = 0$. Definamos a

$$f(x) = c(x)\exp\left\{\int_1^x \frac{\epsilon(t)}{t}\,\mathrm{d}t\right\}.$$

Probar que f varía lentamente en $+\infty$.

Es importante señalar que esta es la forma general de una función que varía lentamente.

7. Proposición (Convergencia uniforme)**:**
Si f varía regularmente en $+\infty$ (respectivamente en 0^+) con índice α, entonces para todo compacto $K \subset (0,\infty)$,

$$\lim_{x\to+\infty/0^+} \frac{f(\lambda x)}{f(x)} = \lambda^\alpha, \quad \textit{uniformemente en } K.$$

Antes de probar el teorema abeliano/tauberiano, vamos a recordar algunos conceptos y resultados. Sea μ una medida de proba en $[0,\infty)$, vamos a denotar por $\mathcal{L}\mu$ a la *transformada de Laplace de μ; i.e.*

$$\mathcal{L}\mu(\lambda) = \int_{[0,\infty)} e^{-\lambda x}\,\mu(\mathrm{d}x).$$

Recordemos que la función $\Gamma : (0,\infty) \to (0,\infty)$ está definida por

$$\Gamma(\beta) = \int_0^\infty e^{-t} t^{\beta-1}\,\mathrm{d}t.$$

Sea $\{\rho_n\}_{n\geqslant 1}$ una sucesión de medidas de probabilidad en $[0,\infty)$. Decimos que $\{\rho_n\}_{n\geqslant 1}$ *converge en distribución* hacia una probabilidad ρ si y solamente si $\{\mathcal{L}\rho_n\}_{n\geqslant 1}$ converge hacia $\mathcal{L}\rho$.

11. Teorema (de Lévy)**:**
Sea $\{\rho_n\}_{n\geqslant 1}$ una sucesión de medidas de probabilidad en $[0,\infty)$. Si $\lim_n \mathcal{L}\rho_n = f$, donde $f(0) = 1$ y f es continua por la derecha en 0, entonces $f = \mathcal{L}\rho$ para alguna medida de probabilidad ρ en $[0,\infty)$.

Este resultado se puede extender de manera inmediata al caso en que las medidas ρ_n satisfacen $\rho_n([0,\infty)) = \infty$ para toda n. En este caso, trabajamos con

$$\frac{e^{-x}\rho_n(\mathrm{d}x)}{\mathcal{L}\rho_n(1)}$$

y con su transformada de Laplace

$$\frac{\mathcal{L}\rho_n(\lambda+1)}{\mathcal{L}\rho_n(1)}, \qquad \lambda \geqslant 0.$$

12. Teorema (Tauberiano/Abeliano)**:**
Sea μ una medida en $[0,\infty)$ con transformada de Laplace $\mathcal{L}_\mu$. Entonces son equivalentes

(i) *$\mu(x) := \mu([0,x])$ varía regularmente en $+\infty$ con índice α y*
(ii) *$\mathcal{L}\mu$ varía regularmente en 0^+ con índice $-\alpha$.*

En cualquier caso,

$$\mu(x) \sim \frac{1}{\Gamma(\alpha+1)}\mathcal{L}\mu(1/x) \quad \textit{conforme } x \to \infty.$$

Demostración: **Parte abeliana.** Supongamos que $\mu(\cdot)$ varía regularmente en $+\infty$ con índice α. Recordemos que para toda $y > 0$,

$$\mu_x(y) := \frac{\mu(xy)}{\mu(x)} \to y^\alpha,$$

uniformemente sobre todo compacto de $(0,\infty)$ y definamos la medida de probabilidad

$$\rho_x(\mathrm{d}y) := \frac{e^{-y}\mu_x(\mathrm{d}y)}{\int_{(0,\infty)} e^{-z}\mu_x(\mathrm{d}z)}.$$

De la convergencia en distribución, tenemos

$$\rho_x(\mathrm{d}y) \to \frac{e^{-y}y^{\alpha-1}\,\mathrm{d}y}{\int_{(0,\infty)} e^{-z}z^{\alpha-1}\,\mathrm{d}z}$$

conforme $x \to \infty$. En particular la transformada de Laplace de ρ_x satisface

$$\int_{(0,\infty)} e^{-\lambda y}\rho_x(\mathrm{d}y) = \frac{\mathcal{L}\mu_x(\lambda+1)}{\mathcal{L}\mu(1)} \to \frac{\alpha\int_0^\infty e^{-(\lambda+1)y}y^{\alpha-1}\,\mathrm{d}y}{\alpha\int_0^\infty e^{-y}y^{\alpha-1}\,\mathrm{d}y}. \tag{$\star$}$$

Usando el cambio de variable $x = (\lambda+1)y$, vemos

$$\int_0^\infty e^{-(\lambda+1)y}y^{\alpha-1}\,\mathrm{d}y = \frac{\Gamma(\alpha)}{(1+\lambda)^\alpha}.$$

Por otro lado

$$\int_0^\infty e^{-(\lambda+1)y}\,\mathrm{d}\mu(xy) = \int_0^\infty e^{-(\lambda+1)t/x}\,\mathrm{d}\mu(t) = \mathcal{L}\mu((\lambda+1)/x).$$

Por lo tanto, de ($\star$) y los cálculos anteriores vemos

$$\frac{\mathcal{L}\mu((\lambda+1)/x)}{\mathcal{L}\mu(1/x)} \to \frac{1}{(\lambda+1)^\alpha} \quad \text{conforme } x \to \infty.$$

Esto es, $\mathcal{L}\mu$ varía regularmente en 0^+ con índice $-\alpha$.

Parte tauberiana. Supongamos que $\mathcal{L}\mu$ varía regularmente en 0^+ con índice $-\alpha$. Veamos que la aplicación

$$\lambda \mapsto \frac{\mathcal{L}\mu(\lambda/x)}{\mathcal{L}\mu(1/x)}$$

es la transformada de Laplace de $\mu(xy)/\mathcal{L}\mu(1/x)$. En efecto,

$$\int_{[0,\infty)} e^{-\lambda y} \,\mathrm{d}_y \frac{\mu(xy)}{\mathcal{L}\mu(1/x)} = \frac{1}{\mathcal{L}\mu(1/x)} \int_{[0,\infty)} e^{-\lambda t/x} \,\mathrm{d}\mu(t) = \frac{\mathcal{L}\mu(\lambda/x)}{\mathcal{L}\mu(1/x)}.$$

Por hipótesis cuando $x \to \infty$, esta transformada de Laplace converge a

$$\Gamma(\alpha)\lambda^{-\alpha} = \int_0^\infty e^{-\lambda t} t^{\alpha-1} \,\mathrm{d}t.$$

De la convergencia vaga de medidas, tenemos

$$\frac{\mu(xy)}{\mathcal{L}(1/x)} \to \frac{1}{\Gamma(\alpha)} \int_0^y t^{\alpha-1} \,\mathrm{d}t = \frac{y^\alpha}{\alpha\Gamma(\alpha)} = \frac{y^\alpha}{\Gamma(\alpha+1)}$$

conforme $x \to \infty$. Por tanto

$$\mu(x) \sim \frac{\mathcal{L}\mu(1/x)}{\Gamma(1+x)} \quad \text{conforme } x \to \infty,$$

y por ende para $c > 0$,

$$\frac{\mu(cx)}{\mu(x)} \sim \frac{\mathcal{L}\mu(1/cx)}{\mathcal{L}\mu(1/x)} \to c^\alpha, \quad \text{conforme } x \to \infty.$$

En otras palabras μ varía regularmente en $+\infty$ con índice α. ■

7. Lema (Densidades monótonas):
Supongamos que $\mathrm{d}\mu(x) = g(x)\,\mathrm{d}x$ *con* g *monótona. Si además suponemos que* μ *varía regularmente en* $+\infty$ *con índice* $\alpha > 0$, *entonces*

$$g(x) \sim \frac{\alpha\mu(x)}{x} \quad \textit{conforme } x \to \infty.$$

En particular g *es de variación regular en* $+\infty$ *con índice* $\alpha - 1$.

Demostración: Supongamos que g es creciente. Por hipótesis,

$$\frac{\mu((1+\varepsilon)x)}{\mu(x)} \to (1+\varepsilon)^\alpha \quad \text{conforme } x \to \infty.$$

Entonces

$$\frac{\mu((1+\varepsilon)x) - \mu(x)}{\varepsilon\mu(x)} \xrightarrow{x\to\infty} \frac{(1+\varepsilon)^\alpha - 1}{\varepsilon} \xrightarrow{\varepsilon\to 0} \alpha.$$

Pero

$$\mu((1+\varepsilon)x)-\mu(x)=\int_x^{(1+\varepsilon)x} g(y)\,\mathrm{d}y \in [g(x)x\varepsilon, g((1+\varepsilon)x)x\varepsilon].$$

De este modo

$$\limsup_{x\to\infty}\frac{g(x)x}{\mu(x)} \leqslant \lim_{x\to\infty}\frac{\mu((1+\varepsilon)x)-\mu(x)}{\varepsilon\mu(x)} \leqslant \alpha.$$

De manera similar,

$$\liminf_{x\to\infty}\frac{g((1+\varepsilon)x)}{\mu(x)} \geqslant \lim_{x\to\infty}\frac{\mu((1+\varepsilon)x)-\mu(x)}{\varepsilon\mu(x)} = \frac{(1+\varepsilon)^\alpha-1}{\varepsilon}.$$

Al tomar ε suficientemente pequeño tenemos que

$$\liminf_{x\to\infty}\frac{g(x)}{\mu(x)} \geqslant \alpha. \qquad \blacksquare$$

Pruebas mucho más completas y explicadas del teorema 12 y el lema 7 se pueden consultar en las páginas 37–39 de Bingham y col. (1987). De forma respectiva son el teorema 1.7.1 y teorema 1.7.2 de la referencia dada.

24 de febrero

13. Teorema (de Dynkin–Lamperti)**:**
Sean A y R los procesos de edad y exceso de vida de un proceso de renovación con media infinita, i.e. $\mathbb{E}[\xi_1]=\infty$*. Entonces son equivalentes:*

(i) *para alguna $\alpha\in(0,1)$ tenemos $\mathbb{E}[A_t]\sim\alpha t$, $t\to\infty$;*
(ii) *A_t/t tiene un límite en ley no degenerado conforme $t\to\infty$;*
(iii) *$(A_t,R_t)/t$ converge en ley, conforme $t\to\infty$, hacia la ley*

$$\frac{1-\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}(1-a)^{-\alpha}(a+r)^{\alpha-2}, \quad 0<a<1, \quad r>0;$$

(iv) *$\overline{F}(x)$ varía regularmente en $+\infty$ con índice $\alpha-1$.*

2. *Observación:*
La primera marginal en (iii) es la ley arcoseno generalizada o Beta$(\alpha,1-\alpha)$ que tiene densidad

$$\frac{\sin\pi\alpha}{\pi}a^{\alpha-1}(1-a)^{-\alpha}, \quad 0<a<1.$$

Demostración: Primero notemos que se satisfacen las siguientes implicaciones: (iii) + la observación implica (ii), la cual implica (i). Probemos pues que (i) implica (iv). Supongamos que $\mathbb{E}[A_t]\sim\alpha t$ e introduzcamos la notación,

$$\phi(q):=1-\int_{[0,\infty)} e^{-qx}\,\mathrm{d}F(x), \qquad q>0.$$

Sea U la medida de renovación, la cual recordamos está dada por

$$U(x) = \sum_{n\geqslant 0} F^{*n}(x).$$

Entonces la transformada de Laplace de U satisface

$$\mathcal{L}U(q) := \int_{[0,\infty)} e^{-qx}\,\mathrm{d}U_x = \sum_{n\geqslant 0} \mathcal{L}F^{*n}(q) = \sum_{n\geqslant 0}(1-\varphi(q))^n = \frac{1}{\varphi(q)}.$$

También tenemos

$$\int_0^\infty e^{-qx}\overline{F}(x)\,\mathrm{d}x = \int_0^\infty \mathrm{d}x\,e^{-qx}\int_{(x,\infty)}\mathrm{d}F(y) = \frac{1}{q}\int_{(0,\infty)}\mathrm{d}F(y)(1-e^{-qy}) = \frac{\varphi(q)}{q}.$$

Ahora verifiquemos que para $f \geqslant 0$,

$$\begin{aligned}\mathbb{E}[f(A_t)] &= \int_{[0,t]} f(t-x)\overline{F}(t-x)\,\mathrm{d}U_x = \mathbb{E}\left[\sum_{n\geqslant 0}\mathbb{1}_{\{T(n)\leqslant t<T(n)+\xi_{n+1}\}}f(t-T(n))\right]\\ &= \mathbb{E}\left[\sum_{n\geqslant 0}\mathbb{1}_{\{T(n)\leqslant t\}}f(t-T(n))\overline{F}(t-T(n))\right].\end{aligned}$$

En particular, para $f(x) = x$ se tiene

$$\mathbb{E}[A_t] = \int_{[0,t]}(t-x)\overline{F}(t-x)\,\mathrm{d}U_x.$$

Estamos interesados en la transformada de Laplace de $t \mapsto \mathbb{E}[A_t]$, la cual satisface

$$\begin{aligned}\int_0^t e^{-qt}\mathbb{E}[A_t]\,\mathrm{d}t &= \int_0^\infty \mathrm{d}t\,e^{-qt}\int_{[0,t]}(t-x)\overline{F}(t-x)\,\mathrm{d}U_x\\ &= \int_{[0,\infty)}\mathrm{d}U_x e^{-qx}\int_0^\infty \mathrm{d}t\,e^{-q(t-x)}\mathbb{1}_{\{0\leqslant x\leqslant t\}}(t-x)\overline{F}(t-x) \qquad (\star)\\ &= \frac{1}{\varphi(q)}\int_0^\infty \mathrm{d}y\,e^{-qy}\overline{F}(y)y.\end{aligned}$$

Calculemos

$$\begin{aligned}\int_0^\infty \mathrm{d}y\,y\,e^{-qy}\overline{F}(y) &= -\frac{\mathrm{d}}{\mathrm{d}q}\int_0^\infty e^{-qx}\overline{F}(x)\,\mathrm{d}x\\ &= -\frac{\varphi'(q)}{q} + \frac{\varphi(q)}{q},\end{aligned}$$

de modo que por $(\star)$,

$$\int_0^t e^{-qt}\mathbb{E}[A_t]\,\mathrm{d}t = \frac{1}{\varphi(q)}\left(-\frac{\varphi'(q)}{q} + \frac{\varphi(q)}{q}\right) = \frac{1}{q^2} - \frac{1}{q}\frac{\varphi'(q)}{\varphi(q)}.$$

Por hipótesis sabemos que $\mathbb{E}[A_t] \sim \alpha t$ conforme $t \to \infty$. Por lo tanto

$$\int_0^t \mathbb{E}[A_s]\,\mathrm{d}s \sim \frac{\alpha t^2}{2}, \quad t\to\infty,$$

y por ende varía regularmente en $+\infty$ con índice 2. Del teorema 12,

$$\int_0^\infty e^{-qt}\mathbb{E}[A_t]\,\mathrm{d}t \sim \frac{2\alpha}{2}q^{-2}, \quad q \to 0^+.$$

Esto implica $\alpha/q^2 \sim q^{-2} - q^{-1}\varphi'(q)/\varphi(q)$, $q \to 0^+$, lo que a su vez implica que

$$\frac{\varphi'(q)}{\varphi(q)} \sim \frac{1-\alpha}{q}, \quad q \to 0^+.$$

Definamos

$$\frac{\epsilon(q)}{q} := \frac{\varphi'(q)}{\varphi(q)} - \frac{1-\alpha}{q} = \mathrm{o}(1/q).$$

Entonces

$$\int_s^1 \frac{\epsilon(q)}{q}\,\mathrm{d}q = \log\varphi(1) - \log\varphi(s) + 1 - \alpha\log s,$$

lo cual implica

$$\varphi(q) = q^{1-\alpha}\varphi(1)\exp\left\{-\int_q^1 \frac{\epsilon(u)}{u}\,\mathrm{d}u\right\}.$$

Por algún ejercicio, sabemos que

$$\ell(q) := \exp\left\{-\int_q^1 \frac{\epsilon(u)}{u}\,\mathrm{d}u\right\}$$

es de variación lenta en 0^+. Esto implica que $\varphi(q)$ varía regularmente con índice $\alpha - 1$ y por ende

$$\int_0^\infty e^{-qx}\overline{F}(x)\,\mathrm{d}x = \frac{\varphi(q)}{q}$$

varía regularmente en 0^+ con índice $-\alpha$. Del teorema 12tenemos que la medida $\int_0^x \overline{F}(y)\,\mathrm{d}y$ varía regularmente en $+\infty$ con índice α y además,

$$\int_0^x \overline{F}(y)\,\mathrm{d}y \sim \frac{1}{\Gamma(1+\alpha)}\frac{\varphi(1/x)}{1/x}, \quad x \to \infty.$$

Como $\overline{F}$ es una función decreciente, entonces el lema 7 implica

$$\overline{F}(x) \sim \frac{\alpha}{x}\frac{x}{\Gamma(1+\alpha)}\varphi(1/x) = \frac{1}{\Gamma(\alpha)}\varphi(1/x), \quad x \to \infty. \tag{$\star\star$}$$

De manera recíproca, si (iv) se satisface, el teorema 12 nos garantiza que ($\star\star$) se cumple. Así, hemos probado la equivalencia entre (i) y (iv).

Para concluir mostremos que (i) implica (iii). Para ello consideremos $0 < a < a' < 1$ y $r > 1$, entonces

$$\begin{aligned}\mathbb{P}(at \leqslant A_t < a't, R_t > rt) &= \mathbb{E}\left[\sum_{n\geqslant 0} \mathbb{1}_{\{T(n)\leqslant t<T(n)+\xi_{n+1}\}}\mathbb{1}_{\{\xi_{n+1}>rt+t-T(n)\}}\mathbb{1}_{\{at\leqslant t-T(n)<a't\}}\right]\\
&= \mathbb{E}\left[\sum_{n\geqslant 0} \overline{F}((r+1)t-T(n))\mathbb{1}_{\{at\leqslant t-T(n)<a't\}}\right]\\
&= \int_{(1-a')t}^{(1-a)t} \frac{\overline{F}((r+1)t-x)}{\varphi(1/t)}\,\mathrm{d}(U_x\varphi(1/t))\\
&= \int_{(1-a')t}^{(1-a)t} \frac{\overline{F}(r+1-t)t}{\varphi(1/t)}\,\mathrm{d}(U_{ty}\varphi(1/t)).\end{aligned}$$

Por un lado recordemos que $\mathcal{L}U(q) = 1/\varphi(q)$ varía regularmente en 0^+ con índice $\alpha-1$. Del teorema 12 tenemos que U_t varía regularmente en $+\infty$ con índice $1-\alpha$ y además

$$U_t \sim \frac{1}{\Gamma(2-\alpha)}\frac{1}{\varphi(1/t)}, \quad t\to\infty.$$

Como

$$\frac{U_{ty}}{U_t} \to y^{1-\alpha}, \quad t\to\infty,$$

lo que implica que

$$U_{ty}\varphi(1/t) \sim \frac{1}{\Gamma(2-\alpha)}y^{1-\alpha}, \quad t\to\infty.$$

De la misma manera, tenemos

$$\frac{\overline{F}((r+1-y)t)}{\overline{F}(t)} \to (r+1-y)^{\alpha-1}, \quad t\to\infty,$$

de modo que, por (⋆⋆)

$$\overline{F}(t(r+1-y))\varphi(1/t) \sim \frac{1}{\Gamma(\alpha)}(r+1-y)^{\alpha-1}, \quad t\to\infty,$$

uniformemente en $y\in[1-a',1-a]$. Esto implica la convergencia en distribución de la integral de arriba, es decir

$$\begin{aligned}\int_{1-a'}^{1-a}\frac{\overline{F}(t(r+1-y))}{\varphi(1/t)}\,\mathrm{d}(U_{ty}\varphi(1/t)) &\to \frac{1}{\Gamma(\alpha)}\frac{1-\alpha}{\Gamma(2-\alpha)}\int_{1-a'}^{1-a}(r+1-y)^{\alpha-1}y^{-\alpha}\,\mathrm{d}y\\
&= \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_a^{a'}(r+x)(1-x)^{-\alpha}\,\mathrm{d}x. \qquad \blacksquare\end{aligned}$$

3.4. Aplicaciones a las variables aleatorias estables en $\mathbb{R}_+$

1 de marzo

El objetivo es extender la ley de los grandes números para sumas de variables aleatorias independientes e idénticamente distribuidas con media infinita. Tomemos $\{\xi_n\}_{n\geq 1}$ una sucesión de variables aleatorias independientes idénticamente distribuidas con ley común $\mathrm{d}F$ y recordemos que

$$T(n) := \xi_1 + \cdots + \xi_n, \quad \overline{F} = 1 - F \quad \text{y} \quad \mathcal{L}F(q) = \int_{[0,\infty)} e^{-qx}\, \mathrm{d}F(x).$$

14. Teorema:
Bajo las hipótesis del teorema de Dynkin–Lamperti tenemos:

(i) *Para $\alpha \in (0,1)$, $\overline{F}$ varía regularmente en $+\infty$ con índice $-\alpha$ si y solamente si $1 - \mathcal{L}F$ varía regularmente en 0 con índice α.*

(ii) *Bajo las hipótesis de la parte (i), consideremos a $a_n > 0$ tal que $1 - \mathcal{L}F(1/a_n) = 1/n$. Entonces $a_n \nearrow \infty$ (de hecho varía regularmente en $+\infty$ con índice $1/\alpha$) y*

$$\frac{T(n)}{a_n} \Rightarrow_n \sigma(\alpha),$$

donde $\sigma(\alpha)$ es una variable aleatoria estable estándar de índice α; i.e.

$$\mathbb{E}[e^{-\lambda\sigma(\alpha)}] = e^{-\lambda^\alpha}.$$

Demostración: La parte (i) forma parte del enunciado del teorema de Dynkin–Lamperti. Recordemos que $\varphi(q) = 1 - \mathcal{L}F(q)$, la cual es una función continua y creciente ya que $\mathcal{L}F$ es transformada de Laplace. Entonces la función $\varphi : \mathbb{R}_+ \to [0,1]$ es una biyección, de aquí que exista una sucesión monótona $\{a_n\}_{n\geq 1}$ que satisface

$$\varphi(1/a_n) = 1/n.$$

Ahora es claro que

$$\mathbb{E}[e^{-qT(n)}] = \left(\mathcal{L}F(q)\right)^n,$$

lo cual implica

$$\mathbb{E}\left[e^{-qT(n)/a_n}\right] = \left(\mathcal{L}F(q/a_n)\right)^n = \left(1 - \varphi(q/a_n)\right)^n.$$

Por otro lado, sabemos que φ varía regularmente en 0^+ con índice α, por lo tanto

$$\frac{\varphi(q/a_n)}{\varphi(1/a_n)} = n\varphi(q/a_n) \to q^\alpha, \quad n \to \infty.$$

Eso va a implicar

$$n\log(1-\varphi(q/a_n)) = n\varphi(q/a_n)\frac{\log(1-\varphi(q/a_n))}{\varphi(q/a_n)} \xrightarrow[n\to\infty]{} -q^\alpha.$$

Por ende,

$$\mathbb{E}\left[e^{-qT(n)/a_n}\right] \to e^{-q^\alpha}.$$

La función $q \mapsto e^{-q^\alpha}$ es continua en 0 y vale 1 en 0. Del teorema de Lévy, $T(n)/a_n$ converge débilmente a una variable aleatoria con transformada de Laplace e^{-q^α}. ∎

Nota: El concepto de estabilidad se da por la igualdad en distribución

$$\sigma(\alpha) = a\tilde{\sigma}(\alpha) + b\sigma'(\alpha)$$

con a y b distintas de cero y $\tilde{\sigma}(\alpha), \sigma'(\alpha)$ son copias independientes de $\sigma(\alpha)$.

4. Procesos y medidas aleatorias de Poisson

18. Definición:
Vamos a llamar proceso de Poisson con parámetro $c > 0$ a un proceso de renovación $\{N_t : t \geqslant 0\}$, el cual está asociado a una variable aleatoria exponencial de parámetro $c > 0$; *i.e.* $\mathrm{d}F(x) = c\,e^{-cx}\mathbb{1}_{\mathbb{R}_+}(x)\,\mathrm{d}x$.

8. Proposición (Propiedad de Markov)**:**
El proceso de Poisson $N = \{N_t : t \geqslant 0\}$ *tiene incrementos estacionarios e independientes;* i.e. *para* $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_n$ *las variables aleatorias* $N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}}$ *son independientes y* $N_{t_i} - N_{t_{i-1}} \stackrel{(d)}{=} N_{t_i - t_{i-1}}$. *Además es un proceso de Markov homogéneo;* i.e. *si* $\mathcal{F}_t = \sigma(N_s, s \leqslant t)$, $\Lambda \in \mathcal{F}_t$ *y toda función* $f : \mathbb{N} \to \mathbb{R}$ *acotada,*

$$\mathbb{E}[\mathbb{1}_\Lambda f(N_{t+s}) \mid N_t = k] = \mathbb{P}(\Lambda \mid N_t = k)\mathbb{E}[f(k + N_s)].$$

Demostración: Basta probar la segunda parte del enunciado, ya que la primera parte se sigue de un argumento inductivo. De hecho si admitimos la propiedad de Markov tenemos que para $\Lambda \in \mathcal{F}_t$,

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_\Lambda f(N_{t+s} - N_t)] &= \sum_{k\geqslant 0} \mathbb{P}(N_t = k)\mathbb{E}[\mathbb{1}_\Lambda f(N_{t+s} - k) \mid N_t = k] \\
&= \sum_{k\geqslant 0} \mathbb{P}(N_t = k)\mathbb{P}(\Lambda \mid N_t = k)\mathbb{E}[f(N_s)] \\
&= \mathbb{P}(\Lambda)\mathbb{E}[f(N_s)],
\end{aligned}$$

lo cual implica que $N_{t+s} - N_t$ es independiente de $\mathcal{F}_t$ y en adición $N_{t+s} - N_t \stackrel{(d)}{=} N_s$.

Para probar la propiedad de Markov, consideremos

$$\mathbb{E}\left[\mathbb{1}_\Lambda f(N_{t+s})\mathbb{1}_{\{N_t = k\}}\right]$$

con $\Lambda \in \mathcal{F}_t$. Bajo el evento $\{N_t = k\}$, $\mathbb{1}_\Lambda$ se puede ver como una función medible de $T(1), \dots, T(k)$. A dicha función la vamos a denotar por $\varphi(T(1), \dots, T(k))$. Ahora bajo el evento $\{N_t = k\}$,

$$N_{t+s} = k + \#\{p \geqslant 1 : T(k) + \xi_{k+1} + \cdots + \xi_{k+p} \leqslant t + s\}.$$

Gracias a la ausencia de memoria de la ley exponencial, al condicionar con respecto al evento $\{\xi_{k+1} > t - T(k)\}$, la variable aleatoria $\xi_{k+1} - (t - T(k))$ se distribuye como una variable aleatoria exponencial de parámetro c y por lo tanto, al condicionar con respecto a $\{N_t = k\}$ tenemos que $N_{t+s} = k + \widetilde{N}_s$, donde $\widetilde{N}_s := \#\{p \in \mathbb{N} : \widetilde{\xi}_1 \cdots + \widetilde{\xi}_p \leqslant s\}$, con

$$\widetilde{\xi}_1 = \xi_{k+1} - (t - T(k)) \quad \text{y} \quad \widetilde{\xi}_p = \xi_{k+p}, \ p \geqslant 2.$$

Además bajo $\{N_t = k\}$, las variables aleatorias $\{\xi_i\}_{1 \leqslant i \leqslant k}$ y $\{\widetilde{\xi}_i\}_{1 \leqslant i \leqslant p}$ son independientes. Entonces

$$\begin{aligned}\mathbb{E}[\mathbb{1}_\Lambda f(N_{t+s})\mathbb{1}_{\{N_t=k\}}] &= \mathbb{E}[\varphi(T(1), \dots, T(k)) f(\widetilde{N}_s + k)\mathbb{1}_{\{N_t=k\}}] \\ &= \mathbb{E}[\varphi(T(1), \dots, T(k))\mathbb{1}_{\{N_t=k\}}]\mathbb{E}[f(\widetilde{N}_s + k)]. \qquad \blacksquare\end{aligned}$$

De ahora en adelante vamos a denotar por $\mathbb{P}_x$, $x \in \mathbb{N}$, a la ley del proceso $x + N$; *i.e.* la ley del proceso de Poisson que parte de x. La propiedad de Markov se puede enunciar de la siguiente manera: para $t \geqslant 0$, si definimos $\{N'_s = N_{t+s}, s \geqslant 0\}$, entonces condicionando con respecto a $\{N_t = k\}$, el proceso N' es independiente de $\mathcal{F}_t$ y tiene por ley $\mathbb{P}_k$.

3 de marzo

9. Proposición (Propiedad de Markov fuerte):
Sea T un $\{\mathcal{F}_t\}_{t \geqslant 0}$-tiempo de paro finito casi seguramente. Sean $N'_s := N_{T+s}$ y $\widetilde{N}_s = N_{T+s} - N_T$, $s \geqslant 0$. Entonces $\widetilde{N}$ es independiente de $\mathcal{F}_T$ y tiene la misma ley que N y condicionando con respecto a $\{N_T = k\}$, el proceso N' es independiente de $\mathcal{F}_T$ y tiene por ley $\mathbb{P}_k$.

Demostración: Si $T = t$ es la propiedad de Markov simple. Supongamos ahora que T es un tiempo de paro *simple*, es decir que toma una cantidad numerable de valores $\{t_1, t_2, \dots\}$. Entonces para $\Lambda \in \mathcal{F}_T$ y $f : \mathbb{N}^n \to \mathbb{R}$ acotada,

$$\begin{aligned}\mathbb{E}[\mathbb{1}_\Lambda f(\widetilde{N}_{s_1}, \dots, \widetilde{N}_{s_n})] &= \sum_{k \geqslant 1} \mathbb{E}\Big[\mathbb{1}_{\Lambda \cap \{T = t_k\}} f(\widetilde{N}_{s_1}, \dots, \widetilde{N}_{s_n})\Big] \\ &= \sum_{k \geqslant 1} \mathbb{E}[\mathbb{1}_{\Lambda \cap \{T=t_k\}}]\mathbb{E}[f(N_{s_1}, \dots, N_{s_n})] \\ &= \mathbb{P}(\Lambda)\mathbb{E}[f(N_{s_1}, \dots, N_{s_n})],\end{aligned}$$

donde la segunda igualdad se da por la propiedad de Markov simple. Para el caso general consideramos a $T < \infty$ con probabilidad uno. Definamos para $n \geqslant 1$,

$$T^{(n)} = 2^{-n}\lceil 2^n T + 1\rceil.$$

Notemos que $T^{(n)}$ es una sucesión de tiempos de paro simples tal que $T^{(n)} \searrow T$ casi seguramente. Si $\Lambda \in \mathcal{F}_T$, *a fortiori* $\Lambda \in \mathcal{F}_{T^{(n)}}$. Si denotamos $\widetilde{N}_s^{(n)} := N_{s+T^{(n)}} - N_{T^{(n)}}$ para $s \geqslant 0$, entonces $\widetilde{N}_s^{(n)} \to \widetilde{N}_s$, conforme $n \to \infty$ gracias a que N tiene trayectorias càdlàg. Por lo tanto, para f acotada, tenemos

$$\mathbb{E}[f(\widetilde{N}_{s_1}^{(n)}, \ldots, \widetilde{N}_{s_k}^{(n)})\mathbb{1}_\Lambda] \xrightarrow[n\to\infty]{} \mathbb{E}[f(\widetilde{N}_{s_1}, \ldots, \widetilde{N}_{s_k})\mathbb{1}_\Lambda].$$

Por otra parte,

$$\mathbb{E}[f(\widetilde{N}_{s_1}^{(n)}, \ldots, \widetilde{N}_{s_k}^{(n)})\mathbb{1}_\Lambda] = \mathbb{P}(\Lambda)\mathbb{E}[f(N_{s_1}, \ldots, N_{s_k})].$$

Entonces,

$$\mathbb{E}[f(\widetilde{N}_{s_1}, \ldots, \widetilde{N}_{s_k})\mathbb{1}_\Lambda] = \mathbb{P}(\Lambda)\mathbb{E}[f(N_{s_1}, \ldots, N_{s_k})]. \qquad \blacksquare$$

3. *Observación:*
La prueba de la propiedad de Markov fuerte para N también se puede extender a procesos con incrementos independientes y estacionarios; *i.e.* los **procesos de Lévy**. En particular la satisface el movimiento browniano.

Vamos a definir al semigrupo del proceso Poisson P_t que actúa sobre funciones $f : \mathbb{N} \to \mathbb{R}$, medibles y acotadas, de la manera siguiente:

$$P_t f(x) := \mathbb{E}_x[f(N_t)] = \mathbb{E}[f(N_t + x)].$$

Notemos que P_t es un operador en $L_\infty(\mathbb{N})$. Además es un *operador de convolución*, esto es

$$P_t f(x) = \sum_{k\geqslant 0} f(k+x)e^{-ct}\frac{(ct)^k}{k!} = \int_{\mathbb{R}} f(x+y)\,\mathbb{P}(N_t \in \mathrm{d}y).$$

El operador P_t satisface la *propiedad de semigrupo*, es decir que para $s, t \geqslant 0$ se tiene

$$P_{t+s} = P_t P_s.$$

En efecto,

$$\begin{aligned} P_{t+s}f(x) &= \mathbb{E}_x[f(N_{t+s})] = \mathbb{E}_x[f(N_{t+s} - N_t + N_t)] \\ &= \mathbb{E}[P_s f(N_t + x)] = P_t P_s f(x). \end{aligned}$$

4.1. Martingalas asociadas a los procesos de Poisson

Sean $N = \{N_t : t \geqslant 0\}$ un proceso de Poisson de parámetro $c > 0$ y $\{\mathcal{F}_t\}_{t\geqslant 0}$ su filtración natural. Definamos, para $t \geqslant 0$,

$$\begin{aligned} M_t^{(1)} &:= N_t - ct, \\ M_t^{(2)} &:= (M_t^{(1)})^2 - ct \qquad \text{y} \\ \mathcal{E}_t^{(q)} &:= \exp\left\{-qN_t + c(1 - e^{-q})\right\}. \end{aligned}$$

4. *Observación:*
Nuevamente, bajo ciertas hipótesis de momentos, si reemplazamos a N por un proceso de Lévy, estas tres familias siguen siendo martingalas con respecto a su filtración natural.

19. Definición:
Decimos que $H = \{H_t : t \geqslant 0\}$ es *predecible* si es medible con respecto a la σ-álgebra generada por los procesos adaptados continuos por la izquierda.

5. *Observación:*
Todo proceso predecible puede aproximarse por procesos adaptados y continuos por la izquierda.

Ahora introduzcamos a la integral estocástica de un proceso predecible con respecto a un proceso Poisson. Para H predecible tenemos

$$\int_0^t H_s \,\mathrm{d}N_s = \sum_{s\leqslant t} H_s \Delta N_s,$$

donde $\Delta N_s = N_s - N_{s-}$. De la definición anterior tenemos la identidad

$$\int_0^t H_s \,\mathrm{d}N_s = \sum_{n\geqslant 0} H_{T(n)} \mathbb{1}_{\{T(n)\leqslant t\}}.$$

10. Proposición:
Sea H un proceso predecible tal que

$$\mathbb{E}\left[\int_0^t |H_s| \,\mathrm{d}s\right] < \infty$$

para todo $t \geqslant 0$. Entonces

$$\int_0^t H_s \,\mathrm{d}N_s - c\int_0^t H_s \,\mathrm{d}s, \quad t \geqslant 0,$$

es una martingala. Además se tiene la fórmula de compensación

$$\mathbb{E}\left[\int_0^t H_s \,\mathrm{d}N_s\right] = c\mathbb{E}\left[\int_0^t H_s \,\mathrm{d}s\right].$$

Demostración: Supongamos que H es un proceso simple, *i.e.*

$$H_s = \sum_{i=0}^{n-1} H_{t_i} \mathbb{1}_{(t_i,t_{i+1}]}(s),$$

donde $0 \leqslant t_0 < t_1 < \cdots < t_n = t$ y H_{t_i} es $\mathcal{F}_{t_i}$-medible. Por lo pronto supongamos que H es acotada. Entonces es claro que

$$\mathbb{E}\left[H_{t_i}\left[(N_s - cs) - (N_{t_i} - ct_i)\right] \,\middle|\, \mathcal{F}_{t_i}\right] = 0$$

para $s \in (t_i, t_{i+1}]$, lo cual implica $\int_0^\cdot H_s \, \mathrm{d}M_s^{(1)}$ es martingala.

Ahora supongamos que H es un proceso adaptado, continuo por la izquierda y acotado. Para $n \geqslant 1$, vamos a definir

$$H_s^{(n)} = H_{2^{-n}k}^{(n)} \quad \text{para} \quad k2^{-n} < s \leqslant (k+1)2^{-n},$$

de modo que $\{H^{(n)}\}$ es una sucesión de procesos simples acotados y continuos por la izquierda tal que $H_s^{(n)} \to H_s$, para toda $s \geqslant 0$, casi seguramente. Del teorema de convergencia dominada deducimos

$$I_t^{(n)} = \int_0^t H_s^{(n)} \, \mathrm{d}M_s^{(1)} \xrightarrow[n\to\infty]{} I_t = \int_0^t H_s \, \mathrm{d}M_s^{(1)}$$

con probabilidad uno. Dado que $I^{(n)}$ es martingala entonces I también lo es y el resultado se sigue para procesos adaptados, continuos por la izquierda y acotados.

Para extender e resultado a procesos predecibles y acotados vamos a usar el teorema de clases monótonas. Sea

$$\mathcal{A} := \left\{ H \text{ predecible y acotado y satisfacen que } \int_0^\cdot H_s \, \mathrm{d}M_s^{(1)} \text{ es martingala} \right\}.$$

Por lo ya realizado $\mathcal{A}$ contiene a los procesos adaptados y continuos por la izquierda (obviamente acotados). Además del teorema de convergencia monótona tenemos que si $\{H^{(n)}\}$ es una sucesión creciente en $\mathcal{A}$ que converge a H con H acotado entonces $H \in \mathcal{A}$. Por lo tanto el teorema de clases monótonas, vemos que $\mathcal{A}$ contiene a todos los procesos predecibles acotados.

Sin pérdida de generalidad supongamos H positiva y consideremos al siguiente tiempo de paro:

$$T_c := \text{ínf}\left\{ t \geqslant 0 : H_t \geqslant c \right\}.$$

Entonces

$$I_\cdot^c := \int_0^{\cdot \wedge T_c} H_s \, \mathrm{d}M_s^{(1)}$$

es martingala. Del Teorema de paro opcional se obtiene la fórmula de compensación. Por otro lado, cuando $c \to \infty$, $T_c \to \infty$, de modo que I_c converge, casi seguramente, a $\int_0^\cdot H_s \, \mathrm{d}M_s^{(1)}$. Dado que I^c es martingala, $\int_0^\cdot H_s \, \mathrm{d}M_s^{(1)}$ también lo es.

Finalmente probemos la fórmula de compensación. Del teorema de convergencia monótona,

$$\mathbb{E}\left[\int_0^{t\wedge T_c} H_s \, \mathrm{d}N_s \right] \xrightarrow[c\to\infty]{} \mathbb{E}\left[\int_0^t H_s \, \mathrm{d}N_s \right].$$

A su vez, por el teorema de convergencia dominada

$$\mathbb{E}\left[\int_0^{t\wedge T_c} H_s \, \mathrm{d}s \right] \xrightarrow[c\to\infty]{} \mathbb{E}\left[\int_0^t H_s \, \mathrm{d}s \right].$$

Por ende obtenemos

$$\mathbb{E}\left[\int_0^t H_s \, \mathrm{d}N_s \right] = c\mathbb{E}\left[\int_0^t H_s \, \mathrm{d}s \right]. \qquad \blacksquare$$

8 de marzo

11. Proposición:
Sea h una función medible y positiva, entonces el proceso definido por

$$\exp\left\{-\int_0^t h(s)\,\mathrm{d}N_s + c\int_0^t (1-e^{-h(s)})\,\mathrm{d}s\right\}, \quad t \geqslant 0,$$

es una $\{\mathcal{F}_t\}_{t\geqslant 0}$-martingala. En particular tenemos

$$\mathbb{E}\left[\exp\left\{-\int_0^t h(s)\,\mathrm{d}N_s\right\}\right] = \exp\left\{-x\int_0^t (1-e^{-h(s)})\,\mathrm{d}s\right\}.$$

Demostración: Supongamos que h es simple. Bajo esta hipótesis, el resultado es claro gracias a la martingala exponencial $\mathcal{E}_t^{(q)}$, pues bastará tomar a q como el valor de h en un paso. Para extender dicho resultado al caso en que h es medible y positiva basta usar el teorema de convergencia monótona. ■

6. *Observación:*
Este resultado *claramente*[a] se puede extender a procesos predecibles y acotados.

[a] ¿Sí es claro?

20. Definición:
Sea $\{\mathcal{F}_t\}_{t\geqslant 0}$ una filtración. Un $\{\mathcal{F}_t\}_{t\geqslant 0}$-*proceso de Poisson* con parámetro $c > 0$ es un proceso N adaptado y continuo por la derecha tal que

- $N_0 = 0$ casi seguramente y
- para cualesquiera $t > s \geqslant 0$ y $k \in \mathbb{N}_0$,

$$\mathbb{P}(N_t - N_s = k \mid \mathcal{F}_s) = c^k \frac{(t-s)^k}{k!} \exp\{-c(t-s)\}.$$

12. Proposición:
Sean N y N' dos $\{\mathcal{F}_t\}_{t\geqslant 0}$-procesos de Poisson. Los procesos N y N' son independientes si y sólo si no saltan simultáneamente; i.e.

$$N_t - N_{t-} = 0 \quad o \quad N'_t - N'_{t-} = 0 \quad \textit{para toda } t \textit{ casi seguramente.}$$

7. *Observación:*
Es importante suponer en la proposición anterior que N y N' son adaptados a la misma filtración, ya que el resultado es falso en otro caso. En efecto, basta tomar $N'_t = N_{2t}$ como contraejemplo.

Para probar este resultado haremos uso del siguiente lema.

8. Lema:
Sean M una martingala càdlàg de variación acotada y M' una martingala càdlàg acotada en la misma filtración. Si M y M' no *saltan simultáneamente entonces MM' es martingala.*

Demostración: Para probar este resultado basta ver que para todo tiempo de paro T acotado con probabilidad uno, se satisface que

$$\mathbb{E}[M_T M'_T] = \mathbb{E}[M_0 M'_0].$$

Sea

$$0 = t_0 < t_1 < \cdots < t_k = A$$

una partición del intervalo $[0, A]$ con $T \leqslant A$ casi seguramente. Entonces

$$M_T M'_T - M_0 M'_0 = \sum_{t_i < T} M_{t_i}(M'_{t_{i+1}} - M'_{t_i}) + \sum_{t_i < T} M'_{t_i}(M_{t_{i+1}} - M_{t_i}) + \sum_{t_i < T} (M'_{t_{i+1}} - M'_{t_i})(M_{t_{i+1}} - M_{t_i}).$$

Al tomar esperanzas en ambos lados, vemos lo siguiente:

$$\begin{aligned}\mathbb{E}[M_T M'_T - M_0 M'_0] &= \mathbb{E}\left[\sum_{t_i < T} M_{t_i}(M'_{t_{i+1}} - M'_{t_i})\right] + \mathbb{E}\left[\sum_{t_i < T} M'_{t_i}(M_{t_{i+1}} - M_{t_i})\right] \\ &\quad + \mathbb{E}\left[\sum_{t_i < T} (M'_{t_{i+1}} - M'_{t_i})(M_{t_{i+1}} - M_{t_i})\right] \\ &= \mathbb{E}\left[\sum_{t_i < T} (M'_{t_{i+1}} - M'_{t_i})(M_{t_{i+1}} - M_{t_i})\right],\end{aligned}$$

con la última igualdad dándose gracias a

$$\mathbb{E}[M_{t_i}(M'_{t_{i+1}} - M'_{t_i})] = \mathbb{E}[M_{t_i}\mathbb{E}[M'_{t_{i+1}} - M'_{t_i} \mid \mathcal{F}_{t_i}]] = 0.$$

Recordemos que M' es acotada y M es de variación finita, de modo que

$$\begin{aligned}\left|\sum_{t_i < T} (M'_{t_{i+1}} - M'_{t_i})(M_{t_{i+1}} - M_{t_i})\right| &\leqslant C \sum_{t_i < T} |M_{t_{i+1}} - M_{t_i}| \\ &\leqslant C V_0^A(M) < C'',\end{aligned}$$

donde $V_0^A(M)$ es la variación de M en $[0, A]$. Cuando la norma de la partición se va a 0, vemos que

$$\sum_{t_i < T} (M'_{t_{i+1}} - M'_{t_i})(M_{t_{i+1}} - M_{t_i}) \to \sum_{s < T} (\Delta M'_s)(\Delta M_s) = 0.$$

Entonces del teorema de convergencia dominada se obtiene el resultado. ■

Demostración de la proposición 12: Primero vamos a suponer que $N \perp\!\!\!\perp N'$. Sea $\{T(n)\}_{n\geqslant 1}$ los tiempos de salto del proceso N. Entonces, con probabilidad uno,

$$\sum_{s>0} (\Delta N_s)(\Delta N'_s) = \sum_{n\geqslant 1} \Delta N'_{T(n)}.$$

Por otro lado, como la let de los saltos no tiene átomos es claro que a cada tiempo fijo t, $\Delta N'_t = 0$ casi seguramente. De la independencia de N' y $\{T(n)\}_{n\geqslant 1}$, vemos que $\Delta N'_{T(n)} = 0$ casi seguramente, lo cual sucede para cada $n \geqslant 1$. Consecuentemente los procesos no saltan de forma simultánea.

En sentido inverso, tomamos h y h' dos funciones simples y definamos las martingales exponenciales

$$M_t := \exp\left\{-\int_0^t h(s)\,\mathrm{d}N_s + c\int_0^t (1-e^{-h(s)})\,\mathrm{d}s\right\}$$

y

$$M'_t := \exp\left\{-\int_0^t h'(s)\,\mathrm{d}N'_s + c'\int_0^t (1-e^{-h'(s)})\,\mathrm{d}s\right\},$$

ambas definidas sobre la misma filtración. Como N y N' no saltan simultáneamente, entonces M y M' tampoco lo hacen. Aplicando el lema 8, MM' es una martingala, y por consiguiente

$$\mathbb{E}[M_t M'_t] = 1 \quad \text{para cualquier } t \geqslant 0,$$

lo que implica

$$\begin{aligned}
&\mathbb{E}\left[\exp\left\{-\int_0^t h(t)\,\mathrm{d}N_s - \int_0^t h'(s)\,\mathrm{d}N'_s\right\}\right] \\
&\quad = \exp\left\{-c\int_0^t (1-e^{-h(s)})\,\mathrm{d}s\right\}\exp\left\{-c'\int_0^t (1-e^{-h'(s)})\,\mathrm{d}s\right\} \\
&\quad = \mathbb{E}\left[\exp\left\{-\int_0^t h(t)\,\mathrm{d}N_s\right\}\right]\mathbb{E}\left[\exp\left\{-\int_0^t h'(t)\,\mathrm{d}N'_s\right\}\right].
\end{aligned}$$

■

4.2. Medidas aleatorias de Poisson y procesos puntuales

21. Definición:
Sea $(E, \mathcal{E}, \mu)$ un espacio medible, μ una medida σ-finita. Una *medida aleatoria de Poisson* es una colección de variables aleatorias

$$M = \{M(A) : A \in \mathcal{E}\}$$

definidas en un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ tales que:

(i) Si $B \in \mathcal{E}$ y tal que $\mu(B) < \infty$ entonces $M(B) \sim \text{Poisson}(\mu(B))$. Si $\mu(B) = \infty$ entonces $M(B) = \infty$ casi seguramente.

(ii) Si $B_0, B_1, \ldots, B_n$ es una sucesión finita de conjuntos mutuamente ajenos de $\mathcal{E}$

entonces $M(B_0), M(B_1), \ldots, M(B_n)$ son independientes.

(Ejercicio) Tomemos $E = \mathbb{R}_+$, $\mathcal{E} = \mathcal{B}(\mathbb{R}_+)$ y $\mu = c\lambda$. Probar que $t \mapsto M((0,t])$ define a un proceso de Poisson de parámetro c.

Propiedades:

Superposición: Sean $\{\mu_n\}_{n\geqslant 1}$ una sucesión de medidas σ-finitas y $\mu = \sum_i \mu_i$. Si el resultado es una medida σ-finita y $M^{(n)}$ es la medida aleatoria de Poisson con intensidad μ_n, $M = \sum_i M^{(i)}$ es una medida aleatoria de Poisson con intensidad μ.

Divisibilidad: Sea M una medida aleatoria de Poisson con intensidad μ y $\{B_i\}_{i\geqslant 1}$ una sucesión de conjuntos mutuamente disjuntos de $\mathcal{E}$. Entonces las restricciones $\{M\restriction_{B_n}\}_{n\geqslant 1}$ son medidas aleatorias de Poisson con intensidades $\{\mu(\cdot \cap B_n)\}_{n\geqslant 1}$.

Imagen: Sean $f : (E,\mathcal{E}) \to (G,\mathcal{G})$ una función medible, μ una medida σ-finita en $(E,\mathcal{E})$ y γ la medida imagen de μ con respecto a f, esto es $\gamma = \mu \circ f^{-1}$. Supongamos que γ es una medida σ-finita. Si M es una medida aleatoria de Poisson en $(E,\mathcal{E})$ con intensidad μ y si definimos a

$$M \circ f^{-1}(C) = M(f^{-1}(C)) \quad \text{para } C \in \mathcal{G},$$

entonces obtenemos una medida aleatoria de Poisson con intensidad γ.

10 de marzo

Existencia (Construcción)

Supongamos que $\mu(E) < \infty$ y definamos la medida de proba

$$\rho(B) = \frac{\mu(B)}{\mu(E)} \quad \text{para } B \in \mathcal{E}.$$

Sean $\{\xi_n\}_{n\geqslant 1}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución común ρ y N una variable aleatoria Poisson con parámetro $\mu(E)$, la cual es independiente de $\{\xi_n\}_{n\geqslant 1}$.

Consideremos a la medida aleatoria

$$M(\mathrm{d}x) = \sum_{i=1}^{N} \delta_{\xi_i}(\mathrm{d}x),$$

donde δ_x es la *medida de Dirac centrada en x*. Se tiene que M es una medida de conteo, *i.e.*

$$M(B) = \#\{i \leqslant N : \xi_i \in B\}.$$

Veremos que M es una medida aleatoria de Poisson de parámetro μ.

En el caso σ-finito se puede construir una medida aleatoria de Poisson usando la propiedad de superposición. En otras palabras, vamos a tomar una partición $\{B_n\}_{n\geqslant 1}$ de E, de tal forma que cada elemento de la partición cumple que $\mu(B_n) < \infty$. Después construiremos, como en el caso finito, una sucesión independiente de medidas aleatorias de Poisson cuyas intensidades están dadas por las restricciones de μ en cada una de las B_n. Finalmente la propiedad de superposición nos da el resultado.

Sin pérdida de generalidad consideraremos $\mu(E) = 1$. Primero calculemos la distribución de M. Para $B \in \mathcal{E}$,

$$\begin{aligned}\mathbb{P}(M(B) = k) &= \sum_{j=k}^{\infty} \mathbb{P}\left(\sum_{i=1}^{N} \delta_{\xi_i}(B) = k \,\middle|\, N = j\right)\mathbb{P}(N = j)\\ &= \sum_{j=k}^{\infty} \binom{j}{k} \mu(B)^k (1-\mu(B))^{j-k} \frac{e^{-1}}{j!}\\ &= \frac{e^{-1}\mu(B)^k}{k!} \sum_{j=k}^{\infty} \frac{(1-\mu(B))^{j-k}}{(j-k)!} = \frac{e^{-\mu(B)}\mu(B)^k}{k!}.\end{aligned}$$

Ahora probemos la propiedad de independencia. Consideremos a B y B' dos conjuntos disjuntos de $\mathcal{E}$. Definamos al proceso

$$X_t := \xi_{N_t}, \quad t \geqslant 0,$$

donde N es un proceso Poisson de intensidad $\mu(E) = 1$, el cual es independiente de la sucesión $\{\xi_n\}_{n\geqslant 1}$. Consideremos a $\{\mathcal{F}_t\}_{t\geqslant 0}$ la filtración generada por $\{X_t : t \geqslant 0\}$. Tomemos a los procesos de conteo

$$N_t^B := \#\{i \leqslant N_t : \xi_i \in B\} \quad \text{y} \quad N_t^{B'} := \#\{i \leqslant N_t : \xi_i \in B'\}.$$

Los dos procesos N^B y $N^{B'}$ son procesos de conteo de Poisson con parámetro $\mu(B)$ y $\mu(B')$ respectivamente en la filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$. Como B y B' son disjuntos, no saltan simultáneamente y por ende son independientes. Puesto que

$$N_1^B = (B) \quad \text{y} \quad N_1^{B'} = M(B')$$

tenemos que $M(B)$ y $M(B')$ son independientes.

Ahora vamos a considerar la integral de $f : E \to \mathbb{R}_+$, medible, con respecto a M. Primero supongamos que f es simple, es decir

$$f = \sum_{i=1}^{\infty} c_i \mathbb{1}_{B_i},$$

donde $c_i \geqslant 0$ y $\{B_i\}_{i\geqslant 1}$ es una partición medible de E. Entonces es claro que

$$\langle M, f\rangle := \int_E f \, M(\mathrm{d}x) = \sum_{i=1}^{\infty} c_i M(B_i).$$

Para poder definir la integral con respecto a cualquier función positiva y medible, haremos uso del método de aproximación. Considerando

$$f_n = 2^{-n}\lfloor 2^n f \rfloor,$$

es claro que f_n converge uniformemente y por debajo a f y al aplicar el teorema de convergencia monótona, vemos que

$$\langle M, f\rangle = \lim_{n\to\infty} \langle M, f_n\rangle = \int_E f(x)\, M(\mathrm{d}x).$$

13. Proposición (Fórmula de Campbell):
Sean $f : E \to \mathbb{R}_+$ una función medible y M una medida aleatoria de Poisson con intensidad μ. Consideremos a la integral de f con respecto a M definida por $\langle M, f\rangle = \int_E f(x)\, M(\mathrm{d}x)$. Entonces

$$\mathbb{E}[\exp\{-\lambda\langle M,f\rangle\}] = \exp\left\{-\int_E (1-e^{-\lambda f(x)})\,\mu(\mathrm{d}x)\right\}.$$

Demostración: Supongamos $\mu(E) < \infty$. De la construcción de medidas aleatorias de Poisson, tenemos

$$\langle M,f\rangle = \sum_{i=1}^{N} f(\xi_i),$$

donde $\{\xi_i\}_{i\geqslant 1}$ es una sucesión de variables aleatorias independientes e idénticamente distribuidas con ley común $\rho(\cdot) = \mu(\cdot)/\mu(E)$ y N es una variable aleatoria Poisson con parámetro $\mu(E)$ e independiente de $\{\xi_i\}_{i\geqslant 1}$. Entonces

$$\begin{aligned}
\mathbb{E}[\exp\{-\lambda\langle M,f\rangle\}] &= \mathbb{E}\left[\exp\left\{-\lambda\sum_{i=1}^{N} f(\xi_i)\right\}\right]\\
&= \sum_{k=0}^{\infty}\frac{\mu(E)^k e^{-\mu(E)}}{k!}\mathbb{E}\left[\exp\left\{-\lambda\sum_{i=1}^{k} f(\xi_i)\right\}\right]\\
&= \sum_{k=0}^{\infty}\frac{\mu(E)^k e^{-\mu(E)}}{k!}\left(\int_E e^{-\lambda f(x)}\,\frac{\mu(\mathrm{d}x)}{\mu(E)}\right)^k\\
&= e^{-\mu(E)}\sum_{k=0}^{\infty}\frac{1}{k!}\left(\int_E e^{-\lambda f(x)}\,\mu(\mathrm{d}x)\right)^k\\
&= \exp\left\{\int_E (e^{-\lambda f(x)}-1)\,\mu(\mathrm{d}x)\right\}.
\end{aligned}$$

Ahora consideremos el caso $\mu(E) = \infty$. Sea $\{B_n\}_{n\geqslant 1}$ una partición de E tal que para cada $n \geqslant 1$ se cumple $\mu(B_n) < \infty$. Después, para cada $n \geqslant 1$ definamos $E_n = \bigcup_{k\leqslant n} B_k$ y la restricción $M^{(n)} = M\!\restriction_{E_n}$. Sea $f \geqslant 0$ medible, entonces

$$\langle M^{(n)},f\rangle = \int_E f(x)\mathbb{1}_{E_n}(x)\,M(\mathrm{d}x),$$

la cual satisface la fórmula de Campbell. Como $E_n \nearrow E$, $f\mathbb{1}_{E_n}$ y $(1-e^{-\lambda f})\mathbb{1}_{E_n}$ crecen a f y $1-e^{-\lambda f}$ de forma respectiva. Nuevamente aplicando el teorema de convergencia monótona, tenemos

$$\langle M,f\rangle = \lim_{n\to\infty}\langle M^{(n)},f\rangle \quad \text{y} \quad \int_E (1-e^{-\lambda f(x)})\,\mu(\mathrm{d}x) = \lim_{n\to\infty}\int_E (1-e^{-\lambda f(x)})\mathbb{1}_{E_n}(x)\,\mu(\mathrm{d}x),$$

lo que implica la fórmula de Campbell en el caso general. ■

Ahora vamos a considerar a una medida de Poisson en $[0,\infty)\times E$ con intensidad $\lambda\otimes\mu$, donde λ es la medida de Lebesgue y μ una medida σ-finita en E.

9. Lema:
Casi seguramente para toda $t \geqslant 0$, $M(\{t\} \times E) \in \{0,1\}$.

Demostración: Supongamos $\mu(E) < \infty$. Para $n \geqslant 1$, por la propiedad de estacionariedad e independencia de M, tenemos

$$\mathbb{P}\Big(\text{Existe } k \leqslant 2^n \text{ tal que } M([(k-1)2^{-n}, k2^{-n}) \times E) \geqslant 2\Big) \leqslant 2^n \mathbb{P}\Big(M([0,2^{-n}) \times E) \geqslant 2\Big).$$

Por otro lado,

$$\mathbb{P}\Big(M([0,2^{-n}) \times E) \geqslant 2\Big) = 1 - e^{-2^{-n}\mu(E)} - 2^{-n}\mu(E)e^{-2^{-n}\mu(E)} \leqslant 2(2^{-n}\mu(E))^2,$$

lo cual implica

$$\mathbb{P}\Big(\text{Existe } k \leqslant 2^n \text{ tal que } M([(k-1)2^{-n}, k2^{-n}) \times E) \geqslant 2\Big) \leqslant 2^{-n+1}\mu(E)^2.$$

El resultado se obtiene al hacer tender $n \to \infty$.

Caso $\mu(E) = \infty$. Usaremos E_n como se definieron en la prueba anterior. Definamos a

$$f_n(x) = \mathbb{1}_{\{t\}\times E}(x)$$

y notemos que casi seguramente, para toda $t \geqslant 0$,

$$\langle M, f_n \rangle = M(\{t\} \times E_n) \leqslant 1.$$

Por otro lado, la sucesión $\{f_n\}$ es creciente y converge a $f = \mathbb{1}_{\{t\}\times E}$. Gracias al teorema de convergencia monótona, tenemos

$$M(\{t\} \times E) = \langle M, f \rangle = \lim_{n\to\infty} \langle M, f_n \rangle \leqslant 1.$$

Dicha propiedad se cumple casi seguramente para toda $t \geqslant 0$ ya que

$$A = \Big\{\omega \in \Omega : t \geqslant 0, \langle M, f \rangle \leqslant 1\Big\}$$

satisface

$$A = \bigcup_{n\geqslant 1} \Big\{\omega \in \Omega : \text{para toda } t \geqslant 0, \langle M, f_n \rangle \leqslant 1\Big\}$$

y cada uno de los conjunto en la unión tiene probabilidad uno. ∎

15 de marzo

Si $M(\{t\} \times E) = 1$ entonces existe un único punto $\Delta_t \in E$ tal que

$$M\restriction_{\{t\}\times E} = \delta_{(t,\Delta_t)}.$$

Si, en cambio, $M(\{t\} \times E) = 0$ entonces no definimos al punto Δ_t en E.

22. Definición:
El proceso definido por $\Delta = \{\Delta_t : t \geqslant 0\}$ es un *proceso puntual de Poisson* con medida característica μ.

10. Lema:
Sea $B \in \mathcal{E}$ tal que $0 < \mu(B) < \infty$ y definamos

$$T_B = \operatorname{ínf}\{t \geqslant 0 : \Delta_t \in B\}.$$

Entonces T_B y Δ_{T_B} son variables aleatorias independientes. La variable aleatoria T_B se distribuye como una variable aleatoria exponencial de parámetro $\mu(B)$ y la distribución de Δ_{T_B} está dada por $\mu(\cdot \cap B)/\mu(B)$.

Demostración: Sea $A \subset B$. Por un lado es claro que

$$\mathbb{P}(T_B \leqslant t, \Delta_{T_B} \in A) = \mathbb{P}(T_A \leqslant T_{B\setminus A}, T_A \wedge T_{B\setminus A} \leqslant t).$$

Por otro lado, la variable aleatoria T_A es el primer tiempo de salto del proceso de Poisson N^A definido por

$$N_t^A = \#\{s \leqslant t : \Delta_s \in A\},$$

de modo que T_A debe ser exponencial de parámetro $\mu(A)$. Como A y $B \setminus A$ son disjuntos, entonces T_A y $T_{B\setminus A}$ son variable aleatorias exponenciales independientes de parámetro $\mu(A)$ y $\mu(B) - \mu(A)$ respectivamente. Luego,

$$\begin{aligned}\mathbb{P}(T_B \leqslant t, \Delta_{T_B} \in A) &= \mathbb{E}[\mathbb{1}_{\{T_A \leqslant t\}} e^{-\mu(B\setminus A)T_A}] \\ &= \int_0^t \mu(A) e^{-\mu(B)s}\, \mathrm{d}s = \frac{\mu(A)}{\mu(B)}(1 - e^{-\mu(B)t}).\end{aligned}$$ ∎

5. Subordinadores

La noción de subordinador es una extensión natural de los procesos puntuales de Poisson y de manera general se definen como sigue.

23. Definición:
Un *subordinador* es un proceso estocástico que toma valores en $\mathbb{R}_+$ con trayectorias càdlàg y con incrementos independientes y estacionarios.

Es importante señalar que los subordinadores *tienen trayectorias no decrecientes*. Además estaremos interesados en la noción de subordinador matado. En otras palabras, sea $\sigma = \{\sigma_t : t \geqslant 0\}$ un subordinador y e_q una variable aleatoria exponencial de parámetro $q > 0$ independiente de σ, el proceso $\sigma^{(q)}$ toma valores en $[0, \infty]$ y se encuentra definido por

$$\sigma_t^{(q)} = \begin{cases} \sigma_t & \text{si } t < e_q, \\ \infty & \text{si } t \geqslant e_q. \end{cases}$$

Al proceso $\sigma^{(q)}$ se le conoce como *subordinador matado con tasa de muerte* q.

De hecho cualquier proceso continuo por la derecha con trayectorias no decrecientes $X = \{X_t : t \geqslant 0\}$ que toma valores en $[0,\infty]$ es un subordinador matado con tasa q si y solamente si

$$\mathbb{P}(X_t < \infty) = e^{-qt}$$

y bajo el evento $\{X_t < \infty\}$, el incremento $X_{t+s} - X_t$ es independiente de $\sigma(X_u, u \leqslant t)$ y tiene la misma ley que X_s.

Ejemplos:

(i) El proceso Poisson.

(ii) Tiempos de pasada del movimiento browniano, *i.e.*

$$\sigma_x = \text{ínf}\{t \geqslant 0 : B_t = x\}, \quad x \geqslant 0.$$

Usando la descomposición

$$\sigma_n = \sigma_1 + \sum_{i=1}^{n} (\sigma_i - \sigma_{i-1})$$

y la independencia y homogeneidad de los incrementos de σ, observamos

$$\mathbb{E}[e^{-q\sigma_n}] = \mathbb{E}[e^{-q\sigma_1}]^n, \quad q \geqslant 0.$$

Si $t = m/n$ es racional, con m, n primos relativos, entonces

$$\mathbb{E}[e^{-q\sigma_t}] = \mathbb{E}[e^{-q\sigma_{1/n}}]^m = \mathbb{E}[e^{-q\sigma_1}]^{m/n} = \mathbb{E}[e^{-q\sigma_1}]^t.$$

Gracias a que σ tiene trayectorias càdlàg, y que para $t \in \mathbb{R}$ existe una sucesión $\{s_n\}$ de racionales que converge a t, tenemos

$$\mathbb{E}[e^{-q\sigma_t}] = \lim_{n\to\infty} \mathbb{E}[e^{-q\sigma_{s_n}}] = \lim_{n\to\infty} \mathbb{E}[e^{-q\sigma_1}]^{s_n} = \mathbb{E}[e^{-q\sigma_1}]^t.$$

Ahora, si escribimos

$$\mathbb{E}[e^{-q\sigma_1}] = e^{-\varphi(q)}$$

para $q \geqslant 0$ con $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ llamada *el exponente de Laplace*, entonces tenemos

$$\mathbb{E}[e^{-q\sigma_t}] = e^{-t\varphi(q)}, \quad q \geqslant 0, t \geqslant 0.$$

Una pregunta natural es *¿cuáles funciones φ son exponentes de Laplace de un subordinador?*

15. Teorema (de Finetti, Lévy, Khintchine)**:**

(i) *Si φ es el exponente de Laplace de un subordinador σ entonces existe una única pareja (k, d) de reales no negativos y una medida π con soporte en $(0,\infty)$ que cumple la condición de integrabilidad*

$$\int_{(0,\infty)} (1 \wedge x)\, \pi(\mathrm{d}x) < \infty,$$

tal que para toda $\lambda \geqslant 0$,

$$\varphi(\lambda) = k + d\lambda + \int_{(0,\infty)} (1 - e^{-\lambda x})\, \pi(\mathrm{d}x). \tag{$\star$}$$

(ii) *Recíprocamente cualquier función φ que puede expresarse como* $(\star)$ *es el exponente de Laplace de un subordinador.*

Demostración del teorema 15: Para $t \geqslant 0$ sabemos que $\mathbb{E}[e^{-q\sigma_t}] = e^{-t\varphi(q)}$. Notemos que $\varphi(q)$ lo podemos escribir como

$$\varphi(q) = \lim_{n\to\infty} n\Big(1 - e^{-\varphi(q)/n}\Big) \quad \text{y} \quad 1 - e^{-\varphi(q)/n} = \mathbb{E}[1 - e^{-q\sigma_{1/n}}].$$

Ahora definamos

$$\overline{F}_n(x) = n\mathbb{P}(\sigma_{1/n} > x), \quad x \geqslant 0,$$

y observemos que

$$\frac{\varphi(q)}{q} = \lim_{n\to\infty} \int_{[0,\infty]} e^{-qx}\overline{F}_n(x)\, \mathrm{d}x.$$

De la igualdad anterior tenemos

$$\frac{\varphi(q+1)}{q+1} \times \frac{1}{\varphi(1)} = \lim_{n\to\infty} \frac{1}{n(1 - e^{-\varphi(1)/n})} \int_{[0,\infty]} e^{-qx}\overline{F}_n(x)\, \mathrm{d}x,$$

por lo que no es muy difícil de ver que

$$\mu_n(\mathrm{d}x) = \frac{\overline{F}_n(x)}{n(1 - e^{-\varphi(1)/n})}\, \mathrm{d}x$$

define una sucesión de medidas de probabilidad en $[0,\infty]$ cuyas transformadas de Laplace convergen; es decir

$$\lim_{n\to\infty} \mathcal{L}_{\mu_n}(q) = \frac{\varphi(q+1)}{q+1} \times \frac{1}{\varphi(1)}.$$

Además $\lim_{n\to\infty} \mathcal{L}_{\mu_n}(0) = 1$ y el lado derecho de la igualdad es continua por la derecha. Por consiguiente, del teorema de Lévy, $\{\mu_n\}$ converge débilmente a una medida de probabilidad en $[0,\infty]$, la cual vamos a denotar por Λ. En particular para f continua y acotada,

$$\int_0^\infty f(s)\overline{F}_n(s)\, \mathrm{d}s \xrightarrow[n\to\infty]{} \varphi(1) \int_0^\infty f(s)e^s\, \Lambda(\mathrm{d}s),$$

y por ende $\{\overline{F}_n(s)\, \mathrm{d}s\}$ converge vagamente conforme $n \to \infty$ hacia la medida

$$\overline{\Lambda}(\mathrm{d}x) = \varphi(1)e^x \Lambda(\mathrm{d}x).$$

Ahora definamos

$$G_n(x) = \int_0^x \overline{F}_n(s)\ \mathrm{d}s, \quad x \geqslant 0.$$

Observemos que para $n \geqslant 1$, G_n tiene densidad decreciente, por lo que G_n tiene que ser cóncava. En otras palabras, $\overline{\Lambda}([0,x])$ también debe ser cóncava. Esto implica que debe tener una densidad decreciente en $(0,\infty)$ y en 0 puede tener un átomo. En otras palabras, el límite de $\{\overline{F}_n(x)\ \mathrm{d}x\}$ es de la forma

$$\Lambda(\mathrm{d}s) = d\delta_0(\mathrm{d}s) + \overline{\pi}(s)\ \mathrm{d}s,$$

donde $d \geqslant 0$ y $\overline{\pi}$ es una función decreciente en $(0,\infty)$. En otras palabras,

$$\frac{\varphi(q)}{q} = d + \int_{(0,\infty)} e^{-qx}\overline{\pi}(x)\ \mathrm{d}x.$$

Al integrar por partes obtenemos el resultado con

$$k = \overline{\pi}(\infty) \quad \text{y} \quad \pi(\mathrm{d}x) = -\mathrm{d}\overline{\pi}(x) \quad \text{en} \quad (0,\infty).$$

Finalmente, probemos que $\int_0^\infty (1\wedge x)\,\pi(\mathrm{d}x) < \infty$. En este sentido vemos que

$$\int_0^\infty (1\wedge x)\,\pi(\mathrm{d}x) = \int_0^1 \overline{\pi}(x)\ \mathrm{d}x < \infty.$$

Hay algunos detallitos en la prueba y falta concluirla. □

17 de marzo

Continuación de la demostración del teorema 15: Por la convergencia mostrada previamente

$$\int_0^1 \overline{F}_n(x)\ \mathrm{d}x \to d + \int_0^1 \overline{\pi}(x)\ \mathrm{d}x.$$

Procedamos con el inciso (ii). Tomemos una medida π en $(0,\infty)$ que satisface

$$\int_{(0,\infty)} (1\wedge x)\,\pi(\mathrm{d}x) < \infty,$$

y construyamos a un proceso de Poisson puntual $\{\Delta_t : t \geqslant 0\}$ con medida característica $\pi + \kappa\delta_\infty$. Ahora definamos al tiempo aleatorio

$$T_\infty = \inf\{t \geqslant 0 : \Delta_t = \infty\}$$

y para toda $t \leqslant T_\infty$ definiremos a

$$\Sigma_t = \sum_{s\leqslant t} \Delta_s.$$

De la fórmula de Campbell es claro

$$\begin{aligned}\mathbb{E}[\exp\{-q\Sigma_t\}] &= \exp\left\{-t\int_{(0,\infty]}(1-e^{-qx})\,\mu(\mathrm{d}x)\right\}\\ &= \exp\left\{-t\kappa - t\int_{(0,\infty)}(1-e^{-qx})\,\pi(\mathrm{d}x)\right\}.\end{aligned}$$

De nuestros supuesto se tiene que $\int_{(0,\infty)}(1-e^{-qx})\,\pi(\mathrm{d}x)<\infty$. Como

$$\lim_{q\to 0}\int_{(0,\infty)}(1-e^{-qx})\,\pi(\mathrm{d}x)=0,$$

tenemos que $\Sigma_t<\infty$ casi seguramente para $t<T_\infty$. Por lo tanto, como Δ es un proceso de Poisson puntual, el proceso Σ es càdlàg por definición con tiempo de vida T_∞ y sus incrementos son independientes estacionarios en $[0,T_\infty)$. Entonces Σ es un subordinador. Para el caso $d>0$ tomamos $\Sigma_t^{(d)}=dt+\Sigma_t$, el cual no es muy difícil ver que es un subordinador matado con exponente de Laplace

$$\varphi_{\Sigma^{(d)}}(\lambda)=d\lambda+\varphi_\Sigma(d).$$ ■

La prueba de la fórmula de Lévy–Khintchine nos da una interpretación probabilística de los subordinadores.

4. Corolario (Descomposición de Lévy–Khintchine):
Casi seguramente tenemos que para $t\geqslant 0$,

$$\sigma_t=dt+\sum_{s\leqslant t}\Delta_s,$$

donde $\Delta=\{\Delta_t:t\geqslant 0\}$ es un proceso de Poisson puntual en $[0,\infty]$ con medida característica $\pi+k\delta_\infty$. El tiempo de vida de σ está dado por $\xi=\inf\{t\geqslant 0:\Delta_t=\infty\}$.

Recordemos que en la sección 3.4 introducimos a las leyes estables en $(0,\infty)$ con índice $\alpha\in(0,1)$. Para ser más precisos, una variable aleatoria estable $\sigma(\alpha)$ positiva y con índice $\alpha\in(0,1)$ satisface la relación

$$\mathbb{E}[e^{-q\sigma(\alpha)}]=\exp\{-q^\alpha\},\quad q\geqslant 0.$$

Por otro lado, observemos que

$$q^{\alpha-1}=\frac{1}{\Gamma(\alpha-1)}\int_0^\infty e^{-qx}x^{-\alpha}\,\mathrm{d}x,$$

por lo que un procedimiento por integración por partes nos da

$$q^\alpha=\frac{\alpha}{\Gamma(1-\alpha)}\int_0^\infty(1-e^{-qx})\frac{\mathrm{d}x}{x^{1+\alpha}}.$$

Del teorema 15 deducimos que q^α es el exponente de Laplace de un subordinador sin deriva y con medida de Lévy

$$\frac{\alpha}{\Gamma(1-\alpha)}\frac{\mathrm{d}x}{x^{1+\alpha}},$$

al cual denotaremos por $\sigma^{(\alpha)} = \{\sigma_t^{(\alpha)} : t \geqslant 0\}$ y llamaremos subordinador estable de índice $\alpha \in (0,1)$. A este tipo de funciones φ se les conoce como *funciones de Bernstein*.

Otro ejemplo **importante** es el *subordinador gamma* $\{\sigma_t : t \geqslant 0\}$, cuyo exponente de Laplace está dado por

$$\varphi(q) = a\log\Big(1+\frac{q}{b}\Big).$$

En otras palabras, tenemos

$$\mathbb{E}[e^{-q\sigma_t}] = \Big(1+\frac{q}{b}\Big)^{-at},$$

la cual es la transformada de Laplace de una variable aleatoria gamma con parámetros (at, b), de modo que

$$\mathbb{P}(\sigma_t \in \mathrm{d}x) = \frac{bat}{\Gamma(at)} x^{t\alpha-1} e^{bx}\, \mathrm{d}x.$$

Por otro lado, haciendo uso de la *integral de Frullani*, tenemos

$$\varphi(q) = a\log\Big(1+\frac{q}{b}\Big) = a\int_0^\infty (1-e^{-qx})\frac{e^{-bx}}{x}\, \mathrm{d}x,$$

por lo que el subordinador gamma no tiene deriva y su medida de Lévy es tal que

$$\pi(\mathrm{d}x) = a\frac{e^{-bx}}{x}\, \mathrm{d}x.$$

5.1. Propiedad de Markov fuerte y de Feller

Un subordinador es un proceso de Markov fuerte y esto se sigue de los mismos argumentos usados para probar dicha propiedad para el proceso de Poisson. En sí, esto es heredado de la propiedad de incrementos independientes y estacionarios. Para enunciar la propiedad de Markov fuerte, primero vamos a introducir para $x \geqslant 0$, la ley $\mathbb{P}_x$ bajo la cual el subordinador $\{\sigma_t : t \geqslant 0\}$ empieza en x. Además vamos a considerar a τ un tiempo de paro con respecto a la filtración natural de σ, $\mathcal{F}_t = \sigma(\sigma_s, s \leqslant t)$. Si el subordinador σ empieza en x entonces este tiene la misma ley que $\sigma + x$ en $\mathbb{P}_0 =: \mathbb{P}$.

16. Teorema (Propiedad de Markov fuerte):
Supongamos que $\tau < \infty$ casi seguramente. Definamos a $\sigma' = \{\sigma'_t : t \geqslant 0\}$, donde

$$\sigma'_t = \sigma_{\tau+t} - \sigma_\tau$$

para $t \geqslant 0$. Entonces σ' es independiente de $\mathcal{F}_\tau$ y tiene la misma ley que σ. Además si tomamos $\widetilde{\sigma} = \{\widetilde{\sigma}_t : t \geqslant 0\}$ con $\widetilde{\sigma}_t = \sigma_{\tau+t}$, condicionado con respecto a $\{\sigma_\tau = x\}$ es independiente de $\mathcal{F}_\tau$ y tiene por ley $\mathbb{P}_x$.

Como en el caso del proceso de Poisson tenemos un semigrupo de convolución asociado a un subordinador $\sigma = \{\sigma_t : t \geqslant 0\}$, el cual está definido para f boreliana y positiva como sigue:

$$P_t f(x) = \mathbb{E}_x[f(\sigma_t)] = \mathbb{E}[f(\sigma_t + x)].$$

Recordemos que $\{P_t : t \geqslant 0\}$ es de convolución por la propiedad de Markov, es decir

$$P_{t+s} f(x) = P_t P_s f(x), \quad t, s \geqslant 0.$$

Como veremos en el siguiente resultado, el semigrupo $\{P_t : t \geqslant 0\}$ satisface la *propiedad de Feller*, es decir que satisface las propiedades:

(i) $P_t \mathcal{C}_0 \subset \mathcal{C}_0$, donde $\mathcal{C}_0$ son las funciones continuas en $[0, \infty)$ y cuyo límite en ∞ es igual a 0.

(ii) Para $f \in \mathcal{C}_0$ se cumple que

$$\lim_{t \to 0} P_t f(x) = f(x), \quad \text{para toda } x \geqslant 0.$$

14. Proposición:
El semigrupo $\{P_t : t \geqslant 0\}$ asociado a un subordinador σ es de Feller.

Demostración: Notemos que $P_t f(x) = \mathbb{E}[f(\sigma_t + x)]$ y además $\sigma_t \to 0$ conforme $t \to 0$ casi seguramente con respecto a $\mathbb{P}$. Por convergencia dominada tenemos $P_t f(x) \to f(x)$ conforme $t \to 0$. La continuidad de $x \mapsto P_t f(x)$ y la propiedad

$$\lim_{x \to \infty} P_t f(x) = 0$$

se sigue del mismo argumento. ∎

5.2. Ley débil de los grandes números

22 de marzo

Ley débil de los grandes números para subordinadores

15. Proposición:
Sea $\sigma = \{\sigma_t : t \geqslant 0\}$ un subordinador con coeficiente de deriva $d \geqslant 0$, entonces

$$\frac{\sigma_t}{t} \xrightarrow[t \searrow 0]{} d \quad \text{en probabilidad.}$$

Demostración: Recordemos que

$$\mathbb{E}[e^{-q\sigma_t}] = \exp\{-t\varphi(q)\},$$

con φ una función de Bernstein, por lo que

$$\mathbb{E}[e^{-q\sigma_t/t}] = \exp\{-t\varphi(q/t)\}.$$

De la fórmula de Lévy–Khintchine, tenemos

$$t\varphi\Big(\frac{q}{t}\Big) = dq + t\int_{(0,\infty)} (1 - e^{-qx/t})\,\pi(\mathrm{d}x).$$

Como $1 - e^{-ax} \leqslant 1 \wedge x$ para $a \geqslant 0$ y $x \geqslant 0$, en particular observamos $t(1 - e^{-qx/t}) \leqslant 1 \wedge x$ para $t \leqslant 1$. Entonces del teorema de convergencia dominada obtenemos

$$\mathbb{E}[e^{-q\sigma_t/t}] \xrightarrow[t\searrow 0]{} \exp\{-dq\},$$

lo cual implica el resultado. ■

16. Proposición:
Sea $\sigma = \{\sigma_t : t \geqslant 0\}$ un subordinador con coeficiente de deriva $d \geqslant 0$ y tal que

$$\int_{(0,\infty)} x\,\pi(\mathrm{d}x) < \infty,$$

entonces

$$\frac{\sigma_t}{t} \xrightarrow[t\to\infty]{} d + \int_{(0,\infty)} x\,\pi(\mathrm{d}x) \quad \textit{en probabilidad.}$$

Demostración: De tarea. ■

La segunda propiedad que veremos relaciona la distribución de un subordinador con su medida de Lévy.

17. Proposición:
Sea $\overline{\pi}(x) = \pi(x,\infty)$, entonces para toda $x > 0$ tal que $\overline{\pi}$ es continua en x, es decir que x no es un átomo de π, tenemos

$$\overline{\pi}(x) = \lim_{\varepsilon\searrow 0} \frac{1}{\varepsilon}\mathbb{P}(\sigma_\varepsilon > x).$$

Demostración: Recordemos que

$$\frac{\varphi(q)}{q} = d + \int_0^\infty e^{-qx}\overline{\pi}(x)\,\mathrm{d}x.$$

Definamos a

$$\mu(\mathrm{d}x) = d\delta_0(\mathrm{d}x) + \overline{\pi}(x)\,\mathrm{d}x,$$

de modo que

$$\frac{\varphi(q)}{q} = \int_{(0,\infty)} e^{-qx}\,\mu(\mathrm{d}x).$$

Por otro lado,

$$\begin{aligned}\frac{1}{\varepsilon}\int_0^\infty e^{-qx}\mathbb{P}(\sigma_\varepsilon > x)\,\mathrm{d}x &= \frac{1}{\varepsilon}\int_0^\infty e^{-qx}\int_{(x,\infty)}\mathbb{P}(\sigma_\varepsilon \in \mathrm{d}z)\,\mathrm{d}x\\ &= \frac{1}{\varepsilon q}\int_{(0,\infty)}\mathbb{P}(\sigma_\varepsilon \in \mathrm{d}x)(1-e^{-qx})\\ &= \frac{1}{\varepsilon q}(1-e^{-\varepsilon\varphi(q)}).\end{aligned}$$

En otras palabras, deducimos

$$\frac{\varphi(q)}{q} = \int_{(0,\infty)} e^{-qx}\,\mu(\mathrm{d}x) = \lim_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\int_0^\infty e^{-qx}\mathbb{P}(\sigma_\varepsilon > x)\,\mathrm{d}x.$$

Esto implica la siguiente convergencia débil:

$$\frac{1}{\varepsilon}\mathbb{P}(\sigma_\varepsilon > x)\,\mathrm{d}x \xrightarrow[\varepsilon\to 0]{\mathrm{w}} d\delta_0(\mathrm{d}x) + \overline{\pi}(x)\,\mathrm{d}x.$$

Entonces para toda $0 < a < b$ se tiene

$$\begin{aligned}\frac{1}{\varepsilon}(b-a)\mathbb{P}(\sigma_\varepsilon > a) &\geqslant \frac{1}{\varepsilon}\int_a^b \mathbb{P}(\sigma_\varepsilon > x)\,\mathrm{d}x\\ &\xrightarrow[\varepsilon\searrow 0]{} \int_a^b \overline{\pi}(x)\,\mathrm{d}x \geqslant (b-a)\overline{\pi}(b).\end{aligned}$$

En consecuencia se deduce que

$$\liminf_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{P}(\sigma_\varepsilon > b) \geqslant \overline{\pi}(b).$$

De manera similar se puede probar

$$\limsup_{\varepsilon\searrow 0}\frac{1}{\varepsilon}\mathbb{P}(\sigma_\varepsilon > b) \leqslant \overline{\pi}(a).$$

Si x no es un átomo de π, llegamos al resultado deseado tomando en la primera desigualdad $a = x$ y b de tal forma que se acerque a a, mientras que en la segunda se toma $b = x$ y a creciente a b. ■

24. Definición:
Decimos que un semigrupo $\{P_t : t \geqslant 0\}$ es de *Feller fuerte* si para toda función medible y acotada f, se tiene que $P_t f$ es continua y acotada para cada $t > 0$.

18. Proposición:
Sea $\{P_t : t \geqslant 0\}$ el semigrupo asociado al subordinador σ. Entonces $\{P_t : t \geqslant 0\}$ es de Feller fuerte precisamente cuando para toda $t \geqslant 0$ la ley de σ_t es absolutamente continua con respecto a la medida de Lebesgue.

Demostración: Primero probemos la suficiencia. Consideremos $\Lambda \in \mathcal{B}(\mathbb{R}_+)$ de medida de Lebesgue cero, y observemos

$$\begin{aligned}\int_0^\infty P_t \mathbb{1}_\Lambda(x)\ \mathrm{d}x = \int_0^\infty \mathbb{E}[\mathbb{1}_\Lambda(x+\sigma_t)]\ \mathrm{d}x = \mathbb{E}\left[\int_0^\infty \mathbb{1}_\Lambda(x+\sigma_t)\ \mathrm{d}x\right] \\ = \mathbb{E}\left[\int_0^\infty \mathbb{1}_{\Lambda'}(x)\ \mathrm{d}x\right] = 0,\end{aligned}$$

donde $\Lambda' = \{x - \sigma_t : x \in \Lambda\}$. De esta última igualdad deducimos $P_t\mathbb{1}_\Lambda = 0$ casi dondequiera con respecto a la medida de Lebesgue. De la propiedad de Feller fuerte, $P_t\mathbb{1}_\Lambda \equiv 0$. Esto implica

$$\mathbb{E}[\mathbb{1}_\Lambda(\sigma_t)] = 0,$$

por lo que la densidad de σ_t es absolutamente continua con respecto a la medida de Lebesgue.

Continuemos con la necesidad de la afirmación. Consideremos p_t como la densidad de σ_t, es decir $\mathbb{P}(\sigma_t \in \mathrm{d}x) = p_t(x)\ \mathrm{d}x$. Lo que deseamos probar es que para toda f boreliana y acotada,

$$P_t f(x) = \mathbb{E}[f(\sigma_t + x)] = \int_0^\infty f(x+y)p_t(y)\ \mathrm{d}y \xrightarrow[x\to x_0]{} \int_0^\infty f(x_0+y)p_t(y)\ \mathrm{d}y = P_t f(x_0).$$

Por simplicidad notacional supondremos $x_0 = 0$. Notemos que

$$\begin{aligned}&\int_0^\infty f(x+y)p_t(y)\ \mathrm{d}y \\ &= \int_0^a f(x+y)p_t(y)\ \mathrm{d}y + \int_a^b f(x+y)p_t(y)\ \mathrm{d}y + \int_b^\infty f(x+y)p_t(y)\ \mathrm{d}y.\end{aligned}$$

Como f es acotada, vemos

$$\int_0^a f(x+y)p_t(y)\ \mathrm{d}y \leqslant \|f\|_\infty \int_0^a p_t(y)\ \mathrm{d}y \xrightarrow[a\searrow 0]{} 0,$$

y de manera similar

$$\int_b^\infty f(x+y)p_t(y)\ \mathrm{d}y \leqslant \|f\|_\infty \int_b^\infty p_t(y)\ \mathrm{d}y \xrightarrow[b\nearrow\infty]{} 0.$$

Por otro lado, es claro que $f \in L^2([a,b], \mathrm{d}x)$. Definamos $f_x(y) = f(x+y)$ y notemos que la transformada de Fourier satisface que

$$\widehat{f_x}(\lambda) = e^{-\mathrm{i}x\lambda}\widehat{f}(\lambda) \xrightarrow[x\searrow 0]{L^2([a,b],\mathrm{d}x)} \widehat{f}(\lambda)$$

gracias al Teorema de Plancherel. En particular

$$\|\widehat{f}\|_{L^2([a,b],\mathrm{d}x)} = \|f\|_{L^2([a,b],\mathrm{d}x)}.$$

Por Hölder tenemos

$$\int_a^b |f(xy)-f(y)|p_t(y)\ \mathrm{d}y \leqslant \left(\int_a^b p_t(y)^2\ \mathrm{d}y\right)^{1/2} \|f_x - f\|_{L^2([a,b],\mathrm{d}x)},$$

cota que converge a 0 conforme $x \to 0$. En otras palabras tenemos

$$\begin{aligned}
&\int_0^\infty |f(x+y)-f(y)|p_t(y)\ \mathrm{d}y \\
&\leqslant 2\|f\|_\infty \int_{\mathbb{R}_+\setminus(a,b)} p_t(y)\ \mathrm{d}y + \int_a^b |f(x+y)-f(y)|p_t(y)\ \mathrm{d}y \\
&\xrightarrow[x\searrow 0]{} 2\|f\|_\infty \int_{\mathbb{R}_+\setminus(a,b)} p_t(y)\ \mathrm{d}y,
\end{aligned}$$

de donde se obtiene el resultado al hacer tender $a \searrow 0$ y $b \nearrow \infty$. ■

6. Teoría de procesos de Markov

24 de marzo

De manera informal vamos a introducir la propiedad de Markov como sigue. Sea $(E,\mathcal{E})$ un espacio medible al cual llamaremos espacio de estados en el cual toma valores un proceso estocástico $X = \{X_t : t \geqslant 0\}$ y consideremos una filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$ bajo la cual el proceso X es adaptado. (En general se usará la filtración canónica.) Diremos que el proceso estocástico X es *markoviano* bajo la filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$ si *el futuro es independiente del pasado condicionado al presente*. En otras palabras, si para $s,t \geqslant 0$, $f : E \to \mathbb{R}$ medible y acotada entonces

$$\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_t] = H(s,X_t)$$

se escribe como una función medible de X_t (y debe depender de s).

6.1. Semigrupos y probabilidades de transición

Primero veamos el concepto de probabilidades de transición.

25. Definición:
Sea $K : E \times \mathcal{E} \to [0,1]$. Decimos que K es *probabilidad de transición* si satisface las siguientes condiciones:

(i) para $x \in E$, $K(x,\cdot)$ es una medida de probabilidad en E;
(ii) para $B \in \mathcal{E}$, $x \mapsto K(x,B)$ es medible.

Observemos que si $f : E \to \mathbb{R}$ es medible y acotada entonces

$$x \mapsto \int f(y)\, K(x, \mathrm{d}y)$$

es aún medible. Esta última propiedad nos garantiza que la composición de probabilidades de transición vuelve a ser una probabilidad de transición. En efecto, consideremos K y $\overline{K}$ dos probabilidades de transición. Entonces podemos definir $M = K \circ \overline{K}$ como sigue:

$$M(x, B) = \int_E K(x, \mathrm{d}y)\overline{K}(y, B) \quad \text{para } x \in E \text{ y } B \in \mathcal{E}.$$

26. Definición:
Sea $\{K_t : t \geqslant 0\}$ una familia de probabilidades de transición. Decimos que es *semigrupo* si satisface $t, s \geqslant 0$ se tiene que

$$K_{t+s} = K_s \circ K_t. \tag{$\star$}$$

A la igualdad ($\star$) se le conoce como la *ecuación de Chapman–Kolmogorov*.

Nosotros solo trabajaremos con semigrupos homogéneos. No obstante, en la literatura podemos encontrar semigrupos inhomogéneos, es decir

$$K_{s,t} \circ K_{t,u} = K_{s,u}, \quad s \leqslant t \leqslant u.$$

27. Definición:
Sean $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geqslant 0}, \mathbb{P})$ un espacio de probabilidad filtrado y $\{X_t : t \geqslant 0\}$ un proceso adaptado. Se dice que X es *markoviano de semigrupo* $\{P_t\}_{t\geqslant 0}$ si para toda $s, t \geqslant 0$ y $f : E \to \mathbb{R}$ medible y acotada se tiene

$$\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_t] = P_s f(X_t).$$

Ejemplos:

(i) Semigrupo de Poisson:

$$P_t(x, \mathrm{d}y) = \mathbb{P}(N_t + x \in \mathrm{d}y).$$

(ii) Semigrupo del calor (movimiento browniano) para $d \geqslant 1$:

$$P_t(x, \mathrm{d}y) = \frac{1}{(2\pi t)^{d/2}} \exp\left\{-\frac{|x-y|^2}{2t}\right\} \mathrm{d}y, \quad x, y \in \mathbb{R}^d.$$

(iii) Semigrupo de Cauchy:

$$P_t(x, \mathrm{d}y) = \frac{1}{\pi(t^2 + (x-y)^2)} \mathrm{d}y, \quad x, y \in \mathbb{R}.$$

Este semigrupo se puede construir vía un movimiento browniano en $\mathbb{R}^2$. Sea (X, Y) un movimiento browniano en $\mathbb{R}^2$ y definamos

$$T_t = \text{ínf}\{s \geqslant 0 : X_s \geqslant t\}.$$

Entonces el proceso $\{Y_{T_t} : t \geqslant 0\}$ tiene asociado al semigrupo de Cauchy.

(iv) Semigrupo asociado a un subordinador es markoviano.

11. Lema:
Si X es markoviano con semigrupo $\{P_t : t \geqslant 0\}$ entonces para una partición

$$0 \leqslant t_0 < t_1 < \cdots < t_n \quad y \quad f_0, \ldots, f_n \text{ medibles y acotadas}$$

tenemos

$$\begin{aligned}
&\mathbb{E}[f_0(X_{t_0})f_1(X_{t_1})\cdots f_n(X_{t_n})] \\
&= \int_E f(x_0)\mathbb{P}(X_{t_0} \in \mathrm{d}x_0) \times \int_E f_1(x_1)P_{t_1-t_0}(x_0, \mathrm{d}x_1) \times \cdots \times \int_E f_1(x_1)P_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n).
\end{aligned} \tag{$\star\star$}$$

Recíprocamente, si tenemos ($\star\star$), entonces X es markoviano de semigrupo $\{P_t : t \geqslant 0\}$ con respecto a su filtración canónica.

8. *Observación:*
Otra forma de expresar ($\star\star$) es la siguiente:

$$\begin{aligned}
&\mathbb{P}(X_{t_0} \in \mathrm{d}x_0, X_{t_1} \in \mathrm{d}x_1, \ldots, X_{t_n} \in \mathrm{d}x_n) \\
&\quad = \mathbb{P}(X_{t_0} \in \mathrm{d}x_0)P_{t_1-t_0}(x_0, \mathrm{d}x_1)\cdots P_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n).
\end{aligned}$$

Demostración del lema 11: La necesidad se sigue por un argumento inductivo. Para $0 \leqslant t_0 < t_1 < \cdots < t_n$ y $f_0, \ldots, f_n$ medibles y acotadas, se tiene que

$$\begin{aligned}
&\mathbb{E}[f_0(X_{t_0})f_1(X_{t_1})\cdots f_n(X_{t_n})] \\
&\quad = \mathbb{E}\Big[\mathbb{E}[f_0(X_{t_0})f_1(X_{t_1})\cdots f_n(X_{t_n}) \mid \mathcal{F}_{t_{n-1}}]\Big] \\
&\quad = \mathbb{E}[f_0(X_{t_0})f_1(X_{t_1})\cdots f_n(X_{t_{n-1}})P_{t_n-t_{n-1}}f_n(X_{t_{n-1}})] \\
&\quad = \mathbb{E}[f_0(X_0)P_{t_1-t_0}f_1(X_{t_0})\cdots P_{t_n-t_{n-1}}f_n(X_{t_{n-1}})].
\end{aligned}$$

Para el recíproco usaremos un argumento de clases monótonas. Consideremos al conjunto

$$\mathcal{H} = \Big\{Y \in L^\infty(\mathbb{P}, \mathcal{F}_t) : \mathbb{E}[Y f(X_{t+s})] = \mathbb{E}[Y P_s f(X_t)]\Big\},$$

donde $s, t \geqslant 0$ y $f \in L^\infty(E)$. Consideremos a la clase $\mathcal{C}$ de variables aleatorias de la forma

$$f_0(X_{t_0})\cdots f_m(X_{t_m}) \quad \text{para } 0 \leqslant t_0 < t_1 < \cdots < t_m \leqslant t \text{ y } f_i \in L^\infty(E).$$

Por hipótesis es claro que $\mathcal{C} \subset \mathcal{H}$ y es un espacio vectorial estable bajo la convergencia monótona acotada, esto es si $\{Y_n\}_{n\geqslant 1} \subset \mathcal{H}$ es una sucesión tal que $Y_n \leqslant Y_{n+1} < C$, donde C es constante, entonces

$$Y_\infty = \lim_{n\to\infty} Y_n \in \mathcal{H}$$

gracias al teorema de convergencia dominada. Entonces del teorema de clases monótonas funciones, $L^\infty(\mathbb{P}, \mathcal{F}_t) \subset \mathcal{H}$, de modo que

$$\mathbb{E}[f(X_{t+s}) \mid \mathcal{F}_t] = P_s f(X_s) \quad \text{casi seguramente.}$$

■

6.2. Construcción de procesos de Markov

29 de marzo

En esta sección vamos a suponer que el espacio de estados E es un espacio de Lusin y sea $\mathcal{E}$ la σ-álgebra de Borel asociada a E.

19. Proposición:
Sea $\{P_t : t \geqslant 0\}$ un semigrupo de transición markoviano definido en $(E, \mathcal{E})$, entonces para $x \in E$ existe una única medida de probabilidad $\mathbb{P}_x$ en $(E^{[0,\infty)}, \mathcal{E}^{[0,\infty)})$ tal que bajo $\mathbb{P}_x$, el proceso canónico X definido por $X_t(\omega) = \omega(t)$ es markoviano de semigrupo $\{P_t : t \geqslant 0\}$ y $\mathbb{P}_x(X_0 = x) = 1$.

Recordemos que $E^{[0,\infty)}$ es el espacio de todas las aplicaciones que van de $[0,\infty)$ a E y $\mathcal{E}^{[0,\infty)}$ es la σ-álgebra generada por los cilíndros de $E^{[0,\infty)}$.

Demostración: La prueba de esta proposición es una simple aplicación del teorema de Daniell–Kolmogorov. Primero vamos a definir a $\mathbb{P}_x$ bajo la σ-álgebra generada por los cilindros $\mathcal{E}^{[0,\infty)}$ como sigue: para $0 \leqslant t_0 < t_1 < \cdots < t_n$ definamos

$$\mathbb{P}_x(X_{t_0} \in \mathrm{d}x_0, \ldots, X_{t_n} \in \mathrm{d}x_n) = P_{t_0}(x, \mathrm{d}x_0) P_{t_1 - t_0}(x_0, \mathrm{d}x_1) \cdots P_{t_n - t_{n-1}}(x_{n-1}, \mathrm{d}x_n).$$

La propiedad de semigrupo garantiza la propiedad de consistencia. Por lo tanto, el teorema 4 nos da la existencia y unicidad de $\mathbb{P}_x$. ■

De manera más general, sea ν una medida de probabilidad en E. Si $Z \in L^\infty(\mathcal{F}_\infty)$, la aplicación $x \mapsto \mathbb{E}_x[Z]$ es medible. En efecto, este es el caso para variables aleatorias Z de la forma $f_0(X_{t_0}) f_1(X_{t_1}) \cdots f_n(X_{t_n})$ para $0 \leqslant t_0 < t_1 < \cdots < t_n$ y $f_i \in L^\infty(E)$ con $i \in [[0, n]]$. Esto se deduce del hecho de que $x \mapsto P_t f(x)$ es medible y un procedimiento inductivo. Las variables aleatorias de este tipo forman una clase que denotaremos por $\mathcal{C}$, la cual es estable bajo la operación producto y que contiene a las constantes. Por otro lado, la clase

$$\mathcal{H} = \{Z \in L^\infty(\mathcal{F}_\infty) : x \mapsto \mathbb{E}_x[Z] \text{ es medible}\}$$

es un espacio vectorial estable bajo la convergencia monótona acotada y que contiene a la clase $\mathcal{C}$. Por lo tanto, este espacio $\mathcal{H}$ contiene a todas las variables aleatorias acotadas y medibles bajo la σ-álgebra generada por la clase $\mathcal{C}$. Por lo tanto podemos definir una nueva medida de probabilidad $\mathbb{P}_\nu$ en $(E^{[0,\infty)}, \mathcal{E}^{[0,\infty)})$ como sigue:

$$\mathbb{E}_\nu[Z] = \int_E \mathbb{E}_x[Z]\, \nu(\mathrm{d}x).$$

20. Proposición (Propiedad de Markov simple)**:**
Sea $t \geqslant 0$, definamos a $X'_s = X_{t+s}$ para $s \geqslant 0$. Entonces, condicionando bajo el evento $\{X_t = x\}$, el proceso X' es independiente de $\mathcal{F}_t$ y tiene por ley a $\mathbb{P}_x$. En otras palabras para $F : E^{[0,t]} \to \mathbb{R}$ y $G : E^{[0,\infty)} \to \mathbb{R}$ medibles y acotadas, para toda ley inicial ν se tiene

$$\mathbb{E}_\nu[F(X_u, u \in [0,t])G(X'_u, u \geqslant 0)] = \mathbb{E}_\nu\Big[F(X_u, u \in [0,t])\mathbb{E}_{X_t}[G(X_u, u \geqslant 0)]\Big].$$

Demostración: Sin pérdida de generalidad podemos tomar $\nu = \delta_x$. Primero condicionamos con respecto a $\mathcal{F}_t$ y vamos a suponer que G es de la forma

$$G(X'_u, u \geqslant 0) = f_0(X'_{t_0}) \cdots f_n(X'_{t_n}),$$

donde $0 \leqslant t_0 < t_1 < \cdots < t_n$ y $f_i \in L^\infty(E)$. Después aplicamos las fórmulas obtenidas anteriormente y concluimos con un argumento de clases monótonas. ■

Hasta ahora hemos supuesto que $P_t(x, E) = 1$, pero existen casos muy interesantes en donde $P_t(x, E) < 1$ para algunas $x \in E$ y $t \geqslant 0$. En el primer caso, llamaremos al semigrupo $\{P_t : t \geqslant 0\}$ *markoviano* y en el segundo *submarkoviano*.

Para entender mejor el caso submarkoviano, pensemos en un proceso de Markov como el modelo de una partícula que se mueve aleatoriamente. Entonces el caso submarkoviano corresponde a que la partícula desaparezca o muera en un tiempo finito. Hay un método que nos permite convertir el caso submarkoviano en markoviano. Al espacio E le vamos a añadir un punto Δ, el cual llamaremos *punto cementerio* y vamos a denotar $E^\Delta = E \cup \{\Delta\}$. En adición introduciremos a $\mathcal{E}^\Delta = \sigma\{\mathcal{E}, \Delta\}$. El punto Δ es aislado. Ahora vamos a definir a un nuevo semigrupo $\overline{P}$ en $(E^\Delta, \mathcal{E}^\Delta)$, donde

$$\overline{P}_t(x, A) = \begin{cases} P_t(x, A) & \text{si } x \in E \text{ y } A \subset E, \\ 1 - P_t(x, E) & \text{si } x \in E \text{ y } A = \{\Delta\}, \\ 1 & \text{si } x = \Delta \text{ y } A = \{\Delta\}. \end{cases}$$

Generalmente no vamos a distinguir entre P y $\overline{P}$ y en la mayoría de los casos en que estamos interesados Δ es un estado absorbente. Por convención todas las funciones f en E serán extendidas a E^Δ con $f(\Delta) = 0$. Entonces la propiedad de Markov se debe establecer como sigue:

$$\mathbb{E}_\nu[Z \circ \theta_t \mid \mathcal{F}_t] = \mathbb{E}_{X_t}[Z], \quad \mathbb{P}_\nu\text{-c.s. en el conjunto } \{X_t \neq \Delta\},$$

ya que el lado derecho toda el valor cero en $\{X_t = \Delta\}$ y, en cambio, el lado izquierdo puede no ser 0.

6.3. Procesos de Hunt

31 de marzo

Vamos a suponer la medida inicial $\nu = \delta_0$, donde 0 es un punto específico en E. Además vamos a suponer que E es polaco, *i.e.* métrico, completo y separable. Vamos a denotar por d a la métrica.

Sea $X = \{X_t : t \geqslant 0\}$ un proceso estocástico en E con trayectorias continuas por la derecha y tal que $\mathbb{P}_0(X_0 = 0) = 1$. Consideremos a $\{\mathcal{F}_t\}_{t\geqslant 0}$ una filtración continua por la derecha y completa, *i.e.* que satisface las condiciones *usuales/habituales*. También supondremos que X es adaptado a la filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$.

28. Definición (Propiedad de Markov fuerte)**:**
Para todo tiempo de para τ finito $\mathbb{P}_0$-casi seguramente, bajo la ley $\mathbb{P}_0(\cdot \mid X_\tau = x)$, el proceso X' definido por $X'_t = X_{\tau+t}$ es independiente de $\mathcal{F}_\tau$ y tiene por ley $\mathbb{P}_x$.

29. Definición (Procesos de Hunt)**:**
Se dice que un proceso de Markov fuerte es *cuasi-continuo por la izquierda* si para toda sucesión $\{T_n : n \geqslant 1\}$ creciente de tiempos de paro con respecto $\{\mathcal{F}_t\}_{t\geqslant 0}$ que convergen a un tiempo de paro T finito $\mathbb{P}_0$-c.s., se tiene

$$X_{T_n} \xrightarrow[n\to\infty]{} X_T \quad \text{casi seguramente.}$$

La clase de estos procesos se les conoce como *procesos de Hunt*.

21. Proposición:
Sea B un conjunto abierto o cerrado de E y X un proceso de Hunt. Entonces

$$T_B = \inf\{t \geqslant 0 : X_t \in B\}$$

es un $\{\mathcal{F}_t\}$-tiempo de paro.

Demostración: Sale gracias al **??**. ■

17. Teorema:
Sea X un proceso de Hunt, entonces, con probabilidad uno, las trayectorias de X tienen límites por la izquierda en $(0,\infty)$.

Demostración: Sea $\varepsilon > 0$ fijo. Definamos de manera recursiva a la siguiente familia de tiempos de paro: $T_0 = 0$, $T_1 = T^{(\varepsilon)} = \inf\{t \geqslant 0 : d(X_t, X_0) > \varepsilon\}$ y

$$T_{n+1} = \inf\{t > T_n : d(X_t, X_{T_n}) > \varepsilon\}.$$

De la proposición 21 sabemos que $T^{(\varepsilon)}$ es tiempo de paro y además $T_{n+1} = T_n + T_1 \circ \theta_{T_n}$. La sucesión $\{T_n : n \geqslant 1\}$ es creciente y por ende el límite $\tau^{(\varepsilon)} = \lim_n T_n$ también es tiempo de paro. Gracias a la cuasi-continuidad por la izquierda de X, bajo el evento $\{\tau^{(\varepsilon)} < \infty\}$, tenemos $X_{T_n} \to X_{\tau^{(\varepsilon)}}$. Por la continuidad por la derecha de X deducimos $d(X_{T_{n+1}}, X_{T_n}) > \varepsilon$ para $n \geqslant 1$. Ambas afirmaciones nos garantizan que $\tau^{(\varepsilon)} = \infty$. En este caso $\mathbb{R}_+ = \bigcup_{n\geqslant 0}[T_n, T_{n+1})$. Notemos que si $\{T_n = \infty\}$, $[T_n, T_{n+1}) = \emptyset$. En caso contrario, en $[T_n, T_{n+1})$, el proceso no oscila con una amplitud mayor que 2ε. Entonces hemos probado que para cada $\varepsilon > 0$ existe $\Omega_\varepsilon \subset \Omega$ con $\mathbb{P}(\Omega_\varepsilon) = 1$ y tal que X no oscila con amplitud mayor a 2ε en $[T_n^{(\varepsilon)}, T_{n+1}^{(\varepsilon)})$.

Ahora sea $\widetilde{\Omega} = \bigcap_{m\geqslant 1} \Omega_{1/m}$, el cual resulta ser un conjunto de probabilidad uno. Con ello afirmamos que para $\omega \in \widetilde{\Omega}$f, la aplicación $t \mapsto X_t(\omega)$ debe tener límites por la izquierda en $(0, \infty)$. De no ser el caso, existirán $t \in (0, \infty)$ y $m \geqslant 1$ tal que $X(\omega)$ oscila con una amplitud mayor a $2/m$ en $(t-\delta, t)$ para toda $\delta > 0$, lo cual implicaría $t \notin [T_n^{(1/m)}, T_{n+1}^{(1/m)})$ para toda $n \geqslant 0$. Esto último es imposible por la partición de $\mathbb{R}_+$ dada previamente. ■

6.4. Procesos de Feller

De ahora en adelante vamos a suponer que E es un espacio métrico localmente compacto. Vamos a denotar por ∞ al punto cementerio, permitiéndonos compactificar el espacio, y en adición tomaremos a $\mathcal{C}_0$ como el espacio de las funciones definidas en $\overline{E} = E \cup \{\infty\}$, continuas y que se anulan en el ∞. Recordemos que el espacio $\mathcal{C}_0$ es un espacio métrico completo bajo la distancia uniforme.

30. Definición:
Llamaremos *semigrupo de Feller* a una familia de operadores lineales $\{T_t : t \geqslant 0\}$ en $\mathcal{C}_0$; es decir $T_t : \mathcal{C}_0 \to \mathcal{C}_0$ tal que

(i) T_t es positivo, esto es $T_t f \geqslant 0$ siempre que $f \geqslant 0$;
(ii) si $f_n \nearrow 1$ entonces $T f_n \nearrow 1$;
(iii) $T_t \circ T_s = T_{s+t}$ para $s, t \geqslant 0$;
(iv) si $T_0 = \mathrm{id}$, donde id es el operador identidad;
(v) para $f \in \mathcal{C}_0$, $T_t f \to f$ conforme $t \to 0$.

Un semigrupo de Feller se puede extender a un semigrupo markoviano. En efecto, si $\{P_t : t \geqslant 0\}$ es felleriano, entonces $P_t : \mathcal{C}_0 \to \mathcal{C}_0$ es positivo y para $x \in E$, $f(x) \mapsto P_t f(x)$ es una forma lineal, continua y positiva. Del Teorema de representación de Riesz, tenemos que existe una medida $P_t(x, \cdot)$ tal que

$$P_t f(x) = \int_E f(y)\, P_t(x, \mathrm{d}y).$$

Además como para $P_t f_n \nearrow 1$ siempre que $f_n \nearrow 1$, $P_t(x, \cdot)$ es una proba. Definamos por extensión a $P_t f(x)$ para $f \in L^\infty(E)$. Sabemos que $x \mapsto P_t f(x)$ es continua y $f \in \mathcal{C}_0$ y por ende medible. Por un argumento de clases monótonas, $x \mapsto P_t f(x)$ es medible para toda $f \in L^\infty(E)$.

Acabamos de extender a $\{P_t : t \geqslant 0\}$ a un semigrupo markoviano en $L^\infty(E)$. La propiedad de semigrupo la tenemos en $\mathcal{C}_0$ y también se puede extender a $L^\infty(E)$.

22. Proposición:
Si $X = \{X_t : t \geqslant 0\}$ es un proceso de Markov con semigrupo felleriano y si $\mathbb{P}_x$ denota la ley de X iniciando en x, entonces para $\{x_n\}_{n\geqslant 1}$ una sucesión que converge a x, tenemos que $\mathbb{P}_{x_n}$ converge a $\mathbb{P}_x$ en el sentido de distribuciones finito dimensionales.

Demostración: Para $f \in \mathcal{C}_0$ tenemos $P_t f(x) = \mathbb{E}_x[f(X_t)]$. Como $P_t f \in \mathcal{C}_0$ tenemos que conforme $x_n \to x$, se satisface

$$\mathbb{E}_{x_n}[f(X_t)] = P_t f(x_n) \to P_t f(x) = \mathbb{E}_x[f(X_t)].$$

El caso general es una aplicación de Chapman–Kolmogorov y recursión. Para ilustrar el procedimiento, observemos que para $t < s$ y $f, g \in \mathcal{C}_0$,

$$\mathbb{E}_x[f(X_t)g(X_s)] = \mathbb{E}_x[f(X_t)P_{t-s}g(X_t)] = \mathbb{E}_x[h(X_t)],$$

donde $h(x) = f(x)P_{s-t}g(x)$. De aquí el resultado es evidente. ■

18. Teorema:
Para toda ley de probabilidad γ en E, existe un proceso estocástico $\widetilde{X}$ tal que para toda t, $\mathbb{P}_\gamma(X_t = \widetilde{X}_t) = 1$ y

$$\mathbb{P}_\gamma(t \mapsto \widetilde{X}_t \text{ tiene trayectorias càdlàg}) = 1.$$

Demostración: Ver páginas 91–92 de Revuz y Yor (2005) o ver capítulo III.2 de Rogers y col. (1994). ■

Este teorema permite ver a $\mathbb{P}_x$ o $\mathbb{P}_\gamma$ como una medida de probabilidad en $D(\mathbb{R}_+, E)$, *el espacio de Skorokhod*, con $X = \{X_t : t \geqslant 0\}$ el proceso canónico en dicho espacio.

Vamos a denotar $\mathcal{F}_t^0 = \sigma(X_s, s \leqslant t)$ la filtración canónica, y sea ν una medida de probabilidad en E. Además consideremos a $\mathcal{F}_t^\nu$ la σ-álgebra $\mathcal{F}_t^0$ completada con los conjuntos $\mathbb{P}_\nu$-nulos. Finalmente vamos a definir a

$$\mathcal{F}_t = \bigcap_\nu \mathcal{F}_t^\nu,$$

donde las ν son medidas de probabilidad. A la filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$ se le conoce como la *filtración universal*.

12. Lema:
La filtración $\{\mathcal{F}_t^\nu\}_{t\geqslant 0}$ es continua por la derecha. Esta propiedad también la satisface la filtración universal.

Demostración: Para probar este resultado basta ver que para toda variable aleatoria $Z \in L^\infty(\mathcal{F}_\infty^0)$ se tiene

$$\mathbb{E}_\nu[Z \mid \mathcal{F}_t^\nu] = \mathbb{E}_\nu[Z \mid \mathcal{F}_{t+}^\nu]. \qquad (\star)$$

Por un argumento de clases monótonas, basta verificar $(\star)$ para variables aleatorias en $L^\infty(\mathcal{F}^0_\infty)$ de la forma $Z = f_1(X_{t_1}) \cdots f_n(X_{t_n})$, donde $0 \leqslant t_1 < \cdots < t_n$ y $f_1, \ldots, f_n \in \mathcal{C}_0$.

Sin pérdida de generalidad vamos a tomar $t < t_1$. Para probar este resultado haremos uso de la propiedad de Markov y solo lo probaremos para $n = 1$, ya que el caso general es claro. Entonces $(\star)$ se reduce a probar

$$\mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_t] = \mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_{t+}].$$

De la propiedad de Markov al tiempo $t + \varepsilon < t_1$ tenemos

$$\mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_{t+\varepsilon}] = P_{t_1-t-\varepsilon} f_1(X_{t+\varepsilon}), \quad \mathbb{P}_\nu\text{-casi seguramente.}$$

Por otro lado, aplicando la propiedad de Markov al tiempo t, vemos

$$\mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_t] = P_{t_1-t} f(X_t), \quad \mathbb{P}_\nu\text{-casi seguramente.}$$

Cuando $\varepsilon \to 0$, tenemos $X_{t+\varepsilon} \to X_t$, $\mathbb{P}_\nu$-casi seguramente, gracias a la continuidad por la derecha y de las trayectorias y para $f_1 \in \mathcal{C}_0$ es claro que

$$P_{t_1-t-\varepsilon} f_1 \xrightarrow[\varepsilon \to 0]{} P_{t_1-t} f_1.$$

En otras palabras se tiene

$$\mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_{t+\varepsilon}] \xrightarrow[\varepsilon \to 0]{\text{c.s.}} P_{t_1-t} f_1(X_t) = \mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_t].$$

Para concluir observemos que el término de la izquierda, gracias al teorema de convergencia de martingalas retrógradas, converge $\mathbb{P}_\nu$-casi seguramente a $\mathbb{E}_\nu[f_1(X_{t_1}) \mid \mathcal{F}^\nu_{t+}]$, de donde deducimos el resultado. ■

Sea $\{\mathcal{F}_n\}_{n \leqslant 0}$ una sucesión de sub-σ-álgebras tales que $\mathcal{F}_n \subset \mathcal{F}_m$ para $n \leqslant m \leqslant 0$. Una *submartingala retrógrada* con respecto a $\{\mathcal{F}_n\}_{n \leqslant 0}$ es un proceso estocástico adaptado $\{X_n : n \leqslant 0\}$ tal que $\mathbb{E}[X_n^+] < \infty$ para $n \leqslant 0$ y además $X_n \leqslant \mathbb{E}[X_m \mid \mathcal{F}_n]$ para $n \leqslant m \leqslant 0$.

19. Teorema (Martingalas retrógradas)**:**
Si $\{X_n : n \leqslant 0\}$ *es una submartingala entonces* $\lim_{n \to -\infty} X_n$ *existe casi seguramente. Si en adición*

$$\sup_{n \leqslant 0} \mathbb{E}[|X_n|] < \infty,$$

entonces $\{X_n : n \leqslant 0\}$ *es uniformemente integrable y la convergencia se satisface en* L^1 *y para toda n, tenemos*

$$\lim_{k \to -\infty} X_k = \mathbb{E}[X_n \mid \mathcal{F}_{-\infty}],$$

donde $\mathcal{F}_{-\infty} = \bigcap_{k \leqslant 0} \mathcal{F}_k$.

Una aplicación interesante de este resultado es la célebre ley 0-1 de Blumenthal.

5. Corolario:
Para $\nu = \delta_x$, $\mathcal{F}_{0+}^{\nu}$ es trivial, es decir todo evento en $\mathcal{F}_{0+}^{\nu}$ tiene probabilidad cero o uno.

Demostración: Del lema 12 sabemos que $\mathcal{F}_t^{\nu} = \mathcal{F}_{t+}^{\nu}$. Por otro lado sabemos que $\mathcal{F}_0^{\nu}$ es la σ-completa de $\mathcal{F}_0^0$ con respecto a $\mathbb{P}_x$ y como $\mathcal{F}_0^0$ no es más que la σ-álgebra generada por X_0, la cual toma $\mathbb{P}_x$-casi seguramente el valor x. ∎

Por ejemplo, sea X un proceso de Feller que empieza en E. Si A es un abierto en E, tenemos

$$\mathbb{P}(\text{para toda } \varepsilon > 0 \text{ existe } t < \varepsilon \text{ tal que } X_t \in A) = 0 \text{ o } 1.$$

Otro ejemplo se da al considerar B abierto o cerrado y $T_B = \inf\{t \geqslant 0 : X_t \in B\}$, entonces el evento $\{T_B = 0\}$ es $\mathbb{P}_x$-trivial.

Decimos que x es *regular en B* si

$$\mathbb{P}_x(T_B = 0) = 1$$

e *irregular en el conjunto B* si

$$\mathbb{P}_x(T_B = 0) = 0.$$

5 de abril

20. Teorema (Propiedad de Markov fuerte)**:**
Sea T un $\{\mathcal{F}_t^{\nu}\}_{t\geqslant 0}$-tiempo de paro tal que sea finito casi seguramente. Definamos $X'_t = X_{T+t}$ para $t \geqslant 0$, entonces bajo $\mathbb{P}_\nu$ y condicionando bajo el evento $\{X_T = x\}$, el proceso X' es independiente de $\mathcal{F}_T^{\nu}$ y tiene por ley $\mathbb{P}_x$. En otras palabras para $Z \in L^{\infty}(\mathcal{F}_T^{\nu})$ y $F : D(\mathbb{R}_+, E) \to \mathbb{R}$ una funcional acotada, se tiene

$$\mathbb{E}_\nu[ZF(X')] = \mathbb{E}_\nu[Z\mathbb{E}_{X_T}[F(X)]].$$

Demostración: Supongamos primero que T es determinista. Este caso es claro gracias a que todo proceso de Feller satisface la propiedad de Markov simple. Ahora supongamos que T es simple, caso en el cual tenemos

$$\mathbb{E}_\nu[ZF(X')] = \sum_{n\geqslant 0} \mathbb{E}_\nu[ZF(X')\mathbb{1}_{\{T=t_n\}}] = \sum_{n\geqslant 0} \mathbb{E}_\nu[Z\mathbb{1}_{\{T=t_n\}}\mathbb{E}_{t_n}[F(X)]],$$

donde en la segunda igualdad se ha usado la propiedad de Markov simple y que $Z\mathbb{1}_{\{T=t_n\}}$ es $\mathcal{F}_{t_n}^{\nu}$-medible. Por consiguiente,

$$\mathbb{E}_\nu[ZF(X')] = \mathbb{E}_\nu[Z\mathbb{E}_{X_T}[F(X)]].$$

Para el caso general tomaremos una sucesión de tiempos de paro que decrezcan a T, denotada por $\{T_n\}_{n\geqslant 1}$. Apliquemos la propiedad de Markov fuerte en estos tiempos de paro simples y notemos que si $f : E^d \to \mathbb{R}$ es una función continua y acotada, se tiene

$$\mathbb{E}_\nu[Zf(X_{T_n+t_1}, \ldots, X_{T_n+t_d})] \xrightarrow[n\to\infty]{} \mathbb{E}_\nu[Zf(X'_{t_1}, \ldots, X'_{t_d})]$$

y en adición

$$\mathbb{E}_\nu[Z\mathbb{E}_{X_{T_n}}[f(X_{t_1},\dots,X_{t_d})]] \xrightarrow[n\to\infty]{} \mathbb{E}_\nu[Z\mathbb{E}_{X_\tau}[f(X_{t_1},\dots,X_{t_d})]],$$

donde las convergencias se satisfacen gracias a que $X_{T_n} \to X_T$ casi seguramente por la continuidad por la derecha de las trayectorias y el hecho que $\mathbb{P}_{x_n}$ converge a $\mathbb{P}_x$ conforme $x_n \to x$ en el sentido finito dimensional. Para completar la prueba basta usar el teorema de clases monótonas. ■

A continuación veamos que una consecuencia de la propiedad de Markov fuerte, junto con la propiedad de Feller, se tiene que todo proceso de Feller es de Hunt.

6. Corolario (Cuasi-continuidad por la izquierda):
Si $\{T_n\}_{n\geqslant 1}$ es una sucesión creciente de tiempos de paro tal que $T = \lim_{n\to\infty} T_n$, donde $T < \infty$ $\mathbb{P}_\nu$-casi seguramente, entonces

$$X_{T_n} \xrightarrow[n\to\infty]{} X_T \quad \mathbb{P}_\nu\text{-casi seguramente.}$$

Demostración: Supongamos $T_n < T$ para toda $n \geqslant 1$, si no el resultado es trivial. Además por hipótesis tenemos que $T < \infty$ $\mathbb{P}_\nu$-casi seguramente. Por un lado sabemos que $X_{T_n} \to X_{T-}$ casi seguramente con respecto a $\mathbb{P}_\nu$ gracias a la existencia del límite por la izquierda. Luego, para toda $\varepsilon > 0$ y $f, g \in \mathcal{C}_0$ tenemos

$$\mathbb{E}_\nu[f(X_{T_n})g(X_{T_n+\varepsilon})] \xrightarrow[n\to\infty]{} \mathbb{E}_\nu[f(X_{T-})g(X_{(T+\varepsilon)-})] \xrightarrow[\varepsilon\searrow 0]{} \mathbb{E}[f(X_{T-})g(X_T)].$$

Por otro lado, de la propiedad de Markov fuerte y de Feller tenemos

$$\begin{aligned}\mathbb{E}_\nu[f(X_{T_n})g(X_{T_n+\varepsilon})] &= \mathbb{E}_\nu[f(X_{T_n})P_\varepsilon g(X_{T_n})] \\ &\xrightarrow[n\to\infty]{} \mathbb{E}_\nu[f(X_{T-})P_\varepsilon g(X_{T-})] \xrightarrow[\varepsilon\searrow 0]{} \mathbb{E}_\nu[f(X_{T-})g(X_{T-})].\end{aligned}$$

De este modo,

$$\mathbb{E}_\nu[f(X_{T-})g(X_T)] = \mathbb{E}_\nu[f(X_{T-})g(X_{T-})].$$

Por un argumento de clases monótonas tenemos que para $h : E^2 \to \mathbb{R}$ medible se tiene

$$\mathbb{E}_\nu[h(X_{T-}, X_T)] = \mathbb{E}_\nu[h(X_{T-}, X_{T-})].$$

Tomando $h(x,y) = \mathbb{1}_{\{x\neq y\}}$ se obtiene que $X_{T-} = X_T$ con probabilidad uno bajo $\mathbb{P}_\nu$. ■

Una aplicación de la cuasi-continuidad por la izquierda es que las trayectorias de un proceso de Feller no tienen *tiempos de salto accesibles*, es decir que no se pueden obtener como límite estricto de tiempos de paro. En efecto, si $T_n < T$, por la cuasi-continuidad por la izquierda tenemos $X_{T-} = X_T$ se da $\mathbb{P}_\nu$-casi seguramente. En particular no hay tiempos de saltos fijos y en consecuencia

$$P_{t-\varepsilon} f(x) = \mathbb{E}_x[f(X_{t-\varepsilon})] \xrightarrow[\varepsilon\searrow 0]{} \mathbb{E}_x[f(X_t)] = P_t f(x) \tag{1}$$

para $f \in \mathcal{C}_0$.

6.5. Resolvente y generador infinitesimal

Sea $\{X_t : t \geqslant 0\}$ un proceso de Feller. Recordemos que tiene trayectorias càdlàg. Primero observemos que para $f \in \mathcal{C}_0$ y $x \in E$ fijo, la aplicación $t \mapsto P_t f(x)$ es continua. En efecto, recordemos que

$$P_t f(x) = \mathbb{E}_x[f(X_t)] \quad \text{y} \quad P_{t+\varepsilon} f(x) = P_\varepsilon P_t f(x).$$

Como $\{P_t : t \geqslant 0\}$ es felleriano, tenemos $P_t g \in \mathcal{C}_0$ y además

$$P_t g(x) \xrightarrow[t\to 0]{} g(x),$$

siempre y cuando $g \in \mathcal{C}_0$. Entonces claramente

$$P_{t+\varepsilon} f(x) \xrightarrow[\varepsilon\to 0]{} P_t f(x).$$

Por otro lado,

$$\begin{aligned} P_{t-\varepsilon} f(x) &= P_{t-\varepsilon} P_\varepsilon f(x) + P_{t-\varepsilon}(f(x) - P_\varepsilon f(x)) = P_t f(x) + P_{t-\varepsilon}(f(x) - P_\varepsilon f(x)) \\ &\xrightarrow[\varepsilon \searrow 0]{} P_t f(x), \end{aligned}$$

con el segundo término convergiendo a cero por tratarse de un semigrupo de contracciones. Forzosamente, la aplicación también es medible y por un argumento de clases monótonas vemos que si $f \in L^\infty(E)$ entonces $t \mapsto P_t f(x)$ es medible. Esto implica que podemos definir, para $f \in L^\infty(E)$ al operador *q-resolvente*,

$$V^{(q)} f(x) = \int_0^\infty e^{-qs} P_s f(x)\, \mathrm{d}s. \tag{q-res}$$

Si restringimos a $V^{(q)}$ en $\mathcal{C}_0$ entonces $V^{(q)}$ también es un operador felleriano, es decir $V^{(q)}\mathcal{C}_0 \subset \mathcal{C}_0$. Vamos a observar que para $f \in \mathcal{C}_0$ tenemos

$$\lim_{q\to\infty} q V^{(q)} f = f.$$

Para ello basta usar el cambio de variable $u = qs$ y la continuidad fuerte del semigrupo $\{P_t : t \geqslant 0\}$.

13. Lema (Ecuación resolvente)**:**
La propiedad de semigrupo implica la célebre ecuación resolvente, es decir

$$V^{(q)} - V^{(r)} + (q-r) V^{(q)} \circ V^{(r)} = 0 \quad \textit{para } q, r > 0.$$

Demostración: De la propiedad de semigrupo tenemos

$$\begin{aligned} V^{(q)} \circ V^{(r)} f(x) &= \int_0^\infty e^{-qs} P_s \int_0^\infty e^{-ru} P_u f(x)\, \mathrm{d}u\, \mathrm{d}s \\ &= \int_0^\infty e^{-qs} \int_0^\infty e^{-ru} P_{s+u} f(x)\, \mathrm{d}u\, \mathrm{d}s. \end{aligned}$$

Ahora, sin pérdida de generalidad, supongamos que $q > r$ y observemos

$$\begin{aligned}\int_0^\infty \mathrm{d}s e^{-qs} \int_0^\infty \mathrm{d}u e^{-ru} P_{s+u} f(x) &= \int_0^\infty \mathrm{d}s e^{-rs} e^{-(q-r)s} \int_0^\infty \mathrm{d}u e^{-ru} P_{s+u} f(x) \\ &= \int_0^\infty \mathrm{d}s e^{-(q-r)s} \int_0^\infty \mathrm{d}u e^{-r(u+s)} P_{s+u} f(x) \\ &= \int_0^\infty \mathrm{d}s e^{-(q-r)s} \int_s^\infty \mathrm{d}u e^{-ru} P_u f(x) \\ &= \int_0^\infty \mathrm{d}u e^{-ru} P_u f(x) \int_0^t \mathrm{d}s e^{-s(q-r)} \\ &= \frac{1}{q-r}\Big(V^{(r)} f(x) - V^{(q)} f(x)\Big). \qquad \blacksquare\end{aligned}$$

26 de abril

Una de las propiedades interesantes del operador resolvente es que es *conmutativa*, es decir que

$$V^{(q)} \circ V^{(r)} = V^{(r)} \circ V^{(q)}, \quad q, r > 0.$$

(Para darle una interpretación probabilista al operador resolvente, consideremos a e_q, una variable aleatoria exponencial de parámetro $q > 0$, independiente del proceso $\{X_t : t \geqslant 0\}$. Entonces se tiene que

$$qV^{(q)} f(x) = \int_0^\infty q e^{-qt} \mathbb{E}_x[f(X_t)] \, \mathrm{d}t = \mathbb{E}_x[f(X_{e_q})].$$

Hemos visto que $V^{(q)}\mathcal{C}_0 \subset \mathcal{C}_0$ y como veremos, la imagen no depende de q, gracias a la ecuación resolvente. De hecho a esta imagen se le conocerá como el generador infinitesimal. En efecto, sea $f \in \mathcal{C}_0$,

$$V^{(r)} f(x) = V^{(q)} f(x) + (q-r) V^{(q)} \circ V^{(r)} f(x) = V^{(q)}\Big(f + (q-r)V^{(r)} f\Big)(x) = V^{(q)} g(x),$$

donde $g = f + (q-r)V^{(r)} f \in \mathcal{C}_0$ debido a que $V^{(r)} f \in \mathcal{C}_0$. Vamos a definir a esta imagen

$$\boxed{\mathcal{D} = V^{(q)}\mathcal{C}_0, \quad q > 0.}$$

31. Definición:
Llamaremos a $\mathcal{D}$ el *dominio del generador infinitesimal* de X o $\{P_t : t \geqslant 0\}$.

El subespacio $\mathcal{D}$ es denso en $\mathcal{C}_0$ gracias a la convergencia uniforme, es decir que para $f \in \mathcal{C}_0$ tenemos $qV^{(q)} f \to f$ fuertemente.

Otra consecuencia de la ecuación resolvente es $V^{(q)} : \mathcal{C}_0 \to \mathcal{D}$ es una biyección. Para verificar esta afirmación basta ver que la función es inyectiva, ya que por definición de $\mathcal{D}$ es claro que es sobreyectiva. Para checar que es inyectiva consideremos a $f \in \mathcal{C}_0$ tal que $V^{(q)} f \equiv 0$. De la ecuación resolvente se deduce que $V^{(r)} f \equiv 0$ para toda $r > 0$. Por lo tanto vemos que $rV^{(r)}f \equiv 0$ y en consecuencia $f \equiv 0$ pues $rV^{(r)}f \to f$. Esto prueba que

$$\ker V^{(q)} = \{0\}.$$

Esto prueba la inyectividad y la afirmación realizada.

32. Definición:
Llamaremos *generador infinitesimal del proceso* X, o del semigrupo $\{P_t : t \geqslant 0\}$, al operador $\mathcal{G} : \mathcal{D} \to \mathcal{C}_0$ definido por

$$\mathcal{G}f = f - (V^{(1)})^{-1} f,$$

donde $(V^{(1)})^{-1}$ es el inverso de $V^{(1)}$.

9. *Observación:*
Observemos $V^{(q)}$ es el inverso de $q\,\mathrm{id} - \mathcal{G}$. Para ello tomemos $f \in \mathcal{D}$, de modo que $f = V^{(1)}g$ con $g \in \mathcal{C}_0$ y en consecuencia

$$\mathcal{G}f = f - g.$$

Así,

$$V^{(q)}(qf - f + g) = (q-1)V^{(q)}V^{(1)}g + V^{(q)}g = V^{(1)}g = f.$$

23. Proposición:
Si $f \in \mathcal{D}$ entonces

$$\mathcal{G}f = \lim_{t \to 0} \frac{P_t f - f}{t}.$$

De manera más general, para $t \geqslant 0$, se cumple

$$\mathcal{G}P_t f = \lim_{\varepsilon \to 0} \frac{P_{t+\varepsilon} f - P_t f}{\varepsilon},$$

la cual es conocida como ecuación backward de Kolmogorov. *Recíprocamente si $f \in \mathcal{C}_0$ y si además $\varepsilon^{-1}(P_\varepsilon f - f)$ tiene un límite conforme $\varepsilon \to 0$, digamos $g \in \mathcal{C}_0$, entonces $f \in \mathcal{D}$ y $\mathcal{G}f = g$.*

Demostración: Supongamos que $f \in \mathcal{D}$, de modo que $f = V^{(1)}g$ para alguna $g \in \mathcal{C}_0$. Entonces para $\varepsilon > 0$ vemos

$$f - P_\varepsilon f = V^{(1)}f - P_\varepsilon V^{(1)}g = V^{(1)}g - e^{-\varepsilon}P_\varepsilon V^{(1)}g + (e^{-\varepsilon} - 1)P_\varepsilon V^{(1)}g.$$

Por un lado, de la propiedad de semigrupo tenemos

$$\begin{aligned} V^{(1)}g - e^{-\varepsilon}P_\varepsilon V^{(1)}g &= \int_0^\infty e^{-t}P_t g \,\mathrm{d}t - e^{-\varepsilon}P_\varepsilon \int_0^\infty e^{-t}P_t g \,\mathrm{d}g \\ &= \int_0^\varepsilon e^{-t}P_t g \,\mathrm{d}t. \end{aligned}$$

Como $e^{-t}P_t g$ es continua en $t = 0$ deducimos que

$$\frac{V^{(1)}g - e^{-\varepsilon}P_\varepsilon V^{(1)}g}{\varepsilon} \xrightarrow[\varepsilon\to 0]{} g.$$

Por otro lado, es claro que

$$\frac{1}{\varepsilon}(e^{-\varepsilon}-1)P_\varepsilon V^{(1)}g \xrightarrow[\varepsilon\to 0]{} -V^{(1)}g.$$

Con esto se sigue que

$$\frac{f - P_\varepsilon f}{\varepsilon} \xrightarrow[\varepsilon\to 0]{} g - V^{(1)}g = (V^{(1)})^{-1}f - f = -\mathcal{G}f.$$

Para obtener el caso general basta notar que $h_t = P_t f \in \mathcal{D}$ siempre que $f \in \mathcal{D}$, lo cual se obtiene de que $P_t V^{(1)}g = V^{(1)}P_t g$ para $g \in \mathcal{C}_0$.

De manera recíproca, supongamos que $f \in \mathcal{C}_0$ y que

$$g = \lim_{\varepsilon\to 0} \frac{P_\varepsilon f - f}{\varepsilon}$$

existe y es un elemento de $\mathcal{C}_0$. Entonces

$$V^{(1)}g = \lim_{\varepsilon\to 0} V^{(1)}\left(\frac{P_\varepsilon f - f}{\varepsilon}\right) = \lim_{\varepsilon\to 0} \frac{P_\varepsilon V^{(1)}f - V^{(1)}f}{\varepsilon},$$

donde $V^{(1)}f \in \mathcal{D}$. De la primera parte y la definición de generador tenemos que

$$\mathcal{G}V^{(1)}f = -f + V^{(1)}f = V^{(1)}g,$$

de donde deducimos que

$$f = V^{(1)}(f-g),$$

implicando $f \in \mathcal{D}$ y $\mathcal{G}f = g$. ∎

28 de abril

Formalmente, la ecuación de Kolmogorov es una ecuación diferencial lineal

$$\frac{\mathrm{d}}{\mathrm{d}t}P_t = \mathcal{G}P_t,$$

cuya solución está dada por $P_t = e^{t\mathcal{G}}$ cuando el espacio de estados E es finito o cuando $\mathcal{G}$ es un operador autoadjunto (ver Yoshida, 1980). En otras palabras se tiene

$$P_t P_s = e^{t\mathcal{G}}e^{s\mathcal{G}} = e^{(s+t)\mathcal{G}}, \quad V^{(q)} = \int_0^\infty e^{-qt}e^{t\mathcal{G}} \,\mathrm{d}t = \frac{1}{q-\mathcal{G}},$$

entre otras cosas.

24. Proposición:
Sea $f \in \mathcal{D}$, entonces para $x \in E$,

$$f(X_t) - \int_0^t \mathcal{G}f(X_s)\,\mathrm{d}s, \quad t \geqslant 0,$$

es una $\mathbb{P}_x$-martingala. Recíprocamente, si $f, g \in \mathcal{C}_0$ son tales que para toda $x \in E$ tenemos

$$f(X_t) - \int_0^t g(X_s)\,\mathrm{d}s, \quad t \geqslant 0,$$

es una $\mathbb{P}_x$-martingala, entonces $f \in \mathcal{D}$ y $g = \mathcal{G}f$.

Demostración: Si $f \in \mathcal{D}$, entonces $\mathcal{G}f = \lim_{\varepsilon\to 0} \varepsilon^{-1}(P_\varepsilon f - f)$ y de manera general se tiene la ecuación de Kolmogorov,

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = \mathcal{G} P_t f.$$

Denotemos

$$M_t = f(X_t) - \int_0^t \mathcal{G}f(X_s)\,\mathrm{d}s, \quad t \geqslant 0$$

y veamos que es martingala. Para ello veamos que

$$\begin{aligned}
\mathbb{E}_x[M_{t+s} \mid \mathcal{F}_t] &= \mathbb{E}_x[f(X_{t+s}) \mid \mathcal{F}_s] - \int_0^t \mathcal{G}f(X_u)\,\mathrm{d}u - \mathbb{E}_x\left[\int_t^{t+s} \mathcal{G}f(X_u)\,\mathrm{d}u \,\middle|\, \mathcal{F}_s\right] \\
&= P_s f(X_t) - \int_0^t \mathcal{G}f(X_u)\,\mathrm{d}u - \int_0^s P_u \mathcal{G}f(X_t)\,\mathrm{d}u \\
&= M_t + P_s f(X_t) - f(X_t) - \int_0^s P_u \mathcal{G}f(X_t)\,\mathrm{d}u,
\end{aligned}$$

por lo que la propiedad de martingala se reduce a probar que para toda $y \in E$,

$$P_s f(y) - f(y) - \int_0^t P_u \mathcal{G}f(y)\,\mathrm{d}u = 0,$$

lo cual se da por la *ecuación forward de Kolmogorov*:

$$P_t \mathcal{G}f = \lim_{\varepsilon\to 0} P_t\left(\frac{P_\varepsilon f - f}{\varepsilon}\right) = \lim_{\varepsilon\to 0} \frac{P_{t+\varepsilon} f - P_t f}{\varepsilon} = \frac{\mathrm{d}}{\mathrm{d}t} P_t f.$$

En sentido inverso supongamos $f, g \in \mathcal{C}_0$ con la propiedad de que

$$M_t = f(X_t) - \int_0^t g(X_s)\,\mathrm{d}s, \quad t \geqslant 0,$$

es una martingala bajo $\mathbb{P}_x$, con $x \in E$. En tal caso

$$\mathbb{E}_x[M_t] = P_t f(x) - \mathbb{E}_x\left[\int_0^t g(X_s)\,\mathrm{d}s\right] = f(x),$$

de donde

$$\frac{P_t f(x) - f(x)}{t} = \frac{1}{t}\mathbb{E}_x\left[\int_0^t g(X_s)\ \mathrm{d}s\right].$$

Al tomar t suficientemente pequeña, el término de la derecha converge a $g(x)$ gracias al teorema de convergencia dominada. Por lo tanto,

$$\frac{P_t f(x) - f(x)}{t} \xrightarrow[t\to 0]{} g(x),$$

convergencia que nos permite concluir. ■

7. Corolario (Fórmula de Dynkin):
Si $f \in \mathcal{D}$ y T es un tiempo de paro tal que $\mathbb{E}_x[T] < \infty$, tenemos

$$\mathbb{E}_x\left[\int_0^T \mathcal{G}f(X_u)\ \mathrm{d}u\right] = \mathbb{E}_x[f(X_T)] - f(x).$$

Demostración: Si T es un tiempo de paro acotado entonces el resultado se sigue del teorema de paro opcional de Doob a la martingala

$$f(X_t) - \int_0^t \mathcal{G}f(X_s)\ \mathrm{d}s, \quad t \geqslant 0.$$

En el caso general tomamos $T \wedge n$ y al aplicar el teorema de paro opcional observamos

$$\mathbb{E}_x\left[\int_0^{T\wedge n} \mathcal{G}f(X_s)\ \mathrm{d}s\right] = \mathbb{E}_x[f(X_{T\wedge n})] - f(x).$$

Al tomar n suficientemente grande, lo primero que deducimos del teorema de convergencia dominada es que

$$\mathbb{E}_x[f(X_{T\wedge n})] \xrightarrow[n\to\infty]{} \mathbb{E}_x[f(X_T)].$$

Mientras tanto, podemos notar que $\|\mathcal{G}f\|_\infty < \infty$ y que

$$\left|\int_0^{T\wedge n} \mathcal{G}f(X_s)\ \mathrm{d}s\right| \leqslant T\,\|\mathcal{G}f\|_\infty \in L^1(\mathbb{P}_x),$$

de modo que

$$\mathbb{E}_x\left[\int_0^{T\wedge n} \mathcal{G}f(X_s)\ \mathrm{d}s\right] \xrightarrow[n\to\infty]{} \mathbb{E}_x\left[\int_0^T \mathcal{G}f(X_s)\ \mathrm{d}s\right]$$

gracias al teorema de convergencia dominada. ■

3 de mayo

Ejemplo 1. Sea $\{P_t : t \geqslant 0\}$ un semigrupo markoviano en $(E, \mathcal{E})$ donde E es un espacio métrico localmente compacto. Supongamos que $P_t : \mathcal{C}_0 \to \mathcal{C}_0$ y que para toda $f \in \mathcal{C}_0$ tenemos

$$P_t f(x) \xrightarrow[t\to 0]{} f(x), \quad \text{para } x \in E.$$

Deseamos probar que la convergencia es uniforme, *i.e.* $\{P_t : t \geqslant 0\}$ es Feller.

Hints: (i) probar que $\mathcal{D}$ es denso en $\mathcal{C}_0$; (ii) ver que $\|P_t g - g\| \to 0$ para $g \in \mathcal{D}$; (iii) concluir.

Del teorema de representación de Riesz, para probar la densidad de $\mathcal{D}$ basta ver que para cualquier medida finita μ definida en E, si

$$\int_E f \, \mathrm{d}\mu = 0 \quad \text{para } f \in \mathcal{D}$$

implica que $\mu = 0$. Para verificar que esto es cierto, observemos que para $g \in \mathcal{C}_0$ y $q > 0$,

$$\int_E qV^{(q)} g \, \mathrm{d}\mu = 0, \quad \lim_{q\to\infty} qV^{(q)} g(x) = g(x) \quad \text{y} \quad \|qV^{(q)} g\| \leqslant \|g\|,$$

de modo que por el teorema de convergencia dominada $\int_E g \, \mathrm{d}\mu = 0$. Ahora bien, para $g \in \mathcal{C}_0$ sea $f = V^{(1)} g$, pues entonces

$$e^{-\varepsilon} P_\varepsilon f(x) = \int_\varepsilon^\infty P_u g(x) e^{-u} \, \mathrm{d}u$$

y en consecuencia

$$f - e^{-\varepsilon} P_\varepsilon f = \int_0^t P_u g\, e^{-u} \, \mathrm{d}u \xrightarrow[\varepsilon\to 0]{} 0$$

ya que g es acotada. Por último resta notar que

$$f - P_\varepsilon f = (f - e^{-\varepsilon} P_\varepsilon f) + (e^{-\varepsilon} - 1) P_\varepsilon f \xrightarrow[\varepsilon\to 0]{} 0.$$

Ejemplo 2. (Generador de un subordinador) Sea σ un subordinador con coeficiente de deriva $d \geqslant 0$ y medida de Lévy π. Recordemos que

$$\int_{(0,\infty)} (1 \wedge x)\, \pi(\mathrm{d}x) < \infty$$

y su transformada de Laplace satisface

$$\mathbb{E}[e^{-q\sigma_t}] = e^{-t\phi(q)} \quad \text{con} \quad \phi(q) = dq + \int_{(0,\infty)} (1 - e^{-qx})\, \pi(\mathrm{d}x).$$

Para $f(x) = e^{-qx}$ calcular $P_t f(x)$, verificar que $f \in \mathcal{D}$ y

$$\mathcal{G} f(x) = d f'(x) + \int_{(0,\infty)} (f(x+y) - f(x))\, \pi(\mathrm{d}y).$$

Finalmente extender para $f \in \mathcal{C}^1$ y $f, f' \in \mathcal{C}_0$.

6.6. Teorema de Hille–Yosida

10 de mayo

Nuestro objetivo principal es construir un semigrupo a partir de un generador o de un resolvente. Aquí veremos la versión de Yosida, *i.e.* la construcción de un semigrupo a partir de una resolvente. Vamos a trabajar en $\mathcal{C}_0(E,\mathbb{R})$. Llamaremos *resolvente en el espacio* $\mathcal{C}_0(E,\mathbb{R})$ a una familia de operadores $\{V^{(q)} : q \geqslant 0\}$ tal que

$$V^{(q)} - V^{(r)} + (q-r)V^{(q)} \circ V^{(r)} = 0.$$

Es importante señalar que si empezamos nuestro estudio a partir de un generador $\mathcal{G}$ y definimos a $V^{(q)}$ como el inverso de $q\,\mathrm{id} - \mathcal{G}$, la ecuación resolvente se verifica fácilmente.

Vamos a suponer que $\mathcal{D} = V^{(q)}(\mathcal{C}_0)$ es denso en $\mathcal{C}_0$ y que el operador resolvente es de contracción, es decir $\|qV^{(q)}f\| \leqslant \|f\|$ y positivo, es decir que si $f \geqslant 0$ entonces $V^{(q)}f \geqslant 0$. En otras palabras, si vemos esto en términos del generador es equivalente al *principio del máximo*:

$$\boxed{\text{Para } f \in \mathcal{D} \text{ tal que } \sup_{x\in E} f(x) = f(x_0) \geqslant 0, \text{ entonces } \mathcal{G}f(x_0) \leqslant 0.}$$

Deseamos construir un semigrupo cuya resolvente sea $V^{(q)}$, en cierto sentido una inversión de la transformada de Laplace. La idea es trabajar con una aproximación del generador. Para $\lambda > 0$ definimos al operador

$$\mathcal{G}_\lambda = \lambda(\lambda V^{(\lambda)} - \mathrm{id}).$$

Observemos que

$$\mathcal{G}_\lambda V^{(q)} = \lambda(\lambda V^{(\lambda)} V^{(q)} - V^{(q)}),$$

mientras que de la ecuación resolvente, vemos

$$\lambda V^{(\lambda)} V^{(q)} = \frac{\lambda}{\lambda - q}(V^{(q)} - V^{(\lambda)}).$$

Juntando ambos resultados,

$$\mathcal{G}_\lambda V^{(q)} = \frac{\lambda}{\lambda - q}(qV^{(q)} - \lambda V^{(\lambda)}).$$

Así, si suponemos que $\lambda V^{(\lambda)} \to \mathrm{id}$ conforme $\lambda \to \infty$, tendremos

$$\mathcal{G}_\lambda V^{(q)} \xrightarrow[\lambda\to\infty]{} qV^{(q)} - \mathrm{id} = \mathcal{G}V^{(q)}.$$

Verifiquemos que $\lambda V^{(\lambda)} \to \mathrm{id}$ conforme $\lambda \to \infty$. Para ello tomemos $f \in \mathcal{D}$ y $f = V^{(1)}g$. Gracias a la ecuación resolvente tenemos

$$\begin{aligned}\lambda V^{(\lambda)}f &= (\lambda-1)V^{(\lambda)}f + V^{(\lambda)}f\\ &= (\lambda-1)V^{(\lambda)}V^{(1)}g + V^{(\lambda)}f\\ &= V^{(1)}g - V^{(\lambda)}g + V^{(\lambda)}V^{(1)}g \xrightarrow[\lambda\to\infty]{} V^{(1)}g = f,\end{aligned}$$

donde la convergencia se da gracias a que $\|\lambda V^{(\lambda)}\|_{\mathsf{op}} \leqslant 1$, de modo que $\|V^{(\lambda)}\|_{\mathsf{op}} \to 0$ cuando $\lambda \to \infty$.

Ahora extenderemos el resultado a $\mathcal{C}_0$ usando la densidad de $\mathcal{D}$. Sea entonces $f \in \mathcal{C}_0$ y consideremos a $\{f_n\}_{n\geqslant 1} \subset \mathcal{D}$ tal que $f_n \to_n f$. Entonces

$$\|\lambda V^{(\lambda)}f - f\| \leqslant \|\lambda V^{(\lambda)}f - \lambda V^{(\lambda)}f_n\| + \|\lambda V^{(\lambda)}f_n - f_n\| + \|f_n - f\|.$$

Tomemos $\varepsilon > 0$ fijo y n tal que $\|f - f_n\| \leqslant \varepsilon$. Por un lado, para λ suficientemente grande tenemos

$$\|\lambda V^{(\lambda)}f_n - f_n\| \leqslant \varepsilon.$$

Por otro lado $\|\lambda V^{(\lambda)}\|_{\mathsf{op}} \leqslant 1$, y por ende

$$\|\lambda V^{(\lambda)}f - f\| \leqslant 3\varepsilon.$$

Hecho lo anterior, definamos

$$P_{t,\lambda} = e^{t\mathcal{G}_\lambda} = e^{-t\lambda}e^{t\lambda^2 V^{(\lambda)}} = e^{-t\lambda}\sum_{n\geqslant 0}\frac{(t\lambda)^n}{n!}(V^{(\lambda)})^n.$$

Observemos que $\{P_{t,\lambda} : t \geqslant 0\}$ es un semigrupo, esto es

$$P_{t,\lambda} \circ P_{s,\lambda} = e^{t\mathcal{G}_\lambda}e^{s\mathcal{G}_\lambda} = e^{(t+s)\mathcal{G}_\lambda}$$

y además

$$\lim_{h\searrow 0}\frac{1}{h}(P_{h,\lambda} - \mathrm{id}) = \mathcal{G}_\lambda.$$

Como $V^{(q)}$ es un operador de contracción, aún más $\|\lambda V^{(\lambda)}\|_{\mathsf{op}} \leqslant 1$, tenemos que $\|P_{t,\lambda}\|_{\mathsf{op}} \leqslant 1$. Finalmente, si $f \geqslant 0$ entones $P_{t,\lambda}f \geqslant 0$. Falta ver que cuando $\lambda \to \infty$, $P_{t,\lambda} \to P_t$, donde $\{P_t : t \geqslant 0\}$ es un semigrupo markoviano con resolvente en $V^{(q)}$. Para establecer la convergencia vamos a verificar un criterio de Cauchy. Sean $\lambda, \mu \geqslant 0$, entonces

$$\begin{aligned}P_{t,\lambda}f - P_{t,\mu}f &= \sum_{k=1}^{n} P_{(k-1)t/n,\lambda}P_{(n-k)t/n,\mu}\Big(P_{t/n,\lambda}f - P_{t/n,\mu}f\Big)\\ &= \sum_{k=1}^{n} P_{kt/n,\lambda}P_{(n-k)t/n,\mu}f - \sum_{l=0}^{n-1} P_{tl/n,\lambda}P_{(n-l)t/n,\mu}f.\end{aligned}$$

Por un lado tenemos que $\|P_{s,\lambda}\|_{\mathrm{op}} \leqslant 1$, de modo que

$$\|P_{t,\lambda}f - P_{t,\mu}f\| \leqslant n\|P_{t/n,\lambda}f - P_{t/n,\mu}f\|.$$

Por otro lado sabemos que

$$n\Big(P_{t/n\lambda}f - f\Big) \xrightarrow[n\to\infty]{} t\mathcal{G}_\lambda f,$$

por lo cual

$$\|P_{t,\lambda}f - P_{t,\mu}f\| \leqslant t\|\mathcal{G}_\lambda f - \mathcal{G}_\mu f\| \xrightarrow[\lambda,\mu\to\infty]{} 0$$

siempre y cuando $f \in \mathcal{D}$ pues

$$\lim_{\lambda\to\infty} \mathcal{G}_\lambda f = \mathcal{G}f.$$

12 de mayo

El criterio de Cauchy nos permite deducir la existencia de un semigrupo en $\mathcal{C}_0$. De hecho tenemos que $\lim_{\lambda\to\infty} P_{t,\lambda}f$ existe para toda $t \geqslant 0$ uniformemente en intervalos acotados para $f \in \mathcal{D}$.

$\{P_t : t \geqslant 0\}$ es semigrupo Basta ver que

$$\begin{aligned} P_{t+s}f - P_tP_sf &= P_{t+s}f - P_{t+s,\lambda}f + P_{t+s,\lambda}f - P_{t,\lambda}P_sf + P_{t,\lambda}P_s - P_tP_sf \\ &= P_{t+s}f - P_{t+s,\lambda}f + P_{t,\lambda}(P_{s,\lambda}f - P_sf) + (P_{t,\lambda} - P_t)P_sf \end{aligned}$$

es válido para toda $\lambda > 0$, pues entonces

$$P_{t+s}f = P_tP_sf.$$

$\{P_t : t \geqslant 0\}$ es de contracción En este caso se tiene

$$\|P_tf\| \leqslant \|P_{t,\lambda}f\| + \|P_{t,\lambda}f - P_tf\| \leqslant \|f\| + \|P_{t,\lambda}f - P_tf\| \xrightarrow[\lambda\to\infty]{} 1.$$

$\{P_t : t \geqslant 0\}$ es de Feller y genera a $\mathcal{G}$ Notemos que

$$P_{t,\lambda}f - f = \int_0^t P_{u,\lambda}\mathcal{G}_\lambda f \,\mathrm{d}u, \quad t \geqslant 0,\ f \in \mathcal{D} \text{ y } \lambda > 0. \tag{$\star$}$$

Ahora, para cada $f \in \mathcal{D}$ y $t \geqslant 0$ se tiene la identidad

$$P_{s,\lambda}\mathcal{G}_\lambda f - P_s\mathcal{G}f = P_{s,\lambda}(\mathcal{G}_\lambda f - \mathcal{G}f) + (P_{s,\lambda} - P_s)\mathcal{G}f.$$

Dado que para $f \in \mathcal{D}$ se tiene que $\lim_{\lambda\to\infty} \mathcal{G}_\lambda f = \mathcal{G}f$, de lo anterior se sigue que

$$\lim_{\lambda\to\infty} P_{s,\lambda}\mathcal{G}_\lambda f = P_t\mathcal{G}f$$

uniformemente en $0 \leqslant s \leqslant t$. Esto implica uqe $(\star)$ se satisface para P y $\mathcal{G}$ cuando $f \in \mathcal{D}$, es decir

$$P_t f - f = \int_0^t P_u \mathcal{G} f \, \mathrm{d}u.$$

En particular esto implica

$$\mathcal{A} f = \lim_{h \to 0} \frac{P_h f - f}{h} = \lim_{h \to 0} \frac{1}{h} \int_0^h P_u \mathcal{G} f \, \mathrm{d}u = \mathcal{G} f,$$

de modo que $\mathcal{D} \subset \mathcal{D}_{\mathcal{A}}$ y $\mathcal{A} f = \mathcal{G} f$ para $f \in \mathcal{D}$. En otras palabras, $\mathcal{A}$ es una extensión de $\mathcal{G}$. Como $\mathcal{A}$ es el generador de $\{P_t : t \geqslant 0\}$, $\lambda \operatorname{id} - \mathcal{A}$ mapea a $\mathcal{D}_{\mathcal{A}}$ en $\mathcal{C}_0$ uno a uno. Pero $\lambda \operatorname{id} - \mathcal{G}$ coincide con $\lambda \operatorname{id} - \mathcal{A}$ en $\mathcal{D}$ y mapea este dominio en $\mathcal{C}_0$, lo cual implica que $\mathcal{D} = \mathcal{D}_{\mathcal{A}}$ y, en consecuencia, $\mathcal{A} = \mathcal{G}$.

6.7. Procesos de Markov de saltos puros

Sea E un espacio métrico, $K(x, A)$ una probabilidad de transición sobre $E \times \mathcal{B}(E)$ y $\lambda : E \to \mathbb{R}$ acotada. Definamos al operador

$$\boxed{\mathcal{A} f(x) = \lambda(x) \int_E (f(y) - f(x)) \, K(x, \mathrm{d}y)}. \qquad (\clubsuit)$$

Este operador resulta ser un operador lineal acotado en $B(E)$, el espacio de funciones reales y acotadas.

Objetivo: construir un proceso de Markov en E con generador infinitesimal dado por $\mathcal{A}$ como en $(\clubsuit)$.

Sea $Y = \{Y_k : k \geqslant 0\}$ una cadena de Markov en E con distribución inicial ν y probabilidad de transición $K(x, A)$, con $A \in \mathcal{B}(E)$. Es decir que $\mathbb{P}(Y_0 \in A) = \nu(A)$ mientras que $\mathbb{P}(Y_{k+1} \in A \mid Y_0, \dots, Y_k) = K(Y_k, A)$ para $A \in \mathcal{B}(E)$ y $k \geqslant 0$.

Sea $\{\mathrm{e}^{(k)} : k \geqslant 0\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución exponencial de parámetro 1, las cuales las tomaremos independientes de Y. Entonces

$$X_t = \begin{cases} Y_0 & \text{si } 0 \leqslant t < \dfrac{\mathrm{e}^{(0)}}{\lambda(Y_0)}, \\ Y_k & \text{si } \displaystyle\sum_{i=1}^{k-1} \frac{\mathrm{e}^{(i)}}{\lambda(Y_i)} \leqslant t < \sum_{i=1}^{k} \frac{\mathrm{e}^{(i)}}{\lambda(Y_i)}, \end{cases}$$

define un proceso de Markov en E con distribución inicial ν y generador $\mathcal{A}$.

Sea $\lambda = \sup_{x \in E} \lambda(x)$ y para evitar trivialidades supondremos que $\lambda > 0$. Definamos $\kappa(x, A)$ a una probabilidad de transición en $E \times \mathcal{B}(E)$ que satisfaga:

$$\kappa(x, A) = \left(1 - \frac{\lambda(x)}{\lambda}\right) \delta_x(A) + \frac{\lambda(x)}{\lambda} K(x, A)$$

y observamos que $\mathcal{A}$ se puede escribir como

$$\mathcal{A} = \lambda \int_E (f(y) - f(x))\, \kappa(x, \mathrm{d}y).$$

Sea ahora $\{Y'_k : k \geqslant 0\}$ una cadena de Markov en E con distribución inicial ν y probabilidad de transición $\kappa(x, A)$. En adición sea $N = \{N_t : t \geqslant 0\}$ un proceso de Poisson de independiente de parámetro λ. Definamos $X'_t = Y'_{N_t}$ para $t \geqslant 0$.

(Ejercicio) Probar que X y X' tienen las mismas distribuciones finito-dimensionales.

17 de mayo

Observemos que

$$Pf(x) = \int_E f(y)\, \kappa(x, \mathrm{d}y)$$

define un operador lineal de contracción en $B(E)$ con el cual obtenemos la relación $\mathcal{A} = \lambda(P - \mathrm{id})$. Por ende, el semigrupo $\{P_t : t \geqslant 0\}$ generado por $\mathcal{A}$ está dado por

$$P_t = \sum_{k \geqslant 0} e^{-\lambda t} \frac{(\lambda t)^k}{k!} P^k.$$

Sea ahora $f \in B(E)$. Por la propiedad de Markov de Y' tenemos

$$\mathbb{E}[f(Y'_{k+l}) \mid Y'_0, \ldots, Y'_l] = P^k f(Y'_l) \quad \text{para } k \in \{0, 1, \ldots\}.$$

Afirmamos que

$$\mathbb{E}[f(Y'_{k+N_t}) \mid \mathcal{F}_t] = P^k f(X'_t) \quad \text{para } k \in \{0, 1, \ldots\} \text{ y } t \geqslant 0,$$

donde $\mathcal{F}_t = \sigma(N_s, s \leqslant t) \vee \sigma(X'_s, s \leqslant t)$. Para ver esto, consideremos a $A \in \sigma(N_s, s \leqslant t)$ y $B \in \sigma(Y'_k, k \leqslant l)$, y veamos que

$$\begin{aligned}
\mathbb{E}[1_{A \cap B \cap \{N_t = l\}} f(Y'_{k+N_t})] &= \mathbb{E}[1_{A \cap B \cap \{N_t = l\}} f(Y'_{k+l})] \\
&= \mathbb{P}(A \cap \{N_t = l\}) \mathbb{E}[f(Y'_{k+l}) \mathbb{1}_B] \\
&= \mathbb{P}(A \cap \{N_t = l\}) \mathbb{E}[\mathbb{1}_B P^k f(Y'_l)] \\
&= \mathbb{E}[\mathbb{1}_{A \cap B \cap \{N_t = l\}} P^k (Y'_l)] \\
&= \mathbb{E}[\mathbb{1}_{A \cap B \cap \{N_t = l\}} P^k f(X'_t)].
\end{aligned}$$

Como

$$\left\{A \cap B \cap \{N_t = l\} : A \in \sigma(N_s, s \leqslant t), B \in \sigma(Y'_k, k \leqslant l), l \in \mathbb{N}_0\right\}$$

es cerrado bajo intersecciones finitas y genera a $\mathcal{F}_t$, por un argumento de clases monótonas tenemos

$$\boxed{\mathbb{E}[\mathbb{1}_\Lambda f(Y'_{k+N_t})] = \mathbb{E}[\mathbb{1}_\Lambda P^k f(X'_t)] \quad \text{para } \Lambda \in \mathcal{F}_t.}$$

Finalmente, como N tiene incrementos independientes tenemos

$$\begin{aligned}\mathbb{E}[f(X'_{t+s}) \mid \mathcal{F}_t] = \mathbb{E}[f(Y'_{N_{t+s}-N_t+N_t}) \mid \mathcal{F}_t] &= \sum_{k\geqslant 0} \frac{e^{-\lambda s}(\lambda s)^k}{k!} \mathbb{E}[f(Y'_{N_t+k}) \mid \mathcal{F}_t] \\ &= \sum_{k\geqslant 0} \frac{e^{-\lambda s}(\lambda s)^k}{k!} P^k f(X'_t) = P_s f(X'_t)\end{aligned}$$

para $s, t \geqslant 0$. Luego, X' es un proceso de Markov en E con distribución inicial ν, semigrupo $\{P_t : t \geqslant 0\}$ y generador $\mathcal{A}$.

6.8. Difusiones

Llamaremos *difusión* a un proceso de Markov con trayectorias en $\mathbb{R}^d$ (o de manera más general en una variedad) con trayectorias continuas. La forma más general del generador infinitesimal de una difusión es

$$\boxed{\mathcal{G}f = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\frac{\partial^2}{\partial x_i \partial x_j} f + \sum_{i=1}^{d} b_i \frac{\partial}{\partial x_j} f, \quad f \in \mathcal{C}^2(\mathbb{R}^d),}$$

y los coeficientes a_{ij} y b_j son continuos y $\{a_{ij}\}_{i,j}$ es simétrica positiva. Decimos que $\mathcal{G}$ es un *operador elíptico*. En la práctica construimos a las difusiones resolviendo ecuaciones diferenciales estocásticas a las cuales están asociadas de forma natural.

Recordemos el siguiente teorema.

21. Teorema:
Sean $\{\sigma_{ij}\}_{i,j}$ y $\{b_j\}_j$ funciones globalmente Lipschitz y para $x \in \mathbb{R}^d$, $\mathbb{P}_x$ la ley de la solución de

$$X_t^{(j)} = x_j + \sum_{i=1}^{d} \int_0^t \sigma_{ij}(X'_s)\, \mathrm{d}B_s^{(i)} + \int_0^t b_j(X_j)(X_s)\, \mathrm{d}s,$$

donde $B^{(1)}, \dots, B^{(d)}$ son movimientos brownianos estándar independientes. El proceso solución $\{X, \mathbb{P}_x\}_{x\in\mathbb{R}^d}$ es fuertemente markoviano (incluso Feller) y con generador $\mathcal{G}$ donde $a_{ij} = (\sigma\sigma^t)_{ij}$.

Recordemos que podemos identificar al generador infinitesimal gracias a la fórmula de Itô, ya que para $f \in \mathcal{C}^2(\mathbb{R}^d)$ tenemos que

$$f(X_t) - f(x) - \int_0^t \mathcal{G}f(X_t)\, \mathrm{d}s \quad \text{es una martingala local.}$$

Estamos interesados en el caso $d = 1$, y para ello vamos a considerar un intervalo I con interior I°. Llamaremos *difusión* en I a un proceso de Markov con trayectorias continuas en I. En general vamos a suponer que la difusión es regular en el sentido de que para toda $x, y \in I$, $\mathbb{P}_x(T_y < \infty) > 0$, donde $T_y = \inf\{t > 0 : X_t = y\}$. Como estamos considerando difusiones que son fuertemente markovianas, en consecuencia tendremos que los puntos son instantáneos; es decir que para $x \in I$,

$$\mathbb{P}_x(T_{I\setminus\{x\}} > 0) = 0.$$

De hecho, esta probabilidad vale 0 o 1 gracias a la ley 0-1 de Blumenthal. Cuando aplicamos la propiedad de Markov fuerte al tiempo de paro $T_{I\setminus\{x\}}$, como las trayectorias son continuas, el proceso aún se encuentra en x en dicho tiempo. De la identidad

$$T_{I\setminus\{x\}} = T_{I\setminus\{x\}} + T_{I\setminus\{x\}} \circ \theta_{T_{I\setminus\{x\}}} \tag{$\star$}$$

y suponiendo $\mathbb{P}(T_{I\setminus\{x\}} > 0) = 1$, vemos claramente que ($\star$) se satisface si $T_{I\setminus\{x\}} = \infty$, lo cual contradice la hipótesis de regularidad.

19 de mayo

25. Proposición:
Si $K \subset I^\circ$ es un compacto, entonces para toda $x \in I$,

$$\mathbb{P}\Big(X_t \in K \textit{ para } t \textit{ suficientemente grande}\Big) = 0.$$

Demostración: Definamos

$$\begin{aligned}\Lambda &= \Big\{\omega : X_t(\omega) \in K \text{ para } t \text{ suficientemente grande}\Big\}\\ &= \bigcap_{n\geqslant 0}\{\omega : X_t(\omega) \in K,\ \text{para toda } t \geqslant n\}.\end{aligned}$$

Sea $\mathcal{F}_t = \sigma(X_u, u \leqslant t)$ y observemos que

$$M_t = \mathbb{P}_x(\Lambda \mid \mathcal{F}_t)$$

es martingala positiva y acotada. Además, al aplicar la propiedad de Markov al tiempo t, tenemos $f(X_t) = M_t$, donde $f(x) = \mathbb{P}_x(\Lambda)$. Del teorema de convergencia de martingalas deducimos

$$M_t \xrightarrow[t\to\infty]{} \mathbb{1}_\Lambda \quad \mathbb{P}_x\text{-casi seguramente}$$

y en particular

$$f(X_t) \xrightarrow[t\to\infty]{} \mathbb{1}_\Lambda \quad \mathbb{P}_x\text{-casi seguramente.}$$

Supongamos que $\mathbb{P}_x(\Lambda) > 0$, digamos $\mathbb{P}_x(\Lambda) \geqslant \varepsilon > 0$. Esto implica que podemos encontrar puntos $y \in I$ tales que $f(y)$ es cercano a 1. Si definimos

$$f_n(y) = \mathbb{P}_y(X_t \in K \quad \text{para toda } t \geqslant n),$$

buscaremos probar que tal cantidad es mayor o igual que $1-\varepsilon$. Al aplicar la propiedad de Markov al tiempo n tenemos que existe $z \in K$ tal que

$$\mathbb{P}_z(X_t \in K \text{ para toda } t \geqslant 0) \geqslant 1 - 2\varepsilon. \qquad \blacksquare$$

24 de mayo

33. Definición:
Llamaremos *función de escala* a una función $S : I^\circ \to \mathbb{R}$ continua estrictamente creciente y tal que para todo intervalo $[a,b] \Subset I$, $S(X_{t\wedge T_{[a,b]^c}})$ con $t \geqslant 0$, es una martingala, donde

$$T_{[a,b]^c} = \inf\Big\{s : X_s \notin [a,b]\Big\}.$$

26. Proposición:
La función de escala permite solucionar el problema de doble salida, i.e. *para* $x \in [a,b]$*,*

$$\mathbb{P}_x(T_b < T_a) = \frac{S(x)-S(a)}{S(b)-S(a)}.$$

Demostración: Aplicamos el teorema de paro de Doob y deducimos que

$$S(x) = \mathbb{E}_x[S(X_{t\wedge T_{[a,b]^c}})],$$

lo cual implica, haciendo t tender a infinito, que

$$S(x) = \mathbb{E}_x[S(X_{T_{[a,b]^c}})] = s(a)\mathbb{P}_x(T_a < T_b) + s(b)\mathbb{P}_x(T_b < T_a).$$ ■

Observemos que si S es función de escala, entonces para toda $\alpha > 0$ y $\beta \in \mathbb{R}$, la función $\widetilde{S}(x) = \alpha S(x) + \beta$ es también de escala. No puede haber otra ya que si $\widetilde{S}$ es función de escala, resuelve el problema de doble salida, entonces

$$\frac{\widetilde{S}(x)-\widetilde{S}(a)}{\widetilde{S}(b)-\widetilde{S}(a)} = \frac{S(x)-S(a)}{S(b)-S(a)},$$

de modo que

$$\widetilde{S}(x) = \underbrace{\frac{\widetilde{S}(b)-\widetilde{S}(a)}{S(b)-S(a)}}_{\alpha} S(x) + \underbrace{\left(\widetilde{S}(a) - \frac{\widetilde{S}(b)-\widetilde{S}(a)}{S(b)-S(a)} S(a)\right)}_{\beta}.$$

Si $\widetilde{S}$ es continua, estrictamente creciente y resuelve el problema de doble salida, entonces $\widetilde{S}$ es función de escala. En este caso tenemos que

$$\mathbb{P}_x(T_b < T_a \mid \mathcal{F}_t)$$

es una martingala que satisface

$$\mathbb{P}_x(T_b < T_a \mid \mathcal{F}_t) = \begin{cases} 0 & \text{si } T_a < t\wedge T_b, \\ 1 & \text{si } T_b < t \wedge T_a, \\ \dfrac{\widetilde{S}(X_t)-\widetilde{S}(a)}{\widetilde{S}(b)-\widetilde{S}(a)} & \text{en otro caso.} \end{cases}$$

De lo anterior se sigue que

$$\widetilde{S}(X_{t\wedge T_{[a,b]^c}}) \quad \text{es martingala.}$$

Supongamos que la difusión X es solución de la ecuación diferencial estocástica

$$\mathrm{d}X_t = \sigma(X_t)\ \mathrm{d}B_t + b(X_t)\ \mathrm{d}t,$$

con $b, \sigma : \mathbb{R} \to \mathbb{R}$ dos funciones borelianas y localmente acotadas. Supongamos además que existe una función de clase C^2, $S : \mathbb{R} \to \mathbb{R}$ tal que satisface la ecuación diferencial

$$\boxed{\frac{1}{2}S''\sigma^2 + S'b = 0.}$$

Definimos a $Y_t = S(X_t)$ y apliquemos Itô para ver

$$\begin{aligned}
\mathrm{d}Y_t &= S'(X_t)\ \mathrm{d}X_t + \frac{1}{2}S''(X_t)\ \mathrm{d}\langle X\rangle_t \\
&= S'(X_t)\sigma(X_t)\ \mathrm{d}B_t + S'(X_t)b(X_t)\ \mathrm{d}t + \frac{1}{2}S''(X_t)\sigma^2(X_t)\ \mathrm{d}t \\
&= S'(X_t)\sigma(X_t)\ \mathrm{d}B_t,
\end{aligned}$$

por lo cual deducimos que Y es martingala local. Si además suponemos que σ es continua y que σ y S' no se anulan, entonces para toda $a < x < b$,

$$T_{[a,b]^c} < \infty \quad \text{casi seguramente.}$$

Esto se sigue del teorema de Dubin–Schwarz, el cual nos dice que existe $C = \{C_t : t \geqslant 0\}$ un movimiento browniano, tal que

$$Y_t = S(X_t) = C_{\langle Y\rangle_t}, \quad \text{donde} \quad \langle Y\rangle_t = \int_0^t (S'(X_u)\sigma(X_u))^2\ \mathrm{d}u.$$

Si $\omega \in \{T_{[a,b]^c} = \infty\}$, entonces $\langle Y\rangle_\infty(\omega) = \infty$ y $\limsup_{t\to\infty} Y_t(\omega) = \infty$, lo cual contradice el que S sea acotada en $[a, b]$. Ahora sean

$$H_{S(a)} = \inf\{t : C_t = S(a)\} \quad \text{y} \quad H_{S(b)} = \inf\{t : C_t = S(b)\},$$

entonces

$$\mathbb{P}_x(X_{T_{[a,b]^c}} = a) = \mathbb{P}(H_{S(a)} < H_{S(b)}) = \frac{S(x) - S(a)}{S(b) - S(a)}.$$

34. Definición:
Decimos que una difusión está en *escala natural* si $S(x) = x$ es una función de escala.

Algunos ejemplos de difusiones que están en escala natural son el movimiento browniano y la difusión dada por

$$\mathrm{d}X_t = \sigma(X_t)\ \mathrm{d}B_t.$$

27. Proposición:
Si S es una función de escala para X, entonces el proceso $Y = S(X)$ es una difusión en escala natural.

Demostración: Observemos que $Y = S(X)$ es una martingala local . Por otro lado, Y tiene trayectorias continuas y verifica la propiedad de Markov fuerte, *i.e.* para T tiempo de paro con respecto a la filtración $\{\mathcal{F}_t\}_{t\geqslant 0}$, entonces

$$\begin{aligned}\mathbb{E}_x[f(Y_{T+t}) \mid \mathcal{F}_T] &= \mathbb{E}_x[f \circ S(X_{T+t}) \mid \mathcal{F}_T] \\ &= P_t(f \circ S)(X_T) = P_t(f \circ s) \circ s^{-1}(Y_T),\end{aligned}$$

donde $\{P_t : t \geqslant 0\}$ es el semigrupo de X. De la propiedad anterior vemos que Y satisface la propiedad de Markov fuerte. Lo único que resta probar es la regularidad de Y, pero esta propiedad la hereda de la difusión X ya que s es invertible. ■

26 de mayo

22. Teorema:
Si X es una difusión regular, entonces existe una función de escala.

Demostración: Vamos a construir una función de escala en $[a,b] \Subset I$ con a, b arbitrarios. Definamos para $x \in [a,b]$ a $S(x) = \mathbb{P}_x(T_b < T_a)$. De la propiedad de Markov deducimos que $S(X_{t\wedge T_{[a,b]^c}})$ es una martingala. Para ello notemos que

$$\begin{aligned}\mathbb{E}_x\left[S(X_{t\wedge T_{[a,b]^c}}) \mid \mathcal{F}_{s\wedge T_{[a,b]^c}}\right] &= \mathbb{E}_x\left[\mathbb{P}_{X_{t\wedge T_{[a,b]^c}}}(T_b < T_a) \mid \mathcal{F}_{s\wedge T_{[a,b]^c}}\right] \\ &= \mathbb{E}_x\left[\mathbb{E}[\mathbb{1}_{\{T_b<T_a\}} \mid \mathcal{F}_{t\wedge T_{[a,b]^c}}] \mid \mathcal{F}_{s\wedge T_{[a,b]^c}}\right] \\ &= S(X_{s\wedge T_{[a,b]^c}}).\end{aligned}$$

Además S es claramente creciente, es decir si $a \leqslant x < y \leqslant b$ entonces

$$S(x) = \mathbb{P}_x(T_b < T_a) = \mathbb{P}_x(T_y < T_a)\mathbb{P}_y(T_b < T_a) < S(y).$$

En lo anterior se usó que $\mathbb{P}_x(T_y < T_a)$ debe ser menor que 1 pues en caso contrario no podríamos llegar a a partiendo de x debido a la regularidad y la propiedad de Markov.

Por último, únicamente probaremos la continuidad por la derecha ya que la continuidad por la izquierda se deduce mediante un argumento similar. Sea $\{x_n\} \subset (a,b)$ una sucesión decreciente que converge a x. Por la monotonía sabemos que

$$\liminf_{n\to\infty} s(x_n) > s(x).$$

Gracias a la ley 0-1 de Blumenthal se tiene que

$$\mathbb{P}_x(T_{(x,b)} = 0) \in \{0,1\} \quad \text{con} \quad T_{(x,b)} = \inf\{t : X_t \in (x,b)\}.$$

Dicha probabilidad no puede ser cero, ya que en tal caso se tendría que $\mathbb{P}_x(T_b < \infty) = 0$ por un argumento similar al de la monotonía. Como $x_n \downarrow x$, si fijamos t, ε arbitrariamente pequeñas, podemos encontrar un n_ε tal que para cada $n > n_\varepsilon$,

$$\mathbb{P}_x(T_{x_n} < T_a) \geqslant 1 - \varepsilon.$$

En particular tenemos $\mathbb{P}_x(T_{x_n} < T_a) \geqslant 1 - 2\varepsilon$ para ε suficientemente pequeña de tal forma que $\mathbb{P}_x(T_a < t) \geqslant \varepsilon$. Lo último se debe gracias a la inclusión

$$\{T_b < T_a\} \supset \{T_{x_n} < t\} \setminus \{T_a < t\}.$$

Finalmente vemos que

$$S(x) = \mathbb{P}_x(T_b < T_a) = \mathbb{P}_x(T_{x_n} < T_a)\mathbb{P}_{x_n}(T_b < T_a) \geqslant (1 - 2\varepsilon)S(x_n),$$

lo cual nos da el resultado deseado. ■

Ahora vamos a considerar a I como un intervalo abierto y X una difusión que toma valores en I. Veremos que la ley de la difusión está completamente caracterizada por la funciones de la forma

$$x \mapsto \mathbb{E}_x[T_{[a,b]^c}].$$

Para lo anterior supondremos que X es Feller. Gracias a la fórmula de Dynkin tenemos que

$$f(X_t) - \int_0^t g(X_s)\,\mathrm{d}s \quad \text{es martingala cuando} \quad g = \mathcal{G}f.$$

Del teorema de paro de Doob tenemos que

$$\begin{aligned} f(x) &= \mathbb{E}_x\left[f(X_{T_{[a,b]^c}}) - \int_0^{T_{[a,b]^c}} g(X_s)\,\mathrm{d}s\right] \\ &= \frac{S(x) - S(a)}{S(b) - S(a)} f(b) + \frac{S(b) - S(x)}{S(b) - S(a)} f(a) - \mathbb{E}_x\left[\int_0^{T_{[a,b]^c}} g(X_s)\,\mathrm{d}s\right]. \end{aligned}$$

Al tomar $a = x - \varepsilon$ y $b = x + \varepsilon$ y consideramos $\varepsilon \to 0$, entonces obtenemos la relación asintótica

$$\mathbb{E}_x\left[\int_0^{T_{[a,b]^c}} g(X_s)\,\mathrm{d}s\right] \sim g(x)\mathbb{E}_x[T_{[x-\varepsilon,x+\varepsilon]^c}];$$

en otras palabras tenemos

$$g(x) = \lim_{\varepsilon \to 0} \frac{1}{\mathbb{E}_x[T_{[x-\varepsilon,x+\varepsilon]^c}]} \times \left(\frac{(S(x) - S(x-\varepsilon))f(x+\varepsilon) + (S(x+\varepsilon) - S(x))f(x+\varepsilon)}{S(x+\varepsilon) - S(x-\varepsilon)} - f(x)\right).$$

Dados $a < b$ en I introducimos a la *función de Green*:

$$\boxed{G_{a,b}(x,y) = \begin{cases} \dfrac{2(x-a)(b-y)}{b-a} & \text{si } a \leqslant x \leqslant y \leqslant b, \\ \dfrac{2(y-a)(b-x)}{b-a} & \text{si } a \leqslant y \leqslant x \leqslant b, \\ 0 & \text{en otro caso.} \end{cases}}$$

35. Definición:
Decimos que una medida $m(\mathrm{d}x)$ en I es una medida de velocidad si para toda $x \in I$ se satisface que

$$\boxed{\mathbb{E}_x[T_{[a,b]^c}] = \int_I G_{a,b}(x,y)\ m(\mathrm{d}y).}$$

Se puede ver que la función

$$x \mapsto \int_I G_{a,b}(x,y)\ m(\mathrm{d}y)$$

es una función cóncava que se anula en a y b. Para esto último se tiene que en el sentido distribucional se da la relación

$$\partial_{xx} G_{a,b}(x,y) = \delta_y,$$

de modo que

$$\partial_{xx} \int_I G_{a,b}(x,y)\ m(\mathrm{d}y) = -2m(x).$$

Ejemplo 1 (Movimiento browniano con barreras absorbentes) En este caso $X^2 - \mathrm{id}$ es una martingala, de modo que al aplicar el teorema de paro de Doob deducimos que

$$x^2 = \frac{x-a}{b-a}b^2 + \frac{b-x}{b-a}a^2 - \mathbb{E}_x[T_{[a,b]^c}],$$

y por consiguiente

$$\mathbb{E}_x[T_{[a,b]^c}] = (b-x)(b-a) = \int_I G_{a,b}(x,y)\ \mathrm{d}y.$$

23. Teorema:
La medida de velocidad existe.

Ejemplo 2 $\mathrm{d}X_t = \sigma(X_t)\ \mathrm{d}B_t$. Supondremos que $\sigma \geqslant \varepsilon > 0$. Debemos encontrar una función f para la cual $f(X) - \mathrm{id}$ sea una martingala. Al aplicar $I\,t$ obtenemos

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\sigma(X_s)\ \mathrm{d}B_s + \underbrace{\frac{1}{2}\int_0^t f''(X_s)\sigma^2(X_s)\ \mathrm{d}t}_{\star}.$$

Buscamos que $\star = t$, de donde debe ser el caso que

$$f''(x)\sigma^2(x) = 2.$$

Creyendo ciegamente en el corazón de las cartas obtenemos

$$\mathbb{E}_x[T_{[a,b]^c}] = \frac{(f(b)-f(x))(x-a) - (f(x)-f(a))(b-x)}{b-a},$$

de donde

$$\partial_{xx}\mathbb{E}_x[T_{[a,b]^c}] = -f''(x),$$

de modo que la medida de velocidad resulta ser

$$m(\mathrm{d}x) = \frac{\mathrm{d}x}{\sigma^2(x)}.$$

II. Referencias

Principales

Bingham, N. H., Goldie, C. M., & Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press.

Ethier, S. N., & Kurtz, T. G. (2005). *Markov Processes: Characterization and Convergence*. Wiley Interscience.

Grimmett, G., & Stirzaker, D. (2001). *Probability and Random Processes* (3.ª ed.). Oxford University Press.

Kingman, J. F. C. (1993). *Poisson Processes*. Clarendon Press ; Oxford University Press.

Norris, J. R. (1998). *Markov Chains* (1.ª ed.). Cambridge University Press.

Revuz, D., & Yor, M. (2005). *Continuous Martingales and Brownian Motion*. Springer.

Rogers, L. C. G., Williams, D., & Williams, D. (1994). *Diffusions, Markov Processes, and Martingales: Foundations* (2.ª ed.). Wiley.

Suplementarias

Asmussen, S. (2003). *Applied Probability and Queues* (2.ª ed.). Springer.

Bauer, H. (2001). *Measure and integration theory*. W. de Gruyter.

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

Bertoin, J., Martinelli, F., Peres, Y., & Bernard, P. (1999). *Lectures on Probability Theory and Statistics: Ecole d'Eté de Probabilités de Saint-Flour, XXVII - 1997*. Springer.

Dawson, D. A. (2017). Introductory Lectures on Stochastic Population Systems. *arXiv:1705.03781 [math]*. http://arxiv.org/abs/1705.03781

Dudley, R. M. (2002). *Real Analysis and Probability*. Cambridge University Press.

Goldstein, J. A. (1985). *Semigroups of Linear Operators and Applications*. Oxford University Press.

Lamperti, J. (1977). *Stochastic Processes: A Survey of the Mathematical Theory*. Springer-Verlag.

Munkres, J. R. (2014). *Topology* (2.ª ed.). Pearson.

Pazy, A. (1983). *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer New York. Consultado el 15 de septiembre de 2021, desde https://doi.org/10.1007/978-1-4612-5561-1

Resnick, S. I. (1992). *Adventures in Stochastic Processes*. Birkhäuser.

Schilling, R. L., & Partzsch, L. (2012). *Brownian Motion: An Introduction to Stochastic Processes* (1.ª ed.). De Gruyter.
Schilling, R. L., Song, R., & Vondraček, Z. (2012). *Bernstein Functions: Theory and Applications* (2.ª ed.). De Gruyter.
Tao, T. (2011). *An Introduction to Measure Theory*. American Mathematical Society.
Yoshida, K. (1980). *Functional Analysis* (6.ª ed.). Springer.