PROBLEMAS DE PROCESOS ESTOCÁSTICOS I
SEMESTRE 2013-II
POSGRADO EN CIENCIAS MATEMÁTICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

GERÓNIMO URIBE BRAVO

Problema 1. Sean $\mathscr{F}_1, \mathscr{F}_2, \ldots$ y $\mathscr{G}$ subσ-fields de $\mathscr{F}$. Decimos que $\mathscr{F}_1, \mathscr{F}_2, \ldots$ son condicionalmente independientes dada $\mathscr{G}$ si para cualquier H_i que sea $\mathscr{F}_i$ medible y acotada se tiene que

$$\mathbb{E}(\, H_1 \cdots H_n \mid \mathscr{G}) = \mathbb{E}(\, H_1 \mid \mathscr{G}) \cdots \mathbb{E}(\, H_n \mid \mathscr{G})\, .$$

1. ¿Qué quiere decir la independencia condicional cuando $\mathscr{G} = \{\Omega, \emptyset\}$?

   Demostración. Si $\mathscr{G}$ es la σ-álgebra trivial, entonces $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}(X)$, de modo que la propiedad de independencia condicional se reduce a

   $$\mathbb{E}(H_1 \cdots H_n) = \mathbb{E}(H_1) \cdots \mathbb{E}(H_n).$$

   Sea $A_i \in \mathscr{F}_i$ para toda $i \leq n$. Si elegimos a $H_i = \mathbb{1}_{A_i}$, entonces la anterior condición se traduce a

   $$\mathbb{P}(A_1 \cdots A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n).$$

   Lo anterior es válido para todo $A_i \in \mathscr{F}_i$ para todo $i \leq n$, lo cual corresponde a la definición de independencia (normal) de $\mathscr{F}_1, \ldots, \mathscr{F}_n$, que se puede extender a la independencia de toda la secuencia $\mathscr{F}_1, \mathscr{F}_2, \ldots$. □

2. Pruebe que F_1 y $\mathscr{F}_2$ son condicionalmente independientes dada $\mathscr{G}$ (denotado $\mathscr{F}_1 \perp_{\mathscr{G}} \mathscr{F}_2$) si y sólo si para cualquier H que sea $\mathscr{F}_1$-medible y acotada se tiene que

   $$\mathbb{E}(\, H \mid \mathscr{F}_2, \mathscr{G}) = \mathbb{E}(\, H \mid \mathscr{G})\, .$$

   Demostración. Notemos que el hecho de que H_1 y H_2 sean acotadas, permite trabajar con la esperanza condicional de su producto.

   $\Rightarrow$)Supongamos que $\mathscr{F}_1 \perp_{\mathscr{G}} \mathscr{F}_2$. Sea H una función $\mathscr{F}_1$ -medible y acotada. Definamos $Y := \mathbb{E}(H|\mathscr{G})$, de manera que Y es una función $\mathscr{G}$-medible y acotada.

Entonces bastará demostrar que para todo $A \in \sigma(\mathscr{F}_2 \cup \mathscr{G})$,

$$\mathbb{E}(H\mathbb{1}_A) = \mathbb{E}(Y\mathbb{1}_A).$$

Para generar correctamente a la familia $\sigma(\mathscr{F}_2 \cup \mathscr{G})$, consideremos a

$$\mathcal{C} := \{C \in \sigma(\mathscr{F}_2 \cup \mathscr{G}) : C = F_2 \cap G \text{ donde } F_2 \in \mathscr{F}_2 \text{ y } G \in \mathscr{G}\}.$$

Como $\Omega \in \mathscr{G}$, entonces todos los elementos de $\mathscr{F}_2$ pueden ser expresados mediante su intersección con Ω y por lo tanto, $\mathscr{F}_2 \subset \mathcal{C}$. El mismo argumento sigue para $\mathscr{G}$, de manera que $\sigma(\mathcal{C}) = \sigma(\mathscr{F}_2 \cup \mathscr{G})$. Además, es muy sencillo verificar que $\mathcal{C}$ es una π-sistema.

Después, definamos

$$\mathcal{L} := \{A \in \sigma(\mathcal{C}) : \mathbb{E}(H\mathbb{1}_A) = \mathbb{E}(Y\mathbb{1}_A)\}.$$

La finalidad es demostrar que $\mathcal{L}$ es un λ-sistema y que $\mathcal{C} \subset \mathcal{L}$ para concluir que $\mathcal{L} = \sigma(\mathcal{C}) = \sigma(\mathscr{F}_2 \cup \mathscr{G})$, es decir, que la propiedad deseada es válida para todo $A \in \sigma(\mathscr{F}_2 \cup \mathscr{G})$. A continuación demostremos que $\mathcal{L}$ es un λ-sistema. La primera condición ($\Omega \in \mathcal{L}$) se cumple al notar que

$$\mathbb{E}(H\mathbb{1}_\Omega) = \mathbb{E}(H) = \mathbb{E}(\mathbb{E}(H|\mathscr{G})) = \mathbb{E}(Y) = \mathbb{E}(Y\mathbb{1}_\Omega).$$

Para la segunda condición, sea $A, B \in \mathcal{L}$ tal que $B \subset A$, es decir $\mathbb{1}_{A-B} = \mathbb{1}_A - \mathbb{1}_B$. Entonces

$$\mathbb{E}(H\mathbb{1}_{A-B}) = \mathbb{E}(H\mathbb{1}_A) - \mathbb{E}(H\mathbb{1}_B) = \mathbb{E}(Y\mathbb{1}_A) - \mathbb{E}(Y\mathbb{1}_B) = \mathbb{E}(Y\mathbb{1}_{A-B}),$$

por lo que $A - B \in \mathcal{L}$. Para la tercera y última condición, sea $A_1, A_2, \dots \in \mathcal{L}$ tal que $A_i \subset A_{i+1}$, de manera que $\{\mathbb{1}_{A_i}\}_{i\in\mathbb{N}}$ es una sucesión creciente que converge a $\mathbb{1}_{\cup A_i}$. Entonces usando convergencia monótona,

$$\mathbb{E}(H\mathbb{1}_{\cup A_i}) = \lim_{i\to\infty} \mathbb{E}(H\mathbb{1}_{A_i}) = \lim_{i\to\infty} \mathbb{E}(Y\mathbb{1}_{A_i}) = \mathbb{E}(Y\mathbb{1}_{\cup A_i}),$$

por lo que $\cup A_i \in \mathcal{L}$, y queda demostrado que $\mathcal{L}$ es un λ-sistema.

Para finalizar esta parte de la demostración, resta demostrar que $\mathcal{C} \subset \mathcal{L}$. Sea $C \in \mathcal{C}$ tal que $C = F_2 \cap G$ donde $F_2 \in \mathscr{F}_2$ y $G \in \mathscr{G}$. Entonces

$$\begin{aligned}
\mathbb{E}(H\mathbb{1}_C) &= \mathbb{E}(H\mathbb{1}_{F_2}\mathbb{1}_G) = \mathbb{E}(\mathbb{E}(H\mathbb{1}_{F_2}\mathbb{1}_G|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H\mathbb{1}_{F_2}|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{1}_G\left(\mathbb{E}(H|\mathscr{G})\mathbb{E}(\mathbb{1}_{F_2}|\mathscr{G})\right)) \qquad \text{(Por ind. cond.)} \\
&= \mathbb{E}(\mathbb{1}_G Y\mathbb{E}(\mathbb{1}_{F_2}|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}_G Y\mathbb{1}_{F_2}|\mathscr{G})) \qquad \text{(Por } \mathscr{G} \text{ medibilidad de } Y \text{ y } G) \\
&= \mathbb{E}(\mathbb{1}_G Y\mathbb{1}_{F_2}) = \mathbb{E}(Y\mathbb{1}_C),
\end{aligned}$$

demostrando que $\mathcal{C} \subset \mathcal{L}$ y concluyendo que la propiedad deseada es válida para todo $A \in \sigma(\mathscr{F}_2 \cup \mathscr{G})$.

$\Leftarrow$) A continuación supongamos que $\mathbb{E}(H_1|\mathscr{F}_2, \mathscr{G}) = \mathbb{E}(H_1|G)$ para toda función H_1 que sea $\mathscr{F}_1$-medible. Entonces

$$\begin{aligned}
\mathbb{E}(H_1 H_2|\mathscr{G}) &= \mathbb{E}(\mathbb{E}(H_1 H_2|\sigma(\mathscr{F}_2 \cup \mathscr{G}))|\mathscr{G}) \qquad \text{(Por propiedad de torre)} \\
&= \mathbb{E}(H_2 \mathbb{E}(H_1|\sigma(\mathscr{F}_2 \cup \mathscr{G}))|\mathscr{G}) \\
&\quad (H_2 \text{ es } \mathscr{F}_2\text{-medible y entonces será } \sigma(\mathscr{F}_2 \cup \mathscr{G})\text{-medible}) \\
&= \mathbb{E}(H_2 \mathbb{E}(H_1|\mathscr{G})|\mathscr{G}) \qquad \text{(Por hipótesis)} \\
&= \mathbb{E}(H_1|\mathscr{G})\mathbb{E}(H_2|\mathscr{G}) \qquad (\text{Por } \mathscr{G}\text{-medibilidad de } \mathbb{E}(H_1|\mathscr{G})),
\end{aligned}$$

finalizando la demostración. □

3. Pruebe que $\mathscr{F}_1, \mathscr{F}_2, \ldots,$ son condicionalmente independientes dada $\mathscr{G}$ si y sólo si para cada $n \geq 1$, $\mathscr{F}_{n+1}$ es condicionalmente independiente de $\mathscr{F}_1, \ldots, \mathscr{F}_n$ dada $\mathscr{G}$.

Demostración. $\Rightarrow$) Supongamos que $\mathscr{F}_1, \mathscr{F}_2, \ldots$ son cond. ind. dada $\mathscr{G}$. Definamos $\mathscr{F}^n := \mathscr{F}_1 \cup \cdots \cup \mathscr{F}_n$: Por el inciso 2), solo será necesario demostrar que para cualquier función H_{n+1} que sea $\mathscr{F}_{n+1}$-medible y acotada, se cumple que

$$\mathbb{E}(H_{n+1}|\sigma(\mathscr{F}^n \cup \mathscr{G})) = \mathbb{E}(H_{n+1}|\mathscr{G}).$$

Definamos $Y := \mathbb{E}(H_{n+1}|\mathscr{G})$.

$\mathcal{C} := \{C \in \sigma(\mathscr{F}^n \cup \mathscr{G}) : C = F_1 \cap \cdots \cap F_n \cap G$ donde $F_i \in \mathscr{F}_i$ y $G \in \mathscr{G}\}$.

Claramente $\mathcal{C}$ es un π sistema, y por los mismos argumentos que en el inciso 2), $\sigma(\mathcal{C}) = \sigma(\mathscr{F}^n \cup \mathscr{G})$. Siendo así, definamos

$$\mathcal{L} := \{A \in \sigma(\mathcal{C}) : \mathbb{E}(H_{n+1}\mathbb{1}_A) = \mathbb{E}(Y\mathbb{1}_A)\}.$$

En el inciso 2) se demostró para un conjunto casi idéntico que este es un λ-sistema. Por lo tanto, si demostramos que $\mathcal{C} \subset \mathcal{L}$, entonces se tendrá como conclusión que $\mathcal{L} = \sigma(\mathcal{C})$ y por lo tanto que la propiedad deseada es válida en todo $\sigma(\mathcal{C})$. Entonces, sea $C \in \mathcal{C}$ tal que $C = F_1 \cap \cdots \cap F_n \cap G$ donde $F_i \in \mathscr{F}_i$ y $G \in \mathscr{G}$.

Entonces

$$\begin{aligned}
\mathbb{E}(H_{n+1}\mathbb{1}_C) &= \mathbb{E}(H_{n+1}\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}\mathbb{1}_G) = \mathbb{E}(\mathbb{E}(H_{n+1}\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}\mathbb{1}_G|\mathscr{G}))\\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H_{n+1}\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}|\mathscr{G}))\\
&= \mathbb{E}(\mathbb{1}_G\left(\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{E}(\mathbb{1}_{F_1}|\mathscr{G})\cdots\mathbb{E}(\mathbb{1}_{F_n}|\mathscr{G})\right)) \qquad \text{(Por ind. cond.)}\\
&= \mathbb{E}(\mathbb{1}_G Y\mathbb{E}(\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}|\mathscr{G})) \qquad \text{(Por def. de } Y \text{ y por ind. cond.)}\\
&= \mathbb{E}(\mathbb{E}(\mathbb{1}_G Y\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}|\mathscr{G}))\\
&\quad \text{(Por } \mathscr{G} \text{ medibilidad de } Y \text{ y } G)\\
&= \mathbb{E}(\mathbb{1}_G Y\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n})\\
&= \mathbb{E}(Y\mathbb{1}_C),
\end{aligned}$$

es decir, $C \in \mathcal{L}$, entonces $\mathcal{L} = \sigma(\mathcal{C}) = \sigma(\mathscr{F}^n \cup \mathscr{G})$ y la propiedad deseada se cumplirá en todo $\sigma(\mathscr{F}^n \cup \mathscr{G})$.

$\Leftarrow$)Ahora supongamos que para todo $n \geq 2$, $\mathscr{F}_n \perp_{\mathscr{G}} \sigma(\mathscr{F}^{n-1} \cup \mathscr{G})$, es decir, que para cualquier función H_n que sea $\mathscr{F}_n$-medible, se cumple

$$\mathbb{E}(H_{n+1}|\sigma(\mathscr{F}^n \cup \mathscr{G})) = \mathbb{E}(H_{n+1}|\mathscr{G}),$$

y queda por demostrar que si H_i es una función $\mathscr{F}_i$-medible para todo $i \leq n$, entonces

$$\mathbb{E}(\,H_1\cdots H_n\;|\mathscr{G}) = \mathbb{E}(\,H_1\;|\mathscr{G})\cdots\mathbb{E}(\,H_n\;|\mathscr{G})\,, \tag{1}$$

lo cual será demostrado por inducción. La base para el caso $n = 2$ está demostrado en el inciso 2). Supongamos entonces que (1) es válido para n. Entonces

$$\begin{aligned}
\mathbb{E}(H_1\cdots H_{n+1}|\mathscr{G}) &= \mathbb{E}(\mathbb{E}(H_1\cdots H_{n+1}|\sigma(\mathscr{F}^n\cup\mathscr{G}))|\mathscr{G}) \qquad \text{(Por propiedad de torre)}\\
&= \mathbb{E}(H_1\cdots H_n\mathbb{E}(H_{n+1}|\sigma(\mathscr{F}_2\cup\mathscr{G}))|\mathscr{G})\\
&\quad (H_1\cdots H_n \text{ son } \sigma(\mathscr{F}^n\cup\mathscr{G})\text{-medibles})\\
&= \mathbb{E}(H_1\cdots H_n\mathbb{E}(H_{n+1}|\mathscr{G})|\mathscr{G}) \qquad \text{(Por hipótesis (no de ind.))}\\
&= \mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{E}(H_1\cdots H_n|\mathscr{G}) \qquad \text{(Por } \mathscr{G}\text{-medibilidad de } \mathbb{E}(H_1|\mathscr{G}))\\
&= \left(\mathbb{E}(H_1|\mathscr{G})\cdots\mathbb{E}(H_n|\mathscr{G})\right)\mathbb{E}(H_{n+1}|\mathscr{G}) \qquad \text{(Por hip. ind.)},
\end{aligned}$$

finalizando la inducción y la prueba.

□

Categorías: Esperanza condicional, Independencia condicional.

Problema 2. Sea μ una distribución de progenie y defina $\tilde{\mu}_j = \mu_{j+1}$. Sea $S = (S_n)$ una caminata aleatoria con distribución de salto $\tilde{\mu}$. Sea k un entero no-negativo y defina recursivamente

$$Z_0 = k = C_0, \quad Z_{n+1} = k + S_{C_n} \quad \text{y} C_{n+1} = C_n + Z_{n+1}.$$

1. Pruebe que $Z_n \geq 0$ para toda n y que si $Z_n = 0$ entonces $Z_{n+1} = 0$.

   Demostración. En el caso en que $Z_n = 0$, se tiene que $0 = k + S_{C_{n-1}}$ y que $C_n = C_{n-1} + 0$, que conjuntamente dan como resultado $S_{C_n} = S_{C_{n-1}} = k$, de manera que por definición, $Z_{n+1} = k + S_{C_n} = k - k = 0$. Es decir, si en algún momento el proceso Z alcanza el nivel 0, se quedará ahí en los siguientes pasos.

   Antes de continuar, notemos que la caminata aleatoria S tiene incrementos, digamos $\tilde{\xi}_i \sim \tilde{\mu}$, con soporte en $\{-1, 0, 1, \dots\}$, o equivalentemente con incrementos $\xi_i - 1$ donde $\xi_i \sim \mu$ con soporte en $\{0, 1, \dots\}$. Para demostrar que $Z_n \geq 0$ para toda n, se hará por inducción. $Z_0 = K \geq 0$, demostrando la base. Ahora supongamos que $Z_n \geq 0$. De nuevo, si Z_n es 0, entonces $Z_{n+1} = 0$. Si $Z_n > 0$, entonces

$$
\begin{aligned}
Z_{n+1} &= k + S_{C_n} = k + S_{C_{n-1}+Z_n} = k + S_{C_{n-1}+Z_n} - Z_n + Z_n \\
&= k + S_{C_{n-1}+Z_n} - (k - S_{C_{n-1}}) + Z_n \\
&= (S_{C_{n-1}+Z_n} - S_{C_{n-1}}) + Z_n \\
&= \sum_{i=1}^{C_{n-1}+Z_n} \tilde{\xi}_i - \sum_{i=1}^{C_{n-1}} \tilde{\xi}_i + Z_n \\
&= \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} \tilde{\xi}_i + Z_n \qquad (2) \\
&\geq \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} (-1) + Z_n \qquad \text{(Por el soporte de } \tilde{\xi}_i\text{)} \\
&= -Z_n + Z_n = 0,
\end{aligned}
$$

   y se concluye que en ese caso, $Z_{n+1} \geq 0$. □

2. Pruebe que C_n es un tiempo de paro para la filtración canónica asociada a S.

   Demostración. Se demostrará por inducción. El caso $C_0 = k$ es un tiempo de paro respecto a cualquier filtración, quedando demostrada la base. Entonces supongamos que C_n es un tiempo de paro respecto a la filtración asociada a S, digamos $\mathscr{F}_m = \sigma(S_i, i \leq m)$. Notemos que

$$
\begin{aligned}
\{C_{n+1} = j\} = \{C_n + Z_{n+1} = j\} &= \cup_{i=0}^{j} \{C_n = i, Z_{n+1} = j - i\} \\
&= \cup_{i=0}^{j} \{C_n = i, k + S_{C_n} = j - i\} \\
&= \cup_{i=0}^{j} \{C_n = i, k + S_i = j - i\} \\
&= \cup_{i=0}^{j} \left[\{C_n = i\} \cap \{S_i = j - i - k\}\right]
\end{aligned}
$$

donde $\{C_n = i\} \in \mathscr{F}_i \subset \mathscr{F}_j$ por hipótesis de inducción y $\{S_i = j - i - k\} \in \mathscr{F}_i \subset \mathscr{F}_j$ por definición de $\mathscr{F}_m$. La unión de estos eventos para todo $i \geq j$ se mantiene en $\mathscr{F}_j$, por lo tanto $\{C_{n+1} = j\} \in \mathscr{F}_j$, entonces C_{n+1} es un tiempo de paro respecto a $\mathscr{F}_m$, finalizando la inducción. □

3. Pruebe que Z es un proceso de Galton-Watson con ley de progenie μ.

Demostración. Z es un proceso de Galton-Watson si y sólo si

$$Z_0 = k \text{ y } Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{i,n},$$

para alguna colección $\{\xi_{i,n}\}_{i,n\in\mathbb{N}}$ v.a.i.i.d. con distribución μ y con soporte en $\mathbb{Z}_+$. Para comprobar la segunda propiedad, basta sustituir en S los incrementos $\tilde{\xi}_i \sim \tilde{\mu}$ por $\xi_i - 1$ donde $\xi_i \sim \mu$, descritos en el inciso 1). Es decir, por (2) se tendrá que

$$\begin{aligned}
Z_{n+1} &= \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} \tilde{\xi}_i + Z_n \\
&= \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} (\xi_i - 1) + Z_n \\
&= \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} \xi_i - Z_n + Z_n \\
&= \sum_{i=C_{n-1}+1}^{C_{n-1}+Z_n} \xi_i \\
&= \sum_{i=1}^{Z_n} \xi_{i+C_{n-1}},
\end{aligned}$$

de manera que los incrementos de la caminata aleatoria S, en lugar de tomarlos de una sucesión lineal (que es lo que se está haciendo), los podríamos tomar de una colección de dimensión 2. Entonces haciendo $\xi_{i,n} = \xi_{i+C_{n-1}}$, para todo $n \in \mathbb{N}$ y $i \leq Z_n$ (que son justamente las v.a. que usamos y necesitamos), entonces

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{i,n},$$

que tendrá incrementos con distribución μ y soporte en $\{0, 1, \dots\}$, finalizando la demostración. □

4. Pruebe que si S alcanza -1 entonces existe n tal que $Z_n = 0$. Deduzca que si la media de μ es 1 entonces Z se extingue. (Sugerencia: utilice un ejercicio anterior sobre martingalas con saltos acotados hacia abajo.)

   Demostración. Si $\mathbb{E}(\xi) = 1$, entonces $\mathbb{E}(\tilde{\xi}) = 0$. Es decir, S es una caminata aleatoria con saltos integrables cuya media es 0, por lo que usando el problema 9 de la primera tarea, esta caminata oscila y en particular, $\liminf_n S_n = -\infty$. Para demostrar que en algún momento Z alcanza el nivel 0, supongamos que no lo hace, es decir, que $Z_n > 0$ para todo $n \in \mathbb{N}$. Entonces $C_n = C_{n-1} + Z_n > C_{n-1}$, por lo que C es una sucesión estrictamente creciente y por lo tanto tiende a ∞. Además, en estos puntos sabemos que $k + S_{C_n} = Z_{n+1} > 0$, es decir, que $S_{C_n} > -k$ para toda $n \in \mathbb{N}$. Solamente falta estudiar los tiempos entre los instantes C_n y C_{n+1} en la caminata aleatoria S. Notemos que existirán Z_{n+1} saltos entre estos tiempos, entonces para $j \in \{1, \ldots, Z_{n+1}\}$, se tendrá que

$$
\begin{aligned}
S_{C_n+j} &= S_{C_n} + \sum_{i=C_n+1}^{C_n+j} \tilde{\xi}_i \\
&\geq S_{C_n} + \sum_{i=C_n+1}^{C_n+j} (-1) \qquad \text{(Por el soporte de } \tilde{\xi}_i\text{)} \\
&= S_{C_n} - j \leq S_{C_n} - Z_{n+1} = S_{C_n} - (k + S_{C_n}) = -k,
\end{aligned}
$$

   es decir, la cadena S será mayor que $-k$ en los pasos entre los tiempos C_n y C_{n+1} (para toda $n \in \mathbb{N}$), y por lo tanto se tendrá que $S_n \geq -k$ para todo $n \in \mathbb{N}$, lo cual es una contradicción ya que S oscila, demostrando que Z debe tomar el valor 0 en algún punto. $\square$

Categorías: Caminatas aleatorias, Procesos de Galton-Watson

Problema 3. El objetivo de este ejercicio es ver ejemplos de cadenas de Markov X y de funciones f tales que $f(X) = (f(X_n), n \in \mathbb{N})$ sean o no cadenas de Markov.

1. Considere el hipercubo n-dimensional $E = \{0,1\}^n$. A E lo pensaremos como la composición de la primera de dos urnas que tienen en total n bolas etiquetadas del 1 al n. Si $x = (x_1, \ldots, x_n) \in E$, interpretaremos $x_i = 1$ como que la bola i está en la urna 1. Considere el siguiente experimento aleatorio: inicialmente la composición de las urnas está dada por x y a cada instante de tiempo escogemos una bola al azar y la cambiamos de urna. Modele esta situación por medio de una cadena de Markov X en E. Sea $f : E \to \{0, \ldots, n\}$ dada por $f(x) = \sum_i x_i$. Pruebe que $f(X) = (f(X_n), n \in \mathbb{N})$ es una cadena de Markov cuya matriz de transición determinará.

Demostración. Sean $y = (y_1, \ldots, y_n)$ y $z = (z_1, \ldots, z_n)$ elementos de E. Lo primero que hay que notar, es que la transición de un estado a otro depende solamente de la constitución actual de las urnas (y no de como estuvieron compuestas anteriormente), es decir, es razonable pensar que se puede modelar mediante una cadena de Markov. El experimento dice que se puede cambiar una bola a la vez en cada paso, es decir, para ir de y a z en un paso, se debe cumplir que haya exactamente un cambio en alguna coordenada de estos vectores, o en otras palabras, que $\sum_{i=1}^{n} |y_i - z_i| = 1$. Las demás transiciones serán imposibles. También hay que notar que para y fijo, existen n vectores z en total que cumplen tal condición (cambiando cada una de las coordenadas), y que la elección de la coordenada a cambiar es equiprobable. Por lo tanto, la matriz de transición para el proceso que deseamos construir tendrá entradas

$$p_{y,z} = \frac{1}{n} \mathbb{1}_{\sum_{i=1}^{n} |y_i - z_i| = 1}$$

para todo $y, z \in E$, tendiendo como distribución inicial $\nu = \delta_x$. Por las notas, se tiene que dados estos parámetros existirá una cadena de Markov, digamos X, que tiene esta matriz de transición y distribución inicial, finalizando la primera parte.

Para ver si el proceso $Y = f(X)$ es una cadena de Markov, basta ver que este modelo es el de las urnas de Ehrenfest, el cual ya fue estudiado en clase y concluimos que su matriz de transición estaba dada por

$$p_{i,j} = \frac{n-i}{n} \delta_{j,i+1} + \frac{i}{n} \delta_{j,i-1}.$$

□

2. Sea $(S_n)_{n \in \mathbb{N}}$ una cadena de Markov con espacio de estados $\mathbb{Z}$ y matriz de transición

$$P_{i,i+1} = p \quad P_{i,i-1} = 1 - p$$

donde $p \in [0, 1]$. Dé una condición necesaria y suficiente para que $(|S_n|, n \in \mathbb{N})$ sea una cadena de Markov.

Demostración. Antes de empezar, calculemos $\mathbb{P}(S_n = i | S_0 = 0)$. Hay que notar que esta probabilidad es positiva solo si n y i son pares o impares y si $n \geq |i|$ (los demás casos tendrán probabilidad 0). Bajo estos supuestos, para que suceda que $S_n = i$, tuvo que suceder que el proceso subió en total $(n+i)/2$ veces y bajo $(n-i)/2$ veces, en cualquier orden. Pero existen justamente

$$\binom{n}{(n+i)/2} = \binom{n}{(n-i)/2}$$

de estas ordenaciones, por lo tanto se tiene que

$$\mathbb{P}(S_n = i | S_0 = 0) = \binom{n}{(n+i)/2} p^{(n+i)/2} q^{(n-i)/2},$$

y al ser una caminata aleatoria, se puede generalizar a

$$\mathbb{P}(S_n = i | S_0 = x_0) = \mathbb{P}(S_n = i - x_0 | S_0 = 0) = \binom{n}{(n+i-x_0)/2} p^{(n+i-x_0)/2} q^{(n-i+x_0)/2}.$$

Sigamos con el estudio de $|S_n|$, al ver que este proceso solamente se puede mover, al igual que S_n, una unidad para arriba o una para abajo en cada paso; estudiaremos el caso cuando este se mueve una unidad para arriba y así también concluiremos, por complementación, el estudio del proceso cuando este se mueve un paso para abajo. Demostraremos que bajo la condición $|S_0| = 0$, la cadena $|S_n|$ es efectivamente una cadena de Markov homogénea. Pero mientras, supongamos que $|S_0| = x_0$. Entonces

$$\begin{aligned}
\mathbb{P}_{x_0}(|S_n| = i+1 || S_n| = i) &= \frac{\mathbb{P}_{x_0}(|S_n| = i+1, |S_n| = i)}{\mathbb{P}_{x_0}(|S_n| = i)} \\
&= \frac{\mathbb{P}_{x_0}(S_n = i+1, S_n = i) + \mathbb{P}_{x_0}(S_n = -i-1, S_n = -i)}{\mathbb{P}_{x_0}(S_n = i) + \mathbb{P}_{x_0}(S_n = -i)} \\
&= \frac{\mathbb{P}_{x_0}(S_n = i)p + \mathbb{P}_{x_0}(S_n = -i)q}{\mathbb{P}_{x_0}(S_n = i) + \mathbb{P}_{x_0}(S_n = -i)} \\
&= \frac{\left(\binom{n}{(n+i-x_0)/2} p^{(n+i-x_0)/2} q^{(n-i+x_0)/2}\right) p + \left(\binom{n}{(n-i-x_0)/2} p^{(n-i-x_0)/2} q^{(n+i+x_0)/2}\right) q}{\left(\binom{n}{(n+i-x_0)/2} p^{(n+i-x_0)/2} q^{(n-i+x_0)/2}\right) + \left(\binom{n}{(n-i-x_0)/2} p^{(n-i-x_0)/2} q^{(n+i+x_0)/2}\right)} \\
&= \frac{\left(\binom{n}{(n+i-x_0)/2} p^{i+1}\right) + \left(\binom{n}{(n-i-x_0)/2} q^{i+1}\right)}{\left(\binom{n}{(n+i-x_0)/2} p^{i}\right) + \left(\binom{n}{(n-i-x_0)/2} q^{i}\right)}
\end{aligned}$$

(Factorizando $p^{(n-i-x_0)/2} q^{(n-i+x_0)/2}$).

Hasta ese punto es donde se puede llegar suponiendo que $S_0 = x_0$; sin embargo, si suponemos que $x_0 = 0$, todas las combinaciones de la expresión anterior son iguales, por lo tanto se cancelan y tendríamos que

$$\mathbb{P}_0(|S_n| = i+1 || S_n| = i) = \frac{p^{i+1} + q^{i+1}}{p^i + q^i},$$

que no depende de n y por lo tanto $|S_n|$ es una cadena de Markov homogénea. □

Categorías: Proyecciones de cadenas de Markov

Problema 4. Sean $\mathbb{P}$ y $\mathbb{Q}$ dos medidas de probabilidad en el espacio canónico $E^{\mathbb{N}}$ para sucesiones con valores en un conjunto a lo más numerable E. Decimos que $\mathbb{Q}$ es **localmente absolutamente continua** respecto de $\mathbb{P}$ si para cada $n \in \mathbb{N}$, $\mathbb{Q}|_{\mathscr{F}_n} \ll \mathbb{P}|_{\mathscr{F}_n}$. Sea

$$D_n = \frac{d\mathbb{Q}|_{\mathscr{F}_n}}{d\mathbb{P}|_{\mathscr{F}_n}}.$$

1. Pruebe que D es una martingala bajo $\mathbb{P}$. Pruebe que si D es uniformemente integrable entonces $\mathbb{Q} \ll \mathbb{P}$.

   Demostración. Por definición de derivada de Radon-Nikodym, D_n es $\mathscr{F}_n$-medible y como $\mathbb{Q}$ es una medida no negativa, entonces D_n es no negativa c.s. y entonces

$$\mathbb{E}(|D_n|) = \mathbb{E}(D_n) = \int D_n d\mathbb{P} = \int D_n d\mathbb{P}|_{\mathscr{F}_n} = \mathbb{Q}|_{\mathscr{F}_n}(\Omega) = \mathbb{Q}(\Omega) = 1,$$

   por lo tanto es integrable. Falta demostrar la condición de martingala, es decir, demostrar que para todo $A \in \mathscr{F}_n$, se tiene que $\mathbb{E}(D_n \mathbb{1}_A) = \mathbb{E}(D_{n+1}\mathbb{1}_A)$ bajo $\mathbb{P}$. Entonces

$$\begin{aligned}
\mathbb{E}(D_n \mathbb{1}_A) &= \int D_n \mathbb{1}_A d\mathbb{P} \\
&= \int D_n \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_n} \qquad (D_n \mathbb{1}_A \text{ es } \mathscr{F}_n\text{-medible}) \\
&= \mathbb{Q}|_{\mathscr{F}_n}(A) = \mathbb{Q}(A) = \mathbb{Q}|_{\mathscr{F}_{n+1}}(A) \qquad (A \in \mathscr{F}_n \subset \mathscr{F}_{n+1}) \\
&= \int D_{n+1} \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_{n+1}} \\
&= \int D_{n+1} \mathbb{1}_A d\mathbb{P} \qquad (D_{n+1} \mathbb{1}_A \text{ es } \mathscr{F}_{n+1}\text{-medible}) \\
&= \mathbb{E}(D_{n+1}\mathbb{1}_A),
\end{aligned}$$

   demostrando que efectivamente, D es una martingala bajo $\mathbb{P}$.

   Si suponemos que D es uniformemente integrable, entonces sabemos que converge c.s. y en L_1 a D_∞, y además $D_n = \mathbb{E}(D_\infty | \mathscr{F}_n)$. Si demostramos que para todo $A \in \mathscr{F}$ se cumple que

$$\mathbb{Q}(A) = \int_A D_\infty d\mathbb{P},$$

   habremos demostrado que $\mathbb{Q} \ll \mathbb{P}$. Lo haremos por el método usual de clases monótonas (no Dynkin). Sea $\mathcal{C} = \cup_n \mathscr{F}_n$. Claramente $\mathcal{C}$ es un álgebra. Sea $A \in \mathcal{C}$, entonces existe $k \in \mathbb{N}$ tal que $A \in \mathscr{F}_k$. Pero como $D_k = \mathbb{E}(D_\infty | F_k)$, en este caso se cumple que

$$\begin{aligned}\mathbb{Q}(A) = \mathbb{Q}|_{\mathscr{F}_k}(A) &= \int D_k \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_k} \\ &= \int D_\infty \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_k} \\ &= \int D_\infty \mathbb{1}_A d\mathbb{P},\end{aligned}$$

y por lo tanto la propiedad deseada es válida en todo $\mathcal{C}$. Definimos

$$\mathcal{L} := \{A \in \sigma(\cup_n \mathscr{F}_n) : \mathbb{Q}(A) = \int_A D_\infty d\mathbb{P}\}.$$

Se acaba de demostrar que el álgebra $\mathcal{C} \subset \mathcal{L}$, por lo tanto si demostramos que $\mathcal{L}$ es una clase monótona, entonces se tendrá que $\mathcal{L} = \sigma(\cup_n \mathscr{F}_n)$ y entonces la propiedad deseada se cumplirá para toda $A \in \sigma(\cup_n \mathscr{F}_n)$, demostrando que efectivamente $q \ll \mathbb{P}$. Entonces sea $A_1, A_2, \cdots \in \mathcal{L}$ una sucesión creciente, es decir, se tiene que $\mathbb{1}_{\cup_i A_i} = \lim_{i\to\infty} \mathbb{1}_{A_i}$. Entonces,

$$\begin{aligned}\mathbb{Q}(\cup_i A_i) = \lim_{i\to\infty} q(A_i) &= \lim_{i\to\infty} \int D_\infty \mathbb{1}_{A_i} d\mathbb{P} \\ &= \int D_\infty \mathbb{1}_{\cup_i A_i} d\mathbb{P} \qquad \text{(Por conv. monótona).}\end{aligned}$$

Lo mismo sucede si consideramos una suceción decreciente (ahí se usa convergencia acotada), de manera que $\mathcal{L}$ es una clase monotona y $\mathcal{L} = \sigma(\cup_n \mathscr{F}_n)$, finalizando la demostración. □

2. Pruebe que si T es un tiempo de paro finito entonces $\mathbb{Q}|_{\mathscr{F}_T} \ll \mathbb{P}|_{\mathscr{F}_T}$.

Demostración. De nuevo, bastará con demostrar que para todo $A \in \mathscr{F}_T$,

$$\mathbb{Q}|_{\mathscr{F}_T}(A) = \int_A D_T d\mathbb{P}|_{\mathscr{F}_T}.$$

Por el problema 4.3 de la primera tarea, sabemos que si T es un tiempo de paro finito, entonces D_T es $\mathscr{F}_T$-medible. Entonces

$$\begin{aligned}
\int D_T \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_T} &= \int D_T \mathbb{1}_A d\mathbb{P} \\
&= \int \sum_{i=1}^{\infty} D_T \mathbb{1}_{A\cap\{T=i\}} d\mathbb{P} \\
&= \int \sum_{i=1}^{\infty} D_i \mathbb{1}_{A\cap\{T=i\}} d\mathbb{P} \\
&= \sum_{i=1}^{\infty} \int D_i \mathbb{1}_{A\cap\{T=i\}} d\mathbb{P} \\
&= \sum_{i=1}^{\infty} \int D_i \mathbb{1}_{A\cap\{T=i\}} d\mathbb{P}|_{\mathscr{F}_i} \\
&\quad (\text{Por } \mathscr{F}_i\text{-medibilidad de } D_i \mathbb{1}_{A\cap\{T=i\}}) \\
&= \sum_{i=1}^{\infty} \mathbb{Q}_{\mathscr{F}_i}(A\cap\{T=i\}) = \sum_{i=1}^{\infty} \mathbb{Q}(A\cap\{T=i\}) \\
&= \mathbb{Q}(A\cap\{T<\infty\}) = \mathbb{Q}(A) = \mathbb{Q}|_{\mathscr{F}_T}(A),
\end{aligned}$$

finalizando la prueba. □

3. Sea $\mathbb{P}^p$ la distribución de una caminata aleatoria simple que comienza en 0 y va de k a $k+1$ con probabilidad p, donde $p \in (0,1)$. Pruebe que $\mathbb{P}^p$ es localmente absolutamente continua respecto de $\mathbb{P}^{1/2}$ y encuentre la martingala D_n asociada.

Demostración. Consideremos una trayectoria válida del proceso hasta el paso n, digamos A_n, que sube $k = k(A_n)$ veces y baja $n-k$. Entonces

$$\mathbb{P}^p|_{\mathscr{F}_n}(A_n) = p^k q^{n-k} \quad \text{y} \quad \mathbb{P}^{1/2}|_{\mathscr{F}_n}(A_n) = 1/2^n,$$

de manera que podemos expresar a

$$\mathbb{P}^p|_{\mathscr{F}_n}(A_n) = (2p)^k (2q)^{n-k} \frac{1}{2^n} = 2^n (p/q)^k (q)^n \mathbb{P}^{1/2}|_{\mathscr{F}_n}(A_n).$$

Sin embargo, k depende de la trayectoria elegida, pero bajo A_n, el numero de saltos hacia arriba es exactamente $(X_n+n)/2$, por lo tanto

$$\mathbb{P}^p|_{\mathscr{F}_n}(A_n) = 2^n (p/q)^{(X_n+n)/2} (q)^n \mathbb{P}^{1/2}|_{\mathscr{F}_n}(A_n).$$

La función $2^n (p/q)^{(X_n+n)/2} (q)^n$ será nuestro candidato para D_n, que de nuevo está definido para toda trayectoria hasta el paso n válida,es decir, estará definida para todo el conjunto

$$\mathcal{C} := \{A_n \in \mathscr{F}_n : A_n \text{ es una trayectoria válida hasta } n\} \cup \emptyset,$$

el cual es un π-sistema que genera a $\mathscr{F}_n$. Consideremos también a la familia

$$\mathcal{L} := \{A_n \in \mathscr{F}_n : \mathbb{P}^p|_{\mathscr{F}_n}(A_n) = \int D_n d\ p^{1/2}|_{\mathscr{F}_n}.$$

Por lo argumentado inicialmente, $\mathcal{C} \subset \mathcal{L}$. Por lo tanto, si demostramos que $\mathcal{L}$ es un λ-sistema, lo cual es bastante parecido (y se omitirá) al procedimiento usado en los ejercicios anteriores (usando propiedades de la integral), concluiremos que efectivamente, la función D_n propuesta es la derivada de Radon-Nikodym de $\mathbb{P}^p|_{\mathscr{F}_n}(A_n)$ respecto a $\mathbb{P}^{1/2}|_{\mathscr{F}_n}(A_n)$, y en conclusión $\mathbb{P}^p|_{\mathscr{F}_n} \ll \mathbb{P}^{1/2}|_{\mathscr{F}_n}$ para todo $n \in \mathbb{N}$, es decir, $\mathbb{P}^p$ es localmente absolutamente continua respecto de $\mathbb{P}^{1/2}$. $\square$

4. Para $a, b > 0$, sea $T = \text{mín}\{n \in \mathbb{N} : X_n \in \{-a, b\}\}$. Pruebe que T y X_T son independientes bajo $\mathbb{P}^{1/2}$. Al utilizar la continuidad absoluta local, pruebe que T y X_T también son independientes bajo $\mathbb{P}^p$. Utilice alguna martingala de ejercicios anteriores para calcular $\mathbb{E}(T^2)$.

Demostración. Claramente en el caso $a \neq b$, T y X_T no son independientes; si $a < b$, basta pensar en $\mathbb{P}^{1/2}(T = a, X_T = b) = 0 \neq \mathbb{P}^{1/2}(T = a)\mathbb{P}^{1/2}(X_T = b) > 0$. Consideraremos entonces el caso $a = b$. Ya que $\mathbb{P}^{1/2}(X_T = a) = \mathbb{P}^{1/2}(X_T = -a)$ (por resultado de clase), para demostrar independencia bajo $\mathbb{P}^{1/2}$, bastará con demostrar que para todo $i \in \mathbb{N}$, se cumple que $\mathbb{P}^{1/2}(X_T = a, T = i) = \mathbb{P}^{1/2}(X_T = -a, T = i)$. Pero esto es casi inmediato al notar que al ser una caminata simétrica, el proceso X es igual en distribución a $-X$ y por lo tanto,

$$\begin{aligned}
\mathbb{P}^{1/2}(X_T = a, T = i) &= \mathbb{P}^{1/2}(X_T = a, \text{mín}\{n \in \mathbb{N} : X_n \in \{-a, a\}\} = i) \\
&= \mathbb{P}^{1/2}(-X_T = a, \text{mín}\{n \in \mathbb{N} : -X_n \in \{-a, a\}\} = i) \\
&\qquad \text{(Por que } X = -X \text{ en dist.)} \\
&= \mathbb{P}^{1/2}(X_T = -a, \text{mín}\{n \in \mathbb{N} : X_n \in \{-a, a\}\} = i) \\
&= \mathbb{P}^{1/2}(X_T = -a, T = i),
\end{aligned}$$

de esta manera la independencia queda demostrada. Para ver que también bajo $\mathbb{P}^p$ son independientes, basta calcular la probabilidad bajo $\mathbb{P}^p$ de los eventos $\{X_T = a, T = i\}$, $\{X_T = a\}$,$\{X_T = -a\}$ y $\{T = i\}$ para toda i razonable, es decir, que suceda con probabilidad positiva. Dado que el evento $\{T < \infty\}$ sucede con probabilidad 1, podemos decir que los anteriores eventos son $\mathscr{F}_T$ medibles, por lo tanto, usando el resultado del último inciso, tenemos que

$$\begin{aligned}
\mathbb{P}^p|_{\mathscr{F}_T}(X_T=a,T=i) &= \mathbb{E}^{1/2}(2^T(p/q)^{(X_T+T)/2}(q)^T\mathbb{1}_{X_T=a,T=i}) \\
&= \mathbb{E}^{1/2}(2^i(p/q)^{(a+i)/2}(q)^i\mathbb{1}_{X_T=a,T=i}) \\
&= 2^i(p/q)^{(a+i)/2}(q)^i\mathbb{P}^{1/2}(X_T=a,T=i) \\
&= 2^i(p/q)^{(a+i)/2}(q)^i\mathbb{P}^{1/2}(X_T=a)\mathbb{P}^{1/2}(T=i) \qquad \text{(Por ind. bajo } \mathbb{P}^{1/2}\text{)} \\
&= 2^i(p/q)^{(a+i)/2}(q)^i\frac{1}{2}\mathbb{P}^{1/2}(T=i)
\end{aligned}$$

y análogamente, se tiene que

$$\mathbb{P}^p|_{\mathscr{F}_T}(X_T=a,T=i) = 2^i(p/q)^{(-a+i)/2}(q)^i\frac{1}{2}\mathbb{P}^{1/2}(T=i).$$

De estas 2 probabilidades, se puede concluir que

$$\mathbb{P}^p|_{\mathscr{F}_T}(T=i) = 2^i(p/q)^{(i)/2}(q)^i\frac{1}{2}\mathbb{P}^{1/2}(T=i)\left((p/q)^{a/2}+(p/q)^{-a/2}\right).$$

Bajo el mismo método,

$$\begin{aligned}
\mathbb{P}^p|_{\mathscr{F}_T}(X_T=a) &= \mathbb{E}^{1/2}(2^T(p/q)^{(X_T+T)/2}(q)^T\mathbb{1}_{X_T=a}) \\
&= \mathbb{E}^{1/2}(2^T(p/q)^{(a+T)/2}(q)^T\mathbb{1}_{X_T=a}) \\
&= (p/q)^{a/2}\mathbb{E}^{1/2}(2^T(pq)^{T/2})\mathbb{P}^{1/2}(X_T=a) \qquad \text{(Por ind. bajo } \mathbb{P}^{1/2}\text{)} \\
&= (p/q)^{a/2}\mathbb{E}^{1/2}(2^T(pq)^{T/2})\frac{1}{2}
\end{aligned}$$

Análogamente,

$$\mathbb{P}^p|_{\mathscr{F}_T}(X_T=a) = (p/q)^{-a/2}\mathbb{E}^{1/2}(2^T(pq)^{T/2})\frac{1}{2}.$$

Resolviendo para $\mathbb{E}^{1/2}(2^T(pq)^{T/2})$ y utilizando ley de la probabilidad total sobre la v.a. X_T, se tiene que

$$\mathbb{E}^{1/2}(2^T(pq)^{T/2}) = \frac{2}{(p/q)^{a/2}+(p/q)^{-a/2}}.$$

De modo que

$$\mathbb{P}^p|_{\mathscr{F}_T}(X_T=a) = \frac{(p/q)^{a/2}}{(p/q)^{a/2}+(p/q)^{-a/2}}$$

y

$$\mathbb{P}^p|_{\mathscr{F}_T}(X_T=-a) = \frac{(p/q)^{-a/2}}{(p/q)^{a/2}+(p/q)^{-a/2}}.$$

Es sencillo verifivar, ya con estas 5 probabilidades, que se cumple que

$$\mathbb{P}^p(X_T=a,T=i) = \mathbb{P}^p(X_T=a)\mathbb{P}^p(T=i)$$

y que

$$\mathbb{P}^p(X_T = -a, T = i) = \mathbb{P}^p(X_T = -a)\mathbb{P}^p(T = i),$$

finalizando la prueba. $\square$

Categorías: Cambio de medida, Caminata aleatoria simple.

Problema 5. Sea N un proceso Poisson de parámetro λ y sea T_n el tiempo de su enésimo salto.

1. Pruebe que condicionalmente a T_2, T_1 es uniforme en $[0, T_2]$.

   Demostración. Como N es un proceso Poisson, $T_1, T_2 - T_1 \sim Exp(\lambda)$ y son independientes, por lo que

   $$f_{T_1,T_2-T_1}(x,y) = \lambda^2 e^{-\lambda(x+y)}\mathbb{1}_{x\geq 0, y\geq 0},$$

   y haciendo una transformación lineal $g : (T_1, T_2 - T_1) \to (T_1, T_2)$, que puede ser representado mediante la matriz

   $$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},$$

   tiene como determinante a 1. Entonces se tendrá que

   $$f_{T_1,T_2}(x,y) = f_{T_1,T_2-T_1}(g^{-1}(x,y)) = f_{T_1,T_2-T_1}(x, y-x) = \lambda^2 e^{-\lambda(y)}\mathbb{1}_{y\geq x\geq 0}.$$

   Además, $T_2 \sim Gamma(2, \lambda)$, es decir,

   $$f_{T_2}(y) = \lambda^2 y e^{-\lambda y}\mathbb{1}_{y\geq 0},$$

   por lo que

   $$f_{T_1|T_2}(x|y) = \frac{f_{T_1,T_2}(x,y)}{f_{T_2}(y)} = \frac{\lambda^2 e^{-\lambda(y)}\mathbb{1}_{y\geq x\geq 0}}{\lambda^2 y e^{-\lambda y}\mathbb{1}_{y\geq 0}} = \frac{\mathbb{1}_{y\geq x\geq 0}}{y},$$

   que se traduce a que bajo el evento $\{T_2 = y\}$, T_1 tendrá una distribución uniforme en el intervalo $[0, y]$. $\square$

2. Pruebe que si W_1 y W_2 son exponenciales de parámetro λ independientes entre si y de una variable uniforme U, entonces $U(W_1 + W_2)$ es una variable aleatoria exponencial de parámetro λ.

   Demostración. Se tiene que $W_1 + W_2 \sim Gamma(2, \lambda)$ y es independiente de U. Entonces consideremos la transformación $g : (a, b) \to (a, ab)$ y su inversa

$g^{-1} : (a, b) \to (a, b/a)$, cuyo Jacobiano es $1/a$. Entonces

$$\begin{aligned}
f_{U(W_1+W_2)}(x) &= \int f_{W_1+W_2,U(W_1+W_2)}(y, x)dy \\
&= \int_0^\infty f_{W_1+W_2}(y) f_U(x/y)(1/y)dy \\
&= \int_0^\infty \mathbb{1}_{1\geq x/y\geq 0} \lambda^2 y e^{-\lambda y} \mathbb{1}_{y\geq 0}(1/y)dy \\
&= \int_0^\infty \lambda^2 e^{-\lambda y} \mathbb{1}_{y\geq x\geq 0} dy \\
&= \int_x^\infty \lambda^2 e^{-\lambda y} dy \\
&= \lambda e^{-\lambda x},
\end{aligned}$$

finalizando la demostración. □

3. Conjeture cómo se generaliza lo anterior con T_n y T_1.

Demostración. De nuevo, consideremos una función $g : (T_1, T_2 - T_1, \ldots, T_n - T_{n-1}) \to (T_1, T_2, \ldots, T_n)$, que puede ser representada mediante la matriz

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix},$$

cuyo Jacobiano vale 1. Entonces

$$\begin{aligned}
f_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t_n) &= f_{T_1,T_2-T_1,\ldots,T_n-T_{n-1}}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1}) \\
&= \prod_{i=1}^n \lambda e^{-\lambda(t_i - i_{i-1})} \mathbb{1}_{t_i - t_{i-1}\geq 0} \qquad \text{(Haciendo } t_0 = 0) \\
&= \lambda^n e^{-\lambda t_n} \mathbb{1}_{0\leq t_1\leq\cdots\leq t_n}.
\end{aligned}$$

Además, como $T_n \sim Gamma(n, \lambda)$,

$$\begin{aligned}
f_{T_1,\ldots,T_{n-1}|T_n}(t_1, \ldots, t_{n-1}|t_n) &= \frac{f_{T_1,\ldots,T_{n-1},T_n}(t_1, \ldots, t_{n-1}, t_n)}{f_{T_n}(t_n)} \\
&= \frac{\lambda^n e^{-\lambda t_n} \mathbb{1}_{0\leq t_1\leq\cdots\leq t_n}}{(1/(n-1)!)\lambda^n t_n^{n-1} e^{-\lambda t_n} \mathbb{1}_{t_n\geq 0}} \\
&= \frac{(n-1)! \mathbb{1}_{0\leq t_1\leq\cdots\leq t_n}}{t_n^{n-1}},
\end{aligned}$$

y entonces

$$\begin{aligned}
f_{T_1|T_n}(t_1|t_n) &= \int\cdots\int \frac{(n-1)!\mathbb{1}_{0\le t_1\le\cdots\le t_n}}{t_n^{n-1}} dt_2\cdots dt_{n-1}\\
&= \frac{(n-1)!}{t_n^{n-1}}\int\cdots\int \mathbb{1}_{0\le t_1\le\cdots\le t_n} dt_2\cdots dt_{n-1}\\
&= \frac{(n-1)!}{t_n^{n-1}}\int\cdots\int \mathbb{1}_{0\le t_1\le\cdots\le t_n} dt_2\cdots dt_{n-1}\\
&= \frac{(n-1)!}{t_n^{n-1}}\mathbb{1}_{0\le t_1\le t_n}\int_{t_1}^{t_n}\cdots\int_{t_1}^{t_3} 1 dt_2\cdots dt_{n-1}\\
&= \frac{(n-1)!}{t_n^{n-1}}\mathbb{1}_{0\le t_1\le t_n}\frac{(t_n-t_1)^{n-2}}{(n-2)!}\\
&= (n-1)\frac{(t_n-t_1)^{n-2}}{t_n^{n-1}}\mathbb{1}_{0\le t_1\le t_n}\\
&= (n-1)\left(1-\frac{t_1}{t_n}\right)\left(\frac{1}{t_n}\right)^{n-2}\mathbb{1}_{0\le t_1\le t_n}.
\end{aligned}$$

□

4. Escriba dos programas en Octave que simulen al proceso de Poisson de parámetro λ en el intervalo $[0,1]$. En uno utilizará sólo variables exponenciales y en el otro puede utilizar una variable Poisson.

Demostración. El primer programa, más que ser simulado en el intervalo $[0,t]$, simula un proceso Poisson que hace n saltos en total. Esto lo hace simulando n v.a. exponenciales e incrementando el valor de N de uno en uno cada que ocurra un salto.

```
n=20;
l=.7;
T=zeros(1,n);
T(1)=exprnd(l);
N=1:n;
fori=2:n
T(i)=T(i-1)+exprnd(l);
end
[Ts, Ns] = stairs (T, N );
plot(Ts,Ns)
```

.

Una realización del anterior código es la Figura 1. Para el segundo código si-

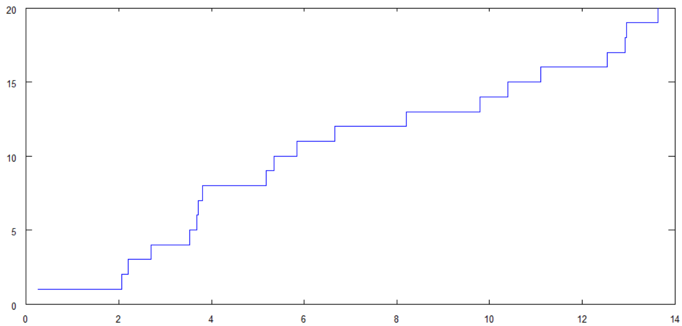

FIGURA 1. Poisson modo 1

mulamos una v.a. $Poi(\lambda)$; con esto condicionamos a que al tiempo t ocurra tal cantidad de saltos. Después se simulan v.a. uniformes sobre $[0, t]$ y se ordenan, indicando los puntos donde N tendrá los saltos.

```
t=30;
l=.7;
n=poissrnd(t*l);
T=zeros(1,n);
T=unifrnd(0,20,1,n);
T=sort(T);
N=1:n;
[Ts, Ns] = stairs (T, N );
plot(Ts,Ns)
```

.

Una realización de este código se encuentra en la Figura 2.

Nota: En un principio creí que se pedía $\lambda \in (0, 1)$, no el proceso N en $[0, 1]$. Cuando ví mi error, ya había cerrado Octave y ya tenía todo listo para integrar la prueba a LATEX; siento la modificación que le hice al ejercicio, pero basta con cambiar en el código a una λ mayor (para que haya más saltos y los podamos notar) y hacer $t = 1$ en el segundo modelo, y para el primero meter un ciclo de simulación de v.a. $Exp(\lambda)$ (e irlas sumando) hasta que el proceso haya superado el tiempo 1, y entonces tomarlo hasta el paso anterior. □

Problema 6. Sea Ξ una medida de Poisson aleatoria en $(0, \infty) \times (0, \infty)$ cuya medida de intensidad ν está dada por $\nu(ds, dx) = \mathbf{1}_{x>0} C/x^{1+\alpha}\, ds\, dx$.

1. Determine los valores de α para los cuales $\int 1 \wedge x\, \nu(dx) < \infty$.

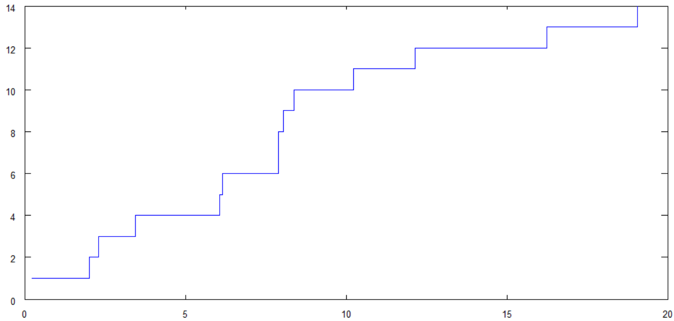

FIGURA 2. Poisson modo 2

Demostración. Se tomará $\nu(\mathrm{d}s, \mathrm{d}x) = \mathbb{1}_{s\geq 0, x>0} C/x^{1+\alpha}\mathrm{d}s\mathrm{d}x$. Para $s < 0$ cualquier integral será exactamente igual a 0. Por lo tanto, elijamos $s \geq 0$. Entonces,

$$\begin{aligned}\int 1 \wedge x\nu(s, \mathrm{d}s) &= \int (1 \wedge x)\mathbb{1}_{s\geq 0, x>0}\frac{C}{x^{1+\alpha}}\mathrm{d}x \\ &= \int_0^\infty (1 \wedge x)\frac{C}{x^{1+\alpha}}\mathrm{d}x \\ &= \int_0^1 x\frac{C}{x^{1+\alpha}}\mathrm{d}x + \int_1^\infty 1\frac{C}{x^{1+\alpha}}\mathrm{d}x \\ &= \int_0^1 \frac{C}{x^\alpha}\mathrm{d}x + \int_1^\infty \frac{C}{x^{1+\alpha}}\mathrm{d}x.\end{aligned}$$

Estudiemos estas dos integrales a continuación. Si $\alpha = 1$, entonces

$$\int_0^1 \frac{C}{x^\alpha}\mathrm{d}x = [C\log(x)]_0^1 = -C \lim_{x\to 0}\log(x),$$

que converge a ∞ y por lo tanto no es un valor posible. Supongamos que $\alpha \neq 1$. Entonces

$$\begin{aligned}\int_0^1 \frac{C}{x^\alpha}\mathrm{d}x &= \int_0^1 Cx^{-\alpha}\mathrm{d}x \\ &= \left[\frac{Cx^{1-\alpha}}{1-\alpha}\right]_0^1 = \frac{C}{1-\alpha} - \lim_{x\to 0}\frac{Cx^{1-\alpha}}{1-\alpha},\end{aligned}$$

de manera que esta integral existe (y es igual a $C/(1-\alpha)$) si y solo si $\alpha < 1$.

Por otro lado, si $\alpha = 0$,

$$\int_1^\infty \frac{C}{x}\mathrm{d}x = C\log_{x\to\infty}\log(x),$$

el cual converge a ∞ y tamoco es una valor posible. Por lo tanto, supongamos que $\alpha \neq 0$. Entonces

$$\int_1^\infty \frac{C}{x^{1+\alpha}}\mathrm{d}x = \int_1^\infty Cx^{-(1+\alpha)}\mathrm{d}x = \left[\frac{Cx^{-\alpha}}{-\alpha}\right]_1^\infty$$
$$= \lim_{x\to\infty}\frac{Cx^{-\alpha}}{-\alpha} + \frac{C}{\alpha},$$

de manera que la integral existe (y es igual a C/α si y solo si $\alpha > 0$.

Por lo tanto, la integral a investigar es finita si y solo si $\alpha \in (0,1)$. $\square$

Nos restringimos ahora a valores de α para los cuales la integral anterior sea finita. Sean $f_t(s,x) = \mathbf{1}_{s\le t}x$ y $X_t = \Xi f_t$.

2. Determine los valores de α para los cuales $X_t < \infty$ para toda $t \geq 0$ casi seguramente.

Demostración. En clase se demostró que Ξf_t es finito c.s. si y sólo si $\int(1\wedge f_t)\mathrm{d}\nu$ es finito. Entonces,

$$\begin{aligned}\int(1\wedge f_t)\mathrm{d}\nu &= \int\int(1\wedge \mathbb{1}_{s\geq t}x)\mathbb{1}_{s\geq 0,x>0}\frac{C}{x^{1+\alpha}}\mathrm{d}x\mathrm{d}s\\ &= \int_0^t\int_0^\infty(1\wedge x)\frac{C}{x^{1+\alpha}}\mathrm{d}x\mathrm{d}s\\ &= \int_0^t\left(\int_0^1\frac{C}{x^\alpha}\mathrm{d}x + \int_1^\infty\frac{C}{x^{1+\alpha}}\mathrm{d}x\right)\mathrm{d}s\\ &= \int_0^t\left(\frac{C}{1-\alpha}+\frac{C}{\alpha}\right)\mathrm{d}s \qquad \text{(Por inciso anterior)}\\ &= t\left(\frac{C}{1-\alpha}+\frac{C}{\alpha}\right) = \frac{Ct}{\alpha(1-\alpha)}.\end{aligned}$$

En conclusión, la restricción $\alpha \in (0,1)$ es necesaria y suficiente. $\square$

Nos restringiremos a dichos valores de α.

3. Calcule $\mathbb{E}\big(e^{-\lambda X_t}\big)$ y pruebe que X_t tiene la misma distribución que $t^{1/\alpha}X_1$.

Demostración. Por clase, sabemos que

$$\begin{aligned}\mathbb{E}(e^{-\lambda X_t}) &= \mathbb{E}(e^{-\Xi[\lambda f_t]}) \qquad (\Xi \text{ es un operador lineal})\\ &= \exp(-\int(1-e^{-\lambda f_t})\mathrm{d}\nu).\end{aligned}$$

Calculando la integral,

$$\begin{aligned}
\int (1-e^{-\lambda f_t})\mathrm{d}\nu &= \int\int (1-e^{-\lambda \mathbb{1}_{s\leq t}x})\mathbb{1}_{s\geq 0, x>0}\frac{C}{x^{1+\alpha}}\mathrm{d}s\mathrm{d}x \\
&= \int_0^\infty \int_0^t (1-e^{-\lambda x})\frac{C}{x^{1+\alpha}}\mathrm{d}s\mathrm{d}x \\
&= Ct\int_0^\infty \frac{1-e^{-\lambda x}}{x^{1+\alpha}}\mathrm{d}x \\
&= Ct\int_0^\infty \frac{\int_0^x \lambda e^{-\lambda r}\mathrm{d}r}{x^{1+\alpha}}\mathrm{d}x \\
&= Ct\lambda\int_0^\infty \int_r^\infty \frac{e^{-\lambda r}}{x^{1+\alpha}}\mathrm{d}x\mathrm{d}r \\
&= Ct\lambda\int_0^\infty e^{-\lambda r}\left[\frac{1}{-\alpha x^\alpha}\right]_r^\infty \mathrm{d}r \\
&= \frac{Ct\lambda}{\alpha}\int_0^\infty \frac{e^{-\lambda r}}{r^\alpha}\mathrm{d}r \\
&= \frac{Ct\lambda}{\alpha}\frac{\lambda^\alpha}{\lambda}\int_0^\infty \frac{e^{-\lambda r}}{(\lambda r)^\alpha}\lambda\mathrm{d}r \qquad \text{(Multiplicando por un 1 conveniente)} \\
&= \frac{Ct\lambda^\alpha}{\alpha}\int_0^\infty (\lambda r)^{(-\alpha+1)-1}e^{-\lambda r}\mathrm{d}(\lambda r) \\
&= \frac{Ct\lambda^\alpha}{\alpha}\Gamma(1-\alpha),
\end{aligned}$$

y por lo tanto,

$$\mathbb{E}(e^{-\lambda X_t}) = \exp\left(-\frac{Ct\lambda^\alpha\Gamma(1-\alpha)}{\alpha}\right),$$

que justamente corresponde a su transformada de Laplace. Tomando $t = 1$ y $\lambda = \lambda' t'^{1/\alpha}$, entonces

$$\mathbb{E}(\exp(-\lambda' t'^{1/\alpha}X_1)) = \exp\left(-\frac{C\cdot 1\cdot(-\lambda' t'^{1/\alpha})^\alpha\Gamma(1-\alpha)}{\alpha}\right) = \exp\left(-\frac{Ct'\lambda'^\alpha\Gamma(1-\alpha)}{\alpha}\right),$$

es decir $\mathbb{E}(\exp(-\lambda(t^{1/\alpha}X_1))) = \mathbb{E}(\exp(-\lambda X_t))$, es decir, sus transformadas de Laplace coinciden, y como son variables aleatorias no negativas c.s., éstas deben tener la misma distribución. □

4. Diga por qué el siguiente código en Octave simula la trayectoria aproximada del proceso X en el intervalo $[0, 1]$.

```
C=1;
```

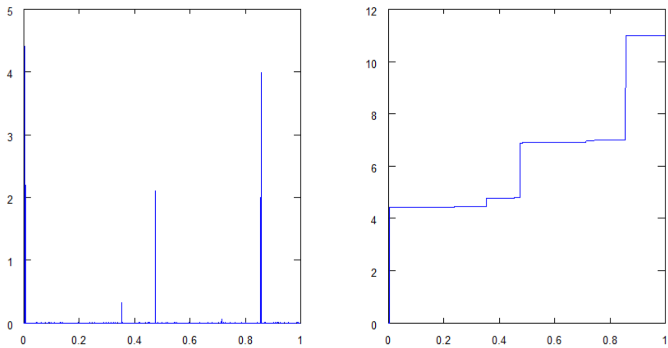

Figura 3. Subordinador

```
e=.000001;
alpha=1/2;
lambda=C/e^alpha;
T=1;
N=poissrnd(lambda*T, 1);
u=T*rand(N,1);
dx=e./rand(N,1).^(1/(alpha));
s=[0;cumsum(dx)];
t=[0;sort(u)];
subplot(1,2,1)
plot(t(2:length(t)),dx)
subplot(1,2,2)
plot(t,s)
```

.

Demostración. Una realización del código anterior está en la Figura 3. Lo que el código hace es primero ubicar los puntos donde habrá saltos, para después decidir el tamaño de estos saltos, los cuales pueden ser chicos o grandes. Sin embargo, los saltos serán de un tamaño mínimo de .000001. $\square$

Problema 7. Pruebe que si X tiene incrementos independientes entonces el proceso X^t dado por $X^t_s = X_{t+s} - X_t$ es independiente de $\mathscr{F}^X_t = \sigma(X_s : s \geq 0)$.

Demostración. Supongamos que X^t tiene incrementos independientes. La σ-álgebra $\mathscr{F}_t = \sigma(X_s, 0 \leq s \leq t)$ es generada por los subconjuntos de la forma

$$\{\omega \in \Omega : X_s(\omega) \in A\}$$

para todo $s \in [0,t]$ y $A \in \mathbb{B}_{\mathbb{R}}$. Más aún, $\mathscr{F}_t$ también es generada por los subconjuntos de la forma

$$\{X_{s_0} \in A_0, X_{s_1} - X_{s_0} \in A_1, \ldots, X_{s_n} - X_{s_{n-1}} \in A_n\}$$

para todo $n \in \mathbb{N}$, $0 \leq s_0 < s_1 < \cdots < s_n \leq t$ y $A_0, \ldots, A_n \in \mathbb{B}_{\mathbb{R}}$. Para ver esto, notemos que los subconjuntos del primer tipo se pueden ver como subconjuntos del segundo tipo si hacemos $A_1, \ldots, A_n = \mathbb{R}$. Ademas, los conjuntos del segundo tipo son $\mathscr{F}_t$-medibles y por lo tanto, generan justamente a $\mathscr{F}_t$. Llamaremos a la familia de subconjuntos del segundo tipo (uniendo al elemento $\{\emptyset\}$) como $\mathcal{C}_t$. Por la discusión anterior se tendrá que $\sigma(\mathcal{C}_t) = \mathscr{F}_t$; además, notemos que $\mathcal{C}_t$ es un π-sistema.

Lo mismo se puede concluir para $\mathscr{F}^t = \sigma(X^t_r, r \geq 0) = \sigma(X_{t+r} - X_t, r \geq 0)$; es decir, que $\mathscr{F}^t$ es generada por subconjuntos de la forma

$$\{X_{t+r_0} - X_t \in B_0, X_{t+r_1} - X_{t+r_0} \in B_1, \ldots, X_{t+r_m} - X_{t+r_{m-1}} \in B_m\}$$

para todo $m \in \mathbb{N}$, $0 \leq r_0 < r_1 < \cdots < r_m \leq t$ y $B_0, \ldots, B_m \in \mathbb{B}_{\mathbb{R}}$. A la familia de subconjuntos de este tipo (unión el elemento $\{\emptyset\}$) le llamaremos $\mathcal{C}^t$, donde también se tendrá que $\sigma(\mathcal{C}^t) = \mathscr{F}^t$ y que $\mathcal{C}^t$ es un π-sistema.

Sea $D_0 \in \mathcal{C}^t$ fijo, es decir,

$$D_0 = \{X_{t+r^0_0} - X_t \in B^0_0, X_{t+r^0_1} - X_{t+r^0_0} \in B^0_1, \ldots, X_{t+r^0_{m^0}} - X_{t+r^0_{m^0-1}} \in B_{m^0}\}.$$

para algún $m^0 \in \mathbb{N}$, $0 \leq r^0_0 < r^0_1 < \cdots < r^0_{m^0} \leq t$ y $B^0_0, \ldots, B^0_{m^0} \in \mathbb{B}_{\mathbb{R}}$. Definamos

$$\mathcal{L}_t := \{C \in \sigma(\mathcal{C}_t) = \mathscr{F}_t : \mathbb{P}(C \cap D_0) = \mathbb{P}(C)\mathbb{P}(D_0)\}.$$

Sea $C_0 \in \mathcal{C}_t$, es decir,

$$C_0 = \{X_{s^0_0} \in A^0_0, X_{s^0_1} - X_{s^0_0} \in A^0_1, \ldots, X_{s^0_{n^0}} - X_{s^0_{n^0-1}} \in A^0_{n^0}\}$$

para algún $n^0 \in \mathbb{N}$, $0 \leq s_0 < s_1 < \cdots < s_{n^0} \leq t$ y $A_0, \ldots, A_{n^0} \in \mathbb{B}_{\mathbb{R}}$. Entonces

$$\begin{aligned}
\mathbb{P}(C_0 \cap D_0) &= \mathbb{P}(\{X_{s_0^0} \in A_0^0, X_{s_1^0} - X_{s_0^0} \in A_1^0, \ldots, X_{s_{n^0}^0} - X_{s_{n^0-1}^0} \in A_{n^0}^0\} \cap \\
&\quad \{X_{t+r_0^0} - X_t \in B_0^0, X_{t+r_1^0} - X_{t+r_0^0} \in B_1^0, \ldots, X_{t+r_{m^0}^0} - X_{t+r_{m^0-1}^0} \in B_{m^0}\}) \\
&= \mathbb{P}(\{X_{s_0^0} \in A_0^0, X_{s_1^0} - X_{s_0^0} \in A_1^0, \ldots, X_{s_{n^0}^0} - X_{s_{n^0-1}^0} \in A_{n^0}^0\} \cap \\
&\quad \{X_t - X_{s_{n^0}^0} \in \mathbb{R}\} \cap \\
&\quad \{X_{t+r_0^0} - X_t \in B_0^0, X_{t+r_1^0} - X_{t+r_0^0} \in B_1^0, \ldots, X_{t+r_{m^0}^0} - X_{t+r_{m^0-1}^0} \in B_{m^0}\}) \\
&= \mathbb{P}(X_{s_0^0} \in A_0^0, X_{s_1^0} - X_{s_0^0} \in A_1^0, \ldots, X_{s_{n^0}^0} - X_{s_{n^0-1}^0} \in A_{n^0}^0) \times \\
&\quad \mathbb{P}(X_t - X_{s_{n^0}^0} \in \mathbb{R}) \times \\
&\quad \mathbb{P}(X_{t+r_0^0} - X_t \in B_0^0, X_{t+r_1^0} - X_{t+r_0^0} \in B_1^0, \ldots, X_{t+r_{m^0}^0} - X_{t+r_{m^0-1}^0} \in B_{m^0}) \\
&= \mathbb{P}(C_0) \times 1 \times \mathbb{P}(D_0),
\end{aligned}$$

de manera que $\mathcal{C}_t \subset \mathcal{L}_t$. Si demostramos que $\mathcal{L}_t$ es un λ-sistema, se tendría que $\mathcal{L}_t = \sigma(\mathcal{C}_t)$; esta prueba es muy sencilla y se ha realizado en repetidas ocasiones; esta vez bastará con utilizar la propiedad de aditividad finita continuidad de la función de probabilidad, sin embargo, la demostración formal será omitida. A continuación, tomemos un $C_0' \in \mathscr{F}_t$ fijo. Definamos

$$\mathcal{L}^t := \{D \in \sigma(\mathcal{C}^t) = \mathscr{F}_t : \mathbb{P}(C_0' \cap D) = \mathbb{P}(C_0')\mathbb{P}(D)\}.$$

Por el resultado anterior, se tiene que $\mathcal{C}^t \subset \mathcal{L}^t$ y de igual manera, se omite la prueba de que $\mathcal{L}^t$ es un λ-sistema (es prácticamente la misma). Así pues, se concluye que $\mathcal{L}_2 = \sigma(\mathcal{C}^t) = \mathscr{F}^t$, lo que quiere decir que para todo $C \in \mathscr{F}_t$ y $D \in \mathscr{F}^t$, se tiene que $\mathbb{P}(C \cap D) = \mathbb{P}(C)\mathbb{P}(D)$, es decir, X^t es independiente de $\mathscr{F}_t$. $\square$

Calcular la esperanza y varianza del proceso de Poisson y de Poisson compuesto (en términos de la intensidad y la distribución de salto). Probar que si X es

$$\mathbb{E}\left(e^{iuZ_t}\right) = e^{-\lambda t(1-\psi(u))} \quad \text{donde} \quad \psi(u) = \mathbb{E}\left(e^{iu\xi_1}\right).$$

Demostración. Sea N_t el proceso Poisson de intesidad λ. Entonces, como $N_t \sim Poi(\lambda t)$, se tiene que

$$\mathbb{E}(N_t) = \sum_{i=1}^{\infty} ie^{\lambda t}\frac{(\lambda t)^i}{i!} = \sum_{i=0}^{\infty} ie^{\lambda t}\frac{(\lambda t)^{i+1}}{i!} = \lambda t.$$

De la misma manera,

$$\begin{aligned}
\mathbb{E}(N_t^2) &= \sum_{i=1}^{\infty} i^2 e^{\lambda t} \frac{(\lambda t)^i}{i!} = \sum_{i=1}^{\infty} i e^{\lambda t} \frac{(\lambda t)^i}{i-1!} \\
&= \sum_{i=0}^{\infty} (i+1) e^{\lambda t} \frac{(\lambda t)^{i+1}}{i!} \\
&= (\lambda t)\mathbb{E}(N_t) + (\lambda t) = (\lambda t)^2 + (\lambda t),
\end{aligned}$$

y por lo tanto

$$Var(N_t) = \mathbb{E}(N_t^2) - \mathbb{E}(N_t)^2 = (\lambda t)^2 + (\lambda t) - (\lambda t)^2 = \lambda t.$$

Sea Z_t el proceso Poisson compuesto con saltos $\{\xi_i\}_i \in \mathbb{N}$ independientes e idénticamente distribuidos, es decir,

$$Z_t = \sum_{i=1}^{N_t} \xi_i.$$

Entonces

$$\begin{aligned}
\mathbb{E}(Z_t) &= \mathbb{E}\left(\sum_{i=1}^{N_t} \xi_i\right) = \mathbb{E}\left(\sum_{i=1}^{\infty} \xi_i \mathbb{1}_{T_i \leq t}\right) \\
&= \sum_{i=1}^{\infty} \mathbb{E}\left(\xi_i \mathbb{1}_{T_i \leq t}\right) = \sum_{i=1}^{\infty} \mathbb{E}(\xi_i)\mathbb{E}\left(\mathbb{1}_{T_i \leq t}\right) \qquad \text{(Por independencia)} \\
&= \sum_{i=1}^{\infty} \mathbb{E}(\xi_1)\mathbb{E}\left(\mathbb{1}_{T_i \leq t}\right) \qquad \text{(Las } \xi_i\text{'s son v.a.i.i.d.)} \\
&= \mathbb{E}(\xi_1) \sum_{i=1}^{\infty} \mathbb{E}\left(\mathbb{1}_{T_i \leq t}\right) = \mathbb{E}(\xi_1)\mathbb{E}\left(\sum_{i=1}^{\infty} \mathbb{1}_{T_i \leq t}\right) \\
&= \mathbb{E}(\xi_1)\mathbb{E}\left(N_t\right) = \mathbb{E}(\xi_1)(\lambda t).
\end{aligned}$$

Para calcular su varianza,

$$
\begin{aligned}
\mathbb{E}(Z_t^2) &= \mathbb{E}\left(\mathbb{E}\left(Z_t^2|N_t\right)\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(\left(\sum_{i=1}^{N_t}\xi_i\right)^2|N_t\right)\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(\sum_{j=0}^{\infty}\mathbb{1}_{N_t=j}\left(\sum_{i=1}^{j}\xi_i\right)^2|N_t\right)\right)\\
&= \mathbb{E}\left(\sum_{j=0}^{\infty}\mathbb{1}_{N_t=j}\mathbb{E}\left(\left(\sum_{i=1}^{j}\xi_i\right)^2|N_t\right)\right) \qquad \text{(Por medibilidad y Fubini)}\\
&= \mathbb{E}\left(\sum_{j=0}^{\infty}\mathbb{1}_{N_t=j}\mathbb{E}\left(\left(\sum_{i=1}^{j}\xi_i\right)^2\right)\right) \qquad \text{(Por independencia de los } \xi_i\text{'s con } N_t)\\
&= \mathbb{E}\left(\sum_{j=0}^{\infty}\mathbb{1}_{N_t=j}\left(Var\left(\sum_{i=1}^{j}\xi_i\right)+\mathbb{E}\left(\sum_{i=1}^{j}\xi_i\right)^2\right)\right)\\
&= \mathbb{E}\left(\sum_{j=0}^{\infty}\mathbb{1}_{N_t=j}\left(j\times Var\left(\xi_1\right)+j^2\times\mathbb{E}\left(\xi_1\right)^2\right)\right) \qquad \text{(Por que las } \xi_i\text{'s son v.a.i.d.d.)}\\
&= \mathbb{E}\left(N_t\times Var\left(\xi_1\right)+N_t^2\times\mathbb{E}\left(\xi_1\right)^2\right)\\
&= \mathbb{E}(N_t)Var\left(\xi_1\right)+\mathbb{E}(N_t^2)\mathbb{E}\left(\xi_1\right)^2\\
&= \mathbb{E}(N_t)Var\left(\xi_1\right)+(Var(N_t)+\mathbb{E}(N_t^2))\mathbb{E}\left(\xi_1\right)^2\\
&= (\lambda t)Var\left(\xi_1\right)+(\lambda t)\mathbb{E}\left(\xi_1\right)^2+\left(\mathbb{E}(N_t)\mathbb{E}\left(\xi_1\right)\right)^2\\
&= (\lambda t)(\mathbb{E}(\xi_1^2))+\left(\mathbb{E}(N_t)\mathbb{E}\left(\xi_1\right)\right)^2,
\end{aligned}
$$

y por lo tanto,

$$Var(Z_t)=\mathbb{E}(Z_t^2)-\mathbb{E}(Z_t)^2=(\lambda t)(\mathbb{E}(\xi_1^2)).$$

Para calcular $\mathbb{E}(e^{iuZ_t})$, se sigue un procedimiento similar, de manera que

$$
\begin{aligned}
\mathbb{E}(e^{iuZ_t}) &= \mathbb{E}\left(\mathbb{E}\left(\exp\left(iu\sum_{j=1}^{N_t}\xi_j\right)|N_t\right)\right)\\
&= \mathbb{E}\left(\mathbb{E}\left(\sum_{k=0}^{\infty}\mathbb{1}_{N_t=k}\exp\left(iu\sum_{j=1}^{k}\xi_j\right)|N_t\right)\right)\\
&= \mathbb{E}\left(\sum_{k=0}^{\infty}\mathbb{1}_{N_t=k}\mathbb{E}\left(\exp\left(iu\sum_{j=1}^{k}\xi_j\right)|N_t\right)\right)\\
&\quad\text{(Medibilidad y Fubini)}\\
&= \mathbb{E}\left(\sum_{k=0}^{\infty}\mathbb{1}_{N_t=k}\mathbb{E}\left(\exp\left(iu\sum_{j=1}^{k}\xi_j\right)\right)\right)\\
&\quad\text{(Independencia con } N_t\text{)}\\
&= \mathbb{E}\left(\sum_{k=0}^{\infty}\mathbb{1}_{N_t=k}\mathbb{E}\left(\exp\left(iu\xi_1\right)\right)^k\right)\\
&\quad\text{(Las } \xi_j\text{'s son v.a.i.i.d.)}\\
&= \mathbb{E}\left(\mathbb{E}\left(\exp\left(iu\xi_1\right)\right)^{N_t}\right) = \mathbb{E}(\psi(u)^{N_t})\\
&= \sum_{i=0}^{\infty}\psi(u)^i \times e^{-\lambda t}\frac{(\lambda t)^i}{i!}\\
&= e^{-\lambda t}\sum_{i=0}^{\infty}\frac{(\psi(u)\lambda t)^i}{i!}\\
&= e^{-\lambda t}e^{\psi(u)\lambda t} = e^{-\lambda t(1-\psi(u))}
\end{aligned}
$$

□

Sea N un proceso de Lévy tal que N_t tiene distribución de parámetro λt.

1. Pruebe que casi seguramente las trayectorias de N son no-decrecientes.

   Demostración. Tendríamos que probar que para todo $s, t \geq 0$, $N_{t+s} - N_t \geq 0$. Sin embargo, esto es directo al utilizar los incrementos estacionarios del proceso de Levy, ya que tenemos que $N_{t+s} - N_t$ es igual en distribución a $N_s \sim Poi(\lambda)$, la cual toma valores en $\mathbb{Z}_+$ con probabilidad 1, demostrando entonces que $N_{t+s} - N_t \geq 0$. Además, este proceso N que toma valores en una cantidad numerable de estados ($\mathbb{Z}_+$) y que es no decreciente, debe tener trayectorias constante a pedazos. Notemos que lo único que falta para afirmar que éste es un proceso Poisson, es que sus incrementos sean de una unidad en una unidad, pero esa es la finalidad de los siguientes incisos. $\square$

2. Sea Ξ la única medida en $\mathscr{B}_{\mathbb{R}_+}$ tal que $\Xi([0,t]) = N_t$. Pruebe que Ξ es una medida de Poisson aleatoria de intensidad $\lambda \cdot \text{Leb}$.

   Demostración. La prueba de este inciso es exactamente la misma que la de la Proposición 4.5 (pág 105); en tal prueba se hace la suposición de que N s un proceso Poisson (cosa que no tenemos como hipótesis) sin embargo, las únicas propiedades que se utiliza para tal demostración son los incrementos independientes y estacionarios (que tenemos porque N es un proceso de Levy), que $N_t \sim Poi(\lambda)$ y que las trayectorias son crecientes y constantes a pedazos (que tenemos por la discusión del inciso anterior). Dicha prueba será omitida en esta tarea por que básicamente sería una copia de la que se encuentra en las notas (con el intercambio de las palabras 'proceso Poisson' por la propiedad que se esté utilizando de las mencionadas anteriormente). $\square$

3. Concluya que N es un proceso de Poisson de intensidad λ.

   Demostración. En la pág 108 de las notas, se demuestra que cualquier medida de Poisson aleatoria $\Xi \in \mathbb{R}_+$ con intensidad $\lambda \times Leb$ corresponde a un proceso de Poisson. Por lo tanto, ya que se demostro que la medida Ξ en el anterior inciso cumple con tales propiedades, debe de suceder que N es un proceso Poisson. $\square$

Problema 8. Sea P_t la probabilidad de transición en t unidades de tiempo para el proceso de Poisson de parámetro λ.

Al utilizar el teorema del biniomio, pruebe directamente que las probabilidades de transición del proceso de Poisson satisfacen las ecuaciones de Chapman-Kolmogorov $P_{t+s} = P_t P_s$. Dé además un argumento probabilístico, basado en condicionar con lo que sucede al tiempo s, para probar dicha ecuación.

Sea

$$Q(i,j) = \begin{cases} -\lambda & j = i \\ \lambda & j = i+1 \\ 0 & j \neq i, i+1 \end{cases}.$$

Pruebe directamente que se satisfacen las ecuaciones de Kolmogorov

$$\frac{d}{dt} P_t(i,j) = QP_t(i,j) = P_t Q(i,j)\,,$$

donde QP_t es el producto de las matrices Q y P_t.

Demostración. Elijamos $i, j \in \mathbb{N}$. Sabemos que

$$P_r(i,j) = \mathbb{P}(N_r = j - i) = e^{-\lambda r} \frac{(\lambda r)^{j-i}}{(j-i)!}$$

cuando $i \leq j$ y 0 en otro caso. Sea entonces $i \leq j$. Entonces,

$$\begin{aligned}
P_{t+s}(i,j) &= e^{-\lambda(t+s)} \frac{(\lambda(t+s))^{j-i}}{(j-i)!} \\
&= e^{-\lambda(t+s)} \frac{\sum_{l=0}^{j-i} \binom{j-i}{l} t^l s^{j-i-l}}{(j-i)!} \\
&= e^{-\lambda(t+s)} \sum_{l=0}^{j-i} \frac{t^l s^{j-i-l}}{l!(j-i-l)!}.
\end{aligned}$$

Si multiplicamos las matrices P_t y P_s, tendríamos que

$$\begin{aligned}
(P_t P_s)(i,j) &= \sum_{l=0}^{\infty} P_t(i,l) P_s(l,j) \\
&= \sum_{l=i}^{j} P_t(i,l) P_s(l,j) \\
&\qquad \text{(Se tiene que } P_t(i,l) = P_s(l,j) = 0 \text{ para todo } l < i \text{ y } j < l) \\
&= \sum_{l=i}^{j} e^{-\lambda t} \frac{(\lambda t)^{l-i}}{(l-i)!} \times e^{-\lambda s} \frac{(\lambda s)^{j-l}}{(j-l)!} \\
&= e^{-\lambda(t+s)} \lambda^{j-i} \sum_{l=i}^{j} \frac{t^{l-i} s^{j-l}}{(l-i)!(j-l)!} \\
&= e^{-\lambda(t+s)} \lambda^{j-i} \sum_{l'=0}^{j-1} \frac{t^{l'} s^{j-i-l'}}{l'!(j-i-l')!} \qquad \text{(Haciendo } l' = l - i).
\end{aligned} \tag{3}$$

Notemos que en el caso $i > j$, los sumandos de (3) son exactamente igual a 0, ya que para todo l, se tendrá que $i > l$ o que $j < l$, y en cualquier caso la multiplicación de probabilidades se anula. Por lo tanto se ha comprobado que las entradas (i,j) para todo $i, j \in \mathbb{N}$ de las matrices P_{t+s} y $P_t P_s$ coinciden y por lo tanto, son idénticas. Probabilisticamente, de nuevo fijandonos sólo en el caso $i \leq j$ (el caso $i > j$ da como resultado probabilidades nulas), se tiene que

$$\begin{aligned}
P_{t+s}(i,j) &= \mathbb{P}(N_{t+s} = j-i) \\
&= \sum_{l=0}^{j-i} \mathbb{P}(N_{t+s} = j-i, N_t = l) \\
&= \sum_{l=0}^{j-i} \mathbb{P}(N_{t+s} - N_t = j-(i+l), N_t = l) \\
&= \sum_{l=0}^{j-i} \mathbb{P}(N_{t+s} - N_t = j-(i+l))\mathbb{P}(N_t = l) \qquad \text{(Por incr. ind.)} \\
&= \sum_{l=0}^{j-i} P_s(i+l,j)P_t(0,l) \\
&= \sum_{l=0}^{j-i} P_s(i+l,j)P_t(i,i+l) \\
&= \sum_{l'=i}^{j} P_s(l',j)P_t(i,l') \qquad \text{(Haciendo } l' = l+i)
\end{aligned}$$

$$= (P_tP_s)(i,j),$$

demostrando las ecuaciones de Chapman-Kolmogorov para el proceso Poisson. Para demostrar las ecuaciones diferenciales de Kolmogorov, veamos que para todo $i \leq j$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}r}P_r(i,j) &= \frac{\mathrm{d}}{\mathrm{d}r}e^{-\lambda r}\frac{(\lambda r)^{j-i}}{(j-i)!} \\
&= e^{-\lambda r}\frac{\lambda^{j-i}}{(j-i)!}(j-i)r^{j-i-1} - \lambda e^{-\lambda r}\frac{(\lambda r)^{j-i}}{(j-i)!} \\
&= \lambda\left(e^{-\lambda r}\frac{(\lambda r)^{j-(i+1)}}{(j-(i+1))!}\right) - \lambda e^{-\lambda r}\frac{(\lambda r)^{j-i}}{(j-i)!} \\
&= \lambda P_r(i+1,j) - \lambda P_r(i,j).
\end{aligned}$$

Por otro lado,

$$\begin{aligned}
(QP_r)(i,j) &= \sum_{l=0}^{\infty} Q(i,l)P_r(l,j) \\
&= Q(i,i)P_r(i,j) + Q(i,i+1)P_r(i+1,j) \\
&= -\lambda P_r(i,j) + \lambda P_r(i+1,j).
\end{aligned}$$

Además,

$$(4) \qquad (P_rQ)(i,j) = \sum_{l=0}^{\infty} P_r(i,l)Q(l,j)$$
$$= P_r(i,j)Q(j,j) + P_r(i,j-1)Q(j-1,j)$$
$$= -\lambda P_r(i,j) + \lambda P_r(i,j-1)$$
$$= -\lambda P_r(i,j) + \lambda P_r(i+1,j).$$

Nuevamente, el caso $j < i$ da probabilidades nulas en todas las sumas (4), lo cual tiene sentido ya que estamos derivando una probabilidad constante (cero). Por lo tanto, se tendrá que

$$\frac{\mathrm{d}}{\mathrm{d}r} P_r = QP_r = P_rQ.$$

$\square$

Problema 9 (Tomado del examen general de probabilidad del Posgrado en Ciencias Matemáticas, UNAM, Febrero 2011)**.** Una planta de producción toma su energía de dos generadores. La cantidad de generadores al tiempo t está representado por una cadena de Markov a tiempo continuo $\{X_t, t \geq 0\}$ con espacio de estados $E = \{0,1,2\}$ y matriz infinitésimal Q dada por

$$Q = \begin{pmatrix} -6 & 6 & 0 \\ 1 & -7 & 6 \\ 0 & 2 & -2 \end{pmatrix}.$$

1. Encuentre la matriz de transición de la cadena de Markov de los estados distintos que toma X, clasifique los estados, diga si existe una única distribución invariante y en caso afirmativo, encuéntrela. Calcule explícitamente las potencias de la matriz de transición. (Recuerde que de ser posible diagonalizar, esta es una buena estrategia.)

   Demostración. Notemos que no hay ningun estado absorbente, por lo tanto,

   $$P(i,j) = \frac{Q(i,j)}{c(i)}(1-\delta_{ij}),$$

   es decir,

   $$P = \begin{pmatrix} 0 & 1 & 0 \\ 1/7 & 0 & 6/7 \\ 0 & 1 & 0 \end{pmatrix}.$$

   Es claro que la trayectoria $0 \to 1 \to 2 \to 1 \to 0$ es válida, y por lo tanto todos los estados se comunican y el proceso será irreducible. Al tener un número finito de estados, será recurrente positiva y por lo tanto existirá una distribución estacionaria $\nu = (\nu_1, \nu_2, \nu_3)$ que resuelva el sistema

$$\nu P = \nu,$$

es decir, querems resolver el sistema

$$\begin{pmatrix}\nu_1 & \nu_2 & \nu_3\end{pmatrix} = \begin{pmatrix}\nu_1 & \nu_2 & \nu_3\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\ 1/7 & 0 & 6/7\\ 0 & 1 & 0\end{pmatrix} = \begin{pmatrix}\nu_2/7 & \nu_1+\nu_3 & 6\nu_2/7\end{pmatrix}.$$

Fijando $\nu_2 = 7$, se tiene que $(\nu_1, \nu_2, \nu_3) = (1, 7, 6)$, y normalizando para que $\nu_1 + \nu_2 + \nu_3 = 1$, tendremos que $(1/14, 1/2, 3/7)$ es la distribución invariante.

Para calcular las potencias de P, haremos una diagonalización. Entonces

$$\det\begin{pmatrix}-6-\lambda & 6 & 0\\ 1 & -7-\lambda & 6\\ 0 & 2 & -2-\lambda\end{pmatrix} = -\lambda(\lambda+1)(\lambda-1),$$

es decir, sus eigenvalores serán $0, -1$ y 1.

Para $\lambda_1 = 0$, el sistema

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}0 & 1 & 0\\ 1/7 & 0 & 6/7\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}x_2\\ x_1/7+6x_3/7\\ x_2\end{pmatrix}$$

tiene como solución al vector $(6, 0, -1)^T$.

Para $\lambda_2 = -1$, el sistema

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}-1 & 1 & 0\\ 1/7 & -1 & 6/7\\ 0 & 1 & -1\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\end{pmatrix} = \begin{pmatrix}-y_1+y_2\\ y_1/7-y_2+6y_3/7\\ y_2-y_3\end{pmatrix}$$

tiene como solución al vector $(1, 1, 1)^T$.

Para $\lambda_3 = 1$, el sistema

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}1 & 1 & 0\\ 1/7 & 1 & 6/7\\ 0 & 1 & 1\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\end{pmatrix} = \begin{pmatrix}z_1+z_2\\ z_1/7+z_2+6z_3/7\\ z_2+z_3\end{pmatrix}$$

tiene como solución al vector $(1, -1, 1)^T$. Entonces $P = ADA^{-1}$, donde

$$A = \begin{pmatrix}6 & 1 & 1\\ 0 & 1 & -1\\ -1 & 1 & 1\end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix}0 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Mediante el método de Gauss (hecho en papel pero omitido en este archivo), es fácil concluir que

$$A^{-1} = \frac{1}{14}\begin{pmatrix} 2 & 0 & -2 \\ 1 & 7 & 6 \\ 1 & -7 & 6 \end{pmatrix},$$

de manera que

$$\begin{aligned} P^n = AD^nA^{-1} &= \frac{1}{14}\begin{pmatrix} 6 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1^n & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 0 & -2 \\ 1 & 7 & 6 \\ 1 & -7 & 6 \end{pmatrix} \\ &= \frac{1}{14}\begin{pmatrix} (-1)^n + 1 & (-1)^n 7 - 7 & (-1)^n 6 + 6 \\ (-1)^n - 1 & (-1)^n 7 + 7 & (-1)^n 6 - 6 \\ (-1)^n + 1 & (-1)^n 7 - 7 & (-1)^n 6 + 6 \end{pmatrix}. \end{aligned}$$

Notemos que varios de esos términos se eliminan si hacemos casos cuando n es par o impar. De cualquier manera, es sencillo notar que esta matriz diverge, hecho que era claro al notar que la matriz de transición P no es aperiódica. $\square$

2. ¿Cuál es la probabilidad de que ambos generadores estén trabajando al tiempo t si sólo uno trabaja al tiempo cero?

Demostración. Necesitamos calcular P_t de manera que se diagonalizará la matriz Q. Entonces

$$\det\begin{pmatrix} -6 & 6 & 0 \\ 1 & -7 & 6 \\ 0 & 2 & -2 \end{pmatrix} = -\lambda(\lambda+5)(\lambda+10),$$

por lo que sus eigenvalores son $0, -5$ y -10.

Para $\lambda_1 = 0$, el sistema

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -6 & 6 & 0 \\ 1 & -7 & 6 \\ 0 & 2 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -6x_1 + 6x_2 \\ x_1 - 7x_2 + 6x_3 \\ 2x_2 - 2x_3 \end{pmatrix}$$

tiene como solución al vector $(1, 1, 1)^T$.

Para $\lambda_2 = -5$, notemos que el sistema

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 & 6 & 0 \\ 1 & -2 & 6 \\ 0 & 2 & 3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} -y_1 + 6y_2 \\ y_1 - 2y_2 + 6y_3 \\ 2y_2 + 3y_3 \end{pmatrix}$$

tiene como solución (al fijar $y_2 = -3$) al vector $(-18, -32)$.

Para $\lambda_3 = -10$, el sistema

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 0 \\ 1 & 3 & 6 \\ 0 & 2 & 8 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 4z_1 + 6z_2 \\ z_1 + 3z_2 + 6z_3 \\ 2z_2 + 8z_3 \end{pmatrix}$$

tiene como solución (al fijar $z_3 = 1$) al vector $(6, -4, 1)^T$.

Entonces $Q = ADA^{-1}$, donde

$$A = \begin{pmatrix} 1 & -18 & 6 \\ 1 & -3 & -4 \\ 1 & 2 & 1 \end{pmatrix} \quad \text{y} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -10 \end{pmatrix}.$$

Usando el método de Gauss, se tendrá que

$$A^{-1} = \frac{1}{25} \begin{pmatrix} 1 & 6 & 18 \\ -1 & -1 & 2 \\ 1 & -4 & 3 \end{pmatrix}.$$

Entonces

$$\begin{aligned} P_t = e^{Qt} &= \sum_{n=0}^{\infty} \frac{Q^n t^n}{n!} = \sum_{n=0}^{\infty} \frac{(ADA^{-1})^n t^n}{n!} \\ &= A \left(\sum_{n=0}^{\infty} \frac{D^n t^n}{n!} \right) A^{-1} = Ae^{Qt} A^{-1} \\ &= \frac{1}{25} \begin{pmatrix} 1 & -18 & 6 \\ 1 & -3 & -4 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} e^{0t} & 0 & 0 \\ 0 & e^{-5t} & 0 \\ 0 & 0 & e^{-10t} \end{pmatrix} \begin{pmatrix} 1 & 6 & 18 \\ -1 & -1 & 2 \\ 1 & -4 & 3 \end{pmatrix} \end{aligned}$$

Entonces para encontrar la probabilidad de que 2 estén funcionando al tiempo t dado que iniciamos con 1, basta con obervar la entrada $(1, 2)$ de la matriz P_t, es decir

$$P_t(1, 2) = \frac{1}{25} \begin{pmatrix} 1 & -3e^{-5t} & -4e^{-10t} \end{pmatrix} \begin{pmatrix} 18 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{25}(18 - 6e^{-5t} - 12e^{-10t}).$$

□

3. Si ρ_2 denota la primera vez que ambos generadores están trabajando al mismo tiempo, encuentre la distribución de ρ_2 cuando sólo un generador está trabajando al tiempo cero.

Demostración. En este caso utilizaré el método de distribuciones tipo fase para ahorrarme cálculos (no sé si esto sea válido en el examen).

Para calcular la distribución del tiempo de entrada al estado 2, basta con modificar las entradas de la fila de ese estado para que éste sea absorbente y así para saber si al tiempo t aún no hemos entrado al estado 2, bastará con preguntarnos si al tiempo t seguimos en el estado 0 ó 1. Sea Q_0 esta modificación y P_t^0 su matriz de transición. Q_0 será de la forma

$$Q_0 = \begin{pmatrix} T & s \\ \begin{pmatrix} 0 & 0 \end{pmatrix} & 0 \end{pmatrix} \text{ donde } T = \begin{pmatrix} -6 & 6 \\ 1 & -7 \end{pmatrix} \text{ y } s = -T\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 6 \end{pmatrix}.$$

Es sencillo verificar, mediante la definición de la expansión de $e^{Q_0 t}$, que

$$P_t^0 = e^{Q_0 t} = \begin{pmatrix} e^{Tt} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} - e^{Tt}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 0 & 0 \end{pmatrix} & 1 \end{pmatrix}.$$

A pesar de parecer que la matriz original se complicó, para obtener la solución basta concentrarnos en la matriz e^{Tt}, ya que es ésta la que dicta el comportamiento del proceso cuando éste está en los estados transitorios 0 ó 1. Procederemos a diagonalizar T. Entonces

$$\det \begin{pmatrix} -6-\lambda & 6 \\ 1 & -7-\lambda \end{pmatrix} = (\lambda+9)(\lambda+4),$$

es decir, sus eigenvalores son -9 y -4.

Para $\lambda_1 = -9$,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 + 6x_2 \\ x_1 + 2x_2 \end{pmatrix},$$

se tiene como solución al eigenvector $(2,-1)^T$.

Para $\lambda_2 = -4$,

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 & 6 \\ 1 & -3 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2y_1 + 6y_2 \\ y_1 - 3y_2 \end{pmatrix},$$

se tiene como solución al eigenvector $(3,1)^T$.

Después de invertir la matriz de eigenvalores, se tendrá que

$$T = \frac{1}{5}\begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} -9 & 0 \\ 0 & -4 \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix}$$

y por el mismo argumento que en el inciso anterior,

$$\begin{aligned} e^{Tt} &= \frac{1}{5}\begin{pmatrix} 2 & 3 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} e^{-9t} & 0 \\ 0 & e^{-4t} \end{pmatrix}\begin{pmatrix} 1 & -3 \\ 1 & 2 \end{pmatrix} \\ &= \frac{1}{5}\begin{pmatrix} 2e^{-9t} + 3e^{-4t} & 6(-e^{-9t} + e^{-4t}) \\ -e^{-9t} + e^{-4t} & 3e^{-9t} + 2e^{-4t} \end{pmatrix} \end{aligned}$$

Sea τ_2 el tiempo hasta la absorción al estado 2. Entonces, considerando una distribución inicial entre los estados transitorios del proceso de la forma $\pi = (0, 1)$ (deseamos que inicie en el estado 1), entonces

$$\begin{aligned}\mathbb{P}_\pi(\tau_2 > t) &= \mathbb{P}_\pi(\text{Al tiempo } t \text{ el proceso esté en 0 ó 1}) \\ &= \begin{pmatrix} 0 & 1 \end{pmatrix} e^{Tt} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{5}(2e^{-9t} + 3e^{-4t}),\end{aligned}$$

concluyendo la demostración al notar que $F_{\tau_2}(t) = 1 - \mathbb{P}_\pi(\tau_2 > t)$. □

4. Encuentre la proporción de tiempo asintótica en que los dos generadores están trabajando. Si cada generador produce 2.5 MW de energía por unidad de tiempo, ¿Cuál es la cantidad promedio de energía producida a largo plazo por unidad de tiempo?

   Demostración. Al ser una cadena irreducible, para obtener la proporción asintótica, debemos fijarnos en el comportamiento límite de cualquier fila de la matriz P_t descrita en el inciso 2 cuando $t \to \infty$ (digamos, la fila 0). Entonces, desarrollando la multiplicación de matrices, la matriz P_t tendrá en su fila 0 a

$$\frac{1}{25}\begin{pmatrix} 1 + 18e^{-5t} + 6e^{-10t} \\ 6 + 18e^{-5t} - 24e^{-10t} \\ 18 - 36e^{-5t} + 18e^{-10t} \end{pmatrix}^T \to \frac{1}{25}\begin{pmatrix} 1 & 6 & 18 \end{pmatrix},$$

   de manera que la proporción de tiempo que trabajarán 2 generadores es de 18/25.

   Para calcular el total de energía producida, basta con hacer el cálculo

$$2.5\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}\begin{pmatrix} 1/25 & 6/25 & 18/25 \end{pmatrix} = 4.2,$$

   es decir, en promedio se producen 4.2 MW de energía por unidad de tiempo. □

Problema 10 (Procesos de ramificación a tiempo continuo)**.** Sea μ una distribución en $\mathbb{N}$. A μ_k lo interpretamos como la probabilidad de que un individuo tenga k hijos. Nos imaginamos la dinámica de la población como sigue: a tasa λ, los individuos de una población se reproducen. Entonces tienen k hijos con probabilidad μ_k. Se pueden introducir dos modelos: uno en que el individuo que se reproduce es retirado de la población (nos imaginamos que muere) y otro en que no es retirado de la población (por ejemplo cuando se interpreta a la población como especies y a sus descendientes como mutaciones). En el caso particular del segundo modelo en que $\mu_1 = 1$, se conoce como proceso de Yule.

1. Especifique un modelo de cadenas de Markov a tiempo continuo para cada uno de los modelos anteriores. A estos procesos se les conoce como procesos de ramificación a tiempo continuo.

Demostración. A continuación se describirá el proceso asociado al primer caso, cuando el individuo que se reproduce, muere. Para los descendientes y para la población total consideraremos como espacio de estados a $\mathbb{Z}_+$. Sin embargo, se pedirá que $\mu_1 = 0$, ya que si $\mu_1 > 0$, pueden existir cambios en la población que no afectan el número total de individuos; sin embargo, al morir y nacer un individuo nuevo técnicamente hubo un cambio. Este problema puede ocasionar que los tiempos entre saltos del número de población se distribuyan Erlang y no exponenciales, y por lo tanto no sería posible modelarlo con procesos de Markov constante por pedazos (o por lo menos no mediante el modelo convencional, aunque al parecer hay modelos que si permiten saltos al mismo estado, como por ejemplo, el usado para aplicarle uniformización a un proceso de Markov constante por pedazos). Bajo esta restricción, supongamos que tenemos inicialmente k individuos, y cada uno espera un tiempo exponencial, digamos de parámetro λ, para reproducirse y tener hijos según la ley de μ. Es decir, para la población general, habrá un salto hasta que el primero de los k individuos se reproduzca, lo cual vimos en clase, tiene una distribución exponencial de parámetro λk, y dado que ya se reprodujo, habrá un incremento de la población en i individuos con probabilidad μ_{i+1} o un decremento con probabilidad μ_0, de manera que el proceso descrito se repetirá con la nueva población hasta la saiguiente reproducción y así sucesivamente. Obviamente si la población llega a ser 0, el proceso se estancará ahí. Todo esto puede ser resumido en la matriz de intensidades

$$Q_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ \lambda\mu_0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ 0 & 2\lambda\mu_0 & -2\lambda & 2\lambda\mu_1 & 2\lambda\mu_2 & \cdots \\ 0 & 0 & 3\lambda\mu_0 & -3\lambda & 3\lambda\mu_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

y por clase, sabemos que existe un proceso de Markov constante por pedazos que tiene a Q_1 como matriz de intensidad.

El segundo modelo, cuando el individuo no muere, requiere de algunos cambios mínimos. Por el mismo argumento que en el modelo anterior, se restringirá a que $\mu_0 = 0$, es decir, que cuando individuo se reproduzca, este debe tener 1 o más hijos. También por el mismo argumento que en el modelo anterior, para una población inicial de k individuos, los saltos entre el cambio del número de individuos será una v.a. exponencial de parametro $k\lambda$; este cambio podrá ser solamente para incrementarlo, y será de i unidades con probabilidad μ_i. Después esta nueva población de nuevo esperará a un nuevo nacimiento y así sucesivamente. En este caso, si $k > 0$, entonces será imposible que la población alguna vez se extinga, sin embargo, añadiremos ese estado para que coincida con el espacio de estados del modelo anterior. La matriz de intensidades de este proceso será

$$Q_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ 0 & 0 & -2\lambda & 2\lambda\mu_1 & 2\lambda\mu_2 & \cdots \\ 0 & 0 & 0 & -3\lambda & 3\lambda\mu_1 & \cdots \\ 0 & 0 & 0 & 0 & -4\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

y de nuevo, por clase sabemos que existe un proceso de Markov constante por pedazos que tiene a Q_2 como matriz de intensidad. □

Nuestro primer objetivo será encontrar una relación entre procesos de ramificación a tiempo continuo y procesos de Poisson compuestos. Sea N un proceso de Poisson y S una caminata aleatoria independiente de N tal que $\mathbb{P}(S_1 = j) = \mu_{j-1}$ ó μ_j dependiendo de si estamos en el primer caso o en el segundo. Sea $k \geq 0$ y definamos a $X_t = k + S_{N_t}$.

2. Diga brevemente por qué X es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal para ambos modelos.

Demostración. Bajo las restricciones impuestas en los modelos anteriores, es claro que los tiempos entre cambios de X corresponden a los tiempos entre saltos del proceso N_t, los cuales son exponenciales de parámetro λ (notemos que en este caso, los tiempos entre saltos no son dependientes del número de individuos 'actual', es el mismo para todos, sea cual sea el número). Al ser un proceso Poisson, los saltos se producirán de una unidad en una unidad y cada salto será finito c.s.; esto quiere decir que el proceso X condicionado en los instantes de salto, es decir, el proceso $\{k + S_{N_{T_i}}\}_{i\in\mathbb{N}}$ es exactamente igual al proceso $\{k + S_i\}_{i\in\mathbb{N}}$, que como es una caminata aleatoria (trasladada), será también una cadena de Markov. En conclusión, se tiene que la sucesión de estados visitados por X es una cadena de Markov con tiempos entre saltos exponenciales de parámetro λ, y por lo tanto, X es una cadena de Markov constante por pedazos. Toda esta discusión tendrá como resultado las matrices infinitesimales

$$Q_1^0 = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \lambda\mu_4 & \cdots \\ \cdots & \lambda\mu_0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ \cdots & 0 & \lambda\mu_0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \cdots \\ \cdots & 0 & 0 & \lambda\mu_0 & -\lambda & \lambda\mu_1 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

y

$$Q_2^0 = \begin{pmatrix} -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \lambda\mu_4 & \cdots \\ 0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ 0 & 0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \cdots \\ 0 & 0 & 0 & -\lambda & \lambda\mu_1 & \cdots \\ 0 & 0 & 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

para el modelo 1 y 2, respectivamente. Para el modelo 1, el proceso X toma todo los valores de $\mathbb{Z}$, debido a que puede suceder que $\mu_0 > 0$ y como se retira un individuo de la población, técnicamente ese proceso puede tomar cualquier valor de los enteros negativos. Para el modelo 2, bajo el supuesto que $k \geq 0$, esto no es así, de manera que el proceso X toma valores sólo en $\mathbb{Z}_+$. □

Sea ahora $\tau = \min\{t \geq 0 : X_t = 0\}$ y $Y_t = X_{t\wedge\tau}$.

3. Argumente por qué Y es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal.

Demostración. Bajo el evento $\{\tau = \infty\}$ es claro que $Y_t = X_t$ y por lo tanto es una cadena de Markov. Bajo el evento $\{\tau < \infty\}$ y para todo $t_0 < \cdots < t_k < \tau \leq t_{k+1} < \cdots < t_n$, se tiene que

$$\begin{aligned} &\mathbb{P}(Y_{t_0} = i_0, \ldots, Y_{t_k} = i_k, Y_{t_{k+1}} = i_{k+1}, \ldots, Y_{t_n} = i_n) \\ &\quad = \mathbb{P}(X_{t_0} = i_0, \ldots, X_{t_k} = i_k)\mathbb{1}_{i_{k+1}=\cdots=i_n=0} \\ &\quad = P^X_{t_0}(k, i_0)P^X_{t_1-t_0}(i_0, i_1)\cdots P^X_{t_k-t_{k-1}}(i_{k-1}, i_k)\mathbb{1}_{i_{k+1}=0}\cdots\mathbb{1}_{i_n=0}, \end{aligned}$$

y por lo tanto, Y es una cadena de Markov a tiempo continuo. Su matriz de intensidad (suponiendo que $k \geq 0$) es sencilla de deducir, al notar que ahora 0 es un estado absorbente. éstas están dadas por

$$Q_1^1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ \lambda\mu_0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ 0 & \lambda\mu_0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \cdots \\ 0 & 0 & \lambda\mu_0 & -\lambda & \lambda\mu_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

y por

$$Q_2^1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \lambda\mu_3 & \cdots \\ 0 & 0 & -\lambda & \lambda\mu_1 & \lambda\mu_2 & \cdots \\ 0 & 0 & 0 & -\lambda & \lambda\mu_1 & \cdots \\ 0 & 0 & 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

para los modelos 1 y 2 respectivamente. □

4. Argumente por qué existe un único proceso Z que satisface

$$Z_t = Y_{\int_0^t Z_s\,ds}$$

y que dicho proceso es un proceso de ramificación a tiempo continuo. Sugerencia: Recuerde que las trayectorias de Y son constantes por pedazos.

Demostración. Supongamos que nos encontramos en el modelo 2. Esto quiere decir que $Y_t \geq 1$ y por lo tanto, $Z_t \geq 1$ también. Definamos la función

$$G(t) := \int_0^t Z(s).$$

Notemos que por la propiedad de $Z_t \geq 1$, G es una función estrictamente creciente además de contínua, por lo que $G^{-1}(u) = \text{ínf}\{t : G(t) = u\}$ existe y está bien definida. La continuidad nos dice que efectivamente los estados visitados por Z son los mismos que Y y estos forman una cadena de Markov (por que Y es de Markov constante a pedazos). Sean $\{T_i\}_{i\in\mathbb{N}}$ los saltos que realiza Y. El objetivo será demostrar que los tiempos entre los saltos de Z son variables aleatorias exponenciales independientes.

Observemos que $Y_0 = k$, lo cual implica que $Z_0 = k$ y dado que el primer salto de Y ocurre en $T_1 = G(T_1/k) = G(T_1/Y_0)$, entonces el proceso sucederá que $Z_r = k$ para todo $r \in [0, T_1/Y_0)$, o en otras palabras, el primer cambio de estado del proceso Z sucede al instante $\mathscr{G}^{-1}(T_1) = T_1/Y_0 =\sim Exp(\lambda Y_0)$. Ahora, sea $\mathscr{G}_1 = \sigma(Y_t : t \leq T_1)$. Por la propiedad fuerte de Markov, el proceso $\{Y_{T_1+t}\}_{t\geq 0}$ tiene la misma distribución que Y bajo $\mathbb{P}_{y_1}$ donde $y_1 = Y_{T_1}$. Así, entonces condicionado a $\mathscr{G}_i$, $T_2 - T_1$ es exponencial de parámetro λ. Ya que estamos considerando el proceso Y trasladado al primer salto, para Z el tiempo entre los 2 saltos es de $(T_2 - T_1)/Y_{T_1}$, lo cual es sencillo de verificar al evaluar esa cantidad en G y ver que es igual a $T_2 - T_1$. Además, $(T_2 - T_1)/Y_{T_1} \sim Exp(\lambda Y_{T_1})$ que por la markovaniedad de Y, es independiente del primer salto. Si definimos $\mathscr{G}_2$ y continuamos con el mismo procedimiento, veremos que esta construcción de Z da como resultado un proceso que salta según una variable aleatoria exponencial de intensidad proporcional al número de individuos que se encuentran presentes, y ya que salta, se incrementará según la distribución de μ; es decir, se describió exactamente el mismo modelo propuesto en el inciso 1.

Suponiendo que nos encontramos en el primer modelo, basta ver que si definimos $\tau_0 = \text{ínf}\{t : Y_t = 0\}$ y sucede que $\{\tau_0 < \infty\}$, entonces por definición, $Z_t = 0$ para todo $t \geq G^{-1}(\tau_0)$, es decir, el proceso Z se absorberá adecuadamente en 0 una vez que llega a ese estado. □

Ahora nos enfocaremos en el proceso de Yule.

5. Escriba las ecuaciones backward de Kolmogorov para las probabilidades de transición $P_t(x, y)$. Al argumentar por qué $P_t(x, x) = e^{-\lambda x}$, resuelva las ecuaciones backward por medio de la técnica de factor integrante (comenzando con $P_t(x, x+1)$) y pruebe que

$$P_t(x, y) = \binom{y-1}{y-x} e^{-\lambda x t} \left(1 - e^{-\lambda t}\right)^{y-x}. \tag{5}$$

Demostración. El proceso de Yule tiene una matriz infinitesimal Q de la forma

$$Q(x, y) = \begin{cases} -\lambda x & y = x \\ \lambda x & y = x + 1 \\ 0 & \text{e.o.c.} \end{cases}$$

También recordemos que este proceso es creciente y por lo tanto, sus probabilidades serán no nulas cuando $x \leq y$, y nos referiremos a este caso a partir de ahora. Por lo tanto, la ecuación backward de Kolmogorov será igual a

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t(x, y) = (QP_t)(x, y) = -\lambda x P_t(x, y) + \lambda x P_t(x + 1, y).$$

Deseamos calcular la solución a esta ecuación diferencial para todo $y \geq x$, así que se realizará por inducción. Para la base, observemos que

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t(x, x) = -\lambda x P_t(x, x) + \lambda x P_t(x + 1, x) = -\lambda x P_t(x, x).$$

La condición inicial de esta ecuación diferencial es $P_0(x, x) = 1$, y de ahí se concluye que $P_t(x, x) = e^{-\lambda x t}$, lo cual coincide con (5). Lo anterior es válido para todo x. Ahora, supongamos que la fórmula es válida para $y = x + k \geq x$ para cualquier x, es decir que la fórmula es válida cuando la diferencia entre los valores x y y es k. Necesitamos demostrar que también es válida para $x + k + 1$ para cualquier x. Entonces,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} P_t(x, x + k + 1) &= -\lambda x P_t(x, x + k + 1) + \lambda x P_t(x + 1, x + k + 1) \\
&= -\lambda x P_t(x, x + k + 1) + \lambda x P_t((x + 1), (x + 1) + k) \\
&= -\lambda x P_t(x, x + k + 1) \\
&\quad + \lambda x \binom{(x + 1) + k - 1}{(x + 1) + k - (x + 1)} e^{-\lambda (x+1) t} \left(1 - e^{-\lambda t}\right)^{(x+1)+k-(x+1)} \\
&\qquad \text{(Por hipótesis de inducción)} \\
&= -\lambda x P_t(x, x + k + 1) + \lambda x \binom{x + k}{k} e^{-\lambda (x+1) t} \left(1 - e^{-\lambda t}\right)^{k}.
\end{aligned}$$

Usando a $e^{\lambda x t}$ como factor integrante, obtenemos que

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{\lambda x t}P_t(x,x+k+1)\right) = \lambda x\binom{x+k}{k}e^{-\lambda x t}\left(1-e^{-\lambda t}\right)^k.$$

Integrando de 0 a t y notando que $P_0(x,x+k+1)=0$,

$$\begin{aligned}
e^{\lambda x t}P_t(x,x+k+1) &= \int_0^t \lambda x\binom{x+k}{k}e^{-\lambda x s}\left(1-e^{-\lambda s}\right)^k \mathrm{d}s\\
&= \left[x\binom{x+k}{k}\frac{\left(1-e^{-\lambda s}\right)^{k+1}}{k+1}\right]_0^t\\
&\qquad (\text{Haciendo } u = 1-e^{-\lambda s},\ \mathrm{d}u = \lambda e^{-\lambda s}\mathrm{d}s)\\
&= \left[x\binom{x+k}{k}\frac{\left(1-e^{-\lambda s}\right)^{k+1}}{k+1}\right]_0^t\\
&= \binom{x+k}{k+1}\left(1-e^{-\lambda t}\right)^{k+1},
\end{aligned}$$

de manera que se concluye que

$$P_t(x,x+k+1) = \binom{x+k}{k+1}e^{\lambda x t}\left(1-e^{-\lambda t}\right)^{k+1},$$

que es justamente la fórmula que queríamos demostrar (recordemos que estábamos demostrando la fórmula (5) para el caso $y=x+k+1$). □

6. Al utilizar la fórmula para la esperanza de una variable binomial negativa, pruebe que

$$\mathbb{E}_x(Z_t) = xe^{\lambda t}.$$

Demostración. Del inciso anterior es fácil notar que entonces $Z_t \sim BinNeg(x, e^{-\lambda t})$ (en este caso, $e^{-\lambda t}$ es la probabilidad de éxito y x el número de éxitos requeridos). Sabemos que la esperanza de una $BinNeg(r,p)$ es igual a r/p, por lo tanto,

$$\mathbb{E}_x(Z_t) = xe^{\lambda t}.$$

□

7. Pruebe que $e^{-\lambda t}Z_t$ es una martingala no-negativa y que por lo tanto converge casi seguramente a una variable aleatoria W.

Demostración. Asumamos que $\mathscr{F}_t = \sigma(Z_t)$. Entonces $e^{-\lambda t}Z_t$ es $\mathscr{F}_t$-adaptada, y además por el inciso anterior

$$\mathbb{E}_x(|e^{-\lambda t}Z_t|) = \mathbb{E}_x(e^{-\lambda t}Z_t) = x.$$

Por lo tanto, sólo falta checar la propiedad de martingala. Sea $s < t$. Haciendo uso de la propiedad de Markov de Z,

$$\mathbb{E}_x(e^{-\lambda t} Z_t | \mathscr{F}_s) = e^{-\lambda t} \mathbb{E}_{Z_s}(Z_{t-s}) = e^{-\lambda t} Z_s e^{\lambda(t-s)} = e^{-\lambda s} Z_s.$$

Como es una martingala no negativa, por el teorema de convergencia de martingalas, ésta debe converger c.s. a una v.a. $W \in L_1$. $\square$

8. Al calcular la transformada de Laplace de $e^{-\lambda t} Z_t$, pruebe que W tiene distribución exponencial. Por lo tanto, argumente que casi seguramente Z crece exponencialmente.

   Demostración. La transformada de Laplace en función de θ de una variable aleatoria $BinNeg(r, p)$ esta dada por

$$\left(\frac{pe^{\theta}}{1 - (1-p)e^{\theta}} \right)^r,$$

   la cual es sencilla de recordar si se recuerda que $BinNeg(r, p)$ es la convolución de r v.a. $Geo(p)$. Sin embargo, nuestro objetivo es calcular $\mathbb{E}(\exp(-\theta e^{-\lambda t} Z_t))$ donde $Z_t \sim BinNeg(x, e^{-\lambda t})$. Haciendo los ajustes necesarios, tendremos que

$$\mathbb{E}(\exp(-\theta e^{-\lambda t} Z_t)) = \left(\frac{e^{-\lambda t} e^{\theta e^{-\lambda t}}}{1 - (1 - e^{-\lambda t}) e^{\theta e^{-\lambda t}}} \right)^x = \left(\frac{e^{-\lambda t}}{e^{-\theta e^{-\lambda t}} - (1 - e^{-\lambda t})} \right)^x.$$

   Calculando el límite,

$$\begin{aligned}
\lim_{t \to \infty} \left(\frac{e^{-\lambda t}}{e^{-\theta e^{-\lambda t}} - (1 - e^{-\lambda t})} \right)^x &= \lim_{s \to 0} \left(\frac{s}{e^{-\theta s} - (1 - s)} \right)^x \\
&= \lim_{s \to 0} \left(\frac{1}{-\theta e^{-\theta s} + 1} \right)^x \\
&\qquad \text{(Por regla de L'Hopital y por continuidad de la función } f(a) = a^x) \\
&= \left(\frac{1}{1 - \theta} \right)^x.
\end{aligned}$$

   En el caso $x = 1$, esto es la transformada de Laplace de una distribución $Exp(1)$. Como ya demostramos que en el limite $e^{-\lambda t} Z_t$ tiende a una v.a. W c.s., esta debe ser entonces $Exp(1)$, finalizando la demostración. $\square$

Problema 11. (Tomado del examen general de conocimientos del área de Probabilidad del Posgrado en Ciencias Matemáticas, UNAM, Agosto 2011)

Sea N un proceso de Poisson homogéneo de parámetro λ. Sea $E = (-1, 1)$ y X_0 una variable aleatoria con valores en E independiente de N. Se define el proceso

$$X_t = X_0 \times (-1)^{N_t}, \quad t \geq 0.$$

1. Explique por qué X es una cadena de Markov a tiempo continuo con valores en E.

   Demostración. Es claro que los tiempos de salto de estado del proceso X es cuando el proceso N salta, digamos la sucesión $\{T_i\}_{i\in\mathbb{N}}$. Los tiempos entre estos saltos, que podemos llamar $\{S_i\}_{i\in\mathbb{N}}$ son v.a.i.i.d. $Exp(\lambda)$, debido a que N es un proceso Poisson. Por lo tanto, para demostrar que X es una cadena de Markov a tiempo continuo basta con verificar que las v.a. $Y_n = X_{T_n}$ formen una cadena de Markov.

   Para la v.a. X_0, consideremos una distribucion μ sobre el espacio de estados E. Primero veamos que como $N_0 = 0$ c.s., entonces X_0 (como un valor del proceso, no como v.a.) si está bien definido, y ademas se tendrá que $Y_0 = X_{T_0} = X_0$. Al ser N un proceso Poisson sus incrementos serán de una unidad en una unidad, y entonces

$$Y_1 = X_{T_1} = X_0 \times (-1)^{N_{T_1}} = X_0 \times (-1)^1 = -X_0,$$
$$Y_2 = X_{T_2} = X_0 \times (-1)^{N_{T_2}} = X_0 \times (-1)^2 = X_0,$$

   y así sucesivamente, de manera que el proceso Y es alternante entre X_0 y $-X_0$. Sea $i_0, \ldots, i_n \in E$. Entonces

$$\begin{aligned}
&\mathbb{P}_\mu(Y_0 = i_0, Y_1 = i_1, \ldots, Y_n = i_n)\\
&\quad= \mathbb{P}_\mu(X_0 = i_0)\mathbb{1}\{\text{La sucesión } i_0, \ldots, i_n \text{ es alternante}\}\\
&\quad= \mu(i_0)\mathbb{1}_{i_0=-i_1}\mathbb{1}_{i_1=-i_2}\cdots\mathbb{1}_{i_{n-1}=-i_n},
\end{aligned}$$

   de manera que se acaba de exhibir que Y es una cadena de Markov con distribución inicial μ. Por clase, entonces sabemos que existe una cadena de Markov continua, que será justamente X, que tiene a Y como cadena de Markov asociada y a tiempos $Exp(\lambda)$ independientes como longitudes entre saltos. ☐

2. Calcule sus probabilidades de transición y su matriz infinitesimal. La matriz de transición de la cadena asociada Y, al ser alternante entre -1 y 1, es de la forma

$$P^Y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

   La matriz de intensidad del proceso X entonces será

$$Q^X = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix}.$$

   Para calcular las probabilidades de transición del proceso X, conviene diagonalizar la matriz Q^X, de manera que

$$\det\begin{pmatrix} -\lambda-\delta & \lambda \\ \lambda & -\lambda-\delta \end{pmatrix} = (\lambda+\delta)^2 - \lambda^2 = \delta(2\lambda+\delta),$$

por lo que sus eigenvalores son 0 y -2λ.

Para $\delta_1 = 0$, el sistema

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -\lambda & \lambda \\ \lambda & -\lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\lambda x_1 + \lambda x_2 \\ \lambda x_1 - \lambda x_2 \end{pmatrix},$$

tiene como solución a $(1,1)^T$.

Para $\delta_1 = -2\lambda$, el sistema

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda & \lambda \\ \lambda & \lambda \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \lambda y_1 + \lambda y_2 \\ \lambda y_1 y \lambda y_2 \end{pmatrix},$$

tiene como solución a $(1,-1)^T$.

Por lo tanto,

$$Q^X = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & -2\lambda \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

y en conclusión,

$$P_t^X = e^{Q^X t} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{-2\lambda t} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + e^{-2\lambda t} & 1 - e^{-2\lambda t} \\ 1 - e^{-2\lambda t} & 1 + e^{-2\lambda t} \end{pmatrix}.$$

3. ¿Existe una distribución estacionaria para esta cadena? En caso afirmativo ?'Cuál es?

   Demostración. Claramente X es irreducible, y como el número de estados en E es finito, si existe distribución estacionaria. Para esto, primero encontremos la de Y. Es decir, se desea encontrar $\nu = (\nu_{-1}, \nu_1)$ tal que $\nu P^Y = \nu$, entonces

$$\begin{pmatrix} \nu_{-1} & \nu_1 \end{pmatrix} = \begin{pmatrix} \nu_{-1} & \nu_1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \nu_1 & \nu_{-1} \end{pmatrix},$$

   es decir, $\nu_{-1} = \nu_1 = 1/2$.

   Entonces por un teorema de clase, una medida invariante de X está dada por $\nu^X = C \cdot \nu^Y$, donde $C = (\lambda, \lambda)$, el vector de intensidades de salto. Así pues, se concluye que $\nu^X = (\lambda/2, \lambda/2)$ es una medida invariante. Normalizandola, la distribución invariante que buscamos es exactamente $(1/2, 1/2)$. □

Problema 12. Sea

$$Q = \begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix}.$$

1. Haga un programa en octave que permita simular las trayectorias de una cadena de Markov a tiempo continuo X con matriz infinitesimal Q.

   Demostración. A continuación se presenta el código con una trayectoria en la Figura 4.

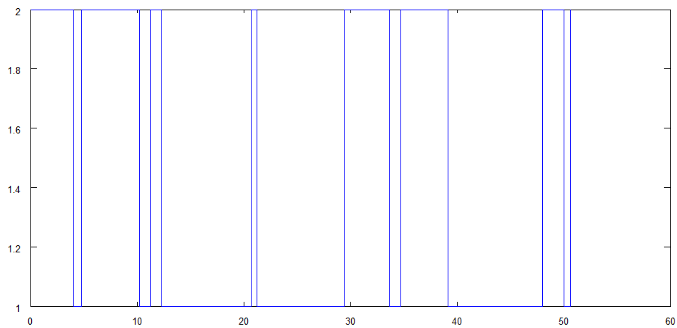

FIGURA 4. Simulación Markov

```
t=10;
T=[];
N=[];
s=0;
l(1)=2;
l(2)=3;
edo=1;
ifunifrnd(0,1)<0.5
edo=edo+1;
endif
i=1;
N(i)=edo;
T(i)=exprnd(l(edo));
while s<t
N(i+1)=abs(N(i)-2) +1;
T(i+1)=T(i) + exprnd(l(N(i+1)));
s=T(i+1);
i=i+1;
endwhile
[Ts, Ns] = stairs (T, N );
plot(Ts,Ns)
```

.

□

2. Utilice su programa para generar 10000 trayectorias en el intervalo de tiempo $[0, 10]$ comenzando con probabilidad $1/2$ en cada estado y obtenga la distribución empírica de $X_1 0$.

   Demostración. El código utilizado será el siguiente,

```
l(1)=2;
l(2)=3;
t=10;
edos=zeros(1,2);
for j=1:1000
T=[];
N=[];
s=0;
edo=1;
ifunifrnd(0,1)<0.5
edo=edo+1;
endif
i=1;
N(i)=edo;
T(i)=exprnd(l(edo));
while s<t
N(i+1)=abs(N(i)-2) +1;
T(i+1)=T(i) + exprnd(l(N(i+1)));
s=T(i+1);
i=i+1;
endwhile
edos(N(i-1))=edos(N(i-1))+1;
endfor
```

   .

   El número total de veces que al tiempo t está en el estado k es la variable $edos(k)$ para $k = 1, 2$. En la trayectoria realizada, se obtuvo que $edos(1) = 5933$ y $edos(2) = 4067$. □

3. Calcule e^{10Q} (utilizando algún comando adecuado) y contraste con la distribución empírica del inciso anterior.

   Demostración.

```
Q=[-2,2;3,-3];
pi=[0.5, 0.5];
t=10;
```

```
Q=Q*t;
Q=expm(Q);
prob=pi*Q;
```

.

, dando como resultado que

$$e^{tQ} = \begin{pmatrix} 0.6 & 0.4 \\ 0.6 & 0.4 \end{pmatrix}.$$

□

4. Codifique el siguiente esquema numérico, conocido como método de Euler, para aproximar a e^{10Q}: escoja $h > 0$ pequeño, defina a P_0^h como la matriz identidad y recursivamente
$$P_{i+1}^h = P_i^h + hQP_i^h.$$
corra hasta que $i = \lfloor 10/h \rfloor$ y compare la matriz resultante con e^{10Q}. Si no se parecen escoja a h más pequeño. ¿Con qué h puede aproximar a e^{10Q} a 6 decimales?

Demostración. El código será

```
Q=[-2,2;3,-3];
P=eye(2);
h=0.33;
i=1;
while i<= floor(10/h)
P=P+h*Q*P;
i=i+1;
endwhile
```

.

Este algoritmo se acerca al valor real cuando $h = 0.3$ (de hecho, Octave señala que es el mismo). Sin embargo, se probaron valores como 0.4 que daba valores que ni siquiera correspondían a una probabilidad. El valor que se detecto era bueno sin un cambio tan brusco era justamente 0.33. □