PROBLEMAS DE PROCESOS ESTOCÁSTICOS
SEMESTRE 2013-II
POSGRADO EN CIENCIAS MATEMÁTICAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
HÉCTOR MANUEL TÉLLEZ GÓMEZ

Part 1. Tarea 1

1.1. Problema 1.1

Problema 1.1.1. Sea $(X_{n n \in \mathbb{N}})$ un proceso estocástico con valores reales y $A \subset \mathbb{R}$ un boreliano. Pruebe que si

$$T_0 = 0 \quad \text{y} \quad T_{n+1} = \min\{k > T_n : X_k \in A \subset \mathbb{R}\}$$

entonces T_n es un tiempo de paro para toda n y $T_n \to \infty$ puntualmente conforme $n \to \infty$.

Categorías:. Tiempos de paro.

Demostración:

Sea $(\mathscr{F}_n = \sigma(X_0, X_1, \ldots, X_n))_{n \in \mathbb{N}}$ una filtración.

No es una filtración, se trata de la filtración generada por X.
Veamos que T_0 es tiempo de paro. $\{T_0 = 0\} = \Omega \in \mathscr{F}_n$ y $\{T_0 = c\} = \emptyset \in \mathscr{F}_n$ para $c \neq 0 \in \mathbb{N}$. Recordemos que:

$$\{T_0 = n\} \in \mathscr{F}_n \ \forall n \in \mathbb{N} \iff \{T_0 \leq n\} \in \mathscr{F}_n \ \forall n \in \mathbb{N}. \tag{1.1.0.1}$$

Y por lo tanto T_0 es variable aleatoria sobre la sigma álgebra $\sigma(\cup_{n \in \mathbb{N}} \mathscr{F}_n)$.

Para $n = 1$, veamos que:

$$\{T_1 = 1\} = \{X_1 \in A\} \in \mathscr{F}_1$$

$$\{T_1 = 2\} = \{X_1 \notin A, X_2 \in A\}$$

$$= \{X_1 \notin A\} \cap \{X_2 \in A\} \in \mathscr{F}_2$$

$$\begin{aligned}\{T_1 = 3\} &= \{X_1 \notin A, X_2 \notin A, X_3 \in A\} \\ &= \{X_1 \notin A\} \cap \{X_2 \notin A\} \cap \{X_3 \in A\} \in \mathscr{F}_3\end{aligned}$$

$$\vdots$$

$$\begin{aligned}\{T_1 = n\} &= \{X_1 \notin A, X_2 \notin A, \ldots, X_{n-1} \notin A, X_n \in A\} \\ &= \bigcap_{i=1}^{i=n-1} \{X_i \notin A\} \cap \{X_n \in A\} \in \mathscr{F}_n\end{aligned}$$

Por hipótesis de inducción supongamos que $\{T_k = n\} \in \mathscr{F}_n$ para cierta $k > 1$ y para toda $n \in \mathbb{N}$.

Para demostrar la primera parte del ejercicio, basta probar que $\{T_{k+1} = n\} \in \mathscr{F}_n$ para toda $n \in \mathbb{N}$. Pues por lo dicho en (1.1.0.1) resultarían ser variables aleatorias en $\sigma(\cup_{n\in\mathbb{N}}\mathscr{F}_n)$.Esto es consecuencia de que sean tiempos de paro. No hay necesidad de repetirlo,

Dado que

$$0 = T_0 < T_1 < T_2 < \cdots < T_k \tag{1.1.0.2}$$

tenemos que $\{T_k = n\} = \emptyset \in \mathscr{F}_l$ para $n < k$.

Ahora, sea $n \geq k$

$$\{T_{k+1} = n\} = \left(\bigcup_{i<n} \{T_k = i\}\right) \cap \left(\bigcap_{i<n} \{X_i \notin A\}\right) \cap \{X_n \in A\} \tag{1.1.0.3}$$

No: si $T_k = i$, debes imponer que entre $i+1$ y n no se visite A y que en n se visite A.

Donde

- $\{T_k = i\} \in \mathscr{F}_n \ \forall i < n \in \mathbb{N}$ (Por hipótesis de inducción).
- $\{X_i \notin A\} \in \mathscr{F}_n \ \forall i < n \in \mathbb{N}$ (Por ser $\mathscr{F}$ filtración).
- $\{X_n \in A\} \in \mathscr{F}_n$ (Por ser X_n variable aleatoria en $\mathscr{F}_n$)

Y por lo tanto, (1.1.0.3) pertenece a $\mathscr{F}_n$. Con lo que termina la demostración de la primera parte del ejercicio.

Ahora, de (1.1.0.2) se sigue inmediatamente que que $T_n \to \infty$.

1.2. Problema 1.2

Problema 1.2.1. Suponga que T es un tiempo de paro tal que para algún $N \in \mathbb{N}$ y $\varepsilon > 0$ se tiene que para toda $n \in \mathbb{N}$:

$$\mathbb{P}(T \leq N + n | F_n) > \varepsilon \text{ casi seguramente} \tag{1.2.0.4}$$

Al verificar la desomposición

$$\mathbb{P}(T > kN) = \mathbb{P}(T > kN, T > (k-1)N) \tag{1.2.0.5}$$

pruebe por inducción que para cada $k = 1, 2, \ldots$:

$$\mathbb{P}(T > kN) \leq (1-\varepsilon)^k.$$

Pruebe que $\mathbb{E}(T) < \infty$.

Categorías:. Tiempos de paro.

Demostración: Tenemos que:

$$\begin{aligned}
(T > kN \Rightarrow T > (k-1)N \Rightarrow (T > kN) \subset (T > (k-1)N) \\
\Rightarrow (T > kN) \cap (T > (k-1)N) = (T > kN) \\
\Rightarrow \mathbb{P}(T > kN, T > (k-1)N) = \mathbb{P}(T > kN)
\end{aligned}$$

Base de inducción. $k = 1$. Usando (1.2.0.4), con $n = 0$

$$\mathbb{P}(T \leq 1N | \mathscr{F}_0) > \epsilon \Rightarrow \mathbb{P}(T > N | \mathscr{F}_0) < 1 - \epsilon$$

Las dos lineas siguientes no parecen ser necesarias. Sustituyendo por la definición de probabilidad condicional tenemos:

$$\mathbb{E}(\mathbb{1}_{T>N} | \mathscr{F}_0) = \mathbb{P}(T > N | \mathscr{F}_0) \qquad < 1 - \epsilon$$

Aplicando esperanza en ambos lados tenemos:

$$\mathbb{P}(T > N) = \mathbb{E}(\mathbb{E}(\mathbb{1}_{T>N} | \mathscr{F}_0)) \qquad < \mathbb{E}(1-\epsilon) = 1 - \epsilon.$$

Hipótesis de induccion. Supongamos que $\mathbb{P}(T > k_0 N) \leq (1-\epsilon)^{k_0}$ para algún $k_0 \geq 1$.

Paso inductivo. Utilizando (1.2.0.5) tenemos que

$$\mathbb{P}(T > (k_0 + 1)N) = \mathbb{P}(T > (k_0 + 1)N, T > k_0N) = \mathbb{E}(\mathbb{1}_{T>(k_0+1)N} \cdot \mathbb{1}_{T>k_0N}).$$

Ahora, dado que $T > k_0N$ es un conjunto $\mathscr{F}_{k_0N}$-medible tenemos:

$$\begin{aligned} \mathbb{E}(\mathbb{1}_{T>(k_0+1)N} \cdot \mathbb{1}_{T>k_0N}) &= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{T>(k_0+1)N} \cdot \mathbb{1}_{T>k_0N} | \mathscr{F}_{k_0N}\right)\right) \\ &= \mathbb{E}\left(\mathbb{1}_{T>k_0N}\mathbb{E}\left(\mathbb{1}_{T>(k_0+1)N} | \mathscr{F}_{k_0N}\right)\right) \end{aligned} \tag{1.2.0.6}$$

Utilizando $n = k_0N$ en (1.2.0.4) tenemos

$$\mathbb{P}\left(T > k_0N + N | \mathscr{F}_{k_0N}\right) = \mathbb{E}\left(\mathbb{1}_{T>(k_0+1)N} | \mathscr{F}_{k_0N}\right) < 1 - \epsilon$$

Sustituyendo esto último en (1.2.0.6) obtenemos:

$$\begin{aligned} E\left(\mathbb{1}_{T>k_0N}\mathbb{E}\left(\mathbb{1}_{T>(k_0+1)N} | \mathscr{F}_{k_0N}\right)\right) &< E\left(\mathbb{1}_{T>k_0N}(1 - \epsilon)\right) \\ &= (1 - \epsilon)E\left(\mathbb{1}_{T>k_0N}\right) \\ &= (1 - \epsilon)\mathbb{P}(T > k_0N) \\ &\leq (1 - \epsilon)(1 - \epsilon)^{k_0} \\ &= (1 - \epsilon)^{k_0+1}. \end{aligned}$$

Con lo que concluimos

$$\mathbb{P}(T > (k_0 + 1)N) \leq (1 - \epsilon)^{k_0+1}.$$

Terminando así la demostración por inducción.

Para la siguiente prueba notemos que si X es una variable aleatoria y $r > s \in \mathbb{R}$ entonces $(X > s) \subset (X > r)$ y por lo tanto $\mathbb{P}(X > s) \leq \mathbb{P}(X > r)$.

En particular para nuestro tiempo de paro T, si dado $n \in \mathbb{N}$, $k \in \mathbb{N}$ es tal que $kN \geq n < (k + 1)N$, entonces $\mathbb{P}(T > n) \leq \mathbb{P}(T > kN)$.

Trabajando por bloques de tamaño N tenemos que

$$\sum_{n=kN}^{(k+1)N-1} \mathbb{P}(T > n) \leq N\mathbb{P}(T > kN).$$

También recordemos que si T es una variable aleatoria positiva con valores en los enteros entonces

$$\mathbb{E}(T) = \sum_{n=0}^{\infty} \mathbb{P}(T > n).$$

Sustituyendo nuestro penúltimo razonamiento en esta ultima fórmula tenemos:

$$\begin{aligned}
\mathbb{E}(T) &= \sum_{n=0}^{\infty} \mathbb{P}(T > n) \\
&= \sum_{k=0}^{\infty} \sum_{n=kN}^{(k+1)N-1} \mathbb{P}(T > n) \\
&\leq N \sum_{k=0}^{\infty} \mathbb{P}(T > kN)
\end{aligned}$$

Sustituyendo nuetstro resultado de que $\mathbb{P}(T > kN) < (1-\epsilon)^k$ resulta que:

$$\mathbb{E}(T) \leq N \sum_{k=0}^{\infty} \mathbb{P}(T > kN) \leq N \sum_{k=0}^{\infty} (1-\epsilon)^k.$$

De lado derecho de esta última desigualdad tenemos una serie geométrica. Dado que ϵ es mayor que 0 y entonces que $1-\epsilon$ es menor que 1, tenemos que es una serie geométrica que converge en los reales y por lo tanto $\mathbb{E}(T)$ está acotada por un número real.

1.3. Problema 1.3

Problema 1.3.1. *Tomado de Mathematical Tripos, Part III, Paper 33, 2012,* http://www.maths.cam.ac.uk/postgrad/mathiii/pastpapers/ Sean $(X_i, i \in \mathbb{N})$ variables aleatorias independientes con $\mathbb{P}(X_i = \pm 1) = 1/2$. Sean $S_0 = 0$ y $S_n = \sum_{i=1}^{n} X_i$.

(i) [1.3.1] Sea $T_1 = \min\{n \geq 0 : S_n = 1\}$. Explique por qué T_1 es un tiempo de paro y calcule su esperanza.
(ii) [1.3.2] Mediante el inciso anterior, construya una martingala que converge casi seguramente pero no lo hace en L_1.
(iii) [1.3.3] Sea M_n la martingala obtenida al detener a $-S$ en T_1. Utilice la solución al Problema de la Ruina para probar que $\mathbb{P}(max_n M_n \geq M) = 1/(M+1)$ para todo $M \geq 1$. Concluya que $\mathbb{E}(max_m M_n) = \infty$ y que por lo tanto $\mathbb{E}(max_{m \leq n} M_m) \to \infty$ conforme $n \to \infty$. Finalmente, deduzca que no puede haber una desigualdad de tipo Doob cuando $p = 1$.
(iv) [1.3.4] Sea $T = \min\{n \geq 2 : S_n = S_{n-2} + 2\}$ y $U = T - 2$. ¿Son T y U tiempos de paro? Justifique su respuesta.
(v) [1.3.1] Para la variable T que hemos definido, calcule $\mathbb{E}(T)$.

Categorías: Tiempos de paro, problema de la ruina

Demostración:

1.3.1. Inciso (i).

Sea $T_1 = \min\{n \geq 0 : S_n = 1\}$. Explique por qué T_1 es un tiempo de paro y calcule su esperanza.

Consideremos a la filtración $(\mathscr{F}_n)_{n \in \mathbb{N}}$ como la filtración generada por $X_1, X_2, \ldots$.

Es decir, $F_0 = \{\emptyset, \Omega\}$, $\mathscr{F}_n = \sigma(X_1, X_2, \ldots, X_n)$

Nótese que S_0 es medible bajo cualquier sigma álgebra por ser constante, en particular bajo $\mathscr{F}_0$.

La idea de un problema anterior funciona, no hay necesidad de volver a demostrarlo. Basta demostrar que $(T_1 = n) \in \mathscr{F}_n$ para ver que T_1 es tiempo de paro. T_1 representa el primer tiempo en que la suma es igual a 1. Es decir, para cualquier momento anterior, la suma no es 1.

Eso escrito en símbolos significa:

$$(T_1 = n) = \bigcap_{i=0}^{n-1} (S_i \neq 1) \cup (S_n = 1).$$

Para $n = 0$, $(S_0 = 1) = \omega \in \mathscr{F}_0$.

Como $(S_i \neq 1) \in \mathscr{F}_j$ siempre que $i \leq j$. Para $n > 0$, $(T_1 = n)$ es el resultado de unir e intersectar conjuntos $\mathscr{F}_n$-medibles, lo cual resulta $\mathscr{F}_n$-medible.

Para $m \in \mathbb{N}$. Definamos $T_m = min\{n \geq 0 : S_n = m\}$ (Nótese que para el caso $m = 1$, esta definición coincide con la definición previa de T_1).

Para $a, b \in \mathbb{N}$, podemos definir el tiempo de paro $T_{a,b} = T_{-a} \wedge T_b$, y corresponde al tiempo de paro del problema de la ruina. Para este tiempo de paro ya conocemos la esperanza y es

$$\mathbb{E}(T_{a,b}) = ab.$$

Ahora, tomemos la sucesion de variables aleatorias

$$T_{1,1}, T_{2,1}, T_{3,1}, \ldots, T_{n,1}, \ldots$$

Notemos que si $a > a' \in N$ entonces $T_{-a} > T_{-a'}$, pues T_{-a} es la primera vez que se llega a $-a$, y para poder alcanzar $-a$ era necesario haber pasado por $-a'$. De aqui tenemos que si $a > a'$, entonces $T_{a,1} \geq T_{a',1}$. De donde nuestra sucesión es no decreciente.

Por otro lado, que si $a > a' \in N$ entonces $T_{-a} > T_{-a'}$ implica que $T_{-n} \geq n \in \mathbb{N}$ es una sucesión extrictamente creciente y por lo tanto $\lim_{n\to\infty} T_{-n} = \infty$, con esto tenemos que el límite de nuestra sucesión es

$$\begin{aligned}
\lim_{n\to\infty} T_{n,1} &= \lim_{n\to\infty} T_{-n} \wedge T_1 \\
&= \infty \wedge T_1 \\
&= T_1
\end{aligned}$$

Tenemos todos los ingredientes para usar Teorema de convergencia monótona sobre nuestra sucesión y la variable T_1. Nuestra sucesion es monótona y converge puntualmente a T_1. Utilizando dicho teorema obtenemos:

$$
\begin{aligned}
\mathbb{E}(T_1) &= \mathbb{E}(\lim_{n\to\infty} T_{n,1}) \\
&= \lim_{n\to\infty} \mathbb{E}(T_{n,1}) \\
&= \lim_{n\to\infty} n \cdot 1 \\
&= \infty.
\end{aligned}
$$

Lo cual era intuitivo. Si $\mathbb{E}(T_1)$ fuese finito, diría que existe un número de volados donde uno puede apostar con mucha certeza que ganará un peso después de jugar "cerca" de esa cantidad de volados. Intuitivamente, esto vuelve injusto un juego de volados donde la moneda es justa.

1.3.2. **Inciso (ii).** *Mediante el inciso anterior, construya una martingala que converge casi seguramente pero no lo hace en L_1.*

En [1.4] se probará que si T y S son tiempos de paro, entonces $T \wedge S$ también es tiempo de paro. Con esto tenemos que si T_1 es tiempo de paro, entonces $T_1 \wedge n$ con $n \in \mathbb{N}$ también es tiempo de paro. Definimos entonces

$$M_n = S_{T_1 \wedge n}.$$

Veamos que los M_n forman una martingala.

- M_n es adaptada a la filtración.

$$M_n(w) = S_{T_1 \wedge n}(w) = S_{T_1 \wedge n(w)}(w) = \sum_{k=1}^{T_1 \wedge n(w)} X_k = \sum_{k=1}^{n} (X_k \cdot \mathbb{1}_{T_1 \geq k})(w).$$

De donde, podemos escribir:

$$M_n = \sum_{k=1}^{n} (X_k \cdot \mathbb{1}_{T_1 \geq k}). \tag{1.3.2.1}$$

Recordemos que X_k es $\mathscr{F}_n$-medible para toda $k \leq n$. Por ser T_1 tiempo de paro, los conjuntos $A_k = \{T_1 = k\}$ y $B_k = \{T_1 \leq k\}$ son F_k medibles y por lo tanto $A_k \cup B_k^c = \{T_1 \geq k\}$ también lo es. De aquí que $\mathbb{1}_{T_1 \geq k}$ es $\mathscr{F}_k$-medible y por lo tanto también $\mathscr{F}_n$-medible para toda n tal que $n \geq k$.

Entonces M_n es suma y productos de funciones $\mathscr{F}_n$-medibles y por lo tanto F_n-medible. Que es lo que queríamos demostrar.

- $M_n \in L_1$

De (1.3.2.1) podemos ver que M_n es suma finita de variables acotadas. Por lo tanto $M_n \in L_1$.

- Ahora probaremos que $\mathbb{E}(M_{n+1} | \mathscr{F}_n) = M_n$

Primero:

$$\begin{aligned}\mathbb{E}(M_{n+1} | \mathscr{F}_n) &= \mathbb{E}(S_{T_1 \wedge (n+1)} | \mathscr{F}_n) \\ &= \mathbb{E}\left(\sum_{k=1}^{n+1} (X_k \cdot \mathbb{1}_{T_1 \geq k}) \middle| \mathscr{F}_n \right)\end{aligned}$$

$$= \mathbb{E}\left(\sum_{k=1}^{n} (X_k \cdot \mathbb{1}_{T_1 \geq k}) \middle| \mathscr{F}_n \right) + \mathbb{E}\left((X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1}) \middle| \mathscr{F}_n \right)$$

(Este paso es gracias a que X_k y $\mathbb{1}_{T_1 \geq k}$ son $\mathscr{F}_n$-medibles)

$$= \sum_{k=1}^{n} (X_k \cdot \mathbb{1}_{T_1 \geq k}) + \mathbb{E}((X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1})|\mathscr{F}_n)$$
$$= S_{T_1 \wedge n} + \mathbb{E}((X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1})|\mathscr{F}_n)$$
$$= M_n + \mathbb{E}((X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1})|\mathscr{F}_n)$$

Entonces, nos basta probar que $\mathbb{E}(X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1}|\mathscr{F}_n) = 0$ para terminar nuestra demostración.

Sean $A = \{T_1 = n\}$ y $B = \{T_1 \leq n\}$. Por ser T_1 tiempo de paro, A y B son $\mathscr{F}_n$-medibles. Por lo tanto $B \setminus A$ también es $\mathscr{F}_n$-medible. Notemos que $\{T_1 \geq n+1\} = (B \setminus A)^c$. Por lo tanto $\{T_1 \geq n+1\}$ es $\mathscr{F}_n$-medible. De donde $\mathbb{1}_{T_1 \geq n+1})$ es $\mathscr{F}_n$-medible.

Con esto, ahora tenemos que:

$$\mathbb{E}((X_{n+1} \cdot \mathbb{1}_{T_1 \geq n+1})|\mathscr{F}_n) = \mathbb{1}_{T_1 \geq n+1} \cdot \mathbb{E}(X_{n+1}|\mathscr{F}_n)$$

(Este paso es gracias a que los X_n son independientes)

$$= \mathbb{1}_{T_1 \geq n+1} \cdot \mathbb{E}(X_{n+1})$$
$$= \mathbb{1}_{T_1 \geq n+1} \cdot 0$$
$$= 0$$

Como queríamos demostrar.

Ahora que tenemos que $(M_n)_{n \in \mathbb{N}}$ es martingala, confirmemos que converge casi seguramente.

Notemos que $(T_1 \wedge n)_{n \to \infty} \to T_1$ c.s.

De aquí que $(M_n)_{n \to \infty} = (S_{T_1 \wedge n})_{n \to \infty} = S_{T_1}$ c.s.

Veamos que la convergencia no ocurre en L_1.

Dado que $T_1 \wedge n$ es un tiempo de paro acotado para toda $n \in \mathbb{N}$, podemos aplicar el Teorema de Muestreo Opcional de Doob. El cual nos dice que $\mathbb{E}(M_n) = \mathbb{E}(S_{T_1 \wedge n}) = \mathbb{E}(S_0) = 0$.

Por otro lado, por definición de T_1, $S_{T_1} = 1$ c.s. De donde $\mathbb{E}(S_T) = 1$.

$$\mathbb{E}(M_n) = 0 \not\to 1 = \mathbb{E}(S_{T_1}).$$

Y con esto, queda demostrado que la convergencia no se da en L_1.

1.3.3. **Inciso (iii).** *Sea M_n la martingala obtenida al detener a $-S$ en T_1. Utilice la solución al Problema de la Ruina para probar que $\mathbb{P}(max_n M_n \geq M) = 1/(M+1)$ para todo $M \geq 1$. Concluya que $\mathbb{E}(max_m M_n) = \infty$ y que por lo tanto $\mathbb{E}(max_{m\leq n} M_m) \to \infty$ conforme $n \to \infty$. Finalmente, deduzca que no puede haber una desigualdad de tipo Doob cuando $p = 1$.*

Definimos $M_n = -S_{T_1 \wedge n}$. Notemos que M_n únicamente toma valores en $[-1, \infty]$. Para calcular $\mathbb{P}(max_n M_n \geq M)$ notemos primero que:

$$\mathbb{P}(max_n M_n \geq M) = 1 - \mathbb{P}(max_n M_n < M).$$

$\max_n M_n < M$ significa que M_n nunca alcanza el valor M.

Intentando hacer analogía con el problema de la ruina, pensemos en dos concursantes, uno con 1 peso y otro con M pesos. Nunca alcanzar M significa que nunca gana el que tiene 1 peso.

Esta probabilidad ya la conocemos y es

$$\mathbb{P}(max_n M_n < M) = \frac{M}{M+1}$$

Por lo tanto

$$\begin{aligned}
\mathbb{P}(max_n M_n \geq M) &= 1 - \mathbb{P}(max_n M_n < M) \\
&= 1 - \frac{M}{M+1} \\
&= \frac{M+1}{M+1} - \frac{M}{M+1} \\
&= \frac{1}{M+1}
\end{aligned}$$

Utilizando este resultado:

$$\mathbb{E}(max_n M_n) = -\mathbb{P}(max_n M_n = -1) + \sum_{M=1}^{\infty} \mathbb{P}(max_n M_n \geq M) \tag{1.3.3.1}$$

$$
\begin{aligned}
&= -\mathbb{P}(max_n M_n = -1) + \sum_{M=1}^{\infty} \frac{1}{M+1} \\
&= -\mathbb{P}(max_n M_n = -1) + \infty \\
&= \infty
\end{aligned}
$$

Ahora, tenemos que:

$$
\begin{aligned}
\|\overline{M_n^+}\|_1 &= \mathbb{E}\overline{M_n^+} \\
&= \mathbb{E}\max_{m\leq n} M_m^+ \\
&\geq \mathbb{E}\max_{m\leq n} M_m
\end{aligned}
$$

Donde, el último término, tiende a infinito en base al resultado (1.3.3.1).

Por otro lado:

$$\|M_n^+\|_1 = \| - S_{T_{1\wedge n}}^+\|_1 \longrightarrow \| - S_{T_1}^+\|_1 = 0 < \infty$$

Por lo tanto, no existe número K, tal que

$$\|\overline{M_n^+}\|_1 \leq K\|M_n^+\|_1$$

En otras palabras, no tenemos una desigualdad de tipo Doob para $p = 1$.

1.3.4. **Inciso (iv).** *Sea* $T = \min\{n \geq 2 : S_n = S_{n-2} + 2\}$ *y* $U = T - 2$. *¿Son* T *y* U *tiempos de paro? Justifique su respuesta.*

Intuitivamente, T significa, el primer tiempo tal que ganamos en dos volados consecutivos. También intuitivamente, esto debería ser un tiempo de paro.

Veamos que efectivamente así ocurre. Utilizando la siguiente prueba por inducción:

Base de inducción:

$$\begin{aligned}
\{T = 0\} &= \emptyset && \in \mathscr{F}_0 \\
\{T = 1\} &= \emptyset && \in \mathscr{F}_1 \\
\{T = 2\} &= \{X_1 = 1, X_2 = 1\} && \in \mathscr{F}_2
\end{aligned}$$

Hipótesis de inducción:

Supongamos que $\{T = n\} \in \mathscr{F}_n$ para cierto $n \geq 2$.

Paso inductivo:

$$\{T = n+1\} = \{X_n = 1, X_{n+1} = 1\} \setminus \bigcup_{i=0}^{n} \{T = i\}.$$

Es claro que $\{X_n = 1, X_{n+1} = 1\} \in \mathscr{F}_{n+1}$ y que por hipótesis de inducción $\bigcup_{i=0}^{n} T = i \in \mathscr{F}_n \subset \mathscr{F}_{n+1}$. Por lo tanto $\{T = n+1\} \in \mathscr{F}_{n+1}$ para toda $n \geq 2$ y con esto termina la demostración.

Ahora, intuitivamente U significa el momento justo antes de ganar dos volados consecutivos. Esto, quedría decir que tenemos información sobre eventos que aún no ocurren. Así que intuitivamente esto no debería ser un tiempo de paro.

Efectivamente, si tomamos como ejemplo el conjunto:

$$\{U = 1\} = \{T - 2 = 1\} = \{T = 3\} = \{X_1 = -1, X_2 = 1, X_3 = 1\}$$

Es fácil notar que es un conjunto que pertenece a $\mathscr{F}_3$, pero no a $\mathscr{F}_1$. Pues $\mathscr{F}_1$ no contiene información alguna sobre X_2 y X_3. Así que el conjunto más pequeño de $\mathscr{F}_1$ que contiene a $\{U = 1\}$ es $\{X_1 = -1\}$.

1.3.5. **Inciso (v).** *Para la variable T que hemos definido, calcule $\mathbb{E}(T)$.*

Primero, platicaré de manera intuitiva cómo vamos a proceder para solucionar este problema.

Imaginemos un juego de casino a base de un juego de volados con las siguientes reglas:

- En cada turno, cada jugador tiene que apostar todo el dinero que tiene.
- Si un jugador se queda sin dinero, tiene que abandonar el juego.
- Si cae "sol" cada jugador recibe el doble de lo que había apostado en ese turno.

El casino tiene dinero infinito y cada habitante cuenta con exactamente 1 peso antes de iniciar el juego.

Además, en cada nuevo turno entra exactamente un nuevo jugador al juego.

Para nuestro problema, supongamos que $X_n = 1$ significa que en el n-ésimo turno, salió sol. Entonces, T nos indica cuando es la primera vez que caen dos soles consecutivos.

Sea D_n la variable que indica cuánto dinero ha ganado el casino para el tiempo n.

Observemos que en el momento que cae "águila", todo jugador pierde todo su dinero y abandona el juego. Y que por cada jugador que pierde, el casino gana exáctamente 1 peso (pues cada jugador en cada turno apuesta todo el dinero que posee, es decir el peso con el que empezó y todo lo que le habia ganado al casino).

Entonces, al tiempo $T-2$, todo mundo había perdido. Es decir que al tiempo $T-2$ el casino ha ganado $T-2$ pesos.

Luego, al tiempo $T-1$, ha caido un sol y hay exactamente un jugador al que el casino tuvo que pagar 1 peso.

Al tiempo T, al jugador del turno pasado el casino tuvo que darle 2 pesos y al jugador del nuevo turno tuvo que darle 1 peso.

Entonces, ya podemos decir cuanto dinero ha ganado el casino al tiempo T.

$$D_T = T - 2 - 1 - 3 = T - 6. \tag{1.3.5.1}$$

Notemos que además el juego es justo, en cada turno cada jugador tiene $1/2$ de probabilidad de ganar 2^t y $1/2$ de probabilidad de perder 2^t. Es decir, la esperanza es 0.

D_n es suma de este tipo de variables y por lo tanto su esperanza también será 0.

Esto, nos da la intuición de que D_n es martingala, pero esa es la parte que demostraremos más adelante.

Si logramos demostrar que $\mathbb{E}(D_T) = 0$. De (1.3.5.1) concluimos

$$0 = \mathbb{E}(T - 6) = \mathbb{E}(T) - 6$$

De donde $\mathbb{E}(T) = 6$.

Ahora, para terminar con las formalidades, definamos bien a D y comprobemos que es martingala y que podemos utilizar el Teoremoa del muestreo opcional de Doob como lo hemos hecho.

Sean entonces $(Y_n)_{n\in\mathbb{N}}$ variables aleatorias Bernulli Bernoulli!!!!! de parámetro $1/2$ independientes. Y sea Z_n^m la cantidad de dinero que el jugador m ha dado al casino definidas como:

- Si $n < m$ entonces $Z_n^m = 0$. (El jugador m no participa en el juego sino hasta el turno m).
- $Z_{n+1}^m = (Z_n^m - 1) \cdot 2(Y_{n+1}) + 1$. (Si $Y_{n+1} = 1$ [El jugador gana el volado], entonces el casino pierde la cantidad apostada, que para el turno $n+1$ es $Z_n^n + 1$). Nótese que en cuanto un Y_{n_0} se hace cero, $Z_{n_0}^n$ y todos los que le sigan son todos iguales a 1 (Como el jugador deja el juego después de haber perdido un volado, deja su peso en el casino y entonces de ahí en adelante la cantidad que ha dado al casino es exactamente 1).

Veamos que para cada $m \in \mathbb{N}$, $(Z_n^m)_{n\in\mathbb{N}}$ forma una martingala con respecto a la filtración $(\mathscr{G}_n)_{n\in\mathbb{N}}$ definida por los Y_n.

Cada Z_n^m es $\mathscr{G}_m$-medible por definición.

Cada Z_n^m es suma finita de variables acotadas. Por lo tanto cada una pertenece a L_1.

Sólo nos falta verificar la propiedad de martingala.

$$
\begin{aligned}
\mathbb{E}(Z_{n+1}^m|\mathscr{G}_n) &= \mathbb{E}((Z_n^m - 1)\cdot 2(Y_{n+1}) + 1|\mathscr{G}_n) \\
&= \mathbb{E}((Z_n^m - 1)\cdot 2(Y_{n+1})|\mathscr{G}_n) + \mathbb{E}(1|\mathscr{G}_n) \\
&\quad \text{(Por ser } (Z_n^m - 1) \text{ una variable } \mathscr{G}_n\text{-medible.)} \\
&= (Z_n^m - 1)\cdot \mathbb{E}(2(Y_{n+1})|\mathscr{G}_n) + \mathbb{E}(1|\mathscr{G}_n) \\
&= (Z_n^m - 1)\cdot 2\mathbb{E}(Y_{n+1}|\mathscr{G}_n) + \mathbb{E}(1|\mathscr{G}_n) \\
&\quad \text{(Por ser } Y_{n+1} \text{ independiente } \mathscr{G}_n.\text{)} \\
&= (Z_n^m - 1)\cdot 2\mathbb{E}(Y_{n+1}) + \mathbb{E}(1|\mathscr{G}_n) \\
&= (Z_n^m - 1)\cdot 2\frac{1}{2} + \mathbb{E}(1|\mathscr{G}_n) \\
&= (Z_n^m - 1)\cdot +1 \\
&= Z_n^m.
\end{aligned}
\tag{1.3.5.2}
$$

Ahora definamos a $D_n = \sum_{i=1}^n (Z_n^i)$. Que significa, La cantidad de dinero que el casino ha ganado al tiempo n. Justo como lo habíamos dicho en la "demostración intuitiva".

Ver que D es martingala es fácil. Cada D_n es suma finita de variables finitas y por lo tanto pertenece a L_1. D_n es suma de variables $\mathscr{G}_n$-medibles y por lo tanto también lo es. Y la propiedad de martingala se sigue directamente de (1.3.5.2).

Es cierto que podemos aplicar Doob sobre el tiempo $T \wedge n$ por ser acotado y de aquí que: $\mathbb{E}(D_{T\wedge n}) = \mathbb{E}(D_1) = 0$.

Notemos que $D_{T\wedge n} \longrightarrow D_T$ $c.s.$

Nos gustaría poder decir lo mismo de sus esperanzas y para eso utilizaremos teroema de convergencia dominada.

Es claro que al tiempo $T \wedge n$ el casino a lo más pudo haber ganado $T \wedge n$ pesos. De aquí que $D_{T\wedge n} \geq T \wedge n \geq T$.

Además, por definición de T, para el tiempo T el casino a lo más ha perdido 4 pesos y es la primera vez que pierde tanto. Así que $-4 \leq D_{T\wedge n}$.

Entonces, nuestra martingala D esta dominada por $\max(T,4) < T+4$. Bastaría demostrar que $\mathbb{E}(T+4) < \infty$ para poder utilizar el teorema de convergencia dominada. Para esto, puedes definir a $T' = \min\{n \geq 1 : X_{2n-1} = 1, X_{2n} = 1\}$, notar que T' es geométrica y notar que $T \leq 2T'$.

Notemos que

$$\mathbb{P}(T > n) = \frac{1}{2}\mathbb{P}(T > n-1) + \frac{1}{4}\mathbb{P}(T > n-2). \tag{1.3.5.3}$$

Pues queremos garantizar que en los primeros n turnos, no pierde dos veces consecutivas el casino.

Si en el turno 1, gana el casino (de aquí el $\frac{1}{2}$), en el resto de los $n-1$ turnos tenemos que garantizar que el casino no pierde dos veces consecutivas, como las variables están idénticamente distribuidas, esto es equivalente a que $T > n-1$. (De aqui el $\mathbb{P}(T > n-1)$.

Si el casino pierde en el turno 1, necesariamente tiene que ganar en el turno 2 (de aquí el $\frac{1}{4}$). Y la probabilidad de que no pierda dos veces consecutivas en los siguientes $n-2$ turnos es $\mathbb{P}(T > n-2)$.

De (1.3.5.3) podemos notar fácilmente que $\mathbb{P}(T > 2) \leq \frac{3}{4}$

También notemos que $\mathbb{P}(T > n) \geq \mathbb{P}(T > n+1)$ (El primer evento contiene al segundo).

Ahora,

$$\begin{aligned}
\mathbb{P}(T > 2(2)) &= \frac{1}{2}\mathbb{P}(T > 2+1) + \frac{1}{4}\mathbb{P}(T > 2) \\
&\leq \frac{1}{2}\mathbb{P}(T > 2) + \frac{1}{4}\mathbb{P}(T > 2) \\
&\leq \frac{1}{2}\cdot\frac{3}{4} + \frac{1}{4}\cdot\frac{3}{4} = \left(\frac{3}{4}\right)^2.
\end{aligned}$$

De manera recursiva tenemos que

$$
\begin{aligned}
\mathbb{P}(T > 2(n+1)) &= \frac{1}{2}\mathbb{P}(T > 2(n+1)-1) + \frac{1}{4}\mathbb{P}(T > 2(n+1)-2) \\
&= \frac{1}{2}\mathbb{P}(T > 2(n)+1) + \frac{1}{4}\mathbb{P}(T > 2(n)) \\
&\leq \frac{1}{2}\mathbb{P}(T > 2n) + \frac{1}{4}\mathbb{P}(T > 2n) \\
&\leq \frac{1}{2}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{4}\cdot\left(\frac{3}{4}\right)^n = \left(\frac{3}{4}\right)^{n+1}.
\end{aligned}
$$

Entonces,

$$
\begin{aligned}
\mathbb{E}(T) &= \sum_{n=0}^{\infty}\mathbb{P}(T > n) \\
&\leq \sum_{n=0}^{\infty} 2\mathbb{P}(T > 2n) \\
&= 2\sum_{n=0}^{\infty}\mathbb{P}(T > 2n) \\
&\leq 2\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n < \infty.
\end{aligned}
$$

Por lo tanto $\mathbb{E}(T+4) \leq \infty$. Y entonces nuestra martingala D está dominada por una variable integrable y finalmente podemos aplicar teorema de convergencia dominada para concluir que:

$$
0 = \lim_{n\longrightarrow\infty}\mathbb{E}(D_{T\wedge n}) = \mathbb{E}(D_T).
$$

Y con esto terminamos de demostrar todas las formalidades que nos hacían falta.

1.4. Problema 1.4

Problema 1.4.1 (Extensiones del teorema de paro opcional)**.** Sea $M = (M_n, n \in \mathbb{N})$ una (super)martingala respecto de una filtración $(\mathscr{F}_n, n \in \mathbb{N})$ y sean S y T tiempos de paro.

(i) [1.4.1] Pruebe que $S \wedge T$, $S + T$ y $S \vee T$ son tiempos de paro.

(ii) [1.4.2]

$$\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leq n\} \in \mathscr{F}_n \text{ para toda } n\}$$

es una σ-álgebra, a la que nos referimos como la σ-álgebra detenida en τ. Comente qué puede fallar si T no es tiempo de paro. Pruebe que T es F_T-medible.

(iii) [1.4.3] Pruebe que si T es finito, entonces M_T es $\mathscr{F}_T$-medible.

(iv) [1.4.4] Pruebe que si $S \leq T \leq n$ entonces $\mathscr{F}_S \subset \mathscr{F}_T$. Si además T es acotado entonces $M_S, M_T \in L_1$ y

$$\mathbb{E}(M_T \mid \mathscr{F}_S) = M_S \quad \text{(En el caso de que M sea martingala)}$$
$$\mathbb{E}(M_T \mid \mathscr{F}_S) \leq M_S \quad \text{(En el caso de que M sea supermartingala)}$$

(v) [1.4.5] Si $X = (X_n, n \in \mathbb{N})$ es un proceso estocástico $(\mathscr{F}_n)$-adaptado y tal que $X_n \in L_1$ y tal que para cualesquiera tiempos de paro acotados S y T se tiene que $\mathbb{E}(X_S) = \mathbb{E}(X_T)$ entonces X es una martingala. Sugerencia: considere tiempos de paro de la forma $n\mathbf{1}_A + (n+1)\mathbf{1}_{A^c}$ con $A \in \mathscr{F}_n$.

(vi) [1.4.6] Pruebe que el proceso M^T obtenido al detener a M al instante T y dado por $M_n^T = M_{T \wedge n}$ es una martingala respecto de $(\mathscr{F}_{T\wedge n}, n \geq 0)$ pero también respecto de $(\mathscr{F}_n, n \geq 0)$. Sugerencia: basta probar el resultado respecto de $(\mathscr{F}_n)$ y para esto es útil el inciso anterior.

Categorías: Tiempos de paro, Muestreo opcional

Demostración:

1.4.1. Inciso (i). *Pruebe que $S \wedge T$, $S + T$ y $S \vee T$ son tiempos de paro.*

- Comprobemos que:

$$\{S \wedge T \leq n\} = \{T \leq n\} \cup \{S \leq n\}.$$

Sea $\omega \in \{S \wedge T \leq n\}$. Entonces $T(\omega) \leq n$ ó $S(\omega) \leq n$. Por lo tanto:

$$\{S \wedge T \leq n\} \subset \{T \leq n\} \cup \{S \leq n\}.$$

Por otro lado, si $\omega \in \{T \leq n\} \cup \{S \leq n\}$ entonces $T(\omega) \leq n$ ó $S(\omega) \leq n$. En particular, el mínimo tendrá que ser menor que n y por lo tanto:

$$\{T \leq n\} \cup \{S \leq n\} \subset \{S \wedge T \leq n\}.$$

Por último, $\{T \leq n\} \in \mathscr{F}_n$ y $\{S \leq n\} \in \mathscr{F}_n$. Por lo tanto $\{T \leq n\} \cup \{S \leq n\} = \{S \wedge T \leq n\} \in \mathscr{F}_n$ y con esto demostramos que $S \wedge T$ es tiempo de paro.

- Comprobemos que:

$$\{S \vee T \leq n\} = \{T \leq n\} \cap \{S \leq n\}.$$

Sea $\omega \in \{S \vee T \leq n\}$. Entonces $T(\omega) \leq n$ y $S(\omega) \leq n$. Por lo tanto:

$$\{S \vee T \leq n\} \subset \{T \leq n\} \cap \{S \leq n\}.$$

Por otro lado, si $\omega \in \{T \leq n\} \cap \{S \leq n\}$ entonces $T(\omega) \leq n$ y $S(\omega) \leq n$. En particular, el máximo tendrá que ser menor que n y por lo tanto:

$$\{T \leq n\} \cap \{S \leq n\} \subset \{S \vee T \leq n\}.$$

Por último, $\{T \leq n\} \in \mathscr{F}_n$ y $\{S \leq n\} \in \mathscr{F}_n$. Por lo tanto $\{T \leq n\} \cap \{S \leq n\} = \{S \vee T \leq n\} \in \mathscr{F}_n$ y con esto demostramos que $S \vee T$ es tiempo de paro.

- Comprobemos que:

$$\{S + T = n\} = \bigcup_{i=1}^{n-1} \Big(\{S = n - i\} \cap \{T = i\} \Big).$$

Lost tiempos de paro podrían tomar el valor cero, por lo que el índice i debe comenzar en cero.

Si $\omega \in \{S + T = n\}$, entonces $S(\omega) + T(\omega) = n$, como S y T son positivas, entonces $1 \leq S(\omega) \leq n - 1$. Entonces basta elegir $i = S(\omega)$ para afirmar que $\omega \in \{S = n - i\} \cap \{T = i\}$ y por lo tanto

$$\{S + T = n\} \subset \bigcup_{i=1}^{n-1} \Big(\{S = n - i\} \cap \{T = i\} \Big).$$

Por otro lado, si $\omega \in \bigcup_{i=1}^{n-1} \Big(\{S = n - i\} \cap \{T = i\} \Big)$ significa que existe un i tal que $1 \leq i \leq n - 1$ y $\omega \in \{S = n - i\} \cap \{T = i\}$. Y por lo tanto: $S(\omega) = n - i$ y $T(\omega) = i$. De donde $(T + S)(\omega) = n$.

De aquí que

$$\bigcup_{i=1}^{n-1}\Big(\{S=n-i\}\cap\{T=i\}\Big)\subset\{S+T=n\}.$$

Aohra, para cada i tal que $1\leq i\leq n-1$, $\{S=n-i\}\cap\{T=i\}\in\mathscr{F}_n$. Por lo tanto $\{S+T=n\}\in\mathscr{F}_n$ y con esto queda demostrado que $S+T$ es tiempo de paro.

1.4.2. **Inciso (ii).**

$$\mathscr{F}_T = \{A \in \mathscr{F} : A \cap \{T \leq n\} \in \mathscr{F}_n \textit{ para toda } n\}$$

es una σ-álgebra, a la que nos referimos como la σ-álgebra detenida en τ T. Comente qué puede fallar si T no es tiempo de paro. Pruebe que T es $\mathscr{F}_T$-medible.

Primero hay que demostrar que $\mathscr{F}_T$ es σ-algebra.

- $\Omega \in \mathscr{F}_T$.

  Notemos que

  $$\Omega \cap \{T \leq n\} = \{T \leq n\} \in \mathscr{F}_n.$$

  Donde la pertenencia a $\mathscr{F}_n$ es gracias a que T es tiempo de paro. (Esta parte podria fallar si T no fuera tiempo de paro).

- $\mathscr{F}_T$ es cerrado bajo complementación.

  Sea $A \in \mathscr{F}_T$. Eso significa que para todo $n \in \mathbb{N}$ ocurre que $B = A \cap \{T \leq n\} \in \mathscr{F}_n$. Por ser $\mathscr{F}_n$ una σ-algebra tenemos que el complemento de B también debe estar en $\mathscr{F}_n$. Escrito en símbolos:

  $$\begin{aligned} B^c &= (A \cap \{T \leq n\})^c \\ &= A^c \cup \{T > n\} \in \mathscr{F}_n \end{aligned}$$

  Dado que B^c y $\{T \leq n\}$ se encuentran en $\mathscr{F}_n$, también $B^c \cap \{T \leq n\} \cap \{T \leq n\}$.

  Escribiendo esto último de otra manera:

$$\begin{aligned} B^c \cap \{T \leq n\} \cap \{T \leq n\} &= A^c \cup \{T > n\} \cap \{T \leq n\} \cap \{T \leq n\} \\ &= A^c \cap \{T \leq n\} \cup \{T > n\} \cap \{T \leq n\} \\ &= \Big(A^c \cap \{T \leq n\}\Big) \bigcup \Big(\{T > n\} \cap \{T \leq n\}\Big) \\ &= \Big(A^c \cap \{T \leq n\}\Big) \bigcup \emptyset \end{aligned}$$

$$= A^c \cap \{T \leq n\} \in \mathscr{F}_n$$

Y por lo tanto $A^c \in \mathscr{F}_T$.

- $\mathscr{F}_T$ es cerrado bajo uniones numerables.
  Sean $A_m \in \mathscr{F}_T$ para $m \in \mathbb{N}$.
  $$\left(\bigcup_{k=1}^{\infty} A_k\right) \cap \{T \leq n\} = \bigcup_{k=1}^{\infty} \Big(A_k \cap \{T \leq n\}\Big)$$
  La última parte, por definición de $\mathscr{F}_T$ es una unión numerable de elementos de $\mathscr{F}_n$ y por lo tanto, dicha unión tambien pertenece a $\mathscr{F}_n$. Y por último $\cup A_k \in \mathscr{F}_T$.

Ahora veamos que T es $\mathscr{F}_T$-medible. Dado un $k \in \mathbb{N}$, ocurre que:

$$\{T \leq k\} \cap \{T \leq n\} = \{T \leq min(k, n)\} \in \mathscr{F}_{min(k,n)} \subset \mathscr{F}_n \forall n \in \mathbb{N}.$$

Y por lo tanto $\{T \leq k\} \in \mathscr{F}_T$.

1.4.3. **Inciso (iii).** *Pruebe que si T es finito, entonces M_T es $\mathscr{F}_T$-medible.*

Sea $A \in \mathscr{B}(\mathbb{R})$. Por ser T finito, tenemos que $\{T \in \mathbb{N}\} = \Omega$. Entonces podemos descomponer a $\{M_T \in A\}$ como:

$$\begin{aligned}\{M_T \in A\} &= \bigcup_{i\in\mathbb{N}} \Big(\{M_T \in A\} \cap \{T = i\}\Big)\\ &= \bigcup_{i\in\mathbb{N}} \Big(\{M_i \in A\} \cap \{T = i\}\Big).\end{aligned}$$

Ahora sea $n \in \mathbb{N}$.

$$\begin{aligned}\{M_T \in A\} \cap \{T \leq n\} &= \bigcup_{i\in\mathbb{N}} \Big(\{M_i \in A\} \cap \{T = i\}\Big) \cap \{T \leq n\}\\ &= \bigcup_{i\leq n} \Big(\{M_i \in A\} \cap \{T = i\}\Big).\end{aligned}$$

Por ser M_i martingala con respecto a la filtración $(\mathscr{F}_n)_{n\in\mathbb{N}}$ y T tiempo de paro. Tenemos que $\{T = i\} \in \mathscr{F}_i$ y $\{M_i \in A\} \in \mathscr{F}_i$. Por lo tanto $\{T = i\} \cap \{M_i \in A\} \in \mathscr{F}_i \subset \mathscr{F}_n$ para todo $i \leq n$. Por lo que $\{M_T \in A\} \cap \{T \leq n\} \in \mathscr{F}_n$. Como esto es cierto para toda $n \in \mathbb{N}$, tenemos que $\{M_T \in A\} \in \mathscr{F}_T$. Con lo que terminamos de demostrar que M_T es $\mathscr{F}_T$-medible.

1.4.4. **Inciso (iv).** *Pruebe que si $S \leq T \leq n$ entonces $\mathscr{F}_S \subset \mathscr{F}_T$. Si además T es acotado entonces $M_S, M_T \in L_1$ y*

(1.4.4.1) $$\mathbb{E}(M_T \mid \mathscr{F}_S) = M_S$$

(En el caso de que M sea martingala)

(1.4.4.2) $$\mathbb{E}(M_T \mid \mathscr{F}_S) \leq M_S$$

(En el caso de que M sea supermartingala)

Primero probaremos que $\mathscr{F}_S \subset \mathscr{F}_T$. Sea $n \in \mathbb{N}$. Notemos que $\{T \leq n\} \subset \{S \leq n\}$ pues si $\omega \in \{T \leq n\}$, entonces $S(\omega) \leq T(\omega) \leq n$ y por lo tanto $\omega \in \{S \leq n\}$.

Ahora sea $A \in \mathscr{F}_S$. Entonces

$$\begin{aligned} A \cap \{T \leq n\} &= A \cap \{T \leq n\} \cap \{S \leq n\} \\ &= (A \cap \{S \leq n\}) \cap \{T \leq n\} \in \mathscr{F}_n \end{aligned}$$

Como escogimos n arbitrariamente, esto es cierto para toda $n \in \mathbb{N}$. Y por lo tanto $A \in \mathscr{F}_T$ y $\mathscr{F}_S \subset \mathscr{F}_T$.

Ahora supongamos que T es acotado y sea N una cota para T. Demostraremos que $M_S, M_T \in L_1$.

$$\begin{aligned} \mathbb{E}(|M_T|) &= \mathbb{E}\left(\left|\sum_{i\in\mathbb{N}} M_i \cdot \mathbb{1}_{T=i}\right|\right) \\ &\leq \mathbb{E}\left(\sum_{i\in\mathbb{N}} |M_i \cdot \mathbb{1}_{T=i}|\right) \\ &\leq \mathbb{E}\left(\sum_{i\in\mathbb{N}} |M_i|\right) \\ &= \mathbb{E}\left(\sum_{i\leq N} |M_i|\right) \\ &< \infty \end{aligned}$$

Y con esto queda demostrado que $M_T \in L_1$. La demostración de que $M_S \in L_1$ es completamente análoga.

Nos queda por demostrar (1.4.4.1) y (1.4.4.2).

Demostraremos primero (1.4.4.1).

Dado que S es finito (en particular, es acotado), por lo visto en [1.4.3], tenemos que S es $\mathscr{F}_S$-medible. Así que únicamente falta demostrar que para todo $A \in \mathscr{F}_S$ se tiene que $\mathbb{E}(M_T \cdot \mathbb{1}_A) = \mathbb{E}(M_S \cdot \mathbb{1}_A)$ lo cual es equivalente a que $\mathbb{E}((M_T - M_S) \cdot \mathbb{1}_A) = 0$. Intentaremos No es suficiente intentarlo. demostrar esto último.

Sea $A \in \mathscr{F}_S$.

Notemos que

$$M_T - M_S = \sum_{i \in \mathbb{N}} (M_{i+1} - M_i) \cdot \mathbb{1}_{S \leq i < T}.$$

También notemos que

$$\begin{aligned}
\{S \leq i < T\} \cap A &= \{S \leq i < T\} \cap A \\
&= \{S \leq i\} \cap \{i < T\} \cap A \\
&= (\{S \leq i\} \cap A) \cap \{i < T\}. \\
&= (\{S \leq i\} \cap A) \cap \{T \leq i\}^c.
\end{aligned}$$

De esta última parte hay que puntualizar que $(\{S \leq i\} \cap A) \in \mathscr{F}_i$ por que $A \in \mathscr{F}_S$ y también hay que puntualizar que $\{T \leq i\}^c \{T \leq i+1\}^c \in F_i$.

Con estos detalles listos procedemos a la parte que nos interesa:

$$\begin{aligned}
\mathbb{E}((M_T - M_S) \cdot \mathbb{1}_A) &= \mathbb{E}\left(\sum_{i \in \mathbb{N}} (M_{i+1} - M_i) \cdot \mathbb{1}_{S \leq i < T} \cdot \mathbb{1}_A \right) \\
&= \mathbb{E}\left(\sum_{i \in \mathbb{N}} (M_{i+1} - M_i) \cdot \mathbb{1}_{(\{S \leq i\} \cap A)} \cdot \mathbb{1}_{\{T \leq i\}^c} \right) \\
&= \sum_{i \in \mathbb{N}} \mathbb{E}\Big((M_{i+1} - M_i) \cdot \mathbb{1}_{(\{S \leq i\} \cap A)} \cdot \mathbb{1}_{\{T \leq i\}^c} \Big)
\end{aligned}$$

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{E}\Bigg((M_{i+1} - M_i)\cdot \mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c}\Bigg| F_i\Bigg)\Bigg)$$

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c} \cdot \mathbb{E}\Bigg((M_{i+1} - M_i)\Bigg| F_i\Bigg)\Bigg)$$

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c} \cdot \mathbb{E}(M_{i+1}|F_i) - \mathbb{E}(M_i|F_i)\Bigg)$$

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c} \cdot M_i - \mathbb{E}(M_i|F_i)\Bigg) \tag{1.4.4.3}$$

(Esto último por ser M una martingala.)

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c} \cdot M_i - M_i\Bigg)$$

(Esto último por que M_i es F_i-medible.)

$$= \sum_{i\in\mathbb{N}} \mathbb{E}\Bigg(\mathbb{1}_{(\{S\leq i\}\cap A)} \cdot \mathbb{1}_{\{T\leq i\}^c} \cdot 0\Bigg)$$

$$= 0$$

Que es justo lo que queríamos demostrar.

Para demostrar (1.4.4.2). Hay que notar que $\mathbb{E}(M_T \mid \mathscr{F}_S) \leq M_S$ es cierto sí y sólo sí para todo $A \in \mathscr{F}_S$ se tiene que $\mathbb{E}((M_T - M_S)\mathbb{1}_A) \leq 0$.

Demostrar esto último lo podemos hacer de manera análoga a como demostramos el caso en el que M es martingala. En las cuentas previamente hechas, utilizando el hecho de que M es supermartingala en vez de martingala, la igualdad de (1.4.4.3) se transforma en una desigualdad de tipo "$\leq$". El resto se queda exactamente igual y entonces tenemos la demostración que necesitabamos.

1.4.5. **Inciso (v).** *Si* $X = (X_n, n \in \mathbb{N})$ *es un proceso estocástico* $(\mathscr{F}_n)$*-adaptado y tal que* $X_n \in L_1$ *y tal que para cualesquiera tiempos de paro acotados* S *y* T *se tiene que* $\mathbb{E}(X_S) = \mathbb{E}(X_T)$ *entonces* X *es una martingala. Sugerencia: considere tiempos de paro de la forma* $n\mathbf{1}_A + (n+1)\mathbf{1}_{A^c}$ *con* $A \in \mathscr{F}_n$.

Ya tenemos dos de las hipótesis necesarias para demostrar que algo es martingala.

(1) $(X_n)_{n\in\mathbb{N}}$ es $(\mathscr{F}_n)_{n\in\mathbb{N}}$-adaptado.
(2) $X_n \in L_1$ para toda $n \in \mathbb{N}$.

Sólo falta demostrar que $\mathbb{E}(\, X_{n+1} \mid \mathscr{F}_n) = X_n$ para toda $n \in \mathbb{N}$.

Es decir, dada una $n \in \mathbb{N}$, queremos demostar que para todo $A \in \mathscr{F}_n$, ocurre que $\mathbb{E}(X_{n+1} \cdot \mathbb{1}_A) = \mathbb{E}(X_n \cdot \mathbb{1}_A)$.

Sean entonces $n \in \mathbb{N}$ y $A \in \mathscr{F}_n$. Definimos $T_{A,n} = n \cdot \mathbb{1}_A + (n+1) \cdot \mathbb{1}_{A^c}$. Veamos que $T_{A,n}$ es tiempo de paro.

Sea $k \in \mathbb{N}$.

- Si $k = n$, entonces

$$\begin{aligned}\{T_{A,n} = k\} &= \{T_{A,n} = n\} \\ &= A \in F_n \mathscr{F}_n = F_k \mathscr{F}_k.\end{aligned}$$

- Si $k = n+1$, entonces

$$\begin{aligned}\{T_{A,n} = k\} &= \{T_{A,n} = n+1\} \\ &= A^c \in F_n \subset \mathscr{F}_{n+1} = F_k.\end{aligned}$$

- Si $k \neq n$ y $k \neq n+1$, entonces

$$\{T_{A,n} = k\} = \emptyset \in \mathscr{F}_k.$$

Entonces, $\{T_{A,n} = k\} \in \mathscr{F}_k$ para toda $k \in \mathbb{N}$ y por lo tanto $T_{A,n}$ es tiempo de paro.

Ahora notemos lo siguiente

$$
\begin{aligned}
X_{T_{A,n}} &= \sum_{i\in\mathbb{N}} X_i \cdot \mathbb{1}_{\{T=i\}} \\
&= X_n \cdot \mathbb{1}_{\{T=n\}} + X_{n+1} \cdot \mathbb{1}_{\{T=n+1\}} \\
&= X_n \cdot \mathbb{1}_A + X_{n+1} \cdot \mathbb{1}_{A^c}.
\end{aligned}
$$

Eso quiere decir que

$$
\begin{aligned}
\mathbb{E}(X_{T_{A,n}}) &= \mathbb{E}(X_n \cdot \mathbb{1}_A + X_{n+1} \cdot \mathbb{1}_{A^c}) \\
&= \mathbb{E}(X_n \cdot \mathbb{1}_A) + \mathbb{E}(X_{n+1} \cdot \mathbb{1}_{A^c}).
\end{aligned}
$$

Ahora tomemos la función constante $S_n = n+1$. Por ser constante $\{S_n = k\} \in \{\Omega, \emptyset\} \subset \mathscr{F}_k$ para toda $k \in \mathbb{N}$ y por lo tanto S_n es tiempo de paro.

Notemos entonces que

$$
\begin{aligned}
\mathbb{E}(X_{S_n}) &= \mathbb{E}(X_{n+1}) \\
&= \mathbb{E}(X_{n+1} \cdot \mathbb{1}_A + X_{n+1} \cdot \mathbb{1}_{A^c}) \\
&= \mathbb{E}(X_{n+1} \cdot \mathbb{1}_A) + \mathbb{E}(X_{n+1} \cdot \mathbb{1}_{A^c}).
\end{aligned}
$$

Dado que tanto $T_{a,n}$ como S_n son variables que toman una cantidad finita de valores y por lo tanto son acotadas, nuestra hipótesis dice que $\mathbb{E}(X_{S_n}) = \mathbb{E}(X_{T_{A,n}})$.

Entonces

$$
\mathbb{E}(X_{n+1} \cdot \mathbb{1}_A) + \mathbb{E}(X_{n+1} \cdot \mathbb{1}_{A^c}) = \mathbb{E}(X_n \cdot \mathbb{1}_A) + \mathbb{E}(X_{n+1} \cdot \mathbb{1}_{A^c}).
$$

De donde

$$
\mathbb{E}(X_{n+1} \cdot \mathbb{1}_A) = \mathbb{E}(X_n \cdot \mathbb{1}_A).
$$

Como esto lo hicimos para $n \in \mathbb{N}$ y $A \in \mathscr{F}_n$ arbitrarios, con esto terminamos la demostración de que $(X_n)_{n\in\mathbb{N}}$ es martingala.

1.4.6. **Inciso (vi).** *Pruebe que el proceso M^T obtenido al detener a M al instante T y dado por $M_n^T = M_{T\wedge n}$ es una martingala respecto de $(\mathscr{F}_{T\wedge n}, n \geq 0)$ pero también respecto de $(\mathscr{F}_n, n \geq 0)$. Sugerencia: basta probar el resultado respecto de $(\mathscr{F}_n)$ y para esto es útil el inciso anterior.*

Es claro que M^T es $(\mathscr{F}_n)_{n\in\mathbb{N}}$-adaptado. Pues

$$\begin{aligned} M_n^T &= M_{n\wedge T} \\ &= \sum_{i\in\mathbb{N}} M_i \cdot \mathbb{1}_{n\wedge T=i} \\ &= \sum_{i\leq n} M_i \cdot \mathbb{1}_{T=i}. \end{aligned}$$

Que resulta ser suma finita de variables que son F_n-medibles y por lo tanto, es $\mathscr{F}_n$-medible.

De aquí mismo podemos notar que M_n^T es suma finita de variables en L_1 y por lo tanto, pertenece a L_1.

Ahora, sean S y U dos tiempos de paro acotados.

Por [1.4.1] sabemos que $S\wedge T$ y $S\wedge U$ también son tiempos de paro.

Además por ser S y U acotados, $S\wedge T$ y $S\wedge U$ son acotados también.

Sabiendo que M es martingala, con el teorema del muestreo opcional de Doob, podemos notar que

$$\begin{aligned} \mathbb{E}(M_S^T) &= \mathbb{E}(M_{S\wedge T}) \\ &= \mathbb{E}(M_0) \end{aligned}$$

y

$$\begin{aligned} \mathbb{E}(M_U^T) &= \mathbb{E}(M_{U\wedge T}) \\ &= \mathbb{E}(M_0). \end{aligned}$$

Por lo tanto, para cualesquiera dos tiempo de paro acotados S y U tenemos que $\mathbb{E}(M_S^T) = \mathbb{E}(M_U^T)$. Entonces, por lo visto en [1.4.5], M^T es una martingala con respecto de la filtración $(F_n)_{n\in\mathbb{N}}$.

Para probar que M_T es una martingala con respecto a $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$, primero notemos que $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$ es efectivamente una filtración.

Dado $n \in \mathbb{N}$, gracias a [1.4.2], tenemos que $\mathscr{F}_{T\wedge n}$ es una σ-álgebra.

Además tenemos que $T \wedge n \leq T \wedge (n+1)$. El resultado de [1.4.4] implica que $\mathscr{F}_{T\wedge n} \subset \mathscr{F}_{T\wedge(n+1)}$ y por lo tanto $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$ sí es una filtración.

Ahora veamos que M^T es adaptado a la filtración $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$.

Dado $n \in \mathbb{N}$ como $T \wedge n$ es acotado, por [1.4.3] $M_n^T = M_{T\wedge n}$ es $\mathscr{F}_{T\wedge n}$-medible. Y con esto tenemos que M^T sí es adaptado a la filtración $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$.

Ya sabíamos que M_n^T pertenecía a L_1.

Gracias [1.4.5] y a la primera parte que se demostró en este ejercicio, se sigue que M_T también es martingala con respecto a la filtración $(\mathscr{F}_{T\wedge n})_{n\in\mathbb{N}}$.

Part 2. Tarea 2

2.1. Problema 2.1

Problema 2.1.1. Sea $S_n = X_1 + \cdots + X_n$ una caminata aleatoria con saltos $X_i \in \{-1, 0, 1, \ldots\}$. Sea C_p una variable aleatoria geométrica de parámetro p independiente de S y definimos

$$M_p = -\min_{n \leq C_p} S_n.$$

El objetivo del ejercicio es determinar la distribución de M_p.

(A las caminatas aleatorias como S se les ha denominado Skip-free random walks Para aplicaciones de este tipo de procesos, ver [**?**]. También aparecen en el estudio de Procesos Galton-Watson. Este ejercicio es el resultado básico del estudio de sus extremos, denominado teoría de fluctuaciones.)

(i) [2.1.1] Sea

$$g(\lambda) = E(e^{-\lambda X_1}).$$

Pruebe que $g(\lambda) \in (0, \infty)$ y que

$$M_n = e^{-\lambda S_n} g(\lambda)^{-n}, n \geq 0$$

es una martingala.

(ii) [2.1.2] Pruebe que g es log-convexa al aplicar la desigualdad de Hölder. Pruebe que si $P(X_1 = -1) > 0$ (hipótesis que se utilizará desde ahora) entonces $g(\lambda) \to \infty$ conforme $\lambda \to \infty$. Utilice esta información para esbozar la gráfica de g. Defina $f(s) = \inf\{\lambda > 0 : g(\lambda)^{-1} < s\}$. Note que $1/g \circ f = Id$ en $(0, 1)$. Pruebe que si $g(\lambda) > 1$, la martingala M es acotada hasta el tiempo de arribo de S a $-k$ dado por

$$T_k = \min\{n \in \mathbb{N} : S_n = -k\}$$

(donde se utiliza la convención $\inf \emptyset = \infty$). Aplique el teorema de muestreo opcional de Doob para mostrar que

$$E(s^{T_k}) = e^{-kf(s)}.$$

Justifique MUY bien por qué la fórmula es válida aún cuando T_k puede tomar el valor ∞ y deduzca que de hecho $\mathbb{P}(T_k = \infty) = 0$.

(iii) [2.1.3] Argumente que

$$\mathbb{P}(M_p \geq n) = \mathbb{P}(T_n \leq C_p) = \mathbb{E}((1-p)^{T_n})$$

para demostrar que M_p tiene distribución geométrica de parámetro $1 - e^{-f(1-p)}$.

(iv) [2.1.4] Tome el límite conforme $p \to 0$ para mostrar que la variable aleatoria

$$M = -\min_{n \geq 0} S_n$$

tiene una distribución geométrica de parámetro $1 - e^{-f(1)}$. Interprete esto cuando $f(1) = 0$.

Categorías: Caminatas aleatorias, muestreo opcional, fluctuaciones.

Demostración:

2.1.1. **Inciso (i).** *Sea*

$$g(\lambda) = E(e^{-\lambda X_1}).$$

Pruebe que

(2.1.1.1)
$$g(\lambda) \in (0, \infty)$$

y que

$$M_n = e^{-\lambda S_n} g(\lambda)^{-n}, n \geq 0$$

es una martingala.

Para probar (2.1.1.1) notemos que

$$e^x > 0 \ \ \forall x \in \mathbb{R}.$$

por lo tanto

$$e^{-\lambda X_1} > 0$$

y entonces

$$\mathbb{E}(e^{-\lambda X_1}) > \mathbb{E}(0) = 0$$

Falta ver que $\mathbb{E}(e^{-\lambda X_1})$ está acotada. Recordemos que e^x es creciente y por lo tanto

$$\begin{aligned}
X_1 &\geq -1 && \Rightarrow \\
1 &\geq -X_1 && \Rightarrow \\
\lambda &\geq -\lambda X_1 && \Rightarrow \\
e^{\lambda} &\geq e^{-\lambda X_1} && \Rightarrow \\
\mathbb{E}(e^{\lambda}) &\geq \mathbb{E}(e^{-\lambda X_1})
\end{aligned}$$

Pero $\mathbb{E}(e^{\lambda}) = e^{\lambda}$. De donde

$$e^{\lambda} \geq \mathbb{E}(e^{-\lambda X_1})$$

Entonces $0 < \mathbb{E}(e^{-\lambda X_1}) < \infty$, y con esto queda demostrado (2.1.1.1).

Probemos ahora que $(M_n)_{n\in\mathbb{N}}$ es martingala.

- Veamos que M_n es $\mathscr{F}_n$-medible.

  Para esto, recordemos que si $\mathscr{G}$ es una σ-algebra, f es $\mathscr{G}$-medible y g es continua con la topología usual de $\mathbb{R}$ entonces $g \circ f$ es $\mathscr{G}$-medible.

  S_n es $\mathscr{F}_n$-medible. Multiplicar por constantes es función continua. Dividir entre constantes es función continua. La función exponencial es continua. Y por lo tanto $M_n = e^{-\lambda S_n} g(\lambda)^{-n}$ es composición de funciones continuas con S_n. De donde M_n es F_n-medible.

- Ahora veamos que $M_n \in L_1$.

  De la misma demostración que dimos para probar la primera parte, se sigue que $0 < e^{-\lambda X_n} < \infty$ y que $e^{-\lambda X_n} \in \mathbb{N}$ para toda $n \in \mathbb{N}$. Por lo tanto

  $$\begin{aligned}
  M_n &= e^{-\lambda S_n} g(\lambda)^{-n} \\
  &= \prod_{1\leq i\leq n} e^{-\lambda X_i} g(\lambda)^{-n}.
  \end{aligned}$$

  Entonces, M_n es producto finito de funciones en L_1, de donde tenemos que $M_n \in L_1$. Mucho cuidado, esto es incorrecto. La idea es acotar a S_n por $-n$ como lo hiciste en algún inciso anterior con S_1.

- Únicamente falta demostrar que $\mathbb{E}(M_{n+1}|\mathscr{F}_n) = M_n$.

$$
\begin{aligned}
\mathbb{E}(M_{n+1}|\mathscr{F}_n) &= \mathbb{E}(e^{-\lambda S_{n+1}} g(\lambda)^{-(n+1)}|\mathscr{F}_n) \\
&= g(\lambda)^{-(n+1)}\mathbb{E}(e^{-\lambda S_{n+1}}|\mathscr{F}_n) \\
&\quad \text{(Esto último es gracias a que } g(\lambda)^{-(n+1)} \text{ es constante)} \\
&= g(\lambda)^{-(n+1)}\mathbb{E}(e^{-\lambda S_{n+1}}|\mathscr{F}_n) \\
&= g(\lambda)^{-(n+1)}\mathbb{E}(e^{-\lambda S_n \cdot + X_{n+1}}|\mathscr{F}_n) \\
&= g(\lambda)^{-(n+1)}\mathbb{E}(e^{-\lambda S_n} \cdot e^{-\lambda X_{n+1}}|\mathscr{F}_n) \\
&= g(\lambda)^{-(n+1)} e^{-\lambda S_n}\mathbb{E}(e^{-\lambda X_{n+1}}|\mathscr{F}_n) \\
&\quad \text{(Esto último es gracias a que } S_n \text{ es } F_n\text{-medible)} \\
&= g(\lambda)^{-(n+1)} e^{-\lambda S_n}\mathbb{E}(e^{-\lambda X_{n+1}}) \\
&\quad \text{(Esto último es gracias a que } X_{n+1} \text{ es independiente de } F_n\text{)} \\
&= g(\lambda)^{-(n+1)} e^{-\lambda S_n}\mathbb{E}(e^{-\lambda X_1}) \\
&\quad \text{Esto último es gracias a que } X_{n+1} \text{ y } X_1 \text{ tienen la)} \\
&\quad \text{(misma distribución} \\
&= g(\lambda)^{-(n+1)} e^{-\lambda S_n} g(\lambda) \\
&= g(\lambda)^{-(n)} e^{-\lambda S_n} \\
&= M_n.
\end{aligned}
$$

Y con esto terminamos de demostrar que $(M_n)_{n\in\mathbb{N}}$ es martingala.

2.1.2. **Inciso (ii).** *Pruebe que g es* log*-convexa al aplicar la desigualdad de Hölder. Pruebe que si $P(X_1 = -1) > 0$ (hipótesis que se utilizará desde ahora) entonces $g(\lambda) \to \infty$ conforme $\lambda \to \infty$. Utilice esta información para esbozar la gráfica de g. Defina $f(s) = \inf\{\lambda > 0 : g(\lambda)^{-1} < s\}$. Note que $1/g \circ f = Id$ en $(0,1)$. Pruebe que si $g(\lambda) > 1$, la martingala M es acotada hasta el tiempo de arribo de S a $-k$ dado por*

$$T_k = \min\{n \in \mathbb{N} : S_n = -k\}$$

(donde se utiliza la convención $\inf \emptyset = \infty$). Aplique el teorema de muestreo opcional de Doob para mostrar que

$$E(s^{T_k}) = e^{-kf(s)}.$$

Justifique MUY bien por qué la fórmula es válida aún cuando T_k puede tomar el valor ∞ y deduzca que de hecho $\mathbb{P}(T_k = \infty) = 0$.

Primero probemos que g es log-convexa, es decir, que $\log \circ g$ es convexa.Sean $a < b$ en el dominio de g y sea $t \in [0,1]$.

Queremos demostrar que:

$$\log \circ g((1-t)a + (t)b) \leq (1-t)(\log \circ g(a)) + (t)(\log \circ g(b)).$$

Si $t = 0$ tenemos:

$$\begin{aligned}\log \circ g((1-0)a + (0)b) &= \log \circ g(a) \\ &= (1-0)(\log \circ g(a)) + (0)(\log \circ g(b))\end{aligned}$$

y por lo tanto no hay nada que demostrar. Análogamente ocurre con $t = 1$.

Entonces concentrémonos en el caso donde $t \in (0,1)$.

Recordemos que la desigualdad de Hölder dice que si $f \in L_p$ y $g \in L_q$ con $p, q \in (1, \infty)$ y $\frac{1}{p} + \frac{1}{q} = 1$. Entonces

$$\mathbb{E}(|fg|) \leq (\mathbb{E}(|f|^{\,p}))^{\frac{1}{p}} (\mathbb{E}(|g|^{\,q}))^{\frac{1}{q}}. \tag{2.1.2.1}$$

Sean entonces $p = \frac{1}{1-t}$ y $q = \frac{1}{t}$. Tenemos que

$$\begin{aligned}\frac{1}{p} + \frac{1}{q} &= \frac{1}{\frac{1}{1-t}} + \frac{1}{\frac{1}{t}} \\ &= t + 1 - t\end{aligned}$$

$$= 1.$$

Veamos que $e^{-a(1-t)X_1}$ pertenece a L_p.

Como $e^x > 0$ para todo $x \in \mathbb{R}$. $|e^x| = e^x$. Entonces

$$\begin{aligned}
\left(\mathbb{E}\left(\left|e^{-a(1-t)X_1}\right|^{p}\right)\right)^{\frac{1}{p}} &= \left(\mathbb{E}\left(\left|e^{-a(1-t)X_1}\right|^{\frac{1}{1-t}}\right)\right)^{1-t}\\
&= \left(\mathbb{E}\left(\left|e^{-aX_1}\right|\right)\right)^{1-t}\\
&= \left(\mathbb{E}\left(e^{-aX_1}\right)\right)^{1-t}\\
&= (g(a))^{1-t}\\
&< \infty
\end{aligned}$$

Donde la última desigualdad es gracias a que a fue tomado en el dominio de g y a lo demostrado en [2.1.1]. Con esto hemos demostrado que $e^{-a(1-t)X_1}$ pertenece a L_p.

De manera análoga se puede demostrar que $e^{-b(t)X_1}$ pertenece a L_q.

Ahora que tenemos todas las hipótesis para la desigualdad de Hölder, basta aplicarla.

$$\begin{aligned}
\log \circ g((1-t)a + (t)b) &= \log\left(\mathbb{E}\left(e^{-a(1-t)+b(t)X_1}\right)\right)\\
&= \log\left(\mathbb{E}\left(e^{-a(1-t)X_1} \cdot e^{-b(t)X_1}\right)\right)\\
&= \log\left(\mathbb{E}\left(\left|e^{-a(1-t)X_1} \cdot e^{-b(t)X_1}\right|\right)\right)\\
&\leq \log\left(\mathbb{E}\left(\left|e^{-a(1-t)X_1}\right|^{p}\right)^{\frac{1}{p}} \cdot \mathbb{E}\left(\left|e^{-b(t)X_1}\right|^{q}\right)^{\frac{1}{q}}\right)\\
&\qquad \text{(Esta desigualdad es gracias a la desigualdad de}\\
&\qquad \text{Hölder y a que log es una función creciente)}\\
&= \log\left(\mathbb{E}\left(e^{-a(1-t)X_1 p}\right)^{\frac{1}{p}} \cdot \mathbb{E}\left(e^{-b(t)X_1 q}\right)^{\frac{1}{q}}\right)
\end{aligned}$$

$$
\begin{aligned}
&= \log\left(\mathbb{E}\left(e^{-a(1-t)X_1\frac{1}{1-t}}\right)^{\frac{1}{\frac{1}{1-t}}} \cdot \mathbb{E}\left(e^{-b(t)X_1\frac{1}{t}}\right)^{\frac{1}{\frac{1}{t}}}\right) \\
&= \log\left(\mathbb{E}\left(e^{-aX_1}\right)^{1-t} \cdot \mathbb{E}\left(e^{-bX_1}\right)^{t}\right) \\
&= \log\left(\mathbb{E}\left(e^{-aX_1}\right)^{1-t}\right) + \log\left(\mathbb{E}\left(e^{-bX_1}\right)^{t}\right) \\
&= \log\left(g(a)^{1-t}\right) + \log\left(g(b)^{t}\right) \\
&= (1-t)\log(g(a)) + (t)\log(g(b)) \\
&= (1-t)(\log \circ g(a)) + (t)(\log \circ g(b))
\end{aligned}
$$

Que es lo que necesitabamos mostrar para probar que g es log-convexa.

Ahora supongamos que $\mathbb{P}(X_1 = -1) > 0$.

Para ver que g tiende a infinito conforme λ crece, descompongamos a $g(\lambda)$ de la siguiente manera.

$$
\begin{aligned}
g(\lambda) &= \mathbb{E}(e^{-\lambda X_1}) \\
&= \sum_{-1 \leq i} e^{-\lambda i}\mathbb{P}(X_1 = i)
\end{aligned}
$$

De donde obtenemos que $g(\lambda) \geq e^{\lambda}\mathbb{P}(X_1 = -1)$. Dado que e^{λ} tiende a infinito comforme λ crece, tenemos que $g(\lambda)$ también lo hace.

Probemos ahora que $(\frac{1}{g}) \circ f(s) = s$ para toda $s \in (0, 1)$.

Ahora notemos que g es convexa también. $\log \circ g$ es convexa por la primera parte de este inciso y e^x es convexa y creciente porque su primera y segunda derivada siempre son mayor que cero. Dado esto, notemos que $g = e^{\log(g)}$ y entonces por ser g una composición de una función convexa con una convexa creciente tenemos que g es convexa.

Uno de los resultados de los cursos de cálculo de la licencuatura es que una función convexa con dominio abierto, es continua. g está definida sobre todo $\mathbb{R}$, que es abierto, y por lo tanto g es continua.

Ahora sea $s \in (0, 1)$. Y sea $\lambda_0 = f(s)$. Por definición, $\lambda_0 = \inf\{\lambda > 0 : g(\lambda)^{-1} < s\}$. Por definición de ínfimo, para cualquier $n \in \mathbb{N}$, existe $\lambda_n > 0$ tal que

(2.1.2.2) $$\lambda_0 \leq \lambda_n \leq \lambda_0 + \frac{1}{n}.$$

y que

(2.1.2.3) $$g(\lambda_n)^{-1} < s.$$

Por (2.1.2.2) sabemos que $\lambda_n \to \lambda_0$. De (2.1.2.3) obtenemos $\frac{1}{s} < g(\lambda_n)$. Por continuidad de g tenemos entonces que $\frac{1}{s} \leq g(\lambda_0)$.

Supongamos que $\frac{1}{s} < g(\lambda_0)$. Entonces, por continuidad de g, existe $\varepsilon > 0$ tal que $\varepsilon < \frac{\lambda_0}{2}$ y tal que para toda $x \in \mathbb{R}$ es cumple que $|\lambda_0 - x| < \varepsilon$ entonces $\frac{1}{s} < g(x)$.

Sea $x_0 = \lambda_0 - \frac{\varepsilon}{2}$. Entonces $|\lambda_0 - x_0| = \frac{\varepsilon}{2} < \varepsilon$ y por lo tanto $\frac{1}{s} < g(x_0)$. Entonces $\frac{1}{s} < g(x_0)$ y por lo tanto $x_0 \in \{\lambda > 0 : g(\lambda)^{-1} < s\}$. Pero como $x_0 < \lambda_0$, esto contradice que λ_0 sea el ínfimo de dicho conjunto. Por lo tanto $g(\lambda_0) = \frac{1}{s}$.

De esto último, tenemos que $g(\lambda_0) = g(f(s)) = \frac{1}{s}$ y por lo tanto $\frac{1}{g(f(s))} = (\frac{1}{g}) \circ f(s) = s$ y con esto terminamos la demostración que buscábamos.

Ahora supongamos que $g(\lambda) > 1$.

Sea $n \in \mathbb{N}$ y $A_n = \{n \leq T_k\}$. Entonces $S_n(A_n) \geq -k$ (abusando de la notación Busca en la red la notación del Iverson bracket.) y por lo tanto $-\lambda S_n(A_n) \leq \lambda k$. Por ser e^x una función estrictamente creciente, $e^{-\lambda S_n(A_n)} \leq e^{\lambda k}$.

Entonces

$$\begin{aligned} M_n(A_n) &= e^{-\lambda S_n} g(\lambda)^{-1}(A_n) \\ &\leq e^{-\lambda k} g(\lambda)^{-1} \\ &\leq e^{-\lambda k} \end{aligned}$$

Donde en el último paso se usa la hipótesis de que $g(\lambda) > 1$.

Esto nos dice que la martingala M_n es acotada hasta el tiempo de paro T_k (Antes de T_k, M_n es acotada por $e^{-\lambda k}$). Que es lo que buscábamos demostrar. Nótese que para esta demostración nunca asumimos que $T_k < \infty$ y por lo tanto es perfectamente válida cuando $T_k = \infty$.

Para la siguiente parte, sea $n \in \mathbb{N}$ y sea $T_k^n = T_k \wedge n$. Sabemos que T_k^n es tiempo de paro y por definición es acotado. Entonces el teorema de muestreo opcional de Doob nos dice que:

$$\begin{aligned}\mathbb{E}(M_{T_k^n}) &= \mathbb{E}(M_1) \\ &= \mathbb{E}(e^{-\lambda S_1} g(\lambda)^{-1}) \\ &= g(\lambda)^{-1}\mathbb{E}(e^{-\lambda S_1}) \\ &= g(\lambda)^{-1} g(\lambda) \\ &= 1\end{aligned}$$

Como T_k^n no rebasa a T_k, entonces, por la parte anterior de este ejercicio $M_{T_k^n}$ es acotada.

Eso quiere decir que podemos aplicar el teorema de convergencia dominada y entonces.

$$1 = \lim_{n\to\infty} \mathbb{E}(M_{T_k^n}) = \mathbb{E}(M_{T_k}).$$

Ahora, recordando la definición de M_n

$$\begin{aligned}1 &= \mathbb{E}(M_{T_k}) \\ &= \mathbb{E}(e^{-\lambda S_{T_k}} g(\lambda)^{-T_k}) \\ &= \mathbb{E}(e^{\lambda k} g(\lambda)^{-T_k}) \\ &= e^{\lambda k}\mathbb{E}(g(\lambda)^{-T_k}).\end{aligned}$$

Sea ahora $\lambda = f(s)$. Entonces

$$\begin{aligned}1 &= e^{\lambda k}\mathbb{E}(g(\lambda)^{-T_k}) \\ &= e^{f(s)k}\mathbb{E}(g(f(s))^{-T_k}) \\ &= e^{f(s)k}\mathbb{E}((g(f(s))^{-1})^{T_k}) \\ &= e^{f(s)k}\mathbb{E}(s^{T_k}).\end{aligned}$$

De donde $e^{-f(s)k} = \mathbb{E}(s^{T_k})$. Como queríamos demostrar.

Ahora, si $T_k = \infty$, ya que $s \in (0,1)$, entonces $s^(T_k)s^{T_k} = 0$.

Los dos párrafos siguientes son incorrectos.Lo que puedes argumentar es que $M_{n\wedge T_k}$ converge a $e^{\lambda k}g(\lambda)^{-T_k}$ si $T_k \to \infty$ pero que sobre el conjunto $T_k = \infty$, el límite es cero pues el primer factor es acotado y el segundo tiende a cero. Por otro lado si $T_k = \infty$, entonces $S_n > -k$ para toda n. Entonces $-\lambda S_n < \lambda k$ y por ser e^x una función estrictamente creciente entonces $e^{-\lambda S_n} < e^{\lambda k}$. Tomando esperanza tenemos

$$\begin{aligned}
(g(\lambda))^n &= \mathbb{E}(e^{-\lambda X_1})^n \\
&= \mathbb{E}(e^{-\lambda S_n}) \\
&\quad \text{(Por ser las } X_i \text{ independientes entre sí)} \\
&< \mathbb{E}(e^{\lambda k}) \\
&= e^{\lambda k}.
\end{aligned}$$

De donde $g(\lambda) < e^{\frac{\lambda k}{n}}$ para toda n. pero $e^{\frac{\lambda k}{n}}$ tiende a 1 conforme n tiende a infinito. Por lo tanto $g(\lambda) \leq 1$. Entonces, si $s \in (0,1)$, $\frac{1}{s} \in (1, \infty)$. Por lo tanto, no existe λ tal que $\frac{1}{s} < g(s)$. Es decir $f(s) = \inf\{\lambda > 0 : \frac{1}{s} < g(s)\} = \inf \emptyset = \infty$.

Entonces, si $T_k = \infty$, ocurre que

$$\begin{aligned}
\mathbb{E}(s^{T_k}) &= \mathbb{E}(0) \\
&= 0 \\
&= e^{-f(s)k}.
\end{aligned}$$

Y con esto queda cubierto el caso en que $T_k = \infty$.

2.1.3. **Inciso (iii).** *Argumente que*

$$\begin{aligned}\mathbb{P}(M_p \geq n) &= \mathbb{P}(T_n \leq C_p) \\ &= \mathbb{E}((1-p)^{T_n})\end{aligned}$$

para demostrar que M_p *tiene distribución geométrica de parámetro* $1-e^{-f(1-p)}$.

Recordemos que

$$M_p = -\min_{n \leq C_p} S_n.$$

Donde C_p es una variable aleatoria con distribución geométrica de parámetro p.

Entonces

$$\begin{aligned}\{M_p \geq n\} &= \{-\min_{k \leq C_p} S_k \geq n\} \\ &= \{\min_{k \leq C_p} S_k \leq -n\} \\ &= \{\min\{k \in \mathbb{N} : S_k = -n\} \leq C_p\} \\ &= \{T_n \leq C_p\}.\end{aligned}$$

De donde tenemos $\mathbb{P}(M_p \geq n) = \mathbb{P}(T_n \leq C_p)$.

Ahora descompongamos $\mathbb{P}(T_n \leq C_p)$.

$$\begin{aligned}\mathbb{P}(T_n \leq C_p) &= \sum_{i \in \mathbb{N}} \mathbb{P}(T_n \leq C_p | T_n = i)\mathbb{P}(T_n = i) \\ &= \sum_{i \in \mathbb{N}} \mathbb{P}(i \leq C_p)\mathbb{P}(T_n = i) \\ &= \sum_{i \in \mathbb{N}} \sum_{j=i}^{\infty} \mathbb{P}(C_p = j)\mathbb{P}(T_n = i) \\ &= \sum_{i \in \mathbb{N}} \sum_{j=i}^{\infty} (1-p)^{k-1} p\mathbb{P}(T_n = i)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i\in\mathbb{N}} p\mathbb{P}(T_n = i) \sum_{j=i}^{\infty} (1-p)^{k-1} \\
&= \sum_{i\in\mathbb{N}} p\mathbb{P}(T_n = i) \frac{(1-p)^{i-1}}{p} \\
&= \sum_{i\in\mathbb{N}} \mathbb{P}(T_n = i)(1-p)^{i-1} \\
&= \mathbb{E}((1-p)^{T_n}).
\end{aligned}
$$

Como que hay un -1 que falta en la última igualdad. Supongo que el problema está en si la geométrica comienza en cero o en uno.

Ahora calculemos $\mathbb{P}(M_p = n)$.

$$
\begin{aligned}
\mathbb{P}(M_p = n) &= \mathbb{P}(M_p \geq n) - \mathbb{P}(M_p \geq n+1) \\
&= \mathbb{P}(T_n \leq Cp) - \mathbb{P}(T_{n+1} \leq n+1) \\
&= \mathbb{E}((1-p)^{T_n}) - \mathbb{E}((1-p)^{T_{n+1}}) \\
&= e^{-nf(1-p)} - e^{-(n+1)f(1-p)} \\
&= e^{-nf(1-p)}(1 - e^{-f(1-p)}) \\
&= (e^{-f(1-p)})^n(1 - e^{-f(1-p)}).
\end{aligned}
$$

De donde tenemos que M_n tiene la distribución geométrica de parámetro $(1 - e^{-f(1-p)})$.

2.1.4. **Inciso (iv).** *Tome el límite conforme $p \to 0$ para mostrar que la variable aleatoria*

$$M = -\min_{n\geq 0} S_n$$

tiene una distribución geométrica de parámetro $1 - e^{-f(1)}$. Interprete esto cuando $f(1) = 0$.
Cuando hablamos de C_p, sólo nos interesa su distribución.

Así que para ejemplificar cómo se comporta una sucesión de geométricas con sus parámetros tendien a 0, mientras conservemos a C_p con distribución geométrica, somos libres de elegir de que manera se comportan.

Definamos $C_p : [0,1] \to \mathbb{N}$ (donde en $[0,1]$ usamos la medida de Lebesgue) de la siguiente manera:

$$C_p(x) = \begin{cases} 1 & \text{if } x \in [0,p) \\ 2 & \text{if } x \in [p, p+(1-p)p) \\ \vdots & \\ n & \text{if } x \in \left[\sum_{i=1}^{n-1}(1-p)^{i-1}p, \sum_{i=1}^{n}(1-p)^{i-1}p\right) \\ \vdots & \end{cases}$$

Claramente, C_p definida de esta manera tiene distribución de una geométrica de parámetro p (Porque no le dejamos de otra).

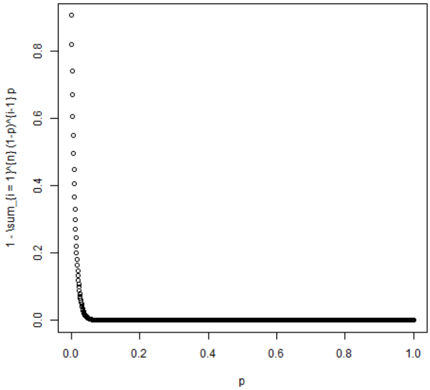
Ahora podemos ver con más claridad que

$$\mathbb{P}(C_p > n) = 1 - \sum_{i=1}^{n}(1-p)^{i-1}p$$

Analicemos la derivada del término $(1-p)^i p$ (i fija) con respecto a p. Su derivada es $(1-p)^i - pi(1-p)^{i-1}$. Evaluando en 0 tenemos
$(1-0)^i - 0i(1-0)^{i-1} = 1 > 0$. Es decir que cerca de 0, cada uno de los términos de este tipo, es estrictamente creciente.

Es decir, que si p se va acercando hacia 0, cada uno de estos términos, decrece.

A continuación una gráfica de cómo se comporta $1-\sum_{i=1}^{100}(1-p)^{i-1}p$ conforme variamos p. (Se incluye el script de R que se implementó para realizar esta gráfica, junto con algunas otras gráficas para distintos valores de n).



Con todo esto dicho, podemos garantizar que C_p diverge a ∞ en distribución. Pues para todo rango $[m, n]$ con $m < n, \in \mathbb{N}$ $\mathbb{P}(C_p \in [m, n]) \to 0$.

Entonces

$$- \min_{n \leq C_p} \to - \min_{n \leq \infty} S_n = \min_{n \geq 0} S_n = M$$

Ahora recordemos que g era una función continua y que $g > 0$. Por lo tanto $1/g$ es una función continua, con inverso derecho f en $(0, 1)$. Es decir
Por lo tanto

$$\begin{aligned} \mathbb{P}(M = n) &= \lim_{p \to 0} \mathbb{P}(M_p = n) \\ &= \lim_{p \to 0} e^{-nf(1-p)}(1 - e^{-f(1-p)}) \\ &= e^{-nf(1)}(1 - e^{-f(1)}) \qquad (2.1.4.1) \end{aligned}$$

Lo cual corresponde a la distribución de una geométrica de parámetro $1 - e^{-f(1)}$.

El caso donde $f(1^-) = 0$, implica que $\lim_{p \to 0}(1 - e^{-f(1-p)}) = 0$. El cual es complétamente análogo al caso anterior donde considerábamos $p \to 0$ en geométricas de parámetro p.

Donde, habíamos dejado claro que conforme el parámetro disminuía hacia 0, las distribuciones de las geométricas divergía a la de ∞ y que por lo tanto la probabilidad de que nuestras geométricas tomaran cierto valor n disminuía hacia 0 conforme el parámetro tendía a 0.

Incluso para este caso, la ecuación (2.1.4.1) es válida según nuestro análisis, pues

$$\begin{aligned}\mathbb{P}(M = n) &= \lim_{p\to 0} \mathbb{P}(M_p = n)\\ &= \lim_{p\to 0} e^{-nf(1-p)}(1 - e^{-f(1-p)})\\ &= e^{-n0}(1 - e^{-0})\\ &= 1(1-1)\\ &= 0.\end{aligned}$$

Y esto, para toda $n \in \mathbb{N}$, justo como nuestro análisis de las distribuciones geométricas nos dijo.Una posible conclusión es que el mínimo es menos infinito casi seguramente.

2.2. Problema 2.2

Problema 2.2.1. -

(i) Instale Octave en su computadora
(ii) Échele un ojo a la documentación
(iii) Ejecute el siguiente código linea por linea:

```
u = rand(600, 1);
x = [1/2];

for i=1:600
    x(i+1) = (2+i)/(2+i+1)*x(i) + (u(i)<x(i))/(2+i+1);
endfor

plot(x)
```

(iv) Lea las secciones sobre simple examples, ranges, random number generation y comparison operators y escriba su interpretación de lo que hace el código anterior. Nota: está relacionado con uno de los ejemplos del curso.
(v) Vuelva a correr el código varias veces y escriba sus impresiones sobre lo que está sucediendo.

En el código de arriba, se implementó una ejecución del proceso de las urnas de Poyla.

El vector `x` representa la proporción de las bolas "verdes" en la urna conforme el tiempo avanza. Que al tiempo inicial se tenga `x=[1/2]` quiere decir que al inicio había tantas bolas "verdes" como "rojas".

Para octave los booleanos pueden operarse numéricamente. `true` es equivalente a `1` y `false` es equivalente a `0`. Entonces, la parte que dice `+(u(i)<x(i))` significa que se está sumando uno o cero dependiendo de si la condición se satisface. Lo cual quiere decir que la constante de bolas que se agregan a la urna es 1.

El vector `u` representa el resultado de sacar una bola. Si `u(i)<x(i)` significa que en el turno `i`, se obtuvo una bola "verde".

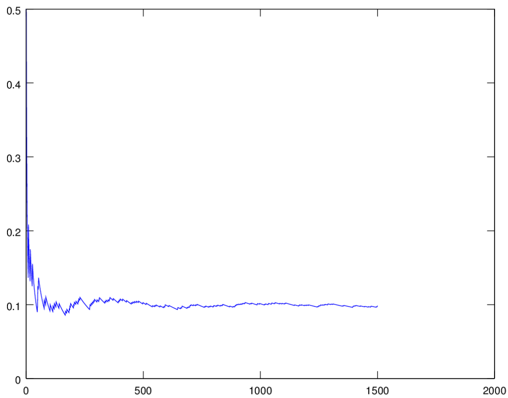
Notemos que en la implementación de arriba, `x(2) = 3/8` ó `x(2) = 5/8`. Lo que quiere decir que para el turno `2` existen al menos `8` bolas. Aún más, podemos asegurar que para el turno `2` existe un número par de bolas. Es decir que en el turno 1 el número de bolas era impar, pero la proporción de bolas "verdes" con respecto al total era `1/2`. Esto revela que existe un error de desfase en la implementación.

Imagino, que lo que se intentó hacer fue una urna con una bola "verde" y una bola "roja" inicialmente. En este caso, lo que debería decir en el ciclo `for` es lo siguiente:

```
for i = 1:600
    x(i+1) = (2+i-1)/(2+i) x(i) + (u(i)<x(i))/(2+i);
endfor
```

Con una chequeo rápido uno puede comprobar que en esta versión modificada de la implementación, `x(2) = 1/3` ó `x(2) = 2/3`. Como al inicio la proporción era `1/2`, esto significa que al inicio existían una bola "verde" y una "roja" exactamente.

A continuación una gráfica obtenida de ejcutar el código (cambiando `600` por `1500` para que sea más clara la estabilización de la proporción de bolas).



Gráfica de una ejecución de urnas de Poyla
1 bola verde inicial, 1 bola roja inicial y constante 1. 1500 iteraciones.

Se puede apreciar que conforme avanza el tiempo, la proporción de bolas "verdes" se estabiliza. Justo como se demostró en clase para martingalas positivas.

2.3. Problema 2.3

Problema 2.3.1. (Ejercicios sueltos sobre martingalas)

(i) [2.3.1] Sea $(X_n, n \geq 0)$ una sucesión $(\mathscr{F}_n)$-adaptada.
Pruebe que

$$\sum_{k=1}^{n} X_k - \mathbb{E}(X_k \mid \mathscr{F}_{k-1}), \quad n \geq 0$$

es una $(\mathscr{F}_n)$-martingala.

(ii) [2.3.2] Descomposición de Doob para submartingalas:
Sea $X = (X_n)_{n\in\mathbb{N}}$ una submartingala. Pruebe que X se puede descomponer de manera única como $X = M + A$ donde M es una martingala y A es un proceso previsible con $A_0 = 0$. Sugerencia: Asuma que ya tiene la descomposición y calcule esperanza condicional de X_{n+1} dada X_n.

(iii) [2.3.3] Sea $S_n = \xi_1 + \cdots + \xi_n$ donde las variables ξ son independientes y ξ_i tiene media cero y varianza finita σ_i^2. Pruebe que si $\sum_i \sigma_i^2 < \infty$ entonces S_n converge casi seguramente y en L_2 conforme $n \to \infty$. Construya un ejemplo de variables aleatorias ξ_i tales que la serie $\sum_i \xi_i$ sea casi seguramente absolutamente divergente y casi seguramente condicionalmente convergente (considere ejemplos simples!). Explique heurísticamente por qué cree que suceda esto.

(iv) [2.3.4] Sean X y Y dos martingalas (respecto de la misma filtración) y tales que $\mathbb{E}(X_i), \mathbb{E}(Y_i) < \infty$ para toda i. Pruebe la siguiente fórmula de integración por partes:

$$\mathbb{E}(X_n Y_n) - \mathbb{E}(X_0 Y_0) = \sum_{i=1}^{n} \mathbb{E}((X_i - X_{i-1})(Y_i - Y_{i-1})).$$

(v) [2.3.5] Desigualdad de Azema-Hoeffding, tomado de [?, E14.2, p.237]

(v.i) [2.3.5.1] Muestre que si Y es una variable aleatoria con valores en $[-c, c]$ y media cero entonces, para $\theta \in \mathbb{R}$

$$\mathbb{E}\left(e^{\theta Y}\right) \leq \cosh(\theta c) \leq \exp\left(\frac{1}{2}\theta^2 c^2\right).$$

(v.ii) [2.3.5.2] Pruebe que si M es una martingala nula en cero tal que para algunas constantes $(c_n, n \in \mathbb{N})$ se tiene que

$$|M_n - M_{n-1}| \leq c_n \quad \forall n$$

entonces, para $x > 0$

$$\mathbb{P}\left(\max_{k\leq n} M_k \geq x\right) \leq \exp\left(\frac{x^2}{2\sum_{k=1}^n c_k^2}\right).$$

Demostración:

2.3.1. **Inciso (i).** *Sea* $(X_n, n \geq 0)$ *una sucesión* $(\mathscr{F}_n)$*-adaptada. Pruebe que*

$$\sum_{k=1}^n X_k - \mathbb{E}(X_k \mid \mathscr{F}_{k-1}), \quad n \geq 0$$

es una $(\mathscr{F}_n)$*-martingala.*

Nombremos $M_n = \sum_{k=1}^n X_k - \mathbb{E}(X_k \mid \mathscr{F}_{k-1}), n \in \mathbb{N}$.

- Veamos que $(M_n)_{n\in\mathbb{N}}$ es $(\mathscr{F}_n)_{n\in\mathbb{N}}$-adaptada.

  $\sum_{k=1}^n X_k$ es $\mathscr{F}_n$-medible por definición. $\mathbb{E}(X_k|\mathscr{F}_{k-1})$ es $\mathscr{F}_{k-1}$-medible por definición de esperanza condicional.

  Por lo tanto $\sum_{k=1}^n \mathbb{E}(X_k|\mathscr{F}_{k-1})$ es $\mathscr{F}_{n-1}$-medible y por lo tanto $\mathscr{F}_n$-medible. Entonces

$$\begin{aligned}\sum_{k=1}^n X_k - \sum_{k=1}^n \mathbb{E}(X_k|\mathscr{F}_{k-1}) &= \sum_{k=1}^n X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1}) \\ &= M_n\end{aligned}$$

  Es suma finita de variables $\mathscr{F}_n$-medibles y por lo tanto $\mathscr{F}_n$-medible. Como queríamos ver.

- Veamos que $M_n \in L_1$.

  Como no fue mencionado. Supondremos que $X_n \in L_1$. Si logramos demostrar que $\mathbb{E}(X_k|F_{k-1}) \in L_1$ habremos terminado. Puesto que M_n sería suma finita de variables aleatorias en L_1 y por lo tanto pertenecería a L_1.

$$\mathbb{E}(|\mathbb{E}(X_k|F_{k-1})|) \leq \mathbb{E}(\mathbb{E}(|X_k|\,|F_{k-1}))$$

$$= \mathbb{E}(|X_k|).$$

Donde $\mathbb{E}(|X_k|)$ es finita por supocisiónCuidar la ortografía y utilización de la lengua: es mejor decir po hipótesis..

Por lo tanto $\mathbb{E}(|\mathbb{E}(X_k|F_{k-1})|)$ también lo es, como queríamos demostrar.

- Veamos por último que $\mathbb{E}(M_n|\mathscr{F}_{n-1}) = M_{n-1}$.

$$\begin{aligned}
\mathbb{E}(M_n|\mathscr{F}_{n-1}) &= \mathbb{E}\left(\sum_{k=1}^{n} X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1})|\mathscr{F}_{n-1}\right) \\
&= \sum_{k=1}^{n} \mathbb{E}\left(X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1})|\mathscr{F}_{n-1}\right) \\
&= \sum_{k=1}^{n} \mathbb{E}\left(X_k \middle| \mathscr{F}_{n-1}\right) - \mathbb{E}\left(\mathbb{E}(X_k|\mathscr{F}_{k-1})|\mathscr{F}_{n-1}\right) \\
&= \sum_{k=1}^{n} \mathbb{E}\left(X_k \middle| \mathscr{F}_{n-1}\right) - \mathbb{E}(X_k|\mathscr{F}_{k-1}) \\
&= \sum_{k=1}^{n-1} X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1}) + \mathbb{E}(X_n|\mathscr{F}_{n-1}) - \mathbb{E}(X_n|\mathscr{F}_{n-1}) \\
&= \sum_{k=1}^{n-1} X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1}) \\
&= M_{n-1}.
\end{aligned}$$

Y con esto terminamos la demostración de que $(M_n)_{n\in\mathbb{N}}$ es martingala con respecto a la filtración $(F_n)_{n\in\mathbb{N}}$.

2.3.2. **Inciso (ii).** *Descomposición de Doob para submartingalas: Sea $X = (X_n)_{n\in\mathbb{N}}$ una submartingala. Pruebe que X se puede descomponer de manera única como $X = M + A$ donde M es una martingala y A es un proceso previsible con $A_0 = 0$. Sugerencia: Asuma que ya tiene la descomposición y calcule esperanza condicional de X_{n+1} dada X_n.*

$(X_n)_{n\in\mathbb{N}}$ es $(\mathscr{F}_n)_{n\in\mathbb{N}}$-adaptada por definición de submartingala. Utilizando el inciso anterior tenemos que $M_n = \sum_{k=1}^{n} X_k - \mathbb{E}(\, X_k \mid \mathscr{F}_{k-1}\,)\,, n \in \mathbb{N}$ es martingala.

Si definimos a $A_n = X_n - M_n$ tendremos que $M_n + A_n = M_n + X_n - M_n = X_n$. Probemos ahora que $(A_n)_{n\in\mathbb{N}}$ es previsible.

$$
\begin{aligned}
A_n &= X_n - M_n \\
&= X_n - \sum_{k=1}^{n} X_k - \mathbb{E}(\, X_k \mid \mathscr{F}_{k-1}) \\
&= X_n - X_n + \mathbb{E}(X_n|\mathscr{F}_{n-1}) + \sum_{k=1}^{n-1} X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1}) \\
&= \mathbb{E}(X_n|\mathscr{F}_{n-1}) + \sum_{k=1}^{n-1} X_k - \mathbb{E}(X_k|\mathscr{F}_{k-1}).
\end{aligned}
$$

Donde por ser $(X_n)n \in \mathbb{N}$ submartingala tenemos que $\mathbb{E}(X_k|\mathscr{F}_{k-1})$ es F_{n-1} siempre que $0 \leq k \leq n-1$. Esto no se sigue de que X sea submartingala sino por el hecho de que la esperanza condicional sea medible respecto de la σ-álgebra condicionante. Si se sigue la sugerencia, se obtiene:

$$
A_n = \sum_{i=1}^{n} \mathbb{E}(\, X_k - X_{k-1} \mid \mathscr{F}_{k-1})
$$

que ayuda a establecer la existencia. Y por lo tanto A_n es suma finita de variables F_{n-1}-medibles y por lo tanto F_{n-1}-medible.

Hemos demostrado hasta ahora que una submartingala se puede descomponer como la suma de una martingala y un proceso previsible. Demostremos ahora la unicidad.

Sean entonces M'_n y A'_n una martingala y un proceso previsible respectivamente tal que $X_n = M'_n + A'_n$.

Entonces para toda $n \in \mathbb{N}$ tenemos que $M'_n + A'_n = M_n + A_n$. Por lo tanto

$$(M'_n + A'_n) - (M'_{n+1} + A'_{n+1}) = (M_n + A_n) - (M_{n+1} + A_{n+1})$$

Tomemos esperanzas condicionales sobre el lado derecho.

$$\begin{aligned}
&\mathbb{E}((M_n + A_n) - (M_{n+1} + A_{n+1})|\mathscr{F}_n) \\
&= \mathbb{E}((M_n + A_n) - (M_{n+1} + A_{n+1})|\mathscr{F}_n) \\
&= \mathbb{E}(M_n - M_{n+1}|\mathscr{F}_n) + \mathbb{E}(A_n - A_{n+1}|\mathscr{F}_n) \\
&= M_n - M_n + \mathbb{E}(A_n - A_{n+1}|\mathscr{F}_n) \qquad (2.3.2.1) \\
&= \mathbb{E}(A_n - A_{n+1}|\mathscr{F}_n) \\
&= \mathbb{E}(A_n|\mathscr{F}_n) - \mathbb{E}(A_{n+1}|\mathscr{F}_n) \\
&= A_n - A_{n+1} \qquad (2.3.2.2)
\end{aligned}$$

Donde (2.3.2.1) es gracias a que $(M_n)_{n\in\mathbb{N}}$ es martingala y (2.3.2.2) es gracias a que $(A_n)_{n\in\mathbb{N}}$ es un proceso previsible.

De manera análoga obtenemos que $\mathbb{E}((M'_n + A'_n) - (M'_{n+1} + A'_{n+1})|\mathscr{F}_n) = A'_n - A'_{n+1}$ y por lo tanto $A_n - A_{n+1} = A'_n - A'_{n+1}$.

Lo que esto último dice, es que $(A_n)_{n\in\mathbb{N}}$ y $(A'_n)_{n\in\mathbb{N}}$ crecen igual. Recordemos que por hipótesis $A'_0, A_0 = 0$.

Entonces

$$A_0 - A_1 = A'_0 - A'_1 \qquad \Rightarrow A_1 = A'_1$$

Inductivamente, si $A_n = A'_n$, entonces

$$A_n - A_{n+1} = A'_n - A'_{n+1} \qquad \Rightarrow A_{n+1} = A'_{n+1}$$

Y con esto queda demostrada la unicidad de $(A_n)_{n\in\mathbb{N}}$. Para la demostración de la unicidad de M_n, sólo recordemos que $M_n = X_n - A_n$ y que $M'_n = X_n - A'_n$, por ser A_n y A'_n iguales, entonces M_n y M'_n también lo son.

2.3.3. **Inciso (iii).** *Sea $S_n = \xi_1 + \cdots + \xi_n$ donde las variables ξ son independientes y ξ_i tiene media cero y varianza finita σ_i^2. Pruebe que si $\sum_i \sigma_i^2 < \infty$ entonces S_n converge casi seguramente y en L_2 conforme $n \to \infty$. Construya un ejemplo de variables aleatorias ξ_i tales que la serie $\sum_i \xi_i$ sea casi seguramente absolutamente divergente y casi seguramente condicionalmente convergente (considere ejemplos simples!). Explique heurísticamente por qué cree que suceda esto.*

Sea $\mathscr{F}_n = \sigma(\xi_0, \xi_1, \ldots, \xi_n)$.

Probemos ahora que $(S_n)_{n\in\mathbb{N}}$ es martingala con respecto de la filtración $(\mathscr{F}_n)_{n\in\mathbb{N}}$.

Por como se definió $\mathscr{F}_n$, S_n es $\mathscr{F}_n$-medible. Por otra parte, S_n es suma finita de variables en L_1 (pues las ξ_i's tienen esperanza finita).

Ahora

$$\begin{aligned}
\mathbb{E}(S_{n+1}|\mathscr{F}_n) &= \mathbb{E}(\xi_{n+1}|\mathscr{F}_n) + \mathbb{E}(S_n|\mathscr{F}_n) \\
&= \mathbb{E}(\xi_{n+1}|\mathscr{F}_n) + S_n \\
&\quad \text{Porque } S_n \text{ es } F_n\text{-medible} \\
&= \mathbb{E}(\xi_{n+1}) + S_n \\
&\quad \text{(Porque } \xi_{n+1} \text{ es independiente de} F_n) \\
&= 0 + S_n.
\end{aligned}$$

Y con esto hemos demostrado que S_n es martingala.

Ahora, como $\mathbb{E}(\xi_n) = 0$ para todo $n \in \mathbb{N}$, $\mathbb{E}(S_n) = 0$ para todo $n \in \mathbb{N}$.

$$\begin{aligned}
\mathrm{Var}(S_n) &= \mathbb{E}(S_n^2) - \mathbb{E}(S_n)^2 \\
&= \mathbb{E}(S_n^2) - 0 \\
&= \sum_{i\le n} \sigma_i^2 \\
&\quad \text{(Esto último gracias a que los } \xi_i\text{'s son independientes).}
\end{aligned}$$

Podrías utilizar la independencia antes al afirmar que la varianza de la suma es la suma de las varianzas.

Ahora $\|S_n\|_2 = \sqrt{\mathrm{Var}(S_n)} = \sqrt{\sum_{i\leq n}\sigma_i^2}$. Recordemos que por ser $\sqrt{\cdot}$ función continua, manda sucesiones convergentes en sucesiones convergentes. Como $\lim\limits_{n\to\infty}\sum\limits_{i<n}\sigma_i^2 < \infty$, entonces $\lim\limits_{n\to\infty}\|S_n\|_2 < \infty$.

En particular, $\sup\limits_n \|S_n\|_2 < \infty$.

El teorema de convergencia en L_p de Doob nos garantiza que S_n converge casi seguramente y en L_2. Como queríamos hacer ver.

Definamos ahora ξ_i con $\mathbb{P}(\xi_i = \pm\frac{1}{i+1}) = \frac{1}{2}$.
Dado que la serie $\sum_{n\geq 1}\frac{1}{n}$ es divergente. S_n es absolutamente divergente casi seguramente.

Ahora, $\mathbb{E}(\xi_i) = \frac{1}{2(i+1)} - \frac{1}{2(i+1)} = 0$ y $\mathrm{Var}(()\,\xi_i) = \frac{1}{(i+1)^2}$. Entonces, por lo dicho en la primera parte de este ejercicio, S_n converge casi seguramente y en L_2. Es decir que la serie S_n converge condicinalmente casi seguramente.
Heurísticamente, para que la serie diverja es necesario mantener el mismo signo por periodos "muy largos". La probabilidad de mantener el mismo signo por mucho tiemp decrece exponencialmente. Por eso es razonable que si elegimos en base a volados los signos de los términos, la serie terminará convergiendo.

2.3.4. **Inciso (iv).** *Sean X y Y dos martingalas (respecto de la misma filtración) y tales que $\mathbb{E}(X_i), \mathbb{E}(Y_i) < \infty$ para toda i. Pruebe la siguiente fórmula de integración por partes:*

$$\mathbb{E}X_nY_n - \mathbb{E}X_0Y_0 = \sum_{i=1}^{n} \mathbb{E}(X_i - X_{i-1})(Y_i - Y_{i-1}).$$

Descompongamos $\sum_{i\leq n} \mathbb{E}((X_i - X_{i-1})(Y_i - Y_{i-1}))$

$$\sum_{1\leq i\leq n} \mathbb{E}((X_i - X_{i-1})(Y_i - Y_{i-1}))$$

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i - X_iY_{i-1} - X_{i-1}Y_i + X_{i-1}Y_{i-1})$$

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i) - \mathbb{E}(X_iY_{i-1}) - \mathbb{E}(X_{i-1}Y_i) + \mathbb{E}(X_{i-1}Y_{i-1})$$

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i) - \mathbb{E}(\mathbb{E}(X_iY_{i-1})|\mathscr{F}_{i-1}) - \mathbb{E}(\mathbb{E}(X_{i-1}Y_i)|\mathscr{F}_{i-1}) + \mathbb{E}(X_{i-1}Y_{i-1})$$

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i) - \mathbb{E}(Y_{i-1}\mathbb{E}(X_i|\mathscr{F}_{i-1})) - \mathbb{E}(X_{i-1}\mathbb{E}(Y_i|\mathscr{F}_{i-1})) + \mathbb{E}(X_{i-1}Y_{i-1})$$

(Esto último gracias a que X_{i-1}, Y_{i-1} son F_{i-1} medibles)

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i) - \mathbb{E}(Y_{i-1}X_{i-1}) - \mathbb{E}(X_{i-1}Y_{i-1}) + \mathbb{E}(X_{i-1}Y_{i-1})$$

(Esto último gracias a que X, Y son martingalas)

$$= \sum_{1\leq i\leq n} \mathbb{E}(X_iY_i) - \mathbb{E}(Y_{i-1}X_{i-1})$$

En esta última suma podemos notar que se trata de una telescópica, y por lo tanto

$$\sum_{1\leq i\leq n} \mathbb{E}((X_i - X_{i-1})(Y_i - Y_{i-1})) = \mathbb{E}(X_nY_n) - \mathbb{E}(Y_0X_0).$$

Que es precisamente lo que buscábamos demostrar.

2.3.5. **Inciso (v).** Desigualdad de Azema-Hoeffding, tomado de [?, E14.2, p.237]

(v.i) [2.3.5.1] Muestre que si Y es una variable aleatoria con valores en $[-c, c]$ y media cero entonces, para $\theta \in \mathbb{R}$

$$\mathbb{E}\big(e^{\theta Y}\big) \leq \cosh(\theta c) \leq e^{(\frac{1}{2}\theta^2 c^2)}.$$

(v.ii) [2.3.5.2] Pruebe que si M es una martingala nula en cero tal que para algunas constantes $(c_n, n \in \mathbb{N})$ se tiene que

$$|M_n - M_{n-1}| \leq c_n \quad \forall n$$

entonces, para $x > 0$

$$\mathbb{P}\left(\max_{k \leq n} M_k \geq x\right) \leq \exp\left(\frac{x^2}{2\sum_{k=1}^n c_k^2}\right).$$

2.3.5.1. *Subinciso (v.i). Muestre que si Y es una variable aleatoria con valores en $[-c, c]$ y media cero entonces, para $\theta \in \mathbb{R}$*

$$\mathbb{E}\big(e^{\theta Y}\big) \leq \cosh(\theta c) \leq e^{(\frac{1}{2}\theta^2 c^2)}.$$

Notemos primero que la función $e^{\theta y}$ es convexa. Esto es fácil viendo que su segunda derivada es $\theta^2 e^{\theta y} \geq 0$.

Ahora

$$\begin{aligned} Y &= \frac{2cY}{2c} \\ &= \frac{cY + cY}{2c} \\ &= \frac{c^2 + cY - c^2 + cY}{2c} \\ &= \frac{c(c+Y) - c(c-Y)}{2c} \\ &= c\frac{(c+Y)}{2c} - c\frac{(c-Y)}{2c} \end{aligned}$$

De donde $\theta Y = c\theta\frac{(c+Y)}{2c} - c\theta\frac{(c-Y)}{2c}$.

Entonces

$$e^{\theta Y} = e^{\left(c\theta\frac{(c+Y)}{2c} - c\theta\frac{(c-Y)}{2c}\right)}$$

Ya que $|Y| \leq c$ tenemos que $0 \leq c+Y \leq 2c$ y $0 \leq c-Y \leq 2c$. Y por lo tanto $0 \leq \frac{c+Y}{2c} \leq 1$ y $0 \leq \frac{c-Y}{2c} \leq 1$. Además $\frac{c+Y}{2c} + \frac{c-Y}{2c} = \frac{c+Y+c-Y}{2c} = 1$. Entonces podemos utilizar la convexidad de $e^{\theta y}$ sobre los puntos $-c$ y c para obtener lo siguiente

$$e^{\theta Y} \leq e^{c\theta}\frac{(c+Y)}{2c} + e^{-c\theta}\frac{(c-Y)}{2c}.$$

Tomando esperanza en ambos lados obtenemos y recordando que $\cosh x = \frac{e^x+e^{-x}}{2}$

$$\begin{aligned}
\mathbb{E}(e^{\theta Y}) &\leq \mathbb{E}(e^{c\theta}\frac{(c+Y)}{2c} + e^{-c\theta}\frac{(c-Y)}{2c}) \\
&= e^{c\theta}\frac{(c+\mathbb{E}(Y))}{2c} + e^{-c\theta}\frac{(c-\mathbb{E}(Y))}{2c} \\
&= e^{c\theta}\frac{c}{2c} + e^{-c\theta}\frac{c}{2c} \\
&= e^{c\theta}\frac{1}{2} + e^{-c\theta}\frac{1}{2} \\
&= \frac{e^{c\theta} + e^{-c\theta}}{2} \\
&= \cosh(c\theta).
\end{aligned}$$

Con lo que terminamos de demostrar la primera desigualdad.

Para la segunda desigualdad recordemos que la expansión en serie de Taylor de e^x es

$$e^x = \sum_{n\in\mathbb{N}} \frac{x^n}{n!}. \tag{2.3.5.1}$$

De donde

$$e^{-x} = \sum_{n\in\mathbb{N}} \frac{-x^n}{n!}$$

$$= \sum_{n\in\mathbb{N}} (-1)^n \frac{x^n}{n!}.$$

Sumando estos dos y dividendo entre 2

$$\begin{aligned}
\cosh(x) &= \frac{e^x + e^{-x}}{2} \\
&= \sum_{n\in\mathbb{N}} \frac{\frac{x^n}{n!} + (-1)^n \frac{x^n}{n!}}{2} \\
&= \sum_{n\in\mathbb{N}} \frac{x^{2n}}{(2n)!}. \qquad (2.3.5.2)
\end{aligned}$$

Ahora notemos lo siguiente.

$$\begin{aligned}
2^0(0!) &= 1 \leq 1 = (2\cdot 0)! \\
2^1(1!) &= 2 \leq 2 = (2\cdot 1)! \\
2^2(2!) &= 8 \leq 24 = (2\cdot 2)! \\
2^3(3!) &= 48 \leq 720 = (2\cdot 3)!
\end{aligned}$$

Apliquemos inducción sobre esto. Supongamos que para cierto n ocurre que $2^n(n!) \leq (2n)!$. Puedes hacer una prueba directa al notar que estás multiplicando por dos cada factor del factorial y que eso te da los factores pares del factorial del doble. Como $n \geq 0$, entonces $2 \leq 2(n+1)$ y $n+1 \leq 2n+1$, por lo tanto $2(n+1) \leq (2n+1)(2n+2)$. De donde

$$2^{n+1}(n+1)! = 2(n+1)2^n(n!) \leq (2n)!(2n+1)(2n+2) = (2(n+1))!.$$

Con esto, podemos acotar (2.3.5.2) de la siguiente manera

$$\begin{aligned}
\cosh(x) &= \sum_{n\in\mathbb{N}} \frac{x^{2n}}{(2n)!} \\
&\leq \sum_{n\in\mathbb{N}} \frac{x^{2n}}{2^n(n)!} \\
&= \sum_{n\in\mathbb{N}} \frac{(x^2)^n}{2^n(n)!}
\end{aligned}$$

$$= \sum_{n\in\mathbb{N}} \frac{(\frac{x^2}{2})^n}{(n)!}$$
$$= e^{\frac{x^2}{2}}$$

Y utilizando esto tenemos que $\cosh(c\theta) \leq e^{\frac{c^2\theta^2}{2}}$.

2.3.5.2. *Subinciso (v.ii). Pruebe que si M es una martingala nula en cero tal que para algunas constantes $(c_n, n \in \mathbb{N})$ se tiene que*

$$|M_n - M_{n-1}| \leq c_n \quad \forall n \tag{2.3.5.3}$$

entonces, para $x > 0$

$$\mathbb{P}\left(\max_{k\leq n} M_k \geq x\right) \leq \exp\left(\frac{x^2}{2\sum_{k=1}^n c_k^2}\right). \tag{2.3.5.4}$$

Lo primero que sugiere (2.3.5.4) es utilizar la desigualdad maximal de Doob que dice que si $N = (N_n)_{n\in\mathbb{N}}$ es una $(\mathscr{G}_n)_{n\in\mathbb{N}}$-submartingala, entonces para toda $\lambda > 0$ se tiene

$$\lambda\mathbb{P}\left(\overline{N}_n^+ > \lambda\right) \leq \mathbb{E}(N_n^+) \tag{2.3.5.5}$$

Donde $\overline{N}_n^+ = \max\limits_{1\leq i\leq n} N_n^+$.

Ahora construiremos una submartingala conveniente. En clase se vio que por ser M es una martingala, al aplicarle la función convexa $e^{\theta y}$, el proceso resultante $(e^{\theta M_n})_{n\in\mathbb{N}}$ es una submartingala positiva con respecto de la filtración $(\mathscr{F}_n)_{n\in\mathbb{N}}$.

Aplicando (2.3.5.5) con $\lambda = e^{\theta x}$ obtenemos

$$e^{\theta x}\mathbb{P}\left(\max_{1\leq k\leq n} e^{\theta M_k} > x\right) \leq \mathbb{E}(e^{\theta M_n})$$

Y por lo tanto

$$\mathbb{P}\left(\max_{1\leq k\leq n} e^{\theta M_k} > e^{\theta x}\right) \leq e^{-\theta x}\mathbb{E}(e^{\theta M_n})$$

Ahora notemos que si $\theta > 0$, entonces $e^{\theta y}$ es inyectiva y por lo tanto los eventos $\{\max_{1\leq i\leq n} e^{\theta M_k} > e^{\theta x}\}$ y $\{\max_{1\leq i\leq n} M_k > x\}$ son iguales. Supongamos entonces que $\theta > 0$ y por lo tanto

$$\mathbb{P}\left(\max_{1\leq k\leq n} M_k > x\right) \leq e^{-\theta x}\mathbb{E}(e^{\theta M_n})$$

Ahora, gracias a (2.3.5.3) tenemos que

$$\begin{aligned}|M_n| &= \left| \sum_{1\leq k\leq n} M_k - M_{k-1} \right| \\ &\leq \sum_{1\leq k\leq n} \left| M_k - M_{k-1} \right| \\ &\leq \sum_{1\leq k\leq n} c_k.\end{aligned}$$

Entonces, gracias a lo demostrado en [2.3.5.1] y usando que $\sum_{1\leq k\leq n} c_k^2 \leq (\sum_{1\leq k\leq n} c_k)^2$ tenemos

$$\begin{aligned}e^{\theta M_n} &\leq e^{\left(\frac{1}{2}\theta^2\left(\sum_{1\leq k\leq n} c_k\right)^2\right)} \\ &\leq e^{\left(\frac{1}{2}\theta^2\left(\sum_{1\leq k\leq n} c_k^2\right)\right)}.\end{aligned}$$

Definamos entonces

$$\theta = \frac{x}{\sum_{1\leq k\leq n} c_k^2}.$$

Definido así, $\theta > 0$ y por lo tanto, todo lo mencionado anteriormente es válido. Conectemos todo lo dicho hasta ahora.

$$\begin{aligned}\mathbb{P}\left(\max_{1\leq k\leq n} M_k > x\right) &\leq e^{-\theta x}\mathbb{E}(e^{\theta M_n}) \\ &\leq e^{-\theta x} e^{\left(\frac{1}{2}\theta^2\left(\sum_{1\leq k\leq n} c_k^2\right)\right)} \\ &= e^{-\left(\frac{x}{\sum_{1\leq k\leq n} c_k^2}\right)x} e^{\left(\frac{1}{2}\left(\frac{x}{\sum_{1\leq k\leq n} c_k^2}\right)^2\left(\sum_{1\leq k\leq n} c_k^2\right)\right)} \\ &= e^{-\left(\frac{x^2}{2\sum_{1\leq k\leq n} c_k^2}\right)}.\end{aligned}$$

Que es precisamente lo que búscabamos demostrar.

2.4. Problema 2.4

Problema 2.4.1. Sea $S_n = \sum_{i=1}^n X_i$ donde $X_1, X_2, \ldots$ son iid. Sea

$$\phi(\lambda) = \mathbb{E}\left(e^{\lambda S_n}\right) \in (0, \infty].$$

(i) [2.4.1] Pruebe que si existen $\lambda_1 < 0 < \lambda_2$ tales que $\phi(\lambda_i) < \infty$ entonces $\phi(\lambda) < \infty$ para toda $\lambda \in [\lambda_1, \lambda_2]$. Sugerencia: escriba $\lambda = a\lambda_1 + (1 - a)\lambda_2$ para algún $a \in [0, 1]$ y aplique la desigualdad de Hölder. A partir de ahora se asume la premisa de este inciso.

(ii) [2.4.2] Pruebe que $\mathbb{E}\left(|S_n|^k\right) < \infty$ para toda $k \geq 0$.

(iii) [2.4.3] Sea $M_t^\lambda = e^{\lambda S_t}/\phi(\lambda)$. Argumente que si M^n es el proceso dado por

$$M_t^n = \left.\frac{\partial^n}{\partial \lambda^n}\right|_{\lambda=0} M_t^\lambda,$$

entonces M^n es una martingala para toda n.

(iv) [??] Calcule las primeras 4 martingalas resultantes si $\mathbb{P}(X_i = \pm 1) = 1/2$. Utilícelas para calcular el valor de $\mathbb{E}\left(T^2\right)$ donde

$$T = \min\{n \geq 0 : S_n \in \{-a, b\}\}$$

y $a, b > 0$.

Categorías: Caminatas aleatorias, muestreo opcional, ejemplos de martingalas.

Demostración:

2.4.1. **Inciso (i).** *Pruebe que si existen $\lambda_1 < 0 < \lambda_2$ tales que $\phi(\lambda_i) < \infty$ entonces $\phi(\lambda) < \infty$ para toda $\lambda \in [\lambda_1, \lambda_2]$. Sugerencia: escriba $\lambda = a\lambda_1 + (1 - a)\lambda_2$ para algún $a \in [0, 1]$ y aplique la desigualdad de Hölder. A partir de ahora se asume la premisa de este inciso.*

Sea $\lambda = \lambda_1(1 - t) + \lambda_2 t$ con $t \in [0, 1]$. Por lo tanto

$$\begin{aligned}\phi(\lambda) &= \mathbb{E}(e^{\lambda S_n}) \\ &= \mathbb{E}\left(e^{(\lambda_1(1-t)+\lambda_2 t)S_n}\right)\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left(e^{(\lambda_1(1-t)S_n+\lambda_2 t S_n)}\right)\\
&= \mathbb{E}\left(e^{\lambda_1(1-t)S_n}e^{\lambda_2 t S_n)}\right)\\
&= \mathbb{E}\left(e^{(\lambda_1 S_n)^{1-t}}e^{(\lambda_2 S_n)^{t}}\right)
\end{aligned}
$$

Si definimos $p = 1/(1-t)$ y $q = 1/t$, tenemos que $e^{(\lambda_1 S_n)^{1-t}} \in L_p$ pues $\left(e^{(\lambda_1 S_n)^{1-t}}\right)^p = e^{(\lambda_1 S_n)}$ quien pertenece a L_1 por hipótesis. Análgomanente tenemos que $e^{(\lambda_2 S_n)^t} \in L_q$.

Ahora, recordando que $|e^x| = e^x$ y utilizando la desigualdad de Hölder (2.1.2.1)

$$
\begin{aligned}
\phi(\lambda) &= \mathbb{E}\left(e^{(\lambda_1 S_n)^{1-t}}e^{(\lambda_2 S_n)^{t}}\right)\\
&\leq \mathbb{E}\left(e^{(\lambda_1 S_n)}\right)^{1-t}\mathbb{E}\left(e^{(\lambda_2 S_n)}\right)^{t}
\end{aligned}
$$

Donde cada uno de los factores es finito por hipótesis y por lo tanto $\phi(\lambda) < \infty$. Como queríamos demostrar.

2.4.2. **Inciso (ii).** *Pruebe que* $\mathbb{E}\big(|S_n|^{\,k}\big) < \infty$ *para toda* $k \geq 0$.

Ahora que sabemos que si $\lambda \in [\lambda_1, \lambda_2]$ entonces $\phi(\lambda) < \infty$. Elijamos $\lambda = \min(-\lambda_1, \lambda_2)$ y de esta manera, tenemos que $\phi(-\lambda), \phi(\lambda) < \infty$.

Entonces

$$\begin{aligned}
\mathbb{E}(e^{|\lambda S_n|}) &= \mathbb{E}(e^{\lambda S_n}\mathbb{1}_{\lambda S_n \geq 0} + e^{-\lambda S_n}\mathbb{1}_{\lambda S_n < 0}) \\
&= \mathbb{E}(e^{\lambda S_n}\mathbb{1}_{\lambda S_n \geq 0}) + \mathbb{E}(e^{-\lambda S_n}\mathbb{1}_{\lambda S_n < 0}) \\
&\leq \mathbb{E}(e^{\lambda S_n}) + \mathbb{E}(e^{-\lambda S_n}) \\
&= \phi(\lambda) + \phi(-\lambda) \\
&< \infty.
\end{aligned}$$

Este hecho lo necesitaremos más adelante.

Ahora utilizaremos la expansión en serie de Taylor de e^x (2.3.5.1) para escribir

$$\begin{aligned}
e^{|\lambda S_n|} &= \sum_{k \in \mathbb{N}} \frac{|\lambda S_n|^k}{k!} \\
&= \sum_{k \in \mathbb{N}} \frac{\lambda^k |S_n|^k}{k!}
\end{aligned}$$

Tomando esperanzas obtenemos

$$\begin{aligned}
\mathbb{E}(e^{|\lambda S_n|}) &= \mathbb{E}\left(\sum_{k \in \mathbb{N}} \frac{\lambda^k |S_n|^k}{k!}\right) \\
&= \sum_{k \in \mathbb{N}} \frac{\lambda^k \mathbb{E}(|S_n|)^k}{k!} \\
&\quad \text{(Aquí se utilizó el teorema de convergencia} \\
&\quad \text{monótona para poder meter el operador esperanza)} \\
&\leq \mathbb{E}(\phi(\lambda) + \phi(-\lambda)) \\
&< \infty.
\end{aligned}$$

Esto último implica que cada uno de los sumandos de $\sum_{i\in\mathbb{N}} \frac{\lambda^i \mathbb{E}(|S_n|)^i}{i!}$ es necesariamente finito y por lo tanto $\mathbb{E}(|S_n|^k) < \infty$ para todo $n \in \mathbb{N}$. Como se quería demostrar.

2.4.3. **Inciso (iii).** *Sea* $M_t^\lambda = e^{\lambda S_t}/\phi(\lambda)$*. Argumente que si* M^n *es el proceso dado por*

$$M_t^n = \left.\frac{\partial^n}{\partial \lambda^n}\right|_{\lambda=0} M_t^\lambda,$$

entonces M^n *es una martingala para toda* n*.*

Análogo a la parte de la demostración de que M_t^λ es martingala en [6.4.1], salvo que para este caso nunca se consiguió dar una función integrable que domine al valor absoluto de la derivada de orden n de M_t^λ. For the records, falta el detalle. La idea es notar que $x^n e^{-\lambda x} \leq C_n e^{-\lambda' x}$ para alguna constante C_n si $\lambda < \lambda'$.

Fuera de ese detalle, el resto es análogo.

•___•••_•___•_•_•__•____•__••_••••_••••••__••_•_•___••_••__••••_•___•_••••_••••••__••_••

Part 3. Tarea 3

3.1. Problema 3.1

Problema 3.1.1. Sea M una $(\mathscr{F}_n)$-martingala. Pruebe que si T es un tiempo de paro finito entonces $\mathbb{E}(M_T) = \mathbb{E}(M_0)$ bajo cada una de las siguientes condiciones:

(i) [3.1.1] M es acotada.
(ii) [3.1.2] T es integrable y la sucesión $(M_n - M_{n-1})$ es acotada.
(iii) [3.1.3] $(M_{n\wedge T})$ es uniformemente integrable.

Categorías: Muestreo opcional.

Demostración:

3.1.1. Inciso (i). *M es acotada.*

Que M sea acotada significa que existen $a \in \mathbb{R}$ tal que $|M_n| < r$ para todo $n \in \mathbb{N}$. En particular tenemos que $|M_{T\wedge n}| < r$.

Por otro lado, que T sea finito implica que $T \wedge n \underset{c.s.}{\longrightarrow} T$ conforme $n \to \infty$. (Para todo $\omega \in \Omega$, $T(\omega) = n_0 < \infty$ y por lo tanto $T \wedge n(\omega) = T(\omega) = n_0$ para todo $n \geq n_0$). De donde afirmamos que $M_{T\wedge n} \underset{c.s.}{\longrightarrow} M_T$.

Ahora, $T \wedge n$ es tiepo de paro acotado, y por Teorema del Muestreo Opcional de Doob tenemos que $\mathbb{E}(M_{T\wedge n}) = \mathbb{E}(M_0)$.

Tenemos todas las hipótesis para aplicar el Teorema de Convergencia Acotada y por lo tanto:

$$\begin{aligned}
\mathbb{E}(M_T) &= \mathbb{E}(\lim_{n\to\infty} M_{T\wedge n}) \\
&= \lim_{n\to\infty} \mathbb{E}(M_{T\wedge n}) \\
&= \mathbb{E}(M_0).
\end{aligned}$$

Con lo que termina la demostración.

3.1.2. **Inciso (ii).** *T es integrable y la sucesión $(M_n - M_{n-1})$ es acotada.* Que la sucesión $(M_n - M_{n-1})$ sea acotada, significa que existe $r \in R$ tal que $|M_n - M_{n-1}| < r$ para todo $n \geq 1$.

Que T sea integrable significa que $\mathbb{E}(T) < \infty$. Recordemos que el que T sea finito implica que $T \wedge n \underset{c.s.}{\longrightarrow} T$ (la justificación está en 3.1.1).

Otra vez, $T \wedge n$ es tiepo de paro acotado, y por Teorema del Muestreo Opcional de Doob tenemos que $\mathbb{E}(M_{T\wedge n}) = \mathbb{E}(M_0)$. En este caso nos interesa escribirlo de la siguiente manera, $\mathbb{E}(M_{T\wedge n} - M_0) = 0$.

Ahora,

$$\begin{aligned}
|M_{T\wedge n} - M_0| &= \left| \sum_{1\leq i\leq (T\wedge n)} (M_i - M_{i-1}) \right| \\
&\leq \sum_{1\leq i\leq (T\wedge n)} |(M_i - M_{i-1})| \\
&\leq (T \wedge n) r \\
&\leq Tr
\end{aligned}$$

De donde

$$\begin{aligned}
\mathbb{E}(|M_{T\wedge n} - M_0|) &\leq \mathbb{E}(Tr) \\
&= r\mathbb{E}(T) \\
&< \infty.
\end{aligned}$$

Con esto tenemos que la sucesión $(M_{T\wedge n} - M_0)_{n\in\mathbb{N}}$ es dominada por Tr y que cada elemento de la sucesión es integrable. Entonces tenemos todas las hipótesis del Teorema de convergencia dominada. Y por lo tanto

$$\begin{aligned}
\mathbb{E}(M_T - M_0) &= \mathbb{E}(\lim_{n\to\infty} M_{T\wedge n} - M_0) \\
&= \lim_{n\to\infty} \mathbb{E}(M_{T\wedge n} - M_0) \\
&= \lim_{n\to\infty} 0 \\
&= 0.
\end{aligned}$$

Lo cual implica $\mathbb{E}(M_T) = \mathbb{E}(M_0)$, como queríamos demostrar.

3.1.3. **Inciso (iii).** $(M_{n\wedge T})$ *es uniformemente integrable.*
Sabemos que $M_{T\wedge n} \xrightarrow[c.s.]{} M_T$ gracias a que T es finito (la justificación está en 3.1.1).

Existe un teorema que garantiza que si $M_{T\wedge n} \xrightarrow[c.s.]{} M_T$ y $(M_{n\wedge T})$ es uniformemente integrable, entonces $M_T \in L_1$ y $M_{T\wedge n} \xrightarrow[L_1]{} M_T$.

Es decir que $\mathbb{E}(M_{T\wedge n}) \longrightarrow \mathbb{E}(M_T)$, y de nueva cuenta, por Teorema del Muestreo Opcional de Doob, en vista que $T \wedge n$ es tiempo de paro acotado, tenemos que $\mathbb{E}(M_{T\wedge n}) = \mathbb{E}(M_0)$ y por lo tanto

$$\begin{aligned}\mathbb{E}(M_T) &= \lim_{n\to\infty} \mathbb{E}(M_{T\wedge n})\\ &= \lim_{n\to\infty} \mathbb{E}(M_0)\\ &= \mathbb{E}(N_0).\end{aligned}$$

Que es lo que buscábamos demostrar.

3.2. Problema 3.2

Problema 3.2.1. Sea M una $(\mathscr{F}_n)$-martingala con saltos acotados. Sean

$$\begin{aligned} C &= \{\limsup M_n = \liminf M_n \in \mathbb{R}\} \\ y& \\ D &= \{\limsup M_n = -\infty \text{ y } \limsup M_n = \infty\}. \end{aligned}$$

Pruebe que $\mathbb{P}(C \cup D) = 1$. Deduzca que las caminatas aleatorias centradas con saltos acotados oscilan. Sugerencia: Para cada $K > 0$ defina

$$T = \min\{n \geq 0 : |M_n| \geq K\}$$

y aplique el teorema de convergencia de martingalas a M^T.

Sea M una caminata aleatoria no trivial con saltos integrables en $-1, 0, 1, \ldots$ y media cero.

Pruebe que $\mathbb{P}(M \text{ converge en } \mathbb{N}) = 0$ y concluya que $\liminf M_n = -\infty$ casi seguramente. (Este resultado permitirá dar una prueba adicional de que un Galton-Watson crítico se extingue). Sugerencia: proceda como en el párrafo anterior y pruebe la integrabilidad uniforme de $(M_{T\wedge n})_{n\in\mathbb{N}}$.
Categorías: Teoremas de convergencia de martingalas

Definamos

$$T_k = \min\{n \geq 0 : |M_n| \geq k\}.$$

Que significa "la primera vez que la martingala se aleja de cero al menos k unidades". Veamos que se trata de un tiempo de paro.

$$\begin{aligned} \{T_k = n\} &= \left(\bigcup_{i \leq n-1} \{|M_i| < k\} \right) \cap (\{|M_n| \geq k\}) \\ &= \left(\bigcap_{i \leq n-1} \{M_i < k\} \cap \{M_i > -k\} \right) \cap (\{M_n \geq k\} \cup \{M_n \leq -k\}). \end{aligned}$$

Donde, $\{M_i < k\}, \{M_i > -k\} \in \mathscr{F}_i \subset \mathscr{F}_n$ y $\{M_n \geq k\}, \{M_n \leq -k\} \in \mathscr{F}_n$. Y por lo tanto $\{T_k = n\} \in \mathscr{F}_n$, de donde concluimos que T_K es efectivamente tiempo de paro.

Sea C una cota para los saltos de M. Por definición de T_k, es claro que para el tiempo T_k (cuando T_k es finito), la martingala se aleja en al menos k unidades de 0, pero forzosamente menos que $k+C$, pues su último salto no puede superar C. Es decir $|M_{T_k}| < k+C$. Pero esto únicamente será cierto en los casos en que T_k resulte ser finito. Tomemos en cuenta entonces al tiempo $T_k \wedge n$, qué satisface ser finito y que $T_k \wedge n \leq T_k$. Para este tiempo, nuestro razonamiento anterior es válido y entonces tenemos $|M_{T_k \wedge n}| < k + C$.

Entonces $\mathbb{E}(|M_{T_k \wedge n}|) < \mathbb{E}(k + C) = k + C$ para toda $n \in N$. Sabemos de [1.4] que $(M_{T_k \wedge n})_{n \in \mathbb{N}}$ es martingala. Por el Teorema de Convergencia de Martingalas, tenemos que $M_{T_k \wedge n} \longrightarrow M_{T_k}$ y es finita c.s.

Estudiemos ahora los casos de convergencia de la martingala. Definimos

$$A_1 = \{\limsup M_n < \infty\}$$
$$A_2 = \{\liminf M_n > -\infty\}$$

Estos son los casos donde la martingala no se dispara hacia arriba indefinidamente o hacia abajo indefinidamente (respectivamente).

Los casos que queremos estudiar son $B_1 = A_1 \cap A_2$, $B_2 = A_1^c \cap A_2$, $B_3 = A_1 \cap A_2^c$ y $B_4 = A_1^c \cap A_2^c$. Es decir que B_1 es el caso donde la caminata se queda acotada. B_2 Es el caso donde la caminata alcanza todos los niveles arriba de cero. B_3 el caso donde la caminata alcanza todos los niveles abajo de cero. Y B_4 es el caso en el que la caminata oscila (alcanza todos los niveles).

Del análisis anterior, podemos ver que B_2 y B_3 tienen probabilidad 0, pues en estos casos, el límite de M_n es infinito y nuesto análisis mostró que la variable aleatoria a la que se converge es finita casi seguramente.

De esto tenemos que $\mathbb{P}(D \cup C) = 1$, pues $C \subset B_1$ y $D = B_4$ en la manera que definimos a B_1 y B_4.

3.3. Problema 3.3

Problema 3.3.1. Sean $X_1, X_2, \ldots$ variables aleatorias intercambiables:

$$(X_1, \ldots, X_n) \stackrel{d}{=} (X_{\pi_1}, \ldots, X_{\pi_n})$$

para cada permutación σ de $\{1, \ldots, n\}$.

(1) Para $\mathscr{G}, \mathscr{H}$ subσ-álgebras de $\mathscr{F}$ definimos a $\mathscr{G} \vee \mathscr{H} = \sigma(\mathscr{G} \cup \mathscr{H})$. Sea
$$\mathscr{G}^n = \sigma(f(X_1, \ldots, X_n) : f : \mathbb{R}^n \to \mathbb{R} \text{ es simétrica}) \vee \sigma(X_{n+1}, X_{n+2}, \ldots)\,.$$
Pruebe que $\mathscr{G}^n, n \geq 1$ es una filtración al revés. Sea $\mathscr{G}$ su intersección.

(2) Para cada $A \in \mathscr{B}_{\mathbb{R}}$, defina a

$$\Xi_n(A) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{X_i \in A}.$$

Pruebe que

$$\mathbb{P}(X_1 \in A \,|\mathscr{G}^n) = \Xi_n(A)\,.$$

¿Por qué puede definir a $\Xi(A) = \lim_{n\to\infty} \Xi_n(A)$?

(3) Al considerar a la martingala

$$\frac{1}{n\,(n-1)} \sum_{1 \leq i < j \leq n} \mathbf{1}_{X_i \in A} \mathbf{1}_{X_j \in A},$$

pruebe que $\mathbb{P}(X_1 \in A, X_2 \in A \,|\mathscr{G}) = \mathbb{P}(X_1 \in A \,|\mathscr{G})\, \mathbb{P}(X_2 \in A \,|\mathscr{G})$. Extienda la afirmación de independencia condicional anterior a $X_1, \ldots, X_n$.

Cagegorías: Teorema de convergencia de martingalas, variables intercambiables, teorema de de Finetti. For the records: falta.

3.4. Problema 3.4

Problema 3.4.1.

(i) [3.4.1] Ejecute y explique la función del siguiente código en Octave. Comente qué teoremas del curso (y del curso de probabilidad) son importantes para interpretar la figura.

```
tic;
n = 1000;
m = 10000;
u = rand(n, m);
r = 2;
v = 3;
c = 1;
x = ones(1,m)*(r/(r+v));

for i=1:n
        x(i+1,:)=(r+v+(i-1)*c)/(r+v+i*c).*x(i,:)+(u(i,:)<x(i,:))./(r+v+i*c);
endfor

y = sort(x(n+1,:));
plot(y,(1:m)./m,y,betacdf(y,r/c,v/c))
toc
```

(Conmigo se negó a brincar de linea. Tuve que hacerlo diminuto para que apareciera el código completo.)

(ii) [3.4.2] Ejecute y explique la función del siguiente código en Octave. Incluya una gráfica en la que la longitud de la variable k sea mayor a 1000. (Puede modificar el programa...) En la gráfica observara un esbozo de la trayectoria de un proceso de ramificación continuo (en una escala distinta...).

```
k   = [10];
aux = k(length(k));

while (aux>0 && length(k)<1000)
    k   = [k;2*binornd(aux, .5)];
    aux = k(length(k));
endwhile

plot(k)
```

Demostración:

3.4.1. **Inciso (i).** *Ejecute y explique la función del siguiente código en Octave. Comente qué teoremas del curso (y del curso de probabilidad) son importantes para interpretar la figura.*

```
tic;
n = 1000;
m = 10000;
u = rand(n, m);
```

```
r = 2;
v = 3;
c = 1;
x = ones(1,m)*(r/(r+v));

for i=1:n
        x(i+1,:)=(r+v+(i-1)*c)/(r+v+i*c).*x(i,:)+(u(i,:)<x(i,:))./(r+v+i*c);
endfor

y = sort(x(n+1,:));
plot(y,(1:m)./m,y,betacdf(y,r/c,v/c))
toc
```

En el código se simulan `10000` procesos de urnas de Póyla con `1000` iteraciones cada uno.

`n` es el número de juegos por proceso.

`m` es el número de simulaciones.

`u` es una matriz de probabilidades. En este caso, la entrada `(i,j)`, es el "resultado" del turno `j` en el proceso `i`.

`r` y `v` son el número de pelotas rojas y verdes (respectivamente) con las que comienza cada proceso. Esto se deduce de la manera en que se obtienen las proporciones iniciales `r/(r+v)`.

`c` es la constante de bolas que se agregan en cada paso del proceso. Esto se deduce en la manera que se obtiene la `i`-ésima proporción (se divide entre `r+v+i*c` es decir, el número de pelotas en el `i`-ésimo paso es el número de bolas iniciales más `i` veces `c`).

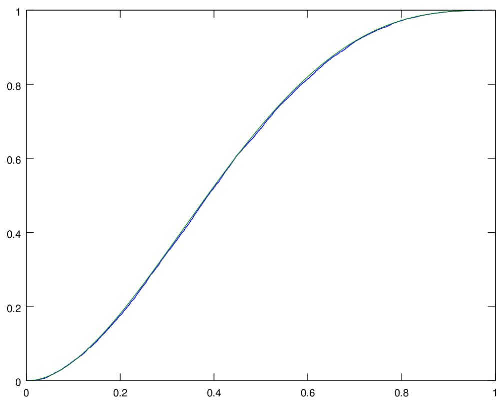
`x` es una matriz en cuyas columnas se representan la sucesión de proporciones de bolas rojas en cada proceso. La instrucción `ones(1,m)` genera una matriz de `1`×`m` con unos en todas las entradas. Al multiplicar por `r/(r+v)` se obtiene una matriz con el valor `r/(r+v)` en todas sus entradas. En otras palabras, `x` es el vector de proporciones de bolas rojas iniciales para cada proceso.

El ciclo `for` es para calcular las sucesiones de proporciones. Se puede comparar con el del problema [2.2] o su versión corregida que se encuentra en la solución de dicho problema.

`y` es la colección de proporciones de bolas rojas finales de cada proceso. Para fines de graficación, se ordenan con el comando `sort`.

En la penúltima linea se ve el comando `plot` a quien se le pasan la colección de proporciones finales y una representación de una función beta con parámetros de forma (shape parameters) `r/c` y `v/c`.

La siguiente gráfica es el resultado de una ejecución del algoritmo anterior. En azul, se muestra el resultado experimental de las `10000` iteraciones del proceso. En verde la gráfica de una distribución Beta con parámetros de forma `r/c` y `v/c`.



Gráfica del histagrama de los radios "finales" de `10000` iteraciones del proceso de urnas de Póyla (azul). En contraste con una distribución Beta (verde).

En clase se demostró que los radios Radios? Supongo que es un anglicismo proveniente de ratios. La palabra correcta es cociente. esperados tenían distribución Beta con los parámetros ya mencionados. En la gráfica se puede apreciar el parecido del resultado experimental con el teórico.

P.S. `tic`, `toc` son funciones de octave para medir el tiempo de ejecución. En este caso la ejecución duró cerca de un minuto. Lo cual era de esperarse, se realizaron poco más de `120,000,000` operaciones y además, construir un gráfico con una retícula de `10000` puntos también es un proceso computacionalmente caro.

3.4.2. **Inciso (ii).** *Ejecute y explique la función del siguiente código en Octave. Incluya una gráfica en la que la longitud de la variable k sea mayor a 1000. (Puede modificar el programa...) En la gráfica observara un esbozo de la trayectoria de un proceso de ramificación continuo (en una escala distinta...).*

```
k   = [10];
aux = k(length(k));

while (aux>0 && length(k)<1000)
    k   = [k;2*binornd(aux, .5)];
    aux = k(length(k));
endwhile

plot(k)
```

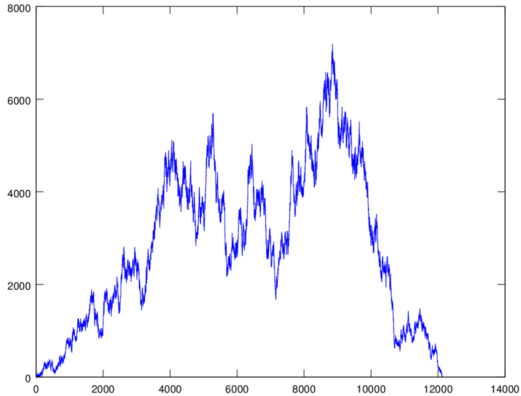
El código simula un proceso de Galton-Watson.

`k = [10]` significa que al inicio habrá 10 individuos.

`aux` es la última población obtenida. `binornd(aux, .5)` es una instrucción de Octave en la que de un conjunto de `aux` elementos, escoge a cada uno con probabilidad `.5`. El resultado es el número de elementos seleccionados. Multiplicar esto por `2` se puede interpretar como que antes de morir, cada individuo seleccionado tiene dos hijos.

`k = [k; 2*binornd(aux, .5)]` está agregando al vector `k` una nueva entrada. La entrada, es la nueva población, resultante de que cada individuo seleccionado tiene dos hijos antes de morir y cada elemento no seleccionado muere sin tener decendientes.

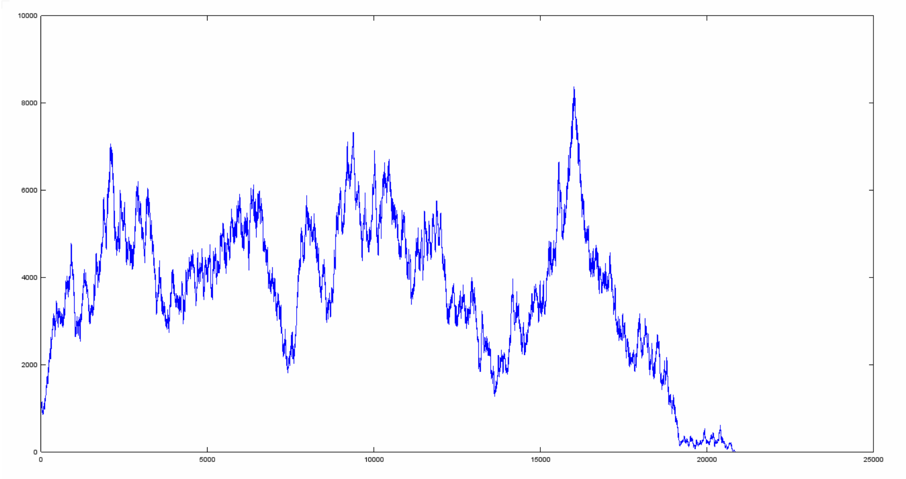
A continuación, una gráfica con una ejecución de este algoritmo comenzando con `10` individuos y terminando con `0` individuos después de un poco más de `12000` iteraciones.



Gráfica de una ejecución del proceso de extinción de Galton-Watson
Parámetros:
10 individuos
La probabilidad de que un individuo tenga 2 hijos es $\frac{1}{2}$
La probabilidad de que un individuo tenga 0 hijos es $\frac{1}{2}$.

La esperanza de número de decendientes que tiene un individuo, es `1`. Es decir, este proceso de Galton-Watson es crítico. En clase se demostró que los procesos Galton-Watson críticos se extinguen con probabilidad `1`. La gráfica arriba mostrada es de una corrida del algoritmo que más iteraciones duró antes de extinguirse, el algoritmo, se corrió `200,000` veces y, en todos los casos se extingió la población. Esto es consistente con con el resultado teórico (en todos los casos, de un conjunto de `200,000`, la población se extinguió).

A continuación, una gráfica resultante del mismo algoritmo, pero con una población inicial de `1000` individuos. que se extingue después de `20,000` iteraciones. Cabe mencionar que esta corrida fue más "suertuda" que la representada en la gráfica anterior. Muchas de las corridas del algoritmo sin modificar superaron una población de `1000` individuos en algún momento y después de eso, nunca duraron tanto como este otro.



Gráfica de una ejecución del proceso de extinción de Galton-Watson
Parámetros:
1000 individuos iniciales
La probabilidad de que un individuo tenga 2 hijos es $\frac{1}{2}$
La probabilidad de que un individuo tenga 0 hijos es $\frac{1}{2}$.

P.S. El algoritmo modificado se incluye bajo el nombre de `BinaryGWModificado.R`. (Incluí un candado para que el algoritmo terminara si superaba un millón de iteraciones, nunca se alcanzó dicho límite).

Part 4. Tarea 4

4.1. Problema 4.1

Problema 4.1.1. 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

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

(i) [4.1.1] ¿Qué quiere decir la independencia condicional cuando $\mathscr{G} = \{\Omega, \emptyset\}$?

(ii) [4.1.2] 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

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

(iii) [4.1.3] 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 $\mathscr{G}$.

Categorías: Esperanza condicional, Independencia condicional.

Demostración:

4.1.1. Inciso (i). *¿Qué quiere decir la independencia condicional cuando $\mathscr{G} = \{\Omega, \emptyset\}$?*

Sabemos que cualquier σ-álgebra es independiente de la σ-álgebra trivial $\mathscr{G}$. Y entonces $\mathbb{E}(X|\mathscr{G}) = \mathbb{E}(X)$, para toda X.

Si $\mathscr{F}_1, \ldots, \mathscr{F}_n$ son condicionalmente independietnes dada $\mathscr{G}$ y H_i es $\mathscr{F}_i$-medible para toda i, entonces tenemos

$$\begin{aligned}
\mathbb{E}(H_1 \cdots H_n) &= \mathbb{E}(H_1 \cdots H_n|\mathscr{G}) \\
&= \mathbb{E}(H_1|\mathscr{G}) \cdots \mathbb{E}(H_n|\mathscr{G}) \\
&= \mathbb{E}(H_1) \cdots \mathbb{E}(H_n|\mathscr{G}).
\end{aligned}$$

Entonces, podemos concluir que cualesquiera funciones H_i así definidas son siempre independientes.

Esto ocurre si y sólamente si las σ-álgebras $\mathscr{F}_1, \mathscr{F}_2, \ldots$ son independientes.

Entonces podemos concluir que cuando una sucesión de σ-álgebras $\mathscr{F}_1, \mathscr{F}_2, \ldots$ son condicionalmente independientes dada la σ-álgebra trivial, entonces son independientes.

4.1.2. **Inciso (ii).** *Pruebe que $\mathscr{F}_1$ 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*

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

Supongamos que $\mathscr{F}_1 \perp_{\mathscr{G}} \mathscr{F}_2$ y, sea H una función $\mathscr{F}_1$-medible y acotada.

Sea $P = \{G \cap F : G \in \mathscr{G}, F \in \mathscr{F}_2\}$. Es claro ver que este conjunto no es vacío (el total pertenece a él) y es fácil notar que es cerrado bajo intersecciones finitas. En otras palabras P es un π-sistema. Nótese que $\mathscr{G}, \mathscr{F}_2 \subset P$ y que $\sigma(P) = \sigma(\mathscr{G}, \mathscr{F}_2)$. Esto será importante más adelante.

Sea ahora $D = \{A \in \sigma(\mathscr{G}, \mathscr{F}_2) : \mathbb{E}(H\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_A)\}$. Veamos que D así definido es un λ-sistema (o sistema de Dynkin).

- Es claro que $\Omega \in D$.
- Veamos que D es cerrado bajo sucesiones crecientes.

  Sean $A_1, A_2, A_3, \dots \in D$ tales que $A_i \subset A_{i+1}$. Entonces

$$\begin{aligned}
\mathbb{E}\left(H\mathbb{1}_{\bigcup A_i}\right) &= \lim_{n\to\infty} \mathbb{E}(H\mathbb{1}_{A_n}) \\
&\qquad \text{(Gracias a que } H \text{ es acotada)} \\
&= \lim_{n\to\infty} \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_{A_n}) \\
&= \mathbb{E}\left(\mathbb{E}(H|\mathscr{G})\mathbb{1}_{\bigcup A_i}\right).
\end{aligned}$$

  Es decir, $\bigcup A_i \in D$.

- Sean $A, B \in D$ con $B \subset A$. Veamos ahora que D es cerrado bajo diferencias de este tipo de conjuntos.

$$\begin{aligned}
\mathbb{E}(H\mathbb{1}_{(A\setminus B)} &= \mathbb{E}(H(\mathbb{1}_A - \mathbb{1}_B)) \\
&\qquad \text{(Esta descomposición es válida gracias} \\
&\qquad \text{a la contención de } B \text{ en } A) \\
&= \mathbb{E}(H\mathbb{1}_A) - \mathbb{E}(H\mathbb{1}_B) \\
&= \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_A) - \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_B) \\
&\qquad \text{(Hipótesis de que } A, B \in D)
\end{aligned}$$

$$= \mathbb{E}(E(H|\mathscr{G})\mathbb{1}_{A\setminus B})$$

Es decir, $A \setminus B \in D$.

Con esto queda demostrado que D es un sistema de Dynkin. Si $P \subset D$ entonces el lema de clases de Dynkin asegura que $\sigma(\mathscr{G}, \mathscr{F}_2) = \sigma(P) \subset D$ y entonces habríamos terminado la demostración de la suficiencia.

Sea entonces $G \cap F \in P$ con $G \in \mathscr{G}$ y $F \in \mathscr{F}_2$. Entonces

$$\begin{aligned}
\mathbb{E}(H\mathbb{1}_{G\cap F}) &= \mathbb{E}(H\mathbb{1}_G\mathbb{1}_F) \\
&= \mathbb{E}(\mathbb{E}(H\mathbb{1}_G\mathbb{1}_F)|\mathscr{G}) \\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H\mathbb{1}_F\mathscr{G})) \\
&\quad \text{(Gracias a que } \mathbb{1}_G \text{ es } \mathscr{G}\text{-medible)} \\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H|\mathscr{G})\mathbb{E}(\mathbb{1}_F|\mathscr{G})) \\
&\quad \text{(Hipótesis de que } F_1, F_2 \text{ son} \\
&\quad\ \text{condicionalmente independientes dada } \mathscr{G}) \\
&= \mathbb{E}\Big(\mathbb{E}\big(\mathbb{1}_G\mathbb{E}(H|\mathscr{G})\mathbb{1}_F|\mathscr{G}\big)\Big) \\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H|\mathscr{G})\mathbb{1}_F) \\
&= \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_G\mathbb{1}_F) \\
&= \mathbb{E}(\mathbb{E}(H|\mathscr{G})\mathbb{1}_{G\cap F}).
\end{aligned}$$

Con lo que queda demostrado que $G \cap F \in D$, es decir que $P \subset D$ y por lo tanto la suficiencia queda demostrada.

Demostremos ahora la necesidad. Sean H_1, H_2 funciones $\mathscr{F}_1, \mathscr{F}_2$-medibles (respectivamente) y acotadas. Nuestra hipótesis dice que

$$\mathbb{E}(H_1|\mathscr{G}, \mathscr{F}_2) = \mathbb{E}(H_1|\mathscr{G}). \tag{4.1.2.1}$$

Sea entonces $A \in G$

$$\begin{aligned}
\mathbb{E}(H_1H_2\mathbb{1}_A) &= \mathbb{E}(\mathbb{E}(H_1H_2\mathbb{1}_A|\sigma(\mathscr{G}, \mathscr{F}_2))) \\
&= \mathbb{E}(H_2\mathbb{E}(H_1\mathbb{1}_A|\sigma(\mathscr{G}, \mathscr{F}_2))) \\
&\quad \text{(gracias a que } H_2 \text{ es } \mathscr{F}_2\text{-medible)} \\
&= \mathbb{E}(H_2\mathbb{1}_A\mathbb{E}(H_1|\sigma(\mathscr{G}, \mathscr{F}_2))) \\
&\quad \text{(gracias a que } A \in \mathscr{G}) \\
&= \mathbb{E}(H_2\mathbb{1}_A\mathbb{E}(H_1|\mathscr{G})) \\
&\quad \text{(gracias a la hipótesis (4.1.2.1))}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}(\mathbb{E}(H_2 \mathbb{1}_A \mathbb{E}(H_1|\mathscr{G})|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{1}_A \mathbb{E}(H_2 \mathbb{E}(H_1|\mathscr{G})|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{1}_A \mathbb{E}(H_1|\mathscr{G}) \mathbb{E}(H_2|\mathscr{G})) \\
&\qquad \text{(porque } \mathbb{E}(H_1|\mathscr{G}) \text{ es } \mathscr{G}\text{-medible)}
\end{aligned}
$$

Y con esto queda demostrado que cualquier versión de $\mathbb{E}(H_1|\mathscr{G})\mathbb{E}(H_2|\mathscr{G})$ es también una versión de $\mathbb{E}(H_1 H_2|\mathscr{G})$ y entonces, $\mathscr{F}_1$ y $\mathscr{F}_2$ son condicionalmente independientes dada $\mathscr{G}$.

4.1.3. **Inciso (iii).** *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 $\mathscr{G}$.*

Primero demostraremos la suficiencia. Para ello emplearemos el ejercicio anterior. Sean $H_1, \ldots, H_{n+1}$ variables aleatorias $\mathscr{F}_1, \ldots, \mathscr{F}_{n+1}$-medibles (respectivamente) y acotadas.

Sea $P = \{A : A = F_1 \cap \cdots \cap \mathscr{F}_n F_n \cap \mathscr{G} G$ donde $F_i \in \mathscr{F}_i$ y $G \in \mathscr{G}\}$, Es claro que $P \subset \sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n, \mathscr{G})$. También es fácil ver que no es vacía (Ω es uno de sus elementos), y que es cerrada bajo intersecciones finitas. En otras palabras C es un π-sistema. Además, justo igual que antes, tenemos que $\sigma(P) = \sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n, \mathscr{G})$.

Ahora sea $D = \{A \in \sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n, \mathscr{G}) : \mathbb{E}(H_{n+1}\mathbb{1}_A) = \mathbb{E}(\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_A)\}$
D es un sistema de Dynkin. Verificar que $\Omega \in D$ es trivial. Y las propiedades de ser cerrado bajo sucesiones crecientes y bajo diferencias (diferencias donde un conjunto está contenido en el otro) se demuestran exactamente igual que en el inciso anterior de este mismo problema [4.1.2].

Entonces, como en el ejercicio anterior, basta demostrar que $P \subset D$, y entonces el lema de clases de Dynkin nos asegurará que

$$\mathbb{E}(H_{n+1}|\sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n, \mathscr{G})) = \mathbb{E}(H_{n+1}|\mathscr{G}) \tag{4.1.3.1}$$

Sea entonces $A \in P$. Donde $A = F_1 \cap \cdots \cap F_n \cap G$ con $F_i \in \mathscr{F}_i$ y $G \in \mathscr{G}$.

$$\begin{aligned}
\mathbb{E}(H_{n+1}\mathbb{1}_A) &= \mathbb{E}(H_{n+1}\mathbb{1}_{F_1} \cdots \mathbb{1}_{F_n}\mathbb{1}_G) \\
&= \mathbb{E}(\mathbb{E}(H_{n+1}\mathbb{1}_{F_1} \cdots \mathbb{1}_{F_n}\mathbb{1}_G|\mathscr{G})) \\
&= \mathbb{E}(\mathbb{1}_G\mathbb{E}(H_{n+1}\mathbb{1}_{F_1} \cdots \mathbb{1}_{F_n}|\mathscr{G})) \\
&\quad \text{(gracias a que } G \in \mathscr{G}\text{)} \\
&= \mathbb{E}\Big(\mathbb{1}_G\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{E}\big(\mathbb{1}_{F_1} \cdots \mathbb{1}_{F_n}|\mathscr{G}\big)\Big) \\
&\quad \text{(gracias a que } \mathscr{F}_1, \mathscr{F}_2, \ldots \text{ son} \\
&\quad\ \text{condicionalmente independientes dada } \mathscr{G}\text{)} \\
&= \mathbb{E}\Big(\mathbb{E}\big(\mathbb{1}_G\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_{F_1} \cdots \mathbb{1}_{F_n}|\mathscr{G}\big)\Big) \\
&\quad \text{(nótese que las variables que entraron son todas} \\
&\quad\ \mathscr{G}\text{-medibles)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\Big(\mathbb{E}\big(\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_G\mathbb{1}_{F_1}\cdots\mathbb{1}_{F_n}|\mathscr{G}\big)\Big)\\
&= \mathbb{E}\Big(\mathbb{E}\big(\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_A|\mathscr{G}\big)\Big)\\
&= \mathbb{E}\big(\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_A\big).
\end{aligned}
$$

Con esto queda demostrado que $A \in D$ y por lo tanto $P \subset D$. Utilizando el lema de clases de Dynkin obtenemos (4.1.3.1), y aplicando el resultado del inciso anterior [4.1.2], obtenemos que $\mathscr{F}_{n+1}$ es condicionalmente independiente de $\mathscr{F}_1, \ldots, \mathscr{F}_n$ dada $\mathscr{G}$.

Ahora demostraremos la necesidad. Nuestra hipótesis es que para cada $n \geq 1$, $\mathscr{F}_{n+1}$ es condicionalemente independiente de $\mathscr{F}_1, \ldots, \mathscr{F}_n$ dada $\mathscr{G}$. La demostración se hará por inducción sobre n.

- **Base De Inducción.**
  Si $n = 1$, no hay nada que demostrar, pues por hipótesis $\mathscr{F}_2$ es condicionalmente independiente de $\mathscr{F}_1$ dada $\mathscr{G}$.

- **Hipótesis De Inducción.**
  Sea $n \geq 2$ tal que $\mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_n$ son condicionalmente independientes (nótese que la base de inducción asegura que existe tal $n \geq 2$).

- **Paso Inductivo.**
  Sean H_i variables $\mathscr{F}_i$-medibles y acotadas con $1 \leq i \leq n+1$.

  Dado que $\mathscr{F}_{n+1}$ es condicionalmente independiente de $\mathscr{F}_1, \ldots, \mathscr{F}_n$ dada $\mathscr{G}$, la equivalencia de [4.1.2] nos dice

$$
\mathbb{E}(H|\sigma(\mathscr{F}_1, \ldots, \mathscr{F}_n, \mathscr{G})) = \mathbb{E}(H|\mathscr{G}). \tag{4.1.3.2}
$$

Sea entonces $A \in \mathscr{G}$.

$$
\begin{aligned}
\mathbb{E}(H_1\cdots H_{n+1}\mathbb{1}_A) &= \mathbb{E}(\mathbb{E}(H_1\cdots H_{n+1}\mathbb{1}_A|\sigma(\mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_n, \mathscr{G})))\\
&= \mathbb{E}(H_1\cdots H_n\mathbb{1}_A\mathbb{E}(H_{n+1}|\sigma(\mathscr{F}_1, \mathscr{F}_2, \ldots, \mathscr{F}_n, \mathscr{G})))\\
&= \mathbb{E}(H_1\cdots H_n\mathbb{1}_A\mathbb{E}(H_{n+1}|\mathscr{G}))\\
&\qquad \text{(ésto último gracias a (4.1.3.2))}\\
&= \mathbb{E}(\mathbb{E}(H_1\cdots H_n\mathbb{1}_A\mathbb{E}(H_{n+1}|\mathscr{G})|\mathscr{G}))\\
&= \mathbb{E}(\mathbb{E}(H_1\cdots H_n|\mathscr{G})\mathbb{1}_A\mathbb{E}(H_{n+1}|\mathscr{G}))
\end{aligned}
$$

(se sacaron las variables $\mathscr{G}$-medibles)

$$= \mathbb{E}(\mathbb{E}(H_1 \cdots H_n|\mathscr{G})\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_A)$$
$$= \mathbb{E}(\mathbb{E}(H_1|\mathscr{G}) \cdots \mathbb{E}(H_n|\mathscr{G})\mathbb{E}(H_{n+1}|\mathscr{G})\mathbb{1}_A)$$

(por hipótesis de inducción)

Por lo tanto, sobre $\mathscr{G}$, cualquier versión de $\mathbb{E}(H_1|\mathscr{G}) \cdots \mathbb{E}(H_n|\mathscr{G})\mathbb{E}(H_{n+1}|\mathscr{G})$ siempre integra lo mismo que $H_1 \cdots H_{n+1}$.

Con esto queda demostrado que $F_1, F_2, \cdots$ son condicionalmente independientes dada $\mathscr{G}$.

4.2. Problema 4.2

Problema 4.2.1. 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

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

(i) [4.2.1] Pruebe que $Z_n \geq 0$ para toda n y que si $Z_n = 0$ entonces $Z_{n+1} = 0$.

(ii) [4.2.2] Pruebe que C_n es un tiempo de paro para la filtración canónica asociada a S.

(iii) [4.2.3] Pruebe que Z es un proceso de Galton-Watson con ley de progenie μ.

(iv) [4.2.4] 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.)

Categorías: Caminatas aleatorias, Procesos de Galton-Watson

Demostración:

4.2.1. Inciso (i).

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

Sean $(\xi_i)_{i\in\mathbb{N}}$ v.a.i.i.d con distribución $\tilde{\mu}$ (es decir, $\mathbb{P}(\xi_i = s) = \tilde{\mu}_s$ donde $-1 \leq s$).

Es decir que podemos denotar a $S_n = \sum_{i\leq n} \xi_i$.

Primero demostraremos que $Z_n \geq 0$. La demostración se hará por inducción sobre n.

- **Base De Inducción.**
  Para $n = 0$ tenemos $Z_0 = k \geq 0$.

- **Hipótesis De Inducción.**
  Sea n tal que $Z_i \geq 0$ para toda $i \leq n$.

- **Paso Inductivo.**

$$\begin{aligned}
Z_{n+1} &= k + S_{C_n} \\
&= k + \sum_{i \le C_n} \xi_i \\
&= k + \sum_{i \le C_{n-1}+Z_n} \xi_i \\
&= k + \sum_{i \le C_{n-1}} \xi_i + \sum_{C_{n-1}<i\le C_{n-1}+Z_n} \xi_i \\
&= k + S_{C_n-1} + \sum_{C_{n-1}<i\le C_{n-1}+Z_n} \xi_i \\
&= Z_n + \sum_{C_{n-1}<i\le C_{n-1}+Z_n} \xi_i
\end{aligned} \tag{4.2.1.1}$$

En la suma del lado derecho tenemos exáctamente Z_n sumandos, y cada sumando es mayor o igual a -1. Entonces en el peor de los casos la suma de la derecha es igual a $-Z_n$. Es decir,

$$\sum_{C_{n-1}<i\le C_{n-1}+Z_n} \xi_i \ge -Z_n.$$

Y combinando estos resultados tenemos

$$\begin{aligned}
Z_{n+1} &= Z_n + \sum_{C_{n-1}<i\le C_{n-1}+Z_n} \xi_i \\
&\ge Z_n - Z_n \\
&= 0.
\end{aligned}$$

Y con esto termina la demostración.

Ahora demostraremos que si $Z_n = 0$, entonces $Z_{n+1} = 0$.

Notemos los siguientes hechos.

- $Z_n = k + S_{C_{n-1}} = 0$ implica que $S_{C_{n-1}} = -k$.
- $C_n = C_{n-1} + Z_n = C_{n-1}$.

Con esto presente, basta recordar la definición de Z_{n+1}:

$$\begin{aligned}
Z_{n+1} &= k + S_{C_n} \\
&= k + S_{C_{n-1}}
\end{aligned}$$

$$= k - k$$
$$= 0.$$

Que era lo que buscábamos demostrar.

4.2.2. **Inciso (ii).** *Pruebe que C_n es un tiempo de paro para la filtración canónica asociada a S.*

Sea $(\mathscr{F}_n)_{n\in\mathbb{N}}$ la filtración canónica asociada a $(S_n)_{n\in\mathbb{N}}$. Tenemos que demostrar que para $m \in \mathbb{Z}$, $\{C_n = m\} \in \mathscr{F}_m$.

Usemos la definición de C_n para descomponerlo,

$$
\begin{aligned}
C_n &= C_{n-1} + Z_n \\
&= C_{n-2} + Z_{n-1} + Z_n \\
&\vdots \\
&= k + \sum_{1\le i\le n} Z_i
\end{aligned} \tag{4.2.2.1}
$$

Ahora, recordando que $0 \le Z_i$, tenemos que $k \le C_n$, es decir que $\{C_n < k\} = \emptyset$, en particular, si $m < k$, $\{C_n = m\} = \emptyset \in \mathscr{F}_m$.

Supongamos entonces que $k \le m$. De nuevo gracias a que $0 \le Z_i$, tenemos que la sucesión $(C_i)_{i\in\mathbb{N}}$ es creciente. Entonces, dentro del conjunto Usualmente se dice sobre el conjunto en vez de en el conjunto. $\{C_n = m\}$ tenemos que $C_i \le m$ para $i \le n$. Y por lo tanto, dentro del conjunto $\{C_n = m\}$, S_{C_i} es $\mathscr{F}_m$-medible (donde dentro del conjunto se debería interpretar como al intesectar con dicho conjunto, las preimagenes de S_{C_i} pertenecen a $\mathscr{F}_m$).

Entonces $Z_i = S_{C_{i-1}}$ también resulta ser $\mathscr{F}_m$-medible (dentro de $\{C_n = m\}$ y por lo tanto. Entonces nos basta con descomponer $\{C_n = m\}$ en conjuntos que sean $\mathscr{F}_m$-medibles usando lo que hemos demostrado ahora. La prueba puede ser un tanto más explícita en este punto. Utilizando (4.2.2.1) tenemos que

$$
\{C_n = m\} = \left\{k + \sum_{1\le i\le n} Z_i = m\right\}.
$$

Y por todo lo anteriormente dicho, el lado derecho de la igualdad resulta ser $\mathscr{F}_m$-medible.

4.2.3. **Inciso (iii).** *Pruebe que Z es un proceso de Galton-Watson con ley de progenie μ.*

Tenemos que encontrar variables $\zeta_{i,n}$ con distribución μ tales que

$$Z_{n+1} = \sum_{1 \leq i \leq Z_n} \zeta_{i,n}. \tag{4.2.3.1}$$

No: lo que hay que mostrar es que Z tiene las mismas distribuciones finito-dimensionales que el proceso de Galton-Watson. Sean ξ_i con $-1 \leq i$ definidas como en el primer inciso de este ejercicio [4.2.1], estas estaban definidas bajo una distribución $\tilde{\mu}$ que era exactamente igual a μ pero desfasada por -1. Es decir $\tilde{\mu}_{-1} = \mu_0, \tilde{\mu}_0 = \mu_1, \ldots, \tilde{\mu}_j = \mu_{j+1}$. Entonces, $\xi_i + 1$ tendrá distribución μ.

Ya teníamos una descomposición útil de Z_{n+1} en (4.2.1.1). La reescribiremos cambiando ξ_i por $\xi_i + 1 - 1$.

$$\begin{aligned}
Z_{n+1} &= Z_n + \left(\sum_{C_{n-1} < i \leq C_{n-1} + Z_n} \xi_i \right) \\
&= Z_n + \left(\sum_{C_{n-1} < i \leq C_{n-1} + Z_n} (\xi_i + 1 - 1) \right) \\
&= Z_n + \left(\sum_{C_{n-1} < i \leq C_{n-1} + Z_n} (\xi_i + 1) \right) - Z_n \\
&= \left(\sum_{C_{n-1} < i \leq C_{n-1} + Z_n} (\xi_i + 1) \right)
\end{aligned}$$

Ahora que tenemos a Z_n expresado como suma de variables con distribución μ, bastará reescribirlo de manera que tenga la forma (4.2.3.1).

Definamos entonces $\zeta_{i,n} = \xi_{C_{n-1}} + 1$, y entonces

$$Z_{n+1} = \left(\sum_{C_{n-1} < i \leq C_{n-1} + Z_n} (\xi_i + 1) \right)$$

$$= \left(\sum_{1 \leq i \leq Z_n} (\zeta_{i,n}) \right).$$

Y con esto hemos demostrado que Z es un proceso Galton-Watson con distribución de progenie μ y, como $Z_0 = k$, se trata de un proceso Galton-Watson con población inicial k.

4.2.4. **Inciso (iv).** *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.)*

Suponiendo que la media de μ es 1, significa que $\mathbb{E}(\xi_i + 1) = 1$ y por lo tanto $\mathbb{E}(\xi_i) = 0$.

Esto convierte a $(S_n)_{n\in\mathbb{N}}$ en una caminata aleatoria no trivial y centrada (la media de sus saltos es 0). Por lo dicho en [3.2] tenemos que es una caminata que oscila y por lo tanto $\liminf S_n = -\infty$.

Esto quiere decir que existe un n tal que $S_{C_n} = -k$ y entonces $Z_{n+1} = k + S_{C_n} = k - k = 0$. Lo cual se traduce en que la población se extinge casi seguramente.Este punto no es claro: como C_n se puede saltar muchos índices, podría ser (aunque no suena plausible, pero en eso consiste el ejercicio que $S_{C_n} > -k$ para cualquier n.

4.3. Problema 4.3

Problema 4.3.1. 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.

(i) [4.3.1] 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.

(ii) [4.3.2] 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.

Categorías: proyecciones de cadenas de Markov

Demostración:

4.3.1. **Inciso (i).** *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.*

Sean $i, j \in E$. Definamos las siguientes entradas de una matriz de transición:

$$p_{i,j} = \begin{cases} \frac{1}{n}, & \|i - j\| = 1 \quad \text{(donde } \|\cdot\| \text{ es la norma euclidiana)} \\ 0, & \text{en cualquier otro caso} \end{cases}$$

Dado $i \in E$, sólo existen n vectores que distan exáctamente 1 de i. Entonces $\sum_{j\in E} = n\frac{1}{n} = 1$ y nuestra matriz sí es de transición.

Ya que nuestro proceso está restringido a comenzar en x, nuestra distribución inicial será $\lambda = (\lambda_i : i \in E)$ tal que $\lambda_x = 1$ y $\lambda_i = 0$ si $i \neq x$.

Es fácil notar que la evolución del proceso depende únicamente del presente. Para llegar a una configuración de bolas en particular, únicamente es importante en qué configuración se encuentra actualmente.

Dado que tenemos una matriz de transición y una distribución inicial bien definidas, existe una cadena de Markov X con dichas matriz de transcición y distribución inicial que describen la evolución del proceso.

El proceso $(f(X_n))_{n \in N}$ es el proceso que cuenta la cantidad de bolas en la urna 1 que equivale al proceso de Ehrenfest, el cual es una cadena de Markov ¿Por qué? (El objetivo del ejercicio era calcular las distribuciones finito-dimensionales de $f(X_n)$). y las entradas de su matriz de transición están dadas por

$$p_{i,j} = \begin{cases} \frac{n-i}{n}, & j = i+1 \\ \frac{i}{n}, & j = i-1 \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

4.3.2. **Inciso (ii).** *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.*

Tenemos que $|S_{n-1} - S_n| = 1$ con probabilidad 1. Esto significa, los saltos del proceso siempre son de tamaño 1. Esto significa que, la paridad de un estado al siguiente siempre cambia con probabilidad 1. En otras palabras, la paridad se preserva después de una cantidad par de pasos.

Es decir que si $r_1, r_2 \in \mathbb{Z}$ entonces:

$$\mathbb{P}(S_{2n+n_0} = r_2 | S_{n_0} = r_1) = \begin{cases} 1, \text{si } 2 \text{ divide a } r_1 - r_2 \\ 0, \text{si } 2 \text{ no divide a } r_1 - r_2 \end{cases}$$

La paridad se preserva bajo valor absoluto, así que lo mismo ocurre en el caso de $|S_n|$For the records: falta.

4.4. Problema 4.4

Problema 4.4.1. 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

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

(i) [4.4.1] Pruebe que D es una martingala bajo $\mathbb{P}$. Pruebe que si D es uniformemente integrable entonces $\mathbb{Q} \ll \mathbb{P}$.

(ii) [4.4.2] Pruebe que si T es un tiempo de paro finito entonces $\mathbb{Q}|_{\mathscr{F}_T} \ll \mathbb{P}|_{\mathscr{F}_T}$.

(iii) [4.4.3] 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.

(iv) [4.4.4] Para $a, b > 0$, sea $T = \min\{n \in \mathbb{N} : X_n \in \{-a, b\}\}$. 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 $\mathbb{E}(T^2)$.

Categorías: Cambio de medida, Caminata aleatoria simple.

Demostración:

4.4.1. **Inciso (i).** *Pruebe que D es una martingala bajo $\mathbb{P}$. Pruebe que si D es uniformemente integrable entonces $\mathbb{Q} \ll \mathbb{P}$.*

Primero veamos que se trata de una martingala bajo $\mathbb{P}$.

- (**Adaptada a la filtración**) Por definición de derivada de Radon-Nykodim, D_n es $\mathscr{F}_n$-medible.

- (**Cada término es integrable**) Por ser $\mathbb{Q}$ una medida de probabilidad, es no negativa. Por lo tanto cualquier versión que se escoja para la derivada de Radon-Nykodim, es no negativa casi seguramente y por lo tanto $\mathbb{E}(|D_n|) = \mathbb{E}(D_n)$. Ahora

$$\begin{aligned}
\mathbb{E}(D_n) &= \int D_n d\mathbb{P} \\
&= \int D_n d\mathbb{P}|_{\mathscr{F}_n} \\
&\quad \text{(gracias a que } D_n \text{ es } \mathscr{F}_n\text{-medible)} \\
&= \mathbb{Q}|_{\mathscr{F}_n}(\Omega) \\
&= \mathbb{Q}(\Omega) \\
&= 1 \\
&< \infty.
\end{aligned}$$

- (**Condición de martingala**) Sea $A \in \mathscr{F}_n$. Entonces

$$\begin{aligned}
\mathbb{E}(D_{n+1}\mathbb{1}_A) &= \int D_{n+1}\mathbb{1}_A d\mathbb{P} \\
&= \int D_{n+1}\mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_{n+1}} \\
&\quad \text{(porque } D_{n+1}, \mathbb{1}_A \text{ son } \mathscr{F}_{n+1}\text{-medibles)} \\
&= \mathbb{Q}|_{\mathscr{F}_{n+1}}(A) \\
&= \mathbb{Q}|_{\mathscr{F}_n}(A) \\
&\quad \text{(porque } D_n, \mathbb{1}_A \text{ son } \mathscr{F}_n\text{-medibles)} \\
&= \int D_n\mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_n} \\
&= \int D_n\mathbb{1}_A d\mathbb{P} \qquad = \mathbb{E}(D_n\mathbb{1}_A).
\end{aligned}$$

Lo que demuestra, que bajo $\mathscr{F}_n$, tanto D_{n+1} como D_n integran lo mismo. Es decir $\mathbb{E}(D_{n+1}|\mathscr{F}_n) = \mathscr{F}_n$, como queríamos demostrar.

Ahora supongamos que D es uniformemente integrable. Tenemos un teorema que dice que si un martingala es uniformemente integrable entonces existe una variable D_∞ tal que $\mathbb{E}(D_\infty|\mathscr{F}_n)$. (Este teorema es el 1.9 de la versión de las notas de la clase que se incluye junto con esta tarea, ver [7.2.4]).

Definimos $\mathscr{F}_\infty = \sigma\left(()\bigcup \mathscr{F}_i\right)$. El mismo teorema asegura que D_n converge casi seguramente y en L_1 a $\mathbb{E}(D_\infty|\mathscr{F}_\infty)$.
Ahora definimos

$$P = \bigcup_{i \in N} \mathscr{F}_i.$$

Es fácil ver que P es un álgebra. Claramente P es no vacío, todos los $\mathscr{F}_i$ son subconjuntos de él. Si $A_1, A_2 \in P$ entonces existen $i, j \in \mathbb{N}$ tales que $A_1 \in \mathscr{F}_i$ y $A_2 \in \mathscr{F}_j$ y por lo tanto $A_1 \cap A_2, A_1 \cup A_2 \in \mathscr{F}_{\max(i,j)}$. Que sea cerrado bajo complementos se hereda directamente de que cada $\mathscr{F}_i$ es una álgebra.

Ahora sea

$$M = \left\{ A \in \mathscr{F}_i nfty : \mathbb{Q}(A) = \int_A D_\infty d\mathbb{P} \right\}$$

Veamos que M es una clase monótona. Es decir, cerrada bajo sucesiones crecientes y decrecientes.

Sean $A_1, A_2, \cdots \in M$ tales que $A_i \subset A_{i+1}$.

$$\begin{aligned}
\mathbb{Q}\left(\bigcup A_i\right) &= \lim_{i\to\infty} \mathbb{Q}(A_i) \\
&= \lim_{i\to\infty} \int_{A_i} D_\infty d\mathbb{P} \\
&= \lim_{i\to\infty} \int D_\infty \mathbb{1}_{A_i} d\mathbb{P} \\
&= \int D_\infty \lim_{i\to\infty} \mathbb{1}_{A_i} d\mathbb{P} \\
&\quad \text{(por que } A_n \text{ es creciente} \\
&\quad \text{y por teorema de convergencia mónotona)} \\
&= \int D_\infty \mathbb{1}_{\cup A_i} d\mathbb{P} \\
&= \int_{\cup A_i} D_\infty d\mathbb{P}
\end{aligned}$$

Tomando complementos, y haciendo una demostración análoga, se tiene que M también es cerrada bajo sucesiones decrecientes.

Ahora veamos que $P \subset M$. Sea $A \in P$. Entonces existe $n \in \mathbb{N}$ tal que $A \in \mathscr{F}_n$. Entonces

$$\begin{aligned}
\mathbb{Q}(A) &= \mathbb{Q}|_{F_n}(A) \\
&= \int D_n \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_n} \\
&= \int \mathbb{E}(D_\infty|\mathscr{F}_n) \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_n}
\end{aligned}$$

$$= \int D_\infty \mathbb{1}_A d\mathbb{P}|_{\mathscr{F}_n}$$

(la integración ya se hace sobre $\mathscr{F}_n$)

$$= \int D_\infty \mathbb{1}_A d\mathbb{P}.$$

Y por lo tanto $A \in M$. El teorema de clases monótonas garantiza que $\sigma(P) \subset M$. Pero $\sigma(P) = \mathscr{F}_\infty$ y $M \subset \mathscr{F}_\infty$, por lo tanto $\sigma(P) = \mathscr{F}_\infty = M$. Entonces, tenemos que para todo $A \in \mathscr{F}_\infty$

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

Por lo tanto $\mathbb{Q} \ll \mathbb{P}$. Como buscábamos demostrar.

4.4.2. **Inciso (ii).** *Pruebe que si T es un tiempo de paro finito entonces* $\mathbb{Q}|_{\mathscr{F}_T} \ll \mathbb{P}|_{\mathscr{F}_T}$.

Gracias a [1.4.3] tenemos que D_T es $\mathscr{F}_T$-medible.

Recordemos que $D_T = \sum_{n\in\mathbb{N}} D_n \mathbb{1}_{T=n}$.

Sea entonces $A \in \mathscr{F}_T$.

$$
\begin{aligned}
\int_A D_T d\mathbb{P}|_{\mathscr{F}_T} &= \int_A D_T d\mathbb{P} \\
&\quad \text{(porque } D_T \text{ es } \mathscr{F}_T\text{-medible)} \\
&= \int_A \sum_{n\in\mathbb{N}} D_n \mathbb{1}_{\{T=n\}} d\mathbb{P} \\
&= \sum_{n\in\mathbb{N}} \int_A D_n \mathbb{1}_{\{T=n\}} d\mathbb{P} \\
&\quad \text{(por teorema de convergencia monótona} \\
&\quad\ \text{dado que } D_n \text{ es no negativa)} \\
&= \sum_{n\in\mathbb{N}} \int D_n \mathbb{1}_{\{T=n\}\cap A} d\mathbb{P} \\
&= \sum_{n\in\mathbb{N}} \int D_n \mathbb{1}_{\{T=n\}\cap A} d\mathbb{P}|_{\mathscr{F}_n} \\
&\quad \text{(porque } \{T=n\}\cap A \in \mathscr{F}_n \text{ por definición de } \mathscr{F}_T\text{)} \\
&= \sum_{n\in\mathbb{N}} \mathbb{Q}(\{T=n\}\cap A) \\
&= \mathbb{Q}(\bigcup_{n\in\mathbb{N}} \{T=n\}\cap A) \\
&\quad \text{(porque los eventos \{T=n\} son disjuntos)} \\
&= \mathbb{Q}(A) \\
&\quad \text{(porque } T \text{ es finito)} \\
&= \mathbb{Q}|_{\mathscr{F}_T}(A) \\
&\quad \text{(porque } A \in \mathscr{F}_T\text{)}
\end{aligned}
$$

En resumen tenemos

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

para todo $A \in \mathscr{F}_T$ y, por lo tanto $\mathbb{Q}|_{\mathscr{F}_T} \ll \mathbb{P}|_{\mathscr{F}_T}$.

4.4.3. **Inciso (iii).** *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.*

4.4.4. **Inciso (iv).** *Para* $a, b > 0$*, sea* $T = \min\{n \in \mathbb{N} : X_n \in \{-a, b\}\}$*. 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* $\mathbb{E}(T^2)$*.*

Part 5. Tarea 5

5.1. Problema 5.1

Problema 5.1.1. Sea N un proceso Poisson de parámetro λ y sea T_n el tiempo de su enésimo salto.

(i) [5.1.1] Pruebe que condicionalmente a T_2, T_1 es uniforme en $[0, T_2]$.

(ii) [5.1.2] 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.

(iii) [5.1.3] Conjeture cómo se generaliza lo anterior con T_n y T_1.

(iv) [5.1.4] 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:

5.1.1. Inciso (i). *Pruebe que condicionalmente a T_2, T_1 es uniforme en $[0, T_2]$.*

Por definición de proceso Poisson de parámetro λ, la distribución de los saltos S_n es exponencial de parámetro λ. Como $T_1 = S_1$ y $T_2 = S_1 + S_2$, entonces $T_1 = S_1 \sim exp(\lambda)$ y $T_2 - T_1 = S_2 \sim exp(\lambda)$.

También por definición de proceso de Poisson, los saltos son independientes y por lo tanto T_1 y $T_2 - T_1$ son independientes.

Podemos dar su distribución conjunta como sigue:

$$\begin{aligned} f_{T_1, T_2 - T_1}(u, v) &= (\lambda e^{-\lambda u})(\lambda e^{-\lambda v})\mathbb{1}_{\{u>0, v>0\}} \\ &= \lambda^2 e^{-\lambda(u+v)}\mathbb{1}_{\{u>0, v>0\}}. \end{aligned}$$

Hacemos el cambio de variable dado por la transformación lineal $l : (u, v - u) \longrightarrow (u, v)$. La matriz asociada a esta transformación lineal tiene la forma

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

cuyo determinante es $(1 \cdot 1) - (1 \cdot 0) = 1$ y por lo tanto tiene inversa (es fácil comprobar que $l^{-1}(u, v) = (u, v - u)$). Entonces, podemos escribir

$$\begin{aligned} f_{T_1,T_2}(u, v) &= f_{T_1,T_2-T_1}(l^{-1}(u, v)) \\ &= f_{T_1,T_2-T_1}(u, v - u) \\ &= \lambda^2 e^{-\lambda(u+v-u)} \mathbb{1}_{\{u>0, v-u>0\}}. \\ &= \lambda^2 e^{-\lambda(v)} \mathbb{1}_{\{v>u>0\}}. \end{aligned}$$

Ahora, $T_2 = (T_1) + (T_2 - T_1)$, de donde $T_2 \sim gamma(2, \lambda)$ y por lo tanto su función de densidad esta dada por

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

Con todo esto, ahora podemos calcular la densidad condicional $f_{T_1|T_2}$.

$$\begin{aligned} f_{T_1|T_2}(u|v) &= \frac{f_{T_1,T_2}(u, v)}{f_{T_2}(v)} \\ &= \frac{\lambda^2 e^{-\lambda(v)} \mathbb{1}_{\{v>u>0\}}}{\lambda^2 v e^{-\lambda v} \mathbb{1}_{v \geq 0}} \\ &= \frac{\mathbb{1}_{\{v>u>0\}}}{v \mathbb{1}_{v \geq 0}} \\ &= \frac{\mathbb{1}_{\{v>u>0\}}}{v}. \end{aligned}$$

Y esto no es otra cosa que la distribución uniforme sobre $[0, v]$, es decir, la distribución uniforme sobre $[0, T_2]$.

5.1.2. **Inciso (ii).** *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.*

La idea era deducirlo del inciso anterior. Como W_1, W_2 son exponenciales de parámetro λ, entonces $W_1+W_2 \sim gamma(2,\lambda)$. Entonces

$$\begin{aligned}
f_{U(W_1+W_2)}(v) &= \int f_{W_1+W_2,U(W_1+W_2)}(u,v)du \\
&= \int f_{W_1+W_2}(u) f_U(v/u)\frac{1}{u}du \\
&\quad \text{(esto último por la independencia de} \\
&\quad U \text{ y } W_1+W_2\text{)} \\
&= \int \lambda^2 u e^{-\lambda u} \mathbb{1}_{\{0<u\}} f_U(v/u)\frac{1}{u}du \\
&= \int \lambda^2 u e^{-\lambda u} \mathbb{1}_{\{0<u\}} \mathbb{1}_{\{0\le v/u\le 1\}}\frac{1}{u}du \\
&= \int \lambda^2 e^{-\lambda u} \mathbb{1}_{\{0<u\}} \mathbb{1}_{\{0\le v/u\le 1\}}du \\
&= \int_v^\infty \lambda^2 e^{-\lambda u}du \\
&= \lambda^2 \int_v^\infty e^{-\lambda u}du \\
&= \lambda^2 \frac{1}{\lambda} e^{-\lambda v} \\
&= \lambda e^{-\lambda v}
\end{aligned}$$

Y con esto queda demostrado que $U(W_1+W_2)$ tiene distribución exponencial de parámetro λ.

5.1.3. **Inciso (iii).** *Conjeture cómo se generaliza lo anterior con T_n y T_1.* Por definición de proceso de Poisson de parámetro λ, los saltos S_n se distribuyen de manera exponencial de parámetro λ y son independientes entre sí.

Es decir que $T_1 = S_1, T_2 - T_1 = S_2, \ldots, T_n - T_{n-1} = S_n$ son independientes y todos tienen distribución exponencial de parámetro λ. Por lo tanto, su función de densidad conjunta está dada por

$$(5.1.3.1)\quad f_{T_1,T_2-T_1,\ldots,T_n-T_{n-1}}(u_1, u_2, \ldots, u_n) = \prod_{1\le i\le n} \lambda e^{-\lambda u_i} \mathbb{1}_{\{0<u_i\}}.$$

También recordemos que T_n es suma de n variables independientes con distribución exponencial de parámetro λ y por lo tanto, T_n tiene distribución *gamma*(n, λ). Por lo tanto su función de densidad está dada por

$$\begin{aligned} f_{T_n}(u) &= \frac{\lambda^n u^{n-1} e^{-\lambda u}}{\Gamma(n)} \\ &= \frac{\lambda^n u^{n-1} e^{-\lambda u}}{(n-1)!} \end{aligned}$$

De manera similar a como se procedió en [5.1.1] definimos la transformación lineal $l : (t_1, t_2 - t_1, \ldots, t_n - t_{n-1}) \longrightarrow (t_1, t_2, \ldots, t_n)$ cuya matriz asociada es la matriz triangular superior

$$\begin{pmatrix} 1 & 1 & 1 & \ldots & 1 \\ 0 & 1 & 1 & \ldots & 1 \\ 0 & 0 & 1 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}$$

cuyo determinante es 1 (esto úlitmo no es difícil de verificar). Es decir, tiene inversa y también es fácil de verificar que su inversa está dada por
$l^{-1} : (x_1, x_2, \ldots, x_n) \longrightarrow (x_1, x_2 - x_1, \ldots, x_n - x_{n-1})$.

Entonces podemos escribir

$$\begin{aligned} f_{T_1,T_2,\ldots,T_n}(u_1, u_2, \ldots, u_n) &= f_{T_1,T_2-T_1,\ldots,T_n-T_{n-1}}(l^{-1}(u_1, u_2, \ldots, u_n)) \\ &= f_{T_1,T_2-T_1,\ldots,T_n-T_{n-1}}(u_1, u_2 - u_1, \ldots, u_n - u_{n-1}). \end{aligned}$$

Para comodidad de notación, definimos $u_0 = 0$ y entonces esta última ecuación la podemos escribir como

$$f_{T_1,T_2,\ldots,T_n}(u_1,u_2,\ldots,u_n) = \\ f_{T_1,T_2-T_1,\ldots,T_n-T_{n-1}}(u_1-u_0,u_2-u_1,\ldots,u_n-u_{n-1}).$$

Utilizando (5.1.3.1) obtenemos

$$\begin{aligned} f_{T_1,T_2,\ldots,T_n}(u_1,u_2,\ldots,u_n) &= \prod_{1\leq i\leq n} \lambda e^{-\lambda u_i - u_{i-1}} \mathbb{1}_{\{0<u_i-u_{i-1}\}} \\ &= \prod_{1\leq i\leq n} \lambda e^{-\lambda u_i - u_{i-1}} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}} \\ &= \lambda^n e^{-\lambda u_n} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}}. \end{aligned}$$

Ahora tenemos todo lo necesario para calcular la densidad de $T_1, T_2, \ldots, T_{n-1}$ dado T_n.

$$\begin{aligned} f_{T_1,dots\ldots,T_{n-1}|T_n}(u_1,\ldots,u_{n-1}|u_n) &= \frac{f_{T_1,T_2,\ldots,T_n}(u_1,u_2,\ldots,u_n)}{f_{T_n}(u_n)} \\ &= \frac{\lambda^n e^{-\lambda u_n} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}}}{\frac{\lambda^n {u_n}^{n-1} e^{-\lambda u_n}}{(n-1)!}} \\ &= \frac{\mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}}}{\frac{{u_n}^{n-1}}{(n-1)!}} \\ &= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}}. \end{aligned}$$

Y a partir de aquí podemos obtener la densidad de T_1 dado T_n de la siguiente manera

$$f_{T_1|T_n}(u_1|u_n)$$

$$= \int \cdots \int \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}} du_2 \ldots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \int \cdots \int \mathbb{1}_{\{0<u_1<u_2<\cdots<u_n\}} du_2 \ldots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_3} 1 du_2 \dots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_4} (u_3 - u_1) du_3 \dots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_5} \frac{1}{2}[(u_4 - u_1)^2 - (u_1 - u_1)^2] du_4 \dots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \frac{1}{2} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_5} (u_4 - u_1)^2 du_4 \dots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \frac{1}{2} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_6} \frac{1}{3}[(u_5 - u_1)^3 - (u_1 - u_1)^3] du_5 \dots du_{n-1}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \frac{1}{2\cdot 3} \int_{u_1}^{u_n} \dots \int_{u_1}^{u_6} (u_5 - u_1)^3 du_5 \dots du_{n-1}$$

$$\vdots$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}} \frac{1}{2\cdot 3 \cdots (n-2)} (u_n - u_1)^{n-2}$$

$$= \frac{(n-1)!}{{u_n}^{n-1}} \frac{(u_n - u_1)^{n-2}}{(n-2)!} \mathbb{1}_{\{0<u_1<u_n\}}$$

$$= (n-1) \frac{(u_n - u_1)^{n-2}}{{u_n}^{n-1}} \mathbb{1}_{\{0<u_1<u_n\}}$$

$$= (n-1) \frac{1}{u_n} \cdot \frac{(u_n - u_1)^{n-2}}{{u_n}^{n-2}} \mathbb{1}_{\{0<u_1<u_n\}}$$

$$= (n-1) \frac{1}{u_n} \cdot \left(1 - \frac{u_1}{u_n}\right)^{n-2} \mathbb{1}_{\{0<u_1<u_n\}}.$$

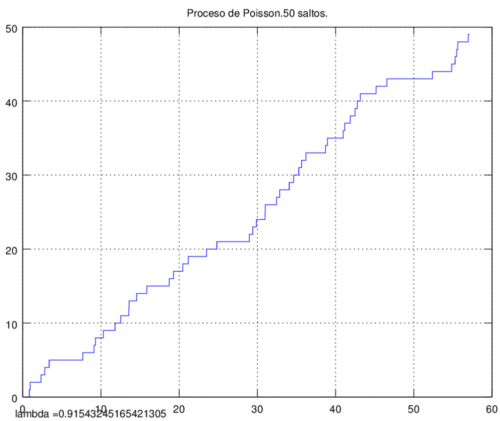
En resumen, la distribución de T_1 dado T_n se generaliza de la siguiente manera

$$f_{T_1|T_n}(u_1|u_n) = (n-1)\frac{1}{u_n} \cdot \left(1 - \frac{u_1}{u_n}\right)^{n-2} \mathbb{1}_{\{0<u_1<u_n\}}.$$

5.1.4. **Inciso (iv).** *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.*
Utilizando variables de distribución exponencial para los saltos:

```
clear;
#Se define el parametro lambda en [0, 1].
        lambda = rand();

#Numero de saltos.
        n = 50;

#Muestra para S_1, S_2, .....
        S = exprnd(1/lambda, n, 1);

#Aqui se calculan los tiempos
#de salto.
        for i=1:n
                T(i, 1) = sum(S(1:i));
        endfor

GraficaPoisson(T, lambda);
```

A continuación una gráfica obtenida de ejecutar este código:

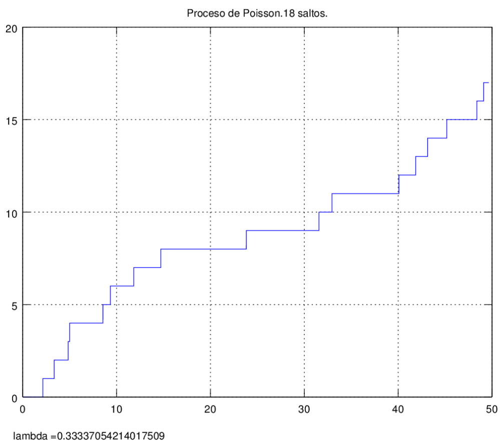


Se pide la simulación en un intervalo de tiempo fijo, no con una cantidad de saltos fija. La implementación anterior fue hecha intentando representar el proceso de Poisson de parámetro λ durante los 50 primeros saltos.

Utilizando variables de distribución Poisson para determinar la cantidad de saltos hasta el tiempo t:

```
clear;
#Se define el parametro lambda en [0, 1].
        lambda = rand();
        t = 50;

#Se calcula la cantidad de saltos al tiempo t
        N_t = poissrnd(lambda*50);
        T(N_t) = t;
        T(1)=0;
        for i = 1:(N_t-1)
                T(i+1) = t1_dado_tn_rnd(N_t - i + 1, t - T(i));
                T(i+1) = T(i+1) + T(i);
        endfor;
        T = T(2:end);

GraficaPoisson(T, lambda);
```

A continuación una gráfica obtenida de ejecutar este código:



A diferencia de la implementación anterior, esta vez intentamos representar el proceso de Poisson hasta un tiempo dado y no hasta un número de saltos dado.

¿Qué es t1 dado tn rnd? La idea es que condicionalmente a la cantidad de saltos, estos ocurren en puntos uniformes.
Como se puede apreciar, en esta última ejecución hasta el tiempo 50 únicamente hubo 18 saltos. A diferencia de que en la ejecución de la primera implementacción hubo 43 saltos para el tiempo 50. Cabe notar, que el parámetro lambda de la primera implementación era aproximadamente 0.915432, mientras que el de la segunda a penas aproximadamente 0.333371. Hay que recordar que los "holding times" tienen distribución exponencial, tienden a ser más largos conforme λ es más pequeño, y por lo tanto, entre más pequeño sea, habrá saltos con menos frecuencia.

Nota: Para poder implementar esta versión del algoritmo se necesitó implementar una función que se deduce del resultado de [5.1.3]. Esta parte está en el archivo con nombre `t1_dado_tn_rnd.m`. Esta función se usa para calcular los tiemos de salto y lo único que hace es simular una variable aleatoria con la densidad del resultado de [5.1.3].

5.2. Problema 5.2

Problema 5.2.1. Sea Ξ una medida de Poisson aleatoria en $(0,\infty)\times(0,\infty)$ cuya medida de intensidad ν está dada por $\nu(ds,dx) = \mathbf{1}_{0<x} C/x^{1+\alpha}\, ds\, dx$.

(i) [5.2.1] Determine los valores de α para los cuales $\int 1 \wedge x\, \nu(dx) < \infty$.

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

(ii) [5.2.2] Determine los valores de α para los cuales $X_t < \infty$ para toda $t \geq 0$ casi seguramente.

Nos restringiremos a dichos valores de α.

(iii) [5.2.3] Calcule $\mathbb{E}\big(e^{-\lambda X_t}\big)$ y pruebe que X_t tiene la misma distribución que $t^{1/\alpha} X_1$.

(iv) [5.2.4] Diga por qué el siguiente código en Octave simula la trayectoria aproximada del proceso X en el intervalo $[0,1]$.

```
C=1;
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:

5.2.1. Inciso (i). *Determine los valores de α para los cuales*

$\int 1 \wedge x\, \nu(dx) < \infty$.

Separemos la integral de manera conveniente:

$$\begin{aligned}
\int 1 \wedge x \, \nu(dx) &= \int (1 \wedge x) \mathbf{1}_{0<x} C/x^{1+\alpha} \, dx \\
&= \int_0^\infty (1 \wedge x) C/x^{1+\alpha} \, dx \\
&= \int_0^1 (1 \wedge x) C/x^{1+\alpha} \, dx + \int_1^\infty (1 \wedge x) C/x^{1+\alpha} \, dx \\
&= \int_0^1 x C/x^{1+\alpha} \, dx + \int_1^\infty C/x^{1+\alpha} \, dx \qquad (5.2.1.1)
\end{aligned}$$

Para que la integral que nos interesa sea finita es necesario y suficiente que los dos sumandos de la última parte de la ecuación lo sean.

Veamos qué ocurre en el primer termino si $\alpha = 1$.

$$\begin{aligned}
\int_0^1 x C/x^{1+\alpha} \, dx &= \int_0^1 x C/x^2 \, dx \\
&= \int_0^1 C/x \, dx \\
&= \lim_{t \to 0^+} \int_t^1 C/x \, dx \\
&= C \lim_{t \to 0^+} \int_t^1 1/x \, dx \\
&= C \lim_{t \to 0^+} [\log(x)]_t^1 \\
&= C \lim_{t \to 0^+} \log(1) - \log(t) \\
&= C \lim_{t \to 0^+} - \log(t) \\
&= C \cdot \infty.
\end{aligned}$$

Por lo tanto $\alpha \neq 1$. Veamos ahora qué ocurre con el primer sumando de (5.2.1.1) suponiendo que $\alpha \neq 1$.

$$\begin{aligned}
\int_0^1 x C/x^{1+\alpha} \, dx &= \int_0^1 C/x^\alpha \, dx \\
&= C \int_0^1 x^{-\alpha} \, dx
\end{aligned}$$

$$
\begin{aligned}
&= C\left[\frac{x^{(1-\alpha)}}{(1-\alpha)}\right]_0^1 \\
&= C\left(\frac{1^{(1-\alpha)}}{(1-\alpha)} - \lim_{t\to 0}\frac{t^{(1-\alpha)}}{(1-\alpha)}\right)
\end{aligned}
$$

El límte $\lim_{t\to 0}\frac{t^{(1-\alpha)}}{(1-\alpha)}$ exíste únicamente si $0 < 1-\alpha$, es decir, si $\alpha < 1$. Entonces basta con que $\alpha < 1$ para que el primer sumando de (5.2.1.1) sea finito. Ahora, si $\alpha = 0$ y analizamos el segundo sumando de (5.2.1.1) vemos que

$$
\begin{aligned}
\int_1^\infty C/x^{1+\alpha}\,dx &= C\int_1^\infty 1/x\,dx \\
&= C\lim_{t\to\infty}\int_1^t 1/x\,dx \\
&= C\lim_{t\to\infty}[\log(x)]_1^t \\
&= C\lim_{t\to\infty}\log(t) - \log(1) \\
&= C\lim_{t\to\infty}\log(t) \\
&= C\cdot\infty.
\end{aligned}
$$

Por lo cual necesitamos $\alpha \neq 0$. Supongamos entonces $\alpha \neq 0$ y analicemos el segundo sumando de (5.2.1.1):

$$
\begin{aligned}
\int_1^\infty C/x^{1+\alpha}\,dx &= C\int_1^\infty 1/x^{1+\alpha}\,dx \\
&= C\int_1^\infty x^{-1-\alpha}\,dx \\
&= C\lim_{t\to\infty}\int_1^t x^{-1-\alpha}\,dx \\
&= C\lim_{t\to\infty}\left[\frac{x^{-\alpha}}{-\alpha}\right]_1^t \\
&= C\left(\lim_{t\to\infty}\frac{t^{-\alpha}}{-\alpha} - \frac{1}{-\alpha}\right)
\end{aligned}
$$

Donde el límite $\lim_{t\to\infty}\frac{t^{-\alpha}}{-\alpha}$ existe únicamente si $-\alpha < 0$, es decir, si $0 < \alpha$.

Entonces podemos concluir que los valores de α para los que (5.2.1.1) es finito están en $(0, 1)$.

5.2.2. **Inciso (ii).** *Determine los valores de α para los cuales $X_t < \infty$ para toda $t \geq 0$ casi seguramente.*

Sean $f_t(s, x) = \mathbb{1}_{\{s \leq t\}} x$ y $X_t = \Xi f_t$.

De la segunda afirmación del teorema 4.6 de la versión de las notas que se incluye en este documento (ver [7.2.4]) sabemos que la variable aleatoria Ξf_t es casi seguramente finita o casi seguramente infinita de acuerdo a si la integral $\int 1 \wedge f_t \, d\nu$ es finita o no.

$$
\begin{aligned}
\int 1 \wedge f_t \, d\nu &= \int 1 \wedge \left(\mathbb{1}_{\{s \leq t\}} x\right) \, d\nu \\
&= \int \int 1 \wedge \left(\mathbb{1}_{\{s \leq t\}} x\right) \mathbf{1}_{0<x} C/x^{1+\alpha} \, ds \, dx \\
&= \int_0^\infty \int 1 \wedge \left(\mathbb{1}_{\{s \leq t\}} x\right) C/x^{1+\alpha} \, ds \, dx \\
&= \int_0^\infty \int_0^t (1 \wedge x) C/x^{1+\alpha} \, ds \, dx \\
&= \int_0^\infty t(1 \wedge x) C/x^{1+\alpha} \, dx \\
&= tC \left(\int_0^\infty (1 \wedge x) 1/x^{1+\alpha} \, dx \right) \\
&= tC \left(\int_0^1 1/x^\alpha \, dx + \int_1^\infty 1/x^{1+\alpha} \, dx \right).
\end{aligned}
$$

En [5.2.1] vimos que para que las integrales entre paréntesis tengan suma finita es neceario y suficiente que $\alpha \in (0, 1)$. Entonces basta que $\alpha \in (0, 1)$ para que $\int 1 \wedge f_t \, d\nu < \infty$.

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

La primera parte de la proposición 4.6 de la versión de las notas que se adjuntó en este documento (ver [7.2.4]) dice que si f es medible y no negativa entonces la integral de f con respecto de Ξ es una variable aleatoria y

$$\mathbb{E}\left(e^{-\Xi f}\right) = e^{-\int (1-e^{-f})\, d\nu}.$$

Usando esto mismo pero con $\Xi\lambda f_t = \lambda X_t$, tenemos

$$\mathbb{E}\left(e^{-\lambda X_t}\right) = e^{-\int (1-e^{-\lambda f_t})\, d\nu}. \tag{5.2.3.1}$$

.
Analicemos qué ocurre con $\int (1-e^{-\lambda f_t})\, d\nu$.

$$\begin{aligned}
\int (1-e^{-\lambda f_t})\, d\nu &= \int\int (1-e^{-\lambda f_t})\frac{C}{x^{1+\alpha}}\, ds\, dx \\
&= \int\int (1-e^{-\lambda \mathbb{1}_{\{s\leq t\}}x})\frac{C}{x^{1+\alpha}}\, ds\, dx \\
&= \int_0^\infty \int_0^t (1-e^{-\lambda x})\frac{C}{x^{1+\alpha}}\, ds\, dx \\
&= \int_0^\infty t(1-e^{-\lambda x})\frac{C}{x^{1+\alpha}}\, dx \\
&= tC\int_0^\infty (1-e^{-\lambda x})\frac{1}{x^{1+\alpha}}\, dx
\end{aligned}$$

Ahora reescribimos a $(1-e^{-\lambda x})$ como $\int_0^x \lambda e^{-\lambda y}\, dy$ y entonces la ecuación anterior se transforma en

$$\begin{aligned}
&= tC\int_0^\infty \left(\int_0^x \lambda e^{-\lambda y}\, dy\right)\frac{1}{x^{1+\alpha}}\, dx \\
&= tC\int_0^\infty \frac{\int_0^x \lambda e^{-\lambda y}}{x^{1+\alpha}}\, dy\, dx \\
&= tC\int_0^\infty \int_0^x \frac{\lambda e^{-\lambda y}}{x^{1+\alpha}}\lambda\, dy\, dx
\end{aligned}$$

La función que se encuentra en la integral es positiva, por lo tanto podemos aplicar el Teorema de Tonelli para intercambiar las integrales de la siguiente manera:

$$= tC \int_0^\infty \int_y^\infty \frac{\lambda e^{-\lambda y}}{x^{1+\alpha}}\, dx\, dy$$

Recordando que $\alpha \in (0,1)$, tenemos que $\int_y^\infty \frac{1}{x^{1+\alpha}}\, dx = \lim_{x\to\infty} \frac{x^{-\alpha}}{-\alpha} - \frac{y^{-\alpha}}{-\alpha} = -\frac{y^{-\alpha}}{-\alpha} = \frac{y^{-\alpha}}{\alpha}$. Sustituyendo esto tenemos

$$\begin{aligned}
&= tC \int_0^\infty \frac{y^{-\alpha}}{\alpha} \lambda e^{-\lambda y}\, dy \\
&= \frac{tC}{\alpha} \int_0^\infty y^{-\alpha} e^{-\lambda y} \lambda\, dy \\
&= \frac{tC\lambda^\alpha}{\alpha} \int_0^\infty (\lambda y)^{-\alpha} e^{-\lambda y}\, dy
\end{aligned}$$

Recordemos que $\int_0^\infty (\lambda y)^{-\alpha} e^{-\lambda y}\, dy$ se puede escribir como $\Gamma(1-\alpha)$ y entonces nuestra ecuación se puede escribir como

$$= \frac{tC\lambda^\alpha}{\alpha}\Gamma(1-\alpha).$$

Ahora, regresando a (5.2.3.1), tenemos que

$$\begin{aligned}
\mathbb{E}\left(e^{-\lambda X_t}\right) &= e^{-\int (1-e^{-\lambda f_t})\, d\nu} \\
&= e^{-\frac{tC\lambda^\alpha}{\alpha}\Gamma(1-\alpha)}
\end{aligned}$$

Ahora de la ecuación anterior, remplacemos t por 1 y entonces obtenemos

$$\mathbb{E}\left(e^{-\lambda X_1}\right) = e^{-\frac{C\lambda^\alpha}{\alpha}\Gamma(1-\alpha)}$$

De esto último, remplacemos ahora λ por $\lambda t^{1/\alpha}$

$$\begin{aligned}
\mathbb{E}\left(e^{-\lambda t^{1/\alpha} X_1}\right) &= e^{-\frac{C(\lambda t^{1/\alpha})^\alpha}{\alpha}\Gamma(1-\alpha)} \\
&= e^{-\frac{C\lambda^\alpha t}{\alpha}\Gamma(1-\alpha)}
\end{aligned}$$

De donde podemos observar fácilmente que

$$\mathbb{E}\left(e^{-\lambda X_t}\right) = \mathbb{E}\left(e^{-\lambda t^{1/\alpha} X_1}\right).$$

Lo que acabamos de escribir no son mas que las funciones generadoras de momentos de X_t y $t^{1/\alpha}X_1$ evaluadas en $-\lambda$. Lo que acabamos de demostrar es que sus funciones generadoras de momentos son idénticas para cada λ y por lo tanto, sus distribuciones son iguales.

5.2.4. **Inciso (iv).** *Diga por qué el siguiente código en Octave simula la trayectoria aproximada del proceso X en el intervalo* $[0,1]$.

```
C=1;
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)
```

Primero se está definiendo el valor de la constante `C` como `1`.

`alpha=1/2;` está definiendo un valor para nuestra variable α (entre 0 y 1, como tendría que ser para que las variables las variables X_t sean finitas casi seguramente según lo demostrado en [5.2.2]).

`lambda=C/e^alpha;` define un valor para λ.

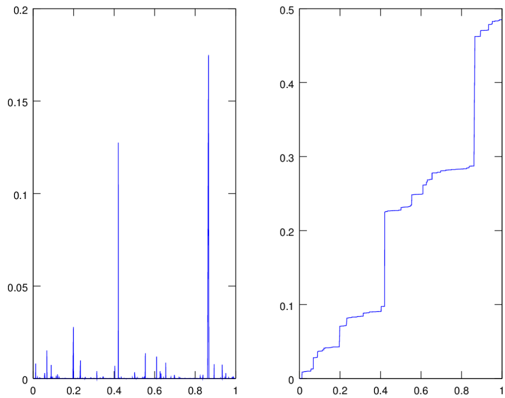
`T=1;` se define así porque sólo nos interesa saber que ocurre con nuestro proceso hasta el tiempo 1.

`N=poissrnd(lambda*T, 1);` está estimando la cantidad de saltos al tiempo `T`.

La siguiente parte es la interesante. `dx=e./(rand(N,1).^(1/(alpha)));` está estimando los incrementos de X. Se está utilizando el resultado de [5.2.3]. Donde `rand(N,1).^(1/(alpha))` no es otra cosa más que $t^{1/\alpha}$. Se está suponiendo que los incrementos "infinitesimales" de $t^{1/\alpha}X_1$ son tan chiquitos como `e = 0.000001`. Utilizando esto, se estiman los incrementos "infinitesimales" de X.

El resto del código es únicamente para propósitos de graficación.

A continuación una gráfica que resultó de ejecutar el código.



5.3. Problema 5.3

Problema 5.3.1. .

(i) [5.3.1] 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 $\mathscr{F}^X_t = \sigma(X_s : s \geq 0)$.
(ii) [5.3.2] 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 proceso Poisson compuesto con salto ξ_i

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

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

(iii) [5.3.3] Pruebe que casi seguramente las trayectorias de N son no-decrecientes.
(iv) [5.3.4] 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 $\lambda \cdot \mathrm{Leb}$.
(v) [5.3.5] Concluya que N es un proceso de Poisson de intensidad λ.

Demostración:

5.3.1. **Inciso (i).** *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 $\mathscr{F}^X_t = \sigma(X_s : s \leq t)$.*

Recordemos que tener incrementos independientes significa que siempre que $0 \leq r_1 < r_2 < r_3 < \cdots < r_m$ se tiene que $X_{r_1} - X_{r_2}, X_{r_2} - X_{r_3}, \ldots, X_{r_m} - X_{t_{m-1}}$ son variables aleatorias independientes.

Veamos quien es la σ-álgebra generada por los incrementos de X hasta t. En otras palabaras, la σ-álgebra generada por conjuntos de la forma

$$\bigcap_{1\leq i\leq n} \{X_{s_i} - X_{s_{i-1}} \in B_i\} \tag{5.3.1.1}$$

Donde $0 \leq s_1 < s_2 < \cdots < s_n \leq t$, $B_i \in \mathscr{B}_{\mathbb{R}}$ y, por comodida de notación, $X_{s_0} = 0$.

Notemos primero que estos conjuntos son todos $\mathscr{F}^X_t$ medibles. También notemos que los conjuntos $\{X_s \in B : B \in \mathscr{B}_{\mathbb{R}}\}$ con $0 \leq s \leq t$ también son de esta forma. Basta con escoger $s_1 = s$, $B_1 = B$ y $B_i = \mathbb{R}$ para $i \geq 2$.

Lo que acabamos de mostrar es que la σ-álgebra generada por los conjuntos definidos en (5.3.1.1) es igual a $\mathscr{F}_t^X$. Llamemos P a la clase de los conjuntos de esta forma. Es fácil ver que P es un π-sistema. Es no vacío ($\mathbb{R}$ y $\emptyset$ pertenecen a él) y es fácil ver que la intersección de dos de ellos está otra vez en P.

Ahora analicemos a la σ-álgebra generada por los incrementos de X^t. Llamemos P' a la clase de los conjuntos de la forma

$$\bigcap_{1\leq i\leq n} \{X_{t+r_i} - X_{t+r_{i-1}} \in B_i\}$$

Donde $0 \leq r_1 < r_2 < \cdots < r_n$, $B_i \in \mathscr{B}_\mathbb{R}$ y, por comodida de notación $r_0 = 0$. Otra vez, es fácil notar que P' es un pi-sistema. Es no vacío ($\mathbb{R}$ y $\emptyset$ pertenecen a él) y es fácil ver que la intersección de dos de ellos está otra vez en P'.

P' en la forma que lo construimos, es el álgebra de los incrementos de X^t, por lo tanto, la σ-álgebra que genera es la σ-álgebra de los incrementos de X^t.

La idea ahora es encontrar un sistema de Dynkin que contenga a P y que sea independiente de P'. De esta manera, utilizando el Lema de clases de Dynkin podemos extender la propieadad de ser independiente de P'.

$$D = \{S \in \sigma(P) : \mathbb{P}(S \cap C') = \mathbb{P}(S)\mathbb{P}(C')\ \forall C' \in P'\}.$$

Veamos que $P \subset D$. Para ello, sea $C \in P$ con

$$C = \bigcap_{1\leq i\leq n} \{X_{s_i} - X_{s_{i-1}} \in B_i\}.$$

Y sea $C' \in P$, con

$$C' = \bigcap_{1\leq i'\leq n'} \{X_{t+r'_i} - X_{t+r_{i'-1}} \in B'_i\}$$

Entonces

$$\mathbb{P}(C \cap C') =$$

$$\mathbb{P}\left(\left[\bigcap_{1\leq i\leq n} \{X_{s_i} - X_{s_{i-1}} \in B_i\}\right] \cap \left[\bigcap_{1\leq i'\leq n'} \{X_{t+r'_i} - X_{t+r_{i'-1}} \in B'_i\}\right]\right) =$$

$$\mathbb{P}\left(\left[\bigcap_{1\leq i\leq n}\{X_{s_i}-X_{s_{i-1}}\in B_i\}\right]\cap\{X_t-X_{s_n}\in\mathbb{R}\}\cap\left[\bigcap_{1\leq i'\leq n'}\{X_{t+r'_i}-X_{t+r_{i'-1}}\in B'_i\}\right]\right)=$$

(Se está intersectando con un conjunto que es igual a $\mathbb{R}$, por eso la igualdad se preserva)

$$\mathbb{P}\left(\left[\bigcap_{1\leq i\leq n}\{X_{s_i}-X_{s_{i-1}}\in B_i\}\right]\right)\cdot\mathbb{P}\left(\{X_t-X_{s_n}\in\mathbb{R}\}\right)\cdot\mathbb{P}\left(\left[\bigcap_{1\leq i'\leq n'}\{X_{t+r'_i}-X_{t+r_{i'-1}}\in B'_i\}\right]\right)=$$

(Se utilizó la hipótesis de que X tiene incrementos independientes)

$$\mathbb{P}\left(\left[\bigcap_{1\leq i\leq n}\{X_{s_i}-X_{s_{i-1}}\in B_i\}\right]\right)\cdot\mathbb{P}\left(\left[\bigcap_{1\leq i'\leq n'}\{X_{t+r'_i}-X_{t+r_{i'-1}}\in B'_i\}\right]\right)=$$

(El termino que desapareció era la probabilidad del total, es decir 1)

$\mathbb{P}(C)\mathbb{P}(C')$.

Con esto hemos demostrado que $C\in D$ y por lo tanto $P\subset D$. Para ver que D es sistema de Dynkin, notemos primero que $\mathbb{R}\in D$ (el total siempre es un conjunto independiente de cualquier otro).

5.3.2. **Inciso (ii).** *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 proceso Poisson compuesto con salto ξ_i*

$$\mathbb{E}\big(e^{iuX_t}\big) = e^{-\lambda t(1-\psi(u))} \quad \textit{donde} \quad \psi(u) = \mathbb{E}\big(e^{iu\xi_1}\big).$$

Sea N_t un proceso de Poisson de intensidad λ. Es decir, que la distribución de N_t tiene distribución de Poisson de parámetro λt. Entonces:

$$\begin{aligned}
\mathbb{E}(N_t) &= \sum_{i\in\mathbb{N}} i\mathbb{P}(N_t = i) \\
&= \sum_{i\in\mathbb{N}} ie^{-\lambda t}\frac{(\lambda t)^i}{i!} \\
&= e^{-\lambda t}\sum_{i\in\mathbb{N}} i\frac{(\lambda t)^i}{i!} \\
&= e^{-\lambda t}(\lambda t)\sum_{i\in\mathbb{N}} \frac{(\lambda t)^{i-1}}{(i-1)!} \\
&= e^{-\lambda t}(\lambda t)e^{\lambda t} \\
&= \lambda t.
\end{aligned}$$

Ahora

$$\begin{aligned}
\mathbb{E}(N_t^2) &= \sum_{i\in\mathbb{N}} i^2\mathbb{P}(N_t = i) \\
&= e^{-\lambda t}\sum_{i\in\mathbb{N}} i^2\frac{(\lambda t)^i}{i!} \\
&= e^{-\lambda t}\sum_{i\in\mathbb{N}} i\frac{(\lambda t)^i}{(i-1)!} \\
&= e^{-\lambda t}(\lambda t)\sum_{i\in\mathbb{N}} i\frac{(\lambda t)^{i-1}}{(i-1)!} \\
&= e^{-\lambda t}(\lambda t)\sum_{i\in\mathbb{N}} (i-1+1)\frac{(\lambda t)^{i-1}}{(i-1)!} \\
&= e^{-\lambda t}(\lambda t)\left(\sum_{i\in\mathbb{N}} (i-1)\frac{(\lambda t)^{i-1}}{(i-1)!} + \sum_{i\in\mathbb{N}} \frac{(\lambda t)^{i-1}}{(i-1)!}\right)
\end{aligned}$$

$$
\begin{aligned}
&= e^{-\lambda t}(\lambda t)\left(\sum_{i\in\mathbb{N}}(i-1)\frac{(\lambda t)^{i-1}}{(i-1)!}+e^{\lambda t}\right)\\
&= \left((\lambda t)e^{-\lambda t}\sum_{i\in\mathbb{N}}(i-1)\frac{(\lambda t)^{i-1}}{(i-1)!}+e^{-\lambda t}(\lambda t)e^{\lambda t}\right)\\
&= (\lambda t)\mathbb{E}(N_t)+(\lambda t)\\
&= (\lambda t)^2+(\lambda t).
\end{aligned}
$$

Por lo tanto

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

Sea ahora X un proceso de Poisson compuesto. Es decir

$$X_t=\sum_{i\leq N_t}\xi_i.$$

Con ξ_i variables aleatorias independientes, idénticamente distribuidas e independientes del proceso de Poisson N_t. Entonces

$$
\begin{aligned}
\mathbb{E}(X_t) &= \mathbb{E}\left(\sum_{i\leq N_t}\xi_i\right)\\
&= \mathbb{E}\left(\sum_{i\in\mathbb{N}}\xi_i\mathbb{1}_{\{T_i\leq t\}}\right)\\
&= \left(\sum_{i\in\mathbb{N}}\mathbb{E}(\xi_i\mathbb{1}_{\{T_i\leq t\}})\right)\\
&\quad \text{(por teorema de la convergencia monótona)}\\
&= \left(\sum_{i\in\mathbb{N}}\mathbb{E}(\xi_i)\mathbb{E}(\mathbb{1}_{\{T_i\leq t\}})\right)\\
&\quad \text{(por independencia de las } \xi \text{ del proceso de Poisson)}\\
&= \left(\sum_{i\in\mathbb{N}}\mathbb{E}(\xi_1)\mathbb{E}(\mathbb{1}_{\{T_i\leq t\}})\right)\\
&\quad \text{(por que las } \xi \text{ son idénticamente distribuidas)}\\
&= \mathbb{E}(\xi_1)\left(\sum_{i\in\mathbb{N}}\mathbb{E}(\mathbb{1}_{\{T_i\leq t\}})\right)
\end{aligned}
$$

$$= \mathbb{E}(\xi_1)\mathbb{E}(N_t)$$
$$= \mathbb{E}(\xi_1)(\lambda t).$$

[ahora más cuentas].

$$\begin{aligned}
\mathbb{E}(X_t^2) &= \mathbb{E}(\mathbb{E}(X_t^2 \Big| N_t)) \\
&= \mathbb{E}\left(\mathbb{E}\left[\left(\sum_{i \le N_t} \xi_i\right)^2 \Bigg| N_t\right]\right) \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \mathbb{E}\left[\left(\sum_{i \le N_t} \xi_i\right)^2 \Bigg| N_t = j\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \mathbb{E}\left[\left(\sum_{i \le j} \xi_i\right)^2\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&\quad \text{(otra vez por independenia de las } \xi \text{ del proceso de Poisson)} \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \mathbb{E}\left[\sum_{i, i' \le j} \xi_i \xi_{i'}\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \left[\sum_{i, i' \le j} \mathbb{E}(\xi_i \xi_{i'})\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \left[\sum_{i \le j} \mathbb{E}(\xi_i \xi_i) + \sum_{i < i' \le j} \mathbb{E}(\xi_i \xi_{i'})\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \left[\sum_{i \le j} \mathbb{E}(\xi_i^2) + \sum_{i < i' \le j} \mathbb{E}(\xi_i)\mathbb{E}(\xi_{i'})\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&\quad \text{(por independencia de los } \xi_i \text{)} \\
&= \mathbb{E}\left(\sum_{j \in \mathbb{N}} \left[\sum_{i \le j} \mathbb{E}(\xi_1^2) + \sum_{i < i' \le j} \mathbb{E}(\xi_1)\mathbb{E}(\xi_1)\right] \mathbb{1}_{\{N_t = j\}}\right) \\
&\quad \text{(porque las } \xi_i \text{ son indénticamente distribuidas)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left(\sum_{j\in\mathbb{N}}\left[j\mathbb{E}(\xi_1^2)+j(j-1)\mathbb{E}(\xi_1)^2\right]\mathbb{1}_{\{N_t=j\}}\right)\\
&= \mathbb{E}\left(\sum_{j\in\mathbb{N}} j\left[(\mathbb{E}(\xi_1^2)+(j-1)\mathbb{E}(\xi_1)^2)\right]\mathbb{1}_{\{N_t=j\}}\right)\\
&= \mathbb{E}\left(\sum_{j\in\mathbb{N}} j\left[j\mathbb{E}(\xi_1)^2+(\mathbb{E}(\xi_1^2)-\mathbb{E}(\xi_1)^2)\right]\mathbb{1}_{\{N_t=j\}}\right)\\
&= \mathbb{E}\left(\sum_{j\in\mathbb{N}} j\left[j\mathbb{E}(\xi_1)^2+Var(\xi_1)\right]\mathbb{1}_{\{N_t=j\}}\right)\\
&= \mathbb{E}\left(\sum_{j\in\mathbb{N}}\left[j^2\mathbb{E}(\xi_1)^2+jVar(\xi_1))\right]\mathbb{1}_{\{N_t=j\}}\right)\\
&= \mathbb{E}\left(N_t^2\mathbb{E}(\xi_1)^2+N_tVar(\xi_1)\right)\\
&= \mathbb{E}(N_t^2)\mathbb{E}(\xi_1)^2+\mathbb{E}(N_t)Var(\xi_1)\\
&= ((\lambda t)^2+(\lambda t))\mathbb{E}(\xi_1)^2+(\lambda t)Var(\xi_1)\\
&= ((\lambda t)^2+(\lambda t))\mathbb{E}(\xi_1)^2+(\lambda t)(\mathbb{E}(\xi_1^2)-\mathbb{E}(\xi_1)^2)\\
&= (\lambda t)^2\mathbb{E}(\xi_1)^2+(\lambda t)\mathbb{E}(\xi_1)^2+(\lambda t)\mathbb{E}(\xi_1^2)-(\lambda t)\mathbb{E}(\xi_1)^2\\
&= (\lambda t)^2\mathbb{E}(\xi_1)^2+(\lambda t)\mathbb{E}(\xi_1^2)
\end{aligned}
$$

Entonces

$$
\begin{aligned}
Var(X_t) &= \mathbb{E}(X_t^2)-\mathbb{E}(X_t)^2\\
&= (\lambda t)^2\mathbb{E}(\xi_1)^2+(\lambda t)\mathbb{E}(\xi_1^2)-\mathbb{E}(\xi_1)^2(\lambda t)^2\\
&= (\lambda t)\mathbb{E}(\xi_1^2).
\end{aligned}
$$

Para la última parte, sólo haran falta más cuentas

$$
\begin{aligned}
\mathbb{E}(e^{iuX_t}) &= \mathbb{E}\left(e^{iu\sum_{j\le N_t}\xi_j}\right)\\
&= \mathbb{E}\left(\mathbb{E}\left[e^{iu\sum_{j\le N_t}\xi_j}\middle| N_t\right]\right)\\
&= \mathbb{E}\left(\sum_{k\in\mathbb{N}}\mathbb{1}_{N_t=k}\mathbb{E}\left[e^{iu\sum_{j\le k}\xi_j}\middle| N_t=k\right]\right)
\end{aligned}
$$

$$= \mathbb{E}\left(\sum_{k\in\mathbb{N}} \mathbb{1}_{N_t=k}\mathbb{E}\left[e^{iu\sum_{j\leq k}\xi_j}\right]\right)$$

(otra vez por independenia de las ξ del proceso de Poisson)

$$= \mathbb{E}\left(\sum_{k\in\mathbb{N}} \mathbb{1}_{N_t=k}\mathbb{E}\left[e^{iu\xi_1}\right]^k\right)$$

(por propiedades de la función característica y por

que las ξ_i son independientes e idénticamente distribuidas)

$$= \mathbb{E}\left(\mathbb{E}\left[e^{iu\xi_1}\right]^{N_t}\right)$$

$$= \mathbb{E}\left(\psi(u)^{N_t}\right)$$

$$= \sum_{j\in\mathbb{N}} \psi(u)^j e^{-\lambda t}\frac{(\lambda t)^j}{j!}$$

$$= e^{-\lambda t}\sum_{j\in\mathbb{N}} \frac{(\psi(u)\lambda t)^j}{j!}$$

$$= e^{-\lambda t}e^{\psi(u)\lambda t}$$

$$= e^{\psi(u)\lambda t-\lambda t}$$

$$= e^{\lambda t(\psi(u)-1)}$$

$$= e^{-\lambda t(1-\psi(u))}.$$

Con lo que termina la demostración.

5.3.3. **Inciso (iii).** *Pruebe que casi seguramente las trayectorias de N son no-decrecientes.*

Por ser un proceso de Lévy, sabemos que sus incrementos son estacionarios. Es decir, que para $s, t \geq 0$. $N_{t+s} - N_t$ tiene la misma distirbución que N_s. La cual tiene distribución $Poisson(\lambda s)$ y por lo tanto toma valores enteros, mayores o iguales a 0, con probabilidad 1. Es decir, casi seguramente N_{t+s} es mayor o igual a N_t. Ahora, recordemos que las N_t casi seguramente toman valores en los entereos y por otra parte, los procesos de Lévy tienen trayectorias continuas por la derecha y con límites por la izquierda (cádlág).

La continuidad por la derecha y que las alturas son enteras nos definen trayectorias constantes a pedazos. Es decir que si ocurriera que $N_t > N_{t+s}$, ocurriría para toda una vecindad (por la derecha) de $t + s$, pero, como ya vimos, este suceso tiene probabilidad 0.

5.3.4. **Inciso (iv).** *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 $\lambda \cdot Leb$.*

Sea

$$M = \{A \in \mathscr{B}_{\mathbb{R}+} : \Xi(A) \sim Poisson(\lambda Leb(A))\}$$

Sean $A_1, A_2, \dots \in M$ tales que $A_i \subset A_{i+1}$ y sea $A = \bigcup A_i$. Por ser Ξ medida aleatoria, tenemos que $\Xi(A_i) \longrightarrow \Xi(A)$. Y por lo tanto $\lim_{n\to\infty} \mathbb{P}(\Xi(A_i) = n) = \mathbb{P}(\Xi(A) = n)$.

Como $A_1, A_2, \dots inM$, podemos escribir lo siguiente

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(\Xi(A_i) = n) &= \lim_{n\to\infty} e^{-\lambda Leb(A_i)} \frac{(\lambda Leb(A_i))^n}{n!} \\
&= e^{-\lambda Leb(A)} \frac{(\lambda Leb(A))^n}{n!} \\
&\text{(pues } A_i \to A \text{ y } Leb(A_i) \to Leb(A)).
\end{aligned}$$

Por lo tanto $A \in M$ y esto demuestra que M es una clase monótona.

Ahroa sea

$$P = \left\{ A \in \mathscr{B}_{\mathbb{R}+} : A = \bigcup_{i\leq n} (a_i, b_i] \text{ tal que } b_i < a_{i+1} \right\}$$

Es fácil notar que P es un álgebra y que $\sigma(P) = \mathscr{B}_{\mathbb{R}+}$.

Sea $A \in P$. Es decir, A tiene la forma

$$A = \bigcup_{i\leq n} (a_i, b_i]$$

con $b_i < a_{i+1}$. Nombremos $A_i = (a_i, b_i]$. Entonces

$$\begin{aligned}
\Xi(A) &= \Xi\left(\bigcup (a_i, b_i]\right) \\
&= \sum_{i\leq n} \Xi((a_i, b_i]) \\
&= \sum_{i\leq n} \Xi([0, b_i]) - \Xi([0, a_i]) \\
&= \sum_{i\leq n} N_{b_i} - N_{a_i}
\end{aligned}$$

$$\sim \sum_{i \leq n} N_{b_i - a_i}$$

Los $N_{b_i-a_i}$ tiene distribución $Poisson(\lambda(b_i - a_i))$ y son independientes. Por lo tanto la distribución de su suma es $Poisson\left(\lambda \sum(b_i - a_i)\right)$. Pero $Leb(A) = \sum(b_i - a_i)$. Así que $\Xi(A)$ tiene distribución $Poisson(\lambda Leb(A))$ y por lo tanto pertenece a M.

Como P es un álgebra y M es una clase monótona que la contiene, el teorema de clases monotonas nos dice que $\sigma(P) \subset M$. Pero como $\sigma(P) = \mathscr{B}_{\mathbb{R}^+}$ y $M \subset \mathscr{B}_{\mathbb{R}^+}$, tenemos que $\sigma(P) = M$ y por lo tanto $\Xi(A) \sim Poisson(\lambda Leb(A))$ para todo $A \in \mathscr{B}_{\mathbb{R}^+}$.

Con esto hemos demostrado la primera condición necesaria para ser variable aleatoria de Poisson de intensidad λLeb. La condición que nos falta es que si $A_1, A_2, \ldots, A_n \in \mathscr{B}_{\mathbb{R}^+}$ son ajenos dos a dos, entonces las variables aleatorias $\Xi(A_1), \Xi(A_2), \ldots, \Xi(A_n)$ son independientes. Esta demostración es idéntica a la segunda parte de la demostración de la propocisión 4.5 de las notas de la clase (véase [7.2.4]).

5.3.5. **Inciso (v).** *Concluya que N es un proceso de Poisson de intensidad λ.*

Tenemos que N es un proceso de contéo. ¿Por qué? La idea era utilizar un resultado de las notas que afirmaba que las medidas aleatorias de Poisson son medidas atómicas y que cada átomo tiene masa unitaria.Y por hipótesis es un proceso de Lévy. El teorema 4.1 dice que un proceso de contéo es un proceso de Poisson si y sólo si es un proceso de Lévy (véase [7.2.4]).

5.4. Problema 5.4

Problema 5.4.1. 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

$$Q(i,j) = \begin{cases} -\lambda & j = i \\ \lambda & j = i+1 \\ 0 & j \neq i, i+1 \end{cases}. \tag{5.4.0.1}$$

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.

Sean $i \leq j \in \mathbb{N}$. Por definición

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

(si $j < i$, P_r es cero). Sean ahora, $t, s \in \mathbb{R}^+$.

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

Ahora veamos que ocurre con la entrada (i, j) de la matriz $P_t P_s$.

$$\begin{aligned}
(P_t P_s)(i,j) &= \sum_{k\geq 0} P_t(i,k)P_t(k,j) \\
&= \sum_{i\leq k\leq j} P_t(i,k)P_t(k,j) \\
&\quad \text{(Pues } P_r(i,j) = 0 \text{ siempre que } j < i\text{)} \\
&= \sum_{i\leq k\leq j} e^{-\lambda t}\frac{(\lambda t)^{k-i}}{(k-i)!}e^{-\lambda s}\frac{(\lambda s)^{j-k}}{(j-k)!} \\
&= e^{-\lambda(t+s)} \sum_{i\leq k\leq j} \frac{(\lambda)^{j-i}t^{k-i}s^{j-k}}{(j-k)!(k-i)!} \\
&= e^{-\lambda(t+s)}\lambda^{j-i} \sum_{i\leq k\leq j} \frac{t^{k-i}s^{j-k}}{(j-k)!(k-i)!}
\end{aligned}$$

Si definimos $k' = k - i$ (nótese que $0 \leq k' \leq j - i$) y hacemos la sustitución teniendo cuidado con los límites de las sumas obtenemos

$$\begin{aligned}
(P_t P_s)(i,j) &= e^{-\lambda(t+s)}\lambda^{j-i} \sum_{i\leq k\leq j} \frac{t^{k-i}s^{j-k}}{(j-k)!(k-i)!} \\
&= e^{-\lambda(t+s)}\lambda^{j-i} \sum_{k'\leq j-i} \frac{t^{k'}s^{j-i-k'}}{(j-i-k')!(k')!}
\end{aligned}$$

Que es equivalente a lo que habíamos conseguido antes y por lo tanto $P_{t+s} = P_t P_s$. Ahora una prueba utilizando los argumentos de que los incrementos son estacionarios e independientes.

$$\begin{aligned}
P_{t+s}(i,j) &= \mathbb{P}(N_{t+s} = i - j) \\
&= \sum_{0\leq k\leq j-i} \mathbb{P}(N_{t+s} = j - i, N_s = k) \\
&= \sum_{0\leq k\leq j-i} \mathbb{P}(N_{t+s} - N_s = j - i - k, N_s = k)
\end{aligned}$$

Hacemos $k' = j - k$ y entonces podemos escribir

$$P_{t+s}(i,j) = \sum_{i\leq k'\leq j} \mathbb{P}(N_{t+s} - N_s = j - i - k, N_s = k)$$

$$
\begin{aligned}
&= \sum_{i \le k' \le j} \mathbb{P}(N_{t+s} - N_s = k' - i, N_s = j - k') \\
&= \sum_{i \le k' \le j} \mathbb{P}(N_{t+s} - N_s = k' - i)\mathbb{P}(N_s = j - k') \\
&\quad \text{(Esto úlitmo por la independencia de los incrementos)} \\
&= \sum_{i \le k' \le j} \mathbb{P}(N_t = k' - i)\mathbb{P}(N_s = j - k') \\
&\quad \text{(Esto último porque los incrementos son estacionarios)} \\
&= \sum_{i \le k' \le j} P_t(i,k)P_s(k',j)
\end{aligned}
$$

Por otro lado

$$
(P_t P_s)(i,j) = \sum_{i \le k \le j} i \le k \le j P_t(i,k)P_s(k',j)
$$

Dado que ests son equivalentes, tenemos que $P_{t+s} = P_t P_s$, que es la propiedad de semigrupo que describen las ecuaciones de Chapman-Kolmogorov.

Ahora las ecuaciones "forward" y "backward".

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

Derivando con respecto a r obtenemos:

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

Ahora, recordando (5.4.0.1), veamos que

$$
\begin{aligned}
QP_r(i,j) &= \sum_{k \in \mathbb{N}} Q(i,k)P_r(k,j) \\
&= Q(i,i)P_r(i,j) + Q(i,i+1)P_r(i+1,j) \\
&= Q(i,i)P_r(i,j) + Q(i,i+1)P_r(i+1,j) \\
&= -\lambda P_r(i,j) + \lambda P_r(i+1,j).
\end{aligned}
$$

Con lo que queda demostrado que $\frac{d}{dr}P_r(i,j) = QP_r$. De manera similar

$$\begin{aligned}
P_rQ(i,j) &= \sum_{k\in\mathbb{N}} P_r(i,k)Q(k,j) \\
&= P_r(i,j)Q(j,j) + P_r(i,j-1)Q(j-1,j) \\
&= P_r(i,j)(-\lambda) + P_r(i,j-1)(\lambda) \\
&= P_r(i,j)(-\lambda) + P_r(i+1,j)(\lambda) \\
&\quad \text{(porque } j-1-i = j-(i+1)\text{)} \\
&= -\lambda P_r(i,j) + \lambda P_r(i+1,j)
\end{aligned}$$

Con lo que queda demostrado que $\frac{d}{dr}P_r(i,j) = QP_r = P_rQ$.

5.5. Problema 5.5

Problema 5.5.1 (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

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

(i) [5.5.1] 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.)

(ii) [5.5.2] ¿Cuál es la probabilidad de que ambos generadores estén trabajando al tiempo t si sólo uno trabaja al tiempo cero?

(iii) [5.5.3] 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.

(iv) [5.5.4] 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:

5.5.1. **Inciso (i).** *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.)*

Como no hay estados absorbentes, podemos utilizar la fórmula

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

donde $c(i)$ es la suma de las entradas positivas de la i-ésima fila de Q y $\delta_{i,j} = 1$ si $i = j$ y $\delta_{i,j} = 0$ en cualquier otro caso.

Después de algunas cuentas se obtiene:

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

De P se puede todos los estados son alcanzables desde cualquier otro (de los estados 0 y 2 se puede llegar al 1, y del 1 a cualquiera de los otros). Es decir, es irreducible.

Existe un teorema que nos asegura que por ser irreducible y de estados finitos, los estados son positivos recurrentes y existe una distribución invariante. También existe otro teorema que nos asegura que por tener estados positivos recurrentes, la distribución invariante es única.

Entonces, buscar una distribución invariante se reduce a resolver el sistema de ecuaciones

$$\begin{aligned} \frac{1}{7}x_2 &= x_1 \\ x_1 + x_3 &= x_2 \\ \frac{6}{7}x_2 &= x_3 \end{aligned}$$

Donde $x_1 + x_2 + x_3 = 1$. Después de algunas cuentas, se llega a que la solución del sistema es $x_1 = \frac{1}{14}$, $x_2 = \frac{7}{14}$, $x_3 = \frac{6}{14}$.

Para calcular explícitamente las potencias de P, primero encontramos el polinomio característico de P.

$$p(x) = -\lambda(\lambda + 1)(\lambda - 1).$$

De donde, los valores propios son $0, -1, 1$. Después de resolver los sistemas de ecuaciones correspondientes (lo cual son muchas cuentas) se obtiene que los vectores propios correspondientes son $(-6, 0, 1), (1, 1, 1), (1, -1, 1)$.
Por lo tanto, P se puede diagonalizar como

$$A^{-1}PA = D$$

Donde

$$A = \begin{pmatrix} -6 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}.$$

y

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

y Entonces

$$P^n = AD^n A^{-1}$$

Ahora notemos que

$$D^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1^n \end{pmatrix}.$$

Por lo tanto

$$P^{2n+1} = AD^{2n+1}A^{-1} = ADA^{-1} = P$$

Y

$$P^{2n} = AD^2A^{-1} = \begin{pmatrix} \frac{1}{7} & 0 & \frac{6}{7} \\ 0 & 1 & 0 \\ \frac{1}{7} & 0 & \frac{6}{7} \end{pmatrix}.$$

5.5.2. **Inciso (ii).** *¿Cuál es la probabilidad de que ambos generadores estén trabajando al tiempo t si sólo uno trabaja al tiempo cero?*

Ahora la estrategia consiste en diagonalizar Q. Su polinomio característico es

$$q(x) = \lambda(\lambda + 5)(\lambda + 10).$$

Por lo tanto sus valores propios son $0, -5, -10$. Después de resolver los sistemas de ecuaciones correspondientes se obtiene que los vectores propios correspondientes son $(1, 1, 1), (18, 3, -2), (-6, 4, -1)$.
Con lo que definimos a la matriz

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

la cual tiene inversa

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

Ahora definimos

$$D = A^{-1}QA = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & -10 \end{pmatrix}.$$

Para calcular la probablidad de que 2 generadores estén funcionando al tiempo t dado que se empezó con 1, necesitamos encontrar la entrada $(2, 3)$ de P_t (recordemos que el estado representado en el renglón 1 significa ningún generador está funcionando, el representado por el renglón 2 significa que un generador está funcionando y el representado por el 3 significa que ambos generadores están funcionando).

Para ello utilizamos que $P_t = e^tQ = e^{t(ADA^{-1})} = Ae^{tD}A^{-1}$ y que

$$e^{tD} = \begin{pmatrix} e^{0t} & 0 & 0 \\ 0 & e^{-5t} & 0 \\ 0 & 0 & e^{-10t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-5t} & 0 \\ 0 & 0 & e^{-10t} \end{pmatrix}.$$

Entonces,

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

5.5.3. **Inciso (iii).** *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.*

Este problema es similar al anterior. En el inciso anterior $P_t(2,3)$ representa la probabilidad de pasar del estado 2 al 3 en t, pero en la cadena cualquier estado es alcanzable desde cualquier otro, es decir que $P_t(2,3)$ concidera la posibilidad de que se haya pasado varias veces por 3 antes de que terminara t. Lo único que hay que suponer entonces, es que el estado "dos generadores están funcionando" es absorbente y entonces lo que buscamos es la $\mathbb{P}(\rho_2 \leq t)$.

Entonces, sustituimos el tercer renglón de Q por puros ceros para convertir al estado "dos generadores están funcionando" en un estado absorvente La tarea no está en portugués ¿o si? y dejar a todas las distribuciones que no implican salir del estado "dos generadores están funcionando" sin cambios.

Definimos entonces Q' como

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

De ahora en adelante, la el problema es idéntico al del inciso anterior. Buscamos el polinomio característico de Q'.

$$q'(\lambda) = \lambda(4+\lambda)(9+\lambda)$$

Vemos que sus valores propios son $0, -4y-9$. Después de cuentas, encontramos los vectores propios asociados $(1,1,1), (3,1,0), (-2,1,0)$ y entonces escribimos

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

De donde

$$P'_t = \frac{1}{5}\begin{pmatrix} 1 & 3 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-4t} & 0 \\ 0 & 0 & e^{-9t} \end{pmatrix}\begin{pmatrix} 0 & 0 & 5 \\ 1 & 2 & -3 \\ -1 & 3 & -2 \end{pmatrix}.$$

Lo que buscamos es $\mathbb{P}(\rho_2 \leq t) = P'_t(2,3)$, es decir

$$\begin{aligned} P'_t(2,3) &= \frac{1}{5}(1,1,1)\cdot(5, -3e^{-4t}, -2e^{-9t}) \\ &= 1 - \frac{3}{5}e^{-4t} - \frac{2}{5}e^{-9t}. \end{aligned}$$

5.5.4. **Inciso (iv).** *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?*

Cuando hay 2, generadores, utilizando el resultado de [5.5.2]. Tenemos

$$\lim_{t\to\infty} P_t(2,3) = \lim_{t\to\infty} \frac{1}{25}(18 - 6e^{-5t} - 2e^{-10t}) \qquad = \frac{18}{25}.$$

Para calcular la cantidad de energía promedio necesitamos el análogo para $P_t(2,2)$, que haciendo el mismo análisis que en [5.5.2] resulta ser

$$\lim_{t\to\infty} P_t(2,2) = \lim_{t\to\infty} \frac{1}{25}(6 + 3e^{-5t} + 16e^{-10t}) \qquad = \frac{6}{25}.$$

Entonces el promedio de mega Watts producidos es:

$$\begin{aligned}\left(\frac{18}{25}\right)5MW + \left(\frac{6}{25}\right)2.5MW &= \left(\frac{36}{50}\right)5MW + \left(\frac{6}{50}\right)5MW \\ &= \left(\frac{42}{10}\right)MW \\ &= 4.2MW.\end{aligned}$$

5.6. Problema 5.6

Problema 5.6.1 (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.

(i) [5.6.1] 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.

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 $X_t = k + S_{N_t}$.

(ii) [5.6.2] Diga brevemente por qué X es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal para ambos modelos.

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

(iii) [5.6.3] Argumente por qué Y es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal.

(iv) [5.6.4] 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.

Ahora nos enfocaremos en el proceso de Yule.

(v) [5.6.5] 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

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

(vi) [5.6.6] Al utilizar la fórmula para la esperanza de una variable binomial negativa, pruebe que

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

(vii) [5.6.7] 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.

(viii) [5.6.8] 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:

5.6.1. **Inciso (i).** *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.*

Para el primer caso, en el que los padres mueren, notemos que cualquier población es posible de un paso al siguiente. Si i es la cantidad de individuos en un momento determinado, cada individuo puede tener como descendencia cualquier cantidad de individuos (mayor o igual que cero). Sin embargo, si i es cero, el proceso se estaciona porque ya no existen individuos que dejen descendencia. Entonces la matriz de tasas de cambio α la podemos expresar como:

$$\alpha(i, j) = \begin{cases} 0 \ \text{ si } i = 0 \text{ ó si } j - i < -1 \\ \lambda i \mu_{k+1} \ \text{ si } j - i = k \end{cases}$$

Donde $\mu_1 = 0$, puesto que tener un individuo como descendencia y morir no afecta la población. Notemos que esta definición hace que $\sum_j \alpha(0, j) = 0$, es decir, el estado "población 0" es absorbente. Justo como necesitamos.

Con esto ahora podemos definir P como:

$$P(i,j) = \begin{cases} \frac{\alpha(i,j)}{\sum_{j'} \alpha(i,j')} = \mu_{j-i+1} \text{ si } \sum_{j'} \alpha(i,j') \neq 0 \\ 1 \text{ en cualquier otro caso} \end{cases}$$

Donde por comodidad de notación hacemos $\mu_{j-i+1} = 0$ si $j - i + 1 < 0$.

Para el segundo caso, donde nadie muere y puede tener cualquier cantidad de descendientes, la población nunca tiende a decrecer. Entonces la matriz de tasas de cambio α se puede expresar como:

$$\alpha(i,j) = \begin{cases} 0 \text{ si } i = 0 \text{ ó si } j < i \\ \lambda i \mu_k \text{ si } j - i = k > 0 \end{cases}$$

Donde hacemos $\mu_0 = 0$, puesto que no tener descendientes no afecta a la población. Entonces podemos calcular P de la misma manera que antes

$$P(i,j) = \begin{cases} \frac{\alpha(i,j)}{\sum_{j'} \alpha(i,j')} = \mu_{j-i+1} \text{ si } \sum_{j'} \alpha(i,j') \neq 0 \\ 1 \text{ en cualquier otro caso} \end{cases}$$

En las notas, en la sección 3 “Matrices infinitesimales y construcción de procesos de Markov” del capítulo 5 “Procesos de Markov constantes por pedazos” (véase 7.2.4), se explica como a partir de las tasas de cambio se puede construir una cadena de Markov a tiempo continuo con dichas tasas. Aplicándo dicho método, conseguimos las cadenas de Markov que buscamos.

5.6.2. **Inciso (ii).** *Diga brevemente por qué X es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal para ambos modelos.*

Sea T_n definido como en el inciso anterior y, sea $W_n = X_{T_n} = k + S_n$. Sean ξ_i las variables independientes que definen a la caminata aleatoria S.

Veamos que W cumple la condición de Markov.

$$\begin{aligned}
&\mathbb{P}\left(W_n = i_n | W_{n-1} = i_{n-1}, \ldots, W_1 = i_1\right) \\
&= \left(\frac{\mathbb{P}\left(W_n = i_n, W_{n-1} = i_{n-1}, \ldots, W_1 = i_1\right)}{\mathbb{P}\left(W_{n-1} = i_{n-1}, \ldots, W_1 = i_1\right)}\right) \\
&= \left(\frac{\mathbb{P}\left(W_n - W_{n-1} = i_n - i_{n-1}, \ldots, W_1 = i_1\right)}{\mathbb{P}\left(W_{n-1} - W_{n-2} = i_{n-1} - i_{n-2}, \ldots, W_1 = i_1\right)}\right) \\
&= \left(\frac{\mathbb{P}\left(\xi_n = i_n - i_{n-1}, \ldots, \xi_1 = i_1 - k\right)}{\mathbb{P}\left(\xi_{n-1} = i_{n-1} - i_{n-2}, \ldots, \xi_1 = i_1 - k\right)}\right) \\
&= \left(\frac{\mathbb{P}\left(\xi_n = i_n - i_{n-1}\right) \cdots \mathbb{P}\left(\xi_1 = i_1 - k\right)}{\mathbb{P}\left(\xi_{n-1} = i_{n-1} - i_{n-2}\right) \cdots \mathbb{P}\left(\xi_1 = i_1 - k\right)}\right) \\
&= \mathbb{P}\left(\xi_n = i_n - i_{n-1}\right) \\
&= \mathbb{P}\left(W_n = i_n | W_{n-1} = i_{n-1}\right).
\end{aligned}$$

Dado que los incrementos de T_n están dados por el proceso de Poisson, tenemos que tienen distribución exponencial de parámetro λ. Por lo tanto X_t es una cadena de Markov a tiempo continuo.

Para el caso en donde los padres mueren, recordando la suposición de que $\mu_1 = 0$ porque es el caso que no afecta a la población, tenemos que la matriz infinitesimal es

$$Q(i,j) = \begin{cases} \lambda i \mu_{j-i+1} & \text{si } i \leq j+1,\ j \neq i \text{ e } i \neq 0 \\ -\lambda i & \text{si } i = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Para el caso donde los padres no mueren, recordando la suposición de que $\mu_0 = 0$ porque es el caso que no afecta a la población, la matriz infenitesimal está dado por

$$Q(i,j) = \begin{cases} \lambda i \mu_{j-i} & \text{si } j > i \text{ e } i > 0 \\ -\lambda i & \text{si } i = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

5.6.3. **Inciso (iii).** *Argumente por qué Y es una cadena de Markov a tiempo continuo e identifique su matriz infinitesimal.*

Para el caso en el que los padres mueren. Detener a la martingala en τ, hace que el proceso se detenga en 0, es decir, el estado 0 es absorbente. Entonces, la matriz infinitesimal de este caso es

$$Q(i,j) = \begin{cases} \lambda i \mu_{j-i+1} & \text{si } i \leq j+1,\ j \neq i \text{ e } i \neq 0 \\ -\lambda i & \text{si } i = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Para el caso en el que los padres no mueren. Notemos que $\tau = \infty$. Puesto que la población nunca dismiuye. Entonces $t \wedge \tau = t$ y $Y_t = X_t$. Es decir, que la matriz infinitesimal de Y sería:

$$Q(i,j) = \begin{cases} \lambda i \mu_{j-i} & \text{si } j > i \text{ e } i > 0 \\ -\lambda i & \text{si } i = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

5.6.4. **Inciso (iv).** *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.

Definamos T_n como en [5.6.1]. Nos interesa ver cómo se comporta Y_t en los intervalos $[T_{n-1}, T_n)$, pues en estos intervalos Y_t es constante.

Analizemos el primer segmento, sea $t \in [0, T_1)$, donde la población no cambia, tenemos que $int_0^t Y_t = tk$. Supongamos que $Z_t = k$ tenemos que la integral de $\int_0^t Z_s ds = kt$ y entonces si $t \in [0, \frac{T_1}{k})$, tenemos que $\int_0^t Z_s ds < T_1$ y entonces $Y_{\int_0^t Z_s ds} = k = Z_t$. Entonces, al menos para $t \in [0, \frac{T_1}{k})$, $Z_t = k$. Por continuidad de la integral, para $t \in [0, \frac{T_1}{k})$, cualquier otro porceso Z que cumpla con la misma condición, es casi seguramente igual a k.

Analizemos el siguiente segmento, sea $t \in [T_1, T_2)$. Aquí $Y_t = k + S_1$. Queremos que $\int_{T_1}^t Z_s\, ds$ recorra todos los valores entre T_1 y T_2. Es decir

$$T_1 \leq \int_0^t k + \int_{\frac{T_1}{k}}^t S_1 < T_2$$

Que resulta equivalente a pedir:

$$\frac{T_1}{k} \leq t < \frac{T_1}{k} + \frac{T_2 - T_1}{k + S_1}$$

Es decir que basta definir a $Z_t = k + S_1$ cuando $t \in [\frac{T_1}{k}, \frac{T_1}{k} + \frac{T_2 - T_1}{k+S_1})$ para que Z se comporte como queremos. De nuvo, la continuidad de la integral nos dice que cualquier otra función que cumpla la condición, es casi seguramente igual a $k + S_1$ dentro del intervalo mencionado.

Generalizando esta idea, si $t \in [T_{n-1}, T_n)$, definimos $I_n = \sum_{0 \leq i}^{n-1} \frac{T_{i+1} - T_i}{k + S_i}$, donde $T_0 = S_0 = I_0 = 0$ por comodidad de notación. Para que $Z_t = Y_{\int_0^t Z_s\, ds}$ necesitamos

$$T_{n-1} \leq \sum_{0 \leq i}^{n-2} \int_{I_i}^{I_{i+1}} k + S_i + \int_{I_{n-1}}^t k + S_{n-1} < T_n$$

Después de muchas cuentas, esto significa que

$$I_{n-1} \leq t < I_n$$

E igual que antes, si $t \in [I_{n-1}, I_n)$ se cumple con la condición $Z_t = \int_0^t Z_s \, ds$ se satisface. Otra vez argumentando la continuidad de la integral, cualquier otro proceso que cumpla lo que Z, es casi seguramente igual a Z.

Ahora, por ser N un proceso de Poisson, $I_n - I_{n-1}$ resulta tener distribución exponencial y en estos intervalos Z es constante. No: es condicionalmente exponencial. Además hay que discutir la independencia condicional. Así que Z es una cadena de Markov a tiempo continuo. Y entonces es un proceso de ramificación.

5.6.5. **Inciso (v).** *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*

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

Esta vez, cada individuo tiene exáctamente un descendiente y no muere. Entonces, su matriz infinitesimal es como en [5.6.2], sólo que $\mu_n = 0$ si $n \neq 1$ y $\mu_1 = 1$. Es decir

$$Q(i,j) = \begin{cases} \lambda i \mu_1 & \text{si } j = i+1 \text{ e } i \neq 0 \\ -\lambda i & \text{si } i = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Entonces

$$\begin{aligned} \frac{d}{dt} P_t(i,j) &= \sum_{k \in \mathbb{N}} Q(i,k) P_t(k,j) \\ &= Q(i,i+1) P_t(i+1,j) + Q(i,i) P_t(i,j) \\ &= \lambda i P_t(i+1,j) - \lambda i P_t(i,j) \end{aligned}$$

Supongamos entonces $i = j$ y entonces

$$\begin{aligned} \frac{d}{dt} P_t(i,i) &= \lambda i P_t(i+1,i) - \lambda i P_t(i,i) \\ &= -i\lambda P_(i,i). \end{aligned}$$

Donde, la solución a la ecuación diferencial

$$\frac{d}{dt} P_t(i,i) = -\lambda i P_(i,i).$$

es

$$P_t(i,i) = e^{-ti\lambda}$$

Supongamos ahora $j = i+1$ y entonces

$$\begin{aligned} \frac{d}{dt} P_t(i,i+1) &= \lambda i P_t(i+1,i+1) - \lambda i P_t(i,i+1) \\ &= \lambda i e^{-t(i+1)\lambda} - \lambda i P_t(i,i+1) \end{aligned}$$

(la sustitución es por el caso recién demostrado)

De donde

$$\begin{aligned}
\frac{d}{dt}P_t(i,i+1)+\lambda i P_t(i,i+1) &= \lambda i e^{-t(i+1)\lambda} \\
(e^{\lambda it})\frac{d}{dt}P_t(i,i+1)+(e^{\lambda it})\lambda i P_t(i,i+1) &= (e^{\lambda it})\lambda i e^{-t(i+1)\lambda}
\end{aligned}$$

(se multiplicó en ambos lados por lo mismo)

$$\begin{aligned}
(e^{\lambda it})\frac{d}{dt}P_t(i,i+1)+(e^{\lambda it})\lambda i P_t(i,i+1) &= \lambda i e^{-t\lambda} \\
\frac{d}{dt}\Big((e^{\lambda it})P_t(i,i+1)\Big) &= \lambda i e^{-t\lambda}
\end{aligned}$$

(la suma se escribió como la derivada de un producto)

Este último término, lo podemos integrar suponiendo la condición inicial $P_0(i,i+1)=0$

$$\begin{aligned}
(e^{\lambda it})P_t(i,i+1) &= \int \frac{d}{dt}\Big((e^{\lambda it})P_t(i,i+1)\Big)dt \\
&= \int_0^t \lambda i e^{-t\lambda}dt \\
&= i(1-e^{\lambda t})
\end{aligned}$$

Y por lo tanto

$$P_t(i,i+1)=i(1-e^{\lambda t})(e^{-\lambda it})$$

Hagamos entonces la hipótesis de inducción (que mágicamente se ha cumplido hasta ahora, pero que para encontrar fue necesario hacer más cuentas)

$$P_t(i,i+k)=\binom{i+k-1}{k}e^{-\lambda it}(1-e^{-\lambda t})^k$$

Probemos entonces que para $j=i+k+1$ también se cumple:

$$\begin{aligned}
\frac{d}{dt}P_t(i,i+k+1) &= \lambda i P_t(i+1,i+1+k)-\lambda i P_t(i,i+1+k) \\
&= \lambda i\binom{i+k}{k}e^{-\lambda(i+1)t}(1-e^{-\lambda t})^k-\lambda i P_t(i,i+1+k)
\end{aligned}$$

Y entonces

$$\frac{d}{dt}P_t(i,i+k+1) - \lambda i P_t(i,i+1+k) = \lambda i \binom{i+k}{k} e^{-\lambda(i+1)t}(1-e^{-\lambda t})^k$$

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

(se multiplicó en ambos lados por lo mismo, otra vez)

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

(la suma se escribió como la derivada de un producto)

Por lo tanto

$$\begin{aligned}
(e^{\lambda i t})P_t(i,i+k+1) &= \int \frac{d}{dt}\left((e^{\lambda i t})P_t(i,i+k+1)\right)dt \\
&= \int_0^t \lambda i \binom{i+k}{k} e^{-\lambda t}(1-e^{-\lambda t})^k dt \\
&= \binom{i+k}{k} i \int_0^t \lambda e^{-\lambda t}(1-e^{-\lambda t})^k dt \\
&= \binom{i+k}{k} \frac{i}{k+1}(1-e^{-\lambda t})^{k+1} \\
&= \binom{i+k+1}{k+1}(1-e^{-\lambda t})^{k+1}
\end{aligned}$$

y entonces

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

como queríamos demostrar.
Entonces basta sustituir a $k = j - i$ para tener una fórmula en términos de i, j.

$$P_t(i,j) = \binom{j-1}{j-i} e^{-\lambda it}(1-e^{-\lambda t})^k$$

5.6.6. **Inciso (vi).** *Al utilizar la fórmula para la esperanza de una variable binomial negativa, pruebe que*

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

Del inciso anterior, tenemos que $P_t(i,j) = \mathbb{P}(W = i - j)$ donde W es una variable aleatoria con distribución binomial negativa de parámetros $r = x$ y $p = e^{-\lambda t}$ [véase distribución binomial negativa].

Entonces

$$\begin{aligned}
\mathbb{E}_x(Z_t) &= \sum_j j\mathbb{P}_x(Z_t = j) \\
&= \sum_j jP_t(x,j) \\
&= \mathbb{E}(W) \\
&= \frac{r}{p} \\
&= \frac{x}{e^{-\lambda t}} \\
&= xe^{\lambda t}.
\end{aligned}$$

Como buscábamos demostrar.

5.6.7. **Inciso (vii).** *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.*

De [5.6.4] tenemos que Z_t es positiva.
Sea $(\mathscr{F}_t)_{t\geq 0}$ la filtración natural para Z_t, es decir, $\mathscr{F}_t = \sigma(Z_s : s \leq t)$. De tal manera que $e^{-\lambda t}Z_t$ es adaptada a $(\mathscr{F}_t)_{t\geq 0}$.

En [5.6.6] se demostró que $Z_t \in L_1$, como también $e^{-\lambda t} \in L_1$, entonces $e^{-\lambda t}Z_t \in L_1$.

Únicamente falta demostrar la propiedad de martingala. Sean entonces $s \leq t$.

$$\begin{aligned}
\mathbb{E}(e^{-\lambda t}Z_t|\mathscr{F}_s) &= e^{-\lambda t}\mathbb{E}(Z_t|\mathscr{F}_s) \\
&= e^{-\lambda t}\mathbb{E}_{Z_s}(Z_{t-s}) \\
&\quad \text{(propiedad de Markov fuerte)} \\
&= e^{-\lambda t}Z_s e^{\lambda(t-s)} \\
&\quad \text{(por [5.6.6])} \\
&= e^{\lambda(t-s)-\lambda t}Z_s \\
&= e^{\lambda s}Z_s
\end{aligned}$$

De donde $e^{-\lambda t}Z_t$ es martingala.

Por ser no negativa, el teorema de convergencia de martingalas nos garantiza que el proceso converge casi seguramente a una variable aleatoria W.

5.6.8. **Inciso (viii).** *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.*

La transformada de Laplace de una distribución binomial negativa de parámetros p, r está dada por

$$\mathbb{E}(e^{uZ_t}) = \left(\frac{pe^u}{1-(1-p)e^u}\right)^r$$

[véase propiedades de la distribución binomial negativa]

En nuestro caso, $r = x$ y $p = e^{-\lambda t}$, suponiendo $x = 1$ como población inicial.

Tenemos entonces

$$\begin{aligned}
\mathbb{E}(e^{uZ_t}) &= \left(\frac{e^{-\lambda t}e^u}{1-(1-e^{-\lambda t})e^u}\right)^1 \\
&= \frac{e^{-\lambda t}e^u}{1-(1-e^{-\lambda t})e^u} \\
&= \frac{e^{-\lambda t}}{e^{-u}-(1-e^{-\lambda t})}
\end{aligned}$$

Hacemos un cambio de variable $u = ve^{-\lambda t}$ y entonces obtenemos

$$\mathbb{E}(e^{ve^{-\lambda t}Z_t}) = \frac{e^{-\lambda t}}{e^{-ve^{-\lambda t}}-(1-e^{-\lambda t})}$$

De donde

$$\begin{aligned}
\lim t \to \infty \mathbb{E}(e^{ve^{-\lambda t}Z_t}) &= \lim t \to \infty \frac{e^{-\lambda t}}{e^{-ve^{-\lambda t}}-(1-e^{-\lambda t})} \\
&= \lim s \to 0 \frac{s}{e^{-vs}-(1-s)} \\
&\qquad \text{(se hizo un cambio de variable } s = e^{-\lambda t}\text{, el} \\
&\qquad \text{cual tiende a 0 conforme } t \text{ a infinito)} \\
&= \lim s \to 0 \frac{1}{-ve^{-vs}+1} \\
&\qquad \text{(se aplicó L'Hopital, pues quedaba un límite del estilo } \tfrac{0}{0}\text{)} \\
&= \frac{1}{1-v}
\end{aligned}$$

Lo cual corresponde a la función característica de una exponencial de parámetro 1 y por lo tanto W tiene distribución exponencial.

5.7. Problema 5.7

Problema 5.7.1. (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

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

(i) [5.7.1] Explique por qué X es una cadena de Markov a tiempo continuo con valores en E.

(ii) [5.7.1] Calcule sus probabilidades de transición y su matriz infinitesimal.

(iii) [5.7.1] ¿Existe una distribución estacionaria para esta cadena? En caso afirmativo ¿Cuál es?

Demostración:

5.7.1. **Inciso (i).** *Explique por qué X es una cadena de Markov a tiempo continuo con valores en E.*

Sea T_n el tiempo en el que ocurre el n-ésimo salto del proceso de Poisson ($T_0 = 0$). Definimos $W_{T_n} = X_{T_n}$. Veamos que W_n es una cadena de Markov.

$$\begin{aligned}
&\mathbb{P}\left(W_n = i_n | W_{n-1} = i_{n-1}, \ldots, W_0 = i_0\right) \\
&= \mathbb{P}\left(X_0 \cdot (-1)^n = i_n | X_0 \cdot (-1)^{n-1} = i_{n-1}, X_0 = i_0\right) \\
&= \mathbb{P}\left(-(i_{n-1}) = i_n | X_0 \cdot (-1)^{n-1} = i_{n-1}, X_0 = i_0\right) \\
&= \mathbb{1}_{-(i_{n-1}) = i_n} \\
&= \mathbb{P}\left(X_0 \cdot (-1)^n = i_n | X_0 \cdot (-1)^{n-1} = i_{n-1}\right) \\
&= \mathbb{P}\left(W_n = i_n | W_{n-1} = i_{n-1}\right)
\end{aligned}$$

Con esto hemos demostrado que W_n es una cadena de Markov a tiempo discreto. Como además, el tiempo entre saltos tiene distribución exponencial (porque está definido por un proceso de Poisson), concluimos que X, es una cadena de Markov.

5.7.2. **Inciso (ii).** *Calcule sus probabilidades de transición y su matriz infinitesimal.*

Recordemos que los tiempos de cambio están dados por un proceso de Poisson de parámetro λ y que únicamente tenemos 2 estados. Es decir que para pasar de un estado al otro, la tasa de cambio es precisamente λ y que la tasa de cambio de un estado a sí mismo es 0. Dicho esto, tenemos que la matriz infinitesimal tiene que ser

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

Esto no le deja más opciones a la matriz de la cadena de Markov asociada más que ser:

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

Para poder calcular P_t, escribimos a Q como conjugada de una matriz diagonal:

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

Utilizando de nuevo que $P_t = e^{tQ}$ y que si $Q = ADA^{-1}$ entonces $e^{tQ} = Ae^{tD}A^{-1}$, obtenemos que las matrices de transición son

$$\begin{aligned} P_t &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & e^{-2\lambda} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & e^{-2\lambda} \\ 1 & -e^{-2\lambda} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1+e^{-2\lambda} & 1-e^{-2\lambda} \\ 1-e^{-2\lambda} & 1+e^{-2\lambda} \end{pmatrix} \end{aligned}$$

5.7.3. **Inciso (iii).** *¿Existe una distribución estacionaria para esta cadena? En caso afirmativo ¿Cuál es?*

Sí existe. Dado que se trata de un proceso de estados finitos y recurrentes, sabemos que es positivo recurrente y por lo tanto existe una distribución invariante que además es única.

Existe un teorema que nos asegura que una distribución ν es invariante para una cadena continua si y solo si $cv = (c_x \nu_x, x \in E)$ es invariante para la cadena discreta asociada.

Entonces estamos buscando un vector (π_1, π_2) tal que $(\pi_1, \pi_2)P = (\pi_1, \pi_2)$.

Pero

$$\begin{aligned}(\pi_1, \pi_2)P &= (\pi_1, \pi_2)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= (\pi_2, \pi_1)\end{aligned}$$

Así que $\pi_1 = \pi_2$. Además $\pi_1 + \pi_2 = 1$ por tratarse de una distribución. Así que $\pi_1 = \pi_2 = \frac{1}{2}$.

5.8. Problema 5.8

Problema 5.8.1. Sea

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

(i) [5.8.1] Haga un programa en octave que permita simular las trayectorias de una cadena de Markov a tiempo continuo X con matriz infinitesimal Q.

(ii) [5.8.2] 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_{10}.

(iii) [5.8.3] Calcule e^{10Q} (utilizando algún comando adecuado) y contraste con la distribución empírica del inciso anterior.

(iv) [5.8.4] 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:

5.8.1. **Inciso (i).** *Haga un programa en octave que permita simular las trayectorias de una cadena de Markov a tiempo continuo X con matriz infinitesimal Q.*

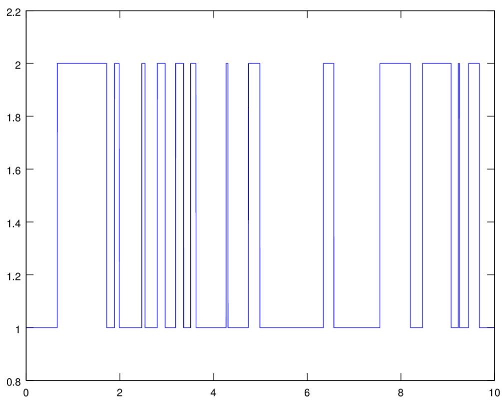
———————— . ————————

Para poder reutilizar el código en el siguiente inciso, esta parte se programó como una función de `Tmax`. `Tmax` que representa una cota para el tiempo que se representa en la ejecución de la Cadena de Markov.

```
function [T,X]    = SimMarkov(Tmax)
    X(1) = ceil(2*rand());
    T(1) = 0;
```

```
    i=2;
    while max(T) <= Tmax;
        estado_anterior = X(i-1);

        #cambia el estado.
        X(i) = estado_anterior + (-1)^(estado_anterior + 1);

        #calcula el tiempo de salto siguiente, aprovechando
        #que los estados son 1 y 2. Y las tasas 2 y 3
        #(por eso el +1 en los parametros
        # de la variable exponencial.
        T(i) = exprnd(1/(estado_anterior + 1)) + T(i-1);
        i = i+1;
    endwhile
endfunction
```

A continuación una imagen que resultó de ejecutar esta función



5.8.2. **Inciso (ii).** *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_{10}.

El siguiente es el código que se utilizó para simular 10000 trayectorias de la cadena de Markov. Las últimas lineas comentadas son el resultado de la distribución empírica de X_{10}

```
for k=[1:10000]
        [T,X]   = SimMarkov(10);
        j = sum(T <= 10);
        x(k) = X(j);
endfor

p1 = mean(x-1)
p2 = 1-p1

##resultado devuelto en consola
##p1 = 0.39940
##p2 = 0.60060
```

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

El siguiente código es el que se utilizó para calcular el resultado teórico de e^{10Q}. Las últimas lineas comentadas son el resultado.

```
Q = [-2 2; 3 -3];
expm(10*Q)

## 0.600000000000002    0.400000000000001
## 0.600000000000002    0.400000000000001
```

El código para obtener el resultado empírico se ejecutó varias veces. Lo que se mostró es el resultado con el error (con respecto al resultado teórico) más alto conseguido.

5.8.4. **Inciso (iv).** *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?*

El siguiente código es la implementación.

```
function P = AproxE10Q(h)
    Q = [-2 2; 3 -3];

    P = eye(2);
    lim = floor(10/h);
    for i=[1:lim]
        R = h*(Q*P);
        P = P + R;
    endfor
endfunction
```

En realidad, la fórmula recursiva no es necesaria. Podríamos escribir:

$$P_i^h = (I + hQ)^i$$

Y de aquí, que si el máximo de los valores absolutos de las entradas de $I + hQ$ es menor o igual que 1, la convergencia de P_i^h está garantizada.

Haciendo algunas cuentas sencillas, se encuentra que si $h = 1/3$ ocurre lo comentado en el párrafo anterior. `(I + hQ)^33` resulta estar a menos de 6 decimales de diferencia de e^{10Q}. Sin embargo, el tope que se le pone al algoritmo es `floor(10/(1/3)) = 30`.

Con el siguiente algoritmo, se buscó una h chiquita que hiciera que el resultado del algoritmo se acercara a e^{10Q} lo suficiente.

```
h = 1/3;
r=1;
i=1;
E10Q = expm(10*[-2 2; 3 -3]);
while( r > 0.000001)
```

```
    r = max(max(abs(AproxE10Q(h) - E10Q)));
    h = h - h/10000;
    i=i+1;
endwhile
h
```

El resultado fue `h = 0.328304729443698`.

•___•••_•___•_•_•__•____•__••_•••_•••••__••_•_•___••_••__•••_•___•_•••_•••••__••_••

Part 6. Tarea 6

6.1. Problema 6.1

Problema 6.1.1. Un proceso estocástico $B = (B_t, t \geq 0)$ es un movimiento browniano en ley si y sólo si es un proceso gaussiano centrado y $\mathbb{E}(B_s B_t) = s \wedge t$.

Necesidad. Sea $B = (B_t, t \geq 0)$ un movimiento browniano en ley. Sabemos que los incrementos de un movimiento browniano conforman un vector gaussiano y que son independientes. También sabemos que $\mathbb{E}(B_t) = 0$ (que el proceso es centrado), y que $\mathbb{E}(B_s B_t) = s \wedge t$.

Sólo falta ver que se trata de un proceso gaussiano.

Sean $0 \geq t_1 < t_2 < \cdots < t_n$.

Para cualquier combinación lineal

$$\sum_{1 \leq i \leq n} \alpha_i B_{t_i}$$

Podemos encontrar β_i's $(1 \leq i \leq n)$ tales que:

$$\sum_{1 \leq i \leq n} \alpha_i B_{t_i} = \sum_{1 \leq i \leq n} \beta_i (B_{t_i} - B_{t_{i-1}})$$

(Donde por comodidad de notación $B_{t_0} = 0$). Para el argumento de que podemos encontrar a las β_i's, considérese el sistema de ecuaciones lineales $\alpha_i = \beta_i - \beta_{i+1}$, donde $\beta_{n+1} = 0$ y entonces tenemos un sistema lineal de n ecuaciones y n variables.

Entonces, cualquier combinación lineal del estilo $\sum_{1 \leq i \leq n} \alpha_i B_{t_i}$ es una combinación linea de variables aleatorias normales y por lo tanto normal. Es decir que $(B_{t_1}, B_{t_2}, \ldots, B_{t_n})$, es un vector gaussiano. Que es lo único que hacía falta.

Suficiencia. Sea $B = (B_t, t \geq 0)$ un proceso gaussiano centrado tal que $\mathbb{E}(B_s B_t) = s \wedge t$.

Veamos que el proceso comienza en 0. Que el proceso sea centrado nos dice que $\mathbb{E}(B_0) = 0$. Como la esperanza es 0, entonces la varianza es $Var(B_0) = \mathbb{E}(B_0^2) = \mathbb{E}(B_0 B_0) = 0 \wedge 0 = 0$. Entonces $B_0 = 0$ c.s.

De manera similar, como $\mathbb{E}(B_t) = 0$, tenemos que $Var(B_t) = \mathbb{E}(B_t B_t) = t$. Como B_t es una combinación linea de un proceso gaussiano, entonces tiene distribución normal. Y como acabamos de ver, con media 0 y varianza t. Igual que antes, para cualquier combinación lineal

$$\sum_{1 \leq i \leq n} \beta_i (B_{t_i} - B_{t_{i-1}})$$

(Donde por comodiad de notación $B_{t_0} = 0$). Podemos encontrar α_i's ($1 \leq i \leq n$) tales que

$$\sum_{1 \leq i \leq n} \beta_i (B_{t_i} - B_{t_{i-1}}) = \sum_{1 \leq i \leq n} \alpha_i B_{t_i}.$$

(El argumento es el mismo que arriba). Y entonces tenemos que los incrementos siempre forman un vector gaussiano. Para ver la independencia de los incrementos sean $1 \leq i < j \leq n$ y entonces

$$\begin{aligned}
&\mathbb{E}((B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}})) \\
&= \mathbb{E}(B_{t_i} B_{t_j} - B_{t_i} B_{t_{j-1}} - B_{t_{i-1}} B_{t_j} + B_{t_{i-1}} B_{t_{j-i}}) \\
&= \mathbb{E}(B_{t_i} B_{t_j}) - \mathbb{E}(B_{t_i} B_{t_{j-1}}) - \mathbb{E}(B_{t_{i-1}} B_{t_j}) + \mathbb{E}(B_{t_{i-1}} B_{t_{j-i}}) \\
&= t_i \wedge t_j - t_i \wedge t_{j-1} - t_{i-1} \wedge t_j + t_{i-1} \wedge t_{j-i} \\
&= t_i - t_i - t_{i-1} + t_{i-1} \\
&= 0.
\end{aligned}$$

Y por lo tanto, la correlación entre los incrementos, siempre es 0. Por lo que son independientes.

Ahora sean $s, t \geq 0$. Sabemos que $\mathbb{E}(B_{t+s} - B_t) = \mathbb{E}(B_{t+s}) - \mathbb{E}(B_t) = 0 - 0 = 0$. Así que $Var(B_{t+s} - B_t) = \mathbb{E}((B_{t+s} - B_t)^2)$. Desarrollando de manera similar a hace un momento

$$\begin{aligned}
E((B_{t+s} - B_t)^2) &= \mathbb{E}(B_{t+s} B_{t+s}) - 2\mathbb{E}(B_{t+s} B_t) + \mathbb{E}(B_t B_t) \\
&= t + s - 2t + t \\
&= s.
\end{aligned}$$

Además $B_{t+s} - B_t$ es una combinación lineal de variables de un proceso gaussiano, y por lo tanto tiene distribución gaussiana con media 0 y varianza s (por lo que acabamos de demostrar). Por lo tanto $(B_{t+s} - B_t) \sim B_s$. Entonces tenemos B cumple con todas las hipótesis para ser movimiento browniano en ley.

6.2. Problema 6.2

Problema 6.2.1. El objetivo de este problema es construir, a partir de movimientos brownianos en $[0,1]$, al movimiento browniano en $[0,\infty)$.

(1) Pruebe que existe un espacio de probabilidad $(\Omega,\mathscr{F},\mathbb{P})$ en el que existe una sucesión $B^1,B^2,\ldots$ de movimientos brownianos en $[0,1]$ independientes. (Sugerencia: utilice la construcción del movimiento browniano de Lévy para que la solución sea corta.)

(2) Defina a $B_t = B^1_1 + \cdots + B^{\lfloor t\rfloor}_1 + B^{\lceil t\rceil}_{t-\lfloor t\rfloor}$ para $t \geq 0$. Pruebe que B es un movimiento browniano.

Demostración:

6.2.1. **Inciso (i).** *Pruebe que existe un espacio de probabilidad* $(\Omega,\mathscr{F},\mathbb{P})$ *en el que existe una sucesión* $B^1,B^2,\ldots$ *de movimientos brownianos en* $[0,1]$ *independientes. (Sugerencia: utilice la construcción del movimiento browniano de Lévy para que la solución sea corta.)*

La idea principal es utilizar que $\aleph_0\times\aleph_0=\aleph_0$ y aprovechar la construcción que se dió en las notas [véase 7.2.4].

Se sabe que existe un espacio de probabilidad $(\Omega,\mathscr{F},\mathbb{P})$ donde están definidas $\xi_{i,n}$ $(0\leq i\leq 2^n)$ con distribuciónes normales de media 0 y varianza 1 e independientes [ver 7.2.4].

Entonces, podemos para cada $m\in\mathbb{N}$ dar una sucesión de variables $\xi^m_{i,n}$ $(0\leq i\ 2^n)$ con distribuciones normales de media 0 y varianza 1 tales que son independientes entre sí y de las otras $\xi^{m'}_{i,n}$.

Entonces, para cada m podemos construir un movimiento browniano B^m en $[0,1]$ utilizando las $\xi^m_{i,n}$. Dichos movimientos serán independientes entre sí por la selección de las $(\xi^m_{i,n})_{m\in\mathbb{N}}$.

6.2.2. **Inciso (ii).** *Defina a* $B_t = B_1^1 + \cdots + B_1^{\lfloor t \rfloor} + B_{t-\lfloor t \rfloor}^{\lceil t \rceil}$ *para* $t \geq 0$. *Pruebe que* B *es un movimiento browniano.*

$B_0 = B_0^1 = 0$.
Veamos que B_t es un proceso gaussiano. Sean $0 \leq t_1 < t_2 < \cdots < t_n$ y sea

$$\sum_{i \leq n} \lambda_i B_{t_i}$$

una combinación lineal. Por ser los B^m's movimientos brownianos independientes, entonces cada B_{t_i} es una combinación lineal de normales independientes, por lo tanto cada B_{t_i} es normal y por construcción, independientes. Si sucediera que más de un B_{t_i} pertenecieran al mismo B^m, basta argumentar que cada B^m es por sí mismo un proceso gaussiano. Entonces nuestra combinación lineal también es normal y con esto hemos dicho que B_t es un proceso gaussiano.

B_t es suma de movimientos brownianos, los cuales son centrados. Por linealidad de la esperanza tenemos que $\mathbb{E}(B_t) = 0$.

Ahora,

$$\begin{aligned}
Var(B_t) &= Var\left(B_1^1 + \cdots + B_1^{\lfloor t \rfloor} + B_{t-\lfloor t \rfloor}^{\lceil t \rceil}\right) \\
&= \sum_{i \leq \lfloor t \rfloor} Var(B_1^i) + Var(B_{t-\lfloor t \rfloor}^{\lceil t \rceil}) \\
&\quad \text{(por independencia de los movimientos)} \\
&= \lfloor t \rfloor + Var(B_{t-\lfloor t \rfloor}^{\lceil t \rceil}) \\
&\quad \text{(porque las } B_1^i \text{ tienen distribución normal(0,1))} \\
&= \lfloor t \rfloor + t - \lfloor t \rfloor \\
&\quad \text{(por que la distribución de } B_{t-\lfloor t \rfloor}^{\lceil t \rceil} \text{ es normal}(0,\, t - \lfloor t \rfloor)) \\
&= t.
\end{aligned}$$

Entonces, B_t tiene distribución normal(0,t).

Veamos que $\mathbb{E}(B_t B_s) = t \wedge s$. Si $t = s$, aprovechamos que B_t tiene distribución normal de media 0 y varianza 1 y entonces $\mathbb{E}(B_t B_t) = Var(B_t) = t$. Si $t < s$, partiremos en dos casos, uno cuando $s \leq \lceil t \rceil$ y otro para $\leq \lceil t \rceil < s$. ¿Cuál es el segundo caso?

Para el primer caso

$$B_t = B_1^1 + \cdots + B_1^{\lfloor t\rfloor} + B_{t-\lfloor t\rfloor}^{\lceil t\rceil}$$
$$B_s = B_1^1 + \cdots + B_1^{\lfloor t\rfloor} + B_{s-\lfloor t\rfloor}^{\lceil t\rceil}$$

Entonces

$$\begin{aligned}
\mathbb{E}(B_tB_s) &= \mathbb{E}\left(\left(B_1^1 + \cdots + B_1^{\lfloor t\rfloor} + B_{t-\lfloor t\rfloor}^{\lceil t\rceil}\right)\cdot\left(B_1^1 + \cdots + B_1^{\lfloor t\rfloor} + B_{s-\lfloor t\rfloor}^{\lceil t\rceil}\right)\right)\\
&= \sum_{i\le\lfloor t\rfloor}\mathbb{E}\left((B_1^i)^2\right) + \mathbb{E}\left(B_{t-\lfloor t\rfloor}^{\lceil t\rceil}B_{s-\lfloor t\rfloor}^{\lceil t\rceil}\right)\\
&\quad\text{(por independencia de los movimientos } B^m)\\
&= \lfloor t\rfloor + \mathbb{E}\left(B_{t-\lfloor t\rfloor}^{\lceil t\rceil}B_{s-\lfloor t\rfloor}^{\lceil t\rceil}\right)\\
&= \lfloor t\rfloor + (t-\lfloor t\rfloor)\wedge(s-\lfloor t\rfloor)\\
&= \lfloor t\rfloor + (t-\lfloor t\rfloor)\\
&= t.
\end{aligned}$$

Para el segundo caso

$$B_t = B_1^1 + \cdots + B_1^{\lfloor t\rfloor} + B_{t-\lfloor t\rfloor}^{\lceil t\rceil}$$
$$B_s = B_1^1 + \cdots + B_1^{\lfloor s\rfloor} + B_{s-\lfloor s\rfloor}^{\lceil s\rceil}$$

Entonces

$$\begin{aligned}
\mathbb{E}(B_tB_s) &= \sum_{i\le\lfloor t\rfloor}\mathbb{E}\left((B_1^i)^2\right) + \mathbb{E}\left(B_{t-\lfloor t\rfloor}^{\lceil t\rceil}B_1^{\lceil t\rceil}\right)\\
&\quad\text{(por independencia de los movimientos } B^m)\\
&= \lfloor t\rfloor + \mathbb{E}\left(B_{t-\lfloor t\rfloor}^{\lceil t\rceil}B_1^{\lceil t\rceil}\right)\\
&= \lfloor t\rfloor + (t-\lfloor t\rfloor)\\
&= t.
\end{aligned}$$

Con todo esto dicho, basta utilizar [6.1] para conculir que B_t es un movimiento browniano en ley. Y la continuidad de las trayectorias se dá por la construcción de B_t puesto que los B^m que lo componen son continuos.

6.3. Problema 6.3

Problema 6.3.1. Pruebe que si $\tilde{X}$ es una modificación de X entonces ambos procesos tienen las mismas distribuciones finito-dimensionales. Concluya que si B es un movimiento browniano en ley y $\tilde{B}$ es una modificación de B con trayectorias continuas entonces $\tilde{B}$ es un movimiento browniano.

Que $\tilde{X}$ sea una modificación de X significa que para todo $t \geq 0$ se tiene que $\mathbb{P}(\tilde{X}_t = X_t) = 1$.

Sean $0 \leq t_1 < t_2 < \cdots < t_n$ y sea $A = \bigcap_{i \leq n} \{\tilde{X}_{t_i} = X_{t_i}\}$, como A es intersección de sucesos con probabilidad 1, se tiene que $\mathbb{P}(A) = 1$. Y por lo tanto $\mathbb{P}(A^c) = 0$.

Sean ahora $x_1, x_2, \ldots, x_n \in \mathbb{R}$, $B = \bigcap_{i \leq n} \{X_{t_i} \leq x_i\}$ y $C = \bigcap_{i \leq n} \{\tilde{X}_{t_i} \leq x_i\}$. Notemos que

$$
\begin{aligned}
A \cap B &= \left(\bigcap_{i \leq n} \{\tilde{X}_{t_i} = X_{t_i}\} \right) \cap \left(\bigcap_{i \leq n} \{X_{t_i} \leq x_i\} \right) \\
&= \bigcap_{i \leq n} \Big(\{\tilde{X}_{t_i} = X_{t_i}\} \cap \{X_{t_i} \leq x_i\} \Big) \\
&= \bigcap_{i \leq n} \Big(\{\tilde{X}_{t_i} = X_{t_i}\} \cap \{\tilde{X}_{t_i} \leq x_i\} \Big) \\
&= \left(\bigcap_{i \leq n} \{\tilde{X}_{t_i} = X_{t_i}\} \right) \cap \left(\bigcap_{i \leq n} \{\tilde{X}_{t_i} \leq x_i\} \right) \\
&= A \cap C
\end{aligned}
$$

y por lo tanto

$$
\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left((B \cap A) \cup (B \cap A^c) \right) \\
&= \mathbb{P}\left(B \cap A \right) + \mathbb{P}\left(B \cap A^c \right) \\
&= \mathbb{P}\left(B \cap A \right) + 0 \\
&= \mathbb{P}\left(C \right)
\end{aligned}
$$

y por lo tanto $\mathbb{P}(X_{t_1} \leq x_1, \ldots, X_{t_n} \leq x_n) = \mathbb{P}(\tilde{X}_{t_1} \leq x_1, \ldots, \tilde{X}_{t_n} \leq x_n)$, lo cual significa que $\tilde{X}$ y X tienen las mismas distribuciones finito dimensionales.

Si B es un movimiento browniano, y $\tilde{B}$ es una modificación de B, por lo que se acaba de demostrar, cualquier vector de la forma $(B_{t_1}, \ldots, B_{t_n})$ (el cual es gaussiano centrado en 0), tiene la misma distribución que $(\tilde{B}_{t_1}, \ldots, \tilde{B}_{t_n})$ y por lo tanto, este segundo también es un vector gaussiano centrado en 0. Además $\mathbb{E}(\tilde{B}_t \tilde{B}_s) = \mathbb{E}(B_t B_s) = t \wedge s$ así que, por [6.1], tenemos que se trata de un movimiento browniano en ley. Además, por hipótesis, $\tilde{B}$ tiene trayectorias continuas, y por lo tanto es un movimiento browniano.

6.4. Problema 6.4

Problema 6.4.1. Sea
$$M_t^\lambda = e^{\lambda B_t - \lambda^2 t/2}.$$

(i) [6.4.1] Explique y pruebe formalmente por qué, para toda $n \geq 1$, $\partial^n M_t^\lambda / \partial \lambda^n$ es una martingala.

(ii) [6.4.2] Sea $H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$. A H_n se le conoce como enésimo polinomio de Hermite. Calcúlelo para $n \leq 5$. Pruebe que H_n es un polinomio para toda $n \in \mathbb{N}$ y que $\partial^n M_t^\lambda / \partial \lambda^n = t^{n/2} H_n\big(B_t/\sqrt{t}\big) M_t^\lambda$.

(iii) [6.4.3] Pruebe que $t^{n/2} H_n\big(B_t/\sqrt{t}\big)$ es una martingala para toda n y calcúlela para $n \leq 5$.

(iv) [6.4.4] Aplique muestreo opcional a las martingalas anteriores al tiempo aleatorio $T_{a,b} = \min\{t \geq 0 : B_t \in \{-a, b\}\}$ (para $a, b > 0$) con $n = 1, 2$ para calcular $\mathbb{P}\big(B_{T_{a,b}} = b\big)$ y $\mathbb{E}(T_{a,b})$, ËQué concluye cuando $n = 3, 4$? ¿ Cree que $T_{a,b}$ tenga momentos finitos de cualquier orden? Justifique su respuesta.

(v) [6.4.5] Aplique el teorema de muestreo opcional a la martingala M^λ al tiempo aleatorio $T_a = \inf\{t \geq 0 : B_t \geq a\}$ si $\lambda > 0$. Diga por qué es necesaria la última hipótesis y calcule la transformada de Laplace de T_a.

Demostración:

6.4.1. **Inciso (i).** *Explique y pruebe formalmente por qué, para toda $n \geq 1$, $\partial^n M_t^\lambda / \partial \lambda^n$ es una martingala.*

Analicemos las primeras derivadas.

$$\begin{aligned}
\frac{d}{d\lambda} M_t^\lambda &= \frac{d}{d\lambda}\left(e^{\lambda B_t - \lambda^2 t/2}\right) \\
&= \frac{d}{d\lambda}\left(e^{\lambda B_t} e^{-\lambda^2 t/2}\right) \\
&= B_t e^{\lambda B_t} e^{-\lambda^2/2} - \lambda t e^{\lambda B_t} e^{-\lambda^2 t/2} \\
&= B_t M_t^\lambda - \lambda t M_t^\lambda \\
&= M_t^\lambda (B_t - \lambda t).
\end{aligned}$$

La de segundo orden

$$\begin{aligned}
\frac{d^2}{d\lambda^2} M_t^\lambda &= \frac{d}{d\lambda}\left(M_t^\lambda (B_t - \lambda t)\right) \\
&= M_t^\lambda (B_t - \lambda t)(B_t - \lambda t) - t M_t^\lambda \\
&= M_t^\lambda \left[(B_t - \lambda t)^2 - t\right] \\
&= t M_t^\lambda \left[(B_t - \lambda t)^2 (t^{-1/2})^2 - 1\right]
\end{aligned}$$

La de tercer orden

$$\begin{aligned}
&\frac{d^3}{d\lambda^3} M_t^\lambda \\
&= \frac{d}{d\lambda}\left(t M_t^\lambda \left[(B_t - \lambda t)^2 (t^{-1/2})^2 - 1\right]\right) \\
&= t \frac{d}{d\lambda}\left(M_t^\lambda \left[(B_t - \lambda t)^2 (t^{-1/2})^2 - 1\right]\right) \\
&= t M_t^\lambda \left[(B_t - \lambda t)^3 (t^{-1/2})^2 - (B_t - \lambda t)\right] + M_t^\lambda \frac{d}{d\lambda}\left[(B_t - \lambda t)^2 (t^{-1/2})^2 - 1\right] \\
&= t M_t^\lambda \left(\left[(B_t - \lambda t)^3 (t^{-1/2})^2 - (B_t - \lambda t)\right] + \frac{d}{d\lambda}\left[(B_t - \lambda t)^2 (t^{-1/2})^2 - 1\right]\right) \\
&= t M_t^\lambda \left(\left[(B_t - \lambda t)^3 (t^{-1/2})^2 - (B_t - \lambda t)\right] + \left[-2t(B_t - \lambda t)(t^{-1/2})^2\right]\right) \\
&= t M_t^\lambda \left(\left[(B_t - \lambda t)^3 (t^{-1/2})^2 - (B_t - \lambda t)\right] + [-2(B_t - \lambda t)]\right) \\
&= t M_t^\lambda \left[(B_t - \lambda t)^3 (t^{-1/2})^2 - 3(B_t - \lambda t)\right] \\
&= t\, t^{1/2} t^{-1/2} M_t^\lambda \left[(B_t - \lambda t)^3 (t^{-1/2})^2 - 3(B_t - \lambda t)\right] \\
&= t^{3/2} M_t^\lambda \left[(B_t - \lambda t)^3 (t^{-1/2})^3 - 3(t^{-1/2})(B_t - \lambda t)\right]
\end{aligned}$$

(Estamos haciendo un poquito de trampa para facilitar el inciso siguiente). Piénsese en el factor que se encuentra entre corchetes como un polinomio en $(B_t - \lambda t)(t^{-1/2})$ (en la parte anterior, se trata de un polinomio de grado 3, con sólo los términos de grado 3 y 1, donde los coeficientes son 1 y 3 respectivamente).

Para facilitar la notación, definimos $B_t^\lambda = (B_t - \lambda t)(t^{-1/2})$ y sólo hay que tener presente que $\frac{d}{d\lambda} B_t^\lambda = -t^{1/2}$ y en general, que $\frac{d}{d\lambda}(B_t^\lambda)^n = -t^{1/2} n (B_t^\lambda)^{n-1}$.

Con esta nueva notación, tenemos

$$\begin{aligned}\frac{d}{d\lambda}M_t^\lambda &= t^{1/2}M_t^\lambda\left[B_t^\lambda\right]\\ \frac{d^2}{d\lambda^2}M_t^\lambda &= tM_t^\lambda\left[(B_t^\lambda)^2-1\right]\\ \frac{d^3}{d\lambda^3}M_t^\lambda &= t^{3/2}M_t^\lambda\left[(B_t^\lambda)^3-3B_t^\lambda\right]\end{aligned}$$

Hagamos la siguiente hipótesis de inducción. Para cierta $n \geq 1$, se tiene que

$$\frac{d^n}{d\lambda^n}M_t^\lambda = t^{n/2}M_t^\lambda\left[r_n(B_t^\lambda)^n + r_{n-1}(B_t^\lambda)^{n-1} + \cdots + r_1 B_t^\lambda + r_0\right]$$

Donde $r_0, r_1, \ldots, r_n \in \mathbb{R}$ y $r_n \neq 0$.

Derivemos una vez más para ver qué ocurre con la derivada de orden $n+1$

$$\begin{aligned}&\frac{d^{n+1}}{d\lambda^{n+1}}M_t^\lambda\\ &= t^{n/2}\frac{d}{d\lambda}\left(M_t^\lambda\left[r_n(B_t^\lambda)^n + \cdots + r_1 B_t^\lambda + r_0\right]\right)\\ &= t^{n/2}M_t^\lambda\left(t^{1/2}\left[r_n(B_t^\lambda)^{n+1} + \cdots + r_1(B_t^\lambda)^2 + r_0 B_t^\lambda\right] - t^{1/2}\left[r_n n(B_t^\lambda)^{n-1} + \cdots + r_1\right]\right)\\ &= t^{(n+1)/2}M_t^\lambda\left(\left[r_n(B_t^\lambda)^{n+1} + \cdots + r_1(B_t^\lambda)^2 + r_0 B_t^\lambda\right] - \left[r_n n(B_t^\lambda)^{n-1} + \cdots + r_1\right]\right)\\ &= t^{(n+1)/2}M_t^\lambda\left(\left[r_n(B_t^\lambda)^{n+1} + \cdots + (r_{m-2} - r_m)(B_t^\lambda)^{m-1} + \cdots + r_1(B_t^\lambda)^2 + r_0 B_t^\lambda\right]\right)\end{aligned}$$

En resumen, tenemos la siguiente relación recursiva:

$$\begin{aligned}&\frac{d^{n+1}}{d\lambda^{n+1}}M_t^\lambda\\ &= t^{(n+1)/2}M_t^\lambda\left(\left[r_n(B_t^\lambda)^{n+1} + \cdots + (r_{m-2} - r_m)(B_t^\lambda)^{m-1} + \cdots + r_1(B_t^\lambda)^2 + r_0 B_t^\lambda\right]\right)\end{aligned} \tag{6.4.1.1}$$

De esta fórmula recursiva, vemos que basta que $M_t^\lambda(B_t)^m$ sea integrable (para toda m, es decir, cada sumando de la parte de arriba) para que $\frac{d^n}{d\lambda^n}M_t^\lambda$ lo sea. Veamos que para que $M_t^\lambda(B_t)^m = e^{\lambda B_t}e^{-\lambda^2 t/2}(B_t)^m$ sea integrable, basta que $e^{\lambda B_t}(B_t)^m$ lo sea.

Recordemos entonces que B_t es una variable aleatoria con distribución normal de media 0 y varianza 1, entonces, tiene momentos de todos los ordenes y además su función generadora de momentos $\mathbb{E}(e^{\alpha B_t}) = e^{\frac{1}{2}\alpha^2}$ (véase Normal distribution : Moment and culminant generating functions) es finita para toda α. Haciendo la elección $\alpha = \lambda + 1$ tenemos

$$\mathbb{E}(e^{(\lambda+1)B_t}) = \mathbb{E}(e^{\lambda B_t}e^{B_t})$$

$$
\begin{aligned}
&= \mathbb{E}\left(e^{\lambda B_t} \sum \frac{(B_t)^m}{m!} \right) \\
&= \sum \mathbb{E}\left(e^{\lambda B_t} \frac{(B_t)^m}{m!} \right) \\
&< \infty.
\end{aligned}
$$

De donde se tiene que cada sumando tiene que ser finito y por lo tanto las derivadas de todos los órdenes son integrables.

El mismo argumento aplica para la integrabilidad de $e^{|\lambda B_t|}\ |B_t|^{\ m}$. Entonces, los valores absolutos de las derivadas de cada orden están acotadas por una función integrable (es decir la función $e^{|\lambda B_t|}\ |B_t|^{\ m}$).

Existe un resultado que nos permite intercambiar los operadores derivada y esperanza condicional si el valor absoluto de la derivada de la función está dominado por una función integrable. Entonces ahora podemos hacer:

$$
\mathbb{E}\left(\frac{d^n}{d\lambda^n} M_t^\lambda \middle| \mathscr{F}_s \right) = \frac{d}{d\lambda} \mathbb{E}\left(\frac{d^{n-1}}{d\lambda^{n-1}} M_t^\lambda \middle| \mathscr{F}_s \right)
$$

Donde $s < t$ y $(\mathscr{F}_s)_{s \geq 0}$ es la filtración generada por $\sigma(B_t : t \leq s)$.

Suponiendo la hipótesis de inducción

$$
\mathbb{E}\left(\frac{d^{n-1}}{d\lambda^{n-1}} M_t^\lambda \middle| \mathscr{F}_s \right) = \frac{d^{n-1}}{d\lambda^{n-1}} M_s^\lambda
$$

tendremos que

$$
\mathbb{E}\left(\frac{d^n}{d\lambda^n} M_t^\lambda \middle| \mathscr{F}_s \right) = \frac{d^n}{d\lambda^n} M_s^\lambda
$$

Dado que las derivdas de todos los ordenes son integrables y que las derivadas de M_t^λ son límite de funciones $\mathscr{F}_t$-medibles (y por lo tanto $\mathscr{F}_t$-medibles), basta demostrar la base de inducción para concluir que se trata de una martingala.

La base de inducción la haremos sobre $n = 0$, entonces sean $s < t$

$$
\begin{aligned}
\mathbb{E}\left(\frac{d^0}{d\lambda^0} M_t^\lambda \middle| \mathscr{F}_s \right) &= \mathbb{E}\left(M_t^\lambda \middle| \mathscr{F}_s \right) \\
&= \mathbb{E}\left(e^{-\lambda^2 t/2} e^{\lambda B_t} \middle| \mathscr{F}_s \right) \\
&= e^{-\lambda^2 t/2} \mathbb{E}\left(e^{\lambda B_t} \middle| \mathscr{F}_s \right)
\end{aligned}
$$

$$
\begin{aligned}
&= e^{-\lambda^2 t/2}\mathbb{E}\left(e^{\lambda(B_t - B_s + B_s)}\middle|\mathscr{F}_s\right)\\
&= e^{-\lambda^2 t/2}\mathbb{E}\left(e^{\lambda(B_t - B_s)}e^{\lambda B_s}\middle|\mathscr{F}_s\right)\\
&= e^{-\lambda^2 t/2}e^{\lambda B_s}\mathbb{E}\left(e^{\lambda(B_t - B_s)}\middle|\mathscr{F}_s\right)\\
&= e^{-\lambda^2 t/2}e^{\lambda B_s}\mathbb{E}\left(e^{\lambda B_{t-s}}\middle|\mathscr{F}_s\right)\\
&\quad \text{(los incrementos son estacionarios)}\\
&= e^{-\lambda^2 t/2}e^{\lambda B_s)}\mathbb{E}\left(e^{\lambda B_{t-s}}\right)\\
&\quad \text{(los incrementos son independientes)}\\
&= e^{-\lambda^2 t/2}e^{\lambda B_s}e^{\lambda^2 (t-s)/2}\\
&\quad \text{(de la igualdad de la función generadora de momentos)}\\
&= e^{-\lambda^2 s/2}e^{\lambda B_s}\\
&= M_s^\lambda.
\end{aligned}
$$

Y con esto se termina la demostración.

6.4.2. **Inciso (ii).** *Sea* $H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$. *A* H_n *se le conoce como enésimo polinomio de Hermite. Calcúlelo para* $n \leq 5$. *Pruebe que* H_n *es un polinomio para toda* $n \in \mathbb{N}$ *y que* $\partial^n M_t^\lambda / \partial \lambda^n = t^{n/2} H_n\big(B_t/\sqrt{t}\big)\, M_t^\lambda$.

$$
\begin{aligned}
H_1(x) &= x \\
H_2(x) &= x^2 - 1 \\
H_3(x) &= x^3 - 3x \\
H_4(x) &= x^4 - 6x^2 + 3 \\
H_5(x) &= x^5 - 10x^3 + 15x
\end{aligned}
$$

Supongamos que para cierto $n \in \mathbb{N}$ tenemos que H_n es un polinomio de orden n. Es decir, exísten $r_0, r_1, r_2, \ldots, r_n \in \mathbb{R}$ tales que

$$H_n = r_n x^n + \cdots + r_1 x + r_0.$$

Entonces

$$
\begin{aligned}
H_{n+1} &= (-1)^{n+1} e^{x^2/2} \frac{d^n}{dx^{n+1}} e^{-x^2/2} \\
&= (-1)^{n+1} e^{x^2/2} \frac{d}{dx} \left(\frac{d^n}{dx^n} e^{-x^2/2} \right) \\
&= (-1)^{n+1} e^{x^2/2} \frac{d}{dx} \Big((-1)^n e^{-x^2/2} (r_n x^n + \cdots + r_1 x + r_0) \Big) \\
&\quad \text{(hemos despejado la derivada de orden } n \text{ en terminos del supuesto polinomio } H_n\text{)} \\
&= (-1)^{n+1} e^{x^2/2} (-1)^n \frac{d}{dx} \Big(e^{-x^2/2} (r_n x^n + \cdots + r_1 x + r_0) \Big) \\
&= (-1)^{n+1} e^{x^2/2} (-1)^n \Big(-x e^{-x^2/2} (r_n x^n + \cdots + r_1 x + r_0) + e^{-x^2/2} (r_n n x^{n-1} + \cdots + r_1) \Big) \\
&= (-1)^{n+1} e^{x^2/2} e^{-x^2/2} (-1)^n \Big(-x (r_n x^n + \cdots + r_1 x + r_0) + (r_n n x^{n-1} + \cdots + r_1) \Big) \\
&= (-1)^{n+1} (-1)^n \Big(-x (r_n x^n + \cdots + r_1 x + r_0) + (r_n n x^{n-1} + \cdots + r_1) \Big) \\
&= (-1) \Big(-x (r_n x^n + \cdots + r_1 x + r_0) + (r_n n x^{n-1} + \cdots + r_1) \Big) \\
&= (-1) \Big(-(r_n x^{n+1} + \cdots + r_1 x^2 + r_0 x) + (r_n n x^{n-1} + \cdots + r_1) \Big) \\
&= (r_n x^{n+1} + \cdots + r_1 x^2 + r_0 x) - (r_n n x^{n-1} + \cdots + r_1) \\
&= r_n x^{n+1} + \cdots + (r_{m-2} - m r_m) x^{m-1} + \cdots - r_1
\end{aligned}
$$

Hemos demostrado que si H_n es un polinomio de grado n, entonces H_{n+1} es uno de grado $n+1$. Lo que demostramos al inicio de este inciso no es otra cosa sino la base de inducción. Por lo tanto tenemos que para todo $n \in \mathbb{N}$, H_n es un polinomio de grado n.

Para verificar la equivalencia $d^n M_t^\lambda / d\lambda^n = t^{n/2} H_n\big((B_t - \lambda t)/\sqrt{t}\big)\, M_t^\lambda$, basta sustituir en la ecuación anterior B_t^λ (definido en el inciso anterior) en vez de x, y comparar con (6.4.1.1).

Nótese que está mal escrito el enunciado del problema, en vez de escribir $B_t - \lambda t$ se escribe únicamente B_t. Es necesario hacer la aclaración $d^n M_t^\lambda / d\lambda^n \Big|_{\lambda=0} = t^{n/2} H_n\big((B_t)/\sqrt{t}\big)$ (la cual es fácil de verificar al evaluar lo ya obtenido).

6.4.3. **Inciso (iii).** *Pruebe que* $t^{n/2}H_n\big(B_t/\sqrt{t}\big)$ *es una martingala para toda* n *y calcúlela para* $n \leq 5$.

Que $t^{n/2}H_n\big(B_t/\sqrt{t}\big)$ es una martingala para toda n se sigue de lo dicho en [6.4.1] y [6.4.2].

Las primeras cino son

$$\begin{aligned}
t^{1/2}H_1(B_t/\sqrt{t}) &= t^{1/2}\left[(B_t/\sqrt{t})\right] &&= B_t \\
tH_2(B_t/\sqrt{t}) &= t\left[(B_t/\sqrt{t})^2 - 1\right] &&= B_t^2 - t \\
t^{3/2}H_3(B_t/\sqrt{t}) &= t^{3/2}\left[(B_t/\sqrt{t})^3 - 3(B_t/\sqrt{t})\right] &&= B_t^3 - 3tB_t \\
t^2H_4(B_t/\sqrt{t}) &= t^2\left[(B_t/\sqrt{t})^4 - 6(B_t/\sqrt{t})^2 + 3\right] &&= B_t^4 - 6tB_t^2 + 3t^2 \\
t^{5/2}H_5(B_t/\sqrt{t}) &= t^{5/2}\left[(B_t/\sqrt{t})^5 - 10(B_t/\sqrt{t})^3 + 15(B_t/\sqrt{t})\right] &&= B_t^5 - 10tB_t^3 + t^2B_t.
\end{aligned}$$

6.4.4. **Inciso (iv).** *Aplique muestreo opcional a las martingalas anteriores al tiempo aleatorio* $T_{a,b} = \min\{t \geq 0 : B_t \in \{-a, b\}\}$ *(para* $a, b > 0$*) con* $n = 1, 2$ *para calcular* $\mathbb{P}\big(B_{T_{a,b}} = b\big)$ *y* $\mathbb{E}(T_{a,b})$*, ¿Qué concluye cuando* $n = 3, 4$*? ¿ Cree que* $T_{a,b}$ *tenga momentos finitos de cualquier orden? Justifique su respuesta.*

Primero, trabajemos con $n = 1$. Como de costumbre, trabajaremos con los tiempos acotados $T_{a,b} \wedge s$, donde podemos aplicar el teorema del muestreo opcional de Doob y obtener que $\mathbb{E}(B_{T_{a,b}\wedge s}) = \mathbb{E}(B_0) = 0$.

Ahora, $B_{T_{a,b}\wedge s}$ está acotado por $\max\{a, b\}$ y además $B_{T_{a,b}\wedge s} \longrightarrow B_{T_{a,b}}$ conforme s crece. Así que podemos utilizar el teorema de convergencia acotada para obtener que $\mathbb{E}(B_{T_{a,b}}) = \lim_{s\to\infty} \mathbb{E}(B_{T_{a,b}\wedge s}) = 0$.

Al igual que en el problema de la ruina, como $B_{T_{a,b}} \in \{-a, b\}$ casi seguramente (porque la martingala B oscila), tenemos que

$$\begin{aligned}
0 &= \mathbb{E}(B_{T_{a,b}}) \\
&= -a\mathbb{P}(B_{T_{a,b}} = -a) + b\mathbb{P}(B_{T_{a,b}} = b) \\
&= -a\mathbb{P}(B_{T_{a,b}} = -a) + b(1 - \mathbb{P}(B_{T_{a,b}} = -a)) \\
&= -a\mathbb{P}(B_{T_{a,b}} = -a) + b - b\mathbb{P}(B_{T_{a,b}} = -a)) \\
&= b - (a + b)\mathbb{P}(B_{T_{a,b}} = -a))
\end{aligned}$$

De donde $\mathbb{P}(B_{T_{a,b}} = -a) = \frac{b}{a+b}$ y análogamente $\mathbb{P}(B_{T_{a,b}} = b) = \frac{a}{a+b}$.

De manera complétamente análoga, tomando $n = 2$, la martingala $(B_t)^2 - t$, tenemos que

$$0 = \mathbb{E}((B_{T_{a,b}})^2 - T_{a,b}).$$

Descomponiendo la esperanza tenemos

$$\begin{aligned}
0 &= \mathbb{E}((B_{T_{a,b}})^2 - T_{a,b}) \\
&= a^2\mathbb{P}(B_{T_{a,b}} = -a) + b^2\mathbb{P}(B_{T_{a,b}} = b) - \mathbb{E}(T_{a,b}) \\
&= a^2\frac{b}{a+b} + b^2\frac{a}{a+b} - \mathbb{E}(T_{a,b})
\end{aligned}$$

De donde concluimos que

$$\mathbb{E}(T_{a,b}) = \frac{a^2b + b^2a}{a+b}$$

$$
\begin{aligned}
&= \frac{ab(a+b)}{a+b} \\
&= ab.
\end{aligned}
$$

es finito.
Para $n = 3$, con los mismos argumentos tenemos que

$$
\begin{aligned}
0 &= \mathbb{E}((B_{T_{a,b}})^3 - 3B_{T_{a,b}}T_{a,b}) \\
&= \frac{b^3a - a^3b}{a+b} - 3\mathbb{E}(B_{T_{a,b}}T_{a,b})
\end{aligned}
$$

De donde $\mathbb{E}(B_{T_{a,b}}T_{a,b})$ es finito.
Para $n = 4$, con los mismos argumentos de nuevo, tenemos que

$$
\begin{aligned}
0 &= \mathbb{E}((B_{T_{a,b}})^4 - 6(B_{T_{a,b}})^2T_{a,b} + 3(T_{a,b})^2) \\
&= \mathbb{E}((B_{T_{a,b}})^4) - 6\mathbb{E}((B_{T_{a,b}})^2T_{a,b}) + 3\mathbb{E}((T_{a,b})^2)) \\
&= \frac{a^4b + b^4a}{a+b} - 6\mathbb{E}((B_{T_{a,b}})^2T_{a,b}) + 3\mathbb{E}((T_{a,b})^2))
\end{aligned}
$$

Luego, argumentando que de nuevo que $B_{T_{a,b}} \in \{-a, b\}$ casi seguramente, podemos partir la esperanza $\mathbb{E}((B_{T_{a,b}})^2T_{a,b})$ en los casos $B_{T_{a,b}} = -a$ y $B_{T_{a,b}} = b$ y luego aplicar el resultado del caso $n = 3$ para concluir que $\mathbb{E}((T_{a,b})^2))$ es finito.

De (6.4.1.1) podemos ver que para cada n par, en $H_n(x)$ queda un término constante que en $t^{n/2}H_n(B_tt^{-1/2})$ se convierte en un término de $t^{n/2}$ por una constante. Para el resto de los términos, recursivamente se conoce el momento que resulta ser siempre finito (para los términos del tipo $(B_{T_{a,b}})^j(T_{a,b})^k$ basta con descomponer en los casos $B_{T_{a,b}} = -a$ y $B_{T_{a,b}} = b$ e iterar el procedimiento sobre lo que queda). Despejando, obtendremos que $\mathbb{E}((T_{a,b})^{n/2})$ resulta ser finito y por lo tanto, tenemos momentos de todos los ordenes para $T_{a,b}$.

6.4.5. **Inciso (v).** *Aplique el teorema de muestreo opcional a la martingala* M^λ *al tiempo aleatorio* $T_a = \inf\{t \geq 0 : B_t \geq a\}$ *si* $\lambda > 0$. *Diga por qué es necesaria la última hipótesis y calcule la transformada de Laplace de* T_a

Sabemos que T_a es finito casi seguramente dado que la martingala B oscila.

Aplicando el teorema de muestreo opcional de Doob con el tiempo de paro acotado $T_a \wedge s$ tenemos que

$$\mathbb{E}(M^\lambda_{T_a \wedge s}) = \mathbb{E}(M^\lambda_0) = 1$$

Por otro lado, $M^\lambda_{T_a} = e^{\lambda B_{T_a} - \lambda^2 B_{T_a}/2}$. Argumentando que B empieza en 0, para el caso $s \leq T_a$, tenemos que $B_s \leq B_{T_a} = a$ tenemos

$$\begin{aligned} M^\lambda_{T_a \wedge s} &= M^\lambda_s \\ &= e^{\lambda B_s - \lambda^2 s/2} \\ &= e^{\lambda a - \lambda^2 s/2} \\ &\leq e^{\lambda a}. \end{aligned}$$

Aquí fue necesaria la suposición $\lambda > 0$.

Para el caso $T_a < s$, tenemos

$$\begin{aligned} M^\lambda_{T_a \wedge s} &= M^\lambda_{T_a} \\ &= e^{\lambda a - \lambda^2 T_a/2} \\ &\leq e^{\lambda a}. \end{aligned}$$

De este análisis de casos, tenemos que $M^\lambda_{T_a \wedge s}$ es una martingala acotada por $e^{\lambda a}$. Dado que $M^l_{T_a \wedge s} ambda\lambda \to M^\lambda_{T_a}$, el teorema de convergencia acotada nos dice que

$$\begin{aligned} \mathbb{E}(M^\lambda_{T_a}) &= \lim_{s \to \infty} \mathbb{E}(M^l_{T_a \wedge s} ambda\lambda) \\ &= 1. \end{aligned}$$

Es decir

$$\begin{aligned} \mathbb{E}(M^\lambda_{T_a}) &= \mathbb{E}(e^{\lambda a - \lambda^2 T_a/2}) \\ &= \mathbb{E}(e^{\lambda a} e^{-\lambda^2 T_a/2}) \\ &= e^{\lambda a} \mathbb{E}(e^{-\lambda^2 T_a/2}) \\ &= 1 \end{aligned}$$

De donde $\mathbb{E}(e^{-\lambda^2 T_a/2}) = e^{-\lambda a}$. Para concular la transformada de Laplace de T_a,

$$L_{T_a}(\alpha) = \mathbb{E}(e^{-\alpha T_a})$$

(véase Laplace transform: Formal definition) Haciendo $\alpha = \lambda^2/2$ y por lo tanto, $\lambda = (2\alpha)^{1/2}$ tenemos

$$\begin{aligned} L_{T_a}(\alpha) &= \mathbb{E}(e^{-\alpha T_a}) \\ &= \mathbb{E}(e^{-\lambda^2 T_a/2}) \\ &= e^{-\lambda a} \\ &= e^{-(2\alpha)^{1/2} a}. \end{aligned}$$

6.5. Problema 6.5

Problema 6.5.1.

(1) Al aplicar la desigualdad maximal de Doob sobre los racionales de orden n y pasar al límite conforme $n \to \infty$, pruebe que $\sup_{t\leq} |B_t - B_1|$ es cuadrado integrable.

(2) Pruebe que la sucesión de variables aleatorias

$$\left(\sup_{t\in[0,1]} |B_{n+t} - B_n|, n \in \mathbb{N}\right)$$

son independientes, idénticamente distribuidas y de media finita. (Utilice la propiedad de Markov.)

(3) Al utilizar Borel-Cantelli, pruebe que, para cualquier $C > 0$ fija

$$\limsup_{n\to\infty} \sup_{t\in[0,1]} |B_{n+t} - B_n| / n \leq C$$

casi seguramente.

(4) Pruebe que $(B_n/n, n \geq 1)$ converge casi seguramente a 0 y deduzca que

$$\lim_{t\to\infty} B_t/t = 0.$$

Demostración:

6.5.1. **Inciso (i).** *Al aplicar la desigualdad maximal de Doob sobre los racionales de orden n y pasar al límite conforme* $n \to \infty$*, pruebe que* $\sup_{t\leq 1} |B_t - B_1|$ *es cuadrado integrable.*

La idea es partir el proceso en martingalas finitas que al irse refinando nos den un conjunto denso en $[0,1]$ y luego utilizar el argumento de continuidad para llenar los huecos.

Llamemos

$$D_n = \left\{\frac{i}{2_n} : i = 0, 1, \ldots, 2^n\right\}$$

Con estos conjuntos, discretizamos a B haciendo las martingalas finitas $(B_j^n : j \in D_n)$ [donde $B_j^n = B_j$]. Por propiedades de la esperanza condicional y el hecho de que $(B_j^n : j \in D_n)$ son martingalas. Tenemos que $(\left|B_j^n\right| : j \in D_n)$ son sub-marginalas.

La desigualdad maximal de Doob para el caso L_2 nos dice que

$$\| \max_{j \in D_n} \left|B_j^n\right| \|_2 \leq 2\|B_1^n\|_2$$

De donde, elevando al cuadrado y usando que $\mathbb{E}(B_1^n) = 0$.

$$\begin{aligned}
\mathbb{E}\left(\left(\max_{j \in D_n} \left|B_j^n\right|\right)^2\right) &= 4\mathbb{E}((B_1^n)^2) \\
&= 4Var(B_1^n) \\
&= 4Var(B_1) \\
&= 4.
\end{aligned}$$

Ahora tenemos que conforme $n \to \infty$, $\max_{j \in D_n} \left|B_j^n\right|$ converge a $\sup |B_t|$ c.s. El teorema de convergencia monótona nos asegura que

$$E\left((\sup |B_t|)^2\right) = \lim_{n \to \infty} \mathbb{E}\left(\left(max_{j \in D_n} \left|B_j^n\right|\right)^2\right)$$

Como cada término del límite de la derecha está acotado por 4 por la demostración de arriba.

Entonces $\sup_{t \leq 1} |B_t - B_1|$ es cuadrado integrable.

6.5.2. **Inciso (ii).** *Pruebe que la sucesión de variables aleatorias*

$$\left(\sup_{t\in[0,1]} |B_{n+t} - B_n| \,, n \in \mathbb{N}\right)$$

son independientes, idénticamente distribuidas y de media finita. (Utilice la propiedad de Markov.)

La propiedad de Markov nos asegura que para todo $n \in N$ el proceso dado obtenido "desfasar" el proceso en n es independiente del pasado, es decir: $B_{n+t} - B_n$ es independiente de $\mathscr{F}_n = \sigma(B_t : t \leq n)$ y por lo tanto también $|B_{n+t} - B_n|$ es independiente de $\mathscr{F}_n$ para toda n y por lo tanto son independientes.

Además, sabemos que desfasar es únicamente volver a empezar el proceso, es decir, $B_{n+t} - B_n \sim B_t$ y entonces $|B_{n+t} - B_n| \sim |B_t|$. Tomando supremos $\sup_{t\leq 1} |B_{n+t} - B_n| \sim \sup_{t\leq 1} |B_t|$. Es decir que $\sup_{t\leq 1} |B_{n+t} - B_n|$ tienen todas la misma distribución que $\sup_{t\leq 1} |B_t|$.

El en [6.5.1] se demostró que $\sup_{t\leq 1} |B_t - B_1|$ es cuadrado integrable, y por lo tanto tiene medida finita. Ese es el caso cuando $n = 1$, pero como para toda n se tiene la misma distribución, todas tienen media finita.

6.5.3. **Inciso (iii.** *Al utilizar Borel-Cantelli, pruebe que, para cualquier $C > 0$ fija*

$$\limsup_{n\to\infty} \sup_{t\in[0,1]} |B_{n+t} - B_n| / n \leq C$$

casi seguramente.

Sea $C > 0$ y sean $A_n = \left\{\sup_{t\in[0,1]} |B_{n+t} - B_n| \leq nC\right\}$. Entonces tenemos $A_n^c = \left\{\sup_{t\in[0,1]} |B_{n+t} - B_n| > nC\right\}$ y por desigualdad de Markov.

$$\begin{aligned}\mathbb{P}(A_n^c) &= \mathbb{P}(\sup_{t\in[0,1]} |B_{n+t} - B_n| > nC) \\ &= \frac{\mathbb{E}(\sup_{t\in[0,1]} |B_{n+t} - B_n|)}{nC}\end{aligned}$$

De [6.5.2] tenemos que para todo $n \in \mathbb{N}$, $\mathbb{E}(\sup_{t\in[0,1]} |B_{n+t} - B_n|) = r < \infty$ y por lo tanto $\lim_{n\to\infty} \mathbb{P}(A_n^c) = 0$. Es decir que $\lim_{n\to\infty} \mathbb{P}(A_n) = 1$.

Una condición necesaria para la convergencia de una serie es que el límite de la sucesión sea 0. Por lo tanto la serie $\sum \mathbb{P}(A_n)$ diverge. Borel-Cantelli nos dice que entonces $\limsup_n A_n$ tiene probabilidad uno, o sea que con probabilidad uno, $\sup_{t\in[0,1]} |B_{n+t} - B_n| \leq nC$ para una cantidad infinita de índices. Lo que quieres es que sea a partir de un cierto índice, por lo que el argumento no es correcto.

De [6.5.2] tenemos que los A_n son independientes. Y por lo tanto, Borel-Cantelli nos asegura que

$$\mathbb{P}(\limsup A_n) = 1.$$

Con lo que queda demostrado que

$$\limsup_{n\to\infty} \sup_{t\in[0,1]} |B_{n+t} - B_n| / n \leq C$$

casi seguramente. Es decir

$$\limsup_{n\to\infty} \sup_{t\in[0,1]} |B_{n+t} - B_n| / n = 0.$$

6.5.4. **Inciso (iv).** *Pruebe que* $(B_n/n, n \geq 1)$ *converge casi seguramente a* 0 *y deduzca que*

$$\lim_{t\to\infty} B_t/t = 0.$$

Escribimos

$$B_n = \sum_{i\leq n}(B_i - B_{i-1})$$

Recordemos que gracias a [6.5.2] tenemos que los sumandos de la derecha son variables aleatorias idependientes e idénticamente distribuidas y que además, su distribución es normal de media 0 y varianza 1.

De aquí, aplicando ley fuerte de los grandes números, tenemos que

$$\begin{aligned}
\lim_{n\to\infty}\frac{B_n}{n} &= \lim_{n\to\infty}\frac{\sum_{i\leq n}(B_i - B_{i-1})}{n} \\
&= \mathbb{E}(B_1) \\
&= 0.
\end{aligned}$$

Sean ahora $t \geq 0$, $n = \lfloor t \rfloor$ y $s = t - n$ y entonces acotamos a B_t/t por algo que ya conozcamos.

$$\begin{aligned}
\frac{|B_t|}{t} &= \frac{|B_{n+s}|}{n+s} \\
&\leq \frac{|B_{n+s}|}{n} \\
&= \frac{|B_{n+s} - B_n + B_n|}{n} \\
&= \frac{|B_{n+s} - B_n|}{n} + \frac{|B_n|}{n} \\
&= \frac{\sup_{s\leq 1}|B_{n+s} - B_n|}{n} + \frac{|B_n|}{n}
\end{aligned}$$

Ahora tomaremos límite superior, el primer sumando se hace 0 por la conclusión de [6.5.3], y el sumando de la derecha tiene límite y es 0, por lo tanto tiene límite superior y también es 0.

•___•••_•___•_•_•__•____•__••_••••_••••••__••_•_•___••_••__••••_•___•_••••_••••••__••_••

Part 7. Parcial 2

Los problemas 1 y 2 se encuentran resueltos en [5.3.2] y [5.7].

7.1. Problema 3

Sean $B_1, B_2, \ldots$ variables aleatorias independientes con distribución Bernoulli de parámetro $p \in (0,1)$. Sea $T_0 = 0$ y

$$T_{n+1} = \min\{k > T_n : B_k = 1\}$$

7.1.1. Inciso 1.

Pruebe que $T_0, T_1, \ldots$ son tiempos de renovación de un proceso de renovación aritmético.

Pensemos en $B_1, B_2, \ldots$ como volados con monedas cargadas donde $B_i = 1$ significa que ganamos el i-ésimo volado.

Con esta interpretación, T_n no es otra cosa sino el tiempo en que ganamos por n-ésima vez. Por lo tanto

$$\mathbb{P}(T_n = k) = \binom{k-1}{n-1} p^n (1-p)^{k-n}$$

Es decir, T_n tiene distribución binomial negativa de parámetros n y p.

Definimos entonces a los tiempos de vida $S_0 = 0$ y $S_n = T_n - T_{n-1}$. Por tener T_n y T_{n-1} distribución binomial negativa de parámetros n, p y $n-1, p$ respectivamente, tenemos que S_n tiene distribución geométrica de parámetro p.

Otra manera de escribir a S es

$$S_n = \sum_{i=T_{n-1}+1}^{T_n} iB_i$$

De donde tenemos que S_n es suma de variables aleatorias independientes, todas distintas de las que suman S_{n-1}. Por lo tanto las S_n son independientes y dado que S_n tiene distribución geométrica de parámetro p para toda n, tenemos que son idénticamente distribuidas.Se debe verificar que $B_{T_n+1}, B_{T_n+2}, \ldots$ son independientes de T_n y entre si y de distribución Bernoulli de parámetro p.

Con esto, tenemos un proceso de renovación con tiempos de vida S_n y tiempos de renovación T_n.

7.1.2. **Inciso 2.** *Defina al proceso de contéo asociado N, expréselo en términos de las variables $B_1, B_2, \ldots$ y pruebe diréctamente que N_n/n converge casi seguramente conforme n crece. Identifíque el límite.*

Sea
$$N_n = \sum_{i=1}^{n} B_i.$$
Y notemos que
$$\min\{m : T_{m+1} > n\}$$
significa, el tiempo de volado más alto antes de pasarnos de n.

Interpretado de esta forma tenemos que coinciden
$$\min\{m : T_{m+1} > n\} = \sum_{i=1}^{n} B_i = N_n.$$

Por lo tanto tenemos que N_n sí es un proceso de conteo para nuestro proceso de renovación.

Por otra parte N_n es suma de n variables idependientes e idénticamente distribuidas con esperanza p. N_n/n es su promedio y la ley fuerte de los grandes números nos dice que
$$N_n/n \to p \;\; c.s.$$

7.1.3. **Inciso 3.** *Calcule la distribución del tiempo residual al tiempo n*

$$R_n = \min\{k > n : B_k = 1\} - n$$

Asuma que dicho proceso es una cadena de Markov e identifique las probabilidades de transición. Pruebe que la distribución geométrica de parámetro p es invariante para dicho proceso.

$R_n = k$ significa que si han pasado n volados, en k volados ganaremos uno, y antes de eso ninguno. Dado que los B_i son idénticamente distribuidos, tenemos que eso significa que

$$\mathbb{P}(R_n = k) = p(1-p)^r - 1-1?$$

Creo que faltaron paréntesis en el superíndice.

Es decir, R_n tiene distribución geométrica de parámetro p.

Supongamos ahora que R_n es una cadena de Markov. Y supongamos que al tiempo n, sabemos que $R_n = k \neq 0$, es decir, nos faltan k volados para volver a ganar. Entonces $R_{n+1} = k-1$ forzosamente.

Entonces, en nuestra matriz de transición tendremos $P_{k,k-1} = 1$ cuando $k \neq 0$.

Para el caso $R_n = 0$, la probabilidad de que pasemos de 0 a $k > 0$, es equivalente a perder los siguientes $k-1$ volados y luego ganar el k-ésimo. Es decir, $P_{0,k} = p(1-p)^{k-1}$.

Entonces la matriz de transición está determinada por

$$P_{i,j} = \begin{cases} 1 & \text{si } i > 0 \text{ y } j = i-1 \\ p(1-p)^{j-1} & \text{si } i = 0 \text{ y } j > 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

Veamos que la distribución π, con $\pi_i = p(1-p)^{i-1}$ si $i > 0$ y $\pi_0 = 0$ resulta invariante para este proceso.

$$\begin{aligned} (\pi \times P)_j &= \sum_{k \geq 0} \pi_k P_{k,j} \\ &= \pi_0 P_{0,j} + \pi_j P_{j+1,j} \\ &= 0 + \pi_j 1. \end{aligned}$$

Con lo que terminamos que π es una distribución invariante para el proceso.

7.1.4. **Inciso 4.** *Calcule la distribución del proceso de edad al tiempo n:*

$$A_n = n - \max k \leq n : B_k = 1$$

Calcula la distribución límite de A_n conforme $n \to \infty$.

Supongamos $r < n$. Entonces $A_n = r$ significa que hace r volados ganamos y que desde entonces hemos perdido. Eso quiere decir que $\mathbb{P}(A_n = r) = p(1-p)^r$. Es decir que para el caso $r < n$, A_n se comporta como una variable con distribución geométrica de parámetro p.

Para el caso $r = n$. $A_n = r$ significa que nunca hemos ganado un volado. Por lo tanto $\mathbb{P}(A_n = r) = (1-p)^r$

Recordemos para el caso $n < r$, tenemos que $\mathbb{P}(A_n = r) = 0$, pues estaríamos preguntando por la probabilidad de haber perdido más veces que el número de volados jugados.

Eso quiere decir que $A_n \sim Geo(p) \wedge n$.

De esta última fórmula es fácil verificar que $A_n \to A_\infty \sim Geo(p)$.

7.2. Problema 4

Una matriz de transición P se llama reversible si existe una distribución inicial π tal que

$$\pi_i \mathbb{P}_{i,j} = \pi_j P_{j,i}.$$

A las identidades anteriores se les conoce como condición de balance detallado.

7.2.1. **Inciso 1.** *Muestre que si π y P satisfacen la condición de balance detallado entonces π es una distribución invariante para P.*

Sencillamente verifiquemos la igualdad necesaria

$$\begin{aligned}
(\pi \times P)_j &= \sum i \in E \pi_i P_{i,j} \\
&= \sum i \in E \pi_j P_{j,i} \\
&= \pi_j \sum i \in E P_{j,i} \\
&= \pi_j(1) \\
&= \pi_j.
\end{aligned}$$

Y con esto termina la demostración.

7.2.2. **Inciso 2.** *Muestre que si X es una cadena de Markov con distribución inicial π y matriz de transición P en balance detallado entonces*

$$(X_0, \dots, X_n) \sim (X_n, \dots, X_0).$$

Por ser cadena de Markov con distribución inicial π, tenemos que

$$\mathbb{P}(X_0 = x_0, \dots, X_n = x_n) = \pi_{x_0} P_{x_0,x_1} P_{x_1,x_2} \cdots P_{x_{n-1},x_n}.$$

Utilizando la hipótesis de que π y P están en balance detallado tenemos que

$$\begin{aligned}
&\pi_{x_0} P_{x_0,x_1} P_{x_1,x_2} \cdots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} \pi_{x_1} P_{x_1,x_2} \cdots P_{x_{n-1},x_n} \\
&= P_{x_1,x_0} P_{x_2,x_1} \pi_{x_2} \cdots P_{x_{n-1},x_n} \\
&\vdots \\
&= P_{x_1,x_0} P_{x_2,x_1} \cdots P_{x_n,x_{n-1}} \pi_{x_n} \\
&= P(X_n = x_n, \cdots, X_0 = x_0).
\end{aligned}$$

Con lo que termina la demostración.

7.2.3. **Inciso 3.** *Sea $G = (V, E)$ una gráfica y defina la matriz de transición P sobre V al estipular que $P_{x,y}$ es igual a 1 sobre el grado de x, denotado como δ_x. Muestre que la distribución π Dada por $\pi_x = \frac{\delta_x}{\sum_{v\in V}\delta_v}$.*

Si x, y son tales que $\{x, y\} \notin E$, tenemos que $P_{x,y} = 0 = P_{y,x}$. Por lo tanto la condición de balance se da trivialmente en estos casos.

Para, x, y tales que $\{x, y\} \in E$, sencillamente verifiquemos la igualdad pertinente

$$\begin{aligned}
\pi_x P_{x,y} &= \frac{\delta_x}{\sum_{v\in V}\delta_v}\frac{1}{\delta_x} \\
&= \frac{1}{\sum_{v\in V}\delta_v} \\
&= \frac{\delta_y}{\delta_y}\frac{1}{\sum_{v\in V}\delta_v} \\
&\quad \text{(es válido porque } \delta_y \geq 1\text{, pues } x \text{ es vecino de } y) \\
&= \frac{1}{\delta_y}\frac{\delta_y}{\sum_{v\in V}\delta_v} \\
&= P_{y,x}\pi_y.
\end{aligned}$$

Con lo que termina la demostración.

7.2.4. **Inciso 5.** *Muestre que si $\tilde{X}_n = X_{-n}$, entonces $\tilde{X}$ y X tienen la misma distribución bajo $\mathbb{P}$.*

Dadas las condiciones de balance detallado, y la "condición de Markov" dada para $\mathbb{P}$ en el inciso anterior, este ejercicio es exáctamente el mismo que el del segundo inciso de este problema y la demostración es completamente análoga (la nombrada "condición de Markov" fue la única parte de ser cadena de Markov que se utilizó en 2).

•___•••_•___•_•_•__•____•__••_••••_••••••__••__••_•_•___••_••__••••_•___•_••••_••••••__••_••

Referencias

Notas

Se ha incluido una versión de las notas del curso como archivo adjunto en este `pdf`. No todos los lectores de `pdf` soportan archivos adjuntos o no se comportan como es esperado. Adobe Reader y Evince (versiones para Windows) no se comportan como es esperado, pero sí son capaces de accesar a los archivos adjuntos de alguna manera (la manera comprobada [en Windows] es extraer el archivo diciendo "guardar como" y luego, como se hace con cualquier otro `pdf`, abrir el archivo recién creado).

El motivo de adjuntar el archivo, es porque se han hecho algunas citas a dichas notas. Si las notas cambian (lo cual es probable), es posible que las referencias dejen de ser correctas.

Si no es posible recompilar este documento gracias a la librería para manejar archivos adjuntos, comente las lineas 20 y 100 del archivo principal (`tareas.tex`) y vuelva a intentar.

Para el creador de las notas: Sería muy útil antes de cada sección mencionar en qué otra fuente se puede consultar el material de la sección. A veces se omiten detalles que para los novatos en probabilidad (como yo) no son obvios. Personalmente, añadiría recomendaciones a los libros de $J.\,R.\,Norris$ y $Krishna\,B.\,Athreya$.