



Procesos estocásticos

Maria Jaimes¹, Carlos Nosa¹

¹Facultad de Ciencias, Departamento de Matemáticas
Universidad Nacional de Colombia

2025

Contenidos

1. Probabilidad
2. Introducción a procesos estocásticos
3. Caminatas aleatorias
4. Cadenas de Markov
5. Procesos de Poisson
6. Cadenas de Markov a tiempo continuo
7. Procesos de renovación y confiabilidad
8. Martingalas
9. Movimiento browniano
10. Procesos Gaussianos
11. Cálculo estocástico
12. Referencias

Probabilidad

Espacio muestral

Definición

El conjunto Ω de todos los posibles resultados de un experimento aleatorio se llama espacio muestral. Los elementos $\omega \in \Omega$ son llamados puntos muestrales. Un experimento se dice aleatorio si su resultado no puede ser determinado de antemano.

Ejemplo.

Experimento: lanzamiento de una moneda corriente. Los posibles resultados, en este caso, son “cara” = c y “sello” = s . Esto es,

$$\Omega = \{c, s\}$$

Experimento: se observa el número de veces que es necesario lanzar una moneda corriente hasta obtener por primera vez cara. En este caso:

$$\Omega = \{1, 2, \dots, \infty\}$$

Definición

Sea $\Omega \neq \Phi$. Una colección $\mathfrak{S}$ de subconjuntos de Ω es una σ -álgebra sobre Ω , si:

- i) $\Omega \in \mathfrak{S}$.
- ii) Si $A \in \mathfrak{S}$ entonces $A^c \in \mathfrak{S}$.
- iii) Si $A_1, A_2, \dots \in \mathfrak{S}$ entonces $\bigcup_{i=1}^{\infty} A_i \in \mathfrak{S}$.

Los elementos de $\mathfrak{S}$ se llaman eventos.

Ejemplo

Sea $\Omega \neq \emptyset$. Entonces $\mathfrak{S}_0 = \{\Phi, \Omega\}$ y $\wp(\Omega) := \{A : A \subseteq \Omega\}$ son σ -álgebras sobre Ω . $\mathfrak{S}_0$ recibe el nombre de σ -álgebra trivial y $\wp(\Omega)$ recibe el nombre de σ -álgebra total.

Ejercicio

Probar que si $\Omega \neq \emptyset$ y $\mathfrak{S}_1, \mathfrak{S}_2, \dots$ son σ -álgebras sobre Ω entonces $\bigcap_{i=1}^{\infty} \mathfrak{S}_i$ es una σ -álgebra sobre Ω .

σ -álgebra

σ -álgebra generada

Sean $\Omega \neq \Phi$ y L una colección de subconjuntos de Ω . Sea $\mathcal{M} = \{\mathfrak{S} : \mathfrak{S} \text{ es una } \sigma\text{-álgebra sobre } \Omega \text{ que contiene a } L\}$. Entonces, se tiene que $\sigma(L) := \bigcap_{\mathfrak{S} \in \mathcal{M}} \mathfrak{S}$ es la menor σ -álgebra sobre Ω que contiene a L . Esta σ -álgebra se llama σ -álgebra generada por L .

σ -álgebra de Borel

La menor σ -álgebra sobre $\mathbb{R}$ que contiene todos los intervalos de la forma $(-\infty, a]$ con $a \in \mathbb{R}$ se llama σ -álgebra de Borel y se denota por $\mathcal{B}$.

σ -álgebra de Borel en $\mathbb{R}^n$

Sean $a = (a_1, \dots, a_n)$ y $b = (b_1, \dots, b_n)$ elementos de $\mathbb{R}^n$ con $a \leq b$, es decir, $a_i \leq b_i$ para todo $i = 1, \dots, n$. La σ -álgebra generada por todos los intervalos de la forma $(a, b] := \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : a_i < x_i \leq b_i, i = 1, \dots, n\}$ se llama σ -álgebra de Borel en $\mathbb{R}^n$.

Espacio medible

Definición

Sean $\Omega \neq \emptyset$ y $\mathfrak{F}$ una σ -álgebra sobre Ω . La pareja $(\Omega, \mathfrak{F})$ se llama espacio medible.

Decir que un evento A ocurre significa que el resultado obtenido, al realizar el experimento aleatorio cuyo espacio muestral es Ω , es un elemento de A . Por lo tanto si A y B son eventos entonces:

1. $A \cup B$ es un evento que ocurre, si y sólo si, A o B o ambos ocurren.
2. $A \cap B$ es un evento que ocurre, si y sólo, si A y B ocurren.
3. A^c es un evento que ocurre, si y sólo si, A no ocurre.
4. $A - B$ es un evento que ocurre, si y sólo si, A ocurre pero B no.

Espacio de probabilidad

Sea $(\Omega, \mathfrak{S})$ un espacio medible. Una función P definida sobre $\mathfrak{S}$ y de valor real que satisface las siguientes condiciones:

1. $P(A) \geq 0$ para todo $A \in \mathfrak{S}$.
2. $P(\Omega) = 1$
3. Si $A_1, A_2, \dots$ son elementos de $\mathfrak{S}$ mutuamente excluyentes, esto es

$$A_i \cap A_j = \emptyset \text{ para todo } i \neq j,$$

entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

se llama medida de probabilidad sobre $(\Omega, \mathfrak{S})$. La tripla $(\Omega, \mathfrak{S}, P)$ se llama espacio de probabilidad.

Espacio de probabilidad

Ejemplo. Sean $\Omega = \{1, 2\}$, $\mathfrak{F} = \wp(\Omega)$ y P la aplicación definida sobre $\mathfrak{F}$ por:

$$P(A) = \begin{cases} 0 & \text{si } A = \emptyset \\ \frac{1}{3} & \text{si } A = \{1\} \\ \frac{2}{3} & \text{si } A = \{2\} \\ 1 & \text{si } A = \{1, 2\} \end{cases}$$

P es una medida de probabilidad.

Espacio de probabilidad

Teorema. Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad. Entonces:

1. $P(\emptyset) = 0$.
2. Si $A, B \in \mathfrak{F}$ y $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$.
3. $P(A^c) = 1 - P(A)$.
4. Si $A \subseteq B$ entonces $P(A) \leq P(B)$ y $P(B - A) = P(B) - P(A)$. En particular, se tiene que $P(A) \leq 1$ para todo $A \in \mathfrak{F}$.
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Probabilidad condicional

Definición

Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad. Si $A, B \in \mathfrak{F}$ con $P(A) > 0$, entonces se define la probabilidad del evento B bajo la condición de A como sigue:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

Medida de probabilidad condicional

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $A \in \mathcal{F}$ con $P(A) > 0$. Entonces:

1. $P(\cdot | A)$ es una medida de probabilidad sobre Ω , que está concentrada en A , esto es, $P(A | A) = 1$.
2. Si $A \cap B = \emptyset$ entonces $P(B | A) = 0$.
3. $P(B \cap C | A) = P(B | A \cap C) P(C | A)$, si $P(A \cap C) > 0$.
4. Si $A_1, A_2, \dots, A_n \in \mathcal{F}$ con $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$, entonces

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1 \cap A_2) \dots P(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}).$$

Teorema de probabilidad total

Sea $A_1, A_2, \dots$ una partición finita o numerable de Ω , es decir, $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \Omega$, tal que $P(A_i) > 0$ para todo i . Entonces, para cualquier $B \in \mathcal{F}$, se satisface

$$P(B) = \sum_n P(B \mid A_n) P(A_n).$$

Demostración.

$$\begin{aligned} P(B) &= P(B \cap \Omega) \\ &= P\left(\bigcup_n (B \cap A_n)\right) \\ &= \sum_n P(B \cap A_n) \\ &= \sum_n P(B \mid A_n) P(A_n). \end{aligned}$$

Eventos independientes

Definición

Dos eventos A y B son independientes, si y sólo si,

$$P(A \cap B) = P(A)P(B).$$

Sean A y B eventos independientes. Entonces:

1. A y B^c son independientes.
 2. A^c y B son independientes.
 3. A^c y B^c son independientes.
- ★ ¿Cómo probar estos resultados?

Ejercicio

Sean Ω y $\tilde{\Omega}$ conjuntos no vacíos, $\tilde{\mathfrak{S}}$ una σ -álgebra en $\tilde{\Omega}$ y $T : \Omega \rightarrow \tilde{\Omega}$ una función. Demostrar que

$$T^{-1}(\tilde{\mathfrak{S}}) := \{ T^{-1}(\tilde{A}) : \tilde{A} \in \tilde{\mathfrak{S}} \}$$

es una σ -álgebra en Ω .

Ejercicio

Sean $(P_n)_{n \in \mathbb{N}}$ una sucesión de medidas de probabilidad definidas sobre el espacio medible $(\Omega, \mathfrak{F})$, $(\alpha_n)_{n \in \mathbb{N}}$ una sucesión de números reales no negativos tales que $\sum_{n=1}^{\infty} \alpha_n = 1$ y $P : \mathfrak{F} \rightarrow \mathbb{R}$ definida por:

$$P(A) := \sum_{n \in \mathbb{N}} \alpha_n P_n(A) \quad \text{para todo } A \in \mathfrak{F}$$

Demostrar que P es una medida de probabilidad definida sobre $(\Omega, \mathfrak{F})$.

Ejercicio

Los coeficientes a, b y c de la ecuación cuadrática $ax^2 + bx + c = 0$ son determinados lanzando un dado corriente tres veces consecutivas. ¿A qué es igual la probabilidad de que las dos raíces de la ecuación sean reales? ¿A qué es igual la probabilidad de que las dos raíces sean complejas?

Ejercicio

Un número aleatorio N de dados corrientes es lanzado. Sea A_i el evento de que $N = i$. Suponga que $P(A_i) = 2^{-i}$ con $i \geq 1$. Sea S la suma de los resultados. Determinar la probabilidad de que:

- $N = 2$ dado que $S = 4$.
- $S = 4$ dado que N es par.
- $N = 2$ dado que $S = 4$ y el resultado del primer dado es 1.

Variable aleatoria

Definición

Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y $(\tilde{\Omega}, \tilde{\mathfrak{F}})$ un espacio medible.

Una $\mathfrak{F}$ - $\tilde{\mathfrak{F}}$ -variable aleatoria es una aplicación $X : \Omega \longrightarrow \tilde{\Omega}$ tal que, para todo $A \in \tilde{\mathfrak{F}}$ se tiene que $X^{-1}(A) \in \mathfrak{F}$. Si $(\tilde{\Omega}, \tilde{\mathfrak{F}}) = (\mathbb{R}, \mathcal{B})$, entonces, se dice que X es una variable aleatoria real.

Ejemplo. Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad y $A \in \mathfrak{F}$ fijo. La aplicación $\mathcal{X}_A : \Omega \longrightarrow \mathbb{R}$ dada por:

$$\mathcal{X}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$$

es una variable aleatoria real. En efecto, si $\alpha \in \mathbb{R}$ entonces:

$$\{\omega \in \Omega : \mathcal{X}_A(\omega) \leq \alpha\} = \begin{cases} \emptyset & \text{si } \alpha < 0 \\ A^c & \text{si } 0 \leq \alpha < 1 \\ \Omega & \text{si } \alpha \geq 1 \end{cases}$$

La aplicación $\mathcal{X}_A$ se llama la función indicadora de A . Otras notaciones de uso frecuente, para esta aplicación, son: I_A y 1_A .

Distribución de una variable aleatoria

Definición

Supóngase que X es una variable aleatoria definida sobre el espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ y con valores en el espacio medible $(\tilde{\Omega}, \tilde{\mathfrak{F}})$. La función P_X definida sobre la σ -álgebra $\tilde{\mathfrak{F}}$ por medio de:

$$P_X(B) := P(\{X \in B\}); \text{ para todo } B \in \tilde{\mathfrak{F}}$$

es una medida de probabilidad sobre $(\tilde{\Omega}, \tilde{\mathfrak{F}})$ llamada distribución de la variable aleatoria X .

Demostración. Se debe verificar que P_X satisface las tres condiciones que definen una medida de probabilidad.

1. Es claro que para todo $B \in \tilde{\mathfrak{F}}, P_X(B) \geq 0$.
2. $P_X(\tilde{\Omega}) = P(\{\omega \in \Omega : X(\omega) \in \tilde{\Omega}\}) = P(\Omega) = 1$.
3. Sean $A_1, A_2, \dots$ elementos de $\tilde{\mathfrak{F}}$ con $A_i \cap A_j = \Phi$ para todo $i \neq j$. Entonces,

$$\begin{aligned} P_X \left(\bigcup_{i=1}^{\infty} A_i \right) &= P \left(\left\{ X \in \bigcup_{i=1}^{\infty} A_i \right\} \right) \\ &= P \left(\bigcup_{i=1}^{\infty} \{X \in A_i\} \right) \\ &= \sum_{i=1}^{\infty} P(\{X \in A_i\}) \\ &= \sum_{i=1}^{\infty} P_X(A_i) \end{aligned}$$

Función de distribución de una variable aleatoria

Definición

Sea X una variable aleatoria real. La función F_X definida sobre $\mathbb{R}$ por medio de:

$$\begin{aligned} F_X(x) &:= P_X((-\infty, x]) \\ &= P(X \leq x) \end{aligned}$$

se llama función de distribución (o de distribución acumulativa) de la variable aleatoria X .

Ejemplo. Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad arbitrario y $A \in \mathfrak{F}$ fijo. La función de distribución F de la variable aleatoria $\mathcal{X}_A$ está dada por:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ P(A^c) & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Corolario. Sean X una variable aleatoria real definida sobre $(\Omega, \mathfrak{F}, P)$, F_X su función de distribución y $a, b \in \mathbb{R}$ con $a < b$, entonces:

1. $F_X(x^-) := \lim_{h \rightarrow 0^+} F_X(x - h) = P(X < x)$
2. $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$
3. $P(a < X \leq b) = F_X(b) - F_X(a)$
4. $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$
5. $P(a < X < b) = F_X(b^-) - F_X(a)$
6. $P(X = a) = F_X(a) - F_X(a^-)$
7. Si $P(a < X < b) = 0$ entonces F_X es constante en el intervalo (a, b) .

Variable aleatoria discreta

Definición

Sean X una variable aleatoria real y F_X su función de distribución. Se dice que F_X presenta un salto en el punto $a \in \mathbb{R}$ si

$$F_X(a) - F_X(a^-) \neq 0$$

La diferencia $F_X(a) - F_X(a^-)$ se llama magnitud del salto y por las propiedades desarrolladas anteriormente se tiene que es igual a $P(X = a)$.

Si X es una variable aleatoria real continua entonces el conjunto de saltos de F_X es el conjunto vacío.

Función de densidad discreta

Definición

Sea X una variable aleatoria real discreta con valores $x_1, x_2, \dots$ (todos diferentes). La función f_X , definida sobre $\mathbb{R}$, mediante:

$$f_X(x) = \begin{cases} P(X = x_i) & \text{si } x = x_1, x_2, \dots \\ 0 & \text{en otro caso} \end{cases}$$

se llama función de densidad de la variable aleatoria X .

Se tiene que para cualquier $x \in \mathbb{R}$ se satisface:

$$P(X = x) = F_X(x) - F_X(x^-),$$

esto es, la función de distribución de la variable aleatoria X determina completamente su función de densidad.

Variables aleatorias continuas

Definición

Sea X una variable aleatoria real definida sobre el espacio de probabilidad $(\Omega, \mathfrak{F}, P)$. Se dice que X es absolutamente continua, si y sólo si, existe una función real no negativa e integrable f_X tal que, para todo $x \in \mathbb{R}$, se satisface:

$$F_X(x) = \int_{-\infty}^x f_X(t) dt$$

La función f_X recibe el nombre de función de densidad (fdd) de la variable aleatoria X .

Supóngase que X es una variable aleatoria absolutamente continua con función de densidad f_X . Se tiene:

$$\begin{aligned}\int_{-\infty}^{\infty} f_X(t) dt &= \lim_{x \rightarrow \infty} \int_{-\infty}^x f_X(t) dt \\ &= \lim_{x \rightarrow \infty} F_X(x) = 1\end{aligned}$$

Si $a, b \in \mathbb{R}$ entonces:

$$\begin{aligned}P(a < X \leq b) &= F_X(b) - F_X(a) \\ &= \int_{-\infty}^b f_X(x) dx - \int_{-\infty}^a f_X(x) dx \\ &= \int_a^b f_X(x) dx\end{aligned}$$

Valor esperado

Sea X una variable aleatoria real definida sobre el espacio de probabilidad $(\Omega, \mathfrak{S}, P)$.

1. Si X es una variable aleatoria discreta, con valores $x_1, x_2, \dots$, se dice que X posee un valor esperado si:

$$\sum_{k=1}^{\infty} |x_k| P(X = x_k) < \infty$$

En tal caso, se define el valor esperado $E(X)$ (esperanza matemática, media) de X como:

$$E(X) = \sum_{k=1}^{\infty} x_k P(X = x_k).$$

Valor esperado

2. Si X es una variable aleatoria continua con función de densidad f_X , se dice que X posee un valor esperado si

$$\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$$

En tal caso se define el valor esperado $E(X)$ (esperanza matemática, media) de X como:

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx$$

Sea Y una variable aleatoria absolutamente continua con función de densidad f . Si $E(Y)$ existe, entonces:

$$E(Y) = \int_0^{\infty} [1 - F_Y(y)] dy - \int_0^{\infty} F_Y(-y) dy$$

Demostración.

$$\begin{aligned} E(Y) &= \int_0^{\infty} x f(x) dx + \int_{-\infty}^0 x f(x) dx \\ &= \int_0^{\infty} \left(\int_0^x dy \right) f(x) dx - \int_{-\infty}^0 \left(\int_0^{-x} dy \right) f(x) dx \\ &= \int_0^{\infty} \int_y^{\infty} f(x) dx dy - \int_0^{\infty} \int_{-\infty}^{-y} f(x) dx dy \\ &= \int_0^{\infty} P(Y > y) dy - \int_0^{\infty} P(Y \leq -y) dy \\ &= \int_0^{\infty} [1 - F_Y(y)] dy - \int_0^{\infty} F_Y(-y) dy \end{aligned}$$

Ley del estadístico inconsciente

Sea X una variable aleatoria real con función de densidad f_X y $g : \mathbb{R} \rightarrow \mathbb{R}$ una función tal que $Y = g(X)$ es una variable aleatoria. Entonces:

$$E(g(X)) = \begin{cases} \sum_x g(x)f_X(x) & \text{si } X \text{ es una variable aleatoria discreta} \\ \int_{-\infty}^{\infty} g(x)f_X(x)dx & \text{si } X \text{ es una variable aleatoria continua} \end{cases}$$

siempre y cuando la suma, en el caso discreto, o la integral, en el caso continuo, converjan absolutamente.

Definición

Sean X una variable aleatoria, cuyo valor esperado existe. Se define la varianza de X como:

$$\text{Var}(X) = E(X - E(X))^2,$$

la cual mide la dispersión de los valores de la variable X alrededor de su media.

Sean X una variable aleatoria, cuyo valor esperado existe, y $\alpha, \beta \in \mathbb{R}$ constantes. Entonces:

1. $E(\alpha) = \alpha$
2. Si g, h son funciones, tales que $g(X)$ y $h(X)$ son variables aleatorias cuyos valores esperados existen, $E(\alpha g(X) + \beta h(X)) = \alpha E(g(X)) + \beta E(h(X))$
3. $\text{Var}(X) \geq 0$
4. $\text{Var}(\alpha) = 0$
5. $\text{Var}(\alpha X) = \alpha^2 \text{Var}(X)$
6. $\text{Var}(X + \beta) = \text{Var}(X)$

Ejercicio

(Desigualdad de Markov) Sea X una variable aleatoria real con $X \geq 0$ y tal que $E(X)$ existe. Demostrar que para todo $\alpha > 0$ se satisface:

$$P(X \geq \alpha) \leq \frac{E(X)}{\alpha}.$$

Ejercicio

(Desigualdad de Chebyshev) Sea X una variable aleatoria con media μ y varianza σ^2 . Demostrar que para todo $\epsilon > 0$ se satisface:

$$P(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

Ley débil de los grandes números

Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con media μ y varianza finita σ^2 . Entonces, para todo $\epsilon > 0$; se satisface que:

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2},$$

de donde se deduce que, para todo $\epsilon > 0$; se tiene que:

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| > \epsilon\right) = 0.$$

Demostración. Sea $\overline{X}_n$ la media aritmética de las n variables. $E(\overline{X}_n) = \mu$ y $\text{Var}(\overline{X}_n) = \frac{\sigma^2}{n}$. Por la desigualdad de Chebyshev se tiene que, para todo $\epsilon > 0$

$$P\left(\left|\overline{X}_n - E\overline{X}_n\right| \geq \epsilon\right) \leq \frac{\text{Var}(\overline{X}_n)}{\epsilon^2}$$

que es lo que se quería demostrar.

Convergencias

Convergencia en probabilidad

Sean $X, X_1, X_2, \dots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, se dice que $(X_n)_n$ converge en probabilidad a X

$$X_n \xrightarrow[n \rightarrow \infty]{P} X,$$

si, para todo $\epsilon > 0$; se satisface que:

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0.$$

Nota. Haciendo uso de esta definición, se puede expresar la ley débil de los grandes números como sigue: Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con media μ y varianza finita σ^2 . Entonces:

$$\overline{X_n} \xrightarrow[n \rightarrow \infty]{P} \mu,$$

Convergencia casi seguramente

Sean $X, X_1, X_2, \dots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \mathfrak{F}, P)$, se dice que $(X_n)_n$ converge casi seguramente a X

$$X_n \xrightarrow[n \rightarrow \infty]{c.s.} X,$$

si se satisface que $P(\lim_{n \rightarrow \infty} X_n = X) = 1$.

En otras palabras,

$$X_n \xrightarrow[n \rightarrow \infty]{c.s.} X, \text{ si y sólo si, } P(A) = 1,$$

donde $A := \{\omega \in \Omega : X_n(\omega) \longrightarrow X(\omega)\}$.

Ejemplo. Sean $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad arbitrario y X_n las variables aleatorias definidas sobre $(\Omega, \mathfrak{F}, P)$ como sigue:

$$\begin{aligned} X_n : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longmapsto \frac{1}{n} + 1 \end{aligned}$$

Es claro que, para todo $\omega \in \Omega$; se satisface que:

$$\lim_{n \rightarrow \infty} X_n(\omega) = 1.$$

por lo tanto,

$$P(\{\omega \in \Omega : X_n(\omega) \longrightarrow 1\}) = P(\Omega) = 1$$

esto es,

$$X_n \xrightarrow[n \rightarrow \infty]{c.s.} 1.$$

Ley fuerte de los grandes números

Definición

Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con media μ y varianza finita σ^2 . Entonces,

$$\overline{X_n} = \frac{X_1 + X_2 + \dots + X_n}{n} \xrightarrow[n \rightarrow \infty]{c.s.} \mu$$

Convergencias

Convergencia en distribución

Sean $X, X_1, X_2, \dots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \mathfrak{F}, P)$ con funciones de distribución $F, F_1, \dots$, respectivamente. Se dice que $(X_n)_n$ converge en distribución a X :

$$X_n \xrightarrow[n \rightarrow \infty]{d} X,$$

si se satisface que:

$$\lim_{n \rightarrow \infty} F_n(x) = F(x); \text{ para todo } x \text{ punto de continuidad de } F$$

o, equivalentemente,

$$X_n \xrightarrow[n \rightarrow \infty]{d} X, \text{ si y sólo si, } \lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x)$$

para todo x punto de continuidad de F .

Convergencias

Probar que

$$X_n \xrightarrow[n \rightarrow \infty]{c.s.} X \implies X_n \xrightarrow[n \rightarrow \infty]{P} X \implies X_n \xrightarrow[n \rightarrow \infty]{d} X$$

Teoremas límite

Teorema de continuidad de Levy & Cramer

Sean $X, X_1, X_2, \dots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \mathcal{F}, P)$ con funciones características $\phi, \phi_1, \phi_2, \dots$, respectivamente.

1. Si $X_n \xrightarrow[n \rightarrow \infty]{d} X$, entonces,

$$\lim_{n \rightarrow \infty} \phi_n(t) = \phi(t), \quad \text{para todo } t \in \mathbb{R}.$$

2. Si $\lim_{n \rightarrow \infty} \phi_n(t) = \phi(t)$ para todo $t \in \mathbb{R}$, y si $\phi(\cdot)$ es continua en $t = 0$, entonces,

$$X_n \xrightarrow[n \rightarrow \infty]{d} X.$$

Teoremas límite

Teorema del límite central

Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con media μ y varianza finita y positiva σ^2 . Sean $S_n := \sum_{j=1}^n X_j$ y $Y_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}$. Entonces, la sucesión de variables $Y_1, Y_2, \dots$ converge en distribución a una variable aleatoria Y , donde Y tiene distribución normal estándar.

Esto es,

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x); \text{ para todo } x \in \mathbb{R}.$$

Demostración. Se puede suponer, sin pérdida de generalidad, que $\mu = 0$. Sea ϕ la función característica de las variables aleatorias $X_1, X_2, \dots$. Como las variables aleatorias son independientes se obtiene que:

$$\begin{aligned}\phi_{Y_n}(t) &= E(\exp[itY_n]) \\&= E\left(\exp\left[it\frac{X_1 + \dots + X_n}{\sigma\sqrt{n}}\right]\right) \\&= E\left(\prod_{j=1}^n \exp\left[it\frac{X_j}{\sigma\sqrt{n}}\right]\right) \\&= \prod_{j=1}^n E\left(\exp\left[it\frac{X_j}{\sigma\sqrt{n}}\right]\right) \\&= \prod_{j=1}^n \phi\left(\frac{t}{\sigma\sqrt{n}}\right) \\&= \left[\phi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n\end{aligned}$$

Al hacer el desarrollo en serie de Taylor de ϕ alrededor de cero se obtiene:

$$\phi(t) = 1 + 0 - \frac{\sigma^2 t^2}{2} + t^2 o(t).$$

Por lo tanto,

$$\begin{aligned}\phi_{Y_n}(t) &= \left[1 - \frac{\sigma^2}{2} \left(\frac{t}{\sigma\sqrt{n}} \right)^2 + \left(\frac{t}{\sigma\sqrt{n}} \right)^2 o(t) \right]^n \\ &= \exp \left[n \ln \left(1 - \frac{\sigma^2}{2} \left(\frac{t}{\sigma\sqrt{n}} \right)^2 + \left(\frac{t}{\sigma\sqrt{n}} \right)^2 o(t) \right) \right].\end{aligned}$$

Tomando límite, cuando $n \rightarrow \infty$, se obtiene:

$$\lim_{n \rightarrow \infty} \phi_{Y_n}(t) = \exp \left[-\frac{t^2}{2} \right].$$

Por el teorema de continuidad de Levy & Cramer se concluye que

$$Y_n \xrightarrow{d} Y,$$

donde Y es una variable aleatoria con distribución normal estándar.

Introducción a procesos estocásticos

¿Qué es un proceso estocástico?

Definición (Proceso estocástico)

Un proceso estocástico es una colección de variables aleatorias $\{X_t \in S : t \in T\}$ sobre el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ parametrizada por un conjunto T llamado *espacio de índices* y un conjunto S en donde las variables toman valores, este conjunto es llamado *espacio de estados*.

¿Qué es un proceso estocástico?

¿Qué es un proceso estocástico?

Definición (Proceso estocástico)

Un proceso estocástico o aleatorio es una función de dos variables $X : T \times \Omega \rightarrow S$ tal que la pareja (t, ω) se le asocia el valor o el estado $X(t, \omega)$.

- Para cada valor fijo de t en T , el mapeo $\omega \mapsto X_t(\omega)$ es una variable aleatoria.
- Para cada valor de ω en Ω fijo, la función $t \mapsto X_t(\omega)$ es una trayectoria o realización del proceso.

¿Qué es un proceso estocástico?

Definición (Función aleatoria)

Sea L un subconjunto del espacio de funciones $S^T = \{f : T \rightarrow S\}$ y un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Una función aleatoria es una función de la forma $X : \Omega \rightarrow L$ que es medible para la σ -álgebra $\mathcal{F}$ definida sobre Ω y para alguna σ -álgebra definida sobre L .

- ★ ¿La definición de *proceso estocástico* es equivalente a la definición de *función aleatoria*? ¿Una es caso especial de la otra?

Algunos tipos de procesos estocásticos

- **Proceso de ensayos independientes.** El proceso a tiempo discreto $\{X_n : n = 0, 1, \dots\}$ puede estar constituido por variables aleatorias independientes.
- **Procesos de Markov.** Estos tipos de procesos son modelos en donde, suponiendo conocido el estado presente del sistema, los estados anteriores no tienen influencia en los estados futuros del sistema. En formulación matemática,

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

Algunos tipos de procesos estocásticos

- **Procesos con incrementos independientes.** Se dice que un proceso estocástico a tiempo continuo $\{X_t \mid t \geq 0\}$ tiene incrementos independientes si para cualesquiera tiempos $0 \leq t_1 < t_2 < \dots < t_n$ las variables $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes.
- **Procesos estacionarios.** Se dice que un proceso estocástico a tiempo continuo $\{X_t \mid t \geq 0\}$ es estacionario si para cualesquiera tiempos $t_1, \dots, t_n$, la distribución del vector $(X_{t_1}, \dots, X_{t_n})$ es la misma que la del vector $(X_{t_1+h}, \dots, X_{t_n+h})$ para cualquier $h > 0$.

Algunos tipos de procesos estocásticos

- **Procesos con incrementos estacionarios.** Un proceso estocástico a tiempo continuo $\{X_t \mid t \geq 0\}$ tiene incrementos estacionarios si para cualesquiera tiempos $s < t$ y para cualquier $h > 0$, las variables $X_{t+h} - X_{s+h}$ y $X_t - X_s$ tienen la misma distribución de probabilidad.
- **Martingalas.** Un martingala a tiempo discreto es un proceso $\{X_n \mid n = 0, 1, \dots\}$ que cumple la condición

$$E(X_{n+1} \mid X_n = x_n, \dots, X_0 = x_0) = x_n.$$

Algunos tipos de procesos estocásticos

- **Procesos de Lèvy.** Un proceso estocástico a tiempo continuo $\{X_t \mid t \geq 0\}$ es un proceso de Lèvy si sus incrementos son independientes y estacionarios. El proceso de Poisson y el movimiento Browniano son ejemplos de este tipo de proceso.
- **Procesos Gaussianos.** Un proceso estocástico a tiempo continuo $\{X_t \mid t \geq 0\}$ es un proceso Gaussiano si para cualesquiera colección finita de tiempos $t_1, \dots, t_n$ el vector $(X_{t_1}, \dots, X_{t_n})$ tiene distribución Gaussiana multivariada.

Ejercicio

Sean $X_0, X_1, \dots$ los resultados de una sucesión de ensayos independientes Bernoulli. Determine si el proceso $\{X_n : n = 0, 1, \dots\}$

1. tiene incrementos independientes.
2. tiene incrementos estacionarios.
3. es un martingala.
4. cumple la propiedad de Markov.

Ejercicio

Sea X una variable aleatoria con distribución Bernoulli(p). Para cada $t \geq 0$ defina la variable

$$X_t = \begin{cases} \cos \pi t & \text{si } X = 0, \\ \sin \pi t & \text{si } X = 1. \end{cases}$$

1. Dibuje todas las trayectorias del proceso $\{X_t \mid t \geq 0\}$.
2. Calcule la distribución de X_t .
3. Calcule $E(X_t)$.

Ejercicio

Demuestre que todo proceso a tiempo discreto $\{X_n \mid n = 0, 1, \dots\}$ con incrementos independientes cumple la propiedad de Markov.

Caminatas aleatorias

¿Qué es una caminata aleatoria?

Definición

Consideremos la variable aleatoria ξ con $P(\xi = 1) = p$ y $P(\xi = -1) = q$ donde $p + q = 1$. Una **caminata aleatoria** es un proceso estocástico $\{X_n : n = 0, 1, 2, \dots\}$ donde $X_0 = 0$ y para $n > 0$ se define como $X_n = X_0 + \xi_1 + \dots + \xi_n$ con ξ_j distribuida de la misma forma que ξ y además para $i \neq j$, $\xi_i \perp \xi_j$.

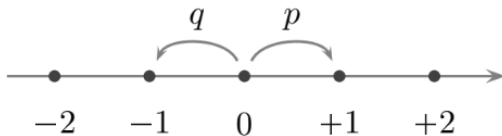


Figura: Caminata aleatoria en $\mathbb{Z}$. Figura tomada de Rincon, 2012

Caminata aleatoria

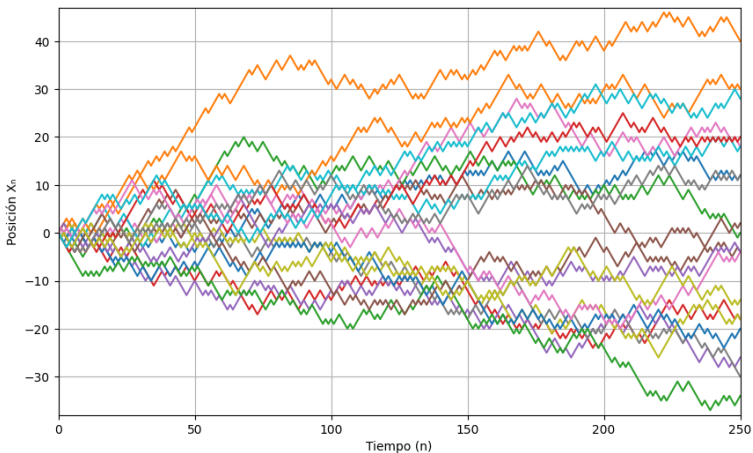


Figura: Caminatas aleatorias simétricas en $\mathbb{Z}$.

Tablero de Galton

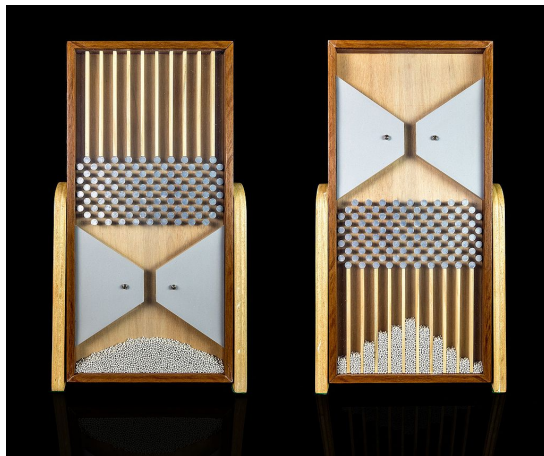


Figura: Orden a partir del caos: tablero de Galton

Momentos de una caminata aleatoria

Proposición

Para cualquier entero $n \geq 0$,

- $E(X_n) = n(p - q)$
- $\text{Var}(X_n) = 4npq$

Demostración. En efecto,

- Para la esperanza:

$$E(X_n) = \sum_{i=1}^n E(\xi_i) = \sum_{i=1}^n p(1) + q(-1) = n(p - q).$$

- Por otra parte, se tiene que $E(\xi_i^2) = p(1)^2 + q(-1)^2 = p + q = 1$, por lo tanto,

$$\text{Var}(X_n) = \sum_{i=1}^n \text{Var}(\xi_i) = \sum_{i=1}^n E(\xi_i^2) - E(\xi_i)^2 = n(1 - (p - q)^2) = 4npq.$$

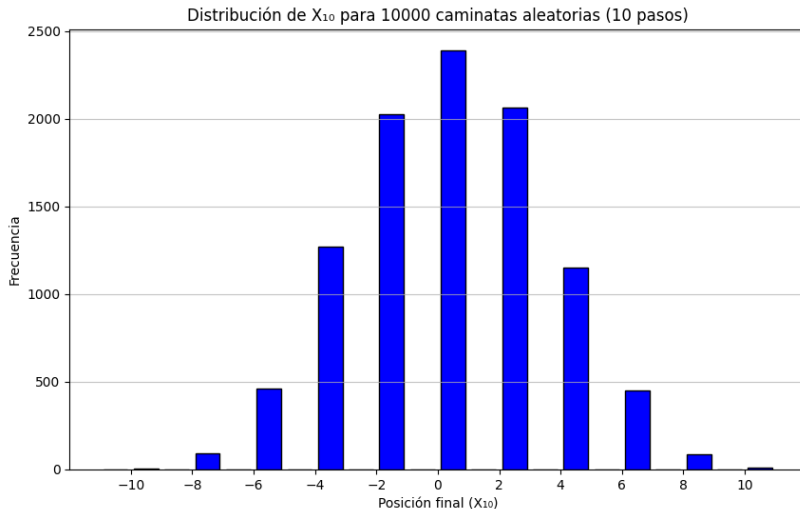


Figura: Generamos 10000 caminatas aleatorias simétricas partiendo desde el origen y con 10 pasos, luego de esto aproximamos la distribución de X_{10} por medio de un histograma.

Distribución de X_n

Proposición

Para cualesquiera números enteros x y n tales que $-n \leq x \leq n$ y para el caso cuando x y n son ambos pares o ambos impares se tiene que,

$$P(X_n = x \mid X_0 = 0) = \binom{n}{\frac{1}{2}(n+x)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)}$$

Para valores de x y n que no cumplen las condiciones indicadas la probabilidad en cuestión vale cero.

★ ¿Qué ocurre cuando $X_0 \neq 0$?

Distribución de X_n

Demostración. Se tiene que $X_n = R_n - L_n$ donde R_n es la variable aleatoria que representa el número de pasos hacia la derecha en el tiempo n y L_n es la variable aleatoria que representa el número de pasos hacia la izquierda en el tiempo n . También es cierto que $R_n + L_n = n$, luego $R_n = \frac{1}{2}(n + X_n)$, y,

$$R_n = \frac{1}{2}(n + X_n) = \sum_{i=1}^n \frac{1}{2}(1 + \xi_i)$$

Por la definición de ξ se tiene que $\frac{1}{2}(1 + \xi) \sim \text{Bernoulli}(p)$, y por la independencia de las variables aleatorias $\{\xi_i\}_i$ se tiene que $R_n \sim \text{Binomial}(n, p)$. Por tanto,

$$P(X_n = x \mid X_0 = 0) = P\left(R_n = \frac{1}{2}(n + x)\right) = \binom{n}{\frac{1}{2}(n + x)} p^{\frac{1}{2}(n+x)} q^{\frac{1}{2}(n-x)}.$$

Probabilidad de regreso a la posición de origen

Proposición

Para una caminata aleatoria sobre $\mathbb{Z}$, la probabilidad de un eventual regreso al punto de partida es

$$1 - |p - q| \begin{cases} 1, & \text{si } p = q, \\ < 1, & \text{si } p \neq q \end{cases}$$

Probabilidad de regreso a la posición de origen

Proposición

Para una caminata aleatoria sobre $\mathbb{Z}$, la probabilidad de un eventual regreso al punto de partida es

$$1 - |p - q| \begin{cases} 1, & \text{si } p = q, \\ < 1, & \text{si } p \neq q \end{cases}$$

Demostración. Considere las siguientes definiciones:

$$p_n := P(X_n = 0 \mid X_0 = 0)$$

$$f_k := P(X_k = 0, X_{k-1} \neq 0, \dots, X_1 \neq 0 \mid X_0 = 0)$$

(...)

Revisar Rincon, 2012.

El problema del jugador

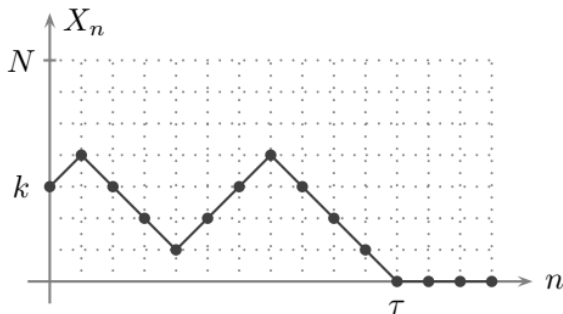
Planteamiento del problema

Suponga que un jugador A apuesta sucesivamente una unidad monetaria a un jugador B . Inicialmente el jugador A tiene k unidades y B tiene $N - k$ unidades.

En cada ronda de apuesta, el jugador A tiene probabilidad de p de ganar y probabilidad $q = 1 - p$ de perder, además no hay empates. El juego termina cuando un jugador no tiene dinero.

Sea X_n la cantidad de unidades monetarias que tiene el jugador A en el tiempo n , con $X_0 = k$. El conjunto $\{X_n : n = 0, 1, \dots\}$ es una caminata aleatoria sobre el conjunto de estados $\{0, 1, \dots, N\}$.

¿Cuál es la probabilidad de que el jugador A se arruine?



El problema del jugador

Solución del problema

Considere el tiempo τ en donde por primera vez se llega a un estado donde termina el juego:

$$\tau = \min\{n \geq 0 : X_n = 0 \text{ ó } X_n = N\}.$$

La pregunta se reduce a encontrar la probabilidad

$$P(X_\tau = 0 \mid X_0 = k) =: u_k.$$

Notemos que para $k = 1, 2, \dots, N - 1$,

$$\begin{aligned} u_k &= P(X_\tau = 0 \mid X_0 = k) \\ &= p P(X_\tau = 0 \mid X_0 = k + 1) + q P(X_\tau = 0 \mid X_0 = k - 1) \\ &= p u_{k+1} + q u_{k-1}, \end{aligned}$$

para las fronteras $u_0 = 1$ y $u_N = 0$.

El problema del jugador

De la anterior ecuación se obtiene que

$$u_{k+1} - u_k = \frac{q}{p}(u_k - u_{k-1}),$$

resolviendo iterativamente y usando que $u_0 = 1$:

$$u_2 - u_1 = \left(\frac{q}{p}\right)(u_1 - 1)$$

$$u_3 - u_2 = \left(\frac{q}{p}\right)^2(u_1 - 1)$$

$$\vdots$$

$$u_k - u_{k-1} = \left(\frac{q}{p}\right)^{k-1}(u_1 - 1)$$

El problema del jugador

Definiendo $S_k = \sum_{l=1}^k \left(\frac{q}{p}\right)^l = \begin{cases} k+1 & \text{si } p = 1/2, \\ \frac{1 - \left(\frac{q}{p}\right)^{k+1}}{1 - \left(\frac{q}{p}\right)} & \text{si } p \neq 1/2. \end{cases}$ y sumando las anteriores

ecuaciones llegamos a que

$$u_k = 1 + S_{k-1}(u_1 - 1).$$

En particular $0 = u_N = 1 + S_{N-1}(u_1 - 1)$ así, $u_1 - 1 = \frac{-1}{S_{N-1}}$. Sustituyendo y simplificando se llega a la solución

$$u_k = 1 - \frac{S_{k-1}}{S_{N-1}}.$$

El problema del jugador

Por lo tanto,

$$u_k = \begin{cases} \frac{N-k}{N} & \text{si } p = 1/2, \\ \frac{\left(\frac{q}{p}\right)^k - \left(\frac{q}{p}\right)^N}{1 - \left(\frac{q}{p}\right)^N} & \text{si } p \neq 1/2. \end{cases}$$

El problema del jugador

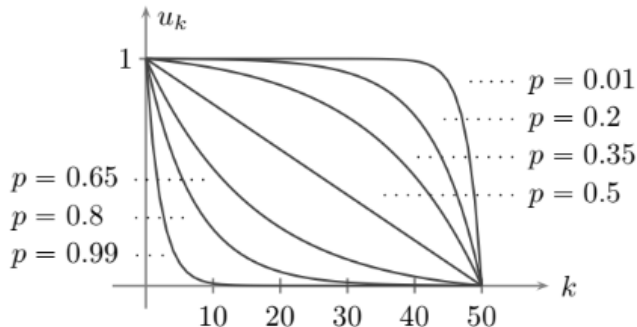


Figura: Probabilidad u_k como función del parámetro k para diferentes valores de p y $N = 50$. Imagen de Rincon, 2012.

El problema del jugador

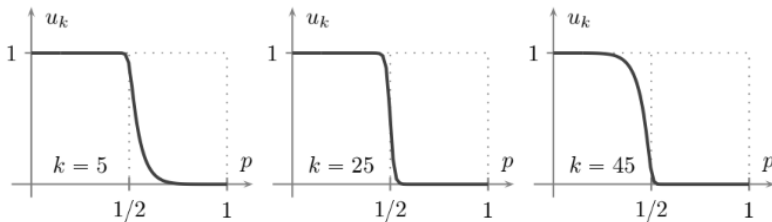


Figura: Probabilidad u_k como función de p para diferentes valores de dinero inicial k y $N = 50$. Imagen de Rincon, 2012.

El problema del jugador

Número de apuestas antes de la ruina

El número esperado de apuestas antes de la ruina es

$$m_k = \begin{cases} k(N - k) & \text{si } p = q, \\ \frac{1}{q-p} \left(k - N \frac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N} \right) & \text{si } p \neq q \end{cases}$$

Demostración. Condicionando sobre la primera apuesta se obtiene que m_k satisface

$$m_k = 1 + p m_{k+1} + q m_{k-1} \quad k = 1, 2, \dots, N - 1$$

$$m_0 = 0$$

$$m_N = 0$$

$$(\dots)$$

Revisar Rincon, 2012.

El problema del jugador

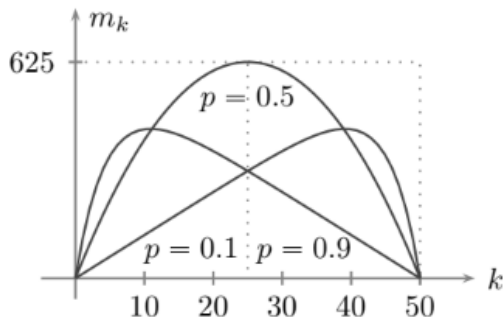


Figura: Forma en que cambia m_k a medida que cambia k cuando $N = 50$. Imagen tomada de Rincon, 2012.

Hombre borracho vs Ave borracha

“Un hombre borracho encontrará eventualmente su camino a casa, pero un ave borracha puede perderse por siempre” – Kakutani

- En una o dos dimensiones la caminata aleatoria vuelve infinitas veces al punto de partida con probabilidad 1. En tres dimensiones o más, hay una probabilidad positiva de no volver nunca al punto de inicio.

Ejercicio

Demuestre que una caminata aleatoria simple $\{X_n : n \geq 0\}$ sobre $\mathbb{Z}$ cumple la propiedad de Markov, es decir, muestre que para cualquier valor natural de n y para cualesquiera enteros $x_0, x_1, \dots, x_{n+1}$, se cumple que

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

Solución: $X_{n+1} = X_n + \xi_{n+1}$. Por lo que $\xi_{n+1} = X_{n+1} - X_n$, siendo $\xi_1, \xi_2, \dots$ variables aleatorias independientes e idénticamente distribuidas. Así, la caminata aleatoria simple es un proceso con incrementos independientes, y como se demostró anteriormente, todo proceso con incrementos independientes cumple la propiedad de Markow.

Ejercicio

Para una caminata aleatoria simple $\{X_n : n \geq 0\}$ sobre $\mathbb{Z}$ demuestre que

$$P(X_{n+1} = x) = p P(X_n = x - 1) + q P(X_n = x + 1).$$

Solución:

Por el teorema de probabilidad total,

$$\begin{aligned} P(X_{n+1} = x) &= P(X_{n+1} = x | X_n = x - 1)P(X_n = x - 1) \\ &\quad + P(X_{n+1} = x | X_n = x + 1)P(X_n = x + 1) \\ &= p P(X_n = x - 1) + q P(X_n = x + 1). \end{aligned}$$

Ejercicio

Una partícula realiza una caminata aleatoria simétrica sobre $\mathbb{Z}$ empezando en cero.

- ¿Cuál es la probabilidad de que la partícula se encuentre nuevamente en el origen en el sexto paso?

$$\frac{\binom{6}{3}}{2^6} = \frac{5}{16} = 0,3125$$

- ¿Cuál es la probabilidad de que la partícula regrese al estado cero por primera vez en el sexto paso?

De las 64 opciones, sólo hay 4 posibles que cumplen la condición; LLRLRR, LLLRRR, RRLLLL, RRLRLL.

$$\frac{4}{2^6} = \frac{1}{16} = 0,0625$$

Ejercicio

Recurrencia. Sea $\{X_n : n \geq 0\}$ una caminata aleatoria simple sobre $\mathbb{Z}$ que inicia en $X_0 = 0$. Sea $f_{i,j}$ la probabilidad de que eventualmente la caminata visite el estado j a partir del estado i , es decir,

$$f_{i,j} = P(X_n = j \text{ para algún } n \geq 1 \mid X_0 = i).$$

- Demuestre que $f_{-1,1} = f_{-1,0}f_{0,1} = f_{0,1}^2$.
- Condicionando sobre $X_0 = i$ demuestre que $f_{0,0} = p f_{1,0} + q f_{-1,0} = p f_{1,0} + q f_{0,1}$.
- Demuestre que $f_{0,1} = p + q f_{0,1}^2$.
- Resuelva la ecuación cuadrática del ítem anterior y obtenga las dos raíces $f_{0,1} = 1, \frac{p}{q}$, y de manera análoga $f_{1,0} = 1, \frac{q}{p}$.
- Analice los casos $p > q$ y $p < q$ por separado para encontrar que $f_{0,0} = 2q$ y $f_{0,0} = 2p$, respectivamente. Concluya que $f_{0,0} = 2 \min\{p, q\} = 1 - |p - q|$.

Ejercicio

Demuestre que si $p = \frac{1}{2}$ entonces $E(X_{n+1} | X_n) = X_n$. ¿La afirmación contraria es verdadera?

Solución:

$$\rightarrow E(X_{n+1} | X_n) = \sum_x xP(X_{n+1} = x|X_n) = p(X_n + 1) + q(X_n - 1) = X_n.$$

$$\leftarrow E(X_{n+1}|X_n) = X_n + p - q. \text{ Esto es igual a } X_n, \text{ sólo si } p = q = \frac{1}{2}.$$

Cadenas de Markov

Andrey Andreyevich Markov

Cadena de Markov

Definición

Una cadena de Markov es un proceso estocástico a tiempo discreto $\{X_n : n = 0, 1, \dots\}$ con espacio de estados discreto y que satisface la propiedad de Markov, esto es, para cualquier entero $n \geq 0$, y para cualesquiera estados $x_0, \dots, x_{n+1}$, se cumple

$$P(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = P(X_{n+1} = x_{n+1} \mid X_n = x_n).$$

Una condición equivalente a la anterior propiedad es poder calcular la distribución conjunta de las variables aleatorias $X_0, X_1, \dots, X_n$ de la siguiente manera:

$$\begin{aligned} P(X_0 = x_0, \dots, X_n = x_n) \\ = P(X_0 = x_0)p(X_1 = x_1 \mid X_0 = x_0) \cdots P(X_n = x_n \mid X_{n-1} = x_{n-1}). \end{aligned}$$

Usualmente vamos a considerar que el conjunto de estados de la cadena de Markov es el conjunto $\{0, 1, 2, \dots\}$.

Matriz de transición

$p_{ij}^{(n,n+1)} = P(X_{n+1} = j \mid X_n = i) \Rightarrow$ Depende del tiempo: No homogénea

$p_{ij} = P(X_{n+1} = j \mid X_n = i) \Rightarrow$ No depende del tiempo: Homogénea

$$P = \begin{pmatrix} p_{00} & p_{01} & \cdots \\ p_{10} & p_{11} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Proposición

La matriz de probabilidades de transición $P = (p_{ij})$ cumple las siguientes dos propiedades:

1. $p_{ij} \geq 0$.
2. $\sum_j p_{ij} = 1$.

Otros elementos

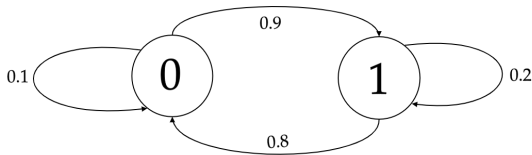
- **Distribución inicial de estados.** Una cadena de Markov inicia partiendo de un estado cualquiera o más generalmente considerando una distribución de probabilidad inicial sobre el espacio de estados.
- **Existencia.** Dada una matriz estocástica y una distribución de probabilidad inicial existe un espacio de probabilidad y una cadena de Markov con matriz de probabilidades de transición y distribución inicial especificadas.
- **Probabilidades de transición en n pasos.** Dada una cadena de Markov homogénea en el tiempo, se define:

$$p_{ij}(n) = \begin{cases} P(X_{n+m} = j \mid X_m = i) & n > 0 \\ \delta_{ij} & n = 0 \end{cases}$$

$$P(n) = \begin{pmatrix} p_{00}(n) & p_{01}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}$$

Otros elementos

- **Cadena de Markov como grafo.** Si el conjunto de estados de una cadena de Markov es finito y además la cadena es homogénea respecto al tiempo entonces es posible construir una representación como un grafo dirigido con pesos en donde cada nodo del grafo corresponde a un posible estado ℓ de la cadena de Markov, hay una arista (i, j) si del estado i hay una probabilidad $p_{ij} > 0$ de hacer una transición al estado j , tal probabilidad corresponde al peso de cada arista.



Kernel de transición

Definición

Sea $(\Omega, \mathcal{F})$ un espacio medible. Una función $P : \Omega \times \mathcal{F} \rightarrow [0, 1]$ es un kernel de transición si:

- Para cada $B \in \mathcal{F}$, la función $x \mapsto P(x, B)$ es una función medible sobre Ω .
- Para cada $x \in \Omega$, la función $B \mapsto P(x, B)$ es una función de distribución.

★ ¿Esta definición es una generalización de la matriz de transición?

Ejemplo: Kernel de transición

Supongamos que tenemos una cadena de Markov $\{X_n \mid n = 0, 1, \dots\}$ con tiempo discreto y espacio de estados continuo $\mathbb{R}$. Si tenemos que $X_{n+1}|X_n \sim \text{Normal}(\mu = X_n, \sigma^2 = 1)$ entonces el kernel de transición de esta cadena está dado por:

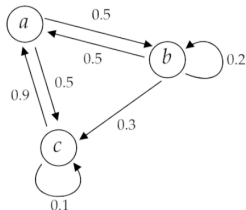
$$P(x, B) = \int_B \frac{1}{(2\pi)^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(x' - x)^2\right) dx'.$$

Por otra parte, si $X_{n+1}|X_n \sim \text{Normal}(\mu = \frac{1}{2}X_n, \sigma^2 = e^{-2n})$ entonces el kernel de transición de esta cadena está dado por:

$$P_n(x, B) = \int_B \frac{1}{(2\pi e^{-2n})^{\frac{1}{2}}} \exp\left(-\frac{1}{2e^{-2n}} \left(x' - \frac{1}{2}x\right)^2\right) dx'.$$

★ ¿Cómo es el comportamiento de las cadenas de Markov cuando el tiempo va para infinito?

Ejercicio



Considere la cadena de Markov en tiempo homogéneo $\{X_n \in \{a, b, c\} : n = 0, 1, \dots\}$ descrita por la figura anterior. Sea $\Omega = \{a, b, c\}$ el conjunto de estados y $\mathcal{S} = \{X \mid X \subseteq \Omega\}$ la σ álgebra definida sobre Ω . Para cada $x \in \Omega$ y para cada $B \in \mathcal{S}$ considere la función $P : \Omega \times \mathcal{S} \rightarrow [0, 1]$ definida por $P(x, B) = P(X_{n+1} \in B \mid X_n = x)$ para cualquier valor entero de n . Demuestre que:

- Si $B = \{a, b\}$ entonces la función $x \mapsto P(x, B)$ es una función medible.
- Si $x = a$ entonces la función $B \mapsto P(x, B)$ es una medida de probabilidad.

Ejemplo: Cadena de juguete

Considere la cadena de Markov definida por la matriz de transición

$$P = \begin{pmatrix} 0,8 & 0,2 \\ 0,3 & 0,7 \end{pmatrix}.$$

- ★ ¿Cómo se puede representar esta cadena de Markov con un grafo?
- ★ ¿Existe $\underline{\pi} \in \mathbb{R}^2$ de tal manera que $\underline{\pi}^\top \underline{1} = 1$ y que $\underline{\pi} = P^\top \underline{\pi}$?

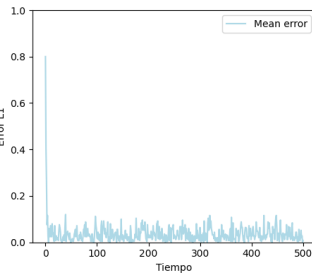
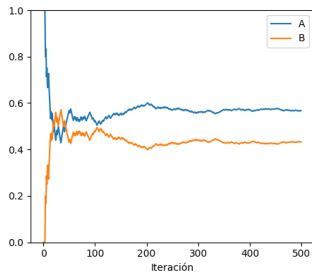
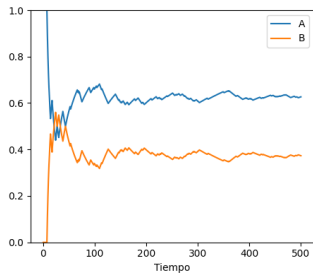
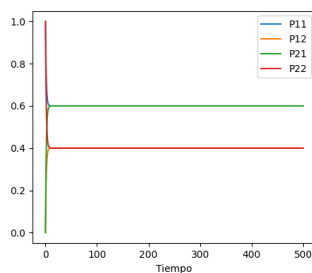
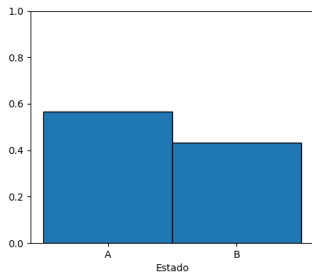
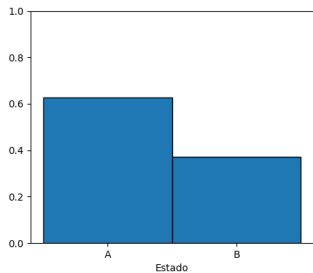


Figura: Diferentes visualizaciones gráficas para la cadena de Markov.

El gráfico resume el comportamiento de una cadena de Markov de dos estados con matriz de transición $P = \begin{pmatrix} 0,8 & 0,2 \\ 0,3 & 0,7 \end{pmatrix}$.

- En la primera fila:
 - El histograma de una trayectoria larga ilustra las frecuencias empíricas de visitas a los estados.
 - La distribución de X_{500} basada en muchas simulaciones muestra la estabilidad de la cadena al final del horizonte.
 - La evolución de las potencias P^n evidencia la convergencia hacia la **distribución estacionaria**.
- En la segunda fila:
 - La evolución temporal de las frecuencias empíricas en una sola trayectoria se estabiliza alrededor de la distribución límite.
 - La proporción acumulada de estados en simulaciones independientes confirma la convergencia media.
 - El error L^1 entre la distribución empírica y la estacionaria disminuye con el tiempo, reflejando la **mezcla de la cadena**.

Ejemplo: Cadena de dos estados

Considere una cadena de Markov (homogénea) con espacio de estados $\{0, 1\}$.

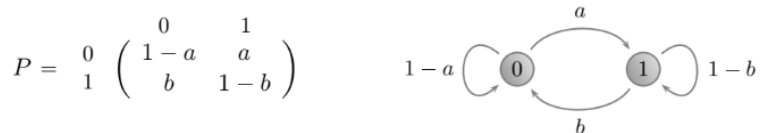


Figura: Representación de cadena de dos estados. Tomado de Rincon, 2012.

- $0 \leq a \leq 1$ y $0 \leq b \leq 1$.
- ★ Si $a = 1 - b$ las variables $X_1, X_2, \dots$ son independientes e idénticamente distribuidas con $P(X_n = 0) = 1 - a$ y $P(X_n = 1) = a$ para $n \geq 1$.
- ★ ¿Qué ocurre cuando $a \neq 1 - b$?

Ejemplo: Cadena de variables aleatorias independientes

Considere una sucesión de variables aleatorias independientes $\{\xi_i\}_{i \geq 1}$ con valores en el conjunto $\{0, 1, \dots\}$ y con idéntica distribución dada por las probabilidades $\{a_i\}_{i \geq 0}$.

Considere $X_n := \xi_n$. El proceso estocástico $\{X_n \mid n \geq 1\}$ es una cadena de Markov con espacios de estados $\{0, 1, \dots\}$ y con probabilidades de transición

$$p_{ij} = P(X_n = j \mid X_{n-1} = i) = P(X_n = j) = a_j.$$

La matriz de transición de esta cadena de Markov es:

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & \dots \\ a_0 & a_1 & a_2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

- ★ Las probabilidades de transición en n pasos están dados por $p_{ij}(n) = a_j$ para cualesquiera $i, j \in \{0, 1, \dots\}$.

Ejemplo: Cadena de variables aleatorias independientes

Considere una sucesión de variables aleatorias independientes $\{\xi_i\}_{i \geq 1}$ con valores en el conjunto $\{0, 1, \dots\}$ y con idéntica distribución dada por las probabilidades $\{a_i\}_{i \geq 0}$. En este caso $X_n := \max_{1 \leq j \leq n} \xi_j$. El proceso estocástico $\{X_n \mid n \geq 1\}$ es una cadena de Markov con espacios de estados $\{0, 1, \dots\}$ y la matriz de transición de esta cadena de Markov es:

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \dots \\ 0 & a_0 + a_1 & a_2 & a_3 & \dots \\ 0 & 0 & a_0 + a_1 + a_2 & a_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Ejemplo: Cadena de variables aleatorias independientes

Considere una sucesión de variables aleatorias independientes $\{\xi_i\}_{i \geq 1}$ con valores en el conjunto $\{0, 1, \dots\}$ y con idéntica distribución dada por las probabilidades $\{a_i\}_{i \geq 0}$. En este caso $X_n := \sum_{1 \leq j \leq n} \xi_j$. El proceso estocástico $\{X_n \mid n \geq 1\}$ es una cadena de Markov con espacios de estados $\{0, 1, \dots\}$ y la matriz de transición de esta cadena de Markov es:

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \dots \\ 0 & a_0 & a_1 & a_2 & \dots \\ 0 & 0 & a_0 & a_1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Ejemplo: Cadena de rachas de éxitos

Sea $\xi_1, \xi_2, \dots$ una sucesión de ensayos independientes Bernoulli con probabilidad de éxito p y probabilidad de fracaso $q = 1 - p$. Sea X_n el número de éxitos consecutivos previos al tiempo n incluyendo el tiempo n . Se dice que una racha de éxitos de longitud r ocurre en el tiempo n si en el ensayo $n - r$ se obtiene un fracaso y los resultados de los ensayos $n - r + 1$ al n son todos éxitos. La colección de variables aleatorias $\{X_n \mid n = 1, 2, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$.

★ Las probabilidades de transición están dadas por

$$p_{ij} = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = 0, \\ 0 & \text{en otro caso} \end{cases}$$

para $i, j \in \{0, 1, \dots\}$.

★ La matriz de transición es $P = \dots$

★ El grafo (infinito) de la cadena de Markov es ...

Ejemplo: Cadena de la caminata aleatoria

Una caminata aleatoria sobre el conjunto de número enteros constituye una cadena de Markov con espacio de estados $\mathbb{Z}$ y con probabilidades de transición

$$p_{ij} = \begin{cases} p & \text{si } j = i + 1, \\ q & \text{si } j = i - 1, \\ 0 & \text{en otro caso} \end{cases}$$

en donde $p + q = 1$.

- ★ La matriz de transición es $P = \dots$
- ★ El grafo de la cadena de Markov es $\dots$

Ejemplo: Cadena del jugador

El modelo usado para el problema del jugador estudiado anteriormente corresponde a una caminata aleatoria sobre el conjunto de números enteros $\{0, 1, \dots, N\}$ en donde los estados 0 y N son absorbentes- Las proabilidades de transición son:

$$p_{ij} = \begin{cases} p & \text{si } j = i + 1 \text{ y } i \in \{1, \dots, N - 1\}, \\ q & \text{si } j = i - 1 \text{ y } i \in \{1, \dots, N - 1\}, \\ 1 & \text{si } i = j \text{ y } i \in \{0, N\} \\ 0 & \text{en otro caso.} \end{cases}$$

- ★ La matriz de transición para $N = 5$ es $P = \dots$
- ★ El grafo de la cadena de Markov para $N = 5$ es ...

Ejemplo: Cadena de Ehrenfest

Sean A y B dos urnas dentro de las cuales se encuentran distribuidas un total de N bolas de acuerdo a una configuración inicial, por ejemplo, en la urna A hay i bolas y en la urna B hay $N - i$ bolas. En cada unidad de tiempo se selecciona una bola al azar y se cambia de urna.

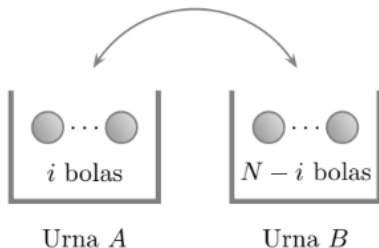


Figura: Figura tomada de Rincon, 2012.

Ejemplo: Cadena de Ehrenfest

Sea X_n el número de bolas en la urna A luego de n extracciones o unidades de tiempo. Entonces la colección $\{X_n \in \{0, \dots, N\} \mid n = 0, 1, \dots\}$ constituye una cadena de Markov con espacio de estados finito.

A partir del estado $i \in \{1, 2, \dots, N-1\}$ sólo se puede pasar al estado $i-1$ o $i+1$ con probabilidades $p_{i,i-1} = \frac{i}{N}$ y $p_{i,i+1} = \frac{(N-i)}{N}$; por otra parte $p_{0,1} = 1$ y $p_{N,N-1} = 1$, en este caso decimos que los estados 0 y N son *reflejantes*.

- ★ ¿Cuál es la esperanza de la variable aleatoria X_n dada una distribución inicial de bolas en cada urna?
- ★ ¿Cuál es la matriz de transición de esta cadena?

Ejemplo: Cadena de ramificación

Sea ξ_k la variable aleatoria que modela el número de descendientes de la k -ésima partícula. Para cada $n \geq 0$ defina X_n como el número de partículas en la generación n . Entonces $\{X_n : n = 0, 1, \dots\}$ es una cadena de Markov con espacio de estados $\{0, 1, \dots\}$.

★ Las probabilidades de transición $p_{ij} = P(\xi_1 + \dots + \xi_i = j)$.

Si en algún momento $X_n = 0$ entonces se dice que la población se ha extinguido y el estado 0 es absorbente.

★ Suponga que $\xi_k \in \{0, 1, 2\}$ con probabilidades $\{\nu_0, \nu_1, \nu_2\}$. Dado que $X_0 = 1$, ¿cuál es la probabilidad de que $X_2 = 0$?

Ejemplo: Cadena de la fila de espera

Considere una línea de espera de clientes que solicitan algún tipo de servicio a un servidor. Cuando hay algún cliente esperando el servicio al inicio del un periodo, el cliente al frente de la fila es atendido terminando el servicio al final del periodo. Si no existiera ningún cliente en la fila al inicio de algún periodo, entonces ningún cliente es atendido.

Para cada entero $n \geq 1$ defina ξ_n como el número de nuevos clientes que se incorporan a la fila durante el periodo n , supongamos que las variables aleatorias $\{\xi_\ell\}$ son independientes e idénticamente distribuidas con valores enteros no negativos y $P(\xi_n = k) =: a_k$ con $a_k \geq 0$ y $\sum_k a_k = 1$.

Sea X_0 el número de clientes iniciales y defina X_{n+1} para $n \geq 0$ como el número de clientes en la fila al final del periodo n . Entonces

$$X_{n+1} = \begin{cases} \xi_{n+1} & \text{si } X_n = 0, \\ X_n + \xi_{n+1} - 1 & \text{si } X_n \geq 1. \end{cases}$$

★ ¿Cómo calcular p_{ij} ?

Ejemplo: Cadena de la fila de espera

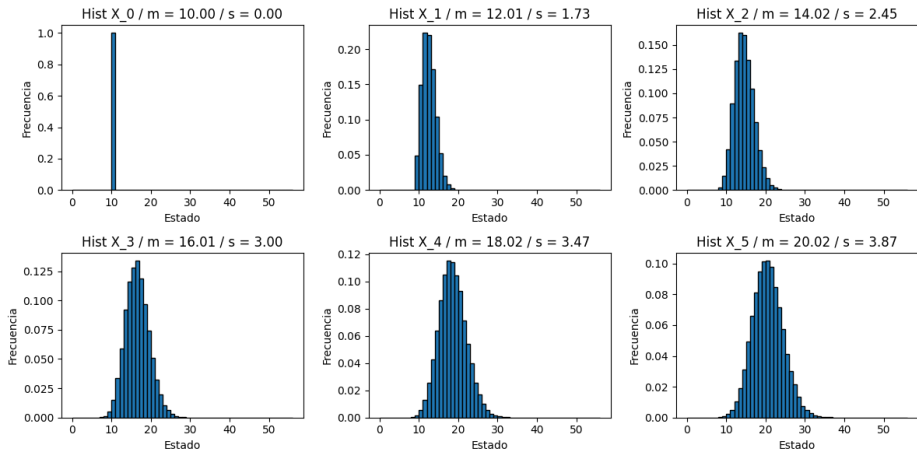


Figura: Simulación de una cadena de fila de espera para 6 tiempos.

El gráfico muestra los histogramas normalizados de los estados X_0 hasta X_5 en una cadena de fila de espera donde:

$$X_{n+1} = \begin{cases} \xi_{n+1}, & X_n = 0, \\ X_n + \xi_{n+1} - 1, & X_n \geq 1, \end{cases} \quad \xi_n \sim \text{Poisson}(\lambda), \lambda = 3.$$

- Se simularon 50,000 trayectorias independientes partiendo de $X_0 = 10$.
- Cada panel del arreglo 2×3 muestra el histograma empírico de X_n , junto con su media y desviación estándar.
- La distribución de X_n se desplaza hacia valores más pequeños en los primeros pasos, debido al efecto del servicio (se atiende un cliente por periodo).
- A medida que avanza el tiempo, la influencia del estado inicial disminuye y las distribuciones comienzan a concentrarse alrededor de un equilibrio determinado por el balance entre servicio y llegadas Poisson.

Ejemplo: Cadena de inventarios

Suponga que se almacena un cierto número de bienes en una bodega, y que se requiere una demanda aleatoria ξ_n del bien en el periodo n . Sea $P(\xi_n = k) = a_k$, para $k = 0, 1, \dots$ con $a_k \geq 0$ y $\sum_k a_k = 1$. El número de bienes en el almacén es revisado al final de cada periodo y se aplica la siguiente política de reabastecimiento: si al final del periodo la cantidad del bien es menor o igual a un cierto nivel s , entonces se reabastece la bodega inmediatamente hasta un nivel máximo S . Si al final del periodo el número de bienes es mayor a s , entonces no hay ningún reabastecimiento. Sea X_n el número de bienes al final del periodo n y justo antes de aplicar la política de reabastecimiento. Entonces $\{X_n : n \geq 0\}$ es una cadena de Markov con espacio de estados $\{\dots, -2, -1, 0, 1, 2, \dots, S\}$. Entonces

$$X_{n+1} = \begin{cases} X_n - \xi_{n+1}, & \text{si } s < X_n < S, \\ S - \xi_{n+1}, & \text{si } X_n \leq s. \end{cases}$$

Ejemplo: Cadena de inventarios

Cuando el inventario cae por debajo de s , el sistema reabastece hasta S y luego descuenta la demanda del siguiente período. Cuando está por encima de s , el inventario simplemente baja según la demanda.

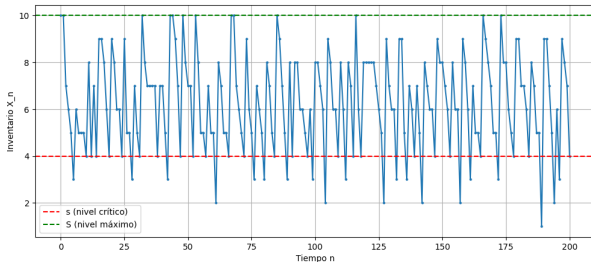


Figura: Simulación de una cadena de inventarios con $S = 10$ y $s = 4$.

★ ¿Cómo calcular p_{ij} ?

Ecuación de Chapman-Kolmogorov

Proposición

Para cualquier par de número enteros r y n tales que $0 \leq r \leq n$, y para cualesquiera estados i y j se cumple que

$$p_{ij}(n) = \sum_k p_{ik}(r) p_{kj}(n-r).$$

Demostración. Por el teorema de probabilidad total y la propiedad de Markov:

$$\begin{aligned} p_{ij}(n) &= P(X_n = j \mid X_0 = i) \\ &= \sum_k \frac{P(X_n = j, X_r = k, X_0 = i)}{P(X_0 = i)} \\ &= \sum_k P(X_n = j \mid X_r = k) P(X_r = k \mid X_0 = i) \\ &= \sum_k p_{kj}(n-r) p_{ik}(r). \end{aligned}$$

Ejercicios: Consecuencias

Proposición ★

Para cualesquiera estados i, j y k , y para $0 \leq r \leq n$, se tiene que

$$p_{ij}(n) \geq p_{ik}(r)p_{kj}(n-r).$$

Proposición ★

La probabilidad de transición en n pasos, $p_{ij}(n)$, está dada por $p_{ij}(n) = (P^n)_{ij}$.

Proposición ★

Si $p_i = P(X_0 = i)$ es una distribución inicial, entonces la distribución de la variable X_n es

$$P(X_n = j) = \sum_i p_i p_{ij}(n)$$

Transición en n pasos

Ejemplo. Consideremos la cadena general de dos estados dada por la matriz de transición

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}.$$

Esta matriz se puede descomponer de la siguiente manera:

$$\underbrace{\begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}}_P = \underbrace{\begin{pmatrix} 1 & a \\ 1 & -b \end{pmatrix}}_Q \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1-a-b \end{pmatrix}}_D \underbrace{\frac{1}{a+b} \begin{pmatrix} b & a \\ 1 & -1 \end{pmatrix}}_{Q^{-1}}$$

Por lo tanto, $P^n = QD^nQ^{-1}$, y de esta manera se pueden encontrar las probabilidades de transición de n pasos.

Transición en n pasos

- ★ ¿Cuál es la expresión explícita de la matriz de transición de n pasos?
- ★ Cuando $a + b < 1$ (o $a + b = 1$ o $a + b > 1$), ¿cómo se comporta P^n cuando $n \rightarrow \infty$?

Transición en n pasos

Ejemplo. Toda matriz estocástica P con renglones idénticos es idempotente, es decir, para cualquier entero $n \geq 1$, se cumple que $P^n = P$. Este es el caso de la *cadena de variables aleatorias independientes*.

Demostración. Sea P una matriz estocástica de tamaño $k \times k$ cuyos renglones son iguales a un mismo vector $\pi = (\pi_1, \dots, \pi_k)$, $\pi_j \geq 0$, $\sum_j \pi_j = 1$. Entonces $P = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}$.

Calculemos la entrada (i, j) de P^2 :

$$(P^2)_{ij} = \sum_{\ell=1}^k P_{i\ell} P_{\ell j} = \sum_{\ell=1}^k \pi_{\ell} \pi_j = \pi_j \sum_{\ell=1}^k \pi_{\ell} = \pi_j.$$

Pero π_j es exactamente P_{ij} , así que $P^2 = P$. Repetiendo el mismo argumento inductivamente, $P^n = P$ para todo $n \geq 1$.

Definición

Se dice que el estado j es accesible desde el estado i si existe un entero $n \geq 0$ tal que $p_{ij}(n) > 0$, esto se denota como $i \rightarrow j$. Además, los estados i y j son comunicantes si se cumple que $i \rightarrow j$ y $j \rightarrow i$, se escribe como $i \leftrightarrow j$.

Proposición

La relación $\leftrightarrow$ es una relación de equivalencia sobre el espacio de estados de la cadena de Markov.

Ejemplo. Considere la cadena de Markov definida por la matriz de transición

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

★ ¿Cuáles son las clases de equivalencias generadas por la relación de comunicación para esta cadena?

Definición

Una cadena de Markov es *irreducible* si todos los estados se comunican entre sí.

Periodo

Definición

El periodo de un estado i es un número entero no negativo denotado por $d(i)$, y definido por

$$d(i) = \text{m.c.d.}\{n \geq 1 \mid p_{ii}(n) > 0\}$$

en donde m.c.d. significa máximo común divisor. Cuando $p_{ii}(n) = 0$ para toda $n \geq 1$, se define $d(i) = 0$. En particular, se dice que un estado i es *aperiódico* si $d(i) = 1$. Cuando $d(i) = k \geq 2$ se dice que i es periódico de periodo k .

★ ¿Cuál es el periodo de cada estado de una cadena de Markov definida por la matriz

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}?$$

Proposición ★

Si los estados i y j pertenecen a la misma clase de comunicación entonces tienen el mismo periodo.

Demostración. El resultado es válido para $i = j$. Suponga que $i \neq j$, dado que los estados están en la misma clase de comunicación, existen enteros $n, m \geq 1$ tales que $p_{ij}(n) > 0$ y $p_{ji}(m) > 0$. Sea $s \geq 1$ un entero cualquier tal que $p_{ii}(s) > 0$ (...)

★ ¿El recíproco de esta proposición es verdadero?

Primeras visitas

Definición

Sea A un subconjunto del espacio de estados de una cadena de Markov $\{X_n \mid n \geq 0\}$. El tiempo de primera visita al conjunto A es la variable aleatoria

$$\tau_A = \begin{cases} \min\{n \geq 1 : X_n \in A \mid X_0 = i\} & \text{si } X_n \in A \text{ para algún } n \geq 1, \\ \infty & \text{en otro caso.} \end{cases}$$

Definición

Para cada $n \geq 1$, el número $f_{ij}(n)$ denota la probabilidad de que una cadena que inicia en el estado i llegue al estado j por primera vez en n pasos, es decir,

$$f_{ij}(n) = P(X_n = j, X_{n-1} \neq j, \dots, X_1 \neq j \mid X_0 = i).$$

Adicionalmente se define $f_{ij}(0) = 0$, incluyendo el caso $i = j$.

- ★ ¿Es cierto que $f_{ij}(n) = P(\tau_{ij} = n)$?
- ★ ¿Es cierto que $f_{ij}(1) = p_{ii}(1)$?
- ★ Considere la cadena de dos estados. Pruebe que $f_{01}(n) = (1 - a)^{n-1}a$ y $f_{10}(n) = (1 - b)^{n-1}b$ para $n \geq 1$, y, $f_{00}(n) = a(1 - b)^{n-2}b$ y $f_{11}(n) = b(1 - a)^{n-2}a$ para $n \geq 2$.

Proposición ★

Para cada $n \geq 1$,

$$p_{ij}(n) = \sum_{k=1}^n f_{ij}(k)p_{jj}(n-k).$$

Primeras visitas

Definición

Se define $f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$ como la probabilidad de una eventual visita al estado j a partir del estado i .

★ ¿El número f_{ij} se encuentra bien definido, es decir, es una probabilidad?

Recurrencia y transitoriedad

Definición

Un estado i es *recurrente* si $f_{ii} = 1$, es decir si la probabilidad de regresar a él en un tiempo finito es uno. Por otra parte, un estado i es *transitorio* si $f_{ii} < 1$.

Recurrencia y transitoriedad

Proposición

El estado i es:

1. recurrente si y sólo si $\sum_{n \geq 1} p_{ii}(n) = \infty$.
2. transitorio si y sólo si $\sum_{n \geq 1} p_{ii}(n) < \infty$.

Demostración. Sea N_i la variable aleatoria que cuenta el número de veces que el proceso está en el estado i a partir del primer paso, así, $N_i = \sum_{n=1}^{\infty} 1_{\{X_n=i\}}$, cuando $X_0 = i$. Entonces N_i tiene distribución geométrica, pues

$$P(N_i \geq k \mid X_0 = i) = \sum_{m=1}^{\infty} P(N_i \geq k-1 \mid X_0 = i) f_{ii}(m) = (f_{ii})^k$$

La esperanza de N_i , posiblemente infinita, puede calcularse de la siguientes dos maneras. Primero,

$$\begin{aligned} E(N_i \mid X_0 = i) &= \sum_{k=1}^{\infty} P(N_i \geq k \mid X_0 = i) \\ &= \sum_{k=1}^{\infty} (f_{ii})^k \\ &= \begin{cases} \frac{f_{ii}}{1-f_{ii}} & \text{si } 0 \leq f_{ii} < 1, \\ \infty & \text{si } f_{ii} = 1. \end{cases} \end{aligned}$$

Por otro lado, por el teorema de convergencia monótona de Lebesgue,

$$\begin{aligned} E(N_i \mid X_0 = i) &= \sum_{n=1}^{\infty} E(1_{\{X_n=i\}} \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} P(X_n = i \mid X_0 = i) \\ &= \sum_{n=1}^{\infty} p_{ii}(n). \end{aligned}$$

El resultado se sigue de igualar las anteriores dos expresiones.

□

Hombre borracho vs Ave borracha

Vamos a probar que en dos dimensiones para una caminata aleatoria simétrica el estado inicial es recurrente, y para tres dimensiones una caminata aleatoria simétrica el estado inicial es transitorio.

Consideremos una caminata aleatoria en el conjunto $\mathbb{Z} \times \mathbb{Z}$ con inicio en $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ con el siguiente conjunto de pasos $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}$ cada uno con la misma probabilidad.

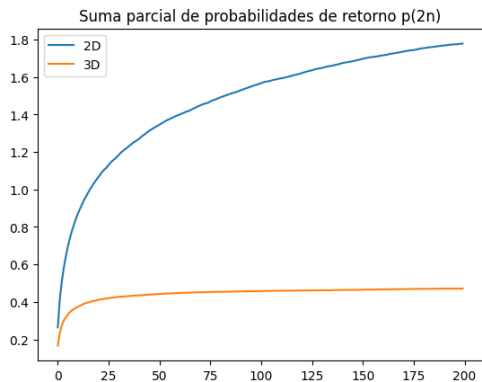
★ Para una trayectoria de longitud $2n$, se cumple que

$$p_{\begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}}(2n) = \left(\frac{1}{4}\right)^{2n} \sum_{i=1}^n \frac{(2n)!}{i! i! (n-i)!(n-i)!},$$

¿cómo se deduce esta fórmula?

Hombre borracho vs Ave borracha

- ★ Use la fórmula de Stirling $n! \approx \sqrt{2\pi} n^{\frac{n+1}{2}} e^{-n}$ para aproximar $p\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix}(2n)$.
- ★ Compruebe que $\sum_{n=1}^{\infty} p\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix}(2n) = \infty$.
- ★ Siga el mismo esquema de demostración para el caso en tres dimensiones y pruebe que $\sum_{n=1}^{\infty} p\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}(2n) < \infty$.



- Se definen dos funciones de caminata aleatoria: una en 2D y otra en 3D, cada una con pasos equiprobables en las direcciones básicas.
- Para un número par de pasos $2n$, se estima la probabilidad de retornar al origen mediante simulación Monte Carlo: se generan muchas trayectorias y se cuenta cuántas terminan en cero.
- Con estas estimaciones se construye la suma parcial $S_d(N) = \sum_{n=1}^N p_{00}(2n)$, que permite visualizar la diferencia fundamental: en 2D la suma crece sin límite (recurrencia), mientras que en 3D se estabiliza (transitoriedad).

Recurrencia y transitoriedad

Proposición

La recurrencia es una propiedad de clase, es decir,

1. Si i es recurrente y $i \leftrightarrow j$ entonces j es recurrente.
2. Si i es transitorio y $i \leftrightarrow j$ entonces j es transitorio.

Demostración. Si $i \leftrightarrow j$ entonces existen enteros $m, n \geq 1$ tales que $p_{ij}(n), p_{ji}(m) > 0$. Entonces $p_{jj}(m+n+r) \geq p_{ji}(m)p_{ii}(r)p_{ij}(n)$, de modo que,

$$\sum_{r=1}^{\infty} p_{jj}(m+n+r) \geq p_{ji}(m)p_{ij}(n) \sum_{r=1}^{\infty} p_{ii}(r)$$

luego, como i es recurrente entonces $\sum_{r=1}^{\infty} p_{ii}(r) = \infty$ lo que implica que $\sum_{r=1}^{\infty} p_{jj}(m+n+r) = \infty$. La segunda afirmación se demuestra por contradicción usando el primer resultado. □

Recurrencia y transitoriedad

Ejemplo. La cadena de dos estados es irreducible y recurrente cuando $a, b \in (0, 1)$. En efecto, tenemos que $f_{00}(1) = 1 - a$ y $f_{00}(n) = a(1 - b)^{n-2}b$ para $n \geq 2$. Por tanto,

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = (1 - a) + ab \sum_{n=2}^{\infty} (1 - b)^{n-2} = (1 - a) + \frac{ab}{b} = 1.$$

Recurrencia y transitoriedad

Ejemplo. Considere la cadena de rachas de éxitos. En este caso es sencillo demostrar que el estado 0 es recurrente pues

$$f_{00} = \sum_{n=1}^{\infty} f_{00}(n) = q(1 + p + p^2 + \dots) = \frac{q}{1-p} = 1.$$

Dado que la cadena es irreducible y la recurrencia es una propiedad de clase, cualquier otro estado es recurrente. Por lo tanto, la cadena es recurrente.

Recurrencia y transitoriedad

Proposición

Sea j un estado transitorio. Para cualquier estado inicial i se cumple que $\sum_{n=1}^{\infty} p_{ij}(n) < \infty$. En consecuencia $\lim_{n \rightarrow \infty} p_{ij}(n) = 0$.

Demostración. Dado que todos los términos son no negativos:

$$\begin{aligned}\sum_{n=1}^{\infty} p_{ij}(n) &= \sum_{n=1}^{\infty} \sum_{k=0}^{n-1} f_{ij}(n-k) p_{jj}(k) \\ &= \sum_{k=0}^{\infty} \sum_{n=k+1}^{\infty} f_{ij}(n-k) p_{jj}(k) = \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} f_{ij}(m) p_{jj}(k) \\ &= f_{ij} \sum_{k=0}^{\infty} p_{jj}(k) \leq \sum_{k=0}^{\infty} p_{jj}(k) < \infty.\end{aligned}$$



Recurrencia y transitoriedad

Proposición

Toda cadena de Markov finita tiene por lo menos un estado recurrente.

Demostración. Por contradicción, suponga que todos los estados son transitorios. Entonces para cualesquiera estados i y j , se cumple que la suma $\sum_{n=1}^{\infty} p_{ij}(n)$ es finita. Sumando sobre el conjunto finito de todos los posibles estados j se obtiene:

$$\sum_j \sum_{n=1}^{\infty} p_{ij}(n) < \infty.$$

Por otro lado, intercambiando las sumatorias se llega a la afirmación contraria:

$$\sum_j \sum_{n=1}^{\infty} p_{ij}(n) = \sum_{n=1}^{\infty} \sum_j p_{ij}(n) = \sum_{n=1}^{\infty} 1 = \infty.$$

Lo cual es una contradicción.



Tiempo medio de recurrencia

Definición

El tiempo medio de recurrencia de un estado recurrente j , a partir del estado i , se define como la esperanza de la variable aleatoria

$$\tau_{ij} = \min\{n \geq 1 : X_n = j \mid X_0 = i\},$$

y se denota por μ_{ij} , es decir,

$$\mu_{ij} = E(\tau_{ij}) = \sum_{n=1}^{\infty} n f_{ij}(n).$$

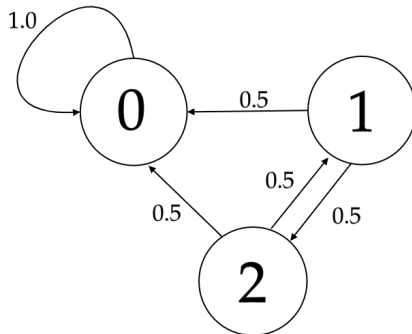
Tiempo medio de recurrencia

Ejemplo. La cadena de dos estados es irreducible y recurrente cuando $a, b \in (0, 1)$. Tenemos que $f_{00}(1) = 1 - a$ y $f_{00}(n) = a(1 - b)^{n-2}b$ para $n \geq 2$. Por tanto,

$$\begin{aligned}\mu_{00} &= \sum_{n=1}^{\infty} n f_{00}(n) = (1 - a) + ab \sum_{n=2}^{\infty} n(1 - b)^{n-2} \\ &= (1 - a) + ab \left(\frac{b + 1}{b^2} \right) \\ &= \frac{a + b}{b}.\end{aligned}$$

- ★ Demuestre que $\frac{1}{\mu_{00}} + \frac{1}{\mu_{11}} = 1$.
- ★ ¿El tiempo medio de recurrencia es una propiedad de clase?

Ejercicio



Considere la cadena de Markov representada por el grafo en la figura anterior. Para cualesquiera estados i, j halle el valor de f_{ij} , $\sum_{n=1}^{\infty} p_{ij}(n)$ y de μ_{ij} , luego de esto, haga una clasificación de estados.

El problema del mono

Un mono escribe caracteres al azar, cada uno elegido uniformemente de un alfabeto de tamaño m . Queremos saber si el mono es capaz de reproducir con exactitud una obra de Shakespeare. Las obras completas de este autor tienen N caracteres.

Definimos X_n como el número de caracteres correctos consecutivos alcanzados hasta el tiempo n . Entonces (X_n) es una cadena de Markov de **racha de éxitos** con estados $\{0, 1, \dots, N\}$:

$$P(X_{n+1} = x + 1 \mid X_n = x) = \frac{1}{m}, \quad P(X_{n+1} = 0 \mid X_n = x) = \frac{m - 1}{m}.$$

La cadena es finita, irreducible y recurrente. Por tanto, cada estado es visitado infinitamente muchas veces con probabilidad uno.

En particular, el estado N (escribir toda la obra sin error) ocurre infinitas veces con probabilidad uno.

Definición

El número de visitas al estado j de la cadena de Markov partiendo de un estado i se define como la variable aleatoria $N_{ij}(n)$ cuando se han dado $n \geq 1$ pasos:

$$N_{ij}(n) = \sum_{k=1}^n 1_{\{X_k=j|X_0=i\}}.$$

Cuando $n \rightarrow \infty$ se define:

$$N_{ij} = \sum_{k=1}^{\infty} 1_{\{X_k=j|X_0=i\}}.$$

Teorema ergódico para cadenas de Markov

Proposición

Para cualesquiera estados i y j de una cadena de Markov irreducible se cumple que

$$\lim_{n \rightarrow \infty} \frac{N_{ij}(n)}{n} = \frac{1}{\mu_j} \quad \text{casi seguramente}$$

donde $N_{ij}(n)$ es el número de visitas al estado j iniciando desde i hasta el tiempo n , y μ_j es el tiempo de retorno esperado al estado j . Este límite es cero cuando $\mu_j = \infty$.

★ ¿Cómo se interpreta este teorema?

Recurrencia positiva y nula

Definición

Se dice que un estado recurrente i es:

1. recurrente positivo si $\mu_{ii} < \infty$.
2. recurrente nulo si $\mu_{ii} = \infty$.

Proposición ★

La recurrencia positiva y la recurrencia nula es una propiedad de clase.

Recurrencia positiva y nula

Proposición ★

No existen estados recurrentes nulos en cadenas de Markov finitas.

Ejemplo. La caminata aleatoria sobre $\mathbb{Z}$ satisface que

$$\mu_{00} = \sum_{n=1}^{\infty} n f_{00}(n) = \frac{4pq}{\sqrt{1-4pq}}.$$

Por ende, si la caminata es simétrica entonces el estado origen es recurrente positivo.

Demostración. Sea $f_{00}(n)$ la probabilidad de *primer* retorno al origen en tiempo n y

$$F(s) = \sum_{n \geq 1} f_{00}(n) s^n.$$

Sea también $u_n = \mathbb{P}(X_n = 0 \mid X_0 = 0)$. Para la caminata simple en $\mathbb{Z}$:

$$u_{2n} = \binom{2n}{n} (pq)^n, \quad u_{2n+1} = 0.$$

Su serie generadora es

$$U(s) = \sum_{n \geq 0} u_n s^n = \sum_{n \geq 0} \binom{2n}{n} (pq s^2)^n = \frac{1}{\sqrt{1-4pq s^2}}.$$

La relación de renovación entre visitas y primeros retornos da:

$$U(s) = \frac{1}{1 - F(s)} \quad \Rightarrow \quad F(s) = 1 - \frac{1}{U(s)} = 1 - \sqrt{1 - 4pq s^2}.$$

Como la esperanza existe cuando la probabilidad de retorno es 1, se tiene

$$\mu_{00} = F'(1).$$

Derivando:

$$F'(s) = \frac{4pq s}{\sqrt{1 - 4pq s^2}},$$

y evaluando en $s = 1$ se obtiene

$$\mu_{00} = F'(1) = \frac{4pq}{\sqrt{1 - 4pq}}.$$

En el caso simétrico $p = q = \frac{1}{2}$, el denominador se anula y $\mu_{00} = +\infty$; el origen es recurrente pero no positivo recurrente.

i	f_{ii}	$\sum_{n=1}^{\infty} p_{ii}(n)$	μ_{ii}
Recurrente positivo			
Recurrente nulo			
Transitorio			

Evolución de distribuciones

Una matriz estocástica define una dinámica sobre el conjunto de distribuciones de probabilidad de una cadena de Markov.

Consideremos un espacio de estados finito $\{0, 1, \dots, n\}$, y una distribución inicial

$$\pi^{(0)} = (\pi_0^{(0)}, \pi_1^{(0)}, \dots, \pi_n^{(0)}).$$

Después de un paso, la cadena tiene una nueva distribución $\pi^{(1)}$, donde la componente j satisface

$$\pi_j^{(1)} = \sum_{i=0}^n \pi_i^{(0)} p_{ij}.$$

Evolución de distribuciones

La actualización anterior puede escribirse de manera compacta como

$$\pi^{(1)} = \pi^{(0)} P,$$

donde $P = (p_{ij})$ es la matriz de probabilidades de transición.

$$(\pi_0^{(1)}, \dots, \pi_n^{(1)}) = (\pi_0^{(0)}, \dots, \pi_n^{(0)}) \begin{pmatrix} p_{00} & \cdots & p_{0n} \\ \vdots & \ddots & \vdots \\ p_{n0} & \cdots & p_{nn} \end{pmatrix}.$$

Esta regla expresa que las distribuciones evolucionan linealmente bajo la acción de P .

Evolución de distribuciones

Aplicando la misma regla en cada paso se obtiene

$$\pi^{(2)} = \pi^{(1)}P = \pi^{(0)}P^2,$$

y en general, para cualquier $m \geq 1$,

$$\pi^{(m)} = \pi^{(m-1)}P = \pi^{(0)}P^m.$$

Así, la matriz P gobierna completamente la evolución temporal de la distribución de estados de la cadena.

Ejemplo. Consideremos la matriz estocástica

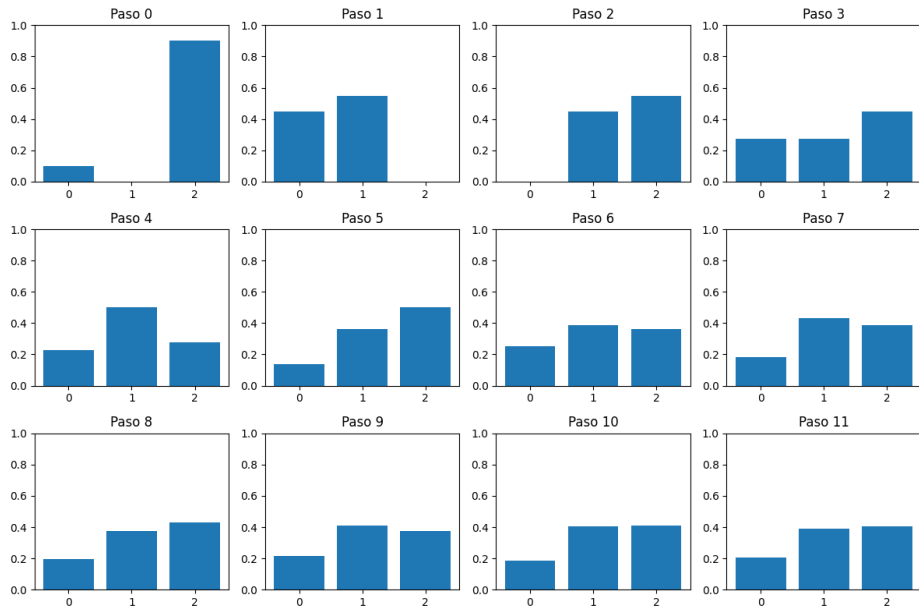
$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \quad \pi^{(0)} = \left(\frac{1}{10}, 0, \frac{9}{10} \right).$$

Cada distribución evoluciona como

$$\pi^{(m)} = \pi^{(m-1)}P = \pi^{(0)}P^m.$$

Los primeros vectores son:

$$\begin{aligned} \pi^{(1)} &= (0,45, 0,55, 0), & \pi^{(2)} &= (0, 0,45, 0,55), \\ \pi^{(3)} &= (0,275, 0,275, 0,45), & \pi^{(4)} &= (0,225, 0,50, 0,275). \end{aligned}$$



Ejemplo. Considere la matriz estocástica

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

con distribución inicial $\pi^{(0)} = (\alpha, 1 - \alpha)$ con $\alpha \in [0, 1]$. Por tanto, la sucesión de distribuciones de probabilidad es

$$\pi^{(1)} = (1 - \alpha, \alpha)$$

$$\pi^{(2)} = (\alpha, 1 - \alpha)$$

$$\pi^{(3)} = (1 - \alpha, \alpha)$$

$$\pi^{(4)} = (\alpha, 1 - \alpha)$$

$\vdots$

Esta sucesión de distribuciones no es convergente a menos que $\alpha = 1/2$.

Distribuciones estacionarias

Definición

Una distribución de probabilidad $\pi = (\pi_0, \pi_1, \dots)$ es estacionaria o invariante para una cadena de Markov con matriz de probabilidades de transición $P = (p_{ij})$ si $\pi_j = \sum_i \pi_i p_{ij}$.

- Es equivalente a afirmar que $\pi = \pi P$ en el caso finito.
- Si X_0 tiene distribución inicial π entonces X_n tiene distribución π para $n \geq 1$.

Distribuciones estacionarias

Ejemplo (no existencia). Para la caminata aleatoria simétrica sobre $\mathbb{Z}$ no existe ninguna distribución estacionaria pues la condición $\pi = \pi P$ se traduce en el sistema de ecuaciones

$$\pi_j = \frac{1}{2}\pi_{j-1} + \frac{1}{2}\pi_{j+1}, \quad j \in \mathbb{Z}$$

o bien, $\pi_{j+1} - \pi_j = \pi_j - \pi_{j-1}$, lo que implica que para todo entero $n \geq 1$ se tiene que $\pi_n - \pi_0 = n(\pi_1 - \pi_0)$ de lo cual se deduce que $\pi_1 - \pi_0 = 0$. Por tanto π_j es constante para todo $j \in \mathbb{Z}$, pero esto es incompatible con la condición de que $\sum_j \pi_j = 1$.

Distribuciones estacionarias

Ejemplo (múltiple existencia). Considere la cadena de Markov definida sobre el conjunto de estados $\{0, 1, 2\}$ y con probabilidades de transición dada por la siguiente matriz:

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}.$$

Es fácil comprobar que para cualquier $\alpha \in [0, 1]$ la distribución $\pi = (1 - \alpha, 0, \alpha)$ satisface el sistema de ecuaciones $\pi = \pi P$.

Distribuciones estacionarias

Ejemplo (existencia única). La cadena de dos estados dada por la matriz

$$P = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

tiene una única distribución estacionaria dada por

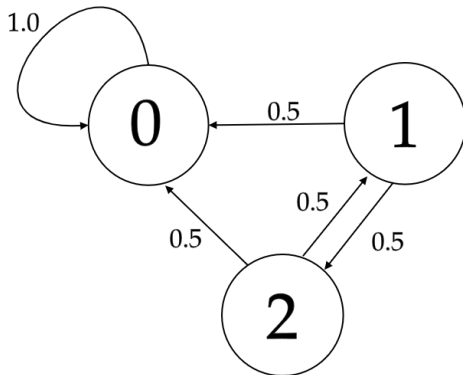
$$\pi = (\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b} \right)$$

cuando $a+b > 0$.

★ ¿Qué ocurre cuando $a+b = 0$?

Ejercicio

Halle una distribución estacionaria para la cadena de Markov representada por el grafo a continuación.



Distribuciones estacionarias

- ★ Demuestre que el conjunto de distribuciones estacionarias es convexo.
- ★ Demuestre que el conjunto de distribuciones estacionarias es vacío ó solo consta de un elemento ó es infinito.

Distribuciones estacionarias

Proposición (Soporte de una distribución estacionaria).

Sea π una distribución estacionaria para una cadena de Markov. Si j es un estado transitorio o recurrente nulo, entonces $\pi_j = 0$.

Demostración. Si j es un estado transitorio o recurrente nulo, entonces para cualquier estado i se tiene que $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n p_{ij}(k) = 0$. Como π es una distribución estacionaria entonces

$$\pi_j = \sum_i \pi_i p_{ij} \leq \sum_i \pi_i \left(\frac{1}{n} \sum_{k=1}^n p_{ij}(k) \right),$$

así, cuando $n \rightarrow \infty$ se tiene que $\pi_j \rightarrow 0$. □

Distribuciones estacionarias

Proposición (Existencia y unicidad de una distribución estacionaria).

Toda cadena de Markov irreducible y recurrente positiva tiene una única distribución estacionaria dada por $\pi_j = \frac{1}{\mu_{jj}} > 0$, en donde μ_{jj} es el tiempo medio de recurrencia al estado j . En particular toda cadena finita e irreducible tiene una única distribución estacionaria.

Demostración. Sea $\pi_j = \frac{1}{\mu_{jj}}$. Demostraremos que

1. $\pi_j = \sum_i \pi_i p_{ij}$.
 2. $\sum_i \pi_i = 1$.
 3. π_j es única.
- (...)

Distribuciones estacionarias

Ejemplo. La cadena de Markov de dos estados es finita e irreducible cuando $a + b > 0$ y por lo tanto recurrente positiva. Sabemos que por la proposición anterior existe una única distribución estacionaria para esta cadena que está dada por $\pi = \pi P$ y $\sum_j \pi_j = 1$, o de manera explícita,

$$\pi = (\pi_0, \pi_1) = \left(\frac{b}{a+b}, \frac{a}{a+b} \right).$$

Por otra parte, por el resultado anterior también podemos calcular

$$(\mu_{00}, \mu_{11}) = \left(\frac{1}{\pi_0}, \frac{1}{\pi_1} \right) = \left(\frac{a+b}{b}, \frac{a+b}{a} \right).$$

Distribuciones estacionarias

Ejemplo. La cadena de Ehrenfest es finita e irreducible y en consecuencia recurrente positiva. Por lo tanto existe una única distribución estacionaria. Resolviendo el sistema de ecuaciones $\pi = \pi P$ junto con $\sum_j \pi_j = 1$ se encuentra que la distribución estacionaria π tiene distribución Binomial(n, p) con $p = \frac{1}{2}$, es decir,

$$\pi_j = \binom{N}{j} \frac{1}{2^N}$$

Distribuciones estacionarias

Para fines ilustrativos tomemos $N = 2$, es decir, se tienen únicamente dos bolas distribuidas en las dos urnas de la cadena de Ehrenfest, entonces la distribución estacionaria es $\pi = (\pi_0, \pi_1, \pi_2) = \left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right)$.

En vista del teorema ergódico, esto significa que, a largo plazo, la cadena de Ehrenfest se encontrará en el estado 1 el 50 % del tiempo, y en los estados 0 y 2 el 25 % del tiempo en cada uno de ellos. Los tiempos medios de recurrencia son $\mu = (\mu_{00}, \mu_{11}, \mu_{22}) = (4, 2, 4)$, es decir, si por ejemplo la cadena se encuentra en algún momento en el estado 0, entonces tardará en promedio 4 tiempos para regresar nuevamente a ese estado.

Ejercicio

Considere la cadena de rachas de éxitos. Halle explícitamente una distribución estacionaria para esta cadena de Markov y calcule los tiempos medios de recurrencia.

Distribuciones límite

Definición

Considere una cadena de Markov con matriz de probabilidades de transición $P = (p_{ij})$ y distribución inicial π^0 . Se llama distribución límite de la cadena a

$$\pi = \lim_{n \rightarrow \infty} \pi^0 P^n, \quad \pi_j = \sum_i \pi_i^0 \left[\lim_{n \rightarrow \infty} p_{ij}(n) \right].$$

★ Pruebe que toda distribución límite es estacionaria. ¿El recíproco es cierto?

Distribuciones límite

Ejemplo. Se puede comprobar que la distribución $\left(\frac{1}{2}, \frac{1}{2}\right)$ es estacionaria para la cadena con la matriz de probabilidades de transición

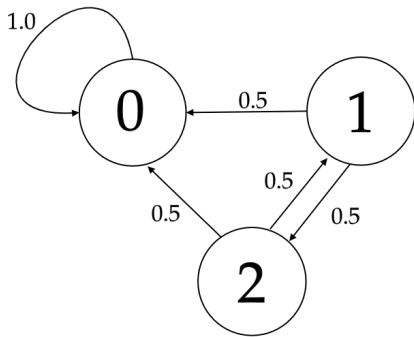
$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Sin embargo, las potencias de P no convergen puesto que

$$P^{2n+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad y \quad P^{2n} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Ejercicio

Halle el límite de la sucesión $\{P^n\}_{n \in \mathbb{N}}$ para la cadena de Markov representada por el grafo a continuación. Luego de hallar la matriz límite $\hat{P}$ de la sucesión calcule el espacio propio de $\hat{P}^\top$ asociado al autovalor 1. ¿La distribución estacionaria de esta cadena de Markov se encuentra en este espacio?



Proposición (Límite $\rightarrow$ Estacionaria)

Considere una cadena de Markov con probabilidades de transición p_{ij} tales que $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$ existen para cada j y no dependen del estado i . Entonces

1. $\sum_j \pi_j = 1$.
2. $\pi_j = \sum_i \pi_i p_{ij}$.

Cuando el espacio de estados es finito se cumple la igualdad en el primer resultado obteniendo una distribución de probabilidad verdadera.

Proposición ★

Considere una cadena de Markov que es irreducible, aperiódica y con distribución estacionaria π . Entonces para cualesquiera estados i y j se tiene que $\lim_{n \rightarrow \infty} p_{ij}(n) = \pi_j$.

Distribuciones límite

Proposición

Considere una cadena de Markov que es irreducible, recurrente positiva y aperiódica. Entonces las probabilidades límite $\pi_j = \lim_{n \rightarrow \infty} p_{ij}(n)$ existen y están dadas por $\pi_j = 1/\mu_{jj}$ y constituyen la única solución al sistema de ecuaciones $\pi_j = \sum_i \pi_i p_{ij}$, sujeto a las condiciones $\pi_j \geq 0$ y $\sum_j \pi_j = 1$.

Demostración. Como la cadena es irreducible y recurrente positiva, tiene una distribución estacionaria dada por $\pi_j = \frac{1}{\mu_{jj}}$. Es decir, es la única solución al sistema de ecuaciones $\pi = \pi P$ con $\pi_j \geq 0$ y $\sum_j \pi_j = 1$. Además por aperiodicidad, $p_{ij}(n) \rightarrow \pi_j$. $\square$

Cadenas regulares

Definición

Cadenas regulares

Proposición

Demostración.

Cadenas regulares

Proposición

Demostración.

Cadenas reversibles

Ejercicio

Demuestre que

1. Si P y Q son dos matrices estocásticas (o doblemente estocásticas) de la misma dimensión, entonces PQ también es estocástica (o doblemente estocástica).
2. Si P es estocástica (o doblemente estocástica) entonces para cualquier entero positivo n se tiene que P^n es estocástica (o doblemente estocástica).

Ejercicio

Sea $\{X_n : n \geq 0\}$ una cadena de Markov con dos estados $\{0, 1\}$ con distribución inicial $P(X_0 = 0) = \frac{1}{2}$ y $P(X_0 = 1) = \frac{1}{2}$ y con matriz de probabilidades de transición

$$P = \begin{pmatrix} 0,4 & 0,6 \\ 0,9 & 0,1 \end{pmatrix}.$$

Calcule

1. $P(X_0 = 0, X_1 = 1, X_2 = 0)$.
2. $P(X_0 = 0 \mid X_1 = 1)$.
3. $P(X_1 = 0)$.
4. $P(X_2 = 1)$.
5. $P(X_0 = 0, X_1 = 0, X_2 = 1, X_3 = 1)$.
6. La distribución de X_2 .
7. La distribución conjunta de X_1 y X_2 .
8. La esperanza condicional $E(X_2 \mid X_1)$.

Ejercicio

Sea $\{\xi_1, \xi_2, \dots\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, con valores en el conjunto $\{1, 2, \dots\}$. Interpretamos estas variables como los tiempos de vida de artículos puestos en operación uno después de otro en un proceso de renovación en tiempo discreto. Sea $X_0 = 0$ y para cada $n \geq 1$ definimos $X_n = \xi_1 + \dots + \xi_n$. Para cada $k \in \{1, 2, \dots\}$ definimos $N_k = \max\{n \geq 1 : X_n \leq k\}$. Si el conjunto es vacío, tomamos $N_k = 0$. Así, N_k cuenta el número de renovaciones efectuadas hasta el tiempo k . Definimos también $A_k = k - X_{N_k}$, es decir, la edad del artículo en operación en el instante k , o el tiempo transcurrido desde la última renovación.

Demuestre que el proceso $\{A_k : k \geq 1\}$ es una cadena de Markov con espacio de estados $\{0, 1, 2, \dots\}$ y probabilidades de transición

$$p_{i,0} = \frac{P(X = i + 1)}{P(X \geq i + 1)}, \quad p_{i,i+1} = \frac{P(X \geq i + 2)}{P(X \geq i + 1)}.$$

Ejercicio

Sea $\{X_n : n \geq 0\}$ una cadena de Markov de dos estados. Demuestre que el proceso $Y_n = (X_{n+1}, X_n)$ es una cadena de Markov.

Determine:

- El espacio de estados del proceso $\{Y_n\}$.
- La matriz de probabilidades de transición de esta nueva cadena.

Finalmente, generalice este resultado para cualquier cadena con espacio de estados $\{0, 1, \dots\}$.

Ejercicio

Se colocan N bolas negras en una primera urna, y N bolas blancas en una segunda urna. Se selecciona al azar una bola de cada urna y se intercambian. Este ensayo se repite varias veces.

Sea X_n el número de bolas blancas en la primera urna después del n -ésimo ensayo.

- Justifique que $\{X_n : n \geq 1\}$ es una cadena de Markov.
- Encuentre la matriz de probabilidades de transición.

Ejercicio

Sea $\xi_0, \xi_1, \dots$ una sucesión de variables independientes con idéntica distribución $\text{Ber}(p)$. Determine si el proceso $\{X_n : n \geq 1\}$ definido a continuación es una cadena de Markov.

En caso afirmativo, determine su espacio de estados y la matriz de probabilidades de transición.

- a) $X_n = \xi_n + \xi_{n-1}$.
- b) $X_n = X_{n-1} + \xi_n$.
- c) $X_n = X_{n-1} + \xi_n \pmod{2}$.

Ejercicio

Dibuje un diagrama de transición, determine las clases de comunicación y calcule el periodo de cada uno de los estados para la cadena con matriz de transición

a)

$$P = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

b)

$$P = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}.$$

Ejercicio

Una persona se traslada todos los días de su casa a la oficina en la mañana, y de la oficina a su casa por la tarde. Tiene un coche que usa en cualquiera de estos dos viajes en caso de lluvia, siempre y cuando el coche esté disponible. A veces el coche no está disponible porque la persona decide dejarlo en la casa o en la oficina y caminar cuando al salir no está lloviendo. La probabilidad de lluvia en la mañana o en la tarde es $p > 0$, independiente entre sí.

- a) Demuestre que la proporción de viajes a largo plazo en los cuales la persona se moja es $\frac{p(1-p)}{p(2-p)}$.
- b) Demuestre que cuando la persona posee r coches, la proporción a largo plazo de viajes en los cuales se moja es $\frac{p(1-p)}{1-(1-p)^r}$.

Métodos de Monte Carlo basados en cadenas de Markov (MCMC)

Convergencia de métodos MCMC

Conteo aproximado

Algoritmo Propp-Wilson

Sandwiching

Propp-Wilson y aleatoriedad de lectura única

Recocido simulado

- Hidden Markov chains
- Higher order Markov chains
- Multivariate Markov chains
- Markov decision processes
- Applications

Procesos de Poisson

Proceso de Poisson

Primera definición

Sea $T_1, T_2, \dots$ una sucesión de variables aleatorias independientes con distribución Exponencial(λ). El proceso de Poisson de parámetro λ es el proceso a tiempo continuo $\{X_t : t \geq 0\}$ definido de la siguiente manera:

$$X_t = \max\{n \geq 1 : T_1 + T_2 + \dots + T_n \leq t\}.$$

- El proceso inicia en cero y para ello se define $\max \emptyset = 0$.
- La variable X_t es el entero n máximo tal que $T_1 + \dots + T_n$ es menor o igual a t .
- $T_1, T_2, \dots$ se llaman tiempos de estancia.
- $T_1, T_2, \dots \stackrel{\text{iid}}{\sim} \text{Exponencial}(\lambda)$, en consecuencia la variable $W_n = T_1 + \dots + T_n$ tiene distribución $W_n \sim \text{Gamma}(n, \lambda)$. Luego, $(X_t \geq n) = (W_n \leq t)$.

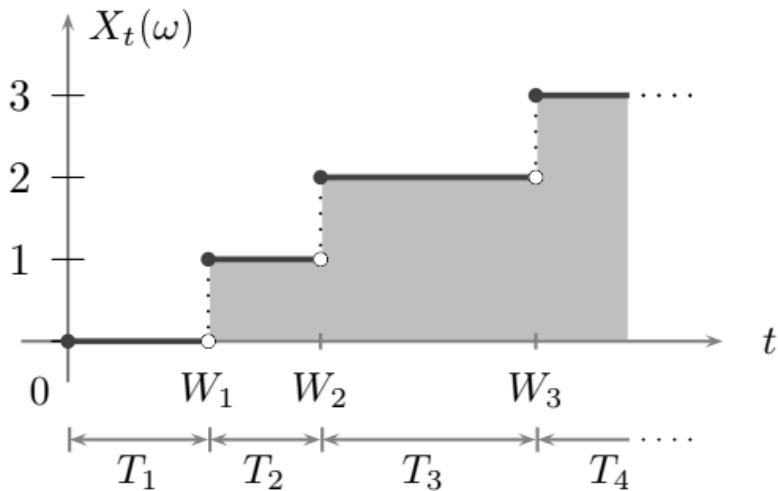


Figura: El tiempo de Poisson y los tiempos de ocurrencia del evento. Tomado de Rincon, 2012.

Proposición

La variable X_t tiene distribución Poisson(λt). ($E(X_t) = \text{Var}(X_t) = \lambda t$)

Demostración. Como $W_n \sim \text{Gamma}(n, \lambda)$, su función de distribución es, para $t > 0$,

$$P(W_n \leq t) = 1 - e^{-\lambda t} \sum_{k=0}^{n-1} \frac{(\lambda t)^k}{k!}.$$

Entonces para cualquier $t > 0$ y $n = 0, 1, \dots$

$$\begin{aligned} P(X_t = n) &= P(X_t \geq n) - P(X_t \geq n-1) \\ &= P(W_n \leq t) - P(W_{n-1} \leq t) \\ &= e^{-\lambda t} \frac{(\lambda t)^n}{n!}. \end{aligned}$$



Ejemplo. Se realiza una simulación completa del proceso de Poisson homogéneo con parámetro $\lambda = 1$ y se producen cuatro visualizaciones:

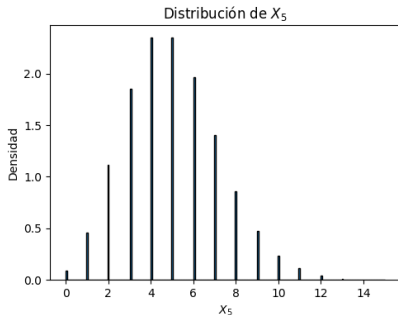
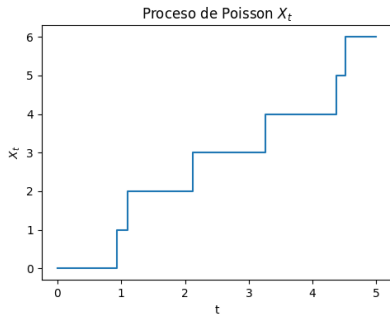
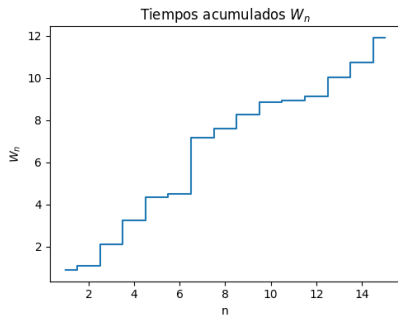
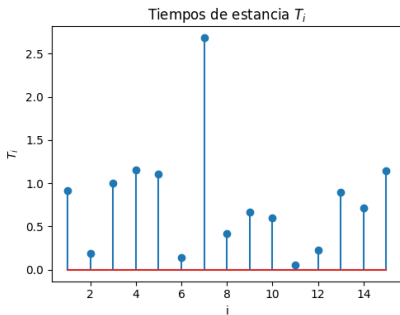
- Se generan $T_1, \dots, T_{15}$ como variables Exponencial(1) y se calcula $W_n = \sum_{i=1}^n T_i$.
- Se construye una trayectoria del proceso X_t usando la relación

$$X_t = \max\{n : W_n \leq t\}.$$

- Se realizan 50,000 simulaciones independientes para estimar la distribución empírica de X_5 .

La figura final es un panel 2×2 :

- (1,1) Tiempos de estancia T_i .
- (1,2) Tiempos de llegada acumulados W_n .
- (2,1) Trayectoria simulada del proceso X_t .
- (2,2) Histograma normalizado que aproxima la distribución de X_5 .



Pérdida de memoria. Si $T \sim \text{Exponencial}(\lambda)$ entonces para cualesquiera tiempo $s, t > 0$ se cumple la igualdad

$$P(T > s + t \mid T > s) = P(T > t).$$

Proposición

Para cualesquiera tiempos $0 \leq s < t$ y para $n = 0, 1, \dots$

$$P(X_t - X_s = n) = P(X_{t-s} = n) = e^{-\lambda(t-s)} \frac{[\lambda(t-s)]^n}{n!}.$$

Demostración.

$$\begin{aligned} P(X_t - X_s = n) &= \sum_{k=0}^{\infty} P(X_t - X_s = n \mid X_s = k) P(X_s = k) \\ &= \sum_{k=0}^{\infty} P(X_{t-s} = n) P(X_s = k) \\ &= P(X_{t-s} = n). \end{aligned}$$



Proposición ★

El proceso de Poisson $\{X_t : t \geq 0\}$ satisface las siguientes propiedades:

1. Es un proceso de Markov.
2. Tiene incrementos independientes.
3. Tiene incrementos estacionarios.
4. Para cualesquiera $s, t \geq 0$ y enteros $0 \leq i \leq j$, las probabilidades de transición son

$$P(X_{t+s} = j \mid X_s = i) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}.$$

- Las probabilidades de transición no depende del tiempo. Son estacionarias en el tiempo. Se escriben como $p_{ij}(t)$.
- Las probabilidades de transición son también estacionarias en el espacio puesto que $p_{ij}(t) = p_{0,j-i}(t)$.

Proposición ★

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de parámetro λ y sea S una variable aleatoria continua con soporte en el intervalo $(0, \infty)$ e independiente del proceso de Poisson. Entonces para cualquier $t > 0$, el incremento $X_{S+t} - X_t$ tiene distribución $\text{Poisson}(\lambda t)$.

Proposición ★

Sean $\{X_t : t \geq 0\}$ y $\{Y_t : t \geq 0\}$ dos procesos de Poisson independientes con parámetro λ_1 y λ_2 respectivamente. El proceso $\{X_t + Y_t : t \geq 0\}$ es un proceso de Poisson de parámetro $\lambda_1 + \lambda_2$.

Proposición

Sean s y t tiempos tales que $0 < s < t$, y sean k y n enteros tales que $0 \leq k \leq n$. Entonces

$$P(X_s = k \mid X_t = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}$$

.

Proposición

Sean $X_1(t)$ y $X_2(t)$ dos procesos de Poisson independientes con parámetros λ_1 y λ_2 respectivamente, y sean k y n enteros tales que $0 \leq k \leq n$. Entonces,

$$P(X_1(t) = k \mid X_1(t) + X_2(t) = n) = \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^k \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-k}.$$

Proposición

Dado el evento $\{X(t) = n\}$, el vector de tiempos reales $(W_1, \dots, W_n)$ tiene la misma distribución que el vector de estadísticos de orden $(Y_{(1)}, \dots, Y_{(n)})$ de una muestra aleatoria $Y_1, \dots, Y_n$ de la distribución uniforme en el intervalo $[0, t]$. En particular,

$$f_{W_1, \dots, W_n | X(t)}(w_1, \dots, w_n | n) = \frac{n!}{t^n} \quad \text{si } 0 < w_1 < \dots < w_n < t,$$

y cero en otro caso.

Ejercicio

Los clientes de una tienda entran al establecimiento de acuerdo a un proceso de Poisson $\{X_t : t \geq 0\}$ de parámetro $\lambda = 4$. Calcule e interprete:

- a) $P(X_2 = 1)$.
- b) $P(X_1 = 3, X_2 = 6)$.
- c) $P(X_1 = 0 \mid X_3 = 4)$.
- d) $P(X_2 = 4 \mid X_1 = 2)$.
- e) $P(X_2 \leq 3)$.
- f) $P(X_1 \geq 4 \mid X_1 \geq 2)$.

Ejercicio

Los pasajeros llegan a una parada de autobús de acuerdo a un proceso de Poisson $\{X_t : t \geq 0\}$ de parámetro $\lambda = 2$. Sea τ el instante en el que llega el primer autobús. Suponga que τ es una variable aleatoria con distribución $\text{Unif}(0, 1)$ e independiente del proceso. Calcule:

- a) $E[X_\tau \mid \tau = t]$.
- b) $E[X_\tau]$.
- c) $E[X_\tau^2 \mid \tau = t]$.
- d) $\text{Var}(X_\tau)$.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de parámetro $\lambda = 2$. Sea W_n el instante del n -ésimo evento. Calcule:

- a) $E[X_5]$.
- b) $E[X_5 \mid X_2 = 1]$.
- c) $E[W_2]$.
- d) $E[W_7 \mid X_2 = 3]$.
- e) $E[W_7 \mid W_5 = 4]$.
- f) $E[W_7]$.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de parámetro λ . Sea T una variable aleatoria con distribución Exponencial(θ) e independiente del proceso de Poisson. Encuentre la función de densidad de la variable X_T .

Ejercicio

Suponga que un cierto aparato eléctrico sufre desperfectos a lo largo del tiempo de acuerdo a un proceso de Poisson de parámetro λ . Cada vez que el aparato se descompone es enviado a reparación y luego es puesto en funcionamiento nuevamente. El aparato se reemplaza por uno nuevo cuando el tiempo entre dos descomposturas sucesivas es menor o igual a una constante $a > 0$, incluyendo el tiempo antes de la primera reparación. Encuentre:

1. La función de densidad del tiempo de vida útil del equipo antes de ser reemplazado.
2. El número de reparaciones antes del reemplazo.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de parámetro λ . Demuestre que, condicionado a que el proceso tiene exactamente un salto en el intervalo $[s, t]$, el instante en el que ocurre ese salto se distribuye uniformemente en dicho intervalo.

Ejercicio

Suponga que los mensajes llegan a una cuenta de correo electrónico de acuerdo a un proceso de Poisson de parámetro λ . Cada mensaje es de tipo “basura” con probabilidad α y de tipo “no basura” con probabilidad $1 - \alpha$.

1. Dado que hasta el tiempo t han llegado n mensajes, encuentre la distribución del número de mensajes tipo basura.
2. Suponiendo que se recibe un mensaje basura en algún momento, encuentre la distribución del número total de mensajes recibidos hasta la llegada del siguiente mensaje basura.
3. Encuentre la probabilidad de que al tiempo t hayan llegado más mensajes basura que no basura.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de tasa λ y sean dos tiempos $0 < s < t$. Defina $X_{(s,t]}$ como el número de eventos del proceso que ocurren en el intervalo $(s, t]$. Demuestre que, condicionado a la ocurrencia del evento $\{X_t = n\}$, la variable $X_{(s,t]}$ tiene distribución Binomial($n, 1 - s/t$).

Definiciones alternativas

Definición

Un proceso de Poisson de parámetro $\lambda > 0$ es un proceso a tiempo continuo $\{X_t : t \geq 0\}$, con espacio de estados $\{0, 1, 2, \dots\}$, que cumple:

- a) $X_0 = 0$.
- b) Tiene incrementos independientes y estacionarios.
- c) Para cualquier $t \geq 0$ y cuando $h \rightarrow 0$,
 - I) $P(X_{t+h} - X_t = 1) = \lambda h + o(h)$,
 - II) $P(X_{t+h} - X_t \geq 2) = o(h)$.

Ejemplo. A partir de la definición anterior vamos a demostrar que la variable $X_{t+s} - X_s$ tiene distribución Poisson(λt). Se define $p_n(t) = P(X_t = n)$ y se considera cualquier $h > 0$. Denotaremos por $p_n(t, t+h)$ a la probabilidad $P(X_{t+h} - X_t = n)$. Por la hipótesis de independencia, para $t \geq 0$ y cuando $h \rightarrow 0$,

$$p_0(t+h) = p_0(t) p_0(t, t+h) = p_0(t) (1 - \lambda h + o(h)).$$

Haciendo $h \rightarrow 0$ se obtiene la ecuación diferencial

$$p'_0(t) = -\lambda p_0(t),$$

cuya solución es $p_0(t) = ce^{-\lambda t}$, donde la constante c es uno por la condición inicial $p_0(0) = 1$.

Ahora encontraremos $p_n(t)$ para $n \geq 1$. Nuevamente por independencia,

$$\begin{aligned} p_n(t+h) &= p_n(t) p_0(t, t+h) + p_{n-1}(t) p_1(t, t+h) + o(h) \\ &= p_n(t) (1 - \lambda h + o(h)) + p_{n-1}(t) (\lambda h + o(h)) + o(h). \end{aligned}$$

Haciendo $h \rightarrow 0$ se obtiene

$$p'_n(t) = -\lambda p_n(t) + \lambda p_{n-1}(t),$$

con condición inicial $p_n(0) = 0$ para $n \geq 1$.

Definiendo $q_n(t) = e^{\lambda t} p_n(t)$, la ecuación diferencial se transforma en

$$q'_n(t) = \lambda q_{n-1}(t), \quad q_n(0) = 0, \quad q_0(t) = 1.$$

Esta ecuación se resuelve iterativamente primero para $q_1(t)$, después para $q_2(t)$, y así sucesivamente; en general,

$$q_n(t) = \frac{(\lambda t)^n}{n!}.$$

Por lo tanto,

$$p_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Esto significa que X_t tiene distribución $\text{Poisson}(\lambda t)$. Debido al postulado de incrementos estacionarios, la variable $X_{t+s} - X_s$ tiene la misma distribución que X_t .

Definiciones alternativas

Definición

Un proceso de Poisson de parámetro $\lambda > 0$ es un proceso a tiempo continuo $\{X_t : t \geq 0\}$, con espacio de estados $\{0, 1, 2, \dots\}$, con trayectorias no decrecientes y que satisface:

- a) $X_0 = 0$.
- b) Tiene incrementos independientes.
- c) Para cualesquiera $s \geq 0$ y $t > 0$,

$$X_{t+s} - X_s \sim \text{Poisson}(\lambda t).$$

Ejercicio

Demuestre que las tres definiciones del proceso de Poisson dadas anteriormente son equivalentes.

Proceso de Poisson no homogéneo

Definición

Un proceso de Poisson no homogéneo es un proceso a tiempo continuo $\{X_t : t \geq 0\}$, con espacio de estados $\{0, 1, 2, \dots\}$, cuyo parámetro es una función positiva y localmente integrable $\lambda(t)$, y que cumple:

- a) $X_0 = 0$.
- b) Los incrementos son independientes.
- c) Para cualquier $t \geq 0$ y cuando $h \rightarrow 0$,
 - I) $P(X_{t+h} - X_t = 1) = \lambda(t) h + o(h)$,
 - II) $P(X_{t+h} - X_t \geq 2) = o(h)$.

★ ¿Qué cambia respecto al proceso de Poisson homogéneo?

Proposición

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson no homogéneo con parámetro $\lambda(t)$. Entonces la variable X_t tiene distribución $\text{Poisson}(\Lambda(t))$, donde

$$\Lambda(t) = \int_0^t \lambda(s) ds.$$

En particular, para $n = 0, 1, 2, \dots$,

$$P(X_t = n) = e^{-\Lambda(t)} \frac{\Lambda(t)^n}{n!}.$$

Proposición

Para el proceso de Poisson no homogéneo, la variable incremento $X_{t+s} - X_s$ tiene distribución $\text{Poisson}(\Lambda(t+s) - \Lambda(s))$.

Demostración. Se escribe $X_t = X_s + (X_t - X_s)$, en donde, por el axioma de incrementos independientes, en el lado derecho aparece la suma de dos variables aleatorias independientes. Recordando que la función generadora de momentos de la distribución $\text{Poisson}(\lambda)$ es $M(r) = \exp(\lambda(e^r - 1))$, al aplicar este resultado a la ecuación anterior se obtiene $\exp(\Lambda(t) - \Lambda(s))(e^r - 1) = \exp(\Lambda(s)(e^r - 1)) M_{X_t - X_s}(r)$. Por lo tanto, $M_{X_t - X_s}(r) = \exp((\Lambda(t) - \Lambda(s))(e^r - 1))$. □

- A la función $\Lambda(t)$ se le llama *función de intensidad* y a $\lambda(t)$ se le conoce como *función de valor medio*. A un proceso de Poisson no homogéneo en donde $\{\Lambda(t) : t \geq 0\}$ es un proceso estocástico se le llama proceso de Cox.

Proceso de Poisson compuesto

Definición

Sea $\{N_t : t \geq 0\}$ un proceso de Poisson y sea $Y_1, Y_2, \dots$ una sucesión de variables aleatorias independientes, idénticamente distribuidas e independientes del proceso de Poisson. Sea $Y_0 = 0$. El proceso de Poisson compuesto se define como:

$$X_t = \sum_{n=0}^{N_t} Y_n.$$

Proposición

El proceso de Poisson compuesto cumple las siguientes propiedades:

1. Tiene incrementos independientes y estacionarios.
2. $E(X_t) = \lambda t E(Y)$.
3. $\text{Var}(X_t) = \lambda t E(Y^2)$.
4. $\text{Cov}(X_t, X_s) = \lambda E(Y^2) \min\{t, s\}$.
5. La función generadora de momentos de X_t es

$$M_{X_t}(u) = E\left(e^{uX_t}\right) = \exp\left[\lambda t (M_Y(u) - 1)\right].$$

Ejemplo (Monto de reclamaciones en seguros).

- **Proceso de reclamaciones:**

$$\{N_t : t \geq 0\}, \quad N_t \sim \text{Poisson}(\lambda t),$$

donde $\lambda > 0$ es la intensidad de siniestros.

- **Montos individuales:**

$$Y_1, Y_2, \dots$$

son i.i.d., positivas, e independientes de $\{N_t\}$.

- **Costo total pagado:**

$$X_t = \sum_{n=1}^{N_t} Y_n,$$

que define el **proceso de Poisson compuesto**.

Propiedades útiles en modelos actuariales

- **Valor esperado del costo agregado:**

$$E[X_t] = \lambda t E[Y].$$

- **Varianza del costo agregado:**

$$\text{Var}(X_t) = \lambda t E[Y^2].$$

- **Capital del asegurador (modelo clásico de riesgo):**

$$R_t = u + ct - X_t,$$

donde u es el capital inicial y c la tasa de primas.

Interpretación y estructura del proceso

- ¿Cuántos siniestros han ocurrido? N_t .
- ¿Cuánto cuesta cada siniestro? Y_n .
- ¿Cuál es el costo agregado hasta tiempo t ? $X_t = Y_1 + \cdots + Y_{N_t}$.
- Si $Y_n \in \{1, 2, \dots\}$: proceso de Poisson generalizado. Si $Y_n \equiv 1$: se reduce al proceso de Poisson estándar.

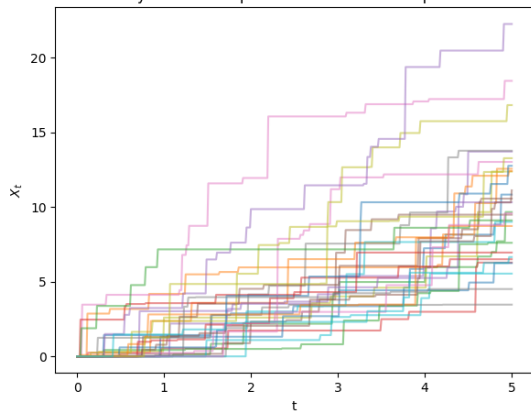
Se generan múltiples trayectorias del proceso de Poisson compuesto $X_t = \sum_{n=1}^{N_t} Y_n$, donde $\{N_t\}$ es un proceso de Poisson homogéneo ($\lambda = 1$) y $\{Y_n\}$ son montos independientes e idénticamente distribuidos (Exponencial($\lambda = 1$)).

- Se fija una tasa de llegadas λ y una distribución para los saltos Y_n .
- Para cada simulación, se generan los tiempos de llegada mediante variables exponenciales y los montos mediante la distribución elegida.
- Se construye la trayectoria acumulando los montos en los tiempos de salto.
- Se repite el procedimiento muchas veces para obtener varias trayectorias.
- Finalmente, se calcula la distribución aproximada de X_T mediante un histograma normalizado.

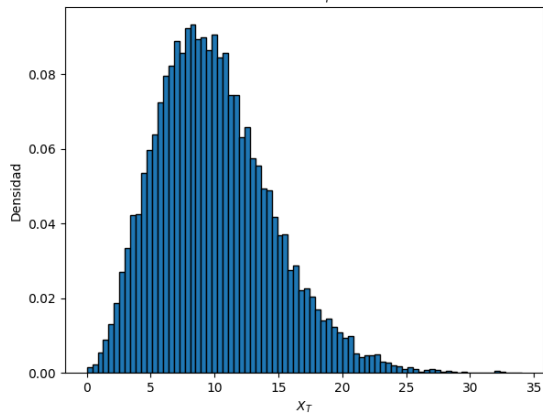
La figura final contiene dos paneles:

1. Múltiples trayectorias del proceso compuesto.
2. Histograma de la distribución empírica de X_T .

Trayectorias del proceso de Poisson compuesto



Distribución de X_T con $T=5.0$



Ejercicio

Considere un proceso de costos agregados por reclamaciones dado por $X_t = \sum_{n=1}^{N_t} Y_n$, donde $\{N_t : t \geq 0\}$ es un proceso de Poisson de intensidad $\lambda > 0$ y $\{Y_n\}_{n \geq 1}$ es una sucesión i.i.d. de montos de siniestros, independientes de N_t . Suponga que los montos individuales tienen distribución exponencial de parámetro $\theta > 0$.

1. Obtenga la función generadora de momentos de X_t y determine su valor esperado y varianza.
2. Halle la probabilidad del evento “*el monto total de siniestros en $[0, t]$ supera una cantidad $a > 0$* ”, es decir, calcule $P(X_t > a)$.
3. Simule este proceso para distintos valores de t y compare empíricamente $E(X_t)$ y $\text{Var}(X_t)$ con los valores teóricos obtenidos.
4. Sea el capital del asegurador $R_t = u + ct - X_t$, donde $u > 0$ es el capital inicial y $c > 0$ es la prima ingresada por unidad de tiempo. Estudie (analíticamente o por simulación) la probabilidad de ruina $P(\{R_t < 0 \text{ para algún } t > 0\})$.

Ejercicio

Suponga que las variables $Y_1, Y_2, \dots$ en un proceso de Poisson compuesto tienen distribución común $\text{Bernoulli}(p)$. Demuestre que el proceso se reduce a un proceso de Poisson de parámetro λp .

Ejercicio

Suponga que los usuarios de cuentas de correo electrónico solicitan acceso al servidor de acuerdo a un proceso de Poisson homogéneo de parámetro λ . Además, cada usuario se mantiene conectado un tiempo aleatorio con función de distribución $F(x)$, independiente entre usuarios. Sea C_t el número de usuarios conectados al servidor al tiempo t , y defina

$$\alpha(t) = \int_0^t [1 - F(x)] dx.$$

Demuestre que, para $k \geq 0$,

$$P(C_t = k) = e^{-\lambda\alpha(t)} \frac{[\lambda\alpha(t)]^k}{k!}.$$

Proceso de Poisson mixto

Definición

Sea Λ una variable aleatoria positiva con función de distribución $F(\lambda)$. Se dice que el proceso de conteo $\{X_t : t \geq 0\}$ es un *proceso de Poisson mixto* con variable mezclante Λ si, para cada entero $n \geq 1$, toda sucesión de enteros no negativos $k_1, \dots, k_n$, y cualesquiera tiempos $0 \leq a_1 \leq b_1 \leq a_2 \leq b_2 \leq \dots \leq a_n \leq b_n$, se cumple

$$\begin{aligned} &P(X_{b_1} - X_{a_1} = k_1, X_{b_2} - X_{a_2} = k_2, \dots, X_{b_n} - X_{a_n} = k_n) \\ &= \int_0^\infty \prod_{i=1}^n \frac{[\lambda(b_i - a_i)]^{k_i}}{k_i!} e^{-\lambda(b_i - a_i)} dF(\lambda). \end{aligned}$$

Cuando la variable aleatoria mezclante Λ es constante e igual a λ , el proceso de Poisson mixto se reduce al proceso de Poisson.

Proposición ★

El proceso de Poisson mixto cumple las siguientes propiedades:

1. Tiene incrementos estacionarios.
2. En general, los incrementos no son independientes; sí lo son en el caso del proceso de Poisson.
3. $E[X_t] = t E[\Lambda]$.
4. $\text{Var}(X_t) = t^2 \text{Var}(\Lambda) + t E[\Lambda]$.
5. $\text{Cov}(X_s, X_t) = \min(s, t) \text{Var}(\Lambda), \quad s, t \geq 0.$

Cadenas de Markov a tiempo continuo

Procesos de saltos

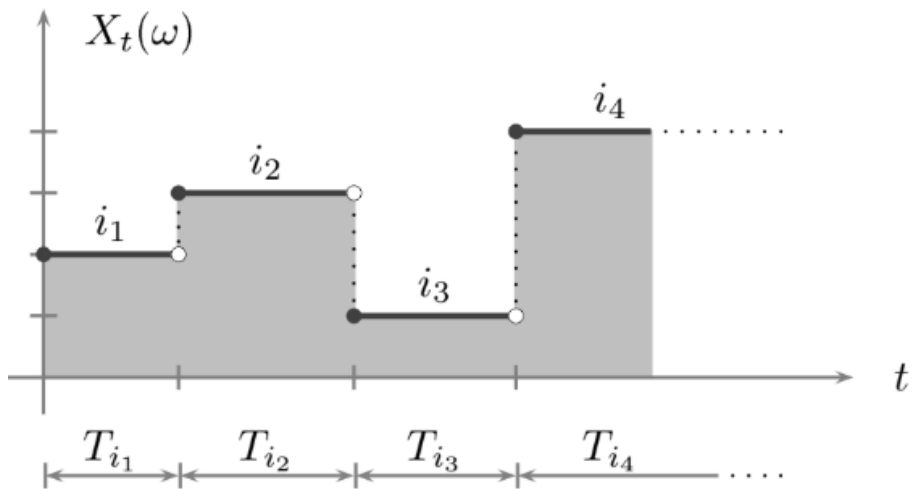
Estudiaremos cadenas de Markov en tiempo continuo donde las variables toman valores enteros. Consideremos un proceso $\{X_t : t \geq 0\}$ que inicia en un estado i_1 , permanece un tiempo aleatorio T_{i_1} , luego salta a un estado $i_2 \neq i_1$, permanece un tiempo T_{i_2} , y así sucesivamente. Los saltos ocurren en los tiempos acumulados

$$W_n = T_{i_1} + T_{i_2} + \cdots + T_{i_n}.$$

El proceso se expresa como

$$X_t = \begin{cases} i_1, & 0 \leq t < W_1, \\ i_2, & W_1 \leq t < W_2, \\ i_3, & W_2 \leq t < W_3, \\ \vdots & \end{cases}$$

- ¿Por qué esto es un proceso de Markov? ¿El proceso está definido para todo t ?



Puede ocurrir que los tiempos de estancia T_{i_k} sean cada vez más pequeños, de modo que

$$\lim_{n \rightarrow \infty} W_n < \infty.$$

En ese caso el proceso realiza infinitos saltos en un intervalo finito y no está bien definido. Para evitar esto suponemos:

$$\lim_{n \rightarrow \infty} W_n = \infty \quad \text{c.s.}$$

Usaremos un espacio de estados $S = \{0, 1, 2, \dots\}$.

Cada estado i tiene un tiempo de estancia T_i con distribución $F_i(t)$. Además:

- T_i es finito con probabilidad 1 $\Rightarrow$ estado no absorbente.
- $T_i = \infty$ con probabilidad 1 $\Rightarrow$ estado absorbente.

Probabilidades de salto

Sea p_{ij} la probabilidad de saltar del estado i al estado j .

Condiciones:

- $p_{ij} \geq 0$,
- $p_{ii} = 0$,
- $\sum_j p_{ij} = 1$.

La matriz de probabilidades de salto es:

$$P = (p_{ij}) = \begin{pmatrix} 0 & p_{01} & p_{02} & \cdots \\ p_{10} & 0 & p_{12} & \cdots \\ p_{20} & p_{21} & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Supondremos:

- Los tiempos de estancia $T_{i_1}, T_{i_2}, \dots$ son independientes.
- Son independientes de la elección del próximo estado.
- Cada estado es absorbente o no absorbente.

Propiedad de Markov

Un proceso de saltos como el descrito es Markoviano si y solo si:

$$T_i \sim \text{Exponencial}(\lambda_i), \quad \lambda_i > 0,$$

para los estados no absorbentes. En estados absorbentes se toma $\lambda_i = 0$ (equivalente a $T_i = \infty$).

Se define para $t > 0$ la probabilidad de transición $p_{ij}(t) = P(X_t = j \mid X_0 = i)$ y la matriz $P(t) = (p_{ij}(t))$. Para $t = 0$: $p_{ij}(0) = \delta_{ij}$.

Definición

Un proceso de saltos con:

- tiempos de estancia exponenciales en los estados no absorbentes,
- matriz de saltos P ,
- ausencia de explosión,

se llama **cadena de Markov en tiempo continuo**.

- Un proceso de Markov a tiempo continuo queda especificado completamente por tres elementos: una distribución inicial sobre el espacio de estados, el conjunto de parámetros no negativos λ_i y las probabilidades de salto p_{ij} .

Ejemplo: Proceso de Poisson

El proceso de Poisson es una cadena de Markov a tiempo continuo que empieza en cero, es decir, la distribución de probabilidad inicial tiene el valor uno en el estado cero, los tiempos de estancia son exponenciales de parámetro λ y las probabilidades de salto de un estado a otro son

$$p_{ij} = \begin{cases} 1 & \text{si } j = i + 1, \\ 0 & \text{si } j \neq i + 1. \end{cases}$$

Ejemplo: Cadena de primera ocurrencia

Sea $\{X_t : t \geq 0\}$ una cadena de Markov a tiempo continuo con espacio de estados $\{0, 1\}$. Suponga que $X_0 = 0$ y que el proceso cambia al estado 1 después de un tiempo aleatorio con distribución Exponencial(λ) y permanece allí el resto del tiempo. Este proceso modela la situación de espera exponencial para la primera ocurrencia de un evento de interés. Las probabilidades de transición son:

$$P(t) = \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \end{pmatrix} = \begin{pmatrix} e^{-\lambda t} & 1 - e^{-\lambda t} \\ 0 & 1 \end{pmatrix}$$

Ejemplo: Cadena de dos estados

Considere el proceso $\{X_t : t \geq 0\}$ con espacio de estados $\{0, 1\}$ y definido por la siguiente dinámica: cuando el proceso entra al estado 0 permanece en él un tiempo Exponencial(λ) y luego de esto va al estado 1 y permanece un tiempo Exponencial(μ). Se postula además que los tiempos de estancia para cada estado son variables aleatorias independientes.

- ★ Haga una representación de este proceso en forma de grafo.
- ★ Dibuje una posible trayectoria de esta cadena.
- ★ ¿Cuál es la distribución inicial sobre el espacio de estados, el conjunto de parámetros no negativos λ_i y las probabilidades de salto de esta cadena?

Probabilidades de transición

Proposición

Sean i y j dos estados. Para cualquier $t \geq 0$:

$$p_{ij}(t) = \delta_{ij}e^{-\lambda_i t} + \lambda_i e^{-\lambda_i t} \int_0^t e^{\lambda_i s} \left(\sum_{k, k \neq i} p_{ik} p_{kj}(s) \right) ds.$$

★ ¿Cuál es la interpretación de este resultado?

Ecuación de Chapman-Kolmogorov

Proposición

Para cualquier par de estados i y j , y para todo $t \geq 0$ y $s \geq 0$,

$$p_{ij}(t+s) = \sum_k p_{ik}(t)p_{kj}(s).$$

En notación matricial, $P_{t+s} = P_t P_s$.

Generador infinitesimal

Proposición

Para cualquier par de estados i y j , y para cualquier $t > 0$,

$$\frac{d}{dt}p_{ij}(t) = p'_{ij}(t) = -\lambda_i p_{ij}(t) + \lambda_i \sum_{k, k \neq i} p_{ik} p_{kj}(t).$$

Cuando $t \rightarrow 0^+$ se tiene que:

$$\begin{aligned} p'_{ij}(0) &= -\lambda_i \delta_{ij} + \lambda_i \sum_{k, k \neq i} p_{ik} \delta_{kj} \\ &= -\lambda_i \delta_{ij} + \lambda_i p_{ij} \\ &= \begin{cases} -\lambda_i, & \text{si } i = j \\ \lambda_i p_{ij}, & \text{si } i \neq j. \end{cases} \end{aligned}$$

Generador infinitesimal

Definición

A las cantidades $p'_{ij}(0)$ se les denotará por g_{ij} y se les conoce con el nombre de parámetros infinitesimales del proceso. Es decir,

$$g_{ij} = \begin{cases} -\lambda_i, & \text{si } i = j \\ \lambda_i p_{ij}, & \text{si } i \neq j. \end{cases}$$

La matriz $G := (g_{ij})$ es llamada el generador infinitesimal del proceso de Markov.

- ★ Se cumple que $g_{ij} \geq 0$ si $i \neq j$, $g_{ii} \leq 0$ y $\sum_j g_{ij} = 0$.
- ★ ¿Cómo es g_{ii} si el estado i es absorbente?
- ★ El generador infinitesimal y una distribución inicial del espacio de estados determinan de manera única a la cadena de Markov a tiempo continuo.

Generador infinitesimal

Ejemplo. El generador infinitesimal para el proceso de Poisson de parámetro λ es

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Generador infinitesimal

Ejemplo. El generador infinitesimal para la cadena de Markov de dos estados es

$$G = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}$$

Ejercicio

Para la cadena de Markov de dos estados, demuestre que para cualquier $t \geq 0$,

$$p_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t}.$$

Ejercicio

Para la cadena de Markov de dos estados, demuestre que el tiempo de estancia promedio en el estado cero, a largo plazo, es

$$\lim_{t \rightarrow \infty} E\left(\frac{1}{t} \int_0^t \mathbf{1}_{\{0\}}(X_s) ds\right) = \frac{\mu}{\lambda + \mu}.$$

Observe que no es necesario especificar el estado inicial. La fracción complementaria corresponde al estado uno.

Ecuaciones de Kolmogorov

Ecuaciones retrospectivas. En términos de los parámetros infinitesimales, el sistema de ecuaciones diferenciales puede escribirse como

$$p'_{ij}(t) = \sum_k g_{ik} p_{kj}(t).$$

En forma matricial:

$$P'(t) = G P(t).$$

Explícitamente, para una cadena genérica:

$$\begin{pmatrix} p'_{00}(t) & p'_{01}(t) \\ p'_{10}(t) & p'_{11}(t) \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} -\lambda_0 & \lambda_0 p_{01} \\ \lambda_1 p_{10} & -\lambda_1 \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} p_{00}(t) & p_{01}(t) \\ p_{10}(t) & p_{11}(t) \\ \vdots & \vdots \end{pmatrix}.$$

A este sistema se le conoce como las ecuaciones retrospectivas de Kolmogorov.

Ecuaciones de Kolmogorov

Ejemplo. Para un proceso de Poisson de parámetro λ , las ecuaciones retrospectivas son:

$$p'_{ii}(t) = -\lambda p_{ii}(t), \quad p'_{ij}(t) = -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t) \quad (i < j).$$

La solución conocida es

$$p_{ij}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}, \quad i \leq j,$$

y $p_{ij}(t) = 0$ si $i > j$.

Ecuaciones de Kolmogorov

Ejemplo. Para la cadena de dos estados con tasas λ y μ :

$$p'_{00}(t) = -\lambda p_{00}(t) + \lambda p_{10}(t),$$

$$p'_{01}(t) = -\lambda p_{01}(t) + \lambda p_{11}(t),$$

$$p'_{10}(t) = -\mu p_{10}(t) + \mu p_{00}(t),$$

$$p'_{11}(t) = -\mu p_{11}(t) + \mu p_{01}(t).$$

Los pares de ecuaciones son independientes. Definiendo

$$q_{00}(t) = e^{\lambda t} p_{00}(t), \quad q_{10}(t) = e^{\mu t} p_{10}(t),$$

el sistema se reduce a uno de segundo orden cuya solución final es

$$p_{00}(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda + \mu)t}.$$

★ ¿Cómo obtener las probabilidades que restan por calcular?

Ecuaciones de Kolmogorov

Ecuaciones prospectivas. El sistema de ecuaciones diferenciales

$$P'(t) = P(t) G$$

recibe el nombre de ecuaciones prospectivas de Kolmogorov.

Explícitamente,

$$p'_{ij}(t) = \sum_k p_{ik}(t) g_{kj}.$$

En general este sistema no siempre se satisface para una cadena de Markov a tiempo continuo, a diferencia del sistema retrospectivo. Sin embargo, para procesos de nacimiento y muerte sí se cumple bajo hipótesis adicionales.

Ecuaciones de Kolmogorov

Ejemplo. Para un proceso de Poisson de parámetro λ , las ecuaciones prospectivas son:

$$p'_{ii}(t) = -\lambda p_{ii}(t), \quad p'_{ij}(t) = -\lambda p_{ij}(t) + \lambda p_{i,j-1}(t), \quad i < j.$$

La solución es

$$p_{ij}(t) = e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}, \quad i \leq j,$$

y $p_{ij}(t) = 0$ si $i > j$.

Ecuaciones de Kolmogorov

Ejemplo. Para la cadena con tasas de salto λ y μ :

$$p'_{00}(t) = -\lambda p_{00}(t) + \mu p_{01}(t),$$

$$p'_{01}(t) = -\mu p_{01}(t) + \lambda p_{00}(t),$$

$$p'_{10}(t) = -\lambda p_{10}(t) + \mu p_{11}(t),$$

$$p'_{11}(t) = -\mu p_{11}(t) + \lambda p_{10}(t).$$

- ★ Puede verificarse que la solución coincide exactamente con la obtenida en el ejemplo de las ecuaciones retrospectivas para esta cadena de dos estados.

Ejercicio

Escriba el sistema de ecuaciones diferenciales retrospectivas de Kolmogorov para un proceso de Poisson de parámetro λ y compruebe que la distribución Poisson satisface este sistema de ecuaciones.

Ejercicio

Sea $\{N_t : t \geq 0\}$ un proceso de Poisson de parámetro λ y sean $Y_1, Y_2, \dots$ v.a.i.i.d. con distribución Bernoulli(p). Considere el proceso de Poisson compuesto

$$X_t = \sum_{j=1}^{N_t} Y_j.$$

Encuentre las ecuaciones de Kolmogorov retrospectivas y prospectivas del proceso X_t . Resuelva cualquiera de estos sistemas y demuestre que, para $j \geq i$,

$$p_{ij}(t) = e^{-\lambda pt} \frac{(\lambda pt)^{j-i}}{(j-i)!}.$$

Procesos de nacimiento y muerte

Definición

Un proceso de nacimiento y muerte es una cadena de Markov a tiempo continuo con espacio de estados $\{0, 1, 2, \dots\}$ y matriz generadora infinitesimal

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1 + \mu_1) & \lambda_1 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2 + \mu_2) & \lambda_2 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

donde λ_i son tasas de nacimiento y μ_i son tasas de muerte, con $\mu_0 = 0$.

Comportamiento del proceso.

Del generador se obtiene que el tiempo de estancia en el estado $i \geq 0$ es exponencial con parámetro

$$\lambda_i + \mu_i.$$

Al finalizar ese tiempo, el proceso salta a:

$$p_{i,i+1} = \frac{\lambda_i}{\lambda_i + \mu_i}, \quad p_{i,i-1} = \frac{\mu_i}{\lambda_i + \mu_i}.$$

Interpretación:

- salto hacia arriba $\Rightarrow$ nacimiento,
- salto hacia abajo $\Rightarrow$ muerte,
- no hay saltos de tamaño mayor a 1.

Interpretación como tamaño de población.

Si X_t denota el número de individuos en una población en el tiempo t :

- cada nacimiento ocurre con tasa λ_i cuando la población está en i ,
- cada muerte ocurre con tasa μ_i ,
- los individuos nacen y mueren de uno en uno,
- como $\mu_0 = 0$, la población nunca baja de cero.

Una trayectoria típica alterna entre tiempos de espera exponenciales y saltos ± 1 .

Definición mediante postulados.

Una caracterización equivalente del proceso se obtiene con:

1. Incrementos independientes y estacionarios.
2. Probabilidades de transición estacionarias:

$$p_{ij}(t) = P(X_{s+t} = j \mid X_s = i).$$

3. Para $h \rightarrow 0$:

$$p_{i,i+1}(h) = \lambda_i h + o(h), \quad i \geq 0,$$

$$p_{i,i-1}(h) = \mu_i h + o(h), \quad i \geq 1,$$

$$p_{ii}(h) = 1 - (\lambda_i + \mu_i)h + o(h).$$

Estos postulados llevan nuevamente a los tiempos de estancia exponenciales y a la estructura de saltos de la matriz generadora.

Ejemplo: Proceso de Poisson.

Es un caso particular donde:

$$\mu_i = 0, \quad \lambda_i = \lambda > 0 \text{ para todo } i.$$

Entonces:

- sólo hay nacimientos,
- cada nacimiento ocurre con tasa λ ,
- X_t aumenta de uno en uno.

La matriz generadora se simplifica a

$$G = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots \\ 0 & -\lambda & \lambda & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{pmatrix}.$$

Proceso de nacimiento puro.

Es un caso particular donde todas las tasas de muerte son cero:

$$\mu_0 = \mu_1 = \mu_2 = \dots = 0.$$

El generador toma la forma:

$$G = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \dots \\ 0 & -\lambda_1 & \lambda_1 & 0 & \dots \\ 0 & 0 & -\lambda_2 & \lambda_2 & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Propiedades:

- En el estado i , el tiempo de estancia es $\text{Exp}(\lambda_i)$.
- El único salto posible es $i \rightarrow i + 1$ con probabilidad 1.
- Incrementos independientes, pero no necesariamente estacionarios.

Probabilidades infinitesimales:

$$P(X_{t+h} = i + 1 \mid X_t = i) = \lambda_i h + o(h), \quad P(X_{t+h} = i \mid X_t = i) = 1 - \lambda_i h + o(h).$$

Proceso de muerte puro.

Es el proceso simétrico al de nacimiento puro, ahora con todas las tasas de nacimiento iguales a cero:

$$\lambda_0 = \lambda_1 = \dots = 0.$$

El proceso inicia típicamente en un estado $k \geq 1$.

- Desde el estado $i \geq 1$, el tiempo de estancia es Exponencial(μ_i).
- El único salto posible es $i \rightarrow i - 1$.
- El proceso puede llegar eventualmente a la extinción: 0.

Interpretación:

- La población solo decrece.
- Los tiempos entre muertes son exponenciales.
- Es un caso especial de cadena de nacimiento y muerte con absorción en 0.

Proceso de Yule.

Es un proceso de nacimiento puro con tasas lineales: $\lambda_n = \lambda n$, $n \geq 1$.

Interpretación:

- Cada individuo da origen a nuevos individuos a tasa λ .
- Si $X_t = n$, hay n "posibles progenitores", por lo que la tasa total es λn .
- El tiempo de estancia en n es Exponencial(λn).

Sistema prospectivo:

$$\begin{aligned} p'_{kk}(t) &= -\lambda k p_{kk}(t), \\ p'_{kn}(t) &= \lambda(n-1) p_{k,n-1}(t) - \lambda n p_{kn}(t), \quad n \geq k+1. \end{aligned}$$

La ecuación recursiva para $n \geq k+1$ es:

$$p_{kn}(t) = \lambda(n-1) e^{-\lambda n t} \int_0^t e^{\lambda n s} p_{k,n-1}(s) ds.$$

Resultado fundamental: $X_t - X_0 \sim \text{BinomialNegativa}(r = k, p = e^{-\lambda t})$, es decir,

$$P(X_t = n) = \binom{n-1}{k-1} (1 - e^{-\lambda t})^{n-k} (e^{-\lambda t})^k, \quad n \geq k.$$

Ejercicio

Considere una cadena de Markov a tiempo continuo de dos estados en donde la estancia en el estado 0 tiene distribución Exponencial(λ) y en el estado 1 distribución Exponencial(μ). Demuestre que:

$$\text{a) } p_{00}(t) = \frac{\mu + \lambda e^{-(\lambda+\mu)t}}{\lambda + \mu},$$

$$\text{b) } p_{01}(t) = \frac{\lambda - \lambda e^{-(\lambda+\mu)t}}{\lambda + \mu},$$

$$\text{c) } p_{10}(t) = \frac{\mu - \mu e^{-(\lambda+\mu)t}}{\lambda + \mu},$$

$$\text{d) } p_{11}(t) = \frac{\lambda + \mu e^{-(\lambda+\mu)t}}{\lambda + \mu}.$$

Ejercicio

Considere una cadena de Markov a tiempo continuo de dos estados en donde el tiempo de estancia en cada uno tiene distribución Exponencial(λ). Defina N_t como el número de saltos efectuados por la cadena hasta el tiempo t . Demuestre que $\{N_t : t \geq 0\}$ es un proceso de Poisson de parámetro λ .

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Yule con estado inicial $X_0 = k \geq 1$. Demuestre que:

- a) $E[X_t] = ke^{\lambda t}$,
- b) $\text{Var}(X_t) = ke^{2\lambda t} (1 - e^{-\lambda t})$.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de muerte puro con $X_0 = N \geq 1$ y tasas $\mu_1, \mu_2, \dots, \mu_N$. Demuestre que el área promedio bajo la trayectoria del proceso, hasta alcanzar el estado absorbente 0, es

$$\sum_{k=1}^N \frac{k}{\mu_k}.$$

Comunicación entre estados.

Para estados i y j :

$$i \rightarrow j \quad \text{si existe } t \geq 0 \text{ tal que } p_{ij}(t) > 0.$$

$$i \leftrightarrow j \quad \text{si } i \rightarrow j \text{ y } j \rightarrow i.$$

La comunicación es una relación de equivalencia y divide el espacio de estados en clases de comunicación. La cadena es irreducible si todos los estados pertenecen a la misma clase.

Tiempos de primera visita.

$$\tau_{ij} = \inf\{t \geq 0 : X_t = j \mid X_0 = i\}, \quad \mu_{ij} = E[\tau_{ij}].$$

Cuando $i = j$ se escribe τ_i y su esperanza μ_i se conoce como tiempo medio de recurrencia.

Recurrencia y transitoriedad.

El estado i es:

- *transitorio* si, con probabilidad 1, el conjunto $\{t \geq 0 : X_t = i\}$ es acotado,
- *recurrente* si, con probabilidad 1, este conjunto es no acotado,
- *recurrente positivo* si $E[\tau_i] < \infty$,
- *recurrente nulo* si $E[\tau_i] = \infty$.

Distribuciones estacionarias.

Una distribución π es estacionaria si

$$\pi P(t) = \pi, \quad \forall t \geq 0.$$

Un criterio útil:

$$\pi \text{ es estacionaria} \iff \pi G = 0.$$

Ejemplo (cadena de dos estados):

$$G = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix} \Rightarrow \pi = \left(\frac{\mu}{\lambda + \mu}, \frac{\lambda}{\lambda + \mu} \right).$$

Cadena a tiempo discreto asociada.

Dada una cadena a tiempo continuo $\{X_t\}_{t \geq 0}$, se asocia una cadena a tiempo discreto $\{X_n\}_{n \geq 0}$ observando el proceso únicamente en los instantes en los que ocurren los saltos.

Propiedades:

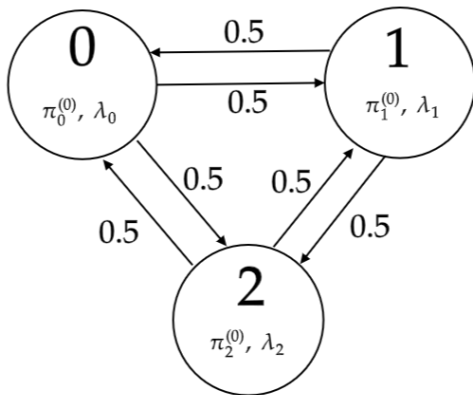
- conserva la estructura de comunicación,
- preserva recurrencia y transitoriedad,
- permite analizar el proceso continuo mediante técnicas de tiempo discreto.

Construcción inversa:

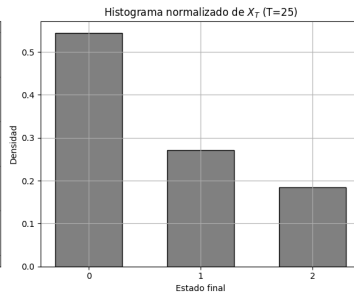
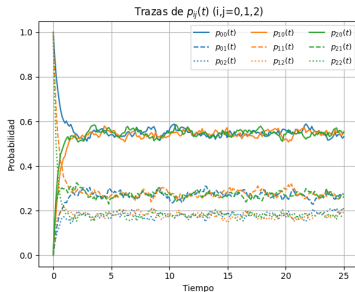
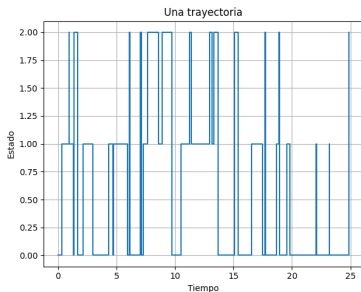
- dada una cadena a tiempo discreto con matriz Q ,
- se le asocian tiempos exponenciales independientes entre saltos,
- obteniendo así una cadena de Markov a tiempo continuo con generador G .

Ejercicio

Considere la cadena de Markov a tiempo continuo $\{X_t : t \geq 0\}$ con espacio de estados $\{0, 1, 2\}$, distribución inicial $\pi^{(0)} = (1, 0, 0)$, parámetros $\lambda_i = i + 1$ para $i = 0, 1, 2$ y probabilidades de saltos como lo muestra el grafo a continuación.



1. Halle el generador G , la distribución estacionaria de la cadena y las probabilidades de transición $p_{ij}(t)$ usando las ecuaciones retrospectivas.
2. Interprete cada uno de los siguientes gráficos. ¿Cómo se realiza una simulación de una cadena de Markov a tiempo continuo?



Procesos de renovación y confiabilidad

Suponga que se pone en operación un cierto componente o artículo cuya duración o vida útil se modela mediante una variable aleatoria T_1 . Una vez que el componente falla, se reemplaza o se *renueva* con otro componente cuyo tiempo de vida es T_2 , y así sucesivamente. La colección de variables aleatorias $T_1, T_2, \dots$ representa la sucesión de tiempos de vida de componentes puestos en operación uno tras otro.

Definición

Un proceso de renovación es una sucesión infinita de variables aleatorias $T_1, T_2, \dots$ que son no negativas, independientes e idénticamente distribuidas.

Definición

Dado un proceso de renovación $\{T_1, T_2, \dots\}$, se definen los tiempos reales de renovación $W_0 = 0$ y $W_n = T_1 + \dots + T_n$, para $n \geq 1$. El proceso de conteo de renovaciones es

$$N_t = \max\{n \geq 0 : W_n \leq t\}$$

para cada $t \geq 0$.

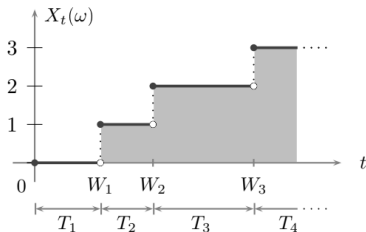


Figura: El tiempo de Poisson y los tiempos de ocurrencia del evento. Tomado de Rincon, 2012.

- La variable aleatoria W_n representa el tiempo real en el que se realiza la n -ésima renovación, mientras que la variable N_t indica el número de renovaciones efectuadas hasta el tiempo t .
- Sea $F(t)$ la función de distribución de los tiempos de vida. Por lo tanto, la función de distribución de W_n es la convolución de $F(t)$ consigo misma n veces, es decir,

$$F_{W_n}(t) = P(T_1 + \dots + T_n \leq t) = F^{*n}(t) = (F * \dots * F)(t).$$

Como casos particulares $F^{*1}(t) = F(t)$ y $F^{*0}(t) = \begin{cases} 0 & \text{si } t \leq 0, \\ 1 & \text{si } t > 0. \end{cases}$

- ★ ¿Es posible conocer la distribución de la variable aleatoria N_t ?
- ★ ¿Hay una dependencia de la distribución de N_t con la distribución de los tiempos de vida?

Proposición

Para cualquier $n \geq 0$

$$P(N_t = n) = F^{*n}(t) - F^{*(n+1)}(t)$$

- En muy pocos casos es posible encontrar la distribución explícita del número de renovaciones N_t .

Ejemplo. Un proceso de Poisson homogéneo de tasa λ es un proceso de renovación en donde los tiempos de vida tienen distribución Exponencial(λ). En consecuencia, los tiempos reales de renovación W_n tienen distribución Gamma, cuya función de distribución es, para $t \geq 0$,

$$F_{W_n}(t) = \int_0^t \frac{(\lambda x)^{n-1}}{\Gamma(n)} \lambda e^{-\lambda x} dx = \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

A partir de esta expresión se puede recuperar la distribución Poisson: para todo $n \geq 0$

$$P(N_t = n) = F^{*n}(t) - F^{*(n+1)}(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}.$$

Función y ecuación de renovación

Definición

La función de renovación $\Lambda(t)$ se define como el número promedio de renovaciones efectuadas hasta un tiempo t , es decir, $\Lambda(t) := E(N_t)$.

Proposición

La función de renovación $\Lambda(t)$ satisface

$$\Lambda(t) = F(t) + \int_0^t \Lambda(t-s) dF(s) = F(t) + (\Lambda * F)(t).$$

- ★ En un proceso de Poisson homogéneo, demuestre que $\Lambda(t) = \lambda t$ cumple la anterior ecuación integral (ecuación de renovación) con $F(t) = 1 - e^{-\lambda t}$ para $t \geq 0$.
- ★ ¿Cómo encontrar $\Lambda(t)$?

Proposición

Para cualquier $t > 0$,

$$\Lambda(t) = \sum_{n=1}^{\infty} F^{*n}(t).$$

- ★ Demuestre el anterior resultado por definición de $\Lambda(t)$.
- ★ Pruebe este resultado en el caso del proceso de Poisson.
 - La serie $\sum_{n=1}^{\infty} F^{*n}(t)$ es convergente. Además, es posible demostrar que el proceso de conteo de renovaciones tiene momentos ordinarios finitos de cualquier orden.
 - El proceso de conteo de renovaciones $\{N_t : t \geq 0\}$ está determinado por la función de renovación $\Lambda(t)$

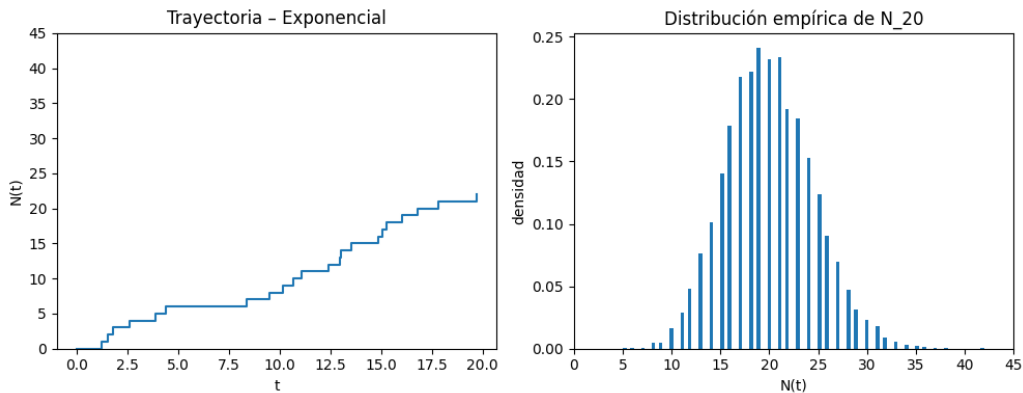


Figura: A la derecha se observa una realización de un proceso de renovación con distribución Exponencial($\lambda = 1$). A la izquierda se observa la distribución de $N_{t=20}$ usando un histograma normalizado con 10000 simulaciones.

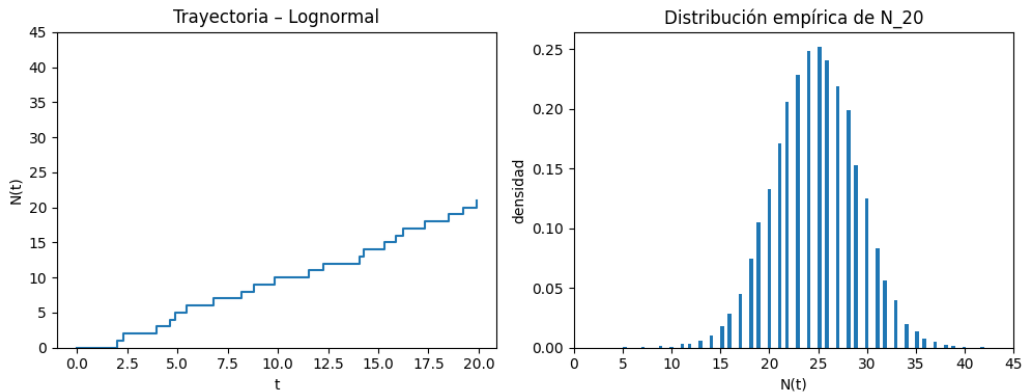


Figura: A la derecha se observa una realización de un proceso de renovación con distribución $\text{LogNormal}(\mu = -0,5, \sigma^2 = 0,75)$. A la izquierda se observa la distribución de $N_{t=20}$ usando un histograma normalizado con 10000 simulaciones.

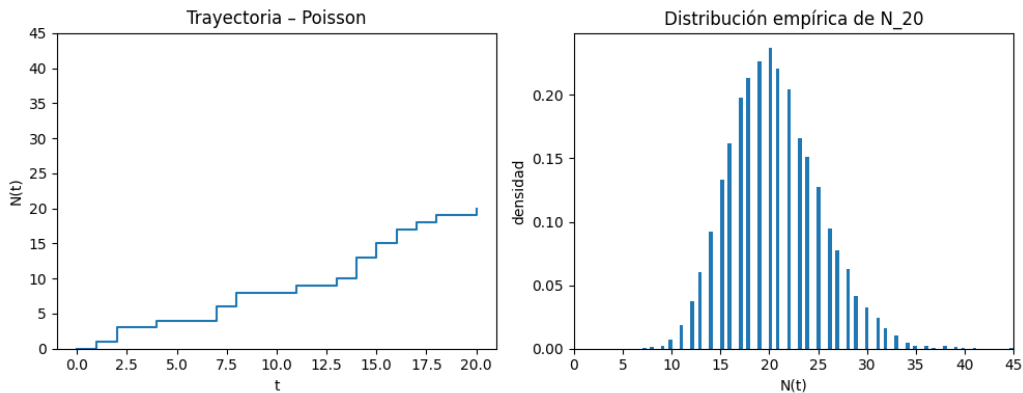


Figura: A la derecha se observa una realización de un proceso de renovación con distribución Poisson($\lambda = 1$). A la izquierda se observa la distribución de $N_{t=20}$ usando un histograma normalizado con 10000 simulaciones.

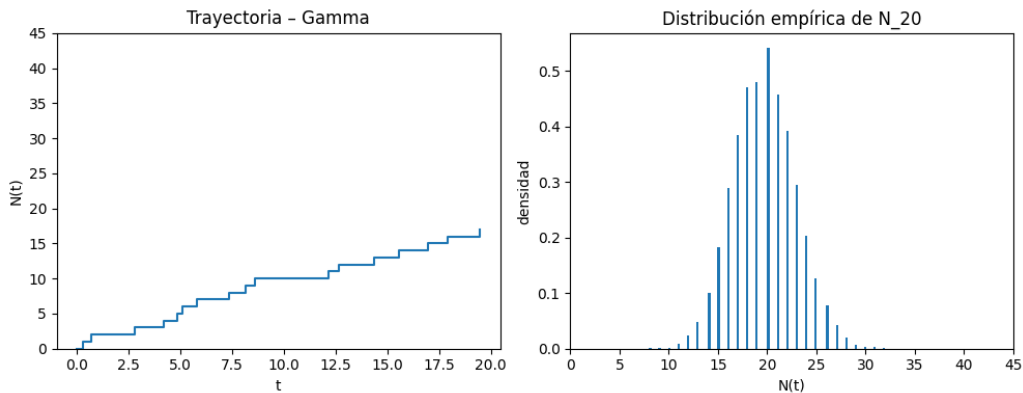


Figura: A la derecha se observa una realización de un proceso de renovación con distribución $\text{Gamma}(a = 2, b = 0,5)$. A la izquierda se observa la distribución de $N_{t=20}$ usando un histograma normalizado con 10000 simulaciones.

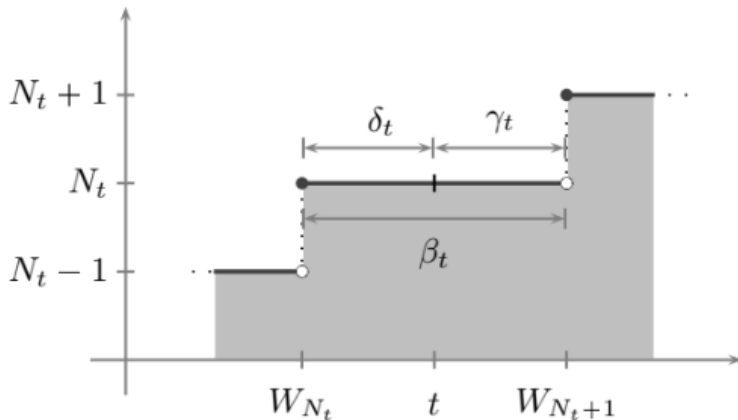


Figura: Representaciones de tiempos de vida de un proceso de renovación. Tomado de Rincon, 2012.

Definición

A continuación se definen tres tiempos de vida:

- **Tiempo de vida restante.** Es el tiempo de vida útil que le resta al elemento que se encuentra en operación al tiempo t , y está dado por la variable $\gamma_t = W_{N_t+1} - t$.
- **Tiempo de vida transcurrido.** Es el tiempo que ha estado en uso el elemento que se encuentra en operación al tiempo t , y está dado por la variable $\delta_t = t - W_{N_t}$.
- **Tiempo de vida total.** Es el tiempo de vida útil que tiene el elemento que se encuentra en operación al tiempo t , y está dado por la variable

$$\beta_t = \delta_t + \gamma_t = W_{N_t+1} - W_{N_t} = T_{N_t+1}.$$

- ★ Halle la distribución de cada uno de los tiempos de vida definidos anteriormente para el caso de un proceso de Poisson homogéneo de parámetro λ .

Teoremas de renovación

Proposición

Para todo proceso de renovación $\lim_{n \rightarrow \infty} N_t \stackrel{\text{c.s.}}{=} \infty$.

Proposición

Para un proceso de renovación $\{N_t : t \geq 0\}$ en donde $E(T) = \mu$ con $0 < \mu < \infty$ se cumple que

$$\lim_{t \rightarrow \infty} \frac{W_{N_t}}{N_t} \stackrel{\text{c.s.}}{=} \mu.$$

★ ¿Cuál es la interpretación de estos resultados?

Teoremas de renovación

Proposición (Teorema elemental de renovación)

Para un proceso de renovación en donde $E(T) = \mu$ con $0 < \mu < \infty$ se tiene que

$$\lim_{t \rightarrow \infty} \frac{N_t}{t} \stackrel{\text{c.s.}}{=} \frac{1}{\mu}.$$

Proposición

Para un proceso de renovación en donde $E(T) = \mu$ con $0 < \mu < \infty$ se tiene que

$$\lim_{t \rightarrow \infty} \frac{\Lambda(t)}{t} = \frac{1}{\mu}.$$

- ★ ¿Cuál es la interpretación de estos resultados?
- ★ Haga una simulación computacional para validar estos resultados.

Confiabilidad

Suponga que $T > 0$ es una variable aleatoria continua que registra el tiempo de falla de un cierto componente que se encuentra en operación al tiempo cero. En la teoría de la confiabilidad interesa conocer la probabilidad de que el componente funcione correctamente por lo menos hasta un tiempo t cualquiera o la probabilidad de que deje de funcionar en el siguiente instante de tiempo, dado que ha funcionado bien hasta un cierto momento.

Definición

La función de confiabilidad se define como

$$R(t) := P(T > t) = 1 - F(t),$$

mientras que la función de tasa de falla se define como

$$r(t) = \frac{f(t)}{1 - F(t)}.$$

La **tasa de falla** o *hazard* se interpreta como la probabilidad de que, dado que el componente ha sobrevivido hasta el tiempo t , falle en el intervalo $(t, t + \Delta t)$:

$$P(t < T \leq t + \Delta t \mid T > t) = r(t) \Delta t + o(\Delta t).$$

En efecto,

$$P(t < T \leq t + \Delta t \mid T > t) = \frac{F(t + \Delta t) - F(t)}{1 - F(t)} = \frac{f(t)}{R(t)} \Delta t + o(\Delta t).$$

Relación entre $R(t)$ y $r(t)$

La tasa de falla y la función de confiabilidad están relacionadas por

$$r(t) = \frac{f(t)}{R(t)} = -\frac{d}{dt} \ln R(t), \quad R(t) = \exp\left(-\int_0^t r(s) ds\right).$$

Propiedades importantes:

- $R(t)$ es siempre decreciente.
- $r(t)$ no tiene por qué ser monótona.

Además, como $T \geq 0$, el tiempo medio de falla puede escribirse como

$$E[T] = \int_0^\infty R(t) dt.$$

Ejemplo. Suponga que el tiempo de falla T tiene distribución exponencial con parámetro λ . Entonces:

$$R(t) = e^{-\lambda t}, \quad r(t) = \lambda.$$

La tasa de falla constante implica que la probabilidad de fallar en un intervalo pequeño $(t, t + \Delta t)$, dado que el componente sigue funcionando en t , es aproximadamente $\lambda \Delta t$, independiente de t .

Esto refleja la propiedad de pérdida de memoria de la distribución exponencial, que es la única distribución absolutamente continua con tasa de falla constante. En este caso, $E[T] = \frac{1}{\lambda}$.

★ Sea $T \sim \text{Exponencial}(\lambda)$. Calcule explícitamente $P(t < T \leq t + \Delta t \mid T > t)$ y muestre que

$$P(t < T \leq t + \Delta t \mid T > t) = \lambda \Delta t + o(\Delta t).$$

Martingalas

Motivación

Existen varias acepciones para el término *martingala*. Una de ellas se refiere a un proceso estocástico en tiempo discreto que satisface la identidad

$$E[X_{n+1} \mid X_0 = x_0, \dots, X_n = x_n] = x_n.$$

- Este tipo de procesos modela la evolución del capital de un jugador que realiza una sucesión de apuestas justas: en promedio, no hay ganancia ni pérdida. Parte de la motivación para el estudio de este tipo de procesos fue buscar demostrar la inexistencia de estrategias ventajosas en este juego de apuestas.

- Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Una σ -álgebra representa la información disponible sobre un experimento aleatorio.
- Si $\mathcal{F}_1 \subset \mathcal{F}_2$, entonces $\mathcal{F}_2$ contiene más información que $\mathcal{F}_1$, en el sentido de que, en general, un mayor número de conjuntos son considerados como eventos.
- Para un proceso estocástico a tiempo discreto $\{X_n : n \geq 1\}$ se define

$$\mathcal{F}_n = \sigma(X_1, \dots, X_n).$$

- La colección $\{\mathcal{F}_n\}_{n \geq 1}$ es una sucesión no decreciente de σ -álgebras y representa la historia del proceso hasta el tiempo n .

Ejemplo. Si X_n representa los resultados sucesivos de lanzar un dado y

$$A = \{\text{la suma de los dos primeros resultados es mayor que } 3\},$$

entonces:

- $A \notin \mathcal{F}_1$,
- $A \in \mathcal{F}_2$.

El evento solo puede determinarse con la información disponible hasta el segundo lanzamiento.

Definición

Una filtración es una colección de σ -álgebras $\{\mathcal{F}_n\}_{n \geq 1}$ tal que

$$\mathcal{F}_n \subset \mathcal{F}_m \quad \text{si } n \leq m.$$

- La filtración natural de un proceso $\{X_n\}$ es $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$.
- A tiempo continuo se define una filtración como una colección no numerable de σ -álgebras $\{\mathcal{F}_t\}_{t \geq 0}$ tal que $\mathcal{F}_s \subseteq \mathcal{F}_t$ cuando $0 \leq s \leq t$. La filtración natural o canónica a tiempo continuo $\{X_t : t \geq 0\}$ es la colección de σ -álgebras dadas por $\mathcal{F}_t = \sigma\{X_s : 0 \leq s \leq t\}$, esto es, $\mathcal{F}_t$ es la mínima σ -álgebra que hace que cada una de las variables X_s para valores de s es el intervalo $[0, t]$ sean medibles. A la σ -álgebra $\mathcal{F}_t$ se le denomina la historia del proceso al tiempo t .

Definición

Un proceso $\{X_n : n \geq 1\}$ es *adaptado* a la filtración $\{\mathcal{F}_n\}$ si X_n es $\mathcal{F}_n$ -medible para todo n .

- ★ Todo proceso es adaptado a su filtración natural, la cual es la más pequeña respecto de la cual el proceso es adaptado.

Definición

Un proceso $\{X_n : n \geq 1\}$ es *predecible* respecto de la filtración $\{\mathcal{F}_n\}_{n \geq 0}$ si para cada $n \geq 1$, la variable X_n es $\mathcal{F}_{n-1}$ -medible.

★ Todo proceso predecible es adaptado. La filtración debe iniciar en $n = 0$.

Tiempos de paro

Definición

Una variable aleatoria τ con valores en $\{1, 2, \dots\} \cup \{\infty\}$ es un tiempo de paro respecto a la filtración $\{\mathcal{F}_n\}_{n \geq 1}$ si para cada $n \geq 1$ se cumple que el evento $(\tau \leq n)$ pertenece a $\mathcal{F}_n$.

- Interpretación: la pregunta de si el evento de interés ha ocurrido al tiempo n o antes, debe poder ser respondida con la información dada por la filtración al tiempo n .
- ★ La condición que aparece en la definición es equivalente a la condición $(\tau = n) \in \mathcal{F}_n$, para cada $n \geq 1$.

Ejemplo (Tiempo de primer arribo). Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias adaptada a la filtración $\{\mathcal{F}_n\}_{n \geq 1}$. Sea A un conjunto de Borel de $\mathbb{R}$ y defina

$$\tau = \min\{n \geq 1 : X_n \in A\},$$

en donde conviene definir $\min \emptyset = \infty$.

Es decir, τ es el primer momento en el que la sucesión toma un valor dentro del conjunto A , si acaso ello sucede. La variable aleatoria τ es un tiempo de paro, pues para cualquier valor entero de n ,

$$\{\tau = n\} = \{X_1 \notin A\} \cap \dots \cap \{X_{n-1} \notin A\} \cap \{X_n \in A\} \in \mathcal{F}_n.$$

En particular, y recordando el problema de la ruina del jugador, tomando $A = \{0, N\}$, el primer momento en el que la variable $X_n = \xi_1 + \dots + \xi_n$ toma uno de los dos valores del conjunto A es un tiempo de paro. Este tiempo de paro corresponde al tiempo aleatorio en el cual el juego finaliza, ya sea porque el jugador se arruina o porque gana todo el capital.

Ejemplo (Tiempo de segundo arribo). Sea nuevamente $X_1, X_2, \dots$ una sucesión de variables aleatorias adaptada a la filtración $\{\mathcal{F}_n\}_{n \geq 1}$, y sea A un conjunto de Borel de $\mathbb{R}$. Dado un tiempo de paro finito τ_1 , defina ahora un segundo tiempo de paro de la siguiente forma:

$$\tau_2 = \min\{n > \tau_1 : X_n \in A\}.$$

Es decir, τ_2 es el primer momento, después de τ_1 , en el que el proceso toma un valor dentro del conjunto A . Naturalmente se cumple $\tau_1 \leq \tau_2$. Esta nueva variable resulta también ser un tiempo de paro, pues el siguiente conjunto es un elemento de $\mathcal{F}_n$:

$$\{\tau_2 = n\} = \bigcup_{k=1}^{n-1} \left(\{\tau_1 = k\} \cap \{X_{k+1} \notin A\} \cap \dots \cap \{X_{n-1} \notin A\} \cap \{X_n \in A\} \right).$$

Martingalas

Definición

Se dice que un proceso a tiempo discreto $\{X_n : n \geq 1\}$ es una martingala respecto de una filtración $\{\mathcal{F}_n\}_{n \geq 1}$ si cumple las siguientes tres condiciones:

1. Es integrable.
2. Es adaptado a la filtración.
3. Para cualesquiera $n \leq m$, $E(X_m | \mathcal{F}_n) \stackrel{\text{c.s.}}{=} X_n$.

Cuando se cumple la desigualdad $E(X_m | \mathcal{F}_n) \stackrel{\text{c.s.}}{\geq} X_n$ entonces el proceso es una submartingala, y si $E(X_m | \mathcal{F}_n) \stackrel{\text{c.s.}}{\leq} X_n$ entonces es una supermartingala.

- ★ ¿Cómo se interpreta esta definición en el ejemplo del capital de un jugador que apuesta su dinero?
- ★ La condición $E(X_m | \mathcal{F}_n) \stackrel{\text{c.s.}}{=} X_n$ es equivalente a $E(X_{n+1} | \mathcal{F}_n) \stackrel{\text{c.s.}}{=} X_n$.

- Cuando la filtración es la natural, es decir, cuando $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$, la condición de martingala puede escribirse en la forma

$$E(X_{n+1} \mid X_1, \dots, X_n) \stackrel{\text{c.s.}}{=} X_n.$$

- Toda martingala es al mismo tiempo una submartingala y una supermartingala.
- Si $\{X_n : n \geq 1\}$ es una submartingala entonces $\{-X_n : n \geq 1\}$ es una supermartingala.
- ★ Todas las variables aleatorias que conforman una martingala tienen la misma esperanza. ¿Qué ocurre con la esperanza de las variables aleatorias en una submartingala?

Ejemplo (Martingala del juego de apuestas).

Sea $\xi_1, \xi_2, \dots$ una sucesión de variables aleatorias independientes, idénticamente distribuidas y con esperanza finita. Defina

$$X_n = \xi_1 + \dots + \xi_n,$$

y considere la filtración $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$.

La variable aleatoria X_n puede interpretarse como la fortuna de un jugador después de n apuestas sucesivas, donde la ganancia o pérdida en la k -ésima apuesta es ξ_k . Claramente el proceso $\{X_n : n \geq 1\}$ es integrable y adaptado.

Además, para cada $n \geq 1$,

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[X_n + \xi_{n+1} \mid \mathcal{F}_n] = X_n + \mathbb{E}[\xi_{n+1}].$$

Si $\mathbb{E}[\xi_{n+1}] = 0$, el proceso es una martingala. Si $\mathbb{E}[\xi_{n+1}] \geq 0$, es una submartingala. Si $\mathbb{E}[\xi_{n+1}] \leq 0$, es una supermartingala.

Ejemplo (Martingala del proceso de Poisson centrado).

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson de parámetro λ . Entonces el proceso centrado $\{X_t - \lambda t : t \geq 0\}$ es una martingala.

Este proceso es integrable y adaptado. Además, para $0 \leq s \leq t$,

$$E[X_t - \lambda t \mid \mathcal{F}_s] = E[X_t \mid \mathcal{F}_s] - \lambda t.$$

Usando incrementos independientes,

$$E[X_t \mid \mathcal{F}_s] = X_s + \lambda(t - s),$$

y por lo tanto,

$$E[X_t - \lambda t \mid \mathcal{F}_s] = X_s - \lambda s.$$

- ★ Si un proceso a tiempo continuo es integrable y tiene incrementos independientes entonces el proceso centrado es un martingala.

Ejemplo (Martingala de la esperanza condicional).

Sea X una variable aleatoria integrable, y sean $\mathcal{F}_1 \subseteq \mathcal{F}_2$ dos sub σ -álgebras. Se cumple la identidad

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_2] \mid \mathcal{F}_1] = \mathbb{E}[X \mid \mathcal{F}_1].$$

Sea $\{\mathcal{F}_n\}_{n \geq 1}$ una filtración dada. Defina

$$X_n = \mathbb{E}[X \mid \mathcal{F}_n].$$

Cada X_n es integrable y el proceso es adaptado a la filtración.

Además,

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n] = \mathbb{E}[X \mid \mathcal{F}_n] = X_n.$$

Por lo tanto, $\{X_n\}$ es una martingala.

Ejemplo (Martingala de la urna de Polya).

Una urna contiene inicialmente una bola blanca y una bola negra. En cada experimento se escoge una bola al azar, se regresa a la urna y se agrega otra bola del mismo color.

Sea X_n el número de bolas blancas después del n -ésimo ensayo.

Se tiene $X_0 = 1$ y después del n -ésimo ensayo hay $n + 2$ bolas en la urna. Las probabilidades de transición son

$$P(X_{n+1} = k + 1 \mid X_n = k) = \frac{k}{n + 2}, \quad P(X_{n+1} = k \mid X_n = k) = \frac{n + 2 - k}{n + 2}.$$

Se puede escribir $X_{n+1} = X_n + Y_{n+1}$, donde Y_{n+1} vale 1 si se escoge una bola blanca y 0 si es negra. Entonces, $E[X_{n+1} \mid X_n] = \frac{n+3}{n+2}X_n$, y el proceso $\{X_n\}$ no es una martingala.

Defina $M_n = \frac{X_n}{n+2}$, la fracción de bolas blancas en la urna. Entonces,

$E[M_{n+1} \mid \mathcal{F}_n] = \frac{X_n}{n+2} = M_n$. Por lo tanto, el proceso $\{M_n : n \geq 0\}$ es una martingala.

★ ¿Se modifica el resultado si la configuración inicial es distinta?

★ ¿Cambia el resultado si se agregan $r > 1$ bolas del mismo color?

Ejercicio

Sea $\xi_1, \xi_2, \dots$ una sucesión de variables aleatorias tal que el proceso $X_n = \xi_1 + \dots + \xi_n$ es un martingala. Demuestre que $E(\xi_i \xi_j) = 0$ para $i \neq j$.

Ejercicio

Sea $\{X_n : n \geq 0\}$ una cadena de Markov con espacio de estados $\{0, 1, \dots\}$. Suponga que las probabilidades de transición son $p_{00} = 1$ y $p_{ij} = e^{-i} \frac{i^j}{j!}$ en los otros casos. Demuestre que $\{X_n : n \geq 0\}$ es una martingala.

Procesos detenidos

Sea $\{X_n\}_{n \geq 0}$ un proceso estocástico adaptado a una filtración $\{\mathcal{F}_n\}_{n \geq 0}$, y sea τ un tiempo de paro respecto de la misma filtración.

Definición

Se define un proceso detenido $\{X_{\tau \wedge n}\}_{n \geq 0}$ donde $\tau \wedge n := \min\{\tau, n\}$, es decir,

$$X_{\tau \wedge n} = \begin{cases} X_n & \text{si } n \leq \tau, \\ X_\tau & \text{si } n > \tau. \end{cases}$$

Esto significa que como función del parámetro n el proceso se vuelve constante a partir del tiempo aleatorio τ .

- **Medibilidad y adaptabilidad de $\{X_{\tau \wedge n}\}_{n \geq 0}$.** Las variables del proceso detenido son efectivamente variables aleatorias y el proceso mismo es adaptado a la filtración. Si el proceso original es integrable, entonces el proceso detenido también lo es.
- Por ejemplo, considere que $\{X_n : n \geq 1\}$ es un martingala del juego de apuestas. Suponga que el jugador decide dejar de jugar cuando pierda todo su capital o cuando consiga ganar el doble de su capital inicial. El momento aleatorio en el que ocurre alguno de estos dos eventos es un tiempo de paro τ , así, el capital del jugador en cualquier momento n se puede expresar como $X_{\tau \wedge n}$.

Proposición

Si $\{X_n : n \geq 0\}$ es una martingala, submartingala o supermartingala, y τ es un tiempo de paro respecto de la misma filtración, entonces el proceso detenido $\{X_{\tau \wedge n} : n \geq 0\}$ también lo es.

Estrategias de juego

Considere la sucesión de variables aleatorias independientes e idénticamente distribuidas $\{\xi_n\}_{n \geq 1}$ tal que $P(\xi = -1) = P(\xi = 1) = \frac{1}{2}$ y con filtración natural $\{\mathcal{F}_n\}_{n \geq 1}$. Considere las sumas $X_n = \xi_1 + \dots + \xi_n$. Sabemos que $\{X_n : n \geq 1\}$ es una martingala que representa el total de ganancias en una serie de n apuestas justas de una unidad monetaria.

Ahora, suponga que el monto de cada apuesta no es uno sino una cantidad a_n para la n -ésima apuesta. Supondremos que a_n es una variable aleatoria que el jugador determina a partir de la información de las $n - 1$ apuestas previas y por lo tanto es $\mathcal{F}_{n-1}$ -medible, es decir, se trata de un proceso predecible.

A la colección de variables aleatorias $\{a_n : n \geq 1\}$ descrita anteriormente se le denomina una estrategia de juego. Bajo una de estas estrategias, la ganancia del jugador después de la n -ésima apuesta es ahora la variable aleatoria $\mathcal{F}_n$ -medible

$$A_n = a_1\xi_1 + \dots + a_n\xi_n.$$

- ★ Bajo la hipótesis de que la estrategia de juego consta de variables aleatorias acotadas, demuestre que el proceso $\{A_n : n \geq 1\}$ es integrable y además cumple la propiedad de martingala.

Es decir, bajo cualquier estrategia de juego, el proceso de ganancias $\{A_n : n \geq 1\}$ es una martingala siempre y cuando el proceso original $\{X_n : n \geq 1\}$ lo sea.

- ★ ¿Cuál es la importancia de esto para los apostadores? ¿Existe una estrategia que convierta un juego justo en un juego favorable o desfavorable al jugador?
- ★ Haga una simulación del proceso de ganancias dado que tiene un proceso original justo, luego de esto, compruebe que el proceso de ganancias es una martingala.

Ejercicio

Sea $\{X_t : t \geq 0\}$ un proceso de Poisson homogéneo de parámetro $\lambda > 0$. Demuestre que los siguiente procesos son martingalas:

1. $Y_t = (X_t - \lambda t)^2 - \lambda t$.
2. $Y_t = \exp(-\theta X_t + \lambda t(1 - e^{-\theta}))$ para cualquier $\theta \in \mathbb{R}$.

Movimiento browniano

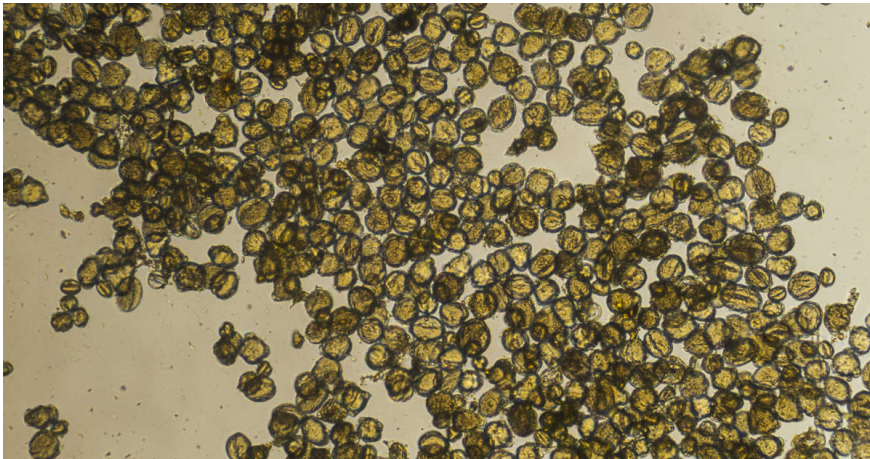


Figura: Granos de polen suspendidos en ciertas sustancia siguen trayectorias irregulares e impredecibles de movimiento (Link).

Definición

Un movimiento Browniano unidimensional de parámetro σ^2 es un proceso estocástico $\{B_t : t \geq 0\}$ con valores en $\mathbb{R}$ que cumple las siguientes condiciones:

1. $B_0 = 0$ casi seguramente.
2. Las trayectorias son continuas.
3. El proceso tiene incrementos independientes.
4. Para cualesquiera tiempos $0 \leq s < t$, la variable incremento $B_t - B_s$ tiene distribución $N(0, \sigma^2(t - s))$.

- La variable B_t representa la posición de una partícula en el tiempo t .
- Debido al gran número de colisiones del grano de polen con las moléculas circundantes en longitudes de tiempo no pequeñas, y teniendo en cuenta el teorema central del límite, los incrementos suelen modelarse como variables aleatorias independientes con distribución Gaussiana.
- En 1923 el matemático N. Wiener demostró la unicidad y la existencia de un proceso estocástico con las propiedades mencionadas anteriormente, por ende este proceso también es conocido como proceso de Wiener. En sentido estricto, el movimiento Browniano es el fenómeno físico y el proceso de Wiener es el modelo matemático que lo representa.
- Este proceso es un ejemplo importante en cadenas de Markov a tiempo continuo con espacio de estados continuo.
- La definición anterior implica que todos los procesos de Wiener tiene incrementos estacionarios.
- Un proceso Browniano es estándar si y solo si $\sigma^2 = 1$.

Ejercicio

Dado un movimiento Browniano $\{B_t : t \geq 0\}$ unidimensional con parámetro σ^2 , multiplique cada una de las variables aleatorias por un valor que convierta el proceso en un movimiento Browniano estándar.

Ejercicio

Dado un movimiento Browniano $\{B_t : t \geq 0\}$ unidimensional con parámetro σ^2 , demuestre que los siguientes procesos estocásticos son movimientos Brownianos y halle sus respectivos parámetros:

1. $\{W_t = -B_t : t \geq 0\}$.
2. $\{W_t = \frac{1}{c}B_c^2 t : t \geq 0\}$, con $c > 0$ constante.
3. $\{W_t = tB_{\frac{1}{t}} : t \geq 0\}$, con $W_0 \stackrel{\text{c.s.}}{=} 0$.
4. $\{W_t = B_{t_0+t} - B_{t_0} : t \geq 0\}$, con $t_0 \geq 0$ fijo.

Definición (alternativa)

Un movimiento Browniano unidimensional de parámetro σ^2 es un proceso estocástico $\{B_t : t \geq 0\}$ con valores en $\mathbb{R}$ que cumple las siguientes propiedades:

- $B_0 = 0$ casi seguramente.
- Las trayectorias $t \mapsto B_t$ son continuas.
- Para cualesquiera tiempos $0 < t_1 < \dots < t_n$ y para cualesquiera conjuntos de Borel $A_1, \dots, A_n$ de $\mathbb{R}$ se cumple que la probabilidad

$$P(B_{t_1} \in A_1, \dots, B_{t_n} \in A_n)$$

es igual a

$$\int_{A_1} \dots \int_{A_n} p(t_1, 0, x_1) p(t_2 - t_1, x_1, x_2) \dots p(t_n - t_{n-1}, x_{n-1}, x_n) dx_n \dots dx_1$$

en donde $p(t, x, y) = \frac{1}{\sqrt{2\pi\sigma t}} e^{-\frac{(x-y)^2}{2\sigma^2 t}}$. En particular, B_t tiene distribución $N(0, \sigma^2 t)$.

Transición

Definimos la función de densidad de transición $p : \mathbb{R}_{\geq 0} \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ del movimiento Browniano de parámetro σ^2 como

$$p(t, x, y) = \frac{1}{\sqrt{2\pi\sigma t}} e^{-\frac{(x-y)^2}{2\sigma^2 t}}.$$

En particular, la probabilidad de que un movimiento Browniano que inicia en x se encuentre en un conjunto medible $A \subseteq \mathbb{R}$ después de t unidades de tiempo es:

$$p(t, x, A) = \int_A p(t, x, y) dy.$$

Ejercicio

- Demuestre que la función de densidad de transición cumple la ecuación de Chapman-Kolmogorov:

$$p(t + s, x, y) = \int_{\mathbb{R}} p(t, x, u)p(s, u, y) du.$$

- Demuestre que la función de densidad de transición cumple la ecuación de difusión o ecuación del calor:

$$\frac{\partial p}{\partial t} = \frac{1}{2} \frac{\partial^2 p}{\partial x^2}.$$

Proposición

El movimiento Browniano es un proceso de Markov.

Demostración. Para cualesquiera tiempos $0 \leq t_1 < \dots < t_n < t_{n+1}$ y para cualquier evento $A \subseteq \mathbb{R}$,

$$\begin{aligned} & P(B_{t_{n+1}} \in A \mid B_{t_n} = x_n, \dots, B_{t_1} = x_1) \\ &= \frac{p(t_1, 0, x_1)p(t_2 - t_1, x_1, x_2) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n)p(t_{n+1} - t_n, x_n, A)}{p(t_1, 0, x_1)p(t_2 - t_1, x_1, x_2) \cdots p(t_n - t_{n-1}, x_{n-1}, x_n)} \\ &= p(t_{n+1} - t_n, x_n, A) \\ &= p(t_{n+1} - t_n, x_n, A) \frac{p(t_n, 0, x_n)}{p(t_n, 0, x_n)} \\ &= P(B_{t_{n+1}} \in A \mid B_{t_n} = x_n). \end{aligned}$$

□

Proposición

El movimiento Browniano es una martingala continua en el tiempo.

Demostración. El proceso es adaptado a su filtración natural por definición y cada variable aleatoria del proceso es integrable. Por otro lado, para cualesquiera tiempos s y t tales que $0 \leq s < t$:

$$\begin{aligned} \mathbb{E}(B_t \mid \mathcal{F}_s) &= \mathbb{E}(B_t - B_s + B_s \mid \mathcal{F}_s) \\ &= \mathbb{E}(B_t - B_s \mid \mathcal{F}_s) + \mathbb{E}(B_s \mid \mathcal{F}_s) \\ &= \mathbb{E}(B_t - B_s) + B_s \\ &= B_s. \end{aligned}$$

□

Proposición (Teorema de caracterización de Lèvy)

Un proceso $\{X_t : t \geq 0\}$ es un movimiento Browniano estándar si, y solo si, tiene trayectorias continuas, empieza en cero, y tanto $\{X_t : t \geq 0\}$ como $\{X_t^2 - t : t \geq 0\}$ son martingalas.

- ★ Demuestre que si $\{B_t : t \geq 0\}$ es un movimiento Browniano estándar entonces $\{B_t^2 - t : t \geq 0\}$ es una martingala.

Caminata aleatoria $\rightarrow$ Movimiento Browniano

- Dada una sucesión de variables aleatorias $\{\xi_i\}_{i \geq 1}$ con $P(\xi = -1) = P(\xi = 1) = \frac{1}{2}$ se construye una caminata aleatoria simétrica como $X_n = X_0 + \xi_1 + \dots + \xi_n$ para $n \geq 1$ donde $X_0 \stackrel{c.s.}{=} 0$. Este proceso estocástico a tiempo discreto cumple que $E(X_n) = 0$ y $\text{Var}(X_n) = n$.
- Para que una caminata aleatoria simétrica converja a un movimiento Browniano debemos hacer un reescalamiento en el tiempo y en el espacio y luego hacer un paso al límite: Definimos el proceso a tiempo continuo para cada $n \geq 0$, $\{W_{n,t} : t \geq 0\}$ como $W_{n,t} = \frac{1}{\sqrt{n}} X_{[nt]}$ donde $[nt]$ representa el número de pasos completados hasta el tiempo t . Se puede probar que la sucesión de procesos estocásticos $\{W_{n,t}\}_{n \in \mathbb{N}}$ tiende débilmente al movimiento Browniano de parámetro $\sigma^2 = 1$ $\{B_t : t \geq 0\}$ dado que $W_{n,t}$ tiende en distribución a $N(0, t)$ por el **teorema del límite central**.
- Esta es la forma más usual en la que se simulan trayectorias del movimiento Browniano a nivel computacional.
- Para más información se invita a consultar el Teorema de Donsker.

Aproximación de Movimiento Browniano ($\sigma^2 = 1$)

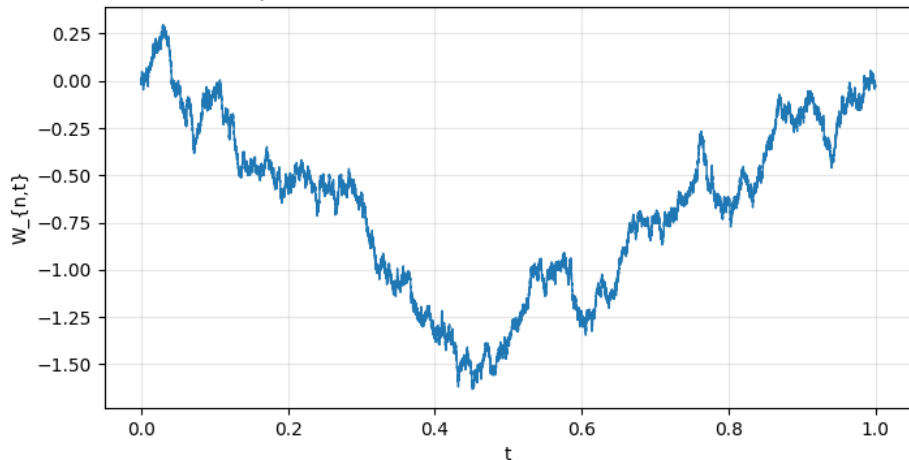


Figura: Simulación de una trayectoria del movimiento Browniano usando una aproximación por medio de caminata aleatoria tomando $n = 50000$ y el tiempo de simulación $t \in [0, 1]$.

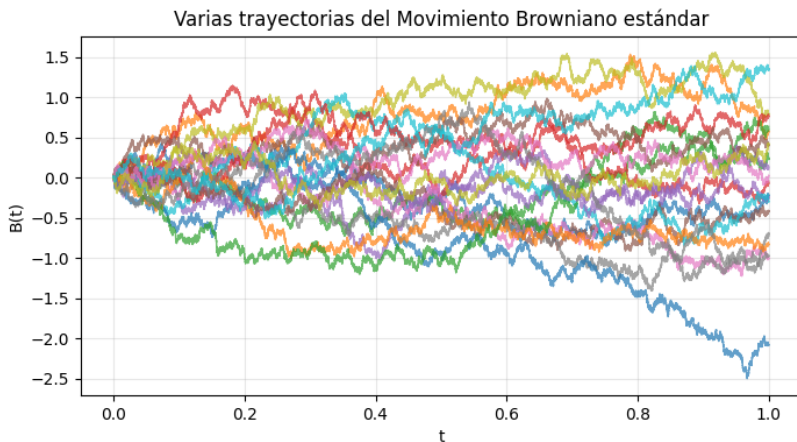


Figura: Simulación de $M = 20$ trayectorias del movimiento Browniano por medio de una aproximación de caminata aleatoria cada una con tiempo de simulación $t \in [0, 1]$ y $n = 50000$.

Convergencia de caminata aleatoria reescalada a Browniano estándar

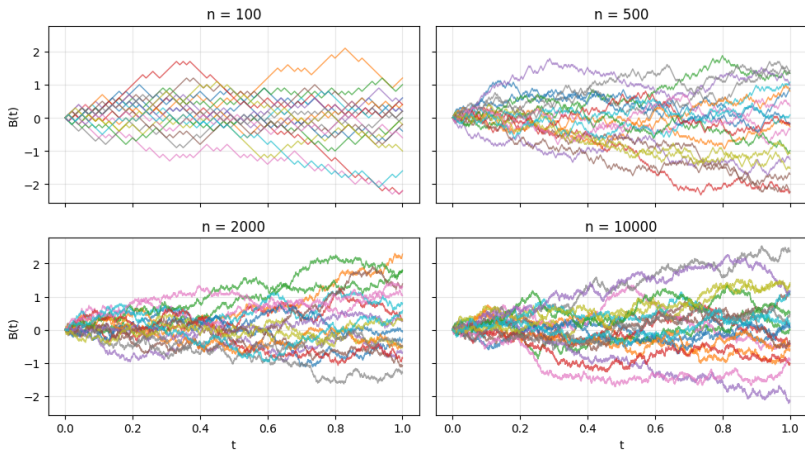


Figura: Simulación de caminatas aleatorias $W_{n,t}$ para diferentes valores de n y para $t \in [0, 1]$. A medida que n aumenta, la caminata aleatoria se asemeja cada vez más al movimiento Browniano.

- Suponga que se coloca una partícula al azar en la recta real de acuerdo a una función de densidad de probabilidad $f(y)$. Suponga que esta partícula se mueve siguiendo un movimiento Browniano estándar unidimensional. Entonces la función de densidad de probabilidad de la posición de la partícula en el tiempo $t \geq 0$ es la función $f(t, x)$ dada por

$$f(t, x) = \int_{\mathbb{R}} f(y)p(t, y, x)dy = \mathbb{E}(f(x + B_t)).$$

- Se puede demostrar que la función $f(t, x)$ satisface la ecuación diferencial parcial de difusión

$$\frac{\partial f}{\partial t}(t, x) = \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, x).$$

Variación de las trayectorias

Definición

Sea $a = t_0 < t_1 < \dots < t_n = b$ una partición de intervalo $[a, b] \subseteq \mathbb{R}$ y defina $\Delta t = \max_{i=0, \dots, n-1} |t_{i+1} - t_i|$. La variación de una función $g : [a, b] \rightarrow \mathbb{R}$ es el número

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)|.$$

Cuando este número existe se dice que la función tiene variación finita. Análogamente la variación cuadrática se define como

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |g(t_{i+1}) - g(t_i)|^2.$$

Proposición

La variación cuadrática sobre el intervalo $[a, b]$ de una trayectoria del movimiento Browniano es la longitud del intervalo, es decir,

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |B_{t_{i+1}} - B_{t_i}|^2 = b - a \quad \text{en el sentido } L^2(P).$$

Demostración. Denote por $\Delta t_i = t_{i+1} - t_i$ y $\Delta B_i = B_{i+1} - B_i$. Entonces

$$\mathbb{E} \left| \sum_i (\Delta B_i)^2 - (b - a) \right|^2 = \sum_i 2(\Delta t_i)^2 \leq 2(b - a) \max_{0 \leq i < n} (\Delta t_i) \rightarrow 0.$$

□

Proposición

La variación sobre el intervalo $[a, b]$ de una trayectoria del movimiento Browniano es infinita, casi seguramente, es decir,

$$\limsup_{\Delta t \rightarrow 0} \sum_{i=0}^{n-1} |B_{t_{i+1}} - B_{t_i}| \stackrel{\text{c.s.}}{=} \infty.$$

Demostración.

No diferenciabilidad

Proposición

Sea $t_0 \geq 0$ fijo. Con probabilidad uno, el movimiento Browniano $\{B_t : t \geq 0\}$ no es diferenciable en t_0 , es decir, para casi todo $\omega \in \Omega$ se tiene que la función $t \mapsto B_t(\omega)$ es no diferenciable en t_0 .

Demostración.

No diferenciabilidad

Proposición

Con probabilidad uno, el movimiento Browniano $\{B_t : t \geq 0\}$ no es diferenciable ningún $t \geq 0$.

Demostración.

Movimiento Browniano multidimensional

Definición

Sean $\{B_{1,t}\}_{t \geq 0}, \dots, \{B_{n,t}\}_{t \geq 0}$ movimientos Brownianos unidimensionales independientes. El movimiento Browniano en $\mathbb{R}^n$ se define como

$$\mathbf{B}_t = (B_{1,t}, \dots, B_{n,t}).$$

- Igual que en el caso unidimensional se tiene que $\mathbf{B}_0 \stackrel{\text{c.s.}}{=} (0, \dots, 0)$, las trayectorias $t \mapsto \mathbf{B}_t$ son continuas y el proceso tiene incrementos independientes.
- Para cualesquiera tiempo $0 \leq s < t$ se tiene que el vector $\mathbf{B}_t - \mathbf{B}_s$ tiene distribución normal multivariada con media en el origen y matriz de covarianza diagonal $\text{Var}(\mathbf{B}_t - \mathbf{B}_s) = \text{diag}((t-s)\sigma_i^2)$.

Trayectorias de Movimiento Browniano estándar en $\mathbb{R}^2$
($T = 1$, $n = 50000$)

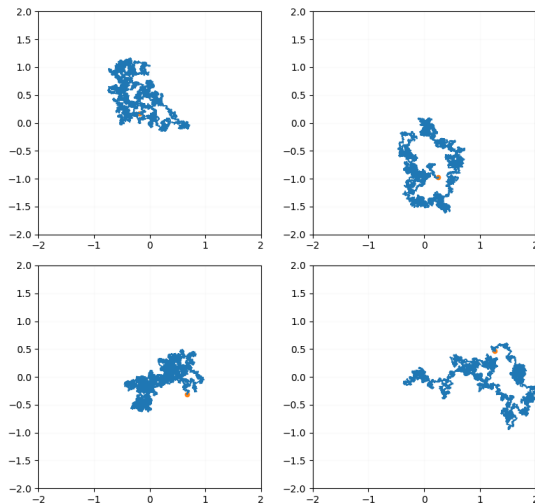


Figura: Cuatro simulaciones de trayectorias del movimiento Browniano estándar en $\mathbb{R}^2$ para el tiempo $t \in [0, 1]$.

El proceso de Bessel. Sea $\{\mathbf{B}_t : t \geq 0\}$ un movimiento Browniano multidimensional. El proceso de Bessel está dado por

$$R_t = \|\mathbf{B}_t\| = (B_{1,t}^2 + \cdots B_{n,t}^2)^{\frac{1}{2}}.$$

El proceso $\{R_t : t \geq 0\}$ es la distancia Euclideana respecto al origen que tiene un movimiento Browniano n -dimensional en el tiempo t . El proceso toma valores en $[0, \infty)$ y tiene trayectorias continuas.

Puede demostrarse que este proceso cumple la propiedad de Markov y que la función de densidad de transición $p(t, x, y)$ puede expresarse en términos de las funciones de Bessel.

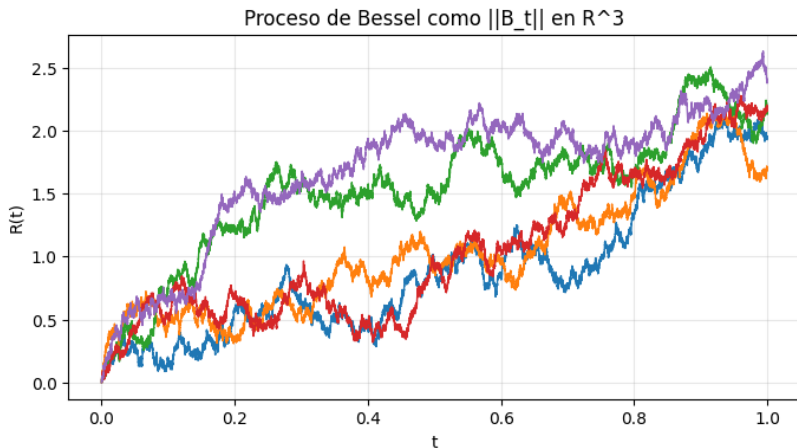


Figura: Diferentes trayectorias del proceso de Bessel luego de simular el movimiento Browniano multidimensional en $\mathbb{R}^3$ en el tiempo $t \in [0, 1]$.

Ecuación del calor en dominios acotados. Considere una región abierta y acotada $D \subseteq \mathbb{R}^n$ con frontera ∂D y suponga que la función $u(t, x)$ representa la temperatura en el punto $x \in D$ al tiempo $t \geq 0$.

La evolución en el tiempo de esta función está dada por la ecuación del calor:

$$\frac{\partial}{\partial t} u(t, x) = \frac{d}{2} \Delta u(t, x),$$

donde Δ represental el operador Laplaciano y d es una constante real positiva. Suponga además que $u(x, 0) = f(x)$ para $x \in D$ y la condición de frontera $u(t, x) = g(x)$ para $x \in \partial D$.

Sea $x \in D$ y defina el tiempo $\tau = \inf\{t > 0 : B_t^x \in \partial D\}$ donde $\{B_t^x : t \geq 0\}$ es un movimiento Browniano de parámetro $\sigma^2 = d$ que inicia en x . La solución a la ecuación del calor con condiciones iniciales y de frontera se puede expresar en términos del movimiento Browniano:

$$u(t, x) = \mathbb{E}(f(B_t^x) \mathbf{1}_{(\tau > t)} + g(B_t^x) \mathbf{1}_{(\tau \leq t)}).$$

El principio de reflexión

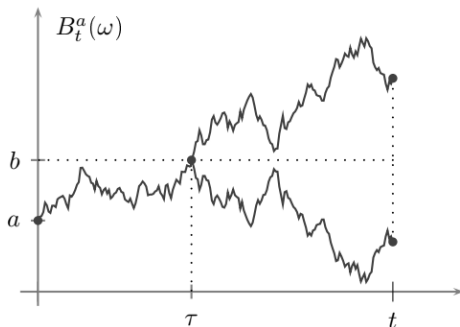


Figura: Ilustración del principio de reflexión del movimiento Browniano unidimensional. Tomado de Rincon, 2012.

Proposición

Sea $\{B_t^a : t \geq 0\}$ un movimiento Browniano que empieza en a , y sea $b > a$. Para cualquier $t > 0$,

$$P(B_s^a \geq b \text{ para algún } s \in [0, t]) = 2P(B_t^a \geq b).$$

- El conjunto de trayectorias que en algún tiempo $\tau \in [0, t]$ tocan b se descompone en dos conjuntos que son simétricos y tienen la misma probabilidad: aquel conjunto de trayectorias que finalizan en el algún punto mayor a b y el conjunto de trayectorias que terminan por debajo de b .
- Una vez que una trayectoria toca el punto b es igualmente probable que finalice al tiempo t arriba o abajo de b .

Demostración. Sea τ el primer momento en el que el movimiento Browniano es igual a b , es decir, $\tau = \inf\{s \geq 0 : B_s^a = b\}$. Esta variable aleatoria puede tomar el valor infinito si esto nunca ocurre. Entonces:

$$P(B_s^a \geq b \text{ para algún } s \in [0, t]) = P(\tau \leq t) = P(\tau < t).$$

Por otro lado,

$$P(B_t^a \geq b) = P(B_t^a \geq b \mid \tau < t)P(\tau < t) = P(B_t^a - b \geq 0 \mid \tau < t)P(\tau < t),$$

en donde, por la propiedad de Markov, y condicionada a la ocurrencia del evento $(\tau < t)$, la variable $B_t^a - b = B_t^a - B_\tau^a$ tiene distribución $N(0, (t - \tau)\sigma^2)$. Por lo tanto $P(B_t^a - b \geq 0 \mid \tau < t) = 1/2$.

Sustituyendo se obtiene que $P(B_t^a \geq b) = 2P(\tau < t)$. □

Recurrencia y transitoriedad

- Así como se estudió en el capítulo de Cadenas de Markov el posible retorno del proceso al punto de partida, también vamos a estudiar el retorno del movimiento Browniano al punto de origen. Veremos que esto depende de la dimensión del proceso, así como el caso de una cadena de Markov en tiempo discreto y estado discreto.

Proposición

Sea $\{\mathbf{B}_t : t \geq 0\}$ un movimiento Browniano multidimensional que inicia en el origen y considere dos tiempos tales que $0 < t_1 < t_2$. Entonces

$$P(\mathbf{B}_t = 0 \text{ para algún } t \in [t_1, t_2]) = 1 - \frac{2}{\pi} \arctan \left(\frac{\sqrt{t_1}}{\sqrt{t_2 - t_1}} \right)$$

Demostración.

Proposición (Recurrencia puntual del movimiento Browniano unidimensional)

Con probabilidad uno el movimiento Browniano unidimensional es recurrente puntual, es decir, regresa a su posición de partida una infinidad de veces casi seguramente.

Demostración.

Proposición (Recurrencia y transitoriedad del movimiento Browniano)

Sea $\{\mathbf{B}_t : t \geq 0\}$ un movimiento Browniano n -dimensional que inicia en el origen, y sea el disco $D = \{x \in \mathbb{R}^n : \|x\| < r\}$, para algún $r > 0$.

1. Cuando $n = 2$, con probabilidad uno el movimiento Browniano visita la vecindad D en una infinidad de tiempos no acotados. Esto es, el proceso es recurrente por vecindades. Sin embargo no es recurrente puntual, pues la probabilidad de que regrese exactamente al punto de partida es cero.
2. Cuando $n \geq 3$, el proceso es transitorio por vecindades, es decir, existe una probabilidad positiva de que el proceso nunca regrese a la vecindad del punto de partida.

Demostración.

Ejercicio

Sea una constante $\sigma > 0$. Demuestre que si $\{B_t : t \geq 0\}$ es un movimiento Browniano estándar, entonces los siguientes son movimientos Brownianos de parámetro σ^2 :

1. $\{W_t = \sigma B_t : t \geq 0\}$.
2. $\{W_t = B_{\sigma^2 t} : t \geq 0\}$.

Ejercicio

Sean las funciones $a(t) = e^{2t}$ y $b(t) = e^{-t}$. Calcule la función de covarianza del proceso $\{X_t = b(t)B_{a(t)} : t \geq 0\}$ en donde $\{B_t : t \geq 0\}$ es un movimiento Browniano estándar.

Procesos Gaussianos

Procesos estacionarios

Definición

Sea $\{X_t\}$ un proceso estocástico. Se dice que $\{X_t\}$ es estrictamente estacionario o estacionario en el sentido estricto si

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$$

para todo $n \in \mathbb{N}^*$, todos los puntos $(t_1, t_2, \dots, t_n) \in \mathcal{T}^n$ y todos los saltos $h \in \mathbb{Z}$.

Lo que se entiende de la anterior definición es que la distribución de cualquier vector aleatorio finito dimensional que se puede formar con las variables del proceso $\{X_t\}$, tiene distribución invariante ante traslaciones en el tiempo.

Procesos estacionarios

Función de medias

La función de medias del proceso estocástico $\{X_t\}$ se define como

$$\mu_t = E[X_t]$$

asumiendo que el $E[X_t] < \infty$.

Función de autocovarianza

Si $\{X_t\}$ es un proceso estocástico tal que $E|X_t|^2 < \infty$ para cada $t \in \mathbb{Z}$, entonces la función de autocovarianza $\gamma_X(\cdot, \cdot)$ de $\{X_t\}$ es definido como el segundo momento

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - E[X_r])(X_s - E[X_s])]$$

para cada $r, s \in \mathbb{Z}$.

Procesos estacionarios

Definición

El proceso estocástico $\{X_t\}$ es llamado estacionario (en el sentido débil o de segundo orden) si

1. $E |X_t|^2 < \infty$ para todo t .
2. $\mu_t = E[X_t] = m$ para todo t .
3. $\gamma_X(r, s) = \gamma_X(r + t, s + t)$ para todo r, s, t .

Así para $t = -s$, por el numeral 3 tenemos que

$\gamma_X(r, s) = \gamma_X(r - s, s - s) = \gamma_X(r - s, 0) := \gamma(r - s)$, es decir, la función $\gamma_X(,)$ depende sólo de r, s a través de su diferencia $r - s = h$.

Ejercicio

Considere un proceso estocástico $\{Z_t\}$. Diremos que el proceso es ruido blanco si

1. $E[Z_t] = 0$ para todo t .
2. La función de autocovarianza es

$$\gamma(s, t) = \begin{cases} 0 & \text{si } s \neq t, \\ \sigma_Z^2 & \text{si } s = t. \end{cases}$$

Mostrar que el proceso es estacionario en el sentido débil.

Ejercicio

Si $\{X_t\}$ es tal que las variables aleatorias son independientes y

$$X_t \sim \begin{cases} \exp(\lambda = 1), & t \text{ impar}, \\ N(\mu = 1, \sigma = 1), & t \text{ par}. \end{cases}$$

Mostrar que el proceso es estacionario en el sentido débil pero no es estrictamente estacionario.

Ejercicio

Demostrar que todo proceso estrictamente estacionario es también estacionario en el sentido débil.

Proceso Gaussiano

Definición

Un proceso $\{X_t\}$ es llamado Gaussiano si los vectores n -dimensionales $\mathbf{X}_{\tilde{t}} = (X_{t_1}, X_{t_2}, \dots, X_{t_n})'$, para cada colección de puntos distintos $t_1, t_2, \dots, t_n$, y para cada entero positivo n , tienen una distribución normal multivariada.

Ejercicio

Si un proceso $\{X_t\}$ es estacionario en el sentido débil y también es Gaussiano, entonces $\{X_t\}$ es estacionario en el sentido estricto.

Propiedades de la función de autocovarianza

Considere que tenemos un proceso estocástico estacionario $\{X_t\}$, y que $\gamma(h) = \text{Cov}(X_{t+h}, X_t) = E[(X_{t+h} - \mu)(X_t - \mu)]$, entonces $\gamma(\cdot)$ satisface:

1. $\gamma(0) \geq 0$
2. $|\gamma(h)| \leq \gamma(0)$ para todo $h \in \mathbb{Z}$
3. Es una función par, es decir,

$$\gamma(h) = \gamma(-h), \text{ para todo } h \in \mathbb{Z}$$

Demostración.

1. Como $\gamma(0) = \text{Cov}(X_t, X_t) = \text{Var}(X_t) \geq 0$.
2. Por definición de la función de covarianza tenemos que $\gamma(h) = \text{Cov}(X_{t+h}, X_t) = E[(X_{t+h} - \mu)(X_t - \mu)]$. Ahora, usando la desigualdad de Cauchy-Schwarz $(|E[XY]| \leq \sqrt{E[X^2]}\sqrt{E[Y^2]})$ tenemos que

$$\begin{aligned} |\gamma(h)| &= |E[(X_{t+h} - \mu)(X_t - \mu)]| \leq \sqrt{E[(X_{t+h} - \mu)^2]} \sqrt{E[(X_t - \mu)^2]} \\ &= \sqrt{\text{Var}(X_t) \text{Var}(X_t)} = \text{Var}(X_t) = \gamma(0) \end{aligned}$$

Con lo cual $|\gamma(h)| \leq \gamma(0)$.

3. Ahora, $\gamma(-h) = \text{Cov}(X_{t-h}, X_t) = \text{Cov}(X_t, X_{t+h}) = \text{Cov}(X_{t+h}, X_t) = \gamma(h)$.

Ejercicio

Mostrar que el movimiento Browniano $\{B_t : t \geq 0\}$ es un proceso Gaussiano con:

- $\mu_t = 0$.
- $\gamma(s, t) = \min(s, t)$ para todo s, t .

Proceso Ornstein–Uhlenbeck

Definición

El proceso Ornstein–Uhlenbeck es un proceso Gaussiano estacionario definido por:

$$\mathbb{E}[X_t] = 0, \quad \gamma(s, t) = \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|}.$$

donde:

- $\sigma^2 > 0$ es el parámetro de intensidad del ruido.
- $\alpha > 0$ es el parámetro de velocidad de reversión a la media.

Es usado en física, finanzas y modelos de reversión a la media.

Cálculo estocástico

Objetivo. Definir la integral de Itô $\int_0^t X_s dB_s$ de un proceso $\{X_t\}$ respecto del movimiento Browniano.

Primeras hipótesis. Consideraremos como elementos iniciales un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, y un movimiento Browniano estándar unidimensional $\{B_t : t \geq 0\}$, junto con su filtración natural $\{\mathcal{F}_t\}_{t \geq 0}$. Supondremos dado un proceso $\{X_t\}$ con espacio parametral el intervalo $[0, T]$, con $T > 0$ fijo, y que visto como función

$$X : \Omega \times [0, T] \rightarrow \mathbb{R},$$

es $\mathcal{F}_T \otimes \mathcal{B}([0, T])$ -medible, en donde el término $\mathcal{F}_T \otimes \mathcal{B}([0, T])$ corresponde a la mínima σ -álgebra generada por el espacio producto $\mathcal{F}_T \times [0, T]$. Supondremos además que el proceso es adaptado, es decir, para cada t en el intervalo $[0, T]$, la variable aleatoria X_t es medible respecto de $\mathcal{F}_t$.

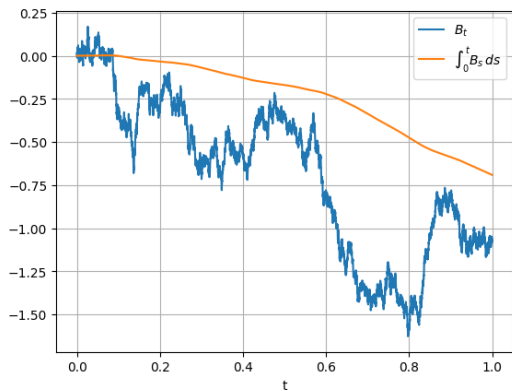


Figura: Simulación de una trayectoria del movimiento Browniano y de su respectiva integral.

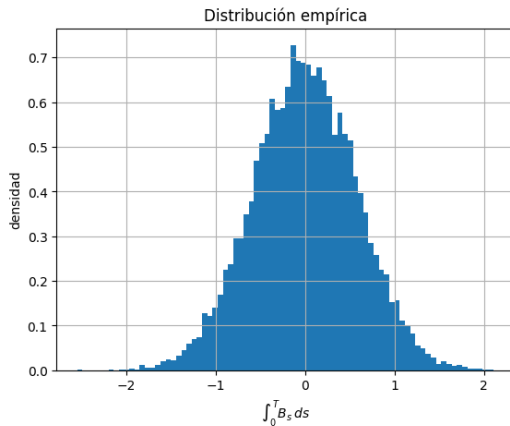


Figura: Distribución empírica de la variable aleatoria $\int_0^{T=1} B_s ds$.

El espacio $L^2(P)$. Definimos:

$$L^2(P) = \{X \text{ variable aleatoria} : \|X\|_{L^2(P)} := \mathbb{E}(|X|^2)^{\frac{1}{2}} < \infty\},$$

es decir, $L^2(P)$ es el espacio de variables aleatorias que son cuadrado integrables.

- El espacio $L^2(P)$ es un espacio vectorial respecto a $\mathbb{R}$.
- La función $\|\cdot\|_{L^2(P)}$ define una norma en $L^2(P)$.
- El espacio $L^2(P)$ es un espacio completo respecto su norma, es decir, toda sucesión de Cauchy en este espacio tiene límite en él. A la convergencia en el espacio $L^2(P)$ se le llama convergencia en media cuadrática.
- Cada una de las variables aleatorias B_t del movimiento Browniano pertenecen a este espacio, pues,

$$\|B_t\|_{L^2(P)} = \mathbb{E}(|B_t|^2)^{\frac{1}{2}} = \sqrt{t} < \infty.$$

El espacio $L^2(P \times dt)$. Se denotará como $L^2(P \times dt)$ al espacio lineal de todos los procesos $X = \{X_t : 0 \leq t \leq T\}$ que cumplan la condición

$$\|X\|_{L^2(P \times dt)} = \left(\mathbb{E} \left(\int_0^T |X_t|^2 dt \right) \right)^{\frac{1}{2}} < \infty.$$

- La función $\|\cdot\|_{L^2(P \times dt)}$ es una norma, y el espacio $L^2(P \times dt)$ es un espacio completo respecto a esta norma.
- El movimiento Browniano $B = \{B_t : 0 \leq t \leq T\}$ pertenece a este espacio, pues,

$$\|B\|_{L^2(P \times dt)} = \left(\mathbb{E} \left(\int_0^T |X_t|^2 dt \right) \right)^{\frac{1}{2}} = \left(\int_0^T \mathbb{E}(|X_t|^2) dt \right)^{\frac{1}{2}} = \left(\int_0^T t dt \right)^{\frac{1}{2}} < \infty.$$

Procesos simples. Definiremos la integral de Itô para procesos simples.

Procesos simples

Sea $0 = t_0 < t_1 < \dots < t_n = T$ una partición finita del intervalo $[0, T]$. Un proceso estocástico simple es un proceso de la forma

$$X_t = \sum_{k=0}^{n-1} X^{(k)} 1_{[t_k, t_{k+1})}(t),$$

en donde $X^{(0)}, \dots, X^{(n-1)}$ es una colección de variables aleatorias adaptadas a la filtración $\{\mathcal{F}_{t_k}\}_{k=0}^{n-1}$ y que son cuadrado integrables.

- Un proceso simple es un proceso constante por trozos con trayectorias continuas por la derecha y con límite por la izquierda.
- Sea $\mathcal{H}_0^2$ el espacio de todos los procesos simples. Este es un espacio vectorial.

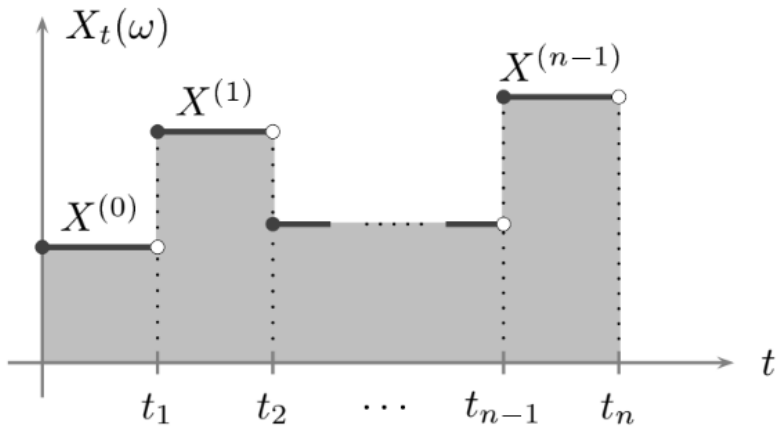


Figura: Representación de una posible trayectoria de un proceso simple. Imagen tomada de Rincon, 2012.

Integral para procesos simples.

Definición

La integral estocástica de Itô $I(X)$ de un proceso simple X de la forma

$$X_t = \sum_{k=0}^{n-1} X^{(k)} 1_{[t_k, t_{k+1})}(t),$$

con variables aleatorias $X^{(k)}$ independientes del movimiento Browniano estándar $\{B_t : t \geq 0\}$, se define como la variable aleatoria

$$I(X) = \int_0^T X_s dB_s = \sum_{k=0}^{n-1} X^{(k)} (B_{t_{k+1}} - B_{t_k}).$$

- La variable aleatoria $I(X)$ es integrable pues siendo las variables aleatorias $X^{(k)}$ y $B_{t_{k+1}} - B_{t_k}$ independientes, cada sumando tiene esperanza cero, y por lo tanto la esperanza de la integral es cero.
- ★ La integral es además cuadrado integrable y se cumple la igualdad fundamental llamada Isometría de Itô:

$$\|I(X)\|_{L^2(P)} = \|X\|_{L^2(P \times dt)}.$$

- La integral estocástica asigna a cada elemento del espacio $\mathcal{H}_0^2$ una variable aleatoria del espacio $L^2(P)$. De esta forma se tiene que la transformación lineal $I : \mathcal{H}_0^2 \rightarrow L^2(P)$ resulta ser continua por la isometría de Itô.

Ejercicio

Se puede tomar el movimiento Browniano discretizado

$$\{B_{t_k} : 0 = t_0 < t_1 < \cdots < t_n = T\}$$

que es un proceso simple, y luego definir la integral de Itô de este proceso simple:

$$I(X) = \sum_{k=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}).$$

Calcule la media y la varianza de $I(X)$ y luego tome el límite cuando $\max(t_{k+1} - t_k) \rightarrow 0$.

Extensión por aproximación. Ahora extendemos la integral estocástica a procesos un poco más generales. Sea $\mathcal{H}^2$ el espacio de todos los procesos $\{X_t\}$ medibles y adaptados, tales que

$$\mathbb{E} \left(\int_0^T |X_t|^2 dt \right) < \infty.$$

- El espacio $\mathcal{H}^2$ resulta ser un subespacio lineal y cerrado de $L^2(P \times dt)$.
- Se cumple la contención $\mathcal{H}_0^2 \subseteq \mathcal{H}^2 \subseteq L^2(P \times dt)$.
- $\mathcal{H}_0^2$ es un subespacio denso de $\mathcal{H}^2$ respecto a la norma $\|\cdot\|_{L^2(P \times dt)}$, es decir, para cualquier proceso $X \in \mathcal{H}^2$ existe una subsucesión de procesos $\{X^k\}$ en $\mathcal{H}_0^2$ tales que

$$\lim_{k \rightarrow \infty} \|X^k - X\|_{L^2(P \times dt)} = 0.$$

- Todo proceso de $\mathcal{H}^2$ puede ser aproximado por procesos simples de $\mathcal{H}_0^2$.
- Usando la isometría de Itô se puede comprobar que la sucesión $I(X^k)$ es una sucesión de Cauchy en el espacio $L^2(P)$, en efecto,

$$\begin{aligned}\|I(X^k) - I(X^l)\|_{L^2(P)} &= \|I(X^k - X^l)\|_{L^2(P)} \\ &= \|X^k - X^l\|_{L^2(P \times dt)} \\ &\leq \|X - X^k\|_{L^2(P \times dt)} + \|X - X^l\|_{L^2(P \times dt)} \\ &\rightarrow 0.\end{aligned}$$

Definición

Sea X un proceso en $\mathcal{H}^2$ y sea $\{X^k\}_k$ una sucesión de procesos en $\mathcal{H}_0^2$ aproximante a X . Se define la integral estocástica de X como

$$I(X) = \lim_{k \rightarrow \infty} I(X^k),$$

en donde el límite debe entenderse dentro del espacio $L^2(P)$, es decir, se trata de convergencia en media cuadrática de una sucesión de variables aleatorias.

- Esto significa que la variable aleatoria $I(X)$ es un elemento de $L^2(P)$ y es tal que $\lim_{n \rightarrow \infty} \|I(X^k) - I(X)\|_{L^2(P)} = 0$.

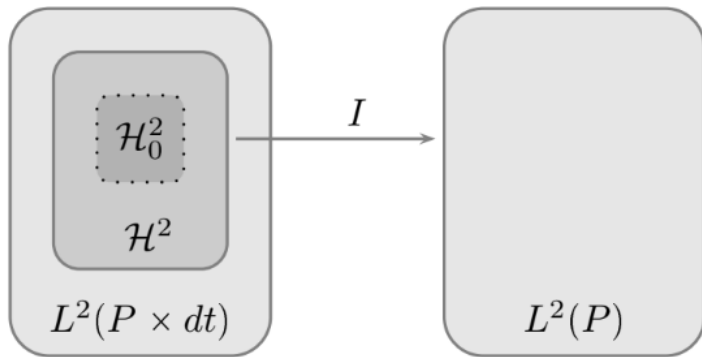


Figura: Definición de la integral de Itô I para el espacio $\mathcal{H}^2$. Figura tomada de Rincon, 2012.

- ★ La definición de integral de Itô del espacio $\mathcal{H}^2$ no depende de la sucesión que se seleccione en el espacio $\mathcal{H}_0^2$.
- ★ Para esta definición por aproximación también se cumple la isometría de Itô.
- ★ Para la integral de Itô en el espacio $\mathcal{H}^2$ se cumple la propiedad de la esperanza nula.
- ★ La función $I : \mathcal{H}^2 \rightarrow L^2(P)$ es una transformación lineal y continua.

Ejemplo. Observe que el movimiento Browniano $\{B_t : t \geq 0\}$ es un ejemplo de un proceso en el espacio $\mathcal{H}^2$ y es posible demostrar que tal proceso puede ser aproximado en el sentido de la norma del espacio $L^2(P \times dt)$ por el proceso simple

$$X^n = \{X_t^n : t \in [0, T]\} \text{ donde } X_t^n = \sum_{k=0}^{n-1} B_{t_k} 1_{[t_k, t_{k+1})}(t),$$

en donde $0 = t_0 < t_1 < \dots < t_n = T$ es una partición de $[0, T]$. Se tiene entonces que la siguiente integral estocástica particular, y su aproximación como límite en media cuadrática

$$\int_0^T B_t dB_t = \lim_{n \rightarrow \infty} \int_0^T X^n dB_t = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}),$$

en donde el límite debe entenderse en el sentido de que la distancia máxima entre dos puntos sucesivos de la partición tiende a cero.

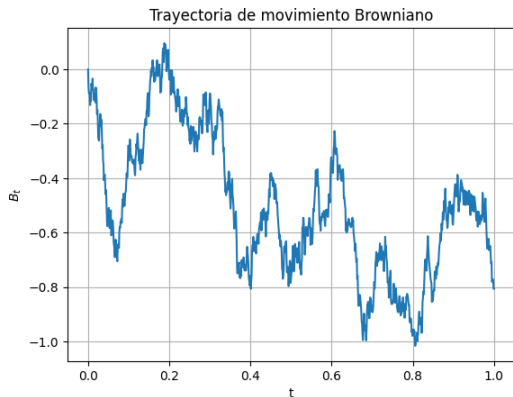


Figura: Una trayectoria del movimiento Browniano simulado desde una caminata aleatoria con $n = 1000$ pasos en el tiempo $t \in [0, 1]$.

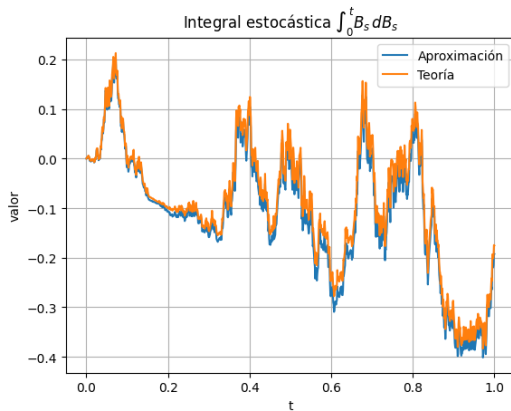


Figura: Integral estocástica del movimiento Browniano usando una partición uniforme del intervalo $[0, 1]$ y aproximando la integral por el límite.

La integral como proceso. Haremos una pequeña extensión. Para cada $t \in [0, T]$ y para cualquier X en $\mathcal{H}^2$ se define el proceso $\{I_t(X) : t \in [0, T]\}$ donde

$$I_t(X) = \int_0^T X_s 1_{[0,t]}(s) dB_s = \int_0^t X_s dB_s.$$

Este artificio permite ver la integral estocástica no como una variable aleatoria sino como un proceso estocástico.

Extensión por localización. Considere el espacio L^2_{loc} como el conjunto de todos los procesos medibles y adaptados que cumplen la condición de que

$$P \left(\int_0^T |X_t|^2 dt < \infty \right) = 1.$$

Ejemplo. Para el movimiento Browniano unidimensional $\{B_t : t \geq 0\}$ y para cualquier función continua f , el proceso $\{f(B_t) : t \geq 0\}$ tiene trayectorias continuas y acotadas, por lo tanto, este proceso pertenece a L^2_{loc} , y tiene sentido la expresión

$$\int_0^t f(B_s) dB_s, \quad 0 \leq t \leq T.$$

Se puede demostrar que esta integral puede ser calculada mediante el siguiente límite en el espacio $L^2(P)$, aunque también se verifica la convergencia en probabilidad:

$$\int_0^t f(B_s) dB_s = \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} f(B_{t_k})(B_{t_{k+1}} - B_{t_k}),$$

en donde $0 = t_0 < t_1 < \cdots < t_n = t$ es una partición del intervalo $[0, t]$, y nuevamente el límite se entiende como que la distancia máxima de dos puntos consecutivos de la partición tiende a cero.

Ejemplo. Calcularemos la siguiente integral con ayuda del límite del ejemplo anterior

$$\int_0^t B_s dB_s.$$

Consideremos $0 = t_0 < t_1 < \cdots < t_n = t$ una partición uniforme de $[0, t]$, es decir, $t_{k+1} - t_k = \frac{1}{n}$. Usando la identidad

$$a(b - a) = \frac{1}{2}(b^2 - a^2) - \frac{1}{2}(b - a)^2$$

se obtiene

$$\begin{aligned}\int_0^t B_s dB_s &= \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} B_{t_k} (B_{t_{k+1}} - B_{t_k}) \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} \left[\frac{1}{2} (B_{t_{k+1}}^2 - B_{t_k}^2) - \frac{1}{2} (B_{t_{k+1}} - B_{t_k})^2 \right] \\ &= \frac{1}{2} B_t^2 - \frac{1}{2} t.\end{aligned}$$

- La primera suma es telescópica mientras que la segunda suma corresponde a la variación cuadrática del movimiento Browniano.
- La esperanza de $\int_0^t B_s dB_s$ es nula y su varianza es igual a $\frac{1}{2}t^2$.

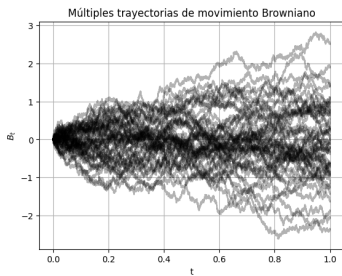


Figura: Simulación de 50 trayectorias del movimiento Browniano utilizando una aproximación de caminata aleatoria con $n = 100000$ pasos y $t \in [0, 1]$.

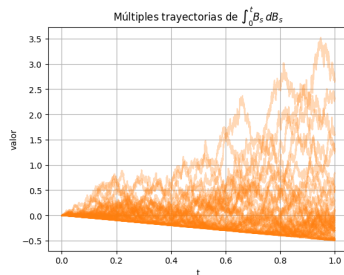


Figura: Integración estocástica teórica de las 50 trayectorias simuladas usando la expresión $\frac{1}{2} B_t^2 - \frac{1}{2} t$ (anteriormente demostrada).

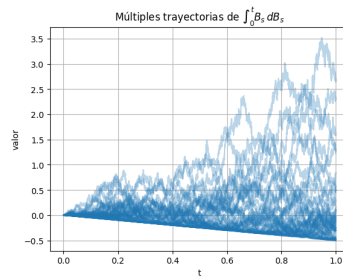


Figura: Integración estocástica aproximada de las 50 trayectorias simuladas usando aproximación al límite con una partición uniforme.

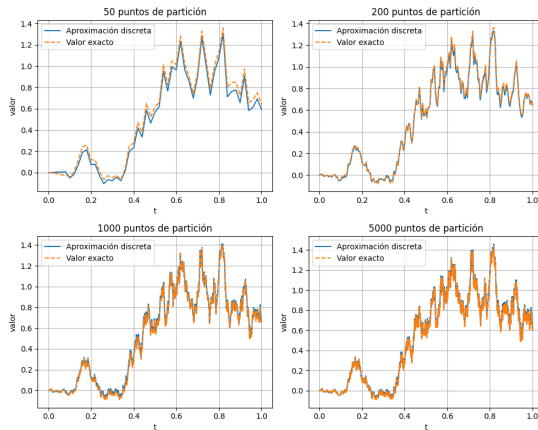


Figura: Aproximación de la integral estocástica a medida que aumenta el número de puntos de la partición. En esta figura se simula un movimiento Browniano con aproximación de caminata aleatoria con 10^6 pasos y se calcula la integral de Itô con diferentes tamaños de partición ($n \in \{50, 200, 1000, 5000\}$).

Ejercicio

Usando la identidad $ab = \frac{1}{2}(a+b)^2 - \frac{1}{2}(a-b)^2$ y la isometría de Itô, demuestre que para dos procesos $\{X_t : t \geq 0\}$ y $\{Y_t : t \geq 0\}$ en $\mathcal{H}^2$ se cumple

$$\mathbb{E} \left(\int_0^t X_s dB_s \int_0^t Y_s dB_s \right) = \int_0^t \mathbb{E}(X_s Y_s) dB_s.$$

Propiedades de la integral.

- La integral $I_t : L^2_{\text{loc}} \rightarrow L^2(P)$ es lineal, es decir, para una constante c y para cualesquiera procesos $\{X_s\}$ y $\{Y_s\}$ en L^2_{loc} , se cumple que

$$\int_0^t (cX_s + Y_s) dB_s \stackrel{\text{c.s.}}{=} c \int_0^t X_s dB_s + \int_0^t Y_s dB_s.$$

- Tiene esperanza cero, es decir, para cualquier proceso $\{X_s\}$ en L^2_{loc} ,

$$\mathbb{E} \left(\int_0^t X_s dB_s \right) \stackrel{\text{c.s.}}{=} 0.$$

- Cuando se restringe al espacio $\mathcal{H}^2$, se cumple la isometría de Itô, es decir,

$$\mathbb{E} \left(\left| \int_0^t X_s dB_s \right|^2 \right) = \mathbb{E} \left(\int_0^t |X_s|^2 ds \right).$$

- Cuando se restringe al espacio $\mathcal{H}^2$, la integral es una martingala, es decir, es integrable y adaptable y para $0 \leq s \leq t$, se satisface

$$\mathbb{E} \left(\int_0^t X_u dB_u \mid \mathcal{F}_s \right) = \int_0^s X_u dB_u.$$

Fórmula de Itô

Teorema

Sea una función de clase C^2 (con segunda derivada continua). Entonces

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) dB_s + \frac{1}{2} \int_0^t f''(B_s) ds$$

Ejercicio

- Calcule $\int_0^t B_s dB_s$ usando la función $f(x) = \frac{1}{2}x^2$ en la fórmula de Itô.
- Aplique a la función $f(x) = ax^2 + bx + c$ para constantes $a, b, c \in \mathbb{R}$ y obtenga una expresión general.
- Calcule la integral $\int_0^t B_s^n dB_s$.

Ecuaciones diferenciales estocásticas

Definición

Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sea $\{B_t : t \geq 0\}$ un movimiento Browniano unidimensional adaptado a la filtración $\{\mathcal{F}_t\}_{t \geq 0}$. Sean $b(t, x)$ y $\sigma(t, x)$ dos funciones de $[0, T] \times \mathbb{R}$ en $\mathbb{R}$. Una ecuación diferencial estocástica es una ecuación de la forma

$$(*) \quad dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t,$$

definida para valores de $t \in [0, T]$ y con condición inicial la variable aleatoria X_0 que se supone $\mathcal{F}_0$ -medible e independiente del movimiento Browniano. La ecuación $(*)$ se puede interpretar como la ecuación integral

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s,$$

en donde la primera es una integral de Riemann, mientras que la segunda es una integral de Itô. Al proceso $\{X_t\}_{t \geq 0}$ se le llama proceso de Itô.

Proposición (Existencia y unicidad)

Si los coeficientes de la ecuación $b(t, x)$ y $\sigma(t, x)$ de la ecuación (*) satisfacen la condición Lipschitz en la variable x

$$|b(t, x) - b(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K|x - y|^2$$

y la condición de crecimiento en x ,

$$|b(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2),$$

para alguna constante $K > 0$, entonces existe un proceso estocástico $\{X_t\}$ solución de (*) que es adaptado a la filtración, tiene trayectorias continuas, es uniformemente acotado $L^2(P)$, es decir, $\sup_{0 \leq t \leq T} \mathbb{E}(X_t)^2 < \infty$, y además es único en el sentido de la indistinguibilidad.

Proposición (Fórmula de Itô)

Si $\{X_t\}$ es un proceso de Itô dado por (*) y $f(t, x)$ es una función de clase C^1 en t y de clase C^2 en x , entonces el proceso $\{Y_t = f(t, X_t)\}$ es también un proceso de Itô y satisface la ecuación estocástica

$$dY_t = f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)(dX_t)^2.$$

- Teniendo en cuenta que $dt \times dt = dt \times dB_t = dB_t \times dt = 0$ y $dB_t \times dB_t = dt$ entonces usando la fórmula (*) se puede reemplazar en la fórmula dada en la anterior proposición:

$$\begin{aligned} dY_t = & \left(f_t(t, X_t) + f_x(t, X_t)b(t, X_t) + \frac{1}{2}f_{xx}(x, X_t)\sigma(t, X_t)^2 \right) dt \\ & + (f_x(t, X_t)\sigma(t, X_t)) dB_t. \end{aligned}$$

obteniendo una forma explícita del proceso de Itô.

Ejemplo. Considere el proceso $X_t := B_t$ y la función $f(t, x) = tx$. Entonces

$$\begin{aligned}d(f(t, B_t)) &= f_t(t, B_t)dt + f_x(t, B_t)dB_t + \frac{1}{2}f_{xx}(t, B_t)(dB_t)^2 \\d(tB_t) &= B_t dt + t dB_t,\end{aligned}$$

de esta forma diferencial podemos pasar a la forma integral

$$tB_t = \int_0^t B_s ds + \int_0^t s dB_s.$$

Ejercicio

1. ¿Es cierto que $B_t \stackrel{\text{c.s.}}{=} \int_0^t dB_s$ para $t > 0$?
2. (**Movimiento Browniano geométrico.**) Aplique la Fórmula de Itô a la función $f(x) = e^x$ y al proceso $X_t = B_t$ donde $\{B_t : t \geq 0\}$ es el movimiento Browniano unidimensional, luego, deduzca la forma integral y la forma diferencial del proceso.
3. Demuestre que el proceso $\{X_0 \stackrel{\text{c.s.}}{=} 0, X_t = \frac{B_t}{1+t} : t > 0\}$ es solución de la ecuación diferencial estocástica

$$dX_t = -\frac{X_t}{1+t} dt + \frac{1}{1+t} dB_t.$$

4. Usando el método de igualación de coeficientes resuelva la siguiente ecuación diferencial estocástica

$$dX_t = -X_t dt + e^{-t} dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} 0.$$

Simulación de ecuaciones diferenciales estocásticas

- Las ecuaciones diferenciales estocásticas (EDE) modelan sistemas con ruido aleatorio y aparecen en relevantes áreas de ingeniería, finanzas y física.
- En general no admiten solución explícita, por lo que se usan métodos numéricos adaptados del caso determinista.

- Para la EDE

$$dX_t = b(t, X_t) dt + \sigma(t, X_t) dB_t, \quad X_0 = x_0,$$

se simulan trayectorias mediante el método de Euler–Maruyama.

- Se discretiza $[0, t]$ en pasos $\Delta t = t/N$, $t_j = j\Delta t$, y se define

$$X_{t_{j+1}} = X_{t_j} + b(t_j, X_{t_j}) \Delta t + \sigma(t_j, X_{t_j}) \sqrt{\Delta t} Y_j,$$

donde $Y_j \sim \mathcal{N}(0, 1)$ i.i.d.

- Este esquema permite aproximar trayectorias del proceso solución y es fácil de implementar computacionalmente.

Ejemplo: Simulación del movimiento Browniano geométrico (GBM).

- Se considera el modelo estocástico

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 > 0,$$

cuyo resultado teórico es

$$X_t = X_0 \exp\left((\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\right).$$

- Se fija un horizonte temporal $[0, T]$ y se discretiza en N pasos de tamaño $\Delta t = T/N$.
- El movimiento Browniano se aproxima generando incrementos

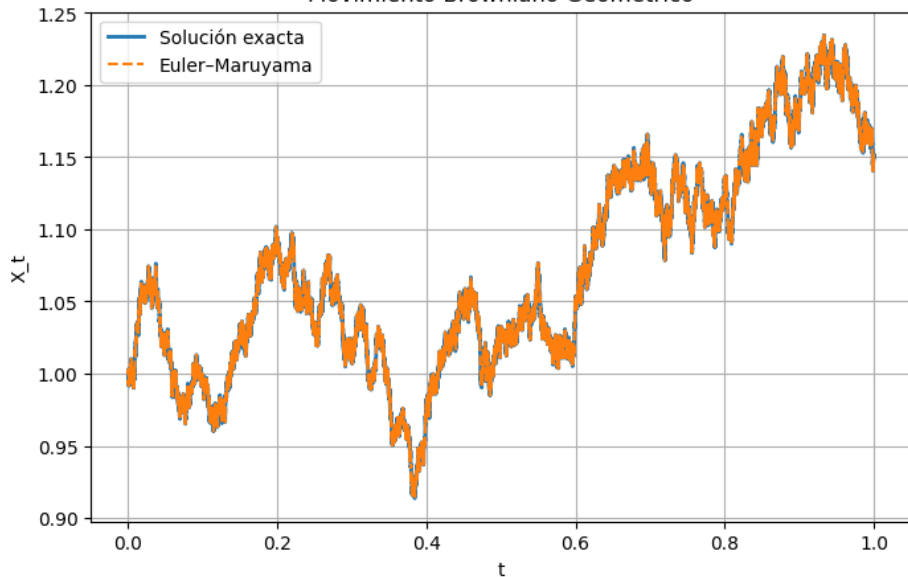
$$\Delta B_n = \sqrt{\Delta t} Y_n, \quad Y_n \sim \mathcal{N}(0, 1) \text{ i.i.d.}$$

- El esquema de Euler–Maruyama se implementa como

$$X_{n+1} = X_n + \mu X_n \Delta t + \sigma X_n \Delta B_n.$$

- Finalmente, se comparan gráficamente la trayectoria aproximada y la solución exacta, y se calcula el error absoluto al tiempo final T .

Movimiento Browniano Geométrico



Movimiento Browniano geométrico

Definición

Sean μ y σ dos constantes y $x_0 > 0$. El movimiento Browniano geométrico es el proceso $\{X_t : t \geq 0\}$ solución de la ecuación diferencial estocástica:

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} x_0,$$

y puede escribirse de manera explícita como:

$$X_t = x_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma B_t \right).$$

- En ausencia del término estocástico $X_t = x_0 \exp(\mu t)$. Esta función representa el comportamiento en el tiempo de un capital positivo x_0 que crece de manera continua y determinista a una tasa efectiva del $100\mu\%$. Por otro lado, la parte estocástica representa la volatilidad de los mercados financieros y supone que es proporcional al valor de la inversión.

Ejercicio

Obtenga la fórmula explícita de X_t por medio del método de igualación de coeficientes en la fórmula de Ito.

Proposición

Para el movimiento Browniano geométrico se cumple lo siguiente:

1. $E(X_t) = x_0 \exp(\mu t)$.
2. $\text{Var}(X_t) = x_0^2 \exp(2\mu t)(\exp(\sigma^2 t) - 1)$.
3. $\text{Cov}(X_t, X_s) = x_0^2 \exp(\mu(s + t))(\exp(\sigma^2 s) - 1)$ para $0 \leq s \leq t$.

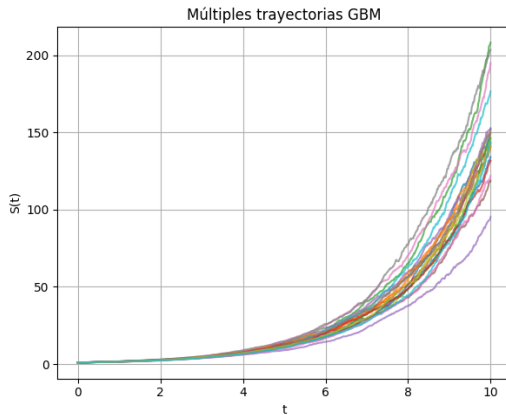
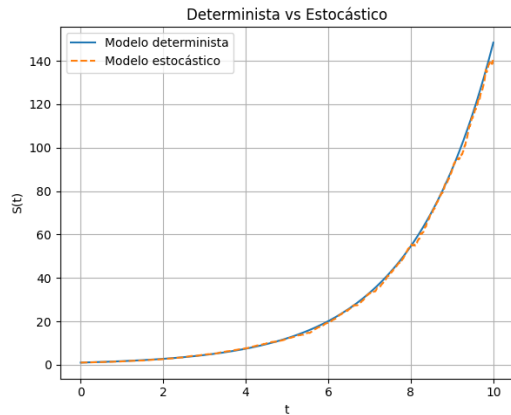


Figura: Movimiento Browniano geométrico con $\mu = 0,5$, $\sigma = 0,05$ y $x_0 = 1$.

Proceso de Ornstein-Uhlenbeck

Definición

Sean α y σ dos constantes positivas. El proceso de Ornstein-Uhlenbeck es aquel proceso $\{X_t : t \geq 0\}$ solución de la ecuación diferencial estocástica:

$$dX_t = -\alpha X_t dt + \sigma dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} x_0,$$

y dado por

$$X_t = x_0 \exp(-\alpha t) + \sigma \int_0^t \exp(-\alpha(t-s)) dB_s.$$

- La variable X_t se interpreta como la velocidad de una partícula en el tiempo t . La parte determinista corresponde a la fuerza de fricción y el sumando estocástico es el ruido aleatorio.

Demuestre que para el proceso de Ornstein-Uhlenbeck se cumple:

1. $E(X_t) = x_0 \exp(-\alpha t)$.
2. $\text{Var}(X_t) = \frac{\sigma^2}{2\alpha}(1 - \exp(-2\alpha t))$.
3. $\text{Cov}(X_t, X_s) = \frac{\sigma^2}{\alpha}(\exp(-\alpha(t-s)) - \exp(-\alpha(t+s)))$ para $0 \leq s \leq t$.

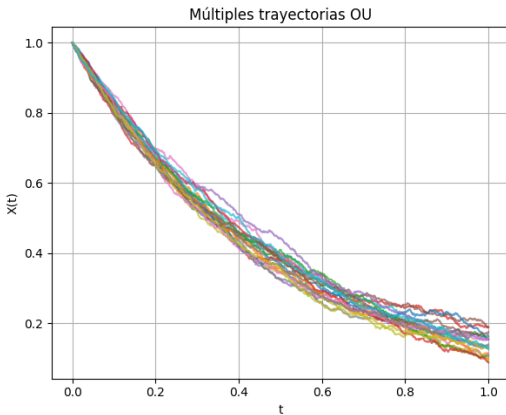
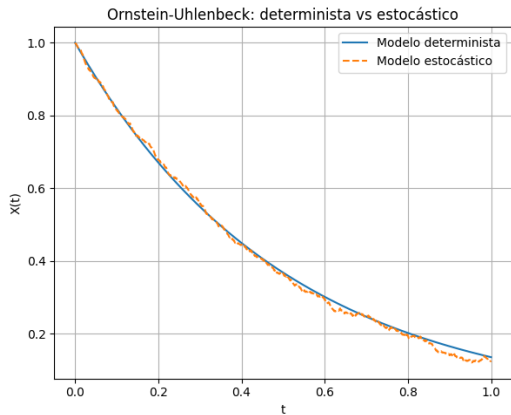


Figura: Proceso de Ornstein-Uhlenbeck con parámetros $x_0 = 1$, $\alpha = 2$ y $\sigma = 0,05$.

Puente Browniano

Definición

El puente Browniano en el intervalo $[0, 1]$ es aquel proceso $\{X_t : t \in [0, 1]\}$ solución de la ecuación diferencial estocástica

$$dX_t = -\frac{X_t}{1-t} dt + dB_t, t \in [0, 1), \quad X_0 \stackrel{\text{c.s.}}{=} 0,$$

y puede ser representado con

$$X_t = (1-t) \int_0^t \frac{1}{1-s} dB_s.$$

- El puente Browniano puede interpretarse como un movimiento Browniano con espacio parametral $[0, 1]$ y es tal que en los extremos de este intervalo el proceso se hace cero.

Proposición

Para el puente Browniano se cumple

1. $E(X_t) = 0$.
2. $\text{Var}(X_t) = t(1 - t)$.
3. $\text{Cov}(X_t, X_s) = s(1 - t)$ para $0 \leq s \leq t \leq 1$.

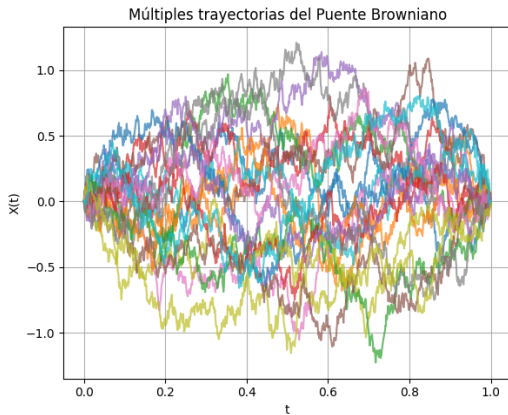
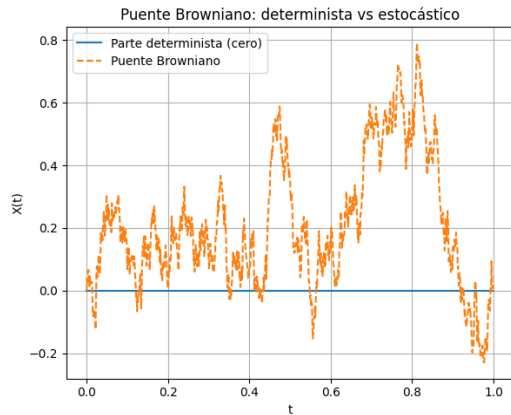


Figura: Puente Browniano.

Ejercicio

Demuestre que el proceso $X_t = \exp(B_t - t/2)$ satisface la ecuación diferencial estocástica

$$dX_t = X_t dB_t.$$

Ejercicio

Demuestre que cada una de las siguientes ecuaciones diferenciales estocásticas tiene única solución. En cada caso encuentre dicha solución. Los términos b , σ y x_0 son constantes.

1. $dX_t = -bX_t dt + \sigma dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} 0.$

2. $dX_t = -b dt + \sigma X_t dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} 0.$

3. $dX_t = -bX_t dt + \sigma X_t dB_t, \quad X_0 \stackrel{\text{c.s.}}{=} 0.$

Ejercicio

Demuestre que el proceso $X_t = B_t - tB_1$ para $t \in [0, 1]$ es un puente Browniano.

Referencias

Referencias



Rincon, L. (2012). *Introducción a procesos estocásticos*. UNAM.