

Procesos de ramificación y árboles aleatorios.

Juan Carlos Pardo Millán

CIMAT, Guanajuato

Programa:

- Preliminares de probabilidad.
- Cadenas de Markov.
- Caminatas aleatorias.
- Procesos de Bienaymé-Galton-Watson.

Preliminares de probabilidad

Consideremos a un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Preliminares de probabilidad

Consideremos a un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Recordemos la definición y algunas propiedades de la **probabilidad condicional**:

Preliminares de probabilidad

Consideremos a un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Recordemos la definición y algunas propiedades de la **probabilidad condicional**:

- i) Sean $A, B \in \mathcal{F}$ tal que $\mathbb{P}(B) > 0$, la probabilidad del evento A dado B esta definida por

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

Preliminares de probabilidad

Consideremos a un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$.

Recordemos la definición y algunas propiedades de la **probabilidad condicional**:

- i) Sean $A, B \in \mathcal{F}$ tal que $\mathbb{P}(B) > 0$, la probabilidad del evento A dado B esta definida por

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

- ii) Sea $(B_i, i \geq 1)$ una sucesión de elementos de \mathcal{F} tal que $\Omega = \cup_{i=1}^{\infty} B_i$ y $\mathbb{P}(B_i) > 0$ para toda i . Para toda $A \in \mathcal{F}$ tenemos

$$\mathbb{P}(A) = \sum_{i=1}^{\infty} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{\infty} \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

- iii) Sea $(B_i, 1 \leq i \leq n)$ una sucesión de elementos de \mathcal{F} tal que $\Omega = \cup_{i=1}^n B_i$ y $\mathbb{P}(B_i) > 0$. Para toda $A, B \in \mathcal{F}$ tales que $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, tenemos

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B \cap B_i)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B|B_i)\mathbb{P}(B_i)}.$$

- iii) Sea $(B_i, 1 \leq i \leq n)$ una sucesión de elementos de \mathcal{F} tal que $\Omega = \cup_{i=1}^n B_i$ y $\mathbb{P}(B_i) > 0$. Para toda $A, B \in \mathcal{F}$ tales que $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, tenemos

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B \cap B_i)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B|B_i)\mathbb{P}(B_i)}.$$

- iv) $A, B \in \mathcal{F}$ son independientes si

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

- iii) Sea $(B_i, 1 \leq i \leq n)$ una sucesión de elementos de \mathcal{F} tal que $\Omega = \cup_{i=1}^n B_i$ y $\mathbb{P}(B_i) > 0$. Para toda $A, B \in \mathcal{F}$ tales que $\mathbb{P}(A) > 0$ y $\mathbb{P}(B) > 0$, tenemos

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B \cap B_i)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\sum_{i=1}^n \mathbb{P}(B|B_i)\mathbb{P}(B_i)}.$$

- iv) $A, B \in \mathcal{F}$ son independientes si

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

- v) Sean \mathcal{A} y \mathcal{B} dos colecciones de subconjuntos de \mathcal{F} . Decimos que \mathcal{A} y \mathcal{B} son independientes si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \quad \forall A \in \mathcal{A} \quad \text{y} \quad \forall B \in \mathcal{B}.$$

Definición

Decimos que $X : \Omega \rightarrow E$ es una **variable aleatoria** si para todo $A \subset E$, tenemos

$$B := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}.$$

Definición

Decimos que $X : \Omega \rightarrow E$ es una **variable aleatoria** si para todo $A \subset E$, tenemos

$$B := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}.$$

Intuitivamente: $\omega \in \Omega$ lo vamos a considerar como un experimento y a $X(\omega)$ como una característica del experimento ω . Entonces B es el evento de los experimentos con características en el evento A . Para poder calcular la probabilidad de este evento es indispensable que viva en \mathcal{F} .

Definición

Decimos que $X : \Omega \rightarrow E$ es una **variable aleatoria** si para todo $A \subset E$, tenemos

$$B := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}.$$

Intuitivamente: $\omega \in \Omega$ lo vamos a considerar como un experimento y a $X(\omega)$ como una característica del experimento ω . Entonces B es el evento de los experimentos con características en el evento A . Para poder calcular la probabilidad de este evento es indispensable que viva en \mathcal{F} .

Dos variables aleatorias X, Y son independientes si para toda $A, B \in E$, los eventos

$$\{\omega \in \Omega : X(\omega) \in A\} \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \in B\}$$

son independientes.

Definición

Un *proceso estocástico discreto* es una colección de variables aleatorias $(X_n, n \geq 0)$ tales que

$$X_n : \Omega \rightarrow E, \quad \text{para toda } n \geq 0.$$

El conjunto E es llamado *espacio de estados*.

Definición

Un *proceso estocástico discreto* es una colección de variables aleatorias $(X_n, n \geq 0)$ tales que

$$X_n : \Omega \rightarrow E, \quad \text{para toda } n \geq 0.$$

El conjunto E es llamado *espacio de estados*.

En nuestro contexto E siempre será un conjunto numerable, para ser mas precisos vamos a considerar a $E = \mathbb{Z}$.

Definición

Un *proceso estocástico discreto* es una colección de variables aleatorias $(X_n, n \geq 0)$ tales que

$$X_n : \Omega \rightarrow E, \quad \text{para toda } n \geq 0.$$

El conjunto E es llamado *espacio de estados*.

En nuestro contexto E siempre será un conjunto numerable, para ser mas precisos vamos a considerar a $E = \mathbb{Z}$.

Estamos interesados en procesos estocásticos que cumplen una propiedad muy interesante: *el futuro es independiente del pasado dado el presente*. Esta propiedad es conocida como **la propiedad de Markov**.

Definición

Un *proceso estocástico discreto* es una colección de variables aleatorias $(X_n, n \geq 0)$ tales que

$$X_n : \Omega \rightarrow E, \quad \text{para toda } n \geq 0.$$

El conjunto E es llamado *espacio de estados*.

En nuestro contexto E siempre será un conjunto numerable, para ser mas precisos vamos a considerar a $E = \mathbb{Z}$.

Estamos interesados en procesos estocásticos que cumplen una propiedad muy interesante: *el futuro es independiente del pasado dado el presente*. Esta propiedad es conocida como **la propiedad de Markov**.

En otras palabras

$$\mathbb{P}(\{\text{futuro}\} \cap \{\text{pasado}\} | \{\text{presente}\}) = \mathbb{P}(\{\text{futuro}\} | \{\text{presente}\}) \mathbb{P}(\{\text{pasado}\} | \{\text{presente}\}).$$

Cadenas de Markov

Definición

Una *cadena de Markov* $X = (X_n, n \geq 0)$ es un proceso estocástico discreto con espacio de estados E numerable, que satisface la propiedad de Markov. En otras palabras, para toda $n \geq 1$ y toda sucesión $x_0, x_1, \dots, x_{n+1} \in E$ se tiene

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) \\ = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n).\end{aligned}$$

Cadenas de Markov

Definición

Una *cadena de Markov* $X = (X_n, n \geq 0)$ es un proceso estocástico discreto con espacio de estados E numerable, que satisface la propiedad de Markov. En otras palabras, para toda $n \geq 1$ y toda sucesión $x_0, x_1, \dots, x_{n+1} \in E$ se tiene

$$\begin{aligned}\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0) \\ = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n).\end{aligned}$$

La distribución X_0 se le llama *distribución inicial de la cadena* y la vamos a denotar por

$$\pi(x) = \mathbb{P}(X_0 = x), \quad x \in E.$$

La familia

$$\left\{ \mathbb{P}(X_{n+1} = y | X_n = x), n \geq 0, x, y \in E \right\},$$

se le llama la familia de probabilidades de transición de la cadena. Dicha familia describe la evolución de la cadena en el tiempo.

La familia

$$\left\{ \mathbb{P}(X_{n+1} = y | X_n = x), n \geq 0, x, y \in E \right\},$$

se le llama la familia de probabilidades de transición de la cadena. Dicha familia describe la evolución de la cadena en el tiempo.

Observemos que para todo $n \geq 0$ y $x \in E$ se tiene

$$\begin{aligned} 1 &= \mathbb{P}(X_{n+1} \in E | X_n = x) \\ &= \mathbb{P}\left(\bigcup_{y \in E} \{X_{n+1} = y\} \mid X_n = x \right) \\ &= \sum_{y \in E} \mathbb{P}(X_{n+1} = y | X_n = x). \end{aligned}$$

Si las probabilidades de ir de un estado a otro en una unidad de tiempo no depende del instante en el cual se encuentra la cadena, i.e.

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \mathbf{P}_{x,y} \quad \forall x, y \in E,$$

no depende de n diremos que la cadena es **homogénea** en el tiempo.

Si las probabilidades de ir de un estado a otro en una unidad de tiempo no depende del instante en el cual se encuentra la cadena, i.e.

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \mathbf{P}_{x,y} \quad \forall x, y \in E,$$

no depende de n diremos que la cadena es **homogénea** en el tiempo.

Nosotros solo estudiaremos cadenas de Markov homogéneas.

Si las probabilidades de ir de un estado a otro en una unidad de tiempo no depende del instante en el cual se encuentra la cadena, i.e.

$$\mathbb{P}(X_{n+1} = y | X_n = x) = \mathbf{P}_{x,y} \quad \forall x, y \in E,$$

no depende de n diremos que la cadena es **homogénea** en el tiempo.

Nosotros solo estudiaremos cadenas de Markov homogéneas.

La familia de probabilidades de transición

$$\mathbf{P} = (\mathbf{P}_{x,y})_{x,y \in E}$$

es llamada matriz de transición de la cadena. La matriz \mathbf{P} cumple

$$\sum_{y \in E} \mathbf{P}_{x,y} = 1, \quad \text{para } x \in E.$$

A \mathbf{P} generalmente se le conoce como matriz estocástica.

Diremos que el proceso estocástico $(X_n, n \geq 0)$ es una cadena de Markov (π, \mathbf{P}) , donde π es una distribución inicial y \mathbf{P} es la matriz de transición.

Diremos que el proceso estocástico $(X_n, n \geq 0)$ es una cadena de Markov (π, \mathbf{P}) , donde π es una distribución inicial y \mathbf{P} es la matriz de transición.

Ejemplo (La cadena de Ehrenfest) Ehrenfest describió un experimento con dos urnas U_1 y U_2 , dentro de las cuales están distribuidas N moléculas. En cada paso del experimento, vamos a escoger al azar una y solo una molécula. Dicha molécula es removida de la urna en la cual se encuentra y es colocada en la otra urna. El proceso estocástico $(X_n, n \geq 0)$ definido por

$X_n =$ no. de moléculas presentes en la urna U_1 al instante n ,

es una cadena de Markov. ¿Por qué?

Diremos que el proceso estocástico $(X_n, n \geq 0)$ es una cadena de Markov (π, \mathbf{P}) , donde π es una distribución inicial y \mathbf{P} es la matriz de transición.

Ejemplo (La cadena de Ehrenfest) Ehrenfest describió un experimento con dos urnas U_1 y U_2 , dentro de las cuales están distribuidas N moléculas. En cada paso del experimento, vamos a escoger al azar una y solo una molécula. Dicha molécula es removida de la urna en la cual se encuentra y es colocada en la otra urna. El proceso estocástico $(X_n, n \geq 0)$ definido por

$X_n =$ no. de moléculas presentes en la urna U_1 al instante n ,

es una cadena de Markov. ¿Por qué? Las probabilidades de transición estan dadas por

$$\mathbf{P}_{0,j} = \begin{cases} 1 & \text{si } j = 1, \\ 0 & \text{en otro caso} \end{cases}$$

$$\mathbf{P}_{N,j} = \begin{cases} 1 & \text{si } j = N - 1, \\ 0 & \text{en otro caso} \end{cases}$$

$$\mathbf{P}_{N,j} = \begin{cases} 1 & \text{si } j = N - 1, \\ 0 & \text{en otro caso} \end{cases}$$

y para $i \in \{1, 2, \dots, N - 1\}$

$$\mathbf{P}_{i,j} = \begin{cases} 0 & \text{si } |i - j| > 1, \\ i/N & \text{si } j = i - 1, \\ (N - i)/N & \text{si } j = i + 1, \\ 0 & \text{si } i = j. \end{cases}$$

$$\mathbf{P}_{N,j} = \begin{cases} 1 & \text{si } j = N - 1, \\ 0 & \text{en otro caso} \end{cases}$$

y para $i \in \{1, 2, \dots, N - 1\}$

$$\mathbf{P}_{i,j} = \begin{cases} 0 & \text{si } |i - j| > 1, \\ i/N & \text{si } j = i - 1, \\ (N - i)/N & \text{si } j = i + 1, \\ 0 & \text{si } i = j. \end{cases}$$

Para $x, y \in E$, se define $\mathbf{P}_{x,y}^{(0)}$ como

$$\mathbf{P}_{x,y}^{(0)} = \delta_x(y) = \begin{cases} 1 & \text{si } y = x, \\ 0 & \text{en otro caso.} \end{cases}$$

A la función δ se le conoce como la delta de Kronecker. A la matriz cuyas entradas están dadas por $\{\delta_x(y), x, y \in E\}$ se le denotará por I .

Para $m \geq 1$ y para $x, y \in E$ denotaremos por $\mathbf{P}_{x,y}^{(m)}$ a la probabilidad de ir al estado x al estado y en m tiempos, i.e

$$\mathbf{P}_{x,y}^{(m)} = \mathbb{P}(X_{n+m} = y | X_n = x), \quad \text{para } n \geq 0.$$

Para $m \geq 1$ y para $x, y \in E$ denotaremos por $\mathbf{P}_{x,y}^{(m)}$ a la probabilidad de ir al estado x al estado y en m tiempos, i.e

$$\mathbf{P}_{x,y}^{(m)} = \mathbb{P}(X_{n+m} = y | X_n = x), \quad \text{para } n \geq 0.$$

Observemos que

$$\begin{aligned} \mathbf{P}_{x,y}^{(m)} &= \mathbb{P}\left(\{X_{n+m} = y\} \cap \left\{\cup_{z \in E} \{X_{n+m-1} = z\}\right\} \middle| X_n = x\right) \\ &= \sum_{z \in E} \mathbb{P}(X_{n+m} = y, X_{n+1} = z | X_n = x) \\ &= \sum_{z \in E} \frac{\mathbb{P}(X_{n+m} = y, X_{n+1} = z, X_n = x)}{\mathbb{P}(X_n = x)} \\ &= \sum_{z \in E} \frac{\mathbb{P}(X_{n+m} = y, X_{n+1} = z, X_n = x)}{\mathbb{P}(X_{n+1} = z, X_n = x)} \frac{\mathbb{P}(X_{n+1} = z, X_n = x)}{\mathbb{P}(X_n = x)} \\ &= \sum_{z \in E} \mathbb{P}(X_{n+m} = y | X_{n+1} = z, X_n = x) \mathbb{P}(X_{n+1} = z | X_n = x). \end{aligned}$$

Por lo tanto, de la propiedad de Markov vemos

$$\begin{aligned}\mathbf{P}_{x,y}^{(m)} &= \sum_{z \in E} \mathbb{P}(X_{n+m} = y | X_{n+1} = z) \mathbb{P}(X_{n+1} = z | X_n = x) \\ &= \sum_{z \in E} \mathbf{P}_{z,y}^{(m-1)} \mathbf{P}_{x,z}^{(1)}.\end{aligned}$$

Por lo tanto, de la propiedad de Markov vemos

$$\begin{aligned} \mathbf{P}_{x,y}^{(m)} &= \sum_{z \in E} \mathbb{P}(X_{n+m} = y | X_{n+1} = z) \mathbb{P}(X_{n+1} = z | X_n = x) \\ &= \sum_{z \in E} \mathbf{P}_{z,y}^{(m-1)} \mathbf{P}_{x,z}^{(1)}. \end{aligned}$$

Proposición (Chapman-Kolmogorov)

Sea $(X_n, n \geq 0)$ una cadena de Markov (π, \mathbf{P}) . Entonces

$$\mathbf{P}_{x,y}^{(n)} = \sum_{z \in E} \mathbf{P}_{x,z}^{(s)} \mathbf{P}_{z,y}^{(r)},$$

donde r, s son dos enteros positivos cualesquiera tales que $r + s = n$. En particular, $\mathbf{P}^{(n)}$ es la n -ésima potencia de la matriz \mathbf{P} y

$$\mathbb{P}(X_n = y) = \sum_{x \in E} \pi(x) \mathbf{P}_{x,y}^{(n)}.$$

Caminatas aleatorias.

Las caminatas aleatorias es un ejemplo de una cadena de Markov y posiblemente es uno de los ejemplos mas estudiados.

Caminatas aleatorias.

Las caminatas aleatorias es un ejemplo de una cadena de Markov y posiblemente es uno de los ejemplos mas estudiados.

Sea $(Y_i, i \geq 1)$ una sucesión de variable aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) con valores en \mathbb{Z} . Consideremos a una v.a. S_0 independiente de $(Y_i, i \geq 1)$ y definamos al proceso estocástico $S = (S_n, n \geq 0)$ de manera recursiva como sigue

$$S_{n+1} = S_n + Y_{n+1}, \quad n \geq 0.$$

Caminatas aleatorias.

Las caminatas aleatorias es un ejemplo de una cadena de Markov y posiblemente es uno de los ejemplos mas estudiados.

Sea $(Y_i, i \geq 1)$ una sucesión de variable aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) con valores en \mathbb{Z} . Consideremos a una v.a. S_0 independiente de $(Y_i, i \geq 1)$ y definamos al proceso estocástico $S = (S_n, n \geq 0)$ de manera recursiva como sigue

$$S_{n+1} = S_n + Y_{n+1}, \quad n \geq 0.$$

El proceso estocástico S es llamado caminata aleatoria.

Observemos que es una cadena de Markov homogénea. Para ello consideremos

$$\begin{aligned}\mathbb{P}(S_n = y | S_{n-1} = x) &= \frac{\mathbb{P}(S_n = y, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} \\ &= \frac{\mathbb{P}(S_{n-1} + Y_n = y, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} \\ &= \frac{\mathbb{P}(Y_n = y - x, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} = \mathbb{P}(Y_1 = y - x),\end{aligned}$$

gracias a la independencia entre S_{n-1} y Y_n y a que esta última tiene la misma distribución que Y_1 .

Observemos que es una cadena de Markov homogénea. Para ello consideremos

$$\begin{aligned} \mathbb{P}(S_n = y | S_{n-1} = x) &= \frac{\mathbb{P}(S_n = y, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} \\ &= \frac{\mathbb{P}(S_{n-1} + Y_n = y, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} \\ &= \frac{\mathbb{P}(Y_n = y - x, S_{n-1} = x)}{\mathbb{P}(S_{n-1} = x)} = \mathbb{P}(Y_1 = y - x), \end{aligned}$$

gracias a la independencia entre S_{n-1} y Y_n y a que esta última tiene la misma distribución que Y_1 .

En otras palabras S es una cadena de Markov con distribución inicial, la ley de S_0 , y matriz de transición, la ley de Y_1 .

Ejemplos:

- El primer ejemplo que vamos a considerar es el caso en que $S_0 = 0$ y las v.a. Y_i 's se distribuyen como una v.a. Bernoulli, i.e.

$$\mathbb{P}(Y_1 = i) = \begin{cases} p & \text{si } i = 1, \\ 1 - p & \text{si } i = -1, \end{cases}$$

donde $p \in (0, 1)$. En este caso, no es muy complicado de ver que

$$\mathbf{P}_{0,0}^{(2n)} = \binom{2n}{n} p^n (1-p)^n \quad \text{y} \quad \mathbf{P}_{0,0}^{(2n-1)} = 0.$$

Intuitivamente esto nos dice que el punto 0 es visitado una infinidad de veces si $p = 1/2$.

Ejemplos:

- El primer ejemplo que vamos a considerar es el caso en que $S_0 = 0$ y las v.a. Y_i 's se distribuyen como una v.a. Bernoulli, i.e.

$$\mathbb{P}(Y_1 = i) = \begin{cases} p & \text{si } i = 1, \\ 1 - p & \text{si } i = -1, \end{cases}$$

donde $p \in (0, 1)$. En este caso, no es muy complicado de ver que

$$\mathbf{P}_{0,0}^{(2n)} = \binom{2n}{n} p^n (1-p)^n \quad \text{y} \quad \mathbf{P}_{0,0}^{(2n-1)} = 0.$$

Intuitivamente esto nos dice que el punto 0 es visitado una infinidad de veces si $p = 1/2$.

- El segundo ejemplo es el conocido como **skip-down free**. En este caso

$$\mathbb{P}(Y_1 \geq -1) = 1.$$

En otras palabras, la caminata tiene saltos positivos de magnitud mayor a 1 y saltos negativos de magnitud igual a 1.

- El último ejemplo que veremos es la caminata aleatoria con barreras absorbentes. Supongamos que

$$\mathbb{P}(Y_1 = i) = \begin{cases} p & \text{si } i = 1, \\ 1 - p & \text{si } i = -1, \end{cases}$$

y que 0 y N son barreras absorbentes, i.e. que cuando la caminata llega al estado 0 o N , ya no puede salir de ahí. Supongamos que la caminata inicia en $0 < x < N$. En este caso

$$\mathbf{P}_{i,j} = \begin{cases} 1 & \text{si } i, j = 0, \\ p & \text{si } 0 < i < N, j = i + 1 \\ 1 - p & \text{si } 0 < i < N, j = i - 1, \\ 1 & \text{si } i = j = N, \\ 0 & \text{en otro caso.} \end{cases}$$

Procesos de Bienaymé-Galton-Watson

Fue en 1874 cuando Francis Galton y H. W. Watson trabajaron en el problema de extinción de familias, mostrando como la probabilidad puede ser aplicada al estudio de los efectos del azar en el desarrollo de familias o poblaciones.

Procesos de Bienaymé-Galton-Watson

Fue en 1874 cuando Francis Galton y H. W. Watson trabajaron en el problema de extinción de familias, mostrando como la probabilidad puede ser aplicada al estudio de los efectos del azar en el desarrollo de familias o poblaciones.

Lamentablemente este primer trabajo tenía una hipótesis que no podía ser aceptada, ya que asumía que las familias distinguidas son más dadas a desaparecer que las ordinarias. Es así como Galton vuelve a plantear el problema de la siguiente manera:

Procesos de Bienaymé-Galton-Watson

Fue en 1874 cuando Francis Galton y H. W. Watson trabajaron en el problema de extinción de familias, mostrando como la probabilidad puede ser aplicada al estudio de los efectos del azar en el desarrollo de familias o poblaciones.

Lamentablemente este primer trabajo tenía una hipótesis que no podía ser aceptada, ya que asumía que las familias distinguidas son más dadas a desaparecer que las ordinarias. Es así como Galton vuelve a plantear el problema de la siguiente manera:

Sean p_0, p_1, p_2, \dots , las respectivas probabilidades de que un hombre tenga 0, 1, 2, \dots , hijos varones; donde cada hijo varón tiene las mismas probabilidades que sus padres de tener hijos varones; y así sucesivamente. ¿Cuál es la probabilidad de que la línea masculina se extinga después de r -generaciones? O de una manera más general ¿Cuál es dicha probabilidad para un número de descendientes en la línea masculina de una generación dada?

Este planteamiento da nacimiento al estudio de los procesos de ramificación.

Este planteamiento da nacimiento al estudio de los procesos de ramificación.

Estos trabajos fueron olvidados por varios años, pero fueron retomados extensivamente entre los años veinte y treinta del siglo pasado por su gran interés matemático y como una base teórica para estudios de poblaciones de individuos como serian: genes, neutrones o rayos cósmicos.

Este planteamiento da nacimiento al estudio de los procesos de ramificación.

Estos trabajos fueron olvidados por varios años, pero fueron retomados extensivamente entre los años veinte y treinta del siglo pasado por su gran interés matemático y como una base teórica para estudios de poblaciones de individuos como serian: genes, neutrones o rayos cósmicos.

La primera correcta y completa determinación de la probabilidad de extinción para dicho proceso fue dada por J. F. Steffenson (1930, 1932). Después el modelo fue retomado por Kolmogorov y Dimitriev (1938), los cuales aumentaron notablemente su estudio y le dieron el nombre de Procesos de Ramificación.

Este planteamiento da nacimiento al estudio de los procesos de ramificación.

Estos trabajos fueron olvidados por varios años, pero fueron retomados extensivamente entre los años veinte y treinta del siglo pasado por su gran interés matemático y como una base teórica para estudios de poblaciones de individuos como serian: genes, neutrones o rayos cósmicos.

La primera correcta y completa determinación de la probabilidad de extinción para dicho proceso fue dada por J. F. Steffenson (1930, 1932). Después el modelo fue retomado por Kolmogorov y Dimitriev (1938), los cuales aumentaron notablemente su estudio y le dieron el nombre de Procesos de Ramificación.

Aquí estudiaremos el proceso de ramificación más simple como una introducción al estudio de los procesos de ramificación más generales.

Definición.

Para empezar con el estudio de este modelo, es necesario imaginar individuos que puedan generar nuevos individuos del mismo tipo, por ejemplo: el hombre, las bacterias reproduciéndose o neutrones en una reacción en cadena.

Definición.

Para empezar con el estudio de este modelo, es necesario imaginar individuos que puedan generar nuevos individuos del mismo tipo, por ejemplo: el hombre, las bacterias reproduciéndose o neutrones en una reacción en cadena.

Primero vamos a empezar con un conjunto inicial de individuos, al cual llamaremos generación cero, si estos individuos tienen descendientes, estos nuevos integrantes van a conformar la primera generación y los descendientes de esta primera generación van a conformar a la segunda generación y así sucesivamente.

Definición.

Para empezar con el estudio de este modelo, es necesario imaginar individuos que puedan generar nuevos individuos del mismo tipo, por ejemplo: el hombre, las bacterias reproduciéndose o neutrones en una reacción en cadena.

Primero vamos a empezar con un conjunto inicial de individuos, al cual llamaremos generación cero, si estos individuos tienen descendientes, estos nuevos integrantes van a conformar la primera generación y los descendientes de esta primera generación van a conformar a la segunda generación y así sucesivamente.

Hay que hacer notar que seguiremos con atención el tamaño de las generaciones sucesivas, no el tiempo al cual cada individuo nace; así como tampoco la relación familiar de cada individuo.

Definición.

Para empezar con el estudio de este modelo, es necesario imaginar individuos que puedan generar nuevos individuos del mismo tipo, por ejemplo: el hombre, las bacterias reproduciéndose o neutrones en una reacción en cadena.

Primero vamos a empezar con un conjunto inicial de individuos, al cual llamaremos generación cero, si estos individuos tienen descendientes, estos nuevos integrantes van a conformar la primera generación y los descendientes de esta primera generación van a conformar a la segunda generación y así sucesivamente.

Hay que hacer notar que seguiremos con atención el tamaño de las generaciones sucesivas, no el tiempo al cual cada individuo nace; así como tampoco la relación familiar de cada individuo.

Vamos a denotar por Z_0, Z_1, Z_2, \dots , como el número de individuos de la generación cero, primera generación, segunda generación, ...

También vamos a agregar las siguientes dos hipótesis:

- i) Si el tamaño de la n -ésima generación es conocido, entonces la ley de probabilidades que gobierna a las generaciones siguientes, no depende del tamaño de las generaciones previas a la n -ésima; en otras palabras Z_0, Z_1, Z_2, \dots forma una cadena de Markov.
- ii) La cadena de Markov considerada aquí, tiene una propiedad especial, la cual consiste en que diferentes individuos no interfieren con los otros; en el sentido de que el número de descendientes de un individuo no depende de cuántos otros individuos están presentes.

También vamos a agregar las siguientes dos hipótesis:

- i) Si el tamaño de la n -ésima generación es conocido, entonces la ley de probabilidades que gobierna a las generaciones siguientes, no depende del tamaño de las generaciones previas a la n -ésima; en otras palabras Z_0, Z_1, Z_2, \dots forma una cadena de Markov.
- ii) La cadena de Markov considerada aquí, tiene una propiedad especial, la cual consiste en que diferentes individuos no interfieren con los otros; en el sentido de que el número de descendientes de un individuo no depende en cuantos otros individuos están presentes.

Sean Z_0, Z_1, Z_2, \dots , v.a. Interpretaremos a Z_n como el número de individuos en la n -ésima generación; durante todo este curso vamos a suponer que el proceso inicia con un solo individuo, en otras palabras tenemos que $Z_0 = 1$.

También vamos a agregar las siguientes dos hipótesis:

- i) Si el tamaño de la n -ésima generación es conocido, entonces la ley de probabilidades que gobierna a las generaciones siguientes, no depende del tamaño de las generaciones previas a la n -ésima; en otras palabras Z_0, Z_1, Z_2, \dots forma una cadena de Markov.
- ii) La cadena de Markov considerada aquí, tiene una propiedad especial, la cual consiste en que diferentes individuos no interfieren con los otros; en el sentido de que el número de descendientes de un individuo no depende en cuantos otros individuos están presentes.

Sean Z_0, Z_1, Z_2, \dots , v.a. Interpretaremos a Z_n como el número de individuos en la n -ésima generación; durante todo este curso vamos a suponer que el proceso inicia con un solo individuo, en otras palabras tenemos que $Z_0 = 1$.

Además vamos a definir $\{X_{i,n} : n \geq 0, i \geq 1\}$ una sucesión de v.a.i.i.d., las cuales representan el número de descendientes que puede tener el i -ésimo individuo de la n -ésima generación.

De esta manera vamos a escribir a Z_{n+1} como:

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

De esta manera vamos a escribir a Z_{n+1} como:

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

Por esta última relación vemos que si $Z_n = 0$, la suma siempre va a ser cero, en consecuencia vamos a afirmar que el cero es un estado absorbente.

De esta manera vamos a escribir a Z_{n+1} como:

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

Por esta última relación vemos que si $Z_n = 0$, la suma siempre va a ser cero, en consecuencia vamos a afirmar que el cero es un estado absorbente.

Denotaremos por \mathbb{P} como la medida de probabilidad del proceso. Entonces la función de densidad de $X_{i,n}$ esta dada por

$$\mathbb{P}(X_{i,n} = k) = \rho_k \quad \text{para toda } k \in \{0, 1, 2, 3, \dots\} \quad \text{y} \quad \sum_{k=0}^{\infty} \rho_k = 1,$$

donde ρ_k es la probabilidad de que un individuo de la n -ésima generación tenga k hijos.

Con toda esta información podemos obtener una probabilidad de transición para el proceso que denotaremos por \mathbf{P}_{ij} la cual representa la probabilidad de que en la $(n + 1)$ -ésima generación haya j individuos dado que en la generación pasada (n -ésima) la población total era de i individuos.

Con toda esta información podemos obtener una probabilidad de transición para el proceso que denotaremos por \mathbf{P}_{ij} la cual representa la probabilidad de que en la $(n + 1)$ -ésima generación haya j individuos dado que en la generación pasada (n -ésima) la población total era de i individuos.

La probabilidad de transición para toda $i, j, n \in \{0, 1, 2, 3, \dots\}$ esta dada por

$$\mathbf{P}_{ij} = \mathbb{P}\left(Z_{n+1} = j \mid Z_n = i\right) = \mathbb{P}\left(\sum_{k=1}^i X_{k,n} = j\right).$$

Con toda esta información podemos obtener una probabilidad de transición para el proceso que denotaremos por \mathbf{P}_{ij} la cual representa la probabilidad de que en la $(n + 1)$ -ésima generación haya j individuos dado que en la generación pasada (n -ésima) la población total era de i individuos.

La probabilidad de transición para toda $i, j, n \in \{0, 1, 2, 3, \dots\}$ esta dada por

$$\mathbf{P}_{ij} = \mathbb{P}\left(Z_{n+1} = j \mid Z_n = i\right) = \mathbb{P}\left(\sum_{k=1}^i X_{k,n} = j\right).$$

Definiremos a la función generadora para la variable aleatoria $X_{1,0}$, como

$$f(s) = \sum_{k=0}^{\infty} \rho_k s^k \quad \text{donde} \quad s \in [-1, 1] \subset \mathbb{R}.$$

De manera similar, la función generadora para la variable aleatoria Z_n se define como

$$f_n(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_n = k) s^k \quad \text{para toda } n \in \{0, 1, 2, 3, \dots\}.$$

De manera similar, la función generadora para la variable aleatoria Z_n se define como

$$f_n(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_n = k) s^k \quad \text{para toda } n \in \{0, 1, 2, 3, \dots\}.$$

Como $\mathbb{P}(Z_0 = 1) = 1$, entonces

$$f_0(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_0 = k) s^k = s$$

y podemos notar que la variable aleatoria Z_1 se distribuye igual que la variable aleatoria $X_{1,0}$ entonces

$$f_1(s) = \sum_{k=0}^{\infty} \mathbb{P}(Z_1 = k) s^k = f(s).$$

Calculemos la función generadora del proceso en la $(n + 1)$ -ésima generación, i.e. f_{n+1} .

Calculemos la función generadora del proceso en la $(n + 1)$ -ésima generación, i.e. f_{n+1} . Por definición tenemos

$$\begin{aligned}
 f_{n+1}(s) &= \sum_{k=0}^{\infty} \mathbb{P}(Z_{n+1} = k) s^k = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \mathbb{P}(Z_{n+1} = k | Z_n = j) \mathbb{P}(Z_n = j) s^k \\
 &= \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j) \sum_{k=0}^{\infty} \mathbb{P}(X_{1,n} + X_{2,n} + \cdots + X_{j,n} = k) s^k \\
 &= \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j) \mathbb{E} \left[s^{\left(\sum_{i=1}^j X_{i,n} \right)} \right] = \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j) \prod_{i=1}^j \mathbb{E}[s^{X_{i,n}}] \\
 &= \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j) (\mathbb{E}[s^{X_{1,0}}])^j = \sum_{j=0}^{\infty} \mathbb{P}(Z_n = j) [f(s)]^j \\
 &= f_n(f(s)).
 \end{aligned}$$

Ahora podemos iterar esta función y obtener el siguiente resultado

$$f_{n+1}(s) = f_n(f(s)) = f_{n-1}(f(f(s))) = f_{n-1}(f_2(s))$$

Ahora podemos iterar esta función y obtener el siguiente resultado

$$f_{n+1}(s) = f_n(f(s)) = f_{n-1}(f(f(s))) = f_{n-1}(f_2(s))$$

Siguiendo con este procedimiento por inducción obtenemos,

$$f_{n+1}(s) = f_{n-k}(f_{k+1}(s)) \quad \text{para cualquier } k \in \{0, 1, 2, \dots, n-1\}.$$

Ahora podemos iterar esta función y obtener el siguiente resultado

$$f_{n+1}(s) = f_n(f(s)) = f_{n-1}(f(f(s))) = f_{n-1}(f_2(s))$$

Siguiendo con este procedimiento por inducción obtenemos,

$$f_{n+1}(s) = f_{n-k}(f_{k+1}(s)) \quad \text{para cualquier } k \in \{0, 1, 2, \dots, n-1\}.$$

Ejercicio 1. Sean $\mu = \mathbb{E}[X_1]$ y $\sigma^2 = \text{Var}[X_1]$ y supongamos que son finitas. Entonces

$$\mathbb{E}[Z_n] = \mu^n \quad \text{y} \quad \text{Var}[Z_n] = \begin{cases} n\sigma^2 & \text{si } \mu = 1, \\ \sigma^2 \mu^{(n-1)} \frac{\mu^n - 1}{\mu - 1} & \text{si } \mu \neq 1. \end{cases}$$