



ZIENTZIA
ETA TEKNOLOGIA
FAKULTATEA
FACULTAD
DE CIENCIA
Y TECNOLOGÍA

50 URTE
AÑOS
1968 - 2018
Biba Zientzia!
Ciencia Viva

Procesos de ramificación: un caso particular

Trabajo Fin de Grado
Grado en Matemáticas

Jennifer González Blanco

Trabajo dirigido por
Ana María Valle Martín

Leioa, 22 de junio de 2022

Índice general

Introducción	v
1. Función generatriz de probabilidad	1
1.1. Preliminares	1
1.2. Función generatriz de probabilidad: definición y ejemplos . . .	8
1.3. Propiedades de la función generatriz	11
1.4. Sumas aleatorias: definición y función generatriz	16
2. Procesos de ramificación	19
2.1. Introducción: definición de proceso estocástico	19
2.2. Introducción histórica	20
2.3. Definición matemática de un proceso de ramificación	21
2.4. Distribución de Z_n	25
2.5. Momentos de Z_n	27
2.5.1. Esperanza de Z_n	27
2.5.2. Varianza de Z_n	28
2.6. Probabilidad de extinción	30
2.7. Ejemplo en la población americana	36
Bibliografía	41

Introducción

Aunque el término de procesos de ramificación apareció posteriormente, la idea surgió a mediados del siglo XIX. Se quería buscar solución a la extinción de apellidos de familias influyentes de la aristocracia inglesa. Francis Galton formuló el problema de la siguiente forma:

Sea N el número de adultos varones de una población. En cada generación, el a_0 por ciento tiene 0 hijos varones, el a_1 por ciento tiene 1 hijo varón, etcétera... Encontrar:

- (i) *la proporción de apellidos que se extinguirá en r generaciones,*
- (ii) *el número de personas, dada una generación, que tendrán un determinado apellido.*

Posteriormente, Henry Watson abordó el problema. De esta forma, surgieron los procesos de Galton-Watson con los que se concluyó, erróneamente, que todas las familias acabarían extintas.

El objetivo de este trabajo es realizar un estudio detallado de un tipo de proceso de ramificación: el proceso de Galton-Watson con espacio de estados discreto y tiempo discreto e ilustrar los conceptos y propiedades consideradas.

Entre las diferentes técnicas que se pueden considerar para analizar el comportamiento de estos procesos estocásticos, hemos elegido trabajar con la función generatriz de probabilidad. La definición de esta función se introduce en los estudios del Grado en Matemáticas, pero, no se realiza un análisis de sus propiedades. Por ello, se han introducido y demostrado aquellas propiedades de la función generatriz de probabilidad que permiten obtener características de los procesos de ramificación. Por ejemplo, deducir la distribución y momentos del tamaño de la población en un determinado instante de tiempo. Por otra parte, se consideran diferentes ejemplos con el objetivo de ilustrar los conceptos y mostrar la aplicación de las propiedades que se consideran a lo largo de la memoria.

El trabajo se ha estructurado en dos capítulos.

En el Capítulo 1, en primer lugar, recordamos la definición de algunos conceptos básicos necesarios en la memoria, como espacio de probabilidad,

variable aleatoria, distribución de probabilidad e independencia y demostramos algunas propiedades que se necesitarán posteriormente. A continuación, establecemos la noción de función generatriz de probabilidad y obtenemos dicha función en algunos casos particulares. También enunciamos y demostramos propiedades de la función generatriz de probabilidad útiles a la hora de estudiar los procesos de ramificación. Por último, definimos las sumas aleatorias de variables aleatorias y expresamos su función generatriz de probabilidad en términos de las funciones generatrices de probabilidad de las variables aleatorias que intervienen en dichas sumas.

En el Capítulo 2, en primer lugar, definimos el concepto de proceso estocástico y realizamos una breve introducción histórica de los procesos de ramificación. A continuación, definimos un caso particular de proceso de ramificación que es el que analizaremos. Consideramos simulaciones de algunos casos particulares de estos procesos de ramificación con el objetivo de ilustrar su comportamiento. Estas simulaciones se han realizado con RStudio. Posteriormente, analizamos la distribución de probabilidad de las variables aleatorias que representan el tamaño de la población en cada instante de tiempo y sus momentos, la esperanza y la varianza. Finalmente, introducimos el concepto de probabilidad de extinción de un proceso de ramificación. Demostramos propiedades útiles que permiten obtener este valor y las aplicamos al estudio de ejemplos concretos, relacionándolas con las simulaciones realizadas anteriormente. Por último, desarrollamos un ejemplo práctico aplicando los resultados destacados del capítulo.

La memoria se completa con una Bibliografía que contiene las referencias que se han utilizado para elaborarla. Hemos utilizado la norma APA en su formato.

Capítulo 1

Función generatriz de probabilidad

En este capítulo, introduciremos conceptos y propiedades de la Teoría de la probabilidad que serán útiles a la hora de estudiar o analizar en el siguiente capítulo un tipo de proceso estocástico, los procesos de ramificación.

Concretamente, en el apartado 1, recordaremos algunos conceptos y propiedades de la Teoría de la probabilidad que se utilizarán a lo largo de la memoria. En los siguientes apartados, se recordará la definición de función generatriz de probabilidad y se introducirán y analizarán propiedades de la función generatriz de probabilidad útiles para el estudio de los procesos de ramificación. A la hora de realizar las demostraciones, se ha intentado utilizar la definición de la función generatriz de probabilidad en lugar de expresar la función generatriz de probabilidad como una esperanza matemática y utilizar las propiedades de la esperanza.

Para la realización de este capítulo, nos hemos basado fundamentalmente en las referencias [2, 3, 4] de la bibliografía.

1.1. Preliminares

En este primer apartado, vamos a definir conceptos básicos y establecer algunas propiedades que utilizaremos a lo largo de la presente memoria.

Comenzamos recordando los conceptos de espacio de probabilidad, probabilidad condicionada, tribu generada, π -clase, variable aleatoria y distribución de probabilidad de una variable aleatoria.

Definición 1.1.1. Un espacio de probabilidad es una terna (Ω, \mathcal{F}, P) tal que Ω es un conjunto, \mathcal{F} es una familia de subconjuntos de Ω ($\mathcal{F} \subset \mathcal{P}(\Omega)$) que tiene estructura de tribu o σ -álgebra y P es una probabilidad sobre el espacio medible (Ω, \mathcal{F}) . Es decir, \mathcal{F} satisface:

- (i) $\Omega \in \mathcal{F}$.

- (ii) $\forall A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$. Es decir, la familia es cerrada para la complementación.
- (iii) $\forall A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$. Es decir, la familia es cerrada para uniones numerables.

y $P : \mathcal{F} \rightarrow [0, 1]$ es una función que satisface los siguientes axiomas:

- (i) $P(\Omega) = 1$.
- (ii) $\forall A_1, A_2, \dots \in \mathcal{F} : A_i \cap A_j = \emptyset, \forall i, j \in \{1, 2, \dots\}, i \neq j \Rightarrow P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, donde $\sum_{i=1}^{\infty} A_i$ denota la unión disjunta de los sucesos A_1, A_2, \dots

Definición 1.1.2. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces, dado un suceso $B \in \mathcal{F}$ tal que $P(B) > 0$, para cada suceso $A \in \mathcal{F}$ se define la probabilidad condicionada de A dado B como:

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

Definición 1.1.3. Sea Ω un conjunto y $\mathcal{C} \subset \mathcal{P}(\Omega)$. Entonces la tribu generada por \mathcal{C} es la mínima tribu que contiene a \mathcal{C} y se denota por $\sigma(\mathcal{C})$.

Observación 1.1.1. Se puede demostrar que esta tribu siempre existe.

Definición 1.1.4. Sea Ω un conjunto y $\mathcal{C} \subset \mathcal{P}(\Omega)$. Entonces se dice que \mathcal{C} es una π -clase si $\forall A, B \in \mathcal{C}$, entonces $A \cap B \in \mathcal{C}$. Es decir, \mathcal{C} es una familia cerrada para intersecciones finitas.

Definición 1.1.5. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces una variable aleatoria sobre (Ω, \mathcal{F}, P) es una aplicación $X : \Omega \rightarrow \mathbb{R}$ tal que

$$\forall B \in \beta \Rightarrow X^{-1}(B) = \{w \in \Omega : X(w) \in B\} = (X \in B) \in \mathcal{F},$$

siendo β la tribu generada por todos los abiertos de \mathbb{R} asociados a la topología usual, $\beta = \sigma(\tau_u)$. Es decir, es la menor tribu que contiene a todos los abiertos.

Observación 1.1.2. Notar que una variable aleatoria X sobre un espacio de probabilidad (Ω, \mathcal{F}, P) es una función \mathcal{F}/β -medible.

Definición 1.1.6. Sea X una variable aleatoria sobre el espacio de probabilidad (Ω, \mathcal{F}, P) . Entonces la distribución de probabilidad de X la denotamos mediante P_X y está definida de la siguiente forma:

$$P_X : \beta \rightarrow [0, 1]$$

$$B \mapsto P_X(B) := P(X^{-1}(B)) = P(\{w \in \Omega : X(w) \in B\}) = P(X \in B).$$

Esta función determina el comportamiento probabilístico de la variable aleatoria.

Definición 1.1.7. Dos variables aleatorias $X : (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}, \beta)$ e $Y : (\Omega', \mathcal{F}', P') \rightarrow (\mathbb{R}, \beta)$, con funciones de distribución de probabilidad P_X y P'_Y respectivamente, se dice que están idénticamente distribuidas si y solo si $P_X \equiv P'_Y$. Es decir, están idénticamente distribuidas si tienen la misma distribución de probabilidad.

En los procesos de ramificación considerados en el Capítulo 2 intervienen variables aleatorias que toman valores enteros no negativos. Estas variables aleatorias son variables aleatorias discretas cuya definición damos a continuación.

Definición 1.1.8. Sea X una variable aleatoria. Si el conjunto de valores que toma X ($Im(X)$) es un conjunto finito o infinito numerable, entonces se dice que X es una variable aleatoria discreta.

Si consideramos una variable aleatoria tal que el conjunto $Im(X)$ es finito o infinito numerable, entonces la siguiente propiedad nos proporciona una condición sencilla necesaria y suficiente para justificar que es una variable aleatoria.

Proposición 1.1.1. Si $Im(X)$ es un conjunto finito o infinito numerable, entonces X es una variable aleatoria sobre el espacio de probabilidad (Ω, \mathcal{F}, P) si y solo si $(X = k) \in \mathcal{F}, \forall k \in Im(X)$.

Demostración. Suponemos primero que X es una variable aleatoria sobre el espacio de probabilidad (Ω, \mathcal{F}, P) . Es decir, $X^{-1}(B) \in \mathcal{F}, \forall B \in \beta$.

Sea $k \in Im(X)$. Entonces como $\{k\}$ es un cerrado en (\mathbb{R}, τ_u) , $\{k\} \in \beta$. Por tanto, $X^{-1}(\{k\}) = (X = k) \in \mathcal{F}, \forall k \in Im(X)$.

Por otro lado, suponemos que $(X = k) \in \mathcal{F}, \forall k \in Im(X)$.

Sea $B \in \beta$. Como $\Omega = \cup_{k \in Im(X)} (X = k)$, entonces:

$$\begin{aligned} X^{-1}(B) &= \{w \in \Omega : X(w) \in B\} \cap (\cup_{k \in Im(X)} (X = k)) \\ &= \cup_{k \in Im(X)} (\{w \in \Omega : X(w) \in B\} \cap \{w \in \Omega : X(w) = k\}) \quad (1.1) \\ &= \cup_{k \in Im(X) \cap B} (X = k) \in \mathcal{F}, \end{aligned}$$

dado que $Im(X) \cap B \subset Im(X)$ es un conjunto finito o infinito numerable y \mathcal{F} , por ser tribu, es cerrada para uniones numerables. \square

Definición 1.1.9. Sea X una variable aleatoria discreta. Entonces se denomina función de masa (de probabilidad) de X a la siguiente aplicación:

$$\begin{aligned} p_X : \mathbb{R} &\rightarrow [0, 1] \\ x &\mapsto p_X(x) := P(X = x) = P(X^{-1}(\{x\})) = P_X(\{x\}). \end{aligned}$$

La función de masa de X satisface las siguientes propiedades:

- $\forall x \notin \text{Im}(X), p_X(x) = P(X = x) = P(\emptyset) = 0.$
- $\sum_{x \in \text{Im}(X)} p_X(x) = \sum_{x \in \text{Im}(X)} P(X = x) = P(\cup_{x \in \text{Im}(X)} (X = x)) = P(\Omega) = 1.$

Observamos que la función de masa de probabilidad determina la distribución de probabilidad de la variable aleatoria. En efecto, $\forall B \in \beta$:

$$\begin{aligned} P_X(B) &= P(X^{-1}(B)) = P(\cup_{k \in \text{Im}(X) \cap B} (X = k)) \\ &= \sum_{k \in \text{Im}(X) \cap B} P(X = k) = \sum_{k \in \text{Im}(X) \cap B} p_X(k). \end{aligned} \quad (1.2)$$

En la segunda igualdad se ha utilizado (1.1).

A continuación, vamos a definir la esperanza y la varianza que son valores que proporcionan información sobre el comportamiento de una variable aleatoria.

La definición de esperanza matemática (media o valor medio) de una variable aleatoria la daremos solamente para variables aleatorias discretas no negativas que son las variables aleatorias que intervienen en los procesos de ramificación que consideramos en el Capítulo 2.

Definición 1.1.10. Sea X una variable aleatoria discreta no negativa sobre el espacio de probabilidad (Ω, \mathcal{F}, P) . Entonces la esperanza matemática de X (media de X o valor medio de X) se denota mediante la expresión $E(X)$ (o μ) y se define de la siguiente forma:

$$E(X) := \sum_{x \in \text{Im}(X)} x \cdot P(X = x).$$

La esperanza nos da información sobre el valor promedio de la variable aleatoria.

Teorema 1.1.2. Si X es una variable aleatoria discreta no negativa y $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ es una función medible entonces:

$$E(g(X)) = \sum_{x \in \text{Im}(X)} g(x)P(X = x).$$

Demostración. Sabemos que $g(X)$ tiene imagen $g(\text{Im}(X))$. Entonces:

$$\begin{aligned} E(g(X)) &= \sum_{y \in g(\text{Im}(X))} yP(g(X) = y) = \sum_{y \in g(\text{Im}(X))} yP(X \in g^{-1}(y)) \\ &\stackrel{(1)}{=} \sum_{y \in g(\text{Im}(X))} y \sum_{x \in \text{Im}(X): g(x)=y} P(X = x) \stackrel{(2)}{=} \sum_{x \in \text{Im}(X)} g(x)P(X = x), \end{aligned}$$

donde (1) es cierta por (1.2) y (2) se obtiene reordenando los términos de la serie. \square

Definición 1.1.11. Sea X una variable aleatoria no negativa con esperanza finita $E(X) = \mu$. Entonces, se define la varianza de X como:

$$Var(X) = E((X - \mu)^2).$$

Su raíz cuadrada positiva se denomina desviación típica de X , $\sigma = +\sqrt{Var(X)}$.

La varianza nos da información sobre la dispersión de los valores que toma la variable aleatoria alrededor de su media.

Observación 1.1.3. Recordamos que utilizando las propiedades de la esperanza se puede deducir la siguiente expresión que proporciona una fórmula alternativa a la definición para el cálculo de la varianza:

$$Var(X) = E(X^2) - (E(X))^2.$$

Finalizamos este apartado recordando el concepto de independencia de sucesos, de tribus y de variables aleatorias. También mostraremos algunas propiedades que se utilizarán posteriormente.

Definición 1.1.12. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $\{A_i : i \in I\}$ una familia de sucesos. Es decir, $A_i \in \mathcal{F}, \forall i \in I$. Entonces los sucesos $\{A_i : i \in I\}$ son independientes si y solo si $\forall J \subset I$ tal que J es finito se cumple:

$$P(\cap_{i \in J} A_i) = \prod_{i \in J} P(A_i).$$

Es decir, la probabilidad de cualquier intersección finita de sucesos de la familia es igual al producto de las probabilidades.

Para definir el concepto de tribus independientes se requiere previamente definir el concepto de selección de sucesos.

Definición 1.1.13. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $\{\mathcal{F}_i : i \in I\}$ una familia de colecciones no vacías de sucesos. Es decir, $\mathcal{F}_i \subset \mathcal{F}, \forall i \in I$. Entonces una selección de $\{\mathcal{F}_i : i \in I\}$ es una familia de sucesos $\{F_i : i \in I\}$ tal que $F_i \in \mathcal{F}_i, \forall i \in I$.

Definición 1.1.14. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $\{\mathcal{F}_i : i \in I\}$ una familia de subtribus de \mathcal{F} . Entonces las subtribus $\{\mathcal{F}_i : i \in I\}$ son independientes si y solo si cada selección de sucesos $\{F_i : i \in I\}$ de $\{\mathcal{F}_i : i \in I\}$ está formada por sucesos independientes.

Para definir el concepto de variables aleatorias independientes, previamente se define el concepto de tribu generada por una variable aleatoria.

Definición 1.1.15. Dada X una variable aleatoria sobre el espacio de probabilidad (Ω, \mathcal{F}, P) , la tribu generada por X se denota por $\sigma(X)$ y se define de la siguiente forma:

$$\sigma(X) = \{X^{-1}(B) : B \in \beta\}.$$

Observación 1.1.4. $\sigma(X)$ es la menor tribu que hace que X sea una variable aleatoria. Es decir, una función $\sigma(X)/\beta$ -medible.

En efecto:

(i) $\sigma(X)$ es una tribu por satisfacer los axiomas de la definición de tribu:

- $\Omega \in \sigma(X)$ dado que $\Omega = X^{-1}(\mathbb{R})$ con $\mathbb{R} \in \beta$.
- Si $A \in \sigma(X)$, entonces $A = X^{-1}(B)$ para algún $B \in \beta$. Por lo tanto, $A^c = (X^{-1}(B))^c = X^{-1}(B^c)$. Como $B \in \beta$ y β es una tribu, entonces $B^c \in \beta$. En definitiva, $A^c \in \sigma(X)$.
- Sean $A_1, A_2, \dots, A_n, \dots \in \sigma(X)$. Entonces, $\forall n \geq 1$ existe $B_n \in \beta$ tal que $A_n = X^{-1}(B_n)$. Luego, $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} X^{-1}(B_n) = X^{-1}(\cup_{n=1}^{\infty} B_n)$. Como $B_n \in \beta, \forall n \geq 1$, y β es una tribu, entonces $\cup_{n=1}^{\infty} B_n \in \beta$. Así, $\cup_{n=1}^{\infty} A_n \in \sigma(X)$.

(ii) Sea $X : (\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ una variable aleatoria. Entonces $\forall B \in \beta$ tenemos que $X^{-1}(B) \in \mathcal{F}$. Por tanto, $\sigma(X) \subset \mathcal{F}$.

Definición 1.1.16. Sea $\{X_i : i \in I\}$ una familia de variables aleatorias sobre un espacio de probabilidad (Ω, \mathcal{F}, P) . Entonces las variables aleatorias X_i son independientes si y solo si las tribus $\sigma(X_i) = \{X_i^{-1}(B) : B \in \beta\}$ son independientes.

Teniendo en cuenta el siguiente teorema cuya demostración se puede ver en la página 55 de [2] y el lema que establecemos a continuación deducimos la siguiente proposición que resulta de gran utilidad a la hora de garantizar la independencia de variables aleatorias.

Teorema 1.1.3. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $\mathcal{C}_1, \mathcal{C}_2, \dots \subset \mathcal{P}(\Omega)$. Si $\mathcal{C}_1, \mathcal{C}_2, \dots$ son independientes y cada \mathcal{C}_i es una π -clase, entonces las tribus $\sigma(\mathcal{C}_1), \sigma(\mathcal{C}_2), \dots$ son independientes.

Lema 1.1.4. Sean X_1, \dots, X_n variables aleatorias sobre el espacio de probabilidad (Ω, \mathcal{F}, P) y $(X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ el vector aleatorio sobre (Ω, \mathcal{F}, P) (es decir, una función $\mathcal{F}/\sigma(\tau_u^n) = \beta^n$ -medible). Entonces:

$$\sigma(\cup_{i=1}^n \sigma(X_i)) = \sigma(X_1, \dots, X_n),$$

donde $\sigma(X_1, \dots, X_n)$ denota la tribu generada por el vector aleatorio (X_1, \dots, X_n) .

Proposición 1.1.5. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y

$$\begin{array}{cccc} X_{11} & X_{12} & \dots & X_{1n_1} \\ X_{21} & X_{22} & \dots & X_{2n_2} \\ \cdot & \cdot & \dots & \cdot \\ X_{k1} & X_{k2} & \dots & X_{kn_k} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \end{array}$$

variables aleatorias sobre (Ω, \mathcal{F}, P) independientes. Entonces las variables aleatorias:

$$\begin{aligned} Y_1 = f_1(X_{11}, X_{12}, \dots, X_{1n_1}) : \Omega &\xrightarrow{(X_{11}, X_{12}, \dots, X_{1n_1})} \mathbb{R}^{n_1} \xrightarrow{f_1} \mathbb{R} \\ Y_2 = f_2(X_{21}, X_{22}, \dots, X_{2n_2}) : \Omega &\xrightarrow{(X_{21}, X_{22}, \dots, X_{2n_2})} \mathbb{R}^{n_2} \xrightarrow{f_2} \mathbb{R} \\ &\dots \\ Y_k = f_k(X_{k1}, X_{k2}, \dots, X_{kn_k}) : \Omega &\xrightarrow{(X_{k1}, X_{k2}, \dots, X_{kn_k})} \mathbb{R}^{n_k} \xrightarrow{f_k} \mathbb{R} \\ &\dots \end{aligned}$$

siendo f_1, \dots, f_k, \dots funciones β^{n_1}/β -medible, \dots , β^{n_k}/β -medible, \dots , respectivamente, son independientes.

Demostración. Las variables aleatorias $\{X_{ij} : i \in \mathbb{N}, j \in \{1, \dots, n_i\}\}$ son independientes si y solo si las tribus $\{\sigma(X_{ij}) : i \in \mathbb{N}, j \in \{1, \dots, n_i\}\}$ son independientes.

Consideramos las tribus $\sigma(\cup_{j=1}^{n_i} \sigma(X_{ij})) = \sigma(X_{i1}, \dots, X_{in_i})$ para $i \in \mathbb{N}$. Estas tribus se pueden reescribir como $\sigma(\cup_{j=1}^{n_i} \sigma(X_{ij})) = \sigma(\mathcal{C}_i)$ con $\mathcal{C}_i = \{X_{i1}^{-1}(B_1) \cap \dots \cap X_{in_i}^{-1}(B_{n_i}) : B_1, \dots, B_{n_i} \in \beta\}$ para $i \in \mathbb{N}$.

Las familias $\{\mathcal{C}_i : i \in \mathbb{N}\}$ son π -clases. En efecto, sean $C_1 = X_{i1}^{-1}(B_1) \cap \dots \cap X_{in_i}^{-1}(B_{n_i})$, $C_2 = X_{i1}^{-1}(B'_1) \cap \dots \cap X_{in_i}^{-1}(B'_{n_i}) \in \mathcal{C}_i$, con $B_j, B'_j \in \beta$, $\forall j \in \{1, \dots, n_i\}$. Entonces:

$$\begin{aligned} C_1 \cap C_2 &= (X_{i1}^{-1}(B_1) \cap X_{i1}^{-1}(B'_1)) \cap \dots \cap (X_{in_i}^{-1}(B_{n_i}) \cap X_{in_i}^{-1}(B'_{n_i})) \\ &= X_{i1}^{-1}(B_1 \cap B'_1) \cap \dots \cap X_{in_i}^{-1}(B_{n_i} \cap B'_{n_i}) \in \mathcal{C}_i, \end{aligned}$$

dado que $B_j \cap B'_j \in \beta$, $\forall j \in \{1, \dots, n_i\}$, por ser β una tribu.

Las familias $\{\mathcal{C}_i : i \in \mathbb{N}\}$ son independientes. En efecto, sea cualquier selección de sucesos $\{C_i \in \mathcal{C}_i : i \in \mathbb{N}\}$, con $C_i = X_{i1}^{-1}(B_{i1}) \cap \dots \cap X_{in_i}^{-1}(B_{in_i})$, $B_{ij} \in \beta$, $\forall i \in \mathbb{N}, j \in \{1, \dots, n_i\}$. Entonces:

$$\begin{aligned} P(\cap_{i \in J} C_i) &= P(\cap_{i \in J} (\cap_{j=1}^{n_i} X_{ij}^{-1}(B_{ij}))) \stackrel{(1)}{=} \prod_{i \in J} \prod_{j=1}^{n_i} P(X_{ij}^{-1}(B_{ij})) \\ &\stackrel{(1)}{=} \prod_{i \in J} P(\cap_{j=1}^{n_i} X_{ij}^{-1}(B_{ij})) = \prod_{i \in J} P(C_i), \quad \forall J \subset \mathbb{N} \text{ finito,} \end{aligned}$$

siendo (1) cierta porque las tribus $\sigma(X_{ij})$, $i \in \mathbb{N}, j \in \{1, \dots, n_i\}$ son independientes.

Por tanto, por el Teorema 1.1.3, las tribus $\sigma(\cup_{j=1}^{n_i} \sigma(X_{ij})) = \sigma(X_{i1}, \dots, X_{in_i})$, $i \in \mathbb{N}$, son independientes.

Por otro lado, $\forall i \in \mathbb{N}$, $\sigma(Y_i) = \sigma(f_i(X_{i1}, \dots, X_{in_i})) \subset \sigma(X_{i1}, \dots, X_{in_i})$. En efecto, sea $B \in \beta$. Entonces:

$$(f_i(X_{i1}, \dots, X_{in_i}))^{-1}(B) = (X_{i1}, \dots, X_{in_i})^{-1}(f_i^{-1}(B)) \in \sigma(X_{i1}, \dots, X_{in_i}),$$

dato que como f_i es una función medible, $f_i^{-1}(B) \in \beta^{n_i}$.

Por tanto, toda selección de sucesos de las tribus $\sigma(Y_i)$, $i \in \mathbb{N}$, es una selección de sucesos de las tribus $\sigma(X_{i1}, \dots, X_{in_i})$, $i \in \mathbb{N}$, que son independientes. Luego dicha selección de sucesos está formada por sucesos independientes. En consecuencia, las variables aleatorias $\{Y_i : i \in \mathbb{N}\}$ son independientes. \square

1.2. Función generatriz de probabilidad: definición y ejemplos

En este apartado definimos la función generatriz de probabilidad de una variable aleatoria que toma valores enteros no negativos. Esta función va a ser una herramienta sencilla y muy útil a la hora de analizar el comportamiento de los procesos de ramificación.

Definición 1.2.1. Sea X una variable aleatoria que toma valores en $\{0, 1, \dots\}$, con función de masa de probabilidad determinada por $p_k = P(X = k)$, $k \in \{0, 1, 2, \dots\}$. Entonces, la función generatriz de probabilidad de X se denota por G_X o G y está definida mediante la siguiente expresión:

$$G_X(s) := \sum_{k=0}^{\infty} p_k s^k$$

para todo $s \in \mathbb{R}$ tal que $|s| < \rho_X$, siendo ρ_X el radio de convergencia de la serie.

Observación 1.2.1. Sea X una variable aleatoria con función generatriz de probabilidad $G_X(s) = \sum_{k=0}^{\infty} p_k s^k$. Si $|s| < 1$, como $\sum_{k=0}^{\infty} p_k = 1$, entonces

$$\left| \sum_{k=0}^{\infty} p_k s^k \right| \leq \sum_{k=0}^{\infty} p_k |s|^k < \sum_{k=0}^{\infty} p_k = 1.$$

Es decir, el radio de convergencia de la serie es siempre mayor o igual que 1 ($\rho_X \geq 1$).

Proposición 1.2.1. Sea X una variable aleatoria con función generatriz de probabilidad $G_X(s) = \sum_{k=0}^{\infty} p_k s^k$. Entonces, en el intervalo $[0, 1]$:

- (i) G_X es una función no decreciente.
- (ii) G_X es una función convexa.

Demostración. (i) Si $s \in [0, 1]$, $G'_X(s) = \sum_{k=1}^{\infty} k p_k s^{k-1} \geq 0$. Por tanto, G_X es una función no decreciente en $[0, 1]$.

- (ii) Si $s \in [0, 1]$, $G_X''(s) = \sum_{k=2}^{\infty} k(k-1)p_k s^{k-2} \geq 0$. Así, G_X es una función convexa en $[0, 1]$. □

A continuación vamos a calcular la función generatriz de probabilidad asociada a algunas distribuciones de probabilidad discretas habituales.

Ejemplo 1.2.1. (i) Sean $n \in \mathbb{N}$, $p \in (0, 1)$ y X una variable aleatoria tal que $Im(X) = \{0, 1, \dots, n\}$ y

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \forall k \in Im(X). \quad (1.3)$$

Es decir, X tiene *distribución binomial* de parámetros n y p ($X \sim Bin(n, p)$). Entonces:

$$\begin{aligned} G_X(s) &= \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} s^k = \sum_{k=0}^n \binom{n}{k} (ps)^k (1-p)^{n-k} \\ &\stackrel{(1)}{=} ((1-p) + ps)^n, \end{aligned}$$

siendo la igualdad (1) cierta por el desarrollo del binomio de Newton. Por tanto, $\rho_X = +\infty$.

- (ii) Sea $\lambda \in (0, \infty)$ y X una variable aleatoria tal que $Im(X) = \{0, 1, 2, \dots\}$ y

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \forall k \in Im(X). \quad (1.4)$$

Es decir, X tiene *distribución de Poisson* de parámetro λ ($X \sim P(\lambda)$). Entonces:

$$G_X(s) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} s^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{1}{k!} (\lambda s)^k \stackrel{(1)}{=} e^{-\lambda} e^{\lambda s} = e^{\lambda(s-1)},$$

siendo la igualdad (1) cierta por el desarrollo en serie de la función exponencial. Luego, $\rho_X = +\infty$.

- (iii) Sea $p \in (0, 1)$ y X una variable aleatoria tal que $Im(X) = \{0, 1, 2, \dots\}$ y

$$P(X = k) = p(1-p)^k, \forall k \in Im(X). \quad (1.5)$$

Es decir, X tiene *distribución geométrica* de parámetro p . Entonces:

$$G_X(s) = \sum_{k=0}^{\infty} p(1-p)^k s^k = p \sum_{k=0}^{\infty} ((1-p)s)^k \stackrel{(1)}{=} \frac{p}{1 - (1-p)s},$$

siendo (1) cierta por ser la suma de la serie geométrica de razón $(1-p)s$. Por tanto, se necesita que $|(1-p)s| < 1$ para la convergencia de la serie, es decir, $|s| < \frac{1}{1-p}$. Luego $\rho_X = \frac{1}{1-p} (> 1)$.

(iv) Sea $n \in \mathbb{N}$, $p \in (0, 1)$ y X una variable aleatoria tal que $Im(X) = \{0, 1, 2, \dots\}$ y

$$P(X = k) = \binom{n+k-1}{k} p^n (1-p)^k, \forall k \in Im(X). \quad (1.6)$$

Es decir, X tiene *distribución binomial negativa* de parámetros n y p . Entonces:

$$\begin{aligned} G_X(s) &= \sum_{k=0}^{\infty} \binom{n+k-1}{k} p^n (1-p)^k s^k \stackrel{(1)}{=} p^n \sum_{k=0}^{\infty} (-1)^k \binom{-n}{k} ((1-p)s)^k \\ &= p^n \sum_{k=0}^{\infty} \binom{-n}{k} (-(1-p)s)^k \stackrel{(2)}{=} \left(\frac{p}{1-(1-p)s} \right)^n, \end{aligned}$$

siendo (1) cierta dado que los coeficientes binomiales satisfacen la propiedad $\binom{n+k+1}{k} = (-1)^k \binom{-n}{k}$. Por otro lado, teniendo en cuenta el desarrollo del binomio de Newton generalizado $(1+x)^r = \sum_{k=0}^{\infty} \binom{r}{k} x^k$ si $|x| < 1$ y $r \in \mathbb{R}$, (2) es cierta si $|-(1-p)s| < 1$. Por tanto, $|s| < \frac{1}{1-p}$.

Luego, $\rho_X = \frac{1}{1-p} (> 1)$.

Observamos que en todos los ejemplos descritos el radio de convergencia es mayor que 1.

A continuación mostramos las representaciones gráficas de las funciones generatrices de probabilidad obtenidas para algunos valores fijados de los parámetros de las distribuciones binomial, Poisson, geométrica y binomial negativa.

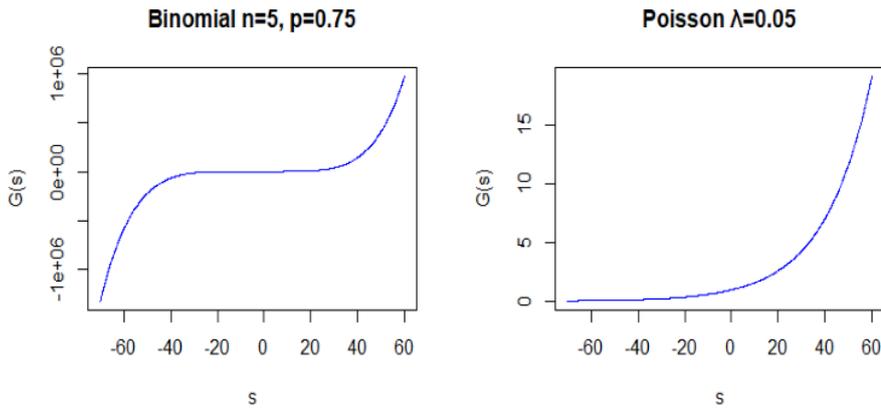


Figura 1.1: Funciones generatrices de probabilidad asociadas a las distribuciones binomial $n=2$, $p=0.75$ y de Poisson $\lambda=0.05$.

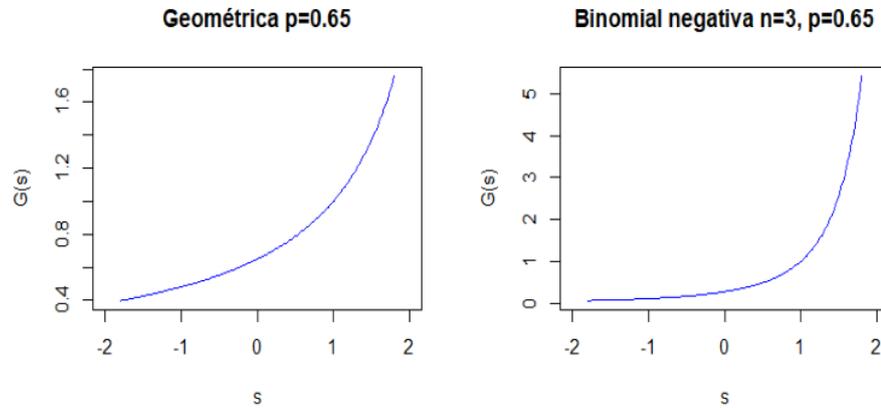


Figura 1.2: Funciones generatrices de probabilidad asociadas a las distribuciones geométrica $p=0.65$ y binomial negativa $n=3$, $p=0.65$.

Al realizar las gráficas en RStudio ([10]), hemos observado que la forma general de la función generatriz de probabilidad asociada a una variable aleatoria con distribución binomial dependía de si n era par o impar. Sin embargo, todos los ejemplos, incluidos estos dos casos, presentan el mismo patrón cuando s es positivo.

1.3. Propiedades de la función generatriz

Entre las propiedades de las funciones generatrices, en este apartado consideraremos las que necesitaremos a lo largo de la memoria.

En primer lugar, podemos observar que aplicando el Teorema 1.1.2 la función generatriz de probabilidad se puede expresar como la esperanza matemática de una variable aleatoria. En efecto, $G_X(s) = E(s^X)$. Esto permite deducir propiedades de la función generatriz de probabilidad a partir de propiedades de la esperanza. Sin embargo, en el presente trabajo se ha realizado un esfuerzo por demostrar estas propiedades utilizando la Definición 1.2.1.

Teorema 1.3.1. *Sean X e Y variables aleatorias discretas que toman valores en $\{0, 1, 2, \dots\}$ con funciones generatrices de probabilidad G_X y G_Y , respectivamente. Entonces son equivalentes las siguientes afirmaciones:*

- (i) $G_X(s) = G_Y(s)$ para todo s .
- (ii) $P(X = k) = P(Y = k)$ para $k = 0, 1, 2, \dots$

Es decir, dos variables aleatorias discretas que toman valores enteros no negativos tienen la misma función generatriz de probabilidad si y solo si tienen la misma función de masa de probabilidad.

Demostración. (ii) \Rightarrow (i) Evidente a partir de la Definición 1.2.1.

(i) \Rightarrow (ii) Sean $G_X(s) = \sum_{k=0}^{\infty} P(X = k)s^k$, para $|s| < \rho_X$ y $G_Y(s) = \sum_{k=0}^{\infty} P(Y = k)s^k$, para $|s| < \rho_Y$. En la Observación 1.2.1 hemos visto que los radios de convergencia de ambas series cumplen que $\rho_X, \rho_Y \geq 1$.

De esta forma, ambas funciones generatrices tienen expansiones únicas en serie de potencias alrededor del origen. Por lo tanto, si $G_X(s) = \sum_{k=0}^{\infty} P(X = k)s^k = G_Y(s) = \sum_{k=0}^{\infty} P(Y = k)s^k$, $|s| < 1$, sus coeficientes son iguales. \square

Así, deducimos que la función generatriz de probabilidad determina la distribución de probabilidad de una variable aleatoria.

A continuación, queremos estudiar la relación entre la función generatriz de la suma de dos variables aleatorias, G_{X+Y} , y sus respectivas funciones generatrices de probabilidad, G_X y G_Y .

Previamente, recordamos cómo realizar el producto de dos series de potencias.

Lema 1.3.2. Sean $\sum_{k=0}^{\infty} a_k s^k$ y $\sum_{l=0}^{\infty} b_l s^l$ dos series de potencias centradas en $s = 0$. Entonces:

$$\sum_{k=0}^{\infty} a_k s^k \cdot \sum_{l=0}^{\infty} b_l s^l = \sum_{n=0}^{\infty} c_n s^n,$$

con $c_n = \sum_{k=0}^n a_k b_{n-k}$, $\forall n \geq 0$.

Además, consideraremos la siguiente propiedad que expresa la distribución de probabilidad de $X + Y$ en términos de la de X y la de Y cuando X e Y son variables aleatorias independientes.

Proposición 1.3.3. Sean X e Y dos variables aleatorias independientes que toman valores en $\{0, 1, 2, \dots\}$. Entonces:

$$P(X + Y = k) = \sum_{l=0}^k P(X = l)P(Y = k - l), \quad \forall k \in \{0, 1, 2, \dots\}.$$

Demostración. Sean X e Y variables aleatorias independientes sobre el espacio de probabilidad (Ω, \mathcal{F}, P) . Entonces:

$$\begin{aligned} (X + Y = k) &= (X + Y = k) \cap \Omega = (X + Y = k) \cap \bigcup_{l=0}^{\infty} (X = l) \\ &= \bigcup_{l=0}^{\infty} (X + Y = k) \cap (X = l) = \bigcup_{l=0}^k (X = k - l, Y = l), \end{aligned}$$

siendo la última igualdad cierta porque X toma valores enteros no negativos y, por tanto, $k \geq l$.

Tomamos probabilidades a ambos lados de la igualdad, obtenemos:

$$\begin{aligned} P(X + Y = k) &= P(\cup_{l=0}^k (X = k - l, Y = l)) = \sum_{l=0}^k P(X = k - l, Y = l) \\ &= \sum_{l=0}^k P(X = k - l)P(Y = l), \end{aligned}$$

donde hemos utilizado que la unión de los sucesos $(X = k - l, Y = l)$ es disjunta y en la última igualdad que X e Y son variables aleatorias independientes. \square

Teorema 1.3.4. *Si X e Y son variables aleatorias independientes que toman valores en $\{0, 1, 2, \dots\}$, entonces:*

$$G_{X+Y}(s) = G_X(s)G_Y(s), \quad |s| < \min\{\rho_X, \rho_Y\}.$$

Demostración. Vamos a desarrollar la expresión para ver que se da la igualdad:

$$\begin{aligned} G_X(s)G_Y(s) &= \left(\sum_{k=0}^{\infty} P(X = k)s^k \right) \left(\sum_{k=0}^{\infty} P(Y = k)s^k \right) \\ &\stackrel{(1)}{=} \sum_{k=0}^{\infty} \sum_{l=0}^k P(X = l)P(Y = k - l)s^k \\ &\stackrel{(2)}{=} \sum_{k=0}^{\infty} P(X + Y = k)s^k \\ &= G_{X+Y}(s), \quad |s| < \min\{\rho_X, \rho_Y\}, \end{aligned}$$

siendo (1) cierto por el Lema 1.3.2 y (2) por la Proposición 1.3.3. \square

En el siguiente ejemplo, mostramos cómo utilizar esta propiedad para obtener fácilmente la distribución de probabilidad de una suma de dos variables aleatorias.

Ejemplo 1.3.1. Sean X e Y dos variables aleatorias independientes que toman valores enteros no negativos y tienen distribución de probabilidad de Poisson de parámetros λ y μ respectivamente, con $\lambda \in (0, +\infty)$ y $\mu \in (0, +\infty)$.

Entonces, por el Teorema 1.3.4, la función generatriz de probabilidad de la variable aleatoria $X + Y$ es:

$$G_{X+Y}(s) = G_X(s)G_Y(s) \stackrel{(1)}{=} e^{\lambda(s-1)}e^{\mu(s-1)} = e^{(\lambda+\mu)(s-1)}, \quad \rho_{X+Y} = +\infty,$$

siendo (1) cierta por el Ejemplo 1.2.1.

Además, deducimos, por el Teorema 1.3.1, que $X + Y$ es una variable aleatoria con distribución de probabilidad de Poisson de parámetro $\lambda + \mu$, con $\lambda + \mu \in (0, +\infty)$.

Por último, vamos a ver una propiedad que podremos utilizar para calcular la esperanza y la varianza de una variable aleatoria a partir de su función generatriz de probabilidad.

Teorema 1.3.5. *Si X es una variable aleatoria que toma valores enteros no negativos y tiene función generatriz de probabilidad $G(s)$, con radio de convergencia $\rho_X > 1$, entonces:*

$$E(X(X-1)\dots(X-r+1)) = G^{(r)}(1).$$

Demostración. Sabemos que dentro del intervalo de convergencia, la derivada de la serie es la serie de las derivadas. Además, el radio de convergencia de la serie de las derivadas sigue siendo el mismo. Entonces:

$$G'(s) = \frac{d}{ds} \sum_{k=0}^{\infty} s^k p_k = \sum_{k=0}^{\infty} \frac{d}{ds} s^k p_k = \sum_{k=1}^{\infty} k s^{k-1} p_k, |s| < \rho_X.$$

Así, como $\rho_X > 1$, podemos sustituir:

$$G'(1) = \sum_{k=1}^{\infty} k p_k = E(X).$$

Si seguimos calculando sus derivadas sucesivas, podemos escribir:

$$G^{(r)}(s) = \sum_{k=r}^{\infty} k(k-1)\dots(k-r+1) s^{k-r} p_k, |s| < \rho_X.$$

De nuevo, al sustituir, obtenemos:

$$G^{(r)}(1) = \sum_{k=r}^{\infty} k(k-1)\dots(k-r+1) p_k = E(X(X-1)\dots(X-r+1)).$$

□

Observación 1.3.1. Señalamos que si X toma valores enteros no negativos, entonces $E(X(X-1)\dots(X-r+1))$ siempre existe.

Corolario 1.3.6. *Si X es una variable aleatoria que toma valores enteros no negativos y tiene función generatriz de probabilidad $G(s)$, con radio de convergencia $\rho_X > 1$, entonces:*

- (i) $E(X) = G'(1)$.
- (ii) $Var(X) = G''(1) + G'(1) - G'(1)^2$.

Demostración. La primera afirmación es consecuencia directa del resultado del Teorema 1.3.5 tomando $r = 1$.

Veamos la segunda parte:

$$\begin{aligned} \text{Var}(X) &= E(X^2) - (E(X))^2 = E(X(X-1) + X) - (E(X))^2 \\ &= E(X(X-1)) + E(X) - (E(X))^2 = G''(1) + G'(1) - G'(1)^2. \end{aligned}$$

En la tercera igualdad se ha utilizado que la esperanza de la suma de dos variables aleatorias es la suma de las esperanzas. \square

Vamos a aplicar el resultado anterior para obtener la esperanza y la varianza de las variables aleatorias del Ejemplo 1.2.1.

Ejemplo 1.3.2. (i) Sea X una variable aleatoria con *distribución binomial* de parámetros $n \in \mathbb{N}$ y $p \in (0, 1)$ y función generatriz $G(s) = ((1-p) + ps)^n$, $\rho_X = +\infty$.

Primero, calculamos explícitamente $G'(s)$ y $G''(s)$:

$$G'(s) = np((1-p) + ps)^{n-1} \quad \text{y} \quad G''(s) = n(n-1)p^2((1-p) + ps)^{n-2}.$$

Entonces:

$$\begin{aligned} E(X) &= np((1-p) + p) = np \quad \text{y} \\ \text{Var}(X) &= n(n-1)p^2((1-p) + p) + np - (np)^2 \\ &= (np)^2 - np^2 + np - (np)^2 = np(1-p). \end{aligned}$$

(ii) Sea X una variable aleatoria con *distribución de Poisson* de parámetro $\lambda \in (0, \infty)$ y función generatriz $G(s) = e^{\lambda(s-1)}$, $\rho_X = +\infty$.

Primero, calculamos explícitamente $G'(s)$ y $G''(s)$:

$$G'(s) = \lambda e^{\lambda(s-1)} \quad \text{y} \quad G''(s) = \lambda^2 e^{\lambda(s-1)}.$$

Entonces:

$$\begin{aligned} E(X) &= \lambda e^{\lambda(1-1)} = \lambda \quad \text{y} \\ \text{Var}(X) &= \lambda^2 e^{\lambda(1-1)} + \lambda - \lambda^2 = \lambda. \end{aligned}$$

(iii) Sea X una variable aleatoria con *distribución geométrica* de parámetro $p \in (0, 1)$ y función generatriz $G(s) = \frac{p}{1-(1-p)s}$, $\rho_X = \frac{1}{1-p} > 1$.

Primero, calculamos explícitamente $G'(s)$ y $G''(s)$:

$$\begin{aligned} G'(s) &= \frac{-p(-(1-p))}{(1-(1-p)s)^2} = \frac{p(1-p)}{(1-(1-p)s)^2} \quad \text{y} \\ G''(s) &= \frac{-p(1-p) \cdot 2(1-(1-p)s)(-(1-p))}{(1-(1-p)s)^4} = \frac{2p(1-p)^2}{(1-(1-p)s)^3}. \end{aligned}$$

Entonces:

$$E(X) = \frac{p(1-p)}{p^2} = \frac{1-p}{p} \quad y$$

$$Var(X) = \frac{2p(1-p)^2}{p^3} + \frac{1-p}{p} - \left(\frac{1-p}{p}\right)^2 = \frac{(1-p)^2}{p^2} + \frac{(1-p)}{p} = \frac{1-p}{p^2}.$$

(iv) Sea X una variable aleatoria con *distribución binomial negativa* de parámetros $n \in \mathbb{N}$ y $p \in (0, 1)$ y función generatriz $G(s) = \left(\frac{p}{1-(1-p)s}\right)^n$, $\rho_X = \frac{1}{1-p} > 1$.

Primero, calculamos explícitamente $G'(s)$ y $G''(s)$:

$$G'(s) = n \left(\frac{p}{1-(1-p)s}\right)^{n-1} \frac{p(1-p)}{(1-(1-p)s)^2} = \frac{np^n(1-p)}{(1-(1-p)s)^{n+1}} \quad y$$

$$G''(s) = \frac{(n+1)np^n(1-p)^2(1-(1-p)s)^n}{(1-(1-p)s)^{2n+2}} = \frac{(n+1)np^n(1-p)^2}{(1-(1-p)s)^{n+2}}.$$

Entonces:

$$E(X) = \frac{np^n(1-p)}{p^{n+1}} = \frac{n(1-p)}{p} \quad y$$

$$\begin{aligned} Var(X) &= \frac{(n+1)np^n(1-p)^2}{p^{n+2}} + \frac{n(1-p)}{p} - \left(\frac{n(1-p)}{p}\right)^2 \\ &= \frac{(n+1)n(1-p)^2}{p^2} + \frac{np(1-p)}{p^2} - \frac{n^2(1-p)^2}{p^2} = \frac{n(1-p)}{p^2}. \end{aligned}$$

1.4. Sumas aleatorias: definición y función generatriz

Definición 1.4.1. Sean N, X_1, X_2, \dots variables aleatorias que toman valores en $\{0, 1, 2, \dots\}$, definidas sobre el mismo espacio de probabilidad (Ω, \mathcal{F}, P) . Entonces definimos la función Y que representa la suma aleatoria de las variables $X_1, X_2, \dots, X_n, \dots$ como la siguiente aplicación:

$$Y := \sum_{l=1}^N X_l : \Omega \rightarrow \mathbb{R}$$

$$w \mapsto Y(w) = \begin{cases} 0 & , \text{ si } N(w) = 0 \\ X_1(w) & , \text{ si } N(w) = 1 \\ \vdots & \\ X_1(w) + \dots + X_n(w) & , \text{ si } N(w) = n \\ \vdots & \end{cases}$$

Podemos comprobar que Y realmente es una variable aleatoria sobre el espacio medible (Ω, \mathcal{F}) .

Sea $B \in \beta$. Entonces:

$$\begin{aligned} (Y \in B) &= (Y \in B) \cap \Omega = (Y \in B) \cap \cup_{l=0}^{\infty} (N = l) = \cup_{l=0}^{\infty} (Y \in B, N = l) \\ &= ((Y \in B, N = 0)) \cup (\cup_{l=1}^{\infty} (Y \in B, N = l)) \\ &= ((0 \in B) \cap (N = 0)) \cup (\cup_{l=1}^{\infty} (X_1 + \dots + X_l \in B) \cap (N = l)). \end{aligned}$$

Ahora, $(N = k) \in \mathcal{F}$, $\forall k \in \{0, 1, \dots\}$, por ser N una variable aleatoria sobre (Ω, \mathcal{F}) . Además, $(0 \in B) = \emptyset$ si $0 \notin B$ y $(0 \in B) = \Omega$ si $0 \in B$, por lo que como \mathcal{F} es una tribu, tenemos $(0 \in B) \in \mathcal{F}$. Finalmente, como $X_1, X_2, \dots, X_n, \dots$ son variables aleatorias sobre (Ω, \mathcal{F}) , $X_1 + \dots + X_l$ también lo es $\forall l \in \{1, 2, \dots\}$. Por lo tanto $(X_1 + \dots + X_l \in B) \in \mathcal{F}$.

En definitiva, tenemos intersecciones y uniones finitas o numerables de elementos de \mathcal{F} , que es una tribu. Luego deducimos que:

$$(Y \in B) \in \mathcal{F}.$$

Es decir, Y es una variable aleatoria sobre el espacio medible (Ω, \mathcal{F}) .

Teorema 1.4.1. *Sea N una variable aleatoria que toma valores enteros no negativos con función generatriz de probabilidad G_N , X_1, X_2, \dots una sucesión de variables aleatorias que toman valores enteros no negativos idénticamente distribuidas y con función generatriz de probabilidad G_X . Además, N, X_1, X_2, \dots son variables aleatorias independientes. Entonces:*

$$Y = X_1 + X_2 + \dots + X_N$$

tiene función generatriz de probabilidad

$$G_Y(s) = G_N(G_X(s)), \quad |G_X(s)| < \rho_N,$$

siendo ρ_N el radio de convergencia de la función generatriz de probabilidad G_N .

Demostración. Calculamos

$$\begin{aligned} G_Y(s) &= \sum_{k=0}^{\infty} P(Y = k) s^k = \sum_{k=0}^{\infty} P((Y = k) \cap \Omega) s^k \\ &= \sum_{k=0}^{\infty} P((Y = k) \cap \cup_{l=0}^{\infty} (N = l)) s^k = \sum_{k=0}^{\infty} P\left(\bigcup_{l=0}^{\infty} (Y = k, N = l)\right) s^k \\ &= \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} P(Y = k, N = l) s^k \\ &= \sum_{k=0}^{\infty} P(Y = k, N = 0) s^k + \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} P(Y = k, N = l) s^k \end{aligned}$$

$$\begin{aligned}
&= P(Y = 0, N = 0) + \sum_{k=0}^{\infty} \sum_{l=1}^{\infty} P(X_1 + \cdots + X_l = k, N = l) s^k \\
&\stackrel{(1)}{=} P(N = 0) + \sum_{l=1}^{\infty} P(N = l) \sum_{k=0}^{\infty} P(X_1 + \cdots + X_l = k) s^k \\
&= P(N = 0) + \sum_{l=1}^{\infty} P(N = l) G_{X_1 + \cdots + X_l}(s) \\
&\stackrel{(2)}{=} P(N = 0) + \sum_{l=1}^{\infty} P(N = l) G_{X_1}(s) \cdots G_{X_l}(s) \\
&\stackrel{(3)}{=} P(N = 0) + \sum_{l=1}^{\infty} P(N = l) G_X(s)^l = \sum_{l=0}^{\infty} P(N = l) G_X(s)^l \\
&= G_N(G_X(s))
\end{aligned}$$

donde (1) es cierto por la Proposición 1.1.5, que garantiza que N y $X_1 + \cdots + X_l$ son variables aleatorias independientes. (2) es cierto por el Teorema 1.3.4 y (3) por el Teorema 1.3.1. \square

Ejemplo 1.4.1. Sea N una variable aleatoria con distribución de probabilidad de Poisson de parámetro λ , con $\lambda \in (0, +\infty)$. Entonces, su función generatriz de probabilidad es:

$$G_N(s) = e^{\lambda(s-1)}, \quad \rho_N = +\infty.$$

Sea X_1, X_2, \dots una sucesión de variables aleatorias idénticamente distribuidas con distribución probabilidad binomial de parámetros 1, p , con $p \in (0, 1)$. Entonces, su función generatriz de probabilidad, G_X , es:

$$G_X(s) = (1 - p) + ps, \quad \rho_X = +\infty.$$

Además, N, X_1, X_2, \dots son variables aleatorias independientes.

Consideramos la variable aleatoria $Y = X_1 + \cdots + X_N$. Vamos a calcular su función generatriz de probabilidad utilizando el Teorema 1.4.1:

$$\begin{aligned}
G_Y(s) &= G_N(G_X(s)) = e^{\lambda((1-p)+ps-1)} = e^{\lambda p(s-1)} \\
&= e^{\lambda p(s-1)}, \quad |(1-p) + ps| < +\infty.
\end{aligned}$$

Por tanto, deducimos, por el Teorema 1.3.1, que Y tiene distribución de probabilidad de Poisson de parámetro λp , con $\lambda p \in (0, +\infty)$. Así, de una forma sencilla, hemos deducido la distribución de probabilidad de una suma aleatoria de variables aleatorias independientes.

Capítulo 2

Procesos de ramificación

En este capítulo, estudiamos un caso particular de proceso de ramificación, el proceso de Galton-Watson [8]. Este proceso proporciona un modelo matemático que representa la evolución de una población a lo largo del tiempo bajo determinadas condiciones.

Para la realización de este capítulo, nos hemos basado en las referencias [1, 3, 4, 5, 6, 7, 8, 9] de la bibliografía.

2.1. Introducción: definición de proceso estocástico

Antes de dar la definición del proceso de ramificación que consideraremos, introducimos un concepto más general.

Definición 2.1.1. Un proceso estocástico es una colección de variables aleatorias $\{X_t\}_{t \in T}$ todas ellas definidas sobre un espacio de probabilidad (Ω, \mathcal{F}, P) .

En un proceso estocástico intervienen los siguientes elementos:

- El conjunto de valores que toman las variables aleatorias $X_t, \forall t \in T$. Este conjunto se denomina espacio de estados del proceso. Lo denotamos por E .
- El conjunto T que se denomina espacio parametral y representa el tiempo.

Así, $\forall t \in T$, X_t representa el estado del proceso en el instante t y $\{X_t\}_{t \in T}$ representa la evolución del proceso a lo largo del tiempo.

Según si T es discreto o continuo y E es discreto o continuo, surgen procesos estocásticos con espacio de estados discreto, en tiempo discreto; con espacio de estados discreto, en tiempo continuo; con espacio de estados continuo y tiempo discreto y con espacio de estados continuo y tiempo continuo.

Además, también se tienen en cuenta las relaciones entre las variables aleatorias que forman el proceso, apareciendo así diferentes tipos de procesos estocásticos.

Por ejemplo, si $\{X_n\}_{n \in \mathbb{N}}$ es un proceso estocástico con espacio de estados discreto tal que $P(X_n = x_n | X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1})$, $\forall x_1, \dots, x_n \in E$, se dice que el proceso $\{X_n\}_{n \in \mathbb{N}}$ es una Cadena de Markov en tiempo discreto.

En el siguiente ejemplo, vemos un caso particular sencillo de Cadena de Markov.

Ejemplo 2.1.1. Sean $X_0, X_1, \dots, X_n, \dots$ variables aleatorias independientes e idénticamente distribuidas. X_0 toma valores en \mathbb{Z} y X_i toma el valor 1 o -1 con probabilidad p y $1-p$ respectivamente, $\forall i \in \mathbb{N}$.

Sea $Y_n = X_0 + \dots + X_n$, $n \geq 0$. Entonces, se define el camino aleatorio en los enteros como el proceso estocástico en tiempo discreto y espacio de estados discreto $\{Y_n\}_{n \geq 0}$.

$\{Y_n\}_{n \geq 0}$ es una Cadena de Markov en tiempo discreto. En efecto, sean $y_0, \dots, y_n \in \mathbb{Z}$:

$$\begin{aligned} P(Y_n = y_n | Y_0 = y_0, \dots, Y_{n-1} = y_{n-1}) &= \frac{P(Y_0 = y_0, \dots, Y_{n-1} = y_{n-1}, Y_n = y_n)}{P(Y_0 = y_0, \dots, Y_{n-1} = y_{n-1})} \\ &= \frac{P(X_0 = y_0, \dots, X_0 + \dots + X_{n-1} = y_{n-1}, X_0 + \dots + X_n = y_n)}{P(X_0 = y_0, \dots, X_0 + \dots + X_{n-1} = y_{n-1})} \\ &= \frac{P(X_0 = y_0, \dots, X_{n-1} = y_{n-1} - y_{n-2}, X_n = y_n - y_{n-1})}{P(X_0 = y_0, \dots, X_{n-1} = y_{n-1} - y_{n-2})} \\ &\stackrel{(1)}{=} \frac{P(X_0 = y_0) \cdots P(X_{n-1} = y_{n-1} - y_{n-2}) P(X_n = y_n - y_{n-1})}{P(X_0 = y_0) \cdots P(X_{n-1} = y_{n-1} - y_{n-2})} \\ &= P(X_n = y_n - y_{n-1}) \end{aligned}$$

$$\begin{aligned} P(Y_n = y_n | Y_{n-1} = y_{n-1}) &= \frac{P(Y_n = y_n, Y_{n-1} = y_{n-1})}{P(Y_{n-1} = y_{n-1})} \\ &= \frac{P(X_n = y_n - y_{n-1}, Y_{n-1} = y_{n-1})}{P(Y_{n-1} = y_{n-1})} \\ &\stackrel{(2)}{=} \frac{P(X_n = y_n - y_{n-1}) P(Y_{n-1} = y_{n-1})}{P(Y_{n-1} = y_{n-1})} = P(X_n = y_n - y_{n-1}), \end{aligned}$$

siendo (1) cierta por la independencia de las variables aleatorias X_0, \dots, X_n y (2) por la Proposición 1.1.5.

2.2. Introducción histórica

A mediados del siglo XIX, existía una preocupación social por la disminución de apellidos provenientes de familias de la aristocracia inglesa. En general,

inquietaba la posible desaparición de los apellidos de familias influyentes históricamente en la sociedad.

Francis Galton intentó abordar el problema. Tras algunas propuestas imprecisas, decidió publicarlo en un periódico matemático, *Educational Times*, en 1873. Relacionó la desaparición de apellidos con el tamaño de la descendencia de estas familias, en concreto, el número de hijos varones. Galton se preguntó cuántos hijos varones (de media) debería tener cada familia en una generación para que el apellido no desapareciera.

Henry William Watson fue el primero en intentar resolver el problema formulado por Galton. Posteriormente, juntos estudiaron estos procesos, conocidos como los procesos de Galton-Watson, y escribieron un artículo al respecto, [8], en 1875. Se consideraron las siguientes suposiciones:

- Inicialmente, cada individuo tiene un apellido distinto.
- La probabilidad de que un adulto varón tenga k hijos varones, con $k \in \mathbb{N} \cup \{0\}$, permanece constante.
- Las generaciones no se superponen.

Con este modelo, se concluyó que los apellidos de todas las familias se extinguirían, incluso cuando de media, el tamaño de la población aumentaba de generación en generación.

Aunque las conclusiones de Galton y Watson no fueron muy acertadas, supusieron un incentivo para otros científicos que posteriormente abordaron el problema. Además, a partir de sus ideas y suposiciones, nace el concepto de procesos de ramificación, que definiremos matemáticamente en el siguiente apartado.

2.3. Definición matemática de un proceso de ramificación

Ahora, para definir los procesos de ramificación se consideran las siguientes hipótesis:

- Los individuos se reproducen independientemente unos de otros.
- El número de descendientes de cada individuo son variables aleatorias independientes e idénticamente distribuidas.

Consideremos un proceso que se inicia con un individuo (0-ésima generación). Cada individuo vive una unidad de tiempo en la que da lugar a un número de descendientes y muere. El número de descendientes de un individuo toma los valores $0, 1, 2, \dots$ y la probabilidad de tener k descendientes es p_k .

Definición 2.3.1. Un procesos de ramificación es un proceso estocástico $\{Z_n\}_{n \geq 0}$ tal que

$$\begin{aligned} Z_0 &= 1 \\ Z_1 &= X_{01} \\ Z_2 &= \sum_{j=1}^{Z_1} X_{1j} \\ &\vdots \\ Z_n &= \sum_{j=1}^{Z_{n-1}} X_{n-1,j} \\ &\vdots \end{aligned}$$

donde X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$ son variables aleatorias independientes, idénticamente distribuidas que toman valores en $\{0, 1, 2, \dots\}$ y representan el número de descendientes del j -ésimo individuo de la i -ésima generación. Por tanto, Z_n es el tamaño de la población en el instante n (n -ésima generación).

Observación 2.3.1. Un proceso de ramificación es un proceso estocástico en tiempo discreto $T = \{0, 1, 2, \dots\}$ y con espacio de estados discreto $E = \{0, 1, \dots\}$.

A continuación, vamos a estudiar algunos ejemplos de procesos de ramificación.

Ejemplo 2.3.1. En estos ejemplos vamos a considerar algunas simulaciones de procesos de ramificación, realizados con RStudio ([10]).

- (i) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad binomial de parámetros 2 y p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$ definida en (1.3). Es decir, cada individuo tiene 0, 1 o 2 descendientes.

Para realizar las simulaciones, hemos utilizado el siguiente código de [11] que comentamos a continuación:

```
Z<-1
Z[0]=1
m=2
p=0.7
n=20
Z[1]=rbinom(Z[0],m,p)
for (i in 2:n)
{
```

```

if (Z[i-1]==0) {Z[i]=0} else
{
  x=rbinom(Z[i-1],m,p)
  Z[i]=sum(x)
}
}
Z
plot(Z,type="l",col="blue",ylab="Z_n",xlab="n")

```

Partimos de un individuo. Elegimos los valores correspondientes a la distribución de probabilidad elegida, en este caso, $m=2$, $p=0.7$. Elegimos el número de generaciones, n . A continuación, guardamos en cada componente de un vector, Z , el valor de Z_k , $k \in \{0, \dots, n\}$. Tenemos dos opciones:

- Si $Z_{k-1} = 0$, entonces $Z_k = 0$ porque la población ya se encuentra extinta.
- Si $Z_{k-1} \neq 0$, entonces calculamos Z_{k-1} observaciones aleatorias de una distribución binomial 2, p . Z_k será la suma de todas ellas. Es decir, se suman los descendientes de cada individuo de la generación anterior.

Por último, obtenemos la representación gráfica del resultado.

Al realizar simulaciones, hemos obtenido diferentes situaciones.

Por una parte, para $p = 0.3$, el tamaño de la población Z_n siempre tendía a cero a lo largo del tiempo. En todas las simulaciones, la población se acababa extinguiendo en los primeros instantes de tiempo.

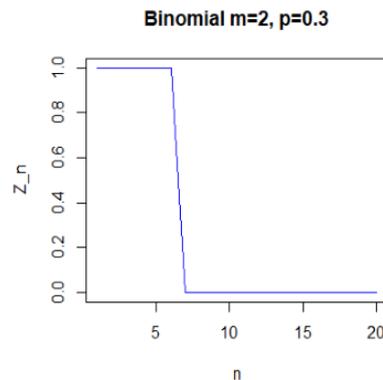


Figura 2.1: Evolución del tamaño de la población ($X_{ij}; \text{Bin}(2,0.3)$).

Por otra parte, para $p = 0.7$, obteníamos resultados más variados. Algunas veces, la población se extinguía en los primeros instantes. Sin embargo, cuando esto no ocurría, el tamaño de la población crecía indefinidamente.

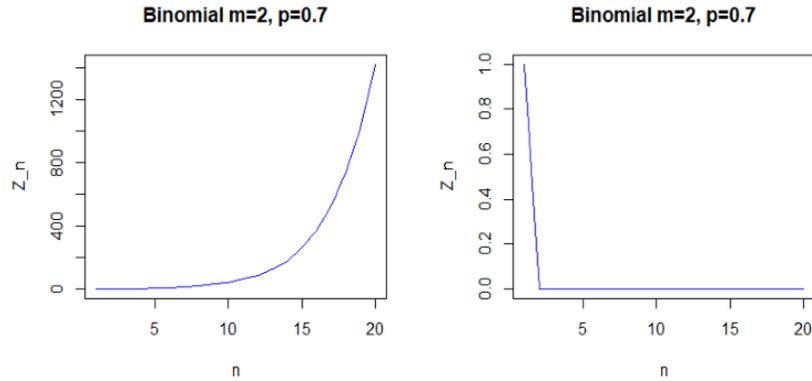


Figura 2.2: Evolución del tamaño de la población ($X_{ij}:\text{Bin}(2,0.7)$).

- (ii) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad geométrica de parámetro p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$ definida en (1.5). Es decir, el número de descendientes de un individuo puede ser 0, 1, 2, ...

Al realizar simulaciones, hemos obtenido de nuevo diferentes resultados.

Por un lado, para $p = 0.3$, a veces, la población se extinguía en los primeros instantes. Sin embargo, cuando esto no ocurría, el tamaño de la población crecía indefinidamente.

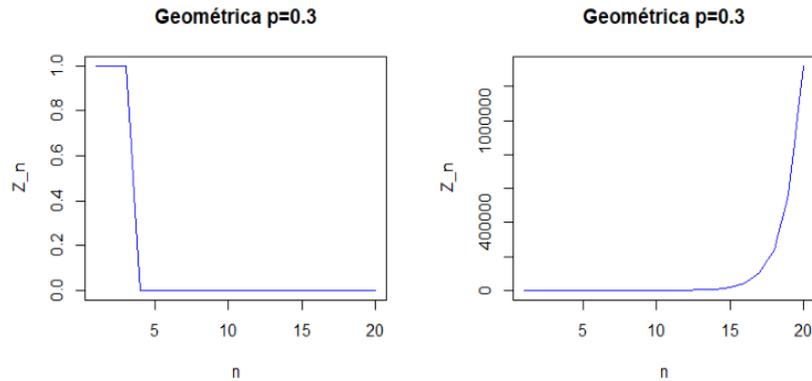


Figura 2.3: Evolución del tamaño de la población ($X_{ij}:\text{Geométrica}(0.3)$).

Por otro lado, para $p = 0.7$, el tamaño de la población Z_n siempre tendía a cero a lo largo del tiempo. En todas las simulaciones, la población se extingue en los primeros instantes de tiempo.

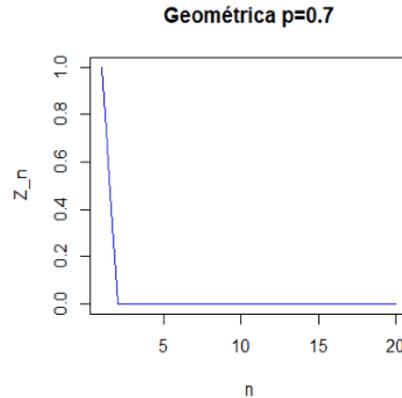


Figura 2.4: Evolución del tamaño de la población (X_{ij} :Geométrica(0.7)).

Es decir, obteníamos un resultado simétrico al primer ejemplo.

En definitiva, vemos que el comportamiento de los procesos de ramificación depende del valor de los parámetros escogidos.

2.4. Distribución de Z_n

A continuación vamos a estudiar la distribución de probabilidad de la variable aleatoria que representa el tamaño de la población en cada instante de tiempo, Z_n .

En algunos casos puede ser complicado obtener su distribución de probabilidad directamente. Por esta razón, vamos a obtener su función generatriz de probabilidad.

El siguiente resultado proporciona una manera sencilla de calcular la función generatriz de probabilidad de Z_n .

Teorema 2.4.1. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Sea*

$$G_n(s) = \sum_{k=0}^{\infty} P(Z_n = k) s^k, \quad n \geq 0,$$

la función generatriz de probabilidad de Z_n y

$$G(s) = \sum_{k=0}^{\infty} p_k s^k$$

la función generatriz de probabilidad de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$. Entonces:

$$G_0(s) = s, \quad G_n(s) = G_{n-1}(G(s)), \quad n = 1, 2, \dots$$

Demostración. Tenemos que $Z_0 = 1$, así:

$$G_0(s) = \sum_{k=0}^{\infty} P(Z_0 = k)s^k = P(Z_0 = 1)s = s.$$

Por otro lado, Z_n es una suma aleatoria de Z_{n-1} variables aleatorias independientes e idénticamente distribuidas $X_{n-1,j}$, $j \in \mathbb{N}$ con función generatriz de probabilidad $G(s) = \sum_{k=0}^{\infty} p_k s^k$. Por lo tanto, aplicando el Teorema 1.4.1 con $N = Z_{n-1}$ y $X_{n-1,j} = X_j$ se tiene:

$$G_n(s) = G_{n-1}(G(s)).$$

□

A partir del teorema anterior, deducimos el siguiente corolario.

Corolario 2.4.2. Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Sea $G_n(s) = \sum_{k=0}^{\infty} P(Z_n = k)s^k$, $n \geq 0$ la función generatriz de probabilidad de Z_n y $G(s) = \sum_{k=0}^{\infty} p_k s^k$ la función generatriz de probabilidad de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$. Entonces:

$$G_n(s) = (G \circ \dots \circ G)(s), \quad n = 0, 1, 2, \dots$$

Es decir, la función generatriz de probabilidad de Z_n , G_n , se puede obtener componiendo la función generatriz de probabilidad de las variables X_{ij} , G , n veces.

A continuación, vamos a ver un ejemplo de cálculo de la función generatriz de probabilidad de Z_n que puede realizarse explícitamente.

Ejemplo 2.4.1. Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad binomial de parámetros 1 y p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$ (cada individuo tiene 0 o 1 descendientes).

Sabemos que la función generatriz de probabilidad de estas variables aleatorias es $G(s) = (1 - p) + ps$ (Ejemplo 1.2.1). Queremos calcular la función generatriz de probabilidad de Z_n . Utilizando el Corolario 2.4.2, obtenemos:

$$\begin{aligned} G_1(s) &= G(s) = (1 - p) + ps \\ G_2(s) &= G(G(s)) = (1 - p) + p((1 - p) + ps) \\ &= (1 - p) + p(1 - p) + p^2s \\ G_3(s) &= G(G_2(s)) = (1 - p) + p(1 - p) + p^2((1 - p) + ps) \\ &= (1 - p) + p(1 - p) + p^2(1 - p) + p^3s \\ &\vdots \end{aligned}$$

Suponemos por hipótesis de inducción que $G_{n-1}(s) = (1-p) + p(1-p) + \dots + p^{n-2}(1-p) + p^{n-1}s$. Entonces, como $G_n(s) = G(G_{n-1}(s))$:

$$\begin{aligned} G_n(s) &= (1-p) + p((1-p) + p(1-p) + \dots + p^{n-2}(1-p) + p^{n-1}s) \\ &= (1-p) + p(1-p) + p^2(1-p) + \dots + p^{n-1}(1-p) + p^n s \\ &= (1-p) \sum_{k=0}^{n-1} p^k + p^n s = (1-p) \frac{1-p^n}{1-p} + p^n s = (1-p^n) + p^n s \end{aligned}$$

Luego, por el Teorema 1.3.1 deducimos que Z_n tiene distribución binomial de parámetros 1 y p^n .

2.5. Momentos de Z_n

Utilizando la expresión de la función generatriz de probabilidad de Z_n obtenida en el apartado anterior vamos a deducir el valor de la esperanza y la varianza de la variable aleatoria Z_n .

2.5.1. Esperanza de Z_n

Teorema 2.5.1. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Entonces la esperanza de Z_n es*

$$E(Z_n) = \mu^n,$$

donde μ es la esperanza de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$.

Demostración. Vamos a demostrar la igualdad por inducción sobre n .

Si $n = 0$, tenemos que $Z_0 = 1$. Por lo tanto, $E(Z_0) = E(1) = 1 = \mu^0$.

Suponemos que se da la igualdad hasta $n-1$. Es decir, $E(Z_k) = \mu^k, \forall k \in \{0, \dots, n-1\}$. Por el Corolario 1.3.6 tenemos que:

$$E(Z_n) = G'_n(1).$$

Por otro lado, por el Teorema 2.4.1 podemos expresar la función generatriz de probabilidad de Z_n como $G_n(s) = G_{n-1}(G(s))$. Derivando respecto de s , por la regla de la cadena, obtenemos:

$$G'_n(s) = G'_{n-1}(G(s))G'(s).$$

Por lo tanto:

$$\begin{aligned} E(Z_n) &= G'_{n-1}(G(1))G'(1) = G'_{n-1}(1)G'(1) \stackrel{(1)}{=} G'_{n-1}(1)\mu \\ &\stackrel{(2)}{=} E(Z_{n-1})\mu \stackrel{(3)}{=} \mu^{n-1}\mu = \mu^n, \end{aligned}$$

siendo (1) y (2) ciertas por el Corolario 1.3.6, y (3) por la hipótesis de inducción. \square

A partir de este teorema, deducimos el siguiente corolario.

Corolario 2.5.2. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Entonces, cuando $n \rightarrow \infty$:*

$$E(Z_n) \rightarrow \begin{cases} 0 & , \text{ si } \mu < 1, \\ 1 & , \text{ si } \mu = 1, \\ \infty & , \text{ si } \mu > 1, \end{cases}$$

donde μ es la esperanza de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$.

Podemos ver que el comportamiento del proceso de ramificación a lo largo del tiempo puede ser muy distinto dependiendo del valor de la esperanza de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$. Más adelante, veremos como influye en la probabilidad de extinción de la población.

2.5.2. Varianza de Z_n

Teorema 2.5.3. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Entonces la varianza de Z_n es*

$$\text{Var}(Z_n) = \begin{cases} n\sigma^2 & , \text{ si } \mu = 1, \\ \sigma^2 \mu^{n-1} \frac{\mu^{n-1}}{\mu-1} & , \text{ si } \mu \neq 1, \end{cases}$$

donde μ y σ^2 son la esperanza y la varianza de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$, respectivamente.

Demostración. Por el Corolario 1.3.6 y el Teorema 2.5.1, podemos escribir las siguientes igualdades:

$$\text{Var}(Z_n) = G''_n(1) + G'_n(1) - G'_n(1)^2 \Rightarrow G''_n(1) = \text{Var}(Z_n) - \mu^n + \mu^{2n}, \quad (2.1)$$

$$G''_{n-1}(1) = \text{Var}(Z_{n-1}) - \mu^{n-1} + \mu^{2n-2}, \quad (2.2)$$

$$G''(1) = \sigma^2 - \mu + \mu^2, \quad (2.3)$$

$$G'_{n-1}(1) = \mu^{n-1}, \quad (2.4)$$

$$G'(1) = \mu. \quad (2.5)$$

Además, si derivamos dos veces la igualdad del Teorema 2.4.1, obtenemos:

$$G''_n(s) = G''_{n-1}(G(s))G'(s)^2 + G'_{n-1}(G(s))G''(s).$$

A continuación, evaluamos la expresión anterior en $s = 1$:

$$G''_n(1) = G''_{n-1}(1)G'(1)^2 + G'_{n-1}(1)G''(1). \quad (2.6)$$

Sustituimos (2.1), (2.2), (2.3), (2.4) y (2.5) en (2.6):

$$\text{Var}(Z_n) - \mu^n + \mu^{2n} = (\text{Var}(Z_{n-1}) - \mu^{n-1} + \mu^{2n-2})\mu^2 + \mu^{n-1}(\sigma^2 - \mu + \mu^2)$$

y obtenemos la siguiente relación de recurrencia:

$$\text{Var}(Z_n) = \text{Var}(Z_{n-1})\mu^2 + \mu^{n-1}\sigma^2.$$

Iterando, tenemos:

$$\begin{aligned} \text{Var}(Z_n) &= (\text{Var}(Z_{n-2})\mu^2 + \mu^{n-2}\sigma^2)\mu^2 + \mu^{n-1}\sigma^2 \\ &= \text{Var}(Z_{n-2})\mu^4 + \mu^n\sigma^2 + \mu^{n-1}\sigma^2 \\ &= (\text{Var}(Z_{n-3})\mu^2 + \mu^{n-3}\sigma^2)\mu^4 + \mu^n\sigma^2 + \mu^{n-1}\sigma^2 \\ &= \text{Var}(Z_{n-3})\mu^6 + \mu^{n+1}\sigma^2 + \mu^n\sigma^2 + \mu^{n-1}\sigma^2 \\ &= \dots \\ &= \text{Var}(Z_{n-n})\mu^{2n} + \sigma^2(\mu^{n+n-2} + \dots + \mu^{n+1} + \mu^n + \mu^{n-1}) \\ &= \text{Var}(Z_0)\mu^{2n} + \sigma^2(\mu^{2n-2} + \dots + \mu^{n+1} + \mu^n + \mu^{n-1}). \end{aligned}$$

Como $\text{Var}(Z_0) = \text{Var}(1) = 0$:

$$\begin{aligned} \text{Var}(Z_n) &= \sigma^2(\mu^{2n-2} + \dots + \mu^{n+1} + \mu^n + \mu^{n-1}) \\ &= \sigma^2\mu^{n-1}(\mu^{n-1} + \dots + \mu^2 + \mu + 1) = \sigma^2\mu^{n-1} \sum_{k=0}^{n-1} \mu^k. \end{aligned}$$

Diferenciamos dos casos:

(i) Si $\mu = 1$, entonces:

$$\text{Var}(Z_n) = \sigma^2 \sum_{k=0}^{n-1} 1 = n\sigma^2.$$

(ii) Si $\mu \neq 1$, entonces:

$$\text{Var}(Z_n) = \sigma^2\mu^{n-1} \sum_{k=0}^{n-1} \mu^k = \sigma^2\mu^{n-1} \frac{\mu^n - 1}{\mu - 1}.$$

Así:

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

□

Aplicando estos resultados podemos calcular la esperanza y la varianza de las variables aleatorias de algunos procesos de ramificación.

Ejemplo 2.5.1. (i) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad binomial de parámetros 2 y p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$.

En el apartado (i) del Ejemplo 1.3.2 vimos que:

$$\mu = 2p \quad \text{y} \quad \sigma^2 = 2p(1-p).$$

Por tanto, por el Teorema 2.5.1 y el Teorema 2.5.3 deducimos:

$$E(Z_n) = (2p)^n,$$

$$Var(Z_n) = \begin{cases} n2p(1-p) & , \text{ si } p = \frac{1}{2}, \\ (2p)^n(1-p)\frac{(2p)^n-1}{2p-1} & , \text{ si } p \neq \frac{1}{2}, \end{cases}$$

ya que $\mu = 1$ si y solo si $2p = 1$. Es decir, $\mu = 1$ si y solo si $p = \frac{1}{2}$.

(ii) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad geométrica de parámetro p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$.

En el apartado (iii) del Ejemplo 1.3.2 vimos que:

$$\mu = \frac{1-p}{p} \quad \text{y} \quad \sigma^2 = \frac{1-p}{p^2}.$$

Por tanto, por el Teorema 2.5.1 y el Teorema 2.5.3 deducimos:

$$E(Z_n) = \left(\frac{1-p}{p}\right)^n,$$

$$Var(Z_n) = \begin{cases} n\frac{1-p}{p^2} & , \text{ si } p = \frac{1}{2}, \\ \frac{(1-p)^{n-1}}{p^{n-1}} \frac{(1-p)}{p^2} \frac{(\frac{1-p}{p})^{n-1}-1}{\frac{1-p}{p}-1} = \frac{(1-p)^n}{p^{2n}} \frac{(1-p)^n-p^n}{1-2p} & , \text{ si } p \neq \frac{1}{2}, \end{cases}$$

ya que $\mu = 1$ si y solo si $\frac{1-p}{p} = 1$. Es decir, $\mu = 1$ si y solo si $p = \frac{1}{2}$.

2.6. Probabilidad de extinción

Por último, en este apartado, vamos a estudiar, dado un proceso de ramificación, qué ocurre con el tamaño de la población a lo largo del tiempo.

Definición 2.6.1. Dado un proceso de ramificación $\{Z_n\}_{n \geq 0}$, definimos la extinción como el siguiente suceso:

$$E := \{\omega \in \Omega : Z_n(\omega) = 0, \text{ para algún } n \in \mathbb{N}\} = \cup_{n \in \mathbb{N}} (Z_n = 0).$$

Definición 2.6.2. Dado un proceso de ramificación $\{Z_n\}_{n \geq 0}$, definimos la probabilidad de extinción como:

$$e := P(\{\omega \in \Omega : Z_n(\omega) = 0, \text{ para algún } n \in \mathbb{N}\}) = P(E).$$

Por otro lado, consideramos para $n \in \mathbb{N}$:

$$E_n := \{\omega \in \Omega : Z_n(\omega) = 0\} = (Z_n = 0),$$

los sucesos tales que el proceso de ramificación está extinto en la n -ésima generación. Denotamos por $e_n := P(E_n)$.

Proposición 2.6.1. *La sucesión $\{e_n : n \in \mathbb{N}\}$ es monótona no decreciente. Es decir, $e_n \leq e_{n+1}$, $\forall n \in \mathbb{N}$.*

Demostración. Dado $\omega \in \Omega$, si $Z_n(\omega) = 0$, entonces $Z_m(\omega) = 0$, $\forall m \geq n$. Es decir, si el tamaño de la población es igual a cero en un instante n , la población seguirá extinta posteriormente. Por lo tanto, tenemos que $E_n \subset E_{n+1}$, y como la probabilidad es una función monótona no decreciente, $e_n \leq e_{n+1}$. \square

A continuación establecemos un lema que utilizaremos en la demostración del siguiente teorema.

Lema 2.6.2. *Dado un proceso de ramificación $\{Z_n\}_{n \geq 0}$, tenemos que:*

$$e = \lim_{n \rightarrow \infty} e_n.$$

Demostración. Inicialmente, podemos escribir la siguiente igualdad:

$$E = \cup_{n=0}^{\infty} E_n.$$

Entonces:

$$P(E) = P(\cup_{n=0}^{\infty} E_n).$$

Consideramos los sucesos $B_0 = E_0$ y $B_n = E_n \setminus E_{n-1} = \{\omega \in \Omega : \omega \in E_n \text{ y } \omega \notin E_{n-1}\}$ para $n \geq 1$. Como $E_{n-1} \subset E_n$, tenemos que $B_i \cap B_j = \emptyset$, $\forall i, j \geq 0$. Además, se cumple que $\cup_{n=0}^{\infty} E_n = \cup_{n=0}^{\infty} B_n$. Por tanto:

$$\begin{aligned} P(E) &= P(\cup_{n=0}^{\infty} E_n) = P(\cup_{n=0}^{\infty} B_n) = P(B_0) + \sum_{n=1}^{\infty} P(B_n) \\ &= P(E_0) + \sum_{n=1}^{\infty} (P(E_n) - P(E_{n-1})) = \lim_{n \rightarrow \infty} P(E_n). \end{aligned}$$

En definitiva, $e = \lim_{n \rightarrow \infty} e_n$. \square

Teorema 2.6.3. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. La probabilidad de extinción e es la menor raíz no negativa solución de la ecuación*

$$x = G(x),$$

donde G es la función generatriz de probabilidad de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$.

Demostración. Primero, vamos a comprobar que e es una raíz de la ecuación $x = G(x)$. Por el Teorema 2.4.1 y el Corolario 2.4.2, escribimos la siguiente igualdad:

$$G_n(s) = G_{n-1}(G(s)) = (G \circ \cdots \circ G)(s) = G(G_{n-1}(s)). \quad (2.7)$$

Además, por la definición de función generatriz de probabilidad, sabemos que $e_n = P(Z_n = 0) = G_n(0)$. Evaluando la expresión (2.7) en $s = 0$, obtenemos que:

$$e_n = G(e_{n-1}), \quad \text{para } n = 1, 2, \dots \quad (2.8)$$

con la condición inicial $e_0 = P(Z_0 = 0) = 0$, ya que $Z_0 = 1$ por definición del proceso de ramificación.

Ahora tomamos el límite cuando $n \rightarrow \infty$ en (2.8). Por el Lema 2.6.2, sabemos que

$$\lim_{n \rightarrow \infty} e_n = e.$$

Por otro lado, sabemos que G tiene un radio de convergencia mayor o igual a 1. Por tanto, G es una función continua en $[0, 1]$ y utilizando de nuevo el Lema 2.6.2 obtenemos

$$\lim_{n \rightarrow \infty} G(e_{n-1}) = G(\lim_{n \rightarrow \infty} e_{n-1}) = G(e).$$

En definitiva, tenemos que

$$e = G(e).$$

Es decir, e es una raíz de la ecuación $x = G(x)$.

En segundo lugar, vamos a comprobar que e es la raíz no negativa más pequeña de $x = G(x)$. Suponemos que η es otra raíz no negativa de $x = G(x)$. Queremos ver que $e \leq \eta$.

Por la Proposición 1.2.1, G es una función no decreciente en $[0, 1]$. Es decir, $G(x_1) \leq G(x_2)$, $\forall x_1, x_2 \in [0, 1]$, $x_1 < x_2$. Por tanto:

$$e_1 = G(e_0) = G(0) \leq G(\eta) = \eta,$$

siendo la primera igualdad cierta por (2.8).

Ahora, aplicando G de nuevo en la desigualdad anterior, obtenemos:

$$e_2 = G(e_1) \leq G(\eta) = \eta,$$

siendo la primera igualdad cierta por (2.8).

Aplicando G sucesivas veces, tenemos la siguiente desigualdad:

$$e_n \leq \eta, \quad \text{para } n = 1, 2, \dots$$

Tomamos límites a ambos lados y utilizando el Lema 2.6.2 obtenemos:

$$e = \lim_{n \rightarrow \infty} e_n \leq \eta.$$

Luego queda demostrado. \square

Este teorema proporciona una herramienta para encontrar la probabilidad de extinción de la población, a partir de la función generatriz de probabilidad G .

Por último, vamos a enunciar un teorema que nos ayudará a saber cuando se producirá la extinción de la población.

Teorema 2.6.4. *Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación. Suponemos que $p_1 \neq 1$. La probabilidad de extinción e cumple que $e = 1$ si y solo si $\mu \leq 1$, siendo μ la esperanza de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$.*

Demostración. Sabemos por el Teorema 2.6.3 que e es la menor raíz no negativa de la ecuación $x = G(x)$, siendo G la función generatriz de probabilidad de las variables aleatorias X_{ij} , $i, j \in \mathbb{N} \cup \{0\}$.

En el intervalo $[0, 1]$, G es continua, porque su radio de convergencia es mayor o igual que 1. Además, por la Proposición 1.2.1, G es no decreciente y convexa en $[0, 1]$. Con estas características, podemos dibujar un boceto de la representación gráfica de G en el intervalo $[0, 1]$:

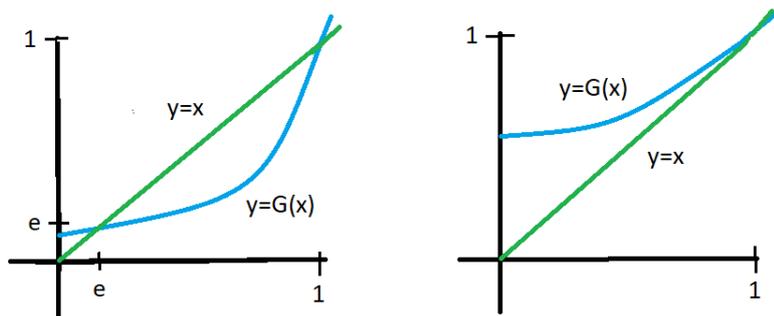


Figura 2.5: Función generatriz de probabilidad y probabilidad de extinción.

Diferenciamos dos casos. En el primer caso, $y = G(x)$ corta a $y = x$ en dos puntos: $x = e < 1$ y $x = 1$. En el segundo caso, solo se cortan en $x = e = 1$. Observamos que en el primer caso, $G'(1) > 1$, mientras que en el segundo caso, $G'(1) \leq 1$.

Por tanto, como $\mu = G'(1)$, tenemos que $e = 1$ si y solo si $\mu \leq 1$. □

Observación 2.6.1. En el teorema anterior, partimos de que $p_1 \neq 1$ porque si $p_1 = 1$, el número de descendientes de cada individuo siempre es igual a 1 y por tanto $\mu = 1$ y $e = 0$.

Finalmente, vamos a calcular las probabilidades de extinción de algunos procesos de ramificación.

Ejemplo 2.6.1. (i) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad binomial de parámetros 2 y p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$.

En el apartado (i) del Ejemplo 1.2.1, vimos que la función generatriz de probabilidad correspondiente es:

$$G(s) = ((1 - p) + ps)^2.$$

Queremos calcular la probabilidad de extinción del proceso de ramificación. Como dice el Teorema 2.6.3, igualamos las expresiones $x = G(x)$:

$$\begin{aligned} x &= ((1 - p) + px)^2 \Leftrightarrow x = (1 - p)^2 + 2p(1 - p)x + p^2x^2 \Leftrightarrow \\ &\Leftrightarrow 0 = p^2x^2 + (2p(1 - p) - 1)x + (1 - p)^2. \end{aligned}$$

Ahora, calculamos sus raíces:

$$\begin{aligned} x &= \frac{1 - 2p(1 - p) \pm \sqrt{4p^2(1 - p)^2 + 1 - 4p(1 - p) - 4p^2(1 - p)^2}}{2p^2} \\ &= \frac{1 - 2p + 2p^2 \pm \sqrt{1 - 4p + 4p^2}}{2p^2} = \frac{1 - 2p + 2p^2 \pm \sqrt{(2p - 1)^2}}{2p^2} \\ &= \frac{(1 - 2p + 2p^2) \pm (2p - 1)}{2p^2}. \end{aligned}$$

Obtenemos:

$$\begin{aligned} x_1 &= \frac{(1 - 2p + 2p^2) + (2p - 1)}{2p^2} = \frac{2p^2}{2p^2} = 1, \\ x_2 &= \frac{(1 - 2p + 2p^2) - (2p - 1)}{2p^2} = \frac{2p^2 - 4p + 2}{2p^2} = \frac{(1 - p)^2}{p^2}. \end{aligned}$$

La primera raíz es igual a 1. La segunda raíz será más pequeña que 1 si y solo si $1 - p < p$. Es decir, si $p > \frac{1}{2}$. La probabilidad de extinción será:

$$e = \begin{cases} 1 & , \text{ si } 0 \leq p \leq \frac{1}{2}, \\ \frac{(1-p)^2}{p^2} & , \text{ si } \frac{1}{2} < p \leq 1, \end{cases}$$

Hemos deducido que la probabilidad de extinción será igual a 1 si y solo si $p \leq \frac{1}{2}$. Esto coincide con el Teorema 2.6.4. En nuestro caso, $\mu = 2p$. Por tanto, $\mu \leq 1$ si y solo si $p \leq \frac{1}{2}$.

En el apartado (i) del Ejemplo 2.3.1, vimos que con $p = 0.3$, la población siempre se extingüía. En este caso, la probabilidad de extinción es $e = 1$. Por otro lado, con $p = 0.7$, la población se extingüía o crecía indefinidamente. En este caso, la probabilidad de extinción es $e = \frac{(0,7-1)^2}{0,7^2} = 0,1836$.

- (ii) Sea $\{Z_n\}_{n \geq 0}$ un proceso de ramificación tal que las variables aleatorias X_{ij} tienen distribución de probabilidad geométrica de parámetro p , con $p \in (0, 1)$, $\forall i, j \in \mathbb{N} \cup \{0\}$.

En el apartado (iii) del Ejemplo 1.2.1, vimos que la función generatriz de probabilidad correspondiente es:

$$G(s) = \frac{p}{1 - (1-p)s}.$$

Queremos calcular la probabilidad de extinción del proceso de ramificación. Como dice el Teorema 2.6.3, igualamos las expresiones $x = G(x)$:

$$\begin{aligned} x &= \frac{p}{1 - (1-p)x} \Leftrightarrow x - (1-p)x^2 = p \Leftrightarrow \\ &\Leftrightarrow 0 = (1-p)x^2 - x + p. \end{aligned}$$

Ahora, calculamos sus raíces:

$$\begin{aligned} x &= \frac{1 \pm \sqrt{1 - 4p(1-p)}}{2(1-p)} = \frac{1 \pm \sqrt{1 - 4p + 4p^2}}{2(1-p)} \\ &= \frac{1 \pm \sqrt{(2p-1)^2}}{2(1-p)} = \frac{1 \pm (2p-1)}{2(1-p)}. \end{aligned}$$

Obtenemos:

$$\begin{aligned} x_1 &= \frac{1 + 2p - 1}{2(1-p)} = \frac{p}{1-p}, \\ x_2 &= \frac{1 - 2p + 1}{2(1-p)} = \frac{2(1-p)}{2(1-p)} = 1. \end{aligned}$$

La segunda raíz es igual a 1. La primera raíz será más pequeña que 1 si y solo si $p < 1-p$, es decir, si $p < \frac{1}{2}$. La probabilidad de extinción será:

$$e = \begin{cases} \frac{p}{1-p} & , \text{ si } 0 \leq p < \frac{1}{2}, \\ 1 & , \text{ si } \frac{1}{2} \leq p \leq 1, \end{cases}$$

Hemos deducido que la probabilidad de extinción será igual a 1 si y solo si $p \geq \frac{1}{2}$. Esto coincide con el Teorema 2.6.4. En nuestro caso, $\mu = \frac{1-p}{p}$. Por tanto, $\mu \leq 1$ si y solo si $p \geq \frac{1}{2}$.

En el apartado (ii) del Ejemplo 2.3.1, vimos que con $p = 0.3$, la población se extinguía o crecía indefinidamente. En este caso, la probabilidad de extinción es $e = \frac{0.3}{1-0.3} = 0.4286$. Por otro lado, con $p = 0.7$, la población siempre se extinguía. En este caso, la probabilidad de extinción es $e = 1$.

2.7. Ejemplo en la población americana

En este apartado, vamos a estudiar un ejemplo más práctico que analiza el tamaño de la población americana ([6]).

Sea Y una variable aleatoria que describe el número total de hijos que tiene una familia americana cualquiera. Suponemos que la probabilidad de que una familia tenga exactamente n hijos es:

$$P(Y = n) = \begin{cases} 1 - \alpha(p + p^2 + p^3 + \dots) = 1 - \alpha \frac{p}{1-p} & , \text{ si } n = 0, \\ \alpha p^n & , \text{ si } n > 0, \end{cases}$$

donde α y p son números reales positivos, $n = 0, 1, 2, \dots$ y $p < 1$, $1 + \alpha \leq \frac{1}{p}$.

Observamos que estas probabilidades están bien definidas si tenemos en cuenta las restricciones en los valores de α y p . Es decir, tenemos que $P(Y = n) \in [0, 1]$, $\forall n \geq 0$ y $\sum_{n \geq 0} P(Y = n) = 1$.

En este ejemplo, vamos a partir de la suposición de que las probabilidades de descendencia se mantienen constantes en cada generación. Además, los tiempos entre generaciones son fijos y los individuos se reproducen de manera independiente.

Ahora, si suponemos que la probabilidad de tener hijos varones o hembras es la misma, nos preguntamos cuál es la probabilidad de que una familia (americana) tenga exactamente k hijos varones.

Sea X el número total de hijos varones de una familia. Vamos a calcular la probabilidad de que una familia tenga exactamente k hijos varones:

(i) Si $k \geq 1$:

$$\begin{aligned} P(X = k) &= P((X = k) \cap (\cup_{n=0}^{\infty} (Y = n))) \\ &= P(\cup_{n=0}^{\infty} (X = k, Y = n)) \stackrel{(1)}{=} P(\cup_{n=k}^{\infty} (X = k, Y = n)) \\ &= \sum_{n=k}^{\infty} P(X = k, Y = n) = \sum_{n=k}^{\infty} P(X = k | Y = n) P(Y = n), \end{aligned}$$

siendo (1) cierta porque como mucho se pueden tener tantos hijos varones como hijos en total.

Como la probabilidad de tener un hijo varón es igual a la probabilidad de tener una hembra, la probabilidad de tener k hijos varones en una familia de n hijos es $P(X = k | Y = n) = \binom{n}{k} \left(\frac{1}{2}\right)^n$.

Por tanto:

$$\begin{aligned} P(X = k) &= \sum_{n=k}^{\infty} \binom{n}{k} \left(\frac{1}{2}\right)^n \alpha p^n = \alpha \left(\frac{p}{2}\right)^k \sum_{n=k}^{\infty} \binom{n}{n-k} \left(\frac{p}{2}\right)^{n-k} \\ &\stackrel{(2)}{=} \alpha \left(\frac{p}{2}\right)^k \sum_{n=k}^{\infty} \binom{-k-1}{n-k} \left(\frac{-p}{2}\right)^{n-k} \\ &\stackrel{(3)}{=} \alpha \left(\frac{p}{2}\right)^k \left(1 - \frac{p}{2}\right)^{-k-1} = \alpha \frac{p^k (2-p)^{-k-1}}{2^k 2^{-k-1}} = \frac{2\alpha p^k}{(2-p)^{k+1}}, \end{aligned}$$

siendo (2) cierta por la siguiente propiedad de los números combinatorios $\binom{n}{n-k} = (-1)^{n-k} \binom{-n+n-k-1}{n-k} = (-1)^{n-k} \binom{-k-1}{n-k}$, y (3) por el desarrollo del binomio de Newton generalizado.

(ii) Si $k = 0$:

$$\begin{aligned} P(X = 0) &= 1 - P(\cup_{k=1}^{\infty} (X = k)) = 1 - \sum_{k=1}^{\infty} P(X = k) \\ &= 1 - \sum_{k=1}^{\infty} \frac{2\alpha p^k}{(2-p)^{k+1}} = 1 - \frac{2\alpha}{2-p} \sum_{k=1}^{\infty} \left(\frac{p}{2-p}\right)^k \\ &\stackrel{(4)}{=} 1 - \frac{2\alpha}{2-p} \frac{\frac{p}{2-p}}{1 - \frac{p}{2-p}} = 1 - \frac{2\alpha}{2-p} \frac{p}{2-2p} \\ &= 1 - \frac{\alpha p}{(2-p)(1-p)}, \end{aligned}$$

siendo (4) cierta por ser la suma geométrica de razón $\frac{p}{2-p}$.

A continuación, vamos a calcular la esperanza de X :

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} k P(X = k) = \sum_{k=1}^{\infty} k \frac{2\alpha p^k}{(2-p)^{k+1}} = \frac{2\alpha p}{(2-p)^2} \sum_{k=1}^{\infty} k \left(\frac{p}{2-p}\right)^{k-1} \\ &\stackrel{(1)}{=} \frac{2\alpha p}{(2-p)^2} \frac{1}{\left(1 - \frac{p}{2-p}\right)^2} = \frac{2\alpha p(2-p)^2}{(2-p)^2(2-2p)^2} = \frac{\alpha p}{2(1-p)^2}, \end{aligned}$$

siendo (1) cierta porque si $(\sum_{k=0}^{\infty} x^k)' = (\frac{1}{1-x})'$, entonces $\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$.

Suponemos que la población americana se extinguirá cuando se extingan todos los individuos varones, dado que los apellidos desaparecerán.

Por el Teorema 2.6.4, la probabilidad de extinción de la población es $e = 1$ si y solo si $E(X) \leq 1$. Distinguimos dos casos:

▪ $p \leq \frac{1}{2}$

Como $1 + \alpha \leq \frac{1}{p}$, $\alpha \leq \frac{1-p}{p}$. Entonces:

$$E(X) = \frac{\alpha p}{2(1-p)^2} \leq \frac{(1-p)p}{2p(1-p)^2} = \frac{1}{2(1-p)} \leq 1 \Leftrightarrow p \leq \frac{1}{2}.$$

Por tanto, deducimos que si $p \leq \frac{1}{2}$, la población americana se extinguirá. Es decir, la probabilidad de extinción es $e = 1$.

▪ $p > \frac{1}{2}$

En este caso, $E(X) > 1$. Por tanto, la probabilidad de extinción será menor que 1.

Primero, vamos a calcular la función generatriz de probabilidad de X :

$$\begin{aligned}
 G(s) &= \sum_{k=0}^{\infty} P(X = k)s^k = P(X = 0) + \sum_{k=1}^{\infty} P(X = k)s^k \\
 &= 1 - \frac{\alpha p}{(2-p)(1-p)} + \sum_{k=1}^{\infty} \frac{2\alpha p^k}{(2-p)^{k+1}} s^k \\
 &= 1 - \frac{\alpha p}{(2-p)(1-p)} + \frac{2\alpha}{2-p} \sum_{k=1}^{\infty} \left(\frac{ps}{2-p} \right)^k \\
 &\stackrel{(1)}{=} 1 - \frac{\alpha p}{(2-p)(1-p)} + \frac{2\alpha}{2-p} \frac{\frac{ps}{2-p}}{1 - \frac{ps}{2-p}} \\
 &= 1 - \frac{\alpha p}{(2-p)(1-p)} + \frac{2\alpha ps}{(2-p)(2-p-ps)},
 \end{aligned}$$

siendo (1) cierta por ser la suma geométrica de razón $\frac{ps}{2-p}$, convergente para $|s| < \frac{2-p}{p}$.

Para calcular la probabilidad de extinción, vamos a buscar las soluciones de la ecuación $x = G(x)$ (Teorema 2.6.4).

Para simplificar los cálculos, como $1 - \frac{\alpha p}{(2-p)(1-p)}$ no depende de s , la vamos a considerar como una constante c .

$$\begin{aligned}
 x = G(x) &\Rightarrow x - c = \frac{2\alpha px}{(2-p)(2-p-px)} \\
 &\Rightarrow (x-c)(px+p-2) = \frac{-2\alpha px}{2-p} \Rightarrow \\
 &\Rightarrow (x-c) \left(x + \left(1 - \frac{2}{p} \right) \right) = \frac{-2\alpha x}{2-p} \Rightarrow \\
 &\Rightarrow x^2 - \left[-\frac{2\alpha}{2-p} + \frac{2-p}{p} + c \right] x + c \frac{2-p}{p} = 0
 \end{aligned}$$

Nos queda una ecuación de segundo grado del tipo $x^2 + ax + b = 0$. Vamos a desarrollar los coeficientes a y b , sustituyendo c por su valor:

$$\begin{aligned}
 a &= -\frac{2\alpha}{2-p} + \frac{2-p}{p} + c = -\frac{2\alpha}{2-p} + \frac{2-p}{p} + 1 - \frac{\alpha p}{(2-p)(1-p)} \\
 &= 1 + \frac{-2\alpha(1-p)p + (2-p)^2(1-p) - \alpha p^2}{p(1-p)(2-p)} \\
 &= 1 + \frac{(2-p)^2(1-p) - \alpha p(2-p)}{p(1-p)(2-p)} = 1 + \frac{(2-p)(1-p) - \alpha p}{p(1-p)} \\
 b &= c \frac{2-p}{p} = \frac{2-p}{p} - \frac{\alpha}{1-p} = \frac{(2-p)(1-p) - \alpha p}{p(1-p)}
 \end{aligned}$$

Así, la ecuación de segundo grado la podemos reescribir como:

$$(x - 1) \left(x - \frac{(2-p)(1-p) - \alpha p}{p(1-p)} \right) = 0$$

Por tanto, las dos raíces de la ecuación son $x_1 = 1$ y $x_2 = \frac{(2-p)(1-p) - \alpha p}{p(1-p)}$.
Luego si $p > \frac{1}{2}$, la probabilidad de extinción de un apellido en esta población será $e = \frac{(2-p)(1-p) - \alpha p}{p(1-p)}$.

Bibliografía

- [1] Athreya, K. B., y Ney, P. E. (1972). *Branching processes*. Berlin: Springer.
- [2] Billingsley, P. (1995). *Probability and Measure*. 3rd Wiley. New York.
- [3] Grimmett, G., y Stirzaker, D. (2020). *Probability and random processes*. Oxford University Press.
- [4] Grimmett, G., y Welsh, D. (2014). *Probability: an introduction*. Oxford University Press.
- [5] Harris, T. E. (1963). *The theory of branching processes* (Vol. 6). Berlin: Springer.
- [6] Sena, F. G., y i Simorra, X. B. (2011). PROCESOS DE RAMIFICACIÓN: UNA APLICACIÓN EN LA POBLACIÓN AMERICANA. *Reportes Científicos de la FACEN*, 2(1), 12-23.
- [7] Stirzaker, D. (2005). *Stochastic processes and models*. OUP Oxford.
- [8] Watson, H. W., y Galton, F. (1875). On the probability of the extinction of families. *The Journal of the Anthropological Institute of Great Britain and Ireland*, 4, 138-144.
- [9] Žytković, G. (2010). Introduction to stochastic processes-lecture notes. *Department of Mathematics, The University of Texas at Austin*.
- [10] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- [11] <https://github.com/AlyssaYelle/StochasticProcesses>

