

# Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas  
Postgrado en Ciencias Matemáticas

*Proceso de Galton-Watson: Análisis, Estimación y  
Simulación*

Tesis

Que para obtener el título de  
Maestra en Ciencias (Matemáticas)

Presenta  
Laura Magdalena Coutiño Escobar

Director de Tesis  
Dr. Hugo Adán Cruz Suárez

Puebla, Pue.

Octubre 2012



# Dedicatoria

*Tomada de tu mano, inicie mi aprendizaje en la vida.  
Ahora todo lo que soy, se lo debo a tu ejemplo de  
tenacidad y valor, esta tesis te la dedico a ti,  
mamá.*



# Agradecimientos

Primero quiero agradecerle a Dios por dejarme permitir vivir este momento, porque siempre ha sabido guiar mi camino y nunca me ha dejado sola.

A mi mami, porque nunca podré pagar todos tus desvelos, tu amor, tu apoyo incondicional, ni aún con las riquezas más grandes del mundo. Por esto y más gracias.

A mi director de tesis:

Dr. Hugo A. Cruz Suárez, gracias por su paciencia y toda su enseñanza transmitida. Siempre le estaré infinitamente agradecida.

A mis hermanos:

Johana, porque eres mi gran ejemplo para seguir superándome cada día más. Gracias Joy.

Lulú, porque a tu manera siempre estuviste apoyándome. Gracias Lulis.

Ricardo, porque aunque sé que nuestra relación de hermanos no esta muy bien, sé que me deseas lo mejor como yo a ti. Gracias Richard.

A mi papá, porque gracias a ti estoy en este mundo.

A mis amigos:

Marina, por creer en mi, porque me brindaste apoyo y consejos, y en los momentos mas difíciles me alentaste a seguir adelante, anhelando que siempre me preparara para enfrentarme a la vida. Muchas gracias.

Carmen, porque a pesar de la distancia y del tiempo siempre has estado conmigo cuando más te eh necesitado. Gracias amiguis.

Selene, Mónica, Brenda y Daniela, por vivir juntas esta experiencia de la maestría. Muchas gracias a todas.

Dr. Agustín Contreras C., Lic. Mónica Macías P., Dra. Esperanza Guzmán O., porque desde la licenciatura siempre me han apoyado y deseado lo mejor.

Porque ya los considero como mis amigos, muchas gracias.

A los que ya no están presentes físicamente:

Abuelita Magdalena (Q.E.P.D.), porque desde donde estés sé que estas compartiendo mi felicidad y porque siempre estarás presente en mi corazón y pensamiento. Te amo abue.

Padrino Manuel (Q.E.P.D.), porque sé que si estuviera aquí con nosotros estaría muy feliz y deseándome siempre lo mejor. Gracias panzón por todos sus consejos.

A los miembros del jurado:

Dr. Víctor H. Vázquez G., Dra. Hortensia J. Reyes C., Dra. Gladys Linares F., Dr. Francisco S. Tajonar S., Dra. Rosa M. Flores H. y Dr. Fernando Velasco L., por el tiempo tomado para las correcciones y observaciones de la tesis. Muchas gracias.

A CONACYT, mi agradecimiento a esta instancia por la beca otorgada que me permitió cursar la maestría durante estos dos años.

Sé que me faltan mas personas por mencionar, pero a todos: ¡MUCHAS GRACIAS!.

# Introducción

La relevancia de la teoría matemática de ramificación se debe a que trata con uno de los procesos primordiales de la naturaleza: la reproducción (véase [2]).

En este trabajo de tesis se hace un desarrollo acerca del Proceso de Ramificación de Galton-Watson (véase [9, 16, 17, 20]).

Los Procesos de Ramificación suelen describir el comportamiento de una población la cual está compuesta generalmente por individuos (o partículas), las partículas o individuos de un sistema ramificado pueden ser de tipos muy diversos, ya que hay muchas poblaciones cuyos elementos pueden reproducirse, fragmentarse o desaparecer. Por ejemplo: moléculas, genes, virus, células, bacterias, animales, especies de animales, personas, apellidos, grupos étnicos, plantas, enfermedades, lenguajes, religiones, por mencionar algunos (véase [7]). Por ello, las aplicaciones de la teoría de ramificación se encuentran en diferentes áreas como lo son: genética, biología, epidemiología, medicina, economía, demografía, finanzas, física de partículas, sociología, etc, (véase [4, 7]).

En el Siglo XVIII se empezó con el desarrollo y uso de los Procesos de Ramificación con el libro de tres volúmenes del investigador Thomas Malthus, titulado "*Essay on the Principles of Population*", en el cual Malthus relata el problema de extinción de familias burguesas existentes en el lapso de dos siglos (1583-1783). El primero en tratar de explicar el fenómeno relatado por Malthus fue el matemático francés I. J. Bienaymé (1796-1878), quien relacionó correctamente la probabilidad de extinción con el promedio de hijos varones de cada individuo.

Posteriormente Galton y Watson, formularon en 1874 un modelo matemático para el problema de extinción de familias, dicho modelo indica que el número de hijos que cada individuo tiene es independiente de su historia familiar y del número de hijos de los demás individuos, entonces el proceso  $\{Z_n\}$  el cual indica el número de individuos en la  $n$ -ésima generación forma una cadena de Markov (véase [3, 10, 11, 18, 19, 20]).

Si los individuos tienen probabilidad positiva de morir sin dejar descendientes, el sistema se puede extinguir en tiempo finito, es decir, a partir de algún instante aleatorio ya no queda ningún individuo. Por ello, uno de los problemas principales de la teoría es calcular la probabilidad de extinción del sistema.

Se tiene que una de las características del Proceso de Galton-Watson es la distribución de descendientes y el promedio de reproducción. En la práctica, es de interés estimar dichos parámetros por lo tanto, es necesario aplicar técnicas esta-

dísticas que permitan determinar dicha estimación. En la tesis se aplica el método de máxima verosimilitud para la estimación de los parámetros de interés (véase [8, 13, 14]).

Existe una forma de clasificar a los procesos de ramificación de acuerdo a su condición crítica, su parámetro tiempo, el caso simple o múltiple de individuos (partículas), el carácter Markoviano y no-Markoviano. En esta tesis se hace el estudio del caso Markoviano con tiempo y espacio de estados discreto.

También se puede considerar el caso en el que las poblaciones tienen distintos tipos o grupos de individuos y cada uno de estos individuos puede generar descendientes de todos los tipos que coexisten en la población, dicho caso es llamado el Proceso de Galton-Watson Multitipo.

En resumen, el trabajo de tesis se enfoca en los puntos siguientes: un análisis de los conceptos importantes del Proceso de Ramificación de Galton-Watson con tiempo y espacio de estados discreto para que posteriormente se pueda generalizar al caso multitipo. Una vez realizado este análisis, se procede a trabajar con la estimación de los parámetros importantes del Proceso de Galton-Watson y con ayuda de simulación mediante el software R se comparan.

La tesis se ha estructurado en cuatro capítulos, un apartado de conclusiones y tres apéndices.

En el Capítulo 1, se dan algunas definiciones y resultados importantes los cuales son utilizados en el desarrollo del tema.

En el Capítulo 2 se define el Proceso de Galton-Watson  $\{Z_n\}$ , a partir de esta definición se trabaja con su función generadora de probabilidades, sus momentos y se calcula la probabilidad de extinción del sistema.

En el Capítulo 3 se estiman los parámetros probabilidad de descendientes y promedio de reproducción mediante el método de máxima verosimilitud, además se dan estimadores no-paramétricos de la probabilidad de extinción propuestos por algunos autores, los cuales son comparados mediante una simulación.

En el Capítulo 4 se revisa la teoría para el Proceso de Galton-Watson Multitipo y posteriormente se analiza un ejemplo en el cual la teoría es utilizada.

Finalmente se escriben las conclusiones y se incluyen tres apéndices, el Apéndice A en el cual se estudia la distribución límite de  $\{Z_n\}$  cuando  $n$  tiende a infinito para ello es importante analizar los resultados asintóticos del proceso (véase [6]), posteriormente el Apéndice B en el cual se hace un desarrollo breve del Proceso Markoviano de Ramificación a Tiempo Continuo y se da un ejemplo que es usado en el Capítulo 4 y finalmente el Apéndice C que incluye algoritmos de los programas que fueron usados en esta tesis.

# Índice general

<b>Agradecimientos</b>	<b>III</b>
<b>Introducción</b>	<b>V</b>
<b>1. Preliminares</b>	<b>3</b>
1.1. Cadenas de Markov . . . . .	3
1.2. Martingalas . . . . .	5
<b>2. Proceso de Galton-Watson</b>	<b>9</b>
2.1. Análisis del Proceso de Galton-Watson . . . . .	9
2.2. La f.g.p. para el Proceso de Galton-Watson . . . . .	11
2.3. Los momentos del PG-W . . . . .	13
2.4. La probabilidad de extinción . . . . .	15
2.5. Ejemplos . . . . .	19
<b>3. Estimación</b>	<b>25</b>
3.1. Estimación de los parámetros del PG-W . . . . .	25
3.2. Comparación de estimadores por simulación . . . . .	29
<b>4. Proceso de Galton-Watson Multitipo</b>	<b>33</b>
4.1. Análisis del PG-WM . . . . .	33
4.2. Ejemplo del PG-WM . . . . .	36
<b>Conclusiones</b>	<b>39</b>
<b>Apéndice A</b>	<b>41</b>
<b>Apéndice B</b>	<b>45</b>
<b>Apéndice C</b>	<b>51</b>

*ÍNDICE GENERAL*

1

**Bibliografía**

**53**



# Capítulo 1

## Preliminares

Se empieza por dar algunas definiciones y resultados importantes los cuales son necesarios para el desarrollo del tema. En este capítulo se define una Cadena de Markov y la clasificación de estados, posteriormente se da la definición de Martingala (véase [3, 12]) y finalmente se da un resultado el cual es importante para un caso que se verá en el Apéndice A.

### 1.1. Cadenas de Markov

Considérese un sistema que puede estar en cualquiera de un número finito o infinito numerable de estados. Sea  $S$  este conjunto de estados y supóngase que es un subconjunto de los enteros. El conjunto  $S$  es llamado el espacio de estados del sistema. Supóngase que el sistema es observado en tiempo discreto  $t = 0, 1, 2, \dots$  y sea  $X_t$  la variable que denota el estado del sistema en el tiempo  $t$ .

Sean  $X_0, X_1, \dots$  variables aleatorias definidas en un espacio de probabilidad común.

Muchos sistemas (por ejemplo: inventarios, líneas de espera, entre otros) tienen la propiedad de que dada la situación actual, los estados pasados no tienen ninguna influencia sobre el futuro. Esta propiedad es llamada la propiedad de Markov, y los sistemas que cumplen esta propiedad se denominan Cadenas de Markov.

**Definición 1.1.1.** Un proceso estocástico  $\{X_t : t = 0, 1, \dots\}$ , con espacio de estados discreto es una cadena de Markov, si cumple la siguiente propiedad:

$$\begin{aligned} P(X_{t+1} = y \mid X_0 = x_0, X_1 = x_1, \dots, X_{t-1} = x_{t-1}, X_t = x) \\ = P(X_{t+1} = y \mid X_t = x). \end{aligned}$$

$x_0, x_1, \dots, x_{t-1}, x, y \in S$ . Sea  $P(X_{t+1} = y \mid X_t = x) := P(x, y)$ ,  $P(x, y)$  es llamada la probabilidad de transición en un paso de la cadena.

Las probabilidades de transición se pueden representar de forma matricial por:

$$\begin{bmatrix} P(0,0) & \dots & P(0,d) \\ \vdots & & \vdots \\ P(d,0) & \dots & P(d,d) \end{bmatrix}.$$

En este caso el espacio de estados es  $\{0, 1, 2, \dots, d\}$ .

Las probabilidades de transición asociadas a los estados juegan un papel importante en el estudio de las cadenas de Markov. Para describir con mas detalles las propiedades de una cadena de Markov es necesario presentar algunos conceptos y definiciones que se refieren a estos estados.

**Definición 1.1.2.** Sea  $\{X_t, t \geq 0\}$  una cadena de Markov con espacio de estados  $S$  y función de probabilidad de transición  $P$ . Se define el tiempo de alcance a un estado  $y$ , como

$$T_y := \inf\{t : X_t = y\}.$$

Se define,

$$\rho_{xy} := P_x(T_y < \infty),$$

como la probabilidad de que una cadena de Markov empiece en  $x$  y en algún tiempo positivo se encuentre en el estado  $y \in S$ . En particular  $\rho_{yy}$  denota la probabilidad de que una cadena de Markov empiece en  $y$  y que en algún momento regrese a  $y$ .

**Definición 1.1.3.** Un estado  $y$  es llamado recurrente, si  $\rho_{yy} = 1$ . Es decir, si  $y$  es un estado recurrente, una cadena de Markov empieza en  $y$  y regresa a  $y$ , en algún momento, con probabilidad 1.

*Observación 1.1.4.* Un estado  $y \in S$  es absorbente si  $P(y, y) = 1$ . De esta forma un estado absorbente es recurrente ya que

$$P_y(T_y = 1) = P(y, y) = 1,$$

y por lo tanto,  $\rho_{yy} = 1$ .

**Definición 1.1.5.** Un estado  $y$  es llamado transitorio, si  $\rho_{yy} < 1$ . Es decir, si  $y$  es un estado transitorio, una cadena de Markov empieza en  $y$  y tiene probabilidad positiva,  $1 - \rho_{yy}$ , de nunca regresar a  $y$ .

A continuación se hablará de otro proceso estocástico, las Martingalas. La propiedad de Martingala (véase Definición 1.2.2) expresa una relación que se produce en numerosos contextos, y se ha convertido en una herramienta básica en la probabilidad teórica y práctica.

## 1.2. Martingalas

La teoría de la probabilidad tiene sus raíces en juegos de azar por lo cual es frecuente interpretar resultados en términos de situación de juegos. Sea  $\{X_1, X_2, \dots\}$  una sucesión de variables aleatorias, se puede pensar a  $X_n$  en este contexto como la ganancia total después de  $n$  etapas en una sucesión de juegos.

Sea  $(\Omega, \mathcal{F})$  un espacio medible y  $T$  un subconjunto de  $\mathbb{R}$  o  $\mathbb{Z}$ .

**Definición 1.2.1.** Una colección de  $\sigma$ -álgebras  $\{\mathcal{F}_t, t \in T\} \subseteq \mathcal{F}$ , es llamada filtración si  $\mathcal{F}_s \subseteq \mathcal{F}_t$  para todo  $s \leq t$ .

A continuación se da la definición de Martingala.

**Definición 1.2.2.** Sean  $(\Omega, \mathcal{F}, P)$  un espacio de probabilidad,  $\{X_n : n = 0, 1, \dots\}$  una sucesión de variables aleatorias en  $(\Omega, \mathcal{F}, P)$ , y  $\{\mathcal{F}_n\}$  una sucesión no-decreciente de sub  $\sigma$ -álgebras de  $\mathcal{F}$ ;  $X_n$  se supone  $\mathcal{F}_n$ -medible. La sucesión  $\{X_n\}$  se dice ser una martingala relativa a la filtración  $\mathcal{F}_n$  (alternativamente se dirá que  $\{X_n, \mathcal{F}_n\}$  es una martingala) si cumple:

1.  $E[|X_n|] < \infty$ , para cada  $n$ .
2.  $E[X_{n+1} | \mathcal{F}_n] = X_n$ , para cada  $n$  con probabilidad 1.

En las mismas condiciones de la definición anterior, si se cumple 1 y en lugar de 2 se verifica:

1.  $E[X_{n+1} | \mathcal{F}_n] \geq X_n$ , para cada  $n$  con probabilidad 1, se dirá que es una submartingala.
2.  $E[X_{n+1} | \mathcal{F}_n] \leq X_n$ , para cada  $n$  con probabilidad 1, se dirá que es una supermartingala.

**Definición 1.2.3.** Si  $\{X_n : n \geq 1\}$  y  $\{Y_n : n \geq 1\}$  son dos procesos estocásticos entonces se dice que  $\{X_n : n \geq 1\}$  es una martingala (respectivamente super o submartingala) con respecto a  $\{Y_n : n \geq 1\}$ , si  $\{X_n\}$  es una martingala (respectivamente super o submartingala) con respecto a  $\{\mathcal{F}_n : n \geq 1\}$  donde  $\mathcal{F}_n = \sigma(Y_1, Y_2, \dots, Y_n)$ .

A continuación se dan algunas propiedades elementales de las Martingalas, éstas son de gran utilidad para las pruebas del Lema 1.2.4 y Teorema 1.2.5 los cuales se enuncian posteriormente, y serán usados para la demostración del Teorema de Convergencia para Martingalas (véase [21]).

1. Si  $\{X_n : n \geq 1\}$  es una martingala (supermartingala) con respecto a  $\{Y_n\}$  entonces  $E[X_{n+k} | Y_0, Y_1, \dots, Y_n] = (\leq) X_n$  para todo  $k \geq 0$ ,

2. Si  $\{X_n\}$  es una martingala (supermartingala) con respecto a  $\{Y_n\}$ , entonces para  $0 \leq k \leq n$  se satisface

$$E[X_n] = (\leq) E[X_k].$$

**Lema 1.2.4.** Si  $\{X_n : n \geq 1\}$  es una martingala y  $f$  una función convexa, entonces  $\{f(X_n), n \geq 1\}$  es una submartingala.

*Demostración.*

$$\begin{aligned} & E[f(X_{n+1}) \mid Y_1, Y_2, \dots, Y_n] \\ & \geq f(E[(X_{n+1}) \mid Y_1, Y_2, \dots, Y_n]) \\ & = f(X_n). \end{aligned}$$

□

**Teorema 1.2.5.** Si  $\{X_n : n \geq 1\}$  es una submartingala no-negativa, entonces para todos  $a, k > 0$ , se tiene que:

$$P(\{\max(X_1, X_2, \dots, X_k) > a\}) \leq \frac{E[X_k]}{a}.$$

*Demostración.* Se define al tiempo de paro

$$N = \begin{cases} \min\{j \geq 0; X_j > a\} & \text{si } X_j > a, \text{ para algún } j = 0, 1, \dots, k, \\ k & \text{si } X_j \leq a, \text{ para } j = 0, 1, \dots, k. \end{cases}$$

Note que  $\max(X_1, X_2, \dots, X_k) > a$  es equivalente a  $X_N > a$ . Por lo tanto,

$$\begin{aligned} P(\{\max(X_1, X_2, \dots, X_k) > a\}) &= P(X_N > a) \\ &\leq \frac{E[X_N]}{a} \\ &\leq \frac{E[X_k]}{a}, \end{aligned}$$

pués  $N \leq k$ .

□

**Corolario 1.2.6.** Sea  $\{X_n : n \geq 1\}$  una martingala. Entonces para  $a > 0$ , se satisface que:

1.

$$P(\{\max(|X_1|, \dots, |X_k|) > a\}) \leq \frac{E[|X_k|]}{a},$$

2.

$$P(\{\max(|X_1|, \dots, |X_k|) > a\}) \leq \frac{E[|X_k|^2]}{a^2}.$$

*Demostración.* Para probar la parte 1, considérese la función  $\varphi(x) = |x|$  con  $x \in \mathbb{R}$  la cual es convexa. En este caso se tiene, debido a que  $\{X_n, n \geq 1\}$  es una martingala y usando el Lema 1.2.4, se tiene que  $\{|X_n|, n \geq 1\}$  es una submartingala no-negativa. Por el Teorema 1.2.5 se obtiene que

$$P(\{\max(|X_1|, \dots, |X_k|) > a\}) \leq \frac{E[|X_k|]}{a}.$$

Para demostrar el inciso 2, considérese la función  $\psi(x) = x^2$  con  $x \in \mathbb{R}$  la cual es convexa. Por otro lado, se tiene que  $\{X_n, n \geq 1\}$  es una martingala, se sigue de nuevo por el Lema 1.2.4 que  $\{X_n^2, n \geq 1\}$  es una submartingala no-negativa. Por lo tanto,

$$P(\{\max(X_1^2, \dots, X_k^2) > a^2\}) \leq \frac{E[X_k^2]}{a^2},$$

es decir

$$P(\{\max(|X_1|, \dots, |X_k|) > a\}) \leq \frac{E[|X_k|^2]}{a^2}.$$

□

A continuación se enuncia y se prueba el Teorema de Convergencia para Martingalas.

**Teorema 1.2.7.** *Si  $\{X_n, n \geq 1\}$  es una martingala tal que para algún  $M < \infty$  se satisface que  $E[|X_n|] \leq M$  para todo  $n$ , entonces con probabilidad 1,  $\lim_{n \rightarrow \infty} X_n$  existe y es finito.*

*Demostración.* El teorema se prueba bajo la suposición de que  $(E[X_n^2])$  está acotada, considérese la función  $f(x) = x^2$  con  $x \in \mathbb{R}$  la cual es convexa, se sigue del Lema 1.2.4 que  $\{X_n^2, n \geq 1\}$  es una submartingala, por lo tanto,  $(E[X_n^2])$  es no-decreciente. Como  $(E[X_n^2])$  está acotada, se sigue que converge cuando  $n \rightarrow \infty$ . Sea

$$\eta := \lim_{n \rightarrow \infty} E[X_n^2].$$

Para probar que el límite de  $X_n$  existe y es finito se probará que  $\{X_n, n \geq 1\}$  es, con probabilidad 1, una sucesión de Cauchy. Es decir, se demuestra que con probabilidad 1

$$|X_{m+k} - X_m| \rightarrow 0,$$

cuando  $k, m \rightarrow \infty$ .

Ahora

$$\begin{aligned} & P\{|X_{m+k} - X_m| > \epsilon, \text{ para algún } k, m \leq n\} \\ & \leq E[(X_{m+n} - X_m)^2] / \epsilon^2 \\ & = E[X_{m+n}^2 - 2X_{m+n}X_m + X_m^2] / \epsilon^2. \end{aligned}$$

Pero

$$\begin{aligned} E[X_{m+n}X_m] &= E[E[X_{m+n}X_m \mid X_m]] \\ &= E[X_m E[X_{m+n} \mid X_m]] \\ &= E[X_m^2]. \end{aligned}$$

Por lo tanto, se tiene que

$$P\{|X_{m+k} - X_m| > \epsilon, \text{ para algún } k \leq n\} \leq (E[X_{m+n}^2] - E[X_m^2]) / \epsilon^2,$$

cuando  $n \rightarrow \infty$  y recordando la definición de  $\eta$  se obtiene

$$P\{|X_{m+k} - X_m| > \epsilon, \text{ para algún } k\} \leq \frac{\eta - E[X_m^2]}{\epsilon^2},$$

y por lo tanto,

$$P\{|X_{m+k} - X_m| > \epsilon, \text{ para algún } k \leq n\} \rightarrow 0,$$

cuando  $m \rightarrow \infty$ . De lo cual se concluye que con probabilidad 1,  $(X_n)$  es una sucesión de Cauchy y por lo tanto,  $\lim_{n \rightarrow \infty} X_n$  existe y es finito. □

El siguiente es un Corolario inmediato del Teorema 1.2.7.

**Corolario 1.2.8.** *Si  $\{X_n, n \geq 1\}$  es una martingala no-negativa, entonces con probabilidad 1,  $\lim_{n \rightarrow \infty} X_n$ , existe y es finito.*

## Capítulo 2

# Proceso de Galton-Watson

En este capítulo se hace un análisis del Proceso de Galton-Watson (PG-W) con tiempo y espacio de estados discreto, se da una definición formal del proceso, además se trabaja con sus funciones generadoras de probabilidades (f.g.p.) y sus momentos, se dan dos ejemplos, uno de ellos se resuelve mediante la iteración de la f.g.p. y en el último se usa un teorema el cual permite calcular la probabilidad de extinción, cabe mencionar que para el primer método se realizó una simulación con el software Mathematica. El Proceso de Galton-Watson también se puede trabajar con espacio de estados discreto y tiempo continuo el cual es conocido con el nombre de Proceso de Ramificación a Tiempo Continuo (véase Apéndice B), el análisis de este proceso no es tan simple, pero se pueden obtener sus momentos y la probabilidad de extinción (véase [9]).

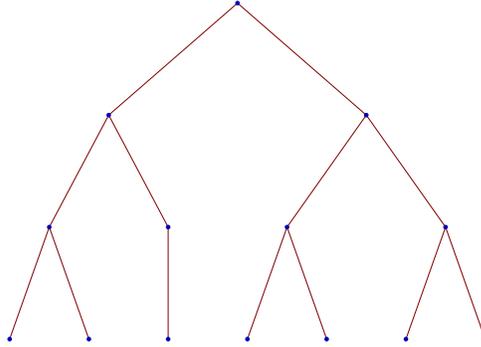
### 2.1. Análisis del Proceso de Galton-Watson

Para introducir el Proceso de Galton-Watson, considérese un proceso que inicia con un individuo, el cual constituye la 0-ésima generación. Supóngase que sus hijos forman la primera generación, sus nietos la segunda y así sucesivamente.

Se denota la variable aleatoria  $Z_n$  (véase Figura 2.1.1) como el número de individuos en la  $n$ -ésima generación de una población, se supone que  $Z_0 = 1$ , a menos que se establezca lo contrario y

$$Z_n = \sum_{i=1}^{Z_{n-1}} X_{n-1,i}, \quad (2.1.1)$$

para todo  $n \in \mathbb{N}$ , donde  $X_{n,i}$  es una familia de variables aleatorias e indica el número de descendientes del  $i$ -ésimo individuo en la  $n$ -ésima generación con  $P[X_{n,i} = k] = p_k$  para cualquier  $k, i \in \mathbb{N}$ , donde  $p_0, p_1, \dots \in [0, 1]$  tales que  $p_k \geq 0$  y  $\sum_{k=0}^{\infty} p_k = 1$ . En este caso  $p_k$  es la probabilidad de que un individuo de la  $n$ -ésima generación tenga  $k$  hijos. Al proceso  $\{Z_n\}$  se le llamará Proceso de Galton-Watson.

Figura 2.1.1: Proceso  $Z_n$ 

Con toda la información anterior se puede obtener una probabilidad de transición para el proceso, la cual se denota con  $P(y, z)$  y representa la probabilidad de que en la  $(n + 1)$ -ésima generación haya  $z$  individuos dado que en la generación pasada ( $n$ -ésima) la población total era de  $y$  individuos. La probabilidad de transición para toda  $y, z, n \in \{0, 1, 2, \dots\}$  esta dada por

$$\begin{aligned} P(y, z) &= P(Z_{n+1} = z \mid Z_n = y) \\ &= P\left(\sum_{i=1}^y X_{n,i} = z\right) := p_z^{*y}, \end{aligned}$$

donde  $p_z^{*y}$  indica la  $z$ -ésima componente de la  $y$ -ésima convolución de  $p$ .

*Observación 2.1.1.* En este caso la matriz de transición está dada por

$$\begin{bmatrix} 1 & 0 & 0 & \dots \\ p_0^{*1} & p_1^{*1} & p_2^{*1} & \dots \\ p_0^{*2} & p_1^{*2} & p_2^{*2} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

A continuación se da un ejemplo en el cual se ve de forma explícita la matriz de transición.

**Ejemplo 2.1.2.** Considérese un proceso de Galton-Watson cuyas variables aleatorias  $X_{n,i} \sim Poi(\lambda)$  con  $\lambda > 0$ . Primero se encuentran las probabilidades de transición para que posteriormente se pueda formar la matriz de transición.

Se tiene que

$$\begin{aligned}
 P(0,0) &= 1 \text{ y para } P(0,n) = 0, n \geq 1, \\
 P(1,0) &= P\left(\sum_{i=1}^1 X_{n,i} = 0\right) = e^{-\lambda} \text{ y para } P(1,n) = e^{-\lambda} \frac{\lambda^n}{n!}, n \geq 1, \\
 P(2,0) &= P\left(\sum_{i=1}^2 X_{n,i} = 0\right) = e^{-(\lambda_1+\lambda_2)} \text{ y para } P(2,n) = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1 + \lambda_2)^n}{n!}, \\
 &\quad n \geq 1,
 \end{aligned}$$

se prosigue de la misma forma para encontrar las demás probabilidades de transición. Así, se tiene que la matriz de transición para este ejemplo está dada por:

$$\begin{bmatrix}
 1 & 0 & 0 & \dots \\
 e^{-\lambda} & e^{-\lambda} \lambda & e^{-\lambda} \frac{\lambda^2}{2} & \dots \\
 e^{-(\lambda_1+\lambda_2)} & e^{-(\lambda_1+\lambda_2)} (\lambda_1 + \lambda_2) & e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^2}{2} & \dots \\
 \vdots & \vdots & \vdots & \vdots
 \end{bmatrix}.$$

**Lema 2.1.3.** *El proceso  $(Z_n)_{n \in \mathbb{N}_0}$  es una Cadena de Markov.*

La demostración del Lema 2.1.3 es directa de la definición de Cadena de Markov.

## 2.2. La f.g.p. para el Proceso de Galton-Watson

Una herramienta importante para la investigación de los procesos de ramificación son las funciones generadoras de probabilidades.

**Definición 2.2.1.** La función generadora de probabilidades de la variable aleatoria  $X_{01} := X$  se define como:

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

donde  $|s| \leq 1$ .

De esta forma la función generadora de probabilidades para la variable aleatoria  $Z_n$  está dada por

$$\psi_n(s) = \sum_{k=0}^{\infty} P(Z_n = k) s^k,$$

para todo  $n \in \mathbb{N}_0$ .

Se supone que  $P(Z_0 = 1) = 1$ , de esta forma se obtiene la siguiente relación

$$\psi_0(s) = \sum_{k=0}^{\infty} P(Z_0 = k) s^k = s.$$

Notando que la variable aleatoria  $Z_1$  se distribuye igual que la variable aleatoria  $X$ , se obtiene que

$$\psi_1(s) = \sum_{k=0}^{\infty} P(Z_1 = k) s^k = \psi(s).$$

Las iteraciones de la función generadora  $\psi(s)$  están definidas por

$$\psi_0(s) = s, \quad \psi_1(s) = \psi(s), \quad (2.2.1)$$

$$\psi_{n+1}(s) = \psi[\psi_n(s)], \quad n = 1, 2, \dots \quad (2.2.2)$$

Se puede verificar que cada una de las iteraciones son una función generadora de probabilidades, y que las siguientes relaciones son una consecuencia de (2.2.1) y (2.2.2)

$$\psi_{m+n}(s) = \psi_m[\psi_n(s)], \quad m, n = 0, 1, 2, \dots,$$

y en particular,

$$\psi_{n+1}(s) = \psi_n[\psi(s)]. \quad (2.2.3)$$

Ahora, iterando la ecuación (2.2.3) se obtiene el siguiente resultado:

$$\begin{aligned} \psi_{n+1}(s) &= \psi_n(\psi(s)) = \psi_{n-1}(\psi(\psi(s))) = \psi_{n-1}(\psi_2(s)) \\ &= \psi_{n-2}(\psi(\psi_2(s))) = \psi_{n-2}(\psi_3(s)). \end{aligned}$$

Siguiendo con este procedimiento por inducción se obtiene que para cualquier  $k \in \{0, 1, 2, \dots, n\}$

$$\psi_{n+1}(s) = \psi_{n-k}(\psi_{k+1}(s)),$$

y si  $k = n - 1$  se tiene

$$\psi_{n+1}(s) = \psi(\psi_n(s)). \quad (2.2.4)$$

A lo largo de este capítulo se toma en cuenta la siguiente suposición: Ninguna de las probabilidades  $p_0, p_1, \dots$  es igual a 1 y  $p_0 + p_1 < 1$ . Por lo tanto, la f.g.p.  $\psi$  es estrictamente convexa en el intervalo unitario.

Note que  $Z_n$  (véase (2.1.1)) es medible ya que:

$$\{Z_n = k\} = \bigcup_{i=0}^{\infty} \{Z_{n-1} = i\} \cap \{X_{n-1,1} + X_{n-1,2} + \dots + X_{n-1,Z_{n-1}} = k\}.$$

**Lema 2.2.2.** *Si las variables aleatorias  $X_{n-1,1}, X_{n-1,2}, \dots$ , son i.i.d entonces la función generadora de probabilidades de  $Z_n$  está dada por*

$$\psi_{Z_n}(s) = \psi_{Z_{n-1}}(\psi(s)).$$

*Demostración.*

$$\begin{aligned}
\psi_{Z_n}(s) &= \sum_{k=0}^{\infty} P[Z_n = k] s^k \\
&= \sum_{k=0}^{\infty} P \left[ \bigcup_{i=0}^{\infty} \{Z_{n-1} = i\} \cap \{X_{n-1,1} + X_{n-1,2} + \dots + X_{n-1,Z_{n-1}} = k\} \right] s^k \\
&= \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} P[Z_{n-1} = i] P[X_{n-1,1} + X_{n-1,2} + \dots + X_{n-1,Z_{n-1}} = k] s^k \\
&= \sum_{i=0}^{\infty} P[Z_{n-1} = i] \sum_{k=0}^{\infty} P[X_{n-1,1} + X_{n-1,2} + \dots + X_{n-1,Z_{n-1}} = k] s^k \\
&= \sum_{i=0}^{\infty} P[Z_{n-1} = i] \psi(s) = \sum_{i=0}^{\infty} P[Z_{n-1} = i] (\psi(s))^i \\
&= \psi_{Z_{n-1}}(\psi(s)).
\end{aligned}$$

□

Del Lema 2.2.2 se puede obtener el siguiente resultado.

**Teorema 2.2.3.**  $\psi_{Z_n}(s) = \psi_n(s)$ , para todo  $n \in \mathbb{N}$ . Es decir, la f.g.p. de  $Z_n$  es la  $n$ -ésima iteración de  $\psi(s)$ .

*Demostración.* Para  $n = 1$ , la afirmación es verdadera por definición. Para  $n \in \mathbb{N}$ , se concluye por el Lema 2.2.2 que  $\psi_{Z_{n+1}} = \psi \circ \psi_{Z_n} = \psi \circ \psi_n = \psi_{n+1}$ .

□

### 2.3. Los momentos del PG-W

Los momentos del proceso, cuando existen, pueden ser expresados en términos de las derivadas de  $\psi(s)$  en  $s = 1$ .

**Lema 2.3.1.** *Sea*

$$m = E[X] = \sum_{k=0}^{\infty} k p_k = \psi'(1),$$

donde  $m$  es el promedio de reproducción,  $\sigma^2 = \text{Var}[X]$  y supóngase que ambas son finitas. Entonces

$$E[Z_n] = m^n \text{ y } \text{Var}[Z_n] = \begin{cases} \sigma^2 m^{n-1} \frac{m^n - 1}{m - 1}, & m \neq 1, \\ n\sigma^2, & m = 1. \end{cases}$$

*Demostración.* Nótese que  $\psi'(1) = m$  y  $\sigma^2 = \psi''(1) + m - m^2$ . Derivando la ecuación (2.2.3) se tiene

$$\psi'_{n+1}(s) = \psi'_n(\psi(s))\psi'(s),$$

y si se toma a  $s = 1$  se obtiene

$$\psi'_{n+1}(1) = \psi'_n(1)\psi'(1).$$

Iterando de nuevo esta última relación se tiene que

$$\psi'_{n+1}(1) = \psi'_n(1)\psi'(1) = \psi'_{n-1}(1)[\psi'(1)]^2 = \psi'_{n-2}(1)[\psi'(1)]^3$$

hasta llegar a que

$$\psi'_{n+1}(1) = [\psi'(1)]^n \psi'(1) = [\psi'(1)]^{n+1}.$$

Como la primera derivada de la función generadora de  $X$  evaluada en uno es igual a  $m$  entonces se tiene que

$$E[Z_{n+1}] = m^{n+1}.$$

Para obtener la  $Var[Z_{n+1}]$  nótese que

$$\psi''_n(1) = \sum_{k=2}^{\infty} k(k-1)P(Z_n = k) = E[Z_n^2] - E[Z_n] = E[Z_n^2] - \psi'_n(1),$$

de esta forma véase que

$$Var[Z_n] = \psi''_n(1) + \psi'_n(1) - (\psi'_n(1))^2. \quad (2.3.1)$$

Ahora, derivando dos veces la ecuación (2.2.4) se obtiene la siguiente relación:

$$\psi''_{n+1}(s) = \psi''(\psi_n(s))[\psi'_n(s)]^2 + \psi'(\psi_n(s))\psi''_n(s),$$

evaluando en  $s = 1$  se tiene

$$\psi''_{n+1}(1) = \psi''(1)[\psi'_n(1)]^2 + \psi'(1)\psi''_n(1),$$

sustituyendo

$$\begin{aligned} \psi''_{n+1}(1) &= (\sigma^2 + m^2 - m)m^{2n} + m\psi''_n(1) \\ &= (\sigma^2 + m^2 - m)m^{2n} + m[(\sigma^2 + m^2 - m)m^{2n-1} + m\psi''_{n-1}(1)] \\ &= (\sigma^2 + m^2 - m)(m^{2n} + m^{2n-1}) \\ &\quad + m^2[\sigma^2 + m^2 - m)m^{2(n-2)} + m\psi''_{n-2}(1)] \\ &= (\sigma^2 + m^2 - m)(m^{2n} + m^{2n-1} + m^{2n-2}) + m^3\psi''_{n-2}(1), \end{aligned}$$

siguiendo la iteración por inducción se llega a

$$\psi''_{n+1}(1) = (\sigma^2 + m^2 - m)(m^{2n} + m^{2n-1} + m^{2n-2} + \dots + m^n).$$

Ahora, sustituyendo en (2.3.1) se obtiene

$$\begin{aligned}
\text{var}[Z_{n+1}] &= (\sigma^2 + m^2 - m)(m^{2n} + m^{2n-1} + m^{2n-2} + \dots + m^n) \\
&\quad + m^{n+1} - m^{2n+2} \\
&= \sigma^2(m^{2n} + m^{2n-1} + m^{2n-2} + \dots + m^n) \\
&\quad + (m^{2n+2} + m^{2n+1} + m^{2n} + \dots + m^{n+2}) \\
&\quad - (m^{2n+1} + m^{2n} + m^{2n-1} + \dots + m^{n+1}) + m^{n+1} - m^{2n+2} \\
&= \sigma^2(m^{2n} + m^{2n-1} + m^{2n-2} + \dots + m^n) \\
&= \sigma^2 m^n (m^n + m^{n-1} + m^{n-2} + \dots + 1).
\end{aligned}$$

Y por lo tanto se concluye que

$$\text{Var}[Z_{n+1}] = \begin{cases} \sigma^2 m^n \frac{m^{n+1}-1}{m-1}, & m \neq 1, \\ \sigma^2(n+1), & m = 1. \end{cases}$$

□

De acuerdo al valor del promedio de reproducción  $m$ , un Proceso de Galton-Watson se puede clasificar de la forma siguiente:

- a) Subcrítico, si  $m < 1$ .
- b) Crítico, si  $m = 1$ .
- c) Supercrítico, si  $m > 1$ .

En el Apéndice A se analiza el comportamiento asintótico de cada uno de los casos anteriores.

## 2.4. La probabilidad de extinción

Considérese el problema originalmente propuesto por Galton, el cual fue encontrar la probabilidad de extinción del proceso  $\{Z_n : n \in \mathbb{N}\}$ . Para ello defínase el evento

$$[\text{extinción}] = \bigcup_{n=1}^{\infty} [Z_n = 0].$$

A continuación se calcula

$$q := P[\text{extinción}],$$

como  $\{Z_n = 0\} \subset \{Z_{n+1} = 0\}$ , se tiene por la propiedad de continuidad que

$$\begin{aligned}
q &= P\left[\bigcup_{n=1}^{\infty} [Z_n = 0]\right] \\
&= \lim_{n \rightarrow \infty} P[Z_n = 0] \\
&= \lim_{n \rightarrow \infty} \psi_n(0) = \lim_{n \rightarrow \infty} q_n, \text{ donde } q_n := \psi_n(0) \\
&= \lim_{n \rightarrow \infty} P[\text{extinción en o antes de la generación } n].
\end{aligned}$$

**Teorema 2.4.1.** *Se tiene que para  $0 \leq s < 1$*

$$\lim_{n \rightarrow \infty} \psi_n(s) = q.$$

*Demostración.* Supóngase que  $s \leq q$ , entonces

$$\psi(s) \leq \psi(q) = q \text{ y } \psi(s) \leq q.$$

Por lo tanto,

$$\psi_2(s) = \psi(\psi(s)) \leq \psi(q) = q,$$

entonces

$$\psi_2(s) \leq q.$$

En general, continuando con este procedimiento se obtiene

$$\psi_n(s) \leq q,$$

de lo cual, para  $0 \leq s \leq q$

$$q_n = \psi_n(0) \leq \psi_n(s) \leq q,$$

y como  $q_n \rightarrow q$ , cuando  $n \rightarrow \infty$  se obtiene que

$$\psi_n(s) \rightarrow q.$$

Para finalizar este teorema se debe suponer que,  $q < 1$  para ver qué pasa en el dominio  $q \leq s < 1$ , para este rango se tiene

$$q = \psi(q) \leq \psi(s) \leq s.$$

La última desigualdad se sigue del hecho de que la gráfica de  $y = \psi(s)$  está bajo la diagonal del conjunto  $\{s : q \leq s < 1\}$ . Ya que  $\psi(s)$  es no-decreciente, de las desigualdades previas se obtiene

$$\psi(q) = q \leq \psi_2(s) \leq \psi(s) \leq s,$$

continuando de esta forma, para el caso general se obtiene

$$q \leq \psi_n(s) \leq \psi_{n-1}(s) \leq \dots \leq \psi(s) \leq s.$$

Se concluye que  $\psi_n(s)$  es no-creciente para  $q \leq s \leq 1$ .

Sea  $\alpha := \lim_{n \rightarrow \infty} \psi_n(s)$ . Supóngase que para algún  $s_0 \in (q, 1)$  se tiene que  $\alpha > q$ , entonces

$$\psi(\alpha) = \lim_{n \rightarrow \infty} \psi(\psi_n(s_0)) = \lim_{n \rightarrow \infty} \psi_{n+1}(s_0) = \alpha,$$

y en el dominio  $(q, 1)$  se tiene que  $\psi(s) \leq s$ .

Si  $\psi(s)$  es lineal entonces  $\psi(s) = s$  en  $(q, 1)$ , pero entonces se debería tener que  $p_k = 0$  para  $k \geq 2$  y  $m \leq 1$  entonces  $q = 1$ , lo cual contradice la suposición de que  $q < 1$ . Si  $\psi(s)$  es no lineal, entonces por convexidad,  $\psi(s) < s$  para  $s \in (q, 1)$ .

Por lo tanto, se tiene que

$$q < \alpha \leq \psi(s_0) < s_0 < 1,$$

pero también

$$\alpha = \psi(\alpha),$$

lo cual contradice el hecho de que no hay soluciones de  $s = \psi(s)$  en  $(q, 1)$ , esta contradicción surge debido a la suposición de que  $\alpha > q$ .  $\square$

En seguida se da un método para calcular la probabilidad de extinción  $q$  sin tener que calcular las iteraciones de la f.g.p.  $\psi(s)$ . Dicho método para resolver lo anterior está dado en el siguiente Teorema 2.4.2. Se supone que  $0 < p_0 < 1$  (si  $p_0 = 0$  entonces  $q = 0$  y si  $p_0 = 1$  entonces  $q = 1$ ).

**Teorema 2.4.2.** *Si  $m = E(Z_1) \leq 1$  entonces  $q = 1$ . Si  $m > 1$ , entonces  $q < 1$  y es la única solución no-negativa para la ecuación*

$$s = \psi(s), \tag{2.4.1}$$

$s < 1$ .

*Demostración.* PASO1.- Primero se prueba que  $q$  es una solución de la ecuación  $s = \psi(s)$ , como

$$\{Z_n = 0\} \subset \{Z_{n+1} = 0\},$$

se tiene que  $q_n = P[Z_n = 0]$  es una sucesión no-decreciente la cual converge a un número real  $q \in [0, 1]$ .

Como

$$\psi_{n+1}(s) = \psi(\psi_n(s)),$$

tomando a  $s = 0$  se tiene que

$$\psi_{n+1}(0) = \psi(\psi_n(0)),$$

$$q_{n+1} = \psi(q_n),$$

cuando  $n \rightarrow \infty$  y usando la continuidad de  $\psi(s)$ , se obtiene

$$q = \psi(q).$$

PASO 2.- Se muestra que  $q$  es la mínima solución de  $s = \psi(s)$  en  $[0, 1]$ .

Supóngase que  $r$  es cualquier solución de la ecuación  $s = \psi(s)$ , satisfaciendo  $0 \leq r \leq 1$ , entonces como

$$0 \leq r,$$

$$q_1 = \psi(0) \leq \psi(r) = r,$$

y por lo tanto,

$$q_2 = \psi_2(0) \leq \psi(r) = r.$$

En general se obtiene que

$$q_n \leq r,$$

tendiendo  $n$  a  $\infty$ , se tiene que  $q \leq r$ , mostrando de esta forma que  $q$  es la solución mínima en  $[0, 1]$ .

PASO 3.- Como  $p_k$  es la densidad de  $Z_1$ , se tiene que

$$\psi(s) = p_0 + p_1s + p_2s^2 + \dots$$

y

$$\psi'(s) = p_1 + 2p_2s + \dots$$

Por lo tanto

$$\psi(0) = p_0, \psi(1) = 1,$$

y

$$\lim_{s \rightarrow 1} \psi'(s) = p_1 + 2p_2 + \dots = m.$$

Supóngase primero que  $m < 1$ . Entonces

$$\lim_{s \rightarrow 1} \psi'(s) < 1,$$

como  $\psi'(s)$  es no decreciente en  $s$ ,  $0 \leq s < 1$ , se concluye que  $\psi'(s) < 1$  para  $0 \leq s < 1$ .

Supóngase que  $m = 1$  y  $P(Z_1 = 1) < 1$  entonces  $P(Z_1 = n) > 0$  para algún  $n > 2$ . Por lo tanto,  $\psi'(s)$  es estrictamente creciente en  $s$ ,  $0 \leq s < 1$ . Como

$$\lim_{s \rightarrow 1} \psi'(s) = 1,$$

nuevamente se concluye que  $\psi'(s) < 1$ , para  $0 \leq s < 1$ .

Ahora supóngase que  $m \leq 1$  y  $P(Z_1 = 1) < 1$ . Se ha demostrado que  $\psi'(s) < 1$  para  $0 \leq s < 1$ . Por lo tanto,

$$\frac{d}{ds} (\psi(s) - s) < 0,$$

$0 \leq s < 1$ , como  $\psi(1) - 1 = 0$  se tiene que  $\psi(s) - s > 0$ ,  $0 \leq s < 1$ , y por lo tanto, (2.4.1) no tiene raíces en  $[0, 1]$ . De lo cual se concluye que  $q = 1$ .

Ahora supóngase que  $m > 1$ , entonces

$$\lim_{s \rightarrow 1} \psi'(s) > 1,$$

por la continuidad de  $\psi'$ , existe  $s_0$  en el intervalo  $(0, 1)$  tal que para toda  $s \in (s_0, 1)$ ,  $\psi'(s) > 1$  y por el Teorema del Valor Medio se tiene que

$$\frac{\psi(1) - \psi(s_0)}{1 - s_0} > 1,$$

como  $\psi(1) = 1$ , se concluye que  $\psi(s_0) - s_0 < 0$ .

Ahora,  $\psi(s) - s$  es continua en  $s$  y no-negativa en  $s = 0$ , entonces por el Teorema del Valor Intermedio se va a tener un cero en  $q \in [0, s_0)$ . Por lo tanto, se afirma que se tiene una raíz en  $q$ .

Por último se demostrará que existe una sola raíz. Supóngase que existe otra raíz  $q_1$  entonces la función  $\psi(s) - s$  vale cero en  $q$ , en  $q_1$  y en 1. Por el Teorema de Rolle la primera derivada tiene al menos dos raíces en  $(0, 1)$ . Aplicando nuevamente el Teorema de Rolle su segunda derivada  $\psi''(s)$  tiene al menos una raíz en  $(0, 1)$ .

Pero si  $m > 1$ , entonces al menos uno de los números  $p_2, p_3, \dots$ , es estrictamente positivo y por lo tanto

$$\psi''(s) = 2p_2 + 6p_3s + \dots,$$

no tiene raíces en  $[0, 1)$ . Esta contradicción muestra que  $s = \psi(s)$  tiene una única raíz en  $[0, 1)$ .

□

*Observación 2.4.3.* Se puede dar el caso en el que las hipótesis del Teorema de Banach del punto fijo se cumplan, es decir, se necesita saber que la función generadora de probabilidades  $\psi$  es una función contráctil ( $|\psi(x) - \psi(y)| \leq \alpha|x - y|$  para algún  $\alpha \in (0, 1)$  y  $x, y \in \text{Dom } \psi$ ), entonces de esta forma se tiene por el Teorema de Banach que  $\psi$  admite un único punto fijo.

## 2.5. Ejemplos

Los siguientes ejemplos son en base a lo que se ha visto hasta este momento, el objetivo es mostrar una aplicación de los Teoremas 2.4.1 y 2.4.2.

**Ejemplo 2.5.1.** Considérese una población que se desarrolla según un Proceso de Galton-Watson, bajo una ley de reproducción inicial Geométrica con parámetro  $r \in (0, 1)$ . El primer método que se utiliza es el del Teorema 2.4.1, entonces se empieza por calcular la f.g.p. para la distribución Geométrica (véase 2.5.1),

$$\psi(s) = \frac{r}{1 - s(1 - r)}. \quad (2.5.1)$$

Posteriormente, se calculan las iteraciones de la f.g.p., para dicho cálculo se implementó un programa en el software Mathematica (véase Apéndice C), cuyo algoritmo que se utiliza para su realización es el siguiente:

1. Dar un valor y error inicial  $(x_0, e_0)$ .
2. Generar las iteraciones en base a la f.g.p. ( $x_i = F[x_{i-1}]$ ).
3. Obtener los errores de aproximación.
4. Si el error obtenido es menor que el error inicial, se imprime a  $(x_i)$  y termina.

5. Si no, regresar al paso 2.

Cabe mencionar que dicho programa además de calcular las iteraciones de la f.g.p. también calcula la probabilidad de extinción.

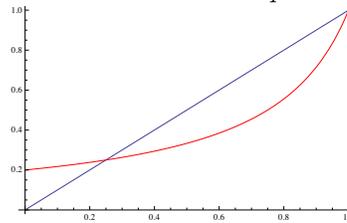
Los datos que se obtuvieron cuando se implementó el algoritmo antes mencionado al software Mathematica son (véase Cuadro 2.1):

Cuadro 2.1: Iteración de la f.g.p. para la distribución Geométrica

r	q
0.2	0.2380
0.5	0.99012
0.7	0.99995
0.9	0.99999

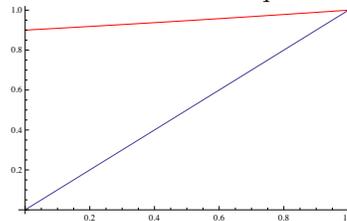
Posteriormente se muestra la gráfica que representa a la función generadora de probabilidades de la distribución Geométrica cuando  $r = 0.2$  (véase Figura 2.5.1), donde la línea azul es la función identidad y la roja es la gráfica de la f.g.p..

Figura 2.5.1: Probabilidad de extinción  $q = 0.2380$ , cuando  $r = 0.2$



Para el caso cuando  $r = 0.9$  la gráfica de la f.g.p. es la siguiente (véase Figura 2.5.2):

Figura 2.5.2: Probabilidad de extinción  $q = 0.99999$ , cuando  $r = 0.9$



Enseguida se va a calcular la probabilidad de extinción usando los mismos datos, con la diferencia de que ahora se aplica el resultado del Teorema 2.4.2, para ello se procede de la siguiente forma:

- Primero se calcula  $E(Z_1) = m$  para la distribución Geométrica( $r$ ) donde  $r$  toma los valores 0.2, 0.5, 0.7 ó 0.9.
- Una vez calculada  $m$ , se aplica el Teorema 2.4.2 es decir, si  $m \leq 1$  entonces  $q = 1$  y si  $m > 1$  se resuelve la ecuación  $s = \psi(s)$ , conociendo que la f.g.p. es la misma que (2.5.1) para la distribución Geométrica.

Se tiene el siguiente Cuadro 2.2 en el cual se ilustra lo descrito anteriormente.

Cuadro 2.2: Teorema 2.4.2

r	m	q
0.2	4	0.25
0.5	1	1
0.7	.42	1
0.9	.11	1

Las gráficas de la f.g.p. en este caso serían las mismas que se presentaron anteriormente, cuando  $r = 0.2$  (véase Figura 2.5.1) se puede observar en esta figura que la gráfica de la f.g.p. esta por abajo de la función identidad y que hay una raíz, caso contrario cuando  $r = 0.9$  (véase Figura 2.5.2) donde se ve que la gráfica de la f.g.p. está por arriba de la función identidad.

Se tiene que utilizando los dos métodos para calcular la probabilidad de extinción  $q$ , es decir en base al cálculo de las iteraciones de la f.g.p. (Teorema 2.4.1) y usando el Teorema 2.4.2 se puede observar que el valor de la probabilidad de extinción para los dos casos son muy aproximados.

**Ejemplo 2.5.2.** El siguiente ejemplo es análogo al Ejemplo 2.5.1, con la diferencia de que ahora se considera una población cuya ley de reproducción inicial es Poisson( $\lambda$ ) donde  $\lambda \in (0, \infty)$ . Se tiene que la f.g.p. para la distribución Poisson es la siguiente (véase 2.5.2)

$$\psi(s) = e^{-\lambda(1-s)}. \quad (2.5.2)$$

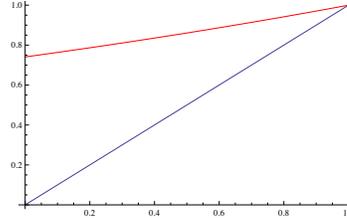
Se calcula la probabilidad de extinción usando el método de las iteraciones de la f.g.p. para encontrar a éstas se usa el algoritmo dado en el Ejemplo 2.5.1 y se obtienen los siguientes resultados para la probabilidad de extinción  $q$  (véase Cuadro 2.3):

Cuadro 2.3: Iteración de la f.g.p. para la Poisson

$\lambda$	q
0.3	0.99997
2	0.203256
5	0.006977

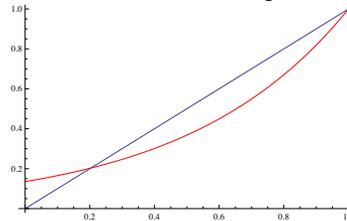
Se muestra la gráfica de la f.g.p. para la distribución Poisson cuando  $\lambda = 0.3$  (véase Figura 2.5.3), en la cual la línea azul representa la función identidad y la roja es la gráfica de la f.g.p.,

Figura 2.5.3: Probabilidad de extinción  $q = 0.99997$ , cuando  $\lambda = 0.3$



la gráfica de la f.g.p. cuando  $\lambda = 2$  es la siguiente (véase Figura 2.5.4).

Figura 2.5.4: Probabilidad de extinción  $q = 0.203256$ , cuando  $\lambda = 2$



Enseguida se calcula la probabilidad de extinción usando los mismos datos, con la diferencia de que ahora se aplica el resultado del Teorema 2.4.2, se procede de la siguiente forma:

- Primero se calcula  $E(Z_1) = m$  para la distribución Poisson( $\lambda$ ) donde  $\lambda = 0.3, 2$  ó  $5$ .
- Una vez calculada  $m$ , se aplica el Teorema 2.4.2 es decir si  $m \leq 1$  entonces  $q = 1$  y si  $m > 1$  se resuelve la ecuación  $s = \psi(s)$ , conociendo que la f.g.p. es la misma que (2.5.2).

Como consecuencia de lo anterior se tiene el siguiente Cuadro 2.4

Cuadro 2.4: Teorema 2.4.2

$\lambda$	$m$	$q$
0.3	0.3	1
2	2	0.203188
5	5	0.006977

Se puede concluir como en el Ejemplo 2.5.1 que los valores de la probabilidad de extinción en base a los dos métodos son muy aproximados.

Se observa que para el caso del Ejemplo 2.5.1 en el cual se trabaja con la distribución Geométrica, calcular la probabilidad de extinción sería más fácil usando el Teorema 2.4.2 ya que la ecuación que se debe resolver es una función cuadrática. Sin embargo, para el Ejemplo 2.5.2 para la distribución Poisson utilizar este método manualmente sería más complicado y es necesario recurrir a la implementación de algoritmos computacionales.



## Capítulo 3

# Estimación

La teoría de muestreo puede emplearse para obtener información acerca de muestras obtenidas aleatoriamente de una población conocida. Sin embargo, desde un punto de vista práctico, suele ser más importante el ser capaz de inferir información acerca de una población a partir de muestras de ellas. Dichos problemas son tratados por la inferencia estadística que utiliza principios de muestreo. Un problema importante de la inferencia estadística es la estimación de parámetros poblacionales o simplemente parámetros (como la media y la varianza poblacionales), a partir de los estadísticos muestrales correspondientes o estadísticos (como la media y varianza muestral). Entonces se dice que la estimación es un conjunto de técnicas que permiten dar un valor aproximado de alguno de los parámetros de una población a partir de los datos proporcionados por una muestra.

Un posible enfoque consiste en considerar que la función de densidad que se desea estimar pertenece a una determinada clase de funciones paramétricas, por ejemplo a algunas de las distribuciones clásicas: Normal, Exponencial, Poisson, etc. Dicha suposición usualmente se basa en informaciones sobre la variable que son externas a la muestra, pero cuya validez puede ser comprobada con posterioridad mediante pruebas de bondad de ajuste. Bajo esta suposición la estimación se reduce a determinar el valor de los parámetros del modelo a partir de la muestra. Esta estimación es la que se denomina estimación paramétrica de la densidad.

En los métodos no paramétricos (véase [22]) se hace inferencia acerca de la distribución o los parámetros de la cual provienen las observaciones, sin hacer suposiciones especiales acerca de la forma de la distribución, es decir, la estimación no paramétrica permite que los datos determinen de forma totalmente libre, sin restricciones, la forma de la densidad que los ha de representar.

### 3.1. Estimación de los parámetros del PG-W

El propósito de este capítulo es considerar el problema de estimación de los diversos parámetros del Proceso de Galton-Watson, a saber: estimación de la media

$m$ , estimación de la probabilidad de extinción  $q$  y la probabilidad de descendientes  $p_k$ .

El Proceso de Galton-Watson se caracteriza por la distribución de descendientes  $p_k$  y por el promedio de reproducción  $m$ ; para calcularlos se usa la estimación vía máxima verosimilitud.

Sea  $p_k$ ,  $k = 0, 1, \dots$  la distribución de descendientes, supongáse que se observa el proceso  $Z_n$  hasta la generación  $n$ , es decir, las variables aleatorias:

$$\{X_{li}, l = 0, 1, \dots, n-1; i = 1, \dots, Z_l\}.$$

Sea

$$Z_l(k) := \sum_{i=1}^{Z_l} I_{\{X_{li}=k\}},$$

$k \geq 0$ ,  $0 \leq l \leq n-1$ , donde  $Z_l(k)$  representa el número de individuos de la  $l$ -ésima generación con exactamente  $k$  descendientes. De esto se deduce que:

$$Z_l = \sum_{k=0}^{\infty} Z_l(k).$$

**Teorema 3.1.1.** *Sea  $\{Z_n\}$  un Proceso de Galton-Watson. El estimador de máxima verosimilitud de  $p_k$ , para  $k \geq 0$ , basado en*

$$\{Z_l(k) : l = 0, 1, \dots, n-1; k = 0, 1, \dots\}$$

es

$$\hat{p}_k := \frac{\sum_{j=0}^{n-1} Z_j(k)}{\sum_{j=0}^{n-1} \sum_{k=0}^{\infty} Z_j(k)}.$$

*Demostración.* La función de verosimilitud basada en las observaciones  $Z_j(k)$ ,  $j = 0, 1, \dots, n-1$  está dada por

$$L = \prod_{j=0}^{n-1} P(Z_{j+1} = z_{j+1} \mid Z_j = z_j),$$

donde la distribución conjunta de  $Z_j(k)$ ,  $k = 0, 1, \dots$  dado  $Z_j$  es

$$P(Z_{j+1} = z_{j+1} \mid Z_j = z_j) = \frac{z_j!}{\prod_{k=0}^{\infty} i_k!} \prod_{k=0}^{\infty} p_k^{i_k},$$

con  $\sum_{k=0}^{\infty} i_k = z_j$ .

Tomando logaritmo a la función de verosimilitud se obtiene

$$\begin{aligned}
\ln L &= \sum_{j=0}^{n-1} \ln \left( \frac{z_j!}{\prod_{k=0}^{\infty} i_k!} \prod_{k=0}^{\infty} p_k^{i_k} \right) \\
&= \sum_{j=0}^{n-1} \left( \ln \left( \frac{z_j!}{\prod_{k=0}^{\infty} i_k!} \right) + \sum_{k=0}^{\infty} (\ln p_k) i_k \right) \\
&= \sum_{j=0}^{n-1} \ln \left( \frac{z_j!}{\prod_{k=0}^{\infty} i_k!} \right) + \sum_{j=0}^{n-1} \sum_{k=0}^{\infty} (\ln p_k) i_k \\
&= \sum_{j=0}^{n-1} \ln \left( \frac{z_j!}{\prod_{k=0}^{\infty} i_k!} \right) + \sum_{j=0}^{n-1} \sum_{k=0}^{\infty} (\ln p_k) z_j(k).
\end{aligned}$$

Para determinar el valor que maximiza la función  $\ln L$  se utiliza el método de los multiplicadores de Lagrange.

Sea

$$f(p) := \sum_{j=0}^{n-1} \sum_{k=0}^{\infty} (\ln p_k) Z_j(k),$$

donde  $p := (p_k : k = 0, 1, \dots)$ , sujeta a:

$$g(p) = \sum_{k=0}^{\infty} p_k = 1,$$

se empieza por construir la función de Lagrange

$$\begin{aligned}
H(p, \lambda) &= f(p) - \lambda g(p) \\
&= \sum_{j=0}^{n-1} \sum_{k=0}^{\infty} (\ln p_k) Z_j(k) - \lambda \left( \sum_{k=0}^{\infty} p_k - 1 \right),
\end{aligned}$$

con  $\lambda \neq 0$ . Derivando la función de Lagrange respecto a  $p$  y  $\lambda$  e igualando a cero se tiene

$$\frac{\partial H(p)}{\partial p} = \sum_{j=0}^{n-1} \frac{Z_j(k)}{p_k} - \lambda = 0 \quad (3.1.1)$$

y

$$\frac{\partial H(p)}{\partial \lambda} = \sum_{k=0}^{\infty} p_k - 1 = 0 \quad (3.1.2)$$

de la ecuación (3.1.1)

$$\frac{1}{p_k} \sum_{j=0}^{n-1} Z_j(k) = \lambda$$

entonces

$$p_k = \frac{\sum_{j=0}^{n-1} Z_j(k)}{\lambda},$$

de la ecuación (3.1.2)

$$\sum_{k=0}^{\infty} p_k = \frac{\sum_{k=0}^{\infty} \sum_{j=0}^{n-1} Z_j(k)}{\lambda} = 1,$$

entonces

$$\lambda = \sum_{k=0}^{\infty} \sum_{j=0}^{n-1} Z_j(k),$$

y por lo tanto

$$\hat{p}_k = \frac{\sum_{j=0}^{n-1} Z_j(k)}{\sum_{j=0}^{n-1} \sum_{k=0}^{\infty} Z_j(k)}.$$

□

*Observación 3.1.2.* El estimador de máxima verosimilitud  $\hat{p}_k$  es un estimador intuitivamente razonable porque estima la probabilidad de que un individuo dé lugar a  $k$  descendientes por la proporción relativa de progenitores con  $k$  descendientes.

Ahora se quiere estimar el promedio de reproducción  $m$  (véase Sección 2.3), para esto nótese que

$$m = \sum_{k=0}^{\infty} k p_k,$$

utilizando  $\hat{p}_k$ , se puede estimar a  $m$ , de la forma siguiente

$$\begin{aligned} \hat{m} &= \sum_{k=0}^{\infty} k \frac{\sum_{j=0}^{n-1} Z_j(k)}{\sum_{j=0}^{n-1} \sum_{k=0}^{\infty} Z_j(k)} \\ &= \frac{\sum_{k=0}^{\infty} \sum_{j=0}^{n-1} k Z_j(k)}{\sum_{j=0}^{n-1} \sum_{k=0}^{\infty} Z_j(k)} \\ &= \frac{\sum_{j=0}^{n-1} Z_{j+1}(k)}{\sum_{j=0}^{n-1} \sum_{k=0}^{\infty} Z_j(k)}, \end{aligned}$$

es decir  $\hat{m}$  es el número total de hijos dividido por el número total de progenitores.

Otro de los parámetros importantes para el Proceso de Galton-Watson es  $q$  el cual representa la probabilidad de extinción.

A continuación se proponen estimadores no paramétricos de la probabilidad de extinción del Proceso de Galton-Watson. Dichos estimadores los cuales son propuestos por Martín-Andrés (véase [13]), Crump y Howe (véase [14]) son comparados mediante simulación.

Sean  $Z_r, Z_{r+1}, \dots, Z_l$  los tamaños de las generaciones sucesivas  $r, r + 1, \dots, l$ , observadas en donde se supone que  $Z_l > 0$ . También se supone que  $m > 1$ , pues de otra forma la población se extingue con probabilidad uno.

Crump y Howe (véase [14]) proponen como estimador no-paramétrico de  $q$  el siguiente:

$$\hat{q}_1 = \exp\{-2(\hat{m} - 1)/\hat{\sigma}^2\}, \quad (3.1.3)$$

con

$$\hat{\sigma}^2 = \frac{1}{l - r - 1} \sum_{k=r+1}^l Z_{k-1} \left( \frac{Z_k}{Z_{k-1}} - \hat{m} \right)^2,$$

donde  $\hat{\sigma}^2$  es un estimador de  $\sigma^2$  que corresponde aproximadamente a una varianza de la muestra. De modo que para conocer a  $\hat{q}_1$ , hacen falta al menos tres generaciones sucesivas, donde  $\hat{m}$  es el estimador del promedio de reproducción que se calculó anteriormente.

Martín-Andrés (véase [13]) propone como estimador de  $q$  (para el caso general en que  $Z_0 = t$ , donde  $t$  indica el número de individuos con el que inicia el proceso), el siguiente

$$\hat{q} = \left( 1 + \frac{\hat{m}^l - 1}{Z_l - p\hat{m}^l} \right)^{-1},$$

y, cuando  $t = 1$ :

$$\hat{q}_2 := (Z_l - \hat{m}^l) / (Z_l - 1). \quad (3.1.4)$$

## 3.2. Comparación de estimadores por simulación

Para la comparación de los estimadores del parámetro  $q$  presentados en la Sección 3.1 se implementó un programa en el software R (véase Apéndice C), para una población que tiene distribución Poisson ( $\lambda$ ) donde  $\lambda = 1.3$  ó  $2$ . Para ello se toman las realizaciones  $Z_r, Z_{r+1}, \dots, Z_l$  las cuales son los tamaños de las generaciones sucesivas  $r, r + 1, \dots, l$ , observadas.

Entonces, tomando una muestra de tamaño  $n = 100$ , para  $r = 6, l = 10$ , conociendo que la población se distribuye de acuerdo a una Poisson con parámetro  $\lambda = 1.3$ , se tiene que  $m = 1.3$  y calculando la probabilidad de extinción usando el Teorema 2.4.2 en el cual se obtuvo que  $q = 0.52$ . Posteriormente se calculan los estimadores  $\hat{q}_1$  (véase (3.1.3)) y  $\hat{q}_2$  (véase (3.1.4)), y finalmente se calcula la raíz del error cuadrado medio para ambos estimadores (véase Cuadro 3.1):

Cuadro 3.1:

		Raíz del error cuadrado medio
$\hat{q}_1$	0.61	0.093
$\hat{q}_2$	0.69	0.174

Para  $r = 10$ ,  $l = 20$ ,  $m = 1.3$  y  $q = 0.52$ , se obtuvieron los valores para los estimadores  $\hat{q}_1$  y  $\hat{q}_2$  con raíz del error cuadrado medio respectivamente mostrados en el Cuadro 3.2:

Cuadro 3.2:

		Raíz del error cuadrado medio
$\hat{q}_1$	0.59	0.075
$\hat{q}_2$	0.88	0.364

Para  $r = 6$ ,  $l = 10$ , cuya población tiene una distribución Poisson con  $\lambda = 2$  se tiene que  $m = 2$  y  $q = 0.36$ , se obtuvieron los siguientes resultados véase Cuadro 3.3:

Cuadro 3.3:

		Raíz del error cuadrado medio
$\hat{q}_1$	0.37	0.0067
$\hat{q}_2$	0.66	0.303

Para  $r = 10$ ,  $l = 20$ ,  $m = 2$  y  $q = 0.36$ , los resultados son los siguientes véase Cuadro 3.4:

Cuadro 3.4:

		Raíz del error cuadrado medio
$\hat{q}_1$	0.44	0.086
$\hat{q}_2$	0.59	0.232

Se puede observar en los cuadros anteriores que el estimador  $\hat{q}_1$  fue mejor estimador que el estimador  $\hat{q}_2$ , al menos para una población que se distribuye de acuerdo a una Poisson con  $\lambda = 1.3$  ó  $2$ .



## Capítulo 4

# Proceso de Galton-Watson Multitipo

Un primer paso en la generalización del Proceso de Galton-Watson es la consideración de los procesos que involucran varios tipos de individuos.

El objetivo de este capítulo es considerar poblaciones con distintos tipos o grupos de individuos, cada uno de estos individuos puede generar descendientes de todos los tipos que coexisten en la población, se hace un desarrollo acerca del Proceso de Galton-Watson Multitipo (PG-WM) y, finalmente se analiza un ejemplo en base a esta teoría.

### 4.1. Análisis del PG-WM

Para analizar el Proceso de Galton-Watson Multitipo es necesario introducir notación, definiciones y algunos resultados importantes.

**Notación:**  $T$  denota el conjunto de todos los vectores  $k$ -dimensionales cuyos componentes son enteros no-negativos.  $e_i$ ,  $1 \leq i \leq k$ , denota el vector cuyo  $i$ -ésimo componente es 1 y cuyos otros componentes son 0.

**Definición 4.1.1.** El Proceso de Galton-Watson multitipo (o vectorial) es un vector temporalmente homogéneo del proceso de Markov  $\mathbf{Z}_0, \mathbf{Z}_1, \mathbf{Z}_2, \dots$ , cuyos estados son vectores en  $T$ . Se supondrá que  $\mathbf{Z}_0$  es no aleatorio. Se interpreta a  $Z_n^i$ , el  $i$ -ésimo componente de  $\mathbf{Z}_n$ , como el número de objetos de tipo  $i$  en la  $n$ -ésima generación.

La ley de transición para el proceso es como sigue. Si  $\mathbf{Z}_0 = e_i$ , entonces  $\mathbf{Z}_1$  tiene la función generadora

$$\psi^i(s_1, \dots, s_k) = \sum_{r_1, \dots, r_k}^{\infty} p^i(r_1, \dots, r_k) s_1^{r_1} \dots s_k^{r_k}, \quad (4.1.1)$$

$|s_1|, |s_2|, \dots, |s_k| \leq 1$ , donde  $p^i(r_1, r_2, \dots, r_k)$  es la probabilidad de que un objeto de tipo  $i$  tenga  $r_1$  hijos de tipo 1,  $\dots$ ,  $r_k$  de tipo  $k$ . En general, si  $\mathbf{Z}_n = (r_1, r_2, \dots, r_k) \in T$ , entonces  $\mathbf{Z}_{n+1}$  es la suma de  $r_1 + r_2 + \dots + r_k$  vectores aleatorios independientes, donde  $r_1$  tiene la función generadora  $\psi^1$ ,  $r_2$  la función generadora  $\psi^2, \dots, r_k$  la función generadora  $\psi^k$ .

**Definición 4.1.2.** La función generadora de probabilidades de  $\mathbf{Z}_n$ , cuando  $\mathbf{Z}_0 = \mathbf{e}_i$ , es denotada por

$$\psi_n^i(s_1, \dots, s_k) = \psi_n^i(\mathbf{s}),$$

$i = 1, \dots, k, n = 0, 1, \dots$ . Entonces,  $\psi_1^i$  es la función  $\psi^i$  de la ecuación (4.1.1). El vector  $(\psi^1(s), \dots, \psi^k(s))$  se denota por  $\boldsymbol{\psi}(\mathbf{s})$ .

Directamente de lo anterior, se puede deducir el siguiente Teorema

**Teorema 4.1.3.** Las funciones generadoras de probabilidades  $\psi_n^i$  son funciones iteradas, definidas por las relaciones

$$\begin{aligned} \psi_{n+1}^i(\mathbf{s}) &= \psi^i[\psi_n^1(\mathbf{s}), \dots, \psi_n^k(\mathbf{s})], \\ \psi_0^i(\mathbf{s}) &= s_i, \end{aligned} \quad (4.1.2)$$

$n = 0, 1, \dots, i = 1, 2, \dots, k$ .

En forma vectorial se tiene que

$$\boldsymbol{\psi}_{n+N}(\mathbf{s}) = \boldsymbol{\psi}_n[\boldsymbol{\psi}_N(\mathbf{s})],$$

$n, N = 0, 1, 2, \dots$

**Definición 4.1.4.** Sea  $\mathbf{M}=(m_{ij})$  la matriz de primeros momentos,

$$m_{ij} = E\left(Z_1^j \mid \mathbf{Z}_0 = \mathbf{e}_i\right) = \frac{\partial \psi^i(1, \dots, 1)}{\partial s_j},$$

$i, j = 1, \dots, k$ .  $\mathbf{M}$  es la matriz del número esperado de progenitores de todos los tipos de individuos. Se supondrá que todos los primeros momentos  $m_{ij}$  son finitos y no todos iguales a 0. Mediante el uso de la regla de la cadena en (4.1.2), se obtiene que  $E(\mathbf{Z}_{n+1} \mid \mathbf{Z}_n) = \mathbf{Z}_n \mathbf{M}$ .

En forma más general,

$$E(\mathbf{Z}_{n+N} \mid \mathbf{Z}_N) = \mathbf{Z}_N \mathbf{M}^n. \quad (4.1.3)$$

Los siguientes son resultados básicos de la teoría de Perron-Frobenius de matrices positivas (véase [1]). Esta teoría demuestra que la iteración de matrices regulares no-negativas pueden ser aproximadas usando las potencias del valor propio dominante de la matriz, lo cual demuestra que son positivas. Como una consecuencia de ello, las propiedades asintóticas del Proceso de Galton-Watson Multitipo en el caso regular positivo pueden ser expresadas usando las potencias de esos valores propios.

**Definición 4.1.5.** Un vector ó matriz positiva, no-negativa ó  $\mathbf{0}$  es llamada así si todos sus componentes tienen esas propiedades. Si  $\mathbf{u}$  y  $\mathbf{v}$  son vectores o matrices, entonces  $\mathbf{u} > \mathbf{v}$  ( $\mathbf{u} \geq \mathbf{v}$ ) significa que  $\mathbf{u} - \mathbf{v}$  es positiva. Se usa signos de valor absoluto que encierran un vector ó matriz para denotar la suma del valor absoluto de los elementos; por lo tanto,  $|\mathbf{Z}_n| = \sum_i |Z_n^i|$ , donde  $|Z_n|$  indica el número total de individuos en la  $n$ -ésima generación.

**Teorema 4.1.6.** Sea  $\mathbf{M}$  una matriz no-negativa de orden  $k$ , tal que  $\mathbf{M}^N$  es positiva para algún entero positivo  $N$ . Entonces  $\mathbf{M}$  tiene una raíz positiva  $\rho$  la cual es simple y mayor en valor absoluto que cualquier otra raíz positiva;  $\rho$  corresponde a los vectores propios por la derecha y por la izquierda  $\boldsymbol{\mu} = (\mu^i)$  y  $\boldsymbol{\nu} = (\nu^i)$ ,  $i = 1, 2, \dots, k$ , los cuales son los únicos vectores propios no-negativos. Por otra parte, se tiene que:

$$\mathbf{M}^n = \rho^n \mathbf{M}_1 + \mathbf{M}_2^n, \quad (4.1.4)$$

$n = 1, 2, \dots$ , donde  $\mathbf{M}_1 = (\mu^i \nu^i)$ , con la normalización  $\sum \mu^i \nu^i = 1$ . Por lo tanto,  $\mathbf{M}_1 \mathbf{M}_1 = \mathbf{M}_1$ . Además

- $\mathbf{M}_1 \mathbf{M}_2 = \mathbf{M}_2 \mathbf{M}_1 = \mathbf{0}$ .
- $|\mathbf{M}_2^n| = O(\alpha^n)$  para cualquier  $\alpha$ ,  $0 < \alpha < \rho$ .
- Para cada  $r$ ,  $1 \leq r \leq k$ , se tiene  $\min_{1 \leq i \leq r} \sum_{j=1}^r m_{ij} \leq \rho$ .
- Si  $j$  es un entero positivo entonces  $\rho^j$  corresponde a  $\mathbf{M}^j$ , sólo cuando  $\rho$  corresponde a  $\mathbf{M}$ .

**Definición 4.1.7.** Un proceso de Galton-Watson Multitipo es llamado positivamente regular o irreducible si  $\mathbf{M}^N$  es positiva para algún entero positivo  $N$ .

Se puede observar como en el Capítulo 2 que los estados del proceso diferentes de cero son transitorios.

El estado  $\mathbf{z}$  es transitorio si  $P(\mathbf{Z}_n = \mathbf{z} \text{ para algún } n = 1, 2, \dots \mid \mathbf{Z}_0 = \mathbf{z}) < 1$  ó equivalentemente si  $P(\mathbf{Z}_n = \mathbf{z} \text{ infinitamente frecuente} \mid \mathbf{Z}_0 = \mathbf{z}) = 0$ .

**Definición 4.1.8.** Un proceso de Galton-Watson Multitipo es llamado singular si las f.g.p.  $\psi^1(s_1, \dots, s_k)$ ,  $\psi^2(s_1, \dots, s_k)$ ,  $\dots$ ,  $\psi^k(s_1, s_2, \dots, s_k)$  son todas lineales en  $s_1, s_2, \dots, s_k$ , con términos no constantes; es decir, cada objeto tiene exactamente un individuo.

**Teorema 4.1.9.** Supóngase que el proceso es positivamente regular y no singular, y suponga que  $\mathbf{z} \in T$ ,  $\mathbf{z} \neq \mathbf{0}$ . Entonces  $P(\mathbf{Z}_n = \mathbf{z} \text{ infinitamente frecuente}) = 0$ .

#### Probabilidad de extinción

El valor propio  $\rho$  de la matriz  $\mathbf{M} = (m_{ij})$  juega un rol similar al de la esperanza  $m$  utilizada en el Capítulo 2, para determinar si la extinción ocurre.

**Definición 4.1.10.** Sea  $q^i$  la probabilidad de extinción si inicialmente existe un objeto de tipo  $i$ ,  $i = 1, 2, \dots, k$ , es decir,

$$q^i = P(\mathbf{Z}_n = 0 \text{ para algún } n \mid \mathbf{Z}_0 = \mathbf{e}_i).$$

El vector  $(q^1, q^2, \dots, q^k)$  es denotado por  $\mathbf{q}$ . Sea  $\mathbf{1}$  denotando el vector  $(1, 1, \dots, 1)$ .

Ahora se enuncia un Teorema el cual indica la generalización del Teorema 2.4.2 del Capítulo 2.

**Teorema 4.1.11.** *Supongáse que el proceso es positivamente regular y no singular. Si  $\rho \leq 1$ , entonces  $\mathbf{q} = \mathbf{1}$ . Si  $\rho > 1$ ,  $\mathbf{q}$  satisface la ecuación*

$$\mathbf{q} = \psi(\mathbf{q}). \quad (4.1.5)$$

*Demostración.* Del Teorema 4.1.9 se sigue que si  $N$  es algún número positivo entonces

$$\begin{aligned} P(Z_n = z \text{ infinitamente frecuente}) &= P\left(\limsup_{n \rightarrow \infty} \{\mathbf{Z}_n = z\}\right) \\ &= P\left(\bigcup_{n=1}^{\infty} \bigcap_{k \geq n} \{\mathbf{Z}_k = z\}\right) = 0, \end{aligned}$$

existe  $n \geq 1$  tal que para todo  $k \geq n$ :  $\mathbf{Z}_k = z$ , entonces

$$P(0 < |\mathbf{Z}_n| < N \text{ para una infinidad de valores de } n) = 0.$$

Por lo tanto,  $P(|\mathbf{Z}_n| \rightarrow 0) + P(|\mathbf{Z}_n| \rightarrow \infty) = 1$ . Ahora de (4.1.3), del Teorema 4.1.6 y la suposición de que  $\rho \leq 1$  se sigue que  $E(|\mathbf{Z}_n|)$  es una función acotada de  $n$ . Por consiguiente  $P(|\mathbf{Z}_n| \rightarrow \infty) = 0$  y por lo tanto,  $P(|\mathbf{Z}_n| \rightarrow 0) = 1$ , probando de esta forma el caso cuando  $\rho \leq 1$ .

Suponga que  $\rho > 1$ , se tiene que

$$\psi_{n+N}(0) = \psi_n(\psi_N(0)), \quad (4.1.6)$$

$n, N = 0, 1, 2, \dots$  Por otra parte, al igual que en el caso simple del Capítulo 2, se tiene que  $\psi_n^i$  es no-decreciente en  $n$  y que se aproxima a  $q^i$ ,  $i = 1, 2, \dots$ , cuando  $n \rightarrow \infty$ . Tomando a  $n = 1$  en la ecuación (4.1.6) y  $N \rightarrow \infty$ , se tiene que la ecuación (4.1.5) se satisface. □

## 4.2. Ejemplo del PG-WM

En 2009 fue identificada una nueva cepa de influenza A (H1N1) la cual causó una pandemia mundial (véase [5, 15]). En el presente estudio se examina la probabilidad

de extinción de la pandemia Influenza A(H1N1), dicho estudio se basa en datos de México (véase [5]).

Se emplea la teoría del Proceso de Galton-Watson Multitipo para aproximar la probabilidad de extinción de la rama de la infección que se deriva de un solo índice de casos. Considérese una población muy grande la cual es completamente susceptible, dividida en dos subpoblaciones, niños y adultos.

Por simplicidad se ignorará inmunidad pre-existente entre adultos. En este capítulo, se denota a los niños como tipo 1 y los adultos como tipo 2.

Sea  $\gamma_i$  ( $i = 1, 2$ ) la tasa de recuperación de individuos infectados de tipo  $i$  y  $\beta_{ij}$  ( $1 \leq i, j \leq 2$ ) la tasa de natalidad (es decir, la tasa de una nueva infección) de individuos infectados de tipo  $i$  causado por un individuo infectado particular de tipo  $j$  durante la etapa inicial de una epidemia.

Se supone que el tiempo de generación tiene una distribución exponencial, y por lo tanto,  $\mathbf{R}_{ij} = \beta_{ij}/\gamma_j$ . Se supondrá que la edad de la matriz de la siguiente generación  $\mathbf{R}_{ij}$  se atribuye completamente a la tasa de infección  $\beta_{ij}$ , y por lo tanto, el período de infección  $\gamma_i$  se supondrá que es una constante  $\gamma$ , independiente del tipo de huésped.

El vector aleatorio  $Z_n = (Z_n^1, Z_n^2)$  representa el número de niños y adultos respectivamente infectados de la población en la  $n$ -ésima generación; se considera al proceso  $\{Z_n\}$  como un proceso de ramificación Multitipo. Suponiendo que un individuo de tipo  $j$  tiene probabilidad  $p^j(\mathbf{r} = r_1, r_2)$  de infectarse, en la siguiente generación,  $r_1$  niños y  $r_2$  adultos, se define a la función generadora de probabilidades como

$$\psi^j(s_1, s_2) = \sum_{\mathbf{r}} p^j(r_1, r_2) s_1^{r_1} s_2^{r_2}, \quad (4.2.1)$$

$j = 1, 2$ . Se tiene por la ecuación (4.2.6) (véase Apéndice B) que la función  $\psi^j(\mathbf{s})$  con un tiempo de generación distribuida exponencialmente está dada por

$$\psi^j(\mathbf{s}) = \frac{\gamma_j}{\gamma_j + \sum_{k=1}^2 \beta_{kj}(1 - s_k)}, \quad (4.2.2)$$

para  $j = 1, 2$ . Puesto que  $\gamma_j$  se supone que es independiente del tipo de huésped  $j$ ,  $\mathbf{M} := \mathbf{R}_{ij} = \beta_{ij}/\gamma$ ,  $\gamma_j = \gamma$  con  $j = 1, 2$  entonces (4.2.2) se simplifica a

$$\psi^j(\mathbf{s}) = \frac{1}{1 + R_{1j}(1 - s_1) + R_{2j}(1 - s_2)}.$$

Las ramas de las infecciones,  $\{Z_n\}$ , que se deriva del índice de casos iniciales llegan a extinguirse con probabilidad 1 si y sólo si el valor propio dominante de  $\mathbf{M}$  es menor que o igual a la unidad, es decir,  $\rho(\mathbf{M}) < 1$ .

Sea  $q^i$  la probabilidad de extinción dado que un individuo particular infectado de tipo  $i$  se introduce en la población. De acuerdo al Teorema 4.1.11 si  $\rho > 1$  la probabilidad de extinción es la raíz no negativa de las ecuaciones

$$q^j = \psi^j(\mathbf{q}),$$

$j = 1, 2$ . Como es usual en los modelos de procesos de ramificación, cada uno de los casos secundarios de tipo  $i$  generado por un caso primario se convierte en un antepasado de un subproceso independiente (el cual se reinicia con un individuo de tipo  $i$ ) comportándose idénticamente entre los del mismo tipo  $i$ . Debido a este carácter multiplicativo, se tiene que la probabilidad de extinción es

$$p(\mathbf{a}) = \prod_{j=1}^2 \{q^{a_j}, q^{a_j}\},$$

donde  $\mathbf{a}=(a_1, a_2)$ .

En la población de dos huéspedes, es decir, una población compuesta de niños y adultos, las probabilidades de extinción dado un solo niño o la persona adulta infectada,  $q_1$  y  $q_2$  satisfacen

$$\begin{aligned} q^1 &= \frac{1}{1 + R_{11}(1 - q_1) + R_{21}(1 - q_2)}, \\ q^2 &= \frac{1}{1 + R_{12}(1 - q_1) + R_{22}(1 - q_2)}. \end{aligned} \quad (4.2.3)$$

En otras palabras, dada la matriz de la siguiente generación  $\mathbf{M}$ , el problema de calcular la probabilidad de extinción dado un cierto número de individuos infectados de tipo  $i$  y  $j$  en la 0-ésima generación es reemplazado por el problema de resolver dos ecuaciones cuadráticas con dos parámetros desconocidos. Existen cuatro posibles combinaciones de las soluciones para (4.2.3) incluyendo números complejos, pero iterativamente sólo se encuentran los números reales no-negativos en el rango de  $0 < q^1, q^2 < 1$ , excepto para la combinación  $(q^1, q^2) = (1, 1)$ .

Se tiene que para encontrar la probabilidad de extinción de  $q^1$  y  $q^2$  para la pandemia A(H1N1), se usan estimadores publicados de la matriz  $\mathbf{M}$  para México (véase [5])

$$\mathbf{M} = \begin{pmatrix} 1.41 & 0.34 \\ 0.35 & 0.87 \end{pmatrix},$$

el valor propio de la matriz  $\mathbf{M}$  estimado es de 1.58, es decir  $\rho(\mathbf{M}) > 1$  lo cual indica que la infección no se extingue con probabilidad 1. Entonces, por el Teorema 4.1.11 se procede a resolver las ecuaciones (4.2.3) simultáneamente con la información anteriormente obtenida y resulta que

$$q^1 = 0.60 \text{ y } q^2 = 0.72.$$

Las cuales son las probabilidades de extinción para niños y adultos respectivamente.

Se observa que la probabilidad de extinción de la infección considerando homogeneidad dado un caso de tipo 2, es decir, de adultos es mayor que la de niños y por lo tanto, resulta que la población en general es infectada por el caso de tipo 1.

# Conclusiones

Los Procesos de Ramificación, además de plantear una serie de desarrollos teóricos interesantes, tienen un buen número de aplicaciones prácticas en áreas como medicina, biología, demografía, etc.

En resumen las aportaciones del trabajo de tesis son las siguientes:

- Análisis detallado del Proceso de Ramificación de Galton-Watson con tiempo y espacio de estados discreto, en esta parte se trabajó con la función generadora de probabilidades del proceso y momentos del mismo.
- Cálculo de la probabilidad de extinción para una población que se desarrolla según un Proceso de Galton-Watson, primero bajo una ley de reproducción inicial Geométrica y posteriormente para la distribución Poisson, donde se consideran diferentes valores de los parámetros de cada distribución. En este caso se usó un enfoque de punto fijo para garantizar la existencia de la probabilidad de extinción (Teorema 2.4.2), así como un procedimiento iterativo de la función generadora de probabilidades (Teorema 2.4.1) para la aproximación del mismo, se puede notar que en algunos ejemplos el uso del Teorema 2.4.2 es más fácil de emplear como en el caso de la distribución Geométrica, para ello se consideraron los dos ejemplos con distribuciones diferentes.
- Estimación del promedio de reproducción y de la probabilidad de descendientes vía máxima verosimilitud. Posteriormente se dieron estimadores no paramétricos propuestos por algunos autores los cuales fueron comparados mediante una simulación. La conclusión fue que el estimador  $\hat{q}_1$  tiene un error cuadrado medio menor que el estimador  $\hat{q}_2$  en este sentido se obtiene que el mejor estimador es  $\hat{q}_1$ , al menos para una población que se distribuye de acuerdo a una Poisson.
- Finalmente, se desarrolló la teoría para el Proceso de Galton-Watson Multitipo, y posteriormente es aplicada a un ejemplo, en el cual se investigó la probabilidad de extinción de la infección causada por la pandemia influenza A(H1N1-2009) basado en datos de México; en esta parte se consideró dividir la población en niños y adultos. Este ejemplo considera homogeneidad en la población, es decir, no pueden entrar individuos ajenos infectados a la población. Se concluye que la probabilidad de extinción de la infección para niños

es 0.60 y para adultos 0.72 con  $\rho = 1.58$ , con lo cual se puede observar que los niños son más propensos a contraer esta infección.

Como trabajos a futuro se tienen los siguientes:

- Analizar y estimar el tiempo de extinción del Proceso de Ramificación de Galton-Watson.
- Desarrollar la teoría para el Proceso de Ramificación de Galton-Watson con espacio de estados discreto y tiempo continuo, para que posteriormente se pueda aplicar a ejemplos.

# Apéndice A

## Resultados Asintóticos

En este primer apéndice se estudia la inestabilidad de  $Z_n$  cuando  $n \rightarrow \infty$ , y el comportamiento de la sucesión aleatoria  $Z_1, Z_2, \dots$ . La distribución límite no puede en general ser obtenida explícitamente, excepto cuando  $m = 1$ .

Los conceptos usuales de convergencia de una sucesión de variables aleatorias  $\{W_n\}$  a una variable  $W$  son los siguientes:

1. Para la convergencia en probabilidad se tiene que, para cada  $\epsilon > 0$

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

2. Convergencia en media cuadrada significa que

$$\lim_{n \rightarrow \infty} E(W_n - W)^2 = 0.$$

3. En la convergencia casi segura se cumple que

$$P\left(\lim_{n \rightarrow \infty} W_n = W\right) = 1.$$

4. Convergencia en distribución significa que

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x),$$

para todo punto de continuidad  $x \in \mathbb{R}$ , donde  $F_X(x)$  es continua.

Cualquiera de los tres primeros modos de convergencia implican el cuarto.

### Inestabilidad de $Z_n$

Se tiene que

$$P[Z_n = 0] \rightarrow q,$$

cuando  $n \rightarrow \infty$ , y

$$P[Z_n = k] \rightarrow 0,$$

cuando  $n \rightarrow \infty$ , para todo  $k \geq 1$ , de hecho se puede obtener el siguiente resultado

$$P[Z_n \rightarrow 0 \text{ ó } Z_n \rightarrow \infty] = 1,$$

y

$$q = P[Z_n \rightarrow 0] = 1 - P[Z_n \rightarrow \infty].$$

El proceso de ramificación exhibe una inestabilidad, es decir, el proceso se extingue o explota.

Esto se puede resumir en el siguiente teorema.

**Teorema 4.2.1.** *Sea  $m = E(Z_1) < \infty$ , como  $\lim_{n \rightarrow \infty} P(Z_n = k) = 0$ ,  $k = 1, 2, \dots$ . Entonces  $Z_n \rightarrow \infty$  con probabilidad  $1 - q$  y  $Z_n \rightarrow 0$  con probabilidad  $q$ .*

*Demostración.* Primero se muestra que cada uno de los estados  $k = 1, 2, \dots$  son transitorios.

Supóngase para evitar una situación degenerada que  $p_1 \neq 1$ , como 0 es un estado absorbente entonces 0 es recurrente. Se tiene, sin embargo que,  $1, 2, \dots$  son transitorios. Si  $p_0 = 0$ , entonces el número de descendientes por individuos es por lo menos 1, y por lo tanto,  $\{Z_n\}$  es no decreciente. Si se empieza en el estado  $k$ , el único camino posible para regresar al estado  $k$  es si cada uno de los  $k$  miembros de la generación actual tiene exactamente un descendiente, por lo tanto, para  $k \geq 1$

$$\begin{aligned} \rho_{kk} &= P[\text{regresar eventualmente a } k] \\ &= P[Z_{n+1} = k \mid Z_n = k] \\ &= P[Z_{n+1,j} = 1, j = 1, \dots, k] = p_1^k < 1, \end{aligned}$$

se tiene que en el caso  $p_0 = 0$ ,  $k$  es transitorio. Para el siguiente caso se considera que  $p_0 = 1$ , entonces  $p_{k0} = 1$ , así que  $\rho_{kk} = 0 < 1$  y otra vez el estado  $k$  es transitorio. Finalmente se considera el caso donde  $0 < p_0 < 1$ . Como 0 es absorbente,

$$\rho_{kk} \leq P_k[Z_1 \neq 0] = 1 - P_k[Z_1 = 0] = 1 - p_0^k < 1.$$

Así que para este caso también,  $k \geq 1$  es transitorio.

Por lo tanto, para cada  $k = 1, 2, \dots$  se tiene que  $\lim_{n \rightarrow \infty} P(Z_n = k) = 0$  y  $P(Z_n = k \text{ para un número infinito de valores de } n) = 0$ . Dado que  $Z_n$  no toma el mismo valor positivo infinitas veces, la sucesión converge a 0 ó  $\infty$ . Del Teorema 2.4.2 se sabe que  $q$  es la probabilidad de que  $Z_n \rightarrow 0$ . Con lo cual se concluye la demostración.  $\square$

A continuación se describen brevemente los 3 casos para el promedio de reproducción  $m$  los cuales se mencionaron en la Sección 2.3.

### Caso supercrítico

Se tiene que una característica adicional de los procesos de ramificación es la siguiente. Sea

$$W_n := Z_n / E[Z_n] = Z_n / m^n,$$

$n \geq 1$ . Se demostrará que  $\{W_n : n \in \mathbb{N}\}$  es una martingala adaptada a la filtración  $\{\mathcal{F}_n = \sigma(Z_0, Z_1, \dots, Z_n) : n \geq 1\}$ . La propiedad de martingala del proceso anterior es de gran ayuda en el estudio asintótico del proceso original.

**Lema 4.2.2.**  $\{W_n\}$  es una martingala.

*Demostración.* Es necesario probar que se cumplen las dos condiciones de la Definición 1.2.2.

El inciso 1 de la Definición 1.2.2 se cumple, ya que

$$\begin{aligned} E[|W_n|] &= E[|Z_n/m^n|] \\ &= \frac{1}{m^n} |E[Z_n]| = 1 < \infty. \end{aligned}$$

Ahora se verificará que se cumple 2 de la Definición 1.2.2.

$$\begin{aligned} E[W_{n+1} | \mathcal{F}_n] &= m^{-(n+1)} E[Z_{n+1} | \mathcal{F}_n] \\ &= m^{-(n+1)} E\left[\sum_{i=1}^{Z_n} X_{n,i} | \mathcal{F}_n\right] \\ &= m^{-(n+1)} \sum_{k=1}^{\infty} E[I_{\{Z_n=k\}} k \cdot X_{n,i} | \mathcal{F}_n] \\ &= m^{-n} \sum_{k=1}^{\infty} E[k \cdot I_{\{Z_n=k\}} | \mathcal{F}_n] \\ &= m^{-n} Z_n = W_n. \end{aligned}$$

□

Entonces se tiene que  $W_n$  es una martingala con respecto a  $\{\mathcal{F}_n\}$ , más aún, como  $W_n \geq 0$  entonces por el Corolario 1.2.8 existe una variable aleatoria no negativa  $W$  tal que con probabilidad 1,  $\lim_{n \rightarrow \infty} W_n = W$ .

Bajo las condiciones  $m > 1$ ,  $\sigma^2 < \infty$  y  $Z_0 = 1$  se obtiene:

1.

$$\lim_{n \rightarrow \infty} E(W_n - W)^2 = 0.$$

2.

$$E(W) = 1,$$

y

$$\text{var}[W] = \sigma^2/m(m-1).$$

3.

$$P(W = 0) = q.$$

**Caso subcrítico**

En el caso subcrítico  $m < 1$ , se sabe que con probabilidad uno la población se extingue, por lo cual la distribución límite de  $Z_n$  no es interesante. Sin embargo, se está interesado en analizar el comportamiento asintótico del proceso  $Z_n$  condicionado por la hipótesis  $Z_n \neq 0$ .

**Teorema 4.2.3.** *Supóngase que  $m < 1$  y  $E(Z^2) < \infty$ . Entonces, para cada  $j = 1, 2, \dots$ ,*

$$\lim_{n \rightarrow \infty} P(Z_n = j \mid Z_n \neq 0) := b_j$$

existe y  $\sum_{k=0}^{\infty} b_k = 1$ . Además, la f.g.p.  $g(s) = \sum b_j s^j$  satisface la ecuación funcional

$$g[f(s)] = mg(s) + 1 - m,$$

$$|s| \leq 1.$$

*Demostración.* Véase [9].

□

**Caso crítico**

Si  $m = 1$  se tiene que  $P(Z_n \rightarrow 0) = 1$  o equivalentemente  $P(Z_n > 0) \rightarrow 0$  cuando  $n \rightarrow \infty$ . En este caso se está interesado en analizar la tasa de convergencia a cero. Se deduce que para  $m = 1$ ,  $\sigma^2 < \infty$  y  $Z_0 = 1$ .

1.  $nP[Z_n > 0] \rightarrow \frac{2}{\sigma^2}$ .
2.  $E[Z_n \mid Z_n > 0] \rightarrow \frac{\sigma^2}{2}$ .
3.  $\lim_{n \rightarrow \infty} P[(Z_n/n) > x \mid Z_n > 0] = e^{-\frac{2x}{\sigma^2}}$ .

El resultado anterior indica que el decrecimiento de una población que se comporta según un Proceso de Galton-Watson con  $m = 1$  es exponencial.

# Apéndice B

En esta parte se ve una extensión del Proceso de Galton-Watson anteriormente estudiado (véase Capítulo 2), en este caso se trabaja con espacio de estados discreto y tiempo continuo, se le denomina Proceso Markoviano de Ramificación a Tiempo Continuo (PMRTC), su análisis es muy breve y se puede ver de forma mas detallada en [9].

Este proceso también modela el crecimiento de una población, salvo que ahora el tiempo de vida de cada individuo es una variable aleatoria continua. Posteriormente se da un ejemplo el cual es utilizado en el Capítulo 4, Sección 4.2.

## Proceso de Ramificación a Tiempo Continuo

En el modelo de Galton-Watson de tiempo discreto cada individuo vive una unidad de tiempo, y por lo tanto, todos los individuos que coexisten pertenecen a la misma generación. Al contrario, en el PMRTC  $Z = Z(t), t \geq 0$ , donde  $Z(t)$  es el número de individuos al tiempo  $t$  cada uno de esos individuos tiene una duración de vida aleatoria, al término de la cual se muere o reproduce un número aleatorio de individuos descendientes. Así, coexisten siempre varias generaciones. Este proceso en general no es de Markov, al menos que los tiempos de vida sean independientes y se distribuyan como variables aleatorias exponenciales, lo cual se va a suponer.

Para empezar con la construcción del PMRTC se definen a las probabilidades de transición

$$P_{ij}(\tau, t) = P(Z(\tau + t) = j \mid Z(\tau) = i),$$

para toda  $i, j \in \{0, 1, 2, \dots\}$ ,  $0 \leq \tau \leq t$ ,  $t \geq 0$ .

Debido a la hipótesis de homogeneidad del tiempo, estas satisfacen que

$$P_{ij}(\tau, t) = P_{ij}(t).$$

**Definición 4.2.4.** Un proceso estocástico  $\{Z(t) : t \geq 0\}$  definido en un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  es llamado el PMRTC unidimensional si:

1. Su espacio de estados esta dado por el conjunto de enteros no-negativos.
2. Es una cadena de Markov estacionaria con respecto a  $\mathcal{F}_t = \sigma\{Z(s) : s \leq t\}$ .

3. Las probabilidades de transición  $P_{ij}(t) = P(Z(t) = j \mid Z(0) = i)$  satisfacen:

$$\sum_{j=0}^{\infty} P_{ij}(t) s^j = \left( \sum_{j=0}^{\infty} P_{1j}(t) s^j \right)^i,$$

para toda  $i \geq 0$  y  $|s| \leq 1$ .

Los incisos 1 y 2 indican que  $Z(t)$  es un Proceso Markoviano a Tiempo Continuo en los enteros, la condición 3 dice que considerar un proceso de ramificación con  $i$  individuos iniciales, es exactamente lo mismo que considerar la suma de  $i$  procesos de ramificación independientes con un individuo inicial.

Las ecuaciones de Kolmogorov son las siguientes:

$$\begin{cases} \frac{d}{dt} P_{ij}(t) = & -jaP_{ij}(t) + a \sum_{k=1}^{j+1} kP_{ik}(t) p_{j-k+1}, \\ P_{ij}(\tau, \tau + 0) = & \delta_{ij}, \end{cases}$$

llamada ecuación prospectiva, y

$$\begin{cases} \frac{d}{dt} P_{ij}(t) = & -iaP_{ij}(t) + ia \sum_{k=i-1}^{\infty} P_{kj}(t) p_{k-i+1}, \\ P_{ij}(t-0, t) = & \delta_{ij}, \end{cases}$$

denominada como la ecuación retrospectiva, donde

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

**Definición 4.2.5.** El conjunto de funciones  $\{P_{ij}(t)\}$  es llamado una solución de las ecuaciones retrospectivas y prospectivas de Kolmogorov si son no negativas, absolutamente continuas en 0 y en  $t$ , satisfacen las ecuaciones de Kolmogorov y las siguientes condiciones

1.  $P_{ij}(t) = \sum_{k=0}^{\infty} P_{ik}(t_1) P_{kj}(t)$ , con  $i, j = 0, 1, \dots$ ,  $0 \leq t_1 \leq t$ , (ecuación de Chapman-Kolmogorov).
2.  $\sum_{j=0}^{\infty} P_{ij}(t) \leq 1, i = 0, 1, 2, \dots$

### Función generadora

Se supone que  $\{P_{ij}(t)\}$  es una única solución de las ecuaciones de Kolmogorov, como es única se tiene que satisface

$$\sum_{j=0}^{\infty} P_{ij}(t) = 1.$$

La función generadora juega un papel importante en el análisis de los procesos continuos. Antes de definir a la función generadora del proceso, se define una función  $\psi(s)$  como

$$\psi(s) = \sum_{j=0}^{\infty} p_j s^j,$$

donde  $|s| \leq 1$ , y otra función  $u(s)$  como

$$u(s) = a[\psi(s) - s].$$

El número  $a$ , que toma valores en el intervalo  $(0, \infty)$ , y la función  $\{p_j : j = 0, 1, 2, \dots\}$  son los datos totales del proceso  $\{Z(t) : t \geq 0\}$ . Se define a la función generadora del proceso  $\{Z(t) : t \geq 0\}$  como

$$\begin{aligned} F_k(s, t) &= \sum_{j=0}^{\infty} P(Z(t) = j \mid Z(0) = k) s^j \\ &= \sum_{j=0}^{\infty} P_{kj}(t) s^j \\ &= E[s^{Z(t)}]. \end{aligned}$$

Si  $k = 1$  entonces la función generadora se denota por  $F(s, t)$ . Por las ecuaciones de Kolmogorov se puede ver que  $F(s, t)$  satisface las siguientes relaciones

$$\frac{\partial}{\partial t} F(s, t) = u(F(s, t)) \quad (\text{ecuación retrospectiva}), \quad (4.2.4)$$

y

$$\frac{\partial}{\partial t} F(s, t) = u(s) \frac{\partial}{\partial s} F(s, t) \quad (\text{ecuación prospectiva}), \quad (4.2.5)$$

con condición de frontera

$$F(s, 0) = s.$$

Para obtener la iteración de la función generadora es necesario utilizar la ecuación

de Chapman-Kolmogorov, de esta manera se tiene que

$$\begin{aligned}
 F(s, t + u) &= \sum_{k=0}^{\infty} P_{1k}(t + u) s^k \\
 &= \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} P_{1j}(u) P_{jk}(t) s^k \\
 &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} P_{1j}(u) P_{jk}(t) s^k \\
 &= \sum_{j=0}^{\infty} P_{1j}(u) \sum_{k=0}^{\infty} P_{jk}(t) s^k \\
 &= \sum_{j=0}^{\infty} P_{1j}(u) [F(s, t)]^j \\
 &= F(F(s, t), u).
 \end{aligned}$$

Dicha relación es análoga a la fórmula iterada para  $\psi_n(s)$  en el Proceso de Galton-Watson.

Resolver las ecuaciones de Kolmogorov resulta en la mayoría de los casos una tarea muy difícil y por ello la distribución de  $Z(t)$  muy pocas veces puede determinarse de manera explícita. Sin embargo, existen ejemplos importantes que si se pueden resolver, a continuación se resolverá uno de ellos.

**Ejemplo 4.2.6.** El Proceso de Yule de Fisión binaria.

Sea  $\{Z(t) : t \geq 0\}$  un PMRTC, en donde cada partícula vive una longitud de tiempo que se distribuye de manera exponencial y enseguida se divide en dos partículas. Lo que se quiere ver es que la función generadora es de la forma

$$F(s, t) = \frac{se^{-at}}{1 - s(1 - e^{-at})},$$

con  $F(s, 0) = s$ . Como cada partícula al morir se divide en dos partículas se puede afirmar que  $p_2 = 1$ , como consecuencia se ve que  $\psi(s) = s^2$ . Con todos estos resultados la ecuación retrospectiva de Kolmogorov para la función generadora queda de la siguiente forma

$$\frac{\partial F(s, t)}{\partial t} = a [F^2(s, t) - F(s, t)].$$

Para obtener la solución a esta ecuación diferencial parcial se va a fijar a  $s$ , entonces el problema se restringe a solucionar una ecuación diferencial ordinaria. Primero se va a proponer una función  $\mu(t)$  con  $\mu(0) = 1$ , entonces

$$\mu(t) \frac{dF(s, t)}{dt} + a\mu(t) F(s, t) = a\mu(t) F^2(s, t),$$

y

$$\mu(t) \frac{dF(s, t)}{dt} \frac{1}{F^2(s, t)} + \frac{a\mu(t)}{F(s, t)} = a\mu(t).$$

Entonces se propone a  $\frac{d\mu(t)}{dt}$  como

$$\frac{d\mu(t)}{dt} = -a\mu(t),$$

cuya solución se obtiene de la siguiente manera

$$\begin{aligned} \frac{d\mu(t)}{\mu(t)} &= -adt \\ \ln|\mu(t)| &= -at \\ \mu(t) &= e^{-at}. \end{aligned}$$

De esta manera se tiene que

$$\begin{aligned} \frac{d}{dt} \left( \frac{-e^{-at}}{F(s, t)} \right) &= ae^{-at} \\ \left. \frac{-e^{-ay}}{F(s, y)} \right|_0^t &= \left. -e^{-ax} \right|_0^t \\ \frac{-e^{-at}}{F(s, t)} - \frac{1}{s} &= -e^{-at} + 1, \end{aligned}$$

multiplicando por menos uno y al desarrollar esta relación se tiene

$$\frac{e^{-at}}{F(s, t)} = \frac{1 - (1 - e^{-at})s}{s},$$

despejando

$$F(s, t) = \frac{se^{-at}}{1 - s(1 - e^{-at})}. \quad (4.2.6)$$

La función generadora (4.2.6) es de gran ayuda para el ejemplo que se presentó en el Capítulo 4.



# Apéndice C

## Algoritmos

El siguiente programa se realizó en el software Mathematica y se utilizó en el Capítulo 2, Sección 2.5 para encontrar el valor de la probabilidad de extinción en los ejemplos estudiados.

```
For[i=1 ,x0 = 0.2; i++, xi = F[xi-1];  
e0=0.0001; y[i]=abs[xi - xi-1];  
If [y[i]<e0, Print[xi];Break[]]
```

En el Capítulo 3 para realizar la comparación de los diferentes estimadores de la probabilidad de extinción  $\hat{q}_1$  y  $\hat{q}_2$  se tiene el siguiente código que se implementó en el software R.

El siguiente código es para encontrar el valor del estimador  $\hat{q}_1$ .

```
gwpoi <- function(n){  
z <- numeric()  
m<-10  
l<-20  
m1<-0  
m2<-0  
z[1] <- 1  
z[2] <- rpois(1,2)  
for (j in 3:(n+1))  
{  
if (z[j-1]==0) {z[j]<-z[j-1] & break}  
else {z[j] <- rpois(1,z[j-1]*2)}  
}  
j<-m+1  
for(j in j:1)  
{  
m1<-m1+z[j]}}
```

```

}
k<-m
for(k in k:1-1)
{
  m2<-m2+z[k]
}
mu<-m1/m2
i<-m+1
s<-0
for(i in i:1)
{
  s<-s+z[i-1]*(((z[i]/z[i-1])-mu)**2)
}
s1<-(1-m-1)**(-1)
s2<-s1*s
pi1<-exp(-2*(mu-1)/s2)
print(pi1)
pi<-0.36
recm<-(pi1-pi)
print(recm)
}

```

Después para calcular el estimador  $\hat{q}_2$  se tiene el siguiente código.

```

gwpoi <- function(n){
z <- numeric()
m<-10
l<-20
m1<-0
m2<-0
z[1] <- 1
z[2] <- rpois(1,2)
  for (j in 3:(n+1))
  {
    if (z[j-1]==0) {z[j]<-z[j-1] & break}
    else {z[j] <- rpois(1,z[j-1]*2)}
  }
j<-m+1
for(j in j:1)
{
  m1<-m1+z[j]
}
k<-m
for(k in k:1-1)
{
  m2<-m2+z[k]
}

```

```
}  
mu<-m1/m2  
pi2<-(z[n]-mu**n)/(z[n]-1)  
print(pi2)  
pi<-0.36  
recm<-(pi2-pi)  
print(recm)  
}
```



# Bibliografía

- [1] Alston S. H., The theory of matrices in numerical analysis. Dover Publications, Inc. (1964).
- [2] Athreya K.B., Ney P.E., Branching Processes. Springer Verlag (1972).
- [3] Basu A.K., Introduction to Stochastic Process. Narosa (2003).
- [4] Becker N., Estimation for Discrete Time Branching Processes with Application to epidemics. Biometrics, Vol. 33, No.3, pp. 515-522 (1977).
- [5] C. Fraser, A. Donnelly, Pandemic potencial of a strain of influenza A(H1N1): early findings. Science, Vol. 324, No.5934, pp. 1557-1561 (2009).
- [6] C.C Heyde, Extension of a Result of Seneta for the Super-Critical Galton-Watson Process. Springer (2010).
- [7] Delgado de la Torre R., Probabilidad y Estadística para ciencias e Ingenieras. Delta Publicaciones (2007).
- [8] Guttorp P., Statistical Inference for Branching Processes. John Wiley & Sons, Inc. (1991).
- [9] Harris Theodore E., The Theory of Branching Processes. Dover Publications, Inc. (1989).
- [10] Karlin S., A First Course in Stochastic Processes. Academic Press (1975).
- [11] Klenke A., Probability Theory a Comprehensive Course. Springer Verlag (2006).
- [12] Lawler G. F., Introduction to Stochastics Processes. Chapman & Hall (2006).
- [13] Martín Andrés A., Estimación no Paramétrica de la Edad y de la Probabilidad de Extinción en Procesos de Galton-Watson, Trabajos de estadística y de investigación operativa, Vol. 32, No. 2, pp. 55-67 (1981).
- [14] Miñarro A., Estimación no paramétrica de la función de densidad. Tesis Licenciatura. Barcelona, Enero (1998).

- [15] Nishiura H., R. Cook Alex, J. Cowling B., Assortativity and the Probability of Epidemic Extinction: A case Study of Pandemic Influenza A (H1N1-2009), Hindawi Publishing Corporation, Volume 2011, Article 194507, pp. 1-9 (2010).
- [16] Ospina Martínez R., Proceso de Galton-Watson. Revista Colombiana de Estadística, Vol. 24, No. 1, pp. 13-26 (2001).
- [17] Port y Stone H., Introduction to Stochastics Processes. Waveland Pres (1986).
- [18] Prahbo N., Stochastics Processes; Basic Theory and Applications. World Scientific Publishing Co. (2007).
- [19] Resnick S., Adventures in Stochastic Processes. Birkhäuser (1992).
- [20] Sheldon M. R., Introduction to Probability Models. Academic Press (2007).
- [21] Sheldon M. R., Stochastics Processes. John Wiley & Sons, Inc. (1996).
- [22] Wasserman L., All of Nonparametric Statistics. Springer (2006).