



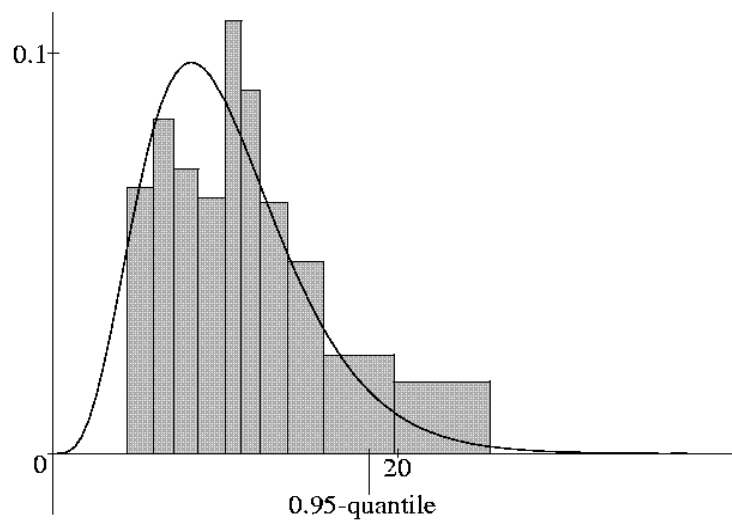
# UNIVERSITÄT POTSDAM

## Institut für Mathematik

### Unas propiedades basicas de procesos de ramificación

*Lectures held at ICIMAF  
La Habana, Cuba, 2009 and 2010*

Sylvie Roelly



Mathematische Statistik und  
Wahrscheinlichkeitstheorie

**Universität Potsdam – Institut für Mathematik**

Mathematische Statistik und Wahrscheinlichkeitstheorie

Unas propiedades basicas de  
procesos de ramificación

Sylvie Roelly

Institut für Mathematik der Universität Potsdam

e-mail: [roelly@math.uni-potsdam.de](mailto:roelly@math.uni-potsdam.de)

Preprint 2010/07

April 2010

## **Impressum**

**© Institut für Mathematik Potsdam, April 2010**

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie  
am Institut für Mathematik

Adresse: Universität Potsdam  
Am Neuen Palais 10  
14469 Potsdam

Telefon: +49-331-977 1500

Fax: +49-331-977 1578

E-mail: [neisse@math.uni-potsdam.de](mailto:neisse@math.uni-potsdam.de)

ISSN 1613-3307

# Unas propiedades basicas de procesos de ramificación

Notas de los cursos impartidos en el ICIMAF, La Habana, Cuba, en abril 2009 y marzo 2010.

Sylvie RÆLLY

Institut für Mathematik der Universität Potsdam

Am Neuen Palais 10

D-14469 Potsdam

e-mail : [roelly@math.uni-potsdam.de](mailto:roelly@math.uni-potsdam.de)

# Índice

<b>1. Unas propiedades de los procesos de Bienaymé-Galton-Watson de tiempo discreto (BGW)</b>	<b>4</b>
1.1. Definición . . . . .	4
1.2. Propiedad de ramificación . . . . .	6
1.3. Propiedad de Markov y clasificación de los estados . . . . .	6
1.4. Función generatriz y momentos . . . . .	7
1.5. Unas martingalas asociadas . . . . .	10
1.6. La probabilidad de extinción . . . . .	12
1.7. Estimación práctica del tiempo de extinción . . . . .	14
1.7.1. Los cálculos . . . . .	14
1.7.2. Aplicación a la extinción de las ballenas negras en el Atlántico Norte . . . . .	16
1.7.3. El valor medio del tiempo de extinción $T_0$ en el caso (sub)crítico . . . . .	17
1.8. Distribución cuasi-estacionaria de BGW . . . . .	18
1.9. El proceso condicionado a la no extinción en el futuro lejano . . . . .	19
<b>2. Unas propiedades del proceso BGW de tiempo continuo</b>	<b>22</b>
2.1. El BGWc como proceso markoviano de salto puro . . . . .	22
2.2. Las ecuaciones de Chapman-Kolmogorov . . . . .	23
2.2.1. El generador infinitesimal . . . . .	25
2.2.2. Las ecuaciones de Kolmogorov hacia adelante y hacia atrás . . . . .	26
2.3. La ley del proceso al tiempo $t$ . . . . .	27
2.3.1. La función generatriz del proceso al tiempo $t$ . . . . .	27
2.3.2. Los primeros momentos . . . . .	28
2.4. El "problema" de martingalas . . . . .	28
2.5. Tiempo de extinción y probabilidad de extinción . . . . .	29
2.6. El proceso BGWc condicionado . . . . .	31
<b>3. Límites de procesos de BGW cuando la población es numerosa</b>	<b>32</b>
3.1. Límite determinista . . . . .	32
3.1.1. BGW de tiempo discreto renormalizado . . . . .	32
3.1.2. BGWc renormalizado . . . . .	32
3.2. Límite difusión . . . . .	33
3.2.1. BGW de tiempo discreto renormalizado . . . . .	33
3.2.2. BGWc renormalizado . . . . .	37
3.2.3. Unas propiedades de la difusión de Feller . . . . .	40

# Introducción

La teoría sobre Procesos de ramificación surge a principios del siglo XX: el origen se atribuye a los estudios realizados, independientemente, por Bienaymé en Francia y por Galton y Watson en Inglaterra, con objeto de investigar las razones de la extinción de determinadas líneas familiares de la aristocracia. Se considere como punto de partida el siguiente problema planteado por F. Galton en la revista “Educational Times”, en 1874:

(Problema 4001) *Consideremos una población de la que sólo nos interesan los varones adultos,  $i$  en número  $y$  que representan apellidos distintos e independientes. Su ley de población es tal que, en cada generación, un  $r_0$  por ciento de los varones adultos tienen hijos varones que no alcanzan ninguno la edad adulta, un  $r_1$  por ciento tienen un sólo hijo varón que alcanza la edad adulta, un  $r_2$  por ciento tiene dos,  $y$  así sucesivamente hasta un  $r_5$  por ciento que tienen cinco. Encontrar (1) ¿qué proporción de apellidos se extinguirán después de  $n$  generaciones?;  $y$  (2) ¿cuántos individuos habrá del mismo apellido en una generación formada por  $j$  personas?*

La primera respuesta a este problema, fue enviada a Galton por el reverendo H.W. Watson. La respuesta aportada por Watson no era totalmente satisfactoria  $y$  el problema de la probabilidad de extinción permaneció, al menos en parte, sin solución durante otro medio siglo. En 1922, R.A. Fisher consideró dicho problema en un contexto genético, trasladándolo al estudio de la progenie de un gen mutante. La importancia de su trabajo reside en que por primera vez los Procesos de Galton-Watson fueron aplicados en Genética. A partir de entonces, los procesos de ramificación han sido utilizados como modelos matemáticos para describir procesos empíricos, no sólo relacionados con la biología, sino también con la Física nuclear, la Medicina, las ciencias de la Computación, la Demografía, etc.

Durante el siglo XX se ha desarrollado ampliamente, ocupando en la actualidad un lugar destacado dentro del contexto general de la teoría sobre Procesos estocásticos, habiéndose publicado numerosas monografías sobre su teoría  $y$  aplicaciones, entre las que citaremos por su especial significación las de Harris [13], Athreya & Ney [1], Jagers [15], Sewastjanow [23], Pakes [22], Haccou, Jagers, Vatutin [12].

Las notas son estructuradas en 3 capítulos. El primer capítulo incluye varias propiedades (Markov, martingala, extinción, cuasi-estacionaridad, condicionamiento a la no extinción) del proceso de ramificación simple de Bienaymé-Galton-Watson (BGW), un proceso de tiempo discreto. En el capítulo dos se introduce el proceso de ramificación BGW de tiempo continuo  $y$  se estudian sus propiedades como proceso de Markov,  $y$  como martingala. También, se analiza lo que resulta cuando se condiciona la dinámica a la no extinción en el futuro lejano. Se discute la similitud con el caso de tiempo discreto. El tercer capítulo incluye límites de procesos de BGW para densidad de población grande. Así se construye -entre otras cosas- el proceso de difusión de Feller. Se mencionará unas propiedades cualitativas típicas.

Deseo expresar mi agradecimiento especialmente a mis compañeros Sylvie Méléard, Maria Emilia Caballero, Begoña Fernandez, Luis Gorostiza, Manuel Molina Fernández  $y$  Miguel González Velasco por su material [20], [3], [8], [10], [19]  $y$  [21] que me prestaron muy generosamente. Quiero hacer patente mi especial reconocimiento a Sophie Péniisson  $y$  Michael Högele por haberme ayudado a concevir estas notas,  $y$  entre otras cosas haber dibujado algunas figuras  $y$  haber corregido mis innumerables faltas en español.

Berlín, abril de 2010

# 1. Unas propiedades de los procesos de Bienaymé-Galton-Watson de tiempo discreto (BGW)

## 1.1. Definición

Consideramos una población sujeta a reproducción. En el instante  $n = 0$  hay  $X_0$  individuos. Al transcurrir el tiempo, después de un tiempo de vida igual a 1, los individuos se reproducen independientemente según una ley de ramificación  $\mathbf{r} := (r_j)_{j \in \mathbb{N}}$ . Los nuevos individuos - después de un tiempo de vida igual a 1 - se reproducen independientemente con las mismas reglas.

El proceso de BGW es la sucesión de variables aleatorias  $(X_n)_{n \in \mathbb{N}}$  -con espacio de estados  $\mathbb{N}$ - representando el número de individuos de las generaciones sucesivas, definida del siguiente modo:

$$\begin{cases} X_{n+1} = \sum_{k=1}^i Y_{n,k} & \text{si } X_n = i \geq 1 \\ X_{n+1} = 0 & \text{si } X_n = 0. \end{cases}$$

Las variables aleatorias  $(Y_{n,k})_{n,k}$  son independientes e idénticamente distribuidas de ley  $\mathbf{r}$ . Intuitivo,  $Y_{n,k}$  representa el número de descendientes del individuo  $k$ -ésimo de la  $n$ -ésima generación.

En estas notas, supondremos siempre que la ley  $\mathbf{r}$  no es degenerada, es decir :  $r_0 > 0$  y  $r_0 + r_1 < 1$ .

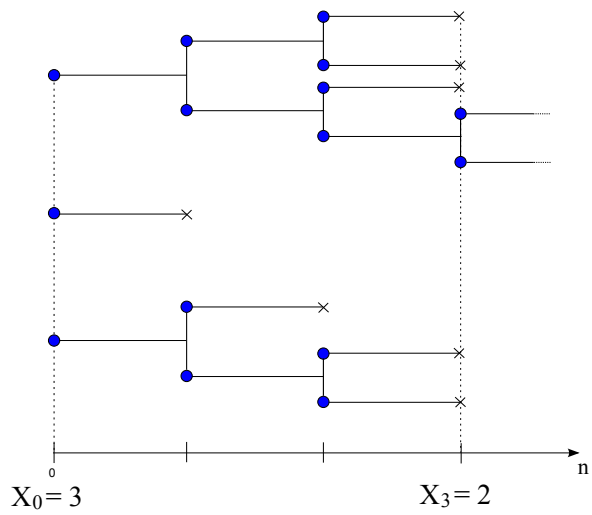
### Ejemplos

- BGW con reproducción binaria: ver figura 1a  
 $r_2 = 1 - r_0, r_j = 0$  por  $j \notin \{0, 2\}$ . Cuando  $r_0 = r_2 = 1/2$  se llama crítica.
- Reproducción de segmentos de ADN de largo  $N$ :  
 $r_0 = b, r_1 = (1 - b)(1 - a^N), r_2 = (1 - b)a^N, r_j = 0$  para todos  $j \geq 3$ . ( $a, b \in ]0, 1[$ )
- propuestos por Watson:
  - (W1):  $\mathbf{r} = U_{\{0,1,2\}}$ , así  $r_0 = r_1 = r_2 = \frac{1}{3}$  ver [9], p. 141
  - (W2):  $\mathbf{r} = \mathcal{B}_{5,1/4}$ , ver [9], p.142 y figura 1b.

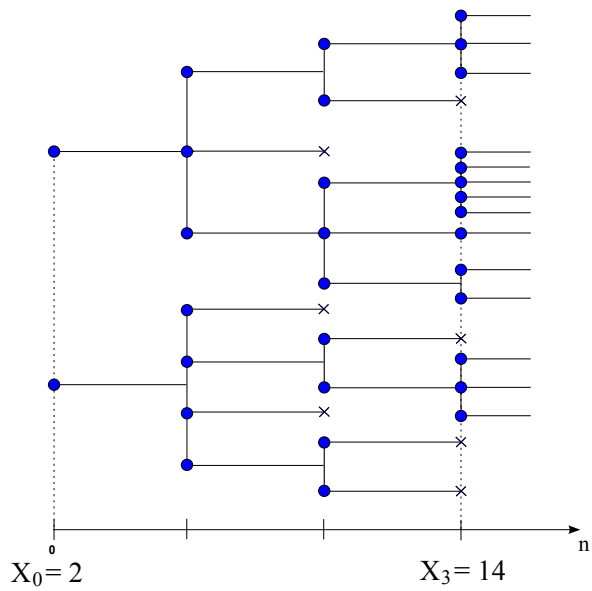
Así

$r_0$	$r_1$	$r_2$	$r_3$	$r_4$	$r_5$
$(\frac{3}{4})^5$	$5 \frac{3^4}{4^5}$	$10 \frac{3^3}{4^5}$	$10 \frac{3^2}{4^5}$	$5 \frac{3}{4^5}$	$\frac{1}{4^5}$
$\approx 0,237$	0,396	0,264	0,088	0,014	0,001

- Modelo geométrico de parámetro  $a \in [0, 1]$  :  $\mathbf{r} = \mathcal{G}_a$   
 $r_j = a(1 - a)^j$  para  $j \geq 0$ .
- Modelo cuasi-geométrico  $(a, b)$  - o geométrico modificado en 0- :  
 $r_0 = b, r_j = (1 - b)a(1 - a)^{j-1}$  para  $j \geq 1$ . (Si  $a = b$ , es geométrico.)  
 Ejemplo de A. Lotka [18] y figura 2. Estadísticas estimadas:  
 $r_0 = 0,4981; r_1 = 0,2103; r_2 = 0,1270; r_3 = 0,0730; r_4 = 0,0418; r_5 = 0,0241 \dots r_{10} =$



(a) BGW binário



(b) Proceso BGW propuesto por Watson (W2)

Figura 1: Procesos de ramificación de tiempo discreto



0,0005... aproximado por

$r_0 = 0,4828; r_1 = 0,2289; r_2 = 0,1279; r_3 = 0,0714; r_4 = 0,0399; r_5 = 0,0233 \dots r_{10} = 0,0012$  que se interpretan como un modelo cuasi-geométrico con los parámetros  $a \approx 0,4414$  y  $b \approx 0,4828$ .

## 1.2. Propiedad de ramificación

- Los descendientes de un individuo de cualquier generación forman un proceso equivalente al total.
- El proceso iniciado en  $X_0 = i$  es la suma de  $i$  copias independientes de  $X$  iniciadas en 1:

$$\forall i, k \quad \mathbb{P}((X_{n+k})_{n \geq 0} \in \cdot | X_k = i) = \mathbb{P}((\sum_{j=1}^i X_n^{(j)})_{n \geq 0} \in \cdot)$$

donde  $\mathbb{P}(X^{(j)} \in \cdot) = \mathbb{P}(X \in \cdot | X_0 = 1)$

## 1.3. Propiedad de Markov y clasificación de los estados

Un proceso de BGW  $X$  es una cadena de Markov porque su comportamiento futuro sólo depende del presente independientemente del pasado: para todos  $n \in \mathbb{N}$ ,  $i, j \geq 1$ ,  $x_0, x_1, \dots \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{P}(X_{n+1} = j | X_0 = x_0, X_1 = x_1, \dots, X_n = i) &= \mathbb{P}\left(\sum_{k=1}^i Y_{n,k} = j | X_0 = x_0, X_1 = x_1, \dots, X_n = i\right) \\ &= \mathbb{P}\left(\sum_{k=1}^i Y_{n,k} = j\right) = \sum_{k_1 + \dots + k_i = j} r_{k_1} \dots r_{k_i} = (\mathbf{r}^{\star i})_j = \mathbb{P}(X_{n+1} = j | X_n = i) \end{aligned}$$

y

$$\mathbb{P}(X_{n+1} = j | X_0 = x_0, X_1 = x_1, \dots, X_n = 0) = 0 \quad \text{si } j \neq 0 \text{ y } 1 \text{ si no.}$$

Su matriz de transición  $P = (P_{ij})_{i,j \geq 0}$  verifica

$$P_{ij} := \mathbb{P}(X_{n+1} = j | X_n = i) = (\overbrace{\mathbf{r} \star \dots \star \mathbf{r}}^i)_j.$$

El estado 0 es absorbente, porque si  $X_n = 0$  entonces  $X_m = 0$  para todos  $m \geq n$ .

### Teorema 1 : Clasificación de estados

Sea  $X$  un proceso de BGW no degenerado. El estado 0 es el único estado recurrente, es decir

$$\mathbb{P}(X_n = 0 \text{ para algún } n \in \mathbb{N} | X_0 = 0) = 1.$$

El resto de estados son transitorios: sea  $i \geq 1$ ;

$$\mathbb{P}(X_n = i \text{ para algún } n \in \mathbb{N} | X_0 = i) < 1.$$

La única distribución estacionaria (o de equilibrio)  $\pi$  es trivial:  $\pi_0 = 1$  y  $\pi_i = 0$  si  $i \geq 1$ .

*Demostración:* 0 es absorbente entonces es recurrente:

Sea  $T_i$  el primer tiempo positivo de visita a  $i$ .  $\mathbb{P}(T_0 < +\infty | X_0 = 0) = \mathbb{P}(T_0 = 1 | X_0 = 0) = 1$ , de donde 0 es recurrente. El tiempo de vuelta a 0 es casi seguramente finito.

Sea  $i \geq 1$ .

$$\mathbb{P}(T_i = +\infty | X_0 = i) \geq \mathbb{P}(X_1 = 0 | X_0 = i) \geq (r_0)^i > 0,$$

de donde  $i$  es transitorio. El tiempo de vuelta a  $i$  puede ser infinito.

Sea  $\pi$  una distribución estacionaria. Es una ley que es invariante sobre la dinamica de reproducción:

Si la ley de  $X_0$  es  $\pi$ , entonces para cada tiempo  $n$  la ley de  $X_n$  se queda igual a  $\pi$ ; en particular,

$$\pi_0 = \mathbb{P}(X_1 = 0) = \sum_{i \geq 0} \mathbb{P}(X_1 = 0 | X_0 = i) \mathbb{P}(X_0 = i) = \sum_{i \geq 0} (r_0)^i \pi_i > \pi_0$$

excepto si  $\pi_i = 0$  para cada  $i \geq 1$ . ◇

## 1.4. Función generatriz y momentos

Se toma la notación  $f$  para la función generatriz de la ley  $\mathbf{r}$ :

$$f(s) := \mathbb{E}(s^{Y_{n,k}}) = \mathbb{E}(s^{X_1} | X_0 = 1) = \sum_{j \geq 0} r_j s^j, \quad 0 \leq s \leq 1.$$

La función  $f$  es continua, crece y es estrictamente convexa en el caso no degenerado.

### Ejemplos :

- Reproducción binaria:  $f(s) = r_0 + r_2 s^2 = 1 + r_2 (s^2 - 1)$ .  
En el caso llamado crítico  $r_0 = r_2 = 1/2$ ,  $f(s) = \frac{1}{2}(1 + s^2)$
- (W1):  $f(s) = \frac{1}{3}(1 + s + s^2) = \frac{1}{3} \frac{1 - s^3}{1 - s}$
- (W2):  $f(s) = \left(\frac{3 + s}{4}\right)^5$ .
- Modelo geométrico ( $a$ ):  $f(s) = \frac{a}{1 - (1 - a)s}$ .
- Reproducción cuasi-geométrica ( $a, b$ ):  $f(s) = b + (1 - b) \frac{as}{1 - (1 - a)s}$ . (Ver la figura 2)

El valor medio  $m$  de la ley  $\mathbf{r}$  representa el número promedio de individuos generados en una reproducción:

$$m = \sum_{j \geq 0} j r_j = \mathbb{E}(X_1 | X_0 = 1) = f'(1^-).$$

Supondremos siempre en estas notas que  $m < +\infty$ . Se consideran tres casos:

$m < 1$  : ramificación subcrítica

$m = 1$  : ramificación crítica

$m > 1$  : ramificación supercrítica

TABLE . THE PROBABILITY  $P_{2j}^{(n)}$  PER 1,000 THAT A NEWBORN MALE SHALL HAVE  $j$  MALE DESCENDANTS IN THE  $n$ -TH GENERATION

$\frac{j}{n}$	None	One or More	1	2	3	4	5	6	7	8	9	10
$(r_j) \rightarrow 1$	482.80	517.20	228.95	127.89	71.44	39.91	22.29	12.45	6.95	3.88	2.17	1.21
2	634.16	365.84	98.29	72.12	52.92	38.83	28.49	20.90	15.33	11.25	8.25	6.05
3	707.65	292.35	53.97	44.19	36.18	29.62	24.25	19.85	16.25	13.30	10.89	8.92
4	750.74	249.26	33.80	29.36	25.50	22.15	19.24	16.71	14.51	12.60	10.94	9.50
5	778.83	221.17	22.95	20.68	18.63	16.79	15.13	13.63	12.28	11.07	9.97	8.98
6	798.45	201.55	16.48	15.22	14.06	12.99	12.00	11.09	10.24	9.46	8.74	8.07
7	812.80	187.20	12.29	11.56	10.87	10.22	9.61	9.04	8.50	7.99	7.51	7.06
8	823.68	176.32	9.46	9.01	8.59	8.19	7.80	7.43	7.08	6.75	6.43	6.13
9	832.13	167.87	7.41	7.13	6.86	6.60	6.35	6.11	5.88	5.66	5.45	5.25
10	838.81	161.19	5.90	5.73	5.56	5.40	5.24	5.08	4.93	4.78	4.64	4.50

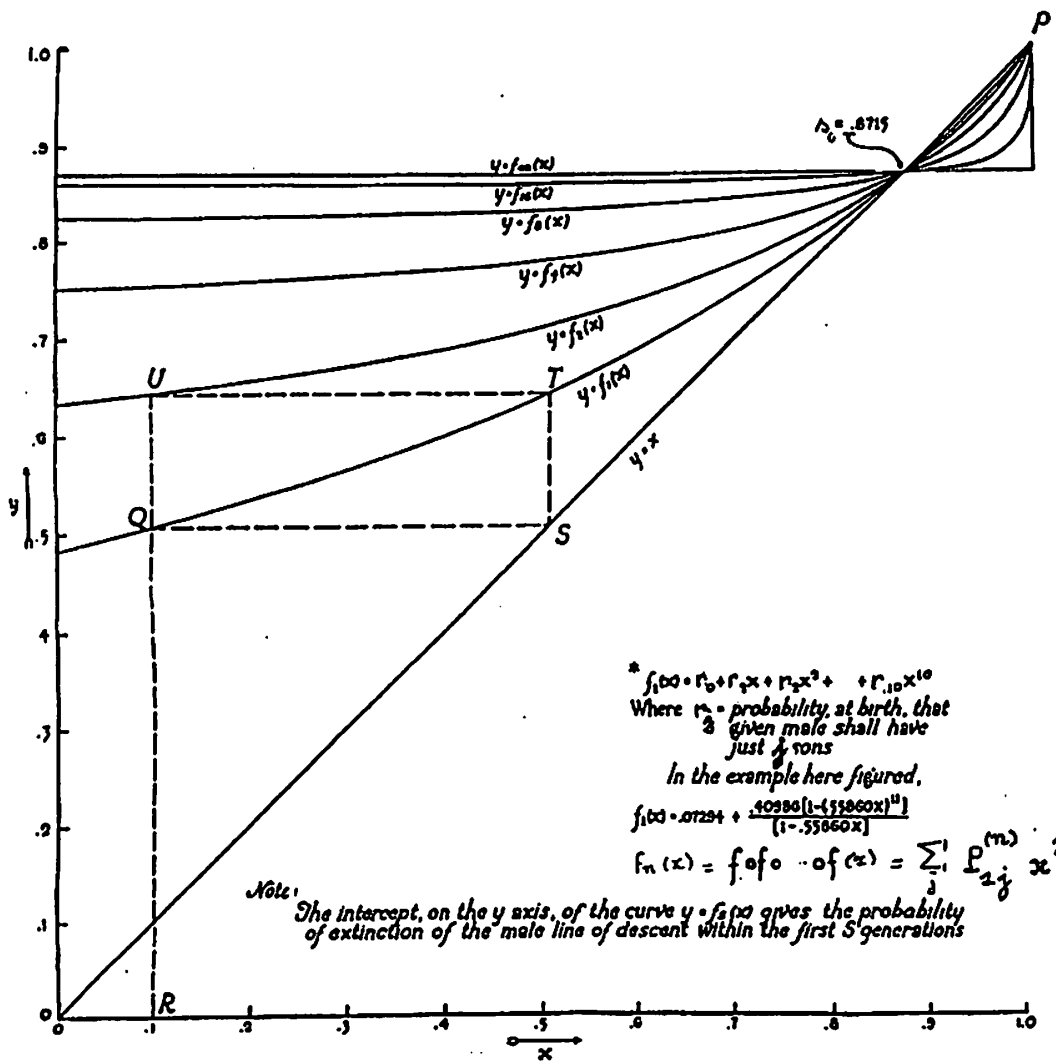


Figura 2: A. Lotka: *The extinction of families* (1931). Tabla y iteración de la función generatriz  $f$

Supondremos también que existe el segundo momento de la ley de reproducción :

$$\sigma^2 := \sum_{j \geq 0} (j - m)^2 r_j = \text{Var}(X_1 | X_0 = 1).$$

**Ejemplos :**

- Reproducción binaria:  $m = 2r_2$ . Es crítica solo si  $r_2 = 1/2$  ( $\Rightarrow r_0 = 1/2$ ).
- (W1):  $m = 1$  y  $\sigma^2 = 2/3$ .
- (W2):  $m = \frac{5}{4}$  y  $\sigma^2 = \frac{15}{16}$ .
- Modelo geométrico ( $a$ ) :  $m = \frac{1-a}{a}$  y  $\sigma^2 = \frac{1-a}{a^2}$ .
- Reproducción cuasi-geométrica ( $a, b$ ):  $m = \frac{1-b}{a}$   
(En el caso de Lotka  $m \approx \frac{0,5172}{0,4414} = 1,17$ . El modelo es supercrítico.)

**Proposición 2 : Estructura iterativa de la función generatriz de BGW**

Si se nota  $f_n(s) := \mathbb{E}(s^{X_n} | X_0 = 1)$  la función generatriz de  $X_n$  cuando  $X_0 = 1$ , entonces

$$f_n = f \circ f \circ \dots \circ f \quad (n \text{ veces}) \quad \text{y} \quad \mathbb{E}(s^{X_n} | X_0 = i) = (f_n(s))^i, \quad i \geq 1.$$

*Demostración:* por inducción

$$\begin{aligned} \mathbb{E}(s^{X_{n+1}} | X_0 = 1) &= \sum_{i \geq 0} \mathbb{E}(\mathbf{1}_{\{X_n=i\}} s^{X_{n+1}} | X_0 = 1) = \sum_{i \geq 0} \mathbb{E}(\mathbf{1}_{\{X_n=i\}} s^{\sum_{k=1}^i Y_{n,k}} | X_0 = 1) \\ &= \sum_{i \geq 0} \mathbb{E}(\mathbf{1}_{\{X_n=i\}} | X_0 = 1) (\mathbb{E}(s^Y | X_0 = 1))^i = \sum_{i \geq 0} \mathbb{P}(X_n = i | X_0 = 1) f(s)^i = f_n \circ f(s). \quad \diamond \end{aligned}$$

Así la respuesta a la pregunta (2) de Galton:

$$\mathbb{P}(X_n = j | X_0 = i) = \frac{1}{j!} \frac{\partial^j}{\partial s^j} (f_n(s)^i) |_{s=0}$$

**Ejemplos**

- Reproducción geométrica ( $a$ ): se prueba que

$$f_n(s) = \begin{cases} \frac{m^n(1-s) + ms - 1}{m^{n+1}(1-s) + ms - 1} & \text{si } a \neq 1/2, \\ \frac{n(1-s) + s}{(n+1) - ns} & \text{si } a = 1/2 \text{ (caso crítico)}. \end{cases}$$

con  $m = (1-a)/a$ .

- Reproducción cuasi-geométrica  $(b, a)$ : (Ver la figura 2)  
Se prueba por inducción que, en el caso no crítico  $a \neq 1 - b$ ,

$$f_n(s) = b_n + (1 - b_n) \frac{a_n s}{1 - (1 - a_n)s}$$

$$\text{donde } b_n := 1 - m^n a_n \text{ y } a_n := \frac{a + b - 1}{b - (1 - a)m^n} \quad (m = \frac{1 - b}{a}).$$

### Los momentos

Se verifican por inducción las siguientes igualdades:

$$\forall n \in \mathbb{N}, \quad \mathbb{E}(X_n) = m^n \mathbb{E}(X_0)$$

$$\text{y } \text{Var}(X_n | X_0 = 1) = \begin{cases} n\sigma^2 & \text{si } m = 1 \\ m^{n-1} \frac{m^n - 1}{m - 1} \sigma^2 & \text{si } m \neq 1. \end{cases}$$

*Demostración:* por inducción

$$\mathbb{E}(X_{n+1} | X_0 = 1) = f'_{n+1}(1) = (f_n \circ f)'(1) = f'(1) f'_n \circ f(1) = m \mathbb{E}(X_n | X_0 = 1). \quad \diamond$$

### Ejemplos

- (W1):  $\mathbb{E}(X_n) = \mathbb{E}(X_0)$  y  $\text{Var}(X_n) = \frac{2}{3} n \mathbb{E}(X_0)$
- (W2):  $\mathbb{E}(X_n) = \mathbb{E}(X_0) (\frac{5}{4})^n$  y  $\text{Var}(X_n) = \frac{15}{4} (\frac{5}{4})^{n-1} ((\frac{5}{4})^n - 1) \mathbb{E}(X_0)$

## 1.5. Unas martingalas asociadas

Normalizando el proceso BGW (se  $m \in ]0, +\infty[$ ), se define la sucesión  $W_n := \frac{1}{m^n} X_n$ .

### Proposición 3 : Primera propiedad de martingala

Sea  $\mathcal{F}_n$  la  $\sigma$ -álgebra definida por  $\mathcal{F}_n := \sigma(X_0, \dots, X_n)$ . Entonces, el proceso normalizado

$$W_n := \frac{X_n}{m^n}$$

es una martingala no negativa con respecto a  $(\mathcal{F}_n)_n$ . Por lo tanto  $W := \lim_n W_n$  existe casi seguramente y  $0 \leq W < +\infty$ .

*Demostración.* Una martingala debe verificar:

$$\forall n \in \mathbb{N}, \quad \mathbb{E}(W_{n+1} | \mathcal{F}_n) = W_n.$$

$$\begin{aligned} \text{Pero } \mathbb{E}(X_{n+1} | \mathcal{F}_n) &= \mathbb{E}(X_{n+1} | X_n). \\ \text{Ademas } \mathbb{E}(s^{X_{n+1}} | X_n = i) &= f(s)^i \\ \Rightarrow \frac{\partial}{\partial s} \mathbb{E}(s^{X_{n+1}} | X_n = i) &= \mathbb{E}(X_{n+1} s^{X_{n+1}-1} | X_n = i) = i f'(s) f(s)^{i-1}; \\ (\text{por } s = 1) \mathbb{E}(X_{n+1} | X_n = i) &= i f'(1) = m i; \end{aligned}$$

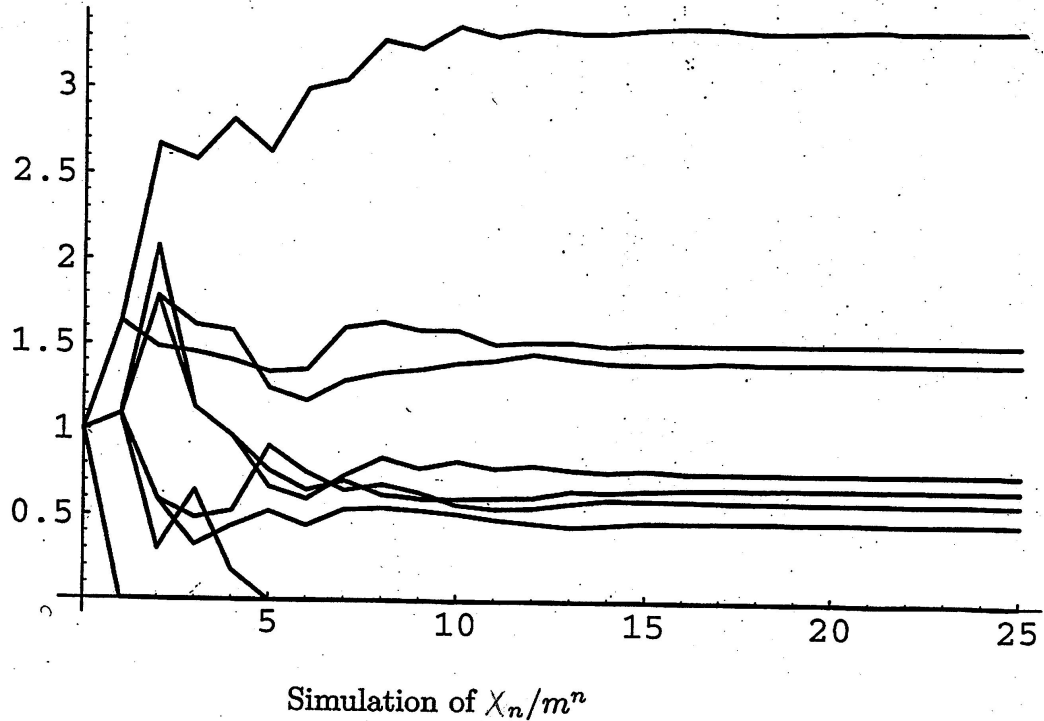


Figura 3: Convergencia de  $W_n = \frac{X_n}{m^n}$

por lo tanto

$$\mathbb{E}(W_{n+1}|\mathcal{F}_n) = \frac{1}{m^{n+1}}\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \frac{1}{m^{n+1}}mX_n = W_n \quad c.s.$$

Se sabe que cada martingala no negativa converge c.s. (ver [24] Teorema 2.21 o [28] §11.5) entonces existe  $W$ .

Si  $m \leq 1$ ,  $X_n \rightarrow 0$  c.s. y  $W \equiv 0$ .

Si  $m > 1$ ,  $\{W > 0\} = \{X_n \rightarrow +\infty\}$ .  $\diamond$

#### Proposición 4 : Una familia de martingalas

Para cada función  $h$  acotada, el proceso  $(h(X_n) - \sum_{k=0}^{n-1} (P - Id)h(X_k))_n$  es una martingala con respecto a  $(\mathcal{F}_n)_n$ .

*Demostración.* Sea  $Z_n := h(X_n) - \sum_{k=0}^{n-1} (P - Id)h(X_k)$ .

$$\mathbb{E}(Z_{n+1} - Z_n|\mathcal{F}_n) = \mathbb{E}(h(X_{n+1})|X_n) - h(X_n) - (P - Id)h(X_n).$$

$$\text{Pero } \mathbb{E}(h(X_{n+1})|X_n = i) = \sum_{j=0}^{+\infty} P_{ij} h(j) =: Ph(i)$$

$$\Rightarrow \mathbb{E}(h(X_{n+1})|X_n) = Ph(X_n)$$

$$\text{y } \forall n \in \mathbb{N} \quad \mathbb{E}(Z_{n+1} - Z_n|\mathcal{F}_n) = 0 \quad \diamond$$

**Comentario:** En el caso crítico  $m = 1$ , la función  $h = Id$  es armónica en el sentido que  $(P - Id)h \equiv 0$ :

$$((P - Id)h)(i) = \sum_{j \geq 0} P_{ij}h(j) - h(i) = \sum_{j \geq 0} jP_{ij} - i = \mathbb{E}(X_1|X_0 = i) - i = mi - i = 0.$$

Entonces, se verifica otra vez que  $X_n$  es una martingala.

## 1.6. La probabilidad de extinción

La probabilidad de extinción de la  $n$ -ésima generación  $\mathbb{P}(X_n = 0|X_0 = 1)$  vale  $f_n(0)$ . (Respuesta a la prejunta (1) de Watson). La probabilidad de extinción es el límite de  $f_n(0)$  y representa la probabilidad del suceso de  $\{X_n \rightarrow 0\}$ . Ella constituyó el punto de arranque del estudio del BGW. Se buscan condiciones para que el proceso no se extinga. Dichas condiciones dependen en gran medida del valor esperado  $m$  de la ley de reproducción. Esto originó la clasificación de los BGW en los tres tipos supercrítico, crítico y subcrítica, según  $m > 1$ ,  $m = 1$  o  $m < 1$ .

### Teorema 5 : Probabilidad de extinción

- *Caso (sub)crítico,  $m \leq 1$ :*

*La población se extingue casi seguramente, es decir,*

$$X_\infty = \lim_n X_n = 0 \text{ c.s.} \Leftrightarrow \forall i \in \mathbb{N}, \mathbb{P}(X_n = 0 \text{ para algún } n|X_0 = i) = 1$$

- *Caso supercrítico,  $m > 1$ :*

*La población sobrevive con probabilidad positiva, es decir,*

$$X_\infty = \lim_n X_n \in \{0, +\infty\} \text{ existe c.s. y } \mathbb{P}(X_\infty = +\infty|X_0 = i) = 1 - q^i$$

*donde  $q$  es la única raíz en  $[0, 1[$  de la ecuación  $f(s) = s$ .*

*Demostración:* El tiempo aleatorio de extinción  $T_0$  es el primer tiempo positivo de visita a 0,  $T_0 = \inf\{n \in \mathbb{N} : X_n = 0\}$ . Entonces  $\{T_0 < +\infty\} = \cup_{n \in \mathbb{N}} \{X_n = 0\} = \{X_\infty = 0\}$ . Así

$$\mathbb{P}(T_0 < +\infty|X_0 = i) = \lim_n \mathbb{P}(X_n = 0|X_0 = i) = \lim_n \mathbb{P}(X_n = 0|X_0 = 1)^i = \lim_n (f_n(0))^i.$$

La prueba del teorema es inmediata si se toma en cuenta el siguiente lema.

**Lema 1**  $\lim_n f_n(0) = q$  donde  $q$  es la raíz la mas pequeña en  $[0, 1]$  de la ecuación  $f(s) = s$ . Cuando  $m \leq 1$ ,  $q = 1$  y cuando  $m > 1$ ,  $q < 1$ .

*Demostración de este lema:*

- Si  $m \leq 1$ ,  $s \mapsto f'(s) - 1 \leq 0$  crece para todo  $s \in [0, 1]$ , luego  $s \mapsto f(s) - s$  decrece estrictamente desde el valor  $r_0 > 0$  hasta el valor 0. Es decir  $q = 1$ .

Dado que  $f$  crece,

$$0 < f(0) \leq f_2(0) \leq f_n(0) \leq \dots,$$

entonces  $f_n(0) \in [0, 1]$  crece y tiende a  $q$  cuando  $n \rightarrow \infty$ , donde  $q = 1$  es el único punto fijo de  $f$  en  $[0, 1]$ .

- Si  $m > 1$ ,  $s \mapsto f'(s) - 1$  toma valores negativos y después positivos, entonces  $s \mapsto f(s) - s$  decrece a partir de  $r_0$  y después crece hasta el valor 0; De donde existen valores negativos y (teorema de los valores intermedias) un valor  $q \in ]0, 1[$  tale que  $f(q) - q = 0$ .

Dado que  $f$  crece,

$$0 < r_0 = f(0) \leq f_2(0) \cdots \leq f_n(0) \cdots ,$$

entonces  $f_n(0)$  crece y tiende a la raíz más pequeña de la ecuación  $f(s) = s$ , que es exacto  $q$ .

◇

### Ejemplos:

- Reproducción binaria supercrítica,  $r_2 > 1/2$ :

$q$  es la raíz la mas pequeña de  $r_2 q^2 - q + 1 - r_2 = 0$ . Entonces  $q = \frac{1 - r_2}{r_2}$ .

- (W1): BGW crítico. 

$f_1(0)$	$f_2(0)$	$f_3(0)$	$f_4(0)$	$f_5(0)$
$\approx 0,333$	0,481	0,571	0,641	0,675

$q = \lim_n f_n(0)$  es raíz de  $\frac{1}{3}(1 - 2x + x^2) = 0$ , entonces  $q = 1$ .

- (W2):

$f_1(0)$	$f_2(0)$	$f_3(0)$	$f_4(0)$	$f_5(0)$	$f_6(0)$	$f_7(0)$	$f_8(0)$	$f_9(0)$	$f_{10}(0)$
$\approx 0,237$	0,346	0,410	0,450	0,477	0,496	0,510	0,520	0,527	0,533

BGW supercrítico,  $q = \lim_n f_n(0)$  es la única raíz en  $[0, 1)$  de

$$x^5 + 15x^4 + 90x^3 + 270x^2 - 619x + 243 = 0.$$

Se calcula:  $q \approx 0,553$ .

- Reproducción geométrica ( $a$ ).

La probabilidad de extinción de la  $n$ -ésima generación vale

$$f_n(0) = \begin{cases} \frac{m^n - 1}{m^{n+1} - 1} & \text{si } a \neq 1/2, \\ \frac{n}{n+1} & \text{si } a = 1/2 \text{ (caso critico)}. \end{cases}$$

En el caso supercrítico  $m > 1 \Leftrightarrow a < 1/2$ , la probabilidad de sobrevivir de la  $n$ -ésima generación es de

$$\mathbb{P}(X_n > 0 | X_0 = i) = 1 - \left(1 - m^n \frac{m-1}{m^{n+1}-1}\right)^i.$$

Entonces

$$q = \lim_n \frac{m^n - 1}{m^{n+1} - 1} = \frac{1}{m} = \frac{a}{1-a}$$

Ver la figura 4 con  $a = 1/3, m = 2$ . Luego

$$\mathbb{P}(X_n > 0 | X_0 = 20) = 1 - \left(1 - \frac{2^n}{2^{n+1}-1}\right)^{20} \approx 1 - (1/2)^{20}$$

a partir de  $n \geq 3$ .



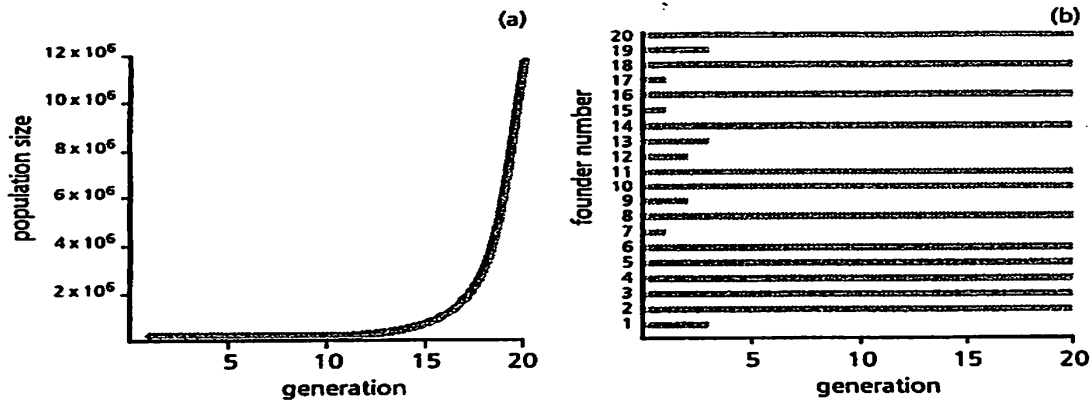


Figura 4:  $r$ =geométrica ( $\frac{1}{3}$ ) y  $X_0 = i = 20$  ( $m = 2$ ). (a) Tamaño de la población hasta  $n = 20$ . (b) Tiempo de supervivencia de cada familia. Cf [12] p.114

En el caso subcrítico  $m < 1$ ,

$$\mathbb{P}(X_n = 0 | X_0 = i) = \left(1 - m^n \frac{1 - m}{1 - m^{n+1}}\right)^i$$

que converge a 1 cuando  $n \rightarrow \infty$  a velocidad exponencial.

Ver la figura 5, con  $m = 0,95$  y  $i = 10$ .

- Reproducción cuasi-geométrica  $(b, a)$ : ( $m = (1 - b)/a$ ).  
La probabilidad de extinción de la  $n$ -ésima generación vale

$$f_n(0) = b_n = 1 - \frac{(a + b - 1)m^n}{b - (1 - a)m^n}.$$

En el caso supercrítico  $1 - a > b$ :

$$q = 1 - \lim_n \frac{m^n(a + b - 1)}{b - (1 - a)m^n} = 1 - \frac{a + b - 1}{a - 1} = \frac{b}{1 - a}$$

Aplicación al modelo de Lotka:  $q \approx 0,86$  (ver la figura 2).

## 1.7. Estimación práctica del tiempo de extinción

### 1.7.1. Los cálculos

Se puede estimar, solo con los dos primeros momentos de la ley  $r$ , la probabilidad de extinción de la  $n$ -ésima generación cuando la población inicial es  $i \geq 1$ .

**Teorema 6** Sea  $X$  un proceso de BGW donde la ley de reproducción tiene  $m$  como valor medio y  $\sigma^2$  como varianza.

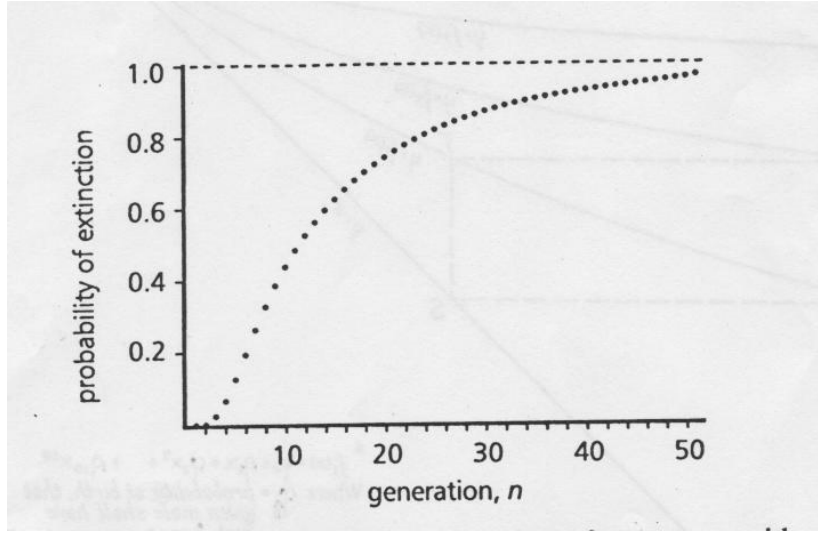


Figura 5:  $r$ =geométrica (0,51) y  $X_0 = i = 10$  ( $m = 0,95$ ). Probabilidad de extinción de las generaciones hasta  $n = 50$ . Cf [12] p.117

- Para cada  $i \geq 1$

$$i \mathbb{P}(X_n > 0 | X_0 = 1) \mathbb{P}(X_n = 0 | X_0 = 1)^{i-1} \leq \mathbb{P}(X_n > 0 | X_0 = i) \leq i \mathbb{P}(X_n > 0 | X_0 = 1)$$

- En particular, en el caso subcrítico  $m < 1$  y cuando  $n \rightarrow +\infty$ ,

$$i(1-m) \frac{m^{n+1}}{\sigma^2} \leq \mathbb{P}(X_n > 0 | X_0 = i) = \mathbb{P}(T_0 > n | X_0 = i) \leq i m^n$$

*Demostración:* Primera desigualdad:

$$\begin{aligned} \mathbb{P}(X_n > 0 | X_0 = i) = 1 - f_n(0)^i &= (1 - f_n(0))(1 + f_n(0) + f_n(0)^2 + \dots + f_n(0)^{i-1}) \\ &\geq (1 - f_n(0)) i f_n(0)^{i-1} \\ &= i \mathbb{P}(X_n > 0 | X_0 = 1) \mathbb{P}(X_n = 0 | X_0 = 1)^{i-1} \end{aligned}$$

Segunda desigualdad:

$$\begin{aligned} \mathbb{P}(X_n > 0 | X_0 = i) = 1 - f_n(0)^i &= (1 - f_n(0))(1 + f_n(0) + f_n(0)^2 + \dots + f_n(0)^{i-1}) \\ &\leq (1 - f_n(0)) i \\ &= i \mathbb{P}(X_n > 0 | X_0 = 1) \end{aligned}$$

Cuarta desigualdad:

$$\begin{aligned} \mathbb{P}(X_n > 0 | X_0 = 1) = \sum_{j \geq 1} \mathbb{P}(X_n = j | X_0 = 1) &\leq \sum_{j \geq 1} j \mathbb{P}(X_n = j | X_0 = 1) \\ &= \mathbb{E}(X_n | X_0 = 1) = m^n \end{aligned}$$

Tercera desigualdad:

$$\mathbb{P}(X_n > 0 | X_0 = 1) \mathbb{P}(X_n = 0 | X_0 = 1)^{i-1} \geq \mathbb{P}(X_n > 0 | X_0 = 1) (1 - m^n)^{i-1}.$$

Para  $n$  grande y  $m < 1$  vale  $(1 - m^n)^{i-1} \approx 1$ . Entonces queda por probar el siguiente lema.

**Lema 2** La estimación para  $n$  grande vale  $\mathbb{P}(X_n > 0 | X_0 = 1) \geq \frac{1-m}{\sigma^2} m^{n+1}$ .

*Demostración.* (En esta prueba supondremos la condición inicial  $X_0 = 1$ ).

Primero,  $\mathbb{E}^2(X_n \mathbf{1}_{\{X_n > 0\}}) \leq \mathbb{E}(X_n^2) \mathbb{E}(\mathbf{1}_{\{X_n > 0\}}) = \mathbb{E}(X_n^2) \mathbb{P}(X_n > 0)$ .

Entonces  $\mathbb{P}(X_n > 0) \geq \mathbb{E}^2(X_n) / \mathbb{E}(X_n^2) = m^{2n} / \mathbb{E}(X_n^2)$ .

Estimación del segundo momento  $\mathbb{E}(X_n^2)$ :  $\mathbb{E}(X_n^2) = f_n''(1) + \mathbb{E}(X_n) = f_n''(1) + m^n$ . Además

$$\begin{aligned} f_n''(1) &= f''(1)(f_{n-1}'(1))^2 + f'(1)f_{n-1}''(1) = f''(1)m^{2(n-1)} + mf_{n-1}''(1) \\ &= f''(1)(m^{2n-2} + m^{2n-3} + m^{2n-4} + \dots + m^{n-1}) \\ &= (\sigma^2 + m^2 - m)m^{n-1} \frac{1-m^n}{1-m}, \end{aligned}$$

luego

$$f_n''(1) + m^n = m^{n-1} \left( \sigma^2 \frac{1-m^n}{1-m} + m^{n+1} \right).$$

Entonces

$$\mathbb{P}(X_n > 0 | X_0 = 1) \geq \frac{m^{2n}}{m^{n-1} \left( \sigma^2 \frac{1-m^n}{1-m} + m^{n+1} \right)} \approx \frac{1-m}{\sigma^2} m^{n+1} \quad \diamond$$

### 1.7.2. Aplicación a la extinción de las ballenas negras en el Atlántico Norte

Cf [4], [12], [25] Lecture 3 y <http://www.rightwhale.ca>

La unidad de tiempo representa un año. La ley de reproducción en este modelo es:

$$r_0 = b, \quad r_1 = (1-b)(1-a), \quad r_2 = (1-b)a,$$

donde  $b$  es la probabilidad que una ballena hembra se muera durante el año siguiente, en el año 1994 esta probabilidad fue estimada en 0,06 [4]; y  $a$  es la probabilidad que una ballena engendre una ballena hembra pequeña durante el año siguiente, esta probabilidad fue estimada en 0,038. Entonces,

$$m = r_1 + 2r_2 = (1-b)(1+a) \approx 0,976 < 1 : \text{reproducción subcrítica!}$$

Cálculo de  $\sigma^2$ :  $\sigma^2 = r_1 + 4r_2 - m^2 = m(1-m) + 2(1-b)a \approx 0,095$ .

En el año 1994, la población de ballenas negras hembras en el Atlántico Norte era de  $i \approx 150$ . Si sus condiciones de vida no cambian, aplicando el teorema 6, esta población no se extinguirá antes de  $n$  años con probabilidad superior a 0,99, si

$$i(1-m) \frac{m^{n+1}}{\sigma^2} \geq 0,99 \Leftrightarrow 150 \cdot 0,024 \frac{(0,976)^{n+1}}{0,095} \geq 0,99 \Leftrightarrow n \leq 150,$$

entonces no antes del año 2144.

Al contrario, esta población se extinguirá antes de  $n$  años con probabilidad superior a 0,99 si

$$i m^n \leq 0,01 \Leftrightarrow 150 \cdot (0,976)^n \leq 0,01 \Leftrightarrow n \geq 395 \quad ,$$

entonces antes del año 2389, con probabilidad superior a 0,99, no habrá más ballenas negras hembras en el Atlántico Norte.

### 1.7.3. El valor medio del tiempo de extinción $T_0$ en el caso (sub)crítico

Ya se sabe que en este caso  $T_0 < +\infty$  c.s. Nos interesa estimar el valor medio de  $T_0$ , en particular para valores iniciales  $i$  grandes.

**Teorema 7** ■ En el caso crítico  $m = 1$ , para cada  $i \geq 1$

$$\mathbb{E}(T_0|X_0 = i) = +\infty.$$

■ Al contrario, en el caso subcrítico  $m < 1$ , para cada  $i \geq 1$ ,

$$\mathbb{E}(T_0|X_0 = i) < +\infty.$$

Además, si la ley  $\mathbf{r}$  satisface  $\sum_j j \log(j+1)r_j < +\infty$ , cuando  $i$  es grande:

$$\mathbb{E}(T_0|X_0 = i) \approx \frac{\log i}{|\log m|}$$

*Demostración.*

$$\mathbb{E}(T_0|X_0 = i) = \sum_{n \geq 0} \mathbb{P}(T_0 > n|X_0 = i) = \sum_n \mathbb{P}(X_n > 0|X_0 = i) = \sum_n 1 - f_n(0)^i$$

Caso crítico: Se puede probar que para  $n$  grande,

$$f_n(0) \approx 1 - \frac{2}{\sigma^2 n}.$$

Luego

$$1 - f_n(0)^i \approx 1 - \left(1 - \frac{2}{\sigma^2 n}\right)^i \approx \frac{2i}{\sigma^2 n}.$$

La serie  $\sum_n 1/n$  diverge, entonces  $\mathbb{E}(T_0|X_0 = i) = +\infty$ .

Caso subcrítico: Se ha visto en el teorema 6 que  $f_n(0) \geq 1 - m^n$ . Entonces

$$\mathbb{E}(T_0|X_0 = i) \leq \sum_n 1 - (1 - m^n)^i.$$

Para  $n$  grande,  $1 - (1 - m^n)^i \approx im^n$  y la serie  $\sum_n m^n$  converge; luego  $\mathbb{E}(T_0|X_0 = i) < +\infty$ .

En [12] Theorem 5.4 (ver los apéndices) los autores demostraron (con las notaciones  $N$  en lugar de  $i$  y  $\tau$  en lugar de  $T_0$ ) las estimaciones

$$\left(\frac{\log i - \log \log i}{|\log m|} - 1\right)\left(1 - \frac{1}{ic_1}\right) \leq \mathbb{E}(T_0|X_0 = i) \leq \frac{\log i}{|\log m|} + \frac{2 - m}{1 - m},$$

donde  $c_1 > 0$ . Se completa la prueba del teorema observando que los términos de izquierda y de derecha son equivalentes a  $\frac{\log i}{|\log m|}$  para  $i$  grande.  $\diamond$

#### Ejemplos:

■ Reproducción geométrica crítica:

$$f_n(0) = 1 - \frac{1}{n+1} \Rightarrow \mathbb{E}(T_0|X_0 = i) = \sum_n 1 - \left(1 - \frac{1}{n+1}\right)^i \approx \sum_n \frac{i}{n+1} = +\infty$$

■ Tiempo medio de extinción de las ballenas negras:

$$i = 150 \text{ grande y } m \approx 0,976 < 1 \Rightarrow \mathbb{E}(T_0|X_0 = i) \approx -\frac{\log 150}{\log 0,976} \approx 206 \in [150, 395]$$

## 1.8. Distribución cuasi-estacionaria de BGW

Antes de extinguirse, a menudo, el proceso subcrítico muestra un comportamiento cuasi-estacionario. Su distribución es invariante sobre la dinámica de reproducción condicionada a la no extinción.

### Teorema 8 : Existencia y cuasi-estacionaridad de la distribución de Yaglom

Sea  $X = (X_n)_n$  un proceso de BGW subcrítico y no degenerado, de ley de reproducción  $\mathbf{r}$ . Existe una ley de probabilidad sobre  $\mathbb{N}^*$ , que denotamos  $\bar{\mathbf{r}} = (\bar{r}_i)_{i \geq 1}$ , llamada **distribución de Yaglom**, definida por

$$\forall i \geq 1, \quad \bar{r}_i = \lim_{n \rightarrow +\infty} \mathbb{P}(X_n = i | X_n > 0);$$

además, si  $\bar{f}$  es la función generadora de la ley  $\bar{\mathbf{r}}$ ,

$$\forall s \in [0, 1], \quad 1 - \bar{f} \circ f(s) = m(1 - \bar{f}(s)). \quad (1)$$

Esta ley es cuasi-estacionaria, es decir

$$X_0 \sim \bar{\mathbf{r}} \quad \Rightarrow \quad \forall n \geq 1, \quad \mathbb{P}(X_n = \cdot | X_n > 0) = \bar{\mathbf{r}}.$$

*Demostración.* En primer lugar, se muestra que existe un límite de la función generadora condicionada  $\bar{f}_n$ , definida como:  $\bar{f}_n(s) := \mathbb{E}(s^{X_n} | X_n > 0)$ .

$$\begin{aligned} \bar{f}_n(s) &= \frac{\mathbb{E}(\mathbf{1}_{X_n > 0} s^{X_n})}{\mathbb{P}(X_n > 0)} = \frac{\mathbb{E}((1 - \mathbf{1}_{X_n = 0}) s^{X_n})}{1 - \mathbb{P}(X_n = 0)} \\ &= \frac{f_n(s) - f_n(0)}{1 - f_n(0)} = 1 - \frac{1 - f_n(s)}{1 - f_n(0)} \in [0, 1]. \end{aligned}$$

Es claro que  $\frac{1 - f_n(s)}{1 - f_n(0)} \leq 1$ . Demostramos que  $n \mapsto \frac{1 - f_n(s)}{1 - f_n(0)}$  crece.

$$\frac{1 - f_{n+1}(s)}{1 - f_{n+1}(0)} = \frac{1 - f(f_n(s))}{1 - f_n(s)} \cdot \frac{1 - f_n(0)}{1 - f(f_n(0))} \cdot \frac{1 - f_n(s)}{1 - f_n(0)} \geq \frac{1 - f_n(s)}{1 - f_n(0)},$$

ya que la función  $u \mapsto \frac{1 - f(u)}{1 - u}$  crece (hasta el valor  $m$ ) y entonces tiene un valor más grande para  $u = f_n(s)$  que para  $u = f_n(0)$ . Así, para cada  $s \in [0, 1]$  existe el límite de la sucesión  $n \mapsto \frac{1 - f_n(s)}{1 - f_n(0)}$ ; definimos  $\bar{f}(s)$  igual al complemento a 1 de este límite.

Para probar la ecuación (1), se nota que:

$$1 - \bar{f}(s) = \lim_n \frac{1 - f_{n+1}(s)}{1 - f_{n+1}(0)} = \lim_n \frac{1 - f_n(f(s))}{1 - f_n(0)} \lim_n \frac{1 - f_n(0)}{1 - f(f_n(0))},$$

donde el segundo límite es  $1 - \bar{f} \circ f(s)$  y  $\lim_n \frac{1 - f_n(0)}{1 - f(f_n(0))} = \lim_{u \rightarrow 1} \frac{1 - u}{1 - f(u)} = \frac{1}{m}$ .

Mostramos ahora que  $\bar{\mathbf{r}}$  es cuasi-estacionaria: Se nota  $\mathbb{P}_{\bar{\mathbf{r}}}$  la ley  $\mathbb{P}(\cdot | X_0 \sim \bar{\mathbf{r}})$ . Note que

$$\begin{aligned} \mathbb{P}_{\bar{\mathbf{r}}}(X_1 = 0) &= \sum_{i \geq 1} \mathbb{P}(X_1 = 0 | X_0 = i) \bar{r}_i = \sum_{i \geq 1} (r_0)^i \bar{r}_i = \bar{f}(r_0) \\ \text{y} \quad \mathbb{E}_{\bar{\mathbf{r}}}(s^{X_1}) &= \sum_{i \geq 1} f(s)^i \bar{r}_i = \bar{f} \circ f(s). \end{aligned}$$

Ahora,

$$\mathbb{E}_{\bar{\mathbf{r}}}(s^{X_1} | X_1 > 0) = \frac{\mathbb{E}_{\bar{\mathbf{r}}}(\mathbf{1}_{X_1 > 0} s^{X_1})}{\mathbb{P}_{\bar{\mathbf{r}}}(X_1 > 0)} = \frac{\mathbb{E}_{\bar{\mathbf{r}}}(s^{X_1}) - \mathbb{P}_{\bar{\mathbf{r}}}(X_1 = 0)}{1 - \mathbb{P}_{\bar{\mathbf{r}}}(X_1 = 0)} = \frac{\bar{f} \circ f(s) - \bar{f}(r_0)}{1 - \bar{f}(r_0)}.$$

Se puede calcular  $\bar{f}(r_0)$  usando la ecuación (1):

$$1 - \bar{f} \circ f(0) = m(1 - \bar{f}(0)) \Rightarrow \bar{f}(r_0) = 1 - m(1 - 0) = 1 - m.$$

$$\text{Entonces } \mathbb{E}_{\bar{\mathbf{r}}}(s^{X_1} | X_1 > 0) = \frac{\bar{f} \circ f(s) - 1 + m}{m} = 1 - \frac{1 - \bar{f} \circ f(s)}{m} \stackrel{(1)}{=} \bar{f}(s),$$

es decir  $\mathbb{P}(X_1 = \cdot | X_1 > 0, X_0 \sim \bar{\mathbf{r}}) = \bar{\mathbf{r}}$ . La prueba para  $n > 1$  es semejante.  $\diamond$

## Ejemplos

- Reproducción geométrica  $\mathcal{G}_a$  subcrítica ( $a > 1/2$ ).

Sea  $\bar{f}_n(s) := \mathbb{E}(s^{X_n} | X_n > 0)$ .

$$\bar{f}_n(s) = \frac{f_n(s) - f_n(0)}{1 - f_n(0)} = \frac{a_n s}{1 - (1 - a_n)s} \Rightarrow \lim_n \bar{f}_n(s) = \left(1 - \frac{1-a}{a}\right) \frac{s}{1 - \frac{1-a}{a}s} =: \bar{f}(s)$$

ya que  $\lim_n a_n := 1 - \frac{1-a}{a}$ .

En este caso, la distribución de Yaglom  $\bar{\mathbf{r}}$  es una ley geométrica de parámetro  $2 - \frac{1}{a} = 1 - m$  sobre  $\mathbb{N}^*$  definida por

$$\bar{r}_i = \left(1 - \frac{1-a}{a}\right) \left(\frac{1-a}{a}\right)^{i-1}, \quad i \geq 1.$$

- Reproducción cuasi-geométrica subcrítica,  $1 - b < a$ . (Por ejemplo  $a = 0,65$  y  $b = 0,4$ ).

$$\bar{f}_n(s) = \frac{f_n(s) - f_n(0)}{1 - f_n(0)} = \frac{a_n s}{1 - (1 - a_n)s} \Rightarrow \lim_n \bar{f}_n(s) = \left(1 - \frac{1-a}{b}\right) \frac{s}{1 - \frac{1-a}{b}s} =: \bar{f}(s)$$

ya que  $\lim_n \alpha_n := 1 - \frac{1-a}{b}$ .

En este caso, la distribución de Yaglom  $\bar{\mathbf{r}}$  es una ley geométrica de parámetro  $\frac{a+b-1}{b}$  sobre  $\mathbb{N}^*$  definida por

$$\bar{r}_i = \left(1 - \frac{1-a}{b}\right) \left(\frac{1-a}{b}\right)^{i-1}, \quad i \geq 1.$$

## 1.9. El proceso condicionado a la no extinción en el futuro lejano

Al lugar de condicionar la dinámica de reproducción a la no extinción en el tiempo presente, se puede condicionar a la no extinción en un tiempo futuro lejano. Como se comporta

$$\mathbb{P}(X_n = \cdot | X_{n+k} > 0)$$

cuando  $k$  es grande?

**Teorema 9 : Existencia del proceso condicionado**

Sea  $X = (X_n)_n$  un proceso de BGW subcrítico y no degenerado, de ley de reproducción  $\mathbf{r}$  verificando  $\sum_j (j \log j) r_j < +\infty$ . Existe un proceso de Markov  $X^* = (X_n^*)_n$  definido por

$$\mathbb{P}(X_n^* = j) = \lim_{k \rightarrow +\infty} \mathbb{P}(X_n = j | X_{n+k} > 0).$$

Su matriz de transición  $P^*$  verifica:

$$(P_{ij}^*)^n := \mathbb{P}(X_n^* = j | X_0^* = i) = \frac{j}{i m^n} P_{ij}^n, \forall i, j \geq 1. \quad (2)$$

A veces, este proceso se llama Q-proceso en la literatura.

*Demostración:*

$$\begin{aligned} \mathbb{P}(X_n = j | X_0 = i, X_{n+k} > 0) &= \frac{\mathbb{P}(X_0 = i, X_n = j, X_{n+k} > 0)}{\mathbb{P}(X_{n+k} > 0 | X_0 = i) \mathbb{P}(X_0 = i)} \\ &= \frac{\mathbb{P}(X_{n+k} > 0 | X_n = j, X_0 = i) \mathbb{P}(X_n = j | X_0 = i) \mathbb{P}(X_0 = i)}{(1 - f_{n+k}(0)^i) \mathbb{P}(X_0 = i)} \\ &= \frac{\mathbb{P}(X_{n+k} > 0 | X_n = j) \mathbb{P}(X_n = j | X_0 = i)}{1 - f_{n+k}(0)^i} \\ &= \frac{1 - f_k(0)^j}{1 - f_{n+k}(0)^i} P_{ij}^n. \end{aligned}$$

Kolmogorov demostró la siguiente estimación:

$$\text{Si } \sum_j (j \log j) r_j < +\infty, \exists c > 0, \mathbb{P}(X_n = 0) = f_n(0) \approx_n 1 - c m^n.$$

Ver [25] Chapter 3 Theorem 1 por una demostración. (De hecho, en el ejemplo  $\mathbf{r} = \mathcal{G}_a$ ,

$$f_n(0) = \frac{1 - m^n}{1 - m^{n+1}} \approx_n 1 - (1 - m)m^n.)$$

Entonces,

$$\begin{aligned} \frac{1 - f_k(0)^j}{1 - f_{n+k}(0)^i} &\approx_k \frac{1 - (1 - c m^k)^j}{1 - (1 - c m^{n+k})^i} \\ &\approx_k \frac{1 - (1 - j c m^k)}{1 - (1 - i c m^{n+k})} = \frac{j}{i m^n}. \end{aligned}$$

Se puede verificar la propiedad de Markov para  $X^*$  :

$$\mathbb{P}(X_{n+l}^* = j | X_0^*, \dots, X_{l-1}^*, X_l^* = i) = (P_{ij}^*)^n \quad \diamond$$

Ademas, el proceso condicionado muestra un comportamiento estable cuando  $n$  es grande.

**Proposición 10** *Sobre las mismas hipótesis que en el Teorema antes, existe una distribución estacionaria  $\mathbf{r}^*$  para  $X^*$  igual a la distribución de Yaglom sesgada:*

$$r_j^* = c j \bar{r}_j.$$

*Demostración:*

$$\lim_n (P_{1j}^*)^n = j \lim_n \frac{P_{1j}^n}{m^n}.$$

$$\text{Pero } \bar{r}_j = \lim_n \mathbb{P}(X_n = j | X_n > 0) = \lim_n \frac{1}{1 - f_n(0)} P_{1j}^n$$

$$= \lim_n \frac{1}{1 - (1 - cm^n)} P_{1j}^n = \frac{1}{c} \lim_n \frac{P_{1j}^n}{m^n}$$

$$\Rightarrow \lim_n (P_{1j}^*)^n = c j \bar{r}_j. \quad \diamond$$



## 2. Unas propiedades del proceso BGW de tiempo continuo

En el modelo de Bienaymé-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 proceso de ramificación BGW de tiempo continuo (BGWc)  $X = (X_t)_t$  cada uno de esos individuos tiene una duración de vida aleatoria, al término de la cual se muere o se produce un número aleatorio de individuos descendientes. Así coexisten siempre varias generaciones. Como en el modelo discreto, los nuevos individuos viven, mueren y se reproducen independientemente con las mismas reglas. Cada individuo inicial genera un árbol aleatorio, como se muestra en la Figura 6a. La ramificación se refiere a la estructura de árbol. Las ramas representan las líneas de descendencia y su evolución temporal. El proceso estocástico al tiempo  $t$ ,  $X_t$ , es el número de individuos presentes al tiempo  $t$ .

Además, supongamos que las leyes probabilísticas de vida de cada individuo son las mismas, exponenciales de parámetro  $\alpha$  - que notamos  $\mathcal{Exp}(\alpha)$ ; la reproducción de cada individuo tiene la misma ley  $\mathbf{r}$  (supongamos que  $r_1 = 0$ ).

### 2.1. El BGWc como proceso markoviano de salto puro

**Definición 1** *Un proceso  $X = (X_t)_{t \geq 0}$  a valores en un espacio  $\mathbf{E}$  numerable es un proceso de salto puro si existe una sucesión  $(Z_n)_n \in \mathbf{E}$  con  $Z_{n+1} \neq Z_n$  y una sucesión de tiempos aleatorios  $\tau_0 = 0 < \tau_1 < \dots < \tau_n < \dots$  tal que*

$$\forall t > 0, \quad X_t = \sum_{n \geq 0} \mathbb{I}_{\tau_n \leq t < \tau_{n+1}} Z_n.$$

De esta forma  $\tau_n$  es el tiempo en el que ocurre el  $n$ -ésimo salto, y  $Z_n$  es el  $(n+1)$ -ésimo estado en  $\mathbf{E}$  visitado por el proceso  $X$ .

Además,  $X$  verifica la propiedad markoviana si  $\forall 0 \leq t_1 < \dots < t_k < t_{k+1}$  y  $x_1, \dots, x_{k+1} \in \mathbf{E}$ ,

$$\mathbb{P}(X_{t_{k+1}} = x_{k+1} | X_{t_1} = x_1, \dots, X_{t_k} = x_k) = \mathbb{P}(X_{t_{k+1}} = x_{k+1} | X_{t_k} = x_k).$$

El proceso es homogéneo si  $\mathbb{P}(X_t = \cdot | X_s) = \mathbb{P}(X_{t-s} = \cdot | X_0)$ .

La propiedad de Markov dice que la probabilidad condicional de un futuro estado en el tiempo  $t_{k+1}$ , dado el estado presente en el tiempo  $t_k$  y todos los estados pasados, solo depende del presente estado y es independiente del pasado. Al proceso de tiempo discreto  $Z_n = X_{\tau_n}$  se llamara el **esqueleto** del proceso de tiempo continuo  $X$ . Si  $X$  es markoviano, su esqueleto  $(Z_n)_n$  es una cadena de Markov.

**Teorema 11** *Sea  $X = (X_t)_{t \geq 0}$  un proceso BGW de tiempo continuo. Es un proceso markoviano homogéneo de salto puro a valores en  $\mathbb{N}$ . El proceso esqueleto  $(Z_n)_n$  de estados visitados es un BGW y los tiempos aleatorios  $(\tau_{n+1} - \tau_n)_{n \geq 0}$  son independientes. Además, si  $Z_n = i$ , la distribución del tiempo  $\tau_{n+1} - \tau_n$  transcurrido en  $i$  es  $\mathcal{Exp}(i\alpha)$ .*

Ver la figura 6b.

*Demostración.* Por la construcción hecha, el esqueleto  $(Z_n)_n$  del proceso  $X$  es un BGW. Si  $Z_0$  vale  $i$ , es decir que el número de individuos en la población al tiempo inicial es  $i$ . El primero salto de la población ocurre cuando el primero de estos  $i$  individuos muere y se reproduce, entonces en el tiempo  $\min\{V_1, \dots, V_i\}$  donde  $V_k$  es la duración aleatoria de la vida del  $k$ -ésimo individuo. Se sabe que  $\forall k \geq 1, V_k \sim \mathcal{Exp}(\alpha)$ .

**Lema 3** Sean  $V_1$  y  $V_2$  dos v.a.i. con distribución exponencial con parámetros  $\alpha_1$  y  $\alpha_2$  respectivamente. Entonces  $\min\{V_1, V_2\} \sim \mathcal{Exp}(\alpha_1 + \alpha_2)$ .

Por la demostración (ver también Proposición 3.3.2 en [26]), se observa que para  $x > 0$ ,

$$\{\min\{V_1, V_2\} > x\} = \{V_1 > x\} \cap \{V_2 > x\} \Rightarrow \mathbb{P}(\min\{V_1, V_2\} > x) = e^{-\alpha_1 x} e^{-\alpha_2 x} = e^{-(\alpha_1 + \alpha_2)x}. \diamond$$

Esto implica que el primero salto del proceso de ramificación BGWc dado  $Z_0 = i$ , será en el instante

$$\tau_1 = \min\{V_1, \dots, V_i\} \sim \mathcal{Exp}(\alpha + \dots + \alpha) = \mathcal{Exp}(i\alpha).$$

La propiedad de Markov es consecuencia del hecho que  $(Z_n)_n$  es una cadena de Markov y de la falta de memoria de la distribución exponencial: si  $V$  es una v.a. exponencial con parámetro  $\alpha$ , para  $x, y > 0$ ,

$$\frac{\mathbb{P}(\{V > x + y\} \cap \{V > x\})}{\mathbb{P}(V > x)} = \frac{e^{-\alpha(x+y)}}{e^{-\alpha x}} = e^{-\alpha y} \Rightarrow \mathbb{P}(V > x + y | V > x) = \mathbb{P}(V > y).$$

Luego,

$$\mathbb{P}(\tau_{n+1} - \tau_n > x | X_{\tau_n} = i) = e^{-i\alpha x} \Rightarrow \mathbb{P}(\tau_{n+1} - \tau_n \in \cdot | Z_n = i) = \mathcal{Exp}(i\alpha). \diamond$$

## 2.2. Las ecuaciones de Chapman-Kolmogorov

Desde ahora en adelante  $P(t)$  denotará la matriz (infinita) con entradas las probabilidades de transición del proceso BGWc:

$$P_{ij}(t) = \mathbb{P}(X_t = j | X_0 = i), \quad i, j \in \mathbb{N}.$$

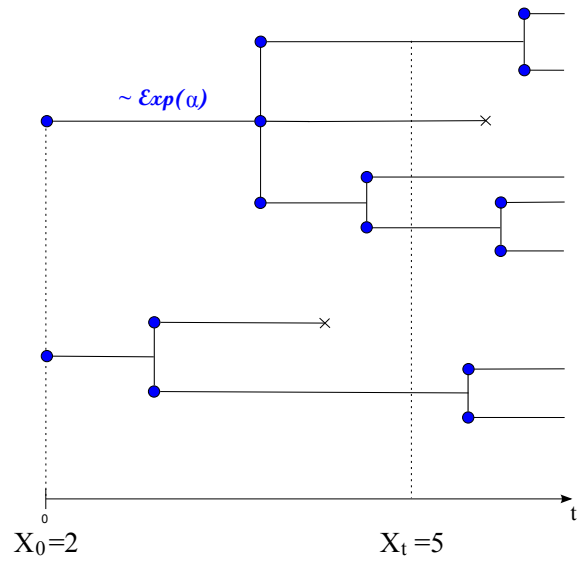
**Teorema 12** Sea  $X = (X_t)_{t \geq 0}$  un proceso BGW de tiempo continuo y  $(P(t))_t$  sus matrices de transición. Ellas verifican la propiedad de semigrupo u ecuación de Chapman-Kolmogorov

$$\forall s, t > 0, \quad P(t+s) = P(s)P(t) = P(t)P(s). \quad (3)$$

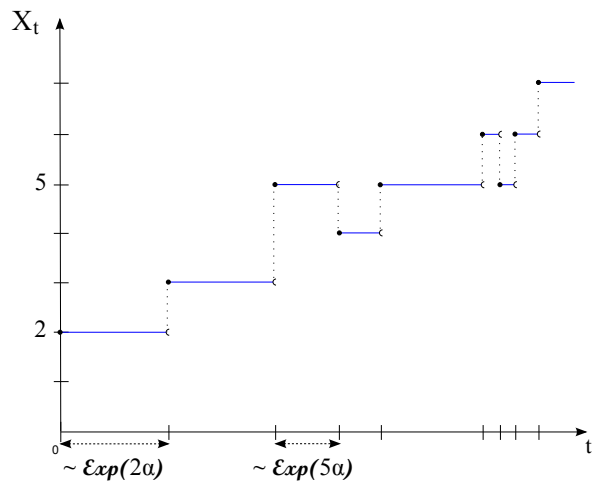
*Demostración.* (ver también Teorema 3.4.1 en [26])

Utilizando las propiedades de uno proceso de Markov homogéneo de tiempo continuo tenemos que:

$$\begin{aligned} P_{ij}(t+s) = \mathbb{P}(X_{t+s} = j | X_0 = i) &= \sum_{k \in \mathbb{N}} \mathbb{P}(X_{t+s} = j, X_t = k | X_0 = i) \\ &= \sum_{k \in \mathbb{N}} \mathbb{P}(X_{t+s} = j | X_t = k) \mathbb{P}(X_t = k | X_0 = i) \\ &= \sum_{k \in \mathbb{N}} P_{kj}(s) P_{ik}(t) = (P(t)P(s))_{ij} \quad \diamond \end{aligned}$$



(a) La genealogía



(b) El proceso de salto puro

Figura 6: El proceso BGWc

### 2.2.1. El generador infinitesimal

**Proposición 13** *Existen las derivadas respecto de  $t$  por  $t = 0$  de la matriz de transición  $P(t)$ :*

▪

$$\frac{\partial}{\partial t} P_{ii}(t)|_{t=0} = -\lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} =: Q_{ii}, \quad i \in \mathbb{N}.$$

$q_i := -Q_{ii}$  es la tasa (o velocidad) de salida del proceso  $X$  del estado  $i$ .

En el caso del BGWc  $q_i = i\alpha$ . En particular,  $q_0 = 0$  y el estado 0 es absorbente.

▪ Para cada  $j \neq i \in \mathbb{N}$ ,

$$\frac{\partial}{\partial t} P_{ij}(t)|_{t=0} = \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} =: Q_{ij}.$$

$Q_{ij}$  es la tasa (o velocidad) de salto del proceso  $X$  de un estado  $i$  a otro  $j$ .

En el caso del BGWc

$$\text{si } j \geq i - 1, j \neq i, \quad Q_{ij} = i\alpha r_{j-i+1}; \quad \text{si } j < i - 1, Q_{ij} = 0 \text{ y } Q_{ii} = -i\alpha.$$

La matriz  $\frac{\partial}{\partial t} P(t)|_{t=0} =: Q$ , cuyas entradas son  $(Q_{ij})_{i,j \in \mathbb{N}}$ , se llama generador infinitesimal del proceso markoviano  $X$ . Ella verifica para cada  $i \in \mathbb{N}$ ,  $\sum_j Q_{ij} = 0$ .

*Demostración.*

$$\begin{aligned} 1 - P_{ii}(t) &= \mathbb{P}(X_t \neq i | X_0 = i) \stackrel{t \rightarrow 0}{\approx} \mathbb{P}(\tau_1 < t | X_0 = i) = 1 - e^{-i\alpha t} \stackrel{t \rightarrow 0}{\approx} i\alpha t \\ \Rightarrow \quad \lim_{t \rightarrow 0} \frac{1 - P_{ii}(t)}{t} &= \lim_{t \rightarrow 0} \frac{i\alpha t}{t} = i\alpha. \end{aligned}$$

Sea  $j \neq i$ .

$$\begin{aligned} P_{ij}(t) &= \mathbb{P}(X_t = j | X_0 = i) \stackrel{t \rightarrow 0}{\approx} \mathbb{P}(\tau_1 < t, Z_1 = j | X_0 = i) \\ &= \mathbb{P}(\tau_1 < t) \mathbb{P}(Z_1 = j | Z_0 = i) = (1 - e^{-i\alpha t}) \mathbb{P}(Y = j - (i - 1)) \\ \Rightarrow \quad \lim_{t \rightarrow 0} \frac{P_{ij}(t)}{t} &= \lim_{t \rightarrow 0} \frac{1 - e^{-i\alpha t}}{t} r_{j-i+1} = i\alpha r_{j-i+1}. \end{aligned}$$

$$\text{y} \quad \sum_j Q_{ij} = \sum_j \frac{\partial}{\partial t} P_{ij}(t)|_{t=0} = \frac{\partial}{\partial t} \sum_j P_{ij}(t)|_{t=0} = \left(\frac{\partial}{\partial t} 1\right) = 0. \quad \diamond$$

**Forma de la matriz  $Q$ , generador infinitesimal de un BGWc general.**

$$Q = \alpha \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ r_0 & -1 & r_2 & r_3 & \dots \\ 0 & 2r_0 & -2 & 2r_2 & \dots \\ 0 & 0 & 3r_0 & -3 & \dots \\ \vdots & & & \ddots & \dots \end{pmatrix}$$

En particular, para cada función  $h$  acotada,

$$Qh(i) = i\alpha \left( \sum_{j \geq i-1, j \neq i} r_{j-i+1} h(j) - h(i) \right) = i\alpha \sum_{k \geq 0} r_k (h(i+k-1) - h(i)) \quad (4)$$

**Ejemplo** con reproducción binaria crítica:

$$Q = \alpha \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1/2 & -1 & 1/2 & 0 & \dots \\ 0 & 1 & -2 & 1 & \dots \\ 0 & 0 & 3/2 & -3 & \dots \\ \vdots & & & \ddots & \dots \end{pmatrix}$$

### 2.2.2. Las ecuaciones de Kolmogorov hacia adelante y hacia atrás

**Teorema 14** *Se cumplen las siguientes ecuaciones diferenciales que describen la evolución temporal de la ley de probabilidad del proceso.*

- *Ecuación de Kolmogorov hacia atrás:*

$$\forall t \geq 0 \quad \frac{\partial}{\partial t} P(t) = QP(t) \quad y \quad P(0) = Id. \quad (5)$$

Además, para cada función  $h$  acotada, la función  $u(i, t) := \mathbb{E}(h(X_t) | X_0 = i)$  resuelve la ecuación diferencial

$$\frac{\partial}{\partial t} u(\cdot, t) = Qu(\cdot, t), \quad u(0) = h. \quad (6)$$

- *Ecuación de Kolmogorov hacia adelante:*

$$\forall t \geq 0 \quad \frac{\partial}{\partial t} P(t) = P(t)Q \quad y \quad P(0) = Id. \quad (7)$$

Además, la ley de probabilidad de  $X_t$ , notada  $\mu_t$ , resuelve la ecuación diferencial de Fokker-Planck

$$\frac{\partial}{\partial t} \mu(t) = \mu(t)Q. \quad (8)$$

*Demostración.* (Ver también Teorema 3.4.4 en [26]) Probaremos (5) ya que (7) se demuestra razonando de manera análoga. Utilizando la ecuación de Chapman-Kolmogorov (3):

$$\frac{\partial}{\partial t} P(t) = \lim_{h \rightarrow 0} \frac{P(t+h) - P(t)}{h} = \lim_{h \rightarrow 0} \frac{P(h)P(t) - P(t)}{h} = QP(t).$$

Por otro lado, 
$$\begin{aligned} \frac{\partial}{\partial t} \mu_j(t) &= \frac{\partial}{\partial t} \sum_i \mu_i(0) P_{ij}(t) = \sum_i \mu_i(0) \frac{\partial}{\partial t} P_{ij}(t) \stackrel{(7)}{=} \sum_i \mu_i(0) \sum_k P_{ik}(t) Q_{kj} \\ &= \sum_k \mu_k(t) Q_{kj} = (\mu(t)Q)_j \quad \diamond \end{aligned}$$

**Corolario 15** *La matriz de transición  $P(t)$  del proceso BGWc resuelve las ecuaciones diferenciales de Kolmogorov:*

$$\forall i, j \in \mathbb{N}, \quad \frac{\partial}{\partial t} P_{ij}(t) = i\alpha r_0 P_{i-1,j}(t) - i\alpha P_{ij}(t) + i\alpha \sum_{k \geq 1} r_{k+1} P_{i+k,j}(t), \quad P_{ij}(0) = \delta_{ij}. \quad (9)$$

Por ejemplo, si la reproducción es binaria crítica ( $r_0 = r_2 = 1/2$ ),

$$\forall i, j \in \mathbb{N}, \quad \frac{\partial}{\partial t} P_{ij}(t) = \frac{i\alpha}{2} (P_{i+1,j}(t) + P_{i-1,j}(t) - 2P_{ij}(t)).$$

## 2.3. La ley del proceso al tiempo $t$

La ley del proceso BGWc  $X_t$  al tiempo  $t$  se define por su función generatriz:

$$F(s, t) := \mathbb{E}(s^{X_t} | X_0 = 1) = \sum_{j \geq 0} \mathbb{P}(X_t = j | X_0 = 1) s^j = \sum_{j \geq 0} P_{1j}(t) s^j.$$

La serie converge uniformemente con  $t$ , al menos para  $|s| < 1$  (Ver [26], Apéndice B).

### 2.3.1. La función generatriz del proceso al tiempo $t$

#### Proposición 16 Estructura diferencial de la función generatriz de BGWc

La función generatriz  $F(s, t)$  verifica para todos  $0 \leq s \leq 1, t > 0$  la siguiente ecuación diferencial:

$$\frac{\partial}{\partial t} F(s, t) = \tilde{f}(F(s, t)), \quad F(s, 0) = s \quad (10)$$

donde  $\tilde{f}(s) := \alpha(f(s) - s)$ .

*Demostración.* De la ecuación (6) por  $u(i, t) = \mathbb{E}(s^{X_t} | X_0 = i)$  ( $h(j) = s^j$ ), es fácil ver que

$$\frac{\partial}{\partial t} F(s, t) = \frac{\partial}{\partial t} u(1, t) = (Qu(\cdot, t))_1 = \sum_{j \geq 0} Q_{1j} u(j, t) = \sum_{j \geq 0} Q_{1j} u(1, t)^j.$$

Utilizando la Proposición 13,  $Q_{1j} = \alpha r_j, j \neq 1$  y  $Q_{11} = -\alpha$ . Entonces

$$\frac{\partial}{\partial t} F(s, t) = \alpha \left( \sum_{j \geq 0} r_j F(s, t)^j - F(s, t) \right) = \tilde{f}(F(s, t)). \quad \diamond$$

#### Ejemplos:

- Reproducción binaria crítica ( $m = 1$ ):  $\tilde{f}(s) = \frac{\alpha}{2}(1 - s)^2$ .

$F$  verifica

$$\frac{\partial}{\partial t} F(s, t) = \frac{\alpha}{2}(1 - F(s, t))^2, \quad F(s, 0) = s.$$

Entonces

$$F(s, t) = 1 - \frac{1 - s}{1 + \frac{\alpha}{2}t(1 - s)}.$$

- Reproducción binaria no crítica:

$$\tilde{f}(s) = \alpha(r_2(1 - s)^2 + (1 - 2r_2)(1 - s)) = \alpha(1 - s) \left( \frac{m}{2}(1 - s) + 1 - m \right)$$

Entonces

$$F(s, t) = 1 - \frac{1 - s}{e^{-\alpha(m-1)t} + \frac{m}{2(m-1)}(1 - e^{-\alpha(m-1)t})(1 - s)}.$$

- (W1):  $\tilde{f}(z) = \frac{\alpha}{3}(1 - z)^2$  y

$$F(s, t) = 1 - \frac{1 - s}{1 + \frac{\alpha}{3}t(1 - s)}.$$

- (W2):  $\tilde{f}(z) = \alpha\left(\left(\frac{3+z}{4}\right)^5 - z\right)$  y  $F$  verifica

$$\frac{\partial}{\partial t}F(s, t) = \alpha\left(\left(\frac{3 + F(s, t)}{4}\right)^5 - F(s, t)\right), \quad F(s, 0) = s.$$

- Reproducción crítica de tipo  $\gamma$ -estable :  $\tilde{f}(s) = 1 - s + c(1 - s)^\gamma$  con  $\gamma \in ]1, 2]$  y  $c \geq 0$ .  
Entonces

$$F(s, t) = 1 - \frac{1 - s}{(1 + c(\gamma - 1)t(1 - s)^{\gamma-1})^{1/\gamma-1}}.$$

### 2.3.2. Los primeros momentos

Se sabe que el valor esperado de la variable  $X_t$ , si existe, está dado por:  $\mathbb{E}(X_t) = \frac{\partial}{\partial s}F(s, t)|_{s=1}$ . Calculando la derivada respecto de  $s$  en la ecuación (10), concluimos que  $m(t) := \mathbb{E}(X_t)$  es solución de la ecuación

$$\begin{aligned} \frac{\partial}{\partial t}m(t) &= \frac{\partial}{\partial s}\tilde{f}(F(s, t))|_{s=1} = \tilde{f}'(F(1, t))m(t) = \tilde{f}'(1) m(t) = \alpha(m - 1) m(t) \\ \Rightarrow \mathbb{E}(X_t) &= e^{\alpha(m-1)t} \mathbb{E}(X_0). \end{aligned}$$

La constante  $\rho := \alpha(m - 1)$  se llama **parámetro de Malthus**.

Así, de manera semejante al modelo discreto BGW asociado, se encuentran tres casos distintos:

$\rho = 0 \Rightarrow (X_t)_t$  crítico y  $\forall t > 0, \mathbb{E}(X_t) = \mathbb{E}(X_0)$ ;

$\rho < 0 \Rightarrow (X_t)_t$  subcrítico y  $\lim_{t \rightarrow +\infty} \mathbb{E}(X_t) = 0$ ;

$\rho > 0 \Rightarrow (X_t)_t$  supercrítico y  $\lim_{t \rightarrow +\infty} \mathbb{E}(X_t) = +\infty$ .

Además, cuando  $X_0 = 1$ ,  $\mathbb{E}(X_t^2 - X_t) = \frac{\partial^2}{\partial s^2}F(s, t)|_{s=1}$ , y calculando la segunda derivada respecto de  $s$  en la ecuación (10), concluimos que  $v(t) := \mathbb{E}(X_t^2|X_0 = 1)$  es solución de la ecuación

$$\begin{cases} \frac{\partial}{\partial t}v(t) = \tilde{f}''(F(1, t)) m^2(t) + \rho v(t) = (\alpha\sigma^2 + m\rho)e^{2\rho t} + \rho v(t) \\ v(0) = 1 \end{cases}$$

$$\Rightarrow \mathbb{E}(X_t^2|X_0 = 1) = \begin{cases} (\alpha\sigma^2 + m\rho) \frac{e^{2\rho t} - e^{\rho t}}{\rho} + e^{\rho t} & \text{si } \rho \neq 0, \\ \alpha\sigma^2 t + 1 & \text{si } \rho = 0 \end{cases}$$

y

$$\text{Var}(X_t|X_0 = 1) = \begin{cases} (\alpha\sigma^2 + \rho(m - 1)) \frac{e^{2\rho t} - e^{\rho t}}{\rho} & \text{si } \rho \neq 0, \\ \alpha\sigma^2 t & \text{si } \rho = 0. \end{cases}$$

## 2.4. El "problema" de martingalas

### Proposición 17 : Una familia de martingalas

Para cada función  $h$  acotada, el proceso  $\left(h(X_t) - h(X_0) - \int_0^t (Qh)(X_\tau) d\tau\right)_{t>0}$  es una martingala respecto a la filtración  $(\mathcal{F}_t)_t$ , donde  $\mathcal{F}_t := \sigma(X_s, 0 \leq s \leq t)$ .

*Demostración.* (Ver [5], Chapter 9-1) Se debe mostrar que, si  $Z_t = h(X_t) - h(X_0) - \int_0^t (Qh)(X_\tau) d\tau$ ,

$$\forall s < t, \quad \mathbb{E}(Z_t|\mathcal{F}_s) = Z_s.$$

Pero,

$$\begin{aligned}
\mathbb{E}(Z_t|\mathcal{F}_s) &= \mathbb{E}(h(X_t)|\mathcal{F}_s) - h(X_0) - \int_0^s \mathbb{E}(Qh(X_\tau)|\mathcal{F}_s)d\tau - \int_s^t \mathbb{E}(Qh(X_\tau)|\mathcal{F}_s)d\tau \\
&= (P(t-s)h)(X_s) - h(X_0) - \int_0^s Qh(X_\tau)d\tau - \int_s^t (P(\tau-s)Q)(h)(X_s)d\tau \\
&= (P(t-s)h)(X_s) - \int_0^{t-s} (P(\tau)Q)(h)(X_s)d\tau - h(X_0) - \int_0^s Qh(X_\tau)d\tau \\
&= \left( P(t-s) - \int_0^{t-s} P(\tau)Qd\tau \right) (h)(X_s) - h(X_0) - \int_0^s Qh(X_\tau)d\tau \\
&\stackrel{(7)}{=} h(X_s) - h(X_0) - \int_0^s Qh(X_\tau)d\tau.
\end{aligned}$$

**Comentarios:**

- Además, esta familia de martingalas caracteriza el proceso  $(X_t)_t$  como un proceso markoviano asociado al generador infinitesimal  $Q$ .
- Existe una versión un poco más general: si  $h : \mathbb{R}^+ \times \mathbb{N} \rightarrow \mathbb{R}$  es acotada y diferenciable en  $t$ ,  $\left( h(t, X_t) - h(0, X_0) - \int_0^t \left( \frac{\partial h}{\partial t}(\tau, X_\tau) + Qh(\tau, X_\tau) \right) d\tau \right)_{t>0}$  es una martingala.

Así se puede mostrar que  $\frac{X_t}{e^{\rho t}}$  es una martingala: se toma  $h(t, i) := e^{-\rho t}i$ . De un lado  $\frac{\partial h}{\partial t} = -\rho h$ . Pero, de otro lado,

$$Qh(t, i) = e^{-\rho t} \left( i\alpha \sum_{j \geq i-1, j \neq i} r_{j-i+1} j - i^2 \alpha \right) = e^{-\rho t} i\alpha \left( \sum_{k \geq 0} r_k (k+i-1) - i \right) = e^{-\rho t} i\alpha (m-1) = \rho h(i).$$

Entonces  $\frac{\partial h}{\partial t} + Qh \equiv 0$ .

## 2.5. Tiempo de extinción y probabilidad de extinción

Asumamos en esta sección que  $X_0 = 1$ . la probabilidad  $e(t)$  que la población sea ya extinguida en el tiempo  $t$  es:

$$e(t) := \mathbb{P}(X_t = 0) = F(0, t).$$

### Proposición 18 : Relación implícita de la probabilidad de extinción

Sea  $e(t), t > 0$  la función de distribución del tiempo aleatorio de extinción  $T_0$ ,  $e(t) := \mathbb{P}(T_0 \leq t)$ . En otras palabras,  $e(t)$  es la probabilidad que la población sea ya extinguida en el tiempo  $t$ . La función  $e(\cdot)$  verifica la relación implícita:

$$\forall t > 0, \quad \int_0^{e(t)} \frac{1}{\tilde{f}(s)} ds = t.$$

*Demostración.* Dada la definición de la función generatriz  $F$

$$e(t) = \mathbb{P}(X_t = 0) = F(0, t), \quad e(0) = 0 \text{ y } \lim_{t \rightarrow +\infty} e(t) = s_0.$$



Ya que  $F$  verifica la ecuación (10),  $\frac{\partial}{\partial t}e(t) = \tilde{f}(e(t))$ ,  $e(0) = 0$ . Sea  $\Phi(t)$  una acumulada de la función  $\frac{1}{\tilde{f}}$ :  $\Phi(t) = \int_0^t \frac{1}{\tilde{f}(s)} ds$ . Es bien definida por  $t < s_0$ , porque  $f(s) > s \Rightarrow \tilde{f}(s) \neq 0$ .

$$\frac{\partial}{\partial t}(\Phi(e(t)) - t) = \Phi'(e(t))\frac{\partial}{\partial t}e(t) - 1 = \frac{1}{\tilde{f}(e(t))}\frac{\partial}{\partial t}e(t) - 1 = 1 - 1 = 0.$$

Entonces la función  $t \mapsto \Phi(e(t)) - t$  es constante y  $\Phi(e(t)) - t \equiv \Phi(e(0)) - 0 = 0$ .  $\diamond$

### Ejemplos:

- Reproducción binaria crítica ( $r_0 = r_2 = \frac{1}{2}$ ,  $\rho = 0$ ):

$$F(s, t) = 1 - \frac{1-s}{1 + \frac{\alpha}{2}t(1-s)} \Rightarrow e(t) = 1 - \frac{1}{1 + \frac{\alpha}{2}t} \stackrel{t \text{ grande}}{\approx} 1 - \frac{2}{\alpha t}.$$

En particular, análogamente a lo establecido en la sección 1.7.3, el valor medio del tiempo de extinción es infinito:

$$\mathbb{E}(T_0) = \int_0^{+\infty} \mathbb{P}(T_0 > t) dt = \int_0^{+\infty} (1 - e(t)) dt \approx \int_0^{+\infty} \frac{2}{\alpha t} dt = +\infty.$$

- Reproducción binaria no crítica:

$$F(s, t) = 1 - \frac{1-s}{e^{-\rho t} + \frac{m}{2(m-1)}(1 - e^{-\rho t})(1-s)} \Rightarrow e(t) = 1 - \frac{2(m-1)}{(m-2)e^{-\rho t} + m}$$

*Caso subcrítico*,  $\rho < 0$ :  $\lim_{t \rightarrow +\infty} e(t) = 1$ , es decir que la población se extingue casi seguramente. Ahora, el valor medio del tiempo de extinción es finito:

$$\mathbb{E}(T_0) = \int_0^{+\infty} (1 - e(t)) dt \approx \int_0^{+\infty} \frac{2(m-1)}{m-2} e^{\rho t} dt < +\infty.$$

*Caso supercrítico*,  $\rho > 0$ :

$$\lim_{t \rightarrow +\infty} e(t) = \frac{2}{m} - 1 = q:$$

la población sobrevive con una probabilidad igual a  $\frac{2(m-1)}{m}$ .

- Reproducción crítica de tipo  $\gamma$ -estable:

$$F(s, t) = 1 - \frac{1-s}{(1 + c(\gamma-1)t(1-s)^{\gamma-1})^{1/\gamma-1}} \Rightarrow e(t) = 1 - \frac{1}{(1 + c(\gamma-1)t)^{1/\gamma-1}}$$

y

$$\mathbb{E}(T_0) \approx \left(\frac{1}{c(\gamma-1)}\right)^{1/\gamma-1} \int_0^{+\infty} \frac{1}{t^{1/(\gamma-1)}} dt < +\infty$$

dado que  $1/(\gamma-1) > 1$ .

Entonces este proceso se extingue mas rápidamente que el binario crítico.

## 2.6. El proceso BGWc condicionado

Nos interesamos a la situación del BGWc (sub)critico antes de su (c.s.) extinción. Como en el caso discreto, se puede condicionar la dinámica a la no extinción en un futuro lejano. Así, para cada tiempo  $t > 0$  y cada suceso  $B_{[0,t]} \in \mathcal{F}_t$  que depende de los valores del proceso hasta el tiempo  $t$ , mostramos que

$$\mathbb{P}(B_{[0,t]} | X_0 = i, X_{t+\theta} > 0)$$

converge cuando  $\theta$  es grande.

**Teorema 19 : Existencia del proceso condicionado a la no extinción en un tiempo futuro cuasi-infinito.**

Sea  $X = (X_t)_t$  un proceso de BGWc subcrítico y no degenerado, de ley de reproducción  $\mathbf{r}$  con momento segundo finito. Existe un proceso de Markov  $X^* = (X_t^*)_t$  definido por

$$\mathbb{P}(X_t^* = \cdot | X_0^* = i) = \lim_{\theta \rightarrow +\infty} \mathbb{P}(X_t = \cdot | X_0 = i, X_{t+\theta} > 0).$$

Ademas, su ley tiene una densidad sobre  $\mathcal{F}_t$  con respecto a la ley de  $X$ :  
para cada tiempos  $t_1 < \dots < t_k < t$ :

$$\mathbb{P}(X_{t_1}^* = i_1, \dots, X_{t_k}^* = i_k | X_0^* = i) = \mathbb{E}(e^{-\rho t} \frac{X_t}{i} \mathbb{1}_{\{X_{t_1}=i_1, \dots, X_{t_k}=i_k\}} | X_0 = i) \quad (11)$$

*Demostración:*

$$\begin{aligned} \mathbb{P}(X_t = j | X_0 = i, X_{t+\theta} > 0) &= \frac{\mathbb{P}(X_t = j, X_{t+\theta} > 0 | X_0 = i)}{\mathbb{P}(X_{t+\theta} > 0 | X_0 = i)} \\ &= \frac{\mathbb{P}(X_{t+\theta} > 0 | X_t = j) \mathbb{P}(X_t = j | X_0 = i)}{1 - \mathbb{P}(X_{t+\theta} = 0 | X_0 = i)} \\ &= \frac{(1 - \mathbb{P}(X_\theta > 0))^j \mathbb{P}(X_t = j | X_0 = i)}{1 - F(0, t + \theta)^i} \\ &= \frac{1 - F(0, \theta)^j}{1 - F(0, t + \theta)^i} \mathbb{P}(X_t = j | X_0 = i) \end{aligned}$$

De manera semejante al caso tiempo discreto, se busca una estimación de  $F(0, t)$  para  $t$  grande.

$$\text{Si } \sum_j j^2 r_j < +\infty, \quad \exists c > 0, \quad \mathbb{P}(X_t = 0) \approx_t 1 - c \exp(\rho t).$$

(Ver [13] p. 109.) Entonces,

$$\frac{1 - F(0, \theta)^j}{1 - F(0, t + \theta)^i} \approx_\theta \frac{1 - (1 - cj e^{\rho \theta})}{1 - (1 - cie^{\rho(t+\theta)})} = e^{-\rho t} \frac{j}{i}$$

y

$$\lim_{\theta \rightarrow +\infty} \mathbb{P}(X_t = j | X_0 = i, X_{t+\theta} > 0) = e^{-\rho t} \frac{j}{i} \mathbb{P}(X_t = j | X_0 = i).$$

Para verificar (11) es simple de mostrar que, para  $t_1 < \dots < t_k < t$ ,

$$\mathbb{P}(X_{t_1}^* = i_1, \dots, X_{t_k}^* = i_k | X_0^* = i) = \mathbb{E}\left(e^{-\rho t_k} \frac{X_{t_k}}{i} \mathbb{1}_{\{X_{t_1}=i_1\}} \cdots \mathbb{1}_{\{X_{t_k}=i_k\}} | X_0 = i\right)$$

y de utilizar la propiedad de martingala de  $e^{-\rho t} X_t$   $\diamond$ .

### 3. Límites de procesos de BGW cuando la población es numerosa

Se muestran varias ventajas del estudio del sistema límite.

- Primero, a menudo la estructura del límite es más sencilla y se puede interpretar mejor el comportamiento cualitativo.
- Muchas veces el número de parámetros que determinan el proceso límite se reduce en relación al proceso original.

Cuando una población es muy numerosa, en lugar de considerar el número de individuos conviene estudiar la "densidad" de individuos. Este enfoque, que se usa en física estadística, se emplea para obtener los procesos con estado continuo.

#### 3.1. Límite determinista

Consideremos el proceso de Bienaymé-Galton-Watson, pero en lugar de uno solo individuo inicial, ahora habrá muchos individuos iniciales,  $X_0^{(N)}$ , con  $X_0^{(N)} \approx N x$ ,  $x \in \mathbb{R}^+$ . Antes supusimos implícitamente que cada individuo tiene una masa unitaria, y ahora los individuos tendrán masas muy pequeñas  $\frac{1}{N}$ .

##### 3.1.1. BGW de tiempo discreto renormalizado

Sea  $(\check{X}_n^{(N)})_n := (\frac{1}{N} X_n^{(N)})_n$  el nuevo proceso de densidad de población. El proceso  $\check{X}^{(N)}$  toma sus valores en  $\frac{1}{N}\mathbb{N} \subset \mathbb{R}^+$ .

Suponemos que los primero y segundo momentos de la ley de reproducción existen y estudiamos en primer lugar el comportamiento de los dos primeros momentos de  $\check{X}_n^{(N)}$ . Según §1.4

$$\begin{aligned} \mathbb{E}(\check{X}_n^{(N)}) &= \frac{1}{N} \mathbb{E}(X_n^{(N)}) = \frac{X_0^{(N)}}{N} m^n \xrightarrow{N \rightarrow +\infty} x m^n \\ \text{Var}(\check{X}_n^{(N)} | X_0^{(N)}) &= \frac{1}{N^2} \text{Var}(X_n^{(N)} | X_0^{(N)}) = \frac{X_0^{(N)}}{N^2} \text{Var}(X_n | X_0 = 1) \approx_N \frac{x}{N} \text{Var}(X_n | X_0 = 1) \xrightarrow{N \rightarrow +\infty} 0 \end{aligned}$$

Así, el límite de  $\check{X}_n^{(N)}$  para  $N$  grande parece ser determinista, entonces igual a su esperanza, y tiene un desarrollo en el tiempo dado por  $x m^n$ . No es muy interesante.

##### 3.1.2. BGWc renormalizado

Sea  $(\check{X}_t^{(N)})_{t \geq 0} := (\frac{1}{N} X_t^{(N)})_{t \geq 0} \in \frac{1}{N}\mathbb{N} \subset \mathbb{R}^+$  el nuevo proceso de densidad de población.

Se puede estudiar el comportamiento en  $N$  de los primeros momentos, como en el caso discreto.

Se hará en **ejercicio**.

Se puede también estudiar el comportamiento de la transformada de Laplace de la variable aleatoria renormalizada  $\check{X}_t^{(N)}$ , definida por

$$\begin{aligned} \check{L}_t^N(\lambda) &:= \mathbb{E}(\exp(-\lambda \check{X}_t^{(N)})) \\ &= \mathbb{E}((e^{-\lambda/N})^{X_t^{(N)}}) = \left( \mathbb{E}((e^{-\lambda/N})^{X_t} | X_0 = 1) \right)^{X_0^{(N)}} = (F(e^{-\lambda/N}, t))^{X_0^{(N)}} \end{aligned}$$

Porque

$$F(e^{-\lambda/N}, t) = \mathbb{E}\left((e^{-\lambda/N})^{X_t}\right) \approx_N \mathbb{E}\left(\left(1 - \frac{\lambda}{N}\right)^{X_t}\right) \approx_N \mathbb{E}\left(1 - \frac{\lambda X_t}{N}\right) \approx_N 1 - \frac{\lambda \mathbb{E}(X_t)}{N} = 1 - \frac{\lambda e^{\rho t}}{N},$$

$$\begin{aligned} \text{si } X_0^{(N)} \approx_N N x &\Rightarrow \check{L}_t^N(\lambda) \approx_N \left(1 - \frac{\lambda e^{\rho t}}{N}\right)^{Nx} \\ &\Rightarrow \lim_{N \rightarrow +\infty} \check{L}_t^N(\lambda) = \exp(-\lambda x e^{\rho t}), \end{aligned}$$

que se interprete como la transformada de Laplace de la función determinista  $\check{X}_t^{(\infty)} \equiv x e^{\rho t}$ . Es solución de la ecuación diferencial de Malthus:

$$\frac{\partial}{\partial t} \check{X}_t^{(\infty)} = \rho \check{X}_t^{(\infty)}, \quad \check{X}_0^{(\infty)} = x. \quad (12)$$

Si el parámetro de Malthus  $\rho$  es positivo,  $\check{X}_t^{(\infty)}$  explota asintóticamente, si  $\rho = 0$   $\check{X}_t^{(\infty)}$  es constante, y si  $\rho < 0$   $\check{X}_t^{(\infty)}$  converge a 0 por  $t$  grande.

## 3.2. Limite difusión

Para evitar un limite determinista, se toma muchos individuos iniciales, cada uno de masa pequeña, pero se acelera además el tiempo de modo que en una unidad de tiempo ocurrirán muchas ramificaciones. También se hace que las leyes de ramificación dependen de  $N$  (tienen primero y segundo momento finidos, notados  $m^{(N)}$  y  $(\sigma^{(N)})^2$ ).

Ver [7] cuando la ley de ramificación no depende de  $N$ , [12] Capitulo 4 para explicaciones heurísticas y [17] para una descripción completa de las condiciones suficientes y necesarias para la convergencia de BGW renormalizados hacia procesos de ramificación muy generales con estado continuo.

### 3.2.1. BGW de tiempo discreto renormalizado

Para cada  $N$  definimos el proceso estocástico en tiempo continuo

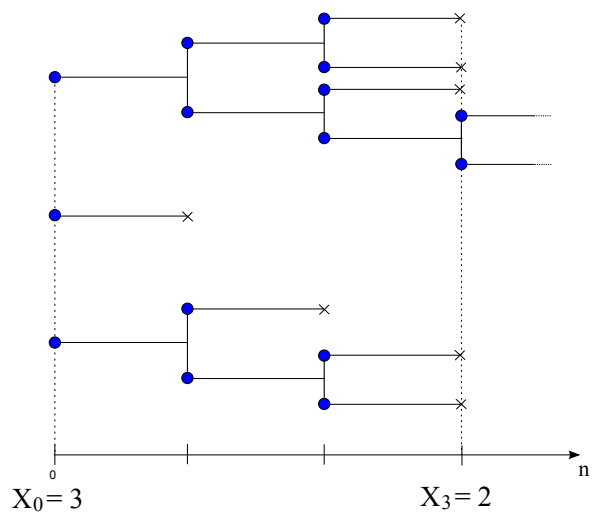
$$\tilde{X}_t^{(N)} := \frac{1}{N} X_{[Nt]}^{(N)} \quad \text{con } X_0^{(N)} \approx_N N x, \quad ([y] := \sup\{n \in \mathbb{N}, n \leq y\}).$$

Así, las ramificaciones ocurren en los instantes  $\frac{i}{N}, i = 1, 2, \dots$ , y cada individuo tiene masa  $1/N$ .

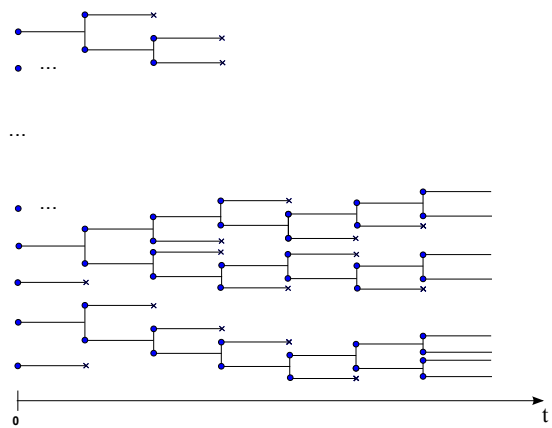
**Ejemplo: el caso  $f^{(N)}$  geométrica no crítico con parámetro  $a^{(N)}$  :**

Como antes, se estudia la convergencia de la transformada de Laplace del proceso renormalizado.

$$\begin{aligned} \tilde{L}_t^N(\lambda) &:= \mathbb{E}(\exp(-\lambda \tilde{X}_t^{(N)})) \\ &= \mathbb{E}((e^{-\lambda/N})^{X_{[Nt]}^{(N)}}) = \left( \mathbb{E}\left((e^{-\lambda/N})^{X_{[Nt]}^{(N)}} | X_0^{(N)} = 1\right) \right)^{X_0^{(N)}} \\ &= (f_{[Nt]}^{(N)}(e^{-\lambda/N}))^{X_0^{(N)}} \end{aligned}$$



(a) BGW binário



(b) BGW binário renormalizado

Figura 7: BGW renormalizado

Se sabe que, en el caso no crítico,

$$f_{[Nt]}^{(N)}(s) = \frac{(m^{(N)})^{[Nt]}(1-s) + m^{(N)}s - 1}{(m^{(N)})^{[Nt]+1}(1-s) + m^{(N)}s - 1} \text{ donde } m^{(N)} = (1 - a^{(N)})/a^{(N)}.$$

**Ejercicio :** Verifiquen que cuando  $m^{(N)} = 1 + \frac{b}{N}$  la función  $\tilde{L}_t^N(\lambda)$  converge. Identificar la forma de la transformada de Laplace límite.

Volvemos al caso general. Aquí también estudiamos en primer lugar el comportamiento del primero momento:

$$\mathbb{E}(\tilde{X}_t^{(N)}) = \frac{X_0^{(N)}}{N} (m^{(N)})^{[Nt]} \approx_{N \rightarrow +\infty} x (m^{(N)})^{[Nt]}.$$

Una condición necesaria para la convergencia de  $(m^{(N)})^{[Nt]}$  es que  $\mathbf{r}^{(N)}$  sea asintóticamente crítica:  $m^{(N)} \stackrel{N \rightarrow +\infty}{\approx} 1 + \frac{b}{N}$ ,  $b \in \mathbb{R}$ . En este caso  $\lim_{N \rightarrow +\infty} \mathbb{E}(\tilde{X}_t^{(N)}) = xe^{bt}$ .

Analizando el segundo momento:

$$\text{Var}(\tilde{X}_t^{(N)}) = \frac{1}{N^2} \text{Var}(X_{[Nt]}^{(N)}) = \frac{X_0^{(N)}}{N^2} \text{Var}(X_{[Nt]}^{(N)} | X_0^{(N)} = 1) \approx_N \frac{x}{N} \text{Var}(X_{[Nt]}^{(N)} | X_0^{(N)} = 1)$$

Peró, si  $b \neq 0$ ,

$$\begin{aligned} \text{Var}(X_{[Nt]}^{(N)} | X_0^{(N)} = 1) &= (\sigma^{(N)})^2 (m^{(N)})^{[Nt]-1} \frac{(m^{(N)})^{[Nt]} - 1}{m^{(N)} - 1} \\ &\approx_N (\sigma^{(N)})^2 \left(1 + \frac{b}{N}\right)^{[Nt]} \frac{\left(1 + \frac{b}{N}\right)^{[Nt]} - 1}{\left(1 + \frac{b}{N}\right) - 1} \\ &\approx_N N (\sigma^{(N)})^2 e^{bt} \frac{e^{bt} - 1}{b} \end{aligned}$$

Así, si  $\lim_N \sigma^{(N)} =: \sigma$ ,  $\lim_{N \rightarrow +\infty} \text{Var}(\tilde{X}_t^{(N)}) = x\sigma^2 e^{bt} \frac{e^{bt} - 1}{b}$ .

De otro lado, si  $b = 0$  ( $\Leftrightarrow \lim_N N(m^{(N)} - 1) = 0$ ),

$$\text{Var}(\tilde{X}_t^{(N)}) \approx_N \frac{1}{N} x (\sigma^{(N)})^2 [Nt] \rightarrow_N x \sigma^2 t.$$

### **Teorema 20 Construcción del proceso límite via su transformada de Laplace**

Si las tres condiciones  $m^{(N)} \approx_N 1 + \frac{b}{N}$ ,  $\lim_N \sigma^{(N)} = \sigma$  y  $\lim_N \sum_{j \geq tN} j^2 r_j^{(N)} = 0$  son realizadas, y si  $\lim_N \tilde{X}_0^{(N)} = x$ , entonces por cada  $t > 0$ ,  $\tilde{X}_t^{(N)}$  converge hacia un proceso  $\tilde{X}_t^\infty$  definido por su transformada de Laplace

$$\tilde{L}_t(\lambda) := \mathbb{E}(\exp(-\lambda \tilde{X}_t^\infty) | \tilde{X}_0^\infty = x) = \exp(-xu_t(\lambda))$$

donde la función  $u_t$  verifica (15).

Esbozo de la demostración: ([7], Appendix II o [11] Theorem 3.2)

$$\begin{aligned}\tilde{L}_t^N(\lambda) &:= \mathbb{E}(\exp(-\lambda\tilde{X}_t^{(N)})) \\ &= \mathbb{E}\left((e^{-\lambda/N})^{X_{[Nt]}^{(N)}}\right) = \left(\mathbb{E}\left((e^{-\lambda/N})^{X_{[Nt]}^{(N)}}|X_0^{(N)} = 1\right)\right)^{X_0^{(N)}} \\ &= \left(f_{[Nt]}(e^{-\lambda/N})\right)^{X_0^{(N)}}\end{aligned}$$

$$\begin{aligned}\text{Como } f(e^{-\lambda/N}) &\approx_N 1 + (e^{-\lambda/N} - 1)\left(1 + \frac{b}{N}\right) + \frac{1}{2}(e^{-\lambda/N} - 1)^2((\sigma^{(N)})^2 - m^{(N)} + (m^{(N)})^2) + O(1/N^3) \\ &\approx_N 1 - \frac{\lambda}{N} - (b\lambda - \frac{1}{2}(\sigma^2 + 1)\lambda^2)\frac{1}{N^2} + O(1/N^3)\end{aligned}$$

$$\Rightarrow \log f(e^{-\lambda/N}) \approx_N -\frac{\lambda}{N} - \frac{b\lambda - \frac{\sigma^2}{2}\lambda^2}{N^2} + O(1/N^2);$$

$$\begin{aligned}\text{luego } \tilde{L}_{t+1/N}^N(\lambda) &\approx_N (f_{[Nt]}(f(e^{-\lambda/N})))^{N^x} \approx \tilde{L}_t^N(-N \log f(e^{-\lambda/N})) \\ &\approx \tilde{L}_t^N\left(\lambda + \frac{b\lambda - \frac{\sigma^2}{2}\lambda^2}{N} + O(1/N^2)\right).\end{aligned}$$

Si  $\lim_N \tilde{L}_t^N(\lambda) =: \tilde{L}_t(\lambda)$  ella verifica

$$\frac{\partial \tilde{L}_t}{\partial t}(\lambda) = (b\lambda - \frac{\sigma^2}{2}\lambda^2) \frac{\partial \tilde{L}_t}{\partial \lambda}(\lambda). \quad (13)$$

Notando  $\tilde{L}_t(\lambda) =: \exp(-xu_t(\lambda))$ , la función  $u_t$  verifica tambien

$$\frac{\partial}{\partial t}u_t = (b\lambda - \frac{\sigma^2}{2}\lambda^2) \frac{\partial}{\partial \lambda}u_t, \quad u_0(\lambda) = \lambda. \quad (14)$$

La unica solución de este ecuación es

$$u_t(\lambda) = \begin{cases} \frac{\lambda e^{bt}}{1 + \lambda\sigma^2(e^{bt} - 1)/2b} & \text{si } b \neq 0, \\ \frac{\lambda}{1 + \lambda\sigma^2 t/2} & \text{si } b = 0 \text{ (caso critico)}. \end{cases} \quad \diamond \quad (15)$$

La función  $u_t$  se llama a veces cumulante y es también solución de la ecuación diferencial no lineal:

$$\frac{\partial}{\partial t}u_t = bu_t - \frac{\sigma^2}{2}u_t^2.$$

Utilizando la propiedad de Markov, se puede tambien mostrar que las distribuciones finito-dimensionales  $(\tilde{X}_{t_1}^{(N)}, \dots, \tilde{X}_{t_k}^{(N)})$  convergen.

En general, la convergencia débil de las distribuciones finito-dimensionales no es una condición suficiente para la convergencia de los procesos, considerados como funciones aleatorias a valores en  $\mathbf{D}([0, T]; \mathbb{R})$ , el espacio de todos los mapeos de  $[0, T]$  en  $\mathbb{R}$ , continuos por la derecha, con límites por la izquierda. En  $\mathbf{D}([0, T]; \mathbb{R})$ , se define la métrica

$$d(f, g) := \inf_{\theta \in \Theta} \left( \sup_{t \in [0, T]} |f(t) - g(\theta(t))| + \sup_{s, t \in [0, T], s \neq t} \left| \log \frac{\theta(t) - \theta(s)}{t - s} \right| \right)$$

donde  $\Theta$  es la clase de todas las funciones continuas, estrictamente crecientes, del intervalo  $[0, T]$  sobre sí mismo. La métrica  $d$  induce una topología llamada de Skorohod. El espacio  $\mathbf{D}([0, T]; \mathbb{R})$  con esta topología es un espacio métrico, separable y completo. Ver [2].

Una condición adicional suficiente para la convergencia de la sucesión de los procesos es su compacidad relativa. El milagro con los procesos de BGW renormalizados consiste en que su convergencia en  $\mathbf{D}([0, T]; \mathbb{R})$  ya es una consecuencia de la convergencia de sus distribuciones uni-dimensionales. Ver [11] Theorem 3.4.

### 3.2.2. BGWc renormalizado

Para cada  $N$  definimos el proceso estocástico en tiempo continuo

$$\hat{X}_t^{(N)} := \frac{1}{N} X_t^{(N)}$$

donde  $X_t^{(N)}$  es un BGWc con parametros  $\alpha^{(N)}$  y  $\mathbf{r}^{(N)}$ . Cada individuo tiene masa  $1/N$ . Acortamos la duración de vida aleatoria de cada individuo en función de  $N$  para crecer el numero medio de ramificaciones en una unidad de tiempo (es decir:  $\lim_N \alpha^{(N)} = +\infty$ ).

Como antes, se controla la convergencia de  $\hat{X}_t^{(N)}$  utilizando su transformada de Laplace.

$$\begin{aligned} \hat{L}_t^N(\lambda) &:= \mathbb{E}(\exp(-\lambda \hat{X}_t^{(N)})) \\ &= \mathbb{E}((e^{-\lambda/N})^{X_t^{(N)}}) = \left( \mathbb{E}((e^{-\lambda/N})^{X_t^{(N)}} | X_0^{(N)} = 1) \right)^{X_0^{(N)}} \\ &= (F^{(N)}(e^{-\lambda/N}, t))^{X_0^{(N)}} \approx_N (F^{(N)}(e^{-\lambda/N}, t))^{Nx} \end{aligned}$$

En primar lugar analisamos un

**Ejemplo : reproducción binaria no crítica.** Ver figura (8).

Se calcula

$$F^{(N)}(s, t) = 1 - \frac{1 - s}{e^{-\alpha^{(N)}(m^{(N)}-1)t} + \frac{m^{(N)}}{2(m^{(N)}-1)}(1 - e^{-\alpha^{(N)}(m^{(N)}-1)t})(1 - s)}$$

para  $s = e^{-\lambda/N}$ .

**Ejercicio:** Mostrar que, cuando  $m^{(N)} = 1 + \frac{b}{N}$  ( $\Rightarrow \mathbf{r}^{(N)}$  es asintoticamente crítica:  $r_0^{(N)} = \frac{1}{2} - \frac{b}{N}$ ,  $r_2^{(N)} = \frac{1}{2} + \frac{b}{N}$ ,  $\sigma^{(N)} \rightarrow 1$ ) y cuando  $\alpha^{(N)} = N\alpha$ , entonces

$$\lim_N \hat{L}_t^N(\lambda) = \begin{cases} \exp\left(-x \frac{\lambda e^{\alpha b t}}{1 + \lambda(e^{\alpha b t} - 1)/2b}\right) & \text{si } b \neq 0, \\ \exp\left(-x \frac{\lambda}{1 + \lambda \alpha t/2}\right) & \text{si } b = 0 \text{ (caso critico)}. \end{cases}$$

De hecho, las condiciones  $\rho^{(N)} = \alpha^{(N)}(m^{(N)} - 1) \xrightarrow{N} \rho (= \alpha b)$  y  $\sigma^{(N)} \xrightarrow{N} \sigma$  son cuasi necesarias y suficientes para la convergencia de los procesos  $\hat{X}^{(N)}$  en general.



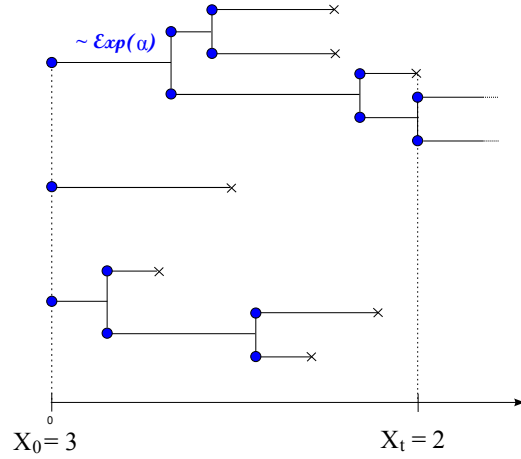


Figura 8: BGWc binario

**Teorema 21 Construcción de la difusión de Feller como proceso limite**

Si las tres condiciones  $m^{(N)} \approx_N 1 + \frac{b}{N}$ ,  $\lim_N \sigma^{(N)} = \sigma$  y  $\sup_N \sum_{j \geq 0} j^3 r_j^{(N)} < +\infty$  son realizadas, si

$\alpha^{(N)} = N\alpha$  y  $\lim_N \hat{X}_0^{(N)} = x$ , entonces los procesos  $\hat{X}^{(N)}$  convergen hacia un proceso markoviano  $Y$  de generador infinitesimal  $A$  donde

$$\forall h \in C_b^2(\mathbb{R}^+), \quad Ah(x) = \alpha b x h'(x) + \alpha \frac{\sigma^2}{2} x h''(x).$$

Llamaremos al proceso  $Y$  : difusión de Feller, con parámetros  $(\alpha b, \alpha \sigma^2)$ .

Esbozo de la demostración: (Ver Theorem 1.3 in [5])

Se hace en tres pasos:

- (i) Se muestra la convergencia de los generadores infinitesimales  $\hat{Q}^{(N)}$  asociados a  $\hat{X}^{(N)}$  hacia  $A$
- (ii) Se muestra que las leyes de los procesos  $\hat{X}^{(N)}$  son relativamente compactas.
- (iii) Se verifica que el problema de martingala asociado a  $A$  tiene una única solución.

Sobre (i)

Sea  $h \in \mathcal{C}_b^2(\mathbb{R}^+)$  y  $x = \frac{i}{N} \in \frac{1}{N}\mathbb{N}$ .

$$\begin{aligned}
\hat{Q}^{(N)}h(x) &= \frac{\partial}{\partial t}\Big|_{t=0} \left( \hat{P}^{(N)}(t)h \right) \left( \frac{i}{N} \right) = \frac{\partial}{\partial t}\Big|_{t=0} \left( P^{(N)}(t)h \left( \frac{\cdot}{N} \right) \right) (i) \\
&= \alpha^{(N)}i \sum_{k \geq 0} r_k^{(N)} \left( h \left( \frac{i+k-1}{N} \right) - h \left( \frac{i}{N} \right) \right) \\
&= N^2 \alpha x \sum_{k \geq 0} r_k^{(N)} \left( h \left( x + \frac{k-1}{N} \right) - h(x) \right) \\
&= N^2 \alpha x \sum_{k \geq 0} r_k^{(N)} \left( \frac{k-1}{N} h'(x) + \frac{1}{2} \left( \frac{k-1}{N} \right)^2 h''(x) + O \left( \frac{k^3}{N^3} \right) \right) \\
&= \alpha x N (m^{(N)} - 1) h'(x) + \frac{1}{2} \alpha x \sum_{k \geq 0} r_k^{(N)} (k - m^{(N)} + m^{(N)} - 1)^2 h''(x) + O \left( \frac{1}{N} \right) \\
&\xrightarrow{N} \alpha b x h'(x) + \frac{\sigma^2}{2} \alpha x h''(x) = Ah(x).
\end{aligned}$$

*Sobre (ii)*

Las leyes de los procesos  $\hat{X}^{(N)}$  son relativamente compactas como probabilidades sobre  $\mathbf{D}([0, T]; \mathbb{R})$  si (Ver en [5], el criterio de Aldous)

- $\forall t \leq T, \forall \varepsilon > 0, \exists C_\varepsilon \subset \mathbb{R}$  compacto :  $\inf_N \mathbb{P}(\hat{X}_t^{(N)} \in C_\varepsilon) \geq 1 - \varepsilon$
- Para cada familia  $(\tau_N)_N$  de tiempos de paro limitados pour  $T, \forall \varepsilon > 0, \exists \delta > 0$  y  $N_0$  asi que

$$\sup_{N \geq N_0} \sup_{\theta \leq \delta} \mathbb{P}(|\hat{X}_{\tau_N + \theta}^{(N)} - \hat{X}_{\tau_N}^{(N)}| > \varepsilon) \leq \varepsilon.$$

Para mostrar esta ultima desigualdad, se verifica para la variación cuadrática de la parte martingala de  $\hat{X}^{(N)}$  y para su parte a variación limitada.

*Sobre (iii)* (Ver [24] §5.6.2)

Primero, verificamos que cada solución del problema de martingalas asociado con  $A$  es también solución debil de la ecuación diferencial estocástica

$$Y_t = x + \int_0^t \sigma \sqrt{\alpha Y_s} dB_s + \int_0^t \alpha b Y_s ds, \quad (16)$$

donde  $B$  es un movimiento Browniano.

Aplicando el problema de martingala a la función  $h = Id$ , se consigue:

$$\bar{Y}_t := Y_t - \int_0^t A(Id)(Y_s) ds = Y_t - \int_0^t \alpha b Y_s ds \quad \text{es una martingala locale.}$$

Como es ademas integrable, es una martingala.

Se quiere probar que la variación cuadrática de la martingala  $\bar{Y}_t$  es igual a  $\int_0^t \alpha \sigma^2 Y_s ds$ , para poder representarla como  $\int_0^t \sigma \sqrt{\alpha Y_s} dB_s$ . Probamos de manera equivalente que

$$\bar{Y}_t^2 - \int_0^t \alpha \sigma^2 Y_s ds = Y_t^2 - 2Y_t \int_0^t \alpha b Y_s ds + \left( \int_0^t \alpha b Y_s ds \right)^2 - \int_0^t \alpha \sigma^2 Y_s ds \quad \text{es una martingala;}$$

Se sabe ya, aplicando el problema de martingala a la función  $h(x) = x^2$ , que

$$Y_t^2 - \int_0^t A(h)(Y_s) ds = Y_t^2 - \int_0^t (2\alpha b Y_s^2 + \alpha \sigma^2 Y_s) ds \quad \text{es una martingala.}$$

Entonces, es suficiente probar que la diferencia

$$2Y_t \int_0^t \alpha b Y_s ds - \left( \int_0^t \alpha b Y_s ds \right)^2 - 2 \int_0^t \alpha b Y_s^2 ds \quad \text{es una martingala.}$$

Se consigue, utilizando la formula de integración por partes:

$$\begin{aligned} Y_t \int_0^t Y_s ds &= \int_0^t Y_s Y_s ds + \int_0^t \int_0^s Y_r dr dY_s = \int_0^t Y_s^2 ds + \int_0^t \int_0^s Y_r (\alpha b Y_s) dr ds + \text{mart.} \\ &= \int_0^t Y_s^2 ds + \frac{1}{2} \int_0^t \int_0^t \alpha b Y_r Y_s dr ds + \text{mart.} = \int_0^t Y_s^2 ds + \frac{1}{2} \alpha b \left( \int_0^t Y_s ds \right)^2 + \text{mart.} \end{aligned}$$

Ahora, como el coeficiente de difusión  $x \mapsto \sigma \sqrt{\alpha x}$  de la ecuación (16) es una función Hölder-continua, los resultados clásicos permiten concluir que la solución es única.  $\diamond$

### 3.2.3. Unas propiedades de la difusión de Feller

#### Proposición 22 Universalidad de la difusión de Feller

El proceso  $\tilde{X}^\infty$  construido como límite de BGW renormalizados es también una difusión de Feller  $Y$  con parámetros  $(b, \sigma^2)$ .

*Demostración*

Verificamos que, por  $\alpha = 1$ , la transformada de Laplace de  $Y_t$  es también solución de la ecuación (13).

Aplicando el problema de martingala a la función  $h(x) = e^{-\lambda x}$  se consigue

$$e^{-\lambda Y_t} - \int_0^t A(h)(Y_s) ds = e^{-\lambda Y_t} - \int_0^t \left( \frac{\sigma^2}{2} \lambda^2 - b\lambda \right) Y_s e^{-\lambda Y_s} ds \quad \text{es una martingala.}$$

Entonces

$$\begin{aligned} \mathbb{E}(e^{-\lambda Y_t}) &= e^{-\lambda x} + \int_0^t \left( \frac{\sigma^2}{2} \lambda^2 - b\lambda \right) \mathbb{E}(Y_s e^{-\lambda Y_s}) ds \\ \Rightarrow \frac{\partial}{\partial t} \mathbb{E}(e^{-\lambda Y_t}) &= (b\lambda - \frac{\sigma^2}{2} \lambda^2) \frac{\partial}{\partial \lambda} \mathbb{E}(e^{-\lambda Y_t}) \end{aligned}$$

que es exactamente la ecuación diferencial (13). Como la transformada de Laplace caracteriza la ley de una variable aleatoria, la demostración es finita.  $\diamond$

- Así, los procesos BGW y BGWc renormalizados convergen hacia el mismo proceso límite (si se toma  $\alpha = 1$ ).
- La difusión de Feller verifica la propiedad de ramificación: el proceso  $Y$  iniciado en  $Y_0 = x_1 + x_2$  es la suma de dos copias independientes de  $Y$  iniciadas en  $x_1$  y en  $x_2$ , ya que la transformada de Laplace es multiplicativa con respecto a  $Y_0$ .

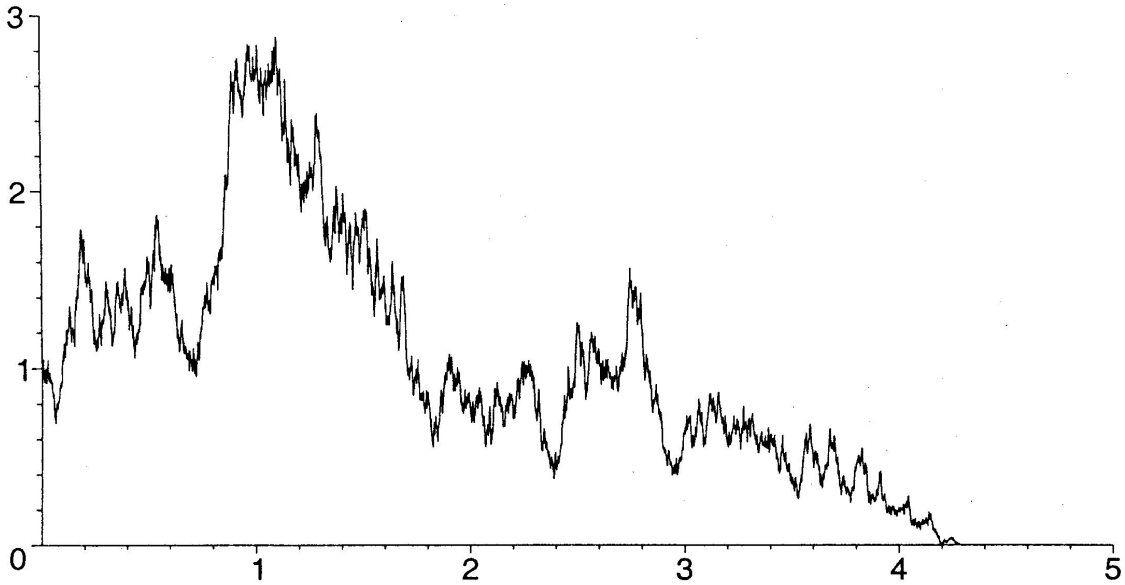


Figura 9: Simulación de  $t \mapsto Y_t$  con el parámetro  $b = 0$ .

- ( $\alpha = 1$ ) El parámetro  $b$ , coeficiente de la deriva del proceso, se interprete como un coeficiente de velocidad condicionada:

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E} \left( \frac{Y_{t+\varepsilon} - Y_t}{\varepsilon} | Y_t \right) = b Y_t$$

- ( $\alpha = 1$ ) El segundo parámetro  $\sigma^2$  aparece como el coeficiente de la fluctuación cuadrática condicionada del proceso:

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \text{Var} \left( Y_{t+\varepsilon} - Y_t | Y_t \right) = \sigma^2 Y_t.$$

Así se presenta que las trayectorias del proceso  $Y$  ya que tienen una variación cuadrática finita, no tienen variación finita. (ver figura 9).

- Cuando  $b = 0$  (resp.  $b \leq 0$ ) el proceso se llama crítico (resp. subcrítico). En estos casos, c.s. el proceso toma el valor 0 y después, se queda en 0. Es la extinción del proceso.
- En el caso subcrítico, se puede también condicionar la difusión  $Y$  a la no extinción en un futuro lejano (Ver [16]). Se obtiene una otra difusión  $Y^*$ , de generador infinitesimal

$$A^*h(x) = Ah(x) + \sigma^2 h'(x) = (bx + \sigma^2)h'(x) + \frac{\sigma^2}{2} xh''(x).$$

$Y^*$  es una solución débil de la ecuación diferencial estocástica

$$Y_t = x + \int_0^t \sigma \sqrt{\alpha Y_s} dB_s + \int_0^t (\alpha b Y_s + \sigma^2) ds. \quad (17)$$

**Box 5.2 Deduction of the survival chance approximation**

Turning now to the proofs, we first show that the sequence in Equation (5.28) does not increase. Indeed, the process survives to time  $n$  if the offspring of at least one child of the initial individual survives to time  $n$ , that is

$$P_1(Z_n > 0) = P_1\left(\bigcup_{i=1}^{Z_1} \{Z_{n-1}^{(i)} > 0\}\right). \quad (a)$$

where  $Z_{n-1}^{(i)}$  is the number of individuals at time  $n$  that stem from the  $i$ th child of the ancestor of the population. However,

$$\begin{aligned} P_1\left(\bigcup_{i=1}^{Z_1} \{Z_{n-1}^{(i)} > 0\}\right) &= E_1\left[P\left(\bigcup_{i=1}^{Z_1} \{Z_{n-1}^{(i)} > 0\} \mid Z_1\right)\right] \\ &\leq E_1\left[\sum_{i=1}^{Z_1} P_1(Z_{n-1}^{(i)} > 0 \mid Z_1)\right]. \end{aligned} \quad (b)$$

and by the Wald identity (see the Appendix) we have

$$E_1\left[\sum_{i=1}^{Z_1} P_1(Z_{n-1}^{(i)} > 0 \mid Z_1)\right] = E_1[Z_1] E_1[P_1(Z_{n-1} > 0)] = m P_1(Z_{n-1} > 0) \quad (c)$$

since all the children behave independently and the survival probability up to time  $n$  of a daughter population equals the survival probability up to time  $n - 1$  of the original population.

Combining Equation (b) and Equation (c), we obtain

$$c_1 \leq \frac{P_1(Z_n > 0)}{m^n} \leq \frac{P_1(Z_{n-1} > 0)}{m^{n-1}}. \quad (d)$$

In other words, the expected population size, given non-extinction,

$$E_1[Z_n \mid Z_n > 0] = m^n / P_1(Z_n > 0) \quad (e)$$

increases toward  $1/c_1$  as generations pass.

If  $Z_0 = N > 1$ , the arguments above fail. However, if  $Q_n = P_1(Z_n = 0)$  is the probability of extinction in the first  $n$  generations, with one single ancestor, then

$$P_N(Z_n > 0) = 1 - Q_n^N \leq N(1 - Q_n) = N P_1(Z_n > 0), \quad (f)$$

which shows that  $c_N \leq c_1 N$ . In the geometric case equality holds here,  $c_N = (1 - m)N$ , as can be calculated from Equation (5.17).

More generally, in the biological case of finite reproduction variance  $\text{Var}_N(Z_n) < \infty$  also. A useful, though not so well known inequality, ascribed to Lyapunov (see the Appendix), shows that

$$P(X > 0) \geq \frac{(E[X])^2}{E[X^2]} \quad (g)$$

for any random variable  $X \geq 0$ . The choice of  $X = Z_n$  and  $Z_0 = 1$  gives

$$P_1(Z_n > 0) \geq \frac{(E_1[Z_n])^2}{E_1[Z_n^2]} = \frac{(E_1[Z_n])^2}{\text{Var}_1[Z_n] + (E_1[Z_n])^2} = \frac{m^{2n}}{\text{Var}_1[Z_n] + m^{2n}}. \quad (h)$$

From Chapter 2, Equation (2.16),

$$\text{Var}_1[Z_n] = \frac{\sigma^2 m^n (m^n - 1)}{m(m - 1)}. \quad (i)$$

Thus

$$\begin{aligned} P_1(Z_n > 0) &\geq \frac{m^{2n+1}(m - 1)}{\sigma^2 m^n (m^n - 1) + m^{2n+1}(m - 1)} \\ &= \frac{(1 - m)m^{n+1}}{\sigma^2(1 - m^n) + m^{n+1}(1 - m)}. \end{aligned} \quad (j)$$

Now, suppose that the population is initiated by  $N \geq 1$  individuals. In the Appendix we derive the inequality

$$1 - (1 - x)^k \geq kx - \frac{k(k - 1)}{2} x^2 \quad (k)$$

for non-negative integers  $k$  and  $0 \leq x \leq 1$ . This yields

$$\begin{aligned} P_N(Z_n > 0) &= 1 - (1 - P_1(Z_n > 0))^N \geq N P_1(Z_n > 0) \left(1 - \frac{N - 1}{2} P_1(Z_n > 0)\right) \\ &\geq \frac{N(1 - m)m^{n+1}}{\sigma^2(1 - m^n) + m^{n+1}(1 - m)} \left(1 - \frac{(N - 1)m^n}{2}\right) \\ &= \frac{N(1 - m)m}{\sigma^2(1 - m^n) + m^{n+1}(1 - m)} \left(1 - \frac{(N - 1)m^n}{2}\right) m^n, \end{aligned} \quad (l)$$

where we used the fact that  $P_1(Z_n > 0) \leq m^n$ . Thus, letting  $n \rightarrow \infty$ ,

$$c_N \geq N(1 - m)m/\sigma^2, \quad (m)$$

in the general situation.

**Theorem 5.4** Consider a Galton–Watson process with mean  $m < 1$  and starting from  $Z_0 = N$  individuals, and assume that the logarithmic moment condition of Theorem 5.3 holds. Then, for  $N \geq 3$ , the time  $\tau$  to extinction satisfies

$$\left( \frac{\ln N - \ln \ln N}{|\ln m|} - 1 \right) \left( 1 - \frac{1}{Nc_1} \right) \leq \mathbb{E}_N[\tau] \leq \frac{\ln N}{|\ln m|} + \frac{2 - m}{1 - m}, \quad (5.34)$$

where

$$c_1 \geq (1 - m)m / \sigma^2 \quad (5.35)$$

if  $\sigma^2 < \infty$  [see Equation (m) in Box 5.2].

**Proof.** From Equation (d) in Box 5.2,

$$c_1 m^n \leq \mathbb{P}_1(Z_n > 0) = \mathbb{P}_1(\tau > n) \leq m^n, \quad (5.36)$$

and

$$\mathbb{P}_N(\tau > n) \leq Nm^n. \quad (5.37)$$

Set

$$\phi(N) = \frac{\ln N}{|\ln m|}, \quad \psi(N) = \frac{\ln \ln N}{|\ln m|}. \quad (5.38)$$

Observe that  $Nm^{\phi(N)} = Nm^{-(\ln N)/\ln m} = 1$ ,

$$\exp\{-c_1 Nm^{\phi(N) - \psi(N)}\} = \exp\{-c_1 m^{-\psi(N)}\} = \exp\{-c_1 \ln N\} = 1/Nc_1. \quad (5.39)$$

Further (see the Appendix),

$$\mathbb{E}_N[\tau] = \sum_{n=0}^{\infty} \mathbb{P}_N(\tau > n), \quad (5.40)$$

and therefore, by Equation (5.37),

$$\begin{aligned} \mathbb{E}_N[\tau] &\leq \sum_{0 \leq n < \phi(N)} \mathbb{P}_N(\tau > n) + N \sum_{n \geq \phi(N)} m^n \\ &\leq \phi(N) + 1 + \frac{Nm^{\phi(N)}}{1 - m} = \frac{\ln N}{|\ln m|} + \frac{2 - m}{1 - m}. \end{aligned} \quad (5.41)$$

However,

$$\mathbb{E}_N[\tau] \geq \sum_{0 \leq n < \phi(N) - \psi(N)} \mathbb{P}_N(\tau > n) = \sum_{0 \leq n < \phi(N) - \psi(N)} (1 - \mathbb{P}_N(\tau \leq n)). \quad (5.42)$$

Clearly,

$$\begin{aligned} \mathbb{P}_N(\tau \leq n) &= \mathbb{P}_1^N(\tau \leq n) = (1 - \mathbb{P}_1(\tau > n))^N \\ &\leq e^{-N\mathbb{P}_1(\tau > n)} \leq e^{-c_1 Nm^n}, \end{aligned} \quad (5.43)$$

where we used first the inequality  $1 - x \leq e^{-x}$ ,  $x > 0$  and then Equation (5.36). Hence, for  $0 \leq n \leq \phi(N) - \psi(N)$ ,

$$\begin{aligned} \mathbb{P}_N(\tau > n) &\geq \mathbb{P}_N(\tau > \phi(N) - \psi(N)) \\ &= 1 - \mathbb{P}_N(\tau \leq \phi(N) - \psi(N)) \\ &\geq 1 - e^{-c_1 m^{\phi(N) - \psi(N)}} \geq 1 - \frac{1}{Nc_1}. \end{aligned} \quad (5.44)$$

Therefore,

$$\begin{aligned} \min_{0 \leq n \leq \phi(N) - \psi(N)} \mathbb{P}_N(\tau > n) &\geq (\phi(N) - \psi(N) - 1) \left( 1 - \frac{1}{Nc_1} \right) \\ &\geq \left( \frac{\ln N - \ln \ln N}{|\ln m|} - 1 \right) \left( 1 - \frac{1}{Nc_1} \right) \end{aligned} \quad (5.45)$$

Combining Equation (5.41) and Equation (5.45) gives Equation (5.34).

## Referencias

- [1] K.B. Athreya y P. Ney. *Branching processes*, Springer (1972).
- [2] P. Billingsley. *Convergence of probability measures*, second ed. Wiley, 1999
- [3] M.E. Caballero, V.M. Rivero, G. Uribe Bravo y C. Velarde. *Cadenas de Markov. Un enfoque elemental*, Aportaciones Matemáticas 29, Sociedad Matematica Mexicana (2004)
- [4] H. Caswell, M. Fujiwara y S. Brault. *Declining survival probability threatens the North Atlantic right whale* Proc. Nat. Acad. Sci. USA 96, Pop. Biology, 3308-3313 (1999)
- [5] S.N. Ethier y T.G. Kurtz. *Markov processes: characterization and convergence*, Wiley, 1986
- [6] L. Euler. *Recherches générales sur la mortalité et la multiplication du genre humain*. Mémoires de l'académie des sciences de Berlin 16, 1760
- [7] W. Feller. *Diffusion processes in genetics*, 2nd. Berkeley Symp. on Math. Stat. Prob., 227-246 (1951)
- [8] B. Fernandez. *Teoremas limites de alta densidad para campos aleatorios ramificados*, Aportaciones Matemáticas Serie Comunicaciones 2, Sociedad Matematica Mexicana (1986)
- [9] F. Galton y H.W. Watson. *On the probability of the extinction of families*, J. Roy. Anthropol. Inst. 4, 138-144 (1874)
- [10] L.G. Gorostiza. *Ramificación y superprocesos*, Notas, 2009  
<http://www.cimat.mx/Eventos/mesd09/gorostiza.pdf>
- [11] A. Grimvall *On the convergence of sequences of branching processes*, Annals Probab. 2-6, 1027-1045 (1974)
- [12] P. Haccou, P. Jagers y V.A. Vatutin. *Branching processes. Variation, growth and extinction of populations*, Cambridge Univ. Press (2005)
- [13] T.E. Harris. *The Theory of Branching Processes*, Springer-Verlag (1963).
- [14] C.C. Heyde y E. Seneta. *Statistical Theory Anticipated, I.J. Bienaymé*. Springer-Verlag (1977)
- [15] P. Jagers. *Branching processes with biological applications*, J. Wiley (1975)
- [16] A. Lambert. *Quasi-stationary distributions and the continuous-state branching process conditioned to be never extinct*, Elec. J. Prob. 12 420–446 (2007).
- [17] J. Lamperti. *The limit of a sequence of branching processes*, Zeit. für Wahrschein. Geb.7 (1967), 271-288.
- [18] A.J. Lotka. *The extinction of families* Journal of the Washington Academy of Sciences, Vol. 21 No 16,18 (1931)
- [19] R. Martinez Qin tana. Tesis doctoral *Proceso de ramificación de Galton-Watson multitypo*, Universidad de Extremadura, 2001

- [20] S. Méléard. *Quasi-stationary distributions for population processes*, VI Escuela de probabilidad y procesos estocásticos- Guanajuato(2009)
- [21] M. Mota Medina. Tesis doctoral *Aportaciones a la teoría de los Procesos de Ramificación Modificados*, Universidad de Extremadura, 1998
- [22] A.G. Pakes *Biological applications of branching processes*, in Stochastic Processes: Modelling and Simulation. Eds: Shanbhag, Rao. Handbook of Statistics Vol. 21. Elsevier (2003), 693-773
- [23] B.A. Sewastjanow *Verzweigungsprozesse*, Akademie Verlag Berlin, 1974
- [24] C. Tudor *Procesos estocásticos*, Aportaciones Matemáticas Serie Textos 2, Sociedad Matemática Mexicana, 3a edición (2002)
- [25] V.A. Vatutin. *Branching Processes and their Applications*, Notas, 2005, Cf. <http://ismi.math.uni-frankfurt.de/vatutin/BranchingProcIndex.shtml>
- [26] M.V. Vega. *Cadenas de Markov de tiempo continuo y aplicaciones*, Trabajo monográfico, Licenciatura en Matemática, Universidad de la Republica Montevideo (Uruguay) 2004
- [27] A. Wakolbinger. *Poblaciones aleatorias ramificadas y sus equilibrios*. Notas de investigacion 9, Sociedad Matematica Mexicana (1994)
- [28] D. Williams. *Probability with martingales*, Cambridge Univ. Press, 1991