



Centro de Investigación en Matemáticas, A.C.

CIMAT

**Estimación Unificada del
Proceso de Galton-Watson con
Inmigración**

T E S I S

que para obtener el grado de

Maestro en Ciencias con Especialidad en
Probabilidad y Estadística

Presenta:

Isabel Cristina García Arboleda

Director de Tesis:

Dr. Víctor Manuel Rivero Mercado



Centro de Investigación en Matemáticas, A.C.

CIMAT

Estimación Unificada del Proceso de Galton-Watson con Inmigración

T E S I S

que para obtener el grado de

Maestro en Ciencias
con Especialidad en
Probabilidad y Estadística

P r e s e n t a

Isabel Cristina García Arboleda

Comité de Evaluación:

Dr. Joaquín Ortega Sánchez
(Presidente)

Dr. Rogelio Ramos Quiroga
(Secretario)

Dr. Víctor Manuel Rivero Mercado
(Vocal y director de Tesis)

Dr. José Miguel Ponciano Castellanos
(Lector Especial)

Enero de 2008, Guanajuato, Gto. México.

Agradecimientos

A Dios, por permitirme vivir esta experiencia.

A mi asesor el Dr. Víctor Rivero, por confiar en mi, por brindarme su apoyo y por todo lo que he aprendido de él.

A mis profesores, por sus consejos y dedicación, en especial a mis sinodales, Dr. Joaquín Ortega y Dr. Rogelio Ramos, quienes con su ejemplo han contribuido a mi formación y con sus comentarios han mejorado el contenido de este trabajo.

A el Dr. José Miguel Ponciano, por sus sugerencias que aunque no están plasmadas en esta tesis me han ayudado a visualizar cosas de manera diferente, así mismo por su dedicación.

A CIMAT, por el apoyo económico para la realización de mis estudios, parte fundamental para mi formación académica, en especial a los investigadores: Dra. Graciela González-Farías, Dr. José Andrés Christen, Dra. Eloisa Díaz-Francés, Dr. Adolfo Sánchez y Dr. Pedro Luis del Angel, quienes permitieron mi llegada y mi estancia en Guanajuato.

Resumen

Una manera simple de modelar la evolución de una población es a través de un proceso de Galton-Watson con Inmigración. Dentro del contexto de Inferencia Estadística se ha probado para este proceso que las distribuciones límites de los estimadores no-secuenciales de la media de reproducción difieren drásticamente para el caso subcrítico, crítico y supercrítico. Con base en una cierta regla de paro, *Qi* y *Reeves* construyen una clase de estimadores de dos etapas para la media de reproducción y demuestran para los tres casos que éstas son Normales de manera asintótica, permitiendo esto una estimación y un procedimiento de inferencia estadística unificado para este parámetro sin ningún conocimiento a priori acerca de la clasificación del proceso.

En este trabajo se presentan bases teóricas del proceso de Galton-Watson con Inmigración, los fundamentos teóricos del método de estimación esbozado anteriormente y se comparan, por medio de un estudio de simulaciones, el procedimiento secuencial con otros métodos de estimación, entre ellos un nuevo método computacional conocido como *Data cloning*.

A Yamidt, vos hacés que todo sea más fácil.

A Fabiola, Franci y Lucelly.

Índice general

1. Preliminares	8
1.1. Proceso de Galton-Watson	9
1.1.1. Extinción y caso crítico	10
1.1.2. Propiedades asintóticas	13
1.2. Proceso de Galton-Watson con Inmigración	16
1.2.1. Función generadora y momentos	18
1.2.2. Comportamiento asintótico	23
2. Algunos Métodos de Estimación	34
2.1. Método de Máxima Verosimilitud	36
2.2. Método de Mínimos Cuadrados	41
2.3. Estimación Bootstrap	44
2.4. Otro enfoque: Estimación Bayesiana	48
2.5. Aproximación <i>Data cloning</i>	50
3. Estimación Secuencial del Proceso de G.W.I.	53
3.1. Introducción	53
3.2. Estimador secuencial de una etapa	54
3.2.1. Propiedades	56
3.3. Estimador secuencial de dos etapas	61
3.4. Comportamiento asintótico	63
4. Simulaciones	71
4.1. Características	71
4.2. Resultados	72
Conclusiones	82

Apéndice	84
A.1. Condición	84
A.2. Convergencia de σ^2	86
A.3. Difusión	88
Bibliografía	93

Introducción

En muchas situaciones es de interés modelar la evolución de una población de individuos, y una manera de hacerlo es a través de un proceso estocástico, específicamente con un proceso de ramificación. Por ejemplo, desde realizar el estudio de la propagación de cierta enfermedad infecto-contagiosa hasta la identificación de personas (cadáveres) e investigación científica sobre el ADN amplificado. Más ejemplos prácticos en los que se usan los procesos de ramificación se pueden consultar en [23].

La teoría de los procesos de ramificación la motivó inicialmente explicar el fenómeno de extinción de la descendencia de cierta familia de la aristocracia Inglesa¹. En la actualidad, éstos proveen una aproximación de un modelo matemático para describir probabilísticamente la evolución de sistemas cuyas componentes (células, partículas e individuos en general) viven un tiempo aleatorio y, que en algún momento de su tiempo de vida o al tiempo de su muerte, producen un número aleatorio de descendientes.

Adicional a su interés teórico, estos procesos tienen una mayor dimensión práctica por sus aplicaciones potenciales en diversas áreas tales como Biología, Demografía, Ecología, Epidemiología, Genética, Medicina, entre otros. El término procesos de ramificación aparece gracias a *A. N. Kolmogorov* en 1938 y *Dmitriev* en 1947 quienes lo introdujeron para describir procesos estocásticos en la teoría de probabilidad para caracterizar la dinámica de poblaciones; entre los autores que han contribuido con textos que ahora son clásicos para el estudio de dichos procesos están *S. Asmussen* y *H. Hering* [2], *K. Athreya* y *P. Ney* [3], *P. Guttorp* [15], *T. Harris* [17], *P. Jagers* [23] y *G. Sankaranarayanan* [39].

Una forma particular de modelar procesos de ramificación ocurre cuando el proceso estocástico en cuestión está indexado por un tiempo discreto, lo conforman variables aleatorias independientes e idénticamente distribuidas, donde los parámetros claves dentro del proceso, pueden llevar a la caracterización del tamaño de la población. Es por esto, que considerando la complejidad que se puede presentar en el estudio de

¹Los detalles históricos del problema pueden ser consultados en [26].

los procesos de ramificación, la estructura del caso particular del proceso de Galton-Watson con Inmigración ha sido un desafío desde el punto de vista de la inferencia estadística; por tanto, muchos autores han trabajado en el mismo, y por ello existen diferentes enfoques para atacar dicho problema.

El objetivo principal de este documento es presentar en forma autocontenida resultados concernientes al proceso de Galton-Watson con Inmigración (G.W.I.) y el fundamento teórico de un enfoque de inferencia estadística alrededor del mismo proceso. En dicho enfoque nos interesa presentar una propiedad del estimador puntual: la distribución límite del estimador de reproducción, sin dependencia de la clasificación del proceso, esto es, sin tener en cuenta si el proceso es subcrítico, crítico o supercrítico, y es a lo que nos referiremos como estimación unificada. En efecto, lo que se encuentra en la literatura es que la distribución de este estimador es diferente de acuerdo a la misma.

Debido a que nos gustaría obtener conclusiones a partir de observaciones, provenientes de un proceso de G.W.I., que nos ayuden a aproximarnos a la evolución de la población, nos concentramos en encontrar un valor adecuado (estimación puntual) o valores adecuados (estimación por intervalo) de un parámetro de interés que se puedan considerar dignos de credibilidad y a su vez, que podamos cuantificar la calidad de estas estimaciones; así obtener la distribución límite de tal estimador es vital para éste propósito. Observemos que conocer ésta distribución es de mucha utilidad desde el punto de vista de las aplicaciones puesto que si en determinado contexto nos presentan un fenómeno que se ajusta a un proceso de G.W.I. y nos preguntan qué información del proceso podemos inferir a partir de los datos, sería muy ilógico preguntar: ¿Cómo está clasificado tu proceso?.

Esta tesis se fundamenta en [36], y se estructura como sigue:

En el Capítulo 1, presentamos resultados clásicos de los procesos de Galton-Watson enfatizando en la inestabilidad del modelo y definimos una modificación del mismo; nuestro proceso de interés, el proceso de Galton-Watson con Inmigración, el cual nos permite hacer una descripción más clara de la evolución de una población. Por otro lado, en el Capítulo 2 describimos algunos enfoques de estimación puntual que se han trabajado alrededor de este proceso, así como algunas propiedades de los mismos estimadores. En el Capítulo 3, presentamos las bases teóricas de una forma más simplificada de hacer inferencia estadística paramétrica para el proceso de G.W.I. En el Capítulo 4, comprobamos las propiedades de los estimadores anteriores con un estudio basado en simulaciones, presentamos las características de dicho estudio y los resultados obtenidos en el mismo. Finalmente, en el Apéndice demostramos resultados fundamentales en las demostraciones dadas a lo largo del documento.

Capítulo 1

Preliminares

La teoría de los procesos de ramificación ha hecho importantes contribuciones para la dinámica de poblaciones y con el desarrollo de la ciencia de la computación, estos procesos se han encontrado en áreas tales como algoritmos, estructura de datos, combinatoria, biología molecular, inmunología, entre otros. Para los experimentadores, los procesos de ramificación han ayudado en el entendimiento de observaciones que obtuvieron sin tener una cuantificación de las mismas, ellos han desarrollado a partir de éstos nuevos experimentos y protocolos clínicos y también han dado predicciones que han sido probadas en situaciones de la vida real. Para los matemáticos, en cambio el hecho de entender nuevas observaciones clínicas y biológicas ha motivado el desarrollo de nuevas matemáticas en el campo de los procesos de ramificación.

Dentro del grupo de los procesos de ramificación, el proceso de Galton-Watson (G.W.) es el más viejo, simple y bien conocido proceso a tiempo discreto. El proceso de G.W. tiene una larga historia y como es de esperarse está estrechamente relacionada con un gran número de aplicaciones en las ciencias físicas y biológicas. El problema original fue introducido por *Francis Galton* en 1822 y resuelto por el Reverendo *Henry Watson* el mismo año (ver [3]) cuyo interés inicial fue la extinción del apellido de una familia de la realeza Inglesa. Históricamente, éste proceso está mal identificado en la literatura, *Heyde* y *Seneta* en 1972 descubrieron una discusión más completa de la solución del problema dada por *Bienaymé* en 1793.

En la siguiente sección, daremos una breve descripción de este proceso y algunos resultados de interés para nuestro estudio, en vista de que éstos resultados son estándar, se pueden encontrar en varios libros de probabilidad y sólo se incluirán las pruebas que son novedosas o los resultados que no se incluyen en estas referencias.

1.1. Proceso de Galton-Watson

Supongamos que una población evoluciona en generaciones y denotemos por $\{Z_n\}$ el número de individuos en la n -ésima generación. Cada individuo muere o da nacimiento a una familia, posiblemente vacía, de miembros de la $n + 1$ generación, asumimos que todos los individuos viven el mismo período de tiempo y que los tamaños de las familias de cada individuo $\zeta_{n,i}$ forman una sucesión de variables aleatorias independientes e idénticamente distribuidas (v.a.i.i.d.) con ley

$$\pi_k = \mathbb{P}[\zeta = k], \quad k = 0, 1, 2, \dots$$

Para cualquier entero no-negativo k , sea $Z_0 = k$ y definimos recursivamente

$$Z_{n+1} = \sum_{i=1}^{Z_n} \zeta_{n,i}, \quad \text{para } n \geq 0 \quad (1.1)$$

La sucesión $\{Z_n\}_0^\infty$ se conoce como el proceso de Galton-Watson con k el tamaño inicial de la población y $\{\pi_k\}$ la distribución del número de descendientes de una familia típica o “offspring distribution”.

El proceso puede pensarse como la representación de una población envolvente¹, éste inicia en el tiempo 0 con Z_0 individuos, cada individuo (después de una unidad de tiempo) se descompone independientemente de los otros individuos en un número aleatorio de descendientes, de acuerdo a la ley π_k . El número total Z_1 de individuos así producido es la suma de Z_0 v.a. y esto constituye la primera generación, con ésta, se produce la segunda generación independientemente de las anteriores y así sucesivamente.

El número de descendientes producidos por un individuo padre en un tiempo cualquiera (dado), es independiente de la historia del proceso y de los otros individuos existentes hasta el presente. Por tanto, $\{Z_n\}$ es una cadena de Markov con espacio de estados $\mathbb{N}^+ = \{0, 1, 2, \dots\}$ y las componentes de la matriz de transición \mathbf{P} están dadas por

$$p(k, j) = \mathbb{P}[Z_{n+1} = j | Z_n = k] = \sum_{i_1 + \dots + i_k = j} \pi_{i_1} \dots \pi_{i_k} = \pi_j^{*k}, \quad j=0, 1, \dots$$

donde $\pi_{(\cdot)}^{*k}$ representa la convolución de π consigo misma k -veces.

Notemos que si $Z_n = 0$ entonces $Z_{n+k} = 0$ para todo $k \geq 0$, es decir, 0 es un estado absorbente y si este estado se alcanza la población se extingue.

¹Debido a la suma de v.a.i.i.d., “convolución”.

El evento $\{Z_n = 0 \text{ para algún } n \geq 0\}$ se conoce como el evento de extinción y el evento $\{Z_n \rightarrow \infty \text{ cuando } n \rightarrow \infty\}$ se conoce como el evento de explosión. Así dos preguntas que surgen son: ¿Cuál es la probabilidad de extinción y con qué razón Z_n va a infinito en el evento de explosión?.

Para dar una respuesta parcial, denotemos por q la probabilidad de extinción del proceso $\{Z_n\}$, esto es,

$$q = \mathbb{P}[Z_n = 0 \text{ para algún } n \geq 1 | Z_0 = 1],$$

específicamente nos interesa saber cuando $q = 1$.

Observemos que el hecho de que los individuos actúen de manera independiente implica que para $k \geq 0$

$$\mathbb{P}[Z_n = 0 \text{ para algún } n \geq 1 | Z_0 = k] = q^k$$

y de aquí

$$\begin{aligned} q &= \sum_{k=0}^{\infty} \mathbb{P}[Z_n = 0 \text{ para algún } n \geq 1, Z_1 = k | Z_0 = 1] \\ &= \sum_{k=0}^{\infty} \mathbb{P}[Z_n = 0 \text{ para algún } n \geq 1 | Z_1 = k] \mathbb{P}[Z_1 = k | Z_0 = 1] \\ &= \sum_{k=0}^{\infty} q^k \pi_k \\ &= f(q) \end{aligned}$$

donde

$$f(s) \equiv \sum_{k=0}^{\infty} \pi_k s^k, \quad \text{para } 0 \leq s \leq 1 \quad (1.2)$$

es la función generadora de probabilidad de π_k con $k \geq 0$. Notemos que q es una de las raíces del polinomio $f(s)$ de grado k , así en general no se puede determinar analíticamente, sin embargo, daremos una caracterización en series de potencias la cual nos permite tener más información acerca de éste valor.

1.1.1. Extinción y caso crítico

Una herramienta importante en el análisis del proceso de Galton-Watson es la función generadora, presentamos un breve resumen de algunas propiedades que usaremos más adelante.

Función Generadora

Sea s un número real y definimos por $f_n : \mathbb{R} \rightarrow \mathbb{R}$ la función generadora (f.g.) de Z_n , la cual está dada por

$$f_n(s) = \mathbb{E}[s^{Z_n}] = \sum_{k=0}^{\infty} s^k \mathbb{P}[Z_n = k] = \sum_{k=0}^{\infty} s^k \pi_k^{*n},$$

para aquellos s para los cuales $\mathbb{E}[s^{Z_n}]$ es finita (ver [16]).

Una de las ventajas de trabajar con la función generadora es que hereda las propiedades de las series de potencias:

- Existe un radio de convergencia $r > 0$ tal que la suma es absolutamente convergente si $|s| < r$, y divergente si $|s| > r$. La suma es uniformemente convergente sobre conjuntos de la forma $\{s : |s| < r'\}$, con $r' < r$.
- f_n puede ser derivada o integrada término a término cualquier número de veces, dentro del radio de convergencia, es decir, para $|s| < r$.
- Si $f_n(s) = f_Y(s)$, para todo $s < r' \leq r$, entonces $\mathbb{P}[Z_n = k] = \mathbb{P}[Y = k]$, para todo $k \geq 0$. Además,

$$\mathbb{P}[Z_n = k] = \frac{1}{k!} f_n^{(k)}(0), \quad k \geq 0,$$

donde $f^{(r)}$ denota la r -ésima derivada de la función f .

- $f_n^{(1)}(1) = \mathbb{E}[Z_n]$, y para $k \geq 2$,

$$f_n^{(k)}(1) = \mathbb{E}[Z_n(Z_n - 1)(Z_n - 2) \dots (Z_n - k + 1)]$$

y de esto se obtiene que la varianza de Z_n esta dada por

$$\text{Var}(Z_n) = \mathbb{E}[Z_n - \mathbb{E}[Z_n]]^2 = f_n^{(2)}(1) + f_n^{(1)}(1) - (f_n^{(1)}(1))^2.$$

Cuando el radio de convergencia sea $r = 1$, las expresiones anteriores se deben tomar como

$$\lim_{s \uparrow 1} f_n^{(k)}(s).$$

Si suponemos que $Z_0 = 1$, es bien conocido que aplicando las propiedades anteriores (ver [16]), obtenemos que la media y la varianza de Z_n están dadas por

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

donde $m = \mathbb{E}[\zeta_1]$ y $\sigma^2 = \text{Var}(\zeta_1)$. Para excluir el caso trivial en que la población se extingue inmediatamente o es constante igual a 1, supondremos que $\pi_0 + \pi_1 < 1$ y $\pi_k \neq 1$ para cualquier k .

Desde la definición dada en (1.2) de $f(s)$, como una serie de potencias con coeficientes no-negativos $\{\pi_k\}$ los cuales suman 1, tenemos otras consecuencias fundamentales son:

- $f(s)$ es estrictamente convexa y creciente en $[0, 1]$.
- $f(0) = \pi_0$ y $f(1) = 1$.
- Si $Z_0 = 1$, $f_0(s) = s$, $f_1(s) = f(s)$ y $f_{n+1}(s) = f(f_n(s))$.

Ahora, como denotamos por m el parámetro que representa la media de reproducción de la población, esto es, $m = f^{(1)}(1-) = \sum_{k=0}^{\infty} k\pi_k$; por las propiedades dadas arriba, tenemos el siguiente resultado, el cual puede ser consultado en [3]-Lema 1-.

Teorema 1.1.1. *La ecuación $f(s) = s$ tiene exactamente una única raíz en $[0, 1]$ si $m > 1$ y ninguna si $m \leq 1$. Sea q tal raíz y, si $m \leq 1$ entonces $q = 1$ y si $m > 1$ entonces $q < 1$.*

Una manera de visualizar éste resultado es a partir de las siguientes gráficas:

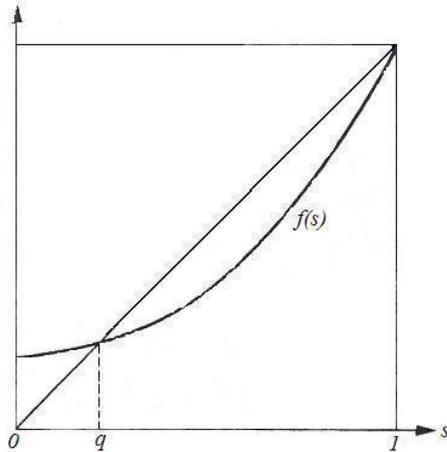


Figura 1.1: $m > 1$

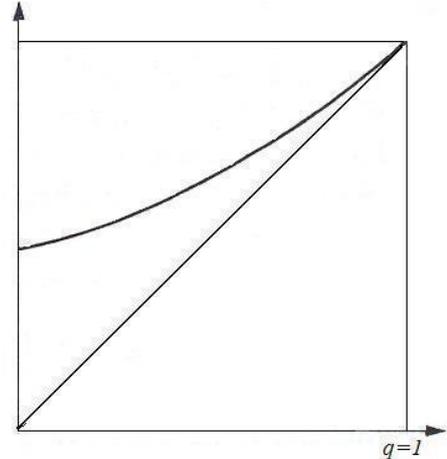


Figura 1.2: $m \leq 1$

También se puede demostrar que iterando $f(s)$ ésta converge a q lo cual se enuncia en el siguiente lema y su demostración de puede ver con detalle en [3] -Lema 2-.

Lema 1.1.1. *Si $s \in [0, q)$, entonces $f_n(s) \uparrow q$ cuando $n \rightarrow \infty$, si $s \in (q, 1)$, entonces $f_n(s) \downarrow q$ cuando $n \rightarrow \infty$ y si $s = q$ ó 1 , entonces $f_n(s) = s$ para todo n .*

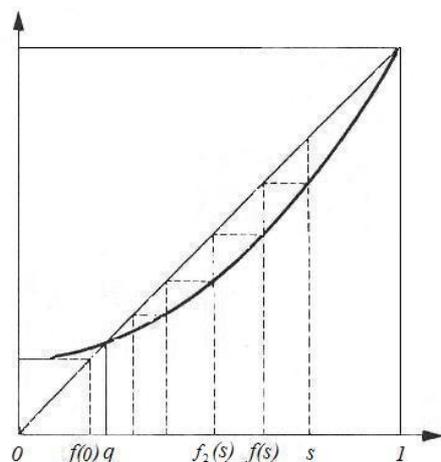


Figura 1.3: $m > 1$

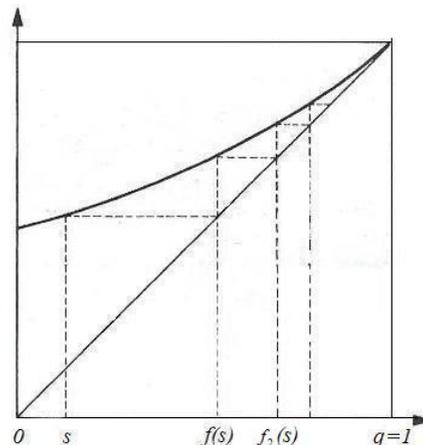


Figura 1.4: $m \leq 1$

El resultado anterior es sumamente importante debido a que podemos establecer cuando nuestro proceso se extingue o no, observando la media de reproducción m y notemos que éste parámetro es más fácil de aproximar.

Además, combinando el resultado del teorema anterior con las expresiones para la probabilidad de extinción, la media y la varianza, tenemos que cuando $m = 1$ y $n \rightarrow \infty$, $Z_n \rightarrow 0$ c.s., $\mathbb{E}[Z_n] = 1$ y $Var[Z_n] \rightarrow \infty$; indicando ésto una sustancial inestabilidad dentro del modelo. Lo cual es una buena razón para diferenciar entre tres tipos de procesos de Galton-Watson, el subcrítico, el crítico y el supercrítico cuando $m < 1$, $m = 1$ y $m > 1$, respectivamente. El caso en que $m = \infty$ no se aborda a lo largo de este documento.

1.1.2. Propiedades asintóticas

Los teoremas límites para los procesos de G.W. son importantes en muchas aplicaciones y del mismo modo desde el punto de vista teórico, puesto que representan la diversidad de resultados que se pueden obtener según la clasificación del proceso; también éstos sugieren qué se debe esperar para procesos más complicados.

En la sección anterior vimos que algunas de las propiedades del proceso de Galton-Watson las podemos deducir a partir de las propiedades de $f_n(s)$; en particular, el comportamiento asintótico de $\{f_n(s)\}$ nos introduce a los teoremas límites del proceso $\{Z_n\}$. A continuación presentamos un pequeño resumen sobre el comportamiento asintótico del proceso de G.W., donde dichos resultados se pueden consultar en [3].

En el caso supercrítico, el resultado principal es que el crecimiento es asintóticamente exponencial y con probabilidad 1 la v.a. Z_n/m^n tiende a un límite W , lo cual trae como consecuencia que Z_n se aproxima por Wm^n para n suficientemente grande; ésta es una extensión de la ley de crecimiento exponencial -*Malthusian Law*-.

Al contrario de los casos subcrítico y crítico, cuya probabilidad de extinción es igual a 1 y el límite Z_n/m^n es 0.

Debido a lo anterior, la ley exponencial no es una buena aproximación; en el caso subcrítico ésta se reemplaza por leyes límites condicional a la no-extinción del proceso, es decir, para el proceso $\{Z_n|Z_n > 0\}$, y en el caso crítico la distribución límite de $\{Z_n/n|Z_n > 0\}$ es exponencial.

De acuerdo a todo lo anterior, los procesos subcrítico y crítico eventualmente se extinguen con probabilidad 1, lo cual parece contradictorio en el caso crítico pues el valor esperado de $\{Z_n\}$ permanece constante.

Un aporte significativo en el estudio de los procesos de ramificación es el detallado por *Harris* en [17], él observó que el comportamiento anterior² de los procesos de G.W. es contrario al comportamiento de las poblaciones biológicas, las cuales tienden a rechazar un estado de balance con su ambiente, lo cual fue reiterado por *Athreya* y *Ney* en [3].

En conclusión, estamos en un escenario en el que las poblaciones tienen un crecimiento o decrecimiento de acuerdo a las leyes multiplicativas de un proceso de ramificación, pero éste tipo de contratiempo lo podemos evitar con la modificación alternativa que veremos en la siguiente sección.

Antes de pasar a definir precisamente dicha modificación al proceso de G.W. damos paso a uno de los múltiples resultados que este proceso posee (ver [3]), el cual se desprende de la definición de $f(\cdot)$ y será útil posteriormente.

Teorema 1.1.2. *Supongamos que $m = 1$ y $\sigma^2 < \infty$, entonces*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left[\frac{1}{1 - f_n(s)} - \frac{1}{1 - s} \right] = \frac{\sigma^2}{2} \quad \text{uniformemente en } 0 \leq s < 1$$

Demostración. Sea $s \in [0, 1)$, por la expansión de series de Taylor para $f(s)$ alrededor

²En Biología, se conoce como la Inestabilidad de los procesos de ramificación, es decir, mientras el tiempo tiende a infinito, el tamaño de la población se puede extinguir o ser infinitamente grande. Ésta inestabilidad se debe a los supuestos de independencia inherentes en la definición del proceso (el número de reproductores y el tiempo de vida de una nueva partícula es independiente de cualquier otra característica presente en la partícula a éste tiempo o en el futuro).

de 1, obtenemos

$$f(s) = s + \frac{\sigma^2}{2}(1-s)^2 + r(s)(1-s)^2 \quad \text{donde} \quad \lim_{s \uparrow 1} r(s) = 0.$$

Notemos que

$$\begin{aligned} \frac{1}{1-f(s)} - \frac{1}{1-s} &= \frac{f(s) - s}{(1-f(s))(1-s)} \\ &= \frac{(\sigma^2/2)(1-s)^2 + r(s)(1-s)^2}{(1-f(s))(1-s)} \\ &= \frac{1-s}{1-f(s)} \left[\frac{\sigma^2}{2} + r(s) \right] \\ &= \frac{\sigma^2}{2} + p(s) \end{aligned}$$

siendo

$$p(s) = \left(\frac{\sigma^2}{2} \right) \frac{f(s) - s}{1-f(s)} + \frac{1-s}{1-f(s)} r(s)$$

y nuevamente, $\lim_{s \uparrow 1} p(s) = 0$.

Iterando la igualdad anterior, resulta que

$$\begin{aligned} \frac{1}{n} \left[\frac{1}{1-f_n(s)} - \frac{1}{1-s} \right] &= \frac{1}{n} \sum_{j=0}^{n-1} \left\{ \frac{1}{1-f(f_j(s))} - \frac{1}{1-f_j(s)} \right\} \\ &= \frac{1}{n} \sum_{j=0}^{n-1} \left\{ \frac{\sigma^2}{2} + p(f_j(s)) \right\} \\ &= \frac{\sigma^2}{2} + \frac{1}{n} \sum_{j=0}^{n-1} p(f_j(s)) \end{aligned}$$

y como $f_n(0) \leq f_n(s) \leq 1$ y $f_n(0) \uparrow 1$, tenemos que $f_n(s) \rightarrow 1$ uniformemente para $s \in [0, 1)$ y por tanto se obtiene que

$$\frac{1}{n} \sum_{j=0}^{n-1} p(f_j(s)) \longrightarrow 0 \quad \text{cuando } n \rightarrow \infty,$$

lo cual prueba el resultado. □

Anteriormente, vimos que $f_n(s) \uparrow 1$ y el teorema anterior nos permite conocer la tasa de convergencia a la que ésto sucede, en particular, nos interesa la siguiente consecuencia:

Cuando el proceso se clasifica como crítico teníamos que $\mathbb{P}[\exists n \in \mathbb{N}, Z_n = 0] = 1$, así que $\mathbb{P}[Z_n > 0] = 0$ cuando $n \rightarrow \infty$ y si en el Teorema 1.1.2 sustituimos $s = 0$, teniendo en cuenta que $\mathbb{P}[Z_n > 0] = 1 - f_n(0)$ obtenemos una estimación de ésta tasa; éste resultado se resume en el siguiente corolario:

Corolario 1.1.3.

$$[1 - f_n(0)] \sim \frac{2}{n\sigma^2} \quad \text{cuando } n \rightarrow \infty$$

Demostración. Usando el teorema anterior y el hecho de que

$$n[1 - f_n(0)] = \left[\frac{1}{n} \left(\frac{1}{1 - f_n(0)} - 1 \right) + \frac{1}{n} \right]^{-1},$$

se sigue el resultado. □

1.2. Proceso de Galton-Watson con Inmigración

Dentro del contexto de un proceso de Galton-Watson, observamos que las poblaciones para las cuales se describe la evolución de su reproducción, se extinguen o bien crecen sin límites, esto asume implícitamente que las poblaciones permanecen aisladas. Una útil y realista modificación de este esquema es la adición de la posibilidad de inmigración a la población desde una fuente externa.

Desde el punto de vista de las aplicaciones a procesos biológicos o ecológicos tal inmigración es un factor de gran importancia. Por otra parte desde un punto de vista matemático se verá que el proceso con la componente de inmigración hace más claro uno de los teoremas límite básicos del proceso de Galton-Watson y también deja una nueva ley límite interesante. En resumen, la inmigración en el proceso subcrítico podría ser vista como un artificio pues se estabiliza el tamaño de la población pero ésta a su vez es un objeto natural de estudio en sí misma y no sólo en el caso subcrítico, como lo veremos más adelante.

Consideremos que una población evoluciona en generaciones y sea $\{Z_n; n \geq 0\}$ una sucesión de variables aleatorias tal que Z_n representa el tamaño de la n -ésima generación de la población. Esta evolución, se dá como un proceso de Galton-Watson pero simultáneamente al tiempo del nacimiento de la n -ésima generación (al tiempo n), existe una inmigración de Y_n partículas a la población.

Gráficamente,

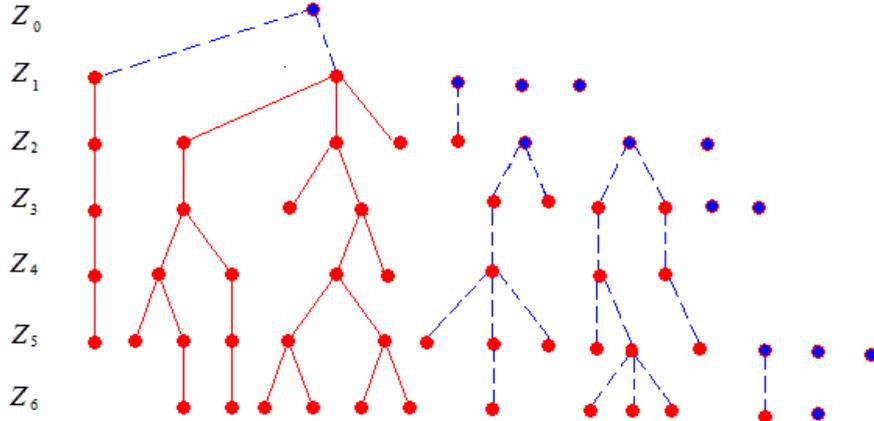


Figura 1.5: Familia típica (—) e Inmigrantes (- - -)

La dinámica del modelo está dada por las siguientes reglas:

- La generación n se conforma con los descendientes de los individuos provenientes de la generación $n - 1$ y con un número aleatorio Y_n de inmigrantes, donde los Y_i 's son independientes e idénticamente distribuidos como Y .
- Los ζ_i 's son independientes e idénticamente distribuidos como ζ .
- Condicional en Z_{n-1} , el i -ésimo individuo proveniente de la generación $n - 1$ da nacimiento a un número ζ_i de descendientes para cualquier $1 \leq i \leq Z_{n-1}$.

Lo anterior, puede formalizarse con la siguiente definición.

Definición 1.2.1. *El proceso de Galton-Watson con Inmigración (G.W.I.) está definido por una sucesión de variables aleatorias $\{Z_n\}$ las cuales están determinadas por la relación*

$$Z_n = \sum_{k=1}^{Z_{n-1}} \zeta_{n-1,k} + Y_n, \quad n \geq 1, \quad Z_0 = Y_0 \quad (1.3)$$

donde $\zeta_{n-1,k}$ denota el número de descendientes del k -ésimo individuo perteneciente a la generación $n - 1$ y Y_n denota el número de inmigrantes en la n -ésima generación.

La sucesión de v.a. $\{\zeta_{n-1,k}\}$ es i.i.d. con función generadora $f(s)$ y la sucesión de v.a. $\{Y_n\}$ es i.i.d. con función generadora $h(s)$, ambas toman valores en los enteros y las ζ 's y Y 's son independientes, además, Z_0 el tamaño inicial es una v.a. no-negativa, con valores en los enteros, la cual es independiente de $\{\zeta_{n-1,k}\}$ y $\{Y_n\}$.

De nuevo, evitamos el caso trivial así que suponemos $0 < \pi_0$, $0 < \pi_0 + \pi_1 < 1$ y $h(0) < 1$.

De lo anterior, $\{Z_n\}$ es una cadena de Markov y con la condición de que $Z_0 = 1$, su matriz de transición está caracterizada por su función generadora mediante la siguiente expresión

$$\mathbb{E}[s^{Z_n}] = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} s^j \mathbb{P}[Z_n = j | Z_{n-1} = i] \mathbb{P}[Z_{n-1} = i],$$

a partir de la f.g. de Z_n se puede hacer un mejor análisis del proceso, considerando el caracter Markoviano del mismo.

1.2.1. Función generadora y momentos

Como vimos en el caso de los procesos de G.W., la f.g. nos ayuda a conocer más acerca de la evolución de la población ya que nos permite calcular de manera simplificada los momentos del proceso, específicamente calculamos el primer y segundo momento, y a su vez el comportamiento asintótico del mismo. Los siguientes dos lemas, nos identifican éstas dos características de interés y son una consecuencia directa de la genealogía del proceso.

En el siguiente lema, consideramos el caso en que $Z_0 = 1$ y debido a la estructura de independencia de las variables se puede realizar un desarrollo similar para la derivación de la siguiente expresión cuando $Z_0 = k$ para algún $k \geq 0$.

Lema 1.2.1. *Para $n \geq 0$, sea g_n la función generadora de $\{Z_n\}$ entonces*

$$g_{n+1}(s) = h(s)g_n[f(s)] \quad \text{para } s \in [0, 1]. \quad (1.4)$$

Demostración. Como $g_n(s) = \mathbb{E}[s^{Z_n}]$ para $s \in [0, 1]$ entonces

$$\begin{aligned} g_{n+1}(s) &= \mathbb{E}[s^{Z_{n+1}}] \\ &= \mathbb{E}[\mathbb{E}[s^{Z_{n+1}} | Z_n]] \\ &= \mathbb{E}[\mathbb{E}[s^{Y_{n+1}} s^{\sum_{i=1}^{Z_n} \zeta_{n,i}} | Z_n]] \\ &= \mathbb{E}[s^{Y_{n+1}}] \mathbb{E}[\mathbb{E}[s^{\sum_{i=1}^{Z_n} \zeta_{n,i}} | Z_n]] \\ &= h(s) \sum_j [f(s)]^j \mathbb{P}[Z_n = j] \\ &= h(s) \mathbb{E}[f(s)^{Z_n}] \\ &= h(s)g_n(f(s)), \end{aligned}$$

lo anterior gracias a que Y_{n+1} es independiente de Z_n y de $\sum_{i=1}^{Z_n} \zeta_{n,i}$. \square

Observación 1.2.1. A partir del Lema 1.2.1, tenemos que

$$g_n(s) = \prod_{k=0}^n h[f_k(s)] \quad \text{para } n \geq 1,$$

veamos por inducción que esta identidad es válida.

Recordemos que para $s > 0$, $f_0(s) = s$, además, $g_0(s) = \mathbb{E}[s^{Z_0}] = \mathbb{E}[s^{Y_0}] = h(s)$, luego del Lema 1.2.1 tenemos que $g_1 = h(s)g_0(f(s)) = h(f_0(s))h(f(s))$ equivalentemente $g_1(s) = \prod_0^1 h(f_k(s))$ y por tanto, el resultado se cumple para $n = 1$.

Ahora, si suponemos cierto el resultado para n :

$$\begin{aligned} g_{n+1}(s) &= h(s)g_n[f(s)] \\ &= h(s) \prod_{k=0}^n h[f_k(f(s))] \\ &= h(s) \prod_{k=0}^n h[f_{k+1}(s)] \\ &= h(s) \prod_{k=1}^{n+1} h[f_k(s)] \\ &= h(f_0(s)) \prod_{k=1}^{n+1} h[f_k(s)] \\ &= \prod_{k=0}^{n+1} h(f_k(s)), \end{aligned}$$

lo cual conlleva a verificar la afirmación.

Por otro lado, recordemos que los momentos de una v.a. constituyen un conjunto de características numéricas que se asocian a la distribución de probabilidad de la misma y cuando éstos existen pueden expresarse en términos de las derivadas de $g_n(s)$, en $s = 1$.

En adición, se denotarán por $m = \mathbb{E}[\zeta]$, $\sigma^2 = \mathbb{E}[\zeta - m]^2$, $\lambda = \mathbb{E}[Y]$ y $\sigma_Y^2 = \mathbb{E}[Y - \lambda]^2$, los primeros y segundos momentos del número de descendientes e inmigrantes, respectivamente.

A continuación se presentan la media y la varianza de $\{Z_n\}$, éstas nos dan una descripción de la localización central y de la dispersión de la distribución de probabilidad y están dadas por las siguientes expresiones:

Lema 1.2.2. *Sea $\{Z_n\}$ el tamaño de la n -ésima generación de un proceso de G.W.I., tenemos que*

$$\mathbb{E}[Z_n] = \begin{cases} (n+1)\lambda & \text{si } m = 1, \\ \left(\frac{m^{n+1}-1}{m-1}\right)\lambda & \text{si } m \neq 1. \end{cases}$$

$$\text{Var}[Z_n] = \begin{cases} (n+1) \left[\sigma_Y^2 + \sigma^2 \lambda \frac{n+1}{2} \right] & \text{si } m = 1, \\ \sigma_Y^2 \left[\frac{m^{2(n+1)}-1}{m^2-1} \right] + \sigma^2 \lambda \left[\frac{m^{2n+1}-m^{n+1}-m^{n+1}}{(m-1)(m^2-1)} \right] & \text{si } m \neq 1. \end{cases}$$

Demostración. Como $g_{n+1}(s) = h(s)g_n(f(s))$ al derivar y evaluar en $s = 1$, se tiene que

$$g_{n+1}^{(1)}(1) = h^{(1)}(1)g_n(f(1)) + h(1)g_n^{(1)}(f(1))f^{(1)}(1)$$

esto es,

$$\mathbb{E}[Z_{n+1}] = \lambda + m \mathbb{E}[Z_n], \text{ para todo } n \geq 1.$$

Si $m = 1$, $\mathbb{E}[Z_{n+1}] = \lambda + \mathbb{E}[Z_n]$ para todo $n \geq 1$ y la solución de ésta ecuación en recurrencias es $\mathbb{E}[Z_{n+1}] = (n+1)\lambda + \mathbb{E}[Z_0]$ y como $Z_0 = Y_0$, se sigue el resultado.

En el caso en que $m \neq 1$,

$$\begin{aligned} \mathbb{E}[Z_{n+1}] &= \lambda + m \mathbb{E}[Z_n] \\ &= \lambda + m(\lambda + m \mathbb{E}[Z_{n-1}]) \\ &= \lambda + m\lambda + m^2 \mathbb{E}[Z_{n-1}] \\ &\vdots \\ &= \lambda + m\lambda + m^2\lambda + \dots + m^{n+1} \mathbb{E}[Z_0] \\ &= \lambda[1 + m + m^2 + \dots + m^{n+1}] \\ &= \lambda \left[\frac{m^{n+2} - 1}{m - 1} \right] \end{aligned}$$

Para probar la segunda afirmación, se deriva dos veces y al evaluar en $s = 1$, se obtiene

$$g_{n+1}^{(2)}(1) = h^{(2)}(1) + 2m\lambda \mathbb{E}[Z_n] + m^2 g_n^{(2)}(1) + f^{(2)}(1) \mathbb{E}[Z_n],$$

usando la expresión para la varianza y realizando algunas operaciones algebraicas se llega a que

$$\text{Var}(Z_{n+1}) = \sigma_Y^2 + m^2 \text{Var}(Z_n) + \sigma^2 \mathbb{E}[Z_n].$$

De manera recursiva,

$$\text{Var}(Z_n) = \sigma_Y^2 \left(\sum_{i=0}^n m^{2i} \right) + \sigma^2 \left(\sum_{i=0}^{n-1} m^{2i} \mathbb{E}[Z_{n-1-i}] \right),$$

y al desarrollar cada suma geométrica, se concluye que si $m = 1$,

$$\text{Var}(Z_n) = (n+1) \left[\sigma_Y^2 + \frac{n\sigma^2\lambda}{2} \right]$$

y cuando $m \neq 1$,

$$\text{Var}(Z_n) = \sigma_Y^2 \left[\frac{m^{2(n+1)} - 1}{m^2 - 1} \right] + \sigma^2\lambda \left[\frac{m^{2n+1} - m^{n+1} - m^n + 1}{(m-1)(m^2-1)} \right].$$

□

Cabe notar que si estamos en el caso crítico o supercrítico, asintóticamente el valor esperado del tamaño de la población es infinito, en cambio, en el caso subcrítico tenemos que

$$\lim_{n \rightarrow \infty} \mathbb{E}[Z_n] = \lim_{n \rightarrow \infty} \left(\frac{m^{n+1} - 1}{m - 1} \right) \lambda = \frac{\lambda}{1 - m}.$$

En otras palabras, asintóticamente en el caso subcrítico el tamaño promedio de la población permanece constante, esto nos induce a pensar que la distribución del tamaño de la población también se estabiliza. Es importante notar que el proceso de G.W.I. se puede construir como una suma de procesos de ramificación independientes, cada uno iniciado al tiempo de una inmigración, y si cada uno de estos es subcrítico, es razonable pensar que $\{Z_n\}$ puede converger a un límite, es decir, el proceso podría tener una distribución estacionaria.

Si hacemos $n \rightarrow \infty$ en (1.4), observamos que si $\{Z_n\}$ converge a una distribución límite, la f.g. $g(s)$ de ésta distribución debe satisfacer

$$g(s) = h(s)g(f(s)).$$

Lo anterior, origina la pregunta: ¿Cuándo g es función generadora de una v.a. no-degenerada?, la cual responde el siguiente teorema, demostrado por *Foster* y *Williamson* en [12].

Teorema 1.2.1. *Si $m \leq 1$ entonces Z_n tienen una distribución límite propia cuando $n \rightarrow \infty$ sí y sólo sí*

$$\int_0^1 \frac{1 - h(s)}{f(s) - s} ds < \infty$$

Demostración. Por la expansión en series de Taylor para $h(\cdot)$, alrededor de 1 y sustituyendo $s = f_k(s)$ se obtiene

$$h(f_k(s)) = 1 - h^{(1)}(1)(1 - f_k(s)) - o(1 - f_k(s))$$

ya que $r(f_k(s))(1 - f_k(s)) = o(1 - f_k(s))$ para $s \in (0, 1)$.

Como

$$g_n(s) = \prod_{k=0}^n h(f_k(s)), \quad s \in (0, 1]$$

y $g_{n+1}(s) \leq g_n(s) \leq 1$ entonces, $g(s) = \lim_{n \rightarrow \infty} g_n(s)$ existe para $0 \leq s \leq 1$, ya que cada uno de los factores del producto es menor o igual que 1, y para $s > 0$ fijo,

$$g(s) > 0 \quad \text{sí y sólo sí} \quad \sum_{k=0}^{\infty} [1 - h(f_k(s))] < \infty.$$

Además,

$$\sum_{k=0}^{\infty} [1 - h(f_k(s))] < \infty \quad \text{sí y sólo sí} \quad \sum_{k=0}^{\infty} [1 - h(f_k(0))] < \infty,$$

pues existe un entero positivo $l = l(s)$ tal que $s \leq f_l(0)$ y así $f_k(0) \leq f_k(s) \leq f_{k+l}(0)$.

Para el límite $g(s)$, se dan los siguientes casos:

- $g(s) = 0$ para $0 < s < 1$
- $g(s) > 0$ para $0 < s < 1$

en el segundo caso, por el *Teorema de Convergencia Dominada*

$$\lim_{s \uparrow 1} g(s) = \lim_{n \rightarrow \infty} g(f_n(0)) = \lim_{n \rightarrow \infty} \prod_{k=0}^{\infty} h(f_k(f_n(0))) = \lim_{n \rightarrow \infty} \prod_{k=n}^{\infty} h(f_k(0)) = 1$$

y por *Teorema de Continuidad* (ver [10]), se tiene que $g(\cdot)$ es la función generadora de una distribución propia.

Ahora, sea $s_k = f_k(0)$ lo cual constituye una partición del $[0, 1]$, y la convexidad de las funciones generadoras implica que

$$\frac{1 - h(s)}{1 - s} \quad \text{y} \quad \frac{1 - s}{f(s) - s} = \left[1 - \frac{1 - f(s)}{1 - s} \right]^{-1} \quad \text{son crecientes}$$

y de aquí,

$$\begin{aligned}
\sum_{k=0}^{\infty} [1 - h(s_k)] &= \sum_{k=0}^{\infty} \frac{1 - h(s_k)}{f(s_k) - s_k} (s_{k+1} - s_k) \\
&\leq \int_0^1 \frac{1 - h(s)}{f(s) - s} ds \\
&\leq \sum_{k=0}^{\infty} \frac{1 - h(s_{k+1})}{f(s_{k+1}) - s_{k+1}} (s_{k+1} - s_k) \\
&= \sum_{k=0}^{\infty} [1 - h(s_{k+1})] \frac{f(s_k) - s_k}{f(f(s_k)) - f(s_k)},
\end{aligned}$$

puesto que

$$\frac{f(s) - s}{f(f(s)) - f(s)} = \frac{1}{f^{(1)}(t)} \quad \text{para algún } s < t < f(s),$$

y $\frac{1}{f^{(1)}(t)} \uparrow \frac{1}{f^{(1)}(1)}$ cuando $s \uparrow 1$, la convergencia de nuestra condición, es equivalente a la convergencia de

$$\sum_{k=0}^{\infty} [1 - h(s_k)].$$

□

Observación 1.2.2. Recordemos que una cadena de Markov irreducible y aperiódica es positivo recurrente o ergódica sí y sólo sí las probabilidades de transición convergen a una ley de probabilidad no-degenerada conforme el tiempo tiende a infinito. Por lo tanto, el teorema anterior nos dá una condición suficiente y necesaria para garantizar la ergodicidad de nuestro proceso, en los casos subcrítico y crítico.

1.2.2. Comportamiento asintótico

Como habíamos dicho, nos interesa tener información acerca del comportamiento límite de $\{Z_n\}$ cuando $n \rightarrow \infty$. A continuación, para el proceso de G.W.I., se presentan varios resultados pertinentes en los casos, subcrítico, crítico y supercrítico.

Antes de dar paso a dichos resultados, necesitamos primero algunos lemas que se cumplen en general (ver [37]) y nos serán útiles para la demostración de los mismos.

Inicialmente, recordemos que para una sucesión de conjuntos $\{A_n, n \geq 1\}$

$$\limsup_{n \rightarrow \infty} \{A_n\} = \bigcap_{n \geq 1} \{ \bigcup_{k \geq n} A_k \} = \{w : w \in A_n \text{ para infinitos valores de } n\},$$

así en adelante, denotaremos

$$\limsup_{n \rightarrow \infty} A_n = \{A_n, i.v.\}.$$

Lema 1.2.3. Sean X_1, X_2, \dots una sucesión de v.a.i.i.d. no-negativas con la misma ley de X , entonces c.s.

$$\limsup_{n \rightarrow \infty} \frac{X_n}{n} = \begin{cases} 0 & \text{si } \mathbb{E}[X] < \infty, \\ \infty & \text{si } \mathbb{E}[X] = \infty. \end{cases}$$

Demostración. Sea $x > 0$ notemos que

$$\sum_{n=1}^{\infty} \mathbb{P} \left[\frac{X_n}{n} \geq x \right] = \sum_{n=1}^{\infty} \mathbb{P} \left[\frac{X_n}{x} \geq n \right] = \frac{\mathbb{E}[X_n]}{x}.$$

En el caso en que $\mathbb{E}[X] = \infty$,

$$\sum_{n=1}^{\infty} \mathbb{P} \left[\frac{X_n}{n} \geq x \right] = \infty, \quad \text{para } x > 0,$$

como los eventos $\{\frac{X_1}{1} \geq x\}, \{\frac{X_2}{2} \geq x\}, \dots$, son independientes, el *Lema de Borel-Cantelli* (ver [37]) implica que

$$\limsup_{n \rightarrow \infty} \frac{X_n}{n} \geq x \quad \text{c.s. para todo } x > 0,$$

y haciendo $x \rightarrow \infty$ se prueba el resultado.

Por otro lado, si $\mathbb{E}[X] < \infty$ entonces

$$\sum_{n=1}^{\infty} \mathbb{P} \left[\frac{X_n}{n} \geq x \right] < \infty$$

y por el *Lema de Borel-Cantelli*

$$\mathbb{P} \left[\frac{X_n}{n} > x, i.v. \right] = 0$$

o equivalentemente

$$\limsup_{n \rightarrow \infty} \frac{X_n}{n} \leq x \quad \text{c.s. } \forall x > 0,$$

lo cual implica

$$\limsup_{n \rightarrow \infty} \frac{X_n}{n} = 0 \quad \text{c.s.}$$

□

Lema 1.2.4. Sean X_1, X_2, \dots v.a.i.i.d. con la misma ley de X , que toman valores en los enteros. Para cualquier $c > 1$,

$$\text{si } \mathbb{E}[\log^+ X] = \infty \text{ entonces } \limsup_{n \rightarrow \infty} \frac{X_n}{c^n} = \infty \text{ c.s.,}$$

$$\text{ó si } \mathbb{E}[\log^+ X] < \infty \text{ entonces } \sum_{n \geq 1} \frac{X_n}{c^n} < \infty \text{ c.s.}$$

Demostración. Sean $W_n = \log^+ X_n$ para $n \geq 1$ y $W = \log^+ X$. Supongamos que $\mathbb{E}[W] = \infty$ y como

$$\begin{aligned} c^{-n} X_n &= \exp(\log(c^{-n} X_n)) \\ &= \exp(\log X_n - \log c^n) \\ &= \exp(W_n - n \log c) \\ &= \exp\left[(-n) \left(\log c - \frac{W_n}{n}\right)\right], \end{aligned}$$

si aplicamos el Lema 1.2.3, obtenemos

$$\limsup_{n \rightarrow \infty} \frac{W_n}{n} = \infty \text{ c.s.,}$$

por lo tanto,

$$\limsup_{n \rightarrow \infty} \frac{X_n}{c^n} = \infty \text{ c.s.}$$

Demostraremos ahora la segunda afirmación, para lo cual usaremos que para $1 < \mu \leq c$ se tiene que

$$\sum_{n=1}^{\infty} \mathbb{P}[X_n > \mu^n] = \sum_{k=1}^{\infty} \sum_{n=1}^{\infty} \mathbf{1}_{\{k > \mu^n\}} \mathbb{P}[X_n = k],$$

y además

$$\sum_{n=1}^{\infty} \mathbf{1}_{\{k > \mu^n\}} = \left\lfloor \frac{\log k}{\log \mu} \right\rfloor$$

donde $\lfloor \cdot \rfloor$ denota la parte entera.

Dado que $\mathbb{E}[W] = \sum_k \log k \mathbb{P}[X = k]$ y teniendo en cuenta lo anterior, se tiene que

$$\sum_{n=1}^{\infty} \mathbb{P}[X_n > \mu^n] = \sum_{k=1}^{\infty} \left\lfloor \frac{\log k}{\log \mu} \right\rfloor \mathbb{P}[X = k] = \mathbb{E} \left(\left\lfloor \frac{\log^+ X}{\log \mu} \right\rfloor \right)$$

y por hipótesis $\sum_{n=1}^{\infty} \mathbb{P}[X_n > \mu^n] < \infty$, lo cual aplicando el *Lema de Borel-Cantelli* implica que $\mathbb{P}[X_n > \mu^n, i.v.] = 0$, eligiendo $\mu < c$ tenemos que la cola de la serie $\sum_n c^{-n} X_n$ eventualmente será dominada por la cola de $\sum_n c^{-n} \mu^n < \infty$. \square

Caso subcrítico

Este caso, lo estudió inicialmente *Heathcote* [18], luego *Seneta* [40] y posteriormente lo retomaron *Foster* y *Williamson* [12].

El siguiente teorema es el resultado fundamental dentro de dicho proceso para este caso, la siguiente es una prueba desarrollada en [30] con teoría más moderna a la propuesta inicialmente.

Teorema 1.2.2. *Para $m \in (0, 1)$, se satisfacen de manera dicotómica, las siguientes afirmaciones*

- $\mathbb{E}[\log^+ Y] = \infty$ implica que $\{Z_n\}$ converge en probabilidad a $+\infty$.
- $\mathbb{E}[\log^+ Y] < \infty$ implica que $\{Z_n\}$ converge en distribución a una variable aleatoria no-degenerada.

Demostración. Observemos que $\{Z_n, n \geq 0\}$, el tamaño de la n -ésima generación de la población, se puede escribir como

$$Z_n = \sum_{k=0}^n \sum_{l=1}^{Y_k} \zeta_{n-k}^{k,l}$$

donde $\{\zeta_{(\cdot)}^{k,l}\}$ son procesos de G.W. i.i.d. iniciando al tiempo k ; debido a que sus distribuciones dependen sólo de $n - k$, es decir, no depende de l y $(Y_0, Y_1, \dots, Y_n) \stackrel{d}{=} (Y_n, Y_{n-1}, \dots, Y_0)$ entonces

$$Z_n \stackrel{d}{=} \hat{Z}_n = \sum_{k=0}^n \sum_{l=1}^{Y_k} \zeta_k^{k,l}$$

de aquí $\{Z_n\}$ converge en distribución sí y sólo sí $\{\hat{Z}_n\}$ lo hace, y esta última suma crece en n a algún límite \hat{Z}_∞ , por lo tanto, es suficiente determinar cuando \hat{Z}_∞ es finito o infinito con probabilidad mayor o igual a cero ya que por la *Ley 0 - 1 de Kolmogorov*, \hat{Z}_∞ es finito o infinito c.s., respectivamente.

En efecto, sea \mathcal{G} la σ -álgebra generada por $\{Y_k, k \geq 1\}$ y sea $\mathbb{E}[\log^+ Y] < \infty$,

entonces

$$\begin{aligned}
\mathbb{E}[\hat{Z}_\infty | \mathcal{G}] &= \mathbb{E} \left[\sum_{k=0}^{\infty} \sum_{l=1}^{Y_k} \zeta_k^{k,l} | \mathcal{G} \right] \\
&= \sum_{k=0}^{\infty} \mathbb{E} \left[\sum_{l=1}^{Y_k} \zeta_k^{k,l} | \mathcal{G} \right] \\
&= \sum_{k=0}^{\infty} Y_k m^k
\end{aligned}$$

y por el Lema 1.2.4 esta esperanza condicional es finita c.s., lo cual implica que \hat{Z}_∞ es finita con probabilidad mayor que cero y por tanto, finita c.s.

Para la segunda afirmación, supongamos que $\mathbb{E}[\log^+ Y] = \infty$ y si $\hat{Z}_\infty < \infty$ c.s., escribiendo $W_k = \sum_{l=1}^{Y_k} \zeta_k^{k,l}$, note que para cada k , W_k es una v.a. con valores en los enteros donde sólo un número finito de estos es diferente de cero. Por el *Lema de Borel-Cantelli*, condicional en \mathcal{G} se tiene que

$$\sum_{k=0}^{\infty} \mathbb{P}[W_k \neq 0 | \mathcal{G}] < \infty \quad \text{c.s.},$$

luego, por la independencia de los procesos y $\{\zeta_{(\cdot)}^{k,l}\} \stackrel{d}{=} \{\zeta_k\}$

$$\mathbb{P}[W_k \neq 0 | \mathcal{G}] \geq Y_k \mathbb{P}[\zeta_k \neq 0] \geq Y_k \mathbb{P}[\zeta \neq 0]^k,$$

de donde $\sum_{k=0}^{\infty} Y_k \mathbb{P}[\zeta \neq 0]^k < \infty$ c.s. lo cual es una contradicción ya que por hipótesis y por el Lema 1.2.4,

$$\limsup_{k \rightarrow \infty} Y_k \mathbb{P}[\zeta \neq 0]^k = \infty \quad \text{c.s.}$$

□

Observemos que a diferencia del proceso de G.W., la adición de la componente de inmigración trae como consecuencia la no-extinción de la población, lo cual es de gran importancia desde el punto de vista práctico en la utilización del modelo. Además, este resultado reafirma las conclusiones citadas en la Observación 1.2.2.

De hecho, aunque en este caso se evita la extinción de la población, en la práctica es fundamental conocer la distribución a la cual converge el tamaño de la población, esta distribución fue identificada como una distribución de tipo *Pearson*, la cual la determinaron *González y Molina* en [13] a través de las primeras cumulantes para las leyes de reproducción e inmigración.

Caso crítico

Seneta en [41], fue quien trabajó desde un principio en este caso para el proceso de G.W.I., otras contribuciones las hicieron *Foster* y *Pakes* (ver [11] y [35]).

Teorema 1.2.3. *Para $m = 1$, si $\sigma^2 < \infty$ entonces $Z_n \xrightarrow{d} \infty$. Si además, $0 < \lambda < \infty$ entonces $\{Z_n/n\}$ converge en distribución a una variable aleatoria con densidad Gamma*

$$\frac{1}{\Gamma(\alpha)\beta^\alpha} u^{\alpha-1} \exp^{-\frac{u}{\beta}}, \quad u \in \mathbb{R}^+$$

donde $\alpha = \frac{2\lambda}{\sigma^2}$ y $\beta = \frac{\sigma^2}{2}$.

Demostración. Para la primera parte, de nuevo, como en el caso subcrítico,

$$Z_n \stackrel{d}{=} \hat{Z}_n = \sum_{k=0}^n \sum_{l=1}^{Y_k} \zeta_k^{k,l}$$

de aquí $\{Z_n\}$ converge en distribución sí y sólo sí $\{\hat{Z}_n\}$ lo hace, ya que \hat{Z}_n es creciente, la convergencia en distribución se da sí y sólo sí $\hat{Z}_\infty < \infty$ c.s., pero para que esto ocurra se debe tener que $\zeta_k^{k,l} > 0$ para sólo un número finito k, l con $l \leq Y_k$. Condicionando sobre $\mathcal{G} = \sigma(Y_0, \dots, Y_n)$ y usando el *Lema de Borel-Cantelli*, la convergencia en distribución de $\{Z_n\}$ se cumple sí y sólo sí $\sum_{n=0}^{\infty} \mathbb{P}[\zeta_n > 0] Y_n = \sum_{n=0}^{\infty} (1 - f_n(0)) Y_n < \infty$, siendo $f_n(0)$ la probabilidad de extinción para cualquier $\{\zeta_{(\cdot)}^{k,l}\}$.

Si $\sigma^2 < \infty$, por el Corolario 1.1.3 tenemos que

$$\lim_{n \rightarrow \infty} n(1 - f_n(0)) = \frac{2}{\sigma^2}$$

esto es, $1 - f_n(0) = O(1/n)$ y como las Y_i 's son independientes, se tiene que

$$\sum_{n=0}^{\infty} \frac{Y_n}{n} = \infty \quad \text{c.s.},$$

lo cual implica que,

$$\sum [1 - f_n(0)] Y_n = \infty \quad \text{c.s.}$$

y de lo anterior, $\{Z_n\}$ no converge en distribución.

Para la segunda parte, recordemos que $g_n(\exp^{-s}) = l_n(s)$ siendo $l_n(s) = \mathbb{E}[\exp^{-sZ_n}]$ la Transformada de Laplace para la v.a. Z_n y como la función $\log(\cdot)$ es continua, si

usamos el Teorema de Continuidad para la *Transformada de Laplace* (ver [10]) es suficiente probar que

$$-\log g_n(\exp^{-s/n}) \longrightarrow \alpha \log(1 + s\beta) \quad \text{para } s > 0 \text{ fijo.} \quad (1.5)$$

Inicialmente, encontremos una expresión para el logaritmo de la f.g. $g_n(\cdot)$:

Como $g_n(s) = \prod_{j=0}^n h(f_j(s))$ entonces

$$-\log g_n(\exp^{-s/n}) = -\sum_{j=0}^n \log h(f_j(\exp^{-s/n})) \quad (1.6)$$

de la expansión de Taylor de $-\log h(s)$ alrededor de 1,

$$-\log h(s) = \lambda(1-s) + (1-s) \left[\frac{\log^{(2)} h(1)}{2!} + \dots \right] \quad \text{si denotamos por}$$

$$p(s) = \sum_{k=1}^{\infty} (1-s)^k [\log^{(k+1)} h(1)], \quad \text{por propiedad del residuo,} \quad \lim_{s \rightarrow 1} p(s) = 0.$$

Notemos que por el Lema 1.1.2,

$$\frac{1}{j} \left[\frac{1}{1-f_j(s)} - \frac{1}{1-s} \right] = \beta + \frac{1}{j} \sum_{k=0}^{j-1} w(f_k(s))$$

si denotamos por

$$r_j(s) = \frac{1}{j} \sum_{k=0}^{j-1} w(f_k(s)), \quad r_j(s) \longrightarrow 0 \text{ uniformemente para } s \in [0, 1)$$

y despejando $1 - f_j(s)$, tenemos,

$$1 - f_j(s) = \frac{1-s}{1 + j(1-s)[\beta + r_j(s)]}.$$

Después de sustituir en (2.19), obtenemos

$$-\log g_n(\exp^{-s/n}) = \sum_{j=0}^n \frac{(\lambda + p(f_j(\exp^{-s/n}))(1 - \exp^{-s/n}))}{1 + j(\beta + r_j(\exp^{-s/n}))(1 - \exp^{-s/n})}.$$

Ahora, veamos la convergencia de la ecuación (1.5):

Dado $0 < \epsilon < \beta$, ya que $f_j \geq f$ existe N_1 tal que $|p(f_j(\exp^{-s/n}))| < \epsilon$ para $n \geq N_1$, además, también existe N_2 tal que $|r_j(\exp^{-s/n})| < \epsilon$ para todo n y $j \geq N_2$.

Sea $N = \min\{N_1, N_2\}$, así

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{j=0}^{N-1} \frac{(\lambda + p(f_j(\exp^{-s/n}))(1 - \exp^{-s/n}))}{1 + j(\beta + r_j(\exp^{-s/n}))(1 - \exp^{-s/n})} &= 0 \quad \text{y para } n \geq N \\ \sum_{j=N}^n \frac{(\lambda - \epsilon)(1 - \exp^{-s/n})}{1 + j(\beta + \epsilon)(1 - \exp^{-s/n})} &\leq \sum_{j=N}^n \frac{(\lambda + p(f_j(\exp^{-s/n}))(1 - \exp^{-s/n}))}{1 + j(\beta + r_j(\exp^{-s/n}))(1 - \exp^{-s/n})} \\ &\leq \sum_{j=N}^n \frac{(\lambda + \epsilon)(1 - \exp^{-s/n})}{1 + j(\beta - \epsilon)(1 - \exp^{-s/n})} \end{aligned}$$

denotemos por

$$S(n) = \sum_{j=N}^n \frac{(\lambda - p(f_j(\exp^{-s/n}))(1 - \exp^{-s/n}))}{1 + j(\beta + r_j(\exp^{-s/n}))(1 - \exp^{-s/n})}.$$

Usando, una desigualdad entre sumas e integrales ³, obtenemos

$$\begin{aligned} \frac{\lambda - \epsilon}{\beta + \epsilon} \log \frac{1 + (\beta + \epsilon)(1 - \exp^{-s/n})(n - 1)}{1 + (\beta + \epsilon)(1 - \exp^{-s/n})N} &\leq S(n) \\ &\leq \frac{\lambda + \epsilon}{\beta - \epsilon} \log \frac{1 + (\beta - \epsilon)(1 - \exp^{-s/n})n}{1 + (\beta - \epsilon)(1 - \exp^{-s/n})(N - 1)} \end{aligned}$$

y de lo anterior,

$$\left\{ \frac{\lambda - \epsilon}{\beta + \epsilon} \right\} \log(1 + (\beta + \epsilon)s) \leq \liminf_{n \rightarrow \infty} S(n) \leq \limsup_{n \rightarrow \infty} S(n) \leq \left\{ \frac{\lambda + \epsilon}{\beta - \epsilon} \right\} \log(1 + (\beta - \epsilon)s)$$

y como ϵ es arbitrario, entre 0 y β , se sigue el resultado. \square

³Para cualquier a y b no-negativos, se cumple que

$$(a/b) \log \frac{1 + b(n - 1)}{1 + bN} \leq \sum_{j=N}^n \frac{a}{1 + bj} \leq (a/b) \log \frac{1 + bn}{1 + b(N - 1)}.$$

Una interpretación del resultado anterior es que, en cada punto en el tiempo sobre la media de los individuos inmigrantes, por el Corolario 1.1.3 se esperaría que λ/β procesos sobrevivan a largo plazo donde cada uno de estos tiene tamaño distribuido Exponencialmente con media β .

Por otro lado, como lo habíamos dicho la componente de inmigración nos deja una nueva ley límite dentro del proceso, es decir, bajo el supuesto de que $\lambda < \infty$ se evita la extinción de la población y más aún, sabemos a cual distribución crece la misma, lo cual es muy significativo en el campo de las aplicaciones.

Caso supercrítico

Nuevamente, *Seneta* fue uno de los principales contribuyentes en este caso del proceso de G.W.I. A continuación, de la misma manera que en el caso subcrítico la prueba que se presenta está desarrollada con base en una teoría moderna (ver [30]).

Teorema 1.2.4. *Para $m > 1$, se satisfacen de manera dicotómica, las siguientes afirmaciones*

- $\mathbb{E}[\log^+ Y] < \infty$ implica que $\lim_{n \rightarrow \infty} \frac{Z_n}{m^n}$ existe y es finito c.s.
- $\mathbb{E}[\log^+ Y] = \infty$ implica que $\limsup_{n \rightarrow \infty} \frac{Z_n}{c^n} = \infty$ para cualquier c positivo c.s.

Demostración. Supongamos que $\mathbb{E}[\log^+ Y] = \infty$ y como $Y_n \stackrel{d}{=} Y$, se tiene por el Lema 1.2.4,

$$\limsup_{n \rightarrow \infty} \frac{Y_n}{c^n} = \infty \quad \text{c.s.}$$

ya que $Z_n \geq Y_n$, se sigue el resultado.

Ahora, supongamos que $\mathbb{E}[\log^+ Y] < \infty$, sea $\mathcal{F}_n = \sigma(Z_0, \dots, Z_n)$ y $\mathcal{G} = \sigma(Y_0, Y_1, \dots)$ para $n \geq 0$, entonces

$$\mathbb{E} \left[\frac{Z_{n+1}}{m^{n+1}} \middle| \mathcal{F}_n \wedge \mathcal{G} \right] = \frac{1}{m^{n+1}} \mathbb{E} \left[\sum_{i=1}^{Z_n} \zeta_i + Y_{n+1} \middle| \mathcal{F}_n \wedge \mathcal{G} \right] = \frac{Z_n}{m^n} + \frac{Y_{n+1}}{m^{n+1}}. \quad (1.7)$$

Notemos que de la igualdad anterior y de la propiedad de suavizamiento

$$\begin{aligned}
\mathbb{E} \left[\frac{Z_n}{m^n} | \mathcal{F}_0 \wedge \mathcal{G} \right] &= \mathbb{E} \left[\mathbb{E} \left[\frac{Z_n}{m^n} | \mathcal{F}_{n-1} \right] | \mathcal{F}_0 \wedge \mathcal{G} \right] \\
&= \mathbb{E} \left[\frac{Z_{n-1}}{m^{n-1}} + \frac{Y_n}{m^n} | \mathcal{F}_0 \wedge \mathcal{G} \right] \\
&= \mathbb{E} \left[\mathbb{E} \left[\frac{Z_{n-1}}{m^{n-1}} + \frac{Y_n}{m^n} | \mathcal{F}_{n-2} \right] | \mathcal{F}_0 \wedge \mathcal{G} \right] \\
&= \mathbb{E} \left[\frac{Z_{n-2}}{m^{n-2}} + \frac{Y_n}{m^n} + \frac{Y_{n-1}}{m^{n-1}} | \mathcal{F}_0 \wedge \mathcal{G} \right] \\
&\vdots \\
&= \mathbb{E} \left[\frac{Z_1}{m^1} + \frac{Y_n}{m^n} + \frac{Y_{n-1}}{m^{n-1}} + \dots + \frac{Y_2}{m^2} | \mathcal{F}_0 \wedge \mathcal{G} \right]
\end{aligned}$$

Además,

$$\mathbb{E} \left[\frac{Z_{n+1}}{m^{n+1}} | \mathcal{F}_n \wedge \mathcal{G} \right] \geq \frac{Z_n}{m^n}, \quad \text{para } n \geq 1.$$

De lo anterior, se tienen las siguientes dos afirmaciones:

- $\mathbb{E} \left[\frac{Z_n}{m^n} | \mathcal{F}_0 \right] = Z_0 + \sum_{k=1}^n \frac{Y_k}{m^k}$ para $n \geq 1$.
- $\{Z_n/m^n\}$ es una submartingala respecto a $\{\mathcal{F}_n \wedge \mathcal{G}\}$.

La primera afirmación se sigue de manera inmediata y para justificar la segunda, falta demostrar que $\{Z_n/m^n\}$ es una v.a. integrable, lo cual se puede ver a partir de

$$\mathbb{E} \left[\frac{Z_n}{m^n} | \mathcal{G} \right] = \mathbb{E} \left[\frac{1}{m^n} \sum_{k=0}^n \sum_{l=1}^{Y_k} \zeta_k^{k,l} | \mathcal{G} \right] = \sum_{k=0}^n \frac{1}{m^k} \mathbb{E} \left[\sum_{l=1}^{Y_k} \frac{\zeta_k^{k,l}}{m^{n-k}} | \mathcal{G} \right]$$

para $k \leq n$ la v.a. $\sum_{l=1}^{Y_k} \zeta_k^{k,l}/m^{n-k}$ es el $n-k$ elemento de una martingala respecto al proceso de Galton-Watson, sin embargo, con Y_k partículas.

Por tanto, se tiene que

$$\mathbb{E} \left[\frac{Z_n}{m^n} | \mathcal{G} \right] = \sum_{k=0}^n \frac{Y_k}{m^k} \leq \sum_{k=0}^{\infty} \frac{Y_k}{m^k}$$

y por el Lema 1.2.4 tenemos que $\mathbb{E} [Z_n/m^n | \mathcal{G}] < \infty$, esto implica por el *Lema de Fatou* que $\liminf Z_n/m^n < \infty$ c.s.

Finalmente, se puede concluir que $\{Z_n/m^n\}_n$ tiene esperanzas acotadas y converge en el conjunto $\{\sup \mathbb{E}[Z_n/m^n | \mathcal{F}_n \wedge \mathcal{G}] < \infty\}$ luego por *Teorema de Convergencia de Martingalas* ésta converge c.s. a una v.a. finita. \square

De nuevo, como en el proceso de G.W. en el caso supercrítico la Ley de Crecimiento Exponencial se conserva, lo cual era de esperarse ya que este caso no tenía el inconveniente de extinción.

De la misma manera que en el caso subcrítico, *González y Molina* en [13] identifican la distribución de la variable asociada al proceso, de nuevo tratando de dar mayor precisión al tamaño de la población.

Capítulo 2

Algunos Métodos de Estimación

En el capítulo anterior, en el proceso G.W.I. definimos a $\{\zeta_n\}$ y $\{Y_n\}$ como el número de descendientes y el número de inmigrantes en la población, respectivamente. Así que en adelante, a las correspondientes distribuciones de estas v.a.'s, nos referiremos como la distribución de reproducción y de inmigración.

Una manera de aproximarnos a analizar la evolución de $\{Z_n\}$ es considerando algunos funcionales de las distribuciones de reproducción e inmigración, tales como m , σ^2 , λ , σ_Y^2 , la probabilidad de extinción y la edad del proceso; sin embargo, en adelante consideramos a m y λ como nuestros parámetros de interés, sin olvidar que una caracterización completa de estas componentes sólo es posible conociendo los vectores de probabilidad que los generan.

Existen varios enfoques para abordar el problema de inferencia estadística de un modelo matemático, debido a los objetivos de este documento, nosotros hemos elegido el enfoque clásico conocido comunmente como *frecuentista*¹, aunque también haremos algunos apartes para el enfoque Bayesiano.

El problema de estimación para el proceso de G.W.I. se ha considerado extensivamente en la literatura. Los pioneros en trabajar en ésta área son *Heyde* y *Seneta* en [21] y [22], más adelante, *Bhat* y *Adke* en [5], *Venkataraman* en [46] también consideraron dicho problema pero a diferencia de los primeros, éstos se concentraron en modelos paramétricos.

Varios estimadores puntuales para los parámetros m y λ fueron propuestos por

¹El hecho de considerar éste enfoque se identifica con la actitud que se debe tomar ante la incertidumbre y ante los problemas de inferencia, cabe notar que aunque éste es el que más se usa en las aplicaciones, esto no indica que sea el único enfoque que exista, ni que sea el correcto e indicado en todos los problemas de inferencia.

Klimko y *Nelson* en [27] y *Wei* y *Winnicky* en [48] bajo diferentes casos; cuando denotamos como \hat{m} y $\hat{\lambda}$ sus correspondientes estimadores, la distribución límite de \hat{m} es diferente drásticamente en los tres casos $m < 1$ (subcrítico), $m = 1$ (crítico) y $m > 1$ (supercrítico), es decir, m juega un papel crucial para el conocimiento de la distribución límite de su estimador.

A continuación, presentamos una recopilación-a manera histórica-de algunos métodos de estimación que tienen como objetivo obtener propiedades de estos estimadores sin considerar la clasificación del proceso de Galton-Watson con Inmigración. Los resultados principales concernientes a cada método se presentan sin demostración, las cuales junto con más detalles relacionados se pueden encontrar en los artículos en los cuales se fundamentan los mismos.

Antes de dar paso a la descripción de los métodos de estimación, hay que tener presente que se puede tener como principio que las distribuciones de reproducción e inmigración son conocidas **(a)**, es decir, contamos con una muestra que proviene de un modelo estadístico específico, o el caso en el que no conocemos de que familia proviene la muestra **(b)**, esto es, de acuerdo al método se parte de alguno o de ambos de los siguientes dos supuestos:

- (a)** En el caso paramétrico (ver [24]), digamos que las distribuciones de reproducción e inmigración pertenecen a la familia de series de potencias².
- (b)** Ahora, en el caso no-paramétrico nos interesaría responder a, si la evolución de una población sigue un proceso de Galton-Watson con Inmigración, pero la distribución de reproducción es desconocida, ¿qué podemos decir acerca del proceso dados los datos de las primeras n generaciones?.

²Las cuales son un caso particular de la familia exponencial e incluyen una variedad de distribuciones como *Poisson*, *Geométrica*, *Binomial*, *Binomial Negativa*,..., ésta familia se caracteriza de la siguiente manera,

$$\pi_k(\theta) = \frac{a_k \theta^k}{a(\theta)} \quad \text{con} \quad a(\theta) = \sum a_k \theta^k \quad \text{para} \quad a_k \theta^k \geq 0, \quad (2.1)$$

a_k es una función de k o una constante y θ garantiza que $a(\theta)$ converja, para $k = 0, 1, \dots$,

$$\rho_j(\varphi) = \frac{b_j \varphi^j}{b(\varphi)} \quad \text{con} \quad b(\varphi) = \sum b_j \varphi^j, \quad \text{y} \quad b_j \varphi^j \geq 0, \quad (2.2)$$

donde b_j es una función de j o una constante y φ garantiza que $b(\varphi)$ converja, $j = 0, 1, \dots$

2.1. Método de Máxima Verosimilitud

Los primeros en investigar acerca del problema de estimación puntual para los parámetros m y λ , vía el método de máxima verosimilitud fueron *Bartlett y Patankar* (ver [4]), sin embargo, como lo puntualizaron *Heyde y Seneta* en [21] las expresiones para dichos estimadores bajo modelos paramétricos eran muy complejas y difíciles de manipular puesto que los primeros calcularon sus estimadores con base en la muestra $\{Z_0, Z_1, \dots, Z_n\}$.

Si se considera que es posible incluir en la muestra el número de inmigrantes en cada generación, la estimación por máxima verosimilitud produce resultados un poco más útiles. A continuación, se presenta un desarrollo con base en el artículo de *Bhat y Adke* en [5] y un resumen acerca de los resultados obtenidos en dicho trabajo.

Las distribuciones de reproducción y de inmigración para las v.a's $\{\zeta_n, n \geq 1\}$ y $\{Y_n, n \geq 1\}$, las denotaremos respectivamente por

$$\pi_k(\theta) = \mathbb{P}[\zeta_n = k], \quad k = 0, 1, \dots \quad \rho_j(\varphi) = \mathbb{P}[Y_n = j], \quad j = 0, 1, \dots \quad (2.3)$$

donde θ y φ son parámetros reales, suponemos que estos parámetros son funcionalmente independientes y que el espacio paramétrico en el que están definidos es compacto y contiene un rectángulo abierto (ver [10]).

Dado que en el contexto de estimación conviene tener información más específica de la muestra, los autores representan la información que poseen de un proceso de G.W.I de la siguiente manera:

El proceso de G.W.I. $\{Z_n, n \geq 0\}$ se reescribe como sigue: $Z_0 = Y_0$ y para $n \geq 1$

$$\begin{aligned} Z_1 &= \zeta_1 + \dots + \zeta_{Y_0} + Y_1 \\ Z_2 &= \zeta_{Y_0+1} + \dots + \zeta_{Y_0+Z_1} + Y_2 \\ &\vdots \\ Z_n &= \zeta_{Y_0+Z_1+\dots+Z_{n-2}+1} + \dots + \zeta_{Y_0+Z_1+\dots+Z_{n-1}} + Y_n \end{aligned}$$

con las modificaciones apropiadas cuando algunos de sus subíndices son 0.

La formulación anterior, implica que la muestra, que cubre las primeras n generaciones, consiste de dos partes:

- La muestra aleatoria $\zeta_1, \zeta_2, \dots, \zeta_{Y_0+Z_1+\dots+Z_{n-1}}$ desde la distribución de reproducción $\{\pi_k(\theta)\}$ la cual es de tamaño aleatorio $N_{n-1} = Y_0 + Z_1 + \dots + Z_{n-1}$.

- La muestra aleatoria Y_1, Y_2, \dots, Y_n de tamaño n fijo desde la distribución $\{\rho_j(\varphi)\}$.

Por lo tanto, la verosimilitud $\mathcal{L}_n(\theta, \varphi)$ de θ y φ basada en la muestra anterior es el producto de

- (i) La verosimilitud correspondiente de una muestra aleatoria de tamaño i desde la distribución de reproducción $\{\pi_k(\theta)\}$ con i reemplazado por N_{n-1} , y
- (ii) La verosimilitud correspondiente a una muestra aleatoria de tamaño n desde la distribución de inmigración $\{\rho_j(\varphi)\}$.

De esto se sigue que la verosimilitud está dada por

$$\mathcal{L}_n(\theta, \varphi) = \mathcal{L}(\theta, \varphi | z_0, z_1, y_1, \dots, z_n, y_n) = \prod_{k=1}^{N_{n-1}} \pi_k(\theta) \prod_{j=1}^n \rho_j(\varphi), \quad (2.4)$$

además, en el caso en que las distribuciones de reproducción y de inmigración pertenecen a la familia de series de potencias, afirmamos que

$$\left\{ \sum_{i=1}^{N_{n-1}} \zeta_i, \sum_{j=1}^n Y_j \text{ y } N_{n-1} \right\}$$

son estadísticos suficientes para los parámetros (θ, φ) , ésto puede verse a partir de tomar dos resultados del experimento $x = \{(z_k, y_k)\}$ y $x' = \{(z'_k, y'_k)\}$ con $k = 0, \dots, n$ y ver que el cociente de las funciones de verosimilitud correspondientes a éstas muestras

$$\frac{\mathcal{L}_n(\theta, \varphi, x)}{\mathcal{L}_n(\theta, \varphi, x')}$$

es independiente de (θ, φ) sí y sólo sí

$$\sum_{i=1}^{N_{n-1}} \zeta_i = \sum_{i=1}^{N_{n-1}} \zeta'_i \text{ y } \sum_{j=1}^n y_j = \sum_{j=1}^n y'_j.$$

Ésta es una condición para la cual x y x' permiten que las respectivas verosimilitudes sean proporcionales en función de (θ, φ) y por el principio de suficiencia (ver [24]), se sigue la afirmación.

Por lo tanto, no es necesario asumir que $\{\zeta'_i s\}$ y $\{Y'_j s\}$ son variables aleatorias observables sino que es suficiente observar Z_0, Z_1, \dots, Z_n y Y_1, Y_2, \dots, Y_n .

Se supone que las distribuciones de reproducción y de inmigración son lo suficientemente suaves para asegurar que bajo el supuesto de i.i.d., los estimadores de

máxima verosimilitud $\hat{\theta}_n, \hat{\varphi}_n$ de θ y φ respectivamente, existen y son asintóticamente normales y eficientes.

Notemos que hay que tener en cuenta que para la estimación vía máxima verosimilitud de θ se tiene un número aleatorio N_{n-1} de observaciones independientes de $\{\pi_k(\theta)\}$ entonces el comportamiento límite, es decir, las condiciones de regularidad se deben verificar cuando n es un tiempo de paro que depende de Z_j 's. Así bajo estas restricciones, los autores utilizan la ley fuerte de los grandes números para un número aleatorio de observaciones y el Teorema Central del límite para Martingalas (ver [1]), para establecer la consistencia y la normalidad asintótica de los estimadores de máxima verosimilitud en los casos subcrítico y supercrítico del proceso G.W.I.

Es decir, a través de éste método demuestran la consistencia de los estimadores en el caso subcrítico y supercrítico³, así mismo como la distribución asintótica de éstos.

Estimadores

Ahora, bajo el supuesto de que las distribuciones $\{\pi_k(\theta)\}, \{\rho_j(\varphi)\}$ de reproducción e inmigración, tienen soportes finitos, presentamos la estimación para las probabilidades $\pi_k(\theta)$ y $\rho_j(\varphi)$ por medio del método de máxima verosimilitud.

Para ello, denotamos por

$$M_j(n) = \#\{i \in \{1, \dots, n\} : Y_i = j\} \quad \text{y} \quad N_k(n) = \sum_{j=1}^n w_j(k), \quad (2.5)$$

con $w_j(k) = \#\{i \in \{1, 2, \dots, Z_{j-1}\} : \zeta_{j-1,i} = k\}$, es decir, $N_k(n)$ representa el número de ancestros en las primeras $n-1$ generaciones que contribuyen con k descendientes a la siguiente generación fuera del total $N_{n-1} = \sum_{j=0}^{n-1} Z_j$ padres disponibles en las primeras $(n-1)$ generaciones y $M_j(n)$ es el número de veces que j inmigrantes se unen al proceso de G.W.I. durante las primeras n generaciones.

Teniendo en cuenta la descripción anterior y los supuestos de independencia de la población, si consideramos a $\delta_n = \{w_j(k) : j = 1, \dots, n, k = 0, \dots, \}$, esto es, $\delta_n = \{W_1, W_2, \dots, W_n\}$ con $W_j = (w_j(0), w_j(1), \dots,)^T$, entonces

$$\mathbb{P}[\delta_n | \theta, \varphi] = \prod_{j=1}^n \mathbb{P}[W_j | w_j, \theta],$$

³En el caso crítico, no pudieron extender los resultados que les permitieran hacer afirmaciones similares ya que el resultado de convergencia de *Seneta*, Teorema 1.2.4, se cumple en ley y no casi seguramente como hubiera sido lo ideal para los autores.

con $\mathbb{P}[W_j|w_j, \theta]$ una distribución multinomial con índice w_j y parámetro θ , por lo tanto la función de verosimilitud para la reproducción, basada en las primeras n generaciones es

$$\begin{aligned}
\mathcal{L}(\pi|\delta_n) &\propto \prod_{j=1}^n \prod_k \pi_k^{w_j(k)} \\
&= \prod_{j=1}^n \left[\pi_0^{w_j(0)} \pi_1^{w_j(1)} \dots \pi_l^{w_j(l)} \dots \right] \\
&= \left[\pi_0^{w_1(0)} \pi_1^{w_1(1)} \dots \pi_l^{w_1(l)} \dots \right] \dots \left[\pi_0^{w_n(0)} \pi_1^{w_n(1)} \dots \pi_l^{w_n(l)} \dots \right] \\
&= \left[\pi_0^{\sum_{j=1}^n w_j(0)} \pi_1^{\sum_{j=1}^n w_j(1)} \dots \pi_l^{\sum_{j=1}^n w_j(l)} \dots \right] \\
&= \prod_k \pi_k^{N_k(n)}
\end{aligned}$$

Por tanto la función de verosimilitud apropiada, cuando $\zeta_1, \dots, \zeta_{Y_0+Z_1+\dots+Z_{n-1}}$ y Y_1, \dots, Y_n son v.a. observables, está dada por

$$\mathcal{L}(\theta, \varphi|\{z_0, (z_{j-1}, y_j), k = 1, \dots, z_{j-1}, j = 1, \dots, n\}) = \prod_{k=0}^{\infty} \pi_k^{N_k(n)} \prod_{j=0}^{\infty} \rho_j^{M_j(n)}. \quad (2.6)$$

Notemos que $N_k(n)$ y $M_j(n)$, $k, j = 0, 1, \dots$, son conjuntamente suficientes si consideramos el espacio paramétrico general donde se satisface

$$\pi_k \geq 0, \quad \sum \pi_k = 1, \quad \rho_j \geq 0 \quad \text{y} \quad \sum \rho_j = 1, \quad (2.7)$$

y observemos también que como una consecuencia de la descripción anterior, afirmamos

$$\sum_{j=0}^{\infty} M_j(n) = n \quad \text{y} \quad \sum_{k=0}^{\infty} N_k(n) = N_{n-1} \quad (2.8)$$

gracias a que los eventos $\{Y_i = j\}$ son disjuntos, así mismo los eventos $\{\zeta_{l-1, i} = k\}$ puesto que

$$\begin{aligned}
\sum_{j=0}^{\infty} M_j(n) &= \sum_{j=0}^{\infty} \sum_{i=1}^n \mathbf{1}_{\{Y_i=j\}} = \sum_{i=1}^n \sum_{j=0}^{\infty} \mathbf{1}_{\{Y_i=j\}} = \sum_{i=1}^n \mathbf{1}_{\{\cup_{j=0}^{\infty} Y_i=j\}} = \sum_{i=1}^n 1 \\
\text{y} \quad \sum_{k=0}^{\infty} N_k(n) &= \sum_{k=0}^{\infty} \sum_{l=1}^n w_l(k) = \sum_{k=0}^{\infty} \sum_{l=1}^n \sum_{i=1}^{Z_{l-1}} \mathbf{1}_{\{\zeta_{l-1, i}=k\}} = \sum_{l=1}^n \sum_{i=1}^{Z_{l-1}} \sum_{k=0}^{\infty} \mathbf{1}_{\{\zeta_{l-1, i}=k\}},
\end{aligned}$$

donde

$$\sum_{l=1}^n \sum_{i=1}^{Z_{l-1}} \mathbf{1}_{\{\cup_{k=0}^{\infty} \zeta_{l-1,i}=k\}} = \sum_{l=1}^n Z_{l-1}$$

Lema 2.1.1. *Los estimadores vía máxima verosimilitud $\hat{\pi}_k$ y $\hat{\rho}_j$ de π_k y ρ_j , respectivamente están dados por*

$$\hat{\pi}_k = \frac{N_k(n)}{N_{n-1}} \quad \hat{\rho}_j = \frac{M_j(n)}{n} \quad \text{para } k, j = 0, 1, \dots$$

Demostración. Al tomar el logaritmo en la ecuación (2.6) tenemos que

$$\log(\mathcal{L}) = \sum_{k=0}^{\infty} N_k(n) \log(\pi_k) + \sum_{j=0}^{\infty} M_j(n) \log(\rho_j)$$

y usando *Multiplicadores de Lagrange*, tenemos que el siguiente sistema de ecuaciones

$$\frac{d}{d\pi_k} \left[\sum_{k=0}^{\infty} N_k(n) \log(\pi_k) + \sum_{j=0}^{\infty} M_j(n) \log(\rho_j) \right] = \nu \left[\sum_{k=0}^{\infty} \pi_k - 1 \right] + \iota \left[\sum_{j=0}^{\infty} \rho_j - 1 \right] \quad (2.9)$$

$$\frac{d}{d\rho_j} \left[\sum_{k=0}^{\infty} N_k(n) \log(\pi_k) + \sum_{j=0}^{\infty} M_j(n) \log(\rho_j) \right] = \nu \left[\sum_{k=0}^{\infty} \pi_k - 1 \right] + \iota \left[\sum_{j=0}^{\infty} \rho_j - 1 \right] \quad (2.10)$$

$$\sum_k \pi_k = 1 \quad \text{y} \quad \sum_j \rho_j = 1 \quad (2.11)$$

luego, para k, j fijos, al derivar e igualar a cero, tenemos que

$$N_k(n) \left(\frac{1}{\pi_k} \right) = \nu \quad M_j(n) \left(\frac{1}{\rho_j} \right) = \iota, \quad (2.12)$$

y usando las ecuaciones (2.7), (2.8) y (2.11), resulta que $\nu = N_{n-1}$, $\iota = n$. Finalmente, sustituyendo en la ecuación (2.12) se prueba el resultado. \square

Corolario 2.1.1. *Si la media de reproducción $m = \sum_{k=0}^{\infty} k\pi_k$ y la media de inmigración $\lambda = \sum_{j=0}^{\infty} j\rho_j$ son finitas, entonces sus estimadores vía máxima verosimilitud \hat{m} y $\hat{\lambda}$, respectivamente, están dados por*

$$\hat{m} = \sum_{k=0}^{\infty} k\hat{\pi}_k \quad \text{y} \quad \hat{\lambda} = \sum_{j=0}^{\infty} j\hat{\rho}_j \quad (2.13)$$

Demostración. Debido a la propiedad de invarianza funcional de la función de verosimilitud, se sigue el resultado (ver [24]). \square

Por otro lado, notemos que

$$\begin{aligned}\hat{m} &= \frac{1}{N_{n-1}} \sum_{k=0}^{\infty} k N_k(n) = \frac{1}{N_{n-1}} \sum_{k=1}^{\infty} k \sum_{i=1}^n w_i(k) \\ &= \frac{1}{N_{n-1}} \sum_{i=1}^n \sum_{l=1}^{Z_{i-1}} \sum_{k=1}^{\infty} k \mathbf{1}_{\{\zeta_{i-1,l}=k\}} = \frac{1}{N_{n-1}} \sum_{i=1}^n \sum_{l=1}^{Z_{i-1}} \zeta_{i-1,l}\end{aligned}$$

así

$$\hat{m} = \frac{1}{N_{n-1}} \sum_{i=1}^n (Z_i - Y_i)$$

y

$$\begin{aligned}\hat{\lambda} &= \frac{1}{n} \sum_{j=0}^{\infty} j M_j(n) = \frac{1}{n} \sum_{j=1}^{\infty} j \sum_{i=1}^n \mathbf{1}_{\{Y_i=j\}} \\ &= \frac{1}{n} \sum_{i=1}^n \left[\sum_{j=1}^{\infty} j \mathbf{1}_{\{Y_i=j\}} \right] = \frac{1}{n} \sum_{i=1}^n Y_i\end{aligned}$$

En conclusión, tenemos que dada la muestra $\{Z_0, Z_1, Y_1, \dots, Z_n, Y_n\}$ correspondiente a un proceso de G.W.I, los estimadores de máxima verosimilitud para las medias de reproducción e inmigración están dados respectivamente por

$$\hat{m}_n = \frac{1}{\sum_{i=1}^n Z_{i-1}} \sum_{i=1}^n (Z_i - Y_i) \quad \text{y} \quad \hat{\lambda}_n = \frac{1}{n} \sum_{i=1}^n Y_i.$$

2.2. Método de Mínimos Cuadrados

Inicialmente, *Heyde* y *Seneta* en [21] y [22], fueron los primeros en obtener resultados de estimación para m y λ sin imponer una distribución específica a las variables aleatorias $\{\zeta\}$ e $\{Y\}$.

Para el caso $m > 1$, *Heyde* y *Seneta* en [20] y *Heyde* en [19] mostraron que para estimar m , se pueden usar las siguientes expresiones, respectivamente,

$$\frac{Z_n}{Z_{n-1}} \quad \text{y} \quad \frac{\sum_{i=1}^n Z_i}{\sum_{i=1}^n Z_{i-1}}.$$

Para el caso $m < 1$ estos autores, con el fin de encontrar estimadores fuertemente consistentes y distribuidos asintóticamente Normal, usaron la analogía entre los procesos de ramificación con inmigración y los procesos autoregresivos de orden 1.

Por este camino, ellos encontraron que dichos estimadores son asintóticamente equivalentes a los estimadores de mínimos cuadrados, los cuales habían sido obtenidos por *Klimko* y *Nelson* en [27] y están fundamentados en lo siguiente:

La ecuación (1.3) se reescribe como

$$Z_n = mZ_{n-1} + \lambda + \epsilon_n \quad (2.14)$$

donde $\epsilon_n = Z_n - mZ_{n-1} - \lambda$ es una sucesión de diferencia de martingala con respecto a la σ -álgebra generada por $\{Z_0, \dots, Z_n\}$ (ver [37]), además la ecuación (2.14) es una ecuación de regresión estocástica y a partir de ésta ellos demuestran que los estimadores condicionales de mínimos cuadrados de m y λ resultantes son

$$\tilde{m}_n = \left\{ \sum_{i=1}^n Z_i \sum_{i=1}^n Z_{i-1} - n \sum_{i=1}^n Z_i Z_{i-1} \right\} \left\{ \left[\sum_{i=1}^n Z_{i-1} \right]^2 - n \sum_{i=1}^n Z_{i-1}^2 \right\}^{-1},$$

$$\tilde{\lambda}_n = \left\{ \sum_{i=1}^n Z_{i-1} Z_i \sum_{i=1}^n Z_{i-1} - \sum_{i=1}^n Z_{i-1}^2 \sum_{i=1}^n Z_i \right\} \left\{ \left[\sum_{i=1}^n Z_{i-1} \right]^2 - n \sum_{i=1}^n Z_{i-1}^2 \right\}^{-1}$$

En el caso subcrítico, las propiedades asintóticas de estos estimadores las trabajó *Venkataraman* en [46].

Estos estimadores, no cumplieron con la condición de ser consistentes cuando no se conoce la clasificación del proceso, así en busca de una teoría unificada de acuerdo a esto, *Wei* y *Winnicky* en [48] consideraron los mismos estimadores para el caso crítico y supercrítico bajo el supuesto de que los segundos momentos de las v.a. $\{\zeta\}$ y $\{Y\}$ son finitos, con lo cual tampoco lograron que fueran consistentes. Así que de nuevo, continuó la pregunta abierta, ¿cuándo existen otros estimadores que sean siempre consistentes?

Nuevamente, *Wei* y *Winnicky* preocupados por unificar esta estimación respecto a la propiedad de consistencia en [49] y con base en la idea de mínimos cuadrados ponderados dada inicialmente por *Nelson* en [34], proponen estimadores tales que, el estimador \hat{m}_n es consistente en todos los casos y que asintóticamente en el caso $m \neq 1$, la ley límite es Normal y si $m = 1$, la distribución asintótica se puede expresar en términos de la distribución límite de $W(t)$, donde $\{W(t)\}$ es un proceso de difusión no-negativo con generador $Af(x) = \lambda f^{(1)}(x) + \frac{1}{2}x\sigma^2 f^{(2)}(x)$ para $f \in C^\infty[0, \infty]$, que se obtiene como límite débil del proceso $\frac{1}{n}Z_{[nt]}$ cuando $n \rightarrow \infty$ (ver Apéndice).

Estimadores

De nuevo, reescriben la ecuación (1.3) como

$$\frac{Z_n}{(Z_{n-1} + 1)^{1/2}} = m(Z_{n-1} + 1)^{1/2} + (\lambda - m)(Z_{n-1} + 1)^{-1/2} + \delta_n \quad (2.15)$$

donde $\delta_n = \epsilon_n / (Z_{n-1} + 1)^{1/2}$.

A partir de esta expresión, *Wei* y *Winnicky* estudiaron los estimadores de mínimos cuadrados, a los cuales llamaron *estimadores de mínimos cuadrados ponderados* y están dados por

$$\tilde{m}_n = \left\{ \sum_{i=1}^n Z_i \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n \sum_{i=1}^n \frac{Z_i}{Z_{i-1} + 1} \right\} \left\{ \sum_{i=1}^n (Z_{i-1} + 1) \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n^2 \right\}^{-1} \quad (2.16)$$

$$\tilde{\lambda}_n = \left\{ \sum_{i=1}^n Z_{i-1} \sum_{i=1}^n \frac{Z_i}{Z_{i-1} + 1} - \sum_{i=1}^n Z_i \sum_{i=1}^n \frac{Z_{i-1}}{Z_{i-1} + 1} \right\} \left\{ \sum_{i=1}^n (Z_{i-1} + 1) \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n^2 \right\}^{-1} \quad (2.17)$$

aquí, $\tilde{\lambda}_n = \tilde{m}_n + \tilde{p}_n$, siendo \tilde{p}_n el estimador de mínimos cuadrados de $p = \lambda - m$, en la ecuación (2.15)

De acuerdo a la clasificación del proceso de Galton-Watson con Inmigración, en dicho artículo se presentan las demostraciones de las siguientes propiedades asintóticas:

En el caso subcrítico, recordemos que bajo el supuesto de que $\mathbb{E}[\log^+ Y_n] < \infty$, $\{Z_n\}$ tiene una única distribución estacionaria, además, si la distribución de Z es la distribución de $\{Z_n\}$, entonces $\{Z\}$ es estacionaria y ergódica, por tanto, podemos asumir que $\{Z_n\}$ es estacionaria y ergódica (ver Capítulo 1).

- Si $m < 1$, entonces \tilde{m}_n y $\tilde{\lambda}_n$ son fuertemente consistentes y

$$\left(\left[\sum_{i=1}^n (Z_{i-1} + 1) \right]^{1/2} (\tilde{m}_n - m), \left[\sum_{i=1}^n (Z_{i-1} + 1)^{-1} \right]^{1/2} (\tilde{\lambda}_n - \lambda) \right)^T$$

converge en distribución a $\mathcal{N}(\mathbf{0}, V^{-1} W V^T)$, donde Q^T denota la transpuesta de una matriz Q y

$$V = \begin{pmatrix} \frac{\mathbb{E}[Z]}{(\mathbb{E}[Z+1])^{1/2}} & (\mathbb{E}[\frac{1}{Z+1}])^{-1/2} \\ \frac{\mathbb{E}[Z/Z+1]}{(\mathbb{E}[Z+1])^{1/2}} & (\mathbb{E}[\frac{1}{Z+1}])^{1/2} \end{pmatrix}.$$

$$W = \begin{pmatrix} \mathbb{E}[\sigma^2 Z + \sigma_Y^2] & \mathbb{E}[\frac{\sigma^2 Z + \sigma_Y^2}{Z+1}] \\ \mathbb{E}[\frac{\sigma^2 Z + \sigma_Y^2}{Z+1}] & \mathbb{E}[\frac{\sigma^2 Z + \sigma_Y^2}{(Z+1)^2}] \end{pmatrix}.$$

siendo Z una variable aleatoria con la distribución estacionaria del proceso $\{Z_n\}$.

- Si $m > 1$, entonces $\tilde{m}_n \xrightarrow{c.s.} m$, mientras $\tilde{\lambda}_n$ es débilmente consistente, más aún,

$$\left(\sum_{i=1}^n (Z_{i-1} + 1) \right)^{1/2} (\tilde{m}_n - m) \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

- Si $m = 1$, entonces $\tilde{m}_n \xrightarrow{\mathbb{P}} m$ y $\tilde{\lambda}_n \xrightarrow{\mathbb{P}} \lambda$, además

$$\left(\sum_{i=1}^n (Z_{i-1} + 1) \right)^{1/2} (\tilde{m}_n - m) \xrightarrow{d} \frac{1}{\left(\int_0^1 W(t) dt \right)^{1/2}} (W(1) - \lambda).$$

Más aún, bajo el supuesto de que $2\lambda > \sigma^2$, $\mathbb{E}[|\zeta|^{2+\delta}] < \infty$ y $\mathbb{E}[|Y|^{2+\delta}] < \infty$ para algún $\delta > 0$, se tiene que

$$\left(\sum_{i=1}^n (Z_{i-1} + 1)^{-1} \right)^{1/2} (\tilde{\lambda}_n - \lambda) \xrightarrow{d} \mathcal{N}(0, \sigma^2).$$

2.3. Estimación Bootstrap

Supongamos que tenemos una muestra $\{(Z_i, Y_i), i = 1, \dots, n\}$, y recordemos que los estimadores vía máxima verosimilitud son

$$\hat{m}_n = \left(\sum_{i=1}^n Z_{i-1} \right)^{-1} \sum_{i=1}^n (Z_i - Y_i)$$

y

$$\hat{\lambda}_n = n^{-1} \sum_{i=1}^n Y_i.$$

Para hacer inferencia acerca de los parámetros m y λ , se puede considerar la cantidad pivotal dada por

$$V_n = \left(\sum_{i=1}^n Z_{i-1} \right)^{1/2} (\hat{m}_n - m),$$

sin embargo, la forma de la distribución de V_n depende de m y específicamente, *Sriram* en [43] ($m < 1$, $m = 1$) y *Wei* y *Winnicky* en [49] ($m > 1$) obtuvieron que cuando $n \rightarrow \infty$

$$V_n \xrightarrow{d} \begin{cases} N(0, \sigma^2) & \text{si } m \neq 1, \\ \frac{W(1)-\lambda}{(\int_0^1 W(t)dt)^{1/2}} & \text{si } m = 1, \end{cases}$$

donde $\{W(t)\}$ es un proceso de difusión no-negativo con generador dado por $Af(x) = \lambda f^{(1)}(x) + \frac{1}{2}x\sigma^2 f^{(2)}(x)$, para $f \in C^\infty[0, \infty]$ que se obtiene como límite débil del proceso $\frac{1}{n}Z_{[nt]}$ cuando $n \rightarrow \infty$.

Con el fin de obtener la distribución muestral de V_n , *Sriram* en [44] consideró un Bootstrap paramétrico asumiendo que las distribuciones de reproducción y de inmigración pertenecen a la familia de series de potencias y demostró que la aproximación no es válida para el caso crítico.

En este artículo (ver [44]), proponen una modificación al Bootstrap paramétrico donde el procedimiento modificado provee una aproximación a la distribución de V_n no sólo para el caso $m = 1$ sino también para $m \neq 1$, además el resultado se muestra para una clase general de distribuciones discretas y no necesariamente para la clase de series de potencias.

Procedimiento Bootstrap

Para el propósito del Bootstrap paramétrico, suponemos que las v.a.'s de reproducción e inmigración tienen funciones de densidad de probabilidad

$$\pi_k(\theta) = \mathbb{P}[\zeta_n = k] \quad \text{y} \quad \rho_j(\varphi) = \mathbb{P}[Y_n = j] \quad \text{para } k, j = 0, 1, \dots,$$

dependiendo de los parámetros θ y φ , así $m = \mathbb{E}_\theta[\zeta] = f(\theta)$ y $\lambda = \mathbb{E}_\varphi[Y] = g(\varphi)$ para algunas funciones f, g las cuales son un mapeo uno a uno de los respectivos espacios a \mathbb{R}^+ .

Un procedimiento Bootstrap paramétrico para aproximar la distribución muestral de V_n se puede describir como sigue:

Dada una muestra $\chi_n = \{(Z_i, Y_i), i = 1, \dots, n\}$, primero se estima la media de reproducción m por algún estimador \tilde{m}_n , basado en χ_n y la media de inmigración λ por $\tilde{\lambda}_n$. Luego, se reemplaza θ y φ por sus respectivos estimadores $\tilde{\theta}_n = f^{-1}(\tilde{m}_n)$ y $\tilde{\varphi}_n = g^{-1}(\tilde{\lambda}_n)$, así condicional en χ_n , $\{\zeta_{i,j}^*\}$ es una sucesión de v.a.'s con f.d.p. $\pi(\tilde{\theta}_n)$ y $\{Y_i^*\}$ es una sucesión de v.a.'s con f.d.p. $\rho(\tilde{\varphi}_n)$.

La muestra Bootstrap $\chi_n^* = \{(Z_i^*, Y_i^*), i = 1, \dots, n\}$ nos permite obtener

$$\begin{aligned}\hat{m}_n^* &= \left(\sum_{i=1}^n Z_{i-1}^* \right)^{-1} \sum_{i=1}^n (Z_i^* - Y_i^*), \\ \hat{\lambda}_n^* &= n^{-1} \sum_{i=1}^n Y_i^*, \\ V_n^* &= \left(\sum_{i=1}^n Z_{i-1}^* \right)^{1/2} (\hat{m}_n^* - m),\end{aligned}$$

donde la distribución (condicional) de V_n^* (dada la muestral original χ_n) constituye una aproximación Bootstrap a la distribución de V_n .

Sin embargo, bajo el supuesto de que las distribuciones pertenecen a la familia de series de potencias, Sriram en [44], mostró que para el caso $m = 1$, la distribución límite condicional del pivotal Bootstrap V_n^* no coincide con la distribución límite de V_n . En otras palabras, la validéz asintótica no se cumple para el Bootstrap paramétrico estándar. Un análisis profundo muestra que la principal razón de ésta falla es que cuando $m = 1$, la distribución estimada $\pi(f^{-1}(\hat{m}_n))$ no es una buena elección para la población Bootstrap, ya que \hat{m}_n converge a m muy lentamente.

Por tal razón, a continuación presentamos un breve resumen de los resultados obtenidos por *Sriram* y *Datta* en [8], con dicha propuesta se busca evitar el problema anterior.

Estimador

En [8], los autores proponen un estimador modificado \tilde{m}_n , el cual converge suficientemente rápido a m , la selección de este estimador modificado se hace con base en el estimador de *Hodges* (ver [31]), dicho trabajo se fundamentó en un adaptive shrinkage (datos-dependientes):

Para \hat{m}_n definimos

$$\tilde{m}_n = \begin{cases} 1 - \eta_n & \text{si } 1 - \delta_n \leq \hat{m}_n \leq 1 + \eta_n \\ 1 + \eta_n & \text{si } 1 + \eta_n \leq \hat{m}_n \leq 1 + \delta_n \\ \hat{m}_n & \text{e.o.c.} \end{cases} \quad (2.18)$$

donde $\{\eta_n\}$ es una sucesión no-aleatoria tal que $n\eta_n \rightarrow 0$ cuando $n \rightarrow \infty$. Además, $\delta_n = h(\sum_{i=1}^n Z_{i-1})$ siendo h una función positiva en $[0, \infty)$ tal que se satis-

facen $\lim_{z \rightarrow \infty} \left(\frac{z}{\log \log z} \right)^{1/2} h(z) = \infty$ y $\lim_{z \rightarrow \infty} h(z) = 0$ (por ejemplo, $h(z) = z^{-1/3}$).

La construcción de \tilde{m}_n garantiza que éste se encuentra entre η_n a una distancia 1 cuando \hat{m}_n está entre δ_n a una distancia 1 (una indicación de que el verdadero valor m es 1 o muy cercano a 1).

El resultado principal es que con esta selección de \tilde{m}_n se satisface que la distribución condicional del pivotal Bootstrap V_n^* es asintóticamente cercana a la distribución del pivotal V_n , pero antes de pasar a enunciar éste, necesitamos los siguientes supuestos:

- (i) Las varianzas σ^2 y σ_Y^2 son funciones continuas de θ y φ , respectivamente.
- (ii) La función $\theta \mapsto \mathbb{E}_\theta[|\zeta|^{2+\delta}]$ es continua en θ para algún $\delta > 0$.
- (iii) Para cualquier sucesión $\{\theta_n\} \in \Theta$, tal que $\theta_n \rightarrow \theta$

$$\limsup_{n \rightarrow \infty} |\theta_n - \theta|^{-1} \sum_{i=0}^{\infty} |\gamma_{\theta_n}(i) - \gamma_\theta(i)| < \infty,$$

donde $\gamma_\theta(i) = \sum_{j=0}^i \pi_j(\theta)$ y para cualquier sucesión $\{\varphi_n\} \in \Phi$ tal que $\lim_{n \rightarrow \infty} \varphi_n = \varphi$ y

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{\infty} |s_{\varphi_n}(i) - s_\varphi(i)| = 0,$$

donde $s_\varphi(i) = \sum_{j=0}^i \rho_j(\varphi)$.

Teorema 2.3.1. *Si se cumple (i) y (ii)*

$$\sup_x |\mathbb{P}[V_n \leq x] - \mathbb{P}^*[V_n^* \leq x]| \longrightarrow 0, \quad \text{cuando } n \rightarrow \infty \quad (2.19)$$

casi seguramente cuando $m \leq 1$. Además, cuando $m > 1$ la convergencia de la ecuación (2.19) se cumple bajo las condiciones (i) y (iii).

La demostración de este teorema, se basa en una serie de resultados relacionados con arreglos de procesos de ramificación, además de diferentes técnicas de acuerdo a la clasificación del proceso.

Propiedades del estimador

- Para $m < 1$, $\tilde{m}_n \xrightarrow{c.s.} m$ cuando $n \rightarrow \infty$.
- Para $m > 1$, $n(\tilde{m}_n - m) \xrightarrow{c.s.} 0$ cuando $n \rightarrow \infty$.
- Para $m = 1$, si $\mathbb{E}[|\zeta|^{2+s}] < \infty$ para algún $s > 0$ entonces $n(\tilde{m}_n - m) \xrightarrow{c.s.} 0$ cuando $n \rightarrow \infty$.

2.4. Otro enfoque: Estimación Bayesiana

En esta sección, nos salimos del enfoque frecuentista y damos paso a un enfoque diferente como el análisis Bayesiano. Los siguientes resultados, se encuentran en el artículo de *González, Molina y Mota* (ver [14]).

Recordemos que desde el punto de vista Bayesiano, nuestros parámetros de interés son vistos como variables aleatorias, lo cual implica que para utilizar la regla de Bayes necesitamos una distribución a priori para los parámetros del modelo. El objetivo de los autores en este artículo fue proponer un análisis conjugado para hacer inferencia en el modelo G.W.I., ellos plantean su objetivo desde la situación paramétrica y no-paramétrica y encuentran estimadores Bayesianos bajo la función de pérdida mínima para las medias y las varianzas de las distribuciones de reproducción e inmigración. A continuación presentamos un bosquejo de los resultados obtenidos en dicho artículo:

Situación paramétrica

Consideremos el modelo del proceso de Galton-Watson con Inmigración con distribuciones de reproducción y de inmigración independientes. Ambas pertenecientes a la familia de series de potencias.

Dentro de las propiedades, de esta familia tenemos que

$$m = \mathbb{E}_\theta[\zeta] = \theta \frac{d}{d\theta} \log(a(\theta)), \quad \lambda = \mathbb{E}_\varphi[Y] = \varphi \frac{d}{d\varphi} \log(b(\varphi)). \quad (2.20)$$

La verosimilitud cuando se cuenta con la muestra $Z_0, \{Z_i, Y_i, i = 1, 2, \dots, \}$ está dada por

$$\begin{aligned} \mathcal{L}(\theta, \varphi | z_0, z_1, y_1, \dots, z_n, y_n) &= \prod_{j=0}^n \mathbb{P}[Z_j = z_j] \\ &= \prod_{j=0}^n \sum_{y_j} \mathbb{P} \left[\sum_{i=1}^{z_{j-1}} \zeta_{j-1,i} = z_j - y_j \right] \mathbb{P}[Y_j = y_j] \\ &\propto b_{z_0} \varphi^{z_0} b(\varphi)^{-1} \prod_{j=1}^n \theta^{z_j - y_j} a(\theta)^{-z_{j-1}} b_{y_j} \varphi^{y_j} b(\varphi)^{-1} \\ &\propto \theta^{\sum_{j=1}^n (z_j - y_j)} \varphi^{z_0 + \sum_{j=1}^n y_j} a(\theta)^{-\sum_{j=1}^n z_{j-1}} b(\varphi)^{-(n+1)} \end{aligned}$$

Teniendo en cuenta la forma de la distribución, una clase razonable de distribuciones a priori conjugada para (θ, φ) es

$$f(\theta, \varphi) = \theta^{\alpha_1} \varphi^{\alpha_2} a(\theta)^{-\beta_1} b(\varphi)^{-\beta_2} [\omega(\alpha_1, \alpha_2, \beta_1, \beta_2)]^{-1}, \quad (2.21)$$

donde $\alpha_1, \alpha_2, \beta_1, \beta_2$ son constantes tales que

$$\omega(\alpha_1, \alpha_2, \beta_1, \beta_2) = C(\alpha_1, \beta_1)C(\alpha_2, \beta_2) < \infty,$$

siendo

$$C_1(\alpha_1, \beta_1) = \int_{\theta} \theta^{\alpha_1} a(\theta)^{-\beta_1} d\theta \quad \text{y} \quad C_2(\alpha_2, \beta_2) = \int_{\varphi} \varphi^{\alpha_2} b(\varphi)^{-\beta_2} d\varphi.$$

Luego, la distribución a posteriori es

$$f[(\theta, \varphi) | z_0, z_1, y_1, \dots, z_n, y_n] = \theta^{\alpha'_1} \varphi^{\alpha'_2} a(\theta)^{-\beta'_1} b(\varphi)^{-\beta'_2} [\omega(\alpha'_1, \alpha'_2, \beta'_1, \beta'_2)]^{-1},$$

con $\alpha'_1 = \alpha_1 + \sum_{j=1}^n (z_j - y_j)$, $\alpha'_2 = \alpha_2 + z_0 + \sum_{j=1}^n y_j$, $\beta'_1 = \beta_1 + \sum_{j=0}^{n-1} z_j$ y $\beta'_2 = \beta_2 + (n+1)$.

Si se denota por $Z^*_n = \sum_{j=0}^n Z_j$ para $n = 0, 1, \dots$, y $Y^*_n = \sum_{j=1}^n Y_j$ para $n = 1, 2, \dots$ y $\mathcal{F}_n = \sigma\{Z_0, Z_1, Y_1, \dots, Z_n, Y_n\}$ entonces

$$f[(\theta, \varphi) | \mathcal{F}_n] = \theta^{\alpha^*_{*1}} \varphi^{\alpha^*_{*2}} a(\theta)^{-\beta^*_{*1}} b(\varphi)^{-\beta^*_{*2}} [\omega(\alpha^*_{*1}, \alpha^*_{*2}, \beta^*_{*1}, \beta^*_{*2})]^{-1},$$

siendo $\alpha^*_{*1} = \alpha_1 + Z^*_n - (Z_0 + Y^*_n)$, $\alpha^*_{*2} = \alpha_2 + Z_0 + Y^*_n$, $\beta^*_{*1} = \beta_1 + Z^*_{n-1}$ y $\beta^*_{*2} = \beta_2 + (n+1)$.

Finalmente, los autores demuestran que los estimadores de Bayes, bajo la función de pérdida mínima y considerando la clase conjugada dada en la ecuación (2.21), para la media de las distribuciones de reproducción e inmigración están dados por:

$$\hat{m} = \sum_k ka_k \frac{C_1(\alpha^*_1 + k, \beta^*_1 + 1)}{C_1(\alpha^*_1, \beta^*_1)} \quad \text{y} \quad \hat{\lambda} = \sum_j jb_j \frac{C_2(\alpha^*_2 + j, \beta^*_2 + 1)}{C_2(\alpha^*_2, \beta^*_2)} \quad (2.22)$$

Situación, no-paramétrica

En el caso no-paramétrico, la función de verosimilitud está dada por

$$\mathcal{L}[(\pi, \varphi) | z_0, (z_{j-1,i}, y_j), i = 1, \dots, z_{j-1}, j = 1, 2, \dots, n] \propto \prod_k \prod_j \pi_k^{N_k(n)} \rho_j^{M_n(j)},$$

con $M_j(n) = \#\{i \in \{0, 1, \dots, n\} : Y_i = j\}$, $Z_0 = Y_0$ y $N_k(n) = \sum_{j=1}^n w_j(k)$ siendo $w_j(k) = \#\{i \in \{1, 2, \dots, Z_{j-1}\} : \zeta_{j-1,i} = k\}$.

En éste caso, la distribución a priori según la propuesta conjugada es

$$f(\pi, \rho) = D(\alpha)D(\beta) \prod_k \prod_j \pi_k^{\alpha_k - 1} \rho_j^{\beta_j - 1},$$

con $\alpha = (\alpha_k)$ y $\beta = (\beta_j)$, $\alpha_k > 0$, $\beta_j > 0$, $D(\alpha) = \Gamma(\alpha_*)[\prod_k \Gamma(\alpha_k)]^{-1}$ y $D(\beta) = \Gamma(\beta_*)[\prod_j \Gamma(\beta_j)]^{-1}$, $\alpha_* = \sum_k \alpha_k$, $\beta_* = \sum_j \beta_j$ así π y ρ tienen distribuciones de Dirichlet independientes con vectores de parámetros α , β respectivamente.

Si denotamos por $\mathcal{F}'_n = \sigma\{Z_0, (\zeta_{j-1,i}, Y_j), i = 1, \dots, Z_{j-1}, j = 1, \dots, n\}$, $n \geq 1$ se tiene que la distribución a posteriori es

$$f(\pi, \rho | \mathcal{F}'_n) = D(\gamma)D(\delta) \prod_k \prod_j \pi_k^{\gamma_k-1} \rho_j^{\delta_j-1},$$

siendo $\gamma = \{\gamma_k\}$, $\delta = \{\delta_j\}$ siendo $\gamma_k = \alpha_k + N_k(n)$, $\delta_j = \beta_j + M_j(n)$.

Así, los estimadores de Bayes no-paramétricos, bajo la función de pérdida mínima y el análisis conjugado, para las medias de reproducción y de inmigración, están dados por

$$\hat{m} = \frac{1}{\gamma^*} \sum_k k \gamma_k = \frac{\hat{\alpha} + Z_n^* - (Z_0 + Y_n^*)}{\alpha_* + Z_{n-1}^*} \quad \text{y} \quad \hat{\lambda} = \frac{1}{\delta^*} \sum_j j \delta_j = \frac{\hat{\beta} + Z_0 + Y_n^*}{\alpha_* + n + 1}. \quad (2.23)$$

con $\hat{\alpha} = \sum_k k \alpha_k$ y $\hat{\beta} = \sum_j j \beta_j$.

Hay que tener en cuenta que para implementar este método de estimación no se necesita conocer la clasificación del proceso, sin embargo, si hay que tener presente las dificultades computacionales que se pueden dar, así mismo, otro aspecto a considerar es la selección de la distribución a priori cuando no se tiene información del fenómeno que nos pueda servir para proponer la misma. Varios métodos para derivar a priori no-informativas han sido propuestos y se pueden consultar en un texto clásico de Estadística Bayesiana, por ejemplo, *Bernardo* (1979) y *Berger y Bernardo* (1992).

2.5. Aproximación *Data cloning*

En vista de la complejidad computacional que pueden presentar los problemas de inferencia desde el enfoque Bayesiano y de la subjetividad presente desde la selección de la distribución a priori, algunos autores han tratado de evitar dichas restricciones en modelos dinámicos complejos. A continuación presentamos un resumen de los resultados obtenidos por *Lele, Dennis y Lutscher* en [32].

En este artículo, los autores introducen un nuevo método computacional, conocido como *Data cloning*, el cual sirve para calcular los estimadores de máxima verosimilitud y sus errores estándar para modelos ecológicos complejos. Desde este contexto,

la inferencia es completamente invariante respecto a la elección de las distribuciones a priori y por lo tanto evita la subjetividad inherente de la aproximación Bayesiana. *Data cloning* es útil para analizar situaciones ecológicas en las cuales la modelación adecuada es a través de modelos estadísticos anidados, tales como modelos espacio-estado y modelos de efectos fijos. En este caso, lo usamos para el proceso de Galton-Watson con Inmigración, el cual aunque no es un modelo anidado, sin embargo, esta aproximación nos permite hacer una buena estimación de manera unificada para la distribución límite de la media reproducción y de la media de inmigración.

Método

El método funciona de la siguiente manera: Primero se construye el modelo Bayesiano del problema, completando con una distribución a priori para los parámetros desconocidos, además, se usa la verosimilitud correspondiente para K copias (*clones*) de los datos, donde K es suficientemente grande y las copias se asumen independientes la una de la otra. La distribución a posteriori se calcula con una aproximación MCMC usual. La media de la distribución posterior es igual al estimador de máxima verosimilitud, y K veces la varianza de la a posteriori es igual a la varianza asintótica del estimador de máxima verosimilitud.

La formulación Bayesiana y los algoritmos MCMC pueden ser redireccionados para el método *Data cloning* para producir los estimadores de máxima verosimilitud, sus errores estándar e intervalos de confianza para modelos ecológicos complejos.

Técnicamente, lo anterior se realiza en los siguientes pasos:

1. Se crean K conjuntos de datos clonados $W^K = (W, W, \dots, W)$ donde el vector de datos observados se repite K veces.
2. Usando un algoritmo MCMC, se generan B números aleatorios desde la distribución a posteriori, donde virtualmente cualquier distribución a priori se puede seleccionar.
3. Se calcula los valores medios y las varianzas muestrales de estos $(\theta, \varphi)_j$ para $j = 1, \dots, B$ generados desde la marginal a posteriori.

La distribución a posteriori para los K clones converge a una distribución Normal multivariada con media igual a los estimadores de máxima verosimilitud de los parámetros de interés y varianza igual a $\frac{1}{K}I^{-1}$, siendo I la correspondiente matriz de información de Fisher, la demostración de este resultado aún no ha sido publicada por *S.R. Lele*.

Suponemos que tenemos las observaciones $\{Z_0, Z_1, Y_1, \dots, Z_n, Y_n\}$ y denotamos como $W_n = \{(Z_i, Y_i)\}_i$, que provienen de un proceso de Galton-Watson con inmigración:

Sea $\Psi(\theta, \varphi)$ la distribución a priori en el espacio paramétrico Θ , asumimos que la distribución a posteriori es positiva sobre el espacio paramétrico entero. La distribución a posteriori para (θ, φ) está dada por

$$\Psi^{(1)}(\theta, \varphi) \propto \mathcal{L}(\theta, \varphi; W_n)\Psi(\theta, \varphi),$$

luego sustituimos la distribución a posteriori anterior por la distribución a priori para obtener

$$\Psi^{(2)}(\theta, \varphi) \propto \mathcal{L}(\theta, \varphi; W_n)\Psi^{(1)}(\theta, \varphi)$$

así obtenemos

$$\Psi^{(2)}(\theta, \varphi) \propto \mathcal{L}^{(2)}(\theta, \varphi; W_n)\Psi(\theta, \varphi)$$

y continuando de la misma manera, tenemos que

$$\Psi^{(K)}(\theta, \varphi) \propto \mathcal{L}^{(K)}(\theta, \varphi; W_n)\Psi(\theta, \varphi).$$

Ahora, sean $(\hat{\theta}, \hat{\varphi})$ tales que $\mathcal{L}(\hat{\theta}, \hat{\varphi}; W_n) > \mathcal{L}(\theta, \varphi; W_n)$ para todo (θ, φ) . Por definición, éste es el estimador de máxima verosimilitud para (θ, φ) y como $\Psi(\theta, \varphi)$ es positiva en todo el espacio paramétrico, se sigue que

$$\frac{\Psi^K(\theta, \varphi; W_n)}{\Psi^K(\hat{\theta}, \hat{\varphi}; W_n)} \rightarrow \begin{cases} 0 & \text{si } (\theta, \varphi) \neq (\hat{\theta}, \hat{\varphi}), \\ 1 & \text{si } (\theta, \varphi) = (\hat{\theta}, \hat{\varphi}). \end{cases}$$

Notemos que la verosimilitud de los datos clonados tiene la misma localización del máximo, conocidos como los estimadores de máxima verosimilitud $(\hat{\theta}, \hat{\varphi})$ como la verosimilitud original.

Aunque de nuevo, como método de estimación, éste también tiene ciertas restricciones, por ejemplo, tomar K suficientemente grande lo cual implica que hay que tener cuidado con las conclusiones en el momento de la implementación.

Sin embargo, con un uso debido del método, éste permite tener un aproximación muy útil ya que evita uno de los grandes paradigmas de la Estadística Bayesiana: ¿Cómo proponer la distribución a priori cuando no se tiene información del fenómeno en cuestión?.

En la mayoría de métodos vistos hasta aquí, vemos que debemos conocer la clasificación del proceso para encontrar la distribución límite para \hat{m}_n o tener en cuenta otro tipo de restricciones, lo cual no siempre es cierto en las aplicaciones, excepto si se considera un estudio de simulación, como es nuestro caso. A continuación, presentamos una manera de hacer inferencia mejorando este tipo de dificultades.

Capítulo 3

Estimación Secuencial del Proceso de G.W.I.

3.1. Introducción

En este capítulo, el resultado principal es una estimación unificada y una forma de hacer inferencia para la media de reproducción, obtenida por *Qi* y *Reeves* en [36]. A diferencia de los trabajos de otros autores, este desarrollo no necesita tener algún conocimiento a priori acerca del proceso, es decir, no hay distinción si el proceso es subcrítico, crítico o supercrítico, lo cual es de gran utilidad en la práctica. Dicho resultado, se obtiene a partir de una estimación secuencial para este parámetro.

Los experimentos secuenciales difieren de los otros procedimientos estadísticos en que el tamaño de muestra no está fijo sino que incrementa. El experimentador tiene la opción de mirar la sucesión de observaciones (o un número fijo) a un tiempo determinado y decidir si parar el muestreo y tomar una decisión ó continuar el muestreo y tomar la decisión más adelante. Un hecho esencial de un procedimiento secuencial¹ es que el número de observaciones requerido para terminar el experimento es una variable aleatoria pues depende del resultado de las mismas, por eso cuando se implementan éste tipo de experimentos, se debe especificar:

- El tamaño de muestra inicial.
- Una regla para terminar el experimento.
- El número adicional de observaciones a tomar si el experimento es necesario continuarlo.

¹Los procedimientos secuenciales son muy interesantes por el hecho de que son económicos, en el sentido de que se puede rechazar una decisión tomada a la ligera vía un procedimiento secuencial que vía un procedimiento con tamaño de muestra fijo.

- Una regla de decisión terminal, si es necesaria.

Qi y *Reeves* en su artículo presentan la construcción de este estimador secuencial (de dos etapas), el cual se fundamenta en un cierto tiempo de paro y en una clase específica de funciones que dependen de dicha regla, y demuestran que este estimador es fuertemente consistente y asintóticamente Normal para $m > 0$, la idea de explorar por esta vía surge del trabajo de *Sriram* y *Shete* en [42], los cuales se basaron en un tratamiento similar para los procesos autoregresivos de orden 1, resuelto por *Lai* y *Siegmund* en [29].

Antes de iniciar el marco teórico necesario para nuestro desarrollo, introduzcamos una notación que nos será útil en el momento de trabajar con algunos resultados concernientes a caminatas aleatorias.

Para cualquier entero t no-negativo definimos

$$S_1(t) = Z'_0(t) = t \quad \text{y} \quad Z'_1(t) = \sum_{i=1}^{Z'_0(t)} \zeta_i + Y_1$$

y recursivamente para $n \geq 2$, definimos

$$S_n(t) = \sum_{i=1}^n Z'_{i-1}(t) \quad \text{y} \quad Z'_n(t) = \sum_{i=S_{n-1}(t)+1}^{S_n(t)} \zeta_i + Y_n.$$

Notemos que $\{Z_{n-1}, Y_n, n \geq 1\}$ es igual en distribución a $\{Z'_{n-1}(Z_0), Y_n, n \geq 1\}$ así resulta que el estimador de máxima verosimilitud de m (ver Capítulo 2), se puede reescribir como

$$\hat{m}_n = \frac{1}{\sum_{i=1}^n Z'_{i-1}} \left\{ \sum_{i=1}^n (Z'_i - Y_i) \right\}. \quad (3.1)$$

En adelante, Z'_i y S_n denotarán $Z'_i(Z_0)$ y $S_n(Z_0)$ respectivamente, así que continuamos suponiendo que el tamaño de la población inicial es Z_0 .

3.2. Estimador secuencial de una etapa

Como mencionamos anteriormente para la construcción del estimador secuencial de dos etapas es necesario iniciar con algunos resultados para un caso particular obtenidos por *Sriram* (ver [43]); en este artículo él mostró que el estimador propuesto se distribuye asintóticamente Normal para cada m fijo, $m \in (0, 1]$.

Debido a que *Sriram* fundamentó su trabajo en la estimación secuencial, en lugar de usar un número fijo de observaciones, definió la siguiente regla de paro:

$$N_c = \inf \left\{ n \geq 1 : \sum_{i=1}^n Z_{i-1} \geq c\sigma^2 \right\}, \quad \text{donde } c > 0 \text{ se elige apropiadamente.} \quad (3.2)$$

De manera análoga al estimador de máxima verosimilitud, el estimador secuencial de una etapa (*esue*) para m lo definimos como \hat{m}_{N_c} , es decir,

$$\hat{m}_{N_c} = \frac{1}{\sum_{i=1}^{N_c} Z'_{i-1}} \left\{ \sum_{i=1}^{N_c} (Z'_i - Y_i) \right\}, \quad (3.3)$$

su nombre se debe a que N_c es un tiempo de paro como lo veremos a continuación.

A partir de la notación dada anteriormente, podemos ver que la ecuación (3.2) es equivalente a

$$N_c = \begin{cases} \inf\{n \geq 1 : S_n \geq c\sigma^2\} \\ +\infty \end{cases} \quad \text{si } \{n \geq 1 : S_n \geq c\sigma^2\} = \phi. \quad (3.4)$$

Observemos que, bajo el supuesto de que σ^2 es un parámetro conocido, N_c está bien definido ya que en el caso no-vacío es subconjunto de \mathbb{N} y por tanto ésta acotado inferiormente. Además, notemos que la variable aleatoria N_c , toma valores en $\{1, 2, \dots\} \cup \{\infty\}$ y es un tiempo de paro respecto a $\mathcal{F}_n = \{Z'_0, \dots, Z'_{n-1}\}$ ya que para $B = [c\sigma^2, \infty)$ de la ecuación (3.4) tenemos

$$\{N_c = n\} = \{S_n \in B\} \cap \{S_k \notin B, k < n\}$$

y como $\{S_n \in B\} \in F_n$ y $\{S_k \in B\} \in F_{k-1} \subseteq F_n$, se tiene que $\{N_c = n\} \in \mathcal{F}_n$.

Observación 3.2.1. En el caso en que el parámetro σ^2 es desconocido, dicho parámetro lo absorbe la constante c y nuevamente definimos la regla para terminar el experimento como

$$N(c) = \inf \left\{ n \geq 1 : \sum_{i=1}^n Z_{i-1} \geq c \right\}, \quad (3.5)$$

de nuevo, usamos $\hat{m}_{N(c)}$ como el estimador secuencial para m .

En cuanto a este tiempo de paro, en la siguiente subsección, veremos algunas propiedades en relación a las caminatas aleatorias y a su comportamiento asintótico.

3.2.1. Propiedades

- (A) $S_{N_c-1} < c\sigma^2 \leq S_{N_c}$: Observemos que por definición $c\sigma^2 \leq S_{N_c}$, y partir de la definición de ínfimo como éste es la máxima cota inferior se satisface la otra desigualdad.
- (B) $S_n \xrightarrow{c.s.} \infty$: La razón de esta propiedad es,

$$S_n(Z_0) = \sum_{i=1}^n Z'_{i-1}(Z_0) = \sum_{i=1}^n \sum_{j=S_{i-2}(Z_0)+1}^{S_{i-1}(Z_0)} \zeta_j + \sum_{i=1}^n Y_{i-1},$$

dado que Y_0, ζ_i son v.a. no-negativas, resulta que $S_n \geq \sum_{i=1}^{n-1} Y_i$ y aplicando la *Ley Fuerte de los Grandes Números* tenemos que $S_n \rightarrow \infty$, cuando $n \rightarrow \infty$.

A partir de lo anterior, $\mathbb{P}[N_c < \infty] = 1$ y $N_c \xrightarrow{c.s.} \infty$ cuando $c \rightarrow \infty$.

- (C) Para $m \leq 1$, $S_{N_c}/c\sigma^2 \xrightarrow{c.s.} 1$ cuando $c \rightarrow \infty$: Si dividimos por S_{N_c-1} , la desigualdad dada en la propiedad (A), obtenemos que

$$1 < c\sigma^2 S_{N_c-1}^{-1} \leq 1 + \frac{Z_{N_c-1}}{\sum_{i=1}^{N_c-1} Z_{i-1}},$$

y como $N_c \xrightarrow{c.s.} \infty$ para $m \leq 1$ se cumple por Teorema A.1.1 (ver Apéndice), que

$$\frac{Z_{N_c-1}}{\sum_{i=1}^{N_c-1} Z_{i-1}} \xrightarrow{c.s.} 0, \quad \text{cuando } c \rightarrow \infty.$$

- (D) Si $m = 1$, N_c/\sqrt{c} converge en distribución, ésta propiedad se cumple a partir del Teorema A.3.2 (ver Apéndice).

Antes de dar paso a otras propiedades fundamentales para nuestro estudio, veamos el siguiente teorema, el cuál nos servirá para hacer una aproximación al orden de crecimiento de nuestro proceso en el caso crítico.

Teorema 3.2.1. Sean $m = 1$, $0 < \sigma^2 < \infty$ y $\lambda < \infty$, entonces para cualquier sucesión $\{a_n\}$ creciente y positiva tal que $\sum_{n=1}^{\infty} na_n^{-2} < \infty$ se tiene que $\lim_{n \rightarrow \infty} Z_n a_n^{-1} = 0$ c.s.

Demostración. Notemos que en la ecuación (1.3), al sustituir $\{Z_i\}$ para $1 \leq i \leq n-1$,

se obtiene

$$\begin{aligned}
Z_n &= Z_0 - Z_0 + Z_1 - Z_1 + \dots + Z_{n-1} - Z_{n-1} + \sum_{k=1}^{Z_{n-1}} \zeta_{n-1,k} + Y_n \\
&= Z_0 - Z_0 + \left(\sum_{k=1}^{Z_0} \zeta_{0,k} + Y_1 \right) - Z_1 + \dots + \left(\sum_{k=1}^{Z_{n-2}} \zeta_{n-2,k} + Y_{n-1} \right) - Z_{n-1} + \sum_{k=1}^{Z_{n-1}} \zeta_{n-1,k} + Y_n \\
&= Z_0 + \left(\sum_{k=1}^{Z_0} \zeta_{0,k} - Z_0 \right) + \dots + \left(\sum_{k=1}^{Z_{n-2}} \zeta_{n-2,k} - Z_{n-2} \right) + \left(\sum_{k=1}^{Z_{n-1}} \zeta_{n-1,k} - Z_{n-1} \right) + \sum_{k=1}^n Y_k
\end{aligned}$$

Por tanto, la ecuación (1.3) se puede reescribir como

$$Z_n = Z_0 + \sum_{i=1}^n W_i + \sum_{i=1}^n Y_i \quad \text{donde} \quad W_i = \sum_{k=1}^{Z_{i-1}} (\zeta_{i-1,k} - 1). \quad (3.6)$$

Como Y_1, Y_2, \dots , son v.a.i.i.d. con $\mathbb{E}[Y_i] = \lambda < \infty$, por la *Ley Fuerte de los Grandes Números* se satisface que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n Y_i = \lambda, \quad c.s.$$

y por hipótesis $\sum_{n=1}^{\infty} na_n^{-2} < \infty$, lo cual implica que $\lim_{n \rightarrow \infty} na_n^{-1} = 0^2$, luego obtenemos

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \sum_{i=1}^n Y_i = 0, \quad c.s. \quad (3.7)$$

Ahora, como $Z_0 = Y_0$ y $a_n \uparrow \infty$ entonces

$$\lim_{n \rightarrow \infty} \frac{Z_0}{a_n} = 0. \quad (3.8)$$

Para finalizar, basta ver

$$\lim_{n \rightarrow \infty} \frac{1}{a_n} \sum_{i=1}^n W_i = 0 \quad c.s. \quad (3.9)$$

Definamos por $H_k = \sum_{n=1}^k W_n^2 a_n^{-2}$ donde $\{H_k\}_k$ es una submartingala respecto a la filtración natural \mathcal{F}_k , ya que

$$\mathbb{E}[H_{k+1} | \mathcal{F}_k] = \mathbb{E}[H_k + W_{k+1}^2 a_{k+1}^{-2} | \mathcal{F}_k] \leq H_k$$

²Esto es consecuencia de denotar, para $n \geq 1$ $b_n = n/a_n^2$, en el siguiente resultado, *Criterio de Abel* (ver [33]): Si $\sum b_n < \infty$ siendo $\{b_n\}$ una sucesión de términos positivos y decrecientes entonces $\lim_{n \rightarrow \infty} nb_n = 0$.

además, afirmamos que

$$\sup_k \mathbb{E}[H_k^+] = \sum_{n=1}^{\infty} \mathbb{E}[W_n^2] a_n^{-2} = \sum_{n=1}^{\infty} \sigma^2 n \lambda a_n^{-2} < \infty, \quad (3.10)$$

dicha afirmación la demostraremos más adelante, y aplicando el *Teorema de Convergencia de Martingalas*, $\lim_{k \rightarrow \infty} H_k$ existe y es finito c.s., por tanto, $\sum_{n=1}^{\infty} W_n^2 a_n^{-2} < \infty$ c.s. de donde $\sum_{n=1}^{\infty} W_n a_n^{-1} < \infty$ c.s. y por el *Lema de Kronecker* (ver [37]), la ecuación (3.9) es válida.

Finalmente, de las ecuaciones (3.7), (3.8) y (3.9) se obtiene el resultado.

Veamos ahora que, por la independencia de $\{\zeta_{n-1,i}\}$, la ecuación (3.10) es cierta:

$$\begin{aligned} \mathbb{E}[W_n^2] &= \mathbb{E}[\mathbb{E}[W_n^2 | F_{n-1}]] \\ &= \sum_k \mathbb{E}[\mathbb{E}[W_n^2 | Z_{n-1} = k]] \mathbb{P}[Z_{n-1} = k] \\ &= \sum_k \mathbb{E} \left[\mathbb{E} \left[\left(\sum_{i=1}^{Z_{n-1}} (\zeta_{n-1,i} - 1) \right)^2 \mid Z_{n-1} = k \right] \right] \mathbb{P}[Z_{n-1} = k] \\ &= \sum_k k \sigma^2 \mathbb{P}[Z_{n-1} = k] \\ &= \sigma^2 n \lambda. \end{aligned}$$

□

Corolario 3.2.2. *Sea $m = 1$ entonces*

$$\lim_{n \rightarrow \infty} \frac{Z'_n}{n \log n} = 0.$$

Demostración. Para $n \geq 1$, sea $a_n = n \log n$ donde $\sum_{n=2}^{\infty} (n \log^2 n)^{-1}$ converge³, luego de aplicar el Teorema 3.2.1 a dicha sucesión, tenemos que $\lim_{n \rightarrow \infty} Z'_n (n \log n)^{-1} = 0$ c.s. □

Es importante notar que el comportamiento de N_c difiere en los casos subcrítico, crítico y supercrítico del proceso G.W.I. como se puede ver en el siguiente teorema.

³La convergencia de esta serie se sigue por el criterio de la Integral.

Teorema 3.2.3. Para N_c definido en (3.2), se tienen las siguientes afirmaciones.

(1) Si $m \in (0, 1)$, entonces

$$\lim_{c \rightarrow \infty} \frac{N_c}{c} = \frac{(1-m)\sigma^2}{\lambda} \quad c.s.$$

(2) Si $m = 1$, entonces

$$\liminf_{c \rightarrow \infty} \frac{N_c}{\sqrt{c/\log c}} = \infty, \quad \limsup_{c \rightarrow \infty} \frac{N_c}{c} < \infty \quad c.s.$$

(3) Si $m > 1$, entonces

$$\lim_{c \rightarrow \infty} \frac{N_c}{\log c} = \frac{1}{\log m} \quad c.s.$$

Demostración. Para demostrar (1) sabemos, a partir de la propiedad de ergodicidad del proceso (ver Capítulo 1) que

$$\frac{S_n}{n} \xrightarrow{c.s.} \frac{\lambda}{1-m} \quad \text{para todo } n \geq 1,$$

también, por la propiedad (C) tenemos $S_{N_c}/c\sigma^2 \xrightarrow{c.s.} 1$, cuando $c \rightarrow \infty$ y puesto que

$$\frac{N_c}{c} = \frac{N_c}{S_{N_c}} \frac{S_{N_c}}{c},$$

se sigue el resultado.

Ahora, para demostrar la primera afirmación de (2), por la definición de S_n , como Y_0 y ζ_i son v.a. no-negativas entonces $S_n \geq \sum_{i=1}^{n-1} Y_i$, además, $\mathbb{E}[Y_i] = \lambda < \infty$ así por la *Ley Fuerte de los Grandes Números* se sigue que

$$\liminf_{n \rightarrow \infty} \frac{S_n}{n} \geq \lambda \quad c.s. \text{ para todo } n \geq 1,$$

equivalentemente,

$$\liminf_{c \rightarrow \infty} \frac{S_{N_c-1}}{N_c-1} \geq \lambda \quad c.s.$$

De nuevo, por la propiedad (A) del *esue*,

$$\liminf_{c \rightarrow \infty} \frac{c\sigma^2}{N_c-1} \geq \lambda, \quad c.s.$$

y tomando el recíproco tenemos,

$$\limsup_{c \rightarrow \infty} \frac{N_c}{c} \leq \frac{\sigma^2}{\lambda} < \infty, \quad \text{c.s.}$$

Para la segunda afirmación de **(2)** por el Corolario 3.2.2 para todo $n \geq 1$ tenemos que $Z'_n (n \log n)^{-1} \xrightarrow{\text{c.s.}} 0$ cuando $n \rightarrow \infty$, y por propiedades de la notación $o(\cdot)$, tenemos que $S_n = o(\max_{n \geq 1} n \log n)$ por tanto, se sigue que $S_n = o(n^2 \log n)$ y como $S_n \xrightarrow{\text{c.s.}} \infty$ cuando $n \rightarrow \infty$, así

$$\frac{N_c^2 \log N_c}{c} \xrightarrow{\text{c.s.}} \infty, \quad \text{cuando } c \rightarrow \infty$$

y entonces resulta que

$$\frac{N_c \sqrt{\log c}}{c^{1/2}} \xrightarrow{\text{c.s.}} \infty.$$

Finalmente, para demostrar **(3)**, notemos que

$$\frac{S_n}{m^n} = \frac{1}{m^n} \left[\sum_{i=1}^n Z'_{i-1} \right] = \frac{1}{m^n} \left[\sum_{i=1}^n m^{i-1} \frac{Z'_{i-1}}{m^{i-1}} \right] \quad (3.11)$$

como en este caso, $m^{-1} < 1$ por el Teorema 1.2.4

$$\frac{Z'_n}{m^n} \xrightarrow{\text{c.s.}} W, \quad \text{siendo } W \text{ una variable aleatoria positiva,}$$

de donde en la ecuación (3.11)

$$\frac{S_n}{m^n} \xrightarrow{\text{c.s.}} \frac{W}{m-1}, \quad \text{cuando } n \rightarrow \infty. \quad (3.12)$$

Recordemos que $S_{N_c-1} < c\sigma^2 \leq S_{N_c}$, luego de tomar la función $\log(\cdot)$ y dividir por $\log(c)$ obtenemos

$$\frac{\log(S_{N_c-1})}{\log(c)} < 1 + \frac{\log(\sigma^2)}{\log(c)} \leq \frac{\log(S_{N_c})}{\log(c)}$$

lo que implica que

$$\limsup_{c \rightarrow \infty} \frac{\log(S_{N_c-1})}{\log(c)} \leq 1 \quad \text{y} \quad \liminf_{c \rightarrow \infty} \frac{\log(S_{N_c})}{\log(c)} \geq 1,$$

equivalentemente, tenemos que

$$\limsup_{c \rightarrow \infty} \frac{\log\left(\frac{S_{N_c-1}}{m^{N_c-1}} m^{N_c-1}\right)}{\log(c)} \leq 1 \quad \text{y} \quad \liminf_{c \rightarrow \infty} \frac{\log\left(\frac{S_{N_c}}{m^{N_c}} m^{N_c}\right)}{\log(c)} \geq 1,$$

así finalmente después de aplicar la ecuación (3.12) se obtiene que

$$\limsup_{c \rightarrow \infty} \frac{(N_c - 1) \log(m)}{\log(c)} \leq 1 \quad \text{y} \quad \liminf_{c \rightarrow \infty} \frac{N_c \log(m)}{\log(c)} \geq 1$$

y por tanto se sigue el resultado. \square

3.3. Estimador secuencial de dos etapas

Como una extensión al estimador de una etapa *esue*, definimos $G(c) = G(c, N_c) \geq N_c$ una función con valores en los enteros que satisface:

$$\lim_{c \rightarrow \infty} \frac{G(c, N_c) - N_c}{c^\alpha} = 0 \quad \text{c.s. para algún } \alpha \in (0, 1/2) \text{ si } m \in (0, 1] \quad (3.13)$$

$$\lim_{c \rightarrow \infty} \{G(c, N_c) - N_c\} = \infty \quad \text{c.s. si } m > 1. \quad (3.14)$$

Observemos que $G(c)$ es una función que depende de dos tiempos de paro, esto implica que la muestra no está fija, es decir, el tamaño de la segunda muestra depende del resultado de la primera, esta función nos permite seleccionar el tamaño de muestra óptimo⁴, lo cual contribuye a una reducción de costos en la realización del experimento. Este tipo de procedimiento, fué introducido por *Stein* (ver [45]), y su planteamiento se ha modificado en múltiples contextos (ver [28]).

Ahora, con la misma forma del estimador de máxima verosimilitud definimos el estimador secuencial de dos etapas por \hat{m}_G , esto es,

$$\hat{m}_G = \frac{1}{\sum_{i=1}^G Z'_{i-1}} \left\{ \sum_{i=1}^G (Z'_i - Y_i) \right\}.$$

Las condiciones (3.13) y (3.14) las satisfacen una extensa clase de funciones, sin embargo, nos interesan las enunciadas a continuación:

Sea $\kappa(c)$ cualquier función con valores en los enteros que cumple

$$\lim_{c \rightarrow \infty} \kappa(c) = \infty \quad \text{y} \quad \lim_{c \rightarrow \infty} \frac{\kappa(c)}{c^\alpha} = 0 \quad \text{para algún } \alpha \in (0, 1/2), \quad (3.15)$$

observemos que si $G(c) = \kappa(c) + N_c$ entonces las ecuaciones (3.13) y (3.14) se cumplen automáticamente.

Por otro lado, si definimos $G(c) = G(c, N_c) = \text{máx}\{\kappa(c), N_c\}$, esto es,

$$G(c) = \inf \left\{ n \geq \kappa(c) : \sum_{i=1}^n Z_{i-1} \geq c\sigma^2 \right\}$$

donde $\kappa(c)$ además de las condiciones dadas en la ecuación (3.15) satisface que

$$\lim_{c \rightarrow \infty} \frac{\kappa(c)}{\log c} = \infty. \quad (3.16)$$

⁴Desde el punto de vista teórico y en términos del diseño del muestreo

Notemos que a partir del Teorema 3.2.3 y las condiciones (3.15) y (3.16) obtenemos:

(i) Para $m < 1$ y $\alpha \in (0, 1/2)$,

$$\lim_{c \rightarrow \infty} \frac{N_c}{\kappa(c)} = \lim_{c \rightarrow \infty} \frac{N_c}{c^\alpha} \frac{c^\alpha}{\kappa(c)} = \lim_{c \rightarrow \infty} \frac{N_c}{c^\alpha} \lim_{c \rightarrow \infty} \frac{c^\alpha}{\kappa(c)} = \infty, \quad \text{c.s.}$$

(ii) Para $m = 1$,

$$\liminf_{c \rightarrow \infty} \frac{N_c}{\kappa(c)} = \liminf_{c \rightarrow \infty} \frac{N_c}{\sqrt{c/\log(c)}} \frac{\sqrt{c/\log(c)}}{\kappa(c)} = \infty, \quad \text{c.s.}$$

(iii) Para $m > 1$,

$$\lim_{c \rightarrow \infty} \frac{N_c}{\kappa(c)} = \lim_{c \rightarrow \infty} \frac{N_c}{\log(c)} \frac{\log(c)}{\kappa(c)} = \lim_{c \rightarrow \infty} \frac{N_c}{\log(c)} \lim_{c \rightarrow \infty} \frac{\log(c)}{\kappa(c)} = 0. \quad \text{c.s.}$$

De lo anterior, podemos concluir:

1. Para $m \leq 1$ y para algún $\alpha \in (0, 1/2)$, $\lim_{c \rightarrow \infty} N_c/\kappa(c) = \infty$ c.s., así N_c crece más rápido que $\kappa(c)$ y como $G(c) = \max\{\kappa(c), N_c\}$ entonces $G(c) = N_c$ c.s..
2. Para $m > 1$, $\lim_{c \rightarrow \infty} N_c/\kappa(c) = 0$ c.s., lo cual implica que $G(c) = \kappa(c)$ c.s.

Por lo tanto, con $G(c) = \max\{\kappa(c), N_c\}$, las condiciones (3.13) y (3.14), también se cumplen. La relevancia de ésta forma particular de la función $G(c)$, está en que nos permite reducir el tamaño de muestra adecuado ya que nos implica que posiblemente con al menos $\kappa(c)$ se puede aplicar nuestro procedimiento y obtener los resultados deseados.

De la misma manera, observemos que en el caso en que $\kappa(c) = c$, tenemos que $G(c) = \max\{\kappa(c), N_c\}$ es nuevamente un tiempo de paro, lo cual nos dice que para seleccionar el tamaño de muestra adecuado, necesitamos que dicho tamaño sea al menos c .

Observación 3.3.1. Notemos que los resultados vistos en las secciones anteriores para el *esue* también se cumplen para $N(c)$ definido en la ecuación (3.5).

3.4. Comportamiento asintótico

Si de la misma manera que el estimador de máxima verosimilitud, definimos como $\hat{\lambda}_G = \frac{1}{G} \sum_{i=1}^G Y_i$ al estimador secuencial de dos etapas para la media de inmigración y bajo los supuestos que tenemos acerca de los momentos del proceso, podemos concluir a partir del *Teorema Central del Límite* que cuando $G \rightarrow \infty$,

$$\sqrt{G} \left(\frac{\hat{\lambda}_G - \lambda}{\sigma_Y} \right) \xrightarrow{d} \mathcal{N}(0, 1).$$

Por otro lado, una vez definido el estimador secuencial de dos etapas \hat{m}_G es importante estudiar su comportamiento asintótico ya que esto nos permite hacer inferencia estadística sobre la media de reproducción, específicamente nos interesa la distribución límite de su estimador puntual, lo cual se logra con el siguiente teorema y a diferencia de la distribución límite de $\hat{\lambda}_G$, la demostración del siguiente resultado utiliza herramientas un poco más complejas.

Teorema 3.4.1. *El estimador \hat{m}_G converge casi seguramente a m , cuando c tiende a ∞ . Más aún, si $G(c) = G(c, N_c)$ satisface las ecuaciones (3.13) y (3.14), entonces*

$$\left(\sum_{i=1}^G Z_{i-1} \right)^{1/2} \frac{\hat{m}_G - m}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1) \quad (3.17)$$

Para su demostración, usaremos dos resultados técnicos que nos permitirán llevar a cabo la justificación del resultado principal de este capítulo.

Lema 3.4.1. *Para $c > 0$, las sucesiones $\{\zeta_{N_c+j}, j \geq 1\}$ y (S_{N_c}, Z'_{N_c}, N_c) son independientes, más aún, $\{\zeta_{N_c+j}, j \geq 1\}$ son idénticamente distribuidas con la misma ley que ζ .*

Demostración. Por la definición de S_{N_c} y una propiedad de tiempo de paro (ver [37]), se garantiza la independencia; además por la construcción de $\{\zeta\}$ (ver Capítulo 1) se sigue la igualdad en distribución. \square

Por otro lado, el siguiente teorema (ver [6]-Teorema 17.1-) nos permite hacer una generalización del *Teorema Central del Límite* teniendo en cuenta una parametrización dependiendo del tiempo; inicialmente, introduzcamos una notación necesaria para el entendimiento del mismo.

Sean $S_n = \tau_1 + \dots + \tau_n$ para cada n y v_n una variable aleatoria no-negativa que toma valores en los enteros la cual está definida en el mismo espacio de probabilidad de $\{\tau_n\}$, y definamos

$$X_n(t, w) = \frac{1}{\sigma\sqrt{n}} S_{[nt]}(w) \quad \text{y} \quad Y_n(t, w) = \frac{1}{\sigma\sqrt{v_n(w)}} S_{[v_n(w)t]}(w).$$

Teorema 3.4.2. Si $v_n/a_n \xrightarrow{\mathbb{P}} \theta$, donde θ es una constante positiva y a_n son constantes tendiendo a ∞ , entonces $X_n \xrightarrow{d} W$ implica que $Y_n \xrightarrow{d} W$.

En nuestro caso, es indispensable, tomar la misma idea del Teorema 3.4.2 y extenderla a una función de las sucesiones anteriores como lo veremos en el siguiente lema.

Lema 3.4.2. Sean $\{\eta_j, j \geq 1\}$ una sucesión de v.a.i.i.d. con $E[\eta_1] = 0$ y $E[\eta_1^2] = 1$ y H_n es una sucesión de vectores aleatorios. Para cada $n \geq 1$, $\{\eta_j\}$ y H_n son independientes. Si v_n y u_n son dos sucesiones de variables aleatorias con valores en los enteros tales que v_n es medible respecto a $\sigma(H_n)$ y $u_n = \tau_n(H_n, \eta_j, j \geq 1)$ con $u_n/v_n \xrightarrow{\mathbb{P}} 1$ entonces

$$\mathbb{P} \left[\frac{\sum_{j=1}^{u_n} \eta_j}{\sqrt{u_n}} \leq x | H_n \right] \xrightarrow{\mathbb{P}} \Phi(x), \quad \forall x \in \mathbb{R}, \quad (3.18)$$

donde $\Phi(\cdot)$ denota la función de distribución Normal estándar.

Demostración. En un espacio de probabilidad grande, podemos definir una sucesión de v.a. $\{H'_n\}$ independientes de $\{\eta_j\}$ para cada $n \geq 1$ donde $\{H_n\}$ y $\{H'_n\}$ son iguales en distribución, si definimos $u'_n = \tau_n(H'_n, \eta_j, j \geq 1)$, tenemos que

$$\mathbb{P} \left[\sum_{j=1}^{u'_n} \eta_j / \sqrt{u'_n} \leq x | H'_n \right] \stackrel{d}{=} \mathbb{P} \left[\sum_{j=1}^{u_n} \eta_j / \sqrt{u_n} \leq x | H_n \right]. \quad (3.19)$$

Sin pérdida de generalidad, podemos suponer que los H'_n s son independientes.

Sea $\mathcal{H} = \sigma(H_n, n \geq 1)$, entonces $\{\eta_j\}$ son independientes de \mathcal{H} , como $\mathbb{E}[\eta_1] = 0$, $\mathbb{E}[\eta_1^2] = 1$ por el *Teorema Central del Límite* tenemos que

$$\mathbb{P} \left[\frac{\sum_{j=1}^n \eta_j}{\sqrt{n}} \leq x | \mathcal{H} \right] \xrightarrow{c.s.} \Phi(x), \quad \forall x \in \mathbb{R}. \quad (3.20)$$

Para demostrar la ecuación (3.18), es suficiente probar que dada una sucesión de enteros, digamos $\{n'\}$, existe una subsucesión para la cuál (3.18) se cumple⁵:

Como $u_n/v_n \xrightarrow{\mathbb{P}} 1$, existe $\{\epsilon_n\}$ tal que

$$\lim_{n \rightarrow \infty} \mathbb{P} \left[\left| \frac{u_n}{v_n} - 1 \right| > \epsilon_n \right] = 0 \quad \text{donde} \quad \lim_{n \rightarrow \infty} \epsilon_n = 0,$$

⁵Recordemos que $X_n \xrightarrow{\mathbb{P}} X$ sí y sólo sí cada subsucesión $\{X_{n_k}\}$ contiene una subsucesión $\{X_{n_{k(i)}}\}$ la cual converge casi seguramente a X (ver [37]).

por lo tanto,

$$\mathbb{P} \left[\left| \frac{u_n}{v_n} - 1 \right| > \epsilon_n | \mathcal{H} \right] \xrightarrow{\mathbb{P}} 0$$

es decir, existe una subsucesión $\{n''\}$ de $\{n'\}$ para la cual

$$\mathbb{P} \left[\left| \frac{u_{n''}}{v_{n''}} - 1 \right| > \epsilon_{n''} | \mathcal{H} \right] \xrightarrow{c.s.} 0$$

lo cual implica que condicional en \mathcal{H} , $u_{n''}/v_{n''} \xrightarrow{\mathbb{P}} 1$, como v_n es \mathcal{H} -medible y se cumple (3.20), por Teorema 3.4.2 tenemos que

$$\mathbb{P} \left[\frac{\sum_{j=1}^{u_{n''}} \eta_j}{\sqrt{u_{n''}}} \leq x | \mathcal{H} \right] \xrightarrow{c.s.} \Phi(x), \quad \forall x \in \mathbb{R},$$

además

$$\mathbb{P} \left[\sum_{j=1}^{u_n} \eta_j / \sqrt{u_n} \leq x | H_n \right] = \mathbb{P} \left[\sum_{j=1}^{u_n} \eta_j / \sqrt{u_n} \leq x | \mathcal{H} \right], \quad c.s.$$

y de esto se sigue que

$$\mathbb{P} \left[\frac{\sum_{j=1}^{u_{n''}} \eta_j}{\sqrt{u_{n''}}} \leq x | H_n \right] \xrightarrow{\mathbb{P}} \Phi(x), \quad \forall x \in \mathbb{R}.$$

□

Ahora, veamos que podemos escribir \hat{m}_G de la siguiente manera:
Notemos que para todo $n \geq 1$,

$$\begin{aligned} \hat{m}_n - m &= \frac{1}{\sum_{i=1}^n Z_{i-1}} \left\{ \sum_{i=1}^n (Z'_i - Y_i) \right\} - m \\ &= \frac{1}{S_n} \left\{ \sum_{i=1}^n (Z'_i - Y_i) - m S_n \right\} \\ &= \frac{1}{S_n} \left\{ \sum_{i=1}^n \left(\sum_{j=S_{i-1}+1}^{S_i} \zeta_j + Y_i \right) - \sum_{i=1}^n Y_i - m S_n \right\} \\ &= \frac{1}{S_n} \sum_{i=1}^{S_n} (\zeta_i - m) \end{aligned}$$

En particular para S_G , tenemos

$$\hat{m}_G - m = \frac{1}{S_G} \sum_{i=1}^{S_G} (\zeta_i - m),$$

lo cual implica que

$$S_G^{1/2} \left\{ \frac{\hat{m}_G - m}{\sigma} \right\} = \frac{\sqrt{S_G}}{\sigma S_G} \sum_{i=1}^{S_G} (\zeta_i - m) = \frac{1}{\sigma \sqrt{S_G}} \sum_{i=1}^{S_G} (\zeta_i - m). \quad (3.21)$$

Demostración del Teorema 3.4.1. La convergencia casi segura, se sigue de la *Ley Fuerte de los Grandes Números* puesto que \hat{m}_n converge casi seguramente a m y como una consecuencia de esto, la convergencia de \hat{m}_G se satisface gracias a que $G \rightarrow \infty$ c.s.

A continuación, veamos la normalidad asintótica de \hat{m}_G :

En el caso $m \in (0, 1]$, por el *Teorema Central del Límite* tenemos que

$$\frac{1}{\sigma \sqrt{n}} \sum_{i=1}^n (\zeta_i - m) \xrightarrow{d} \mathcal{N}(0, 1)$$

y bastaría probar que

$$\frac{S_G}{c\sigma^2} \xrightarrow{\mathbb{P}} 1 \quad (3.22)$$

para que se cumpla el resultado a través de la ecuación (3.21) y a partir del Lema 3.4.2.

Ahora, para demostrar la ecuación (3.22) es suficiente probar que

$$\frac{S_G}{c\sigma^2} \xrightarrow{c.s.} 1. \quad (3.23)$$

(i) Para $m < 1$, por el Teorema 3.2.3,

$$\frac{N_c}{c} \xrightarrow{c.s.} \frac{(1-m)\sigma^2}{\lambda} \quad \text{cuando } c \rightarrow \infty$$

además, por la ecuación (3.13), para $\alpha \in (0, 1/2)$

$$\frac{G - N_c}{c} = \frac{1}{c^{1-\alpha}} \left(\frac{G - N_c}{c^\alpha} \right) \xrightarrow{c.s.} 0 \quad \text{cuando } c \rightarrow \infty,$$

y por el *Teorema de Slutsky* tenemos que

$$\frac{G}{c} \xrightarrow{c.s.} \frac{(1-m)\sigma^2}{\lambda} \quad \text{cuando } c \rightarrow \infty. \quad (3.24)$$

También por la condición de ergodicidad (ver Capítulo 1), sabemos que

$$\frac{S_n}{n} \xrightarrow{c.s.} \frac{\lambda}{1-m} \quad \text{para todo } n \geq 1 \quad (3.25)$$

así finalmente por las ecuaciones (3.24) y (3.25)

$$\frac{S_G}{c\sigma^2} = \frac{1}{\sigma^2} \frac{S_G}{G} \frac{G}{c} \xrightarrow{c.s.} 1, \quad c \rightarrow \infty.$$

(ii) Para $m = 1$, observemos por la definición de $S(\cdot)$, que

$$\frac{S_G - S_{N_c}}{c} = \frac{1}{c} \sum_{i=N_c+1}^G Z'_{i-1}$$

así mismo, por el Corolario 3.2.2 $Z'_n = o(n \log n)$ luego por propiedad de $o(\cdot)$

$$\frac{S_G - S_{N_c}}{c} = (G - N_c) \frac{o(G \log G)}{c}$$

y a partir de la ecuación (3.13),

$$\frac{S_G - S_{N_c}}{c} = \frac{o(N_c \log N_c) c^\alpha}{c},$$

además por la propiedad **(D)** del *esue* se sigue que

$$\frac{S_G - S_{N_c}}{c} = o\left(\frac{\log c}{c^{1/2-\alpha}}\right) = o(1)$$

y de esto se obtiene la ecuación (3.23) ya que $S_{N_c}/c\sigma^2 \xrightarrow{c.s.} 1$.

(iii) Para $m > 1$, y para aplicar el *Teorema de Slutsky*, debemos demostrar

$$\frac{1}{\sigma\sqrt{S_G}} \sum_{i=1}^{S_{N_c}} (\zeta_i - m) \xrightarrow{\mathbb{P}} 0 \quad \text{y} \quad (3.26)$$

$$\frac{1}{\sigma\sqrt{S_G}} \sum_{i=S_{N_c}+1}^{S_G} (\zeta_i - m) \xrightarrow{d} \mathcal{N}(0, 1). \quad (3.27)$$

Definamos por

$$T_n = \sum_{i=1}^n (\zeta_i - m),$$

como una consecuencia del *Teorema de Mapeo Continuo* (ver Apéndice), tenemos que

$$\frac{1}{n^{1/2}} \max_{1 \leq i \leq n} |T_i| \quad \text{converge en distribución.} \quad (3.28)$$

Como $S_n/m^n \xrightarrow{c.s.} W/(m-1)$, tenemos que $S_{N_c}/S_{N_c-1} \xrightarrow{c.s.} m$, cuando $c \rightarrow \infty$, lo cual junto con la propiedad **(A)** del *esue* implica que $S_{N_c} < [2mc\sigma^2]$ donde $[\cdot]$ denota la parte entera.

Observemos que para probar la ecuación (3.26), basta demostrar que

$$\frac{1}{S_G^{1/2}} \max_{1 \leq i \leq [2mc\sigma^2]} |T_i| \xrightarrow{\mathbb{P}} 0. \quad (3.29)$$

Por otro lado, a partir del Teorema 1.2.4 tenemos que

$$\frac{S_G}{m^{G-N_c} S_{N_c}} \xrightarrow{c.s.} 1 \quad (3.30)$$

lo cual implica que

$$\liminf_{c \rightarrow \infty} \frac{S_G}{[2mc\sigma^2]} = \liminf_{c \rightarrow \infty} \frac{S_G}{S_{N_c}} \frac{S_{N_c}}{[2mc\sigma^2]} \geq \liminf_{c \rightarrow \infty} m^{G-N_c} \frac{c}{[2mc\sigma^2]} = \infty,$$

o equivalentemente,

$$\limsup_{c \rightarrow \infty} \frac{[2mc\sigma^2]^{1/2}}{S_G^{1/2}} = 0. \quad (3.31)$$

De las ecuaciones (3.28), (3.31) se sigue que se satisface la condición (3.29).

Para finalizar, veamos que la ecuación (3.27) es cierta, notemos que de las ecuaciones (3.14) y (3.30) tenemos

$$\frac{S_G - S_{N_c}}{m^{G-N_c} S_{N_c}} = \frac{S_G}{m^{G-N_c} S_{N_c}} - \frac{S_{N_c}}{m^{G-N_c} S_{N_c}} \xrightarrow{c.s.} 1 \quad \text{cuando } c \rightarrow \infty \quad (3.32)$$

así probar la ecuación (3.27) es equivalente a demostrar que

$$\frac{1}{\sigma \sqrt{S_G - S_{N_c}}} \sum_{i=S_{N_c+1}}^{S_G} (\zeta_i - m) \xrightarrow{d} \mathcal{N}(0, 1).$$

En efecto, lo anterior lo podemos demostrar para cualquier subsucesión dada, digamos $\{c_n\}$ con $c_n \rightarrow \infty$:

Sean $H_n = (R_n, N_n, Q_n)$ iguales en distribución que $(S_{N_{c_n}}, N_{c_n}, Z'_{N_{c_n}})$ e independientes de $\{\zeta_j\}$ y $\{Y_j\}$ así aplicando el Lema 3.4.1 obtenemos

$$\frac{1}{\sigma\sqrt{S_{G(c_n, N_{c_n})} - S_{N_{c_n}}}} \sum_{i=S_{N_{c_n}}+1}^{S_G} (\zeta_i - m) \stackrel{d}{=} \frac{1}{\sigma\sqrt{S_{\Delta_n}(Q_n)}} \sum_{i=1}^{S_{\Delta_n}(Q_n)} (\zeta_i - m),$$

siendo $\Delta_n := G(c_n, N_n) - N_n$, cuyo término sólo depende de c_n y N_n , ya que $S_{\Delta_n}(Q_n) \stackrel{d}{=} S_G - S_{N_c}$, de la ecuación (3.32) se tiene que

$$\frac{S_{\Delta_n}(Q_n)}{m^{\Delta_n} R_n} \xrightarrow{\mathbb{P}} 1,$$

y por los Lemas 3.4.1 y 3.4.2, se prueba el resultado. □

Observación 3.4.1. Por otro lado, en el caso de que el parámetro σ^2 es desconocido N_c se puede reemplazar por $N(c)$ y se procede de la misma manera, sin embargo, para obtener un intervalo de confianza para m necesitamos un estimador de σ^2 , *Sriram y Shete* en [42], definieron un estimador de este parámetro como

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n \left[\frac{(Z_i - \hat{m}_n Z_{i-1} - Y_i)^2}{Z_{i-1}} \right]$$

y probaron que $\hat{\sigma}_n^2 \xrightarrow{c.s.} \sigma^2$ bajo el supuesto $\mathbb{E}[\zeta^4] < \infty$ en Teorema A.2.1 (ver apéndice).

Como $G(c) \rightarrow \infty$ c.s. se obtiene que $\hat{\sigma}_G^2 \xrightarrow{c.s.} \sigma^2$ cuando $c \rightarrow \infty$ y por lo tanto se sigue:

Teorema 3.4.3. *Supongamos que $G(c) = G(c, N(c))$, satisface las ecuaciones (3.13) y (3.14), cuando N_c se reemplaza por $N(c)$, además $\mathbb{E}[\zeta^4] < \infty$ entonces*

$$\left(\sum_{i=1}^G Z_{i-1} \right)^{1/2} \frac{\hat{m}_G - m}{\hat{\sigma}_G} \xrightarrow{d} \mathcal{N}(0, 1). \quad (3.33)$$

En conclusión, con este resultado podemos inferir cual es el valor o valores de m , además aparte de tener que la distribución del estimador puntual es aproximadamente Normal también sabemos que el error estándar se aproxima a cero a medida que G crece, lo cual implica que podemos construir intervalos que aproximadamente cubren de manera adecuada el parámetro de interés.

Ahora, otra pregunta que resulta interesante a la hora de hacer inferencia estadística bajo éste enfoque es si el valor de m está clasificado adecuadamente⁶, aunque ese no será nuestro caso a lo largo de este documento puesto que consideraremos simulaciones, sí podríamos aprovechar el resultado anterior y plantear las hipótesis adecuadas, por ejemplo, contrastar $H_0 : m = 1$ contra $H_1 : m \neq 1$.

Finalmente, como lo dijimos al inicio del capítulo, el resultado principal de esta sección es de suma importancia desde el punto de vista de las aplicaciones, cabe notar que hay que tener presente que es un resultado asintótico lo cual puede ser un pequeño inconveniente en la aplicación del mismo, respecto al tamaño de muestra a considerar. Sin embargo, la estimación secuencial permite una muy buena aproximación al tamaño de muestra óptimo tanto teóricamente como experimentalmente, ésta afirmación la comprobaremos en el siguiente capítulo.

⁶Es un error común, confundir un problema de estimación con un problema de prueba de hipótesis, a pesar de que los resultados encontrados en uno pueden servir para responder el otro, la nociones de precisión no son equivalentes.

Capítulo 4

Simulaciones

4.1. Características

En los capítulos anteriores, explicamos un enfoque para encontrar la distribución límite de los estimadores del valor esperado de reproducción (m) y del valor esperado de inmigración (λ) sin tener en cuenta la clasificación del proceso de Galton-Watson con Inmigración, así mismo en el Capítulo 2 presentamos algunos estimadores para dichos parámetros y comentamos algunas de sus propiedades. A continuación presentamos un estudio basado en simulaciones para observar el comportamiento límite de los mismos, es decir corroboramos algunos de los resultados vistos anteriormente.

Consideramos que las distribuciones de inmigración y reproducción pertenecen a la familia paramétrica *Poisson*, donde en cada caso se asumió, para el parámetro para la distribución de reproducción fuera $\theta = 0.5$, $\theta = 1$ y $\theta = 2$, respectivamente. En cambio, el parámetro para la distribución de inmigración siempre se consideró como $\varphi = 2$.

Recordemos que por distribución límite entendemos un comportamiento aproximado y no exacto, es decir, bajo muestras grandes, las probabilidades con la distribución muestral exacta y con la Normal estándar nos darán cantidades muy similares, así dicha aproximación depende tanto de la forma de la distribución de donde se obtiene la muestra y del tamaño de muestra, cabe notar que la distribución *Poisson* con éstos parámetros tiene una forma asimétrica.

Para el procedimiento secuencial la muestra va en avance, y en la práctica la implementación del procedimiento depende de esto, sin embargo, en nuestro caso se generó un vector c de longitud 100 iniciado en 100 con cada componente espaciada en 5 y de acuerdo a los valores de G obtenidos en cada caso se generaron muestras para los demás estimadores. Se asumió que la función de dos etapas está dada como

$G(c) = \log(c) + N_c$, la cuál satisface las ecuaciones (3.13) y (3.14).

En cada caso se presentan las gráficas de acuerdo al siguiente orden, recordemos que para la media de reproducción y de inmigración se tiene para cada enfoque que:

Estimadores Secuenciales:

$$\hat{m}_G = \frac{1}{\sum_{i=1}^G Z_{i-1}} \left\{ \sum_{i=1}^G (Z_i - Y_i) \right\} \quad \text{y} \quad \hat{\lambda}_G = \frac{1}{G} \left\{ \sum_{i=1}^G Y_i \right\}.$$

Estimadores de Mínimos Cuadrados:

$$\tilde{m}_n = \left\{ \sum_{i=1}^n Z_i \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n \sum_{i=1}^n \frac{Z_i}{Z_{i-1} + 1} \right\} \left\{ \sum_{i=1}^n (Z_{i-1} + 1) \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n^2 \right\}^{-1} \quad \text{y}$$

$$\tilde{\lambda}_n = \left\{ \sum_{i=1}^n Z_{i-1} \sum_{i=1}^n \frac{Z_i}{Z_{i-1} + 1} - \sum_{i=1}^n Z_i \sum_{i=1}^n \frac{Z_{i-1}}{Z_{i-1} + 1} \right\} \left\{ \sum_{i=1}^n (Z_{i-1} + 1) \sum_{i=1}^n \frac{1}{Z_{i-1} + 1} - n^2 \right\}^{-1}.$$

Estimadores de Bayes Conjugado:

En el caso particular de la distribución *Poisson* los estimadores de Bayes, considerando el análisis conjugado visto en la Sección 2.4, están dados por

$$\hat{m} = \frac{\alpha_1 + Z_n^* - Y_n^* + 1}{\beta_1 + Z_{n-1}^*}, \quad \text{y} \quad \hat{\lambda} = \frac{\alpha_2 + Z_n^* - Y_n^* + 1}{\beta_2 + Z_{n-1}^*}.$$

Aproximación *Data cloning*:

La aproximación *Data cloning* nos genera al igual que el procedimiento Bootstrap los estimadores de máxima verosimilitud pero esta última aproximación fue modificada buscando una convergencia más rápida a los estimadores, por tal razón no presentamos los resultados obtenidos con la misma puesto que la modificación se hace de manera muy elaborada. En cambio, en el caso del *Data cloning* no es necesario ningún tipo de modificación ni de supuestos alternos acerca de las distribuciones. En los tres casos, usamos $k = 10$.

4.2. Resultados

En algunos casos, al final de cada grupo de gráficas (Secuencial, Mínimos cuadrados, Bayes conjugado y *Data cloning*), se incluyen las generadas sin reescalamiento, lo cual indica los valores reales para dichos estimadores. También, se incluyen tablas con algunos valores para los estimadores en el procedimiento secuencial, sin ningún tipo de selección predeterminada sino simplemente escogida para valores de c igualmente espaciados.

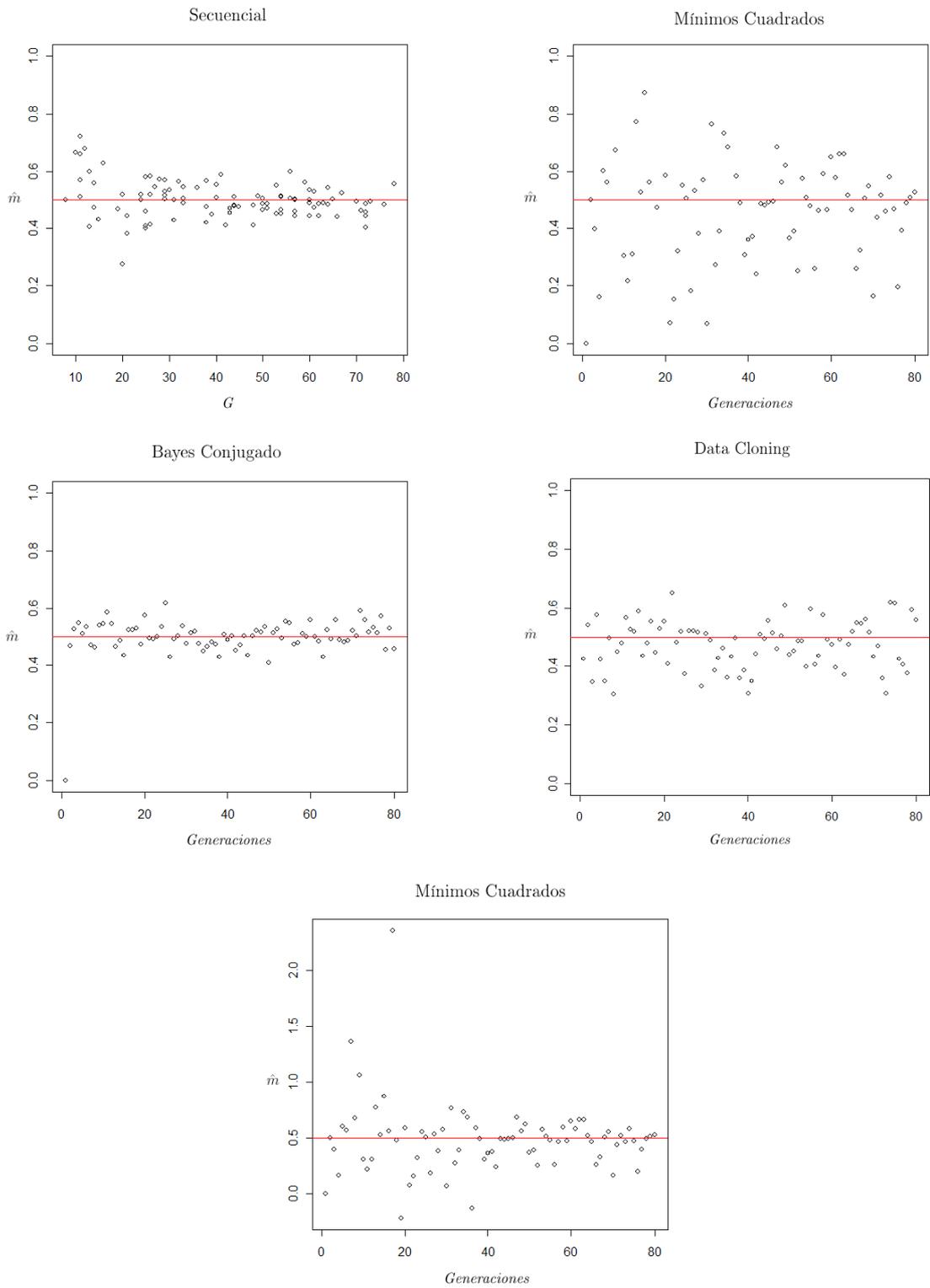


Figura 4.1: Estimadores de \hat{m} , caso subcrítico

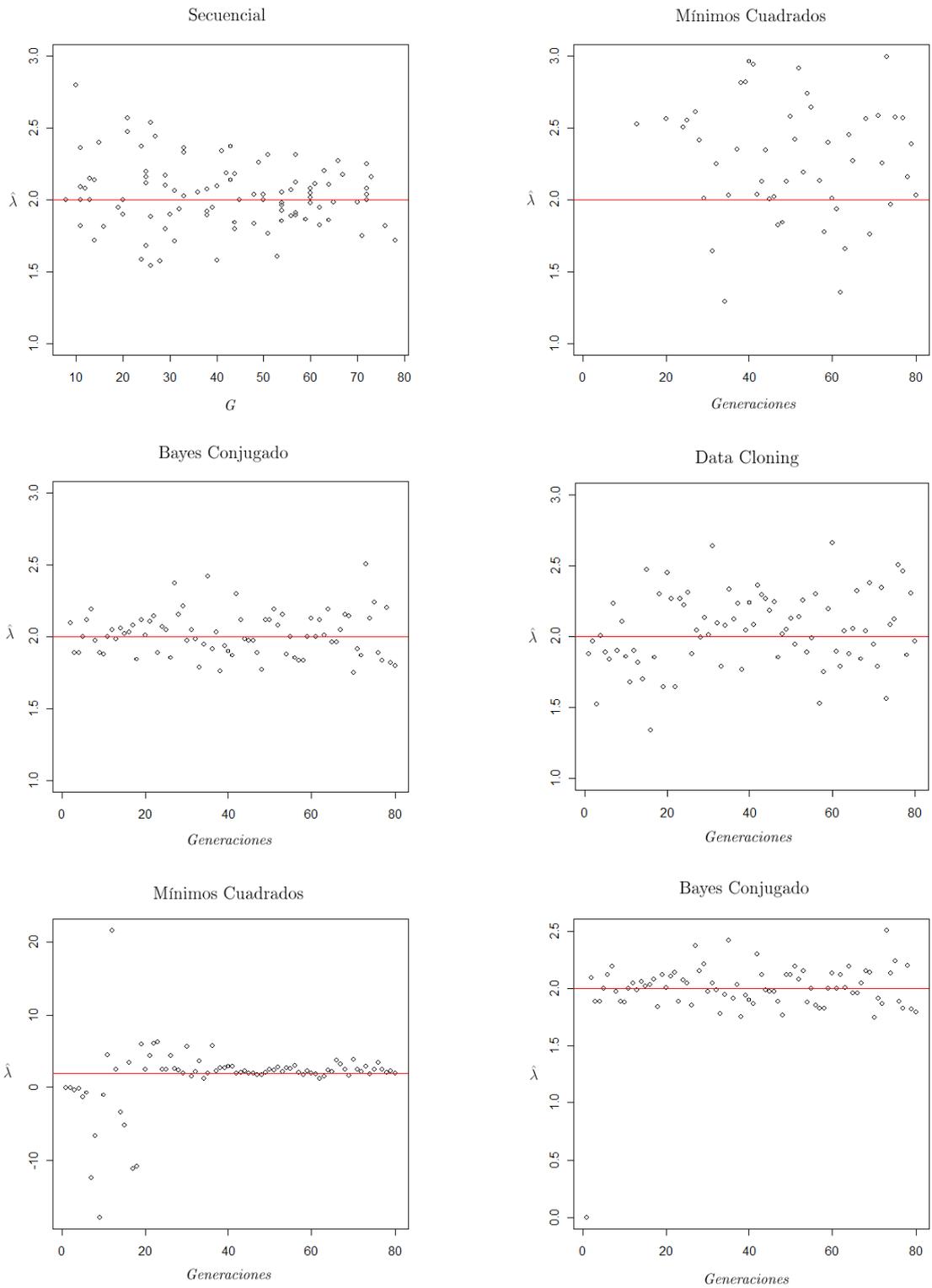


Figura 4.2: Estimadores de $\hat{\lambda}$, caso subcrítico

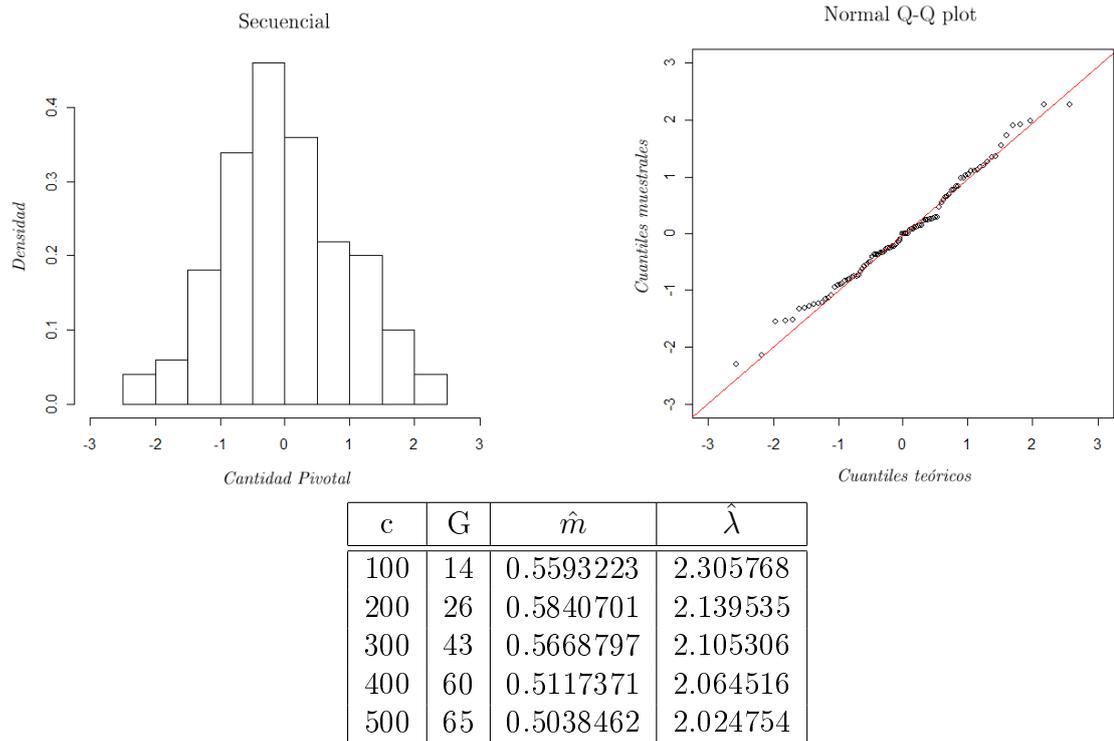


Figura 4.3:
Valores secuenciales, caso subcrítico

Notemos que en cada método el comportamiento de los estimadores es adecuado teniendo en cuenta el tamaño de muestra, lo cual indica que el estimador secuencial tiene un buen comportamiento tanto en términos de la distribución asintótica como de la propiedad de consistencia.

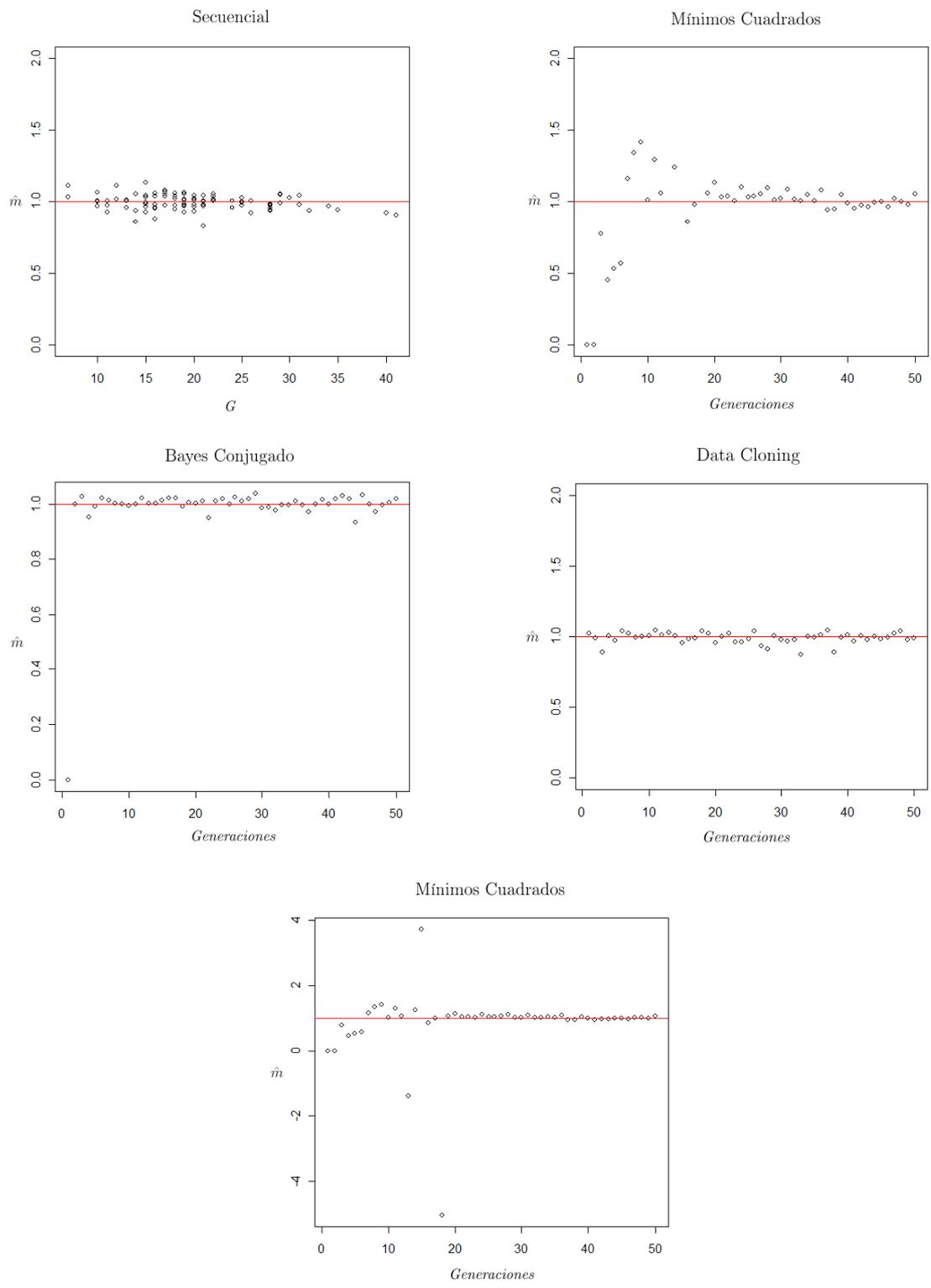


Figura 4.4: Estimadores de \hat{m} , caso crítico

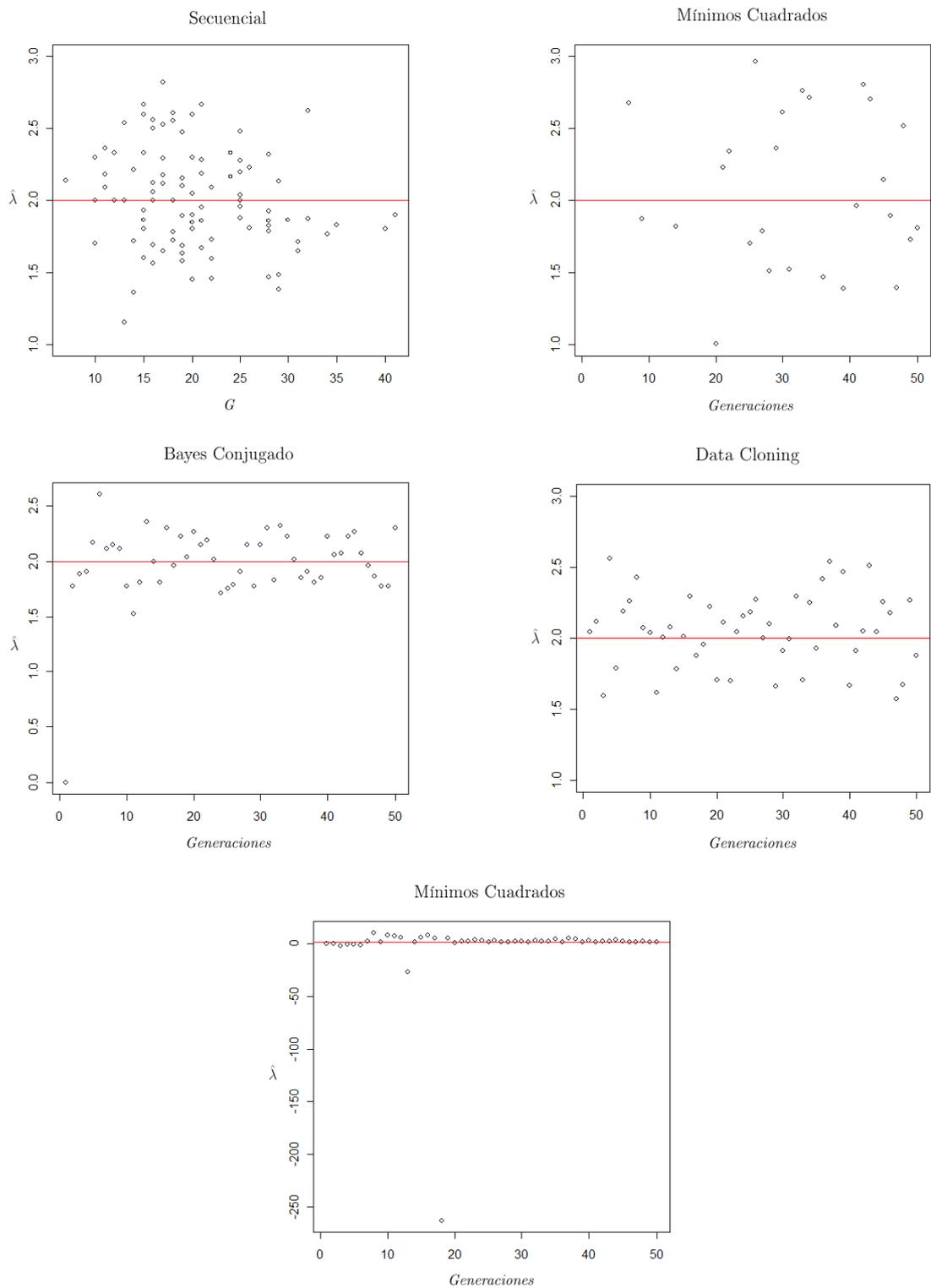
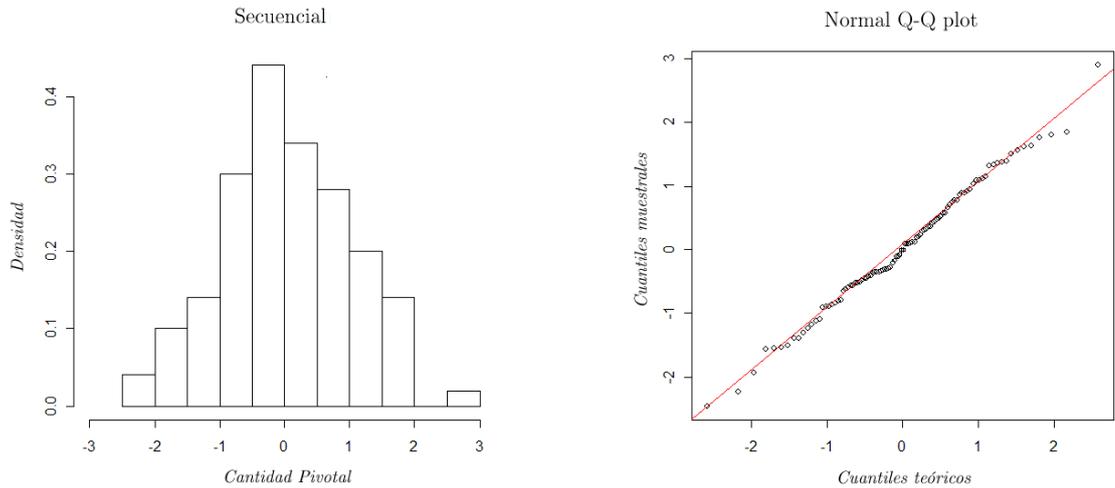


Figura 4.5: Estimadores de $\hat{\lambda}$, caso crítico



c	G	\hat{m}	$\hat{\lambda}$
100	11	0.9291339	2.090909
200	16	0.9960938	2.176471
300	19	1.0688891	2.230769
400	26	1.0403592	2.230769
500	41	1.0131583	2.090909

Figura 4.6:
Valores secuenciales, caso crítico

En este caso, observamos de nuevo el buen comportamiento del estimador secuencial para m , donde el tamaño de muestra mejora respecto al caso subcrítico y los resultados favorecen de nuevo a dicho estimador. Aunque ésto era de esperarse debido al sustento teórico que soporta este procedimiento, el hecho de obtener buenos resultados con un tamaño de muestra relativamente pequeño indica una buena calidad del mismo.

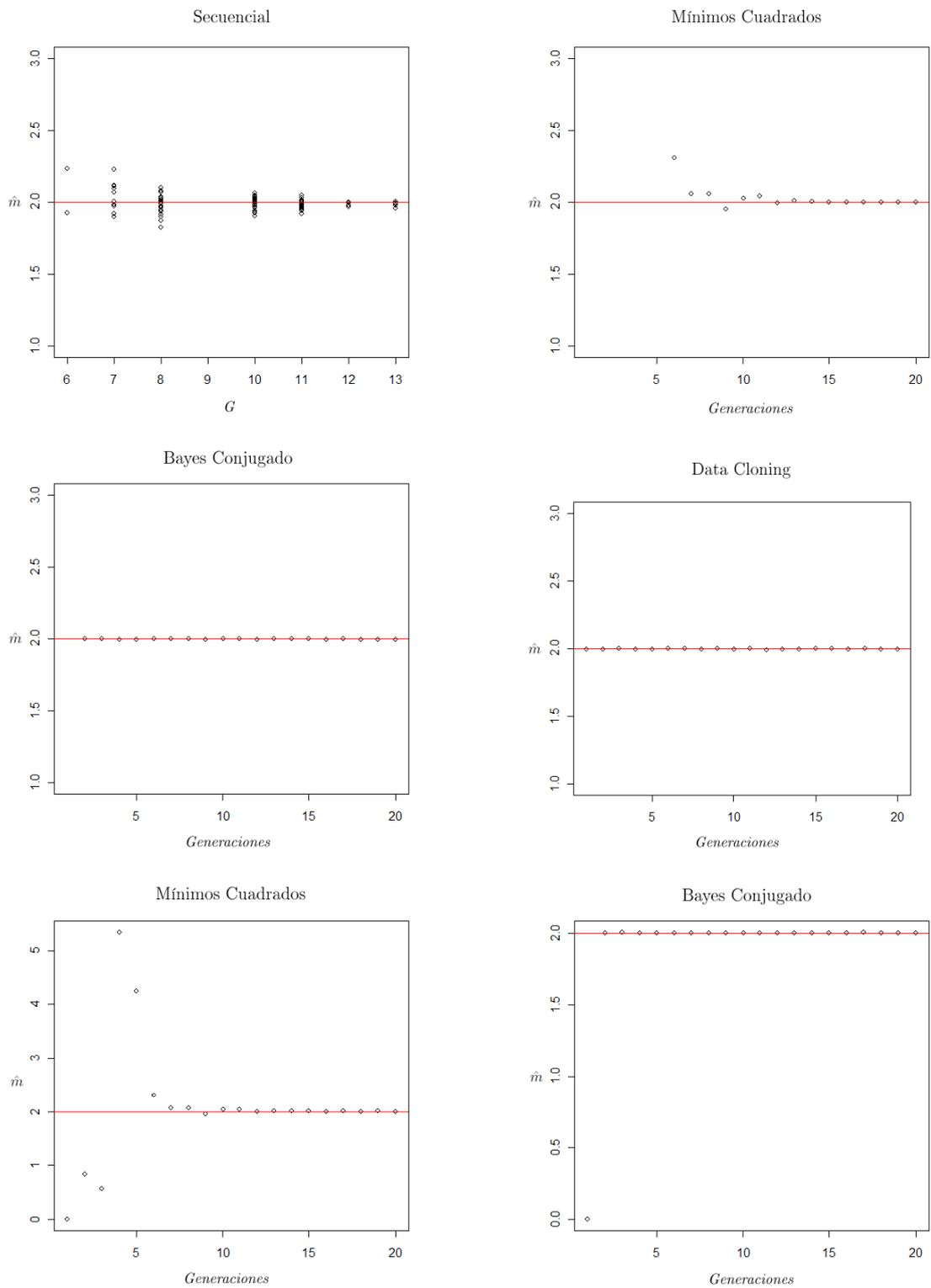


Figura 4.7: Estimadores de \hat{m} , caso supercrítico

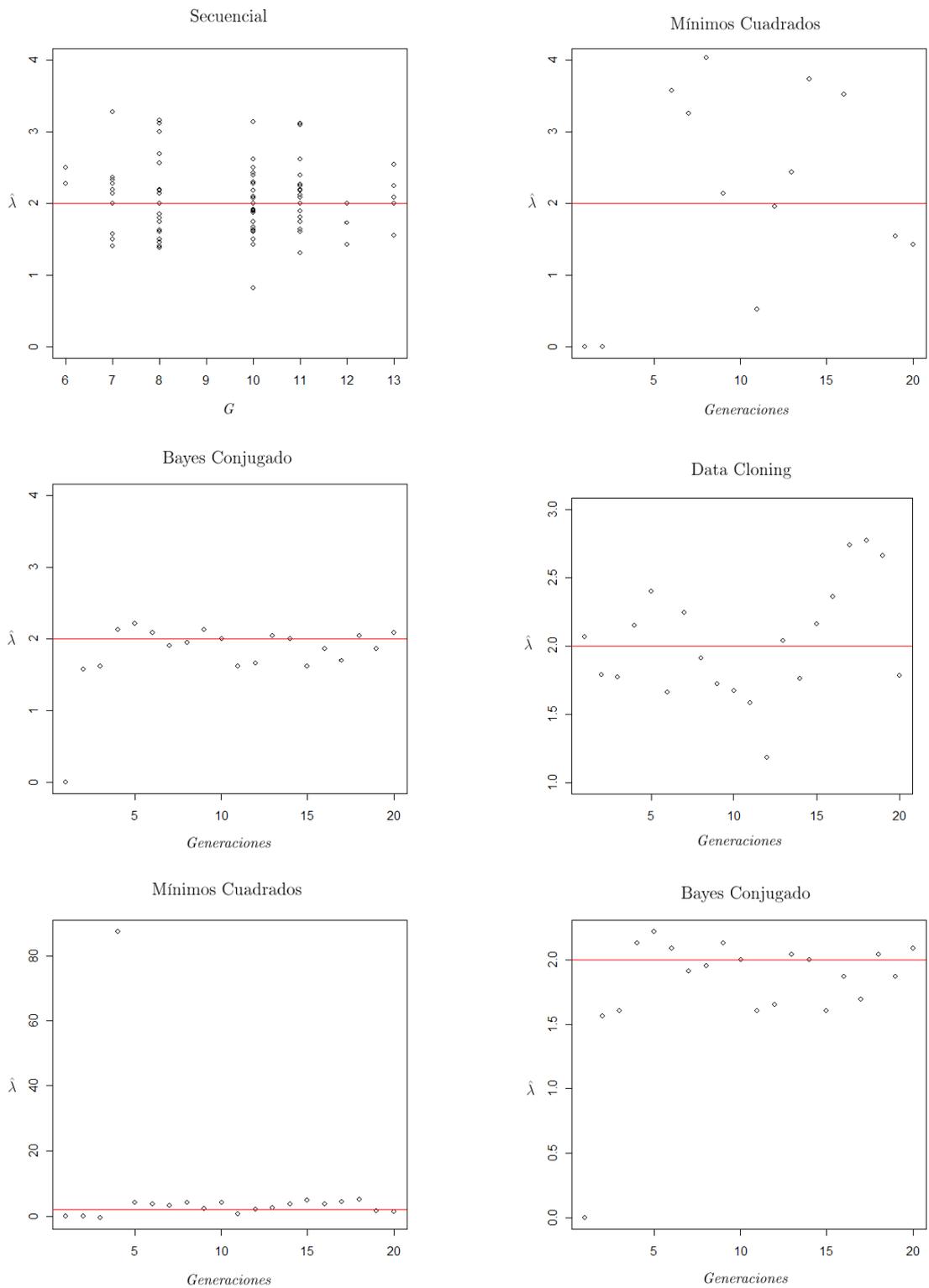
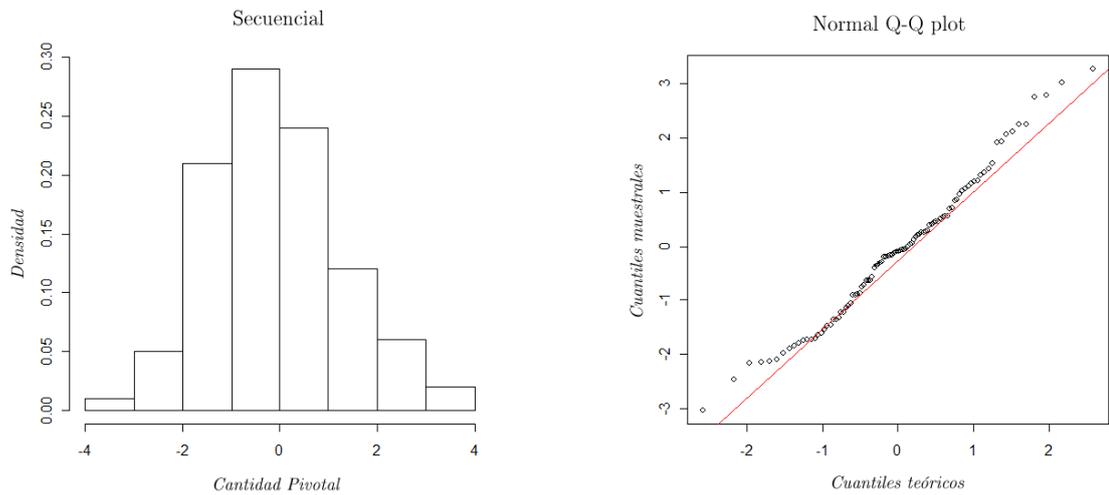


Figura 4.8: Estimadores de $\hat{\lambda}$, caso supercrítico



c	G	\hat{m}	$\hat{\lambda}$
100	8	2.011594	2.142857
200	8	1.971618	2.625478
300	10	2.072121	1.759673
400	11	2.012824	1.9
500	11	2.035477	2.2

Figura 4.9:
Valores secuenciales, caso supercrítico

De nuevo, el tamaño de muestra necesario para obtener estos resultados del estimador secuencial tanto en términos de la distribución asintótica como de la propiedad de consistencia es muy bueno, aunque los demás estimadores también se aproximan adecuadamente. Lo anterior nos permite concluir acerca de la eficiencia del procedimiento secuencial comparado con los demás métodos de estimación, puesto que en este último contamos con la distribución límite.

Conclusiones

Con este trabajo, observamos varios aspectos relacionados al tema inferencia de procesos estocásticos:

Inicialmente, resaltamos la importancia de la media de reproducción en el comportamiento límite del proceso de G.W. y vimos como éste puede modificarse adicionando una componente de inmigración la cual permite establecer un comportamiento asintótico diferente, específicamente en el caso subcrítico del proceso, recordando resultados clásicos de dichos procesos.

Observamos, en el ámbito de inferencia estadística las dificultades que se pueden presentar en el momento de hacer estimación de una cantidad de interés, específicamente en un modelo estocástico a pesar de la simplicidad del mismo, llevando ésto a un desarrollo de la teoría que usa herramientas probabilísticas cuya aplicación a nivel práctico se realiza de manera sencilla, como lo es la implementación del procedimiento secuencial y la aproximación *Data cloning*.

Aunque sin olvidar que en el momento de hacer inferencia estadística hay que tener en cuenta algunos aspectos relevantes que pueden causar otro tipo de inconvenientes y que nosotros los tuvimos como supuesto a lo largo de éste documento:

- Selección del modelo adecuado para el experimento, ésto incluye tanto el conocimiento del proceso para determinar que se comporta como un proceso de G.W.I como de la familia de distribuciones a considerar para las v.a.'s de reproducción e inmigración.
- No siempre tenemos que las v.a.'s son i.i.d., son segundos momentos finitos.
- No siempre el tamaño de muestra es grande, aunque ésta es una de las ventajas que posee el procedimiento secuencial.

En resumen, logrando cumplir de manera adecuada con estos requisitos tenemos que en el caso del procedimiento secuencial, podemos obtener buena información acerca de la evolución de la población, correspondiente a un modelo de G.W.I., puesto que

tenemos estimaciones de buena calidad sin tener que conocer la clasificación de dicho proceso.

También, simulamos variaciones del modelo, encontrando un buen comportamiento del estimador secuencial y de los estimadores de máxima verosimilitud extraídos del *Data cloning*.

Por último, a lo largo del tema hay varios aspectos a revisar, entre ellos están:

- La información contenida en la muestra $\{Z_n, Y_n, \dots, Z_1, Y_1, Z_0\}$ no siempre está disponible así que se propone estimar los parámetros de las v.a.'s de reproducción e inmigración a partir de $\{Z_n, \dots, Z_1, Z_0\}$, considerando que la estructura del modelo es anidada.
- Para el ámbito de las aplicaciones es importante estudiar el comportamiento límite de las cantidades ζ_n/Z_n y Y_n/Z_n , puesto que nos interesa poder observar si hay alguna población dominante, lo cual puede intuirse a través de la modificación que involucra el hecho de considerar la componente de inmigración.

Apéndice A

A.1. Condición

En [38], *Robbins y Siegmund* dan un tratamiento unificado a varios teoremas de convergencia casi segura explotando el hecho de que algunos procesos poseen en común el método de usar una *casi supermartingala*, a continuación daremos la definición de éste concepto y enunciaremos una propiedad que corresponde a la Proposición 2 de dicho artículo en el cual se encuentra su demostración.

Definición A.1.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\mathcal{F}_1 \subset \mathcal{F}_2, \dots$ una sucesión de σ -álgebras de \mathcal{F} . Para cada $n = 1, 2, \dots$, sean X_n, β_n, ξ_n y ς_n variables no-negativas \mathcal{F}_n -medibles, se dice que $\{X_n\}$ es una casi supermartingala si cumple

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n(1 + \beta_n) + \xi_n - \varsigma_n. \quad (\text{A.1})$$

Proposición 2. Si $\{X_n\}$ es una casi supermartingala, es decir, se cumple la condición (A.1) entonces para cualquier $a > 0, m = 1, 2, \dots$, y $n \geq m$

$$\mathbb{P} \left[\max_{m \leq k \leq n} X_k \geq a | \mathcal{F}_m \right] \leq a^{-1} \left\{ \mathbb{E}[X_m] + \mathbb{E} \left[\sum_{k=m}^{n-1} \xi_k | \mathcal{F}_m \right] \right\} + \mathbb{E} \left[\sum_{k=m}^{n-1} \beta_k | \mathcal{F}_m \right].$$

En adelante, todos los resultados se demuestran considerando que $\{Z_n\}$ es un proceso de Galton- Watson con Inmigración definido en el Capítulo 1.

Teorema A.1.1. Para cada $\delta > 0, m \in (0, 1]$ y con $S_n = \sum_{i=1}^n Z_{i-1}$ tenemos que

$$\lim_{k \rightarrow \infty} \mathbb{P} [Z_n \geq \delta S_n \quad \text{para algún } n \geq k] = 0.$$

Demostración. Notemos que podemos reescribir

$$\frac{Z_n^2}{S_n^2} = \frac{Z_n^2}{(S_n + n)^2} \left[1 + \frac{n}{S_n} \right]^2$$

de ésta manera se satisface la siguiente desigualdad

$$\frac{Z_n^2}{(S_n + n)^2} \left[1 + \frac{n}{S_n} \right]^2 \leq \frac{Z_n^2}{(S_n + n)^2} \left[1 + \frac{n}{\sum_{i=1}^{n-1} Y_i} \right]^2$$

además, recordemos que la *Ley fuerte de los grandes números* nos dice que cuando $\mathbb{E}[Y] = \lambda < \infty$

$$1 + n \left(\sum_{i=1}^{n-1} Y_i \right)^{-1} \xrightarrow{c.s.} 1 + \lambda^{-1} \quad \text{cuando } n \rightarrow \infty.$$

A partir de lo anterior, para obtener el resultado, es suficiente mostrar que

$$\lim_{k \rightarrow \infty} \mathbb{P} \left[Z_n^2 \geq \delta^2 [S_n + n]^2, \quad \text{para algún } n \geq k \right] = 0. \quad (\text{A.2})$$

Para esto, primero definamos por $M_n = Z_n^2 / (S_n + n)^2$, recordemos que en el Lema 3.2.1, escribimos a

$$Z_n = \sum_{i=1}^n W_i + \sum_{i=0}^n Y_i \quad \text{con} \quad W_i = \sum_{k=1}^{Z_{i-1}} \zeta_{i-1,k} - Z_{i-1},$$

así $Z_{n+1} = Z_n + W_{n+1} + Y_{n+1}$ por lo tanto, $Z_{n+1}^2 \leq Z_n^2 + 2Y_{n+1}Z_n + Y_{n+1}^2$, entonces con algunas operaciones algebraicas, tenemos que

$$\mathbb{E}[M_{n+1} | \mathcal{F}_n] \leq M_n + \xi_n,$$

siendo

$$\xi_n = \frac{2\lambda Z_n + \mathbb{E}[Y^2]}{[S_{n+1} + (n+1)]^2},$$

observemos que se cumple la condición (A.1) y por lo tanto por la **Proposición 2**,

$$\mathbb{P} \left[\max_{n \geq k} M_n \geq \delta^2 \right] \leq \delta^{-2} \left\{ \mathbb{E}[M_k] + \mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] \right\}$$

y como

$$\begin{aligned} \mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] &\leq 2\lambda \mathbb{E} \left[\sum_{n=k}^{\infty} \frac{S_{n+1} - S_n}{[S_{n+1} + (k+1)]^2} \right] + \mathbb{E}[Y^2] \sum_{n=k}^{\infty} (n+1)^{-2} \\ &\leq 2\lambda \int_0^{\infty} [(k+1) + x]^{-2} dx + \mathbb{E}[Y^2] \int_k^{\infty} x^{-2} dx \end{aligned}$$

se obtiene que

$$\mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] \longrightarrow 0 \quad \text{cuando } k \rightarrow \infty,$$

y

$$\begin{aligned} \mathbb{E}[M_k] &\leq \mathbb{E}[M_{k-1}] + \frac{2\lambda \mathbb{E}[Z_{k-1}]}{[S_k + k]^2} + \frac{\mathbb{E}[Y^2]}{k^2} \\ &\vdots \\ &\leq 2\lambda \mathbb{E} \left[\sum_{i=1}^k \frac{Z_{i-1}}{[S_i + i]^2} \right] + \frac{\mathbb{E}[Y^2]}{k} \\ &\leq 2\lambda \mathbb{E} \left[\sum_{i=1}^k \frac{Z_{i-1}}{[S_i + k]^2} \right] + \frac{\mathbb{E}[Y^2]}{k} \end{aligned}$$

de donde

$$\mathbb{E}[M_k] \longrightarrow 0 \quad \text{cuando } k \rightarrow \infty,.$$

Así finalmente, obtenemos que

$$\mathbb{P} \left[\max_{n \geq k} M_n \geq \delta^2 \right] \longrightarrow 0 \quad \text{cuando } k \rightarrow \infty,$$

y por tanto se sigue el resultado. □

A.2. Convergencia de σ^2

Sriram en [42] definió como

$$\hat{\sigma}_n^2 = \frac{1}{n \sum_{i=1}^n Z_{i-1}} \sum_{i=1}^n [Z_i - \hat{m}_n Z_{i-1} - Y_i],$$

el siguiente resultado nos garantiza la convergencia de $\hat{\sigma}_n^2$ a σ^2 .

Teorema A.2.1. *Si $\mathbb{E}[\zeta^4] < \infty$, entonces para todo $\delta > 0$*

$$\lim_{k \rightarrow \infty} \mathbb{P}[|\hat{\sigma}_n^2 - \sigma^2| \geq \delta \quad \text{para algún } n \geq k] = 0.$$

Demostración. Observemos que

$$\hat{\sigma}_n^2 - \sigma^2 = \frac{1}{n} \sum_{i=1}^n (\omega_i - \sigma^2), \quad \text{con } \omega_i = \frac{1}{\sqrt{Z_{i-1}}} \{Z_i - m Z_{i-1} - Y_i\}$$

donde ω_i está definido para $\sqrt{Z_{i-1}} \neq 0$ y 0 en otros casos.

Podemos definir a

$$X_n = \frac{1}{n^2} \left[\sum_{i=1}^n (\omega_i^2 - \sigma^2) \right]^2,$$

y realizando algunas operaciones algebraicas obtenemos que

$$\mathbb{E}[X_n | \mathcal{F}_{n-1}] \leq X_{n-1} + \xi_n$$

con $\xi_n = \frac{1}{n^2} \mathbb{E}[(\omega_n^2 - \sigma^2)^2 | \mathcal{F}_{n-1}]$ esto es, $\{X_n\}$ es casi una supermartingala y por la **Proposición 2** se tiene que,

$$\mathbb{P}[\max_{n \geq k} X_n \geq \delta^2] \leq \delta^{-2} \left\{ \mathbb{E}[X_k] + \mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] \right\},$$

notemos que

$$\begin{aligned} \mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] &= \mathbb{E} \left[\sum_{n=k}^{\infty} n^{-2} \right] \mathbb{E}[(\omega_n^2 - \sigma^2)^2 | \mathcal{F}_{n-1}] \\ &\leq \mathbb{E} \left[\sum_{n=k}^{\infty} n^{-2} \right] \mathbb{E}[\omega_n^4 | \mathcal{F}_{n-1}] - \sigma^4 \sum_{n=k}^{\infty} n^{-2} \\ &\leq K_0 \mathbb{E}[\zeta^4] \sum_{n=k}^{\infty} n^{-2} - \sigma^4 \sum_{n=k}^{\infty} n^{-2} \end{aligned}$$

de donde la última desigualdad se cumple por el Corolario 10.32 en [7], con K_0 constante, así

$$\mathbb{E} \left[\sum_{n=k}^{\infty} \xi_n \right] \longrightarrow 0, \quad \text{cuando } k \rightarrow \infty.$$

Además,

$$\begin{aligned}
\mathbb{E}[X_k] &= \frac{1}{k^2} \mathbb{E} \left[\left[\sum_{i=1}^{k-1} (\omega_i^2 - \sigma^2) \right]^2 + (\omega_k^2 - \sigma^2)^2 \right] \\
&\leq \frac{1}{[k/2 + (k-1)/2]^2} \mathbb{E} \left[\sum_{i=1}^{k-1} (\omega_i^2 - \sigma^2) \right]^2 + \frac{1}{[k/2 + k/2]^2} \mathbb{E}[\omega_k^2 - \sigma^2]^2 \\
&\vdots \\
&\leq \mathbb{E} \left[\sum_{i=1}^k (\omega_i^2 - \sigma^2)^2 / [k/2 + i/2]^2 \right] \\
&\leq [K_0 \mathbb{E}[\zeta^4] + \sigma^4] \sum_{i=1}^k \frac{1}{[k/2 + i/2]^2} \\
&\leq C_0 \int_0^\infty \frac{1}{(k/2 + x)^2} dx
\end{aligned}$$

de donde la última desigualdad se satisface por el Corolario 10.32 en [7], siendo C_0 y K_0 constantes,

$$\mathbb{E}[X_k] \longrightarrow 0, \quad \text{cuando } k \rightarrow \infty.$$

□

A.3. Difusión

Sea $D^+[0, \infty)$ el espacio de las funciones no-negativas en $[0, \infty)$ las cuales son continuas por la derecha y tienen límites por la izquierda. Dotamos éste espacio con la topología de *Skorohod* (ver [9]). Ahora, sea $W_n(t) = \frac{1}{n} Z_{[nt]}$, luego W_n es una sucesión de elementos aleatorios que toman valores en $D^+[0, \infty)$. También denotamos, por D un proceso de difusión no-negativo con generador

$$Af(x) = \frac{1}{2} \sigma^2 x f^{(2)}(x) + \lambda f^{(1)}(x), \quad \text{con } f \in C_c^\infty[0, \infty), \quad (\text{A.3})$$

donde $C_c^\infty[0, \infty)$ es el espacio de las funciones infinitamente diferenciables en $[0, \infty)$ las cuales tienen soporte compacto (ver [9]).

El siguiente resultado se demuestra de manera similar que en *Either* y *Kurtz* (1986, Capítulo 9, Teorema 1.3), él cual se conoce como la aproximación por difusión

de *Feller*. En nuestro caso, consideramos el proceso de Galton-Watson con Inmigración definido en el Capítulo 1.

Teorema A.3.1. *Supongamos que $m = 1$, $\sigma^2 < \infty$ y $\sigma_Y^2 < \infty$ entonces $W_n \rightarrow W$ (débilmente en $D^+[0, \infty)$), donde W es un proceso de difusión con generador dado por (A.3) y $W(0)=0$.*

Demostración. Por la definición de $\{Z_n\}$, observemos que $\{Z_j/n, j \geq 0\}$ es una cadena de Markov con valores en $E_n = \{l/n : l = 0, 1, 2, \dots, \}$ y para cada $f \in C_c^\infty[0, \infty)$ definimos

$$A_n f(x) = \mathbb{E} \left[f \left(n^{-1} \sum_{j=1}^{nx} \zeta_j + Y \right) \right], \quad x \in E_n.$$

Por el Teorema 6.5 del Cap. 1 y el Corolario 8.9 del Cap. 4 en [9] y puesto que $W_n(0) = Z_0/n \xrightarrow{c.s.} 0$ es suficiente mostrar que

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\epsilon_n(x)| = 0,$$

donde

$$\epsilon_n(x) = n(A_n f(x) - f(x)) - \frac{1}{2} \sigma^2 x f^{(2)}(x) - \lambda f^{(1)}(x) - \frac{1}{2n} f^{(2)}(x) \mathbb{E}[Y^2].$$

Definamos $S_n = n^{-1/2} \left\{ \sum_{j=1}^n (\zeta_j - 1) + Y \right\}$ y para cada $x \in E_n$, afirmamos que

$$\epsilon_n(x) = \mathbb{E} \left[\int_0^1 S_{nx}^2 x (1-v) \{ f^{(2)}(x + v \sqrt{x/n} S_{nx}) - f^{(2)}(x) \} \delta v \right], \quad (\text{A.4})$$

con $x \mathbb{E}[S_{nx}^2] = \sigma^2 x + \frac{1}{n} \mathbb{E}[Y^2]$, y lo demostramos al final.

Supongamos que el soporte de f está en $[0, c]$, como

$$x + v \sqrt{\frac{x}{n}} S_{nx} = x + v \left(\frac{1}{n} \right) \left[\sum_{j=1}^{nx} (\zeta_j - 1) + Y \right] \geq x(1-v),$$

si $v \leq 1 - \frac{c}{x}$ entonces la integral es cero, así

$$\begin{aligned} \left| \int_0^1 S_{nx} x (1-v) \{ f^{(2)}(x + v \sqrt{\frac{x}{n}} S_{nx}) - f^{(2)}(x) \} dv \right| &\leq \int_{0 \vee \{1-c/x\}}^1 S_{nx}^2 x (1-v) 2 \|f^{(2)}\| dv \\ &= x \|f^{(2)}\| (\{c/x\} \wedge 1)^2 S_{nx}^2 \end{aligned} \quad (\text{A.5})$$

Para probar que

$$\lim_{n \rightarrow \infty} \sup_{x \in E_n} |\epsilon_n(x)| = 0,$$

es suficiente mostrar que $\lim_{n \rightarrow \infty} \epsilon_n(x_n) = 0$ para toda sucesión convergente x_n tal que $\lim_{n \rightarrow \infty} x_n < \infty$. En efecto, como

$$\mathbb{E}[S_{nx}^2] = \sigma^2 + \frac{(\sigma_Y^2 + \lambda^2)}{nx} \quad \text{para todo } n \text{ y } x,$$

se sigue que $\lim_{n \rightarrow \infty} x_n = 0$ si $\lim_{n \rightarrow \infty} \epsilon_n(x_n) = 0$. Por lo tanto, sólo necesitamos considerar el caso en que $x_n \rightarrow x$, con $0 < x < \infty$:

Al sustituir, x por x_n en la desigualdad (A.5), S_{nx_n} converge en distribución a una \mathcal{N} y de aquí el lado izquierdo de la desigualdad converge a 0 en probabilidad y el lado derecho en distribución a $x \|f^{(2)}\| (\{c/x\} \wedge 1)^2 \mathcal{N}^2$ y como

$$\lim_{n \rightarrow \infty} \mathbb{E} [x_n \|f^{(2)}\| (\{c/x_n\} \wedge 1)^2 S_{nx_n}^2] = \mathbb{E}[x \|f^{(2)}\| (\{c/x\} \wedge 1)^2 \mathcal{N}^2],$$

el *Teorema de Convergencia Dominada* implica que $\lim_{n \rightarrow \infty} \epsilon_n(x_n) = 0$.

Finalmente, veamos que se cumpla la ecuación (A.4):

Notemos primero que realizando algunos cálculos y denotando por $a = (\frac{x}{n})^{1/2} S_{nx}$, $y = x + va$ obtenemos

$$(\Delta) = \int_0^1 S_{nx}^2 x (1-v) \{f^{(2)}(x + v\sqrt{x/n} S_{nx}) - f^{(2)}(x)\} dv$$

$$\begin{aligned} (\Delta) &= \int_x^{x+a} S_{nx}^2 x \left[\frac{a-y+x}{a} \right] (f^{(2)}(y) - f^{(2)}(x)) \frac{dy}{a} \\ &= n \int_x^{x+a} [a - (y-x)] (f^{(2)}(y) - f^{(2)}(x)) dy \\ &= n \left[-af^{(1)}(x) + f(x+a) - f(x) - \frac{a^2}{2} f^{(2)}(x) \right] \end{aligned}$$

donde

$$\blacksquare f(x+a) = f \left[\frac{1}{n} \sum_{j=1}^{nx} \zeta_j + Y \right]$$

- $-anf^{(1)}(x) = f^{(1)}(x) \mathbb{E} \left[\sum_{j=1}^{nx} (\zeta_j - 1) + Y \right] = -\lambda f^{(1)}(x)$
- $-n \frac{\alpha^2}{2} f^{(2)}(x) = -\frac{1}{2} f^{(2)}(x) \sigma^2 x - \frac{1}{2n} f^{(2)}(x) \mathbb{E}[Y^2],$

y considerando éstas igualdades se prueba el resultado.

Para la segunda afirmación,

$$\mathbb{E}[S_{nx}^2] = \mathbb{E} \left[\frac{1}{nx} \left(\sum_{j=1}^{nx} (\zeta_j - 1) + Y \right)^2 \right],$$

así

$$x \mathbb{E}[S_{nx}^2] = \mathbb{E} \left[\frac{1}{n} \left(\sum_{j=1}^{nx} (\zeta_j - 1) \right)^2 \right] + \frac{1}{n} \mathbb{E}[Y^2] = x\sigma^2 + \frac{1}{n} \mathbb{E}[Y^2]$$

notemos que la igualdad anterior se cumple gracias a la independencia de $\{\zeta_j\}_{j \geq 1}$ y a que $\mathbb{E}[\zeta] = 1$.

□

A continuación se presentan un resultado que es consecuencia de la convergencia anterior y de la aplicación del *Teorema de Mapeo Continuo*:

Lema A.3.1. *Para $f \in C[0, 1]$, $h(f) = \inf \{t > 0 : f(t) > 1\}$ es una función continua bajo la métrica uniforme, la cual ésta dada por $d_k(f, g) = \sup_{t \leq k} |f(t) - g(t)|$ con $g \in C[0, 1]$.*

Demostración. Por definición de $h(\cdot)$, digamos que $\tau = h(t)$. Sea $\epsilon > 0$, como una función continua alcanza su máximo en un intervalo compacto, entonces en $[0, \tau - \epsilon]$, $f(t) \leq \max\{f(t)\} < 1$ así existe $\delta > 0$ tal que $f(t) \leq 1 - \delta$ para $t \leq \tau - \epsilon$ y $f(\tau + \epsilon) \geq 1 + \delta$, entonces afirmamos que para cualquier $g \in C[0, 1]$ con $d_{\tau+\epsilon}(f, g) < \delta$ se cumple $|h(f) - h(g)| < \epsilon$.

Veamos que se cumple la afirmación anterior:

Como $d_{\tau+\epsilon}(f, g) < \delta$ entonces $-\delta < f(\tau + \epsilon) - g(\tau + \epsilon) < \delta$ y $-\delta < f(t) - g(t) < \delta$ para $t \leq \tau - \epsilon$ y de nuevo, afirmamos que $g(\tau + \epsilon) \geq 1$ y $g(t) < 1$ para $t \leq \tau - \epsilon$ lo cual implica que $h(g) > \tau - \epsilon$ y $h(g) < \tau - \epsilon$.

Observemos que en los casos contrarios se generan dos contradicciones; si $g(\tau + \epsilon) < 1$ como $f(\tau + \epsilon) \geq 1 + \delta$ entonces $f(\tau + \epsilon) - g(\tau + \epsilon) > \delta$ ($\rightarrow \leftarrow$) y si $g(t) > 1$ para $t \in [0, \tau - \epsilon]$ como $f(t) \leq 1 - \delta$ entonces $f(t) - g(t) < -\delta$ ($\rightarrow \leftarrow$). □

Análogamente, se cumple que $h_1(f) = \inf \{t > 0 : |f(t)| > 1\}$ es una función continua bajo la métrica anterior.

Distribución del Máximo

Recordemos que en el Capítulo 3, definimos $T_n = \sum_{i=1}^n (\zeta_i - m)$, veamos que $\frac{1}{\sqrt{n}} \max_{1 \leq i \leq n} |T_i|$ converge en distribución:

Sea $\tau_b = \inf \{n \geq 1 : |T_n| > b\}$ luego para $x \in \mathcal{C}$, siendo \mathcal{C} el conjunto de los puntos de continuidad de la función de distribución buscada, tenemos que

$$\mathbb{P} \left[\max_{1 \leq i \leq n} |T_i| > \sqrt{nx} \right] = \mathbb{P} \left[\tau_{\sqrt{nx}} \leq n \right] = \mathbb{P} \left[\frac{1}{(\sqrt{nx})^2} \tau_{\sqrt{nx}} < \frac{1}{x^2} \right],$$

además,

$$\begin{aligned} \frac{\tau_b}{b^2} &= \frac{1}{b^2} \inf \{n \geq 1 : |T_n| > b\} \\ &\approx \inf \{m \geq 1 : |T_{[mb^2]}| > b\} \\ &= \inf \left\{ m \geq 1 : \left| \frac{T_{[mb^2]}}{b} \right| > 1 \right\} \end{aligned}$$

Como consecuencia de la aproximación por difusión de *Feller*, por el *Teorema de Mapeo Continuo* obtenemos que

$$\inf \left\{ m \geq 1 : \left| \frac{T_{[mb^2]}}{b} \right| > 1 \right\} \xrightarrow{d} \inf \{m > 0 : |B_m| > 1\},$$

siendo B_m el Movimiento Browniano.¹

Ahora, si definimos $h_2(f) = \inf \left\{ t > 0 : \int_0^t f(s) ds > 1 \right\}$, y reescribimos a $h_2(\cdot)$ como $h_2(f) = \inf \{t > 0 : p(s) > 1\}$ siendo $p(s) = \int_0^t f(s) \partial s$, notemos que por el *Teorema Fundamental del Cálculo* la función $p(\cdot)$ es continua y aplicando el Lema A.3.1, obtenemos que $h_2(\cdot)$ también es continua.

Teorema A.3.2. Para $N_c = \inf \{n \geq 1 : \sum_{i=1}^n Z_{i-1} \geq c\sigma^2\}$, $c > 0$,

$$\frac{N_c}{c^{1/2}} \xrightarrow{d} \inf \left\{ t > 0 : \int_0^t W(s) ds = 1 \right\},$$

donde $W(t)$ es un proceso de difusión no-negativo con un generador, el cual se obtiene como límite débil del proceso $W_n(t) = \frac{1}{n} Z_{[nt]}$ cuando $n \rightarrow \infty$.

¹Observemos que el símbolo \approx indica la aproximación por difusión puesto que aunque las trayectorias del Mvto. Browniano en $[0, 1]$ son continuas también son muy irregulares.

Demostración. Notemos que

$$\begin{aligned}\frac{N_c}{c^{1/2}} &= \frac{1}{c^{1/2}} \inf \left\{ n \geq 1 : \sum_{i=1}^n Z_{i-1} \geq c\sigma^2 \right\} \\ &\approx \inf \left\{ m \geq 1 : \sum_{i=1}^{\lfloor mc^{1/2} \rfloor} Z_{i-1} \geq c\sigma^2 \right\} \\ &= \inf \left\{ m \geq 1 : \frac{1}{c\sigma^2} \sum_{i=1}^{\lfloor mc^{1/2} \rfloor} Z_{i-1} \geq 1 \right\}\end{aligned}$$

Ahora, usando el Lema A.3.1 y aplicando el *Teorema de Mapeo Continuo* se sigue el resultado. \square

Bibliografía

- [1] Aldous, D. J. and Eagleson, G. K. (1978). *On mixing and stability of limits theorems*. Ann. Prob. **6**, 325-351.
- [2] Asmussen, S. y Hering, H. (1976). *Strong limit theorems for general supercritical branching processes with applications to branching diffusions*. Z. Wahr. **36**, 195-212.
- [3] Athreya, K. B. and Ney, P. E. (1972). *Branching Process*. Springer, New York.
- [4] Bartlett, M. S. (1955) *An introduction to stochastic processes*. Cambridge Univ. Press, Cambridge.
- [5] Bhat, B. R. and Adke, S. R. (1981). *Maximum likelihood estimation for branching processes with immigration*. Advances in Applied Probability, **13**, No. 3, 498-509.
- [6] Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- [7] Chow, Y. S. and Teicher, H. (1978). *Probability Theory: Independence, Interchangeability, Martingales*. Springer, New York.
- [8] Datta, S. and Sriram, T. N. (1995). *A modified bootstrap for Branching processes with immigration*. Stochastic Processes and their Applications, **56**, 275-294.
- [9] Either, S. N. and Kurtz, T. G. (1986) *Markov Processes: Characterization and Convergence*. Wiley, New York.
- [10] Feller, W. (1957) *An Introduction to Probability Theory and Its Applications*. **1**, Second edition. New York: John Wiley and Sons.
- [11] Foster, J. *Branching process involving immigration* Ph. D. Thesis, Univ. of Wisconsin.
- [12] Foster, J. and Williamson, J. A. (1971) *Limit theorems for the Galton-Watson process with time-immigration*. ZW **20**, 227-235.

- [13] Gonzalez M. and Molina M. (1993). *On the limit behaviour in Galton-Watson branching processes with immigration*. Revista Brasileira de Probabilidade e Estadística. **7**, 121-134.
- [14] Gonzalez M., Molina M. and Mota M. *Bayesian Estimation for Galton-Watson with Immigration*. Publicación en el Departamento de Matemáticas, Universidad de Extremadura, España.
- [15] Guttorp, P. (1991) *Statistical Inference for Branching Processes*. John Wiley Sons.
- [16] Grimmett G. R. and Stirzaker D. R. (2002) *Probability and Random Processes*. OXFORD University Press, New York.
- [17] Harris T. E. (1963) *The Theory of Branching Processes*. Berlin: Springer.
- [18] Heathcote, C. R. *A branching process allowing immigration*. JRSS **27**, 138-143.
- [19] Heyde, C. C. (1970) *Extension of a result of Seneta for the supercritical Galton-Watson process*. Ann. Math. Statist. **41**, 739-742.
- [20] Heyde, C. C. and Seneta, E. (1971) *Analogues of classical limit theorems for the supercritical Galton-Watson process with immigration*. Math. Biosci. **11**, 249-259.
- [21] Heyde, C. C. and Seneta, E. (1972) *Estimation theory for growth and immigration rates in a multiplicative process*. J. Appl. Probab. **9**, 235-258.
- [22] Heyde, C. C. and Seneta, E. (1974) *Notes on 'Estimation theory for growth and immigration rates in a multiplicative process'*. J. Appl. Probab. **11**, 572-577.
- [23] Jagers, P. (1975) *Branching Processes with Biological Applications*. London: John Wiley Sons.
- [24] Kalbfleisch, J. G. *Probability and Statistical Inference*. **2**. Springer-Verlag. New York.
- [25] Karlin, S. (1966) *A first course in stochastic processes*. Academic Press, New York, London.
- [26] Kendall, D. G. (1975) *The genalogy of genealogy: branching processes before (and after) 1873*. Bull. London Math. Soc. **7** part 3 (21), 225-253.
- [27] Klimko, L. A. and Nelson, P. I. (1978) *On conditional least squares estimation for stochastic processes*. Ann. Statist. **6**, 629-642.

- [28] Lai, T. L. *Sequential analysis: some classical problems and new challenges*. *Statistica Sinica*, **11**, 303-408.
- [29] Lai, T. L. and Siegmund, D. (1983) *Fixed accuracy estimation of an autoregressive parameter*. *Ann. Statist.* **6**, 629-642.
- [30] Lambert, A. *Population Dynamics and Random Genealogies*. Notas del curso Dinámica de poblaciones, dictado en IX Simposio de Probabilidad y Procesos Estocásticos, Noviembre de 2006.
- [31] LeCam, L. *On some asymptotic properties of maximum likelihood estimates and related Bayes estimates*. *Univ. California Publications* **1**, 277-330.
- [32] Lele, S., Dennis, B. and Lutscher, F. (2007). *Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods*. *Ecology Letters*, **10**, 551-563.
- [33] Marsden, J. F. *Elementary Classical Analysis*. San Francisco, CA: W. H. Freeman.
- [34] Nelson, P. I. (1980) *A note on strong consistency of least squares estimators in regression models with martingale difference errors*. *Ann. Statist.* **8**, 1057-1064.
- [35] Pakes, A. G. (1972). *Further results on the critical Galton-Watson process with immigration*. *J. Austral. Math. Soc.* **13**, 277-290.
- [36] Qi, Y. and Reeves J. (2002) *On sequential estimation for branching processes with immigration*. *Stochastic Processes and their Applications* **100**, 41-51.
- [37] Resnick, S. I. (1998) *A Probability Path*. Springer Verlag, New York.
- [38] Robbins, H. E. and Siegmund D. (1971) *A convergence theorem for nonnegative almost super martingales and some applications*. In: J. S. Rustagi, *Optimizing Methods in Statistics*. Academic, New York, 223-258.
- [39] Sankaranarayanan, G. (1989) *Branching processes and its estimation theory*. Wiley Eastern Limited, New Delphi, India.
- [40] Seneta, E. (1967). *The Galton-Watson process with mean one*. *JAP* **4**, 489-495.
- [41] Seneta, E. (1970). *A note on the supercritical Galton-Watson process with immigration*. *Math. Biosci.* **6**, 305-311.
- [42] Shete, S. and Sriram T. N. (1998) *Fixed precision estimator of the offspring mean in Branching Processes*. *Stochastic Process. Appl.* **77**, 17-33.

- [43] Sriram, T. N., Basawa, I. V. and Huggins, R. M. (1991) *Sequential estimation for branching processes with immigration*. Ann. Statist. 19, 2232-2243.
- [44] Sriram, T. N. (1994) *Invalidity of bootstrap for critical branching processes with immigration*. Ann. Statist. **22**, 1013-1023.
- [45] Govindaradajulu, Z. (2004) *Sequential Statistics*
- [46] Venkataraman, K. N. (1982) *A limit theorem on a subcritical Galton-Watson process with immigration*. Ann. Probab. **10**, 1069-1074.
- [47] Wei, C. Z. and Winnicki, J. (1990). *Estimation of the means in the branching process with immigration*. Annals of statistics. **18**, No. 4, 1757-1773.
- [48] Wei, C. Z. and Winnicky, J. (1987) *A unified estimation theory for the branching process with immigration*. Technical Report. Uni. Maryland.
- [49] Wei, C. Z. and Winnicky, J. (1989) *Some asymptotic results for the branching process with immigration*. Stochastic Process Appl. **31**, 231-282.