

*Frecuencias relativas y estimación de parámetros en el
proceso de Bienaymé - Galton - Watson multitypo*

CAMILO JOSÉ TORRES JIMÉNEZ
MATEMÁTICO



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
JUNIO DE 2010

*Frecuencias relativas y estimación de parámetros en el
proceso de Bienaymé - Galton - Watson multitypo*

CAMILO JOSÉ TORRES JIMÉNEZ
MATEMÁTICO

TRABAJO DE TESIS PARA OPTAR AL TÍTULO DE
MAESTRÍA EN CIENCIAS - ESTADÍSTICA

DIRECTOR
LILIANA BLANCO CASTAÑEDA, DR. RER. NAT.
DOCTOR RERUM NATURALIUM IN MATHEMATIK



UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE ESTADÍSTICA
BOGOTÁ, D.C.
JUNIO DE 2010

Título en español

Frecuencias relativas y estimación de parámetros en el proceso de Bienaymé - Galton - Watson multitipo

Title in English

Relative frequencies and parameter estimation in multi-type Bienaymé - Galton - Watson processes

Resumen: Se exponen los procesos de ramificación de Bienaymé - Galton - Watson multitipo, incluyendo su definición, la matriz de medias, las matrices de varianzas y covarianzas y la definición de las frecuencias relativas asociadas al proceso. Se hace una revisión sobre los estimadores empírico y máximo verosímil de la matriz de medias y los estimadores de las matrices de varianzas y covarianzas asociados, con su comportamiento asintótico cuando el tiempo tiende a infinito. Se proponen dos subgrupos de este tipo de procesos, caracterizados por cierto tipo de distribuciones de descendencia. Se plantean e implementan tres algoritmos distintos para realizar simulaciones. Se ilustran los resultados teóricos expuestos, mediante los resultados numéricos obtenidos por simulación para tres situaciones que se pueden modelar por medio de estos procesos. Por último, se proponen condiciones necesarias para la convergencia de las frecuencias relativas y de los estimadores de la matriz de medias, cuando el número de individuos iniciales y el tiempo tienden a infinito simultáneamente.

Abstract: We present some characteristics of Bienaymé - Galton - Watson multi-type processes, including its definition, the means matrix, the variance and covariance matrices, the relative frequencies definition. A review on the empirical estimator and the maximum likelihood estimator of the reproduction law means and the related variance and covariance matrices estimators, with its asymptotic behaviour when the time tend to infinity. We propose two subgroups of this type of process, characterized by certain types of offspring distributions, and implement three different algorithms for simulations. Theoretical results are illustrated by the numerical results obtained from simulation for three situations can be modeled through these processes. Finally, we propose conditions for the convergence of relative frequencies and means estimators, when the initial number of individuals and time tend to infinity simultaneously.

Palabras clave: procesos estocásticos, procesos de ramificación multitipo, proceso de Bienaymé - Galton - Watson multitipo, comportamiento asintótico, propiedades de estimadores, frecuencias relativas

Keywords: stochastic processes, multi-type branching processes, multi-type Bienaymé - Galton - Watson processes, limit theorems, estimators properties, relative frequencies

Nota de aceptación

Trabajo de tesis
APROBADO

Jurado
Edilberto Cepeda Cuervo

Jurado
Fabio Humberto Nieto Sánchez

Jurado
José Alberto Vargas Navas

Director
Liliana Blanco Castañeda

Bogotá, D.C., Junio 23 de 2010

Dedicatoria

A mis padres,
por sus enseñanzas,
que aún hoy en día,
se siguen ramificando.

Agradecimientos

A Diana, Carlos, Flor Alba, Nelson, Wilmer, Mónica, Dagoberto, Andrés, Guillermo, Claudia, a todos mis compañeros y amigos, con los que trabajamos para obtener los mejores resultados en nuestros cursos de la maestría, gracias a ellos por su colaboración y apoyo.

A los profesores: Luis Alberto López, Oscar Melo, José Alberto Vargas, Luz Mery González y Campo Elías Pardo, quienes con sus clases, me hicieron sentir que había tomado una excelente decisión al ingresar a la Maestría en Estadística, y gracias a sus enseñanzas pude entender varios conceptos que se requirieron para el desarrollo de este trabajo. Quiero destacar al profesor Campo Elías Pardo por sus correcciones en los textos y su motivación a que las implementaciones que fuera a realizar, no se perdieran en un anaquel, sino que fueran visibles y estuvieran disponibles para su uso o modificación.

A Mónica María Peña y a mi familia, por su paciencia, comprensión y apoyo, a ellos les debo la estabilidad que me permitió concentrarme en mis estudios.

Por último, mis mayores agradecimientos, a la profesora Liliana Blanco, por su tiempo, su disposición y su colaboración, fue una fortuna para mi, que aceptará ser mi directora de tesis, su apoyo es incondicional e invaluable, simplemente no hay palabras suficientes para describir su calidad como profesora y como directora de tesis.

Índice general

Índice general	I
Índice de tablas	III
Índice de figuras	IV
Introducción	VI
1. Procesos de ramificación	1
1.1. Proceso de ramificación de Bienaymé - Galton - Watson	1
1.1.1. Comportamiento asintótico.	3
1.2. Proceso de ramificación de Bienaymé - Galton - Watson multitypo (BGWM)	4
1.2.1. Comportamiento asintótico.	6
1.3. Otros Procesos de Ramificación	8
1.4. Estimación de parámetros para procesos de BGWM	8
1.4.1. Estimación de la matriz de medias	11
1.4.2. Estimación de las matrices de varianzas y covarianzas	12
2. Frecuencias relativas en procesos de BGWM	14
2.1. Preliminares	15
2.2. Resultados de Yakovlev y Yanev (2009)	16
2.3. Aplicación del teorema anterior	22
3. Algoritmos y ejemplos numéricos	24
3.1. Caracterización de algunos procesos de BGWM	25
3.1.1. Con distribuciones independientes.	25

3.1.2. Con distribución multinomial.	26
3.2. Algoritmos e implementación	27
3.3. Ejemplos numéricos	30
3.3.1. Estimación de matriz de medias y matrices de covarianzas	34
3.3.2. Comportamiento asintótico de las frecuencias relativas	41
4. Resultados empíricos	55
4.1. Resultados para estimadores de la matriz de medias	56
4.2. Resultados para las frecuencias relativas	60
A. Notación	62
B. Conceptos Previos	65
C. Paquete <i>Branching</i>	67
C.1. Instalación de <i>R</i>	68
C.2. Instalación del paquete <i>Branching</i>	69
C.3. Documentación del paquete <i>Branching</i>	70
rBGWM	70
BGWM.mean	74
BGWM.mean.estim	76
BGWM.covar	78
BGWM.covar.estim	81
Conclusiones	84
Propuestas de investigación	85
Bibliografía	86

Índice de tablas

3.1. Parámetros de las distribuciones de descendencia, para los modelos A, B y C.	31
3.2. Matrices de medias y mayor valor propio respectivo, para los modelos A, B y C. . .	31
3.3. Media muestral del tiempo de extinción para el modelo A (tamaño de muestra = 1000).	34
3.4. Varianza muestral del tiempo de extinción para el modelo A (tamaño de muestra = 1000).	34
3.5. Parámetros de las distribuciones de descendencia, para los modelos D, E y F.	42
3.6. Matrices de medias y mayor valor propio respectivo, para los modelos D, E y F. . .	42
3.7. Vectores $\mathbf{m}(3; \mathbf{e}_1)$ y $\mathbf{p}(3; \mathbf{e}_1)$, para los modelos D, E y F.	42
3.8. Distribuciones de descendencia seleccionadas, para los modelos D, E y F.	42
3.9. Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo D - Caso subcrítico (tamaño de muestra = 1000).	45
3.10. Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo E - Caso crítico (tamaño de muestra = 1000).	46
3.11. Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo F - Caso supercrítico (tamaño de muestra = 1000).	47
3.12. Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000).	50
3.13. Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000).	52
3.14. Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000).	54
4.1. Relación entre N_1 y n , para los modelos D, E y F.	56
4.2. Vector propio a izquierda de la matriz \mathbf{m} , asociado a ρ , para los modelos D, E y F, normalizado con la norma euclidiana y la norma Manhattan.	60

Índice de figuras

3.1. Trayectorias de los tipos de individuos a través del tiempo para el Modelo A.	32
3.2. Trayectorias de los tipos de individuos a través del tiempo para el Modelo B.	32
3.3. Trayectorias de los tipos de individuos a través del tiempo para el Modelo C.	33
3.4. Trayectorias del total de individuos de la población e histograma del tiempo de extinción para el modelo A. Tamaño de muestra = 1000 (Prueba 1 - con distribuciones independientes).	34
3.5. Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias cuando el tiempo crece y el vector de individuos iniciales permanece fijo.	38
3.6. Convergencia del estimador $\check{V}_k(n; \mathbf{N})$ cuando el tiempo crece y el vector de individuos iniciales permanece fijo.	39
3.7. Convergencia del estimador $\hat{V}_k(n; \mathbf{N})$ cuando el tiempo crece y el vector de individuos iniciales permanece fijo.	40
3.8. Convergencia de las frecuencias relativas cuando el tiempo permanece fijo y el número de individuos iniciales de tipo 1 crece.	43
3.9. Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo D - Caso subcrítico (tamaño de muestra = 1000).	45
3.10. Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo E - Caso crítico (tamaño de muestra = 1000).	46
3.11. Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo F - Caso supercrítico (tamaño de muestra = 1000).	47
3.12. Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).	49
3.13. Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000).	50
3.14. Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).	51

3.15. Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000).	52
3.16. Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).	53
3.17. Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000).	54
4.1. Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo D - caso subcrítico ($\lceil N_1 = 5000 * n / \rho^n \rceil$).	57
4.2. Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo E - caso crítico ($N_1 = 10 n^2$).	58
4.3. Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo F - caso supercrítico ($N_1 = n$).	59
4.4. Convergencia de las frecuencias relativas cuando el tiempo y el número de individuos iniciales de tipo 1 crece, para los modelos D, E y F.	61

Introducción

Un proceso de ramificación se puede ver como una representación matemática del comportamiento de una población en términos de la cantidad de individuos que la componen, en donde cada uno de ellos se reproduce y muere de acuerdo a un esquema aleatorio. Dichos procesos, además de plantear una serie de desarrollos teóricos interesantes, tiene un buen número de aplicaciones prácticas. Se podría pensar que la utilidad de estos modelos se limita a temas relacionados con poblaciones biológicas, pero en realidad, tienen aplicación en una gran variedad de contextos, como lo son: biología molecular, biología celular, inmunología, evolución, medicina, ecología, economía, física de partículas, demografía, entre otros (Epps 1996, Kimmel & Axelrod 2002, Dorman, Sinsheimer & Lange 2004, Mitov, Rachev, Kim & Fabozzi 2008, Seidl, Etinski, Uiberacker & Jakubetz 2008, Laredo, David & Garnier 2009).

Históricamente el interés principal de los investigadores con respecto a los procesos de ramificación ha sido analizar el comportamiento asintótico de la población, cuando el tiempo tiende a infinito. Un proceso de ramificación está determinado por una o varias funciones de distribución, para poder conocerlas a partir de información experimental, es necesaria la construcción de estimadores de los parámetros que las caracterizan y sus propiedades asintóticas.

La información experimental u observada asociada a un modelo de ramificación, en donde se tienen diferentes tipos de individuos, puede ser:

- Un conteo del número de descendientes de cada tipo, producidos por cada tipo de ancestro.
- El número total de individuos de cada tipo que componen la población.
- La proporción de individuos de cada tipo con respecto al total de individuos de la población.

para cada momento en el tiempo. Dependiendo del tipo de información obtenida se debe contar, por lo menos, con un estimador asociado, para poder determinar los parámetros del modelo.

Los procesos de ramificación, cómo se ha mencionado, tienen un amplio número de aplicaciones, sin embargo, la teoría estadística desarrollada para este tipo de modelos es relativamente escasa. Yakovlev & Yanev (2009) presentan un método de estimación para experimentos en los cuales sólo es posible observar las frecuencias relativas por tipo, obtenidos a partir de algunos resultados asintóticos cuando el número de individuos iniciales tiende a infinito mientras el tiempo permanece fijo, y proponen investigar el comportamiento de las frecuencias relativas cuando el número de individuos iniciales y el tiempo tienden a infinito de manera simultanea.

El objetivo de este trabajo, es analizar en un proceso de Bienaymé - Galton - Watson multitypo,

el comportamiento asintótico de algunos estimadores de los parámetros del proceso y de las frecuencias relativas de cada tipo de individuo, cuando el tiempo y el número de individuos iniciales tienden a infinito simultáneamente. Para tal fin se realizó una revisión bibliográfica, seleccionando un conjunto de artículos de interés, analizando y verificando los resultados de los artículos por medio de simulaciones, luego, se observó cómo se relacionan dichos resultados con la estimación de los parámetros del proceso, y finalmente, a partir de un conjunto de datos simulados asociados a una situación modelada mediante un proceso de ramificación de Bienaymé - Galton - Watson multitypo se realizó una exploración para posibles nuevos resultados teóricos.

Este documento está organizado de la siguiente manera:

En el Capítulo 1 se estudiaron los procesos de ramificación de Bienaymé - Galton - Watson y de Bienaymé - Galton - Watson multitypo, se revisaron varios temas y enfoques de interés para este tipo de procesos, inicialmente sus definiciones y sus características básicas, detallando los resultados del caso simple como introducción y luego enfocándose en los resultados similares para el caso multitypo. La fuente principal de consulta fueron Harris (1963), Mode (1971) y Athreya & Ney (1972).

En el Capítulo 1.4 se mencionan: los estimadores empírico y máximo verosímil de la matriz de medias del proceso, algunos resultados asintóticos cuando el tiempo tiende a infinito y los estimadores, obtenidos a partir de los anteriores resultados, para las matrices de varianzas y covarianzas con su respectiva convergencia. La fuente principal de consulta fue Maaouia & Touati (2005).

En el Capítulo 2 se describen con detalle los resultados teóricos relacionados con la estimación de parámetros del proceso a partir de los resultados asintóticos de las frecuencias relativas asociadas al mismo, cuando el tiempo se deja fijo y el número de individuos iniciales de tipo 1 tiende a infinito. La fuente principal de consulta fue Yakovlev & Yanev (2009).

En el Capítulo 3, a partir de los resultados de los capítulos anteriores, como parte de los desarrollos propios de este trabajo, se propone una caracterización o tipificación de ciertos procesos de Bienaymé - Galton - Watson multitypo y se desarrollan e implementan los algoritmos necesarios para su simulación, luego, al plantear y simular dos situaciones experimentales modeladas con este tipo de procesos, se obtienen resultados numéricos, que efectivamente coinciden con los resultados teóricos estudiados.

En el Capítulo 4, se realizan las simulaciones correspondientes a la exploración del comportamiento asintótico de los estimadores trabajados en los capítulos anteriores y de las frecuencias relativas de cada tipo de individuo, cuando el tiempo y el número de individuos iniciales tienden a infinito simultáneamente, y a partir de lo obtenido en la simulaciones se proponen las condiciones necesarias para su convergencia.

En el Apéndice A se resume la notación que fue utilizada constantemente a lo largo del documento.

En el Apéndice B se mencionan algunos resultados que fueron utilizados o referenciados en las demostraciones que se realizaron en los primeros capítulos.

En el Apéndice C se encuentran las instrucciones para: la instalación del software R, la instalación del paquete `Branching`, y el uso de las funciones que componen dicho paquete, desarrolladas para realizar las tareas de simulación y cálculo requeridas en este trabajo.

Procesos de ramificación

1.1. Proceso de ramificación de Bienaymé - Galton - Watson

Watson & Galton (1875) publicaron la primera representación matemática de un proceso de ramificación, mediante la cual analizaron la extinción de los apellidos de las familias aristocráticas inglesas, por esta razón, a dichos investigadores se debe el nombre del proceso de ramificación más simple, el proceso de Galton - Watson. Según algunos estudios históricos (Heyde & Seneta 1972), Irénée-Jules Bienaymé también aportó al planteamiento de dicho proceso, y aún cuando no hay constancia escrita de dicha contribución, a este también se le suele llamar proceso de Bienaymé - Galton - Watson.

El proceso de Bienaymé - Galton - Watson modela la evolución de una población de partículas o individuos, que comienza en un tiempo inicial $n = 0$ con un número de individuos inicial $Z(0; N) = N$, cada uno de los cuales (después de una unidad de tiempo), genera de manera independiente un número aleatorio de descendientes, de acuerdo, con una función dada de probabilidad $p(k) = \{p_k, k = 0, 1, 2, \dots\}$ la cual indica las probabilidades de producir k hijos. Por lo tanto, el número total de individuos de la primera generación ($n = 1$), es la suma de $Z(0; N)$ variables aleatorias $Y^{(i)}, i = 1, 2, \dots, Z(0; N)$, con función de probabilidad $p(k) = \{p_k, k = 0, 1, 2, \dots\}$. De esta manera, los hijos de los individuos iniciales constituirán la primera generación, sus nietos la segunda, y así sucesivamente. La variable aleatoria $Z(n; N)$ representa el número de individuos en la n -ésima generación, la cual es igual a la suma de los hijos de los individuos de la $(n - 1)$ -ésima generación. En este modelo se asume que el número de hijos o descendientes producidos por un individuo de una generación, es independiente de la historia familiar y de los demás individuos.

Teniendo en cuenta lo anterior, se tiene la siguiente definición formal (Athreya & Ney 1972, Harris 1963):

Definición 1.1 (Proceso de Bienaymé - Galton - Watson (BGW)). Un proceso de Bienaymé - Galton - Watson (BGW) es una cadena de Markov, homogénea, definida en los enteros no negativos $(Z(n; N))_{n \in \mathbb{N}_0}$, con matriz de transición definida, en términos de una función de probabilidad dada $p(k) = \{p_k \geq 0; k = 0, 1, 2, \dots\}$, $\sum_k p_k = 1$, por

$$\pi_{ij} = P(i, j) = P(Z(n + 1; N) = j | Z(n; N) = i) = \begin{cases} p_j^{*i} & \text{si } i \geq 1, j \geq 0, \\ \delta_{ij} & \text{si } i = 0, j \geq 0, \end{cases} \quad (1.1)$$

con δ_{ij} el delta de Kronecker y p_k^{*i} , $k = 0, 1, 2, \dots$ la i -ésima convolución de $p_k \geq 0$, $k = 0, 1, 2, \dots$.

A partir de la definición, es claro que el proceso de BGW está determinado por la función de probabilidad $p(k) = p_k \geq 0$; $k = 0, 1, 2, \dots$, la cual generalmente se llama distribución de descendencia.

Por otro lado, el número de individuos en la n -ésima generación de un proceso de BGW que tiene N individuos iniciales, es igual a la suma del número de individuos en la n -ésima generación de N procesos idénticos e independientes que tienen 1 individuo inicial,

$$Z(n; N) = \sum_{j=1}^N Z(n; 1)$$

por lo tanto, para los desarrollos teóricos, se suele asumir sin pérdida de generalidad, que el número de individuos iniciales es igual a 1.

Observación 1.1. Dado que el proceso de BGW es una cadena de Markov homogénea, las probabilidades de transición no dependen de n . Teniendo eso en cuenta se tendrá la siguiente notación:

$$\pi_{ij} := P(Z(n+1) = j | Z(n) = i) = P(Z(1) = j | Z(0) = i)$$

$$\pi_{ij}^{(n)} := P(Z(n) = j | Z(0) = i)$$

donde π_{ij} es la probabilidad de transición en un paso y $\pi_{ij}^{(n)}$ es la probabilidad de transición en n pasos.

Notación 1.1. Cuando se quiere resaltar el número inicial de partículas N , el proceso de ramificación se escribe de la siguiente manera:

$$Z(n; N); n = 0, 1, 2, \dots$$

Cuando se asuma que $Z(0) = 1$, a menos que se diga lo contrario, no hay diferencia en escribir $Z(n; 1)$ y $Z(n)$.

Función Generadora

Una herramienta muy útil para el trabajo con procesos estocásticos es la función generadora de probabilidades (f.g.p.), la cual se define como sigue:

$$F(s) = \sum_{k=0}^{\infty} p_k s^k, \quad |s| \leq 1$$

se observa que,

$$\sum_{j=0}^{\infty} \pi_{1j} s^j = F(s); \quad \sum_{j=0}^{\infty} \pi_{ij} s^j = (F(s))^i, \quad i \geq 1$$

por lo tanto,

$$\sum_{j=0}^{\infty} \pi_{1j}^{(n+1)} s^j = \sum_{k=0}^{\infty} \pi_{1k}^{(n)} (F(s))^k,$$

y si se define $F(n; s) := \sum_{j=0}^{\infty} \pi_{1j}^{(n)} s^j$, entonces se tiene que,

$$F(n+1; s) = F(n; F(s)) \quad \text{y} \quad F(1; s) = F(s)$$

con lo cual se concluye que,

$$\sum_{j=0}^{\infty} \pi_{ij}^{(n)} s^j = (F(n; s))^i.$$

Lo que nos indica que la función generadora de probabilidades del proceso en el momento n partiendo del estado i se puede escribir en términos de la función generadora de probabilidades de la distribución de descendencia.

Momentos

Los momentos del proceso, cuando existen, se pueden expresar en términos de las derivadas de $F(s)$ evaluadas en 1. Para la media se tiene que,

$$m := E(Z(1)) = \sum_{j=0}^{\infty} j \pi_{1j} = F'(1)$$

y por la regla de la cadena,

$$E(Z(n)) = \sum_{j=0}^{\infty} j \pi_{1j}^{(n)} = F'(n; 1) = F'(n-1; 1)F'(1) = \dots = (F'(1))^n = m^n. \quad (1.2)$$

De manera similar, usando que $F''(n+1; 1) = F''(1)(F'(n; 1)) + F'(1)F''(n; 1)$, se puede mostrar que,

$$F''(n; 1) = F''(1)(m^{2n-2} + m^{2n-3} + \dots + m^{n-1})$$

y por consiguiente, tomando $\sigma^2 = Var(Z(1))$, se concluye que,

$$Var(Z(n)) = \begin{cases} \frac{\sigma^2 m^{n-1}(m^n-1)}{m-1} & m \neq 1, \\ n\sigma^2 & m = 1 \end{cases} \quad (1.3)$$

1.1.1. Comportamiento asintótico.

Extinción

Para un proceso de ramificación, si existe algún n para el cual $Z(n) = 0$, entonces,

$$Z(n+k) = 0 \quad \text{para todo } k \geq 0$$

lo que indica que 0 es un estado absorbente del proceso. Esto es natural teniendo en cuenta la situación que origina el proceso de BGW, ya que si en algún momento la población se queda sin individuos, no habrá quien tenga descendencia y por lo tanto, a partir de ese momento no se generaran individuos que compongan la población.

Alcanzar dicho n determina la extinción de la población de estudio. Por consiguiente la probabilidad de extinción del proceso $(Z(n))_{n \in \mathbb{N}_0}$ está dada por,

$$\begin{aligned} P(Z(i) = 0 \text{ para algún } i \geq 1) &= \lim_n P(Z(i) = 0 \text{ para algún } 1 \leq i \leq n) \\ &= \lim_n P(Z(n) = 0) \\ &= \lim_{n \rightarrow \infty} F(n; 0) \end{aligned} \quad (1.4)$$

Teorema 1.1. *La probabilidad de extinción del proceso $(Z(n))_{n \in \mathbb{N}_0}$ es la menor raíz, no negativa (q) de la ecuación $t = F(t)$. Toma el valor 1 si $m \leq 1$ y toma un valor estrictamente menor que 1 si $m > 1$.*

Demostración. Sea q la menor solución de $F(t) = t$ en $[0, 1]$. Si $0 \leq s \leq q$ entonces es claro que,

$$0 \leq s \leq F(s) \leq F(q) = q$$

aplicando de nuevo F se obtiene,

$$0 \leq s \leq F(s) \leq F(F(s)) \leq F(q) = q$$

siguiendo este proceso se llega a,

$$0 \leq s \leq F(s) \leq F(2; s) \leq \dots \leq F(n; s) \leq \dots \leq q$$

por lo tanto existe $s_1 := \lim_{n \rightarrow \infty} F(n; s) \leq q$. Como F es continua se sigue que,

$$s_1 = \lim_{n \rightarrow \infty} F(n+1, s) = \lim_{n \rightarrow \infty} F(F(n; s)) = F(s_1)$$

esto es, s_1 es solución de la ecuación $s = F(s)$, como q es la menor solución, se obtiene que $q \leq s_1$. Por consiguiente $q = s_1$. Si $q \leq s \leq 1$ entonces se deduce análogamente que,

$$s \geq F(s) \geq \dots \geq F(n; s) \geq \dots \geq q$$

por lo tanto existe $s_2 := \lim_{n \rightarrow \infty} F(n; s) \geq q$.

Con un razonamiento similar al hecho en la primera parte, se llega a que $s_2 = q$.

Entonces, q es la probabilidad de extinción del proceso, ya que $\lim_{n \rightarrow \infty} F(n; 0) = q$, donde q es la menor solución de $t = F(t)$. Como F es una función convexa y $F(0) > 0$, los gráficos de las funciones $y = F(x)$ y $y = x$ tienen a lo sumo dos puntos en común, uno de ellos es 1, ya que $F(1) = 1$.

Si $m \leq 1$, entonces, a la izquierda de 1, F crece más lentamente que $y = x$, por lo tanto, como F es convexa, no existe una raíz de $F(t) = t$ menor que 1, de modo que $q = 1$. En cambio si $m > 1$, hay una raíz menor que 1, y ésta tiene que ser igual a q , en ese caso, $q < 1$. \square

Criticidad

Se dice que el proceso de Bienaymé - Galton - Watson es subcrítico, crítico o supercrítico, dependiendo de si $m < 1$, $m = 1$ o $m > 1$ respectivamente.

1.2. Proceso de ramificación de Bienaymé - Galton - Watson multitypo (BGWM)

Una posible generalización de un proceso de ramificación, radica en considerar que la población de estudio, está compuesta por individuos que pueden pertenecer a d tipos diferentes. Teniendo esto en cuenta, el desarrollo de la población a través del tiempo queda ahora descrito por el vector aleatorio:

$$(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0} = (Z_1(n; \mathbf{N}), Z_2(n; \mathbf{N}), \dots, Z_d(n; \mathbf{N}))_{n \in \mathbb{N}_0}$$

donde se tiene un vector de individuos iniciales $\mathbf{N} = (N_1, \dots, N_d)$ siendo N_k el número de individuos iniciales del tipo k y $Z_k(n; \mathbf{N}) :=$ “número de individuos de la n -ésima generación pertenecientes al tipo k ”, para $k = 1, 2, \dots, d$ y $n \in \mathbb{N}_0$. El proceso con esta generalización se llama proceso de Bienaymé - Galton - Watson Multitipo.

Cada uno de los individuos iniciales del vector \mathbf{N} (después de una unidad de tiempo), genera de manera independiente un número aleatorio de descendientes, de acuerdo, con un vector de funciones de probabilidad multivariadas $\{\mathbf{p} = (\mathbf{p}_{(1)}, \dots, \mathbf{p}_{(d)})\}$, donde $\mathbf{p}_{(k)}(r_1, r_2, \dots, r_d)$ indica la probabilidad de que un individuo del tipo k produzca r_1 individuos del tipo 1, r_2 individuos del tipo 2, \dots , r_d individuos del tipo d , es decir, corresponde a la distribución de descendencia de los individuos de tipo k . Es claro que, para todo $k = 1, \dots, d$, se tiene:

$$\sum_{(r_1, \dots, r_d) \in \mathbb{N}_0^k} \mathbf{p}_{(k)}(r_1, \dots, r_d) = 1$$

Si $\mathbf{Z}(n; \mathbf{N}) = (r_1, \dots, r_d)$, es decir, en la n -ésima generación hay r_1 individuos del tipo 1, r_2 individuos del tipo 2, \dots , r_d individuos del tipo d , y, si $Y_{(k,j)} :=$ “número de individuos del tipo j que genera un individuo del tipo k ”, entonces, teniendo en cuenta que $Z_j(n+1; \mathbf{N}) :=$ “el número de individuos de tipo j en la $(n+1)$ -ésima generación”, con $j = 1, 2, \dots, d$ y $n \in \mathbb{N}_0$, se tiene que:

$$Z_j(n+1; \mathbf{N}) = \sum_{k=1}^d \sum_{i=1}^{r_k} Y_{(k,j)}^{(i)}$$

De manera análoga al caso del proceso de Bienaymé - Galton - Watson (BGW), se tiene el supuesto, de que los individuos se reproducen independientemente unos de otros e independientemente de la historia del proceso.

Una definición formal para el proceso de Bienaymé - Galton - Watson multitipo es la siguiente (Athreya & Ney 1972):

Definición 1.2 (Procesos de Bienaymé - Galton - Watson multitipo (BGWM)). Un proceso de Bienaymé - Galton - Watson multitipo es una cadena de Markov homogénea, definida en \mathbb{N}_0^d :

$$(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0} = (Z_1(n; \mathbf{N}), Z_2(n; \mathbf{N}), \dots, Z_d(n; \mathbf{N}))_{n \in \mathbb{N}_0}$$

con matriz de transición dada por,

$$\begin{aligned} \pi_{\mathbf{i}\mathbf{j}} &= P(\mathbf{i}, \mathbf{j}) = P(\mathbf{Z}(n+1; \mathbf{N}) = \mathbf{j} | \mathbf{Z}(n; \mathbf{N}) = \mathbf{i}) \\ &= \text{coeficientes de } \mathbf{s}^{\mathbf{j}} \text{ en } (\mathbf{F}(\mathbf{s}))^{\mathbf{i}}, \quad \mathbf{i}, \mathbf{j} \in \mathbb{N}_0^d \end{aligned}$$

donde,

$$\mathbf{F}(\mathbf{s}) = \sum_{\mathbf{r} \in \mathbb{N}_0^d} \mathbf{p}(\mathbf{r}) \mathbf{s}^{\mathbf{r}}, \quad \mathbf{s} = (s_1, s_2, \dots, s_d), 0 \leq s_k \leq 1, k = 1, 2, \dots, d$$

siendo $\mathbf{p}(\mathbf{r}) = (p_{(1)}(\mathbf{r}), p_{(2)}(\mathbf{r}), \dots, p_{(d)}(\mathbf{r}))$, donde $p_{(k)}(\mathbf{r})$ es la distribución de descendencia de los individuos padres de tipo k .

Funciones Generadoras

Obsérvese que $\mathbf{F}(\mathbf{s}) = (F_1(\mathbf{s}), F_2(\mathbf{s}), \dots, F_d(\mathbf{s}))$, donde

$$F_k(s_1, s_2, \dots, s_d) = \sum_{r_1, \dots, r_d \in \mathbb{N}_0} p_k(r_1, r_2, \dots, r_d) s_1^{r_1} s_2^{r_2} \dots s_d^{r_d} \quad 0 \leq s_k \leq 1, k = 1, 2, \dots, d$$

siendo F_k la función generadora de la distribución de descendencia de un individuo de tipo k .

Momentos

- $\mathbf{m}_k := E(\mathbf{Z}(1; \mathbf{e}_k) \mid \mathbf{Z}(0; \mathbf{e}_k))$ para $k = 1, 2, \dots, d$ denota el vector de valores esperados de la primera generación de individuos de los d tipos, dado un solo individuo inicial de tipo k (\mathbf{e}_k es el k -ésimo vector de la base canónica de \mathbb{N}_0^d).
- \mathbf{m} es la matriz de medias del proceso, donde sus d filas corresponden a los d vectores de medias:

$$\mathbf{m} := (m_{ij})_{1 \leq i, j \leq d} = (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_d)'$$
- $\mathbf{V}_k := \text{Var}(\mathbf{Z}(1; \mathbf{e}_k) \mid \mathbf{Z}(0; \mathbf{e}_k))$ para $k = 1, 2, \dots, d$ denota la matriz de varianzas y covarianzas de la primera generación de individuos de los d tipos, dado un solo individuo inicial de tipo k .

Se puede verificar que el vector de valores esperados del número de individuos de la población, en el tiempo n , dada una población inicial de N individuos, está dada por:

$$\begin{aligned} \mathbf{m}(n; \mathbf{N}) &:= E(\mathbf{Z}(n; \mathbf{N}) \mid \mathbf{Z}(0; \mathbf{N})) \\ &= (m_1(n; \mathbf{N}), m_2(n; \mathbf{N}), \dots, m_d(n; \mathbf{N})) \\ &= \mathbf{N} \mathbf{m}^n \end{aligned} \quad (1.5)$$

y que la correspondiente matriz de varianzas y covarianzas en el tiempo n está dada por:

$$\begin{aligned} \mathbf{V}(n; \mathbf{N}) &:= \text{Var}(\mathbf{Z}(n; \mathbf{N}) \mid \mathbf{Z}(0; \mathbf{N})) \\ &= (V_{ij}(n; \mathbf{N}))_{1 \leq i, j \leq d} \\ &= \sum_{l=1}^n \left[(\mathbf{m}^{n-l})' \left(\sum_{k=1}^d m_k(l-1; \mathbf{N}) \mathbf{V}_k \right) \mathbf{m}^{n-l} \right] \end{aligned} \quad (1.6)$$

1.2.1. Comportamiento asintótico.

En esta sección se supondrá que el proceso de BGWM es no singular y regular positivo, conceptos que se exponen a continuación (Athreya & Ney 1972).

Definición 1.3. Un proceso de BGWM se dice que es un proceso singular, si y sólo si, la función generadora del proceso está dada por, $\mathbf{F}(\mathbf{s}) = \mathbf{A} \mathbf{s}'$, donde \mathbf{A} es una matriz no negativa de dimensión $d \times d$.

En otras palabras, un proceso de BGWM singular es aquel en el cual cada individuo de la población tiene exactamente un descendiente. Por otra parte:

Definición 1.4. Se dice que un proceso de BGWM es regular positivo o irreducible, si y sólo si, existe $n_0 \in \mathbb{N}$, tal que, las componentes de la matriz \mathbf{m}^n son positivas, para todo $n \geq n_0$.

Extinción

Los siguientes teoremas presentan los resultados con respecto a la probabilidad de extinción y al comportamiento asintótico de un proceso de BGWM, cuando el tiempo n tiende a infinito. Las demostraciones de dichos resultados se pueden consultar en Harris (1963), Mode (1971) y Athreya & Ney (1972).

Teorema 1.2 (Explosión o extinción del proceso). *Dado un proceso de BGWM*

$$(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0}$$

entonces:

$$P\left(\mathbf{Z}(n; \mathbf{N}) \xrightarrow[n \rightarrow \infty]{} \infty \mid \mathbf{Z}(0; \mathbf{N})\right) + P\left(\mathbf{Z}(n; \mathbf{N}) \xrightarrow[n \rightarrow \infty]{} \mathbf{0} \mid \mathbf{Z}(0; \mathbf{N})\right) = 1$$

para cada $\mathbf{N} \in \mathbb{N}_0^k$.

El anterior teorema concluye que cuando el tiempo n tiende a infinito, sólo existen dos situaciones posibles, o la población se extingue o crece indefinidamente. De manera adicional, el siguiente teorema plantea una condición de extinción para el caso múltiple, similar a la dada para el caso simple:

Teorema 1.3 (Probabilidad de extinción). *Sea $(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0}$ un proceso de BGWM, si ρ es el mayor valor propio de la matriz \mathbf{m} de valores esperados, se tiene que:*

- Si $\rho \leq 1$ entonces $\mathbf{q} = \mathbf{1}$.
- Si $\rho > 1$ entonces $\mathbf{0} \leq \mathbf{q} < \mathbf{1}$.

donde $\mathbf{q} = (q_1, q_2, \dots, q_d)$ y q_k es la probabilidad de la eventual extinción del proceso dado un único individuo inicial de tipo k .

Criticidad

Definición 1.5. Sea $(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0}$ un proceso de BGWM, se dice que el proceso es supercrítico, crítico o subcrítico si $\rho > 1$, $\rho = 1$ o $\rho < 1$ respectivamente, donde ρ es el mayor valor propio de la matriz de valores esperados \mathbf{m} .

Teorema 1.4 (Caso subcrítico). *Si $\rho < 1$ y $E(\|\mathbf{Z}(1; \mathbf{N})\|_{max}^2) < \infty$ entonces, condicionada por $\mathbf{Z}(n; \mathbf{N}) \neq \mathbf{0}$:*

$$\mathbf{Z}(n, \mathbf{N}) \xrightarrow[n \rightarrow \infty]{\text{distribución}} \mathbf{W}$$

y,

$$\mathbf{W} \sim f(\mathbf{z}) = \lim_{n \rightarrow \infty} P(\mathbf{Z}(n, \mathbf{N}) = \mathbf{z} \mid \mathbf{Z}(n; \mathbf{N}) \neq \mathbf{0})$$

límite que existe y es independiente de $\mathbf{Z}(0; \mathbf{N})$.

Teorema 1.5 (Caso crítico). Si $\rho = 1$, $E(\|\mathbf{Z}(1; \mathbf{N})\|_{max}^2) < \infty$ y si \mathbf{w} es un vector tal que $\mathbf{w}' \mathbf{u} > 0$, entonces, condicionada por $\mathbf{Z}(n; \mathbf{N}) \neq \mathbf{0}$:

$$\frac{\mathbf{w}' \mathbf{Z}(n, \mathbf{N})}{n} \xrightarrow[n \rightarrow \infty]{\text{distribución}} W$$

y,

$$W \sim f(x) = \frac{1}{\gamma_1} e^{-\frac{x}{\gamma_1}}, x \geq 0 \text{ donde } \gamma_1 = \frac{\mathbf{u}' \mathbf{w}}{\mathbf{u}' \mathbf{Q}[\mathbf{v}]}$$

\mathbf{u} y \mathbf{v} denotan los vectores propios a izquierda y derecha, correspondientes al valor propio ρ , respectivamente.

$\mathbf{Q}[\mathbf{v}] := (Q^{(1)}[\mathbf{v}], Q^{(2)}[\mathbf{v}], \dots, Q^{(k)}[\mathbf{v}])$, siendo

$$Q^{(k)}[\mathbf{v}] := \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d v_i (q^{(k)}(i, j)) v_j, k = 1, 2, \dots, d.$$

y donde,

$$q^{(k)}(i, j) := \left. \frac{\partial^2}{\partial s_i \partial s_j} F_k(\mathbf{s}) \right|_{\mathbf{s}=\mathbf{1}}, k, i, j = 1, 2, \dots, d.$$

Teorema 1.6 (Caso supercrítico). Si $\rho > 1$, entonces:

$$\frac{\mathbf{Z}(n; \mathbf{N})}{\rho^n} \xrightarrow[n \rightarrow \infty]{\text{casi siempre}} \mathbf{u} W$$

donde W es una variable aleatoria no negativa y \mathbf{u} es el vector propio a izquierda de la matriz de valores esperados \mathbf{m} .

Es así como dependiendo del valor de ρ se tienen tres diferentes comportamientos asintóticos cuando el tiempo n tiende a infinito.

1.3. Otros Procesos de Ramificación

Existen otros procesos de ramificación aparte de los expuestos anteriormente, están los llamados: procesos de ramificación dependientes de la edad, procesos de ramificación dependientes del tamaño de la población, procesos de ramificación en ambientes variables o aleatorios, entre otros (Mode 1971, Athreya & Ney 1972, Kimmel & Axelrod 2002).

Un proceso de ramificación general es un proceso estocástico que busca modelar una población de individuos, en donde, cada individuo tiene un tiempo de vida (aleatorio), dicho individuo puede tener un cierto número de descendientes (aleatorio), en varios momentos del tiempo a lo largo de toda su vida (aleatorios).

1.4. Estimación de parámetros para procesos de BGWM

Asmussen & Keiding (1978) consideraron el problema de estimar la tasa de crecimiento ρ (valor propio dominante de la matriz de medias) de un proceso de BGWM regular positivo y

supercrítico, la estimación fue desarrollada, primero aplicando el teorema central del límite para martingalas, y luego, estimando la matriz de medias por máxima verosimilitud, suponiendo conocidas todas las combinaciones tipo padre- tipo hijo en las primeras n generaciones.

Nanthi (1982) planteó dos estimadores para las matrices de covarianza de un proceso, regular positivo y supercrítico de BGWM y analizó su comportamiento asintótico. El primero, es un estimador por máxima verosimilitud, basado en el conocimiento del número de descendientes por individuo, el segundo, es un estimador que está basado únicamente en los conteos correspondientes a las combinaciones de los tipos padre-hijo.

Integrando tres aspectos diferentes: estimación, comportamiento asintótico y número de individuos iniciales aleatorio, Dion & Yanev (1997) realizan un análisis extenso de las propiedades asintóticas de un proceso de ramificación de BGW, cuando el número de individuos iniciales (aleatorio) y el tiempo tienden a infinito, obteniendo, estimadores consistentes y asintóticamente normales para la media y la distribución de descendencia del proceso.

Maaouia & Touati (2005) realizaron una revisión bibliográfica acerca de los desarrollos de inferencia estadística en cuanto a las matrices de medias y/o matrices de covarianzas de las distribuciones de descendencia en un proceso de BGWM, regular positivo y supercrítico, a partir de la cual desarrollaron regiones de confianza asintóticas para dichas medias y covarianzas.

Sea $(Z(n; N))_{n \in N_0}$ un proceso de BGW supercrítico o lo que es lo mismo, un proceso de BGWM con un solo tipo de individuos ($d = 1$), entonces $m := E(Z(1; 1)) \in (1, \infty)$ es la media y $\sigma^2 := Var(Z(1; 1)) \in (0, \infty)$ es la varianza de la distribución de descendencia del proceso. Los siguientes resultados para la estimación de la media se pueden encontrar en Dion (1974) y Maaouia & Touati (2005).

El *estimador de Lotka-Nagaev o estimador empírico* de m , está dado por,

$$\check{m}(n; N) = \frac{Z(n; N)}{Z(n-1; N)} 1_{\{Z(n-1; N) > 0\}} \quad (1.7)$$

el cual, es fuertemente consistente en el conjunto de no extinción $\mathcal{T} = \{\lim_{n \rightarrow \infty} Z(n; N) = \infty\}$, y se tiene que,

$$(Z(n-1; N))^{1/2} (\check{m}(n; N) - m) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}(0, \sigma^2) \quad (1.8)$$

donde $\mathcal{N}(0, \sigma^2)$ hace referencia a una distribución normal de media 0 y varianza σ^2 .

El *estimador máximo verosímil* de m , está dado por,

$$\hat{m}(n; N) = \frac{Z(1; N) + Z(2; N) + \cdots + Z(n; N)}{Z(0; N) + Z(1; N) + \cdots + Z(n-1; N)} \quad (1.9)$$

el cual también, es fuertemente consistente en el conjunto de no extinción \mathcal{T} o condicionado sobre el conjunto $\mathcal{T}_n = \{Z(n; N) > 0\}$, y además, entre otros resultados se tiene que,

$$\left(\sum_{i=0}^{n-1} Z(i; N) \right)^{1/2} (\hat{m}(n; N) - m) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}(0, \sigma^2) \quad (1.10)$$

A partir de las definiciones de los anteriores estimadores, se observa que sus valores coinciden para $n = 1$.

De manera similar a lo realizado con la estimación de la media, se observan las siguientes resultados para la estimación de la varianza, los cuales se pueden encontrar en Heyde (1974), Dion (1975) y Maaouia & Touati (2005).

A partir del estimador empírico de la media ($\check{m}(n; N)$) se obtiene el siguiente estimador de la varianza,

$$\check{\sigma}^2(n; N) = \frac{1}{n} \sum_{i=1}^n Z(i-1; N) (\check{m}(i; N) - \check{m}(n; N))^2 \xrightarrow[n \rightarrow \infty]{c.s.} \sigma^2 \quad (1.11)$$

sobre el conjunto de no extinción $\mathcal{T} = \{\lim_{n \rightarrow \infty} Z(n; N) = \infty\}$, y adicionalmente, bajo el supuesto de que el momento de orden cuatro de la distribución de descendencia sea finito, entonces,

$$\sqrt{n}(\check{\sigma}^2(n; N) - \sigma^2) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}(0, 2\sigma^4) \quad (1.12)$$

Por otro lado, a partir del estimador máximo verosímil de la media ($\hat{m}(n; N)$) se obtiene el siguiente estimador de la varianza,

$$\hat{\sigma}^2(n; N) = \frac{1}{n} \sum_{i=1}^n \left(\sum_{j=0}^{i-1} Z(j; N) \right) (\hat{m}(i; N) - \hat{m}(n; N))^2 \xrightarrow[n \rightarrow \infty]{c.s.} \sigma^2 \quad (1.13)$$

sobre el conjunto de no extinción \mathcal{T} , y adicionalmente, bajo el supuesto de que el momento de orden cuatro de la distribución de descendencia sea finito, entonces,

$$\sqrt{n}(\hat{\sigma}^2(n; N) - \sigma^2) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}\left(0, 2 \left(\frac{m+1}{m-1}\right) \sigma^4\right) \quad (1.14)$$

En el caso de las ecuaciones (1.12) y (1.14), el momento de orden cuatro de las distribución de descendencia se puede ver como una medida de variabilidad del estimador de la varianza y para obtener la convergencia en distribución de dicho estimador se requiere que su variabilidad sea finita.

Se tienen así, como mínimo, dos estimadores para la media y dos para la varianza, de la distribución de descendencia de un proceso de BGW, en donde se puede destacar que el estimador empírico de la media (1.7) es insensible a cambios en las observaciones asociadas a los tiempos $0, 1, \dots, n-2$ y que el estimador $\check{\sigma}^2(n; N)$ es asintóticamente más eficiente que el estimador $\hat{\sigma}^2(n; N)$, debido a las ecuaciones (1.12) y (1.14).

En el caso de un proceso de BGWM, éste está constituido por el vector aleatorio $\mathbf{Z}(n; \mathbf{N}) = (Z_1(n; \mathbf{N}), Z_2(n; \mathbf{N}), \dots, Z_d(n; \mathbf{N}))$, definido como el número de individuos en la n -ésima generación, a partir de \mathbf{N} individuos iniciales para los d tipos, este vector aleatorio generalmente se asume observable, es decir, que a partir de una situación experimental puedo obtener los valores resultantes asociados a dichas variables aleatorias, en cuyo caso, a un conjunto de observaciones de este tipo a través del tiempo, se le denominará *muestra de totales por tipo*.

Sin embargo, los estimadores que se presentan en este capítulo requieren que sean observables otras variables aleatorias, tales como el número de individuos de tipo j en la n -ésima generación, provenientes de individuos de tipo k de la $(n-1)$ -ésima generación, a partir de \mathbf{N} individuos iniciales, es así que en adelante se supondrá que este grupo de variables aleatorias requeridas son observables, su notación será $Y_{(k,j)}(n; \mathbf{N})$ y a un conjunto de observaciones de la misma, se le denominará *muestra completa*.

Para las siguientes secciones, las definiciones y teoremas con sus respectivas demostraciones, se pueden encontrar en Maaouia & Touati (2005).

1.4.1. Estimación de la matriz de medias

En esta sección se presentan, los estimadores de la matriz de medias de un proceso de BGWM regular positivo y supercrítico, y algunas propiedades asintóticas cuando el tiempo tiende a infinito.

Definición 1.6 (Estimador empírico de la matriz de medias (EE)). El estimador empírico del vector de medias de la distribución de descendencia de los individuos de tipo k , a partir de, los valores observados de $Y_{(k,j)}(n; \mathbf{N})$ en n generaciones, y de los \mathbf{N} individuos iniciales, está dado por,

$$\check{\mathbf{m}}_k(n; \mathbf{N}) = \frac{\mathbf{Y}_k(n; \mathbf{N})}{Z_k(n-1; \mathbf{N})} \quad (1.15)$$

el cual es, el análogo multidimensional del estimador empírico definido en (1.7) y donde,

$$\mathbf{Y}_k(n; \mathbf{N}) = (Y_{(k,1)}(n; \mathbf{N}), Y_{(k,2)}(n; \mathbf{N}), \dots, Y_{(k,d)}(n; \mathbf{N}))$$

Es así como, el *estimador empírico (EE) de la matriz de medias del proceso de BGWM* es,

$$\check{\mathbf{m}}(n; \mathbf{N}) = (\check{\mathbf{m}}_1(n; \mathbf{N}), \check{\mathbf{m}}_2(n; \mathbf{N}), \dots, \check{\mathbf{m}}_d(n; \mathbf{N}))' \quad (1.16)$$

es decir, el vector fila k -ésimo de esta matriz corresponde a (1.15).

Definición 1.7 (Estimador máximo verosímil de la matriz de medias (MLE)). El estimador máximo verosímil del vector de medias de la distribución de descendencia de los individuos de tipo k , a partir de, los valores observados de $Y_{(k,j)}(n; \mathbf{N})$ en n generaciones, y, de los \mathbf{N} individuos iniciales, está dado por,

$$\hat{\mathbf{m}}_k(n; \mathbf{N}) = \frac{\sum_{i=1}^n \mathbf{Y}_k(i; \mathbf{N})}{\sum_{i=0}^{n-1} Z_k(i; \mathbf{N})} \quad (1.17)$$

Es así como, el *estimador máximo verosímil (MLE) de la matriz de medias del proceso de BGWM* es,

$$\hat{\mathbf{m}}(n; \mathbf{N}) = (\hat{\mathbf{m}}_1(n; \mathbf{N}), \hat{\mathbf{m}}_2(n; \mathbf{N}), \dots, \hat{\mathbf{m}}_d(n; \mathbf{N}))' \quad (1.18)$$

es decir, el vector fila k -ésimo de esta matriz corresponde a (1.17).

Maaouia & Touati (2005) demuestran que el estimador máximo verosímil efectivamente es obtenido, al maximizar la función de verosimilitud y maximizando cada componente del k -ésimo vector de medias \mathbf{m}_k . Además, ambos estimadores tienen una idea de proporción, ya que se está considerando el número de individuos que se generaron en la(s) última(s) generación(es) de todos los tipos, a partir de cierta cantidad de individuos de tipo k presentes en la(s) anterior(es) generación(es).

El siguiente teorema establece las condiciones bajo las cuales, los anteriores estimadores de la matriz de medias, son fuertemente consistentes y presentan una distribución normal asintótica, cuando el tiempo n tiende a infinito.

Teorema 1.7 (Propiedades asintóticas de EE y MLE). Sea $(\mathbf{Z}(n; \mathbf{N}))_{n \in \mathbb{N}_0}$ un proceso de ramificación de BGWM, con $d < \infty$ tipos distintos, regular positivo, supercrítico, con vector de individuos iniciales igual a $(1, 1, \dots, 1)$, con momentos de cuarto orden finitos y con matrices de varianzas y covarianzas invertibles, donde $\mathbf{T} = \{\lim_{n \rightarrow \infty} \mathbf{Z}(n; \mathbf{N}) = \infty\}$ designa el conjunto de no extinción del proceso y $\mathbf{T}_n := \cup_{k=1}^d \{Z_k(n, \mathbf{N}) > 0\}$, entonces, se tienen las siguientes propiedades:

- (i) El estimador empírico $\check{\mathbf{m}}(n; \mathbf{N})$ es un estimador fuertemente consistente de la matriz de medias \mathbf{m} sobre el conjunto \mathbf{T} .
- (ii) El estimador empírico $\check{\mathbf{m}}(n; \mathbf{N})$ asintóticamente tiene una distribución normal, condicionado sobre el conjunto \mathbf{T} o \mathbf{T}_n se tiene que,

$$(Z_k(n-1; \mathbf{N}))^{1/2} (\check{\mathbf{m}}_k(n; \mathbf{N}) - \mathbf{m}_k) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}_d(\mathbf{0}, \mathbf{V}_k) \quad (1.19)$$

- (iii) El estimador máximo verosímil $\hat{\mathbf{m}}(n; \mathbf{N})$ es un estimador fuertemente consistente de la matriz de medias \mathbf{m} sobre el conjunto \mathbf{T} .
- (iv) El estimador máximo verosímil $\hat{\mathbf{m}}(n; \mathbf{N})$ asintóticamente tiene una distribución normal, en efecto, condicionado sobre el conjunto \mathbf{T} o \mathbf{T}_n , se tiene que,

$$\left(\sum_{i=0}^{n-1} Z_k(i; \mathbf{N}) \right)^{1/2} (\hat{\mathbf{m}}_k(n; \mathbf{N}) - \mathbf{m}_k) \xrightarrow[n \rightarrow \infty]{dist} \mathcal{N}_d(\mathbf{0}, \mathbf{V}_k) \quad (1.20)$$

donde $\mathcal{N}_d(\mathbf{0}, \mathbf{V}_k)$ hace referencia a una distribución normal multivariada de dimensión d , con vector de medias $\mathbf{0}$ y matriz de varianzas y covarianzas \mathbf{V}_k .

1.4.2. Estimación de las matrices de varianzas y covarianzas

En esta sección se indicarán los estimadores de las matrices de varianzas y covarianzas de un proceso de BGWM regular positivo y supercrítico.

Definición 1.8 (Estimador de las matrices de varianzas y covarianzas a partir de EE). El estimador de la matriz de varianzas y covarianzas de la distribución de descendencia de los individuos de tipo k , calculado a partir del estimador empírico del vector de medias de la misma distribución de descendencia, está dado por,

$$\check{\mathbf{V}}_k(n; \mathbf{N}) = \frac{1}{n} \sum_{j=1}^n Z_k(j-1; \mathbf{N}) (\check{\mathbf{m}}_k(j; \mathbf{N}) - \check{\mathbf{m}}_k(n; \mathbf{N})) (\check{\mathbf{m}}_k(j; \mathbf{N}) - \check{\mathbf{m}}_k(n; \mathbf{N}))' \quad (1.21)$$

Definición 1.9 (Estimador de las matrices de varianzas y covarianzas a partir de MLE). El estimador de la matriz de varianzas y covarianzas de la distribución de descendencia de los individuos de tipo k , calculado a partir del estimador máximo verosímil del vector de medias de la misma distribución de descendencia, está dado por,

$$\hat{\mathbf{V}}_k(n; \mathbf{N}) = \frac{1}{n} \sum_{j=1}^n \left(\sum_{i=0}^{j-1} Z_k(i; \mathbf{N}) \right) (\hat{\mathbf{m}}_k(j; \mathbf{N}) - \hat{\mathbf{m}}_k(n; \mathbf{N})) (\hat{\mathbf{m}}_k(j; \mathbf{N}) - \hat{\mathbf{m}}_k(n; \mathbf{N}))' \quad (1.22)$$

Los estimadores de las matrices de varianzas y covarianzas son obtenidos a partir de los resultados (ii) y (iv) del Teorema 1.7. Además se tiene el siguiente teorema:

Teorema 1.8 (Propiedades asintóticas de los estimadores de las matrices de varianzas y covarianzas). *Bajo las mismas hipótesis del Teorema 1.7, sobre el conjunto de no extinción \mathbf{T} se tiene que,*

- (i) *El estimador $\check{\mathbf{V}}_k(n; \mathbf{N})$ es un estimador fuertemente consistente de la matriz de varianzas y covarianzas (de la distribución de descendencia de los individuos de tipo k) \mathbf{V}_k .*
- (ii) *El estimador $\hat{\mathbf{V}}_k(n; \mathbf{N})$ es un estimador fuertemente consistente de la matriz de varianzas y covarianzas \mathbf{V}_k .*

Frecuencias relativas en procesos de BGWM

Probablemente el primero en considerar las frecuencias relativas asociadas a un proceso de ramificación multitypo fue Jagers (1969), quien estudió las propiedades asintóticas (cuando el tiempo tiende a infinito) de un proceso de ramificación dependiente de la edad, con dos tipos de individuos, suponiendo que los individuos del primer tipo generan individuos de los dos tipos y que los individuos del segundo tipo sólo generan individuos de su propio tipo. Mostró que, bajo ciertas condiciones, la frecuencia relativa asociada a los individuos del primer tipo converge a un valor constante que depende de la distribución de descendencia correspondiente.

A partir de su interés en las aplicaciones biológicas, Mode (1971) considera la utilidad de las frecuencias relativas y menciona que para el proceso de BGWM supercrítico, si el tiempo tiende a infinito entonces las frecuencias relativas convergen a una constante, asociada al vector propio a izquierda del mayor valor propio de la matriz de medias del proceso.

Después de estos avances iniciales, el concepto de frecuencias relativas ha sido poco explorado por parte de los investigadores, tanto desde el punto de vista matemático, como desde el punto de vista de la inferencia estadística, a pesar de tener una clara aplicación en estudios biológicos de proliferación y diferenciación celular, en donde, es posible medir las frecuencias relativas y no los conteos de diferentes tipos de células.

En un artículo reciente, Yakovlev & Yanev (2009), consideran las frecuencias relativas de distintos tipos de individuos en procesos de ramificación multitypo y prueban que dichas frecuencias tienen una distribución asintótica normal multivariada cuando el número de individuos o partículas iniciales es grande y el tiempo de observación es fijo. El resultado que obtienen es interesante, ya que es válido para todo tipo de procesos de ramificación con un número finito de tipos, siempre y cuando, se mantenga la hipótesis de independencia con respecto a la historia del proceso y entre los individuos.

En las siguientes secciones se presentarán los resultados teóricos expuestos por Yakovlev & Yanev (2009), desglosando las demostraciones respectivas.

2.1. Preliminares

Sin pérdida de generalidad se puede asumir que existe al menos un tipo (por ejemplo, el tipo 1) que genera o produce todos los d (número finito de) tipos, y adicionalmente, que el proceso de ramificación comienza con $\mathbf{Z}(0; \mathbf{N}) = \mathbf{N}$ individuos (por ejemplo, $\mathbf{N}_1 = (N_1, 0, \dots, 0)$).

Sea $\mathbf{Z}(t; \mathbf{N}) = (Z_1(t; \mathbf{N}), Z_2(t; \mathbf{N}), \dots, Z_d(t; \mathbf{N}))$ un proceso de ramificación multitipo, donde $Z_k(t; \mathbf{N})$ es el proceso estocástico que corresponde al número de individuos de tipo k ($k = 1, 2, \dots, d$) en el tiempo t . Obsérvese que en esta notación, se resalta la dependencia que tiene el proceso con respecto al número inicial de individuos $Z(0; \mathbf{N}) = \mathbf{N}$. Por otra parte el tiempo puede ser discreto ($t \in N_0 = 0, 1, 2, \dots$) o continuo ($t \in R^+ = [0, \infty)$), y se debe mantener la condición de independencia, entre individuos y con respecto a la historia del proceso.

La llamada propiedad aditiva de los procesos de BGWM dice que:

$$Z_k(t; \mathbf{N}_1) = \sum_{i=1}^{N_1} Z_k^{(i)}(t, \mathbf{e}_1), \quad k = 1, 2, \dots, d, \quad (2.1)$$

donde $Z_k^{(i)}(t; \mathbf{e}_1)$ con $1 \leq i \leq N$ son N copias i.i.d. del proceso de ramificación $Z_k(t; \mathbf{e}_1)$ con $k = 1, 2, \dots, d$, en donde, $\mathbf{Z}(0) = \mathbf{e}_1$, es decir, que cada copia comienza con un individuo de tipo 1.

Tomando $\mathbf{Z}(t) = (Z_1(t), Z_2(t), \dots, Z_d(t)) = (Z_1(t; \mathbf{e}_1), Z_2(t; \mathbf{e}_1), \dots, Z_d(t; \mathbf{e}_1))$, su función generadora de probabilidad (f.g.p.) está dada por:

$$F(t, \mathbf{s}) := F(t, \mathbf{s}; \mathbf{e}_1) = E \left[\mathbf{s}^{\mathbf{Z}(t)} \mid \mathbf{Z}(0) = \mathbf{e}_1 \right] = E \left[s_1^{Z_1(t)} s_2^{Z_2(t)} \dots s_d^{Z_d(t)} \mid \mathbf{Z}(0) = \mathbf{e}_1 \right] \quad (2.2)$$

donde $\mathbf{s} = (s_1, s_2, \dots, s_d)$ y $|s_k| \leq 1, k = 1, 2, \dots, d$.

Por las características propias de un proceso de ramificación y la hipótesis de independencia entre individuos, se tiene que:

$$F_{N_1}(t, \mathbf{s}) := F_{N_1}(t, \mathbf{s}; \mathbf{e}_1) = E \left[\mathbf{s}^{\mathbf{Z}(t)} \mid Z(0) = \mathbf{N}_1 \right] = (F(t, \mathbf{s}))^{N_1} \quad (2.3)$$

Todas las características de un proceso de ramificación se pueden obtener a partir de su f.g.p., los siguientes resultados se tendrán en cuenta con su correspondiente notación:

$$m_k(t) := m_k(t; \mathbf{e}_1) = E [Z_k(t) \mid \mathbf{Z}(0) = \mathbf{e}_1] = \frac{\partial}{\partial s_k} F(t, \mathbf{s}) \Big|_{\mathbf{s}=\mathbf{1}} \quad k = 1, 2, \dots, d, \quad (2.4)$$

$$b_{ij}(t) := b_{ij}(t; \mathbf{e}_1) = \frac{\partial^2}{\partial s_i \partial s_j} F(t, \mathbf{s}) \Big|_{\mathbf{s}=\mathbf{1}} \quad i, j = 1, 2, \dots, d, \quad (2.5)$$

donde $\mathbf{1} = (1, 1, \dots, 1)$.

A partir de (2.4) y (2.5) se obtiene que la matriz de varianzas y covarianzas

$$\mathbf{V}_{(d \times d)}(t) := \mathbf{V}_{(d \times d)}(t; \mathbf{e}_1)$$

está dada por:

$$V_{ij}(t) = \begin{cases} \text{Var}(Z_i(t)) = b_{ii}(t) + m_i(t) - m_i^2(t) & i = j = 1, 2, \dots, d, \\ \text{Cov}(Z_i(t), Z_j(t)) = b_{ij}(t) - m_i(t)m_j(t) & i \neq j = 1, 2, \dots, d, \end{cases} \quad (2.6)$$

Se supondrá que $\sigma_i^2(t) := V_{ii}(t) > 0$, para $i = 1, 2, \dots, d$ y para todo t , evitando así casos triviales y permitiendo definir la matriz de correlación $\mathbf{R}_{(d \times d)}(t) := \mathbf{R}_{(d \times d)}(t; \mathbf{e}_1)$ dada por:

$$r_{ij}(t) = \frac{V_{ij}(t)}{\sigma_i(t) \sigma_j(t)} = \text{Cor}(Z_i(t), Z_j(t))$$

Adicionalmente, se define el proceso estocástico $U(t; \mathbf{N}_1)$ que representa el número total de individuos en el tiempo t , dados N_1 individuos iniciales de tipo 1, $U(0; \mathbf{N}_1) = N_1$. Por lo tanto:

$$\begin{aligned} U(t; \mathbf{N}_1) &= \sum_{k=1}^d Z_k(t; \mathbf{N}_1) \\ &= \sum_{k=1}^d \sum_{i=1}^{N_1} Z_k^{(i)}(t, \mathbf{e}_1) \quad \text{por (2.1)} \end{aligned} \tag{2.7}$$

Las frecuencias relativas o proporciones, de cada tipo, en el momento t , dados N_1 individuos iniciales de tipo 1, están dadas por $\Delta_k(t; \mathbf{N}_1)$, de la siguiente forma:

$$\Delta_k(t; \mathbf{N}_1) = \frac{Z_k(t; \mathbf{N}_1)}{U(t; \mathbf{N}_1)}, \quad k = 1, 2, \dots, d \tag{2.8}$$

en el conjunto de no extinción $\{U(t; \mathbf{N}_1) > 0 : t \in \mathbb{R}^+ \text{ y } N_1 \in \mathbb{N}_0\}$.

En la siguiente proposición se van a analizar algunas convergencias cuando N_1 tiende a infinito, para ello se introduce la siguiente notación. Sea $p_k(t; \mathbf{N}_1)$ definido de la siguiente manera:

$$p_k(t; \mathbf{N}_1) = \frac{m_k(t; \mathbf{N}_1)}{M(t; \mathbf{N}_1)}, \quad k = 1, 2, \dots, d \tag{2.9}$$

donde,

$$\begin{aligned} m_k(t; \mathbf{N}_1) &= E(Z_k(t; \mathbf{N}_1) \mid \mathbf{Z}(0) = \mathbf{N}_1) \\ M(t; \mathbf{N}_1) &= E(U(t; \mathbf{N}_1)) = \sum_{k=1}^d m_k(t; \mathbf{N}_1) \end{aligned} \tag{2.10}$$

2.2. Resultados de Yakovlev y Yanev (2009)

A continuación se enuncia y demuestra una proposición tal que cuando el número inicial de individuos de tipo 1 tiende a infinito ($N_1 \rightarrow \infty$) mientras se deja fijo el tiempo (t), se obtienen tres resultados, el primero, la probabilidad de extinción del proceso tiende a 0, el segundo, las variables aleatorias correspondientes a las frecuencias relativas convergen a un valor constante que depende del número inicial de individuos de tipo 1 y del momento fijo que se consideró en el tiempo, y el tercero, la varianza de dichas variables aleatorias converge a 0, en conclusión, el proceso no se extingue y sus frecuencias relativas convergen a valores constantes que se pueden calcular a partir de los valores de N_1 , de t y de la matriz de medias.

Proposición 2.1. Sea $m_k(t; \mathbf{N}_1) < \infty$, $k = 1, 2, \dots, d$, entonces para $N_1 \rightarrow \infty$ y t fijo tal que $U(t; \mathbf{N}_1) > 0$

$$q(t; \mathbf{N}_1) := P(U(t; \mathbf{N}_1) = 0) \rightarrow 0, \quad (2.11)$$

$$\Delta_k(t; \mathbf{N}_1) \xrightarrow{c.s.} p_k(t; \mathbf{e}_1) \quad y \quad E(\Delta_k(t; \mathbf{N}_1)) \rightarrow p_k(t; \mathbf{e}_1) \quad (2.12)$$

$$Var(\Delta_k(t; \mathbf{N}_1)) \rightarrow 0. \quad (2.13)$$

Demostración. De (2.1)-(2.3) y (2.7)-(2.8), se tiene que $q(t; \mathbf{e}_1) = P(U(t; \mathbf{e}_1) = 0) = F(t, \mathbf{0})$, lo que corresponde a la probabilidad de extinción dado un solo individuo inicial de tipo 1, donde $\mathbf{0} = (0, 0, \dots, 0)$ vector de ceros, y como para definir las frecuencias relativas $\Delta_k(t; \mathbf{N}_1)$ (2.8) sólo se considera t en el conjunto de no extinción de $U(t; \mathbf{e}_1)$, entonces, $q(t; \mathbf{e}_1) < 1$ para todo tiempo t fijo. Por lo tanto cuando $N_1 \rightarrow \infty$, entonces, $q(t; \mathbf{N}_1) = P(U(t; \mathbf{N}_1) = 0) = (q(t; \mathbf{e}_1))^{N_1} \rightarrow 0$, con lo cual queda demostrado (2.11)

De (2.1), (2.4) y la ley fuerte de los grandes números, cuando $N_1 \rightarrow \infty$:

$$\frac{1}{N_1} Z_k(t; \mathbf{N}_1) = \frac{1}{N_1} \sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) \xrightarrow{c.s.} m_k(t; \mathbf{e}_1), \quad k = 1, 2, \dots, d \quad (2.14)$$

Por otro lado, de (2.7), (2.10) y la ley fuerte de los grandes números, cuando $N_1 \rightarrow \infty$:

$$\begin{aligned} \frac{1}{N_1} U(t; \mathbf{N}_1) &= \frac{1}{N_1} \sum_{k=1}^d \sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) \\ &= \frac{1}{N_1} \sum_{i=1}^{N_1} \left(\sum_{k=1}^d Z_k^{(i)}(t; \mathbf{e}_1) \right) \xrightarrow{c.s.} \sum_{k=1}^d m_k(t; \mathbf{e}_1) = M(t; \mathbf{e}_1) \end{aligned} \quad (2.15)$$

Se obtiene de (2.8), (2.9), (2.14) y (2.15), cuando $N_1 \rightarrow \infty$:

$$\Delta_k(t; \mathbf{N}_1) = \frac{\frac{1}{N_1} Z_k(t; \mathbf{N}_1)}{\frac{1}{N_1} U(t; \mathbf{N}_1)} \xrightarrow{c.s.} \frac{m_k(t; \mathbf{e}_1)}{M(t; \mathbf{e}_1)} = p_k(t; \mathbf{e}_1), \quad k = 1, 2, \dots, d. \quad (2.16)$$

En otras palabras, las proporciones $\Delta_k(t; \mathbf{N}_1)$, $k = 1, 2, \dots, d$ cuando se consideran como funciones del número inicial de ancestros N_1 , son estimadores fuertemente consistentes para $p_k(t; \mathbf{e}_1)$.

Dado que $0 \leq \Delta_k(t; \mathbf{N}_1) \leq 1$ y por el teorema de convergencia dominada, de (2.16) se tiene que $E(\Delta_k(t; \mathbf{N}_1))$ converge a $p_k(t; \mathbf{e}_1)$ cuando $N_1 \rightarrow \infty$, lo cual implica que, $\Delta_k(t; \mathbf{N}_1)$ es un estimador asintóticamente insesgado de $p_k(t; \mathbf{e}_1)$, $k = 1, 2, \dots, d$. De manera similar a (2.16), se obtiene que $\Delta_k^2(t; \mathbf{N}_1)$ converge casi siempre a $p_k^2(t; \mathbf{e}_1)$ y por el teorema de convergencia dominada, $E(\Delta_k^2(t; \mathbf{N}_1))$ converge a $p_k^2(t; \mathbf{e}_1)$. Así que:

$$Var(\Delta_k(t; \mathbf{N}_1)) = E(\Delta_k^2(t; \mathbf{N}_1)) - [E(\Delta_k(t; \mathbf{N}_1))]^2 \xrightarrow{N_1 \rightarrow \infty} p_k^2(t; \mathbf{e}_1) - p_k^2(t; \mathbf{e}_1) = 0 \quad (2.17)$$

lo que concluye la demostración. \square

El siguiente teorema es el resultado principal del artículo (Yakovlev & Yanev 2009), en él, se seguirá usando la misma notación para $\sigma_k^2(t) := Var(Z_k(t))$, y con el fin de simplificar la escritura, se utilizarán las siguientes notaciones adicionales:

$$\begin{aligned} p_k(t) &:= p_k(t, \mathbf{e}_1) \\ m_k(t) &:= m_k(t, \mathbf{e}_1) \\ M(t) &:= M(t, \mathbf{e}_1) \end{aligned}$$

Teorema 2.1. *Suponga que $\sigma_k^2(t) < \infty$, $k = 1, 2, \dots, d$, sea la matriz*

$$\mathbf{A}_{(d \times d)}(t) = (a_{ij}(t))_{1 \leq i, j \leq d},$$

donde,

$$a_{ij}(t) = \begin{cases} \sigma_i(t) (1 - p_i(t)) & i = j = 1, 2, \dots, d, \\ -\sigma_i(t) p_j(t) & i \neq j = 1, 2, \dots, d, \end{cases} \quad (2.18)$$

y sean las v.a.

$$W_k(t; \mathbf{N}_1) = M(t) \sqrt{N_1} (\Delta_k(t; \mathbf{N}_1) - p_k(t)), \quad k = 1, 2, \dots, d, \quad (2.19)$$

entonces cuando $N_1 \rightarrow \infty$, se cumple lo siguiente:

(i) Para todo $k = 1, 2, \dots, d$,

$$W_k(t; \mathbf{N}_1) \xrightarrow{\text{dist}} X_k(t), \quad (2.20)$$

donde $X_k(t)$ es una v.a. que tiene una distribución normal con $E(X_k(t)) = 0$ y

$$\text{Var}(X_k(t)) = \sum_{i,j=1}^d r_{ij}(t) a_{ik}(t) a_{jk}(t). \quad (2.21)$$

(ii) Para todo $k = 2, \dots, d-1$ y todo subconjunto $\{\alpha_1, \alpha_2, \dots, \alpha_k\}$ con elementos no repetidos del conjunto $\{1, 2, \dots, d\}$,

$$(W_{\alpha_1}(t; \mathbf{N}_1), W_{\alpha_2}(t; \mathbf{N}_1), \dots, W_{\alpha_k}(t; \mathbf{N}_1)) \xrightarrow{\text{dist}} (X_{\alpha_1}(t), X_{\alpha_2}(t), \dots, X_{\alpha_k}(t)) \quad (2.22)$$

donde $(X_{\alpha_1}(t), X_{\alpha_2}(t), \dots, X_{\alpha_k}(t))$ tiene una distribución conjunta normal multivariada.

(iii) La matriz de covarianza asociada al vector $\mathbf{X}_{(k)}(t) = (X_1(t), X_2(t), \dots, X_k(t))$ está dada por:

$$\begin{aligned} \mathbf{D}_{(k \times k)}(t) &:= (\text{Cov}(X_i(t), X_j(t)))_{1 \leq i, j \leq k} \\ &= (\mathbf{A}_{(d \times k)}(t))' \mathbf{R}_{(d \times d)}(t) (\mathbf{A}_{(d \times k)}(t)) \end{aligned} \quad (2.23)$$

donde $\mathbf{A}_{(d \times k)}(t) = (a_{ij}(t))_{1 \leq i \leq d, 1 \leq j \leq k}$ es la submatriz de tamaño $d \times k$, tomando las primeras k columnas de la matriz $\mathbf{A}_{(d \times d)}(t)$, y $(\mathbf{A}_{(d \times k)}(t))'$ es la matriz transpuesta correspondiente, de tamaño $k \times d$. La matriz de covarianza de cualquier subvector

$$(X_{\alpha_1}(t), X_{\alpha_2}(t), \dots, X_{\alpha_k}(t))$$

se puede obtener de manera similar.

Demostración. Partiendo de (2.1), (2.7) y (2.8), se tiene que para todo $k = 1, 2, \dots, d$,

$$\begin{aligned}
& \Delta_k(t; \mathbf{N}_1) - p_k(t) \\
&= \frac{Z_k(t; \mathbf{N}_1)}{U(t; \mathbf{N}_1)} - p_k(t) \\
&= \frac{Z_k(t; \mathbf{N}_1) - U(t; \mathbf{N}_1) p_k(t)}{U(t; \mathbf{N}_1)} \\
&= \frac{1}{U(t; \mathbf{N}_1)} \left[\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - p_k(t) \left(\sum_{j=1}^d \sum_{i=1}^{N_1} Z_j^{(i)}(t; \mathbf{e}_1) \right) \right] \\
&= \frac{1}{U(t; \mathbf{N}_1)} \left[\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - m_k(t) - p_k(t) \left(\sum_{j=1}^d \sum_{i=1}^{N_1} Z_j^{(i)}(t; \mathbf{e}_1) \right) + m_k(t) \right]
\end{aligned}$$

por (2.9) se tiene que $m_k(t; \mathbf{N}_1) = p_k(t; \mathbf{N}_1) M(t; \mathbf{N}_1)$ y por lo tanto,

$$\begin{aligned}
&= \frac{1}{U(t; \mathbf{N}_1)} \left[\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - m_k(t) - p_k(t) \left(\sum_{j=1}^d \sum_{i=1}^{N_1} Z_j^{(i)}(t; \mathbf{e}_1) \right) + p_k(t) M(t) \right] \\
&= \frac{\sqrt{N_1}}{U(t; \mathbf{N}_1)} \left[\frac{\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - m_k(t)}{\sqrt{N_1}} - p_k(t) \frac{\sum_{j=1}^d \sum_{i=1}^{N_1} Z_j^{(i)}(t; \mathbf{e}_1) - M(t)}{\sqrt{N_1}} \right] \\
&= \frac{\sqrt{N_1}}{U(t; \mathbf{N}_1)} \left[\sigma_k(t) \frac{\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - m_k(t)}{\sqrt{N_1} \sigma_k(t)} - p_k(t) \sum_{j=1}^d \sigma_j(t) \frac{\sum_{i=1}^{N_1} Z_j^{(i)}(t; \mathbf{e}_1) - m_j(t)}{\sqrt{N_1} \sigma_j(t)} \right]
\end{aligned}$$

tomando,

$$\xi_k(t; \mathbf{N}_1) = \frac{\sum_{i=1}^{N_1} Z_k^{(i)}(t; \mathbf{e}_1) - m_k(t)}{\sqrt{N_1} \sigma_k(t)} \quad k = 1, 2, \dots, d.$$

se tiene,

$$\begin{aligned}
&= \frac{\sqrt{N_1}}{U(t; \mathbf{N}_1)} \left[\sigma_k(t) \xi_k(t; \mathbf{N}_1) - p_k(t) \sum_{j=1}^d \sigma_j(t) \xi_j(t; \mathbf{N}_1) \right] \\
&= \frac{\sqrt{N_1}}{U(t; \mathbf{N}_1)} \left[\sigma_k(t) [1 - p_k(t)] \xi_k(t; \mathbf{N}_1) - p_k(t) \sum_{j \neq k}^d \sigma_j(t) \xi_j(t; \mathbf{N}_1) \right]
\end{aligned} \tag{2.24}$$

además,

$$\begin{aligned}
& E(\xi_k(t; \mathbf{N}_1)) = 0 \\
& Var(\xi_k(t; \mathbf{N}_1)) = 1 \\
& Cor(\xi_k(t; \mathbf{N}_1), \xi_l(t; \mathbf{N}_1)) = Cor(Z_k(t; \mathbf{e}_1), Z_l(t; \mathbf{e}_1)) = r_{kl}(t) = \frac{V_{kl}(t)}{\sigma_k(t) \sigma_l(t)}
\end{aligned} \tag{2.25}$$

entonces, por el teorema central del límite para vectores aleatorios, cuando $N_1 \rightarrow \infty$,

$$(\xi_1(t; \mathbf{N}_1), \xi_2(t; \mathbf{N}_1), \dots, \xi_d(t; \mathbf{N}_1)) \xrightarrow{dist} (\psi_1(t), \psi_2(t), \dots, \psi_d(t)), \quad (2.26)$$

donde el vector aleatorio $\psi_{(d)}(t) = (\psi_1(t), \psi_2(t), \dots, \psi_d(t))$ tiene una distribución conjunta normal multivariada d -dimensional con,

$$\begin{aligned} E(\psi_k(t)) &= 0 \\ Var(\psi_k(t)) &= 1 \\ Cov(\psi_k(t), \psi_l(t)) &= Cor(\psi_k(t), \psi_l(t)) = r_{kl}(t) = \frac{V_{kl}(t)}{\sigma_k(t)\sigma_l(t)} \end{aligned} \quad (2.27)$$

De (2.19), (2.24) y (2.26), cuando $N_1 \rightarrow \infty$, se concluye que

$$\begin{aligned} W_k(t; \mathbf{N}_1) &= M(t) \sqrt{N_1} (\Delta_k(t; \mathbf{N}_1) - p_k(t)) \\ &= M(t) \sqrt{N_1} \frac{\sqrt{N_1}}{U(t; \mathbf{N}_1)} \left[\sigma_k(t)[1 - p_k(t)] \xi_k(t; \mathbf{N}_1) - p_k(t) \sum_{j \neq k}^d \sigma_j(t) \xi_j(t; \mathbf{N}_1) \right] \end{aligned}$$

lo cual, teniendo en cuenta que por la ley fuerte de los grandes números,

$$\frac{U(t; \mathbf{N}_1)}{N_1} \xrightarrow[N_1 \rightarrow \infty]{c.s.} M(t).$$

queda de la siguiente forma,

$$W_k(t; \mathbf{N}_1) \xrightarrow{dist} \sigma_k(t)[1 - p_k(t)]\psi_k(t) - p_k(t) \sum_{l \neq k}^d \sigma_l(t)\psi_l(t) \quad (2.28)$$

Tomando

$$X_k(t) = \sigma_k(t)[1 - p_k(t)]\psi_k(t) - p_k(t) \sum_{l \neq k}^d \sigma_l(t)\psi_l(t)$$

y a partir de (2.18) y (2.28), para todo $k = 1, 2, \dots, d$, $X_k(t)$ se puede escribir de la siguiente manera,

$$X_k(t) = \sum_{j=1}^d a_{jk}(t)\psi_j(t) \quad (2.29)$$

lo que indica que los $X_k(t)$ son combinaciones lineales de variables aleatorias con distribución conjunta normal multivariada, entonces, cada $X_k(t)$ es una v.a. que tiene una distribución normal con $E(X_k(t)) = 0$ y de (2.29) con,

$$\begin{aligned} Var(X_k(t)) &= \sum_{i=1}^d \sum_{j=1}^d E(\psi_k(i)\psi_k(j)) a_{ik}(t)a_{jk}(t) \\ &= \sum_{i=1}^d \sum_{j=1}^d r_{ij}(t)a_{ik}(t)a_{jk}(t) \end{aligned} \quad (2.30)$$

con lo que se tiene demostrado el punto (i) del teorema y con los resultados (2.26), (2.28) y (2.29) queda demostrado el punto (ii).

Por otro lado, teniendo las diferentes combinaciones lineales de (2.29), éstas se pueden escribir en forma matricial de la siguiente manera,

$$\mathbf{X}_{(k)}(t) = (\mathbf{A}_{(d \times k)}(t))' \boldsymbol{\Psi}_{(d)}(t), \quad k = 1, 2, \dots, d. \quad (2.31)$$

y dado que $\mathbf{X}_{(k)}(t)$ es una transformación lineal del vector $\boldsymbol{\Psi}_{(d)}(t)$, que tiene distribución normal multivariada con matriz de covarianza $\mathbf{R}_{(d \times d)}(t) = (r_{ij}(t))_{1 \leq i, j \leq d}$, entonces la matriz de covarianza de $\mathbf{X}_{(k)}(t)$ es,

$$\mathbf{D}_{(k \times k)}(t) = (\text{Cov}(X_i(t), X_j(t)))_{1 \leq i, j \leq k} = (\mathbf{A}_{(d \times k)}(t))' \mathbf{R}_{(d \times d)}(t) (\mathbf{A}_{(d \times k)}(t)) \quad (2.32)$$

con lo que se obtiene el resultado (iii) del teorema y se concluye la demostración. \square

Corolario 2.1. *Se tiene la siguiente igualdad,*

$$d_{kk}(t) = (1 - p_k(t))^2 \sigma_k^2(t) + p_k^2 \sum_{i, j \neq k} V_{ij}(t) - 2p_k(t)(1 - p_k(t)) \sum_{i \neq k} V_{ki}(t) \quad (2.33)$$

Demostración. Partiendo de (2.18) y (2.21), para probar el corolario anterior, basta con observar que

$$r_{ij}(t) a_{ik}(t) a_{jk}(t) = \begin{cases} \sigma_i^2(t) (1 - p_i(t))^2 & i = j = k, \\ -p_i(t) (1 - p_i(t)) V_{ik}(t) & i \neq k, j = k, \\ -p_i(t) (1 - p_i(t)) V_{il}(t) & i = k, j \neq k \end{cases}$$

\square

Corolario 2.2. *Si $i \neq j$, entonces*

$$d_{ij}(t) = V_{ij}(t) + p_i(t)p_j(t) \sum_{k, l} V_{kl}(t) - p_i(t) \sum_k C_{kj}(t) - p_j(t) \sum_l V_{il}(t) \quad (2.34)$$

Demostración. Por (2.23) se tiene que

$$d_{ij} = \sum_{k, l=1}^d r_{kl}(t) a_{ki}(t) a_{lj}(t)$$

y por otro lado, (2.18) implica que

$$r_{kl}(t) a_{ki}(t) a_{lj}(t) = \begin{cases} p_i(t)p_j(t)V_{kl}(t) & k \neq i, l \neq j \\ -p_i(t)(1 - p_j(t))V_{kj}(t) & k \neq i, l = j \\ -(1 - p_i(t))p_j(t)V_{il}(t) & k = i, l \neq j \\ (1 - p_i(t))(1 - p_j(t))V_{il}(t) & k = i, l = j \end{cases}$$

a partir de lo cual, realizando las operaciones, se obtiene el resultado del corolario. \square

Dados los resultados expuestos y por la forma en que se definieron las proporciones o frecuencias relativas $\Delta_k(t; \mathbf{N}_1)$, $k = 1, 2, \dots, d$, se cumple que $\sum_{k=1}^d \Delta_k(t; \mathbf{N}_1) = 1$, por lo tanto los $\Delta_k(t; \mathbf{N}_1)$ son linealmente dependientes, esto quiere decir que existen $d - 1$ vectores $d - 1$ dimensionales que asintóticamente tienen una distribución conjunta normal multivariada (no degenerada). Por ejemplo, sea el vector

$$\mathbf{\Delta}_{(d-1)}(t; \mathbf{N}_1) = (\Delta_1(t; \mathbf{N}_1), \Delta_2(t; \mathbf{N}_1), \dots, \Delta_{d-1}(t; \mathbf{N}_1))$$

el cual tiene distribución conjunta asintótica normal multivariada $(d - 1)$ -dimensional con vector de medias

$$E(\mathbf{\Delta}_{(d-1)}(t; \mathbf{N}_1)) = (p_1(t), p_2(t), \dots, p_{d-1}(t))$$

y matriz de covarianza

$$Var(\mathbf{\Delta}_{(d-1)}(t; \mathbf{N}_1)) = \frac{\mathbf{D}_{(d-1)}(t)}{N_1 [M(t)]^2}$$

donde la matriz $\mathbf{D}_{(d-1)}(t) = (d_{ij}(t))_{1 \leq i, j \leq d-1}$ está dada por (2.23) con $k = d - 1$.

2.3. Aplicación del teorema anterior

En estudios de proliferación, diferenciación y muerte de células con su tipología, a partir de un conjunto inicial (por ejemplo, del mismo tipo), más que el conteo, se puede realizar la identificación de las proporciones en distintos momentos del tiempo. A partir de esta situación experimental es posible obtener n observaciones independientes de la forma:

$$\zeta_j = \mathbf{\Delta}(j; \mathbf{N}_1) = (\Delta_1(j; \mathbf{N}_1), \Delta_2(j; \mathbf{N}_1), \dots, \Delta_d(j; \mathbf{N}_1)) \quad (2.35)$$

con $j = 1, 2, \dots, n$, donde $\Delta_k(j; \mathbf{N}_1)$ para $k = 1, 2, \dots, d$, sigue representando a la frecuencia relativa del proceso, correspondiente al tipo k .

De acuerdo con el Teorema 2.1, si en dicho experimento el número inicial de células de un cierto tipo que produce los demás tipos, es suficientemente grande, entonces cada vector ζ_k tiene asintóticamente una distribución normal multivariada, y en consecuencia, se tiene que la función asintótica de log-verosimilitud de dichos vectores está dada por,

$$\Lambda_n(\zeta_1, \zeta_2, \dots, \zeta_n) = \sum_{j=1}^n L_j(\zeta_j; j, \mathbf{N}_1) \quad (2.36)$$

donde $L_j(\zeta_j; j, \mathbf{N}_1)$ denota la correspondiente función de log-verosimilitud de cada vector.

A manera de ejemplo Yakovlev y Yanev analizan la función de log-verosimilitud en el caso de un proceso de BGWM de dos tipos $\mathbf{Z}(n; \mathbf{N}_1) = (Z_1(n; \mathbf{N}_1), Z_2(n; \mathbf{N}_1))$, $n = 0, 1, 2, \dots$, modelo que es utilizado en el análisis de evoluciones celulares.

Como se ha mencionado anteriormente el proceso queda determinado por sus funciones generadoras de probabilidades (f.g.p.)

$$h_k(s_1, s_2) = E\left(s_1^{Z_1(1)}, s_2^{Z_2(1)} \mid Z_k(0) = 1\right), \quad k = 1, 2. \quad (2.37)$$

A partir de (2.37) se puede obtener el primer y segundo momento de la distribución de descendencia, mediante la primera y segunda derivada de las f.g.p evaluadas en el vector $(1, 1)$:

$$m_{ij} = \frac{\partial}{\partial s_j} h_i(s_1, s_2)|_{s_1=s_2=1}, \quad i, j = 1, 2. \quad (2.38)$$

$$b_{jk}^i = \frac{\partial^2}{\partial s_j \partial s_k} h_i(s_1, s_2)|_{s_1=s_2=1}, \quad i, j, k = 1, 2. \quad (2.39)$$

y por (2.6), se tiene que,

$$\begin{aligned} \sigma_i^2(j) &= b_{ii}^1(j) + m_{ii}(j) - m_{ii}^2(j), \quad i = 1, 2 \\ V_{12}(j) &= b_{12}^1(j) - m_{11}(j)m_{12}(j). \end{aligned}$$

Por otra parte, del Teorema 2.1 y la fórmula (2.36) se tiene que,

$$\Lambda_n(\zeta_1, \zeta_2, \dots, \zeta_n) = -\frac{n}{2} \log 2\pi - \frac{1}{2} \sum_{j=1}^n \log S^2(j; \mathbf{N}_1) - \frac{1}{2} \sum_{j=1}^n \frac{(\zeta_j - p(j))^2}{S^2(j; \mathbf{N}_1)}, \quad (2.40)$$

donde,

$$\begin{aligned} \zeta_j &= \Delta_1(j; \mathbf{N}_1), \\ p(j) &= \frac{m_{11}(j)}{M(j)}, \\ M(j) &= m_{11}(j) + m_{12}(j), \end{aligned}$$

y teniendo en cuenta el resultado (2.33), se llega a que,

$$S^2(j; \mathbf{N}_1) = \frac{1}{N_1 M^2(j)} \left(\sigma_1^2(j)(1-p(j))^2 + \sigma_2^2(j)p^2(j) - 2V_{12}(j)p(j)(1-p(j)) \right)$$

En este punto, por todo lo anterior y despejando adecuadamente, se observa que la función de log-verosimilitud queda en términos de las observaciones $\zeta_1, \zeta_2, \dots, \zeta_n$ y de $m_{lk}(j), b_{lk}^i(j)$ con $i, l, k = 1, 2$.

Sean $m_{lk} := m_{lk}(1)$ y $b_{lk}^i := b_{lk}^i(1)$, para cada generación $j = 1, 2, \dots$ se deducen los momentos definidos en (2.4) y (2.5) y a partir de (2.38) y (2.39), se tiene que

$$\mathbf{m}(j) = (m_{lk}(j))_{1 \leq l, k \leq d} = \mathbf{m}^j$$

donde $m_{lk}(j) = E(Z_k(j) | Z_l(0) = 1)$ y $\mathbf{m} = (m_{lk})_{1 \leq l, k \leq d}$.

Además, como $b_{lk}^i(j) = E(Z_l(j)(Z_k(j) - \delta_{lk}) | Z_i(0) = 1)$, con δ_{lk} delta de Kronecker, se obtiene la siguiente fórmula de recurrencia,

$$b_{lk}^i(j+1) = \sum_{o=1}^2 \sum_{r=1}^2 b_{or}^i m_{ol}(j) m_{rk}(j) + \sum_{o=1}^2 m_{io} b_{lk}^o(j), \quad j = 1, 2, \dots, \quad (2.41)$$

Con lo cual Yakovlev y Yanev concluyen que es posible construir la función de log-verosimilitud Λ_n de tal forma que (por (2.40)) quede en términos de las observaciones $\zeta_1, \zeta_2, \dots, \zeta_n$, y de $\{m_{lk}\}$ y $\{b_{lk}^i\}$, $i, l, k = 1, 2$, con lo cual se podría obtener los estimadores de máxima verosimilitud asintóticos de dichos parámetros.

Algoritmos y ejemplos numéricos

Uno de los objetivos de este trabajo consiste en ilustrar y observar, por medio de simulaciones, la estimación de parámetros y el comportamiento de las frecuencias relativas, en un proceso de Bienaymé - Galton - Watson Multitipo (BGWM). En el presente capítulo se plantean y analizan algunos ejemplos numéricos con el fin de ilustrar los resultados teóricos enunciados en los capítulos anteriores.

Al revisar la bibliografía acerca de simulaciones de procesos de BGWM, se encontró que Ospina (2001) describió las características de un proceso de Bienaymé - Galton - Watson (BGW), y observó el comportamiento de dicho proceso, mediante simulaciones con diferentes tamaños iniciales de la población, medias de reproducción, tiempos máximos para las trayectorias y diferentes distribuciones, como la Binomial, la Geométrica y la Poisson. En dicho trabajo también se analizó, desde la parte teórica y numérica, el comportamiento de algunos estimadores de la media de reproducción.

Por otra parte, Stefanescu (1998), trabajó con una colección particular de procesos de BGWM, cuyo modelo teórico, según se menciona en dicho artículo, es un caso especial del modelo de Wright para describir el desarrollo de ciertos genes dentro de una población, a partir de ahí, construyó un algoritmo para la simulación de las trayectorias del proceso a través del tiempo y luego analizó los resultados de las simulaciones, enfocándose en el efecto que tiene la distribución inicial y la fertilidad relativa de cada tipo de alelo, en el tiempo de extinción de la población.

Stefanescu (1998) presenta una colección de procesos de BGWM que están determinados por el siguiente vector de funciones de densidad discretas multivariadas,

$$\mathbf{p}(r_1, r_2, \dots, r_d) := (p_{(1)}(r_1, r_2, \dots, r_d), p_{(2)}(r_1, r_2, \dots, r_d), \dots, p_{(d)}(r_1, r_2, \dots, r_d))$$

donde,

$$p_{(k)}(r_1, r_2, \dots, r_d) = e^{-d a_k} \frac{a_k^{r_1+r_2+\dots+r_d}}{r_1! r_2! \dots r_d!} \quad (3.1)$$

para todo $r_1, r_2, \dots, r_d \in \mathbb{N}_0$, con $k = 1, 2, \dots, d$, cada una de las funciones,

$$p_{(k)}(r_1, r_2, \dots, r_d)$$

corresponde a la distribución de descendencia de los individuos de tipo k , y a_k es un valor no

negativo que representa la media de descendientes por tipo (igual para todos los tipos), que generan los individuos de tipo k .

Cuando se tiene un proceso de BGWM con estas distribuciones de descendencia, en el caso particular de la evolución de ciertos genes dentro de una población, entonces, los parámetros tienen la siguiente interpretación, d es el número de posibles tipos de alelos, es decir, cada una de las formas alternativas que puede tener el gen y la probabilidad de que un alelo de tipo k con fertilidad relativa a_k tenga r_1 descendientes del tipo 1, r_2 descendientes del tipo 2, \dots , r_d descendientes del tipo d , está dada por la función $p_{(k)}(r_1, r_2, \dots, r_d)$.

Al reescribir (3.1) se obtiene lo siguiente:

$$\begin{aligned} p_{(k)}(r_1, r_2, \dots, r_d) &= e^{-d a_k} \frac{a_k^{r_1+r_2+\dots+r_d}}{r_1! r_2! \dots r_d!} \\ &= \left(e^{-a_k} \frac{a_k^{r_1}}{r_1!} \right) \left(e^{-a_k} \frac{a_k^{r_2}}{r_2!} \right) \dots \left(e^{-a_k} \frac{a_k^{r_d}}{r_d!} \right) \end{aligned} \quad (3.2)$$

con lo que se observa que cada una de las distribuciones de descendencia es el producto de d distribuciones Poisson, independientes y con media a_k , por lo tanto, la matriz de medias de reproducción de este proceso es:

$$\mathbf{m} = \begin{pmatrix} a_1 & a_1 & \dots & a_1 \\ a_2 & a_2 & \dots & a_2 \\ \vdots & \vdots & \dots & \vdots \\ a_d & a_d & \dots & a_d \end{pmatrix} \quad (3.3)$$

Al observar lo anterior, antes de continuar con los desarrollos algorítmicos, se propone una tipificación de ciertos procesos de BGWM particulares, lo que permite tener diferentes implementaciones dependiendo de las características propias de la respectiva tipificación.

3.1. Caracterización de algunos procesos de BGWM

3.1.1. Con distribuciones independientes.

Teniendo en cuenta el modelo analizado por Stefanescu (1998), es posible considerar una colección más amplia de procesos de BGWM, en donde, cada distribución de descendencia es el producto de d distribuciones independientes, pero no necesariamente distribuciones Poisson, ni distribuciones idénticas. En esta colección de procesos de BGWM, la distribución de descendencia para un individuo de cierto tipo k , es de la forma,

$$p_{(k)}(r_1, r_2, \dots, r_d) = g_{(k,1)}(r_1) g_{(k,2)}(r_2) \dots g_{(k,d)}(r_d) \quad (3.4)$$

en donde $\{g_{(k,j)}(\cdot) \text{ con } 1 \leq k, j \leq d\}$, es el conjunto de funciones de densidad asociado a un conjunto de variables aleatorias discretas e independientes,

$$\{\eta_{kj} \text{ con } 1 \leq k, j \leq d\}$$

Los procesos con este tipo de distribuciones de descendencia, serán denominados *procesos de BGWM con distribuciones independientes*.

Para esta clase de procesos, tenemos que el vector de medias asociado a la k -ésima distribución de descendencia, está dado por,

$$\mathbf{m}_k = (E(\eta_{k1}), E(\eta_{k2}), \dots, E(\eta_{kd}))' \quad (3.5)$$

y la matriz de varianzas y covarianzas respectiva, es de la forma,

$$\mathbf{V}_k = \begin{pmatrix} Var(\eta_{k1}) & 0 & \dots & 0 \\ 0 & Var(\eta_{k2}) & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & Var(\eta_{kd}) \end{pmatrix} \quad (3.6)$$

3.1.2. Con distribución multinomial.

En el caso de los procesos de BGWM con distribuciones independientes, se observa que cada individuo de tipo k , tiene un cierto número de descendientes de acuerdo a un conjunto de distribuciones univariadas independientes, una distribución por cada tipo de descendiente. Al considerar un enfoque diferente, si se supone que un individuo de tipo k tiene un número total de descendientes aleatorio, y a partir de ahí, cada uno de estos descendientes tiene una probabilidad de ser de cierto tipo, entonces, se tiene una colección distinta de procesos de BGWM, con distribuciones de descendencia que no necesariamente son el producto de funciones de densidad univariadas independientes.

Sí se supone que $g_{(k)}(\cdot)$ es la función de densidad discreta asociada al número aleatorio total de descendientes que genera un individuo de tipo k , y sí la probabilidad de que al tener $s = r_1 + r_2 + \dots + r_d$ descendientes, r_1 son de tipo 1, r_2 son de tipo 2, \dots , r_d son de tipo d , está dada por:

$$h_{(k)}(r_1, r_2, \dots, r_d) = \frac{s!}{r_1! r_2! \dots r_d!} b_{k1}^{r_1} b_{k2}^{r_2} \dots b_{kd}^{r_d}$$

entonces el proceso de BGWM asociado, quedaría determinado por la siguiente distribución de descendencia para el tipo k ,

$$\begin{aligned} p_{(k)}(r_1, r_2, \dots, r_d) &= (h_{(k)}(r_1, r_2, \dots, r_d)) (g_{(k)}(s)) \\ &= \frac{s!}{r_1! r_2! \dots r_d!} b_{k1}^{r_1} b_{k2}^{r_2} \dots b_{kd}^{r_d} (g_{(k)}(s)) \end{aligned} \quad (3.7)$$

con $1 \leq k \leq d$.

Escrito de otra manera, sea $\mathbf{Y}_{(k)} = (Y_{(k,1)}, Y_{(k,2)}, \dots, Y_{(k,d)})$ el vector aleatorio correspondiente al número de descendientes por tipo, generados por un individuo de tipo k , entonces,

$$\mathbf{Y}_{(k)} \sim Multinomial(\eta_k, b_{k1}, b_{k2}, \dots, b_{kd})$$

donde, el número total de descendientes de un individuo de tipo k está dado por la variable aleatoria discreta η_k , con función de densidad $g_{(k)}(\cdot)$, y las constantes b_{kj} corresponden a la probabilidad de que un individuo de tipo k tenga un descendiente de tipo j . Los procesos con este tipo de distribuciones de descendencia, serán denominados *procesos de BGWM con distribución multinomial*.

Sean $E(\eta_k) = \mu_k$ y $Var(\eta_k) = \sigma_k^2$, entonces, para la clase de procesos de BGWM con distribución multinomial, el vector de medias asociado a la k -ésima distribución de descendencia, está dado por,

$$\mathbf{m}_k = E(\mathbf{Y}_{(k)}) = E(E(\mathbf{Y}_{(k)} | \eta_k))$$

por lo tanto, el j -ésimo término de dicho vector de medias es,

$$\begin{aligned} m_{kj} &= E(Y_{(k,j)}) \\ &= E(E(Y_{(k,j)} | \eta_k)) \\ &= E(b_{kj} \eta_k) \\ &= b_{kj} E(\eta_k) \\ &= b_{kj} \mu_k \end{aligned} \tag{3.8}$$

por otro lado, se obtiene el siguiente resultado para la matriz de varianzas y covarianzas respectiva,

$$\mathbf{V}_k = Var(\mathbf{Y}_{(k)}) = E(Var(\mathbf{Y}_{(k)} | \eta_k)) + Var(E(\mathbf{Y}_{(k)} | \eta_k))$$

donde,

$$\begin{aligned} Var(Y_{(k,j)}) &= E(Var(Y_{(k,j)} | \eta_k)) + Var(E(Y_{(k,j)} | \eta_k)) \\ &= E(b_{kj} (1 - b_{kj}) \eta_k) + Var(b_{kj} \eta_k) \\ &= b_{kj} (1 - b_{kj}) E(\eta_k) + b_{kj}^2 Var(\eta_k) \\ &= b_{kj} (1 - b_{kj}) \mu_k + b_{kj}^2 \sigma_k^2 \\ &= b_{kj} \mu_k + b_{kj}^2 (\sigma_k^2 - \mu_k) \end{aligned} \tag{3.9}$$

y,

$$\begin{aligned} Cov(Y_{(k,j)}, Y_{(k,l)}) &= E(Cov(Y_{(k,j)}, Y_{(k,l)} | \eta_k)) + Cov(E(Y_{(k,j)} | \eta_k), E(Y_{(k,l)} | \eta_k)) \\ &= E(-b_{kj} b_{kl} \eta_k) + Cov(b_{kj} \eta_k, b_{kl} \eta_k) \\ &= -b_{kj} b_{kl} E(\eta_k) + b_{kj} b_{kl} Var(\eta_k) \\ &= -b_{kj} b_{kl} \mu_k + b_{kj} b_{kl} \sigma_k^2 \\ &= b_{kj} b_{kl} (\sigma_k^2 - \mu_k) \end{aligned} \tag{3.10}$$

3.2. Algoritmos e implementación

Los siguientes son los tipos de datos o variables que se utilizarán en los algoritmos para las simulaciones:

- d es el número de tipos del proceso de BGWM.
- n es el contador para el tiempo o las generaciones.
- n^* es el valor más grande de n que se tendrá en cuenta para la simulación.

- $\mathbf{N} = (N_1, N_2, \dots, N_d)'$ es el vector de individuos iniciales, donde N_1 es el número inicial de individuos del tipo 1, N_2 es el número inicial de individuos del tipo 2, \dots , N_d es el número inicial de individuos del tipo d .
- $\mathbf{C}(n) := (c_{ij}(n))_{1 \leq i, j \leq d}$, es una matriz cuadrada de tamaño $d \times d$, de tal forma que el valor $c_{ij}(n)$, corresponde al número de individuos de tipo j de la n -ésima generación, que generaron los individuos de tipo i de la $(n - 1)$ -ésima generación.

Teniendo en cuenta las tipificaciones que se presentaron en la sección anterior, se observa que cada una de ellas requiere diferentes datos de entrada para determinar las distribuciones de descendencia, por lo tanto, es necesario desarrollar un algoritmo para simular un proceso de BGWM con distribuciones independientes, uno para el caso con distribución multinomial y uno para el caso en que se tiene un vector de distribuciones de descendencia en general.

Para cada uno de los algoritmos se debe tener como datos de entrada d , n , n^* y \mathbf{N} , de manera adicional, para el caso general se deben tener las d distribuciones d -multivariadas discretas $p_{(k)}(\cdot, \dots, \cdot)$, con $1 \leq k \leq d$ de acuerdo a la Definición 1.2.

Algoritmo 3.1. Simulación de un proceso de BGWM general.

Paso 0. Entrada d , n^* , \mathbf{N} y $p_{(k)}(\cdot, \dots, \cdot)$ con $1 \leq k \leq d$.

Paso 1. $\mathbf{C}(1) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 2. Para $k \leftarrow 1$ hasta d haga:

Paso 2.1. Si $N_k = 0$ pase a la siguiente iteración.

Paso 2.2. Para $l \leftarrow 1$ hasta N_k haga:

Paso 2.2.1. $\mathbf{v} \leftarrow$ generar un vector aleatorio de $p_{(k)}(\cdot, \dots, \cdot)$.

Paso 2.2.2. Vector fila k -ésimo de $\mathbf{C}(1) \leftarrow$ vector fila k -ésimo de $\mathbf{C}(1) + \mathbf{v}$.

Paso 3. Para $n \leftarrow 2$ hasta n^* haga:

Paso 3.1. $\mathbf{C}(n) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 3.2. Para $k \leftarrow 1$ hasta d , para $j \leftarrow 1$ hasta d , haga:

Paso 3.2.1. Si $c_{jk}(n - 1) = 0$ entonces pase a la siguiente iteración.

Paso 3.2.2. Para $l \leftarrow 1$ hasta $c_{jk}(n - 1)$ haga:

Paso 3.2.2.1. $\mathbf{v} \leftarrow$ generar un vector aleatorio de $p_{(j)}(\cdot, \dots, \cdot)$.

Paso 3.2.2.2. Vector fila k -ésimo de $\mathbf{C}(n) \leftarrow$ vector fila k -ésimo de $\mathbf{C}(n) + \mathbf{v}$.

Paso 4. Salida $\mathbf{C}(1), \mathbf{C}(2), \dots, \mathbf{C}(n^*)$.

Para los procesos de BGWM con distribución multinomial, se deben tener, d distribuciones univariadas discretas $g_{(k)}(\cdot)$ con $1 \leq k \leq d$, asociadas a las variables aleatorias η_k , y los d vectores de tamaño d , asociados a los parámetros $(b_{k1}, b_{k2}, \dots, b_{kd})$ con $1 \leq k \leq d$, correspondientes a cada una de las distribuciones multinomiales de descendencia planteadas en (3.7).

Algoritmo 3.2. Simulación de un proceso de BGWM con distribución multinomial.

Paso 0. Entrada $d, n^*, \mathbf{N}, g_{(k)}(\cdot)$ con $1 \leq k \leq d$ y $\mathbf{B} = (b_{kj})_{1 \leq k, j \leq d}$.

Paso 1. $\mathbf{C}(1) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 2. Para $k \leftarrow 1$ hasta d haga:

Paso 2.1. Si $N_k = 0$ entonces pase a la siguiente iteración.

Paso 2.2. Para $l \leftarrow 1$ hasta N_k haga:

Paso 2.2.1. $nDesc \leftarrow$ generar un número aleatorio de $g_{(k)}(\cdot)$.

Paso 2.2.2. $\mathbf{v} \leftarrow$ generar un vector aleatorio de
 $Multinomial(nDesc, b_{k1}, b_{k2}, \dots, b_{kd})$.

Paso 2.2.3. Vector fila k -ésimo de $\mathbf{C}(1) \leftarrow$ vector fila k -ésimo de $\mathbf{C}(1) + \mathbf{v}$.

Paso 3. Para $n \leftarrow 2$ hasta n^* haga:

Paso 3.1. $\mathbf{C}(n) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 3.2. Para $k \leftarrow 1$ hasta d , para $j \leftarrow 1$ hasta d , haga:

Paso 3.2.1. Si $c_{jk}(n-1) = 0$ entonces pase a la siguiente iteración.

Paso 3.2.2. Para $l \leftarrow 1$ hasta $c_{jk}(n-1)$ haga:

Paso 3.2.2.1. $nDesc \leftarrow$ generar un número aleatorio de $g_{(j)}(\cdot)$.

Paso 3.2.2.2. $\mathbf{v} \leftarrow$ generar un vector aleatorio de
 $Multinomial(nDesc, b_{j1}, b_{j2}, \dots, b_{jd})$.

Paso 3.2.2.3. Vector fila k -ésimo de $\mathbf{C}(n) \leftarrow$ vector fila k -ésimo de $\mathbf{C}(n) + \mathbf{v}$.

Paso 4. Salida $\mathbf{C}(1), \mathbf{C}(2), \dots, \mathbf{C}(n^*)$.

Para los procesos de BGWM con distribuciones independientes se deben tener d^2 distribuciones univariadas discretas $g_{(k,j)}(\cdot)$ con $1 \leq k, j \leq d$, asociadas a las variables aleatorias η_{kj} , de acuerdo a la ecuación (3.4).

Algoritmo 3.3. Simulación de un proceso de BGWM con distribuciones independientes.

Paso 0. Entrada $d, n^*, \mathbf{N}, g_{(k,j)}(\cdot)$ con $1 \leq k, j \leq d$.

Paso 1. $\mathbf{C}(1) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 2. Para $k \leftarrow 1$ hasta d haga:

Paso 2.1. Si $N_k = 0$ entonces pase a la siguiente iteración.

Paso 2.2. Para $l \leftarrow 1$ hasta N_k , para $j \leftarrow 1$ hasta d , haga:

Paso 2.2.1. $aux \leftarrow$ generar un número aleatorio de $g_{(k,j)}(\cdot)$.

Paso 2.2.2. $c_{kj}(1) \leftarrow c_{kj}(1) + aux$.

Paso 3. Para $n \leftarrow 2$ hasta n^* haga:

Paso 3.1. $\mathbf{C}(n) \leftarrow \mathbf{0}_{(d \times d)}$.

Paso 3.2. Para $k \leftarrow 1$ hasta d , para $i \leftarrow 1$ hasta d , haga:

Paso 3.2.1. Si $c_{ik}(n-1) = 0$ entonces pase a la siguiente iteración.

Paso 3.2.2. Para $l \leftarrow 1$ hasta $c_{ik}(n-1)$, para $j \leftarrow 1$ hasta d , haga:

Paso 3.2.2.1. $aux \leftarrow$ generar un número aleatorio de $g_{(k,j)}(\cdot)$.

Paso 3.2.2.2. $c_{kj}(1) \leftarrow c_{kj}(1) + aux$.

Paso 4. Salida $C(1), C(2), \dots, C(n^*)$.

Se decidió que la implementación de estos algoritmos se realizara, no sólo para los análisis de este trabajo, sino de tal forma que fuera posible su uso, distribución o modificación por otras personas interesadas en este tema, por tal razón, se escogió como plataforma de desarrollo, el software R (R Development Core Team 2009a), el cual es multiplataforma, es decir, funciona bajo diferentes sistemas operativos, como lo son, diferentes versiones de Windows, MacOS, UNIX, incluyendo para el último caso, diferentes distribuciones de Linux; por otro lado, es software de uso libre bajo la licencia GNU GPL (Free Software Foundation 2007), licencia que permite cualquier tipo de uso del software creado, su modificación y distribución, conservando los derechos de autor, exigiendo la correspondiente citación de sus autores y en caso de distribución de alguna modificación sobre dicho software, ésta estará obligada a tener una licencia que, como mínimo, conserve las libertades garantizadas por la GNU GPL.

En el caso de R, es posible formar parte de su desarrollo mediante la construcción de los llamados paquetes (R Development Core Team 2009b), que amplían su configuración básica. Teniendo en cuenta las condiciones y estándares definidos para la construcción de dichos paquetes, se creó uno llamado `Branching`, en donde se incluyeron las funciones básicas necesarias para el desarrollo de la parte numérica de este trabajo, de tal manera que puede usarse para diversas aplicaciones que sean modeladas mediante procesos de BGWM.

En el Apéndice C, se describen las instrucciones para instalar R, las instrucciones para instalar el paquete `Branching` y las funciones que componen dicho paquete, con un ejemplo de su uso.

En la siguiente sección se analizan los resultados de algunas simulaciones para aplicaciones específicas, modeladas mediante procesos de BGWM, que fueron obtenidos mediante el uso de las funciones del paquete `Branching`, permitiendo así, ilustrar con estos ejemplos numéricos, los resultados teóricos de los capítulos anteriores.

3.3. Ejemplos numéricos

Los primeros ejemplos numéricos se realizan a partir del proceso de BGWM con las distribuciones de descendencia (3.1) expuestas por Stefanescu (1998). Dicho proceso en particular, se puede ver como un proceso de BGWM con distribuciones independientes, con distribuciones multinomial o con distribuciones en general, por lo tanto, puede simularse mediante cualquiera de los Algoritmos 3.1, 3.2 y 3.3.

Para el Algoritmo 3.1, el caso general, se conocen las distribuciones de descendencia, por lo tanto, para una colección finita de vectores (r_1, r_2, \dots, r_d) con $r_1, r_2, \dots, r_d \in \mathbb{N}_0$ es posible calcular sus probabilidades respectivas; para el Algoritmo 3.3, el caso con distribuciones independientes, se tiene (3.2), con lo cual podemos ver que cada una de las distribuciones de descendencia es el producto de d distribuciones Poisson idénticas; por último, para el Algoritmo 3.2, el caso con distribución multinomial, basta tomar $\eta_k \sim \text{Poisson}(da_k)$ y $b_{k1} = b_{k2} = \dots = b_{kd} = \frac{1}{d}$, de

esa forma, de (3.7) se tiene que,

$$\begin{aligned}
 p_{(k)}(r_1, r_2, \dots, r_d) &= \frac{s!}{r_1! r_2! \dots r_d!} b_{k1}^{r_1} b_{k2}^{r_2} \dots b_{kd}^{r_d} (g_{(k)}(s)) \\
 &= \frac{s!}{r_1! r_2! \dots r_d!} \left(\frac{1}{d}\right)^{r_1+r_2+\dots+r_d} \left(e^{-d a_k} \frac{(d a_k)^s}{s!}\right) \\
 &= e^{-d a_k} a_k^s \frac{d^s}{s!} \frac{s!}{r_1! r_2! \dots r_d!} \frac{1}{d^{r_1+r_2+\dots+r_d}} \\
 &= e^{-d a_k} \frac{a_k^{r_1+r_2+\dots+r_d}}{r_1! r_2! \dots r_d!}
 \end{aligned} \tag{3.11}$$

resultado que coincide con la distribución en (3.1).

Se simularon los procesos mediante cada una de las implementaciones respectivamente asociadas a los Algoritmos 3.1, 3.2 y 3.3 (ver Apéndice C), tomando los siguientes datos de entrada, similares a los utilizados por Stefanescu (1998), $d = 2$, $n^* = 30$, $\mathbf{N} = (90, 10)$ y los que se indican en la Tabla 3.1.

Tabla 3.1: Parámetros de las distribuciones de descendencia, para los modelos A, B y C.

	a_1	a_2
Modelo A.	0.2	0.3
Modelo B.	0.4	0.6
Modelo C.	0.6	0.9

Teniendo en cuenta (3.3) y los parámetros de la Tabla 3.1, en la Tabla 3.2 se muestran, la matriz de medias de cada modelo con su respectivo mayor valor propio ρ .

Tabla 3.2: Matrices de medias y mayor valor propio respectivo, para los modelos A, B y C.

	Modelo A	Modelo B	Modelo C.
\mathbf{m}	$\begin{pmatrix} 0.2 & 0.2 \\ 0.3 & 0.3 \end{pmatrix}$	$\begin{pmatrix} 0.4 & 0.4 \\ 0.6 & 0.6 \end{pmatrix}$	$\begin{pmatrix} 0.6 & 0.6 \\ 0.9 & 0.9 \end{pmatrix}$
ρ	0.5	1	1.5

Como resultado de las mencionadas simulaciones, se obtuvieron las Figuras 3.1, 3.2 y 3.3, en donde cada una de ellas, representa uno de los modelos (A, B y C respectivamente) y está compuesta por tres gráficos que muestran las trayectorias resultantes para los dos tipos de individuos, al realizar una simulación con cada uno de los algoritmos (3.1, 3.2 y 3.3, respectivamente).

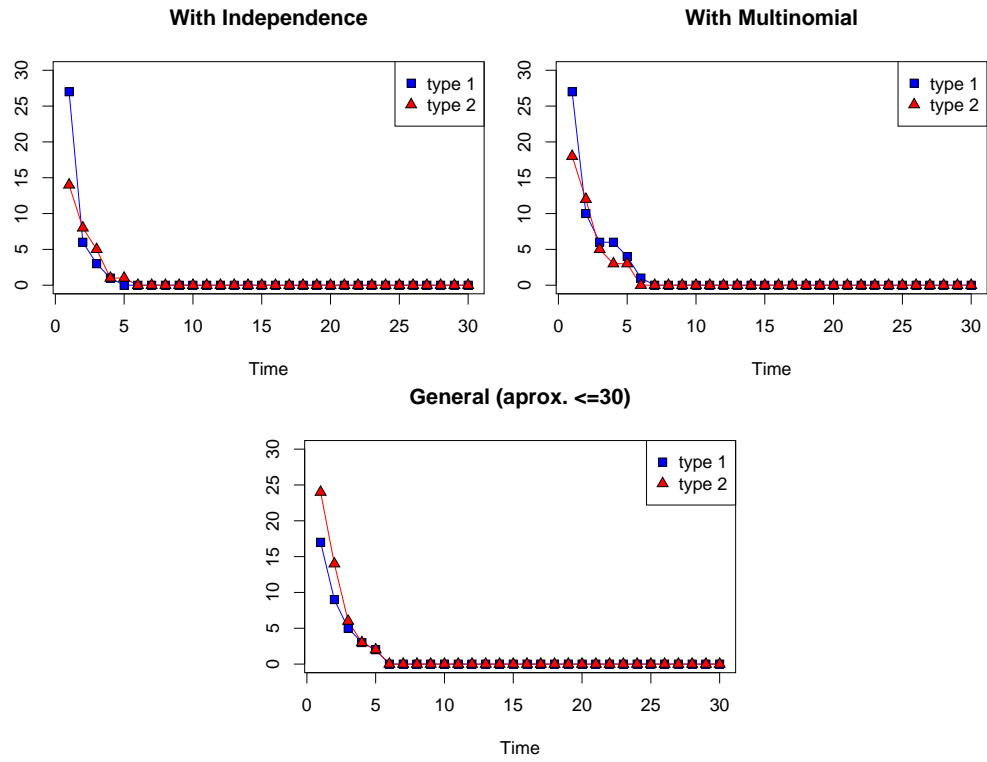


Figura 3.1: Trayectorias de los tipos de individuos a través del tiempo para el Modelo A.

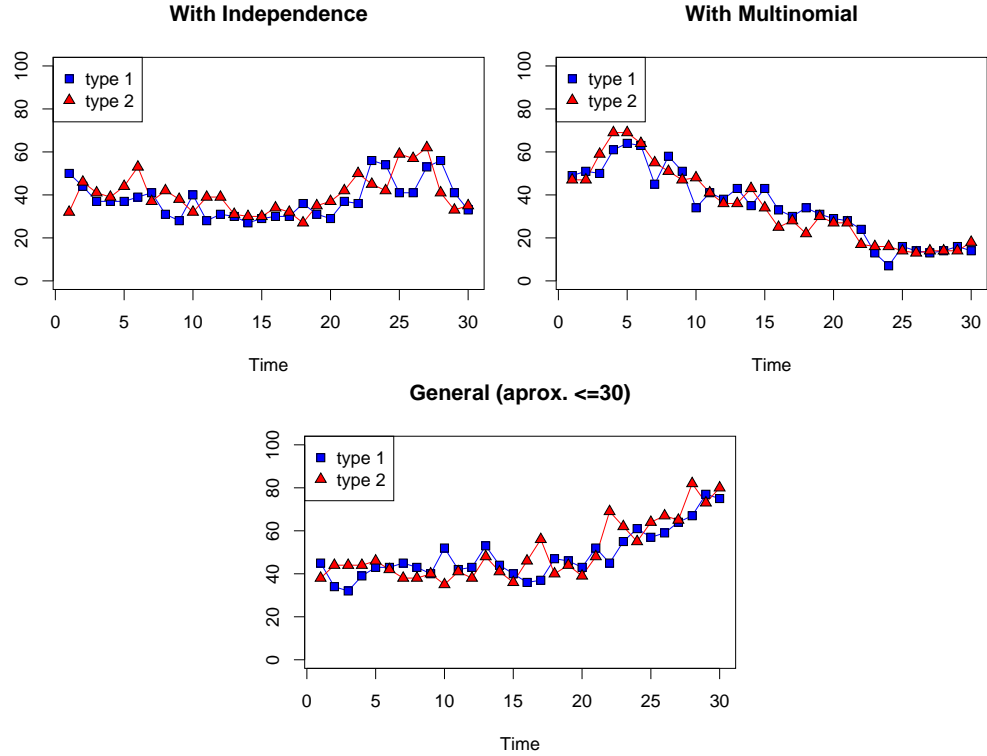


Figura 3.2: Trayectorias de los tipos de individuos a través del tiempo para el Modelo B.

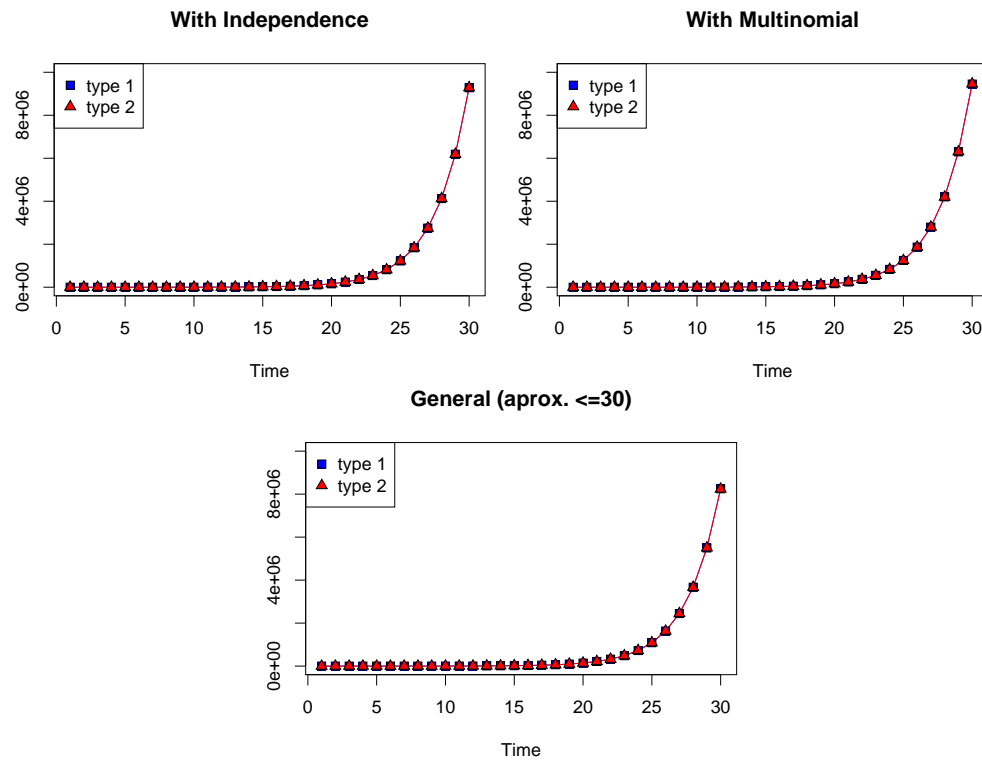


Figura 3.3: Trayectorias de los tipos de individuos a través del tiempo para el Modelo C.

Se observa que:

- En el modelo A (Figura 3.1), el número de individuos de cada tipo va disminuyendo a través del tiempo hasta extinguirse.
- En el modelo C (Figura 3.2), el número de individuos de cada tipo crece de manera exponencial, a medida que pasa el tiempo.
- En el modelo B (Figura 3.3), el número de individuos de cada tipo se mantiene relativamente estable en comparación con el modelo A y el modelo C.
- En los tres modelos, el número de individuos del tipo 1 se mantiene, a través del tiempo, relativamente cercano al número de individuos del tipo 2

dicho comportamiento coincide con lo expuesto en la sección 1.2.1 de acuerdo a los valores de la Tabla 3.2, en donde se concluye a partir de los valores de ρ que, los modelos A, B y C son procesos de BGWM subcrítico, crítico y supercrítico, respectivamente. De manera adicional, como era de esperarse, las figuras no muestran diferencias significativas con respecto al algoritmo utilizado para la simulación.

Stefanescu (1998) estimó la media y la varianza de la distribución asociada al tiempo de extinción τ para el modelo A, por lo cual, con el objetivo de obtener valores que fueran comparables, se realizaron tres pruebas por algoritmo, cada una de ellas compuesta por 1000 simulaciones, tomando los parámetros utilizados para obtener la Figura 3.1 asociada al modelo subcrítico A. Se obtuvo la Figura 3.4 a modo de ilustración del análisis del tiempo de extinción realizado, luego,

se calcularon las estimaciones de la media y la varianza, valores que aparecen en las Tablas 3.3 y 3.4, de manera comparativa a los obtenidos por Stefanescu (1998).

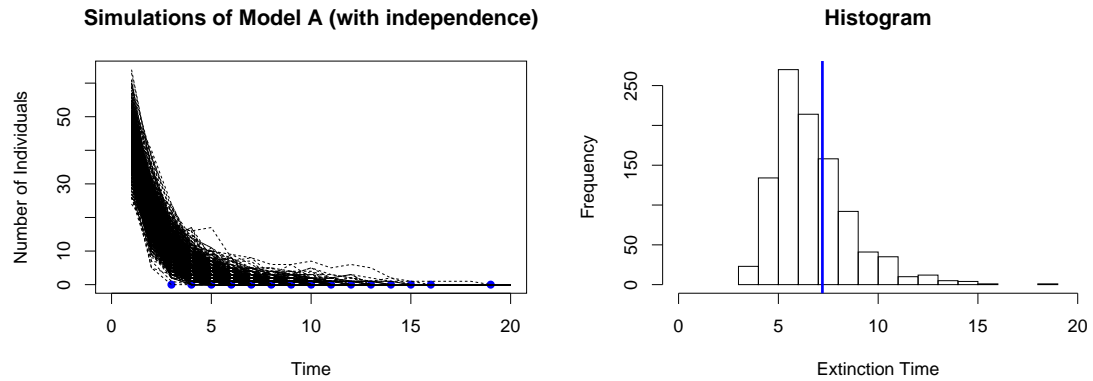


Figura 3.4: Trayectorias del total de individuos de la población e histograma del tiempo de extinción para el modelo A. Tamaño de muestra = 1000 (Prueba 1 - con distribuciones independientes).

Tabla 3.3: Media muestral del tiempo de extinción para el modelo A (tamaño de muestra = 1000).

	independientes	multinomial	general
Prueba 1	7.207	7.191	7.186
Prueba 2	7.127	7.187	7.113
Prueba 3	7.071	7.042	7.1
Stefanescu		7.15	

Tabla 3.4: Varianza muestral del tiempo de extinción para el modelo A (tamaño de muestra = 1000).

	independientes	multinomial	general
Prueba 1	3.8440	3.8263	3.7536
Prueba 2	3.4443	4.1762	3.2435
Prueba 3	3.6196	3.2615	3.3113
Stefanescu		3.84	

La consistencia entre los resultados se puede tomar como un indicador del buen funcionamiento de los algoritmos implementados para realizar las simulaciones.

3.3.1. Estimación de matriz de medias y matrices de covarianzas

En esta sección se observa numéricamente, el comportamiento de los estimadores de la matriz de medias y las matrices de varianzas y covarianzas de un proceso de BGWM supercrítico, cuando el tiempo crece y el número de individuos iniciales es constante, de acuerdo a los resultados teóricos presentados en el Capítulo 1.4. Para tal fin se realizarán las simulaciones a partir de la siguiente aplicación de un proceso de BGWM:

Ejemplo 3.1 (Un modelo demográfico). Uno de los puntos de interés en los estudios demográficos consiste en observar la distribución de las edades de una población. En cuanto a un proceso

de ramificación multitypo se puede considerar que cada individuo, de acuerdo al correspondiente rango de edad al que pertenece dentro de la población, pertenece a un tipo. Si se consideran sólo los individuos de género femenino dentro de la población humana, se dividen en $d > 1$ grupos de edad y si se observa el comportamiento de la población en intervalos de tiempo sucesivos $[n, n+1)$ con $n = 0, 1, 2, \dots$ para cierta unidad de tiempo, entonces, teniendo en cuenta que durante cada intervalo de tiempo cierta cantidad de individuos sobrevive y algunos de ellos generan descendientes, el comportamiento a través del tiempo de la población descrita con estas características, puede ser descrito por un proceso de BGWM.

En un intervalo de tiempo, una mujer del grupo de edad k , con $k = 1, 2, \dots, d - 1$, puede tener descendientes que pertenecerían al grupo de edad 1, además, si ella sobrevive el intervalo de tiempo definido y el tamaño de dicho intervalo coincide con el tamaño del rango de edad, entonces, en el siguiente intervalo de tiempo, ella pertenecería al grupo $k + 1$. Por otro lado, en el caso de que una mujer pertenezca al grupo de edad d , entonces, podría tener descendientes que irían al grupo 1, pero al final del intervalo de tiempo ya no se consideraría dentro de la población. Un proceso de BGWM como el descrito, generaliza el modelo determinista propuesto por Leslie (Leslie 1945), donde surgen las llamadas matrices de Leslie, matrices que coinciden, en el caso del proceso de BGWM, con la matriz de medias.

Para plantear las simulaciones con datos cercanos a la realidad se tomo el conjunto de datos `goodman` incluidos dentro del paquete `demogR` (Jones 2007) del software R, que corresponden a los datos analizados por Goodman, Keyfitz & Pullum (1974). La matriz de Leslie calculada a partir de dichos datos, asociada al comportamiento de la población de Estados Unidos, con los rangos de edad, en años, dados por los siguientes valores de corte $\{0, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55\}$, es la siguiente:

$$\begin{pmatrix} 0 & 0.9972 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0010 & 0 & 0.9984 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0820 & 0 & 0 & 0.9978 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2884 & 0 & 0 & 0 & 0.9968 & 0 & 0 & 0 & 0 & 0 \\ 0.3777 & 0 & 0 & 0 & 0 & 0.9961 & 0 & 0 & 0 & 0 \\ 0.2647 & 0 & 0 & 0 & 0 & 0 & 0.9949 & 0 & 0 & 0 \\ 0.1405 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9924 & 0 & 0 \\ 0.0586 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9886 & 0 \\ 0.0134 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.9829 \\ 0.0003 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Los valores de la anterior matriz se interpretan de la siguiente manera, la primera columna se refiere a la fertilidad de cada rango de edad, por ejemplo, el valor 0.2647 en la posición (1, 6), corresponde al promedio de hijas que tienen mujeres de 30 a 35 años de edad, por otro lado, los valores que se encuentran sobre la diagonal de la matriz están asociados a la supervivencia de cada individuo por cada rango, por ejemplo, el valor 0.9968 de la posición (4, 5) indica que en promedio el 99.68 % de las mujeres de 20 a 25 años sobrevivirán para llegar a los 25 a 30 años de edad.

Con el objetivo de simplificar los gráficos relacionados con el comportamiento de los estimadores, se decidió construir una matriz de dimensión 3 en donde los rangos de edad tuvieran el

mismo tamaño. Es así como, la matriz de Leslie que se tendrá en cuenta para las simulaciones, es la siguiente:

$$\begin{pmatrix} 0.08 & 0.99 & 0 \\ 1.06 & 0 & 0.98 \\ 0.07 & 0 & 0 \end{pmatrix} \quad (3.12)$$

en donde, los rangos de edad, en años, considerados son: $[0, 20)$, $[20, 40)$ y $[40, 60)$, es decir, de 0 a 20 años, de 20 a 40 años y de 40 a 60 años, respectivamente.

La lectura de la matriz en la ecuación (3.12), incluyendo el punto de vista de los procesos de BGWM, es la siguiente, las mujeres de 0 a 19 años pertenecen al tipo 1, tienen una fertilidad media de 0.08 y una media de supervivencia de 0.99, las mujeres de 20 a 39 años pertenecen al tipo 2, tienen una fertilidad media de 1.06 y una media de supervivencia de 0.98, por último, las mujeres de 40 a 59 años pertenecen al tipo 3 y tienen una fertilidad media de 0.07. Para este proceso en particular, la unidad de tiempo es de 20 años.

Se supondrá que la fertilidad para cada mujer está dada por una distribución Poisson, la supervivencia por una distribución Bernoulli, y que además, se tiene independencia entre las dos. De esta manera las distribuciones de descendencia están dadas por:

$$p_{(1)}(r_1, r_2, r_3) = \left(e^{-0.08} \frac{(0.08)^{r_1}}{r_1!} \right) \left((0.99)^{r_2} (0.01)^{1-r_2} I_{\{0,1\}}(r_2) \right) (I_{\{0\}}(r_3)) \quad (3.13)$$

$$p_{(2)}(r_1, r_2, r_3) = \left(e^{-1.06} \frac{(1.06)^{r_1}}{r_1!} \right) (I_{\{0\}}(r_2)) \left((0.98)^{r_3} (0.02)^{1-r_3} I_{\{0,1\}}(r_3) \right) \quad (3.14)$$

$$p_{(3)}(r_1, r_2, r_3) = \left(e^{-0.07} \frac{(0.07)^{r_1}}{r_1!} \right) (I_{\{0\}}(r_2)) (I_{\{0\}}(r_3)) \quad (3.15)$$

La matriz de medias asociada a dichas distribuciones de descendencia coincide con (3.12) y el mayor valor propio de esta matriz es $\rho = 1.0950$, lo que quiere decir que el proceso es supercrítico, por otro lado, las matrices de varianzas y covarianzas son:

$$\mathbf{V}_1 = \begin{pmatrix} 0.08 & 0 & 0 \\ 0 & (0.99)(0.01) & 0 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.08 & 0 & 0 \\ 0 & 0.0099 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.16)$$

$$\mathbf{V}_2 = \begin{pmatrix} 1.06 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & (0.98)(0.02) \end{pmatrix} = \begin{pmatrix} 1.06 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.0196 \end{pmatrix} \quad (3.17)$$

$$\mathbf{V}_3 = \begin{pmatrix} 0.07 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (3.18)$$

Teniendo definido el proceso de BGWM de interés, se realizó una simulación con $d = 3$, $n^* = 120$, $\mathbf{N} = (10, 10, 10)$, y al ser un proceso con distribuciones independientes, se tomaron η_{kj} , de la siguiente manera:

$$\begin{aligned} \eta_{11} &\sim \text{Poisson}(\lambda = 0.08) & \eta_{12} &\sim \text{Bernoulli}(p = 0.99) & \eta_{13} &= 0 \\ \eta_{21} &\sim \text{Poisson}(\lambda = 1.06) & \eta_{22} &= 0 & \eta_{23} &\sim \text{Bernoulli}(p = 0.98) \\ \eta_{31} &\sim \text{Poisson}(\lambda = 0.07) & \eta_{32} &= 0 & \eta_{33} &= 0 \end{aligned}$$

luego, para $n = 1, 2, \dots, 120$, se calculó el estimador empírico $\tilde{\mathbf{m}}_k(n; \mathbf{N})$ y el estimador máximo verosímil $\hat{\mathbf{m}}_k(n; \mathbf{N})$, estimadores del vector de medias de la distribución de descendencia de los individuos de tipo $k = 1, 2, 3$, obtenidos a partir de los valores de $Y_{(i,j)}(n; \mathbf{N})$ resultantes de la simulación.

En la Figura 3.5 se observan dos gráficos, en ellos se muestra cómo cambian, respectivamente, los valores del estimador empírico y del estimador máximo verosímil de la matriz de medias \mathbf{m} , a medida que el tiempo crece. Las líneas punteadas corresponden a los valores de la matriz de medias (3.12), los gráficos ilustran la convergencia de los dos tipos de estimadores a dichos valores y se observa que, el estimador máximo verosímil $\hat{\mathbf{m}}(n; \mathbf{N})$ tiene menor variabilidad y converge más rápido que el estimador empírico $\tilde{\mathbf{m}}(n; \mathbf{N})$, lo cual coincide con los resultados del Teorema 1.7, ya que la varianza del estimador empírico está dada por,

$$(Z_k(n-1; \mathbf{N}))^{-1/2} \mathbf{V}_k,$$

la varianza del estimador máximo verosímil está dada por,

$$\left(\sum_{i=0}^{n-1} Z_k(i; \mathbf{N}) \right)^{-1/2} \mathbf{V}_k,$$

y es claro que,

$$(Z_k(n-1; \mathbf{N}))^{-1/2} \geq \left(\sum_{i=0}^{n-1} Z_k(i; \mathbf{N}) \right)^{-1/2}$$

Las Figuras 3.6 y 3.7 ilustran el comportamiento de los estimadores $\tilde{\mathbf{V}}_k(n; \mathbf{N})$ y $\hat{\mathbf{V}}_k(n; \mathbf{N})$, de las matrices de varianzas y covarianzas \mathbf{V}_k , $k = 1, 2, 3$, de manera similar a como se realizó con la matriz de medias. Se muestran seis gráficos, dos por cada una de las d distribuciones; a la izquierda aparecen las matrices de varianzas y covarianzas $\tilde{\mathbf{V}}_k(n; \mathbf{N})$ estimadas a partir del estimador empírico de la matriz de medias y a la derecha $\hat{\mathbf{V}}_k(n; \mathbf{N})$ estimadas a partir del estimador máximo verosímil de la matriz de medias. De nuevo la líneas punteadas indican los valores de las matrices de varianzas y covarianzas de las distribuciones 1, 2 y 3, (3.16), (3.17) y (3.18), respectivamente.

En las Figuras 3.6 y 3.7 se observa que, al menos en esta simulación específica aparentemente, el estimador $\tilde{\mathbf{V}}_k(n; \mathbf{N})$ presenta una más rápida convergencia y no tiende a subestimar la varianza como ocurre con el estimador $\hat{\mathbf{V}}_k(n; \mathbf{N})$, por otro lado, desde el punto de vista teórico, una de las conclusiones de Maaouia & Touati (2005) es que el análisis de cual es el mejor estimador para las matrices de varianzas y covarianzas hace parte de una tesis en desarrollo, sin embargo, si se llega a un escenario similar al de la estimación de la varianza para un proceso de BWG, se esperaría que el estimador $\tilde{\mathbf{V}}_k(n; N)$ sea asintóticamente más eficiente que el estimador $\hat{\mathbf{V}}_k(n; N)$, coincidiendo con lo observado.

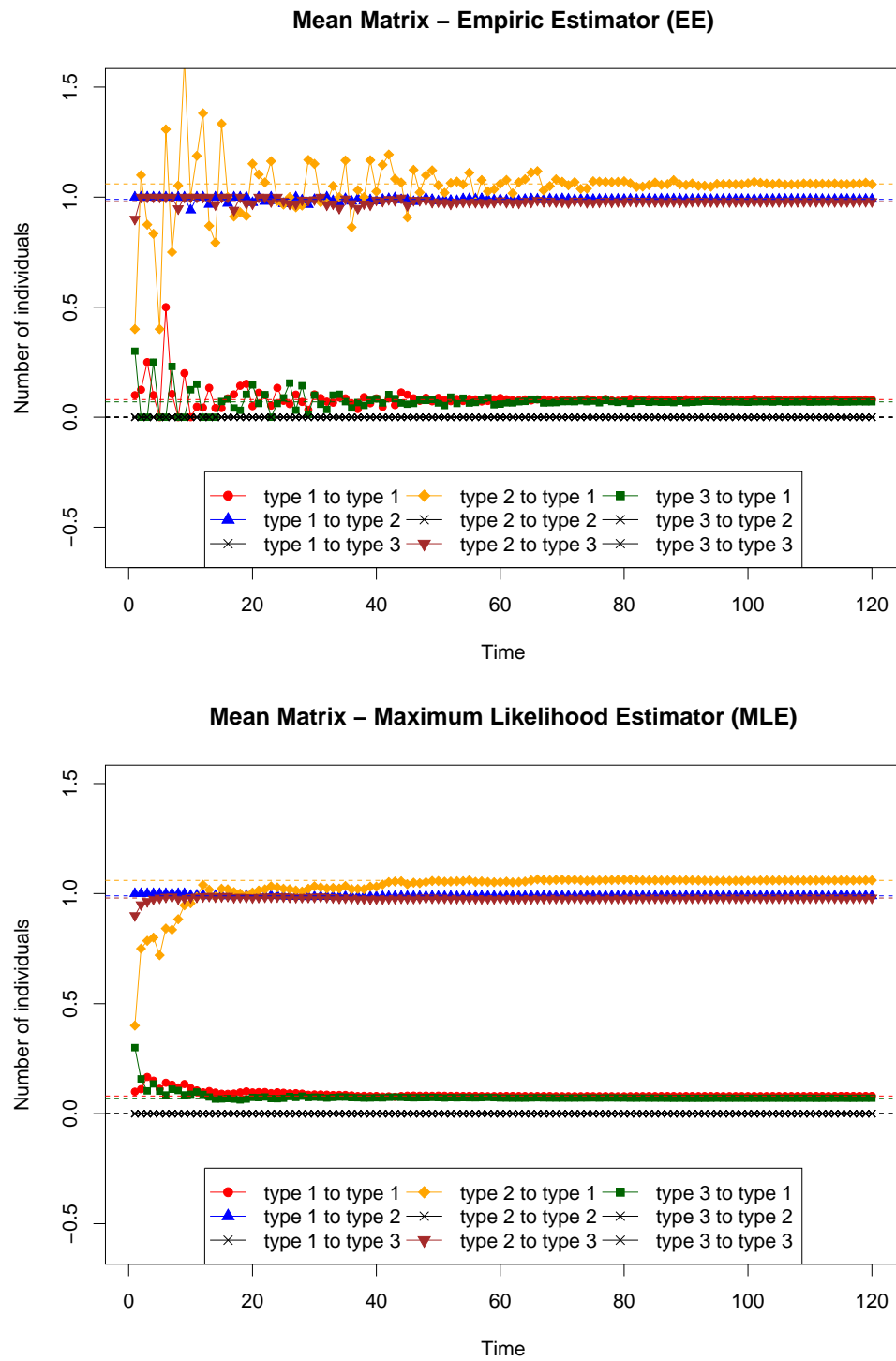


Figura 3.5: Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias cuando el tiempo crece y el vector de individuos iniciales permanece fijo.

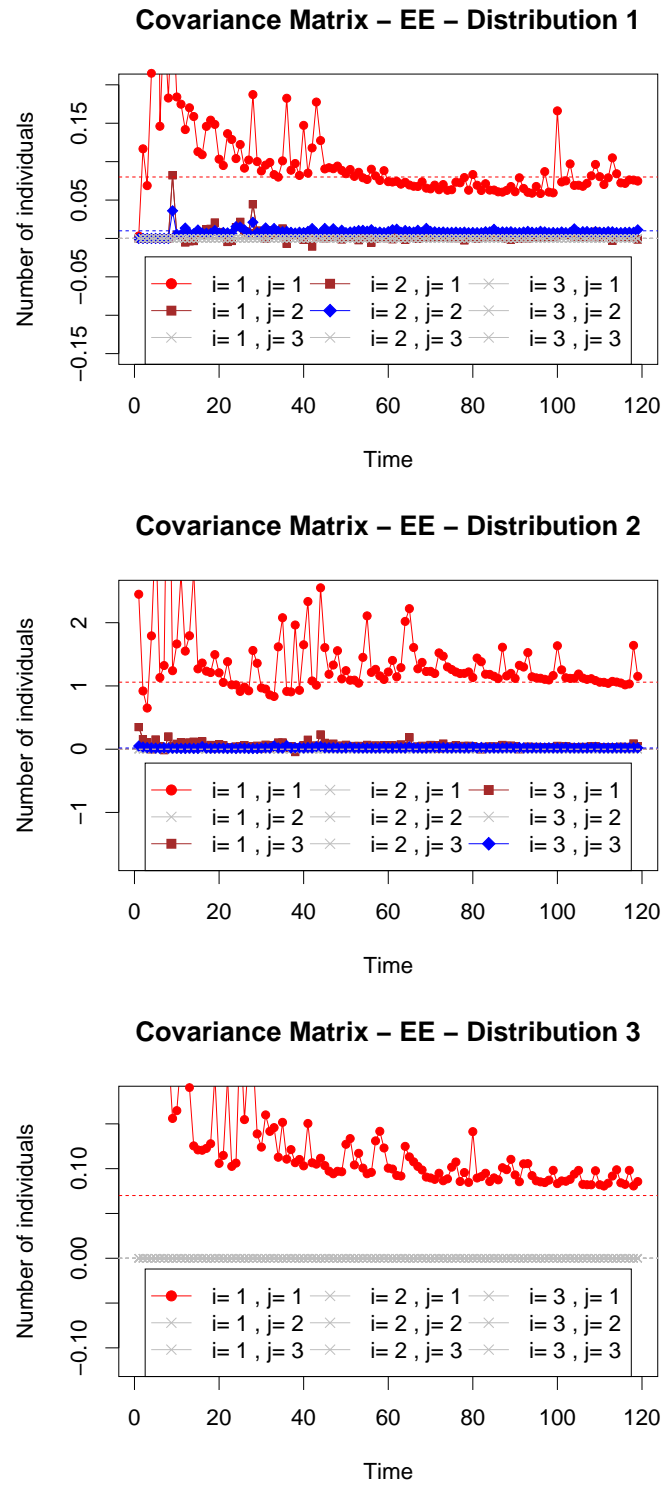


Figura 3.6: Convergencia del estimador $\tilde{V}_k(n; \mathbf{N})$ cuando el tiempo crece y el vector de individuos iniciales permanece fijo.

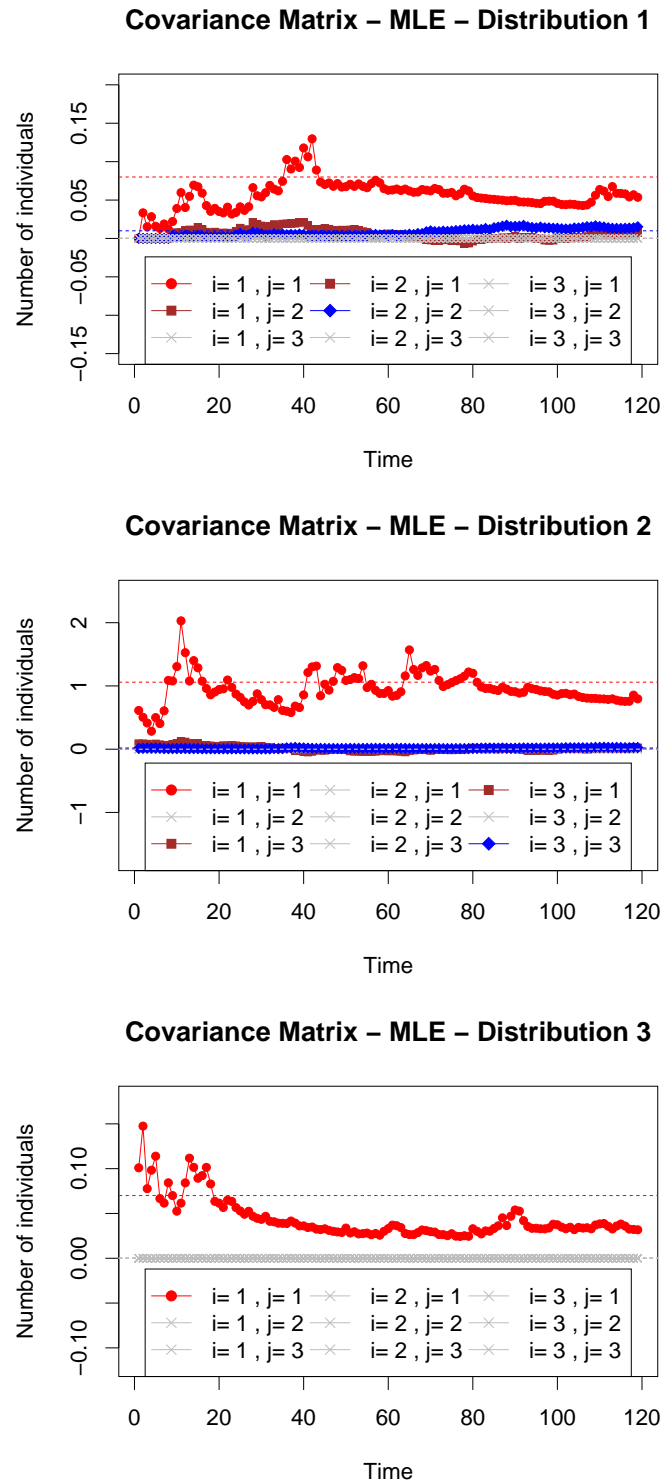


Figura 3.7: Convergencia del estimador $\hat{V}_k(n; \mathbf{N})$ cuando el tiempo crece y el vector de individuos iniciales permanece fijo.

3.3.2. Comportamiento asintótico de las frecuencias relativas

En esta sección se muestran numéricamente, los resultados teóricos del Capítulo 2, observando el comportamiento asintótico de las frecuencias relativas, asociadas a un proceso de BGWM supercrítico, crítico y subcrítico, cuando el tiempo permanece fijo y el número de individuos iniciales crece. Las simulaciones sobre las cuales se trabajará se construirán en función de la siguiente aplicación:

Ejemplo 3.2 (Un modelo de genética de poblaciones). En el campo de genética de poblaciones los procesos de ramificación tienen muchas aplicaciones diferentes, los procesos de BGWM tienen una relación cercana con algunos modelos de genética de poblaciones propuestos por S. Wright y R. A. Fisher los cuales estudian la deriva genética ¹ en poblaciones en donde su tamaño permanece constante por generaciones. Para el modelo de interés, se considera la evolución de una familia proveniente de un mismo ancestro, dentro de una población humana grande. El ancestro de la familia se empareja con un individuo escogido de manera aleatoria dentro de la población y produce un número aleatorio de descendientes, cada uno de estos descendientes se empareja con un individuo aleatorio por fuera de la familia y produce un número aleatorio de descendientes. El proceso continua en la medida que la familia siga teniendo algún individuo. Si adicionalmente, se considera la familia con respecto a un par de alelos auto-somales ² alternativos, (A , a), entonces, surge un proceso de ramificación multitempo con tres tipos de individuos, de acuerdo, a los genotipos AA , Aa y aa que poseen.

Supongamos que las frecuencias de los alelos auto-somales alternativos A y a dentro de la población general son p y q conocidas, con $p, q \geq 0$ y $p + q = 1$. Adicionalmente se asume que cada miembro de la familia se empareja de manera aleatoria con un individuo que no pertenece a la familia, lo cual, es equivalente a escoger un gameto de manera aleatoria dentro de la colección de genes de la población. Al considerar los tres genotipos AA , Aa y aa como los tipos 1, 2 y 3 del proceso de BGWM, y si se supone que no existen mutaciones, entonces las probabilidades de que un individuo con el genotipo AA , produzca un descendiente del genotipo AA , Aa o aa , son p , q y 0 respectivamente, las probabilidades de que un individuo con el genotipo Aa , produzca un descendiente del genotipo AA , Aa o aa , son $\frac{1}{2}p$, $\frac{1}{2}$ y $\frac{1}{2}q$ respectivamente, y, las probabilidades de que un individuo con el genotipo aa , produzca un descendiente del genotipo AA , Aa o aa , son 0, p y q . Además si μ_k con $k = 1, 2, 3$ es el promedio de descendientes que produce un individuo de tipo 1, 2 y 3 respectivamente, entonces la matriz de medias del proceso está dada por:

$$\mathbf{m} = \begin{pmatrix} \mu_1 p & \mu_1 q & 0 \\ \frac{1}{2} \mu_2 p & \frac{1}{2} \mu_2 & \frac{1}{2} \mu_2 q \\ 0 & \mu_3 p & \mu_3 q \end{pmatrix} \quad (3.19)$$

Cuando $\mu_i = \mu_j$ para todo i, j , se dice que no hay selección natural, ya que en este caso no hay reproducción diferencial de unas variantes genéticas respecto de otras ³.

¹“La deriva genética es la fluctuación al azar, de las frecuencias por cada uno de los alelos, entre generaciones, que resulta del sorteo de genes, que se da durante la transmisión de gametos de los padres, para producir los genotipos hijos en poblaciones finitas. El resultado de la deriva suele ser la pérdida de variabilidad genética, siendo un proceso que contrarresta la entrada de variabilidad genética por mutaciones.”(Barbadilla 2010)

²auto-somales se refiere a alelos de cromosomas diferentes a los que definen el sexo del individuo.

³“La selección natural se define como el proceso que resulta del cumplimiento de las siguientes tres condiciones: (1) variación fenotípica entre los individuos de una población, (2) supervivencia o reproducción diferencial asociada a la variación, y (3) herencia de la variación. Si en una población de organismos se dan estas tres condiciones, entonces se sigue necesariamente un cambio en la composición genética de la población por selección natural.”(Barbadilla 2010)

Los resultados expuestos en la Proposición 2.1 y el Teorema 2.1, se tienen, sin importar si el proceso es subcrítico, crítico o supercrítico, por tal razón se simulan tres modelos distintos, en donde se toman $p = 0.6$, $q = 0.4$ y los valores de los μ_k de cada modelo de acuerdo a la Tabla 3.5.

Tabla 3.5: Parámetros de las distribuciones de descendencia, para los modelos D, E y F.

	Modelo D.	Modelo E.	Modelo F.
μ_1	0.8	1	1.2
μ_2	0.7	1	1.1
μ_3	0.6	1	1

Teniendo en cuenta (3.19) y los parámetros de la Tabla 3.5, en la Tabla 3.6 se muestran la matriz de medias de cada modelo con su respectivo mayor valor propio ρ .

Tabla 3.6: Matrices de medias y mayor valor propio respectivo, para los modelos D, E y F.

	Modelo D	Modelo E	Modelo F.
\mathbf{m}	$\begin{pmatrix} 0.48 & 0.32 & 0 \\ 0.21 & 0.35 & 0.14 \\ 0 & 0.36 & 0.24 \end{pmatrix}$	$\begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.6 & 0.4 \end{pmatrix}$	$\begin{pmatrix} 0.72 & 0.48 & 0 \\ 0.33 & 0.55 & 0.22 \\ 0 & 0.6 & 0.4 \end{pmatrix}$
ρ	0.7264	1	1.1242

Por la Proposición 2.1 se tiene que la frecuencias relativas definidas en (2.8), con $t = 3$, convergen, cuando $N_1 \rightarrow \infty$, al vector $\mathbf{p}(3; \mathbf{e}_1)$, cuyos valores puntuales para cada modelo, se encuentran en la Tabla 3.7.

Tabla 3.7: Vectores $\mathbf{m}(3; \mathbf{e}_1)$ y $\mathbf{p}(3; \mathbf{e}_1)$, para los modelos D, E y F.

	Modelo D	Modelo E	Modelo F.
$\mathbf{m}(3; \mathbf{e}_1)$	(0.1986, 0.2043, 0.0479)	(0.42, 0.46, 0.12)	(0.6885, 0.7235, 0.1764)
$M(3; \mathbf{e}_1)$	0.4509	1	1.5883
$\mathbf{p}(3; \mathbf{e}_1)$	(0.4405, 0.4537, 0.1063)	(0.42, 0.46, 0.12)	(0.4334, 0.4555, 0.1110)

Tabla 3.8: Distribuciones de descendencia seleccionadas, para los modelos D, E y F.

	Modelo D.	Modelo E.	Modelo F.
$\eta_1 \sim$	Poisson(0.8)	Poisson(1)	Poisson(1.2)
$\eta_2 \sim$	Binomial(7, 0.1)	Binomial(4, 0.25)	Binomial(11, 0.1)
$\eta_3 \sim$	Geométrica(0.625)	Geométrica(0.5)	Geométrica(0.5)

El modelo del Ejemplo 3.2 tiene las características propias de un proceso de BGWM con distribución multinomial, por lo tanto, las simulaciones se realizaron mediante el Algoritmo 3.2 con los siguientes parámetros: $d = 3$, $n^* = 3$, $\mathbf{N} = (N_1, 0, 0)$ con N_1 tomando los valores 10 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 y 10^7 ,

$$B = \begin{pmatrix} 0.6 & 0.4 & 0 \\ 0.3 & 0.5 & 0.2 \\ 0 & 0.6 & 0.4 \end{pmatrix}$$

y en la tabla 3.8 se indican las distribuciones que fueron seleccionadas, dicha selección no tuvo una razón en particular más allá del interés de utilizar diferentes distribuciones discretas.

Es así como al realizar las simulaciones planteadas, se obtuvo la Figura 3.8, en donde se observa la convergencia de las frecuencias relativas de los tres modelos (D, E y F), a los valores del vector $\mathbf{p}(3; \mathbf{e}_1)$ correspondiente (ver Tabla 3.7), a medida que los valores de N_1 aumentaron.

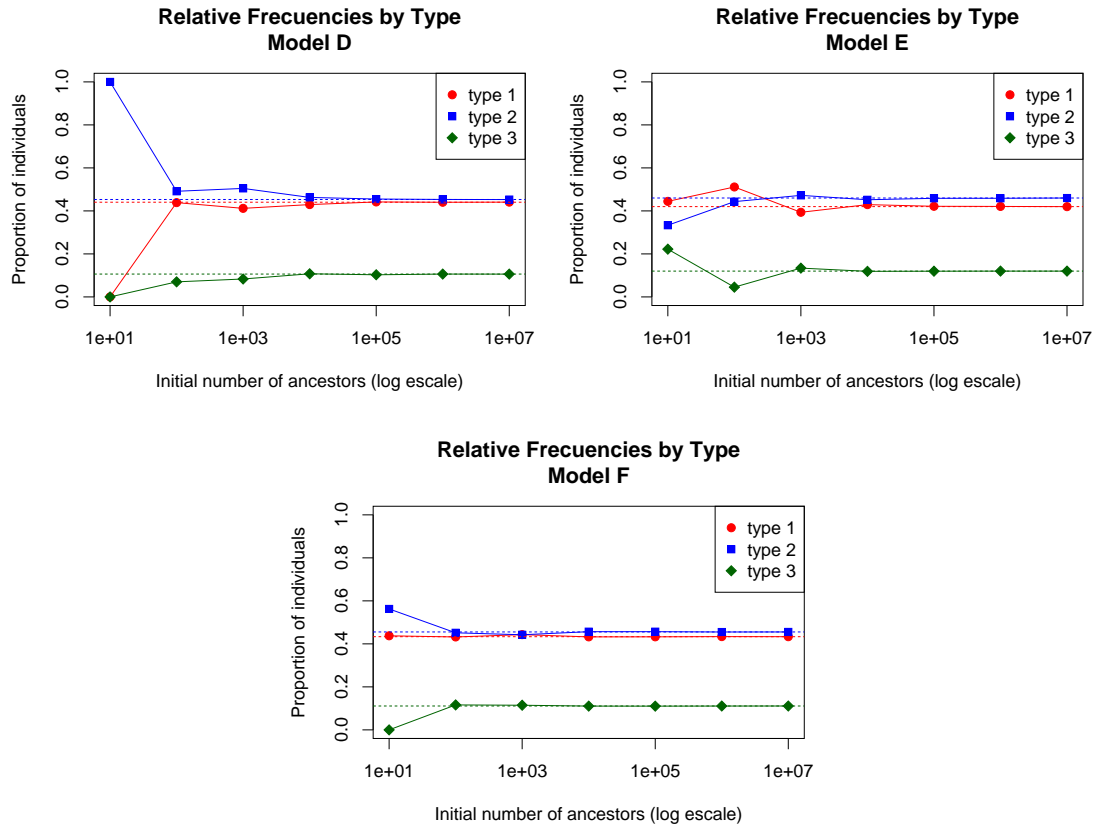


Figura 3.8: Convergencia de las frecuencias relativas cuando el tiempo permanece fijo y el número de individuos iniciales de tipo 1 crece.

Del Teorema 2.1 se tiene que, cuando $N_1 \rightarrow \infty$, las frecuencias relativas $\Delta_k(t; N_1)$ convergen en distribución a una normal, y que cualquier vector formado por hasta $d - 1$ frecuencias relativas asociadas a diferentes tipos, converge en distribución a una normal multivariada. Para verificar que lo anterior se cumpliera para los modelos D, E y F, se realizaron 1000 simulaciones por modelo, tomando $N_1 = 10^7$ y manteniendo los demás parámetros sin modificaciones.

Las mil simulaciones por modelo, se consideran como mil observaciones o realizaciones de variables aleatorias relacionadas con el proceso de BGWM, es así como, para cada modelo se tienen mil observaciones del vector aleatorio,

$$\left(\Delta_1(3; (10^7, 0, 0)), \Delta_2(3; (10^7, 0, 0)), \Delta_3(3; (10^7, 0, 0)) \right)$$

que corresponde a las frecuencias relativas, en la tercera generación a partir de diez millones de individuos iniciales de tipo 1. Por ejemplo, la primera observación del vector aleatorio en cuestión, para el modelo D - caso subcrítico, mediante la primera simulación, fue:

$$(0.4410189, 0.4527683, 0.1062128).$$

A partir de estas observaciones, se produjeron las Figuras 3.9, 3.10 y 3.11, que corresponden a los histogramas de las frecuencias relativas por tipo de individuo, para cada uno de los tres modelos, y las Tablas 3.9, 3.10 y 3.11, en donde se indican los valores del estadístico W y el p -valor asociado, obtenidos al realizar una prueba de normalidad de Shapiro-Wilk, en R, sobre dichas observaciones, mediante la función `shapiro.test` del paquete `stats` (R Development Core Team 2009a).

En los histogramas mencionados se observa una forma acampanada y simétrica, por otra parte, a partir de los tres p -valores de cada una de las tres tablas, se puede decir que no se rechaza la hipótesis nula de que cada una de las frecuencias relativas de cada uno de los modelos, proviene de una distribución normal univariada, tomando un nivel de significancia del 5%, ya que el menor de los nueve p -valores es 0.2138 (ver Tabla 3.11), obtenido en el modelo F - caso supercrítico, con las observaciones asociadas a la frecuencia relativa del tercer tipo de individuos, es decir, a la variable aleatoria $\Delta_3(3; (10^7, 0, 0))$.

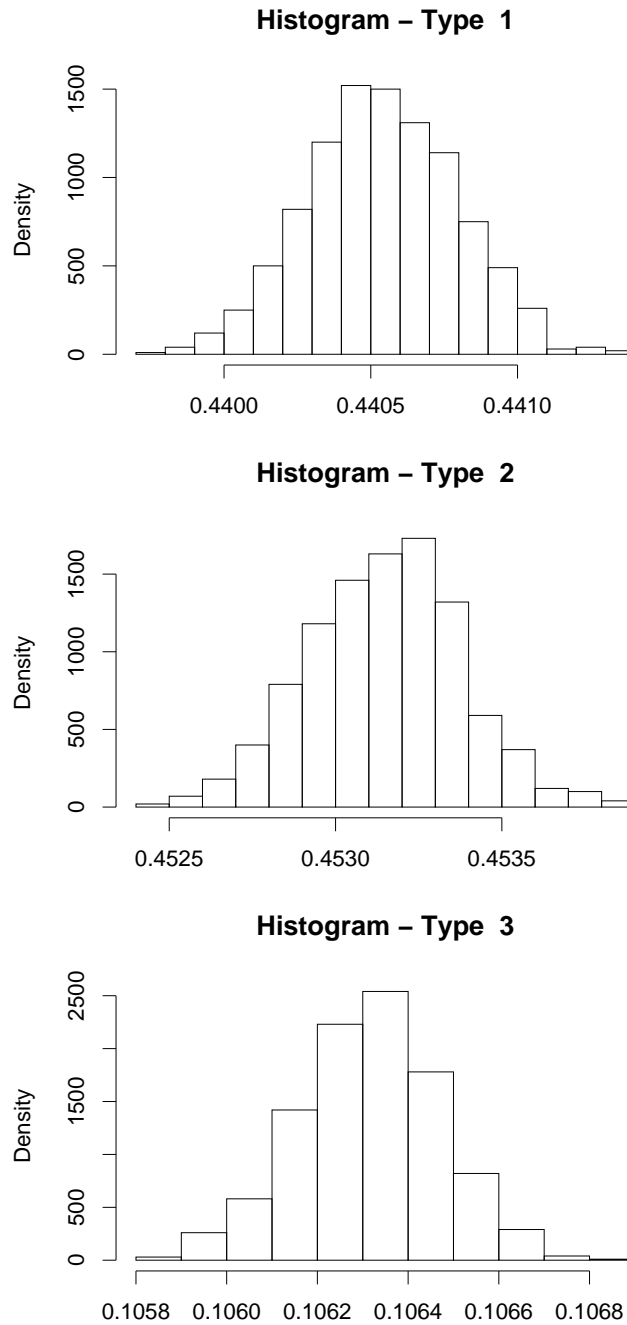


Figura 3.9: Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo D - Caso subcrítico (tamaño de muestra = 1000).

Tabla 3.9: Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo D - Caso subcrítico (tamaño de muestra = 1000).

	Tipo 1	Tipo 2	Tipo 3
Estadístico W	0.9989	0.9985	0.9989
p -valor	0.8601	0.5408	0.7899

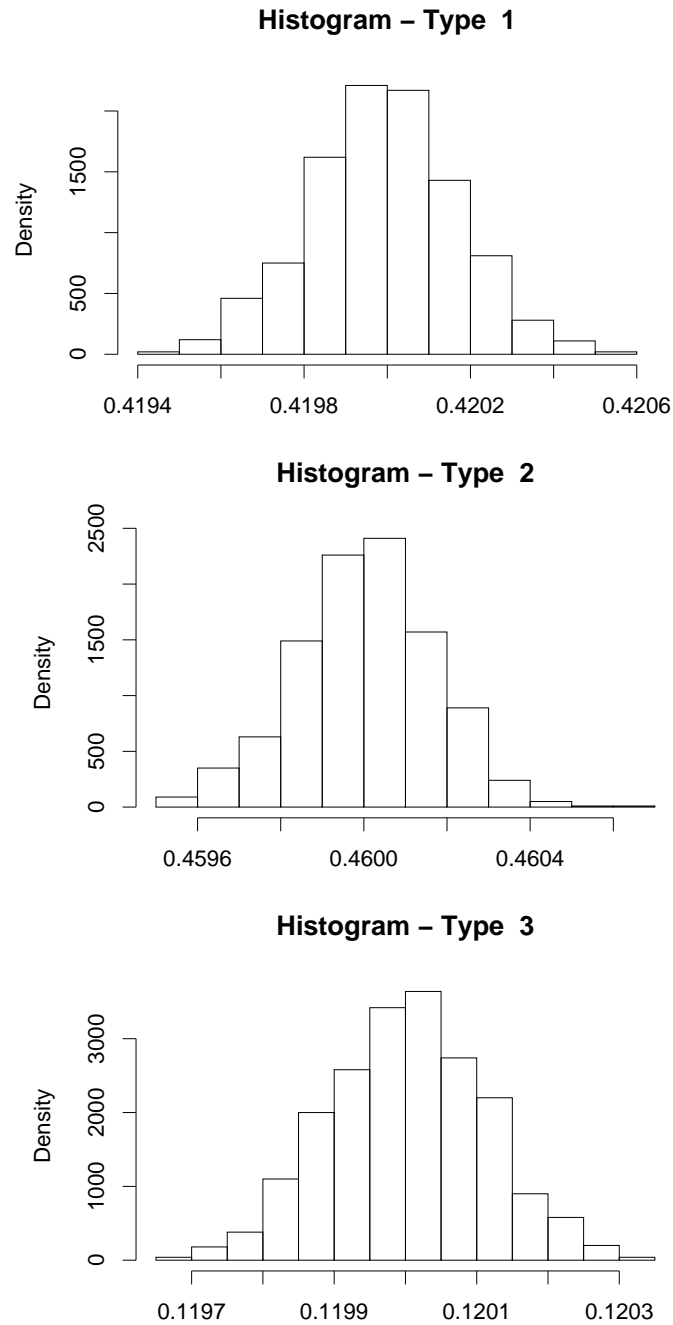


Figura 3.10: Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo E - Caso crítico (tamaño de muestra = 1000).

Tabla 3.10: Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo E - Caso crítico (tamaño de muestra = 1000).

	Tipo 1	Tipo 2	Tipo 3
Estadístico W	0.9992	0.9980	0.9988
p -valor	0.9517	0.2878	0.7876

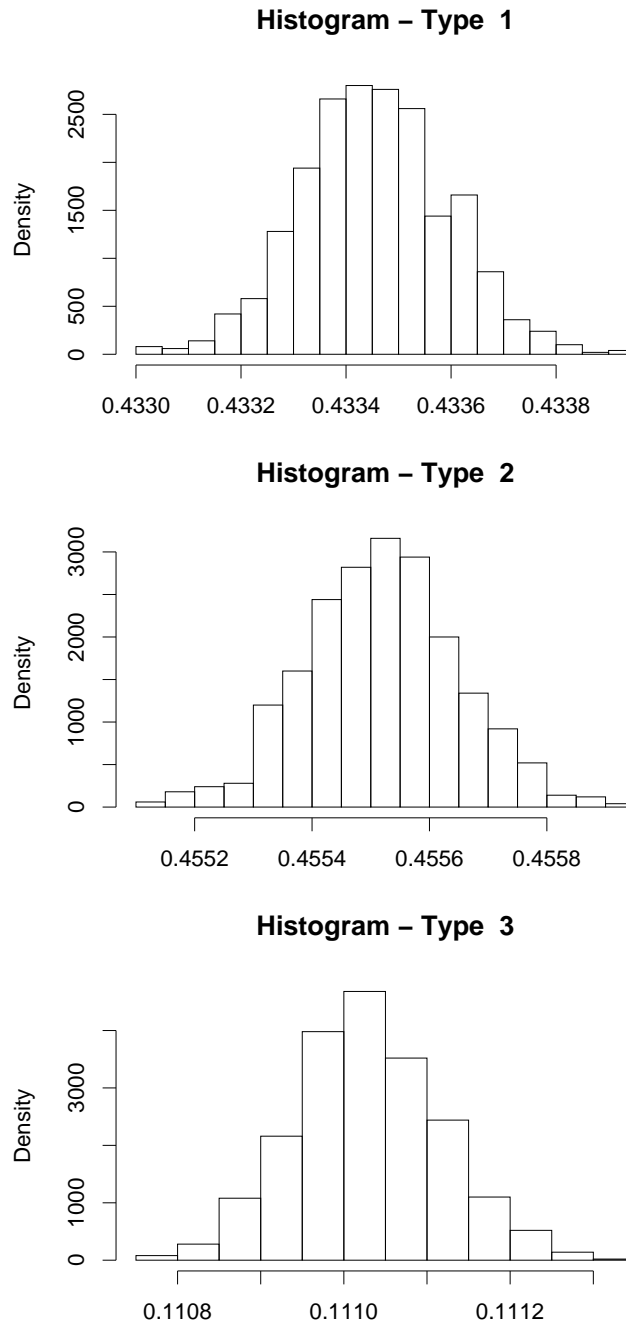


Figura 3.11: Histogramas de las frecuencias relativas de cada tipo de individuo para el modelo F - Caso supercrítico (tamaño de muestra = 1000).

Tabla 3.11: Pruebas de normalidad univariada de Shapiro-Wilk, por tipo de individuos, para el modelo F - Caso supercrítico (tamaño de muestra = 1000).

	Tipo 1	Tipo 2	Tipo 3
Estadístico W	0.9989	0.9988	0.9978
p -valor	0.7640	0.7369	0.2138

Con las anteriores pruebas de normalidad solo se ha observado numéricamente la parte (i) del Teorema 2.1, para la parte (ii) se debe ver, a partir de las mil simulaciones, que cada uno de los vectores aleatorios:

$$\Delta_{1,2} := \left(\Delta_1(3; (10^7, 0, 0)), \Delta_2(3; (10^7, 0, 0)) \right)$$

$$\Delta_{1,3} := \left(\Delta_1(3; (10^7, 0, 0)), \Delta_3(3; (10^7, 0, 0)) \right)$$

$$\Delta_{2,3} := \left(\Delta_2(3; (10^7, 0, 0)), \Delta_3(3; (10^7, 0, 0)) \right)$$

tiene una distribución normal multivariada.

De manera similar a lo expuesto para el caso univariado, se muestra el equivalente bivariado de un diagrama de polígono, correspondiente a los modelos D, E y F, (Figuras 3.12, 3.14 y 3.16, respectivamente).

Adicionalmente, se obtienen las Figuras 3.13, 3.15 y 3.17, en donde, para cada modelo, se tiene un diagrama de dispersión de cada uno de los conjuntos de observaciones asociados a los vectores aleatorios $\Delta_{1,2}$, $\Delta_{1,3}$ y $\Delta_{2,3}$. Por ejemplo, la primera de las mil observaciones para dichos vectores aleatorios, para el modelo D - caso subcrítico, obtenida mediante la primera simulación, es:

$$(0.4410189, 0.4527683), (0.4410189, 0.1062128) \text{ y } (0.4527683, 0.1062128)$$

respectivamente, y estos valores quedan representados como un punto en cada uno de los planos, en su orden, de la Figura 3.13. Además, cada plano tiene cuatro curvas de nivel de la distribución normal multivariada teórica asociada a las observaciones, cada una de las cuatro curvas de nivel, de afuera hacia adentro, corresponde a una elipse con centro en el vector de medias, que en su interior para la distribución teórica, tiene una probabilidad acumulada de, 0.99, 0.95, 0.8 y 0.5, respectivamente. Estos planos se conocen como los diagramas de contorno y buscan contrastar los valores observados contra la distribución teórica asociada.

En las Tablas 3.12, 3.13 y 3.14 se indican los valores del estadístico W y el p -valor asociado, obtenidos al realizar una prueba de normalidad multivariada de Shapiro-Wilk, en \mathbb{R} , sobre las tres observaciones en dos dimensiones, mediante la función `mshapiro.test` del paquete `mvnormtest` (Jarek 2009).

De nuevo, no se rechaza la hipótesis nula con un nivel de significancia del 5%, lo que quiere decir que, a partir de las observaciones, no hay evidencia que muestre que los vectores aleatorios $\Delta_{1,2}$, $\Delta_{1,3}$ y $\Delta_{2,3}$, no tienen una distribución normal multivariada. El menor de los nueve p -valores es 0.0630 (ver Tabla 3.14), obtenido en el modelo F - caso supercrítico, con el vector de observaciones asociado a las frecuencias relativas del primero y segundo tipo de individuos, es decir al vector aleatorio $\Delta_{1,2}$.

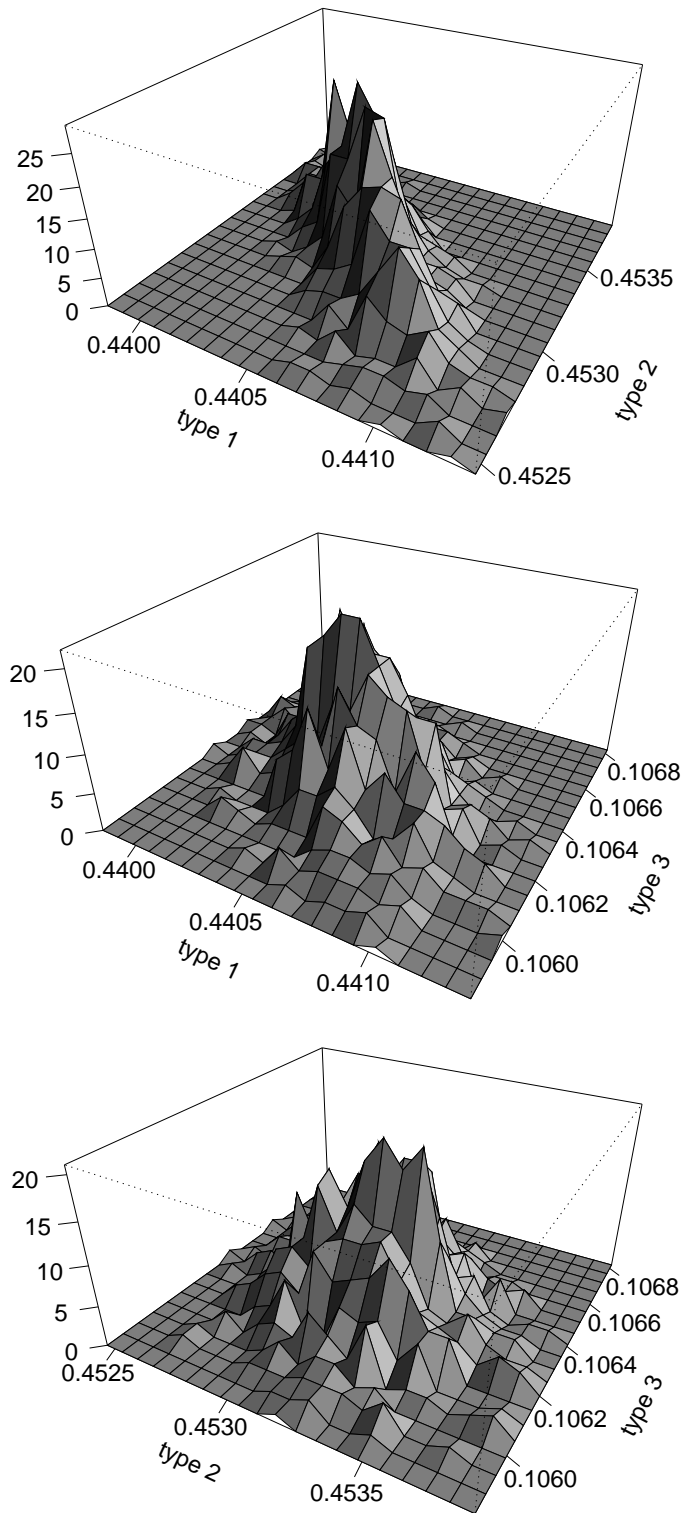


Figura 3.12: Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).

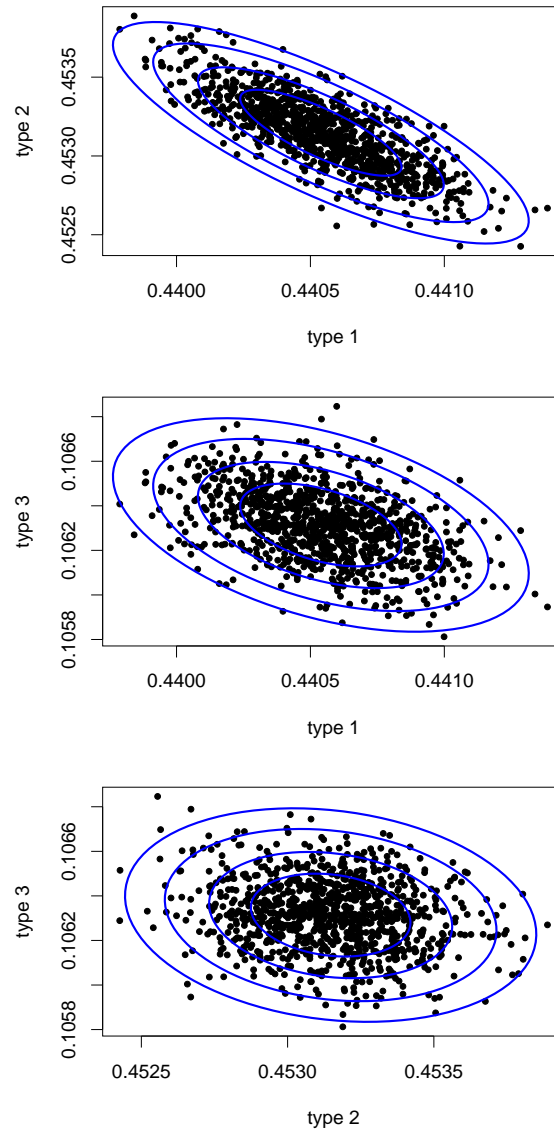


Figura 3.13: Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000).

Tabla 3.12: Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo D - Caso subcrítico (tamaño de muestra = 1000).

	1-2	1-3	2-3
Estadístico W	0.9971	0.9971	0.9971
Valor p	0.0728	0.0726	0.0729

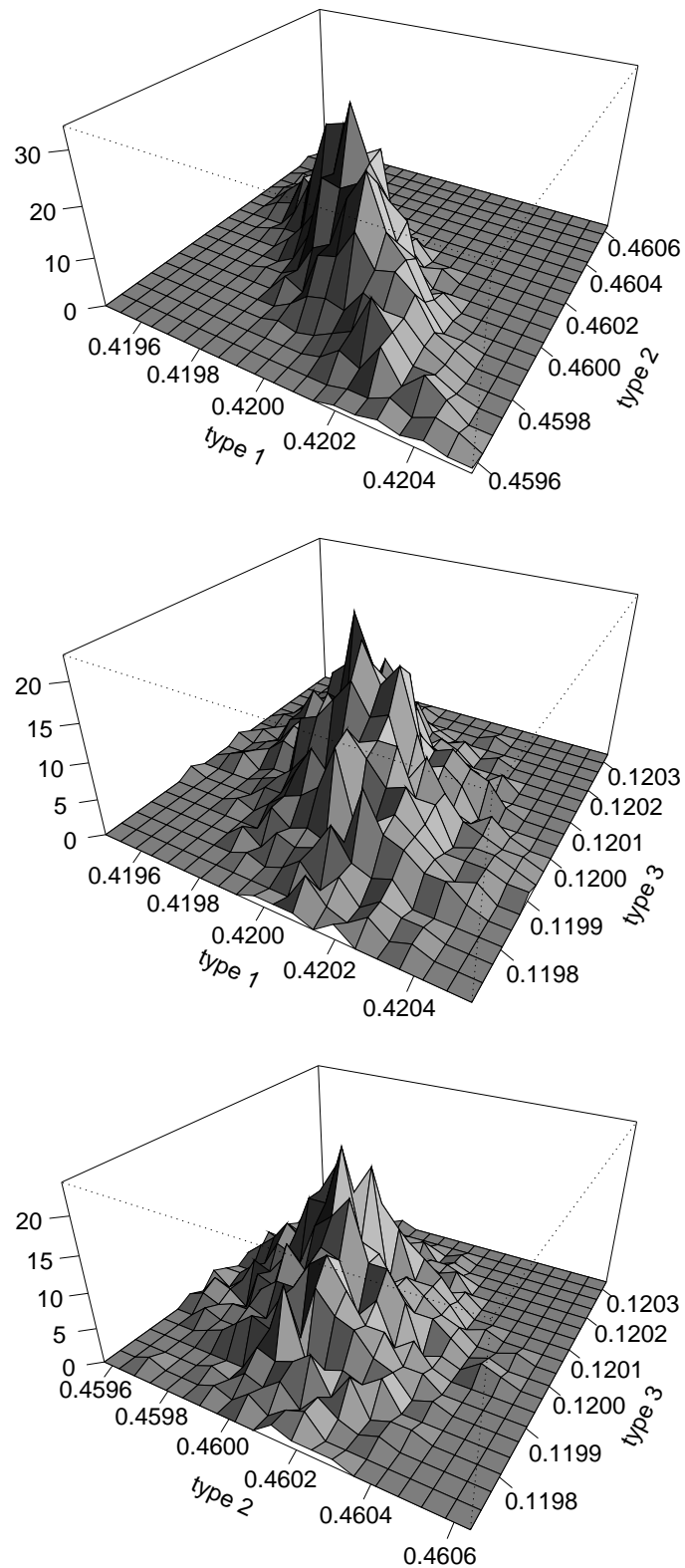


Figura 3.14: Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).

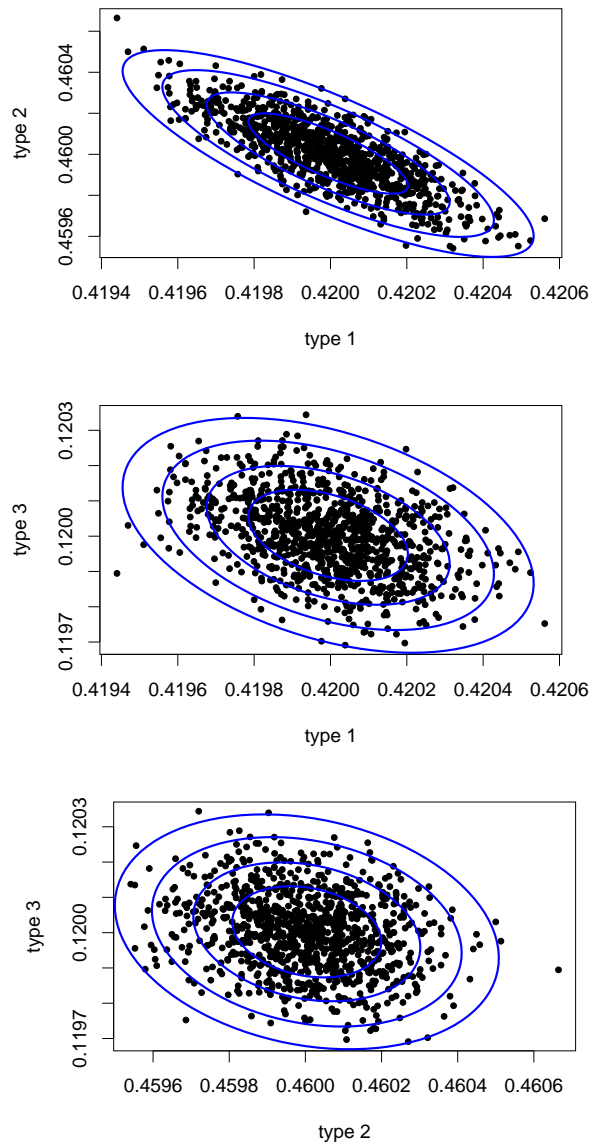


Figura 3.15: Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000).

Tabla 3.13: Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo E - Caso crítico (tamaño de muestra = 1000).

	Tipo 1-2	Tipo 1-3	Tipo 2-3
Estadístico W	0.9981	0.9981	0.9981
p -valor	0.3257	0.3235	0.3255

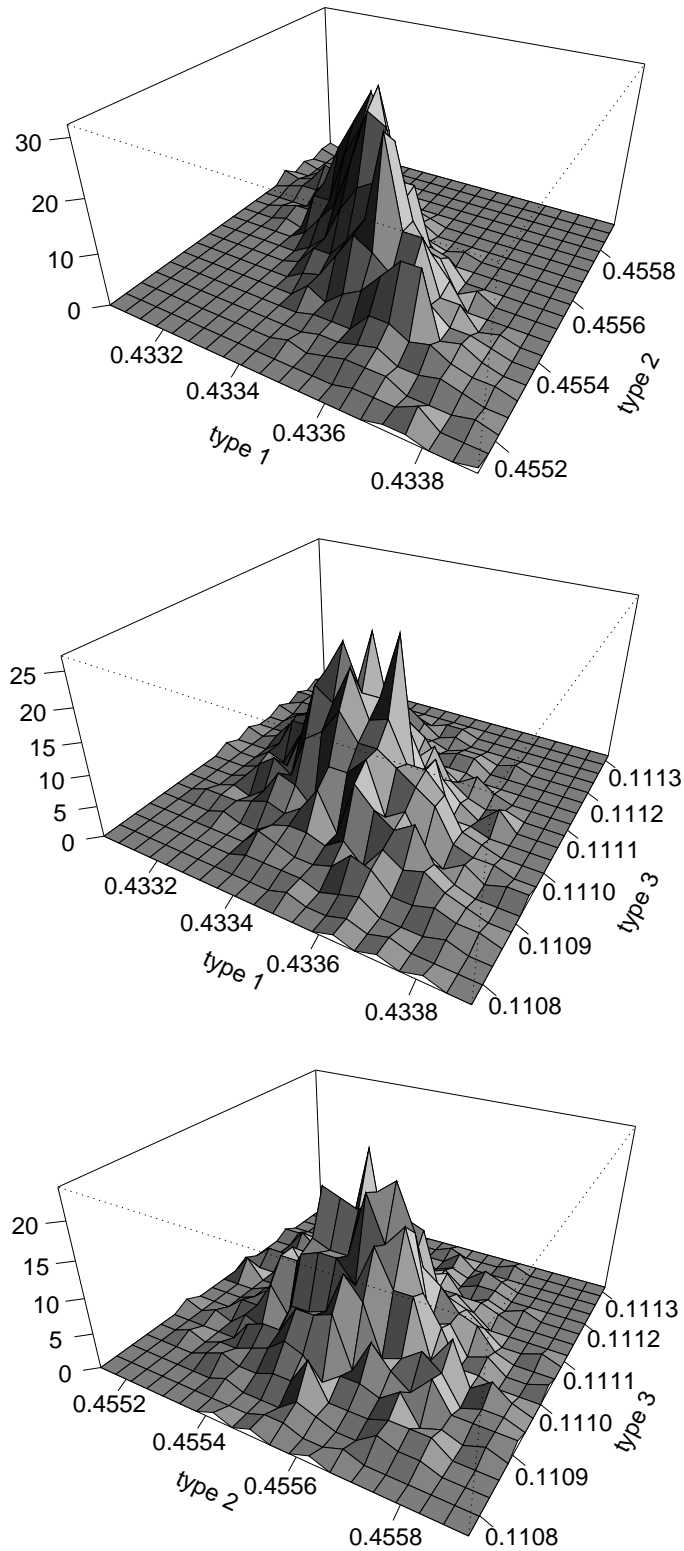


Figura 3.16: Diagrama de poliedro de la distribución empírica bivariada de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000, número de intervalos = 20 x 20).

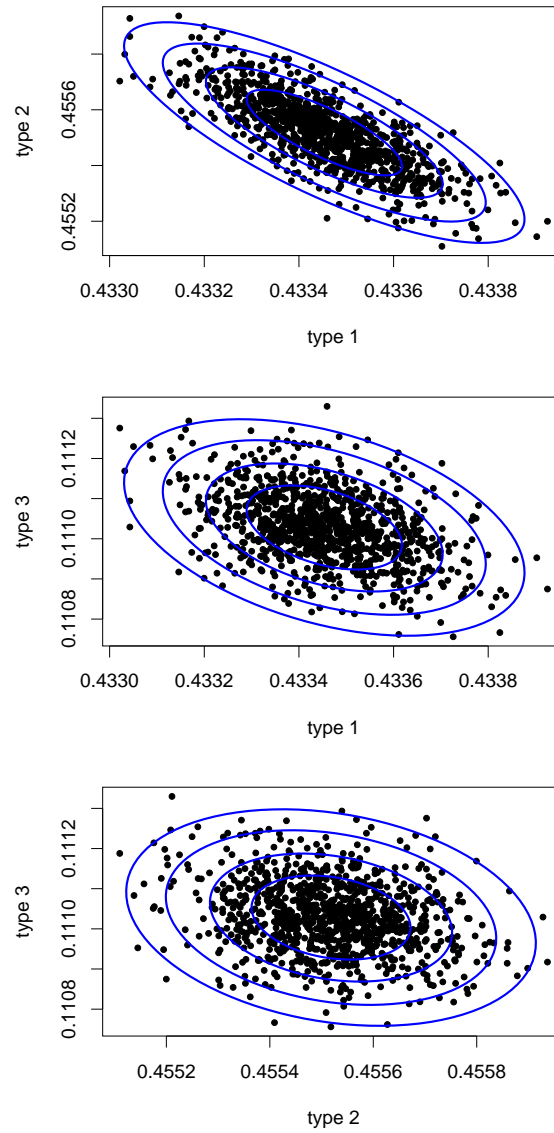


Figura 3.17: Diagrama de contorno de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000).

Tabla 3.14: Pruebas de normalidad multivariada de Shapiro-Wilk de la combinación por tipos de las frecuencias relativas, modelo F - Caso supercrítico (tamaño de muestra = 1000).

	Tipo 1-2	Tipo 1-3	Tipo 2-3
Estadístico W	0.9971	0.9971	0.9971
p -valor	0.0630	0.0710	0.0711

Resultados empíricos

Además de ilustrar y observar, por medio de simulaciones, los resultados existentes acerca de la estimación de parámetros y el comportamiento de las frecuencias relativas en un proceso de Bienaymé - Galton - Watson Multitipo (BGWM), este trabajo tiene como objetivo explorar mediante simulaciones, cuál sería el comportamiento asintótico de los estimadores y de las frecuencias relativas cuando el número de individuos iniciales y el tiempo tienden a infinito. En este capítulo se muestran los resultados de dicha exploración.

El comportamiento de los procesos de ramificación cuando el tiempo tiende a infinito ha tenido un constante interés por parte de los investigadores, pero al realizar una búsqueda de artículos que hayan considerado el comportamiento asintótico cuando el número de individuos tiende a infinito, se observa que este otro enfoque no ha tenido un estudio tan amplio.

Yakovlev & Yanev (2009), como se explica en el capítulo 2 de este trabajo, analizan el comportamiento de las frecuencias relativas, para cualquier proceso de ramificación, cuando el número de individuos iniciales tiende a infinito, mientras el tiempo se deja fijo, por otra parte, en la sección de discusión y conclusiones de este artículo, se menciona que, podría ser interesante investigar el comportamiento asintótico de las frecuencias relativas cuando el número de individuos iniciales y el tiempo tienden de manera simultánea a infinito, y anticipan que dicho comportamiento va a depender del tipo de proceso de ramificación y de si el proceso es subcrítico, crítico o supercrítico.

Dion & Yanev (1994), Dion & Yanev (1997) obtienen resultados para el proceso de Bienaymé - Galton - Watson (BGW) o el proceso de Bienaymé - Galton - Watson con Inmigración (BGWI), acerca del comportamiento asintótico de algunos estimadores, tanto de la media, como de la varianza, cuando se tiene un número creciente y aleatorio de individuos iniciales, a la vez que el tiempo tiende a infinito. Una de las conclusiones a las que llegaron es que, al tener el siguiente supuesto:

$$\left(m > 1\right), \text{ o, } \left(m = 1 \text{ y } n/N \rightarrow 0\right), \text{ o, } \left(m < 1 \text{ y } N m^n \rightarrow \infty\right) \quad (4.1)$$

los estimadores de un proceso de BGW para la media y la varianza, son consistentes y normalmente distribuidos, cuando el número inicial de individuos y el tiempo tienden a infinito.

Adicionalmente, Dion & Yanev (1994) mencionan que, Badalbaev y Mukhitdinov en 1990, analizaron el comportamiento del estimador empírico, cuando el número de individuos iniciales y el tiempo tienden a infinito, para el proceso de Bienaymé - Galton - Watson simple y multitipo.

La pregunta en este punto sería, ¿es posible extender la anterior conclusión para los estimadores que se han venido utilizando a lo largo de este trabajo, en el caso de un proceso de BGWM?

4.1. Resultados para estimadores de la matriz de medias

En esta sección se exponen los resultados de las simulaciones en cuanto a la convergencia del estimador empírico (EE) y el estimador máximo verosímil (MLE) de la matriz de medias del proceso de BGWM, cuando el número inicial de individuos del tipo 1 (N_1) y el tiempo (n) tienden a infinito.

Intuitivamente, es claro que los estimadores empírico y máximo verosímil de la matriz de medias de un proceso de BGWM, subcrítico o crítico, no convergen cuando el tiempo tiende a infinito. Esto debido a que el proceso cuando el tiempo tiende a infinito va a extinguirse. A medida que pasa el tiempo se pierden individuos, es decir, se pierde información a nivel numérico. Sin embargo, si se imponen condiciones de tal forma que no se llegue a esta extinción y, adicionalmente, el número de individuos con los cuales se estiman los parámetros es suficientemente grande, entonces, se espera que la estimación convergerá al parámetro respectivo.

Es por lo anterior que si se considera que el tiempo y el número de individuos iniciales tiende a infinito, en el caso subcrítico y crítico, se puede tener en cuenta cierta relación entre estas dos variables, para garantizar que la pérdida de individuos por la extinción, sea menor que el incremento debido al crecimiento de los individuos iniciales, de esa forma no se pierde información a nivel numérico y los estimadores convergerán a los parámetros deseados.

Se asumirá un supuesto similar al presentado en la ecuación 4.1, para el caso multitypo:

$$\left(\rho > 1\right), \text{ o, } \left(\rho = 1 \text{ y } n/N_1 \rightarrow 0\right), \text{ o, } \left(\rho < 1 \text{ y } N_1 \rho^n \rightarrow \infty\right) \quad (4.2)$$

Siguiendo la anterior idea y retomando los modelos D, E y F descritos en el Capítulo 3 planteados a partir del Ejemplo 3.2, se realizaron simulaciones con los mismos parámetros anteriormente utilizados, pero considerando adicionalmente una relación entre N_1 y n para cada modelo, como se describe en la Tabla 4.1, dichas relaciones hacen que para cada valor de n se tenga asociado un valor de N_1 y que cuando $n \rightarrow \infty$ se cumpla el supuesto 4.2. Adicionalmente, las relaciones fueron escogidas de tal forma que el número de individuos en el tiempo n del modelo, fueran suficientemente grandes para dar una buena precisión o aproximación a los valores de los parámetros.

Tabla 4.1: Relación entre N_1 y n , para los modelos D, E y F.

Modelo D.	Modelo E.	Modelo F.
$N_1 = \lceil 5000 * n / \rho^n \rceil$	$N_1 = 10 n^2$	$N_1 = n$

Para el modelo D, tenemos el caso subcrítico, $\rho < 1$ y,

$$\lim_{n \rightarrow \infty} N_1 \rho^n = \lim_{n \rightarrow \infty} \frac{5000 * n}{\rho^n} \rho^n = \lim_{n \rightarrow \infty} 5000 * n = \infty$$

Para el modelo E, tenemos el caso crítico, $\rho = 1$ y,

$$\lim_{n \rightarrow \infty} \frac{n}{N_1} = \lim_{n \rightarrow \infty} \frac{n}{10 n^2} = \lim_{n \rightarrow \infty} \frac{1}{10 n} = 0$$

De los valores resultantes de las simulaciones, se calcularon los estimadores empírico y máximo verosímil de la matriz de medias y se produjeron las Figuras 4.1, 4.2 y 4.3, para los modelos D, E y F, respectivamente. En dichas figuras se muestra el cambio de los valores estimados, a medida que el tiempo (n), y por la relación entre N_1 y n , el número de individuos iniciales de tipo 1 (N_1) crecen. Las líneas punteadas corresponden a los valores de la matriz de medias respectiva (3.6).

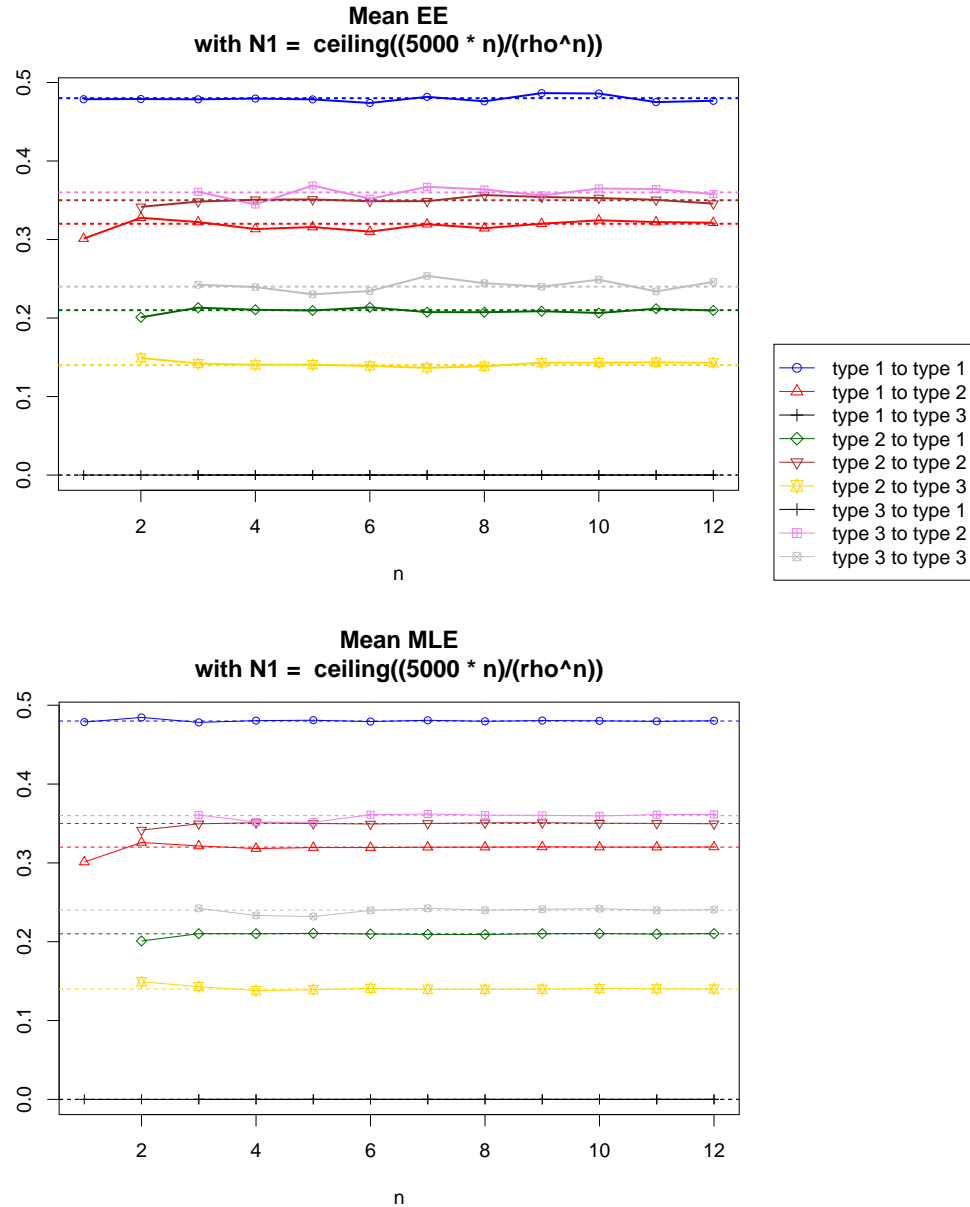


Figura 4.1: Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo D - caso subcrítico ($[N_1 = 5000 * n/\rho^n]$).

La Figura 4.1 muestra que en los primeros momentos en el tiempo, algún tipo de individuo se extinguía y por lo tanto no era posible estimar las medias asociadas a ese tipo, pero a medida que aumento el tiempo y el número de individuos iniciales no se presento de nuevo la situación de

extinción y las estimaciones siempre estuvieron cerca de los valores de los parámetros.

La Figura 4.2 muestra con mayor claridad una convergencia de los valores estimados. Cabe resaltar que en la matriz de medias para el modelo E, se tiene que: $m_{1,1} = m_{3,2} = 0.6$ y $m_{1,2} = m_{3,3} = 0.4$.

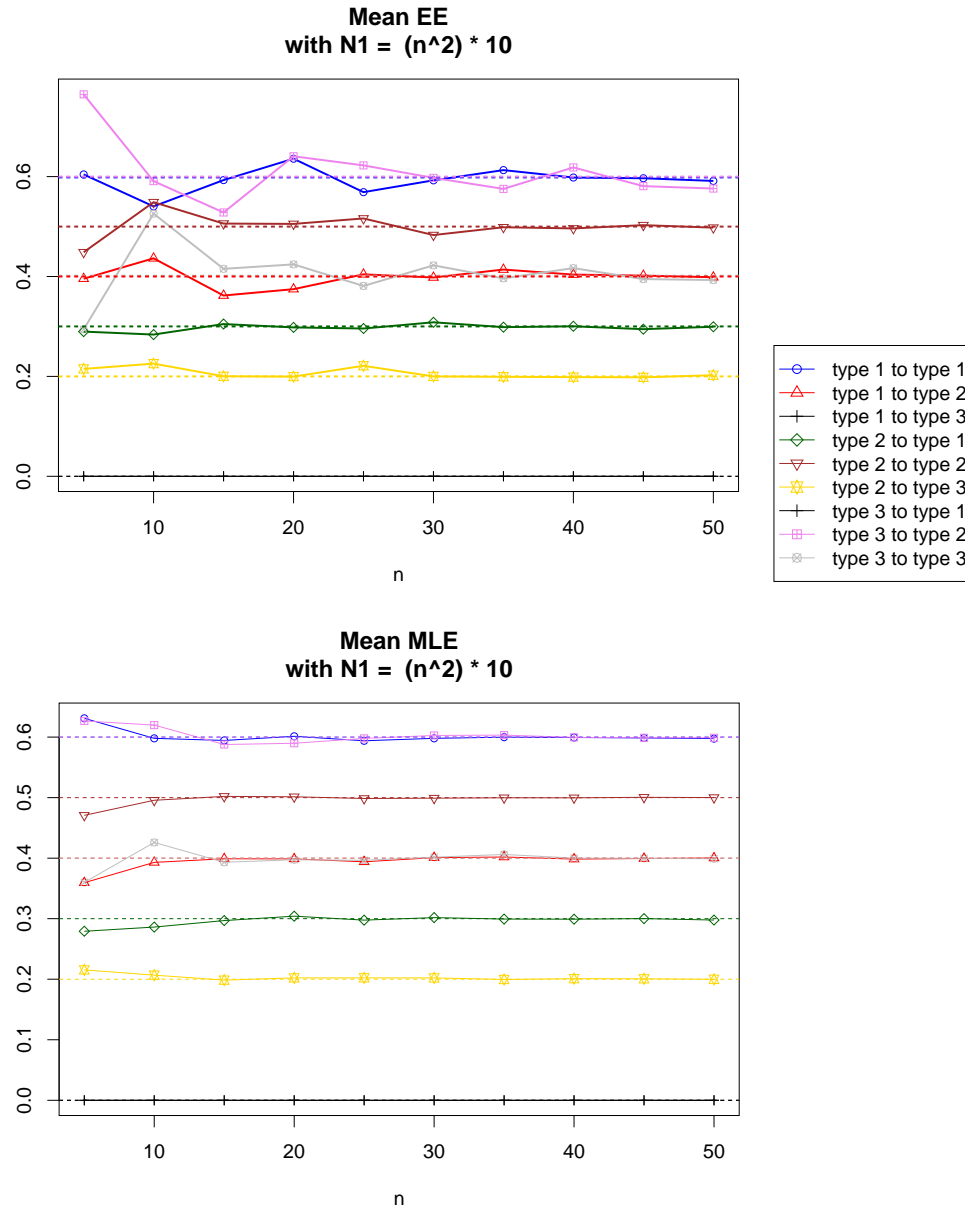


Figura 4.2: Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo E - caso crítico ($N_1 = 10n^2$).

La Figura 4.2 es prácticamente innecesaria porque la convergencia de los estimadores ha sido formalmente demostrada para el caso supercrítico (ver (Maaouia & Touati 2005)), en todo caso, dicha convergencia se evidencia para este ejemplo en particular, es decir, para el modelo F - caso

supercrítico.

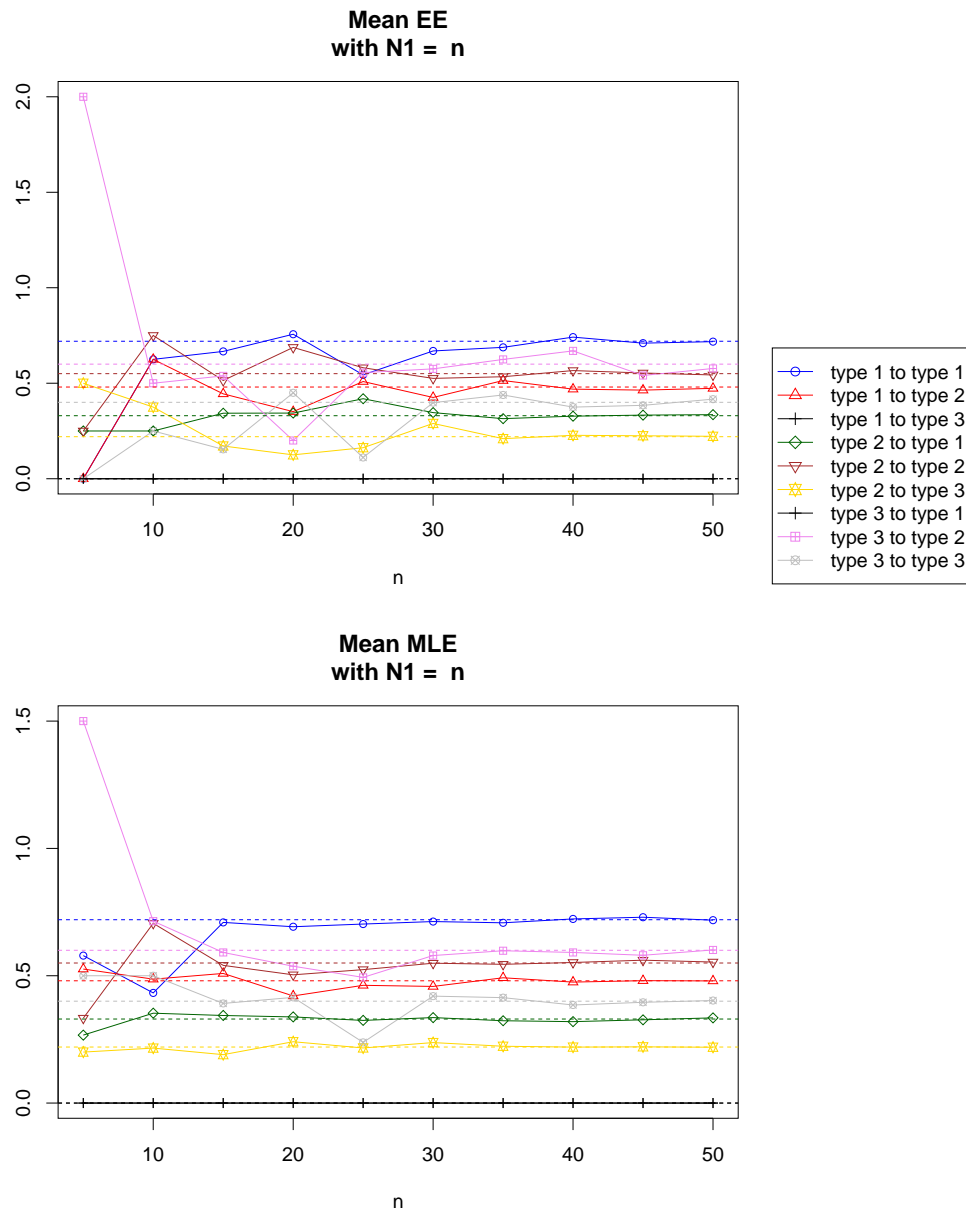


Figura 4.3: Convergencia de los estimadores empírico y máximo verosímil de la matriz de medias, cuando el tiempo y el número de individuos iniciales crece, para el modelo F - caso supercrítico ($N_1 = n$).

Dado lo que se observa mediante las simulaciones, en este trabajo se propone el siguiente nuevo resultado para que pueda ser demostrado o refutado, y el cual generalizaría para $0 < \rho < \infty$, lo obtenido en procesos de BGWM para $\rho > 1$.

Resultado 4.1. *Bajo las mismas hipótesis del Teorema 1.7, pero con $\mathbf{N}_1 = (N_1, 0, \dots, 0)$ vector de individuos iniciales, y con,*

$$\left(\rho > 1\right), o, \left(\rho = 1 \text{ y } n/N_1 \rightarrow 0\right), o, \left(\rho < 1 \text{ y } N_1 \rho^n \rightarrow \infty\right),$$

entonces, cuando n y N_1 tienden a infinito, $\tilde{\mathbf{m}}_k(n; \mathbf{N}_1)$ y $\hat{\mathbf{m}}_k(n; \mathbf{N}_1)$, son estimadores fuertemente consistentes de la matriz de medias \mathbf{m} , en el conjunto de no extinción,

$$\mathbf{T} = \left\{ \lim_{n, N_1 \rightarrow \infty} \mathbf{Z}(n; \mathbf{N}_1) = \infty \right\}$$

4.2. Resultados para las frecuencias relativas

En esta sección se muestran los resultados de las simulaciones en cuanto a la convergencia de las frecuencias relativas de un proceso de BGWM, cuando el número inicial de individuos de tipo 1 (N_1) y el tiempo (n) tienden a infinito.

De manera similar a lo que ocurre con los estimadores expuestos, las frecuencias relativas de los procesos de BGWM críticos y subcríticos, cuando el tiempo tiende a infinito, no convergen, debido básicamente a la extinción de dichos procesos.

Dada la similitud con el caso de los estimadores de la matriz de medias, se realizan las simulaciones exploratorias con las mismas características que las simulaciones de la sección anterior. La única diferencia radica, en que ya no es de interés calcular los estimadores, sino que únicamente interesan, las frecuencias relativas a medida que aumenta el tiempo, y por la relación que existe entre el número inicial de individuos del tipo 1 (N_1) y el tiempo (n) (ver Tabla 4.1), en consecuencia, a medida que aumenta el número de individuos iniciales simultáneamente.

Tomando las frecuencias relativas resultantes en las simulaciones se elaboró la Figura 4.4, en la cual aparecen los gráficos asociados al cambio de las frecuencias relativas del i -ésimo tipo de individuo ($\Delta_i(n; \mathbf{N}_1)$), con $i = 1, 2, 3$, a medida que n y N_1 crecen, para los modelos D, E y F. Las líneas punteadas corresponden al vector propio a izquierda de la matriz de medias \mathbf{m} , asociado al valor propio ρ , normalizado con la norma Manhattan¹. En la Tabla 3.6 se encuentran las matrices de medias \mathbf{m} y los valores de ρ , para cada uno de los tres modelos, y en la Tabla 4.2 se muestra el vector propio en mención, para cada modelo, normalizado para las normas Euclidiana y Manhattan.

Tabla 4.2: Vector propio a izquierda de la matriz \mathbf{m} , asociado a ρ , para los modelos D, E y F, normalizado con la norma euclidiana y la norma Manhattan.

Norma	Modelo D.	Modelo E.	Modelo F.
Euclidiana	(0.6337, 0.7434, 0.2140)	(0.5797, 0.7730, 0.2577)	(0.6156, 0.7540, 0.2291)
Manhattan	(0.3983, 0.4672, 0.1345)	(0.36, 0.48, 0.16)	(0.3851, 0.4716, 0.1433)

¹La norma Manhattan o del taxista se define como: $\|\mathbf{v}\|_1 := \sum_{i=1}^n |v_i|$

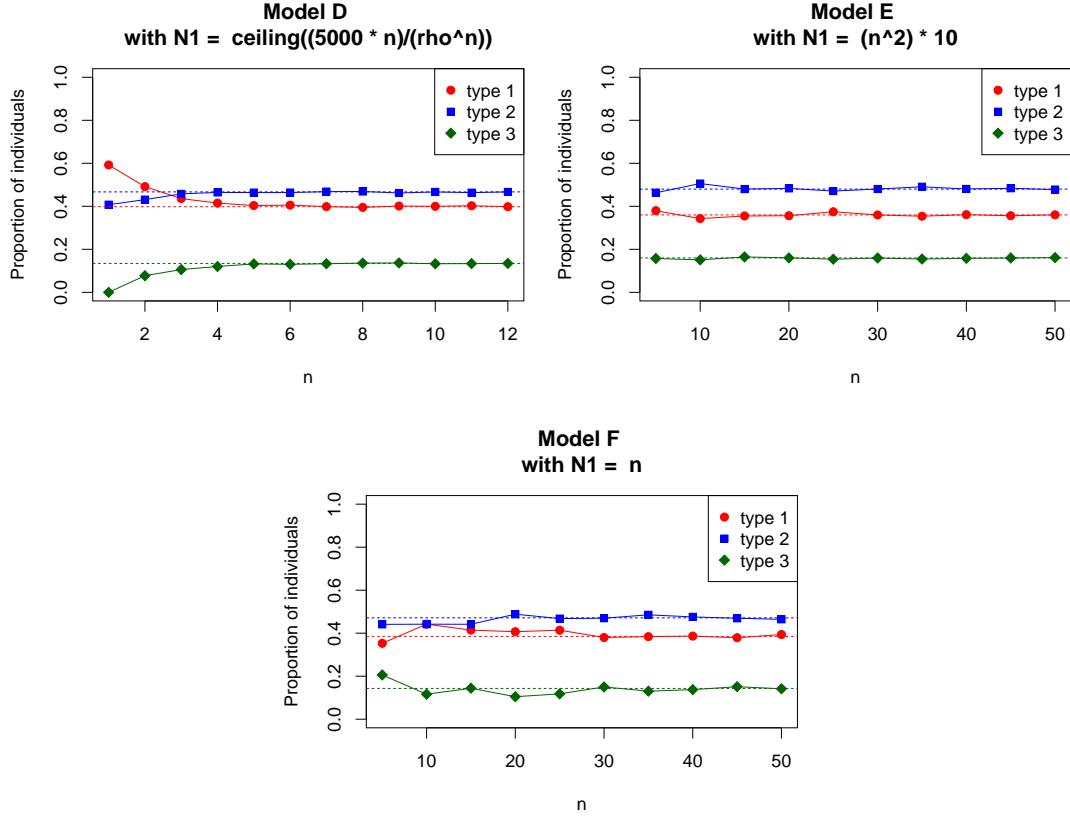


Figura 4.4: Convergencia de las frecuencias relativas cuando el tiempo y el número de individuos iniciales de tipo 1 crece, para los modelos D, E y F.

En la Figura 4.4 se muestra como, para los modelos D y E (caso subcrítico y crítico), se podría decir que se observa como las frecuencias relativas convergen. Para el modelo F (caso supercrítico), el resultado de la simulación coincide con el resultado teórico expuesto en el Teorema 8.3 que se encuentra en (Mode 1971, pág. 30).

A partir del resultado teórico para el caso supercrítico y buscando extenderlo, para los casos crítico y subcrítico, a partir de lo observado en las simulaciones, en este trabajo se propone el siguiente segundo nuevo resultado para ser demostrado o refutado.

Resultado 4.2. Sea $(\mathbf{Z}(n; \mathbf{N}_1))_{n \in \mathbb{N}_0}$ un proceso de ramificación de BGWM, con $d < \infty$ tipos distintos, regular positivo, con momentos de segundo orden finitos y con $\mathbf{N}_1 = (N_1, 0, \dots, 0)$, tal que,

$$\left(\rho > 1\right), \text{ o, } \left(\rho = 1 \text{ y } n/N_1 \rightarrow 0\right), \text{ o, } \left(\rho < 1 \text{ y } N_1 \rho^n \rightarrow \infty\right),$$

entonces, cuando n y N_1 tienden a infinito,

$$\Delta(n; \mathbf{N}_1) = \frac{\mathbf{Z}(n; \mathbf{N}_1)}{\|\mathbf{Z}(n; \mathbf{N}_1)\|_1} \xrightarrow{\text{c.s.}} \frac{\mathbf{u}}{\|\mathbf{u}\|_1} \quad (4.3)$$

en el conjunto de no extinción, $\left\{\sum_{k=1}^d Z_k(t; \mathbf{N}_1) > 0 : n \text{ y } N_1 \in \mathbb{N}_0\right\}$, y donde \mathbf{u} es el vector propio a izquierda correspondiente al mayor valor propio de la matriz de medias del proceso.

Notación

- δ_{ij} Delta de Kronecker.
- $1_A(\cdot)$ Función indicadora del conjunto A .
- $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ Conjunto de números naturales incluyendo el cero.
- n Variable de tiempo discreta $n \in \mathbb{N}_0$.
- t Variable de tiempo continua $t \in [0, \infty)$.
- d Número de tipos distintos.
- N Número inicial de individuos.
- $p(\cdot)$ Función de densidad discreta univariada (asociada a la distribución de descendencia).
- $p_r := p(r)$ Probabilidad de que un individuo tenga r descendientes.
- p_k^{*i} k -ésima componente de la i -ésima convolución de $p(\cdot)$.
- $F(s) := \sum_{k=0}^{\infty} p_k s^k$ Función generadora de probabilidades univariada asociada a $p(\cdot)$.
- $Z(n; N)$ Variable aleatoria definida como: “Número de individuos en la n -ésima generación a partir de N individuos iniciales”.
- $Z(n) := Z(n; 1)$ Variable aleatoria definida como: “Número de individuos en la n -ésima generación a partir de un individuo inicial”.
- $m := E(Z(1))$ media de la distribución de descendencia de los individuos.
- $\sigma^2 := Var(Z(1))$ varianza de la distribución de descendencia de los individuos.
- $\mathbb{N}_0^d := \mathbb{N} \cup \{0\} \times \mathbb{N} \cup \{0\} \times \dots \times \mathbb{N} \cup \{0\}$ Producto cartesiano de d conjuntos de números naturales incluyendo el cero.
- \mathbf{e}_k k -ésimo vector de la base canónica de \mathbb{N}_0^d .

- $\mathbf{N} = (N_1, N_2, \dots, N_d)$ Vector inicial de individuos de d tipos.
- $\mathbf{p}(\cdot, \dots, \cdot) := (p_{(1)}(\cdot, \dots, \cdot), p_{(2)}(\cdot, \dots, \cdot), \dots, p_{(d)}(\cdot, \dots, \cdot))$ Vector de funciones de densidad discretas multivariadas (asociada a las distribuciones de descendencia para los d tipos).
- $p_{(k)}(r_1, r_2, \dots, r_d)$ Probabilidad de que un individuo de tipo k produzca r_1 individuos de tipo 1, r_2 individuos de tipo 2, \dots , r_d individuos de tipo d .
- $\mathbf{F}(\mathbf{s}) := (F_1(\mathbf{s}), F_2(\mathbf{s}), \dots, F_d(\mathbf{s}))$ Vector de funciones generadoras de probabilidades multivariadas, asociado a $\mathbf{p}(\cdot, \dots, \cdot)$.
- $\mathbf{Z}(n; \mathbf{N}) = (Z_1(n; \mathbf{N}), Z_2(n; \mathbf{N}), \dots, Z_d(n; \mathbf{N}))$ Vector aleatorio definido como: “Número de individuos en la n -ésima generación, a partir de \mathbf{N} individuos iniciales, para los d tipos”.
- $Y_{(k,j)}(n; \mathbf{N})$ Variable aleatoria definida como: “Número de individuos de tipo j en la n -ésima generación, provenientes de individuos del tipo k de la $(n - 1)$ -ésima generación, a partir de \mathbf{N} individuos iniciales”.
- $Y_{(k,j)} := Y_{(k,j)}(1; \mathbf{e}_k)$ Variable aleatoria definida como: “Número de individuos de tipo j , provenientes de un solo individuo del tipo k ”.
- $U(n; \mathbf{N})$ Variable aleatoria definida como: “Número total de individuos de todos los tipos en la n -ésima generación, a partir de \mathbf{N} individuos iniciales”.
- $\Delta(n; \mathbf{N}) = (\Delta_1(n; \mathbf{N}), \Delta_2(n; \mathbf{N}), \dots, \Delta_d(n; \mathbf{N}))$ Vector aleatorio definido como: “Frecuencias relativas en la n -ésima generación, a partir de \mathbf{N} individuos iniciales, para los d tipos”.
- \mathbf{m}_k Vector de medias de la distribución de descendencia de los individuos de tipo k .
- \mathbf{V}_k Matriz de varianzas y covarianzas de la distribución de descendencia de los individuos de tipo k .
- $\mathbf{m} := (\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_d)'$ Matriz de medias de las distribuciones de descendencia de los individuos de los d tipos (el vector fila k -ésimo de ésta matriz corresponde a \mathbf{m}_k).
- ρ El mayor valor propio de la matriz \mathbf{m} .
- $\mathbf{m}(n; \mathbf{N})$ Vector de medias de la población en la n -ésima generación, a partir de \mathbf{N} individuos iniciales, para los d tipos.
- $M(n; \mathbf{N})$ Suma de los elementos del vector de medias $\mathbf{m}(n; \mathbf{N})$.
- $\mathbf{p}(n; \mathbf{N})$ Vector de frecuencias relativas o proporciones de la población en la n -ésima generación, a partir de \mathbf{N} individuos iniciales, para los d tipos.
- $\mathbf{V}(n; \mathbf{N})$ Matriz de varianzas y covarianzas de la población en la n -ésima generación, a partir de \mathbf{N} individuos iniciales, para los d tipos.
- $\hat{\mathbf{m}}_k(n; \mathbf{N})$ Estimador máximo verosímil del vector de medias de la distribución de descendencia de los individuos de tipo k , a partir de los valores observados de $Y_{(k,j)}(n; \mathbf{N})$, en n generaciones y de \mathbf{N} individuos iniciales.

-
- $\tilde{\mathbf{m}}_k(n; \mathbf{N})$ Estimador empírico del vector de medias de la distribución de descendencia de los individuos de tipo k , a partir de los valores observados de $Y_{(k,j)}(n; \mathbf{N})$, en n generaciones y de \mathbf{N} individuos iniciales.
 - $\hat{\mathbf{V}}_k(n; \mathbf{N})$ Estimador de la matriz de varianzas y covarianzas de la distribución de descendencia de los individuos de tipo k , a partir de los valores observados de $Y_{(k,j)}(n; \mathbf{N})$, en n generaciones y de \mathbf{N} individuos iniciales, calculado a partir del estimador $\hat{\mathbf{m}}_k(n; \mathbf{N})$.
 - $\check{\mathbf{V}}_k(n; \mathbf{N})$ Estimador de la matriz de varianzas y covarianzas de la distribución de descendencia de los individuos de tipo k , a partir de los valores observados de $Y_{(k,j)}(n; \mathbf{N})$, en n generaciones y de \mathbf{N} individuos iniciales, calculado a partir del estimador $\check{\mathbf{m}}_k(n; \mathbf{N})$.

Conceptos Previos

Teorema B.1 (Teorema de convergencia dominada para esperanza condicional). *Sea un espacio de probabilidad (Ω, \mathcal{F}, P) , una sucesión de variables aleatorias (v.a) $(X_n)_{n \geq 1}$ sobre dicho espacio y \mathcal{G} una sub- σ -álgebra de \mathcal{F} . Si Y es una variable aleatoria sobre el espacio de probabilidad, \mathcal{G} -medible con $E(|Y|) < \infty$, y si se satisface $|X_n| \leq Y$ y X_n converge casi siempre a X ($X_n \xrightarrow{c.s.} X$), entonces,*

$$\lim_{n \rightarrow \infty} E(X_n | \mathcal{G}) = E(X | \mathcal{G}) = E \left(\lim_{n \rightarrow \infty} X_n \mid \mathcal{G} \right)$$

Teorema B.2 (Ley fuerte de los grandes números). *Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias (v.a.) independientes idénticamente distribuidas (i.i.d.), definidas sobre el mismo espacio de probabilidad con $\mu = E(X_1)$ y $\sigma^2 = Var(X_1) < \infty$. Si $S_n = \sum_{i=1}^n X_i$ entonces, cuando $n \rightarrow \infty$,*

$$\frac{S_n}{n} \rightarrow \mu, \quad \text{tanto casi siempre, como en } L^2$$

Teorema B.3 (Teorema central del límite para vectores aleatorios). *Sea $(\mathbf{X}_n)_{n \geq 1}$ una sucesión de vectores aleatorios (v.a.), k dimensionales independientes idénticamente distribuidos (i.i.d.), con vector de medias μ y, matriz de varianzas y covarianzas Σ . Supóngase $\bar{\mathbf{X}}_n = \sum_{i=1}^n \frac{\mathbf{X}_i}{n}$ el vector de medias aritméticas, entonces, cuando $n \rightarrow \infty$,*

$$\sqrt{n} (\bar{\mathbf{X}}_n - \mu) \xrightarrow{dist} \mathbf{Z}$$

donde \mathbf{Z} tiene una distribución normal multivariada de dimensión k ($\mathbf{Z} \sim \mathcal{N}_k(\mathbf{0}, \Sigma)$).

Definición B.1. Proceso estocástico Una familia $(X(t))_{t \in T}$ de v.a. definidas sobre un espacio de probabilidad (Ω, \mathcal{F}, P) y con valores en un espacio medible (S, \mathcal{A}) se llama proceso estocástico. Si el conjunto T es discreto, el proceso estocástico se denomina proceso de parámetro discreto, y si T es continuo, el proceso se denomina proceso de parámetro continuo.

Definición B.2. Cadena de Markov con parámetro de tiempo discreto Una sucesión aleatoria $(X(n))_{n \in N_0}$ con un conjunto de estados discreto, se llama cadena de Markov con parámetro de tiempo discreto, si y sólo si, se satisface la siguiente condición:

$$P(X(n+1) = j | X(n) = i, X(n-1) = i_{n-1}, \dots, X(0) = i_0) = P(X(n+1) = j | X(n) = i)$$

para todo $n \in N_0$ y para todo $i_0, i_1, \dots, i_{n-1}, i, j \in S$, siempre y cuando las probabilidades condicionales estén definidas.

Teorema B.4 (Teorema central del límite para martingalas). Sea $(X(n))_{n \geq 1}$ una sucesión de variables aleatorias, con $\mathcal{F}(n)$ la filtración generada por $X(1), \dots, X(n)$, tales que:

- $E(X(n) | \mathcal{F}(n-1)) = 0$
- $E(X^2(n) | \mathcal{F}(n-1)) = 1$
- $E(|X(n)|^3 | \mathcal{F}(n-1)) \leq k < \infty$

y sea,

$$S(n) := \sum_{i=1}^n X(i),$$
$$S(0) = 0$$

entonces, cuando $n \rightarrow \infty$,

$$\frac{1}{\sqrt{n}} S(n) \xrightarrow{\text{dist}} Z$$

donde Z es una variable aleatoria con distribución normal estándar.

Paquete *Branching*

Este apéndice tiene como objetivo describir los pasos a seguir para el uso de la herramientas informáticas desarrolladas en este trabajo, esto incluye: las instrucciones para instalar R, las instrucciones para instalar el paquete *Branching* y la ayuda para el uso de cada una de las funciones que componen dicho paquete.

R es una implementación gratuita del programa de computación estadística S, algunas de sus características son:

- En general se trabaja por medio de línea de comandos.
- Incluye un limitado uso de GUI's (Interfaces Gráficas), sin embargo existen varias propuestas, siendo *Rcmdr* (R Commander) el más usado.
- Tiene excelentes capacidades gráficas.
- Se distribuye bajo la licencia GNU GPL y esta disponible para los sistemas operativos: Unix/Linux, MacOS y Windows.
- Tiene muy buena documentación y ayuda en línea.

y se pueden destacar las siguientes ventajas de su uso:

- Los desarrollos más recientes en el campo de la estadística generalmente están implementados en R antes que en otra herramienta informática.
- Existen muchas librerías disponibles para aplicar diversos métodos estadísticos.
- Crea excelente gráficas con relativa facilidad.
- Puede leer datos de diferente sistemas de bases de datos (SQL, Oracle, etc).
- Su uso no implica algún costo de licenciamiento o de otro tipo.

C.1. Instalación de *R*

Al ingresar a la página oficial de *R* (<http://cran.r-project.org/>) lo primero que aparece es un recuadro en donde dice “Download and Install *R*”, allí se encuentran tres vínculos, uno por cada sistema operativo soportado, cada uno de ellos lleva a una página en donde se encuentran los instaladores de la última versión disponible.

Aunque este trabajo se ha realizado en su totalidad en un equipo con sistema operativo Ubuntu - Linux, se describirá la instalación para un equipo con sistema operativo Windows, por ser el sistema operativo más usado en el mundo. En el momento en el que se escribió esta sección, del siguiente vínculo:

`http://cran.r-project.org/bin/windows/base/R-2.10.1-win32.exe`

se descargaba el instalador para Windows en su versión 2.10.1, dicho instalador corresponde a la distribución en archivo binario del software *R* en su versión 2.10.1, destinado a ser instalado en un equipo con procesadores *ix86* o *x86_64*, y con sistema operativo Windows 2000 o superior, incluyendo las versiones de 64 bits.

Al hacer doble clic al instalador, se pueden seguir los siguientes pasos:

1. Aparece un cuadro de dialogo en donde se debe seleccionar el idioma que se usará en la instalación. Se debe seleccionar el idioma deseado y hacer clic en **OK**.
2. En seguida aparece, el cuadro usual de instalación de programas de Windows, solicitando seleccionar entre las opciones **Siguiente**, para continuar, o **Cancelar** para salir de la instalación. Se hace clic en **Siguiente**.
3. Aparece la licencia de uso GNU GENERAL PUBLIC LICENSE. Version 2, June 1991. Se hace clic en **Siguiente**.
4. Aparece un cuadro de texto con la ruta predeterminada en donde quedará instalado el programa. Se cambia la ruta si se desea y se hace clic en **Siguiente**.
5. Aparece un cuadro de lista y una serie de cuadros de verificación que permiten seleccionar los componentes que se desean instalar. Se sugiere seleccionar en el cuadro de lista la opción **Instalación completa** y hacer clic en **Siguiente**.
6. Aparece un cuadro de selección que indica si se desea utilizar las opciones de configuración. Se puede dejar en No y hacer clic en **Siguiente**.
7. Si se desea se puede cambiar la configuración de los accesos directos. Se hace clic en **Siguiente**.
8. Si se desea se pueden cambiar algunas configuraciones adicionales. Se hace clic en **Siguiente**.
9. Por último, se hace clic en **Finalizar**.

C.2. Instalación del paquete *Branching*

En la dirección <http://cran.r-project.org/web/packages/> se encuentra todo lo relacionado con los paquetes de R, allí se encuentran los manuales que describen la instalación de paquetes existentes y la construcción de nuevos paquetes, adicionalmente, aparece el listado de paquetes disponibles para su uso.

Para la instalación del paquete *Branching*, se tienen dos opciones:

■ Descarga e instalación utilizando R:

1. Después de abrir R, en el menú que se encuentra en la parte superior de la ventana, seleccionar **Paquetes** y seleccionar **Instalar paquete(s)...**
2. Aparece una ventana emergente en donde se debe seleccionar el lugar desde el cual se desea descargar, de manera automática, el paquete. Seleccione **Colombia** y haga clic en **OK**.
3. Si fue posible realizar la conexión con el servidor asociado al lugar seleccionado¹, aparece una nueva ventana emergente con el listado de todos los paquetes que se pueden descargar e instalar. Seleccione **Branching** y haga clic en **OK**.
4. Por último, sale una barra de progreso indicando el avance de la descarga del paquete y luego una serie de mensajes que informan acerca del proceso de descarga e instalación.

■ Descarga independiente de la instalación:

1. En un navegador ingresar a la página:
<http://cran.r-project.org/web/packages/>
2. En el listado de paquetes disponibles, buscar a la izquierda el nombre del paquete (**Branching**) y hacer clic en él.
3. En la página que aparece, descargar el archivo zip asociado al vínculo que se encuentra en la parte de **Downloads**, en frente de **Windows binary**.
4. Después de abrir R, en el menú que se encuentra en la parte superior de la ventana, seleccionar **Paquetes** y seleccionar **Instalar paquete(s) a partir de archivos zip locales...**
5. Aparece la ventana usual de selección de archivos. Ubicar el archivo zip del paquete *Branching* en la ruta en donde se descargó, luego seleccionarlo y dar clic en **abrir**.
6. Por último, solo aparece un mensaje de verificación de integridad del archivo.

Para verificar que el paquete *Branching* a quedado correctamente instalado o para el uso de las funciones que lo componen, es necesario escribir en la línea de comandos (una sola vez por sesión de trabajo):

```
library(Branching)
```

¹Una falla en la conexión puede ser causada por la configuración del acceso a internet, como lo es el uso de proxy. En este caso se sugiere realizar la descarga independiente de la instalación o leer la documentación detallada.

En cuanto a la instalación de paquetes en general, se puede encontrar información detallada de las siguientes formas:

- Después de abrir R, escribir en línea de comandos `help("INSTALL")`
- Después de abrir R, escribir en línea de comandos `help("install.packages")`
- Se puede encontrar un manual de instalación y administración de paquetes, en la dirección: <http://cran.r-project.org/doc/manuals/R-admin.pdf>

C.3. Documentación del paquete *Branching*

rBGWM

Generar un proceso de Bienaymé - Galton - Watson multitypo

Descripción

Realiza una simulación de un proceso de Bienaymé - Galton - Watson multitypo a partir de sus distribuciones de descendencia, dando como resultado tres posibles tipos de muestras asociadas al proceso.

Uso

```
rBGWM(dists, type=c("general", "multinomial", "independents"),
      d, n, z0=rep(1,d), c.s=TRUE, tt.s=TRUE, rf.s=TRUE,
      file=NULL)
```

Argumentos

<code>dists</code>	distribuciones de descendencia para cada tipo de individuo. Su estructura depende del tipo de distribuciones.
<code>type</code>	tipo de distribuciones de descendencia.
<code>d</code>	número entero positivo; cantidad de tipos de individuos distintos.
<code>n</code>	número entero positivo; tiempo máximo o longitud de la trayectoria del proceso.
<code>z0</code>	vector numérico, con valores enteros no negativos, de tamaño <code>d</code> ; vector de individuos iniciales por tipo.
<code>c.s</code>	valor lógico; si es <code>TRUE</code> , el objeto de salida incluirá los valores simulados del número de individuos por combinación, tipo padre - tipo hijo, asociados al proceso.

<code>tt.s</code>	valor lógico; si es TRUE, el objeto de salida incluirá los valores simulados del número total de individuos por tipo asociados al proceso.
<code>rf.s</code>	valor lógico; si es TRUE, el objeto de salida incluirá los valores simulados de las frecuencias relativas por tipo asociadas al proceso.
<code>file</code>	cadena de caracteres; archivo de salida en donde se irán almacenando los valores simulados del número de individuos por combinación tipo padre - tipo hijo asociados al proceso.

Detalles

Esta función realiza una simulación de un proceso de Bienaymé - Galton - Watson multitypo (BGWM) a partir de las distribuciones de descendencia asociadas al mismo. Teniendo en cuenta algunas distribuciones de descendencia particulares, se propone una tipificación de dichos procesos de la siguiente manera:

`general` Corresponde a todos los procesos de BGWM sin condición alguna sobre las distribuciones de descendencia, para este caso, se requieren como datos de entrada para cada distribución, los vectores d dimensionales con su respectiva probabilidad.

`multinomial` Corresponde a todos los procesos de BGWM en donde cada distribución de descendencia es una distribución multinomial con número de repeticiones aleatorio, para este caso, se requieren como datos de entrada d distribuciones univariadas que modelan el número de repeticiones aleatorio de cada distribución de descendencia y una matriz $d \times d$ en donde cada fila contiene los d parámetros de cada distribución de descendencia multinomial.

`independents` Corresponde a todos los procesos de BGWM en donde cada distribución de descendencia es la distribución conjunta de d variables aleatorias discretas independientes, una por cada tipo de individuo, para este caso, se requieren como datos de entrada d^2 distribuciones univariadas.

La estructura en que se maneja cada uno de los tipos de distribuciones de descendencia se ilustra en los ejemplos.

Las distribuciones univariadas disponibles son:

`unif` Distribución Uniforme Discreta, parámetros min y max . Los valores enteros entre min y max tienen la misma probabilidad de ocurrir.

`binom` Distribución Binomial, parámetros n y p .

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}$$

para $x = 0, \dots, n$.

`hyper` Distribución Hipergeométrica, parámetros m (número de balotas con la característica de interés), n (número de balotas sin la característica), k (número de balotas extraídas).

$$p(x) = \binom{m}{x} \binom{n}{k-x} / \binom{m+n}{k}$$

para $x = 0, \dots, k$.

geom Distribución Geométrica, parámetro p .

$$p(x) = p(1 - p)^x$$

para $x = 0, 1, 2, \dots$

nbinom Distribución Binomial Negativa, parámetros n y p .

$$p(x) = \frac{\Gamma(x + n)}{\Gamma(n)x!} p^n (1 - p)^x$$

para $x = 0, 1, 2, \dots$

pois Distribución Poisson, parámetro λ .

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

para $x = 0, 1, 2, \dots$

norm Distribución Normal redondeada a valores enteros y los valores negativos asociados al valor 0, parámetros μ y σ .

$$p(x) = \int_{x-0.5}^{x+0.5} \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-\mu)^2/2\sigma^2} dt$$

para $x = 1, 2, \dots$

$$p(x) = \int_{-\infty}^{0.5} \frac{1}{\sqrt{2\pi}\sigma} e^{-(t-\mu)^2/2\sigma^2} dt$$

para $x = 0$

lnorm Distribución Log-normal redondeada a valores enteros, parámetros $\log\text{mean} = \mu$ y $\log\text{sd} = \sigma$.

$$p(x) = \int_{x-0.5}^{x+0.5} \frac{1}{\sqrt{2\pi}\sigma t} e^{-(\log(t)-\mu)^2/2\sigma^2} dt$$

para $x = 1, 2, \dots$

$$p(x) = \int_0^{0.5} \frac{1}{\sqrt{2\pi}\sigma t} e^{-(\log(t)-\mu)^2/2\sigma^2} dt$$

para $x = 0$

gamma Distribución Gamma redondeada a valores enteros, parámetros $\text{shape} = \alpha$ y $\text{scale} = \sigma$.

$$p(x) = \int_{x-0.5}^{x+0.5} \frac{1}{\sigma^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-t/\sigma} dt$$

para $x = 1, 2, \dots$

$$p(x) = \int_0^{0.5} \frac{1}{\sigma^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{-t/\sigma} dt$$

para $x = 0$

Valor

Un objeto lista `list` con lo siguiente:

<code>i.d</code>	cantidad de tipos de individuos distintos.
<code>i.dists</code>	distribuciones de descendencia.
<code>i.n</code>	número entero positivo; tiempo máximo o longitud de la trayectoria del proceso.
<code>i.z0</code>	vector numérico, con valores enteros no negativos, de tamaño <code>d</code> ; vector de individuos iniciales por tipo.
<code>o.c.s</code>	valores simulados del número de individuos por combinación tipo padre - tipo hijo asociados al proceso.
<code>o.tt.s</code>	valores simulados del número total de individuos por tipo asociados al proceso.
<code>o.rf.s</code>	valores simulados de las frecuencias relativas por tipo asociadas al proceso.

Autor(es)

Camilo José Torres Jiménez <cjtorresj@unal.edu.co>

Ejemplos

```
## Simulacion a partir de un modelo presentado en Stefanescu(1998)

# parametros
d <- 2
n <- 15
N <- c(90, 10)
a <- c(0.2, 0.3); caso <- "subcritico"

# con distribuciones independientes
Dists.i <- data.frame( name=rep( "pois", d*d ),
                      param1=rep( a, rep(d,d) ),
                      stringsAsFactors=FALSE )

rA <- rBGWM(Dists.i, "independents", d, n, N)

# con distribucion multinomial
dist <- data.frame( name=rep( "pois", d ),
                  param1=a*d,
                  stringsAsFactors=FALSE )
matrix.b <- matrix( rep(0.5, 4), nrow=2 )
Dists.m <- list( dists.eta=dist, matrix.B=matrix.b )

rB <- rBGWM(Dists.m, "multinomial", d, n, N)

# con distribucion general (aproximacion)
max <- 30
A <- t(expand.grid(c(0:max),c(0:max)))
aux1 <- factorial(A)
```

```

aux1 <- apply(aux1,2,prod)
aux2 <- apply(A,2,sum)
distp <- function(x,y,z){ exp(-d*x)*(x^y)/z }
p <- sapply( a, distp, aux2, aux1 )
prob <- list( dist1=p[,1], dist2=p[,2] )
size <- list( dist1=ncol(A), dist2=ncol(A) )
vect <- list( dist1=t(A), dist2=t(A) )
Dists.g <- list( sizes=size, probs=prob, vects=vect )

rC <- rBGWM(Dists.g, "general", d, n, N)

# comparacion grafica
dev.new()
plot.ts(rA$o.tt.s,main="with independents")
dev.new()
plot.ts(rB$o.tt.s,main="with multinomial")
dev.new()
plot.ts(rC$o.tt.s,main="with general (aprox)")

```

BGWM.mean

Medias de un proceso de Bienaymé - Galton - Watson multitypo

Descripción

Calcula la matriz de medias de un proceso de Bienaymé - Galton - Watson multitypo a partir de sus distribuciones de descendencia, adicionalmente, es posible calcular la matriz de medias del número de individuos por tipo, en el tiempo n y el vector de medias respectivo, si se indica el número de individuos iniciales por tipo.

Uso

```

BGWM.mean(dists,
           type=c("general", "multinomial", "independents"),
           d, n=1, z0=NULL, maxiter=1e5)

```

Argumentos

<code>dists</code>	distribuciones de descendencia para cada tipo de individuo. Su estructura depende del tipo de distribuciones.
<code>type</code>	tipo de distribuciones de descendencia.
<code>d</code>	número entero positivo; cantidad de tipos de individuos distintos.
<code>n</code>	número entero positivo; tiempo máximo o longitud de la trayectoria del proceso.
<code>z0</code>	vector numérico con valores enteros no negativos de tamaño d ; vector de individuos iniciales por tipo.

`maxiter` número entero positivo; tamaño de la muestra con la cual se estiman los parámetros de las distribuciones univariadas que no tengan definida una fórmula analítica para su cálculo directo.

Detalles

Esta función realiza el cálculo de la matriz de medias de un proceso de Bienaymé - Galton - Watson multitypo (BGWM) a partir de las distribuciones de descendencia asociadas al mismo.

Teniendo en cuenta algunas distribuciones de descendencia particulares, se propone una tipificación de dichos procesos de la siguiente manera:

`independents`, `general` y `multinomial`

La estructura en que se maneja cada uno de los tipos de distribuciones de descendencia se ilustra en los ejemplos.

Las distribuciones univariadas disponibles son: `unif`, `binom`, `hyper`, `geom`, `nbinom`, `pois`, `norm`, `lnorm` y `gamma`.

Al utilizar `norm`, `lnorm` y `gamma`, la media de la distribución se estima calculando la media aritmética de un total de `maxiter` números aleatorios generados de la respectiva distribución.

En `rBGWM` se encuentra la descripción de las tipificaciones de los procesos y las funciones de densidad utilizadas para cada distribución univariada mencionada.

Valor

Un objeto matriz `matrix` correspondiente a la matriz de medias del proceso en la n -ésima generación, o, un vector `vector` en el caso en que se ha indicado el número de individuos iniciales por tipo (z_0), correspondiente al vector de medias del proceso en la n -ésima generación que parte de z_0 individuos iniciales.

Autor(es)

Camilo José Torres Jiménez <cjtorresj@unal.edu.co>

Ejemplos

```
## Calculo de medias a partir de un modelo presentado en Stefanescu(1998)

# parametros
d <- 2
n <- 15
N <- c(90, 10)
a <- c(0.2, 0.3)
caso <- "subcritico"

# con distribuciones independientes
Dists.i <- data.frame( name=rep( "pois", d*d ),
                      param1=rep( a, rep(d,d) ),
                      stringsAsFactors=FALSE )
```

```

# matriz de medias del proceso
I.matrix.m <- BGWM.mean(Dists.i, "independents", d)

# vector de medias del numero de individuos por tipo, en la generaci3n
# n y partiendo del vector N de ancestros
I.vector.m.n_N <- BGWM.mean(Dists.i, "independents", d, n, N)

# con distribuci3n multinomial
dist <- data.frame( name=rep( "pois", d ),
                    param1=a*d,
                    stringsAsFactors=FALSE )
matrix.b <- matrix( rep(0.5, 4), nrow=2 )
Dists.m <- list( dists.eta=dist, matrix.B=matrix.b )

# matriz de medias del proceso
M.matrix.m <- BGWM.mean(Dists.m, "multinomial", d)

# vector de medias del numero de individuos por tipo, en la generaci3n
# n y partiendo del vector N de ancestros
M.vector.m.n_N <- BGWM.mean(Dists.m, "multinomial", d, n, N)

# con distribuci3n general (aproximacion)
max <- 30
A <- t(expand.grid(c(0:max),c(0:max)))
aux1 <- factorial(A)
aux1 <- apply(aux1,2,prod)
aux2 <- apply(A,2,sum)
distp <- function(x,y,z){ exp(-d*x)*(x^y)/z }
p <- sapply( a, distp, aux2, aux1 )
prob <- list( dist1=p[,1], dist2=p[,2] )
size <- list( dist1=ncol(A), dist2=ncol(A) )
vect <- list( dist1=t(A), dist2=t(A) )
Dists.g <- list( sizes=size, probs=prob, vects=vect )

# matriz de medias del proceso
G.matrix.m <- BGWM.mean(Dists.g, "general", d)

# vector de medias del numero de individuos por tipo, en la generaci3n
# n y partiendo del vector N de ancestros
G.vector.m.n_N <- BGWM.mean(Dists.g, "general", d, n, N)

# comparacion de los resultados
I.vector.m.n_N
I.vector.m.n_N - M.vector.m.n_N
M.vector.m.n_N - G.vector.m.n_N
G.vector.m.n_N - I.vector.m.n_N

```

Descripción

Realiza una estimación de la matriz de medias a partir de valores observados en un experimento que puede modelarse mediante un proceso de Bienaymé - Galton - Watson.

Uso

```
BGWM.mean.estim(sample, method=c("EE", "MLE"), d, n, z0)
```

Argumentos

<code>sample</code>	matriz numérica de enteros no negativos, con <code>d</code> columnas; datos observados a partir de los cuales se realiza la estimación.
<code>method</code>	método de estimación (EE Estimación empírica, MLE Estimación máximo verosímil).
<code>d</code>	número entero positivo; cantidad de tipos de individuos distintos.
<code>n</code>	número entero positivo; cantidad de observaciones realizadas a través del tiempo.
<code>z0</code>	vector numérico con valores enteros no negativos de tamaño <code>d</code> ; vector de individuos iniciales por tipo.

Detalles

Esta función realiza una estimación de la matriz de medias utilizando dos posibles estimadores, el primero corresponde al estimador empírico y el segundo al estimador máximo verosímil, en ambos casos dichos estimadores requieren la denominada muestra completa asociada al proceso, es decir, se requiere observar el número de individuos por cada combinación tipo padre - tipo hijo asociado al proceso en cada momento en el tiempo, dicha muestra debe ir en `sample`. Para más detalles ver Torres(2010) o Maaouia & Touati(2005).

Valor

Un objeto lista `list` con lo siguiente:

<code>method</code>	método de estimación utilizado.
<code>m</code>	Un objeto matriz <code>matrix</code> <code>d x d</code> correspondiente a una estimación de la matriz de medias del proceso.

Autor(es)

Camilo José Torres Jiménez <cjtorresj@unal.edu.co>

Ejemplos

```
## Calculo de la matriz de medias estimada a partir de datos simulados

# parametros
d <- 3
n <- 30
N <- c(10,10,10)
LeslieMatrix <- matrix( c(0.08, 1.06, 0.07,
                        0.99, 0, 0,
                        0, 0.98, 0), 3, 3 )

# distribuciones independientes a partir de la matriz de Leslie
Dists.pois <- data.frame( name=rep( "pois", d ),
                        param1=LeslieMatrix[,1],
                        param2=NA,
                        stringsAsFactors=FALSE )
Dists.binom <- data.frame( name=rep( "binom", 2*d ),
                        param1=rep( 1, 2*d ),
                        param2=c(t(LeslieMatrix[,-1])),
                        stringsAsFactors=FALSE )
Dists.i <- rbind(Dists.pois,Dists.binom)
Dists.i <- Dists.i[c(1,4,5,2,6,7,3,8,9),]
Dists.i

# matriz de medias del proceso a simular
m <- BGWM.mean(Dists.i,"independents",d)

# datos simulados
simulated.data <- rBGWM(Dists.i, "independents", d, n, N,
                        TRUE, FALSE, FALSE)$o.c.s

# matriz de medias estimada por medio del estimador empírico
m.EE <- BGWM.mean.estim( simulated.data, "EE", d, n, N )$m

# matriz de medias estimada por medio del estimador máximo verosímil
m.MLE <- BGWM.mean.estim( simulated.data, "MLE", d, n, N )$m

# comparacion de matriz de medias del proceso contra estimaciones
m
m - m.EE
m - m.MLE
```

BGWM.covar

Varianzas y covarianzas de un proceso de Bienaymé - Galton - Watson multitipo

Descripción

Calcula las matrices de varianzas y covarianzas de un procesos de Bienaymé - Galton - Watson multitipo a partir de sus distribuciones de descendencia, adicionalmente, es posible calcular

las matrices de varianzas y covarianzas del número de individuos por tipo, en el tiempo n y la matriz de varianzas y covarianzas respectiva, si se indica el número de individuos iniciales por tipo.

Uso

```

BGWM.covar(dists,
            type=c("general", "multinomial", "independents"),
            d, n=1, z0=NULL, maxiter=1e5)

```

Argumentos

<code>dists</code>	distribuciones de descendencia para cada tipo de individuo. Su estructura depende del tipo de distribuciones.
<code>type</code>	tipo de distribuciones de descendencia.
<code>d</code>	número entero positivo; cantidad de tipos de individuos distintos.
<code>n</code>	número entero positivo; tiempo máximo o longitud de la trayectoria del proceso.
<code>z0</code>	vector numérico con valores enteros no negativos de tamaño <code>d</code> ; vector de individuos iniciales por tipo.
<code>maxiter</code>	número entero positivo; tamaño de la muestra con la cual se estiman los parámetros de las distribuciones univariadas que no tengan definida una fórmula analítica para su cálculo directo.

Detalles

Esta función realiza el cálculo de las matrices de varianzas y covarianzas de un proceso de Bienaymé - Galton - Watson multitypo (BGWM) a partir de las distribuciones de descendencia asociadas al mismo.

Teniendo en cuenta algunas distribuciones de descendencia particulares, se propone una tipificación de dichos procesos de la siguiente manera:

`independents`, `general` y `multinomial`.

La estructura en que se maneja cada uno de los tipos de distribuciones de descendencia se ilustra en los ejemplos.

Las distribuciones univariadas disponibles son: `unif`, `binom`, `hyper`, `geom`, `nbinom`, `pois`, `norm`, `lnorm` y `gamma`.

Al utilizar `norm`, `lnorm` y `gamma`, la media y la varianza de la distribución se estima calculando la media aritmética y la varianza muestral, respectivamente, de un total de `maxiter` números aleatorios generados de la respectiva distribución.

En `rBGWM` se encuentra la descripción de las tipificaciones de los procesos y las funciones de densidad utilizadas para cada distribución univariada mencionada.

Valor

Un objeto matriz $matrix(d^2 \times d)$ correspondiente a las d matrices $d \times d$ de varianzas y covarianzas del proceso en la n -ésima generación, o, una matriz $matrix(d \times d)$ en el caso en que se ha indicado el número de individuos iniciales por tipo (z_0), correspondiente a la matriz de varianzas y covarianzas del proceso en la n -ésima generación que parte de z_0 individuos iniciales.

Autor(es)

Camilo José Torres Jiménez <cjtorresj@unal.edu.co>

Ejemplos

```
## Calculo de varianzas y covarianzas a partir de un modelo presentado en
## Stefanescu(1998)

# parametros
d <- 2
n <- 15
N <- c(90, 10)
a <- c(0.2, 0.3); caso <- "subcritico"

# con distribuciones independientes
Dists.i <- data.frame( name=rep( "pois", d*d ),
                      param1=rep( a, rep(d,d) ),
                      stringsAsFactors=FALSE )

# matrices de varianzas y covarianzas del proceso
I.matriz.V <- BGWM.covar(Dists.i, "independents", d)

# matriz de varianzas y covarianzas del numero de individuos por tipo,
# en la generación n y partiendo del vector N de ancestros
I.matrix.V.n_N <- BGWM.covar(Dists.i, "independents", d, n, N)

# con distribución multinomial
dist <- data.frame( name=rep( "pois", d ),
                   param1=a*d,
                   stringsAsFactors=FALSE )
matrix.b <- matrix( rep(0.5, 4), nrow=2 )
Dists.m <- list( dists.eta=dist, matrix.B=matrix.b )

# matrices de varianzas y covarianzas del proceso
M.matriz.V <- BGWM.covar(Dists.m, "multinomial", d)

# matriz de varianzas y covarianzas del numero de individuos por tipo,
# en la generación n y partiendo del vector N de ancestros
M.matrix.V.n_N <- BGWM.covar(Dists.m, "multinomial", d, n, N)

# con distribución general (aproximacion)
max <- 30
A <- t(expand.grid(c(0:max),c(0:max)))
aux1 <- factorial(A)
aux1 <- apply(aux1,2,prod)
```

```

aux2 <- apply(A,2,sum)
distp <- function(x,y,z){ exp(-d*x)*(x^y)/z }
p <- sapply( a, distp, aux2, aux1 )
prob <- list( dist1=p[,1], dist2=p[,2] )
size <- list( dist1=ncol(A), dist2=ncol(A) )
vect <- list( dist1=t(A), dist2=t(A) )
Dists.g <- list( sizes=size, probs=prob, vects=vect )

# matriz de varianzas y covarianzas del proceso
G.matrix.V <- BGWM.covar(Dists.g, "general", d)

# matriz de varianzas y covarianzas del numero de individuos por tipo,
# en la generación n y partiendo del vector N de ancestros
G.matrix.V.n_N <- BGWM.covar(Dists.g, "general", d, n, N)

# comparacion de los resultados
I.matrix.V.n_N
I.matrix.V.n_N - M.matrix.V.n_N
M.matrix.V.n_N - G.matrix.V.n_N
G.matrix.V.n_N - I.matrix.V.n_N

```

BGWM.covar.estim *Estimación de las varianzas y covarianzas para datos provenientes de un proceso de Bienaymé - Galton - Watson multitypo*

Descripción

Realiza una estimación de las matrices de varianzas y covarianzas a partir de valores observados en un experimento que puede modelarse mediante un proceso de Bienaymé - Galton - Watson.

Uso

```
BGWM.covar.estim(sample, method=c("EE-m", "MLE-m"), d, n, z0)
```

Argumentos

sample	matriz numérica de enteros no negativos, con d columnas; datos observados a partir de los cuales se realiza la estimación.
method	método de estimación (m-EE Con la estimación empírica de la matriz de medias, m-MLE Con la estimación máximo verosímil de la matriz de medias).
d	número entero positivo; cantidad de tipos de individuos distintos.
n	número entero positivo; cantidad de observaciones realizadas a través del tiempo.
z0	vector numérico con valores enteros no negativos de tamaño d; vector de individuos iniciales por tipo.

Detalles

Esta función realiza una estimación de las matrices de varianzas y covarianzas utilizando los resultados asintóticos de dos estimadores de la matriz de medias, el primero corresponde al estimador empírico y el segundo al estimador máximo verosímil, en ambos casos dichos estimadores requieren la denominada muestra completa asociada al proceso, es decir, se requiere observar el número de individuos por cada combinación tipo padre - tipo hijo asociado al proceso en cada momento en el tiempo, dicha muestra debe ir en `sample`. Para más detalles ver Torres(2010) o Maaouia & Touati(2005).

Valor

Un objeto lista `list` con lo siguiente:

<code>method</code>	método de estimación utilizado
<code>V</code>	Un objeto matriz <code>matrix</code> ($d^2 \times d$) correspondiente a una estimación de las <code>d</code> matrices de varianzas y covarianzas del proceso.

Autor(es)

Camilo José Torres Jiménez <cjtorresj@unal.edu.co>

Ejemplos

```
## Calculo de las matrices de varianzas y covarianzas estimadas a partir
## de datos simulados

# parametros
d <- 3
n <- 30
N <- c(10,10,10)
LeslieMatrix <- matrix( c(0.08, 1.06, 0.07,
                        0.99, 0, 0,
                        0, 0.98, 0), 3, 3 )

# distribuciones independientes a partir de la matriz de Leslie
Dists.pois <- data.frame( name=rep( "pois", d ),
                        param1=LeslieMatrix[,1],
                        param2=NA,
                        stringsAsFactors=FALSE )
Dists.binom <- data.frame( name=rep( "binom", 2*d ),
                        param1=rep( 1, 2*d ),
                        param2=c(t(LeslieMatrix[,-1])),
                        stringsAsFactors=FALSE )
Dists.i <- rbind(Dists.pois,Dists.binom)
Dists.i <- Dists.i[c(1,4,5,2,6,7,3,8,9),]
Dists.i

# matriz de varianzas y covarianzas del proceso a simular
V <- BGWM.covar(Dists.i,"independents",d)
```

```
# datos simulados
simulated.data <- rBGWM(Dists.i, "independents", d, n, N,
                        TRUE, FALSE, FALSE)$o.c.s

# matriz de medias estimada por medio del estimador empírico
V.EE <- BGWM.covar.estim( simulated.data, "EE-m", d, n, N )$V

# matriz de medias estimada por medio del estimador máximo verosímil
V.MLE <- BGWM.covar.estim( simulated.data, "MLE-m", d, n, N )$V

# comparacion de matriz de medias del proceso contra estimaciones
V
V - V.EE
V - V.MLE
```

Conclusiones

Se destacan las siguientes conclusiones:

- La revisión bibliográfica permite entender el funcionamiento de los procesos de Bienaymé - Galton - Watson multitypo (BGWM), así como la definición tanto de estimadores, como de frecuencias relativas y los resultados existentes de su comportamiento asintótico.
- Se logra definir una categorización de los procesos de BGWM que permite un adecuado desarrollo e implementación de varios algoritmos con los que es posible realizar diferentes simulaciones.
- En las simulaciones se observan diferentes resultados numéricos a partir de dos situaciones que se pueden modelar mediante un proceso de BGWM, y es posible confrontar, como se espera, que los resultados numéricos coinciden con los resultados teóricos.
- Se muestra mediante simulaciones el comportamiento asintótico de las frecuencias relativas, cuando el número de individuos iniciales y el tiempo tienden a infinito, formulando dos resultados empíricos verificados por medio de las simulaciones.

Adicionalmente se tienen las siguientes observaciones:

- Los procesos de BGWM tienen una amplia gama de aplicaciones.
- Los procesos de ramificación han sido estudiados desde la matemática, sin embargo, hay mucho por desarrollar desde la estadística e incluso desde lo computacional.
- Desglosar las demostraciones de los diferentes resultados permite entender el funcionamiento, la estructura y la conexión que existe entre las hipótesis y las conclusiones, lo cual facilita llevar los resultados teóricos a la práctica o demostrar nuevos teoremas.
- El uso de simulaciones permite visualizar con mayor facilidad los resultados teóricos y realizar pruebas iniciales de nuevos resultados teóricos antes de pasar a su demostración formal.
- El desarrollo formal del paquete `Branching` en R asociado a este trabajo, requiere mayor dedicación que un desarrollo informal, pero es un trabajo enriquecedor, tanto a nivel personal, como a la comunidad académica, debido a la posibilidad de difusión, uso y construcción colaborativa.

Propuestas de investigación

- Realizar la demostración analítica formal de los Resultados 4.1 y 4.2, páginas 59 y 61 respectivamente, que surgen a partir de las simulaciones, o por otro lado, encontrar contraejemplos que muestren que los dichos resultados no se cumplen.
- Encontrar las expresiones analíticas de los estimadores de máximo verosimilitud que se obtienen a partir de frecuencias relativas observadas. En la sección 2.3, página 22, se concluye que es posible construir una función de log-verosimilitud que queda en términos de las frecuencias relativas observadas del proceso, y de los parámetros de interés, por lo tanto, en teoría es posible obtener estimadores de máximo verosimilitud de dichos parámetros, sin embargo la complejidad de la función parece indicar que el resultado teórico es difícil de llevar a la práctica.
- Analizar e implementar el estimador de mínimos cuadrados para la matriz de medias y el estimador relacionado para las matrices de varianzas y covarianzas. Maaouia & Touati (2005) mencionan las posibles expresiones de dichos estimadores, sin embargo, su deducción no incluye explícitamente la restricción de que todos los valores de la matriz de medias deben ser no negativos.
- Encontrar y comparar los posibles estimadores no paramétricos de la matriz de medias y las matrices de varianzas y covarianzas del proceso de ramificación de Bienaymé - Galton - Watson. Por ejemplo, aquellos obtenidos mediante métodos bayesianos.
- Comparar los estimadores existentes y propuestos de la matriz de medias y las matrices de varianzas y covarianzas del proceso de ramificación de Bienaymé - Galton - Watson, a partir de datos reales o simulados.
- Realizar un trabajo, similar a este, para otros tipos de procesos de ramificación. Por ejemplo, para procesos de ramificación de entorno aleatorio y para procesos de ramificación con parámetro de tiempo continuo, en lo posible, ambos desde un enfoque multitypo.

Bibliografía

- Asmussen, S. & Keiding, N. (1978), ‘Martingale Central Limit Theorems and Asymptotic Estimation Theory for Multitype Branching Processes’, *Advances in Applied Probability* **10**, 109–129.
- Athreya, K. & Ney, P. (1972), *Branching Processes*, Springer-Verlag, New York.
- Barbadilla, A. (2010), ‘La Genética de Poblaciones’, en línea. citado junio 3 de 2010.
*<http://158.109.215.191/base/base.asp?sitio=geneticapoblaciones&anar=concep>
- Dion, J.-P. (1974), ‘Estimation of the Mean and the Initial Probabilities of a Branching Process’, *Journal of Applied Probability* **11**(4), 687–694.
*<http://www.jstor.org/stable/3212552>
- Dion, J.-P. (1975), ‘Estimation of the Variance of a Branching Process’, *The Annals of Statistics* **3**(5), 1183–1187.
*<http://www.jstor.org/stable/2958373>
- Dion, J.-P. & Yanev, N. (1994), ‘Statistical Inference for Branching Processes with an Increasing Random Number of Ancestors’, *Journal of Statistical Planning and Inference* **39**, 329–352.
- Dion, J.-P. & Yanev, N. (1997), ‘Limit Theorems and Estimation Theory for Branching Processes with an Increasing Random Number of Ancestors’, *Journal of Applied Probability* **34**(2), 309–327.
- Dorman, K., Sinsheimer, J. & Lange, K. (2004), ‘In the Garden of Branching Processes’, *SIAM Review* **46**(2), 202–229.
- Epps, T. W. (1996), ‘Stock Prices as Branching Processes’, *Stochastic Models* **12**(4), 529–558.
- Free Software Foundation (2007), ‘The GNU General Public License’, en línea. citado junio 3 de 2010.
*<http://www.gnu.org/licenses/gpl.html>
- Goodman, L. A., Keyfitz, N. & Pullum, T. W. (1974), ‘Family formation and the frequency of various kinship relationships’, *Theoretical Population Biology* **5**(1), 1–27.
- Harris, T. E. (1963), *The Theory of Branching Processes*, Courier Dover Publications, Berlin.

- Heyde, C. C. (1974), 'On Estimating the Variance of the Offspring Distribution in a Simple Branching Process', *Advances in Applied Probability* **6**(3), 421–433.
*<http://www.jstor.org/stable/1426225>
- Heyde, C. C. & Seneta, E. (1972), 'Studies in the History of Probability and Statistics. XXXI. The Simple Branching Process, a Turning Point Test and a Fundamental Inequality: A Historical Note on I. J. Bienayme', *Biometrika* **59**(3), 680–683.
- Jagers, P. (1969), 'The Proportions of Individuals of Different Kinds in Two-Type Populations. A Branching Process Problem Arising in Biology', *Journal of Applied Probability* **6**(2), 249–260.
- Jarek, S. (2009), *mvnormtest: Normality test for multivariate variables*. R package version 0.1-7.
- Jones, J. H. (2007), 'demogR: A Package for the Construction and Analysis of Age-structured Demographic Models in R', *Journal of Statistical Software* **22**(10), 1–28.
*<http://www.jstatsoft.org/v22/i10>
- Kimmel, M. & Axelrod, D. E. (2002), *Branching Processes in Biology*, Springer, New York.
- Laredo, C., David, O. & Garnier, A. (2009), Inference for Partially Observed Multitype Branching Processes and Ecological Applications. preprint.
*<http://arxiv.org/pdf/0902.4520>
- Leslie, P. H. (1945), 'On the Use Of Matrices in Certain Population Mathematics', *Biometrika* **33**(3), 183–212.
*<http://biomet.oxfordjournals.org>
- Maaouia, F. & Touati, A. (2005), 'Identification of Multitype Branching Processes', *The Annals of Statistics* **33**(6), 2655–2694.
- Mitov, G., Rachev, S., Kim, Y. S. & Fabozzi, F. (2008), Barrier Option Pricing by Branching Processes, Technical report, Universitat Karlsruhe.
*<http://www.statistik.uni-karlsruhe.de/download/BarrierOptions-IJTAF.pdf>
- Mode, C. (1971), *Multitype Branching Processes. Theory and Applications*, American Elsevier Pub. Co., New York.
- Nanthi, K. (1982), 'Estimation of the Variance for the Multitype Galton - Watson Process.', *Journal of Applied Probability* **19**, 408–414.
- Ospina, R. (2001), 'Proceso de Galton - Watson', *Revista Colombiana de Estadística* **24**(1), 13–26.
- R Development Core Team (2009a), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*<http://www.R-project.org>
- R Development Core Team (2009b), *Writing R Extensions*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-11-9.
*<http://cran.r-project.org/doc/manuals/R-exts.pdf>

-
- Seidl, M., Etinski, M., Uiberacker, C. & Jakubetz, W. (2008), 'Pulse-train control of branching processes: Elimination of background and intruder state population', *The Journal of Chemical Physics* **129**(23), 234305.
- Stefanescu, C. (1998), 'Simulation of a Multitype Galton - Watson chain', *Simulation Practice and Theory* **6**(7), 657–663.
- Watson, H. W. & Galton, F. (1875), 'On the Probability of the Extinction of Families', *The Journal of the Anthropological Institute of Great Britain and Ireland* **4**, 138–144.
- Yakovlev, A. & Yanev, N. (2009), 'Relative Frequencies In Multitype Branching Processes', *Annals of Applied Probability* **19**(1), 1–14.