Universidad Nacional Autónoma de México

UNIDAD ACADEMICA DE LOS CICLOS PROFESIONAL Y DE POSGRADO DEL C.C.H.

INSTITUTO DE INVESTIGACIONES EN MATEMATICAS APLICADAS Y EN SISTEMAS

INFERENCIA ESTADISTICA EN LOS PROCESOS RAMIFICADOS DE GALTON-WATSON BISEXUALES

T E S I S

Que Para Obtener el Grado de

MAESTRO EN ESTADISTICA E INVESTIGACION DE OPERACIONES

Presenta el Actuario

Antonio Vicente González Fragoso

México, D. F. 1989

I N D I C E

INTRODUCCION 1

CAPITULO 1 PRELIMINARES 9

1.1 DISTRIBUCION SERIE DE POTENCIAS BIVARIADA 10

1.2 DISTRIBUCION POISSON BIVARIADA 24

1.3 METODO DE MINIMIOS CUADRADOS CONDICIONAL BIVARIADO . 30

1.4 RESULTADOS VARIOS 33

CAPITULO 2 PROCESOS RAMIFICADOS DE GALTON-WATSON BISEXUALES 35

2.1 DEFINICIONES Y RESULTADOS PRINCIPALES 35

2.2 APROXIMACIONES 42

2.3 SUPERMARTINGALAS ASOCIADAS 46

CAPITULO 3 ESTIMACION EN LOS PROCESOS RAMIFICADOS DE GALTON- WATSON BISEXUALES 48

3.1 MAXIMA VEROSIMILITUD 49

3.1.1 DISTRIBUCION SERIE DE POTENCIAS BIVARIADA 50

3.1.2 DISTRIBUCION POISSON BIVARIADA 54

3.1.3 ESTIMACION DE OTROS PARAMETROS 58

3.2 MINIMOS CUADRADOS CONDICIONAL 60

CAPITULO 4 PROPIEDADES ASINTOTICAS DE LOS ESTIMADORES .. 67

4.1 DISTRIBUCION ASINTOTICA 67

4.1.1 ESTIMADOR DEL VECTOR DE MEDIAS 69

4.1.2 ESTIMADOR DE LA TASA DE CRECIMIENTO 78

4.2 EFICIENCIA ASINTOTICA 80

BIBLIOGRAFIA 87

INTRODUCCION

El objetivo de este trabajo es desarrollar inferencia estadistica para los principales parámetros de los procesos ramificados de Galton-Watson bisexuales. Previamente se hace una revisión bibliográfica de estos procesos.

Los procesos ramificados de Galton-Watson bisexuales surgen como modelos más cercanos a la realidad para algunos fenómenos, en los que tradicionalmente han sido usados los procesos ramificados de Galton-Watson simples.

Para objetivos de explicación adoptaremos el lenguaje del crecimiento de una población.

Supongamos que una población de organismos comienza con un número fijo Z_0 de estos, y cada uno de estos seres reproduce en su vida, independientemente de los otros, un número de hijos de un mismo tipo, siendo estos los que forman la primera generación Z_1. Cada uno de los nuevos elementos reproducirá independientemente un número de descendientes, y de esta manera quedará constituida la siguiente generación Z_2 con los nuevos elementos. En general Z_{k+1} esta dada de la siguiente manera $Z_{k+1} = \sum_{i=1}^{Z_k} X_{k_i}$, donde X_{k_i} es definida como el número de hijos producidos por el i-ésimo organismo que pertenece a la k-ésima generación. Cada elemento reproduce, no importando en que generación se encuentre, con una misma distribución P(x) (distribución de descendientes) e independientemente de los otros. De esta manera la sucesión $\{Z_n, n \geq 0\}$ será un proceso de ramificación de Galton-Watson simple. Es importante mencionar que la función generatriz de probabilidad (f.g.p.), f(s) de Z_1 junto con la media del número de hijos que reproduce cada organismo, μ juegan un papel muy importante en el estudio del crecimiento de la población. Esto es, se sabe que la probabilidad de extinción puede ser encontrada como la raíz no negativa más pequeña de la ecuación $f(s) = s$. Además si $\mu \leq 1$, la población estará en vías de extinción, esto es, $P(Z_n=0$ para alguna $n) = 1$, y cuando la media es mayor que uno, la la tendencia del proceso será la de crecer (Athreya y Ney (1972), pág. 7). El conocimiento de la media μ es fundamental para conocer la tendencia de la población, es por eso que surge el interés de estimar este párametro. Las estimaciones de la media de reproducción por máxima verosimilitud y por mínimos cuadrados condicional, se encuentran en Basawa y Prakasa Rao (1980), pág 20-25, así como propiedades asintóticas del estimador de máxima verosimilitud. Aplicaciones a la detección de brotes de epidemias y aspectos estadísticos de los procesos de Galton-Watson simples aparecen en Pérez-Abreu (1987).

Una generalización de estos modelos son los procesos ramificados bitipos (Athreya y Ney (1972), pág 181-199). La diferencia consiste de que en lugar de considerar en la población un solo tipo de organismos se tienen dos, tipo I y tipo II (hembras y machos). Cada organismo del tipo i, i=1,2, reproduce independientemente en el transcurso de su vida, un número de descendientes del tipo I $X_1^{(i)}$ y un número del tipo II $X_2^{(i)}$, de acuerdo a una distribución bivariada $P(X_1^{(i)}= k_1, X_2^{(i)}= k_2)$ (distribución de descendientes). De esta manera, el número de organismos del tipo I y el número de organismos del tipo II en cada generación $\{(Z_{1_n}, Z_{2_n}), n \geq 0\}$ formarán el proceso. Obsérvese que en este proceso, existen cuatro medias de reproducción, por lo que es posible construir una matriz de medias M (de orden 2x2), de la siguiente manera $M = \begin{pmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{pmatrix}$, donde

μ_{ij} = el promedio de individuos del tipo j reproducidos por un elemento del tipo i, i=1,2 , j=1,2. De manera análoga a los procesos de Galton-Watson simples, las funciones generatrices de probabilidades $f(\underline{S}) = f(s_1,s_2) = (f^{(1)}(\underline{S}), f^{(2)}(\underline{S}))$ y la matriz de medias M son fundamentales para el análisis de la tendencia del proceso. Esto es, cuando el valor propio mayor ρ de M es mayor que uno, la población tiende a crecer y el vector de probabilidades de extinción podrá ser calculado como la raíz de $f(\underline{S}) = \underline{S}$ y cuando $\rho \leq 1$ tenderá a extinguirse el proceso, esto es las probabilidades de extinción serán igual a uno (veáse Athreya y Ney (1972), pág.186).

Los procesos ramificados bisexuales son un modelo mejor para aplicarlo al crecimiento de una población, ya que como se verá (capítulo 2), en este modelo se consideran dos tipos de individuos, hembras y machos; y además las parejas que se "casan", son las que reproducen los individuos, tanto hembras como machos, en cada generación. Es importante mencionar que en los procesos ramificados simples, son los organismos los que individualmente reproducen; en cambio en los procesos de Galton-Watson bisexuales, para la reproducción de nuevos seres, es necesario la unión o "casamiento" de dos individuos de diferente tipo. De esta manera los procesos de ramificación de Galton-Watson bisexuales quedarán constituidos por dos procesos diferentes, pero dependientes, uno bitipo el cual estará formado en cada una de las generaciones por el número de hembras y el número de machos $\{(F_n,M_n), n \geq 0\}$ y el otro proceso univariado formado por el número de parejas "casadas", en cada generación $\{Z_n , n \geq 0\}$, el cual esta dado por $Z_n = \mathcal{L}(F_n,M_n)$, donde $\mathcal{L}(\cdot,\cdot)$ es conocida como la función de casamiento (veáse Definición 2.1.1.). Es importante mencionar que existen condiciones para las cuales el proceso ramificado bisexual se encuentra en extinción o tiende a crecer, y estan dadas en función del parámetro tasa de crecimiento r, el cual se presenta en el capitulo 2. Cabe señalar que aunque en los procesos ramificados bisexuales existe la independencia de reproducción entre las parejas, lo que hace pensar que habrá cierta similaridad con los procesos de Galton-Watson simples, cualquier tipo de análisis en los procesos de Galton-Watson bisexuales se dificulta, ya que en estos no se tiene la propiedad aditiva de los primeros, la cual es una consecuencia de la relación entre las funciones generadoras de probabilidades en las sucesivas generaciones (veáse Bagley (1986)). Esto es, en los procesos ramificados simples existe la propiedad de que $f_n(s) = f_{n-1}(f(s))$ donde $f_k(s)$ es la f.g.p. de Z_k y dicha propiedad no se cumple en los procesos ramificados bisexuales. Esta aditividad permitía definir algunas martingalas de interés y aplicar entonces resultados asintóticos de las mismas. Desafortunadamente esta propiedad "martingala" se pierde en los procesos ramificados de Galton-Watson bisexuales.

El objetivo principal de este trabajo es estimar los principales parámetros de interés de los procesos de Galton-Watson bisexuales, así como estudiar las propiedades asintóticas de los estimadores. Ningún análisis estadístico para estos procesos fue encontrado en la literatura.

Para propósitos de estimación es necesario suponer una distribución de descendientes. Es por eso que ciertas propiedades de la distribución serie de potencias bivariada y de la distribución Poisson bivariada, se presentan en el Capítulo 1 de este trabajo. Es importante mencionar que una gran variedad de pares de variables aleatorias pertenecen a la familia serie de potencias bivariada. También en este capítulo se explica en que consiste el método de mínimos cuadrados condicional, el cual es un método de estimación, que a diferencia del método de máxima verosimilitud, no necesita del conocimiento de una distribución de descendientes. Finalmente en el Capítulo 1 se enuncian ciertos resultados aislados que son útiles para encontrar las propiedades asintóticas de nuestros estimadores.

En el capítulo 2 se describen los procesos ramificados bisexuales así como sus principales propiedades entre las que se encuentra una que es fundamental, ya que nos dice bajo que condiciones el proceso se encuentra en extinción o tiende a crecer

(Teorema 2.1.1.). La primera parte de este capítulo es una recopilación bibliográfica de la literatura sobre procesos ramificados bisexuales que cubre del trabajo de Daley (1968) hasta el de Daley, Hull y Taylor (1986). Es importante mencionar que hemos incluido una condición adicional a las funciones de casamiento tradicionalmente usadas, con la introducción de lo que hemos llamado funciones de casamiento naturales; esta condición es satisfecha por todos los ejemplos de funciones de casamiento encontrados en la literatura. Dentro de esta revisión se presentan las aproximaciones que existen para los procesos de Galton-Watson bisexuales. Estas permiten encontrar un rango para la probabilidad de extinción del proceso, ya que como mencionamos anteriormente no se tiene la propiedad aditiva y la f.g.p. no puede usarse como herramienta para encontrar la probabilidad de extinción. Finalmente en este capítulo encontramos algunas supermartingalas de interés relacionadas con estos procesos y que son de gran utilidad para probar propiedades asintóticas de los estimadores. Fue posible obtener estas supermartingalas gracias a la introducción de las funciones de casamiento naturales. Aunque simples, estos últimos resultados no fueron encontrados en la literatura.

La extinción o no de un proceso de Galton-Watson bisexual depende del valor de la tasa de crecimiento r (Teorema 2.1.1.). Así si $r \leq 1$ el proceso se extinguirá y si $r > 1$ tenderá a crecer. Como r es función de μ_1 y μ_2, las medias de reproducción, y no se tiene una función explícita de r en términos de μ_1 y μ_2 en general, la estimación se efectuará sobre las medias. La estimación por máxima verosimilitud del parámetro r no presenta algún problema, ya que es posible aplicar el principio de invarianza de máxima verosimilitud.

En el capítulo 3 abordaremos el problema de estimación de parámetros. Son dos los métodos de estimación que son empleados en este trabajo para la estimación del vector de las medias: el de máxima verosimilitud y el de mínimos cuadrados condicional. Con el de máxima verosimilitud se obtienen dos estimadores, uno donde la distribución de descendientes es la distribución serie de potencias bivariada y el otro donde es la Poisson bivariada. Estos dos estimadores resultaron ser diferentes. Con el método de mínimos cuadrados condicional fueron encontrados tres estimadores, resultando estos de modelos lineales diferentes, más solamente en uno de ellos los residuales cumplen con las propiedades usuales. El estimador encontrado por máxima verosimilitud, suponiendo distribución en serie de potencias bivariada y el estimador obtenido por mínimos cuadrados condicional coinciden. Este es el estimador del que nos ocupamos en el Capítulo 4 donde probamos sus propiedades asintóticas. Es importante mencionar que este no es un estimador asintoticamente normal, más bien la distribución asintótica es una normal mezclada (normal con varianza aleatoria). Sin embargo probamos que la distribución asintótica del producto de un factor aleatorio por la diferencia entre el estimador y el vector de medias es una distribución normal bivariada. Como una consecuencia de este resultado posteriormente se prueba que la distribución asintótica del producto de un elemento aleatorio por la diferencia entre el estimador de la tasa de crecimiento $\hat{r}$ y la tasa de crecimiento r es normal. Otra de las propiedades estudiadas en este capítulo es la eficiencia asintótica del estimador. El tipo de eficiencia que se considera en este trabajo es para modelos no ergódicos (veáse Basawa y Scott (1983)), esto es, la eficiencia según Heyde (1975). La razón por la que es necesario considerar este tipo de eficiencia es que los estimadores de máxima verosimilitud para estos modelos no son asintóticamente normales; por lo que no es posible usar la definición clásica de eficiencia. Probamos en este capítulo que nuestro estimador es eficiente en este nuevo sentido.

CAPITULO I

PRELIMINARES

En esta primera parte presentamos los resultados que no son propios del tema de procesos ramificados de Galton-Watson bisexuales (P.R.G.W.B.), pero que son utilizados dentro del desarrollo de este trabajo. En la primera sección se deducen ciertas propiedades de la distribución serie de potencias bivariada. Muchas variables aleatorias bivariadas pertenecen a esta familia, ya sea que las variables univariadas sean dependientes o independientes. Ejemplos de estas serán dadas en la Sección 1. En la segunda sección presentamos los resultados más importantes de la distribución Poisson bivariada, la cual como se verá no es de la familia serie de potencias bivariada. En la tercera sección de este capítulo se explica brevemente en que consiste el método de estimación de mínimos cuadrados condicional bivariado, y se da una solución para modelos lineales. Finalmente en la última parte se enlistan ciertos resultados de diversos temas de probabilidad necesarios para la parte de propiedades asintóticas de los correspondientes estimadores de los P.R.G.W.B..

1.1 DISTRIBUCION SERIE DE POTENCIAS BIVARIADA

La razón por la que nos interesa estudiar propiedades de la distribución serie de potencias bivariada, es porque se usa como distribución de descendientes más adelante, en la parte de estimación y cuando se estudian propiedades asintóticas.

En primer lugar se dará la definición de distribución serie de potencias bivariada. Posteriormente se dan ejemplos de distribuciones que pertenecen a esta familia, mostrando así lo amplia que es esta. Por último se estudian las propiedades fundamentales de esta familia de distribuciones, referente a la función generatriz de probabilidades (f.g.p.), medias, varianzas, covarianza y momentos factoriales.

DEFINICION 1.1.1 (Khatri (1959)). La distribución serie de potencias bivariada se define como

$$P(X_1=k_1,X_2=k_2) = \frac{a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}}{g(\theta_1,\theta_2)} = P_{k_1,k_2} \; ; \; k_i=0,1,2,\dots ; \; i=1,2$$

donde $g(\theta_1,\theta_2) = \sum_{k_1,k_2} a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}$ es una serie convergente tal que a_{k_1,k_2} puede ser función de k_1 y de k_2 o ser una constante y $a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2} \geq 0$.

Obsérvese que esta distribución no solo depende de θ_1 y θ_2, sino también de la función g, por lo tanto para referirnos a una distribución serie de potencias bivariada con parámetros θ_1 y θ_2, y función g, escribimos serie de potencias bivariada (θ_1, θ_2, g).

Cuando X_1 y X_2 son independientes a_{k_1,k_2} se puede expresar como $a_{k_1,k_2} = b_{k_1} \cdot c_{k_2}$ donde b_{k_1} no depende de k_2 y c_{k_2} no depende de k_1.

Como ejemplos de variables aleatorias bivariadas con distribución en serie de potencias bivariada tenemos a un par de variables aleatorias independientes entre si, cada una con alguna de las distribuciones Poisson, geométrica, binomial, binomial negativa o logaritmica, y en el caso dependiente, las distribuciónes trinomial, trinomial negativa, serie logaritmica bivariada. Asi por ejemplo cuando

1) X_1 y X_2 son variables aleatorias independientes tal que $X_1 \sim$ Poisson (λ) y $X_2 \sim$ geométrica (p), se tiene

$$P(X_1=k_1, X_2=k_2) = \frac{\lambda^{k_1}}{k_1!} e^{-\lambda} p(1-p)^{k_2} \qquad k_i=0,1,2,\ldots \ ; \ i=1,2$$

En este caso $\theta_1 = \lambda$, $\theta_2 = 1-p$, $g(\theta_1, \theta_2) = \frac{e^{\theta_1}}{1-\theta_2}$ y $a_{k_1,k_2} = \frac{1}{k_1!}$

2) $(X_1, X_2) \sim$ trinomial (n, p_1, p_2)

$$P(X_1=k_1, X_2=k_2) = \frac{n!}{k_1!k_2!k_3!} p_1^{k_1} p_2^{k_2} p_3^{k_3}$$

donde $k_i=0,1,\ldots,n$; $i=1,2$; $k_3=n-k_2-k_1$; $\sum_{i=1}^{3} p_i=1$.

Obsérvese que $\theta_1=p_1/p_3$; $\theta_2=p_2/p_3$; $g(\theta_1,\theta_2)=(1+\theta_1+\theta_2)^n$ y

$$a_{k_1,k_2} = \frac{n!}{k_1!k_2!k_3!}$$

3) $(X_1,X_2) \sim$ trinomial negativa (N,P_1,P_2)

$$P(X_1=k_1,X_2=k_2) = \frac{(N+k_1+k_2-1)!}{k_1!\ k_2!\ (N-1)!} Q^{-N-k_1-k_2} P_1^{k_1} P_2^{k_2}$$

donde $k_i=0,1,2,\ldots$; $i=1,2$; $Q - P_1 - P_2 = 1$; $P_i > 0$ y $N > 0$

En este caso se puede observar que $\theta_1 = P_1/Q$; $\theta_2 = P_2/Q$;

$$g(\theta_1,\theta_2) = (1 - \theta_1 - \theta_2)^{-N} \text{ y } a_{k_1,k_2} = \frac{(N-k_1-k_2-1)!}{k_1!\ k_2!\ (N-1)!}$$

4) $(X_1,X_2) \sim$ serie logaritmica bivariada

$$P(X_1=k_1,X_2=k_2) = \frac{(k_1+k_2-1)!}{k_1!\ k_2!\ [-\log(1-\theta_1-\theta_2)]} \theta_1^{k_1}\theta_2^{k_2}$$

donde $k_i=0,1,2,\ldots$, $i = 1,2$, pero no puede ser que $k_1=k_2=0$. En este caso $\theta_1=\theta_1$, $\theta_2=\theta_2$ y $g(\theta_1,\theta_2) = -\log(1-\theta_1-\theta_2)$

Análogo al caso de distribuciones serie de potencias univariadas se tiene el siguiente resultado, que habla sobre la función generatriz de probabilidades de una distribución serie de potencias bivariada.

PROPOSICION 1.1.1 La función generatriz de probabilidad de la distribución serie de potencias bivariada es:

$$f(\underline{S}) = f(s_1,s_2) = \frac{g(\underline{\theta}\cdot\underline{S})}{g(\underline{\theta})} \qquad \text{donde} \quad \underline{\theta}\cdot\underline{S} = (\theta_1 s_1, \theta_2 s_2)$$

y $\underline{\theta}=(\theta_1,\theta_2)$.

Demostración:

$$f(\underline{S}) = \sum_{k_1,k_2} P_{k_1,k_2}\, s_1^{k_1} s_2^{k_2} = \sum_{k_1,k_2} \frac{a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}}{g(\theta_1,\theta_2)}\, s_1^{k_1} s_2^{k_2} =$$

$$\sum_{k_1,k_2} \frac{a_{k_1,k_2}(\theta_1 s_1)^{k_1}(\theta_2 s_2)^{k_2}}{g(\theta_1,\theta_2)} = \frac{g(\theta_1 s_1,\theta_2 s_2)}{g(\theta_1,\theta_2)} = \frac{g(\underline{\theta}\cdot\underline{S})}{g(\underline{\theta})} \qquad \blacksquare$$

El siguiente resultado es importante para obtener la función de verosimilitud, cuando se usa como distribución de descendientes la serie de potencias bivariada.

PROPOSICION 1.1.2 Sean $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ variables aleatorias bivariadas independientes tal que se distribuyen como serie de potencias bivariada (θ_1,θ_2,g). Entonces la distribución de $(X,Y) = \sum_{i=1}^{n}(X_i,Y_i) = (\sum_{i=1}^{n}X_i, \sum_{i=1}^{n}Y_i)$ tiene la siguiente forma

$P(X=k_1,Y=k_2) = A_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}(g(\underline{\theta})^{-n})$, donde A_{k_1,k_2} puede depender de k_1 y k_2 o ser una constante, esto es, dependerá de las a_{k_1,k_2}'s.

Demostración:

Observemos en primer lugar que la función generatriz de probabilidades de (X,Y) es $f(\underline{S})^n$, donde $f(\underline{S})$ es la f.g.p. de (X_i,Y_i). Esto es $f_{X,Y}(\underline{S}) = E(s_1^{X}s_2^{Y}) = E(s_1^{X_1+\dots+X_n}s_2^{Y_1+\dots+Y_n})$

$= E(s_1^{X_1}\ s_2^{Y_1})\ E(s_1^{X_2}\ s_2^{Y_2})\ \cdots\ E(s_1^{X_n}\ s_2^{Y_n}) = \underbrace{f(\underline{S})\ \cdots\ f(\underline{S})}_{n\ \text{veces}}$

$= f(\underline{S})^n$.

De aquí que $P(X=k_1,Y=k_2)$ es el coeficiente de $s_1^{k_1}s_2^{k_2}$ en la expansión de $f_{X,Y}(\underline{S}) = (f(\underline{S}))^n$.

Ahora por inducción será demostrado lo que se quiere: Para $n = 1$

$P(X=k_1,Y=k_2) = P(X_1=k_1,Y_1=k_2) = a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}(g(\underline{\theta})^{-1})$, y

por lo tanto se tiene el resultado para n=1.

Supongamos que se cumple para $n = m$. Esto es

$$P\left(\sum_{i=1}^{m} X_i=k_1, \sum_{i=1}^{m} Y_i=k_2\right) = A_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}(g(\underline{\theta})^{-m})$$

Ahora la f.g.p. de $\sum_{i=1}^{m+1}(X_i,Y_i)$ es $f(\underline{S})^{m+1}$, pero

$$f(\underline{S})^{m+1} = f(\underline{S})^m f(\underline{S}) = \left[\sum_{k_1,k_2} A_{k_1,k_2}(\theta_1 s_1)^{k_1}(\theta_2 s_2)^{k_2}(g(\theta)^{-m})\right]$$

$$\left[\sum_{l_1,l_2} a_{l_1,l_2}(\theta_1 s_1)^{l_1}(\theta_2 s_2)^{l_2}(g(\underline{\theta})^{-1})\right] =$$

$$\sum_{k_1,k_2}\sum_{l_1,l_2} A_{k_1,k_2}\, a_{l_1,l_2}(\theta_1 s_1)^{k_1+l_1}(\theta_2 s_2)^{k_2+l_2}(g(\underline{\theta})^{-(m+1)}) =$$

$$\sum_{n_1,n_2}\left(\sum_{\substack{k_i+l_i=n_i\\ i=1,2}} A_{k_1,k_2}a_{l_1,l_2}\right)(\theta_1 s_1)^{n_1}(\theta_2 s_2)^{n_2}(g(\underline{\theta})^{-(m+1)}) =$$

$$\sum_{n_1,n_2} C_{n_1,n_2}\theta_1^{n_1}\theta_2^{n_2}s_1^{n_1}s_2^{n_2}(g(\underline{\theta})^{-(m+1)})$$

$$\therefore\ P\left(\sum_{i=1}^{m+1} X_i=k_1, \sum_{i=1}^{m+1} Y_i=k_2\right) = C_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2}(g(\underline{\theta})^{-(m+1)})$$

donde C_{k_1,k_2} depende de las a_{k_1,k_2}'s originales y por lo tanto dependerá de k_1 y de k_2 o será una constante ∎

Para una distribución serie de potencias bivariada (θ_1,θ_2,g), las medias de las distribuciones marginales se encuentran dadas en términos de las derivadas parciales de g . Para las varianzas es posible también encontrar una expresión en términos de θ_1 y θ_2, lo cual es dado por el siguiente resultado.

PROPOSICION 1.1.3

a) Si $\frac{\partial}{\partial s_i} f(\underline{S}) \Big|_{s_i=1}$ existe , i=1,2. Entonces

$$\mu_1 = E(X_1) = \frac{\theta_1}{g(\theta_1,\theta_2)} \left(\frac{\partial}{\partial \theta_1} g(\theta_1,\theta_2)\right) \qquad y$$

$$\mu_2 = E(Y_1) = \frac{\theta_2}{g(\theta_1,\theta_2)} \left(\frac{\partial}{\partial \theta_2} g(\theta_1,\theta_2)\right)$$

b) Si $\frac{\partial^2}{\partial s_i^2} f(\underline{S}) \Big|_{s_i=1}$ existe , i=1,2. Entonces

$$\sigma_1^2 = Var(X_1) = \theta_1\left(\frac{\partial\theta_1}{\partial\mu_1}\right)^{-1} \qquad y$$

$$\sigma_2^2 = Var(Y_1) = \theta_2\left(\frac{\partial\theta_2}{\partial\mu_2}\right)^{-1}$$

Demostración:

Bajo las suposiciones en la derivada de f

$$\mu_1 = [\frac{\partial}{\partial s_1} f(\underline{S})]_{\underline{S}=\underline{1}} =$$

$$= [(g(\underline{\theta})^{-1}) \sum_{k_1,k_2} a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2} k_1 s_1^{k_1-1} s_2^{k_2}]_{\underline{S}=\underline{1}} =$$

$$= \theta_1 (g(\underline{\theta})^{-1}) \sum_{k_1,k_2} a_{k_1,k_2} k_1 \theta_1^{k_1-1}\theta_2^{k_2} = \theta_1 (g(\underline{\theta})^{-1}) (\frac{\partial}{\partial \theta_1} g(\underline{\theta})).$$

Analogamente se demuestra que

$$\mu_2 = \theta_2 (g(\underline{\theta})^{-1})(\frac{\partial}{\partial \theta_2} g(\underline{\theta}))$$

Para demostrar las igualdades de (b) se usará el hecho

$$\sigma_1^2 = E[X_1(X_1-1)] + E[X_1] - [E(X_1)]^2 =$$

$$[\frac{\partial^2}{\partial s_1^2} f(\underline{S}) + \frac{\partial}{\partial s_1} f(\underline{S}) - (\frac{\partial}{\partial s_1} f(\underline{S}))^2]_{\underline{S}=\underline{1}}$$

En nuestro caso se tiene

$$[\frac{\partial^2}{\partial s_1^2} f(\underline{S})]_{\underline{S}=\underline{1}} = [(g(\underline{\theta})^{-1}) \sum_{k_1,k_2} a_{k_1,k_2}\theta_1^{k_1}\theta_2^{k_2} k_1(k_1-1) s_1^{k_1-2} s_2^{k_2}]_{\underline{S}=\underline{1}}$$

$$= \theta_1^2 (g(\underline{\theta})^{-1}) (\frac{\partial^2}{\partial \theta_1^2} g(\theta_1,\theta_2))$$

Por lo que σ_1^2 =

$$(g(\underline{\theta})^{-2}) \; \{g(\underline{\theta}) \; \theta_1^2(\frac{\partial^2}{\partial\theta_1^2}g(\underline{\theta})) + g(\underline{\theta}) \; \theta_1(\frac{\partial}{\partial\theta_1}g(\underline{\theta})) - \theta_1^2(\frac{\partial}{\partial\theta_1}g(\underline{\theta}))^2\}$$

Por (a) es posible llegar a $\frac{\partial}{\partial\mu_1} \; \mu_1 = 1 =$

$$(g(\underline{\theta})^{-2}) \; \{ \; g(\underline{\theta}) \; \theta_1 \; (\frac{\partial^2}{\partial\theta_1^2} \; g(\underline{\theta})) \; \frac{\partial\theta_1}{\partial\mu_1} + (\frac{\partial}{\partial\theta_1} \; g(\underline{\theta})) \; g(\underline{\theta}) \; \frac{\partial\theta_1}{\partial\mu_1} -$$

$$\theta_1 \; (-\frac{\partial}{\partial\theta_1}g(\underline{\theta}))^2 \; \frac{\partial\theta_1}{\partial\mu_1} \; \}$$, lo que implica que

$$\theta_1\left(\frac{\partial\theta_1}{\partial\mu_1}\right)^{-1} = (g(\underline{\theta})^{-2}) \; \{\theta_1^2 \; g(\underline{\theta}) \; (\frac{\partial^2}{\partial\theta_1^2} \; g(\underline{\theta})) + \theta_1 \; g(\underline{\theta}) \; (\frac{\partial}{\partial\theta_1} \; g(\underline{\theta}))$$

$$- \; \theta_1^2 \; (\frac{\partial}{\partial\theta_1} \; g(\underline{\theta}))^2\}$$

Pero esta última expresión es σ_1^2 , por lo que

$$\sigma_1^2 = \theta_1\left(\frac{\partial\theta_1}{\partial\mu_1}\right)^{-1} .$$

Analogamente se puede también demostrar que

$$\sigma_2^2 = \theta_2\left(\frac{\partial\theta_2}{\partial\mu_2}\right)^{-1} \qquad \blacksquare$$

Existe también una expresión para la covarianza de X_1 y X_2 en términos de g y sus derivadas parciales.

PROPOSICION 1.1.4

Si $\frac{\partial^2}{\partial s_i \ \partial s_j} f(\underline{S}) \Big|_{s_i=s_j=1}$ existe para i=1,2 y j=1,2. Entonces

$$E(X_1X_2) = \theta_1\theta_2(g(\underline{\theta})^{-1})\left(\frac{\partial^2}{\partial\theta_1\partial\theta_2} g(\underline{\theta})\right) \qquad \text{y por lo tanto}$$

$$\sigma_{2,1} = \sigma_{1,2} = \mathrm{Cov}(X_1,X_2) = \frac{\theta_1\theta_2}{g(\underline{\theta})} \left\{ \left(\frac{\partial^2}{\partial\theta_1\partial\theta_2} g(\underline{\theta})\right) - \right.$$

$$\left. (g(\underline{\theta})^{-1} \left(\frac{\partial}{\partial\theta_1} g(\underline{\theta})\right) \left(\frac{\partial}{\partial\theta_2} g(\underline{\theta})\right) \right\}$$

Demostración:

$$\frac{\partial^2}{\partial s_1 \partial s_2} f(\underline{S}) \Big]_{\underline{S}=\underline{1}} = (g(\underline{\theta})^{-1}) \sum_{k_1,k_2} a_{k_1,k_2} \theta_1^{k_1} \theta_2^{k_2} k_1 s_1^{k_1-1} k_2 s_2^{k_2-1} \Big]_{\underline{S}=\underline{1}}$$

$$= \theta_1 \ \theta_2 \ (g(\underline{\theta})^{-1}) \sum_{k_1,k_2} a_{k_1,k_2} \ k_1 \ \theta_1^{k_1-1} \ k_2 \ \theta_2^{k_2-1} =$$

$$\theta_1 \ \theta_2 \ (g(\underline{\theta})^{-1}) \left(\frac{\partial^2}{\partial\theta_1\partial\theta_2} g(\underline{\theta})\right)$$

Pero $\frac{\partial^2}{\partial s_1 \partial s_2} f(\underline{S})]_{\underline{S}=\underline{1}} = E(X_1 X_2)$.

De esta manera queda demostrada la primera igualdad. Para la segunda solamente se sustituye las expresiones correspondientes en $Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$. ∎

Los momentos factoriales también se pueden expresar en términos de las derivadas parciales de la función g como lo indica el siguiente resultado.

PROPOSICION 1.1.5

$$E(X_1(X_1-1)) = \frac{\theta_1^{\,2}}{g(\underline{\theta})} \left(\frac{\partial^2 g(\underline{\theta})}{\partial \theta_1^{\,2}}\right)$$

$$E(X_2(X_2-1)) = \frac{\theta_2^{\,2}}{g(\underline{\theta})} \left(\frac{\partial^2 g(\underline{\theta})}{\partial \theta_2^{\,2}}\right)$$

Demostración:

$$\frac{1}{g(\underline{\theta})} \quad \frac{\partial^2 g(\underline{\theta})}{\partial \theta_2^{\,2}} =$$

$$\frac{1}{g(\underline{\theta})} \sum_{k_1} \sum_{k_2} a_{k_1,k_2} \theta_1^{k_1} k_2(k_2-1) \theta_2^{k_2-2} = \frac{1}{\theta_2^{\,2}} \sum_{k_1} \sum_{k_2} k_2(k_2-1) a_{k_1,k_2}$$

$$\theta_1^{k_1} \theta_2^{k_2} (g(\theta)^{-1}) = \frac{1}{\theta_2^{\,2}} E(X_2(X_2-1))$$

Para X_1, la demostración es análoga. ∎

La siguiente expresión será de utilidad para encontrar los estimadores de máxima verosimilitud en los procesos de Galton-Watson bisexuales.

PROPOSICION 1.1.6

$$\frac{1}{\theta_1}\left[\frac{\partial\theta_1}{\partial\mu_2}\right] = -\frac{\sigma_{12}}{\sigma_1^{\,2}\sigma_2^{\,2}} = \frac{1}{\theta_2}\left[\frac{\partial\theta_2}{\partial\mu_1}\right]$$

Demostración:

Se tiene que $\mu_2 = \frac{\theta_2}{g(\underline{\theta})}\left[\frac{\partial}{\partial\theta_2}\, g(\underline{\theta})\right]$

$$\therefore \frac{\partial\mu_2}{\partial\mu_1} = 0 = \frac{\theta_2}{g(\underline{\theta})}\left(\frac{\partial}{\partial\theta_1}\,\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial\theta_1}{\partial\mu_1}\right) + \frac{\theta_2}{g(\underline{\theta})}\left(\frac{\partial}{\partial\theta_2}\,\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial\theta_2}{\partial\mu_1}\right)$$

$$+ \quad \theta_2\left(\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial}{\partial\theta_1}\,\frac{1}{g(\underline{\theta})}\right)\left(\frac{\partial\theta_1}{\partial\mu_1}\right) \quad + \quad \theta_2\left(\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial}{\partial\theta_2}\,\frac{1}{g(\underline{\theta})}\right)\left(\frac{\partial\theta_2}{\partial\mu_1}\right) \quad +$$

$$\frac{1}{g(\underline{\theta})}\left(\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial\theta_2}{\partial\mu_1}\right) = \theta_2\,\frac{E(X_1X_2)}{\theta_1\theta_2}\left(\frac{\partial\theta_1}{\partial\mu_1}\right) \quad + \quad \theta_2\quad\left(\frac{E(X_2^{\,2})-\mu_2}{\theta_2^{\,2}}\right)\quad\left(\frac{\partial\theta_2}{\partial\mu_1}\right) \quad +$$

$$\theta_2\left(\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(-\frac{1}{g(\underline{\theta})^2}\,\frac{\partial g(\underline{\theta})}{\partial\theta_1}\right)\left(\frac{\partial\theta_1}{\partial\mu_1}\right) + \theta_2\left(\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(-\frac{1}{g(\underline{\theta})^2}\,\frac{\partial g(\underline{\theta})}{\partial\theta_2}\right)\left(\frac{\partial\theta_2}{\partial\mu_1}\right) +$$

$$\frac{\mu_2}{\theta_2}\left(\frac{\partial\theta_2}{\partial\mu_1}\right) = \frac{1}{\theta_1}\left(\frac{\partial\theta_1}{\partial\mu_1}\right)(E(X_1X_2) \quad - \quad \mu_1\mu_2) \quad + \quad \frac{1}{\theta_2}\left(\frac{\partial\theta_2}{\partial\mu_1}\right)(E(X_2^2)-\mu_2^2) \quad =$$

$$\frac{\sigma_{12}}{\sigma_1^2} + \frac{1}{\theta_2}\left(\frac{\partial\theta_2}{\partial\mu_1}\right)\sigma_2^2 = 0$$

De aquí $\frac{1}{\theta_2}\left(\frac{\partial\theta_2}{\partial\mu_1}\right) = -\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}$

Analogamente se demuestra que $\frac{1}{\theta_1}\left(\frac{\partial\theta_1}{\partial\mu_2}\right) = -\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}$ ■

Lo que se ha hecho hasta el momento es expresar μ_1, μ_2, σ_1^2, σ_2^2, $\sigma_{1,2}$ $(\sigma_{2,1})$ en términos de θ_1, θ_2, $g(\underline{\theta})$ y las derivadas parciales de $g(\underline{\theta})$ con respecto a θ_1 y θ_2. Cabe mencionar que estas expresiones son útiles para encontrar el estimador de máxima verosimilitud del vector de medias, en un P.R.G.W.B. (veáse capítulo 3 §1).

OBSERVACION Cuando $\theta_1=\theta_2=\theta$ no es posible representar las medias en términos de las derivadas parciales de $g(\underline{\theta})$.

Esto es $\mu_1 = \frac{\partial}{\partial s_1} f(\underline{S})\big]_{\underline{S}=\underline{1}} = (g(\underline{\theta})^{-1})\frac{\partial}{\partial s_1}\sum_{k_1,k_2} a_{k_1,k_2}\theta^{k_1+k_2}$

$s_1^{k_1}s_2^{k_2}\big]_{\underline{S}=\underline{1}} = (g(\theta)^{-1})\sum_{k_1,k_2} a_{k_1,k_2}\theta^{k_1+k_2}k_1 s_1^{k_1-1}s_2^{k_2}\big]_{\underline{S}=\underline{1}} =$

$$(g(\theta_1,\theta_2)^{-1}) \sum_{k_1,k_2} a_{k_1,k_2} k_1 \theta^{k_1+k_2}$$

Y lo anterior no es igual a alguna derivada, en términos de θ.

En general las proposiciones a partir de la 1.1.3 no son válidos para cuando $\theta_1 = \theta_2$. Más adelante, en cuestión de estimación, se estudia por separado el caso $\theta_1 = \theta_2$. Como se verá, no siempre es posible obtener el estimador cuando estos dos parámetros son iguales.

1.2 DISTRIBUCION POISSON BIVARIADA

En esta sección se presentan las propiedades fundamentales de la distribución Poisson bivariada, la cual no forma parte de la familia serie de potencias bivariada. Se estudiará lo referente a las distribuciones marginales , f.g.p., medias, varianzas, covarianza y otros resultados propios de esta distribución. El material presentado en esta sección se encuentra en gran parte en Johnson y Kotz (1969).

DEFINICION 1.2.1 Sean u, v_1, v_2 variables aleatorias independientes con distribución Poisson y parámetros λ_1, λ_2, λ_3 respectivamente. Entonces la distribución conjunta de $X_1 = u + v_1$ y $X_2 = u + v_2$ es conocida como la distribución de Poisson bivariada con parámetros λ_1, λ_2, λ_3. Una expresión para la densidad de probabilidades es

$$P(x_1,x_2) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min(x_1,x_2)} \frac{\lambda_1^{\,j}}{j!} \frac{\lambda_2^{\,x_1-j}}{(x_1-j)!} \frac{\lambda_3^{\,x_2-j}}{(x_2-j)!}$$

Obsérvese que las distribuciones marginales de X_1 y X_2 son Poisson con valores esperados $\lambda_1 + \lambda_2$ y $\lambda_1 + \lambda_3$ respectivamente.

También obsérvese que la distribución Poisson bivariada, tiene involucrado tres parámetros independientes, y por este simple hecho, no pertenece a la familia serie de potencias bivariada.

PROPOSICION 1.2.1 La función generadora de probabilidades de la Poisson bivariada es

$$\phi(\underline{S}) = \phi(s_1, s_2) = \exp[-\lambda_1(1-s_1 s_2) - \lambda_2(1-s_1) - \lambda_3(1-s_2)]$$

Demostración:

$$\phi(\underline{S}) = E(s_1^{x_1} s_2^{x_2}) = E((s_1 s_2)^u s_1^{v_1} s_2^{v_2}) = E((s_1 s_2)^u)$$

$$E(s_1^{v_1}) \quad E(s_2^{v_2}) = e^{-\lambda_1(1-s_1 s_2)} \; e^{-\lambda_2(1-s_1)} \; e^{-\lambda_3(1-s_2)} =$$

$$\exp[-\lambda_1(1-s_1 s_2) - \lambda_2(1-s_1) - \lambda_3(1-s_3)] \quad \blacksquare$$

Obsérvese que la varianza de X_1 y de X_2 es $\lambda_1 + \lambda_2$ y $\lambda_1 + \lambda_3$ respectivamente, y la covarianza de estas dos variables es λ_1. Como en el caso de la Poisson univariada, las varianzas coinciden con las medias.

También de la expresión de la función generadora de probabilidades, se obtiene que la suma de n variables aleatorias bivariadas independientes con distribución Poisson bivariada $(\lambda_1, \lambda_2, \lambda_3)$ se distribuye Poisson bivariada $(n\lambda_1, n\lambda_2, n\lambda_3)$.

Las siguientes expresiones son de gran utilidad.

PROPOSICION 1.2.2 Sea $(x_1, x_2) \sim$ Poisson bivariada $(\lambda_1, \lambda_2, \lambda_3)$ y sea $P(x_1, x_2)$ la función de probabilidad entonces

$$x_1 P(x_1, x_2) = \lambda_2 P(x_1-1, x_2) + \lambda_1 P(x_1-1, x_2-1)$$

$$x_2 P(x_1,x_2) = \lambda_3 P(x_1,x_2-1) + \lambda_1 P(x_1-1,x_2-1)$$

Demostración:

$$\lambda_2 P(x_1-1,x_2) + \lambda_1 P(x_1-1,x_2-1) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min(x_1-1,x_2)} \frac{\lambda_1^{j}}{j!}(x_1-j)\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} + e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min(x_1-1,x_2-1)} (j+1)\frac{\lambda_1^{j+1}}{(j+1)!}$$

$$\frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!}\frac{\lambda_3^{x_2-1-j}}{(x_2-1-j)!} = e^{-(\lambda_1+\lambda_2+\lambda_3)}\left\{x_1 \sum_{j=0}^{\min(x_1-1,x_2)} \frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!}\right.$$

$$- \sum_{j=0}^{\min(x_1-1,x_2)} j\,\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} + \sum_{j=0}^{\min(x_1-1,x_2-1)} (j+1)\frac{\lambda_1^{j+1}}{(j+1)!}$$

$$\left.\frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!}\frac{\lambda_3^{x_2-1-j}}{(x_2-1-j)!}\right\} \qquad (1.1)$$

Si $\min(x_1,x_2) = x_1$, la expresión (1.1) es igual a:

$$e^{-(\lambda_1+\lambda_2+\lambda_3)}\left\{x_1 \sum_{j=0}^{x_1-1} \frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} - \sum_{j=0}^{x_1-1} j\,\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\right.$$

$$\frac{\lambda_3^{x_2-j}}{(x_2-j)!} + \sum_{j=0}^{x_1-1} (j+1)\frac{\lambda_1^{j+1}}{(j+1)!}\frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!}\frac{\lambda_3^{x_2-1-j}}{(x_2-1-j)!}\Bigg\} =$$

$$e^{-(\lambda_1+\lambda_2+\lambda_3)}\Bigg\{x_1\sum_{j=0}^{x_1-1}\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} - \sum_{j=0}^{x_1-1} j\,\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!}$$

$$+\sum_{j=0}^{x_1-2}(j+1)\frac{\lambda_1^{j+1}}{(j+1)!}\frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!}\frac{\lambda_3^{x_2-1-j}}{(x_2-1-j)!} + x_1\frac{\lambda_1^{x_1}}{x_1!}\frac{\lambda_3^{x_2-x_1}}{(x_2-x_1)!}\Bigg\}$$

Obsérvese que la segunda y la tercera sumatoria se anulan y por lo tanto la expresión anterior es igual a $x_1P(x_1,x_2)$

Si $\min(x_1,x_2) = x_2$, hay dos subcasos:

1) Cuando $x_2 = x_1-1$. Entra en el caso anterior.

2) Cuando $x_2 < x_1-1$, la expresión (1) queda como:

$$e^{-(\lambda_1+\lambda_2+\lambda_3)}\Bigg\{x_1\sum_{j=0}^{x_2}\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} - \sum_{j=0}^{x_2} j\,\frac{\lambda_1^{j}}{j!}\frac{\lambda_2^{x_1-j}}{(x_1-j)!}\frac{\lambda_3^{x_2-j}}{(x_2-j)!} +$$

$$\sum_{j=0}^{x_2-1}(j+1)\frac{\lambda_1^{j+1}}{(j+1)!}\frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!}\frac{\lambda_3^{x_2-1-j}}{(x_2-1-j)!}\Bigg\} = x_1P(x_1,x_2),$$ ya que las

últimas dos sumatorias se cancelan .

Analogamente se demuestra la segunda igualdad ■

PROPOSICION 1.2.3

$$\frac{\partial P(x_1,x_2)}{\partial \lambda_1} = P(x_1-1,x_2-1) - P(x_1,x_2)$$

$$\frac{\partial P(x_1,x_2)}{\partial \lambda_2} = P(x_1-1,x_2) - P(x_1,x_2)$$

$$\frac{\partial P(x_1,x_2)}{\partial \lambda_3} = P(x_1,x_2-1) - P(x_1,x_2)$$

Demostración:

$$\frac{\partial P(x_1,x_2)}{\partial \lambda_1} = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min(x_1,x_2)} \frac{\lambda_1^{j-1}}{(j-1)!} \frac{\lambda_2^{x_1-j}}{(x_1-j)!}$$

$$\frac{\lambda_3^{x_2-j}}{(x_2-j)!} - e^{-(\lambda_1+\lambda_1+\lambda_2)} \sum_{j=0}^{\min(x_1,x_2)} \frac{\lambda_1^{j}}{j!} \frac{\lambda_2^{x_1-j}}{(x_1-j)!} \frac{\lambda_3^{x_2-j}}{(x_2-j)!}$$

Sea $j' = j-1$, por lo que la expresión anterior es igual

a $$e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j'=0}^{\min(x_1,x_2)-1} \frac{\lambda_1^{j'}}{j'!} \frac{\lambda_2^{x_1-j'-1}}{(x_1-j'-1)!} \frac{\lambda_3^{x_2-j'-1}}{(x_2-j'-1)!} - P(x_1,x_2)$$

$= P(x_1-1,x_2-1) - P(x_1,x_2)$.

$$\frac{\partial P(x_1,x_2)}{\partial \lambda_2} = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{j=0}^{\min(x_1,x_2)} \frac{\lambda_1^{j}}{j!} (x_1-j) \frac{\lambda_2^{x_1-1-j}}{(x_1-j)!}$$

$$\frac{\lambda_3^{x_2-j}}{(x_2-j)!} - P(x_1,x_2) = e^{-(\lambda_1+\lambda_1+\lambda_2)} \sum_{j=0}^{\min(x_1-1,x_2)} \frac{\lambda_1^{j}}{j!} \frac{\lambda_2^{x_1-1-j}}{(x_1-1-j)!} \frac{\lambda_3^{x_2-j}}{(x_2-j)!}$$

$$- P(x_1,x_2) = P(x_1-1,x_2) - P(x_1,x_2).$$

La tercera igualdad se demuestra de forma análoga a la segunda. ∎

Las proposiciones anteriores serán de gran utilidad para encontrar los estimadores de las medias en un P.R.G.W.B., ya que como se verá, no será posible encontrar una forma explicita para la función de verosimilitud cuando la distribución de descendientes es la Poisson bivariada (veáse la sección 3.1.2.).

1.3 METODO DE MINIMOS CUADRADOS CONDICIONAL

En esta sección se explicará brevemente en que consiste el método de estimación mínimos cuadrados condicional, en un proceso estocástico bivariado. Temas adicionales sobre este método se pueden consultar en Klimko y Nelson (1978) y la solución para modelos lineales en Seber (1984).

Supongamos que se tiene $\underline{Y}_t$ $t=1,2,\ldots$, un proceso estocástico bivariado, definido en un espacio de probabilidad $(\Omega,\mathcal{F},P)$, cuya distribución depende de un vector de parámetros desconocidos $\underline{\alpha} = (\alpha_1,\alpha_2)$. Sea $\{\mathcal{F}_t\}_{t=0}^{\infty}$ una sucesión de sub-σ-álgebras de $\mathcal{F}$ tal que $\mathcal{F}_t$ es generada por $\{\underline{Y}_1,\ldots,\underline{Y}_t\}$ $t \geq 1$ y $\mathcal{F}_0$ es la σ-álgebra trivial.

Dado un conjunto de observaciones $\underline{Y}_1,\ldots,\underline{Y}_n$, el método de mínimos cuadrados condicional, consiste en estimar $\underline{\alpha}$, por el valor que minimiza

$$Q_n(\underline{\alpha}) = U'U = (Y-\Theta)'(Y-\Theta)$$

donde Y es la matriz que tiene como i-ésimo renglón a $\underline{Y}_i'$ y Θ es la matriz donde el i-ésimo renglón es $E(\underline{Y}_i' \mid \mathcal{F}_{i-1})$. Es importante observar que U'U es una matriz de orden 2x2, y para minimizarla es necesario definir una relación de orden entre matrices. Así si C y D son dos matrices simétricas de las mismas dimensiones, entonces $C \geq D$ cuando $C - D$ es semipositiva definida, esto es, cuando $C-D \geq 0$.

DEFINICION 1.3.1. Si $C(\theta)$ es una función matriz simétrica, se dice que C tiene un mínimo en $\theta = \hat{\theta}$, si $C(\theta) \geq C(\hat{\theta})$ $\forall$ θ.

Cuando Θ puede ser escrito de la forma $X\underline{\alpha}$, donde X son observaciones que dependen de Y, es posible aplicar la solución del estimador de $\underline{\alpha}$ que da Seber (1984) para modelos lineales, y que tiene la siguiente expresión

$$\hat{\underline{\alpha}} = (X'X)^{-1}X'Y \text{ , cuando } (X'X)^{-1} \text{ existe.}$$

Esta solución se obtiene encontrando el valor Θ que minimiza U'U, sujeto a que las columnas $\theta^{(j)}$ de Θ, pertenezcan a $\Xi = \mathcal{R}[X]$, el espacio columna de X.

Es importante observar que si Θ puede ser escrito como $X\underline{\alpha}$, estaremos hablando que nuestro proceso cumple con el modelo lineal $Y = X\underline{\alpha} + U$, donde U son definidos como los residuales del modelo.

Para esto ciertas propiedades de los residuales U se deben de cumplir: Supongamos que $U = \begin{bmatrix} \underline{U}_1' \\ \underline{U}_2' \\ \vdots \\ \underline{U}_n' \end{bmatrix}$, donde $\underline{U}_i' = (U_{i1}, U_{i2})$, entonces

$E[U] = 0$ y $Cov\left(\underline{U}_h, \underline{U}_i\right) = \delta_{hi}\,\Sigma$; $h,i = 1,2,\ldots,n$, donde $\delta_{hi} = 1$ si $h = i$ y $\delta_{hi} = 0$ si $h \neq i$, y Σ es la matriz de dispersión común (Esto es, $\Sigma = Cov(\underline{Y}_i')$).

Obsérvese que en otros términos, estas propiedades de los residuales pueden escribirse de la siguiente manera

$E[U_{hj}\ U_{hj}] = Var(U_{hj}) = \sigma_j^2 \quad ; \ h = 1,\ldots,n \ ; \ j = 1,2$

$E[U_{h1}\ U_{h2}] = E[U_{h2}\ U_{h1}] = Cov[U_{h1},U_{h2}] = \sigma_{12} \quad ; h = 1,\ldots,n$

$E[U_{hj}\ U_{ik}] = Cov[U_{hj},U_{ik}] = 0 \ ; \ h,i = 1,\ldots,n \ , \text{con } h\neq i \ ; \ j,k = 1,2$

1.4 RESULTADOS VARIOS

Finalmente, y con el objeto de facilitar la lectura, presentamos tres resultados límites que serán útiles para demostrar las propiedades asintóticas de nuestros estimadores.

TEOREMA 1.4.1 Sea $\{\xi_n, \mathcal{F}_n\}_{n=1}^{\infty}$ una supermartingala. Si $\lim_{n\to\infty} E(\xi_n^+) < \infty$, entonces existe una variable aleatoria integrable ξ tal que $\xi_n \to \xi$ casi seguramente.

LEMA 1.4.2. (DE TOEPLITZ) (Loéve (1978)) Sea a_{nk} $k=1,2,\dots,k_n$ números tal que $\forall$ k $a_{nk} \to 0$ y $\forall$ n $\sum_k |a_{nk}| \le c < \infty$. Sea $X_n' = \sum_k a_{nk}X_k$. Entonces $X_n \to 0$ implica $X_n' \to 0$ y si $\sum_k a_{nk} \to 1$, entonces $X_n \to X$ finito implica $X_n' \to X$. En particular, si $b_n = \sum_{k=1}^{n} a_k \uparrow \infty$, entonces $X_n \to X$ finito implica $\frac{1}{b_n} \sum_{k=1}^{n} a_k X_k \to X$.

TEOREMA 1.4.3 (TEOREMA DE LIMITE CENTRAL PARA SUMAS ALEATORIAS) (Billingsley (1968)) Supóngase ξ_1, ξ_2, ... son variables aleatorias independientes e idénticamente distribuidas con media cero y varianza σ^2. Si $\frac{V_n}{a_n} \xrightarrow{P} W$, donde W es una variable aleatoria positiva y a_n son tales que tienden a ∞ , entonces

$$Y_n = \frac{\xi_1+\xi_2+ \ldots +\xi_{V_n}}{\sqrt{V_n}\ \sigma_n} \xrightarrow{\mathcal{D}} N(0,1)$$

CAPITULO 2

PROCESOS RAMIFICADOS DE GALTON-WATSON BISEXUALES

En este capitulo se presentan las definiciones y resultados más importantes de los procesos de Galton-Watson bisexuales. Además se estudian ciertas aproximaciones que existen para este tipo de procesos. Este material se encuentra en Daley, Hull y Taylor (1986). Por último presentamos algunas supermartingalas asociadas a los procesos de Galton-Watson bisexuales.

2.1 DEFINICIONES Y RESULTADOS PRINCIPALES

DEFINICION 2.1.1 Un proceso de ramificación de Galton-Watson bisexual (P.R.G.W.B.) es aquel que tiene la siguiente estructura: consiste de un proceso bitipo, en el cual, sucesivas generaciones $n=0,1,2,...,$ se componen de F_n hembras y M_n machos los cuales forman $Z_n=\mathscr{L}(F_n,M_n)$ unidades de casamiento y donde cada una de estas reproduce, independientemente de las otras, de acuerdo a una misma distribución de descendientes bivariada. Denotaremos por $P_Z(\underline{k})=P(k_1,k_2)$ a la probabilidad de que una unidad de casamiento produzca k_1 hembras y k_2 machos. Esto es, se tienen $\{(X_{n_i},Y_{n_i});$ $n=0,1,...,$ $i=1,2,...,Z_n\}$ variables aleatorias bivariadas independientes e identicamente distribuidas , con entradas enteras no negativas tal que

$$(F_{n+1}, M_{n+1}) = \sum_{i=1}^{Z_n} (X_{n_i}, Y_{n_i}).$$

La función $\mathcal{L}(\cdot,\cdot)$ se llama la función de casamiento, la cual es una función monótona no decreciente en cada argumento, valuada en parejas enteras con entradas no negativas, que toma valores en los enteros no negativos y además $\mathcal{L}(0,0)=0$.

Como ejemplos de funciones de casamiento tenemos las propuestas por Daley (1968), y que dan lugar a lo que llamaremos modelos M1, M2 y M3 respectivamente.

1) $\mathcal{L}(x,y) = \min(x,y) \quad \forall \ x,y \geq 0$ Casamiento con Fidelidad.

2) $\mathcal{L}(x,y) = x \min(x,y) \quad \forall \ x,y \geq 0$ Casamiento Promiscuo.

3) $\mathcal{L}(x,y) = x \quad \forall \ x,y \geq 0$ Casamiento Estándar.

DEFINICION 2.1.2 Un P.R.G.W.B. es superaditivo cuando la función de casamiento $\mathcal{L}(\cdot,\cdot)$ es superaditiva, o sea es tal que $\mathcal{L}(x_1+x_2, y_1+y_2) \geq \mathcal{L}(x_1,y_1) + \mathcal{L}(x_2,y_2) \ \forall \ x_1, y_1, x_2, y_2 \in [0,\infty)$.

Observemos que las funciones de casamiento de los modelos M1, M2 y M3 son superaditivas.

Dada una función de casamiento superaditiva es posible encontrar una familia de funciones de casamiento superaditivas de la siguiente manera.

PROPOSICION 2.1.1 Sea $\mathcal{L}(x,y)$ una función de casamiento superaditiva entonces $\mathcal{L}_o(x,y) = (1-\alpha)(1-\beta)\mathcal{L}(x,y) + (1-\alpha)\beta\mathcal{L}(x,y+1) + \alpha(1-\beta)\mathcal{L}(x+1,y) + \alpha\beta\mathcal{L}(x+1,y+1)$ con $\alpha,\beta \in [0.1]$, es también una función de casamiento superaditiva.

La superaditividad de $\mathcal{L}_o(\cdot,\cdot)$ puede demostrarse a partir de
$$\mathcal{L}_o(x + \alpha, y + \beta) = \alpha\mathcal{L}_o(x + 1, y + \beta) + (1-\alpha)\mathcal{L}_o(x, y + \beta)$$

Hemos incluido la siguiente condición, pues será de suma utilidad en los dos últimos capítulos.

DEFINICION 2.1.3. Una función de casamiento $\mathcal{L}(\cdot,\cdot)$ que cumple con $\mathcal{L}(x,y) \leq x \quad \forall\ (x,y)$, la llamaremos función de casamiento natural.

Que un P.R.G.W.B. tenga función de casamiento natural, significa que el número de parejas que se unen para reproducir, no puede exceder del número de hembras que hay. Es muy lógico pensar que dicha propiedad suceda en la realidad, ya que es en la hembra donde se genera el nuevo organismo. Los modelos M1, M2 y M3 son funciones de casamiento natural.

El supuesto de independencia entre las parejas aleatorias (X_{n_i}, Y_{n_i}) junto con el supuesto de que $(F_{n+1}, M_{n+1}) = \sum_{i=1}^{Z_n} (X_{n_i}, Y_{n_i})$ implica que la sucesión de variables aleatorias $\{Z_n\}$ es una Cadena de Markov con espacio de estados dado por los enteros no negativos. Además como $\mathcal{L}(0,0)=0$, el estado 0 es absorbente.

Las medias de reproducción de hembras y machos por una unidad de casamiento serán denotadas por μ_1 y μ_2 respectivamente.

De ahora en adelante supondremos que $P(Z_{n+1}=j|Z_n=j) < 1$, lo que significa que con probabilidad menor que uno, en dos generaciones sucesivas, las unidades de casamiento permanecerán iguales.

DEFINICION 2.1.4. Las probabilidades de extinción se definen como $q_j = P(Z_n=0 \text{ para alguna } n|Z_0=j)$ $j=1,2,\ldots$; y las tasas medias de crecimiento se definen como $r_j = j^{-1}E(Z_{n+1}|Z_n=j)$.

La interpretación de la tasa media de crecimiento es la siguiente: Dado que estamos en cierta generación, con j unidades de casamiento, la tasa media de crecimiento es el promedio de unidades de casamiento que habrá en la proxima generación, por unidad de casamiento.

Observemos que $P(Z_{n+1}\leq k|Z_n=j) = P(\mathscr{L}(\sum_{i=1}^{j}(X_{n_i},Y_{n_i}))\leq k) \geq P(\mathscr{L}(\sum_{i=1}^{j+1}(X_{n_i},Y_{n_i}))\leq k) = P(Z_{n+1}\leq k|Z_n=j+1)$, por lo que $\{Z_n\}$ es una cadena de Markov estocasticamente monótona (C.M.E.M.).

Recordemos que para dos variables aleatorias X y Y decimos que X es estocasticamente menor que Y si $P(X\leq j) \geq P(Y\leq j)$ $\forall$ j, y se denota como $X \overset{d}{\leq} Y$.

PROPOSICION 2.1.2 Sean $\{Z_n^{\prime}\}$ y $\{Z_n^{\prime\prime}\}$ dos P.R.G.W.B. dados. Si las distribuciones iniciales son tales que $Z_0^{\prime} \overset{d}{\leq} Z_0^{\prime\prime}$, las funciones de casamiento cumplen con $\mathscr{L}^{\prime}(x,y) \leq \mathscr{L}^{\prime\prime}(x,y)$ $\forall$ x,y y las distribuciones de descendientes, satisfacen $E(f(X_{n_i}^{\prime},Y_{n_i}^{\prime})) \leq E(f(X_{n_i}^{\prime\prime},Y_{n_i}^{\prime\prime}))$ donde f es una función creciente, entonces $P(\text{extinción}|Z_0^{\prime}) \geq P(\text{extinción}|Z_0^{\prime\prime})$.

LEMA 2.1.1 Para un P.R.G.W.B. superaditivo, el límite $\lim_{j\to\infty} j^{-1}E(Z_{n+1}|Z_n=j) = \sup_{j>0} j^{-1}E(Z_{n+1}|Z_n=j) = r$ existe

Al parámetro r se le conoce como la tasa de crecimiento, y tiene la siguiente interpretación: Dado que estamos en cierta generación, r es el promedio de unidades de casamiento que habrá en la siguiente generación, por unidad de casamiento. Observemos que dicho límite depende de μ_1 y μ_2 como lo muestra el siguiente resultado.

LEMA 2.1.2 Para una sucesión $\{(X_i,Y_i),\ i=1,2,\dots\}$ de variables aleatorias bivariadas independientes e identicamente distribuidas (v.a.b.i.i.d.) con esperanzas finitas y función de casamiento $\mathscr{L}$ superaditiva

$$j^{-1}\mathscr{L}\left(\sum_{i=1}^{j} X_i,\ \sum_{i=1}^{j} Y_i\right) \xrightarrow[j\to\infty]{c.n.} \lim_{j\to\infty} j^{-1}\mathscr{L}(jE(X),jE(Y)) = r$$

OBSERVACION En el modelo M1 se tiene $r = \min(\mu_1, \mu_2)$, y en los modelos M2 y M3 se tiene $r = \mu_1$.

En M1, el valor es $r = \lim_{j\to\infty} j^{-1} \min(j\,\mu_1, j\,\mu_2) =$

$\lim_{j\to\infty} j^{-1}\, j \min(\mu_1,\mu_2) = \min(\mu_1,\mu_2)$

En M2 se tiene, $r = \lim_{j\to\infty} j^{-1}\, j\mu_1 \min(1, j\mu_2) =$

$\lim_{j\to\infty} \mu_1 \min(1, j\mu_2) = \mu_1$, ya que $j\mu_2$ crece cuando j crece.

En M3, el valor es $r = \lim j^{-1}\, j\,\mu_1 = \mu_1$

Con esta observación se ha encontrado la tasa de crecimiento r, para los modelos M1, M2 y M3; y de esta manera es posible ver, para estos modelos, como r depende de μ_1 y μ_2.

Daley (1968) demostró que para las funciones de casamiento M1 y M2, $q_j = 1\ \forall\ j=1,2,\ldots$ si y solo si $r \le 1$. El siguiente resultado trata de la extinción para todo modelo que es superaditivo.

TEOREMA 2.1.1 Para un P.R.G.W.B. con funciones de casamiento superaditivas $q_j = 1\ \forall\ j=1,2,\ldots$ si y solo si $r \le 1$.

Este teorema es fundamental en los P.R.G.W.B., ya que nos dice bajo que condiciones el proceso tiende a extinguirse o tiende a crecer. Además es interesante observar la analogía que existe con los procesos ramificados simples y con los procesos de Galton-Watson bitipos, en cuanto a las condiciones de extinción del proceso. Esto es, en los tres procesos, un parámetro es el que interviene de manera semejante, en los ramificados univariados es la media de reproducción, en los ramificados bitipos es el valor propio mayor de la matriz de medias de reproducción (veáse Karlin (1966), página 301), y en los P.R.G.W.B. es la tasa de crecimiento.

2.2 A P R O X I M A C I O N E S

En esta sección serán vistos ciertos resultados que tienen que ver con aproximaciones de las Cadenas de Markov. Estas aproximaciones servirán en los P.R.G.W.B. para darnos una visón de como son los valores de las probabilidades q_j. En los procesos de ramificación simples (bitipos) se usaba la f.g.p. (f.g.p.'s) para saber el valor de la probabilidad (las probabilidades) de extinción, gracias a la propiedad aditiva que estos procesos tienen (veáse Bagley (1986) En los P.R.G.W.B. no es posible usar el mismo mecanismo que los procesos de ramificación simples o bitipos, pero sin embargo estas aproximaciones pueden ser bastantes útiles.

DEFINICION 2.2.1 Sea $P = (p_{ij})$ la matriz de probabilidades de transición a un paso, de una cadena de Markov, donde el espacio de estados son los enteros no negativos y el 0 es un estado absorbente, la matriz $P_{(k)}$ definida como

$$P_{(k)} = \left[\begin{array}{c:c} 1 & \underline{0} \\ \hdashline P_{0(k)} & Q_{(k)} \end{array}\right]$$

donde $P_{0(k)} = (p_{10}, p_{20}, \ldots, p_{k0})^T$ y $Q_{(k)}$ es la matriz de transición para los estados $1, 2, \ldots, k$, se llama la truncación noroeste de P para k.

Denotaremos por $p_{ij}^{(n)}$ a la probabilidad de pasar del estado i al estado j en n pasos, y $p_{ij(k)}^{(n)}$ a la probabilidad de pasar del estado i al estado j en n pasos pero sin haber pasado, en ningún momento, por alguno de los estados $k+1$, $k+2$, ...

Es posible observar que $\lim_{k \to \infty} p_{ij(k)}^{(n)} = p_{ij}^{(n)}$

ya que $p_{ij}^{(n)} = \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \cdots \sum_{i_{n-1}=1}^{\infty} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$ y

$$p_{ij(k)}^{(n)} = \sum_{i_1=1}^{k} \sum_{i_2=1}^{k} \cdots \sum_{i_{n-1}=1}^{k} p_{ii_1} p_{i_1 i_2} \cdots p_{i_{n-1} j}$$

Denotaremos por $q_{j(k)}$ a la probabilidad de extinción dado que se comenzó en el estado j del P.R.G.W.B. truncado con k, y por $q_{(k)} = (q_{1(k)}, q_{2(k)}, \ldots, q_{k(k)})^T$.

PROPOSICION 2.2.1 $\quad q_{(k)} = (I - Q_{(k)})^{-1} P_{o(k)}$

Lo anterior sale del hecho que $q_{i(k)} = p_{jo} + \sum_{i=1}^{k} p_{ji} q_{i(k)}$

$j=1,2,\ldots,k$ y $(I - Q_{(k)})^{-1} = I + Q_{(k)} + Q_{(k)}^2 + \ldots$

Recordemos que una matriz no negativa $S = (s_{ij})$ es monótona si para cada $i \le l$ tenemos $\sum_{j=1}^{k} s_{ij} \ge \sum_{j=1}^{k} s_{lj}$, $k=1,2,\ldots$.

DEFINICION 2.2.2 Sean $S = (s_{ij})$ y $T = (t_{ij})$ matrices no negativas de las mismas dimensiones, S es monotonamente menor que T si para cada i se cumple $\sum_{j=1}^{k} s_{ij} \geq \sum_{j=1}^{k} t_{ij}$, y se denota como $S \overset{m}{\leq} T$

OBSERVACIONES

1) La matriz de transición P y sus truncaciones $P_{(k)}$ $k=1,2,\ldots$, son monótonas.

2) El producto de matrices monótonas es monótona.

3) $q_i \geq q_j$ para $i \leq j$ y $q_{i(k)} \leq q_{j(k)}$ $i \leq j \leq k$

4) $q_{i(k)}$ es monotonamente creciente con respecto a k, por lo que $q_{i(k)} \leq q_i$

5) $q_{i+j} \leq q_i q_j$ $i,j=0,1,2,\ldots$

LEMA 2.2.1 Si el proceso estocástico $\{Z_n\}$ valuado en los enteros es estocasticamente monótono, entoces $q_i - q_{i(k)} \leq (1 - q_{i(k)})q_{k+1}$ para todos los enteros i,k tal que $i \leq k$.

Con el siguiente resultado, podemos darnos una idea de los valores que tienen las probabilidades de extinción en los P.R.G.W.B.

TEOREMA 2.2.1 Para cadenas de Markov superaditivas tenemos

$$q_{i(i+j-1)} \leq q_i \leq \min\left\{1, \frac{q_{i(i+j-1)}}{1-q_{j(i+j-1)}}\right\} \quad \text{para } i,j=1,2,\ldots$$

Los subíndices $_1$, $_2$ y $_3$ serán usados para los modelos M1, M2 y M3 respectivamente.

PROPOSICION 2.2.2

1) $\mathcal{L}_1(x,y) \leq \mathcal{L}_2(x,y) \leq \mathcal{L}_3(x,y)$ $\forall$ $x,y \in [0,\infty)$.

2) Si cada unidad de casamiento reproduce independientemente de las otras y se tienen distribuciones de descendientes idénticas, entonces $T_1 \overset{m}{\leq} T_2 \overset{m}{\leq} T_3$, donde T_i es la matriz de transición para los modelos $i=1,2,3$, y de aquí se sigue que $T_1{}^n \overset{m}{\leq} T_2{}^n \overset{m}{\leq} T_3{}^n$, $n=1,2,\ldots$

3) Cuando $n\to\infty$ la probabilidad de extinción usando M1 es menor o igual que la probabilidad de extinción usando M2, la cual es menor o igual que la probabilidad de extinción utilizando M3.

4) $r_{j1} \leq r_{j2} \leq r_{j3}$.

Los resultados anteriores también son válidos para las truncaciones noroestes.

3.1 SUPERMARTINGALAS ASOCIADAS

Para concluir este capitulo presentaremos en esta sección resultados sobre supermartingalas asociadas a los P.R.G.W.B.. Hemos incluido estos resultados y sus demostraciones ya que aunque fáciles, no fueron encontrados en la literatura.

Para P.R.G.W.B. con función de casamiento natural, se tiene el resultado siguiente

PROPOSICION 2.3.1 En un P.R.G.W.B., con función de casamiento natural, $\{\frac{Z_n}{\mu_1^n} ; \mathcal{F}_n , n \geq 1\}$ es una supermartingala, donde $\mathcal{F}_n = \sigma\{(F_0,M_0),(F_1,M_1), \ldots ,(F_n,M_n)\}$es la σ-álgebra generada por $(F_0,M_0),\ldots,(F_n,M_n)$.

Demostración:

$$E\left[\frac{Z_{n+1}}{\mu_1^{n+1}} \middle| \mathcal{F}_n \right] = \frac{1}{\mu_1^{n+1}} E(\mathcal{L}(F_{n+1},M_{n+1}) | \mathcal{F}_n) = \frac{1}{\mu_1^{n+1}}$$

$$E(\mathcal{L}(\sum_{i=1}^{Z_n} X_{n_i}, \sum_{i=1}^{Z_n} Y_{n_i}) | \mathcal{F}_n) \leq \frac{1}{\mu_1^{n+1}} E\left[\sum_{i=1}^{Z_n} X_{n_i} \middle| \mathcal{F}_n \right] =$$

$$\frac{1}{\mu_1^{n+1}} \sum_{i=1}^{Z_n} E(X_{n_i}) = \frac{Z_n}{\mu_1^n} \qquad \blacksquare$$

Recordemos que las funciones de casamiento M1, M2 y M3 son naturales por lo que para estos modelos $\{\frac{Z_n}{\mu_1{}^n} ; \mathcal{F}_n , n \geq 1\}$ es una supermartingala.

Obsérvese que para el modelo M1, $\{\frac{Z_n}{\mu_2{}^n} ; \mathcal{F}_n , n \geq 1\}$, es también una supermartingala. También obsérvese que para el modelo M3 la sucesión $\{\frac{Z_n}{\mu_1{}^n} ; \mathcal{F}_n , n \geq 1\}$, es una martingala. En general para cualquier proceso con función de casamiento que cumple con la propiedad de superaditividad se tiene el siguiente resultado.

PROPOSICION 2.3.2 En un P.R.G.W.B., con función de casamiento superaditiva, la sucesión $\{\frac{Z_n}{\nu^n} ; \mathcal{F}_n , n \geq 1\}$ es una submartingala, donde $\nu = E(\mathcal{L}(X,Y))$.

Demostración:

$$E\left[\frac{Z_{n+1}}{\nu^{n+1}} \middle| \mathcal{F}_n \right] = \frac{1}{\nu^{n+1}} E(\mathcal{L}(F_{n+1}, M_{n+1}) \mid \mathcal{F}_n) = \frac{1}{\nu^{n+1}}$$

$$\frac{1}{\nu^{n+1}} E\left[\mathcal{L}(\sum_{i=1}^{Z_n} (X_{n_i}, Y_{n_i}) \middle| \mathcal{F}_n \right] \geq \frac{1}{\nu^{n+1}} E\left[\sum_{i=1}^{Z_n} \mathcal{L}(X_{n_i}, Y_{n_i}) \middle| \mathcal{F}_n \right] =$$

$$\frac{1}{\nu^{n+1}} \sum_{i=1}^{Z_n} E(\mathcal{L}(X,Y)) = \frac{1}{\nu^{n+1}} Z_n E(\mathcal{L}(X,Y)) = \frac{Z_n}{\nu^n} \quad \blacksquare$$

CAPITULO 3

ESTIMACION EN LOS PROCESOS RAMIFICADOS DE GALTON-WATSON BISEXUALES

En este capítulo nos ocuparemos, en primer lugar, de estimar por máxima verosimilitud el vector de medias $\underline{\mu}$, suponiendo primero que la distribución de descendientes es serie de potencias bivariada y luego que es Poisson bivariada. En seguida, para varias distribuciones bivariadas, se estiman por el mismo método, las varianzas marginales y la covarianza. Posteriormente se estima por el método de mínimos cuadrados condicional el vector de medias, sin necesidad de suponer alguna distribución de descendientes.

En el teorema 2.1.1, se puede ver la gran importancia que tiene estimar el parámetro r, en un P.R.G.W.B.. Más nos preocuparemos primero por estimar el vector $\underline{\mu}$, ya que no se tiene una expresión explícita en general para r, y como se vió en el capítulo 2 §1, r depende de las entradas de $\underline{\mu}$, y de esta manera se podrá dar una estimación, que aunque no explícita, de la tasa de crecimiento.

Cabe mencionar que los resultados que se presentan en este capítulo no fueron encontrados en la literatura. Estos son análogos al caso de procesos de Galton-Watson simples.

3.1 MAXIMA VEROSIMILITUD

En esta sección se consideran dos casos de distribuciones de descendientes: Cuando $(X_{k_i}, Y_{k_i}) \sim$ Serie de Potencias Bivariada (θ_1, θ_2, g) y cuando $(X_{k_i}, Y_{k_i}) \sim$ Poisson Bivariada $(\lambda_1, \lambda_2, \lambda_3)$.

Supongamos que tenemos una muestra de un P.R.G.W.B. de la siguiente forma:

$Z_0, Z_1, \ldots, Z_n$ y $(F_0, M_0), (F_1, M_1), \ldots, (F_n, M_n)$; tal que $Z_0 = 1$ y $(F_1, M_1) = (1,1)$

La probabilidad conjunta de $(F_0, M_0), (F_1, M_1), \ldots, (F_n, M_n)$ se puede expresar como:

$$P((F_0, M_0)=(f_0, m_0), (F_1, M_1)=(f_1, m_1), \ \ldots \ , P(F_n, M_n)=(f_n, m_n)) =$$

$$P((F_0, M_0)=(f_0, m_0)) \quad P((F_1, M_1)=(f_1, m_1) \,|(F_0, M_0)=(f_0, m_0)) \quad \ldots$$

$$P((F_n, M_n)=(f_n, m_n) \,|(F_{n-1}, M_{n-1})=(f_{n-1}, m_{n-1})).$$

Ahora $P((F_k, M_k)=(f_k, m_k) \,|(F_{k-1}, M_{k-1})=(f_{k-1}, m_{k-1})) =$

$P(\sum_{i=1}^{z_{k-1}} (X_{k-1_i}, Y_{k-1_i}) = (f_k, m_k))$, ya que conociendo (f_k, m_k) es posible conocer z_k, y además $(F_k, M_k) = \sum_{i=1}^{z_{k-1}} (X_{k-1_i}, Y_{k-1_i})$

Por lo tanto la distribución conjunta queda como:

$$P\left(\sum_{i=1}^{\vartheta_0} (X_{o_i}, Y_{o_i}) = (f_1, m_1)\right) \; P\left(\sum_{i=1}^{\vartheta_1} (X_{1_i}, Y_{1_i}) = (f_2, m_2)\right)$$

$$\ldots P\left(\sum_{i=1}^{\vartheta_{n-1}} (X_{n-1_i}, Y_{n-1_i}) = (f_n, m_n)\right) \qquad (3.1)$$

3.1.1. DISTRIBUCION SERIE DE POTENCIAS BIVARIADA

En esta sección se hará la estimación con el supuesto de que la distribución de descendientes es serie de potencias bivariada (θ_1, θ_2, g).

PROPOSICION 3.1.1. En un P.R.G.W.B. con distribución de descendientes Serie de Potencias Bivariada (θ_1, θ_2, g), con $Z_o = 1$ y corr(X,Y) = coeficiente de correlación de X y Y $\neq \pm 1$, el estimador de máxima verosimilitud de $\underline{\mu}$ es

$$\hat{\underline{\mu}} = \left(\frac{\sum_{i=1}^{n} F_i}{\sum_{i=0}^{n-1} Z_i} , \frac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i} \right) \qquad (3.2)$$

Demostración:

Utilizando la expresión (3.1) y la Proposición 1.1.2, la función de verosimilitud tiene la siguiente expresión:

$$L = cte\ \theta_1^{F_1}\theta_2^{M_1}\ (g(\theta_1,\theta_2)^{-1})\ \theta_1^{F_2}\theta_2^{M_2}\ (g(\theta_1,\theta_2)^{-Z_1})\ \ldots$$

$\theta_1^{F_n}\theta_2^{M_n}\ (g(\theta_1,\theta_2)^{-Z_{n-1}})$. De aquí

$$\log L = \log(cte) + \log(\theta_1)\sum_{i=1}^{n} F_i + \log(\theta_2)\sum_{i=1}^{n} M_i - \log(g(\theta_1,\theta_2))\sum_{i=0}^{n-1} Z_i$$

De donde $\displaystyle \frac{\partial}{\partial\mu_1}\log L = \left(\frac{1}{\theta_1}\sum_{i=1}^{n} F_i - \frac{1}{g(\underline{\theta})}\frac{\partial g(\underline{\theta})}{\partial\theta_1}\sum_{i=0}^{n-1} Z_i\right)\left(\frac{\partial\theta_1}{\partial\mu_1}\right) +$

$$\left(\frac{1}{\theta_2}\sum_{i=1}^{n} M_i - \frac{1}{g(\underline{\theta})}\frac{\partial g(\underline{\theta})}{\partial\theta_2}\sum_{i=0}^{n-1} Z_i\right)\left(\frac{\partial\theta_2}{\partial\mu_1}\right) = 0 \quad y$$

$$\frac{\partial}{\partial\mu_2}\log L = \left(\frac{1}{\theta_1}\sum_{i=1}^{n} F_i - \frac{1}{g(\underline{\theta})}\frac{\partial g(\underline{\theta})}{\partial\theta_1}\sum_{i=0}^{n-1} Z_i\right)\left(\frac{\partial\theta_1}{\partial\mu_2}\right) +$$

$$\left(\frac{1}{\theta_2}\sum_{i=1}^{n} M_i - \frac{1}{g(\underline{\theta})}\frac{\partial g(\underline{\theta})}{\partial\theta_2}\sum_{i=0}^{n-1} Z_i\right)\left(\frac{\partial\theta_2}{\partial\mu_2}\right) = 0$$

Aplicando la Proposición 1.1.3 y la Proposición 1.1.6, las ecuaciones quedan de la siguiente manera:

$$\left(\sum_{i=1}^{n} F_i - \mu_1\sum_{i=0}^{n-1} Z_i\right)\frac{1}{\sigma_1^2} + \left(\sum_{i=1}^{n} M_i - \mu_2\sum_{i=0}^{n-1} Z_i\right)\left(-\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}\right) = 0 \qquad (3.3)$$

$$\left(\sum_{i=1}^{n} M_i - \mu_2\sum_{i=0}^{n-1} Z_i\right)\frac{1}{\sigma_2^2} + \left(\sum_{i=1}^{n} F_i - \mu_1\sum_{i=0}^{n-1} Z_i\right)\left(-\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}\right) = 0 \qquad (3.4)$$

De la expresión (3.3) obtenemos

$$\sum_{i=1}^{n} F_i - \frac{\sigma_{12}}{\sigma_2^2}\sum_{i=1}^{n} M_i + \mu_2\frac{\sigma_{12}}{\sigma_2^2}\sum_{i=0}^{n-1} Z_i = \mu_1\sum_{i=0}^{n-1} Z_i \qquad (3.5)$$

Sustituyendo en la ecuación (3.4) la expresión (3.5) y simplificando resulta

$$-\frac{\sigma_{12}^{2}}{\sigma_1^{2}\sigma_2^{2}}\sum_{i=1}^{n} M_i + \frac{\sigma_{12}^{2}}{\sigma_1^{2}\sigma_2^{2}}\mu_2\sum_{i=0}^{n-1} Z_i + \sum_{i=1}^{n} M_i - \mu_2\sum_{i=0}^{n-1} Z_i = 0$$

De donde $\left(1 - \frac{\sigma_{12}^{2}}{\sigma_1^{2}\sigma_2^{2}}\right)\left(\sum_{i=1}^{n} M_i - \mu_2\sum_{i=0}^{n-1} Z_i\right) = 0$

Por lo que $\hat{\mu}_2 = \dfrac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i}$

Sustituyendo en (3.5), obtenemos:

$$\sum_{i=1}^{n} F_i - \frac{\sigma_{12}}{\sigma_2^{2}}\sum_{i=1}^{n} M_i + \frac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i}\frac{\sigma_{12}}{\sigma_2^{2}}\sum_{i=0}^{n-1} Z_i = \mu_1\sum_{i=0}^{n-1} Z_i$$

Por lo tanto $\hat{\mu}_1 = \dfrac{\sum_{i=1}^{n} F_i}{\sum_{i=0}^{n-1} Z_i}$

OBSERVACION Los estimadores anteriores son ciertos si $\left(1 - \frac{\sigma_{12}^{2}}{\sigma_1^{2}\sigma_2^{2}}\right) \neq 0$. Esto es $Corr(X_{k-1_i}, Y_{k-1_i})$ = coeficiente de correlación de X_{k-1_i} y $Y_{k-1_i} \neq \pm 1$. En otras palabras no debe de existir dependencia lineal entre las variables X_{k-1_i} y Y_{k-1_i}.

El estimador resultante tiene la siguiente interpretación. El estimador de la media del número de hembras por unidad de casamiento, resultó ser, el número de hembras producidas en las n generaciones entre el número de parejas casadas que produjeron las hembras durante la n generaciones. La interpretación es análoga para la segunda entrada.

OBSERVACION Si X ~ serie de potencias (θ_1, g_1) y Y ~ serie de potencias (θ_2, g_2), y además son independientes, entonces

$$\hat{\mu}_1 = \hat{\mu}_2 = \hat{\mu} = \frac{\sum_{i=1}^{n} (F_i + M_i)}{2 \sum_{i=0}^{n-1} Z_i}$$

Lo anterior es verdad ya que

$$L = \text{cte} \quad \theta^{F_1} \theta^{M_1} (g(\theta) g(\theta))^{-1} \quad \theta^{F_2} \theta^{M_2} (g(\theta) g(\theta))^{-Z_1} \quad \ldots$$

$$\theta^{F_n} \theta^{M_n} (g(\theta) g(\theta))^{-Z_{n-1}}$$

Por lo que

$$\log L = \log(\text{cte}) + \log(\theta) \sum_{i=1}^{n} (F_i + M_i) - 2\log(g(\theta)) \sum_{i=0}^{n-1} Z_i$$

De aqui

$$\frac{\partial \log L}{\partial \mu} = \frac{1}{\theta} \frac{\partial \theta}{\partial \mu} \sum_{i=1}^{n} (F_i + M_i) - \frac{2 \sum_{i=0}^{n-1} Z_i}{g(\theta)} \frac{\partial g(\theta)}{\partial \theta} \frac{\partial \theta}{\partial \mu} = 0$$

Aplicando los resultados para la media y la varianza, en términos de la derivada de $g(\cdot)$ y de la derivada de θ, respectivamente, de la distribución serie de potencias univariada (Basawa y Prakasa Rao (1980)), obtenemos

$$\left(\sum_{i=1}^{n} (F_i+M_i) - 2\sum_{i=0}^{n-1} Z_i \; \mu \right) \left(\frac{1}{\sigma^2}\right) = 0$$

De esta ecuación, el estimador de μ se obtiene como

$$\hat{\mu} = \frac{\sum_{i=1}^{n} (F_i+M_i)}{2 \sum_{i=0}^{n-1} Z_i}$$

Para la interpretación de esta media, será necesario tomar en cuenta toda la información, esto es, dicho estimador resultó ser el número de individuos reproducidos , entre dos veces las unidades de casamiento que produjeron estos individuos.

Obsérvese que cuando $\theta = \theta_1 = \theta_2$ y $g_1 \neq g_2$ o $\theta = \theta_1 = \theta_2$ y X y Y son dependientes, entonces no es posible estimar $\underline{\mu}$.

3.1.2. DISTRIBUCION POISSON BIVARIADA

Supondremos en esta sección que la distribución de descendientes es la Poisson bivariada $(\lambda_1, \lambda_2, \lambda_3)$. Recordemos que esta no es un caso incluido en la distribución serie de potencias bivariada, por lo que la estimación de $\underline{\mu}$ se realizará sin usar los resultados de media en función de la derivada de $g(\cdot,\cdot)$.

PROPOSICION 3.1.2. En un P.R.G.W.B. con distribución de descendientes Poisson bivariada $(\lambda_1,\lambda_2,\lambda_3)$ y que se encuentra en el conjunto de no extinción, el estimador de máxima verosimilitud de $\underline{\mu}$ es

$$\underline{\tilde{\mu}} = \left(\frac{1}{n} \sum_{i=1}^{n} \frac{F_i}{Z_{i-1}} , \frac{1}{n} \sum_{i=1}^{n} \frac{M_i}{Z_{i-1}} \right)$$

Demostración:

La función conjunta de (F_1,M_1), ... ,(F_n,M_n) puede expresarse como

$$\prod_{i=1}^{n} P\left(F_i=f_i, M_i=m_i, z_{i-1}\lambda_1, z_{i-1}\lambda_2, z_{i-1}\lambda_3\right) \quad , \qquad \text{donde}$$

$P\left(F_i=f_i, M_i=m_i, z_{i-1}\lambda_1, z_{i-1}\lambda_2, z_{i-1}\lambda_3\right)$ = la probabilidad de que en la i-ésima generación la reproducción sea de f_i hembras y m_i machos, siendo la distribución de (F_i,M_i) Poisson bivariada $(z_{i-1}\lambda_1, z_{i-1}\lambda_2, z_{i-1}\lambda_3)$. Dicha probabilidad será denotada solamente por $P(f_i,m_i)$.

Así tenemos que $\log L = \sum_{i=1}^{n} \log P(F_i,M_i)$

Por lo que $\dfrac{\partial \log L}{\partial \lambda_1} = \sum_{i=1}^{n} \dfrac{1}{P(F_i,M_i)} \dfrac{\partial}{\partial \lambda_1} P(F_i,M_i) = 0$

$$\frac{\partial \log L}{\partial \lambda_2} = \sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} \frac{\partial}{\partial \lambda_2} P(F_i,M_i) = 0$$

$$\frac{\partial \log L}{\partial \lambda_3} = \sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} \frac{\partial}{\partial \lambda_3} P(F_i,M_i) = 0$$

Aplicando la proposición 1.2.3, las ecuaciones anteriores se pueden expresar como

$$\sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} (P(F_i-1,M_i-1) - P(F_i,M_i)) = 0$$

$$\sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} (P(F_i-1,M_i) - P(F_i,M_i)) = 0$$

$$\sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} (P(F_i,M_i-1)) - P(F_i,M_i) = 0$$

De la primera ecuación se tiene:

$$\frac{1}{n} \sum_{i=1}^{n} \frac{P(F_i-1,M_i-1)}{P(F_i,M_i)} = 1 \qquad (3.6)$$

Aplicando la proposición 1.2.2, la segunda y tercera ecuación quedan como

$$\sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} \frac{F_i P(F_i,M_i) - Z_{i-1}\lambda_1 P(F_{i-1},M_{i-1})}{Z_{i-1}\lambda_2} = n \qquad (3.7)$$

$$\sum_{i=1}^{n} \frac{1}{P(F_i,M_i)} \frac{M_i P(F_i,M_i) - Z_{i-1}\lambda_2 P(F_{i-1},M_{i-1})}{Z_{i-1}\lambda_3} = n \qquad (3.8)$$

Sustituyendo (3.6) en (3.7) y (3.8) y despejando, resultan los siguientes estimadores

$$\hat{\lambda}_2 + \hat{\lambda}_1 = \tilde{\mu}_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{F_i}{Z_{i-1}}$$

$$\hat{\lambda}_3 + \hat{\lambda}_1 = \tilde{\mu}_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{M_i}{Z_{i-1}}$$

Obsérvese que es necesario pedir para estos estimadores que el proceso se encuentre fuera del conjunto de extinción, o sea $Z_{i-1} \neq 0$ para toda $i = 1,\ldots,n$. Cuando $Z_k = 0$ entonces $Z_j = 0$ para $j > k$. Supongamos que la muestra tenga la siguiente forma $Z_1,\ldots,Z_{k-1}$ distintos de cero y $Z_k,\ldots,Z_n$ iguales a cero. Entonces los estimadores quedarán como

$$\tilde{\mu}_1 = \frac{1}{n} \sum_{i=1}^{k} \frac{F_i}{Z_{i-1}} \quad \text{y} \quad \tilde{\mu}_2 = \frac{1}{n} \sum_{i=1}^{k} \frac{M_i}{Z_{i-1}}$$

Así la estimación de μ_1 resultó ser el promedio de las n generaciones del número de hembras producidas, entre el número de parejas casadas que produjeron a dichas hembras (este cociente es por generación). Para $\tilde{\mu}_2$, la explicación es análoga.

En un proceso de ramificación de Galton-Watson simple, el estimador de máxima verosimilitud de μ, cuando se supone distribución de descendientes serie de potencias univariada, es análogo al estimador de máxima verosimilitud de cada entrada de $\underline{\mu}$ en un P.R.G.W.B. suponiendo distribución de descendientes serie de potencias bivariada (véase Basawa y Prakasa Rao (1980), pág. 22).

3.1.3 ESTIMACION DE OTROS PARAMATROS

Para la estimación de $\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$, podemos aprovechar el principio de invarianza del estimador de máxima verosimilitud. Así por ejemplo

1) Cuando X y Y son independientes, X ~ Poisson y Y ~ Geométrica, entonces $\hat{\sigma}_1^2 = \hat{\mu}_1$ y $\hat{\sigma}_2^2 = \hat{\mu}_2(1 + \hat{\mu}_2)$.

2) Cuando (X,Y) ~ Trinomial (n, p_1, p_2), entonces $\hat{\sigma}_i^2 = \hat{\mu}_i(1 - \frac{\hat{\mu}_i}{n})$ $i = 1,2$ y $\hat{\sigma}_{12} = - \frac{\hat{\mu}_1 \hat{\mu}_2}{n}$.

3) $(X,Y) \sim$ Trinomial Negativa (N,P_1,P_2), se tiene que

$$\hat{\sigma}_i^2 = \hat{\mu}_i(1 + \frac{\hat{\mu}_i}{N}) \quad i=1,2 \quad \text{y} \quad \hat{\sigma}_{12} = \frac{\hat{\mu}_1 \ \hat{\mu}_2}{N} .$$

Sin embargo no siempre es posible estimar dichos parámetros por máxima verosimilitud usando el principio de invarianza. Por ejemplo si $(X,Y) \sim$ Poisson Bivariada $(\lambda_1,\lambda_2,\lambda_3)$, se puede ver que $\hat{\sigma}_i^2 = \hat{\mu}_i$ $i=1,2$. Pero como $\sigma_{12} = \lambda_1$, no es posible estimarlo, ya que no es posible tener información por separado del parámetro λ_1 a través de una muestra del P.R.G.W.B..

La estimación de la tasa de crecimiento r por máxima verosimilitud resulta ser

$$\hat{r} = \lim_{j\to\infty} j^{-1} \mathscr{L}(j\ \hat{\mu}_1, j\ \hat{\mu}_2)$$

Dicho resultado se obtiene al aplicar una vez más el principio de invarianza y el Lema 2.1.2.

De esta manera queda resuelto el problema de estimar el parámetro r.

3.2 MINIMOS CUADRADOS CONDICIONAL

En esta sección se obtendrá otra estimación de $\underline{\mu}$ basada en el método de mínimos cuadrados condicional (M.M.C.C.) el cual no supone el conocimiento de la distribución de descendientes.

En el caso de procesos de Galton-Watson, este método se presenta en Basawa y Prakasa Rao (1980), pág 21. Si se tiene un proceso de ramificación simple $\{ Z_0=1, Z_1, Z_2, \ldots \}$ tal que $E(Z_1) = \mu$ y $Var(Z_1) = \sigma^2$, se puede ver que $E(Z_{n+1}|Z_n) = Z_n \ \mu$ y $Var(Z_{n+1}|Z_n) = Z_n \ \sigma^2$. Es por eso que se sugiere el siguiente modelo de tipo autorregresivo para $\{Z_k\}$, $Z_{n+1} = \mu Z_n + Z_n^{1/2} \ U_{n+1}$ donde se puede verificar que $\{U_k\}$, los residuales, cumplen con las propiedades de los residuales: $E(U_k) = 0$, $Var(U_k) = \sigma^2$ y U_k es incorrelacionado con U_j, para cualquier k y j tal que k≠j. El estimador por el M.M.C.C. de μ, resultó ser

$$\underline{\hat{\mu}} = \frac{\sum_{i=1}^{n} Z_i}{\sum_{i=0}^{n-1} Z_i}$$

, el cual coincide con el de máxima verosimilitud, en este caso.

El M.M.C.C para el caso bivariado fue explicado en la sección 1.3 de este trabajo. Aplicando dicha teoría en nuestro caso tenemos que si

$$Y = \begin{pmatrix} Y_1' \\ \vdots \\ Y_n' \end{pmatrix} = (F,M) = \begin{pmatrix} F_1 & M_1 \\ \vdots & \vdots \\ F_n & M_n \end{pmatrix}$$

, entonces

$$E(Y_i' \mid \mathcal{F}_{i-1}) = E((F_i, M_i) \mid \mathcal{F}_{i-1}) = (Z_{i-1}\mu_1, Z_{i-1}\mu_2) = Z_{i-1}\underline{\mu}$$

Aquí $X = Z = (Z_0\ ,\ \ldots\ ,\ Z_{n-1})$ y $\underline{\alpha} = \underline{\mu}$. De esta manera el estimador del vector de medias por el M.M.C.C., que será denotado por $\underline{\mathring{\mu}}$, resulta ser

$$\underline{\mathring{\mu}} = (Z'Z)^{-1}Z'Y$$

Esto es

$$\underline{\mathring{\mu}} = (\mathring{\mu}_1, \mathring{\mu}_2) = \left(\frac{\sum_{i=1}^{n} F_i Z_{i-1}}{\sum_{i=1}^{n} Z_{i-1}^{2}} \ ,\ \frac{\sum_{i=1}^{n} M_i Z_{i-1}}{\sum_{i=1}^{n} Z_{i-1}^{2}} \right)$$

Obsérvese la gran ventaja que tiene este método, al no suponer el conocimiento de una distribución de descendientes, esto es, es un método no paramétrico. Es también importante observar que no suponemos que $(F_0, M_0) = (1,1)$ y $Z_0 = 1$. Sin embargo una desventaja, lo es el hecho de que no es posible estimar por este método la matriz de covarianza Σ.

Veamos ahora que si cambiamos la sucesión $\{\underline{Y}_i\}$ por $\{\underline{Y}_i^{(1)}\} = \left\{ \frac{1}{Z_{i-1}}\ \underline{Y}_i \right\}$, el resultado será diferente. Esto es, para $\underline{Y}_i^{(1)}$,

$E(\underline{Y}_i^{(1)} \mid \mathcal{F}_{i-1}) = E\left(\frac{1}{Z_{i-1}}\ \underline{Y}_i \,\middle|\, \mathcal{F}_{i-1} \right) = (\mu_1, \mu_2)$. Por lo que en este caso

$$Y = \begin{bmatrix} \frac{F_1}{Z_0} & \frac{M_1}{Z_0} \\ \vdots & \vdots \\ \frac{F_n}{Z_{n-1}} & \frac{M_n}{Z_{n-1}} \end{bmatrix} \; ; \; X = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad y \quad \underline{\alpha} = \underline{\mu}$$

De esta manera el estimador de $\underline{\mu}$, que será denotado por $\underline{\tilde{\mu}}$ resulta ser $\underline{\tilde{\mu}} = (\tilde{\mu}_1, \tilde{\mu}_2) = n^{-1}\left(\sum_{i=1}^{n} \frac{F_i}{Z_{i-1}} \, , \, \sum_{i=1}^{n} \frac{M_i}{Z_{i-1}} \right)$, que coincide con el estimador de máxima verosimilitud de $\underline{\mu}$, cuando se uso distribución de descendientes Poisson bivariada. Recordemos que se pedía para este estimador que $Z_i \neq 0$, y cuando $Z_k = 0$ y $Z_i \neq 0$ para $i < k$, se hace la estimación hasta Z_{k-1}.

Otra posible elección de Y es $\underline{Y}_i^{(2)} = \left\{ \frac{1}{Z_{i-1}} \underline{Y}_i \right\}$ de donde se obtiene $E(\underline{Y}_i^{(2)} | \mathcal{F}_{i-1}) = Z_{i-1}^{1/2}\underline{\mu}$. Por lo tanto

$$Y^{(2)} = \begin{bmatrix} \frac{F_1}{Z_0^{1/2}} & \frac{M_1}{Z_0^{1/2}} \\ \vdots & \vdots \\ \frac{F_n}{Z_{n-1}^{1/2}} & \frac{F_n}{Z_{n-1}^{1/2}} \end{bmatrix} \; ; \; X = \begin{bmatrix} Z_0^{1/2} \\ \vdots \\ Z_{n-1}^{1/2} \end{bmatrix} \quad y \quad \underline{\alpha} = \underline{\mu}$$

El estimador resulta ser entonces

$\hat{\underline{\mu}} = (\hat{\mu}_1, \hat{\mu}_2) = \left[\sum_{i=1}^{n} Z_{i-1} \right]^{-1} \left[\sum_{i=1}^{n} F_i \;,\; \sum_{i=1}^{n} M_i \right]$, el cual coincide con el estimador de máxima verosimilitud, cuando se usa distribución de de descendientes serie de potencias bivariada.

Los modelos lineales que se han propuesto para estimar el vector de medias por el M.M.C.C., son

$$(F_i, M_i) = Z_{i-1}\, \underline{\mu} + (U_{i_1}, U_{i_2}) \qquad (3.9)$$

$$(F_i^{(1)}, M_i^{(1)}) = \underline{\mu} + (U_{i_1}^{(1)}, U_{i_2}^{(1)}) \qquad (3.10)$$

$$(F_i^{(2)}, M_i^{(2)}) = Z_{i-1}^{1/2}\, \underline{\mu} + (U_{i_1}^{(2)}, U_{i_2}^{(2)}) \qquad (3.11)$$

El estimador $\mathring{\underline{\mu}}$ resultó de (3.9), $\tilde{\underline{\mu}}$ de (3.10) y $\hat{\underline{\mu}}$ salió de (3.11).

Puede observarse que el tercer modelo y la estimación resultante de $\underline{\mu}$ son análogos respectivamente, al modelo lineal y al estimador resultante de la media en los procesos de ramificación simples.

Es importante recordar que $E((F_i, M_i) | \mathcal{F}_{i-1}) = Z_{i-1}\, \underline{\mu}$, $E((F_i^{(1)}, M_i^{(1)}) | \mathcal{F}_{i-1}) = \underline{\mu}$ y $E((F_i^{(2)}, M_i^{(2)}) | \mathcal{F}_{i-1}) = Z_{i-1}^{1/2}\, \underline{\mu}$, lo que significa que cada uno de los modelos son razonables

Tres estimadores del vector de medias se han obtenido por el M.M.C.C., y los tres han resultado ser razonables, sin embargo antes de tratar de hacer una comparación entre ellos para ver cual es mejor estimador, sería más conveniente ver que modelo cumple con las propiedades de los residuales, por lo que verificaremos que modelos cumplen dichas propiedades de los residuales, las cuales se pueden encontrarse en la Sección 3 del Capítulo 1. En cuanto a la propiedad de la esperanza

$$E[\underline{U}_i] = E[E((U_{i_1},U_{i_2})|\mathcal{F}_{i-1})] = E[E((F_i-Z_{i-1}\mu_1, M_i-Z_{i-1}\mu_2)|\mathcal{F}_{i-1})] = (0,0) .$$

$$E[\underline{U}_i^{(1)}] = E[E((U_{i_1}^{(1)},U_{i_2}^{(1)})|\mathcal{F}_{i-1})] = E[E((Z_{i-1})^{-1}(F_i-Z_{i-1}\mu_1, M_i-Z_{i-1}\mu_2)|\mathcal{F}_{i-1})] = (0,0) .$$

$$E[\underline{U}_i^{(2)}] = E[E((U_{i_1}^{(2)},U_{i_2}^{(2)})|\mathcal{F}_{i-1})] = E[E((Z_{i-1})^{-1/2}(F_i-Z_{i-1}\mu_1, M_i-Z_{i-1}\mu_2)|\mathcal{F}_{i-1})] = (0,0) .$$

Verifiquemos la propiedad de las varianzas

$$Var[U_{i_1}] = E[U_{i_1}^2] = E[E((F_i-Z_{i-1}\mu_1)^2|\mathcal{F}_{i-1})] = \sigma_1^2 E(Z_{i-1}),$$

analogamente $Var[U_{i_2}] = \sigma_2^2 E(Z_{i-1})$.

$$Var[U_{i_1}^{(1)}] = E[(U_{i_1}^{(1)})^2] = E[E((Z_{i-1})^{-2}(F_i-Z_{i-1}\mu_1)^2|\mathcal{F}_{i-1})] =$$

$\sigma_1^2 E((Z_{i-1})^{-1})$, analogamente $Var[U_{i_2}^{(1)}] = \sigma_2^2 E((Z_{i-1})^{-1})$.

$$Var[U_{i_1}^{(2)}] = E[(U_{i_1}^{(2)})^2] = E[E((Z_{i-1})^{-1}(F_i-Z_{i-1}\mu_1)^2|\mathcal{F}_{i-1})] = \sigma_1^2 ,$$

anlogamente $\mathrm{Var}[U_{i_2}^{(2)}] = \sigma_2^2$.

De aquí concluimos que tanto el primer modelo, como el segundo modelo no cumplen con todas las propiedades de los residuales. Verifiquemos que sucede con las correlaciones de los residuales del tercer modelo

$$\mathrm{Cov}[U_{i_1}, U_{i_2}] = E[U_{i_1} U_{i_2}] = E[E((Z_{i-1})^{-1}(F_i - Z_{i-1}\mu_1)(M_i - Z_{i-1}\mu_2) \,|\, \mathcal{F}_{i-1})]$$

$$= \sigma_{12}.$$

$$\mathrm{Cov}[U_{i_1}, U_{j_1}] = E[U_{i_1} U_{j_1}] = E[E((Z_{i-1} Z_{j-1})^{-1/2}(F_i - Z_{i-1}\mu_1)$$

$(F_j - Z_{j-1}\mu_1) \quad \mathcal{F}_{i-1})] = 0$, donde sin pérdida de generalidad

suponemos que $i < j$, analogamente $\mathrm{Cov}[U_{i_2}, U_{j_2}] = 0$.

$$\mathrm{Cov}[U_{i_1}, U_{j_2}] = E[U_{i_1} \quad U_{j_2}] = E[E((Z_{i-1} Z_{j-1})^{-1/2}(F_i - Z_{i-1}\mu_1)$$

$(M_j - Z_{j-1}\mu_2) \,|\, \mathcal{F}_{i-1})] = 0$, donde sin pérdida de generalidad $i < j$.

Con estos resultados sobre la covarianza de los residuales hemos comprobado que efectivamente, el tercer modelo cumple con cada una de las propiedades usuales de los residuales, por lo que el estimador por M.M.C.G. del vector de medias en un P.R.G.W.B. más conveniente es

$$\underset{\sim}{\hat{\mu}} = \left(\sum_{i=0}^{n-1} Z_i\right)^{-1} \left(\sum_{i=1}^{n} F_i, \sum_{i=1}^{n} M_i\right)$$

que coincide con el de máxima verosimilitud suponiendo distribución en serie de potencias bivariada como distribución de descendientes, y cada entrada es análoga al estimador de la media μ en los procesos de ramificación simples, cuando fue utilizado este mismo método.

CAPITULO 4

PROPIEDADES ASINTOTICAS DE LOS ESTIMADORES

En este capítulo desarrollaremos propiedades asintóticas del estimador de medias

$$\underline{\hat{\mu}} = \left[\frac{\sum_{i=1}^{n} F_i}{\sum_{i=0}^{n-1} Z_i} , \frac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i} \right]$$

que resultó ser el de máxima verosimilitud suponiendo distribución de descendientes en serie de potencias bivariadas, así como el dado por el método de mínimos cuadrados condicional.

4.1. DISTRIBUCION ASINTOTICA

En esta sección se estudiará primero la distribución asintótica del producto de un elemento aleatorio por la diferencia $\underline{\hat{\mu}} - \underline{\mu}$, y por otro lado, como una consecuencia de lo primero la distribución asintótica de un factor aleatorio por la diferencia $\hat{r} - r$. Cabe mencionar que $\sqrt{n}\,(\underline{\hat{\mu}} - \underline{\mu})$ no se distribuye asintóticamente normal, para esto obsérvese el siguiente ejemplo:

Supongamos que (X_{k_i}, Y_{k_i}) se distribuye de tal manera que X_{k_i} y Y_{k_i} son independientes y que X_{k_i} se distribuye como una geométrica(p), además la función de casamiento es $\mathcal{L}(x,y) = x$. Obsérvese que $Z_n = \mathcal{L}(F_n, M_n) = F_n$, por lo que $\sum_{i=0}^{n-1} Z_i = \sum_{i=0}^{n-1} F_i$, además $F_n = \sum_{i=1}^{F_{n-1}} X_{n-1_i}$. De esta manera $\{F_n, n \geq 1\}$ es un proceso de ramificación simple con distribución de descendientes geométrica(p). Por lo tanto concluimos que

$$\frac{1}{\sigma_1^{\,2}} \left[\frac{1-\mu_1^{\,n}}{1-\mu_1} \right] (\hat{\mu}_1 - \mu_1) \xrightarrow{\mathcal{D}} t_{(2)}$$

Este resultado se encuentra en Basawa y Prakasa Rao (1980) en la página 17. Y así $\sqrt{n}\,(\hat{\mu}_1 - \mu_1)$ no se distribuye asintóticamente normal, por lo que la distribución conjunta de $\sqrt{n}\,(\hat{\underline{\mu}} - \underline{\mu})$ tampoco es asintóticamente normal.

Es interesante mencionar que cuando se hace inferencia estadística en los procesos de Galton-Watson simples, se sacan propiedades asintóticas del estimador de la media (véase Basawa y Prakasa Rao (1980)). Toda esta teoría ya deducida, no fue posible aplicarla a los P.R.G.W.B., por lo que fue necesario sacar nuevos resultados para los P.R.G.W.B., los cuales no fueron encontrados en la literatura.

4.1.1 ESTIMADOR DEL VECTOR DE MEDIAS

PROPOSICION 4.1.1 Sea un P.R.G.W.B. con función de casamiento natural, tal que se encuentra en el conjunto de no extinción, entonces

$$\begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix} \left[\sum_{i=0}^{n-1} Z_i \right]^{1/2} (\underline{\hat{\mu}} - \underline{\mu}) \xrightarrow{\mathcal{D}} N_2(\underline{0}, \Gamma)$$

donde $\underline{\hat{\mu}} = \left[\frac{\sum_{i=1}^{n} F_i}{\sum_{i=0}^{n-1} Z_i} , \frac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i} \right]$ $\Gamma = \begin{bmatrix} 1 & \frac{\sigma_{12}}{\sigma_1 \sigma_2} \\ \frac{\sigma_{12}}{\sigma_1 \sigma_2} & 1 \end{bmatrix}$

Demostración:

Nótese que $E\left(\frac{Z_n}{\mu_1{}^n} \right) = \frac{1}{\mu_1{}^n} E(E(Z_n \mid \mathcal{F}_{n-1})) \leq \frac{1}{\mu_1{}^n}$

$E(E(\sum_{i=1}^{Z_{n-1}} X_{k-1_i} | \mathcal{F}_{n-1})) = E\left(\frac{Z_{n-1}}{\mu_1{}^{n-1}} \right)$, ya que $\mathcal{L}(\cdot,\cdot)$ es natural

De donde $E\left(\frac{Z_n}{\mu_1{}^n} \right) \leq E(Z_0) < \infty$, ya que F_0 y M_0 los suponemos fijos.

Recordemos que $\{ \frac{Z_n}{\mu_1{}^n} \}$ es una supermartingala, cuando $\mathcal{L}(\cdot,\cdot)$ es función de casamiento natural (veáse Proposición 2.3.1.). Entonces aplicando el teorema 1.4.1, existe una variable aleatoria ξ, integrable tal que $\frac{Z_n}{\mu_1{}^n} \xrightarrow{c.s.} \xi$.

Obsérvese que el P.R.G.W.B. se encuentra en el conjunto de no extinción, esto quiere decir que $Z_n > 0 \ \forall\ n$, y esto implica que $\xi > 0$

Aplicando ahora el Lema de Toeplitz (Lema 1.4.2.), igualando $b_n = \sum_{k=1}^{n} \mu_1{}^{k-1}$, $X_n = \frac{Z_{n-1}}{\mu_1{}^{n-1}}$, podemos observar que $b_n \longrightarrow \infty$, cuando $n \to \infty$, ya que $\mu_1 > 1$ (por $r > 1$) y que $X_n \longrightarrow X = \xi$ finito y positivo. Entonces tenemos que

$$\frac{1}{\sum_{k=1}^{n} \mu_1{}^{k-1}} \sum_{k=1}^{n} \mu_1{}^{k-1} \frac{Z_{k-1}}{\mu_1{}^{k-1}} = \frac{\sum_{k=1}^{n} Z_{k-1}}{\sum_{k=1}^{n} \mu_1{}^{k-1}} \xrightarrow{c.s} \xi \text{ finito y positivo} \quad (4.1)$$

Por otro lado

$$\begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix} \begin{bmatrix} \sum_{i=0}^{n-1} Z_i \end{bmatrix}^{1/2} (\underline{\hat{\mu}} - \underline{\mu}) = \begin{bmatrix} \dfrac{\sum_{i=1}^{n} F_i - \mu_1 \sum_{i=0}^{n-1} Z_i}{\sigma_1 \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2}} \\ \\ \dfrac{\sum_{i=1}^{n} M_i - \mu_2 \sum_{i=0}^{n-1} Z_i}{\sigma_2 \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2}} \end{bmatrix}$$

Pero las F_i's y las M_i's pueden expresarse como:

$F_0 = 1$; $M_0 = 1$

$F_1 = X_1$; $M_1 = Y_1$

$F_2 = X_{1+1} + X_{1+2} + \dots + X_{1+Z_1}$; $M_2 = Y_{1+1} + Y_{1+2} + \dots + Y_{1+Z_1}$

$F_3 = X_{1+Z_1+1} + \dots + X_{1+Z_1+Z_2}$; $M_3 = Y_{1+Z_1+1} + \dots + Y_{1+Z_1+Z_2}$

.

.

.

$F_n = X_{1+Z_1+\dots+Z_{n-2}+1} + \dots + X_{1+Z_1+\dots+Z_{n-2}+Z_{n-1}}$; $M_n =$

$Y_{1+Z_1+\dots+Z_{n-2}+1} + \dots + Y_{1+Z_1+\dots+Z_{n-2}+Z_{n-1}}$ (4.2)

donde las (X_i, Y_i)'s son variables aleatorias bivariadas independientes, con una misma distribución de descendientes.

Por lo que

$$\begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix} \left[\sum_{i=0}^{n-1} Z_i\right]^{1/2} (\underset{\sim}{\hat{\mu}} - \underset{\sim}{\mu}) = \begin{bmatrix} \dfrac{\sum_{i=1}^{Z_0+\dots+Z_{n-1}} X_i - \sum_{i=0}^{n-1} Z_i \; \mu_1}{\sigma_1 \left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}} \\[2ex] \dfrac{\sum_{i=1}^{Z_0+\dots+Z_{n-1}} Y_i - \sum_{i=0}^{n-1} Z_i \; \mu_2}{\sigma_2 \left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}} \end{bmatrix} \quad (4.3)$$

Sean a y b dos números reales cualesquiera, y denótese $R = \sum_{i=o}^{n-1} Z_i$. Obsérvese que

$$a \frac{1}{\sigma_1} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_1-\mu_1) + b \frac{1}{\sigma_2} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_2-\mu_2) =$$

$$\frac{\sum_{i=o}^{R} (\frac{a}{\sigma_1} X_i + \frac{b}{\sigma_2} Y_i) - R (\frac{a}{\sigma_1} \mu_1 + \frac{b}{\sigma_2} \mu_2)}{R^{1/2}} \quad (4.4)$$

Además $E(\frac{a}{\sigma_1} X_i + \frac{b}{\sigma_2} Y_i) = \frac{a}{\sigma_1} \mu_1 + \frac{b}{\sigma_2} \mu_2$ y

$$Var(\frac{a}{\sigma_1} X_i + \frac{b}{\sigma_2} Y_i) = a^2 + b^2 + 2ab \frac{\sigma_{12}}{\sigma_1\sigma_2} = \gamma$$

El Teorema 1.4.3. es posible aplicarlo en (4.4), rescribiendo dicha expresión como

$$\frac{\xi_1 + \xi_2 + \ldots + \xi_{V_n}}{\sqrt{V_n}} \quad \text{donde}$$

$$\xi_i = \frac{a}{\sigma_1} X_i + \frac{b}{\sigma_2} Y_i - (\frac{a}{\sigma_1} \mu_1 + \frac{b}{\sigma_2} \mu_2)$$

$$V_n = Z_o + \ldots + Z_{n-1}$$

$$a_n = \sum_{i=o}^{n} \mu_1^{i-1}$$

Así $\frac{V_n}{a_n} \xrightarrow{c.s.} W = \xi$ y $a_n \longrightarrow \infty$ por (4.1)

Por lo que $\dfrac{\sum_{i=o}^{R} (\frac{a}{\sigma_1} X_i + \frac{b}{\sigma_2} Y_i) - R(\frac{a}{\sigma_1} \mu_1 + \frac{b}{\sigma_2} \mu_2)}{R^{1/2}} \xrightarrow{\mathcal{D}} N(0,\gamma)$

Nótese que

$$\gamma = (a,b) \begin{bmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} \\ \frac{\sigma_{12}}{\sigma_1\sigma_2} & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = (a,b)\, \Gamma \begin{bmatrix} a \\ b \end{bmatrix}$$

De esta manera

$$a \frac{1}{\sigma_1} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_1 - \mu_1) + b \frac{1}{\sigma_2} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_2 - \mu_2) \xrightarrow{\mathcal{D}}$$

$$a X + b Y \sim N(0,\gamma)$$

Y es así como

$$\left(\frac{1}{\sigma_1} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_1 - \mu_1) , \frac{1}{\sigma_2} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_2 - \mu_2) \right) . \xrightarrow{\mathcal{D}}$$

$$(X,Y) \sim N(\underline{0},\Gamma) \qquad \blacksquare$$

OBSERVACION En particular $\frac{1}{\sigma_1} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_1 - \mu_1) \xrightarrow{\mathcal{D}}$

$N(0,1)$ y $\frac{1}{\sigma_2} \left[\sum_{i=o}^{n-1} Z_i\right]^{1/2} (\hat{\mu}_2 - \mu_2) \xrightarrow{\mathcal{D}} N(0,1)$.

Además obsérvese que un P.R.G.W.B. con función de casamiento M1, M2 o M3, cumple con la Proposición 4.1.1..

En términos de la función de verosimilitud es posible reescribir el resultado anterior

COROLARIO 4.1.3. Sea un P.R.G.W.B. tal que tiene función de casamiento natural y distribución de descendientes Serie de Potencias Bivariada, entonces

$$\xi_n^{1/2}(\theta)(\hat{\underline{\mu}} - \underline{\mu}) \xrightarrow{\mathcal{D}} N_2(\underline{0}, \beta(\theta)\Gamma) ,$$

donde $\xi_n(\theta) = \mathrm{diag}\ \{ \sum_{i=1}^{n} E(U_i(\underline{\mu})\ U_i^T(\underline{\mu})\ |\ \mathcal{F}_{n-1}) \} =$

$$\mathrm{diag}\ E\left[\sum_{i=1}^{n} \left(\frac{\partial \ell n L_i}{\partial \underline{\mu}} - \frac{\partial \ell n L_{i-1}}{\partial \underline{\mu}} \right) \left(\frac{\partial \ell n L_i}{\partial \underline{\mu}} - \frac{\partial \ell n L_{i-1}}{\partial \underline{\mu}} \right)^T \middle| \mathcal{F}_{i-1} \right]$$

$$\hat{\underline{\mu}} = \left(\frac{\sum_{i=1}^{n} F_i}{\sum_{i=0}^{n-1} Z_i} , \frac{\sum_{i=1}^{n} M_i}{\sum_{i=0}^{n-1} Z_i} \right) , \quad \beta(\theta) = 1 - \frac{\sigma_{12}^{\ 2}}{\sigma_1^{\ 2}\sigma_2^{\ 2}} \quad y \quad \Gamma = \begin{pmatrix} 1 & \frac{\sigma_{12}}{\sigma_1\sigma_2} \\ \frac{\sigma_{12}}{\sigma_1\sigma_2} & 1 \end{pmatrix}$$

Demostración:

Esto se debe a que

$$U_i(\underline{\mu}) = \frac{\partial \log L_i}{\partial \underline{\mu}} - \frac{\partial \log L_{i-1}}{\partial \underline{\mu}} =$$

$$\begin{bmatrix} \frac{1}{\sigma_1^2}(F_i - \mu_1 Z_{i-1}) - \frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}(M_i - \mu_2 Z_{i-1}) \\ -\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2}(F_i - \mu_1 Z_{i-1}) + \frac{1}{\sigma_2^2}(M_i - \mu_2 Z_{i-1}) \end{bmatrix}$$

Por lo tanto $U_i(\underline{\mu})\ U_i^T(\underline{\mu}) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ donde

$$a_{11} = \frac{1}{\sigma_1^4}(F_i^2 - 2\mu_1 Z_{i-1}F_i + \mu_1^2 Z_{i-1}^2) - 2\frac{\sigma_{12}}{\sigma_1^4\sigma_2^2}$$

$$(F_iM_i - \mu_2 Z_{i-1}F_i - \mu_1 Z_{i-1}M_i + \mu_1\mu_2 Z_{i-1}^2) + \frac{\sigma_{12}}{\sigma_1^4\sigma_1^4}$$

$$(M_i - 2\mu_2 Z_{i-1}M_i + \mu_2^2 Z_{i-1}^2)$$

$$a_{12} = -\frac{\sigma_{12}}{\sigma_1^4\sigma_1^2}(F_i^2 - 2\mu_1 Z_{i-1}F_i + \mu_1^2 Z_{i-1}^2) + \frac{1}{\sigma_1^2\sigma_2^2}$$

$$(F_iM_i - \mu_2 Z_{i-1}F_i - \mu_1 Z_{i-1}M_i + \mu_1\mu_2 Z_{i-1}^2) + \frac{\sigma_{12}^2}{\sigma_1^4\sigma_2^4}$$

$$(F_iM_i - \mu_2 Z_{i-1}F_i - \mu_1 Z_{i-1}M_i + \mu_1\mu_2 Z_{i-1}^2) - \frac{\sigma_{12}}{\sigma_1^2\sigma_2^4}$$

$$(M_i^2 - 2\mu_2 Z_{i-1}M_i + \mu_2^2 Z_{i-1}^2) = a_{21}$$

$$a_{22} = \frac{1}{\sigma_2^4} (M_i^2 - 2\mu_2 Z_{i-1} M_i + \mu_2^2 Z_{i-1}^2) - 2 \frac{\sigma_{12}}{\sigma_1^2 \sigma_2^4}$$

$$(F_i M_i - \mu_1 Z_{i-1} M_i - \mu_2 Z_{i-1} F_i + \mu_1 \mu_2 Z_{i-1}^2) + \frac{\sigma_{12}^2}{\sigma_1^4 \sigma_2^4}$$

$$(F_i^2 - 2\mu_1 Z_{i-1} F_i + \mu_1^2 Z_{i-1}^2)$$

Observemos que:

$$E(F_i^2 \mid \mathcal{F}_{i-1}) = Z_{i-1}\sigma_1^2 + Z_{i-1}^2 \mu_1^2$$

$$E(M_i^2 \mid \mathcal{F}_{i-1}) = Z_{i-1}\sigma_2^2 + Z_{i-1}^2 \mu_2^2$$

$$E(F_i \mid \mathcal{F}_{i-1}) = Z_{i-1}\mu_1 \quad ; \quad E(M_i \mid \mathcal{F}_{i-1}) = Z_{i-1}\mu_2$$

$$E(F_i M_i \mid \mathcal{F}_{i-1}) = Z_{i-1}\sigma_{12} + Z_{i-1}^2 \mu_1 \mu_2$$

Usando dichas expresiones, facilmente se puede llegar a

$$E(a_{11} \mid \mathcal{F}_{i-1}) = \frac{1}{\sigma_1^2} \left[1 - \frac{\sigma_{12}^2}{\sigma_1^2 \sigma_2^2} \right] Z_{i-1}$$

$$E(a_{12} \mid \mathcal{F}_{i-1}) = E(a_{21} \mid \mathcal{F}_{i-1}) = - \frac{\sigma_{12}}{\sigma_1^2 \sigma_2^2} \left[1 - \frac{\sigma_{12}^2}{\sigma_1^2 \sigma_2^2} \right] Z_{i-1}$$

$$E(a_{22} \mid \mathcal{F}_{i-1}) = \frac{1}{\sigma_2^2} \left[1 - \frac{\sigma_{12}^2}{\sigma_1^2 \sigma_2^2} \right] Z_{i-1}$$

$$\therefore \quad E(\sum_{i=1}^{n} U_i(\underline{\mu}) \; U_i^T(\underline{\mu}) \mid \mathcal{F}_{i-1}) =$$

$$\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2} \\ -\frac{\sigma_{12}}{\sigma_1^2\sigma_2^2} & \frac{1}{\sigma_2^2} \end{bmatrix} \left[1 - \frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2} \right] \left(\sum_{i=0}^{n-1} Z_i\right)$$

Por lo tanto $\xi^{1/2}(\theta) = \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix} \left[1 - \frac{\sigma_{12}^2}{\sigma_1^2\sigma_2^2} \right]^{1/2} \left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}$

Y es así como $\xi^{1/2}(\theta)(\underline{\hat{\mu}} - \underline{\mu}) = \beta^{1/2}(\theta) \begin{bmatrix} \frac{1}{\sigma_1} & 0 \\ 0 & \frac{1}{\sigma_2} \end{bmatrix} \left[\sum_{i=0}^{n-1} Z_i \right]^{1/2} (\underline{\hat{\mu}}-\underline{\mu})$

Por lo tanto $\xi^{1/2}(\theta)\ (\underline{\hat{\mu}} - \underline{\mu}) \longrightarrow N_2(0, \beta(\theta)\Gamma)$, por la Proposición anterior. ■

Obsérvese que lo que realmente significa esta última Proposición y Corolario, es que la diferencia $\underline{\hat{\mu}} - \underline{\mu}$, es normalizada por medio de un factor aleatorio. Este factor aleatorio es análogo al que se usa en los procesos de ramificación simples, cuando también se normaliza aleatoriamente la diferencia $\hat{\mu}-\mu$, (veáse Basawa y Prakasa Rao (1980) en la pág. 24). De esta manera, como aplicación a esta normalización en los procesos de Galton-Watson simples, es posible encontrar un intervalo de confianza para el parámetro μ. En los P.R.G.W.B., con el resultado de la Proposición 4.1.1., existirá la posibilidad de encontrar una región de confianza para el vector de medias $\underline{\mu}$, con matriz de covarianza Σ conocida. Esto porque la matriz Γ es definida positiva, y tiene sentido hablar de la raíz cuadrada única de Γ, y de esta manera al premultiplicar la expresión (4.3) por la raíz cuadrada de Γ, se distribuirá asintóticamente $N(\underline{0},I)$, y de aquí será posible encontrar una región de confianza para el vector de medias.

4.1.2. ESTIMADOR DE LA TASA DE CRECIMIENTO

COROLARIO 4.1.4. Sea un P.R.G.W.B. con función de casamiento natural, tal que se encuentra en el conjunto de no extinción, bajo ciertas condiciones de diferenciación de $r = g(\mu_1,\mu_2)$, se tiene que

$$\left(\sum_{i=0}^{n-1} Z_i \right)^{1/2} (\hat{r} - r) \longrightarrow N(0,\eta)$$

donde η es cierto número positivo.

Demostración:

Desarrollando $\hat{r} = g(\underline{\hat{\mu}})$ en series de Taylor obtenemos la siguiente expresión

$g(\underline{\hat{\mu}}) = g(\underline{\mu}) + (g_x(\underline{\mu}),g_y(\underline{\mu}))\ (\underline{\hat{\mu}} - \underline{\mu}) + R_2$, por lo que

$$g(\hat{\underline{\mu}}) - g(\underline{\mu}) = (\hat{\mu}_1-\mu_1)\; g_x(\mu_1,\mu_2) + (\hat{\mu}_2-\mu_2)\; g_y(\mu_1,\mu_2) + R_2 \text{ , así}$$

$$\left(\sum_{i=0}^{n-1} Z_i \right)^{1/2} (\hat{r} - r) = g_x(\mu_1,\mu_2) \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2} (\hat{\mu}_1-\mu_1) +$$

$$g_y(\mu_1,\mu_2) \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2} (\hat{\mu}_2-\mu_2) + R_2'$$

Por la observación anterior que se hizo, podemos afirmar que

$$\left(\sum_{i=0}^{n-1} Z_i \right)^{1/2} (\hat{r} - r) \xrightarrow{\mathcal{D}} N(0,\eta) \quad \blacksquare$$

4.2. EFICIENCIA ASINTOTICA

Otra de las propiedades de los estimadores de máxima verosimilitud, que serán analizadas en este trabajo, es la eficiencia asintótica.

Muchas definiciones de eficiencia se han dado en la literatura estadística, y por lo tanto algo de confusión existe sobre este concepto (véase Rao (1973), en la pág. 346).

En esta sección nos preocuparemos por ver si el estimador de máxima verosimilitud del vector de medias en un P.R.G.W.B., con distribución de descendientes serie de potencias bivariada, es asintóticamente eficiente según Heyde. Cabe mencionar que la definición de eficiencia asintótica que da Heyde (1975), es para modelos no ergódicos (veáse Basawa y Scott (1983), página 15), la cual generaliza la definición de eficiencia asintótica que da Rao (1973) en la pág. 348 o Basawa y Scott (1983) en pág. 8, para estimadores univariados de máxima verosimilitud que son calculados de observaciones que vienen de modelos ergódicos. La razón por la se utiliza la definición de eficiencia según Heyde, es porque los procesos ramificados bisexuales son un ejemplo de los modelos no ergódicos, esto es, como mencionamos en la sección anterior, el estimador del vector de medias considerado en este capítulo, no tiene distribución asintótica normal.

La definición de eficiencia según Heyde, tomando en cuenta una secuencia de variables aleatorias $\{ X_1, X_2, \dots , X_n \}$, de un proceso estocástico univariado es la siguiente

Un estimador consistente T_n de θ es asintóticamente eficiente si $\{I_n(\theta)\}^{1/2} [T_n - \theta - \gamma(\theta) I_n^{-1}(\theta) S_n(\theta)] \longrightarrow 0$ en probabilidad cuando $n \to \infty$, donde $I_n(\theta) = \sum_{k=1}^{n} E[U_k^2(\theta) \mid \mathcal{F}_{n-1}]$, $U_k(\theta) = S_k(\theta) - S_{k-1}(\theta)$, $S_k(\theta) = \dfrac{\partial \log L_k(\theta)}{\partial \theta}$, $L_k(\theta)$ es la función de verosimilitud de $X_1, X_2, \ldots, X_k$ y $\gamma(\theta)$ es una constante que no depende de las observaciones y posiblemente dependa de θ.

Rao también define lo que es un estimador eficiente en el caso multiparamétrico y donde las observaciones vienen de un modelo ergódico. Su definición es la siguiente

Sea $L_n(\underline{\theta})$ la función de verosimilitud de las variables aleatorias $X_1, \ldots, X_n$, un estimador consistente $\underline{T}_n$ de $\underline{\theta}$ se dice eficiente de primer orden si $\sqrt{n} \mid \underline{T}_n - \underline{\theta} - B Z_n \mid \longrightarrow 0$ en probabilidad, donde B es una matriz de constantes la cual puede depender de $\underline{\theta}$, $Z_n = (z_n^1, z_n^2, \ldots, z_n^q)$ y $z_n^i = \dfrac{1}{n} \dfrac{\partial \log L_n(\theta)}{\partial \theta_i}$, $i = 1, 2, \ldots, q$.

Si un estimador es eficiente según alguna de las definiciones de eficiencia asintótica dadas por Heyde o Rao, se tendrá como una consecuencia, que el estimador se estará concentrando cada vez más alrededor del parámetro estimado, mientras n crece. Esta concentración se lleva a cabo en una distribución normal. No se encontró en la literatura una definición de un estimador eficiente para modelos no ergódicos que contemplara el caso multiparamétrico, y es así como generalizamos la definición de un estimador eficiente asintótico según Heyde, de la siguiente manera

DEFINICION 5.2.1. Un estimador consistente $\underline{T}_n$ de $\underline{\theta}$ es asintoticamente eficiente si

$$\left| \xi_n^{1/2}(\underline{\theta}) \, (\underline{T}_n - \underline{\theta}) - B \, \xi_n^{-1/2}(\underline{\theta}) \, S_n(\underline{\theta}) \right| \longrightarrow 0 \qquad (4.5)$$

en probabilidad, cuando $n \to \infty$, para alguna matriz B que no involucra las observaciones y posiblemente dependa de θ, donde

$$\xi_n(\underline{\theta}) = \operatorname{diag} \left\{ \sum_{i=1}^{n} E(U_i(\underline{\theta}) \; U_i^T(\underline{\theta}) \mid \mathcal{F}_{i-1}) \right\} =$$

$$\operatorname{diag} \left\{ \sum_{i=1}^{n} E\left[\left(\frac{\partial \ln L_i}{\partial \underline{\theta}} - \frac{\partial \ln L_{i-1}}{\partial \underline{\theta}}\right)\left(\frac{\partial \ln L_i}{\partial \underline{\theta}} - \frac{\partial \ln L_{i-1}}{\partial \underline{\theta}} \right)^T \middle| \; \mathcal{F}_{i-1}\right]\right\} \quad y$$

$$S_n(\underline{\theta}) = \sum_{i=1}^{n} U_i(\underline{\theta}) = \frac{\partial \ln L_n}{\partial \underline{\theta}}$$

Analizaremos la implicación más inmediata de esta definición, esto para el caso bivariado. Si $\underline{T}_n$ es asintoticamente eficiente según la definición anterior, entonces se tiene que $\xi_n^{1/2}(\underline{\theta}) \, (\underline{T}_n - \underline{\theta}) \xrightarrow{d} N(\underline{0}, \Gamma_1)$, donde Γ_1 es una matriz definida positiva, lo que significa que el estimador $\underline{T}_n$ se irá concentrando alrededor del parámetro $\underline{\theta}$ a través de una distribución normal, mientras n crece. Esto se debe a la condición (4.5) y a que $\xi_n^{-1/2}(\underline{\theta})S_n(\underline{\theta}) \xrightarrow{d} N(\underline{0},\Gamma_2)$ bajo ciertas condiciones de regularidad. Esto último se debe a que

$$\xi_n^{-1/2}(\underline{\theta})\ S_n(\underline{\theta}) =$$

$$\left[\frac{\dfrac{\partial \ln L_n}{\partial \theta_1}}{\left(\sum_{i=1}^{n} E\left\{\left(\dfrac{\partial \ln L_i}{\partial \theta_1}-\dfrac{\partial \ln L_{i-1}}{\partial \theta_1}\right)^2 \middle| \mathcal{F}_{i-1}\right\}\right)^{1/2}},\ \frac{\dfrac{\partial \ln L_n}{\partial \theta_2}}{\left(\sum_{i=1}^{n} E\left\{\left(\dfrac{\partial \ln L_i}{\partial \theta_2}-\dfrac{\partial \ln L_{i-1}}{\partial \theta_2}\right)^2 \middle| \mathcal{F}_{i-1}\right\}\right)^{1/2}}\right]$$

$$= \left(\frac{\sum_{i=1}^{n} U_{1_i}}{\left(\sum_{i=1}^{n} E(U_{1_i}^{\ 2} | \mathcal{F}_{i-1})\right)^{1/2}},\ \frac{\sum_{i=1}^{n} {}_{2_i}}{\left(\sum_{i=1}^{n} E(U_{2}^{\ 2} | \mathcal{F}_{i-1})^{1/2}\right.}\right) \qquad (4.6)$$

Ahora, la serie $\sum_{i=1}^{n} U_{j_i}$ es una martingala de media 0, j=1,2. Por lo tanto $\left(\sum_{i=1}^{n} U_{1_i},\ \sum_{i=1}^{n} U_{2_i}\right)$ es una martingala bivariada de media cero, y de esta manera, bajo ciertas condiciones de regularidad, es posible aplicar el Teorema de Límite Central para martingalas bivariadas, a la expresión (4.6), y es así como $\xi_n^{-1/2}(\underline{\theta})\ S_n(\underline{\theta}) \xrightarrow{d} N(\underline{0},\Gamma_2)$.

En los P.R.G.W.B. es posible mostrar como el estimador del vector de medias $\underline{\hat{\mu}}$ es asintoticamente eficiente bajo ciertas condiciones.

PROPOSICION 5.2.1 Sea un P.R.G.W.B. tal que su función de casamiento es natural y su distribución de descendientes es la serie de potencias bivariada, si el proceso se encuentra en el conjunto de no extinción, entonces el estimador de máxima verosimilitud de $\underline{\mu}$ es asintoticamente eficiente en el sentido de la Definición 5.2.1.

Demostración:

En primer lugar justificaremos que $\hat{\underline{\mu}}$ es fuertemente consistente. Hemos visto que cada F_i y M_i puede expresarse en términos de variables aleatorias independientes, veáse (4.2), por lo que es posible expresar a $\hat{\underline{\mu}}$ como

$$\hat{\underline{\mu}} = (\hat{\mu}_1, \hat{\mu}_2) = \left(\frac{\sum_{i=1}^{Z_0+\dots+Z_{n-1}} X_i}{Z_0+\dots+Z_{n-1}} , \frac{\sum_{i=1}^{Z_0+\dots+Z_{n-1}} Y_i}{Z_0+\dots+Z_{n-1}} \right)$$

Ya que el proceso se encuentra en el conjunto de no extinción $Z_i > 0$, para toda i, y a medida de que n crece $Z_0+\dots+Z_{n-1}$ crece. Ahora como las variables aleatorias X_i's son independientes e identicamente distribuidas con media μ_1, se puede aplicar la ley fuerte de los grandes números para sumas de variables aleatorias, y de esta manera concluir que $\hat{\mu}_1 \longrightarrow \mu_1$, casi seguramente. Para $\hat{\mu}_2$, la consistencia se demuestra analogamente.

Ahora obsérvese que $\xi^{-1/2}(\theta)\ S_n(\theta) =$

$$
\begin{bmatrix}
\beta^{-1/2}(\theta)\dfrac{\dfrac{1}{\sigma_1}\left(\sum_{i=1}^{n} F_i - \mu_1 \sum_{i=0}^{n-1} Z_i\right) - \dfrac{\sigma_{12}}{\sigma_1\ \sigma_2^2}\left(\sum_{i=1}^{n} M_i - \mu_2 \sum_{i=0}^{n-1} Z_i\right)}{\left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}} \\
\\
\beta^{-1/2}(\theta)\dfrac{-\dfrac{\sigma_{12}}{\sigma_1^2\sigma_2}\left(\sum_{i=1}^{n} F_i - \mu_1 \sum_{i=0}^{n-1} Z_i\right) + \dfrac{1}{\sigma_2}\left(\sum_{i=1}^{n} M_i - \mu_2 \sum_{i=0}^{n-1} Z_i\right)}{\left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}}
\end{bmatrix} =
$$

$$
\beta^{-1/2}(\theta)\begin{bmatrix} 1 & -\dfrac{\sigma_{12}}{\sigma_1\sigma_2} \\ -\dfrac{\sigma_{12}}{\sigma_1\sigma_2} & 1 \end{bmatrix}
\begin{bmatrix} \dfrac{\sum_{i=1}^{n} F_i - \mu_1 \sum_{i=0}^{n-1} Z_i}{\sigma_1\ \left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}} \\ \\ \dfrac{\sum_{i=1}^{n} M_i - \mu_2 \sum_{i=0}^{n-1} Z_i}{\sigma_2\ \left(\sum_{i=0}^{n-1} Z_i\right)^{1/2}} \end{bmatrix}
$$

Pero $\xi^{1/2}(\theta)(\underline{\hat{\mu}} - \underline{\mu}) = \beta^{1/2}(\theta)\begin{bmatrix} \dfrac{1}{\sigma_1} & 0 \\ 0 & \dfrac{1}{\sigma_2} \end{bmatrix} \begin{bmatrix} \sum_{i=0}^{n-1} Z_i \end{bmatrix}^{1/2} (\underline{\hat{\mu}}-\underline{\mu}) =$

$$\beta^{1/2}(\theta) \begin{bmatrix} \dfrac{\sum_{i=1}^{n} F_i - \mu_1 \sum_{i=0}^{n-1} Z_i}{\sigma_1 \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2}} \\ \\ \dfrac{\sum_{i=1}^{n} M_i - \mu_2 \sum_{i=0}^{n-1} Z_i}{\sigma_2 \left(\sum_{i=0}^{n-1} Z_i \right)^{1/2}} \end{bmatrix}$$

Si $B = \beta(\theta) \; \Lambda^{-1}$, donde $\Lambda = \begin{bmatrix} 1 & -\dfrac{\sigma_{12}}{\sigma_1 \sigma_2} \\ -\dfrac{\sigma_{12}}{\sigma_1 \sigma_2} & 1 \end{bmatrix}$, entonces

$\xi^{1/2}(\underline{\theta}) \; (\underline{\hat{\mu}} - \underline{\mu}) = B \; \xi_n^{-1/2}(\underline{\theta}) \; S_n(\underline{\theta})$, y de esta manera el estimador $\underline{\hat{\mu}}$ cumple con la condición (14). ■

Obsérvese que la matriz $B = \begin{bmatrix} 1 & \dfrac{\sigma_{12}}{\sigma_1 \sigma_2} \\ \dfrac{\sigma_{12}}{\sigma_1 \sigma_2} & 1 \end{bmatrix}$

De esta manera queda demostrada la eficiencia asintótica del estimador $\underline{\hat{\mu}}$, lo que significa que este se irá concentrando alrededor del parámetro $\underline{\mu}$ a través de una distribución normal cuando n tiende a crecer.

B I B L I O G R A F I A

ATHREYA K.B., NEY P.E. (1972) BRANCHING PROCESSES. SPRINGER - VERLAG.

BAGLEY J.H. (1986) ON THE ASYMPTOTIC PROPIERTIES OF A SUPERCRITICAL BISEXUAL BRANCHING PROCCES. JOURNAL OF APPLIED PROBABILITY 23: 820-826.

BASAWA I.V., PRAKASA RAO B.L.S. (1980) STATISTICAL INFERENCE FOR STOCHASTIC PROCESSES. ACADEMIC PRESS.

BASAWA I.V., SCOTT D.J. (1983) ASYMPTOTIC OPTIMAL INFERENCE FOR NON-ERGODIC MODELS. SPRINGER-VERLAG.

BILLINGSLEY (1968) CONVERGENCE OF PROBABILITY MEASURES. WILEY.

DALEY D.J. (1968) EXTINCTION CONDITIONS FOR CERTAIN BISEXUAL GALTON-WATSON BRANCHING PROCESSES. Z.WAHRSCHEINLICHKEITSTHEORIE 9: 315-322

DALEY D.J., HULL D.M., TAYLOR J.M. (1986) BISEXUAL GALTON-WATSON BRANCHING PROCESSES WITH SUPERADDITTIVE MATING FUNCTIONS. J. APPL. PROB. 23: 585-600.

HEYDE C.C. (1975) REMARKS ON EFFICIENCY IN ESTIMACION FOR BRANCHING PROCESSES. BIOMETRIKA 62, 1: 49-55.

JOHNSON Y KOTZ (1969) DISTRIBUTIONS IN STATISTICS. DISCRETE DISTRIBUTION. ED. HOUGHTON MIFFLIN COMPANY.

KHATRI C.G. (1959) ON CERTAIN PROPIETIES OF POWER-SERIES DISTRIBUTIONS. BIOMETRIKA 46: 486-490.

KLIMKO L.A., NELSON P.I. (1978) ON CONDITIONAL LEAST SQUARES ESTIMATION FOR STOCHASTIC PROCESSES. THE ANNALS OF STATISTICS. VOL. 6, N° 3, 629-642.

LOÉVE M. (1978) PROBABILITY THEORY I. SPRINGER-VERLANG.

PEREZ-ABREU V. (1987) LOS PROCESSOS ESTOCASTICOS RAMIFICADOS COMO MODELOS PARA DETECTAR BROTES DE EPIDEMIAS DE UNA ENFERMEDAD CONTAGIOSA. ASPECTOS ESTADISTICOS. COMUNICACIONES CIMAT.

RAO C.R. (1973) LINEAR STATISTICAL INFERENCE AND ITS APPLICATIONS. SECOND EDITION. WILEY SERIES IN PROBABILITY AND MATHEMATICAL STATISTICS.

SEBER G.A.F. (1984) MULTIVARIATE OBSERVATIONS. JOHN WILEY AND SONS.