ALGUNAS CUESTIONES SOBRE ESTIMACION PARAMETRICA Y NO PARAMETRICA EN PROCESOS DE NACIMIENTO PURO Y DE GALTON-WATSON

MEMORIA que, para optar al grado de Doctor, presenta el licenciado en Ciencias Matematicas D. Antonio Martin Andres.

DIRECTORES DE LA TESIS: Profesor Dr. D. Alfonso Guiraum Martin
Profesor Dr. D. Ramon Gutierrez Jaimez

UNIVERSIDAD DE GRANADA. FACULTAD DE CIENCIAS.

"ALGUNAS CUESTIONES SOBRE ESTIMACION PARAMETRICA Y NO PARAMETRICA EN PROCESOS DE NACIMIENTO PURO Y DE GALTON-WATSON"

Tesis presentada para aspirar al grado de Doctor en Ciencias, Seccion de Matematicas, por:

D. ANTONIO MARTIN ANDRES

realizada bajo la direccion del Catedratico de Estadistica Matematica y Calculo de Probabilidades, profesor Dr. D. ALFONSO GUIRAUM MARTIN y el Agregado de Estadistica Matematica y Calculo de Probabilidades, profesor Dr. D. RAMON GUTIERREZ JAIMEZ, en la Facultad de Ciencias de la Universidad de Granada y juzgada el dia 26 de Junio de 1976 en dicha Facultad por el siguiente Tribunal:

Presidente: Dr. D. SIXTO RIOS GARCIA, Catedratico de Estadistica Matematica y Calculo de Probabilidades de la Universidad Complutense de Madrid, Academico de Ciencias.

Vocales: Dr. D. FRANCISCO SALES VALLES, Catedratico de Estadistica Matematica y Calculo de Probabilidades de la Universidad Central de Barcelona.

Dr. D. ALFONSO GUIRAUM MARTIN, Catedratico de Estadistica Matematica y Calculo de Probabilidades de la Universidad de Granada, PONENTE de la Tesis.

Dr. D. ENRIQUE TRILLAS RUIZ, Catedratico de Matematicas de la Escuela Tecnica Superior de Arquitectura de Barcelona.

Secretario: Dr. D. RAMON GUTIERREZ JAIMEZ, Agregado de Estadistica Matematica y Calculo de Probabilidades de la Universidad de Granada, PONENTE de la Tesis.

Calificacion obtenida: SOBRESALIENTE CUM LAUDE, por unanimidad.

A MI ESPOSA E HIJAS

Mi agradecimiento :

A D.ALFONSO GUIRAUM MARTIN y D.RAMON GUTIERREZ JAIMEZ,catedrático y agregado,respectivamente,del Departamento de Estadistica Matematica e Investigacion Operativa de la Facultad de Ciencias,directores de esta Tesis,por su continuo animo y ayuda.

A los Departamentos de Estadistica de la Facultad de Ciencias y de Fisiologia y Bioquimica de la Facultad de Medicina,que han soportado parte de los gastos de esta Tesis.

A D.Antonio Martinez Molina,profesor de la Facultad de Ciencias,que realizo el programa-en leguaje FORTRAN V - para la simulacion de Monte Carlo.

Al Centro de Calculo de la Universidad de Granada,en el cual se ejecuto el programa de simulacion,y al personal del mismo por sus continuas ayudas en cuantos problemas se presentaron al respecto.

Granada,Primavera de 1976

I N D I C E

- - -

CAPITULO I

F U N D A M E N T O S

1.1.INTRODUCCION.-

En este primer capítulo,no pretendemos dar un estudio exhaustivo de todo lo existente sobre procesos de Galton-Watson y de nacimiento puro,por caer fuera de los limites de este trabajo;unicamente se desea exponer,del modo mas homogeneo posible,aquellos resultados que en los proximos capítulos - sobre estimacion de los parametros de los procesos mencionados - nos seran de utilidad.

1.2.PROCESOS DE NACIMIENTO PURO.-

1.2.1.DEFINICION Y PROPIEDADES BASICAS.-

Sea $X(t)$ un proceso de Markov de estados $0,1,2,\ldots\ldots$ y con probabilidad de transicion $P_{ij}(t)$ estacionaria,esto es,

$$P_{ij}(t) = P\left\{ X(t+s)=j \ / \ X(t)=i \right\}$$

solo depende de t pero no de s. Dicho proceso sera llamado de nacimiento y muerte si se verifican las condiciones siguientes:

1ª) $P_{i,i+1}(h) = \lambda_i . h + o(h)$,, $i \geq 0$

2ª) $P_{i,i}(h) = 1 - (\lambda_i + \mu_i)h + o(h)$,, $i \geq 0$

3ª) $P_{i,i-1}(h) = \mu_i . h + o(h)$,, $i \geq 1$

4ª) $P_{i,j}(0) = \delta_{ij}$

5ª) $\mu_o = 0$,, $\lambda_o \geq 0$,, $\lambda_i, \mu_i > 0$,, $i = 1, 2, \ldots$

en donde λ_i y μ_i son las proporciones infinitesimales de nacimiento y muerte.

Si el intervalo $(0, t+h)$ se divide en los intervalos $(0,h)$ y $(h, t+h)$, entonces se obtienen las ecuaciones atrasadas de Kolmogorov que tienen el estado inicial variable:

$$P'_{oj}(t) = -\lambda_o . P_{oj}(t) + \lambda_o . P_{1j}(t)$$

$$P'_{ij}(t) = \mu_i . P_{i-1,j}(t) - (\lambda_i + \mu_i) P_{ij}(t) +$$

$$+ \lambda_i P_{i+1,j}(t) \quad ,, \ i \geq 1 \quad ,, \ j \geq 0$$

con la condicion de contorno

$$P_{ij}(o) = \delta_{ij}$$

en donde P'indica derivada respecto del tiempo t.

Cuando el intervalo (0.t+h) se divide en los intervalos (0,t) y (t,t+h) obtendremos las ecuaciones adelantadas de Kolmogorov con el estado final variable:

$$P'_{io} = -\lambda_o P_{io}(t) + \mu_1 P_{i1}(t)$$

$$P'_{ij}(t) = \lambda_{j-1} P_{i,j-1}(t) - (\lambda_j + \mu_j) P_{ij}(t) +$$

$$+ \mu_{j+1} P_{i,j+1}(t) \quad ,, \; i \geq 0 \quad ,, \; j \geq 1$$

en donde,en ambos casos,se ha hecho tender h hacia 0.

En particular,cuando $\mu_i = 0$,, $\forall$ i entonces la muerte es imposible y el proceso sera llamado de nacimiento puro.Finalmente,cuando $\lambda_i = i.\lambda$,, $\forall i$,entonces el proceso sera llamado <u>proceso de nacimiento puro lineal de intensidad λ</u> ,de modo que las ecuaciones difenenciales adelantadas son ahora:

$$P'_{ij}(t) = -\, j.\lambda.P_{ij}(t) + (j-1).\lambda.P_{i,j-1}(t) ,, j \geq i+1$$

$$P'_{ii}(t) = -\, i\,\lambda\, P_{ii}(t)$$

cuya solucion es:

$$P_{ij}(t) = P\left\{ X(s+t) = j \;/\; X(s)=i \right\} =$$

$$= \begin{cases} \binom{j-1}{-i+j} \left(e^{-\lambda t}\right)^{i} \cdot \left(1 - e^{-\lambda t}\right)^{j-i} & ,,j \quad i \quad 1 \\ 0 & \text{en otro caso} \end{cases}$$

de modo que $X(t+s)-X(s)$ es una binomial negativa de parametros :

$$p= e^{-\lambda t} \qquad ,, \qquad r=i=X(s)$$

con lo cual:

$$E\left\{ X(t) \;/\; X(o)=q \right\} = q.e^{\lambda t}$$

$$V\left\{ X(t) \;/\; X(o)=q \right\} = q.e^{\lambda t}.\left(e^{\lambda t} -1\right)$$

Este proceso es denominado a veces por proceso de Yule y suele hablarse de el en los terminos siguientes :

Sea una poblacion de individuos en la que un individuo presente en el tiempo t tiene una probabilidad de dar lugar a otro individuo en el intervalo $(t,t+\Delta t)$ de $\lambda\Delta t + o(\Delta t)$ con independencia de su edad,del individuo y de lo que les ocurra al resto de los individuos.Si ademas los individuos nunca mueren,entonces tendremos un proceso de nacimiento puro lineal en donde el parametro λ es llamado por intensidad o tasa de nacimiento o prporcion de nacimientos por individuo.Para que el proceso tenga lugar,hace falta que el numero inicial de individuos presentes en el tiempo cero sea $X(o)=q \geq 1$ puesto que $\lambda_o = 0.\lambda = 0$.

Las pruebas y ampliacion de todo lo expuesto pueden encontrase en Tautu(seccion 2.3.4.7.),Karlin(pgs. 189-206),Cox-Miller(seccion 4.3) ó Parzen (cap.7).

1.2.2.TIEMPOS DE PRIMERA ESTANCIA Y TIEMPOS DE OCUPACION.-

Para un proceso de nacimiento y muerte de intensidades λ_i y μ_i ,el movimiento de X(t)-ver Karlin pgs.190-191 -puede describirse asi:

El proceso esta en un estado i un tiempo aleatorio que sigue una exponencial de parametro $(\lambda_i+\mu_i)$ .Cuando sale de el,lo hace al estado i+1 con probabilidad $\lambda_i(\lambda_i+\mu_i)^{-1}$ y al i-1 con probabilidad $\mu_i(\lambda_i+\mu_i)^{-1}$.

por lo cual su movimiento es analogo al de un recorrido aleatorio salvo que las transiciones ocurren en tiempos al azar.

Con ello, cuando se trate de un proceso de nacimiento puro de intensidad λ, entonces el tiempo Z_i que el proceso gasta en el estado i o tiempo de estancia en el estado i-tiempo transcurrido desde la llegada al estado i hasta la llegada al estado i+1 - es una variable aleatoria que sigue una exponencial de parametro $\lambda .i$.

Sea ahora T_a el tiempo en que el estado a es ocupado por primera vez o tiempo de primera estancia en a; naturalmente que a debe de ser mayor que X(o)=q. Claramente:

$$P(T_a \leq t) = P \{ X(t) \geq a \quad / \quad X(o)=q \}$$

que al ser X(t) creciente no seria dificil de obtener a partir de la funcion generatriz de X(t). Hagamoslo de otro modo alternativo; como Z_i es el tiempo gastado en el estado i, entonces:

$$T_a = Z_q + Z_{q+1} + \ldots\ldots + Z_{a-1}$$

y por lo tanto:

$$E(T_a) = \frac{1}{\lambda} \sum_{h=q}^{a-1} \frac{1}{h}$$

$$V(T_a) = \frac{1}{\lambda^2} \cdot \sum_{h=q}^{a-1} \frac{1}{h^2}$$

Dichas cantidades son laboriosas en su calculo,pero de un modo aproximado pueden usarse las expresiones:

$$E(T_a) \simeq \frac{1}{\lambda} . \ln \frac{a}{q} \quad \text{si } a/q \text{ y } q \text{ son grandes}$$

$$E(T_a) \simeq \frac{1}{\lambda} . \ln(a-1) + \frac{\gamma}{\lambda} - \frac{1}{\lambda} \sum_{h=1}^{q-1} \frac{1}{h}$$

si a es grande y q es pequeño.

en donde $\gamma \simeq 0{,}5772$ es la constante de Euler.Ademas:

$$V(T_a) \simeq \frac{\pi^2}{6\,\lambda^2} - \frac{1}{\lambda^2} \sum_{h=1}^{q-1} \frac{1}{h^2} \quad \text{si a es grande.}$$

Notese que $V(T_a)$ es finita cuando aumente el tamaño de a.

Finalmente,como la funcion generatriz

$$E(e^{-uT_a}) = \prod_{j=q}^{a-1} \frac{j.\lambda}{u + j.\lambda}$$

entonces, cuando $q=1$, la variable aleatoria

$$U_a = \lambda . T_a - \ln a$$

se distribuye, cuando $a \longrightarrow \infty$, como una exponencial, teniendo U_a la densidad limite $\exp.(-x - e^{-x})$ con $-\infty < x < \infty$.(Ver Cox-Miller pgs. 160-161).

1.2.3.LAS LEYES LIMITES.-

Sea $X(t)$ un proceso de nacimiento puro de intensidad λ ; si $X_o=q$ entonces por lo dicho en la seccion 2.1.:

$$E(X(t)) = q.e^{\lambda t}$$

Con ello:

$$W_t = \frac{X(t)}{E(X(t))}$$

es una martingala no negativa, de parametro continuo, que converge c.s. a un limite W finito y positivo. Ademas - ver Ťautu pgs.273-4 o Harris pgs.104-147 - la distribucion de W es una gamma de parametros (q,q) o sea:

$$P\,(\,W < w \;/\; X(o)=q\,) = \frac{1}{(q-1)!} \cdot \int_{o}^{w} e^{-qu}.u^{q-1}.du$$

con $u \geq 0$.Finalmente,$X(t)$ posee una distribucion condicional con W como variable condicionante;ello se vera mas explicitamente en el teorema 2.2.1. del capitulo II.

1. 3.PROCESOS DE GALTON-WATSON.-

1. 3.1.DEFINICION Y PROPIEDADES BASICAS.-

Se denominan por cadenas de ramificacion del tipo Galton-Watson - en adelante las denominaremos abreviadamente por procesos de Galton-Watson - a una cadena de Markov $\{X_n, n \in N\}$ cuyas probabilidades de transicion de un paso estan dadas por:

$$P(i,j) = P(\,X_{n+1}=j \;/\; X_n = i\,) =$$

$$= P(\,\xi_1 + \ldots\ldots + \xi_i = j)$$

endonde ξ_1 , ... , ξ_i son i variables aleatorias discretas e independientes con funcion de probabilidad

$$\{ p_k \ ,, \ k=0,1,2,\ldots\ldots \} \quad p_k \geqslant 0 \ ,, \ \sum p_k = 1$$

La interpretacion de tal proceso es la clasica - y en estos términos hablaremos en todo lo que sigue - de que se dispone de un numero inicial X_o de individuos,cada uno de los cuales,al morir,tiene un numero aleatorio ξ de hijos e independiente de los demas.Dicho número aleatorio tiene la distribucion $\{p_k, k \in N\}$ anterior,a la que llamaremos por <u>distribucion del número de hijos</u> , y pasa a formar parte de la primera generacion del proceso.Con ello,el número de individuos al final de la primera generacion sera:

$$X_1 = \xi_1 + \xi_2 + \ldots\ldots\ldots + \xi_{X_o}$$

en donde ξ_i es el numero de hijos del individuo número i de la generacion 0.De igual modo surgen la 2ª generacion-a partir de los individuos de la 1ª - de tamaño X_2,la 3ª de tamaño X_3,etc.

Es muy importante,en este tipo de procesos, el utilizar la funcion generatriz del numero de hijos:

$$f(s) = \sum_{k=o}^{\infty} p_k . s^k \quad ,, \quad |s| \leq 1$$

en la que cláramente

$$f(1)= 1 \quad ,, \quad f(o) = p_o$$

Supongamos que X_o=1;entonces la funcion generatriz $f_n(s)$ del tamaño de la poblacion en la generacion n-esima es:

$$f_o(s) = s$$

$$f_1(s) = f(s)$$

y en general:

$$f_{n+1}(s) = \sum_{k=o}^{\infty} P(X_{n+1}=k).s^k = f_n(f(s))=$$

$$= f(f_n(s)) \qquad (1.1)$$

relaciones que aun pueden generalizarse mas mediante la expresion:

$$f_{m+n}(s) = f_m (f_n(s)) \quad ,, \quad m,n \in N$$

Finalmente,cuando X_o=q,entonces - por la independencia de las variables aleatorias ξ_i - si llamamos por

$$\varphi_n(s) = \sum_k P(X_n=k / X_o=q).s^k$$

entonces:

$$\varphi_o(s) = s^q = \{ f_o(s) \}^q$$

$$\varphi_1(s) = \{ f(s) \}^q$$

y en general:

$$\varphi_n(s) = \{ f_n(s) \}^q = \varphi(\varphi_{n-1}(s))$$

Supongamos en todo lo que sigue,y mientras que no se diga lo contrario,que $X_o=1$ y que:

$$\mu = E(X_1 / X_o=1) = E(\xi)$$

$$\sigma^2 = V(X_1 / X_o=1) = V(\xi)$$

son finitos y existentes.Como $E(X_n) = f'_n(1)$,por derivacion de la (1.1) puede obtenerse que:

$$E(X_n) = \mu^n \tag{1.2}$$

$$V(X_n) = \begin{cases} \sigma^2 \cdot \mu^{n-1} \cdot \dfrac{\mu^n - 1}{\mu - 1} & \text{si } \mu \neq 1 \\ n \cdot \sigma^2 & \text{se } \mu = 1 \end{cases}$$

Con lo cual la varianza crece o decrece geométricamente si $\mu > 1$ o $\mu < 1$, y linealmente si $\mu = 1$, con la edad n en generaciones. La media y varianza cuando $X_o = q$ es sencilla de obtener por la independencia en la actuacion de los individuos, de modo que:

$$E(X_n / X_o=q) = q.E(X_n / X_o=1)$$

$$V(X_n / X_o=q) = q.V(X_n / X_o=1)$$

Con relacion a la clasificacion de los estados, puede verse fácilmente que el estado k=0 es absorbente, mientras que los estados $k \neq 0$ son todos transitorios si $p_1 \neq 1$. Con ello, si $p_1 \neq 1$:

$$\lim_{n \to \infty} P (X_n=k) = 0 \quad ,, \quad \forall\ k \neq 0$$

de modo que un proceso de Galton-Watson, en el limite, no puede ser moderadamente grande, sino que es cero o es infinito.

Todas las pruebas de esta seccion y de las dos siguientes pueden verse en Tautu (pgs. 90-99), Karlin (pgs. 286-98) y Athreya-Ney (cap. I).

1.3.2. LAS PROBABILIDADES Y LOS TIEMPOS DE EXTINCION.-

Se denomina por probabilidad de extincion de un proceso de Galton-Watson a:

$$\pi = P\ (\ X_n=0 \text{ para algun } n>0\ )$$

es decir, la probabilidad de que la poblacion finalmente se extinga.

Supongamos que $p_o=1$; entonces $X_1=0$ y el problema no tiene interes. Ademas, si $p_o=0$ entonces nunca ocurriria la extincion y por tanto $\pi=0$. Por ello, en adelante, nos limitaremos al caso de que $0<p_o<1$.

Se ha definido π por:

$$\pi = P(\ X_n=0 \text{ para algun } n \in N^*\ ) =$$

$$= P\ \left(\ \bigcup_{n \in N^*} (\ X_n=0\ )\ \right)$$

y como $P(X_{n+1}=0\ /\ X_n=0)=1$, entonces:

$$\pi = \lim_{n \to \infty} P \left\{ (X_1=0) \cup \ldots\ldots \cup (X_n=0) \right\} =$$

$$= \lim_{n \to \infty} P\ (\ X_n=0\ ) = \lim_{n \to \infty} f_n(o)$$

con lo que puede probarse el teorema siguiente:

<u>Teorema 1.1</u>. La probabilidad de extincion de un proceso de Galton-Watson X_n con $X_o=1$ y funcion generatriz de hijos f(s),es la mas pequeña raiz no negativa π de la ecuacion

$$f(s) = s$$

Ademas:

Si $\mu \leq 1$ entonces $\pi = 1$

Si $\mu > 1$ entonces $\pi < 1$

y como consecuencia:

<u>Corolario</u>.

a) $\lim_{n \to \infty} P (X_n=k) = 0 \quad ,, \quad \forall \ k \geq 1$

b) $P (\lim_{n \to \infty} X_n = 0) = 1 - P(\lim_{n \to \infty} X_n = \infty) = \pi$

con lo cual la sucesion $\{ X_n \}$ converge hacia infinito o hacia cero con probabilidades $(1-\pi)$ y π respectivamente,cuando n tiende hacia infinito.

Naturlamente que,cuando $X_o=q$, la probabilidad de extincion sera π^q.

Paralelamente a la probabilidad de extincion π ,definimos por tiempo de extincion al tiempo de primer paso por el estado 0,es decir,al mas pequeño subindice s tal que $X_s=0$.La distribucion de s sera:

$$P(\ s=0) = 0$$

$$P(s=n) = f_n(o) - f_{n-1}(o) \quad ,, \quad n \in N^*$$

de modo que:

<u>Teorema 1.2</u>.-

a) Si $\mu > 1$ entonces $E(s)=\infty$.

b) Si $\mu < 1$ entonces $P(s>n) = 1- f_n(o)$ se aproxima a cero al menos exponencialmente,de modo que todos los momentos de s son finitos.

c) Si $\mu=1$ entonces $E(s)$ y $F = \int_0^1 \frac{1}{f(x)-x} - \frac{x}{} dx$

son finitas o infinitas ambas a la vez.

Podemos estar interesados en saber el numero total de individuos que han existido hasta el tiempo de extincion.Claramente:

$$V = X_o + X_1 + \dots\dots + X_s$$

sera dicho numero total.Si denominamos ahora por c_j a

$$c_j = P(\ V=j\ ) \quad ,, \quad j \in N^*$$

y por c(z) a su funcion generatriz,entonces:

<u>Teorema 1.3</u>.- Si $p_o > 0$, entonces $c(z)$ es la única raiz de la ecuacion funcional:

$$c(z) = z.f\{c(z)\}$$

con $0 < c(z) \leq 1$ para $0 < z \leq 1$. Ademas:

$$P(V < \infty) = c(1) = \begin{cases} 1 & \text{si } \mu \leq 1 \\ \pi & \text{si } \mu > 1 \end{cases}$$

Volviendo a la probabilidad de extincion, recordemos que $f_n(o)$ converge hacia π; veamos a continuacion tres teoremas que nos hablan de la intensidad de tal convergencia:

<u>Teorema 1.4</u>.- Si $\mu > 1$ y $p_o > 0$ entonces $0 < f'(\pi) < 1$ y :

$$f_n(o) = \pi - \left(\sum_{j \in N} \gamma_j . \pi^j \right) . \left[f'(\pi) \right]^n +$$

$$+ o\left\{ (f'(\pi))^{2n} \right\}$$

cuando $n \longrightarrow \infty$, en donde γ_j son no negativas y determinadas univocamente por las ecuaciones:

$$\sum_{i \in N} \gamma_i . P(i,j) = f'(\pi) . \gamma_j \quad ,, \quad j \in N^*$$

con

$$\sum_{i \in N^*} \gamma_i . i . \pi^{i-1} = 1.$$

asi,en el caso fraccional lineal (ver seccion siguiente) $\gamma_i = (1-\pi)^2$ para todo $i \in N^*$.

Teorema 1.5.- Si $\mu < 1$ entonces:

$$\frac{\mu^n}{1 - f_n(o)} \longrightarrow k \neq 0$$

cuando $n \longrightarrow \infty$,con k finita o no segun que la serie

$$\sum_{i \in N^*} i . \ln i . p_i$$

sea convergente o divergente.

Teorema 1.6.- Si $\mu = 1$, $p_1 \neq 1$ y $\sigma^2 < \infty$ entonces:

$$f_n(o) = 1 - \frac{2}{n . \sigma^2} + o(n^{-1})$$

cuando $n \longrightarrow \infty$.

1.3.3.EL CASO FRACCIONAL LINEAL.-

Normalmente,las funciones generatrices de distribucion de hijos tienen el problema de que el calculo de $f_n(s)$ - o n-iterada de $f(s)$ - es complicado; hay sin embargo una funcion generatriz,que denominaremos en adelante por <u>caso fraccional lineal</u>,que,aparte de poder calcularse fácilmente la $f_n(s)$, presenta la ventaja de depender de dos parámetros.Asi,diremos que la funcion generatriz del número de hijos es del tipo fraccional lineal cuando:

$$f(s) = 1 - \frac{b}{1-c} + \frac{b.s}{1-c.s}$$

con $b,c > 0$ y $b+c \leq 1$.Con ello:

$$p_k = P\ (\ X_1 = k\ /\ X_o = 1\ ) = b.c^{k-1} \quad,, \quad k=1,2,\ldots.$$

$$p_o = P(\ X_1 = 0\ /\ X_o = 1\ ) = \frac{1-b-c}{1-c}$$

y por tanto:

$$\mu = \frac{b}{(1-c)^2}$$

$$\sigma^2 = \mu\frac{1+c}{1-c} - \mu^2$$

$$\pi = \min.\left\{ 1, \frac{1-b-c}{c(1-c)} \right\}$$

con lo cual, si $\mu > 1$, $=1$, < 1 entonces $\pi < 1, =1$, ò $= 1$.

Puede verse que la forma de $f_n(s)$ depende del valor de la media μ. Asi, cuando $\mu \neq 1$, entonces:

$$f_n(s) = 1 - \frac{b(n)}{1 - c(n)} + \frac{b(n).s}{1 - c(n).s}$$

en donde:

$$b(n) = \mu^n . \left(\frac{1 - \pi}{\mu^n - \pi} \right)^2$$

$$(1.4)$$

$$c(n) = \frac{\mu^n - 1}{\mu^n - \pi}$$

con lo cual las probabilidades de transicion de n pasos desde el estado 1 son:

$$P(X_n = k) = b(n). \{ c(n) \}^{k-1} ,, \quad k=1,2,....$$

$$(1.5)$$

$$P(X_n = 0) = 1 - \mu^n . \frac{1 - \pi}{\mu^n - \pi}$$

De igual modo, cuando $\mu = 1$, entonces $b = (1-c)^2$ y $\pi = 1$, de modo que:

$$f(s) = \frac{c - (2c - 1).s}{1 - c.s}$$

$$\sigma^2 = \frac{2c}{1 - c}$$

Ahora la n-iterada de f(s) es:

$$f_n(s) = \frac{nc - (nc + c - 1).s}{1 - c + nc - ncs}$$

con lo cual las probabilidades de transicion de n pasos son:

$$P(X_n=k) = \left\{ 1- c(n) \right\}^2 .c(n)^{k-1} \quad ,, \ k=1,2,.....$$

$$P(X_n=0) = c(n) \tag{1.6}$$

en donde:

$$c(n) = \frac{nc}{1 + (n-1)c} \tag{1.7}$$

Como mas adelante necesitaremos la distribucion de probabilidad de la variable aleatoria X_n condicionada en que $X_n > 0$,de las expresiones (1.5) y (1.6) se deduce que :

$$P(X_n = k / X_n > 0) = \left\{ 1 - c(n) \right\} . c(n)^{k-1} \quad ,, \quad k=1,2,......$$

en ambos casos(para μ distinto o igual a 1),en donde c(n) toma los valores (1.4) o (1.7) segun que μ sea $\neq 1$ o =1.Con ello:

$$E(X_n / X_n > 0) = \frac{\mu^n}{1 - P(X_n=0)} = \frac{\mu^n - \pi}{1 - \pi}$$

si $\mu \neq 1$ y :

$$E(X_n / X_n > 0) = 1 + \frac{nc}{1 - c}$$

si $\mu = 1$.

Finalmente,si s es el tiempo de extincion:

$$P(s=n) = \mu^{n-1} . \pi . \frac{(\mu - 1)(1-\pi)}{(\mu^n - \pi)(\mu^{n-1} - \pi)}$$

cuando $\mu > 1$ y:

$$P(s=n) = \frac{c(1-c)}{\{1 + (n-1)c\}.\{1 + (n-2)c\}}$$

si $\mu = 1$.

1. 3.4.ALGUNOS TEOREMAS LIMITES.-

Ya hemos visto mas arriba que X_n tiende hacia cero o infinito,cuando n tiende hacia infinito,con probabilidad 1.¿Cual es la naturaleza de esta divergencia? Veamos de estudiar para ello el comportamiento limite de X_n.

Ya sabemos que por la propiedad de Markov:

$$E(X_{n+k} / X_n=i_n, X_{n-1}=i_{n-1},\ldots\ldots,X_o=i_o)=$$

$$=E(X_{n+k} / X_n=i_n)= i_n.\mu^k$$

con lo cual, si $W_n= X_n.\mu^{-n}$ entonces :

$$E(W_{n+k} / W_o, W_1, \ldots\ldots , W_n) = W_n$$

Con ello se prueba que:

<u>Teorema 1.7</u>.- Si $0 < \mu < \infty$, $W_n = X_n \cdot \mu^{-n}$ y F_n es la σ-algebra generada por $X_0, X_1, \dots,$ X_n, entonces $\{W_n, F_n; n=0,1,2,\dots\}$ es una martingala. Ademas, cuando $W_n \geq 0$, existe una variable aleatoria W tal que :

$$\lim_{n \to \infty} W_n = W \qquad \text{c.s.}$$

Con ello $X_n(w)$ crece como $\mu^n . W(w)$, pero por ahora no sabemos nada de W. Desde luego puede ocurrir que $W \equiv 0$ en cuyo caso el teorema anterior no dice nada excepto que μ^n es demasiado fuerte como factor normalizador. De hecho, si $\mu \leq 1$ entonces $X_n = 0$, para grandes n, con probabilidad 1, de modo que W es degenerada en cero; de ahi que el teorema solo tenga pleno significado cuando $\mu > 1$. Aun en este caso podra ocurrir que $P(W=0)=1$; sin embargo si $\sigma^2 < \infty$ podemos asegurar que W no es degenerada. Mas adelante se dara un teorema al respecto.

Veamos a continuacion algunos teoremas limites para cada uno de los posibles procesos de Galton-Watson segun cual sea el valor de μ ; dichos teoremas limites seran utilizados a menudo en el cap. III y solo reproducimos aqui aquellos que alli nos seran útiles.

Las pruebas para el teorema 1.7, y los teoremas posteriores, pueden encontrarse en Athreya-Ney (cap. I) y en Tǎutu (pgs. 105-8).

1.3.4.1. TEOREMAS LIMITES EN EL CASO SUBCRITICO.-

Cuando $\mu < 1$ el proceso muere con probabilidad 1. Para describir su comportamiento asintótico, Yaglom ha introducido el artificio de condicionar X_n en el suceso $\{X_n > 0\}$. Asi:

<u>Teorema 1.8.</u>- (Teorema de Yaglom)

Si $\mu < 1$, entonces para cada $j \in N^*$ existe

$$\lim_{n \to \infty} P(X_n = j \; / \; X_n > 0) = b_j \text{ con } \sum b_j = 1.$$

Ademas, la funcion generatriz $b(s) = \sum b_j . s^j$ es la única funcion generatriz de probabilidad con $b(o)=0$ que verifica la ecuacion funcional:

$$b\{f(s)\} = \mu . b(s) + (1-\mu)$$

Ademas, $b'(1)$ es la constante k del teorema 1.5. Finalmente:

$$\sum j . b_j < \infty \iff \sum j . p_j . \ln j < \infty$$

1.3.4.2.- TEOREMAS LIMITES PARA EL CASO CRITICO.-

Cuando $\mu = 1$ no basta con considerar el proceso $\{X_n \; / \; X_n > 0\}$ pues puede verse que diverge a infinito. Hace falta una apropiada normalizacion para que el proceso condicionado converja a un limite no degenerado.

En el teorema 1.6. ya vimos que si $\mu = 1$ y $\sigma^2 < \infty$ entonces, cuando $n \longrightarrow \infty$:

$$1 - f_n(o) = P(X_n > 0) \simeq \frac{2}{n.\sigma^2}$$

Con ello, como :

$$E(X_n / X_n > 0) = \frac{1}{P(X_n > 0)}$$

parece conveniente estudiar el proceso

$$\left\{ \frac{X_n}{n} \; / \; X_n > 0 \right\}$$

Asi, Spitzer y otros han probado que

<u>Teorema 1.9</u>.- Si $\mu = 1, p_1 \neq 1$ y $\sigma^2 < \infty$, entonces:

$$\lim_{n \longrightarrow \infty} .P \left\{ \frac{X_n}{n} \cdot \frac{2}{\sigma^2} \geq x \; / \; X_n > 0 \right\} = e^{-x}$$

para $x \geq 0$.

Nótese que dicho teorema limite es válido cualquiera que sea la funcion generatriz f(s) y que su resultado no depende de f(s), cosa que no ocurría en el caso $\mu < 1$ ni ocurrirá en el caso $\mu > 1$.

1.3.4.3.TEOREMAS LIMITES EN EL CASO SUPERCRITICO.-

El teorema 1.7. quedó pendiente de particularización al caso $\mu > 1$.Asi :

<u>Teorema 1.10</u>.- Si $\mu > 1$ y $\sigma^2 < \infty$ entonces:

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

b) $E(W) = 1$,, $V(W) = \dfrac{\sigma^2}{\mu(\mu-1)}$

c) $P(W =0) = \pi$ = probabilidad de extinción.

Con ello,la condición $\sigma^2 < \infty$ es suficiente para que $P(W=0) < 1$,pero no hace falta algo tan fuerte como que $\sigma^2 < \infty$ para la no degeneración de W;asi,el teorema siguiente da una condición algo mas fuerte que la existencia de la media.En toda esta sección se supone que $X_0=1$, $\mu > 1$ y $p_j \neq 1$,, $\forall$ j.

<u>Teorema 1.11</u>.-

a) Si $E(X_1 \ln X_1) < \infty$ entonces $E(W)=1$.

b) Si $E(X_1 \ln X_1) = \infty$ entonces $E(W)=0$ o,equivalentemente, $P(W=0)=1$.

Con ello,cuando $E(X_1 \ln X_1) = \infty$,entonces:

$$P\left\{ \lim_{n\to\infty} \frac{X_n}{\mu^n} = 0 \right\} = 1$$

lo cual indica que μ^n no es una normalización correcta para que X_n converja a una ley limite no degenerada en cero. Sin embargo, Seneta ha probado que siempre existe una sucesión $\{C_n\}$ tal que $C_n^{-1}.X_n$ converge, en distribución, a un limite no degenerado. Asi:

<u>Teorema 1.12</u>.- Sea $\{X_n, n=0,1,2,\ldots\}$ un proceso de Galton-Watson con $1<\mu<\infty$. Entonces siempre existe una sucesión de constantes C_n con $C_n \longrightarrow \infty$ y $C_n^{-1}.C_{n+1} \longrightarrow \mu$ cuando $n \longrightarrow \infty$, tal que la variable aleatoria $W_n = C_n^{-1}.X_n$ converge c.s. a una variable aleatoria W con $P(W>0)=1-\pi$.

Ademas, si $\varphi(s) = E(e^{-sW})$ para $Re.s \geq 0$, entonces $\varphi(s)$ verifica la igualdad:

$$\varphi(s) = f\left\{\varphi\left(\frac{s}{\mu}\right)\right\}$$

Volviendo a la variable aleatoria W del teorema 1.10., veamos de especificar sus propiedades y distribución:

<u>Teorema 1.13</u>.- Si $\mu > 1$ y $\sum k.p_k.\ln k < \infty$ entonces:

$$\frac{X_n}{\mu^n} \longrightarrow W \quad c.s.$$

cuando $n \longrightarrow \infty$, en donde W es una variable aleatoria con un salto de magnitud π en el origen y de densidad continua en $(0,\infty)$. Ademas, si $\varphi(s)=E(e^{-sW})$ entonces $\varphi(s)$ es la única solución de la ecuación funcional:

$$\varphi(s) = f\left\{\varphi\left(\frac{s}{\mu}\right)\right\}$$

verificando la condición $\varphi'(o)=-1$.

Como mas adelante nos hara falta la distribición limite de X_n/μ^n condicionada en el suceso $\{X_n > 0\}$ es inmediata la extensión del teorema anterior en el siguiente:

<u>Teorema 1.13.bis</u>.- Si $\mu > 1$ y $\sum k.p_k.\ln k < \infty$ entonces:

$$\left(\frac{X_n}{\mu^n} \Big/ X_n > 0\right) \dashrightarrow W' \text{c.s.}$$

cuando $n \dashrightarrow \infty$, en donde W' es una variable aleatoria de distribución continua en $(0,\infty)$ y tal que si

$$\varphi(s) = E(e^{-sW})$$

entonces $\varphi(s)$ es la única solución de la ecuación funcional :

$$\varphi(\mu s)(1-\pi) + \pi = f\{\varphi(s)(1-\pi) + \pi\}$$

verificando que $\varphi'(o)=-E(W')=-(1-\pi)^{-1}$.

El cálculo de $\varphi(s)$ no suele ser sencillo, pero en el caso fraccional lineal puede verse que la distribucion de W'es:

$$(1-\pi).\exp.\left\{-(1-\pi).w\right\} \qquad ,, \quad w>0$$

Lo que si puede calcularse explicitamente sin dificultad son los momentos de W';asi,por derivacion:

$$E(W')=(1-\pi)^{-1}$$

$$V(W'^2)=\frac{f''(1)}{\mu.(\mu-1).(1-\pi)}$$

1.4.EL PROCESO DE NACIMIENTO PURO LINEAL COMO UNO DE GALTON-WATSON.-

Sea X_ξ un proceso de nacimiento puro lineal de intensidad λ ,y supongamos que es observado en los tiempos 0,z,2z,.....,kz=t ,en donde k es un entero,t el tiempo de última observacion y z un positivo dado.Sean Z_n las observaciones en dichos tiempos,de modo que :

$$Z_n = X_{nz} \qquad ,, \quad n=0,1,2,\ldots\ldots,k.$$

entonces,dichas observaciones Z_n forman un proceso de

Galton-Watson con distribucion de hijos la geometrica:

$$P(Z_1=i\ /Z_o=1)=e^{-\lambda z}.(1-e^{-\lambda z})^{i-1}\ ,,\ i=1,2,\ldots.$$

$$P(Z_1=0\ /\ Z_o=1)=\ 0$$

Con ello:

$$E(Z_1\ /\ Z_o=1)=\ e^{\lambda z}$$

$$V(Z_1\ /Z_o=1)\ =e^{\lambda z}.(e^{\lambda z}-1)$$

son la media y varianza de la distribucion de hijos.
(Harris,pg.101).

\- \- -

CAPITULO II

ESTIMACION PARAMETRICA EN PROCESOS DE NACIMIENTO PURO.

2.1.INTRODUCCION.-

Sea X_t un proceso de Markov em el cual para i=1, 2,3,...... se tiene:

$$P(X_{t+h} =j / X_t=i) = \begin{cases} i\lambda h + o(h) & \text{si } j=i+1 \\ 1 - i\lambda h + o(h) & \text{si } j=i \\ o(h) & \text{en otro caso.} \end{cases}$$

en donde λ es un parámetro desconocido y tal que $\lambda > 0$.Supongamos ademas que $P(X_0=q)=1$ en donde q es un entero positivo fijado de antemano y conocido o no.A un tal proceso es a lo que se le llamó en el capítulo I por proceso de nacimiento puro lineal,de modo que X_t suele ser denominado por "tamaño de la población en el tiempo t".

Unproblema abordado por diversos autores ha sido el de la estimacion de la intensidad λ bajo diversas esquemas de muestreo.La sección 2.2. se ocupa de resumir brevemente todo lo mas importante que hay dicho al respecto.Asimismo en la sección 2.3. se dan algunas ideas sobre la estimación de λ .

Sin embargo, el problema de estimar la edad de un proceso de nacimiento puro o el número inicial de individuos de que parte el proceso, no ha sido aún tratado y son problemas que asimismo abordamos en la sección 2.3.

Como quiera que las anteriores estimaciones pueden hacerse bajo diversos esquemas de muestreo, y ellos seran citados a menudo en el texto, es conveniente que los llamemos brevemente de algún modo. Asi, denominaremos por esquema de muestreo de tipos I, II, III, IV y V a los siguientes:

TIPO I : observaciones continuadas del proceso X en el intervalo de tiempo-fijado de antemano- $[0,t]$.

TIPO II : observaciones continuadas del proceso desde el tiempo 0 hasta el tiempo T en que el estado p - fijado - es alcanzado por primera vez.

TIPO III: observaciones equidistantes en los tiempos 0, z, 2z, ..., nz con z fijo y n un entero tambien fijo(muestreo equidistante).

TIPO IV : una única observación X_t en un tiempo t fijado de antemano.

TIPO V : observación del tiempo T en que el estado p es ocupado por primera vez.

Nótese que el muestreo de tipo IV es un caso particular del muestreo de tipo III cuando z=t y n=1.

2.2.RESULTADOS CONOCIDOS.-

2.2.1.ESTIMACION DE λ BAJO MUESTREO DE TIPO I.-

Este problema ha sido abordado por diversos autores,pero Keiding(1974) da un resumen de todo lo conocido -salvo en el caso del muestreo de tipo II - y aporta nuevas soluciones.Las pruebas y referencias de todo lo que sigue pueden verse en dicho artículo.

Veamos primero algunos resultados generales a los que haremos referencia posteriormente.Asi,son importantes los siguientes teoremas:

Teorema 2.2.1.- Condicionalmente en W, X_t es un proceso de Poisson de tiempo homogeneo con X_o=q e intensidad $qW\lambda e^{\lambda t}$;con ello:

$$E(X_t-q / W) = qW (e^{\lambda t} - 1)$$

en donde W es la variable a la cual converge casi seguro, cuando t tiende hacia infinito,la variable $X_t/ E(X_t)$ como se indico en el capítulo anterior,y con X_t un proceso de nacimiento de intensidad λ y X_o=q.

Teorema 2.2.2.- Sea $\{X_t,t\geq 0\}$ un proceso de Poisson de intensidad $qw\lambda e^{\lambda t}$ ($\lambda>$ 0) y X_o=q,y sea

$$S_t= \int_o^t X_u du$$

Entonces,casi seguramente,cuando t $\longrightarrow\infty$:

(a) $X_t.e^{-\lambda t} \longrightarrow$ qw

(b) $S_t.e^{-\lambda t} \longrightarrow qw/\lambda$

En particular, para el proceso de naxcimiento, podemos concluir que :

$$q^{-1}.e^{-\lambda t}.\int_0^t X_u du \longrightarrow W/\lambda \quad \text{c.s.}$$

A continuación veremos dos procedimientos de estimación de λ, uno en el proceso X_t y el otro en el proceso condicional en W.

2.2.1.1.INFERENCIA EN EL RPOCESO DE NACIMIENTO.-

La distribución de $\{X_u, 0 \leq u \leq t\}$ está plénamente determinada por X_t y los tiempos aleatorios $T_{q+1},\ldots\ldots,T_{X_t}$, en donde $T_i=\inf.\{u \,/\, X_u=i\}$ es el tiempo en que el proceso salta al estado i. Entonces :

<u>Teorema 2.2.3</u>.- La función de verosimilitud esta dada por:

$$L(\lambda)=(X_t-1)^{(X_t-q)}.\lambda^{X_t-q}.e^{-\lambda S_t}$$

en donde:

$$S_t=\int_0^t X_u du \;,,\; a^{(x)}=a.(a-1)\ldots\ldots(a-x+1)$$

Ademas, (X_t,S_t) es minimal suficiente y el estimador de máxima verosimilitud esta dado por :

$$\hat{\lambda} = \frac{X_t - q}{S_t}$$

Teorema 2.2.4.-

(a) La distribución de S_t, cuando $X_t=x$, tiene por función caracteristica:

$$E(\ e^{ivS_t}\ /\ X_t=x\ )=e^{ivqt}\cdot\left\{\frac{1-e^{(iv-\lambda)t}}{(1-\frac{iv}{\lambda})(1-e^{-\lambda t})}\right\}^{x-q}$$

y densidad:

$$(\lambda t)^{x-q}\cdot e^{-\lambda(s-qt)}\cdot(1-e^{-\lambda t})^{q-x}\cdot g_{x,q}(s)$$

con $qt+Y_1+\dots+Y_{x-q}$ de igual densidad que $g_{x,q}(s)$ en donde Y_i son variables aleatorias independientes y uniformemente distribuidas en el intervalo $[0,t]$,y con $qt\leq s<\infty$.

(b)La distribución de S_t , dado W y $X_t=x$, es la misma que bajo(a).

(c)La distribución del minimal suficiente (X_t,S_t) tiene por función caracteristica :

$$\left\{E\ (\ e^{iuX_t+ivS_t})\right\}^{1/q}=$$

$$=\frac{(1-\frac{iv}{\lambda})\cdot e^{iu+(iv-\lambda)t}}{1-\frac{iv}{\lambda}-e^{iu}+e^{iu+(iv-\lambda)t}}$$

y densidad:

$$\binom{x-1}{q-1}(\lambda t)^{x-q}\cdot e^{-\lambda s}\cdot g_{x,q}(s)$$

con $x=q,q+1,\ldots\ldots,qt \leq s<\infty$.

Teorema 2.2.5.- Cuando $t \longrightarrow \infty$. $\hat{\lambda} \longrightarrow \lambda$ casi seguro.

Teorema 2.2.6.- Sea $\{X_t, t \geq 0\}$ un proceso de Poisson de intensidad $qw\lambda e^{\lambda t}$ y $X_o=q$. Entonces la distribución de

$$(A_t, \lambda . e^{-\lambda t} . S_t) = \left\{ e^{-\lambda t/2}(X_t - q - \lambda S_t), e^{-\lambda t} S_t \lambda \right\}$$

converge débilmente, cuando $t \longrightarrow \infty$,hacia la distribución de (A, qw) en donde A es $N(0,qw)$.

Teorema 2.2.7.- Cuando $t \longrightarrow \infty$ y q esta fijo:

(a) $(\lambda S_t)^{1/2} . (\hat{\lambda}/\lambda - 1)$ es asintóticamente $N(0,1)$

(b) $(q . e^{\lambda t})^{1/2} . (\hat{\lambda}/\lambda - 1)$ converge débilmente hacia una distribución de Student con 2q grados de libertad.

Teorema 2.2.8.- Para un t fijo y $q \longrightarrow \infty$:

(a) $(\lambda S_t)^{1/2} . (\hat{\lambda}/\lambda - 1)$ es asintóticamente $N(0,1)$.

(b) $(q . e^{\lambda t})^{1/2} . (\hat{\lambda}/\lambda - 1)$ es asintóticamente $N(0, 1/(1 - e^{-\lambda t}))$.

Observese el comportamiento del estimador $\hat{\lambda}$ de λ. Los resultados del teorema 2.2.7. pueden obtenerse a partir de los teoremas 2.2.1. y 2.2.6. El teorema 2.2.8. es mas sencillo de obtener pues si X_o=q,entonces X_t es la suma de q procesos de nacimiento independientes y de igual intensidad λ ,y basta aplicar entonces el teorema central del límite.

Observese que si en el teorema 2.2.7. (b) hacemos que q $\longrightarrow \infty$ obtenemos una N(0,1),al igual que si en el teorema 2.2.8.(b) hacemos t $\longrightarrow \infty$,lo cual era lógico de esperar.

2.2.1.2.-INFERENCIA CONDICIONAL.-

Veamos de aprovecharnos,para la estimacion de λ , de lo indicado en el teorema 2.2.1. pero haciendo intervenir el límite w condicionante.

<u>Teorema 2.2.9</u>.- Sea $\{Z_t, t \geq 0\}$ un proceso de Poisson con intensidad exp.$(\mu + \lambda t)$, $(\lambda, \mu) \in R^2$ y sea

$$R_t = \int_o^t Z_u du.$$

Para estimar (λ, μ) la funcion de verosimilitud la da:

$$\ln L(\lambda, \mu) = (Z_t - q)\mu + (t.Z_t - R_t)\lambda - \frac{(e^{\lambda t} - 1)e^{\mu}}{\lambda}$$

y el estadístico minimal suficiente es

(Z_t, R_t). Ademas, excepto cuando $Z_t=q$ (y por tanto $tZ_t-R_t=0$), el estimador de máxima verosimilitud lo da la única solución de las ecuaciones de verosimilitud:

$$Z_t-q= E(Z_t-q)= \frac{e^{\mu}.(e^{\lambda t}-1)}{\lambda}$$

$$tZ_t-R_t=E(tZ_t-R_t)= \frac{e^{\mu}(e^{\lambda t}.\lambda.t+1-e^{\lambda t})}{\lambda^2}$$

denominando a dicha solución por (λ^*, μ^*).

Finalmente:

$$\lambda^* > 0 \Longleftrightarrow 2(tZ_t-R_t) > t(Z_t-q) \Longleftrightarrow$$

$$\Longleftrightarrow 2(R_t-qt) < t(Z_t-q)$$

Con ello, si Z_t es el proceso de Poisson condicional obtenido de un proceso de nacimiento X_t de intensidad $\lambda > 0$ y $X_0=q$ condicionado al limite, casi seguro, $W=\lim. X_t/E(X_t)=w$, entonces Z_t tiene de intensidad $qw\lambda e^{\lambda t}$, de modo que haciendo $\mu=\ln.(qw\lambda)$, el teorema 2.2.9. se vuelve aplicable a nuestro estudio.

<u>Teorema 2.2.10</u>.- Supongamos que $\lambda > 0$ en el teorema 2.2.9. Entonces, cuando $t \dashrightarrow \infty$:

(a) $\lambda^* \dashrightarrow \lambda$ casi seguro.

(b) $(\lambda^*, \mu^*) \dashrightarrow (\lambda, \mu)$ en probabilidad.

(c) $(\lambda R_t)^{1/2} \circ (\lambda^*/\lambda - 1)$ es asintóticamente $N(0,1)$.

(d) $(e^{\lambda t})^{1/2} . (\lambda^*/\lambda - 1)$ es asintóticamente $N(0, \lambda e^{-\mu})$.

Si se comparan los teoremas 2.2.5. y 2.2.6.con el actual 2.2.10. se verá que cuando $t \dashrightarrow \infty$,las distribuciones asintóticas de $\hat{\lambda}$ y λ^* son idénticas,unas para el proceso de nacimiento y la otra en el proceso de Poisson condicionado.

2.2.2.ESTIMACION DE λ BAJO MUESTREO DE TIPO II.-

Sea $\{X_t, t \geq 0\}$ un proceso de nacimiento puro de intensidad λ que comienza con $X_0=q$ individuos.Supongamos que dicho proceso es observado contínuamente hasta la aparición de k -fijado de antemano - nuevos individuos;sea T el tiempo en que aparece el imdividuo número $(q+k)$.Denominemos por $t_1,t_2,\ldots\ldots,t_k$ a los tiempos de primera ocupación de los estados $(q+1),(q+2),\ldots\ldots,(q+k)$;con ello los t_i seran variables aleatorias que tomaran distintos valores en cada realizacion del proceso.Nótese que en el muestreo de tipo I el tiempo T estaba fijado de antemano,mientras que ahora esta fijado k;con ello,en un caso la variable aleatoria sera k mientras que en el otro lo sera el tiempo T.

En estas condiciones,es conocido que las variables aleatorias

$$d_{s+1} = t_{s+1} - t_s \quad ,, \quad s=0,1,2,\ldots\ldots,(k-1)$$

son independientes unas de otras y estan distribuidas como una exponencial de parámetro $(q+s)\lambda$,de modo que el logaritmo de la verosimilitud es:

$$\ln L = \sum_{s=0}^{k-1} \ln . \left\{ \lambda (q+s).e^{-\lambda (q+s).d_{s+1}} \right\} =$$

$$= \sum_{s=0}^{k-1} \left\{ \ln \lambda + \ln(q+s) - \lambda (q+s).d_{s+1} \right\}$$

en donde naturalmente $d_1 = t_1$.Derivando respecto a λ e igualando a cero (ver artículo de Moran) entonces :

$$\frac{\partial \ln . L}{\partial \lambda} = k. \lambda^{-1} - \sum_{s=0}^{k-1} (q+s).d_{s+1} = 0$$

de modo que el estimador $\hat{\lambda}$ de λ que se propone es:

$$\hat{\lambda} = \frac{k}{\sum_{s=0}^{k-1} (q+s).(t_{s+1} - t_s)} \qquad (2.1.)$$

en donde el estimador tiene por denominador,de hecho,el area bajo la curva en el grafo de X_t;con ello:

$$\hat{\lambda} = \frac{X_T - X_o}{\int_o^T X_t dt}$$

estimador que tiene la misma forma que el obtenido en la sección anterior - ver teorema 2.2.3. - si bien sus distribuciones son lógicamente distintas,pues en uno T es variable y en el otro lo es k. Vease el articulo de Moran para una discusión sobre la varianza de ambos estimadores.

Si en ambos lados de la expresion (2.1) tomamos la inversa, entonces:

$$y = \frac{1}{k} \cdot \sum_{s=0}^{k-1} (q+s).d_{s+1}$$

es un estimador insesgado de λ^{-1}. Para ver su distribución, como d_{s+1} seguia una exponencial de parámetro $(q+s)\lambda$, entonces la variable aleatoria

$$2\lambda(q+s).d_{s+1}$$

sigue una X^2 con 2 grados de libertad y por tanto "y" seguirá una $(2k\lambda)^{-1}.X^2$ con 2k g.l., de modo que su varianza sera $(k.\lambda^2)^{-1}$. Tambien, para grandes $(q+k)$, $\hat{\lambda}$ es un estimador de λ que está distribuido aproximádamente como una normal de varianza $\lambda^2.k^{-1}$.

2.2.3. ESTIMACION DE λ BAJO MUESTREO DE TIPO III.-

Sea X_ξ un proceso de nacimiento observado en los tiempos $0, z, 2z, \ldots, nz=t$ y sea $Z_k=X_{kz}$. Si X_ξ tiene de intensidad λ, es conocido que entonces el proceso Z_k es uno de Galton-Watson con distribución de hijos la geométrica, esto es :

$$P(\ Z_1=i \ / \ Z_o=1\ ) = e^{-\lambda z}.(1- e^{-\lambda z})^{i-1} \quad \text{si } i=1,2,\ldots$$

$$P(Z_1=0 \ / \ Z_o=1)=0$$

y por tanto la media y varianza de $(Z_1/Z_o=1)$ es :

$$E(\ Z_1 \ / \ Z_o=1\ ) = e^{\lambda z}=\mu$$

$$V(\ Z_1 \ / \ Z_o=1\ ) = e^{\lambda z}.(\ e^{\lambda z} -1)$$

Los dos teoremas siguientes - ver el artículo de Keiding - dan los resultados apetecidos para el problema de estimar λ :

<u>Teorema 2.2.11</u>.- La función de verosimilitud es:

$$L(\lambda)=\prod_{k=1}^{n}\begin{pmatrix} X_{kz} -1 \\ X_{kz} - X_{(k-1)z}\end{pmatrix}.(1-e^{-\lambda z})^{X_{nz}-X_o}.$$

$$.\exp.\left\{-\lambda z \sum_{k=1}^{n} X_{(k-1)z}\right\}$$

y el estimador de maxima verosimilitud es:

$$\tilde{\lambda} = \frac{1}{z} .\ln\left(\frac{X_z + \ldots\ldots\ldots + X_{nz}}{X_o + \ldots\ldots\ldots + X_{(n-1)z}}\right)$$

Nótese que en definitiva la fraccion puesta entre paréntesis es el estimador de Harris para la media μ

de un proceso de Galton-Watson, o sea, el estimador de $e^{\lambda z}$, y puesto que z es conocido, despejando λ obtenemos tambien el estimador anterior.

Teorema 2.2.12.- Cuando $n \longrightarrow \infty$:

(a) $\tilde{\lambda} \longrightarrow \lambda$ casi seguro y ademas:

$$E(e^{\tilde{\lambda} z}) \longrightarrow e^{\lambda z}$$

(b) $z(X_0 + \ldots\ldots + X_{nz})^{1/2} . (\tilde{\lambda} - \lambda)$ es asintóticamente $N(0,\ 1 - e^{-\lambda z})$.

(c) $\left\{ q.(e^{\lambda(n+1)z} - 1).e^{\lambda z} \right\}^{1/2} . \dfrac{\tilde{\lambda} - \lambda}{e^{\lambda z} - 1} . z$

converge débilmente hacia una distribución de Student con 2q grados de libertad.

En particular, cuando z tiende hacia cero y n tiende hacia infinito de modo que nz tienda hacia t, entonces se prueba que no solo $\tilde{\lambda} \longrightarrow \hat{\lambda}$ sino que tambien las distribuciones asintóticas son asintóticamente iguales.

2.2.4. ESTIMACION DE λ BAJO MUESTREO DE TIPO IV.-

Ahora solo contamos con la única observación X_t. Ello es un caso particular de la estimacion bajo el muestreo de tipo III en que n=1 y z=t. La forma del estimador de λ surge por particularización del teorema 2.2.10., pero ahora las distribuciones límites y la convergencia del teorema

2.2.11. no nos valen, puesto que n no tiende hacia infinito al ser n=1 fijo. La convergencia límite puede darse ahora a partir de la convergencia casi segura ya citada :

$$X_t \cdot q^{-1} \cdot e^{-\lambda t} \longrightarrow W$$

Asi tenemos ahora el :

<u>Teorema 2.2.13</u>.- El estimador λ^+ de maxima verosimilitud de λ esta dado por :

$$\lambda^+ = \frac{1}{t} \cdot \ln \frac{X_t}{X_o} \qquad (2.2)$$

Ademas, cuando $t \longrightarrow \infty$, entonces:

$$t(\lambda^+ - \lambda) \longrightarrow \ln W$$

en donde W se distribuye segun una gamma (q,q) de modo que :

$$E(\ln W) = \psi(q) - \ln q$$

$$V(\ln W) = \psi'(q)$$

siendo ψ la funcion digamma:

$$\psi(x) = d \ln \Gamma(x) \,/\, dx$$

Nótese que :

$\psi(q) - \ln q < 0$ pero $\psi(q) - \ln q \longrightarrow 0$ cuando $q \longrightarrow \infty$ y ademas $q \cdot \psi'(q) \longrightarrow 1$ cuando $q \longrightarrow \infty$.

En particular :

$$\psi(1) = -\gamma \cong -0{,}577$$

$$\psi'(1) = \pi^2/6 \simeq 1{,}645$$

Finalmente, cuando $q \longrightarrow \infty$, entonces :

$$\sqrt{q}\ (\lambda^{+}-\lambda) \text{ es asintóticamente } N\left(0, \frac{\{2sh.(\lambda t/2)/t\}^{2}}{e^{\lambda t}-1}\right)$$

2.3.APORTACIONES.-

2.3.1.EL PROCESO DE NACIMIENTO LINEAL PURO Y EL PROCESO CONJUGADO.-

Sea $\{X_t, t \geqslant 0\}$ un proceso de nacimiento puro de intensidad λ ;para identificación indicaremos por $X_t^{(\lambda)}$ a tal proceso,de modo que el superindice entre paréntesis indica la intensidad del proceso,mientras que el subindice indica el prámetro tiempo.Es conocido que :

$$P_{m,n}^{(\lambda)}(t) = P\left\{X_{s+t}^{(\lambda)} = n \ / \ X_s^{(\lambda)} = m\right\} = \binom{n-1}{n-m}.(e^{-\lambda t})^m.(1-e^{-\lambda t})^{n-m}$$

para $n \geqslant m \geqslant 1$ y 0 en otro caso.

Sea ahora el proceso Y_η ,tambien de nacimiento puro,pero ahora de intensidad t;en el :

$$P_{m,n}^{(t)}(\lambda) = P\left\{Y_{s+\lambda}^{(t)} = n \ / Y_s^{(t)} = m\right\} = \binom{n-1}{n-m}.(e^{-t\lambda})^m.(1-e^{-t\lambda})^{n-m}$$

Con ello:

$$P_{m,n}^{(\lambda)}(t) = P_{m,n}^{(t)}(\lambda)$$

Si ahora denominamos por <u>proceso original</u> al X_ξ y por <u>proceso conjugado</u> al Y_η ,concluiremos en qme,probabilisticamente,es lo mismo hablar de un proceso de nacimiento de intensidad λ que comenzando con m individuos se encuentra en el estado n despues de un tiempo t,que de un proceso de nacimiento de intensidad t,que partiendo de m individuos se encuentra en el estado n despues de un tiempo λ .Ello ha sido asi por el papel simétrico que juegan λ y t en $P_{m,n}^{(\lambda)}(t)$, y nos habla de la existencia de un único proceso de nacimiento de intensidad 1 y un único parámetro tiempo que seria el producto $\lambda . t$.

2.3.2.ESTIMACION DEL TIEMPO z ENTRE OBSERVACIONES EQUIDISTANTES.-

Sea X_ξ un proceso de nacimiento de intensidad λ conocida del cual se hacen las siguientes observaciones: $X_0, X_z, X_{2z}, \ldots\ldots, X_{nz}$,pero en el cual no se sabe el intervalo z entre observaciones.Si ahora consideramos el proceso conjugado Y_η de intensidad z al cual observamos en los tiempos 0, λ ,2λ ,.....,nλ ,el es probabilisticamente identico al anteior,de modo que podemos hablar de un proceso de intensidad z desconocida y que al abservarlo de λ en λ da los tamaños anteriores,o sea :

$$Y_{k\lambda} = X_{kz}$$

Con ello, los teoremas 2.2.11. y 2.2.12. son aplicables aqui pero cambiando los papeles de λ y z. En particular :

$$\tilde{z} = \frac{1}{\lambda} \cdot \ln \left(\frac{X_z + X_{2z} + \cdots\cdots + X_{nz}}{X_o + X_z + \cdots\cdots\cdots + X_{(n-1)z}} \right)$$

2.3.3. ESTIMACION DE λ POR MUESTREO DE TIPO II.-

Sea $\{X_t, t \geq 0\}$ un proceso de nacimiento puro de intensidad λ que es observado desde el tiempo 0 hasta un tiempo T - que es variable aleatoria - en que el estado (q+k) es alcanzado por primera vez, en donde X_o=q. Denominemos por $t_1, t_2, \ldots, t_k$=T a los tiempos de primera estancia en los estados (q+1),(q+2),....,(q+k), con lo cual:

$$X_{t_i} = q+i$$

y sean $d_i = t_{i+1} - t_i$ los tiempos entre nacimientos sucesivos, en donde naturalmente $d_o = t_1$. Con ello, por el proceso conjugado :

$$P^{(\lambda)}_{q+r,q+r+1}(d_r) = P^{(d_r)}_{q+r,q+r+1}(\lambda) \;,, \; r=0,1,\ldots,(k-1)$$

de modo que si en proceso original d_r sigue una exponencial de parámetro $(q+r)\lambda$, en el conjugado λ_r - el tiempo hasta la aparición del primer nacimiento desde el individuo $(q+r)$ en un proceso de nacimiento puro de intensidad d_r - seguirá una exponencial de parámetro $(q+r).d_r$. Esto es, la variable aleatoria (λ_r / d_r) tiene por densidad:

$$(q+r)d_r \cdot e^{-(q+r)d_r \cdot x}$$

de modo que podria tentarnos la idea de obtener, por muestreo de Monte Carlo, el tiempo λ_r que tarda el proceso de nacimiento, de intensidad d_r dada, en ir desde el estado $(q+r)$ al estado $(q+r+1)$ y considerar el valor λ_r obtenido como una estimada de λ. El problema es que de hacerlo asi, entonces (haciendo $q=1$ para simplificar) :

$$P(\lambda_r = x)dx = \int_0^\infty P(d_r = y) \cdot P(\lambda_r = x \,/\, d_r = y)\,dy\,dx =$$

$$= \lambda r^2 \cdot \left[\int_0^\infty y \cdot e^{-y(\lambda r + xr)} \cdot dy \right] dx =$$

$$= \frac{\lambda}{(\lambda + x)^2}\, dx$$

con lo cual

$$E(\lambda_r) = \infty$$

de modo que la doble aleatorización - la de d_r y la de λ_r -

nos es perjudicial. Alternativamente, podemos tomar como valor de λ_r el valor medio que se obtendria de utilizar, como se dijo mas arriba, el proceso conjugado, esto es:

$$\hat{\lambda}_r = E(\lambda_r / d_r) = \frac{1}{(q+r).d_r}$$

de modo que $\hat{\lambda}_r$ seria la estimada propuesta de λ. Para facilitar el cálculo de un estimador conjunto de λ y de su distribución, es conveniente el considerar la variable:

$$y_r = \frac{1}{\hat{\lambda}_r} = (q+r).d_r$$

que es un estimador insesgado de λ^{-1}, pues en efecto:

$$E(y_r) = (q+r).E(d_r) = \frac{(q+r)}{(q+r).\lambda} = \frac{1}{\lambda}$$

Como ademas todos las y_r tienen igual varianza:

$$V(y_r) = (q+r)^2.V(d_r) = \frac{1}{\lambda^2}$$

entonces un estimador conjunto de λ^{-1} seria:

$$\hat{y} = \frac{\sum_{r=o}^{k-1} y_r}{k} = \frac{\sum_{r=o}^{k-1} (q+r).d_r}{k}$$

que es insesgado y de varianza:

$$V(\hat{y}) = (k.\lambda^2)^{-1}$$

Con ello:

$$V(\hat{y}) = \frac{1}{\lambda^2 . k} \longrightarrow 0$$

cuando $k \longrightarrow \infty$, de modo que:

$$\hat{y} \longrightarrow \lambda^{-1} \text{ ,cuando } k \longrightarrow \infty \text{ ,en probabilidad}$$

con lo que $\hat{y}$ es consistente de λ^{-1}.

Como y_r sigue una exponencial de parámetro λ, entonces $\hat{y}$ sigue una gamma$(k, \lambda k)$ y, finalmente, $\hat{y}\lambda$ seguirá una gamma(k,k). Con ello todos los resultados actuales obtenidos mediante el proceso paralelo, coinciden con lo indicado en la seccion 2.2.2. Finalmente, para k grande, $\hat{y}\lambda$ seguirá una normal de media 1 y varianza k^{-1}, de donde se deduce que un intervalo de confianza para λ -en el caso de grandes k - será:

$$\lambda \in \frac{1}{\hat{y}} \pm \frac{t_\alpha}{\sqrt{k} . \hat{y}}$$

con t_α mirado en la $N(0,1)$.

2.3.4. ESTIMACION DE λ Y t POR MUESTREO DE TIPO IV.-

La expresión (2.1) del teorema 2.2.13. nos daba la estimación de la intensidad λ de un proceso de nacimiento bajo el muestreo de tipo IV, a saber:

$$\lambda^{+} = \frac{1}{t} . \ln \frac{X_t}{X_o}$$

asi como la convergencia cuando $t \longrightarrow \infty$.

Consideremos ahora el problema de estimar la edad t de un proceso de nacimiento de intensidad λ conocida, cuando observado en el tiempo t da un tamaño X_t. Ello es equivalente a estimar la intensidad t de un proceso de nacimiento-el proceso conjugado-que en el tiempo λ tiene el tamaño X_t; por el estimador anterior:

$$t^{+} = \frac{1}{\lambda} . \ln \frac{X_t}{X_o} \qquad (2.3.)$$

que es tambien de máxima verosimilitud y le es aplicable el mismo teorema 2.2.13. con el consiguiente intercambio de λ y t. En efecto, como:

$$\frac{X_t}{E(X_t)} \longrightarrow W \quad \text{c.s. cuando } t \longrightarrow \infty$$

entonces:

$$\ln \frac{X_t}{X_o} - \lambda t \longrightarrow \ln W \quad \text{c.s. cuando } t \longrightarrow \infty$$

y finalmente:

$$\lambda (t^{+} - t) \longrightarrow \ln W \quad \text{c.s. cuando } t \longrightarrow \infty \qquad (2.4)$$

en donde W es una gamma(q,q), con $X_o = q$ como hipótesis inicial.

De otro lado, cuando el objeto de nuestro estudio sea estimar la edad t, como es el caso, sera usual el desconocer $X_0=q$, de modo que no quedará otro remedio que suponer $X_0=q=1$. En ese caso W seguira una gamma(1,1) y por tanto su densidad será e^{-w}. Consecuentemente, la densidad de lnW será:

$$e^{-e^y} \cdot e^y$$

y por tanto:

$$P(\ln W < w) = -e^{-e^w} + 1$$

Supongamos que se desea un intervalo para lnW; tomando un error α, sean α_1 y α_2 los dos errores de cola y tales que $\alpha_1 + \alpha_2 = \alpha$. Entonces:

$$P(\ln W < w_1) = \alpha_1 = -e^{-e^{w1}} + 1 \Longrightarrow \alpha_1 = 1 - e^{-e^{w1}}$$

$$P(\ln W < w_2) = 1 - \alpha_2 = -e^{-e^{w2}} + 1 \Longrightarrow \alpha_2 = e^{-e^{w2}} \qquad (2.5)$$

Si finálmente deseamos el intervalo de longitud mínima para un error dado α, el valor α_1 que hay que tomar lo da la única raiz positiva de la igualdad:

$$(\alpha - \alpha_1)\ln(\alpha - \alpha_1) = (1-\alpha_1)\ln(1-\alpha_1)$$

de modo que obtenido α_1 entonces $\alpha_2 = \alpha - \alpha_1$ y finálmente w_1 y w_2 se obtendran de las igualdades (2.5), es decir:

$$w_1 = \ln(\, -\ln(1-\alpha_1)\,)$$

$$w_2 = \ln(-\ln \alpha_2)$$

intervalo que se traslada a t inmediátamente por la (2.4) y que solo vale,como hemos venido suponiendo,para cuando t es razonáblemente grande y de un modo solo aproximado.

Veamos otro modo aún de obtener idénticos resultados. Ya se ha indicado repetidas veces que un proceso de nacimiento puede ponerse como uno de Galton-Watson considerando el proceso:

$$Z_k = X_{kz} \quad ,, \quad k=0,1,2,\ldots\ldots,n \quad ,, \quad nz=t$$

Si suponemos que $X_o=Z_o=1$,entonces:

$$P(Z_1=i \;/\; Z_o=1)=e^{-\lambda z}.(1-e^{-\lambda z})^{i-1} \quad ,,i=1,2,\ldots\ldots$$

$$P(Z_1=0 \;/\; Z_o=1) = 0$$

de modo que :

$$E(Z_1 \;/\; Z_o=1)=e^{\lambda z}=\mu > 1 \quad \text{pues } \lambda > 0$$

con lo cual la función generatriz de hijos es :

$$f(s) = \frac{s}{(1-s).e^{\lambda z} + s}$$

y por tanto π (la probabilidad de extinción) es 0.Como f(s) es del tipo fraccional lineal,si admitimos el estimador de la edad n de un proceso de Galton-Watsom supercrítico(ver sección 3.1.1.1.) entonces:

$$\hat{n} = \frac{\ln(X_t.(1-\pi) + \pi)}{\ln \mu} = \frac{\ln X_t}{\lambda . z}$$

de modo que $\hat{n}=t^{+}$ si se hace $z=1$,y los estimadores son coincidentes.Para ello no hace falta suponer que $X_o=1$;en efecto,cuando $X_o=q>1$,basta fijarnos en la estimacion de n propuesta por nosotros en la sección 3.2.5.,a saber :

$$\hat{n} = \frac{\ln \left\{ \frac{X_t(1-\pi) + \pi}{X_o(1-\pi) + \pi} \right\}}{\ln \mu} = \frac{1}{\lambda . z} . \ln X_t \ / \ X_o$$

para que de nuevo $\hat{n}=t^{+}$ tomando $z=1$.

Volviendo al caso mas usual en que se desconoce X_o y por tanto se le supone igual a uno,los resultados de Stigler para $\hat{n}$ se trasladan de inmediato a t^{+} (ver de nuevo la seccion 3.1.1.1.).En particular t^{+} es α -consistente para todo $\alpha > 0$ en el sentido de Feldman and Fox(1968);es decir:

$$\frac{t^{+} - t}{t^{\alpha}} \longrightarrow 0 \text{ débilmente cuando } t \longrightarrow \infty$$

ademas,los resultados límites e intervalos de la seccion antes indicada siguen valiendo ahora y son coincidentes con los ya expuestos mas arriba.Asi,el intervalo para n con error $\alpha = \alpha_1 + \alpha_2$ es :

$$\hat{n} - \frac{\ln(-\ln \alpha_2)}{\lambda} \quad ; \quad \hat{n} - \frac{\ln(-\ln(1-\alpha_1))}{\lambda}$$

que es el mismo citado antes. Asimismo se indica en dicha sección 3.1.1.1. que, la probabilidad de cobertura de n de dicho intervalo es $(1-\alpha)$ solo en el límite; la probabilidad real es la cantidad P(n) que alli se especifica, cantidad que converge rápidamente hacia $(1-\alpha)$ cuando μ no esta próximo a 1, esto es, cuando λ no esta próximo a 0. Ademas, cuando α =0,05 , entonces :

$$P(n) \geq 0,95 \quad ,, \quad \forall n$$

de modo que el valor nominal dado, es conservativo en todo caso.

Notemos finálmente que admitido el estimador de Stigler $\hat{n}$ anterior, y usando el proceso conjugado, llegaremos a un estimador de λ que es obviamente el mismo dado por el teorema 2.2.13. La construccion y pruebas son las indicadas arriba, intercambiando los papeles de λ y t .

2.3.5. ESTIMACION DE λ POR MUESTREO DE TIPO V.-

Sea X_ξ un proceso de nacimiento puro de intensidad λ que comienza con X_0=q individuos. Supongamos que tal proceso no se ha observado -por alguna razón- hasta el momento en que el tamaño de la población llegó a k individuos, en donde k esta fijado de antemano. Sea T el tiempo de primera entrada en el estado k, de modo que por tanto X_T=k.

Si denominamos por T_i a los tiempos de estancia en los estados i con $i=q,(q+1),\ldots,(k-1)$, es conocido que las T_i son variables aleatorias, independientes unas de otras, de densidad una exponencial de parámetro λi, de modo que :

$$T = T_q + T_{q+1} + \cdots\cdots + T_{k-1}$$

y por tanto :

$$\hat{T} = E(T / X_o=q) = E(T_q + T_{q+1} + \cdots + T_{k-1}) =$$

$$= \frac{1}{\lambda} \cdot \sum_{h=q}^{k-1} \frac{1}{h} =$$

$$\simeq \begin{cases} \dfrac{1}{\lambda} \cdot \ln \dfrac{k}{q} & \text{si } K/q \text{ y } q \text{ son grandes.} \\[2ex] \dfrac{\gamma}{\lambda} + \dfrac{\ln(k-1)}{\lambda} - \dfrac{1}{\lambda} \cdot \displaystyle\sum_{h=1}^{q-1} \frac{1}{h} & \\ \qquad \text{si } K \text{ es grande y } q \text{ es pequeño.} & \end{cases}$$

siendo $\gamma \simeq 0{,}5772$ la constante de Euler. Con ello $\hat{T}$ puede ser una apropiada estimación de T cuando T sea desconocido. En particular, cuando no conozcamos X_o habremos de suponerlo igual a 1, con lo cual:

$$\hat{T} = \frac{\gamma}{\lambda} + \frac{\ln(k-1)}{\lambda} \qquad \text{si k es grande.}$$

Lo mas corriente sera desconocer λ y conocer T, de modo que cuando asi ocurra - por el proceso conjugado -

se tiene :

$$\hat{\lambda} \simeq \begin{cases} \frac{1}{T} \cdot \ln \frac{k}{q} & \text{si } q \text{ y } k/q \text{ son grandes} \\ \frac{\gamma}{T} + \frac{\ln(k-1)}{T} - \frac{1}{T} \sum_{h=1}^{q-1} \frac{1}{h} & \text{si } q \text{ es pequeño y } k \text{ grande.} \end{cases}$$

Para dar intervalos de confianza para λ (o para T) recuérdese que en la seccion 1.2.2. se indico que la variable aleatoria $\lambda T - \ln k$ (con k fijo)tenia por distribucion limite-cuando $k \longrightarrow \infty$ -una exp.$(-x-e^{-x})$ con $-\infty < x < +\infty$. Al igual que en 2.3.3., $\hat{\lambda}^{-1}$ es consistente de λ^{-1}.

Notese que cuando q y k/q son grandes,entonces $\hat{\lambda}$ tiene igual forma que el (2.2).

2.3.6. ESTIMACION DEL NUMERO INICIAL DE INDIVIDUOS.-

Supongase un proceso de nacimiento X_ξ de intensidad λ que partiendo de un numero inicial de individuos $X_0=q$ desconocido,es observado en un tiempo n en el cual $X_n=k$. Ahora se trata de abordar la estimacion del anterior tamaño inicial q de individuos del proceso.

Sea para ello un proceso Y_η de igual intensidad λ que el anterior y tal que :

$$Y_0=1 \quad ,, \quad Y_s=q \quad ,, \quad Y_m=X_n=k \quad ,, \quad m=n+s$$

es decir,un proceso que,partiendo de un solo individuo,alcanza el estado q en un tiempo s desconocido,y en adelante se comporta como el proceso original X pero en la nueva escala de tiempos $\eta = \xi + s$ .Se vio en la seccion 2.3.4. que una estimación de la edad m del proceso Y_η es:

$$\hat{m} = \frac{\ln k}{\lambda}$$

de modo que una estimación de la edad s en la cual pasó por el estado q sera :

$$\hat{s} = \hat{m} - n = \frac{\ln k - \lambda n}{\lambda}$$

Para estimar q parece apropiado utilizar el tamaño medio

$$E(Y_{\hat{s}} / Y_m = k)$$

pero como m es desconocido,ello no es factible.Siguiendo en la misma linea,podria utilizarse la expresion :

$$E(Y_{\hat{s}} / Y_{\hat{s}} \leq k)$$

lo cual da una forma excesivamente complicada para $\hat{q}$;de cualquier modo,cuando k sea suficientemente grande,la expresión puede aproximarse con la que finalmente proponemos:

$$\hat{q} = E(Y_{\hat{s}}) = e^{\lambda \hat{s}} = e^{\ln k - \lambda n}$$

y por tanto :

$$\hat{q} = \frac{X_n}{e^{\lambda n}} \qquad (2.6)$$

es el estimador que se propone.

Observese que:

$$E(\hat{q}) = \frac{E(X_n)}{e^{\lambda n}} = \frac{q.e^{\lambda n}}{e^{\lambda n}} = q$$

de modo que $\hat{q}$ es un estimador insesgado de q.Ademas:

$$V(\hat{q}) = q.\ \frac{(e^{\lambda n} - 1)}{e^{\lambda n}}$$

con lo cual la varianza es tanto mas grande cuanto mas nos alejemos del comienzo del proceso y cuanto mayor sea q,como era logico de esperar.Por esta perdida de memoria- en el sentido anteior -del proceso con el tiempo y con q,y por la escasez de los datos tomados,esta claro que pedir la consistencia de q es excesivo;sin embargo la α -consistencia si debe de ser accesible.Sea :

$$y = \frac{\hat{q}-q}{q^{\alpha}} \qquad \text{con } \alpha > 0,5$$

entonces $E(y) = 0$ y ademas:

$$V(y) = \frac{V(\hat{q})}{q^{2\alpha}} = \frac{q.(e^{\lambda n} - 1)}{e^{\lambda n}.q^{2\alpha}} = \frac{e^{\lambda n} - 1}{e^{\lambda n}.q^{2\alpha - 1}} \longrightarrow 0$$

cuando $q \longrightarrow \infty$ para un tiempo fijo n.Con ello :

$$\frac{\hat{q} - q}{q^{\alpha}} \longrightarrow 0$$ dèbilmente cuando $q \longrightarrow \infty$ y n esta fijo.

de modo que $\hat{q}$ es α-consistente de q , $\forall\alpha > 0,5$.

Ya se indicó que en los procesos de nacimiento puros :

$$x = \frac{X_n}{E(X_n)} \longrightarrow W \text{ c.s. cuando } m \longrightarrow \infty$$

en donde W sigue una gamma(q,q).Con ello :

$$\hat{q} = q.x \longrightarrow W' \text{ c.s. cuando } m \longrightarrow \infty$$

en donde W'sigue una gamma(q,1).

Para dar intervalos de confianza de q,habremos de partir de los intervalos para m.Asi,en la sección 2.3.4., se indicó un intervalo (w_2,w_1) para $\lambda(\hat{m}-m)$,con lo cual:

$$w_1 < \lambda m - \lambda \hat{m} < w_2$$

y por tanto un intervalo para m es:

$$m_1 = \frac{w_1 + \ln k}{\lambda} < m < m_2 = \frac{w_2 + \ln k}{\lambda}$$

de modo que uno para s seria :

$$s_1 = \frac{w_1 + \ln k - \lambda n}{\lambda} < m-n=s < s_2 = \frac{w_2 + \ln k - \lambda n}{\lambda}$$

y finalmente,como :

$$E(Y_{s_1}) = e^{w_1 + \ln k - \lambda n} = \frac{e^{w_1} . k}{e^{\lambda n}}$$

$$E(Y_{s_2}) = e^{w_2 + \ln k - \lambda n} = \frac{e^{w_2} \cdot k}{e^{\lambda n}}$$

y como tambien :

$$w_1 = \left\{ \ln \left\{ -\ln(1-\alpha_1) \right\} \right\}$$
$$w_2 = \ln \left\{ -\ln \alpha_2 \right\}$$

por lo dicho en la sección 2.3.4.,entonces un intervalo de confianza para q es :

$$\left[\frac{-k.\ln(1-\alpha_1)}{e^{\lambda n}} \quad ; \quad \frac{-k_\circ \ln \alpha_2}{e^{\lambda n}} \right]$$

en donde $\alpha_1 + \alpha_2 = \alpha$ es el error del intervalo y α_1 se determina como la raiz positiva mas chica de la igualdad :

$$(\alpha - \alpha_1)\ln(\alpha - \alpha_1) = (1-\alpha_1)\ln(1-\alpha_1)$$

El estimador (2.6) puede obtenerse aún de otro modo.Ya hemos indicado que el proceso X_ξ observado en los tiempos 0,z,2z,.....,rz=n se convierte en uno Z_k que es de Galton-Watson de media de la función generatriz de hijos $e^{\lambda z}$.Con ello tenemos un proceso Z_k de Galton-Watson del tipo fraccional lineal con :

$$Z_o=q \quad ,, \quad Z_r=k \quad ,, \quad \mu = e^{\lambda z} > 1 \quad ,, \quad \pi = 0$$

y por tanto podemos estimar q por la (3.17.bis) del próximo capítulo (ver sección 3.2.5.2.) ; asi :

$$\hat{q} = \frac{k}{\mu^{n}} = \frac{k}{e^{\lambda rz}} = \frac{k}{e^{\lambda n}}$$

que coincide con el anterior.De igual modo con los intervalos de confianza y propiedades de $\hat{q}$.

CAPITULO III

ESTIMACION PARAMETRICA EN LOS PROCESOS DE GALTON-WATSON.-

En los últimos años ha habido un gran crecimiento en el número de trabajos publicados sobre las propiedades de los procesos de ramificacion de Galton-Watson y sus generalizaciones.La gran mayoria de dichos trabajos estudian el comportamiento probabilistico de estos procesos y particularmente su crecimiento asintótico o la extincion de la poblacion cuando la distribucion de hijos esta totalmente especificada.Sin embargo,recientemente,han merecido atencion varios problemas de estimacion sobre los parámetros de los procesos de Galton-Watson,a saber:estimacion de número medio de hijos (Harris),estimacion de la probabilidad de extincion π (Bailey y Stigler),estimacion de la edad del proceso basada en el tamaño actual de la poblacion y en la distribucion de hijos (Stigler).Estas últimas cuestiones de estimacion son las que reproducimos aqui - seccion 3.1. - o en el capitulo IV,segun que se trate de estimadores paramétricos o no paramétricos.

Sin embargo, la estimacion de X_0=q-número inicial de individuos de un proceso de Galton-Watson- o la estimacion de la edad n cuando $X_0=q$, son problemas que no han sido aun abordados. La seccion 3.2. esta dedicada a la obtencion de dichos estimadores, dandose ademas en ella otro camino para la obtencion del estimador de la edad n cuando $X_0=1$ y proponiendose una nueva estimacion de la probabilidad de extincion π que mas adelante-en el capítulo IV-nos sera de utilidad.

3.1. RESULTADOS CONOCIDOS.-

3.1.1. ESTIMACION DE LA EDAD n DE UN PROCESO DE GALTON-WATSON.-

Sean $X_0=1, X_1, X_2, \ldots\ldots$ los tamaños de una poblacion, en generaciones sucesivas, que suponemos sigue un proceso de ramificacion de Galton-Watson con funcion generatriz de hijos :

$$f(s)=\sum_{k=o}^{\infty} p_k \cdot s^k \quad \text{con} \quad p_k= P(\ X_1=k\ /\ X_0=1)$$

Supongamos que para algun $n \geq 0$ observamos el tamaño $X_n > 0$ de la poblacion, pero desconocemos su edad n, en generaciones, en dicho momento. Se tratara pues de estimar dicha edad n. Los procedimientos para ello han sido dados por Stigler y son los que reproducimos a continuacion.

El problema esta pues planteado en el sentido de hacer inferencias sobre el parámetro n (pues como tal sera tratado en adelante) conociendo el tamaño X_n y la funcion generatriz f(s) del número de hijos.

Sea $\mu = f'(1-)$ el numero medio de hijos para un solo individuo;entonces el problema de estimacion anterior sera dividido en tres apartados segun que μ sea mayor,igual o menor que uno,es decir,los llamados procesos supercritico,critico y subcritico.

Hay un tipo de distribucion de hijos que permite simplificaciones matematicas de importancia en el actual problema y que por ello sera la que consideremos en adelante sin perdida,como se vera,de generalidad. Ella es la funcion generatriz que en el capitulo I se denomino por <u>fraccional lineal</u> , es decir:

$$f(s) = 1 - \frac{b}{1-c} + \frac{bs}{1-cs}$$

en donde $b,c > 0$ y $b \leq 1-c$; con ello ahora:

$$p_k = b.c^{k-1} \qquad \text{si } k \geq 1$$

$$p_0 = 1 - \frac{b}{1-c}$$

$$\mu = \frac{b}{(1-c)^2}$$

es evidente que,en adelante, el problema de estimacion sera planteado con la condicion de que $X > 0$,pues en otro caso el problema no surgira.

Notese que hasta ahora hemos venido suponiendo que $X_0=1$;ello es asi por dos posibles situaciones:

a)Porque realmente se tenga constancia de que $X_0=1$.

b)Porque no se sepa nada de cual fue el comienzo del proceso, de modo que no queda otra posibilidad que suponer $X_0=1$.

Una generalizacion de los resultados que siguen se dan en las secciones 3.2.5. y 3.2.6. para el caso en que $X_0=q$ conocido.

3.1.1.1. ESTIMACION EN EL CASO SUPERCRITICO.-

Consideremos el problema de la estimacion de n en el caso de que $\mu > 1$. Para una funcion generatriz f(s) cualquiera, sea $f_n(s)$ su n-esima iterada o funcion generatriz de X_n; entonces:

$$P(X_n=k/\ X_n > 0)= \frac{f_n^{(k)}(0)}{k!\ (1- f_n(0)\ )} \quad ,,\ k=1,2,\dots.$$

en donde $f_n^{(k)}(s)$ es la derivada k-esima de $f_n(s)$. Con ello podriamos estimar n eligiendo $\hat{n}$ que maximiza la anterior probabilidad; el problema es que el calculo explicito de $f_n(s)$ es generalmente dificil de obtener. Sin embargo, cuando f(s) es la fraccional lineal $\varphi(s)$ el problema se simplifica bastante; asi, en el capitulo I ya vimos que:

$$P(X_n=k\ /\ X_n > 0\ ) = \left\{1- c(n)\right\}.c(n)^{k-1} \quad ,,k=1,2,\dots.$$

en donde:

$$c(n) = \frac{\mu^n - 1}{\mu^n - \pi}$$

y:

π = probabilidad de extincion del proceso=

$$= \min. \left\{ 1, \frac{1-b-c}{c\,(1-c)} \right\}$$

Con ello,por derivacion,obtenemos:

$$\hat{n} = \frac{\ln\left\{ (1-\pi)X_n + \pi \right\}}{\ln \mu} \qquad (3.1)$$

que es el estimador buscado.

Notese que estamos considerando a n como un parametro continuo y no como un entero;si se desea,el valor final de la estimacion puede redondearse;entonces, si bien las propiedades asintoticas de $\hat{n}$ permanecen,otras como la maxima verosimilitud valen solo aproximadamente.

El caso fraccional lineal,aparte de la ventaja de tener dos parametros por determinar,nos ha proporcionado el estimador $\hat{n}$ de la edad.Solo falta ya el ver si dicho estimador es apropiado en otras situaciones.Desde luego,pedir la consistencia de $\hat{n}$ en el sentido de que $|\hat{n}-n|$ tienda hacia 0 para grandes n,es demasiado fuerte en vista de la pobreza de los datos observados,sin embargo $\hat{n}$ posee otro tipo de consistencia.Para ver esto,recordemos que cuando $\mu > 1$ y $\sum k(\ln k).p_k < \infty$,entonces,condicionalmente en $X_n > 0$, $X_n.\mu^{-n} \longrightarrow W$ casi seguro cuando n tiene hacia infinito,en donde W es una variable aleatoria de densidad continua en $(0, \infty)$.Con ello:

$$\hat{n}-n = \frac{\ln\left\{(1-\pi)\mu^{-n}.X_n + \pi.\mu^{-n}\right\}}{\ln \mu} \longrightarrow \frac{\ln\left\{(1-\pi)W\right\}}{\ln \mu}$$

cuando $n \longrightarrow \infty$ y en particular $\hat{n}$ es α-consistemte de n $\forall \alpha > 0$ en el sentido de Feldman and Fox de que $\hat{n}$ lo es cuando $(\hat{n}-n)/n^{\alpha}$ converge hacia 0, cuando n tiende hacia infinito, débilmente.

Notese que ello nos ha valido para toda funcion generatriz y no solo para el caso fraccional lineal.

Otra aplicacion de lo anterior es para el calculo de intervalos de confianza para n. En efecto, si $K(u)= P(W \leq u)$ y

$$a_1= (1-\pi)K^{-1}(\alpha_1)$$

$$a_2= (1-\pi)K^{-1}(1-\alpha_2)$$

con $\alpha_1 + \alpha_2 = \alpha$,entonces:

$$\left[\hat{n} - \frac{\ln a_2}{\ln \mu} \quad ; \quad \hat{n} - \frac{\ln a_1}{\ln \mu} \right] \tag{3.2}$$

es un intervalo de confianza, al $(1-\alpha)\%$,aproximado de n. La dificultad esta en calcular $K(u)$. Es conocido que si $h(s)= E(\exp.(-sW))$,ella es la unica solucion de la ecuacion funcional:

$$h(\mu s).(1-\pi) = f\left\{ h(s).(1-\pi) + \pi \right\} \tag{3.3}$$

verificando que:

$$h'(0) = -E(W) = -(1-\pi)^{-1}$$

Pero ello no es facil de resolver siempre. Cuando f(s) es la fraccional lineal $\phi(s)$ se prueba que W tiene la distribucion exponencial de densidad $(1-\pi)\exp.\{-(1-\pi)u\}$ para $u>0$, de modo que en este caso el intervalo (3.2) puede formarse con:

$$a_1 = -\ln(1-\alpha_1)$$

$$a_2 = -\ln \alpha_2$$

La longitud de dicho intervalo es:

$$L = \frac{\ln a_2/a_1}{\ln \mu}$$

y por derivacion puede verse que L es minima si se toma α_1 como la unica riz positiva de la ecuacion:

$$(\alpha-\alpha_1)\ln(\alpha-\alpha_1) = (1-\alpha_1)\ln(1-\alpha_1)$$

Una tabla de valores de α_1, a_1 y a_2 la da Stigler(1970) para algunos valores de α.

Como el intervalo de confianza anterior solo vale asintoticamente en n, es conveniente calcular las probabilidades P(n) reales de cobertura para algunos valores de n. En el mismo articulo, Stigler presenta una tabla de los valores minimos de n para que la probabilidad real P(n) coincida con $(1-\alpha)=0,95$ al menos en dos cifras decimales.

Se prueba que P(n) converge bastante rapido a $(1-\alpha)=0,95$ cuando μ no esta demasiado cercano a uno. Ademas la probabilidad asignada de 0,95 es siempre conservadora pues $\Gamma(n) \geqslant 0{,}95$ para todo n.

Cuando f(s) no es la fraccional lineal, entonces K no suele conocerse explicitamente, de modo que habra que evaluar $K^{-1}(\alpha_1)$ y $K^{-1}(\alpha_2)$ por simulacion de Monte Carlo o bien resolver numericamente la ecuacion (3.3) y usar luego la formula de inversion.

Lo que si puede hacerse es calcular cualquier numero de momentos de W por diferenciacion en la (3.3); asi:

$$E(W) = (1-\pi)^{-1}$$

$$E(W^2) = \frac{f''(1)}{\mu(\mu-1)(1-\pi)}$$

y utilizar asi las desigualdades existentes al efecto como la de Chebychev.

3.1.1.2. ESTIMACION EN EL CASO CRITICO.-

En las mismas condiciones expuestas antes, sea ahora $\mu = 1$ y supongamos de nuevo que f(s) es la fraccional lineal con $b = (1-c)^2$ para que asi la media sea realmente uno. Ahora:

$$P(X_n = k \,/\, X_n > 0) = \{1 - c(n)\} \cdot c(n)^{k-1} \;,,$$

para k=1,2,.... ,en donde:

$$c(n) = \frac{nc}{1-c+nc}$$

de modo que por derivaciom:

$$\hat{n} = \frac{1-c}{c} \cdot (X_n - 1) = \frac{2 (X_n - 1)}{\sigma^2} \qquad (3.4)$$

con $\sigma^2 = \mathrm{Var.}(X_1 / X_0 = 1)$,de modo que $\hat{n}$ es el estimador de maxima verosimilitud de n.Ademas,como:

$$E(X_n / X_n > 0) = \{1 - c(n)\}^{-1} = 1 + \frac{nc}{1-c}$$

entonces $\hat{n}$ es un estimador insesgado de n en el caso fraccional lineal.

Veamos de la utilidad de $\hat{n}$ como estimador de n cuando la funcion generatriz no sea la fraccional lineal sino otra cualquiera.En primer lugar,como

$$E(X_n / X_n > 0) = \{1 - \varphi_n(0)\}^{-1}$$

y es conocido que:

$$n \{1 - \varphi_n(0)\} \longrightarrow 2/\sigma^2 \quad \text{cuando } n \longrightarrow \infty$$

entonces $\hat{n}$ es un estimador asintoticamente insesgado de n en el sentido de que :

$$E\left\{ \frac{\hat{n}}{n} / X_n > 0 \right\} \longrightarrow 1 \quad \text{cuando } n \longrightarrow \infty$$

De otro lado,como:

$$\lim_{n \to \infty} P\left\{ \frac{X_n}{n} \geq z \;\; X_n > 0 \right\} = \exp.\left\{ - \frac{2z}{\sigma^2} \right\} \quad ,, \quad z \geq 0$$

entonces ($\hat{n}/n$ / $X_n > 0$) converge a una exponencial de parametro 1,de modo que el intervalo :

$$S(\hat{n}) = \left[\frac{\hat{n}}{a_2} \, , \, \frac{\hat{n}}{a_1} \right]$$

con:

$$a_1 = -\ln(1-\alpha_1) \quad ,, \quad a_2 = -\ln \alpha_2 \quad ,, \alpha_1 + \alpha_2 = \alpha$$

es un intervalo de confianza al $(1-\alpha)\%$ para n en el sentido de que :

$$P\left\{ n \in S(\hat{n}) \right\} \longrightarrow 1-\alpha \text{ cuando } n \longrightarrow \infty$$

Notese que ahora la longitud del intervalo es proporcional al tamaño de la poblacion.

El mismo articulo de Stigler da algunas indicaciones acerca de la verdadera probabilidad P(n) de que el intervalo $S(\hat{n})$ contenga realmente a n.

3.1.1.3. ESTIMACION EN EL CASO SUBCRITICO.-

Ahora $\mu < 1$ y se prueba que :

$$\lim_{n \to \infty} P\left\{ X_n = k \;/\; X_n > 0 \right\} = b_k \quad ,, \quad k=1,2,.....$$

existen,y ademas $\sum_{k=1}^{\infty} b_k = 1$ de modo que no tiene sentido el calculo de un intervalo de confianza para n.

3.1.2. ESTIMACION DE LA PROBABILIDAD DE EXTINCION π.-

Sea un proceso de Galton-Watson con funcion generatriz de probabilidad $g(s, \theta)$ conocida en su forma, pero no en el valor de sus parametros (θ es un vector parametrico q-dimensional).Se desea estimar la probabilidad de extincion π del proceso a partir de una muestra aleatoria $X_1, X_2, \ldots, X_n$ de variables aleatorias distribuidas segun la funcion generatriz $g(s, \theta)$.Es conocido que π es la raiz no negativa mas chica de la ecuacion $g(s, \theta) = s$,pero el problema es que θ es desconocido.

Supongamos que $\hat{\theta}$ es una estimada de θ ;entonces definimos:

$$\tilde{\pi} = \inf.\left\{ s \,/\, g(s, \hat{\theta}) = s \right\},\, , \quad s \geq 0$$

de modo que si $\hat{\theta}$ es de maxima verosimilitud para θ ,entonces $\tilde{\pi}$ lo es para π .

En estas condiciones Stigler (1971) da el siguiente teorema en el cual θ_0 denota al verdadero valor de $\theta = \left\{ \theta^{(1)}, \ldots, \theta^{(q)} \right\}$ y π es la probabilidad de extincion para cuando la distribucion del numero de hijos es $g(s, \theta_0)$:

<u>Teorema 3.1</u>.- Sea $\left\{ g(s, \theta) ; \theta \in \Omega \right\}$ una familia de funciones generatrices de probabilidad de espacio parametrico $\Omega \subset R^q$.Supongamos que:

1)
$$\left[\frac{\partial}{\partial s} g(s, \theta_0) \right]_{s=1} > 1$$

$g(0,\vartheta_0) > 0$

y que $(\partial / \partial s) g(s,\vartheta)$ es continua en s y ϑ en (π,ϑ_0)

2)

$$\exists \quad a_i = \left[\frac{\partial}{\partial \vartheta^{(i)}} g(\pi,\vartheta) \right]_{\vartheta = \vartheta_0} \qquad i=1,2,\dots,q$$

3)Existe una estimada $\hat{\vartheta}$ de ϑ tal que cuando el verdadero valor de ϑ es ϑ_0 entonces:

$\sqrt{n}\,(\hat{\vartheta}-\vartheta_0)$ es asintoticamente $N(0,\Sigma)$

con $a'\Sigma a > 0$,, $a'=(a_1,\dots,a_q)$

Entonces,bajo todas estas condiciones:

$\sqrt{n}\,(\tilde{\pi}-\pi)$ es asintóticamente normal de media 0 y varianza $h^{-2}.a'\Sigma a$ en donde

$$h=1-\left[\frac{\partial}{\partial s} g(s,\vartheta_0)\right]_{s=\pi}$$

En particular,en el caso de que $g(s,\vartheta)=g(s,\lambda)= =\exp.\{\lambda(s-1)\}$,entonces se tendria la familia de Poisson y:

$$\hat{\lambda} = \bar{X} \quad ,, \quad V(\tilde{\pi}/\lambda) = \lambda \left\{ \frac{\pi(1-\pi)}{(1-\lambda\pi)} \right\}^2$$

En el caso de que $g(s,\vartheta)$ sea la fraccional lineal,entonces llamando:

M_n= nº de X_i que son cero.

$$Y_n = \begin{cases} \dfrac{\sum_{i=1}^{n} X_i}{(n - M_n)} & \text{si } M_n < n \\ 1 & \text{si } M_n = n \end{cases}$$

los estimadores de maxima verosimilitud de b,c y π son:

$$\hat{b} = \frac{(n - M_n)}{nY_n}$$

$$\hat{c} = 1 - Y_n^{-1}$$

$$\tilde{\pi} = \min. \left[\frac{Y_n . M_n}{n(Y_n - 1)} , 1 \right]$$

en donde:

$\sqrt{n} \left\{ (\hat{b} , \hat{c}) - (b,c) \right\}$ esta asintoticamente distribuido segun una $N(0, \Sigma)$ con $\Sigma = (\sigma_{ij})$ y:

$$\sigma_{11} = b(1-b-c) + b^2 . c(1-c)(1-b-c)^{-1}$$

$$\sigma_{22} = c(1-c)^3 . (1-b-c)^{-1}$$

$$\sigma_{12} = \sigma_{21} = -bc(1-c)^2 . (1-b-c)^{-1}$$

de modo que cuando $b > (1-c)^2$:

$$\mathrm{Var.}(\tilde{\pi} / b,c) = \frac{(1-b-c)\left\{(1-b-c)(1-c)^3 + b^2 c\right\}}{b . c^3 . (1-c)^2}$$

por el teorema 3.1.

En el proximo capitulo ya veremos de la conveniencia de no usar estos estimadores, sino unos no paramétricos.

3.2. APORTACIONES.-

3.2.1. EL PROCESO DE GALTON-WATSON FRACCIONAL LINEAL Y SU PROCESO DE NACIMIENTO PURO ASOCIADO.-

Sea X_{ξ} un proceso de Gaton-Watson del tipo fraccional lineal. Cuando $X_0=1$ y $\mu > 1$ ya sabemos que:

$$P\left(X_n=K \;/\; X_0=1 \;,,\; X_n > 0 \right) = \left\{ 1-c(n) \right\} . \left\{ c(n) \right\}^{K-1}$$

en donde

$$c(n) = \frac{\mu^n - 1}{\mu^n - \pi} \qquad ,, \qquad \mu = E(X_1 \;/\; X_0=1)$$

y π es como siempre la probabilidad de extincion. Sustituyendo:

$$P(X_n=K \;/\; X_0=1, X_n > 0) = \frac{1-\pi}{\mu^n - \pi} . \left\{ 1- \frac{1-\pi}{\mu^n - \pi} \right\}^{K-1} \tag{3.5}$$

Como $(1-\pi)/(\mu^n - \pi) < 1$ por ser $\mu > 1$, entonces podemos convenir en llamar por:

$$e^{-\lambda t} = \frac{1-\pi}{\mu^n - \pi}$$

determinandose el producto $\lambda.t$ de modo que tal igualdad se verifique;en particular,si convenimos en que $\lambda=1$, entonces t esta determinado univocamente por

$$t = \ln \frac{\mu^n - \pi}{1 - \pi} \qquad (3.6)$$

Con esta nueva notacion,la (3.5) puede ponerse de un nuevo modo:

$$P(\ X_n=K\ /\ X_o=1, X_n > 0\ ) = e^{-t}.(1 - e^{-t})^{K-1}$$

Sea ahora un proceso Y_η de nacimiento puro y de intensidad $\lambda=1$ con $Y_o=1$; entonces:

$$P(\ Y_t=K\ /\ Y_o=1\ ) = e^{-t}.(1- e^{-t})^{K-1}$$

con lo cual:

$$P(\ X_n=K\ /\ X_o=1, X_n > 0\ ) = P(\ Y_t=K\ /\ Y_o=1\ ) \qquad (3.7)$$

Supongamos ahora el mismo proceso X_ξ del inicio de la seccion con $X_o=1$ tambien,pero ahora $\mu = 1$.Ahora:

$$P(\ X_n=K\ /\ X_o=1, X_n > 0\ ) = \{1- c(n)\}\ c(n)^{K-1}$$

en donde:

$$c(n) = \frac{nc}{1-c + nc} \quad ,, \sigma^2 = V(X_1 \,/\, X_o=1) = \frac{2c}{1-c}$$

con ello:

$$P(\, X_n=K \,/\, X_o=1,\; X_n > 0)= \frac{2}{2 + n\sigma^2} \cdot \left\{ 1- \frac{2}{2 + n\sigma^2} \right\}^{K-1}$$

de nuevo, como $2/\,(2+n\sigma^2) < 1$, podemos igualar dicha cantidad a $e^{-\lambda t}$ y si hacemos $\lambda=1$, entonces t viene dado univocamente por:

$$t = \ln \frac{2 + n\sigma^2}{2} \tag{3.8}$$

y sustituyendo arriba:

$$P(\, X_n=K \,/\, X_o=1,\; X_n > 0 \,) = e^{-t}.(\, 1- e^{-t})^{K-1}$$

Sea ahora un proceso de nacimiento Y_η de intensidad $\lambda=1$ con $Y_o=1$; entonces:

$$P(\, Y_t=K \,/\, Y_o=1)= e^{-t}.(\, 1- e^{-t})^{K-1}$$

de modo que de nuevo:

$$P(\, X_n=K \,/\, X_o=1,\; X_n > 0 \,)= P(\, Y_t=K, / \, Y_o=1) \tag{3.9}$$

Con ello,observando las igualdades (3.7) y (3.9) se deduce que probabilisticamente es lo mismo hablar de un proceso de Galton-Watson del tipo fraccional lineal que partiendo de $X_o=1$ se encuentra en el estado K en la generacion n,que hablar de un proceso de nacimiento de intensidad $\lambda=1$,que partindo de $Y_o=1$ se encuentra en el estado K en un tiempo t relacionado con n por las expresiones (3.6) ó (3.8) segun que se trate del caso supercritico ó critico.A dicho proceso Y_η sera al que llamemos <u>proceso de nacimiento asociado al de Galton-Watson</u>. Se trata pues de un nuevo ejemplo de dos procesos que si bien microscopicamente responden a leyes distintas,macroscopicamente funcionan igual.Dicho proceso asociado nos sera de gran utilidad en los problemas de estimacion que siguen.

Notese que para el caso supercritico:

$$E(Y_t) = e^{\lambda t} = \frac{\mu^n - \pi}{1 - \pi} = E(X_n \;/\; X_o=1, X_n > 0)$$

$$V(Y_t) = e^{\lambda t}(e^{\lambda t} - 1) = \frac{(\mu^n - \pi)(\mu^n - 1)}{(1 - \pi)^2} =$$

$$= V(X_n \;/\; X_o=1 \;,\; X_n > 0)$$

de modo que en ambos procesos las medias y varianzas son coincidentes como era de esperar.De igual modo,en el caso critico:

$$E(Y_t) = \frac{2 + n\sigma^2}{2} = E(X_n \ / \ X_o=1,X_n > 0)$$

$$V(Y_t)= \frac{2 + n\sigma^2}{2} \cdot \frac{n\sigma^2}{2} = V(X_n \ / \ X_o=1,X_n > 0)$$

3.2.2.EL PROCESO PARALELO A UN PROCESO DE GALTON-WATSON.-

Sea X_ξ un proceso de Galton-Watson de funcion generatriz de hijos f(s) y tal que comenzando con $X_o=q$ individuos,es observado en la generacion n en la cual $X_n=K$.Sea Y_η un nuevo proceso de Galton-Watson con igual funcion generatriz f(s) que el anterior y tal que comenzando con $Y_o=1$ individuos,pasa por el estado q en un tiempo t desconocido y observado el el tiempo m = n + t es tal que $Y_m=X_n=K$.Con ello el proceso Y_η se comporta como el proceso X_ξ desde que alcanza el estado q con la previa transformacion de la escala de tiempos.Graficamente:

0	t	m	tiempos η
I	I	I	
1	q	K	tamaños Y_η

0	n	tiempos ξ
I	I	
q	K	tamaños X_ξ

Dicho proceso Y_η sera llamado en adelante <u>proceso paralelo</u> del X_ξ ,y nos sera asimismo de utilidad en las proximas secciones.

3.2.3. ESTIMACION DE LA EDAD n Y LA PROBABILIDAD DE EXTINCION π CUANDO $X_o=1$ EN EL CASO SUPERCRITICO.-

Sea X_ξ un proceso de Galton-Watson del tipo fraccional lineal con $X_o=1$ y $\mu>1$,que observado en la generacion n da $X_n=K$.Sea Y_η su proceso de nacimiento asociado de modo que $Y_o=1$,$Y_t=K$ de intensidad $\lambda=1$,en donde t viene dado por la (3.6).Supongamos que en el proceso original es desconocida la edad n o bien la probabilidad de extincion π ,de modo que t sera desdonocida en el proceso Y_η .Una estimacion de t bajo este tipo de muestreo -muestreo de tipo 4 - la daba el estimador (2.2),es decir:

$$\hat{t} = \frac{\ln Y_t}{\lambda} = \ln K \quad \text{estima a } t = \ln \frac{\mu^n - \pi}{1 - \pi}$$

estimador que ademas era de maxima verosimilitud.Con ello K estimara a $(\mu^n - \pi) / (1-\pi)$ y despejando cada uno de los parámetros:

$$\hat{n} = \frac{\ln\{K(1-\pi) + \pi\}}{\ln \mu} \qquad (3.10)$$

$$\hat{\pi} = \frac{K - \mu^{n}}{K - 1} \qquad (3.11)$$

$$\hat{\mu} = (K(1-\pi) + \pi)^{1/n} \qquad (3.12)$$

Los intervalos de confianza para n, π ó μ pueden darse a partir de los intervalos para el parametro **t** ya estudiados en el capitulo 2.Para que ello sea valido,habremos de comprobar que las leyes limites son la misma en el proceso X_ξ que en el proceso Y_η .En efecto,en el proceso de nacimiento Y_η :

$$\frac{Y_t}{E(Y_t)} = \frac{K(1-\pi)}{\mu^{n}-\pi} \longrightarrow W \text{ c.s.}$$

cuando n tiende hacia infinito,en dohde W sigue una gamma(1,1),o sea, W tiene de densidad e^{-w} .De igual modo, para el proceso de Galton-Watson X_ξ :

$$\left(\frac{X_n}{\mu^{n}} / X_n > 0 \right) \longrightarrow W' \text{ c.s.}$$

cuando n tiende hacia infinito,en donde W'tiene por densidad $(1-\pi) \cdot \exp\{-(1-\pi)w\}$como se dijo en el capitulo 1.

Con ello:

$$\left(\frac{X_n(1-\pi)}{\mu^n} \,/\, X_n > 0 \right) \dashrightarrow W \quad c.s.$$

cuando n tiende hacia infinito,en donde W tiene ya por densidad e^{-w}.Naturalmente,la convergencia no se altera por restar π en el denominador,de modo que las leyes limites para X_ξ e Y_η son coincidentes y los intervalos propuestos arriba son validos.Notese que la validez de los intervalos ha venido condicionada al caso fraccional lineal que es de la hipotesis de que se partio.

Observese que el estimador (3.10) de la edad n del proceso de Galton-Watson es el mismo de Stigler de la seccion 3.1.1.1.,por lo cual le es aplicable el estudio hecho entonces sobre su validez en cualquier proceso-no necesariamente el caso fraccional lineal- y sobre los intervalos de confianza.

El estimador $\hat{\mu}$ requiere un conocimiento previo de n,K y π .El conocer n y X_n=K es factible,pero el conocer π presupone el conocer la funcion generatriz f(s) conociendose por tanto tambien la media μ ,en cuyo caso pierde sentido el problema de la estimacion de dicha media.Podria forzarse la situacion e intentar que $\hat{\mu}$ nos sirva como estimador no parametrico en el sentido de que previamente habriamos de estimar π mediante el conocimiento de varias generaciones de X;el problema es que todos los estimadores de π contienen a μ como incognita y μ

es precisamente el parametro a estimar. En definitiva, el estimador $\hat{\mu}$ no tiene validez salvo en la situacion que se describe en la seccion 4.3.2.

El estimador $\hat{\pi}$ de π presenta el mismo problema de depender de μ, pero ahora puede soslayarse por la existencia de estimadores de μ que no dependen de π. En el proximo capitulo se hara un estudio de tal estimador.

3.2.4. ESTIMACION DE LA EDAD n CUANDO $X_0=1$ EN EL CASO CRITICO.-

Sea de nuevo X_ξ un proceso de Galton-Watson del tipo fraccional lineal con $X_0=1$ pero ahora de media la unidad; supongamos es observado en la generacion n de modo que $X_n=K$. Sea Y_η el proceso de nacimiento asociado al X_ξ que con intensidad $\lambda=1$ e $Y_0=1$ es observado en un tiempo t - dado por la (3.8) - en el cual $Y_t=K$. Supongamos que n o σ^2 son desconocidos; entonces t es tambien desconocido y habra de estimarsele por la expresion (2.2); con ello :

$$\hat{t} = \ln Y_t = \ln K \quad \text{estima a } t = \ln\left\{ \frac{2+n.\sigma^2}{2} \right\}$$

de modo que K estima a $(2+n.\sigma^2)/2$ y despejando cada uno de los parametros, parecen aconsejables cada uno de los estimadores siguientes :

$$\hat{n} = 2 \frac{K-1}{\sigma^2} \qquad (3.13)$$

$$\widehat{\sigma^2} = 2 \frac{K-1}{n} \qquad (3.14)$$

em donde $\hat{t}$ era estimador de maxima verosimilitud de t.

De nuevo $\hat{n}$ es el estimador de Stigler - seccion 3.1.1.2. - y su validez e intervalos fueron dados entonces para cualquier X_ξ ,no necesariamente el caso fraccional lineal.

Aqui tambien puede verse que las leyes limites para los procesos X_ξ e Y_η son coincidentes,de modo que pueden darse intervalos para n y σ^2 a partir de los intervalos para t estudiados en el capitulo anterior(válido solo para el caso fracional lineal que es la hipotesis mantenida hasta ahora).Para n,estos intervalos coinciden con los dados en 3.1.1.2.

De nuevo,el estimador $\widehat{\sigma^2}$ pierde sentido pues esta basado en el comocimiento de que la media μ es la unidad y por las razones expuestas en la seccion anterior.

3.2.5. ESTIMACION DE n , π y q CUANDO $X_0=q$ EN EL CASO SUPERCRITICO.-

El procedimiento usado en las secciones precedentes 3.2.3. y 3.2.4. no puede aplicarse ahora entre otras razones porque en el proceso de Galton-Watson puede suceder que $X_n < q$ mientras que en el de nacimiento $Y_t \geq q$ siempre,si $Y_0=q$.Veamos de adaptar la situacion para seguir aprovechandonos de lo anterior.

Sea X_ξ el proceso de Galton-Watson fraccional lineal original de media μ y probabilidad de extincion π con $X_0=q$ y $X_n=K$,y sea Y_η el proceso paralelo al anterior con $Y_0=1,Y_s=q,Y_m=K,m=n+s$,siendo s desconocido.Ahora,para el proceso Y_η ,ya si es aplicable lo dicho en la seccion 3.2.3. de modo que:

$$\hat{t} = \ln K \quad \text{estima a} \quad t = \ln \frac{\mu^m - \pi}{1 - \pi} = \ln \frac{\mu^n \cdot \mu^s - \pi}{1 - \pi}$$

El problema es que el tiempo s,en que el proceso Y_η esta ocupando el estado q,es desconocido.Ello puede solventarse haciendo notar que el proceso realmente comenzo con q individuos y no con 1,de modo que s parece logico calcularlo como el tiempo medio que tardaria el proceso en recorrer esos estados.Como:

$$E(Y_s / Y_s > 0) = \frac{\mu^s - \pi}{1 - \pi}$$

entonces el tiempo $\bar{s}$ tal que $E(Y_{\bar{s}} / Y_{\bar{s}} > 0) = q$ es:

$$\bar{s} = \frac{\ln \left\{ q.(1-\pi) + \pi \right\}}{\ln \mu}$$

y el sera el que adoptemos como valor de s. Con ello :

$$K \text{ estima } \frac{\mu^{n} . \left\{ q(1-\pi) + \pi \right\} - \pi}{1 - \pi}$$

de modo que despejando cada uno de los parametros obtenemos los estimadores que se proponen :

$$\hat{n} = \frac{\ln \left\{ \frac{K(1-\pi) + \pi}{q(1-\pi) + \pi} \right\}}{\ln \mu} \tag{3.15}$$

$$\hat{\pi} = \frac{1}{1 + \frac{\mu^{n} - 1}{K - q.\mu^{n}}} \tag{3.16}$$

$$\hat{q} = \frac{K}{\mu^{n}} + \frac{\pi}{1 - \pi} \circ \frac{1 - \mu^{n}}{\mu^{n}} \tag{3.17}$$

$$\hat{\mu} = \left(\frac{K.(1-\pi) + \pi}{q.(1-\pi) + \pi} \right)^{1/n} \tag{3.12 bis}$$

en donde $\hat{\mu}$ recibe los mismos comentarios hechos en 3.2.3. con respecto al estimador (3.12).

El estimador $\hat{\pi}$ sera estudiado en el capitulo próximo,mientras que los estimadores $\hat{n}$ y $\hat{q}$ se veran en las dos subsecciones siguientes.

3.2.5.1.ESTUDIO DEL ESTIMADOR $\hat{n}$.-

Veamos antes otra justificacion de la forma del estimador $\hat{n}$ propuesta mas arriba.Sean para ello el proceso original X_ξ con $X_0=q$ y $X_n=K$ y sea Y_η el proceso paralelo tal que:

$Y_0=1$,, $Y_s=q$,, $Y_m=K$,, $m = n + s$,, s desconocida

Entonces,las edades m y s pueden estimarse mediante la (3.10) de Stigler,con lo cual:

$$\hat{m} = \frac{\ln\left\{ K(1-\pi) + \pi \right\}}{\ln \mu}$$

$$\hat{s} = \frac{\ln\left\{ q(1-\pi) + \pi \right\}}{\ln \mu}$$

de modo que un estimador apropiado de n seria:

$$\hat{n} = \hat{m} - \hat{s} = \frac{\ln\left\{ \frac{K(1-\pi) + \pi}{q(1-\pi) + \pi} \right\}}{\ln \mu}$$

en donde μ es la media de la distribucion de hijos y π la probabilidad de extincion.Notese que $\hat{n}$ es el mismo (3.15) anterior;ademas,cuando q=1 obtenemos el mismo estimador (3.10) de Stigler para la edad de un proceso de Galton-Watson que comenzo con un solo individuo.

Para estudiar las propiedades del estimador $\hat{n}$ actual,notemos que:

$$\hat{n}-n = \frac{\ln\left\{\frac{K(1-\pi)+\pi}{q(1-\pi)+\pi}\right\}}{\ln \mu} - \frac{\ln \mu^{m-s}}{\ln \mu} =$$

$$= \frac{\ln\left\{\frac{K(1-\pi)\mu^{-m}+\pi\mu^{-m}}{q(1-\pi)\mu^{-s}+\pi\mu^{-s}}\right\}}{\ln \mu}$$

Si ahora hacemos que n tienda hacia infinito,para un q dado,entonces m tiende hacia infinito quedando s fijo y por tanto como,condicionado en que $K>0$,se sabe que:

$$\frac{Y_m}{E(Y_m)} = \frac{K}{\mu^m} \longrightarrow W \text{ c.s. cuando } m \longrightarrow \infty$$

en donde W es la misma variable aleatoria de la seccion 3.1.1.1. de densidad continua en $(0, \infty)$, se tiene que :

$$\hat{n} - n \longrightarrow \frac{\ln \left\{ \dfrac{W(1-\pi)}{q.\mu^{-s}(1-\pi) + \pi.\mu^{-s}} \right\}}{\ln \mu}$$

c.s. cuando $n \longrightarrow \infty$,de modo que finalmente:

$$\frac{\hat{n}-n}{n^{\alpha}} \longrightarrow 0 \qquad \text{cuando } n \longrightarrow \infty$$

con lo cual $\hat{n}$ es α-consistente de n para todo $\alpha > 0$.

El calculo de intervalos aproximados para n no es ahora directamente accesible pues :

$$\left(\frac{X_n}{\mu^n} \,/\, X_n > 0, , X_0=q \right) \longrightarrow W'' \qquad \text{c.s.}$$

cuando $n \longrightarrow \infty$,en donde W'' es continua en $(0, \infty)$, pero su densidad es suma de q densidades relacionadas con distribuciones gamma que tienen uno de sus parametros distintos,lo cual haria sumamente engorroso el calculo de tales intervalos especialmente para valores grandes de q.

Un intervalo muy conservador,para cuando m y s sean grandes,consistiria en dar intervalos para m y s como en la seccion 3.1.1.1. y combinarlos luego para obtener un intervalo para n.

La misma expresion anterior de $\hat{n}$ -n puede valernos para dar un intervalo de confianza para n doblemente aproximado.Como en dicha expresion aparece el tiempo desconocido s,si convenimos,como antes,en que el valor mas apropiado para s es :

$$\frac{\ln\left\{q.(1-\pi)+\pi\right\}}{\ln \mu}$$

entonces :

$$\hat{n}-n \longrightarrow \frac{\ln\left\{w(1-\pi)\right\}}{\ln \mu}$$

de modo que los mismos intervalos para W y n vistos en la seccion 3.1.1.1. - en la que $X_0=1$ - nos valen ahora salvo que $\hat{n}$ esta dado por la (3.15) y no por la (3.10). Asi pues, la anchura de ambos intervalos es la misma, pero no el punto base $\hat{n}$ de cada intervalo. Como :

$$\hat{n}(q)=\frac{\ln\left\{\frac{K(1-\pi)+\pi}{q(1-\pi)+\pi}\right\}}{\ln \mu} \leq \hat{n}(1)=\frac{\ln\left\{K.(1-\pi)+\pi\right\}}{\ln \mu}$$

entonces el punto base $\hat{n}(q)$ del intervalo actual es inferior al del intervalo de Stigler como era logico que fuera pues es de esperar que el tamaño K se alcance antes partiendo de q individuos que partiendo de uno solo.

La doble aproximacion antes aludida se debe a que por un lado n no es infinito, y por otro se ha sustituido s por su estimacion, al no ser accesible un intervalo para el tiempo s.

Otro camino para lograr intervalos para n,consistiria en dar un intervalo para t -segun la expresion (2.4) de 2.3.4. - y luego trasladarlos a n por los mismos procedimientos descritos en 3.2.5.,intervalo que dara sensiblemente igual al anterior.

3.2.5.2.ESTUDIO DEL ESTIMADOR $\hat{q}$.-

Antes de ver las propiedades de $\hat{q}$,obtengamoslo de otro modo.Sea para ello,en las mismas condiciones del apartado anterior,X_{ξ} e Y_{η} los procesos original y paralelo.Para el proceso Y_{η} una estimacion de la edad m era :

$$\hat{m} = \frac{\ln \left\{ K.(1-\pi) + \pi \right\}}{\ln \mu}$$

de modo que una estimacion del tiempo s es $\hat{s}=\hat{m} - n$.Con ello,parece conveniente estimar q por el tamaño medio que alcanzara el proceso en dicho tiempo $\hat{s}$,condicionado a que comenzo con un solo individuo y a que en el tiempo m este ocupando el estado K.Como m es desconocido,habremos de conformarnos con la siguiente estimacion :

$$\hat{q} = E(Y_{\hat{s}} \ / \ 0 < Y_{\hat{s}}) = \frac{E(Y_{\hat{s}})}{P(Y_{\hat{s}} > 0)}$$

En el caso fraccional lineal,tendremos:

$$\hat{q} = \mu^{\hat{s}} \frac{1 - (\mu^{\hat{s}} - 1)/(\mu^{\hat{s}} - \pi)}{\mu^{\hat{s}} \cdot \frac{(1 - \pi)^2}{(\mu^{\hat{s}} - \pi)^2}} = \frac{\mu^{\hat{s}} - \pi}{1 - \pi}$$

y como:

$$\mu^{\hat{s}} = \frac{\mu^{\hat{m}}}{\mu^{n}} = \frac{K(1 - \pi) + \pi}{\mu^{n}}$$

entonces,finalmente:

$$\hat{q} = \frac{K}{\mu^{n}} + \frac{\pi}{1 - \pi} \cdot \frac{1 - \mu^{n}}{\mu^{n}}$$

que es el mismo estimador (3.17) visto anteriormente.

Naturalmente, cuando $\hat{q} < 1$ entonces haremos $\hat{q}=1$. Notese ademas que para el caso fraccional lineal:

$$E(\hat{q}) = E(Y_s / Y_s > 0)$$

Para estudiar las propiedades de $\hat{q}$, sea $f(s)$ la funcion generatriz del numero de hijos, y sea:

$$\alpha = P(X_n > 0 / X_o=q) = 1 - f_n^q (0)$$

Como estamos en el caso supercritico, entonces $\pi < 1$, de modo que $P(X_n=0) = f_n(0) < 1$; con ello:

$$\alpha \longrightarrow 1 \quad ,, \quad \alpha^2 \longrightarrow 1 \quad ,, \quad \alpha -1 \longrightarrow 0$$

cuando $q \longrightarrow \infty$, y por tanto:

$$E(X_n / X_n > 0) = \frac{E(X_n)}{\alpha} \longrightarrow E(X_n)$$

$$V(X_n / X_n > 0) = \frac{\alpha V(X_n) + (\alpha -1).E^2(X_n)}{\alpha^2} \longrightarrow$$

$$\longrightarrow V(X_n)$$

cuando q $\longrightarrow \infty$ .Finalmente:

$$E(\hat{q}) \longrightarrow q + \frac{\pi}{1-\pi} \cdot \frac{1-\mu^{n}}{\mu^{n}}$$

$$E\left\{\frac{\hat{q}-q}{q^{\alpha}}\right\} \longrightarrow \frac{\frac{\pi}{1-\pi} \cdot \frac{1-\mu^{n}}{\mu^{n}}}{q^{\alpha}}$$

$$\longrightarrow 0 \quad \text{si } \alpha > 0$$

cuando q $\longrightarrow \infty$,y como tambien:

$$V\left\{\frac{\hat{q}-q}{q^{\alpha}}\right\} \longrightarrow \frac{V(X_n)}{\mu^{2n} \cdot q^{2\alpha}} = \frac{\mu^{-n-1} \cdot \sigma^{2} \cdot \frac{\mu^{n}-1}{\mu-1} \cdot q}{q^{2\alpha}}$$

$$\longrightarrow 0 \quad \text{si } \alpha > 0{,}5$$

al tender q hacia infinito,entonces $\hat{q}$ es α -consistente de q para todo $\alpha > 0{,}5$.

Tambien como:

$$E(X_n / X_o = q, X_n > 0) = q \cdot \frac{\mu^{n}}{1 - f_n^{q}(0)}$$

entonces:

$$E\left\{\frac{\hat{q}}{q} \,/\, X_0=q, X_n > 0\right\} = \frac{1}{1 - f_n^q(0)} +$$

$$+ \frac{1}{q} \cdot \frac{\pi}{1-\pi} \cdot \frac{1-\mu^n}{\mu^n} \longrightarrow 1$$

si $q \longrightarrow \infty$,con lo cual $\hat{q}$ es asintoticamente insesgado en ese sentido.

En particular,cuando $f(0)=P(X_1=0 \,/\, X_0=1)=0$, entonces $\pi=0$ y por tanto el estimador $\hat{q}$ se convierte en el siguiente:

$$\hat{q} = \frac{X_n}{\mu^n} \qquad (3.17 \text{ bis})$$

con lo cual:

$$E(\hat{q}) = \frac{q \cdot \mu^n}{\mu^n} = q$$

de modo que en este caso,para toda $f(s)$, $\hat{q}$ es centrado -ya no asintoticamente- de q.

Cuando m sea grande,entonces es posible dar intervalos de confianza para q a partir de los intervalos de Stigler para m. Notese que como $m=n+s$, entonces m es grande cuando n ó s loęs. Supongamos calculado, por los metodos de la seccion 3.1.1.1., un intervalo para m y sea este:

$$\left[\hat{m} - a_1 \; ; \; \hat{m} + a_2 \right]$$

con ello, un intervalo para s seria:

$$\left[s_1=\hat{m} - a_1 - n \quad ; \; s_2= \hat{m} + a_2 - n \right]$$

de modo que si denominamos por $\hat{q}(n)$ al estimador de q cuando la observacion se hace en el tiempo n, entonces un intervalo para q es:

$$\left[q_1= \hat{q}(n+a_1) \quad ; \; \hat{q}(n- a_2)= q_2 \right]$$

es decir:

$$q_1= \hat{q}(n+a_1)= \frac{K}{\mu^{n+a_1}} - \frac{\pi}{1 - \pi} \cdot \frac{1 - \mu^{n+a_1}}{\mu^{n+a_1}}$$

$$q_2= \hat{q}(n-a_2)= \frac{K}{\mu^{n-a_2}} - \frac{\pi}{1 - \pi} \cdot \frac{1 - \mu^{n-a_2}}{\mu^{n-a_2}}$$

3.2.6. ESTIMACION DE n Y q CUANDO $X_0=q$ EN EL CASO CRITICO.-

Sea X_ξ un proceso de Galton-Watson del tipo fraccional lineal con $\mu=1$ y $X_0=q$ y tal que observado en el tiempo n da un tamaño $X_n=K$. Sea Y_η el proceso paralelo con $Y_0=1$, $Y_s=q$, $Y_m=K$, $m=n+s$. Para este proceso Y_η ya si es posible hablar del proceso de nacimiento asociado que con intensidad $\lambda=1$ parte de un solo individuo y en el tiempo t - cuyo valor es el indicado por la expresion (3.8)-se encuentra ocupando el estado K. Como t es desconocido-por desconocerse n ó q -podemos estimarlo por lo indicado en la seccion 2.3.4. de modo que:

$$\hat{t} = \ln K \quad \text{estima a } t = \ln \frac{2 + m\sigma^2}{2} = \ln \frac{2 + n\sigma^2 + s\sigma^2}{2}$$

y repitiendo el argumento de la seccion 3.2.5. anterior, s puede ser asimilado a la cantidad $2(q-1)/\sigma^2$, con lo cual:

$$\hat{t} = \ln K \quad \text{estima a } t = \ln \frac{n\sigma^2 + 2q}{2}$$

de modo que si

$$K \quad \text{estima a} \quad \frac{n\sigma^2 + 2q}{2}$$

entonces parecen aconsejables -despejando-los estimadores siguientes:

$$\hat{n} = 2\,\frac{K - q}{\sigma^2} \tag{3.18}$$

$$\hat{\sigma^2} = 2\,\frac{K - q}{n} \tag{3.19}$$

$$\hat{q} = \frac{2K - n\,\sigma^2}{2} \tag{3.20}$$

en donde, de nuevo, el estimador $\hat{\sigma^2}$ no tiene sentido puesto que al conocerse que $\mu = 1$ entonces se conoce la distribucion de hijos y por tanto debe de conocerse asimismo la varianza σ^2.

Los estimadores $\hat{n}$ y $\hat{q}$ son estudiados a continuacion.

3.2.6.1. ESTUDIO DEL ESTIMADOR $\hat{n}$.-

Veamos otro camino de obtener el mismo estimador $\hat{n}$ anterior. De modo paralelo a lo hecho en la seccion 3.2.5.1., las edades m y s pueden ser estimadas mediante la expresion (3.4) de Stigler :

$$\hat{m} = 2.\frac{K-1}{\sigma^2} \quad , , \quad \hat{s} = 2.\frac{q-1}{\sigma^2}$$

de modo que un estimador de n sera :

$$\hat{n} = \hat{m} - \hat{s} = 2.\frac{K-1}{\sigma^2} - 2.\frac{q-1}{\sigma^2} = 2.\frac{K-q}{\sigma^2}$$

que es el mismo (3.18) propuesto antes. Notese que cuando q=1, entonces $\hat{n}$ es el mismo estimador (3.4) dado entonces.

Al igual que cuando q=1, puede verse que $\hat{n}$ es asintoticamente insesgado. Para ello recuerdese que si f(s) es la funcion generatriz del numero de hijos de un proceso de Galton-watson critico que comienza con un solo individuo, entonces :

$$n\left\{1 - f_n(0)\right\} \longrightarrow \frac{2}{\sigma^2} \quad \text{cuando } n \longrightarrow \infty \tag{3.21}$$

entonces, como :

$$E(X_n / X_n > 0) = \frac{2}{n.\sigma^2} \circ \frac{q}{1 - f_n^q(0)} - \frac{2.q}{n.\sigma^2}$$

y de la (3.21) puede deducirse, por desarrollo del binomio de Newton que se obtiene, que :

$$E\left(\frac{\hat{n}}{n} \,/\, X_n > 0 \right) \longrightarrow 1 \quad \text{cuando } n \longrightarrow \infty$$

de modo que $\hat{n}$ es asintoticamente insesgado en ese sentido.

A igual conclusion se llega mediante otro argumento. Sea el proceso paralelo Y_η en el cual $Y_0=1$; en el :

$$m.\left\{ 1 - f_m(0) \right\} \longrightarrow \frac{2}{\sigma^2} \quad \text{cuando } m \longrightarrow \infty \tag{3.21bis}$$

y como $E(Y_m / Y_m > 0)$ tiende a infinito con m, alguna vez pasara por q, no importando por tanto la condicion de que $Y_s=q$ puesto que s no esta fijado. Con ello :

$$\frac{\hat{n}}{n} = \frac{2}{\sigma^2} \cdot \frac{Y_m - q}{n} = \frac{2}{\sigma^2} \cdot \frac{m}{n} \cdot \frac{Y_m}{m} - \frac{2 \cdot q}{n \cdot \sigma^2}$$

se tiene que :

$$E\left(\frac{\hat{n}}{n} \,/\, Y_m > 0 \right) = \frac{2}{\sigma^2} \cdot \frac{m}{n} \cdot \frac{1}{m.\left\{1-f_m(0)\right\}} - \frac{2 \cdot q}{n \cdot \sigma^2}$$

Con ello, cuando $n \longrightarrow \infty$, dejando fijo s, entonces $m/n \longrightarrow 1$ y $m \longrightarrow \infty$ y finalmente :

$$E\left(\frac{\hat{n}}{n} \,/\, X_n > 0 \right) \longrightarrow 1 \quad \text{cuando } n \longrightarrow \infty$$

de nuevo, con independencia de cual sea la funcion generatriz f(s).

Para dar intervalos de confianza, el metodo mas conservador consiste en combinar los intervalos para m y s y asi dar un intervalo para n;ello solo sera valido cuando s sea grande y para que los intervalos no sean triviales hara falta que tambien n lo sea.Cuando m sea grande pero s sea chico,veamos de ensayar un intervalo aproximado.

Es conocido que :

$$\lim_{m \to \infty} P \left\{ \frac{2}{\sigma^2} \cdot \frac{Y_m}{m} \leq z \;/\; Y_m > 0 \right\} = \int_0^z e^{-x} dx \tag{3.22}$$

de modo que si $\alpha_1 + \alpha_2 = \alpha$ y

$$a_1 = - \ln (1 - \alpha_1)$$

$$a_2 = - \ln \alpha_2$$

entonces (a_2 , a_1) es un intervalo de confianza para

$$(2 Y_m) / (m. \sigma^2)$$

con una confianza del $(1-\alpha)$% cuando $m \longrightarrow \infty$ .Como por otro lado:

$$\frac{\hat{n}}{n} = \frac{2}{\sigma^2} \cdot \frac{K - q}{n} = \frac{2}{\sigma^2} \cdot \frac{K}{m} \cdot \frac{m}{n} - \frac{2}{\sigma^2} \cdot \frac{q}{n}$$

entonces, cuando $n \longrightarrow \infty$ tambien $m \longrightarrow \infty$ y por tanto un intervalo de confianza para $\hat{n} / n$ es :

$$\left[a_2 \cdot \frac{m}{n} \quad ; \quad a_1 \cdot \frac{m}{n} \right]$$

en el cual m aun es desconocido. A este intervalo-exacto para toda funcion generatriz con tal de que $n \longrightarrow \infty$ - podemos hacerle dos aproximaciones:

a) Cuando m sea grande y s chico, entonces $m/n \simeq 1$ y por consiguiente un intervalo para $\hat{n}/n$ es (a_2, a_1) y uno para n sera por tanto

$$(\hat{n}/a_1 \; ; \; \hat{n}/a_2)$$

Notese que de nuevo obtenemos un intervalo de anchura proporcional a $\hat{n}$ en el que la constante de proporcionalidad es la misma que en el intervalo para n cuando $X_0=1$ de la seccion 3.1.1.2. Lo unico que varia es el estimador $\hat{n}$ distinto en cada caso.

b) Cuando n y s sean grandes, entonces como:

$$a_2 \cdot \frac{n + s}{n} < \frac{\hat{n}}{n} < a_1 \cdot \frac{n + s}{n}$$

se tiene que:

$$a_2 n + a_2 s < \hat{n} < a_1 n + a_1 s$$

y finalmente:

$$\frac{\hat{n}}{a_1} - s < n < \frac{\hat{n}}{a_2} - s$$

de modo que estimando s-por el mismo argumento de secciones anteriores- por $2(q-1)/\sigma^2$ tenemos que:

$$\left[\frac{2(K-q)}{a_1 \sigma^2} - \frac{2(q-1)}{\sigma^2} ; \frac{2(K-q)}{a_2 \sigma^2} - \frac{2(q-1)}{\sigma^2} \right]$$

es un intervalo de confianza doblemente aproximado-una aproximacion porque n no es infinito y otra porque s tampoco esta exactamente determinado-de la edad n.Este intervalo es uno alternativo, al mas exigente que primero se propuso,y posiblemente mas conveniente,piensese en efecto que el intervalo para m hacia previsiones para las posibles variaciones de un proceso que comenzo con 1 solo individuo y que posteriormente paso por un estado q dado;como quiera que realmente el proceso comenzo ya con q individuos,entonces la anchura del intervalo para m ha previsto en realidad mas de lo necesario en dichas posibles variaciones del proceso,de modo que el conservar tal anchura no es solo apropiado sino que ademas sera posiblemente hasta estricto.Cuando q=1,dicho intervalo coincide con el de Stigler.

3.2.6.2. ESTUDIO DEL ESTIMADOR $\hat{q}$.-

Veamos ahora tambien otro modo de obtener el mismo estimador $\hat{q}$ de la (3.20). En las mismas condiciones de la subseccion anterior, un estimador de m es:

$$\hat{m} = 2 \cdot \frac{K-1}{\sigma^2}$$

de modo que un estimador del tiempo s - supuesta conocida la edad n - sera:

$$\hat{s} = \hat{m} - n = \frac{2}{\sigma^2}(K-1) - n$$

Con ello, parece logico estimar q por el tamaño medio

$$E(Y_{\hat{s}} \,/\, Y_m = K)$$

pero al ser desconocido m, habremos de conformarnos con estimarlo por :

$$\hat{q} = E(Y_{\hat{s}} \,/\, Y_{\hat{s}} > 0) = \frac{1}{1 - P(Y_{\hat{s}} = 0)}$$

y suponiendo el caso fraccional lineal

$$\hat{q} = \frac{2 + \hat{s}\,\sigma^2}{2} \qquad (3.23)$$

de modo que teniendo en cuenta el valor de $\hat{s}$:

$$\hat{q} = K - \frac{n\sigma^2}{2}$$

que es el mismo estimador dado por la (3.20).Naturalmente que cuando $\hat{q} < 1$ entonces se hara $\hat{q} = 1$.

Ademas,cuando s sea grande y para cualquier funcion generatriz:

$$E(\hat{q}) = E(Y_m / Y_m > 0) - \frac{n\sigma^2}{2} \simeq \frac{m.\sigma^2}{2} - \frac{n.\sigma^2}{2} =$$

$$= \frac{s.\sigma^2}{2} \simeq E(Y_s / Y_s > 0)$$

Cuando se trate del caso fraccional lineal entonces:

$$E(\hat{q}) = E(Y_s / Y_s > 0)$$

exactamente ya.

Veamos una propiedad asintotica de $\hat{q}$.Como:

$$E(X_n / X_n > 0) = \frac{q}{1 - f_m^q(o)}$$

entonces:

$$E\left(\frac{\hat{q}}{q} \,/\, X_n > 0\right) = \frac{1}{1 - f_n^q(o)} - \frac{n.\sigma^2}{2q} \longrightarrow 1$$

cuando $q \longrightarrow \infty$ pues $f_n(o) = P(Y_n \doteq 0 \,/\, Y_o = 1) < 1$ si la edad n se deja fija.Con ello $\hat{q}$ es asintoticamente insesgado en ese sentido.De igual modo puede verse que $\hat{q}$ es α - consistente para todo α mayor que 0,5.

Para dar intervalos de confianza para q,supongamos calculado un intervalo para m por el metodo de la seccion 3.1.1.2.; el era de la forma:

$$\left[\frac{\hat{m}}{a_2} \;;\; \frac{\hat{m}}{a_1}\right]$$

entonces,un intervalo para s sera:

$$\left[\frac{\hat{m}}{a_2} - n \;;\; \frac{\hat{m}}{a_1} - n\right] =$$

$$= \left[s_1 = \frac{\hat{m} - n.a2}{a_2} \;;\; s_2 = \frac{\hat{m} - n.a1}{a_1}\right]$$

de modo que un intervalo para q sera:

$$\left[E(Y_{s_1} / Y_{s_1} > 0) \; ; \; E(Y_{s_2} / Y_{s_2} > 0) \right]$$

o sea, la estimacion de q cuàndo el tiempo s fuera s_1 o s_2. Como:

$$s_1 = \frac{\frac{2(K-1)}{\sigma^2} - a_2 . n}{a_2} = \frac{2K - 2 - a_2 . \sigma^2 . n}{a_2 \sigma^2}$$

$$s_2 = \frac{\frac{2(K-1)}{\sigma^2} - a_1 . n}{a_1} = \frac{2K - 2 - a_1 . \sigma^2 . n}{a_1 \sigma^2}$$

entonces por la(3.23) :

$$\hat{q}_1 = \frac{2 + \sigma^2 . s_1}{2} = \frac{2a_2 + 2K - 2 - a_2 . \sigma^2 . n}{2a_2} =$$

$$= 1 + \frac{K-1}{a_2} - \frac{n . \sigma^2}{2}$$

$$\hat{q}_2 = 1 + \frac{K - 1}{a_1} - \frac{n . \sigma^2}{2}$$

con lo cual finalmente,un intervalo aproximado para q es :

$$\left[1 + \frac{K-1}{a_2} - \frac{n\sigma^2}{2} \; ; \; 1 + \frac{K-1}{a_1} - \frac{n\sigma^2}{2} \right]$$

en donde a_1 y a_2 son los mismos de la seccion 3.1.1.2. Notese que los intervalos dados son doblemente aproximados en cuanto que $\hat{m}$ era asintoticamente insesgado de m y en cuanto que q_1 y q_2 lo son asimismo con respecto a los verdaderos limites que se deberian de haber obtenido.

- - -

CAPITULO IV :

ESTIMACION NO PARAMETRICA EN PROCESOS DE GALTON-WATSON.

4.1.INTRODUCCION.-

En este ultimo capítulo abordaremos problemas de estimacion no paramétrica en los procesos de Galton-Watson.En una primera parte estudiaremos los estimadores clasicos de la media μ y la probabilidad de extincion π junto con el mas reciente estimador de la edad n.En la segunda parte se proponen dos nuevos estimadores de π y un nuevo estimador de n,los cuales se comparan,por simulacion,con los anteriores.Tambien se proponen estimadores no paramétricos del tamaño inicial X_0 de la poblacion.

4.2.RESULTADOS CONOCIDOS.-

4.2.1.ESTIMACION NO PARAMETRICA DE LA MEDIA μ DE UN PROCESO DE GALTON-WATSON.-

El estimador de la media μ es el mas antiguo en su obtencion;asi,dadas varias generaciones sucesivas

$X_0, X_1, X_2, \ldots, X_n$,el estimador

$$\hat{\mu} = \frac{X_1 + \ldots + X_n}{X_0 + \ldots + X_{n-1}}$$

es de maxima verosimilitud de μ.

Aun cuando $\hat{\mu}$ es generalmente sesgado, el siguiente teorema-debido a Crump y Howe - indica que es asintoticamente insesgado y que converge a μ con probabilidad 1 :

<u>Teorema 4.1</u>.- Si $V(X_1 / X_0=1) < \infty$ y en cualquiera de las dos situaciones siguientes

a) m fijo y $n \longrightarrow \infty$

b) n-m fijo y $m \longrightarrow \infty$

entonces:

1) $\hat{\mu} \longrightarrow \mu$ en probabilidad.

2) $E(\hat{\mu}) \longrightarrow \mu$

condicionando en ambos casos en la no extincion del proceso, y siendo $X_m, X_{m+1}, \ldots, X_n$ los tamaños generacionales conocidos.

El primero de los resultados anteriores ya habia sido probado por Harris con anterioridad.

4.2.2. ESTIMACION NO PARAMETRICA DE LA PROBABILIDAD DE EXTINCION π .-

En el capítulo III - seccion 3.1.2. - ya se abordo el problema de la estimacion de π (probabilidad de extincion) de un proceso de Galton-Watson para el caso de conocerse la forma de la funcion generatriz del número de hijos(estimacion paramétrica);supongamos ahora que dicha funcion generatriz no es conocida ni en su forma,ni en sus parámetros,y tratamos de estimar la probabilidad de extincion π de tal proceso.

Denominemos por f(s) a la funcion generatriz -desconocida- del número de hijos del proceso de Galton-Watson en cuestion,de modo que :

$$f(s)= \sum_{k} p_k . s^k$$

en donde

$$p_k= P(X_1=k \;/\; X_o=1 \;) \quad ,, \quad k=0,1,2,......$$

y siendo X_1 la variable aleatoria "número de hijos al final de la vida de un individuo inicial".Tomemos n individuos;ellos al final de su vida daran lugar a $X_1,X_2,$ ....,X_n hijos,de modo que X_i ,, i=1,2,....,n son variables aleatorias independientes e igualmente distribuidas.Con el fin de estimar π ,sean :

$$\hat{p}_k = \frac{\text{nº de } X_i \text{ que son igual a } k}{n} \quad ,, \; k=0,1,2,....$$

y notemos por

$$f_n(s) = \sum_{k} \hat{p}_k . s^k$$

a la funcion generatriz de probabilidad empirica. Como $\{\hat{p}_k\}$ son estimadores de maxima verosimilitud del $\{p_k\}$ entonces f_n es una estimada de maxima verosimilitud de f e igual con cualquier funcion de f_n con respecto a la misma funcion de f.

Sea $f^{(k)}(s)$ la derivada de orden k de f(s), con $f^{(o)} \equiv f$. Similarmente con $f_n(s)$. Veamos primero un teorema que nos da una serie de propiedades de $f_n^{(k)}$ como estimador de $f^{(k)}$; las pruebas de este teorema y de los siguientes pueden encontrarse en el articulo de Stigler (1971).

<u>Teorema 4.2</u>.-

a) Para cualquier $k \geq 0$ y cualquier $0 \leq s \leq s_1$, en donde s_1 es cualquier numero menor que uno, $f_n^{(k)}(s)$ es un estimador uniformemente fuertemente consistente de $f^{(k)}(s)$ en el sentido de que

$$f_n^{(k)}(s) \longrightarrow f^{(k)}(s) \;,, \text{ c.s.}$$

cuando $n \longrightarrow \infty$, uniformemente para $0 \leq s \leq s_1$. Ademas, si $f^{(k)}(1-) < \infty$, lo anterior vale para $s_1=1$ tambien.

b) Si $0 < p_o < 1$ y $p_o + \ldots\ldots + p_{k-1} < 1$, entonces :

$$\frac{f_n^{(k)}(s) - f^{(k)}(s)}{\left(V \left\{ f_n^{(k)}(s) \right\} \right)^{1/2}} \dashrightarrow N(0,1)$$

en distribucion, cuando $n \dashrightarrow \infty$, para cualquier $0 < s < 1$. Ademas :

$$n.V \left\{ f_n(s) \right\} = f(s^2) - \left\{ f(s)^2 \right\}$$

c)Si $f^{(k)}(s) < \infty$, entonces $f_n^{(k)}(s)$ es el estimador insesgado uniformemente de minima varianza.

Sea ahora el problema de estimar π. Para ello usaremos el estimador de maxima verosimilitud, esto es, la raiz no negativa mas chica de la ecuacion $f_n(s)=s$. Sea:

$$\hat{\pi} = \inf. \left\{ s \ / \ f_n(s)=s \ ,, \ s \geq 0 \right\}$$

Con ello, por el teorema 4.2.a), $\hat{\pi}$ es un estimador fuertemente consistente de π. Notese que $\hat{\pi}$ no es insesgado de π al no poder tomar valores mayores que 1. Veamos el comportamiento asintotico de $\hat{\pi}$:

Teorema 4.3.-

a)Si $f'(1-) > 1$ y $p_0 > 0$, entonces:

$$\sqrt{n} \, .(\hat{\pi} - \pi) \dashrightarrow N(0, \sigma^2)$$

en distribucion, cuando $n \dashrightarrow \infty$, en donde :

$$\sigma^2 = \frac{f(\pi^2) - \pi^2}{(1-f'(\pi))^2}$$

b)Si $f'(1-) = 1, p_1 < 1$ y $f''(1-) < \infty$, entonces:

$$P(\sqrt{n}\,.(1-\hat{\pi}) \leq x) \dashrightarrow \phi\left(\frac{1}{2}x.\sqrt{f''(1-)}\right) \qquad \text{si } x \geq 0$$

$$P(\sqrt{n}\,(1-\hat{\pi}) \leq x) \equiv 0 \qquad \text{si } x < 0$$

en donde ϕ es la funcion de distribucion de la normal estandar.

c)Si $f'(1-) < 1$, entonces :

$$P(\hat{\pi} = 1, \text{para } n \text{ suficientemente grande}) = 1$$

El teorema 4.3. puede usarse para dar intervalos de confianza aproximados para π. Por ejemplo, si $f'(1-) > 1$ y llamamos por

$$a_n^2 = \frac{f_n(\hat{\pi}^2) - \hat{\pi}^2}{(1 - f_n'(\hat{\pi}))^2}$$

entonces, por el teorema 4.2.a), T.4.3.a) y el teorema de Slutsky, la variable aleatoria

$$\frac{\sqrt{n}}{a_n} \circ (\hat{\pi} - \pi)$$

esta asintóticamente distribuida como una normal estandar. Con ello

$$\hat{\pi} \pm t_\alpha . \frac{a_n}{\sqrt{n}}$$

es un intervalo de confianza aproximado para π ,al error α ,en donde t_α se mira en una $N(0,1)$.

Otro parametro de gran interes en el estudio de los procesos de ramificacion supercriticos es $f'(\pi)$. Athreya-Ney han probado que si X_k es el tamaño de la poblacion en la generacion k de un proceso de Galton-watson con funcion generatriz de hijos f(s) y $m=f'(1-)>1$,entonces la diferencia entre la funcion de distribucion de $X_k.m^{-k}$ y su limite,es asintoticamente mas pequeño que ρ^{-k}, para cualquier

$$\rho < m^{\delta/(3+\delta)}$$

en donde $\delta = -\ln.f'(\pi)/\ln.m$ .Asimismo,$f'(\pi)$ tambien determina la intensidad en que la probabilidad de extincion para la generacion k se aproxima a π.

Si proponemos el estimar $f'(\pi)$ por el estimador de maxima verosimilitud $f'(\hat{\pi})$,entonces por el teorema 4.2.

$$f_n'(\hat{\pi}) \longrightarrow f'(\pi) \qquad \text{c.s.}$$

y la distribucion limite de $f_n'(\hat{\pi})$ puede calcularse de modo similar al teorema 4.3..Stigler(1971) presenta la solucion para cuando $f'(1-)>1$.

Volviendo al estimador no paramétrico $\hat{\pi}$ de π, supongamos que $f(s) = \exp.\{\lambda(s-1)\}$,es decir,la familia de Poisson.Entonces,si $\lambda>1$,por el teorema 4.3. :

$$V(\hat{\pi}/\lambda) = \frac{\exp.\{\lambda(\pi^2-1)\} - \pi^2}{(1-\lambda\pi)^2}$$

Si recordamos ahora el estimador paramétrico $\tilde{\pi}$ de la seccion 3.1.2.,entonces la eficiencia de $\hat{\pi}$ relativa a $\tilde{\pi}$ cuando el verdadero valor del parametro es λ es :

$$e(\lambda) = \frac{V(\tilde{\pi}/\lambda)}{V(\hat{\pi}/\lambda)}$$

dando Stigler una tabla de $e(\lambda)$ para diferentes valores de λ .El prueba que :

$$e(\lambda) \longrightarrow 1 \text{ cuando } \lambda \downarrow 1$$
$$e(\lambda) \longrightarrow 0 \text{ cuando } \lambda \longrightarrow \infty$$

Cuando f(s) sea la fraccional lineal,entonces por el teorema 4.3. :

$$V(\hat{\pi}/b,c) = \frac{1-b.(1-c)^{-1} + b.\pi^2.(1-c.\pi^2)^{-1} - \pi^2}{1-(1-c)^2.b^{-1}}$$

cuando $b > (1-c)^2$,en donde ahora

$$\pi = \frac{1-b-c}{c(1-c)}$$

Con ello la eficiencia

$$e(b,c) = \frac{V(\tilde{\pi}/b,c)}{V(\hat{\pi}/b,c)}$$

puede calcularse para diferentas valores de b y c (ver articulo de Stigler). Puede probarse que para cualquier $0 < c < 1$ fijo :

$$e(b,c) \longrightarrow 1 \text{ cuando } b \downarrow (1-c)^2 \text{ o } b \uparrow (1-c)$$

En resumen, es preferible usar el estimador no paramétrico $\hat{\pi}$ a menos que la muestra sea chica o bien que los calculos indiquen una eficiencia baja (como pasaba en la distribucion de Poisson con media grande); ha de tenerse tambien en cuenta el peligro de que la funcion generatriz verdadera no sea la propuesta, lo cual refuerza la eleccion del estimador no paramétrico $\hat{\pi}$.

Observese que el estimador no paramétrico $\hat{\pi}$ de π anterior esta basado en el conocimiento de una muestra aleatoria de la distribucion $\{p_k\}$ de hijos, pero es de preveer que ello no sea siempre posible. Supongamos entonces que los unicos datos a nuestro alcance sean el tamaño de la poblacion en varias generaciones sucesivas: $X_m, X_{m+1}, \ldots\ldots, X_n$ y que $X_0=1$. Bailey (1964, pg.64) da la aproximacion :

$$\pi \simeq \exp. \{ -2.(\mu -1) / \sigma^2 \}$$

obtenida por desarrollo en serie, en donde μ y σ^2 son la media y varianza de la distribucion de hijos $\{p_k\}$, la cual actua bien cuando π es cercano a uno. Con ello Crump y Howe (1972) proponen como posible estimador de π :

$$\hat{\pi}_1 = \exp. \{ -2.(\hat{\mu} -1) / \hat{\sigma}^2 \} \qquad (4.1)$$

en donde

$$\hat{\sigma}^2 = \frac{1}{n-m-1} \cdot \sum_{h=m+1}^{n} X_{h-1} \cdot \left(\frac{X_h}{X_{h-1}} - \hat{\mu} \right)^2 \qquad (4.2)$$

es un estimador de σ^2 y $\hat{\mu}$ es el estimador de Harris de la media μ:

$$\hat{\mu} = \frac{\sum_{h=m+1}^{n} X_h}{\sum_{h=m}^{n-1} X_h} \qquad (4.3)$$

Notese que para poder usar el estimador $\hat{\pi}_1$ hace falta el conocimiento de al menos 3 generaciones sucesivas del proceso.Los mismos autores anteriores proponen el estimador $\hat{\pi}_2$ obtenido como la raiz no negativa mas chica de la ecuacion :

$$s = \exp.\left\{ \hat{\mu}\,(s-1) \right\}$$

para lo cual basta con el conocimiento de dos generaciones.El sera el verdadero valor de π si la distribucion de hijos es de Poisson de parametro $\hat{\mu}$.

Notese asimismo que la estimacion de π ,en la hipótesis de muestreo anterior,es dificil por la escasa informacion que intuitivamente se ve que dan los tamaños $X_m, \ldots\ldots, X_n$.

Por simulacion de Monte Carlo para el caso fraccional lineal,Crump y Howe obtuvieron los siguientes resultados(ver tambien seccion 4.2.3.) :

1º)Los valores de $\hat{\pi}_2$ no dependen del verdadero valor de π ,sino unicamente del valor de μ .Esto es asi porque,para una distribucion de Poisson, π esta completamente determinado por μ ,y el estimador $\hat{\pi}_2$ se obtuvo a partir de dicha distribucion.

2º)La estimacion de π por $\hat{\pi}_1$ mejora conforme el numero de generaciones sucesivas observadas aumenta.Ello era logico de esperar pues el numero de observaciones es importante en la estimacion de la variamza dada por la (4.2). Finalmente,el estimador $\hat{\pi}_1$ actua mejor cuando π es grande y tiende a sobrestimar los pequeños valores de π .Ello es tambien logico porque la aproximacion (4.1) es valida para grandes π y tiende a ser demasiado grande para otros valores.

4.2.3.ESTIMACION NO PARAMETRICA DE LA EDAD n DE UN PROCESO DE GALTON-WATSON.-

Esta cuestion es tratada por Crump y Howe,a los cuales nos remitimos en las pruebas y conclusiones.

Sean $X_0=1,X_1,X_2,\ldots\ldots$ los tamaños de las generaciones sucesivas de un proceso de Galton-Watson con distribucion de hijos $\{p_k\}$ que tiene por funcion generatriz f(s).Supongamos que se han observado los tamaños $X_m,\ldots\ldots,X_n$ y que deseamos estimar la edad n del proceso en la ultima observacion.En adelante supondremos que la media μ de la distribucion de hijos es mayor que 1,pues de otro modo la poblacion se extingue con probabilidad uno,en cuyo caso el problema de estimacion lo mas seguro es que no aparezca.

Notese que la hipotesis de que $X_o=1$ no es contrastable y que esta intimamente ligada a la edad de la poblacion.

Ya vimos en la seccion 3.1.1.1. un estimador parametrico de n para cuando solo se observa una generacion,es decir,cuando m=n.El era :

$$\hat{n}_o = \ln.\left\{ (1-\pi).X_n + \pi \right\} / \ln \mu \qquad (4.4.)$$

en donde π era la probabilidad de extincion del proceso. Este estimador era de maxima verosimilitud cuando f(s) es del tipo fraccional lineal.

Notese que el conocimiento de generaciones adicionales no permite mejorar la estimada.En efecto,como X_h (h=0,1,2,...) es una cadena de Markov,entonces la densidad condicional conjunta de $X_{m+1},\ldots.,X_n$ dada X_m no envuelve a m y por tanto el estimador de maxima verosimilitud de n es :

$$\frac{\ln.\left\{ (1-\pi).X_m + \pi \right\}}{\ln \mu} + n - m$$

El problema radica en que para poder usar el estimador $\hat{n}_o$ hace falta el conocimiento previo de μ y π. De otro lado, ya sabemos que si $\sigma^2 < \infty$,entonces :

$$\left(\frac{X_n}{\mu^n} \;/\; X_n > 0 \right)$$

converge con probabilidad uno y en media cuadrada a una variable aleatoria W cuando n $\longrightarrow \infty$,en donde

$$E(\,W \;/\; \text{la poblacion no se extingue}) = \frac{1}{1-\pi}$$

Con ello,

$$E(\,X_n \;/\; X_n > 0) \simeq \mu^n/(1-\pi)$$

para grandes n. Resolviendo esta ecuacion aproximada en n:

$$\frac{\ln (1-\pi).X_n}{\ln \mu} \qquad (4.5)$$

es un estimador de n. A continuacion estudiaremos varios estimadores no paramétricos de n formados reemplazando en las (4.4) y (4.5) μ y π por estadisticos que los estimen.

Como estimador de μ tomamos el (4.3) de Harris, que es de maxima verosimilitud. Aunque $\hat{\mu}$ es generalmente insesgado, el teorema 4.1. nos indica que es asintoticamente insesgado y que converge a μ con probabilidad uno (condicionando en ambos casos en la no extincion).

La estimacion de π es mas dificil, por la poc a informacion que dan,sobre dicho parametro,los tamaños $X_m,\ldots,X_n$.Si observamos las (4.4) y (4.5),vemos que los valores de π tienen relativamente pequeños efectos sobre el estimador de la edad salvo cuando π esta cercano a uno. Asi,podemos usar los estimadores $\hat{\pi}_1$ o $\hat{\pi}_2$ propuestos en la seccion anterior.

Esta claro que n no admite un estimador consistente por la variabilidad de las primeras generaciones del proceso,sin embargo,aun cuando los estimadores de n sean necesariamente inconsistentes,el siguiente resultado debil permanece :

Teorema 4.3.- Supongamos que $E(X_1^{2+\delta}) < \infty$ para algun $\delta > 0$ y sea $\hat{n}$ un estimador de n obtenido de la (4.4) o (4.5) reemplazando μ por X_n / X_{n-1} y π por cualquier estadistico $\hat{\pi}$ no negativo y tal que

$$P(\hat{\pi} < 1 \;/\; X_n > 0) = 1$$

Entonces lim.sup.$(\hat{n}-n) < \infty$ con probabilidad uno(condicionado el proceso en la no extincion),con lo cual :

$$\frac{\hat{n} - n}{n^{\alpha}} \longrightarrow 0 \text{ cuando } n \longrightarrow \infty$$

con probabilidad uno(condicionado en la no extincion del proceso) para $\alpha > 0$,lo cual para $\alpha = 1$ es equivalente a que

$$\frac{\hat{n}}{n} \longrightarrow 1 \text{ cuando } n \longrightarrow \infty$$

con probabilidad uno(condicionado en la no extincion).

Notese que el no paramétrico es α-consistente $\forall \alpha > 0$, cosa que ya fue probada en la seccion 3.1.1.1. para el $\hat{n}_0$ parametrico,

El teorema vale pues para una amplia clase de estimadores de π, lo cual indica que la estimacion de π no es particularmente importante. El teorema esta enunciado para el caso de solo dos generaciones, pero es presumiblemente cierto en general.

Crump y Howe abordan a continuacion el problema de comparar por simulacion de Monte Carlo los diversos estimadores posibles de n. Ellos estudian los tres siguientes :

$$\hat{n}_0 = \frac{\ln\{(1-\pi)X_n + \pi\}}{\ln \mu}$$

$$\hat{n}_1 = \frac{\ln\{(1-\hat{\pi}_1)X_n + \hat{\pi}_1\}}{\ln \hat{\mu}}$$

$$\hat{n}_2 = \frac{\ln\{(1-\hat{\pi}_2)X_n + \hat{\pi}_2\}}{\ln \hat{\mu}}$$

en donde el primero es un estimador paramétrico, mientras que los dos siguientes son no paramétricos. Para la simulacion tomaron como distribucion de hijos el caso fraccional lineal, seleccionado por dos razones. Primero, porque la funcion generatriz fraccional lineal depende de dos parametros, de modo que son posibles todas las combinaciones que se deseen de π y μ .Segundo,porque $\hat{n}_0$ es un estimador de maxima verosimilitud cuando la distribucion de hijos es del tipo fraccional lineal,de modo que actuara bastante bien para este tipo de distribucion.Con ello, fijados los valores de los parametros b,c de la distribucion fraccional lineal,la comparacion de la media cuadrada de error para $\hat{n}_0$ con la de los otros dos estimadores no parametricos,indicara cuando los nuevos estimadores actuaran razonablemente bien.

Otros posibles estimadores,no parametricos,razonables de n serian :

$$\hat{n}_3 = \frac{\ln X_n}{\ln \hat{\mu}}$$

$$\hat{n}_4 = \frac{\ln\{(1-\hat{\pi}_1)X_n\}}{\ln \hat{\mu}}$$

$$\hat{n}_5 = \frac{\ln\{(1-\hat{\pi}_2)X_n\}}{\ln \hat{\mu}}$$

Podemos obtener $\hat{n}_3$ haciendo $\hat{\pi}=0$ en $\hat{n}_1$ o $\hat{n}_2$,de modo que este estimador es previsible que actuara bien cuando π sea cercano a cero.Como X_n y $(1-\pi)X_n$ difieren bastante para grandes π ,es de esperar que $\hat{n}_3$ sobreestime a n cuando π no sea cercano a cero.Los estimadores $\hat{n}_4$ y $\hat{n}_5$ se obtienen quitando la correccion aditiva del estimador $\hat{\pi}$ usado en $\hat{n}_1$ y $\hat{n}_2$ respectivamente.Por ejemplo,$\hat{n}_4$ y $\hat{n}_1$ estan relacionados por :

$$\hat{n}_1 = \hat{n}_4 + \frac{\ln \dfrac{1 + \hat{\pi}_1}{(1-\hat{\pi}_1)X_n}}{\ln \hat{\mu}}$$

de modo que las diferencias entre $\hat{n}_1$ y $\hat{n}_4$ seran pequeñas excepto cuando π sea grande y μ cercano a uno.Un estudio preliminar de Monte Carlo verificaron -indican los autores - estas observaciones.

Volviendo a los estimadores $\hat{n}_0$,$\hat{n}_1$ y $\hat{n}_2$,Crump y Howe han realizado 5.000 simulaciones para cada cuaterna de valores de m,n,π y μ ,calculandose en cada una de ellas las estimaciones $\hat{\pi}_1$,$\hat{\pi}_2$,$\hat{n}_0$,$\hat{n}_1$ y $\hat{n}_2$.En su articulo presentan dos tablas con los valores promedios y sus errores estandar para los estimadores de π y con las medias cuadradas de error y errores estandar para los estimadores de n,en cada una de las situaciones estudiadas.

Las conclusiones para los estimadores de π , $\hat{\pi}_1$ y $\hat{\pi}_2$,ya fueron expuestas en la seccion anterior.Las conclusiones para los estimadores de la edad n se enumeran a continuacion :

a)En todos los casos el error cuadratico medio es mas grande cuando μ esta cercano a uno.

b)En todos los casos,el error cuadratico medio es mas grande para valores de π grandes que para valores de π pequeños.

Con ello,todos los estimadores de la edad actuaran mejor cuando μ es grande y π es chico.

c) En todos los casos, los errores cuadraticos medios de los estimadores no aumentan apreciablemente cuando n aumenta con (n-m) fijado.

d) Los errores cuadraticos medios de $\hat{n}_1$ y $\hat{n}_2$ disminuyen con el aumento del numero de generaciones conocidas.

e) Cuando se conocen varias generaciones, $\hat{n}_1$ y $\hat{n}_2$ pueden competir con el estimador paramétrico $\hat{n}_0$, si bien $\hat{n}_1$ necesita de mas generaciones. Sin embargo, cuando μ es chico y π grande, lo anterior no es siempre cierto.

Con todo ello, como la funcion generatriz de hijos es desconocida normalmente y por tanto no estara a nuestro alcance los verdaderos valores de π y μ, no es descabellado el tomar $\hat{n}_1$ y $\hat{n}_2$ como unos estimadores apropiados. De hecho $\hat{n}_2$ tiene la ventaja de necesitar de solo dos generaciones sucesivas del proceso para ser calculado.

4.3. APORTACIONES.-

4.3.1. ESTIMACION NO PARAMETRICA DE LA PROBABILIDAD DE EXTINCION π DE UN PROCESO DE GALTON-WATSON.-

En la seccion 4.2.2. se propusieron dos estimadores $\hat{\pi}_1$ y $\hat{\pi}_2$, de la probabilidad de extincion, para cuyo calculo hacian falta un minimo de tres y dos generaciones sucesivas, respectivamente, del proceso, $X_m, X_{m+1}, \ldots, X_n = X_{m+r}$, pero no las edades respectivas $m, (m+1), \ldots$

....,m+r=n. El comportamiento de dichos estimadores fue ya estudiado entonces.

Recordemos la expresion (3.11) del capitulo anterior; ella nos proporciona un nuevo estimador no paramétrico de π si se sustituye μ por $\hat{\mu}$ (el estimador de la media de Harris); asi :

$$\hat{\pi}_3 = \frac{X_n - \hat{\mu}^n}{X_n - 1} \qquad \text{si } X_o=1 \qquad (4.6)$$

El estimador $\hat{\pi}_3$ puede justificarse aun de otro modo. Supongamos que se observa la variable aleatoria

$$w = \left(\frac{X_n}{\mu^n} \; / \; X_n > 0 \right)$$

para n grande; entonces la verosimilitud seria:

$$L = (1-\pi).e^{-(1-\pi).w}$$

de modo que cuando $w \geqslant 1$ el estimador de maxima verosimilitud de π para cuando $n \longrightarrow \infty$ es :

$$\hat{\pi} = \frac{X_n - \mu^n}{X_n}$$

que guarda un gran parecido con el $\hat{\pi}_3$. Cuando $w < 1$, la maxima verosimilitud para valores de π entre 0 y 1 es $\pi = 0$. Naturalmente, cuando $\hat{\pi}_3$ sea menor que 0 o mayor que 1, se hara $\hat{\pi}_3$ igual a 0 o 1 respectivamente.

Recordemos ahora el estimador mas general dado por la (3.16) em el capitulo III. Sustituyendo $\hat{\mu}$ por μ :

$$\hat{\pi}_4 = \frac{1}{1 + \dfrac{\hat{\mu}^n - 1}{X_n - q.\hat{\mu}^n}} \quad \text{si } X_0=q \qquad (4.7)$$

en donde $\hat{\pi}_4 = \hat{\pi}_3$ cuando $X_0=1$ como era de esperar.

Notese que el estimador $\hat{\pi}_4$ solo vale para procesos que, comenzando con un numero cualquiera X_0 de individuos, son observados en los tiempos $m,(m+1),\ldots,(m+r)=n$ en donde la edad n debe ser conocida, a diferencia de los estimadores $\hat{\pi}_1$ y $\hat{\pi}_2$ que no requerian el conocimiento de la edad n. Ello puede obviarse tomando X_m como tamaño de partida del proceso, con lo cual :

$$\hat{\pi}_5 = \frac{1}{1 + \dfrac{\hat{\mu}^r - 1}{X_n - X_m \cdot \hat{\mu}^r}} \qquad (4.8)$$

es ya un estimador de π equiparable (en cuanto a los datos que son necesarios para su calculo) con los $\hat{\pi}_1$ y $\hat{\pi}_2$. Notese que dicho estimador $\hat{\pi}_5$ solo es valido cuando se conocen mas de dos generaciones sucesivas del proceso.

En todos los casos, cuando algun estimador $\hat{\pi}$ de valores menores que cero o mayores que uno, se tomara $\hat{\pi}=0$ o $\hat{\pi}=1$ respectivamente.

Notese finalmente que cuando $\hat{\mu} > 1$, entonces $\hat{\pi}_4$

estara entre 0 y 1 si $X_n > q.\hat{\hat{\mu}}$,de modo que cuando ello no ocurra con el ultimo tamaño X_n convendra tomar otro intermedio X_h tal que $X_h > q.\hat{\hat{\mu}}$ .Cuando no haya ninguno, tomar entonces el ultimo tamaño X_n;cuando haya varios, tomar el que tenga mayor subindice h.De igual modo con $\hat{\pi}_5$.

4.3.2.OTROS ESTIMADORES NO PARAMETRICOS.-

El estimador $\hat{\pi}_5$ puede usarse alternativamente a los $\hat{\pi}_1$ y $\hat{\pi}_2$ para estimar la edad n,obteniendose entonces $\hat{n}_5$.

El estimador $\hat{q}$ dado por la (3.17) del capitulo III,puede convertirse en uno no parametrico sustituyendo μ por $\hat{\mu}$ y π por cualquiera de los $\hat{\pi}_1$, $\hat{\pi}_2$ o $\hat{\pi}_5$.

Los estimadores (3.12) y (3.12bis) de μ podrian tener sentido cuando solo se conociera una generacion X_n del proceso-y por tanto el estimador $\hat{\mu}$ de Harris no tendria validez - y el numero de hijos de algunos individuos de la generacion (n-1),pues entonces π podria estimarse no parametricamente como se describio en la seccion (4.2.2).

La proxima seccion se dedica a un estudio comparativo,por simulacion de Monte Carlo,de los estimadores $\hat{\pi}_1$, $\hat{\pi}_3$, $\hat{\pi}_5$, $\hat{n}_0$,$\hat{n}_1$ y $\hat{n}_5$,de los cuales el 1º,4º y 5º ya fueron estudiados por Crump y Howe en el articulo ya citado.

El estudio de cual de los estimadores de π es mas efectivo para estimar q y n - cuando X_0=q - sera fruto de un trabajo posterior.Asimismo queda abierta la comparacion del estimador de la media de Harris con los estimadores -cuando $X_o=1$ y cuando $X_o=q$ - no parametricos arriba propuestos.

4.3.3. ESTUDIO COMPARATIVO DE LOS DISTINTOS ESTIMADORES NO PARAMETRICOS DE π Y n .-

Por razones ya expuestas en otra ocasion, se ha elegido el caso fraccional lineal para realizar el estudio por simulacion de Monte Carlo de los diversos estimadores de π y n.

Para cada pareja de valores de π y μ de las indicadas en las tablas I y II, se han obtenido-por simulacion- los tamaños poblacionales $X_m, X_{m+1}, \ldots\ldots, X_n$ (que en adelante seran llamados por $Y_o, Y_1, \ldots, Y_r$ con r=n-m) para los valores de n y m dados asimismo en dichas tablas. Lo anterior se hizo en 1.000 ocasiones para cada cuaterna π, μ, m, n. A partir de cada serie de valores de $Y_o, Y_1, \ldots\ldots Y_r$ se obtuvieron los diversos estimadores propuestos. El criterio por nosotros seguido para la obtencion de tales estimadores lo sumarizamos a continuacion.

En primer lugar deben de encontrarse dos valores s y h, con s > h, de entre los 0,1,....,r tales que el estimador de la media

$$\hat{\mu} = \frac{Y_{h+1} + Y_{h+2} + \ldots\ldots + Y_s}{Y_h + Y_{h+1} + \ldots\ldots\ldots + Y_{s-1}}$$

sea mayor que uno. Ello debe de hacerse con la condicion de que (s-h) sea maximo y, si hubiese varios, se escogera el de mayor valor de s. Cuando tal pareja no exista, entonces se hara

$$\hat{\pi}_1 = \hat{\pi}_3 = \hat{\pi}_5 = 1$$

$$\hat{n}_1 = \hat{n}_5 = r+1$$

y se calcula $\hat{n}_o$ como mas adelante se indica.

Cuando existan dichos valores de s y h, calcular entonces $\hat{\pi}_1$, $\hat{\pi}_3$, $\hat{\pi}_5$, $\hat{n}_o$, $\hat{n}_1$, $\hat{n}_5$ como se indica mas adelante, a partir, cuando ha lugar, del estimador $\hat{\mu}$ anterior. La unica salvedad es cuando (s-h)=1, en cuyo caso debe de tenerse la precaucion de que la pareja (s,h) que se utilice para el calculo de $\hat{\pi}_5$ no debe de ser la misma que sirvio para el calculo de $\hat{\mu}$ no debe de ser la misma que sirvio para el calculo de $\hat{\mu}$. Si es asi, buscar otra pareja para $\hat{\pi}_5$.

El calculo de los estimadores de π y n a que nos hemos venido refiriendo se realiza del modo siguiente:

a) Calculo de $\hat{\pi}_1$:

$$\hat{\sigma}^2 = \frac{1}{r-1} \cdot \sum_{h=1}^{r} Y_{h-1} \left(\frac{Y_h}{Y_{h-1}} - \hat{\mu} \right)^2$$

$$\hat{\pi}_1 = \exp. \left\{ -2. (\hat{\mu} - 1) / \hat{\sigma}^2 \right\}$$

b) Calculo de $\hat{\pi}_3$:

$$\hat{\pi}_3 = \frac{Y_h - \hat{\mu}^{(h+m)}}{Y_h - 1}$$

en donde h es el mayor natural de entre los 0,1,2,...,r tal que

$$Y_h > \hat{\mu}^{(h+m)}$$

Si no existe un tal valor de h,hacer $\hat{\pi}_3 = 0$.

c)<u>Calculo de $\hat{\pi}_5$</u>:

$$\hat{\pi}_5 = \frac{1}{1 + \dfrac{\hat{\mu}^{(s-h)} - 1}{Y_s - Y_h \cdot \hat{\mu}^{(s-h)}}}$$

en donde s y h son dos naturales de entre los 0,1,2,.. ...,r tales que

$$s > h \quad ,, \quad Y_s > Y_h \cdot \hat{\mu}^{(s-h)}$$

elegidos de modo que (s-h) sea maximo.Si hay varias de tales parejas de valores de s y h,tomar la de mayor valor de s.Cuando no existan tales valores,hacer s=r,h=0 y calcular $\hat{\pi}_5$.Naturalmente,cuando $\hat{\pi}_5 \leq 0$ se hara $\hat{\pi}_5 = 0$, y cuando $\hat{\pi}_5 \geq 1$ se hara $\hat{\pi}_5 = 1$.

d)<u>Calculo de los estimadores de la edad</u>:

Hemos indicado tres:$\hat{n}_0, \hat{n}_1, \hat{n}_5$.Ellos se calculan asi:

$$\hat{n}_0 = \frac{\ln\{Y_h(1-\pi) + \pi\}}{\ln \mu} + (r-h)$$

$$\hat{n}_1 = \frac{\ln\{Y_h(1-\hat{\pi}_1) + \hat{\pi}_1\}}{\ln \hat{\mu}} + (r-h)$$

$$\hat{n}_5 = \frac{\ln\{Y_h(1-\hat{\pi}_5) + \hat{\pi}_5\}}{\ln \hat{\mu}} + (r-h)$$

en donde, en todos los casos, h es el mayor natural de entre los $0,1,2,\ldots,r$ tal que:

$$\hat{n}_x > (r+1)$$

Cuando no exista un tal valor de h, hacer $\hat{n}_x = r+1$.

Con todo ello, obtenidos los $\hat{\pi}_1$, $\hat{\pi}_3$, $\hat{\pi}_5$, $\hat{n}_0$, $\hat{n}_1$, $\hat{n}_5$ en los mil casos, se calcula para cada uno de ellos la media, error de la media y la raiz cuadrada del error cuadratico medio, en donde entendemos por error cuadratico medio del estimador x a la cantidad

$$\frac{\sum_{i=1}^{p} (x_i - \alpha)^2}{p}$$

siendo $x_1, x_2, \ldots, x_p$ las p estimaciones del parametro α.

En las tablas I y II se indican los resultados finales para los estimadores de π y n respectivamente. La simulacion y calculos han sido realizados en un ordenador UNIVAC 1108 a traves de una terminal DCT 2000 del Centro de Calculo de la Universidad de Granada.

Observando la tabla I se llega a las siguientes conclusiones para los estimadores de π:

a) Los estimadores $\hat{\pi}_1$ y $\hat{\pi}_3$ sobreestiman los valores pequeños de π y subestiman los valores grandes.

El estimador $\hat{\pi}_5$ siempre sobreestima a π, tanto mas cuanto menor sea μ.

b) Los estimadores $\hat{\pi}_1$ y $\hat{\pi}_5$ actuan mejor para grandes valores de π, mientras que el $\hat{\pi}_3$ lo hace peor.

Los tres estimadores mejoran con el aumento de la media μ.

c) Si m se deja fijo y se aumenta $(n-m)$:

- $\hat{\pi}_1$ mejora, salvo si μ es grande y π pequeño, en cuyo caso no varia.
- $\hat{\pi}_3$ mejora para valores pequeños de μ.
- $\hat{\pi}_5$ mejora para valores pequeños de π y de μ.

d) Si $(n-m)$ se deja fijo y se aumenta m:

- $\hat{\pi}_1$ mejora algo cuando π y μ son pequeños
- $\hat{\pi}_3$ practicamente no varia
- $\hat{\pi}_5$ empeora para valores pequeños de π.

e) Comparando los estimadores por parejas:

- $\hat{\pi}_3$ es mejor que $\hat{\pi}_1$ para valores grandes de μ, salvo cuando el numero de generaciones observadas y π sean grandes.
- $\hat{\pi}_5$ es mejor que $\hat{\pi}_1$ para valores grandes de π.
- $\hat{\pi}_3$ es mejor que $\hat{\pi}_5$ para valores pequeños de π.

f) Comparando los estimadores en conjunto:

- Cuando π sea grande es preferible el estimador $\hat{\pi}_5$.
- Cuando π y μ sean pequeños, es preferible $\hat{\pi}_1$.
- Cuando π sea pequeño y μ grande, es preferible $\hat{\pi}_3$.

De igual modo, por observacion de la tabla II, se llega a las siguientes conclusiones para los estimadores de n:

a) El estimador $\hat{n}_0$ subestima siempre a n, cada vez menos conforme aumenta el valor de π.

El estimador $\hat{n}_1$ subestima a n cuando π es pequeño. El estimador $\hat{n}_5$ subestima a n, cuando π es chico, de un modo apreciable, y tanto mas conforme aumenta μ.

b) Los estimadores $\hat{n}_1$ y $\hat{n}_5$ actuan mejor cuando π es grande que cuando π es pequeña, mientras que $\hat{n}_0$ no varia demasiado.

De otro lado, los tres estimadores $\hat{n}_0, \hat{n}_1, \hat{n}_5$ mejoran susceptiblemente cuando μ es grande respecto a cuando μ es pequeño.

c) Si m se deja fijo y (n-m) se aumenta :

- $\hat{n}_0$ empeora debilmente cuando π es grande y mejora cuando π es pequeña.
- $\hat{n}_1$ mejora sensiblemente cuando μ es pequeño y tiene poca variacion cuando μ es grande.
- $\hat{n}_5$ empeora debilmente cuando μ es grande y suele mejorar cuando μ es pequeño.

d) Si (n-m) se deja fijo y se aumenta m :

- $\hat{n}_5$ empeora sensiblemente.
- $\hat{n}_0$ y $\hat{n}_1$ suele empeorar sobretodo cuando μ es pequeña.

e) Claramente, el estimador $\hat{n}_0$ es siempre mejor que los $\hat{n}_1$ y $\hat{n}_5$ como era de esperar al ser un estimador parametrico.

Para el caso en que π sea desconocida y haya que usar los estimadores no parametricos (que sera lo mas usual), el estimador $\hat{n}_1$ actua siempre mejor que el $\hat{n}_5$ cuando μ es grande, y son desiguales cuando μ es cercano a uno. Con ello, el estimador $\hat{n}_5$ tiene escasa validez frente a los $\hat{n}_0$ y $\hat{n}_1$.

El que $\hat{\pi}_5$ fuera mejor que $\hat{\pi}_1$ cuando π era grande y no ocurra asi con $\hat{n}_5$ y $\hat{n}_1$, es explicable porque la sobreestimacion de π perjudica notablemente a $\hat{n}$ y el $\hat{\pi}_5$ siempre sobreestima a π .

Granada, Primavera del 76

TABLA I :Valores obtenidos para los estimadores de π

$\mu = 1.3$

m,n		$\pi = 0.2$ Media	E.E.	Raiz de E.C.M.	$\pi = 0.75$ Media	E.E.	Raiz de E.C.M.
6,8	$\hat{\pi}_1$	.42	.0389	.45	.61	.0363	.39
	$\hat{\pi}_3$	.43	.0422	.48	.52	.0393	.45
	$\hat{\pi}_5$	.62	.0324	.53	.76	.0243	.24
6,10	$\hat{\pi}_1$	.39	.0298	.35	.73	.0236	.24
	$\hat{\pi}_3$	.32	.0344	.36	.62	.0340	.36
	$\hat{\pi}_5$	.45	.0269	.37	.77	.0215	.22
10,12	$\hat{\pi}_1$	.33	.0355	.38	.57	.0364	.41
	$\hat{\pi}_3$	.43	.0382	.44	.50	.0402	.47
	$\hat{\pi}_5$	.62	.0302	.52	.77	.0228	.23
10,14	$\hat{\pi}_1$	.35	.0271	.31	.70	.0225	.23
	$\hat{\pi}_3$	.33	.0339	.36	.62	.0360	.38
	$\hat{\pi}_5$	.57	.0277	.46	.77	.0231	.23

E.E.=Error Estandar

E.C.M.=Error Cuadratico Medio

TABLA I : Continuacion

$\mu = 2$

m,n		π =0.2			π =0.75		
		Media	E.E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.
6,8	$\hat{\pi}_1$	.29	.0308	.32	.64	.0348	.36
	$\hat{\pi}_3$	.24	.0282	.28	.62	.0313	.34
	$\hat{\pi}_5$	.69	.0247	.55	.89	.0138	.20
6,10	$\hat{\pi}_1$	.40	.0263	.33	.76	.0224	.22
	$\hat{\pi}_3$	.23	.0274	.28	.57	.0326	.37
	$\hat{\pi}_5$	.75	.0224	.59	.89	.0168	.22
10,12	$\hat{\pi}_1$	.32	.0308	.33	.65	.0346	.36
	$\hat{\pi}_3$	.22	.0269	.27	.60	.0312	.35
	$\hat{\pi}_5$	.88	.0152	.70	.97	.0056	.23
10,14	$\hat{\pi}_1$	.43	.0242	.33	.78	.0183	.19
	$\hat{\pi}_3$	.21	.0267	.27	.50	.0358	.44
	$\hat{\pi}_5$	.88	.0153	.70	.95	.0100	.22

E.E.= Error Estandar

E.C.M.=Error Cuadratico Medio

TABLA II : Valores obtenidos para los estimadores de n.

m,n		$\mu = 1.3$					
		$\pi = 0.2$			$\pi = 0.75$		
		Media	E.E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.
6,8	$\hat{n}_0$	6.79	.3056	3.29	7.14	.2678	2.81
	$\hat{n}_1$	6.92	.5058	5.17	8.76	1.1190	11.22
	$\hat{n}_5$	5.84	.4593	5.08	7.03	.5639	5.72
6,10	$\hat{n}_0$	8.27	.2667	3.18	8.99	.2650	2.84
	$\hat{n}_1$	8.70	.3785	4.00	9.96	.5066	5.07
	$\hat{n}_5$	8.21	.3608	4.03	10.14	.8235	8.24
10,12	$\hat{n}_0$	10.03	.3891	4.36	10.42	.3754	4.07
	$\hat{n}_1$	11.45	.7743	7.76	13.55	1.3965	14.05
	$\hat{n}_5$	8.60	.5934	6.84	11.61	2.0972	20.98
10,14	$\hat{n}_0$	12.17	.4022	4.42	11.83	.4038	4.58
	$\hat{n}_1$	12.89	.4913	5.04	14.96	.7903	7.96
	$\hat{n}_5$	11.04	.4794	5.63	13.24	.8958	8.99

E.E. = Error Estandar

E.C.M. = Error Cuadratico Medio

TABLA II : Continuacion

$\mu = 2$

m,n		$\pi = 0.2$ Media	$\pi = 0.2$ E.E.	$\pi = 0.2$ Raiz de E.C.M.	$\pi = 0.75$ Media	$\pi = 0.75$ E.E.	$\pi = 0.75$ Raiz de E.C.M.
6,8	$\hat{n}_o$	7.11	.1678	1.90	7.23	.1753	1.91
	$\hat{n}_1$	7.12	.1798	2.00	7.33	.2329	2.42
	$\hat{n}_5$	5.63	.1447	2.78	5.54	.1779	3.04
6,10	$\hat{n}_o$	9.35	.1638	1.76	9.12	.1867	2.06
	$\hat{n}_1$	8.81	.1711	2.08	8.83	.2277	2.56
	$\hat{n}_5$	7.28	.1484	3.10	7.37	.2061	3.34
10,12	$\hat{n}_o$	11.17	.1877	2.05	11.60	.1601	1.65
	$\hat{n}_1$	10.76	.2096	2.43	11.28	.2271	2.38
	$\hat{n}_5$	7.66	.1719	4.67	7.50	.1723	4.82
10,14	$\hat{n}_o$	13.42	.1616	1.72	13.01	.1939	2.18
	$\hat{n}_1$	12.81	.1678	2.06	12.43	.2144	2.66
	$\hat{n}_5$	9.63	.1925	4.77	9.25	.1920	5.12

E.E. = Error Estandar

E.C.M. = Error Cuadratico Medio

B I B L I O G R A F I A
Y
R E F E R E N C I A S

I.-ATHREYA,K.B. y NEY,P.E.(1972)."Branching Processes". Berlin:Springer-Verlag.

II.-BAILEY,N.T.J.(1964)."The Elements of Stochastic Processes with applications to the Natural Sciences".New York: Wiley.

III.-BARTLETT,M.S.(1960)."Modelos estocasticos para poblaciones(en Ecologia y Epidemiologia)".Methuen's monographs on applied probability and statistics.

IV.-CHIN LONG CHIANG,(1968)."Introduction to Stochastic Processes in Biostatistics".New York:Wiley.

V.-COX,D.R.(1970)."Renewal Theory".Methuen and Co.Ltd.- Science Paperbacks.

VI.-COX,D.R. y MILLER,H.D.(1970)."The Theory of Stochastic Processes".London:Methuen and Co.Ltd.

VII.-CRUMP,K.S. y HOWE,R.B.(1972)."Nonparametric estimation of the age of a Galton-Watson branching process".Biometrika 59,533-38.

VIII.-FELDMAN,D. y FOX,M.(1968)."Estimation of the parameter n in the binomial distribution".J.Am.Statist.Assoc.63, 150-8.

IX.-HARRIS,T.E.(1963)."The Theory of Branching Processes". Berlin:Springer-Verlag.

X.-IOSIFESCU,M. y TAUTU,P.(1973)."Stochastic Processes and Applications in Biology and Medicine.I-Theory".Biomathematics-Volume 3.Berlin:Springer-Verlag.

XI.-KARLIN,S.(1971)."A First Course in Stochastic Processes". New York:Academic Press.

XII.-KEIDING,N.(1974)."Estimation in the birth process". Biometrika 61,71-80.

XIII.-MORAN,P.A.P.(1951)."Estimation methods for evolutive processes".J.R.Statist.Soc. B13,141-6.

XIV.-PARZEN,E.(1972)."Procesos estocasticos".Madrid:Paraninfo.

XV.-STIGLER,S.M.(1970)."Estimating the age of a Galton-Watson branching process".Biometrika 57,505-12.

XVI.-STIGLER,S.M.(1971)."The estimation of the probability of extinction and other parameters associated with branching processes".Biometrika 58,499-508.

\- - -