TRABAJOS DE ESTADISTICA Y
DE INVESTIGACION OPERATIVA
Vol. 32, Núm. 1, 1981, pp. 52 a 69

ESTIMACION DE LA EDAD Y DEL NUMERO INICIAL DE INDIVIDUOS EN PROCESOS DE NACIMIENTO PURO Y DE GALTON-WATSON

A. Martin Andres
Sección de Bioestadística
Facultad de Medicina
(Universidad de Granada)

SUMARIO

Se dan estimaciones puntuales y por intervalo para la edad y el número inicial de individuos en procesos de nacimiento puro de intensidad conocida y en procesos de Galton-Watson con distribución de descendientes conocida.

SUMMARY

Point estimates and confidence intervals for the age and the individual initial number are given for a pure birth process of known intensity and a Galton-Watson process with known offspring distribution.

Algunas palabras claves: Estimación en procesos de Galton-Watson; estimación en procesos de nacimiento; procesos de Galton-Watson; procesos de nacimiento puro.

1. ESTIMACION EN PROCESOS DE NACIMIENTO PURO LINEAL

1.1. Introducción

Sea X_t el tamaño de la población en el tiempo t en un proceso de nacimiento puro, esto es, en un proceso de Markov en el cual, para $i = 1, 2, ...$

$$pr\,(X_{t+h} = j \mid X_t = i) = \begin{cases} i\,\lambda\, h + 0\,(h) & (j = i + 1) \\ 1 - i\,\lambda\, h + 0\,(h) & (j = i) \\ 0\,(h) & \text{en otro caso} \end{cases}$$

en donde λ, la intensidad del proceso, es mayor que cero. Supongamos que $pr\,(X_0 = q) = 1$, en donde q es un entero positivo fijado de antemano.

Es bien conocido (Harris, 1963, págs. 104-147) que si X_t es un proceso de nacimiento con $X_0 = q$, entonces la esperanza $E\,(X_t) = q \cdot e^{\lambda t}$, y existe una variable aleatoria W tal que $X_t/E\,(X_t) \to W$ casi seguro cuando $t \to \infty$. La distribución de W es una gamma (q, q), esto es, tiene por densidad $q^q \cdot w^{q-1} \cdot e^{-q w}/\Gamma\,(q)$ para $w > 0$ y, por tanto, $E\,(W) = 1$.

Un problema abordado por diversos autores (Keiding, 1974 y Moran, 1951 entre otros) ha sido el de la estimación del parámetro λ bajo diversos esquemas de muestreo; en particular Keiding (1974) prueba el siguiente teorema de interés posterior:

Teorema 1.1. El estimador de máxima verosimilitud $\hat{\lambda}$ basado en la observación de X_t solamente, viene dado por $\hat{\lambda}\, t = \log\,(X_t/q)$. Además, cuando $t \to \infty$, $t\,(\hat{\lambda} - \lambda) \to \log W$ casi seguro, en donde la distribución de W es una gamma (q, q).

Sin embargo, el problema de estimar la edad t de un proceso de nacimiento o el de estimar el número inicial q de individuos de que parte el proceso, son situaciones que no han sido aún tratadas y que se estudian en la presente sección 1.

1.2. Estimación de la edad

Es conocido que si X_ξ es un proceso de nacimiento puro y de intensidad λ, entonces

$$pr\,(X_{s+t} = n \mid X_s = m) = \binom{n-1}{m-1} (e^{-\lambda t})^m \cdot (1 - e^{-\lambda t})^{n-m}$$

de modo que por la simetría existente entre λ y t se deduce que, probabilísticamente, es lo mismo hablar de un proceso de nacimiento de intensidad λ que comenzando con m individuos se encuentra en el estado n después de un tiempo t, que de uno de intensidad t que partiendo de m individuos se encuentra en el estado n después de un tiempo λ. Si el proceso de intensidad λ le llamamos proceso original, al de intensidad t le llamaremos "proceso conjugado" del original.

Si lo que se desea es estimar la edad t de un proceso de intensidad λ que partiendo de q individuos ocupa el estado X_t en el tiempo t, ello será lo mismo que estimar la intensidad t del proceso conjugado; con ello:

Teorema 1.2. El estimador de máxima verosimilitud $\hat{t}$ basado solamente en la observación de X_t viene dado por

$$\hat{t} = \frac{1}{\lambda} \cdot \log\,(X_t/q) \qquad (1)$$

Además $\lambda\,(\hat{t} - t) \to \log W$ casi seguro (a) cuando $\lambda \to \infty$ (b) cuando $t \to \infty$, en donde W es la misma variable descrita anteriormente.

Prueba: El estimador (1) y (a) son consecuencia de aplicar el teorema 1.1 al proceso conjugado. La parte (b) es inmediata por la convergencia casi segura $X_t/E\,(X_t) \to W$.

Cuando el objeto sea estimar la edad t, como es el caso, será usual desconocer $X_0 = q$ de modo que no queda otro remedio que suponer $X_0 = q = 1$. En ese caso W tiene por densidad e^{-w}, de modo que un intervalo de confianza para log W será $[w_1 = \log \{-\log (1 - \alpha_1)\}$, $w_2 = \log (-\log \alpha_2)]$, en donde $\alpha_1 + \alpha_2 = \alpha$ son los errores de cola. Para un valor dado de α, dicho intervalo será de longitud mínima si se toma α_1 como la única raiz positiva de la ecuación $(\alpha - \alpha_1) \cdot \log (\alpha - \alpha_1) = (1 - \alpha_1) \log (1 - \alpha_1)$. Stigler (1970) da una tabla de valores de α_1 para algunos valores de α. Dado el intervalo $[w_1, w_2]$ para log W, uno aproximado para la edad t se obtiene de inmediato a través del teorema 1.2.

Como solo se observa el tamaño actual de la población y no las fluctuaciones aleatorias previas, es demasiado pedir la consistencia del estimador en el sentido de que $|\hat{t} - t|$ tienda a cero para grandes valores de t. Sin embargo, veremos que $\hat{t}$ es un estimador relativamente consistente para grandes valores de t.

Siguiendo a Feldman and Fox (1968), diremos que el estimador $\hat{\theta}$ de θ es α-consistente si $(\hat{\theta} - \theta)/\theta^{\alpha} \to 0$ en probabilidad cuando $\theta \to \infty$. De ahí que, por el teorema 1.2. (b), el estimador $\hat{t}$ de la edad t sea α-consistente para cualquier $\alpha > 0$.

1.3. Estimación del número inicial de individuos

Sea un proceso de nacimiento X_ξ de intensidad λ que, partiendo de un número inicial de individuos $X_0 = q$ desconocido, es observado en el tiempo n en el cual $X_n = K$. Ahora se desea un estimador del tamaño inicial q.

Sea para ello Y_η un proceso de nacimiento de igual intensidad λ que, partiendo de un solo individuo, ocupa el estado q en un tiempo s desconocido, comportándose en adelante como el proceso original X_ξ pero en la nueva escala de tiempos $\eta = \xi + s$. Con ello $Y_0 = 1$, $Y_s = q$, $Y_t = X_n = K$ con $t = n + s$. Un tal proceso será denominado por "proceso paralelo" al original. Por el teorema 1.2, una estimación de la edad t es $\hat{t} = \log K/\lambda$, de modo que una estimación de la edad s será $\hat{s} = \hat{t} - n = (\log K - \lambda n)/\lambda$. Para estimar q parece apropiado utilizar el tamaño medio $E(Y_{\hat{s}} \mid Y_t = K)$, pero como t es desconocido ello no es factible. Siguiendo en la misma línea, podríamos utilizar la expresión $E(Y_{\hat{s}} \mid Y_{\hat{s}} \leqslant K)$, lo cual da una forma excesivamente complicada para $\hat{q}$. De cualquier modo, cuando K sea suficientemente grande, la expresión anterior puede aproximarse a la que finalmente se propone: $\hat{q} = E(Y_{\hat{s}}) = = \exp \cdot (\lambda \hat{s}) = \exp \cdot (\log K - \lambda n)$, por lo cual

$$\hat{q} = \frac{X_n}{e^{\lambda n}} \tag{2}$$

Nótese que $E(\hat{q}) = q$ y que $V(\hat{q}) = q \cdot \{\exp(\lambda n) - 1\} / \exp.(\lambda n)$, con lo cual $\hat{q}$ es insesgado pero de varianza tanto más grande cuanto más nos alejemos del comienzo del proceso y cuanto mayor sea q. Por esta pérdida de memoria, en el sentido anterior, con el tiempo y con q, está claro que pedir la consistencia de $\hat{q}$ es excesivo. Alternativamente la α-consistencia si es accesible. En efecto, si hacemos $x = (\hat{q} - q)/q^{\alpha}$, entonces $E(x) = 0$ y $V(x) = \{\exp \cdot (\lambda n) - 1\}/\exp(\lambda n) \cdot q^{2\alpha - 1} \to 0$ cuando $q \to \infty$; de ahí que x converja en media cuadrada a cero, para cualquier $\alpha > 0{,}5$ y, por tanto, lo hará también en probabilidad. Con ello $\hat{q}$ es un estimador α-consistente de q para cualquier $\alpha > 0{,}5$.

Para dar intervalos de confianza, como $z = X_n / E(X_n)$ converge casi seguro a W, cuando $n \to \infty$, y W tiene por distribución la gamma (q, q), se deduce que $\hat{q} = q \cdot z \to W'$ casi seguro, siendo W' una gamma $(q, 1)$. Con ello pueden construirse intervalos de confianza aproximados de q, para grandes valores de n.

2. ESTIMACION EN PROCESOS DE GALTON-WATSON

2.1. Introducción

En los últimos años ha habido un gran crecimiento en el número de trabajos publicados sobre las propiedades de los procesos de ramificación de Galton-Watson. La gran mayoría de dichos trabajos estudian el comportamiento porbabilístico de estos procesos, interesándose en particular por el crecimiento asintótico o la extinción de la población cuando la edad y la distribución de descendientes están totalmente especificadas. Mas recientemente, han merecido atención varios problemas de estimación sobre los parámetros de tales procesos: estimación del número medio de descendientes por individuo (Harris, 1963, pág. 32), estimación de la probabilidad de extinción del proceso, paramétrico (Stigler, 1970) o no paramétrica (Crump y Howe, 1972), cuando se supone que éste comienza con un sólo individuo. Sin embargo, el problema de estimar la edad del proceso cuando comienza con q individuos o el de estimar dicho número inicial de individuos, son situaciones aún no abordadas y que tratamos de resolver aquí. En este trabajo nos limitamos a dar las estimaciones paramétricas de la edad y del número inicial de individuos, reservando para uno posterior, el estudio de los estimadores no paramétricos.

Sean $X_0 = q$, X_1, X_2, ... los tamaños de las generaciones sucesivas de un proceso de ramificación de Galton-Watson que comienza con q individuos y cuya función generatriz de probabilidad del número de descendientes es $f(s) = \Sigma\, p_k \cdot s^k$ con $p_k = pr$ (el número de descendientes de un individuo dado sea igual a k). Esto es, la población comienza en la generación cero con q individuos, los cuales viven una unidad de longitud de tiempo, o sea una generación; y entonces mueren, produciendo cada uno de ellos, de modo independiente a los demás, un número aleatorio de descendientes de acuerdo con la distribución de probabilidad $\{p_k\}$. Todos los descendientes de la generación cero pasan a formar la generación uno y se comportan como sus padres para así dar lugar a la generación dos, etc. En adelante supondremos que $f(s)$ es conocido, esto es, conocemos la distribución de probabilidad del número de descendientes de un sólo individuo. Tales procesos son discutidos en detalle por Harris (1963).

Si denominamos por $\mu = f'(1-)$ el número medio de descendientes producidos por un sólo individuo, entonces son posibles tres casos:

(i) $\mu > 1$, el caso supercrítico;
(ii) $\mu = 1$, el caso crítico;
(iii) $\mu < 1$, el caso subcrítico.

En cada uno de ellos, la obtención de los estimadores propuestos es distinta.

Otro parámetro de interés es la probabilidad de extinción π de un proceso que comienza con un solo individuo. Es conocido (Harris, 1963, pág. 7) que π es la raiz positiva más pequeña de $f(s) = s$, y que $\pi < 1$ sí y sólo si $\mu > 1$. Tal parámetro representa pues la probabilidad de que la población eventualmente se extinga. Es evidente, por la independencia de actuación de los individuos, que cuando $X_0 = q$ entonces la probabilidad de extinción será π^q.

En adelante, para que los problemas de estimación tengan sentido, se supondrá que las observaciones del tamaño de la población en la generación n son $X_n > 0$, pues únicamente entonces el problema de estimación surgirá.

2.2. El caso fraccional lineal y su proceso de nacimiento asociado

Un caso particular importante es cuando $f(s)$ es la denominada función generatriz fraccional lineal (Harris, 1963, pág. 9). Ella viene definida por

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

para algún $0 < b, c$ con $b \leqslant 1 - c$. En este caso $p_k = b \cdot c^{k-1}$ $(k = 1,2,...)$, $p_0 = 1 - b/(1-c)$ y $\mu = b/(1-c)^2$. Además, en el caso supercrítico $\pi = \text{mín}\,\{1, (1-b-c)/c\,(1-c)\}$ y en el caso crítico $\sigma^2 = f''(1-) = 2\,c/(1-c)$ que es la varianza de la distribución de probabilidad $\{p_k\}$.

Si $X_0 = 1, X_1, X_2, \ldots$ es un proceso de Galton-Watson del tipo fraccional lineal, entonces la función generatriz $f_n(s)$ de X_n, cuando $\mu > 1$, es

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

en donde

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

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

por lo cual

$$pr\,(X_n = K \mid X_0 = 1) = b\,(n) \cdot c\,(n)^{K-1} \qquad (K = 1, 2, \ldots)$$

$$pr\,(X_n = 0 \mid X_0 = 1) = 1 - b\,(n)/\{1 - c\,(n)\}$$

y, por tanto,

$$pr\,(X_n > 0 \mid X_0 = 1) = \mu^n\,(1 - \pi)/(\mu^n - \pi)$$

$$pr\,(X_n = K \mid X_0 = 1, X_n > 0) = \{1 - c\,(n)\} \cdot c\,(n)^{K-1}$$

$$= \frac{1 - \pi}{\mu^n - \pi} \left\{ 1 - \frac{1 - \pi}{\mu^n - \pi} \right\}^{K-1}$$

$$(K = 1, 2, \ldots)$$

y como $(1 - \pi)/(\mu^n - \pi) < 1$, por ser $\mu > 1$, podemos convenir en llamar por $\exp \cdot (-\lambda\, t)$ a dicha cantidad, determinándose el producto $\lambda \cdot t$ de modo que tal igualdad se verifique. En particular, si convenimos en que $\lambda = 1$, entonces t está determinado de un modo único por

$$t = \log\,\{(\mu^n - \pi)/(1 - \pi)\} \tag{3}$$

y así con la nueva notación

$$pr\,(X_n = K \mid X_0 = 1, X_n > 0) = e^{-t}\,(1 - e^{-t})^{K-1},\; K = 1,2,\ldots \tag{4}$$

De modo semejante, cuando $\mu = 1$, entonces:

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

por lo cual

$$pr\,(X_n = K \mid X_0 = 1) = \{1 - c\,(n)\}^2 \cdot c\,(n)^{K-1} \qquad (K = 1, 2, \ldots)$$

$$pr\,(X_n = 0 \mid X_0 = 1) = c\,(n)$$

en donde

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

y, por tanto,

$$pr(X_n > 0 \mid X_0 = 1) = 2/(2 + n\sigma^2)$$

$$pr(X_n = K \mid X_0 = 1, X_n > 0) = \{1 - c(n)\} \cdot c(n)^{K-1}$$
$$= \frac{2}{2 + n\sigma^2}\left\{1 - \frac{2}{2 + n \cdot \sigma^2}\right\}^{K-1}$$
$$(K = 1, 2, \ldots)$$

y como $2/(2 + n\sigma^2) < 1$, podemos igualar dicha cantidad a $\exp(-\lambda t)$. En particular haciendo $\lambda = 1$, entonces t viene dado por la expresión.

$$t = \log\{(2 + n\sigma^2)/2\} \tag{5}$$

de modo único, volviéndose a verificar, por tanto, la expresión (4).

Sea ahora Y_η un proceso de nacimiento puro de intensidad $\lambda = 1$ tal que $Y_0 = 1$. Entonces $pr(Y_t = K \mid Y_0 = 1) = e^{-t} \cdot (1 - e^{-t})^{K-1}$, de modo que por la expresión (4):

$$pr(X_n = K \mid X_0 = 1, X_n > 0) = pr(Y_t = K \mid Y_0 = 1)$$

por lo cual se concluye en que, probabilísticamente, es lo mismo hablar de un proceso de Galton-Watson con una función generatriz del tipo fraccional lineal que, partiendo de un solo individuo, se encuentra en el estado $K \neq 0$ en la generación n, que hablar de un proceso de nacimiento puro de intensidad $\lambda = 1$ que, partiendo de un solo individuo, se encuentra en el estado $K \neq 0$ en un tiempo t relacionado con el n por las expresiones (3) o (5) según que $\mu > 1$ o $\mu = 1$. A un tal proceso Y_η le denominamos en adelante por "proceso de nacimiento puro asociado" al de Galton-Watson original X_ξ.

Es bien conocido (Harris, 1963, pág. 101) que un proceso de nacimiento puro, observado en tiempo equidistantes, se convierte en uno de de Galton-Watson del tipo fraccional lineal. Ahora se ha deducido que todo proceso de Galton-Watson del tipo fraccional lineal que comienza con un solo individuo, se convierte en uno de nacimiento puro de intensidad $\lambda = 1$, en el sentido anterior, con un cambio apropiado en la escala de tiempos, si se le condiciona en la no extinción. Ello no sería factible cuando $X_0 = q \neq 1$ pues el proceso de nacimiento puro siempre crece, mientras que el de Galton-Watson no.

2.3. El proceso paralelo a uno de Galton-Watson

Sea X_ξ un proceso de Galton-Watson con función generatriz $f(s)$ que, partiendo de $X_0 = q$, es observado en la generación n en la cual $X_n = K \neq 0$. Sea Y_η un nuevo proceso, también de Galton-Watson, que, con igual función generatriz que el anterior, parte de $Y_0 = 1$ individuos, pasa por el estado q en una generación s desconocida y observado en el tiempo $m = n + s$ da $Y_m = X_n = K$. Con ello, el proceso Y_η se comporta como el X_ξ, desde que alcanza el estado q, con la previa transformación de la escala de tiempos. En adelante, un tal proceso Y_η será llamado "proceso paralelo" al X_ξ.

2.4. Estimación

Sea X_ξ un proceso de Galton-Watson con función generatriz del tipo fraccional lineal de media $\mu > 1$, caso supercrítico, que comenzando con $X_0 = q$ es observado en la generación n en la cual $X_n = K \neq 0$. Sea Y_η el proceso paralelo al anterior, de modo que $Y_0 = 1$, $Y_s = q$, $Y_m = X_n = K$, $m = n + s$, y es también del tipo fraccional lineal. Finalmente, sea Z_ν el proceso de nacimiento puro asociado al proceso paralelo; en él $Z_0 = 1$, $Z_t = K$, $\lambda = 1$, en donde el tiempo t está dado por la expresión (3) que en las condiciones actuales será

$$t = \log \{(\mu^m - \pi)/(1 - \pi)\} = \log \{(\mu^n \cdot \mu^s - \pi)/(1 - \pi)\}.$$

El problema está en que el tiempo s es desconocido; ello puede solventarse haciendo notar que el proceso realmente comenzó con q individuos y no con uno sólo, de modo que s parece lógico sea calculado como el tiempo medio que tardaría el proceso Y_η en recorrer los estados 1 al q. Como $E(Y_s \mid Y_s > 0) = (\mu^s - \pi)/(1 - \pi)$ entonces se propone como valor de s al $\bar{s}$ tal que $E(Y_{\bar{s}} \mid Y_{\bar{s}} > 0) = q$, con lo cual $\bar{s} = \log\{q(1-\pi)+\pi\}/\log\mu$. Con ello:

$$t = \log \frac{\mu^n\{q(1-\pi)+\pi\}-\pi}{1-\pi} \tag{6}$$

de modo que el tiempo t será desconocido cuando alguna de las cantidades μ, π, q o n sean desconocidas. Si acaso alguna de ellas lo fuera, aún podría estimarse t por la expresión (1), de modo que como el proceso de nacimiento asociado al paralelo tiene de intensidad $\lambda = 1$ y comienza con $Z_0 = 1$, entonces $\hat{t} = \log Z_t = \log K$ estima al tiempo t. Finalmente, teniendo en cuenta la expresión (6) se deduce que

$$K \text{ estima a } \quad \frac{\mu^n\{q(1-\pi)+\pi\}-\pi}{1-\pi}$$

y despejando cada uno de los parámetros μ, π, n y q se obtienen los estimadores que se proponen:

$$\hat{n} = \log\left\{\frac{X_n(1-\pi)+\pi}{X_0(1-\pi)+\pi}\right\}/\log\mu \tag{7}$$

$$\hat{q} = \frac{X_n}{\mu^n} + \frac{\pi}{1-\pi}\,\frac{1-\mu^n}{\mu^n} \tag{8}$$

$$\hat{\pi} = \left(1 + \frac{\mu^n - 1}{X_n - X_0 \cdot \mu^n}\right)^{-1} \tag{9}$$

$$\hat{\mu} = \left(\frac{X_n(1-\pi)+\pi}{X_0(1-\pi)+\pi}\right)^{1/n} \tag{10}$$

Cuando se trate del caso crítico, en que $\mu = 1$, todo es igual salvo que ahora t viene dado por la expresión (5), de modo que, $t = \log\{(2+m\sigma^2)/2\} = \log\{(2+n\sigma^2+s\sigma^2)/2\}$ en donde s se determina como el valor $\bar{s}$ tal que $E\,(Y_{\bar{s}} \mid Y_{\bar{s}} > 0) = q$ y como $E\,(Y_s \mid Y_s >$ $> 0) = (2 + s\,\sigma^2)/2$ entonces $\bar{s} = 2\,(q - 1)/\sigma^2$. Con ello ahora:

$$t = \log\,\{(n\,\sigma^2 + 2\,q)/2\}$$

y como por otro lado una estimación de la edad t dada por la expresión es $\hat{t} = \log K$, entonces

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

de modo que despejando cada uno de los parámetros σ^2, n y q se obtienen los estimadores que se proponen:

$$\hat{n} = 2 \cdot \frac{X_n - X_0}{\sigma^2} \tag{11}$$

$$\hat{q} = \frac{2 \cdot X_n - n\,\sigma^2}{2} \tag{12}$$

$$(\hat{\sigma}^2) = 2 \cdot \frac{X_n - X_0}{n} \tag{13}$$

Nótese que todos los estimadores dados hasta ahora han sido obtenidos bajo la hipótesis de que X_ξ es del tipo fraccional lineal. En lo que sigue se darán propiedades de los estimadores $\hat{n}$ y $\hat{q}$ que valen en situaciones más generales.

Cuando $\mu < 1$, caso subcrítico, la estimación pierde sentido por alcanzarse una situación estacionaria (Stigler, 1970).

2.5. Estudio de los estimadores de la edad

Cuando en los estimadores (7) y (11) se hace $q = 1$, se obtiene

$$\begin{aligned} \hat{n} &= \log\,\{X_n\,(1 - \pi) + \pi\}/\log \mu \\ \hat{n} &= 2\,(X_n - 1)/\sigma^2 \end{aligned} \tag{14}$$

respectivamente, que son los mismos propuestos por Stigler (1970).

Como dicho autor hizo notar, el pedir la consistencia de $\hat{n}$, en el sentido de que $|\hat{n} - n| \to 0$ para grandes n, es excesivo dada la escasez de datos tomados. Sin embargo, la α-consistencia si será accesible. En efecto, como

$$\hat{n} - n = \log \left\{ \frac{X_n (1 - \pi) \mu^{-m} + \pi \cdot \mu^{-m}}{X_0 (1 - \pi) \mu^{-s} + \pi \cdot \mu^{-s}} \right\} / \log \mu$$

si hacemos que $n \to \infty$, para un $X_0 = q$ dado, entonces $m \to \infty$ quedando s fijo. Como por otro lado Stigum (1966) ha probado que bajo las condiciones de que $\mu > 1$ y $\Sigma\, k (\log k) p_k < \infty$, entonces, condicionalmente en la no extinción, $Y_m \cdot \mu^{-m} = X_n \cdot \mu^{-m} \to W$ casi seguro, cuando $m \to \infty$, en donde W es una variable aleatoria con una densidad contínua en $(0, \infty)$, se sigue que casi seguro $\hat{n} - n \to \log [W (1 - \pi)/ / \{q\, \mu^{-s} \cdot (1 - \pi) + \pi \cdot \mu^{-s}\}]/\log \mu$, y en particular $\hat{n}$ es un estimador α-consistente de n para cualquier $\alpha > 0$ y cualquiera que sea la función generatriz $f(s)$.

El cálculo de intervalos aproximados de confianza para n no es directamente accesible pues cuando $X_0 = q$, condicionando en la no extinción, $X_n/\mu^n \to W'$ casi seguro, cuando $n \to \infty$, en donde W' tiene una densidad que es suma de q densidades relacionadas con distribuciones gamma que tienen distintos parámetros, lo cual haría sumamente engorroso el cálculo de tales intervalos, especialmente para valores altos de q. Alternativamente, la expresión anterior de $\hat{n} - n$ puede valernos para dar un intervalo de confianza aproximado para n. Como dicha expresión contiene el tiempo desconocido s, si convenimos como antes en que el valor más apropiado para s es $\bar{s}$, entonces $\hat{n} - n \to \log \{W (1 - \pi)\}/\log \mu$ de modo que los mismos intervalos para W y n dados por Stigler (1970) cuando $X_0 = 1$ valen ahora, salvo que $\hat{n}$ está dado por la expresión (7) y no por la (14). Como además $\hat{n}$, dado por la (7), es menor que $\hat{n}$, dado por la (14), entonces el punto base del intervalo actual es inferior al del intervalo de Stigler como era lógico, pues es de esperar que el tamaño K se alcance antes partiendo de q individuos que cuando se parte de uno solo. Como quiera que el intervalo original hacia previsiones para la edad cuando se parte de un individuo y realmente se ha partido de q, es de suponer que dicho intervalo, dentro de su validez aproximada, sea conservador.

En el caso crítico, veamos que el estimador $\hat{n}$, dado por la expresión (11), es asintóticamente insesgado en el sentido de que $E(\hat{n}/n \mid X_n > 0) \rightarrow 1$ cuando $n \rightarrow \infty$, con independencia de cual sea la función generatriz $f(s)$. En efecto, para un proceso de Galton-Watson que comienza con un solo individuo y que tiene por función generatriz una $f(s)$ cualquiera, Kesten, Ney and Spitzer (1966) han probado que cuando $n \rightarrow \infty$, $n\{1 - f_n(0)\} \rightarrow 2/\sigma^2$, en donde $f_n(s)$ es la función generatriz del tamaño de la población en la generación n. Como

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

se deduce, por el desarrollo del binomio de Newton que se obtiene, que $E(\hat{n}/n \mid X_n > 0) \rightarrow 1$ cuando $n \rightarrow \infty$.

Por otro lado, los mismos autores anteriores han probado que la distribución límite de $(2X_n/n\sigma^2 \mid X_0 = 1, X_n > 0)$ es una exponencial de parámetro uno, de modo que, para el proceso paralelo, la variable $\{2Y_m/(m\sigma^2) \mid Y_m > 0\}$ también tiene por distribución límite dicha exponencial, no importanto la condición de pasar por q para valores grandes de m, como es el caso, pues $E(Y_m \mid Y_m > 0)$ tiende a infinito con m. Con ello, un intervalo de confianza aproximado al $(1-\alpha)\cdot$ $\cdot 100\,\%$ para $2Y_m/(m\sigma^2)$ es $[a_1, a_2]$ en donde $a_1 = -\log(1-\alpha_1)$, $a_2 = -\log \alpha_2$, $\alpha_1 + \alpha_2 = \alpha$ y α_1, α_2 se determinan como en la subsección 1.2. Finalmente, como $m = n + s$, un intervalo aproximado para n es $[-s + 2Y_m/(\sigma^2 a_2), -s + 2Y_m/(a_1\sigma^2)]$, el cual, estimando s por $\bar{s}$, se convierte en el siguiente

$$\left[\frac{2}{\sigma^2}\left\{\frac{Y_m}{a_2} - q + 1\right\}, \frac{2}{\sigma^2}\left\{\frac{Y_m}{a_1} - q + 1\right\}\right]$$

2.6. Estudio de los estimadores del tamaño inicial.

Para ver que los estimadores, dados por las expresiones (8) y (12), del número de individuos con que comienza un proceso de Galton-Watson son α-consistentes para cualquier $\alpha > 0{,}5$ y para cualquier función generatriz $f(s)$, denominemos por $u = pr(X_n > 0 \mid X_0 = q) = 1 - f_n^q(0)$. Como el problema de estimación se ha presentado, entonces $f_n(0) = pr(X_n = 0 \mid X_0 = 1) < 1$ y, por tanto, u y $u^2 \to 1$ cuando $q \to \infty$. Si ahora tenemos en cuenta que

$$E(X_n \mid X_n > 0) = E(X_n)/u$$

$$V(X_n \mid X_n > 0) = V(X_n)/u + \{(u-1)E^2(X_n)\}/u^2$$

y que en los procesos de ramificación de Galton-Watson $E(X_n) = q \cdot \mu^n$ y $V(X_n) = q \cdot h$ en donde h es una cantidad que no depende de q, entonces llamando por $y = (\hat{q} - q)/q^{\alpha}$ con $\alpha > 0{,}5$ y teniendo en cuenta las expresiones (8) y (12), se deduce que $E(y) \to 0$ cuando $q \to \infty$. Por otro lado, en el caso supercrítico:

$$V(y) = \frac{q \cdot h \cdot u - f_n^q(0)\, q^2 \cdot \mu^{2n}}{u^2 \cdot \mu^{2n} \cdot q^{2\alpha}} = \frac{A}{\mu^{2n}}$$

y en el caso crítico $V(y) = A$, de modo que como $q \cdot f_n^q(0) \to 0$ cuando $q \to \infty$, entonces $V(y) \to 0$ en ambos casos. Con todo ello $E(y^2) \to 0$ y, por tanto, $y = (\hat{q} - q)/q^{\alpha} \to 0$ en probabilidad cuando $q \to \infty$.

En particular, cuando $p_0 = 0$ entonces $\pi = 0$ y, por tanto

$$\hat{q} = X_n/\mu^n \tag{16}$$

de modo que $E(\hat{q}) = q$ y $\hat{q}$ será insesgado.

Cuando m sea grande, es posible dar intervalos de confianza aproximados para q. Nótese que como $m = n + s$, entonces m es grande cuando n lo es. Para dar dichos intervalos en el caso supercrítico, supóngase calculado, por los métodos indicados por Stigler (1970), un intervalo $[\hat{m} - b_1,\ \hat{m} + b_2]$ para la edad m, en donde $\hat{m} = \log\{X_n(1-\pi) + \pi\}/\log\mu$; por tanto, un intervalo para s será $[s_1 = \hat{m} - n - b_1,\ s_2 = \hat{m} - n + b_2]$ y, por el sentido que se ha venido dando a s, un intervalo para q tendrá como límite inferior

$$q_1 = E\,(Y_{s_1} \mid Y_{s_1} > 0) = E\,(Y_{s_1})/r_1 = \{X_n\,(1-\pi) + \pi\}/(r_1 \cdot \mu^{n+b_1})$$

en donde $r_1 = pr\,(Y_{s_1} > 0)$; de igual modo el límite superior será

$$q_2 = \{X_n\,(1-\pi) + \pi\}/(r_2\,\mu^{n\text{-}b_2})$$

en donde $r_2 = pr\,(Y_{s_2} > 0)$.

En el caso fraccional lineal los anteriores límites son

$$\frac{X_n}{\mu^{n+b}} + \frac{\pi}{1-\pi} \cdot \frac{1-\mu^{n+b}}{\mu^{n+b}}$$

haciendo $b = b_1$ para obtener q_1 y $b = - b_2$ para obtener q_2.

En el caso crítico el procedimiento es similar. Así, si $[\hat{m}/a_2, \hat{m}/a_1]$ es el intervalo para m dado por Stigler (1970) con $\hat{m} = 2\,(X_n - 1)/\sigma^2$, entonces un intervalo para s será $[s_1 = (\hat{m} - na_2)/a_2,\ s_2 = (\hat{m} - na_1)/a_1]$ por lo cual uno para q es

$$q_1 = E\,(Y_{s_1} \mid Y_{s_1} > 0) = 1/r_1$$

$$q_2 = E\,(Y_{s_2} \mid Y_{s_2} > 0) = 1/r_2$$

En particular, para el caso fraccional lineal, los límites serán:

$$q_1 = 1 + (X_n - 1)/a_2 - (n\,\sigma^2)/2$$

$$q_2 = 1 + (X_n - 1)/a_1 - (n\,\sigma^2)/2$$

2.7. Consideraciones finales

Los estimadores $\hat{\pi}$ y $\hat{\mu}$ dados por las expresiones (9) y (10) no tienen sentido tal como están. En efecto, como $\hat{\pi}$ depende de μ y para conocer μ hace falta conocer $f\,(s)$, la función generatriz, entonces $\hat{\pi}$ no tiene validez pues de conocer $f\,(s)$ se conoce π, al ser π la raiz positiva más pequeña de la ecuación $f\,(s) = s$. De igual modo, el estimador $\hat{\mu}$ no tendrá validez al depender de π. Al estimador (13) de σ^2 le ocurre otro tanto, pues si se conoce que estamos en el caso crítico, es que se sabe que $\mu = 1$ y, por tanto, se conoce $f(s)$ y de rechazo $\sigma^2 = f''(1-)$ también es conocida, no teniendo sentido ya el estimarla.

Sin embargo, los estimadores (9) y (10) pueden tener valor como estimadores no paramétricos. En un trabajo posterior se estudiarán los estimadores no paramétricos de π que se obtienen de la expresión (9), su influencia en la estimación no paramétrica de la edad n y del tamaño inicial q del proceso y su comparación con los distintos estimadores no paramétricos dados por Crump and Howe (1972).

Obsérvese que los estimadores $\hat{n}$ y $\hat{q}$, dados más arriba, han sido obtenidos a partir del estimador $\hat{t}$ de la edad de un proceso de nacimiento. Iguales estimadores se hubieran podido obtener a partir de los estimadores (14) y (15) de la edad n propuestos por Stigler por un procedimiento similar al indicado en la subsección 1.3. De hacerse así, los estimadores (1) de t y (2) de q en un proceso de nacimiento, podrían obtenerse a partir de los anteriores, considerando el proceso de Galton-Watson que se obtiene al hacer observaciones equidistantes en el proceso de nacimiento puro. Así, el estimador (16) daría el estimador (2).

Este trabajo es parte de la tesis, no publicada, del mismo autor. Mi agradecimiento al Dr. D. Alfonso Guiraum Martín y al Dr. D. Ramón Gutiérrez Jaimez, catedrático y agregado respectivamente, del Departamento de Estadística Matemática de la Universidad de Granada, que dirigieron dicha tesis. Este trabajo ha sido financiado por los Departamentos de Fisiología y Bioquímica de la Facultad de Medicina de la misma Universidad y el anteriormente citado.

REFERENCIAS

CRUMP, K. S. and HOWE, R. B. (1972): Nonparametric estimation of the age of a Galton-Watson branching process. Biometrika 59, 533-8.

FELDMAN, D. and FOX, M. (1968): Estimation of the parameter *n* in the binomial distribution. J. Am. Statist. Assoc., 150-8.

HARRIS, T. E. (1963): The Theory of Branching Processes. Berlin: Springer-Verlag.

MORAN, P. A. P. (1951): Estimation methods for evolutive processes. J. R. Statis. Soc. B 13, 141-6.

KEIDING, N. (1974): Estimation in the Birth process. Biometrika 61, 71-80.

KESTEN, H., NEY, P. and SPITZER, F. (1966): The Galton-Watson process with mean one and finite variance. Theory Prob. Appl. 11, 513-40.

STIGLER, S. M. (1970): Estimating the age of a Galton-Watson branching process. Biometrika 57, 505-12.

STIGUM, B. P. (1966): A theorem on the Galton-Watson process. Ann. Mathe. Statist. 37, 695-8.