TRABAJOS DE ESTADISTICA Y
DE INVESTIGACION OPERATIVA
Vol. 32, Núm. 2, 1981, pp. 55 a 67

ESTIMACION NO PARAMETRICA DE LA EDAD Y DE LA PROBABILIDAD DE EXTINCION EN PROCESOS DE GALTON-WATSON

A. Martín Andrés
Departamento de Fisiologia
Sección de Bioestadística
Facultad de Medicina
Universidad de Granada

SUMARIO

Se proponen estimadores no paramétricos de la edad y de la probabilidad de extinción de un proceso de ramificación de Galton-Watson. Dichos estimadores son comparados por simulación de Monte-Carlo, con otros estimadores propuestos por Stigler (1970) y Grump and Howe (1972).

SUMMARY

Nonparametric estimators of the age and the probability of extinction of a Galton-Watson branching process are given. A Monte-Carlo simulation shows that these estimators compare favourably with a estimators studied earlier by Stigler (1970) and Crump and Howe (1972).

CONCEPTOS NECESARIOS

Procesos de Galton-Watson; Edad de la población; Estimación en procesos de Galton-Watson; Comparación de los estimadores por el método de Monte-Carlo.

Conceptos necesarios: Procesos de Galton-Watson; Edad de la población; Estimación en procesos de Galton-Watson; Comparación de los estimadores por el método de Monte-Carlo.

1. Introducción

Sean $X_0 = q$, X_1, X_2, ..., los tamaños de las generaciones sucesivas de un proceso de ramificación de Galton-Watson (Harris, 1963, cap. I). Esto es, la población comienza con q individuos, que forman la generación cero, cada uno de los cuales al final de su vida da lugar a un número aleatorio de descendientes que forman la primera generación. Cada uno de ellos a su vez tiene un número aleatorio de descendientes, con independencia de cuantos tuvieron sus progenitores o sus compañeros, que pasan a engrosar la segunda generación y así sucesivamente. Si notamos por $\{p_k\}$ a la distribución de probabilidad del número de descendientes de cada individuo, entonces $f(s) = \Sigma\, p_k \cdot s^k$ será la función generatriz asociada y $\mu = f'(1)$ será el número medio de descendientes por individuo.

Un parámetro de interés es la probabilidad de extinción π del proceso, la cual es (Harris, 1963; pág. 7) la raiz positiva más pequeña de la ecuación $f(s) = s$ y representa la probabilidad de que la población eventualmente se extinga, cuando ésta comienza por un solo individuo. Naturalmente, por la independencia de actuación de los individuos, cuando $X_0 = q$ entonces la probabilidad de extinción será π^q.

Es reciente el interés por la estimación de los diversos parámetros de dichos procesos, a saber: estimación de la media μ (Harris, 1963; página 32), estimación de la probabilidad de extinción π, paramétrica (Stigler, 1971) y no paramétrica (Stigler, 1971 y Crump and Howe, 1972);

estimación de la edad n del proceso, paramétrica (Stigler, 1970) y no paramétrica (Crump and Howe, 1972). Los dos últimos estimadores fueron calculados para el caso de que $X_0 = q = 1$; posteriormente, Martín-Andrés (1981) ha dado estimadores paramétricos de la edad n cuando $X_0 = q$ y del tamaño inicial q de la población.

En este trabajo se pretenden obtener estimadores no paramétricos de π y n y compararlos por simulación de Monte-Carlo con los ya conocidos.

2. Estimadores conocidos

Sean $X_m, X_{m+1}, \ldots, X_n$ los tamaños de las generaciones sucesivas m, $m + 1, \ldots, n$ observadas, en donde se supone que $X_n > 0$ para que el problema de la estimación de n se presente. También se supondrá en adelante que $\mu > 1$, pues de otro modo la población se extingue con probabilidad uno y el problema no surge casi seguro.

Cuando $m = n$, o sea cuando se observa solo una generación, Stigler (1970) propuso un estimador paramétrico de n, para el caso de que $X_0 = 1$, dado por

$$\hat{n}_0 = \log \{(1 - \pi) X_n + \pi\} / \log \mu \tag{1}$$

que era de máxima verosimilitud, cuando la función generatriz del número de descendientes $f(s)$ es del tipo fraccional lineal, es decir, cuando $f(s) = 1 - b/(1 - c) + bs/(1 - cs)$ para algún $b, c > 0$ con $b \leqslant 1 - c$, en cuyo caso $\mu = b/(1 - c)^2$ y $\pi = (1 - b - c)/c\,(1 - c)$. Tal estimador no mejora por el hecho de conocer el tamaño de generaciones adyacentes, por el carácter de cadena de Markov que tiene el proceso X_ξ (Crump and Howe, 1972).

Los mismos autores convierten el estimador anterior en no paramétrico mediante el uso de estimadores, independientes de la forma de $f(s)$, de π y μ. Así Harris (1963, pág. 32) propone como estimador no paramétrico de μ el siguiente:

$$\hat{\mu} = \sum_{j=m+1}^{n} X_j \Big/ \sum_{j=m}^{n-1} X_j \qquad (2)$$

que es de máxima verosimilitud, asintóticamente insesgado y converge a μ con probabilidad uno bajo ciertas condiciones (Crump and Howe, 1972). Por su parte los mismos autores proponen como estimador no paramétrico de π el siguiente:

$$\hat{\pi}_1 = \exp. \{-2(\hat{\mu} - 1)/\hat{\sigma}^2\} \qquad (3)$$

con

$$\hat{\sigma}^2 = \frac{1}{n - m - 1} \sum_{k=m+1}^{n} X_{k-1} \left(\frac{X_k}{X_{k-1}} - \hat{\mu} \right)^2$$

de modo que para conocer $\hat{\pi}$, hacen falta al menos tres generaciones sucesivas.

Con los dos estimadores anteriores puede construirse un estimador no paramétrico para n así:

$$\hat{n}_1 = \log \{(1 - \hat{\pi}_1) X_n + \hat{\pi}_1\} / \log \hat{\mu} \qquad (4)$$

En donde la oportunidad de tales sustituciones de μ por $\hat{\mu}$ y de π por cualquier estadístico no negativo $\hat{\pi} < 1$ con probabilidad uno (cuando $X_n \neq 0$) la prueben dichos autores viendo que el estimador $\hat{n}$ no paramétrico así obtenido es α-consistente para todo $\alpha > 0$, es decir, que $(\hat{n} - n)/n^a \to 0$ con probabilidad uno cuando $n \to \infty$ para todo $\alpha > 0$.

3. Estimadores propuestos

En el caso general en que $X_0 = q$, Martín-Andrés (1981) propone como estimador de π el siguiente:

$$\hat{\pi} = \left(1 + \frac{\hat{\mu}^n - 1}{X_n - q \cdot \hat{\mu}^n} \right)^{-1} \qquad (5)$$

que se convierte en el

$$\hat{\pi}_2 = (X_n - \hat{\mu}^n) \,/\, (X_n - 1) \tag{6}$$

cuando $q = 1$. Tal estimador puede justificarse de otro modo. Supongamos se observa la variable aleatoria $w = (X_n/\mu^n \mid X_n > 0)$ para un valor de n grande. Entonces, para el caso fraccional lineal, la verosimilitud sería (Stigler, 1970) $L = (1 - \pi) \exp. \{ -(1 - \pi) w \}$ de modo que cuando $w \geqslant 1$ el estimador de máxima verosimilitud de π para cuando $n \to \infty$ es $\hat{\pi} = (X_n - \hat{\mu}^n) \,/\, X_n$ que guarda un gran parecido con el $\hat{\pi}_2$. Cuando $w < 1$, la máxima verosimilitud para valores de π entre 0 y 1 es $\pi = 0$.

Nótese que los estimadores (5) y (6) requieren el conocimiento de la edad n a diferencia del estimador $\hat{\pi}_1$. Ello puede obviarse tomando X_m como tamaño de partida del proceso, con lo cual obtendremos:

$$\hat{\pi}_3 = \left(1 + \frac{\hat{\mu}^r - 1}{X_n - X_m \cdot \hat{\mu}^r}\right)^{-1} \tag{7}$$

en donde $r = m - n$, estimador que es equiparable, en cuanto a los datos que se requieren para su cálculo, con el $\hat{\pi}_1$. De nuevo $\hat{\pi}_3$ requiere, como mínimo, el conocimiento de tres generaciones sucesivas del proceso.

En todos los casos, cuando alguna estimación de π dé valores menores que cero o mayores que uno, se tomarán $\hat{\pi} = 0$ ó $\hat{\pi} = 1$, respectivamente.

Nótese que cuando $\hat{\mu} > 1$, entonces $\hat{\pi}_3$ estará entre 0 y 1, si $X_n > X_m \cdot \hat{\mu}^r$, de modo que cuando esto no ocurra con el último tamaño X_n, convendrá tomar otro intermedio X_h tal que $X_h > X_m \cdot \hat{\mu}^{h-m}$. Cuando no haya ninguno, tomar entonces el último tamaño X_n; cuando haya varios, tomar el de mayor subíndice.

El estimador $\hat{\pi}_3$ puede usarse alternativamente al $\hat{\pi}_1$ para así obtener el $\hat{n}_3$:

$$\hat{n}_3 = \log \{ (1 - \hat{\pi}_3) X_n + \hat{\pi}_3 \} \,/\, \log \hat{\mu} \tag{8}$$

La próxima sección se dedica a un estudio comparativo, por simulación de Monte-Carlo, de los estimadores $\hat{\pi}_1$, $\hat{\pi}_2$ y $\hat{\pi}_3$ por un lado, y los $\hat{n}_0$, $\hat{n}_1$ y $\hat{n}_3$ por otro.

4. Comparación de los estimadores por simulación

Para la simulación de Monte-Carlo se ha tomado como función generatriz el caso fraccional lineal y, ello por dos razones:

(i) Porque depende de dos parámetros, de modo que son posibles todas las combinaciones que se deseen de π y μ.

(ii) Porque $\hat{n}_0$ es de máxima verosimilitud en dicho caso, de modo que tal estimador actuará bastante bien, cuando se use tal distribución y, consecuentemente, la comparación de su eficacia con la de los otros dos estimadores $\hat{n}_1$ y $\hat{n}_3$ nos indicará cuando actúan de un modo razonablemente bueno.

Para cada pareja de valores de π y μ de las indicadas en las tablas I y II, se han obtenido, por simulación, los tamaños poblacionales X_m, X_{m+1}, ..., X_n (para los valores de m y n indicados, asimismo, en dichas tablas) y a su partir se obtuvieron las estimaciones $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$, $\hat{n}_0$, $\hat{n}_1$ y $\hat{n}_3$. Lo anterior se hizo en mil ocasiones para cada cuaterna π, μ, m y n. El criterio seguido para la obtención de tales estimadores, se sumariza a continuación.

En primer lugar, deben encontrarse dos valores de s y h con $s > h$, de entre los 0, 1, ..., r tales que el estimador de la media $\hat{\mu} = (Y_{h+1} + \dots + Y_s) / (Y_h + \dots + Y_{s-1})$ sea mayor que uno, en donde a los tamaños poblacionales les hemos llamado ahora por Y_0, Y_1, ..., Y_r en vez de X_m, ..., X_n. Ello debe hacerse con la condición de que $(s - h)$ sea máximo y, si hubiese varios, se escogerá el de mayor valor de s. Cuando tal pareja no exista, se hará $\hat{\pi}_1 = \hat{\pi}_2 = \hat{\pi}_3 = 1$ y $\hat{n}_1 = \hat{n}_3 = r + 1$, calculándose $\hat{n}_0$ como se indica más adelante. Cuando por el contrario si exista calcular entonces $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$, $\hat{n}_0$, $\hat{n}_1$ y $\hat{n}_3$ como más adelante a partir, cuando ha lugar, del estimador $\hat{\mu}$ anterior. La única salvedad es cuando $s - h = 1$, en cuyo caso debe de tomarse la precaución de que la pareja (s, h) que se utilice para el cálculo de $\hat{\pi}_3$ no debe ser la misma que sirvió para calcular $\hat{\mu}$. Si es así, buscar otra pareja para $\hat{\pi}_3$.

El cálculo de los estimadores de π y n a que nos hemos venido refiriendo, se realiza del modo siguiente:

a) Cálculo de $\hat{\pi}_1$:

$$\hat{\pi}_1 = \exp. \{-2(\hat{\mu}-1)/\hat{\sigma}^2\} \text{ con } \hat{\sigma}^2 = \frac{1}{r-1} \sum_{h=1}^{r} Y_{h-1} \left(\frac{Y_h}{Y_{h-1}} - \hat{\mu} \right)^2$$

b) Cálculo de $\hat{\pi}_2$:

$\hat{\pi}_2 = (Y_h - \hat{\mu}^{(h+m)}) / (Y_h - 1)$ 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 tal valor de h, hacer $\hat{\pi}_2 = 0$.

c) Cálculo de $\hat{\pi}_3$:

$$\hat{\pi}_3 = \left(1 + \frac{\hat{\mu}^{(s-h)} - 1}{Y_s - Y_h \cdot \hat{\mu}^{(s-h)}} \right)^{-1}$$

en donde s y h son dos naturales de entre los 0, 1, ..., r tales que $s > h$ e $Y_s > Y_h \cdot \hat{\mu}^{(s-h)}$, elegidos de modo que $(s-h)$ sea máximo. Si hay varias de tales parejas, tomar la de mayor valor de s. Cuando no exista ninguna, hacer $s = r$, $h = 0$ y calcular $\hat{\pi}_3$. Naturalmente, con el criterio de siempre de que se encuentre entre 0 y 1.

d) Cálculo de los estimadores de la edad:

Se calculan mediante las expresiones (1), (4) y (8), sustituyendo Y_h por el X_n que en ellas aparece y sumándoles a todas ellas la cantidad $(r - h)$, en donde en todos los casos h es el mayor natural de entre los 0, 1, ..., 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 $\hat{\pi}_1$, $\hat{\pi}_2$, $\hat{\pi}_3$, $\hat{n}_0$, $\hat{n}_1$ y $\hat{n}_3$ en los mil casos de cada apartado, se calcula para cada estimador la media, error de la media y raiz cuadrada del error cuadrático medio. Los resultados finales de todo el proceso, se indican en las tablas I y II. La simulación y cálculos han sido realizados en un ordenador UNIVAC 1108, a través de una terminal DCT 2000.

5. Discusión de los resultados

Observando la tabla I, se llega a las siguientes conclusiones para los estimadores de π:

a) Los estimadores $\hat{\pi}_1$ y $\hat{\pi}_2$ sobreestiman los valores pequeños de π y subestiman los valores grandes. El estimador $\hat{\pi}_3$ siempre sobreestima a π, tanto más cuanto menor sea μ.

b) Los estimadores $\hat{\pi}_1$ y $\hat{\pi}_3$ actúan mejor para grandes valores de π, mientras que el $\hat{\pi}_2$ 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 varía.
- $\hat{\pi}_2$ mejora para valores pequeños de μ.
- $\hat{\pi}_3$ 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.
- $\hat{\pi}_2$ prácticamente no varía.
- $\hat{\pi}_3$ empeora para valores pequeños de π.

e) Comparando los estimadores por parejas :
- $\hat{\pi}_2$ es mejor que $\hat{\pi}_1$ para valores grandes de μ, salvo cuando el número de generaciones observadas y π sean grandes.
- $\hat{\pi}_3$ es mejor que $\hat{\pi}_1$ para valores grandes de π.
- $\hat{\pi}_2$ es mejor que $\hat{\pi}_3$ para valores pequeños de π.

f) Comparando los estimadores en conjunto :
- Cuando π sea grande, es preferible el estimador $\hat{\pi}_3$.
- Cuando π y μ sean pequeños, es preferible $\hat{\pi}_1$.
- Cuando π sea pequeño y μ grande, es preferible $\hat{\pi}_2$.

De igual modo, por observación 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, y cada vez menos conforme aumenta el valor de π. El estimador $\hat{n}_1$ subestima a n cuando π es pequeño, mientras que el $\hat{n}_3$ lo hace cuando π es también pequeño y, tanto más cuanto aumenta μ.

b) Los estimadores $\hat{n}_1$ y $\hat{n}_3$ actuan mejor cuando π es grande, que cuando es pequeño, mientras que $\hat{n}_0$ no varía demasiado. Por otro lado, los tres estimadores de n 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 débilmente cuando π es grande y mejora cuando es pequeña.
   - $\hat{n}_1$ mejora sensiblemente cuando μ es pequeña y tiene poca variación, cuando es grande.
   - $\hat{n}_3$ empeora débilmente cuando μ es grande y suele mejorar cuando es pequeño.

d) Si $(n - m)$ se deja fijo y se aumenta m:

   - $\hat{n}_3$ empeora sensiblemente.
   - $\hat{n}_0$ y $\hat{n}_1$ suele empeorar especialmente cuando μ es pequeño.

e) Claramente, el estimador $\hat{n}_0$ es siempre mejor que los $\hat{n}_1$ y $\hat{n}_3$, como era de esperar al ser un estimador paramétrico.

f) Para el caso en que π sea desconocida y haya que usar los estimadores no paramétricos de la edad, que será lo más usual, el estimador $\hat{n}_1$ actúa siempre mejor que el $\hat{n}_3$ cuando μ es grande, y son desiguales cuando μ es cercano a uno.

El que $\hat{\pi}_3$ fuera mejor que el $\hat{\pi}_1$ cuando π era grande y, sin embargo, no ocurra así con los $\hat{n}_3$ y $\hat{n}_1$, es explicable porque la sobreestimación de π perjudica notablemente a $\hat{n}$ y el $\hat{\pi}_3$ siempre sobreestima a π.

Resumiendo, el estimador más aconsejable, salvo que se posea una información específica, de la probabilidad de extinción π es el $\hat{\pi}_3$; mientras que para estimar la edad el más apropiado es el paramétrico o en su defecto el $\hat{n}_1$.

Este trabajo es parte de la tesis, no publicada, del mismo autor. Mi agradecimiento al Dr. Don Alfonso Guiraum Martín y el Dr. Don Ramón Gutiérrez Jaimez, catedrático y agregado respectivamente, del Departamento de Estadística Matemática de la Universidad de Granada, que dirigieron dicha tesis. Asimismo, agradezco la ayuda prestada al Centro de Cálculo y al profesor Don Antonio Martínez Molina de la misma Universidad.

BIBLIOGRAFIA

CRUMP, K. S. and HOWE, R. B. (1972). "Nonparametric estimation of the age of a Galton-Watson branching process" Biometrika 59, 533-38.

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

MARTIN-ANDRES (1981). "Estimación de la edad y del número inicial de individuos en procesos de nacimiento puro y de Galton-Watson". Trabajos de Estadística y de Investigación operativa, vol. 32, Cuad. 1.

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

STIGLER, S. M. (1971). "The estimation of the probability of extinction and other parameters associated with branching processes". Biometrika 58, 499-508.

TABLA I

Valores obtenidos para los estimadores de π

		$\mu = 1.3$						$\mu = 2$					
		$\pi = 0.2$			$\pi = 0.75$			$\pi = 0.2$			$\pi = 0.75$		
m,n		Media	E. E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.
6,8	$\hat{\pi}_1$	.42	.0389	.45	.61	.0363	.39	.29	.0308	.32	.64	.0348	.36
	$\hat{\pi}_2$	.43	.0422	.48	.52	.0393	.45	.24	.0282	.28	.62	.0313	.34
	$\hat{\pi}_3$	.62	.0324	.53	.76	.0243	.24	.69	.0247	.55	.89	.0138	.20
6,10	$\hat{\pi}_1$	.39	.0298	.35	.73	.0236	.24	.40	.0263	.33	.76	.0224	.22
	$\hat{\pi}_2$	.32	.0344	.36	.62	.0340	.36	.23	.0274	.28	.57	.0326	.37
	$\hat{\pi}_3$	.45	.0269	.37	.77	.0215	.22	.75	.0224	.59	.89	.0168	.22
10,12	$\hat{\pi}_1$	.33	.0355	.38	.57	.0364	.41	.32	.0308	.33	.65	.0346	.36
	$\hat{\pi}_2$	.43	.0382	.44	.50	.0402	.47	.22	.0269	.27	.60	.0312	.35
	$\hat{\pi}_3$	.62	.0302	.52	.77	.0228	.23	.88	.0152	.70	.97	.0056	.23
10,14	$\hat{\pi}_1$	.35	.0271	.31	.70	.0225	.23	.43	.0242	.33	.78	.0183	.19
	$\hat{\pi}_2$	.33	.0339	.36	.62	.0360	.38	.21	.0267	.27	.50	.0358	.44
	$\hat{\pi}_3$	.57	.0277	.46	.77	.0231	.23	.88	.0153	.70	.95	.0100	.22

E.E. = Error Estandar

E.C.M. = Error Cuadrático Medio

TABLA II

Valores obtenidos para los estimadores de n

		$\mu = 1.3$						$\mu = 2$					
		$\pi = 0.2$			$\pi = 0.75$			$\pi = 0.2$			$\pi = 0.75$		
m,n		Media	E.E.	Raiz de E.C.M.	Media	E.E.	Raiz de E.C.M.	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	7.11	.1678	1.90	7.23	.1753	1.91
	$\hat{n}_1$	6.92	.5058	5.17	8.76	1.1190	11.22	7.12	.1798	2.00	7.33	.2329	2.42
	$\hat{n}_3$	5.84	.4593	5.08	7.03	.5639	5.72	5.63	.1447	2.78	5.54	.1779	3.04
6,10	$\hat{n}_0$	8.27	.2667	3.18	8.99	.2650	2.84	9.35	.1638	1.76	9.12	.1867	2.06
	$\hat{n}_1$	8.70	.3785	4.00	9.96	.5066	5.07	8.81	.1711	2.08	8.83	.2277	2.56
	$\hat{n}_3$	8.21	.3608	4.03	10.14	.8235	8.24	7.28	.1484	3.10	7.37	.2061	3.34
10,12	$\hat{n}_0$	10.03	.3891	4.36	10.42	.3754	4.07	11.17	.1877	2.05	11.60	.1601	1.65
	$\hat{n}_1$	11.45	.7743	7.76	13.55	1.3965	14.05	10.76	.2096	2.43	11.28	.2271	2.38
	$\hat{n}_3$	8.60	.5934	6.84	11.61	2.0972	20.98	7.66	.1719	4.67	7.50	.1723	4.82
10,14	$\hat{n}_0$	12.17	.4022	4.42	11.83	.4038	4.58	13.42	.1616	1.72	13.01	.1939	2.18
	$\hat{n}_1$	12.89	.4913	5.04	14.96	.7903	7.96	12.81	.1678	2.06	12.43	.2144	2.66
	$\hat{n}_3$	11.04	.4794	5.63	13.24	.8958	8.99	9.63	.1925	4.77	9.25	.1920	5.12

E.E. = Error Estandar

E.C.M. = Error Cuadrático Medio