MAT²
MATerials MATemàtics
Volum 2010, treball no. 4, 29 pp. ISSN: 1887-1097
Publicació electrònica de divulgació del Departament de Matemàtiques
de la Universitat Autònoma de Barcelona
www.mat.uab.cat/matmat

Desapareixerà el teu cognom?

Maria Jolis

1 Introducció

Cognoms més freqüents

La pregunta que ens fa el títol té avui dia una resposta clara: SI. Les xifres de natalitat al nostre país i als països del nostre entorn, cas de mantenir-se, porten a una extinció segura de tots els cognoms. En efecte, una dona a Catalunya té una mitjana de 1,43 fills.[1] Amb aquesta dada no ens és possible saber quin és le nombre mitjà de fills que té un home que passi el seu cognom als seus fills, però no serà massa diferent. Si tenim en compte que, d'aquests fills, els que passen (habitualment) el cognom són els fills de sexe masculí, dels quals n'hi haurà una mitjana d'aproximadament la meitat que la del nombre total de fills, obtindrem que el nombre mitjà de fills homes que passaran a les generacions posteriors un cert cognom serà de l'ordre de 0,7. Com veurem, per tal que no hi hagi extinció segura d'un cognom, el nombre mitjà de fills homes hauria de ser estrictament més gran que 1. I fins i tot en aquest cas, la probabilitat d'extinció és estrictament positiva.

De fet, en països com la Xina o Corea, en què el sistema de cognoms heretats per via masculina porta funcionant molts segles, es produeix un fenomen curiós. Per exemple, en una població tan gran com la xinesa, en

[1]Aquestes són dades de 2005. La mitjana d'Espanya del mateix any és 1,34.

la que s'esperaria una grandíssima varietat de cognoms, els tres cognoms més freqüents (Wang, Li i Zhang) són compartits per més del 21% de la població (més de 270 milions de persones!) i els 100 cognoms més freqüents ja cobreixen gairebé el 85% de la població. El cas de Corea és encara més espectacular, els tres cognoms més típics (Kim, Lee i Park) són compartits per gairebé la meitat de la població i actualment només hi ha uns 250 cognoms. Això s'ha produït pel fet que hem comentat abans, la probabilitat d'extinció d'un cognom és estrictament positiva i molts d'ells acaben desapareixent.

Lotka

Kolmogorov

Hem de recalcar que el model que permet fer aquestes afirmacions és una simplificació de la realitat i fa una suposició crítica que consisteix en què la distribució del nombre de fills es manté constant al llarg de les generacions. Aquesta suposició està ben lluny de satisfer-se en aquests moments, almenys a Catalunya i a Espanya: en aquestes darreres generacions ha anat descendint el nombre de fills i només l'últim any del que disposem de dades (2005) havia augmentat lleugerament la mitjana.

Per estudiar el model esmentat a dalt, necessitarem estudiar primer l'anomenada *funció generatriu de probabilitats* que és l'eina probabilística principal per a tractar el problema. Tractarem també alguns exemples de models "ficticis" i un de real basant-nos en dades del cens de 1920 als Estats Units. Aquests dades van ser usades per A. J. Lotka en un famós article de 1931 (vegi's [6]). Refarem els seus càlculs aprofitant la major potència computacional que tenim avui dia i estudiarem un model teòric aproximat (que també va descobrir Lotka [7]) que s'ajusta força bé a les dades reals.

2 Els processos de ramificació

Un model matemàtic relativament senzill de l'evolució de la mida de certes poblacions en les successives generacions és l'anomenat *procés de Galton-Watson*. Es tracta del tipus més simple de *procés de ramificació*. Aquests processos estocàstics són usats freqüentment en Biologia i Física Nuclear i permeten descriure l'evolució d'una població composta d'uns certs membres, anomenats individus o partícules. Es suposa que aquests membres es reprodueixen o no, independentment uns dels altres i del que han fet les generacions anteriors.

El nom de "processos de ramificació" segurament apareix per primer cop en un article d'A. N. Kolmogorov i N. A. Dmitriev [5] i un altre de A. M. Yaglom [8], ambdós de 1947, tot i que, de fet, en un altre important treball de 1938 sobre processos de ramificació en Genètica, Kolmogorov ja va donar expressions asimptòtiques i aproximacions de la probabilitat d'extinció de les poblacions (veure [4]).

De tota manera, els primers indicis de l'ús d'aquesta classe de processos es troben en el segle XVIII en el famós llibre de tres volums de Thomas Malthus titulat *Essay on the principles of Populations*, en el qual es planteja que una població no controlada hauria de créixer exponencialment. Malthus explica també que a la ciutat de Berna, de les 487 famílies burgeses existents, 379 s'havien extingit en el lapse de dos segles (1583-1783). El primer en tractar d'explicar el fenomen va ser el matemàtic francès I. J. Bienaymé (1796-1878), el qual va relacionar correctament la probabilitat d'extinció amb la mitjana de fills homes de cada home. A l'article [3] de les referències es pot trobar, com un apèndix, l'únic treball (del que s'ha trobat la publicació) de Bienaymé sobre el tema. En ell, Bienaymé no dóna cap manera de trobar les probabilitats d'extinció, però el fet que relacioni l'extinció segura o, contràriament, extinció amb probabilitat estrictament menor que 1, amb la mitjana del número de fills fa pensar que efectivament havia trobat la solució del problema.


Malthus

Independentment de Bienaymé, el famós naturalista britànic F. Galton i el seu compatriota, el matemàtic H. W. Watson, van abordar el mateix problema motivats per la preocupació de la societat victoriana sobre la possibilitat que els cognoms aristocràtics s'estiguessin extingint. En principi, Galton va plantejar la qüestió respecte la probabilitat d'aquesta extinció a la revista Educational Times l'any 1873 i Watson va donar una solució. L'any següent van escriure plegats un article (veure [1]). El model que van introduir Galton i Watson es va aplicar primer a l'extinció de cognoms però aviat es va començar a utilitzar en el camp de la biologia per a modelitzar el creixement de les poblacions d'éssers vius. Per a una història detallada de l'origen i el desenvolupament dels processos de ramificació, vegeu [2] i [3].

3 La funció generatriu de probabilitats

Per tal de tractar el problema de l'extinció de poblacions usarem la *funció generatriu de probabilitats*. Tot i que aquesta funció és una eina molt útil en l'estudi de les variables aleatòries que prenen valors en el conjunt dels nombres naturals, no sempre s'explica en els cursos de probabilitats. Això és degut a la falta de temps i al fet que hi ha tècniques similars (les funcions generatrius de moments o les funcions característiques) que s'apliquen a variables aleatòries amb valors reals arbitraris i permeten resoldre els mateixos tipus de problemes abordats en el primer curs de probabilitats. Per altra banda, les funcions generatrius de probabilitats tenen un caràcter més elemental i són l'eina adequada per al tractament del problema que ens interessa.

Donada una variable aleatòria X que pren valors en $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, es defineix la seva funció generatriu de probabilitats (o simplement funció generatriu), que serà denotada per g_X, de la manera següent:

$$g_X(s) = \sum_{k=0}^{\infty} s^k P\{X = k\}.$$

Aquesta funció (tot i que es pot considerar $s \in \mathbb{C}$) està ben definida, i és finita, en un cert subconjunt dels nombres reals que anomenarem D_X. Una propietat que es comprova immediatament és que aquest conjunt D_X sempre conté l'interval $[-1, 1]$. En efecte, si $|s| \leq 1$, la sèrie que defineix $g_X(s)$ és absolutament convergent ja que

$$\sum_{k=0}^{\infty} |s|^k P\{X = k\} \leq \sum_{k=0}^{\infty} P\{X = k\} = 1.$$

Una altra observació fonamental és que la funció generatriu en cada punt és l'esperança matemàtica d'una certa variable aleatòria:

$$g_X(s) = E[s^X].$$

El nom de *funció generatriu de probabilitats* és degut a que compleix la propietat següent:

$$P\{X=k\}=\frac{g_X^{(k)}(0)}{k!}, \quad \text{per a tot } k\geq 0.$$

I en particular,

$$P\{X=0\}=g_X(0),$$

$$P\{X=1\}=g'_X(0).$$

És a dir, coneixent la funció generatriu podem conèixer totalment les probabilitats de tots els valors de la variable aleatòria. Aquest resultat pot semblar una mica "trampós" ja que, aparentment, per a conèixer la funció generatriu de probabilitats també cal conèixer les probabilitats de tots els valors de la variable aleatòria. Però aquesta propietat també es pot interpretar com que la funció generatriu caracteritza totalment la llei de la variable aleatòria. És a dir, si tenim dues variables X i Y amb la mateixa funció generatriu de probabilitats en un entorn de $s=0$, aleshores X i Y tenen la mateixa llei. Això sí que és un resultat molt útil ja que, usant altres propietats de la funció generatriu, un pot calcular-la per a variables aleatòries tals que no coneixem a priori la seva llei i aquest resultat permet, precisament, identificar aquesta llei. Un cop ressaltada la importància d'aquest resultat, en donarem la demostració. En efecte, el fet que $D_X\subset[-1,1]$ implica que la sèrie de potències que defineix g_X té radi de convergència $r\geq 1$. En particular, això vol dir que en l'interval $(-1,1)$ (com a mínim) la funció és infinitament derivable i com és una sèrie desenvolupada al voltant del punt $s=0$ sabem, a més, que els seus coeficients (que en aquest cas són els $P\{X=k\}$) s'expressen en termes de les successives derivades en $s=0$:

$$P\{X=k\}=\frac{g_X^{(k)}(0)}{k!}.$$

Una altra propietat fonamental de la funció generatriu és el seu comportament respecte la suma de variables aleatòries independents. Concretament, si X i Y són dues variables aleatòries independents tenim que per a tot $s\in D_X\cap D_Y$ es compleix que

$$g_{X+Y}(s)=g_X(s)g_Y(s).$$

Això és molt fàcil de veure ja que les variables aleatòries s^X i s^Y també són variables aleatòries independents i llavors per a qualsevol s en el domini de les dues funcions generatrius tenim

$$g_{X+Y}(s) = E[s^{X+Y}] = E[s^X s^Y] = E[s^X]E[s^Y] = g_X(s)g_Y(s).$$

Una altra propietat que necessitarem usar de la funció generatriu (i amb això també s'assembla a la funció generatriu de moments i a la funció característica) és que determina els moments de la variable aleatòria.

Teorema 3.1. *Si X pren valors a $\mathbb{N}_0$, aleshores*

(i) $E(X) = \lim_{s\uparrow 1} g'_X(s)$,
i això és cert tant si $E[X]$ és finita com si no. (Observem que com X pren sempre valors positius la sèrie que defineix la seva esperança està definida sense ambigüitat però les sumes parcials poden tendir a $+\infty$).

(ii) $E[X(X-1)(X-2)\cdots(X-k+1)] = \lim_{s\uparrow 1} g_X^{(k)}(s)$,
i això és cert tant si aquesta esperança és finita com si no. L'esperança està ben definida ja que $X(X-1)\cdots(X-k+1)$ només pren un nombre finit de valors negatius.

En l'enunciat anterior, la notació $s \uparrow 1$ vol dir que s tendeix a 1 de manera creixent (és a dir, estem fent el límit per l'esquerra). Aquesta notació s'usarà sense més comentaris en la resta del treball.

La demostració d'aquest teorema (només del primer apartat, el segon és molt semblant) es pot trobar a la Secció 7.

Exemple. Calculem la funció generatriu de probabilitats d'una variable aleatòria amb distribució binomial de paràmetres n i p, on $n \in \mathbb{N}$ i $p \in (0,1)$. Recordem que si X té distribució binomial amb els paràmetres anteriors

$$P\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \qquad \text{per a } k = 0, 1, \ldots, n.$$

Aquesta distribució modela una situació en la que fem n experiments aleatoris independents que només poden donar dos resultats, anomenats "èxit" i "fracàs", en què la probabilitat d'èxit a cada prova és p. En aquest context, X denota el nombre d'èxits en les n proves.

El càlcul de la funció generatriu per a la distribució binomial és força senzill:

$$g_X(s) = E[s^X] = \sum_{k=0}^{n} s^k P\{X = k\} = \sum_{k=0}^{n} s^k \binom{n}{k} p^k (1-p)^{n-k} = (q+sp)^n \,,$$

on $q = 1 - p$ és la probabilitat de fracàs a cada prova.

Usant aquest resultat es pot calcular de manera molt fàcil l'esperança i la variància de X.

$$E[X] = \lim_{s \uparrow 1} g'_X(s) = \lim_{s \uparrow 1} \left[np\,(q+sp)^{n-1} \right] = np\,(p+q) = np.$$

Per calcular la variància usem que

$$E[X(X-1)] = \lim_{s \uparrow 1} g''_X(s) = \lim_{s \uparrow 1} \left[n(n-1)p^2 (q+sp)^{n-2} \right] = n(n-1)p^2,$$

d'aquí deduïm que

$$E[X^2] = E[X(X-1)] + E[X] = n(n-1)p^2 + np,$$

i finalment

$$\begin{aligned} \mathrm{Var}(X) = E[X^2] - (E[X])^2 &= n(n-1)p^2 + np - (np)^2 \\ &= np - np^2 = np\,(1-p) = np\,q. \end{aligned}$$

4 El procés de Galton-Watson

Començarem definint de manera formal el model de Galton i Watson. Aquest ve donat per una successió de variables aleatòries $\{X_n,\, n \in \mathbb{N}_0\}$, en què cada variable X_n representa el nombre d'individus *nous* d'una certa població (humans, partícules, bacteris...) a l'n-éssima generació. L'individu k de la n-éssima generació dóna lloc a $Z_{n+1}^{(k)}$ descendents a la generació $n+1$. Fem la hipòtesi que aquestes variables $Z_{n+1}^{(k)}$ són independents entre elles i també independents del nombre d'individus X_n a la generació anterior. A més suposarem que totes les $Z_{n+1}^{(k)}$ tenen la mateixa distribució i, naturalment, prenen valors a $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Així, podem expressar el nombre d'individus nous a la generació $(n+1)$-éssima de la manera següent:

$$X_{n+1} = \sum_{k=1}^{X_n} Z_{n+1}^{(k)}.$$

Potser aquí cal fer alguns comentaris sobre aquesta darrera expressió. Per començar, si la variable X_n val 0 (la població ja està extingida a l'instant n) el sumatori anterior no té sentit i hem de fer el conveni que val 0, cosa totalment lògica, per altra banda. L'altra observació important és que el nombre de sumands és una variable aleatòria, però intuïtivament és clar el que vol dir: si degut a l'atzar a la generació n tenim N individus i cadascun d'aquests pot donar lloc a un nombre, aleatori també, d'individus, el nombre (aleatori) d'individus nous a la generació $n+1$ és la suma dels individus als que han donat lloc els N de la generació anterior. El fet de tenir un sumatori amb un nombre aleatori de sumands farà que haguem d'utilitzar eines de condicionament per al seu tractament, com veurem de seguida.

Una altra suposició que fa el model és que $X_0 = 1$, és a dir, la primera generació parteix d'un individu.

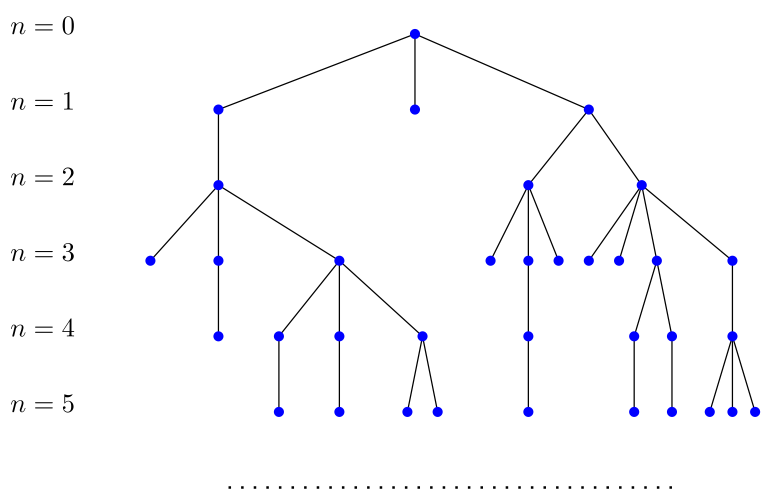


Figura 1: Cinc generacions d'un procés de ramificació.

Com ja s'ha comentat a la introducció històrica, el problema que es va tractar primer (i un dels centrals de la teoria) és el de trobar la probabilitat d'extinció d'una població que segueixi aquest model. I aquest serà el problema que ens ocuparà a nosaltres també. Anomenem doncs ρ a aquesta probabilitat d'extinció que no és més que

$$\rho = P\{X_n = 0,\ \text{per a algun } n \in \mathbb{N}\} = P\Big(\bigcup_{n=1}^{\infty}\{X_n = 0\}\Big). \tag{1}$$

Sota aquest model, es pot donar una solució completa al problema proposat. Només cal conèixer la distribució comuna de totes les $Z_n^{(k)}$. A continuació s'enuncien els dos teoremes que donen aquest resultat. Per simplificar notació, com totes les $Z_n^{(k)}$ tenen la mateixa distribució, Z denotarà una variable amb aquesta llei.

Teorema 4.1. *Sigui g la funció generatriu de probabilitats de Z. Aleshores, la probabilitat d'extinció ρ, definida a (1), és un punt fix de la funció g. És a dir, $s = \rho$ és solució de l'equació*

$$s = g(s). \tag{2}$$

Abans d'enunciar el següent resultat observem que per a qualsevol variable X amb valors en $\mathbb{N}_0$ es té que

$$g_X(1) = \sum_{k=0}^{\infty} 1^k P\{X = k\} = 1.$$

Això vol dir que 1 és sempre una solució de l'equació (2) i el problema és determinar si 1 és o no la solució que ens interessa (si és el cas, la població s'extingirà amb probabilitat 1). El que és interessant és el fet que es pot donar una resposta molt senzilla només en funció del nombre mitjà d'individus en què es reprodueix cadascun dels elements d'una generació, és a dir de l'esperança de les variables $Z_n^{(k)}$.

Teorema 4.2. *Suposem que $P\{Z = 0\} > 0$. Aleshores, en funció del valor de $m = E[Z]$, només tenim aquestes dues possibilitats:*

(i) Si $m \leq 1$, aleshores $\rho = 1$ (és a dir, extinció amb probabilitat 1). De fet, l'equació (2) només té la solució $s = 1$ en tot l'interval $[0,1]$.

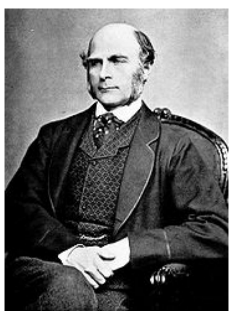
(ii) Si $m > 1$, aleshores $\rho \in (0,1)$. És a dir, no hi ha extinció segura, però la probabilitat d'extinció és sempre estrictament positiva. En aquest cas, a més, ρ és l'única solució de l'equació (2) en l'interval obert $(0,1)$.

Observem que la hipòtesi $P\{Z=0\}>0$ del darrer teorema és raonable ja que si tenim que $P\{Z=0\}=0$ vol dir que cada individu té com a mínim un “fill” i aleshores és clar que la població no s’extingirà, és a dir $\rho=0$. El cas interessant es presenta quan pot ser que un individu no es reprodueixi.

La demostració d’aquests dos teoremes la podeu trobar a la Secció 7.

En la literatura de processos de ramificació els diferents casos segons els valors de m reben els noms següents:

- *Subcrític* quan $m<1$.
- *Crític* si $m=1$.
- *Supercrític* quan $m>1$.


Galton


Watson

5 Uns exemples molt senzills

5.1 Exemple 1

Suposem que tenim un model de Galton-Watson en què els elements es reprodueixen de la manera següent: poden tenir tres descendents o bé no tenir-ne cap, amb probabilitat $1/2$ ambdues possibilitats, i això independentment dels elements i de les successives generacions. És a dir, les variables aleatòries $Z_n^{(k)}$ tenen la mateixa distribució que una variable aleatòria Z que pren els valors 0 i 3 amb probabilitat $1/2$. La funció generatriu de probabilitats de Z ve donada per

$$g(s)=E[s^Z]=s^0P\{Z=0\}+s^3P\{Z=3\}=\frac{1}{2}(1+s^3).$$

Calculem la probabilitat d'extinció de la població generada per un element. Primer hem de veure si la probabilitat d'extinció és 1 o no, per la qual cosa només cal calcular l'esperança de Z. Aquesta esperança es calcula molt fàcilment

$$E[Z] = 0 \cdot P\{Z = 0\} + 3 \cdot P\{Z = 3\} = 3/2.$$

Com $3/2 > 1$ tenim que la probabilitat d'extinció de la nostra població serà l'única solució en l'interval $(0, 1)$ de l'equació

$$s = g(s),$$

és a dir, de

$$s = \frac{1}{2}(1 + s^3).$$

L'equació anterior s'escriu com

$$s^3 - 2s + 1 = 0,$$

però tot i ser de tercer grau la podem reduir immediatament a una de segon grau ja que sabem que $s = 1$ sempre n'és una solució. Dividint per $(s - 1)$, obtenim l'equació

$$(s - 1)(s^2 + s + 1) = 0$$

que efectivament només té una solució en l'interval $(0, 1)$ que és

$$\rho = \frac{\sqrt{5} - 1}{2} \approx 0{,}618.$$

5.2 Exemple 2

Considerem ara un altre exemple per il·lustrar el fet que, malgrat que el valor de $E[Z]$ determina si hi ha extinció segura o no, un mateix valor d'aquesta esperança pot donar probabilitats d'extinció molt diferents. Suposem ara que la variable Z que dóna el nombre de descendents d'un individu té distribució binomial amb paràmetres $n = 3$ i $p = 1/2$. En aquest cas, pels càlculs que hem fet a la Secció 3, $E[Z] = np = 3/2$, com a l'exemple anterior. La distribució de probabilitats de Z ve donada per $P\{Z = 0\} = P\{Z = 3\} = \frac{1}{8}$ i $P\{Z = 1\} = P\{Z = 2\} = \frac{3}{8}$, amb la qual cosa la seva funció generatriu de probabilitats serà:

$$g_Z(s) = \frac{1}{8} + \frac{3}{8}s + \frac{3}{8}s^2 + \frac{1}{8}s^3.$$

La probabilitat d'extinció serà la solució de

$$s = g_Z(s),$$

és a dir:

$$\frac{1}{8}s^3 + \frac{3}{8}s^2 - \frac{5}{8}s + \frac{1}{8} = 0,$$

o equivalentment

$$s^3 + 3s^2 - 5s + 1 = 0,$$

que dividint per $(s-1)$ porta a

$$(s-1)(s^2 + 4s - 1) = 0.$$

Les arrels de l'equació de segon grau resultant són $s_1 = -2 - \sqrt{5}$ i $s_2 = -2 + \sqrt{5}$. Com l'única solució a l'interval $(0, 1)$ és la segona tenim que la probabilitat d'extinció és

$$\rho = -2 + \sqrt{5} \approx 0{,}23607.$$

6 Un exemple degut a Lotka

6.1 La probabilitat d'extinció dels cognoms als Estats Units, usant dades del cens de 1920

A. J. Lotka va ser el primer en usar dades reals dels censos per a calcular la probabilitat d'extinció dels cognoms. A. J. Lotka és ben conegut pel seus treballs sobre dinàmica de poblacions i va publicar el 1924 el primer llibre de biologia matemàtica, *Elements of Mathematical Biology*. Cal remarcar el seu model depredador-presa, desenvolupat simultàniament i de manera independent per V. Volterra.

En un article de 1931 (vegeu [6]), Lotka va aplicar el teorema sobre les probabilitats d'extinció del model de Galton-Watson a les dades obtingudes al cens de 1920 dels Estats Units i va estimar en 0,8797 la probabilitat d'extinció de la línia masculina d'un recent nascut home als Estats Units.

Les dades poblacionals del cens dels Estats Units en aquell 1920 permetien donar la següent distribució dels naixements (independentment del sexe):

n	0	1	2	3	4
probabilitat	0,3686	0,1575	0,1277	0,0972	0,0708
n	5	6	7	8	9
probablitat	0,0490	0,0349	0,0263	0,0216	0,0147
n	10	11	12	13	14
probabilitat	0,0126	0,0075	0,0053	0,0029	0,0017
n	15	16	17	18	19
probabilitat	0,0008	0,0005	0,0002	0,0001	0,0001

En aquesta taula, n indica el nombre de fills d'un home i les probabilitats són les freqüències relatives a la població de cada possible nombre de fills. Com s'ha dit abans, aquesta taula no té en compte el sexe dels fills i aleshores s'ha d'obtenir primer de tot una taula amb les probabilitats dels diferents nombres possibles de fills homes. En l'article citat, Lotka afirmava que la probabilitat de naixement d'un nen de sexe masculí (almenys en aquella època i als Estats Units) era lleugerament superior a la del naixement d'una nena, concretament de 0,515.[2] Usant també que quan una família té n fills el nombre de fills homes tindrà una distribució binomial amb paràmetres n i $p = 0{,}515$, un pot calcular per a cada n la probabilitat que hi hagi k fills homes per a $k = 0, 1, \ldots, n$, que ve donada per

$$\binom{n}{k}(0{,}515)^k(1-0{,}515)^{n-k}.$$

I aleshores, si diem Y a la variable aleatòria que compta el nombre de fills (sense tenir en compte el sexe) i Z el nombre de fills homes d'un home qualsevol, tindrem

$$\begin{aligned} P\{Z=k\} &= \sum_{n=k}^{19} P\{Z=k/Y=n\}P\{Y=n\} \\ &= \sum_{n=k}^{19} \binom{n}{k}(0{,}515)^k(0{,}485)^{n-k}P\{Y=n\}, \end{aligned}$$

[2]Actualment, a Catalunya les dades de naixements dels darrers anys també donen una probabilitat de naixement d'un nen de sexe masculí al voltant del 0,515. Se suposa que aquesta major natalitat dels nens és un resultat de l'evolució humana per compensar la mortalitat infantil, també superior a la de les nenes. En qualsevol cas, per simplificar els càlculs (pensem en què en aquella època no hi havia computadores!), Lotka va acabar fent els càlculs amb $p = 1/2$.

per a $k = 0, \dots, 19$, i on les $P\{Y = n\}$ són les donades a la taula de dalt.

Fent aquests càlculs obtenim les probabilitats següents:

k	0	1	2	3	4
$P\{Z = k\}$	0,4921	0,2089	0,1280	0,0744	0,0432
k	5	6	7	8	9
$P\{Z = k\}$	0,0253	0,0143	0,0076	0,0036	0,0016
k	10	11	12	13	14
$P\{Z = k\}$	0,0006	0,0002	0,0001	$2{,}7 \cdot 10^{-5}$	$7{,}9 \cdot 10^{-6}$
k	15	16	17	18	19
$P\{Z = k\}$	$2 \cdot 10^{-6}$	$4{,}1 \cdot 10^{-7}$	$6{,}4 \cdot 10^{-8}$	$6{,}6 \cdot 10^{-9}$	$3{,}3 \cdot 10^{-10}$

Cal fer notar que en aquesta taula només es mostren les probabilitats arrodonides a les quatre primeres xifres decimals i, quan aquestes xifres eren 0, s'ha usat la notació científica per donar una idea del seu ordre de petitesa. Els càlculs posteriors estan fets usant tot el desenvolupament decimal d'aquestes probabilitats que ens dóna l'ordinador.

Calculem l'esperança d'aquesta distribució:

$$E[Z] = 0 \cdot P\{Z = 0\} + 1 \cdot P\{Z = 1\} + 2 \cdot P\{Z = 2\} + \cdots + 19 \cdot P\{Z = 19\} \\ \approx 1{,}18038$$

Com aquesta esperança és estrictament més gran que 1 ja sabem que la probabilitat d'extinció és l'única solució a l'interval $(0, 1)$ de l'equació

$$s = P\{Z = 0\} + P\{Z = 1\}s + P\{Z = 2\}s^2 + \cdots + P\{Z = 19\}s^{19}.$$

Calculant numèricament les solucions d'aquesta equació, l'única pertanyent a l'interval $(0, 1)$ és:

$$\rho \approx 0{,}8559,$$

que difereix lleugerament de la trobada per Lotka[3].

[3] Les discrepàncies entre aquest resultat i el de Lotka s'han d'atribuir a l'aproximació, comentada abans, de prendre $p = 1/2$ en lloc de $p = 0{,}515$ i també al fet que com a expressió de la funció generatriu de probabilitats va menysprear tots els termes a partir de grau 10, ja que corresponien a probabilitats extremadament petites. Degut a la facilitat de càlcul que tenim actualment, aquí s'han considerat tots els termes.

6.2 Un model matemàtic que s'ajusta a les dades reals

En un altre article del mateix any (vegeu [7]), Lotka va fer una observació molt interessant des del punt de vista de la modelització: va constatar que les probabilitats dels diferents nombres de fills masculins (excepte les molt petites) s'ajustaven molt bé a una distribució *geomètrica modificada*. Una variable aleatòria Y amb aquesta distribució assigna una certa massa al zero i, condicionada a que Y és estrictament positiva, té distribució geomètrica. Això es tradueix en la següent distribució de probabilitats:

$$P\{Y=0\}=\alpha, \qquad P\{Y=n\}=(1-\alpha)(1-\beta)\beta^{n-1}, \text{ si } n\geq 1, \tag{3}$$

on α, β són nombres de l'interval $(0,1)$.

De fet, si aquest model fos correcte (aproximadament) per a la nostra variable Z, en representar gràficament els punts $(k, \log P\{Z=k\})$, amb $k=1,\ldots,19$, hi hauria d'haver una recta que passés molt a prop d'ells. Al següent gràfic podem veure la representació d'aquests punts.

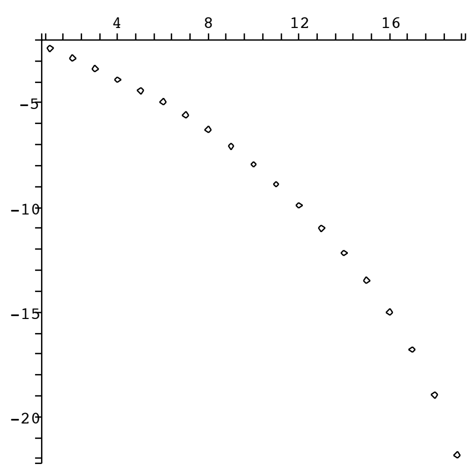


Figura 2: Representació gràfica dels punts $(k, \log P\{Z=k\})$.

És clar que, en realitat, seria més adequat algun altre tipus d'ajust no lineal. Ara bé, si tenim en compte només els punts corresponents a $k=1,2,\ldots,8$ que corresponen als valors amb probabilitats superiors a 4×10^{-3}, si que seria acceptable un ajust lineal. Així, Lotka, va trobar la recta de mínims quadrats que aproximaria bé al núvol de punts format pels $(k, \log P\{Z=k\})$ amb $k=1,\ldots,8$ i d'aquí va calcular les estimacions dels paràmetres[4] α i β .

Recordem que una variable aleatòria U té distribució geomètrica de paràmetre $p=1-\beta$ si

$$P\{U=n\}=(1-\beta)\,\beta^{n-1},\quad \text{per a tot } n\geq 1.$$

És molt fàcil comprovar que la distribució geomètrica comparteix amb la distribució exponencial la propietat de *falta de memòria*, que s'expressa de la manera següent: per a qualssevol n, $m\in\mathbb{N}$

$$P\{U>n+m/U>m\}=P\{U>n\}.$$

Això, en el cas del nostre model voldria dir que si sabem que un home ha tingut més d'un nombre positiu de fills m, la probabilitat de que tingui més de $n+m$ és la mateixa que la probabilitat que des del principi hagi tingut més de n. És a dir, traient a part els homes que, pels motius que siguin, no han tingut fills, el fet de tenir un cert nombre de fills no incrementa ni disminueix les probabilitats de seguir tenint-ne més!! Per altra banda, l'ajust d'una recta als punts $(k, \log P\{Z=k\})$ només era adequat fins a $k=8$, però si pensem que tenir més de 8 fills homes en realitat és molt poc probable, no deixa de ser sorprenent que un model així es pugui ajustar bé. Òbviament, avui en dia, almenys en els països del nostre entorn, un model com aquest seria un disbarat però, en aquella època, podem interpretar que la falta de mètodes anticonceptius (més, possiblement, altres condicionants sociològics, religiosos, etc.) produïa el fenomen de falta de memòria fins a un cert nombre de fills i després l'envelliment dels progenitors sí que disminuïa la probabilitat de tenir més fills.

6.3 Probabilitat d'extinció en el model geomètric modificat

Resoldrem ara el problema general de trobar la probabilitat d'extinció dels cognoms, si la variable Z que dóna el nombre de fills homes de cada individu

[4] Recordem que Lotka treballava en realitat amb aproximacions de $P\{Z=k\}$ prenent com a probabilitat d'un fill de sexe masculí $p=1/2$. Amb les seves dades s'observa un gràfic molt similar.

segueix la distribució descrita en (3). En aquest cas tenim que

$$\begin{aligned} g(s) = E[s^Z] = P\{Z=0\} + \sum_{n=1}^{\infty} P\{Z=n\}s^n \\ &= \alpha + \sum_{n=1}^{\infty}(1-\alpha)(1-\beta)\beta^{n-1}s^n \\ &= \alpha + (1-\alpha)(1-\beta)s\sum_{n=1}^{\infty}\beta^{n-1}s^{n-1} \\ &= \alpha + (1-\alpha)(1-\beta)\frac{s}{1-\beta s}. \end{aligned}$$

Usant el Teorema 3.1 tenim que

$$E[Z] = \lim_{s\uparrow 1} g'(s) = \lim_{s\uparrow 1}\frac{(1-\alpha)(1-\beta)}{(1-\beta s)^2} = \frac{1-\alpha}{1-\beta}$$

Llavors, si $\alpha \geq \beta$ aquesta esperança és menor o igual que 1 i amb probabilitat 1 tenim extinció de la línia masculina de descendents. Si $\alpha < \beta$ aquesta esperança és més gran que 1 i podem calcular la probabilitat d'extinció resolent l'equació

$$s = \alpha + (1-\alpha)(1-\beta)\frac{s}{1-\beta s}.$$

Fent simplificacions elementals arribem a que l'equació anterior és equivalent una equació de segon grau,

$$s^2 - (1+\frac{\alpha}{\beta})s + \frac{\alpha}{\beta} = 0,$$

que té per solucions $s = 1$ i $s = \frac{\alpha}{\beta}$ i per tant, la probabilitat d'extinció serà

$$\rho = \frac{\alpha}{\beta} \in (0,1).$$

6.3.1 Estimació de la probabilitat d'extinció usant el mètode dels moments

Passem ara a fer els càlculs amb les nostres dades. Per tal d'ajustar el model no aplicarem mínims quadrats als punts $(k, \log P\{Z=k\})$, com va fer Lotka, sinó que usarem la idea del mètode dels moments per a l'estimació de paràmetres i després el mètode de màxima versemblança[5] . El mètode dels moments consisteix en expressar primer els moments de la distribució en funció dels paràmetres. En el nostre cas això seria expressar $\mu_1 = E[Z] = f_1(\alpha, \beta)$ i $\mu_2 = E[Z^2] = f_2(\alpha, \beta)$. Aleshores les estimacions $\hat{\alpha}$ i $\hat{\beta}$ seran les solucions del sistema d'equacions

$$\begin{aligned} \hat{\mu}_1 &= f_1(\hat{\alpha}, \hat{\beta}) \\ \hat{\mu}_2 &= f_2(\hat{\alpha}, \hat{\beta}), \end{aligned}$$

on $\hat{\mu}_1$ i $\hat{\mu}_2$ són estimacions de μ_1 i μ_2 a partir de les nostres dades. Així primer de tot necessitarem calcular també el moment de segon ordre de la nostra distribució. Apliquem el Teorema 3.1 i tenim que

$$\begin{aligned} E[Z(Z-1)] = E[Z^2 - Z] &= \lim_{s \uparrow 1} g''(s) \\ &= \lim_{s \uparrow 1} \frac{2\,(1-\alpha)(1-\beta)\beta}{(1-\beta s)^3} = \frac{2\,(1-\alpha)\beta}{(1-\beta)^2}, \end{aligned}$$

i d'aquí deduïm que

$$\begin{aligned} \mu_2 = E[Z^2] &= E[Z] + E[Z(Z-1)] \\ &= \frac{1-\alpha}{1-\beta} + \frac{2\,(1-\alpha)\beta}{(1-\beta)^2} = \frac{(1-\alpha)(1+\beta)}{(1-\beta)^2}. \end{aligned}$$

Així, les equacions del mètode dels moments ens quedaran

$$\begin{aligned} \hat{\mu}_1 &= \frac{1-\hat{\alpha}}{1-\hat{\beta}} \\ \hat{\mu}_2 &= \frac{(1-\hat{\alpha})(1+\hat{\beta})}{(1-\hat{\beta})^2}. \end{aligned}$$

Aquest sistema es pot resoldre fàcilment i s'obté:

$$\hat{\beta} = \frac{\hat{\mu}_2 - \hat{\mu}_1}{\hat{\mu}_1 + \hat{\mu}_2}, \tag{4}$$

$$\hat{\alpha} = 1 - \hat{\mu}_1(1-\hat{\beta}). \tag{5}$$

[5]Hem de tenir en compte que, en realitat, si un calcula la recta de mínims quadrats corresponent a punts de la forma $(x, \log y)$, això no és equivalent a trobar la corba (de tipus exponencial) de mínims quadrats corresponent als punts (x, y).

El que falta ara és trobar unes estimacions de l'esperança i del moment de segon ordre usant les nostres dades. De fet, en el nostre cas les dades són poblacionals, almenys en principi (recordem que les probabilitats dels nombres de fills sense considerar sexe eren les freqüències relatives poblacionals i les probabilitats de fills homes s'han obtingut a partir d'aquestes i aplicant la distribució binomial de paràmetres $n=$ nombre de fills i $p = 0{,}515 =$ probabilitat que un fill sigui home). Així

$$\hat{\mu}_1 = 0 \cdot P\{Z = 0\} + 1 \cdot P\{Z = 1\} + 2 \cdot P\{Z = 2\} + \cdots + 19 \cdot P\{Z = 19\} = E[Z] \approx 1{,}18038,$$

com ja havíem calculat abans, i

$$\hat{\mu}_2 = 0^2 \cdot P\{Z = 0\} + 1^2 \cdot P\{Z = 1\} + 2^2 \cdot P\{Z = 2\} + \cdots + 19^2 \cdot P\{Z = 19\} \approx 4{,}0754.$$

Substituint aquests valors en (4) i (5) obtenim:

$$\hat{\beta} = 0{,}5508,$$
$$\hat{\alpha} = 0{,}4698.$$

Abans d'usar aquests valors per estimar finalment la probabilitat d'extinció sota aquest model, podem veure a la taula següent les probabilitats estimades segons el model geomètric modificat i les probabilitats reals dels diferents nombres de fills. En el càlcul de les probabilitats estimades s'ha aplicat:

$$P\{Z = 0\} \approx \hat{\alpha}, \quad P\{Z = n\} \approx (1 - \hat{\alpha})(1 - \hat{\beta})\hat{\beta}^{n-1}, \text{ per a } n = 1, \ldots, 19.$$

Observem que l'ajust és força bo.

k	0	1	2	3	4
$P\{Z = k\}$	0,4921	0,2089	0,1280	0,0744	0,0432
prob. estimada	0,4698	0,2381	0,1312	0,0723	0,0398
k	5	6	7	8	9
$P\{Z = k\}$	0,0253	0,0143	0,0076	0,0036	0,0016
prob. estimada	0,0219	0,0121	0,0067	0,0020	0,0011
k	10	11	12	13	14
$P\{Z = k\}$	0,0006	0,0002	0,0001	$2{,}7 \cdot 10^{-5}$	$7{,}9 \cdot 10^{-6}$
prob. estimada	0,0006	0,0003	0,0002	0,0001	0,0001
k	15	16	17	18	19
$P\{Z = k\}$	$2 \cdot 10^{-6}$	$4{,}1 \cdot 10^{-7}$	$6{,}4 \cdot 10^{-8}$	$6{,}6 \cdot 10^{-9}$	$3{,}3 \cdot 10^{-10}$
prob. estimada	0,0037	$3{,}1 \cdot 10^{-5}$	$1{,}7 \cdot 10^{-5}$	$9{,}4 \cdot 10^{-6}$	$5{,}2 \cdot 10^{-6}$

Finalment, el càlcul de la probabilitat d'extinció amb el model estimat ens donarà

$$\rho \approx \frac{\hat{\alpha}}{\hat{\beta}} \approx 0{,}8529.$$

6.3.2 Estimació de la probabilitat d'extinció usant el mètode de màxima versemblança

Passem ara a fer els càlculs usant la idea del mètode de màxima versemblança per a l'estimació de paràmetres. Aquest mètode consisteix en prendre com a estimacions d'α i β aquells valors que fan màxima la probabilitat d'obtenir els resultats que s'han obtingut a les dades observades. En el nostre cas, anomenant n_0 al nombre d'homes amb 0 fills homes, n al nombre total d'homes (en realitat no coneixem aquests nombres, però, com veurem de seguida, només necessitarem n_0/n i això si que és conegut) i z_1, $z_2, \ldots, z_n$ les dades diferents de nombre de fills del total de n homes, la probabilitat d'obtenir aquesta successió concreta de nombres de fills és (anomenant $Z_1, \ldots, Z_n$ a les variables aleatòries que donen el nombre de fills dels n homes):

$$\begin{aligned} L(\alpha, \beta; z_1, \ldots, z_n) &= P\{Z_1 = z_1, \ldots, Z_n = z_n\} \\ &= \alpha^{n_0}(1-\alpha)^{n-n_0}(1-\beta)^{n-n_0} \prod_{z_i \neq 0} \beta^{z_i - 1}, \end{aligned}$$

ja que les variables aleatòries Z_i són independents i en la nostra successió hi ha un total de n_0 valors $z_i's$ que són 0 i $n - n_0$ valors z_i que són diferents de zero; a més $P\{Z_i = 0\} = \alpha$ i $P\{Z_i = z_i\} = (1-\alpha)(1-\beta)\beta^{z_i - 1}$, si $z_i \neq 0$.

Observem que la funció L és estrictament positiva sobre l'espai de paràmetres $\Theta = \{(\alpha, \beta) : \alpha \in (0,1), \beta \in (0,1)\}$ i que quan (α, β) tendeix a qualsevol punt de la frontera de Θ, L tendeix a 0 (a la nostra successió hi ha valors $z_i \geq 2$). Això ens diu que aquesta funció tindrà com a mínim un punt de màxim a l'interior de l'espai Θ.

Com la funció logaritme és estrictament creixent, prendrem logaritmes a l'expressió anterior (per tal de facilitar el càlcul de les derivades parcials), i això no modificarà el punt (o punts) on s'assoleix el màxim. Diguem K a la funció obtinguda prenent logaritmes, és a dir:

$$K(\alpha, \beta; z_1, \dots, z_n) = \log L(\alpha, \beta; z_1, \dots, z_n) = n_0 \log \alpha + (n - n_0)\log(1-\alpha) + (n-n_0)\log(1-\beta) + \left[\sum_{i=1}^{n} z_i - (n-n_0)\right]\log\beta.$$

Per trobar els punts de màxim s'ha de resoldre el sistema d'equacions

$$\frac{\partial}{\partial\alpha}K = \frac{n_0}{\alpha} - \frac{n-n_0}{1-\alpha} = 0$$

$$\frac{\partial}{\partial\beta}K = -\frac{n-n_0}{1-\beta} + \frac{\sum z_i - (n-n_0)}{\beta} = 0.$$

Es pot comprovar fàcilment que la solució d'aquest sistema ve donada per

$$\hat{\alpha} = \frac{n_0}{n} \quad \text{i}$$

$$\hat{\beta} = 1 - \frac{1-\hat{\alpha}}{\overline{z}},$$

amb

$$\overline{z} = \frac{\sum_{i=1}^{n} z_i}{n} = E(Z).$$

La darrera igualtat és deguda a que quan les dades són poblacionals, la mitjana no és més que l'esperança de la variable aleatòria Z. També es comprova fàcilment que la matriu Hessiana és definida negativa. Això ens indica que la solució anterior efectivament correspon a l'únic punt de màxim global de la funció L sobre Θ.

En el nostre cas, altre cop perquè les dades són poblacionals

$$\hat{\alpha} = \frac{n_0}{n} = P\{Z = 0\} \approx 0{,}49207$$

i

$$\hat{\beta} = 1 - \frac{1-\hat{\alpha}}{\overline{z}} \approx 1 - \frac{1-0{,}49207}{1{,}18038} \approx 0{,}5697.$$

A la taula següent es mostren les probabilitats estimades usant màxima versemblança segons el model geomètric modificat i les probabilitats reals dels diferents nombres de fills. Observem que l'ajust és força bo per als valors més probables.

k	0	1	2	3	4
$P\{Z=k\}$	0,4921	0,2086	0,1280	0,0777	0,0432
prob. estimada	0,4921	0,2186	0,1245	0,0709	0,0404
k	5	6	7	8	9
$P\{Z=k\}$	0,0253	0,0143	0,0076	0,0036	0,0016
prob. estimada	0,0230	0,0131	0,0074	0,0043	0,0024
k	10	11	12	13	14
$P\{Z=k\}$	0,0006	0,0002	0,0001	$2{,}7\cdot10^{-5}$	$7{,}9\cdot10^{-6}$
prob. estimada	0,0014	0,0008	0,0004	0,0003	0,0001
k	15	16	17	18	19
$P\{Z=k\}$	$2\cdot10^{-6}$	$4{,}1\cdot10^{-7}$	$6{,}4\cdot10^{-8}$	$6{,}6\cdot10^{-9}$	$3{,}3\cdot10^{-10}$
prob. estimada	0,0001	$4{,}7\cdot10^{-5}$	$2{,}7\cdot10^{-5}$	$1{,}5\cdot10^{-5}$	$8{,}7\cdot10^{-6}$

Finalment, el càlcul de la probabilitat d'extinció amb el model estimat ens dóna

$$\rho \approx \frac{\hat{\alpha}}{\hat{\beta}} \approx \frac{0{,}49207}{0{,}5697} \approx 0{,}8637.$$

Com a conclusió, els càlculs sota el model exacte i sota els models aproximats porten a unes probabilitats d'extinció d'un cognom molt semblants entre elles i semblants a les que va obtenir Lotka fent les simplificacions que hem comentat. Aquesta probabilitat es situa al voltant del 85%, cosa que significa que, si s'haguessin mantingut les dades de 1920, a la llarga el 85% dels cognoms dels Estats Units desapareixeria. S'ha de notar que en la distribució del nombre de fills té gran rellevància el fet que gairebé el 50% dels homes d'aquella època no tinguessin cap fill de sexe masculí (i, de fet, gairebé el 37% cap fill).

7 Demostracions dels Teoremes

Demostració del Teorema 3.1.

Com s'ha dit a la secció corresponent, només es provarà l'apartat (i) d'aquest teorema. Per venir definida com una sèrie de potències, la funció g_X és derivable sempre a l'interior de l'interval de convergència i a més es pot derivar terme a terme. Això implica que

$$g_X'(s) = \sum_{k=1}^{\infty} k s^{k-1} P\{X=k\},$$

per a tot $s \in (-r, r)$, on $r \geq 1$ és el radi de convergència de g_X. El problema és que si $r = 1$, la funció en $s = r = 1$ ja no té perquè ser derivable. De tota manera tenim que

$$\sum_{k=0}^{\infty} \lim_{s \uparrow 1} [k\, s^{k-1}\, P\{X = k\}] = \sum_{k=0}^{\infty} k\, P\{X = k\} = E[X].$$

Observem a més que, degut a que $P\{X = k\} \geq 0$, quan $s \uparrow 1$ tenim que

$$k\, s^{k-1}\, P\{X = k\} \uparrow k\, P\{X = k\},$$

és a dir, el límit que hem posat a la sèrie de dalt és creixent. Com és una sèrie de termes positius i el límit és creixent, podem aplicar el Teorema de la Convergència Monòtona per a sèries, que permet intercanviar en aquesta situació el límit amb el sumatori, i tindrem que

$$\lim_{s \uparrow 1} g_X(s) = \lim_{s \uparrow 1} \sum_{k=0}^{\infty} [k\, s^{k-1}\, P\{X = k\}] = \sum_{k=0}^{\infty} \lim_{s \uparrow 1} [k\, s^{k-1}\, P\{X = k\}] = E[X].$$

I això és el que volíem demostrar. $\square$

Demostració del Teorema 4.1

Com $\rho = P\big(\bigcup_{n=1}^{\infty} \{X_n = 0\}\big)$ i tenim la clara inclusió $\{X_n = 0\} \subset \{X_{n+1} = 0\}$ (si a la generació n la població ja s'ha extingit, segur que estarà extingida a la generació $n+1$) ρ és la probabilitat d'una unió creixent d'esdeveniments i, per tant, és el límit de les probabilitats d'aquests esdeveniments:

$$\rho = \lim_{n \to \infty} P\{X_n = 0\}.$$

Diguem h_n a la funció generatriu de cada X_n. Com $h_n(0) = P\{X_n = 0\}$, el que volem és calcular

$$\rho = \lim_{n \to \infty} h_n(0).$$

Ara el que farem és deduir una relació recurrent que satisfan les h_n: Per definició,

$$h_{n+1}(s) = E[s^{X_{n+1}}] = E[s^{\sum_{k=1}^{X_n} Z_{n+1}^{(k)}}],$$

i condicionant a cada possible valor de X_n i, aplicant la fórmula de l'esperança total, la darrera expressió es pot calcular com

$$\sum_{j=0}^{\infty} E[s^{\sum_{k=1}^{X_n} Z_{n+1}^{(k)}}/X_n = j]\, P\{X_n = j\} = \sum_{j=0}^{\infty} E[s^{\sum_{k=1}^{j} Z_{n+1}^{(k)}}/X_n = j]\, P\{X_n = j\}.$$

Usant que les variables $Z_{n+1}^{(k)}$ són independents de les X_n, l'esperança condicionada que apareix a dalt coincideix amb l'esperança sense condicionar i l'expressió anterior val

$$\sum_{j=0}^{\infty} E[s^{\sum_{k=1}^{j} Z_{n+1}^{(k)}}]\, P\{X_n = j\} = \sum_{j=0}^{\infty} (g(s))^j\, P\{X_n = j\} = h_n(g(s)),$$

on hem usat que les $Z_{n+1}^{(k)}$ són independents, totes elles amb funció generatriu g.

Resumint, la successió $\{h_n\}$ satisfà que

$$h_{n+1}(s) = h_n(g(s)),$$

i iterant aquesta relació obtindrem que

$$h_{n+1}(s) = h_n(g(s)) = h_{n-1}(g(g(s)) = \cdots = h_0(g^{n+1}(s)),$$

on estem usant la notació g^{n+1} per a la composició $n+1$ vegades de la funció g amb si mateixa. A més, observem que $h_0(s) = E[s^{X_0}] = E(s) = s$ (recordem la suposició $X_0 = 1$ del nostre model) i, per tant, tenim que

$$h_{n+1}(s) = g^{n+1}(s) = g(g^n(s)) = g(h_n(s)),$$

i, en particular,

$$P\{X_n = 0\} = h_n(0) = g^n(0) \tag{6}$$

i també

$$h_{n+1}(0) = g(h_n(0)). \tag{7}$$

Per altra banda, g és una funció contínua en l'interval $[0,1]$ (com el radi de convergència de g és $r \geq 1$, està clara la continuïtat en l'interval semiobert

$[0, 1)$, però usant un argument de convergència monòtona de sèries, i tenint en compte que la sèrie convergeix en $s = 1$, es veu fàcilment la continuïtat per l'esquerra també en el punt $s = 1$). Usant aquesta continuïtat podem passar al límit quan $n \to \infty$ en ambdós costats de la igualtat (7) i obtindrem el que volíem provar:

$$\rho = g(\rho).$$

□

Demostració del Teorema 4.2. Hem d'estudiar el comportament de g en funció del valor de l'esperança de Z. Observem que en el cas (i) tenim que $g'(1) \leq 1$ (entenent per $g'(1) = \lim_{s\uparrow 1} g'(s)$), mentre que en el cas (ii) tenim que $g'(1) > 1$.

Comencem provant l'enunciat de (i).

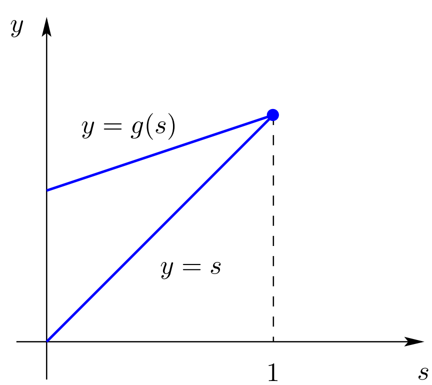
Fem primer un cas particular en què tenim $E(Z) \leq 1$: aquell en què

$$P\{Z \leq 1\} = P\{Z = 0\} + P\{Z = 1\} = 1,$$

llavors, com Z només pren els valors 0 i 1 i, degut a la hipòtesi que hem fet de que $P\{Z = 0\} > 0$,

$$g(s) = P\{Z = 0\} + P\{Z = 1\}\, s$$

té per gràfic una recta amb ordenada a l'origen estrictament positiva i pendent estrictament menor que 1 (vegeu el dibuix).



En aquest cas, està clar que l'únic punt fix de g és 1 i per tant $\rho = 1$. Suposem ara, doncs, que $E(Z) \leq 1$ però $P\{Z \leq 1\} < 1$. En aquest cas, tenim que la funció g és estrictament creixent i convexa en $(0, 1)$. En efecte,

$$g'(s) = \sum_{n=1}^{\infty} n\, P\{Z = n\}\, s^n \geq 0,$$

i a més, $g'(s) = 0$ si i només si per a tot $n \geq 1$ es té que $P\{Z = n\} = 0$, però això implicaria que $P\{Z = 0\} = 1$ i aquesta possibilitat l'hem descartat.

En quant a la convexitat,

$$g''(s) = \sum_{n=2}^{\infty} n(n-1)\, P\{Z = n\}\, s^n \geq 0,$$

i a més $g''(s) = 0$ si i només si per a tot $n \geq 2$ es té que $P\{Z = n\} = 0$, però això implicaria que $P\{Z \leq 1\} = 1$, cosa que estem suposant que no es compleix. Vegem ara si l'equació

$$g(s) - s = 0$$

té o no solucions en l'interval $(0, 1)$. Definim

$$f(s) = g(s) - s$$

i vegem que l'únic zero d'aquesta funció és el punt $s = 1$. Estudiarem el comportament de f'. Com $f''(s) = g''(s) > 0$ en tot l'interval $(0, 1)$ tenim que f' és estrictament creixent. Per tant f' té com a molt un zero en (0,1). Per altra banda, $f'(0) = g'(0) - 1 < 0$ (ja que $g'(0) = P\{Z = 1\} < 1$); aleshores, com tenim que

$$\lim_{s \uparrow 1} g'(s) = E(Z) \leq 1$$

(cosa equivalent a que $\lim_{s\uparrow 1} f'(s) \leq 0$), això implica que $f'(s) < 0$ en tot l'interval $(0, 1)$. Ara, si existís algun punt $s_0 \in (0, 1)$ tal que $f(s_0) = 0$, com $f(1) = 0$, pel teorema de Rolle, f' s'anul·laria en algun punt de $(s_0, 1)$, però acabem de veure que f' és no nul·la en tot $(0, 1)$. Així ja hem provat que si $E(Z) \leq 1$, l'única arrel de l'equació $s = g(s)$ en l'interval $[0, 1]$ és $\rho = 1$.

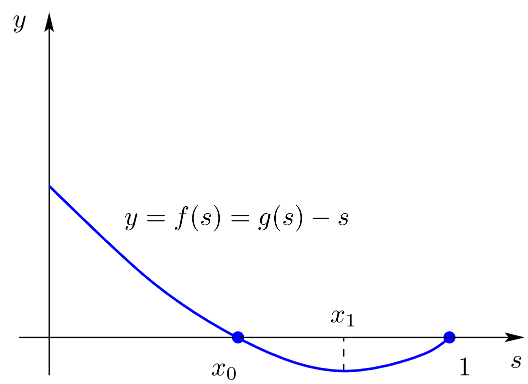
Suposem ara que estem en la situació (ii), és a dir

$$\lim_{s \uparrow 1} g'(s) = E(Z) > 1$$

(això, en particular ens diu que $P\{Z \leq 1\} < 1$ i, usant el raonament de l'apartat anterior, tindrem que g és estrictament creixent i convexa). Definint com abans $f(s) = g(s) - s$, tindrem que

$$\lim_{s \uparrow 1} f'(s) = E(Z) - 1 > 0.$$

Com $f'(0) = g'(0) - 1 = P\{Z = 1\} - 1 < 0$, tenim, aplicant el teorema de Bolzano, que existeix un punt $x_1 \in (0,1)$ tal que $f'(x_1) = 0$. A més, com $g''(s) = f''(s) > 0$ per a tot $s \in (0,1)$, aquest zero de f' és únic i és un mínim de f. Per altra banda, $f(0) = g(0) - 0 = P\{Z = 0\} > 0$ i sabem que $f(1) = 0$. Això ens diu que f tindrà una gràfica com la següent:

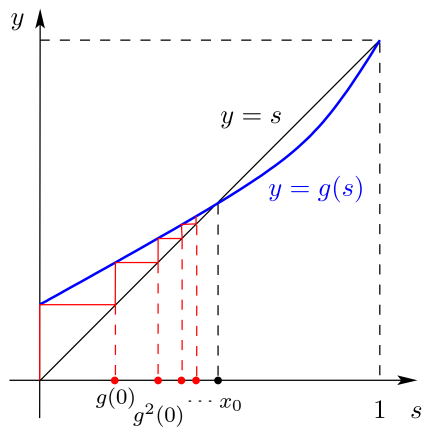


Fixem-nos que $f(x_1) < 0$ forçosament, ja que sabem que f arriba valent 0 en $s = 1$ de forma creixent. Per tant, aplicant altre cop el teorema de Bolzano, existirà un (únic, ja que f' no té més zeros) $x_0 \in (0, x_1) \subset (0,1)$ tal que $f(x_0) = 0$, o el que és el mateix, g té exactament dos punts fixos en l'interval $(0,1]$, aquest x_0 i 1. Ara només cal comprovar que $P\{X_n = 0\}$ no tendeix cap a 1 sinó cap a x_0 i ja haurem acabat.

Tenim que $P\{X_0 = 0\} = 0$ (estem suposant que $P\{X_0 = 1\} = 1$), per tant la nostra successió $\Big(P\{X_n = 0\}\Big)_{n\in\mathbb{N}_0}$ ve donada per (vegeu (6))

$$0,\ g(0),\ g^2(0),\ \dots,\ g^n(0),\ \dots$$

i sabem que té límit i només hem de veure si aquest límit és 1 o x_0. Ara bé, com g és estrictament creixent, g^n ho és també per a tot $n \in \mathbb{N}$, i això vol dir que com $0 < x_0$ tenim $g^n(0) < g^n(x_0) = x_0$ (vegeu el gràfic).



Així, tots els termes de la successió $\{g^n(0)\}_{n\in\mathbb{N}}$ són menors que x_0 i per tant la successió no pot tendir a 1. Això acaba la demostració del teorema.

Referències

[1] Watson, H. W.; Galton, F. (1875), "On the Probability of the Extinction of Families", Journal of the Anthropological Institute of Great Britain **4** 138–144.

[2] Kendall, David G. (1966), "Branching Processes Since 1873", Journal of the London Mathematical Society **41** (1): 385-406, doi:10.1112/jlms/s1-41.1.385.

[3] Kendall, David G. (1975), "The Genealogy of Genealogy Branching Processes before (and after) 1873", Bulletin of the London Mathematical Society **7** (3): 225-253, doi:10.1112/blms/7.3.225

[4] Kolmogorov, A. N. (1938), "Zur Lösung einer biologischen Aufgabe, Isvestiya Nuchno-Issledovatel'skogo Instituta Matematiki i Mekhaniki pri Tomskom Gosudarstvennom Universitete, **2** , 1–12.

[5] Kolmogorov, A. N., Dmitriev, N. A. (1947) "Branching stochastic processes", Dokl. Akad. Nauk U.S.S.R. **56**, 5–8.

[6] Lotka, A. J. (1931) "Population analysis. The extinction of families-I", Journal of the Washington Acaademy of Sciences, **21**(16), 377–380. ISSN 0043-0439.

[7] Lotka, A. J. (1931) "Population analysis. The extinction of families-II",Journal of the Washington Academy of Sciences, **21** (18), 453–459. ISSN 0043-0439.

[8] Yaglom, A. M. (1947), "Certain limit theorems of the theory of branching random processes", Dokl. Akad. Nauk U.S.S.R. **56**, 195–798.

Departament de Matemàtiques
Universitat Autònoma de Barcelona
mjolis@mat.uab.cat

Publicat el 30 de juliol de 2010