Revista Colombiana de Estadística
Nº 5 - 1982

FUNDAMENTOS MATEMATICOS DE LA DEMOGRAFIA

Helmut Knolle

Escuela Superior de Medicina
de Hannover

La demografía es la ciencia que trata de la magnitud del crecimiento de poblaciones humanas y de la frecuencia de los eventos demográficos (nacimiento, muerte, emigración, etc.) en ellas. Un concepto fundamental de la demografía es la distribución de edades en una población, que indica para $n = 1,2,3,\ldots$ el porcentaje de personas, que en un año particular cumplen el n-ésimo año de vida. Conociendo la distribución de las edades de ambos sexos y la frecuencia de fallecimientos y nacimientos dependiendo de la edad, se puede pronosticar el desarrollo futuro de una población cerrada (sin emigración/inmigración). Entre los modelos matemáticos que se usan en la demografía, se distinguen los modelos para un sexo y los modelos para dos sexos. En los modelos para un sexo se supone que existe una abundancia de hombres y que el desarrollo de la población depende únicamente de la subpoblación de las mujeres. Este trabajo considera sólamente modelos para un sexo.

1. La ecuación de Foerster.

Se considera la edad a y el tiempo t como variables continuas. La distribución de edades se sustituye por una función densidad $p(a,t)$ de la manera siguiente:

Para todo $a \geqslant 0$ sea $P(a,t)$ la probabilidad de que una mujer elegida al azar en el tiempo t tenga una edad $\leqslant a$. Sea

$$p(a,t) = \frac{\partial P(a,t)}{\partial a} \quad \text{y} \quad u(a,t) = Np(a,t),$$

donde N es el número total de mujeres. Los cálculos que siguen se efectúan siempre con la función $u(a,t)$. Para todo $a_1 \geqslant 0$ y $a_2 > a_1$ se puede afirmar entonces que

$$\int_{a_1}^{a_2} u(a,t)da = \text{número de mujeres de edad entre } a_1 \text{ y } a_2 ,$$

despreciando la diferencia entre la integral y el número entero vecino.

Además se tiene la función no negativa $\mu(a)$ llamada función de mortalidad. Se supone que la probabilidad de que una mujer de edad a alcance la edad $a+h$ es

$$1 - h\mu(a) + o(h) \qquad \left(\lim_{h\to 0} \frac{o(h)}{h} = 0 \right)$$

Entonces el número de mujeres que en t tienen una edad entre a_1 y a_2 mueren entre t y $t+h$ es

$$M(a_1, a_2, h) = h\int_{a_1}^{a_2} \mu(a)u(a,t)da + o(h)(a_2-a_1)$$

Para a_2-a_1 pequeño se llama el grupo de mujeres de edad entre a_1 y a_2 en t una cohorte de edad, y específicamente la cohorte (a_1,a_2,t). Con esta terminología se puede decir que la cohorte (a_1,a_2,t) perdió $M(a_1,a_2,h)$ miembros en el intervalo $(t,t+h]$. Por otro lado la cohorte tuvo $\int_{a_1}^{a_2} u(a,t)da$ miembros en el tiempo t y tiene $\int_{a_1}^{a_2} u(a+h,t+h)da$ miembros en el tiempo $t+h$. Luego:

$$\int_{a_1}^{a_2} [u(a+h,t+h) - u(a,t)]da = - h \int_{a_1}^{a_2} \mu(a)u(a,t)da + o(h)(a_2-a_1)$$

Como esta identidad es válida para todo a_1,a_2, tenemos

$$u(a+h,t+h) - u(a,t) = -h\mu(a)u(a,t)+o(h)$$

Después de restar y sumar $u(a,t+h)$, de dividir por h y de tomar el límite cuando $h \to 0$, se obtiene la ecuación en derivadas parciales

$$(1) \qquad \frac{\partial u}{\partial a} + \frac{\partial u}{\partial t} = - \mu(a)u$$

que se llama ecuación de Foerster.

A fin de resolver la ecuación (1) se considera

$$v(a,t) = u(a,t)\ \exp(\int_0^a \mu)\ .$$

Para esta función se tiene

$$(2) \qquad \frac{\partial v}{\partial a} + \frac{\partial v}{\partial t} = \left(\frac{\partial u}{\partial a} + \frac{\partial u}{\partial t} + \mu u\right) \exp\left(\int_0^a \mu\right) = 0$$

de donde resulta $v(a,t) = f(a-t)$, con f función diferenciable arbitraria. De hecho, para cualquier f, se tiene

$$\frac{\partial v}{\partial a} = f'(a-t), \quad \frac{\partial v}{\partial t} = -f'(a-t)\ .$$

Por lo tanto $v(a,t) = v(a-t,0) = v(0,t-a)$ o sea

$$u(a,t)\ \exp(\int_0^a \mu) = u(a-t,0)\ \exp(\int_0^{a-t} \mu) = u(0,t-a)$$

Como $u(a,t)$ y $\mu(a)$ están definidas solamente para $a \geqslant 0$, se trabaja la primera identidad con $a \geqslant t$ y la segunda con $a < t$.

La primera identidad da

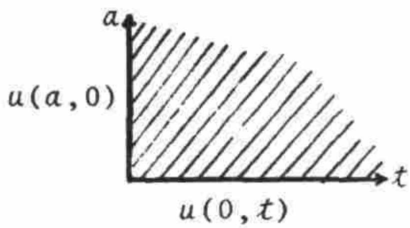
$$u(a,t) = u(a-t,0)\ \exp(-\int_0^a \mu + \int_0^{a-t} \mu) \quad \text{o sea}$$

$$(3a) \qquad u(a,t) = u(a-t,0)\ \exp(-\int_{a-t}^a \mu) \qquad (a \geqslant t)$$

La segunda identidad da

$$\text{(3b)} \qquad u(a,t) = u(0,t-a)\exp\left(-\int_0^a \mu\right) \qquad (a < t)$$

Las fórmulas (3a) y (3b) proporcionan la solución de (1) en el caso que $u(a,0)$ o sea la distribución de edades en $t = 0$ y $u(0,t)$ o sea la frecuencia de los nacimientos en $t \geqslant 0$ estén dadas (Fig.1).



valores dados: ———

valores a calcular: ///

Figura 1

En la práctica $u(0,t)$ no se conoce; se tiene que calcular por medio de la función de maternidad que se define enseguida:

2. La función de maternidad y la ecuación integral de Lotka.

Para todo $a > 0$ sea $H(a)$ el número de hijas de una mujer de edad a. $H(a)$ es una variable aleatoria. Sea $M(a)$ el valor esperado de $H(a)$ y sea $m = \frac{dM}{da}$. Entonces se da la relación

$$\text{(4)} \qquad u(0,t) = \int_0^\omega m(a)u(a,t)\,da$$

donde $m(a)$ es la función de maternidad y ω el límite superior de la edad reproductiva. Si $0 \leqslant t \leqslant \omega$, tenemos por medio de (3a) y (3b)

$$\begin{aligned} u(0,t) &= \int_0^t m(a)u(a,t)da + \int_t^\omega m(a)u(a,t)da \\ &= \int_0^t m(a)u(0,t-a)\exp(-\int_0^a \mu)da \\ &\quad + \int_t^\omega m(a)u(a-t,0)\exp(-\int_{a-t}^a \mu)da \end{aligned}$$

Si $u(a,0)$ está dada, entonces la última integral es una función dada $G(t)$. Además introducimos las notaciones

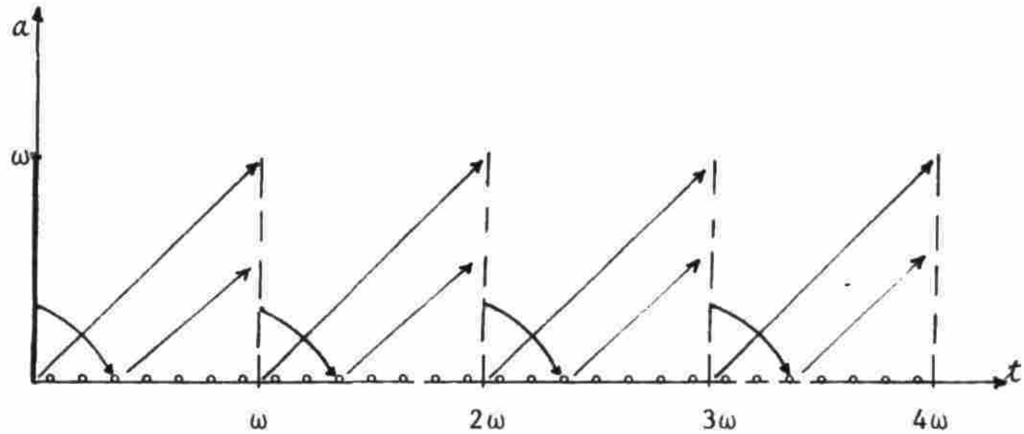
$$\begin{aligned} B(t) &= u(0,t) \\ p(a) &= \exp(-\int_0^a \mu) \\ k(a) &= m(a)p(a) \end{aligned}$$

La función $p(a) \leqslant 1$ es la probabilidad de que una niña recién nacida alcance la edad a. La función $k(a)$ se llama maternidad neta y $\int_0^\omega k(a)da$ es el promedio de hijas de una mujer durante toda su vida.

La ecuación (4) se escribe en la forma

$$(5) \qquad B(t) = G(t) + \int_0^t B(t-a)k(a)da \qquad (0 \leqslant t \leqslant \omega)$$

El método de resolver esta ecuación integral, llamada ecuación de Lotka, será estudiado en el parágrafo 5. Por ahora se supone que es posible resolver la ecuación (5) para $n\omega \leqslant t \leqslant (n+1)\omega$, siempre y cuando $G(t)$ esté dado para t en el mismo intervalo. Veamos como se puede obtener $u(a,t)$ para $t \geqslant 0$, $a-t \leqslant \omega$, si se conoce sólamente $u(a,0)$ para $0 \leqslant a \leqslant \omega$. En la Fig. 2 se muestra este procedimiento.



valores dados: ———
valores calculados por (5): ∘ ∘ ∘ ∘
valores calculados por (3b): — — — —

Figura 2

3. Un modelo discreto.

En la práctica la función $u(a,t)$ se reemplaza por un vector $N(t) = (n_0(t), n_1(t), \ldots, n_K(t))$, donde $n_i(t)$ es el número de mujeres cuya edad está comprendida entre ih y $(i+1)h$, donde h es un intervalo fijo (p.ej. $h = 1, 5, 10$ ó 15 años). Así mismo el eje del tiempo esta discretizado en intervalos de longitud h.

En lo siguiente se toma $h = 1$ año. Sea p_k la probabilidad de que una mujer que ha cumplido el k-ésimo año cumplirá también el $(k+1)$-ésimo año. Entonces

$$n_{k+1}(t+1) = p_k\, n_k(t) \qquad k = 0,1,\ldots,K-1 \tag{6a}$$

Sea β_{k-1} la probabilidad de que una mujer de edad comprendida entre $k-1$ y k año dé a luz una niña durante el año siguiente. Entonces

$$n_0(t+1) = \sum_{k=0}^{K} \beta_k\, n_k(t) \tag{6b}$$

Al formar la matriz

$$A = \begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \cdots & \beta_K \\ \rho_0 & 0 & 0 & & 0 \\ 0 & \rho_1 & 0 & \bigcirc & \vdots \\ & & \ddots & & \vdots \\ & \bigcirc & & \rho_{K-1} & 0 \end{pmatrix}$$

Las ecuaciones (6a) y (6b) se pueden escribir en la forma

$$N(t+1) = AN(t) \tag{7}$$

Para el estudio del crecimiento de la población es posible considerar únicamente la subpoblación de las niñas y de las mujeres en edad reproductiva. Por lo tanto $\beta_K \neq 0$ y $\beta_{K-1} \neq 0$. Además todos los p_i

son $\neq 0$. En estas condiciones (y aũn cuando $\beta_{K-1} = 0$) se puede aplicar el teorema de Frobenius sobre los valores propios de las matrices no negati vas irreducibles. Pero como la matriz A tiene una forma muy especial se pueden derivar de manera direc ta los enunciados de aquel teorema.

Sea C_1 la matriz diagonal con elementos $(1,p_0,1,...,1)$; entonces

$$C_1^{-1}AC_1 = \begin{pmatrix} \beta_0 & \rho_0\beta_1 & & & \vdots & \vdots \\ 1 & 0 & & & \vdots & \vdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Continuando asĩ, se obtiene una matriz $C = C_1C_2 \ldots C_{K-1}$ tal que

$$C^{-1}AC = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{K-1} & a_K \\ 1 & & & & \bigcirc & 0 \\ & 1 & & & & \vdots \\ & & 1 & & & \vdots \\ & \bigcirc & & \ddots & & \vdots \\ & & & & 1 & 0 \end{pmatrix}$$

donde

$$a_0 = a_0$$

$$a_1 = \rho_0\beta_1$$

$$\vdots$$

$$a_K = \rho_0\rho_1 \cdots \rho_{-1}\beta$$

Es bien conocido que $C^{-1}AC$ tiene los mismo valores propios que A.

Para calcular el $\mathrm{Det}(C^{-1}AC - \lambda I)$ se multiplica la primera columna por λ y se suma a la segunda columna. Al final se tiene

$$|C^{-1}AC-\lambda I| = \left|\begin{array}{ccc|c} a_0-\lambda & \lambda(a_0-\lambda)+a_1 & \cdots & \varphi(\lambda) \\ \hline & & & 0 \\ & I_K & & \vdots \\ & & & 0 \end{array}\right| = \varphi(\lambda)$$

donde $\varphi(\lambda) = -\lambda^{K+1} + a_0\lambda^K + a_1\lambda^{K-1} + \ldots + a_{K-1}\lambda + a_K$.

Por consiguiente la ecuación característica de A se lee $\varphi(\lambda) = 0$, o sea

$$(8) \qquad \lambda^{K+1} = a_0\lambda^K + a_1\lambda^{K-1} + \ldots + a_{K-1}\lambda + a_K$$

Como $\varphi(0) = a_k > 0$, $\lim_{\lambda\to\infty} \varphi(\lambda) = -\infty$, existe al menos un valor propio positivo r. Luego vale

$$(8') \qquad r = a_0 + a_1 r^{-1} + \ldots + a_{K-1} r^{-K+1} + a_K r^{-K}$$

Sea $0 < w < r$, entonces

$$w < r = a_0 + a_1 r^{-1} + \ldots + a_K r^{-K} < a_0 + a_1 w^{-1} + \ldots + a_K w^{-K}$$

y en el sentido contrario, si $w > r$. Por lo tanto existe solo un valor propio positivo.

Ahora sea $|z| \geqslant r$ un valor propio complejo o negativo. Como $\beta_{K-1} \neq 0$ y luego $a_{K-1} \neq 0$, se tiene $|a_{K-1}z+a| < a_{K-1}|z| + a_K$ y por lo tanto

$$r < |z| \leqslant a_0 + \frac{a_1}{|z|} + \ldots + \frac{a_K}{|z|^K} \leqslant a_0 + \frac{a_1}{r} + \ldots + \frac{a_K}{r^K}$$

lo que contradice a (8').

El valor propio r es simple. En efecto, tenemos $r\varphi'(r) < (K+1)\varphi(r) = 0$ luego $\varphi'(r) \neq 0$. El vector propio que pertenece a r, tiene las componentes $v_K = 1$, $v_{K-1} = rp_{K-1}^{-1}$, $v_{K-2} = r^2(p_{K-1}p_{K-2})^{-1}$, etc.

Resumiendo lo obtenido:

La matriz A tiene un único valor propio positivo r. Este valor propio es simple y a él corresponde un vector propio con componentes positivas. Los demás valores propios son de valor absoluto menor que r.

Si la matriz A tiene $K+1$ valores propios distintos $\lambda_0, \lambda_1, \ldots, \lambda_K$ con vectores propios $V_0, V_1, \ldots, V_K$ y $r = \lambda_0$, entonces para un vector positivo $N(0)$ cualquiera existen números complejos c_k tales que

$$N(0) = \sum_{k=0}^{K} c_k V_k$$

Por consiguiente $N(1) = AN(0) = \sum_{k=0}^{K} \lambda_k c_k V_k$ y para $m \in \mathbb{N}$

$$N(m) = r^m c_0 V_0 + \sum_{k=1}^{K} \left(\frac{\lambda_k}{r}\right)^m c_k V_k$$

Como $|\lambda_k| < r$ $(k = 1,\ldots,K)$, entonces $N(m) \to r^m c_0 V_0$ cuando $m \to \infty$.

Al vector propio V_0 se le llama distribución de edades estable. Si $N(0) = c_0 V_0$, entonces la población se multiplica cada año por el factor r. En el caso contrario el crecimiento al principio no es uniforme, pero la distribución de edades tiende a la distribución estable y el crecimiento se acerca siempre más al crecimiento geométrico o exponencial.

3.1. Un modelo demográfico de educación.

Todo lo anterior se puede aplicar a poblaciones de animales también. De otro lado se puede reemplazar la variable biológica "edad" por la variable "nivel de educación" y buscar aplicaciones correspondientes. En países en vía de desarrollo se plantea a veces el problema de la falta de maestros, cuando se trata de aumentar el número de alumnos en todo el país. En seguida se propone un modelo muy sencillo, en el cual se identifican las condiciones para un crecimiento apropiado del número de alumnos y maestros. Sea p, s, m_i, el número de alumnos de primaria, secundaria y de profesores de antiguedad $(i-1)$-quinquenios. Además se supone:

a) la fracción a_1 de los alumnos que estuvieron 5 años en primaria pasa a la secundaria que tiene también 5 años,

b) la fracción a_2 de los egresados de secundaria se hace inmediatamente después maestro de primaria,

c) un maestro enseña a n alumnos de primaria,

d) la fracción p_i de los maestros que prestaron servicio durante $5(i-1)$ añossiguen siendo maestros ($i = 1,\ldots,K-1$),

e) el crecimiento de las cantidades p y s no es limitado por el número de niños y de maestros de secundaria.

Por consiguiente se tienenlas ecuaciones:

(9) a) $s(t+5) = a_1 p(t)$

b) $m_1(t+5) = a_2 s(t)$

c) $p(t) = n \sum_{i=1}^{K} m_i(t)$

d) $m_{i+1}(t+5) = p_i m_i(t) \qquad i = 1,\ldots,K-1$

La ecuación (9)c) se puede escribir en la forma

$$p(t+5) = n \sum_{i=1}^{K} m_i(t+5) = na_2\ s(t) + n \sum_{i=2}^{K} m_{i-1}(t) p_{i-1}$$

Después de introducir el vector V y la matriz A :

$$V = \begin{pmatrix} p \\ s \\ m_1 \\ \cdot \\ \cdot \\ \cdot \\ m_{K-1} \end{pmatrix} \qquad A = \begin{pmatrix} 0 & na_2 & np_1 & \cdots & np_{K-1} \\ a_1 & 0 & 0 & & 0 \\ 0 & a_2 & 0 & \bigcirc & \cdot \\ 0 & 0 & p_1 & & \cdot \\ & & & \ddots & \cdot \\ & \bigcirc & & p_{K-2} & 0 \end{pmatrix}$$

Las ecuaciones (9) se pueden escribir en la forma

(10) $$V(t+5) = AV(t)$$

La matriz A tiene todas las características que tenía en el parágrafo 3. Por lo tanto se puede concluír que si $V(0)$ es el vector propio que pertenece al único valor propio positivo r de la matriz A, entonces el número de alumnos de primaria se multiplica por r cada 5 años y se cumple la condición de que a n alumnos corresponda en media un maestro.

El modelo se puede generalizar incluyendo la enseñanza universitaria y los maestros de secundaria. Pero entonces la matriz A sería más complicada y se tendría que aplicar el teorema de Frobenius sobre matrices no negativas irreducibles (ver p.ej. H. Knolle, Matemáticas y Ciencias, en: Matemática-enseñanza universitaria N°10, Bogotá, Marzo 1979).

4. Modelo contínuo: Solución por medio de la transformación de Laplace.

Sea f una función contínua definida sobre $[0,\infty)$ tal que para algún a y c positivo se tenga que $|e^{-at}f(t)| \leqslant c$ para todo $t \geqslant 0$. Entonces se define la transformada de Laplace de f:

$$L(f)\ (p) = \int_0^\infty e^{-pt}\ f(t)dt$$

para todo valor complejo p tal que $Re(p)$ es mayor que algún p_0. Sin discutir las condiciones de existencia de las transformadas correspondientes, afirmamos que después de aplicar la transformación a la ecuación (5) se obtiene

$$L(B) = L(g) + L(B)\ L(k)$$

o sea

$$L(B) = \frac{L(g)}{1-L(k)}$$

Si el denominador tiene los ceros distintos, $\lambda_1,\lambda_2,...$ entonces

(11) $$L(B)(p) = \frac{Q_1}{p-\lambda_1} + \frac{Q_2}{p-\lambda_2} + ...$$

donde

$$Q_i = \lim_{p\to\lambda_i} \frac{(p-\lambda_i)L(g)(p)}{1-L(k)(p)}$$

Por lo tanto se debe analizar la ecuación

$$(12) \qquad 1 = L(k)(p) = \int_0^\infty e^{-pt}\, k(t)dt$$

Como $k(t) \geq 0$ para todo t y $k(t) > 0$ en algún intervalo, es obvio que $L(k)(p)$ es una función estrictamente decreciente de p. Además $L(k)(p)$ tiende a 0 ó ∞ cuando p tiende respectivamente a ∞ ó $-\infty$. Por consiguiente (12) tiene exactamente una solución real ρ. Sea $z = u+iv$ una solución compleja. Entonces

$$Re(e^{-tz}) = e^{-tu}\cos tu \quad \text{y} \quad 1 = \int_0^\infty e^{-tu}\cos(tu)k(t)dt.$$

Como ρ satisface (12) y como $\cos(tu) < 1$ en casi toda parte, se tiene que $u < \rho$. Eso significa que cada solución compleja de (12) tiene parte real menor que ρ.

La transformada inversa de $(\lambda-p)^{-1}$ es $e^{\lambda t}$.

Por lo tanto se obtiene de (11) $B(t) = \sum c_i\, e^{\lambda_i t}$ con coeficientes complejos c_i o bien

$$B(t) = c_1 e^{\rho t} + \sum_{i\geq 2} \exp(t\, Re\lambda_i)\, [c_{i1}\cos(t\, Im\lambda_i) + c_{i2}\operatorname{sen}(t\, Im\lambda_i)]$$

con coeficientes reales c_{i1}, c_{i2} y c_1.

Como $\rho > Re\lambda_i$ $(i = 2,3,\ldots)$ se tiene que

$$\lim_{t\to\infty} e^{-\rho t}\, B(t) = c_1.$$

ρ es asintóticamente el coeficiente de crecimiento de $B(t) = u(0,t)$.

De la teoría de estabilidad para ecuaciones diferenciales ordinarias se conoce el hecho de que un punto que atrae todas las soluciones que pasan por una vecindad de él, es también una solución de la ecuación diferencial. Por lo tanto se considera el caso $B(t) = c_1\, e^{\rho t}$.

De (3b) se deduce

$$u(a,t) = c_1 e^{\rho t}\cdot e^{-\rho a}\,\exp\left(-\int_0^a \mu\right) = c_1 g(t) h(a)$$

donde

$$g(t) = e^{\rho t} \tag{13}$$

$$h(a) = \exp\left(-\rho a - \int_0^a \mu\right) \tag{14}$$

Se verifica inmediatamente que $u = g(t)h(a)$ satisface la ecuación de Foerster y es, salvo un factor constante la única solución de la forma $g(t)h(a)$. En efecto, si $u(a,t) = g(t)h(a)$ es solución de (1) y (4), entonces

$$\frac{\partial u}{\partial a} + \frac{\partial u}{\partial t} = g(t)h'(a) + g'(t)h(a) = -\mu(a)g(t)h(a)$$

y después de dividir por $g(t)h(a)$ se obtiene

$$\frac{g'(t)}{g(t)} = -\frac{h'(a)}{h(a)} - \mu(a) = \text{const} = \lambda \qquad \lambda \in \mathbb{R}$$

y $g(t) = c_1 e^{\lambda t}$, $h(a) = c_2 \exp\left(-\lambda a - \int_0^a \mu\right)$, $u(a,t) = g(t)h(a)$.

Reemplazando en (4) se deduce

$$e^{\lambda t} = \int_0^\infty e^{\lambda t} e^{-\lambda a} m(a) \exp(-\int_0^a \mu) da$$

o bien, teniendo en cuenta la definición de $k(a)$,

(15) $$1 = \int_0^\infty e^{-\lambda a} k(a) da$$

y se vuelve a la ecuación (12) que, como ya se sabe, tiene sólo una solución real. La función $h(a) = \exp(-\rho a - \int_0^a \mu)$ se llama distribución de edades estable.

Luego se puede enunciar el siguiente

Teorema. La ecuación (1) junto con la condición inicial (4) tiene, salvo un factor constante, exactamente una solución u_0 de la forma $u_0(a,t) = g(t)h(a)$. Las funciones g y h son dadas por (13) y (14), donde ρ es la única solución real de (15). Cada solución tiende a u_0, si t tiende a ∞.

De (14) se desprende también el hecho de que en la distribución de edades estable la fracción de niñas de edad menor que algún a_0 es una función creciente de ρ. Es suficiente observar que $h'(a) = (-\rho a - \mu(a))h(a)$ es una función decreciente de ρ y comparar las distribuciones de edades en la Fig. 3.

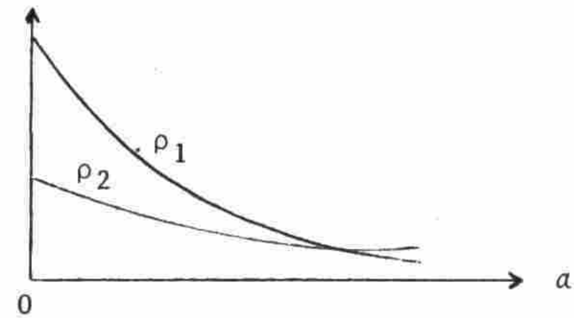


Figura 3

Distribución de edades para coeficientes de crecimiento $\rho_1 > \rho_2$.

*

BIBLIOGRAFIA

Keyfitz, N., *Introduction to the Mathematics of Population*, Addison-Wesley, 2 ed., 1977.

Hoppensteadt, F., *Mathematical Theories of Populations: Demographics, genetics and Epidemics*, SIAM, Philadelphia, 1975.

Se menciona también la bibliografía en el artículo de David Ospina, Revista Colombiana de Estadística, N° 3, Bogotá 1981, p.86-87.

* * *