Simulation, Final project: Demography

Søren Asmussen*

December 2, 2004

Stable population theory

We first recall from week 12 the definition of a Galton-Watson process with a finite set $1, \ldots, q$ of types. The process is vector-valued, $\boldsymbol{X}_n = \big(X_n^1, \ldots, X_n^q\big)$ where X_n^j is the number of type j individuals at time n, and the children of the ith type k individual is a vector $\boldsymbol{Z}_{n,i}^j = \big(Z_{n,i}^{k,1}, \ldots, Z_{n,i}^{k,q}\big)$. Thus

$$X_{n+1}^j \;=\; \sum_{k=1}^{q} \sum_{i=1}^{X_n^k} Z_{n,i}^{k,j}.$$

The expected offspring production is given by the $m_{jk} = \mathbb{E} Z_{n,i}^{j,k}$. Let $\boldsymbol{M}$ be the $q \times q$ matrix with elements m_{jk}. One can show under an irreducibility assumption on $\boldsymbol{M}$ that the largest eigenvalue λ of $\boldsymbol{M}$ is strictly positive and that

$$\mathbb{E}^{(j)}\big[X_n^k\big] \;\sim\; c_j d_k \lambda^n$$

where $\mathbb{E}^{(j)}$ refers to one intitial individual of type j and we will adapt the normalization $d_1 + \cdots + d_q = 1$. One calls c_j the *reproductive value* of an type j individual, $\boldsymbol{d} = \big(d_1, \ldots, d_q\big)$ is the *stable type distribution* and λ (or sometimes $\log\lambda$) is the *Malthusian rate of growth* .

With general initial conditions $\boldsymbol{X}_0 = \boldsymbol{x}$, it follows immediately that

$$\mathbb{E}_{\boldsymbol{x}}\big[X_n^k\big] \;\sim\; c(\boldsymbol{x}) d_k \lambda^n$$

where

$$c(\boldsymbol{x}) \;=\; \boldsymbol{x} \cdot \boldsymbol{c} \;=\; \sum_{j=1}^{q} c_j x_j.$$

Further, one can show that

$$X_n^k \;\sim\; \lambda^n d_k W \tag{1}$$

where W is a r.v. satisfying $\mathbb{E}_{\boldsymbol{x}} W = c(\boldsymbol{x})$. In many arguments to follow, the equations contain $c(\boldsymbol{x})$ on both sides and we will tacitly omit $c(\boldsymbol{x})$ then.

*Department of Theoretical Statistics. Department of Mathematical Sciences, Aarhus University, Ny Munkegade, 8000 Aarhus C, Denmark; asmus@imf.au.dk; http://home.imf.au.dk/asmus

The example we have in mind is indeed from demography. We consider the female part of a human population divided into $q+1$ age groups $0,\ldots,q$ so that the type is age in years. In traditional actuarial notation, we write ${}_a p_x$ for the probability of surviving from age x to $a > x$. Let further β_a denote the probability of giving birth at age a. Then the possibilities for the children of an age a individual are

- one $a+1$ individual w.p. ${}_{a+1}p_a(1-\beta_a)$
- one $a+1$ and one 0 individual w.p. ${}_{a+1}p_a\beta_a$
- one 0 individual w.p. $(1-{}_{a+1}p_a)\beta_a$
- none w.p. $(1-{}_{a+1}p_a)(1-\beta_a)$

What are the Malthusian parameter λ, the stable age distribution and the reproductive value? The present demographic set-up turns out to be simpler than a general multitype GW process, and we will be able to come up with explicit formulas without having to invoke matrix calculus.

The expected number of daughters born to a baby girl is

$$m \;=\; \sum_{a=0}^{q} {}_a p_0 \beta_a \,.$$

Intuitively, the stable age distribution $\boldsymbol{d}$ should have the property that if the expected age distribution in year n is $\boldsymbol{d}$, then it should be so in year $n+1$. Thus, if $m=1$, the expected proportion $d_a = d_a^n$ of the population in age a should not depend on n and satisfy $d_a^{n+1} = d_{a-1}^n {}_a p_{a-1}^n$ which gives

$$d_a \;=\; d_{a-1\,a}p_{a-1} \;=\; d_{a-2\,a}p_{a-1\,a-1}p_{a-2} \;=\; \ldots \;=\; d_{0\,a}p_{a-1}\cdots{}_1p_0 \;=\; d_{0\,a}p_0 \,.$$

Recalling $d_1+\cdots+d_q=1$, this means

$$d_a \;=\; \frac{{}_a p_0}{\sum_{a'=0}^{q} {}_{a'}p_0} \;=\; \frac{\mathbb{P}(T>a)}{\mathbb{E}T} \tag{2}$$

where $T=T_0$ is a r.v. having the distribution of the life length of a baby girl. When $m=1$, it is further clear that one should have $\lambda=1$.

When $m>1$, one should have $\lambda>1$. To compute $\boldsymbol{d}$, note that the age a individuals at time n were born at time $n-a$. For n large, this gives $\lambda^n d_a = \lambda^{n-a} d_{0\,a}p_0$, and we conclude that

$$d_a \;=\; \frac{\lambda^{-a}\,{}_a p_0}{\sum_{a'=0}^{q} \lambda^{-a'}\,{}_{a'}p_0}. \tag{3}$$

Thus, compared to the case $m>1$ the stable age distribution has been biased in favour of small ages which is also intuitively reasonable. What is the precise value of λ? Splitting the expected number of baby girls at time $n+1$ up according to the age of their mother gives

$$\lambda^{n+1} d_0 \;=\; \sum_{a=0}^{q} \lambda^n d_a \beta_a, \tag{4}$$

$$\lambda \;=\; \sum_{a=0}^{q} \lambda^{-a} d_{a\,a}p_0 \tag{5}$$

where we used (3). The l.h.s. of (5) is increasing in λ with limits $1, \infty$ at $1, \infty$, whereas the r.h.s. is decreasing with limits $m > 1, 0$. Thus we conclude that there is a unique solution of (5) and that we must have $\lambda > 1$.

Assignment

You are given a female population of size N, divided into the age groups $0, \ldots, 100$. The mortality is given by the G82 parameters used in week 8's assignment and the probability of giving birth at age a and parity p is

$$b_{a,p} \;=\; \begin{cases} z_0\beta_a & p = 0 \\ z_1 z_2^{p-1}\beta_a & p = 1, 2, \ldots \end{cases} \tag{6}$$

where $\beta_a = (a-15)^3(46-a)^3$ for $a = 16, \ldots, 45$ and $\beta_a = 0$ otherwise. Further, a baby is a boy w.p. 51% and a girl w.p. 49%.

A) Determine z_0, z_1 such that the probability of a woman getting at least one child is 95% and 75% of getting at least two. Use these parameters in the following.

B) By varying z_3, illustrate the connection between the expected number m of children born to a woman and 1) the Malthusian parameter λ, 2) the stable age distribution, using what you consider a relevant range for m.

C) In a population with $m = 3$, the government wants to stop the population exploding by introducing drastic sanctions for getting more than two children. The government expects that this will change (6) to

$$b_{a,p} \;=\; \begin{cases} z_0\beta_a & p = 0 \\ z_1\beta_a & p = 1 \\ \frac{1}{10} z_1 z_2^{p-1}\beta_a & p = 2, 3, \ldots \end{cases} \tag{7}$$

Find the new λ and give projections for the age distribution 5,10,25 and 50 years ahead.

Comments

A) can be done by simulation or, alternatively, by fairly simple non-stochastic numerical analysis.

Use a multitype branching process for the woman only, but note that parity also counts number of sons. You will need (1)

Age distributions may be illustrated by histograms or one-dimensional characteristics like the proportion of old (66 or more) people to working (18–65) people, or children (17 or less) to working people.

Estimates for one-dimensional quantities (λ, the proportions above etc.) should be accompanied by confidence intervals. In some cases, you may need the delta method.

You should of course start with a small N like 1.000, but aim for a population of several million. My MATLAB program could update a population of 1.000.000 one year ahead in a few seconds.

References

[1] J.H. Pollard (1973) *Mathematical Models for the Growth of Human Populations*. Cambridge University Press.

[2] S.H. Preston, P. Heuveline & M. Guillot (2001) *Demography. Measuring and Modeling Population Processes*. Blackwell.