Developments in Statistics and Methodology
A. Ferligoj and A. Kramberger (Editors)
Metodološki zvezki, 9, Ljubljana: FDV 1993

Natural Population Growth Models

Drago Čepar[1], Lojze Gosar[2], Erika Uršič[1]

Abstract

Annual model for natural growth of a population is a specific case of Leslie models. To use it we need one year age specific fertility and survival ratios and one year population age distribution. We describe algorithms for approximation of one year demographic data from five year data. A very simple natural growth model is also described. Its solution is obtained analytically as an explicit function of initial population pyramid, survival and fertility ratios.

Keywords:: Demographic models; Fertility; Mortality; Natural growth.

1 Introduction

Annual model for natural growth of a population is a specific case of Leslie models described in [1]. To use it we need one year age specific fertility and survival ratios and one year population age distribution. We propose an algorithm to approximate these one year data from five year data. The discussion of the behaviour of the annual model solutions is complicated by the fact that the general solution can not be obtained as an explicit function of initial population pyramid and survival and fertility ratios. We describe a very simple model and give its solutions in an explicit form. In section 2, a well-known model of natural population growth based on one-year age-group data is described; in section 3, transformations of five-year to one-year data are described; in section 4, a simple linear-difference-equations model is introduced, solved analytically, and the long term behaviour of its explicit solution discussed.

2 Annual model for natural population growth

It is described by the following equations:

[1]Jožef Stefan Institute, Jamova 39, 61 111 Ljubljana, Slovenia
[2]Urban Planning Institute, Jamova 39, 61 111 Ljubljana, Slovenia

$$N^{t+1}_{s,i+1} = N^t_{s,i} \cdot P_{s,i}, \ i = 0,1,2,..,83, s = 1,2,$$

$$N^{t+1}_{s,85} = N^t_{s,84} \cdot P_{s,84} + N^t_{s,85} \cdot P_{s,85}, \ B = 1/2 \cdot \sum_{j=15}^{49} (N^t_{2,j} + N^{t+1}_{2,j+1}) \cdot R_j, \quad (1)$$

$$N^{t+1}_{1,0} = 0.517 \cdot B \cdot P'_1, \ N^{t+1}_{2,0} = 0.483 \cdot B \cdot P'_2.$$

$N^t_{s,i}$ is the number of males (s=1) or females (s=2) belonging to i^{th} age group (defined as [i,i+1) for i=0,1,..,84 and [i,∞) for i=85) in year t. R_i, i=15,16,..,49, are one-year specific fertility ratios and $P_{s,i}$, s=1,2, i=0,1,..,85 one-year specific survival ratios. P'_s, s=1,2, are survival ratios for infant children. Variable B represents the number of newborns in year t. As the number of women $N^t_{2,j}$ varies throughout the year, we take the arithmetic mean of $N^t_{2,j}$ and $N^{t+1}_{2,j+1}$ as the average number of women belonging to j^{th} age-group in year t. The constants 0.517 and 0.483 represent the ratios of the numbers of live-born boys and girls with respect to the total number of live-born children. Model (1) is taken from [2], for more about it see also [3]. In several sources e.g., [1], [4] and [5] it is given in matrix notation, which has several advantages, but requires some more space for introduction.

3 Transformations

Population data for one-year age groups for different countries in the world are not always available. They are usually collected for five-year age groups. In many demographic yearbooks, population data is given for the following age groups: 0 year, 1-4 years, 5-9 years, 10-14 years,..., 80-84 years and the last age group 85+ years (people aged 85 or over). For the purpose of using model (1), we have tried to approximate one-year age distribution, fertility and survival data by five-year data. The estimation errors caused by those approximations have been observed in specific cases. Age group from i to j will be denoted by index 'i-j'.

3.1 Age distribution

Five-year age group data can be expressed by one-year data as follows:

$$N^t_{s,0} = N^t_{s,0} \ , \quad N^t_{s,1-4} = \sum_{i=1}^{4} N^t_{s,i} \ , \quad N^t_{s,i-(i+4)} = \sum_{j=i}^{i+4} N^t_{s,j} \ , \quad i = 5,10,..,80 \ ,$$

$$N^t_{s,85+} = N^t_{s,85} \ .$$

The simplest approximation of one-year data from five-year data is the following one:

$$N_{s,0}^t = N_{s,0}^t, \quad N_{s,1}^t = N_{s,2}^t = N_{s,3}^t = N_{s,4}^t = \frac{1}{4} \cdot N_{s,1-4}^t, \quad N_{s,85}^t = N_{s,85+}^t \ ,$$

$$N_{s,i}^t = N_{s,i+1}^t = N_{s,i+2}^t = N_{s,i+3}^t = N_{s,i+4}^t = \frac{1}{5} \cdot N_{s,i-(i+4)}^t, \quad i = 5, 10, .., 80 \ . \tag{2}$$

The resulting population pyramid is staircased. We can reduce the approximation errors by reducing the height of steps of step function (2) simultaneously taking into account the conditions $N_{s,1-4}^t = \sum_{i=1}^{4} {'N_{s,i}^t}$ and $N_{s,i-(i+4)}^t = \sum_{j=i}^{i+4} {'N_{s,j}^t}$ for i=5,10,..,80, where ${'N_{s,i}^t}$ are corrected values. The underlying equations are:

$${'N_{s,0}^t} = N_{s,0}^t, \quad {'N_{s,85}} = N_{s,85+}$$

$$a_s = N_{s,1}^t, \quad h_s = (N_{s,0}^t - a_s)/4,$$

$${'N_{s,1}^t} = a_s + h_s, \quad {'N_{s,2}^t} = a_s + \frac{h_s}{3}, \quad {'N_{s,3}^t} = a_s - \frac{h_s}{3}, \quad {'N_{s,4}^t} = a_s - h_s,$$

and for $i = 5, 10, ..., 80$: $\quad H_{s,i} = (N_{s,i-3}^t - N_{s,i+2}^t)/5, \quad {'N_{s,i}^t} = N_{s,i+2}^t + 2H_{s,i},$

$${'N_{s,i+1}^t} = N_{s,i+2}^t + H_{s,i}, \quad {'N_{s,i+2}^t} = N_{s,i+2}^t,$$

$${'N_{s,i+3}^t} = N_{s,i+2}^t - H_{s,i}, \quad {'N_{s,i+4}^t} = N_{s,i+2}^t - 2H_{s,i} \ .$$

3.2 Survival ratios

We denote five-year survival ratios by $P_{s,0}^5$, $P_{s,1-4}^5$, $P_{s,i-(i+4)}^5$, i=5,10,..,80, $P_{s,85+}^5$. Five-year survival ratio is the probability that a person who belongs to the corresponding age group will survive the next five years. Five-year survival ratios can be expressed as follows:

$$P_{s,0}^5 = \frac{N_{s,5}^{t+5}}{N_{s,0}^t}, \quad P_{s,1-4}^5 = \frac{N_{s,6}^{t+5} + N_{s,7}^{t+5} + N_{s,8}^{t+5} + N_{s,9}^{t+5}}{N_{s,1}^t + N_{s,2}^t + N_{s,3}^t + N_{s,4}^t}, \quad P_{s,85+}^5 = \frac{\sum_{i=85}^{100} N_{s,i+5}^{t+5}}{\sum_{i=85}^{100} N_{s,i}^t} \ ,$$

$$P_{s,i-(i+4)}^5 = \frac{N_{s,i+5}^{t+5} + N_{s,i+6}^{t+5} + N_{s,i+7}^{t+5} + N_{s,i+8}^{t+5} + N_{s,i+9}^{t+5}}{N_{s,i}^t + N_{s,i+1}^t + N_{s,i+2}^t + N_{s,i+3}^t + N_{s,i+4}^t} \ , \quad i = 5, 10, .., 80 \ .$$

The number of people who are i years old in calendar year t and will survive the next five years is $N_{s,i+5}^{t+5} = N_{s,i}^t \cdot P_{s,i} \cdot P_{s,i+1} \cdot P_{s,i+2} \cdot P_{s,i+3} \cdot P_{s,i+4}$. So all one-year survival ratios are needed at least for up to the age of 100; they are usually available in complete life tables.

To obtain one-year survival ratios from five year ratios we propose the following equations:

$$P_{s,0} = \sqrt[5]{P^5_{s,0}},\ P_{s,5} = \sqrt[5]{P^5_{s,1-4}},\ P_{s,i+4} = \sqrt[5]{P^5_{s,i-(i+4)}},\ i = 5, 10, .., 80\ ;\ P_{s,85} = \sqrt[5]{P^5_{s,85+}}$$

We obtain the remaining ratios by linear interpolation given by the equations:

$$P_{s,i} = P_{s,5} + (i-5)\cdot d_s\ ,\quad i = 1, 2, .., 8 \quad where \quad d_s = \frac{1}{4}(P_{s,9} - P_{s,5})\ ;$$

$$P_{s,i+k} = P_{s,i-1} + \frac{(k+1)}{5}\cdot (P_{s,i+4} - P_{s,i-1})\ \ for\ \ k = 0,1,2,3,\ \ i = 10, 15, .., 80\ .$$

3.3 Fertility ratios

The five-year ratio $R_{i-(i+4)}$ represents the proportion between the number of babies born by mothers belonging to the age group i-(i+4) in one year t and the number of mothers in that age group.

If we know the *total fertility rate* $T_f = \sum_{i=15}^{49} R_i$ and the five-year fertility ratios, R_i can be approximated as follows:

$$R_i = R_i \cdot (T_f / \textstyle\sum_{i=15}^{49}\ 'R_i)\ ,$$

where $'R_i = R_i^* \cdot (R_{j-(j+4)}/R^*_{j-(j+4)})$ *for* $j = 15, 20, ..., 45$, $i = j, j+1, j+2, j+3, j+4$. T_f is the average number of children one woman gives birth to during her whole life. R_i^* and $R^*_{j-(j+4)}$ are one-year and five-year fertility ratios for women belonging to the population for which we believe, that it has the similar fertility structure as observed population.

3.4 Some empiric results and errors

Results of population projections for some European countries are presented in Table 1. The last year for which a population pyramid is available is taken as the initial year t_0. Ratios R_i and $P_{s,i}$ are one-year ratios calculated from the latest available five-year ratios.

It is clear, that the far future values (more than 200 years ahead) have little or no relevance as predictions. The assumptions of nonexistence of migrations and long term invariability of fertility ratios are too unrealistic. So nobody expects the population of Italy will decrease to 15% of its present size or that of former GFR to 6%. Nevertheless these 'projections' can be precious as possible way of illustrating the values of present survival and fertility ratios.

The error of the forecast of a country's total population in 2200 due to the approximation of one-year data by five-year data, observed on some specific cases, is up to 5%.

It was observed on the case of Slovenia, that errors caused by the approximation of one-year survival ratios, discussed in chapter 3.2, are greater for $i \geq 74$ than for $i \leq 73$. These errors can be reduced by using survival ratios of a specific country

Table 1: Population forecasts for some European countries in millions. The ratios with respect to the year 1990 are given in parenthesis. The years for which fertility and survival ratios are used are given in the second and third columns. Input data are taken from [6].

country$_{t_0}$	fert.	surv.	1990	2000	2100	2200
Austria$_{88}$	'87	'87	7.61	7.56 (0.99)	2.47 (0.32)	0.59 (0.08)
Belgium$_{84}$	'83	'84	9.93	9.89 (0.99)	4.14 (0.42)	4.40 (0.14)
CS$_{86}$	'86	'86	15.75	16.37 (1.04)	15.54 (0.99)	13.30 (0.84)
Denmark$_{87}$	'87	'87	5.13	5.09 (0.99)	1.90 (0.37)	0.57 (0.11)
Finland$_{86}$	'86	'86	4.98	5.01 (1.01)	2.32 (0.47)	0.90 (0.18)
France$_{88}$	'87	'87	56.21	58.12 (1.04)	40.70 (0.72)	24.27 (0.43)
GDR$_{88}$	'85	'85	16.72	16.74 (1.00)	9.36 (0.56)	4.56 (0.27)
GFR$_{86}$	'86	'86	60.95	59.74 (0.98)	16.13 (0.26)	3.41 (0.06)
Hungary$_{87}$	'87	'87	10.60	10.57 (0.99)	6.39 (0.60)	3.51 (0.33)
Iceland$_{84}$	'84	'84	0.26	0.28 (1.10)	0.33 (1.31)	0.33 (1.30)
Ireland$_{86}$	'85	'85	3.67	4.06 (1.11)	7.49 (2.04)	12.04 (3.28)
Italy$_{87}$	'82	'84	57.75	58.63 (1.02)	25.29 (0.44)	8.65 (0.15)
Norway$_{87}$	'87	'87	4.22	4.30 (1.02)	2.65 (0.63)	1.37 (0.33)
Poland$_{87}$	'87	'87	38.42	40.55 (1.06)	47.84 (1.25)	51.50 (1.34)
Portugal$_{87}$	'87	'87	10.36	10.63 (1.03)	4.88 (0.47)	1.62 (0.16)
Sweden$_{87}$	'87	'87	8.46	8.50 (1.01)	5.68 (0.67)	3.54 (0.42)
Switzerland$_{87}$	'87	'87	6.60	6.64 (1.01)	2.61 (0.40)	0.83 (0.13)
UK$_{87}$	'87	'87	57.46	58.52 (1.02)	39.49 (0.69)	23.27 (0.40)
Slovenia$_{89}$	'89	'85	1.99	2.00 (1.01)	0.80 (0.40)	0.25 (0.13)

that has the similar mortality as observed population. $P_{s,i}$ is transformed into $'P_{s,i} = P^*_{s,i} \cdot k_i \;,\; i = 74, ..., 85$, where $P^*_{s,i}$ are one- year survival ratios of a country with similar mortality and $k_i = \sum_{j=i-10}^{i-1} P_{s,i} / \sum_{j=i-10}^{i-1} P^*_{s,i}$.

4 Demonstrative model

We consider a hypothetical population consisting of male and female individuals. Their lives consist of three periods: childhood, fertility period of female individuals, and old age. An individual can survive or die in each life period, but none can survive the last life period. The natural growth of this population is described by the following equations:

$$N^{t+1}_{1,1} = P_{1,0} \cdot N^t_{1,0}, \quad N^{t+1}_{2,1} = P_{2,0} \cdot N^t_{2,0}, \quad N^{t+1}_{1,2} = P_{1,1} \cdot N^t_{1,1}, \quad N^{t+1}_{2,2} = P_{2,1} \cdot N^t_{2,1}, \tag{3}$$

$$N^{t+1}_{1,0} = \frac{1}{2} \cdot D_m \cdot P'_1 \cdot R \cdot (N^t_{2,0} + N^t_{2,1}), \quad N^{t+1}_{2,0} = \frac{1}{2} \cdot D_f \cdot P'_2 \cdot R \cdot (N^t_{2,0} + N^t_{2,1}).$$

$N^t_{s,i}$, s=1,2, i=0,1,2 is the number of male or female individuals belonging to the i^{th} life period in time t. $P_{s,i}$, s=1,2, i=0,1 are survival ratios for male and female individuals belonging to the i^{th} life period; they are probabilities that male/female individuals who belong to i^{th} life period will survive the end of this life period. P'_s, s=1,2 are survival ratios for male and female infant children. Fertility rate R is the average number of children given birth to by one female during her life. Ratios of live-born males and females are denoted by D_m and D_f. These two constants tell us that among 100 live-born individuals there are $D_m \cdot 100$ males and $D_f \cdot 100$ females. $N^t_{2,1}$ is number of females in fertility period and $N^t_{2,0}$ number of females in 'childhood' period. The oldest girls will begin to bear children in the beginning of the period and by the end of the period all of them will be in the fertility life period. The females that are in the fertility life period will gradually leave it and at the end all of them will be in the old age life period. So both groups will equally contribute to the number of newborns. Note that model (3) does not match with model (1) at this point.

From the second and the sixth equations of (3) a difference equation of order two for the number of female individuals in their fertility life period $N^t_{2,1}$ is obtained:

$$N^t_{2,1} - \frac{1}{2} \cdot D_f \cdot R \cdot P'_2 \cdot N^{t-1}_{2,1} - \frac{1}{2} \cdot D_f \cdot R \cdot P'_2 \cdot P_{2,0} \cdot N^{t-2}_{2,1} = 0. \tag{4}$$

Its solution is:

$$N^t_{2,1} = A_1 \cdot G^t_1 + A_2 \cdot G^t_2, \tag{5}$$

where:

$$G_{1,2} = \frac{1}{4} \cdot (D_f \cdot R \cdot P'_2 \pm \sqrt{D_f^2 \cdot R^2 \cdot P'^2_2 + 8 \cdot D_f \cdot R \cdot P'_2 \cdot P_{2,0}}), \tag{6}$$

$$A_1 = \frac{P_{2,0} \cdot N^0_{2,0} - N^0_{2,1} \cdot G_2}{G_1 - G_2}, \quad A_2 = \frac{N^0_{2,1} \cdot G_1 - P_{2,0} \cdot N^0_{2,0}}{G_1 - G_2}. \tag{7}$$

As D_f, R, P'_2 and $P_{2,0}$ are real and nonnegative, it follows from (6): $G_1, G_2 \in \mathcal{R}$; $G_1 \geq 0$, $G_2 \leq 0$; $| G_1 | \geq | G_2 |$. The value of $| G_1 |$ can be smaller or greater than 1, whereas $| G_2 | < 1$. The solution (5) of difference equation (4) is a linear combination of two exponential functions. The asymptotic behaviour (when $t \to \infty$) is determined by the first root G_1, which has a greater absolute value than root G_2. The second root alternately contributes positive and negative values to the solution's value, but as $| G_2 | < 1$, this contribution exponentially disappears with time. If $R = \frac{2}{D_f \cdot P'_2 \cdot (P_{2,0}+1)} = R_0$ then $G_1 = 1$, $G_2 = \frac{1-\sqrt{1+4 \cdot P_{2,0} \cdot (P_{2,0}+1)}}{2 \cdot (P_{2,0}+1)}$, and $\lim_{t\to\infty} N^t_{2,1} = A_1 = \frac{2 \cdot P_{2,0} \cdot (P_{2,0}+1) \cdot N^0_{2,0} - N^0_{2,1} \cdot (1-\sqrt{1+4 \cdot P_{2,0} \cdot (P_{2,0}+1)})}{2 \cdot P_{2,0}+1+\sqrt{1+4 \cdot P_{2,0} \cdot (P_{2,0}+1)}}$. If $R > R_0$, $N^t_{2,1}$ grows exponentially; when $R < R_0$, $\lim_{t\to\infty} N^t_{2,1} = 0$ (see figure 1). For the asymptotic behaviour of other variables see [7].

If $R = R_0$, model (3) will give stable, nonzero solutions. Which parameter in model (1) corresponds to R? It is total fertility rate given by $T_f = \sum_{i=15}^{49} R_i$. The value of T_f that according to the model (1) for 1989 fertility and 1985 survival ratios would lead neither to population explosion nor implosion, is 2.11 for Slovenia, while its actual value in 1990 was 1.4371. Let us take $D_f = 0.483$ (ratio of live-born girls in Slovenia) and $P_2' = 0.9919$ (survival ratio for female infant children in Slovenia). To get $P_{2,0}$ for Slovenia, model (3) is mapped on model (1) in the following way: 0-19 years is 'childhood' period, 20-39 is fertility period, 40-59 is old age. So $P_{2,0} =$ (number of women who will be 20-39 years old by projection of the annual model) / (the total number of women 0-19 years old in Slovenia 1990) = 0.9916. From these values we get $R_0 = 2.10$, which is close to 2.11. Figure 2 shows how the number of women in fertility period $N_{2,1}^t$, according to model (3), would fall in case total fertility rate remained unchanged, and how it would rise if Ireland's T_f was substituted for R. The behaviour of these solutions is very similar to the behaviour of the respective solutions for model (1).

Model (1) is elaborate and realistic, however, its solution cannot be expressed analytically in an explicit form. Model (3) is very simple and can be solved analytically. Its solutions demonstrate all the main features of long term behaviour of the solutions of model (1) and its parameters (especially in the expression for R_0) have meanings evident also to nonprofessionals. Both can show some temporary periodic behaviour and both tend to zero or grow without limits as $t \to \infty$. With an appropriate mapping of one model on another, an analogy or a correspondence between parameters of both models can be obtained. So model (3) can also be used for teaching purposes as an example of a very simple mathematical modeling of a demographic process.

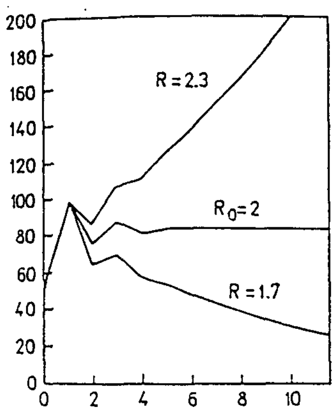

Figure 1: Solutions $N_{2,1}^t$ for $D_f = 0.5$, $P_2' = P_{2,0} = 1$ and $N_{2,0}^0 = 100$, $N_{2,1}^0 = 50$.

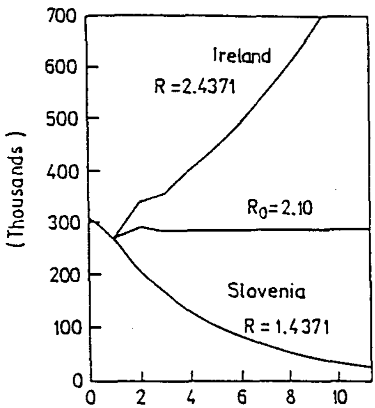

Figure 2: Solutions $N_{2,1}^t$ for $D_f = 0.483$, $P_2' = 0.9919$, $P_{2,0} = 0.9916$ and $N_{2,0}^0 = 272261$, $N_{2,1}^0 = 312756$.

Note. A software package DEMOGRAPHER based on the annual model is available. It is a menu-driven user-friendly software running on an IBM PC or compatible. Its structure is described in [8].

5 Conclusion

Tools for transformation of five year data to one year data and for evident presentation and school discussion of meaning of demographic parameters are described in the paper. Annual model allows us to show what present fertility and mortality really mean in terms of population size and age distribution after 50, 100 or more years. If we do not expect migrations and changes in fertility and mortality (which in most cases is an unrealistic expectation), the estimates obtained could be considered a realistic forecast. In any case, population projections give valuable information for the interpretation of fertility and survival ratios. Model (3) shows the influence of fertility rate, ratio between live-born boys and girls, and mortality of girls on the demographic growth. Population data in data-base of the programme DEMOGRAPHER make possible a quick comparison between different countries.

References

[1] Leslie P.H. (1945): The use of matrices in certain population mathematics. *Biometrika*, **33**, 183-212.

[2] Gosar L. (1980): *Demografske projekcije v prostorskem planiranju*. Ljubljana: Urbanistični institut.

[3] Čepar D., Gosar L., and Uršič E. (1992): Natural growth population projections. In: P. Gritzmann, R. Hettich, R. Horst, and E. Sachs (Eds): *Operations Research '91*, Heidelberg: Physica - Verlag, 305-308.

[4] Pielou E. C. (1977): *Mathematical Ecology*. New York: John Willey and Sons, Inc.

[5] Hoppensteadt F. C. (1983): Population mathematics. In: H. Marcus-Roberts and M. Thompson (Eds.): *Life Science Models*. New York: Springer Verlag.

[6] United Nations (1990): *Demographic Yearbook*.

[7] Čepar D., and Uršič E. (1993): Natural growth population models. In: A. Karmann, K. Mosler, M. Schader, and G. Uebe (Eds): *Operations Research '92*. Heidelberg: Physica - Verlag, 343-345.

[8] Čepar D., Habinc R., Uršič E., and Založnik R. (1993): Demographer. In: A. Karmann, K. Mosler, M. Schader, and G. Uebe (Eds): *Operations Research '92*. Heidelberg: Physica - Verlag, 339-342.

Acknowledgement

We thank Barbara Motnikar for her helpful comments.