PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

Pliska Stud. Math. Bulgar. **22** (2013), 207–224

PLISKA
STUDIA MATHEMATICA
BULGARICA

CRUMP-MODE-JAGERS BRANCHING PROCESS: MODELLING AND APPLICATION FOR HUMAN POPULATION

Plamen Trayanov*

Abstract. The future human population count in a country depends on many factors which influence birth and death. The interaction of birth and death determines the rate at which the population grows or diminishes. Modelling the population can give us information for the current condition of a country. The paper describes a methodology based on Crump-Mode-Jagers branching process theory (see [2]) for modelling human population and shows how the Malthusian parameter can be numerically estimated using the model. A population that has greater Malthusian parameter is expected to have greater population count from some point on in the future. The model results from comparison between Sweden, Greece, Slovenia and Bulgaria using official EUROSTAT data (see [1]) show the Malthusian parameter for Greece is declining for the past few years due to the crisis. Before that Greece was comparable to a country with good social and demographic policy like Sweden. The model General Branching Process (GBP) could also be used for population projection. The results for Bulgaria show the model expects decreasing population count.

Introduction. The paper introduces a population model based on General Branching Process Theory described in Jagers ([16]). The use of General Branching Process (GBP) assumes each woman gives birth in random intervals of time

*This work was supported by the European Social Fund through the Human Resource Development Operational Programme under contract BG051PO001-3.3.06-0052 (2012/2014).

2010 *Mathematics Subject Classification:* 60J85, 92D25

Key words: branching process, human population, Malthusian parameter

and her life length is random too. The births of the woman are modeled as a point process and her life length is a random variable. This way GBP takes us a step closer to the real process of birth and death in a population and allows us to derive some demographic parameters for the conditions in different countries. The GBP gives us the expected future population regardless of the generation number. But if we are interested in the number of women in different generations then the embedded Galton-Watson branching process is a good choice for a model (see Bojkova [3]).

The classic GBP starts from a newborn woman. In order to calculate the expected population given the current age structure it is imperative to use a branching process starting from a grown woman and use the current information for the population like birth and death probabilities that are relevant for the current year as it is well known that these can change in time. Assuming we know how these probabilities change in time or that a certain forecast of them will come true we can incorporate this forecast in the GBP model and get expected future population dynamics conditional on the forecast. Another useful result from the model is that we can derive the Malthusian parameter of the population conditional on the given year and country. This makes a general comparison between different populations possible and allows us to measure the general impact of a crisis to the population growth.

In this paper are compared the demographic conditions in several European countries and the results for Greece show that the crisis had severe impact not only to its economy but to its demographic condition as well. It is also curious to notice that before the crisis it was comparable to very well developed countries with traditional social policy like Sweden.

The structure of the paper is as follows. In section 1 is described the GBP and some demographic interpretations of mathematical terminology. In section 2.1 are presented the assumptions made for constructing the model, some corollaries that follow from them and several simple and yet useful propositions. In section 2.2 is presented methodology for modelling the survivability function and the expected birth count of a woman using penalized smoothing splines with constrictions. In section 2.3 is presented the general methodology for computing the expected contribution of a live birth to the population and the expected future population. In section 3 are presented results from the calculations made in R (see [19]). The demographic conditions in Sweden, Greece, Slovenia and Bulgaria are compared and discussed. The calculations show that Sweden has the best demographic conditions for year 2010. The GBP for Bulgaria was supercritical until 1981 and then changed to subcritical. The population however increased for several years after that because of young age distribution.

1. Preliminary Results.

1.1. General Branching Process (GBP). In this section the point process and the branching process are described from demographic point of view using the notations given in Jagers [16]. The branching process of Crump-Mode-Jagers is also called general branching process (GBP). The theorems and definitions related to GBP can be found in Jagers [16].

A point process ξ on R^+ is a map from some probability space into the set of integer or infinite valued measures on R^+ defined on the Borel $\sigma - algebra$, such that the mass $\xi(A)$ given to a bounded Borel set is a finite random variable (see Jagers [16]). For example A could be $[a, b)$. Let $\xi(t) = \xi[0, t)$ and $\mu(A) = \mathbb{E}(\xi(A))$. Then μ is also a measure on R^+ but not random nor integer.

Let I be the set of all n-tuples for all n. It is an index set for all individuals of the population ever born, not yet born or not born ever at all. We assume in this section that the GBP started from a single individual denoted with (0). The woman (0) is assumed to have age 0 and her birth is at time $t = 0$. The n-th child of $x \in I$ is denoted by (x, n). Though the set I is infinite with cardinality of continuum, the set of the actual individuals ever lived is quite finite for human population. We say the individual $(x, n) \in I$ is realized if $\xi_x(\infty) \geq n$.

Let λ_x be the life length of the individual x and be a random variable. Let ξ_x be a point process describing the births of woman x. For each $x \in I$ is defined a couple (λ_x, ξ_x) and these couples are assumed to be identically and independently distributed. This means the distributions of ξ and λ are time invariant. This quality of a population is referred to as stationary population in this article.

Let $\tau_x(k) = \inf\{t : \xi_x(t) \geq k\}$ is the age of birth of child (x, k). It's allowed to be even infinity if the child is never born. Let $\sigma_x = \tau_0(j_1) + \tau_{j_1}(j_2) + ... + \tau_{(j_1,...,j_{n-1})}(j_n)$, where $x = (j_1, ..., j_n)$. This means σ_x is the birth date of x and can be infinity too. In particular $\sigma_0 = 0$.

An indicator variable $z_t^a(x)$ is defined for the individual x to be alive and younger than $a > 0$ at time $t > 0$ as follows

$$z_t^a(x) = \begin{cases} 1, & \text{when } t - a < \sigma_x \leq t < \sigma_x + \lambda_x \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The GBP is defined as

$$z_t^a = \sum_{x \in I} z_t^a(x). \tag{2}$$

Let $z_t = z_t^a$, for all $a > t$. At time t the oldest individual is of age less than

t, so the definition is correct. So z_t is the total number of individuals in the population.

Theorem 1 (see Jagers [16]). *If $\mu(0) < 1$ and $\mu(t) < \infty$ for some t, then $\mathbb{P}(z_t < \infty, \forall t) = 1$.*

Let us denote $f(s) = \mathbb{E}(s^{\xi(\infty)})$, $|s| \leq 1$, $L(t) = \mathbb{P}(\lambda_x \leq t)$, $\hat{\mu}$ is the Laplace-Stieltjes transformation of μ and $S(t) = 1 - L(t)$.

Theorem 2 (see Jagers [16]). *If $f(s) < \infty, |s| \leq 1$, then $m_t = \mathbb{E}(z_t) < \infty, \forall t$ and $m_t^a = \mathbb{E}(z_t^a)$ satisfies*

$$m_t^a = 1_{[0,a)}(t)\{1 - L(t)\} + \int_0^t m_{t-u}^a \mu(du). \tag{3}$$

If $m = \mu(\infty) < 1$, then $\lim_{t \to \infty} m_t = 0$. If $m = 1$ and μ is non-lattice, then

$$m_t^a \to \frac{\int_0^a \{1 - L(u)\}du}{\int_0^\infty u\mu(du)}.$$

If further $\int_0^\infty tL(dt) < \infty$, then

$$m_t^a \to \frac{\int_0^\infty uL(du)}{\int_0^\infty u\mu(du)}.$$

If $m > 1$, μ is non-lattice and $\alpha > 0$ is the Malthusian parameter defined by $\hat{\mu}(\alpha) = 1$, then for $0 \leq a \leq \infty$

$$m_t^a \sim e^{\alpha t} \frac{\int_0^a e^{-\alpha u}\{1 - L(u)\}du}{\int_0^\infty ue^{-\alpha u}\mu(du)}.$$

In the lattice cases corresponding assertions hold.

In case of human population the conditions of the theorem hold. The mean population m_t is the solution of renewal equation (3). Due to the fact that women can give birth only to finite age the renewal function of this equation is finite. The GBP gives us the expected population at time t regardless of the generation number. If we are interested in the number of women in different generations then the embedded Galton-Watson branching process is a good choice for a model (see Bojkova [3]).

In case of subcritical process we have that $m_t \to 0$ exponentially.

Theorem 3 (see Jagers [16]). *In a non-lattice, subcritical process admitting Malthusian parameter $\alpha < 0$*

$$m_t^a \sim e^{\alpha t} \int_0^a e^{-\alpha t}\{1 - L(t)\}dt / \int_0^\infty t e^{-\alpha t}\mu(dt) \tag{4}$$

for $0 \le a < \infty$, as $t \to \infty$. For $a = \infty$ the relation still holds, provided

$$\int_0^\infty t e^{-\alpha t} L(dt) < \infty. \tag{5}$$

This means m_t^a converges towards zero exponentially. If we have two different branching processes with two different Malthusian parameters then one of them is bound to have greater expected population from one moment on. If two countries are considered then the expected contribution of a live birth in the country with the highest Malthusian parameter will exceed the expected contribution in the other country from one moment on.

1.2. Demographic interpretation of some terms in GBP theory. The point process $\xi(t)$ is the number of children a woman has during her life until age t. Each woman realizes one point process in her life, i.e. makes a sequence of births on particular dates. In the beginning of her life she has no idea of how many children she will have or on what dates and depending on her personal development and many other factors she may or may not make a plan. The fact is however that all women plan (or do not plan) differently so the point process is stochastic in nature, i.e. we can think of $\xi(t)$ as a random variable. The number of children a woman has in the age interval $[a, b]$ is denoted by $\xi[a, b]$ and is a random variable too. The expected number of births is then $\mu[a, b] = \mathbb{E}(\xi[a, b])$ and obviously is not integer. $\mu[a, b]$ is the mean number of children a woman has in this age interval. For human population it is appropriate to consider such point processes $\xi(t)$ that $\mu(t)$ is smooth function. Another property of the human population is that children can't give birth so $\xi(t) = 0$ and $\mu(t) = 0$ when t is less than the lower bound on the fertility interval. In our case we can assume the lower bound is 12 years as the cases of giving birth on lesser ages are quite rare.

The life length of a woman is also varying depending on many factors and can be thought of as stochastic in nature. λ represents the exact age of death and can be modelled as random variable. For human population it is appropriate to model the distribution $L(t) = 1 - S(t)$ of λ as smooth function. $S(t)$ is called survivability function of a live birth so we can assume $S(0) = 1$ and $S(\omega) = 0$, where ω is the oldest age in a life table.

2. Results. In this section is proposed an approach for modelling the GBP for human population. The index x is skipped for more simple notation.

2.1. Assumptions and corollaries. In the proposed model only women are considered because they define the branching process. To make a simple and yet useful model the following assumptions are made:

1) The fertility interval for each woman is $[12, 50]$ and women can't give birth outside it or if they aren't alive. In terms of GBP

$$\mathbb{P}(\xi\,[a,b) = 0)) = 1, when\,[a,b) \cap [12,50] = \varnothing$$

$$\mathbb{P}(\xi[\lambda,\infty) = 0) = 1.$$

2) A woman could have 0 or 1 daughter during a year and each birth is a live birth. This means the number of live births is equal to the number of women who gave birth.

3) It is assumed there is no migration.

For human population $\xi(12) = 0$ because the children don't give birth so $\mu(12) = 0$ and Theorem 1 holds. It tells us something intuitive - the human population is always finite. In the following propositions are described the implications of these assumptions on the GBP model.

Let ${}_bz_t$ is the branching process started from a woman aged b at time $t = 0$, ${}_b\xi$ to be her point process, ${}_b\mu$ to be the expectation of the point process and ${}_bS$ to be her survivability function. Let $n_b = \mathbb{P}(\xi\,[b, b+1) = 1|\lambda \geq b)$ be the probability a woman to give birth at age b if she survived to the beginning of this age interval.

Proposition 1. *For $k \geq 1$ the distribution of ${}_b\xi$ satisfies*

$$\mathbb{P}({}_b\xi[b+k-1, b+k) = 1) = 1 - \mathbb{P}({}_b\xi[b+k-1, b+k) = 0) = {}_bS(b+k-1) \cdot n_{b+k-1}$$

and the expected number of births in $[b+k-1, b+k)$ *of a woman aged* b *is*

$$ {}_b\mu[b+k-1, b+k) = {}_bS(b+k-1) \cdot n_{b+k-1} $$

Proof. If a woman survived to age b, then her point process ${}_b\xi$ satisfies

$$ \mathbb{P}({}_b\xi[b+k-1, b+k) = 0) = \mathbb{P}(\xi[b+k-1, b+k) = 0 | \lambda \geq b) $$
$$ = \mathbb{P}(\lambda < b+k-1 | \lambda \geq b) $$
$$ + \mathbb{P}(\lambda \geq b+k-1 | \lambda \geq b)\mathbb{P}(\xi[b+k-1, b+k) = 0 | \lambda \geq b+k-1) $$
$$ = 1 - \frac{S(b+k-1)}{S(b)} + \frac{S(b+k-1)}{S(b)}(1 - n_{b+k-1}) $$
$$ = 1 - \frac{S(b+k-1)}{S(b)} n_{b+k-1} = 1 - {}_bS(k-1) \cdot n_{b+k-1}, $$

where ${}_bS(k-1)$ is the conditional probability of a woman to survive to age $b+k-1$ if she survived to age b.

By assumption 2) it follows

$$ {}_b\mu[b+k-1, b+k) = \mathbb{E}({}_b\xi[b+k-1, b+k)) $$
$$ = \mathbb{P}(\xi[b+k-1, b+k) = 1 | \lambda \geq b) $$
$$ = {}_bS(b+k-1) \cdot n_{b+k-1} $$

□

When $k = 1$ from proposition 1 we have

$$ \mu\,[b, b+1) = \mathbb{P}(\xi\,[b, b+1) = 1) = n_b \cdot S(b). \tag{6} $$

Proposition 2. *The expected population count at time* t *started from woman aged* b *at time zero is given by the following equation*

$$ {}_bm_t = {}_bS(t) + \int_0^t m_{t-u}\ {}_b\mu(b+du), \tag{7} $$

where ${}_bS(t)$ *denotes the probability a woman of age* b *to survive to* $b+t$*, i.e.*

$$ {}_bS(t) = \frac{S(b+t)}{S(b)}. \tag{8} $$

Proof. Let ${}_bz_t$ be the population at time t started from a woman aged b, z_t be the population at time t started from a woman aged 0. Let $z^{(u)}_{t-u}$ denotes an instance of the branching process that started at time u, so its descendats at time t are $z^{(u)}_{t-u}$. From the assumptions follows that each woman could give no more than one birth in each age interval $[b, b+1)$ for $b = 12, 13, ..., 50$.

$$ {}_bz_t = {}_bz_t(0) + \sum_{k=1}^{n} {}_b\xi\,[b+k-1; b+k)\, z^{(t-t_k)}_{t_k}, \tag{9} $$

where $n = \lceil t-b \rceil$, ${}_b\xi$ is a point process for woman aged b and $t_k \in (t-k, t-k+1)$ is a random variable the time passed from the exact moment of giving birth during the year if it happened. If such event never happened then ${}_b\xi\,[b+k-1; b+k) = 0$ and $z^{(t-t_k)}_{t_k} < +\infty, \forall k$ so it follows ${}_b\xi\,[b+k-1; b+k)\, z^{(t-t_k)}_{t_k} = 0$ and the sum is correctly defined.

Equation (9) can be written as

$$ {}_bz_t = {}_bz_t(0) + \sum_{k=1}^{n} z^{(t-t_k)}_{t_k} [{}_b\xi(b+k) - {}_b\xi(b+k-1)]. \tag{10} $$

$\xi(du)$ is 1 in no more than one point of $[b+k-1; b+k)$, so each realization of the random sum (10) can be written as a Stieltjes integral. This gives us the stochastic integral form of equation (10)

$$ {}_bz_t = {}_bz_t(0) + \int_0^t z^{(u)}_{t-u}\; {}_b\xi(b+du). \tag{11} $$

The integral is actually a finite sum of random variables and the expected population, started from a woman on age b, can be obtained using the linearity of expectation:

$$ {}_bm_t = {}_bS(t) + \int_0^t m_{t-u}\; {}_b\mu(b+du), $$

where ${}_bS(t)$ denotes the probability a woman of age b to survive to age $b+t$, i.e.

$$ {}_bS(t) = \frac{S(b+t)}{S(b)}. \qquad \square $$

Let $P[b, b+1)$ is the number of women of age b by last birthday at time $t = 0$. Let the exact ages be random variables η_i. They can be assumed uniformly distributed because the uniform distribution is a local approximation of any absolutely continuous distribution. Then $\mathbb{E}(\eta_i) = b + 0.5$.

The population at time zero can be thought of as a set of individuals (ancestors) who start or don't start new branching processes. The number of individuals at time t, derived from an ancestor x is referred to as a contribution of the ancestor to the total population at time t.

Proposition 3. *Let the births in the age interval* $[b, b+1]$ *are uniformly distributed. The expected contribution* ${}_bC(t)$ *after time* t *of all ancestors of age* b *by last birthday at time* $t = 0$ *is then given by*

$$ {}_bC(t) = P[b, b+1) \int_b^{b+1} {}_u m_t du $$

P r o o f. Let the exact ages of birth are random variables η_i. They are assumed uniformly distributed in the proposition. Then

$$ \begin{aligned} {}_bC(t) &= \mathbb{E}\left(\sum_{i=1}^{P[b,b+1)} {}_{\eta_i}z_t^{(i)}\right) = \sum_{i=1}^{P[b,b+1)} \mathbb{E}\left({}_{\eta_i}z_t^{(i)}\right) \\ &= \sum_{i=1}^{P[b,b+1)} \mathbb{E}(\mathbb{E}({}_{\eta_i}z_t^{(i)}|\eta_i)) = \sum_{i=1}^{P[b,b+1)} \mathbb{E}\left({}_{\eta_i}m_t\right) \qquad (12) \\ &= \sum_{i=1}^{P[b,b+1)} \int_b^{b+1} {}_u m_t du = P[b, b+1) \int_b^{b+1} {}_u m_t du \end{aligned} $$

□

To obtain approximation of the last equation an assumption for smoothness of $\mu(t)$ is made. The women of age less than 12 don't give birth at all, then they begin to give birth but only a small number of them (the probability is rising slowly from zero). At later age the probability of birth slowly subsides to 0 at age 50 so the model of continuous or even a smooth $\mu(t)$ is adequate. So a possible approximation of the result in (12) is:

$$ \mathbb{E}\left(\sum_{i=1}^{P[b,b+1)} {}_{\eta_i}z_t^{(i)}\right) \approx P[b, b+1)\ {}_{b+0.5}m_t. \qquad (13) $$

In other words each ancestor's age in $[b, b+1)$ can be approximated by $b+0.5$. The total expected population $\mathbb{E}(P_t)$ at time t is the sum of the expected contributions of all individuals in the population. It is then given by

$$\mathbb{E}(P_t) \approx \sum_{b=0}^{\omega} P[b, b+1)_{b+0.5} m_t, \tag{14}$$

where ω denotes the oldest age in the life table.

Finding solution ${}_bm_t$ for equation (7) can be reduced to finding solution m_t for equation (3).

Proposition 4. *If m_t has a continuous second derivative then a third order approximation of equation (7) is given by*

$${}_bm_t \approx {}_bS(t) + \sum_{k=1}^{n} m_{b+k-0.5} \cdot {}_b\mu(b+k-1, b+k).$$

Proof. Consider equation (10) with expectations on both sides. (λ_x, ξ_x) are independent for different x so $z_{t_k}^{(t-t_k)}$ and ${}_b\xi[b+k-1, b+k)$ are independent too. Then

$${}_bm_t = {}_bS(t) + \sum_{k=1}^{n} \mathbb{E}(m_{t_k})\ {}_b\mu(b+k-1, b+k)$$

$$= {}_bS(t) + \sum_{k=1}^{n} (\int_{b+k-1}^{b+k} m_u du)\ {}_b\mu(b+k-1, b+k)$$

The birth times t_k can be approximated with the middle of the years because m_t has a continuous second derivative. It is a third order approximation of the integral with elementary trapezoidal rule. We have that $n < \omega$ and ${}_b\mu(b+k-1, b+k) < 1$, for all b and k, so the total error is less than $\omega \cdot 1 \cdot O(h^3)$, where h is notation for the length of integration interval. This gives us order of approximation $O(h^3)$.

$${}_bm_t = {}_bS(t) + \sum_{k=1}^{n} m_{b+k-0.5} \cdot {}_b\mu(b+k-1, b+k). \tag{15}$$

Equation (15) gives us ${}_bm_t$ in terms of m_t. □

${}_bS(t)$ and ${}_b\mu(t)$ are expressed in terms of $S(t)$ and n_k in equations (8) and (6). So in order to solve (15) models for $S(t)$ and n_k must be found and then a solution to equation (3).

2.2. Modelling $S(t)$ and $\mu(t)$. Methodology. Consider a stationary population. Then the period life table probabilities of birth and death will coincide with the cohort probabilities, because the last don't change in time. This means the period life table from the last available year can be used to model the functions $S(t)$ and $\mu(t)$ and the derived population projection will be a good local forecast (assuming there is no migration). By definition

$$n_b = \mathbb{P}(\xi[b, b+1) = 1 | \lambda > b). \tag{16}$$

This is the probability a woman who survived to age b to give birth in the age interval $[b, b+1)$ and can be statistically evaluated using assumption 2). Let $N[b, b+1)$ be the number of live births of women on age b by last birthday during the year. As a corollary of assumption 2) the probability of live birth during the year is equal to the probability of a woman to give birth during the year. Then the age-specific fertility rate is

$$ASFR_b = n_b / e_b, \tag{17}$$

where e_b is the expected number of years lived in $[b, b+1)$ by a woman. So the ASFR can be approximated, according to the law of large numbers (see Yanev [2]), as

$$ASFR_b = \frac{N[b, b+1)}{E[b, b+1)}, \tag{18}$$

where $E[b, b+1)$ is the total number of years lived by all women aged b during the year and is called population at risk (see [14]). There are methods for computing $E[b, b+1)$ described in Mode [5] and HMD [14] so from equations (17) and (18) the empirical n_b is

$$\hat{n}_b = \hat{ASFR}_b \, \hat{e}_b. \tag{19}$$

The empirical $\hat{p}(b)$, which are the conditional probabilities of surviving, can be calculated as described in Mode [5] and HMD [14].

The empirical evaluations are distorted by noise and it is more appropriate to remove the noise and find the model function n_b and p_b. One possible method of obtaining model functions is smoothing splines with constrictions (see Ramsay [9]).

From the model for p_b can be computed the survivability function $S(t)$ in points $b = 0, 1, 2, \dots, 100+$ and interpolated by spline. From the smoothness of $\mu(t)$ follows the approximation

$$\hat{\mu}'(b + 0.5) \approx \hat{\mu}[b, b+1) = \hat{n}_b \, \hat{S}(b). \tag{20}$$

Again with smoothing splines can be obtained a model function $\mu'(t)$ for all $t > 0$ and as a corollary $\mu(t)$.

2.3. Contribution of a live birth. Expected future population. The Malthusian parameter of a population is closely related to the contribution of a live birth and can be numerically calculated as follows. From Theorem 3

$$\frac{m_{t+h}}{m_t} \sim e^{\alpha h}, t \to \infty$$

which is the same as

$$\frac{log(m_{t+h}) - log(m_t)}{h} \to \alpha, t \to \infty. \tag{21}$$

So α is approximately $(log(m_t))'$ for large t assuming m_t is smooth function. The estimation of the Malthusian parameter is shown in Figure 7.

The Malthusian parameter could be used for comparison between two countries. A country with greater Malthusian parameter has greater expected contribution of a live birth. The expected contribution of a live birth is a measure for the demographic condition of a population.

Using the modelled $S(t)$ and $\mu(t)$ we can find approximate solution to equation (7) by Proposition 4. Solving it for all ages $b = 0, 1, ...\omega$ and substituting in equation (14) gives the total expected contribution of all women in the population in time t or in other words the expected future population in time t (assuming there is no migration). If the contribution of a live birth for a country is greater than in another, the expected future population of the first country is bound to be asymptotically greater.

The period age-specific fertility and mortality coefficients and population at risk are calculated using the methodology of HMD [14]. The Kannisto model of old-age mortality is used for missing mortality data. The survivor ratio method is used for missing population count data. The empiric conditional death probabilities are calculated (see [5], [14]) and smoothed using penalized smoothing spline. The empiric n_b is calculated as discussed above and smoothed using penalized smoothing spline with constriction for positivity. The resulting model functions for $S(t)$ and $\mu(t)$ are used in equations (3) and (15) for the expected contribution to the population. The solution to the renewal equation is numerically calculated. The projected female population is then the sum of all expected contributions of women in the current population.

3. Application The calculations made in this paper are based on data for Sweden, Greece, Slovenia and Bulgaria found in Eurostat Database [4]. There are data for female population count by age, number of deaths by Lexis triangles and number of births by last birthday and by birth order. The life tables are complete and open-ended. The software used is R [19] more specifically the demography [17] and gam [18] package. The figures are presented in Appendix A. The results of smoothing conditional probabilities are shown in Figures 1 and 2. The resulting expected future population of Bulgaria is shown in Figure 5.

A comparison of the expected contributions (EC) of a live birth in different countries is shown on Figure 3. The country with the asymptotically highest EC has the greatest Malthusian parameter and the best demographic condition in that year. The expected contributions interleave but tend to differ for large period of time. They have different Malthusian parameters and because of Theorem 3 they have different rates of convergence toward zero. This means we can be sure they don't interleave again after some time. The visual representation shows that for year 2010 Sweden has the best demographic condition, then Greece, Slovenia and finally Bulgaria. A comparison between the Malthusian parameters for these countries confirms the visual results (see Figure 4). A comparison between histories of this parameter for different countries shows that all four of them have similar tendencies during the years. It shows that the branching process for Bulgaria was mainly supercritical for years 1960-1981, critical around 1981 and subcritical after that. We can also see the demographic conditions in 1960 were better than these in 2009 for Bulgaria (see Figure 6). A change to subcriticality means that after some time the population count will surely decrease due to probability of extinction equal to 100% (assuming the process stays subcritical in time). The change to decreasing population follows after a change to subcriticality and may be slowed down by good age distribution with more young women as is the case of Bulgaria. This is one of the reasons why better long term contribution means better demographic condition.

Appendix A

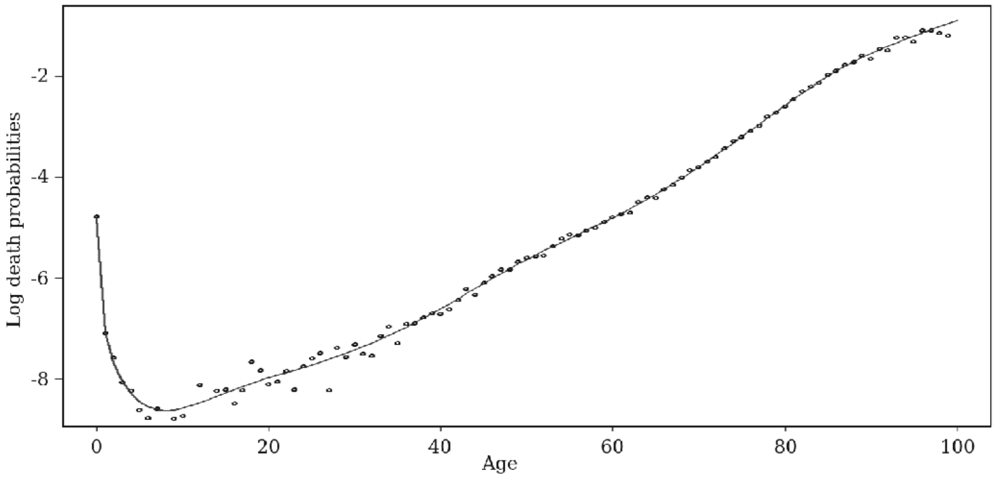

Fig. 1. Smoothed conditional probability of dying (qx) for Bulgaria (derived from data for year 2009)

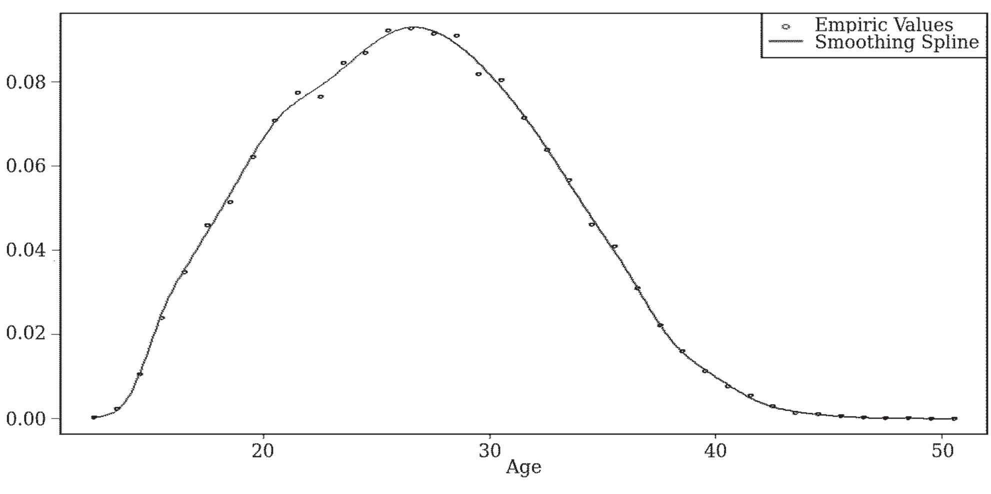

Fig. 2. Spline model of $\mu'(t)$ for Bulgaria (derived from data for year 2009)

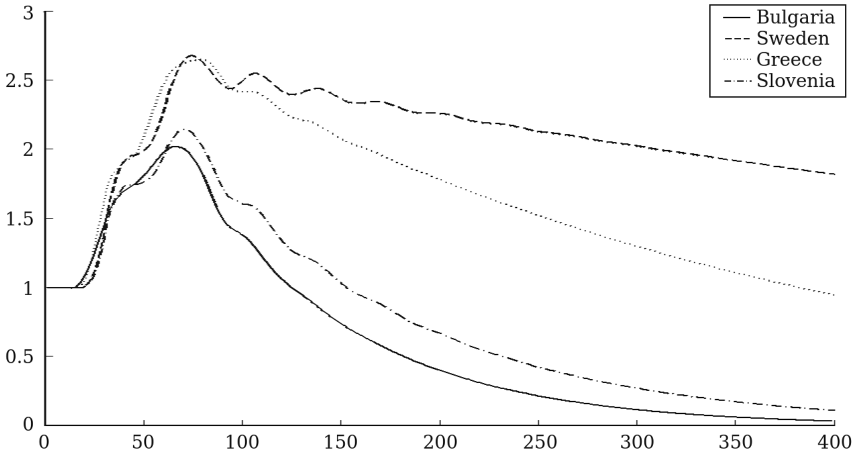

Fig. 3. Expected contributions of a live birth in different countries

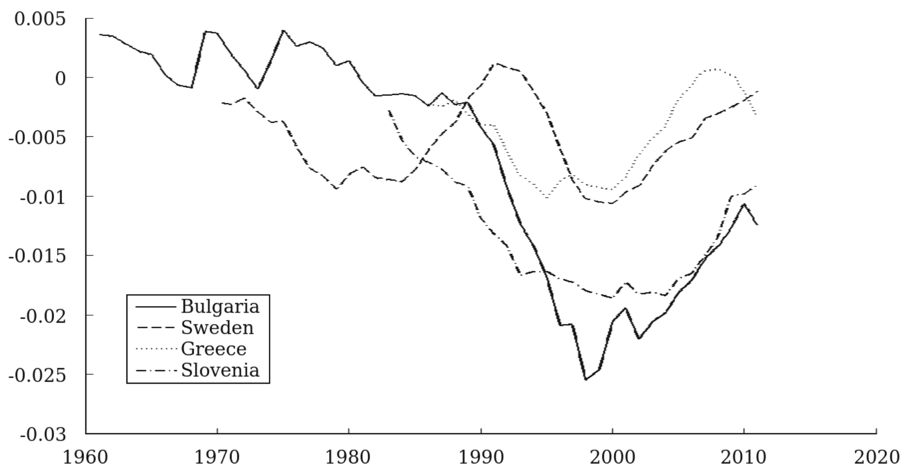

Fig. 4. Malthusian parameter in different countries through time

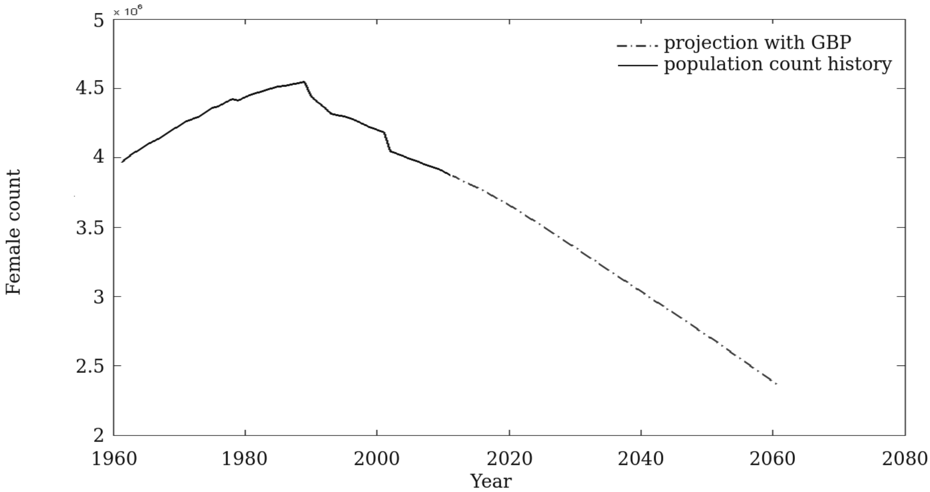

Fig. 5. Expected future female population of Bulgaria

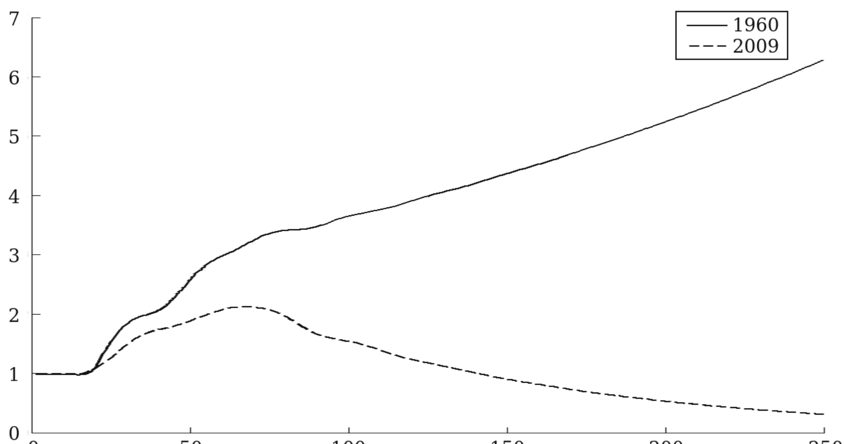

Fig. 6. Expected contribution of a live birth (Bulgaria)

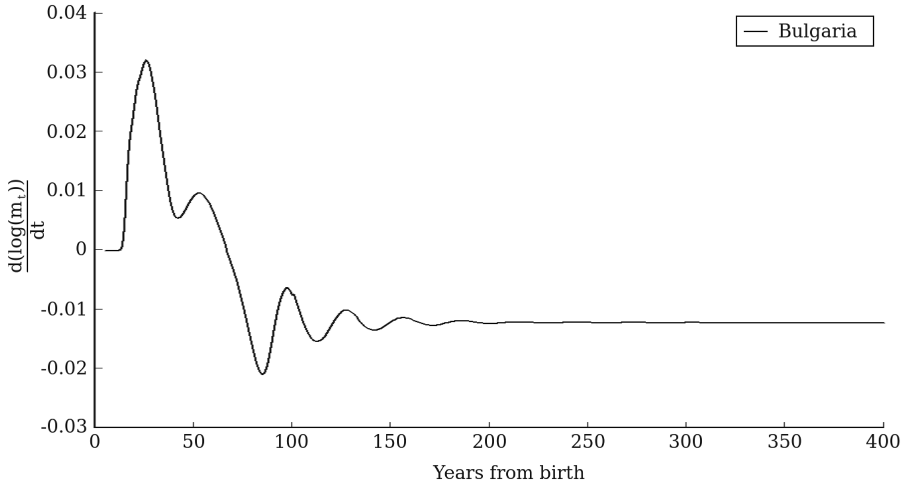

Fig. 7. The Maltusian parameter α is the asymptotic limit of $log'(m_t)$ (derived from data for Bulgaria 2010).

REFERENCES

[1] W. Feller. An introduction to probability theory an its applications, vol. **2**, Wiley, 1971.

[2] B. N. Dimitrov, N. M. Yanev. Probability Theory and Statistics. Sofia, St. Kliment Ohridski University press, 1990.

[3] M. Slavchova-Bojkova, N. M. Yanev. Branching Processes. Sofia, St. Kliment Ohridski University press, 2007.

[4] Eurostat Database. `http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database`.

[5] C. J. Mode. Stohastic Processes In Demography and Their Computer Implementation. Springer, 1985.

[6] C. D. Boor. A Practical Guide to Splines. Springer, 2001.

[7] N. Keyfitz, H. Caswell. Applied Mathematical Demography. Springer, 2005.

[8] C. L. Chiang. Life Table and Its Applications. Krieger Pub Co, 1983.

[9] J. O. Ramsay, B. W. Silverman. Functional Data Analysis. Springer, 2005.

[10] J. M. Alho, B. D. Spencer. Statistical Demography and Forecasting. Springer, 2005.

[11] T. Hastie, R. Tibshirani, J. H. Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, 2009.

[12] G. E. P. Box, G. M. Jenkins, G. C. Reinsel. Time Series Analysis: Forecasting and Control. Prentice-Hall, Inc., 1994.

[13] V. M. Shkolnikov. Methodology Note on the Human Life-Table Database (HLD). Unpublished Manuscript, http://www.lifetable.de

[14] University of California, Berkley, Max Planck Institute for Demographic Research. Human Mortality Database. http://www.mortality.org

[15] P. H. C. Eilers, P. O. Box, B. D. Marx. Splines, Knots and Penalties. Unpublished Manuscript, http://www.stat.lsu.edu/faculty/marx/splines_knots_penalties.pdf 2004.

[16] P. Jagers. Branching Processes with Biological Applications. John Wiley & Sons Ltd, 1975.

[17] R. J. Hyndman, H. Booth, L. Tickle, J. Maindonald. demography: Forecasting mortality, fertility, migration and population data. R package version 1.09-1, http://CRAN.R-project.org/package=demography, 2011.

[18] T. Hastie. gam: Generalized Additive Models. R package version 1.09-1, http://CRAN.R-project.org/package=gam, 2011.

[19] R Development Core Team. R: A Language and Environment for Statistical Computing. http://www.R-project.org/, 2011

Plamen Ivaylov Trayanov
ul. Yanchovska reka 22, kv. Bistritsa
Sofia, Bulgaria
e-mail: plamentrayanov@gmail.com