Provided for non-commercial research and educational use.
Not for reproduction, distribution or commercial use.

PLISKA
STUDIA MATHEMATICA BULGARICA

ПЛИСКА
БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

The attached copy is furnished for non-commercial research and education use only.
Authors are permitted to post this version of the article to their personal websites or institutional repositories and to share with other researchers in the form of electronic reprints.
Other uses, including reproduction and distribution, or selling or licensing copies, or posting to third party websites are prohibited.

For further information on
Pliska Studia Mathematica Bulgarica
visit the website of the journal http://www.math.bas.bg/~pliska/
or contact: Editorial Office
Pliska Studia Mathematica Bulgarica
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: pliska@math.bas.bg

Pliska Stud. Math. Bulgar. **17** (2005), 295–312

PLISKA
STUDIA MATHEMATICA
BULGARICA

PARAMETRIC ESTIMATION IN BRANCHING PROCESSES WITH AN INCREASING RANDOM NUMBER OF ANCESTORS

Vessela Stoimenova[1], Nickolay Yanev[2]

The paper deals with a parametric estimation in branching processes $\{Z_t(n)\}$ having random number of ancestors $Z_0(n)$ as both n and t tend to infinity (and thus $Z_0(n)$ in some sense). The offspring distribution is considered to belong to a discrete analogue of the exponential family - the class of the power series offspring distributions. Consistency and asymptotic normality of the estimators are obtained for all values of the offspring mean m, $0 < m < \infty$, in the subcritical, critical and supercritical case.

1. Introduction

Assume there exists on some probability space a set of i.i.d. r.v. $\{\xi_i(t,n)\}$ with values in the set of nonnegative integers $N = \{0,1,2,\cdots\}$ and $\{\xi_i(t,n),\ i \in N\}$ are independent of $Z_0(n)$. Then for each $n = 1,2,\cdots,\ \ Z(n) = \{Z_t(n), t = 0,1,\cdots\}$ is a Bienayme-Galton-Watson process having a random number of ancestors $Z_0(n) \geq 1$, where

$$(1)\qquad Z_t(n) = \begin{cases} \sum\limits_{i=1}^{Z_{t-1}(n)} \xi_i(t,n) & if \quad Z_{t-1}(n) > 0,\ t = 1,2,\cdots \\ 0, & otherwise. \end{cases}$$

[1]Partialy supported by Pro-ENBIS GTC1 -2001-43031.

[2]The paper is supported by the National Science Fund of Bulgaria, Grant No. MM-1101/2001.

2000 *Mathematics Subject Classification:* 60J80,62M05

Key words: Branching processes, random number of ancestors, power series distribution, paramertic estimation, consistency, asymptotic normality, efficiency.

Such a process will be denoted BGWR.

Let $\{p_k\}$ be the common offspring distribution, i.e.

$$p_k = P(\xi = k) \geq 0, \ \sum p_k = 1, \ p_0 + p_1 < 1$$

and put

$$m = E\xi, \quad \sigma^2 = Var(\xi). \tag{2}$$

We will assume throughout that $0 < \sigma^2 < \infty$.

The individual distribution is said to belong to the class of power series offspring distributions (PSOD) if

$$p_k = P(\xi = k) = \frac{a_k \theta^k}{A(\theta)}, \quad \theta > 0, \ a_k \geq 0, \tag{3}$$

where $A(\theta) = \sum_{k=0}^{\infty} a_k \theta^k$ is a positive function.

Our main concern in this paper is the parametric estimation of a BGWR process with power series offspring distribution based on a sample $\{Z_0(n), \cdots, Z_t(n)\}$ as both n and t tend to infinity (and thus $Z_0(n)$ in some sense). Naturally the relative speed, at which n and $t \to \infty$, will come into play for $0 < m < \infty$.

In contrast to the parametric estimation, the nonparametric estimation of m and σ^2 of a BGWR process is thoroughly considered in several works.

Yakovlev and Yanev (1989) noted that branching processes with a large and often random number of ancestors occur naturally in the study of cell proliferation. Such is also the case in applications to nuclear chain reactions.

Dion and Yanev (1994) suggest two other motivations for the study of a random number of ancestors:

First, let $\{X_t\}$, $X_0 = 1$ a.s. be a classical Bienayme-Galton Watson process (or BGW process) with $p_0 = 0$, $m > 1$, $0 < \sigma^2 < \infty$. By defining $Z_0(n) = X_n, \cdots, Z_t(n) = X_{n+t}$, $t \in N$, one obtains a BGW process $Z_t(n)$ for each $n \geq 1$, such that $Z_0(n) \to \infty$ a.s.. Moreover, in this case $Z_0(n)/m^n \to W > 0$ in L_2 and a.s..

Another motivation, presented in this work, comes from the branching process with immigration (or BGWI). Let $\{X_t\}$ be a BGWI process, $\{Y_t\}$ be i.i.d. r.v. of the independent immigration component added to each generation. Let the process start with the first nonzero Y_t, labelled Y_0. One can consider the tree underlying the $\{X_t\}$ and denote by $Z_t(n)$ the number of individuals among generations $t, t+1, \cdots, t+n$, whose ancestors immigrated exactly t generations ago. That way $Z_0(n) = \sum_{j=0}^{n} Y_j$ is the total number of immigrants from time 0 to time n; $Z_1(n)$ is the total number of their offspring etc. Hence $\{Z_t(n)\}$ is a BGWR process.

Some asymptotic features of the BGWI process may be captured by the BGWR process with an appropriate choice of $n = n(t)$. However, one notable exception is the estimation of $\lambda = EY_t < \infty$. Wei and Winnicki (1990) showed that in the supercritical BGWI process $\{X_t\}$ λ can not be estimated consistently on the basis of $X_0, \cdots, X_{t+n}$ alone. With respect to the BGWR process $\{Z_t(n)\}$ not only $Z_0(n)/n \to \lambda$ a.s., but also $Z_t(n)/nm^t \to \lambda$ a.s. as $n, t \to \infty$. Thus one can obtain an estimator even if $Z_0(n)$ is unobserved, see for instance Dion and Yanev (1992), Dion (1993). In general, the knowledge of $Z_0(n), \cdots, Z_t(n)$ would seem to be asymptotically equivalent to $\{\{X_k\}_0^{t+n}, \sum\limits_{k=0}^{n} Y_k\}$ as $n, t \to \infty$ on the set of nonextinction.

Results about the nonparametric estimation of the offspring mean m and variance σ^2 in the BGWR process have been announced in Dion and Yanev (1991),(1994), (1997) and Dion(1993),where the nonparametric m.l.e. and a family of l.s.e. for σ^2 are concerned and consistency and asymptotic normality of these estimators are obtained for all values of the mean m, $0 < m < \infty$. Results about the parametric and nonparametric robust estimation in BGWR processes are presented in Stoimenova, Atanasov, Yanev (2004), (2005).

The present work is divided into two parts: in Section 2 the parametric maximum likelihood is introduced and new results are presented and in Section 3 the efficiency properties of the maximum likelihood estimators are considered.

2. Parametric estimation in BGWR processes

Further on we will suppose that $n = n(t) \to \infty$ as $t \to \infty$ and will use the following

Condition A: $m > 1$ or $m = 1$, $t/n \to 0$ or $m < 1$, $nm^t \to \infty$.

First we will introduce the following notation, relevant to the likelihood martingale :

Let $Z_0(n), Z_1(n), \cdots, Z_k(n)$ be a sample of consecutive observations from a BGWR stochastic process with power series offspring distribution depending on the unknown single parameter θ. Assume that θ takes values in some open subset Θ of the positive part of the real line and that the distribution of $Z_0(n)$ does not depend on θ. Let $p_{k+1}(z_0, \cdots, z_k; \theta) = P(Z_0(n) = z_0, \cdots, Z_k(n) = z_k)$ be the joint probability function of $(Z_0(n), \cdots, Z_k(n))$. We will assume that

$$\sum_{z_k} p_{k+1}(z_0, \cdots, z_k; \theta)$$

can be differentiated twice with respect to θ under the summation sign.

Let $L_{k+1}(\theta) = p_{k+1}(z_0, \cdots, z_k; \theta)$ be the likelihood function associated with $Z_0(n), Z_1(n), \cdots, Z_k(n)$ and let the σ-algebra $\Im_{k+1} = \sigma(Z_0(n), \cdots, Z_k(n))$ describe the natural history of the process, where $\Im_0 = \{\emptyset, \Omega\}$ and $L_0 = 1$. Define

$$u_{i+1}(\theta) = \frac{d}{d\theta} \log p_{i+1}(Z_0(n), \cdots, Z_i(n)|\Im_i; \theta).$$

Since

$$L_{k+1}(\theta) = \prod_{i=0}^{k} p_{i+1}(Z_0(n), \cdots, Z_i(n)|\Im_i; \theta)$$

then

$$M_{k+1}(\theta) = \frac{d}{d\theta} \log L_{k+1}(\theta) = \sum_{i=0}^{k} u_{i+1}(\theta).$$

Note that almost surely for $k \geq 0$ $E\left(M_{k+1}(\theta)|\Im_k\right) = M_k(\theta)$, and therefore $\{M_{k+1}(\theta), \Im_{k+1},\ k \geq 0\}$ is a square integrable martingale.

Next, we set

$$I_{k+1}(\theta) = \sum_{i=1}^{k} E\left(M_{k+1}(\theta) - M_k(\theta)|\Im_k\right)^2, \quad I_1(\theta) \equiv 0.$$

This is a form of conditional information, which reduces to the standard Fisher information in the case of independent observations. Therefore

$$I_{k+1}(\theta) = E\left(M_{k+1}(\theta) - M_k(\theta)|\Im_k\right)^2 = E\left[\left(M_{k+1}(\theta)\right)^2|\Im_k\right] - \left(M_k(\theta)\right)^2$$

$$= E(u_{k+1}^2(\theta)|\Im_k)$$

can be interpreted as the conditional information contained in $Z_0(n), Z_1(n), \cdots,$ $Z_k(n)$ for given $Z_0(n), Z_1(n), \cdots, Z_{k-1}(n)$, which is not contained in $Z_0(n), Z_1(n),$ $\cdots,\ Z_{k-1}(n)$.

It can be obtained by interchanging integration and differentiation that

$$\begin{aligned} E_\theta(u_{i+1}^2(\theta)|\Im_i) = &\ E_\theta\left(\left[\frac{d}{d\theta}\log p_{i+1}(Z_0(n), \cdots, Z_i(n)|\Im_i; \theta)\right]^2 |\Im_i\right) = \\ = &\ \sum \frac{\dot{p}_{i+1}^2(z_0, \cdots, z_i|\Im_i; \theta)}{p_{i+1}^2(z_0, \cdots, z_i|\Im_i; \theta)} p_{i+1}(z_1, \cdots, z_i|\Im_i; \theta) \\ = &\ \sum \frac{\dot{p}_{i+1}^2(\cdot|\cdot) - \ddot{p}_{i+1}(\cdot|\cdot)p_{i+1}(\cdot|\cdot)}{p_{i+1}^2(\cdot|\cdot)} p_{i+1}(\cdot|\cdot) = \end{aligned}$$

$$
\begin{aligned}
&= -\sum \left[\frac{\dot{p}_{i+1}(\cdot|\cdot)}{p_{i+1}(\cdot|\cdot)}\right]' p_{i+1}(\cdot|\cdot) = -\sum \frac{d}{d\theta}\left[\frac{d}{d\theta}\log p_{i+1}(\cdot|\cdot)\right] p_{i+1}(\cdot|\cdot) \\
&= -\sum \frac{d}{d\theta}[u_{i+1}(\theta)]p_{i+1}(\cdot|\cdot) = E\left(-\frac{du_{i+1}(\theta)}{d\theta}|\Im_i\right).
\end{aligned}
$$

For simplicity we have denoted here by $p_{i+1}(\cdot|\cdot)$ the conditional probabilities $p_{i+1}(z_0,\cdots,z_i|\Im_i;\theta)$.

Define the observed information $J_{k+1}(\theta)=\sum_{i=0}^{k}\nu_{i+1}(\theta),\ \ J_1=0$, where $\nu_{i+1}(\theta)= -\frac{du_{i+1}(\theta)}{d\theta},\ \ \nu_1=0$. Then $\{[I_{k+1}(\theta)-J_{k+1}(\theta)],\ \Im_{k+1},\ k\geq 0\}$ is a martingale, because

$$
\begin{aligned}
E\left(I_{k+1}(\theta)-J_{k+1}(\theta)|\Im_k\right) &= E\left(\sum_{i=1}^{k}E(u_{i+1}^2(\theta)|\Im_i)-\sum_{i=1}^{k}\left(-\frac{du_{i+1}(\theta)}{d\theta}\right)|\Im_k\right) \\
&= \sum_{i=1}^{k}E\Big[E(u_{i+1}^2(\theta)|\Im_i)|\Im_k\Big]-\sum_{i=1}^{k}E\left(-\frac{du_{i+1}(\theta)}{d\theta}|\Im_k\right) \\
&= \sum_{i=1}^{k-1}E(u_{i+1}^2(\theta)|\Im_i)-\sum_{i=1}^{k-1}E\left(-\frac{du_{i+1}}{d\theta}|\Im_k\right)+ \\
&+ E(u_{k+1}^2(\theta)|\Im_k)-E\left(-\frac{du_{k+1}(\theta)}{d\theta}|\Im_k\right)= \\
&= I_k(\theta)-J_k(\theta),
\end{aligned}
$$

and $du_{i+1}(\theta)/d\theta,\ i=0,1,2,\cdots,k-1$, is $\Im_k$-measurable.

The Fisher information is defined as follows:

$$
\zeta_{k+1}(\theta)=E_\theta I_{k+1}(\theta)=E_\theta J_{k+1}(\theta).
$$

The three quantities, $\zeta_{k+1}(\theta),\ \ J_{k+1}(\theta),\ \ I_{k+1}(\theta)$ measure aspects of the information about θ, contained in the sample.

We also draw the attention to the fact, that for the power series distribution the following properties hold: its p.g.f has the form $P(s)=A(\theta s)/A(\theta)$; the mean can be represented as $m=\theta A'(\theta)/A(\theta)$ and the variance $\sigma^2=((d/dm)\log\theta)^{-1}$.

We begin our study by calculating the derivative of the logarithm of the likelihood function and the information quantities in the BGWR process with power series offspring distribution under a suitable reparametrization, using as a new parameter of the individual distribution the offspring mean m.

Proposition 1: For the BGWR processes with power series offspring distribution the following properties of the likelihood hold: for each $k = 0, 1, 2, \cdots$

$$u_{k+1}(m) = \frac{d}{dm} \log p_{k+1}(Z_0(n), Z_1(n), \cdots, Z_k(n) | \Im_k; m) =$$

$$(Z_k(n) - mZ_{k-1}(n)) \frac{d}{dm} \log \theta,$$

$$M_{t+1}(m) = \frac{1}{\sigma^2(m)} \cdot [Y_{t+1}(n) - Z_0(n) - mY_t(n)],$$

where $u_1(m) = \frac{d}{dm} \log p_1(Z_0) = 0$ and $Y_{t+1}(n) = \sum_{i=0}^{t} Z_i(n), \;\; Y_1 = Z_0, \;\; Y_0 = 0.$ ♢

P r o o f. The process is a discrete Markov chain and therefore

$$p_{k+1}(Z_0(n), Z_1(n), \cdots, Z_k(n) | \Im_k; m) =$$

$$\mathbf{P}\left(\sum_{i=1}^{Z_{k-1}(n)} \xi_i = Z_k(n) | Z_{k-1}(n)\right) = \sum_{s_1+\cdots+s_{Z_{k-1}(n)} = Z_k(n)} p_{s_1} \cdots p_{s_{Z_{k-1}(n)}} =$$

$$\sum_{s_1+\cdots+s_{Z_{k-1}(n)} = Z_k(n)} \frac{\left(\prod_{i=1}^{Z_{k-1}(n)} a_{s_i}\right) \theta^{\sum_{i=1}^{Z_{k-1}(n)} s_i}}{[A(\theta)]^{Z_{k-1}(n)}} =$$

$$\frac{\theta^{Z_k(n)}}{[A(\theta)]^{Z_{k-1}(n)}} \left[\sum_{s_1+\cdots+s_{Z_{k-1}(n)} = Z_k(n)} \prod_{i=1}^{Z_{k-1}(n)} a_{s_i}\right]$$

and

$$\log p_{k+1}(\cdot|\cdot) = Z_k(n) \log \theta - Z_{k-1}(n) \log A(\theta) + \log\left(\sum_{s_1+\cdots+s_{Z_{k-1}(n)} = Z_k(n)} \prod_{i=1}^{Z_{k-1}(n)} a_{s_i}\right).$$

Therefore

$$\frac{d}{dm} \log p_{k+1}(\cdot|\cdot) = Z_k(n) \frac{\theta'(m)}{\theta(m)} - Z_{k-1}(n) \frac{A'(\theta(m))\theta'(m)}{A(\theta(m))} =$$

$$\frac{Z_k(n) - \frac{\theta A'(\theta(m))}{A(\theta(m))} Z_{k-1}(n)}{\frac{\theta(m)}{\theta'(m)}} = [Z_k(n) - mZ_{k-1}(n)]/\sigma^2(m)$$

and

$$
\begin{aligned}
M_{t+1}(m) &= \frac{d}{dm}\log L_{t+1}(m) = \sum_{i=0}^{t} u_{i+1}(m) = \\
&= \frac{1}{\sigma^2(m)} \cdot \left(\sum_{i=1}^{t} Z_i(n) - m \sum_{i=1}^{t} Z_{i-1}(n) \right) = \\
&= \frac{1}{\sigma^2(m)} \cdot [Y_{t+1}(n) - Z_0(n) - mY_t(n)]. \qquad \square
\end{aligned}
$$

Proposition 2: For the BGWR processes with power series offspring distribution the following formulae hold: for each $t = 1, 2, \cdots$

$$
\begin{aligned}
I_{t+1}(m) &= \frac{Y_t(n)}{\sigma^2(m)}, \;\; I_1 = 0, \;\; I_2 = \frac{Z_0(n)}{\sigma^2}; \\
\nu_{t+1}(m) &= \frac{Z_{t-1}(n)}{\sigma^2(m)} + \frac{Z_t(n) - mZ_{t-1}(n)}{\sigma^4(m)} \cdot \frac{d}{dm}\sigma^2(m), \;\; \nu_1 = 0; \\
J_{t+1}(m) &= I_{t+1}(m) + M_{t+1}(m) \cdot \frac{d}{dm}\log \sigma^2(m), \;\; J_1 = 0; \\
\zeta_{t+1}(m) &= EI_{t+1}(m) = EJ_{t+1}(m) = EZ_0(n)\frac{m^t - 1}{\sigma^2(m)(m-1)}, \;\; \zeta_1 = 0. \;\; \diamondsuit
\end{aligned}
$$

P r o o f. Let us first calculate the conditional information:

$$
\begin{aligned}
I_{t+1}(m) &= \sum_{i=1}^{t} E_m \left(u_{i+1}^2(m) | \Im_i \right) = \sum_{i=1}^{t} E_m \left(\left(\frac{Z_i(n) - mZ_{i-1}(n)}{\sigma^2(m)} \right)^2 | \Im_i \right) = \\
&= \sum_{i=1}^{t} E_m \left(\frac{\left(\sum\limits_{k=1}^{Z_{i-1}(n)} \xi_k - Z_{i-1}(n) \cdot E_m \xi_i \right)^2}{\sigma^4(m)} \right) = \sum_{i=1}^{t} \frac{Var \sum\limits_{k=1}^{Z_{i-1}(n)} \xi_k}{\sigma^4(m)} = \\
&= \sum_{i=1}^{t} \frac{Z_{i-1}(n) \cdot \sigma^2(m)}{\sigma^4(m)} = \frac{Y_t(n)}{\sigma^2(m)}.
\end{aligned}
$$

For the observed information we need to calculate

$$
\begin{aligned}
\nu_{t+1}(m) &= -\frac{d}{dm} u_{t+1}(m) = -\frac{d}{dm} \frac{Z_t(n) - mZ_{t-1}(n)}{\sigma^2(m)} = \\
&= -\frac{\left(\frac{d}{dm} \left(Z_t(n) - mZ_{t-1}(n) \right) \right) \cdot \sigma^2(m) - \frac{d}{dm}\sigma^2(m) \left(Z_t(n) - mZ_{t-1}(n) \right)}{\sigma^4(m)}
\end{aligned}
$$

$$= \frac{Z_{t-1}(n)}{\sigma^2(m)} + \frac{Z_t(n) - mZ_{t-1}(n)}{\sigma^4(m)} \cdot \frac{d}{dm}\sigma^2(m).$$

For $J_{t+1}(m)$ we have:

$$\begin{aligned} J_{t+1}(m) &= \sum_{i=0}^{t} -\frac{d}{dm} u_{i+1}(m) = \sum_{i=0}^{t} \nu_{i+1}(m) = \\ &= \frac{\sum_{i=1}^{t} Z_{i-1}(n)}{\sigma^2(m)} + \frac{\sum_{i=1}^{t} (Z_i(n) - mZ_{i-1}(n))}{\sigma^2(m)} \cdot \frac{\frac{d}{dm}\sigma^2(m)}{\sigma^2(m)} = \\ &= I_{t+1}(m) + M_{t+1}(m) \cdot \frac{d}{dm} \log \sigma^2(m), \end{aligned}$$

hence $\zeta_{t+1}(m) = EI_{t+1}(m) = EJ_{t+1}(m) = EZ_0(n)\frac{\sum_{i=0}^{t-1} m^i}{\sigma^2(m)}$, because $EM_{t+1}(m) = 0$. □

From Proposition 1 immediately follows

Corollary 1: For the BGWR process with power series offspring distribution the parametric and nonparametric maximum likelihood estimators for the offspring mean coincide.

Proof. Obviously from the equation $M_{t+1}(m) = 0$ and Proposition 1 one obtains that

$$\widehat{m}_t = \sum_{i=1}^{t} Z_i(n) / \sum_{i=0}^{t-1} Z_i(n), \tag{4}$$

which is nothing but the well known Harris estimator for the offspring mean. □

Let us now assume that ϱ is a positive random variable.

Proposition 3: Let $n \to \infty$ in a BGWR process with power series offspring distribution. Then uniformly for t, $1 \le t \le \infty$ it follows that

(i) If $m < \infty$ and $Z_0(n) \xrightarrow{P} \infty$ or $\sigma^2 < \infty$, $Z_0(n)/n \xrightarrow{d} \varrho$ and Condition A holds then the likelihood equation has a weakly consistent root;

(ii) If $\sigma^2 < \infty$, Condition A holds and $Z_0(n)/n \xrightarrow{a.s.} \varrho$ then the likelihood equation has a strongly consistent root.

Proof. We will prove the existence of the strongly consistent root. The proof of the weak consistency is completely analogous. Let m be the true parameter value and let $|m^* - m| < \delta$ for some $\delta > 0$. Let us first calculate $[I_{t+1}(m) - J_{t+1}(m^*)]/I_{t+1}(m)$. It can be written in the form

$$\frac{I_{t+1}(m) - J_{t+1}(m^*)}{I_{t+1}(m)} = \frac{I_{t+1}(m) - J_{t+1}(m)}{I_{t+1}(m)} + \frac{J_{t+1}(m) - J_{t+1}(m^*)}{I_{t+1}(m)}. \tag{5}$$

According to Proposition 2

$$\begin{aligned}
\frac{I_{t+1}(m)-J_{t+1}(m)}{I_{t+1}(m)} &= -\frac{M_{t+1}(m)\cdot\frac{d}{dm}\log\sigma^2(m)}{\frac{Y_t(n)}{\sigma^2(m)}} = \\
&= -\frac{\frac{1}{\sigma^2(m)}\left(Y_{t+1}(n)-Z_0(n)-m\cdot Y_t(n)\right)}{\frac{Y_t(n)}{\sigma^2(m)}}\cdot\frac{d}{dm}\log\sigma^2(m) \\
&= \left(m-\frac{Y_{t+1}(n)-Z_0(n)}{Y_t(n)}\right)\cdot\frac{d}{dm}\log\sigma^2(m) = \\
&= \left(m-\widehat{m}_{t+1}(n)\right)\cdot\frac{d}{dm}\log\sigma^2(m),
\end{aligned}$$

where the Harris estimator $\widehat{m}_{t+1}(n)$ of the offspring mean m of a BGWR process is defined by (4).

Applying Theorem 2.1. from [8] it follows that

$$\frac{M_{t+1}(m)}{I_{t+1}(m)} = \left(\left[\sum_{k=1}^{t} Z_k(n)\right] / \left[\sum_{k=0}^{t-1} Z_k(n)\right] - m\right) \to 0 \ \ a.s.,$$

and the first term in the right part of equality (5) tends to zero a.s.

Let us consider the second term of (5). It can be presented in the form

$$\begin{aligned}
&\frac{J_{t+1}(m)-J_{t+1}(m^*)}{I_{t+1}(m)} = \\
= \ &\frac{I_{t+1}(m)-I_{t+1}(m^*)}{I_{t+1}(m)} + \frac{M_{t+1}(m)}{I_{t+1}(m)}\frac{d\log\sigma^2(m)}{dm} - \frac{M_{t+1}(m^*)}{I_{t+1}(m)}\frac{d\log\sigma^2(m^*)}{dm^*} \\
= \ &\frac{Y_t(n)\left(\frac{1}{\sigma^2(m)}-\frac{1}{\sigma^2(m^*)}\right)}{\frac{Y_t(n)}{\sigma^2(m)}} + o_{a.s.}(1) - \frac{M_{t+1}(m^*)}{I_{t+1}(m^*)}\frac{I_{t+1}(m^*)}{I_{t+1}(m)}\frac{d}{dm^*}\log\sigma^2(m^*) \\
= \ &\left(1-\frac{\sigma^2(m)}{\sigma^2(m^*)}\right) + o_{a.s.}(1) - o_{a.s.}(1)\cdot\frac{Y_t(n)\sigma^2(m)}{Y_t(n)\sigma^2(m^*)}\frac{d}{dm^*}\log\sigma^2(m^*) \\
= \ &\left(1-\frac{\sigma^2(m)}{\sigma^2(m^*)}\right) + o_{a.s.}(1) - o_{a.s.}(1)\frac{\sigma^2(m)}{\sigma^2(m^*)}\frac{d}{dm^*}\log\sigma^2(m^*).
\end{aligned}$$

But $\sigma^2(m)/\sigma^2(m^*)$ can be made arbitrarily close to one by continuity of $\sigma^2(m)$. Hence the second term of (5)tends to zero a.s. and therefore

$$[I_{t+1}(m)-J_{t+1}(m^*)]/I_{t+1}(m) \to 0 \ \ a.s. \ \ as \ \ t\to\infty.$$

On the other hand, expanding $M_{t+1}(m_1)$ around the true parameter value m, where in general $m_1 \neq m$, it follows that

$$\begin{aligned} M_{t+1}(m_1) &= M_{t+1}(m) - (m_1 - m)\sum \nu_{i+1}(m^*) = \\ &= M_{t+1}(m) - (m_1 - m)I_{t+1}(m) + (m_1 - m)(I_{t+1}(m) - J_{t+1}(m^*)). \end{aligned}$$

Here $m^* = m + \varepsilon(m_1 - m)$ for some $\varepsilon = \varepsilon(t, m)$ with $|\varepsilon| < 1$. Solving the likelihood equation $M_{t+1}(m_1) = 0$ yields

$$(\widehat{m}_{t+1}(n) - m)\left(1 - \frac{I_{t+1}(m) - J_{t+1}(m^*)}{I_{t+1}(m)}\right) = \frac{M_{t+1}(m)}{I_{t+1}(m)},$$

whence the claim follows. □

Remark: In the proof we used the interesting technique for nonergodic processes, proposed in Heyde and Feigin (1975) and applied to the classical BGW processes in [10]. Furthermore, Proposition 3 can be considered as an analogy to the results obtained in [8] in the nonparametric situation, where the random sums method is used.

Proposition 4: Let $\{Z_t(n)\}$ be a BGWR process with power series offspring distribution. Let $\sigma^2 < \infty$, $Z_0(n)/n \xrightarrow{P} \varrho$ and $n, t \to \infty$ (with the additional condition $t/n \to 0$ in the critical case $m = 1$) or $n \to \infty$ and t is fixed. Then

$$\sqrt{I_{t+1}(m)}(\widehat{m}_{t+1}(n) - m) \xrightarrow{d} N(0, 1).$$

Proof. The claim can be considered as a particular case of the existence of an asymptotically normal root of the likelihood equation in the nonergodic processes.

We have already proved that

$$\frac{J_{t+1}(m)}{I_{t+1}(m)} = \frac{I_{t+1}(m) + M_{t+1}(m) \cdot \frac{d}{dm}\log\sigma^2(m)}{I_{t+1}(m)} = 1 + \frac{M_{t+1}(m)}{I_{t+1}(m)} \cdot \frac{d}{dm}\log\sigma^2(m) \xrightarrow{P} 1$$

and $I_{t+1}(m) - J_{t+1}(m^*)/I_{t+1}(m) \xrightarrow{P} 0$ under the conditions of Proposition 3. Thus similarly as in Proposition 3 using the equality

$$M_{t+1}(m) = I_{t+1}(m)(m_1 - m)\left(1 - \left[I_{t+1}(m) - J_{t+1}(m^*)\right]/I_{t+1}(m)\right),$$

and applying statements from Corollary 2.2.1 and Theorem 2.3 from [8], which ensure the equality

$$\frac{M_{t+1}(m)}{\sqrt{I_{t+1}(m)}} = \left(\left[\sum_{k=1}^{t} Z_k(n)\right] \Big/ \left[\sum_{k=0}^{t-1} Z_k(n)\right] - m\right)\sqrt{Y_t(n)}/\sigma^2(m) \xrightarrow{d} N(0,1)$$

continuously in m, the claim follows. $\square$

Corollary 2: If $\sigma^2 < \infty$ and $Z_0(n)/n \stackrel{d}{\longrightarrow} \varrho$ then the conclusion is valid in the critical case $m = 1$ when $t/n \to 0$.$\diamondsuit$

Remark: The proposition can be considered as an analogy of Theorems 2.2, 2.3 and Corollary 2.2.1 in [8].

3. Efficiency

The technique, which uses the information quantities, allows the consideration of the efficient properties of $\widehat{m}_{t+1}(n)$.

Let us consider the class K of all asymptotically unbiased and asymptotically normal estimators of the offspring mean m. The estimator m_1^* is asymptotically efficient in K as $t \longrightarrow \infty$ if for an arbitrary estimator $m^* \in K$ and for every parameter value m

$$\limsup_{t\to\infty} \frac{E_m(m_1^* - m)^2}{E_m(m^* - m)^2} \leq 1.$$

Let the estimator m^* belong to the class K and its variance equal $\sigma^2_{m^*}(t)$ (here t is the number of observations over the process). The asymptotic efficiency is defined as

$$e(m^*) = \lim \frac{1}{\sigma^2_{m^*}(t) I_t(m)};$$

the asymptotic relative efficiency of the estimator m_1^* to the estimator m^* is defined by the ratio $\lim \sigma^2_{m^*}(t)/\sigma^2_{m_1^*}(t)$. Those two definitions are established only for the asymptotically normal estimators and refer to the first order asymptotic properties (see Cox and Hinkley 1974) .

C.C.Heyde (1975) gives a definition of asymptotic efficiency, which generalizes that of Rao (1973, pp 348-349). A consistent estimator m_t^* of m is asymptotically (Rao) efficient if

$$I_t(m)^{1/2} \left\{ m_t^* - m - \beta(m) I_t(m)^{-1} M_t(m) \right\} \longrightarrow 0 \tag{6}$$

in probability as $t \to \infty$ for some $\beta(m)$- a real - valued, nonnegative function of the studied parameter m, which does not involve the observations. In the standard case of independent and identically distributed observations this definition reduces to that of Rao (1973). A first order efficient estimator (Rao-efficient estimator) should asymptotically be a linear function of the derivative of the logarithm of the likelihood, since the latter contains all the information available from the data about the unknown parameter.

Hence, provided the conditions of Proposition 4 are satisfied we say that m_t^* has asymptotic relative efficiency of $\beta(m)^{-2}$ with respect to the m.l.e. $\widehat{m}_{t+1}(n)$.

The definition of asymptotic Rao-efficiency ensures the asymptotic normality of the estimator m_t^*: $I_t^{1/2}(m)(m_t^* - m) \to N(0, \beta^2(m))$ in distribution.

If the condition 6 is strengthened to convergence in the mean in order two, then equivalent forms are

$$corr\left[I_t^{1/2}(m)(m_t^* - m), I_t^{-1/2}(m)\frac{d\log L_t(m)}{dm}\right] \to 1, \tag{7}$$

or alternatively,

$$E\left[(m_t^* - m)\cdot\frac{d\log L_t(m)}{dm}\right]\Big/\, E\left[I_t(m)(m_t^* - m)^2\right]\cdot\left[I_t^{-1}(m)\left\{\frac{d\log L_t(m)}{dm}\right\}^2\right] \to 1. \tag{8}$$

Let $\overline{m}_{t+1}(n) = \frac{Z_t(n)}{Z_{t-1}(n)}$ be the Lotka-Nagaev estimator.

Proposition 5: In a BGWR process with PSOD if $\sigma^2 < \infty$, $\frac{Z_0(n)}{n}.\xrightarrow{a.s.} \varrho$ and Condition A holds then $\widehat{m}_{t+1}(n)$ is Rao-efficient. If $m > 1$ the asymptotic relative efficiency of $\overline{m}_{t+1}(n)$ to $\widehat{m}_{t+1}(n)$ is $1 - \frac{1}{m}$. If $m \le 1$ then $\overline{m}_{t+1}(n)$ is not asymptotically efficient. If $m > 1$ $\widehat{m}_{t+1}(n)$ is not L_2- asymptotically Rao-efficient.

Proof. According to Proposition 4 $I_{t+1}^{1/2}(m)(\widehat{m}_{t+1}(n) - m) \xrightarrow{d} N(0,1)$ and

$$I_{t+1}^{1/2}(m)\left(\widehat{m}_{t+1}(n) - m - \beta(m)\frac{M_{t+1}(m)}{I_{t+1}(m)}\right) =$$

$$= \sqrt{\frac{Y_t(n)}{\sigma^2(m)}}\left((\widehat{m}_{t+1}(n) - m) - \frac{\sum_{k=1}^{t} Z_k(n) - m\sum_{k=0}^{t-1} Z_k(n)}{\sum_{k=0}^{t-1} Z_k(n)}\right)$$

$$= \sqrt{\frac{Y_t(n)}{\sigma^2(m)}}\left((\widehat{m}_{t+1}(n) - m) - (\widehat{m}_{t+1}(n) - m)\right) \equiv 0.$$

Here we have replaced $\beta(m)$ in (6) by 1.

Let us now consider $\overline{m}_{t+1}(n)$. It holds that

$$\sqrt{\frac{Z_{t-1}(n)}{\sigma^2}}(\overline{m}_{t+1}(n) - m) \xrightarrow{d} N(0,1)$$

according to [8], Theorem 2.5. Our aim is to use as a normalizing factor the conditional information $I_{t+1}(m)$. In that way

$$\sqrt{\frac{Z_{t-1}(n)}{\sigma^2}\cdot\frac{\frac{m-1}{m}}{\frac{m-1}{m}}\cdot\frac{\sum_{k=0}^{t-1}Z_k(n)}{\sum_{k=0}^{t-1}Z_k(n)}}\cdot(\overline{m}_{t+1}(n)-m)=$$

$$=\sqrt{\frac{Z_{t-1}(n)}{\sum_{k=0}^{t-1}Z_k(n)\cdot\frac{m-1}{m}}}\cdot\sqrt{\frac{m-1}{m}}\cdot\sqrt{I_{t+1}(m)}\cdot(\overline{m}_{t+1}(n)-m).$$

But according to [7], Lemmas 1, 2 and 3

$$\sum_{k=0}^{t-1}Z_k(n)\cdot(m-1)/Z_t(n)\stackrel{a.s.}{\longrightarrow}1\quad if\quad m>1$$

and

$$Z_{t-1}(n)/Z_0(n)m^{t-1}\stackrel{a.s.}{\longrightarrow}1\ .$$

It results that

$$\frac{\sum_{k=0}^{t-1}Z_k(n)\cdot\frac{m-1}{m}}{Z_{t-1}(n)} = \frac{\sum_{k=0}^{t-1}Z_k(n)(m-1)}{Z_t(n)}\frac{Z_t(n)}{Z_0(n)m^t}\frac{Z_0(n)m^{t-1}}{Z_{t-1}(n)}\frac{m}{m}\longrightarrow 1\ a.s.,$$

$$\sqrt{I_{t+1}(m)}\cdot\sqrt{\frac{m-1}{m}}\cdot(\overline{m}_{t+1}(n)-m)\stackrel{d}{\longrightarrow}N(0,1)$$

and the asymptotic relative efficiency of $\overline{m}_{t+1}(n)$ to $\widehat{m}_{t+1}(n)$ is $1-\frac{1}{m}$.

Completely analogous if $m<1$

$$\sqrt{I_{t+1}(m)}\cdot\sqrt{(1-m)m^{t-1}}\cdot(\overline{m}_{t+1}(n)-m)\stackrel{d}{\longrightarrow}N(0,1)$$

and if $m=1$

$$\sqrt{I_{t+1}(m)}\cdot\sqrt{\frac{1}{tm}}\cdot(\overline{m}_{t+1}(n)-m)\stackrel{d}{\longrightarrow}N(0,1)$$

and the relative efficiency $\beta(m)^{-2}$ from (6) has to be equal to zero.

Let us prove the last part of the proposition. Using the fact that $\overline{m}_{t+1}(n)$ is an unbiased estimator, it is valid that

$$E\overline{m}_{t+1}(n)\cdot\frac{d\log L_{t+1}(m)}{dm} = E\overline{m}_{t+1}(n)\cdot\frac{dL_{t+1}(m)}{dm}\cdot\frac{1}{L_{t+1}(m)}$$

$$\begin{aligned} &= \sum \overline{m}_{t+1}(n) \cdot \frac{dL_{t+1}(m)}{dm} \cdot \frac{1}{L_{t+1}(m)} \cdot L_{t+1}(m) \\ &= \sum \overline{m}_{t+1}(n) \cdot \frac{dL_{t+1}(m)}{dm} \\ &= \frac{d}{dm} \sum \overline{m}_{t+1}(n) \cdot L_{t+1}(m) = \frac{d}{dm} m = 1. \end{aligned}$$

Similarly, using the fact that $E\left[d \log L_{t+1}(m)/dm\right] = 0$

$$\begin{aligned} E[\overline{m}_{t+1}(n) - m] I_{t+1}(m)[\hat{m}_{t+1}(n) - m] = \\ E\overline{m}_{t+1}(n) \frac{d \log L_{t+1}(m)}{dm} - mE\frac{d \log L_{t+1}(m)}{dm} = \\ E\overline{m}_{t+1}(n) I_{t+1}(m)[\hat{m}_{t+1}(n) - m] = 1. \end{aligned} \tag{9}$$

It follows that

$$\frac{E\left[I_{t+1}^{1/2}(m)(\overline{m}_{t+1}(n) - m) \cdot I_{t+1}^{1/2}(m)(\hat{m}_{t+1}(n) - m)\right]}{\left\{E\left[I_{t+1}(m)(\hat{m}_{t+1}(n) - m)^2\right] \cdot E\left[I_{t+1}(m)(\overline{m}_{t+1}(n) - m)^2\right]\right\}^{1/2}} =$$

$$= 1 \; / \left\{E\left[I_{t+1}(m)(\overline{m}_{t+1}(n) - m)^2\right] \cdot E\left[I_{t+1}(m)(\hat{m}_{t+1}(n) - m)^2\right]\right\}^{1/2}$$

and, using Theorem 5.3. of Billingsley(1968)

$$\liminf E\left[I_{t+1}(m)(\overline{m}_{t+1}(n) - m)^2\right] \geq \frac{m}{m-1}$$

while

$$\liminf E\left[I_{t+1}(m)(\hat{m}_{t+1}(n) - m)^2\right] \geq 1,$$

so that if $m > 1$

$$\limsup \frac{E\left[I_{t+1}^{1/2}(m)(\overline{m}_{t+1}(n) - m) \cdot I_{t+1}^{1/2}(m)(\hat{m}_{t+1}(n) - m)\right]}{\left\{E\left[I_{t+1}(m)(\hat{m}_{t+1}(n) - m)^2\right] \cdot E\left[I_{t+1}(m)(\overline{m}_{t+1}(n) - m)^2\right]\right\}^{1/2}} \leq \sqrt{\frac{m-1}{m}} < 1. \tag{10}$$

□

Remark: The result can be considered as a generalization of the classical BGW process for $m > 1$, where upon nonextinction the estimator $\widehat{m}_t(1)$ is Rao-efficient and the asymptotic relative efficiency of the Lotka-Nagaev estimator

$\overline{m}_t(1) = Z_t(1)/Z_{t-1}(1)$ to $\widehat{m}_t(1)$ for the PSOD is $1 - 1/m$.

In many situations it is preferable to have an optimality concept that is not asymptotic and does not have to be interpreted conditionally. The proposed by Godambe (1985) finite sample optimality property for estimators is based on ideas, similar to those underlying Rao-efficiency. Under mild regularity conditions the unbiased estimating function $g^* = g^*(Z_0(n), \cdots, Z_t(n))$, $E_m g^* = 0$, is optimal if it is highly correlated with the underlying score function, i.e. $corr(g^*, M_{t+1}(m)) \geq corr(g, M_{t+1}(m))$. In Heyde (1991) is described one important general strategy for finding optimal estimating functions: to compute the score function for some plausible underlying distribution (such as a convenient member of the appropriate exponential family) $\partial \log L_2(m)/\partial m$. Then use the differences in this martingale to form the functions $h_{t+1}(m)$'s when we restrict to the class of linear estimating functions $\mathrm{H} = \left\{ \mathrm{g} = \sum_{\mathrm{i}=1}^{\mathrm{t}} \mathrm{h}_{\mathrm{i}+1}(\mathrm{m})\mathrm{a}_{\mathrm{i}}(\mathrm{m}) \right\}$ with $a_i(m)$ which may depend on m and $Z_0(n), \cdots, Z_{i-1}(n)$, $E(h_{i+1}(m)|\Im_i) = 0$, $Eh_s(m) \cdot h_t(m) = 0$ for $s \neq t$. Finally, one can choose the optimal estimating function within this class suitably specifying the weights a_{i+1}.

Proposition 6: In a BGWR process with power series offspring distribution the optimal estimating equation for the offspring mean m is

$$g^* = \sum_{i=1}^{t} (Z_i(n) - mZ_{i-1}(n)) = 0,$$

which has the solution $\widehat{m}_{t+1}(n)$. The Lotka-Nagaev estimator $\overline{m}_{t+1}(n)$ is not optimal within the class of estimating functions.

P r o o f. Let us first calculate

$$\frac{\partial \log L_2(m)}{\partial m} = u_2(m) = [Z_1(n) - mZ_0(n)] \frac{d}{dm} \log \theta$$

Now it is easy to find the basic martingale from $Z_i(n)$ since

$$Z_i(n) - E(Z_i(n)|\Im_i) = Z_i(n) - mZ_{i-1}$$

are martingale differences. Let

$$\mathrm{H} = \{g : \ g = \sum_{i=1}^{t} a_i(m)(Z_i(n) - mZ_{i-1}), \ \ a_i(m) \ is \ \Im_i - measurable.\}$$

In the class H the optimal estimating function $g^*$is given by choosing

$$a_i^* = E\left(\frac{\partial h_{i+1}(m)}{\partial m}|\Im_i\right)/E(h_i^2(m)|\Im_i).$$

But

$$E\left(\frac{\partial h_{i+1}(m)}{\partial m}|\Im_i\right) = E\left(\frac{\partial (Z_i(n) - mZ_{i-1})}{\partial m}|\Im_i\right) = Z_{i-1}$$

and

$$E\left((Z_i(n) - mZ_{i-1})^2|\Im_i\right) = Var\left(\sum_{k=1}^{Z_{i-1}} X_k|\Im_i\right) = Z_{i-1}\sigma^2(m).$$

Hence $a_i^*(m) = -1/\sigma^2(m)$. Then the O_F-optimal estimating equation is

$$-\frac{1}{\sigma^2(m)}\left[\sum_{i=1}^{t} Z_i(n) - m\sum_{i=1}^{t} Z_{i-1}(n)\right] = 0$$

and the optimal estimator of m is $\widehat{m}_{t+1}(n)$. Since $\overline{m}_{t+1}(n)$ satisfies $g = \frac{Z_t(n)}{Z_{t-1}(n)} - m = 0$ it is not optimal within the class of estimating functions H. □

We will now see that the following property holds:

Proposition 7: In a BGWR process with power series offspring distribution the random vector $(Y_t(n), \widehat{m}_{t+1}(n))$ is a minimal sufficient statistic.

P r o o f. Note that

$$\begin{aligned} M_{t+1}(m) &= \frac{1}{\sigma^2(m)}(Y_{t+1}(n) - Z_0(n) - mY_t(n)) = \\ &= I_{t+1}(m)(\widehat{m}_{t+1}(n) - m) \end{aligned}$$

and the equation holds for all $t \geq 1$ if and only if

$$M_{t+1}(n) = \Phi(m)H_{t+1}(Z_0(n), \cdots, Z_t(n))(\widehat{m}_{t+1}(n) - m)$$

(see f.e. [10], Proposition 2.12.). The Proposition follows from the factorization criterion. □

Remark: It is the case that the m.l.e. is not a sufficient statistic by itself (in the contrast to the i.i.d. model where H_{t+1} is a constant and in our case $H_{t+1} = \sum_{i=0}^{t-1} Z_i = Y_t(n)$). It is known that one may explain generally the good performance of the m.l.e.'s in the i.i.d. case by the fact that they are asymptotically sufficient under mild regularity conditions (see f.e. Cox and Hinkley 1974) even when there is no exactly sufficient statistic. In the general nonergodic case sufficiency may not even occur asymptotically. Even though $\widehat{m}_{t+1}(n)$ is asymptotically efficient, it is not asymptotically sufficient, since the factor multiplying $\Phi(m)$ depends both on $Y_t(n)$ and $\widehat{m}_{t+1}(n)$, which can not be improved by conditioning on the sufficient statistic, because it is already a part of it.

REFERENCES

[1] Billingsley P. Convergence of Probability Measures. New York, Wiley 1968

[2] Dion, J.-P. Estimation of the Variance of a Branching Process *Ann. Statist.* **3** (1975) , 1183–1187.

[3] Dion, J.P. Statistical Inference for Discrete Time Branching Processes. *Proc. 7th International Summer School on Prob.Th.& Math.Statist.* (Varna 1991). *Sci.Cult.Tech.Publ.* Singapore (1993), 60–121.

[4] Dion, J.-P., N. M. Yanev A New Transfer Limit Theorem. *Compt. rend. Acad. bulg. Sci.* **44** (1) (1991) , 19–21.

[5] Dion, J.-P., N. M. Yanev Limiting Distributions of a Galton-Watson Branching Process with a Random Number of Ancestors. *Compt. rend. Acad. bulg. Sci.* **44** (3) (1991), 23-26.

[6] Dion, J.P., N. M. Yanev Estimation Theory for Branching Processes with or without Immigration. *Compt. rend. Acad. bulg. Sci.* **44** (4) (1991), 19–22.

[7] Dion, J.-P., N. M. Yanev Statistical Inference for Branching Processes with an Increasing Number of Ancestors. *J.Statistical Planning & Inference* **39** (1994), 329–359.

[8] Dion, J.-P., N. M. Yanev Limit Theorems and Estimation Theory for Branching Processes with an Increasing Number of Ancestors. *J.Appl.Prob.* **34** (1997), 309–327.

[9] Godambe, V. P. An Optimum Property of Regular Maximum Likelihood Estimation. *Ann. Math. Statist.* **31** (1960), 1208–1211

[10] Guttorp, P. Statistical Inference for branching processes. John Wiley and Sons, New York. 1991

[11] Heyde, C. C. Remarks on Efficiency in Estimation for Branching Processes. *Biometrika* vol.**62** (1975), 49–55.

[12] Heyde, C. C. Quasi-Likelihood And Its Application. A General Approach to Optimal Parameter Estimation. Springer Series in Statistics, New York. 1997

[13] Jagers, P. Branching Processes with biological applications. John Wiley and Sons, New York.1975

[14] Rao, C. R. Linear Statistical Inference and its Applications. 2nd edition. New York, Wiley.1973

[15] Stoimenova, V., D. Atanasov, N. Yanev Robust Estimation and Simulation of Branching Processes *Comptes rendus de l'Academie bulgare des Sciences.* **5** (2004), 19–22

[16] Stoimenova, V., D. Atanasov, N. Yanev Simulation and Robust Modifications of Estimates in Branching Processes. *Pliska Stud. Math. Bulgar.* **16** (2004), 259–271

[17] Stoimenova, V., D. Atanasov, N. Yanev Algorithms for Generation and Robust Estimation of Branching Processes with Random Number of Ancestors (2005), to appear

[18] Wei, C. Z., J. Winnicki Estimation of the Means in the Branching Process with Immigration. *Ann. Statist.* **18** (1990), 1757–1773.

[19] Yanev, N. M. On the Statistics of Branching Processes. *Theory.Probab.Appl.* **20** (1975), 612–622.

[20] Yanev, N. M. Limit Theorems for Estimators in Galton-Watson Branching Processes. *C.R. Acad.Bulg.Sc.* **38** (1985), 683–686.

Vessela Stoimenova
Sofia University,
Faculty of Mathematics
and Informatics
5 J. Boucher Str.
1164 Sofia, Bulgaria,
e-mail: `stoimenova@fmi.uni-sofia.bg`

Nickolay Yanev
Bulgarian Academy of Science,
Acad. G. Bontchev Str.
1113 Sofia, Bulgaria,
e-mail: `yanev@math.bas.bg`