Serdica

Mathematical Journal

Сердика

Математическо списание

ROBUST PARAMETRIC ESTIMATION OF BRANCHING PROCESSES WITH A RANDOM NUMBER OF ANCESTORS[*]

Vessela Stoimenova

Communicated by J.-P. Dion

ABSTRACT. The paper deals with a robust parametric estimation in branching processes $\{Z_t(n)\}$ having a 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. Robust estimators, based on one and several sample paths, are proposed and studied for all values of the offspring mean m, $0 < m < \infty$, in the subcritical, critical and supercritical case.

1. Introduction. There are many situations, which give rise to the Bienayme – Galton – Watson processes having a random number of ancestors (or

BGWR processes). Yakovlev and Yanev [28] noted that branching processes with a large and often a random number of ancestors occur naturally in the study of cell proliferation and in applications to nuclear chain reactions. Dion and Yanev [6] suggest two other motivations for the study of a random number of ancestors relevant to the classical Bienayme – Galton – Watson (or BGW) process and to the branching process with immigration (or BGWI) (see also Dion and Yanev [4], Dion [3]).

Results about the nonparametric estimation of the offspring mean m and variance σ^2 in the BGWR process are announced in Dion and Yanev [5, 6, 7] and Dion [3], where the nonparametric m.l.e. and a family of l.s.e. for σ^2 are studied and consistency and asymptotic normality of these estimators are obtained for all values of the mean m, $0 < m < \infty$. Results about the robust parametric and nonparametric estimation in BGWR processes are presented in Stoimenova, Atanasov, Yanev, [21, 23]. Some features of the parametric estimation in BGWR processes are studied in Stoimenova, Yanev [24].

Our main concern in this paper is the robust parametric estimation of a BGWR process with power series offspring distribution based on one sample path $\{Z_0(n), \ldots, Z_t(n)\}$ as both n and t tend to infinity (and thus $Z_0(n)$ in some sense) and on several realizations of the process. The relative speed, at which n and $t \to \infty$, has its important role for $0 < m < \infty$.

It is known that when there are outliers in the observations (dependent and independent), the standard methods of estimation (maximum likelihood, least squares , method of moments, etc.) may be seriously affected with adverse results. Robustified versions of some of these procedures are discussed in quite a number of sources, for example Huber [13], Gastwirth and Rubin [8], Kulkarni and Heyde [16]; most of the works concern time series, f.e. Denby and Martin [2], Martin [19], Künsch [17] and references therein. We note that the author is not aware of any work in the field of robust statistics in stochastic processes, using the d-fullness technique of Vandev [25].

In Section 2 the definition of a BGWR process is reminded and the parametric maximum likelihood is introduced. In Section 3 an overview of robustness and the theory of d-fullness is considered. In Section 4 the robust estimation is considered in both cases – the parametric situation based on one and several realizations of the process. Robust estimators of the parameter of the BGWR process with power series offspring distribution are introduced. The index of fullness of the set of log-density functions, arising from the sample, and the breakdown point properties of the proposed estimators are studied. Based on the asymptotic properties of the nonparametric estimators of the BGWR process, considered in

Dion and Yanev [7], the consistency of the robust estimators is studied. Section 5 contains numerical results based on generated realizations of the process.

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

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

Such a process is 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

$$(2) \qquad m = E\xi, \ \ \sigma^2 = Var(\xi).$$

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

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

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

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

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

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

Condition A is needed to ensure that $Z_t(n)$ tends to infinity when t is sufficiently large.

We first introduce the following notation, relevant to the likelihood martingale. Let $Z_0(n), Z_1(n), \ldots, Z_k(n)$ be a sample of consecutive observations from a BGWR stochastic process with PSOD 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 $p(z_0, \ldots, z_k; \theta) = P(Z_0(n) = z_0, \ldots, Z_k(n) = z_k)$ be the joint probability function of $(Z_0(n), \ldots, Z_k(n))$. We assume that $\sum\limits_{z_k} p(z_0, \ldots, z_k; \theta)$ can be differentiated twice with respect to θ under the summation sign.

Let $L_{k+1}(\theta) = p(Z_0, \ldots, Z_k; \theta)$ be the likelihood function associated with $Z_0(n), Z_1(n), \ldots, Z_k(n)$, the σ-algebra $\Im_{k+1} = \sigma(Z_0(n), \ldots, Z_k(n))$ describe the natural history of the process, $\Im_0 = \{\emptyset, \Omega\}$ and $L_0 = 1$. Define $u_{i+1}(\theta) = \dfrac{d}{d\theta} \log p(Z_0(n), \ldots, Z_i(n) | \Im_i; \theta)$. Since $L_{k+1}(\theta) = \prod\limits_{i=0}^{k} p(Z_0(n), \ldots, Z_i(n) | \Im_i; \theta)$ then $M_{k+1}(\theta) = \dfrac{d}{d\theta} \log L_{k+1}(\theta) = \sum\limits_{i=0}^{k} u_{i+1}(\theta)$. Note that almost surely for $k \geq 0$ $E\left(M_{k+1}(\theta) \big| \Im_k\right) = M_k(\theta)$, and therefore $\{M_{k+1}(\theta), \Im_{k+1}, \; k \geq 0\}$ is a square integrable martingale.

Calculating the derivative of the logarithm of the likelihood function and the information quantities in the BGWR process with PSOD under the reparametrization, which uses the offspring mean m as a new parameter of the individual distribution , in Stoimenova, Yanev [24] the derivative of the log likelihood is calculated. Hence for the BGWR processes with power series offspring distribution for each $k = 1, 2, \ldots$

$$u_{k+1}(m) = (Z_k(n) - m Z_{k-1}(n)) \frac{d}{dm} \log \theta,$$

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

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

Due to the Markov property of the BGWR process the following equations are obtained:

$$p(Z_0(n), Z_1(n), \ldots, Z_k(n) | \Im_k; m) =$$

(4)
$$= \frac{\theta^{Z_k(n)}}{[A(\theta)]^{Z_{k-1}(n)}} \left[\sum_{s_1 + \cdots + s_{Z_{k-1}(n)} = Z_k(n)} \prod_{i=1}^{Z_{k-1}(n)} a_{s_i} \right]$$

and

$$\log p(Z_0(n), Z_1(n), \ldots, Z_k(n)|\Im_k; m) =$$

(5)
$$= Z_k(n) \log \theta - Z_{k-1}(n) \log A(\theta) + \log \left(\sum_{s_1+\cdots+s_{Z_{k-1}(n)}=Z_k(n)} \prod_{i=1}^{Z_{k-1}(n)} a_{s_i} \right).$$

Consequently, for the BGWR process with power series offspring distribution the parametric and nonparametric maximum likelihood estimators for the offspring mean coincide (see f.e. Keiding and Lauritzen [15]).

From now on we note the nonparametric maximum likelihood estimator for m by

(6)
$$\widehat{m}_{t+1}(n) = \sum_{i=1}^{t} Z_i(n) / \sum_{i=0}^{t-1} Z_i(n),$$

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

Let us now assume that ϱ is a positive random variable. The following property holds (it follows as a consequence of the results from Dion and Yanev [7], see also Stoimenova, Yanev [24]).

Proposition 1. *Let $n \to \infty$ in a BGWR process with PSOD. 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.

3. The concept of d-fullness. The classical maximum likelihood estimator (MLE) can be very sensitive to outliers in the data. In fact, even a single outlier can ruin totally the maximum likelihood estimate. A robust extension of the MLE, called Weighted Least Trimmed (WLT) estimator was proposed by Hadi and Luceño [10] and Vandev and Neykov [26]. This modification considers the likelihood of individual observations as residuals and applies the basic idea of the Least Trimmed Sqwares (LTS) estimators of Rousseeuw [20] using appropriate weights. Depending on the weights choice and the trimming constant, the WLT estimator reduces to the MLE, to the Least Median of Squares and LTS estimators in the normal regression cases, to the Minimum Volume Ellipsoid and Minimum Covariance Determinant estimators of the multivariate

location and scatter parameter in the multivariate normal cases, considered in details by Rousseeuw and Leroy [20] (see also Vandev and Neykov [26, 27]).

Generally speaking, Vandev and Neykov [27] defined the $WLT(k)$ estimators, $\widehat{\theta}$, for the unknown parameter $\theta \in \Theta^p$ as

$$(7) \qquad \widehat{\theta} = \operatorname*{argmin}_{\theta \in \Theta^p} \sum_{i=1}^{k} w_i f_{\nu(i)}\left(\theta\right),$$

where $f_{\nu(1)}\left(\theta\right) \leq f_{\nu(2)}\left(\theta\right) \leq \cdots \leq f_{\nu(n)}\left(\theta\right)$ are the ordered values of $f_i = -\log \varphi\left(x_i, \theta\right)$ at θ, $\varphi\left(x_i, \theta\right)$ is a probability density, θ is an unknown parameter and $\nu = (\nu(1), \ldots, \nu(n))$ is the corresponding permutation of the indices, which may depend on θ. The weights $w_i \geq 0$, $i = 1, \ldots, k$, are such that an index $k = \max\{i : w_i > 0\}$ exists.

The breakdown properties of the $WLT(k)$ estimator were studied by Vandev and Neykov [27], Atanasov and Neykov [1] and others. Vandev and Neykov [27] proved that the finite sample breakdown point of the $WLT(k)$ estimators is not less than $(n-k)/n$ if $n \geq 3d$, $(n+d)/2 \leq k \leq n-d$, when Θ^p is a topological space and the set $F = \{f_i(\theta), i = 1, \ldots, n\}$ is d-full. A finite set F of n functions is called d-full, according to Vandev [25], if for each subset $J \subset \{i = 1, \ldots, n\}$ of cardinality d $(|J| = d)$, the supremum $g\left(\theta\right) = \sup\{f_i(\theta)\}$ is a subcompact function in the sense that its Lesbegue sets $L_g\left(C\right) = \{\theta : g\left(\theta\right) \leq C\}$ are compact for any real constant C (see Vandev and Neykov [26]).

A simpler and easier to apply criterion for subcompactness is given in Atanasov and Neykov [1], where it is proved that the real valued continuous function $g\left(\theta\right)$, defined on an open subset of $D \subset R^n$, is subcompact if and only if for any sequence $\theta_i \to \theta_0$ where θ_0 belongs to the boundary of D, $g\left(\theta_i\right) \to \infty$ when $i \to \infty$. Thus if D is a compact set, any continuous function, defined on D, is subcompact.

For the sake of completeness we draw the attention to the fact that the finite sample breakdown point of an estimator T at the finite sample $X = \{x_i; i = 1, \ldots, n\}$ is defined as the largest fraction m/n of observations, for which the $\sup_{\tilde{X}} \left\| T\left(X\right) - T\left(\tilde{X}\right) \right\|$ is finite, where $\tilde{X}$ is a sample obtained from X by replacing any m of the points in X by arbitrary values (see Hampel et al. [11], Rousseeuw and Leroy [20]).

4. Robust parametric estimation in BGWR processes. In this Section we construct robust estimators of the parameter of the BGWR process with power series offspring distribution, based on the entire family tree and on the generation sizes.

First we draw our attention to the fact that if a parametric maximum likelihood estimator exists, strongly consistent for the offspring mean $m = m(\theta)$ as a function of the unknown distribution parameter θ, then under mild regularity conditions a maximum likelihood estimator for an arbitrary function of m exists and it is strongly consistent when the function is continuous. This holds also for $\theta = \theta(m)$, considered as a function of m.

Note that the definition (1) of the BGWR process with PSOD has the following equivalent form:

Let $\Xi = \{\xi_i(t,n), i = 0, 1, 2, \ldots; t, n = 1, 2, \ldots; \xi\}$ be a set of i.i.d.r.v. with values in the set of nonnegative integers $N = \{0, 1, 2, \ldots\}$ with common probability distribution $p_k = P(\xi = k) = a_k \theta^k / A(\theta)$, $\theta > 0$, $a_k \geq 0$, $k = 0, 1, \ldots$, where $A(\theta) = \sum\limits_{k=0}^{\infty} a_k \theta^k$ is a positive function and $p_0 + p_1 < 1$. Let $\xi(0, n)$ be a r.v., independent of Ξ. Define for each $n = 1, 2, \ldots$ the process $\mathbf{Z}(\mathbf{n}) = \{Z_t(n), t = 0, 1, \ldots\}$ as

$$
\begin{aligned}
Z_0(n) &= \xi(0, n) \\
Z_1(n) &= \xi_1(1, n) + \xi_2(1, n) + \cdots + \xi_{Z_0(n)}(1, n) \\
Z_2(n) &= \xi_{Z_0(n)+1}(2, n) + \xi_{Z_0(n)+2}(2, n) + \cdots + \xi_{Z_0(n)+Z_1(n)}(2, n) \\
&\vdots \\
Z_{t+1}(n) &= \xi_{Z_0(n)+Z_1(n)+\cdots+Z_{t-1}(n)+1}(t+1, n) + \cdots + \xi_{Z_0(n)+Z_1(n)+\cdots+Z_t(n)}(t+1, n)
\end{aligned}
$$

Then $\mathbf{Z}(\mathbf{n})$ is a BGWR process.

Hence if $Y_t(n) = Z_0(n) + \cdots + Z_{t-1}(n)$ then

$$
Z_{t+1}(n) = \xi_{Y_t(n)+1}(t+1, n) + \cdots + \xi_{Y_{t+1}(n)}(t+1, n).
$$

In fact, $\xi_k(t, n)$ is the number of descendants, which live in the t-th generation, of the k-th particle, living in the $t - 1$-th generation.

Let

$$
\begin{aligned}
\vartheta_k(t) = \sharp\{j : \xi_j(s, n) = k; \\
j \in \{Z_0(n) + \cdots + Z_{s-2}(n) + 1, \ldots, Z_0(n) + \cdots + Z_{s-1}(n)\}, \\
s = 0, 1, 2, \ldots t - 1\}
\end{aligned}
$$

be the number of particles with k offspring among generations $0, 1, 2, \ldots, t - 1$.

Let

$$
\begin{aligned}
N(t) = \max\{k : \xi_j(s, n) = k, \\
j \in \{Z_0(n) + \cdots + Z_{s-2}(n) + 1, \ldots, Z_0(n) + \cdots + Z_{s-1}(n)\}; \\
s = 0, 1, 2, \ldots t - 1\}
\end{aligned}
$$

be the maximum number of offspring of one particle, observed in a single finite BGWR process realization of length t and

$$
\begin{aligned}
S(t) = \sharp\{k : \exists(j,s), \\
j \in \{Z_0(n) + \cdots + Z_{s-2}(n) + 1, \ldots, Z_0(n) + \cdots + Z_{s-1}(n)\}; \\
s = 0, 1, 2, \ldots t - 1 : \xi_j(s, n) = k\}
\end{aligned}
$$

the number of of the different observed amount of offspring of one particle.

Then

$$
\sum_{k=0}^{N(t)} k\vartheta_k(t) = \sum_{j=1}^{S(t)} k_j \vartheta_{k_j}(t) = \sum_{i=1}^{t} Z_i(n) = Y_t(n) - Y_1(n),
$$

(8)
$$
\sum_{k=0}^{N(t)} \vartheta_k(t) = \sum_{j=1}^{S(t)} \vartheta_{k_j}(t) = \sum_{i=0}^{t-1} Z_i(n) = Y_t(n).
$$

Now we prove the following

Lemma 1. *For the BGWR process with PSOD the MLE, based on the entire tree, coincides with the MLE, based on the generation sizes only.*

P r o o f. First note that according to (5)

$$
\begin{aligned}
\log L_{t+1}(\theta) &= \log\left(p_1(Z_0) \cdot \prod_{k=1}^{t} p\left(Z_0, \ldots, Z_k | \Im_k; \theta\right) \right) = \\
&= \log p_1(Z_0) + \log \theta^{\sum_{k=1}^{t} Z_k} - \log\left(A(\theta)\right)^{\sum_{k=1}^{t} Z_{k-1}} \\
&\quad + \log\left[\prod_{k=1}^{t} \sum_{s_1 + \cdots + s_{Z_{k-1}(n)} = Z_k(n)} \prod_{i=1}^{Z_{k-1}(n)} a_{s_i} \right]
\end{aligned}
$$

and $M_{t+1}(\theta) = \left[\sum\limits_{k=1}^{t} Z_k \cdot A(\theta) - \theta \sum\limits_{k=0}^{t-1} Z_k \cdot A'(\theta) \right] / [\theta A(\theta)]$

Hence, in order to obtain a MLE for the unknown parameter θ, based on the generation sizes only, one must solve the equation

(9)
$$
\sum_{k=1}^{t} Z_k \cdot A(\theta) - \theta \sum_{k=0}^{t-1} Z_k \cdot A'(\theta) = 0
$$

with respect to θ.

Let us consider the situation, when we are able to observe the entire family tree. The log likelihood has the form

$$L_{Y_t(n)}(\theta) = P(\xi(0,n)) \prod_{s=1}^{t} \prod_{i=Y_{s-1}(n)+1}^{Y_s(n)} P(\xi_i(s,n)) =$$

$$L_{S(t)}(\theta) = P(\xi(0,n)) P(\vartheta_{k_1}, \vartheta_{k_2}, \ldots, \vartheta_{k_{S(t)}}) =$$

$$= P(\xi(0,n)) \left(\frac{a_{k_1} \theta^{k_1}}{A(\theta)} \right)^{\vartheta_{k_1}} \cdot \left(\frac{a_{k_2} \theta^{k_2}}{A(\theta)} \right)^{\vartheta_{k_2}} \cdot \ldots \cdot \left(\frac{a_{k_{S(t)}} \theta^{k_{S(t)}}}{A(\theta)} \right)^{\vartheta_{k_{S(t)}}} =$$

$$= P(\xi(0,n)) \frac{\theta^{\sum_{k=0}^{N} k\vartheta_k}}{(A(\theta))^{\sum_{k=0}^{N} k\vartheta_k}} \cdot \prod_{k=0}^{N} a_k^{\vartheta_k}$$

$$= P(\xi(0,n)) \frac{\theta^{\sum_{k=1}^{t} Z_k}}{(A(\theta))^{\sum_{k=0}^{t-1} Z_k}} \cdot \prod_{k=0}^{N} a_k^{\vartheta_k};$$

$$M_{Y_t(n)}(\theta) = M_{S(t)}(\theta) = \frac{d}{d\theta} \log P(\xi(0,n)) + \frac{d}{d\theta} \log \theta^{\sum_{k=1}^{t} Z_k} -$$

$$- \frac{d}{d\theta} \log (A(\theta))^{\sum_{k=0}^{t-1} Z_k} + \frac{d}{d\theta} \log \prod_{k=0}^{N} a_k^{\vartheta_k}.$$

Here $k_1, \ldots, k_{S(t)}$ are the different numbers of descendants, observed in the sample. Again, in order to obtain a MLE for θ, based on the entire family tree, one must similarly solve the equation (9). This proves the Lemma.

Remark. As we have already shown, the MLE, based on the generation sizes, is strongly consistent under certain conditions. (see Proposition 1). Hence, using the entire family tree, under the same conditions a strongly consistent MLE can be obtained.

It follows from the properties of the $WLT(k)$ estimators that if a robust estimator exists, based on the entire family tree, it is strongly consistent if the classical m.l.e. is strongly consistent.

Note that the quantities $Y_t(n)$ or $S(t)$ describe the number of observations, included in the likelihood $L_{Y_t(n)}(\theta)$ and $L_{S(t)}(\theta)$ respectively. In the first likelihood we have at our disposal the number of offspring of any particle in

generation $0, 1$, or $t - 1$; in the second one the information only about the number of particles with k offspring is available (it does not take into account the number of offspring of a single particle). Let the r.v. $\{\xi_i(t, n)\}$ in the definition of a BGWR process be i.i.d. copies of the r.v. ξ, which takes values in the subset $B \subset N = \{0, 1, 2, \ldots\}$ and let us denote by $|B|$ the cardinality of this subset. Let $B_{SP} \subset B$ be the observed values of ξ in the sample. Denote by $\operatorname{Min} V = \min_i\{i : P(\xi = i) > 0, i \in B\}$ the minimal value of ξ and by $\operatorname{Max} V = \max_i\{i : P(\xi = i) > 0, i \in B\}$ the maximal value of ξ and let $N_{\operatorname{Max} V} = \sharp\{(i, s) : \xi_i(s, n) = \operatorname{Min} V, \; \xi_i(s, n) \in B_{SP}\}$ be the number of observations of the sample equal to $\operatorname{Min} V$ and respectively let $N_{\operatorname{Max} V} = \sharp\{(i, s) : \xi_i(s, n) = \operatorname{Max} V, \; \xi_i(s, n) \in B_{SP}\}$ be the number of observations of the sample equal to $\operatorname{Max} V$. Define the sets of functions $F = \{f_i(\theta) = A(\theta)/\theta^i\}_{i \in B}$, $F_{SP} = \{f_i(\theta) = A(\theta)/\theta^i\}_{i \in B_{SP}}$ and $LF_{SP} = \{\log(A(\theta)/\theta^i)\}_{i \in B_{SP}}$. Let R be the radius of convergence of the power series ($A(\theta) < \infty$ if $\theta \in (0, R)$ and $A(\theta) = \infty$ if $\theta \in (R, \infty)$) and K be the trimming factor of the $WLT(K)$ estimator. We prove the following

Proposition 2. *Let us consider a sample with $Y_t(n)$ observations based on the entire family tree over a BGWR process with PSOD. Let the $WLT(K)$ estimator, defined with (7), be based on the log-likelihood $M_{Y_t(n)}(\theta)$. Then the following statements are valid:*

1. If $\theta \in (0, \infty)$, $R = \infty$, $|B| = \infty$ and $Y_t(n) > N_{\operatorname{Min} V} + 1$ then the set LF_{SP} is $N_{\operatorname{Min} V} + 1$-full, the $WLT(K)$ estimator exists and its breakdownpoint is not less than $[Y_t(n) - K]/Y_t(n)$ if $Y_t(n) \geq 3(N_{\operatorname{Min} V} + 1)$, $[Y_t(n) + N_{\operatorname{Min} V} + 1]/2 \leq K \leq Y_t(n) - N_{\operatorname{Min} V} - 1$.

2. If $\theta \in (0, R)$, $R < \infty$, $|B| = \infty$, $A(R) = \infty$ and $Y_t(n) > N_{\operatorname{Min} V} + 1$ then the conclusion of situation 1. is valid.

3. If $\theta \in (0, R)$, $R < \infty$, $|B| = \infty$ and $A(R) < \infty$ then the set LF_{SP} is not d-full for any $d = 1, 2, \ldots$.

4. If $\theta \in (0, \infty)$ and $|B| < \infty$, $Y_t(n) > \max\{N_{\operatorname{Min} V}, N_{\operatorname{Max} V}\} + 1$ then the set LF_{SP} is $\max\{N_{\operatorname{Min} V}, N_{\operatorname{Max} V}\} + 1$-full, the $WLT(K)$ estimator exists and its breakdownpoint is not less than $[Y_t(n) - K]/Y_t(n)$ if $Y_t(n) \geq 3(\max\{N_{\operatorname{Min} V}, N_{\operatorname{Max} V}\} + 1)$, $[Y_t(n) + \max\{N_{\operatorname{Min} V}, N_{\operatorname{Max} V}\} + 1]/2 \leq K \leq Y_t(n) - \max\{N_{\operatorname{Min} V}, N_{\operatorname{Max} V}\} - 1$.

5. If $\theta \in [a, b] \subset (0, R)$, where R is a positive real number or infinity, then the set LF_{SP} is 1-full, the $WLT(K)$ estimator exists and its breakdownpoint is not less than $[Y_t(n) - K]/Y_t(n)$ if $Y_t(n) \geq 3$, $[Y_t(n) + 1]/2 \leq K \leq Y_t(n) - 1$.

In all these situations except Case 3 the $WLT(K)$ estimator is consistent

under the conditions of Proposition 1.

P r o o f. Note that

$$-\log L_{Y_t(n)} = \sum_{s=1}^{t} \sum_{i=Y_{t-1}(n)}^{Y_t(n)} [\log(A(\theta)) - \xi_i(s,n)\log\theta - \log(a_{\xi_i(s,n)})] - \log P(\xi(0,n)),$$

$$-\log L_{S(t)} = \sum_{i=1}^{S(t)} \vartheta_{j_i}[\log(A(\theta)) - \log(a_{j_i}) - j_i\log\theta] - \log P(\xi(0,n));$$

$$\widehat{\theta}_{Y_t(n)} = \operatorname*{argmin}_{\theta} \sum_{i=1}^{K} \omega_i[\log(A(\theta)) - \xi_{\nu(i)}\log\theta],$$

$$\widehat{\theta}_{S(t)} = \operatorname*{argmin}_{\theta} \sum_{i=1}^{K} \omega_i\vartheta_{\nu(j_i)}[\log(A(\theta)) - \nu(j_i)\log\theta],$$

where ν denotes the permutation of the indices, according to (7). Therefore one must study the subcompactness of the set of functions LF_{SP} using the properties of the functions in the set F_{SP}.

According to the definition of power series offspring distribution

$$f_k(\theta) = \frac{a_k}{p_k} = \frac{\sum_{s=0}^{k-1} a_s\theta^s}{\theta^k} + a_k + \frac{\sum_{s=k+1}^{\infty} a_s\theta^s}{\theta^k}$$

We need to calculate whether $\lim_{\theta\to 0}[\log A(\theta)/\theta^i] = \lim_{\theta\to R}[\log A(\theta)/\theta^i] = \infty$ for every $i \in B_{SP}$, where $R = \infty$ or $R < \infty$. In this case the set LF_{SP} would be 1-full. The possible exceptions may appear in the values $\operatorname{Max} V$ and $\operatorname{Min} V$. We use the fact that if a function is bigger than a subcompact function then the function itself is subcompact. Note that the average of functions is always smaller than their supremum.

In Case 1 $\log A(\theta)/\theta^{\operatorname{Min} V}$ is a strictly increasing function in θ, but $\lim_{\theta\to 0}\frac{1}{2}[\log A(\theta)/\theta^{\operatorname{Min} V} + \log A(\theta)/\theta^i] = \lim_{\theta\to\infty}\frac{1}{2}[\log A(\theta)/\theta^{\operatorname{Min} V} + \log A(\theta)/\theta^i] = \infty$ for every $i \in B/\{\operatorname{Min} V\}$.

Case 2 is analogous to Case 1.

In Case 3 $\lim_{\theta\to R} f_k(\theta) = A(R)/R^k < \infty$ for any $k \in B$.

In Case 4 $A(\theta)/\theta^{\operatorname{Min} V}$ is strictly increasing and $A(\theta)/\theta^{\operatorname{Max} V}$ is strictly

decreasing (but differs from zero), but

$$\lim_{\theta \to 0} 1/2[\log A(\theta)/\theta^{\mathrm{Min}\,V} + \log A(\theta)/\theta^{\mathrm{Max}\,V}] =$$

$$= \lim_{\theta \to \infty} 1/2[\log A(\theta)/\theta^{\mathrm{Min}\,V} + \log A(\theta)/\theta^{\mathrm{Max}\,V}] =$$

$$= \lim_{\theta \to 0} 1/2[\log A(\theta)/\theta^{\mathrm{Min}\,V} + \log A(\theta)/\theta^{i}] =$$

$$= \lim_{\theta \to \infty} 1/2[\log A(\theta)/\theta^{\mathrm{Min}\,V} + \log A(\theta)/\theta^{i}] =$$

$$= \lim_{\theta \to 0} 1/2[\log A(\theta)/\theta^{\mathrm{Max}\,V} + \log A(\theta)/\theta^{i}] =$$

$$= \lim_{\theta \to \infty} 1/2[\log A(\theta)/\theta^{\mathrm{Max}\,V} + \log A(\theta)/\theta^{i}] = \infty$$

for every $i \in B/\{\mathrm{Min}\,V \bigcup \mathrm{Max}\,V\}$.

For the last part of the Proposition one should notice that LF_{SP} consists of continuous functions with values in a closed subset of the positive part of the real line and therefore any function of LF_{SP} is subcompact.

The consistency follows from Lemma. This proves the proposition. □

Example 1.   Poisson offspring distribution $Po(\lambda)$. In this situation $p_k(\lambda) = \lambda^k e^{-\lambda}/k! = [\lambda^k/k!]/[\sum \lambda^k/k!]$, $A(\lambda) = e^\lambda$, $\lambda \in (0, \infty)$, $R = \infty$, $|B| = \infty$, $A(\infty) = \infty$. One obtains Case 1 of Proposition 2 and therefore the set LF_{SP} is $N_0 + 1$-full, where N_0 is the number of observed particles, which do not have any offspring.

Let us suppose that we have not taken into account the requirement in the Proposition about the trimming factor and we have set it to N_0. Then it could happen that among all subsets of the sample the optimization algorithm might choose the subset, consisting of observation values equal to zero. Then the demanded minimum would be zero (note that zero does not belong to the definition set of the studied parameter).

Example 2.   Geometric offspring distribution $Ge(1 - \theta)$. Here $p_k(\theta) = \theta^k(1 - \theta) = \theta^k/[1/(1 - \theta)] = \theta^k/[\sum \theta^s]$, $\theta \in (0, 1)$, $R = 1$, $A(\theta) = 1/(1 - \theta)$, $A(1) = \infty$, $|B| = \infty$. One obtains Case 2 of Proposition 2 and therefore the set LF_{SP} is $N_0 + 1$-full.

Example 3.   Binomial offspring distribution $Bi(n, p)$. Here $p_k(p) = \binom{n}{k}p^k(1 - p)^{n-k} = \binom{n}{k}(p/(1 - p))^k/[1/(1 - p)]^n = \binom{n}{k}(p/(1 - p))^k/[1 + p/(1 - p)]^n = \binom{n}{k}(p/(1 - p))^k/\left[\sum_{s=0}^{n}\binom{n}{s}(p/(1 - p))^s\right]$, $p/(1 - p) \in (0, \infty)$, $|B| < \infty$,

$A(p/(1-p)) = [1 + p/(1-p)]^n$, $1/p_0(p)$ is strictly increasing, $1/p_n(p)$ is strictly decreasing in p and $p/1 - p$, one obtains Case 4 from Proposition 2 and if N_n is the number of particles with n descendants, the set LF_{SP} is $\max\{N_0, N_n\}+1$-full.

Remark. Note that the set LF, consisting of the logarithms of the functions in F, is 2-full in situations 1, 2 and 4 of Proposition 2. We study the index of fullness of the set LF_{SP} due to the fact that the trimming factor is defined only over a concrete sample.

Similarly the following proposition can be proved:

Proposition 3. *Let us consider a sample with $S(t)$ observations from a BGWR process with PSOD. The observations present the number of particles with one and the same number of offspring. Let the WLT(K) estimator, defined with (7), be based on the log-likelihood $M_{S(t)}(\theta)$. Then the following statements are valid:*

1. If $\theta \in (0, \infty)$, $R = \infty$ and $|B| = \infty$ or $\theta \in (0, R)$, $R < \infty$, $|B| = \infty$ and $A(R) = \infty$ or $\theta \in (0, \infty)$ and $|B| < \infty$, then the set $\hat{LF}_{SP} = \left\{\log\left(A(\theta)/\theta^i\right)^{\nu_i}\right\}_{i \in B_{SP}}$ is 2-full, the WLT(K) estimator exists and its breakdownpoint is not less than $[S(t) - K]/S(t)$ if $S(t) \geq 6$, $[S(t) + 2]/2 \leq K \leq S(t) - 2$.

2. If $\theta \in (0, R)$, $R < \infty$, $|B| = \infty$ and $A(R) < \infty$ then the set $\hat{LF}_{SP}$ is not d-full for any $d = 1, 2, \ldots$.

3. If $\theta \in [a, b] \subset (0, R)$, where R is a positive real number or infinity, then the set $\hat{LF}_{SP}$ is 1-full, the WLT(K) estimator exists and its breakdownpoint is not less than $[S(t) - K]/S(t)$ if $S(t) \geq 3$, $[S(t) + 1]/2 \leq K \leq S(t) - 1$.

In situations 1 and 3 the WLT(K) estimator is consistent under the conditions of Proposition 1.

Let us now suppose that we are not able to observe the entire family tree, and that information only about the generation sizes is available. The proposed method for constructing robust estimators uses several sample paths over the process.

Let us have at our disposal r independent realizations $Z^{(i)}(n)$, $i = 1, \ldots, r$, from a BGWR process with one and the same power series offspring distribution and number of generations, equal to t_i. Let $L_{t_i}(\theta)$ be the likelihood function, based on the generation sizes in the $i-$th realization. Define the set $\tilde{F} = \left\{\sum_{i=0}^{t_k-1} Z_i^{(k)}(n) \cdot \log(A(\theta)) - \sum_{i=1}^{t_k} Z_i^{(k)}(n) \cdot \log\theta\right\}_{k=1}^{r}$ and the variable $\tilde{N}_{\text{Min}\,V} = \sharp\{i : \hat{m}_{t_i}(n) \leq \text{Min}\,V; i = 1, \ldots, r\}$, where $\hat{m}_{t_i}(n)$ is the Harris estimator over the i-th sample path. We prove the following

Proposition 4. *Let us consider a set of r sample paths from a BGWR process with PSOD. Let the sample paths be based on the generation sizes of the process. Suppose that the power series has the following properties: $\theta \in (0, \infty)$ and $|B| = \infty$; or $\theta \in (0, R)$, $R < \infty$, $|B| = \infty$ and $A(R) = \infty$. Then the set $\tilde{F}$ is $\tilde{N}_{\text{Min }V} + 1$-full, the WLT(K) estimator over all r realizations of the process exists and its breakdownpoint is not less than $[r - K]/r$ if $r \geq 3(\tilde{N}_{\text{Min }V} + 1)$, $[r + \tilde{N}_{\text{Min }V} + 1]/2 \leq K \leq r - \tilde{N}_{\text{Min }V} - 1$, where K is the trimming factor. If all r sample paths consist of t consecutive generation sizes $(t_1 = t_2 = \cdots = t_r = t)$ then the WLT(K) estimator is consistent when $t \to \infty$ under the conditions of Proposition 1.*

P r o o f. The likelihood function, based on all r realizations, has the form:

$$
L_r(\theta) = P(Z^{(1)}(n), Z^{(2)}(n), \ldots, Z^{(r)}(n)) =
$$
$$
= P(Z_0^{(1)}(n), Z_1^{(1)}(n), \ldots, Z_{t_1}^{(1)}(n)) \ldots P(Z_0^{(r)}(n), Z_1^{(r)}(n), \ldots, Z_{t_r}^{(r)}(n)) =
$$

$$
= P(Z_0^{(1)}(n)) \cdot \frac{\theta^{\sum_{i=1}^{t_1} Z_i^{(1)}(n)}}{(A(\theta))^{\sum_{i=0}^{t_1-1} Z_i^{(1)}(n)}} \cdot \prod_{i=1}^{t_1} \left[\sum_{\substack{s_1+s_2+\cdots+s_{Z_{i-1}^{(1)}(n)} = Z_i^{(1)}(n)}} \prod_{j=1}^{Z_{i-1}^{(1)}(n)} a_{s_j} \right] \ldots
$$

$$
\cdots \mathbf{P}(Z_0^{(r)}(n)) \cdot \frac{\theta^{\sum_{i=1}^{t_r} Z_i^{(r)}(n)}}{(A(\theta))^{\sum_{i=0}^{t_r-1} Z_i^{(r)}(n)}} \cdot \prod_{i=1}^{t_r} \left[\sum_{\substack{s_1+s_2+\cdots+s_{Z_{i-1}^{(r)}(n)} = Z_i^{(r)}(n)}} \prod_{j=1}^{Z_{i-1}^{(r)}(n)} a_{s_j} \right] =
$$

$$
= \prod_{i=1}^{r} \mathbf{P}(Z_0^{(i)}(n)) \cdot \prod_{l=1}^{r} \left\{ \prod_{i=1}^{t_1} \left[\sum_{\substack{s_1+s_2+\cdots+s_{Z_{i-1}^{(l)}(n)} = Z_i^{(l)}(n)}} \prod_{j=1}^{Z_{i-1}^{(l)}} a_{s_j} \right] \right\} \cdot
$$

$$
\cdot \prod_{i=1}^{r} \frac{\theta^{\sum_{j=1}^{t_i} Z_j^{(i)}(n)}}{(A(\theta))^{\sum_{j=0}^{t_i-1} Z_j^{(i)}(n)}}.
$$

But

$$
\prod_{i=1}^{r} \mathbf{P}(Z_0^{(i)}(n)) \cdot \prod_{l=1}^{r} \left\{ \prod_{i=1}^{t_1} \left[\sum_{\substack{s_1+s_2+\cdots+s_{Z_{i-1}^{(l)}(n)} = Z_i^{(l)}(n)}} \prod_{j=1}^{Z_{i-1}^{(l)}(n)} a_{s_j} \right] \right\}
$$

does not depend on θ. Therefore

$$
\begin{aligned}
M_r(\theta) &= \frac{d}{d\theta} \sum_{k=1}^{r} \left[\sum_{i=0}^{t_k-1} Z_i^{(k)}(n) \cdot \log(A(\theta)) - \sum_{i=1}^{t_k} Z_i^{(k)}(n) \cdot \log\theta \right] = \\
&= \frac{d}{d\theta} \left[\sum_{k=1}^{r} \sum_{i=0}^{t_k-1} Z_i^{(k)}(n) \cdot \log(A(\theta)) - \sum_{k=1}^{r} \sum_{i=1}^{t_k} Z_i^{(k)}(n) \cdot \log\theta \right]
\end{aligned}
$$

(10)

and when $t_1 = t_2 = \cdots = t_r = t$ the MLE, based on the r realizations, coincides with the MLE, based on one BGWR process realization, starting with $\sum_{k=1}^{r} Z_0^{(k)}(n)$ ancestors and having t generations. This proves the consistency. The proof of the fullness of the set $\tilde{F}$ is analogous to Cases 1 and 2 of Proposition 2, taking into account that

$$
A(\theta)^{\sum_{k=0}^{t_i} Z_k^{(i)}(n)} \Big/ \theta^{\sum_{k=1}^{t_i-1} Z_k^{(i)}(n)} = \left[\frac{\sum a_k \theta^k}{\theta^{\hat{m}_{t_i}(n)}} \right]^{\sum_{k=0}^{t_i} Z_k^{(i)}(n)} \; ; \;\; i = 1, \dots, r.
$$

Example 4. Let us consider a sample of r sample paths of BGWR processes. Let k of the processes, $0 \leq k \leq r$, have Poisson offspring distribution $Po(\lambda)$ and $r - k$ of the processes have a Geometric distribution $Ge(\lambda)$. Then the set $\tilde{F}$ is $\tilde{N} + 1$-full, where $\tilde{N}$ is the number of sample paths, which become extinct at the first generation after the generation of the ancestors.

5. Computational results. To calculate the value of the estimate, defined with (7), we use an algorithm considered in [1] and applied in the way described in [22] . The algorithm can be summarized as follows:

1. Setting the initial value for the unknown parameter $\theta = \theta_0 \in (0, \infty)$.

2. Sorting the observations according to the log-density function at the current value of the unknown parameter : $f_{\nu(1)}(\theta) \leq f_{\nu(2)}(\theta) \leq \ldots \leq f_{\nu(n)}(\theta)$.

3. The weights are equal to 1.

4. Finding the value which satisfies (7).

5. If the exit conditions are not satisfied than go back to 2.

The optimization algorithm in 4. is based on Golden Section search, parabolic interpolation and Nelder-Mead simplex (direct search) method (see [18]).

To illustrate the advantages of the robust estimator, defined above, we did a number of simulations of BGWR processes with PSOD and known parameter value. We replaced randomly 30% of the generated data with outliers. Then we used the formulated computational procedure, to obtain an estimated value of the parameter.

We set all the weights in (7) equal to 1. In this way we obtain a Least Trimmed Estimator. As is shown in [1], in this case we have a faster and more simple algorithm and a smaller value of the standard error of the estimate.
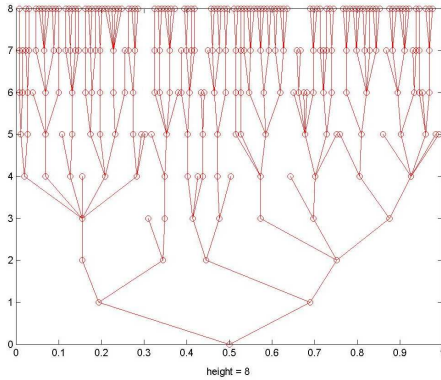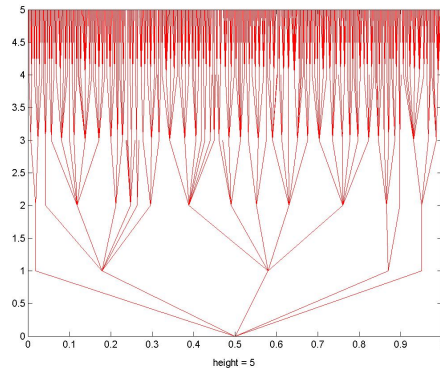

Figure 1

Figure 2

real value	1.5
generations	9
$Y_t(n)$	161
outliers	30%
trimming factor	40
Estimate	1.4425
St. error	0.11298

Result 1

real value	3
generations	6
$Y_t(n)$	247
outliers	30%
trimming factor	40
Estimate	2.867
St. error	0.12878

Result 2

Let us first consider the robust estimator over the entire family tree. On the Figure 1 a simulation of a BGWR process is shown, having 9 generations,

1 ancestor and Poisson offspring distribution $Po(1.5)$ The 30% outliers are put randomly in the family tree and represent 100 offspring of one particle instead of the real number of offspring.

The situation is similar on Figure 2, where a BGWR process is shown, having 6 generations, 1 ancestor, Poisson offspring distribution $Po(3)$ and 30% outliers of size 100.

We have calculated the robust estimates of the parameter of the Poisson distribution and the standard error.

Let us consider the situation, when we are able to observe the genetation sizes only.We have simulated 10 sample paths of BGW processes with Poisson offspring distribution and offspring mean values shown in the first rows of the tables given below and with 5 outliers with Poisson offspring distribution with mean 2 (supercritical situation on Table 1) and mean 0.5 (subcritical on Table 2).

offspring mean	estimate
0.8	0.8139
0.9	0.9500
1.0	1.1020
1.1	1.1257
1.3	1.2169
1.5	1.4103
1.8	1.8052

Table 1.

offspring mean	estimate
0.7	0.7222
0.9	0.8657
1.1	1.0597
1.5	1.4726
1.8	1.8227

Table 2.

Comment: The estimates behave adequately and cope successfully with the presence of outliers. At a comparatively small number of observations and ancestors and a large number of errors reasonable estimates of the needed parameter can be obtained. The standard error lies in the borders of 10% of the estimates, which is to be expected (see f.e. [1]). Relatively 'worse' estimates appear when estimating the parameter values equal to 1.3 and 1.5 from Table 1. This might be due to the fact that the outliers' distribution (and respectively outliers' values) is close to that of the correct data and this can perplex the estimating procedure when eliminating the incorrect data.

It is worth mentioning that the robust parametric estimation over one single realization of a BGWR process, based on the generation sizes only, is still an open question. Due to the fact that the transition probabilities, which take

part in the log likelihood function, depend on two consecutive generation sizes, the technique for calculating the $WLT(k)$ estimators might probably differ from the just proposed one.

REFERENCES

[1] D. Atanasov, N. Neykov. On the Finite Sample Breakdown Point of the $WLTE(k)$ and d-fullness of a Set of Continuous Functions. Proceedings of the VI International Conference "Computer Data Analysis And Modeling", Minsk, Belarus, 2002.

[2] L. Denby, R. L. Martin. Robust Estimation on the First Order Autoregressive Parameter. *J. Amer. Statist. Assoc.* **74** (1979), 140–146.

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

[5] J.-P. Dion, N. M. Yanev. Estimation Theory for Branching Processes with or without Immigration. *C. R. Acad. Bulgare Sci.* **44**, *4* (1991), 19–22.

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

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

[8] J. I. Gastwirth, H. Rubin. The Behaviour of Robust Estimators on Dependent Data. *Ann. Statist.* **3** (1975), 1070–1100.

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

[10] A. HADI, A. LUCEÑO. Maximum Trimmed Likelihood Estimators: A Unified Approach, Examples and Algorithms. *Comput. Statist. Data Anal.* **25** (1997), 251–272.

[11] F. HAMPEL, E. RONCHETTI, P. ROUSSEEUW, W. STAHEL. Robust Statistics: The Approach Based on Influence Functions. John Wiley and Sons, New York, 1986.

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

[13] P. J. HUBER. Robust Statistics. Wiley, New York, 1981.

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

[15] N. KEIDING, S. LAURITZEN. Marginal Maximum Likelihood Estimates and Estimation of the Offspring Mean in Branching Processes. *Scand. J. Statist.* **5** (1978), 106–110.

[16] P. M. KULKARNI, C. C. HEYDE. Optimal Robust Estimation for Discrete Time Stochastic Processes. *Stochastic Process. Appl.* **26** (1987), 267–276.

[17] H. KÜNSCH. Infinitesimal robustness for autoregressive processes. *J. Appl. Prob.* **33** (1984), 1061–1076.

[18] J. LAGARIAS, J. REEDS, M. WRIGHT, P.WRIGHT. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM J. Optim.* **9** (1998), 112–147.

[19] R. D. MARTIN. The Cramer-Rao Bound and Robust M-estimates for autoregressions. *Biometrika* **69** (1982), 437–442.

[20] P. ROUSSEEUW, A. LEROY. Robust Regression and Outlier Detection. John Wiley and Sons, New York, 1981.

[21] V. STOIMENOVA, D. ATANASOV, N. YANEV. Robust Estimation and Simulation of Branching Processes. *C. R. Acad. Bulgare Sci.* **5** (2004), 19–22

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

[23] V. Stoimenova, D. Atanasov, N. Yanev. Algorithms for Generation and Robust Estimation of Branching Processes with a random number of Ancestors. *Math. and Education in Math.* **34** (2005), 196–201.

[24] V. Stoimenova, N. Yanev. Parametric Estimation in Branching Processes with an Increasing Random Number of Ancestors. *Pliska Stud. Math. Bulgar.* (2005), in print.

[25] D. Vandev. A Note on Breakdown Point of the Least Median of Squares and Least Trimmed Estimators. *Statist. Probab. Lett.* **16** (1993), 117–119.

[26] D. Vandev, N. Neykov. Robust Maximum Likelihood in the Gaussian Case. In: New Directions in Statistical Data Analysis and Robustness (Eds S. Morgenthaler, E. Ronchetti, W. A. Stahel) Birkhauser Verlag, Basel, 1993.

[27] D. Vandev, N. Neykov. About Regression Estimators with High Breakdown Point. *Statistics* **32** (1998), 111–129.

[28] A. Yakovlev, N. M. Yanev. Transient processes in cell proliferation kinetics. Lecture Notes in Biomathematics, vol. **82**, Berlin, Springer-Verlag, 1989.

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

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

Sofia University
Faculty of Mathematics and Informatics
5, J. Boucher Str.
1164 Sofia, Bulgaria
e-mail: `stoimenova@fmi.uni-sofia.bg`   *Received April 24, 2005*