PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

Pliska Stud. Math. Bulgar. **20** (2011), 203–220

PLISKA
STUDIA MATHEMATICA
BULGARICA

ROBUST ESTIMATION IN MULTITYPE BRANCHING PROCESSES BASED ON THEIR ASYMPTOTIC PROPERTIES*

Vessela Stoimenova, Dimitar Atanasov

In this work we propose two procedures for robust estimation of the individual distributions of multitype discrete-time Galton-Watson branching processes with an increasing number of ancestors, using the relative frequencies of the process and their asymptotic distributions. The study is based on simulations and numerical results.

1. Introduction

In the present paper we consider some aspects of the robust estimation in discrete-time multitype Galton-Watson branching processes with a large (and increasing) number of ancestors (MGWL processes).

The general formulation and handling of branching processes with several types of particles was first introduced by Kolmogorov and Dmitriev (1947) and Kolmogorov and Sevastyanov (1947) in the Markov case. Since then there is an impressive number of work in the area of branching processes theory and applications (see f.e. the books of Asmussen and Herring, 1983, Athreya and

*The research was partially supported by appropriated state funds for research allocated to Sofia University (contract 112/2010), Bulgaria.

2000 *Mathematics Subject Classification:* 60J80.

Key words: Trimmed likelihood, multitype branching process, asymptotic properties, parameter estimation.

Ney, 1972, Harris, 1963, Jagers, 1975, Sevastyanov, 1971, Yakovlev and Yanev, 1989 and others).

Yakovlev and Yanev (1989) noted that branching processes with a large and often random number of ancestors may be usefull for modelling purposes in the study of cell proliferation as well as in applications to nuclear chain reactions. Statistical inference for Bienaymé-Galton-Watson processes with an increasing random number of ancestors (BGWR processes) was introduced and developed by Yanev (1975) and Dion and Yanev (1991, 1992, 1994, 1997). Robustified versions in the sense of the weighted and trimmed likelihood of the classical estimators are proposed in Stoimenova, Atanasov, Yanev (2004 a, b, 2005), Stoimenova and Atanasov (2006). In the class of the power series offspring distributions some topics of the parametric estimation are considered in Stoimenova, Yanev (2005) and of the robust parametric estimation - in Stoimenova (2005). The effectiveness of the estimators of Dion and Yanev of the individual and immigration mean in discrete-time branching processes with immigration, based on their relationship to the BGWR processes, is studied in Atanasov, Stoimenova, Yanev (2007, 2009), where also their robust modifications are proposed.

The asymptotic behaviour of multitype Markov branching processes is considered in Yakovlev and Yanev (2010) and the usage of the obtained limiting results for cell kinetics studies is shown. Applications in the area of cell biology were also a motivation for Yakovlev and Yanev (2009) to consider and study the relative frequencies of distinct types of particles in MBPR. We base ourselves on asymptotic results from these papers to construct robust estimators of the individual distributions.

Within the present paper under robustness we mean weighted and trimmed likelihood (WLT(k) estimators), defined by Vandev and Neykov (1998) and based on the principle of the maximum likelihood estimation.

2. Multitype Galton-Watson processes – notations and overview of the preliminary results

In the multitype Galton-Watson processes (or as we reffer to MGW processes) we allow for the existence of distinguishable particles (individuals, cells depending on the interpretation) with different probabilistic behaviour. To each particle we assign a type in the set of types $T = \{1, 2, \dots, d\}$ that is assumed to be finite and with cardinality d. Each particle, say the l-th particle of type $k \in T$ living in the $t-$th generation (t= 0,1,2,...), is associated with a random vector $\overrightarrow{\xi}_k(t,l) = (\xi_k^1(t,l), \dots, \xi_k^d(t,l))$, where $\xi_k^j(t,l)$ is a random variable that represents the number of children of type j, $j \in T$, in the generation $t+1$, produced from

the k-th type l particle in the generation t. The distribution of the random vector $\overrightarrow{\xi}_k(t,l)$ does not depend on the generation, where the parent particle lives, and on the index l. The offspring of the particles in the generation t forms the next generation $t+1$. Hence a MGW process is defined as a sequence of random vectors $\{\mathbf{Z}(t) = (Z_1(t), \dots, Z_d(t))\}$, $t \in N_0 = \{0, 1, 2, \dots\}$, where $Z_k(t)$ represents the number of particles of type $k \in T$ in generation t, $Z_k(t+1) = \sum_{j=1}^{d} \sum_{l=1}^{Z_l(t)} \xi_j^k(t,l)$.

We denote by

$$h_i(s) = E[\mathbf{s}^{\mathbf{Z}(1)} | Z_i(0) = 1] = E[s_1^{Z_1(1)} \dots s_d^{Z_d(1)} | Z_1(0) = 1]$$

the offspring probability generating function of the MGW process, starting with one particle of type $i \in T$, and by

$$F^i(t,s) = E[\mathbf{s}^{\mathbf{Z}(t)} | Z_i(0) = 1] = E[s_1^{Z_1(t)} \dots s_d^{Z_d(t)} | Z_1(0) = 1]$$

the probability generating function of the process in the moment t, starting with one particle of type $i \in T$. Here $\mathbf{s} = (s_1, \dots, s_d)$ and $|s_k| \le 1$, $k = 1, 2, \dots, d$.

We use the following notations for the first and second moments of the offspring distribution:

$$\begin{aligned} m_{ij} &= \frac{\partial}{\partial s_j} h_i(\mathbf{s})|_{\mathbf{s}=\mathbf{1}} = E[Z_j(1) | Z_i(0) = 1], \\ b_{jk}^i &= \frac{\partial^2}{\partial s_j \partial s_k} h_i(\mathbf{s})|_{\mathbf{s}=\mathbf{1}} = E[Z_j(1)[Z_k(1) - \delta_{jk}] | Z_i(0) = 1], \end{aligned}$$

where $i, j, k = \overline{1, d}$, $\delta_{jk} = 0$, if $j \neq i$, $\delta_{jk} = 1$, if $j = i$, and $\mathbf{1} = (1, \dots, 1)$.

If we denote by $\mathbf{M} = ||m_{ij}||$, then it is well known that by the branching property and the independence of individual particle evolutions

$$\begin{aligned} \mathbf{M}(t) &:= ||m_{ij}(t)|| = ||E[Z_j(t) | Z_i(0) = 1]|| = ||m_{ij}||^t = \mathbf{M}^t, \\ m_{ij}(t) &:= E[Z_j(t) | Z_i(0) = 1] = \frac{\partial}{\partial s_j} F^i(t, \mathbf{s})|_{\mathbf{s}=\mathbf{1}}, \\ b_{jk}^i(t) &:= \frac{\partial^2}{\partial s_j \partial s_k} F^i(t, \mathbf{s})|_{\mathbf{s}=\mathbf{1}} = E[Z_j(t)[Z_k(t) - \delta_{jk}] | Z_i(0) = 1], \quad i, j, k = \overline{1, d}. \end{aligned}$$

Let us now suppose that the MGW process starts with one particle of type 1. We need the following notations:

$$\sigma_k^2(t) := Var[Z_k(t) | Z_1(0) = 1] = b_{kk}^1(t) + m_{1k}(t) - (m_{1k}(t))^2,$$

$$\begin{aligned} C_{ij}(t) &:= Cov[Z_i(t), Z_j(t)|Z_1(0)=1] = b^1_{ij}(t) - m_{1i}(t)m_{1j}(t), \\ r_{ij}(t) &:= Corr[Z_i(t), Z_j(t)|Z_1(0)=1] = C_{ij}(t)/\sigma_i(t)\sigma_j(t), \;\; i,j,k,= \overline{1,d}. \end{aligned}$$

Hence $C_{ii}(t) = \sigma_i^2(t)$ and $r_{ii}(t) = 1$ for $i = \overline{1,d}$.

We suppose that the covariance matrix $\mathbf{C}(t) = ||C_{ij}(t)||$ and the correlation matrix $\mathbf{R}(t) = ||r_{ij}(t)||$ are finite and well defined.

Let $U(t) = \sum\limits_{k=1}^{d} Z_k(t)$ be the total number of particles at the moment t and the relative frequencies (or fractions, proportions) of types $\Delta_k(t) = \dfrac{Z_k(t)}{U(t)}$ be defined on the set of nonextinction $\{U(t) > 0\}$. One should notice, that there exists the following obvious relationship between the relative frequencies: $\sum\limits_{k=1}^{d} \Delta_k(t) = 1$.

We need also the following notation about the theoretical proportions $p_i(t) := \dfrac{m_{1i}(t)}{M(t)}$, where $M(t) = EU(t) = \sum\limits_{j=1}^{d} m_{1j}(t)$, which, as noted in Yakovlev and Yanev (2009), may be interpreted as the probability that a randomly chosen cell at time t is of type i.

Let us now consider the multitype Galton-Watson branching process starting with $Z_1(0) = N$ initial number of ancestors.

Then the relative frequencies can be written as

$$\Delta_i(t,N) = \frac{Z_i(t,N)}{U(t,N)} = \frac{\sum\limits_{k=1}^{N} Z_i^{(k)}(t)}{\sum\limits_{k=1}^{N} U^{(k)}(t)},$$

where due to the independence of cell evolutions $\{Z_i^{(k)}(t)\}_{k=1}^{N}$ are iid copies of the process $\{Z_i(t),\ i = \overline{1,d}\}$ and $U^{(k)}(t) = \sum\limits_{i=1}^{d} Z_i^{(k)}(t)$.

According to the notations introduced in Yakovlev and Yanev (2009, 2010) let

$$\begin{aligned} a_{ij}(t) &= \begin{cases} \sigma_i(t)(1-p_i(t)) & \text{if} \quad i = j \\ -\sigma_i p_j(t) & \text{if} \quad i \neq j, \;\; i,j = \overline{1,d} \end{cases}, \\ W_i(t,N) &= M(t)\sqrt{N}[\Delta_i(t,N) - p_i(t)], \\ V_i(t,N) &= \sum_{k=1}^{N} \frac{Z_i^{(k)}(t) - m_{1i}(t)}{\sigma_i(t)\sqrt{N}} = \frac{Z_i(t,N) - Nm_{1i}(t)}{\sigma_i(t)\sqrt{N}}. \end{aligned}$$

In Yakovlev and Yanev (2009), Proposition 1, it is shown that if $m_{1i}(t) < \infty$, $i = \overline{1,d}$, then the relative frequencies $\Delta_i(t,N)$ are strongly consistent and asymptotically unbiased estimators for the proportions $p_i(t)$ when the initial number of ancestors $N \to \infty$. Moreover, when $\sigma_i^2(t) < \infty$, $i = \overline{1,d}$, their multivariate normality is proved (Yakovlev and Yanev, 2009, Theorem 1) and as a consequence one has that

$$W_i(t,N) \quad \stackrel{d}{\longrightarrow} \quad Y_i(t), \quad N \to \infty \tag{1}$$

for every $i = \overline{1,d}$, where $Y_i(t)$ is a normally distributed centered random variable with

$$S_i^2(t) \quad = \quad VarY_i(t) = \sum_{k,l=1}^{d} r_{kl}(t)a_{ki}(t)a_{li}(t) \tag{2}$$

and

$$(V_1(t,N),\ldots,V_d(t,N)) \quad \stackrel{d}{\longrightarrow} \quad (X_1(t),\ldots,X_d(t)), \quad N \to \infty, \tag{3}$$

where the random variables $(X_1(t),\ldots,X_d(t))$ have a joint normal distribution with $EX_i(t) = 0$, $VarX_i(t) = 1$, $Cov(X_i(t),X_j(t)) = r_{ij}(t)$.

3. Robust estimation of the individual distribution and algorithms

We apply the concept of the weighted least trimmed estimators in order k ($WLT(k)$) (see Vandev and Neykov, 1998) in order to estimate the offspring distributions in the MGWL processes in the presence of outliers.

Let us suppose that we have two sets of sample paths of a branching process with several types of particles based on the generation sizes and of the entire family tree. This means that we are able to observe correspondingly the frequencies $Z_i(t,N)$ (as already mentioned they represent the number of particles of type i in the t-th generation of a MGWL process starting with N particles of type 1) and the relative frequencies $\Delta_i(t,N) = Z_i(t,N)/U(t,N)$. Using these two sets of observations, over each realization we obtain a number of estimated values for the offspring distributions. As already mentioned, under the appropriate norming $V_i(t,N)$ and $W_i(t,N)$ these values are asymptotically normally distributed. If the required conditions for asymptotic normality are not satisfied, the estimated values are far from the real values of the offspring distributions. The aim is to

propose an algorithm for robust estimation of the offspring distribution, exploring the idea of the weighted and trimmed likelihood, in order to eliminate the cases, which do not satisfy these conditions.

We remind that the robust properties of an estimator can be studied by the measure of robustness, called breakdown point (BP). We adopt the definition of a finite sample breakdown point of Hampel at al. (1986). For a given estimator S it is defined as $BP(S) = \frac{1}{r}\max\{m : \sup\|S(X_m)\| < \infty\}$, where X_m is a sample, obtained from the sample X over r observations by replacing any m of the observations by arbitrary values.

Vandev and Neykov (1993) determined the breakdown point of the $WLT(\alpha)$ estimators in the case of multivariate normal distribution as $BP > (r-\alpha)/r$ if $r \geq 3(d+1)$ and $(r+d+1)/2 \leq \alpha \leq r-d-1$, where d is the space dimensionality and α is the trimming factor.

3.1. Robust estimation based on generation sizes

In this subsection we consider the WLT(k) estimator of the mean and covariance matrix for the fixed generation t based on the observations over the generation sizes in this moment t of several sample paths of a MGWL process. This procedure gives us as a further result the "correct" trees to use and the "outlier" trees to avoid for improvement of the offspring distribution estimates for the different particle types. The estimates of the individual distributions are calculated in the standard way using the information about the evolution of the entire family tree, i.e. we estimate the probabilities $p^i_{(j_1,\ldots,j_d)}$ that a particle of type i produces in the next generation j_1 particles of type 1, j_2 particles of type 2, etc., as the number of particles of type i with the given offspring divided by the total number of particles of type i (thus we obtain a Harris type estimator).

Let us consider the set $\{\mathbf{Z}^{(1)}(N_1),\ldots,\mathbf{Z}^{(r)}(N_r)\}$, where

$$\mathbf{Z}^{(i)}(N_i) = (\mathbf{Z}_1^{(i)}(N_i),\ldots,\mathbf{Z}_d^{(i)}(N_i))$$

is a single realization of a MGWL process with N_i ancestors of the same type and length L,

$$\mathbf{Z}_j^{(i)}(N_i) = (Z_j^{(i)}(0,N_i), Z_j^{(i)}(1,N_i),\ldots,Z_j^{(i)}(L,N_i)),$$

$N_i, L \geq 1$, $i = 1,2,\ldots,r$, $j = 1,2,\ldots,d$, $d \in N^+$ is the number of particle types and $r \in N^+$ is the number of sample paths.

We also need the following notation for the vector of the number of types in the fixed t-th generation of the i-th sample path

$$\mathbf{Z}^{(i)}(t,N_i) = (Z_1^{(i)}(t,N_i),\ldots,Z_d^{(i)}(t,N_i)).$$

In this section we consider the $WLT(\alpha)$ estimator of the mean vector

$$N\mathbf{M}_1(t) = (Nm_{11}(t), Nm_{12}(t), \ldots, Nm_{1d}(t))$$

and covariance matrix

$$N\mathbf{C}(t)$$

of the asymptotic multivariate normal distribution of $(Z_1(t,N), \ldots, Z_d(t,N))$, $t = \overline{1,L}$, (see (3)), obtained when $N = N_1 = \cdots = N_d$, which may be presented in the following way:

$$(4) \quad (\widehat{\mathbf{M}_1(t)}, \widehat{\mathbf{C}(t)}) = \underset{\mathbf{M}_1(t),\mathbf{C}(t)}{argmin} \sum_{i=1}^{\alpha} - \log \phi(\mathbf{Z}^{\nu(\mathbf{i})}(t,N), N\mathbf{M}_1(t), N\mathbf{C}(t)),$$

where in the expression $(\widehat{\mathbf{M}_1(t)}, \widehat{\mathbf{C}(t)})$ $\widehat{\mathbf{M}_1(t)}$ is the estimator of $\mathbf{M}$ and $\widehat{\mathbf{C}(t)}$ is the estimator of $\mathbf{C}$. Here α is a properly chosen trimming factor, $\phi(\mathbf{Z}^{(i)}(t,N)$, $N\mathbf{M}_1(t), N\mathbf{C}(t))$ is the density probability function of the asymptotic multivariate normal distribution of $\mathbf{Z}^{(i)}(t,N)$, ν is a permutation of the indices such that

$$\phi(\mathbf{Z}^{\nu(1)}(t,N), N\mathbf{M}_1(t), N\mathbf{C}(t)) \geq \cdots \geq \phi(\mathbf{Z}^{\nu(\alpha)}(t,N), N\mathbf{M}_1(t), N\mathbf{C}(t)),$$

$\mathbf{M}_1(t)$ and $\mathbf{C}(t)$ are the unknown parameters of the process. This is a $WLT(\alpha)$ estimator, in which all weights are equal to 1.

As a direct corollary of the result of Vandev and Neykov (1993) we see that

Proposition 3.1. *The breakdown point BP of the $WLT(\alpha)$ estimators* (4) *of the mean vector $\mathbf{M}_1(t)$ and covariance matrix $\mathbf{C}(t)$, $t = \overline{1,L}$, in the MGWL process starting with N ancestors of type 1 is $BP > (r-\alpha)/r$, if $r \geq 3(d+1)$ and $(r+d+1)/2 \leq \alpha \leq r-d-1$. Here d is the number of particle types, α is the trimming factor and r is the number of observed trajectories.*

Remark 3.1. It is possible to consider the generalization of (4) when the sample paths that we observe start with different (and large) number of ancestors $N_1, \ldots, N_d$ (or, we observe realizations over a multivariate Galton-Watson process starting with an increasing and random number of ancestors):

$$(5) \quad (\widehat{\mathbf{M}_1(t)}, \widehat{\mathbf{C}(t)}) = \underset{\mathbf{M}_1(t),\mathbf{C}(t)}{argmin} \sum_{i=1}^{\alpha} - \log \phi(\mathbf{Z}^{\nu(\mathbf{i})}(t, N_{\nu(i)}), N_{\nu(i)}\mathbf{M}_1(t), N_{\nu(i)}\mathbf{C}(t)),$$

where again α is a properly chosen trimming factor,

$$\phi(\mathbf{Z}^{(i)}(t, N_{\nu(i)}), N_{\nu(i)}\mathbf{M}_1(t), N_{\nu(i)}\mathbf{C}(t))$$

is the density probability function of the asymptotic multivariate normal distribution of $\mathbf{Z}^{(i)}(t, N_{\nu(i)})$, ν is a permutation of the indices such that

$$\phi(\mathbf{Z}^{\nu(1)}(t, N_{\nu(1)}), N_{\nu(1)}\mathbf{M}_1(t), N_{\nu(1)}\mathbf{C}(t)) \geq \dots$$

$$\cdots \geq \phi(\mathbf{Z}^{\nu(\alpha)}(t, N_{\nu(\alpha)}), N_{\nu(\alpha)}\mathbf{M}_1(\alpha), N_{\nu(\alpha)}\mathbf{C}(\alpha)),$$

$\mathbf{M}_1(t)$ and $\mathbf{C}(t)$ are the unknown parameters of the process.

3.1.1. An algorithm

We propose an algorithm for calculating the estimates of the offspring distributions over the whole family trees excluding the outlier trees. The basis for determining the trees with outlier generation sizes is formula (5).

1. Setting the initial value of the mean vector $\mathbf{M}_1(t)$ and covariance matrix $\mathbf{C}(t)$ of the multivariate normal probability density function $\phi(\circ, N\mathbf{M}_1(t), N\mathbf{C}(t))$.

2. Calculating the values ϕ_k of the of the log-density function of the vector

$$\mathbf{Z}^{(k)}(t, N_k) = ((Z_1^{(k)}(t, N_k), \dots, Z_d^{(k)}(t, N_k)))$$

for the sample path $k, k = 1, \cdots, r$ at the predefined moment t.

3. Sorting the values $\{\phi_k\}$ in a descending way: $\phi_{(1)} \geq \cdots \geq \phi_{(r)}$.

4. Calculating $\mathbf{M}_1(t)$ and $\mathbf{C}(t)$ from $\phi_{(1)}, \cdots, \phi_{(\alpha)}$, where the trimming factor is a proportion of the number of the observed sample paths: $\alpha = int(q.r)$, $q \in (0.5, 1]$.

5. If the the alteration of the sum $\sum_{k=1}^{\alpha} \phi_{(k)}$ is less than an appropriate chosen small value ε than exit and calculate probabilities, else go back to 2.

3.1.2. A numerical example

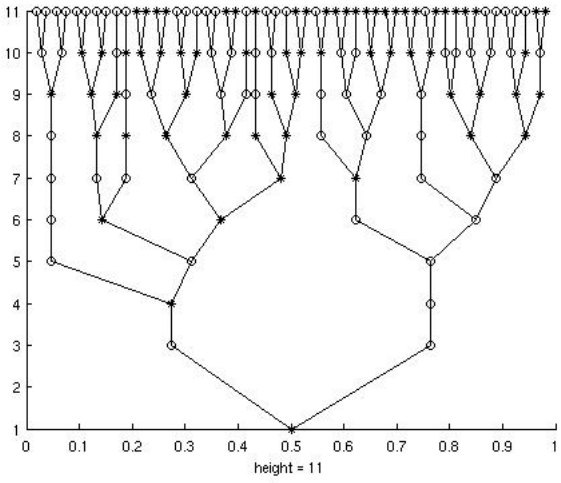
As an example of the proposed procedure we generate 120 sample trees over a multivariate Galton-Watson branching process with a large random number of ancestors. We simulate

- $r_1 = 100$ sample paths $\{\mathbf{Z}^{(1)}(N_1), \ldots, \mathbf{Z}^{(100)}(N_{100})\}$ and
- $r_2 = 20$ outlier sample paths $\{\mathbf{Z}^{(101)}(N_{101}), \ldots, \mathbf{Z}^{(120)}(N_{120})\}$.

The distribution of the regular sample paths is given in the table below

Type	Probability	Offsprings
1	0.2	1 0
1	0.2	0 1
1	0.4	2 0
1	0.2	0 2
2	0.2	1 0
2	0.2	1 1
2	0.3	2 0
2	0.3	0 2

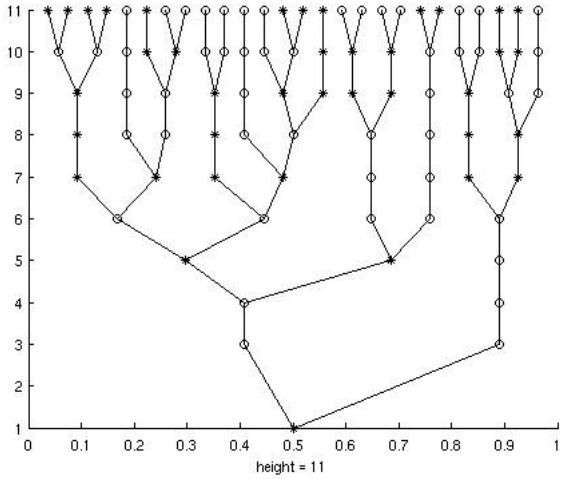
and one simulated family tree of process can be seen on the next figure



The distribution of the outlier sample path is

Type	Probability	Offsprings
1	0.5	1 0
1	0.5	0 2
2	0.5	2 0
2	0.5	0 1

On the next figure one sample path of the outlier process is given



The estimation of the disturbed probabilities gives us the following result, where the influence of the outlier sample paths is seen (note the last row):

Type	Probability	Offsprings
1	0.2297	0 2
1	0.1812	0 1
1	0.2271	1 0
1	0.3621	2 0
2	0.2613	0 2
2	0.1737	1 0
2	0.3251	2 0
2	0.1726	1 1
2	0.0674	0 1

which is far from the true generating mechanism of the process.

The proposed procedure gives the following result, where the influence of the outlier sample paths is reduced:

Type	Probability	Offsprings
1	0.1956	0 1
1	0.2063	1 0
1	0.2042	0 2
1	0.3939	2 0
2	0.1941	1 0
2	0.2933	2 0
2	0.3137	0 2
2	0.1989	1 1

3.2. Robust estimation based on the relative frequencies

In this section we consider robust estimators of the theoretical offspring distribution $\pi = \{p^1_{j_1,j_2}\}$, in a MGWL process with two types of particles. We remind that under $p^1_{j_1,j_2}$ we understand the probability that a particle of type 1 has j_1 children of type 1 and j_2 of type 2. The estimation is based on the asymptotic distribution (1) of the relative frequency $\Delta_1(t,N)$, i.e. the observations in this model are r relative frequencies $\Delta_1^{(1)}(t,N),\ldots,\Delta_1^{(r)}(t,N)$ obtained from r independent realizations of the process starting with N particles of type 1.

The estimator can be expressed in the form:

$$\widehat{\pi} = \underset{\pi}{argmin} \sum_{i=1}^{\alpha} -\log \phi(\boldsymbol{\Delta}_{\mathbf{1}}^{\nu(\mathbf{i})}(t,N), p_1(t), \frac{S_1^2(t)}{M^2(t)N}), \tag{6}$$

where α is a properly chosen trimming factor, $\phi\left(\boldsymbol{\Delta}_{\mathbf{1}}^{\nu(\mathbf{i})}(t,N), p_1(t), \frac{S_1^2(t)}{M^2(t)N}\right)$ is the density probability function of the asymptotically normal distribution of $\Delta_1^{(i)}(t,N)$, ν is a permutation of the indices such that

$$\phi(\boldsymbol{\Delta}_{\mathbf{1}}^{\nu(\mathbf{1})}(t,N), p_1(t), \frac{S_1^2(t)}{M^2(t)N}) \geq \cdots \geq \phi(\boldsymbol{\Delta}_{\mathbf{1}}^{\nu(\alpha)}(t,N), p_1(t), \frac{S_1^2(t)}{M^2(t)N}),$$

$p_1(t)$, $S_1^2(t)$ and $M(t)$ are defined in Section 2 and are functions of the unknown parameters π of the process.

3.2.1. A cell proliferation example

Let us consider the age-dependent two-type reducible Bellman-Harris branching model of oligodendrocyte generation in cell culture, studied first in Yakovlev, M. Mayer-Proschel and M. Noble (1998) and later in Yakovlev, Stoimenova,

Yanev (2008), Yakovlev and Yanev (2009). The oligodendrocyte type-2 astrocyte progenitor cells (O-2A progenitor cells) are known to be precursors of oligodendrocytes in the developing central nervous system. An O-2A progenitor cell either differentiates into an oligodendrocyte, which does not divide under normal conditions, or retains the ability to proliferate into two cells of the same type.

Let $Z_1(t,N)$ be the number of O-2A progenitor cells (cells of type 1) at the moment t and $Z_2(t,N)$ - the number of oligodendrocytes (cells of type 2), supposing that the process starts at time 0 with N progenitor cells. In the above cited papers it is noted that the embedded discrete time process of the considered 2-type Bellman-Harris branching model $(Z_1(t,N), Z_2(t,N))$, $t \geq 0$ is a 2-type Bienaymé-Galton-Watson process with offspring probability generating functions

$$h_1(s_1,s_2) = p_0 + p_1 s_1^2 + p_2 s_2, \;\; h_1(1,1) = p_0 + p_1 + p_2 = 1, \;\; h_2(s_1,s_2) = 1.$$

The interpretation of equations (7) is the following: in this process at the end of its life (mitotic cycle) every cell of type 1 either dies with probability p_0, or differentiates into a new cell of type 2 with probability p_2, or divides into two new type 1 cells with probability p_1. Every type 2 cell at the end of its life dies without any offspring.

In the notations of Section 2 the first and second moments of the offspring distributions are

$$\begin{aligned}
m_{11} &= \frac{\partial}{\partial s_1} h_1(s_1,s_2)|_{s_1=s_2=1} = 2p_1, \;\; m_{12} = \frac{\partial}{\partial s_2} h_1(s_1,s_2)|_{s_1=s_2=1} = p_2, \\
m_{21} &= \frac{\partial}{\partial s_1} h_2(s_1,s_2)|_{s_1=s_2=1} = 0, \quad\;\; m_{22} = \frac{\partial}{\partial s_2} h_2(s_1,s_2)|_{s_1=s_2=1} = 0, \\
b_{11}^1 &= \frac{\partial^2}{\partial s_1^2} h_1(s_1,s_2)|_{s_1=s_2=1} = 2p_1, \;\; b_{12}^1 = b_{21}^1 = b_{22}^1 = b_{11}^2 = b_{12}^2 = b_{21}^2 = b_{22}^2 = 0
\end{aligned}$$

Hence

$$\sigma_1^2 = Var[Z_1(1)|Z_1(0)=1] = b_{11}^1 + m_{11} - (m_{11})^2 = 4p_1[1-p_1],$$

$$\sigma_2^2 = Var[Z_2(1)|Z_1(0)=1] = b_{22}^1 + m_{12} - (m_{12})^2 = p_2[1-p_2],$$

$$C_{12} = b_{12}^1 - m_{11}m_{12} = -2p_1p_2.$$

Using the formula for the mean matrix for generation t

$$\mathbf{M}(t) = \mathbf{M}^t = \begin{bmatrix} 2p_1 & p_2 \\ 0 & 0 \end{bmatrix}^t = \begin{bmatrix} m_{11}(t) & m_{12}(t) \\ m_{21}(t) & m_{22}(t) \end{bmatrix} = \begin{bmatrix} (2p_1)^t & (2p_1)^{t-1}p_2 \\ 0 & 0 \end{bmatrix},$$

we obtain

$$\begin{aligned} M(t) &= m_{11}(t) + m_{12}(t) = (2p_1)^{t-1}[2p_1 + p_2], \\ p_1(t) &= \frac{m_{11}(t)}{m_{11}(t) + m_{12}(t)} = \frac{2p_1}{2p_1 + p_2}. \end{aligned} \tag{7}$$

For the second moments in generation t the following recurrence formula is valid:

$$b^i_{jk}(t+1) = \sum_{l=1}^{2}\sum_{r=1}^{2} b^i_{lr} m_{lj}(t) m_{rk}(t) + \sum_{l=1}^{2} m_{il} b^l_{jk}(t).$$

This yields in our particular case

$$b^2_{jk}(t+1) = \sum_{l=1}^{2}\sum_{r=1}^{2} b^2_{lr} m_{lj}(t) m_{rk}(t) + \sum_{l=1}^{2} m_{2l} b^l_{jk}(t) = 0$$

and

$$\begin{aligned} b^1_{jk}(t+1) &= \sum_{l=1}^{2}\sum_{r=1}^{2} b^1_{lr} m_{lj}(t) m_{rk}(t) + \sum_{l=1}^{2} m_{1l} b^l_{jk}(t) = \\ &= b^1_{11} m_{1j}(t) m_{1k}(t) + m_{11} b^1_{jk}(t). \end{aligned} \tag{8}$$

From (9) one has

$$\begin{aligned} b^1_{11}(t+1) &= b^1_{11}(m_{11}(t))^2 + m_{11} b^1_{11}(t) = (2p_1)^t \frac{(2p_1)^t - 1}{2p_1 - 1}, \\ \Rightarrow C_{11}(t) &= \sigma^2_1(t) = b^1_{11}(t) + m_{11}(t) - (m_{11}(t))^2 \\ &= \frac{(2p_1)^t[(2p_1)^t - 1][2 - 2p_1]}{2p_1 - 1}; \end{aligned} \tag{9}$$

$$\begin{aligned} b^1_{12}(t+1) &= b^1_{11} m_{11}(t) m_{12}(t) + m_{11} b^1_{12}(t) = (2p_1)^t p_2 \frac{(2p_1)^{t-1} - 1}{2p_1 - 1}, \\ \Rightarrow C_{12}(t) &= b^1_{12}(t) - m_{11}(t) m_{12}(t) \\ &= \frac{(2p_1)^t p_2[2(2p_1)^{t-1} - (2p_1)^t - 1]}{2p_1 - 1}; \end{aligned} \tag{10}$$

$$
\begin{aligned}
b_{22}^1(t+1) &= b_{11}^1(m_{12}(t))^2 + m_{11}b_{22}^1(t) = (2p_1)^{t-1}p_2^2\frac{(2p_1)^{t-1}-1}{2p_1-1}, \\
\Rightarrow C_{22}(t) &= \sigma_2^2(t) = b_{22}^1(t) + m_{12}(t) - (m_{12}(t))^2 = \\
&= \frac{(2p_1)^{t-1}p_2[2(2p_1)^{t-1}p_2 - (2p_1)^t p_2 + 2p_1 - p_2 - 1]}{2p_1-1}; \qquad (11)
\end{aligned}
$$

Now we have all we need to calculate the asymptotic variance $S_1^2(t) = VarY_1(t)$, because according to (2)

$$
\begin{aligned}
S_i^2(t) &= \sum_{k,l=1}^{d} r_{kl}(t)a_{ki}(t)a_{li}(t) = \\
&= C_{11}(t)[1-p_1(t)]^2 - 2C_{12}(t)[1-p_1(t)]p_1(t) + C_{22}(t)p_1^2(t) = \\
&= \frac{2(2p_1)^t p_2[p_1+p_2]}{(2p_1+p_2)^2} \qquad (12)
\end{aligned}
$$

and

$$
\frac{S_1^2(t)}{M^2(t)} = \frac{2p_2[p_1+p_2]}{(2p_1)^{t-2}[2p_1+p_2]^4}. \qquad (13)
$$

Finally we can explicitly express the asymptotic distribution of the relative frequency $\Delta_1(t,N)$ as a function of the offspring probabilities p_1 and p_2 in the form

$$
\sqrt{N}\left[\Delta_1(t,N) - \frac{2p_1}{2p_1+p_2}\right] \sim N\left(0, \frac{2p_2[p_1+p_2]}{(2p_1)^{t-2}[2p_1+p_2]^4}\right),
$$

or equivalently,

$$
\Delta_1(t,N) \sim N\left(\frac{2p_1}{2p_1+p_2}, \frac{2p_2[p_1+p_2]}{N(2p_1)^{t-2}[2p_1+p_2]^4}\right).
$$

Hence the robust estimator of the vector of unknown parameters $\pi = (p_1, p_2)$ is obtained from (6), replacing $p_1(t)$ by (9) and $\frac{S_1^2(t)}{M^2(t)}$ by (14).

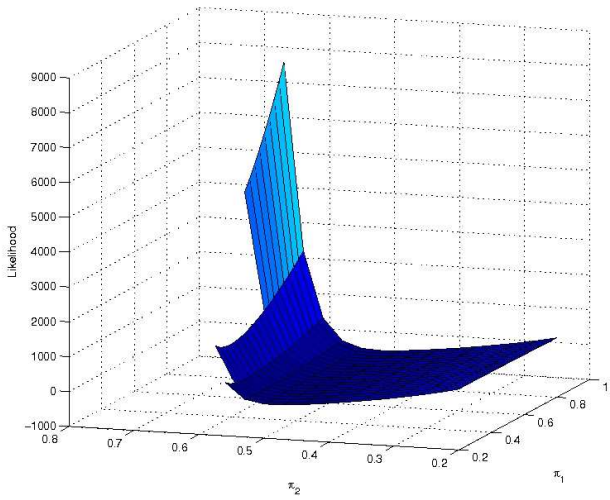
We illustrate the applicability of model (6) for detecting outlier trees by simulating 80 sample paths of the considered process with offspring distribution

Type	Probability	Offsprings
1	0.2	0 0
1	0.6	2 0
1	0.2	0 1
2	1	0 0

and 20 outlier trees distributed according

Type	Probability	Offsprings
1	0.5	2 0
1	0.5	0 1
2	1	0 0

On the following graph the likelihood function of the robust estimator is shown:



The estimates obtained by the robust estimator (6) are

$$\hat{\pi} = (\hat{p}_1 = 0.5906, \hat{p}_2 = 0.2070)$$

$$\hat{p}_0 = 1 - 0.5906 - 0.2070 = 0.2024,$$

while the ML estimation of the disturbed set gives the result

$$\tilde{\pi} = (\tilde{p}_1 = 0.7223, \tilde{p}_2 = 0.1730)$$

$$\tilde{p}_0 = 1 - 0.7223 - 0.1730 = 0.1047.$$

Remark. All calculations are made under MATLAB with "BP Engine Rev. 2" package, available at `http://www.fmi.uni-sofia.bg/fmi/statist/projects/bp`.

Acknowledgements. The authors are grateful to professor Nikolay Yanev from the Institute of Mathematics and Informatics, Bulgarian Academy of Science for the introduction to this interesting topic and the inspiring discussions.

REFERENCES

[1] S. Asmussen, H. Herring. Branching Processes. Birkhauser, Boston, (1983).

[2] K. B. Athreya, P. E. Ney. Branching Processes. Springer-Verlag, Berlin (1972).

[3] D. Atanasov, V. Stoimenova, N. M. Yanev. Estimators in Branching Processes with Immigration. *Pliska Stud. Math. Bulgar.*, **18** (2007), 19–40.

[4] D. Atanasov, V. Stoimenova, N. Yanev. Offspring Mean Estimators in Branching Processes with Immigration. *Pliska Stud. Math. Bulgar.*, **19** (2009), 69–82.

[5] J. P. Dion, N. M. Yanev. Limiting Distributions of a Galton-Watson Branching Process with a Random Number of Ancestors. *C. R. Acad. bulg. Sci.*, **44** (1991), No 3, 23–26.

[6] J. P. Dion, N. M. Yanev. Estimation theory for the variance in a branching process with an increasing random number of ancestors. *C. R. Acad. bulg. Sci.*, **45** (1992), No 11, 27–30.

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

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

[9] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, W. A. Stahel. Robust Statistics: The Approach Based on Influence Functions. John Wiley and Sons, New York, 1986.

[10] T. E. Harris. The Theory of Branching Processes. Springer-Verlag, Berlin, 1963.

[11] P. Jagers. Branching Processes with Biological Applications. Wiley, London, 1975.

[12] A. N. Kolmogorov, N. A. Dmitriev. Branching random processes. *Dokl. Akad. Nauk (Proc. Acad. Sci. USSR)*, **56**, (1947), 7–10 (in Russian).

[13] A. N. Kolmogorov, B. A. Sevastyanov. Calculation of final probabilities of branching random processes. *Dokl. Akad. Nauk (Proc. Acad. Sci. USSR)*, **56**, (1947), 783–786 (in Russian).

[14] B. A. Sevastyanov. Branching Processes. Nauka, Moscow, 1971 (in Russian).

[15] V. Stoimenova. Robust Parametric Estimation of Branching Processes with Random Number of Ancestors. *Serdica Math. J.*, **31** (2005), No 3, 243–262.

[16] V. Stoimenova, D. Atanasov. Nonparametric Robust Estimation of the Individual Distribution in Branching Processes with a Random Number of Ancestors. *Math.and Education in Math.*, **35** (2006), 302–308.

[17] V. Stoimenova, D. Atanasov, N. Yanev. Robust Estimation and Simulation of Branching Processes. *C. R. Acad. bulg. Sci.*, **57** (2004a), No 5, 19–22.

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

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

[20] V. Stoimenova, N. Yanev. Parametric Estimation in Branching Processes with an Increasing Random Number of Ancestors. *Pliska Stud. Math. Bulgar.*, **17** (2005), 295–312.

[21] D. L. Vandev. A Note on Breakdown Point of the Least Median of Squares and Least Trimmed Estimators. *Statistics and Probability Letters*, **16** (1993), 117–119.

[22] D. L. Vandev, N. M. 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), Basel, Birkhauser Verlag, 1993, 257–264.

[23] D. L. Vandev, N. M. Neykov. About Regression Estimators with High Breakdown Point. *Statistics*, **32**, (1998), 111—129.

[24] A. Y. Yakovlev, M. Mayer Proschel, M. Noble. A stochastic model of brain cell differentiation in tissue culture. *J. Math. Biol.*, **37**, (1998), 49–60.

[25] A. Y. Yakovlev, V. K. Stoimenova, N. M. Yanev. Branching Processes as Models of Progenitor Cell Populations and Estimation of the Offspring Distributions. *J. American Statistical Assoc*, **103** (2008), 1357–1366.

[26] A. Y. Yakovlev, N. M. Yanev. Transient Processes in Cell Proliferation Kinetics. Springer Verlag, Berlin, 1989.

[27] A. Y. Yakovlev, N. M. Yanev. Relative frequencies in multitype branching processes. *Ann. Appl. Probab.*, **19** (2009), No 1, 1–14.

[28] A. Y. Yakovlev, N. M. Yanev. Limiting distributions in multitype branching processes. *Stochastic Analysis and Applications*, **28** (2010), No 6, 1040–1060.

[29] N. M. Yanev. On the Statistics of Branching Processes. *Theor. Prob. Appl., XX*, **3**, (1975), 623–633.

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

Institute of Mathematics and Informatics
Bulgarian Academy of Science
Acad. G. Bontchev Str., Bl. 8
1113 Sofia,Bulgaria

Dimitar Atanasov
New Bulgarian University
21 Montevideo Str., Sofia, Bulgaria
e-mail: datanasov@nbu.bg