PLISKA
STUDIA MATHEMATICA
BULGARICA

ПЛИСКА
БЪЛГАРСКИ
МАТЕМАТИЧЕСКИ
СТУДИИ

Pliska Stud. Math. Bulgar. **18** (2007), 19–40

PLISKA
STUDIA MATHEMATICA
BULGARICA

ESTIMATORS IN BRANCHING PROCESSES WITH IMMIGRATION

Dimitar Atanasov, Vessela Stoimenova, Nikolay Yanev [1]

In the present paper we consider the branching process with immigration and its relationship to the Bienayme - Galton - Watson process with a random number of ancestors. Several estimators of the immigration component are considered - the conditional least squares estimator of Heyde - Seneta, the conditional weighted least squares estimator of Wei - Winnicki and the estimator of Dion and Yanev. Their comparison is based on simulations of the entire immigration family trees and computational results. The asymptotic normality of the estimator of Dion and Yanev is combined with the general idea of the trimmed and weighted maximum likelihood. As a result, robust modifications of the immigration component estimator is proposed. They are based on one and several realizations of the entire family tree and are studied via simulations and numerical results.

1. Introduction

The statistical inference for branching processes with immigration (BGWI process) is considered in the papers of Heyde and Leslie (1971), Heyde and Seneta (1972), Heyde (1974), Yanev and Tchoukova-Dancheva (1980), Winnicki (1988), Wei and Winnicki (1989, 1990) and others. In these papers the asymptotic properties of the BGWI process are studied as well as the nonparametric maximum

[1]The paper is supported in part by the National Science Fund of Bulgaria, Grant No VU-MI -105/2005 and by ECO-NET 06 action 12634TJ founded by French Foreign Office.

2000 *Mathematics Subject Classification:* 60J80

Key words: immigration mean, asymptotic normality, robust estimator

likelihood, the conditional least squares and weighted conditional least squares estimation are introduced.

In the present paper we consider the estimation of the immigration component in BGWI processes from another point of view using the relationship between the process with immigration and the process with an increasing random number of ancestors, whose statistical estimation is proposed by Yanev (1975) and studied in the nonparametric situation by Dion and Yanev (1991, 1992, 1994, 1997). We show the advantages of this approach and extend it for the purpose of robustification in the sense of the trimmed and weighted likelihood.

We remind that the BGWI process is a process with two types of particles: the so-called natives and immigrants. They are characterized by the fact that each particle reproduces independently of each other. Each native particle gives rise only to natives, according to an offspring distribution ($\{p_k\}$) with mean m and variance σ^2. The immigrant always produces just one immigrant as well as a random number of "natives" according to an immigration distribution ($\{q_k\}$), whose mean and variance are λ and b^2 respectively. This is an example of a decomposable singular multitype branching process. Because of its importance for applications, it is treated separately and is considered in a different way from the multitype model (Dion, 1993).

Let $\{Y_n\}_{n=0}^{\infty}$ be a BGWI process defined by the recursive formula

$$Y_n = \begin{cases} \sum\limits_{j=1}^{Y_{n-1}} X_{nj} + I_n, & if \quad Y_{n-1} > 0, \\ I_n, & if \quad Y_{n-1} = 0. \end{cases} \tag{1}$$

Here $\{X_{nj}\}$ and $\{I_n\}$ are independent sequences of i.i.d. nonnegative, integer valued random variables with distribution $\{p_k\}$ and $\{q_k\}$ respectively. The r.v. Y_0 is nonnegative and integer valued , which is also independent of $\{X_{nj}\}$ and $\{I_n\}$. Without any loss of generality further on one can assume that $Y_0 = 0$.

Assume further that the offspring and immigration distributions are nondegenerate and the offspring mean $m < \infty$.

As usually the process is called *subcritical* if $m < 1$, *critical* if $m = 1$ and *supercritical* if $m > 1$.

In the supercritical case when $\sum\limits_{1}^{\infty} p_k \log k < \infty$ and $\lambda < \infty$ one has $Y_n/m^n \stackrel{a.s.}{\longrightarrow} W > 0$. The process behaves like a classical Bienayme - Galton - Watson process (BGW process) on the set of non-extinction. In this case only one family of immigrants is included in each generation and is 'assimilated' among the geometrically increasing number of natives. Therefore no parameter of the immigration distribution can have a consistent estimator. A formal proof of this result is provided by Wei & Winnicki (1990).

In the subcritical case, assuming that the offspring distribution and the immigration distribution are such that the Markov chain $\{Y_n\}$ is aperiodic and irreducible, it is well known that $\{Y_n\}$ is positive recurrent if and only if $E(\log^+ I_n) < \infty$. Under these conditions and without loss of generality one may think that Y_n is constructed of many independent branching processes, each of them starting at the time of immigration. If each of these processes are subcritical, they have an a.s. finite length and Y_n converges to a limit, i.e. the process has a stationary distribution. Hence standard results on ergodic stationary Markov chains can be applied to estimators of the parameters of the stationary distribution or of the transition matrix. (Wei& Winnicki (1989, 1990)).

The critical situation is the most difficult case to study, because the process is null recurrent if $2\lambda \leq \sigma^2$ and transient if $2\lambda \geq \sigma^2$.

1.1. Classical nonparametric estimation

Results about the nonparametric estimation in the supercritical BGWI process are announced in Heyde (1974), Heyde and Seneta (1971), Heyde and Leslie (1971) and others.

As already mentioned, the branching process with immigration does not become extinct and therefore one would expect to be able to estimate consistently the mean and variance of the individual distribution and the immigration component. In the subcritical case one has a stationary and ergodic Markov chain, which is usually analyzed using the results on the statistical inference for Markov chains. The other approach is to use the time series method. The supercritical case is similar to the regular BGW situation. Winnicki (1988) tried to unify the theory by introducing the conditional least squares estimators (l.s.e).

The conditional least squares estimators of Heyde - Seneta for m and λ are

$$\overline{m}_n = \frac{n\sum_{k=1}^{n} Y_k Y_{k-1} - \sum_{k=1}^{n} Y_k \sum_{k=1}^{n} Y_{k-1}}{n\sum_{k=1}^{n} Y_{k-1}^2 - \left(\sum_{k=1}^{n} Y_{k-1}\right)^2} \tag{2}$$

and

$$\overline{\lambda}_n = \frac{n\sum_{k=1}^{n} Y_k \sum_{k=1}^{n} Y_{k-1}^2 - \sum_{k=1}^{n} Y_{k-1} \sum_{k=1}^{n} Y_k Y_{k-1}}{n\sum_{k=1}^{n} Y_{k-1}^2 - \left(\sum_{k=1}^{n} Y_{k-1}\right)^2}. \tag{3}$$

In Winnicki (1988) it is noted that the conditional least squares estimators are not satisfactory for the following reasons:

1. The estimator $\overline{m}_n$ has a larger asymptotic variance than the m.l.e. $\sum_{i=1}^n Y_i / \sum_{i=1}^n Y_{i-1}$ in the supercritical case.

2. The estimator $\overline{\lambda}_n$ is not a consistent estimator for λ in the supercritical case.

To avoid these disadvantages Wei and Winnicki (1989) proposed to use the weighted conditional least squared estimators, i.e. estimators, obtained by minimizing

$$\sum_{k=1}^{n} \left(\frac{Y_k - E(Y_n|\Im_{k-1}}{\sqrt{Var(Y_k|\Im_{k-1}))}} \right)^2 .$$

The weighted conditional least squares estimators for the offspring and immigration mean are

$$\tilde{m}_n = \frac{\sum\limits_{k=1}^{n} Y_k \sum\limits_{k=1}^{n} \frac{1}{Y_{k-1}+1} - n \sum\limits_{k=1}^{n} \frac{Y_k}{Y_{k-1}+1}}{\sum\limits_{k=1}^{n} (Y_{k-1}+1) \sum\limits_{k=1}^{n} \frac{1}{Y_{k-1}+1} - n^2} \tag{4}$$

and

$$\tilde{\lambda}_n = \frac{\sum\limits_{k=1}^{n} Y_{k-1} \sum\limits_{k=1}^{n} \frac{Y_k}{Y_{k-1}+1} - \sum\limits_{k=1}^{n} Y_k \sum\limits_{k=1}^{n} \frac{Y_{k-1}}{Y_{k-1}+1}}{\sum\limits_{k=1}^{n} (Y_{k-1}+1) \sum\limits_{k=1}^{n} \frac{1}{Y_{k-1}+1} - n^2}. \tag{5}$$

In Winnicki (1988) it is noted that in the supercritical case $\tilde{m}_n$ is a more efficient estimator than $\overline{m}_n$ in the sense of achieving a lower asymptotic variance. However the estimator $\tilde{\lambda}_n$ is not consistent. The weighted conditional least squares method offers a substantial improvement over the ordinary conditional least squares when the supercritical case is considered, but it does not solve the problem of estimating the parameters λ and b^2 of the immigration distribution.

1.2. The relationship between BGWI and BGWR

As noted in Dion (1993), traditionally branching processes with or without immigration have been treated separately. However the estimation theory for the offspring parameters in a Bienayme - Galton -Watson process having a random number of initial ancestors $Z_0(n)$ (BGWR process) can be transferred to a process with immigration without taking account of the criticality of the processes.

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 and in applications to nuclear chain reactions. Results about the classical nonparametric estimation of the offspring mean m and variance σ^2 in the BGWR process are announced in 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, Yannev (2004 a,b, 2005, 2006). Some aspects of the classical parametric estimation are considered in Stoimenova, Yanev (2005) and of the robust parametric estimation - in Stoimenova (2005).

One may consider the partial tree in (1) underlying $\{Y_0, \dots, Y_t, \dots, Y_{n+t}\}$ and define the r.v. $Z_t(n)$ as the number of individuals among generations $t, t+1, t+2, t+n$, whose ancestors immigrated exactly t generations ago, $n, t = 0, 1, 2, \dots$.

One can see that $Z_0(n) = \sum\limits_{j=0}^{n} I_j$ 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 BGW process having a random number of ancestors and can be presented as follows:

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

where $p_k = P(\xi_i(t,n) = k)$. Such a process is denoted by BGWR. $Z_0(n)$.

On the figures bellow a realization of a BGWI process with Poisson offspring mean 1.1 and Poisson immigration mean 1.5 and the corresponding BGWR process are shown:

It is noted in Dion and Yanev (1994) that in general the knowledge of $\{Z_0(n), \dots, Z_t(n)\}$ would seem to be asymptotically equivalent to

$$\left\{ [Y_k]_0^{t+n}, \sum_{k=0}^{n} I_k \right\}$$

as $n, t \longrightarrow \infty$ on the set of the nonextinction.

The estimators of the immigration component λ of Dion and Yanev are

$$\overline{\lambda}_t(n) = \frac{Z_t(n)}{nm^t} \tag{6}$$

if the offspring mean m is known and

$$\widetilde{\lambda}_t(n) = \frac{Z_t(n)}{n(\widehat{m}_t(n))^t} \tag{7}$$

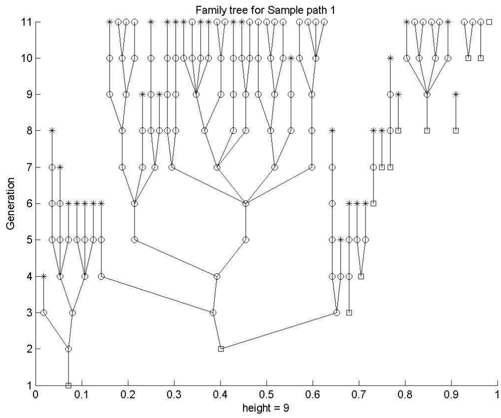


Figure 1: BGWI process with $X \in Po(1.1)$ and $I \in Po(1.5)$.

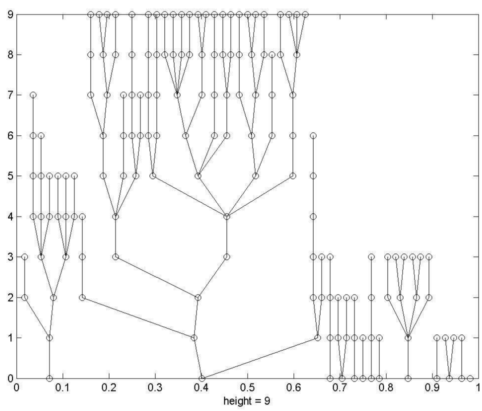


Figure 2: The corresponding BGWR process.

if m is unknown, where $\widehat{m}_t(n) = \left\{\sum\limits_{i=1}^{t} Z_i(n)\right\} / \left\{\sum\limits_{i=0}^{t-1} Z_i(n)\right\}$ is the Harris estimator.

Theorem 1. *[Dion and Yanev (1991, 1992, 1994, 1997)]. Let $\{I_k\}$ be i.i.d. with $\lambda = EI_k < \infty$. Then as $n \to \infty$ uniformly by $0 \le t \le \infty$*

1. *(a)* $\overline{\lambda}_t(n) = Z_t(n)/nm^t \xrightarrow{P} \lambda$;

   (b) $\widetilde{\lambda}_t(n) = Z_t(n)/n[\widehat{m}_t(n)]^t \xrightarrow{P} \lambda$.

2. *Let $b^2, \sigma^2 < \infty$ and the following condition hold:*

$$(m > 1) \vee (m = 1, t/n \to 0) \vee (m < 1, nm^t \to \infty)$$

   (a) If $m < 1$, then

$$\sqrt{nm^t}(\overline{\lambda}_t(n) - \lambda) \xrightarrow{d} N(0, \lambda); \tag{8}$$

   (b) If $m = 1$, then

$$\sqrt{n/t}(\overline{\lambda}_t(n) - \lambda) \xrightarrow{d} N(0, \lambda\sigma^2); \tag{9}$$

   (c) If $m > 1$, then

$$\sqrt{n}(\overline{\lambda}_t(n) - \lambda) \xrightarrow{d} N(0, \frac{\lambda\sigma^2}{m(m-1)} + b^2). \tag{10}$$

   (d) If $m \ne 1$, then $\widetilde{\lambda}_t(n)$has the same asymptotic normality as $\overline{\lambda}_t(n)$;

   (e) If $m = 1$, then

$$\sqrt{n}(\widetilde{\lambda}_t(n) - \lambda) \xrightarrow{d} N(0, b^2). \tag{11}$$

Remark. The statements of the Theorem remain valid if $\{I_t\}$ is a stationary ergodic process, i.e. $Z_0(n)/n = (1/n)\sum_{k=0}^{n} I_k \longrightarrow \nu$ in probability, where ν is a positive r.v.

2. Comparison of the behaviour of the immigration mean estimators

In this section we introduce one quantitative comparison of immigration mean estimators, described in the previous section.

30 family trees are generated for each combination of the parameter values:

- expected number of descendants: $[0.9, 1.0, 1.0]$;

- expected number of immigrants in each generation: from 0.5 to 2.0 with sep 0.1;
- number of generations in the family tree: from 10 to 30 with step 5.

For each of these family trees the estimators of Wei-Winnicki $\widetilde{\lambda}_n$ and Heyde-Seneta $\widehat{\lambda}_t$ for the corresponding value of the immigration parameter, as well as the Dion-Yanev estimators $\overline{\lambda}_t(n)$ and $\widetilde{\lambda}_t(n)$ are calculated with known and unknown value of the expected number of descendants for all possible values of t and n.

For all four sets of the obtained estimators of the immigration parameter the sum of squares of the deviations from the real parameter value (the theoretical immigration mean) are calculated and compared.

We use the following formula for the mean deviation of the estimated values from the theoretical immigration mean:

$$D_n = \frac{1}{n-1}\sum_{i=1}^{n}[\lambda_i(\mathbf{Y_i}) - \lambda]^2. \tag{12}$$

For the case of the Wei - Winnicki and Heyde -Seneta estimators n is the total number of generations of the considered process, $\mathbf{Y_i}$ is the sample path up to the $i-$th generation and $\lambda_i(\mathbf{Y_i})$ is the corresponding estimate. For the case of the estimators of Dion and Yanev $n = t(t-1)/2$ is the number of obtained estimates over a sample path with t generations when passing through all possible values of the two parameters of the process.

The behaviour of the estimators over all values of the alternated parameters may be seen on Figure 3, where the boxplots of the deviations are presented. On the plot it is seen that the Heyde - Seneta estimators have the largest mean value of the deviations. The deviations of the estimators of Wei-Winnicki have a lower variation than those of Heyde-Seneta. This results from the fact that the Wei-Winnicki estimators have a smaller asymptotic variance than the Heyde-Seneta estimators.

The Dion-Yanev estimators have a smaller mean value of the deviations than the estimators of Heyde-Seneta and Wei-Winnicki. The deviations of the estimators of Dion-Yanev with known m have a relatively large deviation. To some extend this can be explained by the fact that for a family tree a set of values of the estimator $\widetilde{\lambda}_t(n)$ for all possible values of t and n may be obtained.

The estimator of Dion - Yanev with unknown offspring mean m behaves best when the four types of estimators are considered. Here the unknown value of the offspring mean is replaced by the Harris estimator

$$\widehat{m}_t(n) = \frac{Z_1(n) + Z_2(n) + \cdots + Z_t(n)}{Z_0(n) + Z_2(n) + \cdots + Z_{t-1}(n)},$$

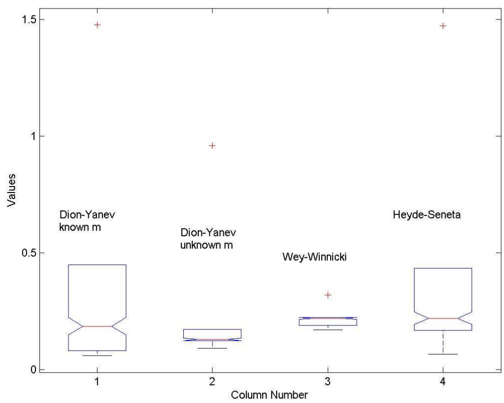


Figure 3: Boxplot of the deviation in the immigration parameter estimators

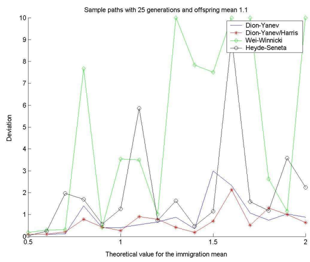
which uses information based on all generation sizes. Consequently the estimator $\widetilde{\lambda}_t(n)$ with m unknown is more informative than that with m known.
On the next figures some typical result are shown. Firstly, the graph of deviation for different values of the immigration mean are presented (Figure 4.1 - 4.4).

It can be seen that the estimators of Dion and Yanev have a relatively lower deviation from the true parameter value than the classical estimators of Heyde - Seneta and Wei - Winnicki. The observed deviations decrease with the increase of the immigration mean value. Comparing these estimators with respect to generation size the following graphs are obtained: (Figure 5.1 - 5.4).

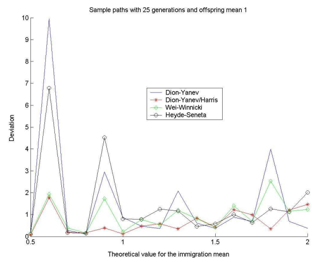
The deviation decreases as the generation size increases. The estimators of Dion and Yanev have a lower deviation than the classical estimators.

Let us now consider the deviation of the estimators with 30 generations for the values of the immigration mean 2.0 and the offspring mean 1.1. On the Figures 4.3 and 4.4 the boxplots of the deviations of Heyde-Seneta, Wei - Winnicki and Dion - Yanev are presented.

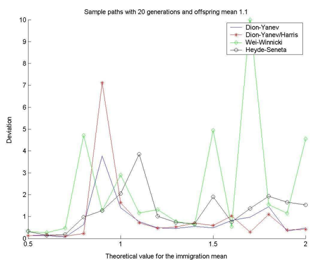
On Figure 4.4 the values of the deviations outside the confidence region of the



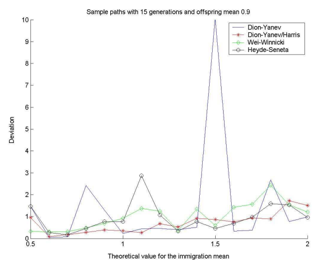
4.1



4.2



4.3



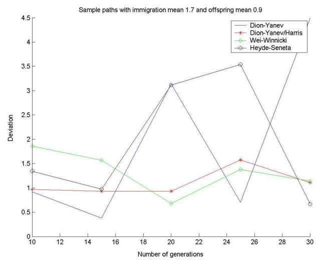
4.4

Figure 4: Deviation for different values of the immigration mean.

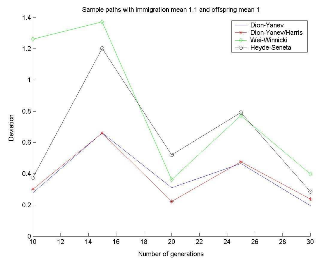
boxplot are denoted by the symbol "+". They may be considered as outliers. It can be seen that the estimators of Heyde-Seneta and Wei-Winnicki do not possess such values. This can be explained by the fact that they have a relative larger deviation than the Dion-Yanev estimators. The presence of deviation values of the Dion - Yanev estimators, which may be considered as outliers, justifies the use of the robust modification of this type of estimators. We will use the robustness of the type of weighted and trimmed likelihood. Our aim is to reduce the influence of the estimators with higher values of the deviation over the immigration mean estimator.

3. Robust modified nonparametric estimators

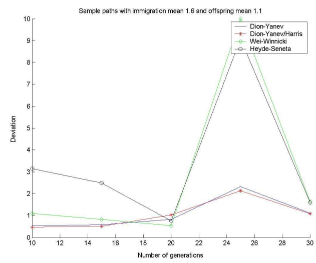
We apply the concept of the weighted least trimmed estimators in order k (WLTE(k)) (see Vandev and Neykov, 1998) in order to estimate the immigration



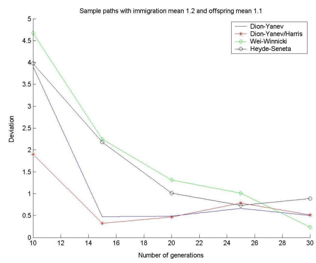
5.1



5.2



5.3



5.4

Figure 5: Comparing the estimators with respect to generation size.

component in the BGWI processes in the presence of outliers.

Let us suppose that we have a set of sample paths of the entire family tree of a branching process with immigration. This means that we are able to observe also the equivalent BGW process, starting with a random number of ancestors (BGWR process). Using this set and the estimators of Dion and Yanvev over each realization we obtain a number of values for the offspring distribution. Under the conditions of *Theorem 1* these values are asymptotically normally distributed. If these conditions are not satisfied, the estimated value is far from the real value of the immigration mean. The aim is to apply the weighted and trimmed likelihood in order to eliminate the cases, which do not satisfy these conditions, and to obtain estimators of the immigration component, closer to the real value.

Let us consider the set $\mathbf{Z} = \{\mathbf{Z^{(1)}(n)}, \cdots, \mathbf{Z^{(r)}(n)}\}$, where $\{\mathbf{Z^{(i)}(n)}\}$ is a single realization of a BGWR process (equivalent to a corresponding realization of the process with immigration) with the same parameters n and t.

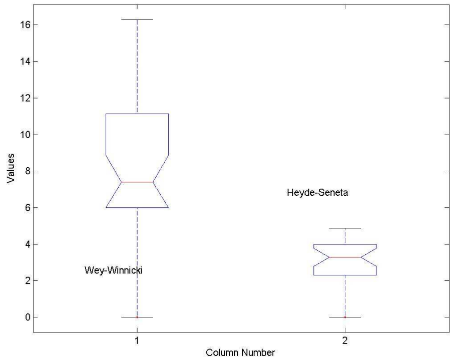


Figure 6: Deviation of Heyde-Seneta and Wei-Winnicki estimators, 30 generations, immigration parameter 2.0, offspring mean 1.1

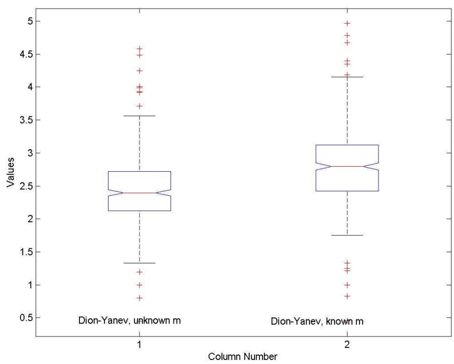


Figure 7: Deviation of Dion-Yanev estimators, 30 generations, immigration mean 2.0, offspring mean 1.1

Let

$$\overline{\lambda}_t^{(i)}(n) = \frac{Z_t^{(i)}(n)}{nm^t} \tag{13}$$

and

$$\widetilde{\lambda}_t^{(i)}(n) = \frac{Z_t^{(i)}(n)}{n(\widehat{m}_t^{(i)}(n))^t}, \tag{14}$$

$i = 1, 2, \ldots, r$, be the estimators of Dion and Yanev for the immigration mean λ for the sample path $\mathbf{Z^{(i)}}(\mathbf{n})$ when the offspring mean m is known or unknown. Here $\widehat{m}_t^{(i)}(n)$ is the Harris estimator for the unknown individual mean based on this sample path.

Let the function $Est(\mathbf{Z^{(i)}}(\mathbf{n}), \lambda)$ be a transformation of the paramether λ, which gives the asymptotic normality:

$$Est(\mathbf{Z^{(i)}}(\mathbf{n}), \lambda) \xrightarrow{d} N(0, a\lambda + c), t \to \infty.$$

Depending on the criticality of the process $Est(\mathbf{Z^{(i)}}(\mathbf{n}), \lambda)$ is expressed as follows:

- in subcritical case $m < 1$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{nm^t}(\overline{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, \lambda); \tag{15}$$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{nm^t}(\widetilde{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, \lambda); \tag{16}$$

- in critical case $m = 1$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{n/t}(\overline{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, \lambda\sigma^2); \tag{17}$$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{n/t}(\widetilde{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, b^2); \tag{18}$$

- in supercritical case $m > 1$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{n}(\overline{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, \frac{\lambda\sigma^2}{m(m-1)} + b^2), \tag{19}$$

$$Est(Z^{(i)}(n), \lambda) = \sqrt{n}(\widetilde{\lambda}_t^{(i)}(n) - \lambda) \xrightarrow{d} N(0, \frac{\lambda\sigma^2}{m(m-1)} + b^2). \tag{20}$$

Let us propose a trimmed estimator based on a sample of Dion and Yanev estimates of the unknown immigration mean for a given set of family trees $\mathbf{Z} = \{\mathbf{Z}^{(\mathbf{1})}(\mathbf{n}),\cdots,\mathbf{Z}^{(\mathbf{r})}(\mathbf{n})\}$. The estimator is presented as follows:

$$(21) \qquad \widehat{\lambda}_t^T(n) = \underset{\lambda>0}{argmin} \sum_{i=1}^{T} -w_i f(Est(\mathbf{Z}^{(\nu(\mathbf{i}))}(\mathbf{n}),\lambda)),$$

where T is properly choosen trimming factor, $f(x)$ is the log-density of the asymptotically normal distribution of the used Dion - Yanev estimators (expressed by the formulas (14)-(19)), ν is a permutation of the indices, such that

$$f(Est(\mathbf{Z}^{(\nu(\mathbf{1}))}(\mathbf{n}),\lambda)) \geq f(Est(\mathbf{Z}^{(\nu(\mathbf{2}))}(\mathbf{n}),\lambda)) \geq \cdots \geq f(Est(\mathbf{Z}^{(\nu(\mathbf{T}))}(\mathbf{n}),\lambda))$$

and λ is the unknown immigration mean. The weights w_i are nonnegative and at least T of them are strictly positive.

Remark. This estimator is defined as $WLTE(k)$ but it can be considered as $R(k)$ estimator (Vandev and Neykov, 1998; Vandev 2003) as well.

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

$$BP(S) = \frac{1}{n}\max\{m : sup||S(X_m)|| < \infty\},$$

where X_m is a sample, obtained from the sample X by replacing any m of the observations by arbitrary values.

Vandev and Neykov (1998) and Vandev (1993) propose a method for determining the value of the BP of a statistical estimator based on the index of fullness of the set of log-density functions in the log-likelihood. According to Vandev (1993) a set of functions F is called d-full, if the supremum of any subset with cardinality d is a subcompact function (i.e. its Lesbegues sets are compacts). A simple criterion for subcompactness of a continuous function is proposed by Atanasov and Neykov (2001). A continuous function $g(x)$, defined on the set D, is subcompact iff $g(x) \underset{x\to\partial D}{\longrightarrow} \infty$ as x tends to the boundary ∂D of D.

Let us denote by

$$N_0 = \sharp\{i = 1, 2, \dots, r : \overset{\simeq}{\lambda}{}_t^{(i)}(n) = 0\}$$

the number of the estimators $\overline{\lambda}_t(n)$ (or correspondingly $\widetilde{\lambda}_t(n)$ when m is not known) of the immigration mean λ, which are equal to zero.

Put

$$F = \left\{ -f(Est(\mathbf{Z^i}(\mathbf{n}), \lambda)) \right\}, i = 1, \cdots, r.$$

Then for the breakdown point properties of the estimator (21) the following theorem is valid.

Theorem 2. *Let for a BGWI stochastic process the equivalent BGWR process have the properties* $0 < \sigma^2 < \infty$, $Z_0(n)/n \xrightarrow{P} \nu$, *where* ν *is a positive r.v.,* $n, t \to \infty$ *and*

$$(m > 1) \vee (m = 1, t/n \to 0) \vee (m < 1, nm^t \to \infty).$$

The estimators $\widehat{\lambda}_t^T(n)$ *of the immigration mean* λ*, defined by (20), exists and its breakdown point is not less than* $(r-T)/r$*, if* $r \geq 3(N_0+1)$*,* $(r+N_0+1)/2 \leq T \leq r - N_0 - 1$ *in the subcritical and critical cases. In the supercritical case the set* F *is not d-full for any* $d = 1, 2, \ldots$.

Proof. We prove the theorem for the estimator $\overline{\lambda}_t(n)$. The proof for $\widetilde{\lambda}_t(n)$ is analogous.

Let us adopt the following notation relevant to the function $Est(\mathbf{Z^i}(\mathbf{n}), \lambda)$:

$$Est(\mathbf{Z^i}(\mathbf{n}), \lambda) = p(\overline{\lambda}_t(n) - \lambda) \to N(0, a\lambda + c).$$

Here p, a and c are nonnegative constants.

Therefore the function $f(Est(\mathbf{Z^i}(\mathbf{n}), \lambda))$ is expressed in the following form:

$$\exp(f(Est(\mathbf{Z^i}(\mathbf{n}), \lambda))) = \frac{1}{\sqrt{2\pi(a\lambda + c)}} e^{-\frac{p^2(\overline{\lambda}_t(n) - \lambda)^2}{a\lambda + c}}.$$

Then any function $g(\lambda) = -f(Est(\mathbf{Z^i}(\mathbf{n}), \lambda))$ from the set F can be written as

$$g(\lambda) = \log 2\pi + \frac{1}{2}\log(a\lambda + c) + \frac{p^2(\overline{\lambda}_t(n) - \lambda)^2}{a\lambda + c}.$$

Let us first consider the case when $\overline{\lambda}_t(n) \neq 0$ and $a \neq 0$. Then

$$\lim_{\lambda \to \infty} \left[\frac{1}{2}\log(a\lambda + c) + \frac{1}{2}\log(a\lambda + c) + \frac{p^2(\overline{\lambda}_t(n) - \lambda)^2}{a\lambda + c} \right] =$$

$$= \lim_{\lambda \to \infty} \frac{p^2(\overline{\lambda}_t(n) - \lambda)^2}{a\lambda + c} \left[\frac{\frac{1}{2}\log(a\lambda + c)}{\frac{p^2(\overline{\lambda}_t(n) - \lambda)^2}{a\lambda + c}} + 1 \right] =$$

$$= \lim_{\lambda\to\infty} \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda+c}\left[\frac{\frac{1}{2}(a\lambda+c)\log(a\lambda+c)}{p^2(\overline{\lambda}_t(n)-\lambda)^2}+1\right] = \infty.$$

Here we use the equality

$$\lim_{\lambda\to\infty} \frac{\frac{1}{2}(a\lambda+c)\log(a\lambda+c)}{p^2(\overline{\lambda}_t(n)-\lambda)^2} = 0. \tag{22}$$

Let now $c = 0$ and $a \neq 0$. Then

$$\lim_{\lambda\to 0} \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda}\left[\frac{\frac{1}{2}(a\lambda)\log(a\lambda+c)}{p^2(\overline{\lambda}_t(n)-\lambda)^2}+1\right] = \infty.$$

Here the equality (22) is used as well as the fact that for $\lambda \neq \overline{\lambda}_t(n)$

$$\lim_{\lambda\to 0} \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda} = \infty.$$

If $c \neq 0$ and $a \neq 0$, then

$$\lim_{\lambda\to 0} = \left[\frac{1}{2}\log(a\lambda+c) + \frac{1}{2}\log(a\lambda+c) + \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda+c}\right] =$$

$$\lim_{\lambda\to 0} = \left[\frac{1}{2}\log(c) + \frac{1}{2}\log(c) + \frac{p^2\overline{\lambda}_t(n)^2}{c}\right] = const.$$

Therefore, according to Atanasov and Neykov (2001), the function is not sub-compact.

Now let $a \neq 0$ and $\overline{\lambda}_t(n) = 0$. Then using (22) again, one has

$$\lim_{\lambda\to\infty} \left[\frac{1}{2}\log(a\lambda+c) + \frac{1}{2}\log(a\lambda+c) + \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda+c}\right] =$$

$$= \lim_{\lambda\to\infty} = \frac{p^2\lambda^2}{a\lambda+c}\left[\frac{\frac{1}{2}\log a\lambda+c}{\frac{p^2\lambda^2}{a\lambda+1}}+1\right] = \infty$$

To study the function when $\lambda \to 0$, let us consider two cases. First let $c = 0$. Then

$$\lim_{\lambda\to 0} \left[\frac{1}{2}\log(a\lambda+c) + \frac{1}{2}\log(a\lambda+c) + \frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda+c}\right] =$$

$$= \lim_{\lambda\to 0} \frac{1}{2}\log(a\lambda) + \frac{p^2\lambda^2}{a} = -\infty.$$

Therefore the function is not subcompact.

By analogy if $c \neq 0$ one has

$$\lim_{\lambda\to 0}\left[\frac{1}{2}\log(a\lambda+c)+\frac{1}{2}\log(a\lambda+c)+\frac{p^2(\overline{\lambda}_t(n)-\lambda)^2}{a\lambda+c}\right]=const.$$

So as in the previous case the function is not subconpact.

Therefore the set F is not 1-full. Let $c=0$ and let $\overline{\lambda}_t^{(1)}(n)=0$ and $\overline{\lambda}_t^{(2)}(n)\neq 0$ be two estimates. Let f_1 and f_2 be the corresponding functions from F.

Then

$$\sup\{f_1,f_2\}\geq\frac{1}{2}(f_1+f_2)$$

and $\sup\{f_1,f_2\}$ is a subcompact function if $\frac{1}{2}(f_1+f_2)$ is a subcompact one.

Studying $\frac{1}{2}(f_1+f_2)$ one has

$$\lim_{\lambda\to\infty}\frac{1}{2}(f_1+f_2)\geq\lim_{\lambda\to\infty}=\infty$$

and

$$\lim_{\lambda\to 0}\frac{1}{2}(f_1+f_2)=2\log a\lambda+\frac{p^2(\overline{\lambda}_t^{(2)}(n)-\lambda)^2}{a\lambda}+\frac{p^2\lambda}{a}=$$

$$=\lim_{\lambda\to\infty}\left(\frac{p^2(\overline{\lambda}_t^{(2)}(n)-\lambda)^2}{a\lambda}+\frac{p^2\lambda}{a}\right)\left(\frac{2a\lambda\log a\lambda}{p^2(\overline{\lambda}_t^{(2)}(n)-\lambda)^2+p^2\lambda^2}+1\right)=\infty.$$

Therefore if the process is subcritical ($c=0$), the set F is (N_0+1)-full if at least one nonzero estimate takes part in the calculation of (21).

As we saw, in the supercritical case the set F is not 1-full. Now we prove that the set F is not d-full for any $d>1$ and $c\neq 0$. Indeed, let $d_1\geq 0$ and $d_2\geq 0$ be such that $d_1+d_2=d$. Let $J_1\subset F$ and $J_2\subset F$ be two subsets of cardinality d_1 and d_2 respectevely. The set J_1 consists of zero estimates and J_2 - of nonzero estimates.

If $c\neq 0$, the functions from F do not converge to $-\infty$. Consequently there exists a constant K, such that $f>K$ for any function $f\in F$ and $f+K>0$. Then

$$\sup(f\in J_1\bigcup J_2)\leq\sum_{f\in J_1\bigcup J_2}f+K_0$$

for a proper constant K_0.

For the right side one has

$$\sum_{f\in J_1\bigcup J_2} f = d_1\left(\log(a\lambda+c)+\frac{p^2\lambda^2}{a\lambda+c}\right)+d_2\left(\log(a\lambda+c)\right)+\sum_{i=1}^{d_2}\frac{p^2(\overline{\lambda}^{(i)}-\lambda)^2}{a\lambda+c}=$$

$$=\left(\sum_{i=1}^{d_2}\frac{p^2(\overline{\lambda}^{(i)}-\lambda)^2}{a\lambda+c}+\frac{d_1p^2\lambda^2}{a\lambda+c}\right)\left(\frac{d(a\lambda+c)\log(a\lambda+c)}{\sum_{i=1}^{d_2}p^2(\overline{\lambda}^{(i)}+d_1(a\lambda+c)}+1\right).$$

Then

$$\lim_{\lambda\to 0}\sum_{f\in J_1\bigcup J_2} f = const$$

as

$$\lim_{\lambda\to 0}\frac{d(a\lambda+c)\log(a\lambda+c)}{\sum_{i=1}^{d_2}p^2(\overline{\lambda}^{(i)}+d_1(a\lambda+c)}=0,$$

$$\lim_{\lambda\to 0}\frac{d_1p^2\lambda^2}{a\lambda+c}=0$$

and

$$\lim_{\lambda\to 0}\sum_{i=1}^{d_2}\frac{p^2(\overline{\lambda}^{(i)}-\lambda)^2}{a\lambda+c}=const.$$

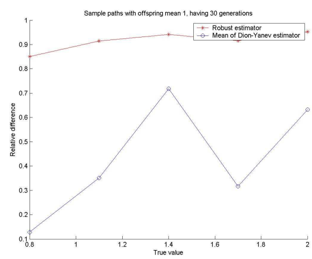
Therefore $\sup(f\in J_1\bigcup J_2)$ is not a subcompact function and the set F is not d-full for any $d>0$.

The lower bound for the breakdown point can be found using the results from Vandev and Neykov (1998). $\square$

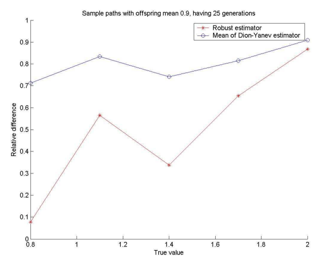
4. Computational results

In this section we compare the classical estimator of Dion and Yanev with unknown value of the offspring mean vs the proposed robust estimator. The Dion - Yanev estimator is chosen because it proves to be the one with relatively lowest deviation of the studied classical estimates. Again, a base for comparison is the deviation of the estimated values from the theoretical immigration mean.

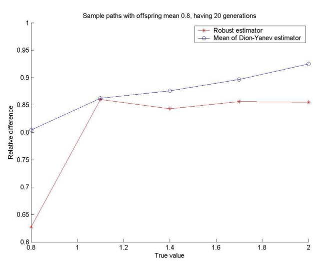
For any value of the immigration mean from 0.8 to 2.0 (with step 0.2) 30 values of the robust estimates are calculated. All of these values are based on 30 realizations of the entire family tree of a BGWI process. For the same set of family trees the Dion - Yanev estimator is calculated. For the two obtained sets of estimates the deviations from the theoretical immigration mean are computed. On the next plots some typical results on the comparison between these two estimators deviations are presented for different parameters of the branching processes. Figure 8.1 presents the the behaviour of these estimators in the critical case with offspring mean equal to 1, based on sample paths with 30 generations. On Figure 8.2 the offspring mean is equal to 0.9 and the number of generations is 25. Comparison for trees with 15 generations and offspring mean 0.6 and 30 generations with offspring mean 0.8 are presented on Figures 8.3 and 8.4 respectively.



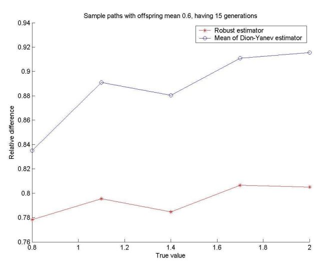
8.1



8.2



8.3



8.4

Figure 8: Comparison between estimators deviations for different parameters of the branching processes.

It is seen that the robust estimates have lower deviation values than those of Dion and Yanev. Only in the critical case (Figure 8.1) the deviation of Dion - Yanev estimate is lower than the robust one.

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`.

REFERENCES

[1] Atanasov D.V. and Neykov N.M On the Finite Sample Breakdown Point of the WLTE(k) and d-fullness of a Set of Continuous Functions In: *Proceedings of the VI International Conference "Computer Data Analysis And Modeling"* Minsk, Belarus (2001), 52–57.

[2] Dion J.P. Statistical Inference for descrete time branching processes. Lecture Notes. 7th International Summer School on Probability Theory and Mathematical Statistics. ed. A. Obretenov, V.T. Stefanov (1993), 60–121.

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

[4] Dion, J.P., Yanev, N.M. Estimation theory for the variance in a branching process with an increasing random number of ancestors *C.R. Acad.Bulg.Sci.* **45(11)** (1992a), 27–30.

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

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

[7] Heyde C.C., Leslie J.R. Improved classical limit analogues for Galton - Watson process with or without immigration. *Bull. Aus tral. Math. Soc.* **5** (1971), 145–156.

[8] Heyde C.C., Seneta E. Estimation theory for growth and immigration rates in a multiplicative process. *J. Appl. Prob* **9** (1972), 235–256.

[9] Heyde C.C. On estimating the variance of the offspring distribution in a simple branching process. Adv. Appl. Prob. **3** (1974), 421–433.

[10] Stoimenova, V. Robust Parametric Estimation of Branching Processes with Random Number of Ancestors *Serdica Math. J.* **31(3)** (2005), 243–262.

[11] Stoimenova V., Atanasov D., Yanev N. Robust Estimation and Simulation of Branching Processes. *C. R. Acad. Bulg. Sci.* **57(5)**, (2004a), 19–22.

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

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

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

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

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

[17] Vandev D.L. and Neykov, N.M. About Regression Estimators with High Breakdown Point. *Statistics* (1998).

[18] Wei C.Z., Winnicki J. Some asymptotic results of branching processes with immigration. *Ann. Statist.* **31** (1989), 261–282.

[19] Wei C.Z., Winnicki J. Estimations of the means in the branching processes with immigration. *Ann. Statist.* **18** (1990), 1757–1773.

[20] Winnicki J. Estimation Theory of the Branching Processes with Immigration. *Contemporary Mathematics* **80** (1988), 301–323.

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

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

[23] Yanev, N. M., Tchoukova-Dancheva, C. On the Statistics of Branching Processes with Immigration. *C. R. Acad. Bulg. Sci.* **33(4)** (1980), 469–571.

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

Vessela Stoimenova
Sofia University,
Faculty of Mathematics and Informatics
5 J. Boucher Str. 1164 Sofia, Bulgaria
and
Institute of Mathematics and Informatics,
Bulgarian Academy of Science,
Acad. G. Bontchev Str. 1113 Sofia,Bulgaria
e-mail: stoimenova@fmi.uni-sofia.bg

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