On the statistics of branching processes with a random number of ancestors

Vessela Stoimenova
Sofia University,
Faculty of Mathematics and Informatics,
Department of Probability, Operations Research and Statistics,
5 J. Boucher Str. 1164 Sofia, Bulgaria

Keywords: offspring mean, individual distribution, robust estimation
AMS: 60J80

Abstract

In the present paper we consider the robust estimation in the Bienayme - Galton - Watson processes with a random number of ancestors (BGWR processes) in the sense of the weighted and trimmed likelihood. Robust modification of the classical estimators of the individual mean and weighted and trimmed estimators of the parameter of the power series offspring distribution are described.

1. Introduction

The branching stochastic processes may be considered as a mathematical model of many real phenomena of reproduction and transformation of different objects in physics, chemistry, biology and so on. Those objects are usually referred to as particles or individuals. They transform or reproduce according to some stochastic laws. Today branching processes represent one of the most developed parts of the stochastic processes with applications in different scientific and practical areas. The focus of this paper is on the robust statistical estimation (in the sense of the trimmed and weighted likelihood) in Bienayme - Galton - Watson processes with an increasing random number of ancestors.

2. The Bienayme - Galton - Watson branching stochastic process with an increasing random number of ancestors

Definition 1 Assume there exists on some probability space a set of i.i.d. r.v. $\{\xi_i(t,n)\}$ with values in the set of nonnegative integers $N = \{0,1,2,\cdots\}$; $\xi_i(t,n),\ i \in N$ are independent of $Z_0(n)$ such that

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

Then for each $n = 1,2,\cdots$ $Z(n) = \{Z_t(n), t = 0,1,\cdots\}$ is a Bienayme-Galton-Watson process having a random number of ancestors $Z_0(n) \geq 1$. Such a process is denoted BGWR.

Here $\xi_i(t+1,n)$ is the number of the particles - descendants in the $(t+1)$-st generation, born from the i-th particle, which exists in the t-th generation.

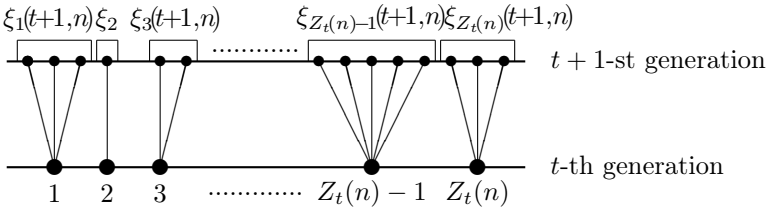

Fig.1 The evolution of a BGWR process

The main local characteristics of the process are the individual distribution $p_k = P(\xi = k) \geq 0$, $\sum_{k=0}^{\infty} p_k = 1$, $p_0 + p_1 < 1$, the offspring mean $m = E\xi = \sum_{k=0}^{\infty} p_k k$ and the offspring variance $\sigma^2 = D\xi = \sum_{k=0}^{\infty} p_k (k-m)^2$. Further on we suppose that $0 < \sigma^2 < \infty$. The asymptotic behaviour of the process depends on these characteristics and that is why the problem of their statistical estimation is so important.

The BGWR discrete time stochastic process is called *subcritical*, if $m < 1$, *critical*, if $m = 1$ and *supercritical*, if $m > 1$.

The most popular models to be studied in the statistics of the branching processes are two: when data about the entire family tree are available $\{\xi_i(t,n)\}$ and when only the consecutive family sizes $\{Z_0(n), Z_1(n), \ldots, Z_t(n)\}$ can be observed.

The two estimators of the individual mean m, which we need later, are the Harris and the Lotka-Nagaev estimators. They are based on observations over the family sizes.

The nonparametric mle (Harris, 1948) of the offspring mean m is given by the formula

$$\widehat{m}_{t+1}(n) \quad = \quad \frac{Y_{t+1}(n) - Z_0(n)}{Y_t(n)}, \tag{1}$$

where $Y_t(n) = \sum_{j=0}^{t-1} Z_j(n)$ is the total number of particles up to the moment $t-1$.

The Lotka-Nagaev estimator depends only on two consecutive generations of the process. It is defined by Lotka (1939):

$$\overline{m}_{t+1}(n) \quad = \quad \begin{cases} \frac{Z_t(n)}{Z_{t-1}(n)} & \text{if} \quad Z_{t-1}(n) > 0, \\ 1 & \text{if} \quad Z_{t-1}(n) = 0. \end{cases} \tag{2}$$

In contrast to the parametric estimation, the nonparametric estimation of m and σ^2 of a BGWR process is thoroughly considered in several works. The nonparametric statistics of BGWR stochastic processes is introduced by Yanev (1975) and thoroughly studied in Dion and Yanev (1991, 1994, 1997).

Yakovlev and Yanev (1989) show that the BGWR processes may occur in the study of cell proliferation and in applications to nuclear chain reactions.

Dion and Yanev (1994) show the duality between the BGWR processes and the Bienayme - Galton - Watson processes with immigration, which results in transferring results on the statistical inference of these processes.

3. Robust parametric estimation of the parameter of the individual distribution in the class of the power series offspring distributions

We remind that the r.v. ξ has a probability distribution, which belongs to the class of the power series offspring distributions (PSOD), if

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

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

3.1. Robust statistical procedures and the d-fullness technique

Vandev and Neykov (1998) defined the trimmed and weighted estimator in order k (the $WLT(k)$ estimator) as

$$\widetilde{\theta} = \min_{\theta} \sum_{i=1}^{k} -w_i \log f(x_{(i)}, \theta), \tag{4}$$

where the weights $w_i \geq 0, i = 1, \cdots n$, are nonnegative and the trimming factor $k = max\{i : w_i > 0\}, \quad k \geq \frac{n}{2}$. The observations $x_1, \cdots, x_n$ are ordered according to the value of their probability density function $f(x, \theta)$: $\log f(x_{(1)}, \theta) \geq \log f(x_{(2)}, \theta) \geq \cdots \geq \log f(x_{(n)}, \theta)$. The aim of this procedure is to find the estimator $\widetilde{\theta}$ using those observations, which have high values of the density function. That way the observations, which lie near the center of the distribution, are included in the model, and those far from it are considered as outliers. An analogous approach is used by Hadi and Luceño (1997). The $WLT(k)$ estimators are a generalization of the maximum likelihood estimator.

Vandev (1993) and Vandev and Neykov (1998) introduce a technique called d-fullness technique, which allows for an easy study of the breakdownpoint of the $WLT(k)$ estimators.

Definition 2 [Hampel et al. (1986)] Let $X = (X_1, \cdots, X_n)$ be a sample of size n, and $\widetilde{X}$ be the sample, obtained by replacing any m observations by

arbitrary values. The breakdownpoint of the statistics T for the sample X is defined as

$$BP(T,X) \quad = \quad max\{\frac{m}{n} : \sup_{\widetilde{X}} |T(X) - T(\widetilde{X})| < \infty)\}.$$

>From now on under robustness we mean weighted and trimmed likelihood.

3.2. Robust parametric estimators, based on the entire family tree

Let the r.v. $\{\xi_i(t,n)\}$ in the definition of the BGWR process be iid copies of the r.v. ξ, which takes values in the subset $B \subset N = \{0,1,2,\cdots\}$ and let us denote by $|B|$ the cardinality of this subset.

Let $B_{SP} = \{k \in B : \exists(i,j) : \xi_i(t,n) = k, i = 1,\ldots,Y_t(n), j = 0,\ldots,t\}$ be the set of all observed values of ξ in the sample.

We denote by $MinV = min\{i : P(\xi = i) > 0, i \in B\}$ the minimal value of ξ, and by $MaxV = max\{i : P(\xi = i) > 0, i \in B\}$ - the maximal value of ξ. Let $N_{MinV} = \sharp\{(i,s) : \xi_i(s,n) = MinV,\ \xi_i(s,n) \in B_{SP}\}$ be the number of the particles which have a minimal number of descendants and let $N_{MaxV} = \sharp\{(i,s) : \xi_i(s,n) = MaxV,\ \xi_i(s,n) \in B_{SP}\}$ be the number of particles with a maximal number of children in the next generation.

Let us define the set of functions : $LF_{SP} = \{\log(A(\theta)/\theta^i)\}_{i \in B_{SP}}$. Let R be the convergence radius of the power series $A(\theta)$ ($A(\theta) < \infty$, 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 (4), where the finction $f(x_{(i)},\theta)$ is replaced by the corresponding probability $p_{\xi_i(j,n)}$, $i = 0,1,\ldots,Y_t(n)$, $j = 0,\ldots,t-1$, of the power series offspring distribution. The following Theorem 1 is proved in Stoimenova(2005).

Theorem 1 Let us consider a sample of size $Y_t(n)$ observations over the entire family tree of the BGWR process with PSOD. For the $WLT(K)$ estimator (4) the following statements hold:

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

2. If $\theta \in (0,R)$, $R < \infty$, $|B| = \infty$, $A(R) = \infty$ and $Y_t(n) > N_{MinV} + 1$, then the statements from 1. are 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 $d = 1,2,\cdots$.

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

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 $Y_t(n) \geq 3$, $[Y_t(n)+1]/2 \leq K \leq Y_t(n)-1$.

4. A robust modification of the Lotka-Nagaev and Harris estimators

Let us define a robust modification of the estimators (1) and (2) for the unknown offspring mean m in a BGWR process $\{Z_t(n)\}$ over the set of sample paths $\mathbf{Z} = \{\mathbf{Z^{(1)}(n)},\cdots,\mathbf{Z^{(r)}(n)}\}$, where $\{\mathbf{Z^{(r)}(n)}\}$ is a realization of a BGWR process, based on the consecutive family sizes, $r = 1,2,\ldots$.

Let $Est(\mathbf{Z^{(i)}(n)},m)$ be a transformation of the estimator (1) or (2) of the offspring mean m, which gives us *asymptotically normal distribution*:

In the case of the Lotka - Nagaev estimator $Est(\mathbf{Z^{(i)}(n)},m)$ can be written as

$$Est(\mathbf{Z^{(i)}(n)},m)\frac{\sqrt{Z_{t_i}^{(i)}(n)}}{\sigma}(\overline{m}_{t_i}^{(i)}(n)-m) \xrightarrow{d} N(0,1), \quad m \in R^+,$$

where $\overline{m}_{t_i}^{(i)}(n)$ is the Lotka-Nagaev estimator (2) in the $i-$th sample path. We have denoted by $Z_{t_i}^{(i)}(n)$ the number of individuals in the t_i-th (the last nonzero observed or the first zero) generation of the trajectory $\mathbf{Z^{(i)}(n)} = \{Z_0^{(i)}(n), Z_1^{(i)}(n),\ldots,Z_{t_i}^{(i)}(n)\}$. The fixed (not estimable) parameter σ^2 is the individual variance.

Analogously in the case of the Harris estimator (1) $\widehat{m}_{t_i}^{(i)}(n)$ in the i-th sample path one has

$$Est(\mathbf{Z^{(i)}(n)},m) = \frac{\sqrt{Y_{t_i}^{(i)}(n)}}{\sigma}(\widehat{m}_{t_i}^{(i)}(n)-m) \xrightarrow{d} N(0,1), \quad m \in R^+,$$

where $Y_{t_i}^{(i)}(n) = Z_0^{(i)}(n)+\cdots+Z_{t_i-1}^{(i)}(n)$.

Let $f(x)$ be the logarithm of the density function of the standard normal distribution $f(x) = -\frac{1}{2}\log(2\pi) - \frac{x^2}{2}$.

Let ν be a permutation of the indices, such that $f(Est(\mathbf{Z}^{(\nu(\mathbf{1}))}(\mathbf{n}),m)) \geq f(Est(\mathbf{Z}^{(\nu(\mathbf{2}))}(\mathbf{n}),m)) \geq \cdots \geq f(Est(\mathbf{Z}^{(\nu(\mathbf{r}))}(\mathbf{n}),m))$.

Denote by k the trimming factor and by $\omega_i \geq 0$, $i = 1,2,\ldots,r$ - the nonnegative weights.

Then the robust modification of the classical estimators of the offspring mean m is defined by

$$\overline{M}_r(m) \quad = \quad \underset{m\in R^+}{argmin}\sum_{i=1}^{k} -w_i f(Est(\mathbf{Z}^{(\nu(\mathbf{i}))}(\mathbf{n}),m)). \tag{5}$$

The following Theorem 2 is stated in Stoimenova, Atanasov, Yanev (2004).

Theorem 2 The estimator $\overline{M}_r(m)$ of the offspring mean m in a BGWR process exists and its brakdownpoint is not less than $(r-k)/r$, if $r \geq 3$, $(r+1)/2 \leq k \leq r-1$, where r is the number of the observed sample paths.

Acknowledgements: This research was partially supported by the National Science Fund of Bulgaria, Grant No VU-MI-105/2005.

5. Bibliography

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

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

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

[4] Hadi A., Luceño A. (1997). *Maximum Trimmed Likelihood Estimators: A Unified Approach, Examples and Algorithms.* Comput. Statist. and Data Analysis, 25, 251 – 272.

[5] Hampel F., Ronchetti E., Rousseeuw P.,Stahel W. (1986). *Robust Statistics: The Approach Based on Influence Functions.* John Wiley and Sons, New York.

[6] Harris, T. E. (1948). *Branching Processes.* Ann. Math. Statist. 19, 474-494.

[7] Lotka, A., (1939). *Theorie analytique des associations biologiques.* Actualités Sci. Ind. 780, 123-136.

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

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

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

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

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

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