JIRSS (2013)
Vol. 12, No. 2, pp 211-233

On the Maximum Likelihood Estimators for some Generalized Pareto-like Frequency Distribution

Davood Farbod[1]**, Karen Gasparian**[2]

[1]Department of Mathematics, Quchan University of Advanced Technologies Engineering, Quchan, Iran.
[2]Chair of Probability Theory and Mathematical Statistics, Yerevan State University, Yerevan, Armenia.

Abstract. In this paper we consider some four-parametric, so-called *Generalized Pareto-like Frequency Distribution*, which have been constructed using stochastic *Birth-Death Process* in order to model phenomena arising in Bioinformatics (Astola and Danielian, 2007). As examples, two "real data" sets on the number of proteins and number of residues for analyzing such distribution are given. The conditions of coincidence of solution for the system of Likelihood Equations with the Maximum Likelihood Estimators (MLE) for the parameters of this distribution are also investigated. In addition, we propose *Accumulation Method* as a recurrence method for approximate computation of the MLE of the parameters. Simulation studies are done.

Keywords. Accumulation method, birth-death process, generalized Pareto-like frequency distribution, Kolmogorov-Smirnov (K-W) test, Markov chain monte carlo (MCMC), MLE.

MSC: Primary: 62F10, 62F30; Secondary: 62P10, 60J28.

Davood Farbod(✉)(d.farbod@gmail.com), Karen Gasparian(kargasp@gmail.com)
Received: May 2012; Accepted: April 2013

1 Introduction

In large-scale biomolecular systems a basic subject of any statistical inference is characterization of the distributions of object frequencies for a population, so-called *frequency distributions* (see, for example: Kuznetsov, 2002; Danielian and Astola, 2007). For empirical frequency distributions some general statistical facts based on many data sets are extracted. From the mathematical point of view these are: skewness to the right, regular variation at infinity, unimodality, stability by parameters, convexity, etc. (see, for example, Astola and Danielian, 2004, 2010). Any distribution satisfying the statistical facts has a chance to be approved by biologist in order to be applied, at least, in one among great variety of large-scale biomolecular systems (Astola and Danielian, 2007, p.1).

We notice that there are different methods for constructing parametric families of frequency distributions. These methods are: usage of stationary distributions of standard stochastic *Birth-Death Process*; discretizations of stable densities; special functions; etc. For a review see, for example, Astola et al. (2010).

The standard stochastic *Birth-Death Process* with various forms of coefficients is an excellent source for obtaining skewed distributions which in turn are important in modeling different phenomena in many large-scale biomolecular systems: for instance, in protein data sets and the number of expressed genes (Danielian and Astola, 2004; Kuznetsov et al., 2002b). Based on the standard *Birth-Death* models several frequency distributions have been considered for biomolecular applications. We would like to point out, for example, the works of Simon (1955), Irwin (1963), Glanzel and Schubert (1995), Bornholdt and Ebel (2001) and Kuznetsov (2001).

According to variety and diversity of biomolecular sequences new frequency parametric families were needed. Kuznetsov et al. (2002b) and Kuznetsov (2003a) suggested three-parametric *Kolmogorov-Waring Distribution*; Astola and Danielian (2004) and Danielian and Astola (2004) gave a wide generalization of the previous models. In the sequel and based on the statistical facts mentioned above, a four-parametric regular frequency distribution (Also, known as Generalized Pareto-type Frequency Distribution) has been introduced by Astola and Danielian (2007). The paper is devoted to investigate the MLE for the parameters of such distribution.

The remainder of the paper is formed as follows. In Section 2, we give some preliminaries about the method of constructing such frequency distribution using stochastic *Birth-Death Process* and introduce four-parametric *Generalized Pareto-like Frequency Distribution.* Moreover, some notaions (are needed in the paper) are presented. The main results of the paper are proposed in Sections 3, 4 and 5. Section 6 concludes.

2 Preliminaries

Let $\{\varrho(t):\ t \geq 0\}$ be a homogeneous *Markov Process* with continuous time and countable number of states $0, 1, 2, \ldots$. The stationary distribution of the process $\{\varrho(t):\ t \geq 0\}$ exists if and only if (see, for example, Danielian and Astola, 2004; Saaty, 1983)

$$\sum_{n=1}^{\infty}\prod_{m=1}^{n}\frac{\lambda_{m-1}}{\mu_m} < \infty, \tag{1}$$

which takes the form ($\frac{\lambda_{m-1}}{\mu_m}$ gives a sequences of ratios of *Birth* and *Death* coefficients):

$$\begin{cases} p_k = p_0 \cdot \prod_{m=1}^{k}\frac{\lambda_{m-1}}{\mu_m}, \qquad k = 1, 2, ..., \\ p_0 = \left(1 + \sum_{n=1}^{\infty}\prod_{m=1}^{n}\frac{\lambda_{m-1}}{\mu_m}\right)^{-1}. \end{cases} \tag{2}$$

With the help of (1) and (2), Astola and Danielian (2007) constructed the following four-parametric *Generalized Pareto-like Frequency Distribution*:

$$\begin{cases} p_\alpha(k) = \mathbb{P}_\alpha(\xi = k) = [g(\alpha)]^{-1} \cdot \frac{\theta^k}{(k+b)^\rho} \cdot \prod_{m=0}^{k-1}(1 + \frac{c-1}{(m+b)^\rho}), \quad k = 1, 2..., \\ p_\alpha(0) = [g(\alpha)]^{-1} = \left[1 + \sum_{n=1}^{\infty}\frac{\theta^n}{(n+b)^\rho} \cdot \prod_{m=0}^{n-1}(1 + \frac{c-1}{(m+b)^\rho})\right]^{-1}, \end{cases} \tag{3}$$

where $\alpha = (\theta, c, b, \rho)$ is unknown parameter and

$$\alpha \in \Omega = \Big\{\alpha : 0 < \theta < 1, 0 < c < \infty, 0 < b < \infty, 1 < \rho < \infty\Big\}, b^\rho > 1 - c.$$

The role of the parameter θ here is explained by Astola and Danielian (2007, Ch. 4, Th. 4.2); the parameter c so-called *non-linear scale* parameter (or *exponential scale* parameter); the parameter b is a *location* parameter and the parameter ρ characterize the *shape* of the probability function.

One of the most important problems for the model (3) is the statistical analysis of the parameters estimators. In this paper, firstly, we give two real data sets for fitting of the distribution (3) and, secondly, we propose some conditions under which the MLE for the parameter α of distribution (3) coincides with the solution of the system of likelihood equations. Thirdly, in order to estimate the parameters, the approximation for the solution based on *Accumulation Method* is given and simulation studies are proposed as well.

Let us use symbols $E_\alpha(.)$, $Var_\alpha(.)$ and $Cov_\alpha(.,.)$, correspondingly, for the expectation, the variance and the covariance with respect to the distribution $\mathbb{P}_\alpha$. Let also $X^n = (X_1, ..., X_n)$ be a sample corresponding to a random variable ξ with the distribution (3). In the sequel we use the following notations:

$$
\begin{aligned}
h_{\gamma,j}(x,\alpha) &= \sum_{m=0}^{x-1}(m+b)^{-\gamma}[(m+b)^\rho + c - 1]^{-j}, \quad x, j \in \mathbb{N},\ \gamma \in \mathbb{R};\\
l_{\gamma,j,k}(x,\alpha) &= \sum_{m=0}^{x-1}(m+b)^{\gamma}\ [(m+b)^\rho + c - 1]^{-j} \cdot [\ln(m+b)]^k,\\
&\qquad\qquad\qquad\qquad\qquad\qquad j,k \in \mathbb{N},\ \gamma \in \mathbb{R};\\
H(x,\alpha) &= (c-1)h_{1,1}(x,\alpha) + (x+b)^{-1};\\
\Lambda(x,\alpha) &= (c-1)\ l_{0,1,1}(x,\alpha) + \ln(x+b);\\
\overline{X^n} &= \frac{1}{n}\sum_{i=1}^n X_i\ ;\quad \overline{\psi_k^n(b)} = \frac{1}{n}\sum_{i=1}^n (X_i+b)^{-k},\ k \in \mathbb{N};\\
\overline{f^n(\alpha)} &= \frac{1}{n}\sum_{i=1}^n f(X_i,\alpha).
\end{aligned}
$$

3 Fitting of the Distribution

As we have said in the Preliminaries, the model (3) has been constructed using stochastic *Birth-Death Process*. But, Astola and Danielian (2007) were not fitted the model (3) with real data. Here, we propose two examples for fitting the model (3) with some real data sets. Comparing to Kuznetosv et al. (2002b, p. 399) and Kuznetsov (2003b, p. 378), in order to use the probability function (3) to the data, we consider the random variable ξ as *doubly-truncated*. Namely, random variable ξ is restricted from 1 to the maximum observed in each data set. Moreover, we present some graphs of the distribution (3) for some different values of the parameters.

Example 3.1. Let us consider 30 biggest protein clusters for Saccharomyces cerevisiaetwo (Apweiler et al., 2000) as a real data set in the following Table:

Table 1.

124	115	83	69	68	66	52	35	34	33	32	32	28	24	24
21	21	20	20	20	19	19	18	17	17	16	16	14	14	14

Let us assume that the data (Table 1) random variable, ξ, follows the model $p_\alpha(k)$. We obtain the MLE of the parameter $\alpha = (\theta, c, b, \rho)$ in the following:

$$\widehat{\theta} = 0.9903;\ \widehat{c} = 96.0111;\ \widehat{b} = 0.7213;\ \widehat{\rho} = 2.2488.$$

The *p-value* of the *K-S Test* is 0.4271, which does not reject the adequacy of the *Generalized Pareto-like Frequency Distribution* for the number of proteins. In order to give an informal goodness of fit test, we plot the empirical cumulative distribution function (ecdf) and fitted cumulative distribution function (cdf) for the number of proteins data in Figure 1.

Example 3.2. We consider the number of residues for 12 electron transports in globular proteins (Kabsch and Sander, 1983) as a real data set in the following Table:

Table 2.

85	103	103	112	134	82
54	98	138	54	125	99

Supposing the data (Table 2) random variable, ξ, follows the model $p_\alpha(k)$. The MLE of the parameter $\alpha = (\theta, c, b, \rho)$ are obtained as follows:

$$\widehat{\theta} = 0.9724;\ \widehat{c} = 16.1974;\ \widehat{b} = 0.102;\ \widehat{\rho} = 1.2465.$$

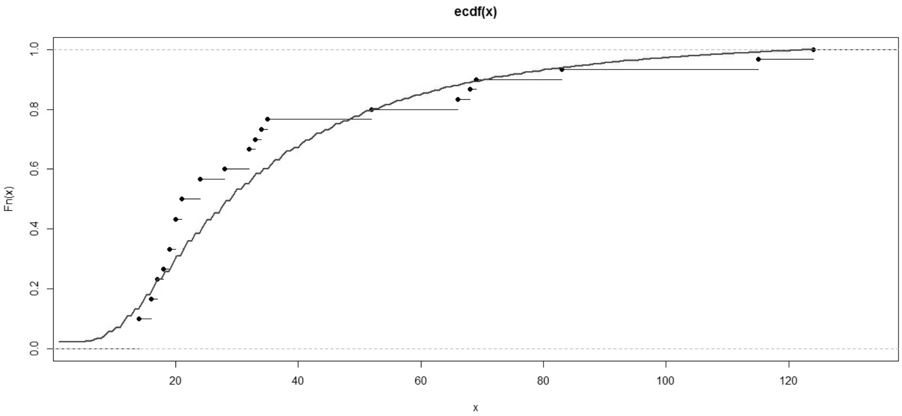

Figure 1: Fitting of the doubly-truncated *Generalized Pareto-like Frequency Distribution* to the data of Table 1. The dashed line is the ecdf of data and the solid line is the fitted cdf.

With the help of *K-S Test* the *p-value* is 0.963, which does not reject the adequacy of the *Generalized Pareto-like Frequency Distribution* for the number of residues. For an informal goodness of fit test, we plot the ecdf and fitted cdf of the number of residues data in Figure 2.

3.1 Figures of the Model

Let us present some plots of the doubly-truncated *Generalized Pareto-like Frequency Distribution* for different values of the parameters in Figure 3. We see that the Plots have right skewness.

4 On the MLE

In this Section, we are going to prove the following Theorems 4.1 and 4.2.

Theorem 4.1. *The system of likelihood equations for finding the MLE of the parameter α for a random variable ξ with the distribution*

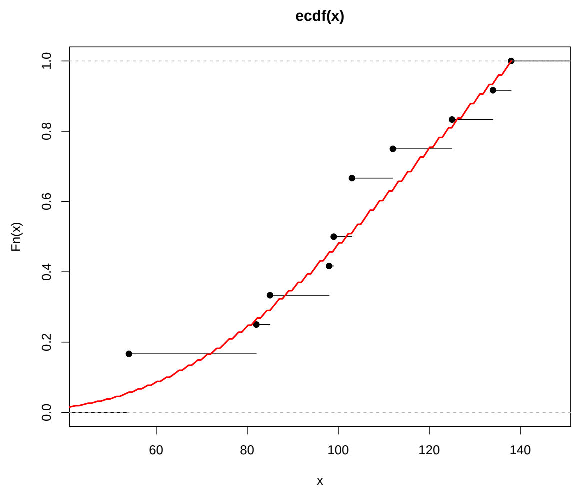

Figure 2: Fitting of the doubly-truncated *Generalized Pareto-like Frequency Distribution* to the data of Table 2. The dashed line is the ecdf of data and the solid line is the fitted cdf.

(3) when $\rho > 3$ *has the form:*

$$\begin{cases} E_\alpha(\xi) = \overline{X^n}, \\ E_\alpha[h_{0,1}(\xi,\alpha)] = \overline{h^n_{0,1}(\alpha)}, \\ E_\alpha[H(\xi,\alpha)] = \overline{H^n(\alpha)}, \\ E_\alpha[\Lambda(\xi,\alpha)] = \overline{\Lambda^n(\alpha)}. \end{cases} \tag{4}$$

Theorem 4.2. *If for* $\rho > 3$ *the solution of the system (4) (if it exists) satisfies the following conditions (we shall call them (A)),*

$$\begin{cases} E_\alpha(\xi+b)^{-k} = \overline{\psi^n_k(b)}, \\ E_\alpha[h_{k,1}(\xi,\alpha)] = \overline{h^n_{k,1}(\alpha)}, \ k=1,2, \\ E_\alpha[h_{k,2}(\xi,\alpha)] = \overline{h^n_{k,2}(\alpha)}, \quad k=0,\ 1-\rho,\ 2-\rho, \\ E_\alpha[l_{\rho,2,k}(\xi,\alpha)] = \overline{l^n_{\rho,2,k}(\alpha)}, \ k=1,2, \\ E_\alpha[l_{\rho-1,2,1}(\xi,\alpha)] = \overline{l^n_{\rho-1,2,1}(\alpha)}, \end{cases}$$

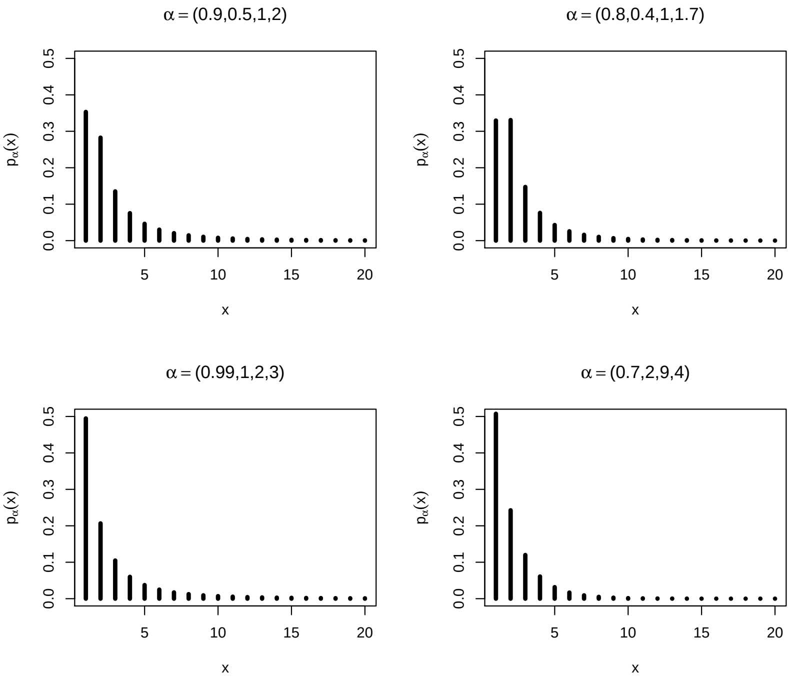

Figure 3: Graphs of the doubly-truncated *Generalized Pareto-like Frequency Distribution* for different values of the parameters θ, c, b, ρ.

then it coincides with the MLE of the parameter α*.*

In order to prove Theorems 4.1 and 4.2 the following Lemma will be used (compare to Astola et. al, 2007; Gasparian and Danielian, 2006):

Lemma 4.1. *a). For the model (3) the following inequalities hold for any* $\alpha \in \Omega$*:*

$$(i) \quad 1 < g(\alpha) < 1 + \Big\{\mathbb{I}_{(0,1]}(c) + \mathbb{I}_{(1,\infty)}(c) \cdot \exp[(c-1)\zeta(\rho, b)]\Big\} \cdot \zeta(\rho, b),$$

where $\zeta(\rho, b) = \sum_{n=0}^{\infty}(n+b)^{-\rho} < \infty$ $\Big(\zeta(\rho, b) < b^{-\rho} + \zeta(\rho), \quad \zeta(\rho) = \sum_{n=1}^{\infty} n^{-\rho} < \infty,$

$\rho > 1,$ *is a Riemann's Zeta Function*$\Big)$*, and* $\mathbb{I}_{(.,.)}(.)$ *is an indicator function;*

(ii) if $\rho > k+1$, $k \in \mathbb{N}$, then

$$E_\alpha(\xi^k) \leq \left\{\mathbb{I}_{(0,1]}(c) + \mathbb{I}_{(1,\infty)}(c) \cdot \exp[(c-1)\zeta_k(\rho,b)]\right\} \cdot \zeta_k(\rho,b),$$

where $\zeta_k(\rho,b) = \sum_{n=1}^{\infty} n^k (n+b)^{-\rho} < \infty$.

b). If $\rho > 3$ then the following means are finite:

(i) $E_\alpha[h_{i,1}(\xi,\alpha)]^2 < \infty, \quad i = 0,1,2;$

(ii) $E_\alpha[h_{i,2}(\xi,\alpha)] < \infty, \quad i = 0, 1-\rho, 2-\rho;$

(iii) $E_\alpha[l_{0,1,1}(\xi,\alpha)]^2 < \infty;$

(iv) $E_\alpha[l_{\rho,2,i}(\xi,\alpha)] < \infty, \ i = 1,2;$

(v) $E_\alpha[l_{\rho-1,2,1}(\xi,\alpha)] < \infty.$

Proof. a). The following inequalities take place when $\alpha \in \Omega$:

(i) according to (3) we have

$$1 < g(\alpha) < 1 + \sum_{n=1}^{\infty} \frac{\theta^n}{(n+b)^\rho} \cdot \exp\left[(c-1) \cdot \sum_{m=0}^{n-1} (m+b)^{-\rho}\right]$$

$$< \begin{cases} 1 + \zeta(\rho,b), & if \quad 0 < c \leq 1, \\ 1 + \zeta(\rho,b) \cdot \exp[(c-1)\zeta(\rho,b)], & if \quad c > 1; \end{cases}$$

(ii) on the other hand, if $\rho > k+1$, $k \in \mathbb{N}$, then

$$E_\alpha(\xi^k) \leq [g(\alpha)]^{-1} \cdot \sum_{n=1}^{\infty} \frac{n^k}{(n+b)^\rho} \cdot \exp\left[(c-1)\sum_{m=0}^{n-1}(m+b)^{-\rho}\right]$$

$$< \begin{cases} \zeta_k(\rho,b), & if \ 0 < c \leq 1, \\ \zeta_k(\rho,b) \cdot \exp[(c-1)\zeta(\rho,b)], & if \ c > 1. \end{cases}$$

b). The following inequalities also hold when $\rho > 3$:

(i) we have for $i = 0,1,2$,

$$E_\alpha[h_{i,1}(\xi,\alpha)]^2 \leq \sum_{n=1}^{\infty} n \cdot \left(\sum_{m=0}^{n-1} \frac{1}{(m+b)^{2i}} \cdot \frac{1}{[(m+b)^\rho + c - 1]^2}\right) \cdot p_\alpha(n)$$

$$< \left(\sum_{m=0}^{\infty} \frac{1}{(m+b)^{2i}} \cdot \frac{1}{[(m+b)^\rho + c - 1]^2}\right) \cdot E_\alpha(\xi) < \infty.$$

The series on the right side is convergent because $\zeta(2(\rho+i),b)<\infty$;

(ii) it is obvious that for $i=0,1-\rho$ *or* $2-\rho$,

$$E_\alpha[h_{i,2}(\xi,\alpha)] < \sum_{m=0}^{\infty}\tfrac{1}{(m+b)^i}\cdot\tfrac{1}{[(m+b)^\rho+c-1]^2}<\infty,$$

where the condition $\zeta(2\rho+i,b)<\infty$ implies the convergence of the series;

(iii) also, we have

$$E_\alpha[l_{0,1,1}(\xi,\alpha)]^2<\Big(\sum_{m=0}^{\infty}\frac{(m+b)^2}{[(m+b)^\rho+c-1]^2}\Big)\cdot E_\alpha(\xi)<\infty,$$

where the series is convergent because $\zeta(2\rho-2,b)<\infty$;

(iv) similarly, there take place $(i=1,2)$

$$E_\alpha[l_{\rho,2,i}(\xi,\alpha)]<\sum_{m=0}^{\infty}\frac{(m+b)^{\rho+i}}{[(m+b)^\rho+c-1]^2}<\infty,$$

in view of the fact that $\zeta(\rho-i,b)<\infty$;

(v) and finally we have

$$E_\alpha[l_{\rho-1,2,1}(\xi,\alpha)]<\sum_{m=0}^{\infty}\frac{(m+b)^{\rho}}{[(m+b)^\rho+c-1]^2}<\infty,$$

as $\zeta(\rho,b)<\infty$.

Thus the proof of Lemma 4.1 is completed. □

Proof of Theorem 4.1. For a random sample $X^n=(X_1,...,X_n)\sim\mathbb{P}_\alpha$ from the distribution (3) the likelihood function will be

$$f_\alpha(X^n)=\prod_{i=1}^{n}p_\alpha(X_i)=[g(\alpha)]^{-n}\cdot\frac{\theta^{n\cdot\overline{X^n}}}{\prod_{i=1}^{n}(X_i+b)^\rho}\prod_{i=1}^{n}\prod_{m=0}^{X_i-1}(1+\frac{c-1}{(m+b)^\rho})$$

such that for the logarithm of likelihood function we obtain

$$\begin{aligned}L_\alpha(X^n)&=\ln f_\alpha(X^n)\\&=-n\ln g(\alpha)+n\cdot\overline{X^n}\ln\theta-\rho\textstyle\sum_{i=1}^{n}\ln(X_i+b)\\&\quad+\textstyle\sum_{i=1}^{n}\sum_{m=0}^{X_i-1}\ln(1+\frac{c-1}{(m+b)^\rho}).\end{aligned}$$

The necessary conditions for existence of the MLE for the parameter α are

$$\frac{\partial L_\alpha(X^n)}{\partial \alpha_i} = 0, \quad i = 1 - 4,$$

where $\alpha_1 = \theta$, $\alpha_2 = c$, $\alpha_3 = b$, $\alpha_4 = \rho$.

The derivatives with respect to parameters are as follows:

$$\frac{\partial L_\alpha(X^n)}{\partial \theta} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \theta} + \frac{n \cdot \overline{X^n}}{\theta}, \tag{5}$$

where (see (3))

$$\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \theta} = \frac{1}{\theta} \cdot E_\alpha(\xi), \tag{6}$$

such that the condition $\frac{\partial L_\alpha(X^n)}{\partial \theta} = 0$ implies

$$E_\alpha(\xi) = \overline{X^n}.$$

Also, we have

$$\frac{\partial L_\alpha(X^n)}{\partial b} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial b} - n\rho \cdot \overline{\psi_1^n(b)} - n\rho(c-1)\overline{h_{1,1}^n(\alpha)}, \tag{7}$$

where

$$\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial b} = -\rho \cdot E_\alpha(\xi + b)^{-1} - \rho(c-1)E_\alpha[h_{1,1}(\xi, \alpha)] = -\rho \cdot E_\alpha[H(\xi, \alpha)], \tag{8}$$

and from the condition $\frac{\partial L_\alpha(X^n)}{\partial b} = 0$ we obtain

$$E_\alpha[H(\xi, \alpha)] = \overline{H^n(\alpha)}.$$

Besides, we have

$$\frac{\partial L_\alpha(X^n)}{\partial c} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial c} + n \cdot \overline{h_{0,1}^n(\alpha)}, \tag{9}$$

where

$$\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial c} = E_\alpha[h_{0,1}(\xi, \alpha)], \tag{10}$$

such that from the condition $\frac{\partial L_\alpha(X^n)}{\partial c} = 0$ we obtain

$$E_\alpha[h_{0,1}(\xi, \alpha)] = \overline{h_{0,1}^n(\alpha)}.$$

Finally, we have

$$\frac{\partial L_\alpha(X^n)}{\partial \rho} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \rho} - n \cdot \overline{\ln(X^n + b)} - n(c-1)\ \overline{l^n_{0,1,1}(\alpha)}, \quad (11)$$

where

$$\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \rho} = -E_\alpha[\ln(\xi + b)] - (c-1) \cdot E_\alpha[l_{0,1,1}(\xi, \alpha)] = -E_\alpha[\Lambda(\xi, \alpha)], \quad (12)$$

and the condition $\frac{\partial L_\alpha(X^n)}{\partial \rho} = 0$ implies

$$E_\alpha[\Lambda(\xi, \alpha)] = \overline{\Lambda^n(\alpha)}.$$

Theorem 4.1 is proved. □

Let us now prove that the solution $\widehat{\alpha} = \widehat{\alpha}^n = (\widehat{\alpha}^n_i)^4_{i=1}$ of the system (4) (if it exists) is the MLE of the parameter α. In order to do that, it is enough to show that the matrix

$$\widehat{L}^n = \left(\widehat{L}^n_{ij}\right)^4_{i,j=1}$$

with $\widehat{L}^n_{ij} = L^n_{ij}(\widehat{\alpha})$, $\quad L^n_{ij}(\widehat{\alpha}) = \frac{\partial^2 L_\alpha(X^n)}{\partial\alpha_i\ \partial\alpha_j}\ |_{\alpha=\widehat{\alpha}}$, is *negative definite.*

Lemma 4.2. *If for $\rho > 3$ the solution $\widehat{\alpha}$ of the system (4) (if it exists) satisfies the Conditions (A), then the elements of the matrix $\widehat{L}^n$ are:*

$\widehat{L}^n_{11} = -\frac{n}{\widehat{\theta}^2} \cdot Var_{\widehat{\alpha}}(\xi)$, $\qquad \widehat{L}^n_{12} = \widehat{L}^n_{21} = -\frac{n}{\widehat{\theta}} \cdot Cov_{\widehat{\alpha}}[\xi, h_{0,1}(\xi, \widehat{\alpha})]$,

$\widehat{L}^n_{13} = \widehat{L}^n_{31} = \frac{n\widehat{\rho}}{\widehat{\theta}} \cdot Cov_{\widehat{\alpha}}[\xi, H(\xi, \widehat{\alpha})]$, $\qquad \widehat{L}^n_{14} = \widehat{L}^n_{41} = \frac{n}{\widehat{\theta}} \cdot Cov_{\widehat{\alpha}}[\xi, \Lambda(\xi, \widehat{\alpha})]$,

$\widehat{L}^n_{22} = -n \cdot Var_{\widehat{\alpha}}[h_{0,1}(\xi, \widehat{\alpha})]$, $\widehat{L}^n_{23} = \widehat{L}^n_{32} = n\widehat{\rho} \cdot Cov_{\widehat{\alpha}}[h_{0,1}(\xi, \widehat{\alpha}), H(\xi, \widehat{\alpha})]$,

$\widehat{L}^n_{24} = \widehat{L}^n_{42} = n \cdot Cov_{\widehat{\alpha}}[h_{0,1}(\xi, \widehat{\alpha}), \Lambda(\xi, \widehat{\alpha})]$, $\quad \widehat{L}^n_{33} = -n\widehat{\rho}^2 \cdot Var_{\widehat{\alpha}}[H(\xi, \widehat{\alpha})]$,

$\widehat{L}^n_{34} = \widehat{L}^n_{43} = -n\widehat{\rho} \cdot Cov_{\widehat{\alpha}}[H(\xi, \widehat{\alpha}), \Lambda(\xi, \widehat{\alpha})]$, $\quad \widehat{L}^n_{44} = -n \cdot Var_{\widehat{\alpha}}[\Lambda(\xi, \widehat{\alpha})]$,

such that the matrix $\widehat{L}^n$ can be represented as $\widehat{L}^n = n \cdot \widehat{M^n}$.

Lemma 4.3. *If the conditions of Lemma 4.2 are valid then the matrix $\widehat{L}^n$ is negative definite.*

Proof of Lemma 4.2.

1. From the representation (5) we obtain

$$L_{11}^n(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial\theta^2} = n\,\Big[\frac{1}{g^2}\cdot(\frac{\partial g}{\partial\theta})^2 - \frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta^2} - \frac{\overline{X^n}}{\theta^2}\Big],$$

$$L_{1j}^n(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial\theta\partial\alpha_j} = n\,\Big[\frac{1}{g^2}\cdot\frac{\partial g}{\partial\theta}\cdot\frac{\partial g}{\partial\alpha_j} - \frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta\partial\alpha_j}\Big],\quad j=2,3,4,$$

where by (6) it follows that

$$\frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta^2} = \frac{1}{\theta^2}\cdot E_\alpha[\xi(\xi-1)],\quad \frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta\partial c} = \frac{1}{\theta}\cdot E_\alpha[\xi\cdot h_{0,1}(\xi,\alpha)],$$

$$\frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta\partial b} = -\frac{\rho}{\theta}\cdot E_\alpha[\xi\cdot H(\xi,\alpha)],\quad \frac{1}{g}\cdot\frac{\partial^2 g}{\partial\theta\partial\rho} = -\frac{1}{\theta}\cdot E_\alpha[\xi\cdot\Lambda(\xi,\alpha)],$$

such that according to (6), (8), (10) and (12) we receive

$$L_{11}^n(\alpha) = \tfrac{n}{\theta^2}\cdot\Big\{(E_\alpha(\xi) - \overline{X^n}) - Var_\alpha(\xi)\Big\},$$

$$L_{12}^n(\alpha) = L_{21}^n(\alpha) = -\tfrac{n}{\theta}\cdot Cov_\alpha[\xi, h_{0,1}(\xi,\alpha)],$$

$$L_{13}^n(\alpha) = L_{31}^n(\alpha) = \tfrac{n\rho}{\theta}\cdot Cov_\alpha[\xi, H(\xi,\alpha)],$$

$$L_{14}^n(\alpha) = L_{41}^n(\alpha) = \tfrac{n}{\theta}\cdot Cov_\alpha[\xi, \Lambda(\xi,\alpha)].$$

2. It follows from the representation (9) that

$$L_{22}^n(\alpha) = \tfrac{\partial^2 L_\alpha(X^n)}{\partial c^2} = n\,\Big[\tfrac{1}{g^2}\cdot(\tfrac{\partial g}{\partial c})^2 - \tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c^2}\Big] - n\cdot\overline{h_{0,2}^n(\alpha)},$$

$$L_{23}^n(\alpha) = \tfrac{\partial^2 L_\alpha(X^n)}{\partial c\,\partial b} = n\,\Big[\tfrac{1}{g^2}\cdot\tfrac{\partial g}{\partial c}\cdot\tfrac{\partial g}{\partial b} - \tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c\,\partial b}\Big] - n\rho\cdot\overline{h_{1-\rho,2}^n(\alpha)},$$

$$L_{24}^n(\alpha) = \tfrac{\partial^2 L_\alpha(X^n)}{\partial c\,\partial\rho} = n\,\Big[\tfrac{1}{g^2}\cdot\tfrac{\partial g}{\partial c}\cdot\tfrac{\partial g}{\partial\rho} - \tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c\,\partial\rho}\Big] - n\cdot\overline{l_{\rho,2,1}^n(\alpha)}.$$

From (10), after not complicated transformations, we obtain

$$\tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c^2} = E_\alpha[h_{0,1}(\xi,\alpha)]^2 - E_\alpha[h_{0,2}(\xi,\alpha)],$$

$$\tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c\,\partial b} = -\rho\,\Big\{E_\alpha[h_{0,1}(\xi,\alpha)\cdot H(\xi,\alpha)] + E_\alpha[h_{1-\rho,2}(\xi,\alpha)]\Big\},$$

$$\tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial c\,\partial\rho} = -\Big\{E_\alpha[h_{0,1}(\xi)\cdot\Lambda(\xi,\alpha)] + E_\alpha[l_{\rho,2,1}(\xi,\alpha)]\Big\}.$$

Then, according to (8), (10) and (12) we have

$$L_{22}^n(\alpha) = n\left\{(E_\alpha[h_{0,2}(\xi,\alpha)] - \overline{h_{0,2}^n(\alpha)}) - Var_\alpha[h_{0,1}(\xi,\alpha)]\right\},$$

$$L_{23}^n(\alpha) = L_{32}^n(\alpha) = n\rho\left\{(E_\alpha[h_{1-\rho,2}(\xi,\alpha)] - \overline{h_{1-\rho,2}^n(\alpha)}) + Cov_\alpha[h_{0,1}(\xi,\alpha), H(\xi,\alpha)]\right\},$$

$$L_{24}^n(\alpha) = L_{42}^n(\alpha) = n\left\{(E_\alpha[l_{\rho,2,1}(\xi,\alpha)] - \overline{l_{\rho,2,1}^n(\alpha)}) + Cov_\alpha[h_{0,1}(\xi,\alpha), \Lambda(\xi,\alpha)]\right\}.$$

3. From (8) it follows that

$$L_{33}^n(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial b^2} = n\Big[\frac{1}{g^2}\cdot(\frac{\partial g}{\partial b})^2 - \frac{1}{g}\cdot\frac{\partial^2 g}{\partial b^2}\Big] + n\rho\Big[(c-1)\cdot\overline{h_{2,1}^n(\alpha)} + \overline{\psi_2^n(b)}\Big] + n\rho^2(c-1)\cdot\overline{h_{2-\rho,1}^n(\alpha)},$$

$$L_{34}^n(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial b\partial\rho} = n\Big[\frac{1}{g^2}\cdot\frac{\partial g}{\partial b}\cdot\frac{\partial g}{\partial \rho} - \frac{1}{g}\cdot\frac{\partial^2 g}{\partial b\partial\rho}\Big] - n\cdot\overline{H^n(\alpha)} + n\rho(c-1)\cdot\overline{l_{\rho-1,2,1}^n(\alpha)}.$$

According to (8) we obtain

$$\frac{1}{g}\cdot\frac{\partial^2 g}{\partial b^2} = \rho^2 E_\alpha[H(\xi,\alpha)]^2 + \rho E_\alpha(\xi+b)^{-2} + \rho(c-1)E_\alpha[h_{2,1}(\xi,\alpha)] + \rho^2(c-1)\cdot E_\alpha[h_{2-\rho,2}(\xi,\alpha)],$$

$$\frac{1}{g}\cdot\frac{\partial^2 g}{\partial b\partial\rho} = \rho\ Cov_\alpha[H(\xi,\alpha),\Lambda(\xi,\alpha)] - E_\alpha[H(\xi,\alpha)] + \rho(c-1)E_\alpha[l_{\rho-1,2,1}(\xi,\alpha)],$$

such that by (8), (12) and with the help of not difficult transformations we receive

$$
\begin{aligned}
L_{33}^n(\alpha) = & -n\rho^2 \cdot Var_\alpha[H(\xi,\alpha)] \\
& + n\rho^2(c-1)\Big\{\overline{h_{2-\rho,2}^n(\alpha)} - E_\alpha[h_{2-\rho,2}(\xi,\alpha)]\Big\} \\
& + n\rho\Big\{\overline{\psi_2^n(b)} - E_\alpha(\xi+b)^{-2}\Big\} \\
& + n\rho(c-1)\Big\{\overline{h_{2,1}^n(\alpha)} - E_\alpha[h_{2,1}(\xi,\alpha)]\Big\},
\end{aligned}
$$

$$
\begin{aligned}
L_{34}^n(\alpha) = & -n\rho\ Cov_\alpha[H(\xi,\alpha),\Lambda(\xi,\alpha)] + n\Big\{E_\alpha[H(\xi,\alpha)] - \overline{H^n(\alpha)}\Big\} \\
& + n\rho(c-1)\Big\{E_\alpha[l_{\rho-1,2,1}(\xi,\alpha)] - \overline{l_{\rho-1,2,1}^n(\alpha)}\Big\}.
\end{aligned}
$$

4. Finally from (11) we have

$$L_{44}^n(\alpha) = \tfrac{\partial^2 L_\alpha(X^n)}{\partial\rho^2} = n\Big[\tfrac{1}{g^2}\cdot(\tfrac{\partial g}{\partial\rho})^2 - \tfrac{1}{g}\cdot\tfrac{\partial^2 g}{\partial\rho^2}\Big] + n(c-1)\cdot\overline{l_{\rho,2,2}^n(\alpha)}.$$

Then by (12) we obtain

$$\frac{\partial^2 g}{\partial\rho^2} = E_\alpha[\Lambda(\xi,\alpha)]^2 + (c-1)\cdot E_\alpha[l_{\rho,2,2}(\xi,\alpha)]$$

such that

$$L_{44}^n(\alpha) = -n\cdot Var_\alpha[\Lambda(\xi,\alpha)] - n(c-1)\Big\{E_\alpha[l_{\rho,2,2}(\xi,\alpha)] - \overline{l_{\rho,2,2}^n(\alpha)}\Big\}.$$

Assuming now that the solution $\widehat{\alpha} = (\widehat{\theta},\widehat{c},\widehat{b},\widehat{\rho})$ of the system (4) satisfies the *Conditions (A)*, the proof of Lemma 4.2 is finished.

Proof of Lemma 4.3. The necessary and sufficient conditions for *negative definiteness* of the matrix $\widehat{L}^n$ are the following conditions (see, for example, Gantmacher, 2000)

$$\widehat{L}_{11}^n < 0, \quad \Delta_2 \equiv \det\Big(\widehat{L}_{ij}^n\Big)_{i,j=1}^2 > 0,$$

$$\Delta_3 \equiv \det\Big(\widehat{L}_{ij}^n\Big)_{i,j=1}^3 < 0, \quad \Delta_4 \equiv \det\widehat{L}^n > 0.$$

By Lemma 4.2 we have

$$\widehat{L}_{11}^n = -\frac{n}{\widehat{\theta}^2}\cdot Var_{\widehat{\alpha}}(\xi) < 0$$

in view of $\xi \neq constant$, $\mathbb{P} - a.s.$ (a.s. means almost surely).

Then,

$$\Delta_2 = \widehat{L}^n_{11} \cdot \widehat{L}^n_{22} - (\widehat{L}^n_{12})^2$$

$$= (\tfrac{n}{\widehat{\theta}})^2 \cdot \Big\{ Var_{\widehat{\alpha}}(\xi) \cdot Var_{\widehat{\alpha}}[h_{0,1}(\xi, \widehat{\alpha})] - Cov^2_{\widehat{\alpha}}[\xi, h_{0,1}(\xi, \widehat{\alpha})] \Big\}.$$

Hence by Cauchy-Schwarz inequality it follows that $\Delta_2 \geq 0$. But, because

$$h_{0,1}(\xi, \alpha) \neq b(\alpha) + c(\alpha)\xi \quad \mathbb{P}_\alpha - a.s.,$$

for some $b(\alpha)$ and $c(\alpha)$, then $\Delta_2 > 0$.

On the other hand, we have

$$\Delta_3 = -\frac{n^3\, \widehat{\rho}^2}{\widehat{\theta}^2} \cdot \det\Big(\widehat{D}_3\Big), \quad \widehat{D}_3 = \Big(\widehat{q}_{ij}\Big)^3_{i,j=1},$$

where

$$\widehat{q}_{ij} \equiv q_{ij}(\widehat{\alpha}) = Cov_{\widehat{\alpha}}(\xi_i, \xi_j), \quad \xi_1 = \xi, \ \xi_2 = h_{0,1}(\xi, \widehat{\alpha}), \ \xi_3 = H(\xi, \widehat{\alpha}).$$

Hence, because $\det\Big(\widehat{D}_3\Big) \geq 0$ ($\widehat{D}_3$ is a *variance-covariance matrix*), then $\Delta_3 \leq 0$. But since there does not exist such $c_i(\alpha) \neq 0$, $i = 1, 2, 3$, and $b(\alpha) \in \mathbb{R}$ that

$$\sum_{i=1}^{3} c_i(\alpha)\xi_i = b(\alpha) \qquad \mathbb{P}_\alpha - a.s.,$$

then $\det\Big(\widehat{D}_3\Big) \neq 0 \ \ \mathbb{P}_\alpha - a.s.$.

Finally, we have

$$\Delta_4 = \frac{n^4\, \widehat{\rho}^2}{\widehat{\theta}^2} \cdot \det\Big(\widehat{D}_4\Big), \quad \widehat{D}_4 = \Big(\widehat{q}_{ij}\Big)^4_{i,j=1},$$

where $\widehat{q}_{ij} = Cov_{\widehat{\alpha}}(\xi_i, \xi_j)$, ξ_1, ξ_2, ξ_3 are as before and $\xi_4 = \Lambda(\xi, \widehat{\alpha})$. So, as above, it can be showed that $\det\Big(\widehat{D}_4\Big) > 0$ and the proof of Lemma 4.3 is finished.

The proof of Theorem 4.2 now immediately follows from Lemmas 4.2 and 4.3.

Remark 4.1. Note that the *Conditions (A)* can be omitted if n is sufficiently large. Really, in general case, if the *Conditions (A)* are not satisfied, one can represent the matrix $\widehat{L}^n$ as (see Lemma 4.2)

$$\begin{aligned}\widehat{L}^n &= n \cdot \widehat{M}^n \\ &= n \cdot \Big\{ \widehat{M}^n(0) - \widehat{\delta}_2^n \cdot (e_{22}) + \widehat{\rho}^n \cdot \widehat{\delta}_3^n \cdot (e_{33}) + (\widehat{c}_n - 1) \cdot \widehat{\delta}_4^n \cdot (e_{44}) \\ &\quad - \widehat{\rho}_n \cdot \widehat{\delta}_5^n \cdot (e_{23}) - \widehat{\delta}_6^n \cdot (e_{24}) - \widehat{\delta}_7^n \cdot (e_{34}) \Big\},\end{aligned}$$

where $\widehat{M}^n(0)$ is the matrix $\widehat{M}^n$ under *the Conditions (A)*; $\Big(e_{ij}\Big),\ i,j = 2,3,4$, are 4×4-dimensional matrices all elements of which are zero except the elements $e_{ij} = e_{ji} = 1$,

$$\begin{aligned}\widehat{\delta}_2^n &= \overline{h_{0,2}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[h_{0,2}(\xi, \widehat{\alpha})], \\ \widehat{\delta}_3^n &= \Big\{ \overline{\psi_2^n(b)} - E_{\widehat{\alpha}}(\xi + \widehat{b})^{-2} \Big\} + (\widehat{c}_n - 1)\Big\{ \overline{h_{2,1}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[h_{2,1}(\xi, \widehat{\alpha})] \Big\} \\ &\quad + \widehat{\rho}_n(\widehat{c}_n - 1)\Big\{ \overline{h_{2-\rho,2}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[h_{2-\rho,2}(\xi, \widehat{\alpha})] \Big\}, \\ \widehat{\delta}_4^n &= \overline{l_{\rho,2,2}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[l_{\rho,2,2}(\xi, \widehat{\alpha})], \\ \widehat{\delta}_5^n &= \overline{h_{1-\rho,2}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[h_{1-\rho,2}(\xi, \widehat{\alpha})], \\ \widehat{\delta}_6^n &= \overline{l_{\rho,2,1}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[l_{\rho,2,1}(\xi, \widehat{\alpha})], \\ \widehat{\delta}_7^n &= \Big\{ \overline{\psi_1^n(b)} - E_{\widehat{\alpha}}(\xi + \widehat{b})^{-1} \Big\} + (\widehat{c}_n - 1)\Big\{ \overline{h_{1,1}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[h_{1,1}(\xi, \widehat{\alpha})] \Big\} \\ &\quad + \widehat{\rho}_n(\widehat{c}_n - 1)\Big\{ \overline{l_{\rho-1,2,1}^n(\widehat{\alpha})} - E_{\widehat{\alpha}}[l_{\rho-1,2,1}(\xi, \widehat{\alpha})] \Big\}.\end{aligned}$$

For n sufficiently large by the Strong Law of Large Numbers, a random variable $\widehat{\delta}_i^n,\ i = 2-7$, becomes a.s. small enough such that all principal minors of the matrix $\widehat{M}^n$ preserves the respective signs of the principal minors of the *negative definite* matrix $\widehat{M}^n(0)$. Thus the matrix $\widehat{M}^n$ (and also $\widehat{L}^n$) for n sufficiently large is also a.s. *negative definite* (see Seber and Lee, 2003, Appendix A.4).

5 Accumulation Method for Approximate Computation of the MLE

As we see from the system (4), it is hard to solve and obtain an explicit form for the solution. To overcome this problem (compare to Astola et al., 2007; Gasparian and Danielian, 2006), we suggest the *Fisher's Accumulation Method* (see, for example, Ivchenko and Medvedev, 1990, p. 88) as an approximate method. In this Section, without loss of generality and only for simplicity, let us assume $c = 1$.

Let $\alpha(0) = (\alpha_i)_{i=1,3,4}$ ($\alpha(0)$ is the initial value of the parameter α) be any consistent estimator for unknown parameter $\alpha = (\alpha_i)_{i=1,3,4} \in \Omega$, where $\alpha_1 = \theta$, $\alpha_3 = b$, $\alpha_4 = \rho$, (here $\alpha_2 = c = 1$). We consider the *Contribution Function* $U(\alpha) = (U_j(\alpha))_{j=1,3,4}$ with components $U_j(\alpha) = \frac{\partial L_\alpha(x^n)}{\partial \alpha_j}$, where $x^n = (x_1, ..., x_n)$ is a realization of a sample $X^n \sim \mathbb{P}_\alpha$.

We expand the components of the function $U(\alpha)$ into a Taylor Series around the point $\alpha(0)$, chosen as the first approximation for $\widehat{\alpha} = (\widehat{\alpha}_j)_{j=1,3,4}$. Then, since $U(\widehat{\alpha}) = 0$, we obtain (compare to Astola et al., 2007)

$$U_j(\alpha(0)) + \sum_{m=1,3,4} \frac{\partial U_j(\widetilde{\alpha})}{\partial \alpha_m} (\widehat{\alpha}_m - \alpha_m(0)) = 0, \quad j = 1, 3, 4, \tag{13}$$

where $\widetilde{\alpha}$ is some intermediate point between $\alpha(0)$ and $\widehat{\alpha}$.
Replacing now $\widetilde{\alpha}$ into $\alpha(0)$ and $\frac{\partial U_j(\widetilde{\alpha})}{\partial \alpha_m}$ into $E_{\alpha(0)}\left[\frac{\partial U_j(\alpha(0))}{\partial \alpha_m}\right] = -n \cdot I_{jm}(\alpha(0))$ where $I(\alpha) = \Big(I_{jm}(\alpha)\Big)_{j,m=1,3,4}$ is a Fisher's Information Matrix $\left(I_{jm}(\alpha) = -E_\alpha\left[\frac{\partial^2 L_\alpha(X_1)}{\partial \alpha_j \partial \alpha_m}\right]\right)$, the system (13) transforms into the following one:

$$\sum_{m=1,3,4} I_{jm}(\alpha(0)) \cdot (\widehat{\alpha}_m - \alpha_m(0)) = \frac{1}{n} \cdot U_j(\alpha(0)), \quad j = 1, 3, 4.$$

The solution of this system gives the first approximation $\alpha(1) = (\alpha_j(1))_{j=1,3,4}$ for $\widehat{\alpha}$:

$$\alpha_j(1) = \alpha_j(0) + \frac{\Delta_j(\alpha(0))}{n \cdot \det I(\alpha(0))}, \quad j = 1, 3, 4,$$

where $\Delta_j(\alpha(0))$ is determinant of the matrix, obtained by the matrix $I(\alpha(0))$, replacing the j_{th} column by the vector $U(\alpha)$.

Then, the $(z+1)_{th}$ approximation can be found recurrently by the formula

$$\alpha_j(z+1) = \alpha_j(z) + \frac{\Delta_j(\alpha(z))}{n \cdot \det I(\alpha(z))}, \quad j = 1,3,4, \quad z = 0,1,2,... \, . \tag{14}$$

The formula (14), in details, may be given as follows:

$$\begin{cases} \theta(z+1) = \theta(z) + \frac{\Delta_1(\alpha(z))}{n \cdot \det I(\alpha(z))}, \\ \\ b(z+1) = b(z) + \frac{\Delta_3(\alpha(z))}{n \cdot \det I(\alpha(z))}, \\ \\ \rho(z+1) = \rho(z) + \frac{\Delta_4(\alpha(z))}{n \cdot \det I(\alpha(z))}. \end{cases} \tag{15}$$

By using the formula (14), we propose the following Algorithm ($c = 1$).

Algorithm I.

1. Generate data based on the MCMC method;

2. Use (14) (or (15)) in order to calculate $\alpha_j(z)$, $j = 1,3,4$, $z = 0,1,2,...$;

3. If $|\alpha_j(z+1) - \alpha_j(z)| < \varepsilon$ (ε is some small positive constant), then $\alpha_j(z+1) = \widehat{\alpha}$ is the MLE, otherwise go to the step 2.

5.1 Simulation study

Until now, to the authors knowledge, there has not been proposed any closed form for the cdf of the model (3) (even for the case $c = 1$). So, it is not possible to generate data based on the cdf. In order to overcome this difficulty, we use MCMC method (see, for example, Given and Hoeting, 2005; Rizzo, 2008). The same as Section 3, let us have the random variable ξ as *doubly-truncated*. Here, random variable ξ is restricted from 1 to 100.

To investigate the behavior of the MLE, obtained by the formula (15), using MCMC we do simulation for 1000 times, i.e. $M = 1000$ (M is the number of iteration) and for $N = 50, 100$ (N is the sample size). We consider $\varepsilon = 0.0005$.

Note. Notice that the following values are considered as true values of the parameters:

$$\alpha = (\theta = 0.9, b = 2, \rho = 3.6), \quad \alpha = (\theta = 0.99, b = 3, \rho = 4.5).$$

Now, from *Algorithm I* the following Table is proposed.

Table 3. The average of estimated mean squared errors (EMSE).

	N=50		N=100	
Case	Mean	EMSE	Mean	EMSE
$\theta = 0.9$	0.87049	0.00087	0.87112	0.00083
$b = 2$	2.26768	0.07165	2.07161	0.00513
$\rho = 3.6$	3.19627	0.16299	3.42955	0.02905
Iteration	104	—	82	—
$\theta = 0.99$	0.94064	0.00244	0.99911	8.3×10^{-5}
$b = 3$	3.11567	0.01338	2.99373	3.9×10^{-5}
$\rho = 4.5$	4.56459	0.00417	4.70897	4.3×10^{-2}
Iteration	51	—	4	—

The average of estimations and the EMSE have obtained. As we expected, with increasing sample size the EMSE decreases. Also, from Table 3, we see that for θ close to 1 the rate of convergence increases and the number of iterations decrease.

6 Conclusions

Two real data sets on the number of proteins and number of residues (Tables 1 and 2) have been analyzed in order to fit the model (3). As it has shown in the Section 3 the *p-values* of *K-S Test* are 0.4271 and 0.963, respectively. It shows that the model (3) fits data well.

Meanwhile, in the Section 4 we obtained some conditions under which the MLE for the parameters of distribution (3) coincides with the solution of the system of likelihood equations. The obtained MLE are the same as some Moment Estimators. The solution of the systems (4) always exists for sufficiently large n when some Regularity Conditions (see, for example, Borovkov, 1998; Lehmann, 1983) are satisfied. The questions relating to the asymptotic properties will be considered later.

In Section 5, using *Accumulation Method* for approximate computation of the MLE, we have proposed a recurrence formula for evaluating the MLE of the parameters θ, b and ρ when $c = 1$. Simulation studies have been illustrated to support our theory. As we saw from Table 3, the *Accumulation Method* works well. We notice that all of the computations have been done using *Statistical Software R (Version 15.2)*.

Acknowledgements

The authors are most grateful to the Editor, Associate Editor and two anonymous referees for their highly constructive suggestions that led to a great improvement to the expression of this paper, in particular for simulation studies. The authors also thankful to Dr. Mahdi Salehi for several discussions on fitting of the distribution and simulation studies. We dedicate this paper to the fond memory of Professor Eduard Danielian (1946-2011).

References

Apweiler, R., Biswas, M., Fleischmann, W., Kanapin, A., Karavidopoulou, Y., Kersey, P., Kriventseva, E., Mittard, V., Mudler, N., Oinn, T., Phan I., and Zdobnov, E. (2000), Proteome analysis: application of InterPro and ClusSTr for the functional classification of proteins in whole genomes. Proceedings of the German Conference on Bioinformatics, 149-158.

Astola, J. and Danielian, E. (2004), Regularly Varying Distributions Generated by Birth-Death Processes. Tampere International Center for Signal Processing (TICSP), Series no. 27, Tampere, Finland, (94 pages; ISBN: 952-15-1270-9; ISSN: 1456-2774).

Astola, J. and Danielian, E. (2006), On regularly varying distributions generated by birth-death processes. Facta Universitatis (NIS): Ser. Elec. Energ., **19**(1), 109-131.

Astola, J. and Danielian, E. (2007), Frequency Distributions in Biomolecular Systems and Growing Networks. TICSP Series no. 31, Tampere, Finland, (251 pages; ISBN: 952-15-1570-8; ISSN: 1456-2774).

Astola, J., Danielian, E., and Arzumanyan, S. (2010), Frequency distributions in bioinformatics, a review. Proceedings of the Yerevan State University: Phys. Math. Sci., **223**(3), 3-22.

Astola, J., Danielian, E., and Gasparian, K. (2007), The maximum likelihood estimators for distributions with moderate growth. Proceedings 6_{th} International Conference on Computer Sciences and Information Technologies (CSIT), September 24-28, Yerevan, Armenia, 91-94.

Bornholdt, S. and Ebel, H. (2001), Wolrd wide web scaling exponent from Simon's 1955 model. Physical Review E, **64**(3-2), 035104(4).

Borovkov, A. A. (1998), Mathematical Statistics. Gordon and Breach Science Publishers, (translated from original Russian edition).

Danielian, E. and Astola, J. (2004), On the steady state of birth-death process with coefficients of moderate growth. Facta Universitatis (NIS): Ser. Elec. Energ., **17**, 405-419.

Danielian, E. and Astola, J. (2007), On generating functions of Pareto and Waring distributions, In: R. Bregovic and A. Gotchev (eds.). Proceedings 2007 International TICSP Workshop on Spectral Methods and Multirate Signal Processing, September 1-2, Moscow, Russia, 235-237.

Gantmacher, F. R. (2000), The Theory of Matrices. American Mathematical Society, vol. 1, (translated from original Russian edition).

Gasparian, K. and Danielian, E. (2006), Maximum likelihood estimators for one class of distributions. Vestnik Russian-Armenian University: Ser. Fiziko-Matemat. Estes. Nauki, **2**, 7-14, (in Russian).

Givens, G. H. and Hoeting, J. A. (2005), Computational Statistics. Wiley and Sons.

Glanzel, W. and Schubert, A. (1995), Predictive aspects of stochastic model for citation processes. Information Processing and Management, **31**(1), 69-80.

Irwin, O. J. (1963), The place of mathematics in the medical and biological statsitics. Journal Royal Statistical Society: A, **126**, 1-45.

Ivchenko, G. I. and Medvedev, Yu. (1990), Mathematical Statistics. Mir Press, Moscow, (translated from original Russian edition).

Kabsch, W. and Sander, C. (1983), Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers, **22**(12), 2577-2637.

Kuznetsov, V. A. (2001), Distributions associated with stochastic processes of gene expression in a single eukaryotic cell. EURASIP Journal on Applied Signal Processing, **4**, 258-296.

Kuznetsov, V. A. (2002), Statistics of the numbers of transcripts and protein sequences encoded in the genome. in Computational and Statistical Methods to Genomics, edited by W. Zhang and I. Shmulevich, Kluwer, Boston, 125-171.

Kuznetsov, V. A. (2003a), Family of skewed distributions associated with the gene expression and proteome evolution. Signal Processing, **33**(4), 889-910.

Kuznetsov, V. A. (2003b), Stochastic model of evolution of conserved protein coding sequences. AIP Conference Proceedings, American Institute of Physics, **665**, 369-380.

Kuznetsov, V. A., Knott, G. D., and Bonner, R. F. (2002a), General Statistics of Stochastic Process of Gene Expression in Eukaryotic Cells. Genetics, Genetics Society of America, **161**, 1321-1332.

Kuznetsov, V. A., Pickalov, V. A., Senko, O. V., and Knott, G. D. (2002b), Analysis of the evolving proteomes: Predictions of the number for protein domains in nature and the number of genes in eukaryotic organisms. Journal of Biological Systems, **10**(4), 381-407.

Lehmann, E. L. (1983), Theory of Point Estimation. Wiley and Sons.

Rizzo, M. L. (2008), Statistical Computing with R. Chapmann & Hall /CRC.

Saaty, T. (1983), Elements of Queuing Theory, Dower Pulications.

Seber, G. A. and Lee, A. J. (2003), Linear Regression Analysis. 2nd edition, Wiley and Sons.

Simon, H. A. (1955), On a class of skew distribution functions. Biometrika, **42**, 425-440.