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

Davood Farbod\textsuperscript{1}, Karen Gasparian\textsuperscript{2}

\textsuperscript{1}Department of Mathematics, Quchan University of Advanced Technologies Engineering, Quchan, Iran.
\textsuperscript{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.


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.
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 notations (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 \( \{\rho(t) : t \geq 0\} \) be a homogeneous Markov Process with continuous time and countable number of states 0, 1, 2, ... . The stationary distribution of the process \( \{\rho(t) : t \geq 0\} \) 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):

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

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

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

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

\[
\alpha \in \Omega = \left\{ \alpha : 0 < \theta < 1, 0 < c < \infty, 0 < b < \infty, 1 < \rho < \infty \right\}, \ b^{\rho} > 1 - c.
\]
The role of the parameter $\theta$ 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 $\rho$ 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 $\alpha$ 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 $E_{\alpha}(\cdot)$, $\text{Var}_{\alpha}(\cdot)$ and $\text{Cov}_{\alpha}(\cdot, \cdot)$, correspondingly, for the expectation, the variance and the covariance with respect to the distribution $P_{\alpha}$. Let also $X^n = (X_1, ..., X_n)$ be a sample corresponding to a random variable $\xi$ with the distribution (3). In the sequel we use the following notations:

$$h_{\gamma,j}(x, \alpha) = \sum_{m=0}^{x-1} (m + b)^{-\gamma}[(m + b)^\rho + c - 1]^{-j}, \quad x, j \in \mathbb{N}, \ \gamma \in \mathbb{R};$$

$$l_{\gamma,j,k}(x, \alpha) = \sum_{m=0}^{x-1} (m + b)^{\gamma} [(m + b)^\rho + c - 1]^{-j} \cdot [\ln(m + b)]^k,$$

$$j, k \in \mathbb{N}, \ \gamma \in \mathbb{R};$$

$$H(x, \alpha) = (c - 1)h_{1,1}(x, \alpha) + (x + b)^{-1};$$

$$\Lambda(x, \alpha) = (c - 1) \ l_{0,1,1}(x, \alpha) + \ln(x + b);$$

$$\overline{X^n} = \frac{1}{n} \sum_{i=1}^{n} X_i; \quad \overline{\psi^n_k(b)} = \frac{1}{n} \sum_{i=1}^{n} (X_i + b)^{-k}, \quad k \in \mathbb{N};$$

$$\overline{f^n(\alpha)} = \frac{1}{n} \sum_{i=1}^{n} f(X_i, \alpha).$$

3 Fitting of the Distribution

As we have said in the Preliminary, 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 Kuznetsov 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 $\xi$ as *doubly-truncated*. Namely, random variable $\xi$ 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 cerevisiae*two (Apweiler et al., 2000) as a real data set in the following Table:

**Table 1.**

<table>
<thead>
<tr>
<th>124</th>
<th>115</th>
<th>83</th>
<th>69</th>
<th>68</th>
<th>66</th>
<th>52</th>
<th>35</th>
<th>34</th>
<th>33</th>
<th>32</th>
<th>28</th>
<th>24</th>
<th>24</th>
<th>21</th>
<th>21</th>
<th>20</th>
<th>20</th>
<th>19</th>
<th>19</th>
<th>18</th>
<th>17</th>
<th>16</th>
<th>14</th>
<th>14</th>
<th>14</th>
</tr>
</thead>
</table>

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

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

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.**

<table>
<thead>
<tr>
<th>85</th>
<th>103</th>
<th>103</th>
<th>112</th>
<th>134</th>
<th>82</th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>98</td>
<td>138</td>
<td>54</td>
<td>125</td>
<td>99</td>
</tr>
</tbody>
</table>

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

$\hat{\theta} = 0.9724; \hat{c} = 16.1974; \hat{b} = 0.102; \hat{\rho} = 1.2465$. 
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 $\alpha$ for a random variable $\xi$ 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:

\[
\begin{align*}
E_\alpha(\xi) &= X^n, \\
E_\alpha[h_{0,1}(\xi, \alpha)] &= h_{0,1}^n(\alpha), \\
E_\alpha[H(\xi, \alpha)] &= H^n(\alpha), \\
E_\alpha[\Lambda(\xi, \alpha)] &= \Lambda^n(\alpha).
\end{align*}
\]

(4)

\[
\begin{align*}
E_\alpha(\xi + b)^{-k} &= \psi_k^n(b), \\
E_\alpha[h_{k,1}(\xi, \alpha)] &= h_{k,1}^n(\alpha), \hspace{1cm} k = 1, 2, \\
E_\alpha[h_{k,2}(\xi, \alpha)] &= h_{k,2}^n(\alpha), \hspace{1cm} k = 0, 1 - \rho, 2 - \rho, \\
E_\alpha[l_{\rho,2,k}(\xi, \alpha)] &= l_{\rho,2,k}^n(\alpha), \hspace{1cm} k = 1, 2, \\
E_\alpha[l_{\rho-1,2,1}(\xi, \alpha)] &= l_{\rho-1,2,1}^n(\alpha),
\end{align*}
\]

**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)),

\[
\begin{align*}
E_\alpha(\xi + b)^{-k} &= \psi_k^n(b), \\
E_\alpha[h_{k,1}(\xi, \alpha)] &= h_{k,1}^n(\alpha), \hspace{1cm} k = 1, 2, \\
E_\alpha[h_{k,2}(\xi, \alpha)] &= h_{k,2}^n(\alpha), \hspace{1cm} k = 0, 1 - \rho, 2 - \rho, \\
E_\alpha[l_{\rho,2,k}(\xi, \alpha)] &= l_{\rho,2,k}^n(\alpha), \hspace{1cm} k = 1, 2, \\
E_\alpha[l_{\rho-1,2,1}(\xi, \alpha)] &= l_{\rho-1,2,1}^n(\alpha),
\end{align*}
\]
then it coincides with the MLE of the parameter $\alpha$.

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 $\alpha \in \Omega$:

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

where $\zeta(\rho, b) = \sum_{n=0}^{\infty} (n + b)^{-\rho} < \infty$, $\zeta(\rho) = \sum_{n=1}^{\infty} n^{-\rho} < \infty$, $\rho > 1$, is a Riemann’s Zeta Function, and $\mathbb{I}_{(-,\infty)}(\cdot)$ is an indicator function;
(ii) if $\rho > k + 1$, $k \in \mathbb{N}$, then
\[
E(\xi^k) \leq \left\{ I_{(0,1]}(c) + I_{(1,\infty)}(c) \cdot \exp[(c - 1)\zeta_k(\rho, b)] \right\} \cdot \zeta_k(\rho, b),
\]
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) $E_\alpha[h_{i,1}(\xi, \alpha)]^2 < \infty$, $i = 0, 1, 2$;
(ii) $E_\alpha[h_{i,2}(\xi, \alpha)] < \infty$, $i = 0, 1 - \rho, 2 - \rho$;
(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
\[
1 < g(\alpha) < 1 + \sum_{n=1}^{\infty} \frac{\theta^n}{(n+b)^\rho} \cdot \exp \left[ (c - 1) \cdot \sum_{m=0}^{n-1} (m+b)^{-\rho} \right]
\leq \left\{ \begin{array}{ll}
1 + \zeta(\rho, b), & \text{if } 0 < c \leq 1, \\
1 + \zeta(\rho, b) \cdot \exp[(c - 1)\zeta(\rho, b)], & \text{if } c > 1;
\end{array} \right.
\]

(ii) on the other hand, if $\rho > k + 1$, $k \in \mathbb{N}$, then
\[
E_\alpha(\xi^k) \leq [g(\alpha)]^{-1} \cdot \sum_{n=1}^{\infty} \frac{n^k}{(n+b)^\rho} \cdot \exp \left[ (c - 1) \sum_{m=0}^{n-1} (m+b)^{-\rho} \right]
\leq \left\{ \begin{array}{ll}
\zeta_k(\rho, b), & \text{if } 0 < c \leq 1, \\
\zeta_k(\rho, b) \cdot \exp[(c - 1)\zeta(\rho, b)], & \text{if } c > 1.
\end{array} \right.
\]

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

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

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

(iii) also, we have

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

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

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

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

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

(v) and finally we have

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

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

Thus the proof of Lemma 4.1 is completed. \( \Box \)

**Proof of Theorem 4.1.** For a random sample \( X^n = (X_1, ..., X_n) \sim P_\alpha \) from the distribution (3) the likelihood function will be

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

such that for the logarithm of likelihood function we obtain

\[
L_\alpha(X^n) = \ln f_\alpha(X^n)
\]

\[
= -n \ln g(\alpha) + n \cdot \bar{X}^n \ln \theta - \rho \sum_{i=1}^{n} \ln(X_i + b)
\]

\[
+ \sum_{i=1}^{n} \sum_{m=0}^{X_i-1} \ln(1 + \frac{c - 1}{(m+b)^\rho}).
\]
The necessary conditions for existence of the MLE for the parameter $\alpha$ 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:
\[
\frac{\partial L_\alpha(X^n)}{\partial \theta} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \theta} + \frac{n \cdot \bar{X}^n}{\theta},
\]
where (see (3))
\[
\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \theta} = \frac{1}{\theta} \cdot E_\alpha(\xi),
\]
such that the condition $\frac{\partial L_\alpha(X^n)}{\partial \theta} = 0$ implies $E_\alpha(\xi) = \bar{X}^n$.

Also, we have
\[
\frac{\partial L_\alpha(X^n)}{\partial b} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial b} - n\rho \cdot \psi_n(b) - n\rho(c-1)\overline{h_{1,1}^n(\alpha)},
\]
where
\[
\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial b} = -\rho \cdot E_\alpha(\xi + b)^{-1} - \rho(c-1)E_\alpha[h_{1,1}(\xi, \alpha)] = -\rho \cdot E_\alpha[H(\xi, \alpha)],
\]
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
\[
\frac{\partial L_\alpha(X^n)}{\partial c} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial c} + n \cdot \overline{h_{0,1}^n(\alpha)},
\]
where
\[
\frac{1}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial c} = E_\alpha[h_{0,1}(\xi, \alpha)],
\]
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
\[
\frac{\partial L_\alpha(X^n)}{\partial \rho} = -\frac{n}{g(\alpha)} \cdot \frac{\partial g(\alpha)}{\partial \rho} - n \cdot \ln(X^n + b) - n(c - 1) \int_{0.1.1}^1 (\alpha), \quad \text{as } bL_n = n \cdot cM_n.
\]

Lemma 4.3. If the conditions of Lemma 4.2 are valid then the matrix \( bL_n \) is negative definite.

\[
\text{Theorem 4.1 is proved. } \Box
\]

Let us now prove that the solution \( \hat{\alpha} = \hat{\alpha}^n = (\hat{\alpha}^n)_1^n \) of the system (4) (if it exists) is the MLE of the parameter \( \alpha \). In order to do that, it is enough to show that the matrix
\[
\hat{L}^n = \left( \hat{L}^n_{ij} \right)_{i,j=1}^4
\]
with \( \hat{L}^n_{ij} = L^n_{ij}(\hat{\alpha}) \), \( L^n_{ij}(\hat{\alpha}) = \frac{\partial^2 L_\alpha(X^n)}{\partial \alpha_i \partial \alpha_j} |_{\alpha=\hat{\alpha}} \), is negative definite.

**Lemma 4.2.** If for \( \rho > 3 \) the solution \( \hat{\alpha} \) of the system (4) (if it exists) satisfies the Conditions (A), then the elements of the matrix \( \hat{L}^n \) are:
\[
\begin{align*}
\hat{L}^n_{11} &= -\frac{n}{\rho} \cdot \text{Var}_\alpha(\xi), & \hat{L}^n_{12} = \hat{L}^n_{21} &= -\frac{n}{\rho} \cdot \text{Cov}_\alpha[\xi, h_{0.1}(\xi, \hat{\alpha})], \\
\hat{L}^n_{13} &= \hat{L}^n_{31} &= \frac{n\rho}{\rho} \cdot \text{Cov}_\alpha[\xi, H(\xi, \hat{\alpha})], & \hat{L}^n_{14} = \hat{L}^n_{41} &= \frac{n}{\rho} \cdot \text{Cov}_\alpha[\xi, \Lambda(\xi, \hat{\alpha})], \\
\hat{L}^n_{22} &= -n \cdot \text{Var}_\alpha[h_{0.1}(\xi, \hat{\alpha})], & \hat{L}^n_{23} &= \hat{L}^n_{32} = n\hat{\rho} \cdot \text{Cov}_\alpha[h_{0.1}(\xi, \hat{\alpha}), H(\xi, \hat{\alpha})], \\
\hat{L}^n_{24} &= \hat{L}^n_{42} &= n \cdot \text{Cov}_\alpha[h_{0.1}(\xi, \hat{\alpha}), \Lambda(\xi, \hat{\alpha})], & \hat{L}^n_{33} &= -n\hat{\rho}^2 \cdot \text{Var}_\alpha[H(\xi, \hat{\alpha})], \\
\hat{L}^n_{34} &= \hat{L}^n_{43} &= -n\hat{\rho} \cdot \text{Cov}_\alpha[H(\xi, \hat{\alpha}), \Lambda(\xi, \hat{\alpha})], & \hat{L}^n_{44} &= -n \cdot \text{Var}_\alpha[\Lambda(\xi, \hat{\alpha})],
\end{align*}
\]
such that the matrix \( \hat{L}^n \) can be represented as \( \hat{L}^n = n \cdot \hat{M}^n \).

**Lemma 4.3.** If the conditions of Lemma 4.2 are valid then the matrix \( \hat{L}^n \) is negative definite.
On the Maximum Likelihood Estimators for some ... 223

Proof of Lemma 4.2.

1. From the representation (5) we obtain

\[ L^n_{11}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \theta^2} = n \left[ \frac{1}{g^2} \cdot \left( \frac{\partial g}{\partial \theta} \right)^2 - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \theta^2} - \frac{X^n}{\theta^2} \right], \]

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

where by (6) it follows that

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

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

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

\[ L^n_{11}(\alpha) = \frac{n}{\theta^2} \cdot \left\{ (E_\alpha(\xi) - X^n) - Var_\alpha(\xi) \right\}, \]

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

\[ L^n_{13}(\alpha) = L^n_{31}(\alpha) = \frac{n\rho}{\theta} \cdot Cov_\alpha[\xi, H(\xi, \alpha)], \]

\[ L^n_{14}(\alpha) = L^n_{41}(\alpha) = \frac{n}{\theta} \cdot Cov_\alpha[\xi, \Lambda(\xi, \alpha)]. \]

2. It follows from the representation (9) that

\[ L^n_{22}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \alpha^2} = n \left[ \frac{1}{g^2} \cdot \left( \frac{\partial g}{\partial \alpha} \right)^2 - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \alpha^2} \right] - n \cdot \hat{h}_{0,2}^n(\alpha), \]

\[ L^n_{23}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \alpha \partial \theta} = n \left[ \frac{1}{g^2} \cdot \frac{\partial g}{\partial \alpha} \cdot \frac{\partial g}{\partial \theta} - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \alpha \partial \theta} \right] - n\rho \cdot \hat{h}_{1-\rho,2}^n(\alpha), \]

\[ L^n_{24}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \alpha \partial \rho} = n \left[ \frac{1}{g^2} \cdot \frac{\partial g}{\partial \alpha} \cdot \frac{\partial g}{\partial \rho} - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \alpha \partial \rho} \right] - n \cdot \hat{h}_{\rho,2}^n(\alpha). \]

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

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

\[ \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \alpha} = -\rho \left\{ E_\alpha[h_{0,1}(\xi, \alpha) \cdot H(\xi, \alpha)] + E_\alpha[h_{1-\rho,2}(\xi, \alpha)] \right\}, \]

\[ \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \alpha} = -\left\{ E_\alpha[h_{0,1}(\xi) \cdot \Lambda(\xi, \alpha)] + E_\alpha[l_{\rho,2,1}(\xi, \alpha)] \right\}. \]
Then, according to (8), (10) and (12) we have

\[ L_{n22}(\alpha) = n \left\{ (E[h_{0,2}(\xi, \alpha)] - h_{0,2}(\alpha)) - Var_{\alpha}[h_{0,1}(\xi, \alpha)] \right\}, \]

\[ L_{n23}(\alpha) = L_{n32}(\alpha) = n\rho \left\{ (E[h_{1,1}(\xi, \alpha)] - h_{1,1}(\alpha)) + Cov_{\alpha}[h_{0,1}(\xi, \alpha), H(\xi, \alpha)] \right\}, \]

\[ L_{n24}(\alpha) = L_{n42}(\alpha) = n \left\{ (E[l_{\rho,1}(\xi, \alpha)] - l_{\rho,1}(\alpha)) + Cov_{\alpha}[h_{0,1}(\xi, \alpha), \Lambda(\xi, \alpha)] \right\}. \]

3. From (8) it follows that

\[ L_{n33}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \rho^2} \]

\[ = n \left[ \frac{1}{g} \cdot \left( \frac{\partial g}{\partial \rho} \right)^2 - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \sigma^2} \right] + n\rho \left[ (c - 1) \cdot h_{2,1}(\alpha) \right. \]

\[ + \left. \psi_2(b) \right] + n\rho^2(c - 1) \cdot h_{n,0,1}(\alpha), \]

\[ L_{n34}(\alpha) = \frac{\partial^2 L_\alpha(X^n)}{\partial \sigma \partial \rho} \]

\[ = n \left[ \frac{1}{g} \cdot \frac{\partial g}{\partial \sigma} \cdot \frac{\partial g}{\partial \rho} - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \sigma \partial \rho} \right] - n \cdot \frac{\partial H_n(\alpha)}{\partial \rho} + n\rho(c - 1) \cdot l_{\rho,1,2,1}(\alpha). \]

According to (8) we obtain

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

\[ + \rho^2(c - 1) \cdot E_{\alpha}[h_{2,1}(\xi, \alpha)], \]

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

\[ + \rho(c - 1)E_{\alpha}[l_{\rho,1,2,1}(\xi, \alpha)], \]

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

\[ L_n^{33}(\alpha) = -n\rho^2 \cdot \text{Var}_{\alpha}[H(\xi, \alpha)] + n\rho^2(c - 1) \left\{ \frac{h_{2\rho,2}(\alpha)}{\rho} - E_\alpha[h_{2\rho,2}(\xi, \alpha)] \right\} \\
+ n\rho \left\{ \psi_{2b}(b) - E_\alpha(\xi + b)^{-2} \right\} \\
+ n\rho(c - 1) \left\{ \frac{h_{2,1}(\alpha)}{\rho} - E_\alpha[h_{2,1}(\xi, \alpha)] \right\}, \]

\[ L_n^{34}(\alpha) = -n\rho \text{Cov}_{\alpha}[H(\xi, \alpha), \Lambda(\xi, \alpha)] + n \left\{ E_\alpha[H(\xi, \alpha)] - \bar{H}_n(\alpha) \right\} \\
+ n\rho(c - 1) \left\{ E_\alpha[l_{\rho-1,2,1}(\xi, \alpha)] - \frac{m_{\rho-1,2,1}(\alpha)}{\rho} \right\}. \]

4. Finally from (11) we have

\[ L_n^{44}(\alpha) = \frac{\partial^2 E_\alpha(\hat{X}_n)}{\partial \rho^2} = n \left[ \frac{1}{g^2} \cdot \left( \frac{\partial g}{\partial \rho} \right)^2 - \frac{1}{g} \cdot \frac{\partial^2 g}{\partial \rho^2} \right] + n(c - 1) \cdot \frac{m_{\rho,2,2}(\alpha)}{\rho}. \]

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

\[ L_n^{44}(\alpha) = -n \cdot \text{Var}_{\alpha}[\Lambda(\xi, \alpha)] - n(c - 1) \left\{ E_\alpha[l_{\rho,2,2}(\xi, \alpha)] - \frac{m_{\rho,2,2}(\alpha)}{\rho} \right\}. \]

Assuming now that the solution \( \hat{\alpha} = (\hat{\theta}, \hat{c}, \hat{b}, \hat{\rho}) \) 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 \( \hat{L}^n \) are the following conditions (see, for example, Gantmacher, 2000)

\[ \hat{L}_{11}^n < 0, \quad \Delta_2 \equiv \det \left( \hat{L}_{ij}^n \right)_{i,j=1}^2 > 0, \]

\[ \Delta_3 \equiv \det \left( \hat{L}_{ij}^n \right)_{i,j=1}^3 < 0, \quad \Delta_4 \equiv \det \hat{L}^n > 0. \]

By Lemma 4.2 we have

\[ \hat{L}_{11}^n = -\frac{n}{\theta^2} \cdot \text{Var}_{\alpha}(\xi) < 0 \]
in view of $\xi \neq \text{constant}$, $P - \text{a.s.}$ (a.s. means almost surely).

Then,
\[
\Delta_2 = \tilde{L}_1 \cdot \tilde{L}_2 - (\tilde{L}_1)^2
\]
\[
= (\frac{a}{b})^2 \cdot \left\{ \text{Var}(\xi) \cdot \text{Var}[h_{0,1}(\xi, \hat{\alpha})] - \text{Cov}_{\hat{\alpha}}^2[\xi, h_{0,1}(\xi, \hat{\alpha})] \right\}.
\]

Hence by Cauchy-Schwarz inequality it follows that $\Delta_2 \geq 0$. But, because
\[
h_{0,1}(\xi, \alpha) \neq b(\alpha) + c(\alpha)\xi \quad P - \text{a.s.},
\]
for some $b(\alpha)$ and $c(\alpha)$, then $\Delta_2 > 0$.

On the other hand, we have
\[
\Delta_3 = -\frac{n^3}{\theta^2} \cdot \det \left( \hat{D}_3 \right), \quad \hat{D}_3 = (\hat{q}_{ij})^3_{i,j=1},
\]
where
\[
\hat{q}_{ij} = q_{ij}(\hat{\alpha}) = \text{Cov}_{\hat{\alpha}}(\xi_i, \xi_j), \quad \xi_1 = \xi, \quad \xi_2 = h_{0,1}(\xi, \hat{\alpha}), \quad \xi_3 = H(\xi, \hat{\alpha}).
\]

Hence, because $\det \left( \hat{D}_3 \right) \geq 0$ ($\hat{D}_3$ 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
\[
\sum_{i=1}^{3} c_i(\alpha)\xi_i = b(\alpha) \quad P - \text{a.s.},
\]
then $\det \left( \hat{D}_3 \right) \neq 0 \quad P - \text{a.s.}$.

Finally, we have
\[
\Delta_4 = \frac{n^4}{\theta^2} \cdot \det \left( \hat{D}_4 \right), \quad \hat{D}_4 = (\hat{q}_{ij})^4_{i,j=1},
\]
where $\hat{q}_{ij} = \text{Cov}(\xi_i, \xi_j), \quad \xi_1, \xi_2, \xi_3$ are as before and $\xi_4 = \Lambda(\xi, \hat{\alpha})$. So, as above, it can be showed that $\det \left( \hat{D}_4 \right) > 0$ 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 \( \hat{L}^n \) as (see Lemma 4.2)

\[
\hat{L}^n = n \cdot \hat{M}^n \\
= n \cdot \left\{ \hat{M}^n(0) - \delta_2^n \cdot (e_{22}) + \rho^n \cdot \delta_3^n \cdot (e_{33}) + (\hat{c}_n - 1) \cdot \delta_4^n \cdot (e_{44}) \\
- \hat{\rho}_n \cdot \delta_5^n \cdot (e_{23}) - \delta_6^n \cdot (e_{24}) - \delta_7^n \cdot (e_{34}) \right\},
\]

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

\[
\delta_2^n = h_{0,2}(\alpha) - E_{\hat{\alpha}}[h_{0,2}(\xi, \hat{\alpha})],
\]
\[
\delta_3^n = \left\{ \psi_2(\hat{\beta}) - E_{\hat{\alpha}}(\xi + \hat{\beta})^{-2} \right\} + (\hat{c}_n - 1) \left\{ h_{2,1}^n(\alpha) - E_{\hat{\alpha}}[h_{2,1}(\xi, \hat{\alpha})] \right\} \\
+ \hat{\rho}_n(\hat{c}_n - 1) \left\{ h_{2-\rho,2}^n(\alpha) - E_{\hat{\alpha}}[h_{2-\rho,2}(\xi, \hat{\alpha})] \right\},
\]
\[
\delta_4^n = l_{\rho,2,2}(\alpha) - E_{\hat{\alpha}}[l_{\rho,2,2}(\xi, \hat{\alpha})],
\]
\[
\delta_5^n = h_{1-\rho,2}(\alpha) - E_{\hat{\alpha}}[h_{1-\rho,2}(\xi, \hat{\alpha})],
\]
\[
\delta_6^n = l_{\rho,1,2}(\alpha) - E_{\hat{\alpha}}[l_{\rho,1,2}(\xi, \hat{\alpha})],
\]
\[
\delta_7^n = \left\{ \psi_1(\hat{\beta}) - E_{\hat{\alpha}}(\xi + \hat{\beta})^{-1} \right\} + (\hat{c}_n - 1) \left\{ h_{1,1}^n(\alpha) - E_{\hat{\alpha}}[h_{1,1}(\xi, \hat{\alpha})] \right\} \\
+ \hat{\rho}_n(\hat{c}_n - 1) \left\{ l_{\rho-1,2,1}^n(\alpha) - E_{\hat{\alpha}}[l_{\rho-1,2,1}(\xi, \hat{\alpha})] \right\}.
\]

For \( n \) sufficiently large by the Strong Law of Large Numbers, a random variable \( \delta_i^n \), \( i = 2 - 7 \), becomes a.s. small enough such that all principal minors of the matrix \( \hat{M}^n \) preserves the respective signs of the principal minors of the negative definite matrix \( \hat{M}^n(0) \). Thus the matrix \( \hat{M}^n \) (and also \( \hat{L}^n \)) 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 $\alpha$) 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 $U_j(\alpha) = \frac{\partial L_n(x_n)}{\partial \alpha_j}$, where $x_n = (x_1, ..., x_n)$ is a realization of a sample $X_n \sim P_\alpha$.

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

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

where $\tilde{\alpha}$ is some intermediate point between $\alpha(0)$ and $\tilde{\alpha}$.

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

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

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

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

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 \((z + 1)_{\text{th}}\) approximation can be found recurrently by the formula

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

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

\[
\begin{align*}
\theta(z + 1) &= \theta(z) + \frac{\Delta_1(\alpha(z))}{n \cdot \det I(\alpha(z))}, \\
b(z + 1) &= b(z) + \frac{\Delta_3(\alpha(z))}{n \cdot \det I(\alpha(z))}, \\
\rho(z + 1) &= \rho(z) + \frac{\Delta_4(\alpha(z))}{n \cdot \det I(\alpha(z))}.
\end{align*}
\] (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 \(\alpha_j(z), \ j = 1, 3, 4, \ z = 0, 1, 2, \ldots;\)

3. If \(|\alpha_j(z + 1) - \alpha_j(z)| < \epsilon\) (\(\epsilon\) is some small positive constant), then \(\alpha_j(z + 1) = \hat{\alpha}\) 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 \(\xi\) as doubly-truncated. Here, random variable \(\xi\) 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 \(\epsilon = 0.0005\).
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>
<thead>
<tr>
<th>Case</th>
<th>N=50</th>
<th></th>
<th>N=100</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>EMSE</td>
<td>Mean</td>
<td>EMSE</td>
</tr>
<tr>
<td>(\theta = 0.9)</td>
<td>0.87049</td>
<td>0.00087</td>
<td>0.87112</td>
<td>0.00083</td>
</tr>
<tr>
<td>(b = 2)</td>
<td>2.26768</td>
<td>0.07165</td>
<td>2.07161</td>
<td>0.00513</td>
</tr>
<tr>
<td>(\rho = 3.6)</td>
<td>3.19627</td>
<td>0.16299</td>
<td>3.42955</td>
<td>0.02905</td>
</tr>
<tr>
<td>Iteration</td>
<td>104</td>
<td>—</td>
<td>82</td>
<td>—</td>
</tr>
<tr>
<td>(\theta = 0.99)</td>
<td>0.94064</td>
<td>0.00244</td>
<td>0.99911</td>
<td>(8.3 \times 10^{-5})</td>
</tr>
<tr>
<td>(b = 3)</td>
<td>3.11567</td>
<td>0.01338</td>
<td>2.99373</td>
<td>(3.9 \times 10^{-5})</td>
</tr>
<tr>
<td>(\rho = 4.5)</td>
<td>4.56459</td>
<td>0.00417</td>
<td>4.70897</td>
<td>(4.3 \times 10^{-2})</td>
</tr>
<tr>
<td>Iteration</td>
<td>51</td>
<td>—</td>
<td>4</td>
<td>—</td>
</tr>
</tbody>
</table>

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 \(\theta\) 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.
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 $\theta$, $b$ and $\rho$ 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


