On Estimating the Parameters of an Extended Form of Logarithmic Series Distribution

C. Satheesh Kumar¹,a, A. Riyaz a

aDepartment of Statistics, University of Kerala

Abstract

We consider an extended version of a logarithmic series distribution and discuss the estimation of its parameters by the method of moments and the method of maximum likelihood. Test procedures are suggested to test the significance of the additional parameter of this distribution and all procedures are illustrated with the help of real life data sets. In addition, a simulation study is conducted to assess the performance of the estimators.

Keywords: Generalized likelihood ratio test, logarithmic series distribution, maximum likelihood estimation, probability generating function, Rao's score test.

1. Introduction

The logarithmic series distribution (LSD), introduced by Fisher et al. (1943), has found application in areas such as biology, ecology, economics, operations research and marine sciences. Fisher et al. (1943) obtained the LSD as the limit of a zero-truncated negative binomial distribution in connection with an investigation of the frequency distribution of number of species of animals obtained from random samples. For a detailed account of the LSD see the Chapter 7 of Johnson et al. (2005). Various generalized versions of the LSD have been proposed in the literature. For example see Tripathi and Gupta (1985, 1988), Ong (2000), Khang and Ong (2007) and Kumar and Riyaz (2013).

We obtain an extended form of the LSD called “the extended logarithmic series distribution (ELSD)” and study some of its important aspects. In Section 2 we present the definition of the ELSD and derive its properties. In Section 3, we discuss the estimation of the parameters of the ELSD by the method of moments and the method of maximum likelihood, and illustrate its usefulness by fitting the model to certain real life data sets. In Section 4 we consider the generalized likelihood ratio test and Rao’s efficient score test to test the significance of the additional parameter of the ELSD. In Section 5 we conduct a simulation study to compare the performance of the estimators obtained by both the methods of estimation.

2. The Extended Logarithmic Series Distribution and Its Properties

Here we define the extended logarithmic series distribution as follows.

Definition 1. Let $X = (X_1, X_2)$ follows bivariate logarithmic series distribution (cf. Kocherlakota and Kocherlakota, 1992, p.192) with probability generating function (pgf)

$$G(t) = \frac{\ln(1 - \alpha t_1 - \alpha t_2 - \alpha_3 t_1 t_2)}{\ln(1 - \alpha)}$$ (2.1)

¹Corresponding author: Department of Statistics, University of Kerala, Trivandrum-695 581, India.
E-mail: drcsatheeshkumar@gmail.com
in which \(\alpha_1, \alpha_2 > 0\) and \(\alpha_3 \geq 0\) such that \(\alpha = \alpha_1 + \alpha_2 + \alpha_3 < 1\). Then the distribution of \(Y = X_1 + X_2\) is known as “the extended logarithmic series distribution (ELSD)” with following pgf, in which \(\theta_1 = \alpha_1 + \alpha_2\) and \(\theta_2 = \alpha_3\).

\[
H(t) = \frac{\ln \left(1 - \theta_1 t - \theta_2 t^2\right)}{\ln(1 - \theta_1 - \theta_2)}. \tag{2.2}
\]

Clearly, when \(\theta_2 = 0\), (2.2) reduces to the pgf of the LSD due to Fisher et al. (1943). The pgf of ELSD can also be written as

\[
H(t) = \left(\frac{\theta_1 t + \theta_2 t^2}{\theta_1 + \theta_2}\right) \sum_{k=0}^{\infty} \frac{a(a+1)(a+2)\cdots(a+r-1)b(b+1)(b+2)\cdots(b+r-1) z^r}{c(c+1)(c+2)\cdots(c+r-1)} \frac{t^r}{r!}, \tag{2.3}
\]

where

\[
\sum_{k=0}^{\infty} a(a+1)(a+2)\cdots(a+r-1)b(b+1)(b+2)\cdots(b+r-1) \frac{z^r}{c(c+1)(c+2)\cdots(c+r-1)} \frac{t^r}{r!}
\]

is the Gauss hypergeometric function. If we replace \(t\), by \(e^{it}\) and \((1 + t)\) in (2.2) we get the corresponding characteristic function \(\psi(t)\) and the factorial moment generating function \(F(t)\) of the ELSD as given below, in which \(C = [-\ln(1 - \theta_1 - \theta_2)]^{-1}\)

\[
\psi(t) = C \left[ -\ln \left(1 - \theta_1 e^{it} - \theta_2 e^{2it}\right) \right]. \tag{2.4}
\]

and

\[
F(t) = C \left[ -\ln \left(1 - \theta_1 (1 + t) - \theta_2 (1 + t)^2\right) \right]. \tag{2.5}
\]

Now, we obtain the following result in the light of the series representation

\[
\sum_{x=0}^{\infty} \sum_{r=0}^{\infty} A(r, x) = \sum_{x=0}^{\infty} \sum_{r=0}^{x} A(r, x - r) \tag{2.6}
\]

and

\[
\sum_{x=0}^{\infty} \sum_{r=0}^{\infty} B(r, x) = \sum_{x=0}^{\infty} \sum_{r=0}^{\lfloor x \rfloor} B(r, x - r). \tag{2.7}
\]

**Result 1.** For \(x = 1, 2, \ldots\) the probability mass function (pmf) \(q_x = P(Y = x)\) of the ELSD with pgf (2.3) is the following, in which \([a]\) denote the integer part of a for any \(a > 0\).

\[
q_x = C \sum_{r=0}^{\lfloor x \rfloor} \frac{(x - r - 1)! \theta_1^{x-2r} \theta_2^r}{(x-2r)! r!}. \tag{2.8}
\]

**Proof:** From (2.2) we have the following

\[
H(t) = \sum_{x=0}^{\infty} q_x t^x \tag{2.9}
\]

\[
= C \left[ -\ln \left(1 - \theta_1 t - \theta_2 t^2\right) \right]. \tag{2.10}
\]
On expanding the logarithmic function in (2.10), we get

\[ H(t) = C \sum_{x=1}^{\infty} \frac{(\theta_1 t + \theta_2 t^2)^x}{x} \]

(2.11)

\[ = C \sum_{x=0}^{\infty} \frac{(\theta_1 t + \theta_2 t^2)^{x+1}}{x + 1}. \]

(2.12)

Now, on applying binomial theorem, we obtain the following from (2.12).

\[ H(t) = C \sum_{x=0}^{\infty} \sum_{r=0}^{x+1} \frac{x!}{r!(x-r+1)!} \theta_1^{r+1} \theta_2^{x+1} t^{x+r+1} \]

(2.13)

\[ = C \sum_{x=0}^{\infty} \sum_{r=0}^{x+1} \frac{x!}{r!(x-r+1)!} \theta_1^{r+1} \theta_2^{x+r+1} + C \sum_{x=0}^{\infty} \frac{(\theta_2 t^2)^{x+1}}{x + 1}. \]

(2.14)

By applying (2.6) in (2.14) to get

\[ H(t) = C \sum_{x=0}^{\infty} \sum_{r=0}^{x+1} \frac{(x+r)!}{r!(x-r+1)!} \theta_1^{r+1} \theta_2^{x+1} t^{r+1} + C \sum_{x=0}^{\infty} \frac{(\theta_2 t^2)^{x+1}}{x + 1} \]

(2.15)

which implies the following, in the light of (2.7)

\[ H(t) = C \sum_{x=0}^{\infty} \sum_{r=0}^{x+1} \frac{(x-r)!}{r!(x-2r+1)!} \theta_1^{r+2} \theta_2^{x+r+1} + C \sum_{x=0}^{\infty} \frac{(\theta_2 t^2)^{x+1}}{x + 1}. \]

(2.16)

On equating the coefficient of \( t^x \) on the right hand side expressions of (2.9) and (2.16) we get (2.8).

The mean and variance of the ELSD are obtained in the following result.

**Result 2.** The mean and variance of the ELSD are the following in which \( \delta = (1 - \theta_1 - \theta_2)^{-1} \) and \( \lambda = \theta_1 + 2\theta_2 \),

\[ E(Y) = C\delta\lambda \]

and

\[ \text{Var}(Y) = C\delta \left[ (\theta_1 + 4\theta_2) + (1 - C)\delta \lambda^2 \right]. \]

Proof is straightforward and omitted.

Now we obtain an expression for \( r \)th raw moment of the ELSD which can be used for the computation of higher order moments of distribution.

**Result 3.** For \( r \geq 1 \), the \( r \)th raw moment \( \mu_r \) of the ELSD with characteristic function (2.4) is

\[ \mu_r = CL(\theta_1, \theta_2; r) + C^2 \theta_2 \phi(\theta_2, 1 - r, 1) \]

(2.17)
From (2.4) we have

\[ L(\theta_1, \theta_2; r) = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (j + 2k + 1)! (j + k)! \frac{\theta_1^{j+1} \theta_2^k}{j + 1! k!} \]  

(2.18)

can be computed for particular values of \( \theta_1, \theta_2, \) and \( r, \) using mathematical software (such as MATHEMATICA and MATHCAD), since the series converges for \( \theta_1 < 1 \) and \( \theta_2 < 1, \) and

\[ \phi(\theta_2, 1 - r, 1) = \sum_{j=0}^{\infty} \theta_2^j (1 + j)^{-r} \]  

(2.19)

is the Lerch function for any real \( \theta_2 \) and \( r \geq 1 \) (Johnson et al., 2005, p.20).

**Proof:** From (2.4) we have

\[ \phi(t) = \sum_{j=1}^{\infty} \frac{(it)^j}{r!} = C \left[ -\ln \left( 1 - \theta_1 e^r - \theta_2 e^{2r} \right) \right]. \]  

(2.20)

On expanding the logarithmic function in (2.21), we get

\[ \phi(t) = C \sum_{j=0}^{\infty} \frac{\left( \theta_1 e^r + \theta_2 e^{2r} \right)^{j+1}}{j + 1}. \]  

(2.22)

Now, on applying binomial theorem, we obtain the following from (2.22)

\[ \phi(t) = C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{j^1}{k!(j-k+1)!} \theta_1^{j+k+1} \theta_2^{k} e^{(j+k+1)}}{j + 1 + 1}. \]  

(2.23)

By applying (2.6) in (2.24) we get

\[ \phi(t) = C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(j+k)!}{k!(j-k+1)!} \theta_1^{j+k+1} \theta_2^{k} e^{(j+k+1)}}{j + 1 + 1 + 1} \]  

(2.25)

which implies the following, in the light of (2.7).

\[ \phi(t) = C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(j-k)!}{k!(j-2k+1)!} \theta_1^{j-2k+1} \theta_2^{k} e^{(j+1)}}{j + 1 + 1} \]  

(2.26)

\[ \phi(t) = C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \frac{(j-k)!}{k!(j-2k+1)!} \theta_1^{j-2k+1} \theta_2^{k} e^{(j+1)}}{j + 1 + 1} \frac{(2i)^j}{j!} \frac{r}{r} + C \sum_{j=0}^{\infty} \frac{\theta_2^{j+1}}{j + 1 + 1} \sum_{r=0}^{\infty} (j+1)^{(2i)} r! \]  

(2.27)
On estimating the parameters of an extended form of logarithmic series distribution

On equating the coefficient of \((it)^r/r!\) the right hand side expressions of (2.20) and (2.27) we get

\[
\mu_r = C \sum_{j=0}^{\infty} \sum_{k=0}^{j} \frac{(j-k)!}{k!(j-2k+1)!} \theta_1^{j-2k+1} \theta_2^k (j+1)^r + C \sum_{j=0}^{\infty} \frac{\theta_1^{j+1}}{j+1} (j+1)^{2r}. \tag{2.28}
\]

Now apply (2.7) in (2.28) to get (2.17).

3. Estimation

In this section we discuss the estimation of the parameters of the ELSD by the method of moments and the method of maximum likelihood and illustrated the procedure using real life data sets.

3.1. Method of moments

Here the moment estimators \(\hat{\theta}_1\) and \(\hat{\theta}_2\) of the parameters \(\theta_1\) and \(\theta_2\) of the ELSD are obtained by solving the following system of equations. These equations are developed by equating the first two raw moment of the ELSD to the corresponding sample raw moments \(\tau_1\) and \(\tau_2\).

\[
C(1 - \theta_1 - \theta_2)^{-1}(\theta_1 + 2\theta_2) = \tau_1, \tag{3.1}
\]

\[
C(1 - \theta_1 - \theta_2)^{-1} \left[ \theta_1 + 4\theta_2 + (1 - \theta_1 - \theta_2)^{-1}(\theta_1 + 2\theta_2)^2 \right] = \tau_2. \tag{3.2}
\]

3.2. Method of maximum likelihood

Let \(a(x)\) be the observed frequency of \(x\) events and let \(y\) be the highest value of \(x\) observed. Then the likelihood function of the sample is

\[
L = \prod_{x=1}^{y} q_x^{a(x)}, \tag{3.3}
\]

where \(q_x\) is the pmf of the ELSD as given in (2.8). Now taking the logarithm on both sides of (3.3), we have

\[
\log L = \sum_{x=1}^{y} a(x) \log(q_x) \tag{3.4}
\]

\[
= \sum_{x=1}^{y} a(x) [\log C + \log \Phi(x; \theta_1; \theta_2)] \tag{3.5}
\]

where

\[
\Phi(x; \theta_1; \theta_2) = \sum_{r=0}^{\lfloor x \rfloor} \frac{(x-r)!}{(x-2r)!} \frac{\theta_1^{x-2r} \theta_2^r}{r!}. \tag{3.6}
\]

Let \(\hat{\theta}_1\) and \(\hat{\theta}_2\) denote the maximum likelihood estimators of the parameter \(\theta_1\) and \(\theta_2\) respectively of the ELSD. On differentiating (3.5) partially with respect to the parameters \(\theta_1\) and \(\theta_2\) respectively and equating to zero, we get the following likelihood equations.

\[
\sum_{x=1}^{y} a(x) \left[ \frac{-C}{(1 - \theta_1 - \theta_2)} + \Phi^{-1}(x; \theta_1; \theta_2) \sum_{r=0}^{\lfloor x \rfloor} \frac{(x-r)!}{(x-2r)!} \frac{\theta_1^{x-2r} \theta_2^r}{r!} \right] = 0 \tag{3.7}
\]
Table 1: Observed frequencies and computed values of expected frequencies of the LSD and the ELSD by the method of moments and the method of maximum likelihood for the first data set.

<table>
<thead>
<tr>
<th>No. of mites per leaf</th>
<th>Leaves observed</th>
<th>LSD</th>
<th>ELSD method of moments</th>
<th>ELSD method of maximum likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>38</td>
<td>52.939</td>
<td>40.72</td>
<td>40.40</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>15.617</td>
<td>16.64</td>
<td>16.80</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>6.143</td>
<td>8.56</td>
<td>8.56</td>
</tr>
<tr>
<td>4</td>
<td>9</td>
<td>2.715</td>
<td>4.96</td>
<td>4.96</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>1.283</td>
<td>3.04</td>
<td>3.12</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0.631</td>
<td>1.92</td>
<td>2.00</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0.315</td>
<td>1.28</td>
<td>1.28</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0.353</td>
<td>2.88</td>
<td>2.88</td>
</tr>
</tbody>
</table>

Total 80 80 80 80

Estimates of $\hat{\theta}_1 = 0.59$ $\hat{\theta}_1 = 0.73$ $\hat{\theta}_1 = 0.73$

$\hat{\theta}_2 = 0.032$ $\hat{\theta}_2 = 0.04$

Chi-square Values 24.506 0.483 0.428

d.f 2 1 1

P-values $< 0.000001$ 0.492 0.513

Table 2: Observed frequencies and computed values of expected frequencies of the LSD and the ELSD by the method of moments and the method of maximum likelihood for the second data set.

<table>
<thead>
<tr>
<th>No. of cases</th>
<th>Observation</th>
<th>LSD</th>
<th>ELSD method of moments</th>
<th>ELSD method of maximum likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>156</td>
<td>179.080</td>
<td>156.282</td>
<td>159.478</td>
</tr>
<tr>
<td>2</td>
<td>55</td>
<td>42.108</td>
<td>54.908</td>
<td>53.724</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>13.310</td>
<td>18.134</td>
<td>17.182</td>
</tr>
<tr>
<td>4</td>
<td>10</td>
<td>4.598</td>
<td>7.223</td>
<td>6.776</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2.904</td>
<td>5.453</td>
<td>4.840</td>
</tr>
<tr>
<td>Total</td>
<td>242</td>
<td>242</td>
<td>242</td>
<td>242</td>
</tr>
</tbody>
</table>

Estimates of $\hat{\theta}_1 = 0.47$ $\hat{\theta}_1 = 0.41$ $\hat{\theta}_1 = 0.40$

$\hat{\theta}_2 = 0.06$ $\hat{\theta}_2 = 0.055$

Chi-square Values 12.051 3.295 0.311

d.f 2 2 1

P-values 0.04 0.298 0.568

and

$$\sum_{x=1}^{y} a(x) \left[ -C \left( 1 - \theta_1 - \theta_2 \right) + \Phi^{-1}(x; \theta_1; \theta_2) \sum_{r=1}^{\left[ \frac{x}{r} \right]} \left( x - r - 1 \right)! \frac{\theta_1^{x-r-\theta_2}}{(x-2r)! (r-1)!} \right] = 0. \quad (3.8)$$

Likelihood equations do not always have a solution because the ELSD is not a regular model; subsequently, likelihood equations do not always have a solution for the maximum of the likelihood function attained at the border of the domain of parameters. We obtained the second order partial derivatives of log $g_x$ with respect to parameters $\theta_1$ and $\theta_2$ and we observed (using MATHCAD software) that these equations give negative values for all $\theta_1 > 0$ and $\theta_2 > 0$ such that $\theta_1 + \theta_2 < 1$. Thus the density of the ELSD is a log-concave and have maximum likelihood estimates where the parameters $\theta_1$ and $\theta_2$ are unique (cf. Puig, 2003). Now on solving these two likelihood equations by using mathematical software such as (MATHLAB, MATHCAD, and MATHEMATICA), one can obtain maximum likelihood estimates of the parameters $\theta_1$ and $\theta_2$ of the ELSD.

For numerical illustration, we considered two data sets where the first data sets is from a zero-
On Estimating the Parameters of an Extended Form of Logarithmic Series Distribution

4. Testing of the Hypothesis

In this section we discuss the testing of the hypothesis \( H_0 : \theta_2 = 0 \) against the alternative hypothesis \( H_1 : \theta_2 \neq 0 \) by using generalized likelihood ratio test and Raos efficient score test.

4.1. Generalized likelihood ratio test

In case of generalized likelihood ratio test, the test statistic is

\[
-2 \log \lambda = 2 \left[ \log L(\hat{\theta}; x) - \log L(\hat{\theta}^*; x) \right],
\]

where \( \hat{\theta} \) is the maximum likelihood estimate of \( \theta = (\theta_1, \theta_2) \) with no restrictions, and \( \hat{\theta}^* \) is the maximum likelihood estimate of \( \theta \) when \( \theta_2 = 0 \). The test statistic \( -2 \log \lambda \) given in (4.1) is asymptotically distributed as \( \chi^2 \) with one degree of freedom (for details see Rao, 1973). We have computed the values of \( \log L(\hat{\theta}; x) \), \( \log L(\hat{\theta}^*; x) \) and the test statistic for the ELSD and present them in Table 3. Since the critical value for the test at 5% level of significance is 3.84 at one degree of freedom, the null hypothesis is rejected in both the cases.

5. Rao’s Efficient Score Test

In case of Raos score test, the statistic is

\[
T = V^\prime \phi^{-1} V,
\]

where \( V^\prime = ((1/\sqrt{n})(\partial \log L/\partial \theta_1), (1/\sqrt{n})(\partial \log L/\partial \theta_2)) \) and \( \phi \) is the Fisher information matrix. The test statistic given in (4.2) follows chi-square distribution with one degree of freedom (for details see Rao, 1973). We have computed the values of \( T \) for (i) the ELSD for the first data set as \( T_1 \) (ii) the ELSD for the second data set as \( T_2 \) are given below

\[
T_1 = \begin{pmatrix} 50.175 & 50.044 \\ 0.0076 & 0.0030 & 50.175 \\ 0.0030 & 0.0061 & 50.044 \end{pmatrix}
= 49.476,
\]

\[
T_2 = \begin{pmatrix} -0.247 & 25.771 \\ 0.045 & -0.05 & -0.247 \\ -0.05 & 0.0097 & 25.771 \end{pmatrix}
= 6.674.
\]
Table 4: Bias and standard errors of the estimators of the parameters $\theta_1$ and $\theta_2$ of the ELSD for the simulated data sets.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Sample size</th>
<th>Method of moments</th>
<th>Method of maximum likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\hat{\theta}_1$</td>
<td>$\hat{\theta}_2$</td>
</tr>
<tr>
<td>1</td>
<td>75</td>
<td>0.12 (0.395)</td>
<td>0.08 (0.350)</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.08 (0.169)</td>
<td>0.07 (0.150)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.07 (0.136)</td>
<td>0.04 (0.125)</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>0.25 (0.167)</td>
<td>0.07 (0.164)</td>
</tr>
<tr>
<td></td>
<td>150</td>
<td>0.23 (0.159)</td>
<td>0.05 (0.143)</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.20 (0.138)</td>
<td>0.04 (0.121)</td>
</tr>
</tbody>
</table>

Since the critical value for the test at 5% level of significance is 3.84 at one degree of freedom, the null hypothesis is rejected in both cases.

6. Simulation

It is difficult to compare the theoretical performance of estimators of different parameters of the ELSD obtained by method of moments and method of maximum likelihood. So in this section we tried a simulation study to compare the performance of estimators obtained by both methods of estimation. We simulated two sets for the following two data sets of parameters: (i) $\theta_1 = 0.33, \theta_2 = 0.12$, (under dispersion) and (ii) $\theta_1 = 0.56, \theta_2 = 0.08$, (over dispersion). Table 4 presents the computed values of the bias and standard errors of each of the estimators are presented in Table 4.

Table 4 shows that both the bias and standard errors of the estimators of both parameters are in decreasing order as the sample size increases.

References


Received July 20, 2013; Revised September 12, 2013; Accepted September 12, 2013