9-2019

Bias-corrected Maximum-Likelihood Estimator for the Parameter of the Logarithmic Series Distribution and its Characterizations

Mahdi Rasekhi

Gholamhossein Hamedani

Bias-corrected Maximum-Likelihood Estimator for the Parameter of the Logarithmic Series Distribution and its Characterizations

Mahdi Rasekhi*,† and Gholamhossein G. Hamedani‡

† Malayer University
‡ Marquette University

Abstract. In this article, we study parameter estimation of the logarithmic series distribution. A well-known method of estimation is the maximum likelihood estimate (MLE) and this method for this distribution resulted in a biased estimator for the small sample size datasets. The goal here is to reduce the bias and root mean square error of MLE of the unknown parameter. Employing the Cox and Snell method, a closed-form expression for the bias-reduction of the maximum likelihood estimator of the parameter is obtained. Moreover, the parametric Bootstrap bias correction of the maximum likelihood estimator is studied. The performance of the proposed estimators is investigated via Monte Carlo simulation studies. The numerical results show that the analytical bias-corrected estimator performs better than bootstrapped-based estimator and MLE for small sample sizes. Also, certain useful characterizations of this distribution are presented. An example via a real dataset is presented for the illustrative purposes.

Keywords. Cox-Snell bias-correction; Bootstrap bias-correction; Logarithmic series distribution; Maximum likelihood estimator; Monte Carlo simulation.

MSC 2010: 62F10, 62F40.

---

* Corresponding author

1   Introduction

The logarithmic series (LS) distribution is one of the popular discrete distributions. This distribution is a limiting case of the zero truncated negative binomial distribution. The probability mass function (PMF) of the LS distribution is given by

$$f(x;p) = \frac{-1}{\log(1-p)}\frac{p^x}{x}, \quad x \in \{1,2,3,...\}. \tag{1}$$

The LS distribution is used for modeling biological, behavioral and ecological datasets, see Fisher et al. (1943), Gupta and Jain (1976) and Izsak and Juhasz-Nagy (1982). Some of the researchers presented certain estimators of the parameter of the LS distribution. For example: Patil (1962) obtained ML and moment estimations of the parameter of the LS distribution. The minimum variance unbiased (UMVU) estimation of the parameter is calculated by Patil and Bildikar (1966). Bohning (1983) presented two numerical methods for obtaining the ML estimation of the parameter of the LS distribution. Many researchers studied other aspects of this distribution, for example: Gower (1961) presented some asymptotic properties of this distribution, Ramalingam and Jagbir (1984) studied a characterization and discussed the goodness of fit test for this distribution. Wani and Lo (1975) obtained large sample interval estimation. Aghababaei Jazi and Alamatsaz (2010) investigated ordering comparison of LS random variables based on the mixtures.

Among all of the methods of parameter estimation, the maximum-likelihood estimation method is the most frequently used since this method is simple, intuitive and posses good properties such as asymptotically unbiasedness, consistent, and asymptotically normally distributed. But, it is well known that in finite samples this method does not include any desirable sampling properties. blueIn particular, the MLE is often biased on small finite sample size in specially for the LS distribution. Cox and Snell (1968) suggested a method to obtain analytical expressions for the bias, with $O(n^{-1})$, of the maximum likelihood estimators. This method can be used for bias-correct of the maximum likelihood estimators with $O(n^{-2})$. Another approach is the parametric Bootstrap resampling method suggested by Efron (1979). In this method bias correction is obtained numerically without analytical expression for the bias function. Some researchers used these methods for other distributions for example: Lemonte et al. (2007), Giles (2012), Wang and Wang

(2017), Reath et al. (2018), Mazucheli et al. (2018) to name a few. blue One of the motivations of this article is to obtain better estimators of p than MLE in term of bias and standard error. Another motivation is to present some novel characterizations of the LS distribution. In this article, based on the above two approaches, we presented two bias corrected maximum likelihood estimations for the parameter of LS distribution.

This article is organized as follows. Section 2 briefly discusses the maximum likelihood estimation of the parameter of the LS distribution. An analytical bias correction estimator for the parameter of this distribution is obtained in Section 3. Section 4 includes the bias-correction mechanism based on Efron's bootstrap resampling. In Section 5, Monte Carlo simulations are used for comparing the performance between two proposed methods with MLE. An application to a real dataset is presented for illustrative purposes in Section 6. Certain characterizations of the LS distribution are provided in Section 7. Finally, Section 8 concludes the article.

2 Maximum Likelihood Estimation

Let $x_1, x_2, ..., x_n$ be observed values of a random sample of size n from the LS distribution with parameter p. The log-likelihood function is

$$l(p) = -n \log(-\log(1-p)) + \log(p) \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \log(x_i). \qquad (2)$$

The maximum likelihood estimate of the p, $\hat{p}$, can be obtained by solving the nonlinear equation:

$$\frac{dl(p)}{dp} = \frac{n}{(1-p)\log(1-p)} + \frac{\sum_{i=1}^{n} x_i}{p}. \qquad (3)$$

To obtain an interval estimation of the parameter via the maximum likelihood estimate $\hat{p}$, we need the expected Fisher information, $nE\left(-\frac{d^2}{dp^2} \log f(x;p)\right)$,which is presented as follows:

$$K(p) = \frac{-n\left(\log(1-p) + 1\right)}{\left((1-p)\log(1-p)\right)^2} - \frac{n}{p(1-p)\log(1-p)}. \qquad (4)$$

The inverse of the expected Fisher information is

$$K^{-1}(p) = \left(\frac{-n\left(\log(1-p)+1\right)}{\left((1-p)\log(1-p)\right)^2} - \frac{n}{p(1-p)\log(1-p)} \right)^{-1}, \qquad (5)$$

and if (5) is evaluated at $\hat{p}$, the estimate of the asymptotic variance of the maximum likelihood estimate is provided. So, the asymptotic $100 \times (1-\alpha)\%$ confidence interval for p is given by

$$\hat{p} \pm z_{1-\frac{\alpha}{2}} \sqrt{K^{-1}(\hat{p})},$$

where $z_{1-\frac{\alpha}{2}}$ is the $100 \times (1-\frac{\alpha}{2})\%$ percentile of the standard normal distribution.

3   Closed form Bias-corrected Maximum Likelihood Estimation

Cox and Snell (1968) proposed a method for deriving closed-form expressions for the second order biases of MLE of the parameters. They illustrated that when the sample data are independent, but not necessarily identically distributed, the bias of the s-th element of the MLE of Θ, $\hat{\Theta}$, can be obtained as:

$$Bias(\hat{\Theta}_s) = \sum_{i=1}^{q}\sum_{j=1}^{q}\sum_{l=1}^{q} \kappa^{si}\kappa^{jl}\left[0.5\kappa_{ijl} + \kappa_{ij,l}\right] + O(n^{-2}), \quad s = 1,2,3,...,q,$$

$$(6)$$

where κ^{ij} is the (i,j)th element of the inverse of the information matrix, K. The joint cumulants of the derivatives of $l(\Theta)$ are

$$k_{ij} = E\left(\frac{d^2}{d\Theta_i d\Theta_j}l(\Theta)\right), \ i,j = 1,2,...,q,$$

$$k_{ijl} = E\left(\frac{d^3}{d\Theta_i d\Theta_j d\Theta_l}l(\Theta)\right), \ i,j,l = 1,2,...,q,$$

$$k_{ij,l} = E\left(\frac{d^2}{d\Theta_i d\Theta_j}l(\Theta).\frac{d}{d\Theta_l}l(\Theta)\right), \ , \ i,j,l = 1,2,...,q.$$

Moreover, Cordeiro and Klein (1994) illustrated that equation (6) can be rewritten as:

$$Bias(\hat{\Theta}_s) = \sum_{i=1}^{q} \kappa^{si} \sum_{j=1}^{q} \sum_{l=1}^{q} \left[\kappa_{ij}^{(l)} - \frac{1}{2} \kappa_{ijl} \right] + O(n^{-2}), \quad s = 1, 2, 3, ..., q, \quad (7)$$

where $\kappa_{ij}^{(l)}$ is $\frac{d\kappa_{ij}}{d\Theta_l}$. Now by defining $a_{ij}^{(l)} = \kappa_{ij}^{(l)} - \frac{1}{2}\kappa_{ijl} \quad \forall i, j, l = 1, 2, 3, ..., q,$ matrix $A^{(l)} = \left[a_{ij}^{(l)} \right], \quad \forall i, j, l = 1, 2, 3, ..., q$ and blockwise matrix $A = \left[A^{(1)} \mid A^{(2)} \mid \cdots \mid A^{(q)} \right]$, we can obtain the bias of $\hat{\Theta}$ as

$$Bias(\hat{\Theta}) = K^{-1} A \, vec(K^{-1}) + O(n^{-2}). \quad (8)$$

Finally the bias corrected of $\hat{\Theta}$ is

$$\hat{\Theta}_{BC} = \hat{\Theta} - \hat{K}^{-1} \hat{A} \, vec(\hat{K}^{-1}),$$

where $\hat{K}$ and $\hat{A}$ are the values of K and A by MLEs and vec (.) denotes the "vectorization" operator.

Since the logarithmic series is considered in this article, we have $q = 1$ with $\Theta = (p)$. The cumulants of the derivatives of the log-likelihood function of LS distribution are provided below. The details on the mathematical derivations are presented in the "Appendix". The joint cumulants are

$$\kappa_{11} = \frac{n(\log(1-p)+1)}{((1-p)\log(1-p))^2} + \frac{n}{p(1-p)\log(1-p)},$$

$$(9)$$

$$\kappa_{111} = \frac{-n(1-p)(\log(1-p))^2 + 2n(1-p)\log(1-p)(\log(1-p)+1)^2}{((1-p)\log(1-p))^4} - \frac{2n}{p^2(1-p)\log(1-p)}.$$

The derivatives of the joint cumulants is

$$\kappa_{11}^{(1)} = \frac{-n(1-p)(\log(1-p))^2 + 2n(1-p)\log(1-p)(\log(1-p)+1)^2}{((1-p)\log(1-p))^4}$$
$$+ \frac{n(p+2p\log(1-p)-\log(1-p))}{(p(1-p)\log(1-p))^2}.$$

$$(10)$$

Using (9) and (10), we have

$$
\begin{aligned}
A = a_{11}^{(1)} &= \kappa_{11}^{(1)} - \tfrac{1}{2}\kappa_{111} \\
&= \frac{-n(1-p)(\log(1-p))^2 + 2n(1-p)\log(1-p)(\log(1-p)+1)^2}{((1-p)\log(1-p))^4} \\
&\quad + \frac{n(p+2p\log(1-p)-\log(1-p))}{(p(1-p)\log(1-p))^2} \\
&\quad - \frac{-\tfrac{n}{2}(1-p)(\log(1-p))^2 + n(1-p)\log(1-p)(\log(1-p)+1)^2}{((1-p)\log(1-p))^4} \\
&\quad + \frac{n}{p^2(1-p)\log(1-p)}.
\end{aligned}
\tag{11}
$$

Thus, from (5) and (11), the bias of the MLE of the LS distribution is given by

$$
Bias(p) = K^{-1}(p)\,A\,vec(K^{-1}(p)) + O(n^{-2}) = \frac{A}{K^2(p)} + O(n^{-2}).
$$

Then, the bias-corrected MLE can be obtained as

$$
\hat{p}_{BC} = \hat{p} - \frac{\hat{A}}{\hat{K}^2(p)},
\tag{12}
$$

where $\hat{A}$ and $\hat{K}(p)$ are A and $K(p)$ with p replaced by $\hat{p}$.

4 Bootstrap form Bias-corrected Maximum Likelihood Estimation

In this section, the bootstrap resampling method (Efron, 1979) is used to reduce the bias of the MLEs. First, by generating pseudo-samples from the original sample, the biases of the MLEs are estimated. Then, the new bias-corrected MLE is obtained by subtraction of the estimated biases from the original MLEs.

Suppose $x = (x_1, x_2, ..., x_n)$ is a sample from the random variable X with the distribution function F. Suppose $\eta = t(F)$ is a function of F similar to a parameter and $\hat{\eta} = s(x)$ is an estimator of η. Based on the Efron's bootstrap method, we select a large number of pseudo-samples $x^* = (x_1^*, x_2^*, ..., x_n^*)$ from x and campute the corresponding bootstrap replicates of $\hat{\eta}$, say $\hat{\eta}^* = s(x^*)$. Then, an estimate of the distribution function of $\hat{\eta}$ is the empirical distribution of $\hat{\eta}$. If F is a known parametric distribution, F_η, then an estimate of F is obtained by using consistent estimator for $F_{\hat{\eta}}$. The bias of

$\hat{\eta} = s(x)$ is

$$B_F(\hat{\eta}, \eta) = E_F(s(X)) - \hat{\eta}(F),$$

where the index F shows that the expectation is performed with respect to the distribution F. The bootstrap bias estimate is calculated by replacing F, based on the original sample was obtained, by $F_{\hat{\eta}}$. Hence, the bias can be obtained as

$$B_{F_{\hat{\eta}}}(\hat{\eta}, \eta) = E_{F_{\hat{\eta}}}(\hat{\eta}) - \hat{\eta}. \tag{13}$$

Upon generating N independent bootstrap samples from the original sample x, the bootstap estimates are $(\hat{\eta}^{*(1)}, \hat{\eta}^{*(2)}, ..., \hat{\eta}^{*(N)})$. When N tends to larger value, the expected value $E_{F_{\hat{\eta}}}(\hat{\eta})$ is approximated by

$$\hat{\eta}^{*(.)} = \frac{1}{N} \sum_{i=1}^{N} \hat{\eta}^{*(i)}.$$

Thus, the bootstrap estimate of (13) is

$$B_{F_{\hat{\eta}}}(\hat{\eta}, \eta) = \hat{\eta}^{*(.)} - \hat{\eta},$$

and finally the second-order bias-corrected MLEs of the LS distribution is obtained as

$$\hat{\eta}_{BBC} = 2\hat{\eta} - \hat{\eta}^{*(.)}. \tag{14}$$

It is worth to mention that the estimator $\hat{\eta}_{BBC}$ should be named the constant bias-corrected MLE since the function is approximated by a constant (MacKinnon and Smith, 1998). So, in this article, we let $\hat{\eta} = \hat{p}$ and $\hat{\eta}_{BBC} = \hat{p}_{BBC}$.

5 Simulation Study

In this section, Monte Carlo simulations are used to compare the behavior of the MLE of p and its proposed bias corrections based on the finite-sample. The data were simulated via the rlogarithmic function in the "actuar" R package created by Goulet (2018). Random samples are generated with the sample size $n = 10, 20, 30$ and real values of the parameter are $p = 0.05, 0.4, 0.5, 0.6, 0.7, 0.8, 0.95$. For each combination of (n, p), the Monte Carlo replications in our simulations is $M = 3000$. Also the bootstrap replications are assumed B=4000. Since the MLE of p has bias with both positive and negative sign, the comparison is performed via mean of

the absolute empirical biases and root mean squared error criteria, i.e.

$$Bias(\hat{p}_i^{est}) = \frac{1}{M} \sum_{i=1}^{M} \left| \hat{p}_i^{est} - p \right| \quad and \quad RMSE(\hat{p}_i^{est}) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left(\hat{p}_i^{est} - p \right)^2},$$

where $\hat{p}^{est}$ is an estimator of p. The MLE of p is calculated by "optim" function in R statistical program. The results are presented in Table 1.

From Table 1, the performance of $BC-MLE$ is better than that of MLE and $BBC-MLE$ based on mean absolute bias and root mean squared error for sample with small size. In most of the cases, the biases of all the estimators of p decrease when sample size increases except for the value of p near zero and one.

6   Illustrative Example

In this section, a real dataset is considered and the proposed methods in the previous sections are illustrated. Best et al. (2008) presented three methods of goodness of fit test for logarithmic series distribution. They used a dataset on insect catches from the Sierra Tarahuma, Mexico, reported by Al-drete (2002) and illustrated that the LS distribution is a good fit for this dataset (Table 2). Table 3 shows the point estimates of p with standard errors of MLE, BC-MLE and BBC-MLE. The standard errors of BC-MLE and BBC-MLE are obtained by parametric bootstrap method with 2000 replications and the same sample size (For example, this method is applied in section 5 of Rasekhi et al. (2017)). It is observed that BC-MLE provides the lowest standard errors for the parameter p. It is worth to mention that, the BC-MLE and BBC-MLE of p are bigger than the MLE. Thus the MLE underestimate the parameter p of the LS distribution.

7   Characterizations Results

The problem of characterizing a distribution is an important problem in applied sciences, where an investigator is vitally interested to know if their model follows the right distribution. To this end, the investigator relies on conditions under which their model would follow specifically the chosen distribution. In this section, we present two characterizations of the LS distribution based on: (i) conditional expectation of certain function of the

Table 1. Estimated mean absolute bias (root mean-squared error) for estimators of p

p	n	Estimators of p		
		MLE	BC-MLE	BBC-MLE
0.05	10	0.2913(0.3788)	0.2586(0.3453)	0.2689(0.3664)
	20	0.2667(0.3446)	0.2503(0.3284)	0.2542(0.3399)
	30	0.3475(0.4189)	0.3348(0.4063)	0.3412(0.4165)
0.4	10	0.2325(0.2495)	0.2164(0.2337)	0.2155(0.2546)
	20	0.2062(0.2298)	0.2017(0.2236)	0.2112(0.2400)
	30	0.2155(0.2339)	0.2152(0.2297)	0.2221(0.2423)
0.5	10	0.2391(0.2665)	0.2317(0.2612)	0.2541(0.2145)
	20	0.2145(0.2404)	0.2072(0.2350)	0.2222(0.2124)
	30	0.1894(0.2145)	0.1865(0.2516)	0.1939(0.2214)
0.6	10	0.3682(0.3972)	0.3481(0.3787)	0.3701(0.4008)
	20	0.3005(0.3426)	0.2918(0.3342)	0.3016(0.3468)
	30	0.2920(0.3358)	0.2859(0.3300)	0.2926(0.3392)
0.7	10	0.3972(0.4438)	0.3778(0.4239)	0.3969(0.4429)
	20	0.4546(0.4963)	0.4433(0.4867)	0.4588(0.5043)
	30	0.3085(0.3602)	0.2981(0.3516)	0.3032(0.3593)
0.8	10	0.5268(0.5662)	0.5026(0.5447)	0.5242(0.5672)
	20	0.4782(0.5268)	0.4628(0.5146)	0.4753(0.5298)
	30	0.2724(0.3500)	0.2695(0.3456)	0.2728(0.3516)
0.95	10	0.5621(0.6308)	0.5326(0.6055)	0.5529(0.6263)
	20	0.4976(0.5909)	0.4842(0.5788)	0.4932(0.5902)
	30	0.6109(0.6619)	0.6003(0.6530)	0.6077(0.6620)

Table 2. Insect dataset.

Times caught	1	2	3	4	5	6	8	10	11	12	13	16	25	69	95
Frequency	10	3	4	2	2	2	1	1	1	1	1	1	2	1	1

Table 3. MLE and bias-corrected MLEs (standard error).

Estimators	Estimate of p (SE)
MLE	0.9743(0.0098)
BC-MLE	0.9777(0.0088)
BBC-MLE	0.9795(0.0096)

random variable ; (ii) the hazard rate function.

The cumulative distribution function (CDF), $F(x;p)$, corresponding PMF, $f(x;p)$ and hazard rate function, $h_F(x;p)$ can be expressed, respectively, as

$$F(x;p) = 1 - CB(p;x+1,0), \quad x = 1, ..., \tag{15}$$

$$f(x;p) = C\{B(p;x,0) - B(p;x+1,0)\}, \quad x = 1, ..., \tag{16}$$

$$h_F(x) = \frac{B(p;x,0)}{B(p;x+1,0)} - 1, \quad x = 1, ..., \tag{17}$$

where $p \in (0,1)$ is a parameter, $C = -\frac{1}{\log(1-p)}$ and

$B(\xi, a, b) = \int_0^\xi u^{a-1}(1-u)^{b-1}du$. We present our characterizations (i) and (ii) via two subsections 7.1 and 7.2.

7.1   Characterizations of LS in terms of the conditional expectation of certain function of the random variable

Proposition 1. Let $X : \Omega \to \mathbb{N}$ be a random variable. The PMF of X is (16) if and only if

$$E\{[B(p;X,0) + B(p;X+1,0)] \mid X > k\} = B(p;k+1,0). \tag{18}$$

Proof. If X has PMF (16), then the left-hand side of (18) will be

$$(1 - F(k;p))^{-1} \sum_{x=k+1}^{\infty} C\left\{[B(p;x,0)]^2 - [B(p;x+1,0)]\right\}$$

$$= \frac{C}{CB(p;k+1,0)} \left\{[B(p;k+1,0)]^2 - 0\right\} = B(p;k+1,0).$$

Conversely, if (18) holds, then

$$\begin{aligned}
\sum_{x=k+1}^{\infty} & \{[B(p;x,0) + B(p;x+1,0)] f(x;p)\} \\
& = (1 - F(k;p))B(p;k+1,0) \\
& = [(1 - F(k+1;p)) + f(k+1)] B(p;k+1,0).
\end{aligned} \tag{19}$$

From (18), we also have

$$\sum_{x=k+2}^{\infty} \{[B(p;x,0) + B(p;x+1,0)]\, f(x;p)\} \\ = (1 - F(k+1;p))B(p;k+2,0). \tag{20}$$

Now, subtracting (20) from (19), we arrive at

$$(1 - F(k+1;p))\,[B(p;k+1,0) - B(p;k+2,0)] = f(k+1)\,B(p;k+2,0).$$

From the last equality, we have

$$h_F(k+1) = \frac{f(k+1)}{1 - F(k+1)} = \frac{B(p;k+1,0)}{B(p;k+2,0)} - 1,$$

which, in view of (17), implies that X has PMF (16). $\qquad\square$

7.2 Characterization of LS based on hazard function

Proposition 2. Let $X : \Omega \to \mathbb{N}$ be a random variable. The PMF of X is (16) if and only if its hazard rate function satisfies the difference equation

$$h_F(k+1) - h_F(k) = \frac{B(p;k+1,0)}{B(p;k+2,0)} - \frac{B(p;k,0)}{B(p;k+1,0)}, \quad k \in \mathbb{N}, \tag{21}$$

with the initial condition $h_F(1) = \frac{B(p,1,0)}{B(p,2,0)} - 1$.

Proof. If X has PMF (16), then clearly (21) holds. Now, if (21) holds, then for every $x \in \mathbb{N}$, we have

$$\sum_{k=1}^{x-1} \{h_F(k+1) - h_F(k)\} \\ = \sum_{k=1}^{x-1} \left\{ \frac{B(p;k+1,0)}{B(p;k+2,0)} - \frac{B(p;k,0)}{B(p;k+1,0)} \right\},$$

or

$$h_F(x) - h_F(1) = \frac{B(p;x,0)}{B(p;x+1,0)} - \frac{B(p;1,0)}{B(p;2,0)}.$$

In view of the fact that $h_F(1) = \frac{B(p;1,0)}{B(p;2,0)} - 1$, from the last equation we have

$$h_F(x) = \frac{B(p;x,0)}{B(p;x+1,0)} - 1,$$

which, in view of (17), implies that X has PMF (16).   □

8   Conclusions

The second-order bias-corrected MLE of the logarithmic series distribution based on Cox and Snell (1968) corrective approach is presented and this estimator is called BC-MLE. This estimator has explicit expressions (assuming MLE is known) and also reduces the bias and the root mean square error (RMSE) of the parameter of the LS distribution. For a comparison study, Efron's bootstrap resampling method is used to obtain another bias-corrected MLE called BBC-MLE. The results of simulation studies and real data application strongly suggest that the BC-MLE is the best choice when the sample size is small. Also, certain characterizations of the LS distribution based on conditional expectation of certain function of the random variable and hazard function are presented.

References

Aghababaei Jazi, M. and Alamatsaz, M.H. (2010). Ordering Comparison of Logarithmic Series Random Variables with Their Mixtures. *Communications in Statistics-Theory and Methods*, **39**, 3252-3263.

Aldrete, A.N.G. (2002). Psocoptera (Insecta) from the Sierra Tarahumara, Chihuahua, Mexico. *Anales del Instituto de Biologia, Universidad Nacional Autonoma d e Mexico, Serie Zoologia*, **73**, 145-156.

Best, D.J., Rayner, J.C. and Thas, O. (2008). Tests o f Fit for the Logarithmic Distribution*Journal of Applied Mathematics and Decision Sciences*, DOI:10.1155/2008/463781.

Bohning, D. (1983). Maximum Likelihood Estimation of the Logarithmic Series Distribution. *Statistische Hefte*, **24**, 121-140.

Cordeiro, G.M. and Klein, R. (1994). Bias Correction in ARMA Models. *Statistics and Probability Letters*, **19**, 169-176.

Cox, D.R., Snell, E.J. (1968). A General Definition of Residuals. *J. Royal Stat. Soc. Ser. B*, **30**, 248-275.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *Annals of Statistics*, **7**, 1-26.

Fisher, R.A., Corbert, A.S. and Williams, C.B. (1943). The Relation between the Number of Species and the Number of Individuals in a Random Sample of an Animal Population.*Journal of Animal Ecology*, **12**, 42-58.

Giles, D.E. (2012). Bias Reduction for the Maximum Likelihood Estimators of the Parameters in the Half-Logistic Distribution. *Communications in Statistics-Theory and Methods*, **41**, 212-222.

Goulet, V. (2018). Actuar: Actuarial Functions and Heavy Tailed Distributions. R Package Version 3.3.0. https://cran.r-project.org/web/packages/actuar/index.html.

Gower, J.C. (1961). A Note on Some Asymptotic Properties of the Logarithmic Series Distribution. *Biometrika*, **48**, 212-215.

Gupta, R.P. and Jain, G.C. (1976). A Generalized Bivariate Logarithmic Series Distribution. *Biometrical Journal*, **18**, 169-173.

Izsak, J. and Juhasz-Nagy, P. (1982). Studies of Lognormality on Mortality Statistics. *Biometrical Journal*, **24**, 731-741.

Lemonte, A.J. , Cribari-Neto, F., Vasconcellos, K.L.P. (2007). Improved Statistical Inference for the Two-parameter Birnbaum-Saunders Distribution. *Computational Statistics and Data Analysis*, **51**, 4656-4681.

MacKinnon, J. and Smith, A. (1998). Approximate Bias Correction in Econometrics. *Journal of Econometrics*, **85**, 205-230.

Mazucheli, J., Menezes, A.F.B. and Dey, S. (2018). Bias-corrected Maximum Likelihood Estimators of the Parameters of the Inverse Weibull Distribution. *Communications in Statistics-Simulation and Computation*, DOI: 10.1080/03610918.2018.1433838.

Patil, G.P. (1962). Some Methods of Estimation for the Logarithmic Series Distribution. *Biometrics*, **18**, 68-75.

Patil, G.P. and Bilidikar, S. (1966). On Minimum Variance Unbiased Estimation for the Logarithmic Series Distribution. *Sankhya: The Indian Journal of Statistics*, **28**, 239-250.

Ramalingam, S. and Jagbir, S. (1984). A Characterization of the Logarithmic Series Distribution and Its Application. *Communications in Statistics - Theory and Methods*, **13**, 865-875.

Rasekhi, M., Hamedani, G.G. and Chinipardaz, R. (2017). A Flexible Extension of Skew Generalized Normal Distribution. *METRON*, **75**, 87-107.

Reath, J., Dong, J. and Wang, M. (2018). Improved Parameter Estimation of the Log-logistic Distribution with Applications. *Computational Statistics*, **33**, 339-356.

Wang, M. and Wang, W. (2017). Bias-Corrected Maximum Likelihood Estimation of the Parameters of the Weighted Lindley Distribution. *Communications in Statistics-Simulation and Computation*, **46**, 530-545.

Wani, J.K. and Lo, H.P. (1975). Large Sample Interval Estimation for the Logarithmic Series Distribution. *Canadian Journal of Statistics*, **3**, 277-284.

Appendix

The derivatives of the log-likelihood function (2) up to three order with respect to parameter p are presented as follows:

$$\frac{dl(p)}{dp} = \frac{n}{(1-p)\log(1-p)} + \frac{\sum\limits_{i=1}^{n} x_i}{p},$$

$$\frac{d^2l(p)}{dp^2} = \frac{n\{-\log(1-p)+1\}}{((1-p)\log(1-p))^2} - \frac{\sum\limits_{i=1}^{n} x_i}{p^2},$$

$$\frac{d^3l(p)}{dp^3} = \frac{-n(1-p)\{\log(1-p)\}^2 + 2(1-p)\log(1-p)\{\log(1-p)+1\}}{((1-p)\log(1-p))^4}$$
$$+ \frac{2\sum\limits_{i=1}^{n} x_i}{p^3}.$$

The expectation of X

$$E(X) = \frac{-p}{(1-p)\log(1-p)},$$

is needed for finding the cumulants.

Mahdi Rasekhi
Department of Statistics,
Malayer university,
Malayer, Iran.
email: *rasekhimahdi@gmail.com*

Gholamhossein G. Hamedani
Department of Statistics,
Marquette University,
Milwaukee, USA.
email: *g.hamedani@mu.edu*