Estimation in Misclassified Size – Biased Log Series Distribution

B. S. Trivedi[1*] M. N. Patel[2]
1. H. L. Institute of Commerce, Ahmedabad University, Navrangpura, Ahmedabad
2. Department of Statistics, School of Sciences, Gujarat University, Ahmedabad
* E-mail: bhaktida.trivedi@ahduni.edu.in

Abstract

A misclassified size-biased Log Series Distribution (MSBLSD) where some of the observations corresponding to $x = c + 1$ are wrongly reported as $x = c$ with probability α, is defined. Various estimation methods like the method of maximum likelihood (ML), method of moments and the Bayes estimation for the parameters of MSBLS distribution are used and verified their efficiency through simulation study. Bayes estimators of the parameters are obtained. An example is presented for the size-biased log series distribution to illustrate the results and a goodness of fit test is also done using the method of maximum likelihood.

Key words: Log Series Distribution, Misclassification, Size – Biased, Method of Moments, Maximum likelihood, Bayes estimation

1. Introduction

The logarithmic series distribution was introduced by Fisher, Corbett and Williams (1943) to investigate the distribution of butterflies in the Malayan Peninsula, and data by Williams (1947) on the number of moths of different species caught in a light-trap in a given period. Chatfield et al (1966) used the logarithmic series distribution (LSD) to represent the distribution of numbers of items of a product purchased by a buyer in a specified time period. In probability and statistics, the logarithmic distribution (also known as logarithmic series distribution) is a discrete probability distribution derived from the Maclaurin series expansion for $0 < \theta < 1$.

$$-\ln(1-\theta) = \theta + \frac{\theta^2}{2} + \frac{\theta^3}{3} + \cdots \tag{1}$$

From this we obtain the identity

$$\sum_{x=1}^{\infty} \frac{-1}{\ln(1-\theta)} \frac{\theta^x}{x} = 1 \tag{2}$$

Because of the identity above, the distribution is properly normalized.

The logarithmic series distribution (LSD) characterized by a parameter θ is given by

$$P[X = x] = -\frac{1}{\log(1-\theta)} \frac{\theta^x}{x} = \beta \frac{\theta^x}{x}; \text{where } \beta = -\frac{1}{\log(1-\theta)} \ \& \ x = 1, 2, 3 \ldots \tag{3}$$

The equation (3) is a limiting form of zero-truncated negative binomial distribution. Sadinle (2008) linked the negative binomial distribution with the logarithmic series and Shanumugam and Singh (1984) studied the characterization of equation (3). A brief list of authors and their works can be seen in Johnson, Kotz and Kemp (2005).

The mean and variance of LSD are given as

$$\mu_1' = \beta \frac{\theta}{1-\theta} \quad , \quad \mu_2 = \beta \frac{\theta}{(1-\theta)^3}(1 - 2\theta + \theta^2) \tag{4}$$

When an investigator records an observation by nature according to certain stochastic model, the recorded observation will not have the original distribution unless every observation is given an equal chance of being recorded. For example, suppose that the original observation x_0 comes from a distribution with p.m.f./p.d.f.

$f_0(x_0)$ and that observation x is recorded according to a probability re-weighted by a weight function $w(x) > 0$, then x comes from a distribution with p.m.f./p.d.f.

$$f(x) = \frac{w(x)}{E[w(X_0)]} f_0(x) \tag{5}$$

The weighted distribution with $w(x) = x$ is called size-biased/length-biased distribution.

The size – biased logarithmic Series distribution (SBLSD) is obtained by taking the weight of the LSD (3) as x. We have from (3) and (4)

$$f(x) = \frac{w(x)}{E[w(X_0)]} f_0(x) = \frac{x}{\beta\frac{\theta}{1-\theta}} \beta \frac{\theta^x}{x} = (1-\theta)\theta^{x-1} \tag{6}$$

From (6) it is known that size-biased logarithmic distribution is a geometric distribution.

2. Miscalsiified Size-biased Logarithmic Series Distribution

Suppose the number of defectives actually present in a sample of size n is a size-biased logarithmic series variable with parameter θ and let α be the probability that a sample which actually contains $(c+1)$ defectives is misclassified by reporting it as containing only c defectives, and in all other cases the observations and reporting of defectives are found correct. Let X denote the number of defectives reported in a sample of k units. Then the probability function of the random variable X following a misclassified size-biased logarithmic series distribution $P_M[X = x]$ will be given in the following way:

$$P_M[X = c] = P_s[X = c] + \alpha P_s[X = c+1] \tag{7}$$

$$P_M[X = c+1] = P_s[X = c+1] - \alpha P_s[X = c+1] \tag{8}$$

from (6), (7) and (8) the p.d.f. of misclassified size-biased logarithmic series distribution is

$$P_M[X = x] = \begin{cases} (1-\theta)\theta^c\left(\frac{1}{\theta} + \alpha\right), & for\ x = c \\ (1-\alpha)(1-\theta)\theta^c, & for\ x = c+1 \\ (1-\theta)\theta^{x-1}, & for\ x \in S \end{cases} \tag{9}$$

$0 < \theta < 1,\ \ 0 \le \alpha \le 1,\ \ c = 1, 2, 3, \dots k-1$

and S is a subset of the set I of non $-$ negative integeres not containing c and $c+1$.

That is $S = I - [c, c+1]$, where I is a set of non negative integers.

Trivedi and Patel (2013) have considered estimation in misclassified size-biased generalized negative binomial distribution. In this paper we have discussed three methods of estimation to make the comparative analysis among the methods for the parameters of the Misclassified Size-Biased Logarithmic Series distribution through computer program.

3. Maximum Likelihood Estimators

To estimate the parameters α and θ involved in (9), we employ the method of ML when the samples are subject to the type of misclassification described above. Let n_i be the frequency corresponding to $x = i$ in n samples and so we have

$$\sum_{i=1}^{k} n_i = n, \text{the number of samples.}$$

Then the likelihood function of the sample of n observations following (9) may be written as

$$L \propto \prod_{i=1}^{k} P_i^{n_i} \tag{10}$$

Taking natural logarithm on both the sides, we have

$$\ln L = \sum_{i=1}^{k} n_i \ln(1-\theta) + \sum_{i=1}^{k} n_i\,(\mathrm{i}-1)\ln\theta + n_c ln\left(\frac{1}{\theta}+\alpha\right) + n_{c+1}\ln(1-\alpha) + n_c \ln\theta \tag{11}$$

Differentiating with respect to α and θ in turn and equating them to zero, we have

$$\alpha = \frac{n_c\theta - n_{c+1}}{\theta(n_{c+1}+n_c)} \quad \text{and} \tag{12}$$

$$\theta^2 n_c(n\bar{x}+n_c) + \theta n_c(n-n_c-n_{c+1}) + n_c[n_{c+1}-n(\bar{x}-1)] = 0 \tag{13}$$

The equation (13) is a quadratic equation, the roots are given by

$$\theta = \frac{-n_c(n-n_c-n_{c+1}) \pm \sqrt{\left(n_c(n-n_c-n_{c+1})\right)^2 - 4n_c(n\bar{x}+n_c)\{n_c[n_{c+1}-n(\bar{x}-1)]\}}}{2n_c(n\bar{x}+n_c)} \tag{14}$$

Here we consider only that value of $\theta \in (0,1)$.

The estimate of α can be obtained from equation (12) using the value of θ obtain in (14).

For $\alpha = 0$, that is where no misclassification has occurred than from (14) we get maximum likelihood estimate of θ as $\hat{\theta} = 1 - \frac{1}{\bar{x}}$

4. Method of Moments

r^{th} row moments of misclassified size-biased log series distribution denoted by $\mu'_{r(M)}$ is defined as

$$\mu'_{r(M)} = \sum_{x=1}^{\infty} x^r P_M(x) \tag{15}$$

Using the distribution in (9), we get

$$\mu'_{r(M)} = \mu'_{r(s)} + \alpha P_s(c+1)[c^r - (c+1)^r] \tag{16}$$

Where $\mu'_{r(s)} = \frac{1-\theta}{\theta\beta}\mu'_{r+1(L)} = r^{th}$ row moment of size biased log series distribution is given by

When $\mu'_{r+1(L)}$ is the $(r+1)^{th}$ row moment of log series distribution.

Using the above relationships of row moments, first four moments of MSBLSD can be derived as

$$\mu'_{1(M)} = \frac{1}{1-\theta}[1 - \alpha(1-\theta)^2\theta^c] \tag{17}$$

$$\mu'_{2(M)} = \frac{(1+\theta)}{(1-\theta)^2} - \alpha(1-\theta)\theta^c[2c+1] \tag{18}$$

$$\mu'_{3(M)} = \frac{1}{(1-\theta)^5}[(\theta^4 + 2\theta^3 - 6\theta^2 + 2\theta + 1) - \alpha(1-\theta)\theta^c(3c^2+3c+1)] \tag{19}$$

$$\mu'_{4(M)} = (1-\theta)^{-4}\{(1+\theta)(1+10\theta+\theta^2) + \beta\theta(1+4\theta+\theta^2)\} - \alpha(1-\theta)\theta[4c^3+6c^2+4c+1] \tag{20}$$

From the first and second row moments (30) and (31) of MSBLSD, the variance of MSBLSD is

$$\mu_{2(M)} = \frac{1}{(1-\theta)^2}\{[(1+\theta) - \alpha(1-\theta)^3\theta^c(2c+1)] - [1 - \alpha(1-\theta)^2\theta^c]^2\} \tag{21}$$

solving (17) and (18) after some manipulation, we deduce

$$\theta^2\left[\mu'_{2(M)}-\mu'_{1(M)}(2c+1)\right]-2\theta\left\{\mu'_{2(M)}-\mu'_{1(M)}(2c+1)-(c+1)\right\}+\left[\mu'_{2(M)}-\mu'_{1(M)}(2c+1)+2c\right]=0 \quad (22)$$

The solution of the above quadratic equation is given by

$$\theta=\frac{-B\pm\sqrt{D}}{2A} \quad (23)$$

5. Bayes Estimation

In Bayesian analysis, we treat the parameter θ as a random variable, with a given probability density function for θ. Bayes theorem leads to posterior distribution of unknown parameters given by the data, all inferential problems concerning the parameters can be solved by means of posterior distributions. Based on these distributions we will estimate the unknown parameters where the expectation is taken over the joint distribution of θ and x.

Here we have considered the informative prior for both the unknown parameters α and θ as,

$$\pi_1(\alpha)\;=g_1(\alpha)=\frac{1}{\beta_{(a,\;b)}}\alpha^{a-1}(1-\alpha)^{b-1} \qquad a,b>0,\;0<\alpha<1 \quad (24)$$

$$\pi_2(\theta)\;=g_2(\theta)=\frac{1}{\beta_{(p,\;q)}}\theta^{p-1}(1-\theta)^{q-1} \qquad p,q>0,\;0<\theta<1 \quad (25)$$

The Bayes estimate of θ is given by

$$\hat{\theta}_B \qquad =\int_0^1\theta\pi(\,\theta|x)d\theta=\frac{\sum_{i=0}^{n_c}\binom{n_c}{i}\beta_{(a+i,\;n_{c+1}+b)}\;\beta_{(n(\bar{x}-1)+p+i+1,\;n+q)}}{\sum_{i=0}^{n_c}\binom{n_c}{i}\beta_{(n+q,\;n(\bar{x}-1)+p+i)}\,\beta_{(a+i,\;n_{c+1}+b)}} \quad (26)$$

And the Bayes estimate of α is

$$\hat{\alpha}_B \qquad =\int_0^1\alpha\pi(\,\alpha|x)d\alpha=\frac{\sum_{i=0}^{n_c}\binom{n_c}{i}\,\beta_{(n+q,\;n(\bar{x}-1)+p+i)}\,\beta_{(a+i+1,n_{c+1}+b)}}{\sum_{i=0}^{n_c}\binom{n_c}{i}\beta_{(n+q,\;n(\bar{x}-1)+p+i)}\,\beta_{(a+i,n_{c+1}+b)}} \quad (27)$$

6. Real Life Application

To illustrate the practical application of results obtained in this paper, data from Singh and Yadav's (1971), classical example on the number of households (f) having at least one migrant according to the number of migrants (X) has been suitable altered. For the purpose of this illustration it has been assumed that ten of the records which should have shown two migrants each were in error by reporting on emigrant. Both the original and the altered data for this example are given in Table 6.1. For the altered data, we fit the misclassified size biased log series distribution. The maximum likelihood estimation is used for estimating the parameters.

Table 6.1: Number of households (f) having at least one migrant according to the number of migrants (X) (Singh and Yadav (1971)

No. of Migrants X	Original Data	Altered Data - Misclassified Size-Biased Log Series Distribution			
		O_i	E_i	χ^2	Estimates of parameters
1	375	385	385.7	0.00127	$\theta = 0.346927$ $\alpha = 0.003154$
2	143	133	133.3	0.000675	
3	49	49	46.4	0.14569	
4	17	17	16.1	0.050311	
5	2	2	5.6	0.735294	
6	2	2	2.0		
7	1	1	0.7		
8	1	1	0.2		
Total	590	590	590	0.93324	
d. f.	2		χ^2	0.93324	

The results are given in Table 6.1, which shows that the altered data fits good with misclassified size biased log series distribution, since ${\chi^2}_{cal} = 0.93324 < {\chi^2}_{(2,0.05)} = 5.991$

Original data fits good with size – biased log series distribution,

since ${\chi^2}_{cal} = 2.305162 < {\chi^2}_{(3,0.05)} = 7.841$

7. **Simulation:**

To study the behavior of the estimates of the parameters we consider a simulation study. Here, we have generated 1000 random samples of size n by using the method of Monte Carlo simulation with different sample size (n), θ *and* α and value of $c = 1$ from the MSBLSD defined in equation (17) and obtained the simulated risk (SR) of estimators of α and θ obtained by the Method of MLE, Method of Moments and Bayes estimation. The simulated results are shown in the following tables. In the Bayes estimation the hyper parameter β is taken as 0.5. The simulated risk is defined as

$$SR = \sqrt{\frac{\sum_{i=1}^{1000}(\hat{\theta}_i - \theta)^2}{1000}}$$

Note:

- The Simulated Risk and Bias of ML, Moment and Bayes estimators are shown in the following tables for different choice of parameters θ, α and sample size n.
- In each table, in each cell the first entry denote the simulated risk (SR) and the entry in the bracket denote Bias of the estimator.

Table 7.1: Simulated Risk and Bias of ML, Moment and Bayes estimators for different values of α and θ for $n = 10, 20$ & 30

		ML		Moments		Bayes	
$n = 10$							
θ	α	(θ)	(α)	(θ)	(α)	(θ)	(α)
0.3	0.01	0.399004 (0.0721172)	0.59724 (0.505499)	0.1938821 (0.0164887)	1.445537 (0.3572701)	0.084188062 (0.011052805)	0.418903518 (0.407770892)
	0.02	0.4016862 (0.0733394)	0.6149005 (0.5296398)	0.1932705 (0.0195291)	1.442425 (0.3532178)	0.085808701 (0.010932404)	0.418002172 (0.405953239)
	0.05	0.3756781 (0.0733394)	0.5930687 (0.4996398)	0.1953939 (0.0090503)	1.453089 (0.3380325)	0.086273056 (0.014165747)	0.384037316 (0.373251748)
	0.1	0.3658907 (0.0729163)	0.5641046 (0.4497586)	0.1936655 (0.0176213)	1.439605 (0.2910491)	0.085937961 (0.020990539)	0.341060892 (0.328482234)
0.5	0.01	0.4955966 (0.0209823)	0.5046846 (0.4146149)	0.1631581 (0.0359248)	1.818949 (0.3756844)	0.109537693 (0.057554603)	0.391354281 (0.379243603)
	0.02	0.5039659 (0.0099472)	0.503231 (0.4061625)	0.1621508 (0.0318394)	1.792378 (0.3604621)	0.116620568 (0.067815989)	0.381769218 (0.370226305)
	0.05	0.4727896 (0.0143286)	0.4763728 (0.3715718)	0.1597018 (0.0211269)	1.828251 (0.3555594)	0.114399639 0.066639364	0.359815949 (0.346913379)
	0.1	0.4675834 (0.0103450)	0.4656599 (0.3509191)	0.1611629 (0.0271445)	1.833101 (0.2906393)	0.115896887 (0.065641195)	0.316176307 (0.301413858)
0.7	0.01	0.7502842 (0.0104145)	0.5221298 (0.4236429)	0.11615 (0.046483)	3.386369 (0.4734417)	0.117128902 (0.081138921)	0.403164628 (0.392388006)
	0.02	0.6466744 (0.0085639)	0.5606665 (0.450756)	0.1109011 (0.0288766)	3.511884 (0.4115277)	0.116049538 (0.083357882)	0.394239658 (0.382547181)
	0.05	0.6600482 (0.0089921)	0.5417621 (0.4233661)	0.1136934 (0.0387534)	3.339942 (0.3521471)	0.11984195 (0.083811951)	0.370519575 (0.357491744)
	0.1	0.6218842 (0.0149871)	0.496126 (0.3702069)	0.1112477 (0.0295937)	3.631248 (0.3462837)	0.123250056 (0.087420714)	0.33363785 (0.318916643)
$n = 20$							
0.3	0.01	0.3017328 (0.038132)	0.4027729 (0.332237)	0.1371913 (0.0171345)	1.023906 (0.3851617)	0.066369176 (0.004659367)	0.363098283 (0.351466408)
	0.02	0.3059354 (0.0282135)	0.3984518 (0.3222679)	0.1380428 (0.0111014)	1.01873 (0.3994549)	0.068034692 (0.002308321)	0.356981642 (0.344837199)
	0.05	0.3141293 (0.0239377)	0.3839981 (0.2936172)	0.1384097 (0.0085895)	1.028735 (0.327202)	0.067387491 (0.002976757)	0.337793579 (0.323183906)
	0.1	0.2915172 (0.0286663)	0.3576095 (0.2626598)	0.1391457 (0.0038261)	1.019742 (0.3125257)	0.068466726 (0.000451487)	0.304253024 (0.376933059)
0.5	0.01	0.5543573 (0.0109824)	0.3533887 (0.2849317)	0.1083589 (0.0056666)	1.292296 (0.3906668)	0.078920319 (0.033207834)	0.333135586 (0.320584019)
	0.02	0.5206028 (0.0006753)	0.3547063 (0.2872428)	0.1076437 (0.0101457)	1.320054 (0.3582872)	0.076335115 (0.027420372)	0.326307976 (0.312724303)
	0.05	0.4944243 (0.0079398)	0.3457509 (0.2719096)	0.1089601 (0.002216)	1.289081 (0.363291)	0.077827582 (0.030955732)	0.303428347 (0.288467687)
	0.1	0.446305 (0.0074838)	0.3282745 (0.2319045)	0.1087706 (0.0032778)	1.289637 (0.3219843)	0.077202708 (0.368965322)	0.26805287 (0.248273247)
0.7	0.01	0.6925815 (0.006641)	0.3760264 (0.3038586)	0.07338303 (0.0058169)	2.609863 (0.4597519)	0.068739925 (0.041136575)	0.348968152 (0.335700383)
	0.02	0.6969855 (0.0039239)	0.3709635 (0.2877676)	0.07586689 (0.0176246)	2.469913 (0.4307284)	0.068556664 (0.040268433)	0.34130748 (0.326854746)
	0.05	0.6681908 (0.0040931)	0.3754791 (0.2741693)	0.07448379 (0.011184)	2.532471 (0.424473)	0.072951682 (0.044918847)	0.313116883 (0.297559452)
	0.1	0.6081768 (0.0067052)	0.3356357 (0.2245442)	0.07464357 (0.0117768)	2.533962 (0.3518326)	0.073152042 (0.043507707)	0.289063711 (0.270658219)
$n = 30$							

0.3	0.01	0.3327296 (0.0188216)	0.3289334 (0.2629785)	0.1110675 (0.0171345)	0.8264615 (0.3382358)	0.05863135 (0.011734557)	0.335270119 (0.321160843)
	0.02	0.3091901 (0.0212382)	0.3379346 (0.2720422)	0.1111489 (0.0111014)	0.8257576 (0.3489639)	0.061403364 (0.009607297)	0.32815521 (0.313652735)
	0.05	0.3047626 (0.0226437)	0.3285323 (0.2541733)	0.1110496 (0.0085895)	0.8320083 (0.2882102)	0.060002821 (0.009506297)	0.301729145 (0.286743802)
	0.1	0.2803672 (0.0217409)	0.3084224 (0.2121447)	0.1121346 (0.0038261)	0.8248268 (0.2810123)	0.060999484 (0.008661163)	0.271686863 (0.253896767)
0.5	0.01	0.4931183 (0.0058786)	0.3037076 (0.2422074)	0.08757983 (0.0119872)	1.055981 (0.3635401)	0.06310101 (0.017106295)	0.297403756 (0.283533027)
	0.02	0.468285 (0.0073714)	0.2808079 (0.2179854)	0.08795311 (0.0091758)	1.042197 (0.3322786)	0.063315374 (0.018589616)	0.293166348 (0.279538016)
	0.05	0.462062 (0.004135)	0.2755829 (0.2015636)	0.08887982 (0.0024347)	1.032292 (0.320333)	0.061225508 (0.017098337)	0.273475658 (0.25763486)
	0.1	0.38625 (0.0028078)	0.2558371 (0.166174)	0.0889461 (0.0020503)	1.038967 (0.2525797)	0.060723208 (0.014270931)	0.241177089 (0.220600659)
0.7	0.01	0.7083158 (0.0004318)	0.3125612 (0.2442503)	0.06052437 (0.0099964)	2.019973 (0.4055967)	0.053872896 (0.029786658)	0.300551562 (0.28703036)
	0.02	0.674432 (0.0075)	0.3013778 (0.2326161)	0.0602727 (0.0083755)	2.050203 (0.3804104)	0.055185621 (0.028565419)	0.300022394 (0.285417776)
	0.05	0.6183718 (0.0020838)	0.2719811 (0.2048467)	0.06006786 (0.0074012)	2.043279 (0.3581949)	0.052551783 (0.02831192)	0.28127253 (0.265294117)
	0.1	0.5386516 (0.0061761)	0.276865 (0.1769128	0.06093138 (0.0122031)	2.017987 (0.3405125)	0.054291198 (0.028807652)	0.254014667 (0.233518356)

Conclusions:

In this paper we have discussed the problem of misclassification in a Size Biased Log Series distribution. We have also applied MSBLSD in real life example and showed that MSBLSD has a good fitting for the altered data. It is clear from Table 7.1 to 7.3 that for any sample size, for given value of θ and α, Bayes estimate performs better than Maximum Likelihood estimation and Method of Moments estimators since simulated risk of the estimates are smaller than that of ML and MOM estimates. Also as n increases simulated risk of Bayes and MOM estimates decreases, but in MLE the pattern do not follow in certain cases.

Bibliography

Best, D.J., Rayner, J.C.W. & Thas, O. (2008), "Test of fit for the logarithmic distributions". *Journal of Applied Mathematics and Decision Sciences*, vol 2008, 1-8

Chatfield, C., Ehrenberg, A. S. C. & Goodhardt, G. J. (1966), Progress on a simplified model of stationary purchasing behavior (with discussion), *Journal of The Royal Statistical Society*, Series A, 129, 317-367

Cressie, N. & Borkent, M. (1986), "The moment generating function has its moments". *Journal of American Statistical Association*, vol. 13, 337-344

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943), The relation between the number of species and the number of individuals in a random sample of an animal population., *J. Animal Ecology* 12, 42-58

Gupta, R.C. (1974), "Modified power series distribution and some of its applications", *Sankhya*, vol.36, 288-298

Johnson, N. L., Kotz, S. & Kemp, A. W. (2005), "Univariate discrete distributions", 3rd ed. *John Wiley & Sons inc*; Hoboken, New Jersey

Kempton, R. A. (1975), "A generalized form of Fisher's logarithmic series", *Biometrika*, vol. 62, 29-38

Patil, G. P. (1962), Some methods of estimation for the logarithmic series distribution, *Biometrics* 18, 68-75

Patil, G. P. & Ord, J. K. (1975), On size-biased sampling and related form-invariant weighted distributions, *Sankhya*, 38, 48-61

Rao, C. R. (1965), On discrete distributions arising out of ascertainment, In: Classical and Contagious discrete distribution; G.P. Patil (ed.), *Pergamon press and Statistical Publishing Society*, Calcutta, 302-332

Sadinle, M. (2008), "Linking the negative binomial and logarithmic series distributions via their associated series". *Revista Colombiana de Estadistica*, vol. 31, pp. 311-319

Shanumugam, R. & Singh J. (1984), "A characterization of the logarithmic series distribution and its applications." *Communications in Statistics*, vol. 13, 865-875

Singh, S.N. & Yadav, R.C. (1971), Trends in rural out-migration at household level. *Rural Demography*.8, 53 – 61

Trivedi, B. S. & Patel, M. N. (2013), "Estimation in Misclassified Size-biased Generalized Negative Binomial Distribution", *Mathematics and Statistics",* Horizon Research Publishing Corporation, Vol 1(2), 74-85

Williams, C.B. (1947), The logarithmic distribution and its applications to biological problems, J.*Ecology*34, 253-271

B. S. Trivedi (MSc, MPhil), is a doctoral student in the Department of Statistics at Gujarat University, Ahmedabad, Gujarat, India. She is working on misclassified size-biased discrete distribution. Currently, she is an Assistant Professor at H. L. Institute of Commerce, Ahmedabad University.

M. N. Patel (MSc, PhD), is a Professor of Statistics at the Department of statistics, Gujarat University, Ahmedabad, Gujarat, India. He has published 60 research papers in journals of National and International repute. His research interests are in statistical inference, life testing and reliability theory and statistical distributions. His E-mail id is mnpatel.stat@gmail.com