INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY EDUCATIONAL RESEARCH
ISSN : 2277-7881; IMPACT FACTOR - 2.735; IC VALUE:5.16
VOLUME 3, ISSUE 6(4), JUNE 2014

ESTIMATION OF PARAMETER IN A NEW DISCRETE DISTRIBUTION ANALOGOUS TO BURR DISTRIBUTION

T. Raveendra Naika
Maharani's Science College for Women, Bangalore, India

Abstract

Estimation of parameter in a new discrete distribution which is analogous to Burr distribution is discussed in this paper. The maximum likelihood and the method of moment estimators are obtained. The asymptotic normality of the moment estimator is established. The asymptotic relative efficiency of the maximum likelihood estimator over the moment estimator is computed.

Keywords: Maximum Likelihood and Moment Estimators, Fisher Information, Asymptotic Relative Efficiency.

INTRODUCTION

sreehari (2010) has characterized a class of discrete distributions which turns out to be the discrete analogue of Burr (1942) family. The probability mass function (pmf) of the random variable X of the d-th distribution in the class characterized by Sreehari (2010) is

$$p(x,\theta) = \begin{cases} (x+1-\theta)\dfrac{\theta^x}{(x+1)!}, x = 0,1,2,\ldots, 0<\theta<1 \\ 0, \ otherwise. \end{cases} \qquad (1.1)$$

The distribution function of X is

$$F(x) = \begin{cases} 0, \ x<0 \\ 1 - \dfrac{\theta^{[x+1]}}{([x+1])!}, \ x \geq 0 \end{cases} \qquad (1.2)$$

We refer to this distribution as *S(d)* distribution. The probability generating function (pgf) of X is

$$f(s) = E(s^X)$$

$$= \frac{1-(1-s)e^{\theta s}}{s}, \ 0 < s \leq 1.$$

The mean and the variance of X are respectively

$$E(X) = f'(1)$$

$$= e^{\theta} - 1 \quad \text{and}$$

$$Var(X) = f''(1) + f'(1) - (f'(1))^2$$

$$= e^{\theta}(2\theta - e^{\theta} + 1).$$

It can easily be seen that *E(X) > Var(X)* and hence *S(d)* – distribution is under dispersed. The pmf of this distribution is similar to that of Poisson distribution in structure but its mean and variance are equal. Therefore S(d) distribution can be a suitable model for the data exhibiting under dispersion and Poisson is not a good fit.

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

If $\underline{X} = (X_1, X_2, \ldots, X_n)$ is a random sample on X having the pmf specified in (1.1), then the likelihood function becomes

$$L(\theta|x) = \prod_{j=1}^{n} P(X = x_j)$$

$$L(\theta|x) = \prod_{j=1}^{n} (x_j + 1 - \theta)\frac{\theta^{x_j}}{(x_j+1)!}.$$

The log-likelihood function is

$$\log L(\theta|x) = \sum_{j=1}^{n} \log(x_j + 1 - \theta) + \sum_{j=1}^{n} x_j \log\theta - \text{constant}.$$

The likelihood equation is

$$\frac{d \log L(\theta|x)}{d\theta} = -\sum_{j=1}^{n} \frac{1}{x_j + 1 - \theta} + \frac{\sum_{j=1}^{n} x_j}{\theta}.$$

The maximum likelihood estimate of θ is the solution of

$$\sum_{j=1}^{n} \frac{\theta}{x_j + 1 - \theta} - n\bar{x} = 0$$

where $\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j$. It is evident that this likelihood equation does not yield a closed form expression for the maximum likelihood estimate (MLE) of θ. Hence a numerical procedure like Newton-Raphson method can be employed to compute the MLE $\hat{\theta}_{mle}$.

FISHER INFORMATION

When X has the pmf specified in (1.1),

$$\log p(x,\theta) = \log(x+1-\theta) + x\log\theta - \log(x+1)!$$

Also $$\frac{d\log p}{d\theta} = \frac{-1}{x+1-\theta} + \frac{x}{\theta}$$

and $$\frac{d^2\log p}{d\theta^2} = \frac{-1}{(x+1-\theta)^2} - \frac{x}{\theta^2}.$$

Hence

$$E\left[\frac{d^2\log p}{d\theta^2}\right] = -\sum_{x=0}^{\infty}\frac{1}{(x+1-\theta)^2}p(x) - \sum_{x=0}^{\infty}\frac{x}{\theta^2}p(x)$$

$$= -\sum_{x=0}^{\infty}\frac{1}{(x+1-\theta)}\frac{\theta^x}{(x+1)!} - \frac{1}{\theta}\sum_{x=0}^{\infty}\frac{\theta^{x+1}}{(x+1)!}$$

$$E\left[\frac{d^2\log p}{d\theta^2}\right] = -\sum_{x=0}^{\infty}\frac{1}{(x+1-\theta)}\frac{\theta^x}{(x+1)!} - \frac{1}{\theta^2}(e^\theta - 1).$$

The Fisher information becomes

$$I(\theta) = -E\left[\frac{d^2\log p}{d\theta^2}\right]$$

$$= \sum_{x=0}^{\infty}\frac{1}{(x+1-\theta)}\frac{\theta^x}{(x+1)!} + \frac{1}{\theta^2}(e^\theta - 1).$$

The infinite series evidently converges for all values of θ in (0, 1) but its sum is not tractable. Hence it can be numerically computed correct to desired number of decimal places.

Note that

i) the support S = $\{x: p(x,\theta) > 0\}$ does not depend on the parameter θ.
ii) the parameter space (0, 1) is an open set.
iii) log $p(x, \theta)$ can be differentiated twice w.r.t. θ.
iv) $\sum_{x=0}^{\infty} p(x,\theta) = 1$ is twice differentiable under summation sign.

Therefore $p(x, \theta)$ satisfies the regularity conditions of Cramer (1966) and it belongs to Cramer family. Hence if $\underline{X} = (X_1, X_2, \ldots, X_n)$ is a random sample on X having the pmf specified in (1.1) and $\hat{\theta}_{mle}$ is the MLE of θ, then

$$\sqrt{n}\left(\hat{\theta}_{mle} - \theta\right) \xrightarrow{L} N\left(0, \frac{1}{I(\theta)}\right), \text{ as } n \to \infty.$$

That is $\hat{\theta}_{mle}$ is consistently and asymptotically normal (CAN) for θ with the asymptotic variance $\frac{1}{I(\theta)}$.

METHOD OF MOMENT ESTIMATION (MME)

When $\underline{X} = (X_1, X_2, \ldots, X_n)$ is a random sample on X having the pmf specified in (1.1), the moment estimator of θ is the solution of the equation $e^\theta - 1 = \bar{X}_n$. Hence the moment estimator of θ is

$$\hat{\theta}_{mme} = \log(\bar{X}_n + 1).$$

Since the pmf of X does not belong to the exponential family, we need to establish the asymptotic normality of the moment estimator separately. The following theorem states the asymptotic normality of the moment estimator.

Theorem: If $\underline{X} = (X_1, X_2, \ldots, X_n)$ is a random sample on X having the pmf specified in (1.1), then

$$\sqrt{n}\left(\hat{\theta}_{mme} - \theta\right) \xrightarrow{L} N\left(0, e^\theta(2\theta - e^\theta + 1)\frac{1}{(\theta+1)^2}\right), \text{as } n \to \infty.$$

Proof:

Since $X_1, X_2, \ldots, X_n$ are iid with $E(X) = e^\theta - 1$ and $Var(X) = e^\theta(2\theta - e^\theta + 1) < \infty$, by Levy - Lindeberg central limit theorem

$$Z_n = \sqrt{n}\left(\bar{X}_n - (e^\theta - 1)\right) \xrightarrow{L} N\left(0, e^\theta(2\theta - e^\theta + 1)\right), \text{as } n \to \infty.$$

Take $g(x) = \log(x+1)$. Then $g'(x) = \frac{1}{x+1}$ is non-vanishing and continuous for $0 < x < 1$. Hence by Mann-Wald theorem stated in Mukhopadhaya (2000),

$$\sqrt{n}\left(\log(\bar{X}_n + 1) - \theta\right) \xrightarrow{L} N(0, \frac{e^{\theta}(2\theta - e^{\theta} + 1)}{(\theta+1)^2}), \ as \ n \to \infty.$$

That is $\hat{\theta}_{mme}$ is CAN with the asymptotic variance $e^{\theta}(2\theta - e^{\theta} + 1)\frac{1}{(\theta+1)^2}$.

The asymptotic relative efficiency of the MLE over the MME is given by

$$ARE = \frac{Asymptotic\ variance\ of\ MME}{Asymptotic\ variance\ of\ MLE}.$$

Since the asymptotic variance $\frac{1}{I(\theta)}$ of the MLE does not have a closed form expression, it is computed for various values of θ and the ARE is shown in the following table.

θ	0.1	0.2	0.3	0.4	0.5
ARE	1.3405	2.3291	3.0257	3.8895	4.934
θ	0.6	0.7	0.8	0.9	0.95
ARE	6.1455	6.1452	7.1479	8.6441	9.0807

The ARE of the MLE over the MME is uniformly greater than unity and therefore the MLE is asymptotically more efficient than the MME.

ILLUSTRATION

Sreehari (2010) has compared the fit of S(d) and Poisson models for a clinical trial data set. A bioequivalence study was conducted for a test drug (T) and a reference drug (R) by administering them to 144 individuals using a two period, two sequence and two treatment crossover design. The time until maximum concentration (T_{max}) was one of the characteristics observed. Let T_{it} and T_{ir} respectively denote the T_{max} values corresponding to the i-th individual for the test and the reference drugs. Take $D_i = |T_{it} - T_{ir}|$. The individuals were administered the drugs and their blood concentrations were measured just prior to medication and at time points (in hours) 1, 2, 4, 6, 8, 10, 12, 14, 16, 20, 22, 24, 30, 48, 60, 72, 96, 120 after medication. Of the 144 individuals 7 did not complete the course of treatment. The observed values of D_i for the 137 individuals are shown in the following table.

x	0	1	2	3	Total
Frequency	75	46	14	2	137

For this observed distribution, mean = 0.583942 > variance = 0.534925 and it exhibits under dispersion. And $\hat{\theta}_{mme} = \log(\bar{x}_n + 1) = 0.4599$. Further,

x	Observed frequency	Expected frequency	
		S(d)	Poisson
0	75	73.9915	76.4043
1	46	48.5192	44.6157
2	14	12.268	13.0265
3	2	2.22129	2.95354
Chi-square value		**0.41111**	**0.64342**

It is evident that S(d) fits better than the Poisson model to this observed distribution. But *S(d)* model cannot be fitted to all under dispersed data. We have $0 < \theta < 1$ and hence $0 < E(X) < e-1$. If $0 < \bar{x} < e-1$ is violated, then it leads to $\hat{\theta}_{mme} = \log(\bar{x}_n + 1) > 1$ which is inadmissible. For example, consider the data on the number of scintillations from radioactive decay of Polonium reported by Rutherford and Geiger (1910). This data set has been reproduced in Santner and Duffy (1989).

x	0	1	2	3	4	5	6	
f	57	203	383	525	532	408	273	
x	7	8	9	10	11	12	13	14
f	139	45	27	10	4	0	1	1

(x: number of scintillations, f: frequency)

For the observed data, mean = 3.871549 > variance = 3.694773. This exhibits under dispersion. But $\hat{\theta}_{mme} = \log(\bar{x}_n + 1) = 1.58341 > 1$, which is inadmissible.

SIMULATION

Random observations on *X* can be simulated using the distribution function (1.2). A modest simulation study has been carried out to study the performance of the MLE and the MME of θ. Using R, 1000 samples of size 50, 100, 150 were simulated for the specified value of $\theta = 0.4$ and the estimates were computed. The MLEs were obtained by solving the likelihood equation by Newton-Raphson method and the MMEs were taken as the initial estimates. The histograms of the estimates are displayed in Figures 6.1-6.2.

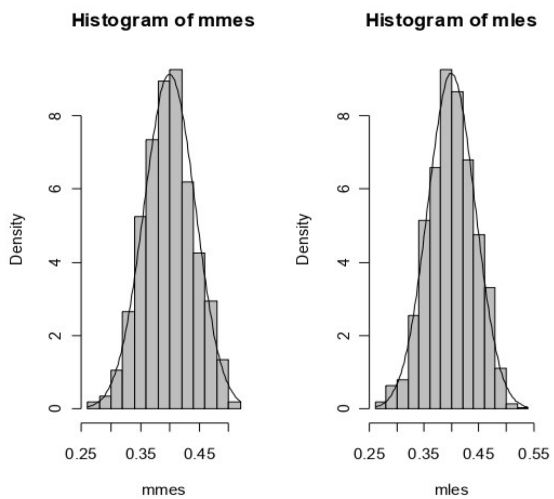

Figure 6.1: Histogram of the MMEs and the MLEs of θ based on 1000 sample of size 100 for θ=0.4.

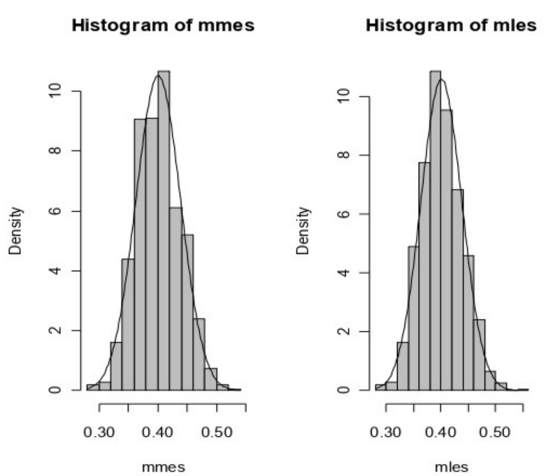

Figure. 6.2: Histogram of the MMEs and the MLEs of θ based on 1000 sample of size 150 for θ=0.4.

These histograms give graphical evidence for the asymptotic normality of both the estimates. But the MLE approaches normality faster than the MME.

CONCLUSION

The S(d) distribution is an appropriate alternate to Poisson model when the observed data exhibit under dispersion. Though the maximum likelihood estimator of the parameter of S(d) has no closed form expression, it can easily be computed by Newton-Raphson method. The moment estimator of the parameter has a closed form expression and easy to compute. Both the estimators are asymptotically normal. When computer facility is available, the MLE can be preferred to the MME since the former approaches normality faster than the later.

ACKNOWLEDGEMENT

The authors are grateful to Prof. M. Sreehari for sending a softcopy of his paper and to Professors K. Suresh Chandra and B. Chandrasekar for useful comments.

REFERENCES

Burr, I.W. 1942. Cumulative frequency functions, USA, Annals of Mathematical Statistics, pp.215-232.

Cramer, H. 1966. Mathematical Methods of Statistics, Princeton, 11th Print, Princeton University Press.

Kale, B.K. 1999. A First Course on Parametric Infe rence, New Delhi, Narosa Publishing House.

Mukhopadhyay, N. 2000. Probability and Statistical Inference, New York, Marcel Dekker, pp-261.

Rao, C.R. 1973. Linear Statistical Inference and Its Applications, New York , John Wiley.

Rutherford, E & Geiger, H. 1910. The probability variations in the distribution of α particles, Phil, Mag. Sixth Ser., 20, pp. 698 – 704.

Santner, T.J. & Duffy, D.E. 1989. The Statistical Analysis of Discrete Data, New York, Springer Verlag.

Sreehari,M. 2010. On a class of discrete distributions analogous to Burr family, Pune, Journal of Indian Statistical Association, Vol. 48, No. 1