Aligarh Journal of Statistics
Vol. 30 (2010), 15-22

A SIZE - BIASED AND ZERO -TRUNCATED GENERALIZED GEOMETRIC SERIES DISTRIBUTION AND THEIR APPLICATIONS

Anwar Hassan and Bilal Ahmad Bhat

ABSTRACT

In this paper, a size-biased generalized geometric series distribution $(SBGGSD)$ has been defined. The moments and recurrence relation of $SBGGSD$ have also been obtained. The estimation of parameters of $SBGGSD$ and Zero truncated geometric series distribution $(ZTGGSD)$ has been discussed. The distributions have been fitted some observed sets of data to test their goodness of fit.

1. INTRODUCTION

Mishra (1982) using the results of the Lattice path analysis obtained a two parameter generalized geometric serried distribution $(GGSD)$ given by its probability function

$$\Pr(X=x)=\frac{\Gamma(\beta x+1)}{x!\,\Gamma(\beta x-x+2)}\alpha^{x}(1-\alpha)^{1+\beta x-x}, \qquad x=0,1,2,\ldots$$

$$0<\alpha<1, \quad |\alpha\beta|<1. \qquad (1.1)$$

At $\beta=0$ this distribution (1.1) reduces to Bernoulli distribution and is a particular case of the Jain and Consul's (1971) generalized negative binomial distribution $(GNBD)$ in the same way as the geometric series distribution is a particular case of negative binomial distribution. It can be seen that geometric series distribution (GSD) is a particular case of (1.1) and can be obtained at $\beta=1$.

The interesting properties and applications of the distribution have been studied by Mishra (1982), Singh (1989), Mishra and Singh (1992), Hassan (1995) and Hassan *et al.* (2002, 2003, 2007, 2008 and 2009). They investigated and found this distribution (1.1) to provide much closer fits to all those observed distributions where the geometric distribution and the various compound geometric series distributions have been fitted earlier by many authors.

The moments of the $GGSD$ (1.1) about origin and variance have been obtained as

$$\mu_1' = \frac{\alpha}{(1-\alpha\beta)} \tag{1.2}$$

$$\mu_2' = \frac{\alpha(1-\alpha)}{(1-\alpha\beta)^3} + \frac{\alpha^2}{(1-\alpha\beta)^2} \tag{1.3}$$

$$\mu_3' = \frac{\alpha^3}{(1-\alpha\beta)^3} + \frac{3\alpha^2(1-\alpha)}{(1-\alpha\beta)^4} + \frac{\alpha(1-\alpha)}{(1-\alpha\beta)^5}(1-2\alpha+\alpha\beta(2-\alpha)) \tag{1.4}$$

$$\mu_2 = \frac{\alpha(1-\alpha)}{(1-\alpha\beta)^3}\,. \tag{1.5}$$

TRUNCATED *GGSD*

A zero-truncated form of the *GGSD* (1.1) is defined as

$$\Pr_1(X=x) = \frac{\Gamma(\beta x+1)}{x!\;\Gamma(\beta x - x + 2)}\alpha^{x-1}(1-\alpha)^{1+\beta x - x}, \quad x = 0,1,2,\ldots$$

$$0<\alpha<1, \quad |\alpha\beta|<1. \tag{1.6}$$

and its mean and variance are

$$\text{Mean} = \mu_1'(z) = \frac{1}{1-\alpha\beta} \tag{1.7}$$

$$\mu_2'(z) = \frac{\alpha}{(1-\alpha\beta)^2} + \frac{1-\alpha}{(1-\alpha\beta)^3} \tag{1.8}$$

$$\text{Variance} = \mu_2(z) = \frac{\alpha\beta(1-\alpha)}{(1-\alpha\beta)^3}\,. \tag{1.9}$$

The method of maximum likelihood estimation (MLE) of the parameters α and β is not simple to apply to the two equations for (1.1) and (1.6). The MLE equations are not found to be solved directly. The estimates given in (1.10) and (1.11) are obtained by the method of moments.

$$\alpha^* = \bar{x}/(1+\beta\bar{x}) \tag{1.10}$$

$$\alpha^{**} = (\bar{x}-1)/\beta\bar{x}\,. \tag{1.11}$$

In this paper, a size-biased *GGSD* has been obtained. The moments of the distribution have also been obtained. The estimation of parameters of *SBGGSD* and *ZTGGSD* has been obtained. The method of moments has been used as a simple alternative to the *MLE* and other methods. The distributions have also been fitted to some observed sets of data to test their goodness of fit.

2. A SIZE - BIASED GGSD

A size-biased *GGSD* is obtained by taking the weights of the *GGSD* (1.1) as x. We have from (1.2) to (1.5)

$$\sum_{x=0} x\,\Pr(X=x)=\frac{\alpha}{1-\alpha\beta}$$

and so $\frac{(1-\alpha\beta)}{\alpha}x\Pr(X=x)$ represents a probability distribution. This gives the size-biased *GGSD* as

$$\Pr_2(X=x)=\frac{(1-\alpha\beta)\,\Gamma(1+\beta x)}{(x-1)!\;\Gamma(\beta x-x+2)}\alpha^{x-1}(1-x)^{1+\beta x-x}, \qquad x=0,1,2,\ldots$$

$$0<\alpha<1, \quad |\alpha\beta|<1. \tag{2.1}$$

The $r-$th moment of the size-biased *GGSD* is obtained as

$$\mu_r'(s)=E(X^r)=\sum_{x=1} x^r \Pr_2(X=x)$$

$$=\frac{1-\alpha\beta}{\alpha}\sum_{x=0} x^{r+1}\Pr(X=x)$$

$$\mu_r'(s)=\frac{1-\alpha\beta}{\alpha}(r+1)-\text{th moment about origin of the } GGSD \text{ (1.1)}. \tag{2.2}$$

The moments of size-biased *GGSD* (2.1) are obtained from (1.3) and (1.4) using (2.2) as

$$\text{Mean}=\mu_1'(s)=\frac{\alpha}{(1-\alpha\beta)}+\frac{(1-\alpha)}{(1-\alpha\beta)^2} \tag{2.3}$$

$$\mu_2'(s)=\frac{\alpha^2}{(1-\alpha\beta)^2}+\frac{3\alpha(1-\alpha)}{(1-\alpha\beta)^3}+\frac{(1-\alpha)}{(1-\alpha\beta)^4}[1-2\alpha+\alpha\beta(2-\alpha)] \tag{2.4}$$

$$\text{Variance}=\mu_2(s)=\frac{\alpha(1-\alpha)}{(1-\alpha\beta)^3}+\frac{\alpha(1-\alpha)}{(1-\alpha\beta)^4}[\beta(2-\alpha)-1] \tag{2.5}$$

The higher moments can be obtained similarly using (2.2) if so desired.

The recurrence relationship of size-biased $GGSD$ about origin are obtained as

$$\mu'_{r+1}(s) = \frac{\alpha(1-\alpha)}{(1-\alpha\beta)}\frac{\partial\mu'_r(s)}{\partial\alpha} + \left[\frac{(1-\alpha)}{(1-\alpha\beta)^2} + \frac{\alpha}{(1-\alpha\beta)}\right]\mu'_r(s),$$

$$r = 1, 2, \ldots \qquad (2.6)$$

3. A METHOD OF MOMENTS OF ESTIMATION OF SIZE-BIASED *GGSD*

Consider the first moment of size-biased $GGSD$ (2.1) about origin is from (2.3) as

$$\mu'_1(s) = \frac{\alpha(1-\alpha\beta) + (1-\alpha)}{(1-\alpha\beta)^2}$$

$$= \frac{\alpha\,\mu'_{-1}(s) + (1-\alpha)}{\mu'^{\,2}_{-1}(s)}, \qquad (3.1)$$

where $\mu'_{-1}(s) = 1 - \alpha\beta$.

$$\mu'_1(s)\,\mu'^{\,2}_{-1}(s) = \alpha\,\mu'_{-1}(s) + (1-\alpha)$$

$$\hat{\alpha} = \frac{\mu'_1(s)\,\mu'^{\,2}_{-1}(s) - 1}{\mu'_{-1}(s) - 1}. \qquad (3.2)$$

From (3.1) we have

$$\hat{\beta} = \frac{1 - \mu'_{-1}(s)}{\hat{\alpha}}. \qquad (3.3)$$

The (3.2) and (3.3) indicate the estimates of the parameters α and β of $SBGGSD$ (2.1).

Replacing $\mu'_1(s)$ and $\mu'_{-1}(s)$ by the corresponding simple values, $\bar{x}$ and m'_{-1} respectively.

$$\hat{\alpha} = \frac{\bar{x}\,m'_{-1} - 1}{\bar{x} - 1} \qquad (3.4)$$

and

$$\hat{\beta} = \frac{1 - m'_{-1}}{\hat{\alpha}}, \qquad (3.5)$$

where $\bar{x} = \dfrac{\Sigma fx}{\Sigma f}$ and $m'_{-1} = \dfrac{\Sigma(f/x)}{\Sigma f}$, f is observed frequency.

4. METHOD OF MOMENTS OF ESTIMATION OF ZERO TRUNCATED *GGSD*

From (1.7) and (1.8) the estimates of α and β are obtained as

$$\frac{\mu_2'(z)}{\mu_1'^2(z)} = \alpha + (1-\alpha)\mu_1'(z) \tag{4.1}$$

$$\hat{\alpha} = \frac{\mu_2'(z)}{(1-\mu_1'(z))\,\mu_1'^2(z)} - \frac{\mu_1'(z)}{1-\mu_1'(z)}. \tag{4.2}$$

From (1.7), we have $1-\alpha\beta = \dfrac{1}{\mu_1'(z)}$, the estimate of β

$$\hat{\beta} = \frac{1}{\hat{\alpha}}\left(1 - \frac{1}{\mu_1'(z)}\right). \tag{4.3}$$

Replacing $\mu_1'(z)$, $\mu_2'(z)$ by the corresponding simple values $\bar{x}$ and m_2' respectively, we get

$$\hat{\alpha} = \frac{m_2'}{(1-\bar{x})\,\bar{x}^2} - \frac{\bar{x}}{(1-\bar{x})} = \frac{1}{(1-\bar{x})}\left(\frac{m_2'}{\bar{x}^2} - \bar{x}\right). \tag{4.4}$$

and

$$\hat{\beta} = \frac{1}{\hat{\alpha}}\left(1 - \frac{1}{\bar{x}}\right), \tag{4.5}$$

where $m_2' = \dfrac{1}{\Sigma f}\Sigma fx^2$, $\bar{x} = \dfrac{\Sigma fx}{\Sigma f}$, f is the observed frequency.

5. GOODNESS OF FIT

We present here the fitting of the size-biased *GGSD* and zero-truncated *GGSD* to the same data set used by Jani and Shah (1981) and Mishra and Singh (1993) to only two other data sets by the moment estimation method. These data sets are regarding migration and publications by authors.

Table 5.1: Number of households (f_x) having at least one migrant according to number of migrants (x) [Singh and Yadev (1971)]

x	f_x	Expected frequency	
		SBGGSD	*ZTGGSD*
1	375	370.32	375.38
2	143	147.70	140.48
3	49	45.95	45.77
4	17	18.19	19.62
5	2	5.23	5.51
6	2	1.74	1.86
7	1	0.54	0.61
8	1	0.18	0.15
Total		590.00	590.00
$\hat{\alpha}$		0.12732	0.37657
$\hat{\beta}$		1.4233	0.93946
df		2	2
χ^2		0.4585	0.8848

Table 5.2: Number of author (f_x) according to number of papers in the Review of Applied Entomology, Vol., 1913 [Williams (1944)]

x	f_x	Expected frequency	
		SBGGSD	*ZTGGSD*
1	285	288.0	284.9
2	70	66.0	70.45
3	32	28.1	28.62
4	10	14.0	10.93
5	4	7.1	8.39
6	3	3.3	3.349
7	3	2.2	1.82
8	1	1.2	1.02
9	2	1.1	0.59
10	1	0.7	0.35
Total	411	411.0	411.0
$\hat{\alpha}$		0.0303	0.037
$\hat{\beta}$		0.8903	0.9093
df		3	3
χ^2		2.613	3.439

It can be seen from the tables that size-biased *GGSD* provides closer fit than that provided by the zero-truncated *GGSD* in both the cases of using the estimates obtained by the method of moments. Moreover, they are relatively very quick to be obtained and so it may be preferred to others where very quick results are required.

Acknowledgement

The authors are grateful to the referee for his/her helpful comments and suggestions that led to this improved version of the paper.

REFERENCES

Hassan, A. (1995): *Problems of Estimation in Lagrangian Probability Distributions.* Unpublished Ph. D. Thesis, Patna University, Patna.

Hassan, A. and Khan, S. (2009): Prediction distribution of generalized geometric series distribution and its different forms. *Pakistan J. Statist.,* **25**, 47-57.

Hassan, A. and Mir, K.A. (2008): On Bayesian estimation of size-biased generalized geometric series distribution and its applications. *J. Statist. Res.*, **42**, 117-126.

Hassan, A., Mir, K.A and Ahmad, M. (2007): On Bayesian analysis of generalized geometric series distribution under different priors. *Pakistan J. Statist.*, **23**, 255-260.

Hassan, A., Mishra, A. and Jan, T.R (2002): A quick method of estimating generalized geometric series distribution. *Studia Sci. Math. Hungar.*, **39,** 291-295.

Hassan, A., Raja, T.A. and. Mishra, A. **(**2003): On estimation of generalized geometric series distribution. *SKUAST (K), Res. J.*, **5**, 55- 60.

Jain, G.C. and Consul, P.C. (1971): A generalized negative binomial distribution; *SIAM J. Appl. Math.*, **21**, 501-513.

Jani, P.N. and Shah, S.M. (1981): The zero-truncated generalized Poisson distribution. *Aligarh J. Statist.*, **1**, 174-182.

Mishra A. and Singh, S.K. (1993): A size-biased Lagrangian Poisson distribution. *J. Bihar Math. Soc.*, **16**, 22-31.

Mishra, A. and Singh, S.K. (1992): On a characterization of the geometric series distribution. *J. Bihar Math. Soc.*, **15**, 32-35.

Mishra, A. (1982): A generalization of geometric series distribution. *J. Bihar Math. Soc*. **6,** 18-22.

Sinha, S.K. (1989): *Some Contributions to the Quasi-binomial and Discrete Lagrangian Probability Distributions,* Unpublished Ph. D. Thesis, Patna University, Patna.

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

Williams, C.B. (1944): Some applications of the logarithmic series and the index of diversity to ecological problems. *Journal of Ecology*, **32**, 1- 44.

Received : 18-10-2006

Revised : 06-11-2009

Anwar Hassan
PG Department of Statistics,
University of Kashmir, Srinagar, India.
e-mail: anwar_hassan2007@hotmail.com ,
anwar.hassan2007@gmail.com

Bilal Ahmad Bhat
Division of Sericulture,
Sher-e-Kashmir University of Agricultural Sciences
and Technology of Kashmir, Srinagar, India.
e-mail: bhat_bilal@rediffmail.com