A NEW DISCRETE FLEXIBLE MODEL: GEETA-KUMARASWAMY DISTRIBUTION WITH APPLICATION

Zahoor Ahmad[1] Adil Rashid[2], and Tariq Rashid Jan[3]

[1,2,3] *P.G Department of Statistics, University of Kashmir, Srinagar(India)*

ABSTRACT

In this paper, we have developed a new and versatile discrete distribution that takes different and flexible functional form for its pmf depending upon the nature of parameter m. The resulting distribution holds several discrete distributions as special for particular parameter setting. Parameter estimation has been performed by employing MLE technique.. Finally the potentiality of proposed distribution is justified by using it to model the real life data set.

***Keywords*: Compound distribution, Factorial moment, Geeta distribution, Kumaraswamy distribution,**

I. INTRODUCTION

From the last few decades researchers are busy to obtain new probability distributions by using different techniques such as compounding, T-X family, transmutation etc. but compounding of probability distribution has received maximum attention which is an innovative and sound technique to obtain new probability distributions. The compounding of probability distributions enables us to obtain both discrete as well as continuous distribution.

Compound distribution arises when all or some parameters of a distributionknown as parent distribution vary according to some probability distribution called the compounding distribution for instance negative binomial distribution can be obtained from Poisson distribution when its parameter λ follows gamma distribution. If the parent distribution is discrete then resultant compound distribution will also be discrete and if the parent distribution is continuous then resultant compound distribution will also be continuous i,e. the support of the original (parent) distribution determines the support of compound distributions.

In several research papers it has been found that compound distributions are very flexible and can be used efficiently to model different types of data sets. With this in mind many compound probability distributions have been constructed. Sankaran (1970) obtained a compound of Poisson distribution with that of Lindley distribution, Zamani and Ismail (2010) constructed a new compound distribution by compounding negative binomial with one parameter Lindley distribution that provides good fit for count data where the probability at zero has a large value. Researchers like Adil and Jan obtained several compound distributions for instance, (2013) a compound of zero truncated generalized negative binomial distribution with generalized beta distribution, (2014a) they obtained compound of Geeta distribution with generalized beta distribution and (2014b) compound of Geeta distribution with generalized beta distribution reecentlyAdiland Jan (2014c) explored a mixture of generalized negative binomial distribution with that of generalized exponential distribution which contains several compound distributions as its sub cases and proved that this particular model is better in comparison to others when it comes to fit observed count data set. Most recently Adil and Jan (2015,) developed a new count data models that can be used as a tool for modeling overdispersion.

II.MATERIAL AND METHODS

2.1 GEETA *DISTRIBUTION (GTD)*

Geeta distribution introduced by Geeta and Shenton (1975) was modified by Islam and Geeta (1990) who derived it as a bunching model in traffic flow through the branching process and also discussed its applications to automobile insurance claims and vehicle bunch size data.

Suppose a queue is initiated with one member and has traffic intensity with binomial arrivals, given by generating function $g(t) = (1 - p + pt)^m$ and constant service time. Then the probability that exactly x members will be served before the queue vanishes is given by Geeta distribution with probability mass function given

$$f_{GTD}(x;\beta,p) = \begin{cases} \dfrac{1}{mx-1}\dbinom{mx-1}{x} p^{x-1}(1-p)^{mx-x} & ; x = 1,2,\ldots \\ 0 & ; otherwise \end{cases} \qquad (2.1)$$

where $0 < p < 1$ and $1 \le m \le \dfrac{1}{p}$. Geeta distribution reduces to the geometric distribution when $m = 1$ in (1).

2.2 *KUMARASWAMY DISTRIBUTION (KSD)*

Kumaraswamy distribution is a two parameter continuous probability distribution that has obtained by Kumaraswamy (1980) but unfortunately this distribution is not very popular among statisticians because researchers have not analyzed and investigated it systematically in much detail. Kumaraswmy distribution is similar to the beta distribution but unlike beta distribution it has a closed form of cumulative distribution function which makes it very simple to deal with. For more detailed properties one can see references Kumaraswamy (1980) and Jones (2009).

A random variable X is said to have a Kumaraswamy distribution (KSD) if its pdf is given by

$$f_2\left(X;\alpha,\beta\right)=\alpha\beta\ x^{\alpha-1}\left(1-x^{\alpha}\right)^{\beta-1},\ 0<x<1 \tag{2}$$

where $\alpha,\beta>0$ are shape parameters. The raw moments of Kumaraswamy distribution are given by

$$E(X^r)=\int_0^1 x^r\ f_2(X;\alpha,\beta)dx$$

$$E(X^r)=\frac{\Gamma(\beta+1)\Gamma\left(1+\frac{r}{\alpha}\right)}{\Gamma\left(1+\beta+\frac{r}{\alpha}\right)} \tag{3}$$

In the Geeta distribution the parameters m and p are fixed but here we have considered a problem in which the parameter mm is fixed but the probability parameter p is itself a random variable following Kumarasawmy distribution (2)

III.DEFINITION OF PROPOSED DISTRIBUTION

If a random variable X follows Geeta distribution with parameters m and p where the parameter m is fixed but p instead of being a fixed constant is also a random variable following Kumarasawamy distribution then determining the distribution that results from marginalizing over p will be known as a compound of Geeta distribution with that of Kumarasawmy distribution.

Theorem 3.1: The probability function of a compound of GTD (m, p) with KSD (α,β) is given by

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=\ f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{mx-x}\binom{mx-x}{j}B\left(\beta,\frac{x+j+\alpha-1}{\alpha}+1\right)$$

where $x=1,2,...,m,\alpha,\beta>0$.

***Proof*:** With the help of definition of proposed distribution the probability function of a compound of GTD(m, p) with KSD(α, β) can be obtained as

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = \int_0^1 f_1\left(x|\, p\right) f_2\left(p\right) dp$$

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = \frac{\alpha\beta}{mx-1}\binom{mx-1}{x} \int_0^1 p^{x+\alpha-2}\left(1-p\right)^{mx-x}\left(1-p^{\alpha}\right)^{\beta-1} dp \tag{4}$$

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = \frac{\alpha\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{\infty}\binom{mx-x}{j}(-1)^j \int_0^1 p^{x+j+\alpha-2}\left(1-p^{\alpha}\right)^{\beta-1} dp$$

Substituting, $1-p^{\alpha} = z$ we get

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = \frac{\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{\infty}\binom{mx-x}{j}\int_0^1 z^{\beta-1}\left(1-z\right)^{\frac{x+j+\alpha-1}{\alpha}} dz$$

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = \frac{\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{\infty}\binom{mx-x}{j} B\left(\beta, \frac{x+j+\alpha-1}{\alpha}+1\right) \tag{5}$$

where $x = 1, 2, \ldots, m, \alpha, \beta > 0$. From here a random X variable following a compound of GTD with KSD will be symbolized by GTKSD(m, α, β).

In the special case if $m \in N$ the above probability function takes thesimpler rearranged form as

$$f_{CKSD}\left(X; m, \alpha, \beta\right) = f_{CKSD}\left(X; m, \alpha, \beta\right) = \frac{\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{mx-x}\binom{mx-x}{j} B\left(\beta, \frac{x+j+\alpha-1}{\alpha}+1\right) \tag{6}$$

where $x = 1, 2, \ldots, \alpha, \beta > 0$ and $m \in N$.

Alternatively we can also proceed from (4) as

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{\alpha\beta}{mx-1}\binom{mx-1}{x}\int_0^1 p^{x+\alpha-2}\left(1-p\right)^{mx-x}\left(1-p^{\alpha}\right)^{\beta-1}dp$$

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{\alpha\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{\infty}\binom{\beta-1}{j}\int_0^1 p^{x+\alpha+\alpha j-2}\left(1-p\right)^{mx-x}dp$$

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{\alpha\beta}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{\infty}\binom{\beta-1}{j}B\left(x+\alpha+\alpha j-1,mx-x+1\right)$$

where $x=1,2,...,m,\alpha,\beta>0$. This gives another form of pmf of GTKSD (m,α,β).

IV.SPECIAL CASES OF GTKSD

In this section it will be shown that GTKSD can be nested to different compound probability distributions for specific parameter setting

Case (i) Since for $m=1$ in GTD we obtain Haight distribution. Hence a compound of Haight distribution with Kumaraswmy distribution is followed from (6) by simply substituting $m=1$ in it. Therefore we have

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{\beta}{x-1}\binom{x-1}{x}B\left(\beta,\frac{x+j+\alpha-1}{\alpha}+1\right)$$

Where $x=1,2,...,\alpha,\beta>0$

***Case (ii)*:**

For $\alpha=\beta=1$, Kumaraswamy distribution reduces to uniform distribution. Hence a compound of GTD with uniform distribution is followed from (6) when we substitute $\alpha=\beta=1$ in it.

$$f_{CKSD}\left(X;m,\alpha,\beta\right)=f_{CKSD}\left(X;m,\alpha,\beta\right)=\frac{1}{mx-1}\binom{mx-1}{x}\sum_{j=0}^{mx-x}\binom{mx-x}{j}B\left(1,x+j+1\right)$$

Where $x=1,2,...,m>0$

Case (3)

If $X \sim$ GTKSD(m, a, β), then by setting $m = a = \beta = 1$ we obtain compound of Haight distribution with uniform distribution

$$f_{CKSD}\left(X;m,a,\beta\right) = f_{CKSD}\left(X;m,a,\beta\right) = \frac{1}{x-1}\binom{x-1}{x}B\left(1,x+j+1\right), \quad x=1,2,...$$

V.PARAMETER ESTIMATION

$$f_{GTKSD}\left(X;m,a,\beta\right) = f_{CKSD}\left(X;m,a,\beta\right) = n\log\beta - \sum_{i=1}^{n}\log(mx-1) + \sum_{i=1}^{n}\log\binom{mx-1}{x} + \sum_{i=1}^{n}\log\left[\sum_{j=0}^{mx-x}\binom{mx-x}{j}B\left(\beta,\frac{x+j+a-1}{a}+1\right)\right]$$

The above log likelihood cannot give simple differential equations so to estimate $\hat{m},\hat{a}$ and $\hat{\beta}$ we maximize the log likelihood function numerically using Newton-Raphson method in R software which is a very powerful technique for solving equations iteratively and numerically.

VI.APPLICATION

In this section we will explore the applicability of the proposed Geeta Kumaraswamy distribution by using a real data set

Table 1: Bunch size frequency distribution of Australian rural highways (Taylor et al.,1974)

Fitted Distribution		Observed Frequency	Number of mites per leaf
GTKSD	GTD		
126.14	129.42	127	1
54.12	60.73	53	2
26.01	21.6	29	3
19.31	15.3	21	4
6.75	9	5	5
4.83	3.12	4	6
2.24	3.12	1	7
6.14	2.21	5	8
245	245	245	Total
$\hat{m} = 0.80$, $\hat{a} = 0.68$, $\hat{\beta} = 0.95$	$\hat{m} = 1.12$, $\hat{p} = 0.45$		Parameter Estimation

2.12	8.93		Chi-Square Estimate
2	3		DF

VII.CONCLUDING REMARKS

We have developed a new discrete distribution that has been named as Geeta Kumaraswamy distribution that takes different functional forsm for pmf depending upon the parameter m. The resulting distribution embodies several discrete distributions for particular parameter substitution. Parameter estimation has been determined by employing MLE technique and there after newton raphson numerical method. In the end the potential of newly proposed distribution has been justified by using it to model a real life traffic data set and statistical results are strictly favouring Geeta-Kumaraswamy distribution.

REFERENCES

[1.] Adil, R and Jan, T.R. A mixture of generalized negative binomial distribution with Generalized exponential Distribution.*J. Stat. Appl. Pro*., **3**(3), 2014c, 451-464.

[2.] Adil, R and Jan, T.R. A compound of negative binomial distribution with two parameter Lindley distribution as a tool for handling overdispersion*Proceedings of International Conference on Advances in Computers, Communication and Electronic Engineering, ISBN: 978-93-82288-54-1*, 2015

[3.] Adil, R and Jan, T.R. A compound of zero truncated generalized negative binomial distribution with the generalized beta distribution. *Journal of Reliability and Statistical studies,* **6**(1), 2013, 11-19.

[4.] Adil, R and Jan, T.R. A compound of Geeta distribution with the Generalized Beta Distribution.*Journal Of Modern Applied Statistical Methods.,***13**(1), 2014a, 278-286.

[5.] Adil, R and Jan, T.R. A compound of Geeta distribution with the Generalized Beta Distribution. *Journal of Applied probability and Statistics Methods***9,** (2), 2014b, 17-24.

[6.] Geeta, P.C and Shenton, L.R On the probabilistic structures and properties of discrete Lagrangian distributions. *A modern course on statistical distributions in scientific work, D.Riedel, Boston,M.A***1**, 1975, 41-57.

[7.] Kumaraswamy, P. A generalized probability density functions for double-bounded random processes. *Journal of hydrology*, ***46****(1),* 1980, *79-88.http//dx.doi.org/10.1016/0022-1694(80)90036-0.*.

[8.] Sankaran, M. (1970) The Discrete Poisson-Lindley Distribution. *Biometrics.*,**26**, 145-149.

[9.] Zamani, H. and Ismail, N. Negative Binomial–Lindley Distribution And Its Application. *J. Mathematics And Statistics*.,**1**, 2010, 4–9.