Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/060118

A New Discrete Compound Distribution with Application

*Zahoor Ahmad**, *Adil Rashid and T. R. Jan*

P.G Department of Statistics, University of Kashmir, Srinagar, India.

Received: 2 Oct. 2016, Revised: 20 Jan. 2017, Accepted: 4 Feb. 2017.
Published online: 1 Mar. 2017.

Abstract: The present paper introduces a discrete compound distribution model, which is obtained by compounding size biased Consul Distribution with generalized beta distribution. The proposed distribution has several properties such as it can be nested to different compound distributions on specific parameter setting. Factorial moments and parameter estimation through maximum likelihood estimation and method of moment have been disused. The potentiality of the proposed model has been tested by chi-square goodness of fit test by modeling the real world count data sets.

Keywords: Consul Distribution, generalized beta distribution, compound distribution, factorial moment.

1 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. The researchers like Adil Rashid 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 Consul distribution with generalized beta distribution recently Adil Rashid and 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, 2016(a), 2016 (b)) constructed a new lifetime distribution and some count data models with wide applications in real life situations.

2 Consul Distribution (CD)

Consul distribution was introduced by Consul and Shenton (1975) was modified by Islam and Consul (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.

*Corresponding author e-mail: zahoor151@gmail.com

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 Consul distribution with probability mass function given by

A discrete r.vX is said to have a Consul distribution if its probability function is given by

$$p(X=x)=\begin{cases}\frac{1}{x}\binom{mx}{x-1}p^{x-1}(1-p)^{mx-x+1}, & ,x=1,2,3...,\\ 0 & otherwise\end{cases} \quad (1)$$

where $0<p<1$ and $1\le m\le p^{-1}$. The mean and the variance of the model exist when $m<p^{-1}$. The Consul distribution reduces to the geometric distribution when $m=1$. Famoye(1997a) obtained the model in (1.3) by using Lagrange expansion on the probability generating function of a geometric distribution and called it a generalized geometric distribution. The mean and variance of Consul distribution are given by

$\mu=(1-pm)^{-1}$ and $\sigma^2=mp(1-p)(1-pm)^{-3}$.

Consul distribution satisfies the dual properties of under-dispersion and over-dispersion. The model is under-dispersion for all values of $m\ge 1$ when $\mu\le(\sqrt{5}+1)/2$ and is over-dispersion for all values of $m\ge 1$ when $\mu>2$. The mean and variance of the Consul distribution are equal when $\mu=2$ and $m=1$.

3 Generalized BetaDistribution (GBD)

A random variable X is said to have a generalized beta distribution if its density function is given by

$$f_{GBD}(x;a,b,w,r)=\begin{cases}\dfrac{ax^{r-1}}{B\left(\frac{r}{a},w\right)(bw)^{\frac{r}{a}}}\left(1-\dfrac{x^a}{bw}\right)^{w-1} & ;0<x<(bw)^{\frac{1}{a}}\\ 0 & ;x\le 0\,or\,x\ge(bw)^{\frac{1}{a}}\end{cases} \quad (2)$$

where $a,b,w,r>0$ and $B\left(\frac{r}{a},w\right)$ is a beta function. Distribution (2.2) is a special limit case of the Bessel distribution investigated by Srodka (1973). It was also analysed by Seweryn (1986) and by Oginiski (1979)was applied in reliability theory. GBD reduces to beta distribution *a=1 b=1/w.*

4 Compounding of Size biased Consul Distribution (SCD) with the Generalized Beta Distribution (GBD)

Here, we shall present a compound of size biased Consul Distribution (SCD) with that of generalized beta distribution (GBD) by treating the success probability parameter in Consul distribution as a generalized beta variate. The resulting distribution so obtained generalizes several distributions. In addition to this, first order factorial moments of some compound distributions will also be discussed.

The pmf associated with the size biased version of (1) is given by

$$P_{SCD}(x;m,p)=\begin{cases}\binom{mx}{x-1}p^{x-1}(1-p)^{mx-x+1}(1-mp); & x=1,2,...\\ 0 & ;otherwise\end{cases} \quad (3)$$

Usually the parameters m and p in (3) are fixed but here we have considered a problem in which the parameter m is fixed but the probability parameter p is itself a random variable following generalized beta distribution , in that case the probability that exactly x members will be served before the queue vanishes is given by the compound of sized biased Consul distribution with that of generalized beta distribution.

Let us now consider SCD (3) that depends on cy:

$$P_{SCD}(x;m,p)=\begin{cases}\binom{mx}{x-1}(cy)^{x-1}(1-cy)^{mx-x+1}(1-mcy); & x=1,2,...\\ 0 & ;otherwise\end{cases} \quad (4)$$

where $0<cy<1$ and $1<m<\frac{1}{cy}$ and Y is a random variable following GBD (2)

4.1 Definition of proposed distribution

If $X \mid p$ be a random variable following SCD $(x,m,p=cy)$, the parameter m is fixed but p instead of being a fixed constant is also a random variable such that $p=cy$ where Y is distributed as GBD $(y;a,b,w,r)$ then determining the distribution that results from marginalizing over Y will be known as a compound of CD with that of GBD.

Theorem 1: The probability function of the compound ofsize biased Consul distribution with generalized beta distribution is given by the expression

$$P_{SCGBD}(x;m,a,b,w,r)=\eta\sum_{k=0}^{\infty}\binom{mx-x+1}{k}(-c)^k(bw)^{\frac{k-1}{a}}B\left(\frac{x+r+k-1}{a},w\right)$$

where, $\eta=\dfrac{\binom{mx}{x-1}c^{x-1}(bw)^{\frac{x}{a}}}{xB\left(\frac{r}{a},w\right)}$ and $x=1,2,...,a,b,w,r>0,\ 0<cy<1$ and $mcy<1$

Proof: If $X \mid p \sim$ CD $(m,p=cy)$ and $Y\sim$ GBD $(Y;a,b,w,r)$, then the probability function of a compound SCD $(m,p=cy)$ with GBD $(Y;a,b,w,r)$ can be obtained with the help of definition of proposed distribution (4.1)

$$P_{SCGBD}(x;m,a,b,w,r)=\int_0^{(bw)^{\frac{1}{a}}}P_1(X\mid p=cy)P_2(y)\,dy$$

$$=a\frac{\binom{mx}{x-1}c^{x-1}}{(bw)^{\frac{r}{a}}B\left(\frac{r}{a},w\right)}\int_0^{(bw)^{\frac{1}{a}}}y^{x+r-2}\left(1-\frac{y^a}{bw}\right)^{w-1}(1-cy)^{mx-x+1}(1-cy)\,dy$$

$$=a\eta_1\left(\int_0^{(bw)^{\frac{1}{a}}}y^{x+r-2}\left(1-\frac{y^a}{bw}\right)^{w-1}(1-cy)^{mx-x+1}dy-c\int_0^{(bw)^{\frac{1}{a}}}y^{x+r-1}\left(1-\frac{y^a}{bw}\right)^{w-1}(1-cy)^{mx-x+1}dy\right)$$

$$=a\eta_1\sum_{k=0}^{\infty}(-c)^k\binom{mx-x+1}{k}\left(\int_0^{(bw)^{\frac{1}{a}}}y^{x+r+k-2}\left(1-\frac{y^a}{bw}\right)^{w-1}dy-c\int_0^{(bw)^{\frac{1}{a}}}y^{x+r+k-1}\left(1-\frac{y^a}{bw}\right)^{w-1}dy\right)$$

Where, $\eta_1 = \dfrac{\binom{mx}{x-1} c^{x-1}}{(bw)^{\frac{r}{a}} B\left(\frac{r}{a}, w\right)}$

Substituting, $\dfrac{y^a}{bw} = t,$ we get

$$P_{CGBD}(x;m,a,b,w,r) = \frac{\binom{mx}{x-1} c^{x-1} (bw)^{\frac{x}{a}}}{B\left(\frac{r}{a}, w\right)} \sum_{k=0}^{\infty} (-c)^k \binom{mx-x+1}{k} \left((bw)^{\frac{k-1}{a}} \int_0^1 t^{\frac{x+r+k-1}{a}-1} (1-t)^{w-1} dt \; - c(bw)^{\frac{k}{a}} \int_0^1 t^{\frac{x+r+k}{a}-1} (1-t)^{w-1} dt \right)$$

Using the definition of beta function we get,

$$P_{CGBD}(x;m,a,b,w,r) = \eta \sum_{k=0}^{\infty} (-c)^k \binom{mx-x+1}{k} \left((bw)^{\frac{k-1}{a}} B\left(\frac{x+r+k-1}{a}, w \right) - c(bw)^{\frac{k}{a}} B\left(\frac{x+r+k}{a}, w \right) \right) \quad (5)$$

Where, $\eta = \dfrac{\binom{mx}{x-1} c^{x-1} (bw)^{\frac{x}{a}}}{B\left(\frac{r}{a}, w\right)}$ and $x = 1,2,\ldots, a,b,w,r > 0,\ 0 < cy < 1$ and $mcy < 1$

4.2 Special cases

***Case I*:** In case when $m=1$ in (3) SCD reduces to the size biased geometric distribution (SGD) and a compound of SGD with GBD is simply followed from (5) when we put $m = 1$

$$P_{SGGBD}(x;1,a,b,w,r) = \eta \sum_{k=0}^{\infty} (-c)^k \binom{1}{k} \left((bw)^{\frac{k-1}{a}} B\left(\frac{x+r+k-1}{a}, w \right) - c(bw)^{\frac{k}{a}} B\left(\frac{x+r+k}{a}, w \right) \right) \quad (6)$$

Where $\eta_2 = \dfrac{c^{x-1} (bw)^{\frac{x}{a}}}{B\left(\frac{r}{a}, w\right)}$ and $x = 1,2,\ldots, a,b,w,r > 0,\ 0 < cy < 1$ and $cy < 1$

***Case II*:** When $b = \dfrac{1}{w}$ and $a = 1$, in (2) the generalized beta distribution reduces to beta distribution and a compound of SCD with beta distribution is simply followed from (5) on $b = \dfrac{1}{w}$ and $a = 1$

$$P_{SCBD}(x;m,a,b,w,r) = \eta_3 \sum_{k=0}^{\infty} (-c)^k \binom{mx-x+1}{k} \left(B(x+r+k-1, w) - cB(x+r+k, w) \right) \quad (7)$$

Where $\eta_3 = \dfrac{\frac{1}{x} \binom{mx}{x-1} c^{x-1}}{B(r,w)}$ and $x = 1,2,\ldots, w,r > 0,\ 0 < cy < 1$ and $mcy < 1$

Case III: When $b=\frac{1}{w}$, $a=1$ and $m=1$ in (2) and (3) respectively, we obtain geometric and beta distribution and a compound of GD with BD is followed from (5) when we substitute $m=1$ and $b=\frac{1}{w}$, $a=1$ in it

$$P_{SGBD}(x;1,1,\frac{1}{w},w,r)=\frac{c^{x-1}}{B(r,w)}\sum_{k=0}^{\infty}(-c)^{k}\binom{1}{k}\Big(B\big(x+r+k-1,w\big)-cB\big(x+r+k,w\big)\Big) \tag{8}$$

Where $x=1,2,...,w,r>0,\ 0<cy<1$

5 Factorial Moments of the Compound of SCD distribution with GBD and Some Special Cases

If $X\mid p\sim$ SCD $(x;p)$, where p follows GBD $(P;a,b,w,r)$ then

$$\mu_{[l]}(X)=E_{\lambda}\left[m_{l}\left(x\mid p\right)\right] \tag{9}$$

is called a factorial moment of order l of a compound of SCD with GBD where $m_l\left(X\mid p\right)$ is theist order factorial moment of SCD.

Theorem 2: The first order factorial moment of a compound SCD with GBD is given by the formula

$$\mu_{[1]SCGBD}(x)=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^{k}(bw)^{\frac{k}{a}}}{B(r/a,w)}\left(B\left(\frac{r+k}{a}+1,w\right)-mc^{2}(bw)^{\frac{z}{a}}B\left(\frac{r+k+z}{a},w\right)\right)$$

where $x=1,2,3,...,a,b,w,r>0$

Proof: The I[st] order factorial moment of SCD is given by

$$m_{[1]}\left(x,p\right)=\frac{1-mp^{2}}{(1-mp)^{2}}$$

Therefore I[st] order factorial moment of a compound of SCD with GBD can be obtained by using the definition (9), if we let $p=cy$,

$$\mu_{[1]SCGBD}(x)=E_{Y}\left[m_{1}\left(X\mid p=cy\right)\right]$$

$$=E_{Y}\left(\frac{1-m(cy)^{2}}{(1-mcy)^{2}}\right)$$

$$=\int_{0}^{(bw)^{\frac{1}{a}}}\left(\frac{1-m(cy)^{2}}{(1-mcy)^{2}}\right)p_{2}(y)\,dy$$

$$=\frac{a}{(bw)^{r/a}B(r/a,w)}\int_{0}^{(bw)^{\frac{1}{a}}}\left(\frac{1-m(cy)^{2}}{(1-mcy)^{2}}\right)y^{r-1}\left(1-\frac{y^{a}}{bw}\right)^{w-1}dy$$

$$=\frac{a\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^{k}}{(bw)^{r/a}B(r/a,w)}\left(\int_{0}^{(bw)^{\frac{1}{a}}}y^{r+k-1}\left(1-\frac{y^{a}}{bw}\right)^{w-1}dy-mc^{2}\int_{0}^{(bw)^{\frac{1}{a}}}y^{r+k+1}\left(1-\frac{y^{a}}{bw}\right)^{w-1}dy\right)$$

Substituting, $\frac{y^a}{bw}=t,$ we get

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^k (bw)^{\frac{k}{a}}}{B(r/a,w)}\left(\int_0^{(bw)^{\frac{1}{a}}} t^{\frac{r+k}{a}}(1-t)^{w-1}dt - mc^2(bw)^{\frac{z}{a}}\int_0^{(bw)^{\frac{1}{a}}} t^{\frac{r+k+z}{a}-1}(1-t)^{w-1}dt\right)$$

Using the definition of beta function we obtain

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^k (bw)^{\frac{k}{a}}}{B(r/a,w)}\left(B\left(\frac{r+k}{a}+1,w\right)-mc^2(bw)^{\frac{z}{a}}B\left(\frac{r+k+z}{a},w\right)\right) \tag{10}$$

Corollary 5.1 The Ist order factorial moment of a compound of size biased geometric distribution GBD is

$$\mu_{[1]SGGBD}(x)=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-c)^k (bw)^{\frac{k}{a}}}{B(r/a,w)}\left(B\left(\frac{r+k}{a}+1,w\right)-c^2(bw)^{\frac{z}{a}}B\left(\frac{r+k+z}{a},w\right)\right)$$

***Proof*:** In case when $m=1$ in (3) SCD reduces to the size biased geometric distribution (SGD) and therefore Istorder factorial moment of a compound of SGD with GBD is simply followed from (10) when we put $m=1$

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-c)^k (bw)^{\frac{k}{a}}}{B(r/a,w)}\left(B\left(\frac{r+k}{a}+1,w\right)-c^2(bw)^{\frac{z}{a}}B\left(\frac{r+k+z}{a},w\right)\right)$$

***Corollary*5.2**The Ist order factorial moment of a compound of size biasedConsul distribution with beta distribution is

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^k}{B(r,w)}\left(B(r+k+1,w)-mc^2B(r+k+z,w)\right)$$

Proof: When $b=\frac{1}{w}$ and $a=1$, in (2) generalized beta distribution reduces to beta distribution and Ist order factorial moment of a compound of CD with beta distribution is simply followed from (10) on $b=\frac{1}{w}$ and $a=1$

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-mc)^k}{B(r,w)}\left(B(r+k+1,w)-mc^2B(r+k+z,w)\right)$$

Corollary 5.3 The Ist order factorial moment of a compound of size biased geometric distribution beta distribution is

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-c)^k}{B(r,w)}\left(B(r+k+1,w)-c^2B(r+k+z,w)\right)$$

Proof: When $b=\frac{1}{w}$, $a=1$ and $m=1$ in (2) and (3) respectively, we obtain geometric and beta distribution and therefore

I[st] order factorial moments of a compound of GD with BD is followed from (10) when we substitute $m=1$ and $b=\frac{1}{w}$, $a=1$ in it

$$=\frac{\sum_{k=0}^{\infty}\binom{-2}{k}(-c)^k}{B(r,w)}\left(B(r+k+1,w)-c^2B(r+k+z,w)\right)$$

6 Applications

In this section we will explore the applicability of the proposed compound distribution by using a real data set on bunching traffic in Australian rural highways which have been taken from Taylor et al. (1974).The data which appears in the first two columns of table 1 gives bunch size with observed corresponding frequency and the data which appears in the 3[rd] and 4[th] column of this table is the fitted Consul distribution, Consul Kumaraswamy distribution and proposed distribution.

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

Number of mites per leaf	Observed Frequency	Fitted Distribution		
		CD	CKSD	SCGBD
1	127	125.42	125.64	127.24
2	53	58.83	59.27	66.99
3	29	29.07	29.60	31.4
4	21	14.84	14.51	23.10
5	5	9.29	6.77	3.50
6	4	4.12	2.93	4.5
7	1	2.22	1.14	1.14
8	5	1.21	5.14	4.5
Total	245	245	245	
Parameter Estimation		$\hat{m}=1.12$ $\hat{p}=0.45$	$\hat{m}=0.80$ $\hat{\alpha}=0.68,\ \hat{\beta}=0.95$	$\hat{m}=2.69$ $\hat{a}=2.10,\ \hat{r}=1.64$ $\hat{w}=201.8$
ChiSquare Estimate		5.93	4.12	4.05

7 Conclusions

In this paper, we have proposed a compound of SCD with GBD by compounding, the SCD with GBD. The new distribution so obtained has some desirable properties that is they can be nested to different compound distributions on specific parameter setting. Moreover, the factorial moments of proposed distributions have also been discussed along with some special cases.In the end it has been shown that proposed distribution provides a adequate fit to the reported real life data set.

References

[1] Adil, R and Jan, T.R. (2014c) A mixture of generalized negative binomial distribution with Generalized exponential Distribution.*J.*

Stat. Appl. Pro., **3**(3) 451-464.

[2] Adil, R and Jan, T.R. (2015) 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*

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

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

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

[6] Adil R, Jan T R. ,(2016b)A new family of lifetime distribution with real life application.*International Journal of Mathematics and Statistics*: **, 17**(1), 23-38

[7] Adil, R., Zahoor, A., and Jan, T.R. (2016a). A new count data model with application in genetics and ecology. *Electronic Journal of Applied Statistical Analysis,* **9**(1), 213-226

[8] Consul, P.C and Shenton, L.R (1975) 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**, 41-57.

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

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

[11] Taylor MAP, Young W. (1988) *Traffic Analysis,* New technology and new solutions. Hargreen Publishing Company: Melbourne, Australia,

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

Zahoor Ahmad is Lecturer of Statistics in Department of School Education, Jammu and Kashmir. He holds Master degree in Statistics from the University of Kashmir and has been a gold medalist. Presently he is doing P.hd in the Department of Statistics, University of Kashmir. He has published work in international and national journals of repute. His interests Distribution theory, Estimation theory and Bio-Statistics.

Adil Rashid was born in 1984. He holds M.Sc. and M. Phil. degrees in Statistics from University of Kashmir, Srinagar, India in 2009 and 2013 respectively. At present he is pursuing his Ph. D degree in Statistics in the Post Graduate Department of Statistics, University of Kashmir, Srinagar, India. He has to his credit several research publications both in national and international journals of repute. His current research interests are in Compound Probability Distributions, Bio-Statistics and Models for Complex Systems.

Tariq Rashid Jan was born in 1966. He did his M.Sc. and M. Phil. degrees in Statistics from Department of Statistics and Operations Research, Aligarh Muslim University, Aligarh, India in 1989 and 1994 respectively. He received his Ph. D. degree in Statistics from department of Statistics, University of Kashmir, Srinagar, India in 2004. At present he is a Sr. Assistant Professor in the Department of Statistics, University of Kashmir, Srinagar. He has published work in international and national journals of repute. His current research interests are in Bio -Statistics, Generalized Distributions and Reliability Theory.