IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728.Volume 5, Issue 2 (Jan. - Feb. 2013), PP 41-48
www.iosrjournals.org

On size biased Generalized Beta distribution of first kind

A.Ahmed[1], K.A.Mir[2], J.A.Reshi[1]

[1] *Department of Statistics, University of Kashmir, Srinagar, India.*

[2] *Department of Statistics, Govt.Degree CollegeBemina Srinagar, India.*

Abstract: *In this paper, we develop the Size-biased form of the weighted Generalized Beta distribution of first kind (WGBD1); a particular case of the weighted Generalized Beta distribution of first kind, taking the weights as the variate values has been defined. The structural property of Size-biased Generalized Beta distribution of first kind (SBGBD1) including moments, variance, mode and harmonic mean, coefficient of variation, skewness and kurtosis has been derived. The estimates of the parameters of Size- biased Generalized Beta distribution of first kind (SBGBD1) is obtained by employing a new method of moments. Also, a test for detecting the size-baisedness is conducted.*

***Keywords**: Generalized Beta distribution of first kind, Beta function, Size-biased generalized beta distribution of first kind, Structural properties and moment estimator.*

I. Introduction

Beta distributions are very versatile and a variety of uncertainties can be usefully modelled by them. Many of the finite range distributions encountered in practice can be easily transformed into the standard distribution. In reliability and life testing experiments, many times the data are modeled by finite range distributions, see for example Barlow[1] and Proschan(1975) . Many generalizations of beta distributions involving algebraic and exponential functions have been proposed in the literature; see in Johnson et al.[2] (1995) and Gupta and NadarSajah (2004)[3] for detailed accounts. The generalized beta distribution of first kind[13] (GBD1) is a very flexible four parameter distribution. It captures the characteristics of income distribution including skewness, peakedness in low-middle range, and long right hand tail. The Generalized Beta distribution of first kind includes several other distributions as special or limiting cases, such as generalized gamma(GGD), Dagum, beta of the second kind(BD2), Sing-Maddala (SM), gamma, Weibull and exponential distributions.

The probability density function (pdf) of the generalized beta distribution of first kind (GBD1) is given by:

$$f(x;a,b,p,q) = \frac{a}{b^{ap}\beta(p,.q)} x^{ap-1}\left(1-\left(\frac{x}{b}\right)^{a}\right)^{q-1} \qquad for\ \ x>0 \qquad (1)$$
$$= 0, otherwise$$

where a, p, q are shape parameters and b is a scale parameter, $\beta(p,q) = \frac{\Gamma p \Gamma q}{\Gamma p+q}$ is a beta function, a,b, p,q are positive real values.

The rth moment of generalized beta distribution of first kind is given by [14]:

$$E(X^r) = \frac{b^r \beta\left(p+\frac{r}{a},q\right)}{\beta(p,q)} \qquad (2)$$

Put r=1 in relation (2), we have

$$E(X) = \frac{b\beta\left(p+\frac{1}{a},q\right)}{\beta(p,q)} \qquad (3)$$

The concept of weighted distributions can be traced to the study of the effect of methods of ascertainment upon estimation of frequencies by Fisher [4].In extending the basic ideas of Fisher, Rao [5,6] saw the need for a unifying concept and identified various sampling situations that can be modeled by what he called weighted distributions. The weighted distributions arise when the observations generated from a stochastic process are not given equal chance of being recorded; instead they are recorded according to some weighted function. When the weight function depends on the lengths of the units of interest, the resulting distribution is called length biased. More generally, when the sampling mechanism selects units with probability proportional to measure of the unit size, resulting distribution is called size-biased. Size biased distributions are a special case of the more general form known as weighted distributions. First introduced by Fisher (1934) to model ascertainment bias, these were later formalized in a unifying theory by Rao (1965). These distributions arise in practice when observations from a sample are recorded with unequal probability and provide unifying approach for the problems when the observations fall in the non –experimental, non –replicated and non –random categories. Van Deusen [11] arrived at size biased distribution theory independently and applied it to fitting distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS) (Grosenbaugh) inventories. Subsequently, Lappi and Bailey[12] used weighted distributions to analyse HPS diameter increment data. Dennis [7] and Patil (1984) used stochastic differential equations to arrive at a weighted gamma distribution as the stationary probability density function (PDF) for the stochastic population model with predation effects. Gove (2003) [8] reviewed some of the more recent results on size-biased distributions pertaining to parameter estimation in forestry. Sandal[9] C.E (1964) derived the method of estimation of parameters of the gamma distribution. Mir [10] (2007) also discussed some of the discrete size-biased distributions.

II. Materials And Methods

Suppose X is a non-negative random variable with its natural probability density function $f(x;\theta)$,where the natural parameter is θ. Suppose a realization x of X under $f(x;\theta)$ enters the investigator's record with probability proportional to $w(x;\beta)$, so that the weight function $w(x;\beta)$ is a non-negative function with the parameter β representing the recording mechanism. Clearly, the recorded x is not an observation on X , but on the random variable X_w, having a pdf

$$f_w\ (x;\theta,\beta)=\frac{w(x;\beta)\ \ f(x;\theta)}{\int\limits_0^\infty w(x)\,f(x)},\ x>0 \tag{4}$$

Assuming that $E(X)=\int\limits_0^\infty w(x)\,f(x)\ dx$ i,e the first moment of w(x) exists.

By taking weight w(x) = x we obtain length biased distribution. where w is the normalizing factor obtained to make the total probability equal to unity by choosing $w=E[w(x,\beta)]$. The variable X_W is called weighted version of X , and its distribution is related to that of X and is called the weighted distribution with weight function w .For example, when $w(x;\beta)=x$, in that case $w=\mu$ is called the size-biased version of X .The distribution of X^* is called the size-biased distribution with pdf

$$f^*(x;\theta)=\frac{x\,f(x;\theta)}{\mu} \tag{5}$$

III. Derivation of Size-Biased Generalized Beta Distribution of first kind:

The probability distribution of Generalized Beta distribution of first kind is:

$$\begin{aligned} f(x;a,b,p,q) &= \frac{a}{b^{ap}\beta(p,.q)}\,x^{ap-1}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1} \qquad for\ \ x>0 \\ &= 0, otherwise \end{aligned}$$

where a,p,q are shape parameters and $\beta(p,q)=\frac{\Gamma p\Gamma q}{\Gamma p+q}$ is a beta function, p,q are positive real values.

A size biased generalized beta distribution of first kind (SBGBD1) is obtained by applying the weights x^c, where c =1 to the weighted Generalized beta distribution of first kind.
We have from relation (1) and (5)

$$f^*(x;a,b,p,q) = \frac{x\, f(x;a,b,p,q)}{\mu}$$

$$f^*(x;a,b,p,q) = \int_0^\infty x \frac{a\ x^{ap-1}}{b^{ap}\,\beta(p,q)}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1} . \frac{\beta(p,q)}{b\,\beta\left(p+\frac{1}{a},q\right)}\ dx$$

$$f^*(x;a,b,p,q) = \int_0^\infty \frac{a\ x^{ap}}{b^{ap+1}\ \beta\left(p+\frac{1}{a},q\right)}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1} dx$$

$$f^*(x;a,b,p,q) = \frac{a}{b^{ap+1}\ \beta\left(p+\frac{1}{a},q\right)} x^{ap}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1}$$

where $f^*(x;a,b,p,q)$ represents a probability density function. This gives the size –biased generalized beta distribution of first kind (SBGBD1) as:

$$f^*(x;a,b,p,q) = \frac{a}{b^{ap+1}\ \beta\left(p+\frac{1}{a},q\right)} x^{ap}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1} \tag{6}$$

where a,p,q are shape parameters and b is a scale parameter, $\beta(p+1,q) = \frac{\Gamma p+1\Gamma q}{\Gamma p+q+1}$ is a beta function, a,b,p,q are positive real values.

Special case: The distribution like the Size-biased beta distribution of first kind as special case ($a=b=1$), then the probability density function is given as:

$$f^*(x;p,q) = \frac{1}{\beta(p+1,q)} x^p\ (1-x)^{q-1},\ p>0, q>0$$

$\beta(p+1,q) = \frac{\Gamma p+1\Gamma q}{\Gamma p+q+1}$ is a beta function, a,b,p,q are positive real values.

IV. Structural properties of Size- biased Generalized beta distribution of first kind:

The rth moment of Size biased generalized beta distribution of first kind (6) about origin is obtained as:

$$\mu_r' = \int_0^\infty x^r f(x;a,b,p,q)dx$$

$$\mu_r' = \int_0^\infty x^r \frac{a\,x^{ap}}{\beta\left(p+\frac{1}{a},q\right)b^{ap+1}}\left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

$$\mu_r' = \int_0^\infty \frac{a\,x^{ap+r}}{\beta\left(p+\frac{1}{a},q\right)b^{ap+1}}\left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

$$\mu_r' = \frac{ab^{r-1}}{\beta\left(p+\frac{1}{a},q\right)}\int_0^\infty \left(\frac{x}{b}\right)^{ap+r}\left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

$$\mu_r' = \frac{ab^{r-1}}{\beta\left(p+\frac{1}{a},q\right)}\int_0^\infty \left[\left(\frac{x}{b}\right)^a\right]^{p+\frac{r}{a}}\left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

Put $\left(\frac{x}{b}\right)^a = t,\ then\quad x = bt^{\frac{1}{a}}, dx = \frac{b}{a}t^{\frac{1}{a}-1}dt$

and $\mu_r' = \frac{ab^{r-1}}{\beta\left(p+\frac{1}{a},q\right)}\int_0^\infty [t]^{p+\frac{r}{a}}[1-t]^{q-1}\frac{b}{a}t^{\frac{1}{a}-1}dt$

$$\mu_r' = \frac{b^r}{\beta\left(p+\frac{1}{a},q\right)}\int_0^\infty [t]^{p+\frac{r}{a}+\frac{t}{a}-1}[1-t]^{q-1}dt$$

$$\mu_r' = \frac{b^r}{\beta\left(p+\frac{1}{a},q\right)}\beta\left(p+\frac{r}{a}+\frac{1}{a},q\right) \tag{7}$$

4.1 Mean of Size- biased Generalized Beta Distribution of first kind.

Using the equation (7), the mean of the SBGBD1 is given by

$$\mu_1' = \frac{b\,\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} \tag{8}$$

4.2 Second moment of Size- biased Generalized Beta Distribution of first kind.

Using the equation (7), the second moment of the SBGBD1 is given by

$$\mu_2' = \frac{b^2\,\beta\left(p+\frac{3}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} \tag{9}$$

4.3 Variance of Size- biased Generalized Beta Distribution of first kind.

Using the equation (07), the variance of the SBGBD1 is given by

$$\mu_2 = b^2\left[\frac{\beta\left(p+\frac{3}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} - \left[\frac{\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}\right]^2\right] \tag{10}$$

4.4 Coefficient of variation of Size- biased Generalized Beta Distribution of first kind.

$$CV = \frac{\sqrt{V(X)}}{E(X)} = \sqrt{\frac{\beta\left(p+\frac{3}{a},.q\right)\beta\left(p+\frac{1}{a},q\right)}{\beta^2\left(p+\frac{2}{a},q\right)}-1} \tag{11}$$

4.5 Coefficient of skewness of Size- biased Generalized Beta Distribution of first kind.

$$CS = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3' - 3\mu_1'\mu_2' + 2\mu_1'}{\sigma^3} \tag{12}$$

4.6Coefficient of Kurtosis of Size- biased Generalized Beta Distribution of first kind.

$$CS = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3' - 4\mu\mu_3' + 6\mu_1'^2\mu_2' - 3\mu^4}{\sigma^3} \tag{13}$$

Where, the first four moments about origin is given as:

$$\mu_1' = \frac{b\,\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)},$$

$$\mu_2' = \frac{b^2\,\beta\left(p+\frac{3}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}$$

$$\mu_2 = b^2\left[\frac{\beta\left(p+\frac{3}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} - \left[\frac{\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}\right]^2\right]$$

$$\mu_3' = \frac{b^3\,\beta\left(p+\frac{4}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}$$

$$\mu_4' = \frac{b^4\,\beta\left(p+\frac{5}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}$$

4.7 The mode of Size-biased generalized beta distribution of first kind is given as:

The probability distribution of Size-biased Generalized Beta distribution of first kind is:

$$f^*(x;a,b,p,q) = \frac{a}{b^{ap+1}\,\beta\left(p+\frac{1}{a},q\right)}\,x^{ap}\left(1-\left(\frac{x}{b}\right)^a\right)^{q-1}$$

In order to discuss monotonicity of size-biased generalized beta distribution of first kind. We take the logarithm of its pdf:

$$\ln(f(x:a,b,p,q)) = \ln\left(\frac{a}{b^{ap+1}\beta\left(p+\frac{1}{a},q\right)}\right) + \ln x^{ap} + \ln\left\{\left[1-\left(\frac{x}{b}\right)^a\right]^{q-1}\right\} \tag{14}$$

Where C is a constant. Note that

$$\frac{\partial \ln f^*(x;a,b,p,q)}{\partial x} = \frac{ap\left(b^a - x^a\right) - (q-1)ab^a x^a}{x\left(b^a - x^a\right)}$$

Where a, p, q are shape parameters and b is a scale parameter, It follows that

$$\frac{\partial \ln f^*(x;a,b,p,q)}{\partial x} > 0 \Leftrightarrow x < \frac{p^{\frac{1}{a}}\,b\left(b^a - x^a\right)^{\frac{1}{a}}}{q-1}$$

$$\frac{\partial \ln f^*(x;a,b,p,q)}{\partial x} < 0 \Leftrightarrow x > \frac{p^{\frac{1}{a}} b \left(b^a - x^a\right)^{\frac{1}{a}}}{q-1}$$

$$\frac{\partial \ln f^*(x;a,b,p,q)}{\partial x} = 0 \Leftrightarrow x = \frac{p^{\frac{1}{a}} b \left(b^a - x^a\right)^{\frac{1}{a}}}{q-1}$$

The mode of size-biased generalized beta distribution of first kind is:

$$x = \frac{p^{\frac{1}{a}} b \left(b^a - x^a\right)^{\frac{1}{a}}}{q-1} \qquad (15)$$

4.8 The harmonic mean of Size-biased generalized beta distribution of first kind is given as:
The probability distribution of Size-biased Generalized Beta distribution of first kind is:

$$f^*(x;a,b,p,q) = \frac{a}{b^{ap+1}\,\beta\left(p+\frac{1}{a},q\right)} x^{ap} \left(1-\left(\frac{x}{b}\right)^a\right)^{q-1}$$

The harmonic mean (H) is given as:

$$\frac{1}{H} = \int_0^\infty \frac{1}{x} f(x;a,b,p,q)dx$$

$$\frac{1}{H} = \int_0^\infty \frac{1}{x} \frac{a\,x^{ap}}{b^{ap+1}\beta\left(p+\frac{1}{a},q\right)} \left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

$$\frac{1}{H} = \frac{a}{b^2\,\beta\left(p+\frac{1}{a},q\right)} \int_0^\infty \left(\frac{x}{b}\right)^{ap-1} \left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

$$\frac{1}{H} = \frac{a}{b^2\,\beta\left(p+\frac{1}{a},q\right)} \int_0^\infty \left[\left(\frac{x}{b}\right)^a\right]^{p-\frac{1}{a}} \left[1-\left(\frac{x}{b}\right)^a\right]^{q-1} dx$$

Put $\left(\frac{x}{b}\right)^a = t,\ then\quad x = bt^{\frac{1}{a}}, dx = \frac{b}{a} t^{\frac{1}{a}-1} dt$

$$\frac{1}{H} = \frac{a}{b^2\,\beta\left(p+\frac{1}{a},q\right)} \int_0^\infty [t]^{p-\frac{1}{a}} [1+t]^{q-1} \frac{b}{a} t^{\frac{1}{a}-1} dt$$

$$\frac{1}{H} = \frac{1}{b\,\beta\left(p+\frac{1}{a},q\right)} \int_0^\infty [t]^{p-1} [1+t]^{q-1} dt$$

$$\frac{1}{H} = \frac{\beta(p,q)}{b\,\beta\left(p+\frac{1}{a},q\right)}$$

$$H = \frac{b\beta\left(p+\frac{1}{a}\right)\beta(q)}{\beta(p,q)} \qquad (16)$$

V. Estimation of parameters in the size-biased Generalized Beta Distribution of first kind.

In this section, we obtain estimates of the parameters for the Size-biased Generalized Beta distribution of first kind by employing the new method of moment (MOM) estimator.

5.1 New Method of Moment Estimators

Let X_1, X_2, X_3… Xn be an independent sample from the SBGBD1with weight c=1. The method of moment estimators are obtained by setting the row moments equal to the sample moments, that is $E(X^r) = M_r$ where is the sample moment M_r corresponding to the .$E(X^r)$ The following equations are obtained using the first and second sample moments.

$$\frac{1}{n}\sum_{j=1}^{n} X_j = \frac{b\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} \qquad (17)$$

$$\frac{1}{n}\sum_{j=1}^{n} X_j^{\,2} = \frac{b^2\beta\left(p+\frac{3}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)} \qquad (18)$$

Case 1.When p and q are fixed and a=1, then

$$\overline{X} = \frac{b\beta\left(p+\frac{2}{a},q\right)}{\beta\left(p+\frac{1}{a},q\right)}$$

$$\overline{X} = \frac{b\Gamma(p+2)\Gamma(p+q+1)}{\Gamma(p+q+2)\Gamma(p+1)}$$

$$\hat{b} = \overline{X}\,\frac{(p+q+1)}{p+1} \qquad (19)$$

Case 2. When p and b are fixed and a=1, then dividing equation (17) by (18), we have:

$$\frac{\overline{X}}{M_2} = \frac{\Gamma(p+2)\Gamma(p+q+3)}{b\Gamma(p+q+2)\Gamma(p+3)}$$

$$\frac{\overline{X}}{M_2} = \frac{p+q+2}{b(p+2)}$$

$$\hat{q} = (p+2)\left[\frac{b\overline{X}}{M_2} - 1\right] \qquad (20)$$

Case 3: When b and q are fixed and a=1, then dividing equation (17) by (18), we have:

$$\frac{\overline{X}}{M_2} = \frac{p+q+2}{b(p+2)}$$

$$\hat{p} = \frac{qM_2}{b\overline{X} - M_2} - 2 \qquad (21)$$

Case 4. When p and q are fixed, b=1 then we can calculate the value of a estimator by numerical methods/

VI. Test for size-biased generalized beta distribution of second kind.

Let X_1, X_2, X_3,…,X_n be a random samples can be drawn from generalized beta distribution of first kind or size-biased generalized beta distribution of first kind. We test the hypothesis $H_o : f(x) = f(x,a,b,p,q)\, againest\; H_1 : f(x) = f_s^{\,*}(a,b,p,q)$.

To test whether the random sample of size n comes from the generalized beta distribution of first kind or size-biased generalized beta distribution of first kind the following test statistic is used.

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{f^*{}_s(x;a,b,p,q)}{f(a,b,p,q)}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{\dfrac{ax^{ap}}{b^{ap+1}\beta\left(p+\frac{1}{a},q\right)}\left[1-\left(\frac{x}{b}\right)^{a}\right]^{q-1}}{\dfrac{ax^{ap-1}}{b^{ap}\beta(p,q)}\left[1-\left(\frac{x}{b}\right)^{a}\right]^{q-1}}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{\beta(p,q)}{b\beta\left(p+\frac{1}{a},q\right)}.x$$

$$\Delta = \left[\frac{\beta(p,q)}{b\beta\left(p+\frac{1}{a},q\right)}\right]^{n} \prod_{i=1}^{n} x_i \tag{22}$$

We reject the null hypothesis.

$$\left[\frac{\beta(p,q)}{b\Gamma\left(p+\frac{1}{a},q\right)}\right]^{n} \prod_{i=1}^{n} x_i > k$$

Equalivalently, we rejected the null hypothesis where

$$\Delta^* = \prod_{i=1}^{n} x_i > k^*, where\, k^* = k\left[\frac{b\beta\left(p+\frac{1}{a},q\right)}{\beta(p,q)}\right]^{n} > 0$$

For a large sample size of n, $2\log\Delta$ is distributed as a Chi-square distribution with one degree of freedom. Thus, the p-value is obtained from the Chi-square distribution.

References

[1] Barlow, R.E and Proschan, F.(1975). Statistical Theory of Reliability and Life Testing:Probobality Models. New York, Rinehart and Winston.

[2] Johnson, N.L., Kotz, S. and Balakrishnan, N. 1995. Continuous Univariate Distributions, Volume 2(second edition). New York: John Wiley and Sons.

[3] Gupta, A.K. and Nadarajah, S.(2004). Handbook of Beta Distribution and its Applications.NewYork:Marcel Dekker.

[4] Fisher, R.A (1934).The effects of methods of ascertainment upon the estimation of frequencies. Ann. Eugenics, 6, 13-25

[5] Rao, C.R (1965). On discrete distributions arising out of method of ascertainment, in clasasical and Contagious Discrete, G.P. Patil .ed ;Pergamon Press and Statistical publishing Society, Calcutta, pp-320-332.

[6] Rao, C. R(1985).Weighted distributions arising out of method of ascertainment , in A Celebration of Statistics, A.C.Atkinson&S.E.Fienberg, eds, Springer-Verlag,New York, Chapter 24,pp.543-569

[7] Dennis, B. and Patil, G. 1984. The gamma distribution and weighted multimodal gamma distributions as models of population abtidance.Math. Biosciences, 68,187-212.

[8] Gove,H.J.2003.Estimation and application of size-biased distributions in forestry, In Modeling Forest systems, A.Amaro,D. Reed and P.Soares CAB International Wallingford UK,201-212.

[9] Sandal C.E.1964. Estimation of the parameters of the gamma distribution, Technometrics,6, 405-414.

[10] Mir, K.A. 2009. Size-Biased Distributions and their Applications. Pakistan Journal of Statistics, 25(3), 283-294.

[11] P.C.VanDeusen, Forest Science,1986,32;146-148.

[12] J. Lappi, R.L Bailey, Forest Science,1987, 33;725-739.

[13] J.B.McDonald, Some Generalized functions for the size Distribution of Income,Econometrica,52(3), (1984), 647-663.

[14] J.B.McDonald and Y.J.Xu, A Generalized of the Beta distribution with Application, Journal of Econometrics, 69(2), (1995), 133-152.