On Size biased Generalized Beta distribution of second kind

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

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

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

Abstract: *In this paper, we develop the Size-biased form of the weighted Generalized Beta distribution of second kind (WGBD2); a particular case of the weighted Generalized Beta distribution of second kind, taking the weights as the variate values has been defined. The structural properties of Size-biased Generalized Beta distribution of second kind (SBGBD2) have been derived. The estimates of the parameters of Size- biased Generalized Beta distributions of second kind (SBGBD2) are obtained by employing a new method of moments. Also, a test for detecting the size-baisedness is conducted.*
Keywords: *Generalized Beta distribution of second kind, Beta function, Size-biased generalized beta distribution, Structural properties and moment estimator. Likelihood ratio test.*

I. Introduction

The probability density function (pdf) of the generalized beta distribution of second kind (GBD2) is given by:

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

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

a, b, p, q are positive real values.

The rth moment of generalized beta distribution of second kind is given by[7]:

$$E(X^r) = \frac{b^r \Gamma\left(p+\frac{r}{a}\right)\Gamma\left(q-\frac{r}{a}\right)}{\Gamma p \ \Gamma q} \tag{2}$$

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

$$E(X) = \frac{b \ \Gamma\left(p+\frac{1}{a}\right)\Gamma\left(q-\frac{1}{a}\right)}{\Gamma p \ \Gamma q} \tag{3}$$

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 and Proschan[1] . Many generalizations of beta distributions involving algebraic and exponential functions have been proposed in the literature; see in Johnson et al.[2] and Gupta and NadarSajah [3] for detailed accounts. The generalized beta distribution of second kind [6] (GBD2) 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 second 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.

II. Size-Biased Generalized Beta Distribution of second kind

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 [4] to model ascertainment bias, these were later formalized in a unifying theory by Rao [5]. 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.

If the random variable X has distribution $f(x;\theta)$, with unknown parameter θ, then the corresponding size –biased distribution is of the form

$$f^*(x;\theta) = \frac{x^c f(x;\theta)}{\mu_c'} \qquad (4)$$

$$\mu_c' = \int x^c f(x;\theta)dx \qquad \text{For continuous series} \qquad (5)$$

$$\mu_c' = \sum_{i=1}^{n} x^c f(x;\theta)dx \qquad \text{For discrete series.}$$

When c = 1 and 2, we get the size –biased and area biased - distributions respectively.

A size biased generalized beta distribution of second kind (SBGBD2) is obtained by applying the weights x^c, where c =1 to the weighted Generalized beta distribution.
We have from relation (1) and (5)

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

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

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

where $f^*(x;a,b,p,q)$ represents a probability density function. This gives the size –biased generalized beta distribution of second kind (SBGBD2) as:

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

$$(6)$$

where a, p, q are shape parameters and b is a scale parameter,

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

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

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

III. Structural properties of Size- biased Generalized beta distribution of second kind:

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

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

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

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

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

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

$$\mu'_r = \frac{b^r}{\beta\left(p+\dfrac{1}{a},q-\dfrac{1}{a}\right)}\beta\left(p+\dfrac{r}{a}+\dfrac{1}{a},q-\dfrac{r}{a}-\dfrac{1}{a}\right)$$

(7)

Using the equation (7), the mean and variance of the SBGBD2 is given by

$$\mu'_1 = \frac{b\,\beta\left(p+\dfrac{2}{a},q-\dfrac{2}{a}\right)}{\beta\left(p+\dfrac{1}{a},q-\dfrac{1}{a}\right)}$$

(8).

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

(9)

The Coefficient of variation of Size- biased Generalized Beta Distribution of second kind.

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

(10)

3.1 The mode of Size-biased generalized beta distribution of second kind is given as:
The probability distribution of Size-biased Generalized Beta distribution of second kind is:

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

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

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

(11)

Where C is a constant. Note that

$$\frac{\partial \ln f^*(x;a,b,p,q)}{\partial x} = \frac{a\,p\,b^a - a\,q\,x^a}{x(b^a+x^a)}$$

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

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

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

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

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

$$x_0 = b\left(\frac{p}{q}\right)^{\frac{1}{a}}$$

(12)

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

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

The harmonic mean (H) is given as:

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

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

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

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

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

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

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

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

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

(13)

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

4.1 New Method of Moment Estimators

Let X_1, X_2, X_3… Xn be an independent random samples from the SBGBD2 with weight c=1. The method of moment estimators are obtained by setting the raw 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.

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

(14)

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

(15)

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

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

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

(16)

Case 2. When p and b are fixed and a=1, then dividing equation (14) by (15), we have:

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

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

(17)

Case 3: When b and q are fixed and a=1, then dividing equation (14) by (15), we have:

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

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

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

(18)

ase 4. When p and q are fixed, b=1 then we can calculate the value of $\hat{a}$ estimator by numerical methods.

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

Let X_1, X_2, X_3… Xn be random samples can be drawn from generalized beta distribution of second kind or size-biased generalized beta distribution of second 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 second kind or size-biased generalized beta distribution of second kind the following test statistic is used.

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

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

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

$$\Delta = \left[\frac{\Gamma p\Gamma q}{b\Gamma\left(p+\dfrac{1}{a}\right)\Gamma\left(q-\dfrac{1}{a}\right)}\right]^{n} \prod_{i=1}^{n} x_i \qquad (19)$$

We reject the null hypothesis.

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

Equalivalently, we rejected the null hypothesis where

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

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] J.B.McDonald, *Some Generalized functions for the size Distribution of Income,Econometrica,52(3), (1984), 647-663.*
[7] J.B.McDonald and Y.J.Xu, *A Generalized of the Beta distribution with Application, Journal of Econometrics, 69(2), (1995), 133-152.*