Journal of Research Administration
Society of Research Administrators International

DEVELOPMENT OF THE SIZE-BIASED DISTRIBUTION FOR THE RAINBOW TYPE DISTRIBUTION

Sabir Ali Siddiqui [1], *, Shradha Dwivedi[2], Ghulam Mustafa[3], Irfan Siddiqui [4], Masood Alam [5], Sanjay Jain[6], Mustafa Kamal[7], and Showkat Ahmad Lone[8]

[1]Department of Mathematics and Sciences, Dhofar University, Salalah, Sultanate of Oman.

[2] IT Department (Mathematics), University of Technology and Applied Sciences (UTAS), Salalah, Sultanate of Oman.

[3]Islamia College Peshawar, Pakistan.

[4]Department of Management Studies, Sinhagad College of Engineering, Pune, India.

[5]Department of Mathematics, Sultan Qaboos University, Muscat, Sultanate of Oman.
[6]Department of Statistics, St. John's College, Agra, India.

[7]Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Dammam 32256, Kingdom of Saudi Arabia

[8]Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Riyadh, 11673, Kingdom of Saudi Arabia

Abstract: The present work is the extension on the previous work referred in [12], here the size-biased form of [12] has been developed and its mathematical properties have been obtained, that includes, **moment generating function (mgf), characteristic function and other important properties like Fisher's information and Shannon's entropy which gives the measure of uncertainty in the system.**

Key Words: Rainbow type **distribution, size-biased Rainbow Distribution (SBRD), Shannon's entropy, Fisher's information.**

1.Introduction

The search for the new probability distributions, their applications has been an interesting topics for the researchers, Bilal et al (2010) proposed a new statistical model which yields a variety of some well-known probability distributions including Exponential, Weibull and Rayleigh distributions. Mohammad and Kumar (2018), Studied generalized exponential distribution and obtained moments generating function for order statistics, Shah et al(2020), studied a new form of alpha skew distribution and obtained its properties. The present work also dealing with a new probability distribution that has been developed as a size biased distribution of the new distribution [12].

1.1 Size-Biasing

If the observations of a population have unequal probability of their selection in a sample and their selection depend on their size or weight and the probability of selecting observations is not equal but it is proportional to the size of the observations, the probability distribution for such populations are named as the size-biased distributions. The Size-biasing of some distribution is the method to introduce a weight on parent distribution in the form of variables. That is why size biased distributions sometimes known as weighted distributions also.

Size biased distribution was first applied by Warren(1975),later Van Deusen (1986) , Grosenbaugh & Director(1958) , Lappi and Bailey (1987) worked on many application part of size biased distributions. Further it was studied by Dennis and Patil (1984) . Gove (2003)reviewed some of the more recent results on size-biased distributions pertaining to parameter estimation in forestry, with special emphasis on the weibull family. Recent work on size biasing is done by Sabir et. al. (2016) in which size biasing of Mukharji – Islam(1983) was studied.

Method:

Suppose X is a non-negative variable with its pdf $f(x,\theta)$, where θ is the parameter. $X^*$is the weighted version of variable X and its distribution related to that of X, is called weighted distribution. The distribution of X^* is weighted by the value, or size of X, that is why we say that X^* has the X size-biased distribution.

Definition 1. Suppose X is discrete distribution with p.m.f, $p(x,\theta)$ then after size biasing its cdf will be given as,

$$dF^* = \frac{x\,p(x,\theta)}{E(X)} \quad ,x = 0,1,\dots,\theta \tag{1}$$

Definition 2. Suppose X is continuous distribution with pad $f(x,\theta)$, then its pdf after size biasing will be,

$$f^*(x) = \frac{xf(x,\theta)}{E(X)}, x \geq 0 \tag{2}$$

Here

$$E(X) = \int_0^\theta f(x,\theta)\,dx$$

This form of distribution was first introduced by Fisher (1934) and later used by Richard Arratia and Larry Goldstein (2019).

1.2. Basic Features of the Rainbow Type Probability Distribution

The cdf of distribution of a random variable X is given as,

$$F(x) = \begin{cases} \frac{x^2}{\theta^3}(3\theta - 2x), & 0 \le x \le \theta \\ 0 & , Otherwise \end{cases} \quad (3)$$

Here θ is the scale parameter.
The pdf in the equation above is

$$f(x) = \frac{6}{\theta^3}(x\theta - x^2) \quad (4)$$

Where $0 \le x \le \theta$ and $\theta \ge 0$

The reliability function is,

$$R(x) = 1 - \frac{x^2}{\theta^3}(3\theta - 2x) \quad (5)$$

Where $0 \le x \le \theta$ and $\theta \ge 0$

1.3. Development of the pdf of Size-Biased Rainbow Type Distribution (SBRD)

SBRD can be obtained by applying the definition of size biasing on Rainbow Type distribution [10],

since

$$E(X) = \int_0^\theta x f(x, \theta)\, dx$$

Now, for Rainbow distribution, x ranges between 0 to θ, and the first moment of Rainbow Type distribution is

$$E(X) = \frac{\theta}{2} \quad (6)$$

From equation (1) and equation (6), we find,

$$f^*(x) = \frac{x\left\{\frac{6}{\theta^3}(x\theta - x^2)\right\}}{\frac{\theta}{2}}$$

$$f^*(x) = \frac{12}{\theta^4} x^2(\theta - x)$$

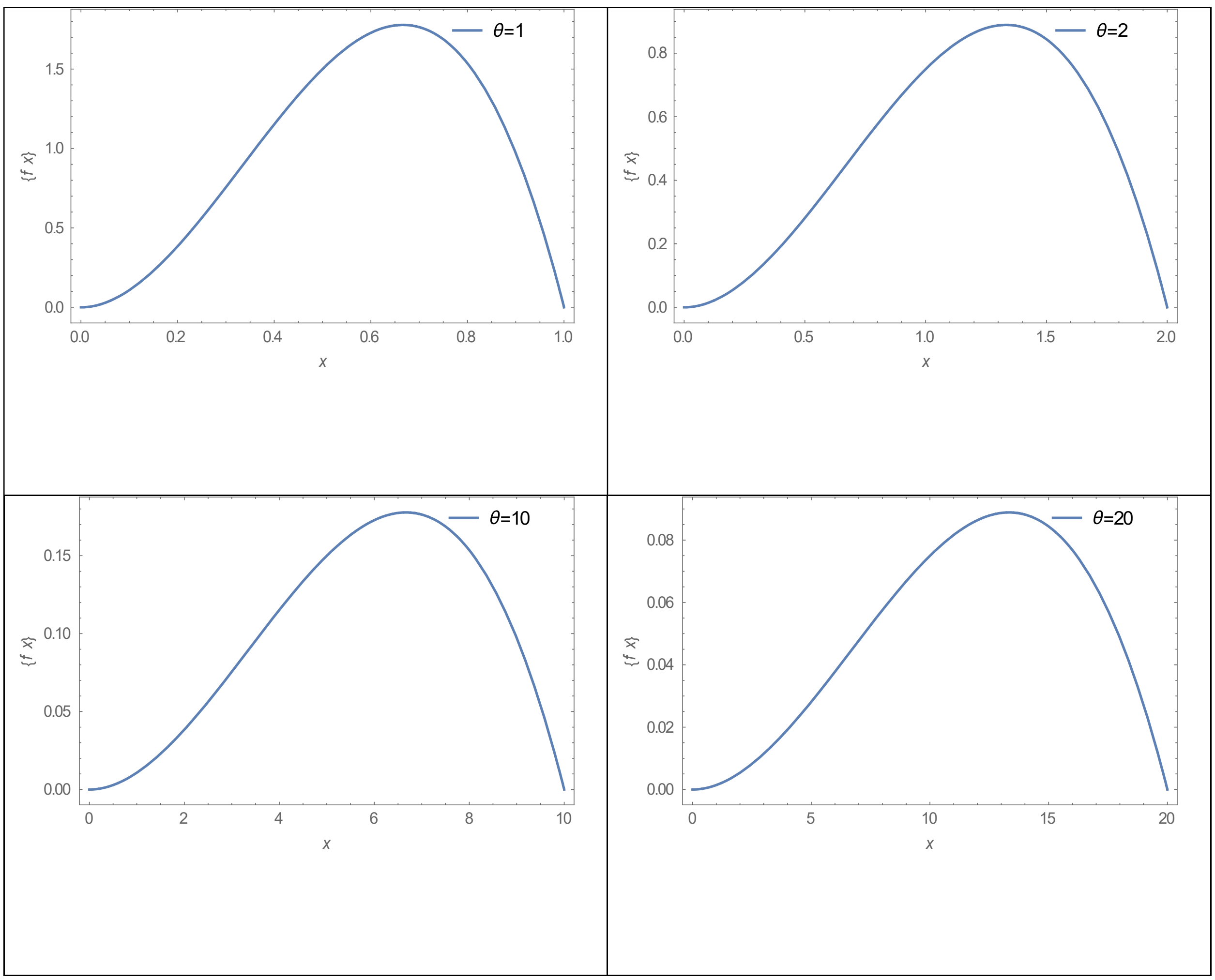
Now, it is easy to verify,

$$\int_0^\theta f^*(x, \theta)\, dx = \int_0^\theta \frac{12}{\theta^4} x^2(\theta - x)\, dx = 1$$

Hence, $f^*(x, \theta)$ represents the pdf of SBRD:

$$f^*(x) = \begin{cases} \frac{12}{\theta^4} x^2(\theta - x), & 0 \le x \le \theta, \quad \theta \ge 0 \\ 0 & Other\ wise \end{cases} \tag{7}$$

Graphs of pdf of SBRD for different values of θ

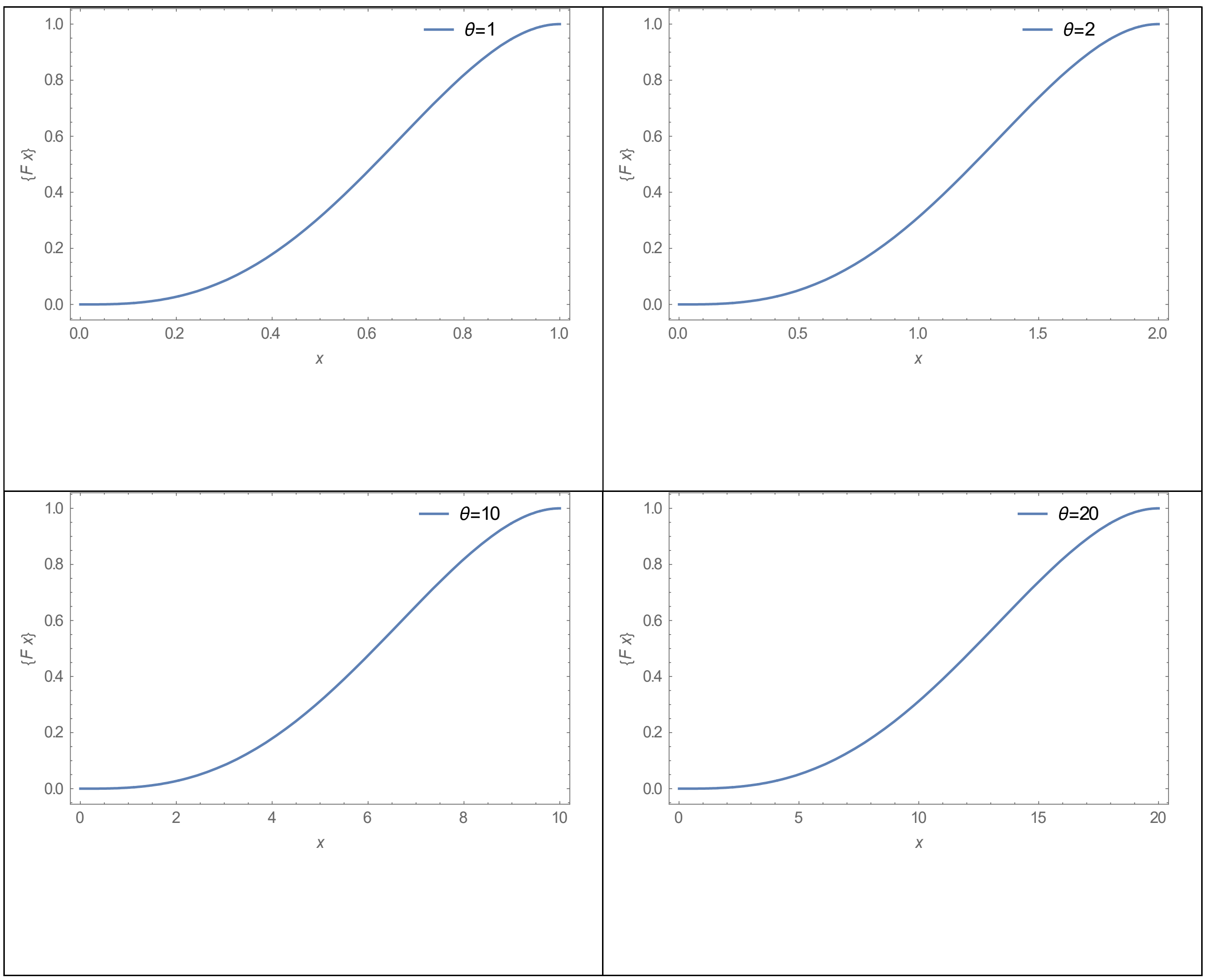


1.4 The cdf of SBRD

Following is the cdf obtained by using (7)

$$F(x) = \frac{x^3}{\theta^3}(4 - \frac{3}{\theta}x) \qquad (8)$$

Graphs of cdf of SBRD for different values of θ



1.5 Reliability and Hazard Rate Functions of SBRD

The Reliability and Hazard rate functions obtained by using (7) and (8),

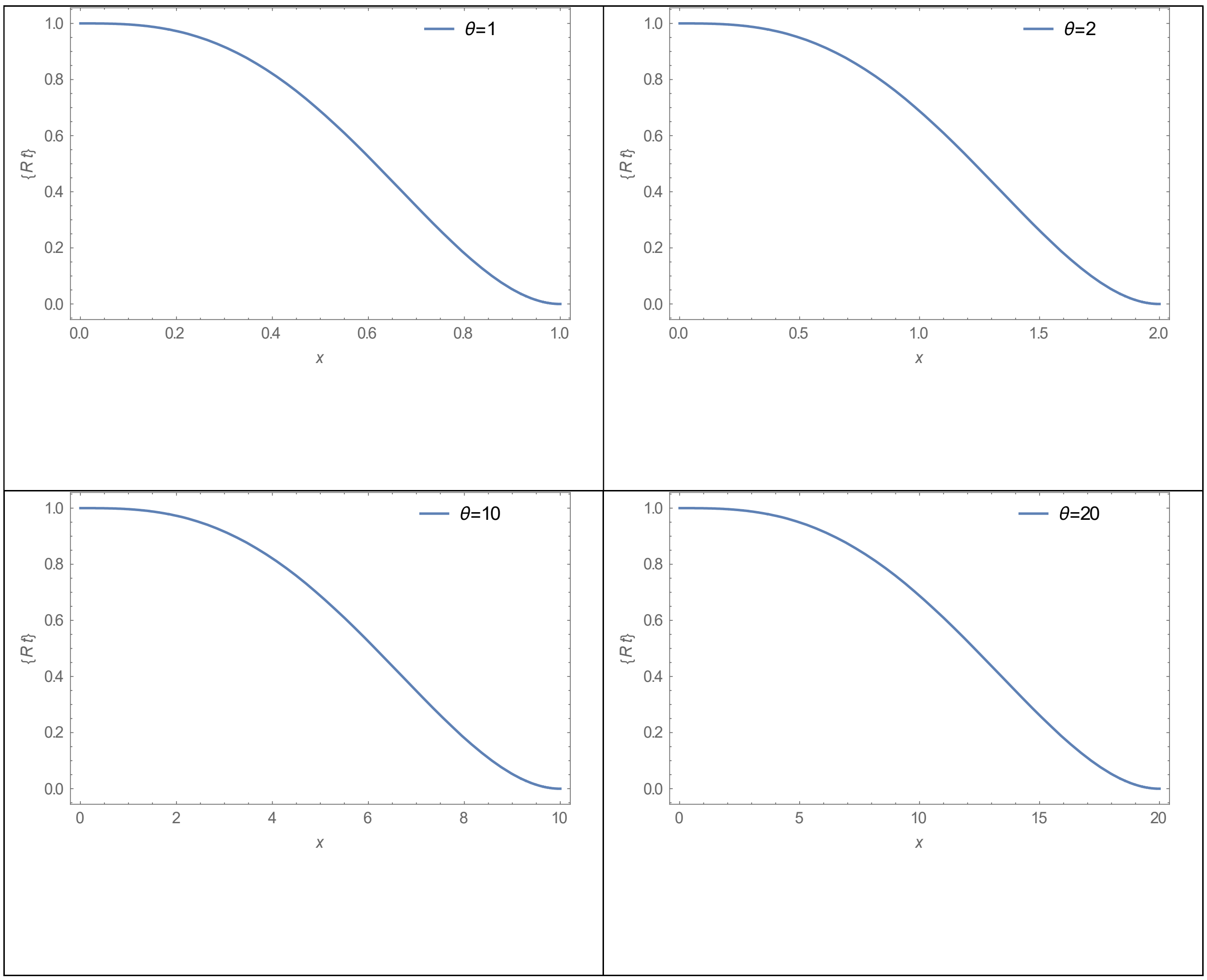
The Reliability Function, R(t) is,

$$R(t) = 1 - \frac{t^3}{\theta^3}\left(4 - \frac{3}{\theta}t\right) \qquad (9)$$

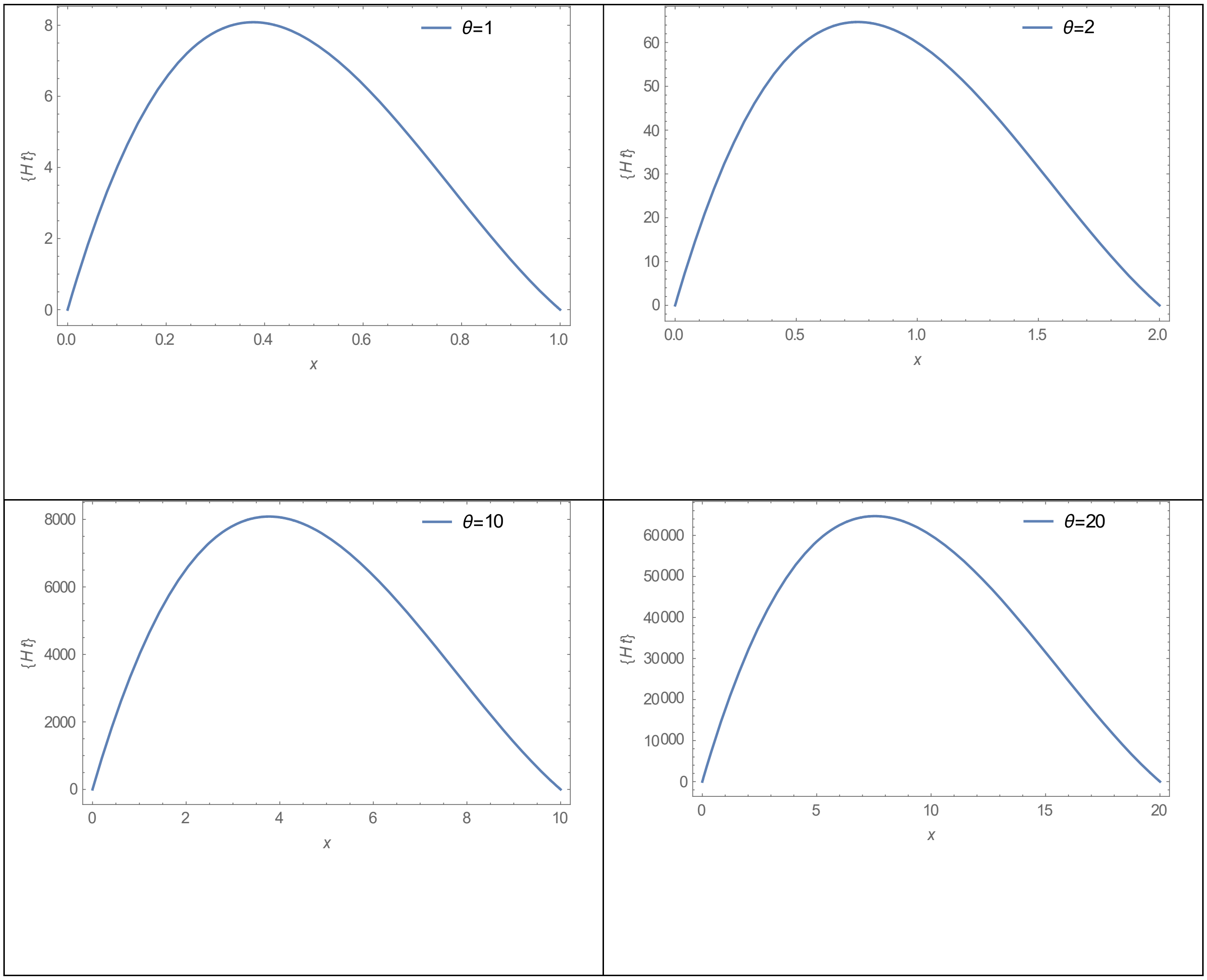
And the Hazard Rate function, H(t) is,

$$H(t) = \frac{12(\theta - x)}{x(4\theta - 3x)} \qquad (10)$$

Graphs of Reliability Function of SBRD for different values of θ



Graphs of Hazard Rate Function of SBRD for different values of θ



2. Mathematical Properties of Size-Biased Rainbow Distribution (SBRD)

Here, we derived some basic characteristics for SBRD, like mean , variance, moments, moment generating function (m.g.f) and characteristic function(cf).

2.1. Mean, Variance and Moments

The r^{th}- moments of SBRD, is

$$\mu_r' = E(x^r) = \int_0^\theta x^r f^*(x)\, dx$$

$$= \int_0^\theta x^r \frac{12}{\theta^4} x^2 (\theta - x)\, dx$$

$$= \frac{12}{\theta^4} \int_0^\theta x^{r+2} (\theta - x)\, dx$$

$$\mu_r^{'} = \frac{12}{\theta^4}\left[\frac{\theta^{r+4}}{(r+3)(r+4)}\right] \tag{11}$$

The mean and the variance of the SBRD are given as

Mean

$$Mean = \mu_1^{'} = \frac{3\theta}{5} \tag{12}$$

Variance:

For variance we need,

$$\mu_2^{'} = \frac{2\theta^2}{5}$$

So variance is,

$$Varinace\ = \frac{\theta^2}{25} \tag{13}$$

Moments

Other useful moments are also here,

$$\mu_3^{'} = \frac{2\theta^3}{7}$$

$$\mu_4^{'} = \frac{3\theta^4}{14}$$

The measure of skewness and Kurtosis of the distribution can be obtained by using above moments, as:

2.2. Moment Generating Function (mgf)

The Moment generating function of SBRD is defined as:

$$M_X(t) = Ee^{tx}$$

$$= \int_0^\theta e^{tx} f^*(x)dx$$

$$= \int_0^\theta e^{tx} \frac{12}{\theta^4} x^2(\theta - x)$$

$$M_X(t) = \frac{12}{\theta^4}\int_0^\theta (\theta x^2 - x^3)e^{tx}dx$$

$$M_x(t) = \frac{12}{\theta^4}\sum \frac{(t)^r}{r!}\left[\frac{\theta^{4+r}}{(3+r)(4+r)}\right] \tag{14}$$

2.3. Characteristic Function(cf)

The characteristic function of SBRD is given as:

$$\varphi_X(t) = Ee^{\iota t}$$

$$= \int_0^\theta e^{itx} f^*(x)dx$$

$$= \int_0^\theta e^{itx}\frac{12}{\theta^4}x^2(\theta - x)$$

$$\varphi_X(t) = \frac{12}{\theta^4}\int_0^\theta (\theta x^2 - x^3)e^{itx}dx$$

$$\varphi_X(t) = \frac{12}{\theta^4}\sum \frac{(\mathrm{i}t)^r}{r!}\left[\frac{\theta^{4+r}}{(3+r)(4+r)}\right] \tag{15}$$

2.4. Shannon's Entropy of SBRD

Shannon's entropy or the measure of uncertainty of a system has two definitions:
Definition 1: For a discrete random variable f , entropy is defined by;

$$H_q(f) = -\sum_{i=1}^{n} P(X = x)\log[P(X = x)] \tag{16}$$

It is obvious that $H_q(f) \geq 0$

Definition 2: Entropy of a continuous variable with pdf isdefined as:

$$H_q(f) = -\int_0^\infty f(x)\log f(x) = \mathrm{E}[-\,log\,f(x)] \tag{17}$$

above is true when integral exists: Shannon's entropy of size-biased Rainbow Type Distribution (where $0 < x \leq \theta$) can be given as:

$$H_q[f^*(\mathrm{x})] = \mathrm{E}[-\log f^*(x)]$$

$$= \mathrm{E}[-\log\left\{\frac{12}{\theta^4}x^2(\theta - x)\right\}]$$

$$= E[-log12 + 4log\theta - 2logx - log(\theta - x)]$$

$$H_q(f) = -log12 + 4log\theta - 2E(logx) - E[log(\theta - x)] \tag{18}$$

$$E(\log x) = \int_0^\theta \log x\, f^*(x)$$

$$= \int_0^\theta \log x \frac{12}{\theta^4}(\theta x^2 - x^3)dx$$

$$= \frac{12}{\theta^4}\left[\theta\int_0^\theta x^2 \log x - \int_0^\theta x^3 \log x\right]$$

$$E(\log x) = \log\theta - \frac{7}{12} \tag{19}$$

$$E[log(\theta - x)] = \int_0^\theta [log(\theta - x)]f^*(x)$$

$$= \int_0^\theta [log(\theta - x)]\frac{12}{\theta^4}(\theta x^2 - x^3)dx$$

$$= \frac{12}{\theta^4}\int_0^\theta (\theta x^2 - x^3)\, log(\theta - x)\, dx$$

$$E[log(\theta - x)] = \frac{13}{12} - \log\theta \tag{20}$$

Now by equ (16) and eqn (17) we have,

$$H_q(f) = -log12 + 4log\theta - 2\left[\log\theta - \frac{7}{12}\right] - \left[\frac{13}{12} - \log\theta\right]$$

$$H_q(f) = 3log\theta - log12 + \frac{1}{12} \tag{21}$$

2.5. Fisher's Information of SBRD

For a random variable X, Fisher Information about the parameter θ is given by,

$$I(\theta) = E\left[\frac{\partial}{\partial\theta}\log\{f^*(x)\}\right]$$

If log f(x, θ) is twice differential wrt θ, Fisher's information is given by,

$$I(\theta) = E\left[\frac{\partial^2}{\partial^2\theta}\log\{f^*(x)\}\right]$$

Now pdf of SBRD is given by:

$$f^*(x) = \frac{12}{\theta^4}x^2(\theta - x)$$

Here θ is a shape parameter.
Taking log on both sides of above pdf, we have

$$log\{f^*(x)\} = \log 12 - 4\log\theta + 2\log x + \log(\theta - x) \tag{22}$$

Differentiating eqn (8) partially w.r.t. θ , we get

$$\frac{\partial}{\partial\theta}\log\{f^*(x)\} = \frac{1}{\theta - x} - \frac{4}{\theta}$$

$$-E\left[\frac{\partial^2}{\partial^2\theta}\log\{f^*(x)\}\right] = -\frac{4}{\theta^2} + E\left[\frac{1}{(\theta - x)^2}\right] \tag{23}$$

$$E\left[\frac{1}{(\theta - x)^2}\right] = E[(\theta - x)^{-2}] = \frac{6}{\theta^2}[3 - 2\ln\theta]$$

Now, Fisher's Information of SBND is given as;

$$-E\left[\frac{\partial^2}{\partial^2\theta}\log\{f^*(x)\}\right] = \frac{14 - 12\ln\theta}{\theta^2} \tag{24}$$

3. Estimation of parameter

The maximum likelihood estimator of the parameter is somehow complicated therefore the method of moments could be preferred for estimating the parameter,

3.1 Estimation by the method of moments

Following relation will provide the estimated value for the parameter θ

$$\mu_r' = \frac{12}{\theta^4}\left[\frac{\theta^{r+4}}{(r+3)(r+4)}\right] = m_r' \tag{25}$$

3.2. Bayesian Estimator of SBRD for Uniform Prior

Likelihood function of SBRD is:

$$L(x;\theta) = \frac{12^n}{\theta^{4\mathrm{n}}}\prod_{i=0}^{n} x^2 \prod_{i=0}^{n}(\theta - x)$$

Since $0 < x \leq \theta$, we assume Uniform prior about θ. Uniform prior $g(\theta) = 1$, and

$$f(\theta \backslash x) = \frac{Lg(\theta)}{\int_0^\theta Lg(\theta)d\theta}$$

$$f(\theta \backslash x) = \frac{\frac{12^n}{\theta^{4\mathrm{n}}}\prod_{i=0}^{n} x^2 \prod_{i=0}^{n}(\theta - x) \times 1}{\int_0^\theta \frac{12^n}{\theta^{4\mathrm{n}}}\prod_{i=0}^{n} x^2 \prod_{i=0}^{n}(\theta - x) \times 1 d\theta} \tag{26}$$

is the posterior distribution for SBRD.

4. Test for Size-Biasedness of Size-Biased Rainbow Distribution (SBRD)

Let $X1, X2, X3, \dots Xn$ is random sample of size n drawn from size-biased New distribution. We test the hypothesis, $H0 = f(x,\theta)$ against $H1 = f^*(x,\theta)$. To check whether sample of size n is drawn from the New distribution or size-biased New distribution, the following test is used.

$$\Delta = \frac{H_1}{H_0} = \prod_{i=1}^{n}\left[\frac{f^*(x,\theta)}{f(x,\theta)}\right]$$

$$\Delta = \frac{H_1}{H_0} = \prod_{i=1}^{n}\left[\frac{\frac{12}{\theta^4}x^2(\theta - x)}{\frac{6}{\theta^3}(x\theta - x^2)}\right]$$

$$\Delta = \frac{H_1}{H_0} = \prod_{i=1}^{n}\left[\frac{\frac{12}{\theta^4}x^2(\theta - x)}{\frac{6}{\theta^3}x(\theta - x)}\right]$$

$$\Delta = \frac{H_1}{H_0} = \left(\frac{2}{\theta}\right)^n \prod_{i=1}^{n} x_i$$

Null hypothesis will be rejected, if we have

$$\left(\frac{2}{\theta}\right)^n \prod_{i=1}^{n} x_i > k$$

or equivalently, we reject the Null hypothesis where,

$$\Delta^* = \prod_{i=1}^{n} x_i > k^*$$

$$k^* = k\left(\frac{2}{\theta}\right)^n > 0$$

5.Conclusion

The aim of this paper is to introduce a new size-biased distribution based on the Rainbow Type Probability Distribution that we called here SBRD, this distribution is based on a new original distribution. Thereafter we discussed it's structural properties, that included mean, variance, moments of the distribution. The generating functions are also derived. Shannon's entropy and Fisher's information matrix has been calculated. We have also obtained estimates of parameters using classical and Bayes' method.

Acknowledgment

The authors are thankful to the colleagues and the research mates who shared their valuable comments and ideas in developing this work.

References

1. Arratia, R., Goldstein, L., & Kochman, F. (2019). Size bias for one and all. *Https://Doi.Org/10.1214/13-PS221*, *16*(none), 1–61. https://doi.org/10.1214/13-PS221
2. Bilal Ahmad Bhat, T. A. Raja, M. S. Pukhta and R. K. Agnihotri (2010): On A Class of Continuous Type Probability Distributions With Application ,Int. J. Agricult. Stat. Sci., Vol. 6, No. 1, pp. 147-156, 2010 ISSN : 0973-190.
3. Dennis, B., & Patil, G. P. (1984). The gamma distribution and weighted multimodal gamma distributions as models of population abundance. *Mathematical Biosciences*, *68*(2), 187–212. https://doi.org/10.1016/0025-5564(84)90031-2.
4. Fisher, RA (1934). The Effect of Methods of Ascertainment Upon The Estimation of Frequencies. *Annals of Eugenics*, *6*(1), 13–25. https://doi.org/10.1111/J.1469-1809.1934.TB02105.

5. Gove, J. H. (2003). Moment and maximum likelihood estimators for Weibull distributions under length- and area-biased sampling. *Environmental and Ecological Statistics. 10: 455-467.*, *10*, 455–467. https://www.fs.usda.gov/research/treesearch/6975
6. Grosenbaugh, L. R., & Director, P. A. B. (1958). Point-Sampling and Line-Sampling Probability Theory, Geometric Implications, Synthesis. *USDA Forest Service, Southern Forest Experiment Station, Occasional Paper 160*. https://www.fs.usda.gov/research/treesearch/2458
7. Lappi, J., & Bailey, R. L. (1987). Estimation of the Diameter Increment Function or Other Tree Relations Using Angle-Count Samples. *Forest Science*, *33*(3), 725–739. https://doi.org/10.1093/FORESTSCIENCE/33.3.725
8. Mohammad Samir Farooqi and Devendra Kumar (2018): Moment Generating Functions of Generalised Exponential Distribution Based on Lower Generalized Order Statistics , Int. J. Agricult. Stat. Sci. Vol. 14, No. 1, pp. 165-174, 2018 ISSN : 0973-1903.
9. Mukherjee and Islam(1983), A finite range distribution of failure times, [*Nav. Res. Log. Quart. Vol.* **30,** 487–491 (1983)].
10. Richard Arratia ,Larry Goldstein and Fred Kochman (2019) "Size bias for one and all." Probab. Surveys 16 1 - 61, 2019. https://doi.org/10.1214/13-PS221
11. S. Shah, P. J. Hazarika and S. Chakraborty (2020): A NEW ALPHA SKEW LAPLACE DISTRIBUTION : PROPERTIES AND ITS APPLICATIONS, Int. J. Agricult. Stat. Sci. Vol. 16, No. 1, pp. 1-10, 2020 ISSN : 0973-1903 .
12. Sabir Ali Siddiqui, S., Dwivedi, S., Kamal, M., Siddiqui, I., Alam, M., & Jain, S. (2022). A Rainbow Type Probability Distribution(RTPD) To Study the Decay in Strength Among the Living Being. *Int. J. Agricult. Stat. Sci*, *18*, 1263–1269.
13. Sabir Ali Siddiqui,, Dwivedi, S., Dwivedi, P., & Siddiqui, I. (2016). Development of size-biased Mukherjee-Islam distribution. *International Journal of Pure and Applied Mathematics*, *107*(2), 505–516. https://doi.org/10.12732/ijpam.v107i2.18.
14. Van Deusen, P. C. (1986). Fitting Assumed Distributions to Horizontal Point Sample Diameters. *Forest Science*, *32*(1), 146–148. https://doi.org/10.1093/FORESTSCIENCE/32.1.146.
15. Warren, W. G. (1975). Statistical Distributions in Forestry and Forest Products Research. *A Modern Course on Statistical Distributions in Scientific Work*, 369–384. https://doi.org/10.1007/978-94-010-1845-6_27.

Mukheerji, S. P. and Islam, A. (1983):"A Finite Range Distribution of Failures Times", Naval Research LogisticsQuarterly, Vol. 30, pp. 487 – 491.Siddiqui, S.A., Dwivedi, S., Dwivedi, P *(2) (PDF) Discrete Mukherjee-Islam Distribution and Its Characterization*. Available from: https://www.researchgate.net/publication/351346025_Discrete_Mukherjee-Islam_Distribution_and_Its_Characterization [accessed May 04 2023].