Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 26 (2023), No. 4, pp. 595–600
DOI : 10.47974/JIM-1464

Multi-double truncated continuous distribution

Kawther F. Al-Hasan *
Kareema Abad Al-Kadim †
Department of Mathematics
College of Education for Pure Sciences
Babylon University
Babil
Iraq

Abstract

Our life contains many scientific, health, and social problems, which are complex and difficult for easy modeling. Therefore, most researchers resorted to studying distributions in different forms to be able to represent these problems, and from these new forms of distributions are called truncated distributions. This paper, suggests a new formula to find the multiple truncations for one or more intervals to a double truncated continuous distribution. Some statistical properties of multi–double truncated are introduced. Weibull distribution is applied for two intervals on the double truncation and introduced some of its statistical properties.

Subject Classification: *62H10.*

Keywords: *Truncation data, Weibull distribution, Hazard function, Multi-double truncated, Moment generating function.*

1. Introduction

Statistical analysis is very important to many fields, including but not limited to prediction and inferences about phenomena in the real world. Several classical probability distributions have been widely used over the past decades for modeling data in various areas. Recently, authors have been attentive to introducing a new probability model adding an extra parameter or two to the classical distributions, which are found to be more flexible in modeling data. Truncation is considered a statistical phenomenon that has been shown to occur in a wide range of applications, including survival analysis, epidemiology, and economics. In addition, a truncation method is an important approach to deleting

* *E-mail:* **pure.kawther.fa@uobabylon.edu.iq (Corresponding Author)**
† *E-mail:* **kareema.kadim@yahoo.com**

intervals from a domain of the random variable which contains incomplete data. Galton (1897) who is the first to use analyzed truncated samples corresponding to the speeds of American trotting horses[1]. Reyah and Kareema the truncated Rayleigh Pareto distribution have been studied in 2020[2]. In (2018), Mohamed et. al.[3] introduced a continuous distribution within [0,1] represented by the truncated Fréchet -Pareto distribution. While (2020) (Mohammed et. al.)[4] a new family of upper truncated distributions. Mohie study on mid-truncated [5]. Zaninetti, presented in 2014, truncated on Pareto[6]. Kawther et. al.)[7], in (2022) study on estimation of one-double Rayleigh and Weibull truncated by considering real data of the infected with COVID-19 in the city of Babylon-Iraq. Data was collected for the period of their infection with the virus until their death at the interval (1/10/2020-30/12/2020). The subject of the research is very useful for representing natural phenomena, which cannot be followed up with all intervals of their growth or development. Therefore we studied the general formula that represents multiple cuts. We introduced a new formula to multi-double intervals truncated from the original distribution and applied on the two-double weibull truncated distribution.

2. Construction of Multi-Double Truncated for Continuous Distribution

Let X be a random variable that has the probability density function, $\mathbb{f}_X(X)$, over $(-\infty, \infty)$, and the cumulative distribution function $\mathbb{F}_X(X)$. such that $x \in (a_1, b_1) \cup (a_2, b_2) \cup \dots\dots\dots \cup (a_n, b_n)$. In this section, we derived the probability density function of the random variable X_T, after multiple double truncations, $\mathbb{g}(X_T)$ the observations of X_T to (a_i, b_i). truncation, that is: deleting some values of the domain of the random variable.

Let $(a_1, b_n) = (-\infty, b_1) \cup (a_2, b_2) \cup \dots\dots \cup (a_n, \infty)$, then the probability density function of a multi-double truncated continuous distribution is:

$$\mathbb{g}_n(X_T) = \frac{1}{n}\left(\frac{\mathbb{f}(x_T)}{\mathbb{F}(b_1)-\mathbb{F}(a_1)} + \frac{\mathbb{f}(x_T)}{\mathbb{F}(b_2)-\mathbb{F}(a_2)} + \cdots + \frac{\mathbb{f}(x_T)}{\mathbb{F}(b_n)-\mathbb{F}(a_n)}\right)$$
$$= \left\{\frac{1}{n}\sum_{i=1}^{n} \frac{\mathbb{f}(\mathrm{x_T})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)} \qquad when \quad a_i < \mathrm{x_T} < b_i \right. \tag{1}$$

Since that $\mathbb{f}(\mathrm{X_T})$ is a probability density function over $(-\infty, \infty)$, $-\infty < x_T < \infty$, $n, \mathbb{F}(b_i) - \mathbb{F}(a_i) > 0$, then $\mathbb{g}_\mathrm{n}(X_T) > 0$ for all $x_T \in (-\infty, \infty)$.And $\int_{-\infty}^{\infty} \mathbb{g}_n(X_T)\, dx_T = 1$. And the cumulative distribution of multi-double-truncated is:

$$\mathbb{G}_n(X_T) = P(X \le x) = \sum_{i=1}^{n}\left(\frac{\mathbb{F}(x_T) - \sum_{j=1}^{n}\mathbb{F}(a_j) + \sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{n(\mathbb{F}(b_i)-\mathbb{F}(a_i))}\right) \tag{2}$$

2.1 *Other Characteristics of the Multi-Double Truncated Distribution*

The survival function $\mathbb{S}_n(X_T) = 1 - \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbb{F}(x_T) - \sum_{j=1}^{n}\mathbb{F}(a_j) + \sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)$ (3)

Hazard function $\mathbb{H}_n(X_T) = \frac{\mathbb{g}_n(X_T)}{\mathbb{S}_n(X_T)} = \dfrac{\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{f}(x_T)}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}}{1-\frac{1}{n}\sum_{i=1}^{n}\left(\frac{(\mathbb{F}(x_T) - \sum_{j=1}^{n}\mathbb{F}(a_j) + \sum_{j=1}^{n}\mathbb{F}(b_{j-1}))}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)}$ (4)

Reverse hazard function $\widetilde{\mathbb{H}}_n(X) = \frac{\mathbb{g}_n(X)}{\mathbb{G}_n(X)} = \frac{\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{f}(x_T)}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}}{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\mathbb{F}(x_T)-\sum_{j=1}^{n}\mathbb{F}(a_j)+\sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)}$ (5)

The r^th^ moment about zero is;

$$E(X_T{}^r) = \frac{1}{n}\sum_{i=1}^{n}\int_{a_i}^{b_i} x^r \mathbb{g}_n(x_T)\, dx_T = \frac{1}{n}\left(\frac{\int_{a_1}^{b_1} x^r \mathbb{f}(x_T)\, dx_T}{\mathbb{F}(b_1)-\mathbb{F}(a_1)} + \frac{\int_{a_2}^{b_2} x^r \mathbb{f}(x_T)\, dx_T}{\mathbb{F}(b_2)-\mathbb{F}(a_2)} \dots + \frac{\int_{a_n}^{b_n} x^r \mathbb{f}(x_T)\, dx_T}{\mathbb{F}(b_n)-\mathbb{F}(a_n)}\right) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{b_i{}^r\mathbb{F}(b_i)- a_i{}^r\mathbb{F}(a_i)-r\int_{a_i}^{b_i} x^{r-1}\mathbb{F}(x_T)\, dx_T}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right) \tag{6}$$

The r^th^ moment about the mean is;

$$E(X_T-\mu)^r = \int_{-\infty}^{\infty}(X_T-\mu)^r \mathbb{g}_n(x_T)\, dx_T = \sum_{j=0}^{r}(-1)^{r-j}\binom{r}{j}(\mu)^{r-j}\mu_j'\ dx_T$$
$$= \sum_{j=0}^{r}(-1)^{r-j}\binom{r}{j}(\mu)^{r-j}\sum_{i=1}^{n}\left(\frac{(b_i)^j\mathbb{F}(b_i)-(a_i)^j\mathbb{F}(a_i)-j\int_{a_i}^{b_i}(x_T)^j\mathbb{F}(x_T)\, dx_T}{n(\mathbb{F}(b_i)-\mathbb{F}(a_i))}\right) \tag{7}$$

2.2 *Order Statistics, Moment Generating, and Characteristic Functions for Multi-Double Truncated*

Let $X_{T_1}, X_{T_2}, \dots, X_{T_m}$ be the independent and identical distribution of random samples with a probability of density function and cumulative distribution function. Then $X_{T[1]} \le X_{T[2]} \le \dots \le X_{T[m]}$, be the order statistic, where $X_{[i]}$ are the X_i arranged in order of increasing magnitude. If the random variable X_T has multi- double truncated probability function as equation (1), then

$$\mathbb{g}_{n[j]}(X_T) = \frac{m!}{n(j-1)!(m-j)!}\sum_{i=1}^{n}\frac{\mathbb{f}(X_T)}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{F}(x_T)-\sum_{j=1}^{n}\mathbb{F}(a_j)+\sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)^{j-1}\left(1-\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{F}(x_T)-\sum_{j=1}^{n}\mathbb{F}(a_j)+\sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)^{m-j} \tag{8}$$

$$\mathbb{g}_{n[1]}(X_T) = \sum_{i=1}^{n}\frac{m\,\mathbb{f}(x_T)}{n(\mathbb{F}(b_i)-\mathbb{F}(a_i))}\left(1-\frac{1}{n}\sum_{i=1}^{n}\frac{\mathbb{F}(x_T)-\sum_{j=1}^{n}\mathbb{F}(a_j)+\sum_{j=1}^{n}\mathbb{F}(b_{j-1})}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)^{m-j} \tag{9}$$

The moment generating and characteristics functions are:

$$\mathcal{M}_{x_T}(t) = \int_{a_i}^{b_i} e^{t x_T}\, \mathbb{g}_N(x_T)\, dx_T$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{e^{b_i t}\ \mathbb{F}(b_i)- e^{a_i t}\ \mathbb{F}(a_i)-t\int_{a_i}^{b_i} e^{t\chi_T}\ \mathbb{F}(x_T)\, dx_T}{\mathbb{F}(b_i)-\mathbb{F}(a_i)}\right)$$
$$\Phi_{x_T}(t) = \frac{1}{n}\sum_{j=1}^{n}\left(\frac{\Phi^{b_j t}\ \mathbb{F}(b_j)- \Phi^{a_j t}\ \mathbb{F}(a_j)-t\int_{a_j}^{b_j}\Phi(x_T)\,\mathbb{F}(x_T)\, dx_T}{\mathbb{F}(b_j)-\mathbb{F}(a_j)}\right) \tag{10}$$

3. Two-Double Truncated Weibull Distribution

Weibull distribution is one of the most essential distributions because it is usually used in reliability theory, such as in manufacturing and electrical engineering, in the study distribution of lifetime[8]. The probability and cumulative functions of Weibull distribution are:

$$f(X)=\frac{\delta}{\omega^{\delta}}\,x^{\delta-1}\;e^{-\left(\frac{x}{\omega}\right)^{\delta}}\;I_{(0,\infty)}(x)\;,\;\mathrm{F}(X)=1-\;e^{-\left(\frac{x}{\omega}\right)^{\delta}},\;x>0\quad\delta\,,\omega>0\qquad(12)$$

3.1 *Two Double Truncated Probability Function of Weibull distribution*

The probability function of two intervals truncation of Weibull distribution double type truncated is;

$$\mathbb{g}_3(\,x_T)=\frac{\frac{\delta}{\omega^{\delta}}x_T{}^{\delta-1}\;e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}}{3\left(e^{-\left(\frac{a1}{\omega}\right)^{\delta}}-e^{-\left(\frac{b1}{\omega}\right)^{\delta}}\right)}+\frac{\frac{\delta}{\omega^{\delta}}x_T{}^{\delta-1}\;e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}}{3\left(e^{-\left(\frac{a2}{\omega}\right)^{\delta}}-e^{-\left(\frac{b2}{\omega}\right)^{\delta}}\right)}+\frac{\frac{\delta}{\omega^{\delta}}x_T{}^{\delta-1}\;e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}}{3\left(e^{-\left(\frac{a3}{\omega}\right)^{\delta}}-e^{-\left(\frac{b3}{\omega}\right)^{\delta}}\right)}\qquad(13)$$

$$-\infty<x_T<b_1,\;a_2<x_T\;<\;b_2,\;\;a_3<x_T<\infty,\text{respectively.}$$

(a) That is, the intervals $(b_1,a_2),(b_2,a_3)$, are deleted from the original interval for distribution. Two intervals (b_1,a_2) = (2,2.5) and (3, 4), (5, 5.5) respectively, are truncated for different values of parameters $\delta\,,\omega$, when drawing the curves of PDF, which are presented in Figure 1 and Figure 2.

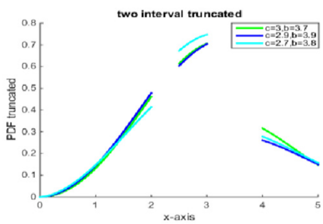

Figure 1

Two intervals truncated of PDF Weibull

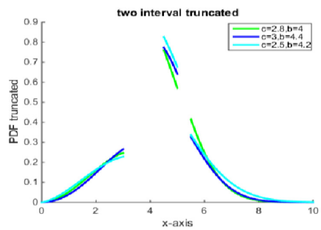

Figure 2

Two intervals truncated of PDF Weibull

3.2 *Two Double Truncated Distribution Function of Weibull distribution*

The cumulative distribution function for two intervals truncation of Weibull distribution double type truncated is;

$$\mathbb{G}_3(\mathrm{x_T})=\frac{1}{3}\left(\frac{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}}-e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}}{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}}\right)+\frac{1}{3}\left(\frac{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}}-e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}+e^{-\left(\frac{a_2}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}}{e^{-\left(\frac{a_2}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_2}{\omega}\right)^{\delta}}}\right)$$

$$+\frac{1}{3}\left(\frac{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}}-e^{-\left(\frac{x_T}{\omega}\right)^{\delta}}+e^{-\left(\frac{a_2}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}+e^{-\left(\frac{a_3}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_2}{\omega}\right)^{\delta}}}{e^{-\left(\frac{a_3}{\omega}\right)^{\delta}}-e^{-\left(\frac{b_3}{\omega}\right)^{\delta}}}\right)\qquad(14)$$

The cumulative function for two double truncated Weibull distribution for different values of the parameter is shown in Figure 3. Such that $(\mathrm{a}_1,\mathrm{b}_1)=(0,5)\;\;(\mathrm{a}_2,\mathrm{b}_2)=(10,15),(\mathrm{a}_3,\mathrm{b}_3)=(20,25\,)$.

3.3 *Survival Function of Two Double Truncated Weibull Distribution*

$$\mathbb{S}_3(x_T) = \left(\frac{3\left(e^{-\left(\frac{a_1}{\omega}\right)^{\delta}}\right)^2 - 6e^{-\left(\frac{a_1}{\omega}\right)^{\delta}} e^{-\left(\frac{b_1}{\omega}\right)^{\delta}} + 3\left(e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}\right)^2}{3\left(1 - e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}\right)}\right) - \left(\frac{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}} - e^{-\left(\frac{x_T}{\omega}\right)^{\delta}} + e^{-\left(\frac{a_2}{\omega}\right)^{\delta}} - e^{-\left(\frac{b_1}{\omega}\right)^{\delta}}}{3\left(e^{-\left(\frac{a_2}{\omega}\right)^{\delta}} - e^{-\left(\frac{b_2}{\omega}\right)^{\delta}}\right)}\right)$$

$$- \left(\frac{e^{-\left(\frac{a_1}{\omega}\right)^{\delta}} - e^{-\left(\frac{x_T}{\omega}\right)^{\delta}} + e^{-\left(\frac{a_2}{\omega}\right)^{\delta}} - e^{-\left(\frac{b_1}{\omega}\right)^{\delta}} + e^{-\left(\frac{a_3}{\omega}\right)^{\delta}} - e^{-\left(\frac{b_2}{\omega}\right)^{\delta}}}{3(e^{-\left(\frac{a_3}{\omega}\right)^{\delta}})}\right) \quad (15)$$

We note in this section, that two intervals from the original survival function of the Weibull distribution were deleted. The two deleted intervals are (1,5), and (7,9) as shown in Figure 4.

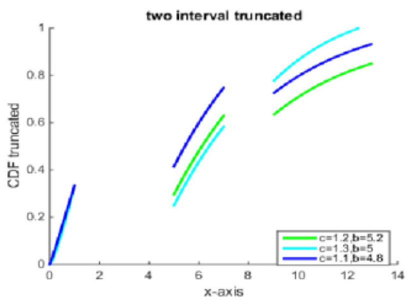

Figure 3

CDF Two double truncated Weibull

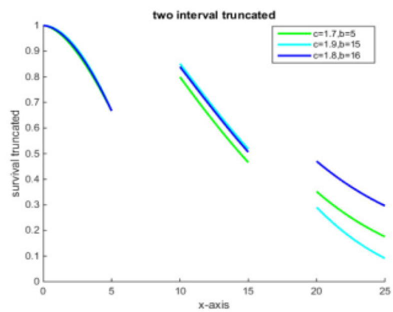

Figure 4

Two intervals truncated of survival Weibull

4. Conclusion

When studying many phenomena, some data was unclear, lost during the study, or deleted under certain circumstances for different intervals. In this case, we need to find a model to describe this problem. The subject of truncation is beneficial for representing natural phenomena, which cannot be followed up with all intervals of their growth or development, or that data cut off. Therefore this paper studies a formula to find the multiple deletions that describe multiple cuts. The probability function and cumulative function for multiple truncations, of type double truncated, have been constructed, and some of its statistical properties. Double truncated Weibull distribution with two intervals has been utilized and it is statistical properties.

References

[1] F. Galton, Examination into the registered speeds of American trotting House. with remarks on their value as Hereditary data, 29 (1898).

[2] Z. K. Reyah A.Kareema, Truncated Rayleigh Pareto distribution *J. Phys.: Conf. Ser*, (2020).

[3] E. A. Mohammad, S. F. Mohammad, [0,1] Truncated Fréchet-Pareto distributions *IQJOSS* 15(61), pp.229-244 (2019).

[4] H.Amal, A.Mohamed, A New family of upper truncated distributions: properties and estimation, *THAIJO*, 18(2), pp. 196-214 (2020).

[5] M.M. Mohie, M.M. Amein, A.M. Abd El-Raheem, On mid truncated distributions and its Applications, jarm. 5(2), pp.20–38 (2013).

[6] L. Zaninetti, M. Ferraro, On the truncated Pareto distribution with applications, *Centr. Eur. J. Phys*, 6, 1 (2002).

[7] Kawther Alhasan, Kareema Abad Al-Kadim, Estimation of multi-truncated Rayleigh and Weibull distributions with the application, *Int. J. Nonlinear Anal. Appl.* 13(1), pp. 3131-3140 (2022).

[8] D. Achim, E. Takeshi, Analysis of doubly truncated data, Springer briefs in statistics, 1st ed (2019).

[9] Shailendra Singh, Kuldeep Kumar, Estimating and testing the significance of correlation coefficient obtained from a truncated bivariate normal distribution, *Journal of Interdisciplinary Mathematics*, 19(2), pp. 245-251 (2016).

Received January, 2022

Revised March, 2022