American Review of Mathematics and Statistics
December 2023, Vol. 11, No. 2, pp. 17-27
ISSN: 2374-2348 (Print), 2374-2356 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v11n2a2
URL: https://doi.org/10.15640/arms.v11n2a2

Bayesian Approach to Estimating the Burr Type III Distribution of Three Parameters and Application to Arthritis Relief Times Data

Alexander K. Kim[1]

Abstract

In this paper, the Metropolis-Hastings algorithm is applied to find Bayes estimates of three parameters of the Burr Type III distribution. The Metropolis-Hastings algorithm is a Markov Chain Monte Carlo (MCMC) method to sample from distributions that cannot easily be sampled from by other means. Using this same algorithm, arthritis relief times data is analyzed.

Key words: Burr III distribution, Bayes estimates, Metropolis-Hastings algorithm, Markov Chain Monte Carlo (MCMC)

Mathematics Subject Classification: 62C10, 62F10, 62F15, 62H10

1. Introduction

In 1942, Burr [2] introduced numerous cumulative frequency functions, which included the two-parameter Burr III distribution. Although the distribution was intended for application to survival and lifetime data, it has been used in various statistical modeling projects due to its flexibility. For example, Mokhlis [12] applied it to discuss the reliability of a system in 2006. Also, Gove [6] used it to analyze rotated sigmoidal diameter distributions in 2008. In 1996, Lindsay [10] introduced and applied the four-parameter Burr III and XII distributions to forestry. The introduced cumulative distribution function (CDF) and probability density function (PDF) of the Burr III

[1] Seneca Valley Senior High School, Harmony, PA, Email: kimak.query@gmail.com

distribution are given by, respectively,

$$F(x) = [1 + (\frac{b}{x-a})^c]^{-k}$$
$$f(x) = \frac{kc}{b}(\frac{b}{x-a})^{c+1}[1 + (\frac{b}{x-a})^c]^{-(k+1)}$$

with location parameter $a < x$, scale parameter $b > 0$, and two shape parameters $c > 0$ and $k > 0$.

There are twelve different Burr distributions. It is noted that substituting X with $1/X$ in the Burr XII distribution equals the Burr III distribution. Burr [3] and Johnson [5] introduced the usefulness and characterized the properties of the Burr distributions. Abd-Elfattah [1] considered a Bayesian estimation for the Burr Type III distribution based on double censoring. Kim [8][9] estimated two parameters of the Burr III distribution using Bayesian methods including the dual generalized order statistics (DGOS), which was introduced by Burkschat [4]. Also, Kim [7] estimated the best linear unbiased estimators and the best linear invariant estimators for the location and scale parameters of the Burr III distribution based on the DGOS.

2. Markov Chain Monte Carlo Simulation

In this section, I will use a Markov Chain Monte Carlo (MCMC) simulation, specifically using the Metropolis-Hastings algorithm, to estimate the scale and shape parameters b, c, and k. It is assumed that $a = 0$ since it does not affect the shape of the distribution. In this case, performing an MCMC simulation requires a joint posterior density function of b, c, and k. So, the following CDF and PDF of the Burr III distribution are used to get the joint posterior density function:

$$F(x) = [1 + (\frac{b}{x})^c]^{-k}$$
$$f(x) = \frac{kc}{b}(\frac{b}{x})^{c+1}[1 + (\frac{b}{x})^c]^{-(k+1)} \tag{2.1}$$

Since each of the parameters are positive and the exact prior distributions are unknown, I decide to assign a gamma distribution with large variance as the prior to each parameter. Then, the prior distributions of parameters b, c, and k are

$$
\begin{aligned}
\pi_1(c) &= \frac{1}{\Gamma(\alpha_c)\beta_c^{\alpha_c}} c^{\alpha_c - 1} e^{-\frac{c}{\beta_c}} \\
\pi_2(k) &= \frac{1}{\Gamma(\alpha_k)\beta_k^{\alpha_k}} k^{\alpha_k - 1} e^{-\frac{k}{\beta_k}} \\
\pi_3(b) &= \frac{1}{\Gamma(\alpha_b)\beta_b^{\alpha_b}} b^{\alpha_b - 1} e^{-\frac{b}{\beta_b}} \qquad (2.2)
\end{aligned}
$$

Given samples $X = (x_1, ..., x_n)$, using (2.1) and (2.2), the joint posterior density function is given by

$$
\begin{aligned}
\pi(c, k, b|X) &= f(x_1) \cdots f(x_n)\pi_1(c)\pi_2(k)\pi_3(b) \\
&\propto [\prod_{i=1}^{n} \frac{kc}{b}(\frac{b}{x_i})^{c+1}[1 + (\frac{b}{x_i})^c]^{-(k+1)}] \\
&\times \; b^{\alpha_b - 1} e^{-\frac{b}{\beta_b}} c^{\alpha_c - 1} e^{-\frac{c}{\beta_c}} k^{\alpha_k - 1} e^{-\frac{k}{\beta_k}} \\
&= [\prod_{i=1}^{n} (\frac{1}{x_i})^{c+1}\{1 + (\frac{b}{x_i})^c\}^{-(k+1)}] e^{-\frac{b}{\beta_b} - \frac{c}{\beta_c} - \frac{k}{\beta_k}} \\
&\times \; b^{\alpha_b + nc - 1} c^{\alpha_c + n - 1} k^{\alpha_k + n - 1} \qquad (2.3)
\end{aligned}
$$

Now, the Metropolis-Hastings algorithm can be used to find Bayes estimates of the parameters in question. Given a parameter vector θ, the Metropolis-Hastings algorithm is defined as

1. Set an initial θ_0.
2. Propose candidate $\theta_t^* \sim q(y|\theta_{t-1})$.
3. Calculate acceptance probability $A = \min(1, \frac{\pi(\theta_t^*)}{\pi(\theta_{t-1})} \cdot \frac{q(\theta_{t-1}|\theta_t^*)}{q(\theta_t^*|\theta_{t-1})})$.
4. Accept θ_t^* with probability A, if accepted $\theta_t = \theta_t^*$, else $\theta_t = \theta_{t-1}$.

Since (2.3) causes exploding values due to $b^{\alpha_b + nc - 1} c^{\alpha_c + n - 1} k^{\alpha_k + n - 1}$, the following log posterior density function has to be used to make calculations feasible:

$$\begin{aligned}\ln \pi(c,k,b|X) \;\propto\; & [\sum_{i=1}^{n} \ln\{(\frac{1}{x_i})^{c+1}(1+(\frac{b}{x_i})^{c})^{-(k+1)}\}] \\ - \;& (\frac{b}{\beta_b}+\frac{c}{\beta_c}+\frac{k}{\beta_k}) + (\alpha_b + nc - 1)\ln b \\ + \;& (\alpha_c + n - 1)\ln c + (\alpha_k + n - 1)\ln k \qquad (2.4)\end{aligned}$$

For this simulation, the proposal density q will be a normal distribution with mean θ_{t-1} and variance $\sigma^2 = 25$. Since the normal distribution is symmetric, $q(\theta_{t-1}|\theta_t^*) = q(\theta_t^*|\theta_{t-1})$ and $A = \min(1, \frac{\pi(\theta_t^*)}{\pi(\theta_{t-1})})$. Now, the algorithm I am going to use has the following steps:

1. Set an initial $\theta_0 = (b_0, c_0, k_0)$.
2. Propose candidate $\theta_t^* \sim N(\theta_{t-1}, \sigma^2)$.
3. Calculate log acceptance probability $\ln A = \min(0, \ln[\pi(\theta_t^*)] - \ln[\pi(\theta_{t-1})])$.
4. Draw $u \sim U(0,1)$ and if $\ln u < \ln A$, then $\theta_t = \theta_t^*$, else $\theta_t = \theta_{t-1}$.

Since (2.3) causes exploding values due to $b^{\alpha_b+nc-1}c^{\alpha_c+n-1}k^{\alpha_k+n-1}$, the following log posterior density function has to be used to make calculations feasible:

$$\begin{aligned}\ln \pi(c,k,b|X) \;\propto\; & [\sum_{i=1}^{n} \ln\{(\frac{1}{x_i})^{c+1}(1+(\frac{b}{x_i})^{c})^{-(k+1)}\}] \\ - \;& (\frac{b}{\beta_b}+\frac{c}{\beta_c}+\frac{k}{\beta_k}) + (\alpha_b + nc - 1)\ln b \\ + \;& (\alpha_c + n - 1)\ln c + (\alpha_k + n - 1)\ln k \qquad (2.4)\end{aligned}$$

For this simulation, the proposal density q will be a normal distribution with mean θ_{t-1} and variance $\sigma^2 = 25$. Since the normal distribution is symmetric, $q(\theta_{t-1}|\theta_t^*) = q(\theta_t^*|\theta_{t-1})$ and $A = \min(1, \frac{\pi(\theta_t^*)}{\pi(\theta_{t-1})})$. Now, the algorithm I am going to use has the following steps:

1. Set an initial $\theta_0 = (b_0, c_0, k_0)$.
2. Propose candidate $\theta_t^* \sim N(\theta_{t-1}, \sigma^2)$.
3. Calculate log acceptance probability $\ln A = \min(0, \ln[\pi(\theta_t^*)] - \ln[\pi(\theta_{t-1})])$.
4. Draw $u \sim U(0,1)$ and if $\ln u < \ln A$, then $\theta_t = \theta_t^*$, else $\theta_t = \theta_{t-1}$.

To execute this algorithm, it is assumed that $a = 0$, $b = 10$, $c = 2$, and $k = 30$. Figure 1 displays the PDF of the proposed Burr III distribution. The CDF is

$$F(x) = [1 + (\frac{10}{x})^2]^{-30} \tag{2.5}$$

Using the inverse CDF technique, 100 samples from the proposed distribution will be drawn. The inverse of (2.5) is

$$x = 10[F(x)^{-\frac{1}{30}} - 1]^{-\frac{1}{2}} \tag{2.6}$$

Since $0 \leq F(x) \leq 1$, 100 uniformly random values are drawn and substituted for $F(x)$ in (2.6). Such would yield 100 random values from the proposed Burr III distribution, which are used to estimate three parameters b, c, and k under the squared error loss function defined as $L(\phi, \hat{\phi}) = (\phi - \hat{\phi})^2$ for a parameter ϕ and an estimate $\hat{\phi}$. Under the squared error loss function, $\mathbb{E}[L]$ is minimized when the Bayes estimates $\hat{b}, \hat{c}$, and $\hat{k}$ are means of the posterior distributions.

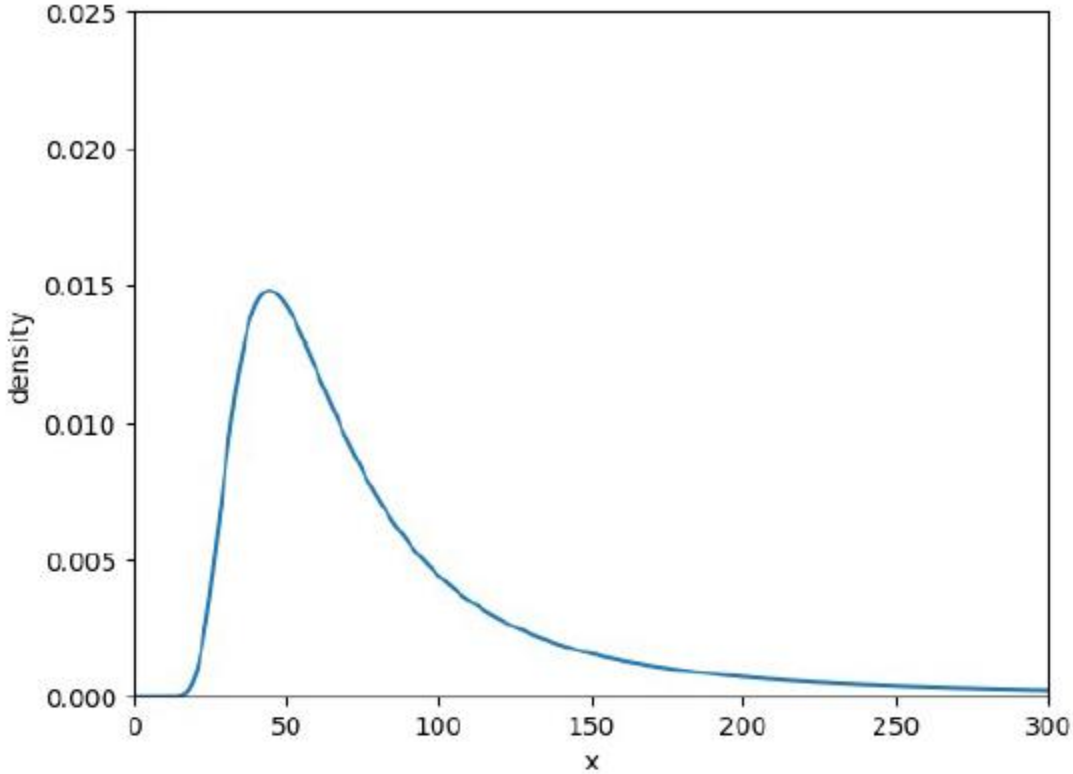

Figure 1: PDF of the Burr III distribution.

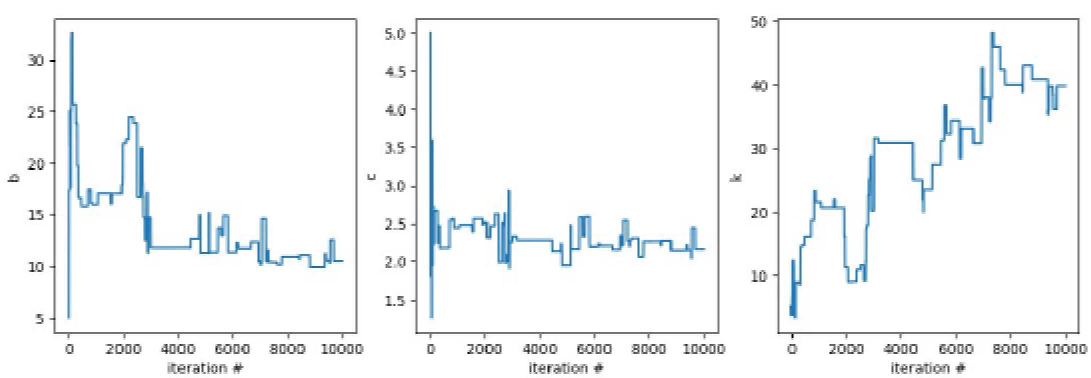

Figure 2: Iteration plots of each parameter from the simulation.

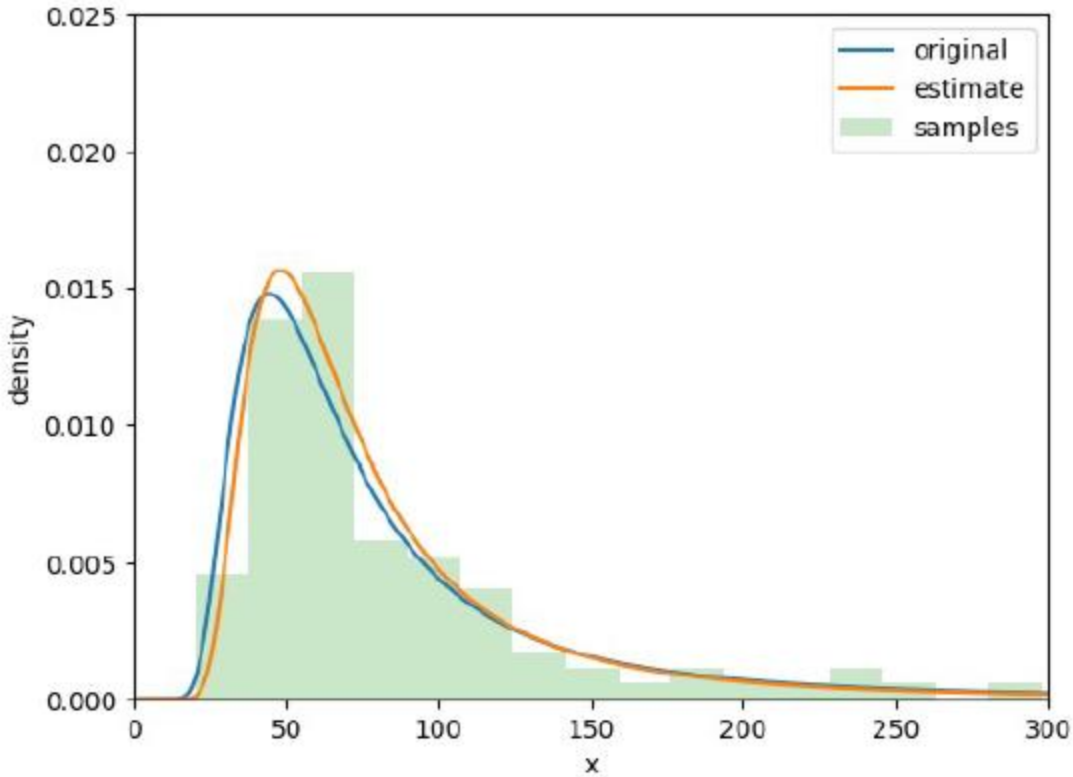

Figure 3: Original PDF and estimated PDF overlaid on random samples.

Figure 2 shows the iteration history of parameters b, c, and k. Assuming the first half of the iterations as the burn-in period, the Bayes estimates of parameters b, c, and k are $\hat{b} = 11.3357, \hat{c} = 2.2394$, and $\hat{k} = 37.0683$, which are similar to the original parameters $b = 10, c = 2, k = 30$. Figure 3 displays the overlay graph with the original density and the approximated curve. The Bayes estimates are seen to well approximate the original curve given only 100 random samples.

3. Real Data Analysis

In Section 3, I analyze arthritis relief times data presented by Wingo [13]. In a clinical trial, 50 arthritis patients received a fixed amount of an analgesic to test the effectiveness of the medication. Table 1 includes data collected from the clinical trial, which is well fit by the Burr XII distribution.

Once again, the Burr III distribution can be obtained from the Burr XII distribution by substituting X with $1/X$. Hence, the Burr III distribution well fits the data in Table 2.

0.70	0.84	0.58	0.50	0.55	0.82	0.59	0.71	0.72	0.61
0.62	0.49	0.54	0.72	0.36	0.71	0.35	0.64	0.85	0.55
0.59	0.29	0.75	0.53	0.46	0.60	0.60	0.36	0.52	0.68
0.80	0.55	0.84	0.70	0.34	0.70	0.49	0.56	0.71	0.61
0.57	0.73	0.75	0.58	0.44	0.81	0.80	0.87	0.29	0.50

Table 1: Arthritis relief times data well fit by the Burr XII distribution.

1.428571	1.190476	1.724138	2.000000	1.818182
1.219512	1.694915	1.408451	1.388889	1.639344
1.612903	2.040816	1.851852	2.777778	2.777778
1.408451	2.857143	1.562500	1.190476	1.818182
1.694915	3.448276	1.333333	2.173913	2.173913
1.666667	1.666667	1.666667	1.923077	1.470588
1.250000	1.818182	1.190476	2.941176	2.941176
1.428571	2.040816	1.785714	1.408451	1.639344
1.754386	1.369863	1.333333	2.272727	2.272727
1.234568	1.250000	1.149425	3.448276	2.000000

Table 2: Reciprocal arthritis relief times data well fit by the Burr III distribution.

Figure 4 shows the iteration history of parameters b, c, and k. Assuming the first half of the iterations as the burn-in period, the Bayes estimates of parameters b, c, and k are $\hat{b} = 0.6958, \hat{c} = 4.8269$, and $\hat{k} = 44.05816$. Figure 5 strongly supports the claim that the Burr III distribution well fits the arthritis relief times data.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

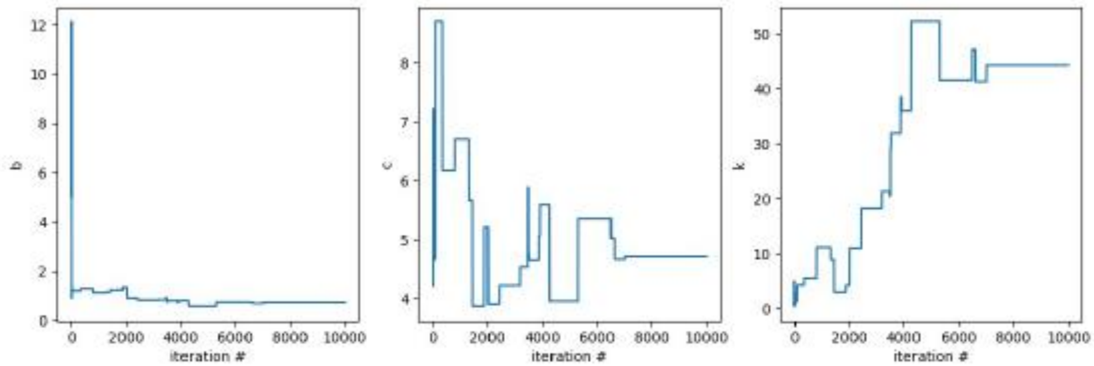

Figure 4: Iteration plots of each parameter from the real data analysis.

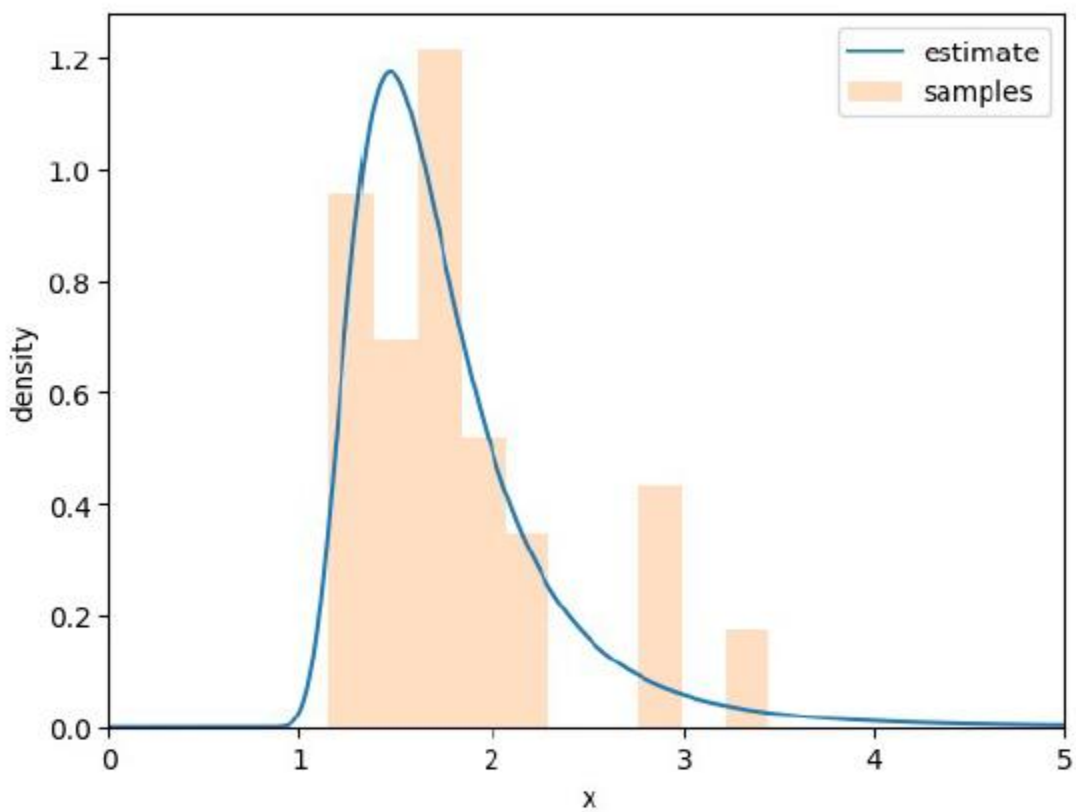

Figure 5: Estimated PDF overlaid on reciprocal arthritis relief times data.

References

[1] Abd-Elfattah, A. M. and Alharbey, A. H. (2012). Bayesian Estimation for Burr Distribution Type III Based on Trimmed Samples, *International Scholarly Research Network Applied Mathematics*, doi:10.5402/2012/0393.

[2] Burr, I. W. (1942). Cumulative frequency functions, *Annals of Mathematical Statistics*, 13, 215-232.

[3] Burr, I. W. and Cislak, P. J. (1968). On a general system of distributions: I Its curve-shape characteristics. II The sample median, *Journal of the American Statistical Association*, 63, 627-635.

[4] Burkschat, M., Cramer, E. and Kamps, U. (2003). Dual generalized order statistics, *METRON*, 1, 13-26.

[5] Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, John Wiley & Sons, New York, NY,USA, 2nd edition.

[6] Gove, J. H., Ducey, M. J., Leak, W. B. and Zhang, L. (2008). Rotated sigmoid structures in managed uneven-aged northern hardwood stands: A look at the Burr type III distribution, *Forestry*, 81, 21-36.

[7] Kim, C. and Kim, W. (2020). General linear regression model approach to estimate location and scale parameters of Burr type III distribution based o dual generalized order statistics, *Journal of Algebra and Applied Mathematics*, Vol. 18(2020), No. 1, 1-14

[8] Kim, W. and Kim, C. (2013). On dual generalized order statistics from Burr Type III distribution and its characterization, *Far East Journal of Mathematical Sciences*, Vol. 81, 1, 21-39

[9] Kim, C., Song, S, and Kim, W. (2016). Statistical Inference for Burr Type III Distribution On dual generalized order statistics and Real Data Analysis, *Applied Mathematical Sciences*, Vol. 10, 2016, no. 14, 683-695

[10] Lindsay, S.R., Wood, G.R. and Woollons, R.C. 1996 Modelling the diameter distribution of forest stands using the Burr distribution. J. Appl. Stat. 23, 609–619

[11] Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data, *Journal of Applied Meteorology*, 12, 275-280.

[12] Mokhlis, N. A. (2005). Reliability of a stress-strength model with Burr type III distributions, *Communications in Statistics-Theory and Methods*, 4, 1643-1657

[13] Wingo, D.R. , Maximum likelihood methods for tting the Burr type XII distribution to life test data, *Biometrics J.*, 25 (1983), 77-84