

A NEW METHOD FOR GENERALIZING BURR AND RELATED DISTRIBUTIONS

TANUJIT CHAKRABORTY*,^c — SUCHISMITA DAS** — SWARUP CHATTOPADHYAY***

(Communicated by Gejza Wimmer)

ABSTRACT. A new method has been proposed to generalize Burr-XII distribution, also called Burr distribution, by adding an extra parameter to an existing Burr distribution for more flexibility. In this method, the exponent of the Burr distribution is modeled using a nonlinear function of the data and one additional parameter. The models of this newly introduced generalized Burr family can significantly increase the flexibility of the former Burr distribution with respect to the density and hazard rate shapes. Families expanded using the method proposed here is heavy-tailed and belongs to the maximum domain of attractions of the Frechet distribution. The method is further applied to yield three-parameter classical Pareto and generalized exponentiated distributions which shows the broader application of the proposed idea of generalization. A relevant model of the new generalized Burr family has been considered in detail, with particular emphasis on the hazard functions, stochastic orders, estimation procedures, and testing methods are derived. Finally, as empirical evidence, the new distribution is applied to the analysis of large-scale heavy-tailed network data and compared with other commonly used distributions available for fitting degree distributions of networks. Experimental results suggest that the proposed Burr distribution with nonlinear exponent better fits the large-scale heavy-tailed networks better than the popularly used Marshall-Olkin generalization of Burr and exponentiated Burr distributions.

©2022
Mathematical Institute
Slovak Academy of Sciences

1. Introduction

The development of statistical distributions is one of the oldest research topic in the field of statistics. There has been a renewed interest in developing more flexible statistical distributions in recent decades. Since the seminal work by Karl Pearson in 1895 [37], several general methods have been developed for generating family of distributions. Pearson presented a systematic approach for generating statistical distribution to model non-symmetric type of data using differential equation. The Pearson system of continuous distributions is a system for which every probability density function (PDF) $f(x)$ satisfies the differential equation of the following form:

$$\frac{1}{f(x)} \cdot \frac{df(x)}{dx} = \frac{a + x}{b_0 + b_1x + b_2x^2}, \quad (1.1)$$

where a, b_0, b_1 , and b_2 are the shape parameters. Different types of distributions correspond to different forms of solution to Eqn. (1.1). The form of solution of Eqn. (1.1) depends on the root

2020 Mathematics Subject Classification: Primary 60E05.

Keywords: Burr distribution, exponentiated distributions, stochastic ordering, reliability properties, maximum likelihood.

^c Corresponding author.

of the equation

$$b_0 + b_1x + b_2x^2 = 0.$$

Pearson presented four types of distributions [14] and characterizations of Pearson type distributions are also available in the literature [4]. Irving W. Burr proposed another important development in this category. In Burr's method [9], system of distributions satisfy the following differential equation:

$$dF = F(1 - F)g(x)dx, \quad (1.2)$$

where $0 \leq F \leq 1$ and $g(x)$ is a non-negative function over x . Twelve different solutions to the Eqn. (1.2) in the form of cumulative distribution functions (CDF) were given and named as Burr Types I–XII distributions. The Burr type-XII distribution, being a member of the Burr system has gained more attention in the last decade due to its potential use in practical situations. The Burr Type XII, simply called Burr distribution, is highly useful for fitting heavy-tailed data sets from the field of reliability, economics, hydrology, actuarial science, and network science among many others [24]. Burr distribution also emerges as a suitable model to describe stationary states of complex and non-equilibrium systems [38, 39]. The main advantage of Burr distribution from the extreme value statistics' point of view is that it has algebraic tails which are effective for modeling failures that occur with lesser frequency than with corresponding models based on exponential tails [41]. The CDF and PDF of the Burr distribution are defined as follows.

DEFINITION 1. A random variable X follows Burr distribution with parameters c , λ , α , if the CDF is of the following form:

$$F(x; c, \lambda, \alpha) = 1 - \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-\alpha}, \quad x > 0, c, \alpha, \lambda > 0, \quad (1.3)$$

where α and c are shape parameters and λ is a scale parameter. The density function of the Burr distribution is given by

$$f(x; c, \lambda, \alpha) = c\alpha\lambda^{-c}x^{c-1} \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-\alpha-1}, \quad x > 0, c, \alpha, \lambda > 0. \quad (1.4)$$

The corresponding survival function is given by

$$S(x; c, \lambda, \alpha) = \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-\alpha}, \quad x > 0. \quad (1.5)$$

The hazard rate function is given by

$$h(x; c, \lambda, \alpha) = \frac{f(x; c, \lambda, \alpha)}{S(x; c, \lambda, \alpha)} = c\alpha\lambda^{-c}x^{c-1} \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-1}, \quad x > 0. \quad (1.6)$$

It is interesting to note that the CDF in (1.3) is regularly varying at infinity, viz. they satisfy for some $\gamma > 0$, called the tail-index, and for all $t > 0$,

$$\lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\gamma}.$$

This result suggests that Burr distribution well-suited for modeling extreme values as a heavy-tailed distribution. The bivariate and multivariate extensions of Burr distribution are available in the past literature [13, 42]. Consequently, numerous modifications or generalizations of the Burr distribution, using different compounding and weighting techniques, have been suggested. See, for instance, the beta Burr distribution introduced by [36], the McDonald Burr distribution introduced by [17], order-statistics based generalized Burr distribution [5], the Marshall-Olkin Burr (MO Burr) distribution developed by [22], modified Burr distribution [21] the exponentiated Burr distribution introduced by [26] and the log-Weibull Burr distribution developed by [1]. Though all these modifications and generalizations resulted well in specific data analysis but all of these

assumed constant exponents for the Burr distribution which causes potential failure when working with large-scale heavy-tailed data sets from various applied domains [24]. In this work, the primary hypothesis is that the exponent α of the Burr distribution is not a constant and varies according to a nonlinear function g which depends on the data.

This article aims to introduce an extra parameter to Burr, Pareto, and exponentiated distributions to bring more flexibility to the given families when dealing with real-world large-scale heavy-tailed data sets. We introduce a new method of generalization, namely the shape-parameter transformation (SPT) method, where the non-negative shape parameter is assumed to be expressible as a nonlinear function of the empirical data and adds an additional parameter to the base distribution. The proposed SPT method is straightforward to use; hence it can effectively be used for data analysis purposes. The proposed SPT method is first applied to a Burr distribution to yield generalized Burr (GBurr) family of distributions. The newly introduced GBurr family has some excellent statistical properties and belongs to the maximum domain of attractions of the Frechet distribution. The method presented in this paper adds an additional parameter in the distribution and can be a competitor to the popularly used Marshall-Olkin (MO) [30] and Lehmann alternatives-based [27] method of generalizing continuous distributions. The SPT method is further applied to the Pareto and the exponentiated family of distributions which shows the broader applicability of the proposed SPT method. Further, a specific nonlinear variant of the GBurr family, we call it NBurr distribution, is discussed in detail. Complementary theoretical aspects are studied, such as shapes, asymptotes, quantiles, stochastic ordering, reliability parameter, and inferential statistics. The application of the proposed NBurr distribution is shown using large-scale heavy-tailed network data sets from various disciplines.

The rest of the paper is organized as follows. We discuss the SPT method and its application to the Burr, Pareto, and exponentiated family of distributions in Section 2 along with some structural properties. In Section 3, we study a simple model from the GBurr family of distributions and discuss its stochastic and inferential characteristics. The empirical application of this new NBurr distribution to various large scale heavy-tailed network data sets is presented in Section 4. Finally, we conclude this paper in Section 5.

2. Shape parameter transformation (SPT) method

In this section, we introduce the SPT method to generalize Burr, Pareto, and exponentiated distributions by modeling the shape parameter of these families as a nonlinear function of the data. It also adds an additional parameter to the families as mentioned above of distributions for flexible modeling of real-life data sets. The SPT method is first applied to yield generalized Burr distributions which are regularly varying distributions at infinity and are heavy-tailed. Furthermore, this new SPT method is applied to Pareto distribution and exponentiated family of distributions to generalize these families.

2.1. The generalized Burr (GBurr) family

DEFINITION 2. A continuous random variable X follows generalized Burr (Burr) family of distributions if and only if it has the following CDF:

$$F_{\text{GBurr}}(x; c, \lambda, \alpha, \beta) = 1 - \left[1 + \left(\frac{x}{\lambda} \right)^c \right]^{-g\left(\frac{x}{\lambda}, \alpha, \beta\right)}, \quad x \geq 0, \alpha, \lambda, c > 0, \quad (2.1)$$

and $F(x) = 0$ if $x < 0$, where the real valued continuous, positive function $g: (0, \infty) \rightarrow \mathbb{R}^+$ is differentiable on $(0, \infty)$. The shape parameter α of the Burr distribution is replaced with

$g(z) = g(x/\lambda; \alpha, \beta) = g(x/\lambda)$, say, where β is an additional shape parameter and $g(z)$ satisfies the following conditions:

- (1) The function $g(z)$ is strictly positive and have finite limit at infinity, viz.

$$\lim_{z \rightarrow \infty} g(z) = \alpha (> 0). \quad (2.2)$$

- (2) $\lim_{z \rightarrow 0^+} (1 + z^c)^{g(z)} = 1$ and $\lim_{z \rightarrow \infty} (1 + z^c)^{g(z)} = \infty$.

- (3) $\frac{g'(z)}{g(z)} \geq -\frac{cz^{c-1}}{(1 + z^c) \log(1 + z)}$, $z > 0$, where $g'(z) = \frac{d}{dz} [g(z)]$.

It is very easy to verify that:

- (1) $F_{\text{GBurr}}(x)$ is non-decreasing since,

$$c \left(\frac{x}{\lambda}\right)^{c-1} g\left(\frac{x}{\lambda}\right) + g'\left(\frac{x}{\lambda}\right) \left(1 + \left(\frac{x}{\lambda}\right)^c\right) \log\left(1 + \frac{x}{\lambda}\right) > 0, \quad x > 0.$$

- (2) $\lim_{x \rightarrow -\infty} F_{\text{GBurr}}(x) = F_{\text{GBurr}}(-\infty) = 0$, $\lim_{x \rightarrow \infty} F_{\text{GBurr}}(x) = 1$.

- (3) $F_{\text{GBurr}}(x)$ is right continuous.

Thus, (2.1) is a standard CDF and it can also be expressed as follows:

$$F_{\text{GBurr}}(x) = 1 - \exp\left[-g\left(\frac{x}{\lambda}\right) \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right)\right], \quad x > 0. \quad (2.3)$$

The corresponding survival function is given by

$$S_{\text{GBurr}}(x) = \begin{cases} \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-g(x/\lambda)}, & x > 0 \\ 1, & x \leq 0. \end{cases} \quad (2.4)$$

The probability density function is

$$f_{\text{GBurr}}(x) = \begin{cases} \frac{1}{\lambda} \left[\frac{c \left(\frac{x}{\lambda}\right)^{c-1}}{1 + \left(\frac{x}{\lambda}\right)^c} g\left(\frac{x}{\lambda}\right) + g'\left(\frac{x}{\lambda}\right) \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right) \right] \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-g(x/\lambda)}, & x > 0 \\ 0, & x \leq 0. \end{cases} \quad (2.5)$$

The hazard rate function is given by

$$h_{\text{GBurr}}(x) = \begin{cases} \frac{c}{\lambda} \cdot \frac{\left(\frac{x}{\lambda}\right)^{c-1}}{1 + \left(\frac{x}{\lambda}\right)^c} g\left(\frac{x}{\lambda}\right) + \frac{1}{\lambda} g'\left(\frac{x}{\lambda}\right) \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right), & x > 0 \\ 0, & x \leq 0. \end{cases} \quad (2.6)$$

The CDF (2.1) of the GBurr family of distributions is a function with regularly varying tails and belongs to the maximum domain of attraction (MDA) of the Frechet distribution with index $\alpha > 0$, viz. $F \in MDA(\Phi_\alpha)$, with $\Phi_\alpha = \exp\{-x^{-\alpha}\}$, $x > 0$, $\alpha > 0$ [15].

THEOREM 2.1. *GBurr family of distributions are:*

- (a) *heavy-tailed,*
- (b) *right tail-equivalent to a Pareto distribution, a*
- (c) *belongs to the MDA of the Frechet distribution.*

Proof. (a) For the GBurr family of distributions, we note that

$$\lim_{x \rightarrow \infty} \exp\{kx\} (1 - F(x)) = \lim_{x \rightarrow \infty} \exp\left[kx - g\left(\frac{x}{\lambda}\right) \log\left(1 + \frac{x^c}{\lambda^c}\right)\right] = \infty,$$

where $k, \lambda, c > 0$.

(b) To show the tail-equivalent property, we show

$$\lim_{x \rightarrow \infty} \frac{1 - F(x)}{1 - G(x)} = \lim_{x \rightarrow \infty} \frac{\exp \left[-g\left(\frac{x}{\lambda}\right) \log \left(1 + \frac{x^c}{\lambda^c}\right) \right]}{\exp \left[-\alpha \log \left(1 + \frac{x}{\lambda}\right) \right]} = 1, \quad \text{for all } \lambda, c > 0,$$

where $G(x)$ is the CDF of the Pareto Type-II distribution.

(c) It should be noted that any function g as defined in Definition 2 satisfying $\lim_{z \rightarrow \infty} g(z) = \alpha > 0$, is slowly varying at infinity:

$$\lim_{z \rightarrow \infty} \frac{g(tz)}{g(z)} = 1, \quad \text{for all } t > 0.$$

Now,

$$\lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = \lim_{x \rightarrow \infty} \frac{\left(1 + \frac{(tx)^c}{\lambda^c}\right)^{-g\left(\frac{tx}{\lambda}\right)}}{\left(1 + \frac{x^c}{\lambda^c}\right)^{-g\left(\frac{x}{\lambda}\right)}} = t^{-\alpha}, \quad \text{for all } t > 0 \text{ and } \lambda, c > 0. \quad \square$$

It can be seen that Burr distribution belongs to this GBurr family and corresponds to the simplest choice $g(z) = \alpha$. We give some examples of nonlinear $g(z)$ satisfying the condition given in Definition 2 and present some distributions belonging to the GBurr family in Table 1. In order to select a model, one can choose the nonlinear exponent $g(z)$ that meets the empirical characteristics of the given data sets. Obviously, this example list of GBurr models can further be expanded to more models by introducing some other forms of $g(z)$ satisfying $\lim_{z \rightarrow \infty} g(z) = \alpha > 0$ or by increasing the number of parameters in the function g .

We can also obtain some popular size distributions as the particular cases of the GBurr family:

- By taking the constant g function, viz., when $g(z) = 1$, then GBurr distributions reduces to the Fisk distribution [16].
- For $\alpha \rightarrow \infty$ and $\beta = 0$, GBurr family of distributions converges to the Weibull distribution [25].
- By taking $c = 1$ and $\beta = 0$, GBurr distribution reduces to the Lomax distribution [29].

2.2. Generalized classical Pareto distribution

In this section, we apply the SPT method to the classical Pareto distribution to yield generalized classical Pareto (GCP) distributions as defined below.

DEFINITION 3. A continuous random variable X follows GCP family of distributions if and only if it has the following CDF:

$$F_{\text{GCP}}(x) = 1 - \left(\frac{x}{\gamma}\right)^{-m\left(\frac{x}{\gamma}\right)}, \quad x \geq \gamma \quad (2.7)$$

and $F(x) = 0$ if $x \leq \gamma$, $\gamma (> 0)$ is a scale parameter, $m: (1, \infty) \rightarrow \mathbb{R}^+$ is a real, continuous, positive function which is differentiable on $(1, \infty)$, and the function m satisfies the following conditions:

- (1) m is strictly positive and have finite limit at infinity, viz.

$$\lim_{z \rightarrow \infty} m(z) = \alpha (> 0). \quad (2.8)$$

- (2) $\lim_{z \rightarrow 1^+} z^{m(z)} = 1$ and $\lim_{z \rightarrow \infty} z^{m(z)} = \infty$.

$$(3) \quad \frac{m'(z)}{m(z)} > -\frac{1}{z \log(z)}, \quad z > 1.$$

It is noted that the condition (3) is equivalent to

$$\frac{d}{dz} \log(g(z)) \geq -\frac{d}{dz} \log[\log(z)], \quad z > 1.$$

It is very easy to verify that $F(x)$ in (2.7) is a standard distribution function and any continuous random variable X satisfying the above-mentioned conditions are called as GCP distributions. An alternative expression for $F(x)$ can also be given as follows:

$$F_{\text{GCP}}(x) = 1 - \exp \left[-m\left(\frac{x}{\gamma}\right) \log\left(\frac{x}{\gamma}\right) \right], \quad x \geq \gamma.$$

The survival function is given by

$$S_{\text{GCP}}(x) = \begin{cases} \left(\frac{x}{\gamma}\right)^{-m} & \text{if } x > \gamma \\ 1 & \text{if } x \leq \gamma. \end{cases} \quad (2.9)$$

The PDF for the family of GCP distributions is given by

$$f_{\text{GCP}}(x) = \begin{cases} \left[m\left(\frac{x}{\gamma}\right) + m'\left(\frac{x}{\gamma}\right) \left(\frac{x}{\gamma}\right) \log\left(\frac{x}{\gamma}\right) \right] \frac{S_{\text{GCP}}(x)}{x} & \text{if } x > \gamma \\ 0 & \text{if } x \leq \gamma. \end{cases} \quad (2.10)$$

The hazard function is given by

$$h_{\text{GCP}}(x) = \begin{cases} \frac{g\left(\frac{x}{\gamma}\right) + m'\left(\frac{x}{\gamma}\right) \left(\frac{x}{\gamma}\right) \log\left(\frac{x}{\gamma}\right)}{x} & \text{if } x > \gamma \\ 0 & \text{if } x \leq \gamma. \end{cases} \quad (2.11)$$

Table 2 shows some examples of GCP distributions satisfying condition (2.8) (i.e., $\lim_{z \rightarrow \infty} m(z) = \alpha > 0$). It is noted that the simplest choice $m(z) = \alpha$ leads the classical Pareto distribution. The rest of the models are completely new. In order to select a model, one can choose the nonlinear exponent $m(z)$ that meets the empirical characteristics of the given data sets. Obviously, this example list of GCP models can further be expanded to more models by introducing some other forms of $m(z)$ satisfying $\lim_{z \rightarrow \infty} m(z) = \alpha > 0$ or by increasing the number of parameters in the function g .

2.3. Generalized exponentiated distributions

The idea of exponentiated family of distributions is based on Lehmann alternatives [27] that can add a non-negative shape parameter to any continuous probability distributions:

$$F^*(x) = [F(x)]^\alpha, \quad \text{where } \alpha > 0 \text{ (real)}, \quad (2.12)$$

where $F(x)$ is a standard CDF. This new family F^* is called ‘exponentiated family’ where one raises the CDF of an existing distribution to a power of an additional parameter. Lehmann motivated it in the following way: $[F(x)]^\alpha$ is the distribution of the maximum of α independent and identically

TABLE 1. Some examples of GBurr family of distributions.

$g(z)$ for all $z > 0$	β	$f(x)$ for all $x > 0$
$\alpha - \frac{\beta}{z+1}$	$\beta \leq \alpha$	$\left[\frac{\beta\lambda}{(x+\lambda)^2} \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c} \right] \left[1 + \left(\frac{x}{\lambda}\right)^c \right]^{\alpha - \frac{\beta}{1+x/\lambda}}$
$\alpha \left[\frac{z}{1+z} \right]^\beta$	$\beta > -1$	$\frac{\alpha}{\lambda} \left[\frac{x}{x+\lambda} \right]^\beta \left[\frac{\beta\lambda^2}{x(x+\lambda)} + \frac{\lambda cx^{c-1}}{x^c + \lambda^c} \right] \exp \left[-\alpha \log\left(1 + \frac{x^c}{\lambda^c}\right) \left(\frac{x}{x+\lambda}\right)^\beta \right]$
$\alpha \left[\frac{\log(1+z)}{1+\log(1+z)} \right]^\beta$	$\beta > -1$	$\frac{\alpha}{\lambda} \left[\frac{\log(1+x/\lambda)}{1+\log(1+x/\lambda)} \right]^\beta \left[\frac{\beta \log\left(1 + \frac{x^c}{\lambda^c}\right)}{(1+x/\lambda)[1+\log(1+x/\lambda)] [\log(1+x/\lambda)]} + \frac{\lambda cx^{c-1}}{x^c + \lambda^c} \right] \times \exp \left[-\alpha \frac{\log^\beta(1+x/\lambda) \log\left(1 + \frac{x^c}{\lambda^c}\right)}{[1+\log(1+x/\lambda)]^\beta} \right]$

TABLE 2. Some examples of GCP family of distributions with $\lim_{z \rightarrow \infty} m(z) = \alpha > 0$.

$m(z)$ for all $z > 1$	β	$f(x)$ for all $x > \gamma$
$\alpha - \frac{\beta}{z}$	$\beta \leq \alpha$	$\left(\frac{x}{\gamma}\right)^{-\alpha+\beta\gamma/x} \left[\frac{\alpha}{x} + \frac{\gamma\beta}{x} \left(\frac{\log(x/\gamma) - 1}{x} \right) \right]$
$\alpha \left[\frac{z-1}{z} \right]^\beta$	$\beta > -1$	$\frac{\alpha}{x} \left(\frac{x}{\gamma}\right)^{-\alpha} \left[\frac{(x/\gamma) - 1}{(x/\gamma)} \right]^\beta \left[\frac{(x/\gamma) - 1 + \beta \log(x/\gamma)}{(x/\gamma) - 1} \right] \left[\frac{(x/\gamma) - 1}{(x/\gamma)} \right]^\beta$
$\alpha \left[\frac{\log(z)}{1+\log(z)} \right]^\beta$	$\beta > -1$	$\frac{\alpha}{x} \left(\frac{x}{\gamma}\right)^{-\alpha} \left[\frac{\log(x/\gamma)}{\log(x/\gamma) + 1} \right]^\beta \left[\frac{\log(x/\gamma) + 1 + \beta}{\log(x/\gamma) + 1} \right] \left[\frac{\log(x/\gamma)}{\log(x/\gamma) + 1} \right]^\beta$

distributed variables with distribution F when α is an integer and when α is a rational number, $[F(x)]^\alpha$ is one-parameter family of nonparametric class of alternatives. In the theory of generalized probability distributions, Lehmann alternatives has been widely used to generate so-called exponentiated Weibull family of distribution by [31]. A systematic treatment of exponentiated weibull, exponentiated gamma, and exponentiated pareto distribution are available in [18]. Gupta and Kundu studied exponentiated exponential distribution [20] and recently Gupta et al. [19] proposed Power Normal distributions using the same idea of Lehmann's alternatives for standard normal distribution. Nadarajah and Kotz [32] studied a list of exponentiated X family of distributions, including exponentiated Gumbel, Fréchet, gamma, etc. distributions. We use our SPT method to generalize any 'exponentiated family' of distributions.

DEFINITION 4. Any continuous random variable X with distribution function $F(x)$ follows generalized exponentiated distribution if and only if it has the following CDF

$$G(x) = [F(x)]^{m(x)}, \quad -\infty < x < \infty,$$

where $m: (-\infty, \infty) \rightarrow (0, \infty)$ is a real, continuous function which is differentiable on $(0, \infty)$. The function m satisfies the following conditions:

- (1) m is strictly positive and have finite limit at infinity, viz.

$$\lim_{x \rightarrow \infty} m(x) = \alpha (> 0).$$

- (2) $\lim_{x \rightarrow -\infty} m(x) = \beta (> \alpha).$

- (3) $m'(x) < 0.$

It is important to note that when $m(x) = \alpha$, then $G(x)$ corresponds to the exponentiated family of distributions [18, 32]. We give some examples of $m(x)$ to be used for the distributions with $x \in \mathbb{R}$.

Example 1. Let $m(x)$ be a function of the following form satisfying $\lim_{x \rightarrow -\infty} \xi(x) = 2$ and $\lim_{x \rightarrow \infty} \xi(x) = 0$:

$$m(x) = \frac{(\beta - \alpha)}{2} \xi(x) + \alpha, \quad \beta > \alpha > 0.$$

Some choices of $\xi(x)$ are as follows:

$$\begin{aligned} (a) \quad \xi(x) &= \begin{cases} e^{-x} & \text{if } x \geq 0, \\ 2 - e^x & \text{if } x < 0, \end{cases} \\ (b) \quad \xi(x) &= \begin{cases} \frac{1}{x^{2k+1} + 1} & \text{if } x \geq 0, \\ 2 + \frac{1}{x^{2k+1} - 1} & \text{if } x < 0, \end{cases} \\ (c) \quad \xi(x) &= \begin{cases} \frac{1}{\log(1+x) + 1} & \text{if } x \geq 0, \\ 2 - \frac{1}{\log(1-x) + 1} & \text{if } x < 0, \end{cases} \end{aligned}$$

where k is any natural numbers.

Example 2. Let $m(x)$ be a function of the following form satisfying $\lim_{x \rightarrow -\infty} \sigma(x) = 0$ and $\lim_{x \rightarrow \infty} \sigma(x) = 1$:

$$m(x) = \alpha \sigma(x) + \beta (1 - \sigma(x)), \quad \beta > \alpha > 0.$$

Some choices of $\sigma(x)$ are as follows:

- $\sigma(x) = \frac{1}{1 + e^{-x}} \quad x \in \mathbb{R}.$
- $\sigma(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x), \quad x \in \mathbb{R}.$
- $\sigma(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-y^2/2) dy, \quad x \in \mathbb{R}.$

It is interesting to see that

$$G(x) = F(x)^{\alpha\Phi(x) + \beta\Phi(-x)}, \quad x \in \mathbb{R}$$

is a standard CDF, where $\Phi(x)$ is a standard normal CDF. For all the above examples when $\alpha = \beta$, we get $G(x) = [F(x)]^\alpha$.

3. NBurr distribution: Definition and properties

In this section, we study a relevant model of the GBurr family of distributions with the choice of $g(z) = \alpha - \frac{\beta}{1+z}$ in Table 1. The CDF is given by

$$F_{\text{NBurr}}(x) = 1 - \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-\left(\alpha - \frac{\beta}{1+x/\lambda}\right)}, \quad x > 0, \quad (3.1)$$

where $\alpha, \lambda, c > 0$, $\beta > -1$ and $\alpha > \beta$. We call this simple generalized form of Burr distribution as NBurr distribution (Burr distribution with nonlinear exponent). The NBurr distribution includes the Burr distribution when $\beta = 0$. The CDF in (3.1) can alternatively be written in the following form:

$$F_{\text{NBurr}}(x) = 1 - \exp\left[-\left(\alpha - \frac{\beta}{1 + \frac{x}{\lambda}}\right) \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right)\right], \quad x > 0. \quad (3.2)$$

The corresponding survival function is given by,

$$S_{\text{NBurr}}(x) = \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-\left(\alpha - \frac{\beta}{1+x/\lambda}\right)}, \quad x > 0. \quad (3.3)$$

The probability density function for $x > 0$ is given by

$$f_{\text{NBurr}}(x) = \left[\frac{\beta\lambda}{(x+\lambda)^2} \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c}\right] S_{\text{NBurr}}(x), \quad x > 0, \quad (3.4)$$

and the hazard rate function is given by

$$h_{\text{NBurr}}(x) = \frac{\beta\lambda}{(x+\lambda)^2} \log\left(1 + \left(\frac{x}{\lambda}\right)^c\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c}, \quad x > 0. \quad (3.5)$$

The proposed NBurr distribution satisfies the extreme value properties as given in Theorem 2.1. Some graphics of the NBurr model derived from this newly introduced GBurr family of distributions are illustrated in Figure 1. Remark that NBurr distribution with parameters $\alpha, \beta, \lambda, c$, as a distribution of the MDA of the Frechet distribution, satisfies the von Mises condition:

$$\lim_{x \rightarrow \infty} \frac{xf_{\text{NBurr}}(x)}{S_{\text{NBurr}}(x)} = \alpha > 0.$$

Next, we study the reliability properties and inferential properties of the new NBurr distribution in the next subsections.

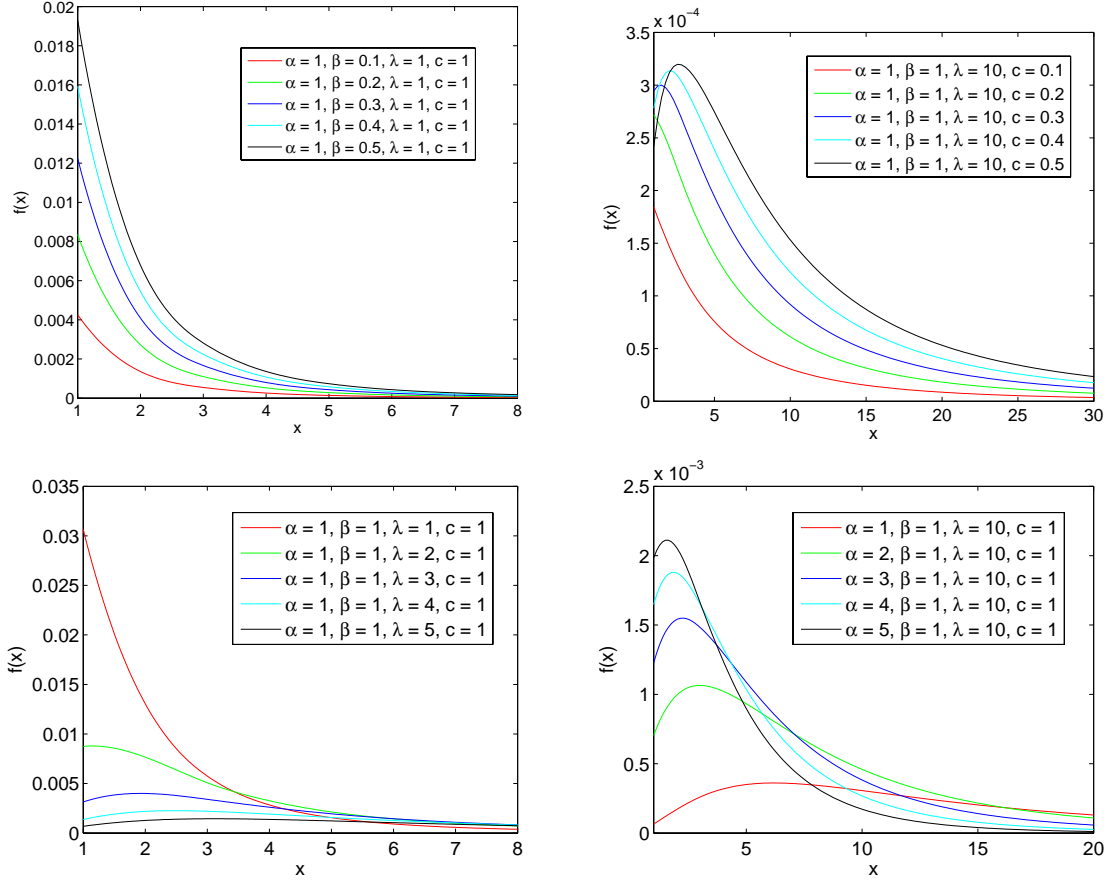


FIGURE 1. Plots of PDFs of the NBurr distribution.

3.1. Reliability properties of NBurr distribution

In this section, we study some reliability properties of the newly introduced NBurr distribution including monotonicity of hazard rates, stochastic orderings, entropy, etc. The following theorem shows that the hazard rate function of the NBurr distribution is increasing and decreasing under certain conditions on c and β .

THEOREM 3.1. *The hazard rate function of the NBurr distribution satisfies the following properties:*

- (1) *If $0 \leq c \leq 1$ and $\beta > 0$ then for all $x > 0$, $h_{\text{NBurr}}(x)$ is decreasing in x ;*
- (2) *If $c > 1$ and $\beta > 0$, then for all $x > 0$,*
 - (a) *$h_{\text{NBurr}}(x)$ is increasing in x , when $c - 1 > \left(\frac{x}{\lambda}\right)^c$;*
 - (b) *$h_{\text{NBurr}}(x)$ is decreasing in x , when $c - 1 < \left(\frac{x}{\lambda}\right)^c$;*
 - (c) *$h_{\text{NBurr}}(x)$ is maximum at $x = \lambda[c - 1]^{1/c}$.*

Proof. Differentiating (3.5) with respect to x , we have

$$\begin{aligned} h'_{\text{NBurr}}(x) &= -\frac{2\beta\lambda}{(x+\lambda)^3} \log\left(1 + \frac{x^c}{\lambda^c}\right) + \frac{2c\beta\lambda x^{c-1}}{(x+\lambda)^2(x^c + \lambda^c)} \\ &\quad + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-2}}{\lambda^c(x^c + \lambda^c)^2} \left[c - 1 - \frac{x^c}{\lambda^c}\right] \\ &= \frac{2\beta\lambda}{(x+\lambda)^3} \left[cx^{c-1} \frac{x+\lambda}{x^c + \lambda^c} - \log\left(1 + \frac{x^c}{\lambda^c}\right) \right] \\ &\quad + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-2}}{(x^c + \lambda^c)^2} \left[c - 1 - \frac{x^c}{\lambda^c}\right]. \end{aligned}$$

Now, $\alpha \geq \frac{\beta\lambda}{x+\lambda}$, $\frac{c\lambda^c x^{c-2}}{(x^c + \lambda^c)^2} \geq 0$ and if $\beta > 0$, $\frac{2\beta\lambda}{(x+\lambda)^3(x^c + \lambda^c)} > 0$.

Thus, $h'_{\text{NBurr}}(x) \geq (\leq) 0$ if

- (i) $A(x) = cx^{c-1} \frac{x+\lambda}{x^c + \lambda^c} - \log\left(1 + \frac{x^c}{\lambda^c}\right) \geq (\leq) 0$;
- (ii) $c - 1 - \frac{x^c}{\lambda^c} \geq (\leq) 0$.

Now,

$$A'(x) = \frac{c\lambda^c x^{c-2}(x+\lambda)}{(x^c + \lambda^c)^2} \left[c - 1 - \frac{x^c}{\lambda^c}\right].$$

We can see that $A'(x) \leq 0$ if $0 \leq c \leq 1$. Therefore, $A(x)$ is decreasing in x and again $A(0) = 0$. Thus, we have $A(x) \leq 0$.

Similarly, we can prove that if $c > 1$, then $A(x) \geq 0$ when $c - 1 > \left(\frac{x}{\lambda}\right)^c$ and $A(x) \leq 0$ when $c - 1 < \left(\frac{x}{\lambda}\right)^c$. This implies that for $0 \leq c \leq 1$, $h_{\text{NBurr}}(x)$ is decreasing in x ; for $c > 1$, $h_{\text{NBurr}}(x)$ is increasing in x , when $c - 1 > \left(\frac{x}{\lambda}\right)^c$ and $h_{\text{NBurr}}(x)$ is decreasing in x , when $c - 1 < \left(\frac{x}{\lambda}\right)^c$. Thus, the hazard function $h_{\text{NBurr}}(x)$ attain its maximum at $x = \lambda[c - 1]^{1/c}$. \square

The following example shows that the NBurr distribution does not preserves the likelihood ratio ordering. It is useful to recall that a random variable X is said to be larger than another random variable Y in likelihood ratio ordering (written as $X \geq_{LR} Y$) if $f_X(x)/f_Y(x)$ is an increasing function for $x > 0$.

Example 3. Let X and Y be two random variables following NBurr distributions with parameters $\alpha_1, \beta_1, \lambda_1, c_1$ and $\alpha_2, \beta_2, \lambda_2, c_2$ respectively. Then, for all $x > 0$, the ratio of the corresponding density functions of X and Y is given by

$$\frac{f_{X_{\text{NBurr}}}(x)}{f_{Y_{\text{NBurr}}}(x)} = \frac{\left[\frac{\beta_1 \lambda_1}{(x+\lambda_1)^2} \log\left(1 + \left(\frac{x}{\lambda_1}\right)^{c_1}\right) + \left(\alpha_1 - \frac{\beta_1 \lambda_1}{x+\lambda_1}\right) \frac{c_1 x^{c_1-1}}{x^{c_1} + \lambda_1^{c_1}} \right] \left[1 + \left(\frac{x}{\lambda_1}\right)^{c_1}\right]^{-\left(\alpha_1 - \frac{\beta_1}{1+x/\lambda_1}\right)}}{\left[\frac{\beta_2 \lambda_2}{(x+\lambda_2)^2} \log\left(1 + \left(\frac{x}{\lambda_2}\right)^{c_2}\right) + \left(\alpha_2 - \frac{\beta_2 \lambda_2}{x+\lambda_2}\right) \frac{c_2 x^{c_2-1}}{x^{c_2} + \lambda_2^{c_2}} \right] \left[1 + \left(\frac{x}{\lambda_2}\right)^{c_2}\right]^{-\left(\alpha_2 - \frac{\beta_2}{1+x/\lambda_2}\right)}}. \quad (3.6)$$

Now, for $\lambda_1 = \lambda_2 = 1$,

Case I, when $c_1 = c_2 = 1$, (3.6) reduces to

$$\frac{f_{X_{\text{NBurr}}}(x)}{f_{Y_{\text{NBurr}}}(x)} = \frac{\left[\frac{\beta_1}{(1+x)^2} \log(1+x) + \left(\alpha_1 - \frac{\beta_1}{1+x}\right) \frac{1}{1+x} \right] (1+x)^{-\left(\alpha_1 - \frac{\beta_1}{1+x}\right)}}{\left[\frac{\beta_2}{(1+x)^2} \log(1+x) + \left(\alpha_2 - \frac{\beta_2}{1+x}\right) \frac{1}{1+x} \right] (1+x)^{-\left(\alpha_2 - \frac{\beta_2}{1+x}\right)}} = P_1(x), \text{ say.}$$

We can see that for $\alpha_1 = 1.5$, $\alpha_2 = 2$, $\beta_1 = 0.5$ and $\beta_2 = 1.5$, $P_1(0.1) = 1.366$, $P_1(2) = 0.889$ and $P_1(6) = 1.426$, which implies that $\frac{f_{X_{NBurr}}(x)}{f_{Y_{NBurr}}(x)}$ is not monotone. Thus $X \not\geq_{LR} Y$.

Case II, when $c_1 \neq c_2$, (3.6) reduces to

$$\frac{f_{X_{NBurr}}(x)}{f_{Y_{NBurr}}(x)} = \frac{\left[\frac{\beta_1}{(1+x)^2} \log(1+x^{c_1}) + \left(\alpha_1 - \frac{\beta_1}{1+x} \right) \frac{c_1 x^{c_1-1}}{1+x^{c_1}} \right] [1+x^{c_1}]^{-(\alpha_1 - \frac{\beta_1}{1+x})}}{\left[\frac{\beta_2}{(1+x)^2} \log(1+x^{c_2}) + \left(\alpha_2 - \frac{\beta_2}{1+x} \right) \frac{c_2 x^{c_2-1}}{1+x^{c_2}} \right] [1+x^{c_2}]^{-(\alpha_2 - \frac{\beta_2}{1+x})}} = P_2(x), \text{ say.}$$

Again, we can see that for $c_1 = 2$, $c_2 = 4$, $\alpha_1 = 1.5$, $\alpha_2 = 2$, $\beta_1 = 0.5$ and $\beta_2 = 1.5$, $P_2(0.5) = 1.5753$, $P_2(1) = 0.4843$ and $P_2(2) = 2.8757$, which implies that $\frac{f_{X_{NBurr}}(x)}{f_{Y_{NBurr}}(x)}$ is not monotone. Thus $X \not\geq_{LR} Y$. Hence, it can be concluded that the proposed NBurr distribution does not preserve the likelihood ratio ordering. \square

In the next theorem, we show that the NBurr distribution preserves the usual stochastic ordering. It is useful to remind that a random variable X is said to be larger than another random variable Y in usual stochastic ordering (written as $X \geq_{ST} Y$) if $S_X(x) \geq S_Y(x)$, for all $x > 0$.

THEOREM 3.2. *Let X and Y be two random variables following NBurr distribution with parameters $\alpha_1, \beta_1, \lambda_1, c_1$ and $\alpha_2, \beta_2, \lambda_2, c_2$, respectively. Then, $X \geq_{ST} (\leq_{ST}) Y$ provided*

- (i) $\lambda_1 \geq (\leq) \lambda_2$,
- (ii) $c_1 \leq (\geq) c_2$,
- (iii) $\alpha_1 \leq (\geq) \alpha_2$, and
- (iv) $\beta_1 \leq (\geq) \beta_2$ for $\beta > 0$.

Proof. $X \geq_{ST} (\leq_{ST}) Y$ if and only if for all $x > 0$, $S_{X_{NBurr}}(x) \geq (\leq) S_{Y_{NBurr}}(x)$, which is equivalent to

$$\frac{\left[1 + \left(\frac{x}{\lambda_2} \right)^{c_2} \right]^{\alpha_2 - \frac{\beta_2 \lambda_2}{x + \lambda_2}}}{\left[1 + \left(\frac{x}{\lambda_1} \right)^{c_1} \right]^{\alpha_1 - \frac{\beta_1 \lambda_1}{x + \lambda_1}}} \geq (\leq) 1.$$

This holds if (i) $\lambda_1 \geq (\leq) \lambda_2$, (ii) $c_1 \leq (\geq) c_2$, (iii) $\alpha_1 \leq (\geq) \alpha_2$ and (iv) $\beta_1 \leq (\geq) \beta_2$, for $\beta > 0$. \square

The following theorem gives the condition under which the NBurr distribution preserves the hazard rate ordering. It is well known that a random variable X is said to be larger than another random variable Y in hazard rate ordering (written as $X \geq_{HR} Y$) if, for all $x > 0$, $h_X(x) \leq h_Y(x)$.

THEOREM 3.3. *Let X and Y be two random variables following NBurr distribution with parameters $\alpha_1, \beta_1, \lambda_1, c_1$ and $\alpha_2, \beta_2, \lambda_2, c_2$, respectively. Then, $X \geq_{HR} (\leq_{HR}) Y$ provided*

- (i) $\lambda_1 \geq (\leq) \lambda_2$,
- (ii) $c_1 \leq (\geq) c_2$,
- (iii) $\alpha_1 \leq (\geq) \alpha_2$ and
- (iv) $\beta_1 \leq (\geq) \beta_2$, for $\beta > 0$.

Proof. $X \geq_{HR} (\leq_{HR}) Y$ if and only if for all $x > 0$, $h_{X_{NBurr}}(x) \leq (\geq) h_{Y_{NBurr}}(x)$, which is equivalent to

$$\frac{\frac{\beta_1}{\lambda_1 \left(1 + \frac{x}{\lambda_1} \right)^2} \log \left(1 + \left(\frac{x}{\lambda_1} \right)^{c_1} \right) + \left(\alpha_1 - \frac{\beta_1}{1 + \frac{x}{\lambda_1}} \right) \frac{c_1 \left(\frac{x}{\lambda_1} \right)^{c_1-1}}{\lambda_1 \left(1 + \left(\frac{x}{\lambda_1} \right)^{c_1} \right)}}{\frac{\beta_2}{\lambda_2 \left(1 + \frac{x}{\lambda_2} \right)^2} \log \left(1 + \left(\frac{x}{\lambda_2} \right)^{c_2} \right) + \left(\alpha_2 - \frac{\beta_2}{1 + \frac{x}{\lambda_2}} \right) \frac{c_2 \left(\frac{x}{\lambda_2} \right)^{c_2-1}}{\lambda_2 \left(1 + \left(\frac{x}{\lambda_2} \right)^{c_2} \right)}} \leq (\geq) 1.$$

$$\leq (\geq) \frac{\beta_2}{\lambda_2 \left(1 + \frac{x}{\lambda_2}\right)^2} \log \left(1 + \left(\frac{x}{\lambda_2}\right)^{c_2}\right) + \left(\alpha_2 - \frac{\beta_2}{1 + \frac{x}{\lambda_2}}\right) \frac{c_2 \left(\frac{x}{\lambda_2}\right)^{c_2-1}}{\lambda_2 \left(1 + \left(\frac{x}{\lambda_2}\right)^{c_2}\right)}.$$

This holds if (i) $\lambda_1 \geq (\leq) \lambda_2$, (ii) $c_1 \leq (\geq) c_2$, (iii) $\alpha_1 \leq (\geq) \alpha_2$ and (iv) $\beta_1 \leq (\geq) \beta_2$, for $\beta > 0$. \square

Remark 1. It is interesting to note that the hazard rate function of the NBurr distribution can be both increasing and decreasing under certain conditions as given in Theorem 3.1. Also, NBurr distribution preserves stochastic and hazard rate orderings as shown in Theorems 3.2 and 3.3, respectively.

The Shannon entropy is an important and well-known concept in information theory as well as engineering sciences. Let X be a random variable that follows NBurr distribution with parameters $\alpha, \beta, \lambda, c$. Then Shannon's entropy for the NBurr distribution is defined as

$$\begin{aligned} H(X) &= -E[\log f_{\text{NBurr}}(X)] = -\int_0^\infty f_{\text{NBurr}}(x) \log f_{\text{NBurr}}(x) dx \\ &= -\int_0^\infty \left(\frac{\beta\lambda}{(x+\lambda)^2} \log \left(1 + \frac{x^c}{\lambda^c}\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c} \right) \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-(\alpha - \frac{\beta}{1+x/\lambda})} \\ &\quad \times \left[\log \left(\frac{\beta\lambda}{(x+\lambda)^2} \log \left(1 + \frac{x^c}{\lambda^c}\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c} \right) \right. \\ &\quad \left. - \left(\alpha - \frac{\beta}{1+x/\lambda}\right) \log \left(1 + \left(\frac{x}{\lambda}\right)^c\right) \right] dx. \end{aligned}$$

When we want to study the system that survived up to an age t , then Shannon's entropy function is not useful in measuring the uncertainty about the residual lifetime of the system. Ebrahimi [3] has introduced residual entropy and defined as

$$H(X; t) = 1 - \int_t^\infty \frac{f_X(x)}{S_X(t)} \log \left(\frac{f_X(x)}{S_X(x)} \right) dx,$$

Then the residual entropy for NBurr distribution with parameters $\alpha, \beta, \lambda, c$ is given by

$$\begin{aligned} H(X; t) &= 1 - \frac{1}{S_{\text{NBurr}}(t)} \int_t^\infty f_{\text{NBurr}}(x) \log (h_{\text{NBurr}}(x)) dx \\ &= 1 - \left[1 + \left(\frac{t}{\lambda}\right)^c\right]^{(\alpha - \frac{\beta}{1+t/\lambda})} \\ &\quad \times \int_t^\infty \left[\frac{\beta\lambda}{(x+\lambda)^2} \log \left(1 + \left(\frac{x}{\lambda}\right)^c\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c} \right] \left[1 + \left(\frac{x}{\lambda}\right)^c\right]^{-(\alpha - \frac{\beta}{1+x/\lambda})} \\ &\quad \times \log \left[\frac{\beta\lambda}{(x+\lambda)^2} \log \left(1 + \left(\frac{x}{\lambda}\right)^c\right) + \left(\alpha - \frac{\beta\lambda}{x+\lambda}\right) \frac{cx^{c-1}}{x^c + \lambda^c} \right] dx. \end{aligned}$$

3.2. Parameter estimation

Let x_1, x_2, \dots, x_n be a sample of size n from $\text{NBurr}(\alpha, \beta, \lambda, c)$ distribution. We give procedure for parameter estimation the including the log-likelihood functions and corresponding normal equations. The log-likelihood function for the vector of parameters $\Theta = (\alpha, \beta, \lambda, c)^T$ corresponding to

NBurr distribution is given by

$$\begin{aligned} \ell \equiv \ell(x; \alpha, \beta, \lambda, c) = & \sum_{i=1}^n \log \left[\frac{\beta \lambda}{(\lambda + x_i)^2} \log \left(1 + \frac{x_i^c}{\lambda^c} \right) + \left(\alpha - \frac{\beta \lambda}{\lambda + x_i} \right) \frac{c x_i^{c-1}}{\lambda^c + x_i^c} \right] \\ & - \alpha \sum_{i=1}^n \log \left(1 + \frac{x_i^c}{\lambda^c} \right) + \beta \lambda \sum_{i=1}^n \frac{1}{\lambda + x_i} \log \left(1 + \frac{x_i^c}{\lambda^c} \right), \end{aligned} \quad (3.7)$$

where n is the sample size, and the maximum likelihood estimates of the unknown parameter vector $(\alpha, \beta, \lambda, c)$ are those that maximize the log-likelihood function ℓ in (3.7). The normal equations can be obtained by taking the partial derivatives of (3.7) w.r.t. $\alpha, \beta, \lambda, c$ and equating them to zero:

$$\begin{aligned} \frac{\partial \ell}{\partial \alpha} = & \sum_{i=1}^n \frac{c \omega_i^2 (\omega_i - 1)^{c-1}}{\beta (1 + (\omega_i - 1)^c) \log (1 + (\omega_i - 1)^c) + c \omega_i (\alpha \omega_i - \beta) (\omega_i - 1)^c} \\ & - \sum_{i=1}^n \log (1 + (\omega_i - 1)^c) \end{aligned} \quad (3.8)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \beta} = & \sum_{i=1}^n \frac{(1 + (\omega_i - 1)^c) \log (1 + (\omega_i - 1)^c) - c \omega_i (\omega_i - 1)^c}{\beta (1 + (\omega_i - 1)^c) \log (1 + (\omega_i - 1)^c) + c \omega_i (\alpha \omega_i - \beta) (\omega_i - 1)^c} \\ & + \sum_{i=1}^n \frac{1}{\omega_i} \log (1 + (\omega_i - 1)^c) \end{aligned} \quad (3.9)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \lambda} = & \sum_{i=1}^n \frac{(\beta x_i - \beta \lambda) (\lambda^c + x_i^c)^2 \log \left(1 + \frac{x_i^c}{\lambda^c} \right) - 2c \beta x_i^c (\lambda + x_i) (\lambda^c + x_i^c) - c^2 \lambda^{c-1} x_i^{c-1} (\lambda + x_i)^2 (\alpha (\lambda + x_i) - \beta \lambda)}{\beta \lambda (\lambda + x_i) (\lambda^c + x_i^c)^2 \log \left(1 + \frac{x_i^c}{\lambda^c} \right) + c x_i^{c-1} (\alpha (\lambda + x_i) - \beta \lambda) (\lambda + x_i)^2 (\lambda^c + x_i^c)} \\ & + \frac{c \alpha}{\lambda} \sum_{i=1}^n \frac{x_i^c}{\lambda^c + x_i^c} + \beta \sum_{i=1}^n \frac{x_i \log \left(1 + \frac{x_i^c}{\lambda^c} \right)}{(\lambda + x_i)^2} - c \beta \sum_{i=1}^n \frac{1}{(\lambda + x_i) (\lambda^c + x_i^c)} \end{aligned} \quad (3.10)$$

$$\begin{aligned} \frac{\partial \ell}{\partial c} = & \sum_{i=1}^n \frac{\beta (\omega_i - 1)^c (1 + (\omega_i - 1)^c) \log (\omega_i - 1) + \omega_i (\omega_i - 1)^{c-1} (\alpha \omega_i - \beta) (1 + c + (\omega_i - 1)^c)}{\beta (1 + (\omega_i - 1)^c)^2 \log (1 + (\omega_i - 1)^c) + c \omega_i (\omega_i - 1)^{c-1} (\alpha \omega_i - \beta) (1 + (\omega_i - 1)^c)} \\ & - \sum_{i=1}^n \frac{(\alpha \omega_i - \beta) (\omega_i - 1)^c \log (\omega_i - 1)}{\omega_i (1 + (\omega_i - 1)^c)}, \end{aligned} \quad (3.11)$$

where $\omega_i = 1 + \frac{x_i}{\lambda}$.

The MLEs of the four parameters for the NBurr distribution with α, β, λ and c are obtained by setting the above partial derivatives to zero and solving them simultaneously. The closed-form solutions are not available for the equations (3.8), (3.9), (3.10) and (3.11). So, an iterative algorithm should be applied to solve these equations numerically. For practical implementation of the model, we fit the NBurr models in the whole range of the data sets with quasi-Newton BFGS numerical algorithm with initial values to be chosen as $(\hat{\alpha}_0, \hat{\beta}_0, \hat{\lambda}_0, \hat{c}_0) = (1, 1, 1, 1)$ to find the MLE estimates of the parameters.

We also present the asymptotic distributions for the NBurr distribution. The Fisher information matrix (I) can be obtained by taking the expected values of the second-order and mixed partial derivatives of $\ell(x; \alpha, \beta, \lambda, c)$ w.r.t. α, β, λ and c . Since the analytical expression is hard to compute, it can be approximated by numerically investigating the $I = (I_{ij})$ matrix. The asymptotic I matrix

can be given as follows:

$$I = \begin{bmatrix} -\frac{\partial^2 \ell}{\partial \alpha^2} & -\frac{\partial^2 \ell}{\partial \alpha \partial \beta} & -\frac{\partial^2 \ell}{\partial \alpha \partial \lambda} & -\frac{\partial^2 \ell}{\partial \alpha \partial c} \\ -\frac{\partial^2 \ell}{\partial \alpha \partial \beta} & -\frac{\partial^2 \ell}{\partial \beta^2} & -\frac{\partial^2 \ell}{\partial \beta \partial \lambda} & -\frac{\partial^2 \ell}{\partial \beta \partial c} \\ -\frac{\partial^2 \ell}{\partial \alpha \partial \lambda} & -\frac{\partial^2 \ell}{\partial \beta \partial \lambda} & -\frac{\partial^2 \ell}{\partial \lambda^2} & -\frac{\partial^2 \ell}{\partial \lambda \partial c} \\ -\frac{\partial^2 \ell}{\partial \alpha \partial c} & -\frac{\partial^2 \ell}{\partial \beta \partial c} & -\frac{\partial^2 \ell}{\partial \lambda \partial c} & -\frac{\partial^2 \ell}{\partial c^2} \end{bmatrix}$$

The second order partial derivatives of $\ell(x; \alpha, \beta, \lambda, c)$ w.r.t. α, β, λ and c can be calculated but the calculations are very tedious. Hence, we omit the calculation part. The variance-covariance matrix is approximated by $M = (M_{ij})$ where $M_{ij} = I_{ij}^{-1}$. The asymptotic distribution of MLEs for α, β, λ , and c can be written as

$$\left[(\hat{\alpha} - \alpha), (\hat{\beta} - \beta), (\hat{\lambda} - \lambda), (\hat{c} - c) \right] \sim N_4 \left(0, I^{-1}(\hat{\theta}) \right).$$

Then the approximate $100(1 - k)\%$ confidence intervals for α, β, λ , and c are given by $\hat{\alpha} \pm \mathcal{Z}_{\frac{k}{2}} \sqrt{\text{Var}(\hat{\alpha})}$, $\hat{\beta} \pm \mathcal{Z}_{\frac{k}{2}} \sqrt{\text{Var}(\hat{\beta})}$, $\hat{\lambda} \pm \mathcal{Z}_{\frac{k}{2}} \sqrt{\text{Var}(\hat{\lambda})}$, and $\hat{c} \pm \mathcal{Z}_{\frac{k}{2}} \sqrt{\text{Var}(\hat{c})}$; where $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda}, \hat{c})$ and $\mathcal{Z}_{\frac{k}{2}}$ is the upper $100 - k$ -th percentile of the standard normal distribution.

3.3. Goodness of fit

The measure of closeness between the hypothesized NBurr distribution and the observed real-world network can be well determined by goodness-of-fit test. We have used the Chi-square statistic test and its corresponding p value to determine the goodness of fit for the NBurr distribution. We calculate the respective p values using the bootstrap resampling computational technique as given below:

- Initially the best fit NBurr distribution can be determined by estimating parameters through the MLE method given network data. Then we calculate the Chi-square statistic value as a measure of goodness-of-fit corresponding to the best-fitted NBurr model.
- Next we generate 50000 synthetic data sets from the NBurr distribution and calculate the Chi-square statistic for each of the synthetic data sets.
- Finally, we obtain the p value for the synthetic data sets as the fraction of NBurr synthetic data sets with a Chi-square value greater than the empirical one. Higher p values signify that the proposed model is ‘most’ suitable for the data set.

In addition, the effectiveness of the proposed NBurr distribution compared to other heavy-tailed distributions, is also verified by computing other well known statistical measures such as Kullback-Leibler divergence (KLDiv), root mean squared error (RMSE), and mean absolute error (MAE).

4. Real-world applications

We show the application of the NBurr distribution in the analysis of large-scale complex network data sets from various disciplines. The examples of such large-scale real-world complex networks include Twitter, Facebook, Orkut, Youtube, Amazon, LinkedIn, Wiki networks, etc. where the number of nodes is of the order of thousands or millions. There has been significant interest and attention devoted toward modeling aspects of such large-scale complex networks. Recent research [23, 28, 44] involved in the analysis of various important structural characteristics of network such as degree distribution, average nearest neighbor, clustering coefficient, community discovery, motif distribution, etc. Most of the interest has been focused on the analysis of the node degree distribution corresponding to these real-world networks [2, 6, 7]. Empirical observations suggest that the

node degree distributions of such real-world networks, for example, collaboration networks, communication networks, social networks, biological networks, etc., follow a heavy-tailed power-law distribution [6, 33]. Previous researchers reported that a baseline power-law, exponential, Pareto, log-normal, and Burr models are insufficient to fit the empirical data properly in its whole range unless some of the lower degree nodes are left out while fitting the model [10–12, 40, 43]. In recent work, Broido et al. [8] pointed out that the recent data concentration on all these networks data shows that they no longer follow the power-law distribution.

4.1. Data

We consider large-scale real-world network data sets from different disciplines, namely social networks, collaboration networks, citation networks, web graphs, product co-purchasing networks, temporal networks, communication networks, and ground-truth networks. We study several individual data sets from each discipline to showcase the general applicability of the proposed NBurr distribution. These data sets are publicly available at <http://snap.stanford.edu/data/index.html>. These are standard network data sets with heavy-tail behaviors and used for modeling in the statistical analysis of networks [33–35]. An overview of these publicly available network data sets along with statistical measures (mean (μ), standard deviation (s), etc.) are presented in Table 3. Another interesting property of these network data sets is their coefficient of variation (s/μ) exceeding unity.

4.2. Experimental results

In this section, we compare the NBurr distribution with the other seven established models, namely Power-law, Pareto, Log-normal, Power-law with cutoff, Burr, exponentiated Burr [26], MO Burr [22] distributions. To estimate the parameters ($\alpha, \beta, \lambda, c$) of the NBurr distribution numerically, we have used ‘optim’ function along with the quasi-Newton L-BFGS-B algorithm in R statistical software by taking the initial parameters value ($\alpha, \beta, \lambda, c$) = (1, 1, 1, 1). The estimated values of the parameters for all the network data sets satisfied the following conditions: $\alpha > 0$, $\beta > -1$, $\lambda > 0$ and $c > 0$ as depicted in Table 3, which characterize the proposed NBurr distribution. Empirically it is observed that in the case of social networks, the estimated value of the parameter λ attains the higher values as compared to the estimated value of α . From Table 3 it is also clear that the proposed NBurr distribution produces higher p values through bootstrapping chi-square test which suggests that the null hypothesis i.e., “the data are drawn from NBurr distribution”, cannot be ruled out at the 0.05 level of significance. This recommends in favor of the use NBurr distribution for fitting the node degree distribution of a network. Experimental results (given in Table 3) suggests that the proposed NBurr distribution is effective in modeling the entire degree distribution of real-world complex networks. Also, we used some other statistical measures, viz. KLDiv, RMSE, and MAE to compare the performance of the proposed NBurr distribution with the competitive heavy-tailed distributions as shown in Tables 4 and 5.

Given a network, the information about the differences between actual and predicted degree frequencies can be determined by calculating the root mean squared error (RMSE) and mean absolute error (MAE). The higher similarity between actual and predicted distributions is achieved by generating smaller values of RMSE and MAE. From Tables 4 and 5, it is clear that the proposed NBurr distribution produces smaller RMSE and MAE values than other competitive distributions. This indicates that the proposed NBurr distribution is competitive in almost all the networks, except a few, where Burr, MO Burr, and power-law cutoff distribution perform better than the proposed one. Power-law and Pareto distributions provide lower performance compared to the other competing distributions in terms of RMSE and MAE measures over all the real-world

TABLE 3. Network data sets and estimated parameters of the proposed NBurr model

Networks		# Nodes	# Edges	Statistical measures			Estimated-Parameters			Bootstrap chi-square value (p)	
				s	μ	$\frac{s}{\mu}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	\hat{c}	
Social	TwitterNet	81,306	1,768,149	57.965	21.747	2.6654	3.3395	0.0918	43.693	0.7324	0.9740
	GplusNet	107,614	13,673,453	1404.8	283.42	4.9568	1.3737	0.4503	24.924	0.5837	0.9890
	DeliciousNet	536,108	1,365,961	39.826	10.673	3.7312	7.6500	3.1845	26.999	0.2197	0.9651
	LiveJournalNet	4,847,571	68,993,773	44.969	15.368	2.926	7.7807	3.7317	79.102	0.3536	0.9720
	AthletesFacebookNet	13,866	86,859	17.978	12.438	1.4453	2.8158	-0.4558	24.053	1.1022	0.9400
Citation	HepThNet	27,770	352,807	43.139	15.220	2.8342	2.9198	-0.4475	26.793	0.7954	0.8300
	PatentsNet	3,774,768	16,518,948	6.9125	5.0687	1.3637	6.7724	3.9977	15.758	0.5793	0.8130
	CiteseerNet	227,320	814,134	9.8260	5.4322	1.8088	3.4008	-0.8157	12.286	0.8391	0.6160
Web	GoogleNet	875,713	5,105,039	43.320	7.1444	6.0634	16.595	5.6714	61.718	0.1137	0.9865
	BerkStanNet	685,230	7,600,595	300.08	12.316	24.364	0.4522	0.1833	1.1626	2.2058	0.6550
	Wikipedia2009Net	1,864,433	4,507,315	12.846	4.8903	2.6268	1.8443	-0.9706	3.0329	0.9147	0.9750
	Amazon0601Net	403,394	3,387,388	15.279	8.3989	1.8191	2.9534	1.8907	9.5274	0.9508	0.6920
CoPurchasing	Amazon0505Net	410,236	3,356,828	15.313	8.1826	1.8714	4.3982	3.3724	12.739	0.6965	0.6670
	Amazon0312Net	400,727	3,200,444	15.073	7.9865	1.8873	6.2523	5.5583	16.563	0.4971	0.6517
Temporal	MathoverflowNet	24,818	506,550	31.476	10.424	3.0195	0.5426	0.7039	1.1057	1.6187	0.9890
	SuperuserNet	194,085	1,443,339	23.782	5.8239	4.0836	0.5987	0.5741	0.9920	1.8685	0.9800
	AskubuntuNet	159,316	964,437	18.404	4.3856	4.1966	0.5997	1.1289	0.7386	2.0061	0.9760
Communication	EmailEnronNet	36,692	183,831	36.100	10.021	3.6027	4.1159	2.6481	6.3332	0.3129	0.9900
	WikiTalkNet	2,394,385	5,021,410	12.259	2.1195	5.7844	0.2091	2.7853	0.1733	3.4821	0.9760
	RecLibimsetiNet	220,970	17,359,346	413.71	102.85	4.0227	4.6667	1.8315	123.52	0.2107	0.9843
GroundTruth	WikiTopcatsNet	1,791,489	28,511,807	283.78	15.915	17.831	1.5987	0.8982	2.8452	0.7260	0.8400
	OrkutNet	3,072,441	117,185,083	154.78	76.281	2.0291	5.4171	5.0492	137.14	0.5217	0.9891
	YoutubeNet	1,134,890	2,987,624	50.754	5.2650	9.6398	2.5928	-0.4403	0.3159	0.5004	0.8520

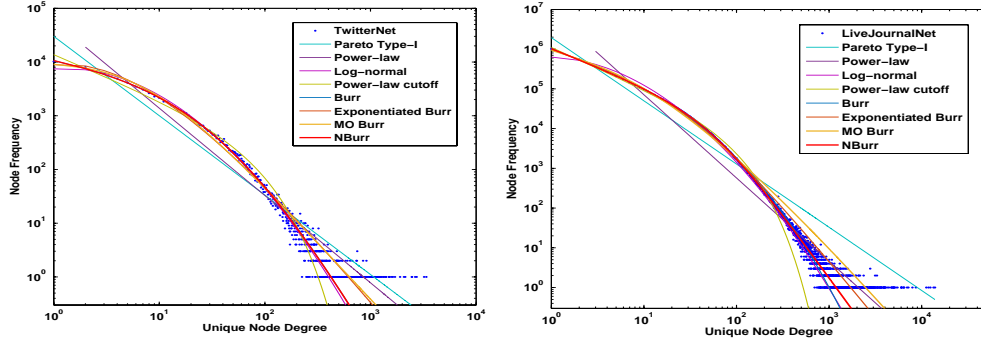


FIGURE 2. Degree distribution of TwitterNet and LiveJournalNet in log-log scale

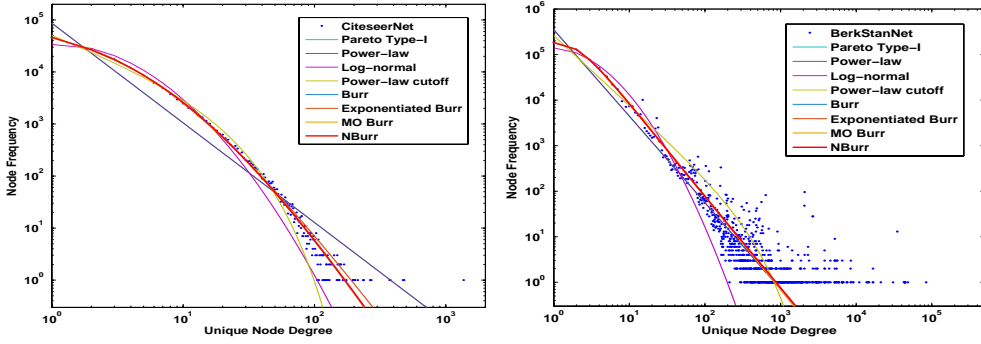


FIGURE 3. Degree distribution of CiteseerNet and BerkStanNet in log-log scale

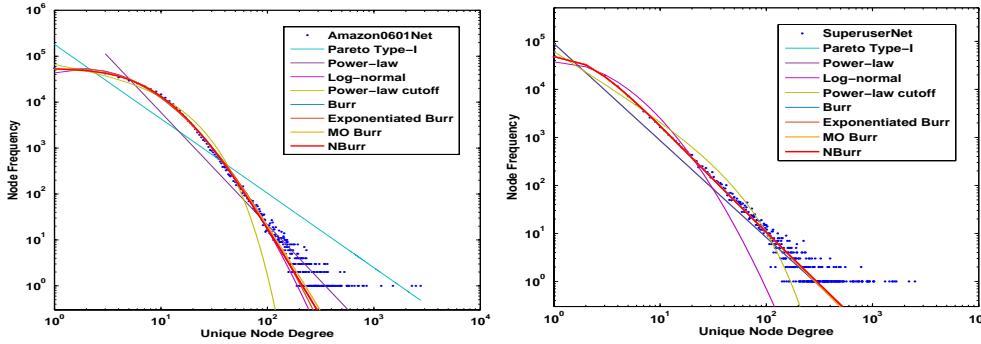


FIGURE 4. Degree distribution of Amazon0601Net and SuperuserNet in log-log scale

networks, as clearly seen from Table 5. Similar performances have been observed through Table 4 in the case of Burr and MO Burr distributions as they produce similar RMSE and MAE values. The dissimilarity between two probability distributions can also be measured by calculating KLD values. The higher similarity between the actual and the predicted distribution is achieved by generating high KLD values. The proposed NBurr distribution produces smaller KLD values compared to others in almost all the networks as clearly seen from Tables 4 and 5. This in-turn suggests that the proposed NBurr distribution is much closure to the observed degree distribution and always superior to the state-of-the-art models in almost all the networks.

TABLE 4. Performances of the proposed NBurr model in terms of RMSE, KLDiv, and MAE compared to the competitive heavy-tailed models over real-world networks

Networks		Burr			NBurr			Exponentiated Burr			MO Burr		
		RMSE	KLDiv	MAE	RMSE	KLDiv	MAE	RMSE	KLDiv	MAE	RMSE	KLDiv	MAE
Social	TwitterNet	13.179	0.00788	1.1663	13.391	0.00787	1.1725	34.754	0.02012	3.1259	28.727	0.01911	2.7780
	GplusNet	2.7411	0.05621	0.1994	1.8217	0.05572	0.1849	9.0526	0.06382	0.3193	4.2128	0.05679	0.2171
	DeliciousNet	29.147	0.00499	1.7952	13.810	0.00453	1.1364	20.291	0.00464	1.4183	44.550	0.00627	2.4668
	LiveJournalNet	384.81	0.00168	13.924	219.73	0.00050	5.0602	464.51	0.00337	19.028	916.55	0.00951	32.342
	AthletesFacebookNet	4.9857	0.00886	1.3876	4.8669	0.00885	1.3721	10.288	0.01361	2.6042	8.0527	0.00965	1.8504
Citation	HepThNet	2.6780	0.01331	0.4718	2.4786	0.01327	0.4561	11.985	0.02091	1.0592	2.8134	0.01346	0.4756
	PatentsNet	626.74	0.00024	64.674	68.176	6.5E-05	10.397	173.55	9.4E-05	22.971	898.55	0.00030	84.917
	CiteseerNet	43.719	0.00200	3.6469	33.484	0.00186	3.1622	13.165	0.00216	1.8519	39.521	0.00192	3.4805
Web	GoogleNet	187.96	0.00699	8.3224	117.15	0.00428	5.8484	197.65	0.00640	8.1009	177.87	0.00708	8.2026
	BerkStanNet	32.321	0.02671	0.5353	32.388	0.02679	0.5355	31.945	0.02670	0.5260	32.198	0.02671	0.5301
	Wikipedia2009net	176.29	0.00115	9.4514	52.148	0.00104	4.7242	91.517	0.00097	6.0324	168.35	0.00115	9.1700
Product CoPurchasing	Amazon0601Net	92.912	0.00315	7.6485	77.387	0.00283	6.4876	188.63	0.00719	14.032	87.137	0.00265	6.9522
	Amazon0505Net	118.17	0.00375	9.1393	78.727	0.00297	6.9574	245.89	0.01183	18.526	112.00	0.00335	8.5741
	Amazon0312Net	105.14	0.00337	8.1901	71.210	0.00260	6.0427	106.63	0.00342	8.2086	80.538	0.00263	6.4877
Temporal	MathoverflowNet	8.8478	0.01432	1.3720	6.1401	0.01381	1.1713	6.5265	0.01421	1.2302	9.2084	0.01432	1.3693
	SuperuserNet	17.943	0.00319	1.3692	11.709	0.00303	1.0642	18.326	0.00319	1.3717	18.797	0.00319	1.3977
	AskubuntuNet	33.041	0.00352	2.1017	27.986	0.00334	1.8650	34.267	0.00375	2.1469	31.861	0.00326	1.9377
Communication	EmailEnronNet	76.787	0.03525	5.2356	73.139	0.03381	5.0662	74.244	0.03497	5.1816	76.118	0.03516	5.2236
	WikiTalkNet	517.42	0.00343	21.725	146.89	0.00109	8.2835	665.73	0.00356	25.758	668.30	0.00356	25.823
	RecLibimsetiNet	41.348	0.03263	0.9010	22.851	0.03400	0.7659	48.044	0.06648	1.4156	44.413	0.06273	1.3570
GroundTruth	WikiTopcatsNet	9.6105	0.00188	0.0913	6.1751	0.00187	0.0767	10.724	0.00189	0.0961	9.2087	0.00188	0.0898
	OrkutNet	194.24	0.00721	6.2222	95.580	0.00298	3.3811	654.44	0.12206	35.536	116.81	0.00329	3.1295
	YoutubeNet	37.006	0.00112	0.5858	33.861	0.00112	0.5684	46.425	0.00116	0.6403	31.298	0.00111	0.5312

TABLE 5. Performances of the proposed NBurr model in terms of RMSE, KLDiv, and MAE compared to the competitive heavy-tailed models over real-world networks

Networks		Burr			NBurr			Exponentiated Burr			MO Burr		
		RMSE	KLDiv	MAE	RMSE	KLDiv	MAE	RMSE	KLDiv	MAE	RMSE	KLDiv	MAE
Social	TwitterNet	204.35	0.1831	10.847	354.25	0.2857	15.603	53.863	0.0169	2.9494	68.004	0.0397	4.1974
	GplusNet	53.064	0.2299	0.9221	86.955	0.3113	1.1847	10.155	0.0678	0.2523	30.925	0.1475	0.6821
	DeliciousNet	349.66	0.2021	14.867	471.02	0.1349	17.874	281.34	0.0579	10.781	66.896	0.0185	4.2366
	LiveJournalNet	5025.2	0.1614	127.98	8100.9	0.1785	164.18	3473.6	0.0355	70.64	808.79	0.0101	24.501
	AthletesFacebookNet	100.16	0.2049	13.387	204.91	0.4164	23.839	15.461	0.0127	2.5674	25.099	0.0324	4.3180
Citation	HepThNet	73.531	0.1741	4.0821	122.79	0.2566	5.997	22.59	0.0255	2.331	25.42	0.0464	2.816
	PatentsNet	27.5K	0.2266	2049.5	34.8K	0.2366	2533.2	9612.7	0.0192	725.71	2424.5	0.0061	271.89
	CiteseerNet	889.88	0.3308	49.467	1156.2	0.2916	62.026	353.26	0.0299	21.921	195.67	0.0131	13.877
Web	GoogleNet	1809.1	0.124	45.023	1809.2	0.124	45.023	1514.5	0.0878	40.067	188.01	0.0157	9.6549
	BerkStanNet	615.03	0.1863	4.0722	615.01	0.1863	4.0721	185.04	0.1002	2.0198	322.63	0.1037	2.8203
	Wikipedia2009Net	4371.9	0.1352	164.58	4371.9	0.1352	164.58	2720.9	0.0798	116.43	781.87	0.0082	35.531
Product CoPurchasing	Amazon0601Net	1495.4	0.2708	70.281	2539.8	0.4022	114.59	286.61	0.0102	16.881	297.39	0.0382	22.199
	Amazon0505Net	1572.9	0.2463	73.003	2494.5	0.3711	111.56	358.59	0.0125	19.123	260.85	0.0342	20.136
	Amazon0312Net	1564.4	0.2425	71.875	2462.9	0.3686	109.21	338.03	0.0116	17.742	273.39	0.0352	20.381
Temporal	MathoverflowNet	213.91	0.2131	13.600	213.82	0.2132	13.612	41.934	0.0634	5.3161	92.603	0.0861	7.9912
	SuperuserNet	900.04	0.1808	33.837	900.33	0.1808	33.839	243.42	0.0616	13.199	354.72	0.0570	16.613
	AskubuntuNet	949.66	0.2091	39.419	949.73	0.2091	39.420	212.91	0.0649	12.451	389.14	0.0719	20.113
Communication	EmailEnronNet	246.51	0.1779	14.886	245.25	0.1778	14.859	121.47	0.0873	8.445	95.468	0.0689	7.664
	WikiTalkNet	9669.4	0.3376	293.63	9669.4	0.3376	293.63	7978.6	0.1902	246.26	672.32	0.0036	25.905
	RecLibimsetiNet	77.081	0.2198	2.1486	133.91	0.2096	2.7441	87.472	0.0755	1.4021	28.059	0.0359	0.6971
Groundtruth	WikiTopcatsNet	565.21	0.1377	2.6145	930.44	0.1612	3.8073	272.99	0.0464	1.5159	389.86	0.0629	2.2289
	OrkutNet	2443.6	0.5498	80.712	4299.3	0.8033	101.64	452.92	0.0459	19.624	261.75	0.0479	16.197
	YoutubeNet	1380.5	0.1342	15.690	1380.5	0.1342	15.691	1422.2	0.1416	17.219	143.79	0.0045	2.2564

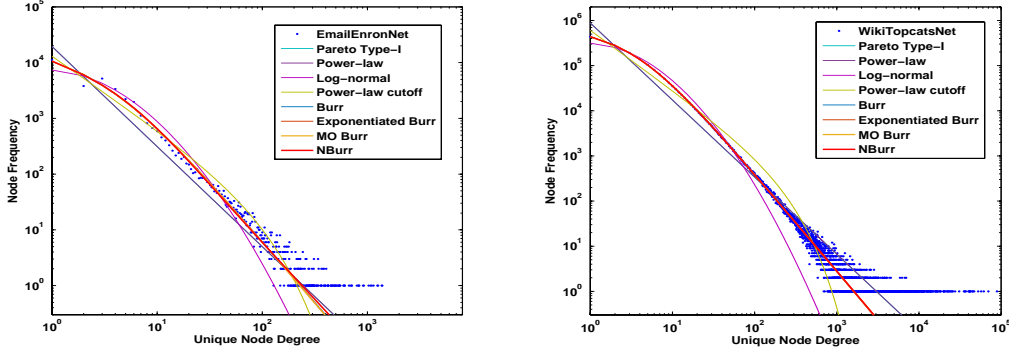


FIGURE 5. Degree distribution of EmailEnronNet and WikiTopcatsNet in log-log scale

Thus in conclusion we can say that the proposed NBurr distribution performs better than the competing distributions by considering all the measures (RMSE, MAE, and KLDiv) together. This suggests/confirms that the observed distribution corresponding to a real-world network plausibly comes from the proposed NBurr distribution.

The scatter plot of the fitted results can be used to verify the effectiveness of the proposed NBurr distribution. To do this, the log-log plots of the original frequency distribution, the estimated frequency by NBurr distribution, and the frequency estimated by Burr, Exponentiated Burr, MO Burr, Power-law, Pareto Type-I, Log-normal and Power-law Cutoff distributions are drawn corresponding to a real-world network. For more clarification, eight such plots are given in Figures 2–5. These are the TwitterNet, LiveJournalNet, CiteseerNet, BerkStanNet, Amazon0601Net, SuperuserNet, EmailEnronNet, and WikiTopcatsNet. From Figures 2–5, it is quite clear that the proposed NBurr curve always passes through the middle of the scatter plot of each of the observed distribution. This signifies that the proposed NBurr distribution provides a better fit compared to the other competitive distributions in almost all of the networks. Thus, it is visually clear, through observing the plotted results, we may now conclude that the entire node degree distribution can be better represented by the NBurr distribution compared to other heavy-tailed distributions. Hence the proposed NBurr distribution, a modification of the Burr distribution with nonlinear exponent in the shape parameter, can be used for effective and efficient modeling of the entire degree distribution of real-world networks without ignoring the lower degree nodes. Finally, it is clear that these heavy-tailed network data sets when modeled in the whole range using the proposed NBurr distribution shows significant improvements in comparison to state-of-the-art models.

5. Conclusion

In this paper, we present a new method called shape parameter transformation (SPT) to extend Burr and related families. The SPT method's idea is to use a nonlinear exponent (depends on data and an additional parameter) instead of using a constant shape parameter in the Burr, Pareto, and exponentiated family of distributions. The method was first applied to the Burr distribution and a generalized Burr (GBurr) model is introduced. These newly introduced GBurr models belong to the maximum domain of attraction of the Frechet distribution and are right-tail equivalent to the Pareto distribution. Further, the SPT method is applied to the classical Pareto and exponentiated

distributions. We also studied a relevant model, namely NBurr distribution, from the generalized Burr family and derived various statistical properties. The practical usefulness of the proposed NBurr distribution was shown using multiple heavy-tailed network data sets from different disciplines. It is interesting to note that the NBurr distribution is a new competitor for popularly used MO Burr and exponentiated Burr distributions. An immediate extension of this work is to apply these newly introduced probability models for modeling survival and lifetime data sets. Another possible extension of this work would be to look for implementing the proposed SPT method for other size distributions available in the statistics paradigm.

Acknowledgement. The authors are thankful to Professor Gopal K. Basak of Indian Statistical Institute, Kolkata for constructive comments and insightful suggestions in this paper.

REFERENCES

- [1] AFIFY, A. Z.—CORDEIRO, G. M.—ORTEGA, E. M.—YOUSOF, H. M.—BUTT, N. S.: *The four-parameter Burr XII distribution: Properties, regression model, and applications*, Comm. Statist. Theory Methods **47** (2018), 2605–2624.
- [2] AMARAL, L. A. N.—SCALA, A.—BARTHELEMY, M.—STANLEY, H. E.: *Classes of small-world networks*, Proc. Natl. Acad. Sci. USA **97** (2000), 11149–11152.
- [3] ASADI, M.—EBRAHIMI, N.: *Residual entropy and its characterizations in terms of hazard function and mean residual life function*, Statist. Probab. Lett. **49** (2000), 263–269.
- [4] ASADL, M.: *Characterization of the pearson system of distributions based on reliability measures*, Statist. Papers **39** (1998), 347–360.
- [5] AUSTIN, J. A.: *Control chart constants for largest and smallest in sampling from a normal distribution using the generalized Burr distribution*, Technometrics **15** (1973), 931–933.
- [6] BARABÁSI, A. L.: *The origin of bursts and heavy tails in human dynamics*, Nature **435** (2005), 207–211.
- [7] BARABÁSI, A. L.—ALBERT, R.: *Emergence of scaling in random networks*, Science **286** (1999), 509–512.
- [8] BROIDO, A. D.—CLAUSET, A.: *Scale-free networks are rare*, Nature Communications **10** (2019), 1–10.
- [9] BURR, I. W.: *Cumulative frequency functions*, Ann. Math. Statist. **13** (1942), 215–232.
- [10] CHATTOPADHYAY, S.—CHAKRABORTY, T.—GHOSH, K.—DAS, A. K.: *Uncovering patterns in heavy-tailed networks: A journey beyond scale-free*. In: 8th ACM IKDD CODS and 26th COMAD, 2021.
- [11] CHATTOPADHYAY, S.—MURTHY, C. A.—PAL, S. K.: *Fitting truncated geometric distributions in large scale real world networks*, Theoret. Comput. Sci. **551** (2014), 22–38.
- [12] CLAUSER, A.—SHALIZI, C. R.—NEWMAN, M. E. J.: *Power-law distributions in empirical data*, SIAM Review **51** (2009), 661–703.
- [13] DOMMA, F.: *Some properties of the bivariate Burr type III distribution*, Statistics **44** (2010), 203–215.
- [14] DUNNING, K. A.—HANSON, J. N.: *Generalized pearson distributions and nonlinear programing*, J. Stat. Comput. Simul. **6** (1977), 115–128.
- [15] EMBRECHTS, P.—KLÜPPPELBERG, C.—MIKOSCH, T.: *Modelling Extremal Events: for Insurance and Finance*, Springer Science & Business, Vol. 33, 2013.
- [16] FISK, P. R.: *The graduation of income distributions*, Econometrica **29** (1961), 171–185.
- [17] GOMES, A. E.—DA SILVA, C. Q.—CORDEIRO, G. M.: *Two extended Burr models: Theory and practice*, Comm. Statist. Theory Methods **44** (2015), 1706–1734.
- [18] GUPTA, R. C.—GUPTA, P. L.—GUPTA, R. D.: *Modeling failure time data by lehman alternatives*, Comm. Statist. Theory Methods **27** (1998), 887–904.
- [19] GUPTA, R. D.—GUPTA, R. C.: *Analyzing skewed data by power normal model*, Test **17** (2008), 197–210.
- [20] GUPTA, R. D.—KUNDU, D.: *Generalized exponential distributions*, Aust. N. Z. J. Stat. **41** (1999), 173–188.
- [21] JAMAL, F.—CHESNEAU, C.—NASIR, M. A.—SABOOR, A.—ALTUN, E.—KHAN, M. A.: *On a modified Burr XII distribution having flexible hazard rate shapes*, Math. Slovaca **70** (2020), 193–212.
- [22] JAYAKUMAR, K.—MATHEW, T.: *On a generalization to Marshall-Olkin scheme and its application to Burr type XII distribution*, Statist. Papers **49** (2008), 421–439.

- [23] KIM, M.—LESKOVEC, J.: *Multiplicative attribute graph model of real-world networks*, Internet Math. **8** (2012), 113–160.
- [24] KLEIBER, C.—KOTZ, S.: *Statistical Size Distributions in Economics and Actuarial Sciences*, John Wiley & Sons **470**, 2003.
- [25] KUMAR, D.: *The Burr type XII distribution with some statistical properties*, J. Data Sci. **16** (2017), 509–534.
- [26] KUMAR, D.—SARAN, J.—JAIN, N.: *The exponentiated Burr XII distribution: moments and estimation based on lower record values*, Sri Lankan J. Appl. Stat. **18** (2017), 1–18.
- [27] LEHMANN, E. L.: *The power of rank tests*, Ann. Math. Statist. **24** (1953), 23–43.
- [28] LESKOVEC, J.—CHAKRABARTI, D.—KLEINBERG, J.—FALOUTSOS, C.—GHAHRAMANI, Z.: *Kronecker graphs: An approach to modeling networks*, J. Mach. Learn. Res. **11** (2010), 985–1042.
- [29] LOMAX, K. S.: *Business failures: Another example of the analysis of failure data*, J. Amer. Statist. Assoc. **49** (1954), 847–852.
- [30] MARSHALL, A. W.—OLKIN, I.: *A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families*, Biometrika **84** (1997), 641–652.
- [31] MUDHOLKAR, G. S.—SRIVASTAVA, D. K.—FREIMER, M.: *The exponentiated Weibull family: A reanalysis of the bus-motor-failure data*, Technometrics **37** (1995), 436–445.
- [32] NADARAJAH, S.—KOTZ, S.: *The exponentiated type distributions*, Acta Appl. Math. **92** (2006), 97–111.
- [33] NEWMAN, M. E. J.: *The structure of scientific collaboration networks*, Proc. Natl. Acad. Sci. USA **98** (2001), 404–409.
- [34] NEWMAN, M. E. J.: *The structure and function of complex networks*, SIAM Review **45** (2003), 167–256.
- [35] NEWMAN, M. E. J.: *Power laws, Pareto distributions and Zipf’s law*, Contemp. Phys. **46** (2005), 323–351.
- [36] PARANAÍBA, P. F.—ORTEGA, E. M.—CORDEIRO, G. M.—PESCIM, R. R.: *The beta Burr XII distribution with application to lifetime data*, Comput. Statist. Data Anal. **55** (2011), 1118–1136.
- [37] PEARSON, K.: *Contributions to the mathematical theory of evolution*, Philosophical Transactions of the Royal Society of London **185** (1894), 71–110.
- [38] RODRIGUEZ, R. N.: *A guide to the Burr type XII distributions*, Biometrika **64** (1977), 129–134.
- [39] SÁNCHEZ, E.: *Burr type XII as a superstatistical stationary distribution*, Physica A: Stat. Mech. Appl. **516** (2019), 443–446.
- [40] STUMPF, M. P.—WIUF, C.—MAY, R. M.: *Subnets of scale-free networks are not scale-free: sampling properties of networks*, Proc. Natl. Acad. Sci. USA **102** (2005), 4221–4224.
- [41] TADIKAMALLA, P. R.: *A look at the Burr and related distributions*, Int. Stat. Rev. **48** (1980), 337–344.
- [42] TAKAHASI, K.: *Note on the multivariate Burr’s distribution*, Ann. Inst. Statist. Math. **17** (1965), 257–260.
- [43] VOITALOV, I.—HOORN, P. V.—HOFSTAD, R. V.—KRIOUKOV, D.: *Scale-free networks well done*, Phys. Rev. Research **1** (2019), Art. 033034.
- [44] YANG, J.—LESKOVEC, J.: *Defining and evaluating network communities based on ground-truth*, Knowl. Inf. Syst. **42** (2015), 181–213.

Received 19. 9. 2020

Accepted 11. 1. 2021

** Department of Science and Engineering*

Sorbonne University Abu Dhabi

Abu Dhabi

UAE

E-mail: tanujit_r@isical.ac.in

*** Department of Data Science*

S P Jain School of Global Management

Mumbai-400070

INDIA

E-mail: suchismita.das@spjain.org

**** Machine Intelligence Unit*

Indian Statistical Institute

203, B.T. Road

Kolkata-700108

INDIA

E-mail: swarup_r@isical.ac.in