

The new Burr distribution and its application

Gholamhossein Yari¹ · Zahra Tondpour¹

Received: 11 August 2016 / Accepted: 14 November 2016 / Published online: 17 January 2017
© The Author(s) 2017. This article is published with open access at Springerlink.com

Abstract This paper derives a new family of Burr-type distributions as new Burr distribution. This particular skewed distribution that can be used quite effectively in analyzing lifetime data. It is observed that the new distribution has modified unimodal hazard function. Various properties of the new Burr distribution, such that moments, quantile functions, hazard function, and Shannon's entropy are obtained. The exact form of the probability density function and moments of i th-order statistics in a sample of size n from new Burr distribution are derived. Estimation of parameters and change-point of hazard function by the maximum likelihood method are discussed. Change-point of hazard function is usually of great interest in medical or industrial applications. The flexibility of the new model is illustrated with an application to a real data set. In addition, a goodness-of-fit test statistic based on the Rényi Kullback–Leibler information is used.

Keywords Burr distributions · Change-point · Goodness-of-fit · Modified unimodal hazard function · Lifetime data analysis · Rényi Entropy

Mathematics Subject Classification 62E15, 62N05, 62F10, 60K10

Introduction

Burr [2] developed the system of Burr distributions. The Burr system of distributions includes 12 types of cumulative distribution functions which yield a variety of density shapes. The attractiveness of this relatively unknown family of distributions for model fitting is that it combines a simple mathematical expression for cumulative frequency function with coverage in the skewness–kurtosis plane. Many standard theoretical distributions, including the Weibull, exponential, logistic, generalized logistic, Gompertz, normal, extreme value, and uniform distributions, are special cases or limiting cases of the Burr system of distributions (see [11]). Family of Burr-type distributions is a very popular distribution family for modelling lifetime data and for modelling phenomenon with monotone and unimodal failure rates (see, for example, [13, 18]).

Analogous to the Pearson system of distributions, the Burr distributions are solutions to a differential equation, which has the form:

$$\frac{dy}{dx} = y(1 - y)g(x, y), \quad (1.1)$$

where y equal to $F(x)$ and $g(x, y)$ must be positive for y in the unit interval and x in the support of $F(x)$. Different functional forms of $g(x, y)$ result in different solutions $F(x)$, which define the families of the Burr system. For example, Burr II distribution is obtained when $g(x, y) = g(x) = \frac{ke^{-x}(1+e^{-x})^{k-1}}{(1+e^{-x})^k - 1}$.

In this paper, we derive a new distribution of Burr-type distributions which is more flexible by replacing $g(x, y)$ with $g(x) = \frac{3px^2e^{-x^3}(1+e^{-x^3})^{p-1}}{(1+e^{-x^3})^p - 1}$, ($p > 0$). We refer to this new distribution as the new Burr distribution. If $g(x, y)$ is taken

✉ Gholamhossein Yari
yari@iust.ac.ir

¹ School of Mathematics, Iran University of Science and Technology, Tehran, Iran

to be $g(x)$, then the solution of the differential Eq. 1.1 is given by

$$F(x) = (e^{-G(x)} + 1)^{-1}, \tag{1.2}$$

where $G(x) = \int g(x)dx$.

Hence, cdf and pdf of new Burr distribution are, respectively, given by

$$F(x; p) = (1 + e^{-x^3})^{-p}, -\infty < x < \infty, (p > 0), \tag{1.3}$$

and

$$f(x; p) = 3px^2 e^{-x^3} (1 + e^{-x^3})^{-p-1}, -\infty < x < \infty. \tag{1.4}$$

If the location parameter μ and the scale parameter σ are introduced in the equation 1.3, we have

$$F(x; \mu, \sigma, p) = (1 + e^{-(\frac{x-\mu}{\sigma})^3})^{-p}, -\infty < x < \infty, (p, \sigma > 0, \mu \in \mathbb{R}) \tag{1.5}$$

and

$$f(x; \mu, \sigma, p) = \frac{3p}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^2 e^{-(\frac{x-\mu}{\sigma})^3} \left(1 + e^{-(\frac{x-\mu}{\sigma})^3}\right)^{-p-1}. \tag{1.6}$$

Hence, Eq. 1.5 is three parameter new Burr distribution. Hazard function associated with the new Burr distribution is

$$h(x; \mu, \sigma, p) = \frac{\frac{3p}{\sigma} \left(\frac{x-\mu}{\sigma}\right)^2 e^{-(\frac{x-\mu}{\sigma})^3} \left(1 + e^{-(\frac{x-\mu}{\sigma})^3}\right)^{-p-1}}{1 - \left(1 + e^{-(\frac{x-\mu}{\sigma})^3}\right)^{-p}}. \tag{1.7}$$

The shapes of density and hazard functions of the new Burr distribution for different values of shape parameter p are illustrated in Fig. 1.

New Burr distribution has unimodal and bimodal pdfs. None of the 12 types of Burr distributions has this feature. Data that exhibit bimodal behavior arises in many different disciplines. In medicine, urine mercury excretion has two

peaks, see, for example, [5]. In material characterization, a study conducted by [4], grain size distribution data reveals a bimodal structure. In meteorology, [19] indicated that, water vapor in tropics, commonly have bimodal distributions. To see more applications of bimodal distributions, see [7–9, 16].

The reminder of the paper is organized as follows: properties of the new Burr distribution, such that moments, quantile functions, hazard function, Shannon’s entropy, and distribution of its order statistics are discussed in Sects. 2, 3, and 4. In Sect. 5, estimation of parameters and change-point of hazard function by the maximum likelihood method are discussed, and in Sect. 6, we establish a goodness-of-fit test statistic based on the Rényi Kullback–Leibler information for testing new Burr model. Finally, in Sect. 7, we present an illustrative example. Section 8 provides conclusions.

Properties of the new Burr distribution

New Burr distribution has unimodal and bimodal pdfs. The modes of distribution are provided by differentiating the density of new Burr distribution in 1.6 with respect to x :

$$f'(x; \mu, \sigma, p) = 3\left(\frac{x-\mu}{\sigma}\right)^3 \left(1 - pe^{-(\frac{x-\mu}{\sigma})^3}\right) - 2\left(1 + e^{-(\frac{x-\mu}{\sigma})^3}\right). \tag{2.1}$$

The derivative $f'(x; \mu, \sigma, p)$ exists every where, hence critical point(s) satisfy equation $f'(x; \mu, \sigma, p) = 0$. In 2.1, set $\mu = 0$ and $\sigma = 1$, because location and scale parameters will not affect the distribution shape. Thus, equation $f'(x; \mu, \sigma, p) = 0$ simplifies to

$$3x^3(1 - pe^{-x^3}) - 2(1 + e^{-x^3}) = 0. \tag{2.2}$$

Analytical solution of 2.2 is not possible. Numerical approximation of modes using the midpoint method is applied to study the modes. The distance between the two

Fig. 1 Graphs of density and hazard functions of the new Burr distribution for different values of shape parameter p

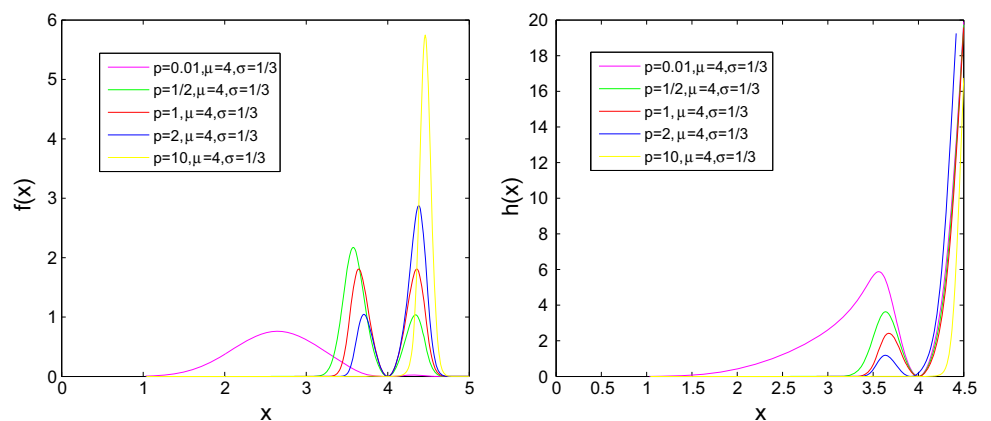


Table 1 Distance between two modes of new Burr distribution

p	Distance
0.3	2.4294
0.5	2.2930
1	2.1388
2	2.0191
3	1.9682
4	1.9409
5	1.9243

modes is demonstrated in Table 1. From Table 1, it is observed that when p increases, the distance between two modes decreases, and for $0 < p < 1$, when p decreases, value of pdf in the second mode decreases to zero and pdf will be almost unimodal, and for $p = 1$, values of pdf in two modes are the same but for $p > 1$, and when p increases, value of pdf in the first mode decreases to zero and pdf will be almost unimodal. Hence, the new Burr distribution can be used to analyse different kinds of lifetime data sets with unimodal and bimodal shapes of pdf.

The new Burr distribution has modified unimodal (unimodal followed by increasing) hazard function, and when p increases, hazard function will be almost increasing.

The main purpose in this paper is to describe and fit the data sets with non-monotonic hazard function, such as the bathtub, unimodal and modified unimodal hazard function. Many modifications of important lifetime distributions have achieved the above purpose, but unfortunately, the number of parameters has increased, the forms of survival and hazard functions have been complicated, and estimation problems have risen. More over some of the modifications do not have a closed form for their cdfs. However, this new distribution with one parameter and simple form of cdf achieves this purpose.

Now, we discuss the reverse hazard function of the new Burr distribution. The reverse hazard function of any distribution function $F(x)$ can be defined as $r(x) = \frac{f(x)}{F(x)}$. Consequently, the reversed hazard function of new Burr distribution with zero location parameter and unit scale parameter is given by

$$r(x; p) = \frac{3x^2 e^{-x^3}}{1 + e^{-x^3}} p.$$

The reversed hazard function has recently attracted considerable interest of researchers (see, for example, [1, 3]). In a reliability setting, the reversed hazard function (multiplied by dx) defines the conditional probability of a failure of an object in $(x - dx, x]$ given that the failure had occurred in $[0, x]$. The reversed hazard function of new Burr distribution with zero location parameter and unit scale parameter is a linear function of p .

The r th moment about origin of the new Burr distribution is given by

$$\mu_r = E(X^r) = \int_{-\infty}^{\infty} x^r \left(\frac{3p}{\sigma} \left(\frac{x-\mu}{\sigma} \right)^2 e^{-\left(\frac{x-\mu}{\sigma}\right)^3} \left(1 + e^{-\left(\frac{x-\mu}{\sigma}\right)^3} \right)^{-p-1} \right) dx,$$

using the change of variable, $t = \frac{1}{1 + e^{-\left(\frac{x-\mu}{\sigma}\right)^3}}$, $0 < t < 1$, we obtain

$$\begin{aligned} E(X^r) &= \int_0^1 p t^{p-1} \left(\sigma \left(-\ln \left(\frac{1}{t} - 1 \right) \right) + \mu \right)^r dt \\ &= p \sum_{i=0}^r \binom{r}{i} (-\sigma)^i \mu^{r-i} \int_0^1 p t^{p-1} \left(\ln \left(\frac{1}{t} - 1 \right) \right)^i dt. \end{aligned}$$

Now, using $\frac{1}{t} - 1 = e^u$, $0 < u < \infty$, we obtain

$$\begin{aligned} E(X^r) &= p \sum_{i=0}^r \binom{r}{i} (-\sigma)^i \mu^{r-i} \int_0^{\infty} (e^u + 1)^{-p-1} u^i e^u du \\ &= p \sum_{i=0}^r \binom{r}{i} (-\sigma)^i \mu^{r-i} E_q(g(X)), \end{aligned}$$

where $E_q(\cdot)$ denotes expectation for $X \sim q$ and q is the standard exponential distribution and $g(x) = (e^x + 1)^{-p-1} x^i e^{2x}$. Using the importance sampling method, the importance sampling estimate of μ_r is given by

$$\hat{\mu}_{rq} = p \sum_{i=0}^r \binom{r}{i} (-\sigma)^i \mu^{r-i} \left(\frac{1}{n} \sum_{k=1}^n g(X_k) \right), \quad X_k \sim q. \tag{2.3}$$

Using $n = 1000$, the importance sampling estimate of mean and variance of the new Burr distribution as $\mu = 0$ and $\sigma = 1$ for different values of p is demonstrated in Table 2. From Table 2, it is observed that when p increases, mean increases and variance decreases. Mean and variance $\hat{\mu}_{rq}$ are given by

$$E(\hat{\mu}_{rq}) = \mu_r, \text{var}(\hat{\mu}_{rq}) = p^2 \sum_{i=0}^r \binom{r}{i} (-\sigma)^i \mu^{r-i} \text{var}(\hat{E}_q(g(X))),$$

where $\text{var}(\hat{E}_q(g(X))) = E_q((g(X) - E_q(g(X)))^2)$.

To form a confidence interval for μ_r , we need to estimate $\text{var}(\hat{E}_q(g(X)))$. Because X_k are sampled from q , the natural variance estimate is

Table 2 Importance sampling estimate of mean and variance of the new Burr distribution

p	Mean	Variance
1	-0.4984	0.3165
2	-0.2091	0.1475
3	-0.0921	0.0683
4	-0.0428	0.0296

$$\text{var}(\hat{E}_q(g(X))) = \frac{1}{n} \sum_{k=1}^n (g(X_k) - \hat{E}_q(g(X)))^2.$$

Then, an approximate 99% confidence interval for μ_r is $\hat{\mu}_{rq} \pm 2.58 \frac{\text{var}(\hat{\mu}_{rq})}{\sqrt{n}}$.

The quantile function, $Q(u)$, $0 < u < 1$, for the new Burr distribution can be computed using the formula:

$$Q(u) = \sigma \left(-\ln \left(u^{-\frac{1}{p}} - 1 \right) \right)^{\frac{1}{3}} + \mu.$$

The median of a new Burr distribution occurs at $\sigma \left(-\ln \left(\left(\frac{1}{2} \right)^{-\frac{1}{p}} - 1 \right) \right)^{\frac{1}{3}} + \mu$, and clearly, it is a decreasing function of p as $p \leq 1$ but an increasing function of p as $p \geq 1$.

Skewness and kurtosis of a parametric distribution are often measured by $\alpha_3 = \frac{\mu_3}{\sigma^3}$ and $\alpha_4 = \frac{\mu_4}{\sigma^4}$, respectively. When the third or fourth moment does not exist, for example, Cauchy, Lévy, and Pareto distributions, α_3 and α_4 , cannot be computed. For the new Burr distribution, skewness and kurtosis can be approximated by approximations of μ_3 and μ_4 or alternative measures for skewness and kurtosis, based on quantile functions. The measure of skewness S defined by [6] and the measure of kurtosis K defined by [12] are based on quantile functions and they are defined as

$$S = \frac{Q\left(\frac{6}{8}\right) - 2Q\left(\frac{4}{8}\right) + Q\left(\frac{2}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)}, \tag{2.4}$$

$$K = \frac{Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right) + Q\left(\frac{3}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{2}{8}\right)}. \tag{2.5}$$

To investigate the effect of the shape parameter p on the new Burr density function, Eqs. 2.4 and 2.5 are used to obtain Galton’s skewness and Moors’ kurtosis. Figure 2 displays the Galton’s skewness and Moors’ kurtosis for the

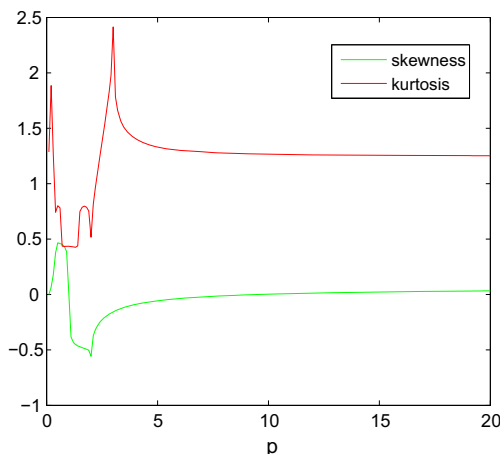


Fig. 2 Galton’s skewness and Moors’ kurtosis for the new Burr distribution

new Burr distribution in terms of the parameter p when $\mu = 0$ and $\sigma = 1$.

Shannon’s entropy

The entropy of a random variable X is a measure of variation of uncertainty. Shannon’s entropy [17] for a random variable X with pdf $f(x)$ is defined as $E(-\log(f(x)))$. In recent years, Shannon’s entropy has been used in many applications in fields of engineering, physics, and economics.

Denote by $H_{sh}(X)$ the well-known Shannon’s entropy. The following theorem gives the Shannon’s entropy of the new Burr distribution.

Theorem 3.1 *The Shannon’s entropy of the new Burr distribution is given by*

$$H_{sh}(X) = -\ln \frac{3p}{\sigma} - \frac{2p}{3} \sum_{i=0}^{\infty} \binom{-p-1}{i} \left(\frac{\Gamma'(1) - \ln(p+i)}{p+i} \right) - p \sum_{i=0}^{\infty} \binom{-p-1}{i} \frac{1}{(p+i)^2} + \frac{1}{p} + 1.$$

Proof

$$H_{sh}(X) = - \int_{-\infty}^{\infty} f(x) \ln f(x) dx = -\ln \frac{3p}{\sigma} - 2E \left(\ln \left(\frac{X-\mu}{\sigma} \right) \right) + E \left(\left(\frac{X-\mu}{\sigma} \right)^3 \right) + (p+1)E \left(\ln \left(1 + e^{-\left(\frac{X-\mu}{\sigma} \right)^3} \right) \right). \tag{3.1}$$

We need to find the expressions $E(\ln(\frac{X-\mu}{\sigma}))$, $E((\frac{X-\mu}{\sigma})^3)$ and $E(\ln(1 + e^{-\left(\frac{X-\mu}{\sigma}\right)^3}))$. First, we calculate the expectation of $(\frac{X-\mu}{\sigma})^r$.

$$E \left(\left(\frac{X-\mu}{\sigma} \right)^r \right) = \int_{-\infty}^{\infty} \frac{3p}{\sigma} \left(\frac{x-\mu}{\sigma} \right)^{r+2} e^{-\left(\frac{x-\mu}{\sigma}\right)^3} \left(1 + e^{-\left(\frac{x-\mu}{\sigma}\right)^3} \right)^{-p-1} dx,$$

using the change of variable, $t = \frac{1}{1+e^{-\left(\frac{x-\mu}{\sigma}\right)^3}}$, $0 < t < 1$, and then, the change of variable, $\frac{1}{t} - 1 = e^u$, $0 < u < \infty$, we obtain,

$$E \left(\left(\frac{X-\mu}{\sigma} \right)^r \right) = (-1)^r p \sum_{i=0}^{\infty} \binom{-p-1}{i} \frac{\Gamma\left(\frac{r}{3} + 1\right)}{(p+i)^{\frac{r}{3}+1}}. \tag{3.2}$$

For $r = 2k$

$$E \left(\left(\frac{X-\mu}{\sigma} \right)^{2k} \right) = p \sum_{i=0}^{\infty} \binom{-p-1}{i} \frac{\Gamma\left(\frac{2k}{3} + 1\right)}{(p+i)^{\frac{2k}{3}+1}}. \tag{3.3}$$

Differentiating both sides of 3.3 with respect to k at $k = 0$ leads to

$$E\left(\ln\left(\frac{X-\mu}{\sigma}\right)\right) = \frac{p}{3} \sum_{i=0}^{\infty} \binom{-p-1}{i} \frac{\Gamma'(1) - \ln(p+i)}{p+i}. \tag{3.4}$$

In the same way, by calculating $E((1 + e^{-\frac{(X-\mu)^3}{\sigma^3}})^r)$ and then differentiating with respect to r at $r = 0$, we obtain

$$E\left(\ln\left(1 + e^{-\frac{(X-\mu)^3}{\sigma^3}}\right)\right) = \frac{1}{p}. \tag{3.5}$$

By replacing 3.2, 3.4, and 3.5 in relation 3.1, the proof is completed. \square

Distribution of order statistics

The pdf of $X_{i:n}$ ($i = 1, \dots, n$) is given by

$$f_{i:n}(x; \mu, \sigma, p) = \frac{n!}{(i-1)!(n-i)!} f(x; \mu, \sigma, p) F^{i-1} \times (x; \mu, \sigma, p) (1 - F(x; \mu, \sigma, p))^{n-i},$$

where $f(x; \mu, \sigma, p)$ and $F(x; \mu, \sigma, p)$ are pdf and cdf given in 1.5 and 1.6, respectively:

$$f_{i:n}(x; \mu, \sigma, p) = \sum_{j=0}^{n-i} d_j(n, i) f(x; \mu, \sigma, p(i+j)), \tag{4.1}$$

where

$$d_j(n, i) = \frac{n!(-1)^j}{(i-1)!j!(n-i-j)!(i+j)}.$$

Note that $d_j(n, i)$ ($j = 0, 1, \dots, n-i$) are coefficients not dependent on p, μ , and σ . This means that $f_{i:n}(x; \mu, \sigma, p)$ is a weighted average of other new Burr.

From 1.6 and 4.1, we get the r^{th} moment of $X_{i:n}$ to be

$$E(X_{i:n}^r) = \sum_{j=0}^{n-i} d_j(n, i) E(X^r) = p \sum_{j=0}^{n-i} (i+j) d_j(n, i) \sum_{k=0}^r \binom{r}{k} (-\sigma)^k \mu^{r-k} E_q(g(Y)),$$

where X has new Burr distribution with parameters μ, σ and $p(i+j)$ and Y has q distribution, standard exponential distribution, and $g(y) = (e^y + 1)^{-p(i+j)-1} y^{\frac{k}{3}} e^{2y}$. Then, the importance sampling estimate of the r th moment about origin of $X_{i:n}$ is given by

$$p \sum_{j=0}^{n-i} (i+j) d_j(n, i) \sum_{k=0}^r \binom{r}{k} (-\sigma)^k \mu^{r-k} \left(\frac{1}{m} \sum_{l=1}^m g(Y_l) \right), Y_l \sim q. \tag{4.2}$$

The cdf of $X_{i:n}$ ($1 \leq i \leq n$) is given by

$$F_{i:n}(x) = I_{F(x)}(i, n-i+1),$$

where $I_x(a, b)$ is lower incomplete gamma function.

There, the 100 u th percentile of $X_{i:n}$ can be obtained by solving

$$F_{i:n}(x) = u. \tag{4.3}$$

The percentage points of $X_{i:n}$ can be evaluated from 4.3 using tables of incomplete beta function (see [15]). However, for $i = 1$, Eq. 4.3 reduces to $(1 - (1 + e^{-\frac{(X-\mu)^3}{\sigma^3}})^{-p})^n = 1 - u$. Thus, the 100 u -percentage point of the smallest order statistic $X_{1:n}$ is given by

$$F_{1:n}^{-1}(u; p, \mu, \sigma) = \mu + \sigma \left(-\ln \left(\left(1 - (1 - u)^{\frac{1}{n}} \right)^{\frac{1}{p}} - 1 \right) \right)^{\frac{1}{3}}.$$

Similarly, for $i = n$, the 100 u -percentage point of the largest order statistic is

$$F_{n:n}^{-1}(u; p, \mu, \sigma) = \mu + \sigma \left(-\ln \left(u^{-\frac{1}{np}} - 1 \right) \right)^{\frac{1}{3}}.$$

Hazard change-point estimation-classical approach

Hazard function plays an important role in reliability and survival analysis. New Burr distribution has modified unimodal (unimodal followed by increasing) hazard function. In some medical situations, for example, breast cancer, the hazard rate of death of breast cancer patients represents a modified unimodal shape.

A modified unimodal shape has three phases: first increasing, then decreasing, and then again increasing. It can be interpreted as a description of three groups of patients, first group is represented by the first phase that contains the weak patients, so the hazard rate of this group is increasing, while the second phase represents the group with strong patients, their bodies have become familiar with the disease and they are getting better. The hazard rate of death of these patients is decreasing. In the third phase, they become weaker and their ability to cope with the disease declines, then the hazard rate of death increases.

For situations, where the hazard function is modified unimodal shaped, usually, we have interest in the estimation of lifetime change-point, that is, the point at which the hazard function reaches to a maximum (minimum) and then decreases (increase). In reliability, the change-point of a hazard function is useful in assessing the hazard in the useful life phase. One of change-points of hazard function of the new Burr distribution is location parameter. In this section, we consider maximum likelihood estimation procedure for change-points of the hazard function.

Let us assume that x_1, \dots, x_n is a random sample of size n of lifetimes generated by a new Burr distribution with parameters μ , σ , and p . The log-likelihood function is given by

$$l(\mu, \sigma, p) = n \log \left(\frac{3p}{\sigma} \right) + 2 \sum_{i=1}^n \log \left(\frac{x_i - \mu}{\sigma} \right) - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^3 - (p+1) \sum_{i=1}^n \log \left(1 + e^{-\left(\frac{x_i - \mu}{\sigma} \right)^3} \right).$$

The maximum likelihood estimates for μ , σ , and p denoted by $\hat{\mu}$, $\hat{\sigma}$, and \hat{p} , respectively, are obtained solving the likelihood equations, $\left(\frac{\partial l}{\partial \mu} = 0, \frac{\partial l}{\partial \sigma} = 0, \text{ and } \frac{\partial l}{\partial p} = 0 \right)$. According to the above, maximum likelihood estimator of one of change-points is $\hat{\mu}$.

From the invariance property of maximum likelihood estimators, we can obtain maximum likelihood estimators for functions of μ , σ and p . For $\phi = g(\mu, \sigma, p)$, a one-to-one function of μ , σ , and p , and we have $\hat{\phi} = g(\hat{\mu}, \hat{\sigma}, \hat{p})$. Taking $\phi = h(x)$, defined in 1.7, the change-point of new Burr hazard function is obtained as solution of equation $\frac{d}{dx} \log(\phi) = 0$. The maximum likelihood estimator of the change-point is the solution of $\frac{d}{dx} \log(\phi) = 0$ with μ , σ , and p replaced by maximum likelihood estimates. We observe that $\frac{d}{dx} \log(\phi) = 0$ is non-linear in x , so the change-point of the hazard function estimate should be obtained using some one-dimensional root finding techniques, such as Newton–Raphson.

Testing new Burr model based on the Rényi Kullback–Leibler information

Test statistics

Suppose that we are interested in a goodness-of-fit test for

$$\begin{cases} H_0 : f(x) = f^0(x; \mu, \sigma, p) = \frac{3p}{\sigma} \left(\frac{x - \mu}{\sigma} \right)^2 e^{-\left(\frac{x - \mu}{\sigma} \right)^3} \left(1 + e^{-\left(\frac{x - \mu}{\sigma} \right)^3} \right)^{-p-1} \\ H_1 : f(x) \neq f^0(x; \mu, \sigma, p), \end{cases}$$

where μ , σ , and p are unknown.

We will denote the complete samples as $X_{1:n} < X_{2:n} < \dots < X_{n:n}$. For a null pdf $f^0(x)$, the Rényi Kullback–Leibler information from complete data is defined as

$$D_r(f; f^0) = \frac{1}{r-1} \log \int_{-\infty}^{\infty} \dots \int_{-\infty}^{x_{2:n}} \frac{(f_{X_{1:n}, \dots, X_{n:n}}(x_{1:n}, \dots, x_{n:n}))^r}{(f_{X_{1:n}, \dots, X_{n:n}}^0(x_{1:n}, \dots, x_{n:n}))^{r-1}} dx_1 \dots dx_n,$$

where $r > 0$ and $r \neq 1$. Because $D_r(f; f^0)$ has the property that $D_r(f; f^0) \geq 0$, and the equality holds if and only if $f = f^0$, the estimate of the Rényi Kullback–Leibler

information can be consider as a goodness-of-fit test statistic. For that purpose, the Rényi Kullback–Leibler information can be estimated by

$$D_r(f; f^0) = -H_r(X_{1:n}, \dots, X_{n:n}) - \sum_{j=1}^n f^0(x_j).$$

Thus, the test statistics based on $\frac{D_r(f; f^0)}{n}$ is given by

$$T_r = \frac{D_r(f; f^0)}{n} = \frac{1}{n} \left(-\hat{H}_r(X_{1:n}, \dots, X_{n:n}) - \sum_{j=1}^n f^0(x_j; \hat{\mu}, \hat{\sigma}, \hat{p}) \right),$$

where $\hat{\mu}$, $\hat{\sigma}$, and \hat{p} are MLEs of μ , σ , and p , respectively, and $\hat{H}_r(X_{1:n}, \dots, X_{n:n})$ is an estimate of Rényi entropy for sample $X_{1:n} < X_{2:n} < \dots < X_{n:n}$. Under the null hypothesis, T_r for r close to 1 will be close to 0, and therefore, large values of T_r will lead to the rejection of H_0 .

In this paper, we use estimation of Rényi entropy based on generalized nearest-neighbor graphs that is introduced by [14]. The basic tool to define their estimator was the generalized nearest-neighbor graph. This graph on vertex set V is a directed graph on V . The edge set of it contains for each $i \in S$ (S is a finite non-empty set of positive integers), an edge from each $x \in V$ to its i^{th} nearest neighbor according to the Euclidean distance to x .

For $p \geq 0$ denote by $L_p(V)$, the sum of the p^{th} powers of Euclidean lengths of its edges. According to proven theorem in [14]

$$\lim_{n \rightarrow \infty} \frac{L_p(X_{1:n}, \dots, X_{n:n})}{n^{1-\frac{p}{d}}} = \gamma > 0 \text{ a.s.},$$

where $p = d(1-r)$ and d is dimension of sample members.

Based on described graph, they estimated Rényi entropy by

$$\hat{H}_r(X_{1:n}, \dots, X_{n:n}) = \frac{1}{1-r} \log \frac{L_p(X_{1:n}, \dots, X_{n:n})}{\gamma n^{1-\frac{p}{d}}}.$$

Application

In this section, we consider an uncensored data set corresponding to remission times (in months) of a random sample of 128 bladder cancer patients. These data were previously reported in [10]. TTT plot for considered data is concave then convex indicating an increasing then decreasing hazard function, and is properly accommodated by new Burr distribution. Because in the system of Burr distributions, only Burr X and Burr XII distributions have unimodal hazard functions, and because of the similarity of cdf of the new Burr distribution with the Burr II distribution compared to the rest of distributions in Burr family,



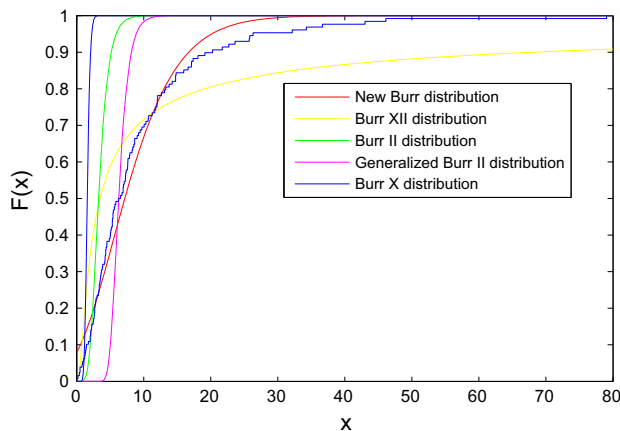


Fig. 3 cdfs of the new Burr, Burr X, Burr XII, Burr II, and generalized Burr II models for the remission times of bladder cancer data

Table 3 Values of test statistics for the remission times of bladder cancer data

Distribution	Test statistic ($T_{0.99}$)
New Burr	2.5732
Burr XII	4.4819
Burr II	6.6133
Generalized Burr II	11.0319
Burr X	193.9448

compare the fits of the new Burr distribution and those of Burr X, Burr XII, and Burr II and generalized Burr II. Plot of the estimated cdfs of models fitted to the data set is given in Fig. 3. Figure 3 and also values of defined test statistics in the previous section that are shown in Table 3 confirm that the new Burr distribution provides a significantly better fit than Burr X, Burr XII, Burr II, and generalized Burr II distributions. The required numerical evaluations are implemented using Matlab (version 2013) and R software (version 3.3.1).

Conclusions

We introduced a new family of Burr-type distributions as new Burr distribution. Various properties of the distribution are investigated. The distribution is found to be unimodal and bimodal. This new distribution with one parameter and simple form of cdf has modified unimodal (unimodal followed by increasing) hazard function. Hence, this new distribution can be used quite effectively in analyzing lifetime data with non-monotonic hazard function.

The method of maximum likelihood is suggested for estimating the parameters and change-points of hazard function of the new Burr distribution. In application to remission times (in months) of a random sample of 128 bladder cancer patients, the new Burr distribution provided a significantly better fit than Burr X, Burr XII, Burr II, and generalized Burr II distributions. This fact is confirmed by goodness-of-fit test statistic based on the Rényi Kullback–Leibler information.

Acknowledgements The authors acknowledge the Department of Mathematics, Iran University of Science and Technology.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Block, H., Savits, T.H., Singh, H.: The reversed hazard rate function. *Probab. Eng. Inf. Sci.* **12**, 69–90 (1998)
- Burr, I.W.: Cumulative frequency distribution. *Ann. Math. Stat.* **13**, 215–232 (1942)
- Chandra, N.K., Roy, D.: Some results on reversed hazard rate. *Probab. Eng. Inf. Sci.* **15**, 95–102 (2001)
- Dierickx, B., Vleugels, J., Van der Biest, O.: Statistical extreme value modeling of particle size distributions: experimental grain size distribution type estimation and parameterization of sintered zirconia D. *Mater. Charect.* **45**, 61–70 (2000)
- Ely, J.T.A., Fudenberg, H.H., Muirhead, R.J., LaMarche, M.G., Krone, C.A., Buscher, D., Stern, E.A.: Urine mercury in micromercurialism: bimodal distribution and diagnostic implications. *Bull. Environ. Contam. Toxicol.* **63**(5), 553–559 (1999)
- Galton, F.: *Enquiries into Human Faculty and its Development*. Macmillan and Company, London (1883)
- Garboś, S., Świecicka, D.: Application of bimodal distribution to the detection of changes in uranium concentration in drinking water collected by random daytime sampling method from a large water supply zone. *Chemosphere* **138**, 377–382 (2015)
- Khodabin, M., Ahmadabadi, A.: Some properties of generalized gamma distribution. *Math. Sci.* **4**, 9–28 (2010)
- Knapp, T.R.: Bimodality revisited. *J. Modern Appl. Stat. Methods* **6**, 8–20 (2007)
- Lee, E.T., Wang, J.W.: *Statistical Methods for Survival Data Analysis*, 3rd edn. Wiley, New York (2003)
- Lewis, A.W.: The Burr distribution as a general parametric family in survivorship and reliability theory applications. Ph.D. Thesis, University of North Carolina at Chapel Hill, USA (1981)
- Moors, J.J.: A quantile alternative for kurtosis. *Statistician* **37**, 25–32 (1988)
- Okasha, M.K., Matter, M.Y.: On the three-parameter Burr type XII distribution and its application to heavy tailed lifetime data. *J. Adv. Math.* **10**, 3429–3442 (2015)
- Pál, D., Szepesvári, Cs, Póczos, B.: Estimation of Rényi entropy and mutual information based on generalized nearest-neighbor graphs (2010). [arXiv:1003.1954](https://arxiv.org/abs/1003.1954)



15. Pearson, E.S., Hartley, H.O.: *Biometrika Tables for Statisticians*, vol. I, 3rd edn. Cambridge University Press, Cambridge (1970)
16. Sewell, M.A., Young, C.M.: Are echinoderm egg size distributions bimodal? *Biol. Bull.* **193**, 297–305 (1997)
17. Shannon, C.E.: The mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 249–428 (1948)
18. Surles, J.G., Padgett, W.J.: Inference for reliability and stress-strength for a scaled Burr Type X distribution. *Lifetime Data Anal.* **7**, 187–200 (2001)
19. Zhang, C., Mapes, B.E., Soden, B.J.: Bimodality in tropical water vapor. *Q. J. R. Meteorol. Soc.* **129**, 2847–2886 (2003)

