KYBERNETIKA — VOLUME *51* (2015), NUMBER 6, PAGES 933–959

COMPOUND GEOMETRIC AND POISSON MODELS

Nooshin Hakamipour, Sadegh Rezaei and Saralees Nadarajah

Many lifetime distributions are motivated only by mathematical interest. Here, eight new families of distributions are introduced. These distributions are motivated as models for the stress of a system consisting of components working in parallel/series and each component has a fixed number of sub-components working in parallel/series. Mathematical properties and estimation procedures are derived for one of the families of distributions. A real data application shows superior performance of a three-parameter distribution (performance assessed with respect to Kolmogorov–Smirnov statistics, AIC values, BIC values, CAIC values, AICc values, HQC values, probability-probability plots, quantile-quantile plots and density plots) versus thirty one other distributions, each having at least three parameters.

Keywords: exponential distribution, exponentiated exponential distribution, maximum likelihood estimation

Classification: 62E99

1. INTRODUCTION

The statistical literature describes numerous distributions for modeling lifetime data. The aim of this paper is to introduce new families of distributions with sound physical motivation. The proposed families encompass several known families of distributions, including Marshall Olkin G distributions due to Marshall and Olkin [15], exponentiated G distributions due to Gupta et al. [5] and exponentiated exponential Poisson G distributions due to Ristic and Nadarajah [26]. As explained in subsequent sections, one of the proposed three-parameter distributions provides better fits (as judged by Kolmogorov–Smirnov statistics, AIC values, BIC values, CAIC values, AICc values, HQC values, probability-probability plots, quantile-quantile plots and density plots) than thirty one other distributions, each having at least three parameters. We feel that this is a remarkable feature.

There are many ways that reliability of a system can be modeled. A common way is to use parallel / series components in one stage, two stages or more stages (Kolowrocki [11]). The use of one stage may be too simplistic. The use of three or more stages may be too complicated. An ideal choice may be a two-stage system with each stage involving

DOI: 10.14736/kyb-2015-6-0933

parallel / series components. In this case, possible configurations of the system can be one of

i) first stage having components working in parallel and each component having sub-components working in parallel, so the system will fail if and only if all the sub-components of every component fail;

ii) first stage having components working in series and each component having sub-components working in series, so the system will fail if and only if at least one sub-component of at least one component fails;

iii) first stage having components working in parallel and each component having sub-components working in series, so the system will fail if and only if at least one sub-component of every component fails;

iv) first stage having components working in series and each component having sub-components working in parallel, so the system will fail if and only if all the sub-components of at least one component fail.

Now consider a data set consisting of one hundred observations on breaking stress of carbon fibers given in Nichols and Padgett [23]. A histogram of the data is shown in Figure 9.

The breaking stress can be described by two stages. The first stage can be described by factors like the length and weight of fibers. Since these factors may vary from one fiber to another, it is reasonable to suppose that the number of components in the first stage is a random variable say N taking values in $\{1, 2, \ldots\}$. The second stage may be described by factors like the type of fiber (vegetable fibers, wood fibers, animal fibers, mineral fibers, etc) and density. Since the data set is on carbon fibers, it is reasonable to suppose that the number of sub-components is a fixed number say α. So, possible configurations of a fiber can be one of

i) N components working in parallel and each component has α sub-components working in parallel, so the fiber breaks if and only if all the sub-components of every component fail;

ii) N components working in series and each component has α sub-components working in series, so the fiber breaks if and only if at least one sub-component of at least one component fails;

iii) N components working in parallel and each component has α sub-components working in series, so the fiber breaks if and only if at least one sub-component of every component fails;

iv) N components working in series and each component has α sub-components working in parallel, so the fiber breaks if and only if all the sub-components of at least one component fail.

Standard models for N are the geometric, zero-truncated Poisson and zero-truncated negative binomial distributions:

$$\Pr(N = n) = (1 - p)p^{n-1} \tag{1}$$

for $0 < p < 1$ and $n = 1, 2, \ldots$;

$$\Pr(N = n) = \frac{\lambda^n}{(e^\lambda - 1)\, n!} \tag{2}$$

for $\lambda > 0$ and $n = 1, 2, \ldots$; and

$$\Pr(N = n) = \frac{1}{1 - (1-p)^r} \binom{n + r - 1}{n} (1-p)^r p^n \tag{3}$$

for $0 < p < 1$, $r > 0$ and $n = 1, 2, \ldots$. We shall stick to the first two distributions as the third one has two parameters and does not have a closed form cumulative distribution function (cdf).

We assume that the N components work independently and the sub-components work independently and independently of N. We also assume that the stress levels of the sub-components have the common probability density function (pdf) $f_0(\cdot)$, common cdf $F_0(\cdot)$ and common survival function $S_0(\cdot)$. Both f_0 and F_0 are assumed to be fully parametric with known shapes.

Let X denote the stress of the system.

Under (i) if N is geometric then X has the cdf

$$F_X(x) = \sum_{n=1}^{\infty} F_0^{\alpha n}(x)(1-p)p^{n-1} = \frac{(1-p)F_0^\alpha(x)}{1 - pF_0^\alpha(x)} \tag{4}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$, where $F_0^{\alpha n}(x)$ represents the cdf of the maximum of $n\alpha$ random variables. The corresponding pdf is:

$$f_X(x) = \frac{\alpha(1-p)F_0^{\alpha-1}(x) f_0(x)}{\left[1 - pF_0^\alpha(x)\right]^2} \tag{5}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$. The corresponding hazard rate function (hrf) is:

$$h_X(x) = \frac{\alpha(1-p)F_0^{\alpha-1}(x) f_0(x)}{\left[1 - F_0^\alpha(x)\right]\left[1 - pF_0^\alpha(x)\right]} \tag{6}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$. We shall refer to the distribution given by (4) and (5) as the *parallel-parallel-geometric-f_0 distribution.*

Under (ii) if N is geometric then the cdf of X is

$$F_X(x) = 1 - \frac{(1-p)S_0^\alpha(x)}{1 - pS_0^\alpha(x)} \tag{7}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$. We shall refer to the distribution given by (7) as the *series-series-geometric-f_0 distribution.*

Under (iii) if N is geometric then the cdf of X is

$$F_X(x) = \frac{(1-p)\left[1 - S_0^\alpha(x)\right]}{1 - p + pS_0^\alpha(x)} \tag{8}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$. We shall refer to the distribution given by (8) as the *parallel-series-geometric-f_0 distribution.*

Under (iv) if N is geometric then the cdf of X is

$$F_X(x) = 1 - \frac{(1-p)\left[1 - F_0^{\alpha}(x)\right]}{1 - p + pF_0^{\alpha}(x)} \tag{9}$$

for $x > 0$, $0 < p < 1$ and $\alpha > 0$. We shall refer to the distribution given by (9) as the *series-parallel-geometric-f_0 distribution.*

Under (i) if N is Poisson then the cdf of X is

$$F_X(x) = \frac{\exp\left[\lambda F_0^{\alpha}(x)\right]}{\exp(\lambda) - 1} \tag{10}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$. We shall refer to the distribution given by (10) as the *parallel-parallel-Poisson-f_0 distribution.*

Under (ii) if N is Poisson then the cdf of X is

$$F_X(x) = 1 - \frac{\exp\left[\lambda S_0^{\alpha}(x)\right]}{\exp(\lambda) - 1} \tag{11}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$. We shall refer to the distribution given by (11) as the *series-series-Poisson-f_0 distribution.*

Under (iii) if N is Poisson then the cdf of X is

$$F_X(x) = \frac{\exp\left[-\lambda S_0^{\alpha}(x)\right]}{1 - \exp(-\lambda)} \tag{12}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$. We shall refer to the distribution given by (12) as the *parallel-series-Poisson-f_0 distribution.*

Under (iv) if N is Poisson then the cdf of X is

$$F_X(x) = 1 - \frac{\exp\left[-\lambda F_0^{\alpha}(x)\right]}{1 - \exp(-\lambda)} \tag{13}$$

for $x > 0$, $\lambda > 0$ and $\alpha > 0$. We shall refer to the distribution given by (13) as the *series-parallel-Poisson-f_0 distribution.*

For the distributions given by (4) − (9), α and p are the shape parameters. For the distributions given by (10) − (13), α and λ are the shape parameters. For continuity and computational ease, we shall assume from now on that α can take any positive real value and not just integers (similar to approximating a discrete distribution by a continuous one). A non-integer value say $\alpha = 2.5$ can be interpreted as having two sub-components working to their full capacity and one sub-component working to half its capacity.

Although (4) appears to be a particular case of the distribution introduced by Marshall and Olkin [15], in all the generality one must note that quite the reverse is true. That is, the cdf, $F_0(x)/\left\{\beta + (1-\beta)F_0(x)\right\}$, introduced by equation (1.1) in Marshall and Olkin [15] is the particular case of (4) for $\alpha = 1$ and $p = 1 - 1/\beta$. Besides, the treatment provided by Marshall and Olkin [15] focused on the particular cases that F_0 is an exponential cdf or a Weibull cdf. The treatment provided here for the generalized class (4) is much more general; that is, most of the properties derived are for general F_0.

Since (4) is the simplest of the distributions given by (4) – (13), the mathematical properties and estimation issues discussed in this paper will be limited to the parallel-parallel-geometric-f_0 distribution. The mathematical properties and estimation issues for other distributions given by (4) – (13) can be derived similarly. The application part of this paper will consider all of the distributions given by (4) – (13).

The contents are organized as follows. In Section 2, we derive general mathematical properties of the parallel-parallel-geometric-f_0 distribution. These include shape properties of (5) and (6), quantile function and moments. The mathematical properties for f_0 being an exponential pdf are considered in detail in Section 3. The maximum likelihood estimation of the parameters of the parallel-parallel-geometric-f_0 distribution including the case of censoring is considered in Section 4. Section 5 assesses of the performance of the maximum likelihood estimates for small samples by simulation. Section 6 gives applications of the distributions given by (4) – (13) to the fiber data set.

Some of the mathematical properties in Section 2 involve single infinite sums, see Sections 2.1 and 2.4. Some other mathematical properties in Section 2 involve double infinite sums, see Section 2.5. Numerical computations not reported here showed that each of these infinite sums can be truncated at 20 to yield a relative error less than 10^{-25} for a wide range of parameter values and for a wide range of choices for f_0 and F_0. This shows that the mathematical properties can be computed for most practical uses with their infinite sums truncated at twenty. The computations were performed using Maple 2015. Maple took only a fraction of a second to compute the truncated versions. The computational times for the truncated versions were significantly smaller than those for the untruncated versions.

2. MATHEMATICAL PROPERTIES

2.1. Expansions for pdf and cdf

Some of the mathematical properties of (4) and (5) cannot be expressed in closed form. In these cases, it is useful to have expansions for the pdf and the cdf. As just mentioned, the truncated versions of these expansions can be of practical use.

Some useful expansions for (4) and (5) can be derived using the concept of exponentiated distributions. A random variable is said to have the exponentiated-F_0 distribution with parameter $a > 0$, if its pdf and cdf are

$$\mathcal{G}_a(x) = aF_0^{a-1}(x)f_0(x) \tag{14}$$

and

$$\mathcal{H}_a(x) = F_0^a(x), \tag{15}$$

respectively. The properties of exponentiated distributions have been studied by many authors in recent years, see Mudholkar and Srivastava [16] for exponentiated Weibull, Gupta et al. [5] for exponentiated Pareto, Gupta and Kundu [6] for exponentiated exponential, Nadarajah [19] for exponentiated Gumbel, Kakde and Shirke [9] for exponentiated lognormal, and Nadarajah and Gupta [21] for exponentiated gamma distributions.

We now provide four expansions for (4) and (5), each in terms of (14) and (15). Expanding the denominator in (4) as a binomial series, we can write (4) and (5) as

$$F_X(x) = (1-p)\sum_{k=0}^{\infty} p^k \mathcal{H}_{\alpha(k+1)}(x) \tag{16}$$

and

$$f_X(x) = (1-p)\sum_{k=0}^{\infty} p^k \mathcal{G}_{\alpha(k+1)}(x), \tag{17}$$

respectively. Similarly, expanding the denominator in (5) as a binomial series, we can write (4) and (5) as

$$F_X(x) = (1-p)\sum_{k=0}^{\infty} \frac{(-p)^k}{k+1}\binom{-2}{k}\mathcal{H}_{\alpha(k+1)}(x) \tag{18}$$

and

$$f_X(x) = (1-p)\sum_{k=0}^{\infty} \frac{(-p)^k}{k+1}\binom{-2}{k}\mathcal{G}_{\alpha(k+1)}(x), \tag{19}$$

respectively.

So, several properties of (4) and (5) can be obtained by knowing those of exponentiated distributions, see, for example, Mudholkar et al. [17], Gupta and Kundu [6], Nadarajah and Kotz [22], among others.

2.2. Asymptotes and shapes

Here, we study asymptotes and shape properties of (4), (5) and (6). Shape properties are important because they allow the practitioner to see if the distribution can be fitted to a given data set (this can be seen by comparing the shape of the histogram of the data with possible shapes of the pdf). Shape properties are also useful to see if the distribution can model increasing failure rates, decreasing failure rates or bathtub shaped failure rates.

The study of asymptotes and shapes is useful to determine if a data set can be modeled by (4) and (5). The asymptotes of (4), (5) and (6) as $x \to 0, \infty$ are given by

$$f_X(x) \sim \alpha(1-p)F_0^{\alpha-1}(x)f_0(x) \tag{20}$$

as $x \to 0$,

$$f_X(x) \sim \frac{\alpha f_0(x)}{1-p} \tag{21}$$

as $x \to \infty$,

$$F_X(x) \sim (1-p)F_0^{\alpha}(x) \tag{22}$$

as $x \to 0$,

$$1 - F_X(x) \sim \frac{1 - F_0^{\alpha}(x)}{1-p} \tag{23}$$

as $x \to \infty$,

$$h_X(x) \sim \alpha(1-p)F_0^{\alpha-1}(x) f_0(x) \tag{24}$$

as $x \to 0$, and

$$h_X(x) \sim \frac{\alpha f_0(x)}{1 - F_0^{\alpha}(x)} \tag{25}$$

as $x \to \infty$. So, $f_X(\cdot)$ behaves like $f_0(\cdot)$ for very large x. Also, $h_X(\cdot)$ behaves like the hrf corresponding to $f_0(\cdot)$ for very large x.

The shapes of (5) and (6) can be described analytically. The critical points of the pdf are the roots of the equation:

$$(\alpha - 1)\frac{f_0(x)}{F_0(x)} + \frac{f_0^{'}(x)}{f_0(x)} + \frac{2\alpha p F_0^{\alpha-1}(x) f_0(x)}{1 - pF_0^{\alpha}(x)} = 0. \tag{26}$$

There may be more than one root to (26). The critical points of the hrf are the roots of the equation:

$$(\alpha - 1)\frac{f_0(x)}{F_0(x)} + \frac{f_0^{'}(x)}{f_0(x)} + \frac{\alpha F_0^{\alpha-1}(x) f_0(x)}{1 - F_0^{\alpha}(x)} + \frac{\alpha p F_0^{\alpha-1}(x) f_0(x)}{1 - pF_0^{\alpha}(x)} = 0. \tag{27}$$

There may be again more than one root to (27).

Figures 1 to 4 illustrate possible shapes of (5) when f_0 is an exponential pdf with unit scale parameter, a gamma pdf with shape parameter $a = 2$ and unit scale parameter, a Weibull pdf with shape parameter $a = 2$ and unit scale parameter and an exponentiated exponential pdf with shape parameter $a = 2$ and unit scale parameter. It is evident that monotonically decreasing, unimodal and bimodal shapes are possible. It is difficult to determine analytically the parameter regions corresponding to these shapes. However, extensive graphical analysis not reported here showed that: monotonically decreasing shapes correspond to small α, small a and small p; unimodal and bimodal shapes correspond to large α or large a or large p.

Extensive graphical analysis examining the corresponding shapes of (6) showed: monotonic increasing, monotonic decreasing and bathtub shaped hazard rates correspond to $\alpha < 1$; monotonic increasing hazard rates correspond to $\alpha \geq 1$. Constant hazard rates do not appear to be possible.

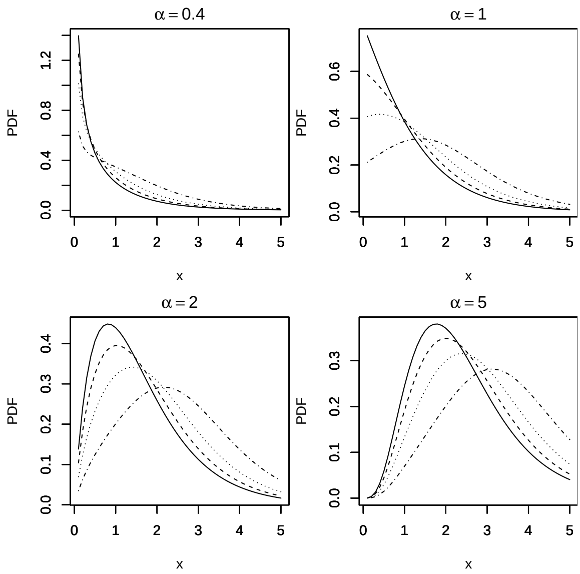

Fig. 1. Plots of (5) for f_0 an exponential pdf with unit scale parameter, $\alpha = 0.4, 1, 2, 5$, $p = 0.2$ (solid curve), $p = 0.4$ (curve of dashes), $p = 0.6$ (curve of dots) and $p = 0.8$ (curve of dots and dashes).

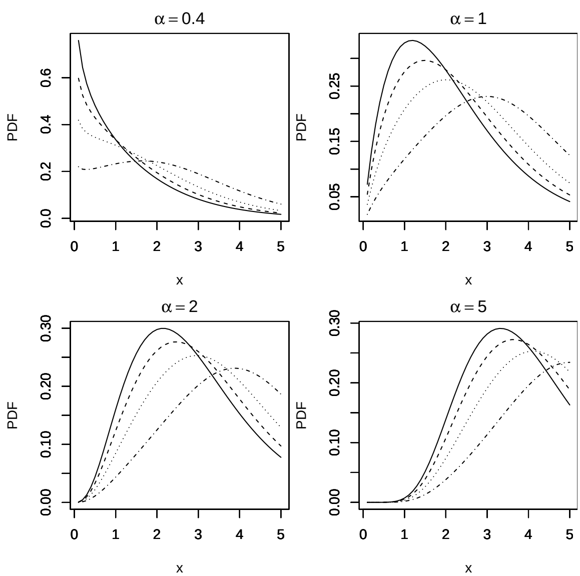

Fig. 2. Plots of (5) for f_0 a gamma pdf with unit scale parameter and shape parameter $a = 2$, $\alpha = 0.4, 1, 2, 5$, $p = 0.2$ (solid curve), $p = 0.4$ (curve of dashes), $p = 0.6$ (curve of dots) and $p = 0.8$ (curve of dots and dashes).

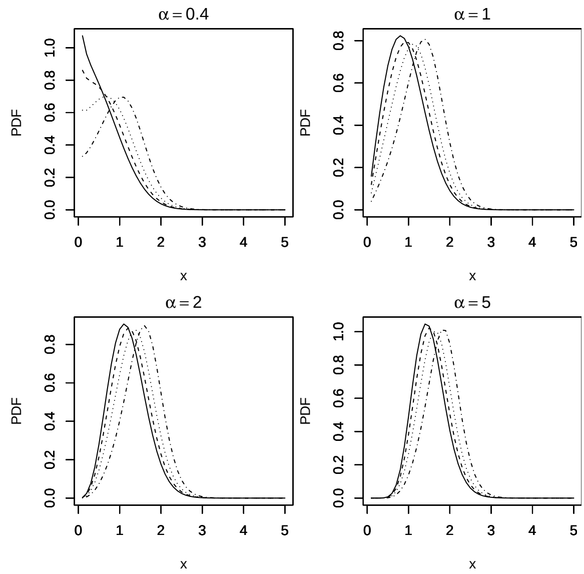

Fig. 3. Plots of (5) for f_0 a Weibull pdf with unit scale parameter and shape parameter $a = 2$, $\alpha = 0.4, 1, 2, 5$, $p = 0.2$ (solid curve), $p = 0.4$ (curve of dashes), $p = 0.6$ (curve of dots) and $p = 0.8$ (curve of dots and dashes).

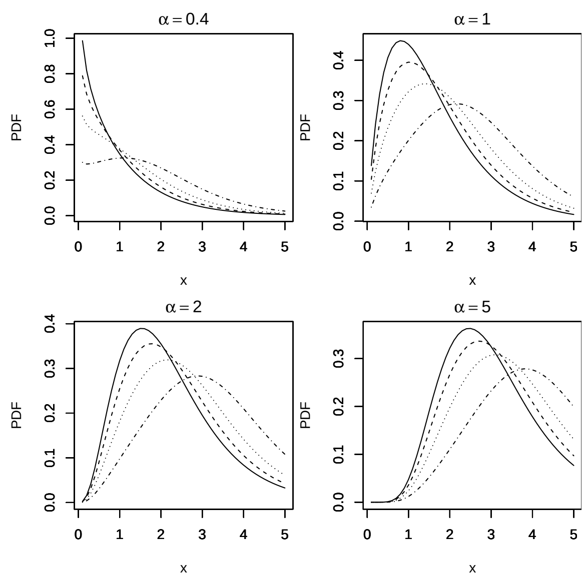

Fig. 4. Plots of (5) for f_0 an exponentiated exponential pdf with unit scale parameter and shape parameter $a = 2$, $\alpha = 0.4, 1, 2, 5$, $p = 0.2$ (solid curve), $p = 0.4$ (curve of dashes), $p = 0.6$ (curve of dots) and $p = 0.8$ (curve of dots and dashes).

2.3. Quantile function

Quantile functions are fundamental for random number generation. They can also be used for estimation (for example, percentile based estimation methods).

Let X denote a random variable with cdf (4). Inverting $F_X(x) = u$, we obtain

$$F_X^{-1}(u) = F_0^{-1}\left(\left(\frac{u}{1-p+pu}\right)^{1/\alpha}\right) \tag{28}$$

for $0 < u < 1$, where $F_0^{-1}(\cdot)$ denotes the inverse function of $F_0(\cdot)$.

2.4. Moments

Moment properties are fundamental for any distribution. For instance, the first four moments can be used to describe any data fairly well. Moments are also useful for estimation (for example, the method of moments).

Let X denote a random variable with cdf (4). Let Z_a denote a random variable with pdf $\mathcal{G}_a(\cdot)$ and cdf $\mathcal{H}_a(\cdot)$. Then, using the expansions, (16) and (17), we have

$$E\left(X^n\right) = (1-p)\sum_{k=0}^{\infty} p^k E\left[Z_{\alpha(k+1)}^n\right]. \tag{29}$$

Using the expansions, (18) and (19), we have

$$E\left(X^n\right) = (1-p)\sum_{k=0}^{\infty} \frac{(-p)^k}{k+1}\binom{-2}{k} E\left[Z_{\alpha(k+1)}^n\right]. \tag{30}$$

Details not reported here showed (29) and (30) yield roughly the same amount of computational accuracy and computational time.

The expressions given by (29) and (30) can be used to compute the mean, variance, skewness and kurtosis of X. Extensive graphical analysis of the variation of these four measures versus a wide range of values of α, β, p and a showed the following:

- The mean is always an increasing function of α and an increasing function of p.
- The skewness is always a decreasing function of α and a decreasing function of p.
- Whenever f_0 is an exponential pdf or an exponentiated exponential pdf or a gamma pdf with small a or a Weibull pdf with small a, the variance is always an increasing function of α.
- Whenever f_0 is an exponential pdf or an exponentiated exponential pdf or a gamma pdf or a Weibull pdf with small a, the kurtosis is always a decreasing function of p.
- Whenever f_0 is an exponential pdf or an exponentiated exponential pdf or a gamma pdf with small a or a Weibull pdf with small a, the variance is always an increasing function of p.

- Whenever f_0 is an exponential pdf or an exponentiated exponential pdf or a gamma pdf or a Weibull pdf with small a, the kurtosis is always a decreasing function of p.
- Whenever f_0 is a gamma pdf with large a or a Weibull pdf with large a, the variance always initially increases before decreasing with respect to α.
- Whenever f_0 is a gamma pdf with large a or a Weibull pdf with large a, the kurtosis always initially decreases before increasing with respect to α.
- Whenever f_0 is a Weibull pdf with large a, the variance always initially increases before decreasing with respect to p.
- Whenever f_0 is a Weibull pdf with large a, the kurtosis always initially decreases before increasing with respect to p.

2.5. Stress-strength parameter

Suppose that the random variable X_1 is the strength of a component which is subjected to a random stress X_2. The component fails whenever $X_1 < X_2$ and there is no hazard when $X_1 > X_2$. In the context of reliability, the stress-strength parameter $R = \Pr\left(X_1 > X_2\right)$ is a measure of component reliability. Its estimation when X_1 and X_2 are independent random variables and follow a specified distribution has been discussed widely in literature.

Suppose X_1 and X_2 are independent random variables distributed according to (5) with parameters (p_1, α_1) and (p_2, α_2), respectively. By using the representations, (18) and (19), we can write $R = \Pr\left(X_2 < X_1\right)$ as

$$R = \left(1 - p_1\right)\left(1 - p_2\right) \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \frac{\left(-p_1\right)^k \left(-p_2\right)^l}{(k+1)(l+1)} \binom{-2}{k} \binom{-2}{l} R_{kl}, \qquad (31)$$

where $R_{kl} = \Pr\left(Z_{\alpha_1(k+1)} < Z_{\alpha_2(l+1)}\right)$ is the stress-strength parameter between two independent exponentiated random variables. Hence, the stress-strength parameter, $R = \Pr\left(X_2 < X_1\right)$, is a linear combination of those for exponentiated random variables. In the particular case $\alpha_1 = \alpha_2$ and $p_1 = p_2$, (31) reduces to $R = 1/2$.

The result (31) also holds if X_1 and X_2 are independent random variables specified by the pdfs

$$f_{X_i}(x) = \frac{\alpha_i \left(1 - p_i\right) F_i^{\alpha_i - 1}(x) f_i(x)}{\left[1 - p_i F_1^{\alpha_i}(x)\right]^2}$$

for $x > 0$ and $i = 1, 2$, where F_i are valid cdfs and $f_i(x) = \mathrm{d}F_i(x)/\mathrm{d}x$. In this case, $R_{kl} = \Pr\left(U_{\alpha_1(k+1)} < V_{\alpha_2(l+1)}\right)$, where U_a and V_a are independent random variables specified by the pdfs $aF_1^{a-1}(x) f_1(x)$ and $aF_2^{a-1}(x) f_2(x)$, respectively.

3. THE CASE OF F_0 EXPONENTIAL

In Section 1, $F_0(\cdot)$ is interpreted as the cdf of the stress of independent and identical sub-components. The oldest and the most popular model for lifetimes is the exponential distribution. In this section, we show how the results in Sections 1 and 2 simplify when $F_0(\cdot)$ is an exponential cdf with scale parameter β.

First, the cdf, pdf and hrf given by (4), (5) and (6) simplify to

$$F_X(x) = \frac{(1-p)\left[1-\exp(-\beta x)\right]^{\alpha}}{1-p\left[1-\exp(-\beta x)\right]^{\alpha}}, \tag{32}$$

$$f_X(x) = \frac{\alpha\beta(1-p)\exp(-\beta x)\left[1-\exp(-\beta x)\right]^{\alpha-1}}{\left\{1-p\left[1-\exp(-\beta x)\right]^{\alpha}\right\}^{2}}, \tag{33}$$

and

$$h_X(x) = \frac{\alpha\beta(1-p)\exp(-\beta x)\left[1-\exp(-\beta x)\right]^{\alpha-1}}{\left\{1-p\left[1-\exp(-\beta x)\right]^{\alpha}\right\}\left\{1-\left[1-\exp(-\beta x)\right]^{\alpha}\right\}}, \tag{34}$$

respectively. It follows from (16) to (19) that (32) and (33) can be expressed as mixtures of the exponentiated exponential distribution due to Gupta and Kundu [6].

The asymptotes of (32), (33) and (34) as $x \to 0, \infty$ are given by

$$f_X(x) \sim \alpha\beta^{\alpha}(1-p)x^{\alpha-1}$$

as $x \to 0$,

$$f_X(x) \sim \alpha\beta\exp(-\beta x)$$

as $x \to \infty$,

$$F_X(x) \sim (1-p)\beta^{\alpha}x^{\alpha}$$

as $x \to 0$,

$$1-F_X(x) \sim \alpha\exp(-\beta x)/(1-p)$$

as $x \to \infty$,

$$h_X(x) \sim \beta^{\alpha}(1-p)x^{\alpha-1}$$

as $x \to 0$, and

$$h_X(x) \to \beta(1-p)$$

as $x \to \infty$. So, the lower tail of the pdf behaves polynomially while its upper tail behaves exponentially. The lower tail of the hrf also behaves polynomially. The hrf approaches a fixed constant, the ultimate hazard rate, as $x \to \infty$.

The quantile function given by (28) simplifies to

$$F_X^{-1}(u) = -\frac{1}{\beta} \log \left[1 - \left(\frac{u}{1-p+pu} \right)^{1/\alpha} \right]$$

for $0 < u < 1$.

As mentioned, the parallel-parallel-geometric-exponential distribution is a mixture of exponentiated exponential distributions. A comprehensive review of known properties of the exponentiated exponential distribution, including known expressions for moments, stress-strength parameter and others, is given in Nadarajah [20]. These results can be used to yield expressions for moments, stress-strength parameter, etc of the parallel-parallel-geometric-exponential distribution.

4. ESTIMATION

4.1. Maximum likelihood estimation

Here, we consider estimation of the unknown parameters of (4) and (5) by the method of maximum likelihood. Let x_1, x_2, ..., x_n be a random sample from (5). Let $\mathbf{\Theta}$ be a q-dimensional parameter vector specifying $F_0(\cdot)$. Then the log-likelihood function, $\log L = \log L\left(\alpha, p, \mathbf{\Theta}\right)$, is

$$\begin{aligned} \log L\left(p, \alpha, \mathbf{\Theta}\right) &= n \log\left[\alpha(1-p)\right] + (\alpha - 1) \sum_{i=1}^{n} \log F_0\left(x_i\right) + \sum_{i=1}^{n} \log f_0\left(x_i\right) \\ &\quad -2 \sum_{i=1}^{n} \log\left[1 - pF_0^{\alpha}\left(x_i\right)\right]. \end{aligned} \tag{35}$$

The first derivatives of the log-likelihood function with respect to the parameters p, α and $\mathbf{\Theta}$ are:

$$\frac{\partial \log L}{\partial p} = -\frac{n}{1-p} + 2 \sum_{i=1}^{n} \frac{F_0^{\alpha}\left(x_i\right)}{1 - pF_0^{\alpha}\left(x_i\right)}, \tag{36}$$

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} + \sum_{i=1}^{n} \log F_0\left(x_i\right) + 2p \sum_{i=1}^{n} \frac{F_0^{\alpha}\left(x_i\right) \log F_0\left(x_i\right)}{1 - pF_0^{\alpha}\left(x_i\right)}, \tag{37}$$

$$\begin{aligned} \frac{\partial \log L}{\partial \mathbf{\Theta}} &= (\alpha - 1) \sum_{i=1}^{n} \frac{\partial F_0\left(x_i\right) / \partial \mathbf{\Theta}}{F_0\left(x_i\right)} + \sum_{i=1}^{n} \frac{\partial f_0\left(x_i\right) / \partial \mathbf{\Theta}}{f_0\left(x_i\right)} \\ &\quad +2p \sum_{i=1}^{n} \frac{F_0^{\alpha-1}\left(x_i\right) \partial F_0\left(x_i\right) / \partial \mathbf{\Theta}}{1 - pF_0^{\alpha}\left(x_i\right)}. \end{aligned} \tag{38}$$

The maximum likelihood estimates of $(p, \alpha, \mathbf{\Theta})$, say $\left(\widehat{p}, \widehat{\alpha}, \widehat{\mathbf{\Theta}}\right)$, are the simultaneous solutions of the equations $\partial \log L / \partial p = 0$, $\partial \log L / \partial \alpha = 0$ and $\partial \log L / \partial \mathbf{\Theta} = \mathbf{0}$.

Alternatively, the maximum likelihood estimates can be obtained by numerical maximization of (35). There are well established routines for numerical maximization like `nlm` or optim in the R statistical package (R Development Core Team [25]). Our numerical calculations showed that the surface of (35) was smooth for given smooth functions $f_0(\cdot)$ and $F_0(\cdot)$. The routines were able to locate the maximum of the likelihood surface for a wide range of smooth functions and for a wide range of starting values. However, to ease computations it is useful to have reasonable starting values. These can be obtained, for example, by the method of moments. For $r = 1, \ldots, q+2$, let $m_r = n^{-1} \sum_{i=1}^{n} x_i^r$ denote the first $q+2$ sample moments. Equating these moments with the theoretical versions given in Section 2.4, we have $m_r = E\left(X^r\right)$ for $r = 1, \ldots, q+2$. These equations can be solved simultaneously to obtain the moments estimates.

For interval estimation of $(p, \alpha, \mathbf{\Theta})$ and tests of hypotheses, one requires the Fisher information matrix. We can express the Fisher information matrix of $\left(\widehat{p}, \widehat{\alpha}, \widehat{\mathbf{\Theta}}\right)$ as

$$E\left(\mathbf{J}\right) = \left(\begin{array}{ccc} E\left(J_{11}\right) & E\left(J_{12}\right) & E\left(\mathbf{J}_{13}\right) \\ E\left(J_{12}\right) & E\left(J_{22}\right) & E\left(\mathbf{J}_{23}\right) \\ E\left(\mathbf{J}_{13}\right) & E\left(\mathbf{J}_{23}\right) & E\left(\mathbf{J}_{33}\right) \end{array} \right). \tag{39}$$

For large n, the distribution of $\sqrt{n}\left(\widehat{p} - p, \widehat{\alpha} - \alpha, \widehat{\mathbf{\Theta}} - \mathbf{\Theta}\right)$ approximates to a $(q+2)$-variate normal distribution with zero means and variance-covariance matrix $\mathbf{J}^{-1}$ evaluated at $\left(\widehat{p}, \widehat{\alpha}, \widehat{\mathbf{\Theta}}\right)$. The properties of $\left(\widehat{p}, \widehat{\alpha}, \widehat{\mathbf{\Theta}}\right)$ can be derived based on this normal approximation.

For the asymptotic normality to hold, certain regularity conditions must be satisfied. The general forms of these conditions are given in Ferguson [4] and pages 461–463 in Lehmann and Casella [13]. For the log likelihood function given by (35), these conditions reduce to

a) the parameter space for $\mathbf{\Theta}$ is compact;

b) f_0 and F_0 are continuous with respect to $\mathbf{\Theta}$;

c) $E\left[\log F_0(X)\right]$, $E\left[-\log f_0(X)\right]$ and $E\left[\log\left[1 - pF_0^{\alpha}(X)\right]\right]$ exist;

d) as $n \to \infty$,

$$\sup_{p, \alpha, \mathbf{\Theta}} \left| \frac{1}{n} \log L\left(p, \alpha, \mathbf{\Theta}\right) - E\left[\log f(X)\right] \right| < \delta$$

for some $\delta > 0$, where $\log L$ is given by (35);

e) the true maximum likelihood estimator of $(p, \alpha, \mathbf{\Theta})$ is in the interior of the parameter space for $(p, \alpha, \mathbf{\Theta})$;

f) f_0 and F_0 are twice differentiable in the neighborhood of the true maximum likelihood estimator of $(p, \alpha, \mathbf{\Theta})$;

g) $E\left(J_{11}\right)$, $E\left(J_{12}\right)$, $E\left(\mathbf{J}_{13}\right)$, $E\left(J_{12}\right)$, $E\left(J_{22}\right)$, $E\left(\mathbf{J}_{23}\right)$, $E\left(\mathbf{J}_{13}\right)$, $E\left(\mathbf{J}_{23}\right)$ and $E\left(\mathbf{J}_{33}\right)$ exist when $(p, \alpha, \mathbf{\Theta})$ is the true maximum likelihood estimator;

h) $E(\mathbf{J})$ is non-singular when $(p, \alpha, \mathbf{\Theta})$ is the true maximum likelihood estimator;

i) that

$$\int \sup_{p,\alpha,\boldsymbol{\Theta}} \left\| \begin{pmatrix} \partial f_X(x)/\partial p \\ \partial f_X(x)/\partial \alpha \\ \partial f_X(x)/\partial \boldsymbol{\Theta} \end{pmatrix} \right\| \mathrm{d}x < \infty,$$

where $f_X(x)$ is given by (5);

j) that

$$\int \sup_{p,\alpha,\boldsymbol{\Theta}} \left\| \begin{pmatrix} \partial^2 f_X(x)/\partial p^2 & \partial^2 f_X(x)/\partial p \partial \alpha & \partial^2 f_X(x)/\partial p \partial \boldsymbol{\Theta} \\ \partial^2 f_X(x)/\partial \alpha \partial p & \partial^2 f_X(x)/\partial \alpha^2 & \partial^2 f_X(x)/\partial \alpha \partial \boldsymbol{\Theta} \\ \partial^2 f_X(x)/\partial \boldsymbol{\Theta} \partial p & \partial^2 f_X(x)/\partial \boldsymbol{\Theta} \partial \alpha & \partial^2 f_X(x)/\partial \boldsymbol{\Theta}^2 \end{pmatrix} \right\| \mathrm{d}x < \infty,$$

where $f_X(x)$ is given by (5);

l) that

$$E \left\{ \sup_{p,\alpha,\boldsymbol{\Theta}} \left\| \begin{pmatrix} \partial^2 \log L/\partial p^2 & \partial^2 \log L/\partial p \partial \alpha & \partial^2 \log L/\partial p \partial \boldsymbol{\Theta} \\ \partial^2 \log L/\partial \alpha \partial p & \partial^2 \log L/\partial \alpha^2 & \partial^2 \log L/\partial \alpha \partial \boldsymbol{\Theta} \\ \partial^2 \log L/\partial \boldsymbol{\Theta} \partial p & \partial^2 \log L/\partial \boldsymbol{\Theta} \partial \alpha & \partial^2 \log L/\partial \boldsymbol{\Theta}^2 \end{pmatrix} \right\| \right\} < \infty,$$

where $\log L$ is given by (35).

4.2. Censored maximum likelihood estimation

Often with lifetime data, one encounters censoring. There are different forms of censoring: type I censoring, type II censoring, etc. Here, we consider the general case of multicensored data: there are n subjects of which

- n_0 are known to have failed at times $t_1, \ldots, t_{n_0}$;
- n_1 are known to have failed in the intervals $[s_{i-1}, s_i]$, $i = 1, \ldots, n_1$;
- n_2 are known to have lived past r_i, $i = 1, \ldots, n_2$ but not observed any longer,

where s_i and r_i are assumed to be deterministic. Note that $n = n_0 + n_1 + n_2$. Note too that type I censoring and type II censoring are contained as particular cases of multicensoring.

In the case of multicensoring, the log-likelihood function is:

$$\log L(p, \alpha, \boldsymbol{\Theta}) = \sum_{i=1}^{n_0} \log f_X(t_i) + \sum_{i=1}^{n_1} \log \left[F_X(s_i) - F_X(s_{i-1}) \right] + \sum_{i=1}^{n_2} \log \left[1 - F_X(r_i) \right], \quad (40)$$

where $F_X(\cdot)$ and $f_X(\cdot)$ are given by (4) and (5), respectively. The first derivatives of the log-likelihood function with respect to the parameters p, α and $\boldsymbol{\Theta}$ are:

$$\frac{\partial \log L}{\partial p} = \sum_{i=1}^{n_0} \frac{\partial f_X(t_i)/\partial p}{f_X(t_i)} + \sum_{i=1}^{n_1} \frac{\partial F_X(s_i)/\partial p - \partial F_X(s_{i-1})/\partial p}{F_X(s_i) - F_X(s_{i-1})} - \sum_{i=1}^{n_2} \frac{\partial F_X(r_i)/\partial p}{1 - F_X(r_i)}, \quad (41)$$

$$\frac{\partial \log L}{\partial \alpha} = \sum_{i=1}^{n_0} \frac{\partial f_X(t_i)/\partial \alpha}{f_X(t_i)} + \sum_{i=1}^{n_1} \frac{\partial F_X(s_i)/\partial \alpha - \partial F_X(s_{i-1})/\partial \alpha}{F_X(s_i) - F_X(s_{i-1})} - \sum_{i=1}^{n_2} \frac{\partial F_X(r_i)/\partial \alpha}{1 - F_X(r_i)}, \quad (42)$$

and

$$\frac{\partial \log L}{\partial \boldsymbol{\Theta}} = \sum_{i=1}^{n_0} \frac{\partial f_X(t_i)/\partial \boldsymbol{\Theta}}{f_X(t_i)} + \sum_{i=1}^{n_1} \frac{\partial F_X(s_i)/\partial \boldsymbol{\Theta} - \partial F_X(s_{i-1})/\partial \boldsymbol{\Theta}}{F_X(s_i) - F_X(s_{i-1})} - \sum_{i=1}^{n_2} \frac{\partial F_X(r_i)/\partial \boldsymbol{\Theta}}{1 - F_X(r_i)}. \quad (43)$$

The first term in (40) is the same as (35) with (x_i, n) replaced by (t_i, n_0). Also the first terms in (41)−(43) are the same as (36)−(38) with (x_i, n) replaced by (t_i, n_0). So, it is sufficient to find explicit expressions for the partial derivatives in (41)−(43). They are

$$\begin{aligned}
\frac{F_X(x)}{\partial p} &= -\frac{F_0^{\alpha}(x)}{1 - pF_0^{\alpha}(x)} + \frac{(1-p)F_0^{2\alpha}(x)}{\left[1 - pF_0^{\alpha}(x)\right]^2}, \\
\frac{F_X(x)}{\partial \alpha} &= \frac{(1-p)F_0^{\alpha}(x)\log F_0(x)}{1 - pF_0^{\alpha}(x)} + \frac{\alpha p(1-p)F_0^{2\alpha}(x)\log F_0(x)}{\left[1 - pF_0^{\alpha}(x)\right]^2}, \\
\frac{F_X(x)}{\partial \boldsymbol{\Theta}} &= \frac{\alpha(1-p)F_0^{\alpha-1}(x)\partial F_0(x)/\partial \boldsymbol{\Theta}}{1 - pF_0^{\alpha}(x)} + \frac{\alpha p(1-p)F_0^{2\alpha-1}(x)\partial F_0(x)/\partial \boldsymbol{\Theta}}{\left[1 - pF_0^{\alpha}(x)\right]^2},
\end{aligned}$$

and

$$\begin{aligned}
\frac{f_X(x)}{\partial p} &= -\frac{\alpha F_0^{\alpha-1}(x)f_0(x)}{\left[1 - pF_0^{\alpha}(x)\right]^2} + \frac{2\alpha(1-p)F_0^{2\alpha-1}(x)f_0(x)}{\left[1 - pF_0^{\alpha}(x)\right]^3}, \\
\frac{f_X(x)}{\partial \alpha} &= \frac{(1-p)F_0^{\alpha-1}(x)f_0(x)\left[1 + \alpha\log F_0(x)\right]}{\left[1 - pF_0^{\alpha}(x)\right]^2} + \frac{2\alpha p(1-p)F_0^{2\alpha-1}(x)f_0(x)\log F_0(x)}{\left[1 - pF_0^{\alpha}(x)\right]^3}, \\
\frac{f_X(x)}{\partial \boldsymbol{\Theta}} &= \frac{\alpha(1-p)F_0^{\alpha-2}(x)\left[(\alpha-1)f_0(x)\partial F_0(x)/\partial \boldsymbol{\Theta} + F_0(x)\partial f_0(x)/\partial \boldsymbol{\Theta}\right]}{\left[1 - pF_0^{\alpha}(x)\right]^2} \\
&\quad + \frac{2\alpha^2 p(1-p)F_0^{2\alpha-2}(x)f_0(x)\partial F_0(x)/\partial \boldsymbol{\Theta}}{\left[1 - pF_0^{\alpha}(x)\right]^3}.
\end{aligned}$$

The maximum likelihood estimates of $(p, \alpha, \boldsymbol{\Theta})$, say $(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}})$, are the simultaneous solutions of the equations $\partial \log L/\partial p = 0$, $\partial \log L/\partial \alpha = 0$ and $\partial \log L/\partial \boldsymbol{\Theta} = \mathbf{0}$.

For large n, the distribution of $\sqrt{n}(\widehat{p} - p, \widehat{\alpha} - \alpha, \widehat{\boldsymbol{\Theta}} - \boldsymbol{\Theta})$ can be approximated by a $(q+2)$-variate normal distribution with zero means and variance-covariance matrix say $\mathbf{K}^{-1}$ evaluated at $(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}})$. The form of $\mathbf{K}$ is too complicated to write here. The regularity conditions for asymptotic normality are the same as those listed in Section 4.1 except for conditions d), g), h) and l). For conditions d) and l), (35) should be replaced by (40). For conditions g) and h), $\mathbf{J}$ should be replaced by $\mathbf{K}$.

5. SIMULATION STUDY

Here, we assess the performance of the maximum likelihood estimates given by (36)–(38) with respect to sample size n when f_0 is an exponential pdf with scale parameter β, a gamma pdf with shape parameter a and scale parameter β, a Weibull pdf with shape parameter a and scale parameter β or an exponentiated exponential pdf with shape parameter a and scale parameter β, i. e., the random sample is from the parallel-parallel-geometric-exponential, parallel-parallel-geometric-gamma, parallel-parallel-geometric--Weibull or the parallel-parallel-geometric-exponentiated exponential distribution. The assessment is based on a simulation study:

1. generate ten thousand samples of size n from (33). The inversion method was used to generate samples, i. e., variates were generated using

$$X = F_0^{-1}\left(\left(\frac{U}{1-p+pU}\right)^{1/\alpha}\right),$$

where $U \sim U(0,1)$ is a uniform variate on the unit interval;

2. compute the maximum likelihood estimates for the ten thousand samples, say $\left(\widehat{p}_i, \widehat{\alpha}_i, \widehat{a}_i, \widehat{\beta}_i\right)$ for $i = 1, 2, \ldots, 10000$;

3. compute the relative biases and coefficients of variation given by

$$\mathrm{RB}_f(n) = \frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{f}_i - f\right)/f$$

and

$$\mathrm{CV}_f(n) = \frac{1}{f}\sqrt{\frac{1}{10000}\sum_{i=1}^{10000}\left(\widehat{f}_i - \frac{1}{10000}\sum_{j=1}^{10000}\widehat{f}_j\right)^2}$$

for $f = p, \alpha, a, \beta$.

We repeated these steps for $n = 10, 11, \ldots, 100$ with $p = 0.5$, $\alpha = 1$, $a = 2$ and $\beta = 1$, so computing $\mathrm{RB}_p(n)$, $\mathrm{RB}_\alpha(n)$, $\mathrm{RB}_a(n)$, $\mathrm{RB}_\beta(n)$ and $\mathrm{CV}_p(n)$, $\mathrm{CV}_\alpha(n)$, $\mathrm{CV}_a(n)$, $\mathrm{CV}_\beta(n)$ for $n = 10, 11, \ldots, 100$.

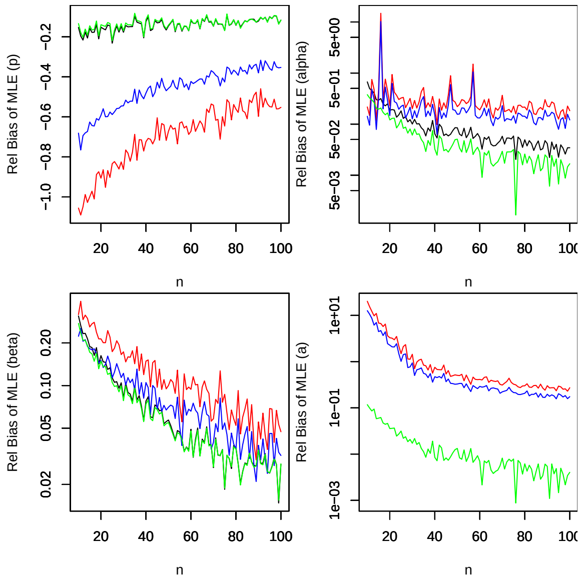

Fig. 5. $\mathrm{RB}_p(n)$ (top left), $\mathrm{RB}_\alpha(n)$ (top right), $\mathrm{RB}_\beta(n)$ (bottom left) and $\mathrm{RB}_a(n)$ (bottom right) versus $n = 10, 11, \ldots, 100$. black for f_0 an exponential pdf, red for f_0 a gamma pdf, blue for f_0 a Weibull pdf and green for f_0 an exponentiated exponential pdf.

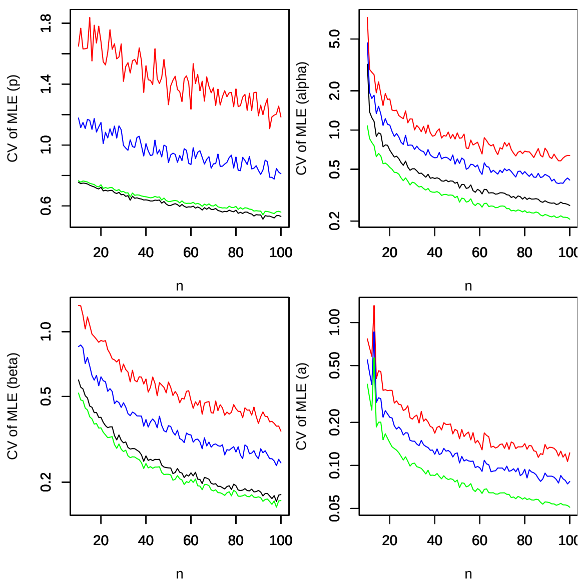

Fig. 6. $\mathrm{CV}_p(n)$ (top left), $\mathrm{CV}_\alpha(n)$ (top right), $\mathrm{CV}_\beta(n)$ (bottom left) and $\mathrm{CV}_a(n)$ (bottom right) versus $n = 10, 11, \ldots, 100$. black for f_0 an exponential pdf, red for f_0 a gamma pdf, blue for f_0 a Weibull pdf and green for f_0 an exponentiated exponential pdf.

Figures 5 and 6 show how the relative biases and the coefficients of variation vary with respect to n for the four distributions. The broken lines in Figure 5 correspond to the relative biases being zero. The y axes are plotted in log scale for the gamma, Weibull and exponentiated exponential distributions. The following observations can be made:

- the relative biases for each parameter and for each distribution either decrease or increase to zero as $n \to \infty$, and the coefficients of variation for each parameter and for each distribution decrease to zero as $n \to \infty$, both being direct consequences of maximum likelihood estimation;
- the relative biases for α, β and a are generally positive;
- the relative biases for p are generally negative;
- the relative biases appear smallest for f_0 the exponential or the exponentiated exponential pdf;
- the relative biases appear largest for f_0 the gamma pdf;
- the relative biases appear second largest for f_0 the Weibull pdf;
- the coefficients of variation appear largest for the parameter, α;
- the coefficients of variation appear smallest for the parameter, a;
- the coefficients of variation appear smallest for f_0 the exponential or the exponentiated exponential pdf;
- the coefficients of variation appear largest for f_0 the gamma pdf;
- the coefficients of variation appear second largest for f_0 the Weibull pdf.

We have presented results for only one choice for (p, α, a, β), namely that $(p, \alpha, a, \beta) = (0.5, 1, 2, 1)$. The simulations were re-run for a wide range of other values for (p, α, a, β). The following observations held in general: the relative biases for each parameter and for each distribution either decreased or increased to zero as $n \to \infty$; the coefficients of variation for each parameter and for each distribution always decreased to zero as $n \to \infty$; the relative biases always appeared smallest for f_0 the exponential or the exponentiated exponential pdf; the relative biases always appeared largest for f_0 the gamma pdf; the relative biases always appeared second largest for f_0 the Weibull pdf; the coefficients of variation always appeared smallest for f_0 the exponential or the exponentiated exponential pdf; the coefficients of variation always appeared largest for f_0 the gamma pdf; the coefficients of variation always appeared second largest for f_0 the Weibull pdf.

The following observations did not hold in general: the relative biases for α, β and a were not always positive, sometimes they were negative depending on the true values for (p, α, a, β) chosen; the relative biases for p were not always negative, sometimes they were positive depending on the true values for (p, α, a, β) chosen; the coefficients of variation for α were not always largest, sometimes they were not the largest depending on the true values for (p, α, a, β) chosen; the coefficients of variation for a were not always smallest, sometimes they were not the smallest depending on the true values for (p, α, a, β) chosen.

Section 6 presents a real data application. The sample size of the data set is one hundred. We shall see that the parallel-parallel-geometric-exponential distribution gives the best to the data among thirty one other distributions having at least the same number of parameters. Based on this fact and the simulation results, we can expect the relative biases for $\widehat{p}$, $\widehat{\alpha}$ and $\widehat{\beta}$ to be less than 0.1, 0.05 and 0.02, respectively. We can expect the coefficients of variation for $\widehat{p}$, $\widehat{\alpha}$ and $\widehat{\beta}$ to be less than 0.6, 0.3 and 0.2, respectively. Hence, the point estimates reported in Section 6 can be expected to be accurate except possibly that for p.

6. REAL DATA APPLICATION

Here, we return to the fiber data set discussed in Section 1. As explained there, the breaking stress of fiber can be modeled as the failure of a system having i) parallel components and parallel sub-components; ii) series components and series sub-components; iii) parallel components and series sub-components; iv) series components and parallel sub-components. We take the number of components as either geometric or Poisson distributed. We fitted the following distributions to the data: parallel-parallel-geometric-f_0, series-series-geometric-f_0, parallel-series-geometric-f_0, series-parallel-geometric-f_0, parallel-parallel-Poisson-f_0, series-series-Poisson-f_0, parallel-series-Poisson-f_0 and series-parallel-Poisson-f_0 with f_0 taken to be an exponential, gamma, Weibull or an exponentiated exponential pdf. In total, thirty two distributions were fitted to the data. Eight of these distributions (parallel-parallel-geometric-exponential, series-series-geometric--exponential, parallel-series-geometric-exponential, series-parallel-geometric-exponential, parallel-parallel-Poisson-exponential, series-series-Poisson-exponential, parallel-series--Poisson-exponential and series-parallel-Poisson-exponential) have three parameters each. The remaining twenty four distributions have four parameters each. The distributions were fitted by the methods of maximum likelihood and moments, see Section 4.

Many of the fitted distributions are not nested. Discrimination among them was performed using various criteria:

- the Akaike information criterion due to Akaike [1] defined by
$$\mathrm{AIC} = 2(q+2) - 2\log L\left(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}}\right);$$
- the Bayesian information criterion due to Schwarz [27] defined by
$$\mathrm{BIC} = (q+2)\log n - 2\log L\left(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}}\right);$$
- the consistent Akaike information criterion (CAIC) due to Bozdogan [2] defined by
$$\mathrm{CAIC} = -2\log L\left(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}}\right) + (q+2)\left(\log n + 1\right);$$
- the corrected Akaike information criterion (AICc) due to Hurvich and Tsai [8] defined by
$$\mathrm{AICc} = \mathrm{AIC} + \frac{2(q+2)(q+3)}{n-q-3};$$

- the Hannan-Quinn criterion due to Hannan and Quinn [7] defined by

$$\mathrm{HQC} = -2\log L\left(\widehat{p}, \widehat{\alpha}, \widehat{\boldsymbol{\Theta}}\right) + 2(q+2)\log\log n;$$

- the p-value of the Kolmogorov–Smirnov statistic (Kolmogorov [10], Smirnov [29]) defined by

$$\sup_x \left| \frac{1}{n}\sum_{i=1}^{n} I\left\{x_i \leq x\right\} - \widehat{F}(x) \right|,$$

  where $I\{\cdot\}$ denotes the indicator function and $\widehat{F}(\cdot)$ the maximum likelihood estimate of $F(x)$;

- the p-value of the Kolmogorov–Smirnov statistic (Kolmogorov [10], Smirnov [29]) defined by

$$\sup_x \left| \frac{1}{n}\sum_{i=1}^{n} I\left\{x_i \leq x\right\} - \widetilde{F}(x) \right|,$$

  where $\widetilde{F}(\cdot)$ is the method of moments estimate of $F(x)$.

The smaller the values of AIC, BIC, CAIC, AICc, and HQC the better the fit. For more discussion on these criteria, see Burnham and Anderson [3].

Since the Kolmogorov–Smirnov test assumes that the fitted distribution gives the “true” parameter values, the p-values were computed by simulation as follows:

(i) fit the distribution to the data and compute the corresponding Kolmogorov–Smirnov statistic;

(ii) generate 10000 samples each of the same size as the data from the fitted model in step (i);

(iii) refit the model to each of the 10000 samples;

(iv) compute the Kolmogorov–Smirnov statistic for the 10000 fits in step (iii);

(v) construct an empirical cdf of the 10000 values of the Kolmogorov–Smirnov statistic obtained in step (iv);

(vi) compare the Kolmogorov–Smirnov statistic obtained in step (i) with the empirical cdf in step (v) to get the p-value.

The values of $-\log L$, AIC, BIC, CAIC, AICc, HQC and the p-values for the thirty two fitted distributions are given in Table 1. The values of AIC, the most commonly used of the criteria, are highlighted. We can see that the smallest AIC, the smallest BIC, the smallest CAIC, the smallest AICc, the smallest HQC and the largest p-values are for the parallel-parallel-geometric-exponential distribution. The second smallest AIC, the second smallest BIC, and the second largest p-values are for the parallel-parallel-geometric-gamma distribution. The second smallest CAIC, the second smallest AICc

Distribution	$-\log L$	AIC	BIC	CAIC	AICc	HQC	KS p-value MLE	KS p-value MME
parallel parallel geo exp	141.279	**288.557**	296.373	299.373	288.807	291.720	0.998	0.970
parallel parallel geo gamma	141.065	**290.129**	300.550	304.550	290.550	294.347	0.909	0.965
parallel parallel geo weibull	141.070	**290.140**	300.561	304.561	290.561	294.357	0.879	0.935
parallel parallel geo ee	141.279	**290.557**	300.978	304.978	290.978	294.775	0.569	0.626
series series geo exp	196.371	**398.743**	406.558	409.558	398.993	401.906	0.022	0.040
series series geo gamma	156.950	**321.901**	332.321	336.321	322.322	326.118	0.150	0.094
series series geo weibull	146.128	**300.255**	310.676	314.676	300.676	304.473	0.189	0.183
series series geo ee	163.791	**335.582**	346.002	350.002	336.003	339.799	0.146	0.080
parallel series geo exp	142.123	**290.246**	298.062	301.062	290.496	293.410	0.830	0.856
parallel series geo gamma	141.068	**290.137**	300.557	304.557	290.558	294.354	0.881	0.945
parallel series geo weibull	141.531	**291.062**	301.483	305.483	291.483	295.279	0.382	0.525
parallel series geo ee	141.075	**290.150**	300.571	304.571	290.571	294.368	0.836	0.885
series parallel geo exp	165.995	**337.990**	345.805	348.805	338.240	341.153	0.117	0.072
series parallel geo gamma	144.701	**297.402**	307.823	311.823	297.824	301.620	0.256	0.313
series parallel geo weibull	143.243	**294.486**	304.907	308.907	294.907	298.703	0.290	0.416
series parallel geo ee	147.034	**302.068**	312.488	316.488	302.489	306.285	0.175	0.110
parallel parallel pois exp	142.837	**291.673**	299.489	302.489	291.923	294.837	0.377	0.524
parallel parallel pois gamma	141.214	**290.429**	300.849	304.849	290.850	294.646	0.665	0.637
parallel parallel pois weibull	141.180	**290.360**	300.780	304.780	290.781	294.577	0.728	0.686
parallel parallel pois ee	142.837	**293.673**	304.094	308.094	294.094	297.891	0.313	0.448
series series pois exp	197.040	**400.080**	407.896	410.896	400.330	403.243	0.020	0.018
series series pois gamma	142.197	**292.394**	302.815	306.815	292.815	296.612	0.346	0.451
series series pois weibull	141.334	**290.669**	301.089	305.089	291.090	294.886	0.518	0.620
series series pois ee	150.516	**309.033**	319.453	323.453	309.454	313.250	0.173	0.096
parallel series pois exp	144.205	**294.410**	302.226	305.226	294.660	297.573	0.307	0.434
parallel series pois gamma	141.175	**290.349**	300.770	304.770	290.770	294.567	0.775	0.710
parallel series pois weibull	141.529	**291.059**	301.479	305.479	291.480	295.276	0.401	0.537
parallel series pois ee	141.174	**290.347**	300.768	304.768	290.768	294.565	0.808	0.805
series parallel pois exp	146.197	**298.395**	306.210	309.210	298.645	301.558	0.243	0.307
series parallel pois gamma	144.639	**297.278**	307.698	311.698	297.699	301.495	0.283	0.398
series parallel pois weibull	141.371	**290.741**	301.162	305.162	291.162	294.958	0.437	0.583
series parallel pois ee	146.337	**300.674**	311.095	315.095	301.095	304.892	0.182	0.139

Tab. 1. Fitted distributions to the fiber data.

and the second smallest HQC are for the parallel-series-geometric-exponential distribution. The largest AIC, the largest BIC, the largest CAIC, the largest AICc, the largest HQC and the smallest p-values are for the series-series-Poisson-exponential distribution.

There is not much difference between the p-values obtained by the methods of maximum likelihood and moments. The relative performances of the thirty two distributions with the respect to the p-values appear the same for both methods. At the five percent level, all of the fitted distributions appear acceptable except for the series-series-geometric-exponential and series-series-Poisson-exponential distributions. However, the best fitting distribution in terms of the seven criteria is the parallel-parallel-geometric-exponential distribution.

The parameter estimates for the best fitting parallel-parallel-geometric-exponential distribution are $\widehat{p} = 0.9209999$, $\widehat{\alpha} = 4.423951$ and $\widehat{\beta} = 1.582465$. This implies that breaking stress of a fiber is that of a system having a geometric number of components working in parallel, the average number being 1. Each component has approximately 4.4 sub-components working in parallel, where the failure time of each sub-component has an exponential distribution with mean equal to 0.632. Another interpretation is that the fiber will break if and only if all the 4.4 sub-components of every component fail.

The probability-probability, quantitle-quantile and density plots for the best fitting parallel-parallel-geometric-exponential distribution are shown in Figures 7, 8 and 9. We see that its fit is reasonable.

The fiber data has been analyzed by several other authors too, including Lemonte and Cordeiro [14], Qian [24] and Shams [28].

Lemonte and Cordeiro [14] fitted the exponentiated generalized inverse Gaussian distribution (equivalent to the parallel-parallel-f_0-generalized inverse Gaussian distribution with $N \equiv 1$) and obtained the estimate $\widehat{\alpha} = 0.127$. This estimate has the interpretation that breaking stress of a fiber is that of a system having 0.1 number of components working in parallel, where the failure time of each component has a generalized inverse Gaussian distribution.

Qian [24] fitted the exponentiated exponential distribution (equivalent to the parallel-parallel-f_0-exponential distribution with $N \equiv 1$) and obtained the estimates $\widehat{\alpha} = 7.788$, $\widehat{\beta} = 1.013$. These estimates have the interpretation that breaking stress of a fiber is that of a system having 7.8 number of components working in parallel, where the failure time of each component has an exponential distribution with mean equal to 0.987.

Shams [28] fitted the Kumaraswamy-generalized exponentiated Pareto distribution (equivalent to the series-parallel-f_0-generalized exponential Pareto distribution with $N \equiv b$) and obtained the estimates $\widehat{\alpha} = 1$, $\widehat{b} = 4.638$. These estimates have the interpretation that breaking stress of a fiber is that of a system having 4.6 number of components working in series, where the failure time of each component has a generalized exponentiated Pareto distribution.

The AIC values for the distributions fitted by Lemonte and Cordeiro [14], Qian [24] and Shams [28] are 291.44, 290.66 and 1969.28, respectively. The corresponding BIC values are 301.86, 301.08 and 1979.70. So, none of the distributions fitted by Lemonte and Cordeiro [14], Qian [24] and Shams [28] provide as good a fit as our parallel-parallel-geometric-exponential distribution.

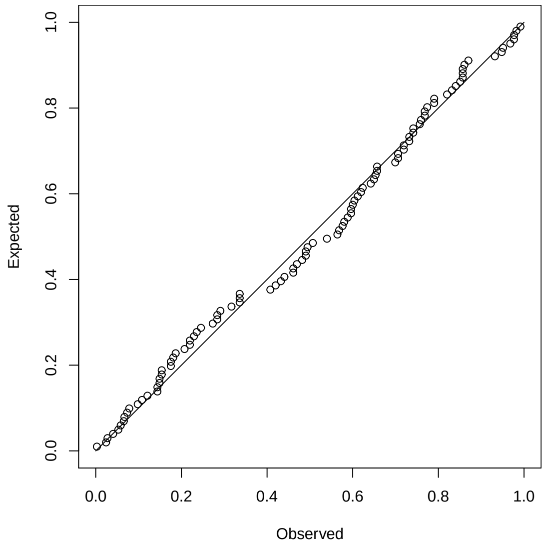

Fig. 7. PP plot for the parallel-parallel-geometric-exponential distribution fitted to the fiber data.

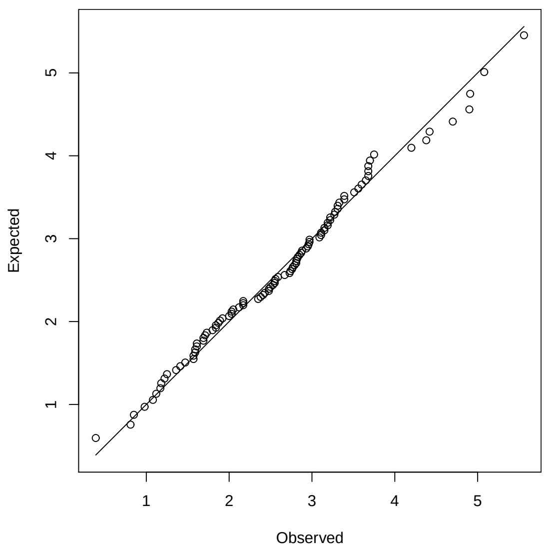

Fig. 8. QQ plot for the parallel-parallel-geometric-exponential distribution fitted to the fiber data.

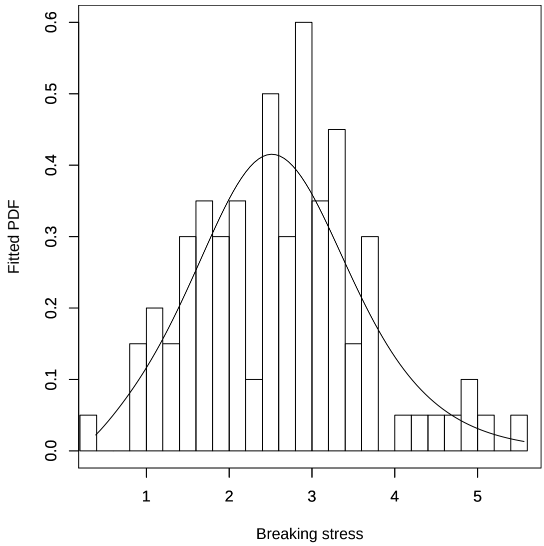

Fig. 9. Fitted pdf of the parallel-parallel-geometric-exponential distribution and the observed histogram for the fiber data.

7. CONCLUSIONS

We have introduced eight new families of distributions, motivated by a data set on breaking stress of fibers. We have studied mathematical properties and estimation issues of one of the families. The properties studied include the quantile function, and moments. Estimation is considered using both the method of maximum likelihood and method of moments.

The data application has shown that a three-parameter distribution gives the best fit for breaking stress. The best fit was determined in terms of AIC, BIC, CAIC, AICc, HQC, p-value of the Kolmogorov–Smirnov statistic based on the method of maximum likelihood and p-value of the Kolmogorov–Smirnov statistic based on the method of moments.

The distributions given by (4) − (13) are the distributions of the maximum/minimum of a random number of independent and identical random variables. A future work is to model the breaking stress as: i) the maximum/minimum of a random number of independent but non-identical random variables; ii) the maximum/minimum of a random number of dependent random variables.

ACKNOWLEDGMENTS

The authors would like to thank the Editors and the referees for careful reading and for comments which greatly improved the paper.

(Received July 30, 2014)

REFERENCES

[1] H. Akaike: A new look at the statistical model identification. IEEE Trans. Automat. Control *19* (1974), 716–723. DOI:10.1109/tac.1974.1100705

[2] H. Bozdogan: Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. Psychometrika *52* (1987), 345–370. DOI:10.1007/bf02294361

[3] K. P. Burnham and D.,R. Anderson: Multimodel inference: Understanding AIC and BIC in model selection. Sociolog. Methods Res. *33* (2004), 261–304. DOI:10.1177/0049124104268644

[4] T. S. Ferguson: A Course in Large Sample Theory. Chapman and Hall, London 1996. DOI:10.1007/978-1-4899-4549-5

[5] R. C. Gupta, P. L. Gupta, and R. D. Gupta: Modeling failure time data by Lehman alternatives. Commun. Statist. — Theory and Methods *27* (1998), 887–904. DOI:10.1080/03610929808832134

[6] R. D. Gupta and D. Kundu: Generalized exponential distributions. Australian and New Zealand J. Statist. *41* (1999), 173–188. DOI:10.1111/1467-842x.00072

[7] E. J. Hannan and B. G. Quinn: The determination of the order of an autoregression. J. Royal Statist. Soc. B *41* (1979), 190–195.

[8] C. M. Hurvich and C.-L. Tsai: Regression and time series model selection in small samples. Biometrika *76* (1989), 297–307. DOI:10.1093/biomet/76.2.297

[9] C. S. Kakde and D. T. Shirke: On exponentiated lognormal distribution. Int. J. Agricult. Statist. Sci. *2* (2006), 319–326.

[10] A. Kolmogorov: Sulla determinazione empirica di una legge di distribuzione. Giornale dell'Istituto Italiano degli Attuari *4* (1933), 83–91.

[11] K. Kolowrocki: Reliability of Large Systems. Elsevier, New York 2004.

[12] M. R. Leadbetter, G. Lindgren and H. Rootzén: Extremes and Related Properties of Random Sequences and Processes. Springer Verlag, New York 1987.

[13] L. E. Lehmann and G. Casella: Theory of Point Estimation. Second edition. Springer Verlag, New York 1998. DOI:10.1007/b98854

[14] A. J. Lemonte and G. M. Cordeiro: The exponentiated generalized inverse Gaussian distribution. Statist. Probab. Lett. *81* (2011), 506–517. DOI:10.1016/j.spl.2010.12.016

[15] A. W. Marshall and I. Olkin: 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. DOI:10.1093/biomet/84.3.641

[16] G. S. Mudholkar and D. K. Srivastava: Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Trans. Reliability *42* (1993), 299–302.

[17] G. S. Mudholkar, D. K. Srivastava and M. Friemer: The exponential Weibull family: Analysis of the bus-motor-failure data. Technometrics *37* (1995), 436–445. DOI:10.2307/1269735

[18] G. S. Mudholkar, D. K. Srivastava and G. D. Kollia: A generalization of the Weibull distribution with application to the analysis of survival data. J. Amer. Statist. Assoc. *91* (1996), 1575–1583. DOI:10.2307/2291583

[19] S. Nadarajah: The exponentiated Gumbel distribution with climate application. Environmetrics *17* (2005), 13–23. DOI:10.2307/2291583

[20] S. Nadarajah: The exponentiated exponential distribution: A survey. Adv. Statist. Anal. *95* (2011), 219–251. DOI:10.1007/s10182-011-0154-5

[21] S. Nadarajah and A. K. Gupta: The exponentiated gamma distribution with application to drought data. Calcutta Statist. Assoc. Bull. *59* (2007), 29–54.

[22] S. Nadarajah and S. Kotz: The exponentiated type distributions. Acta Applic. Math. *92* (2006), 97–111. DOI:10.1007/s10440-006-9055-0

[23] M. D. Nichols and W. J. Padgett: A bootstrap control chart for Weibull percentiles. Qual. Reliab. Engrg. Int. *22* (2006), 141–151. DOI:10.1002/qre.691

[24] L. Qian: The Fisher information matrix for a three-parameter exponentiated Weibull distribution under type II censoring. Statist. Meth. *9* (2012), 320–329. DOI:10.1016/j.stamet.2011.08.007

[25] R Development Core Team: R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria 2014.

[26] M. Ristic and S. Nadarajah: A new lifetime distribution. J. Statist. Comput. Simul. *84* (2014), 135–150. DOI:10.1080/00949655.2012.697163

[27] G. E. Schwarz: Estimating the dimension of a model. Ann. Statist. *6* (1978), 461–464. DOI:10.1214/aos/1176344136

[28] T. M. Shams: The Kumaraswamy-generalized exponentiated Pareto distribution. European J. Appl. Sci. *5* (2013), 92–99.

[29] N. Smirnov: Table for estimating the goodness of fit of empirical distributions. Ann. Math. Statist. *19* (1948), 279–281. DOI:10.1214/aoms/1177730256

Nooshin Hakamipour, Department of Statistics, Amirkabir University of Technology, Tehran. Iran.

Sadegh Rezaei, Department of Statistics, Amirkabir University of Technology, Tehran. Iran.

Saralees Nadarajah, School of Mathematics, University of Manchester, Manchester M13 9PL. United Kingdom.
e-mail: mbbsssn2@manchester.ac.uk