10-2015

The Kumaraswamy Marshal-Olkin Family of Distributions

Morad Alizadeh
Gulf University-Bushehr

M. H. Tahir
Islamia University, Bahawalpur

Gauss M. Cordeiro
Universidade Federal de Pernambuco

M. Mansoor
University of Bahawalpur-Pakistan

Muhammad Zubair
Government Degree College Kahrorpacca

See next page for additional authors

Journal of the Egyptian Mathematical Society (2015) **23**, 546–557

Egyptian Mathematical Society

Journal of the Egyptian Mathematical Society

www.etms-eg.org
www.elsevier.com/locate/joems

ORIGINAL ARTICLE

The Kumaraswamy Marshal-Olkin family of distributions

Morad Alizadeh [a], M.H. Tahir [b,*], Gauss M. Cordeiro [c], M. Mansoor [b], M. Zubair [d], G.G. Hamedani [e]

[a] *Department of Statistics, Ferdowsi University of Mashhad, Mashhad, P.O. Box. 91775-1159, Iran*
[b] *Department of Statistics, The Islamia University of Bahawalpur, Bahawalpur 63100, Pakistan*
[c] *Department of Statistics, Federal University of Pernambuco, 50740-540 Recife, PE, Brazil*
[d] *Department of Statistics, Government Degree College Khairpur Tamewali, Pakistan*
[e] *Department of Mathematics, Statistics and Computer Science, Marquette University, Milwaukee, USA*

Received 5 August 2014; revised 30 November 2014; accepted 18 December 2014
Available online 24 February 2015

KEYWORDS

Characterization;
Estimation;
Fréchet distribution;
Kumaraswamy distribution;
Marshall-Olkin family;
Weibull distribution

Abstract We introduce a new family of continuous distributions called the Kumaraswamy Marshal-Olkin generalized family of distributions. We study some mathematical properties of this family. Its density function is symmetrical, left-skewed, right-skewed and reversed-J shaped, and has constant, increasing, decreasing, upside-down bathtub, bathtub and S-shaped hazard rate. We present some special models and investigate the asymptotics and shapes of the family. We derive a power series for the quantile function and obtain explicit expressions for the moments, generating function, mean deviations, two types of entropies and order statistics. Some useful characterizations of the family are also proposed. The method of maximum likelihood is used to estimate the model parameters. We illustrate the importance of the family by means of two applications to real data sets.

2000 MATHEMATICS SUBJECT CLASSIFICATION: 60E05; 62E15; 62N05

1. Introduction

There has be an increased interest in defining new generators or generalized classes of univariate continuous distributions by introducing additional shape parameter(s) to a baseline model. The extended distributions have attracted several statisticians to develop new models because the computational and analytical facilities available in programming softwares such as R, Maple and Mathematica can easily tackle the problems involved in computing special functions in these extended distributions. Several mathematical properties of the extended distributions

* Corresponding author. Tel.: +92 62 9255454 (Office).
E-mail addresses: mtahir.stat@gmail.com, mht@iub.edu.pk (M.H. Tahir).

Peer review under responsibility of Egyptian Mathematical Society.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.joems.2014.12.002
1110-256X

may easily be explored using mixture forms of the exponentiated-G ("exp-G" for short) distributions. The addition of parameters has been proved useful in exploring skewness and tail properties, and also for improving the goodness-of-fit of the generated family. The well-known generators are the following: beta-G by Eugene et al. [1] and Jones [2], Kumaraswamy-G (Kw-G) by Cordeiro and de Castro [3], McDonald-G (Mc-G) by Alexander et al. [4], gamma-G type 1 by Zografos and Balakrishanan [5], and Amini et al. [6], gamma-G type 2 by Ristić and Balakrishanan [7], and Amini et al. [6], odd-gamma-G type 3 by Torabi and Montazari [8], logistic-G by Torabi and Montazari [9], transformed-transformer (T-X) (Weibull-X and gamma-X) by Alzaatreh et al. [10], discrete T-X by Alzaatreh et al. [11], exponentiated T-X by Alzaghal et al. [12], odd exponentiated generalized (odd exp-G) by Cordeiro et al. [13], odd Weibull-G by Bourguignon et al. [14], exponentiated half-logistic by Cordeiro et al. [15], T-X{Y}-quantile based approach by Aljarrah et al. [16] and T-R{Y} by Alzaatreh et al. [17], Lomax-G by Cordeiro et al. [18], logistic-X by Tahir et al. [19] and new Weibull-G by Tahir et al. [20].

Marshall and Olkin [21] proposed a flexible semi-parametric family of distributions and defined a new survival function $\overline{G}_{MO}(x)$ by introducing an additional parameter $\bar{p}$ such that $p = 1 - \bar{p}$ and $\bar{p} > 0$. They called the parameter $\bar{p}$ the *tilt parameter* and interpreted $\bar{p}$ in terms of the behavior of the hazard rate function (hrf) of $\overline{G}_{MO}(x)$. Their ratio is increasing in x for $\bar{p} \geqslant 1$ and decreasing in x for $\bar{p} \in (0,1)$.

For any arbitrary continuous probability density function (pdf) $g(x;\xi)$ and cumulative distribution function (cdf) $G(x;\xi)$, the cdf and pdf of the Marshall-Olkin (MO) family of distributions are defined by ($x \in \Re$, $\bar{p} > 0$ and $p = 1 - \bar{p}$).

$$G_{MO}(x) = \frac{G(x;\xi)}{1 - p\,\overline{G}(x;\xi)} \tag{1}$$

and

$$g_{MO}(x) = \frac{(1-p)\,g(x;\xi)}{\left[1 - p\,\overline{G}(x;\xi)\right]^2}, \quad x \in \Re,$$

respectively. For $p = 0$, we have $G_{MO}(x) = G(x;\xi)$.

For a baseline random variable having pdf $g(x)$ and cdf $G(x)$, Cordeiro and de Castro [3] defined the two-parameter Kw-G cdf by

$$F(x) = 1 - \{1 - G(x)^a\}^b. \tag{2}$$

The pdf corresponding to (2) becomes

$$f(x) = a\,b\,g(x)\,G(x)^{a-1}\{1 - G(x)^a\}^{b-1}, \tag{3}$$

where $g(x) = dG(x)/dx$ and $a > 0$ and $b > 0$ are two extra shape parameters whose role is to govern skewness and tail weights.

Now, we propose a new extension of the MO family for a given baseline distribution with cdf $G(x;\xi)$, survival function $\overline{G}(x;\xi) = 1 - G(x;\xi)$ and pdf $g(x;\xi)$ depending on a parameter vector ξ. Inserting (1) in (2), we define the cdf of the new *Kumaraswamy Marshal-Olkin* ("KwMO") family of distributions by

$$F(x) = F(x;a,b,p,\xi) = 1 - \left\{1 - \left[\frac{G(x;\xi)}{1 - p\,\overline{G}(x;\xi)}\right]^a\right\}^b, \tag{4}$$

where $a > 0, b > 0$ and $\bar{p} > 0$ are three additional shape parameters. For each baseline G, the "KwMO-G" cdf is given by (4).

Eq. (4) provides a wider family of continuous distributions. It includes the Kw-G family of distributions, the proportional and reversed hazard rate models, the MO family of distributions and other sub-families. In Table 1, we provide some special models of the KwMO family of distributions.

The density function corresponding to (4) is given by

$$f(x) = f(x;a,b,p,\xi) = \frac{ab(1-p)g(x;\xi)G(x;\xi)^{a-1}}{\left[1 - p\overline{G}(x;\xi)\right]^{a+1}}\left\{1 - \left[\frac{G(x;\xi)}{1 - p\overline{G}(x;\xi)}\right]^a\right\}^{b-1}. \tag{5}$$

Eq. (5) will be most tractable when the cdf $G(x)$ and the pdf $g(x)$ have simple analytic expressions. Hereafter, a random variable X with density function (5) is denoted by $X \sim \mathrm{KwMO} - \mathrm{G}(a,b,p,\xi)$. Further, we omit sometimes the dependence on the vector ξ of the parameters and write simply $G(x) = G(x;\xi), F(x) = F(x;a,b,p,\xi)$ and so on.

The hrf of X becomes

$$h(x;a,b,p,\xi) = \frac{a\,b\,(1-p)\,g(x;\xi)\,G(x;\xi)^{a-1}}{\left[1 - p\,\overline{G}(x;\xi)\right]\left\{\left[1 - p\,\overline{G}(x;\xi)\right]^a - G(x;\xi)^a\right\}}. \tag{6}$$

This paper is organized as follows. Four special cases of this family are presented in Section 2. Some mathematical properties are provided in Section 3 such as the shapes of the density and hazard rate functions, useful expansions for the cdf, pdf and quantile function (qf), explicit expressions for the moments, generating function, mean deviations, Rényi and Shannon entropies and order statistics. Section 4 refers to some characterizations of the KwMO family. Estimation of the model parameters by maximum likelihood is performed in Section 5. Two applications to real data sets illustrate the potentiality of the new family in Section 6. The paper is concluded in Section 7.

2. Special models

Here, we provide a few examples of the KwMO-G family of distributions.

2.1. The KwMO-Exponential (KwMO-E) distribution

Let the parent distribution be exponential with parameter $\lambda > 0, g(x;\lambda) = \lambda e^{-\lambda x}, x > 0$ and $G(x;\lambda) = 1 - e^{-\lambda x}$. Then, the pdf of the KwMO-E (for $x > 0$) model is given by

$$f_{KwMOE}(x) = \frac{ab\lambda e^{-\lambda x}(1-p)(1-e^{-\lambda x})^{a-1}}{(1-pe^{-\lambda x})^{a+1}}\left\{1 - \left[\frac{1-e^{-\lambda x}}{1-pe^{-\lambda x}}\right]^a\right\}^{b-1}.$$

2.2. The KwMO-Lomax (KwMO-L) distribution

Consider the parent Lomax distribution with positive parameters α and β and pdf and cdf given by $g(x;\alpha,\beta) = (\alpha/\beta)[1 + (x/\beta)]^{-(\alpha+1)}, x > 0$, $G(x;\alpha,\beta) = 1 - [1 + (x/\beta)]^{-\alpha}$. Then, the pdf of the KwMO-L distribution reduces to

$$f_{KwMOL}(x) = \frac{ab\alpha[1 + (x/\beta)]^{-(\alpha+1)}(1-p)\{1 - [1 + (x/\beta)]^{-\alpha}\}^{a-1}}{\beta\{1 - p[1 + (x/\beta)]^{-\alpha}\}^{a+1}} \times \left\{1 - \left[\frac{1 - [1 + (x/\beta)]^{-\alpha}}{1 - p[1 + (x/\beta)]^{-\alpha}}\right]^a\right\}^{b-1}.$$

Table 1 Some special models.

S. no.	a	b	p	$G(x)$	Reduced model
1	–	–	0	$G(x)$	The Kw-G family of distributions [3]
2	1	1	–	$G(x)$	The MO family of distributions [21]
3	1	1	–	$G(x)$	The exponentiated MO family of distributions [22]
4	1	–	0	$G(x)$	The proportional reversed hazard rate model [23]
5	–	1	0	$G(x)$	The proportional hazard rate model [24]
6	1	1	0	$G(x)$	$G(x)$

2.3. The KwMO-Weibull (KwMO-W) distribution

Consider the parent Weibull distribution with positive parameters λ and β. Then, the pdf and cdf are given by $g(x) = \lambda\beta x^{\beta-1}\mathrm{e}^{-\lambda x^{\beta}}$ and $G(x) = 1 - \mathrm{e}^{-\lambda x^{\beta}}$, respectively. Then, the pdf of KwMO-W distribution becomes

$$f_{KwMOW}(x) = \frac{ab\lambda\beta x^{\beta-1}\mathrm{e}^{-\lambda x^{\beta}}(1-p)\left\{1-\mathrm{e}^{-\lambda x^{\beta}}\right\}^{a-1}}{\left\{1-p\,\mathrm{e}^{-\lambda x^{\beta}}\right\}^{a+1}} \times \left\{1-\left[\frac{1-\mathrm{e}^{-\lambda x^{\beta}}}{1-p\,\mathrm{e}^{-\lambda x^{\beta}}}\right]^{a}\right\}^{b-1}.$$

For $\beta = 2$, we obtain as special case the KwMO-Rayleigh (KwMO-R) distribution.

2.4. The KwMO-Fréchet (KwMO-Fr) distribution

Now, suppose the parent Fréchet distribution with pdf and cdf given by $g(x) = \lambda\delta^{\lambda}x^{-(\lambda+1)}\mathrm{e}^{-(\delta/x)^{\lambda}}$ and $G(x) = \mathrm{e}^{-(\delta/)^{\lambda}}$, $x > 0$, respectively, then the pdf of the KwMO-Fr model reduces to

$$f_{KwMOFr}(x;a,b,p,\lambda,\delta) = \frac{ab(1-p)\lambda\delta^{\lambda}x^{-(\lambda+1)}\mathrm{e}^{-(\delta/x)^{\lambda}}\left[\mathrm{e}^{-(\delta/x)^{\lambda}}\right]^{a-1}}{\left\{1-p\left[1-\mathrm{e}^{-(\delta/x)^{\lambda}}\right]\right\}^{a+1}} \times \left\{1-\left[\frac{\mathrm{e}^{-(\delta/x)^{\lambda}}}{1-p\left(1-\mathrm{e}^{-(\delta/x)^{\lambda}}\right)}\right]^{a}\right\}^{b-1}.$$

Figs. 1 and 2 display some plots of the pdf and hrf of the KwMO-E, KwMO-L, KwMO-W and KwMO-Fr distributions for selected parameter values. Fig. 1 indicates that the KwMO family generates distributions with various shapes such as symmetric, left-skewed, right-skewed and reversed-J. Further, Fig. 2 shows that the KwMO family produces flexible hazard rate shapes such as constant, increasing, decreasing, bathtub, upside-down bathtub and S. This indeed reveals that the KwMO family is very useful in fitting different data sets with various shapes.

3. Mathematical properties

3.1. Asymptotics and shapes

Proposition 1. *The asymptotics of* Eqs. (4)–(6) *as* $G(x) \to 0$ *are given by*

$$F(x) \sim \frac{b\,G(x)^{a}}{(1-p)^{a}} \quad \text{as} \quad G(x) \to 0,$$
$$f(x) \sim \frac{a\,b\,g(x)\,G(x)^{a-1}}{(1-p)^{a}} \quad \text{as} \quad G(x) \to 0,$$
$$h(x) \sim \frac{a\,b\,g(x)\,G(x)^{a-1}}{(1-p)^{a}} \quad \text{as} \quad G(x) \to 0.$$

Proposition 2. *The asymptotics of* Eqs. (4)–(6) *as* $x \to \infty$ *are given by*

$$1 - F(x) \sim \left[a(1-p)\overline{G}(x)\right]^{b} \quad \text{as} \quad x \to \infty,$$
$$f(x) \sim b[a(1-p)]^{b}g(x)\overline{G}(x)^{b-1} \quad \text{as} \quad x \to \infty,$$
$$h(x) \sim \frac{b\,g(x)}{\overline{G}(x)} \quad \text{as} \quad x \to \infty.$$

The shapes of the density and hazard rate functions are described analytically. The critical points of the density of the KwMO-G model are the roots of the equation:

$$\frac{g'(x)}{g(x)} + (a-1)\frac{g(x)}{G(x)} - p(a+1)\frac{g(x)}{1-p\overline{G}(x)} = a(1-b)\frac{g(x)G(x)^{a-1}}{\left[1-p\overline{G}(x,\xi)\right]\left\{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}\right\}}. \tag{7}$$

There may be more than one roots to (7). Let $\lambda(x) = \frac{d^2\log[f(x)]}{dx^2}$. We have

$$\begin{aligned}\lambda(x) = {} & \frac{g''(x)g(x)-g'(x)^2}{g(x)^2} + (a-1)\frac{g'(x)G(x)-g(x)^2}{G(x)^2} \\ & - p(a+1)\frac{g'(x)\left[1-p\overline{G}(x)\right]+pg(x)^2}{\left[1-p\overline{G}(x)\right]^2} \\ & - a(b-1)\frac{g'(x)G(x)^{a-1}}{\left[1-p\overline{G}(x)\right]\left\{\left[1-p\overline{G}(x)\right]^{a}-G(x)^{a}\right\}} \\ & - a(a-1)(b-1)\frac{g(x)^2G(x)^{a-2}}{\left[1-p\overline{G}(x)\right]\left\{\left[1-p\overline{G}(x)\right]^{a}-G(x)^{a}\right\}} \\ & + pa(b-1)\frac{g(x)^2G(x)^{a-1}}{\left[1-p\overline{G}(x,\xi)\right]^2\left\{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}\right\}} \\ & - a^2(b-1)\frac{g(x)G(x)^{a-1}\left\{p\left[1-p\overline{G}(x,\xi)\right]^{a-1}-G(x)^{a-1}\right\}}{\left[1-p\overline{G}(x,\xi)\right]\left\{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}\right\}^2}.\end{aligned}$$

If $x = x_0$ is a root of (7) then it corresponds to a local maximum if $\lambda(x) > 0$ for all $x < x_0$ and $\lambda(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\lambda(x) < 0$ for all $x < x_0$ and $\lambda(x) > 0$ for all $x > x_0$. It refers to an inflexion point if either $\lambda(x) > 0$ for all $x \neq x_0$ or $\lambda(x) < 0$ for all $x \neq x_0$.

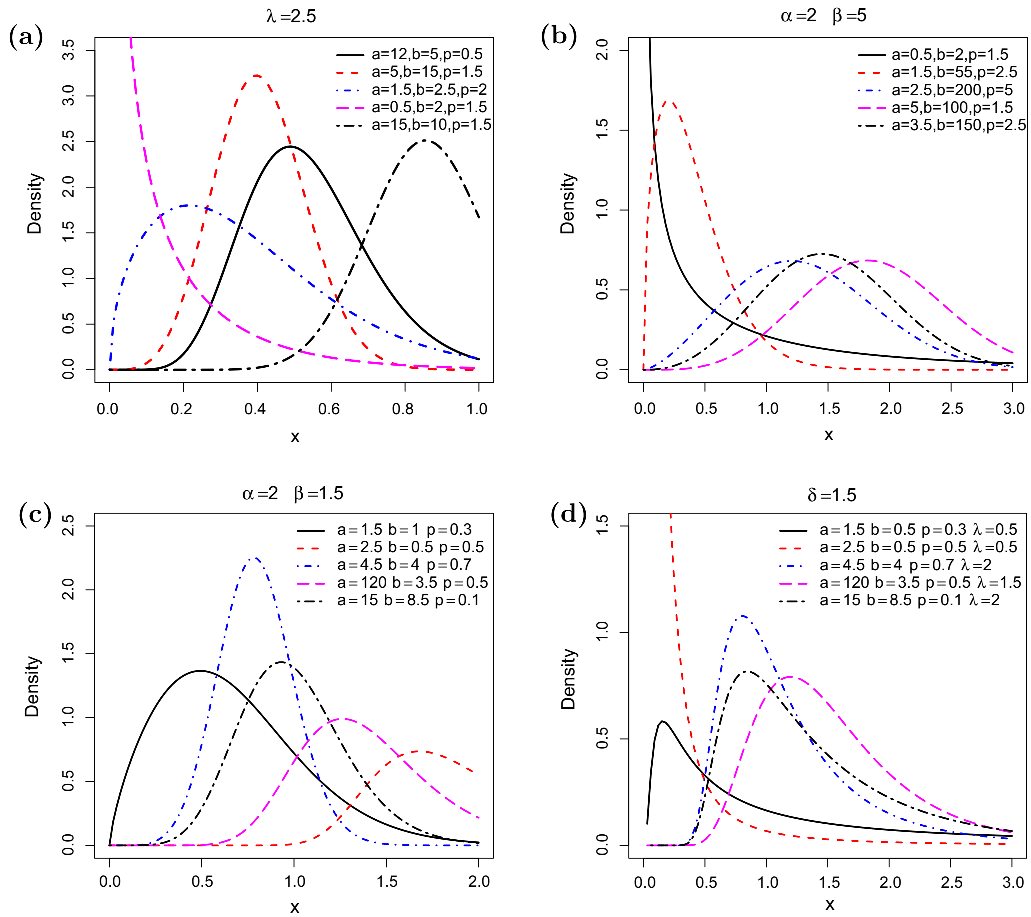

Fig. 1 Density plots: (a) KwMO-E, (b) KwMO-L, (c) KwMO-W and (d) KwMO-Fr models.

The critical point of $h(x)$ is the roots of the equation

$$\frac{g'(x)}{g(x)}+\frac{(a-1)g(x)}{G(x)}-\frac{p\,g(x)}{1-p\overline{G}(x,\xi)} = ag(x)\frac{p\left[1-p\overline{G}(x,\xi)\right]^{a-1}-G(x)^{a-1}}{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}}. \tag{8}$$

There may be more than one roots to (8). Let $\tau(x) = d^2\log[h(x)]/dx^2$. We have

$$\begin{aligned}\tau(x) =& \frac{g''(x)g(x)-g'(x)^2}{g(x)^2}+\frac{(a-1)[g'(x)G(x)-g(x)^2]}{G(x)^2}\\ &-p\frac{g'(x)\left[1-p\overline{G}(x)\right]+pg(x)^2}{\left[1-p\overline{G}(x)\right]^2}\\ &-ag'(x)\frac{p\left[1-p\overline{G}(x,\xi)\right]^{a-1}-G(x)^{a-1}}{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}}\\ &-a(a-1)g(x)^2\,\frac{p^2\left[1-p\overline{G}(x,\xi)\right]^{a-2}-G(x)^{a-2}}{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}}\\ &-\left\{ag(x)\frac{p\left[1-p\overline{G}(x,\xi)\right]^{a-1}-G(x)^{a-1}}{\left[1-p\overline{G}(x,\xi)\right]^{a}-G(x)^{a}}\right\}^2.\end{aligned}$$

If $x = x_0$ is a root of (8) then it refers to a local maximum if $\tau(x) > 0$ for all $x < x_0$ and $\tau(x) < 0$ for all $x > x_0$. It corresponds to a local minimum if $\tau(x) < 0$ for all $x < x_0$ and $\tau(x) > 0$ for all $x > x_0$. It gives an inflexion point if either $\tau(x) > 0$ for all $x \neq x_0$ or $\tau(x) < 0$ for all $x \neq x_0$.

3.2. Useful expansions

We can demonstrate that the cdf (4) admits the expansion

$$F(x) = 1-\sum_{i=0}^{\infty}(-1)^i\binom{b}{i}\frac{G(x)^{a\,i}}{\left[1-p\overline{G}(x)\right]^{a\,i}}. \tag{9}$$

We can obtain an expansion for $G(x)^{\beta}$ ($\beta > 0$ real non-integer) as

$$G(x)^{\beta} = \sum_{r=0}^{\infty}s_r(\beta)\,G(x)^r, \tag{10}$$

where

$$s_r(\beta) = \sum_{j=r}^{\infty}(-1)^{r+j}\binom{\beta}{j}\binom{j}{r}.$$

Then, using (10), we obtain

$$\frac{G(x)^{a\,i}}{\left[1-p\,\overline{G}(x)\right]^{a\,i}} = \frac{\sum_{k=0}^{\infty}\alpha_k\,G(x)^k}{\sum_{k=0}^{\infty}\beta_k\,G(x)^k} = \sum_{k=0}^{\infty}\gamma_k\,G(x)^k, \tag{11}$$

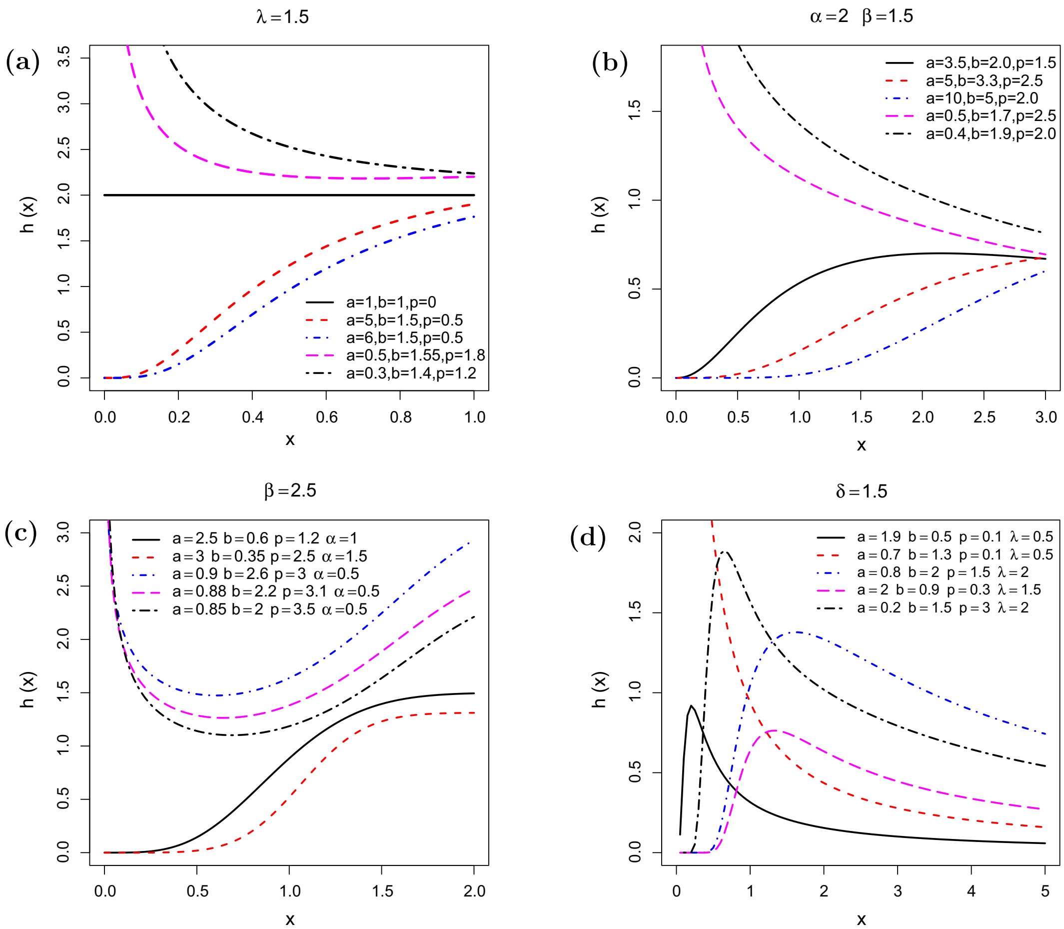

Fig. 2 Hazard plots: (a) KwMO-E, (b) KwMO-L, (c) KwMO-W and (d) KwMO-Fr models.

where

$$\alpha_k = \alpha_k(a,i) = \sum_{j=k}^{\infty}(-1)^{j+k}\binom{a\,i}{j}\binom{j}{k},$$

$$\beta_k = \beta_k(a,p,i) = \sum_{j=k}^{\infty}(-1)^{j+k}p^j\binom{a\,i}{j}\binom{j}{k},$$

and for $k \geq 1$

$$\gamma_k = \gamma_k(a,p,i) = \frac{1}{\beta_0}\left(\alpha_k - \frac{1}{\beta_0}\sum_{r=1}^{k}\alpha_r\,\beta_{k-r}\right),$$

and $\gamma_0 = \alpha_0/\beta_0$. Then, we obtain

$$F(x) = F(x;a,b,p,\xi) = \sum_{k=0}^{\infty} b_k\,G(x)^k, \tag{12}$$

where $a_k = \sum_{i=0}^{\infty}(-1)^i\binom{b}{i}\gamma_k(a,p,i), b_0 = 1 - a_0$ and, for $k \geqslant 1, b_k = -a_k$, and $H_a(x) = G(x)^a$ denotes the exponentiated-G ("exp-G" for short) cdf with power parameter $a > 0$. The last results hold for real non-integer a. For integer a, it is clear that the indices should stop in integers and we can easily update the formula.

The density function of X can be expressed as an infinite linear combination of exp-G densities, namely

$$f(x) = f(x;a,b,p,\xi) = \sum_{k=0}^{\infty} b_{k+1}\,h_{k+1}(x), \tag{13}$$

where $h_{k+1}(x)$ is the exp-G density with power parameter $k+1$. Thus, some mathematical properties of the new distribution can be derived from those properties of the exp-G distribution based on (13). For example, the ordinary and incomplete moments and generating function of X can be obtained from those quantities of the exp-G distribution.

3.3. Quantile power series

Let $Q_G(\cdot) = G^{-1}(\cdot)$ be the baseline qf. The KwMO-G distribution is easily simulated by inverting (4) as follows: if u has a uniform $U(0,1)$ distribution, the solution of the non-linear equation

$$x = Q_G\left(\frac{(1-p)\left[1-(1-u)^{\frac{1}{b}}\right]^{\frac{1}{a}}}{1-p\left\{1-\left[1-(1-u)^{\frac{1}{b}}\right]^{\frac{1}{a}}\right\}}\right) \tag{14}$$

has the density function (5).

The effects of the shape parameters a and b on the skewness and kurtosis can be considered based on quantile measures. Now, we derive a power series for the qf $x = Q(u) = F^{-1}(u)$

of X by expanding (14). First, if $Q_G(u)$ does not have an explicit expression, it can usually be expressed as a power series

$$Q_G(u) = \sum_{i=0}^{\infty} a_i\, u^i, \tag{15}$$

where the coefficients a_i's are suitably chosen real numbers which depend on the parameters of the G distribution. For several important distributions, such as the normal, Student t, gamma and beta distributions, $Q_G(u)$ does not have explicit expressions but it can be expanded as in Eq. (15).

From now on, we use a result by Gradshteyn and Ryzhik [25] for a power series raised to a positive integer n (for $n \geq 1$)

$$Q_G(u)^n = \left(\sum_{i=0}^{\infty} a_i\, u^i\right)^n = \sum_{i=0}^{\infty} c_{n,i}\, u^i, \tag{16}$$

where the coefficients $c_{n,i}$ (for $i = 1, 2, \ldots$) are determined from the recurrence equation (with $c_{n,0} = a_0^n$)

$$c_{n,i} = (i\, a_0)^{-1} \sum_{m=1}^{i} [m(n+1) - i]\, a_m\, c_{n,i-m}. \tag{17}$$

Clearly, the quantity $c_{n,i}$ is obtained from $c_{n,0}, \ldots, c_{n,i-1}$ and then from the quantities $a_0, \ldots, a_i$.

Next, we derive an expansion for the argument of $Q_G(\cdot)$ in (14)

$$A = \frac{(1-p)\left[1-(1-u)^{\frac{1}{b}}\right]^{\frac{1}{a}}}{1-p\left\{1-\left[1-(1-u)^{\frac{1}{b}}\right]^{\frac{1}{a}}\right\}} = \frac{\sum_{k=0}^{\infty} a_k^* u^k}{\sum_{k=0}^{\infty} b_k^* u^k},$$

where $a_k^* = (1-p)\sum_{i=0}^{\infty}(-1)^{i+k}\binom{\frac{1}{a}}{i}\binom{\frac{i}{b}}{k}$, $b_0^* = 1-p$ and

$$b_k^* = p\sum_{i,k=0}^{\infty}(-1)^{i+k+1}\binom{\frac{1}{a}}{i}\binom{\frac{i}{b}}{k}.$$

The ratio of the two power series can be expressed as

$$A = \sum_{k=0}^{\infty} c_k^*\, u^k, \tag{18}$$

where $c_0^* = a_0^*/b_0^*$ and the coefficients c_k^*'s (for $k \geqslant 0$) follow from the recurrence equation

$$c_k = \frac{1}{b_0^*}\left(a_k^* - \frac{1}{b_0^*}\sum_{r=1}^{k} b_r^*\, c_{k-r}^*\right).$$

Then, the qf of X follows from (14) by combining (15) and (18) as

$$Q(u) = Q_G\left(\sum_{k=0}^{\infty} c_k^*\, u^k\right) = \sum_{i=0}^{\infty} a_i \left(\sum_{k=0}^{\infty} c_k^*\, u^k\right)^i. \tag{19}$$

Further, using (16) and (17), we obtain

$$Q(u) = \sum_{k=0}^{\infty} e_k\, u^k, \tag{20}$$

where $e_k = \sum_{i=0}^{\infty} a_i\, d_{i,k}$, $d_{i,0} = c_0^{*i}$ and (for $k > 1$)

$$d_{i,k} = (k\, c_0^*)^{-1} \sum_{m=1}^{k} [m(i+1) - k]\, c_m^*\, d_{i,k-m}.$$

Eq. (20) is the main result of this section since it allows to obtain various mathematical quantities for the KwMO family as demonstrated in the next sections. The formulae derived throughout the paper can be easily handled in most symbolic computation software platforms such as Maple, Mathematica and Matlab.

3.4. Moments

Let $Y_{k+1}(k \geqslant 0)$ be a random variable having the exp-G pdf $h_{k+1}(x)$ with power parameter $k+1$. A first formula for the nth moment of X follows from (13) as

$$E(X^n) = \sum_{k=0}^{\infty} b_{k+1}\, E(Y_{k+1}^n). \tag{21}$$

Moments of some exp-G distributions are given by Nadarajah and Kotz [26], which can be used to obtain $E(X^n)$.

A second formula for $E(X^n)$ follows from (21) as

$$E(X^n) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, \tau(n,k), \tag{22}$$

where $\tau(n,k) = \int_0^1 Q_G(u)^n\, u^k\, du$.

The nth incomplete moment of X is determined as

$$m_n(y) = \int_{-\infty}^{y} x^n f(x)\, dx = \sum_{k=0}^{\infty} (k+1)\, b_{k+1} \int_0^{G(y)} Q_G(u)^n\, u^k\, du, \tag{23}$$

where the integral can be computed numerically for most G distributions.

3.5. Generating function

Here, we provide two formulae for the moment generating function (mgf) $M(t) = E(e^{tX})$ of X. Clearly, the first one can be expressed from (13) as

$$M(t) = \sum_{k=0}^{\infty} b_{k+1}\, M_{k+1}(t), \tag{24}$$

where $M_{k+1}(t)$ is the mgf of Y_{k+1}. Hence, $M(t)$ can be determined from the exp-G generating function. A second formula for $M(t)$ can be derived from (13) as

$$M(t) = \sum_{i=0}^{\infty} (k+1)\, b_{k+1}\, \rho(t,k), \tag{25}$$

where $\rho(t,k) = \int_0^1 \exp[t\, Q_G(u)]\, u^k\, du$.

So, we can obtain the mgf's of several distributions directly from Eqs. (24) and (25).

3.6. Mean deviations

The mean deviations about the mean ($\delta_1 = E(|X - \mu_1'|)$) and about the median ($\delta_2 = E(|X - M|)$) of X are given by $\delta_1 = 2\mu_1' F(\mu_1') - 2m_1(\mu_1')$ and $\delta_2 = \mu_1' - 2m_1(M)$, respectively, where $\mu_1' = E(X)$, $M = Median(X) = Q(0.5)$ is the median, $F(\mu_1')$ is easily calculated from (4) and $m_1(z)$ is the first incomplete moment given by (23) with $n = 1$.

Now, we provide two alternative ways to compute δ_1 and δ_2. First, a general equation for $m_1(z)$ can be derived from (13) as

$$m_1(z) = \sum_{k=0}^{\infty} b_{k+1} J_{k+1}(z), \tag{26}$$

where $J_{k+1}(z) = \int_{-\infty}^{z} x\, h_{k+1}(x) dx$ is the basic quantity to compute the first incomplete moment of the exp-G distribution.

A second general formula for $m_1(z)$ can be derived by setting $u = G(x)$ in (13) as

$$m_1(z) = \sum_{k=0}^{\infty} (k+1)\, b_{k+1}\, T_k(z), \tag{27}$$

where $T_k(z) = \int_0^{G(z)} Q_G(u)\, u^k du$ can be computed numerically.

Eqs. (26) and (27) may be applied to construct Bonferroni and Lorenz curves defined for a given probability π by $B(\pi) = m_1(q)/(\pi\mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $\mu_1' = E(X)$ and $q = Q(\pi)$ is the qf of X at π.

3.7. Entropies

An entropy is a measure of variation or uncertainty of a random variable X. Two popular entropy measures are the Rényi and Shannon entropies [27,28]. The Rényi entropy of a random variable with pdf $f(x)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma} \log\left(\int_0^{\infty} f^{\gamma}(x) dx\right),$$

for $\gamma > 0$ and $\gamma \neq 1$. The Shannon entropy of a random variable X is defined by $E\{-\log[f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$. Direct calculation yields

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} = & -\log[ab(1-p)] - \mathrm{E}\{\log[g(X;\xi)]\} \\ & + (1-a)\mathrm{E}\{\log[G(X;\xi)]\} \\ & + (1+a)\mathrm{E}\{\log[1-p\overline{G}(X;\xi)]\} \\ & + (1-b)\mathrm{E}\left\{\log\left\{1-\left[\frac{G(X,\xi)}{1-p\overline{G}(X,\xi)}\right]^a\right\}\right\}. \end{aligned}$$

First, we define

$$\begin{aligned} A(a_1,a_2,a_3;p,a) &= \int_0^1 \frac{u^{a_1}}{[1-p(1-u)]^{a_2}}\left\{1-\left[\frac{u}{1-p(1-u)}\right]^a\right\}^{a_3} du \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{a_3}{i}\int_0^1 \frac{\sum_{k=0}^{\infty}\alpha_{1,k}u^k}{\sum_{k=0}^{\infty}\beta_{1,k}u^k} \\ &= \sum_{i=0}^{\infty}(-1)^i\binom{a_3}{i}\int_0^1 \sum_{k=0}^{\infty}\gamma_{1,k}u^k du \\ &= \sum_{i,k=0}^{\infty}\frac{(-1)^i\binom{a_3}{i}\gamma_{1,k}}{k+1}, \end{aligned} \tag{28}$$

where

$$\alpha_{1,k} = \sum_{j=k}^{\infty}(-1)^{j+k}\binom{a_1+ai}{j}\binom{j}{k},$$

$$\beta_{1,k} = \sum_{j=k}^{\infty}(-1)^{j+k}p^j\binom{a_2+ai}{j}\binom{j}{k} \quad \text{and}$$

$$\gamma_{1,k} = \frac{1}{\beta_{1,0}}\left[\alpha_{1,k} - \frac{1}{\beta_{1,0}}\sum_{r=1}^{k}\alpha_{1,r}\beta_{1,k-r}\right].$$

After some algebraic manipulations, we obtain

Proposition 3. *Let X be a random variable with pdf (5). Then,*

$$\mathrm{E}\{\log[G(X)]\} = ab(1-p)\frac{\partial}{\partial t}A(a+t-1,a+1,b-1;p,a)|_{t=0},$$

$$\mathrm{E}\{\log[1-p\overline{G}(X;\xi)]\} = ab(1-p)\frac{\partial}{\partial t}A(a-1,a+1-t,b-1;p,a)|_{t=0},$$

$$\mathrm{E}\left\{\log\left\{1-\left[\frac{G(X,\xi)}{1-p\overline{G}(X,\xi)}\right]^a\right\}\right\} = ab(1-p)\frac{\partial}{\partial t}A(a-1,a+1,b+t-1;p,a)|_{t=0},$$

where $A(a+t-1,a+1,b-1;p,a)$, $A(a-1,a+1-t,b-1;p,a)$ and $A(a-1,a+1,b+t-1;p,a)$ are defined by Eq. (28).

The simplest formula for the entropy of X is given by

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} = & -\log[ab(1-p)] - \mathrm{E}\{\log[g(X;\xi)]\} \\ & + (1-a)ab(1-p)\frac{\partial}{\partial t}A(a+t-1,a+1,b-1;p,a)|_{t=0} \\ & + (1+a)ab(1-p)\frac{\partial}{\partial t}A(a-1,a+1-t,b-1;p,a)|_{t=0} \\ & + (1-b)ab(1-p)\frac{\partial}{\partial t}A(a-1,a+1,b+t-1;p,a)|_{t=0}. \end{aligned}$$

After some algebraic developments, we obtain an alternative expression for $I_R(\gamma)$ as

$$\begin{aligned} I_R(\gamma) = & \frac{\gamma}{1-\gamma}\log[ab(1-p)] + \frac{1}{1-\gamma} \\ & \times \log\left\{\sum_{i,k=0}^{\infty} w^*_{i,j,k}\mathrm{E}_{Z_k}\{g^{\gamma-1}[Q_G(Z_k)]\}\right\}, \end{aligned} \tag{29}$$

where Z_k is a beta random variable with parameters $k+1$ and one, $w^*_{i,j,k} = \frac{(-1)^i\gamma_{2,k}}{k+1}\binom{(b-1)\gamma}{i}$ and

$$\alpha_{2,k} = \sum_{j=k}^{\infty}(-1)^{j+k}\binom{\gamma(a-1)+ai}{j}\binom{j}{k},$$

$$\beta_{2,k} = \sum_{j=k}^{\infty}(-1)^{j+k}p^j\binom{\gamma(a+1)+ai}{j}\binom{j}{k} \quad \text{and}$$

$$\gamma_{2,k} = \frac{1}{\beta_{2,0}}\left(\alpha_{2,k} - \frac{1}{\beta_{2,0}}\sum_{r=1}^{k}\alpha_{2,r}\beta_{2,k-r}\right).$$

3.8. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Suppose $X_1, \ldots, X_n$ is a random sample from X, then the pdf $f_{i:n}(x)$ of the ith order statistic, say $X_{i:n}$, is given by

$$f_{i:n}(x) = K\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}f(x)\,F(x)^{j+i-1},$$

where $K = n!/[(i-1)!\,(n-i)!]$.

Following Nadarajah et al. [29], the density function of $X_{i:n}$ can be given as

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k}\, h_{r+k+1}(x), \tag{30}$$

where $h_{r+k+1}(x)$ denotes the exp-G density function with parameter $r+k+1$,

$$m_{r,k} = \frac{n!\,(r+1)\,(i-1)!\,b_{r+1}}{(r+k+1)} \sum_{j=0}^{n-i} \frac{(-1)^j f_{j+i-1,k}}{(n-i-j)!\,j!},$$

and b_k is defined in Eq. (12). Here, the quantities $f_{j+i-1,k}$ are obtained recursively from $f_{j+i-1,0} = b_0^{j+i-1}$ and (for $k \geq 1$)

$$f_{j+i-1,k} = (k\,b_0)^{-1} \sum_{m=1}^{k} [m(j+i) - k]\, b_m f_{j+i-1,k-m}.$$

Based on Eq. (30), we can easily obtain ordinary and incomplete moments and generating function of $X_{i:n}$ for any parent G distribution.

4. Characterization of the KwMO family

Characterizations of distributions are important to many researchers in the applied fields. An investigator will be vitally interested to know whether their model fits the requirements of a particular distribution. To this end, one will depend on the characterizations of this distribution which provide conditions under which the underlying distribution is indeed that particular distribution. Various characterizations of distributions have been established in many different directions. In this section, several characterizations of (KwMO) distribution are presented. These characterizations are based on: (i) a simple relationship between two truncated moments; (ii) a single function of the random variable.

4.1. Characterizations based on truncated moments

In this subsection we present characterizations of (KwMO) distribution in terms of a simple relationship between two truncated moments. Our characterization results presented here will employ an interesting result due to Glänzel [30] (Theorem 1, below). The advantage of the characterizations given here is that, *cdf* F need not have a closed form and are given in terms of an integral whose integrand depends on the solution of a first order differential equation, which can serve as a bridge between probability and differential equation.

Theorem 1. *Let* $(\Omega, \Sigma, \mathbf{P})$ *be a given probability space and let* $H = [a, b]$ *be an interval for some* $a < b$ $(a = -\infty$, $b = \infty$*might as well be allowed). Let* $X : \Omega \to H$ *be a continuous random variable with the distribution function F and let* q_1 *and* q_2 *be two real functions defined on H such that*

$$\mathbf{E}[q_1(X)|X \geqslant x] = \mathbf{E}[q_2(X)|X \geqslant x]\eta(x), \quad x \in H,$$

is defined with some real function η. *Assume that* q_1, $q_2 \in C^1(H), \eta \in C^2(H)$ *and G are twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation* $q_2\eta = q_1$ *has no real solution in the interior of H. Then G is uniquely determined by the functions* q_1 , q_2 *and* η*, particularly*

$$G(x) = \int_a^x C\left|\frac{\eta'(u)}{\eta(u)q_2(u) - q_1(u)}\right| \exp(-s(u))du,$$

where the function s is a solution of the differential equation $s' = \frac{\eta' q_2}{\eta q_2 - q_1}$ *and C is a constant, chosen to make* $\int_H dG = 1$.

Clearly, Theorem 1 can be stated in terms of two functions q_1 and η by taking $q_2(x) \equiv 1$, which will reduce the condition given in Theorem 1 to $E[q_1(X) \mid X \geqslant x] = \eta(x)$. However, adding an extra function will give a lot more flexibility, as far as its application is concerned.

Proposition 4. *Let* $X : \Omega \to \mathbb{R}$ *be a continuous random variable and let* $q_2(x) = \left\{1 - \left[\frac{G(x)}{1 - p\overline{G}(x)}\right]^a\right\}^{1-b} [G(x)]^{1-a}$ *and* $q_1(x) = q_2(x)\left[1 - p\overline{G}(x)\right]^{-a}$ *for* $x \in \mathbb{R}$. *The pdf of X is* (5) *if and only if the function* η *defined in Theorem* 1 *has the form*

$$\eta(x) = \frac{1}{2}\left\{1 + \left[1 - p\overline{G}(x)\right]^{-a}\right\}, \quad x \in \mathbb{R}.$$

Proof. Let X have density (5), then

$$(1 - F(x))\mathbf{E}[q_2(X)|X \geqslant x] = \frac{b(1-p)}{p}\left\{\left[1 - p\overline{G}(x)\right]^{-a} - 1\right\}, \quad x \in \mathbb{R},$$

and

$$(1 - F(x))\mathbf{E}[q_1(X)|X \geqslant x] = \frac{b(1-p)}{2p}\left\{\left[1 - p\overline{G}(x)\right]^{-2a} - 1\right\}, \quad x \in \mathbb{R},$$

and finally

$$\eta(x)q_2(x) - q_1(x) = \frac{1}{2}q_2(x)\left\{1 - \left[1 - p\overline{G}(x)\right]^{-a}\right\} \neq 0 \; for \; x \in \mathbb{R}.$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)q_2(x)}{\eta(x)q_2(x) - q_1(x)} = \frac{-apg(x)\left[1 - p\overline{G}(x)\right]^{-(a+1)}}{\left\{1 - \left[1 - p\overline{G}(x)\right]^{-a}\right\}}, \quad x \in \mathbb{R},$$

and hence

$$s(x) = -\ln\left\{1 - \left[1 - p\overline{G}(x)\right]^{-a}\right\}, \quad x \in \mathbb{R}.$$

Now, in view of Theorem 1, X has density function (5). □

Corollary 1. *Let* $X : \Omega \to \mathbb{R}$ *be a continuous random variable and let* $q_2(x)$*be as in Proposition* 4. *The pdf of X is* (5)*if and only if there exist functions* q_1 *and* η *defined in Theorem* 1 *satisfying the differential equation*

$$\frac{\eta'(x)q_2(x)}{\eta(x)q_2(x) - q_1(x)} = \frac{-apg(x)\left[1 - p\overline{G}(x)\right]^{-(a+1)}}{\left\{1 - \left[1 - p\overline{G}(x)\right]^{-a}\right\}}, \quad x \in \mathbb{R}.$$

Remark 1. (*a*) The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \left\{1 - \left[1 - p\overline{G}(x)\right]^{-a}\right\}^{-1} \left[\int apg(x)\left[1 - p\overline{G}(x)\right]^{-(a+1)} [q_2(x)]^{-1} q_1(x)dx + D\right],$$

where D is a constant. One set of appropriate functions is given in Proposition 4 with $D = \frac{1}{2}$.

(*b*) Clearly there are other triplets of functions (q_2, q_1, η) satisfying the conditions of Theorem 1. We presented one such triplet in Proposition 4.

4.2. Characterizations based on single function of the random variable

In this subsection we employ a single function ψ of X and state characterization results in terms of $\psi(X)$. The following proposition has already appeared in [31] Theorem 2.1.3, so we will just state it here for the sake of completeness.

Theorem 2. $1 - F(x) = [c\psi(x) + d]^e$ *if and only if*

$$E[\psi(X)|X \geqslant x] = \frac{1}{e+1}\left[e\psi(x) - \frac{d}{c}\right], \quad x \in (\delta, \varsigma),$$

where $c \neq 0, d, e > 0$ *are finite constants.*

Remark 2. Taking, e.g., $c = -1$, $d = 1$, $e = b$, $\psi(x) = \left[\frac{G(x)}{1-p\overline{G}(x)}\right]^a$ and $(\delta, \varsigma) = \mathbb{R}$, Theorem 2 provides a characterization of the cdf (4).

5. Estimation

Here, we determine the maximum likelihood estimates (MLEs) of the model parameters of the new family from complete samples only. Let $x_1, \ldots, x_n$ be observed values from the KwMO-G distribution with parameters a, b, p and ξ. Let $\Theta = (a, b, p, \xi)^\top$ be the $(r \times 1)$ parameter vector. The total log-likelihood function for Θ is given by

$$\begin{aligned}\ell_n = \ell_n(\Theta) \; &= n \log[ab(1-p)] + \sum_{i=1}^{n} \log[g(x_i;\xi)] \\ &+ (a-1)\sum_{i=1}^{n} \log[G(x_i;\xi)] \\ &- (a+1)\sum_{i=1}^{n} \log\left[1 - p\overline{G}(x_i,\xi)\right] \\ &+ (b-1)\sum_{i=1}^{n} \log\left\{1 - \left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right]^a\right\}. \qquad (31)\end{aligned}$$

The log-likelihood function can be maximized either directly by using the R (AdequecyModel), SAS (PROC NLMIXED) or the Ox program (sub-routine MaxBFGS) [32] or by solving the nonlinear likelihood equations obtained by differentiating (31).

The score function $U_n(\Theta) = (\partial\ell_n/\partial a, \partial\ell_n/\partial b, \partial\ell_n/\partial p, \partial\ell_n/\partial\xi)^\top$ has components given by

$$\frac{\partial\ell_n}{\partial a} = \frac{n}{a} + \sum_{i=1}^{n}\log\left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right] + (1-b)\sum_{i=1}^{n}\frac{\left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right]^a \log\left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right]}{1-\left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right]^a},$$

$$\frac{\partial\ell_n}{\partial b} = \frac{n}{b} + \sum_{i=1}^{n}\log\left\{1-\left[\frac{G(x_i,\xi)}{1-p\overline{G}(x_i,\xi)}\right]^a\right\},$$

$$\frac{\partial\ell_n}{\partial p} = \frac{-n}{1-p} + (a+1)\sum_{i=1}^{n}\frac{\overline{G}(x_i,\xi)}{1-p\overline{G}(x_i,\xi)},$$

$$\begin{aligned}\frac{\partial\ell_n}{\partial\xi} = &\sum_{i=1}^{n}\frac{g^{(\xi)}(x_i,\xi)}{g(x_i,\xi)} + (a-1)\sum_{i=1}^{n}\frac{G^{(\xi)}(x_i,\xi)}{G(x_i,\xi)} + p(a+1)\sum_{i=1}^{n}\frac{G^{(\xi)}(x_i,\xi)}{1-p\overline{G}(x_i,\xi)} \\ &+ a(1-p)\sum_{i=1}^{n}\frac{G^{(\xi)}(x_i,\xi)G(x_i,\xi)^{a-1}}{\left[1-p\overline{G}(x_i,\xi)\right]\left\{\left[1-p\overline{G}(x_i,\xi)\right]^a - G(x_i,\xi)^a\right\}},\end{aligned}$$

where $h^{(\xi)}(\cdot)$ means the derivative of the function h with respect to ξ. The observed information matrix can be obtained from the authors under request.

6. Applications

In this section, we provide two applications to real data to illustrate the importance of the KwMO-W and KwMO-Fr distributions presented in Section 3. The MLEs of the model parameters are computed and goodness-of-fit statistics for these models are compared with other competing models. The first real data set is a subset of data reported by Bekker et al. [33] which corresponds to the survival times (in years) of a group of patients given chemotherapy treatment alone. The data consisting of survival times (in years) for 46 patients are: 0.047, 0.115, 0.121, 0.132, 0.164, 0.197, 0.203, 0.260, 0.282, 0.296, 0.334, 0.395, 0.458, 0.466, 0.501, 0.507, 0.529, 0.534, 0.540, 0.641, 0.644, 0.696, 0.841, 0.863, 1.099, 1.219, 1.271, 1.326, 1.447, 1.485, 1.553, 1.581, 1.589, 2.178, 2.343, 2.416, 2.444, 2.825, 2.830, 3.578, 3.658, 3.743, 3.978, 4.003, 4.033. The second real data set was originally reported by Badar and Priest [34], which represents the strength measured in GPa for single carbon fibers and impregnated at gauge lengths of 1, 10, 20 and 50 mm. Impregnated tows of 100 fibers were tested at gauge lengths of 20, 50, 150 and 300 mm. Here, we consider the data set of single fibers of 20 mm in gauge with a sample of size 63. The data are: 1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

In the first application, we shall compare the KwMO-W model with other comparative models: the beta-Weibull (BW) [35], the Kumaraswamy-Weibull (KwW) [36], the exponentiated-Weibull (EW) [37], the Marshall-Olkin extended Weibull (MOW) [38] and the Weibull (W). In the second application, we compare the KwMO-Fr model with other comparative models: the beta-Fréchet (BFr) [39], the exponentiated-Fréchet (EFr) [40], the Marshall-Olkin extended Fréchet (MOFr) [41] and the Fréchet (Fr). The MLEs are computed using the Limited-Memory Quasi-Newton Code for Bound-Constrained Optimization (L-BFGS-B) as well as the measures of goodness-of-fit including the log-likelihood function evaluated at the MLEs ($\hat{\ell}$). The measures of goodness-of-fit including the Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Bayesian information criterion (BIC), Hannan-Quinn information criterion (HQIC), Anderson–Darling (A^*), Cramér–von Mises (W^*) and Kolmogorov–Smirnov (K–S) statistics are computed to compare the fitted models. The statistics A^* and W^* are described in details in Chen and Balakrishnan [42]. In general, the smaller the values of these statistics, the better the fit to the data. The required computations are carried out using a script of the R-language [43].

The numerical values of the AIC, CAIC, BIC, HQIC, A^*, W^* and K–S statistics are listed in Tables 2 and 4, whereas Tables 3 and 5 list the MLEs and their corresponding standard errors (in parentheses) of the model parameters.

Table 2 The statistics AIC, CAIC, BIC, HQIC, A^*, W^* and K–S for the survival times of cancer patients data.

Distribution	AIC	CAIC	BIC	HQIC	A^*	W^*	K–S	*P*-value
KwMO-W	119.134	120.672	128.167	122.501	0.217	0.027	0.064	0.988
BW	123.995	124.995	131.222	126.689	0.455	0.066	0.098	0.742
KwW	124.189	125.189	131.416	126.884	0.463	0.068	0.099	0.729
EW	122.087	122.673	127.507	124.108	0.470	0.069	0.100	0.717
MOW	121.716	122.301	127.136	123.736	0.444	0.065	0.087	0.858
W	120.247	120.533	123.861	121.594	0.544	0.081	0.109	0.615

Table 3 MLEs and their standard errors (in parentheses) for survival times of cancer patients data.

Distribution	a	b	p	α	β	λ
KwMO-W	0.461	0.167	0.003	–	3.256	0.143
	(0.412)	(0.046)	(0.002)	–	(0.776)	(0.173)
BW	2.045	4.763	–	–	0.691	0.350
	(3.287)	(21.219)	–	–	(0.633)	(1.381)
KwW	5.023	21.029	–	–	0.302	0.723
	(30.532)	(144.337)	–	–	(1.288)	(3.247)
EW	–	–	–	1.612	0.807	1.124
	–	–	–	(1.977)	(0.527)	(1.117)
MO-W	–	–	0.428	–	0.328	1.278
	–	–	(0.182)	–	(0.282)	(0.214)
EE	–	–	–	–	1.053	0.718
	–	–	–	–	(0.124)	(0.125)

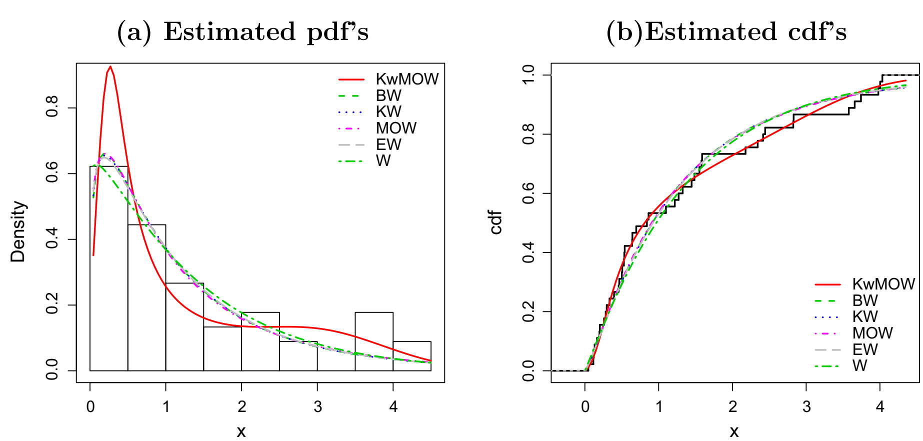

Fig. 3 Plots of the estimated pdf's and cdf's for the KwMO-W, BW, KW, MOW, EW and W models.

Table 4 The statistics AIC, CAIC, BIC, HQIC, A^*, W^* and K–S for single carbon fibers data.

Distribution	AIC	CAIC	BIC	HQIC	A^*	W^*	K–S	*P*-value
KwMO-Fr	121.867	122.920	132.583	126.082	0.236	0.042	0.067	0.941
BFr	120.594	121.283	129.166	123.965	0.324	0.061	0.079	0.822
EFr	118.700	119.107	125.130	121.229	0.330	0.061	0.081	0.803
MOFr	119.746	120.153	126.175	122.275	0.394	0.074	0.083	0.782
Fr	121.804	122.004	126.091	123.490	0.642	0.115	0.100	0.553

Table 5 MLEs and their standard errors (in parentheses) for single carbon fibers data.

Distribution	a	b	p	α	λ	δ
KwMO-Fr	0.053	1.046	0.0203	–	4.066	6.038
	(0.084)	(0.901)	(0.054)	–	(1.941)	(1.612)
BFr	12.547	20.762	–	–	1.167	2.937
	(91.822)	(64.327)	–	–	(1.926)	(10.150)
EFr	–	–	–	7.031	2.364	4.295
	–	–	–	(8.504)	(1.027)	(1.611)
MOFr	–	–	10.343	–	7.906	2.203
	–	–	(12.421)	–	(1.142)	(0.234)
Fr	–	–	–	–	5.433	2.721
	–	–	–	–	(0.508)	(0.067)

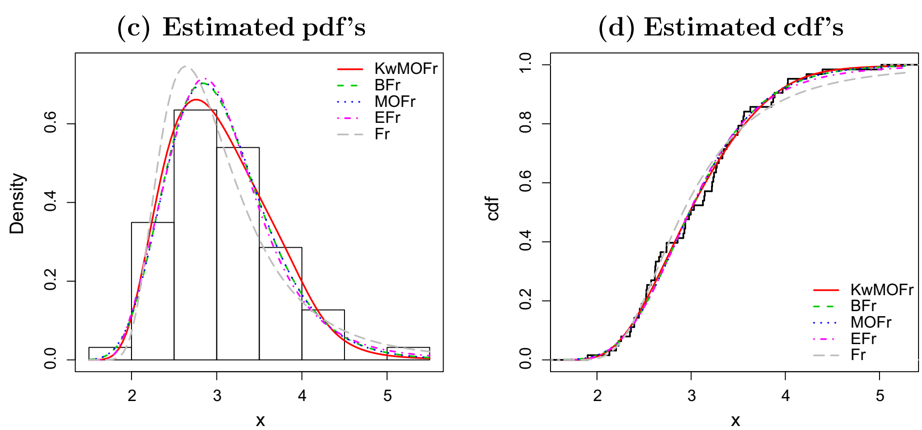

Fig. 4 Plots of the estimated pdf's and cdf's for the KwMO-Fr, BFr, MOFr, EFr and Fr models.

In Table 2, we compare the fits of the KwMO-W model with the BW, KW, MOW, EW and W models. We note that the KwMO-W model gives the lowest values for the AIC, BIC, CAIC, HQIC, A^*, W^* and K–S statistics (for the survival times of cancer patients data) among the fitted models. So, the KwMO-W model could be chosen as the best model. The histogram of the data and the estimated densities and cdfs are displayed in Fig. 3. In Table 4, we compare the fits of the KwMO-Fr model with the BFr, MOFr, EFr and Fr models. We note that the KwMO-Fr model gives the lowest values for the AIC, BIC, CAIC, HQIC, A^*, W^* and K–S statistics (for single carbon fibers data) among all fitted models. So, the KwMO-Fr model can be chosen as the best model. The histogram of the data and the estimated pdfs and cdfs for the fitted models are displayed in Fig. 4. It is very clear from Tables 2 and 4, and Figs. 3 and 4 that the KwMO-W and KwMO-Fr models provide the best fits to the histogram of these data sets.

7. Concluding remarks

In this paper, we propose the new Kumaraswamy Marshall-Olkin family of distributions. We study some of its structural properties including an expansion for the density function and explicit expressions for the moments, generating function, mean deviations, quantile function and order statistics. The maximum likelihood method is employed for estimating the model parameters. We fit two special models of the proposed family to real data sets to demonstrate the usefulness of the new family. These special models provide consistently better fits than other competing models. We hope that the proposed family and its generated models will attract wider applications in several areas such as engineering, survival and lifetime data, hydrology, economics, among others.

Acknowledgments

The authors would like to thank the Editor-in-Chief, the Associate Editor and the two referees for careful reading and for their comments which greatly improved the paper.

References

[1] N. Eugene, C. Lee, F. Famoye, Beta-normal distribution and its applications, Commun. Stat. Theory Methods 31 (2002) 497–512.
[2] M.C. Jones, Families of distributions arising from the distributions of order statistics, Test 13 (2004) 1–43.
[3] G.M. Cordeiro, M. de Castro, A new family of generalized distributions, J. Stat. Comput. Simul. 81 (2011) 883–893.
[4] C. Alexander, G.M. Cordeiro, E.M.M. Ortega, J.M. Sarabia, Generalized beta-generated distributions, Comput. Stat. Data Anal. 56 (2012) 1880–1897.
[5] K. Zografos, N. Balakrishnan, On families of beta- and generalized gamma-generated distributions and associated inference, Stat. Methodol. 6 (2009) 344–362.

[6] M. Amini, S.M.T.K. MirMostafaee, J. Ahmadi, Log-gamma-generated families of distributions, Statistics 48 (2014) 913–932.

[7] M.M. Ristić, N. Balakrishnan, The gamma-exponentiated exponential distribution, J. Stat. Comput. Simul. 82 (2012) 1191–1206.

[8] H. Torabi, N.H. Montazari, The gamma-uniform distribution and its application, Kybernetika 48 (2012) 16–30.

[9] H. Torabi, N.H. Montazari, The logistic-uniform distribution and its application, Commun. Stat. Simul. Comput. 43 (2014) 2551–2569.

[10] A. Alzaatreh, C. Lee, F. Famoye, A new method for generating families of continuous distributions, Metron 71 (2013) 63–79.

[11] A. Alzaatreh, C. Lee, F. Famoye, On the discrete analogues of contineous distributions, Stat. Methodol. 9 (2012) 589–603.

[12] A. Alzaghal, F. Famoye, C. Lee, Exponentiated T-X family of distributions with some applications, Int. J. Probab. Statist. 2 (2013) 31–49.

[13] G.M. Cordeiro, E.M.M. Ortega, D.C.C. da Cunha, The exponentiated generalized class of distributions, J. Data Sci. 11 (2013) 1–27.

[14] M. Bourguignon, R.B. Silva, G.M. Cordeiro, The Weibull-G family of probability distributions, J. Data Sci. 12 (2014) 53–68.

[15] G.M. Cordeiro, M. Alizadeh, E.M.M. Ortega, The exponentiated half-logistic family of distributions: properties and applications, J. Probab. Statist. (2014), 21 p. Art.ID 864396.

[16] M.A. Aljarrah, C. Lee, F. Famoye, On generating T-X family of distributions distributions using quantile functions, J. Stat. Dist. Appl. 1 (2014). Art. 2.

[17] A. Alzaatreh, C. Lee, F. Famoye, T-normal family of distributions: a new approach to generalize the normal distribution, J. Stat. Dist. Appl. 1 (2014). Art. 16.

[18] G.M. Cordeiro, E.M.M. Ortega, B.V. Popović, R.R. Pescim, The Lomax generator of distributions: properties, minification process and regression model, Appl. Math. Comput. 247 (2014) 465–486.

[19] M.H. Tahir, G.M. Cordeiro, A. Alzaatreh, M. Mansoor, M. Zubair, The Logistic-X family of distributions and its applications, Commun. Stat. Theory Methods (2015), in press.

[20] M.H. Tahir, M. Zubair, M. Mansoor, G.M. Cordeiro, M. Alizadeh, G.G. Hamedani, A new Weibull-G family of distributions, Hacet. J. Math. Stat. (2015), in press.

[21] A.N. Marshall, I. Olkin, A new method for adding a parameter to a family of distributions with applications to the exponential and Weibull families, Biometrika 84 (1997) 641–652.

[22] G.M. Cordeiro, M. Alizadeh, C.R.B. Dias, P.R.D. Marinho, Exponentiated Marshall-Olkin family of distributions, Appl. Math. Model. (2014), in preparation.

[23] R.C. Gupta, P.I. Gupta, R.D. Gupta, Modeling failure time data by Lehmann alternatives, Commun. Stat. Theory Methods 27 (1998) 887–904.

[24] R.C. Gupta, R.D. Gupta, Proportional reversed hazard rate model and its applications, J. Stat. Plann. Infer. 137 (2007) 3525–3536.

[25] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, sixth ed., Academic Press, San Diego, 2000.

[26] S. Nadarajah, S. Kotz, The exponentiated type distributions, Acta Appl. Math. 92 (2006) 97–111.

[27] A. Rényi, On Measures of Entropy and Information, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability-I, University of California Press, Berkeley, 1961, pp. 547–561.

[28] C.E. Shannon, A mathematical theory of communication, Bell Sys. Tech. J. 27 (1948) 379–432.

[29] S. Nadarajah, G.M. Cordeiro, E.M.M. Ortega, The Zografos-Balakrishnan–G family of distributions: mathematical properties and applications, Commun. Stat. Theory Methods 44 (2015) 186–215.

[30] W. Glänzel, A characterization theorem based on truncated moments and its application to some distribution families, in: M.L. Puri, P. Révész, W. Wertz, P. Bauer, F. Konecny (Eds.), Mathematical Statistics and Probability Theory, vol. B, D. Reidel Publishing Company, 1987, pp. 75–84.

[31] G.G. Hamedani, Z. Javanshiri, M. Maadooliat, A. Yazdani, Remarks on characterizations of Malinowska and Szynal, Appl. Math. Comput. 246 (2014) 377–388.

[32] J.A. Doornik, Ox 5: An Object-Oriented Matrix Programming Language, fifth ed., Timberlake Consultants, London, 2007.

[33] A. Bekker, J. Roux, P. Mostert, A generalization of the compound Rayleigh distribution: using a Bayesian methods on cancer survival times, Commun. Stat. Theory Methods 29 (2000) 1419–1433.

[34] M.G. Badar, A.M. Priest, Statistical aspects of fiber and bundle strength in hybrid composites, in: T. Hayashi, K. Kawata, S. Umekawa (Eds.), Progress in Science and Engineering Composites, ICCM-IV, Tokyo, 1982, pp. 1129–1136.

[35] F. Famoye, C. Lee, O. Olumolade, The beta-Weibull distribution, J. Stat. Theory Appl. 4 (2005) 121–136.

[36] G.M. Cordeiro, E.M.M. Ortega, S. Nadarajah, The Kumaraswamy Weibull distribution with application to failure data, J. Franklin Inst. 347 (2010) 1174–1197.

[37] G.S. Mudholkar, D.K. Srivastava, Exponentiated Weibull family for analyzing bathtub failure data, IEEE Trans. Reliab. 42 (1993) 299–302.

[38] G.M. Cordeiro, A.J. Lemonte, On the Marshall-Olkin extended Weibull distribution, Stat. Papers 54 (2013) 333–353.

[39] W. Barreto-Souza, G.M. Cordeiro, A. Simas, Some results for beta Fréchet distribution, Commun. Stat. Theory Methods 40 (2011) 798–811.

[40] S. Nadarajah, S. Kotz, The exponentiated Fréchet distribution, InterStat, 2003, <interstat.statjournals.net/YEAR/2003/articles/0312001.pdf>.

[41] E. Krishna, K.K. Jose, M.M. Ristić, Applications of Marshall-Olkin Fréchet distribution, Commun. Stat. Theory Methods 42 (2013) 76–89.

[42] G. Chen, N. Balakrishnan, A general purpose approximate goodness-of-fit test, J. Qual. Tech. 27 (1995) 154–161.

[43] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012.