

# Harmonic Mixture-G Family of Distributions: Survival Regression, Simulation by Likelihood, Bootstrap and Bayesian Discussion with MCMC Algorithm

**Omid Kharazmi**  
Vali-e-Asr University

**Ali Saadati Nik**  
University  
of Mazandaran

**GG. Hamedani**  
Marquette University

**Emrah Altun**  
Bartın University

---

## Abstract

To study the heterogeneous nature of lifetimes of certain mechanical or engineering processes, a mixture model of some suitable lifetime distributions may be more appropriate and appealing than simpler models. In this paper, a new mixture family of the lifetime distributions is introduced via harmonic weighted mean of an underlying distribution and the distribution of the proportional hazard model corresponding to the baseline model. The proposed class of distributions includes the general Marshall-Olkin family of distributions as a special case. Some important properties of the proposed model such as survival function, hazard function, order statistics and some results on stochastic ordering are obtained in a general setting. A special case of this new family is considered by employing Weibull distribution as the parent distribution. We derive several properties of the special distribution such as moments, hazard function survival regression and certain characterizations results. Moreover, we estimate the parameters of the model by using frequentist and Bayesian approaches. For Bayesian analysis, five loss functions, namely the squared error loss function (SELF), weighted squared error loss function (WSELF), modified squared error loss function (MSELF), precautionary loss function (PLF), and K-loss function (KLF) are considered. The beta prior as well as the gamma prior are used to obtain the Bayes estimators and posterior risk of the unknown parameters of the model. Furthermore, credible intervals (CIs) and highest posterior density (HPD) intervals are also obtained. A simulation study is presented via Monte Carlo to investigate the bias and mean square error of the maximum likelihood estimators. For illustrative purposes, two real-life applications of the proposed distribution to Kidney and cancer patients are provided.

*Keywords:* Bayes estimators, credible intervals, loss functions, mixture distribution, posterior risks, survival regression.

---

## 1. Introduction

In recent years, extensive efforts have been made to present new models in the area of distribution theory and related statistical applications. These studies deal mainly with the modeling of various data sources and finding out the probabilistic structure of the model. In connec-

tion with the development of new models, it is worthwhile to note that these new models should have the capability for analyzing a wide range of real observations. Undoubtedly, this is also the most basic concern in the development of the new models from the past to the future. Mixture models are quite versatile and thus have been frequently used in different fields of sciences such as reliability, information theory, economic, engineering, agriculture, to name a few. For example, [Mendenhall and Hader \(1958\)](#), while referring to the practical situations encountered by engineers, pointed out that the failure of a system or a device may be divided into two or more different types of causes. Further, [Acheson and McElwee \(1952\)](#) categorized the failures of the electronic tube into gaseous defects, mechanical defects, and normal deterioration of the cathode in order to find the proportion of failure due to a certain cause. Another example is that of an engineering system which consists of different subsystems. These subsystems may be homogeneous or heterogeneous. Heterogeneity nature of such systems can not be captured by a single probability model but it can be captured through mixture models.

In the following, three well-known mixture distributions are reviewed. The usual arithmetic mixture distribution is defined by the weighted mean of two probability density functions (PDFs):

$$h(x) = \theta f(x) + (1 - \theta)g(x), \quad 0 \leq \theta \leq 1. \quad (1)$$

The geometric mixture is defined by the normalized geometric mean of two PDFs:

$$h(x) = \frac{f^\theta(x)g^{1-\theta}(x)}{k}, \quad k = \int f^\theta(x)g^{1-\theta}(x)dx, \quad 0 \leq \theta \leq 1, \quad (2)$$

provided that the normalizing factor  $k < \infty$  on the support of  $h$ . This model has been derived in statistics, and in physics in different contexts. [Bercher and Vignat \(2009\)](#) called (2) the generalized escort probability distribution. A more general mixture distribution, called here as the power mean mixture, is defined by the following PDF:

$$h(x) = \frac{[\theta f^\alpha(x) + (1 - \theta)g^\alpha(x)]^{\frac{1}{\alpha}}}{k}, \quad 0 \leq \theta \leq 1, \alpha > 0, \quad (3)$$

where

$$k = \int [\theta f^\alpha(x) + (1 - \theta)g^\alpha(x)]^{\frac{1}{\alpha}} dx.$$

The normalizing factor always exists for  $\alpha > 0$  and

$$k \in \begin{cases} [2^{-(1-\alpha)/\alpha}, 1], & 0 < \alpha \leq 1 \\ [1, 2^{(\alpha-1)/\alpha}], & \alpha \geq 1. \end{cases}$$

The  $\alpha$ -mixture family (3) contains (1) as a special case with  $\alpha = 1$  and (2) as a limiting case for  $\alpha \rightarrow 0$ . [Asadi, Ebrahimi, Kharazmi, and Soofi \(2018\)](#) obtained some interesting informational properties for the above mixture models. The above three mixture distributions play important roles in the information theory and shown that they contain optimal information. For more details see [Asadi et al. \(2018\)](#).

Motivated by optimal information properties of these mixture distributions, we propose a new lifetime distribution based on the harmonic mean mixture of two survival functions. The proposed family of distributions is called harmonic mixture  $G$  ( $HMG$ ) distribution. The new  $HMG$  distribution includes Marshall-Olkin family of distributions introduced by Marshall and Olkin (1997).

Our main motivation to introduce this new category of distributions is to provide more flexibility for fitting real data sets compare to the other well-known classic statistical distributions. We first derive certain statistical and reliability properties of the  $HMG$  distribution in a general setting and then we consider a special case of this model by employing the Weibull distribution as the parent distribution  $G$ . It is called  $HMW$  distribution. We provide a

comprehensive discussion about the survival regression of the new *HMW* model. Furthermore, we consider the Maximum likelihood, bootstrap estimation and Bayesian procedures to estimate the unknown parameters of the new model for the real application. In addition, the asymptotic confidence intervals and parametric and non-parametric bootstrap confidence intervals are calculated.

In the Bayesian discussion, we consider different types of symmetric and asymmetric loss functions such as squared error loss, weighted squared error, modified squared error, precautionary, *K*-loss, linear exponential and general entropy loss functions to estimate the four unknown parameters of *HMG* model. Since all of the parameters are positive, we use gamma and beta prior distributions. Bayesian 95% credible and highest posterior density (*HPD*) intervals (see [Chen and Shao \(1999\)](#)) are provided for each parameter of the proposed model. In addition, the asymptotic confidence intervals and parametric and non-parametric bootstrap confidence intervals are calculated for comparison with the corresponding Bayesian intervals. In addition, a simulation study is performed to investigate *MLEs* of consistency.

The rest of the manuscript is organized as follows. In Section 2, we introduce a new class of distributions called harmonic mixture-*G* (*HMG*) distributions and consider the hazard function, quantiles and a discussion about informational properties in a general setting. In Section 3, we consider the Weibull as the parent distribution and introduce generalized harmonic mixture Weibull distribution. This new model is referred to as *HMW* distribution. Also, in this section, we plot the density function, hazard function and 3D plots of skewness and kurtosis for the *HMW* distribution. Some characterization results for the *HMW* distribution are provided in Section 4. In Section 5, the estimation of the parameters of the *HMW* distribution are obtained via three methods: maximum likelihood, Bayesian and bootstrap estimations. Bayesian analysis results are provided by considering five well-known loss functions. Also, we study the performance of the maximum likelihood estimates of the parameters of *HMW* distribution via a simulation study. Section 6 is devoted to a discussion about survival regression of *HMW* model. In Section 7, the superiority of the new model to some competitor statistical models is shown through different criteria of selection model by analyzing a real example. Moreover, Bayesian analysis and associated plots of the posterior samples corresponding to this data are provided in this Section. Also, in this Section we provide numerical analysis for survival regression of *HMW* distribution via a Kidney data set. Finally, the paper is concluded in Section 8.

## 2. Harmonic mixture-*G* (*HMG*) family of distributions

Let  $X$  be a continuous random variable with survival and hazard functions  $\bar{G}(x)$  and  $r(x)$ , respectively. The model with hazard function  $r_{PH} = \alpha r(x); \alpha > 0$ , is known as proportional hazards (*PH*) model, popularized by [Cox \(1972\)](#). The survival function of *PH* model corresponding to the baseline survival function  $\bar{G}(x)$ , is given as  $\bar{G}^\alpha(x)$ . This model is widely used in various applications in many fields especially in survival analysis and economics. In this section, first we propose a new model based on the harmonic mixture mean of two survival functions  $\bar{G}(x)$  and  $\bar{G}^\alpha(x)$  and then study some of its main properties in a general setting.

$$\bar{H}(x) = \frac{1}{\frac{\theta}{\bar{G}(x)} + \frac{1-\theta}{\bar{G}^\alpha(x)}} = \frac{\bar{G}^\alpha(x)}{1 - \theta(1 - \bar{G}^{\alpha-1}(x))}, \quad x \in \mathbb{R}, \quad \alpha \geq 0, \quad 0 < \theta < 1.$$

**Lemma 1.** *If  $\alpha = 0$ , the general Marshall-Olkin family of distribution is obtained as*

$$\bar{H}(x) = \frac{\bar{G}(x)}{1 - (1 - \theta)\bar{G}(x)},$$

where  $G = 1 - \bar{G}$ .

The associated *CDF* and *PDF* are given, respectively, by

$$H(x) = 1 - \frac{\bar{G}^\alpha(x)}{1 - \theta(1 - \bar{G}^{\alpha-1}(x))}; \quad x \in \mathbb{R}, \quad \alpha \geq 0, \quad 0 < \theta < 1, \quad (4)$$

and

$$h(x) = g(x) \bar{G}^{\alpha-1}(x) \frac{\alpha(1 - \theta) + \theta \bar{G}^{\alpha-1}(x)}{\left(1 - \theta(1 - \bar{G}^{\alpha-1}(x))\right)^2}, \quad x \in \mathbb{R}. \quad (5)$$

## 2.1. Hazard rate function

Here, first we obtain the hazard rate function of *HMG* model and then two associated lemmas are given. The hazard rate is a fundamental tool in reliability modeling for evaluation of the aging process. Knowing the shape of the hazard rate is important in the reliability theory, risk analysis and other disciplines. The concepts of increasing, decreasing, bathtub shaped (first decreasing and then increasing) and upside down bathtub shaped (first increasing and then decreasing) hazard rate functions are very useful in the reliability analysis. The lifetime distributions with these aging properties are designated as the IFR, DFR, BUT and UBT distributions, respectively. The hazard function of the *HMG* distribution is

$$r_G(x) = \frac{h(x)}{\bar{H}(x)} = \frac{g(x)}{\bar{G}(x)} \frac{\alpha(1 - \theta) + \theta \bar{G}^{\alpha-1}(x)}{1 - \theta(1 - \bar{G}^{\alpha-1}(x))}. \quad (6)$$

In fact the hazard rate function of the new model is a weighted version of the baseline hazard  $r(x) = \frac{g(x)}{\bar{G}(x)}$  with the weight  $w(x) = \frac{\alpha(1-\theta) + \theta \bar{G}^{\alpha-1}(x)}{1 - \theta(1 - \bar{G}^{\alpha-1}(x))}$ .

**Lemma 2.** *In view of (6), we have:*

- *If  $r(x)$  is increasing and  $\alpha \geq 1$  then  $r_G(x)$  is increasing.*
- *If  $r(x)$  is decreasing and  $0 < \alpha \leq 1$  then  $r_G(x)$  is decreasing.*

*Proof.* The proof is straightforward. □

In the following lemma we provide a result regarding the stochastic ordering of the hazard function to the compare proposed model and baseline distribution. First, we recall the following definition. The random variable  $X$  is said to be less than the random variable  $Y$  in the hazard rate order,  $X \leq_{hr} Y$ , if  $h_X(x) \geq h_Y(x)$ , for all  $x$  in the union of supports of  $X$  and  $Y$ , where  $h_X(x)(h_Y(x))$  is the hazard rate of  $X(Y)$ . For more details see [Shaked and Shanthikumar \(2007\)](#).

**Lemma 3.** *Suppose that  $X$  is a random variable with the baseline distribution  $G$  and  $X_G$  is harmonic mixture random variable associated with  $X$ , and have density function (1).*

- (i) *If  $0 < \alpha \leq 1$ , then  $X \leq_{hr} X_G$ .*
- (ii) *If  $\alpha \geq 1$ , then  $X_G \leq_{hr} X$ .*

*Proof.* The proof is straightforward. □

## 3. Harmonic mixture-Weibull (*HMW*) distribution

In this section, we specialize *HMG* distribution, which is described in the previous Section, by choosing special case for baseline distribution  $G$ . We apply the *HMG* method to a specific case of baseline distribution, namely to a Weibull distribution and call this special model, four-parameter *HMW* distribution.

**Definition 1.** A random variable  $X$  has  $HMW(\alpha, \beta, \lambda, \theta)$  distribution, if its  $PDF$  is given by

$$h(x) = \beta \lambda x^{\beta-1} e^{-\alpha \lambda x^\beta} \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{\left(1 - \theta(1 - e^{-(\alpha-1)\lambda x^\beta})\right)^2}, \quad (7)$$

where,  $x > 0$ ,  $\alpha, \beta, \lambda > 0$  and  $0 < \theta < 1$ .

The CDF corresponding to (7) is

$$H(x) = 1 - \frac{e^{-\alpha \lambda x^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x^\beta})}, \quad x > 0.$$

The survival and hazard rate functions are

$$\bar{H}(x) = \frac{e^{-\alpha \lambda x^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x^\beta})},$$

and

$$r(x) = \beta \lambda x^{\beta-1} \frac{\alpha(1-\theta) + \theta e^{-\lambda(\alpha-1)x^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x^\beta})},$$

respectively. Some plots of  $PDF$  and hazard function for the selected parameter values are given in Figures 1 and 2.

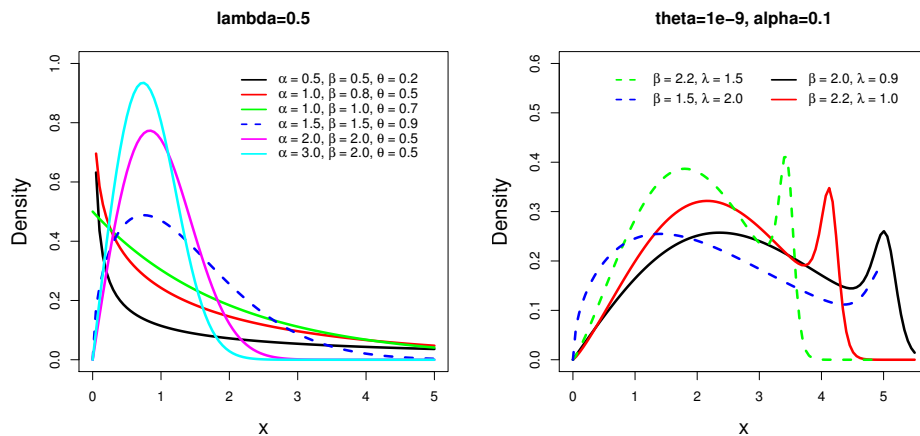


Figure 1: Plots of the  $PDF$  of  $HMW$  for selected parameter values.

### 3.1. Some properties of the $HMW$ distribution

In this section, we obtain some properties of the  $HMW$  distribution, such as the moments and order statistics distribution.

### 3.2. Moments

In this subsection, moments and related measures including coefficients of variation, skewness and kurtosis are presented. Tables of values for the first six moments, standard deviation ( $SD$ ), coefficient of variation ( $CV$ ), coefficient of skewness ( $CS$ ) and coefficient of kurtosis ( $CK$ ) are also presented. The  $r$ th moment of the  $HMW$  distribution, denoted by  $\mu'_r$ , is

$$\mu'_r = E(X^r) = \beta \lambda \int_0^\infty x^{r+\beta-1} e^{-\alpha \lambda x^\beta} \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{\left(1 - \theta(1 - e^{-(\alpha-1)\lambda x^\beta})\right)^2} dx.$$

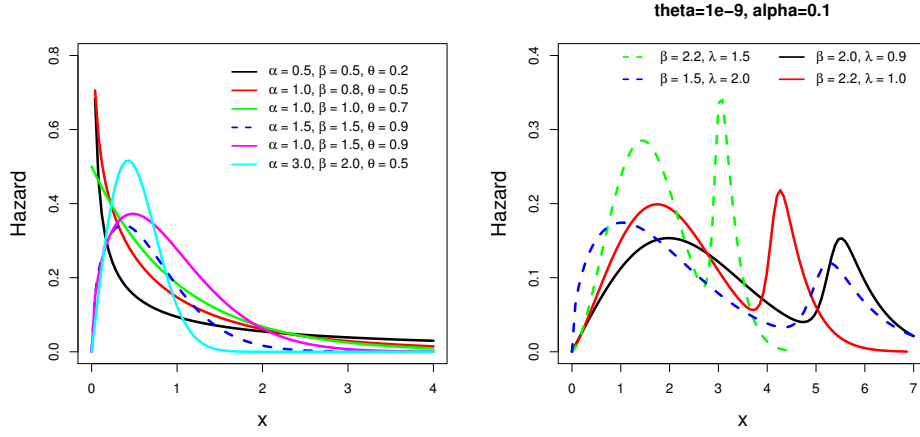


Figure 2: Plots of the failure rate function of  $HMW$  for selected parameter values.

The variance,  $CV$ ,  $CS$ , and  $CK$  are given by

$$\sigma^2 = \mu'_2 - \mu^2, \quad CV = \frac{\sigma}{\mu} = \frac{\sqrt{\mu'_2 - \mu^2}}{\mu} = \sqrt{\frac{\mu'_2}{\mu^2} - 1}, \quad (8)$$

$$CS = \frac{E[(X - \mu)^3]}{[E(X - \mu)^2]^{3/2}} = \frac{\mu'_3 - 3\mu\mu'_2 + 2\mu^3}{(\mu'_2 - \mu^2)^{3/2}}, \quad (9)$$

and

$$CK = \frac{E[(X - \mu)^4]}{[E(X - \mu)^2]^2} = \frac{\mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4}{(\mu'_2 - \mu^2)^2}, \quad (10)$$

respectively.

In order to investigate and analyze the amount of skewness and kurtosis of the new model under the three parameters  $\alpha$ ,  $\lambda$ ,  $\beta$  and  $\theta$ , 3D diagrams are presented in Figures 3-8. These plots indicate that the proposed model can be left skewed, right skewed and symmetrical. Further, it is evident from these plots that all four parameters have a considerable effect on the skewness and Kurtosis of the HMW model. Variety of skewness and Kurtosis shapes make the new model to be useful in modeling various shapes of the real data sets.

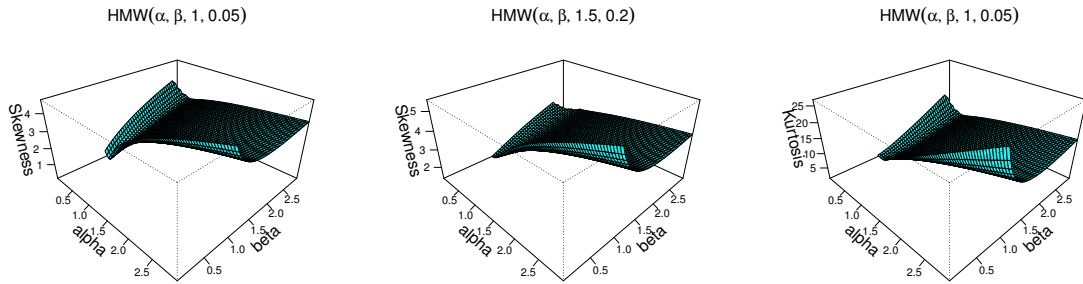


Figure 3: 3D plots of skewness and kurtosis of  $HMW$  distribution for some fixed parameter values  $\lambda$  and  $\theta$ .

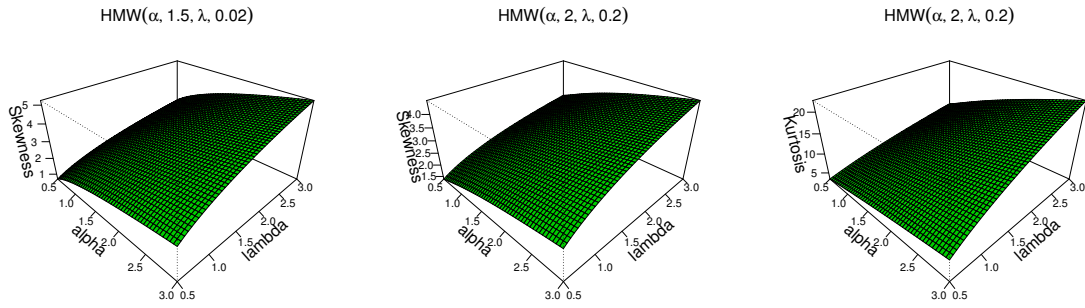


Figure 4: 3D plots of skewness and kurtosis of *HMW* distribution for some fixed parameter values  $\beta$  and  $\theta$ .

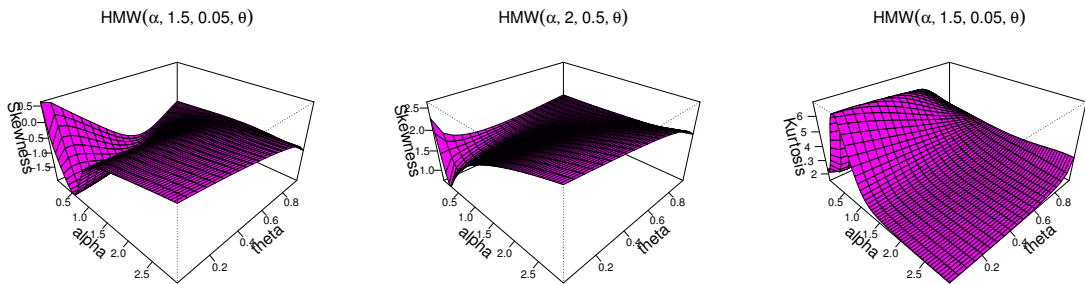


Figure 5: 3D plots of skewness and kurtosis of *HMW* distribution for some fixed parameter values  $\beta$  and  $\lambda$ .

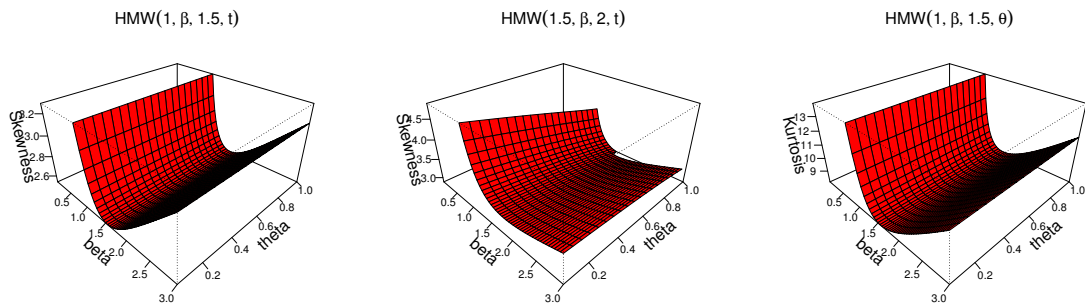


Figure 6: 3D plots of skewness and kurtosis of *HMW* distribution for some fixed parameter values  $\alpha$  and  $\lambda$ .

### 3.3. Order statistics

Order statistics play an important role in probability and statistics. In this subsection, we present the distribution of the  $i$ th order statistic from the *HMW* distribution. The *PDF* of the  $i$ th order statistic from the *HMW PDF*,  $f_{HMW}(x)$ , is given by

$$\begin{aligned}
 f_{i:n}(x) &= \frac{n!}{(i-1)!(n-i)!} f_{HMW}(x) [F_{HMW}(x)]^{i-1} [1 - F_{HMW}(x)]^{n-i} \\
 &= \frac{n!}{(i-1)!(n-i)!} f_{HMW}(x) \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^m [F_{HMW}(x)]^{m+i-1}.
 \end{aligned}$$



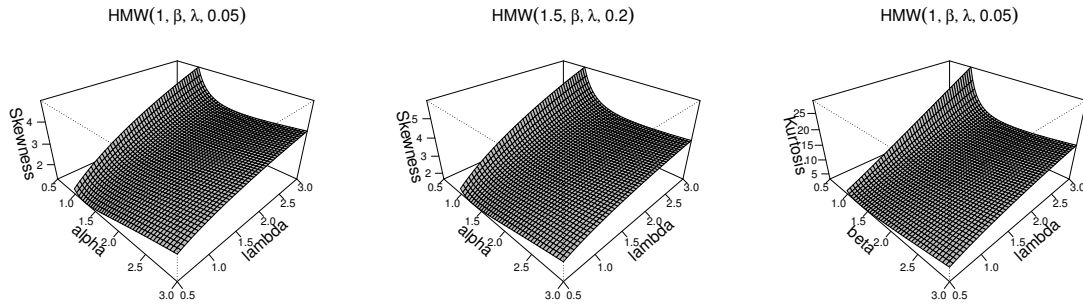


Figure 7: 3D plots of skewness and kurtosis of  $HMW$  distribution for some fixed parameter values  $\alpha$  and  $\theta$ .

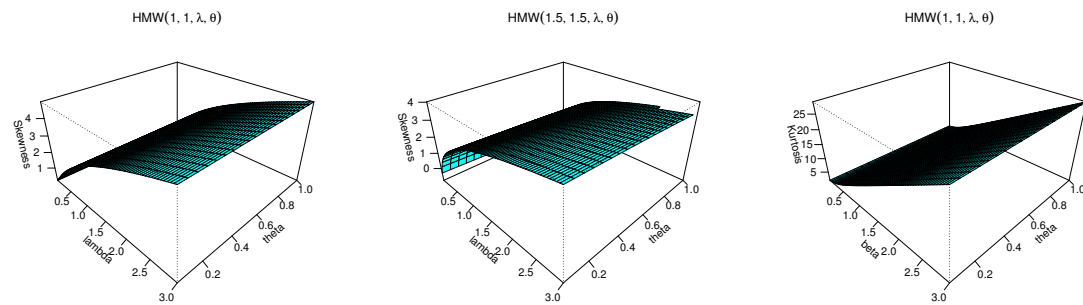


Figure 8: 3D plots of skewness and kurtosis of  $HMW$  distribution for some fixed parameter values  $\alpha$  and  $\beta$ .

Using the binomial expansion  $[1 - F_{HMW}(x)]^{n-i} = \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^m [F_{HMW}(x)]^m$ , we have

$$f_{i:n}(x) = \frac{1}{B(i, n-i+1)} \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^m [F_{HMW}(x)]^{m+i-1} f_{HMW}(x)$$

## 4. Characterization results

In this section we establish certain characterizations of the HMW distribution in three directions: (i) based on two truncated moments; (ii) in terms of the hazard function and (iii) based on the conditional expectation of a function of the random variable. These characterizations will be presented in three subsections.

### 4.1. Characterizations based on two truncated moments

This subsection deals with the characterizations of HMW distribution in terms of a simple relationship between two truncated moments. We will employ Theorem 1 of Glänzel (1987) [1] given in the Appendix A. As shown in [2], this characterization is stable in the sense of weak convergence.

**Proposition 4.1.1.** Suppose  $X$  is a continuous random Let  $q_1(x) = \left( \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{(1-\theta)(1-e^{-\alpha\lambda x^\beta})} \right)^{-1}$  and  $q_2(x) = q_1(x) e^{-\alpha\lambda x^\beta}$  for  $x > 0$ . Then  $X$  has density (7) if and only if the function  $\xi$



defined in Theorem 1 is of the form

$$\xi(x) = \frac{1}{2}e^{-\alpha\lambda x^\beta}, \quad x > 0.$$

*Proof.* If  $X$  has density (7), then

$$(1 - F(x; \alpha, \beta, \lambda, \theta)) E[q_1(X) | X \geq x] = \frac{1}{\alpha}e^{-\alpha\lambda x^\beta}, \quad x > 0,$$

and

$$(1 - F(x; \alpha, \beta, \lambda, \theta)) E[q_2(X) | X \geq x] = \frac{1}{2\alpha}e^{-2\alpha\lambda x^\beta}, \quad x > 0,$$

and hence

$$\xi(x)q_1(x) - q_2(x) = -\frac{1}{2}q_1(x)e^{-\alpha\lambda x^\beta} < 0, \quad \text{for } x > 0.$$

Conversely, if  $\eta$  is of the above form, then

$$s'(x) = \frac{\xi'(x)q_1(x)}{\xi(x)q_1(x) - q_2(x)} = \alpha\lambda\beta x^{\beta-1}, \quad x > 0,$$

and

$$s(x) = \alpha\lambda x^\beta, \quad x > 0.$$

Now, according to Theorem 1,  $X$  has PDF (7). □

**Corollary 4.1.1.** *Suppose  $X$  is a continuous random variable. Let  $q_1(x)$  be as in Proposition 4.1.1. Then  $X$  has density (7) if and only if there exist functions  $q_2$  and  $\xi$  defined in Theorem 1 for which the following first order differential equation holds*

$$\frac{\xi'(x)q_1(x)}{\xi(x)q_1(x) - q_2(x)} = \alpha\lambda\beta x^{\beta-1}, \quad x > 0.$$

**Corollary 4.1.2.** *The differential equation in Corollary 4.1.1 has the following general solution*

$$\xi(x) = e^{\alpha\lambda x^\beta} \left[ - \int \alpha\lambda\beta x^{\beta-1} e^{-\alpha\lambda x^\beta} (q_1(x))^{-1} q_2(x) dx + D \right],$$

where  $D$  is a constant. A set of functions satisfying the above differential equation is given in Proposition 4.1.1 with  $D = 0$ . Clearly, there are other triplets  $(q_1, q_2, \xi)$  satisfying the conditions of Theorem 1.

## 4.2. Characterization based on hazard function

The hazard function,  $h_F$ , of a twice differentiable distribution function,  $F$ , satisfies the following differential equation

$$\frac{f'(x)}{f(x)} = \frac{h'_F(x)}{h_F(x)} - h_F(x).$$

The following proposition establishes a non-trivial characterization HMW distribution based on the hazard function.

**Proposition 4.2.1.** *Suppose  $X$  is a continuous random variable. Then,  $X$  has density (7) if and only if its hazard function  $h_F(x)$  satisfies the following first order differential equation*

$$h'_F(x) - \frac{\beta-1}{x}h_F(x) = \lambda\beta x^{\beta-1} \frac{d}{dx} \left\{ \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{1-\theta(1-e^{-(\alpha-1)\lambda x^\beta})} \right\}, \quad x > 0,$$

with the initial condition  $\lim_{x \rightarrow 0} h_F(x) = 0$  for  $\beta > 1$ .

*Proof.* If  $X$  has density (7), then clearly the above differential equation holds. Now, if the differential equation holds, then

$$\frac{d}{dx} \left\{ x^{-\beta+1} h_F(x) \right\} = \lambda \beta \frac{d}{dx} \left\{ \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{1-\theta(1-e^{-(\alpha-1)\lambda x^\beta})} \right\},$$

or

$$h_F(x) = \lambda \beta x^{\beta-1} \left\{ \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x^\beta}}{1-\theta(1-e^{-(\alpha-1)\lambda x^\beta})} \right\}, \quad x > 0,$$

which is the hazard function corresponding to the PDF (7).  $\square$

### 4.3. Characterizations based on the conditional expectation of a function of the random variable

Hamedani (2013) [3] established the following proposition which can be used to characterize the HMW distribution.

**Proposition 4.3.1.** *Suppose  $X : \Omega \rightarrow (a, b)$  is a continuous random variable with cdf  $F$ . If  $\psi(x)$  is a differentiable function on  $(a, b)$  with  $\lim_{x \rightarrow a^+} \psi(x) = 1$ . Then, for  $\delta \neq 1$ ,*

$$E[\psi(X) \mid X \geq x] = \delta \psi(x), \quad x \in (a, b),$$

if and only if

$$\psi(x) = (1 - F(x))^{\frac{1}{\delta} - 1}, \quad x \in (a, b).$$

**Remark 4.3.1.** *Let  $(a, b) = (0, \infty)$ ,  $\psi(x) = \left\{ \frac{e^{-x^\beta}}{[1-\theta(1-e^{-(\alpha-1)\lambda x^\beta})]^{1/\alpha\lambda}} \right\}$  and  $\delta = \frac{\alpha\lambda}{\alpha\lambda+1}$ , then*

*Proposition 4.3.1 presents a characterization of HMW distribution. Clearly, there are other suitable functions than the one we employed for simplicity.*

## 5. Inference procedure

In this section, we consider estimation of the unknown parameters of the  $HMW(\alpha, \beta, \lambda, \theta)$  distribution via three methods: maximum likelihood method, bootstrap estimation and Bayesian procedure.

### 5.1. Maximum likelihood estimation

Let  $x_1, \dots, x_n$  be a random sample from the  $HMW$  distribution and  $\Delta = (\alpha, \beta, \lambda, \theta)$  be the vector of parameters. The log-likelihood function is given by

$$\begin{aligned} L = L(\Delta) = & n \log \beta + n \log \lambda + (\beta - 1) \sum_{i=1}^n \log x_i - \alpha \lambda \sum_{i=1}^n x_i^\beta \\ & + \sum_{i=1}^n \log \left( \alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta} \right) - 2 \sum_{i=1}^n \log \left( 1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta}) \right). \end{aligned} \quad (11)$$

The elements of the score vector are given by

$$\frac{dL}{d\alpha} = -\lambda \sum_{i=1}^n x_i^\beta + \sum_{i=1}^n \frac{(1-\theta) - \lambda \theta x_i^\beta e^{-(\alpha-1)\lambda x_i^\beta}}{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}} - 2 \sum_{i=1}^n \frac{\lambda \theta x_i^\beta e^{-(\alpha-1)\lambda x_i^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta})} = 0,$$

$$\begin{aligned} \frac{dL}{d\beta} &= \frac{n}{\beta} + \sum_{i=1}^n \log x_i - \alpha \lambda \sum_{i=1}^n x_i^\beta \log x_i - \sum_{i=1}^n \frac{\theta \lambda (\alpha - 1) x_i^\beta \log x_i e^{-(\alpha-1)\lambda x_i^\beta}}{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}} \\ &\quad + 2 \sum_{i=1}^n \frac{\theta \lambda (\alpha - 1) x_i^\beta \log x_i e^{-(\alpha-1)\lambda x_i^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta})} = 0, \end{aligned}$$

$$\begin{aligned} \frac{dL}{d\lambda} &= \frac{n}{\lambda} - \alpha \sum_{i=1}^n x_i^\beta - \sum_{i=1}^n \frac{(\alpha - 1) \theta x_i^\beta e^{-(\alpha-1)\lambda x_i^\beta}}{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}} \\ &\quad + 2 \sum_{i=1}^n \frac{(\alpha - 1) \theta x_i^\beta e^{-(\alpha-1)\lambda x_i^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta})} = 0, \end{aligned}$$

and

$$\frac{dL}{d\theta} = \sum_{i=1}^n \frac{e^{-(\alpha-1)\lambda x_i^\beta} - \alpha}{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}} - 2 \sum_{i=1}^n \frac{e^{-(\alpha-1)\lambda x_i^\beta}}{1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta})} = 0,$$

respectively.

The maximum likelihood estimate,  $\hat{\Delta}$  of  $\Delta = (\alpha, \beta, \lambda, \theta)$  is obtained by solving the nonlinear equations  $\frac{dL}{d\alpha} = 0$ ,  $\frac{dL}{d\beta} = 0$ ,  $\frac{dL}{d\lambda} = 0$ ,  $\frac{dL}{d\theta} = 0$ . These equations are not in closed form and the values of the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$  must be found using iterative methods. Therefore, the maximum likelihood estimate,  $\hat{\Delta}$  of  $\Delta = (\alpha, \beta, \lambda, \theta)$  can be determined using an iterative method such as the Newton-Raphson procedure.

## 5.2. Bootstrap estimation

The parameters of the fitted distribution can be estimated by parametric (resampling from the fitted distribution) or non-parametric (resampling with replacement from the original data set) bootstraps resampling (see [Efron and Tibshirani \(1994\)](#)). These two parametric and nonparametric bootstrap procedures are described below.

### Parametric bootstrap procedure:

- Estimate  $\psi$  (vector of unknown parameters), say  $\hat{\psi}$ , by using the *MLE* procedure based on a random sample.
- Generate a bootstrap sample  $\{X_1^*, \dots, X_m^*\}$  using  $\hat{\psi}$  and obtain the bootstrap estimate of  $\psi$ , say  $\hat{\psi}^*$ , from the bootstrap sample based on the *MLE* procedure.
- Repeat Step 2 *NBOOT* times.
- Order  $\hat{\psi}_1^*, \dots, \hat{\psi}_{NBOOT}^*$  as  $\hat{\psi}_{(1)}^*, \dots, \hat{\psi}_{(NBOOT)}^*$ . Then obtain  $\eta$ -quantiles and  $100(1-\eta)\%$  confidence intervals for the parameters.

In case of the *HMW* distribution, the parametric bootstrap estimators (PBs) of  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ , are  $\hat{\alpha}_{PB}$ ,  $\hat{\beta}_{PB}$ ,  $\hat{\lambda}_{PB}$  and  $\hat{\theta}_{PB}$ , respectively.

### Nonparametric bootstrap procedure

- Generate a bootstrap sample  $\{X_1^*, \dots, X_m^*\}$ , with the replacement from the original data set.
- Obtain the bootstrap estimate of  $\psi$  with *MLE* procedure, say  $\hat{\psi}^*$ , by using the bootstrap sample.

- Repeat Step 2 *NBOOT* times.
- Order  $\widehat{\psi}_1^*, \dots, \widehat{\psi}_{NBOOT}^*$  as  $\widehat{\psi}_{(1)}^*, \dots, \widehat{\psi}_{(NBOOT)}^*$ . Then obtain  $\eta$ -quantiles and  $100(1-\eta)\%$  confidence intervals for the parameters.

In the case of the *HMW* distribution, the nonparametric bootstrap estimators (NPBs) of  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ , are  $\hat{\alpha}_{NPB}$ ,  $\hat{\beta}_{NPB}$ ,  $\hat{\lambda}_{NPB}$  and  $\hat{\theta}_{NPB}$ , respectively.

### 5.3. Bayesian inference

Bayesian inference procedure has been used by many statistical researchers, especially researchers in the field of survival analysis and reliability engineering. In this section, a complete sample data is analyzed through Bayesian point of view. We assume that the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$  of *HMW* distribution have the following independent prior distributions

$$\alpha \sim \text{Gamma}(a, b), \beta \sim \text{Gamma}(c, d), \lambda \sim \text{Gamma}(e, f), \theta \sim \text{Beta}(n_0, n_1),$$

where  $a, b, e, f, n_0$  and  $n_1$  are positive. Hence, the joint prior density function is

$$\pi(\alpha, \beta, \lambda, \theta) = \frac{b^a c^d f^e}{\Gamma(a)\Gamma(c)\Gamma(e)\text{Beta}(n_0, n_1)} \theta^{n_0-1} (1-\theta)^{n_1-1} \alpha^{a-1} \beta^{c-1} \lambda^{e-1} e^{-(b\alpha+d\beta+f\lambda)} \quad (12)$$

In the Bayesian estimation, we do not know the actual value of the parameter, which may be adversely affected by loss when we choose an estimator. This loss can be measured by a function of the parameter and the corresponding estimator. For the Bayesian discussion, we consider different types of symmetric and asymmetric loss functions such as squared error loss function (*SELF*), weighted squared error loss function (*WSELF*), modified squared error loss function (*MSELF*), precautionary loss function (*PLF*) and *K*-loss function (*KLF*). These loss functions, associated Bayesian estimators and posterior risks are presented in Table 1.

Table 1: Bayes estimator and posterior risk under different loss functions

loss function	Bayes estimator	Posterior risk
$SELF = (\psi - d)^2$	$E(\psi x)$	$Var(\psi x)$
$WSELF = \frac{(\psi-d)^2}{\psi}$	$(E(\psi^{-1} x))^{-1}$	$E(\psi x) - (E(\psi^{-1} x))^{-1}$
$MSELF = \left(1 - \frac{d}{\psi}\right)^2$	$\frac{E(\psi^{-1} x)}{E(\psi^{-2} x)}$	$1 - \frac{E(\psi^{-1} x)^2}{E(\psi^{-2} x)}$
$PLF = \frac{(\psi-d)^2}{d}$	$\sqrt{E(\psi^2 x)}$	$2 \left( \sqrt{E(\psi^2 x)} - E(\psi x) \right)$
$KLF = \left( \sqrt{\frac{d}{\psi}} - \sqrt{\frac{\psi}{d}} \right)^2$	$\sqrt{\frac{E(\psi x)}{E(\psi^{-1} x)}}$	$2 \left( \sqrt{E(\psi x)E(\psi^{-1} x)} - 1 \right)$

For more details see [Ahmad, Mahmoudi, Hamedani, and Kharazmi \(2020\)](#) and [Kharazmi, Saadatinik, and Jahangard \(2019\)](#).

Next, we provide the posterior probability distributions for a complete data set. Let us define the function  $\varphi$  as

$$\varphi(\alpha, \beta, \lambda, \theta) = \theta^{n_0-1} (1-\theta)^{n_1-1} \alpha^{a-1} \beta^{c-1} \lambda^{e-1} e^{-(b\alpha+d\beta+f\lambda)}, \quad \alpha > 0, \beta > 0, \lambda > 0, 0 < \theta < 1.$$

The joint posterior distribution in terms of a given likelihood function  $L(data)$  and joint prior distribution  $\pi(\alpha, \beta, \lambda, \theta)$  is defined as

$$\pi^*(\alpha, \beta, \lambda, \theta|data) \propto \pi(\alpha, \beta, \lambda, \theta)L(data). \quad (13)$$

Hence, we obtain the joint posterior density of the parameters  $\alpha, \beta, \lambda, \theta$  for complete sample data by combining the likelihood function and joint prior density (12). Therefore, the joint posterior density function is given by

$$\pi^*(\alpha, \beta, \lambda, \theta|\underline{x}) = K\varphi(\alpha, \beta, \lambda, \theta)L(\underline{x}, \boldsymbol{\xi}) \tag{14}$$

where

$$L(\underline{x}; \boldsymbol{\xi}) = \prod_{i=1}^n \beta \lambda x_i^{\beta-1} e^{-\alpha \lambda x_i^\beta} \frac{\alpha(1-\theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}}{(1-\theta(1-e^{-(\alpha-1)\lambda x_i^\beta}))^2}, \tag{15}$$

$\boldsymbol{\xi} = (\alpha, \beta, \lambda, \theta)$  and  $K$  is given by

$$K^{-1} = \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \varphi(\alpha, \beta, \lambda, \theta)L(\underline{x}, \boldsymbol{\xi})\partial\alpha\partial\beta\partial\lambda\partial\theta.$$

Moreover, the marginal posterior PDF of  $\alpha, \beta, \lambda$  and  $\theta$  assuming that  $\Psi = (\Psi_1, \Psi_2, \Psi_3, \Psi_4) = (\alpha, \beta, \lambda, \theta)$ , can be given

$$\pi(\Psi_i|\underline{x}) = \int_0^\infty \int_0^\infty \int_0^\infty \pi^*(\Psi|\underline{x})\partial\Psi_j\partial\Psi_k\partial\Psi_l, \tag{16}$$

where  $i, j, k, = 1, 2, 3, 4, i \neq j \neq k \neq$  and  $\Psi_i$  is the  $i$ th member of the vector  $\Psi$ .

It is clear from the equations (14) and (16) that there are no closed-form expressions for the Bayesian estimators under the five loss functions described in Table 1. Because of intractable integrals associated with joint posterior and marginal posterior distributions, we need to use a software to solve integral equations numerically via MCMC method. The two most popular MCMC methods are: the Metropolis-Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller 1953; Hastings 1970) and the Gibbs sampling (Geman and Geman 1984). Gibbs Sampling is a special case of the Metropolis-Hastings algorithm which generates a Markov chain by sampling from the full set of conditional distributions.

The Gibbs sampling algorithm can be described as follows:

Suppose that the general model  $f(\underline{x}|\boldsymbol{\psi})$  is associated with parameter vector  $\boldsymbol{\psi} = (\psi_1, \psi_2, \dots, \psi_p)$  and observed data  $\underline{x}$ . Thus, the joint posterior distribution is  $\pi(\psi_1, \psi_2, \dots, \psi_p|\underline{x})$ . We also assume that  $\boldsymbol{\psi}_0 = (\psi_1^{(0)}, \psi_2^{(0)}, \dots, \psi_p^{(0)})$  is the initial values vector to start Gibbs sampler. The Gibbs sampler draws the values for each iteration in  $p$  steps by drawing a new value for each parameter from its full conditional distribution given the most recently drawn values of all other parameters. In symbols, the steps for any iteration, say iteration  $k$ , are as follows:

- Starting with an initial estimate  $(\psi_1^{(0)}, \psi_2^{(0)}, \dots, \psi_p^{(0)})$
- Draw  $\psi_1^k$  from  $\pi(\psi_1|\psi_2^{k-1}, \psi_3^{k-1}, \dots, \psi_p^{k-1}, \underline{x})$
- Draw  $\psi_2^k$  from  $\pi(\psi_2|\psi_1^k, \psi_3^{k-1}, \dots, \psi_p^{k-1}, \underline{x})$ ; and so on down to
- Draw  $\psi_p^k$  from  $\pi(\psi_p|\psi_1^k, \psi_2^k, \dots, \psi_{p-1}^k, \underline{x})$ .

In case of the *HMW* distribution, by considering the parameter vector  $\Psi = (\alpha, \beta, \lambda, \theta)$  and initial parameter vector  $\Psi_0 = c(\alpha^0, \beta^0, \lambda^0, \theta^0)$ , the posterior samples are extracted by the above Gibbs sampler where the full conditional distributions are given as

$$\pi(\alpha|\beta^{k-1}, \lambda^{k-1}, \theta^{k-1}, \underline{x}) \propto \alpha^{a-1} \exp(-\alpha(b + \lambda \sum_{i=1}^n x_i^\beta) + \Phi(\underline{x}, \Psi)),$$

$$\pi(\beta|\alpha^{k-1}, \lambda^{k-1}, \theta^{k-1}, \underline{x}) \propto \beta \exp\left(-\beta(d - \sum_{i=1}^n \log(x_i)) + \alpha\lambda \sum_{i=1}^n x_i^\beta + \Phi(\underline{x}, \Psi)\right),$$

$$\pi(\lambda|\alpha^{k-1}, \beta^{k-1}, \theta^{k-1}, \underline{x}) \propto \lambda^e \exp\left(-\lambda(f + \alpha \sum_{i=1}^n x_i^\beta) + \Phi(\underline{x}, \Psi)\right),$$

and

$$\pi(\theta|\alpha^{k-1}, \beta^{k-1}, \lambda^{k-1}, \underline{x}) \propto \theta^{n_0} (1 - \theta)^{n_1} \exp(\Phi(\underline{x}, \Psi)),$$

where

$$\Phi(\underline{x}, \Psi) = \sum_{i=1}^n \log(\alpha(1 - \theta) + \theta e^{-(\alpha-1)\lambda x_i^\beta}) - 2 \sum_{i=1}^n \log(1 - \theta(1 - e^{-(\alpha-1)\lambda x_i^\beta})).$$

Often Bayesian inference requires computing intractable integrals to generate posterior samples. Using Gibbs sampling, one can obtain samples from the joint posterior distribution. In practice, simulations related to Gibbs sampling are conducted through a special software WinBUGS. WinBUGS software was developed in 1997 to simulate data of complex posterior distributions, where analytical or numerical integration techniques cannot be applied. Also, we can use OpenBUGS software, which is an open source version of WinBUGS. Since there are no prior information about the hyper parameters in (12), one can implement the idea of Congdon (2001) and these parameters can be chosen as  $a = b = c = d = e = f = n_0 = n_1 = 0.0001$ . Hence, we can use *MCMC* procedure to extract posterior samples of (14) by means of Gibbs sampling process in OpenBUGS software.

#### 5.4. Monte Carlo simulation study

In this subsection, we assess the performance of the *MLEs* of the parameters with respect to the sample size  $n$  for the *HMW*( $\alpha, \beta, \lambda, \theta$ ) distribution. The assessment of the performance is based on a simulation study using the Monte Carlo method. First, we generate samples of size  $n$  from (7). The inversion method is used to generate samples of the *HMW* distribution generated as the root of

$$(1 - U)(1 - \theta) + \theta(1 - U)(1 - e^{-\lambda x^\beta})^{\alpha-1} = (1 - e^{-\lambda x^\beta})^\alpha,$$

where  $U \sim U(0, 1)$  is a uniform variate on the unit interval. Let  $\hat{\alpha}$ ,  $\hat{\beta}$ ,  $\hat{\lambda}$  and  $\hat{\theta}$  be the *MLEs* of the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ , respectively. We compute the mean square error (*MSE*) and bias of the *MLEs* of the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ , based on the simulation results of  $N = 1000$  independent replications. results are summarized in Table 2 for selected values of  $n$ ,  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ . From Table 2 the results verify that *MSE* and bias of the *MLEs* of the parameters decrease as sample size  $n$  increases. Hence, the *MLEs* of  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$ , are consistent estimators.

## 6. The LHMW regression model

Let the random variable  $X$  follow a Weibull distribution with *PDF* and *CDF*, respectively as

$$g(x; \lambda, k) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp\left[-(x/\lambda)^k\right], \quad (17)$$

$$G(x; \lambda, k) = 1 - \exp\left[-(x/\lambda)^k\right]. \quad (18)$$

Inserting (17) and (18) in (2), we have

$$f(y, \alpha, \theta, \lambda, k) = \frac{k}{\lambda} \left(\frac{y}{\lambda}\right)^{k-1} \exp\left[-(y/\lambda)^k\right] \times \left(\exp\left[-(y/\lambda)^k\right]\right)^{\alpha-1} \frac{\alpha(1-\theta) + \theta(\exp\left[-(y/\lambda)^k\right])^{\alpha-1}}{(1-\theta)(1 - (\exp\left[-(y/\lambda)^k\right])^{\alpha-1})^2}. \quad (19)$$

Table 2: MSEs and average biases(values in parentheses) of the simulated estimates

		$\alpha = 0.5$	$\beta = 1.0$	$\lambda = 1.5$	$\theta = 0.05$
$n$	30	0.1297 (-0.2797)	0.0548 (-0.0137)	16.4499 (3.3032)	2.2918 (0.0956)
	50	2.7672 (-0.2043)	0.0380 (-0.0263)	9.5995 (2.4463)	2.1054 (0.0462)
	100	0.0893 (-0.2248)	0.0197 (-0.0202)	5.1211 (1.7123)	0.0464 (-0.0055)
	200	0.0687 (-0.2025)	0.0105 (-0.0145)	2.8466 (1.2002)	0.0142 (0.0087)
		$\alpha = 0.5$	$\beta = 1.5$	$\lambda = 1.0$	$\theta = 0.05$
$n$	30	14.5071 (-0.1397)	0.1328 (-0.0288)	7.9946 (2.1579)	0.4992 (0.0452)
	50	0.1186 (-0.2384)	0.0833 (-0.0520)	4.3050 (1.6120)	5.6953 (0.1504)
	100	0.0971 (-0.2182)	0.0472 (-0.0418)	2.4130 (1.1649)	0.0813 (0.0068)
	200	0.0725 (-0.2035)	0.0261 (-0.0184)	1.3738 (0.8408)	0.2781 (0.0324)
		$\alpha = 0.7$	$\beta = 1.0$	$\lambda = 1.5$	$\theta = 0.02$
$n$	30	3.9228 (-0.1715)	0.0450 (-0.0415)	16.3084 (2.9951)	0.3217 (0.0508)
	50	0.4666 (-0.2956)	0.0371 (-0.0635)	10.5131 (2.3362)	0.0213 (0.0234)
	100	0.1670 (-0.2944)	0.0181 (-0.0417)	5.7535 (1.6320)	1.4583 (0.0651)
	200	0.1425 (-0.2815)	0.0095 (-0.0339)	3.6115 (1.2757)	0.0183 (0.0378)
		$\alpha = 0.7$	$\beta = 2.0$	$\lambda = 1.0$	$\theta = 0.1$
$n$	30	2.0483 (-0.1815)	0.2149 (-0.0642)	8.8041 (2.1723)	2.4038 (0.1016)
	50	1.5281 (-0.2362)	0.1305 (-0.0808)	5.6843 (1.6985)	2.3599 (0.0502)
	100	0.1830 (-0.2673)	0.0693 (-0.0749)	3.1972 (1.2292)	0.0759 (-0.0168)
	200	0.1618 (-0.2543)	0.0426 (-0.0731)	2.0204 (0.9420)	0.0219 (-0.0138)
		$\alpha = 0.7$	$\beta = 1.5$	$\lambda = 1.0$	$\theta = 0.1$
$n$	30	0.5811 (-0.2902)	0.1343 (-0.0467)	10.2552 (2.3453)	2.2400 (0.0338)
	50	0.3308 (-0.2888)	0.0791 (-0.0832)	6.3385 (1.8259)	0.8307 (-0.0151)
	100	0.1863 (-0.2718)	0.0367 (-0.0501)	3.6893 (1.3291)	0.1890 (-0.0073)
	200	0.1689 (-0.2482)	0.0256 (-0.0472)	2.1637 (0.9487)	0.0615 (-0.0047)

The *HMW* (19) is considered as different parametrization of the *HMW* distribution. Assume that the random variable  $X$  follows the *HMW* distribution, given in (19). We obtain the log-*HMW* (*LHMW*) distribution by applying  $Y = \log(X)$  transformation and considering the re-parametrization,  $k = 1/\sigma$  and  $\lambda = \exp(\mu)$ . The resulting *PDF* is

$$f(y; \alpha, \theta, \mu, \sigma) = \frac{1}{\sigma} \exp \left[ \left( \frac{y - \mu}{\sigma} \right) - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \left( \exp \left[ - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \right)^{\alpha - 1} \\ \times \frac{\alpha (1 - \theta) + \theta \left( \exp \left[ - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \right)^{\alpha - 1}}{\left( 1 - \theta \left( 1 - \left( \exp \left[ - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \right)^{\alpha - 1} \right) \right)}, \quad (20)$$

where  $\mu \in \mathfrak{R}$  is the location parameter,  $\sigma > 0$  is the scale parameter, and  $\alpha > 0$  and  $\theta > 0$  are the shape parameters. We refer to equation (20) as the *LHMW* distribution, say  $Y \sim \text{LHMW}(\alpha, \theta, \sigma, \mu)$ . The survival function corresponding to (20) is given by

$$S(y) = \frac{\left( \exp \left[ - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \right)^{\alpha}}{1 - \theta \left( 1 - \left( \exp \left[ - \exp \left( \frac{y - \mu}{\sigma} \right) \right] \right)^{\alpha - 1} \right)}, \quad (21)$$

Let  $Z = (Y - \mu)/\sigma$ , then the *PDF* of the standardized random variable is

$$f(y; \alpha, \theta, \mu, \sigma) = \exp [z - \exp(z)] \left( \exp [- \exp(z)] \right)^{\alpha - 1} \frac{\alpha (1 - \theta) + \theta \left( \exp [- \exp(z)] \right)^{\alpha - 1}}{\left( 1 - \theta \left( 1 - \left( \exp [- \exp(z)] \right)^{\alpha - 1} \right) \right)} \quad (22)$$

Figure 9 displays the some possible shapes of the *LHMW* distribution. It is seen that the *LHMW* distribution could be a good choice to model left skewed lifetime dependent variable with some covariates.



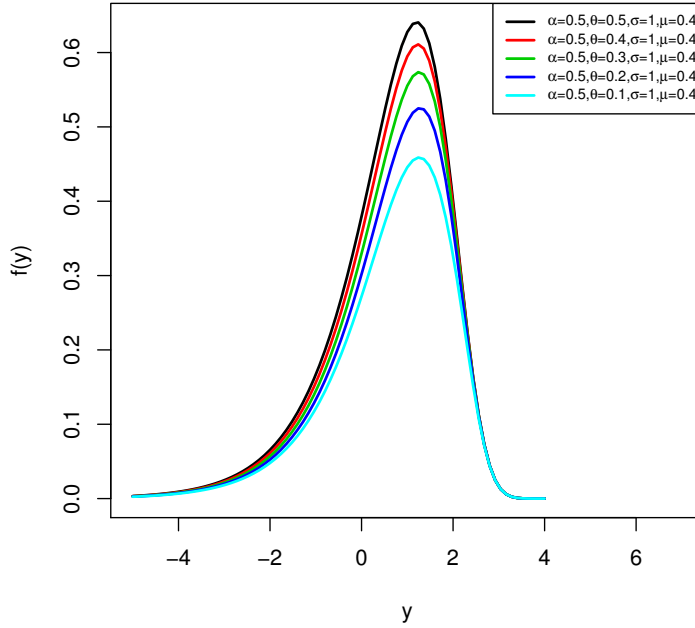


Figure 9: The *PDF* plots of the *LHMW* distribution

Now, using the *LHMW* density, we introduce a new location-scale regression model where the regression structure is

$$y_i = \mathbf{v}_i^T \boldsymbol{\beta} + \sigma z_i, \quad i = 1, \dots, n, \quad (23)$$

and  $\mathbf{v}_i^T$  is the explanatory variable vector and  $z_i$  is the random error following the density in (22). The regression parameters are represented by  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ . The location-scale type regression models are very popular in the recent years. The authors have proposed different location-scale regression models to model different characteristics of the lifetime data sets with some covariates such as Alizadeh, Altun, Cordeiro, and Rasekhi (2018a); Alizadeh, Lak, Rasekhi, Ramires, Yousof, and Altun (2018b); Altun, Yousof, and Hamedani (2018); Korkmaz, Altun, Yousof, and Hamedani (2019b); Korkmaz, Altun, Alizadeh, and Yousof (2019a); Korkmaz, Altun, Yousof, and Hamedani (2020); Esmaeili, Lak, and Altun (2020); Nofal, Altun, Afify, and Ahsanullah (2019); Yousof, Altun, Rasekhi, Alizadeh, Hamedani, and Ali (2019); Yadav, Altun, and Yousof (2019) and among others.

Here, we use the *MLE* method to estimate the unknown parameters of the *LHMW* regression model. First, we define some useful mathematical notations. Let  $y_i$  be a dependent variable, distributed as *LHMW*, given in (19). The dependent variable is defined as  $y_i = \min\{\log(x_i), \log(c_i)\}$  where  $\log(x_i)$  and  $\log(c_i)$  represent the log-lifetime and log-censoring times, respectively. Additionally, we split the observations into the two set. These are  $F$  and  $C$  which indicate the individuals with log-lifetime and log-censoring, respectively. In view of these definitions, the log-likelihood function of the *LHMW* regression model is

$$\begin{aligned} \ell(\tau) &= r \log\left(\frac{1}{\sigma}\right) + \sum_{i \in F} (z_i - u_i) + (1 - \alpha) \sum_{i \in F} u_i + \sum_{i \in F} \log\left(\frac{\alpha(1 - \theta) + \theta(\exp[-u_i])^{\alpha-1}}{(1 - \theta)(1 - (\exp[-u_i])^{\alpha-1})}\right) \\ &+ \sum_{i \in C} \log\left[\frac{(\exp[-u_i])^\alpha}{1 - \theta(1 - (\exp[-u_i])^{\alpha-1})}\right] \end{aligned} \quad (24)$$

where  $\boldsymbol{\tau} = (\alpha, \theta, \sigma, \boldsymbol{\beta}^\top)^\top$  is the parameter vector,  $u_i = \exp(z_i)$ ,  $z_i = (y_i - \mathbf{v}_i^\top \boldsymbol{\beta})/\sigma$  and  $r$  is the number of uncensored observations. The MLEs of  $\boldsymbol{\tau}$  can be obtained by minimizing the negative value of the log-likelihood function in (24). The **optim** function of **R** software is used for this purpose.

### 6.1. Residual analysis

To decide the accuracy of the fitted regression model, we analyze the departure from error distribution by means of the residual analysis. Here, two residuals are used. These are martingale and modified deviance residuals. Residual analysis is an important step of the any regression analysis to check the adequacy of the fitted model. The martingale residuals of the *LHMW* regression model are

$$r_{M_i} = \begin{cases} 1 + \ln \left( \frac{(\exp[-u_i])^\alpha}{1 - \theta(1 - (\exp[-u_i])^{\alpha-1})} \right), & \text{if } i \in F \\ \ln \left( \frac{(\exp[-u_i])^\alpha}{1 - \theta(1 - (\exp[-u_i])^{\alpha-1})} \right), & \text{if } i \in C \end{cases} \quad (25)$$

where  $u_i = \exp((y_i - \mathbf{x}_i^\top \boldsymbol{\beta})/\sigma)$ . Using the martingale residuals, the modified deviance residuals of the *LHMW* regression are

$$r_{D_i} = \begin{cases} \text{sign}(r_{M_i}) \{ -2[r_{M_i} + \log(1 - r_{M_i})] \}^{1/2}, & \text{if } i \in F \\ \text{sign}(r_{M_i}) \{ -2r_{M_i} \}^{1/2}, & \text{if } i \in C, \end{cases} \quad (26)$$

where  $r_{M_i}$  is the martingale residual. The modified deviance residuals are more acceptable and used than martingale residuals. The reason is that the modified deviance residuals are normally distributed with zero mean and unit variance once the fitted regression model is suitable and accurate for the given data.

## 7. Practical data applications

In this section, we provide two applications for modeling the *HMW* distribution to real data sets for illustrative purposes. These applications will show the flexibility and usefulness of the *HMW* distribution. In order to achieve this goal, first we consider the strengths of 1.5 cm glass fibres data set and the parameter estimations are done by means of three methods (maximum likelihood, Bayesian and bootstrap) which are discussed in Section 5. Second we show the performance of survival regression model *LHMW* distribution via maximum likelihood method by analyzing Kidney data set which is associated with covariate variables.

### 7.1. Univariate data modeling

The data set obtained from [Smith and Naylor \(1987\)](#) represents the strengths of 1.5 cm glass fibres, measured at the National Physical Laboratory, England. The observations are as follows;

Table 3: Strengths of glass fibres data set

0.55	0.93	1.25	1.36	1.49	1.52	1.58	1.61	1.64	1.68	1.73	1.81	2.00	0.74	1.04	1.27	1.39
1.49	1.53	1.59	1.61	1.66	1.68	1.76	1.82	2.01	0.77	1.11	1.28	1.42	1.50	1.54	1.60	1.62
1.66	1.69	1.76	1.84	2.24	0.81	1.13	1.29	1.48	1.50	1.55	1.61	1.62	1.66	1.70	1.77	1.84
0.84	1.24	1.30	1.48	1.51	1.55	1.61	1.63	1.67	1.70	1.78	1.89					

**Graphical measure:** The total time test (*TTT*) plot due to Aarset (1987) is an important graphical approach to verify whether the data can be applied to a specific distribution or not. The *TTT* plot for this data set presented in Figure 10 indicates that the empirical hazard rate

functions of the strengths of glass fibres data is increasing. Therefore, the *HMW* distribution is appropriate to fit these data.

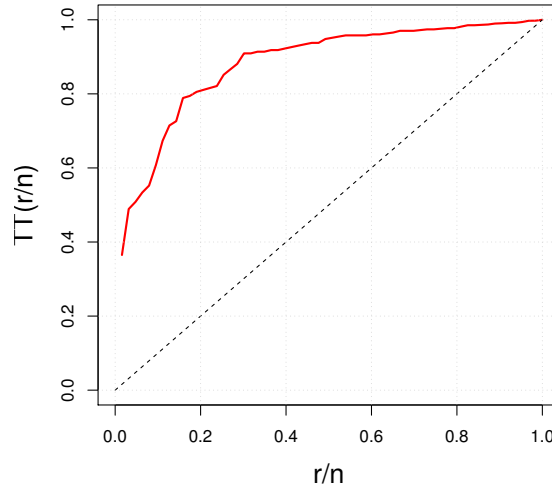


Figure 10: Scaled-*TTT* plot of the strengths of glass fibres data set.

#### *Bootstrap inference for HMW parameters*

Here we obtain point and %95 confidence interval (*CI*) estimation of parameters for the *HMW* distribution by parametric bootstrap method for the real data set. We provide results of bootstrap estimation based on 1000 bootstrap replicates in Table 4. It is interesting to look at the joint distribution of the bootstrapped values in a scatter plot in order to understand the potential structural correlation between parameters. The corresponding plots of the bootstrap estimation are shown in Figure 11.

Table 4: Bootstrap point and interval estimation of the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$

	parametric bootstrap		non-parametric bootstrap	
	point estimation	<i>CI</i>	point estimation	<i>CI</i>
$\alpha$	0.050	(0.000, 0.225)	0.046	(0.000, 0.152)
$\beta$	2.940	(1.854, 5.235)	2.613	(1.827, 4.242)
$\lambda$	1.152	(0.204, 2.519)	1.403	(0.455, 2.469)
$\theta$	0.008	(0.002, 0.105)	0.007	(0.002, 0.073)

#### *MLE estimation and comparison with other models*

Here, we fit the *HMW* distribution to the strengths glass fibres data set and compare it with the Marshall-Olkin Weibull (*MOEW*), Log Logistic (*LL*) and Weibull densities. Table 5 shows the *MLEs* of parameters, log-likelihood, Akaike information criterion (*AIC*), Cramer-von Mises ( $W^*$ ), Anderson-Darling ( $A^*$ ) and *p*-value (*P*) statistics for the data set. The *HMW* distribution provides the best fit for the data set as it shows the lowest *AIC*,  $A^*$  and  $W^*$  than other considered models. The relative histogram, fitted *HMW*, *MOEW*, *LL* and Weibull *PDFs* and corresponding empirical and fitted *CDFs* are plotted in Figure 13 for the current data set. In addition, the *Q*-*Q* and *P*-*P* plots for the *HMW* and other fitted distributions are also displayed in Figure 14. These plots support the results in Table 5. We compare the *HMW* model with a set of competitive models, namely:

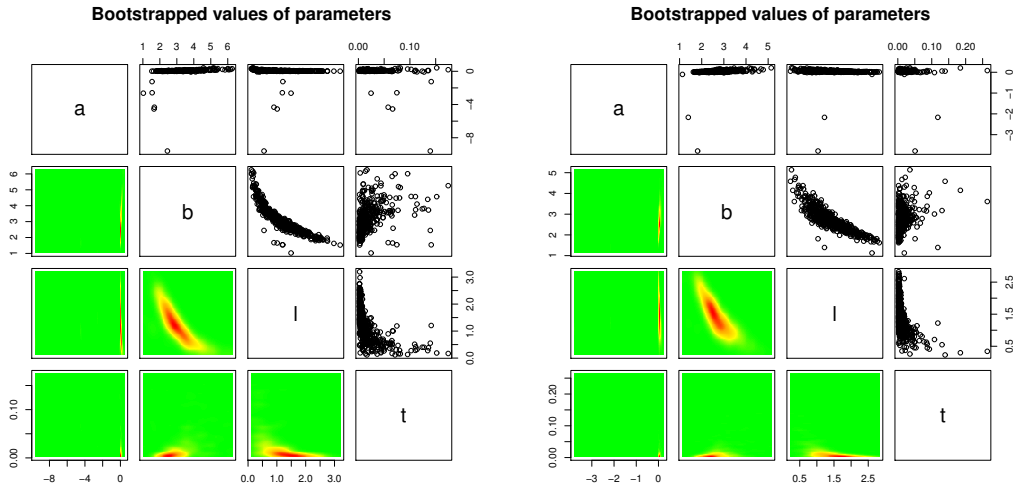


Figure 11: Parametric (left) and non-parametric (right) bootstrapped values of parameters of the HMW distribution for the strengths of glass fibres data.

- (i) The log-logistic distribution: The two parameter log-logistic (*LL*) distribution density function is given by

$$f(x; \alpha, \beta) = \beta \alpha^\beta x^{\beta-1} (\alpha^\beta + x^\beta)^{-2}, \quad x > 0, \alpha, \beta > 0.$$

- (ii) The Marshall-Olkin extended Weibull distribution (Ghitany, Al-Hussaini, and Al-Jarallah 2005): The three parameter Marshall-Olkin extended Weibull (*MOEW*) distribution density function is given by

$$f(x; \alpha, \beta, \lambda) = \frac{\alpha \beta \lambda (\lambda x)^{\beta-1} e^{-(\lambda x)^\beta}}{[1 - \bar{\alpha} e^{-(\lambda x)^\beta}]^2}, \quad x > 0, \alpha, \beta, \lambda > 0, \bar{\alpha} = 1 - \alpha.$$

- (iii) The Weibull distribution: The Weibull (*We*) distribution with shape parameter  $\alpha$  and scale parameter  $\beta$  has density given by

$$f(x, \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{x}{\beta}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\beta}\right)^\alpha\right), \quad x > 0, \alpha > 0, \beta > 0.$$

Table 5: Parameter estimates (standard errors), log-likelihood values and goodness of fit measures

Model	<i>MLEs</i> of parameters (s.e)	Log-likelihood	<i>AIC</i>	<i>BIC</i>	<i>A*</i>	<i>W*</i>	<i>K.S</i>	<i>P</i>
<i>HMW</i>	$\hat{\alpha} = 0.041(0.020), \hat{\beta} = 2.501(0.023)$ $\hat{\lambda} = 1.640(0.420), \hat{\theta} = 0.005(0.003)$	-10.80	29.60	38.18	0.28	0.07	0.07	0.81
<i>MOEW</i>	$\hat{\alpha} = 16.650(20.769), \hat{\beta} = 3.201(0.946)$ $\hat{\lambda} = 0.892(0.196)$	-12.03	30.06	36.49	0.56	0.08	0.09	0.55
<i>LL</i>	$\hat{\alpha} = 1.526(0.040), \hat{\beta} = 7.924(0.873)$	-22.78	49.57	53.86	2.37	0.30	0.15	0.10
<i>Weibull</i>	$\hat{\alpha} = 5.780(0.576), \hat{\beta} = 1.628(0.037)$	-15.20	34.41	38.69	1.24	0.21	0.15	0.10

As mentioned in the inference section, there are no closed expressions for *MLE* estimators of the parameters  $\alpha, \beta, \lambda$  and  $\theta$ . We use numerical methods to obtain *MLE* estimations of

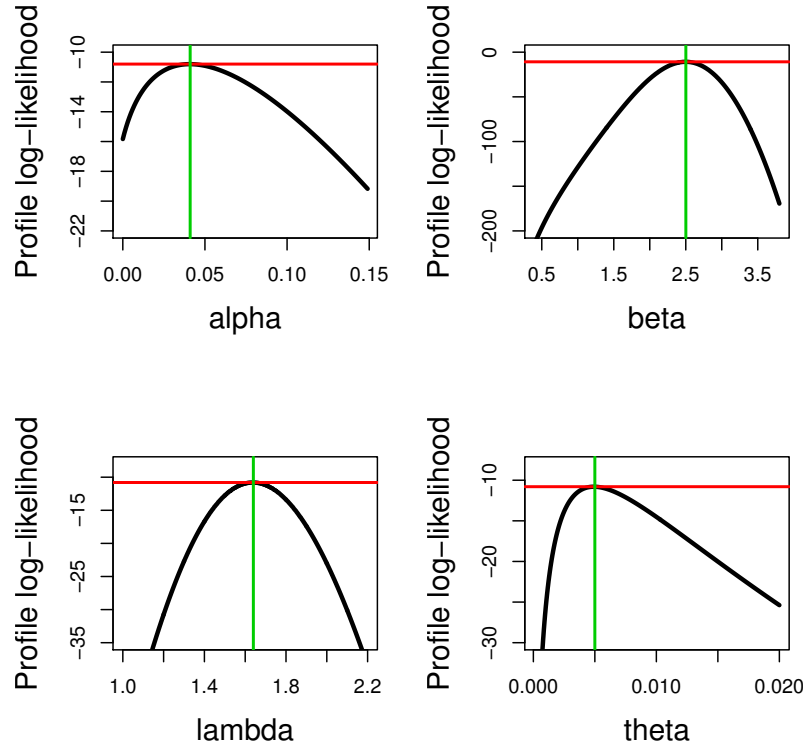


Figure 12: Profile-likelihood plots of parameters of  $HMW$  distribution for strengths of glass fibres.

these parameters. To evaluate the results of  $MLE$  estimation, we provide profile-likelihood plots of  $HMW$  distribution for each parameter in Figure 12.

#### *Bayesian estimation under different loss functions*

Here, we provide Bayesian estimation results for the parameters of  $HMW$  distribution. Bayesian estimators associated with the parameters of  $HMW$  distribution are computed based on the single chain of 20,000 cycles of Gibbs sampler with a conservative burn-in period of the first 5000 iterations. The corresponding Bayesian point and interval estimations and posterior risk are provided in Table 6 for the Strengths of glass fibres data set. Table 7 provides 95% credible and  $HPD$  intervals for each parameter of the  $HMW$  distribution. The convergence of Gibbs sampler process are verified through graphical inspection (Trace, Autocorrelation and Histogram plots) of the posterior sampled values. It is observed that Gibbs samples of all the parameter estimates achieve a good stationary phase for both data sets. We provide the posterior summary plots for 10,000 cycles of Gibbs sampler in Figures 15, 16 and 17 .

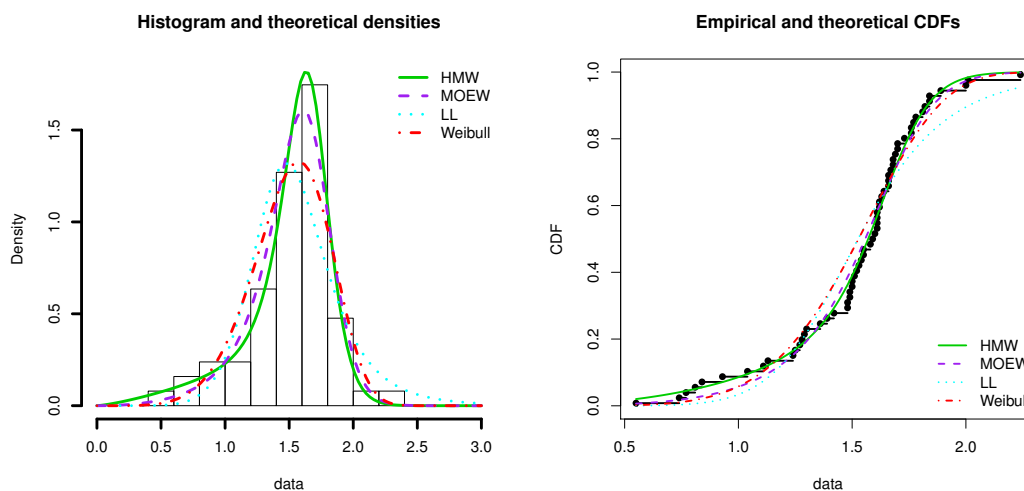


Figure 13: Estimated densities and corresponding CDFs for the data set.

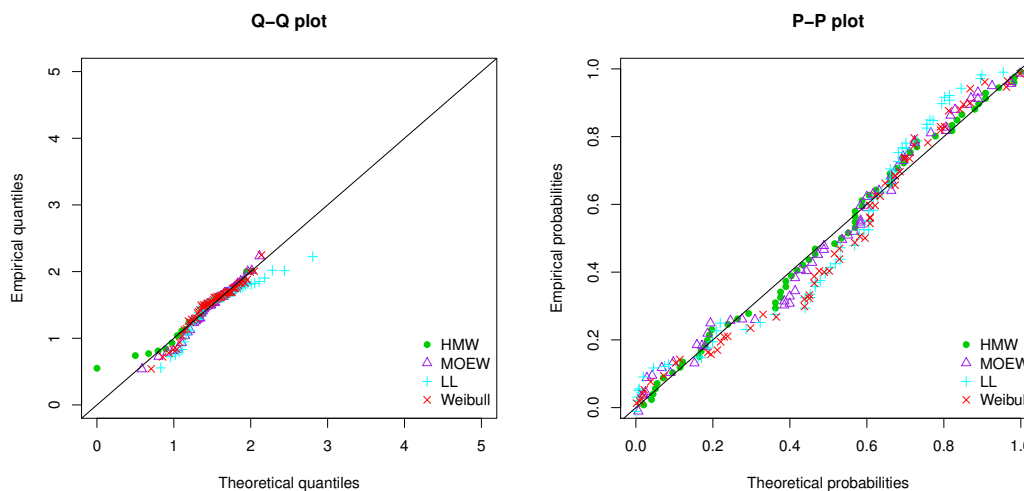


Figure 14: Q-Q and P-P plots for the data set.

## 7.2. Data modeling with covariate variables: Kidney data set

In this application, we assess the performance of *LHMW* regression model via application to a real data set. The data set represents the recurrence times to infection of Kidney patients. The data set is available in **survival** package of **R** software. The censoring rate of the data is approximately 23%. The dependent variable, recurrence times to infection  $y_i$ , is modeled by age  $x_{i1}$  and gender  $x_{i2}$  (0=male, 1=female). The below regression model is fitted

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \sigma z_i,$$

where  $y_i$  is distributed as *LHMW*.

### Parameter estimation

The estimated parameters of the *LHMW* regression model is obtained by *MLE* method using the **optim** function of **R** software. The estimated parameters, standard errors (SEs), confidence intervals (CIs) as well as the corresponding p-values are listed in Table 8. From

Table 6: Bayesian estimates and their posterior risks of the parameters of  $HMW$  distribution under different loss functions based on the Strengths of glass fibres

Data	Strengths of glass fibres			
Bayesian estimation				
Loss function	$\hat{\alpha} (r_{\hat{\alpha}})$	$\hat{\beta} (r_{\hat{\beta}})$	$\hat{\lambda} (r_{\hat{\lambda}})$	$\hat{\theta} (r_{\hat{\theta}})$
SELF	0.0404 (0.0002)	2.51628 (0.0922)	1.65984 (0.1019)	0.0052 (0.0001)
WSELF	0.0345 (0.0059)	2.47944 (0.0368)	1.59789 (0.0619)	0.0043 (0.0009)
MSELF	0.0274 (0.2060)	2.44239 (0.0149)	1.53561 (0.0390)	0.0035 (0.1891)
PLF	0.0429 (0.0051)	2.53454 (0.0365)	1.69027 (0.0609)	0.0056 (0.0008)
KLF	0.0373 (0.1649)	2.49779 (0.0148)	1.62857 (0.0384)	0.0047 (0.1920)

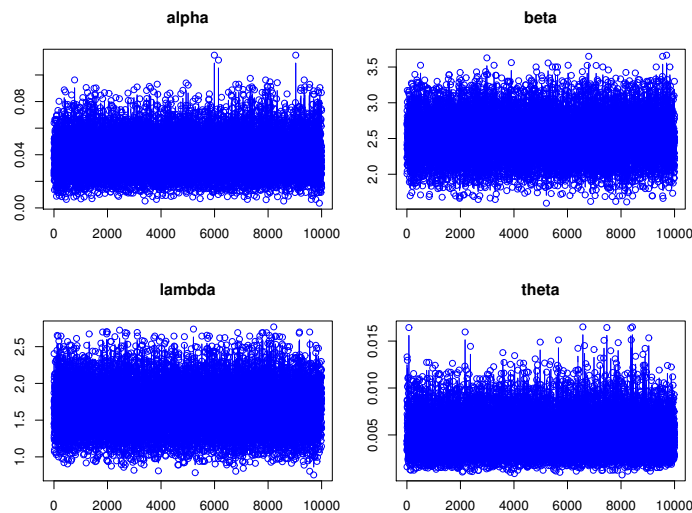


Figure 15: Plots of Bayesian analysis and performance of Gibbs sampling for the Strengths of glass fibres data set. Trace plots of each parameter of  $MHW$  distribution.

Table 7: Credible and  $HPD$  intervals of the parameters  $\alpha$ ,  $\beta$ ,  $\lambda$  and  $\theta$  for the strengths of glass fibers data set

	Credible interval	HPD interval
$\alpha$	(0.030, 0.049)	(0.013, 0.068)
$\beta$	(2.313, 2.710)	(1.952, 3.176)
$\gamma$	(1.435, 1.861)	(1.058, 2.300)
$\lambda$	(0.004, 0.006)	(0.002, 0.009)

results, we conclude that the estimated regression parameters are statistically significant at 1% level. According to the estimated regression parameters, the recurrence times to infection decreases when the age of patient increases. Additionally, the recurrence times to the infection of female patients are higher than those of the male patients.

### Results of residual analysis

The suitability of the fitted  $LHMW$  regression model is evaluated via residual analysis. The plot of the modified deviance residuals and its quantile-quantile ( $Q-Q$ ) plot are displayed in Figure 18 which reveal that the fitted  $LHMW$  regression model provides a good fit to the



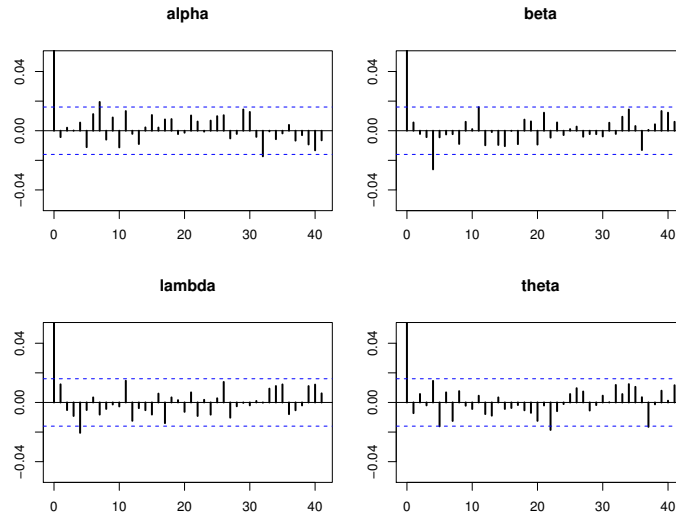


Figure 16: Plots of Bayesian analysis and performance of Gibbs sampling for the Strengths of glass fibres data set. Autocorrelation plots of each parameter of *MHW* distribution.

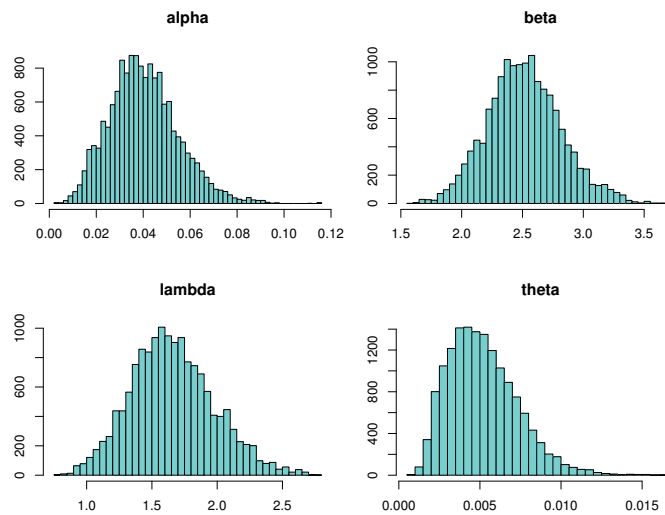


Figure 17: Plots of Bayesian analysis and performance of Gibbs sampling for the Strengths of glass fibres data set. Histogram plots of each parameter of *MHW* distribution.

Table 8: The estimated parameters of the LHMW regression model for the kidney data

Parameters	Estimates	SE	CIs	p-values
$\alpha$	14.389	0.052	[14.287,14.489]	-
$\theta$	$9.266 \times 10^{-9}$	$6.854 \times 10^{-7}$	$[-1.334 \times 10^{-6}, -1.352 \times 10^{-6}]$	-
$\sigma$	1.103	0.106	[0.895,1.310]	-
$\beta_0$	6.255	0.606	[5.067,7.443]	< 0.001
$\beta_1$	-0.003	$3.238 \times 10^{-4}$	[-0.004,-0.003]	< 0.001
$\beta_2$	0.953	0.340	[0.286,1.621]	0.005

data used.

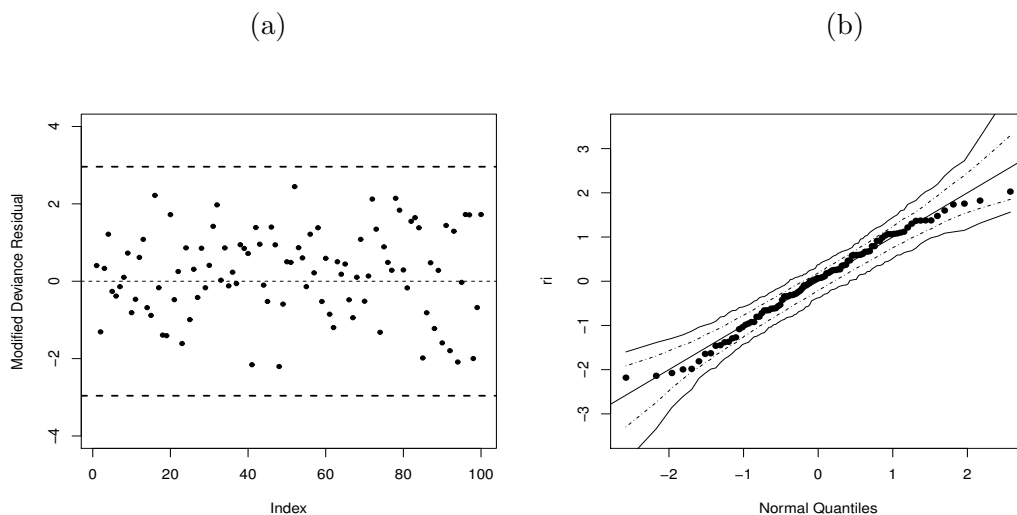


Figure 18: (a) Index plot of the modified deviance residual and (b)  $Q - Q$  plot for modified deviance residual.

## 8. Conclusion

A new family of lifetime distributions is introduced via harmonic mixture mean and its main properties are derived. One of the interesting and important properties of the proposed family is that it includes the Marshall-Olkin family of distributions, as a special case. An example of this family is proposed by considering the Weibull as the baseline distribution. We provide survival regression model of the special model and a comprehensive discussion about Bayesian estimation of the parameters are studied under various loss functions. Numerical results of a maximum likelihood, Bayesian and bootstrap procedures for a univariate real data set is investigated. Moreover, the associated plots to evaluate the results obtained from the the three methods are provided. The data analysis proves, empirically, that the proposed distribution provides a better fit than its sub-models and other competing distributions for the current data. Finally, the performance of the survival regression model of the special distribution is examined by considering a real example with covariate variables.

## References

- Acheson M, McElwee E (1952). “Concerning the Reliability of Electron Tubes.” *Proceeding of the IRE.*, **40**(10), 1204–1206. doi:10.1109/JRPROC.1952.274090.
- Ahmad Z, Mahmoudi E, Hamedani GG, Kharazmi O (2020). “New Methods to Define Heavy-tailed Distributions with Applications to Insurance Data.” *Journal of Taibah University for Science*, **14**(1), 359–382. doi:10.1080/16583655.
- Alizadeh M, Altun E, Cordeiro GM, Rasekhi M (2018a). “The Odd Power Cauchy Family of Distributions: Properties, Regression Models and Applications.” *Journal of Statistical Computation and Simulation*, **88**(4), 785–807. doi:10.1080/00949655.2017.1406938.
- Alizadeh M, Lak F, Rasekhi M, Ramires TG, Yousof HM, Altun E (2018b). “The Odd Log-logistic Topp-Leone G Family of Distributions: Heteroscedastic Regression Models and Applications.” *Computational Statistics*, **33**(3), 785–807. doi:10.1007/s00180-017-0780-9.
- Altun E, Yousof HM, Hamedani GG (2018). “A Flexible Extension of Generalized Half-Normal Distribution: Characterizations and Regression Models.” *International Journal of Applied Mathematics and Statistics*, **57**(3), 27–49.

- Asadi M, Ebrahimi N, Kharazmi O, Soofi ES (2018). “Mixture Models, Bayes Fisher Information, and Divergence Measures.” *IEEE Transactions on Information Theory*, **65**(4), 2316–2321. doi:10.1109/TIT.2018.2877608.
- Bercher JF, Vignat C (2009). “On Minimum Fisher Information Distributions with Restricted Support and Fixed Variance.” *Information Sciences*, **179**(22), 3832–3842. doi:10.1016/j.ins.2009.07.013.
- Chen MH, Shao QM (1999). “Monte Carlo Estimation of Bayesian Credible and HPD Intervals.” *Journal of Computational and Graphical Statistics*, **8**(1), 69–92. doi:10.1080/10618600.
- Congdon P (2001). *Bayesian Statistical Modelling*. John Wiley and Sons New York.
- Cox DR (1972). “Regression Models and Life Tables.” *Journal of the Royal Statistical Society, Series B (Methodological)*, **34**(2), 187–202. doi:10.1111/j.2517-6161.1972.tb00899.x.
- Efron B, Tibshirani RJ (1994). *An Introduction to the Bootstrap*. CRC press.
- Esmaeili H, Lak F, Altun E (2020). “The Ristic-Balakrishnan Odd Log-logistic Family of Distributions: Properties and Applications.” *Statistics, Optimization and Information Computing*, **8**(1), 17–35. doi:10.19139/soic-2310-5070-715.
- Geman S, Geman D (1984). “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images.” *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-6*(6), 1–15. doi:10.1899/15-9834.6.
- Ghitany ME, Al-Hussaini EK, Al-Jarallah RA (2005). “Marshall Olkin Extended Weibull Distribution and Its Application to Censored Data.” *Journal of Applied Statistics*, **32**(10), 1025–1034. doi:10.1080/02664760500165008.
- Glänzel W (1987). “A Characterization Theorem Based on Truncated Moments and Its Application to Some Distribution Families.” *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*, **B**, 75–84. doi:10.1007/978-94-009-3965-3\_8.
- Glänzel W (1990). “Some Consequences of a Characterization Theorem Based on Truncated Moments.” *Statistics: A Journal of Theoretical and Applied Statistics*, **21**(4), 613–618. doi:10.1080/02331889008802273.
- Hamedani GG (2013). “On Certain Generalized Gamma Convolution Distributions II.” *Technical Report, No. 484, MSCS, Marquette University*.
- Hastings WK (1970). “Monte Carlo Sampling Methods Using Markov Chains and Their Applications.” *Biometrika*, **57**(1), 97–109. doi:10.1093/biomet/57.1.97.
- Kharazmi O, Saadatinik A, Jahangard S (2019). “Odd Hyperbolic Cosine Exponential-Exponential (OHC-EE) Distribution.” *Annals of Data Science*, **6**(4), 765–785. doi:10.1007/s40745-019-00200-z.
- Korkmaz MC, Altun E, Alizadeh M, Yousof HM (2019a). “A New Flexible Lifetime Model with Log-location Regression Modeling, Properties and Applications.” *Journal of Statistics and Management Systems*, **22**(5), 871–891. doi:10.1080/09720510.2019.1572980.
- Korkmaz MC, Altun E, Yousof HM, Hamedani GG (2019b). “The Odd Power Lindley Generator of Probability Distributions: Properties, Characterizations and Regression Modeling.” *International Journal of Statistics and Probability*, **8**(2), 70–89.
- Korkmaz MC, Altun E, Yousof HM, Hamedani GG (2020). “The Hjorth’s IDB Generator of Distributions: Properties, Characterizations, Regression Modeling and Applications.” *Journal of Statistical Theory and Applications*, **19**(1), 57–74.

- Mendenhall W, Hader RA (1958). “Estimation of Parameters of Mixed Exponentially Distributed Failure Time Distributions from Censored Life Test Data.” *Biometrika*, **45**, 504–520. doi:10.1093/biomet/45.3-4.504.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953). “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics*, **21**(6), 1087–1092. doi:10.1063/1.1699114.
- Nofal ZM, Altun E, Afify AZ, Ahsanullah M (2019). “The Generalized Kumaraswamy-G Family of Distributions.” *Journal of Statistical Theory and Applications*, **18**(4), 329–342. doi:10.2991/jsta.d.191030.001.
- Shaked M, Shanthikumar JG (2007). *Stochastic Orders*. Springer Science and Business Media.
- Smith RL, Naylor JC (1987). “A Comparison of Maximum Likelihood and Bayesian Estimators for the Three-parameter Weibull Distribution.” *Applied Statistics*, **36**, 358–369. doi:10.2307/2347795.
- Yadav AS, Altun E, Yousof HM (2019). “Burr-Hatke Exponential Distribution: A Decreasing Failure Rate Model, Statistical Inference and Applications.” *Annals of Data Science*, pp. 1–20. doi:10.1007/s40745-019-00213-8.
- Yousof HM, Altun E, Rasekhi M, Alizadeh M, Hamedani GG, Ali MM (2019). “A New Lifetime Model with Regression Models, Characterizations and Applications.” *Communications in Statistics-Simulation and Computation*, **48**(1), 264–286. doi:10.1080/03610918.2017.1377241.

## Appendix A

**Theorem 1.** Let  $(\Omega, \mathcal{F}, \mathbf{P})$  be a given probability space and let  $H = [a, b]$  be an interval for some  $d < b$  ( $a = -\infty$ ,  $b = \infty$  might as well be allowed). Let  $X : \Omega \rightarrow H$  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

$$\mathbf{E}[q_2(X) \mid X \geq x] = \mathbf{E}[q_1(X) \mid X \geq x] \xi(x), \quad x \in H,$$

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

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

where the function  $s$  is a solution of the differential equation  $s' = \frac{\xi' q_1}{\xi q_1 - q_2}$  and  $C$  is the normalization constant, such that  $\int_H dF = 1$ .

**Note:** The goal is to have the function  $\xi(x)$  as simple as possible.

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence (see, Glänzel (1990)), in particular, let us assume that there is a sequence  $\{X_n\}$  of random variables with distribution functions  $\{F_n\}$  such that the functions  $q_{1n}, q_{2n}$  and  $\xi_n$  ( $n \in \mathbb{N}$ ) satisfy the conditions of Theorem 1 and let  $q_{1n} \rightarrow q_1$ ,  $q_{2n} \rightarrow q_2$  for some continuously differentiable real functions  $q_1$  and  $q_2$ . Let, finally,  $X$  be a random variable with distribution  $F$ . Under the condition that  $q_{1n}(X)$  and  $q_{2n}(X)$  are uniformly integrable and the family  $\{F_n\}$  is relatively compact, the sequence  $X_n$  converges to  $X$  in distribution if and only if  $\xi_n$  converges to  $\xi$ , where

$$\xi(x) = \frac{E[q_2(X) \mid X \geq x]}{E[q_1(X) \mid X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions  $q_1$ ,  $q_2$  and  $\xi$ , respectively. It guarantees, for instance, the 'convergence' of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if  $\alpha \rightarrow \infty$ .

A further consequence of the stability property of Theorem 1 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions  $q_1$ ,  $q_2$  and, specially,  $\xi$  should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose  $\xi$  as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

In some cases, one can take  $q_1(x) \equiv 1$ , which reduces the condition of Theorem 1 to  $\mathbf{E}[q_2(X) | X \geq x] = \xi(x)$ ,  $x \in H$ . We, however, believe that employing three functions  $q_1$ ,  $q_2$  and  $\xi$  will enhance the domain of applicability of Theorem 1.

**Affiliation:**

Omid Kharazmi  
Department of Statistics  
Vali-e-Asr university of Rafsanjan  
Rafsanjan, Iran  
E-mail: [omid.kharazmi@vru.ac.ir](mailto:omid.kharazmi@vru.ac.ir)

Ali Saadati Nik  
Department of Statistics  
University of Mazandaran  
Babolsar, Iran  
E-mail: [a.saadatinik@stu.umz.ac.ir](mailto:a.saadatinik@stu.umz.ac.ir)

GG. Hamedani  
Department of Mathematical and Statistical Sciences  
Marquette University  
USA  
E-mail: [gholamhoss.hamedani@marquette.edu](mailto:gholamhoss.hamedani@marquette.edu)

Emrah Altun  
Department of Mathematics  
Bartın University  
Turkey  
E-mail: [emrahaltun@bartin.edu.tr](mailto:emrahaltun@bartin.edu.tr)