Journal of Multivariate Analysis 123 (2014) 19–29

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

On bivariate Weibull-Geometric distribution

Debasis Kundu [a,*], Arjun K. Gupta [b]

[a] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India

[b] Department of Mathematics and Statistics, Bowling Green State University, Bowling Green, OH 43403 - 0221, USA

ARTICLE INFO

Article history:
Received 11 May 2013
Available online 20 August 2013

AMS subject classifications:
primary 62E15
secondary 62H10

Keywords:
Geometric maximum
Weibull distribution
EM algorithm
Fisher information matrix
Monte Carlo simulation

ABSTRACT

Marshall and Olkin (1997) [14] provided a general method to introduce a parameter into a family of distributions and discussed in details about the exponential and Weibull families. They have also briefly introduced the bivariate extension, although not any properties or inferential issues have been explored, mainly due to analytical intractability of the general model. In this paper we consider the bivariate model with a special emphasis on the Weibull distribution. We call this new distribution as the bivariate Weibull-Geometric distribution. We derive different properties of the proposed distribution. This distribution has five parameters, and the maximum likelihood estimators cannot be obtained in closed form. We propose to use the EM algorithm, and it is observed that the implementation of the EM algorithm is quite straightforward. Two data sets have been analyzed for illustrative purposes, and it is observed that the new model and the proposed EM algorithm work quite well in these cases.

1. Introduction

Marshall and Olkin [14] proposed a method to introduce a parameter in a family of distributions. It induces an extra parameter to a model, and hence brings more flexibility. The model has a nice physical interpretation also. They have explored different properties in the case of exponential and Weibull distributions. From now on we call them as univariate exponential geometric (UEG) and univariate Weibull geometric (UWG) distributions, respectively. It is observed that the proposed method brings more flexibility to the well-known exponential and Weibull models, and it has several desirable properties. Since then extensive work has been done on this method exploring different properties and extending to some other lifetime distributions; see for example [7,6,16,2] and the references cited therein.

In the same paper Marshall and Olkin [14] briefly introduced a bivariate extension of the proposed model. They did not provide any properties and also did not discuss any inferential issues. Interestingly, although extensive work has been done on the univariate version of the model, not much work has been done on the bivariate generalization mainly due to its analytical intractability for the general model. We consider the bivariate model with a special emphasis on the Marshall–Olkin bivariate Weibull (MOBW) distribution. We call this new model as the bivariate Weibull-Geometric (BWG) model.

Among several bivariate distributions which can be used to model singular data, Marshall–Olkin's [13] bivariate exponential (MOBE) model plays the most important role. MOBE distribution has exponential marginals; therefore, it has its own limitations. Due to this reason Marshall and Olkin [13] also introduced MOBW distribution which has Weibull marginals. MOBW distribution has four parameters, see for example [9] or [12], and it has a singular component. The proposed BWG model also has a singular component. Moreover, MOBW can be obtained as a special case of the BWG model. Other recent extensions of the MOBE model with a singular component have been discussed in the literature; e.g. see [17,5,11,4].

* Corresponding author.
E-mail addresses: kundu@iitk.ac.in, kundu.debasis@gmail.com (D. Kundu).

0047-259X/$ – see front matter
http://dx.doi.org/10.1016/j.jmva.2013.08.004

Due to the presence of five parameters the joint probability density function of BWG distribution is very flexible, and it can take different shapes depending on the shape parameters. The marginals and conditionals are UWG, and they are also very flexible. It is geometric minimum stable. The generation of random samples from BWG is quite simple; hence, simulation experiments can be performed easily.

The maximum likelihood estimators (MLEs) of the unknown parameters of BWG distribution cannot be obtained in closed forms. It involves solving five non-linear equations simultaneously. Finding initial values and the convergence of the algorithm are important issues. To avoid this problem, we propose to use the expectation maximization (EM) algorithm in computing the MLEs of the unknown parameters. At each 'E'-step, of the EM algorithm, the corresponding 'M'-step involves the maximization of one non-linear function, similar to the EM algorithm proposed by Kundu and Dey [9] in finding the MLEs of the unknown parameters of an MOBW model. For illustrative purposes, we analyze two data sets, and it is observed that the proposed EM algorithm works quite well.

Rest of the paper is organized as follows. In Section 2, we introduce the model. Different properties are discussed in Section 3. The EM algorithm is presented in Section 4. In Section 5 we present the analysis of two data sets, one simulated and one real. Finally in Section 6 we conclude the paper.

2. Preliminaries and model formulation

2.1. Preliminaries and notations

We will be using the following notations through out the paper. A two-parameter Weibull distribution with the shape and scale parameters as $\alpha > 0$ and $\lambda > 0$, respectively, will be denoted by $\text{WE}(\alpha, \lambda)$, and it has the PDF

$$f(x; \alpha, \lambda) = \alpha\lambda x^{\alpha-1} e^{-\lambda x^{\alpha}}; \quad x > 0. \tag{1}$$

In (1) when $\alpha = 1$, it becomes the exponential distribution with parameter λ. A geometric random variable N with probability mass function

$$P(N = n) = p(1-p)^{n-1}; \quad n = 1, 2, \ldots, \ 0 < p < 1 \tag{2}$$

will be denoted by $\text{GM}(p)$. If F is any distribution function, we will use $\bar{F}$ to denote its survival function.

Suppose $\{X_n; n = 1, 2, \ldots\}$ is a sequence of independent and identically distributed (i.i.d.) non-negative random variables with a common distribution function $F(\cdot)$, $N \sim \text{GM}(\theta)$ and it is independent of X_n's. Consider a new random variable U, such that

$$Y = \min\{X_1, \ldots, X_N\}. \tag{3}$$

Then the survival function of Y, say $\bar{G}(y)$, is

$$\bar{G}(y) = P(Y \geq y) = \sum_{n=1}^{\infty} P(Y \geq y | N = n) P(N = n) = \frac{\theta \bar{F}(y)}{1 - (1-\theta)\bar{F}(y)}. \tag{4}$$

Marshall and Olkin [14] proposed this method to introduce a parameter in a family of distributions. It may be easily verified that the above $G(\cdot; \theta)$ is indeed a survival function for all $0 < \theta \leq 1$. It is interesting to note that $\bar{G}$ as defined in (4) is indeed a proper survival function for $\theta \in \mathbb{R}^+$, although in this paper we consider mainly $0 < \theta \leq 1$.

If the distribution function $F(\cdot)$ has a density function $f(\cdot)$, then the density function corresponding to the survival function $\bar{G}(y; \theta)$ becomes

$$g(y; \theta) = \frac{\theta f(y)}{(1 - (1-\theta)\bar{F}(y))^2}, \tag{5}$$

and for $\theta = 1$, $\bar{G} = \bar{F}$.

Marshall and Olkin [14] considered two special cases, namely the one-parameter exponential distribution and two-parameter Weibull distribution, and discussed their properties in detail. When $f(y) = \lambda e^{-\lambda y}$, the PDF (5) takes the following form:

$$g(y; \theta, \lambda) = \frac{\theta \lambda e^{-\lambda y}}{(1 - (1-\theta)e^{-\lambda y})^2}, \tag{6}$$

and it will be denoted by $\text{UEG}(\theta, \lambda)$. In the case of Weibull distribution, *i.e.* when $f(y)$ has the form (1), the PDF (5) takes the form

$$g(y; \theta, \lambda, \alpha) = \frac{\theta \alpha \lambda y^{\alpha-1} e^{-\lambda y^{\alpha}}}{(1 - (1-\theta)e^{-\lambda y^{\alpha}})^2}, \tag{7}$$

and it will be denoted by $\text{UWG}(\theta, \alpha, \lambda)$. The PDF of UWG can be a decreasing, unimodal or some other shapes. The hazard function can be increasing, decreasing or some other shapes.