Hacettepe Journal of Mathematics and Statistics
Volume 46 (6) (2017), 1175 – 1186

Bivariate Weibull-power series class of distributions

Saralees Nadarajah * and Rasool Roozegar †‡

Abstract

We point out that some of the results in Kundu and Gupta [3] in otherwise an excellent paper are incorrect. We propose a more general class of distributions and illustrate its use with two real data sets.

Keywords: EM algorithm, Maximum likelihood estimation, Power series distribution.

2000 AMS Classification: 62E15.

Received : 29.10.2015 *Accepted* : 02.01.2016 *Doi* : 10.15672/HJMS.201610714835

1. Introduction

Kundu and Gupta [3] proposed quite a novel bivariate distribution by compounding a bivariate Weibull distribution with a geometric distribution. The proposal is based on the following construction due to Marshall and Olkin [5]: Suppose $\{(X_{1,n}, X_{2,n}), n = 1, 2, \ldots\}$ are independent and identical random vectors with joint survival function $\overline{F}_{X_1,X_2}$. Let N be a geometric random variable independent of $\{(X_{1,n}, X_{2,n}), n = 1, 2, \ldots\}$ with probability mass function $\Pr(N = n) = \theta(1-\theta)^{n-1}$, $n = 1, 2, \ldots$. Define $Y_1 = \min(X_{1,1}, \ldots, X_{1,N})$ and $Y_2 = \min(X_{2,1}, \ldots, X_{2,N})$. Marshall and Olkin [5] showed that the joint survival function of Y_1 and Y_2 is

$$\overline{F}_{Y_1,Y_2}(y_1, y_2) = \Pr(Y_1 > y_1, Y_2 > y_2) = \frac{\theta \overline{F}_{X_1,X_2}(y_1, y_2)}{1 - (1-\theta)\overline{F}_{X_1,X_2}(y_1, y_2)} \tag{1.1}$$

for $0 < \theta < 1$.

Kundu and Gupta [3] studied the structural properties of (1.1) when $\overline{F}_{X_1,X_2}$ is the joint survival function of Marshall and Olkin [4]'s bivariate Weibull distribution. Kundu and Gupta [3] also showed how the parameters can be estimated by EM algorithm, presented a simulation study and discussed a real data application.

*School of Mathematics, University of Manchester, Manchester, UK, Email: mbbsssn2@manchester.ac.uk
†Department of Statistics, Yazd University, Yazd 89195-741, Iran, Email: rroozegar@yazd.ac.ir
‡Corresponding Author.

Marshall and Olkin [4]'s bivariate Weibull distribution has the joint survival and joint probability density functions specified by

$$\overline{F}_{X_1,X_2}(x_1,x_2) = \begin{cases} \exp\left[-(\lambda_0+\lambda_1)x_1^\alpha - \lambda_2 x_2^\alpha\right], & \text{if } x_1 \geq x_2, \\ \exp\left[-(\lambda_0+\lambda_2)x_2^\alpha - \lambda_1 x_1^\alpha\right], & \text{if } x_1 < x_2 \end{cases} \tag{1.2}$$

and

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \alpha^2\lambda_2(\lambda_0+\lambda_1)x_1^{\alpha-1}x_2^{\alpha-1}\exp\left[-(\lambda_0+\lambda_1)x_1^\alpha - \lambda_2 x_2^\alpha\right], \\ \qquad \text{if } x_1 > x_2, \\ \alpha^2\lambda_1(\lambda_0+\lambda_2)x_1^{\alpha-1}x_2^{\alpha-1}\exp\left[-(\lambda_0+\lambda_2)x_2^\alpha - \lambda_1 x_1^\alpha\right], \\ \qquad \text{if } x_1 < x_2, \\ \alpha\lambda_0 x^{\alpha-1}\exp\left[-(\lambda_0+\lambda_1+\lambda_2)x^\alpha\right], \\ \qquad \text{if } x_1 = x_2 = x, \end{cases} \tag{1.3}$$

respectively, for $x_1 > 0$, $x_2 > 0$, $\alpha > 0$, $\lambda_0 > 0$, $\lambda_1 > 0$ and $\lambda_2 > 0$. Unfortunately, the formula for the latter given in equations (9)-(12) in Kundu and Gupta [3] is

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \alpha^2\lambda_2(\lambda_0+\lambda_1)x_1^{\alpha-1}x_2^{\alpha-1}\exp\left[-(\lambda_0+\lambda_1)x_1^\alpha - \lambda_2 x_2^\alpha\right], \\ \qquad \text{if } x_1 > x_2, \\ \alpha^2\lambda_1(\lambda_0+\lambda_2)x_1^{\alpha-1}x_2^{\alpha-1}\exp\left[-(\lambda_0+\lambda_2)x_2^\alpha - \lambda_1 x_1^\alpha\right], \\ \qquad \text{if } x_1 < x_2, \\ \alpha\lambda_0(\lambda_0+\lambda_1+\lambda_2)^{-1}x^{\alpha-1}\exp\left[-(\lambda_0+\lambda_1+\lambda_2)x^\alpha\right], \\ \qquad \text{if } x_1 = x_2 = x, \end{cases}$$

which is not a valid joint pdf. This error might have been an oversight, but it appears to affect the results in Kundu and Gupta [3] including the estimation procedure, simulation study and real data application.

The aim of this note is not to correct the mistakes in Kundu and Gupta [3]. Instead we present a class of bivariate distributions more general than that introduced in Kundu and Gupta [3]. We show that this general class gives better fits to at least two real data sets, including the data considered in Kundu and Gupta [3]. We also argue that there is no real need for the EM algorithm considered in Kundu and Gupta [3].

The general class of bivariate distributions is proposed in Section 2. Estimation by the method of maximum likelihood is also discussed in Section 2. A simulation study comparing two different algorithms for computing the maximum likelihood estimates is presented in Section 3. Finally, Section 4 presents two real data applications of the general class.

2. New class of distributions

Marshall and Olkin [5] and Kundu and Gupta [3] restricted N to be a geometric random variable. We take N to be a power series random variable (truncated at zero) with probability mass function

$$\Pr(N = n) = \frac{a_n \theta^n}{C(\theta)}$$

for $n = 1, 2, \ldots$ and $0 < \theta < s$ for some s, where

$$C(\theta) = \sum_{n=1}^{\infty} a_n \theta^n < \infty$$

for some a_n and for all $0 < \theta < s$. The power series distribution (truncated at zero) contains many of the standard discrete distributions as particular cases: the binomial distribution (truncated at zero) for $C(\theta) = (\theta + 1)^m - 1$ and $\theta > 0$; the logarithmic distribution for $C(\theta) = -\ln(1-\theta)$ and $0 < \theta < 1$; the Poisson distribution (truncated at zero) for $C(\theta) = \exp(\theta) - 1$ and $\theta > 0$; the negative binomial distribution for $C(\theta) = (1-\theta)^{-m} - 1$ and $0 < \theta < 1$; and so on.

If N is a power series random variable then (1.1) generalizes to

$$\overline{F}_{Y_1,Y_2}(y_1, y_2) = \frac{C\left(\theta \overline{F}_{X_1,X_2}(y_1, y_2)\right)}{C(\theta)}.$$

If we take $\overline{F}_{X_1,X_2}(\cdot, \cdot)$ as that given by (1.2) then

$$\overline{F}_{Y_1,Y_2}(y_1, y_2) = \begin{cases} C\left(\theta \exp\left[-(\lambda_0 + \lambda_1) y_1^\alpha - \lambda_2 y_2^\alpha\right]\right)/C(\theta), \\ \qquad \text{if } y_1 \geq y_2, \\ C\left(\theta \exp\left[-(\lambda_0 + \lambda_2) y_2^\alpha - \lambda_1 y_1^\alpha\right]\right)/C(\theta), \\ \qquad \text{if } y_1 < y_2. \end{cases} \tag{2.1}$$

The corresponding survival functions of Y_1 and Y_2 are

$$\Pr(Y_1 > y_1) = C\left(\theta \exp\left[-(\lambda_0 + \lambda_1) y_1^\alpha\right]\right)/C(\theta)$$

and

$$\Pr(Y_2 > y_2) = C\left(\theta \exp\left[-(\lambda_0 + \lambda_2) y_2^\alpha\right]\right)/C(\theta).$$

The corresponding survival function of $\min(Y_1, Y_2)$ is

$$\Pr(Y_1 > y, Y_2 > y) = C\left(\theta \exp\left[-(\lambda_0 + \lambda_1 + \lambda_2) y^\alpha\right]\right)/C(\theta).$$

The corresponding joint probability density function of (Y_1, Y_2) is

$$\begin{aligned} f_{Y_1,Y_2}(y_1, y_2) &= \frac{\theta^2 C''\left(\theta \overline{F}_{X_1,X_2}(y_1, y_2)\right)}{C(\theta)} \frac{\partial \overline{F}_{X_1,X_2}(y_1, y_2)}{\partial y_1} \frac{\partial \overline{F}_{X_1,X_2}(y_1, y_2)}{\partial y_2} \\ &\quad + \frac{\theta C'\left(\theta \overline{F}_{X_1,X_2}(y_1, y_2)\right)}{C(\theta)} f_{X_1,X_2}(y_1, y_2), \end{aligned} \tag{2.2}$$

where $C'(\theta) = dC(\theta)/d\theta$, $C''(\theta) = d^2C(\theta)/d\theta^2$, $f_{X_1,X_2}(x_1, x_2)$ is given by (1.3),

$$\frac{\partial \overline{F}_{X_1,X_2}(x_1, x_2)}{\partial x_1} = \begin{cases} -\alpha(\lambda_0 + \lambda_1) x_1^{\alpha-1} \exp\left[-(\lambda_0 + \lambda_1) x_1^\alpha - \lambda_2 x_2^\alpha\right], \\ \qquad \text{if } x_1 > x_2, \\ -\alpha \lambda_1 x_1^{\alpha-1} \exp\left[-(\lambda_0 + \lambda_2) x_2^\alpha - \lambda_1 x_1^\alpha\right], \\ \qquad \text{if } x_1 < x_2, \\ -\alpha \lambda_0 x^{\alpha-1} \exp\left[-(\lambda_0 + \lambda_1 + \lambda_2) x^\alpha\right], \\ \qquad \text{if } x_1 = x_2 = x, \end{cases}$$

and

$$\frac{\partial \overline{F}_{X_1,X_2}(x_1,x_2)}{\partial x_2} = \begin{cases} -\alpha\lambda_2 x_2^{\alpha-1}\exp\left[-\left(\lambda_0+\lambda_1\right)x_1^{\alpha}-\lambda_2 x_2^{\alpha}\right], \\ \qquad \text{if } x_1 > x_2, \\ \\ -\alpha\left(\lambda_0+\lambda_2\right)x_2^{\alpha-1}\exp\left[-\left(\lambda_0+\lambda_2\right)x_2^{\alpha}-\lambda_1 x_1^{\alpha}\right], \\ \qquad \text{if } x_1 < x_2, \\ \\ -\alpha\lambda_0 x^{\alpha-1}\exp\left[-\left(\lambda_0+\lambda_1+\lambda_2\right)x^{\alpha}\right], \\ \qquad \text{if } x_1 = x_2 = x. \end{cases}$$

Moreover, $\Pr\left(Y_1 < Y_2\right) = \lambda_1/\left(\lambda_0+\lambda_1+\lambda_2\right)$, $\Pr\left(Y_1 > Y_2\right) = \lambda_2/\left(\lambda_0+\lambda_1+\lambda_2\right)$ and $\Pr\left(Y_1 = Y_2\right) = \lambda_0 \ / \ \left(\lambda_0+\lambda_1+\lambda_2\right)$. We shall refer to the distribution given by (2.1) and (2.2) as the Marshall Olkin Weibull (MOW)-*name* distribution, where *name* is the name of the distribution of N. For example, if N is a geometric random variable we shall refer to the distribution given by (2.1) and (2.2) as the MOW-geometric distribution, the distribution proposed in Kundu and Gupta [3].

If $\left\{\left(Y_{1,n}, Y_{2,n}\right), n = 1, 2, \ldots, n_0\right\}$ is a random sample on $\left(Y_1, Y_2\right)$ then the log-likelihood of $\left(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2\right)$ can be expressed as

$$\begin{aligned} \ln L \quad = \quad & \sum_{y_{1,i} > y_{2,i}} \ln\left[\alpha^2\left(\lambda_0+\lambda_1\right)\lambda_2 y_{1,i}^{\alpha-1} y_{2,i}^{\alpha-1}\right] \\ & + \sum_{y_{1,i} > y_{2,i}} \ln\left\{\theta\exp\left[-\left(\lambda_0+\lambda_1\right)y_{1,i}^{\alpha}-\lambda_2 y_{2,i}^{\alpha}\right]\right\} \\ & + \sum_{y_{1,i} > y_{2,i}} \ln\left\{\theta C^{''}\left(\omega_i\right)\exp\left[-\left(\lambda_0+\lambda_1\right)y_{1,i}^{\alpha}-\lambda_2 y_{2,i}^{\alpha}\right]\right. \\ & \qquad \left. + C^{'}\left(\omega_i\right)\right\} \\ & + \sum_{y_{1,i} < y_{2,i}} \ln\left[\alpha^2\left(\lambda_0+\lambda_2\right)\lambda_1 y_{1,i}^{\alpha-1} y_{2,i}^{\alpha-1}\right] \\ & + \sum_{y_{1,i} < y_{2,i}} \ln\left\{\theta\exp\left[-\left(\lambda_0+\lambda_2\right)y_{2,i}^{\alpha}-\lambda_1 y_{1,i}^{\alpha}\right]\right\} \\ & + \sum_{y_{1,i} < y_{2,i}} \ln\left\{\theta C^{''}\left(\omega_i\right)\exp\left[-\left(\lambda_0+\lambda_2\right)y_{2,i}^{\alpha}-\lambda_1 y_{1,i}^{\alpha}\right]\right. \\ & \qquad \left. + C^{'}\left(\omega_i\right)\right\} \\ & + \sum_{y_{1,i} = y_{2,i}} \ln\left[\alpha\lambda_0 y_{1,i}^{\alpha-1}\right] \\ & + \sum_{y_{1,i} = y_{2,i}} \ln\left\{\theta\exp\left[-\left(\lambda_0+\lambda_1+\lambda_2\right)y_{1,i}^{\alpha}\right]\right\} \\ & + \sum_{y_{1,i} = y_{2,i}} \ln\left\{\theta\alpha\lambda_0 y_{1,i}^{\alpha-1} C^{''}\left(\omega_i\right)\exp\left[-\left(\lambda_0+\lambda_1+\lambda_2\right)y_{1,i}^{\alpha}\right]\right. \\ & \qquad \left. + C^{'}\left(\omega_i\right)\right\} \\ & - n_0 \ln C(\theta), \end{aligned} \tag{2.3}$$

where $\omega_i = \theta\overline{F}_{X_1,X_2}\left(y_{1,i}, y_{1,i}\right)$. The maximum likelihood estimators of $\left(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2\right)$ are the parameter values maximizing (2.3). The maximization was performed using the nlm routine in the R software (R Development Core Team [6]). Extensive numerical computations showed that the surface of (2.3) was smooth for given smooth functions

$C(\cdot)$. The routine was able to locate the maximum of the likelihood surface for a wide range of smooth functions. The routine converged all the time. The solution for the maximum likelihood estimates was unique for a wide range of starting values. Hence, we did not feel the need to use the EM algorithm to find the maximum likelihood estimates.

However, a comparison of the use of the nlm routine and the EM algorithm is made in Section 3. Throughout Sections 3 and 4, the initial values for maximization were taken to correspond to all possible combinations of $\alpha = 0.01, 0.02, \ldots, 10$, $\lambda_0 = 0.01, 0.02, \ldots, 10$, $\lambda_1 = 0.01, 0.02, \ldots, 10$, $\lambda_2 = 0.01, 0.02, \ldots, 10$, $\theta = 0.01, 0.02, \ldots, 10$ for the Poisson and binomial cases and $\theta = 0.01, 0.02, \ldots, 0.99$ for the geometric, logarithmic and negative binomial cases.

3. A simulation study

Here, we assess the efficiency of the nlm routine in the R software in computing the maximum likelihood estimates. We also compare the efficiency with that of using the EM algorithm in Kundu and Gupta [3]. The efficiency is measured in terms of the central processing unit time to compute the maximum likelihood estimates.

We simulated ten thousand samples of size n from each of the MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial and MOW-negative binomial distributions. The following scheme was used:

(1) simulate n from the geometric, (truncated) Poisson, logarithmic, (truncated) binomial or the negative binomial distribution;
(2) simulate independently $U_1, U_2, \ldots, U_n$ from a Weibull distribution with shape parameter α and scale parameter λ_0, $V_1, V_2, \ldots, V_n$ from a Weibull distribution with shape parameter α and scale parameter λ_1 and W_1, W_2, $\ldots$, W_n from a Weibull distribution with shape parameter α and scale parameter λ_2;
(3) set $P_i = \min(U_i, V_i)$ and $Q_i = \min(U_i, W_i)$ for $i = 1, 2, \ldots, n$;
(4) set $Y_1 = \min(P_1, P_2, \ldots, P_n)$ and $Y_2 = \min(Q_1, Q_2, \ldots, Q_n)$.

The maximum likelihood estimates were computed for each of the ten thousand samples by using the nlm routine and the EM algorithm. This process was repeated for $n = 15, 16, \ldots, 100$.

The mean central processing unit time (over the ten thousand samples) versus n is plotted in Figure 1 for each MOW type distribution. We see that the use of the nlm routine is more efficient for every n and for every MOW type distribution. The relative efficiency of using the nlm routine increases with increasing n.

In the simulations, we took the true parameter values to be: $\alpha = 2$, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$ and $\theta = 0.5$ for the MOW-geometric distribution; $\alpha = 2$, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$ and $\theta = 1$ for the MOW-Poisson distribution; $\alpha = 2$, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$ and $\theta = 0.5$ for the MOW-logarithmic distribution; $\alpha = 2$, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$, $\theta = 1$ and $m = 10$ for the MOW-binomial distribution; $\alpha = 2$, $\lambda_0 = 1$, $\lambda_1 = 1$, $\lambda_2 = 1$, $\theta = 0.5$ and $m = 10$ for the MOW-negative binomial distribution. But the pattern in Figure 1 was the same for a wide range of other parameter values. In particular, the nlm routine was always more efficient for every n and for every MOW type distribution and the relative efficiency of using the nlm routine always increased with increasing n.

4. Data applications

Here, we use two real data sets to illustrate the fits of MOW type distributions. The first data set is the same as that used in Kundu and Gupta [3]. The data are on two variables: Y_1 = the time in minutes of the first goal kick scored by any team and Y_2 = the time in minutes of the first goal scored by the home team. The data on Y_1 are 26,

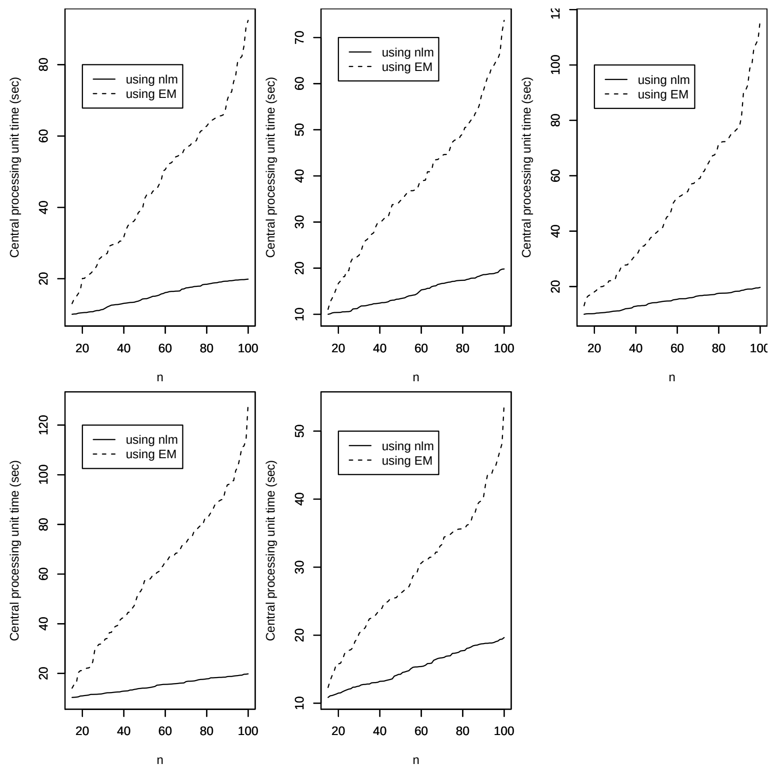

Figure 1. The mean central processing unit time (over the ten thousand samples) versus n for the MOW-geometric distribution (top left), MOW-Poisson distribution (top middle), MOW-logarithmic distribution (top right), MOW-binomial distribution (bottom left) and the MOW-negative binomial distribution (bottom middle).

63, 19, 66, 40, 49, 8, 69, 39, 82, 72, 66, 25, 41, 16, 18, 22, 42, 2, 36, 34, 53, 54, 51, 76, 64, 26, 16, 44, 25, 55, 49, 24, 44, 42, 27, 28. The data on Y_2 are 20, 18, 19, 85, 40, 49, 8, 71, 39, 48, 72, 62, 9, 3, 75, 18, 14, 42, 2, 52, 34, 39, 7, 28, 64, 15, 48, 16, 13, 14, 11, 49, 24, 30, 3, 47, 28. See Kundu and Gupta [3] for details.

The following distributions were fitted to the data: MOW, MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial m for $m = 2, 3, \ldots, 15$, MOW-negative binomial m for $m = 2, 3, \ldots, 15$. Throughout, MOW refers to the distribution given by (1.2) and (1.3). The method of maximum likelihood was used to fit each distribution, i.e., the parameter values are those minimizing the negative of (2.3). Table 1 gives the following for the fitted distributions: the negative log-likelihood, Akaike information criterion (AIC) due to Akaike [1] and the Bayesian information criterion (BIC) due to Schwarz [7]. Table 2 gives the following for the fitted distributions: the Kolmogorov-Smirnov statistic comparing the fitted and empirical cumulative distribution functions of Y_1, the corresponding p-value, the Kolmogorov-Smirnov statistic comparing the fitted and empirical cumulative distribution functions of Y_2, the corresponding p-value, the Kolmogorov-Smirnov statistic comparing the fitted and empirical cumulative distribution functions of $\min(Y_1, Y_2)$ and the corresponding p-value. The Kolmogorov-Smirnov

Model	$-\ln L$	AIC	BIC
MOW	337.153	682.306	688.750
MOW-geo	305.172	620.344	628.399
MOW-Poisson	307.751	625.502	633.557
MOW-log	292.073	594.146	602.201
MOW-bin $m = 2$	310.037	630.074	638.129
MOW-bin $m = 3$	307.604	625.208	633.263
MOW-bin $m = 4$	284.022	578.044	586.099
MOW-bin $m = 5$	302.503	615.006	623.061
MOW-bin $m = 6$	297.923	605.846	613.901
MOW-bin $m = 7$	289.974	589.948	598.003
MOW-bin $m = 8$	288.967	587.934	595.989
MOW-bin $m = 9$	284.888	579.776	587.831
MOW-bin $m = 10$	273.864	557.728	565.783
MOW-bin $m = 11$	273.197	556.394	564.449
MOW-bin $m = 12$	263.481	536.962	545.017
MOW-bin $m = 13$	255.636	521.272	529.327
MOW-bin $m = 14$	252.354	514.708	522.763
MOW-bin $m = 15$	247.437	504.874	512.929
MOW-neg bin $m = 2$	307.033	624.066	632.121
MOW-neg bin $m = 3$	301.347	612.694	620.749
MOW-neg bin $m = 4$	299.947	609.894	617.949
MOW-neg bin $m = 5$	292.352	594.704	602.759
MOW-neg bin $m = 6$	283.846	577.692	585.747
MOW-neg bin $m = 7$	271.521	553.042	561.097
MOW-neg bin $m = 8$	262.389	534.778	542.833
MOW-neg bin $m = 9$	251.878	513.756	521.811
MOW-neg bin $m = 10$	238.214	486.428	494.483
MOW-neg bin $m = 11$	224.912	459.824	467.879
MOW-neg bin $m = 12$	214.145	438.29	446.345
MOW-neg bin $m = 13$	200.722	411.444	419.499
MOW-neg bin $m = 14$	192.481	394.962	403.017
MOW-neg bin $m = 15$	180.946	371.892	379.947

Table 1. Log-likelihood, AIC and BIC values for the MOW type distributions fitted to the data used in Kundu and Gupta (2014).

statistic was computed by

$$D_n = \sqrt{n} \sup_x \left| F_n(x) - F(x) \right|,$$

where F_n denotes the empirical cumulative distribution function and F denotes the fitted cumulative distribution function. The corresponding p-value was computed by using the fact that D_n converges in distribution as $n \to \infty$ to the random variable K with the cumulative distribution function

$$\Pr\left(K \leq x\right) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} \exp\left(-2k^2 x^2\right)$$

for $x > 0$.

According to the p-values, none of the fitted distributions provide satisfactory fits. However, the MOW type distributions improve substantially on the fit of the MOW and

	Y_1		Y_2		$\min(Y_1, Y_2)$	
Model	K-S statistic	p-value	K-S statistic	p-value	K-S statistic	p-value
MOW	0.449	0.000	0.646	0.000	0.646	0.000
MOW-geo	0.626	0.000	0.336	0.000	0.322	0.001
MOW-Poisson	0.604	0.000	0.341	0.000	0.312	0.001
MOW-log	0.611	0.000	0.343	0.000	0.310	0.001
MOW-bin $m = 2$	0.592	0.000	0.338	0.000	0.315	0.001
MOW-bin $m = 3$	0.565	0.000	0.320	0.001	0.337	0.000
MOW-bin $m = 4$	0.534	0.000	0.299	0.002	0.365	0.000
MOW-bin $m = 5$	0.567	0.000	0.304	0.002	0.582	0.000
MOW-bin $m = 6$	0.649	0.000	0.338	0.000	0.640	0.000
MOW-bin $m = 7$	0.577	0.000	0.328	0.000	0.425	0.000
MOW-bin $m = 8$	0.498	0.000	0.369	0.000	0.480	0.000
MOW-bin $m = 9$	0.558	0.000	0.406	0.000	0.495	0.000
MOW-bin $m = 10$	0.453	0.000	0.451	0.000	0.556	0.000
MOW-bin $m = 11$	0.454	0.000	0.493	0.000	0.588	0.000
MOW-bin $m = 12$	0.829	0.000	0.527	0.000	0.812	0.000
MOW-bin $m = 13$	0.418	0.000	0.575	0.000	0.658	0.000
MOW-bin $m = 14$	0.433	0.000	0.613	0.000	0.684	0.000
MOW-bin $m = 15$	0.362	0.000	0.649	0.000	0.721	0.000
MOW-neg bin $m = 2$	0.593	0.000	0.308	0.001	0.594	0.000
MOW-neg bin $m = 3$	0.591	0.000	0.297	0.002	0.393	0.000
MOW-neg bin $m = 4$	0.760	0.000	0.440	0.000	0.735	0.000
MOW-neg bin $m = 5$	0.593	0.000	0.409	0.000	0.485	0.000
MOW-neg bin $m = 6$	0.578	0.000	0.480	0.000	0.546	0.000
MOW-neg bin $m = 7$	0.574	0.000	0.556	0.000	0.608	0.000
MOW-neg bin $m = 8$	0.557	0.000	0.632	0.000	0.673	0.000
MOW-neg bin $m = 9$	0.489	0.000	0.705	0.000	0.740	0.000
MOW-neg bin $m = 10$	0.468	0.000	0.764	0.000	0.789	0.000
MOW-neg bin $m = 11$	0.421	0.000	0.814	0.000	0.839	0.000
MOW-neg bin $m = 12$	0.377	0.000	0.865	0.000	0.886	0.000
MOW-neg bin $m = 13$	0.413	0.000	0.904	0.000	0.919	0.000
MOW-neg bin $m = 14$	0.445	0.000	0.932	0.000	0.941	0.000
MOW-neg bin $m = 15$	0.522	0.000	0.951	0.000	0.957	0.000

Table 2. Goodness of fit measures for the MOW type distributions fitted to the data used in Kundu and Gupta (2014).

MOW-geometric distributions with respect to likelihood and K-S statistic values: the MOW-negative binomial $m = 15$ distribution gives the smallest negative log-likelihood value, the smallest AIC value and the smallest BIC value; the MOW-binomial $m = 15$ distribution gives the smallest K-S statistic with respect to Y_1; the MOW-negative binomial $m = 3$ distribution gives the smallest K-S statistic with respect to Y_2; the MOW-logarithmic distribution gives the smallest K-S statistic with respect to $\min(Y_1, Y_2)$.

The values reported in Table 1 for the MOW-geometric distribution are different from those reported in Kundu and Gupta [3]. Our parameter estimates for the MOW-geometric distribution, $\widehat{\theta} = 9.549 \times 10^{-2}$, $\left(8.453 \times 10^{-2}\right)$, $\widehat{\alpha} = 4.061 \times 10^{-1}$ $\left(4.775 \times 10^{-2}\right)$, $\widehat{\lambda_0} = 7.471 \times 10^{-8}$ $\left(3.405 \times 10^{-2}\right)$, $\widehat{\lambda_1} = 7.940 \times 10^{-2}$ $\left(3.047 \times 10^{-2}\right)$ and $\widehat{\lambda_2} = 1.609 \times 10^{-1}$ $\left(4.324 \times 10^{-2}\right)$, where the numbers in parentheses are standard errors obtained by inverting the observed information matrix, are also different from the estimates reported in Kundu and Gupta [3]. These discrepancies are probably due to the errors in Kundu and Gupta [3] that we pointed out in Section 1.

The second data set is breaking strengths of fibers of lengths 12mm and 30mm. The data were taken from Table 7.2, data set 3, page 144 of Crowder et al. [2]. The data are on the two variables: Y_1 = breaking strength for fibers of length 12mm and Y_2 = breaking strength for fibers of length 30mm. The data on Y_1 are 4, 3.98, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 3.54, 3.4, 3.28, 3.02, 3.06, 3.2, 3.96, 3.73, 3.7, 4, 3.83, 3.92, 4, 4. The data on Y_2 are 4, 3.2, 4, 4, 4, 3.82, 3.4, 4, 3.65, 4, 4, 4, 3.09, 2.32, 2.18, 2.14, 2.14, 2.3, 3.18, 3.22, 3.28, 3.2, 3.16, 3.25, 2.16, 3.22. We have chosen this data in particular because Weibull distributions are most popular models for breaking strength. The results of the fit of the following distributions to this data are shown in Tables 3 and 4: MOW, MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial m for $m = 2, 3, \ldots, 15$, MOW-negative binomial m for $m = 2, 3, \ldots, 15$. According to the p-values with respect to Y_1, none of these distributions provide satisfactory fits. According to the p-values with respect to Y_2 and $\min(Y_1, Y_2)$, all of these distributions provide satisfactory fits. Many of the distributions improve on the fit of the MOW and MOW-geometric distributions:

- all of the fitted distributions have significantly smaller negative log-likelihood values, smaller AIC values and smaller BIC values than those for MOW;
- MOW-logarithmic, MOW-binomial m for $m = 4, 5, \ldots, 15$, MOW-negative binomial m for $m = 2, 3, \ldots, 15$ all have smaller negative log-likelihood values, smaller AIC values and smaller BIC values than those for MOW-geometric;
- MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial m for $m = 2, 3, \ldots, 6$, MOW-negative binomial m for $m = 2, 3, 4$ all have larger p-values with respect to Y_1 than that for MOW;
- MOW-binomial $m = 4$, MOW-binomial $m = 5$ and MOW-negative binomial $m = 2$ all have larger p-values with respect to Y_1 than that for MOW-geometric;
- MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial m for $m = 2, 3, \ldots, 8$, MOW-negative binomial m for $m = 2, 3, \ldots, 6$ all have larger p-values with respect to Y_2 than that for MOW;
- MOW-binomial m for $m = 4, 5, \ldots, 7$, MOW-negative binomial m for $m = 2, 3, \ldots, 5$ all have larger p-values with respect to Y_2 than that for MOW-geometric;
- MOW-geometric, MOW-Poisson, MOW-logarithmic, MOW-binomial m for $m = 2, 3, \ldots, 6$, MOW-negative binomial m for $m = 2, 3, 4$ all have larger p-values with respect to $\min(Y_1, Y_2)$ than that for MOW;
- MOW-binomial $m = 4$, MOW-binomial $m = 5$, MOW-negative binomial $m = 2$ and MOW-negative binomial $m = 3$ all have larger p-values with respect to $\min(Y_1, Y_2)$ than that for MOW-geometric.

Model	$-\ln L$	AIC	BIC
MOW	52.385	112.77	117.802
MOW-geo	-67.015	-124.03	-117.74
MOW-Poisson	-21.945	-33.89	-27.560
MOW-log	-21.945	-33.89	-27.560
MOW-bin $m = 2$	-21.945	-33.89	-27.560
MOW-bin $m = 3$	-66.057	-122.114	-115.824
MOW-bin $m = 4$	-70.257	-130.514	-124.224
MOW-bin $m = 5$	-70.978	-131.956	-125.666
MOW-bin $m = 6$	-71.02	-132.04	-125.75
MOW-bin $m = 7$	-71.048	-132.096	-125.806
MOW-bin $m = 8$	-71.067	-132.134	-125.844
MOW-bin $m = 9$	-71.081	-132.162	-125.872
MOW-bin $m = 10$	-71.094	-132.188	-125.898
MOW-bin $m = 11$	-71.102	-132.204	-125.914
MOW-bin $m = 12$	-71.111	-132.222	-125.932
MOW-bin $m = 13$	-71.118	-132.236	-125.946
MOW-bin $m = 14$	-71.124	-132.248	-125.958
MOW-bin $m = 15$	-71.129	-132.258	-125.968
MOW-neg bin $m = 2$	-71.293	-132.586	-126.296
MOW-neg bin $m = 3$	-71.399	-132.798	-126.508
MOW-neg bin $m = 4$	-71.357	-132.714	-126.424
MOW-neg bin $m = 5$	-71.328	-132.656	-126.366
MOW-neg bin $m = 6$	-71.308	-132.616	-126.326
MOW-neg bin $m = 7$	-71.291	-132.582	-126.292
MOW-neg bin $m = 8$	-71.282	-132.564	-126.274
MOW-neg bin $m = 9$	-71.271	-132.542	-126.252
MOW-neg bin $m = 10$	-71.266	-132.532	-126.242
MOW-neg bin $m = 11$	-71.256	-132.512	-126.222
MOW-neg bin $m = 12$	-71.254	-132.508	-126.218
MOW-neg bin $m = 13$	-71.247	-132.494	-126.204
MOW-neg bin $m = 14$	-71.243	-132.486	-126.196
MOW-neg bin $m = 15$	-71.238	-132.476	-126.186

Table 3. Log-likelihood, AIC and BIC values for the MOW type distributions fitted to the fiber data.

Acknowledgments

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

	Y_1		Y_2		$\min(Y_1, Y_2)$	
Model	K-S statistic	p-value	K-S statistic	p-value	K-S statistic	p-value
MOW	0.286	0.022	0.172	0.382	0.171	0.389
MOW-geo	0.278	0.029	0.170	0.398	0.169	0.406
MOW-Poisson	0.284	0.024	0.171	0.389	0.170	0.396
MOW-log	0.284	0.024	0.171	0.390	0.170	0.397
MOW-bin $m = 2$	0.284	0.024	0.171	0.389	0.170	0.396
MOW-bin $m = 3$	0.279	0.028	0.170	0.397	0.169	0.405
MOW-bin $m = 4$	0.273	0.033	0.169	0.406	0.168	0.414
MOW-bin $m = 5$	0.277	0.030	0.167	0.415	0.167	0.414
MOW-bin $m = 6$	0.282	0.025	0.166	0.424	0.170	0.394
MOW-bin $m = 7$	0.288	0.021	0.168	0.410	0.173	0.374
MOW-bin $m = 8$	0.294	0.018	0.171	0.390	0.176	0.356
MOW-bin $m = 9$	0.299	0.015	0.174	0.370	0.178	0.338
MOW-bin $m = 10$	0.304	0.012	0.176	0.352	0.181	0.321
MOW-bin $m = 11$	0.310	0.010	0.179	0.334	0.184	0.305
MOW-bin $m = 12$	0.315	0.009	0.182	0.318	0.186	0.290
MOW-bin $m = 13$	0.321	0.007	0.184	0.302	0.189	0.276
MOW-bin $m = 14$	0.326	0.006	0.187	0.287	0.191	0.262
MOW-bin $m = 15$	0.331	0.005	0.189	0.272	0.194	0.249
MOW-neg bin $m = 2$	0.272	0.034	0.168	0.408	0.167	0.415
MOW-neg bin $m = 3$	0.278	0.029	0.167	0.417	0.168	0.410
MOW-neg bin $m = 4$	0.284	0.024	0.166	0.427	0.171	0.390
MOW-neg bin $m = 5$	0.290	0.020	0.169	0.405	0.174	0.370
MOW-neg bin $m = 6$	0.295	0.017	0.172	0.384	0.177	0.351
MOW-neg bin $m = 7$	0.301	0.014	0.174	0.364	0.179	0.333
MOW-neg bin $m = 8$	0.307	0.011	0.177	0.346	0.182	0.316
MOW-neg bin $m = 9$	0.313	0.009	0.180	0.328	0.185	0.299
MOW-neg bin $m = 10$	0.318	0.008	0.183	0.310	0.187	0.284
MOW-neg bin $m = 11$	0.324	0.006	0.186	0.294	0.190	0.269
MOW-neg bin $m = 12$	0.329	0.005	0.188	0.279	0.193	0.255
MOW-neg bin $m = 13$	0.335	0.004	0.191	0.264	0.195	0.242
MOW-neg bin $m = 14$	0.340	0.003	0.193	0.251	0.198	0.229
MOW-neg bin $m = 15$	0.346	0.003	0.196	0.237	0.200	0.218

Table 4. Goodness of fit measures for the MOW type distributions fitted to the fiber data.

References

[1] Akaike, H. *A new look at the statistical model identification*, IEEE Transactions on Automatic Control **19**, 716-723, 1974.

[2] Crowder, M.J., Kimber, A., Sweeting, T. and Smith, R. *Statistical Analysis of Reliability Data*, Chapman and Hall, London, 1994.

[3] Kundu, D. and Gupta, A.K. *On bivariate Weibull-geometric distribution*, Journal of Multivariate Analysis **123**, 19-29, 2014.

[4] Marshall, A.W. and Olkin, I. *A multivariate exponential distribution*, Journal of the American Statistical Association **62**, 30-44, 1967.

[5] Marshall, A.W. and Olkin, I. *A new method of adding a parameter to a family of distributions with application to the exponential and Weibull families*, Biometrika **84**, 641-652, 1997.

[6] R Development Core Team. *A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing. Vienna, Austria, 2015.

[7] Schwarz, G.E. *Estimating the dimension of a model*, Annals of Statistics **6**, 461-464, 1978.