A MULTIVARIATE WEIBULL DISTRIBUTION

DAVID D. HANAGAL

Department of Statistics, University of Pune,

Pune-411007, INDIA.

Visiting Professor, Colegio de Postgraduados, ISEI,

Montecillo, Texcoco, CP56230, MEXICO.

ABSTRACT

In this paper, we introduce a new multivariate Weibull (MVW) distribution with many interesting properties. We obtain the maximum likelihood estimate (MLE) of the parameters and their asymptotic multivariate normal (AMVN) distribution in MVW model. We propose large sample studentized tests for testing multivariate exponentiality and also tests for independence and identical marginals of the components.

Key words and Phrases: *Independence, Maximum likelihood estimate, Multivariate weibull model, Multivariate exponentiality, Symmetry.*

1. INTRODUCTION

In reliability theory and life testing experiments, weibull distribution plays an important role. It reduces to exponential distribution when the shape parameter equals one. Weibull distribution has increasing failure rate (IFR) when the shape parameter is greater than one and has decreasing failure rate (DFR) when shape parameter is less than one. The extension of univariate Weibull distribution to multivariate case is desirable in view of the crucial role that Weibull distribution plays in reliability as well as building models for various failure or life time distribution. However, in the exponential case, there does not exist a unique natural extension of the univariate exponential distribution to bivariate or multivariate case. So, we have many bivariate or multivariate extensions of univariate exponential distribution. [See Weinmann(1966), Block(1975) and Hanagal(1993a,1993b)]. In the similar way, we have many bivariate or multivariate Weibull distributions based on these bivariate or multivariate exponential distributions. In this paper we consider the multivariate Weibull (MVW) distribution which can be obtained from multivariate exponential (MVE) model of Marshall - Olkin (1967). This is the only MVE having the marginals as exponentials and this is the main reason to choose this particular MVE model to obtain MVW model.

If $\underline{Y} = (Y_1, ..., Y_n)$ is $(k+1)$ parameter version of MVE distribution of Marshall - Olkin (1967) as stated in Proschan - Sullo (1976) and Hanagal (1991), then by taking the transformation $X_i = Y_i^{1/c}, c > 0, i = 1, ..., k$, we have $\underline{X} = (X_1, ..., X_k)$ follow MVW model which contains singularities. The above transformation can also be done to $(2^k - 1)$ parameter version of MVE of Marshall - Olkin (1967) and from that we obtain 2^k parameter version of MVW model. But we are not interested in 2^k parameter version of MVW and so, we study only $(k+2)$ prameter version of MVW model.

In Section 2, we obtain MVW model and present some interesting properties. In Section 3, we obtian MLEs of the parameters of MVW. In the last Section, we develop large sample studentized test for testing multivariate exponentiality and also test for independence and symmetry or identical marginals of the components.

2. MULTIVARIATE WEIBULL MODEL AND ITS PROPERTIES

The survival function of $\underline{Y}$ of MVE of Marshall - Olkin (1967) is

$$\begin{aligned} \bar{F}_{\underline{Y}}(\underline{y}) &= P[Y_1 > y_1, ..., Y_k > y_k] \\ &= exp[-\lambda_1 y_1 - ... - \lambda_k y_k - \lambda_0 Max(y_1, ..., y_k)] \end{aligned}$$

where $\lambda_0, ..., \lambda_k > 0$. Taking the transformation $X_i = Y_i^{1/c}, c > 0, i = 1, ..., k$, we get the corresponding survival function of $\underline{X}$ of MVW which is given by

$$\begin{aligned} \bar{F}_{\underline{X}}(\underline{x}) &= P[X_1 > x_1, ..., X_k > x_k] \\ &= exp\{-\lambda_1 x_1^c - \lambda_2 x_2^c - ... - \lambda_k x_k^c - \lambda_0 \{Max(x_1, ..., x_k)\}^c\}. \end{aligned}$$

The above MVW model is not absolutely continuous with respect to Lebesgue measure on R^k. As MVW distribution is failure time distribution and derived from MVE of Marshall-Olkin(1967), all real life applications of MVE of Marshall-Olkin(1967) will become real life applications of this proposed MVW. For e.g., simultaneous failure of nuclear power stations, simultaneous failure of hydroelectric pumps in aeroplane etc.

The marginal of $X_i, i = 1, ..., k$ are obtained as

$$P[X_i > x_i] = \bar{F}_{\underline{x}}(0, ..., x_i, 0, ..., 0)$$
$$= exp\{-(\lambda_i + \lambda_0)x_i^c\}, i = 1, ..., k$$

which is the survival function of Weibull with parameters $(\lambda_i + \lambda_0, c)$, i=1,...,k.

The distribution of $Min(X_1, ..., X_k)$ is obtained by

$$P[Min(X_1, ..., X_k) > x] = \overline{F}_{\underline{X}}(x, ..., x)$$
$$= exp\{-\lambda x^c\}, \quad \lambda = \lambda_0 + \lambda_1 + ... + \lambda_k$$

which is the survival function of Weibull with parameters (λ, c).

The random variables $X_i, i = 1, ..., k$ are independent iff $\lambda_0 = 0$ and $X_i, i = 1, ..., k$ are identically distributed iff $\lambda_1 = \lambda_2 = ... = \lambda_k$. The probability that all $X_i, i = 1, ..., k$ are equal to each other is $P[X_1 = X_2 = ... = X_k] = \lambda_0/\lambda$. This MVW model has IFR when $c > 1$ and DFR when $c < 1$.

3. ESTIMATION OF THE PARAMETERS

In this section, we obtain the MLEs of the parameters of MVW model. Let $(x_{1j}, x_{2j}, ..., x_{kj}), j = 1, ..., n$ be i.i.d. random observations of sample of size n. Now we see that there are some similarities in writing the likelihood of MVW and MVE of Marshall - Olkin [See Proschan - Sullo (1976)]. The likelihood of the sample of size n is

$$L = c^p \lambda_0^{n_0} \left[\prod_{i=1}^{k} \lambda_i^{n_i} (\lambda_i + \lambda_0)^{n_i(e)} \prod_{j=1}^{n} x_{ij}^{(c-1)}\right] \left[\prod_{r=2}^{k} \prod_{j \epsilon S_r} x_{(k)j}^{-(r-1)(c-1)}\right]$$
$$exp\{-\sum_{i=1}^{k} \lambda_i \sum_{j=1}^{n} x_{ij}^c - \lambda_0 \sum_{j=1}^{n} x_{(k)j}^c\}$$

where $p = [nk - \sum_{r=2}^{k}(r-1)n_0(r)], n_0(r)$ = number of observations with r of $X_i's, (i = 1, ..., k)$ are equal, $n_0 = \sum_{r=2}^{k} n_0(r), n_i$ = number of observations in which the random variable $X_i < X_{(k)}, n_i(e)$ = number of observations with X_i is strictly the maximum of the $(X_1, ..., X_k)$,

$$S_r = \{X_{i_1} = X_{i_2} = ... = X_{i_r} = X_{(k)}, i_1 \neq i_2 \neq ... \neq i_r = 1, ..., k\}$$

and $X_{(k)} = Max(X_1, ..., X_k)$.

The loglikelihood of the sample of size n is given by

$$\begin{aligned} \mathrm{logL} &= p\mathrm{logc} + \mathrm{n_0log}\lambda_0 + \sum_{\mathrm{i=1}}^{\mathrm{k}} \lambda_\mathrm{i}\mathrm{log}\lambda_\mathrm{i} \\ &+ \sum_{i=1}^{k} n_i(e)\mathrm{log}(\lambda_\mathrm{i} + \lambda_0) + \sum_{\mathrm{i=1}}^{\mathrm{k}}(\mathrm{c}-1)\sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{logx_{ij}} \\ &- \sum_{r=2}^{k}(r-1)(c-1)\sum_{j\epsilon S_r} \mathrm{logx_{(k)j}} - \sum_{\mathrm{i=1}}^{\mathrm{k}} \lambda_\mathrm{i} \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{x_{ij}^c} - \lambda_0 \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{x_{(k)j}^c} \end{aligned}$$

The expected values of $n_i, n_0, n_i(e)$ and $n_0(r)$ are

$E(n_i) = n\lambda_i(1-\phi_i)/(\lambda_i + \lambda_0), i = 1, ..., k$

$E(n_0) = n(1 - \sum_{i=1}^{k} \phi_i)$

$E(n_i(e)) = n\phi_i, i = 1, ..., k$

and

$$E[n_0(r)] = \sum_{i_1\neq} ... \sum_{\neq i_k=1}^{k} \frac{\lambda_{i_1}, ..., \lambda_{i_{k-r}}\lambda_0}{(\lambda_{i_{k-r+1}} + ... + \lambda_{i_k} + \lambda_0)...(\lambda)},$$

$$i_1 \neq ... \neq i_k = 1, ..., k$$

where

$$\begin{aligned} \phi_{i_1} &= P[X_{i_1} > Max(X_{i_2}, ..., X_{i_k})], i_1 \neq ... \neq i_k = 1, ..., k \\ &= \sum_{i_1\neq} .... \sum_{\neq i_k=1}^{k} \frac{\lambda_{i_2}\lambda_{i_3}...\lambda_{i_k}}{(\lambda_{i_1} + \lambda_{i_2} + \lambda_0)...(\lambda)}. \end{aligned}$$

The likelihood equations with respect to the parameters $(\lambda_0, \lambda_1, ..., \lambda_k, c)$ are

$$n_0/\lambda_0 + \sum_{i=1}^{k} n_i(e)/(\lambda_i + \lambda_0) - \sum_{j=1}^{n} x^c_{(k)j} = 0$$

$$n_i/\lambda_i + n_i(e)/(\lambda_i + \lambda_0) - \sum_{j=1}^{n} x^c_{ij} = 0, i = 1, ..., k$$

$$\frac{p}{c} + \sum_{i=1}^{k}\sum_{j=1}^{n} \mathrm{logx_{ij}} - \sum_{\mathrm{r=2}}^{\mathrm{k}} (\mathrm{r} - 1) \sum_{\mathrm{j\epsilon S_r}} \mathrm{logx_{(k)j}}$$

$$- \sum_{i=1}^{k} \lambda_i \sum_{j=1}^{n} x^c_{ij}\mathrm{logx_{ij}} - \lambda_0 \sum_{\mathrm{j=1}}^{\mathrm{n}} \mathrm{x^c_{(k)}logx_{(k)j}} = 0.$$

The likelihood equations are not easy to solve. So one can generate some consistent estimators say $(u_0, ..., u_{k+1})$ of $\underline{\lambda} = (\lambda_0, ..., \lambda_k, c)$ and use $(u_0, ..., u_{k+1})$ as a trial solution in Newton - Raphson procedure or Fisher's method of scoring to obtain MLEs $\underline{\hat{\lambda}} = (\hat{\lambda}_0, \hat{\lambda}_1, ..., \hat{\lambda}_k, \hat{c})$.

So, we choose the consistent estimators $(u_0, ..., u_{k+1})$ of $\underline{\lambda} = (\lambda_0, ..., \lambda_k, c)$ as

$$u_i = r_i / \sum_{j=1}^{n} x^c_{(1)j}, i = 0, 1, ..., k$$

$$u_{k+1} = \frac{\pi}{\sqrt{6}} [\frac{1}{n} \sum_{j=1}^{n} (\mathrm{logx_{(1)j}} - \overline{\mathrm{logx_{(1)}}})^2]^{-1/2}$$

where r_i, i = 1, ..., k be the number of observations with $x_{ij} < Min_{\ell \neq i}(x_{\ell_j})$, $\ell \neq i = 1, ..., k$ and r_0 be the number of observations with $x_{1j} = ... = x_{kj}$ in the sample of size n, $\sum_{i=0}^{k} r_i = n$, $x_{(1)j} = Min(x_{1j}, ..., x_{kj})$, and $\overline{\mathrm{logx_{(1)}}} = \frac{1}{n}\sum_{j=1}^{n} \mathrm{logx_{(1)j}}$. The distribution of $(r_1, ..., r_k)$ is multinomial with parameters $(n, \lambda_1/\lambda, ..., \lambda_k/\lambda)$ and the distribution of $x^c_{(1)j}$ is exponential with failure rate λ and it is easy to check that $u_i \xrightarrow{P} \lambda_i, i = 0, 1, .., k, u_{k+1} \xrightarrow{P} c$. Here the initial estimator u_{k+1} is obtained by the expression $Var(\mathrm{logX_{(1)}}) = (\pi^2/6)\mathrm{c}^{-2}$.

The Fisher information matrix is $nI(\underline{\lambda}) = n(I_{ij}))$ where

$$I_{00} = [1 - \sum_{i=1}^{k} \phi_i]/\lambda_0^2 + \sum_{i=1}^{k} \phi_i/(\lambda_i + \lambda_0)^2$$

$$I_{ii} = [1 - \phi_i]/[\lambda_i(\lambda_i + \lambda_0)] + \phi_i/(\lambda_i + \lambda_0)^2, i = 1, ..., k,$$

$$I_{i0} = \phi_i/(\lambda_i + \lambda_0)^2, i = 1, ..., k, \qquad I_{ij} = 0, i \neq j = 1, ..., k$$

$$I_{cc} = E(p)/[nc^2] + \sum_{i=1}^{k} \lambda_i E[x_{ij}^c (\mathrm{log x_{ij}})^2] + \lambda_0 \mathrm{E}[\mathrm{x^c_{(k)j}}(\mathrm{log x_{(k)_j}})^2]$$

$$I_{ic} = E[x_{ij}^c \mathrm{log x_{ij}}], \mathrm{i = 1, ..., k}, \qquad \mathrm{I_{0c} = E[x^c_{(k)j} log x_{(k)j}]}$$

where

$$E(p) = nk - \sum_{r=2}^{k} (r-1) E[n_0(r)],$$

$$E[x_{ij}^c (\mathrm{log x_{ij}})^2] = [[\log(\lambda_\mathrm{i} + \lambda_0) - \psi(2)]^2 + \psi'(1) - 1]/[(\lambda_\mathrm{i} + \lambda_0)\mathrm{c}^2], \mathrm{i = 1, ..., k}$$

$$E[x_{ij}^c \mathrm{log x_{ij}}] = [\psi(2) - \log(\lambda_\mathrm{i} + \lambda_0)]/[(\lambda_\mathrm{i} + \lambda_0)\mathrm{c}], \mathrm{i = 1, ..., k}$$

$$E[x_{(k)j}^c (\mathrm{log x_{(k)j}})]$$

$$= \frac{1}{c}\left[\sum_{r=1}^{k} (-1)^{r+1} \sum_{i_1<} ... \sum_{<i_r=1:k} \frac{[\psi(2) - \log(\lambda_{\mathrm{i_1}} + .. + \lambda_{\mathrm{i_r}} + \lambda_0)]}{(\lambda_{i_1} + ... + \lambda_{i_r} + \lambda_0)}\right]$$

$$E[x_{(k)j}^c (\mathrm{log x_{(k)j}})^2]$$

$$= \frac{1}{c^2}\left[\sum_{r=1}^{k} (-1)^{r+1} \sum_{i_1<} ... \sum_{<i_r=1:k} \frac{[[\log(\lambda_{\mathrm{i_1}} + ... + \lambda_{\mathrm{i_r}} + \lambda_0) - \psi(2)]^2 + \psi'(1) - 1]}{(\lambda_{i_1} + ... + \lambda_{i_r} + \lambda_0)}\right]$$

where $\psi(\cdot)$ and $\psi'(\cdot)$ are digamma and trigamma functions and are given by

$$\psi(2) = \frac{\partial \mathrm{log \Gamma z}}{\partial z}|_{z=2} = 0.422785,$$

$$\psi'(1) = \frac{\partial^2 \mathrm{log \Gamma z}}{\partial z^2}|_{z=1} = \pi^2/6 = 1.64474.$$

The above Fisher information matrix $nI(\underline{\lambda})$ is positive definite and using multivariate central limit theorem, $\sqrt{n}(\hat{\underline{\lambda}} - \underline{\lambda})$ has an asymptotic multivariate normal (MVN) with mean vector zero and variance - covariance matrix $I^{-1}(\underline{\lambda}) = ((I^{ij})), i, j = 0, 1, ..., k+1$ where I^{ij} are the elements of the inverse of $I(\underline{\lambda})$.

4. LARGE SAMPLE TESTS

TEST FOR MVE :

First we develop large sample studentized test for testing multivariate exponentiality (MVE) based on the MLE of c i.e., $\hat{c}$. The hypothesis of the test for MVE is $H_0 : c = 1$. The asymptotic distribution of $\hat{c}$ is $AN(c, I^{k+1,k+1}/n)$. One can obtain studentized test statistic $\sqrt{n}(\hat{c} - 1)/(\hat{I}^{k+1,k+1})^{1/2}$ which is $AN(0,1)$ under H_0 where $\hat{I}^{k+1,k+1}$ is estimated using MLEs of the parameters under (H_0 U H_1). For the alternatives $H_1 : c \neq 1$, we reject H_0 if $n(\hat{c} - 1)^2/(\hat{I}^{k+1,k+1}) > \chi^2_{1,1-\alpha}$ where $\chi^2_{1,1-\alpha}$ is $100(1-\alpha)\%$ point of the chisquare with 1 d.f.

TEST FOR INDEPENDENCE

We next consider the hypothesis of the test for independence of $(X_1, .., X_k)$ i.e., $H_0 : \lambda_0 = 0$. Here the proposed test is based on the MLE $\hat{\lambda}_0$ which is $AN(\lambda_0, I^{00}/n)$. The studentized test statistic is $\sqrt{n}(\hat{\lambda}_0/(\hat{I}^{00})^{1/2}$ which is $AN(0,1)$ under H_0 where $\hat{I}^{00}$ is the estimate of the variance of $\hat{\lambda}_0$. For the alternatives $H_1 : \lambda_0 > 0$, we reject H_0 if $\sqrt{n}\hat{\lambda}_0/(\hat{I}^{00})^{1/2} > \xi_{1-\alpha}$ where $\xi_{1-\alpha}$ is $100(1-\alpha)\%$ point of the standard normal variate.

TEST FOR SYMMETRY :

We next consider the hypothesis of the test for symmetry or identical marginals or exchangebility of $(X_1, ..., X_k)$ i.e., $H_0 : \lambda_1 = \lambda_2 = ... = \lambda_k$ or $\underline{\mu} = \underline{0}$ where $\underline{\mu} = (\lambda_2 - \lambda_1, \lambda_3 - \lambda_2, ..., \lambda_k - \lambda_{k-1})'$. We develop a test based on MLEs i.e. $\underline{\hat{\mu}} = (\hat{\lambda}_2 - \hat{\lambda}_1, ..., \hat{\lambda}_k - \hat{\lambda}_{k-1})'$ and the studentized test statistic is $\underline{\hat{\mu}}'\hat{\Sigma}^{-1}\underline{\hat{\mu}}$ which is χ^2_{k-1} under H_0 where $\hat{\Sigma}^{-1}$ is the estimate of variance - covariance matrix of $\underline{\hat{\mu}}$. For the alternatives $H_1 : \underline{\mu} \neq 0$, we reject H_0 in favour of H_1 if $\underline{\hat{\mu}}\hat{\Sigma}^{-1}\underline{\hat{\mu}} > \chi^2_{k-1,1-\alpha}$. Under the alternatives $H_1 : \underline{\mu} \neq 0$, $\underline{\hat{\mu}}'\hat{\Sigma}^{-1}\underline{\hat{\mu}}$ is non-cental χ^2_{k-1} with non-centrality parameter $\underline{\mu}'\hat{\Sigma}^{-1}\underline{\mu}$.

ACKNOWLEDGEMENTS

I thank the referee for the constructive suggestions and comments.

REFERENCES

Block, H.W. (1975). Continuous multivariate exponential extensions. *Reliability and Fault Tree Analysis. SIAM*, 285-306.

Hanagal, D.D. (1991). Large sample tests for independence and symmetry in the multivariate exponential distribution. *Journal of the Indian Statistical Association*, **29**, 89-93.

Hanagal, D.D.(1993a). Some inference results in an absolutely continuous multivariate exponential model of Block. *Statistics & Probability Letters*, **16(3)**, 177-80.

Hanagal, D.D.(1993b). Some inference results in several symmetric multivariate exponential models. *Communications in Statistics, Theory & Methods*, **22(9)**, 2549-66.

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

Proschan, F. and Sullo, P. (1976). Estimating the parameters of a multivariate exponential distribution. *Journal of the American Statistical Association*, **71**, 465-72.

Weinmann, D.G.(1966). A multivariate extension of the exponential distribution. **Ph.D thesis**. Arizona State University.