Far East Journal of Mathematical Sciences (FJMS)

http://www.pphmj.com
http://dx.doi.org/10.17654/MS101091977
Volume 101, Number 9, 2017, Pages 1977-1992 **ISSN: 0972-0871**

MULTIVARIATE WEIBULL REGRESSION MODEL

Suyitno[1], Purhadi[2], Sutikno[2] and Irhamah[2]

[1]Department of Statistics
Universitas Mulawarman Samarinda
Indonesia

[2]Department of Statistics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia

Abstract

In this study, a multivariate Weibull regression (MWR) model is proposed. The MWR model is a regression model developed from a multivariate Weibull distribution. The proposed MWR model is derived from the joint survival function of the multivariate Weibull distribution developed by Lee and Wen [5], in which the scale parameters are stated in terms of the regression parameters. The aim of this study is to estimate the MWR model parameters using the maximum likelihood estimation (MLE) method, and to test the regression parameters. The results show that the maximum likelihood estimator can be obtained by using the Newton-Raphson iterative method. The regression parameter testing involves simultaneous and partial tests. The test statistic for simultaneous test is Wilk's likelihood ratio statistic and the test statistic for partial test is Wald statistic. Wilk statistic follows chi-square distribution, which can be derived from the likelihood ratio test (LRT) method. The Wald statistic follows standard normal distribution derived from the asymptotic property of the maximum likelihood estimator.

Received: December 9, 2016; Accepted: February 13, 2017
2010 Mathematics Subject Classification: 62H12, 62H15, 62J12.
Keywords and phrases: MWR, MLE, Wilk statistic, Wald statistic.

1. Introduction

The Weibull distribution is a commonly used model in reliability, life time and environmental data analysis. The univariate Weibull distribution originally depends on three parameters, namely, location parameter, scale parameter and shape parameter. A special Weibull distribution is the scale-shape version. It is a univariate Weibull distribution depending on two parameters, namely, scale and shape parameters. As the Weibull distribution development, it can depend on the covariates or explanatory variables [4, 9]. The scale or shape parameter of the Weibull distribution can be expressed in terms of the regression parameters. Furthermore, a probability density function of the Weibull distribution in which the scale parameter is expressed in terms of the regression parameters is called the *Weibull regression model*. The Weibull distribution has interrelation functions, such as the probability density function (PDF), cumulative distribution function (CDF), survival function and the hazard function. If one of them is known, then the other functions can be found.

Many references in literature discuss the Weibull distribution, but still not much about Weibull regression model. O'Quigley and Roberts [7] proposed a regression model for the study of survival time. The study discussed parameter estimation of the univariate Weibull regression model, and the estimated parameter was computed by Fortran program. Hanagal [2, 3] discussed the bivariate Weibull regression models for survival time (censored samples) derived from bivariate exponential distribution of Marshal-Olkin and by extending Freund's bivariate exponential distribution with identical covariates and non-identical regression parameters. The parameter estimation method of bivariate Weibull regression models proposed by Hanagal was MLE.

Study of both the theory and applications of Weibull regression is still limited to the univariate and bivariate cases, meanwhile many problems in various fields involving more than two responses data need solving using MWR model. Moreover, the study of the Weibull regression model is still limited on the parameter estimation, whereas discussion about the regression parameter testing is rarely done. Therefore, as the Weibull regression model development, in this study, the general model of the MWR and the regression parameter testing procedures are proposed. The proposed MWR model is constructed from the joint survival function of the multivariate Weibull distribution developed by Lee and Wen [5], in which the scale parameters are expressed in terms of the regression parameters with identical covariates and non-identical regression parameters. The study is focused on the MWR model construction, parameter estimation and hypothesis testing for regression parameter. The parameter estimation method is MLE and the test statistic for simultaneous test is Wilk's likelihood ratio statistic which is derived from the likelihood ratio test method. The test statistic for partial test is Wald statistic derived from the asymptotic property of the maximum likelihood estimator.

The remaining part of the paper is organized as follows: In Section 2, we introduce the general model of MWR. Parameter estimation of the MWR model is discussed in Section 3, the regression parameter testing is discussed in Section 4, and finally, we conclude the study in Section 5.

2. The Multivariate Weibull Regression Model

To derive MWR model in this section, first we discuss the interrelation functions in the Weibull distribution, namely, the survival function, CDF and PDF. Let $[Y_1\ Y_2\ \cdots\ Y_m]^T$ be a continuous nonnegative random vector associated with a typical unit. Lawless [4] defined the joint survival function denoted by $S(y_1, ..., y_m)$ as

$$S(y_1, ..., y_m) = P\left(\bigcap_{k=1}^{m}(Y_k > y_k)\right) = P(Y_1 > y_1, Y_2 > y_2, ..., Y_m > y_m), \tag{2.1}$$

and the joint CDF denoted by $F(y_1, ..., y_m)$ is defined by

$$F(y_1, ..., y_m) = P\left(\bigcap_{k=1}^{m}(Y_k \le y_k)\right) = P(Y_1 \le y_1, Y_2 \le y_2, ..., Y_m \le y_m). \tag{2.2}$$

The marginal survival function and distribution function will be denoted by $S_k(y_k) = P(Y_k > y_k)$ and $F_k(y_k) = P(Y_k \le y_k)$, respectively, for $k = 1, 2, \ldots, m$.

Based on the probability properties, the relationship between the joint CDF given by equation (2.2) and the joint survival function (2.1) can be written as

$$F(y_1, \ldots, y_m) = 1 - P\left(\bigcap_{k=1}^{m}(Y_k \le y_k)\right)^c = 1 - P\left(\bigcup_{k=1}^{m}(Y_k > y_k)\right),$$

or

$$F(y_1, \ldots, y_m) = 1 - \sum_{k=1}^{m} S(y_k) + \sum_{k_1<k_2}^{m} S(y_{k_1}, y_{k_2})$$

$$- \sum_{k_1<k_2<k_3}^{m} S(y_{k_1}, y_{k_2}, y_{k_3}) + \cdots + (-1)^m S(y_1, \ldots, y_m) \quad (2.3)$$

[1]. For 3-variates case, the relationship (2.3) can be expressed as

$$F(y_1, y_2, y_3) = 1 - S(y_1) - S(y_2) - S(y_3) + S(y_1, y_2)$$

$$+ S(y_1, y_3) + S(y_2, y_3) - S(y_1, y_2, y_3). \quad (2.4)$$

Let the CDF $F(y_1, \ldots, y_m)$ and the joint survival function $S(y_1, \ldots, y_m)$ be absolutely continuous functions, by differentiating both sides of equation (2.3) with respect to $y_1, y_2, \ldots, y_m$, the joint PDF $f(y_1, \ldots, y_m)$ of the multivariate Weibull distribution can be obtained, that is,

$$f(y_1, y_2, \ldots, y_m) = \frac{\partial^m F(y_1, \ldots, y_m)}{\partial y_1 \, \partial y_2 \cdots \partial y_m} = (-1)^m \frac{\partial^m S(y_1, \ldots, y_m)}{\partial y_1 \, \partial y_2 \cdots \partial y_m}. \quad (2.5)$$

Lee and Wen [5] constructed the joint survival function of the multivariate Weibull distribution for continuous nonnegative random vector $\mathbf{y} = [Y_1 \; Y_2 \cdots Y_m]^T$ which is given by equation (2.6) as follows:

$$S(y_1, y_2, \ldots, y_m) = \exp\left(-\left[\sum_{k=1}^{m}\left(\frac{y_k}{\lambda_k}\right)^{\frac{\gamma_k}{a}}\right]^a\right), \tag{2.6}$$

with $0 < a \leq 1$; $0 < y_1, y_2, \ldots, y_m < \infty$; $0 < \gamma_1, \ldots, \gamma_m < \infty$ and $0 < \lambda_1, \ldots, \lambda_m < \infty$. The parameter a represents the degree of dependence in associations of $Y_1, Y_2, \ldots, Y_m$, the parameters γ_k and λ_k for $k = 1, 2, \ldots, m$ are shape and scale parameters, respectively. The cases $a = 0$ and $a = 1$ correspond to maximal positive dependence and independence, respectively. Lu and Bhattacharyya [6] showed that the correlation coefficient of any pair of the random variables on the multivariate Weibull distribution is nonnegative, which can be obtained from the general moment of random variables of the bivariate Weibull distribution. Furthermore, Lee and Wen [5] derived the general moment of random variables of the multivariate Weibull distribution. The marginal survival functions of (2.6) are

$$S_k(y_k) = \exp\left(-\left[\frac{y_k}{\lambda_k}\right]^{\gamma_k}\right) \text{ for } k = 1, 2, \ldots, m. \tag{2.7}$$

The PDF of the scale-shape version of the univariate Weibull distribution can be derived from the survival function (2.7) by using the relationship (2.5), and it has an expression:

$$f(y) = \frac{\gamma}{\lambda}\left[\frac{y}{\lambda}\right]^{\gamma-1} \exp\left(-\left[\frac{y}{\lambda}\right]^{\gamma}\right).$$

It follows directly (2.5) and (2.6), the joint PDF of the multivariate Weibull distribution can be obtained, and it has an expression as follows:

$$f_0(y_1, y_2, \ldots, y_m) = \left(\prod_{k=1}^{m}\left(\frac{1}{a}\right)\left(\frac{\gamma_k}{\lambda_k}\right)\left(\frac{y_k}{\lambda_k}\right)^{\frac{\gamma_k}{a}-1}\right)\exp[-A_m^a]$$

$$\times A_m^{a-m}\left(\sum_{\ell=1}^{m}(-1)^{\ell+m} C(m, \ell, a) A_m^{(\ell-1)a}\right), \tag{2.8}$$

[5], where $A_m = \sum_{k=1}^{m} \left(\frac{y_k}{\lambda_k}\right)^{\frac{\gamma_k}{a}}$ and

$$C(m, \ell, a) = \sum_{\substack{\ell_1+\ell_2+\cdots+\ell_m=\ell \\ \ell_1+2\ell_2+\cdots+m\ell_m=m \\ \ell_k \text{ is a nonnegative integer}}} P^m_{\ell_1, \ell_2, \ldots, \ell_m} \prod_{k=1}^{m} \binom{a}{k}^{\ell_k}.$$

$P^m_{\ell_1, \ell_2, \ldots, \ell_m} = \binom{m}{\ell_1, \ell_2, \ldots, \ell_m}$ is the extended multinomial coefficient defined by $\binom{m}{\ell_1, \ell_2, \ldots, \ell_m} = \frac{m!}{\ell_1! \ell_2! \cdots \ell_m!}$ and $\binom{a}{k} = \frac{a!}{(a-k)!k!}$ [5, 11]. To compute $C(m, \ell, a)$, the summation is extended over all partitions of m into ℓ parts, that is over all nonnegative integer solutions of the system of equations $\sum_{k=1}^{m} \ell_k = \ell$ and $\sum_{k=1}^{m} k\ell_k = m$. From (2.8), for $m = 3$, we get

$$\sum_{\ell=1}^{3} (-1)^{\ell+3} C(3, \ell, a) A_3^{(\ell-1)a} = a(a-1)(a-2) - 3a^2(a-1)A_3^a + a^3 A_3^{2a},$$

where

$$A_3 = \sum_{k=1}^{3} \left(\frac{y_k}{\lambda_k}\right)^{\frac{\gamma_k}{a}} = \left(\frac{y_1}{\lambda_1}\right)^{\frac{\gamma_1}{a}} + \left(\frac{y_2}{\lambda_2}\right)^{\frac{\gamma_2}{a}} + \left(\frac{y_3}{\lambda_3}\right)^{\frac{\gamma_3}{a}},$$

and for $m = 4$, it is obtained that

$$\sum_{\ell=1}^{3} (-1)^{\ell+4} C(4, \ell, a) A_4^{(\ell-1)a}$$

$$= -a(a-1)(a-2)(a-3) + (7a^4 - 18a^3 + 11a^2)A_4^a$$

$$- 6a^3(a-1)A_4^{2a} + a^4 A_4^{3a},$$

and

$$A_4 = \sum_{k=1}^{4} \left(\frac{y_k}{\lambda_k} \right)^{\frac{\gamma_k}{a}} = \left(\frac{y_1}{\lambda_1} \right)^{\frac{\gamma_1}{a}} + \left(\frac{y_2}{\lambda_2} \right)^{\frac{\gamma_2}{a}} + \left(\frac{y_3}{\lambda_3} \right)^{\frac{\gamma_3}{a}} + \left(\frac{y_4}{\lambda_4} \right)^{\frac{\gamma_4}{a}}.$$

Furthermore, parameter estimation of PDF (2.8) can be done by using the MLE method.

In a univariate Weibull regression, the scale parameter can be expressed in terms of the regression parameters [2-4]. In a similar manner, the scale parameters of the multivariate Weibull distribution can depend on the covariates, that is, the scale parameters of the joint survival function stated in equation (2.6) can be expressed in the regression model. Because the scale parameters are positive valued, then λ_k for $k = 1, 2, ..., m$ can be expressed in terms of the regression parameters in the following way:

$$\lambda_k = \exp[\boldsymbol{\beta}_k^T \mathbf{x}] = \exp[\beta_{k0} + \beta_{k1} X_1 + \cdots + \beta_{kp} X_p], \tag{2.9}$$

where $\boldsymbol{\beta}_k^T = [\beta_{k0} \ \beta_{k1} \ \cdots \ \beta_{kp}]$ is a $1 \times (p+1)$ vector of regression parameters with $-\infty < \beta_{kh} < \infty$ for $k = 1, 2, ..., m$; $h = 0, 1, 2, ..., p$ and $\mathbf{x} = [1 \ X_1 \ X_2 \cdots X_p]^T$ is a $(p+1) \times 1$ vector of covariates or independent variables. Using the relationship (2.9), the joint survival function (2.6) can be written in terms of the regression parameters as follows:

$$S(y_1, ..., y_m) = \exp[-\mathcal{A}_m^a], \tag{2.10}$$

where

$$\mathcal{A}_m = \sum_{k=1}^{m} (y_k)^{\gamma_k / a} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x} \right].$$

Based on the joint survival function given by equation (2.10), the joint PDF stated in terms of the regression parameters can be determined by using a relationship (2.5), and it has an expression as follows:

$$f(y_1, y_2, \dots, y_m) = \left(\prod_{k=1}^{m} \frac{\gamma_k}{a} (y_k)^{(\gamma_k/a)-1} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x} \right] \right) \exp[-\mathcal{A}_m^a]$$

$$\times \mathcal{A}_m^{a-m} \left(\sum_{\ell=1}^{m} (-1)^{\ell+m} C(m, \ell, a) \mathcal{A}_m^{(\ell-1)a} \right). \qquad (2.11)$$

The joint PDF given by equation (2.11) is called a *multivariate Weibull regression* (*MWR*) *model*. It has $m(p+2)+1$ parameters consisted of one parameter of degree of dependence, m shape parameters and $m(p+1)$ regression parameters. The parameters of the MWR model (2.11) can be written in the vector form as $\boldsymbol{\theta} = [a \ \boldsymbol{\gamma}^T \ \boldsymbol{\beta}^T]^T$, with $\boldsymbol{\gamma}^T = [\gamma_1 \ \gamma_2 \ \cdots \ \gamma_m]$; $\boldsymbol{\beta}^T = [\boldsymbol{\beta}_1^T \ \boldsymbol{\beta}_2^T \cdots \boldsymbol{\beta}_m^T]$ and $\boldsymbol{\beta}_k^T = [\beta_{k0} \ \beta_{k1} \ \cdots \ \beta_{kp}]$ for $k = 1, 2, \dots, m$.

For $m = 3$, the model (2.11) is known as a trivariate Weibull regression model which has an expression

$$f(y_1, y_2, y_3) = \left(\prod_{k=1}^{3} \frac{\gamma_k}{a} y_k^{(\gamma_k/a)-1} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x} \right] \right) Q_3 \mathcal{A}_3^{a-3} \exp[-\mathcal{A}_3^a], \qquad (2.12)$$

[10], with

$$\mathcal{A}_3 = \sum_{k=1}^{3} (y_k)^{\gamma_k/a} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x} \right]$$

and

$$Q_3 = a(a-1)(a-2) - 3a^2(a-1)\mathcal{A}_3^a + a^3 \mathcal{A}_3^{2a}.$$

3. Parameter Estimation of the MWR Model

The parameter estimation method proposed for the MWR model in this study is MLE. The initial step of the MLE method is defining the likelihood function. Suppose that there are random samples $(Y_{1i}, Y_{2i}, \dots, Y_{mi})$ of responses taken from the population of the multivariate Weibull distribution which has joint PDF (2.8), and $(X_{1i}, X_{2i}, \ldots, X_{pi})$ for $i = 1, 2, \ldots, n$ are the observation values of covariates. The likelihood function of (2.11) is defined by

$$\mathscr{L}(\boldsymbol{\theta} \mid \mathbf{y}) = \prod_{i=1}^{n} f(\boldsymbol{\theta} \mid \mathbf{y}_i)$$

$$= \left(\prod_{i=1}^{n} \left(\prod_{k=1}^{m} \frac{\gamma_k}{a} (y_{ki})^{(\gamma_k/a)-1} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x}_i \right] \right) \right) \left(\prod_{i=1}^{n} \mathcal{A}_{mi}^{a-m} \right)$$

$$\times \left(\prod_{i=1}^{n} \exp[-\mathcal{A}_{mi}^{a}] \right) \left(\prod_{i=1}^{n} \left(\sum_{\ell=1}^{m} (-1)^{\ell+m} C(m, \ell, a) \mathcal{A}_{mi}^{(\ell-1)a} \right) \right). \quad (3.1)$$

The idea of MLE method is to obtain the estimator $(\hat{\boldsymbol{\theta}})$ maximizing the function (3.1). Because $\mathscr{L}(\boldsymbol{\theta} \mid \mathbf{y})$ has partial derivatives with respect to all components of vector $\boldsymbol{\theta}$, then $\hat{\boldsymbol{\theta}}$ is often obtained by solving the equation $\frac{\partial \mathscr{L}(\mathbf{y} \mid \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}$. In most situations, it is more convenient to work with the natural logarithm of $\mathscr{L}(\boldsymbol{\theta} \mid \mathbf{y})$, because its maxima are attained at the same points as those of $\mathscr{L}(\boldsymbol{\theta} \mid \mathbf{y})$. Taking natural logarithm on both sides of (3.1), we get

$$L(\boldsymbol{\theta} \mid \mathbf{y}) = \log \mathscr{L}(\boldsymbol{\theta} \mid \mathbf{y}) = \sum_{q=1}^{4} L_q(\boldsymbol{\theta} \mid \mathbf{y}), \quad (3.2)$$

where

$$L_1(\boldsymbol{\theta} \mid \mathbf{y}) = \sum_{i=1}^{n} \sum_{k=1}^{m} \left(\log \gamma_k - \log a + \left(\frac{\gamma_k}{a} - 1 \right) \log y_{ki} - \frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x}_i \right), \quad (3.3)$$

$$L_2(\boldsymbol{\theta} \mid \mathbf{y}) = \sum_{i=1}^{n} (a - m) \log \mathcal{A}_{mi}, \quad (3.4)$$

$$L_3(\boldsymbol{\theta}|\mathbf{y}) = -\sum_{i=1}^{n} \mathcal{A}_{mi}^{a}, \tag{3.5}$$

$$L_4(\boldsymbol{\theta}|\mathbf{y}) = \sum_{i=1}^{n} \log Q_i, \tag{3.6}$$

where

$$Q_i = \sum_{\ell=1}^{m} (-1)^{\ell+m} C(m, \ell, a) \mathcal{A}_{mi}^{(\ell-1)a}, \quad \mathcal{A}_{mi} = \sum_{k=1}^{m} (y_{ki})^{\gamma_k/a} \exp\left[-\frac{\gamma_k}{a} \boldsymbol{\beta}_k^T \mathbf{x}_i\right]$$

and

$$\mathbf{x}_i = [1 \ X_{1i} \ X_{2i} \ \cdots \ X_{pi}]^T.$$

Based on (3.2), the maximum likelihood (ML) estimator of MWR model $(\hat{\boldsymbol{\theta}})$ is obtained by solving the equation

$$\frac{\partial L(\boldsymbol{\theta}|\mathbf{y})}{\partial \boldsymbol{\theta}} = \mathbf{0}, \tag{3.7}$$

where $\mathbf{0}$ is $(1 + m(p+2)) \times 1$ vector zero, and $\hat{\boldsymbol{\theta}}$ satisfies $\partial L(\hat{\boldsymbol{\theta}}|\mathbf{y})/\partial \boldsymbol{\theta} = \mathbf{0}$. Equation (3.7) is known as the likelihood equation, and the left hand side of equation (3.7) is the gradient vector. Let the gradient vector $\partial L(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ be a scalar vector which has the general form

$$\mathbf{g}(\boldsymbol{\theta}) = \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial L(\boldsymbol{\theta})}{\partial a} \ \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}^T} \ \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}^T}\right]^T, \tag{3.8}$$

where

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma}^T} = \left[\frac{\partial L(\boldsymbol{\theta})}{\partial \gamma_1} \ \frac{\partial L(\boldsymbol{\theta})}{\partial \gamma_2} \cdots \frac{\partial L(\boldsymbol{\theta})}{\partial \gamma_m}\right] \text{ and } \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}^T} = \left[\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_1^T} \ \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_2^T} \cdots \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_m^T}\right]$$

with

$$\frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}_k^T} = \left[\frac{\partial L(\boldsymbol{\theta})}{\partial \beta_{k0}} \ \frac{\partial L(\boldsymbol{\theta})}{\partial \beta_{k1}} \cdots \frac{\partial L(\boldsymbol{\theta})}{\partial \beta_{kp}}\right] \text{ for } k = 1, 2, ..., m.$$

Considering on the expression of the functions (3.3)-(3.6), the likelihood equation (3.7) is a system of interdependent nonlinear equations which cannot be solved analytically for obtaining the explicit expression (closed form) of the ML estimator. However, it can be solved numerically by the Newton-Raphson procedure, and the ML estimator $(\hat{\boldsymbol{\theta}})$ is estimated by the roots of likelihood equation (3.7). To apply the Newton-Raphson algorithm, computation of the gradient vector and the Hessian matrix is required. The Hessian matrix is the symmetric matrix of the second order partial derivatives of $L(\boldsymbol{\theta})$ with respect to all combinations of the components of vector $\boldsymbol{\theta}$, and it has the general form

$$\mathbf{H}(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial^2 L(\boldsymbol{\theta})}{\partial a^2} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial a \partial \boldsymbol{\gamma}^T} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial a \partial \boldsymbol{\beta}^T} \\ \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial a} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\gamma}^T} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\gamma} \partial \boldsymbol{\beta}^T} \\ \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial a} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\gamma}^T} & \frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^T} \end{bmatrix}. \tag{3.9}$$

Based on the expression of function (3.2), to calculate directly the gradient vector (3.8) and the Hessian matrix (3.9) is not simple, hence to make the calculation easier, these can be split into four parts, so that

$$\mathbf{g}(\boldsymbol{\theta}) = \sum_{q=1}^{4} \mathbf{g}_q(\boldsymbol{\theta}) = \sum_{q=1}^{4} \frac{\partial L_q(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \tag{3.10}$$

and the Hessian matrix is formulated as

$$\mathbf{H}(\boldsymbol{\theta}) = \sum_{q=1}^{4} \mathbf{H}_q(\boldsymbol{\theta}) = \sum_{q=1}^{4} \frac{\partial^2 L_q(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}, \tag{3.11}$$

where $L_q(\boldsymbol{\theta})$ for $q = 1, 2, 3, 4$ is given by equations (3.3)-(3.6), respectively, [10].

Based on the expression of equation (3.6), the component of vector

$\mathbf{g}_4(\boldsymbol{\theta})$ can be formulated as follows:

$$\mathbf{g}_4(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{1}{Q_i}\left[\frac{\partial Q_i}{\partial \boldsymbol{\theta}}\right], \tag{3.12}$$

and the elements of the symmetric matrix $\mathbf{H}_4(\boldsymbol{\theta})$ can be obtained by using the form

$$\mathbf{H}_4(\boldsymbol{\theta}) = \sum_{i=1}^{n} \frac{1}{Q_i^2}\left(\left[\frac{\partial^2 Q_i}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}\right] Q_i - \left[\frac{\partial Q_i}{\partial \boldsymbol{\theta}}\right]\left[\frac{\partial Q_i}{\partial \boldsymbol{\theta}^T}\right]\right). \tag{3.13}$$

Because the gradient vector and the Hessian matrix of log-likelihood (3.2) are now known, then the Newton-Raphson algorithm can be applied to obtain $\hat{\boldsymbol{\theta}}$ that maximizes the function (3.2) or (3.1).

4. Hypothesis Testing for Regression Parameter

Hypothesis testing for regression parameters of the MWR model involves simultaneous and partial tests. The goal of the simultaneous test is to confirm the influence of covariates simultaneously. The hypothesis form of the simultaneous test is

$$\mathrm{H}_0 : \boldsymbol{\beta}_1^* = \boldsymbol{\beta}_2^* = \cdots = \boldsymbol{\beta}_m^* = \mathbf{0}$$

$$\mathrm{H}_a : \text{At least one } \boldsymbol{\beta}_k^* \neq \mathbf{0}, \text{ for } k = 1, 2, ..., m, \tag{4.1}$$

where $\boldsymbol{\beta}_k^* = [\beta_{k1}\ \beta_{k2}\ \cdots\ \beta_{kp}]^T$ and $\mathbf{0}$ is $p \times 1$ vector zero. The test statistic for simultaneous test is derived from the likelihood ratio test (LRT) method given by $LR = \dfrac{\mathscr{L}(\hat{\Omega}_0)}{\mathscr{L}(\hat{\Omega}_1)}$. Based on the LRT method, Wilk's likelihood ratio statistic has a form

$$G = -2\log\left[\frac{\mathscr{L}(\hat{\Omega}_0)}{\mathscr{L}(\hat{\Omega})}\right] = 2(L(\hat{\Omega}) - L(\hat{\Omega}_0)) \tag{4.2}$$

[8], with $L(.) = \log \mathscr{L}(.)$, $\hat{\Omega} = \{\hat{a}, \boldsymbol{\gamma}^T, \boldsymbol{\beta}^T\}$ being the set of parameters under

population maximizing $L(\boldsymbol{\theta})$, namely, the estimator of the MWR model parameters, and $\hat{\Omega}_0 = \{\hat{a}, \hat{\gamma}_0^T, \hat{\beta}_{10}, \hat{\beta}_{20}, ..., \hat{\beta}_{m0}\}$ is the set of parameters under H_0 maximizing $L(\boldsymbol{\theta})$.

Based on the asymptotic distribution of ML estimator, the statistic G can be written in the quadratic form as follows:

$$2(L(\hat{\Omega}) - L(\hat{\Omega}_0)) \approx (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T I(\boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \chi^2_{mp}, \tag{4.3}$$

where $I(\boldsymbol{\theta}) = -E[\mathbf{H}(\boldsymbol{\theta})]$ is the Fisher information matrix [8]. Based on (4.1), equation (4.3) can be written as

$$2(L(\hat{\Omega}) - L(\hat{\Omega}_0)) \approx \hat{\boldsymbol{\theta}}^T I(\hat{\boldsymbol{\theta}})\hat{\boldsymbol{\theta}} \xrightarrow{d} \chi^2_{mp},$$

where under H_0, we have $E(\hat{\boldsymbol{\theta}}) = \mathbf{0}$. Hence, the G statistic follows the asymptotic chi-square distribution with mp degree of freedom. For significant level α with $0 < \alpha < 1$, if the value of G exceeds from $(1 - \alpha)$-quantile of the chi-square distribution, then the null hypothesis is rejected.

Partial test for regression parameters aims to determine any significant parameters affecting the MWR model. The hypothesis for partial test is frequently put in the form

$$H_0 : \beta_{kh} = 0$$

$$H_a : \beta_{kh} \neq 0, \text{ for } k = 1, 2, ..., m \text{ and } h = 1, 2, ..., p. \tag{4.4}$$

Based on the asymptotic normality properties of ML estimator, the asymptotic distribution of the ML estimator $(\hat{\boldsymbol{\theta}})$ is given by $(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathbf{N}(\mathbf{0}, [I(\boldsymbol{\theta})]^{-1})$, and it is equivalent to

$$[I(\boldsymbol{\theta})]^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \mathbf{I}_b), \tag{4.5}$$

where $[I(\boldsymbol{\theta})]^{1/2} \times [I(\boldsymbol{\theta})]^{1/2} = I(\boldsymbol{\theta})$ and $\mathbf{I}_b$ is the $(1 + m(p + 2)) \times (1 + m(p + 2))$

identity matrix. Let $(\hat{\beta}_{kh} - \beta_{kh})/SE(\hat{\beta}_{kh} - \beta_{kh})$ be the jth component of vector $[I(\boldsymbol{\theta})]^{1/2}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$. Then any component of vector (4.5) is distributed as $N(0, 1)$, that is,

$$\frac{\hat{\beta}_{kh} - \beta_{kh}}{SE(\hat{\beta}_{kh} - \beta_{kh})} \sim N(0, 1). \tag{4.6}$$

From (4.6), the test statistic for the individual parameter test $\text{H}_0 : \beta_{kh} = 0$ is given by Wald statistic [8]:

$$W = \frac{\hat{\beta}_{kh}}{SE(\hat{\beta}_{kh})}, \tag{4.7}$$

where under H_0, $E(\hat{\beta}_{kh}) = \hat{\beta}_{kh} = 0$. Under H_0, W statistic follows the standard normal distribution. The standard error of $\hat{\beta}_{kh}$ is given by the estimated standard deviation $SE(\hat{\beta}_{kv}) = \sqrt{d^{jj}}$, where d^{jj} is the jth diagonal term of $[I(\hat{\boldsymbol{\theta}})]^{-1}$. The decision will reject H_0 if the value of $|W|$ is greater than $\left(1 - \frac{\alpha}{2}\right)$-quantile of standard normal distribution, or H_0 will be rejected if $p\_value < \alpha$, where $p\_value = 2\int_{|W|}^{\infty} f_Z(z)dz$ with $Z \sim N(0, 1)$.

To evaluate the goodness of proposed MWR model, the additional information like the mean square error (MSE) given by

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - \hat{f}_i)^2, \tag{4.8}$$

where f is the PDF of the population distribution given by (2.8) and $\hat{f}$ is the MWR model, is needed.

5. Conclusion

The multivariate Weibull regression model is a joint probability density function of the multivariate Weibull distribution in which the scale parameters are stated in terms of the regression parameters. The multivariate Weibull regression model discussed in this study is constructed from the joint survival function of the multivariate Weibull distribution developed by Lee and Wen [5]. The parameter estimation method of the MWR model is maximum likelihood estimation. The result shows that the likelihood equation is a system of interdependent nonlinear equations, which cannot be solved analytically to obtain the closed form of the exactness of the maximum likelihood estimator. The maximum estimator of the multivariate Weibull regression parameters is estimated by the roots of likelihood equation solved by using the Newton-Raphson iterative method. Hypothesis testing for regression parameters of the multivariate Weibull regression model involves simultaneous and partial tests. The test statistic for simultaneous test is Wilk's likelihood ratio statistic. Wilk statistic follows chi-square distribution, which can be derived from the *likelihood ratio test* (*LRT*) method. The test statistic for partial test is Wald statistic. Wald statistic follows standard normal distribution derived from the asymptotic property of the maximum likelihood estimator.

References

[1] B. R. Bhat, Modern Probability Theory, An Introductory Text Book, John Willey & Sons, New York, 1981.

[2] D. D. Hanagal, A bivariate Weibull regression model, Econ. Qual. Control 20 (2005), 143-150. http://dx.doi.org/10.1515/eqc.2005.143.

[3] D. D. Hanagal, Parametric bivariate regression analysis based on censored samples: a Weibull model, Econ. Qual. Control 19 (2004), 83-90. http://dx.doi.org/10.1515/eqc.2004.83.

[4] J. F. Lawless, Statistical Model and Methods for Lifetime Data, John Wiley & Sons, Hoboken, New Jersey, 2003. http://dx.doi.org/10.1002/9781118033005.

[5] C. K. Lee and M. J. Wen, A multivariate Weibull distribution, Pak. J. Stat. Operat. Res. 5 (2009), 55-66. http://dx.doi.org/10.18187/pjsor.v5i2.120.

[6] J. C. Lu and G. K. Bhattacharyya, Some new construction of bivariate Weibull models, Ann. Inst. Statist. Math. 42 (1990), 543-559.

[7] J. O'Quigley and A. Roberts, WEIBULL: a regression model for survival time studies, Computer Programs in Biomedicine 12 (1980), 14-18. http://dx.doi.org/10.1016/0010-468x(80)90106-3.

[8] Y. Pawitan, In All Likelihood: Statistical Modelling and Inference Using Likelihood, Oxford University Press Inc., New York, 2001.

[9] H. Rinne, The Weibull Distribution: A Handbook, CRC Press, Taylor & Francis Group, USA, 2009.

[10] Suyitno, Purhadi, Sutikno and Irhamah, Parameter estimation of geographically weighted trivariate Weibull regression model, Appl. Math. Sci. 10 (2016), 861-878. http://dx.doi.org/10.12988/ams.2016.6129.

[11] I. Wahyudi, Purhadi, Irhamah and Sutikno, The development parameter estimation on hazard rate of trivariate Weibull distribution, Amer. J. Biostat. 2 (2011), 26-35. http://www.doi: 10.3844/ajbssp.2011.26.35.