The Development of Parameter Estimation on Hazard Rate of Trivariate Weibull Distribution

Irfan Wahyudi, Purhadi, Irhamah and Sutikno

Department of Statistics,
Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

ABSTRACT

In this study, the interrelation concepts of trivariate distribution function, trivariate survival function, trivariate probability density function and trivariate hazard rate function of trivariate Weibull distribution are presented. The goal of this contribution is to estimate the trivariate Weibull hazard rate parameters. To reach this goal, we will use an analitical approach in estimating called the Maximum Likelihood Estimation (MLE) method. Using numerical iterative procedure the scale parameters, the shape parameters and the power parameter estimators on trivariate hazard rate of trivariate Weibull distribution must be obtained. The MLE technique estimates accurately the trivariate Weibull hazard rate parameters.

Keywords: Trivariate Weibull Distribution, Parameter Estimation, Trivariate Weibull Hazard Rate, Maximum Likelihood Estimation (MLE)

1. INTRODUCTION

The Weibull distribution is characterized in the class of absolutely continuous distributions by two parameters, one is the shape parameter, say γ and the other is the scale parameter, say η. The Weibull distribution plays a central role in the analysis of survival or failure time data. For survival data, the proportional hazard model is the most popular model. Moreover, the Weibull distribution is good to be used in parameterized the proportional hazard model because of its flexibility, allowing increasing, decreasing or constant hazard rate function.

The literature on parameter estimation for the Weibull distribution is vast. So, we will select only some key papers. The first study, Tuerlinckx (2004), he derived the Laplace tranform of the positive stable density. His proposition had been applied as a useful vehicle to derive a multivariate survival function of Weibull distribution.

The other papers based on analytical methods of parameter estimation are Hanagal (2005), he propose a new bivariate Weibull regression model on censored samples with common covariates and obtaining the maximum likelihood estimators for parameters, Hanagal (2006) developed a maximum likelihood estimation procedure for a bivariate frailty regression model. Lee and Wen (2009) propose a multivariate Weibull model and derived the explicit form of PDF, CDF and general moment.

In this study, we propose a new parameter estimation on trivariate Weibull distribution related to the consideration of hazard rates. The plan of this study is organized as follows. We will introduce the Weibull distribution and its related functions, next we derive estimators for the parameters of the proposed model and the lastly, we present the results, discussion and deduce a few conclusion. Unfortunately, at this stage of the investigations there are no real data for iterating our proposed parameter estimation.

2. MATERIALS AND METHODS

Let T be a continuous nonnegative random variable representing the failure time of an individual from a homogeneous population with probability density function $f_T(t)$ and cummulative distribution functon $F_T(t) = P(T \leq t)$. The complementary cumulative distribution

Corresponding Author: Irfan Wahyudi, Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia

function $S_T(t) = 1 - F_T(t) = P(T > t)$ is called survival function. $S_T(t)$ is the probability of surviving an age of t. The probability density fuction $f_T(t)$ of random variable T is then equal to $dF_T(t)/dt$ or equivalenly $-dS_T(t)/dt$.

The univariate hazard rate $h_T(t)$ of failure time T is defined as the probability of failure during a very small time interval, assuming that the individual has survived to the beginning of the interval:

$$h_T(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t \mid T \ge t)}{\Delta t} = \frac{dF_T(t)/dt}{S_T(t)}$$

$$= \frac{d(1 - S_T(t))/dt}{S_T(t)} = \frac{-dS_T(t)/dt}{S_T(t)} = -\frac{d}{dt}\ln S_T(t)$$

for all t such that $S_T(t) > 0$. Another expression of $h_T(t)$ as defined above is $h_T(t) = f_T(t)/S_T(t)$ (Navarro, 2007). This hazard rate plays an important role in survival data analysis.

Correspondingly, the cumulative hazard rate function of failure time T is defined as:

$$\mathcal{H}_T(t) = \int_0^t h_T(s)\,ds = -\ln S_T(t)$$

hence, the link between the cumulative hazard rate $\mathcal{H}_T(t)$ and the survival function $S_T(t)$ is done by the following relationship:

$$S_T(t) = \exp\left(-\mathcal{H}_T(t)\right)$$

this equation has been noted by (Singpurwalla, 2006).

Unlike the univariate setup, there are various definitions of multivariate hazard (failure) rate functions. One can define the multivariate hazard rate in more than one way. In the case of bivariate hazard rate function (for overviews, see Kundu and Gupta, 2010a).

Let $(T_1, T_2)^T$ be a continuous nonnegative random vector with the joint survival function $S_{T1,T2}(t_1, t_2) = P(T_1 > t_1, T_2 > t_2)$, the joint distribution has the form $F_{T1,T2}(t_1, t_2) = P(T_1 \le t_1, T_2 \le t_2)$. This distribution can be defined in terms of survival functions as:

$$F_{T_1,T_2}(t_1, t_2) = P\left(\bigcap_{i=1}^{2}\{T_i \le t_i\}\right) = 1 - P\left(\bigcap_{i=1}^{2}\{T_i \le t_i\}\right)^c$$

$$= 1 - P\left(\bigcup_{i=1}^{2}\{T_i \le t_i\}^c\right) = 1 - P\left(\bigcup_{i=1}^{2}\{T_i > t_i\}\right)$$

$$= 1 - S_{T_1}(t_1) - S_{T_2}(t_2) + S_{T_1,T_2}(t_1, t_2)$$

where $S_{T1}(t_1) = P(T_1 > t_1)$ and $S_{T2}(t_2) = P(T_2 > t_2)$ are marginal survival functions of T_1 and T_2, respectively.

Similarly, the joint survival function $S_{T1,T2}(t_1, t_2)$ can also be defined in terms of distribution functions as $S_{T1,T2}(t_1, t_2) = 1 - F_{T1}(t_1) - F_{T2}(t_2) + F_{T1,T2}(t_1, t_2)$, where $F_{T1}(t_1) = P(T_1 \le t_1)$ and $F_{T2}(t_2) = P(T_2 \le t_2)$ are marginal distribution functions of T_1 and T_2, respectively. Clearly that the joint survival function $S_{T1,T2}(t_1, t_2)$ and the joint distribution function $F_{T1,T2}(t_1, t_2)$ is not direct as for the univariate case, that is $S_{T1,T2}(t_1, t_2) \neq 1 - F_{T1,T2}(t_1, t_2)$.

Hence, by differentiating $S_{T1,T2}(t_1, t_2)$ with respect to t_1 and t_2, we have:

$$\frac{\partial^2 F_{T_1,T_2}(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{\partial^2 S_{T_1,T_2}(t_1, t_2)}{\partial t_1 \partial t_2}$$

Note that, for the bivariate case, the second order partial derivatives of the joint distribution $F_{T1,T2}(t_1, t_2)$ and the joint survival function $S_{T1,T2}(t_1, t_2)$ are equal.

The joint probability density function $f_{T1,T2}(t_1, t_2)$ of random variable T_1, T_2 is defined as:

$$f_{T_1,T_2}(t_1, t_2) = \lim_{\substack{\Delta t_1 \to 0 \\ \Delta t_2 \to 0}} \frac{P\left(\bigcap_{i=1}^{2}\{t_i \le T_i < t_i + \Delta t_i\}\right)}{\Delta t_1 \Delta t_2}$$

$$= \lim_{\substack{\Delta t_1 \to 0 \\ \Delta t_2 \to 0}} \frac{P\left(\bigcap_{i=1}^{2}\{T_i < t_i + \Delta t_i\}\right) - P\left(\bigcap_{i=1}^{2}\{T_i < t_i\}\right)}{\Delta t_1 \Delta t_2}$$

$$= \frac{\partial^2 F_{T_1,T_2}(t_1, t_2)}{\partial t_1 \partial t_2} = \frac{\partial^2 S_{T_1,T_2}(t_1, t_2)}{\partial t_1 \partial t_2}$$

If we view the density as the unconditional failure rate, we can define a conditional failure rate as being the same quantity after having accounted for the fact that the individual has already survived until the time point t_1 and t_2. Then, the bivariate hazard rate function of the random variable T_1, T_2 can be written as a scalar quantity, given by:

$$h_{T_1,T_2}(t_1, t_2) = \lim_{\substack{\Delta t_1 \to 0 \\ \Delta t_2 \to 0}} \frac{P\left(\bigcap_{i=1}^{2}\{t_i \le T_i < t_i + \Delta t_i\} \mid \bigcap_{i=1}^{2}\{T_i \ge t_i\}\right)}{\Delta t_1 \Delta t_2}$$

$$= \frac{\lim\limits_{\substack{\Delta t_1 \to 0 \\ \Delta t_2 \to 0}} P\left(\bigcap_{i=1}^{2}\{t_i \le T_i < t_i + \Delta t_i\}\right)\frac{1}{\Delta t_1 \Delta t_2}}{P(T_1 \ge t_1, T_2 \ge t_2)}$$

$$= \frac{\partial^2 S_{T_1,T_2}(t_1, t_2)/\partial t_1 \partial t_2}{S_{T_1,T_2}(t_1, t_2)} = \frac{f_{T_1,T_2}(t_1, t_2)}{S_{T_1,T_2}(t_1, t_2)}$$

The other approach in defining bivariate hazard rate is based on a concept of vector-valued bivariate hazard rate as was stated in Kundu and Gupta (2010b.), that is:

$$\mathbf{h}_{T_1,T_2}(t_1,t_2) = \left(h_{T_1,T_2}(t_1,t_2)_1, h_{T_1,T_2}(t_1,t_2)_2\right)^T$$

where, $h_{T1,T2}(t_1,t_2)_i = -\partial \ln S_{T1,T2}(t_1,t_2)/\partial t_i$, i=1,2.

Similarly, in the case of trivariate, the cumulative distribution function of a continuous nonnegative random vector $(T_1,T_2,T_3)^T$ is defined by Eq. 1:

$$F_{T_1,T_2,T_3}(t_1,t_2,t_3) = P(T_1 \le t_1, T_2 \le t_2, T_3 \le t_3)$$

$$= 1 - \sum_{i=1}^{3} S_{T_i}(t_i) + \sum_{1 \le i < j \le 3} S_{T_i,T_j}(t_i,t_j) - S_{T_1,T_2,T_3}(t_1,t_2,t_3) \quad (1)$$

Note from (1) that $S_{Ti,Tj}(t_i,t_j) = P(T_i > T_i, T_j > t_j)$ is the joint survival function of the two failure time random variables T_i, T_j with $1 \le i < j \le 3$ and $S_{Ti}(t_i) = P(T_i > t_i)$ is the marginal survival function of T_i.

The relationship between the joint survival function $S_{T1,T2,T3}(t_1,t_2,t_3)$ and the joint cumulative distribution function $F_{T1,T2,T3}(t_1,t_2,t_3)$ can also be written as:

$$S_{T_1,T_2,T_3}(t_1,t_2,t_3) = P(T_1 > t_1, T_2 > t_2, T_3 > t_3)$$

$$= 1 - \sum_{i=1}^{3} F_{T_i}(t_i) + \sum_{1 \le i < j \le 3} F_{T_i,T_j}(t_i,t_j) - F_{T_1,T_2,T_3}(t_1,t_2,t_3)$$

Which implies that if the cumulative distribution function (1) is absolutely continuous, then the joint probability density function $f_{T1,T2,T3}(t_1,t_2,t_3)$ has the following expression Eq. 2:

$$f_{T_1,T_2,T_3}(t_1,t_2,t_3) = \frac{\partial^3 F_{T_1,T_2,T_3}(t_1,t_2,t_3)}{\partial t_1 \partial t_2 \partial t_3}$$

$$= -\frac{\partial^3 S_{T_1,T_2,T_3}(t_1,t_2,t_3)}{\partial t_1 \partial t_2 \partial t_3} \quad (2)$$

Unlike in the case of bivariate, the joint probability density function in (2) is equal to (-1) times the third partial derivative of the survival function with respect to t_1, t_2 and t_3.

For notational convenience we will drop the triple subscript T_1,T_2,T_3 denoting the failure time random variables. Henceforth, we will denote $S_{T1,T2,T3}(t_1, t_2, t_3)$ by $S_T(t)$, $F_{T1,T2,T3}(t_1,t_2,t_3)$ by $F_T(t)$, $f_{T1,T2,T3}(t_1,t_2,t_3)$ by $f_T(t)$ and $h_{T1,T2,T3}(t_1,t_2,t_3)$ by $h_T(t)$ for short.

The joint survival function for the trivariate Weibull distribution of random variable T_1, T_2 and T_3 proposed by Lee and Wen (2009) for $0 < \alpha \le 1$; $0 \le t_1, t_2, t_3 < \infty$ is given by Eq. 3:

$$S_T(t) = \exp(-\mathcal{K}^\alpha), \text{ where } \mathcal{K} = \sum_{\ell=1}^{3} \left(\frac{t_\ell}{\eta_\ell}\right)^{\frac{\gamma_\ell}{\alpha}} \quad (3)$$

where T_1, T_2, T_3 are three failur times with univariate survival functions $S_{Ti}(t_i) = P(T_i > t_i)$. Each univariate survival function has a shape parameter γ_i and a scale parameter η_i. The parameter α represents the degree of dependence in the association of T_1, T_2 and T_3. The case $\alpha = 0$ and $\alpha = 1$ correspond to maximal positive dependence and independence, respectively.

It follows directly from (2) and (3) we have the joint density function of trivariate Weibull distribution and it takes the form Eq. 4:

$$f_T(t) = \left[\prod_{\ell=1}^{3} \left(\frac{1}{\alpha}\right)\left(\frac{\gamma_\ell}{\eta_\ell}\right)\left(\frac{t_\ell}{\eta_\ell}\right)^{\frac{\gamma_\ell}{\alpha}-1} \right]$$

$$\times \left[\sum_{k=1}^{3} (-1)^{k+3} \mathfrak{C}(3,k,\alpha) \mathcal{K}^{k\alpha-3} \right] \exp(-\mathcal{K}^\alpha) \quad (4)$$

where, $\mathfrak{C}(m,k,\alpha)$ is the generalized factorial coefficient, which for all positif integer m, k with $k \le m$ and for real number α is defined by:

$$\mathfrak{C}(m,k,\alpha) \underset{=}{\text{def}} \sum_{\substack{k_1+k_2+\cdots+k_m=k \\ k_1+2k_2+\cdots+mk_m=m \\ k_i \text{ is nonnegative integer}}} \binom{m}{k_1,k_2,\ldots,k_m} \prod_{i=1}^{m} \binom{\alpha}{i}^{k_i}$$

And $\binom{m}{k_1,k_2,\ldots,k_m}$ is the extended multinomial coefficient defined by:

$$\binom{m}{k_1,k_2,\ldots,k_m} = \frac{m!}{k_1! k_2! \ldots k_m!}$$

Furthermore, the summation is extended over all partitions of m into k parts, that is over all nonnegative integer solutions of the equations $k_1+2k_2+\ldots+mk_m = m$, $k_1 + k_2 +\ldots+ k_m = k$. This implies that expanding $\mathfrak{C}(3,k,\alpha)$ causes the expression (4) may be written, equivalently, as Eq. 5:

$$f_T(t) = \left[\prod_{\ell=1}^{3} \left(\frac{\partial \mathcal{K}}{\partial t_\ell}\right) \right] \begin{bmatrix} \alpha(\alpha-1)(\alpha-2)\mathcal{K}^{\alpha-3} \\ -3\alpha^2(\alpha-1)\mathcal{K}^{2\alpha-3} + \alpha^3\mathcal{K}^{3\alpha-3} \end{bmatrix} \exp(-\mathcal{K}^\alpha) \quad (5)$$

By introducing the relationship of the trivariate Weibull survival function and the trivariate Weibull density function above we have the trivariate expression of the Weibull hazard rate Eq. 6:

$$h_{\mathbf{T}}(\mathbf{t}) = \frac{f_{T}(\mathbf{t})}{S_{T}(\mathbf{t})} = \left[\prod_{\ell=1}^{3}\left(\frac{\partial\mathcal{K}}{\partial t_{\ell}}\right)\right]\left[\sum_{k=1}^{3}(-1)^{k+3}\mathfrak{C}(3,k,\alpha)\mathcal{K}^{k\alpha-3}\right] \quad (6)$$

for all (t_1,t_2,t_3) such that $S_T(t)>0$, or in the concept of vector-valued trivariate hazard rate, it is defined as Eq. 7:

$$\mathbf{h}_{\mathbf{T}}(\mathbf{t}) = \left(h_{\mathbf{T}}(\mathbf{t})_{1}, h_{\mathbf{T}}(\mathbf{t})_{2}, h_{\mathbf{T}}(\mathbf{t})_{3}\right)^{\mathrm{T}} \quad (7)$$

where for $\ell = 1,2,3$

$$h_{\mathbf{T}}(\mathbf{t})_{\ell} = -\partial \ln S_{\mathbf{T}}(\mathbf{t})/\partial t_{\ell} = \frac{\gamma_{\ell}}{\eta_{\ell}}\left(\frac{t_{\ell}}{\eta_{\ell}}\right)^{(\gamma_{\ell}/\alpha)-1}\mathcal{K}^{\alpha-1}$$

Suppose that there are n individuals under investigation and for each individual there are L failure types. Let $T_{j\ell}$ denote the exact failure times for the ℓ-th failure type of the j-th individual (j = 1 ,…, n; ℓ = 1 ,…, L). If their failure times follow the trivariate Weibull distribution and we assume that the observations are n independent and identically distributed (i.i.d), the likelihood function is Eq. 8:

$$\mathscr{L}(\theta) = \prod_{j=1}^{n}\left[\prod_{\ell=1}^{3}\left(\frac{\partial\mathcal{K}_{j}}{\partial t_{\ell}}\right)\right]\left((\alpha)_{3} - 3\alpha(\alpha)_{2}\mathcal{K}_{j}^{\alpha} + \alpha^{3}\mathcal{K}_{j}^{2\alpha}\right) \quad (8)$$
$$\times \mathcal{K}_{j}^{\alpha-3}\exp(-\mathcal{K}_{j}^{\alpha})$$

where, $\theta = (\gamma_1, \gamma_2, \gamma_3, \eta_1, \eta_2, \eta_3, \alpha)^{\mathrm{T}}$ denotes a vector of unknown parameters of trivariate Weibull distribution, $\mathcal{K}_{j} = \sum_{\ell=1}^{3}(t_{j\ell}/\eta_{\ell})^{\gamma_{\ell}/\alpha}$ and $(\alpha)_m$ is descending factorial of α of order m, defined for α a real number and m an integer by $(\alpha)_{m} \underset{=}{\mathrm{def}} \alpha(\alpha-1)(\alpha-2)...(\alpha-m+1)$; m = 1, 2,…,; $(\alpha)_{0} = 1$. The idea of Maximum Likelihood Estimation (MLE) is to find estimator $\hat{\theta}$ that maximizes (8). Taking natural logarithm on both sides of (8) and denote it by $\mathcal{L}(\theta)$, we have Eq. 9:

$$\mathcal{L}(\theta) = \sum_{j=1}^{n}\sum_{\ell=1}^{3}\ln\left(\frac{\partial\mathcal{K}_{j}}{\partial t_{j\ell}}\right) \quad (9)$$
$$+ \sum_{j=1}^{n}\ln\left[(\alpha)_{3} - 3\alpha(\alpha)_{2}\mathcal{K}_{j}^{\alpha} + \alpha^{3}\mathcal{K}_{j}^{2\alpha}\right] + \sum_{j=1}^{3}\ln\mathcal{K}_{j}^{\alpha-3} - \sum_{j=1}^{n}\mathcal{K}_{j}^{\alpha}$$

Note that the equation (9) is known as the log likelihood function of θ.

Maximization of (8) is rarely done by some procedure of direct optimization, but usually by some methods based on derivatives of (8). The process of forming these derivatives is made easier by departing from the log likelihood function (9) which is a sum instead of a product, as the logarithmic transformation is isotonic, the extremal points $\hat{\theta}_i$ of $\mathscr{L}(\theta)$ and $\mathcal{L}(\theta)$ will be the same. The MLE $\hat{\theta}$ of θ is the set of $\{\hat{\gamma}, \hat{\eta}, \hat{\alpha}\}$ that maximizes $\mathscr{L}(\theta)$ or, equivalently, maximizes $\mathcal{L}(\theta)$, i.e., $\mathcal{L}(\hat{\theta}) = \max_{\text{all }\theta}(\mathcal{L}(\theta))$, where $\gamma = (\gamma_1,\gamma_2,\gamma_3)^{\mathrm{T}}$ and $\eta=(\eta_1,\eta_2,\eta_3)^{\mathrm{T}}$.

It is clear that if vector $\hat{\theta}$ is a solution of the following simultaneous equations, which are obtained by taking the partial derivative of (9) with respect to each elements of vector θ and setting the result to zero Eq. 10:

$$\frac{\partial\mathcal{L}(\theta)}{\partial\theta} = \mathbf{0} \quad (10)$$

Then $\hat{\theta}$ is a maximizer of (9). Note that the equation (10) above is known as the log likelihood equation and the 7×1 vector $\partial\mathcal{L}(\theta)/\partial\theta$ in the left hand side of (10) is called the score vector and denoted by $\mathbf{g}(\theta) = \partial\mathcal{L}(\theta)/\partial\theta$, with $\partial\mathcal{L}(\theta)/\partial\theta = (\partial\mathcal{L}(\theta)/\partial\gamma^{\mathrm{T}}, \partial\mathcal{L}(\theta)/\partial\eta^{\mathrm{T}}, \partial\mathcal{L}(\theta)/\partial\alpha)^{\mathrm{T}}$. Where $\partial\mathcal{L}(\theta)/\partial\gamma^{\mathrm{T}} = (\partial\mathcal{L}(\theta)/\partial\gamma_{1}, \partial\mathcal{L}(\theta)/\partial\gamma_{2}, \partial\mathcal{L}(\theta)/\partial\gamma_{3})$ and $\partial\mathcal{L}(\theta)/\partial\eta^{\mathrm{T}} = (\partial\mathcal{L}(\theta)/\partial\eta_{1}, \partial\mathcal{L}(\theta)/\partial\eta_{2}, \partial\mathcal{L}(\theta)/\partial\eta_{3})$.

Let $\mathcal{L}(\theta)$ be a scalar-valued function of a vector θ as in Eq. 9. The second derivative of $\mathcal{L}(\theta)$ with respect to vector θ is a matrix of the partial derivatives of $\mathcal{L}(\theta)$ with respect to the elements of the vector θ. This matrix is called the Hessian and is denoted by $\mathbf{H}(\theta)$:

$$\mathbf{H}(\theta) = \left(\frac{\partial^{2}\mathcal{L}(\theta)}{\partial\theta\partial\theta^{\mathrm{T}}}\right) \quad (11)$$

where the elements of $\mathbf{H}(\theta)$ in (11) are $\partial^{2}\mathcal{L}(\theta)/\partial\theta_{u}\partial\theta_{v}$, $u,v = 1,...,7$.

The Hessian matrix (11) plays a crucial role in solving estimating Eq. 10 with a Newton Raphson algorithm. The MLEs of the trivariate Weibull parameters cannot be given in closed form since Eq. 10 is a system of interdependent non linear equation. Hence, we have to apply iterative methods, carried through to convergence or terminated after reaching some given stopping criterion, to calculate or approximate the MLE resulting in a so called iterated MLE. The MLE $\hat{\theta}$ is the root to the estimating (10), which can be solved by newton raphson algoritm.

Maximization of (9) directly is consuming time. For simplicity in computation equation (10), without loss of generality, we have to split (9) into four parts:

$$\triangleright \mathcal{L}_1(\theta) = \sum_{j=1}^{n} \sum_{\ell=1}^{3} \ln\left(\frac{\partial \mathcal{K}_j}{\partial t_{j\ell}}\right) \qquad (12a)$$

$$\triangleright \mathcal{L}_2(\theta) = \sum_{j=1}^{n} \ln Q_j, \text{ where } Q_j = (\alpha)_3 - 3\alpha(\alpha)_2 \mathcal{K}_j^\alpha + \alpha^3 \mathcal{K}_j^{2\alpha} \quad (12b)$$

$$\triangleright \mathcal{L}_3(\theta) = \sum_{j=1}^{n} \ln\left(\mathcal{K}_j^{\alpha-3}\right) \qquad (12c)$$

$$\triangleright \mathcal{L}_4(\theta) = -\sum_{j=1}^{n} \mathcal{K}_j^\alpha \qquad (12d)$$

The results of analytical calculation for the four functions $\mathcal{L}_i(\theta)$ (i = 1,…, 4) above can be described as follows.

2.1. For the First Part

First of all, by recalling that θ is a seven-dimensional vector of unknown parameters consisting of three shape parameters, three scale parameters and one power parameter, we see that each of Eq.12a-12d is a function of the seven unknown parameters. Moreover, notice that throughout this study, the subscript k or ℓ indicates the k or ℓ -th failure type and therefore the range of k and ℓ are always {1, 2, 3}. Now, let $\mathbf{g}_1(\theta)$ be defined as in the left hand side of (10). The first partial derivative of log likelihood $\mathcal{L}_1(\theta)$ in (12a) with respect to θ, termed score vector $\mathbf{g}_1(\theta)$, is:

$$\mathbf{g}_1(\theta) = \frac{\partial \mathcal{L}_1(\theta)}{\partial \theta} \qquad (13)$$

where, $\mathbf{g}_1(\theta)$ in (13) is 7×1 score vector for θ, which elements of the score vector may be obtained as $\partial \mathcal{L}_1(\theta)/\partial\theta = (\partial\mathcal{L}_1(\theta)/\partial\gamma^T, \partial\mathcal{L}_1(\theta)/\partial\eta^T, \partial\mathcal{L}_1(\theta)/\partial\alpha)^T$. By applying the chain rule in deriving (12a) with respect to θ, we have the following explicit expression:

$$\frac{\partial \mathcal{L}_1(\theta)}{\partial \gamma_\ell} = \frac{n}{\gamma_\ell} + \frac{1}{\alpha} \sum_{j=1}^{n} \ln\left(\frac{t_{j\ell}}{\eta_\ell}\right), \text{ for } \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_1(\theta)}{\partial \eta_\ell} = -\frac{n\gamma_\ell}{\alpha\eta_\ell}, \text{ for } \ell = 1,2,3$$

$$\frac{\partial \mathcal{L}_1(\theta)}{\partial \alpha} = -\frac{3n}{\alpha} - \frac{1}{\alpha} \sum_{j=1}^{n} \sum_{\ell=1}^{3} \ln\left(\frac{t_{j\ell}}{\eta_\ell}\right)^{\frac{\gamma_\ell}{\alpha}}$$

Let $\mathbf{H}_1(\theta)$ be the 7×7 Hessian matrix as given in (11), the second partial derivative of the log likelihood function $\mathcal{L}_1(\theta)$ in (12a) with respect to θ is a matrix of the form:

$$\mathbf{H}_1(\theta) = \left(\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\theta\partial\theta^T}\right) \qquad (14)$$

where the elements of $\mathbf{H}_1(\theta)$ in (14) are $\partial^2\mathcal{L}_1(\theta)/\partial\theta_u\partial\theta_v$, u,v = 1,...,7.

The k ℓ th entries of Hessian matrix (14) , for k, ℓ = 1, 2, 3 with k ≠ ℓ , can be derived easily as:

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_k\partial\gamma_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_\ell\partial\gamma_k} = 0$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_k\partial\eta_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_\ell\partial\eta_k} = 0$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_k\partial\gamma_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_\ell\partial\eta_k} = 0$$

Furthermore, for k $=\ell$, the diagonal entries of Hessian matrix (14) can be displayed as:

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_\ell^2} = -\frac{n}{\gamma_\ell^2}, \text{ for } \ell = 1,2,3$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_\ell^2} = \frac{n\gamma_\ell}{\alpha\eta_\ell^2}, \text{ for } \ell = 1,2,3$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\alpha^2} = \frac{3n}{\alpha^2} + \frac{2}{\alpha^2} \sum_{j=1}^{n} \sum_{\ell=1}^{3} \ln\left(\frac{t_{j\ell}}{\eta_\ell}\right)^{\frac{\gamma_\ell}{\alpha}}$$

In addition, some other off-diagonal entries of Hessian matrix (14) can also be obtained, i.e., for k, ℓ = 1, 2 and 3:

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_\ell\partial\gamma_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_\ell\partial\eta_\ell} = -\frac{n}{\alpha\eta_\ell}$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\alpha\partial\gamma_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\gamma_\ell\partial\alpha} = -\frac{1}{\alpha^2} \sum_{j=1}^{n} \ln\left(\frac{t_{j\ell}}{\eta_\ell}\right)$$

$$\frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\alpha\partial\eta_\ell} = \frac{\partial^2 \mathcal{L}_1(\theta)}{\partial\eta_\ell\partial\alpha} = \frac{n\gamma_\ell}{\alpha^2\eta_\ell}$$

2.2. For the Second Part

In the same way, a vector $\mathbf{g}_2(\theta)$ of the first order partial derivative and a Hessian matrix $\mathbf{H}_2(\theta)$ of mixed second order partial derivatives of (12b), respectively, defined as follows:

$$\mathbf{g}_2(\theta) = \frac{\partial \mathcal{L}_2(\theta)}{\partial \theta} \tag{15}$$

where, $\mathbf{g}_2(\theta)$ in (15) is 7×1 score vector for θ, which elements of the score vector may be obtained as $\partial \mathcal{L}_2(\theta)/\partial\theta = (\partial \mathcal{L}_2(\theta)/\gamma^T, \partial \mathcal{L}_2(\theta)/\partial\eta^T, \partial \mathcal{L}_2(\theta)/\partial\alpha)^T$. By applying the chain rule in deriving (12b) with respect to θ, we have:

$$\frac{\partial \mathcal{L}_2(\theta)}{\partial \theta} = \sum_{j=1}^{n} \frac{\mathbf{Q}_j}{Q_j}, \text{ where } \mathbf{Q}_j = \left(\frac{\partial Q_j}{\partial \theta} \right) \tag{16}$$

Notice here that the vector $\mathbf{Q}_j$ in (16) is the derivative of the scalar-valued function Q_j as in (12b) with respect to vector θ. Hence $\mathbf{Q}_j = (\partial Q_j/\partial\gamma^T, \partial Q_j/\partial\eta^T, \partial Q_j/\partial\alpha)^T$, where $\partial Q_j/\partial\gamma^T = (\partial Q_j/\partial\gamma_1, \partial Q_j/\partial\gamma_2, \partial Q_j/\partial\gamma_3)$ and $\partial Q_j/\partial\eta^T = (\partial Q_j/\partial\eta_1, \partial Q_j/\partial\eta_2, \partial Q_j/\partial\eta_3)$.
And:

$$\mathbf{H}_2(\theta) = \left(\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial\theta\partial\theta^T} \right) \tag{17}$$

Which elements of symmetric Hessian matrix (17) may be formulated as $\partial^2 \mathcal{L}_2(\theta)/\partial\theta_u\partial\theta_v$, u,v = 1,...,7. Then by applying the derivative quotient rule on Eq. 16 with respect to θ, we have:

$$\left(\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial\theta\partial\theta^T} \right) = \sum_{j=1}^{n} \frac{1}{Q_j^2} \left[\left(\frac{\partial \mathbf{Q}_j}{\partial\theta^T} \right) Q_j - \mathbf{Q}_j^{\otimes 2} \right] \tag{18}$$

Note that for r = 0, 1, 2, where, for a column vector $\mathbf{a}$, $\mathbf{a}^{\otimes 2}$ denotes the outer product $\mathbf{a}\mathbf{a}^T$, $\mathbf{a}^{\otimes 1}$ denotes the Vector $\mathbf{a}$ and $\mathbf{a}^{\otimes 0}$ denotes the scalar 1 (Cai et $al.$, 2007).

Hence, by exploiting (16) the elements of vector $\mathbf{Q}_j$ could be found, i.e., for $\ell = 1, 2, 3$

$$\frac{\partial Q_j}{\partial\gamma_\ell} = -\alpha A_{j\ell}\psi_{1j}, \quad \frac{\partial Q_j}{\partial\eta_\ell} = \alpha B_{j\ell}\psi_{1j}, \quad \frac{\partial Q_j}{\partial\alpha} = \vartheta_j \tag{19}$$

Where:

$$A_{j\ell} = \left(t_{j\ell} / \eta_\ell \right)^{\gamma_\ell/\alpha} \ln\left(t_{j\ell} / \eta_\ell \right)$$

$$B_{j\ell} = \left(\gamma_\ell / \eta_\ell \right)\left(t_{j\ell} / \eta_\ell \right)^{\gamma_\ell/\alpha}$$

$$\vartheta_j = 3\left[(\alpha-1)^2 - \alpha(3\alpha-2)\mathcal{K}_j^\alpha + \alpha^2\mathcal{K}_j^{2\alpha} \right] - 1 + \mathcal{W}_j\psi_{1j}$$

$$\psi_{1j} = \left(3(\alpha)_2 - 2\alpha^2\mathcal{K}_j^\alpha \right)\mathcal{K}_j^{\alpha-1}, \text{ and}$$

$$\mathcal{W}_j = \sum_{k=1}^{3}\left[\gamma_k A_{jk} - \left(t_{jk}/\eta_k \right)^{\gamma_k/\alpha} \ln\mathcal{K}_j^\alpha \right]$$

Obviously, $\mathbf{Q}_j^{\otimes 2}$ in (18) is a 7×7 symmetric matrix outer product with $\mathbf{Q}_j^{\otimes 2} = (\partial Q_j/\partial\theta)(\partial Q_j/\partial\theta)^T$, where $(\partial Q_j/\partial\theta)$ is a vector with elements of the vector are as in (19). Moreover, a square matrix $(\partial Q_j/\partial\theta^T)$ in (18) is a 7×7 symmetric matrix with elements of the matrix formulated as $\partial(\partial Q_j/\partial\theta)/\partial\theta^T$, where $\partial Q_j/\partial\theta$ as in (19).

By combining the derivative quotient rule and repeated use of the chain rule on (12b) with respect to θ, we have all elements of the mixed second order partial derivatives symmetric matrix $(\partial^2 Q_j/\partial\theta\partial\theta^T)$ or $(\partial \mathbf{Q}_j/\partial\theta^T)$ in (18) and setting the results to diagonal elements, we have:

$$\left. \begin{aligned} \frac{\partial^2 Q_j}{\partial\gamma_\ell^2} &= -A_{j\ell}\phi_{j\ell}, \text{ for } \ell = 1,2,3 \\ \frac{\partial^2 Q_j}{\partial\eta_\ell^2} &= -B_{j\ell}\zeta_{j\ell}, \text{ for } \ell = 1,2,3 \\ \frac{\partial^2 Q_j}{\partial\alpha^2} &= \varphi_j - 6\left((3\alpha-1)\mathcal{K}_j^\alpha - \alpha\mathcal{K}_j^{2\alpha} \right) \end{aligned} \right\} \tag{20}$$

Where:

$$\varphi_j = 6(\alpha-1) + 3\left[2(3\alpha-2)\mathcal{W}_j - (\alpha)_2\mathcal{V}_j \right]\mathcal{K}_j^{\alpha-1}$$
$$+ 3(\alpha-1)\left[\alpha\mathcal{U}_j^2 - \mathcal{W}_j^2 \right]\mathcal{K}_j^{\alpha-2}$$
$$+ 2\alpha\left[\alpha\mathcal{V}_j - 6\mathcal{W}_j \right]\mathcal{K}_j^{2\alpha-1} - 2\alpha\left[\alpha\mathcal{U}_j^2 - 2\mathcal{W}_j^2 \right]\mathcal{K}_j^{2\alpha-2}$$

$$\phi_{j\ell} = \psi_{1j} \ln\left(t_{j\ell}/\eta_\ell \right) + A_{j\ell}\psi_{2j}$$

$$\zeta_{j\ell} = (\alpha + \gamma_\ell)\psi_{1j}/\eta_\ell + B_{j\ell}\psi_{2j}$$

$$\psi_{2j} = (3(\alpha)_2(\alpha-1) - \alpha(2\alpha)_2\mathcal{K}_j^\alpha)\mathcal{K}_j^{\alpha-2}$$

$$\mathcal{U}_j = \alpha^{-1}\sum_{k=1}^{3}\gamma_k A_{jk}$$

$$\mathcal{V}_j = \alpha^{-1}\sum_{k=1}^{3}\gamma_k A_{jk} \ln\left(t_{jk}/\eta_k \right)^{\gamma_k/\alpha}$$

The diagonal elements of a square matrix $(\partial \mathbf{Q}_j / \partial \theta^T)$ in (18) which have been formulated more explicitly as in Eq. 20 consist of $\partial^2 Q_j / \partial \alpha^2$ itself and the diagonal elements of the following matrices: $(\partial^2 Q_j / \partial \gamma \partial \gamma^T)$ and $(\partial^2 Q_j / \partial \eta \partial \eta^T)$.

The $k \ell^{th}$ elements of matrix $(\partial \mathbf{Q}_j / \partial \theta^T)$ of Q_j as given in (12b), for $k, \ell = 1, 2, 3$ with $k \neq \ell$, can be written in a simple form:

$$\left. \begin{aligned} \frac{\partial^2 Q_j}{\partial \gamma_k \partial \gamma_\ell} &= \frac{\partial^2 Q_j}{\partial \gamma_\ell \partial \gamma_k} = -A_{jk}A_{j\ell}\psi_{2j} \\ \frac{\partial^2 Q_j}{\partial \eta_k \partial \eta_\ell} &= \frac{\partial^2 Q_j}{\partial \eta_\ell \partial \eta_k} = -B_{jk}B_{j\ell}\psi_{2j} \\ \frac{\partial^2 Q_j}{\partial \eta_k \partial \gamma_\ell} &= \frac{\partial^2 Q_j}{\partial \gamma_\ell \partial \eta_k} = A_{j\ell}B_{jk}\psi_{2j} \end{aligned} \right\} \qquad (21)$$

Notice that the elements (21) are off-diagonal elements of the following matrices: $(\partial^2 Q_j / \partial \gamma \partial \gamma^T)$, $(\partial^2 Q_j / \partial \gamma \partial \eta^T)$, $(\partial^2 Q_j / \partial \eta \partial \gamma^T)$ and $(\partial^2 Q_j / \partial \eta \partial \eta^T)$. Moreover, Eq.(21) is not including the elements of the following vectors: $\partial^2 Q_j / \partial \theta \partial \alpha$ and $\partial^2 Q_j / \partial \alpha \, \partial \theta$.

Using the same argument as given early, the remaining elements of matrix $(\partial Q_j / \partial \theta^T)$ for $\ell = 1, 2, 3$ can be express as follow:

$$\left. \begin{aligned} \frac{\partial^2 Q_j}{\partial \eta_\ell \partial \gamma_\ell} &= \frac{\partial^2 Q_j}{\partial \gamma_\ell \partial \eta_\ell} = B_{j\ell}\Phi_{j\ell} \\ \frac{\partial^2 Q_j}{\partial \alpha \partial \gamma_\ell} &= \frac{\partial^2 Q_j}{\partial \gamma_\ell \partial \alpha} = A_{j\ell}\Psi_{j\ell} \\ \frac{\partial^2 Q_j}{\partial \alpha \partial \eta_\ell} &= \frac{\partial^2 Q_j}{\partial \eta_\ell \partial \alpha} = -B_{j\ell}\Psi_{j\ell} \end{aligned} \right\} \qquad (22)$$

Where:

$$\Phi_{j\ell} = \alpha \psi_{1j} / \gamma_\ell + \phi_{j\ell}$$
$$\Psi_{j\ell} = \psi_{1j} \ln\left((t_{j\ell} / \eta_\ell)^{\gamma_\ell / \alpha} / \mathcal{K}_j^\alpha \right) + \psi_{2j} \mathcal{U}_j$$
$$\qquad - 3\alpha(3\alpha - 2)\mathcal{K}_j^{\alpha-1} + 2\alpha^2(3 + \ln \mathcal{K}_j^\alpha)\mathcal{K}_j^{2\alpha-1}$$

Obviously that the elements (22) consisting the diagonal elements of matrix $(\partial^2 Q_j / \partial \gamma \partial \eta^T)$ and $(\partial^2 Q_j / \partial \eta \partial \gamma^T)$. In addition, those elements are also covering elements of the following vectors: $\partial^2 Q_j / \partial \gamma \partial \alpha$, $\partial^2 Q_j / \partial \eta \partial \alpha$, $\partial^2 Q_j / \partial \alpha \partial \gamma$ and $\partial^2 Q_j / \partial \alpha \partial \eta$.

Hence, by plugging the elements of vector $\mathbf{Q}_j$ in (19) into the corresponding elements of score vector $\partial \mathcal{L}_2(\theta) / \partial \theta$ in (16), we can easily verify that:

$$\frac{\partial \mathcal{L}_2(\theta)}{\partial \gamma_\ell} = \sum_{j=1}^{n} \frac{-1}{Q_j} \alpha A_{j\ell}\psi_{1j}, \quad \text{for } \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_2(\theta)}{\partial \eta_\ell} = \sum_{j=1}^{n} \frac{1}{Q_j} \alpha B_{j\ell}\psi_{1j}, \quad \text{for } \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_2(\theta)}{\partial \alpha} = \sum_{j=1}^{n} \frac{1}{Q_j} \vartheta_j$$

Consequently, upon substituting (19), (20), (21) and (22) into (18) and simplifying the resulting expression, we obtain the entries of Hessian matrix $\mathbf{H}_2(\theta)$ as defined in Eq.(17), namely:

- For $\ell = 1, 2, 3$, we have:

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \gamma_\ell^2} = -\sum_{j=1}^{n} \frac{A_{j\ell}}{Q_j^2}\left[\phi_{j\ell}Q_j + \alpha^2 A_{j\ell}\psi_{1j}^2 \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_\ell^2} = -\sum_{j=1}^{n} \frac{B_{j\ell}}{Q_j^2}\left[\zeta_{j\ell}Q_j + \alpha^2 B_{j\ell}\psi_{1j}^2 \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \alpha^2} = \sum_{j=1}^{n} \frac{1}{Q_j}\left(\varphi_j - 6(3\alpha - 1)\mathcal{K}_j^\alpha + 6\alpha \mathcal{K}_j^{2\alpha} \right) - \sum_{j=1}^{n} \frac{\vartheta_j^2}{Q_j^2}$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_\ell \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \gamma_\ell \partial \eta_\ell} = \sum_{j=1}^{n} \frac{B_{j\ell}}{Q_j^2}\left[\Phi_{j\ell}Q_j + \alpha^2 A_{j\ell}\psi_{1j}^2 \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \alpha \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \gamma_\ell \partial \alpha} = \sum_{j=1}^{n} \frac{A_{j\ell}}{Q_j^2}\left[\Psi_{j\ell}Q_j + \alpha \vartheta_j\psi_{1j} \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \alpha \partial \eta_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_\ell \partial \alpha} = -\sum_{j=1}^{n} \frac{B_{j\ell}}{Q_j^2}\left[\Psi_{j\ell}Q_j + \alpha \vartheta_j\psi_{1j} \right]$$

- For $k, \ell = 1, 2, 3$ with $k \neq \ell$, we have:

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \gamma_k \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \gamma_\ell \partial \gamma_k} = -\sum_{j=1}^{n} \frac{A_{jk}A_{j\ell}}{Q_j^2}\left[\psi_{2j}Q_j + \alpha^2 \psi_{1j}^2 \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_k \partial \eta_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_\ell \partial \eta_k} = -\sum_{j=1}^{n} \frac{B_{jk}B_{j\ell}}{Q_j^2}\left[\psi_{2j}Q_j + \alpha^2 \psi_{1j}^2 \right]$$

$$\frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_k \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_2(\theta)}{\partial \eta_\ell \partial \gamma_k} = \sum_{j=1}^{n} \frac{A_{j\ell}B_{jk}}{Q_j^2}\left[\psi_{2j}Q_j + \alpha^2 \psi_{1j}^2 \right]$$

2.3. For the Third Part

Considering Eq. 12c that the first order partial derivatives of $\mathcal{L}_3(\theta)$ with respect to θ can be expressed as:

$$\mathbf{g}_3(\theta) = \frac{\partial \mathcal{L}_3(\theta)}{\partial \theta} \qquad (23)$$

Note that $\mathbf{g}_3(\theta)$ in (23) is the 7×1 score vector for θ, which elements of the score vector may be obtained as $\partial \mathcal{L}_3(\theta)/\partial \theta = (\partial \mathcal{L}_3(\theta)/\partial \gamma^T, \ \partial \mathcal{L}_3(\theta)/\partial \eta^T, \ \partial \mathcal{L}_3(\theta)/\partial \alpha)^T$. Therefore by applying the chain rule in deriving (12c) with respect to θ, the elements of the score vector $\mathbf{g}_3(\theta)$ above can be done easily:

$$\frac{\partial \mathcal{L}_3(\theta)}{\partial \gamma_\ell} = \frac{(\alpha - 3)}{\alpha} \sum_{j=1}^{n} A_{j\ell} \frac{1}{\mathcal{K}_j}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_3(\theta)}{\partial \eta_\ell} = -\frac{(\alpha - 3)}{\alpha} \sum_{j=1}^{n} B_{j\ell} \frac{1}{\mathcal{K}_j}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_3(\theta)}{\partial \alpha} = -\frac{1}{\alpha} \sum_{j=1}^{n} \left[\alpha \mathcal{V}_j^* - 3 \mathcal{U}_j \right] \frac{1}{\mathcal{K}_j}$$

where, $\mathcal{V}_j^* = \mathcal{U}_j - \mathcal{K}_j \ln(\mathcal{K}_j)$. Using similar arguments to (14) and (17), symmetric Hesian matrix $\mathbf{H}_3(\theta)$ of second order partial derivatives of $\mathcal{L}_3(\theta)$ can be represented as:

$$\mathbf{H}_3(\theta) = \left(\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \theta \partial \theta^T} \right) \qquad (24)$$

The diagonal entries of Hessian matrix (24) are found by deriving the right-hand side of Eq. 24 and take the form:

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \gamma_\ell^2} = \frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} A_{j\ell} A_{j\ell}^* \frac{1}{\mathcal{K}_j^2}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \eta_\ell^2} = \frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} B_{j\ell} B_{j\ell}^* \frac{1}{\mathcal{K}_j^2}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \alpha^2} = -\frac{1}{\alpha^2} \left\{ (\alpha - 3) \mathcal{U}_j^* + 6 \mathcal{K}_j \mathcal{U}_j \right\} \frac{1}{\mathcal{K}_j^2}$$

Where:

$$A_{j\ell}^* = \mathcal{K}_j \ln(t_{j\ell} / \eta_\ell) - A_{j\ell}, B_{j\ell}^* = ((\alpha + \gamma_\ell) / \eta_\ell) \mathcal{K}_j - B_{j\ell}.$$
$$\text{and} \ \mathcal{U}_j^* = \mathcal{U}_j^2 - \mathcal{K}_j \mathcal{V}_j$$

Similarly, the $k\ell^{th}$ entries of Hessian matrix (24), for k, $\ell = 1, 2, 3$ with $k \neq \ell$, can be derived to be:

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \gamma_k \partial \gamma_\ell} = -\frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} A_{jk} A_{j\ell} \frac{1}{\mathcal{K}_j^2}$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \eta_k \partial \eta_\ell} = -\frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} B_{jk} B_{j\ell} \frac{1}{\mathcal{K}_j^2}$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \eta_k \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \gamma_\ell \partial \eta_k} = \frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} B_{jk} A_{j\ell} \frac{1}{\mathcal{K}_j^2}$$

After some straightforward calculation, the other entries of Hessian matrix (24) are accomplished, i.e., for $\ell = 1, 2, 3$ satisfy:

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \eta_\ell \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \gamma_\ell \partial \eta_\ell} = -\frac{(\alpha - 3)}{\alpha^2} \sum_{j=1}^{n} B_{j\ell} \left[\frac{\alpha}{\gamma_\ell} \mathcal{K}_j + A_{j\ell}^* \right] \frac{1}{\mathcal{K}_j^2}$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \alpha \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \gamma_\ell \partial \alpha} = \frac{1}{\alpha^2} \sum_{j=1}^{n} A_{j\ell} \xi_{j\ell} \frac{1}{\mathcal{K}_j^2}$$

$$\frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \alpha \partial \eta_\ell} = \frac{\partial^2 \mathcal{L}_3(\theta)}{\partial \eta_\ell \partial \alpha} = -\frac{1}{\alpha^2} \sum_{j=1}^{n} B_{j\ell} \xi_{j\ell} \frac{1}{\mathcal{K}_j^2}$$

Where:

$$\xi_{j\ell} = (\alpha - 3) \left[\mathcal{U}_j - \mathcal{K}_j \ln(t_{j\ell} / \eta_\ell)^{\gamma_\ell / \alpha} \right] + 3 \mathcal{K}_j$$

2.4. For the Fourth Part

Let $\mathbf{g}_4(\theta)$ be defined in accordance with (13), (15) and (23). Then, for the $\mathcal{L}_4(\theta)$ as given in (12d), we get the score vector, which is written as:

$$\mathbf{g}_4(\theta) = \frac{\partial \mathcal{L}_4(\theta)}{\partial \theta} \qquad (25)$$

By employing the right-hand side of (25), the elements of score vector $\mathbf{g}_4(\theta)$ above can be easily calculated:

$$\frac{\partial \mathcal{L}_4(\theta)}{\partial \gamma_\ell} = -\sum_{j=1}^{n} A_{j\ell} \mathcal{K}_j^{\alpha - 1}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_4(\theta)}{\partial \eta_\ell} = \sum_{j=1}^{n} B_{j\ell} \mathcal{K}_j^{\alpha - 1}, \ \text{for} \ \ell = 1, 2, 3$$

$$\frac{\partial \mathcal{L}_4(\theta)}{\partial \alpha} = \sum_{j=1}^{n} \left[\mathcal{U}_j - \mathcal{K}_j \ln(\mathcal{K}_j) \right] \mathcal{K}_j^{\alpha - 1}$$

33

As before, we let $\mathbf{H}_4(\boldsymbol{\theta})$ be defined similar to (14), (17) and (24). Then, for the $\mathcal{L}_4(\boldsymbol{\theta})$ as already mentioned in (12d), we have:

$$\mathbf{H}_4(\boldsymbol{\theta}) = \left(\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathrm{T}}} \right) \tag{26}$$

By differentiating $\mathcal{L}_4(\boldsymbol{\theta})$ in Eq. 12d two times with respect to the same elements of vector $\boldsymbol{\theta}$, the diagonal entries of Hessian matrix (26) can be found immediately and we have the following results:

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \gamma_\ell^2} = -\frac{1}{\alpha} \sum_{j=1}^n A_{j\ell} \left[\alpha A_{j\ell} + A_{j\ell}^* \right] \mathcal{K}_j^{\alpha-2}, \text{ for } \ell = 1,2,3$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \eta_\ell^2} = -\frac{1}{\alpha} \sum_{j=1}^n B_{j\ell} \left[\alpha B_{j\ell} + B_{j\ell}^* \right] \mathcal{K}_j^{\alpha-2}, \text{ for } \ell = 1,2,3$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \alpha^2} = -\frac{1}{\alpha} \sum_{j=1}^n \left[\alpha \mathcal{V}_j^{*2} - \mathcal{U}_j^* \right] \mathcal{K}_j^{\alpha-2}.$$

In a similar manner to those above, we can treat the off-diagonal entries of Hessian matrix (26), for k, $\ell = 1$, 2, 3 with $k \neq \ell$, to that off-diagonal entries and write them in the form of:

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \gamma_k \partial \gamma_\ell} = -\frac{(\alpha-1)}{\alpha} \sum_{j=1}^n A_{jk} A_{j\ell} \, \mathcal{K}_j^{\alpha-2}$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \eta_k \partial \eta_\ell} = -\frac{(\alpha-1)}{\alpha} \sum_{j=1}^n B_{jk} B_{j\ell} \, \mathcal{K}_j^{\alpha-2}$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \eta_k \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \gamma_\ell \partial \eta_k} = \frac{(\alpha-1)}{\alpha} \sum_{j=1}^n B_{jk} A_{j\ell} \, \mathcal{K}_j^{\alpha-2}$$

Analogously, for k = ℓ = 1, 2, 3, it follows that with a simple algebraic manipulation yields the rest entries of the Hessian matrix (26):

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \eta_\ell \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \gamma_\ell \partial \eta_\ell} = \frac{1}{\alpha} \sum_{j=1}^n B_{j\ell} \left[\frac{\alpha}{\gamma_\ell} \mathcal{K}_j + \alpha A_{j\ell} + A_{j\ell}^* \right] \mathcal{K}_j^{\alpha-2}$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \alpha \partial \gamma_\ell} = \frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \gamma_\ell \partial \alpha} = \frac{1}{\alpha} \sum_{j=1}^n A_{j\ell} \Upsilon_{j\ell} \, \mathcal{K}_j^{\alpha-2}$$

$$\frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \alpha \partial \eta_\ell} = \frac{\partial^2 \mathcal{L}_4(\boldsymbol{\theta})}{\partial \eta_\ell \partial \alpha} = -\frac{1}{\alpha} \sum_{j=1}^n B_{j\ell} \Upsilon_{j\ell} \, \mathcal{K}_j^{\alpha-2}$$

where, $\Upsilon_{j\ell} = (\alpha-1)\mathcal{U}_j + \mathcal{K}_j \ln\left((t_{j\ell}/\eta_\ell)^{\gamma_\ell/\alpha} / \mathcal{K}_j^\alpha \right)$. Finally, all calculations are performed by means of analytical expressions.

From (13), (15), (23) and (25) we get the score vector of log likelihood function (9) as below Eq. 27:

$$\mathbf{g}(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^4 \frac{\partial \mathcal{L}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{27}$$

Similarly, from (14), (17), (24) and (26) we have the Hessian matrix of the log likelihood function (9), namely Eq. 28:

$$\mathbf{H}(\boldsymbol{\theta}) = \left(\frac{\partial^2 \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathrm{T}}} \right) = \sum_{i=1}^4 \left(\frac{\partial^2 \mathcal{L}_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^{\mathrm{T}}} \right) \tag{28}$$

Since the score vector and the Hessian matrx of the log likelihood are now known, the Newton Raphson or other gradient-based methods can be applied to calculate $\hat{\boldsymbol{\theta}}$ that optimizes $\mathcal{L}(\boldsymbol{\theta})$ in Eq. 9.

3. RESULTS

Let $\hat{\boldsymbol{\theta}}$ be the maximizer of (9). When we insert the MLE $\hat{\boldsymbol{\theta}}$ into the survival function $\hat{S}_T(t_j) = \exp(-\hat{\mathcal{K}}_j^{\hat{\alpha}})$ with $\hat{\mathcal{K}}_j = \sum_{\ell=1}^3 (t_{j\ell}/\hat{\eta}_\ell)^{\hat{\gamma}_\ell/\hat{\alpha}}$ as in (3), then we will get the MLE of $S_T(t_j)$ and an estimator of $\mathcal{K}_j$ of the j-th individual, because the MLEs are functional invariant: i.e., when $\hat{\boldsymbol{\theta}}$ is the MLE of $\boldsymbol{\theta}$, h($\hat{\boldsymbol{\theta}}$) will also be the MLE of h($\boldsymbol{\theta}$) if h($\boldsymbol{\theta}$) is a finite function and need not be one to one. When we substitute $\hat{\boldsymbol{\theta}}$ to the Hessian matrix (11) we will get the estimated covariance matrix of the MLE $\hat{\boldsymbol{\theta}}$, that is given by $\hat{\mathbf{V}}(\hat{\boldsymbol{\theta}}) = \hat{C}\mathrm{ov}(\hat{\boldsymbol{\theta}}) = \left[-\mathbf{H}(\hat{\boldsymbol{\theta}}) \right]^{-1}$.

Putting the estimators $\hat{\boldsymbol{\theta}}$ and $\hat{\mathcal{K}}_j$ into (6), we obtain the scalar quantity trivariate Weibull hazard rate estimator, that is:

$$\hat{h}_T(t_j) = \left[\prod_{\ell=1}^3 \left(\frac{1}{\hat{\alpha}} \right) \left(\frac{\hat{\gamma}_\ell}{\hat{\eta}_\ell} \right) \left(\frac{t_{j\ell}}{\hat{\eta}_\ell} \right)^{(\hat{\gamma}_\ell/\hat{\alpha})-1} \right]$$
$$\times \sum_{k=1}^3 (-1)^{k+3} \mathcal{C}(3,k,\hat{\alpha}) \, \hat{\mathcal{K}}_j^{k\hat{\alpha}-3}$$

This hazard rate estimator is useful for characterizing trivariate exponential distributions as the special cases of trivariate Weibull distributions.

In the vector-valued trivariate hazard rete function, the trivariate Weibull hazard rate estimator (7) can be written as:

$$\widehat{h}_\mathbf{T}(\mathbf{t_j}) = \left(\widehat{h}_\mathbf{T}(\mathbf{t_j})_1, \widehat{h}_\mathbf{T}(\mathbf{t_j})_2, \widehat{h}_\mathbf{T}(\mathbf{t_j})_3\right)^\mathrm{T}$$

where, $h_\mathbf{T}(\mathbf{t_j})_\ell = -\partial \ln S_\mathbf{T}(\mathbf{t_j}) / \partial t_\ell$ for $\ell = 1, 2, 3$ or it is given explicitly by:

$$\widehat{h}_\mathbf{T}(\mathbf{t_j})_\ell = \frac{\hat{\gamma}_\ell}{\hat{\eta}_\ell}\left(\frac{t_{j\ell}}{\hat{\eta}_\ell}\right)^{(\hat{\gamma}_\ell/\hat{\alpha})-1} \widehat{\mathcal{K}}_j^{\hat{\alpha}-1}$$

This hazard rate estimator plays an important role in parametric baseline hazards function with the trivariate Weibull baseline distribution as the baseline hazard, in particular to parametrize the baseline hazard function of the trivariate Cox regression model.

4. DISCUSSION

The Weibull distributed failure times are often used in practice because they seem to be able to describe the actual evolution of the hazard function in an appropriate way in many circumstances.

5. CONCLUSION

In this study we considered a continuous trivariate Weibull distribution which has seven parameters. It is observed that the maximum likelihood estimator of the unknown parameters can be obtained by solving the interdependent non linear system of the first order partial derivatives of log likelihood after setting to zero. Since the optimization algorithms in general are iterative procedures which start with an initial estimate of the solution and converge to a single solution, one or more local maxima may be missed. This problem can be overcome by performing the optimization several times using a different initial estimate for the solution in each case until the desired accuracy has been reached. Once unknown parameters estimated, trivariate hazard rate of trivariate Weibull distributions are also estimated.

6. REFERENCES

Cai, J., J. Fan, H. Zhou and Y. Zhou, 2007. Hazard models with varying coefficients for multivariate failure time data. Ann. Stat., 35: 324-354. DOI: 10.1214/009053606000001145

Hanagal, D.D., 2005. A bivariate weibull regression model. Econ. Q. Control, 20: 1-8.

Hanagal, D.D., 2006. Weibull extension of bivariate exponential regression model with gamma frailty for survival data. Econ. Q. Control, 21: 261-270. DOI: 10.1515/EQC.2006.261

Kundu, D. and R.D. Gupta, 2010a. A class of absolutely continuous bivariate distributions. J. Stamet., 7: 464-477. DOI: 10.1016/j.stamet.2010.01.004

Kundu, D. and R.D. Gupta, 2010b. Modified Sarhan-Balakrishnan singular bivariate distribution. J. Stat. Plann. Infer., 140: 526-538. DOI: 10.1016/j.jspi.2009.07.026

Lee, C.K. and M.J. Wen, 2009. A multivariate weibull distribution. Pak. J. Stat. Operat. Res., 2: 55-66.

Navarro, J., 2007. Characterizations using the bivariate failure rate function. Stat. Prob. Lett., 78: 1349-1354. DOI: 10.1016/j.spl.2007.12.004

Singpurwalla, N.D., 2006. The hazard potential: Introduction and overview. J. Am. Stat. Assoc., 101: 1705-1716. DOI: 10.1198/016214506000001068

Tuerlinckx, F., 2004. A multivariate counting process with Weibull-distributed first-arrival times. J. Math. Psychol., 48: 65-79. DOI: 10.1016/j.jmp.2003.12.001