A Study on Estimation of Lifetime Distribution with Covariates Using Online Monitoring

Masahiro Yokoyama^{*1}, Watalu Yamamoto², Kazuyuki Suzuki²

1. Chuo University: 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, Japan

2. The University of Electro-Communications: 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan

**contact author's e-mail address: masahiro@indsys.chuo-u.ac.jp*

Abstract:

In these days, online monitoring becomes a common tool for keeping highly reliability of products and systems. The online monitoring information which includes usage history, system conditions, and environmental conditions is reported and stored as big data. On statistical modeling, these variables from the online monitoring are primary candidates for covariates which affect the failure mechanism. There is some literature on modeling by the cumulative exposure model for the products lifetime distribution with covariate effects. The existing literatures require the already known parametric baseline distribution of the cumulative exposure. However such knowledge may be difficult to acquire in advance in some cases. When an incorrect baseline distribution is assumed, it is called misspecification. This paper proposes the strategy which use a likelihood function under a log-normal distribution to estimate parameters which represent covariate effects when the truly underlying baseline distribution is either a Weibull distribution or a log-normal distribution. In this paper, it is derived that the score function of a likelihood function under a log-normal baseline distribution is identified as the approximation for a Weibull cases. Besides, the simulation study and the discussion for the bias of estimation are shown, and this paper clarify the relationship among distribution parameters and the bias of estimation under misspecification.

Keywords

Misspecification, Weibull distribution, Log-normal distribution, Cumulative exposure, Covariates

1. Introduction

Online monitoring in real time using the Internet has become a common tool for keeping products and systems operating reliably. The information gathered includes usage history, system conditions, and environmental conditions and is stored as big data. In the statistical modeling of failure mechanisms, the variables contained in this information are primary candidates for covariates that affect a failure mechanism. Information on covariates can be used to improve product reliability. For example, in the analysis of the lifetime distribution of a truck engine, the lifetime is usually estimated on the basis of the distance traveled (mileage) before catastrophic failure. However, other variables, such as the average slope of the routes traveled and the average weight of the loads carried, affect the load on the engine, so lifetimes can differ even though the mileage is the same. Therefore, mileage is not always sufficient information for optimizing inspection and maintenance times.

2. Previous studies and our proposal

Meeker and Hong (2014) provided a general method that utilizes dynamic covariate information to predict field-failure. In the study of Hong and Meeker (2013), the cumulative exposure (CE) model was used to describe the effect of a dynamic covariate on the failure-time distribution. Nelson (1990, 2001) described the CE model in the context of life tests to incorporate time-varying covariates into failure-time models. The CE model is also known as the cumulative damage (CD) model (Bagdonavicius and Nikulin (2001)). Besides, Meeker and Escobar (1998) called the CE model a “proportional quantities (PQ) model” or a “scale accelerated failure-time

(SAFT) model.” Furthermore, Nelson (2001) remarked that the CE model can be considered as a limit of the step-stress model when the length of each step interval goes to zero.

Given the entire covariate history, the cumulative exposure t^* by time t defined as

$$t^*[t; \boldsymbol{\beta}, \mathbf{z}(t)] = \int_0^t \exp[\boldsymbol{\beta}^{\text{tr}} \mathbf{z}(s)] ds. \quad (1)$$

Here, $\boldsymbol{\beta}$ is a vector of covariate parameters (acceleration coefficients) and $\mathbf{z}(t)$ is a covariate vector acquired at continuous time point t , and tr represents transposed. The value of the covariates are either discrete or continuous. Equation (1) represents a nonlinear transformation of t to quantity t^* , which incorporates dynamic covariate information for the lifetime prediction of individual user’s products.

Let T be the failure time of the unit and $\boldsymbol{\beta}^*$ be the true value of $\boldsymbol{\beta}$. The unit fails at time T when the amount of cumulative exposure reaches a random threshold $T^*(\boldsymbol{\beta}^*)$. Thus, the relationship between the cumulative exposure $T^*(\boldsymbol{\beta}^*)$ and the failure time T is

$$T^*(\boldsymbol{\beta}^*) = \int_0^T \exp[\boldsymbol{\beta}^{*\text{tr}} \mathbf{z}(s)] ds. \quad (2)$$

The cumulative exposure threshold $T^*(\boldsymbol{\beta}^*)$ has the baseline cumulative distribution function (cdf) $F_0(T^*)$ with parameter $\boldsymbol{\theta}_0$. Equation (2) represents an accumulation of damage. Thus, if the product is heavily used or used in a harsh environment, the product fails sooner.

For this approach based on the CE model, the baseline distribution $F_0(T^*)$ is needed to identify in advance and it is usually assumed that $F_0(T^*)$ is a log-location-scale distribution (e.g., Weibull, log-normal). In the study of Hong and Meeker (2013), a Weibull distribution is used for the CE of actual products and it is said that engineering knowledge based on laboratory tests, previous experience with similar failure modes, or knowledge of the physics of failure can also provide useful information to identify the baseline distribution. As a special case, the baseline distribution is the same as the distribution of the product failure time T , when it is used under a constant covariates. However consumer products are used under various environment and the CE value $T^*(\boldsymbol{\beta}^*)$ cannot be observed. There is a research about a statistical test to identify the baseline distribution in the context of accelerated failure time models (Bagdonavicius, Levulienė and Nikulin (2013)). On the other hand, our study focus on the estimation bias of $\boldsymbol{\beta}$ under misspecification for the baseline distribution between a log-normal distribution and a Weibull distribution.

We propose a strategy and its validity to use a likelihood function under a log-normal distribution to estimate the covariate effects parameter $\boldsymbol{\beta}$ when the underlying distribution for $F_0(T^*)$ is either assumed Weibull or log-normal. With this approach, a model can be built that incorporates covariates even if a Weibull baseline distribution is misspecified as a log-normal one. This study makes the lifetime prediction for individual user’s products using the observed covariates easy to use.

In section 3, the derived score function under a log-normal baseline distribution is shown to approximate the score function under assuming a Weibull baseline distribution. The numerical simulation study is shown in section 4 to clarify the relationship among distribution parameters and the estimation of the covariate parameter $\boldsymbol{\beta}$ under misspecification. In section 5, we discuss a bias of estimation from the results of simulation. For the case of a Weibull distributions, if a shape parameter is large, the bias of the estimation is small.

3. Analytical study using likelihood function

3.1. Covariate

Here we deal with the case in which a target covariate is observed at time $t_j^o (j = 1, \dots, J)$ for a discrete interval (Figure 1). The covariate value is assumed to be constant between each pair of t_j^o .

Figure 1. Covariate z_q observed at discrete intervals ($J = 4$)

The symbols here used are defined as follows.

T	Lifetime
t_j^o	Covariate acquisition points ($j = 1, \dots, J$)
J	Number of covariate acquisition points up to failure
q	Covariate type ($q = 1, \dots, Q$, Q : total number of covariates)
$z_q(t_j^o)$	Value of covariate acquired at t_j^o

The nonlinear transformation of lifetime T to quantity $T^*(\beta)$ on the basis (1) is illustrated in Figure 2.

Figure 2. Transformation to $T^*(\beta)$

Let $i, (i = 1, \dots, n)$ be the product number of each user, T_i be the lifetime of each product, $\mathbf{z}(t_{ij}^o)$ be the value of the observed covariate at t_{ij}^o for each product, and $T_i^*(\beta)$ be the amount of transformation for each product.

3.2. Likelihood function and score function

We define μ as the average and σ as the standard deviation of the log-transformed log-normal distribution. The likelihood function for a log-normal distribution is given by

$$\log L_L(\beta, \mu, \sigma) = \sum_{i=1}^n (\beta \mathbf{z}(t_{i,B_i}^o)) - \sum_{i=1}^n \log(\sqrt{2\pi}\sigma) - \sum_{i=1}^n \log(T_i^*(\beta)) + \sum_{i=1}^n \left\{ -\frac{(\log(T_i^*(\beta)) - \mu)^2}{2\sigma^2} \right\}. \quad (3)$$

The score function obtained by differentiating this function with respect to β is given by

$$g_L(\boldsymbol{\beta}) = \sum_{i=1}^n (\mathbf{z}(t_{i,B_i}^o)) - \sum_{i=1}^n \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} - \frac{1}{\hat{\sigma}} \cdot \sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})}{\hat{\sigma}} \right) \cdot \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} \right\}. \quad (4)$$

In the maximum likelihood estimation (MLE) for $T^*(\boldsymbol{\beta}^*)$ assuming a log-normal distribution, the MLE is given by an estimation of $\boldsymbol{\beta}$ in which the score function is 0.

The log-likelihood function for a Weibull distribution (log-location-scale) is given by

$$\log L_W(\boldsymbol{\beta}, \mu, \sigma) = \sum_{i=1}^n (\mathbf{z}(t_{i,B_i}^o)) - \sum_{i=1}^n \log(\sigma) - \sum_{i=1}^n \log(T_i^*(\boldsymbol{\beta})) + \sum_{i=1}^n \left\{ \frac{(\log(T_i^*(\boldsymbol{\beta})) - \mu)}{\sigma} - \exp\left[\frac{(\log(T_i^*(\boldsymbol{\beta})) - \mu)}{\sigma}\right] \right\}. \quad (5)$$

The score function obtained by differentiating this function with respect to $\boldsymbol{\beta}$ is given by

$$g_W(\boldsymbol{\beta}) = \sum_{i=1}^n (\mathbf{z}(t_{i,B_i}^o)) - \sum_{i=1}^n \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} - \frac{1}{\hat{\sigma}} \cdot \sum_{i=1}^n \left\{ \left(\exp\left[\frac{(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})}{\hat{\sigma}}\right] - 1 \right) \cdot \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} \right\}. \quad (6)$$

Let $\exp[(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})/\hat{\sigma}]$ be $\exp[A]$. By Taylor expansion around $A = 0$, we get $\exp[A] \approx \exp(0) + \exp(0) \cdot A + \exp(0) \cdot A^2/2 = 1 + A + A^2/2$. Therefore, the second-order approximation is given by

$$\begin{aligned} g_W(\boldsymbol{\beta}) &\approx \sum_{i=1}^n (\mathbf{z}(t_{i,B_i}^o)) - \sum_{i=1}^n \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} - \frac{1}{\hat{\sigma}} \cdot \sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})}{\hat{\sigma}} \right) \cdot \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} + \frac{1}{2} \cdot \left(\frac{(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})}{\hat{\sigma}} \right)^2 \cdot \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} \right\} \\ &= g_L(\boldsymbol{\beta}) - \frac{1}{2\hat{\sigma}} \cdot \sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\boldsymbol{\beta})) - \hat{\mu})}{\hat{\sigma}} \right)^2 \cdot \frac{T_i^{*'}(\boldsymbol{\beta})}{T_i^*(\boldsymbol{\beta})} \right\} \approx g_L(\boldsymbol{\beta}) \end{aligned} \quad (7)$$

By using this approximate relationship, we can approximately estimate $\boldsymbol{\beta}$ by using MLE assuming a log-normal distribution, even if T follows a Weibull distribution.

4. Simulation study and results

4.1 Simulation setting

In this simulation study, $\boldsymbol{\beta}$ was estimated from covariate $\mathbf{z}(t_{ij}^o)$ and lifetime T_i using MLE under assuming $T^*(\boldsymbol{\beta}^*)$ as a log-normal distribution. These data were generated as explained below.

(1). $T_i^*(\boldsymbol{\beta}^*)$

The values of $T_i^*(\boldsymbol{\beta}^*)$ were generated in accordance with a log-normal distribution or a Weibull distribution. In section 4.3.1, the results of estimation under no misspecification is shown. On the other hand, in section 4.3.2, the results of estimation under misspecification for a Weibull distribution is shown. In the case of section 4.3.2, for a Weibull distribution, the shape parameter m was set as 1, 2, 3, 4, 5, 6 and the location parameter η was set as 50, 100. Here, ($\eta = 100$) means that covariates of double quantity than ($\eta = 50$) are observed by failure occurrence.

(2). $\mathbf{z}(t_{ij}^o)$

As was explained above for Figure 1, the values of $\mathbf{z}(t_{ij}^o)$ was generated as constant during each observation. We assumed two kinds of covariate; covariates $z_1(t_{ij}^o)$ and $z_2(t_{ij}^o)$ were independently generated for each product for each observed time point as following a normal distribution $N(1.0, 0.01)$, that is taking positive values. Thus, in our simulation, it is assumed that each product is used under a normal condition that time-varying covariates have no trend for time. Here, true values of parameter vector $\boldsymbol{\beta}^*$ were set as 1.00 for

both.

(3). T_i

The values of T_i were determined from $T_i^*(\beta^*)$ and $\mathbf{z}(t_{ij}^o)$ generated as described above. The interval between each pair of covariate observations was determined as follows.

$$t_{i1}^o = (t_{i2}^o - t_{i1}^o) = \dots = (t_{i,J_i-1}^o - t_{i,J_i-2}^o) = 1 \quad (\text{Interval is constant.})$$

$$((t_{i,J_i}^o - t_{i,J_i-1}^o) \leq 1 \quad (\text{Interval is less than 1 only for last observation.}))$$

Here,

$$T_i^*(\beta^*) = \sum_{j=1}^{J_i-1} \exp[\beta_1^* \cdot z_1(t_{ij}^o) + \beta_2^* \cdot z_2(t_{ij}^o)] + \exp[\beta_1^* \cdot z_1(t_{i,J_i}^o) + \beta_2^* \cdot z_2(t_{i,J_i}^o)] \cdot (t_{i,J_i}^o - t_{i,J_i-1}^o).$$

Therefore, lifetime T_i for each product was obtained from the summation of t_{i,j_i}^o .

4.2 Comparison of estimations with coefficient of variation

Kordonsky and Gertsbakh (1993) proposed using a time scale that minimizes the coefficient of variation (*c. v.*). A method using this approach does not need a parametric form of the distribution.

$$c. v. [T^*(\beta)] = \frac{\sqrt{\text{var}[T^*(\beta)]}}{\text{mean}[T^*(\beta)]},$$

where $\text{var}[T^*(\beta)]$ is the sample variance of $T^*(\beta)$ and $\text{mean}[T^*(\beta)]$ is the sample mean of $T^*(\beta)$. In section 4.3.1, we compared this method with MLE.

4.3 Simulation results

4.3.1 Maximum likelihood estimate of $T^*(\beta^*)$ that follows a log-normal distribution

First we show that $T^*(\beta^*)$ can be correctly estimated using MLE assuming a log-normal distribution for $T_i^*(\beta^*)$ generated in accordance with a log-normal distribution. Table 1 shows the estimation result of β . The μ and σ represent the set value of the mean value and the standard deviation value of the log-transformed log-normal distribution, respectively. The estimation from simulation data was repeated 500 times. The results shown in the upper rows are the average estimated values, and those shown in the lower rows are the standard deviations. These results show that the value of β was correctly estimated by MLE.

Table 1. Estimations obtained for $T_i^*(\beta^*)$ when $T_i^*(\beta^*)$ was generated using a log-normal distribution

($\beta_1^* = \beta_2^* = 1.00$, $\mu = 2.0$, $n = 1000$, no. of repetitions = 500)

set value		proposed method		minimizing coefficient of variation	
μ	σ	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
2.0	1	0.991	0.997	0.165	0.119
		0.301	0.286	1.305	1.489
2.0	0.7	0.994	0.991	0.619	0.605
		0.209	0.222	0.489	0.493
2.0	0.5	1.001	0.998	0.801	0.795
		0.168	0.163	0.244	0.053
2.0	0.1	1.001	0.999	0.979	0.977
		0.032	0.035	0.032	0.035
2.0	0.01	1.000	1.000	1.000	1.000
		0.003	0.004	0.003	0.004

(Upper rows contain estimate means, and lower rows contain standard deviations.)

On the other hand, the right side of Table 1 shows the estimate results for the coefficient of variation minimization method. Estimation accuracy which bias and variance of estimated value were good when the value of scale parameter σ was large. However, because of non-parametric approach, the bias and variance of estimated value by the coefficient of variation minimization method were larger than that of the MLE assuming a log-normal distribution, especially on a case having low σ .

4.3.2 Maximum likelihood estimate of $T_i^*(\beta^*)$ that follows a Weibull distribution

Here we show that β can be approximately estimated using MLE assuming a log-normal distribution for $T_i^*(\beta^*)$ generated in accordance with a Weibull distribution. That is, in this simulation study, β is estimated when a Weibull distribution is misspecified as a log-normal distribution.

The true value of β was set to 1.00, i.e., $\beta_1^* = \beta_2^* = 1.00$. The estimation of β was repeated 1000 times. Table 2 shows the mean value and the standard deviations of estimated values. The n represents the number of samples, m represents the shape parameter, and η represents the scale parameter of the Weibull distribution. Figure 3 graphically represents the estimates of β_1 for $\eta = 100$ in Table 2.

The results in Table 2 and Figure 3 for different distribution assumptions show that parameter β can be approximately obtained. When the value of the shape parameter m of the Weibull distribution was large, the bias of the resulting estimates of β was small.

Table 2. Estimations obtained for β when $T_i^*(\beta^*)$ was generated using a Weibull distribution ($\beta_1^* = \beta_2^* = 1.00$, no. of repetitions = 1000)

m = 1 $\eta = 50$				m = 1 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.974	0.960	0.987	mean	0.957	0.958	0.955
SD	0.322	0.138	0.085	SD	0.320	0.140	0.099
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	0.945	0.957	0.987	mean	0.947	0.963	0.953
SD	0.326	0.138	0.092	SD	0.310	0.143	0.099
m = 2 $\eta = 50$				m = 2 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.951	0.961	0.966	mean	0.983	0.974	0.975
SD	0.282	0.132	0.085	SD	0.301	0.133	0.093
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	0.958	0.957	0.965	mean	0.972	0.979	0.969
SD	0.275	0.127	0.091	SD	0.301	0.135	0.094
m = 3 $\eta = 50$				m = 3 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.978	0.980	0.984	mean	0.987	0.983	0.987
SD	0.245	0.108	0.078	SD	0.285	0.125	0.092
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	0.977	0.983	0.987	mean	1.001	0.987	0.988
SD	0.249	0.111	0.080	SD	0.275	0.122	0.091
m = 4 $\eta = 50$				m = 4 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.992	0.987	0.991	mean	0.997	0.995	0.994
SD	0.221	0.098	0.069	SD	0.258	0.113	0.081
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	1.011	0.989	0.993	mean	1.001	0.995	0.993
SD	0.215	0.096	0.066	SD	0.256	0.115	0.079
m = 5 $\eta = 50$				m = 5 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.996	0.995	0.994	mean	0.997	0.996	0.996
SD	0.188	0.083	0.059	SD	0.232	0.102	0.073
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	0.998	0.998	0.997	mean	0.997	0.995	0.995
SD	0.185	0.081	0.062	SD	0.228	0.104	0.071
m = 6 $\eta = 50$				m = 6 $\eta = 100$			
$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_1$	n = 1000	n = 5000	n = 10000
mean	0.995	1.000	1.000	mean	0.999	0.999	0.998
SD	0.159	0.074	0.050	SD	0.202	0.093	0.063
$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000	$\hat{\beta}_2$	n = 1000	n = 5000	n = 10000
mean	0.997	0.998	1.000	mean	1.004	0.994	1.000
SD	0.157	0.074	0.052	SD	0.208	0.089	0.066

(mean represents mean value of each estimation; SD represents standard deviation of each estimation)

Figure 3. Estimation results for β_1 for ($\eta = 100$, $n = 10000$) case in Table 2

($\beta_1^* = \beta_2^* = 1.00$, no. of repetitions = 1000)

(Center of plot represents mean value of 1000 repetitions;
upper and lower lines represent mean value $\pm$ standard deviation)

Next, we investigated the variation in the estimated results in each value of shape parameter m . Figure 4 shows a box plot (for $\eta = 100$, $n = 10000$) of estimation result β_1 in Table 2. The greater the value of shape parameter m , the smaller the variation in the estimate.

Figure 4. Box plot of estimation result β_1 for ($\eta = 100$, $n = 10000$) case in Table 2
($\beta_1^* = \beta_2^* = 1.00$, no. of repetitions = 1000)

5. Discussion

5.1 Bias of estimation under misspecification

In this section, in the case of the CE model containing a univariate covariate, we discuss a bias of estimation for a covariate parameter β under misspecification that the true model is a Weibull distribution and the incorrect model is a log-normal distribution. The results in White (1982) and Pascual (2005) can be used to discuss the asymptotic distribution of the MLE when the assumed model is incorrect. We call this incorrect MLE quasi-MLE (QMLE), as is done in White (1982), and use MLE to call the estimators under true assumption, i.e. under the model with no misspecification. We investigate the asymptotic bias of the estimator for a covariate parameter β . This kind of investigation is model-specific and the results given here are for the case of misspecification that the model by a Weibull distribution is true and log-normal distribution is incorrect.

Here, let $\hat{\beta}_L$ be the QMLE which is the value of β that maximizes equation (3) for the data that $T^*(\beta^*)$ follows a Weibull distribution. Additionally, let β_L be the asymptotic mean of the QMLE $\hat{\beta}_L$,

$$E[g_L(\beta_L)] = 0.$$

Here, $E[\cdot]$ represents expected value. Thus, from equation (7), equation (8) holds

$$E[g_w(\beta_L)] \approx E[g_L(\beta_L)] - \frac{1}{2\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_L)) - \mu_L)}{\sigma_L} \right)^2 \cdot \frac{T_i^{*'}(\beta_L)}{T_i^*(\beta_L)} \right\} \right]. \quad (8)$$

Here, $\mu_L = \text{mean}[\log(T^*(\beta_L))]$ and $\sigma_L^2 = \text{var}[\log(T^*(\beta_L))]$. Therefore the term of $\left(\frac{(\log(T_i^*(\beta_L)) - \mu_L)}{\sigma_L} \right)$ represents standardization. Under the case of continuous covariate, the term of $\frac{T_i^{*'}(\beta_L)}{T_i^*(\beta_L)}$ is shown as

$$\frac{T_i^{**}(\beta_L)}{T_i^*(\beta_L)} = \frac{\int_0^T z(s) \cdot \exp[\beta_L \cdot z(s)] ds}{\int_0^T \exp[\beta_L \cdot z(s)] ds}.$$

Additionally, under our condition setting, the value of covariate were set for positive values. Therefore, from equation (8), equation (9) holds under our condition setting,

$$E[g_W(\beta_L)] \approx -\frac{1}{2\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_L)) - \mu_L)}{\sigma_L} \right)^2 \cdot \frac{T_i^{**}(\beta_L)}{T_i^*(\beta_L)} \right\} \right] < 0. \quad (9)$$

Here, let $\hat{\beta}_W$ be the Weibull MLE which is the value of β that maximizes equation (5) for the data that $T^*(\beta^*)$ follows a Weibull distribution. Additionally, let β_W be the asymptotic mean of the MLE $\hat{\beta}_W$,

$$E[g_W(\beta_W)] = 0$$

Therefore the value of $(\beta_W - \beta_L)$ has a positive value under our setting of condition, because $g_W(\beta)$ is a monotone decreasing function. The simulation results show that QMLE $\hat{\beta}_L$ has a lower value than the true value ($\beta_W = 1.00$). Therefore, this result accord with the numerical simulation. Incidentally, when the value of covariate were set for negative values, the value of $(\beta_W - \beta_L)$ has a negative value because of $E[g_W(\beta_L)] > 0$.

5.2 Relationship between scale parameter and bias of estimation under misspecification

From the results of our numerical simulation under misspecification, it is confirmed that, if a shape parameter m of a Weibull distribution is large, the bias of the resulting estimates for parameter β is small. In this section, we discuss the relationship between the scale parameter $\sigma (= 1/m)$ and the asymptotic mean of bias.

Now, from equation (7), equation (10) holds,

$$E[g_L(\beta_W)] \approx E[g_W(\beta_W)] + \frac{1}{2\sigma_W} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right)^2 \cdot \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} \right\} \right]. \quad (10)$$

Here, $\mu_W = \text{mean}[\log(T^*(\beta_W))]$ and $\sigma_W^2 = \text{var}[\log(T^*(\beta_W))]$.

Let $\Delta\beta$ be the bias that $\beta_L = \beta_W - \Delta\beta$. Form equation (10), equation (11) holds,

$$\begin{aligned} E[g_L(\beta_W - \Delta\beta)] - E[g_L(\beta_W)] &= E[g_L(\beta_L)] - E[g_L(\beta_W)] \\ &\approx 0 + \frac{1}{2\sigma_W} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right)^2 \cdot \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} \right\} \right]. \end{aligned} \quad (11)$$

On the other hand, equation (12) holds,

$$\begin{aligned} &E[g_L(\beta_W - \Delta\beta)] - E[g_L(\beta_W)] \\ &= E \left[\sum_{i=1}^n \left(\mathbf{z}(t_{i,B_i}^o) \right) - \sum_{i=1}^n \frac{T_i^{**}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right] - \frac{1}{\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W - \Delta\beta)) - \mu_L)}{\sigma_L} \right) \cdot \frac{T_i^{**}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right\} \right] \\ &\quad - E \left[\sum_{i=1}^n \left(\mathbf{z}(t_{i,B_i}^o) \right) - \sum_{i=1}^n \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} \right] + \frac{1}{\sigma_W} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right) \cdot \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} \right\} \right] \\ &= E \left[\sum_{i=1}^n \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} - \sum_{i=1}^n \frac{T_i^{**}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right] + \frac{1}{\sigma_W} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right) \cdot \frac{T_i^{**}(\beta_W)}{T_i^*(\beta_W)} \right\} \right] \\ &\quad - \frac{1}{\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W - \Delta\beta)) - \mu_L)}{\sigma_L} \right) \cdot \frac{T_i^{**}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right\} \right]. \end{aligned} \quad (12)$$

Therefore, from equation (11) and (12), we obtain following equation (13),

$$\begin{aligned} & \frac{1}{2} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right)^2 \cdot \frac{T_i^{*'}(\beta_W)}{T_i^*(\beta_W)} \right\} \right] - E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W)) - \mu_W)}{\sigma_W} \right) \cdot \frac{T_i^{*'}(\beta_W)}{T_i^*(\beta_W)} \right\} \right] \\ & \approx \sigma_W \cdot E \left[\sum_{i=1}^n \frac{T_i^{*'}(\beta_W)}{T_i^*(\beta_W)} - \sum_{i=1}^n \frac{T_i^{*'}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right] - \frac{\sigma_W}{\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W - \Delta\beta)) - \mu_L)}{\sigma_L} \right) \cdot \frac{T_i^{*'}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right\} \right]. \end{aligned} \quad (13)$$

Right side of equation (13) adjusts the value of $\Delta\beta$. Let the term of $E \left[\sum_{i=1}^n \frac{T_i^{*'}(\beta_W)}{T_i^*(\beta_W)} - \sum_{i=1}^n \frac{T_i^{*'}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right]$ be B and the term of $\frac{\sigma_W}{\sigma_L} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta_W - \Delta\beta)) - \mu_L)}{\sigma_L} \right) \cdot \frac{T_i^{*'}(\beta_W - \Delta\beta)}{T_i^*(\beta_W - \Delta\beta)} \right\} \right]$ be C. If the value of σ_W becomes smaller, the proportion of the term B becomes smaller and the proportion of the term C becomes larger. Besides, we confirm the difference value based on $\Delta\beta$ in equation (13). For the term of B, equation (14) holds,

$$E \left[\frac{\partial \left[\sum_{i=1}^n \frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \right]}{\partial \beta} \right] = E \left[\sum_{i=1}^n \left\{ \frac{T_i^{*''}(\beta) \cdot T_i^*(\beta) - T_i^{*'}(\beta) \cdot T_i^{*'}(\beta)}{(T_i^*(\beta))^2} \right\} \right]. \quad (14)$$

Furthermore, under a normal condition that time-varying covariates take an independent value at each time point, that is having no trend, equation (15) holds,

$$\begin{aligned} & E \left[\frac{\partial \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta)) - \mu)}{\sigma} \right) \cdot \frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \right\} \right]}{\partial \beta} \right] \\ & = \frac{1}{\sigma} \cdot E \left[\sum_{i=1}^n \left\{ \frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \cdot \frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \right\} \right] + E \left[\sum_{i=1}^n \left\{ \left(\frac{(\log(T_i^*(\beta)) - \mu)}{\sigma} \right) \cdot \frac{T_i^{*''}(\beta) \cdot T_i^*(\beta) - T_i^{*'}(\beta) \cdot T_i^{*'}(\beta)}{(T_i^*(\beta))^2} \right\} \right] \\ & = \frac{1}{\sigma} \cdot E \left[\sum_{i=1}^n \left\{ \left(\frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \right)^2 \right\} \right] + E \left[\sum_{i=1}^n \left\{ \frac{T_i^{*''}(\beta) \cdot T_i^*(\beta) - T_i^{*'}(\beta) \cdot T_i^{*'}(\beta)}{(T_i^*(\beta))^2} \right\} \right] > E \left[\frac{\partial \left[\sum_{i=1}^n \frac{T_i^{*'}(\beta)}{T_i^*(\beta)} \right]}{\partial \beta} \right]. \end{aligned} \quad (15)$$

It is shown that the difference value of the term C based on $\Delta\beta$ is larger than that of the term B. Therefore, if the value of σ becomes smaller, $\Delta\beta$ becomes smaller due to the term C.

5.3 Strategy under misspecification for baseline distribution

Finally, from the results of numerical simulation and discussion, we frame the strategy under misspecification for a baseline distribution. For this results, we remark that it is assumed that each product is used under a normal condition which implies time-varying covariates have no trend for time. From the results in White (1982), a bias between MLE and QMLE is independent for a sample size n . Additionally, the value of a location parameter η has no influence for the score function (e.g., equation (4), (6)) because of standardization. On the other hand, from the result of section 5.2, it is shown that a scale parameter $\sigma (= 1/m)$ affect the bias of estimation of covariate parameters β under misspecification. From the result in Table 2, when a shape parameter takes $m \geq 4$, it is confirmed that the rate of the bias for the true value ($= 1.00$) is fall under 0.01% in the case of $n = 10000$.

Therefore, when a shape parameter takes the larger value as somewhere $m \geq 4$ ($\sigma \leq 0.25$), it is seemed that the misspecified baseline distribution has a little effect on the estimation of covariate parameters β for the unique case of misspecification that the model by a Weibull distribution is true and log-normal distribution is incorrect. When it is seems that a shape parameter takes a small value, it is needed to accurately confirm the baseline distribution of the cumulative exposure $T^*(\beta^*)$.

6. Conclusion and future work

In the case of misspecification that the true model is a Weibull distribution and the incorrect model is a log-normal distribution, this paper shows that covariate parameter β can be approximately estimated by maximum likelihood estimation assuming a log-normal distribution for the baseline distribution of cumulative exposure. Besides, the simulation study and discussion for a bias of estimation are shown, and this paper clarifies that the bias of the estimation for parameter β is small, if a shape parameter m of a Weibull distribution is large. Thereby, this result makes strategy to use a likelihood function under a log-normal distribution to estimate covariate effect parameters when the truly underlying baseline distribution is either a Weibull distribution or a log-normal distribution.

Future work includes finding ways to improve reliability by using online condition monitoring and our proposed method. We will extend our study to the case of time-varying covariates having trend for time. In addition, we will build similar model for other distributions which would enable online condition monitoring to be more widely applied.

References:

- Bagdonavicius, V. and Nikulin, M. (2001), *Accelerated Life Models: Modeling and Statistical Analysis*, Chapman & Hall/CRC: Boca Raton, FL.
- Bagdonavicius, V., Levulienė, R. and Nikulin, M. (2013), "Chi-squared goodness-of-fit tests for parametric accelerated failure time models", *Communications in Statistics-Theory and Methods*, Vol.42, pp. 2768-2785.
- Hong, Y. and Meeker, W. Q. (2013), "Field-failure predictions based on failure-time data with dynamic covariate information", *Technometrics*, Vol. 55, pp. 135–149.
- Kordonsky, K. B. and Gertsbakh, I. B. (1993), "Choice of the best time scale for system reliability analysis", *European Journal of Operational Research*, Vol. 65, pp. 235–246.
- Meeker, W. Q. and Escobar, L. A. (1998), *Statistical Methods for Reliability Data*, John Wiley: New York.
- Meeker, W. Q. and Hong, Y. (2014), "Reliability meets big data: Opportunities and challenges", *Quality Engineering*, Vol. 26, pp. 102-116.
- Nelson, W. (1990), *Accelerated Testing: Statistical Model, Test Plans, and Data Analyses*, John Wiley: New York.
- Nelson, W. (2001), "Prediction of field reliability of units, each under differing dynamic stresses, from accelerated test data", *Handbook of Statistics*, Vol. 20, pp. 611-621.
- Pascual, F. G. (2005), "Maximum likelihood estimation under misspecified lognormal and Weibull distributions", *Communications in Statistics—Simulation and Computation*, Vol. 34, pp. 503-524.
- White, H. (1982), "Maximum likelihood estimation of misspecified models", *Econometrica*, Vol. 50, pp. 1-25.

Acknowledgements:

This work was supported by a grant from Japan Ministry of Education, Culture, Sports, Science and Technology (No. 22241038).

Author's biographical notes

Masahiro Yokoyama is an assistant professor at the Department of Industrial and Systems Engineering at Chuo University, Japan. He received the doctor's degree (Ph.D.) from University of Electro-Communications in 2014. His research interests include statistics, reliability engineering and statistical quality control.

Watalu Yamamoto is a lecturer at the Department of Informatics at the University of Electro-Communications, Japan. His research interests include industrial statistics for statistical quality control and reliability engineering and also applied statistics.

Kazuyuki Suzuki is a professor at the Department of Informatics at the University of Electro-Communications, Japan. He has been working on Reliability Engineering and Quality Management for over three decades. He has served as both the president of the Japanese Society for Quality Control (JSQC) in 2010 and 2011,

and the president of Reliability Engineering Association of Japan (REAJ) in 2012 and 2013.

[DOI : 10.17929/tqs.1.89]

Received, May 20, 2014

Revised: November 30, 2014

Accepted: March 31, 2015