Improved Likelihood Inferences for Weibull Regression Model

Yan Shen[a,*] and Zhenlin Yang[b]

[a] *Department of Statistics, School of Economics, Xiamen University, P.R.China*

[b] *School of Economics, Singapore Management University, Singapore*

Email: sheny@xmu.edu.cn, zlyang@smu.edu.sg

June 12, 2016

Abstract

A general procedure is developed for bias-correcting the maximum likelihood estimators (MLEs) of the parameters of Weibull regression model with either complete or right-censored data. Following the bias correction, variance corrections and hence improved t-ratios for model parameters are presented. Potentially improved t-ratios for other reliability-related quantities are also discussed. Simulation results show that the proposed method is effective in correcting the bias of the MLEs, and the resulted t-ratios generally improve over the regular t-ratios.

Key Words: Bias correction; Variance correction; Bootstrap; Improved t-ratios; Stochastic expansion; Right censoring.

1 Introduction

To improve the efficiency of lifetime analysis, people usually incorporate auxiliary information (e.g., group indicators, individual characteristics, or environmental factors, etc.) into lifetime models and treat them as covariates. Parametric regression is such a model that generalizes a parametric probability distribution by treating some or all parameters as functions of covariates [1, 4, 7]. In this paper, we consider a Weibull regression model that is developed based on a Weibull distribution by allowing the scale parameter to depend on covariates. This regression model has a wide application in accelerated life test, and plays an important role as a type of accelerated life model [7]. By further allowing the shape parameter as a function of covariates, a more general Weibull regression model can be obtained.

[*]Corresponding author: South Siming Road, No. 422, Xiamen, Fujian, China. 361005.

A lifetime random variable T is said to follow the Weibull distribution, denoted by $\mathrm{WB}(\alpha, \beta)$, if its probability density function (pdf) has the form,

$$f(t) = \alpha^{-\beta}\beta t^{\beta-1}\exp[-(t/\alpha)^{\beta}], \quad t \geq 0, \tag{1}$$

where $\alpha > 0$ is the scale parameter and $\beta > 0$ is the shape parameter. The survival function (SF) is $S(t) = \exp[-(t/\alpha)^{\beta}]$ and the hazard function (HF) is $\lambda(t) = \alpha^{-\beta}\beta t^{\beta-1}$. By allowing the scale parameter or the shape parameter or both to depend on a $p\times 1$ vector of explanatory variables (or covariates) X, the Weibull distribution is generalized to a Weibull regression model. For example, if $\alpha = \alpha(X)$, then we have a Weibull regression model where the covariates affect the Weibull life through its scale parameter. There is an issue on the choice of the functional form of $\alpha(X)$. The most natural choice may be $\alpha(X) = \exp(a'X)$ as in such a setting, $\alpha(X) > 0$ is guaranteed without restrictions on a and X. In this case, the vector a is referred to as the regression coefficients as in the regular linear regression models. Also, this choice leads a Weibull regression model that can be interpreted as both the *proportional hazards model* and the *accelerated failure time model* (Kalbfleisch and Prentice [7], Cox and Oakes [2], Lawless [9]), see next section for details. Note that this model can be further extended by allowing the shape parameter β to depend on the covariates as well. To ease the exposition, we focus on the former in this paper although all the methods can be extended to the more general model without much technical difficulty.

For estimating the common shape parameter β and the vector of coefficients a, the maximum likelihood estimation (MLE) method remains the popular method. However, similar to the case of a Weibull distribution, the MLEs for the Weibull regression model, especially the MLE of β, can be rather biased, in particular when the sample size is small or data are heavily censored. Undoubtedly, the biased parameter estimates would affect the subsequent statistical inferences, such as constructing a confidence interval for β, estimating a future percentile life given certain covariates values, and predicting a future lifetime at different covariates, etc. It may also affect further experimental design that consists of the determination of values of sample size, censoring times and covariate values [9]. Moreover, since there is a regression part in the Weibull regression model, significance tests would be expected to be carried out on the scale-related parameters for the purpose of model refinement and variable selection. To improve the accuracy of the above mentioned statistical inference problems, one may first consider to correct the bias of the MLEs of the Weibull regression model, and then correct the bias of the variance estimates of the MLEs and further develop improved t-ratios based on these bias-corrected estimates.

In the Weibull literature, several approaches were proposed to deal with the bias problem for the shape parameter of the Weibull distribution, such as the bias-expanding method by Hirose [5], the modified MLE by Yang and Xie [18] and the stochastic expansion method by Shen and Yang [13]. However, the bias problem for the MLEs of the Weibull regression model and other parametric regression models was rarely considered. A possible reason is that, unlike the explicitly expressed MLE for the Weibull distribution's parameter, the MLEs for the Weibull regression parameters can only be obtained numerically, which increases the difficulty in correcting the bias. Furthermore, the variances of corrected MLEs need to be corrected, which was also seldom touched upon in the early works.

In this paper, we attempt to solve the bias-correction and the variance-estimation problems for the Weibull regression model by extending the univariate method in Shen and Yang [13] to a multivariate situation so that the MLEs of all parameters in the model can be bias-corrected simultaneously, and the variances for the corrected MLEs as well as some improved inferences can be obtained. The proposed correction method is developed based on a multivariate third-order stochastic expansion for the MLE [12] and a nonparametric bootstrap procedure for estimating various expectations involved in the expansion [22]. The advantages of the proposed method in this work are that, (i) it requires only the estimating function that is used to generate estimators, i.e. the score function for the MLEs, (ii) it can deal with multivariate models and parameter vectors, and (iii) it can be easily applied to other models. The simulation results show that the new multivariate method is general and effective in correcting the bias of the MLEs regardless of sample size and data type, i.e. complete or censored. Based on the corrected MLEs, the variance estimate can also be corrected and the resulted inference methods (t-ratios) show improved performances.

Our paper is organized as follows. Section 2 describes the Weibull regression model and the maximum likelihood estimation. Section 3 describes the general bias correction methodology, and presents details for the Weibull regression model. Section 4 discusses subsequent model inferences and presents some improved statistics. Section 5 presents Monte Carlo results.

Section 6 presents an illustrative example, and Section 7 concludes the paper.

2 The Model and Maximum Likelihood Estimation

2.1 Weibull regression model

Let $T_1, \ldots, T_n$ be life (failure) times of n patients (items) in a medical (reliability) study. Let $X_1, \ldots, X_n$ be the corresponding values of the $p \times 1$ vector of covariates. The *accelerated life (failure time) model* (see, e.g., [2] and [7]) is to related the logarithms of life (failure) times to their covariates through a loglinear regression equation

$$\log T_i = a' X_i + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2}$$

where a is a $p \times 1$ vector of parameters and ε_i are random errors, independent and identically distributed (i.i.d.) with a specified cumulative distribution function (CDF) G. Exponentiation gives $T_i = \exp(a' X_i)\xi_i$, where $\xi_i = \exp(\varepsilon_i)$ has a hazard function $\lambda_0(\cdot)$. It follows that the hazard function for T_i can be written in terms of this baseline hazard $\lambda_0(\cdot)$ as

$$\lambda(t_i|X_i) = \exp(-a' X_i)\lambda_0(e^{-a' X_i} t_i).$$

This hazard function shows that the accelerated life model specifies the multiplicative effect of the covariates X on time t, and in fact brings in the covariates to alter time scale. This explains why the above model is called ‘accelerated life model’, as in an accelerated life test, a testing process will subject products to severer conditions in an effort to increase the failure rate and to uncover faults in a shorter period of time [7, 17, 20].

It is well known that if T has a Weibull distribution with scale parameter α and shape parameter β, then $\log T$ has a type-I extreme value (EV-I) distribution, or Gumbel distribution, with location parameter $\log \alpha$ and scale parameter $1/\beta$. Thus, for the Weibull regression model with $\alpha(X) = \exp(a' X)$ and constant β, if we let $Y_i = \log T_i$, then Y_i can be written as

$$Y_i = a' X_i + Z_i/\beta, \tag{3}$$

where $\{Z_i\}$ are i.i.d. errors subject to a standard EV-I distribution with location parameter

0 and scale parameter 1. This shows that the Weibull regression model we consider is an accelerated life model as the covariates act additively on Y_i or multiplicatively on T_i. Some basic properties of the standard EV-I distribution are useful for the latter developments. First, the mean, variance, pdf and SF of Z_i are, respectively, $e_0 = -0.5772$, $r_0 = 1.6449$,

$$f_0(z) = \exp(z - e^z), \quad \text{and} \quad S_0(z) = \exp(-e^z), \quad -\infty < z < \infty. \tag{4}$$

As the EV-I distribution is closed under location and scale transformations, Y also follows an EV-I distribution with the SF, $S(y|X) = \exp\{-\exp[\beta(y - a'X)]\}, \quad -\infty < y < \infty$.

Now, the hazard function for T_i, given the covariates X_i such that $\alpha(X_i) = \exp(a'X_i)$, is

$$\lambda(t_i|X_i) = \exp[-\beta(a'X_i)]\beta t_i^{\beta-1} = \lambda_0(t_i)\exp(-\beta a'X_i),$$

where $\lambda_0(t) = \beta t^{\beta-1}$ is the hazard function for the Weibull distribution with $\alpha = 1$. From this, one sees that the Weibull regression model is also a proportional hazards model as the covariates act multiplicatively on the hazard function. It is easy to see that $\lambda(t_i|X_i) = \beta e^{-a'X_i}(e^{-a'X_i}t_i)^{\beta-1} = e^{-a'X_i}\lambda_0(e^{-a'X_i}t_i)$, showing again that the Weibull regression model is a special accelerated life model. The Weibull regression model discussed above is the only accelerated life model that is also a Cox's proportional hazards model [7, 9]. When the CDF G in Model (2) is known, like the Weibull model, the usual maximum likelihood method can be applied for statistical inference. When no parametric assumptions are imposed on G, the rank estimation method (see, e.g., [23]) is often applied.

2.2 Maximum likelihood estimation

In this paper, we propose to estimate the model (3) by the maximum likelihood method based on either complete and censored data. In practice, we may only observe $S_i = T_i \wedge C_i$, $\delta_i = I_{\{T_i<C_i\}}$ and X_i, $i = 1, \ldots, n$, where C_i $(i = 1, \ldots, n)$ are the censoring times. In such a case, we say that the observed lifetimes are right censored. Assume that we have a censored random sample (s_i, δ_i, x_i), $i = 1, \cdots, n$, from a population subject to the Weibull regression model with parameters β and $\alpha(x_i) = \exp(a'x_i)$, where s_i $(i = 1, 2, \ldots, n)$ are the observed lifetimes or censoring times of n randomly selected 'items', δ_i $(i = 1, 2, \ldots, n)$ are

the failure indicators with $\delta_i = 1$ for the actual lifetime and $\delta_i = 0$ for the censoring time, and x_i $(i = 1, 2, \ldots, n)$ are the $p \times 1$ covariates for the ith item. Denote by $r = \sum_{i=1}^n \delta_i$ the total number of observed lifetimes.

Let $\theta = (a', \beta)' = (a_1, a_2, \cdots, a_p, \beta)'$. Clearly, the first element of x_i is one so that when there are no covariate effect, i.e., $a_2 = \cdots = a_p = 0$, the model reduces to a single Weibull distribution. The Weibull loglikelihood function of θ, based on the observed values (s_i, δ_i, x_i), $i = 1, \cdots, n$, is thus

$$\ell_n(\theta) = r \log \beta + \sum_{i=1}^{n} (\delta_i z_i(\theta) - e^{z_i(\theta)}) - \sum_{i=1}^{n} \delta_i \log s_i, \tag{5}$$

where $z_i(\theta) = \beta(\log s_i - a' x_i)$ and $e^{z_i(\theta)} = [s_i \exp(-a' x_i)]^{\beta}$. Maximizing $\ell_n(\theta)$ gives the MLEs $\hat{a}_n$ for a and $\hat{\beta}_n$ for β, and thus $\hat{\theta}_n = (\hat{a}_n', \hat{\beta}_n)'$. Equivalently, $\hat{\theta}_n$ can be obtained by solving the score equation $\frac{\partial}{\partial \theta} \ell_n(\theta) = 0$, where the score function

$$\frac{\partial \ell_n(\theta)}{\partial \theta} = \begin{pmatrix} \frac{\partial}{\partial a} \ell_n(\theta) \\ \frac{\partial}{\partial \beta} \ell_n(\theta) \end{pmatrix} = \begin{pmatrix} \beta \sum_{i=1}^n (e^{z_i(\theta)} - \delta_i) x_i \\ \frac{r}{\beta} - \frac{1}{\beta} \sum_{i=1}^n (e^{z_i(\theta)} - \delta_i) z_i(\theta) \end{pmatrix}. \tag{6}$$

The consistency and asymptotic normality of the MLE $\hat{\theta}_n$ can be established based on the following regularity conditions.

Assumption 1. *The true value θ_0 of θ is an interior point of an open subset of the real $(p+1)$-dimensional space Θ.*

Assumption 2. *The distribution of the covariates X is not concentrated on a $(p-1)$-dimensional affine subspace of $\mathbb{R}^p$.*

Assumption 3. *The (expected) number of observed failures times ($\mathrm{E}(r)$ or r) approaches ∞ at rate n as $n \to \infty$.*

Assumption 2 guarantees the full rank of the covariate matrix $\mathbf{X} = (X_1, \ldots, X_n)'$ and hence the uniqueness of the MLE $\hat{\theta}_n$. Assumption 3 says that, when the data are censored, the amount of available information needs to grow at the same rate as n when $n \to \infty$, see (Lawless [9], Sec. 2.2.3) for an interesting intuitive discussion. Under Assumptions 1-3, we have $\hat{\theta}_n \xrightarrow{p} \theta_0$, and

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\big[0, \ \lim n I^{-1}(\theta_0)\big],$$

where $I(\theta_0) = -\mathrm{E}[\frac{\partial^2}{\partial\theta\partial\theta'}\ell(\theta)]|_{\theta=\theta_0}$ is the Fisher information matrix, which can easily be shown to exist and to be positive definite for the Weibull regression model. These large sample results can be proved by following Example 5.43, Theorem 5.41 and Theorem 5.42 in [15]. Alternatively, they can be proved using the counting process and martingale theory outlined in (Lawless [9], Appendix F). As we are concerning more on the finite sample properties of $\hat{\theta}_n$, detailed proofs of the asymptotic properties of of $\hat{\theta}_n$ are not provided. Inference for θ can be carried out based on the above large sample results, with $I(\theta)$ being estimated by the observed information matrix $J_n(\hat{\theta}_n) = -\frac{\partial^2}{\partial\theta\partial\theta'}\ell_n(\theta)|_{\theta=\hat{\theta}_n}$.

When sample size is small or the censorship is heavy, the MLEs $\hat{a}_n$ and $\hat{\beta}_n$, in particular the latter, can be rather biased, which will likely have serious impacts on the subsequent inferences. Therefore, it is highly desirable to bias-correct the MLEs and their standard error estimates so that inferences concerning a_j's and β can be made more reliable. Moreover, for an accelerated life study, the main purpose is to estimate certain percentile life at a designed operating condition (for more inference issues, see e.g, Nelson [10]). In summary, bias corrections on the MLEs and their variances can potentially improve inferences for all the reliability-related quantities.

3 Bias Correction and Variance Estimation on the MLEs

3.1 Stochastic expansion of the MLEs

Rilstone et al. [12] considered a class of $\sqrt{n}$-consistent estimators identified by estimating equation: $\hat{\theta}_n = \arg\{\psi_n(\theta) = 0\}$, where $\psi_n(\theta)$ is a vector-valued function of the same dimension as θ and normalized to have order $O_p(n^{-1/2})$, and obtained a third-order stochastic expansion for $\hat{\theta}_n$, assuming $\mathrm{E}[\psi_n(\theta_0)] = 0$, where θ_0 is the true value of the parameter vector θ.

In our case, $\psi_n(\theta)$ is a $(p+1)\times 1$ joint estimating function obtained by dividing the score function given in (6) by n, i.e.,

$$\psi_n(\theta) = \begin{cases} \frac{\beta}{n}\sum_{i=1}^n (e^{z_i(\theta)} - \delta_i)x_i, \\ \frac{r}{n\beta} - \frac{1}{n\beta}\sum_{i=1}^n (e^{z_i(\theta)} - \delta_i)z_i(\theta), \end{cases} \tag{7}$$

and $\hat{\theta}_n$ is the MLE. Let $H_{kn}(\theta)$ be the kth-order partial derivative of $\psi_n(\theta)$ with respect to

θ', $k = 1, 2, 3$, obtained sequentially and elementwise. Denote $\psi_n \equiv \psi_n(\theta_0)$, $H_{kn} \equiv H_{kn}(\theta_0)$, $H^\circ_{kn} = H_{kn} - \mathrm{E}(H_{kn})$, $k = 1, 2, 3$, and $\Omega_n = -[\mathrm{E}(H_{1n})]^{-1}$. Under some general smoothness conditions on $\psi_n(\theta)$ (see Rilstone et al. [12]), $\hat{\theta}_n$ possesses the following third-order stochastic expansion at θ_0:

$$\hat{\theta}_n - \theta_0 = a_{-1/2} + a_{-1} + a_{-3/2} + O_p(n^{-2}), \tag{8}$$

where $a_{-1/2} = \Omega_n \psi_n$, $a_{-1} = \Omega_n H^\circ_{1n} a_{-1/2} + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2} \otimes a_{-1/2})$, and $a_{-3/2} = \Omega_n H^\circ_{1n} a_{-1} + \frac{1}{2}\Omega_n H^\circ_{2n}(a_{-1/2} \otimes a_{-1/2}) + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(a_{-1/2} \otimes a_{-1} a_{-1} \otimes a_{-1/2}) + \frac{1}{6}\Omega_n \mathrm{E}(H_{3n})(a_{-1/2} \otimes a_{-1/2} \otimes a_{-1/2})$, representing terms of order $O_p(n^{-s/2})$, $s = 1, 2, 3$, respectively, E denotes the expectation corresponding to θ_0, and $\otimes$ denotes the Kronecker product.

Proof of this third-order expansion in the context of Weibull regression model can easily be done by verifying the conditions of Rilstone et al. [12]. However, we are mostly interested in the finite sample bias of $\hat{\theta}_n$, the detail of this proof is omitted.

As θ is a $(p+1) \times 1$ vector, $H_{1n}(\theta)$, $H_{2n}(\theta)$, $H_{3n}(\theta)$ are matrices of dimensions $(p+1) \times (p+1)$, $(p+1) \times (p+1)^2$, $(p+1) \times (p+1)^3$, respectively. The detailed expressions of $H_{kn}(\theta), k = 1, 2, 3$ are given in Appendix A. It is interesting to note that the elements of $\psi_n(\theta_0)$ and $H_{kn}(\theta_0)$ ($k = 1, 2, 3$) are all functions of only the shape parameter β_0 and the random terms $z_i \equiv z_i(\theta_0)$, which is a realization of the standard EV-I random variable when $\delta_i = 1$, or a censored observation on $\beta_0(\log C_i - a_0' x_i)$, $i = 1, \ldots, n$, when $\delta_i = 0$. Therefore, it is expected that most bias would come from the estimation of the shape parameter β.

Assumption 4. *For θ in a neighborhood of θ_0, $\frac{1}{n}\left|\sum_{i=1}^n (\beta e^{z_i(\theta)} - \beta_0 e^{z_i})\right| = \|\theta - \theta_0\| B_{n,1}$, $\frac{1}{n}\left|\sum_{i=1}^n (\beta e^{z_i(\theta)} z_i(\theta) - \beta_0 e^{z_i} z_i)\right| = \|\theta - \theta_0\| B_{n,2}$, where $\mathrm{E}|B_{n,1}| < c_1$ and $\mathrm{E}|B_{n,2}| < c_2$, for some finite constants c_1 and c_2.*

Theorem 1 *Under Assumptions 1-4, we have the 2nd-order ($O(n^{-1})$) bias and the 3rd-order ($O(n^{-3/2})$) bias for the MLEs $\hat{\theta}_n$ of the model parameters θ_0:*

$$b_2(\theta_0) = \Omega_n \mathrm{E}(H_{1n}\Omega_n \psi_n) + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n}) \mathrm{E}[(\Omega_n \psi_n) \otimes (\Omega_n \psi_n)], \tag{9}$$

$$\begin{aligned} b_3(\theta_0) = {} & \Omega_n \mathrm{E}(H^\circ_{1n} a_{-1}) + \frac{1}{2}\Omega_n \mathrm{E}[H^\circ_{2n}(a_{-1/2} \otimes a_{-1/2})] + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n}) \mathrm{E}(a_{-1/2} \otimes a_{-1} \\ & + a_{-1} \otimes a_{-1/2}) + \frac{1}{6}\Omega_n \mathrm{E}(H_{3n}) \mathrm{E}(a_{-1/2} \otimes a_{-1/2} \otimes a_{-1/2}), \end{aligned} \tag{10}$$

where $\psi_n \equiv \psi_n(\theta_0)$, $H_{kn} \equiv H_{kn}(\theta_0)$, $H_{kn}^{\circ} \equiv H_{kn} - \mathrm{E}(H_{kn})$, $k = 1, 2, 3$, *and* $\Omega_n = -[\mathrm{E}(H_{1n})]^{-1}$.

Note that the second-order bias $b_2 \equiv b_2(\theta_0) = \mathrm{E}(a_{-1})$ noting that $\mathrm{E}(a_{-1/2}) = 0$, and the third-order bias $b_3 \equiv b_3(\theta_0) = \mathrm{E}(a_{-3/2})$. If the estimates of b_2 and b_3 are available, denoted by $\hat{b}_2$ and $\hat{b}_3$, then the second- or third-order bias-corrected MLEs of θ can be obtained by

$$\hat{\theta}_n^{\mathrm{bc2}} = \hat{\theta}_n - \hat{b}_2 \quad \text{and} \quad \hat{\theta}_n^{\mathrm{bc3}} = \hat{\theta}_n - \hat{b}_2 - \hat{b}_3. \tag{11}$$

It will be shown in next section that under some mild conditions, the extra variability introduced by the estimation of the bias is not higher than the remainder.

Naturally, use of the bias corrected MLEs $\hat{\theta}_n^{\mathrm{bc2}}$ or $\hat{\theta}_n^{\mathrm{bc3}}$, and the observed information matrix evaluated at $\hat{\theta}_n^{\mathrm{bc2}}$ or $\hat{\theta}_n^{\mathrm{bc3}}$ should lead improved inferences for θ. Further improvements are possible by using the bias-corrected variance estimate as well. Based on the stochastic expansion (8), a 2nd-order variance expansion of $\hat{\theta}_n$ can be directly obtained,

$$\begin{aligned} \mathrm{V}_2(\hat{\theta}_n) &= \mathrm{Var}(a_{-1/2} + a_{-1}) + O(n^{-2}) \\ &= \mathrm{E}[(a_{-1/2} + a_{-1})(a_{-1/2} + a_{-1})'] + O(n^{-2}), \end{aligned} \tag{12}$$

noting $\mathrm{E}(a_{-1/2}) = 0$, $\mathrm{E}(a_{-1}) = O(n^{-1})$, and $\mathrm{E}(a_{-1/2} + a_{-1})\mathrm{E}(a_{-1/2} + a_{-1})' = O(n^{-2})$. Furthermore, it is easy to see that $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}}) = \mathrm{V}_2(\hat{\theta}_n) + O(n^{-2})$. Thus, further improved inferences for θ can be expected, by using $\hat{\theta}_n^{\mathrm{bc2}}$ in connection with $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}})$. Third-order variance correction can also be carried out by extending the above result. We will concentrate on the second-order results. Monte Carlo results presented in Section 5 show that the second corrections are sufficient for most of the practical situations.

Question remains on the estimation of b_2 and $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}})$. The explicit expressions of these quantities are difficult if not impossible to obtain, and hence the standard plug-in method can not be applied. Alternative methods are thus desired.

3.2 Bootstrap estimates of bias correction and variance

Different from the Weibull distribution, Weibull regression model deals with not only the lifetime data and censoring mechanism, but also the covariates, whose distributions are unknown. Therefore generating bootstrap samples in a censored Weibull regression model can be much tricker. Shen and Yang [13] introduced a parametric bootstrap method to bias-correct the MLE of the common shape parameter of several Weibull populations, based on complete or censored data. Although Weibull regression model is also a parametric model, we consider to adopt a nonparametric bootstrap method here to estimate the bias and variance corrections, concerning the involvement of covariates in model. For the case of complete data, the bootstrap samples are drawn only on the estimated errors, whereas for the censoring case, the bootstrap samples are drawn on the triples: estimated errors, censoring indicators, and covariate values.

Note that the key quantities $\psi_n(\theta_0)$ and $H_{kn}(\theta_0)$ can be written as $\psi_n(\theta_0) = \psi_n(\beta_0, \mathbf{z}_n)$ and $H_{kn}(\theta_0) = H_{kn}(\beta_0, \mathbf{z}_n)$, $k = 1, 2, 3$, where $\mathbf{z}_n = (z_1, \ldots, z_n)'$, and $z_i = z_i(\theta_0) = \beta_0(\log s_i - a_0' x_i)$. When sample data are complete, the error terms z_i $(i = 1, \ldots, n)$ are subject to the standard EV-I distribution. Then the estimates of z_i $(i = 1, \ldots, n)$, which are called ML residuals in the MLE framework, can be resampled by a regular nonparametric bootstrap method and used for estimating the desired quantity expectations. Following the bootstrap steps in Yang [22], the nonparametric bootstrap procedure can be carried out in this way:

(1) Compute the MLEs $\hat{\theta}_n = (\hat{a}_n', \hat{\beta}_n)'$ based on the original data;

(2) Compute ML residuals $\hat{z}_i = z_i(\hat{\theta}_n) = \hat{\beta}_n(\log s_i - \hat{a}_n' x_i)$, $i = 1, \ldots, n$;

(3) Resample $\{\hat{z}_i, \cdots, \hat{z}_n\}$ in a usual way, and denote the resampled vector by $\hat{\mathbf{z}}^*_{n,b}$;

(4) Compute $\hat{\psi}_{n,b} = \psi_{n,b}(\hat{\beta}_n, \hat{\mathbf{z}}^*_{n,b})$, and $\hat{H}_{kn,b} = H_{kn,b}(\hat{\beta}_n, \hat{\mathbf{z}}^*_{n,b})$, $k = 1, 2, 3$, with the original covariate matrix $\mathbf{X} = (X_1, \ldots, X_n)'$ unchanged;

(5) Repeat the steps (3)-(4) B times to get a sequences of bootstrapped values $\{\hat{\psi}_{n,b}, b = 1, \ldots, B\}$ for ψ_n, and $\{\hat{H}_{kn,b}, b = 1, \ldots, B\}$ for $H_{kn}, k = 1, 2, 3$.

When the sample data are right censored, Step (3) and (4) in the above procedure should be changed to Step (3') and (4') as follows,

(3') Resample $\{(\hat{z}_i, x_i, \delta_i), \cdots, (\hat{z}_n, x_n, \delta_n)\}$ in a usual way, and denote the resampled vectors by $\hat{\mathbf{z}}^*_{n,b}$, $\hat{\mathbf{x}}^*_{n,b}$ and $\hat{\boldsymbol{\delta}}^*_{n,b}$ respectively;

(4') Compute $\hat{\psi}_{n,b} = \psi_{n,b}(\hat{\beta}_n, \hat{\mathbf{z}}^*_{n,b}, \hat{\mathbf{x}}^*_{n,b}, \hat{\boldsymbol{\delta}}^*_{n,b})$, and $\hat{H}_{kn,b} = H_{kn,b}(\hat{\beta}_n, \hat{\mathbf{z}}^*_{n,b}, \hat{\mathbf{x}}^*_{n,b}, \hat{\boldsymbol{\delta}}^*_{n,b})$, $k = 1, 2, 3$;

The reason for involving covariates into bootstrapping is that, when $\delta_i = 0$, the censored error observations $z_i = \beta_0(\log C_i - a_0' x_i)$ $(i = 1, \ldots, n)$ rely on the covariates x_i's, although the censoring distribution for C_i's is independent of the covariates and the failure times.

The bootstrap estimates of various expectations in $b_2(\theta_0)$ and $b_3(\theta_0)$ thus are simply the averages of the corresponding B bootstrap values. For example, the bootstrap estimates of $\mathrm{E}(\psi_n \otimes \psi_n)$ and $\mathrm{E}(H_{1n})$ are, respectively,

$$\hat{\mathrm{E}}(\psi_n \otimes \psi_n) = \tfrac{1}{B} \textstyle\sum_{b=1}^{B} \hat{\psi}_{n,b} \otimes \hat{\psi}_{n,b} \ \text{ and } \ \hat{\mathrm{E}}(H_{1n}) = \tfrac{1}{B} \textstyle\sum_{b=1}^{B} \hat{H}_{1n,b}.$$

The latter gives $\hat{\Omega}_n = -[\hat{\mathrm{E}}(H_{1n})]^{-1}$, which leads to the bootstrap estimates of quantities that contain Ω_n, e.g. $\mathrm{E}(H_{1n}\Omega_n\psi_n)$, by repeating the bootstrapping procedure based on the same set of bootstrap data $\hat{\mathbf{z}}_{n,b}^*$ or $(\hat{\mathbf{z}}_{n,b}^*, \hat{\mathbf{x}}_{n,b}^*, \hat{\boldsymbol{\delta}}_{n,b}^*)$, $b = 1, \ldots, B$, obtained in Steps (3) and (4) or (3') and (4') above, i.e.,

$$\hat{\mathrm{E}}(H_{1n}\Omega_n\psi_n) = \tfrac{1}{B} \textstyle\sum_{b=1}^{B} \hat{H}_{1n,b}\hat{\Omega}_n\hat{\psi}_{n,b}.$$

This is a 'two-stage' bootstrap procedure. After getting the estimates of all those expectations, we can calculate $\hat{b}_2$ and $\hat{b}_3$, and thus $\hat{\theta}_n^{\mathrm{bc2}}$ and $\hat{\theta}_n^{\mathrm{bc3}}$.

Note that the nonstochastic matrices such as Ω_n, $\mathrm{E}(H_{1n})$ and $\mathrm{E}(H_{2n})$ are involved in the expectation operator. Pulling these nonstochastic matrices outside the expectation sign could simplify the evaluation of the expectations. Using the properties of Kronecker product $(A \otimes B)(C \otimes D) = AC \otimes BD$ and $\mathrm{vec}(ACB) = (B' \otimes A)\mathrm{vec}(C)$, where 'vec' vectorizes a matrix by stacking its columns (see, e.g., Horn and Johnson [6]), b_2 becomes

$$b_2 = \Omega_n \mathrm{E}(\psi_n' \otimes H_{1n})\mathrm{vec}(\Omega_n) + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(\Omega_n \otimes \Omega_n)\mathrm{E}(\psi_n \otimes \psi_n).$$

Therefore, the bootstrap estimate of b_2 can be realized in 'one-stage', instead of two-stage described above. The same idea may apply to get the bootstrap estimate for the third-order bias b_3, but the expression becomes messy, in particular when the variance correction is involved. We thus recommend the two-stage procedure as the added computation is not at all an issue of concern due to the fact that the introduced bootstrap procedure does not involve 're-estimation' of the model parameters.

To estimate the 2nd-order variance $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}})$ in (12), an additional bootstrap procedure can be carried out after Step (5) but with bootstrap parameters $\hat{\theta}_n^{\mathrm{bc2}}$:

(6) Use the same resampled vector(s) $\hat{\mathbf{z}}_{n,b}^*$ *for complete data or* $(\hat{\mathbf{z}}_{n,b}^*, \hat{\mathbf{x}}_{n,b}^*, \hat{\boldsymbol{\delta}}_{n,b}^*)$ *for censored data;*

(7) Compute the quantities with $\hat{\theta}_n^{\mathrm{bc2}}$*, i.e.* $\hat{\psi}_{n,b} = \psi_{n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*)$*,* $\hat{H}_{1n,b} = H_{1n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*)$*, and* $\hat{H}_{2n,b} = H_{2n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*)$ *for complete data, or* $\hat{\psi}_{n,b} = \psi_{n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*, \hat{\mathbf{x}}_{n,b}^*, \hat{\boldsymbol{\delta}}_{n,b}^*)$*,* $\hat{H}_{1n,b} = H_{1n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*, \hat{\mathbf{x}}_{n,b}^*, \hat{\boldsymbol{\delta}}_{n,b}^*)$*, and* $\hat{H}_{2n,b} = H_{2n,b}(\hat{\beta}_n^{\mathrm{bc2}}, \hat{\mathbf{z}}_{n,b}^*, \hat{\mathbf{x}}_{n,b}^*, \hat{\boldsymbol{\delta}}_{n,b}^*)$ *for censored data,;*

(8) Repeat the steps (6)-(7) B times to get sequences of bootstrapped values $\{\hat{\psi}_{n,b}, b = 1, \ldots, B\}$ *for* ψ_n*, and* $\{\hat{H}_{kn,b}, b = 1, \ldots, B\}$ *for* H_{kn}*,* $k = 1, 2$*.*

After Steps (6)-(8), we can have the estimate of Ω_n and also B values of of $a_{-1/2} + a_{-1} = \Omega_n\psi_n + \Omega_n H_{1n}^{\circ}\Omega_n\psi_n + \frac{1}{2}\Omega_n \mathrm{E}(H_{2n})(\Omega_n\psi_n \otimes \Omega_n\psi_n)$ with $\hat{\theta}_n^{\mathrm{bc2}}$, denoted by $\{a_{-1/2,b} + a_{-1,b}$, $b = 1, \ldots, \mathrm{B}\}$. Then a bootstrap estimate for $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}})$ is

$$\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}}) = \tfrac{1}{B}\textstyle\sum_{b=1}^{B}(a_{-1/2,b} + a_{-1,b})(a_{-1/2,b} + a_{-1,b})'.$$

Remark 1 *In case of complete data, the bootstrap procedure is much simpler, as it depends only on the* $\hat{\beta}_n$ *value, and the resampled samples are from the ML residuals.*

Remark 2 *It is easy to see that* $V_2(\hat{\theta}_n) = V_2(\hat{\theta}_n^{\mathrm{bc2}}) + O(n^{-2})$*. Thus,* $\hat{\theta}_n$ *can be used in estimating the second-order variance of* $\hat{\theta}_n^{\mathrm{bc2}}$ *as well. In this case, the additional steps in the second-order variance estimation are not needed.*

3.3 Validity of the bootstrap method

We now present some results concerning the validity of the bootstrap methods for estimating the bias and the variance of the MLE of θ.

Corollary 1 *Under Assumptions 1-4, if further (i)* $\frac{\partial^r}{\partial\theta_0^r} b_j(\theta_0) \sim b_j(\theta_0), r = 1, 2, j = 2, 3,$ *(ii) a quantity bounded in probability has a finite expectation, then the bootstrap estimates of* *the 2nd- and 3rd-order biases for the MLE $\hat{\theta}_n$ are such that:*

$$\hat{b}_2 = b_2 + O_p(n^{-2}) \;\; and \;\; \hat{b}_3 = b_3 + O_p(n^{-5/2}),$$

where $\sim$ indicates that the two quantities are of the same order of magnitude. It follows that $Bias(\hat{\theta}_n^{\text{bc2}}) = O(n^{-3/2})$ and $Bias(\hat{\theta}_n^{\text{bc3}}) = O(n^{-2})$.

Corollary 2 *Under Assumptions 1-4, if further (i) $\hat{b}_2 - b_2 = O_p(n^{-3/2})$, (ii) a quantity bounded in probability has a finite expectation, then the 2nd-order variances and their the bootstrap estimates are such that:*

$$\begin{aligned} V_2(\hat{\theta}_n^{\text{bc2}}) &= V_2(\hat{\theta}_n) + O(n^{-2}), \\ \hat{V}_2(\hat{\theta}_n^{\text{bc2}}) &= V_2(\hat{\theta}_n^{\text{bc2}}) + O(n^{-2}), \end{aligned}$$

where $\hat{b}_2$ is the estimate of b_2.

The results of Corollary 1 show that using the bootstrap method to estimate the bias terms only (possibly) introduces additional bias of order $O_p(n^{-2})$ or lower. This guarantees the validity of the second-order and the third-order bootstrap bias corrections. Assumption (*ii*) is to ensure $\text{E}[O_p(1)] = O(1)$, $\text{E}[O_p(n^{-2})] = O(n^{-2})$, etc., so that the expectation of a 'stochastic' remainder is of proper order. The proofs of Corollaries 1 and 2 are given in Appendix B.

4 Inferences Following Bias and Variance Corrections

It is well known that inference concerning the covariate effects may be one of the most important types of inference in the context of any regression analysis. In the special case of accelerated life testing model, inference concerning a 'future' percentile life is also of utmost importance. Given the fact that the MLEs of the Weibull regression model can be seriously biased, it is important to study how this bias impacts the subsequent inferences, and how the standard inferences methods can be improved after corrections have been made on the points estimates of the model parameters. In this section, we present some improved inferences for the Weibull regression model following the bias and variance corrections on the MLEs of the model parameters.

4.1 Inferences concerning the covariates effects

Inferences concerning the covariates effects are typically based on the asymptotic t-ratios, constructed based on the results that $\hat{\theta}_n$ is approximately (asymptotically) normal distributed with mean θ_0 and variance $J_n^{-1}(\hat{\theta}_n)$, where $J_n(\hat{\theta}_n)$ is the observed information matrix. Note that $J_n(\theta) = -nH_{1n}(\theta)$. Partitioning $J_n^{-1}(\hat{\theta}_n)$ according to a and β and denoting the partitioned matrix by $J_n^{-1}(\hat{\theta}_n) = (\hat{J}_n^{11}, \hat{J}_n^{12}; \hat{J}_n^{21}, \hat{J}_n^{22})$, then $\hat{J}_n^{11}$ gives an estimate of the asymptotic variance-covariance (VC) matrix of $\hat{a}_n$, and $\hat{J}_n^{22}$ gives an estimate of the asymptotic variance $\hat{\beta}_n$, and $\hat{J}_n^{21}$ gives an estimate of the asymptotic covariance between $\hat{a}_n$ and $\hat{\beta}_n$. Thus, an asymptotic t-statistic for inference for a linear contrast of $c_0' a_0$, a linear contrast of the parameter $a_0 = (a_{10}, \ldots, a_{p0})'$, has the form,

$$t_n = \frac{c_0'\hat{a}_n - c_0' a_0}{\sqrt{c_0' \hat{J}_n^{11} c_0}}. \tag{13}$$

For the 2nd-order bias-corrected MLEs $\hat{\theta}_n^{\text{bc2}}$, there are two available variance estimates: the inverse of the observed information matrix $J_n^{-1}(\hat{\theta}_n^{\text{bc2}})$ and the 2nd-order variance estimate $\widehat{\text{V}}_2(\hat{\theta}_n^{\text{bc2}})$. We normally choose the latter for an obvious reason. Let $\widehat{\text{Var}}(\hat{\theta}_n^{\text{bc2}}) = J_n^{-1}(\hat{\theta}_n^{\text{bc2}})$ or $\widehat{\text{V}}_2(\hat{\theta}_n^{\text{bc2}})$. Denoting the partitioned $\widehat{\text{Var}}(\hat{\theta}_n^{\text{bc2}})$ as $(\hat{V}_{11}, \hat{V}_{12}; \hat{V}_{21}, \hat{V}_{22})$, a 2nd-order corrected t-statistic is thus,

$$t_n^{\text{bc2}} = \frac{c_0'\hat{a}_n^{\text{bc2}} - c_0' a_0}{\sqrt{c_0' \hat{V}_{11} c_0}}. \tag{14}$$

Both t-ratios (13) and (14) can be used for testing the significance of a regression coefficient, or constructing a confidence interval of it. Both refer to the standard normal distribution for the critical values. For example, to test the null hypothesis $H_0 : a_2 = 0$ verses a two-sided alternative, we set $c_0 = (0, 1, \ldots, 0)$ and the reject region is the compliment of the interval $[-u_{\gamma/2}, u_{\gamma/2}]$, where $u_{\gamma/2}$ is the $(\gamma/2)$-quantile of the standard normal distribution. As expected, Monte Carlo results presented in Section 5 show that the second-order corrected t-statistic offers a significant improvement over the large sample t-statistic.

4.2 Inference concerning the shape parameter

Similarly, tests and confidence intervals (CI) for β can be constructed based on two sets of t-ratios. Based on the partitions of $J_n^{-1}(\hat{\theta}_n)$ and $\widehat{\mathrm{Var}}(\hat{\theta}_n^{\mathrm{bc2}})$ discussed above, we have a large sample $100(1-\gamma)\%$ CI for β:

$$\left\{\hat{\beta}_n - u_{\gamma/2}\sqrt{\hat{J}_n^{22}},\ \ \hat{\beta}_n + u_{\gamma/2}\sqrt{\hat{J}_n^{22}}\right\}, \tag{15}$$

and a 2nd-order bias-corrected CI for β:

$$\left\{\hat{\beta}_n^{\mathrm{bc2}} - u_{\gamma/2}\sqrt{\hat{V}_{22}},\ \ \hat{\beta}_n^{bc2} + u_{\gamma/2}\sqrt{\hat{V}_{22}}\right\}. \tag{16}$$

4.3 Confidence intervals for a percentile

As discussed earlier, another interesting problem is to estimate a certain percentile or quantile life under a regular operating condition in an accelerated life test, i.e., to estimate $y_p = a'x_{\mathrm{reg}} + z_p/\beta$ and $T_p = \exp(y_p)$, where z_p is the pth-percentile of the standard EV-I distribution with a predetermined p, and x_{reg} is the values of the covariates corresponding to the regular operating condition. The point estimators for y_p based on the original MLE $\hat{\theta}_n = (\hat{a}_n', \hat{\beta}_n)'$ and the 2nd-order bias-corrected MLE $\hat{\theta}_n^{\mathrm{bc2}} = ((\hat{a}_n^{\mathrm{bc2}})', \hat{\beta}_n^{\mathrm{bc2}})'$ are, respectively,

$$\hat{y}_{n,p} = \hat{a}_n' x_{\mathrm{reg}} + z_p/\hat{\beta}_n \quad \text{and} \quad \hat{y}_{n,p}^{\mathrm{bc2}} = (\hat{a}_n^{\mathrm{bc2}})' x_{\mathrm{reg}} + z_p/\hat{\beta}_n^{\mathrm{bc2}}.$$

Applying the multivariate Delta theorem yields the following large sample results,

$$\begin{aligned}
\hat{y}_{n,p} - y_p &\sim N\left(0,\ c_n' J_n^{-1}(\hat{\theta}_n) c_n\right) \\
\hat{y}_{n,p}^{\mathrm{bc2}} - y_p &\sim N\left(0,\ (c_n^{\mathrm{bc2}})' \widehat{\mathrm{Var}}(\hat{\theta}_n^{\mathrm{bc2}}) c_n^{\mathrm{bc2}}\right)
\end{aligned}$$

where $c_n = (x_{\mathrm{reg}}', -z_p/\hat{\beta}_n^2)'$, $c_n^{\mathrm{bc2}} = (x_{\mathrm{reg}}', -z_p/(\hat{\beta}_n^{\mathrm{bc2}})^2)'$, and $\widehat{\mathrm{Var}}(\hat{\theta}_n^{\mathrm{bc2}}) = J_n^{-1}(\hat{\theta}_n^{\mathrm{bc2}})$ or $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$. The corresponding two $100(1-\gamma)\%$ CIs for y_p are, respectively,

$$\begin{aligned}
\mathrm{CI}_1(y_p) &= \left\{\hat{y}_{n,p} - u_{\gamma/2}\sqrt{c_n' J_n^{-1}(\hat{\theta}_n) c_n},\ \ \hat{y}_{n,p} + u_{\gamma/2}\sqrt{c_n' J_n^{-1}(\hat{\theta}_n) c_n}\right\}, \\
\mathrm{CI}_2(y_p) &= \left\{\hat{y}_{n,p}^{\mathrm{bc2}} - u_{\gamma/2}\sqrt{(c_n^{\mathrm{bc2}})' \widehat{\mathrm{Var}}(\hat{\theta}_n^{\mathrm{bc2}}) c_n^{\mathrm{bc2}}},\ \ \hat{y}_{n,p}^{\mathrm{bc2}} + u_{\gamma/2}\sqrt{(c_n^{\mathrm{bc2}})' \widehat{\mathrm{Var}}(\hat{\theta}_n^{\mathrm{bc2}}) c_n^{\mathrm{bc2}}}\right\}.
\end{aligned}$$

Confidence intervals for other reliability-related quantities, such as the survival function, mean lifetime, failure rate, etc., can be constructed in a similar manner.

5 Monte Carlo Simulations

To investigate the finite sample performances of the proposed method of bias-correcting the MLEs of the Weibull regression parameters and the followed inferences, Monte Carlo simulation experiments are performed along the design in Shen and Yang [13]. Two scenarios are considered, which include (i) complete samples, and (ii) randomly censored samples.

In Monte Carlo experiments, we consider a Weibull regression model with an intercept:

$$\log T = a_1 + a_2 X_2 + a_3 X_3 + Z/\beta.$$

For all the Monte Carlo experiments, $a' = (a_1, a_2, a_3)$ is set at $\{5, 1, 1\}$, β takes values $\{0.5, 0.8, 1, 2, 5\}$, and n takes values $\{20, 50, 100\}$. The two covariates are generated independently, according to $\{x_{i2}\} \overset{iid}{\sim} N(0,1)/\sqrt{2}$ and $\{x_{i3}\} \overset{iid}{\sim} N(0,1)/\sqrt{2}$.

In the entire simulation study, the nonparametric bootstrap is adopted, which (i) fits original data to the above Weibull regression model, (ii) computes the ML residuals, and then (iii) draws random samples with replacement directly from the ML residuals with the size being the same as the original sample size, and also records the corresponding covariates and censoring indicators if the original data are right censored. For all the experiments, 10,000 replications are run in each simulation and the number of bootstrap B is set to be 699.

5.1 Performance of the second-order bias corrected MLEs

Tables 1-2 summarize the empirical mean, root-mean-square-error (rmse) and standard error (se) of the original and 2nd-order bias-corrected MLEs under different combinations of models, data types, and the values of n and β.

Table 1 presents the estimation results of all 4 parameters for the case of complete samples. From the results in the table, we see that the 2nd-order bias-corrected MLE $\hat{\theta}_n^{\text{bc2}}$ is generally nearly unbiased and is much superior to the original MLE $\hat{\theta}_n$ regardless of the values of n. It is also shown that the shape parameter β incurs most bias compared to the scale parameters $a_i, i = 1, 2, 3$, which coincides with our expectation mentioned in Sec. 3.1. Some details are: (i) $\hat{\beta}_n$ always over-estimates the shape parameter, (ii) $\hat{\beta}_n^{\text{bc2}}$ has smaller rmse's and se's compared with those of $\hat{\beta}_n$, (iii) although the improvements of $\hat{a}_{in}^{\text{bc2}}$ over $\hat{a}_{in}$, $i = 1, 2, 3$, are not so significant as that of $\hat{\beta}_n^{\text{bc2}}$, $\hat{a}_{in}^{\text{bc2}}$ is still generally better than $\hat{a}_{in}$ in terms of mean, except some occasional cases for a_3.

We also consider the case of samples with random censoring, which includes Type-I censoring as a special form by treating censoring time fixed. In the random censoring scheme, each item is subject to a different censoring time. For each Monte Carlo replication, two sets of observations $\mathbf{T} = \{T_1, \ldots, T_n\}$ and $\mathbf{C} = \{C_1, \ldots, C_n\}$ are generated, with T_j from a Weibull regression model and C_j from any proper distribution. In this paper, a Uniform distribution $\text{U}(0.5\zeta_{0.9}, 1.5\zeta_{0.9})$, where $\zeta_{0.9} = \exp\{5 + z_{0.9}/\beta\}$ and $z_{0.9}$ is the 90%-percentile of the standard EV-I distribution, is chosen to generate the censoring times C_j's, considering its simple formulation and easy-handling. Then the observed lifetimes $\{S_j = \min(T_j, C_j), j = 1\ldots, n\}$ and the failure indicators $\{\delta_j\}$ are recorded. [1] Based on these original observed lifetimes, the ML residuals can be calculated and bootstrap samples of residuals can be generated by carrying out the nonparametric bootstrap procedure discussed in Sec. 3.2.

The Monte Carlo results are summarized in Table 2. From the results we see that the bias-corrected MLE $\hat{\theta}_n^{\text{bc2}}$ can greatly reduce the bias as well as the variability of $\hat{\theta}_n$ in all combinations under the random censoring mechanism. Moreover, the shape β is shown once more to be the parameter that would incur most bias.

Another important observation is that, different from the results for complete samples, the bias-corrected estimators for the scale-related parameters $\hat{a}_{in}^{\text{bc2}}$ $(i = 1, 2, 3)$ significantly outperform the original MLEs $\hat{a}_{in}$ $(i = 1, 2, 3)$ for randomly censored samples, especially in terms of the reduced bias. Therefore, based on these findings, we may conclude that the proposed method is a desirable choice when dealing with randomly censored data.

[1] With the uniform distributed censoring times, the non-censoring proportions are around 86.22%, 82.28%, 79.88%, 70.13%, 58.60% for $\beta = 0.5,\ 0.8,\ 1.0,\ 2.0,\ 5.0$ respectively.

5.2 Performance of the significance tests

To compare the performances of the two t-ratios t_n and t_n^{bc2}, we reset the covaraite coefficients to $a' = (a_1, a_2, a_3) = (5, 0, 1)$ and test $H_0 : a_2 = 0$. Table 3 and 4 report the simulation results of empirical significance levels of t_n and t_n^{bc2} for complete data and censored data respectively. The variance estimate used in t_n^{bc2} is the inverse of the observed information matrix, $J_n^{-1}(\hat{\theta}_n^{\mathrm{bc2}})$, which was shown leading to a better performance than the 2nd-order variance estimate $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ (experiment results available upon request).

In Tables 3 and 4, we can observe that, (i) the asymptotic test t_n can be very unreliable in the sense that it rejects the true H_0 much too often than it is supposed to. The test t_n^{bc2} offers huge reduction in significance level distortions, with the empirical levels getting close to their nominal levels faster than t_n; (ii) for both data types, the two tests converge in terms of empirical significance level with n increases; (iii) the empirical significance level of t_n is always greater than the nominal level, while t_n^{bc2} does not have obvious pattern. In overall, the improved test t_n^{bc2} outperforms the asymptotic test t_n greatly, regardless of sample size, data type and nominal level. Thus for the purpose of significance test, the test t_n^{bc2} with variance estimate $J_n^{-1}(\hat{\theta}_n^{\mathrm{bc2}})$ is strongly recommended.

5.3 Confidence intervals for the shape parameter

Instead of $J_n^{-1}(\hat{\theta}_n^{\mathrm{bc2}})$ as a variance estimate, the 2nd-order variance estimate $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ is a better choice in constructing confidence interval for β. The simulation results are given in Tables 5 and 6. It is shown that, (i) for complete data, $\mathrm{CI}_2(\beta)$ with $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ is able to provide a more accurate coverage probability than $\mathrm{CI}_1(\beta)$ in almost all situations for the case of small sample size $n = 20$; (ii) for censored data with $n = 20, 1-\gamma = 0.90, 0.95$ and $n = 50, 1-\gamma = 0.90$, $\mathrm{CI}_2(\beta)$ with $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ has the coverage probabilities much closer to the nominal levels and is of shorter length compared to $\mathrm{CI}_1(\beta)$; (iii) for other complete or censored cases, $\mathrm{CI}_2(\beta)$ still has the performances comparable with $\mathrm{CI}_1(\beta)$ in terms of coverage probability but also with even shorter lengths.

Based on the simulation results, we may say that $\mathrm{CI}_2(\beta)$ with $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ is recommended in constructing the confidence interval for β when sample size is small. For median or large-size samples, $\mathrm{CI}_2(\beta)$ may be also preferred as it not only has the coverage probability close to nominal level but also owns the advantage in confidence length in most situations.

5.4 Confidence intervals for percentiles with certain covariates

Monte Carlo simulation experiments are also designed for the confidence intervals for percentile with given covariates, which are set as $x_{\text{reg}} = (1, 0, 0)'$. Thus the percentile we want to estimate is $y_p = 5 + z_p/\beta$ or $T_p = \exp(y_p)$, where the probability $p = 0.5$. The two CIs, $\text{CI}_1(y_p)$ and $\text{CI}_2(y_p)$ are given in Sec. 4.3. Similar to the construction of $\text{CI}_2(\beta)$ for the shape parameter, the 2nd-order variance estimate $\widehat{\text{V}}_2(\hat{\theta}_n^{\text{bc2}})$ is also adopted in constructing $\text{CI}_2(y_p)$.

The results in Tables 7 and 8 show that the confidence interval based on the improved t-ratio, $\text{CI}_2(y_p)$, has a overwhelming superior performance compared to the regular $\text{CI}_1(y_p)$. We can find that the coverage probabilities are greatly improved and get much closer to the nominal levels when using the confidence interval $\text{CI}_2(y_p)$, which has an appropriately enlarged length compared to $\text{CI}_1(y_p)$. The superiority of $\text{CI}_2(y_p)$ is demonstrated in almost all parameter combinations in our experiment, except three cases for censored data with $\beta = 5$ and $y_{0.5} = 4.9267$. This exception may be due to the relatively high censoring proportion for $\beta = 5$, which is 42.41% compared to the less than 30% censoring proportions for other values of β.

Besides the median percentile $y_{0.5}$, two tail percentiles $y_{0.05}$ and $y_{0.95}$ were also taken in to account in our experiment, although the experiment results (available upon request) are not presented here due to the limit of space. Once more, the results substantially support the satisfying performances of $\text{CI}_2(y_p)$ in terms of coverage probability, regardless of sample size, parameter value and type of data. Therefore we may conclude that the improved t-ratio and the resulted confidence interval $\text{CI}_2(y_p)$ should be the choice when concerning the inference for a percentile with certain covariates.

6 An Example

A set of real data from Nelson and Haln [11] or Kalbfleisch and Prentice ([7], p.5) is used to illustrate the application of the proposed bias-corrected method and the subsequent inferences. The data given (replicated in Table 9 for easy reference) describe the number of hours to failure of motorettes operating under various temperatures, and are obtained from an accelerated life test. The test uses temperature as a stress factor to increase the rate of failure so that the exact failure times of at least part of motorettes would be observed during a shorter time period. The interest of such a test is to determine the relationship between failure time and temperature for the purpose of extrapolation to regular operating temperature of 130°C.

From Table 9, we see that (*i*) under each of 4 temperatures, 10 motorettes are subject to test; (*ii*) the data are type-I censored, that is, the failure had not occurred prior to a predestined time at which the test was to be terminated and only censored time was observed; (*iii*) the censoring is server with only 17 of 40 motorettes failing.

Based on the data, we want to model the failure time as a function of operating temperatures of 150°C, 170°C, 190°C, or 220°C. Nelson and Haln [11] adopted a log-normal regression model with the covariates $X_1 = 1$ and $X_2 = 1000/(273.2 + ^\circ\mathrm{C})$. Kalbfleisch and Prentice ([7], p.70) suggested that a Weibll regression model would be preferred, which has the form

$$\log T = a_1 + a_2 X_2 + Z/\beta,$$

where T is the failure time and Z is subject to a standard EV-I distribution. Both estimating results excluded the 150°C data, which are all censored. Here we use all 40 lifetimes and obtain the original and the 2nd-order bias-corrected MLEs for the three parameters as follows,

$$\hat{a}_{1,40} = -13.3553, \quad \hat{a}_{2,40} = 9.7260, \quad \hat{\beta}_{40} = 3.0727;$$
$$\hat{a}_{1,40}^{\mathrm{bc2}} = -13.2948, \quad \hat{a}_{2,40}^{\mathrm{bc2}} = 9.6958, \quad \hat{\beta}_{40}^{\mathrm{bc2}} = 2.8489.$$

The variance estimates $J_n^{-1}(\hat{\theta}_n)$, $J_n^{-1}(\hat{\theta}_n^{\mathrm{bc2}})$ and $\hat{\mathrm{V}}_2(\hat{\theta}_n^{\mathrm{bc2}})$ are, respectively,

$$\begin{pmatrix} 2.2522 & -1.0433 & 0.1283 \\ -1.0433 & 0.4850 & -0.0714 \\ 0.1283 & -0.0714 & 0.4167 \end{pmatrix}, \begin{pmatrix} 2.4330 & -1.1247 & 0.1086 \\ -1.1247 & 0.5216 & -0.0600 \\ 0.1086 & -0.0600 & 0.3435 \end{pmatrix}, \begin{pmatrix} 1.7327 & -0.8068 & 0.1662 \\ -0.8068 & 0.3775 & -0.0908 \\ 0.1662 & -0.0908 & 0.5513 \end{pmatrix}$$

To test the null hypothesis $H_0 : a_2 = 0$, we have $|t_n| = |9.7260|/\sqrt{0.4850} = 13.9657$ and $|t_n^{\mathrm{bc2}}| = |9.6958|/\sqrt{0.5216} = 13.4250$ or $|t_n^{\mathrm{bc2}}| = |9.6958|/\sqrt{0.3775} = 15.7807$. The absolute values of all three statistics are greater than $u_{0.025} = 1.96$, which is the critical value corresponding to 5% significance level. These results indicate that the Weibull regression model should include the covariate variable X_2 into the scale parameter with the expression $a_1 + a_2 X_2$.

At the regular operating temperature 130°C, $X_2 = 1000/(273.2 + 130) = 2.4802$, so that the corresponding log-median lifetime is estimated by,

$$\hat{y}_{40,0.5} = -13.3553 + 9.7260 \times 2.4802 + \log(\log(2))/3.0727 = 10.6498,$$
$$\hat{y}^{\text{bc2}}_{40,0.5} = -13.2948 + 9.6958 \times 2.4802 + \log(\log(2))/2.8489 = 10.6261,$$

where $\log(\log(2))$ is the median of standard EV-I distribution $z_{0.5}$. The two vectors of coefficients are

$$c_n = (1, 2.4802, -\log(\log(2))/\hat{\beta}^2_{40})' \quad \text{and} \quad c^{\text{bc2}}_n = (1, 2.4802, -\log(\log(2))/(\hat{\beta}^{\text{bc2}}_{40})^2)'.$$

Based on the formula in Sec. 4.3, the 90% confidence interval for $y_{0.5}$ is (10.2567, 11.0429) with $\hat{y}_{40,0.5}$ and $J_n^{-1}(\hat{\theta}_n)$, and is (10.2232, 11.0289) with $\hat{y}^{\text{bc2}}_{40,,0.5}$ and $J_n^{-1}(\hat{\theta}^{\text{bc2}}_n)$, or (10.2625, 10.9896) with $\hat{y}^{\text{bc2}}_{40,,0.5}$ and $\hat{\text{V}}_2(\hat{\theta}^{\text{bc2}}_n)$, respectively. Further based on the original MLE $\hat{\theta}_n$, the estimate of the median lifetime is $\hat{T}_{40,0.5} = \exp(10.6498) = 42184$ with an associate approximate 90% confidence interval (28471, 62499). When the 2nd-order bias-corrected MLE $\hat{\theta}^{\text{bc2}}_n$ is used, the estimate of the median life is $\hat{T}^{\text{bc2}}_{40,0.5} = \exp(10.6261) = 41196$ with an approximate 90% confidence interval (27534, 61633) or (28640, 59253), corresponding to the variance estimate $J_n^{-1}(\hat{\theta}^{\text{bc2}}_n)$ or $\hat{\text{V}}_2(\hat{\theta}^{\text{bc2}}_n)$.

Note: In this example, it is type-I censoring with 4 censoring times (i.e., 8064, 5448, 1680, 528) for 4 subgroups respectively. Thus the censoring times $C = (C_1, \ldots, C_{40})$ are, respectively, $\{C_i = 8064,\ i = 1, \ldots, 10\}$, $\{C_i = 5448,\ i = 11, \ldots, 20\}$, $\{C_i = 1680,\ i = 21, \ldots, 30\}$, $\{C_i = 528,\ i = 31, \ldots, 40\}$.

7 Discussion and Conclusion

In this paper, we proposed a general multivariate bias-correction method for correcting the MLE and the variance, and hence improving t-ratios for the parameters of the Weibull regression model for both complete and randomly censored data. The method, based on a third-order stochastic expansion for the MLE, and a simple bootstrap procedure, is an extension of the bias-correction method in [13] by generalizing the method to a multi-dimensional version but adopting a nonparametric bootstrap procedure instead. Asymptotic properties of the proposed bias-corrected estimators are provided under some mild assumptions.

The results of several Monte Carlo simulation experiments show that, (i) the proposed method performs very well in correcting the bias rooted in the parameter MLEs for finite sample size data, and the shape parameter incurs most bias; (ii) with the corrected MLEs and second-order variance estimator, the improved t-ratios are able to greatly enhance the performances of some sequent inferences, including the significance test for scale-related parameters, and the construction of confidence intervals for the shape parameter and percentile. In particular, the superiority of the proposed method with respect to the significance test and the confidence interval for percentile is significant.

Although only the Weibull regression model is considered in this paper, the proposed method can be easily extended to other regression models. Furthermore, the proposed method can be applied to other more complicated censoring mechanisms besides the considered randomly censoring, such as progressively censoring [24], interval censoring [21], etc. It would be interesting to evaluate the performances of the method for different censoring mechanisms and censoring proportions. Also, a possible future work is to compare the proposed bootstrap-based method to some existing likelihood-related approaches, such as profile-kernel likelihood inference [8], penalized maximum likelihood approach [14], etc.

Appendix A: The expressions of $H_{kn}, k = 1, 2, 3$

Recall that $z_i(\theta) = \beta(\log s_i - a'x_i)$ and $e^{z_i(\theta)} = [s_i \exp(-a'x_i)]^\beta$, and that $z_i \equiv z_i(\theta_0)$. The $(p+1)\times(p+1)$ matrix $H_{1n} = \frac{\partial}{\partial\theta_0'}\psi_n(\theta_0)$ has the form:

$$H_{1n} = \frac{1}{n\beta_0^2}\begin{pmatrix} -\beta_0^4\sum_{i=1}^n e^{z_i}x_ix_i' & \beta_0^2\sum_{i=1}^n(e^{z_i}z_i + e^{z_i} - \delta_i)x_i \\ \beta_0^2\sum_{i=1}^n(e^{z_i}z_i + e^{z_i} - \delta_i)x_i' & -r - \sum_{i=1}^n e^{z_i}z_i^2 \end{pmatrix}.$$

The $(p+1)\times(p+1)^2$ matrix, $H_{2n} = \frac{\partial}{\partial\theta_0'}H_{1n}(\theta_0)$, has its first $(p+1)\times[p(p+1)]$ block:

$$\frac{1}{n\beta_0}\sum_{i=1}^n x_i' \otimes \begin{pmatrix} \beta_0^4 e^{z_i}x_ix_i' & -\beta_0^2 e^{z_i}(2+z_i)x_i \\ -\beta_0^2 e^{z_i}(2+z_i)x_i' & e^{z_i}z_i(2+z_i) \end{pmatrix},$$

and the last $(p+1)\times(p+1)$ block:

$$\frac{1}{n\beta_0^3}\begin{pmatrix} -\beta_0^4\sum_{i=1}^n e^{z_i}(2+z_i)x_ix_i' & \beta_0^2\sum_{i=1}^n e^{z_i}z_i(2+z_i)x_i \\ \beta_0^2\sum_{i=1}^n e^{z_i}z_i(2+z_i)x_i' & 2r - \sum_{i=1}^n e^{z_i}z_i^3 \end{pmatrix}.$$

Finally, the $(p+1)\times(p+1)^3$ matrix, $H_{3n}(\theta) = \frac{\partial}{\partial\theta_0'}H_{2n}(\theta_0)$, has the following blocks sorted in a row by $(\mathrm{B}_1^1, \mathrm{B}_1^2, \cdots, \mathrm{B}_p^1, \mathrm{B}_p^2, \mathrm{B}_{p+1}^1, \mathrm{B}_{p+2}^2)$, where

$$\mathrm{B}_j^1 = \frac{1}{n}\sum_{i=1}^n \mathbf{I}_j' x_ix_i' \otimes \begin{pmatrix} -\beta_0^4 e^{z_i}x_ix_i' & \beta_0^2 e^{z_i}(3+z_i)x_i \\ \beta_0^2 e^{z_i}(3+z_i)x_i' & -e^{z_i}(2+4z_i+z_i^2) \end{pmatrix},$$

$$\mathrm{B}_j^2 = \frac{1}{n\beta_0^2}\sum_{i=1}^n \mathbf{I}_j' x_i \begin{pmatrix} \beta_0^4 e^{z_i}(3+z_i)x_ix_i' & -\beta_0^2 e^{z_i}(2+4z_i+z_i^2)x_i \\ -\beta_0^2 e^{z_i}(2+4z_i+z_i^2)x_i' & e^{z_i}z_i^2(3+z_i) \end{pmatrix},$$

$$\mathrm{B}_{p+1}^1 = \frac{1}{n\beta_0^2}\sum_{i=1}^n x_i' \otimes \begin{pmatrix} \beta_0^4 e^{z_i}(3+z_i)x_ix_i' & -\beta_0^2 e^{z_i}(2+4z_i+z_i^2)x_i \\ -\beta_0^2 e^{z_i}(2+4z_i+z_i^2)x_i' & e^{z_i}z_i^2(3+z_i) \end{pmatrix},$$

$$\mathrm{B}_{p+1}^2 = \frac{1}{n\beta_0^4}\begin{pmatrix} -\beta_0^4\sum_{i=1}^n e^{z_i}(2+4z_i+z_i^2)x_ix_i' & \beta_0^2\sum_{i=1}^n e^{z_i}z_i^2(3+z_i)x_i \\ \beta_0^2\sum_{i=1}^n e^{z_i}z_i^2(3+z_i)x_i' & -6r - \sum_{i=1}^n e^{z_i}z_i^4 \end{pmatrix},$$

and $\mathbf{I}_j$ is a $p\times 1$ vector the jth element being 1, and the other elements being 0.

Proof of Theorem 1: Given some regular conditions, the MLE $\hat\theta_n$ is $\sqrt{n}$-consistent. And the differentiability and measureability of $\ell_n(\theta)$ are obvious. These facts lead to the Taylor series expansion of $\psi_n(\theta)$:

$$\begin{aligned} 0 &= \psi_n(\hat{\theta}_n) \\ &= \psi_n + H_{1n}(\hat{\theta}_n - \theta_0) + \frac{1}{2}H_{2n}[(\hat{\theta}_n - \theta_0) \otimes (\hat{\theta}_n - \theta_0)] \\ &\quad + \frac{1}{6}H_{3n}[(\hat{\theta}_n - \theta_0) \otimes (\hat{\theta}_n - \theta_0) \otimes (\hat{\theta}_n - \theta_0)] \\ &\quad + \frac{1}{6}[H_{3n}(\bar{\theta}_n) - H_{3n}][(\hat{\theta}_n - \theta_0) \otimes (\hat{\theta}_n - \theta_0) \otimes (\hat{\theta}_n - \theta_0)], \end{aligned}$$

where $\bar{\theta}_n$ lies between $\hat{\theta}_n$ and θ_0. As $\hat{\theta}_n = \theta_0 + O_p(n^{-1/2})$, we have $\bar{\theta}_n = \theta_0 + O_p(n^{-1/2})$.

For T following the Weibull regression model, it can be shown that the moment $\mathrm{E}T^k|\log T^l|$ is finite for every $k, l \geq 0$. This result further yields that $\frac{\beta}{n}\sum_{i=1}^n e^{z_i} = \mathrm{E}\left[\frac{\beta}{n}\sum_{i=1}^n e^{z_i}\right] + O_p(n^{-1/2})$, $\frac{\beta}{n}\sum_{i=1}^n e^{z_i} z_i = \mathrm{E}\left[\frac{\beta}{n}\sum_{i=1}^n e^{z_i} z_i\right] + O_p(n^{-1/2})$. Together with Assumption 4, we have the following:

1) $\psi_n = O_p(n^{-1/2})$ and $\mathrm{E}(\psi_n) = O(n^{-1})$;
2) $\mathrm{E}(H_{kn}) = O(1)$ and $H_{kn}^\circ = O_p(n^{-\frac{1}{2}})$, k = 1,2,3;
3) $\mathrm{E}(H_{1n})^{-1} = O(1)$ and $H_{1n}^{-1} = O_p(1)$;
4) $\| H_{kn}(\theta) - H_{kn} \| \leq \| \theta - \theta_0 \| B_n$, for θ in a neighborhood of θ_0, $k = 1, 2, 3$, and $\mathrm{E}|B_n| < c < \infty$ for some constant c;
5) $H_{3n}(\bar{\theta}_n) - H_{3n} = O_p(n^{-1/2})$.

The above proofs are straightforward and parallel those of [22] and [13]. The details are available from the authors. These make the stochastic expansion (8) valid. Assumption 4 also guarantees the transition from the stochastic expansion (8) to the results of Theorem 1.

Proof of Corollary 1: Note that the second-order bias $b_2 \equiv b_2(\theta_0)$ is of order $O(n^{-1})$, and the third-order bias $b_3 = b_3(\theta_0)$ is of order $O(n^{-3/2})$. If explicit expressions of $b_2(\theta_0)$ and $b_3(\theta_0)$ exist, then the "plug-in" estimates of b_2 and b_3 would be, respectively, $\hat{b}_2 = b_2(\hat{\theta}_n)$ and $\hat{b}_3 = b_3(\hat{\theta}_n)$, where $\hat{\theta}_n$ is the MLE of θ_0 defined at the beginning of Section 3. Under the additional assumptions in the corollary, we have ,

$$b_2(\hat{\theta}_n) = b_2(\theta_0) + \frac{\partial}{\partial\theta_0} b_2(\theta_0)(\hat{\theta}_n - \theta_0) + O_p(n^{-2}),$$

and $\mathrm{E}[b_2(\hat{\theta}_n)] = b_2(\theta_0) + \frac{\partial}{\partial\theta_0} b_2(\theta_0)\mathrm{E}(\hat{\theta}_n - \theta_0) + \mathrm{E}[O_p(n^{-2})] = b_2(\theta_0) + O(n^{-2})$, noting that $\frac{\partial}{\partial\theta_0} b_2(\theta_0) = O(n^{-1})$ and $\mathrm{E}(\hat{\theta}_n - \theta_0) = O(n^{-1})$. Similarly, $\mathrm{E}[b_3(\hat{\theta}_n)] = b_3(\theta_0) + O(n^{-5/2})$. These show that replacing θ_0 by $\hat{\theta}_n$ only (possibly) imposes additional bias of order $O_p(n^{-2})$ for $b_2(\hat{\theta}_n)$, and an additional bias of order $O_p(n^{-5/2})$ for $b_3(\hat{\theta}_n)$, leading to $\mathrm{Bias}(\hat{\theta}_n^{\mathrm{bc2}}) = O(n^{-3/2})$ and $\mathrm{Bias}(\hat{\theta}_n^{\mathrm{bc3}}) = O(n^{-2})$.

Our bootstrap estimate has two step approximations, one is that described above, and the other is the bootstrap approximations to the various expectations in (9) and (10), given $\hat{\theta}_n$. However, these approximations can be made arbitrarily accurate, for a given $\hat{\theta}_n$, by choosing an arbitrarily large B. The results of Corollary 1 thus follow.

Proof of Corollary 2: The additional assumptions stated in the Corollary 2 ensure that the 2nd-order variance for $\hat{\theta}_n^{\mathrm{bc2}}$ has $\mathrm{V}_2(\hat{\theta}_n^{\mathrm{bc2}}) = \mathrm{V}_2(\hat{\theta}_n - \hat{b}_2) = \mathrm{V}_2(\hat{\theta}_n) - 2\mathrm{Cov}(\hat{\theta}_n, \hat{b}_2) + O(n^{-5/2}) = \mathrm{V}_2(\hat{\theta}_n) + O(n^{-2})$ as the other terms can all be merged into $O(n^{-5/2})$ and $\mathrm{Cov}(\hat{\theta}_n, \hat{b}_2) = O(n^{-2})$. This proves the first equation in Corollary 2.

The bootstrap approximation to the second-order variance for $\hat{\theta}_n^{\mathrm{bc2}}$ actually also requires the approximations to various expectations, as shown by (12). Hence following the proof in Corollary 1, we know that for a given $\hat{\theta}_n^{\mathrm{bc2}}$, choosing an arbitrarily large B would make those approximations arbitrarily accurate. The second equation in Corollary 2 thus holds.

Acknowledgement

This work is supported by research grants from the National Natural Science Foundation of China (Grant No. 71401146), and from MOE (Ministry of Education) Key Laboratory of Econometrics and Fujian Key Laboratory of Statistical Sciences, China.

References

[1] Bobrowski S, Chen, H., Doring, M., Jensen, U., and Schinkothe, W. (2015). Estimation of the lifetime distribution of mechatronic systems in the presence of a covariate: A comparison among parametric, semiparametric and nonparametric models. *Reliability Engineriing & System Safety* 139, 105-112.

[2] Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.

[3] Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and Their Applications.* London: Cambridge University Press.

[4] Hashimoto, E. M., Ortega, E. M. M., Cancho, V. G., and Cordeiro, G. M. (2010). The log-exponentiated Weibull regression model for interval-censored data. *Computational Statistics & Data Analysis* 54(4), 1017-1035.

[5] Hirose, H. (1999). Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution. *IEEE Transactions on Dielectrics and Electrical Insulation* 6, 66-69.

[6] Horn, R. A. and Johnson, C. R. (1985). *Matrix Analysis*. Cambridge: Cambridge University Press.

[7] Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.

[8] Lam, C. and Fan, J. Q.(2008). Profile-kernel likelihood inference with diverging number of parameters. *Annals of Statistics* 36(5), 2232-2260.

[9] Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data, 2nd Ed*. New Jersey: Wiley.

[10] Nelson, W. B. (2004). *Applied Life Data Analysis*. New Jersey: Wiley.

[11] Nelson, W. B. and Hahn, G. J. (1972). Linear estimation of a regression relationship from censored data – Part I. Simple methods and their applications (with discussion). *Technometrics*. 14, 247-276.

[12] Rilstone, P., Srivastava, V. K. and Ullah A. (1996). The second-order bias and mean squared error of nonlinear estimators. *Journal of Econometrics* 75, 369-395.

[13] Shen, Y. and Yang, Z.L. (2015). Bias-Correction for Weibull Common Shape Estimation. *Journal of Statistical Computation and Simulation* 85(15), 3017-3046.

[14] Tran, M. N. (2009). Penalized maximum likelihood principle for choosing ridge parameter. *Communications in Statistics - Simulation and Computation* 38(8), 1610-1624.

[15] Van Der Vaart, A. W. (1998). *Asymptotic Statistics*. Cambridge: Cambridge University Press.

[16] White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*

53, 1-16.

[17] Xu, A., Tang, Y., and Guan, Q. (2014). Bayesian analysis of masked Data in step-stress accelerated life testing. *Communications in Statistics - Simulation and Computation* 43(8), 2016-2030.

[18] Yang, Z. L. and Xie, M. (2003). Efficient estimation of the Weibull shape parameter based on modified profile likelihood. *Journal of Statistical Computation and Simulation* 73, 115-123.

[19] Yang, Z. L. and Lin, D. K. J. (2007). Improved maximum-likelihood estimation for the common shape parameter of several Weibull populations. *Applied Stochastic Models in Business and Industry* 23, 373-383.

[20] Ye, Z. S., Xie, M., and Tang, L. C. (2013). Reliability evaluation of hard disk drive failures based on counting processes. *Reliability Engineering & System Safety* 109, 110-118.

[21] Yang, M. G., Chen, L.H., and Dong, G. H. (2015). Semiparametric Bayesian accelerated failure time model with interval-censored data. *Journal of Statistical Computation and Simulation* 85(10), 2049-2058.

[22] Yang, Z. L. (2015). A general method for third-order bias and variance corrections on a nonlinear estimator. *Journal of Econometrics* 186, 178-200.

[23] Ying, Z. (1993). A large sample study of rank estimation for censored regression data. *The Annals of Statistics* 21, 76-99.

[24] Zhang, M., Ye, Z. S., and Xie, M. (2014). A Stochastic EM Algorithm for Progressively Censored Data Analysis. *Quality and Reliability Engineering International* 30(5), 711-722.

Table 1: Empirical mean [rmse](se) of the estimators of all parameters, complete data

$n = 20$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5833 [.1520](.1271)	0.5228 [.1211](.1190)	5.0	4.8849 [.6886](.6790)	4.9635 [.6898](.6889)
0.8	0.9250 [.2266](.1891)	0.8282 [.1737](.1714)	5.0	4.9058 [.5269](.5184)	4.9541 [.5234](.5214)
1.0	1.1556 [.2812](.2343)	1.0339 [.2145](.2118)	5.0	4.9260 [.4325](.4262)	4.9653 [.4293](.4279)
2.0	2.2946 [.5660](.4833)	2.0532 [.4369](.4337)	5.0	4.9405 [.4109](.4066)	4.9598 [.4102](.4082)
5.0	5.7411 [1.413](1.204)	5.1335 [1.091](1.083)	5.0	4.9717 [.2846](.2832)	4.9794 [.2844](.2837)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{3n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	0.9939 [.7912](.7912)	0.9976 [.7918](.7919)	1.0	1.0018 [.8571](.8572)	0.9954 [.8544](.8544)
1.0	0.9950 [.4703](.4703)	0.9977 [.4699](.4699)	1.0	1.0096 [.4533](.4532)	1.0080 [.4531](.4530)
1.0	1.0048 [.3621](.3621)	1.0027 [.3618](.3618)	1.0	0.9996 [.3803](.3803)	1.0044 [.3802](.3802)
1.0	0.9943 [.1837](.1836)	0.9960 [.1836](.1836)	1.0	0.9979 [.1883](.1882)	0.9997 [.1881](.1881)
1.0	0.9998 [.0762](.0762)	0.9998 [.0761](.0761)	1.0	1.0022 [.0741](.0741)	1.0017 [.0740](.0740)
$n = 50$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5287 [.0686](.0624)	0.5041 [.0599](.0597)	5.0	4.9636 [.3068](.3046)	4.9971 [.3039](.3039)
0.8	0.8439 [.1091](.0998)	0.8047 [.0959](.0958)	5.0	4.9818 [.1933](.1924)	5.0025 [.1920](.1920)
1.0	1.0566 [.1373](.1251)	1.0076 [.1203](.1201)	5.0	4.9812 [.1547](.1536)	4.9978 [.1532](.1532)
2.0	2.1134 [.2727](.2480)	2.0152 [.2380](.2376)	5.0	4.9921 [.0868](.0864)	5.0004 [.0863](.0863)
5.0	5.2728 [.6830](.6262)	5.0288 [.6015](.6008)	5.0	4.9961 [.0308](.0305)	4.9994 [.0304](.0304)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{3n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	0.9987 [.4290](.4290)	0.9982 [.4290](.4290)	1.0	0.9944 [.4291](.4291)	0.9940 [.4291](.4290)
1.0	1.0000 [.2733](.2733)	0.9998 [.2734](.2734)	1.0	0.9961 [.2703](.2703)	0.9951 [.2703](.2703)
1.0	1.0014 [.2120](.2120)	1.0013 [.2121](.2121)	1.0	1.0045 [.2145](.2145)	1.0045 [.2145](.2145)
1.0	0.9972 [.1075](.1075)	0.9976 [.1075](.1075)	1.0	1.0004 [.1090](.1090)	0.9997 [.1090](.1090)
1.0	1.0003 [.0432](.0432)	1.0001 [.0432](.0432)	1.0	0.9987 [.0430](.0430)	0.9988 [.0430](.0430)
$n = 100$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5137 [.0436](.0414)	0.5012 [.0404](.0404)	5.0	4.9860 [.2157](.2153)	5.0029 [.2150](.2150)
0.8	0.8209 [.0691](.0659)	0.8009 [.0645](.0645)	5.0	4.9878 [.1333](.1327)	4.9983 [.1327](.1327)
1.0	1.0268 [.0875](.0833)	1.0017 [.0814](.0814)	5.0	4.9914 [.1066](.1063)	4.9998 [.1061](.1061)
2.0	2.0532 [.1735](.1652)	2.0036 [.1617](.1616)	5.0	4.9952 [.0540](.0538)	4.9995 [.0537](.0537)
5.0	5.1303 [.4333](.4133)	5.0050 [.4051](.4051)	5.0	4.9984 [.0214](.0214)	5.0001 [.0213](.0214)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{3n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	0.9974 [.2931](.2931)	0.9978 [.2933](.2933)	1.0	0.9988 [.2967](.2967)	0.9985 [.2971](.2971)
1.0	0.9998 [.1813](.1813)	0.9997 [.1814](.1814)	1.0	1.0011 [.1840](.1840)	1.0000 [.1842](.1842)
1.0	1.0009 [.1471](.1471)	1.0009 [.1471](.1471)	1.0	1.0011 [.1442](.1442)	1.0012 [.1443](.1443)
1.0	0.9995 [.0738](.0738)	0.9995 [.0738](.0738)	1.0	1.0005 [.0743](.0743)	1.0004 [.0744](.0744)
1.0	1.0004 [.0292](.0292)	1.0003 [.0292](.0292)	1.0	0.9997 [.0292](.0292)	0.9998 [.0292](.0292)

Table 2: Empirical mean [rmse](se) of the estimators of all parameters, censored data

$n = 20$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5704 [.1541](.1371)	0.5142 [.1299](.1292)	5.0	4.8934 [.7363](.7286)	4.9078 [.7397](.7340)
0.8	0.9209 [.2453](.2134)	0.8266 [.1951](.1933)	5.0	4.9035 [.6364](.6290)	4.9131 [.6343](.6284)
1.0	1.1492 [.3132](.2754)	1.0268 [.2479](.2465)	5.0	4.9296 [.5304](.5257)	4.9362 [.5285](.5247)
2.0	2.3642 [.7590](.6660)	2.0737 [.5764](.5717)	5.0	4.9063 [.6074](.6002)	4.9096 [.6071](.6003)
5.0	6.3500 [2.580](2.199)	5.3456 [1.819](1.786)	5.0	4.9461 [.4434](.4401)	4.9488 [.4436](.4407)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{3n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	1.0425 [.9219](.9209)	1.0202 [.9618](.9616)	1.0	1.0590 [.8648](.8628)	1.0441 [.8760](.8749)
1.0	1.0593 [.5626](.5595)	1.0481 [.5673](.5653)	1.0	1.0577 [.5726](.5697)	1.0509 [.5751](.5729)
1.0	1.0456 [.4819](.4798)	1.0299 [.4802](.4792)	1.0	1.0478 [.4803](.4779)	1.0369 [.4803](.4789)
1.0	1.0325 [.2641](.2621)	1.0275 [.2676](.2662)	1.0	1.0356 [.2729](.2706)	1.0298 [.2755](.2739)
1.0	1.0127 [.1343](.1337)	1.0130 [.1391](.1385)	1.0	1.0144 [.1324](.1317)	1.0169 [.1385](.1375)
$n = 50$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5232 [.0721](.0682)	0.5015 [.0655](.0655)	5.0	4.9737 [.3258](.3248)	4.9888 [.3226](.3224)
0.8	0.8418 [.1211](.1136)	0.8054 [.1089](.1088)	5.0	4.9876 [.2086](.2083)	4.9966 [.2070](.2070)
1.0	1.0540 [.1530](.1432)	1.0075 [.1372](.1370)	5.0	4.9874 [.1714](.1710)	4.9942 [.1704](.1703)
2.0	2.1340 [.3339](.3059)	2.0274 [.2920](.2907)	5.0	4.9943 [.1001](.1000)	4.9982 [.0997](.0997)
5.0	5.4230 [.9435](.8434)	5.0876 [.7907](.7859)	5.0	4.9975 [.0527](.0526)	4.9999 [.0528](.0528)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{2n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	1.0317 [.4759](.4749)	1.0196 [.4714](.4710)	1.0	1.0264 [.4844](.4837)	1.0140 [.4808](.4806)
1.0	1.0243 [.3157](.3148)	1.0147 [.3134](.3131)	1.0	1.0233 [.3111](.3102)	1.0133 [.3092](.3089)
1.0	1.0145 [.2519](.2515)	1.0061 [.2506](.2506)	1.0	1.0187 [.2615](.2609)	1.0099 [.2603](.2602)
1.0	1.0102 [.1531](.1528)	1.0054 [.1529](.1528)	1.0	1.0108 [.1474](.1470)	1.0068 [.1473](.1472)
1.0	1.0050 [.0746](.0744)	1.0036 [.0749](.0748)	1.0	1.0050 [.0760](.0759)	1.0037 [.0764](.0763)
$n = 100$					
β	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	a_1	$\hat{a}_{1n}$	$\hat{a}_{1n}^{\text{bc2}}$
0.5	0.5109 [.0479](.0466)	0.5001 [.0456](.0457)	5.0	4.9915 [.2230](.2229)	5.0001 [.2222](.2222)
0.8	0.8198 [.0786](.0761)	0.8016 [.0745](.0745)	5.0	4.9946 [.1451](.1450)	4.9996 [.1446](.1446)
1.0	1.0254 [.0999](.0966)	1.0018 [.0945](.0944)	5.0	4.9951 [.1190](.1189)	4.9992 [.1186](.1186)
2.0	2.0627 [.2096](.2000)	2.0082 [.1949](.1948)	5.0	4.9972 [.0691](.0691)	4.9996 [.0691](.0691)
5.0	5.1994 [.5734](.5376)	5.0319 [.5197](.5188)	5.0	4.9990 [.0349](.0349)	5.0005 [.0350](.0350)
a_2	$\hat{a}_{2n}$	$\hat{a}_{2n}^{\text{bc2}}$	a_3	$\hat{a}_{3n}$	$\hat{a}_{3n}^{\text{bc2}}$
1.0	1.0181 [.3275](.3270)	1.0096 [.3249](.3248)	1.0	1.0170 [.3336](.3332)	1.0076 [.3311](.3310)
1.0	1.0096 [.2122](.2120)	1.0025 [.2113](.2113)	1.0	1.0070 [.2181](.2180)	1.0002 [.2172](.2173)
1.0	1.0082 [.1771](.1769)	1.0022 [.1765](.1765)	1.0	1.0089 [.1778](.1775)	1.0029 [.1771](.1771)
1.0	1.0059 [.1024](.1023)	1.0028 [.1022](.1022)	1.0	1.0034 [.1026](.1025)	1.0002 [.1026](.1026)
1.0	1.0027 [.0478](.0477)	1.0017 [.0479](.0479)	1.0	1.0027 [.0488](.0487)	1.0017 [.0489](.0489)

Table 3: Empirical significance levels: two-sided tests of $H_0 : a_2 = 0$, complete data

β	Test	10%	5%	1%	10%	5%	1%	10%	5%	1%
		$n = 20$			$n = 50$			$n = 100$		
0.5	(1)	0.1481	0.0881	0.0317	0.1207	0.0656	0.0169	0.1094	0.0601	0.0132
0.5	(2)	0.0998	0.0537	0.0147	0.0987	0.0506	0.0115	0.1005	0.0521	0.0107
0.8	(1)	0.1499	0.0893	0.0293	0.1230	0.0677	0.0157	0.1100	0.0533	0.0136
0.8	(2)	0.0987	0.0506	0.0147	0.1013	0.0514	0.0107	0.0975	0.0458	0.0101
1.0	(1)	0.1569	0.0935	0.0310	0.1170	0.0662	0.0188	0.1110	0.0599	0.0160
1.0	(2)	0.1057	0.0557	0.0136	0.0981	0.0533	0.0119	0.1001	0.0511	0.0131
2.0	(1)	0.1490	0.0910	0.0309	0.1230	0.0685	0.0176	0.1071	0.0547	0.0119
2.0	(2)	0.1019	0.0557	0.0152	0.1009	0.0516	0.0118	0.0949	0.0490	0.0089
5.0	(1)	0.1467	0.0904	0.0304	0.1222	0.0679	0.0170	0.1023	0.0556	0.0132
5.0	(2)	0.0975	0.0536	0.0134	0.1012	0.0510	0.0115	0.0919	0.0481	0.0114

Test: (1) t_n with $J_n^{-1}(\hat{\theta}_n)$, (2) t_n^{bc2} with $J_n^{-1}(\hat{\theta}_n^{\text{bc2}})$; nominal significance levels: 10%, 5%, 1%

Table 4: Empirical significance levels: two-sided tests of $H_0 : a_2 = 0$, censored data

β	Test	10%	5%	1%	10%	5%	1%	10%	5%	1%
		$n = 20$			$n = 50$			$n = 100$		
0.5	(1)	0.1396	0.0796	0.0230	0.1186	0.0632	0.0155	0.1129	0.0613	0.0125
0.5	(2)	0.1023	0.0544	0.0133	0.0976	0.0498	0.0100	0.1026	0.0519	0.0101
0.8	(1)	0.1409	0.0795	0.0210	0.1173	0.0601	0.0154	0.1099	0.0574	0.0135
0.8	(2)	0.1026	0.0522	0.0111	0.0987	0.0467	0.0109	0.1003	0.0507	0.0108
1.0	(1)	0.1451	0.0831	0.0250	0.1165	0.0635	0.0160	0.1104	0.0559	0.0133
1.0	(2)	0.1036	0.0566	0.0138	0.0963	0.0511	0.0114	0.0981	0.0487	0.0112
2.0	(1)	0.1467	0.0885	0.0275	0.1254	0.0695	0.0171	0.1105	0.0596	0.0147
2.0	(2)	0.1114	0.0595	0.0143	0.1043	0.0557	0.0128	0.0993	0.0529	0.0119
5.0	(1)	0.1703	0.1125	0.0441	0.1304	0.0706	0.0211	0.1134	0.0614	0.0155
5.0	(2)	0.1281	0.0751	0.0255	0.1083	0.0558	0.0146	0.101	0.0528	0.0111

Test: (1) t_n with $J_n^{-1}(\hat{\theta}_n)$, (2) t_n^{bc2} with $J_n^{-1}(\hat{\theta}_n^{\text{bc2}})$; nominal significance levels: 10%, 5%, 1%

Table 5: Empirical coverage probability (average length) of confidence intervals for β, complete data: $\hat{\beta}_n$ with variance $J_n^{-1}(\hat{\theta}_n)$, $\hat{\beta}_n^{\text{bc2}}$ with variance $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\text{bc2}})$

β_0	$1-\gamma=0.90$		$1-\gamma=0.95$		$1-\gamma=0.99$	
	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$
$n=20$						
0.5	0.8530 (0.3492)	0.9056 (0.3538)	0.9237 (0.4139)	0.9495 (0.4194)	0.9862 (0.5403)	0.9869 (0.5476)
0.8	0.8547 (0.5506)	0.9042 (0.5578)	0.9257 (0.6539)	0.9508 (0.6625)	0.9879 (0.8558)	0.9880 (0.8671)
1.0	0.8594 (0.6861)	0.9080 (0.6981)	0.9282 (0.8160)	0.9522 (0.8304)	0.9885 (1.0701)	0.9878 (1.0889)
2.0	0.8552 (1.3589)	0.9026 (1.3755)	0.9280 (1.6173)	0.9514 (1.6371)	0.9882 (2.1225)	0.9885 (2.1485)
5.0	0.8610 (3.3951)	0.9106 (3.4544)	0.9292 (4.0446)	0.9558 (4.1153)	0.9885 (5.3139)	0.9876 (5.4068)
$n=50$						
0.5	0.8778 (0.1942)	0.8881 (0.1917)	0.9398 (0.2314)	0.9416 (0.2284)	0.9897 (0.3041)	0.9869 (0.3002)
0.8	0.8761 (0.3102)	0.8909 (0.3071)	0.9387 (0.3696)	0.9405 (0.3659)	0.9871 (0.4858)	0.9832 (0.4809)
1.0	0.8809 (0.3884)	0.8934 (0.3841)	0.9377 (0.4628)	0.9432 (0.4577)	0.9871 (0.6083)	0.9869 (0.6015)
2.0	0.8822 (0.7765)	0.8933 (0.7684)	0.9394 (0.9252)	0.9435 (0.9156)	0.9876 (1.2159)	0.9861 (1.2032)
5.0	0.8756 (1.9373)	0.8854 (1.9096)	0.9378 (2.3084)	0.9370 (2.2754)	0.9883 (3.0338)	0.9842 (2.9904)
$n=100$						
0.5	0.8910 (0.1327)	0.8952 (0.1303)	0.9472 (0.1581)	0.9433 (0.1552)	0.9891 (0.2078)	0.9874 (0.2040)
0.8	0.8903 (0.2119)	0.8931 (0.2083)	0.9475 (0.2525)	0.9425 (0.2482)	0.9902 (0.3319)	0.9884 (0.3262)
1.0	0.8905 (0.2651)	0.8884 (0.2606)	0.9452 (0.3159)	0.9433 (0.3106)	0.9887 (0.4152)	0.9862 (0.4082)
2.0	0.8916 (0.5304)	0.8896 (0.5192)	0.9456 (0.6320)	0.9425 (0.6187)	0.9905 (0.8305)	0.9871 (0.8131)
5.0	0.8906 (1.3244)	0.8916 (1.3000)	0.9445 (1.5781)	0.9407 (1.5491)	0.9893 (2.0739)	0.9861 (2.0358)

Table 6: Empirical coverage probability (average length) of confidence intervals for β, censored data: $\hat{\beta}_n$ with variance $J_n^{-1}(\hat{\theta}_n)$, $\hat{\beta}_n^{\text{bc2}}$ with variance $\widehat{\mathrm{V}}_2(\hat{\theta}_n^{\text{bc2}})$

β_0	$1-\gamma=0.90$		$1-\gamma=0.95$		$1-\gamma=0.99$	
	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$	$\hat{\beta}_n$	$\hat{\beta}_n^{\text{bc2}}$
$n=20$						
0.5	0.8841 (0.3826)	0.8921 (0.3810)	0.9439 (0.4530)	0.9387 (0.4511)	0.9886 (0.5907)	0.9816 (0.5881)
0.8	0.8724 (0.6247)	0.8968 (0.6223)	0.9353 (0.7409)	0.9421 (0.7381)	0.9872 (0.9681)	0.9823 (0.9644)
1.0	0.8736 (0.7903)	0.8919 (0.7845)	0.9361 (0.9425)	0.9400 (0.9324)	0.9890 (1.2348)	0.9832 (1.2214)
2.0	0.8629 (1.7076)	0.9014 (1.7135)	0.9312 (2.0306)	0.9456 (2.0377)	0.9880 (2.6620)	0.9844 (2.6713)
5.0	0.8367 (4.9445)	0.9117 (5.1124)	0.9244 (5.8895)	0.9552 (6.0895)	0.9888 (7.7364)	0.9885 (7.9993)
$n=50$						
0.5	0.8949 (0.2178)	0.8963 (0.2155)	0.9522 (0.2596)	0.9474 (0.2568)	0.9913 (0.3411)	0.9875 (0.3375)
0.8	0.8893 (0.3566)	0.8927 (0.3512)	0.9454 (0.4249)	0.9411 (0.4185)	0.9908 (0.5584)	0.9864 (0.5500)
1.0	0.8857 (0.4485)	0.8902 (0.4423)	0.9449 (0.5345)	0.9409 (0.5270)	0.9897 (0.7024)	0.9842 (0.6926)
2.0	0.8818 (0.9374)	0.8899 (0.9222)	0.9375 (1.1170)	0.9390 (1.0989)	0.9894 (1.4680)	0.9837 (1.4442)
5.0	0.8780 (2.5368)	0.8958 (2.5093)	0.9424 (3.0228)	0.9462 (2.9901)	0.9888 (3.9726)	0.9869 (3.9296)
$n=100$						
0.5	0.8963 (0.1502)	0.8926 (0.1491)	0.9451 (0.1790)	0.9417 (0.1776)	0.9881 (0.2353)	0.9856 (0.2335)
0.8	0.8940 (0.2439)	0.8895 (0.2410)	0.9490 (0.2907)	0.9445 (0.2871)	0.9895 (0.3820)	0.9875 (0.3774)
1.0	0.8925 (0.3079)	0.8918 (0.3038)	0.9445 (0.3669)	0.9416 (0.3620)	0.9884 (0.4822)	0.9839 (0.4757)
2.0	0.8938 (0.6381)	0.8970 (0.6255)	0.9479 (0.7603)	0.9447 (0.7454)	0.9892 (0.9993)	0.9859 (0.9795)
5.0	0.8877 (1.6825)	0.8938 (1.6506)	0.9416 (2.0048)	0.9413 (1.9668)	0.9891 (2.6348)	0.9866 (2.5848)

Table 7: Empirical coverage probability (average length) of confidence intervals for $y_{0.5}$, complete data: $\hat{y}_{n,0.5}$ with variance $J_n^{-1}(\hat{\theta}_n)$, $\hat{y}^{\text{bc2}}_{n,0.5}$ with variance $\widehat{\mathrm{V}}_2(\hat{\theta}^{\text{bc2}}_n)$

$y_{0.5}$	$1-\gamma=0.90$		$1-\gamma=0.95$		$1-\gamma=0.99$	
	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$
$n=20$						
4.2670	0.8609 (1.6421)	0.9025 (1.8727)	0.9181 (1.9545)	0.9456 (2.2293)	0.9719 (2.5651)	0.9839 (2.9262)
4.5419	0.8584 (1.0123)	0.9060 (1.1611)	0.9183 (1.2041)	0.9496 (1.3814)	0.9717 (1.5788)	0.9860 (1.8119)
4.6335	0.8469 (0.8170)	0.8941 (0.9358)	0.9082 (0.9721)	0.9436 (1.1136)	0.9706 (1.2752)	0.9845 (1.4611)
4.8167	0.8582 (0.4113)	0.9061 (0.4711)	0.9171 (0.4883)	0.9480 (0.5595)	0.9712 (0.6387)	0.9846 (0.7323)
4.9267	0.8545 (0.1686)	0.8986 (0.1922)	0.9114 (0.2000)	0.9447 (0.2281)	0.9703 (0.2613)	0.9853 (0.2982)
$n=50$						
4.2670	0.8831 (1.0650)	0.9045 (1.1369)	0.9360 (1.2690)	0.9490 (1.3547)	0.9826 (1.6678)	0.9888 (1.7804)
4.5419	0.8839 (0.6684)	0.9011 (0.7137)	0.9301 (0.7965)	0.9452 (0.8504)	0.9823 (1.0467)	0.9880 (1.1176)
4.6335	0.8779 (0.5321)	0.8991 (0.5685)	0.9348 (0.6340)	0.9482 (0.6774)	0.9819 (0.8333)	0.9871 (0.8903)
4.8167	0.8858 (0.2675)	0.9047 (0.2855)	0.9362 (0.3188)	0.9513 (0.3401)	0.9812 (0.4189)	0.9868 (0.4470)
4.9267	0.8840 (0.1068)	0.9054 (0.1139)	0.9370 (0.1272)	0.9506 (0.1358)	0.9826 (0.1672)	0.9876 (0.1784)
$n=100$						
4.2670	0.8857 (0.7630)	0.8975 (0.7909)	0.9405 (0.9092)	0.9492 (0.9424)	0.9869 (1.1949)	0.9889 (1.2386)
4.5419	0.8940 (0.4779)	0.9062 (0.4949)	0.9446 (0.5695)	0.9519 (0.5897)	0.9871 (0.7484)	0.9899 (0.7750)
4.6335	0.8960 (0.3819)	0.9075 (0.3957)	0.9460 (0.4551)	0.9532 (0.4715)	0.9867 (0.5981)	0.9888 (0.6196)
4.8167	0.8868 (0.1911)	0.8992 (0.1980)	0.9414 (0.2277)	0.9483 (0.2359)	0.9860 (0.2992)	0.9882 (0.3100)
4.9267	0.8888 (0.0765)	0.9008 (0.0792)	0.9444 (0.0911)	0.9507 (0.0944)	0.9861 (0.1197)	0.9880 (0.1241)

Note: Each 5 rows of data corresponds to $\beta = 0.5, 0.8, 1.0, 2.0, 5.0$.

Table 8: Empirical coverage probability (average length) of confidence intervals for $y_{0.5}$, censored data: $\hat{y}_{n,0.5}$ with variance $J_n^{-1}(\hat{\theta}_n)$, $\hat{y}^{\text{bc2}}_{n,0.5}$ with variance $\widehat{\text{V}}_2(\hat{\theta}^{\text{bc2}}_n)$

$y_{0.5}$	$1-\gamma=0.90$		$1-\gamma=0.95$		$1-\gamma=0.99$	
	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$	$\hat{y}_{n,0.5}$	$\hat{y}^{\text{bc2}}_{n,0.5}$
$n=20$						
4.2670	0.8638 (1.7367)	0.9096 (1.9856)	0.9251 (2.0666)	0.9559 (2.3631)	0.9778 (2.7112)	0.9890 (3.1010)
4.5419	0.8639 (1.0950)	0.9069 (1.2553)	0.9218 (1.3014)	0.9535 (1.4923)	0.9772 (1.7047)	0.9883 (1.9556)
4.6335	0.8608 (0.8933)	0.9071 (1.0222)	0.9226 (1.0620)	0.9510 (1.2156)	0.9782 (1.3917)	0.9898 (1.5936)
4.8167	0.8647 (0.5088)	0.9025 (0.5828)	0.9244 (0.6021)	0.9512 (0.6903)	0.9768 (0.7846)	0.9856 (0.9005)
4.9267	0.8299 (0.2602)	0.8663 (0.3131)	0.8897 (0.3077)	0.9190 (0.3708)	0.9556 (0.4007)	0.9681 (0.4836)
$n=50$						
4.2670	0.8857 (1.0962)	0.9034 (1.1525)	0.9370 (1.3062)	0.9492 (1.3733)	0.9839 (1.7166)	0.9877 (1.8049)
4.5419	0.8777 (0.6947)	0.8987 (0.7308)	0.9318 (0.8277)	0.9439 (0.8708)	0.9831 (1.0878)	0.9875 (1.1445)
4.6335	0.8854 (0.5673)	0.9035 (0.5967)	0.9384 (0.6760)	0.9499 (0.7110)	0.9864 (0.8884)	0.9896 (0.9345)
4.8167	0.8803 (0.3159)	0.8942 (0.3306)	0.9331 (0.3764)	0.9448 (0.3939)	0.9834 (0.4947)	0.9867 (0.5177)
4.9267	0.8780 (0.1667)	0.8817 (0.1719)	0.9287 (0.1987)	0.9341 (0.2048)	0.9810 (0.2611)	0.9805 (0.2692)
$n=100$						
4.2670	0.8939 (0.7785)	0.9059 (0.7987)	0.9456 (0.9276)	0.9519 (0.9517)	0.9871 (1.2191)	0.9888 (1.2507)
4.5419	0.8903 (0.4954)	0.9007 (0.5083)	0.9431 (0.5903)	0.9494 (0.6057)	0.9884 (0.7758)	0.9900 (0.7960)
4.6335	0.8927 (0.4035)	0.9008 (0.4140)	0.9463 (0.4808)	0.9511 (0.4933)	0.9881 (0.6319)	0.9895 (0.6483)
4.8167	0.8848 (0.2250)	0.8913 (0.2302)	0.9365 (0.2680)	0.9432 (0.2743)	0.9884 (0.3523)	0.9889 (0.3605)
4.9267	0.8855 (0.1145)	0.8870 (0.1157)	0.9418 (0.1364)	0.9405 (0.1378)	0.9864 (0.1793)	0.9858 (0.1811)

Note: Each 5 rows of data corresponds to $\beta = 0.5, 0.8, 1.0, 2.0, 5.0$.

Table 9: Hours to failure of motorettes

150°C	170°C	190°C	220°C
8064+	1764	408	408
8064+	2772	408	408
8064+	3444	1344	504
8064+	3542	1344	504
8064+	3780	1440	504
8064+	4860	1680+	528+
8064+	5196	1680+	528+
8064+	5448+	1680+	528+
8064+	5448+	1680+	528+
8064+	5448+	1680+	528+

'+' indicates a censoring time.