Reliability Engineering and System Safety 212 (2021) 107622

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Product reliability analysis based on heavily censored interval data with batch effects

Liangliang Zhuang [a], Ancha Xu [b,*], Jihong Pang [c]

[a] Department of Mathematics, Wenzhou University, Zhejiang 325035, China
[b] Department of Statistics, Zhejiang Gongshang University, Zhejiang 310018, China
[c] College of Mechanical and Electrical Engineering, Wenzhou University, Zhejiang 325035, China

ARTICLE INFO

Keywords:
Two-stage approach
Fractional-random-weight bootstrap
Interval data
Heavily censored data
Confidence interval

ABSTRACT

In many industries including engineering, biology, and medical science, etc, interval failure data commonly exist. Utilizing the data to estimate product lifetime is often confounded with both heavy censoring and batch effects. To deal with the two characteristics, in this paper, we propose a novel two-stage method called fractional-random-weight bootstrap to help make interval estimation for both model parameters and future failure numbers. By carrying out various simulation studies, our method demonstrates the superiority over two other commonly-used bootstrap methods in terms of the relative bias, root mean squared error, and width of confidence intervals. When extremely heavy censoring is present, the advantage is more significant. In addition, we illustrate the application of the proposed methodology using a real dataset from experiments on printed circuit boards. By comparison, we show that misconsidering the batch effects in the interval data could lead to inaccurate predicted number of failures.

1. Introduction

1.1. Motivation

Product reliability is a common concern to many manufacturers. To assess the reliability, failure time data that originate from either laboratory life tests or field experiments are often collected. Depending on the type of a life test or experiment, various data formats including complete data, censored data, and interval failure data are available. Among them, the interval failure data is commonly seen in many practice. For example, in life tests, practitioners usually perform periodic, say monthly or bi-monthly inspections on sample test units. As a result, the time window of an event of failure instead of the exact failure time is recorded. However, utilizing the interval failure data to make reliability prediction is faced with two difficulties — heavy censoring and batch effects. Censoring happens when the duration of an experiment is less than the typical lifetime of a test unit so that this test unit is still alive when the experiment terminates. If highly-reliable products are tested, a number of test units would not fail until the end of a life test, which results in the case of heavy censoring. On the other hand, due to the variations in raw materials, units that are produced in the same batch behave consistently in terms of failure mechanism. On the contrary, those manufactured in different batches may behave diversely. If assigning different batches of test units to various test groups, the batch effects should be considered as they impact the following failure analysis.

To motivate our study, we provide a real dataset from life tests on a certain type of printed circuit board (PCB). The data comprises eight batches or groups, each containing 2000 products. These PCBs are tested in a chamber to guarantee that the testing environment are the same. Due to the complexity of the product testing procedures, engineers observe the lifetime of PCB at a fixed time every week for a total of 10 weeks, and the data are listed in Table 1. Unlike in conventional tests, the data type corresponds to the interval-censored data. Although there are 2000 units tested in each group, the data are heavily censored. Specifically, in groups 5–8, the cumulative failure numbers of products account for less than 2% of the total number in each group. As an illustration, assume that the lifetime of PBC in each batch independently follows Weibull distribution, we estimate the model parameters of each batch that are represented in Table 2. It can been seen that scale parameters among batches disperse widely, which implies that the batch effects may exist in the dataset. According to the observed patterns in the data, three problems need to be concerned by the manufacturer: (1) How to construct a model for the heavily censored interval data with batch effects (also called block effects in reliability area)? (2) How to obtain the point and interval estimates of the model parameters? (3) How to predict the number of failures in a

* Corresponding author.
E-mail address: xuancha2011@aliyun.com (A. Xu).

https://doi.org/10.1016/j.ress.2021.107622
Received 17 July 2020; Received in revised form 16 February 2021; Accepted 11 March 2021
Available online 26 March 2021

Table 1
Interval failure data of PCB.

Batches	0–1	1–2	2–3	3–4	4–5	5–6	6–7	7–8	8–9	9–10	Censored
1	0	2	4	10	19	21	39	65	71	99	1670
2	0	3	6	4	23	30	44	54	64	81	1691
3	0	2	4	7	5	13	14	17	28	28	1882
4	0	0	2	2	4	8	9	13	13	16	1933
5	0	0	0	0	3	1	2	6	4	4	1980
6	0	1	0	0	0	3	2	4	5	6	1979
7	0	0	0	2	1	0	3	2	3	3	1986
8	0	0	0	0	2	2	0	1	2	2	1991

Table 2
Parameter estimation of each batch when Weibull distribution is assumed.

Parameters	Batches							
	1	2	3	4	5	6	7	8
Scale	16.311	17.640	28.602	31.901	43.893	42.661	65.461	66.013
Shape	3.503	3.144	2.665	2.913	3.111	3.137	2.638	2.848

certain future period? The former two problems are the basis for solving the third problem, while the third problem is related with inventory control. Accurate prediction will help reduce the inventory costs, and thus is very important for the manufacturer.

1.2. Related literature

Conventional reliability analysis are commonly based on the assumption that the data are from a completely random designed experiment [1]. However, when some external effects are considered, such as random effects, subsampling effects, block effects, etc., the data may be not completely randomized anymore. Feiveson and Kulkarni [2], Leon et al. [3] and Lv et al. [4] have emphasized the importance of incorporating the effect of subsampling into the analysis of lifetime data.

Freeman and Vining [5] proposed a two-stage method to analyze accelerated life test (ALT) data with the effect of subsampling. The method can be easily implemented by using common statistical softwares (for example, Minitab or JMP). However, Vining et al. [6] indicated that the two-stage method was prone to bring biases in the estimation of shape parameters when the Weibull distribution was assumed. Furthermore, the confidence intervals (CIs) of percentile lifetimes cannot be computed based on the two-stage method. Zhang et al. [7] proposed a bias reduction method for the estimation of Weibull shape parameters with complete and censored data. Wang et al. [8] developed a two-stage bootstrap method based on an unbiased factor approach. This method included the resampling and subsampling steps, thereby facilitating the computation of CIs. Wang et al. [9] constructed a model interconnecting percentiles and stress factors to derive the likelihood-based (LB) inference on percentile lifetimes, and the simulation showed that the results obtained by [9] were better than those reported by [8]'s. Kensler et al. [10] compared the two-stage method with conventional analysis concerning an experiment with both subsampling and block effects. Wang et al. [11] proposed a bootstrap analysis of designed experiments for reliability improvement with a non-constant scale parameter. Lv et al. [12] compared the performance of the two-stage approach for right-censored data.

In addition to the two-stage method, the nonlinear mixed-model (NLMM) can be considered as the other approach to address subsampling. It was first proposed by Freeman and Vining [13], and extended by [14–18]. Specifically, Lv et al. [14] extended the NLMM method to incorporate the effects of both subsampling and non-constant shape parameters. Kensler et al. [15] adjusted the NLMM method to conduct the experiments with random blocks and subsampling. Seo and Pan [16] proposed a generalized linear mixed model (GLMM) to analyze the ALT data obtained from the experiments with constrained randomization. Seo and Pan [17] modified the GLMM approach to incorporate random effects in step-stress ALT. Medlin et al. [18] developed a NLMM for a split-plot reliability experiment with subsampling and right-censored lifetime data. Recently, Zhu [19] developed a hierarchical model for estimating reliability performances under the assumption that the lifetimes of units are Weibull distributed with block effects.

The aforementioned studies mainly correspond to the analysis of the complete or right-censored data with batch effects in reliability experiments. However, interval censoring has become increasingly common in the areas producing the failure time data. Lindsay and Ryan [20] provided a tutorial on statistical methods for the interval-censored data. Tan [21] studied the interval data problem for the Weibull distribution. Peng et al. [22] proposed a new method to evaluate and predict the dynamic reliability of a repairable system subject to interval-censored problem. García-Mora et al. [23] used a generalized non-linear model for interval-censored data to handle the service life of the pipeline from installation to failure. Sun [24] gave a comprehensive introduction of the methods for analyzing interval failure data.

In the reliability data analysis, the focus is generally emphasized on parameter estimates and corresponding CIs. The original methods based on the right-censored data are usually applicable to the interval data. However, for the interval data with heavily censored and batch effects, these methods will lead to unsatisfactory CIs as shown in the simulation studies of this paper. The contributions of this paper have two aspects. Firstly, to the best of our knowledge, reliability analysis based on heavily censored interval data with batch effects has been not well studied yet. Secondly, a novel two-stage fractional-random-weight (FRW) bootstrap method is proposed to construct the CIs of the model parameters. To evaluate the performance of the proposed method, we compare it with two alternative approaches via simulation studies.

1.3. Overview

This paper is organized as follows. In Section 2, we discuss the assumptions about the proposed model. Then, a new procedure for the two-stage method is presented in Section 3. The results of the simulation studies are reported to compare the performance of the proposed method with the other considered two-stage approaches in Section 4. A real dataset is analyzed with illustrative purposes in Section 5. Finally, we present conclusions and discussions for future work in Section 6.

2. Model

In the present study, we consider an experiment involving k batches so that each of them contains n_i units. The measurement times are fixed as follows: $t_{i_0} < t_{i_1} < t_{i_2} < \cdots < t_{id_i}$, where t_{i_0} is the initial time, and t_{id_i} is the censored time for the ith batch. We assume that $t_{i_0} = 0$, $t_{id_i+1} = \infty$, and m_{ij} is the number of failures in $\left(t_{ij-1}, t_{ij}\right]$, $i = 1, \ldots, k, j = 1, \ldots, d_i$. $m_{id_i+1} = n_i - \sum_{i=1}^{d_i} m_{ij}$ is the number of products that have not failed when the experiment is terminated. Denote that the observed data $\boldsymbol{D} = \{\left(\left(t_{ij-1}, t_{ij}\right], m_{ij}\right), i = 1, \ldots, k, j = 1, \ldots, d_i + 1\}$ (see Fig. 1).

Let T_i be the lifetime of the product in the ith batch, and assume that T_i follows a distribution in the log-location-scale family. Then the cumulative distribution function (CDF) and the probability density function (PDF) of T_i are formulated as:

$$F\left(t; \boldsymbol{\omega}_i\right) = \Phi\left[\frac{\log(t_{ij}) - \mu_i}{\sigma_i}\right] \text{ and } f\left(t; \boldsymbol{\omega}_i\right) = \frac{1}{\sigma_i t_{ij}} \phi\left[\frac{\log(t_{ij}) - \mu_i}{\sigma_i}\right], \quad (1)$$

where Φ and ϕ are the standard CDF and PDF corresponding to the location-scale family of distributions, respectively. Here, $\boldsymbol{\omega}_i = (\mu_i, \sigma_i)'$, where $-\infty < \mu_i < \infty$ and $\sigma_i > 0$ are the location and scale parameters, respectively. The log-location-scale family contains some commonly used distributions, such as exponential, Weibull and lognormal distribution. For the Weibull and lognormal distribution, Φ equals Φ_{sev} and Φ_{nor}, respectively, where $\Phi_{\text{sev}}(z) = 1 - \exp[-\exp(z)]$ and Φ_{nor} are

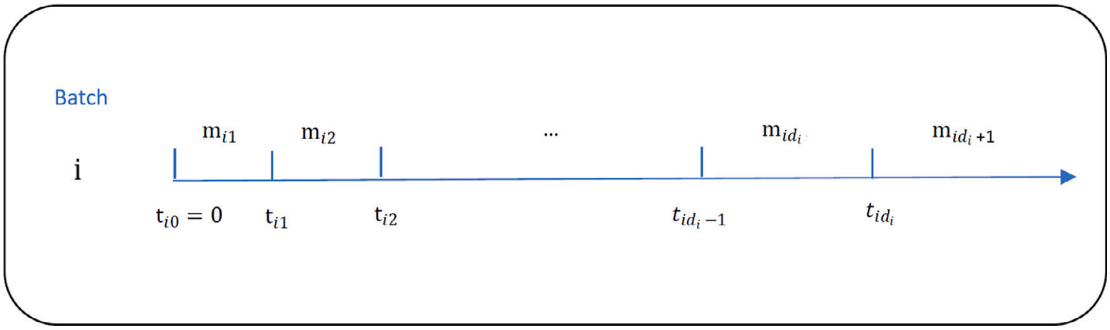

Fig. 1. The observed interval failure data for the ith batch.

the CDFs of the standard smallest extreme value distribution and the standard normal distribution, respectively.

Under the assumption (1), the probabilities that the unit will fail in the interval time $\left(t_{ij-1}, t_{ij}\right], j = 1, \ldots, d_i + 1$, can be formulated as

$$\lambda(t_{ij}; \boldsymbol{\omega}_i) = F\left(t_{ij}; \boldsymbol{\omega}_i\right) - F\left(t_{ij-1}; \boldsymbol{\omega}_i\right), j = 1, \ldots, d_i + 1,$$

where $F\left(t_{id_i+1}; \boldsymbol{\omega}_i\right) = F\left(\infty; \boldsymbol{\omega}_i\right) = 1$. We assume that batch effects are related with the location parameter μ_i through a log linear function given by

$$\log \mu_i = \eta + \varepsilon_i, \quad i = 1, 2, \ldots, k, \tag{2}$$

where η is unknown parameter and $\varepsilon_i \sim N(0, \delta^2)$. In summary, the assumptions used in this article can be specified as below:

1. The lifetime T_i follows a log-location-scale distribution.
2. The failure mechanism for different batches is the same, namely, σ_i is constant for all batches. For convenience, we define $\sigma_1 = \cdots = \sigma_k = \sigma$. Then the impact of batch effects is only addressed through the location parameters $(\mu_1, \mu_2, \ldots, \mu_k)$.
3. Given the location parameters, the units among different batches are considered as independent.

3. The proposed methodology

Two-stage methods are widely used in analyzing the failure data with batch effects [5,8,9]. The two-stage method proposed by [5] is easily implemented and can be extended to deal with interval failure data. Except for this extension, a new two-stage method combined with fractional-random-weight (FRW) bootstrap is also proposed to analyze the interval data with batch effects.

3.1. Two-stage methods

The direct two-stage approach proposed by [5] can be applied to analyze the interval failure data with batch effects, and the procedure can be summarized as follows:

1. The first stage is devoted to obtain the estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ using the maximum likelihood (ML) method, where $\boldsymbol{\theta} = (\sigma, \mu_1, \ldots, \mu_k)$. Thus, μ_is are treated as fixed unknown parameters. Given the observed data $\boldsymbol{D}$, the log-likelihood function can be formulated as follows:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{k} \sum_{j=1}^{d_i+1} m_{ij} \log \left\{ \Phi\left[\frac{\log(t_{ij}) - \mu_i}{\sigma}\right] - \Phi\left[\frac{\log(t_{ij-1}) - \mu_i}{\sigma}\right] \right\}. \tag{3}$$

The ML estimates of $\boldsymbol{\theta}$ can be obtained by maximizing $\ell(\boldsymbol{\theta})$, which can be implemented easily using the function *optim()* provided in *R* software. The asymptotic covariance matrix of $\hat{\boldsymbol{\theta}}$, $\hat{\Sigma}_{\hat{\theta}}$, can be obtained by inverting the observed Fisher information matrix.

2. In the second stage, the parameters in the distribution of batch effects are estimated. $\hat{\mu}_i$s obtained in the first stage are treated as the "observations" from model (2), then the estimates of (η, δ^2) can be easily obtained:

$$\hat{\eta} = \frac{1}{k} \sum_{i=1}^{k} \ln \hat{\mu}_i, \quad \hat{\delta}^2 = \frac{1}{k-1} \sum_{i=1}^{k} \left(\ln \hat{\mu}_i - \frac{1}{k} \sum_{i=1}^{k} \ln \hat{\mu}_i \right)^2, \quad i = 1, \ldots, k. \tag{4}$$

This method has several significant advantages. It is computationally simple and allows incorporating the batch effects. The CI of a function of the parameters μ_i and σ^2 can be computed by using asymptotic normality theorem of ML estimator and delta method. For example, if the p-percentile in the ith group $t_{ip} = \exp(\mu_i + \sigma \Phi^{-1}(p))$ is of interest, then by plug-in method, the point estimate of t_{ip} is

$$\hat{t}_{ip} = \exp(\hat{\mu}_i + \hat{\sigma} \Phi^{-1}(p)). \tag{5}$$

From Meeker and Escobar [1], the $100(1-\alpha)\%$ CI of t_{ip} can be constructed as follows

$$\left[\hat{t}_{ip}/w, \hat{t}_{ip} \times w\right], \tag{6}$$

where $w = \exp\left(z_{(1-\alpha/2)} \widehat{\mathrm{se}}_{\hat{t}_{ip}} / \hat{t}_{ip}\right)$, and $\mathrm{se}_{\hat{t}_{ip}}$ is the standard error of $\hat{t}_{ip}$ that is calculated by delta method:

$$\widehat{\mathrm{se}}_{\hat{t}_{ip}} = \hat{t}_{ip} \left[\widehat{\mathrm{Var}}(\hat{\mu}_i) + 2\Phi^{-1}(p)\widehat{\mathrm{Cov}}(\hat{\mu}_i, \hat{\sigma}) + [\Phi^{-1}(p)]^2 \widehat{\mathrm{Var}}(\hat{\sigma})\right]^{1/2}, \tag{7}$$

where $\widehat{\mathrm{Var}}(\hat{\mu}_i)$, $\widehat{\mathrm{Var}}(\hat{\sigma})$ and $\widehat{\mathrm{Cov}}(\hat{\mu}_i, \hat{\sigma})$ are the corresponding elements in $\hat{\Sigma}_{\hat{\theta}}$ obtained from the first stage. However, the estimates of η and δ^2 are based on "pseudo sample" $\hat{\mu}_i$s, which means that we cannot calculate the Fisher information matrix of (η, δ^2) exactly in the second stage. Thus, the CIs of η and δ^2 cannot be computed directly by the asymptotic normality theorem. In reliability engineering and quality control, bootstrap has been widely used to construct CI of the quantity of interest, for example, process incapability index [25], failure probability [26] and others [27–29]. In our paper, there are two ways to perform bootstrap resampling.

Bootstrap method I
(1) Obtain the estimate $\hat{\boldsymbol{\theta}}$ using the ML method based on the observed data $\boldsymbol{D}$.
(2) Generate bootstrap sample $\boldsymbol{D}^*$ from the log-location-scale distribution when the parameter vector $\boldsymbol{\theta}$ is replaced by $\hat{\boldsymbol{\theta}}$.
(3) Based on the bootstrap sample $\boldsymbol{D}^*$, obtain bootstrap estimates $\hat{\boldsymbol{\theta}}^*$, (η^*, δ^*) and $\hat{t}^*_{ip}$ by performing the direct two-stage method.
(4) Repeat steps 2 and 3 B times, then we have B bootstrap estimates

$$\left\{ \hat{\Theta}^*_s = \left(\hat{\mu}^*_s, \hat{\sigma}^*_s, \hat{\eta}^*_s, \hat{\delta}^*_s, \hat{t}^*_{ip_s} \right), s = 1, \ldots, B \right\}.$$

Then an approximate $100(1-\alpha)\%$ bootstrap CI for the function of the parameters $G(\Theta)$ is given by $\left(G_{\alpha/2}, G_{1-\alpha/2}\right)$, where $G_{\alpha/2}$ and $G_{1-\alpha/2}$ are the $(\alpha/2)$th and $(1-\alpha/2)$th percentiles of $\left\{G\left(\hat{\Theta}^*_s\right), s = 1, \ldots, B\right\}$.

Bootstrap method II

In this method, the generation of bootstrap sample is a little different from that in the first method. Firstly, based on model (2), generate $\log \mu_i^{(B)}$ from $N(\hat{\eta}, \hat{\delta}^2)$, $i = 1, \ldots, k$. Then, generate bootstrap sample in each batch from $F\left(t; \mu_i^{(B)}, \hat{\sigma}\right)$ in (1). Other steps are the same as these in the first bootstrap method.

We denote the direct two-stage method as the "TS method" hereinafter. Since the two bootstrap methods can lead to similar results in the simulation, we mainly use the bootstrap method I below, denoted as "BS method". When dealing with heavily censored interval data with batch effects, the bootstrap samples generated by the BS method may have zero failures, which may induce parameter identifiability problem in the first stage. The FRW bootstrap method could avoid this problem well, and we will introduce this method in the following section.

3.2. FRW bootstrap

The FRW bootstrap method introduced by Newton and Raftery [30] provides an effective and easy-to-use approach to generate bootstrap samples for various complicated problems. Xu et al. [31] gave a comprehensive review of FRW bootstrap and demonstrated many advantages of this method. Compared with the commonly-used resampling bootstrap procedure, the main idea of the FRW bootstrap suggests selecting a random weight vector $\left(\omega_1, \ldots, \omega_n\right)'$ from a uniform Dirichlet distribution for the n observations rather than integer weights. It can be shown that the fractional weights generated from the uniform Dirichlet distribution are equivalent to generating the standardized random weights from the standard exponential distribution. Let $Z_i, i = 1, \ldots, n$ be a random sample of size n from the standard exponential distribution. Then, the random vector

$$\left(\frac{Z_1}{\sum_{i=1}^n Z_i}, \ldots, \frac{Z_i}{\sum_{i=1}^n Z_i} \ldots, \frac{Z_n}{\sum_{i=1}^n Z_i}\right)' \tag{8}$$

has a uniform Dirichlet distribution. Let $X_1, X_2, \ldots, X_n$ be the n observations. The implementation of the FRW bootstrap has two steps: firstly generate a random weights vector $(Z_1, \ldots, Z_n)$, then obtain the bootstrap estimate of the model parameter based on the random weighted log-likelihood

$$\ell^*(\xi) = \sum_{i=1}^n Z_i \ell_i\left(\xi; X_i\right), \tag{9}$$

where $\ell_i(\xi; X_i)$ is the log-likelihood function for the ith observation, and ξ is a general notation for the unknown parameter vector. It should be noted that the term $\sum_{i=1}^n Z_i$ in (8) is omitted, as it has no effects on the estimation.

Now, we incorporate the FRW bootstrap method into the two-stage method based on interval censored data. The procedure for implementing FRW bootstrap to construct CI is as follows:

1. Generate random weights $Z_{i1}, \ldots, Z_{in_i}$ from standard exponential distribution for the ith batch. Denote that $W_{ij} = \sum_{h=M_{ij-1}+1}^{M_{ij}} Z_{ih}$, $j = 1, \ldots, d_i + 1$. Then the random weighted log-likelihood can be written as

$$\ell^*(\boldsymbol{\theta}) = \sum_{i=1}^k \sum_{j=1}^{d_i+1} W_{ij} \log\left\{\Phi\left[\frac{\log(t_{ij}) - \mu_i}{\sigma}\right] - \Phi\left[\frac{\log(t_{ij-1}) - \mu_i}{\sigma}\right]\right\}. \tag{10}$$

Compared with (3), fractional weights W_{ij} are generated for substituting m_{ij} in each bootstrap iteration. Using the function *optim()* provided in *R* software, the estimates of $\boldsymbol{\theta}$, $\boldsymbol{\theta}^*$, can be obtained easily. The parameter estimates of η, δ^2 and t_{ip} are computed by Eqs. (4) and (5).

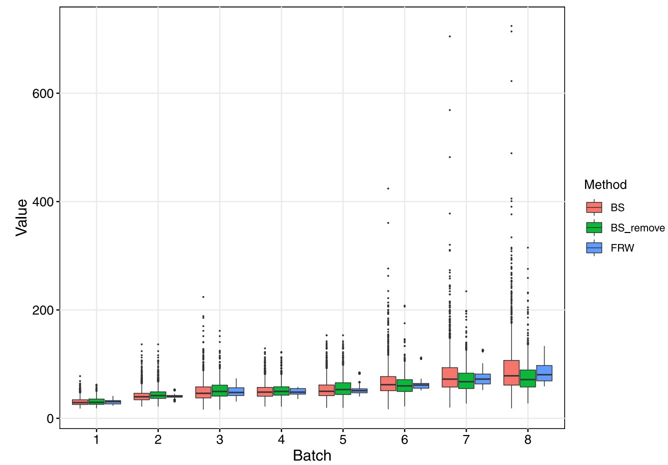

Fig. 2. The results of BS and FRW bootstrap method for the Weibull scale parameters ($\eta = 4, n = 2000$).

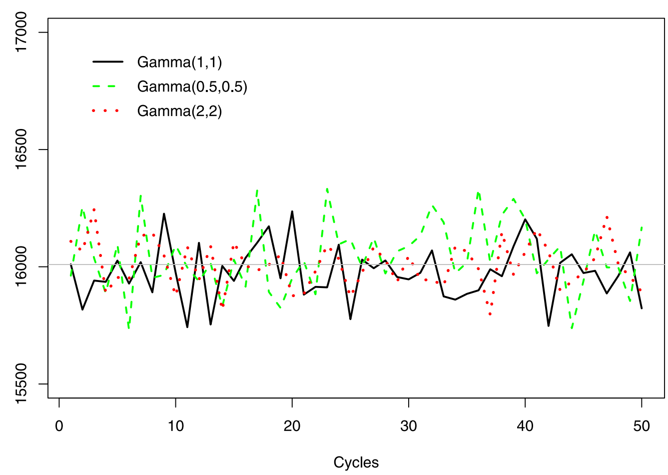

Fig. 3. The values of $\sum_{i=1}^n Z_i$ for different distribution.

2. Repeat the above step B times, and get B bootstrap estimates $\{(\hat{\boldsymbol{\theta}}^{(b)}, \hat{t}_{ip}^{(b)}, \hat{\eta}^{(b)}, \hat{\delta}^{(b)}), b = 1, \ldots, B\}$. Then the point estimates and $100(1-\alpha)\%$ CIs for these parameters can be constructed based on these bootstrap estimates. The detailed procedure is provided in Appendix A.

3.3. Prediction

In this section, we discuss a procedure to give a point prediction as well as the $100(1-\alpha)\%$ prediction interval of the number of product failures in a certain future time. The probability that a product survives at time t_d but has been failed before time t ($t > t_d$) can be formulated as

$$F\left(t|t_d; \boldsymbol{\theta}\right) = \Pr\left(T \leq t | T > t_d\right) = \frac{F(t; \boldsymbol{\theta}) - F\left(t_d; \boldsymbol{\theta}\right)}{1 - F\left(t_d; \boldsymbol{\theta}\right)}, \quad t > t_d. \tag{11}$$

When the future failures in the ith batch are of interest, $F(t; \boldsymbol{\theta})$ takes form of $F\left(t; \mu_i, \sigma\right)$ in (1). While the future failure number in the population is of interest,

$$F(t; \boldsymbol{\theta}) = \int_0^\infty F\left(t; \mu_i, \sigma\right) f_{\mu_i}(\mu_i; \eta, \delta^2) \mathrm{d}\mu_i,$$

where $f_{\mu_i}(\mu_i; \eta, \delta^2)$ is the PDF of μ_i defined in (2). Let n_{t_d} be the number of units that survive until time t_d. Then, a point prediction of the

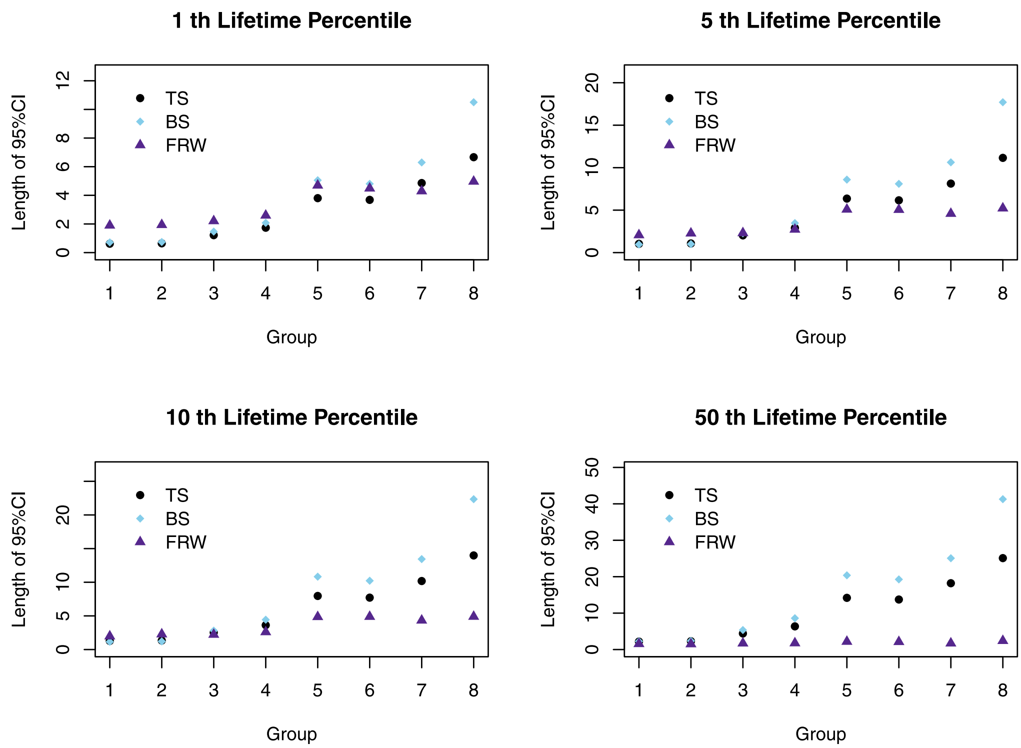

Fig. 4. Length of 95% CI of the 1st, 5th, 10th and 50th percentile based on different method under different batches.

number of product failures at time t is $n_{t_d} F\left(t|t_d;\hat{\theta}\right)$. The $100(1-\alpha)\%$ prediction interval can be constructed as follows:

(1) Based on the FRW bootstrap method in Section 3.2, we obtain B bootstrap estimates $\hat{\theta}_b, b = 1, 2, \ldots, B$.

(2) Compute $U_b = n_{t_d} F(t|t_d;\hat{\theta}_b), b = 1, \ldots, B$.

(3) Let U^l and U^u denote the lower and upper $\alpha/2$ quantiles of $\{U_b, b = 1, \ldots, B\}$, respectively. The $100(1-\alpha)\%$ prediction interval of the number of product failures at time t can be specified as $\left[U^l, U^u\right]$.

3.4. Model selection criteria

There may be several candidate lifetime distributions for fitting the data. We choose the Akaike's information criterion (AIC) for model selection. The AIC is defined as follows:

$$AIC = -2\ell + 2p, \tag{12}$$

where p is the number of parameters in the model, and ℓ is the log-likelihood function. Our goal is to choose the distribution with the smallest AIC value.

4. Simulation study

For illustrative purpose, assume that there are $k = 8$ groups in the experiment, and that the lifetime of the product follows the Weibull distribution. The CDF and PDF of the lifetime T_i of the ith group are respective

$$F(t;\gamma_i,\omega) = 1 - \exp\left[-\left(\frac{t}{\gamma_i}\right)^{\omega}\right] \text{ and } f(t;\gamma_i,\omega) = \left(\frac{\omega}{\gamma_i}\right)\left(\frac{t}{\gamma_i}\right)^{\omega-1}\exp\left[-\left(\frac{t}{\gamma_i}\right)^{\omega}\right],$$

where $\gamma_i = \exp(\mu_i)$ is the scale parameter, and $\omega = 1/\sigma$ is the shape parameter. The batch effects are described by (2), that is, $\ln\mu_i \sim N(\eta, \delta^2)$. The model parameters are ω, δ^2 and η. We set $\omega = 0.8$ and 3 corresponding to decreasing and increasing failure rate, respectively. For the case of $\omega = 0.8$, we choose $\eta = 5$ and 8, and then there are 90% and 99% of the units are censored. When $\omega = 3$, η is set to be 3 and 4 with the censoring rate 85% and 99%, respectively. For all the cases, $\delta^2 = 0.1$. Without loss of generality, the sample size in each group is assumed to be the same, that is, $n_1 = \cdots = n_k = n$. We choose $n = 500, 1000$ and 2000. The inspection time epochs are $(t_1, t_2, \ldots, t_m) = (1, 2, \ldots, 10)$. Three two-stage methods (TS, BS and FRW) are compared for each combination of (ω, η, n). The interval estimates are constructed based on 1000 bootstrap samples for the BS and FRW methods. The parameters of interest are the percentile lifetimes t_p in each group, where $p = 0.01, 0.05, 0.1$ and 0.5. Except for this, we are also interested in the scale parameters $\gamma_1, \ldots, \gamma_8$ and shape parameter ω, and the parameters η and δ^2 for batch effects. For the three two-stage methods, the relative bias (RB) and the root mean square error (RMSE) of the estimates of these parameters, the length of the 95% confidence interval (LCI) are computed based on 1000 random samples. RB and RMSE for the parameter ν based on 1000 repetitions are defined as follows:

$$\mathrm{RB}(\hat{\nu}) = \frac{1}{1000}\sum_{i=1}^{1000}\frac{|\hat{\nu}_i - \nu|}{\nu}, \quad \mathrm{RMSE}(\hat{\nu}) = \left(\frac{1}{1000}\sum_{i=1}^{1000}\left(\hat{\nu}_i - \nu\right)^2\right)^{1/2}. \tag{13}$$

Since the results based on $\omega = 0.8$ and 3 are similar, we just list the results of the percentiles for $\omega = 3$ in Tables 3–5, where RB and RMSE are multiplied by 100. The others simulation results can be found in the supplementary. Some conclusions can be clearly summarized from these tables.

(1) For the heavily censored cases ($\omega = 3, \eta = 3$), the RB, RMSE and LCI based on all the three methods become better and better as the sample size increases. However, the RB and RMSE values based on the FRW bootstrap method are the smallest for nearly all cases, and thus FRW bootstrap method has the best performance. Among the three methods, BS method performs the worst, since some abnormal values exist in the 1000 estimates of the percentile lifetimes which lead to large RBs and RMSEs. We will explain the drawback of BS method via simulation later.

(2) For the extremely heavily censored cases ($\omega = 3, \eta = 4$), the superiority of the FRW bootstrap method become more significant,

Table 3
RB (in percentile) of the estimates of percentiles for $\omega = 3$.

(η, n)	Group	$t_{0.01}$			$t_{0.05}$			$t_{0.1}$			$t_{0.5}$		
		TS	BS	FRW	TS	BS	FRW	TS	BS	FRW	TS	BS	FRW
(3500)	1	4.165	7.764	1.608	2.932	8.869	1.269	2.497	9.089	0.966	2.199	9.581	0.594
	2	3.969	4.909	1.421	2.820	4.707	0.942	2.504	4.936	0.775	2.816	5.981	0.337
	3	4.093	4.534	1.458	3.233	4.184	0.908	2.967	4.173	0.647	2.973	5.200	0.213
	4	4.621	4.584	1.491	4.049	3.841	0.936	3.954	3.797	0.689	4.244	4.385	0.229
	5	3.889	4.079	1.323	3.655	3.875	0.958	3.721	3.978	0.689	4.520	5.120	0.215
	6	5.327	5.433	1.388	5.177	5.270	0.903	5.261	5.437	0.696	5.923	6.327	0.188
	7	5.894	6.236	1.368	5.661	5.930	0.891	5.776	6.126	0.680	6.482	7.051	0.186
	8	5.733	8.174	1.391	5.689	8.205	0.922	5.841	8.399	0.692	6.607	9.273	0.167
(3,1000)	1	2.929	7.734	1.532	2.027	8.256	1.131	1.767	8.520	0.923	1.583	9.114	0.581
	2	2.939	4.653	1.335	2.101	4.650	0.909	1.901	4.740	0.734	2.023	5.787	0.230
	3	2.984	3.775	1.349	2.249	3.611	0.805	2.104	3.781	0.611	2.725	4.820	0.182
	4	2.909	3.703	1.305	2.544	3.124	0.872	2.497	3.228	0.663	2.978	4.130	0.209
	5	3.171	3.921	1.298	2.785	3.173	0.851	2.820	3.069	0.671	3.315	3.910	0.174
	6	3.668	3.902	1.304	3.099	3.229	0.865	3.000	3.300	0.675	3.246	4.023	0.196
	7	3.683	3.948	1.311	3.787	3.739	0.866	3.941	3.917	0.671	4.777	4.996	0.183
	8	5.481	6.555	1.289	5.159	6.171	0.856	5.159	6.223	0.662	5.461	6.857	0.155
(3,2000)	1	2.266	7.029	1.501	1.596	7.829	1.120	1.354	8.216	0.907	1.117	9.003	0.499
	2	2.071	4.088	1.330	1.486	4.084	0.900	1.332	4.192	0.714	1.297	4.818	0.203
	3	2.049	2.908	1.301	1.544	2.750	0.801	1.395	2.815	0.587	1.519	3.646	0.162
	4	2.221	3.463	1.272	1.890	3.152	0.833	1.791	3.158	0.640	1.918	3.651	0.170
	5	2.221	2.857	1.273	1.997	2.562	0.847	1.947	2.573	0.661	2.082	2.953	0.167
	6	2.305	2.722	1.296	2.079	2.369	0.854	2.106	2.406	0.666	2.474	3.232	0.184
	7	3.090	3.176	1.244	2.971	3.127	0.807	2.994	3.299	0.616	3.334	4.215	0.147
	8	3.711	4.155	1.247	3.764	4.122	0.808	3.839	4.208	0.620	4.197	4.839	0.127
(4500)	1	7.539	1032.274	2.921	6.802	898.27	1.563	7.677	845.866	1.279	14.199	724.509	0.639
	2	10.870	48.482	2.925	8.744	38.798	1.568	9.792	36.551	1.285	16.558	37.033	0.403
	3	13.217	1.95E+24	2.948	15.918	7.21E+24	1.589	17.588	6.55E+24	1.256	23.112	5.09E+24	0.429
	4	25.809	1.86E+12	2.938	26.522	1.55E+12	1.573	27.094	1.43E+12	1.186	30.632	1.16E+12	0.388
	5	17.609	5.27E+10	2.982	21.868	4.16E+10	1.615	23.857	3.75E+10	1.127	29.341	2.87E+10	0.387
	6	32.080	8.18E+48	2.919	31.737	5.84E+48	1.558	32.184	5.03E+48	1.172	34.808	3.41E+48	0.366
	7	33.419	7.99E+57	2.960	32.813	5.71E+57	1.601	32.750	4.92E+57	1.116	33.926	3.33E+57	0.371
	8	23.564	8.18E+38	2.938	27.159	1.82E+38	1.578	28.887	1.60E+38	1.194	33.529	1.14E+38	0.351
(4,1000)	1	6.414	12.169	2.761	6.108	10.748	1.356	6.935	12.228	1.113	10.446	17.696	0.496
	2	7.124	39.128	2.712	7.599	36.803	1.300	8.545	35.145	1.056	12.125	32.282	0.389
	3	10.190	8.96E+19	2.736	11.523	1.39E+19	1.319	12.635	1.23E+19	1.075	16.209	8.13E+19	0.383
	4	12.262	1.76E+08	2.755	13.312	1.28E+08	1.345	14.151	1.20E+08	1.102	17.331	1.26E+08	0.344
	5	15.812	1.92E+07	2.758	16.472	1.61E+07	1.337	17.326	1.50E+07	1.091	19.942	1.61E+07	0.373
	6	16.267	3.62E+09	2.741	16.740	3.26E+09	1.325	17.246	2.47E+09	1.082	19.538	1.78E+09	0.328
	7	21.993	2.21E+31	2.756	22.122	1.63E+31	1.341	22.591	1.43E+31	1.096	24.584	1.00E+31	0.353
	8	22.448	5.15E+27	2.749	22.307	3.81E+27	1.331	22.418	3.33E+27	1.085	23.211	2.12E+27	0.340
(4,2000)	1	4.383	4.868	2.094	4.059	4.329	1.160	4.680	4.902	0.821	7.444	7.570	0.432
	2	5.306	6.117	2.090	5.901	6.424	1.155	6.721	7.216	0.814	9.691	10.120	0.218
	3	6.484	7.958	2.096	7.586	8.779	1.159	8.471	9.580	0.816	11.322	12.304	0.218
	4	8.438	11.135	2.085	9.204	11.536	1.161	9.811	12.069	0.818	12.256	14.132	0.234
	5	9.129	2.65E+05	2.078	10.116	2.48E+05	1.163	10.846	2.40E+05	0.814	13.335	2.22E+05	0.219
	6	10.409	108.714	2.099	11.386	92.575	1.167	12.138	86.670	0.819	14.564	74.311	0.206
	7	14.724	9.58E+05	2.106	15.320	7.60E+05	1.168	15.933	6.86E+05	0.825	18.082	5.25E+05	0.229
	8	16.686	2.78E+08	2.098	17.182	2.09E+08	1.160	17.620	1.84E+08	0.817	19.127	1.32E+08	0.211

namely we can still obtain reasonable RB, RMSE and LCI of the estimates by the FRW bootstrap method, while BS method almost fails to deal with extremely heavily censored data, and TS method produces some abnormal values of RMSE and LCI of the parameters $\gamma_1, \ldots, \gamma_8$.

From the above analysis, we see that the FRW bootstrap is effective for constructing interval estimation when the data is heavily censored, in which case BS method usually fails. For the BS method, conventional bootstrap is used in the resampling step, which may induce zero-failure in some batches in the bootstrap samples. The probability of producing zero-failure cases will increase as the censoring rate increases. Then the ML estimates of the scale parameters in these zero-failure batches will diverge to infinity, which explains why there are many abnormal bootstrap estimates using BS method. While for FRW bootstrap method, fractional random weights are generated for the observations, which avoids the zero-failure case and the ML estimates would exist if the original data are reasonable. For illustration, we generate an interval failure data when $\eta = 4, n = 2000$, then record the bootstrap estimates of the scale parameters for BS and FRW bootstrap methods. The estimates are shown in Fig. 2, where "BS_remove" denotes the bootstrap estimates when the zero-failure cases are removed. As can be seen in Fig. 2, the variation among bootstrap estimates has been greatly reduced after removing zero-failure cases. However, its performance is still worse than FRW bootstrap method. For FRW bootstrap method, the variations among bootstrap estimates are induced by random weights for the original observations. We sum these weights, and generate 50 random samples of the summation, which are shown in Fig. 3, where the results based on the two alternative distributions of Z_i, $Gamma(0.5, 0.5)$ and $Gamma(2, 2)$, are also listed. From Fig. 3, we see that the values of $\sum_{i=1}^{n} Z_i$ for different distributions are all stable around 16000 (since the total sample size $\sum_{i=1}^{8} n_i = 16000$). Thus, the variations among the bootstrap estimates that are based on the weighted likelihood functions will be small.

5. Real data analysis

In this section, we will analyze the interval failure data of PCB in Table 1, and provide the solutions of the three problems from the manufacturer.

Table 4
RMSE (in percentile) of the estimates of percentiles for $\omega = 3$.

(η, n)	Group	$t_{0.01}$			$t_{0.05}$			$t_{0.1}$			$t_{0.5}$		
		TS	BS	FRW	TS	BS	FRW	TS	BS	FRW	TS	BS	FRW
(3500)	1	16.812	28.677	5.354	20.109	49.251	6.690	21.625	65.588	7.222	34.624	134.980	8.515
	2	17.709	21.855	5.734	22.376	34.434	6.636	25.753	44.384	6.679	52.950	99.651	5.418
	3	19.554	21.354	6.469	25.123	32.616	7.414	29.115	41.760	7.399	59.758	95.619	4.085
	4	24.565	24.400	7.206	36.838	34.868	8.271	46.250	43.832	8.235	98.893	99.423	5.119
	5	24.788	26.287	7.969	39.205	42.125	9.127	50.340	55.019	9.080	111.016	126.985	4.478
	6	36.018	36.740	8.729	61.462	63.298	9.865	79.984	83.356	9.713	169.601	184.031	5.089
	7	47.900	53.457	10.109	79.505	88.960	12.558	101.839	114.428	11.466	206.466	235.118	5.482
	8	81.566	99.233	12.470	140.030	167.048	14.313	199.480	213.114	14.233	362.683	518.896	5.751
(3,1000)	1	10.546	28.044	5.059	12.278	47.913	6.154	13.559	65.197	6.594	27.340	134.322	8.086
	2	13.112	19.179	5.497	16.154	32.364	6.417	18.387	42.811	6.506	37.816	96.580	4.399
	3	14.681	18.389	5.847	19.991	28.953	6.553	24.380	38.135	6.357	54.664	90.247	3.449
	4	16.141	19.336	7.024	25.129	27.935	8.165	32.196	36.134	8.181	71.838	87.113	4.931
	5	19.239	22.846	7.443	28.329	31.649	8.501	35.873	39.309	8.383	79.581	91.314	3.990
	6	25.909	27.214	8.497	39.606	40.412	9.827	49.770	51.719	9.685	103.351	119.614	4.903
	7	29.431	32.178	9.988	52.925	51.873	11.572	71.526	68.755	11.098	163.776	163.742	5.248
	8	65.402	97.082	11.830	112.230	165.401	13.768	144.956	212.209	13.708	294.220	424.056	5.661
(3,2000)	1	7.816	24.066	4.923	9.674	44.993	6.132	10.726	59.637	6.525	18.457	125.569	6.749
	2	8.440	16.810	5.090	10.766	28.285	5.949	12.281	36.813	5.985	22.732	79.521	3.678
	3	9.989	13.232	5.467	13.588	21.113	6.311	16.079	27.592	6.302	31.670	65.216	3.880
	4	11.519	17.492	6.132	16.485	26.455	6.955	20.184	33.367	6.833	40.457	72.493	3.317
	5	13.802	15.959	6.771	20.867	23.504	7.756	25.891	30.107	7.685	51.062	69.279	4.230
	6	14.792	16.654	7.618	23.598	26.617	8.738	30.130	35.315	8.662	63.763	84.944	4.418
	7	23.426	22.536	8.606	39.994	39.673	9.734	51.675	53.268	9.534	105.993	123.304	4.305
	8	37.019	40.519	10.714	64.668	70.043	12.065	83.755	91.298	11.752	169.418	192.563	4.423
(4500)	1	85.514	7.51E+04	23.902	103.692	1.13E+05	24.818	152.510	1.35E+05	21.658	595.631	2.15E+05	9.083
	2	113.459	3247.502	28.579	174.787	3247.502	26.177	257.597	4877.670	25.111	844.262	1797.478	9.621
	3	151.468	1.48E+26	32.478	297.723	2.05E+26	30.608	419.749	2.37E+26	27.776	1113.122	3.46E+26	9.893
	4	431.158	1.35E+14	36.261	735.363	1.85E+14	34.908	945.576	5.49E+14	31.846	2038.268	8.06E+14	11.119
	5	290.264	2.98E+06	42.250	693.283	4.23E+06	39.122	998.554	4.82E+06	34.628	2489.260	7.52E+06	10.996
	6	394.037	1.87E+50	47.639	819.127	5.02E+50	43.724	1124.826	2.52E+50	39.163	2652.087	8.45E+50	11.484
	7	550.581	2.37E+59	54.774	1004.276	1.51E+59	50.841	1487.194	3.19E+59	45.425	2772.735	2.54E+59	13.644
	8	571.501	6.84E+40	64.439	997.195	6.56E+40	60.025	1350.057	9.83E+40	54.193	2706.635	1.54E+40	15.643
(4,1000)	1	64.006	111.309	23.506	86.227	196.854	22.563	144.815	282.178	19.356	383.607	737.634	8.563
	2	77.717	466.069	28.324	136.136	734.391	25.611	191.828	915.611	23.186	502.859	1165.321	8.587
	3	120.243	1.86E+21	31.421	227.440	2.29E+21	30.314	315.299	2.51E+21	27.117	762.027	3.19E+21	8.618
	4	186.127	1.92E+09	35.646	344.470	2.36E+09	33.465	463.456	2.58E+09	29.516	1035.917	3.28E+09	9.297
	5	308.866	1.57E+04	39.088	517.843	5.33E+04	38.101	663.749	6.17E+04	34.603	1327.124	7.07E+04	10.807
	6	315.282	3.84E+11	44.297	519.007	5.06E+11	43.175	663.982	7.40E+11	38.530	1334.708	7.76E+11	11.156
	7	499.859	3.88E+33	50.924	845.098	5.27E+33	49.873	1078.445	6.03E+33	45.069	2102.923	8.59E+33	12.119
	8	518.472	7.30E+28	61.155	957.540	9.29E+28	59.672	1224.020	1.03E+28	53.322	2408.467	1.37E+28	13.974
(4,2000)	1	44.911	49.574	19.542	81.480	86.558	19.852	117.113	123.385	18.548	311.646	326.115	6.216
	2	71.657	82.382	23.387	117.234	127.385	23.523	159.671	169.685	21.774	396.069	409.890	7.094
	3	109.539	134.098	26.740	206.017	240.838	26.993	278.849	320.294	25.076	626.807	697.464	7.193
	4	124.206	155.031	29.312	231.091	276.877	29.426	311.238	366.593	27.179	695.442	789.840	8.618
	5	148.842	195.125	32.462	289.940	355.226	32.697	392.823	470.115	30.276	871.425	997.190	9.393
	6	240.340	475.433	35.888	419.032	758.710	36.041	542.855	942.267	33.269	1096.862	1710.943	10.923
	7	265.625	448.096	41.115	470.890	724.329	41.353	616.690	906.790	38.199	1292.458	1698.005	9.731
	8	458.598	1009.656	50.194	802.434	1605.973	50.359	1033.625	1982.139	46.429	2035.657	3496.345	11.257

For the first problem, we use the model in Section 2 to fit the data, and choose exponential, Weibull, lognormal and loglogistic distributions as the candidate models. The AIC values for the four distributions are listed in Table 6, from which we can see that Weibull distribution has the smallest AIC value. Thus, we choose Weibull distribution to fit the real data.

For the second problem, three two-stage methods are utilized to obtain the point and interval estimates of the model parameters as well as the percentile lifetimes of each batch. CIs are constructed based on 1000 bootstrap samples, and the results are listed in Table 7. As can be see in Table 7, the point estimates of the parameters based on the three methods are close to each other, while LCI based on FRW method are the shortest that is consistent to the simulation results. For estimating the percentile lifetimes of each batch, the results are similar, for which we draw these lengths in Fig. 4. After obtaining the estimates of the parameters, the cumulative number of failures at different time intervals in each batch are also fitted, and the 95% point-wise CIs of the cumulative number of failures are computed. The results are shown in Fig. 5, where the ith subfigure shows the results of the ith batch. As can be seen in each subfigure, the fitted cumulative numbers of failures by the proposed model are close to the observed data, and almost all the observed data lie in the 95% point-wise CIs, which indicates that the model fits the data well.

For the last problem, we fit the model by using interval failure data before 9 weeks, and predict the number of failures in the period of the tenth week. Then the prediction result can be well assessed because the true number of failures is known in the time interval. As a comparison, the prediction based on the model without batch effects and the model for each batch are also computed, which can be conducted by pooling the data together and separate batch to make statistical inference, respectively. Table 8 lists the points prediction and 95% CI of the number of failures in the tenth weeks based on the three methods as well as the true values for each batch. From Table 8, we can see that the prediction based on the model without batch effects performs badly, which disperses the true values widely. The point predictions between separate batch and considering batch effects are very close. If we are interested in the number of failures in the population, there are actually 239 units failing in the tenth week.

Table 5
LCI of the estimates of percentiles for $\omega = 3$.

(η, n)	Group	$t_{0.01}$			$t_{0.05}$			$t_{0.1}$			$t_{0.5}$		
		TS	BS	FRW	TS	BS	FRW	TS	BS	FRW	TS	BS	FRW
(3500)	1	0.530	0.731	2.003	0.910	0.962	2.321	1.152	1.102	2.358	2.105	1.897	2.163
	2	0.763	0.909	2.337	1.312	1.296	2.703	1.663	1.559	2.735	3.046	2.946	2.295
	3	0.948	1.107	2.575	1.630	1.674	2.967	2.066	2.068	2.971	3.790	4.027	2.317
	4	1.277	1.438	2.894	2.202	2.295	3.321	2.795	2.889	3.310	5.154	5.699	2.316
	5	1.405	1.735	3.191	2.421	2.834	3.643	3.072	3.588	3.611	5.658	7.078	2.360
	6	1.846	2.239	3.507	3.186	3.752	3.997	4.048	4.783	3.945	7.492	9.442	2.510
	7	2.441	3.033	3.894	4.206	5.150	4.444	5.341	6.574	4.379	9.893	12.844	2.583
	8	4.102	5.387	4.662	7.072	9.244	5.288	8.986	11.801	5.185	16.725	22.738	2.694
(3,1000)	1	0.383	0.515	1.667	0.660	0.681	1.958	0.836	0.783	2.012	1.533	1.355	2.040
	2	0.537	0.641	1.924	0.921	0.916	2.226	1.167	1.104	2.229	2.134	2.086	2.016
	3	0.683	0.793	2.175	1.174	1.205	2.503	1.487	1.491	2.530	2.725	2.904	2.183
	4	0.892	1.014	2.429	1.535	1.620	2.788	1.948	2.041	2.772	3.584	4.022	2.215
	5	1.008	1.213	2.640	1.734	1.984	3.012	2.199	2.517	2.991	4.043	4.975	2.229
	6	1.364	1.611	2.971	2.345	2.703	3.388	2.974	3.444	3.346	5.485	6.776	2.231
	7	1.908	2.242	3.393	3.287	3.825	3.850	4.175	4.889	3.777	7.740	9.557	2.402
	8	3.989	4.218	4.042	6.938	7.283	4.555	8.853	9.314	4.461	16.693	18.009	2.516
(3,2000)	1	0.226	0.332	1.418	0.387	0.432	1.664	0.489	0.491	1.698	0.891	0.826	1.653
	2	0.327	0.418	1.641	0.560	0.592	1.879	0.709	0.710	1.895	1.292	1.327	1.611
	3	0.410	0.506	1.848	0.702	0.761	2.121	0.889	0.938	2.125	1.623	1.816	1.731
	4	0.503	0.617	2.049	0.863	0.967	2.340	1.093	1.210	2.335	1.998	2.366	1.742
	5	0.579	0.748	2.228	0.993	1.207	2.541	1.257	1.521	2.511	2.298	2.980	1.806
	6	0.782	0.973	2.501	1.342	1.614	2.845	1.701	2.051	2.807	3.126	4.022	1.786
	7	1.153	1.377	2.873	1.983	2.335	3.254	2.516	2.981	3.195	4.652	5.812	1.895
	8	2.082	2.357	3.550	3.577	4.045	4.014	4.539	5.167	3.931	8.421	9.949	2.103
(4500)	1	9.691	10.238	5.614	16.919	17.606	6.113	21.614	22.702	5.885	40.484	45.702	2.948
	2	19.098	17.591	6.751	31.845	29.225	7.412	40.840	38.986	7.170	77.541	74.774	3.346
	3	24.306	30.799	7.723	42.413	50.809	8.431	54.194	63.445	8.120	102.032	119.924	3.703
	4	60.302	114.263	8.575	100.450	174.005	9.331	126.065	210.300	8.942	228.683	350.513	3.875
	5	59.157	64.379	9.843	106.687	109.534	10.787	138.539	138.512	10.381	274.026	260.188	4.022
	6	187.614	1602.767	11.187	305.396	2112.709	12.186	379.704	2438.594	11.707	675.157	3464.484	4.427
	7	259.670	2101.317	12.691	455.788	2931.135	13.808	585.820	3404.315	13.231	1135.564	5200.364	4.700
	8	125.401	214.168	14.515	229.545	337.629	15.829	299.915	417.646	15.186	603.088	730.785	5.263
(4,1000)	1	8.281	7.111	5.120	13.961	12.191	5.640	17.536	16.621	5.468	31.288	32.176	2.898
	2	18.305	11.792	6.142	25.474	19.072	6.750	38.907	25.659	6.520	66.202	48.246	3.139
	3	19.944	26.141	6.816	34.485	41.215	7.484	43.867	50.324	7.227	81.626	89.055	3.300
	4	19.282	20.026	7.669	32.864	34.436	8.391	41.525	44.673	8.066	75.735	85.130	3.460
	5	34.410	29.929	8.556	58.202	50.398	9.387	73.357	63.714	9.048	133.545	118.981	3.774
	6	35.236	47.379	9.505	59.016	77.214	10.409	74.076	96.874	9.992	133.446	175.678	3.927
	7	51.451	67.043	10.700	86.704	107.781	11.698	109.109	133.810	11.223	197.986	234.431	4.191
	8	67.436	104.400	12.813	113.305	166.393	13.990	142.450	205.386	13.412	258.272	358.464	4.798
(4,2000)	1	5.712	5.128	4.607	9.818	8.798	5.102	12.443	11.367	4.961	22.797	23.227	2.752
	2	7.688	7.226	5.463	13.247	12.577	6.032	16.813	16.270	5.844	30.977	32.929	2.928
	3	9.712	9.680	6.207	16.757	16.895	6.830	21.286	21.817	6.603	39.368	43.642	3.087
	4	12.400	12.652	6.867	21.325	21.983	7.555	27.052	28.277	7.292	49.914	55.735	3.237
	5	13.466	15.969	7.534	23.103	27.642	8.285	29.278	35.449	7.987	53.847	69.148	3.396
	6	16.606	20.659	8.321	28.522	35.513	9.142	36.172	45.349	8.799	66.757	87.272	3.577
	7	31.686	32.721	9.417	54.462	55.456	10.326	69.135	70.293	9.932	128.415	132.198	3.858
	8	38.053	46.835	11.473	65.012	78.407	12.572	82.323	98.787	12.057	152.036	182.604	4.406

Only 152 failures are reported if no batch effects are assumed. For the separate batch method, there are 255 units failing in the tenth week, which is close to 249 predicted by the proposed method. However, the 95% prediction CI with considering batch effects is [171, 325], which is much shorter than [134, 380] by separate batch method. In the batch effects model, a common scale parameter is shared for each batch, and statistical strength shared by the population can be leverage to improve the prediction for each batch. Thus, the precision of prediction can be greatly improved by considering batch effects in the model.

In fact, making accurate future failure numbers prediction benefits the reduction of inventory cost. When products fail under warranty coverage, customers return their products for repair or replacement. For manufacturers, it is necessary to prepare a certain amount of inventory to deal with product maintenance in a certain future period (e.g., next month). The amount of inventory can be easily determined with point prediction or upper bound of 95% CI as described in Table 8. Inventory cost would be greatly saved with the proposed model that has the precise prediction of failure numbers. The proposed model can also be used to estimate field reliability, and to determine whether it is necessary to recall products from certain batches that have critical failure modes. For example, the reliability of PCB from the first two batches in the real case is significantly smaller than the other batches. Then careful detection should be implemented for these failures. If the high failure probability is only due to materials issues, then some targeted solutions can be presented to solve these problems. While, if critical failure modes have been detected, these PCBs would be recalled from the market.

6. Conclusion and discussion

In this paper, we have proposed an improved two-stage method combined with FRW bootstrap to analyze the interval failure data, focusing on the heavily censored data with batch effects. We compare the proposed method with the other two alternatives through simulation studies. The results show that the proposed method demonstrates better performance in terms of RB, RMSE and LCI, especially for extremely

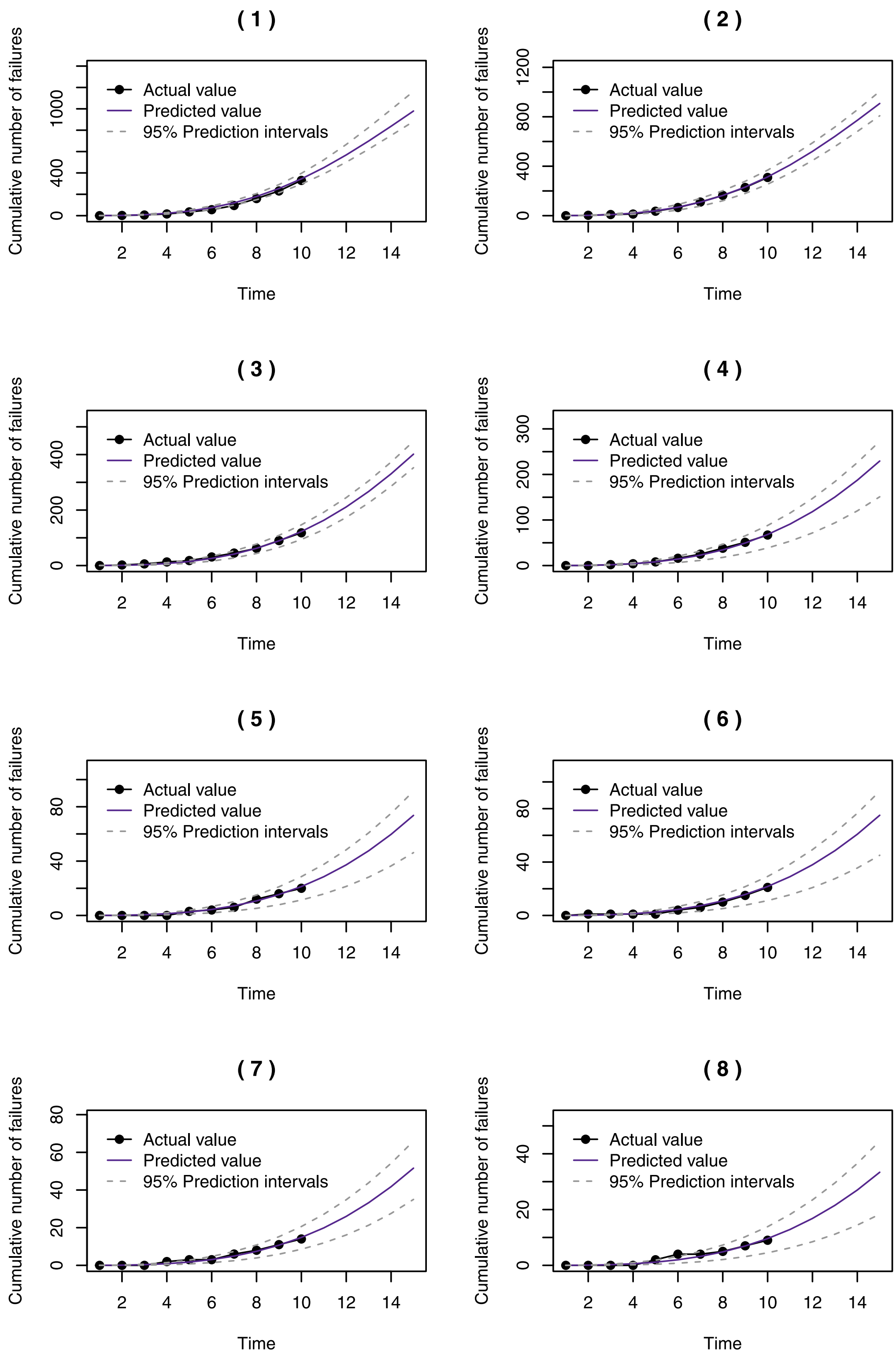

Fig. 5. The fitted cumulative number of failure for each batch based on the proposed model.

heavily censored data. In the data analysis, we utilize AIC to choose the Weibull distribution as the best distribution among four candidates. We use the proposed model and the other two models to predict the number of failures in a certain time period, and find that the predictions based on the proposed model are close to the true values, while the model without batch effects performs badly.

The data considered in this paper are interval censored with batch effects. There are several possible future research directions worthing studying. Firstly, due to the limitation of test conditions, the units may be tested in several chambers, and thus the interval failure data may have both subsamlping and block effects. Secondly, for reducing the experiment time, ALT is a commonly used way in practice. Then covariates should be incorporated into the model. Interval failure data

Table 6
AIC for the different distributions.

Distribution	AIC
Exponential	5102.53
Weibull	3581.65
Lognormal	3981.28
Loglogistic	3860.94

Table 7
The estimates of the parameters based on different methods for the real dataset.

	TS			BS			FRW		
θ	2.5%	97.5%	Mean	2.5%	97.5%	Mean	2.5%	97.5%	Mean
γ_1	15.866	18.579	17.222	16.031	18.618	17.256	16.007	18.356	17.220
γ_2	15.958	18.712	17.335	16.154	18.806	17.398	16.048	18.371	17.326
γ_3	20.907	26.628	23.768	21.177	27.095	23.884	22.323	25.177	23.769
γ_4	24.305	32.925	28.615	24.73	33.772	28.835	26.998	29.667	28.586
γ_5	31.737	51.223	41.480	32.895	54.629	42.061	40.373	42.823	41.487
γ_6	32.151	52.667	42.409	33.409	56.099	42.945	41.066	43.893	42.418
γ_7	33.991	59.462	46.727	35.962	65.311	47.889	45.615	48.255	46.761
γ_8	36.244	72.445	54.345	39.568	83.829	56.264	53.045	55.589	54.346
ω	2.860	3.448	3.154	2.966	3.353	3.155	2.733	3.731	3.171
η				3.348	3.558	3.448	3.423	3.458	3.441
β				0.150	0.321	0.219	0.186	0.231	0.205

Table 8
Point prediction and 95% CI of the number of failures in the tenth weeks for each batch based on different models.

Model		Batch								Total
		1	2	3	4	5	6	7	8	
Separate batch	2.50%	52	58	16	6	1	1	0	0	134
	Mean	100	86	31	19	6	6	4	3	255
	97.50%	141	120	48	30	12	13	10	6	380
Without batch effects	2.50%	7	7	7	7	7	7	7	7	56
	Mean	17	17	17	17	17	17	17	17	136
	97.50%	33	33	33	33	33	33	33	33	264
With batch effects	2.50%	67	67	19	10	3	2	2	1	171
	Mean	98	84	29	18	6	6	5	3	249
	97.50%	120	107	41	24	10	10	8	5	325
	True	99	81	28	16	4	6	3	2	239

with subsamlping and block effects based on ALT needs to be considered in this situation. Finally, for some repairable systems, the data is recurrent for each system. When the interval recurrent data have subsamlping and block effects, finding a suitable model or statistical inference method is more challenging.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Natural Science Foundation of China under Grant 11671303 and 71671130, the Natural Science Foundation of Zhejiang Province, China under Grant LY20G010014, the Foundation of Zhejiang Educational Committee of China under Grant Y202044238 and the Foundation of Wenzhou University, China Graduate Innovation under Grant 3162019061.

All authors approved the version of the manuscript to be published.

Appendix A

Algorithm 1: FRW bootstrap algorithm

Input: Observation data $\mathcal{D}$.
Output: The point estimates and $100(1-\alpha)\%$ CIs for parameters $(\hat{\theta}, \hat{t}_{ip}, \hat{\eta}, \hat{\delta})$.

for b *in* $\{1, 2, \ldots, B\}$ **do**
- Generate random weights $Z_{i1}, \ldots, Z_{in_i}, i = 1, \ldots, k$ from standard exponential distribution;
- Denote M_{i0}, M_{ij} and W_{ij};
- Obtain $\hat{\theta}^{(b)}$, by Eq. (10);
- Calculate $\hat{t}_{ip}^{(b)}, \hat{\eta}^{(b)}, \hat{\delta}^{(b)}$, by Eqs. (4) and (5);

end
Calculate the point estimates and $100(1-\alpha)\%$ CIs for these parameters.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ress.2021.107622.

References

[1] Meeker WQ, Escobar LA. Statistical methods for reliability data. John Wiley & Sons; 1998.
[2] Feiveson AH, Kulkarni PM. Reliability of space-shuttle pressure vessels with random batch effects. Technometrics 2000;42(4):332–44.
[3] León RV, Li Y, Guess FM, Sawhney RS. Effect of not having homogeneous test units in accelerated life tests. J Qual Technol 2009;41(3):241–6.
[4] Lv SS, He Z, Vining GG. Simultaneous optimization of quality and reliability characteristics through designed experiment. Qual Eng 2015;29(3):344–57.
[5] Freeman L, Vining G. Reliability data analysis for life test experiments with subsampling. J Qual Technol 2010;42(3):233–41.
[6] Vining GG, Freeman LJ, Kensler JLK. An overview of designing experiments for reliability data. Springer International Publishing; 2015.
[7] Zhang L, Xie M, Tang L. Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data. Reliab Eng Syst Saf 2006;91(8):930–9.
[8] Wang GD, Niu ZW, Lv SS, Qu L, He Z. Bootstrapping analysis of lifetime data with subsampling. Qual Reliab Eng Int 2016;32:1945–53.
[9] Wang G, Niu Z, Lv S, Qu L, He Z. Calculating confidence intervals for percentiles of accelerated life tests with subsampling. Qual Technol Quant Manag 2018;16(4):424–38.
[10] Kensler J, Freeman L, Vining G. A practitioner's guide to analyzing reliability experiments with random blocks and subsampling. Qual Eng 2014;26(3):359–69.
[11] Wang G, He Z, Xue L, Cui Q, Lv S, Zhou P. Bootstrap analysis of designed experiments for reliability improvement with a non-constant scale parameter. Reliab Eng Syst Saf 2017;160:114–21.
[12] Lv S, Wang Z, He Z, Vining G. Impact of censoring types on the two-stage method for analyzing reliability experiments with random effects. Qual Eng 2019;31(2):302–13.
[13] Freeman L, Vining G. Reliability data analysis for life test designed experiments with sub-sampling. Qual Reliab Eng Int 2013;29(4):509–19.
[14] Lv SS, Niu ZW, Qu L, He SG, He Z. Reliability modeling of accelerated life tests with both random effects and nonconstant shape parameters. Qual Eng 2015;27:329–40.
[15] Kensler JL, Freeman LJ, Vining GG. Analyzing reliability experiments with random blocks and subsampling. J Qual Technol 2015;47:235–51.
[16] Seo K, Rong P. Data analysis for accelerated life tests with constrained randomization. In: Reliability & maintainability symposium. 2016.
[17] Seo K, Pan R. Data analysis of step-stress accelerated life tests with heterogeneous group effects. IISE Trans 2017;49(9):885–98.
[18] Medlin RM, Freeman LJ, Kensler JL, Vining GG. Analysis of split-plot reliability experiments with subsampling. Qual Reliab Eng Int 2019;35(3):738–49.
[19] Zhu T. Reliability estimation for two-parameter Weibull distribution under block censoring. Reliab Eng Syst Saf 2020;203:107071.
[20] Lindsey JC, Ryan LM. Methods for interval-censored data. Stat Med 1998;17(2):219–38.
[21] Tan Z. A new approach to MLE of Weibull distribution with interval data. Reliab Eng Syst Saf 2009;94(2):394–403.

[22] Peng Y, Wang Y, Zi Y, Tsui K-L, Zhang C. Dynamic reliability assessment and prediction for repairable systems with interval-censored data. Reliab Eng Syst Saf 2017;159:301–9.
[23] García-Mora B, Debón A, Santamaría C, Carrión A. Modelling the failure risk for water supply networks with interval-censored data. Reliab Eng Syst Saf 2015;144:311–8.
[24] Sun J. The statistical analysis of interval-censored failure time data. Statistics for biology and health, first ed. Springer-Verlag New York; 2006.
[25] Chou C-Y, Lin Y-C, Chang C-L, Chen C-H. On the bootstrap confidence intervals of the process incapability index Cpp. Reliab Eng Syst Saf 2006;91(4):452–9.
[26] Wang Z, Shafieezadeh A. On confidence intervals for failure probability estimates in Kriging-based reliability analysis. Reliab Eng Syst Saf 2020;196:106758.
[27] Wang D, Hryniewicz O. A fuzzy nonparametric shewhart chart based on the bootstrap approach. Int J Appl Math Comput Sci 2015;25(2):389–401.
[28] Ramos-Guajardo A, Blanco-Fernandez A, Gonzalez-Rodriguez G. Applying statistical methods with imprecise data to quality control in cheese manufacturing. In: Soft modeling in industrial manufacturing. Cham: Springer; 2019.
[29] Grzegorzewski P, Hryniewicz O, Romaniuk M. Flexible resampling for fuzzy data. Int J Appl Math Comput Sci 2020;30(2):281–97.
[30] Newton MA, Raftery AE. Approximate Bayesian inference with the weighted likelihood bootstrap. J R Stat Soc Ser B Stat Methodol 1994;56:3–48.
[31] Xu L, Gotwalt C, Hong Y, King CB, Meeker WQ. Applications of the fractional-random-weight bootstrap. Amer Statist 2020;1–29.