Taylor & Francis
Taylor & Francis Group

Journal of Applied Statistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjas20

Type-II progressive censoring with GLM-based random removal mechanism dependent on the experimental conditions

Fatemeh Hassantabar Darzi, Samaneh Eftekhari Mahabadi & Firoozeh Haghighi

To cite this article: Fatemeh Hassantabar Darzi, Samaneh Eftekhari Mahabadi & Firoozeh Haghighi (2022): Type-II progressive censoring with GLM-based random removal mechanism dependent on the experimental conditions, Journal of Applied Statistics, DOI: 10.1080/02664763.2022.2104230

To link to this article: https://doi.org/10.1080/02664763.2022.2104230

Published online: 06 Aug 2022.

JOURNAL OF APPLIED STATISTICS
https://doi.org/10.1080/02664763.2022.2104230

Taylor & Francis
Taylor & Francis Group

Type-II progressive censoring with GLM-based random removal mechanism dependent on the experimental conditions

Fatemeh Hassantabar Darzi, Samaneh Eftekhari Mahabadi and Firoozeh Haghighi

School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran

ABSTRACT

This article presents a novel stochastic removal mechanism under Type-II progressive random censoring in which removal probabilities are allowed to be dependent on the lifetime conditions through Generalized Linear Models (GLM). These conditions potentially include failure distances (the time required to observe the next failure) or other covariate information available in the experiment. The proposed GLM-based random removal mechanism includes a set of tuning parameters that are determined by the researcher according to the possible failure distance category. These parameters allow flexible determination of the removal probabilities leading to necessary experimental cost and time reductions. To establish the proposed mechanism, the Proportional Hazard Rate (PHR) family of distributions is considered. Also, the maximum likelihood estimators of parameters and their asymptotic variances are derived for the Weibull distributed lifetime data. A simple simulation algorithm for generating Type-II progressive censoring samples with GLM-based dependent removal probabilities is also presented. The expected experiment time required to complete the life test under this censoring scheme is also investigated using the Monte Carlo integration method. Several simulation studies are conducted to evaluate and compare the performance of the proposed mechanism. A sensitivity analysis is also considered to study the effect of misspecification of removal mechanism coefficients. Finally, two real data sets are analyzed for illustrative purposes.

ARTICLE HISTORY
Received 15 July 2021
Accepted 15 July 2022

KEYWORDS
Type-II progressive censoring; stochastic dependent random removal; generalized linear models; proportional hazard rate family; expected experiment time

1. Introduction

Censored observations of different types frequently occur in many applications depending on the setting of the data collection process. In fact, censoring could be either controlled or uncontrolled by the researcher who plans the experiment. For example, an experimenter may terminate the life test study when a determined number of failed products are observed to save time or cost which is referred to as Type-II censoring. Furthermore, some surviving test units may have to be removed from the study at different failure times due to various

CONTACT Samaneh Eftekhari Mahabadi seftekhari@ut.ac.ir School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran 14155–6455, Iran

reasons which would result in Type-II progressive censoring. This procedure works as follows. Life test experiment starts with n units and terminates after mth failure is observed. After observing the first failure, r_1 units are randomly selected among the $n-1$ surviving units and removed. At the time of second failure, which is the smallest lifetime among the $n-1-r_1$ units, r_2 units are randomly chosen from $n-2-r_1$ remained units and withdrawn from the experiment. This process is continued until observing mth failure, then all $n-m-r_1-\cdots-r_{m-1}$ remained units are removed from the experiment. Note that if $r_1=r_2=\cdots=r_m=0$, then $n=m$ which corresponds to the complete sampling. Also, $r_1=r_2=\cdots=r_{m-1}=0$ and $r_m=n-m$ lead to the conventional Type-II right censoring plan.

Inferential issues, for Type-II progressive censoring schemes, have been addressed by several authors. Gibbons and Vance [21] investigated two methods of estimation for progressive censoring Weibull distributed data. Balakrishnan *et al.* [5,6] discussed parameters estimation under Type-II progressive censoring for Gaussian and Extreme value lifetime distributed data, respectively. Ng *et al.* [29] considered three optimality criteria for finding optimal progressive censoring plans. They computed the expected Fisher information and the asymptotic variance–covariance matrix of the maximum likelihood estimates based on a Type-II progressive censoring sample from the Weibull distribution. Optimal censoring designs in terms of minimum variance of best linear unbiased estimators were considered by Burkschat *et al.* [13,14]. Many optimality criteria have been proposed by different researchers among which one can refer to Wu and Haung [41], Cramer and Ensenbach [16] and references cited therein for further study. Moreover, under Type-II progressive censoring scheme, statistical inference of various models has been studied by many authors and some of them are mentioned in the following. Rasouli and Balakrishnan [31] considered two exponential populations for modeling comparative lifetime experiments when joint Type-II progressive censoring is implemented on the two samples. Also, a joint progressive Type-II censoring was developed to the k comparative exponential populations by Balakrishnan *et al.* [8]. Sel *et al.* [?] developed a new six-parameter distribution based on Type-II progressive censoring sample which can be fitted well to complex data. They considered different estimation methods such as maximum likelihood estimation, bootstrap, and Bayesian for comparison purposes. Mondal and Kundu [28] introduced a balanced two-sample Type-II progressive censoring scheme and provided the exact inference when two populations had the Weibull distribution. For elaborate discussion and deep description of the different progressive censoring schemes and their related issues, the reader is encouraged to study the books by Balakrishnan and Aggarwala [9] and Balakrishnan and Cramer [10] or the excellent review article by Balakrishnan [7].

One of the most important questions in the Type-II progressive censoring scheme is how to planning the number of removals. In the above-mentioned progressive censoring studies, it is most often assumed that the removal of units from an experiment is pre-planned and intentional. Actually, in many practical situations, the number of removals cannot be predetermined in advance by an experimenter. For example, the number of patients who drop-out of a clinical test at each failure time is not predetermined or in some industrial experiments, it is too dangerous to determine the number of removal units in advance. In these cases, it is more realistic to assume that censoring numbers are chosen randomly, according to a probability distribution on the set of possible censoring numbers. This type of censoring was defined as Type-II progressive censoring with random removals, denoted

by Type-II PCR. This paper proposes a planning of removal numbers via a mechanism dependent on the experimental conditions using GLM.

First, Yuen and Tse [44] and Tse *et al.* [36] proposed Type-II PCR where the number of units removed at each stage follows a discrete uniform and binomial distribution with a fixed known probability p, respectively. Tse and Yuen [39] considered the expected experiment times to assess the required time to complete a life test for the Weibull distributed lifetimes under Type-II PCR. Later, Tse and Yang [38] and Tse and Xiang [37] studied expected experiment time and investigated the problem of parameters estimation for Type-II PCR model. Also, Singh *et al.* [34] considered situation where the number of removals follows a binomial distribution with success probability following a Beta prior distribution to reflect the uncertainty in the probability of a removal at each failure time. There are numerous literature on random removals using binomial or discrete uniform distributions assuming different lifetime distributions. Readers can find more details referring to Wu *et al.* [42], Amin [4], Yan *et al.* [?], Dey and Dey [19], Singh *et al.* [35], Day *et al.* [18], Gunasekera [22] and Sharafi [33] which considered Gompertz, Pareto, Generalized Exponential, Rayleigh, Poisson-Exponential, Weighted Exponential, Burr XII and two-parameter Lindley as lifetime distributions, respectively.

The assumptions regarding the uniform distribution with equal chance for the number of removals, or binomial distribution with the fixed probability of a removal at each stage, do not seem to be realistic in the practical situations. As in the previous clinical example, if more or less deaths are recorded in the early stages of the test then what would happen at the next upcoming stages. Certainly, the probability of removing in different stages will not be the same. Actually, one of the most challenging issues in Type-II progressive censoring is that removal vector is chosen independent from lifetime distribution. Recently, Ghahramani *et al.* [20] proposed an approach for determining the removal vector based on failure distances. They considered a specific increasing function of difference between the two last failure times divided by the time of the first failure to determine the removal numbers. This may improve random removal setting by allowing the number of removals to be dependent on failure distances but there are two major drawbacks: (i) this approach is only applicable for the specific case of exponential lifetime distribution where the joint probability mass functions of random removals can be obtained based on the result '5' in [7] and (ii) the number of removals is restricted to be a known systematic function of the failure distances which always encourages more removals for the larger previous failure distances.

However, in realistic situations, failure distances and covariate information are available in the experiment which could influence on the removal numbers stochastically based on a flexible unknown relation which should be estimated or determined according to the goals of study. For example, let us consider a survival study where the cancer patients are put under a new clinical test. During such studies, if the failures (patient death) happen in the early stages of the test with short distances then patients are more likely to leave the test (by themselves or the experimenter). On the other hand, if the distances between failures are large enough, then the drop-out chance of a patient is expected to be relatively small. In industries for reliability testing, there is an opposite direction, the experimenter might decide to withdraw more units when larger failure intervals happen.

In this paper, for the first time, we introduce a stochastic removal mechanism which gives the flexibility of updating the removal probability based on the available information at each failure time for a large flexible class of lifetime distributions. We have proposed

some conditional binomial distributions for the removal numbers with success probabilities dependent on the experimental conditions. These conditions may include preceding number of removals, observed failure times and some additional information potentially related to the removal decision at each stage. To define this random removal mechanism, GLMs for binomial distribution assuming different link functions are considered. The presented model includes the binomial removal scheme with fixed probability as a special case. Here after, we will call this terminology as 'Type-II PCR with Random Dependent removal' and denote it by Type-II PCRD. The proposed method has been discussed when the experimental units follow PHR family. The PHR family has been extensively used to model failure time data. This family is flexible enough to accommodate both monotonic as well as non-monotonic failure rates even though the baseline failure rate is monotonic. Gupta and Kunda [24] used this family to introduce a new statistical distribution. In fact, they considered exponential distribution as the base distribution and created extended exponential distribution. Many authors have investigated this family, such as Ahmadi *et al.* [2,3], Kundu and Nanda [25] and Psarrakos and Sordo [30]. The main contributions of this work to the current literature lie in the following aspects. (1) Proposing a Type-II progressive censoring scheme where the random removal numbers follow binomial distributions with different success probabilities at different failure times according to GLMs for binomial family. (2) The proposed GLM-based random removal mechanism includes a set of tuning parameters that could be adjusted by the researcher which allows both increasing or decreasing impact of previous failure distances (according to the study protocols) as well as other available information. (3) The proposed mechanism is not restricted to a special lifetime density and is shown to be applicable for the large family of PHR distributions.

The structure of our paper is organized as follows. Our methodology is presented in Section 2, which includes the introduction of the proposed GLM-based random removal mechanism for Type-II PCR scheme and its application for the PHR family. In Section 3, the maximum likelihood estimators and their asymptotic properties are derived for the Weibull distributed lifetime data, as a special member of PHR family, under Type-II PCRD. The expected experiment time of our proposed censoring scheme is calculated in Section 4. In Section 5, after developing an algorithm to generate Type-II PCRD samples, several simulation studies are performed. In Section 6, a sensitivity analysis is conducted to analyze the effect of misspecification of the tuning parameter values on experiment design. Two real lifetime data about the endurance of deep groove ball bearings and survival times of ovarian cancer patients are analyzed in Section 7. Finally, Section 8 contains some concluding remarks and comments.

2. Methodology

2.1. Proposed GLM-based random removal mechanism

Consider a reliability experiment in which n independent and identical units are put on a life test. Let $T_{1:m:n} < \cdots < T_{m:m:n}$ denote the corresponding progressive sample including m ordered failure times out of n randomly selected items, where $m < n$ is predetermined before testing. As a progressive censoring scheme, assume that at the ith failure, for $i = 1, \ldots, m$, R_i items are randomly removed from the test and the experiment terminates when the mth failure is observed.

To allow for the unintentional as well as intentional random removals and also the flexibility to use the updated lifetime information available after each observed failure time, we assume that R_is have some conditional binomial distributions with success probabilities dependent on the experimental conditions. These conditions, at the ith stage might include preceding number of removals $(R_1, \ldots, R_{i-1})$, observed failure times $(T_{1:m:n}, \ldots, T_{i:m:n})$ and some additional covariates related to the removal decision denoted by a $q \times 1$ vector Z_i. Actually, we will assume that the number of random removals have the following conditional binomial distributions:

$$R_1 \mid T_{1:m:n}, Z_1 \sim b(n-m, p_1),$$

$$R_i \mid R_1, \ldots, R_{i-1}, \quad T_{1:m:n}, \ldots, T_{i:m:n}, \quad Z_i \sim b\left(n - m - \sum_{j=1}^{i-1} R_j, p_i\right), \quad i = 2, \ldots, m-1$$

$$R_m = n - \sum_{i=1}^{m-1} R_i - m,$$

where we propose the conditional removal probability, p_i, to be related to the lifetime conditions available at the ith stage, including the failure distance $FD_i = (T_{i:m:n} - T_{i-1:m:n})$ and the vector of possibly available covariates denoted by $Z_i = (Z_{i1}, \ldots, Z_{iq})$, through a known monotonic differentiable link function $H(\cdot)$. The stochastic random removal mechanism is introduced as follows:

$$H(p_i) = \alpha_0 + \alpha_1 FD_i + \sum_{h=1}^{q} \gamma_h Z_{ih}, \quad i = 1, \ldots, m-1, \tag{1}$$

where $T_{0:m:n} = 0$ and $\{\alpha_0, \alpha_1, \gamma\}$ is the set of tuning parameters leading to different removal probabilities according to the goals of study. The α_0 indicates the transformed removal probability for zero failure distances (when two failures happen exactly at the same time), while α_1 controls the direction of relation between FD_i and the removal probability. For example, if the experimenter desires to remove more (less) items for smaller FD_i's, $\alpha_1 < 0$ ($\alpha_1 > 0$) should be applied. Therefore, the above Type-II PCRD scheme has the ability to model the uncertainty in the removal probability at each stage of a Type-II progressive censoring experiment.

To explore features of the new random removal mechanism, suppose that the removal probability is assumed to be only dependent on the length of time spent to observe current failure. Therefore, the random removal mechanism or the conditional removal probability of (1) reduces to

$$H(p_i) = \alpha_0 + \alpha_1 FD_i. \tag{2}$$

One can apply *logit*, *probit* or complementary *log–log* (*c log–log*) link functions as common choices for $H(\cdot)$. The *logit* link is the inverse cumulative distribution function (cdf) of logistic distribution which is symmetrical, unimodal and similar in the shape to the normal distribution with the same mean and standard deviation with slightly thicker tails. Using the *logit* link function, the logarithm of the removal odds, $log\frac{p_i}{1-p_i}$, is modeled as a

linear combination of the FD_i which leads to the following logistic removal mechanism:

$$logit(p_i) = log\left(\frac{p_i}{1-p_i}\right) = \alpha_0 + \alpha_1 FD_i$$

or

$$p_i = \frac{e^{\alpha_0+\alpha_1 FD_i}}{1+e^{\alpha_0+\alpha_1 FD_i}},$$

where, for each 1 unit increase in the FD_i, the odds of removal would be multiplied by e^{α_1}.

Using *probit* link function, the random removal mechanism has the form,

$$probit(p_i) = \Phi^{-1}(p_i) = \alpha_0 + \alpha_1 FD_i,$$

where Φ is the standard normal cdf. For the *probit* link, removal probability, p_i, as a function of FD_i has the appearance of the normal cdf when $\alpha_1 > 0$. The rate of change in p_i with respect to FD_i is $\frac{\partial p_i}{\partial FD_i} = \alpha_1\phi(\alpha_0 + \alpha_1 FD_i)$, where ϕ is the standard normal probability density function (pdf). This rate achieves its maximum when $\alpha_0 + \alpha_1 FD_i = 0$ (i.e. at $FD_i = -\alpha_0/\alpha_1$). The *logit* and *probit* links are symmetric, so the removal probability curve, p_i, has a symmetric appearance about the point $p_i = 0.5$ and so p_i has the same rate for approaching 0 as well as for approaching 1.

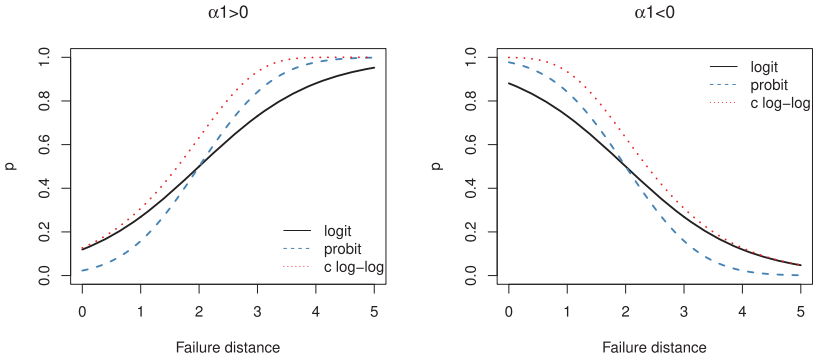
When the removal probability as a function of FD_i is assumed not to be symmetric in the sense that p_i approaches 0 fairly slowly but approaches 1 quite sharply, the *logit* and *probit* models are inappropriate. However, in this situation, the *c log–log* model could give better results. The *c log–log* link function is defined as

$$c\,log - log(p_i) = log[-log(1-p_i)],$$

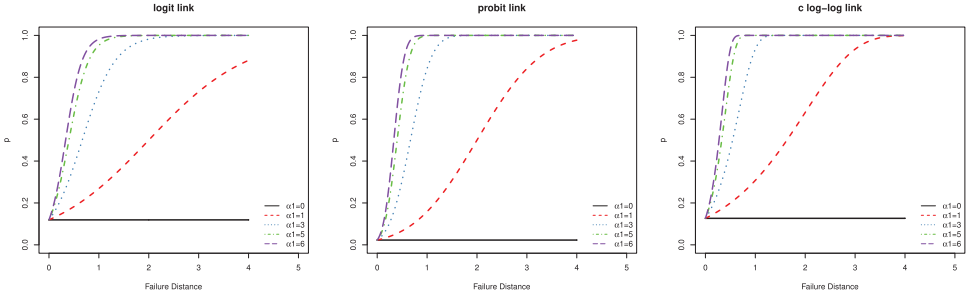
which is the inverse of the cdf of extreme value (log-Weibull) distribution. For more information on the behavior of these links, see Agresti [1].

2.2. Mechanism behavior with respect to the tuning parameters

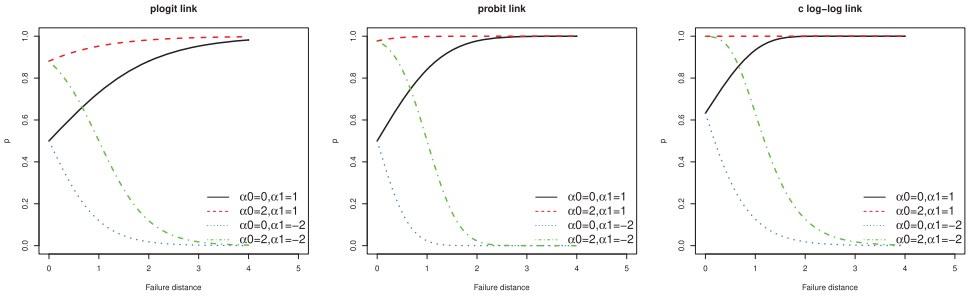
Understanding different behaviors of these link functions along with experimental conditions will allow researcher to intelligently choose when to use which one? The removal probability curve as a function of FD_i in (2), under three different links, when ($\alpha_0 = -2$ and $\alpha_1 = +1$) and ($\alpha_0 = 2$ and $\alpha_1 = -1$) are respectively plotted in the left and right panels of Figure 1(a). The sign of α_1 coefficient in the random removal mechanism determines whether p_i is an increasing or decreasing function of the FD_is. Figure 1(a) shows that the *logit* and *probit* links are both symmetrical but the *probit* link approaches to 0 and 1 faster than the *logit* curve. Therefore, to have a more sensitive removal mechanism to slight variations of failure distances, *probit* link can be used. Figure 1(a) shows that *c log–log* link function is right-skewed which lead to greater removal probability than the other two links for the same distances. When $\alpha_1 > 0$, the three link curves have the appearance of the logistic, normal and extreme value (log-Weibull) cdf with dispersion parameter $\frac{1}{|\alpha_1|}$ (i.e. the variance of *logit*, *probit* and *c log-log* are $\frac{\pi}{\sqrt{3}|\alpha_1|}$, $\frac{1}{|\alpha_1|}$ and $\frac{\pi}{\sqrt{6}|\alpha_1|}$, respectively). Hence, the rate of climb or descent in p_i increases as dispersion parameter decreases which can



(a) The effect of α_1 sign on the removal probability under three different link functions.



(b) Removal probability plot assuming different positive α_1 values and $\alpha_0 = -2$.



(c) Removal probability plot assuming different α_1 and α_0 values.

Figure 1. Removal probability plots against *FD*$_i$s assuming different links and tuning parameter values in the Type-II PCRD scheme: (a) the effect of α_1 sign on the removal probability under three different link functions; (b) removal probability plot assuming different positive α_1 values and $\alpha_0 = -2$; and (c) removal probability plot assuming different α_1 and α_0 values.

be obtained by choosing larger values for $|\alpha_1|$. Figure 1(b) illustrates the effect of positive α_1 values on p_i curves when $\alpha_0 = -2$, for three different link functions. Actually, larger positive values of α_1 return smaller variances and cause the removal probability approach to 1 faster for all three links. Figure 1(c) shows an effect of α_0 on the mechanism's behavior under three different links. Changing the value of α_0 causes the change of baseline removal probability.

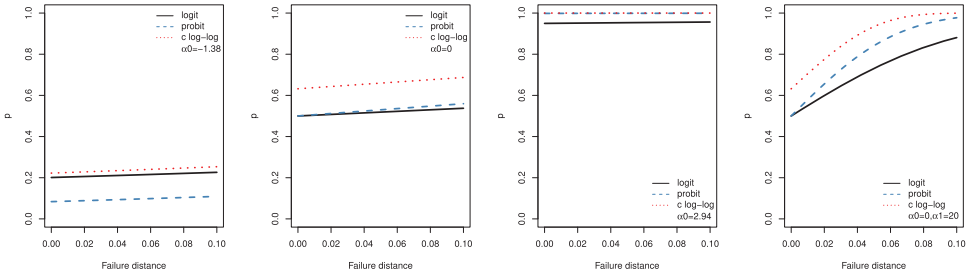
Determining appropriate values for α_0 and α_1 depends on the range of possible FD_i s and researcher's objective. If FD_i s are very small ($FD_i \longrightarrow 0$), the α_0 value has an important role to determine the probability of removal. In this case, to have a significant impact of FD_i s on the removal probability, one could change the scale of failure times or choose larger absolute values of α_1. Figure 2(a–c) shows the removal probability plots (rpps) against small, moderate and large FD_i s, respectively. These plots emphasize the determination of α_1 values with respect to the possible range of FD_i s. For example, choosing small values of α_1 for very small FD_i s leads to the Type-II PCR with $p_i = H^{-1}(\alpha_0)$ (three left panels of Figure 2 a), while increasing α_1 value allows significant impact of FD_i s on the removal probability (the right panel of Figure 2a).

Figure 2(b) illustrates the rpps for different values of α_0, α_1 assuming moderate values of FD_i (i.e. $0.1 \leq FD_i \leq 2$). Changing rate of p_i depends on the value of α_1 as was shown in Figure 1(b). The upper panel of Figure 2(b) gives the rpps for fixed value of $\alpha_0 = 0$ and $\alpha_1 = -3, 1, 3$ and 5. It is obvious that for larger values of α_1, the removal probability approaches to 1 more quickly. Consequently, to investigate the role of failure distances, we must use appropriate α_1 values which make a gradual change in the removal probability. Also, the rpps are given for the values of $\alpha_0 = -2, -1, 0, 1$ and fixed value of $\alpha_1 = 2$ in the bottom panel of Figure 2(b). These two tuning parameters (i.e. α_0 and α_1) can be applied to adjust the rate and the initial location of the removal probability. Figure 2(c) includes larger values of FD_i (i.e. $FD_i > 2$). For $\alpha_1 < 0$, the mechanism leads to small removal probability, $p \longrightarrow 0$, so the test units are likely to be removed at the end of experiment similar to the Type-II censoring. For positive values of α_1, removal probability quickly approaches to 1 and the experiment will be resembled a complete sampling test. Again, adjusting the values of α_0 and α_1 can lead to different removal probabilities in this case.

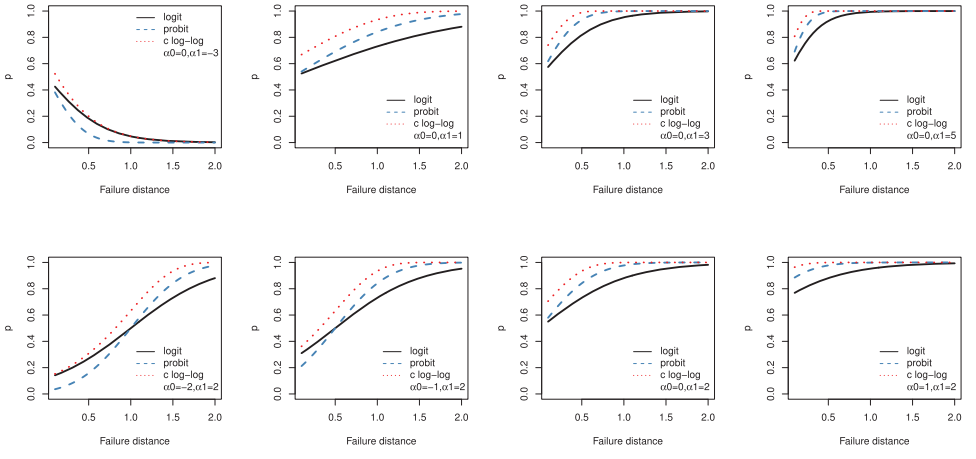
Consequently, these two tuning parameters (i.e. α_0 and α_1) can be applied to adjust the rate and the initial location of the mechanism. More specifically, to have a significant role of failure distances on the removal probabilities, we must use appropriate $\alpha_i, i = 0, 1$ values which make a gradual change in the removal probability. In the next section, appropriate $\alpha_i, i = 0, 1$ values are suggested according to possible range of FD_i s.

2.3. Determination of effective tuning parameters according to possible FD_i range

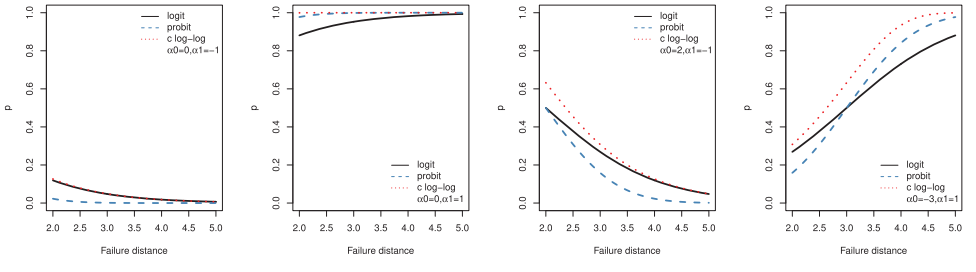
To determine appropriate tuning parameter values which allow the censoring mechanism to remove units according to FD_i, one needs approximately to know the possible range of FD_i s. Actually, the tuning parameters of random removal mechanism are used for emphasizing the role of experimental conditions especially FD_i s and adjusting them. Moreover, the introduced mechanism is based on three links which have the appearance of the logistic, normal and extreme value (log-Weibull) cdf. Therefore, by knowing the behavior of these distributions and possible ranges of FD_i s, the effective $\alpha_i, i = 0, 1$ values which cause various removal probabilities at each stage according to failure distances can be determined. As mentioned in Section 2.2, possible range of FD_i s are categorized to small, moderate and large values, where their corresponding appropriate $\alpha_i, i = 0, 1$ under *logit* link are presented in Table 1. Table 1 reports the mean of removal probabilities for 1000 FD_i s generated in each category, assuming different $\alpha_i, i = 0, 1$. The points which are common in all parts of Table 1 are given as follows:



(a) The rpps for small FD_i's (i.e., $(0, 0.1)$).



(b) The rpps for moderate FD_i's (i.e., $(0.1, 2)$).



(c) The rpps for large FD_i's (i.e., $(2, 5)$)

Figure 2. The rpps for different ranges of FD_is assuming different values of the tuning parameters under three different link functions: (a) the rpps for small FD_i's (i.e. (0, 0.1)); (b) the rpps for moderate FD_i's (i.e. (0.1, 2)); (c) the rpps for large FD_i's (i.e. (2, 5)).

(1) The constant parameter, α_0, determines the initial location of the removal mechanism, where changing its value causes the baseline removal probability changes, as is shown in Figure 2. Large values of $|\alpha_0|$ lead to extreme values of removal probability (removal probability close to 0 or 1) which do not allow the mechanism to operate according to FD_i s. Actually, probability of removal under *logit* link without considering any

covariates (i.e. $p_i = \frac{e^{\alpha_0}}{1+e^{\alpha_0}}$) takes values in $(0.04, 0.9)$ for $\alpha_0 \in (-3, 3)$. It should be noticed that considering a wider interval for α_0 does not significantly affect on the removal vector.

(2) Usually, it is better to consider opposite signs for α_0 and α_1. This is due to the fact that assuming similar negative (positive) signs of α_1 and α_0 leads to $p_i \to 0$ ($p_i \to 1$) where distinguishing between α_0 or FD_i s effect could be difficult.

(3) Large negative values of α_0 cause small removal probabilities ($p_i \to 0$) which is likely to result in a Type-II censoring with random removal vector, $R = (0, \ldots, 0, n - m)$. Also, large positive values of α_0 lead to large removal probability ($p_i \to 1$) which is likely to result in a complete censoring design with random removal vector, $R = (n - m, 0, \ldots, 0)$. Therefore, large absolute values of α_0 do not allow the random removal mechanism to operate according to FD_i s. It is advised to apply α_0 values in $(-3, 3)$; however, considering slightly larger or smaller values when are adjusted with α_1 values, does not cause serious problems.

Consequently, by considering α_0 belongs to $(-3, 3)$ in Table 1, one can select α_1 to have random removal according to possible FD_i categories. Specifically, Table 1(a) shows appropriate values of $\alpha_i, i = 0, 1$ when FD_i s are small (i.e. $0 \leq FD_i < 0.1$). Since the FD_i values are small in this category, we should choose larger $|\alpha_1|$ coefficients (i.e. $|\alpha_1| \geq 10$) to have an effective mechanism which could adjust the removal probability of each stage according to the current FD_i. As shown in Table 1(a), considering opposite sings of α_0 and α_1 prevents the removal probability to approach 0 or 1 quickly.

Table 1(b) shows appropriate $\alpha_i, i = 0, 1$ for moderate FD_i (i.e. $FD_i \in (0.1, 2)$). In this case, α_1 values satisfying the inequality $|\alpha_1| \leq 10$ are appropriate to have random removals according to FD_i s. Larger absolute values of α_1 can be used for smaller values of FD_i s in moderate range. For this category of FD_i, close values of α_1 and α_0 with the same sign can be used.

Table 1(c) shows appropriate $\alpha_i, i = 0, 1$ for large values of FD_i (i.e. $FD_i > 2$). As shown in Figure 2(c), for large values of FD_i removal probability quickly approaches 0 or 1 and does not permit the mechanism to work. In this case, for adjusting the mechanism to handle large failure distances, small values for the FD_i coefficient, α_1, are needed. The effective range of α_1 according to Table 1(c) is considered in $(-3, 3)$. This coefficient should be smaller for larger values of FD_i s at this category (i.e. for FD_i s greater than 5, the effective range of α_1 is $(-0.5, 0.5)$).

The above results guide the researcher to determine appropriate values for the removal mechanism coefficients, $\alpha_i, i = 0, 1$ that allow the random removal mechanism to operate according to FD_i s. Also, it should be noticed that determining exact or optimal values of $\alpha_i, i = 0, 1$ could be dependent on the desired optimality criteria (e.g. to reduce duration or variation of experiment) and the goal of experiment which could be studied as a future work.

Obviously, in real situations, all FD_i s of an experiment might not be in one category, and according to a probability distribution, FD_i s may be in one, two or all three categories. In this case, decision about effective $\alpha_i, i = 0, 1$ values can be made based on the category that most failure distances fall into. On the other hand, determining a neighborhood of $\alpha_i, i = 0, 1$ values is enough to guarantee a random removal scheme with desired level of dependence on failure distances.

Table 1. Removal probabilities with *logit* link assuming different $\alpha_i, i = 0, 1$ values when FD_i s are in small, moderate or large category.

(a) The small FD_i's category

		α_1										
Category of FD_i	α_0	< -40	-40	-30	-20	-10	$(-10,10)$	10	20	30	40	> 40
small FD_i	-3			$p_i < 0.04$			$p_i \simeq \frac{e^{\alpha_0}}{1+e^{\alpha_0}}$	0.08700	0.13325	0.21693	0.31200	> 0.32
$FD_i < 0.1$	-2			$p_i < 0.12$				0.20300	0.39988	0.39988	0.50200	> 0.51
	-1			$p_i < 0.26$				0.40500	0.60973	0.60973	0.68600	> 0.69
	1	< 0.30	0.30987	0.21120	0.49669	0.59510				$p_i > 0.73$		
	2	< 0.49	0.49302	0.60012	0.71471	0.79650				$p_i > 0.88$		
	3	< 0.67	0.67922	0.78307	0.86675	0.91300				$p_i > 0.95$		

(b) The moderate FD_i's category

		α_1													
Category of FD_i	α_0	-10	-8	-6	-4	-3	-2	-1	1	2	3	4	6	8	10
moderate FD_i	-3			$p_i \longrightarrow 0$					0.13457	0.32339	0.51523	0.64187	0.77475	0.83975	0.87872
$.1 \leq FD_i \leq 2$	-2							0.05205	0.28790	0.51257	0.66849	0.75895	0.81417	0.85074	0.92395
	-1						0.06942	0.12746	0.50750	0.69895	0.80212	0.85819	0.91458		
	1					0.19788	0.30105	0.49250	0.87254	0.93210			$p_i \longrightarrow 1$		
	2			0.14926	0.24105	0.33151	0.48743	0.71210	0.94795						
	3	0.09326	0.16025	0.22525	0.35813	0.67661	0.67661	0.86543	0.98002						

(c) The large FD_i's category

		α_1						
Category of FD_i	α_0	-3	-2	-1	-0.5	0.5	1	2	3
large FD_i	-3	$p_i \longrightarrow 0$			0.00941	0.23133	0.60369	0.94661	0.97430
$FD_i > 2$	-2			0.00581	0.02511	0.44023	0.78410	0.97827	
	-1			0.01553	0.06498	0.67171	0.90088		
	0			0.04045	0.15662	0.84338		$p_i \longrightarrow 1$	
	1			0.09912	0.32829	0.93502			
	2		0.02173	0.21590	0.55977	0.97489			
	3	0.03451	0.05339	0.39631	0.76867	0.99059			

2.4. The proportional hazard rate family under type-II PCRD

Let X be a non-negative random variable denoting the lifetime of a component having cdf $F_0(x)$ and pdf $f_0(x)$. Then the hazard rate function of X is given by $h_0(x) = f_0(x)/\bar{F}_0(x)$, where $\bar{F}_0(x) = 1 - F_0(x)$. The family of distributions with hazard rate function of the form $\{\theta h_0(\cdot), \theta > 0\}$ are called PHR, with baseline cdf, $F_0(\cdot)$. If T is a member of the PHR family with the baseline cdf $F_0(\cdot)$, its cdf would be

$$F_T(t;\theta) = 1 - [\bar{F}_0(t)]^{\theta}, \quad t > 0, \ \ \theta > 0, \tag{3}$$

see Gupta *et al.* [23] for more details. The pdf and the hazard rate function of T are respectively given by

$$f_T(x;\theta) = \theta f_0(x)[\bar{F}_0(x)]^{\theta-1}$$

and

$$h_T(t;\theta) = \theta h_0(t).$$

The PHR family of distributions includes several well-known lifetime distributions such as Exponential, Gamma, the Weibull, Pareto (Types I and II), Burr type XII and Beta. For a Type-II progressive censoring with predetermined vector of removals $\underline{R}_m = (R_1, \ldots, R_m)$ in which lifetimes of the units follow a member of PHR family given in (3), the likelihood function would be

$$\begin{aligned} L(\xi, \theta | \underline{t}, \underline{R}_m) &= C_r \Pi_{i=1}^{m} f_T(t_i;\theta)\bar{F}_T(t_i;\theta)^{r_i} \\ &= C_r \Pi_{i=1}^{m} \theta f_0(t_i,\xi)\bar{F}_0(t_i,\xi)^{\theta(r_i+1)-1}, \end{aligned} \tag{4}$$

where $C_r = n\Pi_{i=1}^{m}(n - i + 1 - \sum_{j=1}^{i} r_j)$ and ξ is the vector of parameters of the baseline distribution. The above likelihood is derived conditional on R_is which can be any integer values between 0 and $n - m - (R_1 + \cdots + R_{i-1})$, for $i = 1, \ldots, m$.

Assuming random removals under the proposed Type-II PCRD mechanism, the observed data includes the vector of lifetimes, $\underline{T}_m$, along with the vector of number of removals, $\underline{R}_m$. Let $\underline{T}_k = (T_{1:m:n}, \ldots, T_{k:m:n})$ and $\underline{R}_k = (R_1, \ldots, R_k)$ for $k = 1, \ldots, m$. The joint density function of $(\underline{T}_m, \underline{R}_m)$ can be decomposed as follows:

$$\begin{aligned} f(\underline{T}_m, \underline{R}_m) &= f((T_{1:m:n}, R_1), \ldots, (T_{m:m:n}, R_m)) \\ &= f(T_{1:m:n}, R_1)\Pi_{i=2}^{m} f(T_{i:m:n}, R_i | \underline{T}_{i-1}, \underline{R}_{i-1}). \end{aligned}$$

In addition, we can decompose each conditional density in the above product as

$$f(T_{i:m:n}, R_i | \underline{T}_{i-1}, \underline{R}_{i-1}) = f(T_{i:m:n} | \underline{T}_{i-1}, \underline{R}_{i-1}) f(R_i | \underline{T}_i, \underline{R}_{i-1}).$$

Also, using the following two results presented by Balakrishnan [7] (page 215):

(1) The marginal distribution of $T_{i:m:n}$, $1 \le i \le m$, is free of $(R_i, \ldots, R_m)$.
(2) $\underline{T}_i$ forms a Type-II progressive censoring sample of size i from n units on a life test with the progressive censoring scheme $(R_1, \ldots, R_{i-1}, n - i - \sum_{j=1}^{i-1} R_j)$,

we have,

$$f(T_{i:m:n}|\underline{T}_{i-1:m:n},\underline{R}_{i-1}) = \frac{f(\underline{T}_{i:m:n}|\underline{R}_{i-1})}{f(\underline{T}_{i-1:m:n}|\underline{R}_{i-1})}$$
$$= (n-\sum_{j=1}^{i} r_j - i + 1)\frac{f(t_i)}{1-F(t_{i-1})}\left(\frac{1-F(t_i)}{1-F(t_{i-1})}\right)^{n-\sum_{j=1}^{i-1} r_j - i}.$$

Hence, the likelihood function for the observed sample under Type-II PCRD reduces to

$$L(\Theta;\underline{R}_m,\underline{T}_m) = C_r f(t_1)[1-F(t_1)]^{n-1}\Pi_{i=2}^{m}\frac{f(t_i)}{1-F(t_{i-1})}\left(\frac{1-F(t_i)}{1-F(t_{i-1})}\right)^{n-\sum_{j=1}^{i-1} r_j - i}$$
$$\times \Pi_{i=1}^{m-1}\binom{n-m-\sum_{j=1}^{i-1} r_j}{r_i}\left(H^{-1}(\alpha_0+\alpha_1 FD_i+\sum_{h=1}^{q}\gamma_h Z_{ih})\right)^{r_i}$$
$$\times\left(1-H^{-1}(\alpha_0+\alpha_1 FD_i+\sum_{h=1}^{q}\gamma_h Z_{ih})\right)^{n-m-\sum_{j=1}^{i} r_j},$$

which can be rewritten in terms of the baseline cdf, $F_0(\cdot)$, and pdf, $f_0(\cdot)$, as follows:

$$L(\Theta;\underline{R}_m,\underline{T}_m) = \theta^m C_r f_0(t_1)[\bar{F}_0(t_1)]^{\theta(n+1)-2}\Pi_{i=2}^{m}\frac{f_0(t_i)[\bar{F}_0(t_i)]^{(\theta-1)}}{\bar{F}_0(t_{i-1})}\left(\frac{\bar{F}_0(t_i)}{\bar{F}_0(t_{i-1})}\right)^{\theta(n-\sum_{j=1}^{i-1} r_j - i)}$$
$$\times \Pi_{i=1}^{m-1}\binom{n-m-\sum_{j=1}^{i-1} r_j}{r_i}\left(H^{-1}(\alpha_0+\alpha_1 FD_i+\sum_{h=1}^{q}\gamma_h Z_{ih})\right)^{r_i}$$
$$\times\left(1-H^{-1}(\alpha_0+\alpha_1 FD_i+\sum_{h=1}^{q}\gamma_h Z_{ih})\right)^{n-m-\sum_{j=1}^{i} r_j}, \quad (5)$$

where $\Theta = (\theta,\xi)$.

3. Statistical inference for the Weibull lifetime data under type-II PCRD

In this section, we assume the Weibull distribution as a special case of PHR family with $\theta = 1$. The pdf of the Weibull distribution with parameters $\xi = (\lambda,\beta)$ is given by

$$f(x) = \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-(\frac{x}{\lambda})^{\beta}}, \quad x>0,\ \beta>0,\ \lambda>0, \quad (6)$$

and its baseline survival function is $\bar{F}_0(x) = e^{-(\frac{x}{\lambda})^{\beta}}$. Let $\underline{T} = (T_{1:m:n},\ldots,T_{m:m:n})$ be the m ordered failure times out of n randomly selected items. Under Type-II PCRD, R_i units would be removed at the ith failure according to a binomial distribution with success probability dependent on the failure distances. Consider dependent removal mechanism

assuming *logit* link as follows:

$$p_i = \frac{e^{\alpha_0+\alpha_1 FD_i}}{1+e^{\alpha_0+\alpha_1 FD_i}}.$$

Substituting the Weibull density and *logit* link in (5), the joint likelihood function can be rewritten as

$$L(\xi;\underline{R}_m,\underline{T}) = C_R \frac{\beta}{\lambda}\left(\frac{t_1}{\lambda}\right)^{\beta-1}\exp\left[-\left(\frac{t_1}{\lambda}\right)^{\beta}\right]\prod_{i=2}^{m}\frac{\beta}{\lambda}\left(\frac{t_i}{\lambda}\right)^{\beta-1}\exp\left[\left(\frac{t_{i-1}}{\lambda}\right)^{\beta}-\left(\frac{t_i}{\lambda}\right)^{\beta}\right]^{n-i+1-\sum_{j=1}^{i-1}r_j}$$
$$\times\prod_{i=1}^{m-1}\binom{n-m-\sum_{j=1}^{i-1}r_j}{r_j}e^{(\alpha_0+\alpha_1 FD_i)r_i}\left[1+e^{\alpha_0+\alpha_1 FD_i}\right]^{-(n-m-\sum_{j=1}^{i-1}r_j)}.$$

The maximum likelihood estimators of the model parameters $\xi=(\lambda,\beta)$ for the Weibull distributed data can be obtained by solving the following equations:

$$\frac{\partial \ell}{\partial \lambda} = -m\frac{\beta}{\lambda}+\frac{\beta}{\lambda}\left(\frac{t_1}{\lambda}\right)^{\beta}+\sum_{i=2}^{m}a_i\frac{\beta}{\lambda}\left[\left(\frac{t_i}{\lambda}\right)^{\beta}-\left(\frac{t_{i-1}}{\lambda}\right)^{\beta}\right],$$
$$\frac{\partial \ell}{\partial \beta} = \frac{m}{\beta}+\sum_{i=1}^{m}\ln\frac{t_i}{\lambda}-\left(\frac{t_1}{\lambda}\right)^{\beta}\ln\frac{t_1}{\lambda}+\sum_{i=2}^{m}a_i\left[\left(\frac{t_{i-1}}{\lambda}\right)^{\beta}\ln\frac{t_{i-1}}{\lambda}-\left(\frac{t_i}{\lambda}\right)^{\beta}\ln\frac{t_i}{\lambda}\right],$$

where $a_i = n-i+1-\sum_{i=1}^{i-1}r_j$, the maximum likelihood estimators of $\xi=(\lambda,\beta)$ are the solutions of the following equations:

$$\hat{\lambda} = \left(\frac{1}{m}t_1^{\beta}+\sum_{i=2}^{m}a_i[t_i^{\beta}-t_{i-1}^{\beta}]\right)^{\beta^{-1}},$$
$$\frac{t_1^{\beta}\ln t_1+\sum_{i=2}^{m}a_i[t_{i-1}^{\beta}\ln t_{i-1}-t_i^{\beta}\ln t_i]}{t_1^{\beta}+\sum_{i=2}^{m}a_i[t_i^{\beta}-t_{i-1}^{\beta}]}-\frac{1}{\beta}=\frac{1}{m}\sum_{i=1}^{m}\ln t_i,$$

which can be solved numerically using iterative algorithms such as Newton–Raphson. Also, the entries of the observed information matrix, for Type-II PCRD sample of the Weibull distributed lifetimes are obtained as

$$\frac{\partial^2 \ell}{\partial \lambda^2} = m\frac{\beta}{\lambda^2}-\frac{\beta}{\lambda}\frac{\beta+1}{\lambda}\left(\frac{t_1}{\lambda}\right)^{\beta}-\sum_{i=2}^{m}a_i\frac{\beta}{\lambda}\frac{\beta+1}{\lambda}\left[\left(\frac{t_i}{\lambda}\right)^{\beta}-\left(\frac{t_{i-1}}{\lambda}\right)^{\beta}\right],$$
$$\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{m}{\beta^2}-\left(\frac{t_1}{\lambda}\right)^{\beta}\left(\ln\frac{t_1}{\lambda}\right)^2+\sum_{i=2}^{m}a_i\left[\left(\frac{t_{i-1}}{\lambda}\right)^{\beta}(\ln\frac{t_{i-1}}{\lambda})^2-\left(\frac{t_i}{\lambda}\right)^{\beta}(\ln\frac{t_i}{\lambda})^{\beta}\right].$$

Therefore, the observed Fisher information matrix based on the maximum likelihood estimators of (λ,β) is as follows:

$$I(\hat{\lambda},\hat{\beta}) = -\left[\begin{array}{cc}\frac{\partial^2\ell}{\partial\lambda^2} & \frac{\partial^2\ell}{\partial\beta\partial\lambda}\\ \frac{\partial^2\ell}{\partial\lambda\partial\beta} & \frac{\partial^2\ell}{\partial\beta^2}\end{array}\right]\Bigg|_{\lambda=\hat{\lambda},\beta=\hat{\beta}}.$$

Let $V = lim_{n\longrightarrow\infty} I^{-1}(\hat{\lambda}, \hat{\beta})$ which is a reasonable approximation to the covariance matrix of the vector of maximum likelihood estimators $(\hat{\lambda}, \hat{\beta})$ for large sample sizes. Then, the joint distribution of the maximum likelihood estimators of λ and β is approximately bivariate normal,

$$\begin{pmatrix} \hat{\beta} - \beta \\ \hat{\lambda} - \lambda \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, V\right) = N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \hat{\sigma}^2(\hat{\beta}) & \hat{\sigma}(\hat{\beta}, \hat{\lambda}) \\ \hat{\sigma}(\hat{\lambda}, \hat{\beta}) & \hat{\sigma}^2(\hat{\lambda}) \end{bmatrix}\right).$$

Therefore, the Wald confidence intervals according to asymptotic theory of maximum likelihood estimator can be applied.

4. The expected experiment time

In practical applications, it is important to have an idea of the duration of a life test. The experiment termination time is directly associated with the cost of the experiment. In progressive Type-II censoring plan, the termination time is given by the expectation of the mth order statistic in a sample of size n. From Balakrishnan and Aggarwala [9], the conditional expectation of $T_{m:m:n}$ for a fixed vector of $\underline{R} = (R_1 = r_1, \ldots, R_m = r_m)$ is given by

$$E(T_{m:m:n}|\underline{R} = \underline{r}) = C_R \sum_{l_1=0}^{r_1} \cdots \sum_{l_m=0}^{r_m} (-1)^A \frac{\binom{r_1}{l_1} \cdots \binom{r_m}{l_m}}{\Pi_{i=1}^{m-1} h(l_i)} \int_0^\infty x f(x) F^{h(l_i)-1}(x)\, \mathrm{d}x, \qquad (7)$$

where $A = l_1 + l_2 + \cdots + l_m$ and $h(l_i) = l_1 + l_2 + \cdots + l_i + i$. The expected experiment time (EET) for Type-II progressive censoring with random removal is evaluated by taking the expectation on both sides of (7) with respect to $\underline{R}$. It is given by

$$E(T_{m:m:n}) = E_R[E[T_{m:m:n}|\underline{R}]] = \sum_{r_1=0}^{g(r_1)} \sum_{r_2=0}^{g(r_2)} \cdots \sum_{r_{m-1}=0}^{g(r_{m-1})} E(T_{m:m:n}|\underline{R} = \underline{r}) P(\underline{R}), \qquad (8)$$

where $g(r_i) = n - m - \sum_{j=1}^{i-1} r_j$. Thus (8) gives an expression to compute the EET. A natural way to approximate the above complicated expression is to use the Monte Carlo method introduced by Metropolis and Ulam [27] and Von Neuman [40] which takes advantage of the special nature of (8), namely the fact that $f(\underline{T}_m, \underline{R}_{m-1})$ is a pdf. So that, if it is possible to generate K samples from $f(\underline{T}_m, \underline{R}_{m-1})$ (using the simulation algorithm presented in the next section) the sample average,

$$E(T_m) = \overline{T}_m = \frac{1}{K} \sum_{j=1}^{K} t_m^{(j)},$$

where $t_m^{(j)}$ denotes the mth-order statistics of the jth sample, converges (almost surely) to (8), when K goes to ∞, according to the Law of Large Numbers. Based on this approximation, one can simply obtain values of the EET for different values of parameters without solving such a complicated and time-consuming summation. Consequently, if the expected value of complete sample (i.e. $R_i = 0$, $i = 1, \ldots, m$ and $m = n$) is denoted by $E(T^*)$, then we define the ratio of the EET under the Type-II PCRD over EET for complete sample as REET $= \frac{E(T_m)}{E(T^*)}$ which does not depend on the scale parameter. When REET is close to 1,

termination point is closer to the complete sampling plan. Suppose that an experimenter wants to observe at least m units when the test is anticipated to be conducted under Type-II PCRD. Then the REET provides important information in determining whether the experiment time can be shortened significantly if a much larger sample of n test units is used and the test is stopped once m failures are observed.

5. Simulation study

In this section, several simulation studies are conducted to investigate properties of the proposed mechanism. First, an algorithm to generate Type-II PCRD sample is developed in the following section.

5.1. Algorithm for generating type-II PCRD sample

Previously, an algorithm to generate Type-II progressive censoring samples for different sizes and schemes based on the uniform distribution was proposed by Balakrishnan and Sandhu [11]. When the removal process is considered as random but independent of failure time, one can first simply simulate the removal scheme then perform as the case where the scheme is fixed prior to the study. However, when the random removal process depends on the failure time distances, as is proposed in this article, number of removals have to be generated conditionally at each stage. Consequently, the previous algorithm is not efficient. Here, we take the advantage of using the theorem proposed by Balakrishnan and Sandhu [11], to develop a simulator algorithm for Type-II PCRD sample of an arbitrary continuous distribution with cdf, $F(\cdot)$ as follows:

---

(1) Generate m independent uniform$(0, 1)$ observations, denoted by $W_1, \ldots, W_m$.
(2) Let $U_1 = 1 - W_1^{1/n}$, then $T_{1:m:n} = F^{-1}(U_1)$ is an observation for the first order statistic.
(3) For $i = 2, \ldots, m-1$ do the following steps:
   (a) Generate R_{i-1} from

$$binomial\left(n - m - \sum_{j=1}^{i-2} R_j\,,\; H^{-1}(\alpha_0 + \alpha_1 FD_{i-1} + \sum_{h=1}^{q} \beta_h Z_{ih})\right),$$

   where, $R_0 = 0$.
   (b) Set $V_i = W_i^{(1/(n-i+1-\sum_{j=1}^{i-1} R_j))}$,
   (c) Set $U_i = 1 - \Pi_{k=1}^{i} V_k$, then U_i is a Type-II PCRD sample from uniform $(0, 1)$.
   (d) Finally, let $T_{i:m:n} = F^{-1}(U_i)$.
(4) For $i = m$ the remaining units are removed from the test and

$$T_{m:m:n} = F^{-l}(W_m^{1/n-m+1-\sum_{j=1}^{m-1} R_j}).$$

In the above algorithm, $(T_{1:m:n}, T_{2:m:n}, \ldots, T_{m:m:n})$ is a Type-II PRCD sample from $F(\cdot)$ distribution with dependent removal scheme $\underline{R} = (R_1, \ldots, R_m)$. Notice that the third

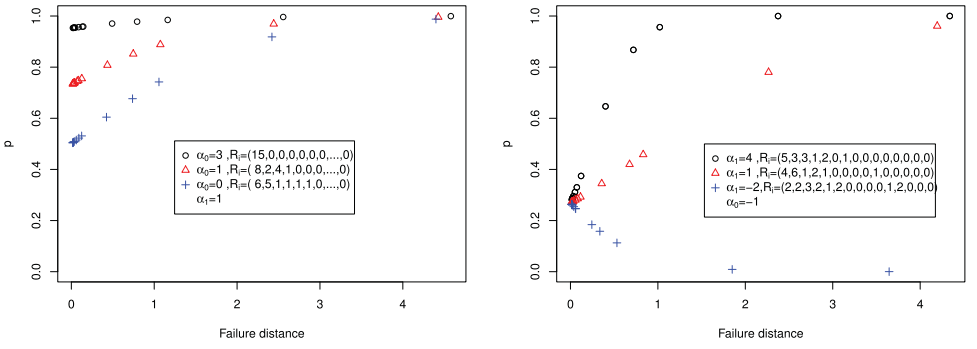
step of this algorithm uses the second result of Balakrishnan [7] (page 215) previously mentioned in Section 2.1 which states that, at the ith stage, $(T_{1:m:n}, \ldots, T_{i:m:n})$ given $(R_1, \ldots, R_{i-1})$ form a Type-II progressive censoring sample of size i, from n units with the vector of removal numbers $(R_1, \ldots, R_{i-1}, n - i - \sum_{j=1}^{i-1} R_j)$. Actually, the above simulation algorithm is established by generating Type-II progressive censoring sample from the uniform $(0, 1)$ distribution, $U_{(1)}, U_{(2)}, \ldots, U_{(m)}$, and defining $V_1 = U_{(1)}$, $V_i = \frac{1-U_{(i)}}{1-U_{(i-1)}}$ for $i = 2, \ldots, n$ and $W_i = V_i^{R_m+\cdots+R_i+m-i+1}$ for $i = 1, \ldots, m$, which leads to the fact that $W_i, i = 1, \ldots, m$, identically and independently follows uniform $(0, 1)$ distribution.

5.2. The effect of the Weibull parameters on the behavior of proposed mechanism

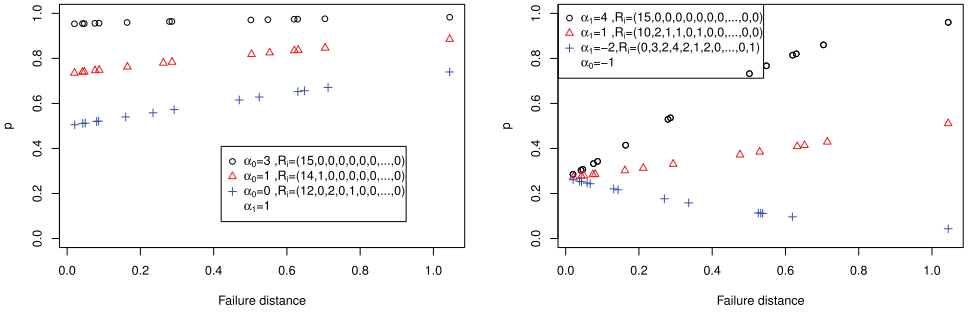
In practical applications, it is very important for the researcher to choose an efficient removal scheme when a Type-II progressive censoring is considered. Here, we study the properties of our proposed GLM-based random removal mechanism based on simulated Type-II PCRD samples from the Weibull distribution. Based on the properties of the lifetime distribution, one can choose appropriate tuning parameter values for the random removal mechanism. Actually, the parameters of the Weibull distribution affect the possible FD_i range. The shape parameter (i.e. β) of the Weibull distribution indicates dispersion of the lifetime data, which consequently affects on the range of FD_i s. The value of $\beta < 1$ indicates that the failure rate decreases over time for example in 'infant mortality'. In this case, the lifetime data, as were illustrated in Figure 2(a,c) could have large variations leading to both very small and large FD_i s, although, most of them fall into one of small or large categories of FD_i s that depends on λ value. A value of $\beta > 1$ indicates that the failure rate increases with time for example in 'aging' process or systems that are more likely to fail as time goes on. In this case, the lifetime data are more homogenous and FD_i s are moderate as shown in Figure 2(b), however, the value of λ also affects it.

On the other hand, the determination of the values of tuning parameters in the presented GLM-based random removal mechanism depends on how and when the researcher is interested to remove sample units from experiment and possible FD_i s. Actually, changing tuning parameters influences on the removal patterns ($\underline{R}$) and lifetime data ($\underline{T}$). Figure 3 plots removal probability against FD_i for simulated Type-II PCRD samples assuming different values of α_1, α_0, m, n and β (to better show the effect of the parameters, the primary failure data is assumed to be fixed so that FD_i s have approximately similar ranges).

For $\beta < 1$, if the researcher is interested in removing units at the beginning of an experiment to avoid additional loss, she/ he can choose higher removal probabilities. To have higher probabilities, the value of α_0 with *logit* link must be large enough, for example $\alpha_0 = 0, 1$ and 2 leads to $p = 0.5$, 0.73 and 0.88, respectively. But, if the researcher wants to remove the units according to FD_i s, she/ he can change the scale of failure lifetime data or choose proper values of α_1 assuming $\alpha_0 < 0$. Figure 3(a) illustrated the removal probability against FD_i s for simulated Weibull sample with shape parameter $\beta = 0.5$ and sample size of $n = 30$ and $m = 15$ censoring items. The left panel of Figure 3(a) is plotted assuming $\alpha_1 = 1$ and different α_0 values as the parameters of the *logit* link of random removal mechanism. According to this figure, choosing larger values of α_0 causes higher removal probabilities and the removal of units in the earlier stages of an experiment. The right panel



(a) Removal probability plot for the Weibull distributed data under Type-II PCRD assuming $\beta = 0.5$.



(b) Removal probability plot for the Weibull distributed data under Type-II PCRD assuming $\beta = 2$.

Figure 3. The effect of the shape parameter of the Weibull distribution on the random removal scheme assuming ($n = 30, m = 15$) and different values of α_0 and α_1: (a) removal probability plot for the Weibull distributed data under Type-II PCRD assuming $\beta = 0.5$ and (b) removal probability plot for the Weibull distributed data under Type-II PCRD assuming $\beta = 2$.

of Figure 3(a) shows the influence of α_1 values on the removal pattern assuming the logistic removal mechanism when $\alpha_0 = -1$. For $\beta > 1$, Figure 3(b) shows removal probability against FD_i for different values of α_1 and α_0. The $\underline{R}$ vectors are also shown in each panel which confirms the behavior of the mechanism. By comparing Figure 3(a,b), it can be seen that the range and dispersion of FD_i s are significantly different.

5.3. Simulation set up and results

One of the most important reason for appropriateness of Type-II PCRD is that the removals are determined dependent on the lifetime conditions. When the pre-fixed removal percentages are assumed, one can simply use maximum likelihood method corresponding to Type-II progressive censoring to estimate the parameters. Also, if the removal process is random following some binomial distributions with fixed removal probability, the maximum likelihood estimators could be obtained in the same manner as the non-random removal case since the removal mechanism is independent of the lifetime distribution.

Table 2 presents the parameter estimations, EET and coverage rate, CR, of parameters assuming $n = 20, 30$ and $m = 10$ with different values of tuning parameters in the logistic random removal mechanism. In this table, a simulation study has been conducted to generate 5000 Type-II PCRD samples from the Weibull distribution with shape parameters $\beta \in (0.75, 2)$ and scale parameter $\lambda = 6$. Then, the parameter estimation has been done using both dependent removal model presented in this article, that is Type-II PCRD, and Type-II PCR model assuming uniform and binomial distributions. As it is shown in Table 2, the EET of Type-II PCRD model is less than Type-II PCR. In addition, the bias (Bias) and the mean squared error (MSE) of the maximum likelihood estimators of the parameters in Type-II PCRD and Type-II PCR models are computed. Bias is computed by averaging the difference between the estimator and true parameter value over simulation samples. Also, the MSE measures the average squared difference between the estimation (i.e. $\hat{\lambda}$ or $\hat{\beta}$ in the Weibull distribution) and the true parameter value over 5000 simulated samples. The maximum likelihood estimator of the parameters and their MSE and Bias are not very different in Type-II PCRD and Type-II PCR. This numerical result coincides with that obtained by Cramer and Iliopoulos [17]. They obtained that the progressive censoring random removal mechanism does not affect on the maximum likelihood estimator of parameters. Therefore, the maximum likelihood estimators are expected to have similar properties in different kinds of Type-II progressive censoring schemes. Also, the CR of the parameters of the Weibull distribution are reported in Table 2. The CR of the estimators is the ratio of times that the confidence intervals overlap the true value of over 5000 simulations. The results of Table 2 show that the CR of scale parameter, CR_λ, is very close to the nominal confidence level (0.95) in all approaches. Also, it can be observed that most of the CRs for the shape parameter, CR_β, are less than the nominal confidence level (0.95) in all approaches.

Tables 3–5 show the approximated values of $E(T_m)$ under Type-II PCRD using *logit* link in the removal mechanism. For investigating properties of the EET, we have considered different model parameter values which are listed below,

$$\begin{aligned}
&\lambda = 4, \quad \beta = .5, \quad \alpha_0 = \{-3, 0\}, \quad \alpha_1 = \{-4, -2, 0, 2, 4\},\\
&\lambda = 5, \quad \beta = 1, \quad \alpha_0 = \{-2, -1, 0, 1\}, \quad \alpha_1 = \{-4, -2, 0, 2, 4\},\\
&\lambda = 6, \quad \beta = 2, \quad \alpha_0 = \{-1.5, -.5\}, \quad \alpha_1 = \{-4, -2, 0, 2, 4\}.
\end{aligned}$$

For each combination, we have taken different sample sizes $n = 6$, 20, 30 and 50 respectively as the tiny, small, moderate and large samples. Also, m is chosen such that the observed sample contains 100%, 90%, ..., 50% of the available sample units. The complete sampling plan is included when $m = n$. In our study, assuming the Weibull distributed lifetime data, the shape parameter β is chosen to be 0.5, 1 and 2. These values contain decreasing (or increasing) failure rates for $\beta < 1$ (or $\beta > 1$). Also, the exponential distribution is included when $\beta = 1$. The results of all the simulation studies in the following are based on 5000 repetitions.

Table 3 gives the EET of the Weibull distributed Type-II PCRD sample assuming $\beta = 0.5$ with two different values for the intercept of the removal mechanism, $\alpha_0 = -3$ and 0. The reduction in the EET comparing with a complete sampling scheme is highly significant for all n and m. Also, the reduction in the EET is very significant for all sample sizes, tiny to large, when the negative values are selected for α_0. When $\alpha_0 = 0$, the experiment time is

Table 2. Comparing various Type-II progressive censoring with random removal schemes by maximum likelihood estimator of parameters and EET for the Weibull distribution with $\beta = 2$ and 0.75 and $\lambda = 6$.

m	n	Approaches	Parameters	β	$\hat{\lambda}$	Bias($\hat{\lambda}$)	MSE($\hat{\lambda}$)	$\hat{\beta}$	Bias($\hat{\beta}$)	MSE($\hat{\beta}$)	$E[T_m]$	CR_λ	CR_β
10	20	Type-II PCRD	$\alpha_0 = 0, \alpha_1 = -3$		5.80961	−0.19039	0.85988	2.29514	0.29514	0.48899	7.70891	0.970	0.891
	20	Type-II PCRD	$\alpha_0 = 0, \alpha_1 = -1.5$		5.82052	−0.17948	0.83414	2.27714	0.27714	0.44187	8.70844	0.966	0.883
	20	Type-II PCR	Binomial with $p = 0.25$		6.28762	0.28762	1.18233	2.18015	0.18015	0.38286	9.03265	0.953	0.916
	20	Type-II PCR	Binomial with $p = 0.5$		6.02693	0.02693	0.85323	2.30019	0.30019	0.42298	9.74708	0.971	0.890
	20	Type-II PCR	Binomial with $p = 0.75$		5.96267	−0.03733	0.83464	2.28591	0.28591	0.40638	9.89503	0.969	0.888
	20	Type-II PCR	uniform	2	6.01753	0.01753	0.86226	2.30328	0.30328	0.42949	9.77080	0.970	0.894
	30	Type-II PCRD	$\alpha_0 = 0, \alpha_1 = -3$		5.75485	−0.24515	0.99193	2.26906	0.26906	0.43520	6.79356	0.964	0.856
	30	Type-II PCRD	$\alpha_0 = 0, \alpha_1 = -1.5$		5.78448	−0.21552	0.87163	2.23336	0.23336	0.34708	7.97426	0.965	0.863
	30	Type-II PCR	Binomial with $p = 0.25$		6.66502	0.66502	2.04536	2.07796	0.07796	0.32098	8.41603	0.927	0.926
	30	Type-II PCR	Binomial with $p = 0.5$		6.09188	0.09188	0.88663	2.28475	0.28475	0.36261	9.62468	0.970	0.891
	30	Type-II PCR	Binomial with $p = 0.75$		5.99611	−0.00389	0.84676	2.27509	0.27509	0.35379	9.82052	0.970	0.893
	30	Type-II PCR	uniform		6.09780	0.09779	0.91054	2.28551	0.28551	0.36692	9.65057	0.969	0.891
10	20	Type-II PCRD	$\alpha_0 = -2, \alpha_1 = 2$		5.87252	−0.12748	6.05814	0.86221	0.11221	0.06157	18.29934	0.972	0.840
	20	Type-II PCRD	$\alpha_0 = -2, \alpha_1 = -1.5$		5.76229	−0.23771	6.55631	0.89224	0.14224	0.10008	5.78459	0.966	0.829
	20	Type-II PCR	Binomial with $p = 0.25$		6.96410	0.96410	11.49934	0.83605	0.08605	0.06148	20.21370	0.945	0.912
	20	Type-II PCR	Binomial with $p = 0.5$		6.39079	0.39079	6.77982	0.86199	0.11199	0.05916	24.17103	0.969	0.887
	20	Type-II PCR	Binomial with $p = 0.75$		6.19916	0.19916	6.37365	0.85703	0.10703	0.05709	24.83280	0.968	0.885
	20	Type-II PCR	uniform	0.75	6.37868	0.37868	6.97183	0.86356	0.11356	0.06101	24.32909	0.969	0.892
	30	Type-II PCRD	$\alpha_0 = -2, \alpha_1 = 2$		5.69397	−0.30603	7.36154	0.87727	0.12727	0.07594	11.18104	0.971	0.857
	30	Type-II PCRD	$\alpha_0 = -2, \alpha_1 = -1.5$		5.71002	−0.28998	9.64616	0.89085	0.14085	0.09848	3.28390	0.966	0.842
	30	Type-II PCR	Binomial with $p = 0.25$		8.17971	2.17971	26.99833	0.79920	0.04920	0.04990	16.93944	0.914	0.934
	30	Type-II PCR	Binomial with $p = 0.5$		6.59727	0.59727	7.69888	0.85687	0.10687	0.05099	23.35260	0.967	0.892
	30	Type-II PCR	Binomial with $p = 0.75$		6.27472	0.27472	6.52800	0.85158	0.10158	0.04885	24.54825	0.970	0.885
	30	Type-II PCR	uniform		6.60539	0.60539	8.26189	0.85899	0.10899	0.05312	23.59731	0.970	0.888

Table 3. EET for the Weibull samples under Type-II PCRD for ($\beta = 0.5, \lambda = 4$) and $\alpha_0 = -3$ and 0 with different values of m and n.

		$\alpha_0 = -3$					$\alpha_0 = 0$				
		α_1					α_1				
n	m	-4	-2	0	2	4	-4	-2	0	2	4
6	6	28.6345	28.6345	28.6345	28.6345	28.6345	29.7914	29.7914	29.7914	29.7914	29.7914
	5	10.9209	11.1324	11.8981	19.5534	21.4435	19.1341	21.1553	24.1464	25.4109	25.2222
	4	4.8732	5.1987	5.9202	9.9415	12.7337	11.3558	12.1100	15.9209	18.2484	19.2485
	3	2.1516	2.1509	2.2911	3.5886	5.3463	4.6213	5.9708	8.7506	10.4875	12.11193
20	20	56.9686	56.9686	56.9686	56.9686	56.9686	59.0374	59.0374	59.0374	59.0374	59.0374
	18	24.9003	26.7045	31.4106	48.5745	50.4042	54.5632	54.9630	55.0786	55.0743	55.1070
	16	14.2154	14.7782	19.1606	42.1317	46.5384	51.8886	52.6859	53.3027	53.3887	53.4151
	14	8.0062	8.5416	10.3696	30.1376	38.2799	43.1737	45.1241	46.4612	46.6357	46.7174
	12	4.5752	5.0218	5.7275	16.7895	26.0151	36.6820	38.8528	42.2385	42.6798	42.7545
	10	2.5690	2.7009	2.9682	7.0476	13.6168	26.1383	29.8840	34.4646	35.6903	35.9428
30	30	70.5584	70.5584	70.5584	70.5584	70.5584	70.5051	70.5051	70.5051	70.5051	70.5051
	27	31.2392	34.3577	43.5971	62.7722	64.1943	66.4911	66.5101	66.5146	66.5254	66.5333
	24	17.1312	18.9296	24.3400	53.7880	57.5144	62.0793	62.2831	62.3175	62.3549	62.3695
	21	10.0949	10.9840	13.9566	42.1444	50.2213	58.0876	58.6064	58.8038	58.8448	58.8906
	18	5.6054	6.0219	7.1612	23.6033	34.9921	50.7447	51.8681	52.9925	53.0923	53.1483
	15	3.2045	3.4364	3.8856	11.1074	21.3656	42.3169	45.8479	47.2915	47.6224	47.7762
50	50	87.8111	87.8111	87.8111	87.8111	87.8111	86.3530	86.3530	86.3530	86.3530	86.3530
	45	44.2449	50.4143	62.6988	81.7618	82.6872	83.2192	83.2234	83.2296	83.2262	83.2288
	40	25.1410	28.0188	40.8431	72.2103	74.8617	81.1335	81.1444	81.1463	81.1460	81.1487
	35	14.8089	16.9003	23.4436	59.9691	65.4054	75.1813	75.1798	75.2028	75.2117	75.2220
	30	8.5235	9.5795	12.2554	44.1061	55.2855	68.5623	68.5733	68.6154	68.6261	68.6331
	25	4.6548	5.0317	6.0886	21.9632	38.1281	65.4674	65.7396	65.8099	65.8367	65.85307

reduced but not as much as $\alpha_0 = -3$. By comparing these two different values of α_0 in removal mechanism, it can be found out that choosing larger positive values for the intercept of the removal mechanism can minimize the effect of FD_i s. As shown in Figure 3(a), assuming the Weibull distribution with $\beta < 1$ under Type-II PCRD results in larger variations in the FD_i values which could be both very small and very large. Mostly, the smaller FD_i s happen in the beginning of the experiment, therefore, choosing the small values of α_0, causes more units to be removed later during the experiment which permits the mechanism to work and causes a reduction of experiment time. By choosing $\alpha_0 \geq 0$, even when the number of test units n is large, most units would be removed at the early of stages and an experiment would resemble a complete sampling test. Therefore, the removal mechanism could not be efficient and the reduction in the experiment time under Type-II PCRD could not be significant.

Table 4 gives the EET for the Weibull distributed Type-II PCRD sample with $\beta = 2$ assuming $\alpha_0 = -1.5$ and -0.5. The reduction in the experiment time comparing with complete sampling ($m = n$) is significant but not as much as when $\beta < 1$. Table 4 assuming $\alpha_0 = -1.5$ leads to smaller removal probabilities comparing with $\alpha_0 = -0.5$. Therefore, more units are stayed in the study and removed during the experiment which causes in a reduction of experiment time.

The results of the numerical studies show that the EET of a Type-II PCRD is highly influenced by the removal probability, p_i. Also, the values of tuning parameters α_0, α_1 and FD_i s are affecting p_i. Smaller values of the removal probability, p_i, means more units are stayed in the study and removed during the experiment that cause in a reduction of the experiment time. The shortest experiment time can happen when all units are removed at

Table 4. EET for the Weibull samples under Type-II PCRD for ($\beta = 2, \lambda = 6$) and $\alpha_0 = -1.5$ and -0.5 with different values of m and n.

		$\alpha_0 = -1.5$					$\alpha_0 = -.5$				
		α_1					α_1				
n	m	-4	-2	0	2	4	-4	-2	0	2	4
6	6	9.0761	9.0761	9.0761	9.0761	9.0761	9.1670	9.1670	9.1670	9.1670	9.1670
	5	7.1758	7.2868	7.8839	8.7054	8.7339	7.2569	7.4594	8.2366	8.5991	8.6082
	4	5.7373	5.8027	6.4021	8.0229	8.1154	5.8648	6.0398	7.1825	8.1054	8.1386
	3	4.5587	4.5977	5.0105	7.0422	7.2743	4.5884	4.6865	5.5967	7.1738	7.2685
20	20	11.1901	11.1901	11.1901	11.1901	11.1901	11.1897	11.1897	11.1897	11.1897	11.1897
	18	10.0021	10.3945	10.8417	11.0247	11.0340	10.7560	10.9353	11.0186	11.0401	11.0426
	16	8.7835	9.3659	10.3248	10.8035	10.8286	10.0622	10.4942	10.7701	10.8307	10.8370
	14	7.6006	8.1907	9.5587	10.5175	10.5614	9.0172	9.7764	10.4197	10.5513	10.5622
	12	6.5163	7.0118	8.5775	10.1991	10.2776	7.8498	8.7737	10.0029	10.2742	10.2942
	10	5.5073	5.8778	7.3017	9.7597	9.9077	6.5340	7.4479	9.3288	9.9226	9.9601
30	30	11.8324	11.8324	11.8324	11.8324	11.8324	11.8404	11.8404	11.8404	11.8404	11.8404
	27	11.1548	11.4284	11.6386	11.7028	11.7103	11.6287	11.6634	11.6848	11.6950	11.6967
	24	10.1216	10.7492	11.3250	11.5199	11.5375	11.2568	11.3862	11.4515	11.4777	11.4824
	21	8.9213	9.6742	10.7634	11.2186	11.2548	10.6686	11.0277	11.1910	11.2468	11.2552
	18	7.5800	8.4395	10.0317	10.9553	11.0221	9.7027	10.4221	10.8616	10.9681	10.9839
	15	6.2911	6.9587	8.8011	10.5198	10.6331	8.2731	9.4182	10.4173	10.6456	10.6715
50	50	12.5977	12.5977	12.5977	12.5977	12.5977	12.5998	12.5998	12.5998	12.5998	12.5998
	45	12.3633	12.4225	12.4571	12.4721	12.4772	12.4759	12.4814	12.4878	12.4923	12.4935
	40	11.8786	12.0807	12.1929	12.2350	12.2479	12.1724	12.1880	12.2044	12.2153	12.2185
	35	11.1554	11.6260	11.9475	12.0525	12.0763	11.9272	11.9667	11.9995	12.0214	12.0275
	30	9.8828	10.7556	11.4978	11.7559	11.7985	11.5882	11.6954	11.7663	11.8058	11.8160
	25	8.2456	9.3607	10.7999	11.3982	11.4776	10.9170	11.2566	11.4254	11.4999	11.5178

the end of experiment (Type-II censoring). Table 5 shows the behavior of EET for some other values of model parameters, assuming $\beta = 1$ and $\lambda = 5$.

Figure 4 plots the REET against m for various values of n and α_1 when $\alpha_0 = -3, 0$ and $\beta = 0.5$. As it is shown in this figure, the smallest REET under Type-II PCRD is related to the value of $\alpha_1 = -4$ for all α_0 values and sample sizes. We also observe when the value of α_0 increases, the value of α_1 does not seem to have any effect on the mechanism for large sample sizes n (the two right panels of Figure 4b). This is because the larger values of α_0 cause more units to be removed earlier so that the EET is very close to complete sampling plan, as were explained before for the results of Table 3.

Also, REET plots for different sample sizes, α_1, $\alpha_0 = -1.5, 2$ and $\beta = 2$ are shown in Figure 5. This figure, similar to Figure 4, shows that the reduction of REET is significant for smaller values of α_1. Actually, by increasing α_1, the EET of Type-II PCRD becomes close to the complete sampling scheme. Figure 5(b) illustrates that assuming $\alpha_0 > 0$, the random removal mechanism could be effective just for tiny sample sizes as were explained before.

Figure 6 shows the REET plots against m for $\beta = 0.5$. It is seen that the value of $\alpha_0 = -3$ has the smallest REET for all sample sizes. The REET increases with increasing the value of α_0. The REET assuming $\alpha_0 = 0$ and 2 are nearly the same for moderate and large sample sizes (i.e. $n = 20, 30$), see two left panels of Figure 6. The reason is the removal of the units earlier in the experiment, such as a complete experiment. The right panel of Figure 6 shows that increasing the value of α_1 leads to smaller slop of REET plot for all α_0 values when other parameters are fixed. This is correct for the other panels of Figure 6 as well.

Table 5. EET for the Weibull samples under Type-II PCRD for ($\beta = 1, \lambda = 5$) and different values of α_0, α_1, m and n.

			α_1				
n	m	α_0	−4	−2	0	2	4
6	6		12.35210	12.35210	12.35210	12.35210	12.35210
	5	1	9.53668	10.28604	11.08475	11.18503	11.23548
	4		6.99684	7.92991	9.68250	9.99193	10.03603
	3		4.47774	5.24663	7.27729	8.03784	8.15290
20	20		17.96959	17.96959	17.96959	17.96959	17.96959
	18		17.17592	17.33489	17.38240	17.39880	17.40540
	16	0	16.05582	16.51136	16.69780	16.74063	16.75696
	14		14.53925	15.53499	16.06286	16.15287	16.18490
	12		12.11501	13.77618	14.93360	15.12191	15.17914
	10		9.46655	11.40311	13.65423	14.05110	14.15376
30	30		19.80353	19.80353	19.80353	19.80353	19.80353
	27		18.90899	19.25191	19.43790	19.47056	19.48509
	24	−1	17.23654	18.18912	18.74717	18.86228	18.89916
	21		14.68193	16.36693	17.64376	17.90741	17.97806
	18		11.63404	13.73631	16.32719	16.94774	17.07330
	15		8.53246	10.61312	14.42099	15.85595	16.08044
50	50		22.49356	22.49356	22.49356	22.49356	22.49356
	45		20.47848	21.22988	21.73539	21.88876	21.92200
	40	−2	17.52288	18.98297	20.54655	21.00463	21.08601
	35		14.15624	16.18088	18.87284	20.16583	20.34438
	30		10.64101	12.55475	16.28568	18.87956	19.23195
	25		7.64604	9.17252	12.73309	17.53305	18.22173

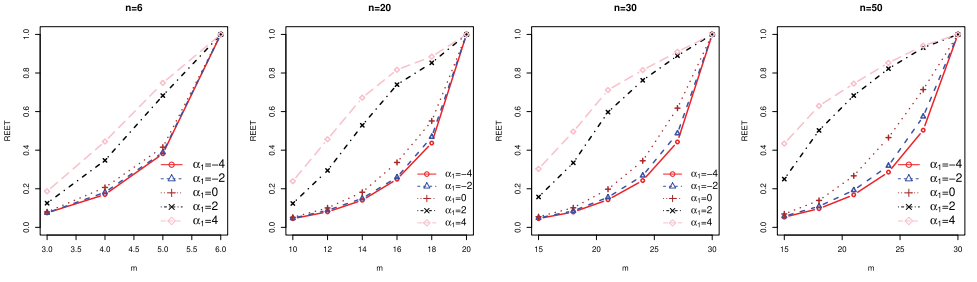
6. Sensitivity analysis

In this section, a sensitivity analysis is conducted to analyze the effect of misspecification of the tuning parameter values on experiment design. The importance of such a study follows from the fact that the design of an experiment depends on the choice of input tuning parameter values. The Relative Efficiency (RE) measure of a set of parameter values $\theta^{(1)}$ compared with the set of true parameter values $\theta^{(0)}$ is defined as

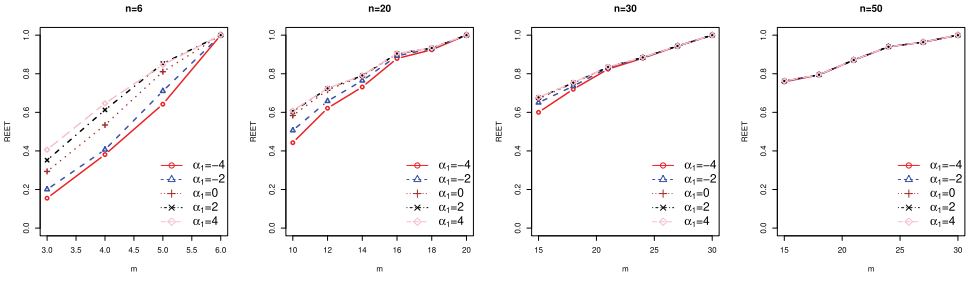
$$RE(\theta) = \frac{E(T_m) \text{ under } \theta^{(0)}}{E(T_m) \text{ under } \theta^{(1)}}.$$

In Type-II PCRD design, we conduct the sensitivity analysis with respect to EET by varying marginally α_1 values in the random removal mechanism. Actually, among the set of input parameter values, θ, the parameter that would be changed marginally is α_1. The RE values measure the sensitivity of the mechanism, where values closer to 1 represent less sensitivity of the mechanism due to misspecification of input parameter values. Here, we want to show that the proposed $\alpha_i, i = 0, 1$ values for moderate FD_i s according to Table 1(b) are effective.

Note, in Type-II progressive scheme with fixed number of removals, $\underline{R}_1^* = (n - m, 0, \ldots, 0)$ and $\underline{R}_2^* = (0, \ldots, 0, n - m)$ are the vector of removals corresponding to minimum and maximum duration of experiment time, where all removals are placed on the first and last stages, respectively. Analogously, in the Type-II PCRD scheme, most of the units are expected to be removed at the first or last stages of the experiment to have a minimum or maximum of $E(T_m)$, respectively. This is confirmed both by numerical results and Bhattacharya's work (2020) [12]. The EET is an increasing

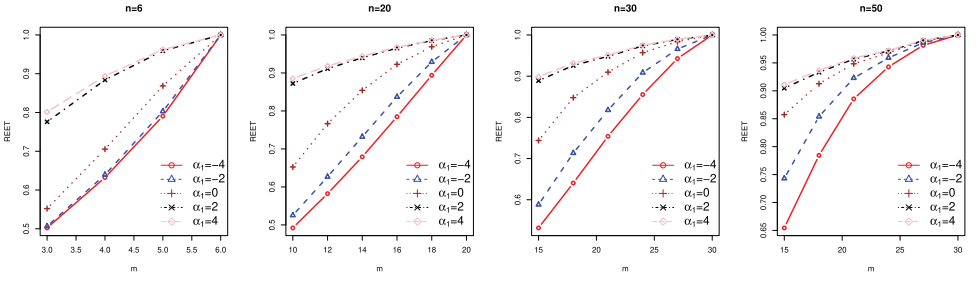


(a) REET plot for $\alpha_0 = -3$.

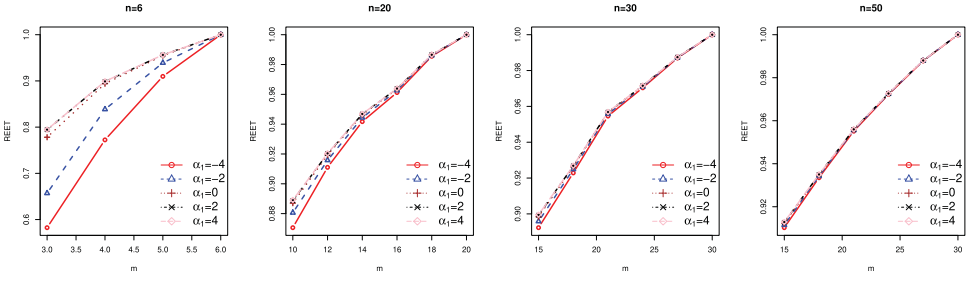


(b) REET plot for $\alpha_0 = 0$.

Figure 4. REET plots when $\beta = 0.5$ assuming different values of n, m, α_0 and α_1: (a) REET plot for $\alpha_0 = -3$ and (b) REET plot for $\alpha_0 = 0$.

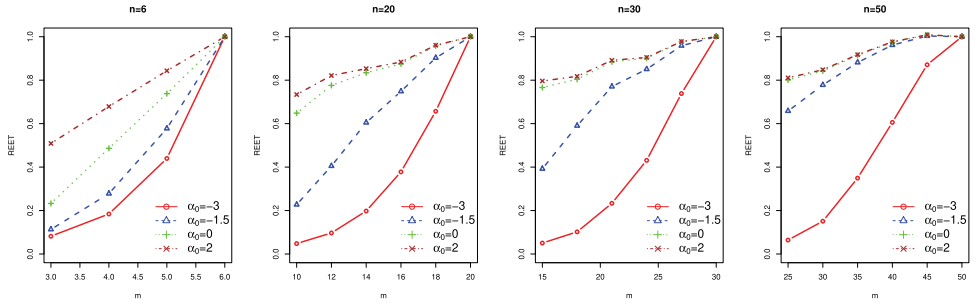


(a) REET plot for $\alpha_0 = -1.5$.

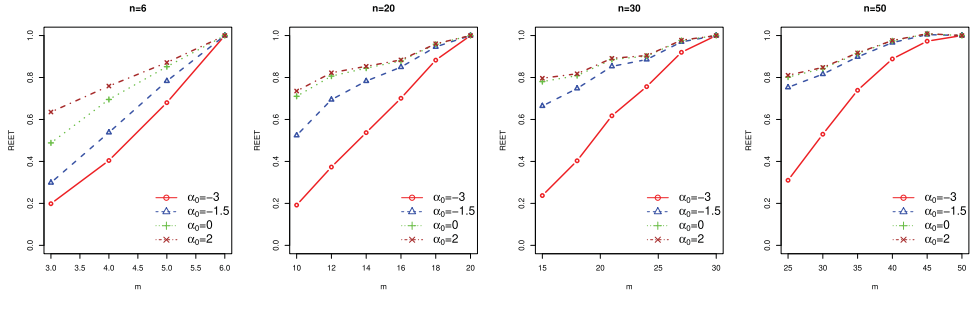


(b) REET plot for $\alpha_0 = 2$.

Figure 5. REET plots when $\beta = 2$ assuming different values of n, m, α_0 and α_1: (a) REET plot for $\alpha_0 = -1.5$ and (b) REET plot for $\alpha_0 = 2$.



(a) REET plot for $\alpha_1 = -4$.



(b) REET plot for $\alpha_1 = 0$.

Figure 6. REET plots when $\beta = 0.5$ assuming different values of n, m, α_0 and α_1: (a) REET plot for $\alpha_1 = -4$ and (b) REET plot for $\alpha_1 = 0$.

function of $\alpha_i, i = 0, 1$. That is, increasing (decreasing) $\alpha_i, i = 0, 1$, result in increasing (decreasing) EET. Hence, in Type-II PCRD an approximate boundary for $E(T_m)$ can be determined.

To perform sensitivity analysis, 1000 samples of size 10 are simulated from the Weibull distribution with the vector of parameters $(\lambda = 6, \beta = 0.75)$. More than %60 of the generated FD_i's are in moderate category. Therefore, according to Table 1(b) the appropriate $\alpha_i, i = 0, 1$ values can be considered $\alpha_0 = -1$ and $\alpha_1 \in (-6, 6)$. Then the sensitivity analysis is performed for 1000 simulated the Weibull samples of size $(n = 20, m = 10)$ under Type-II PCRD scheme with the above-mentioned tuning parameters.

Table 6 shows the results of sensitivity analysis for different α_1 values assuming $\alpha_0 = -1$. In Table 6, the EET, *RE* and the removal vector, $\underline{R}$, are reported in three parts, A, B and C. In the part B of Table 6, the appropriate α_1 is selected according to Table 1(b). The *RE* values are not close to 1, which represent the sensitivity of the mechanism due to changing α_1 values. This sensitivity is also evident according to the significant differences of EET values. Selecting different α_1 values in this interval leads to a new random removal vector. In parts A and C of Table 6, we have considered larger interval of α_1 values, which show that $RE(\alpha_1)$ values are not significant (*RE* are close to 1) and the EETs are very close to each other. Consequently, this sensitivity analysis approximately confirms proposed α_1 values for moderate FD_i in Table 1(b).

Table 6. Sensitivity analysis for assessment of effective range of possible α_1 values assuming $\alpha_0 = -1$ for the Weibull distributed data with $(\lambda = 6, \beta = 0.75)$ under Type-II PCRD with $(n = 20, m = 10)$.

Part	$\alpha_1^{(0)}$	$\alpha_1^{(1)}$	$E(T_m)$	$RE(\alpha_1)$	$\underline{R} = (R_1, \ldots, R_m)$
		−20	4.74019	1.07587	(1, 1, 0, 0, 0, 0, 0, 0, 0, 8)
		−15	5.23074	0.97497	(1, 1, 1, 0, 1, 0, 0, 0, 0, 6)
A	−12	−12	5.09981	1.00000	(0, 1, 0, 1, 1, 0, 0, 0, 0, 7)
		−10	6.00015	0.84995	(1, 1, 1, 1, 1, 0, 0, 0, 0, 5)
		−8	5.95341	0.85662	(1, 1, 1, 0, 1, 1, 0, 0, 0, 5)
		−6	8.32877	1.48219	(2, 2, 1, 1, 1, 0, 0, 0, 0, 3)
		−2	9.74439	1.26686	(2, 1, 1, 1, 1, 1, 1, 0, 0, 2)
B	−1	−1	12.34480	1.00000	(2, 2, 2, 1, 1, 0, 0, 0, 0, 2)
		1	21.74714	0.56765	(3, 4, 1, 1, 0, 0, 0, 0, 0, 1)
		2	22.30385	0.55348	(4, 2, 1, 1, 1, 0, 1, 0, 0, 0)
		6	23.72323	0.52037	(6, 2, 1, 1, 0, 0, 0, 0, 0, 0)
		8	23.69925	1.02033	(7, 2, 1, 0, 0, 0, 0, 0, 0, 0)
		10	24.18103	1.00345	(8, 1, 1, 0, 0, 0, 0, 0, 0, 0)
C	12	12	24.26439	1.00000	(8, 1, 1, 0, 0, 0, 0, 0, 0, 0)
		15	24.25823	1.00025	(8, 2, 0, 0, 0, 0, 0, 0, 0, 0)
		20	24.33276	0.99719	(8, 2, 0, 0, 0, 0, 0, 0, 0, 0)

7. Data analysis

In this section, to illustrate the properties of our proposed random removal mechanism, we analyze two real data sets. The aim of the first data analysis is to compare the performance of Type-II PCRD with other Type-II PCR schemes. The second data set is analyzed to illustrate different scenarios of generating removal vectors by choosing various tuning parameters according to the possible FD_i s.

The first data set is about the endurance of 23 deep groove ball bearings considered by Lawless [26] and Singh *et al.* [34]. Singh *et al.* [34] fitted four lifetime distributions for this data set which one of them was the Weibull distribution with vector of parameters $(\lambda = 82, \beta = 2)$. They showed that the Weibull distribution fits well to this real data set. According to Section 5.2, the Weibull distributed data with these parameters are homogenous and FD_i s fall in large category because of large value of scale parameter, λ. Therefore, small α_1 values (i.e. the values presented in Table 1c) are appropriate for random removal mechanism. On the other hand, we prefer smaller removal probabilities for having smaller EET. Therefore, $\alpha_1 = -0.5$ and $\alpha_0 = 0$ are chosen according to Table 1(c). We compare the performance of Type-II PCRD with other Type-II PCR schemes through generating 10,000 sets of removal vectors under Type-II PCRD assuming $(n, m) = (23, 8)$. The total test time is examined by implementing random removal vectors on the endurance of deep groove ball bearings data set. Table 7 shows the Type-II progressive censoring sample and the random removal vectors of the endurance of deep groove ball bearings for different approaches. The results show that the termination time of the experiment under Type-II PCRD is reduced. However, changing the tuning parameters can also significantly decrease the total time of the experiment.

The second analyzed data set is about survival times (in days) of 26 ovarian cancer patients after their surgical treatment. The data set is taken from Collett [15] which is also discussed by Singh *et al.* [35]. For illustrating the different scenarios of generating removal vectors by the various tuning parameters according to the purpose of research, let us consider the Type-II PCRD sample of $(n, m) = (26, 12)$. As stated in Sections 2.2

Table 7. Comparing the performances of different random removal mechanisms for the endurance of deep groove ball bearings data set.

Approach	Parameters	$\underline{T} = (T_1, \ldots, T_m)$	$\underline{R} = (R_1, \ldots, R_m)$
Type-II PCRD	$\alpha_0 = 0, \alpha_1 = -.5$	(17.88, 28.92, 33.00, 41.52, 42.12, 48.40, 51.96, 55.56)	(0, 0, 2, 0, 5, 1, 1, 6)
Type-II PCR	Binomial $p = 0.25$	(17.88, 28.92, 33.00, 42.12, 48.40, 55.56, 67.80, 68.64)	(4, 3, 2, 2, 1, 1, 0, 2)
Type-II PCR	Binomial $p = 0.5$	(17.88, 33.00, 41.52, 48.40, 55.56, 68.64, 105.84, 128.04)	(8, 4, 2, 1, 0, 0, 0, 0)
Type-II PCR	Binomial $p = 0.9$	(17.88, 33.00, 45.60, 51.96, 67.80, 68.64, 105.84, 128.04)	(11, 3, 1, 0, 0, 0, 0, 0)
Type-II PCR	Uniform	(17.88, 33.00, 41.52, 51.84, 55.56, 68.60, 105.12, 127.92)	(8, 4, 2, 1, 0, 0, 0, 0)

Table 8. Type-II PCRD data ($\underline{T}$) and random removal vectors ($\underline{R}$) obtained from the ovarian cancer data assuming $\alpha_0 = -1$ and different small values of $|\alpha_1|$.

α_1	$\underline{T} = (T_1, \ldots, T_m)$	$\underline{R} = (R_1, \ldots, R_m)$
0.05	(59, 268, 329, 353, 421, 464, 563, 638, 744, 855, 1040, 1206)	(12, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)
0.005	(59, 156, 268, 329, 353, 377, 421, 475, 563, 638, 803, 1040)	(3, 2, 2, 1, 1, 1, 1, 2, 0, 0, 0, 1)
−0.05	(59, 115, 156, 268, 329, 353, 365, 377, 431, 448, 477, 563)	(0, 0, 1, 0, 0, 1, 2, 1, 1, 1, 1, 6)
−0.25	(59, 115, 156, 268, 329, 353, 365, 377, 421, 432, 448, 465)	(0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 13)

and 2.3, determining values of tuning parameters depends on the FD_i s, which are large in this data set. To have different removal probabilities, the large FD_i s have been adjusted by choosing small values of α_1 and negative value for α_0 according to Table 1(c). Since most of FD_i s are too large in this experiment, selecting small $|\alpha_1|$ values allow the mechanism to be dependent on the FD_i s. Therefore, four values are chosen for α_1 in the interval $(-0.25, 0.05)$. As shown in Table 8, selecting smaller $|\alpha_1|$ values allows the mechanism to determine the number of removals at each stage according to failure distances during the experiment, while large $|\alpha_1|$ values resemble the experiment to a complete or Type-II sampling test.

8. Conclusion

In this article, the analysis of a Type-II progressive censoring sample of lifetime data with a novel GLM-based random removal mechanism, where the number of drop-outs at each failure time follows a conditional binomial distribution with dependent success probabilities, is proposed. Actually, for the large flexible PHR family of lifetime distributions, this stochastic removal mechanism has the flexibility of updating the removal probability at each stage via adjusting some tuning parameters. Appropriate GLM (i.e. *logit, probit* and *c log–log*) links are considered to define this random removal mechanism. On the other hand, the number of removals in this mechanism can be determined based on flexible unknown relation of failure distance and covariate information according to the goals of study and possible FD_i category. This stochastic removal mechanism leads to more flexibility in removal patterns which might seem necessary in some applications for cost and time considerations or withdrawal of units.

We have compared the properties of maximum likelihood estimators (MLEs) for the Weibull samples under proposed Type-II PCRD and other Type-II PCR schemes. The simulation results show that MLEs are expected to have similar properties (based on bias, MSE and Coverage Rate) applying different kinds of Type-II progressive censoring schemes.

Also, we have computed the EET under Type-II PCRD. Due to complexity of the EET calculation in all kinds of random censoring plans, the Monte Carlo method is used to

approximate these values. The numerical results indicate that the Type-II PCRD could perform more efficiently compared with the other random censoring based on uniform and binomial distributions. More specifically, the EET due to random removals dependent on the failure distances could considerably be decreased assuming a reverses relationship between the number of removals and the failure distances. Equivalently, the results of REET suggest that the experiment times are dictated by tuning parameters and FD_i s; specially, when FD_i s are fixed, choosing negative values of α_1 leads to smaller experiment times.

Two important advantages of the proposed method are: (i) the implementation of GLM-based random removal mechanism is applicable for a large class of lifetime distributions and (ii) the proposed random removal mechanism includes a set of tuning parameters that could be adjusted by the researcher, which allows both increasing or decreasing impact of pervious failure distances (according to the study protocols) as well as other available information.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] A. Agresti, *Categorical Data Analysis*, 2nd Edition, John Wiley & Sons, Inc., Hoboken, NJ, 2002.
[2] J. Ahmadi, M.J. Jozani, *É.* Marchand, and A. Parsian, *Bayes estimation based on k-record data from a general class of distributions under balanced type loss functions*, J. Stat. Plan. Inference 139 (2009), pp. 1180–1189.
[3] J. Ahmadi, M.J. Jozani, *É.* Marchand, and A. Parsian, *Prediction of k-records from a general class of distributions under balanced type loss functions*, Metrika 70 (2009), pp. 19–33.
[4] Z.H. Amin, *Bayesian inference for the Pareto lifetime model under progressive censoring with binomial removals*, J. Appl. Stat. 35 (2008), pp. 1203–1217.
[5] N. Balakrishnan, N. Kannan, C.T. Lin, and H.T. Ng, *Point and interval estimation for Gaussian distribution, based on progressively type-II censored samples*, IEEE Trans. Reliab. 52 (2003), pp. 90–95.
[6] N. Balakrishnan, N. Kannan, C.T. Lin, and S.J.S. Wu, *Inference for the extreme value distribution under progressive type-II censoring*, J. Stat. Comput. Simul. 74 (2004), pp. 25–45.
[7] N. Balakrishnan, *Progressive censoring methodology: an appraisal*, TEST 16 (2007), pp. 211.
[8] N. Balakrishnan, F. Su, and K.Y. Liu, *Exact likelihood inference for k exponential populations under joint progressive type-II censoring*, Commun. Stat. Simul. Comput. 44 (2015), pp. 902–923.
[9] N. Balakrishnan and R. Aggarwala, *Progressive Censoring: Theory, Methods, and Applications*, Springer Science & Business Media, Boston, 2000.
[10] N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, Springer, New York, 2014.
[11] N. Balakrishnan and R.A. Sandhu, *A simple simulational algorithm for generating progressive type-II censored samples*, Am. Stat. 49 (1995), pp. 229–230.
[12] R. Bhattacharya, *Implementation of compound optimal design strategy in censored life-testing experiment*, TEST 29, (2020), pp. 1–22.
[13] M. Burkschat, E. Cramer, and U. Kamps, *On optimal schemes in progressive censoring*, Stat. Probab. Lett. 76 (2006), pp. 1032–1036.
[14] M. Burkschat, E. Cramer, and U. Kamps, *Optimality criteria and optimal schemes in progressive censoring*, Commun. Stat. Theory Methods 36 (2007), pp. 1419–1431.
[15] D. Collett, *Modelling Survival Data in Medical Research*, 3rd ed., Champman and Hall/ CRC Press, Bristol, 2015.

[16] E. Cramer and M. Ensenbach, *Asymptotically optimal progressive censoring plans based on fisher information*, J. Stat. Plan. Inference 141 (2011), pp. 1968–1980.
[17] E. Cramer and G. Iliopoulos, *Adaptive progressive type-II censoring*, TEST 19 (2010), pp. 342–358.
[18] S. Dey, T. Kayal, and Y.M. Tripathi, *Statistical inference for the weighted exponential distribution under progressive type-II censoring with binomial removal*, Am. J. Math. Manag. Sci. 37 (2018), pp. 188–208.
[19] S. Dey and T. Dey, *Statistical inference for the Rayleigh distribution under progressively type-II censoring with binomial removal*, Appl. Math. Model. 38 (2014), pp. 974–982.
[20] M. Ghahramani, M. Sharafi, and R. Hashemi, *Analysis of the progressively type-II right censored data with dependent random removals*, J. Stat. Comput. Simul. 90 (2020), pp. 1001–1021.
[21] D.I. Gibbons and L.C. Vance, *Estimators for the 2-parameter Weibull distribution with progressively censored samples*, IEEE Trans. Reliab. 32 (1983), pp. 95–99.
[22] S. Gunasekera, *Inference for the Burr XII reliability under progressive censoring with random removals*, Math. Comput. Simul. 144 (2018), pp. 182–195.
[23] R.C. Gupta, P.L. Gupta, and R.D. Gupta, *Modeling failure time data by Lehman alternatives*, Comm. Stat. Theory Methods 27 (1998), pp. 887–904.
[24] R.D. Gupta and D. Kundu, *Exponentiated exponential family: an alternative to gamma and Weibull distributions*, Biom. J. 43 (2001), pp. 117–130.
[25] P. Kundu and A.K. Nanda, *Reliability study of proportional odds family of discrete distributions*, Comm. Stat. Theory Methods 47 (2018), pp. 1091–1103.
[26] J.F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley & Sons, New York, 1982.
[27] N. Metropolis and S. Ulam, *The Monte Carlo method*, J. Am. Stat. Assoc. 44 (1949), pp. 335–341.
[28] S. Mondal and D. Kundu, *Inference on Weibull parameters under a balanced two–sample type II progressive censoring scheme*, Qual. Reliab. Eng. Int. 36 (2020), pp. 1–17.
[29] H.K.T. Ng and P.S. Chan and N. Balakrishnan, *Optimal progressive censoring plans for the Weibull distribution*, Technometrics 46 (2004), pp. 470–481.
[30] G. Psarrakos and M.A. Sordo, *On a family of risk measures based on proportional hazards models and tail probabilities*, Insur. Math. Econ. 86 (2019), pp. 232–240.
[31] A. Rasouli and N. Balakrishnan, *Exact likelihood inference for two exponential populations under joint progressive type-II censoring*, Commun. Stat. Theory Methods 39 (2010), pp. 2172–2191.
[32] S. Sel, M. Jung, and Y. Chung, *Bayesian and maximum likelihood estimations from parameters of McDonald extended Weibull model based on progressive type-II censoring*, J. Stat. Theory Pract. 12 (2018), pp. 231–254.
[33] M. Sharafi, *Inference of the two-parameter Lindley distribution based on progressive type II censored data with random removals*, Comm. Stat. Simul. Comput. 51:4 (2019), pp. 1–15.
[34] S.K. Singh, U. Singh, and V.K. Sharma, *Expected total test time and Bayesian estimation for generalized Lindley distribution under progressively type-II censored sample where removals follow the beta-binomial probability law*, Appl. Math. Comput. 222 (2013), pp. 402–419.
[35] S.K. Singh, U. Singh, and M. Kumar, *Bayesian estimation for Poisson-exponential model under progressive type-II censoring data with binomial removal and its application to ovarian cancer data*, Comm. Stat. Simul. Comput. 45 (2016), pp. 3457–3475.
[36] S.K. Tse, C. Yang, and H.K. Yuen, *Statistical analysis of Weibull distributed lifetime data under type II progressive censoring with binomial removals*, J. Appl. Stat. 27 (2000), pp. 1033–1043.
[37] S.K. Tse and L. Xiang, *Interval estimation for Weibull-distributed life data under type II progressive censoring with random removals*, J. Biopharm. Stat. 13 (2003), pp. 1–16.
[38] S.K. Tse and C. Yang, *Reliability sampling plans for the Weibull distribution under type II progressive censoring with binomial removals*, J. Appl. Stat. 30 (2003), pp. 709–718.
[39] S.K. Tse and H.K. Yuen, *Expected experiment times for the Weibull distribution under progressive censoring with random removals*, J. Appl. Stat. 25 (1998), pp. 75–83.
[40] J. Von Neumann, *Various techniques used in connection with random digits*, John von Neumann, Collected Works 5 (1963), pp. 768–770.

[41] S.J. Wu and S.R. Huang, *Optimal progressive group-censoring plans for exponential distribution in presence of cost constraint*, Stat. Pap. 51 (2010), pp. 431–443.
[42] C.C. Wu and S.F. Wu and H. Y. Chan, *MLE and the estimated expected test time for the two-parameter Gompertz distribution under progressive censoring with binomial removals*, Appl. Math. Comput. 181 (2006), pp. 1657–1670.
[43] W. Yan, Y. Shi, B. Song, and Z. Mao, *Statistical analysis of generalized exponential distribution under progressive censoring with binomial removals*, J. Syst. Eng. Electron. 22 (2011), pp. 707–714.
[44] H.K. Yuen and S.K. Tse, *Parameters estimation for Weibull distributed lifetimes under progressive censoring with random removals*, J. Stat. Comput. Simul. 55 (1996), pp. 57–71.