Optimal Nonparametric Quantile Estimation under Progressive Type-II Censoring

David Han*

Abstract
The optimal progressive censoring schemes are examined for the nonparametric confidence intervals of population quantiles. The results obtained can be universally applied to any continuous probability distribution. By using the interval mass as an optimality criterion, the optimization process is free of the actual observed values from the sample and needs only the initial sample size n and the number of complete failures m. Using several sample sizes combined with various degrees of censoring, the results of the optimization are presented here for the population median at selected levels of confidence (99%, 95% and 90%). With the optimality criterion under consideration, the efficiencies of the worst progressive Type-II censoring scheme and ordinary Type-II censoring scheme are also examined in comparison to the best censoring scheme obtained for fixed n and m.

Key Words: confidence interval, nonparametric inference, optimal censoring scheme, order statistic, progressive Type-II censoring, quantile

1. Introduction

In order to set a warranty period of a new product or even to compare alternative manufacturing designs, the estimation of quantiles is routinely performed in reliability and lifetime analysis. If one applies a parametric procedure to estimate a quantile, an important presumption underlying the method is that the model fits the data well. Unless this is verified in the very first stage of analysis, the inferential results may lose power considerably and lead the analyst to a severely distorted conclusion. One way to overcome this is to apply a nonparametric procedure which does not specify the model structure a priori so that the results of inference are free from violation of the model assumption.

Another common feature frequently encountered by statistical analysts and practitioners is censoring. Censored data arise when the experiments involving lifetimes of testing units have to be terminated earlier. Although a complete collection of data is the most favorable scenario prior to the actual analysis, for the reasons of cost reduction and time constraint, intentional censoring is unavoidable in practice, especially for reliability experiments. One general form of censoring considered in this paper is progressive Type-II right censoring, the importance of which lies in its efficient exploitation of the available resources compared to the traditional sampling. Withdrawn unfailed testing units can typically be used in other experiments in the same or at a different facility.

The focus of this paper is to review the procedure of constructing an exact nonparametric confidence interval for a population quantile of interest under progressive Type-II right censoring and to numerically investigate the associated problem of selecting the optimal censoring schemes using the expected interval mass as an optimality criterion. Recently, the problem of optimal scheduling and optimal censoring has received much attention in the reliability literature. Balakrishnan and Aggarwala (2000) have addressed this problem in general and investigated it using the trace and determinant functions based on the variance-covariance matrix of BLUEs (Best Linear Unbiased Estimators) as optimality criteria for several continuous parametric distributions including exponential, normal, extreme-value, and log-normal.

*Department of Management Science and Statistics, University of Texas, San Antonio, TX 78249

In this paper, we look at the case of progressive Type-II censoring in a nonparametric setting and discuss the optimal progressive censoring schemes in the construction of nonparametric confidence intervals for population quantiles. This work is based on the recent results of Guilbaud (2001, 2004) and Balakrishnan, Childs and Chandrasekar (2002) on some representations of distributions of progressively Type-II right censored order statistics from a continuous distribution.

2. Nonparametric Confidence Intervals for Quantiles

First, we illustrate a simple nonparametric procedure to construct a confidence interval for a given quantile based on complete observations ($viz.$, no censoring). It is assumed that the random sample is from a continuous distribution $F(t) = Pr[X \leq t]$, and ξ_p is the given p-quantile that satisfies

$$Pr[X \leq \xi_p] = F(\xi_p) = p, \qquad 0 < p < 1, \tag{1}$$

where X is a random variable whose cdf is $F(t)$. Then, suppose that a sample of size n is taken from this population and ordered so that $X_{1:n} < X_{2:n} < \ldots < X_{n:n}$ denote the order statistics from the sample. The probability of the interval $(-\infty, X_{j:n}]$ covering ξ_p is then given by the lower tail binomial probability,

$$Pr[X_{j:n} > \xi_p] = Pr[X_{j:n} \geq \xi_p] = \sum_{k=0}^{j-1} \binom{n}{k} p^k (1-p)^{n-k} = b_j, \tag{2}$$

for each $j = 1, 2, \ldots, n$. Using (2), for integers r and s which satisfy $1 \leq r < s \leq n$, the probability of $X_{r:n} \leq \xi_p \leq X_{s:n}$ is obtained as

$$Pr[X_{r:n} \leq \xi_p \leq X_{s:n}] = Pr[X_{s:n} \geq \xi_p] - Pr[X_{r:n} > \xi_p] = b_s - b_r = \sum_{k=r}^{s-1} \binom{n}{k} p^k (1-p)^{n-k}.$$

For a selected confidence coefficient $1 - \alpha$, the values of r and s can be searched so that $Pr[X_{r:n} \leq \xi_p \leq X_{s:n}] \geq 1 - \alpha$ holds. Substituting the observed values $x_{r:n}$ and $x_{s:n}$ for $X_{r:n}$ and $X_{s:n}$, the two-sided $100(1-\alpha)\%$ confidence interval for ξ_p is obtained as $[x_{r:n}, x_{s:n}]$. It should be noted that it is not always possible to find r and s to give the conventional 0.90, 0.95, 0.99 values for $1 - \alpha$. Especially with small sample sizes, it becomes more difficult to produce a confidence interval nonparametrically with a high level of confidence. Nevertheless, the confidence interval constructed as above is free of any particular form of the probability distribution $F(t)$ as long as the underlying parent distribution is continuous.

3. Progressive Type-II Censoring and Some Representations

The two traditional forms of censoring which have been studied extensively in the literature are Type-I and Type-II censoring. Type-I censoring occurs when the experiment is terminated at a prefixed time T, independent of the failure times. While Type-I censoring specifies the time of termination, Type-II censoring restricts the number of failures to be observed. As such, in Type-II right censoring, there would be a pre-fixed number m $(< n)$ so that the experiment is terminated at the time of the mth failure and all the remaining units are removed from the experiment.

In order to introduce flexibility, a more general type of censoring called *progressive censoring* has been discussed in the literature. Progressive censoring can also be of either

Type-I or Type-II, and these do in fact include the conventional Type-I and Type-II censoring as special cases. Progressive Type-I right censored samples are observed when a pre-specified number or proportion of unfailed units are continuously removed during the experiment at pre-specified time points. Similarly, progressive Type-II right censored samples arise when a pre-specified number of surviving units are continuously withdrawn from the experiment at each observed failure time until the pre-fixed number of units have failed.

In this paper, we focus on progressive Type-II right censoring as it possesses more tractable and interesting mathematical properties. Consider a life-testing experiment involving n experimental units, and suppose m complete failures are to be observed for $2 \le m \le n$. Let $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ be the planned progressive censoring scheme to be adopted here, where $R_i \ge 0$ for $i = 1, 2, \ldots, m$ and

$$\sum_{i=1}^{m} R_i + m = n.$$

As mentioned before, the conventional Type-II right censoring is a special case when $R_1 = R_2 = \cdots = R_{m-1} = 0$ and $R_m = n - m > 0$, while the complete sample case ($viz.$, no censoring) corresponds to the case with $m = n$ and $R_1 = R_2 = \cdots = R_m = 0$. The failure times of the testing units since time zero can be viewed as a random sample of size n from cdf $F(t)$, and the corresponding order statistics of the successive failure times are denoted as before by $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$. As the parent distribution $F(t)$ is assumed to be continuous, $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ are distinct with probability 1. Suppose these order statistics are not all observable, and that the implemented progressive censoring leads to m observable uncensored order statistics $Y_1 < Y_2 < \cdots < Y_m$ that are available for inference.

Guilbaud (2001, 2004) has then shown that each observed progressively Type-II right censored order statistic in $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_m)^T$ can be represented as a mixture of underlying ordinary order statistics $\mathbf{X} = (X_{1:n}, X_{2:n}, \ldots, X_{n:n})^T$. Let $\hat{w}_{i,j}$ be the indicator of the event $Y_i = X_{j:n}$, which equals 1 if the event occurs and 0 otherwise. Thus, $\hat{w}_{i,j}$'s simply tell which of the order statistics in $\mathbf{X}$ is selected as one in $\mathbf{Y}$. Then, the $m \times n$ random matrix $\hat{W} = (\hat{w}_{i,j})$ composed of the indicators holds the relationship

$$\mathbf{Y} = \hat{W}\mathbf{X}, \tag{3}$$

with $\hat{w}_{1,1} = 1$ since $Y_1 = X_{1:n}$ by definition. Using the idea of sequential and independent simple random sampling without replacement, $\hat{W}$ can be further decomposed into a product of m matrices that are mutually independent of each other and of $\mathbf{X}$. If the expectation of $\hat{W}$ is denoted by $W = (w_{i,j})$, one can see that for each $i = 2, 3, \ldots, m$, $(\hat{w}_{i,1}, \hat{w}_{i,2}, \ldots, \hat{w}_{i,n})$ has multinomial distribution with parameters 1 and $(w_{i,1}, w_{i,2}, \ldots, w_{i,n})$. Given the independence of $\hat{W}$ and $\mathbf{X}$, the mixture representation in (3) thus holds with the weights $w_{i,j} = Pr[Y_i = X_{j:n}] = Pr[\hat{w}_{i,j} = 1] = E[\hat{w}_{i,j}]$ for $1 \le i \le m$ and $1 \le j \le n$, and the mixture weights are given by the non-random matrix $W = E[\hat{W}]$. It should be noted that W is completely determined by n, m and $\mathbf{R} = (R_1, R_2, \ldots, R_m)$, and the computation of all the elements of W can be carried out via an efficient recursive relation described in Guilbaud (2001).

Another representation that is computationally very efficient is due to Balakrishnan, Childs and Chandrasekar (2002). Adopting the usual conventions that $\prod_{i=1}^{0} u_i \equiv 1$ and $\sum_{i=1}^{0} u_i \equiv 0$, the density function of Y_r is given by

$$f_{Y_r}(y_r) = c_r \sum_{i=0}^{r-1} c_{i,r-1}(R_1 + 1, R_2 + 1, \ldots, R_{r-1} + 1) f(y_r) \Big(1 - F(y_r)\Big)^{R_i''-1}, \tag{4}$$

$-\infty < y_r < \infty$ for $r = 1, 2, \ldots, m$, where

$$\begin{aligned} c_r &= n(n - R_1 - 1) \cdots (n - R_1 - \cdots - R_{r-1} - r + 1), \\ R_i'' &= n - \sum_{j=1}^{r-i-1} (R_j + 1), \end{aligned}$$

and $f(t)$ denotes the pdf corresponding to $F(t)$. The coefficients $c_{i,r-1}(R_1+1, \ldots, R_{r-1}+1)$ in (4) are given by

$$c_{i,q}(\mathbf{u}_q) = \frac{(-1)^i}{\left(\prod_{j=1}^{i} \sum_{k=q-i+1}^{q-i+j} u_k\right)\left(\prod_{j=1}^{q-i} \sum_{k=j}^{q-i} u_k\right)},$$

defined for any real vector $\mathbf{u}_q = (u_1, u_2, \ldots, u_q)$ of length $q \geq 1$. Integrating the density function in (4), the corresponding distribution function of Y_r is obtained as

$$\begin{aligned} F_{Y_r}(y_r) = \int_{-\infty}^{y_r} f_{Y_r}(y_r)\, dy &= c_r \sum_{i=0}^{r-1} c_{i,r-1}(R_1 + 1, R_2 + 1, \ldots, R_{r-1} + 1) \\ &\quad \times \frac{1}{R_i''} \left(1 - \left(1 - F(y_r)\right)^{R_i''}\right), \end{aligned} \tag{5}$$

$-\infty < y_r < \infty$ for $r = 1, 2, \ldots, m$.

4. Confidence Intervals for Quantiles under Progressive Type-II Censoring

In Section 2, a procedure to construct a nonparametric confidence interval for the population quantile ξ_p was illustrated in the complete sample case. Using the mixture representation for $\mathbf{Y}$ given in (3) or the expression of the distribution function of Y_r in (5), a nonparametric confidence interval for ξ_p under progressive Type-II right censoring can be constructed in a similar way as described in Guilbaud (2001, 2004). Now, let $\mathbf{b} = (b_j)$ be the $n \times 1$ matrix whose elements are defined by (2) and $\mathbf{a} = (a_j)$ be the $m \times 1$ matrix whose elements are defined in terms of $W = (w_{i,j})$ through the relationship

$$\mathbf{a} = W\mathbf{b}. \tag{6}$$

Then, the vector $\mathbf{a}$ is simply a collection of the probabilities of covering ξ_p by each Y_r, since

$$Pr[Y_r \geq \xi_p] = Pr[Y_r > \xi_p] = \sum_{j=1}^{n} w_{r,j} Pr[X_{j:n} \geq \xi_p] = \sum_{j=1}^{n} w_{r,j} b_j = a_r, \tag{7}$$

for $r = 1, 2, \ldots, m$. It is clear that $a_1 \leq a_2 \leq \cdots \leq a_m$ as $Y_1 < Y_2 < \cdots < Y_m$ with probability 1. Now, suppose that r and s are some integers satisfying $1 \leq r < s \leq m$. The coverage probability of the interval estimator $[Y_r, Y_s]$ for ξ_p can then be easily expressed in terms of the elements of $\mathbf{a}$ as follows:

$$Pr[Y_r \leq \xi_p \leq Y_s] = Pr[Y_s \geq \xi_p] - Pr[Y_r > \xi_p] = a_s - a_r. \tag{8}$$

Once $\mathbf{a}$ has been evaluated using (6), the integers r and s can be determined so that the confidence level of this interval is at least a specified value $1 - \alpha$ (*i.e.*, $a_s - a_r \geq 1 - \alpha$).

Since the coverage probability in (7) is determined through the matrix operation in (6), the expression in (5) offers another efficient way to compute (7) without going through the direct computation of the matrix W. This is particularly useful when one wants to estimate only one or few elements of $\mathbf{a}$ and to avoid the intensive computation required for determining the whole matrix W. In this case, using (5), the coverage probability in (7) can be explicitly expressed as

$$\begin{aligned} Pr[Y_r \geq \xi_p] &= Pr[Y_r > \xi_p] = 1 - Pr[Y_r \leq \xi_p] = 1 - F_{Y_r}(\xi_p) \\ &= 1 - c_r \sum_{i=0}^{r-1} c_{i,r-1}(R_1+1, R_2+1, \ldots, R_{r-1}+1) \\ &\qquad \times \frac{1}{R_i''}\left(1-(1-p)^{R_i''}\right), \end{aligned}$$

and the last step follows directly from the definition of ξ_p in (1).

5. Optimal Progressive Censoring Schemes

5.1 Optimality Criterion and Optimal Schemes

In the previous section, the exact nonparametric confidence interval has been derived for any population quantile of interest based on a progressively Type-II right censored sample. Then, as pointed out by Balakrishnan and Aggarwala (2000), some natural questions arise here: "how can a practitioner decide on which censoring scheme to be used out of numerous censoring schemes?"; "Is the decision made strictly on the basis of convenience, or can one select a censoring scheme which makes the most sense within some statistical settings?". From a practical point of view, the question of choosing the optimal values for $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ is certainly indispensable and it has to be addressed when one designs a progressive Type-II censoring experiment as the number of distinct censoring schemes becomes very large even for moderate values of n and m.

Before selecting the optimal censoring scheme, one must first devise an optimality criterion or an objective function to be optimized, as done by Balakrishnan and Aggarwala (2000) in the case of point estimation. Consequently, the meaning of the optimal censoring scheme is restricted to the criterion of one's choice. In the case of nonparametric interval estimation for a quantile ξ_p with n and m fixed, a simple optimization with respect to the choice of $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ is to select $\mathbf{R}$ which enables to find r and s in (8) that satisfy $a_s - a_r \approx 1 - \alpha$ and $a_r \approx 1 - a_s$. This constraint is equivalent to finding r and s such that $a_r \approx \alpha/2$ and $a_s \approx 1 - \alpha/2$ in order to yield a symmetric confidence interval when possible.

In the complete sample case, a reasonable objective function for choosing r and s is the index difference $s - r$ corresponding to the interval $[X_{r:n}, X_{s:n}]$ with the level of confidence at least $1 - \alpha$. Minimizing this function is to minimize the expected probability mass $F(X_{s:n}) - F(X_{r:n})$ of the underlying distribution within the interval. This clearly reflects the purpose of the interval estimation, which is to produce the shortest interval with a desired confidence level. Under progressive Type-II right censoring, it follows from (3) that $Y_i = \sum_{j=1}^{n} \hat{w}_{i,j} X_{j:n}$ and the expectation of $F(Y_i) = \sum_{j=1}^{n} \hat{w}_{i,j} F(X_{j:n})$ is therefore equal to

$$e_i = E\left[F(Y_i)\right] = \sum_{j=1}^{n} w_{i,j}\, \frac{j}{n+1}, \tag{9}$$

for $i = 1, 2, \ldots, m$. An alternate expression for e_i in (9), obtained through a generalization of Malmquist's transformation, is simply given by

$$e_i = 1 - \prod_{j=m-i+1}^{m} \frac{\delta_j}{\delta_j + 1},$$

where

$$\delta_j = j + \sum_{k=m-j+1}^{m} R_k,$$

for $i = 1, 2, \ldots, m$. Hence, the expectation of the probability mass within the interval $[Y_r, Y_s]$ is simply equal to $e_s - e_r$ for $1 \le r < s \le m$. Then, for given n and m, it is quite reasonable to determine the optimal progressive Type-II censoring scheme for the nonparametric confidence interval for ξ_p by minimizing the expected interval mass

$$M(\mathbf{R}) = \min_{(r,s)\in\mathcal{S}} \{e_s - e_r\}, \tag{10}$$

where $\mathcal{S}$ is a subset of the binary Cartesian product of positive integers (r, s) such that $1 \le r < s \le m$ and $a_s - a_r \ge 1 - \alpha$. If $\mathcal{S}$ happens to be an empty set for some $\mathbf{R}$'s, $M(\mathbf{R})$ is simply defined to be 1. It should be noted that the objective function $M(\cdot)$ and the index set $\mathcal{S}$ are both depending on the choices of n, m, $\mathbf{R}$, and α for a given quantile ξ_p, but $M(\mathbf{R})$ is minimized with respect to every possible progressive censoring scheme $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ when n, m and α are all pre-fixed along with ξ_p. This provides great flexibility to the practitioners as the number of units to be put on the test and the number of complete failure times to be observed are both to be decided a priori by the experimenter based on the availability of units and experimental facilities. If one or both of these are to be determined in the planning stage, one may also use the tables presented here to decide upon the values of n and m which are feasible given an agreeable value of the objective function.

In this finite sample case, to minimize (10) with all the other values given, one may list every possible choice of censoring scheme and the corresponding value of the objective function. After determining the best value, the value that minimizes the expected interval mass $M(\cdot)$ in (10), or a certain region of satisfactory values from this list, one can pick out either the best censoring scheme or one that gives a value close to the best but may be practically more convenient ($i.e.$, a sub-optimal censoring scheme). For the purpose of comparing different censoring schemes with the selected objective function, a sensible definition of the efficiency of a censoring scheme $\mathbf{R}_A$ with respect to a scheme $\mathbf{R}_B$ can be given by

$$\mathit{Efficiency}(\mathbf{R}_A, \mathbf{R}_B) = \frac{M(\mathbf{R}_A)}{M(\mathbf{R}_B)} \times 100\%. \tag{11}$$

This is simply a ratio of the interval masses and if one is interested in searching for a region containing a number of satisfactory censoring schemes, such a region can be defined in terms of efficiencies, as done by Balakrishnan and Aggarwala (2000). For example, an experimenter may be pleased with any censoring scheme which is at least 95% as efficient as the optimal (best) scheme. Moreover, one can see that for any fixed values of n and m, the efficiency of the conventional Type-II censoring scheme with respect to the optimal censoring scheme would be at most 100%; consequently, by using the optimal censoring scheme, there is no loss in efficiency over the conventional one.

5.2 Numerical Study

A numerical study has been conducted with some selected values of parameters and the results are presented in Tables 1–10. In the original computation performed, the choices of the sample size n ranged from 10 to 100 with an increment of 5 and the pre-determined number of failure observations m ranged from 2 to n with a unit increment for each choice of n. In addition, the choices of m were restricted in such a way that the total number of available censoring schemes $\binom{n-1}{m-1}$ does not exceed 8.0×10^7 in order to keep the computational time and space manageable. To examine a variety of quantiles, p of ξ_p was selected from 0.05 to 0.95 with an increment of 0.05. Moreover, corresponding to the conventional 99%, 95% and 90% confidence intervals, α was chosen to be 0.01, 0.05 and 0.10, respectively.

Due to the constraint on space, only selected tables for the population median $\xi_{0.5}$ are presented here and the tables are intended to shed some light on the general behavior of this discrete optimization. With fixed n and m, each table lists the best progressive censoring scheme and the worst progressive censoring scheme along with conventional Type-II censoring scheme for the sake of comparison. The best censoring scheme is simply the one that minimizes the objective function $M(\mathbf{R})$ given in (10) and the worst censoring scheme is, on the other hand, the one that maximizes

$$\bar{M}(\mathbf{R}) = \max_{(r,s)\in\mathcal{S}} \{e_s - e_r\}.$$

In the tables, the meaning of $0 \star k$ with some positive integer k is to repeat zero k times. Hence, it simply denotes a zero vector of size k embedded in the censoring scheme $\mathbf{R}$, and the meaning of $1 \star k$ is to be interpreted in a similar way. In each table, count denotes the total number of distinct progressive censoring schemes which can produce at least one interval with the confidence level greater than or equal to the nominal level $1 - \alpha$. In other words, it is the number of censoring schemes for which $\mathcal{S}$ in (10) is a non-empty set. Since the nonparametric confidence interval depends on the order or the index of the uncensored observations, the confidence interval is given in the form of $[r, s]$ with the corresponding confidence interval being $[Y_r, Y_s]$. In each table, the actual level of confidence for each interval estimator is also provided along with its expected probability mass. Moreover, efficiency was calculated using (11) with respect to the best censoring scheme found for the pre-fixed values of n, m, p and α. Naturally, in cases where a certain censoring scheme can not yield any interval with the desired level of confidence, the efficiency is simply noted as not available.

6. Discussion

Based on the results of the numerical study presented in Section 5.2 and Tables 1–10, a number of interesting observations can be made. First of all, for certain choices of n and m, the available progressive censoring schemes which can produce confidence intervals for a given quantile with a desired level of confidence can be few. Moreover, the higher the required level of confidence is, the fewer the choices of such censoring schemes. One can also observe that for fixed n, m and α, the actual level of confidence increases initially and then decreases as p increases. Consequently, when the quantile of interest is too small or too large, there may not be any censoring schemes to generate a confidence interval with a selected level of confidence unless n and m are both reasonably large.

Then, how large should m be compared to n? It was found that the size of m related to n is also a significant factor to boost up the available censoring schemes which can yield confidence intervals with a desired level. By examining tables, it was observed that if m is too small compared to n or too close to n for fixed ξ_p and α, the number of censoring schemes which can yield confidence intervals with a desired level decreases. This in turn reduces the number of sub-optimal censoring scheme choices whose efficiencies are close to that of the best scheme. To one's surprise, this finding also implies that the case without any censoring actually performs worse than the case with an appropriate censoring while constructing a nonparametric confidence interval for a quantile. Therefore, the censoring proportion $\frac{n-m}{n}$ is an important factor to consider at the planning stage of a progressive Type-II censoring experiment if one wishes to construct a confidence interval with a desired coverage probability.

Another general observation relating to n and m is that raising n with m fixed enables to find the optimal censoring schemes for small quantiles like $\xi_{0.05}$ and $\xi_{0.10}$ within a limited range. On the other hand, increasing m with fixed n enables to find the optimal censoring schemes for a large range of quantiles of interest. Although it is not shown in the tables, it was observed that the maximum level of confidence could be achieved by the censoring scheme $\mathbf{R} = (n - m, 0 \star (m - 1))$ in any case where none of the censoring schemes examined could produce an interval with at least the nominal level. Thus, withdrawing every unit immediately after observing the first failure time is the best in a sense that the interval $[Y_1, Y_m]$ attains the highest confidence level. Nevertheless, it turned out to be also the worst censoring scheme in some situations when we could in fact locate the optimal censoring scheme. This is because $[Y_1, Y_m]$ bears the highest probability mass within the interval.

Although the tables presented here could be used for reference for practitioners who are designing a progressive Type-II censoring experiment, it is rather difficult to see whether there is a universal pattern or a mathematical trend of the best censoring schemes or the worst censoring schemes. However, it can be clearly pointed out that if one randomly chooses a progressive censoring scheme for an experiment in which a large censoring takes place, it will be less likely to obtain a desired nonparametric confidence interval for the quantile of interest. Hence, a careful planning of an experiment is important from an inferential point of view. Another interesting observation is that in a few cases, the best censoring scheme coincides with the conventional Type-II censoring scheme. For example, for $n = 25$ and $m = 20$, the conventional Type-II censoring scheme turns out to be the best censoring scheme to construct a 90% confidence interval for $\xi_{0.35}$. But, in most cases, the ordinary Type-II censoring scheme could not even produce a single confidence interval with an acceptable level and the efficiency gained by the optimal censoring scheme turns out to be quite substantial compared to the conventional Type-II censoring scheme as well as to the worst censoring scheme.

REFERENCES

Balakrishnan, N. and Aggarwala, R. A. (2000), *Progressive Censoring: Theory, Methods, and Applications*, Boston: Birkhäuser.

Balakrishnan, N., Childs, A., and Chandrasekar, B. (2002), "An efficient computational method for moments of order statistics under progressive censoring," *Statistics & Probability Letters*, 60, 359–365.

Guilbaud, O. (2001), "Exact nonparametric confidence intervals for quantiles with progressive Type-II censoring," *Scandinavian Journal of Statistics*, 28, 699–713.

Guilbaud, O. (2004), "Exact nonparametric confidence, prediction and tolerance intervals with progressive Type-II censoring," *Scandinavian Journal of Statistics*, 31, 265–281.

[1] confidence interval
[2] confidence level
[3] progressive censoring
[4] not available

Table 1: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 15$ and $m = 10$ (total 2002 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI[1]	Actual CL[2]	CI Mass	Efficiency
0.01	212	Best PC[3]	(4, 0⋆8, 1)	[2, 10]	0.9904	0.6818	100.00 %
		Worst PC	(1⋆5, 0⋆5)	[1, 10]	0.9901	0.8328	81.87 %
		Type-II	(0⋆9, 5)	[1, 10]	0.8491	0.5625	NA[4]
0.05	1831	Best PC	(0⋆3, 3, 0⋆5, 2)	[4, 10]	0.9524	0.5000	100.00 %
		Worst PC	(0⋆4, 1, 0, 0, 1, 3, 0)	[1, 10]	0.9501	0.7370	67.84 %
		Type-II	(0⋆9, 5)	[1, 10]	0.8491	0.5625	NA
0.10	1992	Best PC	(0, 0, 1, 0⋆4, 1, 0, 3)	[4, 10]	0.9004	0.4288	100.00 %
		Worst PC	(0⋆3, 2, 0, 0, 2, 1, 0, 0)	[4, 10]	0.9609	0.6100	70.30 %
		Type-II	(0⋆9, 5)	[1, 10]	0.8491	0.5625	NA

Table 2: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 20$ and $m = 10$ (total 92378 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	32	Best PC	(0⋆4,10, 0⋆5)	[4, 10]	0.9922	0.6825	100.00 %
		Worst PC	(8, 1, 1, 0⋆7)	[1, 10]	0.9905	0.8542	79.91 %
		Type-II	(0⋆9,10)	[1, 10]	0.4119	0.4286	NA
0.05	3550	Best PC	(0, 0, 8, 0⋆6, 2)	[4, 10]	0.9507	0.5143	100.00 %
		Worst PC	(4, 1, 1, 2, 1, 0, 1, 0⋆3)	[1, 10]	0.9500	0.8317	61.83 %
		Type-II	(0⋆9,10)	[1, 10]	0.4119	0.4286	NA
0.10	27873	Best PC	(7, 0⋆8, 3)	[4, 10]	0.9143	0.4396	100.00 %
		Worst PC	(0, 1, 0, 3, 1, 1, 4, 0⋆3)	[1, 10]	0.9000	0.8024	54.78 %
		Type-II	(0⋆9,10)	[1, 10]	0.4119	0.4286	NA

Table 3: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 20$ and $m = 16$ (total 3876 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	3876	Best PC	(0⋆12, 1, 0, 0, 3)	[5, 16]	0.9902	0.5442	100.00 %
		Worst PC	(0, 0, 1, 1, 0⋆5, 1, 1, 0⋆5)	[4, 15]	0.9967	0.6655	81.77 %
		Type-II	(0⋆15, 4)	[4, 16]	0.9928	0.5714	95.24 %
0.05	3876	Best PC	(0⋆6, 1, 0⋆3, 1, 0⋆4, 2)	[6, 14]	0.9503	0.4258	100.00 %
		Worst PC	(0, 1, 1, 0⋆8, 1, 1, 0⋆3)	[5, 14]	0.9761	0.5140	82.84 %
		Type-II	(0⋆15, 4)	[6, 15]	0.9586	0.4286	99.36 %
0.10	3876	Best PC	(0⋆8, 1, 0⋆6, 3)	[7, 14]	0.9015	0.3550	100.00 %
		Worst PC	(0⋆5, 2, 0, 2, 0⋆8)	[6, 13]	0.9484	0.4457	79.65 %
		Type-II	(0⋆15, 4)	[6, 14]	0.9216	0.3810	93.18 %

Table 4: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 25$ and $m = 10$ (total 1307504 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	21	Best PC	(0⋆4,15, 0⋆5)	[5, 10]	0.9903	0.6731	100.00 %
		Worst PC	(0,15, 0⋆8)	[2, 10]	0.9979	0.8205	82.03 %
		Type-II	(0⋆9,15)	[1, 10]	0.1148	0.3462	NA
0.05	766	Best PC	(0⋆6,15, 0⋆3)	[7, 10]	0.9577	0.5481	100.00 %
		Worst PC	(11, 0, 1⋆3, 0, 0, 1, 0, 0)	[1, 10]	0.9502	0.8347	65.66 %
		Type-II	(0⋆9,15)	[1, 10]	0.1148	0.3462	NA
0.10	7122	Best PC	(0, 0,13, 0⋆6, 2)	[5, 10]	0.9016	0.4423	100.00 %
		Worst PC	(8, 0, 3, 2, 0, 2, 0⋆4)	[1, 10]	0.9000	0.8406	52.62 %
		Type-II	(0⋆9,15)	[1, 10]	0.1148	0.3462	NA

Table 5: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 25$ and $m = 15$ (total 1961256 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	571742	Best PC	(0⋆5, 9, 0⋆8, 1)	[6, 14]	0.9903	0.5594	100.00 %
		Worst PC	(0, 1, 0, 1, 2, 0, 1⋆3, 2, 0, 1, 0⋆3)	[1, 15]	0.9900	0.8649	64.68 %
		Type-II	(0⋆14,10)	[1, 15]	0.7878	0.5385	NA
0.05	1908500	Best PC	(0⋆7, 6, 0⋆6, 4)	[8, 15]	0.9506	0.4038	100.00 %
		Worst PC	(1, 0⋆7, 1, 0, 1, 0, 4, 3, 0)	[1, 15]	0.9500	0.7686	52.54 %
		Type-II	(0⋆14,10)	[1, 15]	0.7878	0.5385	NA
0.10	1959458	Best PC	(1, 0⋆6, 3, 0⋆6, 6)	[8, 15]	0.9010	0.3405	100.00 %
		Worst PC	(0⋆3, 1, 0⋆6, 1, 1, 0, 3, 4)	[2, 15]	0.9000	0.5680	59.95 %
		Type-II	(0⋆14,10)	[1, 15]	0.7878	0.5385	NA

Table 6: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 25$ and $m = 20$ (total 42504 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	42504	Best PC	(0⋆19, 5)	[7, 20]	0.9906	0.5000	100.00 %
		Worst PC	(0, 1, 1, 2, 1, 0⋆15)	[6, 18]	0.9924	0.5956	83.94 %
		Type-II	(0⋆19, 5)	[7, 20]	0.9906	0.5000	100.00 %
0.05	42504	Best PC	(0⋆10, 2, 0⋆8, 3)	[8, 17]	0.9501	0.3817	100.00 %
		Worst PC	(0⋆4, 1, 0, 3, 1, 0⋆12)	[7, 16]	0.9749	0.4660	81.90 %
		Type-II	(0⋆19, 5)	[8, 18]	0.9567	0.3846	99.23 %
0.10	42504	Best PC	(0⋆13, 1, 0⋆5, 4)	[9, 17]	0.9010	0.3182	100.00 %
		Worst PC	(0, 1, 0, 1⋆4, 0⋆13)	[8, 16]	0.9288	0.4051	78.54 %
		Type-II	(0⋆19, 5)	[8, 17]	0.9245	0.3462	91.92 %

Table 7: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 30$ and $m = 10$ (total 10015005 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	17	Best PC	(19, 0⋆8, 1)	[2, 10]	0.9902	0.7038	100.00 %
		Worst PC	(1, 19, 0⋆8)	[2, 10]	0.9911	0.8306	84.74 %
		Type-II	(0⋆9, 20)	[1, 10]	0.0214	0.2903	NA
0.05	436	Best PC	(0, 19, 0⋆7, 1)	[4, 10]	0.9540	0.5613	100.00 %
		Worst PC	(16, 3, 0⋆4, 1, 0⋆3)	[1, 10]	0.9501	0.8554	65.62 %
		Type-II	(0⋆9, 20)	[1, 10]	0.0214	0.2903	NA
0.10	2725	Best PC	(17, 0⋆8, 3)	[4, 10]	0.9145	0.4467	100.00 %
		Worst PC	(13, 6, 0, 0, 1, 0⋆5)	[1, 10]	0.9000	0.8615	51.85 %
		Type-II	(0⋆9, 20)	[1, 10]	0.0214	0.2903	NA

Table 8: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 30$ and $m = 15$
(total 77558760 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	77558760	Best PC	(12, 0⋆13, 3)	[1, 9]	0.9905	0.4301	100.00 %
		Worst PC	(0⋆9, 1, 0, 1, 3, 10, 0)	[1, 15]	0.9934	0.7029	61.19 %
		Type-II	(0⋆14, 15)	[1, 15]	0.9904	0.4516	95.24 %
0.05	77558760	Best PC	(0, 12, 0⋆5, 1, 0⋆6, 2)	[1, 9]	0.9572	0.4230	100.00 %
		Worst PC	(0⋆11, 1, 2, 12, 0)	[1, 15]	0.9720	0.6976	60.63 %
		Type-II	(0⋆14, 15)	[1, 15]	0.9599	0.4516	93.66 %
0.10	77558760	Best PC	(0⋆4, 9, 0⋆9, 6)	[8, 15]	0.9001	0.3454	100.00 %
		Worst PC	(0⋆9, 1, 2, 5, 7, 0, 0)	[1, 14]	0.9132	0.6007	57.50 %
		Type-II	(0⋆14, 15)	[1, 14]	0.9145	0.4194	82.35 %

Table 9: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 30$ and $m = 20$
(total 20030010 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	20030010	Best PC	(0⋆7, 4, 0⋆10, 1, 5)	[8, 20]	0.9901	0.4742	100.00 %
		Worst PC	(0⋆14, 3, 6, 1, 0⋆3)	[8, 20]	0.9945	0.6427	73.78 %
		Type-II	(0⋆19, 10)	[1, 19]	0.9926	0.5806	81.66 %
0.05	20030010	Best PC	(0⋆11, 1, 0, 1, 0⋆5, 8)	[10, 20]	0.9500	0.3505	100.00 %
		Worst PC	(0⋆15, 1, 3, 4, 2, 0)	[10, 20]	0.9672	0.5157	67.98 %
		Type-II	(0⋆19, 10)	[6, 20]	0.9505	0.4516	77.62 %
0.10	20030010	Best PC	(0⋆19, 10)	[11, 20]	0.9013	0.2903	100.00 %
		Worst PC	(0⋆18, 10, 0)	[11, 20]	0.9031	0.4516	64.29 %
		Type-II	(0⋆19, 10)	[11, 20]	0.9013	0.2903	100.00 %

Table 10: Optimal PC Schemes for Nonparametric CIs of $\xi_{0.50}$ with $n = 30$ and $m = 25$
(total 118755 progressive Type-II censoring schemes examined)

α	Count	Category	Censoring Scheme ($\mathbf{R}$)	CI	Actual CL	CI Mass	Efficiency
0.01	118755	Best PC	(0⋆19, 1, 0⋆4, 4)	[8, 22]	0.9904	0.4581	100.00 %
		Worst PC	(0⋆4, 1, 2, 2, 0⋆18)	[8, 21]	0.9933	0.5259	87.11 %
		Type-II	(0⋆24, 5)	[8, 23]	0.9948	0.4839	94.67 %
0.05	118755	Best PC	(0⋆11, 1, 0, 1, 0⋆10, 3)	[10, 20]	0.9500	0.3505	100.00 %
		Worst PC	(0⋆5, 1, 0, 0, 3, 1, 0⋆15)	[9, 19]	0.9721	0.4141	84.66 %
		Type-II	(0⋆24, 5)	[10, 21]	0.9572	0.3548	98.79 %
0.10	118755	Best PC	(0⋆24, 5)	[11, 20]	0.9013	0.2903	100.00 %
		Worst PC	(0, 1⋆3, 0, 0, 1, 1, 0⋆17)	[8, 17]	0.9066	0.3605	80.53 %
		Type-II	(0⋆24, 5)	[11, 20]	0.9013	0.2903	100.00 %