Empirical quantile process under type-II progressive censoring

Sergio Alvarez-Andrade*, Laurent Bordes

Laboratoire de Mathématiques Appliquées de Compiègne, Université de Technologie de Compiègne, BP 529, Compiègne Cedex, F-90205, France

Received March 2003; accepted January 2004

Abstract

This work deals with asymptotic properties of the $[zm]$th-order statistic of a type-II progressively censored sample of size $m$. Such an order statistic, indexed by $z \in [0, 1]$, is called the quantile process. Our main results concern the normalized version of the quantile process for which a weak convergence result is obtained. This result is applied in order to construct non-parametric estimators of quantiles. Monte-Carlo simulations illustrate the behavior of the estimators for limited sample size.

Keywords: Weak convergence; Martingales; Monte-Carlo study; Progressive censoring; Quantile process; Reliability; Variance estimators

1. Introduction

In many industrial experiments involving lifetimes of machines or units, experiments have to be terminated early or the number of experiments must be limited due to a variety of circumstances (e.g. when expensive items must be destroyed, when experiments are time-consuming and expensive, etc.). In addition, some lifetests require removals of functioning test specimens to collect degradation related information to failure time data. The samples that arise from such experiments are called censored. The planning of experiments with the aim of reducing both the number of failures and the total duration of the experiment leads naturally to the well-known types-I and -II right censoring schemes. For many references and historical notes on this subject we refer to Balakrishnan and Aggarwala (2000). The progressive censoring scheme, as will be introduced hereafter, has the same objectives, but it is constructed with the aim of moderating the loss of information by reducing the number of failures with respect to the full sample approach. Montanari and Cacciari (1988)

* Corresponding author.
E-mail address: sergio.alvarez@utc.fr (S. Alvarez-Andrade).

0167-7152/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
reported results of progressively censored data aging tests on XLPE-insulated cable models under combined thermal—electrical stresses. In this experiment, live specimens were removed at selected times and/or at the time of breakdowns. The progressive censoring sampling plans by Montanari and Cacciari (1988) are considered in Balasooriya et al. (2000) with the aim of defining optimal sampling plans under Weibull estimation.

In a paper presenting a unified approach for many models based on order statistics and record values, Kamps (1995) proposed a generalized form of the joint distribution of \( n \) ordered random variables. One of the models included in this general setup is the type-II progressive censoring scheme as defined by Balakrishnan and Aggarwala (2000), see also Viveros and Balakrishnan (1994) for an interesting review of the background and of developments in this field. Hereafter we recall how to obtain such progressively censored data.

Let \( X_1, \ldots, X_n \) be independent and identically distributed (i.i.d.) random lifetimes of \( n \) items. A type-II progressively right censored sample may be obtained in the following way: at the time of the first failure, noted \( X_{1:m:n} \), \( r_1 \) surviving items are removed at random from the \( n - 1 \) remaining surviving items, at the time of the next failure, noted \( X_{2:m:n} \), \( r_2 \) surviving items are removed at random from the \( n - r_1 - 2 \) remaining items, and so on. At the time of the \( m \)th failure, all the remaining \( r_m = n - m - r_1 - \cdots - r_{m-1} \) surviving items are censored. Therefore a type-II progressive censoring scheme is specified by integers \( n, m \) and \( r_1, \ldots, r_{m-1} \) with the constraints \( n - m - r_1 - \cdots - r_{m-1} \geq 0 \) and \( n \geq m \geq 1 \).

If \( r_1 = \cdots = r_{m-1} = 0 \) we get the usual type-II right censored sample, that is we observe the first \( m \)-order statistics \( X_{1:n}, \ldots, X_{m:n} \) and the \( n - m \) remaining times are right censored by \( X_{m:n} \). If moreover \( m = n \) the usual order statistic is obtained.

Our main development deals with asymptotic behavior of the \([x m]n\)th-order statistic of a type-II progressively censored sample of size \( m \) (\([x]\) is the unique integer satisfying \( [x] \leq x < [x] + 1 \)). Such an order statistic, indexed by \( x \in [0,1] \), is called the quantile process and it is written \( (X_{[zm]:m:n})_{x \in [0,1]} \) with \( X_{0:m:n} \equiv 0 \). In Section 2, we get a weak convergence result for the normalized version of the quantile process by using a martingale method. Section 3 is devoted to the estimation of the asymptotic variance function of the quantile process. One estimator based on kernel estimation of the hazard rate is proposed. These results are applied in Section 4 for constructing non-parametric estimators of quantiles with confidence intervals (note that recently, by another approach, Guilbaud (2001) derived exact non-parametric confidence intervals for quantiles). Finally, in order to illustrate the behavior of the estimators for finite sample size, a Monte-Carlo study is provided in Section 5.

From now on we assume that the \( X_i \)'s have a common distribution function \( F \) with density \( f \). We denote by \( \lambda \) the hazard rate function and \( A \) the cumulative hazard rate function.

2. Asymptotics for the quantile process

In this section we investigate the asymptotic properties of the quantile process \( (X_{[zm]:m:n})_{x \in [0,a]} \). We first introduce assumptions under which these asymptotic properties are obtained. All asymptotic results are given with respect to \( m \to +\infty \).

A1. \( x \in [0,a] \) and \( 0 < a < 1 \).
A2. \( (r_i)_{i \geq 1} \) is a bounded sequence of nonnegative integers \( (r_i \leq K \) for all \( i \geq 1 \)).
A3. \( \bar{r} = m^{-1} \sum_{i=1}^{m} r_i = r + \gamma_m \) where \( r \) is a nonnegative real number and \((\gamma_m)_{m \geq 1}\) is a real valued sequence which satisfies (a) \( \gamma_m = o(1) \) or (b) \( \gamma_m = o(m^{-1/2}) \).

A4. Let us consider \( G = 1 - (1 - F)^{r+1} \). There exists a real number \( \varepsilon \in [0, a] \) such that \( G^{-1}([\varepsilon, a]) \subset (c, b) \subset \mathbb{R}^+ \) and \( \lambda \) is continuous and strictly positive on \((c, b)\).

Note that asymptotic results will be obtained supposing that \((r_i)_{i \geq 1}\) is a deterministic sequence of integers (in fact it could be a triangular array). However, if this sequence is not deterministic (this is the case in our simulations studies) results must be understood as conditional results with respect to this sequence.

We now introduce some notations. Let \((\varkappa_j^m)_{1 \leq j \leq m}\) be a triangular array of nonnegative integers defined by

\[ \varkappa_j^m = r_j + \cdots + r_m + m - j + 1 \quad \text{for} \quad 1 \leq j \leq m, \]

we denote by \( u_x \) the \( x \)-quantile of the distribution function \( G \), \( u_x = G^{-1}(x) \) where \( G^{-1} \) is taken in the generalized inverse sense \((G^{-1}(x) = \inf\{y : G(y) \geq x\}) \) when it is not invertible.

We shall use intensively in the sequel the following technical results.

**Lemma 1.** (a) There exists \((\varkappa_j^m)_{1 \leq j \leq m}\) a triangular array of row i.i.d. random variables exponentially distributed with mean 1, such that

\[ A(X_i;m:n) = \sum_{j=1}^{i} \frac{\varkappa_j^m}{\varkappa_j^m}, \quad a.s. \quad (1) \]

(b) Under A1–A3(a or b), the following sequence has the same rate as \((\gamma_m)_{m \geq 1}\),

\[ \left( \sup_{1 \leq j \leq \lceil am \rceil} \left| \frac{\varkappa_j^m}{m - j + 1} - (r + 1) \right| \right)_{m \geq 1}. \quad (2) \]

(c) Under A1–A3(a) we have

\[ \lim_{m \to \infty} \sup_{x \in [0, a]} \left| \sum_{j=1}^{\lceil 1 \rceil} \frac{1}{\varkappa_j^m} - \log(1/(1 - x)) \right| r + 1 = 0. \quad (3) \]

(d) Under A1–A3(a) we have

\[ \lim_{m \to \infty} \sup_{x \in [0, a]} \left| m \sum_{j=1}^{\lceil 1 \rceil} \frac{1}{(\varkappa_j^m)^2} - \frac{x}{(r + 1)^2(1 - x)} \right| = 0. \quad (4) \]

**Proof.** (a) is a direct result of Balakrishnan and Aggarwala (2000, pp. 17–18). Let us prove (b) which extends Lemma 4.1 in Bordes (2004). First remark that for \( 1 \leq j \leq am \)

\[ \frac{\varkappa_j^m}{m - j + 1} - (r + 1) = \frac{m_j^m}{m - j + 1} - \frac{(j - 1)\gamma_{j-1}}{m - j + 1}. \]
Then we have by A1
\[
\sup_{1 \leq j \leq [am]} \left| \frac{\gamma^m_j}{m - j + 1} - (r + 1) \right| \leq \frac{1}{1 - a} \left\{ |\gamma_m| + \frac{1}{m} \sup_{1 \leq j \leq \theta_m} j |\gamma_j| + \frac{1}{m} \sup_{\theta_m \leq j \leq am} j |\gamma_j| \right\},
\]
where \((\theta_m)_m \geq 1\) satisfies \(\theta_m \to +\infty\) and \(\theta_m/am = o(m^{-1/2})\). Now, by A2 we have \(|\gamma_m| \leq K + r\) for all \(m \geq 1\), thus:
\[
\sup_{1 \leq j \leq [am]} \left| \frac{\gamma^m_j}{m - j + 1} - (r + 1) \right| \leq \frac{1}{1 - a} \left\{ |\gamma_m| + \frac{\theta_m}{m} (K + r) + \frac{1}{m} \sup_{\theta_m \leq j \leq am} j |\gamma_j| \right\}.
\]
Then the right-hand side of the above inequality has the rate:
\[
O(\gamma_m) + O\left(\frac{\theta_m}{m}\right) + O\left(\frac{1}{m} \sup_{\theta_m \leq j \leq am} j |\gamma_j|\right).
\]
Under A3(a) we have \(O(\gamma_m) + O(\theta_m/m) = o(1)\), and
\[
\frac{1}{m} \sup_{\theta_m \leq j \leq am} j |\gamma_j| \leq \frac{am}{m} \sup_{\theta_m \leq j \leq am} |\gamma_j| = o(1),
\]
which gives the result.

Under A3(b) we have \(\gamma_m = m^{-1/2} a_m\) where \(a_m = o(1)\). Then,
\[
\frac{1}{m} \sup_{\theta_m \leq j \leq am} j |\gamma_j| = \frac{1}{m} \sup_{\theta_m \leq j \leq am} j \times j^{-1/2} |a_j| \leq \frac{(am)^{1/2}}{m} \sup_{\theta_m \leq j \leq am} |a_j| = o(m^{-1/2}).
\]
We conclude since \(O(\gamma_m) + O(\theta_m/m) = o(m^{-1/2})\), by assumptions on sequences \((\gamma_m)_m \geq 1\) and \((\theta_m)_m \geq 1\).

The proof of (c) is the same as the proof of (d), and then is omitted. To prove (d) remark that
\[
\sum_{j=1}^{[am]} \frac{1}{(r+1)^2(m-k+1)^2} + \sum_{k=1}^{[am]} \frac{1}{m-k+1^2} \left[ \frac{1}{(\gamma^m_k/(m-k+1))^2} - \frac{1}{(r+1)^2} \right] := I^{(m)} + II^{(m)}.
\]
Note that under A2
\[
II^{(m)} \leq m \frac{K + r + 2}{(r+1)^2} \sum_{k=1}^{[am]} \left[ \frac{r + 1 - \gamma^m_k/(m-k+1)}{(m-k+1)^2} \right].
\]
We have
\[
II^{(m)} \leq m \frac{K + r + 2}{(r+1)^2} \sum_{k=1}^{[am]} \frac{1}{(m-k+1)^2} \sup_{1 \leq j \leq am} \left| \frac{\gamma^m_j}{m - j + 1} - (r + 1) \right|
\]
\[
\leq m \frac{K + r + 2}{(r+1)^2} \int_0^{[am]} \frac{dx}{(m-x)^2} \sup_{1 \leq j \leq am} \left| \frac{\gamma^m_j}{m - j + 1} - (r + 1) \right| = o(1),
\]
from (b). Now, since \((r + 1)^2 \Gamma(m) = m \sum_{k=1}^{[mx]} (m - k + 1)^{-2}\), we have
\[
m \int_0^{[mx]} \frac{dx}{(m - x + 1)^2} \leq (r + 1)^2 \Gamma(m) \leq m \int_0^{[mx]} \frac{dx}{(m - x)^2},
\]
which is equivalent to
\[
\frac{[mx] - 1}{m - [mx] + 1} \leq (r + 1)^2 \Gamma(m) \leq \frac{[mx]}{m - [mx]},
\]
where the upper and the lower bounds have the same limit uniformly in \(\alpha \in [0, a]\). The result (d) follows.

Let \((Z_i^{(m)})_{1 \leq i \leq m}\) be the triangular array of Lemma 1. In the sequel we shall use intensively the above triangular array, thus, for simplicity, we omit from now on the exponent \(m\) in the \(Z_i^{(m)}\)'s.

**Proposition 1.** Let \(\tilde{Y}^{(m)}\) be the process defined on \([0, a]\) by
\[
\tilde{Y}^{(m)}(\alpha) = m^{1/2} \sum_{j=1}^{[zm]} \frac{Z_j - 1}{\alpha_j^m}.
\]
Under A1–A3(a) we have \(\tilde{Y}^{(m)} \succeq \mathcal{A}_\gamma\) in \(D[0, a]\), where \(\mathcal{A}_\gamma\) is a centered Gaussian process on \([0, a]\) with variance function \(\gamma\) defined by
\[
\gamma(\alpha) = \frac{\alpha}{(r + 1)^2(1 - \alpha)}.
\]

**Proof.** Define the natural filtration \(\mathcal{F}_k^{(m)} = \{Z_1, \ldots, Z_k; r_1, \ldots, r_k\}\) (of course since \((r_i)_{i \geq 1}\) is deterministic, the \(r_i\)'s may be removed from the filtration). It is clear that:
\[
M_k^{(m)} = m^{1/2} \sum_{j=1}^k \frac{Z_j - 1}{\alpha_j^m} \quad \text{for } 1 \leq k \leq m,
\]
is an \(\mathcal{F}_k^{(m)}\)-martingale. Moreover, for all \(1 \leq k \leq m\)
\[
E[(M_k^{(m)})^2] = m \sum_{j=1}^k \frac{1}{(\alpha_j^m)^2} < +\infty.
\]
We have
\[
\langle M_k^{(m)} \rangle_{[zm]} = m \sum_{j=1}^{[zm]} \frac{1}{(\alpha_j^m)^2} \to \frac{\alpha}{(1 + r)^2(1 - \alpha)} = \gamma(\alpha),
\]
from Lemma 1(d), where \(\gamma : [0, a] \to \mathbb{R}\) is a continuous, non decreasing, deterministic function on \([0, a]\) satisfying \(\gamma(0) = 0\). Therefore Condition K1 of Theorem 7.4.28 in Dacunha-Castelle and Duflo (1993, p. 226) is satisfied.
Now, since \( M_k^{(m)} - M_{k-1}^{(m)} = m^{1/2}(Z_k - 1)/\alpha_k^{(m)} \), Condition K2 of Theorem 7.4.28 in Dacunha-Castelle and Duflo (1993, p. 226) is true if for all \( \varepsilon > 0 \)

\[
\Gamma^{(m)} = \sum_{j=1}^{[zm]} E \left( m \left( \frac{Z_j - 1}{\alpha_j^{(m)}} \right)^2 1(|Z_j - 1| \geq \varepsilon \alpha_j^{(m)} m^{-1/2}) | \mathcal{F}_{j-1}^{(m)} \right)
\]

\[
= m \sum_{j=1}^{[zm]} \frac{1}{(\alpha_j^{(m)})^2} E((Z_j - 1)^2 1(|Z_j - 1| \geq \varepsilon \alpha_j^{(m)} m^{-1/2})) \to 0.
\]

Note that for all \( 1 \leq j \leq [zm] \) we have the almost sure inequality

\[
1(|Z_j - 1| \geq \varepsilon \alpha_j^{(m)} m^{-1/2}) \leq 1(|Z_j - 1| \geq \varepsilon (1-a)m^{1/2}).
\]

Then

\[
\Gamma^{(m)} \leq m \sum_{j=1}^{[zm]} \frac{1}{(m - j + 1)^2} E((Z_j - 1)^2 1(|Z_j - 1| \geq \varepsilon (1-a)m^{1/2}))
\]

\[
\leq \frac{a}{(1-a)^2} E((Z_1 - 1)^2 1(Z_1 - 1 \geq \varepsilon (1-a)m^{1/2}))
\]

where the last inequality is true whenever \( m \geq (\varepsilon (1-a))^{-2} \). It is then easy to see, using two integration by parts, that the above expectation term tends to 0. Then Condition K2 of Theorem 7.4.28 in Dacunha-Castelle and Duflo (1993, p. 226) is satisfied, and, since \( \tilde{Y}^{(m)}(x) = M_{[zm]}^{(m)} \), by result 1 of that theorem we conclude. \( \square \)

**Corollary 1.** Under A1–A3(a), if \( Y^{(m)} \) is the process defined by

\[
Y^{(m)}(x) = m^{1/2} \left( \sum_{j=1}^{[zm]} \frac{Z_j}{\alpha_j^{(m)}} - \frac{1}{r + 1} \log \left( \frac{1}{1 - x} \right) \right),
\]

we have

(i) \( \sup_{x \in [0,a]} |m^{-1/2}Y^{(m)}(x)| \overset{p}{\to} 0; \)

(ii) Moreover, if A3(b) is satisfied, then \( Y^{(m)} \overset{D}{\to} G \) in \( D[0,a] \), where \( G \) is the Gaussian process defined in Proposition 1.

**Proof.** Let us prove (ii).

\[
\sup_{x \in [0,a]} |Y^{(m)}(x) - \tilde{Y}^{(m)}(x)| \leq \Gamma^{(m)} + \Pi^{(m)},
\]

where

\[
\Gamma^{(m)} = \sup_{x \in [0,a]} \left| m^{1/2} \sum_{i=1}^{[zm]} \left( \frac{1}{\alpha_i^{(m)}} - \frac{1}{(r + 1)(m - i + 1)} \right) \right|
\]

\[
\Pi^{(m)} = \sum_{j=1}^{[zm]} \frac{1}{(\alpha_j^{(m)})^2} E((Z_j - 1)^2 1(|Z_j - 1| \geq \varepsilon \alpha_j^{(m)} m^{-1/2}) | \mathcal{F}_{j-1}^{(m)}).
\]
and
\[ \Pi^{(m)} = \sup_{x \in [0,a]} \left| \frac{m^{1/2}}{r+1} \left( \sum_{i=1}^{[zm]} \frac{1}{m-i+1} - \log \left( \frac{1}{1-x} \right) \right) \right|. \]

We have
\[ \Pi^{(m)} \leq \frac{m^{1/2}}{r+1} \sup_{1 \leq i \leq [zm]} \left| \frac{x^m_i}{m-i+1} - (r+1) \sum_{j=1}^{[zm]} \frac{1}{m-j+1} \right| \]
\[ = O(m^{1/2}) \times o(m^{-1/2}) \times O(1) = o(1), \]
by Lemma 1(b) and (c) under Assumption A3(b). For \( \Pi^{(m)} \) we have
\[ \sup_{x \in [0,a]} \left| \sum_{i=1}^{[zm]} \frac{1}{m-i+1} - \log \left( \frac{1}{1-x} \right) \right| \]
\[ \leq \sup_{x \in [0,a]} \log \left( \frac{m-[zm]+1}{(m+1)(1-x)} \right) \leq \log \left( 1 + \frac{2}{m(1-a)} \right) = O(m^{-1}), \]
where the first inequality is obtained by similar calculations used in the proof of Lemma 1(d). This last result clearly implies the convergence of \( \Pi^{(m)} \) to 0.

Let us prove (i). We have
\[ \sup_{x \in [0,a]} |m^{-1/2} Y^{(m)}(x)| \leq \sup_{x \in [0,a]} |m^{-1/2} \tilde{Y}^{(m)}(x)| + m^{-1/2}(\Pi^{(m)} + \Pi^{(m')}). \]

By Proposition 1 we have \( \sup_{x \in [0,a]} m^{-1/2} |\tilde{Y}^{(m)}| = o_p(1) \) and, similar to the previous arguments for \( \Pi^{(m)} \) and \( \Pi^{(m')} \), we get \( m^{-1/2} (\Pi^{(m)} + \Pi^{(m')}) = o_p(1) \) under A3(a). This completes the proof of (i).

**Theorem 1.** Under A1–A3(a) we have

(i) \( \sup_{x \in [0,a]} |X_{[zm]:m:n} - u_x| \xrightarrow{p} 0; \)

(ii) Moreover, if A3(b) and A4 are satisfied, and \( X^{(m)} \) is the process defined by \( X^{(m)}(x) = m^{1/2} (X_{[zm]:m:n} - u_x) \), then
\[ X^{(m)} \xrightarrow{D} \mathcal{G}_{\sigma} \quad \text{in } D[\varepsilon,a], \]
where \( \mathcal{G}_{\sigma} \) is the centered Gaussian process on \([\varepsilon,a]\) with variance function \( \sigma^2 \) defined by
\[ \sigma^2(x) = E(\mathcal{G}_{\sigma}^2(x)) = \frac{x}{(1+r)^2(1-x)^2(u_x)}, \quad x \in [\varepsilon,a]. \]

**Proof.** For simplicity, we introduce the notations
\[ Y_{i:m:n} = \sum_{j=1}^{i} \frac{Z_j}{x^m_j} \quad \text{and} \quad v_x = \frac{1}{r+1} \log \left( \frac{1}{1-x} \right). \]
Now let us prove (i). There exists $\varepsilon' > 0$ such that $\Lambda^{-1}$ is continuous on $[0, b + \varepsilon']$ (since $\Lambda$ is continuous on $\mathbb{R}^+$). For $0 < \eta < \varepsilon'$ we have, by Lemma 1(a):

$$P \left( \sup_{x \in [0,a]} |X_{[2m];m:n} - u_x| > \varepsilon \right)$$

$$\leq P \left( \sup_{x \in [0,a]} |A^{-1}(Y_{[2m];m:n}) - A^{-1}(v_x)| > \varepsilon; \sup_{x \in [0,a]} |Y_{[2m];m:n} - v_x| \leq \eta \right)$$

$$+ P \left( \sup_{x \in [0,a]} |Y_{[2m];m:n} - v_x| > \eta \right).$$

Choosing $\eta$ small enough the first term in the right-hand side of the above inequality is equal to 0 due to the uniform continuity of $\Lambda^{-1}$ on $[0, b + \varepsilon']$, whereas the second term tends to 0 by simply applying Corollary 1(ii).

Finally, using A4, (ii) is an obvious application of the $\delta$-method (see Gill, 1989, or Andersen et al., 1993). Indeed, by Lemma 1(a) we have

$$X^{(m)}(x) = \sqrt{m}(A^{-1}(Y_{[2m];m:n}) - A^{-1}(v_x)).$$

Now, from Corollary 1(ii) we know that $Y^{(m)}(\cdot) \overset{D}{=} \mathcal{G}_f$. Assumption A4 insures differentiability of $\Lambda^{-1}$ on $(c, b)$ (remark that by A4, if $x \in [\varepsilon, a]$ then $\Lambda^{-1}(v_x) \in G^{-1}([\varepsilon, a]) \subset (c, b)$ where $\Lambda^{-1}$ is differentiable). Then we have

$$X^{(m)}(\cdot) \overset{D}{=} \mathcal{G}_f(\cdot) = \mathcal{G}_\lambda(\cdot),$$

on $D[\varepsilon, a]$. \qed

3. Variance estimation based on a kernel type estimator

Here we need to estimate $\sigma^2(x) = x/(r + 1)^2(1 - x)\lambda^2(u_x))$. Since by Theorem 1 and Assumption A3(a), $u_x$ and $r$ are naturally estimated by $X_{[2m];m:n}$ and $\tilde{r} = \sum_{j=1}^{m} r_j/m$, we propose to estimate $\sigma^2(x)$ by

$$\hat{\sigma}^2(x) = \frac{x}{(\tilde{r} + 1)^2(1 - x)\hat{\lambda}^2(X_{[2m];m:n})},$$

where $\hat{\lambda}$ is an estimator of $\lambda$. Such an estimator, following the ideas of Andersen et al. (1993), may be obtained by smoothing the estimator $\hat{A}^{(m)}$ of the cumulative hazard rate $\Lambda$ proposed by Bordes (2004). Formally, we define the estimator by

$$\hat{\lambda}^{(m)}(t) = b_m^{-1} \int_{[0,\tau]} K \left( \frac{t - s}{b_m} \right) \hat{A}^{(m)}(ds),$$

where $\tau$ is the upper bound of the interval study, and $\hat{A}^{(m)}$ is defined by

$$\hat{A}^{(m)}(t) = \int_{0}^{t} \frac{N^{(m)}(ds)}{Y^{(m)}(s)} \quad \text{for} \quad 0 \leq t \leq \tau.$$
Here processes $N^{(m)}$ and $Y^{(m)}$ are respectively defined by $N^{(m)}(t) = \sum_{j=1}^{m} \mathbb{1}(X_{j;m:n} \leq t)$ and $Y^{(m)}(t) = \sum_{j=1}^{m} (r_j + 1)\mathbb{1}(X_{j;m:n} \geq t)$. The real number $\tau$ is such that $0 < F(\tau) < 1$, $K$ is a bounded kernel function, which vanishes outside $[-1,1]$ and has integral 1. The bandwidth or window size $b_m$ is a positive parameter satisfying $(b_m)_{m \geq 1} \searrow 0$. By Theorem IV.2.2 in Andersen et al. (1993), if $\hat{\lambda}$ is continuous on $(c,b)$ and $F(b) < 1$, we have for all $[t_1,t_2] \subset (c,b)$:

$$\sup_{t \in [t_1,t_2]} |\hat{\lambda}^{(m)}(t) - \hat{\lambda}(t)| \xrightarrow{p} 0, \quad \text{as} \quad m \to +\infty. \quad (6)$$

Other results, concerning estimation of $\lambda$, can be found in Balakrishnan and Bordes (2004).

**Proposition 2.** Under Assumptions A1–A4, if $\lambda$ is continuous on $(c,b)$ and $F(b) < 1$ then, for all $\varepsilon > 0$ and $\eta > 0$ be two real numbers, we have

$$\sigma^2(\varepsilon) \xrightarrow{p} \sigma^2(\varepsilon), \quad \text{as} \quad m \to +\infty.$$

**Proof.** It is sufficient to prove that $\hat{\lambda}(X_{(2m;m:n)}) \xrightarrow{p} \lambda(u_2)$, as $m \to +\infty$. Let $\varepsilon > 0$ and $\eta > 0$ be two real numbers, we have

$$P(\hat{\lambda}(X_{(2m;m:n)}) - \lambda(u_2) > \varepsilon) \leq P(|X_{(2m;m:n)} - u_2| > \eta)$$

$$+ P\left(\sup_{\{t \geq 0; 0 \leq u_2 \leq \eta\}} |\hat{\lambda}^{(m)}(t) - \hat{\lambda}(t)| > \varepsilon/2\right) + P\left(\sup_{\{t \geq 0; 0 \leq u_2 \leq \eta\}} |\lambda(t) - \lambda(u_2)| > \varepsilon/2\right).$$

Now, $u_2 \in (c,b)$ then choosing $[t_1,t_2] \subset (c,b)$ and $\eta > 0$ such that both

$$\{t \geq 0; |u_2 - t| \leq \eta\} \subset [t_1,t_2]$$

and

$$\sup_{\{t \geq 0; 0 \leq u_2 \leq \eta\}} |\hat{\lambda}(t) - \lambda(u_2)| \leq \varepsilon/2,$$

(the later is possible since $\hat{\lambda}$ is uniformly continuous on $[t_1,t_2]$). Then, for such a choice of $\eta > 0$ and $[t_1,t_2]$ we have

$$P(|\hat{\lambda}(X_{(2m;m:n)} - \lambda(u_2)| > \varepsilon)$$

$$\leq P(|X_{(2m;m:n)} - u_2| > \eta) + P\left(\sup_{t \in [t_1,t_2]} |\hat{\lambda}^{(m)}(t) - \hat{\lambda}(t)| > \varepsilon/2\right),$$

where the two probabilities of the right-hand side of the above inequality tend to 0 as $m$ tends to infinity by Theorem 1(i) and (6). \(\square\)

**4. Application to reliability**

It is easy to see, by Theorem 1, that if $g(x) = 1 - (1 - x)^{r+1}$ then

$$q_x^{(m)} = X_{[g(x)m;m:n]} \xrightarrow{p} F^{-1}(x).$$
as \( m \to +\infty \). Since \( g \) is unknown due to the constant \( r \), we define \( \tilde{r} = m^{-1} \sum_{j=1}^{m} r_j \). Then a natural estimator for the \( \alpha \)-quantile of \( F \) is

\[
\hat{q}_x^{(m)}(F) \equiv X_{[g^{(m)}(x)m]_{m:n}}.
\]

**Remark 1.** If \( r = 0 \) then \( g^{(m)}(x) = x \) and our estimator coincides with the usual empirical quantile for a full sample.

**Proposition 3.** Under A1–A3(a) we have

(i) for all \( x \in [0, a] \):

\[
\lim_{m \to \infty} \frac{\hat{q}_x^{(m)}}{F^{-1}(x)} = 1.
\]

(ii) if A3(b) and A4 are satisfied, then, for \( x \in [\epsilon, a] \):

\[
\sqrt{m}(\hat{q}_x^{(m)} - F^{-1}(x)) \overset{D}{\to} N(0, \sigma^2(g(x))).
\]

**Remark 2.** In Proposition 3(ii) we restrict \( x \) to \([\epsilon, a]\). Indeed, such a restriction is necessary since there exists densities for which \( \sigma(x) \to +\infty \) as \( x \to 0 \). This is the case, e.g., for the density \( t \mapsto t^2/21[0,1](t) \), for which \( \sigma^2(x) \) is equivalent to \( x^{-1/3} \) in a neighborhood of 0.

**Remark 3.** Note that results of Proposition 3 can be obtained uniformly under restrictive conditions. For example, if in addition to the assumptions of Proposition 3 we have \( \lambda > 0 \) on \( \mathbb{R}^+ \) and \( |m^{\gamma}_m| < 1 \) then we can show that

\[
\sup_{x \in [\epsilon, a]} \left| \frac{\hat{q}_x^{(m)}(x) - F^{-1}(x)}{\sqrt{m}} \right| \to 0
\]

and \( \sqrt{m}(\hat{q}_x^{(m)}(\cdot) - F^{-1}(\cdot)) \) converges weakly in \( D[\epsilon, a] \) to a Gaussian process with variance function \( \sigma^2 \circ g \).

**Proof of Proposition 3.** First set

\[
\alpha_{\min} = \min(\{g^{(m)}(x)m\}, \{g(x)m\}) \quad \text{and} \quad \alpha_{\max} = \max(\{g^{(m)}(x)m\}, \{g(x)m\}).
\]

For \( b = 1 - (1 - a)^{K+1} \in (0, 1) \) we have

\[
\forall m \geq 1, \quad 0 \leq \alpha_{\max} \leq bm
\]

and

\[
\alpha_{\max} - \alpha_{\min} \leq m|g^{(m)}(x) - g(x)| + 1 \leq m|(1 - x)^{\tilde{r}+1} - (1 - x)^{r+1}| + 1 \leq m(1 - x)|\tilde{r} - r| + 1.
\]

It results that

\[
\alpha_{\max} - \alpha_{\min} = \mathcal{O}(m^{\gamma}_m)
\]

(7)
and
\[
\max_{j_{\min} \leq j \leq j_{\max}} \frac{1}{m} \frac{x_j}{z_j(m)} \leq \frac{1}{m - z_{\max}(m)} \leq \frac{1}{(1 - b)m}.
\]
(8)

Let us prove (i). We have
\[
\tilde{q}_x^{(m)} = \hat{q}_x^{(m)} + (\hat{q}_x^{(m)} - \tilde{q}_x^{(m)}),
\]
since \( \tilde{q}_x^{(m)} = X_{[g(x)m]:m:n} \), by Theorem 1(i) we have
\[
\tilde{q}_x^{(m)} \rightarrow q_{g(x)} = F^{-1}(x).
\]
It remains to show that
\[
\hat{q}_x^{(m)} - \tilde{q}_x^{(m)} = A^{-1} \left( \sum_{j=1}^{[g(x)m]} Z_j \frac{x_j}{z_j(m)} \right) - A^{-1} \left( \sum_{j=1}^{[g(x)m]+1} Z_j \frac{x_j}{z_j(m)} \right) \rightarrow 0.
\]
By Theorem 1 we know that \( \sum_{j=1}^{[g(x)m]} Z_j / x_j^m \) converges in probability to \( q_{g(x)} \), moreover \( A^{-1} \) is continuous, then, the above convergence result is true if
\[
U^{(m)} = \left| \sum_{j=1}^{[g(x)m]} \frac{Z_j}{x_j^m} - \sum_{j=1}^{[g(x)m]+1} \frac{Z_j}{x_j^m} \right| = \sum_{j=\max} \frac{Z_j}{x_j^m} \rightarrow 0.
\]
Since \( \gamma_m = o(1) \), we have by (7) and (8)
\[
EU^{(m)} \leq (z_{\max}^{(m)} - z_{\min}^{(m)}) \times \max_{x_{\min} \leq j \leq x_{\max}^{(m)}} \frac{1}{x_j^m} = o(1),
\]
then (i) follows by the Markov inequality applied to \( U^{(m)} \).

Let us prove (ii). We have
\[
\sqrt{m}(\hat{q}_x^{(m)} - F^{-1}(x)) = \sqrt{m}(X_{[g(x)m]:m:n} - q_{g(x)}) + \sqrt{m}(\tilde{q}_x^{(m)} - \hat{q}_x^{(m)}).
\]
From Theorem 1(ii) we have
\[
\sqrt{m}(X_{[g(x)m]:m:n} - q_{g(x)}) \rightarrow N(0, \sigma^2(g(x))).
\]
To achieve the proof of (ii) it remains to show that
\[
\sqrt{m}(\tilde{q}_x^{(m)} - \hat{q}_x^{(m)}) \rightarrow 0,
\]
or equivalently, by using a Taylor expansion, that
\[
V^{(m)} = \sqrt{m} \sum_{j=\max^{(m)}} \frac{Z_j}{x_j^m} \rightarrow 0.
\]
Again using (7), (8) and the assumption \( \gamma_m = o(m^{-1/2}) \) we get
\[
EV^{(m)} \leq \sqrt{m}(z_{\max}^{(m)} - z_{\min}^{(m)}) \times \max_{x_{\min} \leq j \leq x_{\max}^{(m)}} \frac{1}{x_j^m}
\]
\[
\leq \sqrt{m} O(m \gamma_m) O(m^{-1}) = o(1),
\]
and again (ii) follows by applying the Markov inequality to \( V^{(m)} \).
5. Monte-Carlo studies

We simulate type-II progressively censored data with hazard rate of a Weibull distribution function \( F(t) = (1 - \exp(t/5)^3)1(t \geq 0) \). The \( r_i \)'s are chosen at random in order to get the rate \( \gamma_m = o(m^{-1/2}) \). We provide simulations by generating \( r_i \)'s that satisfy \( r_i = 3 - e_i \), where the random variables \( e_i \)'s are i.i.d. with \( P(e_i = 1) = 1 - P(e_i = 0) = m^{-1/2} \) which leads to the rate \( \gamma_m = o(m^{-1/2}) \).

Here we provided \( N = 1000 \) samples of size \( m = 50 \). For each sample we calculated the quantile estimator, therefore its empirical mean, and standard deviation. Fig. 1 shows this empirical mean of the quantile estimator, and the corresponding 95% pointwise confidence band (obtained by using results of Proposition 3). We can see the poor behavior of the estimator for values of \( \alpha \) that are close to one. Of course it is no longer true for large values of \( m \).

Fig. 2 shows the behavior of the estimator of function \( \sigma^2 \). We compare the mean of \( N \) estimators \( \tilde{\sigma}^2(x) \) with \( \sigma^2(x) \), and with the empirical variance of \( \hat{g}^{(m)}(x) \). We can see that the bias of \( \tilde{\sigma}^2(x) \)
increases with $\alpha$, whereas the empirical variance is far from the true asymptotic value $\sigma^2(\alpha)$ for $\alpha \geq 0.5$. Note that with this kind of censoring scheme, tail data are more likely to be removed during the experiment. This corresponds to a loss of information concentrated on the distribution tail, which explains the behavior of the estimator in this region.

**Acknowledgements**

The authors thank the referee for comments leading to improvements in the paper, and N. Balakrishnan for his helpful comments during the visit of the second author to McMaster University.

**References**


