Ann Inst Stat Math (2013) 65:589–615
DOI 10.1007/s10463-012-0385-6

New estimating equation approaches with application in lifetime data analysis

Keming Yu · Bing Xing Wang · Valentin Patilea

Received: 16 January 2012 / Revised: 8 August 2012 / Published online: 11 November 2012

Abstract Estimating equation approaches have been widely used in statistics inference. Important examples of estimating equations are the likelihood equations. Since its introduction by Sir R. A. Fisher almost a century ago, maximum likelihood estimation (MLE) is still the most popular estimation method used for fitting probability distribution to data, including fitting lifetime distributions with censored data. However, MLE may produce substantial bias and even fail to obtain valid confidence intervals when data size is not large enough or there is censoring data. In this paper, based on nonlinear combinations of order statistics, we propose new estimation equation approaches for a class of probability distributions, which are particularly effective for skewed distributions with small sample sizes and censored data. The proposed approaches may possess a number of attractive properties such as consistency, sufficiency and uniqueness. Asymptotic normality of these new estimators is derived. The construction of new estimation equations and their numerical performance under different censored schemes are detailed via Weibull distribution and generalized exponential distribution.

K. Yu
Department of Mathematical Sciences, Brunel University, London UB8 3PH, UK
e-mail: keming.yu@brunel.ac.uk

K. Yu
Business School, Shihezi University, Wujiaqu, China

B. X. Wang (✉)
Department of Statistics, Zhejiang Gongshang University, Hangzhou 310018, China
e-mail: wangbingxing@163.com

V. Patilea
Ecole Nationale de la Statistique et de l'Analyse de l'Information (Ensai), (CRES),
Rue Blaise Pascal, BP 37203, 35172 BRUZ Cedex, France
e-mail: valentin.patilea@gmail.com

Keywords Estimation equation · Nonlinear combination of order statistics · Asymptotic normality · Weibull distribution · Generalized exponential distribution

1 Introduction

Consider the general problem of inferring a family of distribution $F(x; \lambda, \alpha)$ for various forms of censoring and also for the complete sample case. Statistical inference (point estimation and interval estimation) often include inference on unknown parameters λ and α, even a function of λ and α such as mean and quantile, as well survival function or reliability function.

In this paper, $F(x; \lambda, \alpha)$ belongs either of the family

$$F(x; \lambda, \alpha) = 1 - [1 - G(x; \lambda)]^{C(\alpha,\lambda)}, \tag{1}$$

or the family

$$F(x; \lambda, \alpha) = [G(x; \lambda)]^{C(\alpha,\lambda)}, \tag{2}$$

where $(a, t) \mapsto C(a, t)$ is a bivariate function defined on the parameters set and function $G(\cdot; \lambda)$ is a distribution function dependent only on λ.

The distribution family with the form of $F(x; \lambda, \alpha)$ includes Weibull distribution and many known distributions. It also includes a newly proposed generalized exponential distribution (Gupta and Kundu 1999, 2001, 2006, 2007; Mitra and Kundu 2008; Raqab 2002; Kundu et al. 2005), which can be used as an alternative to Gamma or Weibull distribution in many situations and has attracted much attention in literature recently. The quantity $C(\alpha, \lambda)$ could depend only on λ or could allow for a new parameter α. We focus on the later case and we suppose that, for each fixed t, the map $a \mapsto C(a, t)$ is one-to-one. In other words, the map $(\alpha, \lambda) \mapsto (\eta, \lambda) = (C(\alpha, \lambda), \lambda)$ is a reparametrization of the family of distribution functions (1) and (2). Clearly, one can work with the parameters (η, λ) and not introduce $C(\alpha, \lambda)$. However, there might be no consensus on a unique parametrization of families of laws to be used in applications (for instance because of the interpretation the parameters may have), in which case it could in interesting to allow for our $C(\alpha, \lambda)$. A well-known example is the Weibull distribution which may have different parametrizations in different textbooks.

Estimating equation approaches have been widely used in statistics inference of probability distribution. Important examples of estimating equations are the likelihood equations. However, MLE may produce substantial bias, or even fail for interval estimation as when the distribution is highly skewed, or sample size is not large enough or the data are highly censored. Take the Weibull distribution as an example, which has the distribution function as

$$F(x; \alpha, \lambda) = 1 - \mathrm{e}^{-(\alpha x)^{\lambda}}, \quad x > 0, \tag{3}$$

where $\alpha > 0, \lambda > 0$ are unknown parameters. With this parametrization we have $C(\alpha, \lambda) = \alpha^{\lambda}$ in Eq. (1). The MLEs of its parameters have been discussed by a number of authors. For the point estimation, since the solution is numerical, issues of existence and uniqueness of the estimates have to be addressed, which in the case of censored data can get quite involved. For the interval estimation, many software programs compute confidence limits based on the standard errors of the maximum likelihood estimates, however Dodson (1994) cautions against the interpretation of confidence limits computed from MLEs. In general, when the shape parameter λ is <2, the variance estimates computed for MLEs lack accuracy.

In this paper, based on nonlinear combinations of order statistics, we have explored new estimation equation approaches to inference the family of probability distributions. We expect new estimation equation approaches to provide not only unique and consistent parameter estimates but also exact confidence intervals, including exact confidence interval for reliability, which is challenging for existing methods of estimation equations.

Section 2 illustrates the general structure of the estimation equations for the family of probability distributions with both right censored and left-censored samples. Section 3 features construction details from some of the more typical examples of censored schemes, including exact confidence intervals for reliability function. Section 4 illustrates the numerical performance of new estimation equation approaches under a variety of scenarios. Section 5 derives the asymptotic normality of new parameter estimators. Section 6 concludes with a brief account of new results.

2 The basic estimation equations

Order statistics appears in a natural way in inference procedures when the sample is censored. Linear combinations of order statistics, such as L-statistics which have been shown to exhibit the desirable property of robustness, play an important role in the theory of estimation (Hampel et al. 1986; Giorgi 1999; Jones and Zitikis 2003; among others). In particular, Hosking (1990, 1995) discussed L-moments (the expectations of certain L-statistics) as the analysis and estimation of distributions. But there are virtually no research topics on some sophisticated nonlinear combinations of order statistics (NL-statistics) for statistics inference.

Given a random independent sample $(X_1, X_2, \ldots, X_n)$ from a distribution function $F(x; \lambda, \alpha)$, let $(X_{(1)}, X_{(2)}, \ldots, X_{(n)})$ be the induced order statistics, then a general class of NL-statistics could be defined as $\sum_i a_{ni} f_{ni}(X)$, where $f_{ni}(i = 1, \ldots, n)$ is some selected functionals such that the NL-statistic possess expected properties. In this paper, write $F(x)$ for $F(x; \lambda, \alpha)$ and consider an NL-statistic of the form:

$$\text{NL-statistic} = \sum_{i=1}^{n-1}(-2\log(U_{(i)})), \tag{4}$$

with

$$U_{(i)} = \frac{Y_i}{Y_n}, \quad (i = 1, \ldots, n-1),$$

$$Y_i = \sum_{j=1}^{i} W_j, \quad (i = 1, \ldots, n),$$

$$W_1 = nV_{(1)}, \ \ W_i = (n-i+1)(V_{(i)} - V_{(i-1)}), \quad (i = 2, \ldots, n),$$
$$V_{(i)} = -\log(1 - F(X_{(i)}; \lambda, \alpha)) \text{ with the distribution family (1)}$$
$$\text{or } V_{(i)} = -\log(F(X_{(n-i+1)}; \lambda, \alpha)) \text{ with the distribution family (2).} \tag{5}$$

Clearly, for a generic $F(\cdot)$, $\{V_{(i)}\}_{i=1}^{n}$ are the corresponding order statistics from the standard exponential distribution. However, while $V's$ are not independent $W's$ are. Eventually, the NL-statistic given by (4) has an exact χ^2 distribution. As $\{U_{(i)}\}_{i=1}^{n-1}$ can be regarded as induced order statistics from an independent sample, $\{U_i\}_{i=1}^{n-1}$, from a standard distribution $U(0, 1)$. Thus, the NL-statistic in (5) can be equivalently written as a simple sum of independent, identically distributed random variables with finite moments, then it follows from the Strong Law of Large Numbers strong-consistency exists for the statistic.

Moreover, the ratio Y_i and Y_n for $U_{(i)}$ cancels out parameter α, as $U_{(i)}$ does not contain parameter $C(\alpha, \lambda)$ for $F(x; \alpha, \lambda)$. This idea is also used in the analysis of Weibull distribution analysis with progressive censored data (Wang et al. 2010). It is well known that the accuracy of statistical inference for a particular parameter can be considerably improved if there is no nuisance parameter involved. Existing methods for this aim typically include likelihood-based conditional inference. Modern Gibbs sampling and MCMC, which are developed for this aim, are able to integrate out nuisance parameters. However, MLE-based equation equations or other estimation equation approaches cannot cancel out a nuisance parameter.

2.1 Estimation equation for the family of distributions under censored samples

In reliability analysis and biomedical research, censoring often exists. Depending on the direction of the censoring, censored data can be classified into right censored and left censored. Right censored includes type-I and type-II right censored, depending on whether the failure number is random or fixed in advance. Type I: completely random dropout (e.g. emigration) and/or fixed time of end of study no event having occurred, and type II: study ends when a fixed number of events amongst the subjects has occurred.

We first consider estimation equation approach with a typical right-censored sample.

Let $m < n$ and $X_{(1)} \leq \cdots \leq X_{(m)}$ be a type-II-censored sample from the family (1) with

$$F(x; \lambda, \alpha) = 1 - [1 - G(x; \lambda)]^{C(\alpha, \lambda)},$$

then

$$S_{(1)} = -\log[1 - G(X_{(1)}; \lambda)] \leq \cdots \leq S_{(m)} = -\log[1 - G(X_{(m)}; \lambda)]$$

is a type-II-censored sample from the exponential distribution with mean $1/C(\alpha, \lambda)$.

Write $Y_i = S_{(1)} + \cdots + S_{(i)} + (n-i)S_{(i)}$, $i = 1, \ldots, m$, then along the same line as construction of Eq. (4), we have the estimation equations for the family (1) under a type-II right-censored sample as:

$$\sum_{i=1}^{m-1} \log\left(\frac{Y_m}{Y_i}\right) = m - 2, \tag{6}$$

$$C(\alpha, \lambda) = \frac{m-1}{Y_m},$$

$$\alpha = C^{-1}\left(\frac{m-1}{Y_m}, \lambda\right), \tag{7}$$

where $C(., \lambda)^{-1}$ is the inverse of $a \mapsto C(a, \lambda)$.

Remark The estimation equations (6) and (7) still hold when $m = n$, which correspond to the case with complete sample.

Similarly, let $0 < r < n$ and $X_{(r+1)} \leq \cdots \leq X_{(n)}$ be a type-II left-censored sample from the family (2) with

$$F(x; \lambda, \alpha) = [G(x; \lambda)]^{C(\alpha,\lambda)},$$

then

$$S_{(1)} = -\log[G(X_{(n)}; \lambda)] \leq \cdots \leq S_{(n-r)} = -\log[G(X_{(r+1)}; \lambda)]$$

is a type-II-censored sample from the exponential distribution with mean $1/C(\alpha, \lambda)$.

Write $Y_i = S_{(1)} + \cdots + S_{(i)} + (n-i)S_{(i)}$, $i = 1, \ldots, r$. According to the discussion above, we have the estimation equations for the family (2) as:

$$\sum_{i=1}^{n-r} \log\left(\frac{Y_{n-r}}{Y_i}\right) = n - r - 2, \tag{8}$$

$$C(\alpha, \lambda) = \frac{n-r-1}{Y_{n-r}},$$

$$\alpha = C^{-1}\left(\frac{n-r-1}{Y_{n-r}}, \lambda\right). \tag{9}$$

Remark The estimation Eqs. (8) and (9) still hold when $r = 0$, which correspond to the case with complete sample.

3 Estimation equations and exact confidence intervals in action

3.1 Weibull distribution with type-II censoring

For type-II censoring, a reliability testing is ended when there is a prespecified number of failures. If X denotes the response variable, $X \geq 0$, and we are assuming that every subject follows the same distribution function $F(x)$, then the reliability function is $R(x) = Pr(X > x) = 1 - F(x)$. While confidence interval estimation of reliability function as well as mean time to failure is challenging for MLE and many existing estimation equation approaches, we consider new approach for a type-II-censored data consisting of failure times $X = (X_{(1)}, \ldots, X_{(m)})$, where $m(\leq n)$ is specified in advance. The test ends at the mth failure time $X_{(m)}$, and $(n-m)$ units have survived.

Note that, in the Weibull case, $C(\alpha, \lambda) = \alpha^{\lambda}$ and $S_{(i)} = X_{(i)}^{\lambda}$, hence the estimating Eq. (8) reduces to

$$W(\lambda) = 2\sum_{i=1}^{m-1} \log\left\{\frac{\sum_{j=1}^{m} X_{(j)}^{\lambda} + (n-m)X_{(m)}^{\lambda}}{\sum_{j=1}^{i} X_{(j)}^{\lambda} + (n-i)X_{(i)}^{\lambda}}\right\}.$$

Clearly, $W(\lambda)$ is a function of λ only and *does not depend on* α and has the χ^2 distribution with $2(m-1)$ degrees of freedom. Hence $W(\lambda)/(2m-4)$ converges with probability one to 1. Also $W(\lambda)$ is a strictly increasing function of λ due to that fact below:

$$\frac{Y_m}{Y_i} = 1 + \frac{\sum_{j=i+1}^{m}(X_{(j)}/X_{(i)})^{\lambda} + (n-m)(X_{(m)}/X_{(i)})^{\lambda} - (n-i+1)}{\sum_{j=1}^{i}(X_{(j)}/X_{(i)})^{\lambda} + (n-i+1)},$$

where $(X_{(j)}/X_{(i)})^{\lambda}$ is a strictly increasing function of λ (resp. decreasing) for $j >$ (resp. $<$)i. Therefore, the point estimator of λ from $W(\lambda) = 2(m-2)$ is unique. Let $\hat{\lambda}$ is a solution of $W(\hat{\lambda}) = 2(m-2)$. Similarly, note that

$$2\,C(\alpha, \lambda)\,Y_m = 2\,C(\alpha, \lambda)\left(\sum_{j=1}^{m} X_{(j)}^{\lambda} + (n-m)X_{(m)}^{\lambda}\right)$$

has the χ^2 distribution with $2m$ degrees of freedom, the estimating Eq. (9) reduces to provide the estimator $\hat{\alpha}$ of the parameter α as

$$\hat{\alpha} = \left(\frac{m-1}{\sum_{j=1}^{m} X_{(j)}^{\hat{\lambda}} + (n-m)X_{(m)}^{\hat{\lambda}}}\right)^{1/\hat{\lambda}}.$$

Furthermore, an exact $1-p(0<p<1)$ confidence interval for the parameter λ of the Weibull distribution which does not depend on α is given by:

$$\left[W^{-1}\{\chi^2_{1-p/2}(2(m-1))\},\ \ W^{-1}\{\chi^2_{p/2}(2(m-1))\}\right],$$

where $\chi^2_p(v)$ is the upper p percentile of the χ^2 distribution with v degrees of freedom and $W^{-1}(t)$ is the value of λ satisfying the equation $W(\lambda) = t$.

Let $g(W, X)$ be the unique solution of $W(\lambda) = W$, where $W \sim \chi^2(2(m-1))$. Notice that

$$\alpha = \left(\frac{Y_m}{\sum_{j=1}^{m} X_{(j)}^{\lambda} + (n-m)X_{(m)}^{\lambda}}\right)^{1/\lambda},$$

according to the substitution method given by Weerahandi (2004), we substitute $g(W, X)$ for λ in the expression for α we obtain the following generalized pivotal quantity for the parameter α:

$$T_1 = \left(\frac{Y_m}{\sum_{j=1}^{m} x_{(j)}^{g(W,x)} + (n-m)x_{(m)}^{g(W,x)}} \right)^{1/g(W,x)}, \tag{10}$$

where $x = (x_{(1)}, \ldots, x_{(m)})$ is the observed value of $X = (X_{(1)}, \ldots, X_{(m)})$.

Let $T_{1,p}$ denote the upper pth percentile of T_1, then $T_{1,1-p}$ and $T_{1,p}$ are the $1-p$ generalized lower and upper confidence limits for α, respectively. The values $T_{1,1-p}$ and $T_{1,p}$ can be obtained using Monte Carlo simulations which can be achieved using the following algorithm. For a given data set (n, m, x), generate $W \sim \chi^2(2m-2)$ and $2C(\alpha, \lambda)Y_m \sim \chi^2(2m)$, independently. Using these values, we compute the values of T_1 in (12). This process of generating the value of T_1 is repeated m_1 times for the fixed values of (n, m, x). Based on the generated values of T_1, the values $T_{1,1-p}$ and $T_{1,p}$ can be obtained.

Now notice that the survival function of the Weibull distribution is given by $S(x_0) = \exp[-(\alpha x_0)^\lambda]$, along the same line as the derivation of T_1 for parameter α, we obtain a generalized pivotal quantity for $S(x_0)$ as

$$T_2 = \exp\left(-\frac{Y_m}{\sum_{j=1}^{m} x_{(j)}^{g(W,x)} + (n-m)x_{(m)}^{g(W,x)}} \cdot x_0^{g(W,x)} \right). \tag{11}$$

Let $T_{2,p}$ denotes the upper pth percentile of T_2. Then $T_{2,p}$ is the $1-p$ lower confidence limits for $S(x_0)$.

In the following section, we study the performance of coverage probabilities of these confidence intervals for α and $S(x_0)$ via simulation.

3.2 Generalized exponential distribution with left censoring

Now consider a reliability analysis based on the generalized exponential distribution (Gupta and Kundu 1999). The probability density function and the cumulative distribution function of a generalized exponential distribution are given by

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

and

$$F(x, \alpha, \lambda) = \left(1 - \mathrm{e}^{-\lambda x}\right)^{\alpha}, \quad x > 0,$$

respectively, where $\alpha > 0, \lambda > 0$ are the unknown shape and scale parameters, respectively. Henceforth, for simplicity of notations let the generalized exponential distribution be denoted by GE(α, λ). For $0 < \alpha \leq 1$, the density function is a decreasing function and for $\alpha > 1$, it becomes an uni-modal function.

While its peer members such as Weibull and Gamma distributions are particularly useful in censored data analysis, however, (Gupta and Kundu 2007) noted that not much development has taken place for censored data with the generalized exponential distribution. In particular, interval estimation maybe difficult to obtain for this distribution with existing methods.

Let $X_{(r+1)}, \ldots, X_{(n)}$ be the last $n - r$ order statistics from GE(α, λ) with sample size n. If $r = 0$, then the left-censored sample corresponds to complete sample. Hence complete sample is a special case of left-censored sample.

To derive confidence interval for the parameter λ, the following Lemma 1 is needed and its proof is given in the Appendix.

Lemma 1 *Let*

$$f(\lambda) = \frac{\log(1 - \mathrm{e}^{-b\lambda})}{\log(1 - \mathrm{e}^{-a\lambda})},$$

where $b > a > 0$ are constants. Then $f(\lambda)$ is strictly decreasing on $(0, +\infty)$.

We now discuss the interval estimation of the parameter λ.

Since $X_{(r+1)}, \ldots, X_{(n)}$ is a left-censored sample from GE(α, λ) with sample size n, then

$$-\alpha \log\left(1 - \mathrm{e}^{-\lambda X_{(n)}}\right), -\alpha \log\left(1 - \mathrm{e}^{-\lambda X_{(n-1)}}\right), \ldots, -\alpha \log\left(1 - \mathrm{e}^{-\lambda X_{(r+1)}}\right)$$

is a type-II-censored sample from the standard exponential distribution with sample size n.

Note that, in the generalized exponential distribution, $C(\alpha, \lambda) = \alpha$ and $S_{(i)} = -\log\left(1 - \mathrm{e}^{-\lambda X_{(n-i+1)}}\right)$, hence the estimating Eq. (10) reduces to

$$W(\lambda) = 2 \sum_{i=1}^{n-r-1} \log\left(\frac{Y_{n-r}}{Y_i}\right).$$

This $W(\lambda)$ has the χ^2 distribution with $2(n - r - 1)$ degrees of freedom. Further, from

$$\frac{Y_{n-r}}{Y_i} = 1 + \frac{\frac{S_{(i+1)}}{S_{(i)}} + \cdots + \frac{S_{(n-r)}}{S_{(i)}} + r\frac{S_{(n-r)}}{S_{(i)}} - (n-i)}{\frac{S_{(1)}}{S_{(i)}} + \cdots + \frac{S_{(i-1)}}{S_{(i)}} + (n-i+1)},$$

we have from Lemma 1 that the $W(\lambda)$ is strictly increasing on $(0, +\infty)$, and

$$\lim_{\lambda \to 0+} W(\lambda) = 0, \lim_{\lambda \to +\infty} W(\lambda) = +\infty.$$

Therefore, we have

$$P(W^{-1}(\chi^2_{1-p/2}(2n-2r-2)) \le \lambda \le W^{-1}(\chi^2_{p/2}(2n-2r-2)))$$
$$= P(\chi^2_{1-p/2}(2n-2r-2) \le W(\lambda) \le \chi^2_{p/2}(2n-2r-2)) = 1-p.$$

In all, the exact confidence interval for λ could be described as follows:

Suppose $X_{(r+1)}, \dots, X_{(n)}$ is a left-censored sample from $\mathrm{GE}(\alpha, \lambda)$ with sample size n. Then, for any $0 < p < 1$,

$$[W^{-1}(\chi^2_{1-p/2}(2n-2r-2)), W^{-1}(\chi^2_{p/2}(2n-2r-2))]$$

is a $1-p$ confidence interval for the scale parameter λ.

The point estimators of λ and α can be obtained as below:

The estimator $\hat{\lambda}$ of λ is from the following equation:

$$W(\lambda) = 2n - 2r - 4, \tag{12}$$

and by Lemma 1, the solution of Eq. (14) is unique.

Noting that $V = 2\alpha Y_{n-r}$ has χ^2 distribution with $2n-2r$ degrees of freedom, the estimator $\hat{\alpha}$ of α is from the following equation:

$$\hat{\alpha} = \frac{n-r-1}{Y_{n-r}}. \tag{13}$$

Finally, similar to Eq. (13) for the Weibull distribution, we can obtain the generalized pivotal quantity for the survival function $S(x_0) = 1 - \left(1 - \mathrm{e}^{-\lambda x_0}\right)^{\alpha}$ of generalized exponential distribution as

$$T_3 = 1 - \left(1 - \mathrm{e}^{-g(W,x)x_0}\right)^{V/\left(-2\left[\sum_{i=1}^{n-r} \log(1-\mathrm{e}^{-g(W,x)\,x_{(n-i+1)}})+r\log(1-\mathrm{e}^{-g(W,x)\,x_{(r+1)}})\right]\right)},$$

where $g(W, X)$ is the unique solution of $W(\lambda) = W$, and $x = (x_{(r+1)}, \dots, x_{(n)})$ is the observed value of $X = (X_{(r+1)}, \dots, X_{(n)})$.

4 Numerical performance

To assess the finite sample properties of the proposed estimation equation approaches, in particular, the performance of new confidence intervals, we carry out numerical analysis for Sect. 3 under a variety of scenarios.

4.1 Performance of Weibull distribution

First, a simulation study was conducted to study the coverage probabilities and the average interval lengths of the proposed confidence intervals for survival function under Weibull distribution with type-II right-censored sampling discussed in Sect. 3.1.

Table 1 Coverage probabilities and the average interval lengths (in parentheses) for the confidence intervals of α and $S(1.2)$ with Weibull distribution

(n, m)	GCI				BBCP			
	α		$S(1.2)$		α		$S(1.2)$	
	0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95
(10, 5)	0.911	0.944	0.908	0.945	0.740	0.802	0.796	0.817
	(0.981)	(1.189)	(0.519)	(0.599)	(0.686)	(0.849)	(0.408)	(0.447)
(10, 8)	0.886	0.946	0.887	0.941	0.800	0.866	0.837	0.885
	(0.719)	(0.906)	(0.375)	(0.442)	(0.556)	(0.678)	(0.378)	(0.438)
(20, 10)	0.900	0.957	0.905	0.960	0.819	0.863	0.847	0.878
	(0.645)	(0.779)	(0.398)	(0.464)	(0.530)	(0.644)	(0.349)	(0.394)
(20, 15)	0.914	0.952	0.908	0.952	0.859	0.904	0.872	0.911
	(0.472)	(0.578)	(0.293)	(0.346)	(0.411)	(0.494)	(0.287)	(0.336)
(30, 10)	0.895	0.951	0.898	0.947	0.846	0.908	0.869	0.903
	(0.756)	(0.897)	(0.456)	(0.524)	(0.656)	(0.798)	(0.390)	(0.434)
(30, 15)	0.895	0.951	0.904	0.953	0.868	0.918	0.892	0.933
	(0.508)	(0.611)	(0.339)	(0.397)	(0.449)	(0.541)	(0.318)	(0.364)
(50, 15)	0.902	0.941	0.899	0.946	0.836	0.897	0.863	0.904
	(0.641)	(0.760)	(0.409)	(0.472)	(0.594)	(0.718)	(0.357)	(0.402)
(50, 25)	0.896	0.959	0.910	0.951	0.861	0.910	0.877	0.919
	(0.382)	(0.458)	(0.275)	(0.323)	(0.361)	(0.433)	(0.260)	(0.302)

Since α is the scale parameter and the estimators are appropriately scale equi- and invariant, in our simulation study we take $\alpha = 1$.

In Table 1, we report the coverage probabilities and average lengths of the generalized confidence intervals (GCI) at confidence levels 0.9 and 0.95 for the scale parameter α and the survival function $S(1.2)$ when $\lambda = 2$. We also compare the performance of GCI with bootstrap bias-corrected percentile confidence intervals (BBCP). These values were computed over 1000 replications for each different case using $m_1 = 1000$. The average interval lengths are given in parentheses. Clearly, from so many different combinations of censoring schemes and (small) sample sizes, the simulated probabilities with the proposed GCI for confidence levels 0.9 and 0.95 of are quite close to their corresponding nominal levels, but the simulated probabilities of BBCP for these levels are not close to their corresponding nominal levels in most of cases.

4.2 Performance of generalized exponential distribution

To assess the finite sample properties of the proposed approaches for generalized exponential distribution discussed in Sect. 3.2, including parameter estimation and confidence interval for survival function, a simulation study was conducted to compare the performance of the proposed point estimators with MLEs used by Mitra and Kundu (2008) for the generalized exponential distribution GE(λ, α).

Table 2 the relative biases and relative MSE of the estimators when $\alpha = 1.0$

n	r	Bias				MSE			
		α		λ		α		λ	
		NL	MLE	NL	MLE	NL	MLE	NL	MLE
15	0	0.0247	0.2369	0.0057	0.1795	0.1967	0.3807	0.1519	0.2269
15	1	0.0321	0.2800	0.0050	0.1942	0.2644	0.5639	0.1636	0.2524
15	3	0.0623	0.4232	0.0043	0.2319	0.5306	1.5804	0.1885	0.3150
20	0	0.0116	0.1582	0.0002	0.1255	0.1215	0.1980	0.1043	0.1407
20	1	0.0150	0.1778	0.0000	0.1333	0.1395	0.2383	0.1088	0.1498
20	3	0.0219	0.2259	−0.0018	0.1497	0.1881	0.3625	0.1175	0.1694
20	5	0.0375	0.3042	−0.0023	0.1722	0.3036	0.6886	0.1337	0.2022
30	0	0.0099	0.1018	0.0029	0.0838	0.0707	0.0982	0.0632	0.0779
30	1	0.0103	0.1082	0.0025	0.0866	0.0754	0.1070	0.0648	0.0807
30	5	0.0162	0.1452	0.0017	0.1010	0.1073	0.1678	0.0732	0.0948
30	10	0.0320	0.2291	−0.0001	0.1262	0.1959	0.3693	0.0869	0.1209
50	0	0.0046	0.0568	0.0013	0.0482	0.0385	0.0468	0.0371	0.0417
50	5	0.0068	0.0697	0.0013	0.0539	0.0475	0.0602	0.0400	0.0460
50	15	0.0141	0.1116	−0.0003	0.0686	0.0856	0.1214	0.0485	0.0585
100	0	0.0036	0.0289	0.0011	0.0240	0.0179	0.0198	0.0179	0.0189
100	10	0.0043	0.0346	0.0008	0.0265	0.0216	0.0244	0.0190	0.0204
100	30	0.0075	0.0532	0.0001	0.0336	0.0361	0.0431	0.0233	0.0255

We consider different sample sizes n and left-censored numbers r. Since λ is the scale parameter and all estimators are scale invariant, we take $\lambda = 1$ in our simulation study and consider different values of α. We report the average relative biases and average relative mean square errors (MSE) over 10000 replications for different cases. The simulation results are presented in Table 2 when $\alpha = 1$. The other results not being shown here have consistent result with following.

It is quite clear from the Table 2 that for a fixed sample size, as the left failure number r increases the average relative biases and average relative MSE's also increase as expected. It is also be observed that the average biases and the average MSE's decrease as sample size increases. It follows from simulation results that the proposed estimators outperform the MLEs from both bias and MSE for all cases. According to these simulation results, we suggest the proposed estimators for small and moderate sample sizes.

In Table 3 we report the coverage probabilities and average lengths of the GCI at confidence levels 0.9 and 0.95 for the scale parameter α and survival function of GE(λ, 2.5) when λ is unknown. The average interval lengths are given in parentheses. We also compare the performance of GCI with BBCP. These values were computed over 1,000 replications for each different case using $m_1 = 1000$. Similar to Table 1 under Weibull distribution case, the simulated probabilities of the proposed GCI under generalised exponential distribution are also much closes to their corresponding nominal levels than those of BBCP.

Table 3 Coverage probabilities and average interval lengths (in parentheses) for the confidence intervals of α and survival function of GE(λ, 2.5) with unknown λ

n	r	GCI				BBCP			
		α		$S(1.2)$		α		$S(1.2)$	
		0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95
15	0	0.908	0.948	0.914	0.947	0.870	0.924	0.902	0.937
		(6.260)	(8.307)	(0.314)	(0.370)	(5.585)	(7.572)	(0.343)	(0.411)
15	3	0.905	0.957	0.893	0.951	0.872	0.931	0.901	0.936
		(9.052)	(11.942)	(0.336)	(0.395)	(12.025)	(19.368)	(0.375)	(0.450)
20	0	0.912	0.954	0.896	0.952	0.856	0.921	0.901	0.941
		(3.560)	(4.331)	(0.275)	(0.325)	(4.233)	(5.507)	(0.295)	(0.353)
20	1	0.910	0.950	0.905	0.946	0.875	0.918	0.902	0.942
		(4.143)	(5.085)	(0.278)	(0.329)	(4.674)	(6.165)	(0.300)	(0.359)
20	5	0.889	0.941	0.897	0.942	0.880	0.931	0.909	0.939
		(7.752)	(10.006)	(0.302)	(0.356)	(8.420)	(12.169)	(0.330)	(0.397)
50	0	0.897	0.953	0.919	0.953	0.897	0.946	0.897	0.940
		(2.014)	(2.418)	(0.177)	(0.210)	(2.038)	(2.502)	(0.182)	(0.218)
50	5	0.901	0.953	0.921	0.955	0.880	0.941	0.898	0.942
		(2.279)	(2.746)	(0.181)	(0.215)	(2.361)	(2.920)	(0.188)	(0.224)
50	15	0.891	0.947	0.902	0.947	0.872	0.934	0.890	0.929
		(3.178)	(3.881)	(0.201)	(0.239)	(3.362)	(4.262)	(0.210)	(0.251)
50	25	0.899	0.946	0.899	0.954	0.876	0.925	0.871	0.919
		(5.179)	(6.540)	(0.254)	(0.300)	(6.115)	(8.258)	(0.261)	(0.313)

4.3 Two illustrative examples

Example 1 Consider the pseudo-random data from example 2 of Lawless (1975) which is also reproduced in Lawless (2003). These data were type-II right-censored data with $n = 40, r = 28$. Following Lawless (1975, 2003), we assume that the lifetime follows the Weibull distribution. According to page 223 of Lawless (2003), the (conditional) 90 % confidence intervals for the parameters λ and α are (0.7813, 1.3889) and (0.8958, 1.6653), respectively, and the lower limit of the (conditional) 95 % confidence interval for the survival function $S(x_0)$ with $x_0 = \exp(-1)$ is 0.647.

According to Sect. 3.1, the 90 % exact confidence interval for the parameter λ is (0.7805, 1.3785), 90 % generalized confidence interval for α is (0.8987, 1.6742) and the lower limit of the 95 % confidence interval for the survival function $S(x_0)$ with $x_0 = \exp(-1)$ is 0.6471. These are quite similar to the (conditional) confidence intervals. However, our method can obtain generalized confidence interval for an important parameter in lifetime data analysis, namely, the expected lifetime or the mean time to failure. It is noted that the estimation of this parameter via standard method such as MLE-based methods is very challenging. In fact, the mean of the Weibull distribution is given by $\mu = \Gamma(1 + 1/\lambda)/\alpha$, and along the same lines as the derivation of T_1 and T_2 in Eqs. (12) and (13) for parameter α and survival function $S(x_0)$, we obtain the following generalized pivotal quantity T_4 for μ: $T_4 = \left(\frac{\sum_{j=1}^{m} x_{(j)}^{g(W,x)} + (n-m)x_{(m)}^{g(W,x)}}{Y_m}\right)^{\frac{1}{g(W,x)}} \Gamma\left(1 + \frac{1}{g(W,x)}\right)$, then in this example the 90 % generalized confidence interval for μ is (0.8791, 1.7590).

Example 2 Let us consider the data sets on failure times of the air conditioning system of two different air planes from Gupta and Kundu (2003). Gupta and Kundu (2003) use the data to illustrate closeness of Gamma and generalized exponential distributions.

The left-censored sample generated from the data set 2 with $r = 3, n = 12$ is as follows: 55, 56, 104, 176, 182, 220, 239, 246, 320. The MLEs for the α and λ are 1.8397 and 0.0094, respectively. The NL-statistics-based estimators of α and λ from Eqs. (14) and (15), on the other hand, are $\hat{\alpha} = 1.4542, \hat{\lambda} = 0.0084$, respectively, which are quite different from their MLEs. Further, with the proposed approaches in Sect. 3.2, we are able to explore interval estimation for the parameters and many associated quantities of the distribution. For example, The 90 % exact confidence interval for the parameter λ is (0.0034, 0.0142) and the 90 % generalized confidence interval for the parameter α is (0.5707, 3.4952). The 90 % generalized confidence intervals for the survival function $S(x_0)$ at $x_0 = 100$ and expected lifetime are (0.3722, 0.7694) and (104.4428, 259.8360), respectively.

5 Asymptotic normality of estimations

Classical estimators like MLEs usually have asymptotic normality, which is also a fundamental issue for parameter estimation. In this section, we aim to derive the asymptotic normality of *point estimates* of the parameters λ and α. Although our presentation focuses on families of distributions supported on the nonnegative half-line, the proofs allow the supports to be any kind of bounded or unbounded interval of real numbers.

Let N be such that $N/n \to p \in (0, 1)$. The cases we have in mind are $N = m - 1$ and $N = n - r - 1$. If $F(\cdot; \alpha, \lambda)$ stands for the distribution function of the observations, let

$$W_n(a, t; N) = 2 \sum_{i=1}^{N} \log \left\{ \frac{Z_{N+1}(a, t)}{Z_i(a, t)} \right\} - 2(N - 1), \tag{14}$$

where

$$Z_i(a, t) = -\frac{1}{n} \left\{ \sum_{j=1}^{i} \log(1 - F(X_{(j)}; a, t)) + (n - i) \log(1 - F(X_{(i)}; a, t)) \right\},$$

if the model is the one defined in (1), or $Z_i(a, t)$ are defined in the same way but with $1 - F(X_{(j)}; a, t)$ replaced by $F(X_{(j)}; a, t)$ if model (2) is considered. For simplicity, hereafter we will only use $Z_i(a,t)$ defined as for model (1), but all the statements below obviously adapt for the model (2).

In view of Eqs. (1) and (2) above, we are interested in the cases where $W_n(a,t)$ does not depend on a. This means that $Z_i(a,t)$ can be factorized like $Z_i(a,t) = C(a,t)Z_i(t)$ with $C(a,t)$ depending only on the parameters (and not on the observations), and

$$Z_i(t) = -\frac{1}{n}\left\{\sum_{j=1}^{i}\log(1-G(X_{(j)};t)) + (n-i)\log(1-G(X_{(i)};t))\right\}$$

depending on the parameter t only. For instance, in the Weibull case, we have $C(a,t) = a^t$ and $1-G(x;t) = \exp(-x^t)$ and thus

$$W_n(t) = 2\sum_{i=1}^{N}\log\left\{\frac{\sum_{j=1}^{N+1}X_{(j)}^t + (n-N-1)X_{(N+1)}^t}{\sum_{j=1}^{i}X_{(j)}^t + (n-i)X_{(i)}^t}\right\} - 2(N-1). \tag{15}$$

Keeping in mind the factorization property, hereafter we replace definition (14) by

$$W_n(t) = 2\sum_{i=1}^{N}\log\left\{\frac{Z_{N+1}(t)}{Z_i(t)}\right\} - 2(N-1), \tag{16}$$

where $Z_i(t)$ is the (positive) factor in the decomposition of $Z_i(a,t)$ that depends only on t and the observations, that is $Z_i(t) = Z_i(a,t)/C(a,t)$. Here, and in the following, we simply write $W_n(t)$ instead of $W_n(t;N)$.

For simplicity, we focus on the asymptotic normality aspect and hence in the arguments we provide below we implicitly suppose that $\hat{\lambda}$ is consistent. Moreover, we take $N = [np]$ where $[\cdot]$ is the integer part function. It will be quite clear from below how these two simplifications could be dropped at the expense of more technicalities that will be omitted.

In the cases we have in mind, the function $t \mapsto W_n(t)$ is strictly monotonic. The estimator $\hat{\lambda}$ we propose is then the unique solution of the equation $W_n(t) = 0$, the true value of the parameter being $t = \lambda$. Therefore, for proving asymptotic normality we adopt a variant of the approach of Maritz (1995) and Brown (1985).[1] Following the notation of Maritz (1995), chapter 8, let

$$S_n(t) = \frac{W_n(t)}{2\sqrt{n}}.$$

Suppose that for each n the function $t \mapsto W_n(t)$ is continuously differentiable for almost all samples (that means, except for a set of zero probability under the sampling model) and let

[1] The monotonicity property of $t \mapsto W_N(t)$ is guaranteed for the examples provided in the paper, hence the approach in the textbook Maritz (1995) is sufficiently general in that cases. However, our theoretical arguments for proving asymptotic normality adapt rapidly to nonmonotonic maps $t \mapsto W_N(t)$.

$$c_n(t) = \frac{\nabla_t W_n(t)}{2n} = \frac{\nabla_t S_n(t)}{n^{1/2}},$$

(here ∇_t stands for the partial derivative with respect to the parameter t). Next, following Maritz (1995) let us write $t = \lambda + \delta n^{-1/2}$ and define

$$U(\delta) = S_n(t) - S_n(\lambda) - \delta c_n(\lambda).$$

Since $W_n(\lambda) + 2(N-1)$ has exactly a χ^2 distribution with $2(N-1)$ degrees of freedom, clearly

$$S_n(\lambda) \xrightarrow{D} N(0, p). \tag{17}$$

On the other hand, suppose that

$$c_n = c_n(\lambda) \to c_0, \quad \text{in probability}, \tag{18}$$

where c_0 is some nonzero constant. By a minor modification of the proof of Proposition 8.2 of Maritz (1995), see also the proof of Theorem 2 of Brown (1985), we can deduce that if the function $t \mapsto W_n(t)$ is monotone increasing, and conditions (17) and (18) hold true, and for all fixed δ

$$U_n(\delta) \to 0 \ \text{ as } n \to \infty, \quad \text{in probability}, \tag{19}$$

then

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0, pc_0^{-2}).$$

5.1 The asymptotic normality result for the parameter λ

Let us gather the previous facts in the following statement. For this purpose we need some notation. Define

$$\phi(x) = -\nabla_t[\log(1 - G(x; t))]_{t=\lambda}$$

and let $\phi'(\cdot)$ denote the derivative of $\phi(\cdot)$. Let $F(\cdot)$ be a short notation for $F(x; \alpha, \lambda)$ and $Q(u) = \inf\{x : F(x) \geq u\}$.

Theorem 1 *Suppose that $N/n \to p \in (0, 1)$ and that Property 1 in the Appendix holds true. Let $\hat{\lambda}$ be the unique solution of the equation $W_n(t) = 0$ with $W_n(\cdot)$ defined in (16). Then*

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0, pc_0^{-2}),$$

where

$$c_0 = C(\alpha, \lambda)\left[H(p) - \int_0^p \frac{H(u)}{u}\,\mathrm{d}u\right]$$

and

$$H(u) = \int_{Q(0)}^{Q(u)} \phi'(x)(1 - F(x))\,\mathrm{d}x, \qquad u \geq 0.$$

The technical conditions necessary for proving Theorem 1 are satisfied by common models, like for instance the Weibull model, a case that we consider in the following corollary.

Corollary 1 *Let* $F(x) = 1 - \exp(-(\alpha x)^\lambda)$, $x > 0$, *be the distribution function of a Weibull law of parameters* $\alpha, \lambda > 0$. *Let* $\hat{\lambda}$ *be the unique solution of the equation* $W_n(t) = 0$ *with* $W_n(\cdot)$ *defined in* (15). *Suppose that* $N/n \to p \in (0, 1)$ *and let*

$$c_0 = \widetilde{H}(p) - \int_0^p \frac{\widetilde{H}(u)}{u}\,\mathrm{d}u$$

where

$$\widetilde{H}(u) = \int_0^{\alpha^{-1}\log^{1/\lambda}[1/(1-u)]} \{\log(x) - \lambda^{-1}\}F'(x)\,\mathrm{d}x, \quad 0 < u \leq p.$$

If $c_0 \neq 0$, *then*

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0, pc_0^{-2}).$$

5.2 Asymptotics for α

Suppose that the conditions of Theorem 1 hold true. Consider the family of distributions functions defined in Eq. (1) above [(the arguments for (2) are similar]. Suppose that for each fixed t, the map $a \mapsto C(a, t)$ is one-to-one. Let $v \mapsto C^{-1}(v; t)$ denote the inverse of the map $a \mapsto C(a, t)$. In view of the identity

$$Y_N = -C(\alpha, \lambda)\left\{\sum_{j=1}^{N} \log(1 - G(X_{(j)}; \lambda)) + (n - N)\log(1 - G(X_{(N)}; \lambda))\right\} \tag{20}$$

where Y_N is a sum of N independent standard exponential, let us define an estimator of α as

$$\hat{\alpha} = C^{-1}(a_n(\hat{\lambda}); \hat{\lambda}),$$

where

$$a_n(t) = -\frac{N}{\sum_{j=1}^{N} \log(1 - G(X_{(j)}; t)) + (n - N)\log(1 - G(X_{(N)}; t))}.$$

The limit in probability of $a_n(\lambda)$ is $a(\lambda)$ defined by

$$\begin{aligned} a(\lambda) &= \frac{p}{\left[\int_{Q(0)}^{Q(p)} h_G(x;\lambda)(1 - F(x;\lambda,\alpha))\,\mathrm{d}x\right]} \\ &= -\frac{C(\alpha,\lambda)p}{\left[\int_{Q(0)}^{Q(p)} (1 - F(x;\lambda,\alpha))'\,\mathrm{d}x\right]} = C(\alpha,\lambda), \end{aligned}$$

where $h_G(x;\lambda) = G'(x;\lambda)/(1-G(x;\lambda))$ is the hazard function associated to $G(x;\lambda)$, while the limit in probability of $\nabla_t a_n(\lambda) = \nabla_t a_n(t)|_{t=\lambda}$ is given by

$$\begin{aligned} \nabla_t a_n(t)|_{t=\lambda} &= -p^{-1}\, a^2(\lambda) \int_{Q(0)}^{Q(p)} \nabla_t h_G(x;t)|_{t=\lambda}(1 - F(x;\lambda,\alpha))\,\mathrm{d}x \\ &= -p^{-1}\, C^2(\alpha,\lambda) \int_{Q(0)}^{Q(p)} \nabla_t h_G(x;t)|_{t=\lambda}(1 - G(x;\lambda))^{C(\alpha,\lambda)-1}\,\mathrm{d}x. \end{aligned}$$

The delta-method implies that

$$\sqrt{n}(\hat{\alpha} - \alpha) \xrightarrow{D} N(0, pc_0^{-2}[\nabla_t C^{-1}(a(\lambda);\lambda)]^2),$$

where

$$\nabla_t C^{-1}(a(\lambda);\lambda) = \nabla_t C^{-1}(a(t);t)|_{t=\lambda} = \frac{\nabla_t a(t)|_{t=\lambda}}{C(\alpha,\lambda)} + \nabla_t C^{-1}(C(\alpha,\lambda);t)|_{t=\lambda}.$$

6 Conclusion

This paper proposes new estimation equation approaches for a large class of probability distributions, including exact confidence intervals for parameters, a function of parameters and survival function under a variety of censored schemes. Asymptotic theory associated with the approaches are derived in a novel way. The new inference for both point estimation and interval estimation has been shown to have good performance, even more efficient than maximum likelihood estimation (MLE) for fitting lifetime distributions with censored data and small sample sizes. Next we maybe able to extend the work by introducing covariates into $C(\alpha,\lambda)$ in the distributions families (2) and (3), then consider modeling $C(\lambda,\alpha)$ by a parametric or nonparametric regression model, so that the research would fairly link to Cox model for survival analysis.

7 Appendix

Property 1 1. For almost all samples the map $t \mapsto W_n(t)$ defined in (16) is continuously differentiable and strictly monotonic.
2. For any α the function $t \mapsto C(\alpha, t) = n^{-1}\{E[Z_1(t)]\}^{-1}$ is continuously differentiable.
3. $c_0 = C(\alpha, \lambda)[H(p) - \int_0^p u^{-1} H(u)\, \mathrm{d}u] \neq 0$
4. Conditions (27), (28), (32), (33), (35) through (38), and (39) through (42) below hold true.

7.1 Proof of Theorem 1

To prove Theorem 1, we have to check conditions (18) and (19).

7.1.1 Proof of (18)

Below we only consider the case of model (1). The case of model (2) could be handled with obvious adaptations. In the following $c, C, \ldots$, are constants that may change from line to line. Since

$$c_n(\lambda) = \frac{N}{n}\left\{\frac{\nabla_t Z_{N+1}(\lambda)}{Z_{N+1}(\lambda)} - \frac{1}{N}\sum_{i=1}^{N}\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)}\right\},$$

we investigate the behavior of the quantities $\nabla_t Z_i(\lambda)/Z_i(\lambda)$. By construction $E[Z_i(\lambda)] = i/[C(\alpha, \lambda)n]$ and $var[Z_i(\lambda)] = i/[C(\alpha, \lambda)n]^2$. By the identity $y/x - y/x_0 = y(x_0 - x)/(x_0 x)$ and Cauchy–Schwarz inequality we have

$$\begin{aligned} E\left[\left|\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)} - \frac{\nabla_t Z_i(\lambda)}{E[Z_i(\lambda)]}\right|\right] &= E\left[\left|\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)} - \frac{\nabla_t Z_i(\lambda)}{iC^{-1}(\alpha, \lambda)n^{-1}}\right|\right] \\ &= \frac{C(\alpha, \lambda)}{i/n} E\left[\left|\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)}\{E[Z_i(\lambda)] - Z_i(\lambda)\}\right|\right] \\ &\leq \frac{C(\alpha, \lambda)}{i/n} E^{1/2}\left[\left\{\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)}\right\}^2\right] \mathrm{var}^{1/2}[Z_i(\lambda)] \\ &= \frac{1}{\sqrt{i}} E^{1/2}\left[\left\{\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)}\right\}^2\right]. \end{aligned} \tag{21}$$

Next, we impose the following mild assumption: there exists a sequence $r_n, n \geq 1$ such that, for all $i = 1, \ldots, N$,

$$E^{1/2}\left[\left\{\frac{\nabla_t Z_i(\lambda)}{Z_i(\lambda)}\right\}^2\right] \leq r_n \tag{22}$$

and $r_n/\sqrt{n} \to 0$ as $n \to \infty$. In this case, we have

$$E\left[|c_n(\lambda) - \overline{c}_n(\lambda)|\ \right] = (N/n)\left[\frac{r_n}{\sqrt{N}} - \frac{r_n}{N}\sum_{i=1}^{N}\frac{1}{\sqrt{i}}\right] \to 0, \qquad \text{as } n \to \infty,$$

where

$$\overline{c}_n(t) = \frac{C(\alpha,\lambda)N}{n}\left\{\frac{\nabla_t Z_{N+1}(t)}{(N+1)/n} - \frac{1}{N}\sum_{i=1}^{N}\frac{\nabla_t Z_i(t)}{i/n}\right\} =: \frac{C(\alpha,\lambda)N}{n}\overline{c}_{1n}(t). \tag{23}$$

Hence, condition (18) will be implied by the following one:

$$\overline{c}_n(\lambda) \to c_0, \quad \text{in probability.}$$

Next, we concentrate on the quantity $\overline{c}_{1n}(\lambda)$ defined of the right-hand side of (23). For any $u > 0$, consider the statistic

$$\begin{aligned} H_n(u) &= \frac{1}{n}\sum_{j=1}^{[nu]}\phi(X_{(j)}) + (1 - [nu]/n)\phi(X_{([nu])}) \\ &= \frac{1}{n}\sum_{j=1}^{[nu]}(n+1-j)[\phi(X_{(j)}) - \phi(X_{(j-1)})] \end{aligned}$$

where $\phi(\cdot)$ is some continuously differentiable function of the observations. The values $H_n(u)$ are related to the so-called total time on test, see Csörgő and Yu (1997). By construction, we have

$$\nabla_t Z_i(\lambda) = H_n(i/n)$$

with

$$\phi(x) = -\nabla_t[\log(1 - G(x;t))]_{t=\lambda}\ .$$

Let $F_n(\cdot)$ and $Q_n(\cdot) = F_n^{-1}(u) = \inf\{x : F_n(x) \ge u\}, 0 \le u \le 1$ be the empirical distribution and the left-continuous quantile functions, respectively. If $\phi'(\cdot)$ stands for the derivative of the function $\phi(\cdot)$, we can rewrite

$$H_n(u) = \int_{X_{(1)}}^{Q_n(u)} \phi'(x)(1 - F_n(x))\,\mathrm{d}x, \quad u \ge X_{(1)}.$$

The theoretical counterpart of $H_n(u)$ is defined by

$$H(u) = \int_{Q(0)}^{Q(u)} \phi'(x)(1 - F(x))\,\mathrm{d}x, \quad u \geq 0 \tag{24}$$

where hereafter $F(\cdot)$ is a short notation for $F(x; \alpha, \lambda)$ and $Q(u) = F^{-1}(u) = \inf\{x : F(x) \geq u\}$. Now, let

$$\overline{c}_{2n} = p^{-1} H_n(p) - p^{-1} \int_{1/(n+1)}^{p} \frac{H_n(u)}{u}\,\mathrm{d}u.$$

We claim that under mild conditions on the function $\phi(\cdot)$,

$$\begin{aligned} &\overline{c}_{2n} \to p^{-1} H(p) - p^{-1} \int_0^p \frac{H(u)}{u}\,\mathrm{d}u, \quad \text{in probability,} \\ &\text{and } \overline{c}_{2n} - \overline{c}_{1n}(\lambda) = o_P(1). \end{aligned} \tag{25}$$

Hence

$$c_n(\lambda) \to c_0 = C(\alpha, \lambda)\left[H(p) - \int_0^p \frac{H(u)}{u}\,\mathrm{d}u\right], \quad \text{in probability.} \tag{26}$$

Concerning the conditions on $\phi(\cdot)$, without any loss of generality, we set $\phi(Q(0)) = 0$. Moreover, we assume that

$$\phi(\cdot) \quad \text{is monotonic on an interval } (Q(0), r), \tag{27}$$

for some $Q(0) < r < Q(p)$ and

$$\int_0^{\epsilon} \frac{|\phi(Q(u))|}{u}\,\mathrm{d}u < \infty, \tag{28}$$

for some $\epsilon > 0$. The monotonicity condition (27) is a very mild convenient restriction. By simple algebra, it is easy to check that conditions (27) and (28) and the fact that $\phi'(\cdot)$ is bounded on $[r, Q(p)]$ guarantee that

$$\int_0^p \frac{|H(u)|}{u}\,\mathrm{d}u < \infty$$

and thus c_0 is well defined. A simple way to ensure $c_0 \neq 0$ is to check that the function $u \mapsto H(u)/u$ is monotonic on $(0, p)$. For instance, this is the case when $H(\cdot)$ is concave or convex on $(0, p)$.

Define

$$\widetilde{H}_n(u) = \int_{Q(0)}^{Q(u)} \phi'(x)(1 - F_n(x))\,\mathrm{d}x = \int_{X_{(1)}}^{Q(u)} \phi'(x)(1 - F_n(x))\,\mathrm{d}x$$

and notice that

$$\left|H(u)-\widetilde{H}_n(u)\right| \le \sup_{0\le x<\infty} |F_n(x)-F(x)| \int_{Q(0)}^{Q(u)} \left|\phi'(x)\right| \mathrm{d}x \le \int_{Q(0)}^{Q(u)} \left|\phi'(x)\right| \mathrm{d}x.$$

The integrability of the function $u \mapsto u^{-1} \int_{Q(0)}^{Q(u)} \left|\phi'(x)\right| \mathrm{d}x$ on $(0, p]$ follows from (27) and (28) and the boundedness of $\phi'(\cdot)$ on compact intervals. Next, from this, Glivenko–Cantelli theorem, and the dominated convergence theorem,

$$\int_{1/(n+1)}^{p} \frac{\widetilde{H}_n(u)}{u} \mathrm{d}u - \int_0^p \frac{H(u)}{u} \mathrm{d}u \to 0, \quad \text{almost surely.}$$

Moreover, it is clear that $H(p) - \widetilde{H}_n(p) \to 0$, almost surely. Hence, to prove the first part of (25) it suffices to show that

$$\bar{c}_{2n} - \widetilde{c}_{2n} \to 0, \quad \text{in probability,} \tag{29}$$

where

$$\widetilde{c}_{2n} = p^{-1}\widetilde{H}_n(p) - p^{-1} \int_{1/(n+1)}^{p} \frac{\widetilde{H}_n(u)}{u} \mathrm{d}u.$$

It is easy to see that the regularity of the function $\phi(\cdot)$ and the almost sure convergence of $Q_n(p)$ towards $Q(p)$ imply $\widetilde{H}_n(p) - H_n(p) \to 0$, almost surely. Hence for proving (29) it remains to show

$$\int_{1/(n+1)}^{p} \frac{\widetilde{H}_n(u) - H_n(u)}{u} \mathrm{d}u \to 0, \quad \text{in probability.} \tag{30}$$

Now, fix some x_0 such that $Q(p) < x_0 < Q(1)$. To control the large values of $Q_n(u)$, let us decompose

$$\begin{aligned} H_n(u) &= \int_{X_{(1)}}^{Q_n(u)\wedge x_0} \phi'(x)(1-F_n(x))\,\mathrm{d}x + \int_{Q_n(u)\wedge x_0}^{Q_n(u)} \phi'(x)(1-F_n(x))\,\mathrm{d}x \\ &=: H_{1n}(u) + H_{2n}(u). \end{aligned}$$

It is easy to check that

$$P\left(\int_{1/(n+1)}^{p} \frac{|H_{2n}(u)|}{u} \mathrm{d}u > 0\right) \le P(Q_n(p) > x_0) \to 0.$$

On the other hand, for some constant $C > 0$ we have

$$\begin{aligned} \left|\widetilde{H}_n(u) - H_{1n}(u)\right| &\le C\,|\phi(Q_n(u)\wedge x_0) - \phi(Q(u))| \\ &= C|\phi'(Q(\theta_n(u)))|\,|Q_n(u) - Q(u)| \end{aligned}$$

with $F(Q_n(u)) \wedge u \le \theta_n(u) \le F(Q_n(u) \wedge x_0) \vee u$. To derive (30) it suffices then to show

$$\int_{1/(n+1)}^{p} \frac{|\phi'(Q(\theta_n(u)))|}{f(Q(u))u^{1/2}\eta(u)} |\rho_n(u)| \, \mathrm{d}u = O_P(1) \tag{31}$$

for some continuous function $\eta(\cdot)$ such that $\lim_{u \downarrow 0} \eta(u) = 0$, where $f(\cdot) = F'(\cdot)$ is the density of the observations and, like in Csörgő and Horváth (1990), $\rho_n(u)$ is the quantile process, that is

$$\rho_n(u) = \sqrt{n} f(Q(u))(Q_n(u) - Q_(u)), \quad 0 < u < 1.$$

To show (31) we will apply Theorem 2.1 of Csörgő and Horváth (1990) with $p = 1$. However, before proceeding we have to bound $\phi'(Q(\theta_n(u)))$ for small values of u. For this purpose, let us suppose that there exists some constants $c > 0$ and $a, d \ge 0$ such that

$$|\phi'(x)| \le c f^a(x) F^{-d}(x), \quad \forall \; Q(0) < x \le x_0. \tag{32}$$

At this point, it becomes more clear why we consider $x_0 < Q(1)$: condition (32) may be too restrictive on the whole support of the observations. Now, since

$$F(Q_n(u)) \wedge u \le \theta_n(u) \le F(Q_n(u) \wedge x_0) \vee u \le F(Q_n(u)) \vee u$$

and

$$\begin{aligned}(1 - F(Q_n(u))) \wedge (1 - u) &\le (1 - F(Q_n(u) \wedge x_0)) \wedge (1 - u) \\ &\le 1 - \theta_n(u) \le (1 - F(Q_n(u))) \vee (1 - u),\end{aligned}$$

by minor modifications of the arguments used by Csörgő and Horváth (1990) on page 76, see also Lemma 1 of Csörgő and Révész (1978), deduce that if

$$\sup_{Q(0) < x < Q(1)} F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)} < \infty \tag{33}$$

and $f \ne 0$ on $(Q(0), Q(1))$, then

$$\sup_{1/(n+1) \le u \le n/(n+1)} \frac{f(Q(\theta_n(u)))}{f(Q(u))} = O_P(1).$$

This, condition (32) and Eq. (31) indicate that property (30) holds if

$$\int_{1/(n+1)}^{p} |\rho_n(u)|/q(u) \, \mathrm{d}u = O_P(1) \tag{34}$$

with

$$q(u) = u^{1/2+d}\eta(u) f^{1-a}(Q(u)).$$

To apply Theorem 2.1 of Csörgő and Horváth (1990), the functions $Q(\cdot)$, $f(\cdot)$ and $q(\cdot)$ should satisfy the following additional assumptions:

$$f(Q(\cdot)) \text{ is positive and continuous on } (0, p]; \tag{35}$$

$$\int_0^p u^{1/2}/q(u)\,\mathrm{d}u < \infty; \tag{36}$$

$$1/f(Q(u)) \le w(u), \quad \text{where } w(\cdot) \text{ is a monotone function on } (0, p]; \tag{37}$$

$$\int_0^p f(Q(u))w(u/\nu)u^{1/2}/q(u)\,\mathrm{d}u < \infty, \quad \forall \nu > 1 \text{ (resp. } \nu < 1) \tag{38}$$
$$\text{if } w(\cdot) \text{ is } \nearrow \text{ (resp. } \searrow) \text{ on } (0, p].$$

Theorem 2.1 of Csörgő and Horváth (1990) implies (34) and so eventually we obtain (30) and thus the first part of (25). The arguments for second part of (25) are elementary and hence will be omitted. □

7.1.2 The proof of (19)

By the Taylor expansion, $U(\delta) = S_n(t) - S_n(\lambda) - \delta c_n(\lambda) = \delta[c_n(\widetilde{t}) - c_n(\lambda)]$, where $|\widetilde{t} - \lambda| \le |\delta| n^{-1/2}$. To show (19), let us impose the following convenient assumptions that are satisfied in the common examples encounter in reliability analysis: for any fixed $p \in (0, 1)$, α and λ, one can chose some (small) $\zeta_0 > 0$ such that

- there exists a constant $C > 0$ such that

$$[1 - F(x, \alpha, \lambda)]^{1+C\zeta} \le 1 - F(x, \alpha, \lambda + \zeta) \le [1 - F(x, \alpha, \lambda)]^{1-C\zeta}, \tag{39}$$

  $\forall x \in (Q(0), Q(p))$ and $\forall |\zeta| \le \zeta_0$;
- a uniform version of condition (22) holds true, that is there exists a sequence r_n satisfying $r_n/\sqrt{n} \to 0$ such that $\forall\ 1 \le i \le N$

$$E^{1/2}\left[\sup_{|\zeta| \le \zeta_0} \left\{\frac{\nabla_t Z_i(\lambda + \zeta)}{Z_i(\lambda + \zeta)}\right\}^2\right] \le r_n \tag{40}$$

- for any $\bar{p} \in (0, p)$ there exist a function $h(\cdot) \ge 0$ with $h(\zeta) \to 0$ when $\zeta \to 0$ such that

$$\left|\phi'_{\lambda+\zeta}(x) - \phi'_\lambda(x)\right| \le \Phi(x)h(\zeta), \quad \forall x \in (Q(\bar{p}), Q(p)),\ \forall |\zeta| \le \zeta_0, \tag{41}$$

where $\phi_t(x) = -\nabla_t \log(1 - G(x;t))$ and $\Phi(\cdot)$ is a function independent of $\bar{p}$ which satisfies condition (32) with the same constants a and d like $\phi'_\lambda(\cdot)$ (possibly with a different constant c).

- for any $\bar{p} \in (0, p)$ there exist a real-valued function $\gamma_1(\cdot)$ and a nonnegative function $\gamma_2(\cdot)$ with $\gamma_j(\zeta) \to 0$, $j = 1, 2$, when $\zeta \to 0$ and a constant $c > 0$ such that

$$\begin{aligned} \left|\phi'_{\lambda+\zeta}(x)\right| &\le c\{|\phi'_\lambda(x)|\}^{1+\gamma_1(\zeta)} F(x)^{-\gamma_2(\zeta)}, \\ &\forall x \in (Q(0), Q(\bar{p})), \quad \forall |\zeta| \le \zeta_0. \end{aligned} \tag{42}$$

Condition (39) implies that $\forall\ 1 \le i \le N$ and $\forall\ |\zeta| \le \zeta_0$,

$$\frac{C(\alpha,\lambda)[1 - C\zeta]}{C(\alpha, \lambda+\zeta)} Z_i(\lambda) \le Z_i(\lambda+\zeta) \le \frac{C(\alpha,\lambda)[1 + C\zeta]}{C(\alpha, \lambda+\zeta)} Z_i(\lambda).$$

From this and under the assumption of a continuous differentiable function $t \mapsto C(\alpha, t)$, deduce that $\forall\ 1 \le i \le N$ and $\forall\ |\zeta| \le \zeta_0$, $|1 - E[Z_i(\lambda+\zeta)]C(\alpha,\lambda)n/i| \le C_1\zeta$ for some $C_1 > 0$. Moreover,

$$\begin{aligned} E\left[\sup_{|\zeta|\le\zeta_0} |Z_i(\lambda+\zeta) - E[Z_i(\lambda+\zeta)]|^2\right] &\le C_2\zeta^2\{E[Z_i^2(\lambda)] + E^2[Z_i(\lambda)]\} \\ &\le C_3\zeta^2 i^2 n^{-2}, \end{aligned}$$

for some constants $C_2, C_3 > 0$. Now, we can write

$$\begin{aligned} \frac{\nabla_t Z_i(\lambda+\zeta)}{Z_i(\lambda+\zeta)} - \frac{\nabla_t Z_i(\lambda)}{E[Z_i(\lambda)]} &= \nabla_t Z_i(\lambda+\zeta)\left(\frac{1}{Z_i(\lambda+\zeta)} - \frac{1}{E[Z_i(\lambda+\zeta)]}\right) \\ &\quad + \frac{\nabla_t Z_i(\lambda+\zeta) - \nabla_t Z_i(\lambda)}{E[Z_i(\lambda+\zeta)]} \\ &\quad + \nabla_t Z_i(\lambda)\left(\frac{1}{E[Z_i(\lambda+\zeta)]} - \frac{1}{E[Z_i(\lambda)]}\right) \\ &=: \Delta_1(i,\zeta) + \Delta_2(i,\zeta) + \Delta_3(i,\zeta). \end{aligned}$$

Let $\Delta_j(\zeta) = N^{-1}\sum_{i=1}^N \Delta_j(i,\zeta)$, $j = 1, 2, 3$. The term $\Delta_1(i,\zeta)$ could be bounded using a Cauchy–Schwarz inequality, see also (21), and the uniform bounds on the first- and second-order moments above. From these facts we can deduce that $\sup_{|\zeta|\le\delta n^{-1/2}} \Delta_1(\zeta) = O_P(r_n/\sqrt{n}) = o_P(1)$. For the term $\Delta_3(\zeta)$, take absolute values and, one one hand, use the fact that $|E[Z_i(\lambda+\zeta)] - E[Z_i(\lambda)]| \le C_1 E[Z_i(\lambda)]\zeta$ and, on the other hand, apply the arguments used for proving (18) with $\phi'_\lambda(\cdot)$ replaced by $|\phi'_\lambda(\cdot)|$. Deduce that $\sup_{|\zeta|\le\delta n^{-1/2}} \Delta_3(\zeta) = o_P(1)$. Finally, for any small $\bar{p} > 0$, the sum defining $\Delta_2(\zeta)$ could be split in two parts: $\sum_{1\le i\le N} = \sum_{1\le i\le n\bar{p}} + \sum_{n\bar{p}<i\le N}$. For the second sum, use assumption (41) and apply the arguments used for (18) with $\phi'_\lambda(\cdot)$ replaced by $|\Phi(\cdot)|$. Deduce that the second sum converges in probability to a positive real number that can be made arbitrarily small is ζ is sufficiently

close to zero. For the first sum, take absolute values, use assumption (42), triangle inequality and apply again the arguments used for (18) with $\phi'(\cdot)$ replaced by $|\phi'_\lambda(\cdot)| + c|\phi'_\lambda(x)|^{1+\gamma_1(\zeta)}$. Deduce that the first sum converges in probability to a positive real number that can be made arbitrarily small is $\bar{p}$ is sufficiently close to zero. Conclude that $\sup_{|\zeta|\leq\delta n^{-1/2}} \Delta_2(\zeta) = o_P(1)$. Gather results and deduce that

$$\sup_{|\zeta|\leq\delta n^{-1/2}} |c_n(\lambda+\zeta) - \overline{c}_n(\lambda)| \to 0 \quad \text{in probability,}$$

which implies (19) and thus completes the proof. □

7.2 Proof of Corollary 1

In the Weibull case,[2] the estimator $\hat{\lambda}$ is the solution of the equation $W_n(t) = 0$ with $W_n(t)$ defined in (15). Conditions 1 and 2 of Property 1 are quite obvious. By lengthy but quite elementary calculations, it can be shown that for any α, $\lambda > 0$, condition 3 is satisfied with any $p \in (0, 1)$ except at most two values. Now let us check condition 4 of Property 1. We have, $Q(0) = 0$, $\phi(x) = x^\lambda \log(x)$ so that $\phi(\cdot)$ is strictly decreasing on $(0, \exp(-1/\lambda))$, which guarantees (27). Moreover, $Q(u) = \alpha^{-1} \log^{1/\lambda}(1/(1-u))$. Since $\log(1/(1-u)) \leq u/(1-u)$, deduce that

$$|\phi(Q(u))| \leq C \frac{u}{1-u} \log\left(\frac{u}{1-u}\right), \quad 0 < u \leq p$$

for some constant C, thus condition (28) also holds. Conditions (33) and (35) are clearly fulfilled. Next, $1/f(Q(u)) = cQ^{1-\lambda}(u)(1-u)^{-1}$ for some $c > 0$. Since $\lim_{u\downarrow 0} u^{-1}\log(1/(1-u)) = 1$, deduce that (37) is fulfilled with $w(u) = Cu^{-1+1/\lambda}$ for some $C > 0$. Finally, for the definition of the $q(\cdot)$ function one can take $\eta(u) = u^\gamma$ for some (small) $\gamma > 0$ and thus conditions (36) and (38) are satisfied if

$$\int_0^p f(Q(u))^{a-1} u^{-\gamma-d}\, du < \infty.$$

To check that the last integral is finite it suffices to notice that in the Weibull model with $\lambda \neq 1$ condition (32) holds with $d = 0$ and $a = 1 + b/(1-\lambda)$ and arbitrarily small $b > 0$ (and x_0, $c_0 > 0$ sufficiently large), so that $f(Q(u))^{a-1}u^{-\gamma} \leq Cu^{-\gamma-b/\lambda}$ with $\gamma + b/\lambda < 1$ and some $C > 0$. When $\lambda = 1$, the function $f(Q(\cdot))$ is bounded on $(0, p]$ so that it suffices to take $a = 0$ and $d > 0$ sufficiently small. Clearly, conditions (36) and (38) also hold for $\bar{a}$ and $\bar{d}$ sufficiently close to a and d. Condition (39) obviously holds with $C = 1$. To check condition (40), notice that in the Weibull case, $\forall\, \lambda + \zeta > 0$,

[2] In this section, c and C are constants that may change from line to line.

$$\left| \frac{\nabla_t Z_i(\lambda+\zeta)}{Z_i(\lambda+\zeta)} \right| \le \max\left\{ |\log(X_{(1)})|, |\log(X_{(N+1)})| \right\}.$$

Using the moments of the order statistics from an exponential distribution, it is easy to check that (40) is satisfied for $r_n = C \log n$ with C some positive constant (that depends on λ). For checking condition (41) it suffice to notice that $|\phi'_{\lambda+\zeta}(x) - \phi'_\lambda(x)| \le \phi'_\lambda(x)|1 - x^\zeta[(\lambda+\zeta)\log(x)+1]/[\lambda\log(x)+1]|$ where c is some constant depending on λ, ζ and $\bar{p}$ and x belongs to a compact interval to the right of the origin. Finally, for (42) notice that if $\lambda \neq 1$, $\zeta > 0$ and $\bar{p} < \exp(-1/\lambda)$,

$$|\phi'_{\lambda+\zeta}(x)| = \{x^{\lambda-1}|\lambda\log(x)+1|\}^{1+\gamma_1(\zeta)}\varrho(x) = |\phi'_\lambda(x)|^{1+\gamma_1(\zeta)}\varrho(x)$$

where $\gamma_1(\zeta) = \zeta/(2(\lambda-1))$ and $\varrho(x) = \{x^{\zeta/2}|\lambda\log(x)+1|^{-1-\zeta/(2(\lambda-1))}|(\lambda+\zeta)\log(x)+1|\}$ is a bounded function on $(0, \bar{p})$. Hence (42) holds with $\gamma_2(\zeta) = 0$. The case $\lambda \neq 1$ and $\zeta < 0$ can be handled similarly. When $\lambda = 1$, take $\gamma_1(\zeta) = 1$ and, for instance, $\gamma_2(\zeta) = |\zeta|$. □

7.3 Proof of Lemma 1

$$f'(\lambda) = \frac{\frac{b\mathrm{e}^{-b\lambda}}{1-\mathrm{e}^{-b\lambda}}\log(1-\mathrm{e}^{-a\lambda}) - \frac{a\mathrm{e}^{-a\lambda}}{1-\mathrm{e}^{-a\lambda}}\log(1-\mathrm{e}^{-b\lambda})}{(\log(1-\mathrm{e}^{-a\lambda}))^2}$$

$$= \frac{b(\mathrm{e}^{a\lambda}-1)\log(1-\mathrm{e}^{-a\lambda}) - a(\mathrm{e}^{b\lambda}-1)\log(1-\mathrm{e}^{-b\lambda})}{(\mathrm{e}^{a\lambda}-1)(\mathrm{e}^{b\lambda}-1)(\log(1-\mathrm{e}^{-a\lambda}))^2}.$$

Let

$$g(\lambda) = b(\mathrm{e}^{a\lambda}-1)\log(1-\mathrm{e}^{-a\lambda}) - a(\mathrm{e}^{b\lambda}-1)\log(1-\mathrm{e}^{-b\lambda}),$$

then

$$g'(\lambda) = ab\left[\frac{\log(1-\mathrm{e}^{-b\lambda})}{-\mathrm{e}^{-b\lambda}} - \frac{\log(1-\mathrm{e}^{-a\lambda})}{-\mathrm{e}^{-a\lambda}}\right].$$

Note that $\log(1+x)/x$ is strictly decreasing on $(0, \infty)$ and that $-\mathrm{e}^{-\lambda x}$ is strictly increasing in $x > 0$ for $\lambda > 0$, we have that $\frac{\log(1-\mathrm{e}^{-\lambda x})}{-\mathrm{e}^{-\lambda x}}$ is strictly decreasing in $x > 0$ for $\lambda > 0$. Thus $g'(\lambda) < 0$ on $(0, \infty)$. Hence $g(\lambda)$ is strictly decreasing on $(0, \infty)$. Therefore, for $\lambda > 0$, we have

$$g(\lambda) < \lim_{\lambda\to 0+} g(\lambda) = 0.$$

So we have that $f'(\lambda) < 0$ on $(0, \infty)$, thus $f(\lambda)$ is strictly decreasing on $(0, \infty)$. □

Acknowledgments The authors thank the Editor, the Associate Editor and referees for constructive comments and suggestions that led to substantial improvements in the paper. Keming Yu is partially supported by NSF grant 11261048 of China. Bing Xing Wang is supported by Foundation of Ministry of Education of China under Grant No. 12YJA910005. Valentin Patilea gratefully acknowledges financial support from the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project PN-II-ID-PCE-2011-3-0893.

References

Brown, B. M. (1985). Grouping problems in distribution-free regression. *Australian Journal of Statistics*, *27*, 123–134.

Csörgő, M., Horváth, L. (1990). On the distribution of L_p norms of weighted quantile processes. *Annales de l'I.H.P., section B*, *26*, 65–85.

Csörgő, M., Révész, P. (1978). Strong approximations of the quantile process. *Annals of Statistics*, *6*, 882–894.

Csörgő, M., Yu, H. (1997). Estimation of total time on test transforms for stationary observations. *Stochastic Processes and their Applications*, *68*, 229–253.

Dodson, B. (1994). *Weibull analysis: with software*. Milwaukee: ASQ Quality Press.

Giorgi, G. M. (1999). Income inequality measurement: the statistical approach. In J. Silber (Ed.), *Handbook of income inequality measurement* (pp. 245–267). Boston: Kluwer.

Gupta, R. D., Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*, *41*, 173–188.

Gupta, R. D., Kundu, D. (2001). Generalized exponential distributions: different methods of estimations. *Journal of Statistical Computation and Simulation*, *69*, 315–338.

Gupta, R. D., Kundu, D. (2003). Closeness of gamma and generalized exponential distribution. *Communications in Statistics—Theory and Methods*, *32*, 705–721.

Gupta, R. D., Kundu, D. (2006). On the comparison of Fisher information of the Weibull and GE distributions. *Journal of Statistical Planning and Inference*, *136*, 3130–3144.

Gupta, R. D., Kundu, D. (2007). Generalized exponential distribution: existing results and some recent developments. *Journal of Statistical Planning and Inference*, *137*, 3537–3547.

Hampel, F. R., Rousseeuw, P. J., Ronchetti, E. M., Stahel, W. (1986). *Robust statistics: the approach based on influence functions*. New York: Wiley Interscience.

Hosking, J. R. M. (1990). L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society B*, *52*, 105–124.

Hosking, J. R. M. (1995). The use of L-moments in the analysis of censored data. In N. Balakrishnan (Ed.), *Recent advances in life-testing and reliability* (pp. 545–564). Boca Raton: CRC Press.

Jones, B. L., Zitikis, R. (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal*, *7*, 44–54.

Kundu, D., Gupta, R. D., Manglick, A. (2005). Discriminating between the log-normal and generalized exponential distribution. *Journal of Statistical Planning and Inference*, *127*, 213–227.

Lawless, J. F. (1975). Construction of tolerance bounds for the extreme value and Weibull distributions. *Technometrics*, *17*, 255–261.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data* (2nd ed., pp. 1691–1696). New York: Wiley.

Maritz, J. S. (1995). *Distribution-free statistical methods*. London: Chapman & Hall.

Mitra, S., Kundu, D. (2008). Analysis of the left censored data from the generalized exponential distribution. *Journal of Statistical Computation and Simulation*, *78*, 669–679.

Raqab, M. Z. (2002). Inferences for generalized exponential distribution based on record statistics. *Journal of Statistical Planning and Inference*, *104*, 339–350.

Wang, B. X., Yu, K., Jones, M. C. (2010). Inference under progressively type II right censored sampling for certain lifetime distributions. *Technometrics*, *52*, 453–460.

Weerahandi, S. (2004). *Generalized inference in repeated measures: exact methods in MANOVA and mixed models*. New Jersey: Wiley.