Ann. Inst. Statist. Math.
Vol. 49, No. 3, 447-466 (1997)

EMPIRICAL LIKELIHOOD TYPE CONFIDENCE INTERVALS UNDER RANDOM CENSORSHIP

GIANFRANCO ADIMARI

Dipartimento di Scienze Statistiche, Università di Padova, via S. Francesco 33, 35121 Padova, Italy

(Received January 4, 1996; revised August 6, 1996)

Abstract. In this paper a simple way to obtain empirical likelihood type confidence intervals for the mean under random censorship is suggested. An extension to the more general case where the functional of interest is an M-functional is discussed and the proposed technique is used to construct confidence intervals for quantiles. The results of a simulation study carried out to assess the accuracy of these inferential procedures are also given.

Key words and phrases: Censored data, empirical likelihood, Kaplan-Meier estimator, M-estimator, M-functional, profile likelihood.

1. Introduction

Let $X_1, \dots, X_n$ be a sequence of positive i.i.d. random variables with a common unknown distribution F_0. Let $C_1, \dots, C_n$ be another sequence of positive i.i.d. random variables, independent of the X's and with unknown distribution G. In the random censorship model from the right, one observes only the pairs (T_i, δ_i), $i = 1, \dots, n$, where $T_i = \min\{X_i, C_i\}$ and $\delta_i = I\{X_i \le C_i\}$. The observation T_i is censored when δ_i is equal to 0 and uncensored if δ_i is equal to 1. In survival analysis the variables X's are lifetimes so that an uncensored observation identifies a death time while a censored one represents the random loss of an individual under study. Usually, F_0 is estimated by the Kaplan-Meier (1958) estimator. Let $T_{(1)} \le T_{(2)} \le \cdots \le T_{(n)}$ be the ordered T-values, where ties within censored or uncensored data are ordered arbitrarily and ties among censored and uncensored data are treated as if the former precedes the latter. Let $\delta_{(i)}$ be the concomitant of $T_{(i)}$ and assume that there are n_u distinct uncensored values $T_1^* < T_2^* < \cdots < T_{n_u}^*$ in the sample. The Kaplan-Meier estimator is

$$\bar{F}(t) = \begin{cases} 1 - \prod_{i=1}^{K(t)} \left(1 - \frac{D_i}{N_i}\right) & \text{if} \quad t \le T_{(n)} \\ 1 & \text{if} \quad t > T_{(n)} \text{ and } \delta_{(n)} = 1 \\ \text{undefined} & \text{if} \quad t > T_{(n)} \text{ and } \delta_{(n)} = 0. \end{cases}$$

Using the survival analysis terminology, $K(t)$ denotes the number of (distinct) death times in $(0,t]$, D_i the number of deaths occuring at T_i^* and N_i the number of individuals still at risk at T_i^*, that is the number of individuals neither dead or lost just before time T_i^*. In the following we will often refer to $\bar{F}(t)$ as a distribution function attaching, in particular, positive mass to the uncensored observations. If in the sample there are no censored data, $\bar{F}$ coincides with the usual empirical distribution function, here denoted by $\hat{F}$.

It is well known (see Kaplan and Meier (1958)) that $\bar{F}$ maximizes the nonparametric likelihood

$$\mathcal{L}(F_h) = \prod_{i=1}^{n_u} (1-h_i)^{D_i} h_i^{(N_i - D_i)}$$

defined on the family $\mathcal{F}_h$ of distributions F_h, with support $\{T_1^*, \ldots, T_{n_u}^*\} \cup \mathcal{A}$, $\mathcal{A} \subset [T_{(n)}, +\infty)$, characterized by the vectors $h = (h_1, \ldots, h_{n_u})$ of conditional survival probabilities

$$h_i = h_i(F_h) = pr\{Y > T_i^* \mid Y > T_{i-1}^*; Y \sim F_h\} = \frac{1 - F_h(T_i^*)}{1 - F_h(T_{i-1}^*)},$$

$i = 1, \ldots, n_u$, with $F_h(T_0^*) = T_0^* = 0$. Using this property of $\bar{F}$, Thomas and Grunkemeier (1975) propose a technique to construct confidence intervals for survival probabilities. For a fixed a such that $K(a) \geq 1$, the authors obtain a nonparametric likelihood $\mathcal{L}_\theta(\theta)$ for the functional of interest $\theta(F) = 1 - F(a)$ by profiling $\mathcal{L}(F_h)$. They show heuristically that, when $\theta = \theta_0$, $\theta_0 = \theta(F_0)$, the log likelihood ratio $\ell_\theta(\theta) = -2\log\{\mathcal{L}_\theta(\theta)/\mathcal{L}_\theta(\bar{\theta})\}$, where $\bar{\theta} = \theta(\bar{F}) = 1 - \bar{F}(a)$, has asymptotic χ_1^2 distribution. So, as in the parametric case, a confidence interval for θ_0 with nominal coverage γ is given by the set $\{\theta : \ell_\theta(\theta) \leq c_\gamma\}$, where c_γ is such that $pr\{\chi_1^2 > c_\gamma\} = 1 - \gamma$. A rigorous justification of this method is given by Li (1995). Owen (1988, 1990, 1991) shows that this technique can be successfully applied to a wide range of problems under the usual random sampling model and to linear regression problems. In the uncensored case, that is when the observed data are $X_1, \ldots, X_n$, Owen (1988, 1990) suggests to associate to the sample the nonparametric likelihood

$$L(F_p) = \prod_{i=1}^{n} p_i$$

defined on the family $\mathcal{F}_p$ of the multinomial distributions on the points $X_1, \ldots, X_n$, F_p, characterized by the probability vectors $p = (p_1, \ldots, p_n)$ satisfying the constraints $p_i \geq 0$, $i = 1, \ldots, n$, and $\sum_{i=1}^n p_i = 1$. This likelihood is maximum when $F_p = \hat{F}$. For a given functional of interest $\theta(\cdot)$, the nonparametric likelihood $L_\theta(\theta)$, called empirical likelihood, is deduced from $L(F_p)$ as a profile likelihood. In regular cases, the statistic $l_\theta(\theta_0) = -2\log\{L_\theta(\theta_0)/L_\theta(\hat{\theta})\}$, where $\hat{\theta} = \theta(\hat{F})$, has asymptotic χ^2 distribution. An explicit expression for $l_\theta(\theta)$ is generally obtained using the Lagrange multipliers technique. In particular, when the functional of interest is the mean, i.e. when $\theta(F) = \mu(F) = \int t dF(t)$, the empirical log likelihood ratio function is defined for μ belonging to $(X_{(1)}, X_{(n)})$ and its expression is

$$l_\mu(\mu) = 2\sum_{i=1}^{n} \log\{1 + \lambda(\mu)(X_i - \mu)\}, \tag{1.1}$$

where $\lambda(\mu)$ is the unique solution of the equation

$$\frac{1}{n}\sum_{i=1}^{n} \frac{X_i - \mu}{1 + \lambda(\mu)(X_i - \mu)} = 0.$$

From a practical point of view, outside $(X_{(1)}, X_{(n)})$ it is adequate to set $l_\mu(\mu) = +\infty$. Observe that finding the explicit expression (1.1) constitutes the first important step to establish the asymptotic distribution of the statistic $l_\mu(\mu_0)$, with $\mu_0 = \mu(F_0)$. This way of proceeding is quite common in empirical likelihood frameworks and can be found also in the paper of Li (1995). Moreover, defining the nonparametric likelihood $L(\cdot)$ over a family of distributions which are absolutely continuous with respect to $\hat{F}$ assures that the confidence intervals constructed using l_μ are bounded. Under the random censorship model, an analogous restriction on the domain of $\mathcal{L}(\cdot)$ would be necessary to construct confidence intervals for the mean of F_0. In this case, a natural choice would be to limit the domain of the nonparametric likelihood function to distributions with support on $\{T_1^*, \ldots, T_{n_u}^*\} \cup \{T_{(n)}\}$, that is, treating the largest observation $T_{(n)}$ always as uncensored, to distributions with support on $\{T_1^*, \ldots, T_{n_u}^*\}$. However, even with this device, it might be hard to justify the construction of a confidence interval for μ_0 according to the Thomas and Grunkemeier scheme. In fact, we should maximize $\mathcal{L}(F_h)$ under the constraint

$$T_1^* + \sum_{i=2}^{n_u}(T_i^* - T_{i-1}^*) \prod_{j=1}^{K(T_{i-1}^*)} h_j = \mu,$$

and an explicit form for the log likelihood ratio function seems to us difficult to obtain.

In this paper we suggest a simple way to construct empirical likelihood type confidence intervals for the mean under random censorship. We propose to refer to the function

$$l_\mu^*(\mu) = 2n\sum_{i=1}^{n_u} \tilde{p}_i \log\{1 + \lambda(\mu)(T_i^* - \mu)\}, \tag{1.2}$$

with $\lambda(\mu)$ satisfying

$$\sum_{i=1}^{n_u} \frac{\tilde{p}_i(T_i^* - \mu)}{1 + \lambda(\mu)(T_i^* - \mu)} = 0, \tag{1.3}$$

and where $\tilde{p}_i$ represents the probability mass that the Kaplan-Meier distribution attaches to T_i^*. The Kaplan-Meier distribution is computed on the sample (T_j, δ_j),

$j = 1, \ldots, n$, where the largest observation is always assumed to be uncensored. Thus, except for the proof of Lemma 2.2, we will always assume in the following $T^*_{n_u} = T_{(n)}$. The modified version of the Kaplan-Meier estimator will be denoted by $\tilde{F}(t)$. Observe that the function l^*_μ is the obvious extension of the function l_μ to the censored case and coincides with l_μ when censoring is not present. In Section 2 we justify our choice of (1.2) as a tool to construct confidence intervals for the mean under random censorship. In particular, we show that, under some conditions, the statistic $l^*_\mu(\mu_0)$, suitably corrected, has a χ^2_1 limit distribution. Furthermore, in Section 3, we discuss how this approach can be extended to the more general case where the functional of interest is an M-functional, and in Section 4 this technique is used to obtain confidence intervals for quantiles. Finally, Section 5 gives the results of a simulation study carried out to assess the accuracy of the proposed inferential procedures.

2. Confidence intervals for the mean

Using the function l^*_μ to construct confidence intervals for μ_0 is equivalent to base the inferential procedure on an artificial $L(\cdot)$ type likelihood function defined on the multinomial distributions on the points $T^*_1, \ldots, T^*_{n_u}$ and maximized by $\tilde{F}$. An example can help to explain. Suppose that the sample $T_{(1)} < T_{(2)} < \cdots < T_{(5)}$ is observed with $\delta_{(1)} = \delta_{(2)} = \delta_{(5)} = 1$ and $\delta_{(3)} = \delta_{(4)} = 0$. In this case, the Kaplan-Meier distribution assigns masses equal to 1/5, 1/5 and 3/5, respectively, to the uncensored data $T^*_1 = T_{(1)}$, $T^*_2 = T_{(2)}$ and $T^*_3 = T_{(5)}$. Thus, the nonparametric likelihood function $L(\cdot)$ based on the sample T^*_1, T^*_2, three times T^*_3, thought of as a set of independent observations from F_0, would be maximized by $\tilde{F}$. In general, given the sample data (T_j, δ_j), $j = 1, \ldots, n$, we can always find n_u integers $\nu_1, \ldots, \nu_{n_u}$, such that $\tilde{p}_i = \nu_i/\nu$, with $\nu = \sum_{i=1}^{n_u} \nu_i$, and for which the likelihood $L(\cdot)$ constructed on the artificial sample

$$\nu_1 \text{ times } T^*_1, \nu_2 \text{ times } T^*_2, \ldots, \nu_{n_u} \text{ times } T^*_{n_u},$$

is maximized by $\tilde{F}$. For such a sample, the empirical log likelihood ratio, using (1.1), becomes

$$l^\circ_\mu(\mu) = 2 \sum_{i=1}^{n_u} \nu_i \log\{1 + \lambda(\mu)(T^*_i - \mu)\},$$

with $\lambda(\mu)$ solution of the equation

$$\sum_{i=1}^{n_u} \frac{\nu_i(T^*_i - \mu)}{1 + \lambda(\mu)(T^*_i - \mu)} = 0.$$

It may be easily verified that $l^*_\mu(\mu) = (n/\nu) l^\circ_\mu(\mu)$. This formal link between the function $l^*_\mu(\mu)$ and Owen's empirical likelihood function leads us to estabilish that the former one is well defined and finite over the interval $(T^*_1, T^*_{n_u})$. Of course, in order to justify the use of the function $l^*_\mu(\cdot)$ to construct confidence intervals for the mean of F_0, it is necessary to check its asymptotic behaviour, at least under some suitable conditions on F_0 and the censoring mechanism.

Let $\tau_0 = \sup\{t : F_0(t) < 1\}$, $\tau_1 = \sup\{t : G(t) < 1\}$, and suppose that the following hypotheses hold:

(I) F_0 is a continuous distribution;

(II) $\tau_0 \le \tau_1$;

(A-I) $0 < \sigma_{\tilde{\mu}}^2 = \int_0^{\tau_0}(\int_t^{\tau_0}\{1 - F_0(s)\}ds)^2 \frac{dF_0(t)}{\{1-F_0(t)\}^2\{1-G(t-)\}} < \infty$;

(A-II) $n^{1/2}\int_{T_{(n)}}^{\tau_0}\{1 - F_0(t)\}dt$ converges to zero in probability.

Let σ_X^2 be the variance of F_0, and denote by $\tilde{\mu}$ and $\tilde{\sigma}_X^2$ the estimators for μ_0 and σ_X^2 based on the Kaplan-Meier distribution, that is

$$\tilde{\mu} = \mu(\tilde{F}) = \sum_{i=1}^{n_u} \tilde{p}_i T_i^* \quad \text{and} \quad \tilde{\sigma}_X^2 = \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \tilde{\mu})^2.$$

THEOREM 2.1. *Under conditions* (I)–(II) *and* (A-I)–(A-II), *if* $0 < \sigma_X^2 < \infty$ *then*

$$l_\mu^*(\mu_0) = n\frac{(\tilde{\mu} - \mu_0)^2}{\tilde{\sigma}_X^2} + o_p(1).$$

In order to prove the theorem the following two results are useful.

LEMMA 2.1. *Let* $Z_1, \ldots, Z_n$ *be an i.i.d. sequence of positive random variables from a distribution admitting finite second moment. Then* $\max_{1\le i\le n} Z_i = o(n^{1/2})$ *with probability* 1 *when* $n \to \infty$.

PROOF. See Owen ((1990), pp. 98–99). □

LEMMA 2.2. *Under conditions* (I)–(II), *for each* F_0*-integrable real function* $g(t)$,

$$\sum_{i=1}^{n_u} \tilde{p}_i g(T_i^*) \to \int g(t)dF_0(t)$$

with probability 1 *when* $n \to \infty$.

PROOF. Put $\bar{\bar{F}}(t) = I\{t \le T_{(n)}\}\bar{F}(t) + [1 - I\{t \le T_{(n)}\}]\bar{F}(T_{(n)})$. Furthermore, for each $\varepsilon \ge 0$, let

$$\bar{\bar{S}}(\varepsilon) = \int_0^{\tau_0-\varepsilon} g(t)d\bar{\bar{F}}(t), \qquad \tilde{S}(\varepsilon) = \int_0^{\tau_0-\varepsilon} g(t)d\tilde{F}(t), \qquad S(\varepsilon) = \int_0^{\tau_0-\varepsilon} g(t)dF_0(t).$$

Clearly, the functions $\bar{\bar{F}}(t)$ and $\tilde{F}(t)$ can differ only for $t \ge T_{(n)}$. For each $\varepsilon > 0$ we have

$$\begin{aligned}|\tilde{S}(\varepsilon) - S(0)| &\le |\tilde{S}(\varepsilon) - \bar{\bar{S}}(\varepsilon)| + |\bar{\bar{S}}(\varepsilon) - S(0)| \\ &\le |\tilde{S}(\varepsilon) - \bar{\bar{S}}(\varepsilon)| + |\bar{\bar{S}}(\varepsilon) - S(\varepsilon)| + |S(\varepsilon) - S(0)|.\end{aligned}$$

Under the condition (II) $T_{(n)} \to \tau_0$ with probability 1 when $n \to \infty$. Thus, $|\tilde{S}(\varepsilon) - \bar{\bar{S}}(\varepsilon)| \to 0$ with probability 1. On the other hand, as a consequence of

Theorem 1.1 of Stute and Wang (1993), $|\bar{\bar{S}}(\varepsilon) - S(\varepsilon)| \to 0$ with probability 1 when $n \to \infty$. The result follows from the arbitrariness of ε and the integrability of $g(t)$ with respect to F_0. □

Proof of Theorem 2.1. Under conditions (i)–(ii) and (a-i)–(a-ii), we have, by Gill (1983),

$$n^{1/2}(\tilde{\mu} - \mu_0) \xrightarrow{d} N(0, \sigma^2_{\tilde{\mu}}). \tag{2.1}$$

Then $\tilde{\mu} - \mu_0 = O_p(n^{-1/2})$. On the other hand, under the hypotheses made, T_1^* and $T_{n_u}^*$ converge, with probability 1, to $\inf\{t : F_0(t) > 0\}$ and τ_0, respectively. So, $pr\{T_1^* < \mu_0 < T_{n_u}^*\} \to 1$ when $n \to \infty$. Consequently, $l_\mu^*(\mu_0)$ exists (finite) with probability tending to 1. By Dini's theorem, the function $\lambda(\mu)$ implicitly defined by equation (1.3) is continuous in a neighbourhood of $\tilde{\mu}$ resulting

$$\lambda'(\tilde{\mu}) = \frac{d\lambda}{d\mu}\bigg|_{\mu=\tilde{\mu}} = -\frac{1}{\tilde{\sigma}_X^2}.$$

Since

$$\tilde{\sigma}_X^2 = \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^2 + O_p(n^{-1}), \tag{2.2}$$

and, as a consequence of Lemma 2.2,

$$\sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^2 = \sigma_X^2 + o_p(1), \tag{2.3}$$

we have $\tilde{\sigma}_X^2 = \sigma_X^2 + o_p(1)$. Then it is possible to consider the Taylor series expansion of $\lambda(\mu)$ around $\tilde{\mu}$, $\lambda(\mu) = \lambda(\tilde{\mu}) + \lambda'(\tilde{\mu})(\mu - \tilde{\mu}) + o(|\mu - \tilde{\mu}|)$. Since $\lambda(\tilde{\mu}) = 0$ and putting $\lambda_0 = \lambda(\mu_0)$, at $\mu = \mu_0$ this expression becomes

$$\lambda_0 = \frac{\tilde{\mu} - \mu_0}{\tilde{\sigma}_X^2} + o_p(n^{-1/2}). \tag{2.4}$$

Thus $\lambda_0 = O_p(n^{-1/2})$ and, by Lemma 2.1,

$$|\lambda_0| \max_{1 \le i \le n_u} |T_i^* - \mu_0| \le |\lambda_0| \max_{1 \le j \le n} |T_j - \mu_0| = o_p(1).$$

Then, using the McLaurin series expansion

$$\log(1+z) = z - \frac{1}{2}z^2 + \frac{z^3}{3(1+\ddot{z})^3}, \qquad |\ddot{z}| \le |z|, \tag{2.5}$$

in the expression of $l_\mu^*(\mu_0)$, we obtain, after some algebra,

$$\begin{aligned} l_\mu^*(\mu_0) &= 2n\lambda_0 \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0) - n\lambda_0^2 \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^2 \\ &\quad + 2n\lambda_0^3 \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^3 \zeta_i, \end{aligned} \tag{2.6}$$

where the variables ζ_i are such that $pr\{|\zeta_i| \leq B, i = 1, \dots, n_u\} \to 1$ when $n \to \infty$, with B a suitable positive constant. Using again Lemma 2.1 and (2.3),

$$\left|\sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^3 \zeta_i\right| \leq \max_{1 \leq j \leq n} |T_j - \mu_0| \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^2 |\zeta_i| = o_p(n^{1/2}).$$

So, keeping in mind that $\lambda_0 = O_p(n^{-1/2})$, we have

$$2n\lambda_0^3 \sum_{i=1}^{n_u} \tilde{p}_i (T_i^* - \mu_0)^3 \zeta_i = o_p(1),$$

and, on the basis of relations (2.2) and (2.4), the result follows from (2.6). □

As a consequence of Theorem 2.1 and (2.1), if $\tilde{\sigma}_{\tilde{\mu}}^2$ denotes a consistent estimator for the asymptotic variance of $n^{1/2}(\tilde{\mu} - \mu_0)$, we have that

$$l_\mu^{**}(\mu_0) = \frac{\tilde{\sigma}_X^2}{\tilde{\sigma}_{\tilde{\mu}}^2} l_\mu^*(\mu_0) \xrightarrow{d} \chi_1^2.$$

Thus, under conditions (I)–(II) and (A-I)–(A-II), the set $\mathcal{I}_\mu(\gamma) = \{\mu : l_\mu^{**}(\mu) \leq c_\gamma\}$, constitutes a confidence region for μ_0 with asymptotic coverage γ. The link between $l_\mu^*(\cdot)$ and Owen's empirical likelihood makes us conclude that $\mathcal{I}_\mu(\gamma)$ is an interval. Furthermore, the shape of the confidence intervals constructed in this way is not subject to predetermined symmetry constraints (characterizing, on the contrary, classical methods based on the normal approximation). Such constraints would make them not much adequate in contexts, like the ones we are dealing with, in which the underlying distributions are typically asymmetric. Evidently, in comparison with the use of a procedure based on the likelihood $\mathcal{L}(\cdot)$, the simplification introduced by considering the function $l_\mu^*(\cdot)$ requires to estimate a correction factor constituted by the ratio of two variances. A possible estimator for the asymptotic variance of $n^{1/2}(\tilde{\mu} - \mu_0)$ is

$$\text{(2.7)} \qquad \tilde{\sigma}_{\tilde{\mu}}^2 = n \sum_{i=1}^{n_u - 1} \left(\int_{T_i^*}^{T_{n_u}^*} \{1 - \tilde{F}(t)\} dt \right)^2 \left(\sum_{j=1}^{n} I\{T_j > T_i^*\} \right)^{-2}.$$

Conditions (I)–(II) and (A-I)–(A-II) are substantially necessary to assure the asymptotic normality of $\tilde{\mu}$ and the validity of a strong law under random censorship. As Gill (1983) observes, conditions (A-I)–(A-II) hold, in particular, when F_0 is a distribution with bounded mean residual life function, provided that $(1 - G) \geq \alpha(1 - F_0)^\beta$ close to τ_0 for some constants $\alpha > 0$ and $\beta < 1$, i.e. provided that the censoring mechanism is not too heavy in the tail.

3. Extension to M-functionals

Let $\psi(t,\theta)$ be a real function defined on $\mathbb{R}\times\Theta$, Θ being an open subset of $\mathbb{R}$, Borel measurable on $\mathbb{R}$ for each $\theta\in\Theta$. A functional $\theta(F)$ defined as a solution of the equation in θ

$$\int \psi(t,\theta)dF(t)=0, \tag{3.1}$$

is called M-functional corresponding to ψ. In the uncensored case, when the observed data are $X_1,\ldots,X_n$, the usual estimator $\hat{\theta}=\theta(\hat{F})$ for $\theta_0=\theta(F_0)$ is called M-estimator corresponding to ψ and, under suitable conditions (see Serfling (1980), Chapter 7), is consistent and asymptotically normal. A generalization of the concept of M-estimator to the censored case is given by the statistic $\tilde{\theta}=\theta(\tilde{F})$, i.e. by a solution of the equation $\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta)=0$. In this case, for some particular ψ functions, the asymptotic properties of the corresponding M-estimators are studied by Reid (1981) and Oakes (1986). Furthermore, general results on the strong consistency of $\tilde{\theta}$ can be obtained following a recent paper of Wang (1995) using Lemma 2.2, given in the previous section, instead of his Lemma 1. Regarding the asymptotic distribution of $\tilde{\theta}$, at the best of our knowledge, a general result does not exist. However, if $\theta(\cdot)$ is compactly differentiable, then we can establish the asymptotic normality of $n^{1/2}(\tilde{\theta}-\theta_0)$ using the functional delta-method (see Andersen *et al.* (1993), Section II.8).

For an M-functional $\theta(\cdot)$, Owen's empirical log likelihood ratio is defined (finite) for those values θ such that $\min_{1\le i\le n}\psi(X_i,\theta)<0<\max_{1\le i\le n}\psi(X_i,\theta)$ and its expression is

$$l_\theta(\theta)=2\sum_{i=1}^{n}\log\{1+\lambda(\theta)\psi(X_i,\theta)\},$$

where $\lambda(\theta)$ is the solution of the equation

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\psi(X_i,\theta)}{1+\lambda(\theta)\psi(X_i,\theta)}=0.$$

By analogy to the mean case, extending $l_\theta(\theta)$ to the censored case leads to the function

$$l_\theta^*(\theta)=2n\sum_{i=1}^{n_u}\tilde{p}_i\log\{1+\lambda(\theta)\psi(T_i^*,\theta)\}, \tag{3.2}$$

with $\lambda(\theta)$ satisfying the equation

$$\sum_{i=1}^{n_u}\frac{\tilde{p}_i\psi(T_i^*,\theta)}{1+\lambda(\theta)\psi(T_i^*,\theta)}=0, \tag{3.3}$$

defined for θ such that $\min_{1\le i\le n_u}\psi(T_i^*,\theta)<0<\max_{1\le i\le n_u}\psi(T_i^*,\theta)$. In the following, we study the asymptotic behaviour of l_θ^* assuming the asymptotic normality of $\tilde{\theta}$ and sufficient regularity conditions on the shape of $\psi(\cdot,\cdot)$. In particular, Theorem 3.1 provides an expression for $l_{\theta}^{*}(\theta_0)$ like that given by Theorem 2.1 for $l_{\mu}^{*}(\mu_0)$.

Let θ_0 be the unique solution of (3.1) when $F = F_0$. Let σ_{ψ}^2 denote the variance of $\psi(X_1, \theta_0)$ and suppose that the following hypotheses hold:

(B-I) $\psi(t,\theta)$ admits derivative $\varphi(t,\theta) = (\partial/\partial\theta)\psi(t,\theta)$ at $\theta = \theta_0$;

(B-II) $0 < \sigma_{\psi}^2 < +\infty$ and $0 < E\{\varphi^2(X_1,\theta_0)\} < +\infty$;

(B-III) $E\{\varphi(X_1,\theta_0)\} \neq 0$;

(B-IV) $\varphi(t,\theta)$ is continuous at $\theta = \theta_0$ uniformly with respect to t,
or,
$\psi(t,\theta)$ is twice differentiable at θ_0, and $E\{(\partial^2/\partial\theta^2)\psi(X_1,\theta)\mid_{\theta=\theta_0}\}$ is finite.

THEOREM 3.1. *Under conditions* (I)–(II) *and* (B-I)–(B-IV), *if* $\tilde{\theta} - \theta_0$ *is* $O_p(n^{-1/2})$ (*but not* $o_p(n^{-1/2})$), *then*

$$l_{\theta}^{*}(\theta_0) = n\frac{(\tilde{\theta}-\theta_0)^2}{\tilde{\sigma}_{\psi}^2/\tilde{\xi}^2} + o_p(1),$$

with

$$\tilde{\sigma}_{\psi}^2 = \sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\tilde{\theta}) \quad \textit{and} \quad \tilde{\xi} = \sum_{i=1}^{n_u}\tilde{p}_i\varphi(T_i^*,\tilde{\theta}).$$

PROOF. Let $\lambda_0 = \lambda(\theta_0)$. Since $1 + \lambda_0\psi(T_i^*,\theta_0) \geq 0$, we have

$$|1+\lambda_0\psi(T_i^*,\theta_0)|^{-1} \geq \left\{1 + |\lambda_0| \max_{1\leq j\leq n}|\psi(T_j,\theta_0)|\right\}^{-1}.$$

Furthermore, from (3.3) with $\theta = \theta_0$,

$$\begin{aligned} 0 &= \left|\sum_{i=1}^{n_u}\tilde{p}_i\left\{\frac{\lambda_0\psi^2(T_i^*,\theta_0)}{1+\lambda_0\psi(T_i^*,\theta_0)} - \psi(T_i^*,\theta_0)\right\}\right| \\ &\geq \left|\sum_{i=1}^{n_u}\tilde{p}_i\frac{\lambda_0\psi^2(T_i^*,\theta_0)}{1+\lambda_0\psi(T_i^*,\theta_0)}\right| - \left|\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)\right| \\ &\geq |\lambda_0|\left\{1+|\lambda_0|\max_{1\leq j\leq n}|\psi(T_j,\theta_0)|\right\}^{-1}\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0) - \left|\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)\right|. \end{aligned}$$

Putting $M_0 = \max_{1\leq j\leq n}|\psi(T_j,\theta_0)|$, it follows that

$$|\lambda_0| \leq \frac{|\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)|}{\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0) - M_0|\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)|}. \tag{3.4}$$

By Lemma 2.1 $M_0 = o(n^{1/2})$ and, as a consequence of Lemma 2.2,

$$\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0) = \sigma_{\psi}^2 + o_p(1). \tag{3.5}$$

Moreover, a Taylor series expansion provides the relation

$$0=\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)+\sum_{i=1}^{n_u}\tilde{p}_i\varphi(T_i^*,\theta_0)(\tilde{\theta}-\theta_0)+o_p(n^{-1/2}), \tag{3.6}$$

from which we obtain $\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)=O_p(n^{-1/2})$. Thus, from (3.4), $\lambda_0=O_p(n^{-1/2})$. On the other hand, equation (3.3) can be rewritten as

$$0=\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)-\lambda_0\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)+\lambda_0^2\sum_{i=1}^{n_u}\tilde{p}_i\psi^3(T_i^*,\theta_0)\eta_i, \tag{3.7}$$

where $\eta_i=\{1+\lambda_0\psi(T_i^*,\theta_0)\}^{-1}$, $i=1,\dots,n_u$. As

$$|\lambda_0|\max_{1\le i\le n_u}|\psi(T_i^*,\theta_0)|\le M_0|\lambda_0|=o_p(1),$$

it is easy to verify that $pr\{|\eta_i|\le B_1, i=1,\dots,n_u\}\to 1$ when $n\to\infty$, being B_1 a suitable positive constant. So

$$\left|\lambda_0^2\sum_{i=1}^{n_u}\tilde{p}_i\psi^3(T_i^*,\theta_0)\eta_i\right|\le\lambda_0^2M_0\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)|\eta_i|=o_p(n^{-1/2}),$$

and, from (3.7), using equation (3.6), we have

$$\lambda_0=\frac{(\theta_0-\tilde{\theta})\sum_{i=1}^{n_u}\tilde{p}_i\varphi(T_i^*,\theta_0)}{\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)}+o_p(n^{-1/2}). \tag{3.8}$$

Using the McLaurin series expansion (2.5) in the expression of $l_\theta^*(\theta_0)$ given by (3.2), we obtain

$$\begin{aligned} l_\theta^*(\theta_0)&=2n\lambda_0\sum_{i=1}^{n_u}\tilde{p}_i\psi(T_i^*,\theta_0)-n\lambda_0^2\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)\\ &\quad+2n\lambda_0^3\sum_{i=1}^{n_u}\tilde{p}_i\psi^3(T_i^*,\theta_0)\zeta_i, \end{aligned} \tag{3.9}$$

where the variables ζ_i are such that $pr\{|\zeta_i|\le B_2, i=1,\dots,n_u\}\to 1$ when $n\to\infty$, B_2 being a suitable positive constant. Since $2n\lambda_0^3\sum_{i=1}^{n_u}\tilde{p}_i\psi^3(T_i^*,\theta_0)\zeta_i=o_p(1)$, on the basis also of relations (3.6) and (3.8), equation (3.9) becomes

$$l_\theta^*(\theta_0)=n\frac{(\tilde{\theta}-\theta_0)^2\{\sum_{i=1}^{n_u}\tilde{p}_i\varphi(T_i^*,\theta_0)\}^2}{\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)}+o_p(1). \tag{3.10}$$

Finally, $\sum_{i=1}^{n_u}\tilde{p}_i\psi^2(T_i^*,\theta_0)=\tilde{\sigma}_\psi^2+o_p(1)$, and $\sum_{i=1}^{n_u}\tilde{p}_i\varphi(T_i^*,\theta_0)=\tilde{\xi}+o_p(1)$ from assumption (B-IV). Thus the result follows immediately from (3.10). □

As a direct consequence of Theorem 3.1, we have the following

COROLLARY 3.1. *Assume that $n^{1/2}(\tilde{\theta} - \theta_0)$ is asymptotically normal with* 0 *mean and variance $\sigma^2_{\tilde{\theta}}$. Denote by $\tilde{\sigma}^2_{\tilde{\theta}}$ a consistent estimator for $\sigma^2_{\tilde{\theta}}$. Then, under conditions* (I)–(II) *and* (B-I)–(B-IV),

$$\frac{\tilde{\sigma}^2_{\psi}/\tilde{\xi}^2}{\tilde{\sigma}^2_{\tilde{\theta}}} l^*_{\theta}(\theta_0) \xrightarrow{d} \chi^2_1.$$

Therefore, when the conditions of Corollary 3.1 hold, the function $l^*_{\theta}(\theta)$, suitably corrected, can be used to construct confidence regions for θ_0 with asymptotically exact coverage. Furthermore, on the basis of the link between $l^*_{\theta}(\cdot)$ and Owen's empirical log likelihood ratio, if $\psi(t, \theta)$ is monotone in θ for each t, then the confidence regions obtained in this way are intervals (cfr. Owen (1988), Theorem 2). Observe that, under conditions (I)–(II) and (B-I)–(B-IV), $\tilde{\sigma}^2_{\psi}/\tilde{\xi}^2$ is a consistent estimator for the asymptotic variance of $n^{1/2}(\hat{\theta} - \theta_0)$, that is for the asymptotic variance of the M-estimator in the uncensored case. Thus, as in the particular case of the mean, the correction factor needed to relate $l^*_{\theta}(\theta_0)$ to a known asymptotic distribution, is a ratio between variances. More precisely, it is the ratio between the M-estimator asymptotic variances, in the uncensored and censored case.

In specific cases, the assumptions made so far on the censoring mechanism and the regularity conditions imposed on the function ψ may be not strictly necessary to show that, up to a correction factor that we can estimate consistently, $l^*_{\theta}(\theta_0)$ has asymptotic χ^2_1 distribution. Let us think, for example, to the case where $\psi(t, \theta) = I\{t > a\} - \theta$, that is when the functional of interest is the survival probability $1 - F(a)$ which does not depend on the whole distribution F: in such a situation assumption (II) seems to be excessive. Another case highlighting these aspects, which is strictly connected to the just mentioned one, is discussed in detail in the following section.

4. Confidence intervals for quantiles

For $0 < q < 1$, let $\theta_0 = \inf\{t : F_0(t) \geq q\}$ be the q-th quantile of the distribution F_0. Assume that F_0 is continuous and strictly increasing in a neighbourhood of θ_0. On the basis of the sample (T_i, δ_i), $i = 1, \ldots, n$, θ_0 is usually estimated by the q-th quantile $\tilde{\theta}$ of the Kaplan-Meier distribution,

$$\tilde{\theta} = \inf\{t : \tilde{F}(t) > q\}, \tag{4.1}$$

for $0 < q < F_0(\tau)$, with $\tau = \min\{\tau_0, \tau_1\}$. Under the assumptions made on F_0, for $\theta_0 < \tau$, $\tilde{\theta}$ is a consistent estimator. Moreover, if F_0 has a positive and continuous density f_0 in a neighbourhood of θ_0, then $\tilde{\theta}$ is asymptotically normal (see Shorack and Wellner (1986), p. 657). The estimation of the asymptotic variance of $\tilde{\theta}$ involves the estimation of the density $f_0(\theta_0)$. Thus, in order to construct confidence regions for θ_0, rather than use the asymptotic normality of $n^{1/2}(\tilde{\theta}-\theta_0)$ it is better to follow the idea of Brookmeyer and Crowley (1982) and use the function

$$w(\theta) = n\frac{\{\tilde{F}(\theta)-q\}^2}{\{1-\tilde{F}(\theta)\}^2\tilde{\sigma}^2(\theta)}, \tag{4.2}$$

obtained as a generalization of the sign test for censored data. In expression (4.2) it is

$$\tilde{\sigma}^2(\theta) = \sum_{i=1}^{K(\theta)} \frac{nD_i}{N_i(N_i-D_i)},$$

so that $\{1-\tilde{F}(\theta)\}^2\tilde{\sigma}^2(\theta)$ is the Greenwood's estimator for $\{1-F_0(\theta)\}^2\sigma^2(\theta)$, with $\sigma^2(\theta) = \int_0^\theta \frac{dF_0(t)}{\{1-F_0(t)\}^2\{1-G(t)\}}$, that is for the asymptotic variance of $n^{1/2}\{\tilde{F}(\theta)-F_0(\theta)\}$. Asymptotically, $w(\theta_0)$ has a χ_1^2 distribution.

Alternatively, confidence regions for θ_0 can be obtained following the approach proposed in this paper. In fact, under the assumptions made on F_0, the q-th quantile θ_0 is the unique solution of equation (3.1) when $F=F_0$ and

$$\psi(t,\theta) = \begin{cases} -1 & \text{if } t-\theta\le 0 \\ q/(1-q) & \text{if } t-\theta>0. \end{cases} \tag{4.3}$$

Thus, inference can be based on the extension to the censored case of the empirical log likelihood ratio for M-functionals. Consider the function $l_\theta^*(\theta)$ defined by relations (3.2) and (3.3) with $\psi(\cdot,\cdot)$ given by (4.3). In this case, $l_\theta^*(\cdot)$ is finite on $[T_1^*, T_{n_u}^*)$ and it is easy to show that, when θ belongs to this interval, the solution $\lambda(\theta)$ of equation (3.3) has expression $\lambda(\theta)=\{q-\tilde{F}(\theta)\}/q$. Consequently,

$$l_\theta^*(\theta) = 2n\left\{\tilde{F}(\theta)\log\frac{\tilde{F}(\theta)}{q} + \{1-\tilde{F}(\theta)\}\log\frac{1-\tilde{F}(\theta)}{1-q}\right\}. \tag{4.4}$$

THEOREM 4.1. *Let F_0 be continuous and strictly increasing in a neighbourhood of θ_0. When $\psi(t,\theta)$ is given by (4.3), if $\theta_0<\tau$ then $\kappa l_\theta^*(\theta_0)\xrightarrow{d}\chi_1^2$, where $\kappa = q/\{(1-q)\sigma^2(\theta_0)\}$.*

PROOF. Since $\tilde{F}(\theta_0) = q + O_p(n^{-1/2})$, by the expansion $\log(z) = z-1-(z-1)^2/2 + o((z-1)^2)$, from (4.4) with $\theta=\theta_0$ we obtain, after some algebra,

$$l_\theta^*(\theta_0) = n\{\tilde{F}(\theta_0)-q\}^2\left\{\frac{2q-3q^2-\tilde{F}(\theta_0)+2q\tilde{F}(\theta_0)}{q^2(1-q)^2}\right\} + o_p(1),$$

which gives, using again the consistency of $\tilde{F}(\theta_0)$,

$$l_\theta^*(\theta_0) = n\frac{\{\tilde{F}(\theta_0)-q\}^2}{q(1-q)} + o_p(1).$$

The result follows immediately from the asymptotic normality of $n^{1/2}\{\tilde{F}(\theta_0)-q\}$. □

Theorem 4.1 establishes that, up to the factor κ, $l^*_\theta(\theta_0)$ has an asymptotic χ^2_1 distribution. Of course, from a practical point of view, there are several asymptotically equivalent ways to correct the function $l^*_\theta(\cdot)$. In particular, here we propose to estimate κ by using

$$\tilde{\kappa} = \tilde{F}(\tilde{\theta})/\{(1-q)\tilde{\sigma}^2_*(\tilde{\theta})\}, \tag{4.5}$$

where $\tilde{\theta}$ is the q-th sample quantile, given by (4.1), and

$$\tilde{\sigma}^2_*(\theta) = \sum_{i=1}^{K(\theta)} \frac{nD_i}{N_i^2}$$

is an estimator for $\sigma^2(\theta)$ alternative to $\tilde{\sigma}^2(\theta)$ (see Andersen *et al.* (1993), Section IV.3).

COROLLARY 4.1. *Under the conditions of the Theorem* 4.1, $l^{**}_\theta(\theta_0) = \tilde{\kappa} l^*_\theta(\theta_0) \xrightarrow{d} \chi^2_1$.

PROOF. Using the consistency of $\tilde{\theta}$, the continuity of $F_0(\cdot)$ at θ_0 and the uniform consistency of $\tilde{F}(\theta)$ in a neighbourhood of θ_0, it is easy to show that $\tilde{F}(\tilde{\theta}) = q + o_p(1)$. On the other hand, $|\tilde{\sigma}^2_*(\theta) - \sigma^2(\theta)| \xrightarrow{p} 0$ uniformly in a neighbourhood of θ_0 (see Andersen *et al.* (1993), Section IV.3) and $\sigma^2(\cdot)$ is a continuous function at θ_0. Therefore, in a similar way we can show that $\tilde{\sigma}^2_*(\tilde{\theta}) = \sigma^2(\theta_0) + o_p(1)$. □

In expression (4.5), $\tilde{F}(\tilde{\theta})$ can be substituted by the asymptotic value q, but, in this case, simulation results show a worsening of the χ^2 approximation for the distribution of $l^{**}_\theta(\theta_0)$. Still in expression (4.5), the choice of the estimator $\tilde{\sigma}^2_*(\theta)$ instead of $\tilde{\sigma}^2(\theta)$ is justified by the fact that, unlike $\tilde{\sigma}^2(\tilde{\theta})$, $\tilde{\sigma}^2_*(\tilde{\theta})$ is different from zero also for samples where, due to heavy censoring, the sample quantile $\tilde{\theta}$ coincides with the largest observed death time. Under the conditions of Theorem 4.1, the set $\mathcal{I}_\theta(\gamma) = \{\theta : l^{**}_\theta(\theta) \leq c_\gamma\}$ is a confidence region for θ_0 with nominal coverage γ. Actually, $\mathcal{I}_\theta(\gamma)$ is an interval. More precisely, since $l^*_\theta(\theta)$ is a step function with jumps at the uncensored data T_i^*, $\mathcal{I}_\theta(\gamma)$ is an interval of the kind $[t_1, t_2)$, where t_1 and t_2 are two observed death times.

We conclude this section with an example based on the data on the treatment by radiation therapy of head and neck cancer (Efron (1988)). Figure 1 shows the functions l^{**}_θ and l^{**}_μ for the median and the mean for these data. The horizontal lines shown are at the asymptotically justified 90%, 95% and 99% levels.

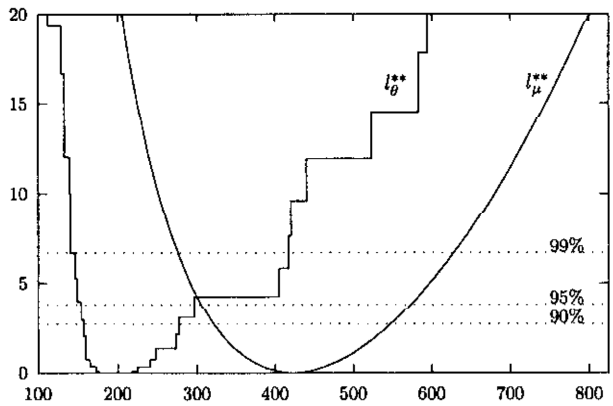

Fig. 1. Functions l_θ^{**} and l_μ^{**} for the median and the mean for the data from the Head-and-Neck-Cancer Study.

5. A simulation study

Sections 2 and 4 provide asymptotic results that justify the use of functions l_μ^* and l_θ^*, given by expressions (1.2) and (4.4), to obtain confidence intervals, respectively, for mean and quantiles under random censorship. To assess the accuracy of the confidence intervals constructed according to this approach (in the following called ELT approach) some simulation experiments have been carried out. This allows us also to compare these inferential procedures with alternative procedures such that based on the normal approximation (in the following called $\mathcal{AN}$ approach), in the mean case, and that proposed by Brookmeyer and Crowley (in the following called $\mathcal{BC}$ approach), in the quantile case.

For the mean case, which is, from a pratical point of view, the less relevant in this context, a unique experiment was carried out assuming that the distribution of interest and the censoring distribution were exponential, the former with mean 1 and the latter with mean b. For b we chose four different values in a way such that the censoring probability $\pi = pr\{C_1 < X_1\}$ was equal to 0.1, 0.15, 0.2 and 0.25, respectively. From each model we generated 5000 samples of size n, with $n = 40$, 60 and 100. On each sample the statistics $l_\mu^{**}(\mu_0)$ and $(n/\tilde{\sigma}_{\tilde{\mu}}^2)^{1/2}(\tilde{\mu} - \mu_0)$ were calculated, with μ_0 the true value of the mean and $\tilde{\sigma}_{\tilde{\mu}}^2$ given by (2.7). Then, on the basis of these 5000 values of the statistics, the actual coverages of the confidence intervals constructed according to ELT and $\mathcal{AN}$ approaches were estimated.

Regarding the quantiles, we considered the median case. We performed four experiments each related to a different choice of the distribution of interest: exponential(1), Weibull(1, 0.5), Weibull(1, 1.5), log-normal(0, 1). The four distributions are characterized by a hazard function respectively constant, decreasing, increasing and non-monotone. In all the experiments we assumed that the censoring distribution was uniform over $[0, \mathcal{T}]$, with $\mathcal{T}$ chosen in a way to obtain censoring rate π equal to 0.2, 0.4 and 0.5. In each experiment, 5000 samples of size n were generated from each model, choosing for n the values 20 and 40 when $\pi = 0.2$ and 20, 40 and 60 in the other two cases. On each sample we computed the statistics

Table 1. Simulation results in the mean case.

			γ					
			0.990	0.975	0.950	0.900	0.700	0.500
$\pi = 0.1$	$n = 40$	$\mathcal{AN}$	0.967	0.946	0.919	0.871	0.680	0.488
		ELT	0.980	0.962	0.932	0.884	0.684	0.485
		ELT_m	0.988	0.972	0.951	0.903	0.713	0.511
	$n = 60$	$\mathcal{AN}$	0.975	0.956	0.929	0.887	0.698	0.496
		ELT	0.983	0.969	0.942	0.891	0.701	0.495
		ELT_m	0.987	0.973	0.949	0.900	0.710	0.504
	$n = 100$	$\mathcal{AN}$	0.980	0.965	0.938	0.890	0.701	0.507
		ELT	0.988	0.970	0.945	0.893	0.700	0.507
		ELT_m	0.990	0.972	0.948	0.897	0.704	0.509
$\pi = 0.15$	$n = 40$	$\mathcal{AN}$	0.958	0.937	0.910	0.861	0.669	0.474
		ELT	0.978	0.954	0.924	0.873	0.669	0.474
		ELT_m	0.989	0.974	0.949	0.903	0.712	0.511
	$n = 60$	$\mathcal{AN}$	0.967	0.945	0.918	0.871	0.682	0.483
		ELT	0.977	0.961	0.931	0.881	0.682	0.485
		ELT_m	0.985	0.971	0.950	0.907	0.715	0.512
	$n = 100$	$\mathcal{AN}$	0.972	0.954	0.931	0.882	0.689	0.502
		ELT	0.984	0.962	0.936	0.891	0.689	0.501
		ELT_m	0.989	0.971	0.946	0.903	0.708	0.518
$\pi = 0.2$	$n = 40$	$\mathcal{AN}$	0.954	0.933	0.906	0.858	0.674	0.478
		ELT	0.967	0.948	0.917	0.867	0.671	0.475
		ELT_m	0.985	0.970	0.946	0.906	0.727	0.521
	$n = 60$	$\mathcal{AN}$	0.965	0.942	0.919	0.875	0.682	0.494
		ELT	0.978	0.960	0.927	0.878	0.683	0.492
		ELT_m	0.985	0.971	0.947	0.902	0.718	0.517
	$n = 100$	$\mathcal{AN}$	0.974	0.956	0.930	0.879	0.688	0.494
		ELT	0.983	0.968	0.940	0.884	0.686	0.492
		ELT_m	0.988	0.974	0.951	0.898	0.704	0.512
$\pi = 0.25$	$n = 40$	$\mathcal{AN}$	0.946	0.924	0.894	0.848	0.665	0.476
		ELT	0.960	0.939	0.912	0.861	0.663	0.476
		ELT_m	0.982	0.967	0.945	0.911	0.738	0.541
	$n = 60$	$\mathcal{AN}$	0.958	0.939	0.911	0.860	0.669	0.477
		ELT	0.970	0.951	0.925	0.871	0.669	0.471
		ELT_m	0.983	0.970	0.949	0.910	0.720	0.511
	$n = 100$	$\mathcal{AN}$	0.967	0.947	0.921	0.872	0.682	0.485
		ELT	0.979	0.957	0.929	0.876	0.682	0.488
		ELT_m	0.988	0.971	0.946	0.898	0.713	0.513

Table 2. Simulation results in the median case. Distribution of interest exponential(1).

			γ						excluded
			0.990	0.975	0.950	0.900	0.700	0.500	samples
$\pi = 0.2$	$n = 20$	$\mathcal{BC}$	0.975	0.957	0.936	0.877	0.694	0.497	0
		ELT	0.989	0.975	0.952	0.899	0.703	0.504	
		ELT_m	0.990	0.975	0.948	0.898	0.699	0.500	
	$n = 40$	$\mathcal{BC}$	0.985	0.967	0.935	0.886	0.688	0.495	0
		ELT	0.991	0.976	0.950	0.895	0.691	0.497	
		ELT_m	0.992	0.975	0.951	0.898	0.696	0.501	
$\pi = 0.4$	$n = 20$	$\mathcal{BC}$	0.972	0.951	0.926	0.873	0.682	0.491	0
		ELT	0.989	0.977	0.950	0.895	0.702	0.506	
		ELT_m	0.989	0.974	0.951	0.901	0.701	0.503	
	$n = 40$	$\mathcal{BC}$	0.981	0.964	0.936	0.885	0.698	0.497	0
		ELT	0.991	0.977	0.954	0.905	0.713	0.507	
		ELT_m	0.989	0.973	0.947	0.898	0.708	0.504	
	$n = 60$	$\mathcal{BC}$	0.986	0.968	0.943	0.893	0.696	0.501	0
		ELT	0.992	0.979	0.955	0.907	0.700	0.505	
		ELT_m	0.991	0.975	0.951	0.903	0.699	0.505	
$\pi = 0.5$	$n = 20$	$\mathcal{BC}$	0.969	0.951	0.920	0.873	0.685	0.480	6
		ELT	0.990	0.980	0.962	0.926	0.762	0.550	
		ELT_m	0.991	0.974	0.949	0.900	0.709	0.496	
	$n = 40$	$\mathcal{BC}$	0.978	0.956	0.928	0.878	0.674	0.484	0
		ELT	0.990	0.976	0.956	0.915	0.720	0.514	
		ELT_m	0.991	0.974	0.949	0.902	0.701	0.499	
	$n = 60$	$\mathcal{BC}$	0.983	0.962	0.937	0.892	0.695	0.491	0
		ELT	0.990	0.973	0.953	0.916	0.723	0.509	
		ELT_m	0.988	0.971	0.948	0.905	0.709	0.497	

$l_\theta^{**}(\theta_0)$ and $w(\theta_0)$, with θ_0 the true value of the median, and on the basis of 5000 values of these statistics we estimated the real coverages of the confidence intervals constructed following ELT and $\mathcal{BC}$ approaches.

Tables 1–5 summarize the simulation results. They provide, for some values of the nominal coverage γ, the empirical coverages of the confidence intervals constructed using the three considered approaches. Furthermore, in each table the rows marked by ELT_m report the estimates of the coverages achieved by correcting the statistics $l_\mu^*(\mu_0)$ and $l_\theta^*(\theta_0)$ through their Monte Carlo mean. In other words, the estimates, in this case, refer to the coverages of the confidence intervals based on the functions $m_\mu^{-1} l_\mu^*(\cdot)$ and $m_\theta^{-1} l_\theta^*(\cdot)$ where m_μ and m_θ are the Monte Carlo means of $l_\mu^*(\mu_0)$ and $l_\theta^*(\theta_0)$ respectively. If the confidence intervals based on the

Table 3. Simulation results in the median case. Distribution of interest Weibull(1,1.5)

			γ						excluded
			0.990	0.975	0.950	0.900	0.700	0.500	samples
$\pi = 0.2$	$n = 20$	$\mathcal{BC}$	0.972	0.947	0.922	0.864	0.682	0.487	0
		ELT	0.989	0.972	0.943	0.889	0.690	0.488	
		ELT_m	0.991	0.974	0.946	0.892	0.697	0.512	
	$n = 40$	$\mathcal{BC}$	0.986	0.967	0.940	0.889	0.688	0.498	0
		ELT	0.992	0.977	0.950	0.900	0.694	0.498	
		ELT_m	0.992	0.978	0.949	0.900	0.694	0.500	
$\pi = 0.4$	$n = 20$	$\mathcal{BC}$	0.965	0.943	0.915	0.856	0.671	0.489	3
		ELT	0.989	0.977	0.950	0.895	0.702	0.506	
		ELT_m	0.989	0.977	0.950	0.899	0.700	0.510	
	$n = 40$	$\mathcal{BC}$	0.981	0.962	0.935	0.887	0.689	0.494	0
		ELT	0.990	0.978	0.952	0.900	0.698	0.501	
		ELT_m	0.990	0.974	0.950	0.901	0.699	0.504	
	$n = 60$	$\mathcal{BC}$	0.983	0.964	0.933	0.883	0.692	0.504	0
		ELT	0.991	0.972	0.948	0.898	0.699	0.508	
		ELT_m	0.990	0.972	0.945	0.897	0.705	0.513	
$\pi = 0.5$	$n = 20$	$\mathcal{BC}$	0.964	0.944	0.915	0.865	0.670	0.482	11
		ELT	0.991	0.976	0.955	0.917	0.728	0.519	
		ELT_m	0.990	0.973	0.949	0.899	0.699	0.502	
	$n = 40$	$\mathcal{BC}$	0.981	0.962	0.933	0.885	0.684	0.483	0
		ELT	0.991	0.977	0.956	0.911	0.713	0.503	
		ELT_m	0.989	0.975	0.950	0.902	0.699	0.494	
	$n = 60$	$\mathcal{BC}$	0.981	0.963	0.932	0.878	0.681	0.488	0
		ELT	0.991	0.976	0.952	0.897	0.698	0.496	
		ELT_m	0.989	0.976	0.949	0.898	0.703	0.506	

functions $l^*_\mu(\cdot)$ and $l^*_\theta(\cdot)$ admitted Bartlett correction, the values provided in the ELT_m rows would reflect this feature and would turn out to be an useful term of comparison. Of course, the Monte Carlo means of $l^*_\mu(\mu_0)$ and $l^*_\theta(\theta_0)$ were computed provided that these statistics assumed finite value. The samples for which this did not occur, all related to the median case with $n = 20$ and high censoring rate, were excluded from the study. Their number, indicated in the tables, is not significant for the evaluation of results. The estimates of the empirical coverage standard errors can be computed through the binomial formula. Their value decreases with the empirical coverage value, varying approximately between 0.0042 and 0.0014 on the interval $[0.90, 0.99]$. In the mean case, the simulation results show that, as one should expect, the ELT confidence intervals are sufficiently accurate for

Table 4. Simulation results in the median case. Distribution of interest Weibull(1,0.5).

			γ						excluded
			0.990	0.975	0.950	0.900	0.700	0.500	samples
$\pi = 0.2$	$n = 20$	$\mathcal{BC}$	0.969	0.956	0.947	0.878	0.711	0.492	0
		ELT	0.991	0.970	0.950	0.894	0.712	0.492	
		ELT_m	0.989	0.984	0.951	0.886	0.711	0.491	
	$n = 40$	$\mathcal{BC}$	0.982	0.963	0.937	0.895	0.700	0.509	0
		ELT	0.990	0.973	0.946	0.899	0.698	0.506	
		ELT_m	0.991	0.977	0.947	0.899	0.705	0.512	
$\pi = 0.4$	$n = 20$	$\mathcal{BC}$	0.968	0.950	0.928	0.867	0.673	0.487	0
		ELT	0.992	0.979	0.962	0.921	0.729	0.539	
		ELT_m	0.992	0.973	0.947	0.899	0.691	0.503	
	$n = 40$	$\mathcal{BC}$	0.983	0.968	0.942	0.898	0.694	0.503	0
		ELT	0.993	0.981	0.962	0.922	0.724	0.520	
		ELT_m	0.989	0.974	0.951	0.905	0.698	0.500	
	$n = 60$	$\mathcal{BC}$	0.985	0.970	0.940	0.889	0.691	0.490	0
		ELT	0.991	0.980	0.952	0.905	0.706	0.493	
		ELT_m	0.990	0.975	0.949	0.897	0.700	0.496	
$\pi = 0.5$	$n = 20$	$\mathcal{BC}$	0.974	0.952	0.928	0.869	0.685	0.503	3
		ELT	0.990	0.979	0.964	0.931	0.800	0.631	
		ELT_m	0.992	0.974	0.947	0.903	0.701	0.509	
	$n = 40$	$\mathcal{BC}$	0.983	0.962	0.936	0.891	0.690	0.492	0
		ELT	0.990	0.980	0.966	0.940	0.775	0.592	
		ELT_m	0.990	0.974	0.947	0.905	0.700	0.496	
	$n = 60$	$\mathcal{BC}$	0.983	0.965	0.934	0.882	0.701	0.508	0
		ELT	0.991	0.978	0.957	0.919	0.761	0.568	
		ELT_m	0.988	0.973	0.942	0.893	0.712	0.516	

medium-large sample sizes provided that the censoring rate is small. However, compared to the $\mathcal{AN}$ approach, the ELT one yields intervals with coverage closer to the nominal one. Also in the median case the ELT approach yields confidence intervals generally more accurate than those obtained by the reference approach $\mathcal{BC}$. In this case, the accuracy is good also for small sample sizes and high censoring rates. To conclude, observe that, in the mean case, the discrepancy between ELT and ELT_m empirical coverages seems to suggest that a more efficient estimator for the correction factor for l^*_μ should improve the ELT confidence intervals accuracy. With regard to this, it could be interesting to assess the accuracy of the confidence intervals constructed by using the function $m_{b\mu}^{-1} l^*_\mu(\cdot)$ where $m_{b\mu}$ denotes the bootstrap mean of $l^*_\mu(\mu_0)$.

Table 5. Simulation results in the median case. Distribution of interest log-normal(0,1).

			γ						excluded
			0.990	0.975	0.950	0.900	0.700	0.500	samples
$\pi = 0.2$	$n = 20$	$\mathcal{BC}$	0.974	0.954	0.931	0.874	0.681	0.481	0
		ELT	0.990	0.974	0.948	0.897	0.686	0.484	
		ELT_m	0.991	0.975	0.950	0.907	0.692	0.506	
	$n = 40$	$\mathcal{BC}$	0.987	0.969	0.943	0.897	0.707	0.510	0
		ELT	0.992	0.978	0.954	0.905	0.711	0.511	
		ELT_m	0.991	0.974	0.950	0.899	0.703	0.503	
$\pi = 0.4$	$n = 20$	$\mathcal{BC}$	0.970	0.949	0.924	0.873	0.687	0.495	2
		ELT	0.992	0.978	0.957	0.916	0.719	0.515	
		ELT_m	0.988	0.974	0.946	0.901	0.703	0.508	
	$n = 40$	$\mathcal{BC}$	0.983	0.964	0.935	0.886	0.689	0.490	0
		ELT	0.993	0.978	0.951	0.904	0.702	0.494	
		ELT_m	0.991	0.975	0.949	0.903	0.700	0.499	
	$n = 60$	$\mathcal{BC}$	0.984	0.967	0.942	0.894	0.694	0.493	0
		ELT	0.988	0.978	0.951	0.904	0.702	0.496	
		ELT_m	0.989	0.975	0.949	0.904	0.703	0.500	
$\pi = 0.5$	$n = 20$	$\mathcal{BC}$	0.969	0.948	0.920	0.867	0.675	0.488	4
		ELT	0.989	0.976	0.958	0.921	0.741	0.538	
		ELT_m	0.988	0.976	0.951	0.901	0.704	0.505	
	$n = 40$	$\mathcal{BC}$	0.981	0.958	0.930	0.879	0.682	0.485	0
		ELT	0.990	0.978	0.957	0.915	0.718	0.510	
		ELT_m	0.991	0.975	0.948	0.900	0.701	0.500	
	$n = 60$	$\mathcal{BC}$	0.985	0.968	0.944	0.893	0.693	0.487	0
		ELT	0.993	0.979	0.959	0.918	0.708	0.499	
		ELT_m	0.990	0.974	0.951	0.903	0.697	0.490	

Acknowledgements

The author is grateful to a referee in particular for suggesting that the results of Section 4 hold under weaker conditions than those stated in an earlier version. The author also thank Guido Masarotto, Francesca Parpinel and Alessandra Salvan for stimulating discussions about this work.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*, Springer-Verlag, Berlin.

Brookmeyer, R. and Crowley, J. (1982). A confidence interval for the median survival time, *Biometrics*, **38**, 29–41.

Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve, *J. Amer. Statist. Assoc.*, **83**, 414–425.

Gill, R. D. (1983). Large sample behavior of the product-limit estimator on the whole line, *Ann. Statist.*, **11**, 49–58.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations, *J. Amer. Statist. Assoc.*, **53**, 457–481.

Li, G. (1995). On nonparametric likelihood ratio estimation of survival probabilities for censored data, *Statist. Probab. Lett.*, **25**, 95–104.

Oakes, D. (1986). An approximate likelihood procedure for censored data, *Biometrics*, **42**, 177–182.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional, *Biometrika*, **75**, 237–249.

Owen, A. B. (1990). Empirical likelihood ratio confidence regions, *Ann. Statist.*, **18**, 90–120.

Owen, A. B. (1991). Empirical likelihood for linear models, *Ann. Statist.*, **19**, 1725–1747.

Reid, N. (1981). Influence functions for censored data, *Ann. Statist.*, **9**, 78–92.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*, Wiley, New York.

Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.

Stute, W. and Wang, J.-L. (1993). The strong law under random censorship, *Ann. Statist.*, **21**, 1591–1607.

Thomas, D. R. and Grunkemeier, G. L. (1975). Confidence interval estimation of survival probabilities for censored data, *J. Amer. Statist. Assoc.*, **70**, 865–871.

Wang, J.-L. (1995). M-estimators for censored data: strong consistency, *Scand. J. Statist.*, **22**, 197–205.