On exact confidence intervals in a competing risks model with generalized hybrid type-I censored exponential data

George Iliopoulos[1]

Abstract

In a recent paper by Mao, Shi and Sun to appear in Journal of Statistical Computation and Simulation, the authors discuss, among other approaches, the construction of exact confidence intervals for the underlying parameters by "pivoting the CDFs" of the corresponding maximum likelihood estimators. The authors assume that this method is applicable without providing the appropriate justification. In this short note the two requirements for the applicability of this method are discussed, namely, the stochastic monotonicity of the maximum likelihood estimators and the existence of solutions to the equations defining the exact confidence interval's endpoints.

Keywords: competing risks, exponential distribution, generalized type-I hybrid censoring, maximum likelihood estimators, stochastic monotonicity, pivoting the CDF

1 Introduction

Mao et al. (2013) analyze a competing risks model under which n items are exposed to two independent risk factors. The lifetimes of the n items are considered independent and under risk factor j they follow an exponential distribution with mean θ_j, denoted by $\mathcal{E}(\theta_j)$, $j = 1, 2$. If X_{ij} denotes the lifetime of item i under factor j, then this item fails at time $X_i = \min\{X_{i1}, X_{i2}\}$. By the independence of the factors and the properties of the exponential distribution, $X_i \sim \mathcal{E}(\theta)$ where $\theta = (\theta_1^{-1} + \theta_2^{-1})^{-1}$. Hence, the failures to be observed are independent and identically distributed random variables $X_1, \dots, X_n$ from $\mathcal{E}(\theta)$. Let $X_{1:n} < \cdots < X_{n:n}$ denote the corresponding ordered lifetimes. Mao et al. (2013) consider the case where the data are obtained under a generalized hybrid type-I censoring sampling scheme. This scheme works as follows. A timepoint $T > 0$ is fixed in advance as well as two integers $1 \leqslant k < r < n$ and the experiment runs until the random time $T_D = \min\{\max\{X_{k:n}, T\}, X_{r:n}\}$. The idea behind this scheme is to obtain at least k and at most r observations while at

[1]Department of Statistics and Insurance Science, School of Finance and Statistics, University of Piraeus; 80 Karaoli & Dimitriou str., 18534 Piraeus, Greece; e-mail: geh@unipi.gr

the same time to keep its duration under control. So, it is initially intended to run until time T unless r items have failed before in which case it stops as soon as the rth failure is observed. If, however, less than r failures have been observed by time T, then the experiment continues beyond this timepoint until the kth failure is observed where it is terminated. This scheme has been originally proposed by Chandrasekar et al. (2004). Let $D = \sum_{i=1}^{n} I(X_i \leqslant T)$ be the number of items that fail by time T. Then, the number of failures observed in the experiment equals $s_D = \min\{\max\{k, D\}, r\}$. Mao et al. (2013) show that the maximum likelihood estimators (MLEs) of θ_1, θ_2 are given by

$$\hat{\theta}_1 = \frac{1}{N_1}\left\{\sum_{i=1}^{s_D} X_{i:n} + (n - s_D)T_D\right\} \quad \text{and} \quad \hat{\theta}_2 = \frac{1}{N_2}\left\{\sum_{i=1}^{s_D} X_{i:n} + (n - s_D)T_D\right\}, \tag{1}$$

respectively, where N_j is the (random) number of failures associated with risk factor j, $j = 1, 2$, and $N_1 + N_2 = s_D$. In order the above estimators to be defined it must hold $N_1 \geqslant 1$ and $N_2 \geqslant 1$. Mao et al. (2013) derive the exact distributions of the MLEs, conditional on the event $N_1 \geqslant 1$, $N_2 \geqslant 1$. More specifically, the cumulative distribution function (CDF) of $\hat{\theta}_1$ is

$$F(y; \theta_1, \theta_2) = \sum_{d=0}^{n} \sum_{i=1}^{s_d-1} \sum_{j=0}^{d} \frac{(-1)^j \binom{n}{d}\binom{d}{j}\binom{s_d}{i} \theta_1^{s_d-i}\theta_2^{i}}{(\theta_1+\theta_2)^{s_d} - \theta_1^{s_d} - \theta_2^{s_d}} e^{-(n-d+j)(1/\theta_1+1/\theta_2)T} \times$$
$$G\left(\left(\frac{1}{\theta_1} + \frac{1}{\theta_2}\right)\{iy - (n-d+j)T\}\,; s_d\right), \quad y > 0, \tag{2}$$

where $G(y; a) = \Gamma(a)^{-1} \int_0^y u^{a-1} e^{-u} \mathrm{d}u$, $y > 0$, is the CDF of the gamma distribution with shape $a > 0$ and scale 1, $\Gamma(\cdot)$ is the gamma function and $s_d = \min\{\max\{k, d\}, r\}$. The expression in (2) looks slightly different than that given by Mao et al. but it can be verified that they are equivalent. On the other hand, the CDF of $\hat{\theta}_2$ is $F(y; \theta_2, \theta_1)$, i.e., (2) but with the two parameters switched.

Mao et al. (2013) discuss, among other methods, the construction of "exact" confidence intervals for θ_1 and θ_2. These intervals are based on pivoting (i.e., inverting with respect to the parameter of interest) the CDF (cf. Barlow et al., 1968, or Casella and Berger, 2002). In particular, the endpoints of the $100(1-\alpha)\%$ "exact" equal-tail confidence interval for θ_1 are the solutions to the equations

$$F(\hat{\theta}_1^{\text{obs}}; \theta_1, \hat{\theta}_2^{\text{obs}}) = 1 - \alpha/2 \quad \text{and} \quad F(\hat{\theta}_1^{\text{obs}}; \theta_1, \hat{\theta}_2^{\text{obs}}) = \alpha/2, \tag{3}$$

with respect to θ_1, where $\hat{\theta}_j^{\text{obs}}$ is the observed value of $\hat{\theta}_j$, $j = 1, 2$. However, in order this construction to be meaningful, the equations in (3) *must have a solution and this solution must be unique.* Usually, people consider the existence of the solutions in such problems for granted and make a suitable assumption which guarantees its uniqueness. More specifically, they assume that the MLE is stochastically monotone in the parameter which in the current context means that the CDF $F(y;\theta_1,\theta_2)$ is decreasing in θ_1 for each fixed y and θ_2 or, equivalently, the tail probability $1-F(y;\theta_1,\theta_2)$ is increasing in θ_1 for each fixed y and θ_2. This is the assumption originally made by Mao et al. (2013) and other researchers in the past for similar problems. Although such an assumption sounds reasonable, it needs to be formally proven. A method which can be used for this purpose has been developed by Balakrishnan and Iliopoulos (2009). On the other hand, Balakrishnan et al. (2014) realized that under several sampling schemes which produce censored exponential lifetimes, there is a positive probability that the CDF of the MLE can not be inverted in the sense of (3). They showed further that whenever the MLE assumes a value such that some (or both) of the equations defining the endpoints of the confidence interval has no solution, the corresponding endpoint must be set infinite so that the nominal confidence level to be maintained.

The paper proceeds as follows. In Section 2 the stochastic monotonicity of the MLE is formally proven by using the approach of Balakrishnan and Iliopoulos (2009). This guarantees the uniqueness of the solutions of the equations in (3) *if* they exist. In Section 3 it is shown that the probability of nonexistence of these solutions is strictly positive. Numerical results illustrate that its magnitude is quite significant for some of the configurations used by Mao et al. (2013) in their simulations. The paper concludes with a discussion.

2 Stochastic monotonicity of the MLEs

The proof that the tail probability of the MLEs is strictly increasing in the corresponding parameters will be based on an application of the Three Monotonicities Lemma introduced by Balakrishnan and Iliopoulos (2009). According to this lemma, if the CDF of a random variable Y has the mixture representation

$$F(y;\theta)=\sum_{d\in\mathcal{D}}\mathsf{P}_\theta(D=d)F(y;\theta|D=d), \tag{4}$$

where $\mathcal{D}$ is a finite set, and (M1) the conditional distribution of Y, given $D=d$, is stochastically increasing in θ; (M2) the conditional distribution of Y, given $D=d$, is stochastically decreasing in d; (M3) D is stochastically decreasing in θ, then Y is (unconditionally) stochastically increasing in θ.

Below only the stochastic monotonicity of $\hat\theta_1$ with respect to θ_1 when θ_2 is fixed is proven. The proof of stochastic monotonicity of $\hat\theta_2$ proceeds similarly. In what follows, $[Y|D=d]$ denotes the conditional distribution of the random variable Y given $D=d$ and the notation $X\overset{\mathrm{d}}{=}Y$ is used to state that X and Y have the same distributions.

The right truncated exponential distribution at T is denoted by $\mathcal{E}(\theta; T)$. Moreover, let $W = \sum_{i=1}^{s_D} X_{i:n} + (n - s_D)T_D$ be the numerator of $\hat{\theta}_1$ and note that, conditional on $D = d$, it is independent of N_1. Indeed, following Mao et al. (2013), let $Z_i = 1$ or 0 according to whether $X_i = X_{i1}$ or X_{i2}. By the independence of X_{i1}, X_{i2} and the properties of the exponential distribution it follows that $[X_i | Z_i = j] \stackrel{\mathrm{d}}{=} X_i$. Hence, the random variables $X_1, \ldots, X_n$ are independent of $Z_1, \ldots, Z_n$ and consequently, conditional on $D = d$, W is independent of $N_1 = \sum_{i=1}^{s_D} Z_{[i:n]}$, where $Z_{[i:n]}$ is associated with $X_{i:n}$. Finally, in order to simplify the presentation, let $U_1, \ldots, U_n \stackrel{\text{iid}}{\sim} \mathcal{E}(\theta; T)$, $V_1, \ldots, V_n \stackrel{\text{iid}}{\sim} \mathcal{E}(\theta)$ denote mutually independent random variables which are also independent of D and N_1. Also notice that they all are stochastically increasing in θ and since θ is stictly increasing in θ_1 they are stochastically increasing in θ_1 as well.

The proofs of the three monotonicities follow.

(M1) When $d \leqslant k - 1$, condition on $D = d$ and write

$$W = \left[\sum_{i=1}^{d} X_{i:n} + (n-d)T\right] + \left[\sum_{i=d+1}^{k} (X_{i:n} - T) + (n-k)(X_{k:n} - T)\right], \tag{5}$$

where, by convention, $\sum_{i=a}^{b} X_{i:n} \equiv 0$ whenever $a > b$. By the block conditional independence of order statistics (cf. Iliopoulos and Balakrishnan, 2009), conditional on $D = d$, the random vectors $(X_{1:n}, \ldots, X_{d:n})$, $(X_{d+1:n}, \ldots, X_{k:n})$ are independent. In particular, when $d \geqslant 1$, $[(X_{1:n}, \ldots, X_{d:n}) | D = d] \stackrel{\mathrm{d}}{=} (U_{1:d}, \ldots, U_{d:d})$, while (for all $d \leqslant k-1$) $[(X_{d+1:n} - T, \ldots, X_{k:n} - T) | D = d] \stackrel{\mathrm{d}}{=} (V_{1:n-d}, \ldots, V_{k-d:n-d})$. Since $\sum_{i=1}^{d} U_{i:d} \equiv \sum_{i=1}^{d} U_i$ and $\sum_{i=1}^{k-d} V_{i:n-d} + \{(n-d) - (k-d)\} V_{k-d:n-d} \stackrel{\mathrm{d}}{=} \sum_{i=1}^{k-d} V_i$ (cf. Arnold et al., 2008), we have $[W | D = d] \stackrel{\mathrm{d}}{=} \sum_{i=1}^{d} U_{i:d} + (n-d)T + \sum_{i=1}^{k-d} V_{i:n-d} + (n-k) V_{k-d:n-d} \stackrel{\mathrm{d}}{=} \sum_{i=1}^{d} U_i + (n-d)T + \sum_{i=1}^{k-d} V_i$ which is the sum of independent random variables that are stochastically increasing in θ_1 and a constant, hence the conditional distribution of W given $D = d$ is stochastically increasing in θ_1.

For $k \leqslant d \leqslant r-1$, conditional on $D = d$, we have $W = \sum_{i=1}^{d} X_{i:n} + (n-d)T$. Since $[(X_{1:n}, \ldots, X_{d:n}) | D = d] \stackrel{\mathrm{d}}{=} (U_{1:d}, \ldots, U_{d:d})$ we have $[W | D = d] \stackrel{\mathrm{d}}{=} \sum_{i=1}^{d} U_i + (n-d)T$ hence the conditional distribution of W given $D = d$ is stochastically increasing in θ_1 for such d too.

For $r \leqslant d \leqslant n$, conditional on $D = d$, we have $W = \sum_{i=1}^{r} X_{i:n} + (n-r) X_{r:n}$. Then, $[(X_{1:n}, \ldots, X_{r:n}) | D = d] \stackrel{\mathrm{d}}{=} (U_{1:d}, \ldots, U_{r:d})$ and so $[W | D = d] \stackrel{\mathrm{d}}{=} \sum_{i=1}^{r-1} U_{i:r} + (n-r+1) U_{r:n}$ which is also stochastically increasing in θ_1.

Finally, conditional on $D = d$, N_1 follows a binomial distribution $\mathcal{B}(s_d, p(\theta_1, \theta_2))$, where $p(\theta_1, \theta_2) = \theta_2 / (\theta_1 + \theta_2)$, but doubly truncated to the set $\{1, \ldots, s_d - 1\}$. Since this truncated distribution has the monotone likelihood ratio property with respect to $p(\theta_1, \theta_2)$ which is strictly decreasing in θ_1, N_1 is stochastically strictly decreasing in θ_1. Therefore, conditional on $D = d$, $1/N_1$ is stochastically increasing in θ_1.

Hence, for any d, $[\hat{\theta}_1|D = d]$ is stochastically increasing in θ_1 because it has the distribution of the product of two independent, strictly positive, random variables which are stochastically increasing in θ_1.

(M2) The proof of this stochastic monotonicity is based on standard coupling. Consider first $d \leqslant k-1$. From the analysis in the previous subsection it turns out that $[\hat{\theta}_1|D = d] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{d} U_i + (n-d)T + \sum_{i=1}^{k-d} V_i\big\}/N_1$ and $[\hat{\theta}_1|D = d+1] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{d+1} U_i + (n-d-1)T + \sum_{i=1}^{k-d-1} V_i\big\}/N_1$. But

$$\frac{\sum\limits_{i=1}^{d} U_i + (n-d)T + \sum\limits_{i=1}^{k-d} V_i}{N_1} - \frac{\sum\limits_{i=1}^{d+1} U_i + (n-d-1)T + \sum\limits_{i=1}^{k-d-1} V_i}{N_1} = \frac{V_{k-d} + T - U_{d+1}}{N_1} > 0$$

almost surely, and this implies that $[\hat{\theta}_1|D = d]$ decreases stochastically in d for $d \leqslant k-1$.

Let now $k \leqslant d \leqslant r-2$. In this case, $[\hat{\theta}_1|D = d] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{d} U_i + (n-d)T\big\}/N_1$, $[\hat{\theta}_1|D = d+1] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{d+1} U_i + (n-d-1)T\big\}/N_1'$, where N_1' equals either N_1 or N_1+1 according to whether the additional random variable U_{d+1} is associated to the risk factor 2 or the risk factor 1, respectively. In any case,

$$\frac{\sum\limits_{i=1}^{d} U_i + (n-d)T}{N_1} - \frac{\sum\limits_{i=1}^{d+1} U_i + (n-d-1)T}{N_1'} \geqslant \frac{T - U_{d+1}}{N_1} > 0$$

almost surely, thus $[\hat{\theta}_1|D = d]$ keeps decreasing stochastically in d for these d's too. For $d = r-1$, $[\hat{\theta}_1|D = d] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{r-1} U_i + (n-r+1)T\big\}/N_1 \equiv \big\{\sum_{i=1}^{r-1} U_{i:r-1} + (n-r+1)T\big\}/N_1$ and $[\hat{\theta}_1|D = d+1] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{r} U_{i:r} + (n-r)U_{r:r}\big\}/N_1'$ where $N_1' = N_1$ or N_1+1 as above. Clearly, $\big\{\sum_{i=1}^{r} U_{i:r} + (n-r)U_{r:r}\big\}/N_1' \leqslant \big\{\sum_{i=1}^{r-1} U_{i:r} + (n-r+1)T\big\}/N_1$. By Zhuang and Hu (2007), $(U_{1:m}, \ldots, U_{m-1:m}) \leqslant_{\text{st}} (U_{1:m-1}, \ldots, U_{m-1:m-1})$ (with respect to the multivariate stochastic order) for all $m \geqslant 2$. So, by taking $m = r$, this implies that $[\hat{\theta}_1|D = r] \leqslant_{\text{st}} [\hat{\theta}_1|D = r-1]$.

Finally, for $r \leqslant d \leqslant n-1$, $[\hat{\theta}_1|D = d] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{r} U_{i:d} + (n-r)U_{r:d}\big\}/N_1$, $[\hat{\theta}_1|D = d+1] \stackrel{\mathrm{d}}{=} \big\{\sum_{i=1}^{r} U_{i:d+1} + (n-r)U_{r:d+1}\big\}/N_1$. Using once more the result of Zhuang and Hu (2007) with $m = d+1$ and considering only the first r entries of the two random vectors we get that $(U_{1:d+1}, \ldots, U_{r:d+1}) \leqslant_{\text{st}} (U_{1:d}, \ldots, U_{r:d})$. This implies that $[\hat{\theta}_1|D = d]$ is stochastically decreasing in d for $d \geqslant r$ as well.

Hence, $[\hat{\theta}_1|D = d]$ is stochastically decreasing in $d \in \{0, 1, \ldots, n\}$, i.e., the second stochastic monotonicity is also established.

(M3) The establishment of this stochastic monotonicity is straightforward since D has a binomial distribution $\mathcal{B}(n, 1 - e^{-T/\theta})$ and the probability of success $1 - e^{-T/\theta}$ is strictly decreasing in θ_1. Hence D is stochastically decreasing in θ_1 and the third monotonicity holds too.

Thus, all required stochastic monotonicities hold and this establishes the stochastic monotonicity of $\hat{\theta}_1$ with respect to θ_1.

3 On the existence of the solutions

Recall the expression of the CDF of $\hat{\theta}_1$ in (2). Note first that for any $y > 0$, all terms corresponding to d, j with $n - d + j > 0$ converge to zero as $\theta_1 \downarrow 0$ since in this case $e^{-(n-d+j)(1/\theta_1+1/\theta_2)T} \to 0$ and everything else involved in the corresponding term is bounded. The only case where $n - d + j = 0$ corresponds to $d = n$, $j = 0$. Then $s_d = r$ and the corresponding term equals

$$\sum_{i=1}^{r-1} \frac{\binom{r}{i}\theta_1^{r-i}\theta_2^{i}}{(\theta_1+\theta_2)^r - \theta_1^r - \theta_2^r} G\Big(\Big(\frac{1}{\theta_1} + \frac{1}{\theta_2}\Big) iy\,; r\Big).$$

It can be easily verified that as $\theta_1 \downarrow 0$, all terms of the above sum for $i \leqslant r - 2$ converge to zero while the term corresponding to $i = r - 1$ converges to one. Hence,

$$\lim_{\theta_1 \downarrow 0} F(y; \theta_1, \theta_2) = 1$$

for all y and θ_2. On the other hand, as $\theta_1 \uparrow \infty$, we have the limits $e^{-(n-d+j)(1/\theta_1+1/\theta_2)T} \to e^{-(n-d+j)T/\theta_2}$, $G\big(\big(\frac{1}{\theta_1} + \frac{1}{\theta_2}\big)\{iy - (n - d + j)T\}\,;\, s_d\big) \to G\big(\frac{iy-(n-d+j)T}{\theta_2}\,;\, s_d\big)$ and

$$\frac{\binom{s_d}{i}\theta_1^{s_d-i}\theta_2^{i}}{(\theta_1+\theta_2)^{s_d} - \theta_1^{s_d} - \theta_2^{s_d}} \longrightarrow \begin{cases} 1, & i = 1, \\ 0, & i > 1. \end{cases}$$

It follows that for any $y > 0$,

$$\lim_{\theta_1 \uparrow \infty} F(y; \theta_1, \theta_2) = \sum_{d=0}^{n}\sum_{j=0}^{d} (-1)^j \binom{n}{d}\binom{d}{j} e^{-(n-d+j)T/\theta_2} G\Big(\frac{y - (n-d+j)T}{\theta_2}\,; s_d\Big).$$

Now, $\lim_{\theta_1 \uparrow \infty} F(y; \theta_1, \theta_2)$ is continuous in y with $\lim_{y \downarrow 0} \lim_{\theta_1 \uparrow \infty} F(y; \theta_1, \theta_2) = 0$ and

$$\begin{aligned}
\lim_{y \uparrow \infty} \lim_{\theta_1 \uparrow \infty} F(y; \theta_1, \theta_2) &= \sum_{d=0}^{n}\sum_{j=0}^{d} (-1)^j \binom{n}{d}\binom{d}{j} e^{-(n-d+j)T/\theta_2} \\
&= \sum_{d=0}^{n} \binom{n}{d} e^{-(n-d)T/\theta_2} \sum_{j=0}^{d} (-1)^j \binom{d}{j} e^{-jT/\theta_2} \\
&= \sum_{d=0}^{n} \binom{n}{d} e^{-(n-d)T/\theta_2} (1 - e^{-T/\theta_2})^d = 1.
\end{aligned}$$

Thus $\lim_{\theta_1 \uparrow \infty} F(y; \theta_1, \theta_2)$ may assume any value in the interval $(0, 1)$, depending on y and θ_2. Since the support of $\hat{\theta}_1$ is the whole interval $(0, \infty)$ (see Table 1), it turns out that for any $\alpha \in (0,1)$ and $\theta_2 > 0$ there exists y_{α,θ_2} with $\mathsf{P}_{\theta_1,\theta_2}(\hat{\theta}_1 > y_{\alpha,\theta_2}) > 0$, such that whenever $\hat{\theta}_1^{\text{obs}} > y_{\alpha,\theta_2}$ at least one of the equations in (3) (with θ_2 in the place of $\hat{\theta}_2^{\text{obs}}$) has no solution. Recently, Balakrishnan et al. (2014) noticed that such a situation arises quite often in the analysis of exponential lifetimes based on experiments that run subject to time constraints. They further showed that in situations like this, whenever the solution of an equation does not exist the corresponding endpoint must be replaced by infinity in order the nominal confidence level to be maintained. This implies that when at least one endpoint is not defined with positive probability then the resulting exact confidence interval has infinite expected length. Moreover, when both endpoints are not defined the confidence interval results in an empty set.

Nevertheless, here θ_2 is unknown and must be replaced by the observed value of its MLE (or some other reasonable estimate). From (1) it is apparent that $\hat{\theta}_2 = N_1\hat{\theta}_1/N_2$. Hence, the actual equations that must be solved for θ_1 are of the form

$$F(y;\theta_1, n_1y/n_2) = \gamma, \tag{6}$$

where y, n_1, n_2 are the observed values of $\hat{\theta}_1^{\text{obs}}, N_1, N_2$ and $\gamma = 1-\alpha/2$ or $\alpha/2$. As shown before, for any values of y and θ_2 the limit of $F(y;\theta_1,\theta_2)$ as $\theta_1 \downarrow 0$ equals 1. On the other hand, as $\theta_1 \uparrow \infty$ the limit of $F(y;\theta_1, n_1y/n_2)$ becomes

$$\sum_{d=0}^{n}\sum_{j=0}^{d}(-1)^j\binom{n}{d}\binom{d}{j}e^{-(n-d+j)n_2T/(n_1y)}G\left(\frac{y-(n-d+j)T}{n_1y/n_2};s_d\right). \tag{7}$$

By the stochastic monotonicity of $\hat{\theta}_1$ with respect to θ_1 it is clear that whenever this limit exceeds (or is equal to) γ, the equation in (6) has no solution.

Table 2 shows the range of the limit in (7) for some configurations of (n,r,k) which have been used in the simulations presented in Mao et al. (2013), for several values d and $n_1 = 1$. In almost all of them the equation in (6) has no solution for $\gamma = 0.025$ and so if $N_1 = 1$ is observed then the upper endpoint of the 95% exact confidence interval is not defined. As a consequence, the corresponding confidence interval for θ_1 has infinite length. Note that in some cases the limit also exceeds .975 in which case instead of a confidence interval we get an empty set. For instance, consider the configuration $(n,r,k) = (20,16,8)$ and suppose that we observe $D \geqslant r = 16$, $N_1 = 1$ and $\hat{\theta}_1$ close to nT which in this case is its maximum possible value; see Table 1. Then the limit in (7) becomes larger than .975 which means that the equation in (6) has no solution for both $\gamma = .025$ and .975. On the other hand, for these particular configurations the limit of $F(y;\theta_1, n_1y/n_2)$ as $\theta_1 \uparrow \infty$ when $n_1 > 1$ is almost zero for all $y > 0$ (yet still positive) and thus (6) can be solved for any reasonable value of α. Table 3 shows the corresponding probabilities that $N_1 = 1$ for $(\theta_1,\theta_2) = (5,2)$ and $(7,3)$ and $T = 1,2,4$ which are the values of (θ_2,θ_1) and T originally used by Mao et al. (2013) in their simulations. We can see that for the two first configurations this probability is quite high especially when $T = 1$. In particular, when $(n, r, k) = (20, 12, 6)$ the probability that $N_1 = 1$ (and consequently that at least one of the exact confidence interval's endpoints is infinite) exceeds 23%. This means that Mao et al. (2013) should have faced this issue in more than one out of five simulated datasets but apparently they did not notice that the corresponding equations admit no solutions.

4 Discussion

The stochastic monotonicity of a statistic with respect to the parameter of interest is necessary for pivoting its CDF in order to obtain confidence intervals for this parameter. When this stochastic monotonicity does not hold, the resulting confidence sets may not be intervals and, more importantly, they are unlikely to have the nominal level. Therefore, whoever intends to use the method must first verify that the required stochastic monotonicity holds. Relying simply to intuition by arguing that it is reasonable the larger the true value of a particular parameter is the more probable it will be for its MLE to exceed a given value is not enough. For instance, the MLE of the scale parameter of a two-parameter exponential distribution is not stochastically monotone in the scale parameter when type-I censoring takes place (cf. Mitra et al., 2013). Fortunately, for the model discussed by Mao et al. (2013) the required stochastic monotonicity holds as is shown in Section 2.

The issue of possible nonexistence of solutions to equations of the form $F(y; \theta) = \gamma$ has been discussed only recently by Balakrishnan et al. (2014). Since Barlow et al. (1968) who formally presented the method of pivoting the CDF until the appearence of the aforementioned work no one seems to have realized that there is such a problem. The main reason is that the method is typically used in situations where the equations can not be solved analytically and everyone thought that the computers' occasional inability to return a solution was due to their insufficient precision. In this respect Mao et al. (2013) are not to be blamed for failing to notice that the equations in (3) have no solution for some values of the MLEs. Note that Balakrishnan et al. (2014) concluded that, at least in the context of estimation of exponential mean, the problem can appear in sampling schemes that run under time constraints like, for example, type-I censoring, hybrid type-I censoring (Epstein, 1954; Childs et al., 2003) and generalized hybrid type-II censoring (Chandrasekar et al., 2004). These sampling schemes involve time constraints since the experiment is terminated at timepoints T, $\min\{X_{r:n}, T\}$, $\min\{\max\{X_{r:n}, T\}, T'\}$, respectively, where $r < n$ and $0 < T < T'$ are fixed in advance. On the contrary, the problem is not present under type-II censoring, hybrid type-II censoring (Childs et al., 2003) and generalized hybrid type-I censoring which do not impose any time constraints. In these schemes the experiment is terminated at $X_{r:n}$, $\max\{X_{r:n}, T\}$, $\min\{\max\{X_{k:n}, T\}, X_{r:n}\}$, respectively, where $k < r < n$ and $T > 0$. Nevertheless, when competing risks are involved the problem of solutions' nonexistence appears under generalized hybrid type-I censoring as well, as it is seen in Section 3. The common aspect of this context and of the schemes that run under time constraints is the requirement of at least one observation related to the parameter of interest in order the corresponding MLE to exist. Therefore it seems that the problem arises due to this requirement rather than the time constraint imposed to the experiment.

One may wonder whether it is feasible the construction of "exact" confidence intervals for θ_1 (or θ_2) which have finite expected length. By setting $\rho = \theta_2/\theta_1$ we can rewrite the CDF of $\hat{\theta}_1$ as

$$F^*(y;\theta_1,\rho) = \sum_{d=0}^{n}\sum_{i=1}^{s_d-1}\sum_{j=0}^{d} \frac{(-1)^j \binom{n}{d}\binom{d}{j}\binom{s_d}{i}\rho^i}{(1+\rho)^{s_d} - 1 - \rho^{s_d}} e^{-(n-d+j)(1+1/\rho)T/\theta_1} \times$$
$$G\left(\frac{1}{\theta_1}\left(1+\frac{1}{\rho}\right)\{iy - (n-d+j)T\}\,; s_d\right), \quad y > 0.$$

It can be easily verified that for fixed $y, \rho > 0$ we have $\lim_{\theta_1 \downarrow 0} F^*(y;\theta_1,\rho) = 1$ and $\lim_{\theta_1 \uparrow \infty} F^*(y;\theta_1,\rho) = 0$. These limits still hold if ρ is replaced by the observed value of its MLE $\hat{\rho} = N_1/N_2$. It turns out that the equation $F^*(\hat{\theta}_1^{\text{obs}};\theta_1,\hat{\rho}^{\text{obs}}) = \gamma$ has always a solution with respect to θ_1 and thus by solving it with $\gamma = 1 - \alpha/2$ and $\alpha/2$ we get a confidence interval with finite endpoints and thus with finite expected length. However, simulations show that the coverage probability of confidence intervals based on $F^*(\hat{\theta}_1^{\text{obs}};\theta_1,\hat{\rho}^{\text{obs}})$ may be much smaller than the nominal confidence level, see Table 4. Therefore the confidence intervals suggested by Mao et al. (2013) should be preferred even if they have infinite length with positive probability.

Acknowledgement

The author would like to thank the associate editor and one reviewer for their comments which helped to improve the paper.

References

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N. (2008). *A first course in order statistics*, Classic Edition, SIAM, Philadelphia.

Balakrishnan, N., Iliopoulos, G. (2009). Stochastic monotonicity of the MLE of exponential mean under different censoring schemes, *Annals of the Institute of*

Statistical Mathematics, 61, 753–772.

Balakrishnan, N., Cramer, E., Iliopoulos, G. (2014). On the method of pivoting the CDF for exact confidence intervals with illustration for exponential mean under life-test with time constraints. *Statistics and Probability Letters*, 89, 124–130.

Barlow, R.E., Madansky, A., Proschan, F., Scheuer, E. (1968). Statistical estimation procedures for the "burn-in" process. *Technometrics*, 10, 51–62.

Casella, G. and Berger, R.L. (2002). *Statistical Inference*, Second edition, Duxbury Press, Boston.

Chandrasekar, B., Childs, A., Balakrishnan, N. (2004). Exact likelihood inference for the exponential distribution under generalized Type-I and Type-II hybrid censoring, *Naval Research Logistics*, 7, 994–1004.

Childs, A., Chandrasekar, B., Balakrisnan N., Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution, *Annals of the Institute of Statistical Mathematics*, 55, 319–330.

Epstein, B. (1954). Truncated life tests in the exponential case, *Annals of Mathematical Statistics*, 25, 555–564.

Iliopoulos, G., Balakrishnan, N. (2009). Conditional independence of blocked ordered data, *Statistics and Probability Letters*, 79, 1008–1015.

Mao, S., Shi, Y.-M., Sun, Y.-D. (2013). Exact inference for competing risks with generalized type-I hybrid censored exponential data, *Journal of Statistical Computation and Simulation* (to appear).

Mitra, S., Ganguly, A., Samanta, D., Kundu, D. (2013). On the simple step-stress model for two-parameter exponential distribution, *Statistical Methodology*, 15, 95–114.

Zhuang, W., Hu, T. (2007). Multivariate stochastic comparisons of sequential order statistics, *Probability in the Engineering and Informational Sciences*, 21, 47-66.

Range of D	Range of N_1	Range of $\hat{\theta}_1$
$d \in \{0, \ldots, k-1\}$	$n_1 \in \{1, \ldots, k-1\}$	$((n-d)T/n_1, \infty)$
$d \in \{k, \ldots, r-1\}$	$n_1 \in \{1, \ldots, d-1\}$	$((n-d)T/n_1, nT/n_1)$
$d \in \{r, \ldots, n\}$	$n_1 \in \{1, \ldots, r-1\}$	$(0, nT/n_1)$

Table 1: Range of $\hat{\theta}_1$ for all possible values of D and N_1. The union of the intervals in the last column is $(0, \infty)$.

$(n,r,k) = (20,12,6)$			$(n,r,k) = (20,16,8)$			$(n,r,k) = (40,24,8)$		
d	min	max	d	min	max	d	min	max
0	.3840	.3840	0	.4013	.4013	0	.4013	.4013
1	.3839	.3840	3	.3937	.4013	5	.1888	.4013
2	.3762	.3840	4	.3431	.4013	8	.0103	.4013
3	.3081	.3840	5	.2282	.4013	9	.0088	.5470
4	.1681	.3840	6	.1167	.4013	10	.0077	.6761
5	.0662	.3840	7	.0522	.4013	11	.0068	.7789
6	.0230	.3840	8	.0218	.4013	12	.0062	.8568
7	.0288	.5543	9	.0236	.5470	16	.0198	.9820
8	.0552	.6993	10	.0308	.6761	17	.0367	.9900
9	.1120	.8087	11	.0515	.7798	18	.0633	.9946
10	.1970	.8843	13	.1556	.9105	20	.1510	.9985
11	.3032	.9329	15	.3306	.9684	22	.2840	.9996
$\geqslant 12$	.4207	.9625	$\geqslant 16$	.4319	.9820	$\geqslant 24$	.4449	.9999

Table 2: The minimum and maximum values that can achieve the limit in (7) for selected configurations of (n, r, k) and values of d when $n_1 = 1$.

		(n,r,k)		
(θ_1, θ_2)	T	$(20,12,6)$	$(20,16,8)$	$(40,24,8)$
$(5,2)$	1	.1602	.1457	.0136
	2	.0877	.0457	.0030
	4	.0862	.0296	.0030
$(7,3)$	1	.2327	.1826	.0411
	2	.0929	.0773	.0032
	4	.0722	.0252	.0020

Table 3: Probability of observing $N_1 = 1$ for selected configurations of (n, r, k), (θ_1, θ_2) and T.

		(n,r,k)					
(θ_1,θ_2)	T	$(20,12,6)$		$(20,16,8)$		$(40,24,8)$	
$(5,2)$	1	.971	.801	.959	.815	.967	.957
	2	.950	.868	.948	.906	.954	.949
	4	.945	.857	.944	.908	.949	.935
$(7,3)$	1	.969	.747	.956	.788	.938	.927
	2	.962	.866	.950	.881	.951	.938
	4	.948	.872	.952	.913	.954	.944

Table 4: Coverage probabilities of confidence intervals of 95% nominal level based on $F(\hat{\theta}_1^{\text{obs}};\theta_1,\hat{\theta}_2^{\text{obs}})$ (left columns) and $F^*(\hat{\theta}_1^{\text{obs}};\theta_1,\hat{\rho}^{\text{obs}})$ (right columns) for selected configurations of (n,r,k), (θ_1,θ_2) and T, estimated from 2000 Monte Carlo iterations.