Stochastic monotonicity of the MLE of exponential mean under different censoring schemes

N. Balakrishnan[1] and **G. Iliopoulos**[2]

Abstract

In this paper, we present a general method which can be used in order to show that the maximum likelihood estimator (MLE) of an exponential mean θ is stochastically increasing with respect to θ under different censored sampling schemes. This propery is essential for the construction of exact confidence intervals for θ via "pivoting the cdf" as well as for the tests of hypotheses about θ. The method is shown for Type-I censoring, hybrid censoring and generalized hybrid censoring schemes. We also establish the result for the exponential competing risks model with censoring.

AMS 2000 subject classifications: 62F10, 60E99

Keywords and phrases: Exponential distribution, maximum likelihood estimation, Type-I censoring, Type-I and Type-II hybrid censoring, Type-I and Type-II generalized hybrid censoring, exact confidence intervals, stochastic ordering, competing risks model.

1 Introduction

A standard method for constructing exact confidence intervals for a real parameter θ based on a statistic $\hat{\theta}$ is "pivoting the cdf", or, equivalently, the survival function; see, for example, Casella and Berger (2002, p. 432). The method is applicable as long as $\hat{\theta}$ is stochastically monotone with respect to θ, that is, $\mathsf{P}_\theta(\hat{\theta} > x)$ is a monotone function of θ for all x. Assuming without loss of generality that it is increasing, the method then proceeds as follows: Choose α_1 and α_2 such that $\alpha_1 + \alpha_2 = \alpha$ (for example, $\alpha_1 = \alpha_2 = \alpha/2$) and solve the equations $\mathsf{P}_\theta(\hat{\theta} > \hat{\theta}_{\text{obs}}) = \alpha_1$, $\mathsf{P}_\theta(\hat{\theta} > \hat{\theta}_{\text{obs}}) = 1 - \alpha_2$ for θ. Here, $\hat{\theta}_{\text{obs}}$ is the observed value of $\hat{\theta}$ determined from the given sample. The existence and uniqueness of the solutions of these equations are then guaranteed by the monotonicity of $\mathsf{P}_\theta(\hat{\theta} > \hat{\theta}_{\text{obs}})$

[1]Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario, Canada L8S 4K1; e–mail: `bala@mcmaster.ca`

[2]Department of Statistics and Insurance Science, University of Piraeus, 80 Karaoli & Dimitriou str., 18534 Piraeus, Greece; e–mail: `geh@unipi.gr`

with respect to θ. Denote by $\theta_L(\hat{\theta}_{\text{obs}}) < \theta_U(\hat{\theta}_{\text{obs}})$ these solutions. Then, $[\theta_L(\hat{\theta}_{\text{obs}}), \theta_U(\hat{\theta}_{\text{obs}})]$ is the realization of an exact $100(1-\alpha)\%$ confidence interval for the parameter θ.

Obviously, the stochastic monotonicity of $\hat{\theta}$ with respect to θ is crucial in the above construction. However, in the literature, a series of papers have been published constructing exact confidence intervals for the parameters of interest by assuming the stochastic monotonicity of the corresponding MLEs and not being able to show it theoretically but only observing it empirically. In particular, Chen and Bhattacharyya (1988), Childs *et al.* (2003), and Chandrasekar *et al.* (2004) derived the MLE of the exponential mean θ as well as its distribution for different censoring schemes, but they did not provide a formal proof that these MLEs are stochastically increasing with respect to the parameter θ. In all these cases, the survival function of the MLE takes on a mixture form

$$\mathsf{P}_\theta(\hat{\theta} > x) = \sum_{d \in \mathcal{D}} \mathsf{P}_\theta(D = d)\mathsf{P}_\theta(\hat{\theta} > x | D = d), \tag{1}$$

where $\mathcal{D}$ is a finite set. They all conjectured that the MLEs are stochastically increasing and supported it by presenting numerical results for some special cases. They then proceeded to the construction of exact confidence limits by "pivoting the survival function". In this paper, we formally prove that these conjectures are indeed true thus validating the exact inferential procedures developed by all these authors.

Another useful need for the stochastic monotonicity of the MLE is in the context of hypothesis testing. Suppose we want to test $H_0 : \theta \leqslant \theta_0$ versus $H_1 : \theta > \theta_0$. It is natural to consider tests of the form $\hat{\theta} > C_\alpha(\theta_0)$, where $C_\alpha(\theta_0)$ denotes the upper α-quantile of the distribution of $\hat{\theta}$ at θ_0. However, in order for such a test to have desirable properties such as unbiasedness and monotone power function, the MLE $\hat{\theta}$ should be stochastically increasing in θ.

This paper is organized as follows. In Section 2, we present a lemma providing three conditions which together are sufficient for a survival function of the form in (1) to be increasing in θ. In other words, successive verification of these conditions would imply that $\hat{\theta}$ is stochastically increasing in θ. In the subsequent sections, we apply this lemma in different censoring scenarios from an exponential distribution. In Section 3, we consider the case of the usual Type-I censoring as an illustrative example, since the application of the lemma in this case is quite straightforward. Moreover, this particular result will be used repeatedly in the sequel. In Section 4, we prove the stochastic monotonicity of the MLE under hybrid censoring, while in Section 5 we establish the result for generalized hybrid censoring. Section 6 summarizes the results and discusses some other potential applications of our approach. In addition, we prove in this section the stochastic monotonicity of the MLE in the setting of exponential competing risks, a result conjectured earlier by Kundu and Basu (2000). Finally, the technical results needed for verifying the conditions of the basic lemma are presented in an Appendix.

2 The basic lemma

Suppose that the survival function of $\hat{\theta}$ has the form in (1). Then, the following lemma holds.

Lemma 2.1. [Three Monotonicities Lemma] *Assume that the following hold true:*

(M1) For all $d \in \mathcal{D}$, the conditional distribution of $\hat{\theta}$, given $D = d$, is stochastically increasing in θ, i.e., the function $\mathsf{P}_\theta(\hat{\theta} > x|D = d)$ is increasing in θ for all x and $d \in \mathcal{D}$;

(M2) For all x and $\theta > 0$, the conditional distribution of $\hat{\theta}$, given $D = d$, is stochastically decreasing in d, i.e., the function $\mathsf{P}_\theta(\hat{\theta} > x|D = d)$ is decreasing in $d \in \mathcal{D}$;

(M3) D is stochastically decreasing in θ.

Then, $\hat{\theta}$ is stochastically increasing in θ.

Proof. It is well-known that if $X \leqslant_{\rm st} Y$, where "$\leqslant_{\rm st}$" means stochastically smaller, then for any integrable decreasing function g we have $\mathsf{E}\{g(X)\} \geqslant \mathsf{E}\{g(Y)\}$; see Shaked and Shanthikumar (2007). Therefore, under the assumptions of the lemma, for any $\theta < \theta'$,

$$\begin{aligned}
\mathsf{P}_\theta(\hat{\theta} > x) &= \sum_{d \in \mathcal{D}} \mathsf{P}_\theta(D = d)\mathsf{P}_\theta(\hat{\theta} > x|D = d) \\
&\leqslant \sum_{d \in \mathcal{D}} \mathsf{P}_{\theta'}(D = d)\mathsf{P}_\theta(\hat{\theta} > x|D = d) \quad \text{(by M2 and M3)} \\
&\leqslant \sum_{d \in \mathcal{D}} \mathsf{P}_{\theta'}(D = d)\mathsf{P}_{\theta'}(\hat{\theta} > x|D = d) \quad \text{(by M1)} \\
&= \mathsf{P}_{\theta'}(\hat{\theta} > x)
\end{aligned}$$

as required. □

Hence, a proof of the stochastic monotonicity of $\hat{\theta}$ with respect to θ may be completed in three steps, that is, establishing the three conditions of Lemma 2.1. We will refer to the above lemma as TML (Three Monotonicities Lemma) in the sequel.

3 Type-I censoring

Type-I censoring is the most practical type of censoring in that the duration of the experiment is fixed in advance by the experimenter. Specifically, let $T > 0$ be a fixed time and let $X_1, \ldots, X_n$ be iid random variables from an exponential distribution $\mathcal{E}(\theta)$, $\theta > 0$. Suppose that the life-test is terminated at time T, and D denotes the number of observed failures. Clearly, D is a random variable. By writing down the likelihood, it can be easily seen that the MLE of θ does not exist if $D = 0$. Hence, in order to make inference about θ, we must condition on the event $D \geqslant 1$. In this case, the MLE of θ is

$$\hat{\theta} = \frac{1}{D}\left\{\sum_{i=1}^{D} X_{i:n} + (n - D)T\right\}. \tag{2}$$

The conditional distribution of $\hat{\theta}$, given $D \geqslant 1$, has been explicitly derived by Bartholomew (1963). He then found the mean and the variance of this distribution and used them in order to make asymptotic inference for θ via the Central Limit Theorem. Later on, Spurrier and Wei (1980) used this conditional distribution in order to make exact inference for θ. They stated that "it can be shown that $\mathsf{P}_\theta(\hat{\theta} \geqslant c)$ is an increasing function of θ", but did not present a proof. The result was formally proved by Balakrishnan *et al.* (2002) by using a coupling argument.

Conditional on $D \geqslant 1$, the survival function of the MLE can be expressed as

$$\mathsf{P}_\theta(\hat{\theta} > x) = \sum_{d=1}^{n} \mathsf{P}_\theta(D = d | D \geqslant 1)\mathsf{P}_\theta(\hat{\theta} > x | D = d), \tag{3}$$

and so it has the form in (1) with $\mathcal{D} = \{1, \ldots, n\}$. Of course, (3) coincides with the expression of Bartholomew (1963), although this is not clear at first glance. Below, we prove once more the stochastic monotonicity of $\hat{\theta}$ with respect to θ using TML. Its application is rather straightforward in this case, and so it will also serve as an illustrative example. Moreover, this result will be used repeatedly in the following sections. Now, we proceed to the verification of the three monotonicities.

(M1) Recall that we have to show that the conditional distribution of $\hat{\theta}$, given $D = d$, is stochastically increasing in θ. To this end, note that conditional on $D = d$, $(X_{1:n}, \ldots, X_{d:n})$ have the same distribution as $(Z_{1:d}, \ldots, Z_{d:d})$, where $Z_1, \ldots, Z_d \stackrel{\text{iid}}{\sim} \mathcal{E}(\theta)I(Z \leqslant T)$, i.e., exponential with parameter θ but right truncated at T; see Arnold *et al.* (1992). Hence, conditional on $D = d$, $\sum_{i=1}^{D} X_{i:n} \stackrel{\text{d}}{=} \sum_{i=1}^{d} Z_{i:d} \equiv \sum_{i=1}^{d} Z_i$. Since the right truncated exponential distribution is stochastically increasing in θ and Z_i's are independent, the required monotonicity follows immediately.

(M2) Next, we have to prove that the conditional distribution of $\hat{\theta}$, given $D = d$, is stochastically decreasing in d. This will be done via standard coupling. For any $d \in \{1, \ldots, n-1\}$, let $Z_1, \ldots, Z_d, Z_{d+1}$ be iid from $\mathcal{E}(\theta)I(Z \leqslant T)$. Then,

$$\hat{\theta}|(D = d) \quad \text{has the same distribution as} \quad \frac{1}{d}\left\{\sum_{i=1}^{d} Z_i + (n - d)T\right\}$$

while

$$\hat{\theta}|(D=d+1) \quad \text{has the same distribution as} \quad \frac{1}{d+1}\left\{\sum_{i=1}^{d+1} Z_i + (n-d-1)T\right\}.$$

But,

$$\begin{aligned}\frac{1}{d}\left\{\sum_{i=1}^{d} Z_i + (n-d)T\right\} &- \frac{1}{d+1}\left\{\sum_{i=1}^{d+1} Z_i + (n-d-1)T\right\} \\ &= \frac{\sum_{i=1}^{d} Z_i + (n-d)T + d(T - Z_{d+1})}{d(d+1)} > 0,\end{aligned}$$

which implies that $\mathsf{P}_\theta(\hat{\theta} > x | D = d) > \mathsf{P}_\theta(\hat{\theta} > x | D = d+1)$ for all $x, \theta > 0$.

(M3) Finally, we should verify that D is stochastically decreasing in θ. However, this is a consequence of the fact that D has the monotone likelihood ratio property with respect to θ. This is proven in Lemma A.2(a) (with D, T in the place of D_1, T_1).

Thus follows the monotonicity of the conditional survival function of the MLE in (3) for the case of Type-I censoring.

4 Hybrid censoring

4.1 Type-I hybrid censoring

Suppose there are n identical units under test, and that $T > 0$ and $r \in \{1, \ldots, n\}$ are fixed. In this particular sampling scheme, the life-test stops at the random time $T_1^* = \min\{X_{r:n}, T\}$. The scheme was introduced first by Epstein (1954). By assuming that the lifetimes $X_1, \ldots, X_n$ are from the exponential distribution $\mathcal{E}(\theta)$, he found the MLE as

$$\hat{\theta} = \begin{cases} \dfrac{1}{D}\left\{\displaystyle\sum_{i=1}^{D} X_{i:n} + (n-D)T\right\}, & \text{if } D = 1, \ldots, r-1, \\ \dfrac{1}{r}\left\{\displaystyle\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\}, & \text{if } D = r, \ldots, n, \end{cases} \tag{4}$$

where again $D = \#\{X\text{'s} \leqslant T\}$. Chen and Bhattacharyya (1988) derived the exact distribution of the MLE of θ, but this was in a very complicated form. It was simplified later by Childs *et al.* (2003) who termed this sampling scheme "Type-I hybrid censoring" since it shares with standard Type-I censoring the feature that the total time under test is no more than the pre-fixed time T. As mentioned earlier, in both these papers, the authors were not able to prove the stochastic monotonicity of the MLE with respect to θ. Here, we shall prove this result using TML.

(M1) As already mentioned in the case of Type-I censoring, conditional on $D = d$, $(X_{1:n}, \ldots, X_{d:n})$ has the same distribution as the order statistics $(Z_{1:d}, \ldots, Z_{d:d})$ in a sample of size d from the right truncated exponential distribution $\mathcal{E}(\theta)I(Z \leqslant T)$. Thus, for any $x \in \mathbb{R}$,

$$\mathsf{P}_\theta(\hat{\theta} > x | D = d) = \mathsf{P}_\theta(w_0 + \textstyle\sum_{i=1}^{d} w_d Z_{i:d} > x),$$

where $w_1 = \cdots = w_d = 1/d$ and $w_0 = (n-d)T/d$ for $d < r$, and $w_1 = \cdots = w_{r-1} = 1/r$, $w_r = (n-r+1)/r$, $w_0 = w_{r+1} = \cdots = w_d = 0$, for $d \geqslant r$. Since $\mathcal{E}(\theta)I(Z \leqslant T)$ is stochastically increasing in θ, the result follows from Lemma B.1.

(M2) For $d \leqslant r-2$, the result is the same as that in Section 3. For $d = r-1$, let $Z_1, \ldots, Z_r$ be iid observations from $\mathcal{E}(\theta)I(Z \leqslant T)$. Then,

$$\hat{\theta}|(D = r-1) \quad \stackrel{\mathrm{d}}{=} \quad \frac{1}{r-1}\left\{\sum_{i=1}^{r-1} Z_i + (n-r+1)T\right\}$$

and

$$\hat{\theta}|(D = r) \quad \stackrel{\mathrm{d}}{=} \quad \frac{1}{r}\left\{\sum_{i=1}^{r} Z_r + (n-r)Z_{r:r}\right\}.$$

Now,

$$\begin{aligned}&\frac{1}{r-1}\left\{\sum_{i=1}^{r-1} Z_i + (n-r+1)T\right\} - \frac{1}{r}\left\{\sum_{i=1}^{r} Z_r + (n-r)Z_{r:r}\right\}\\ &\qquad = \frac{\sum_{i=1}^{r} Z_i + r(Z_{r:r} - Z_r) + (n-r)Z_{r:r} + r(n-r+1)(T - Z_{r:r})}{r(r-1)} > 0,\end{aligned}$$

which implies the result. Finally, for $d \geqslant r$, the result is obtained by applying Lemma B.2.

(M3) The distribution of D is the same as in Section 3.

Thus follows the monotonicity of the conditional survival function of the MLE in (4) for the case of Type-I hybrid censoring.

4.2 Type-II hybrid censoring

In Type-I hybrid censoring, there is a possibility of observing no failures at all. For that reason, Childs *et al.* (2003) proposed an alternative sampling scheme wherein the life-test terminates at the random time $T_2^* = \max\{X_{r:n}, T\}$. This guarantees that at least r failures will be observed. In this case, the MLE of θ is given by

$$\hat{\theta} = \begin{cases} \dfrac{1}{r}\left\{\displaystyle\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\}, & \text{if } D = 0, 1, \ldots, r-1, \\ \dfrac{1}{D}\left\{\displaystyle\sum_{i=1}^{D} X_{i:n} + (n-D)T\right\}, & \text{if } D = r, r+1, \ldots, n. \end{cases} \tag{5}$$

Note here that the MLE is always defined, and so $\mathcal{D} = \{0, 1, \dots, n\}$.

We now proceed to establishing the stochastic monotonicity of $\hat{\theta}$ via TML.

(M1) For $d \geqslant r$, the stochastic monotonicity of $\hat{\theta}$, given $D = d$, has already been proved in Section 3. In order to prove it for $d < r$, use Lemma B.4 to get that, conditional on $D = d$, $(X_{1:n}, \dots, X_{d:n})$ and $(X_{d+1:n}, \dots, X_{r:n})$ are independent. This implies that, conditional on $D = d$, $\sum_{i=1}^{d} X_{d:n}$ and $\sum_{i=d+1}^{r} X_{i:n} + (n-r)X_{r:n}$ are also independent. Moreover, conditional on $D = d$, $(X_{1:n}, \dots, X_{d:n})$ has the same distribution as the order statistics in a sample of size d from the right-truncated exponential distribution $\mathcal{E}(\theta)I(X \leqslant T)$ and $(X_{d+1:n}, \dots, X_{r:n})$ has the same distribution as the first $r - d$ order statistics in a sample of size $n - d$ from the left-truncated exponential distribution $\mathcal{E}(\theta)I(X > T)$. Since both these distributions are stochastically increasing in θ, the stochastic monotonicity of both sums with respect to θ follows from Lemma B.1. By their independence, their sum inherits the stochastic monotonicity.

(M2) For $d \geqslant r$, the situation is the same as in the Type-I censoring case. In order to prove it for $d = r - 1$, let $Z_1, \dots, Z_r$ be iid random variables from the right-truncated exponential distribution $\mathcal{E}(\theta)I(Z \leqslant T)$ and Y an independent random variable having the same distribution as the minimum in a sample of size $n - r + 1$ from the left-truncated exponential distribution $\mathcal{E}(\theta)I(Z > T)$. (This is in fact the conditional distribution of $X_{r:n}$ given $d = r - 1$.) Then,

$$\hat{\theta}|(D = r-1) \quad \stackrel{\mathrm{d}}{=} \quad \frac{1}{r}\left\{\sum_{i=1}^{r-1} Z_i + (n-r+1)Y\right\}$$

and

$$\hat{\theta}|(D = r) \quad \stackrel{\mathrm{d}}{=} \quad \frac{1}{r}\left\{\sum_{i=1}^{r} Z_r + (n-r)T\right\}.$$

But,

$$\left\{\sum_{i=1}^{r-1} Z_i + (n-r+1)Y\right\} - \left\{\sum_{i=1}^{r} Z_r + (n-r)T\right\} = (n-r)(Y-T) + Y - Z_r > 0$$

with probability one, since $Y > T > Z_r$ with probability one.

Finally, let us consider the case $d \leqslant r - 2$. Let $Z_1, \dots, Z_{d+1} \stackrel{\text{iid}}{\sim} \mathcal{E}(\theta)I(Z \leqslant T)$ and

$W_1, \ldots, W_{r-d} \stackrel{\text{iid}}{\sim} \mathcal{E}(\theta)$, independent of Z's. Then, we have

$$\begin{aligned}
\hat{\theta}|(D=d) &= \frac{1}{r}\left\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\right\} \\
&= \frac{1}{r}\left\{\sum_{i=1}^{d} X_{i:n} + (n-d)T + \sum_{i=d+1}^{r}(X_{i:n}-T) + (n-r)(X_{r:n}-T)\}\right\} \\
&\stackrel{\text{d}}{=} \frac{1}{r}\left\{\sum_{i=1}^{d} Z_i + (n-d)T + \sum_{i=1}^{r-d} W_i\right\}.
\end{aligned}$$

The sum of W's appears above since conditional on $D=d$, $(X_{d+1:n}-T, \ldots, X_{r:n}-T)$ has the same distribution as the first $r-d$ order statistics in a sample of size $n-d$ from $\mathcal{E}(\theta)$, and that $\sum_{i=1}^{r-d} W_{i:n-d} + \{(n-d)-(r-d)\}W_{r:n}$ follows a gamma distribution $\mathcal{G}(r-d,\theta)$. Similarly,

$$\hat{\theta}|(D=d+1) \stackrel{\text{d}}{=} \frac{1}{r}\left\{\sum_{i=1}^{d+1} Z_i + (n-d-1)T + \sum_{i=1}^{r-d-1} W_i\right\}.$$

Taking their difference (and omitting $1/r$), we get

$$\left\{\sum_{i=1}^{d} Z_i + (n-d)T + \sum_{i=1}^{r-d} W_i\right\} - \left\{\sum_{i=1}^{d+1} Z_i + (n-d-1)T + \sum_{i=1}^{r-d-1} W_i\right\} = W_{r-d} + T - Z_{d+1} > 0$$

with probability one. Hence, the condition holds in this case as well.

(M3) It is the same as in the previous cases.

Thus follows the monotonicity of the survival function of the MLE in (5) for the case of the Type-II hybrid censoring.

5 Generalized hybrid censoring

Both Type-I and Type-II hybrid censoring schemes have some potential drawbacks. Specifically, in Type-I hybrid censoring, there may be very few or even no failures observed whereas in Type-II hybrid censoring the experiment could last for a very long period of time. In order to overcome these drawbacks, Chandrasekar *et al.* (2004) defined generalized hybrid censoring schemes and derived the MLEs of the exponential mean lifetime θ. However, the stochastic monotonicity of these MLEs was not proved by these authors.

5.1 Generalized Type-I hybrid censoring

Recall the notation of Section 4.1. Now, in addition to T and r, fix $k \in \{1, \ldots, r-1\}$ and terminate the life-test at $T_1^{**} = \max\{X_{k:n}, T_1^*\} = \max\{X_{k:n}, \min\{X_{r:n}, T\}\}$. This censoring scheme guarantees that at least k failures will be observed. If the lifetimes are from $\mathcal{E}(\theta)$, the MLE of θ has been derived by Chandrasekar *et al.* (2004) to be

$$\hat{\theta} = \begin{cases} \frac{1}{k}\Big\{\sum_{i=1}^{k} X_{i:n} + (n-k)X_{k:n}\Big\}, & \text{if } D = 0, 1, \ldots, k-1, \\ \frac{1}{D}\Big\{\sum_{i=1}^{D} X_{i:n} + (n-D)T\Big\}, & \text{if } D = k, \ldots, r-1, \\ \frac{1}{r}\Big\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Big\}, & \text{if } D \geqslant r, \end{cases} \tag{6}$$

where again $D = \{\#X\text{'s} \leqslant T\}$. There appears to be a misprint in Chandrasekar *et al.* (2004) in that in the last case the MLE seems to be defined only for $D = r$ rather than for $D \geqslant r$.

We could again use TML to prove the stochastic monotonicity of the MLE. However, all the work has been done in the previous section since actually the above MLE has a form similar to the MLEs in hybrid censoring. Specifically, for $D \leqslant r-1$, $\hat{\theta}$ is exactly like the MLE in Type-II hybrid censoring case (but with k and $r-1$ instead of r and n, respectively) whereas for $D \geqslant k-1$ it is similar to the MLE in Type-I hybrid censoring case (but with k instead of 1). Hence, the stochastic monotonicity of the survival function of $\hat{\theta}$ in (6) may be proved exactly along the same lines.

5.2 Generalized Type-II hybrid censoring

We shall now slightly change the notation and denote T and D by T_1 and D_1, respectively. This is because under generalized Type-II hybrid censoring a second time point $T_2 > T_1$ is fixed and the life-test is terminated at the random time $T_2^{**} = \min\{T_1^*, T_2\} = \min\{\max\{X_{r:n}, T_1\}, T_2\}$. Under this censoring scheme, it is guaranteed that the total time under test will be at most T_2.

Define $D_2 = \{\#X\text{'s} \leqslant T_2\}$ and $\Delta_2 = T_2 - T_1$. Under exponentiality, the MLE of θ has been derived by Chandrasekar *et al.* (2004) to be

$$\hat{\theta} = \begin{cases} \frac{1}{D_1}\Big\{\sum_{i=1}^{D_1} X_{i:n} + (n-D_1)T_1\Big\}, & \text{if } D_1 = r, r+1, \ldots, n, \\ \frac{1}{r}\Big\{\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Big\}, & \text{if } D_1 = 0, 1, \ldots, r-1,\ D_2 \geqslant r, \\ \frac{1}{D_2}\Big\{\sum_{i=1}^{D_2} X_{i:n} + (n-D_2)T_2\Big\}, & \text{if } D_2 = 1, 2, \ldots, r-1. \end{cases} \tag{7}$$

Note that in Chandrasekar *et al.* (2004) there is a misprint in this case too, in that in the second line the MLE is defined only for $D_2 = r$ rather than for $D_2 \geqslant r$.

In order to express $\hat{\theta}$ in (7) in a suitable form for using TML, we introduce an auxiliary random variable D with pmf

$$\mathsf{P}_\theta(D = d) = \begin{cases} \mathsf{P}_\theta(D_2 = d)/\mathsf{P}_\theta(D_2 \geqslant 1), & d = 1, \ldots, r-1, \\ \mathsf{P}_\theta(D_1 \leqslant r-1, D_2 \geqslant r)/\mathsf{P}_\theta(D_2 \geqslant 1), & d = r', \\ \mathsf{P}_\theta(D_1 = d)/\mathsf{P}_\theta(D_2 \geqslant 1), & d = r, \ldots, n, \end{cases}$$

where r' is some (irrelevant) value between $r-1$ and r. Then, the MLE of θ can be expressed as

$$\hat{\theta} = \begin{cases} \dfrac{1}{D}\Big\{\displaystyle\sum_{i=1}^{D} X_{i:n} + (n-D)T_2\Big\}, & \text{if } D = 1, \ldots, r-1, \\ \dfrac{1}{r}\Big\{\displaystyle\sum_{i=1}^{r} X_{i:n} + (n-r)X_{r:n}\Big\}, & \text{if } D = r', \\ \dfrac{1}{D}\Big\{\displaystyle\sum_{i=1}^{D} X_{i:n} + (n-D)T_1\Big\}, & \text{if } D = r, \ldots, n. \end{cases} \tag{8}$$

The survival function of $\hat{\theta}$ in (8) can be expressed in the form in (1) with $\mathcal{D} = \{1, \ldots, r-1, r', r, \ldots, n\}$.

Before proceeding to verify the three conditions of TML, we need to observe the following facts:

Fact 1. For any any $d_1 = 0, 1, \ldots, r-1$ and $x > 0$,

$$\mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 \geqslant r) = \sum_{d_2=r}^{n} \frac{\mathsf{P}_\theta(D_2 = d_2 | D_1 = d_1)}{\mathsf{P}_\theta(D_2 \geqslant r | D_1 = d_1)} \mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 = d_2)$$

is increasing in θ. This will be proved using TML as follows:

(M1.1) Conditional on $D_1 = d_1 \leqslant r-1$, $D_2 = d_2 \geqslant r$,

$$\begin{aligned} \hat{\theta} &= \frac{1}{r}\Big\{\sum_{i=1}^{d_1} X_{i:n} + \sum_{i=d_1+1}^{r} X_{i:n} + (n-r)X_{r:n}\Big\} \\ &= \frac{1}{r}\Big\{\sum_{i=1}^{d_1} X_{i:n} + (n-d_1)T_1 + \sum_{i=d_1+1}^{r} (X_{i:n} - T_1) + (n-r)(X_{r:n} - T_1)\Big\} \\ &\overset{\mathrm{d}}{=} \frac{1}{r}\Big\{\sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 + \sum_{i=1}^{r-d_1-1} W_{i:d_2-d_1} + (n-r+1)W_{r-d_1:d_2-d_1}\Big\}, \end{aligned} \tag{9}$$

where $Z_1, \ldots, Z_{d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)I(Z \leqslant T_1)$ and $W_1, \ldots, W_{d_2-d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)I(W \leqslant \Delta_2)$, independently of Z's. The sum of Z's in (9) is stochastically increasing in θ. Using Lemma B.1, we have the same to hold true for the sum of W's in (9) as well. By the independence of the two sums, we conclude that the conditional distribution of $\hat{\theta}$ is stochastically increasing in θ.

(M1.2) By Lemma B.2, the sum of W's in (9) is stochastically decreasing in d_2.

(M1.3) This is a consequence of Lemma A.2(b).

Fact 2. For any $x > 0$,

$$\mathsf{P}_\theta(\hat{\theta} > x | D_1 \leqslant r-1, D_2 \geqslant r) = \sum_{d_1=0}^{r-1} \frac{\mathsf{P}_\theta(D_1 = d_1 | D_2 \geqslant r)}{\mathsf{P}_\theta(D_1 \leqslant r-1 | D_2 \geqslant r)} \mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 \geqslant r)$$

is increasing in θ. Once again, we will use TML to prove this result as follows:

(M2.1) This is exactly Fact 1.

(M2.2) For any $d_1 \leqslant r-2$, let $Z_1, \ldots, Z_{d_1+1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta) I(Z \leqslant T_1)$. Conditional on $D_1 = d_1$, $D_2 \geqslant r$, we have

$$\begin{aligned}
&\frac{1}{r}\left\{ \sum_{i=1}^{d_1} X_{i:n} + \sum_{i=d_1+1}^{r} X_{i:n} + (n-r+1) X_{r:n} \right\} \\
&\quad = \frac{1}{r}\left\{ \sum_{i=1}^{d_1} X_{i:n} + (n-d_1)T_1 + \sum_{i=d_1+1}^{r-1} (X_{i:n} - T_1) + (n-r+1)(X_{r:n} - T_1) \right\} \\
&\quad \overset{\text{d}}{=} \frac{1}{r}\left\{ \sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 + \sum_{i=1}^{r-1-d_1} W_{i:n-d_1} + (n-r+1) W_{r-d_1:n-d_1} \right\},
\end{aligned}$$

where $W_1, \ldots, W_{n-d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$ but conditional on the event that at least $r - d_1$ of them are less than Δ_2 and are independent of Z's. Similarly, conditional on $D_1 = d_1 + 1$, $D_2 \geqslant r$, the MLE has the same distribution as

$$\frac{1}{r}\left\{ \sum_{i=1}^{d_1+1} Z_i + (n-d_1-1)T_1 + \sum_{i=1}^{r-2-d_1} W'_{i:n-d_1-1} + (n-r+1) W'_{r-d_1-1:n-d_1-1} \right\},$$

where $W'_1, \ldots, W'_{n-d_1-1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$ but conditional on the event that at least $r - d_1 - 1$ of them are less than Δ_2 and are independent of Z's. Now,

$$\left\{ \sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 \right\} - \left\{ \sum_{i=1}^{d_1+1} Z_i + (n-d_1-1)T_1 \right\} = T_1 - Z_{d+1} \geqslant 0.$$

Moreover, using arguments similar to those in Lemma B.3, we can prove that the sum of W's is stochastically larger than the sum of W''s. Indeed, conditional on $W_{1:n-d_1} = x$ $(\leqslant \Delta_2)$, $(W_{2:n-d_1}, \ldots, W_{n-d_1:n-d_1})$ have the same distribution as the order statistics in a sample of size $n - d_1 - 1$ from $\mathcal{E}(\theta)I(W > x)$ but conditional further on the event that at least $r - d_1 - 1$ of them are less han Δ_2. The rest of the proof is similar. Thus, the conditional distribution of the MLE given $D_1 = d_1$, $D_2 \geqslant r$, is stochastically decreasing in d_1.

(M2.3) This is Lemma A.2(d).

Fact 3. For any $x, \theta > 0$,

$$\mathsf{P}_\theta(\hat{\theta} > x | D_1 \leqslant r-1, D_2 \geqslant r) < \mathsf{P}_\theta(\hat{\theta} > x | D_2 = r-1).$$

Here, we use once more TML but with a slight variation, where the events $\{D_2 \geqslant r\}$ and $\{D_2 = r-1\}$ play the roles of θ and θ', respectively. Before proceeding, note that

$$\mathsf{P}_\theta(\hat{\theta} > x | D_2 = r-1) = \sum_{d_1=0}^{r-1} \mathsf{P}_\theta(D_1 = d_1 | D_2 = r-1) \mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 = r-1).$$

(M3.1) We want to show that for all $d_1 = 0, 1, \dots, r-1$ and $x, \theta > 0$,

$$\mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 \geqslant r) < \mathsf{P}_\theta(\hat{\theta} > x | D_1 = d_1, D_2 = r-1).$$

For any $d_1 = 0, 1, \dots, r-1$, conditional on $D_1 = d_1$, $D_2 = r-1$, we have

$$\begin{aligned}
&\frac{1}{r-1}\left\{\sum_{i=1}^{d_1} X_{i:n} + \sum_{i=d_1+1}^{r-1} X_{i:n} + (n-r+1)T_2\right\} \\
&\quad = \frac{1}{r-1}\left\{\sum_{i=1}^{d_1} X_{i:n} + (n-d_1)T_1 + \sum_{i=d_1+1}^{r-1} (X_{i:n} - T_1) + (n-r+1)\Delta_2\right\} \\
&\quad \overset{\mathrm{d}}{=} \frac{1}{r-1}\left\{\sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 + \sum_{i=1}^{r-1-d_1} W_{i:r-1-d_1} + (n-r+1)\Delta_2\right\}, \qquad (10)
\end{aligned}$$

where $Z_1, \dots, Z_{d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)I(Z \leqslant T_1)$ and $W_1, \dots, W_{r-1-d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)I(W \leqslant \Delta_2)$ independently of Z's. On the other hand, conditional on $D_1 = d_1$, $D_2 \geqslant r$, we have

$$\begin{aligned}
&\frac{1}{r}\left\{\sum_{i=1}^{d_1} X_{i:n} + \sum_{i=d_1+1}^{r} X_{i:n} + (n-r+1)X_{r:n}\right\} \\
&\quad = \frac{1}{r}\left\{\sum_{i=1}^{d_1} X_{i:n} + (n-d_1)T_1 + \sum_{i=d_1+1}^{r-1} (X_{i:n} - T_1) + (n-r+1)(X_{r:n} - T_1)\right\} \\
&\quad \overset{\mathrm{d}}{=} \frac{1}{r}\left\{\sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 + \sum_{i=1}^{r-1-d_1} W_{i:n-d_1} + (n-r+1)W_{r-d_1:n-d_1}\right\},
\end{aligned}$$

where the Z's are as before and $W_1, \ldots, W_{n-d_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$ but conditional on the event that at least $r - d_1$ of them are less than Δ_2. This implies immediately that $W_{r-d_1:n-d_1} \leqslant \Delta_2$ and that the MLE is stochastically smaller than

$$\frac{1}{r-1}\left\{\sum_{i=1}^{d_1} Z_i + (n-d_1)T_1 + \sum_{i=1}^{r-1-d_1} W_{i:n-d_1} + (n-r+1)\Delta_2\right\}. \tag{11}$$

Observe that (10) and (11) differ only in the sum of W's which in both cases are independent of the sum of Z's. Therefore, we will complete the proof if we show that $\sum_{i=1}^{r-1-d_1} W_{i:n-d_1} \leqslant_{\text{st}} \sum_{i=1}^{r-1-d_1} W_{i:r-1-d_1}$. Since D_2 ranges from r to n, $\sum_{i=1}^{r-1-d_1} W_{i:n-d_1}$ has a mixture of distributions; conditional on $D_2 = d_2$, it has the same distribution as $\sum_{i=1}^{r-1-d_1} W_{i:d_2-d_1}$. By Lemma B.2, these distributions are stochastically ordered, the stochastically greatest of which corresponding to $d_2 = r$. Thus, $\sum_{i=1}^{r-1-d_1} W_{i:n-d_1} \leqslant_{\text{st}} \sum_{i=1}^{r-1-d_1} W_{i:r-d_1}$. Further, the latter sum is stochastically smaller than $\sum_{i=1}^{r-1-d_1} W_{i:r-1-d_1}$ and this completes the proof of (M3.1).

(M3.2) Next, we want to show that for any $d_1 = 0, 1, \ldots, r-2$ and $x, \theta > 0$,

$$\mathsf{P}_\theta(\hat\theta | D_1 = d_1 + 1, D_2 \geqslant r) < \mathsf{P}_\theta(\hat\theta | D_1 = d_1, D_2 \geqslant r).$$

But, this has been already proved in (M2.2).

(M3.3) We need to show that $\mathsf{P}_\theta(D_1 = d_1 | D_2 \geqslant r)/\mathsf{P}_\theta(D_1 = d_1 | D_2 = r-1)$ is increasing in $d_1 \in \{0, 1, \ldots, r-1\}$. But this is exactly Lemma A.2(c).

We are now ready to apply TML for proving the stochastic monotonicity of the MLE in (8).

(M1) For $d \neq r'$, the conditional distribution of $\hat\theta$, given $D = d$, is similar to that in Type-I censoring. For $d = r'$, it is Fact 2.

(M2) Except for the cases $d = r - 1$ and r', all other cases are similar to Type-I censoring. For $d = r - 1$, it is Fact 3. Now, we have to show that

$$\mathsf{P}_\theta(\hat\theta > x | D_1 \leqslant r-1, D_2 \geqslant r) > \mathsf{P}_\theta(\hat\theta > x | D_1 = r).$$

The conditional distribution of $\hat\theta$, given $D_1 = r$, is the same as of

$$\frac{1}{r}\left\{\sum_{i=1}^{r} Z_i + (n-r)T_1\right\}, \tag{12}$$

where $Z_1, \ldots, Z_r \overset{\text{iid}}{\sim} \mathcal{E}(\theta) I(Z \leqslant T_1)$. On the other hand, the conditional distribution of $\hat\theta$, given $D_1 \leqslant r-1$, $D_2 \geqslant r$, can be written as a mixture of distributions as D_1 ranges from 0 to $r - 1$. These are the same distributions encountered in (M2.2) wherein we proved that they are stochastically decreasing in d_1, the stochastically smallest arising when $d_1 = r-1$. Hence, $\hat{\theta}|(D_1 = r-1, D_2 \geqslant r) \leqslant_{\rm st} \hat{\theta}|(D_1 \leqslant r-1, D_2 \geqslant r)$. Conditional on $D_1 = r-1$, $D_2 \geqslant r$, $\hat{\theta}$ has the same distribution as

$$\frac{1}{r}\left\{\sum_{i=1}^{r-1} Z_i + (n-r+1)T_1 + (n-r+1)W_{1:n-r+1}\right\}, \tag{13}$$

where $W_{1:n-r+1}$ is the minimum in a sample of size $n-r+1$ from $\mathcal{E}(\theta)$ but conditional on the event that at least one observation is less than Δ_2. The difference between (13) and (12) is proportional to $T_1 - Z_r + (n-r+1)W_{1:n-r+1} > 0$, and this implies the result.

(M3) Since $\mathsf{P}_\theta(D_1 = d_1)/\mathsf{P}_{\theta'}(D_1 = d_1)$ and $\mathsf{P}_\theta(D_2 = d_2)/\mathsf{P}_{\theta'}(D_2 = d_2)$ are both strictly increasing functions for $\theta < \theta'$, it turns out that $\mathsf{P}_\theta(D = d)/\mathsf{P}_{\theta'}(D = d)$ is strictly increasing in $\{1, \ldots, r-1\}$ and $\{r, \ldots, n\}$. Moreover, in Lemma A.2(e), it is shown that

$$\frac{\mathsf{P}_\theta(D_2 = r-1)}{\mathsf{P}_{\theta'}(D_2 = r-1)} \leqslant \frac{\mathsf{P}_\theta(D_1 \leqslant r-1, D_2 \geqslant r)}{\mathsf{P}_{\theta'}(D_1 \leqslant r-1, D_2 \geqslant r)} \leqslant \frac{\mathsf{P}_\theta(D_1 = r)}{\mathsf{P}_{\theta'}(D_1 = r)}$$

as required.

Thus follows the stochastic monotonicity of the survival function of the MLE $\hat{\theta}$ in (8) in the case of generalized Type-II hybrid censoring.

6 Discussion

In this paper, we have presented a lemma which is very useful in establishing the stochastic monotonicity of an estimator in situations wherein its distribution can be expressed as a mixture. By checking the three monotonicities described in this lemma, we were able to present a formal proof for the stochastic monotonicity of the MLE of an exponential mean under different types of censored data. In the case of Type-I hybrid censoring, this monotonicity was in question for nearly two decades since the work of Chen and Bhattacharyya (1988).

Clearly, TML can also be useful outside the censoring context whenever a mixture distribution has the required monotonicities. We shall now present such an example.

Kundu and Basu (2000) considered the following model. Let (X_{1i}, X_{2i}), $i = 1, \ldots, n$, be independent random vectors consisting of independent components such that for $j = 1, 2$ and $i = 1, \ldots, n$, $X_{ji} \sim \mathcal{E}(\theta_j)$. Further, let $X_i = \min\{X_{1i}, X_{2i}\}$ and δ_i be an indicator of whether $X_{1i} < X_{2i}$ or $X_{1i} \geqslant X_{2i}$. Such data arise when n individuals are exposed to two competing risks, so that X_i represents the failure time of the i-th individual and δ_i indicates its cause of failure. For a known fixed $m < n$, the observed data are $(X_1, \delta_1), \ldots, (X_m, \delta_m), (X_{m+1}, *), \ldots, (X_n, *)$. Here, a "$*$" means that the corresponding indicator δ is unobserved, and so there are $n - m$ unallocated failures.

Let D be the number of failures due to cause 1. Then, the MLE of θ_1 is

$$\hat{\theta}_1 = \frac{m\sum_{i=1}^n X_i}{nD},$$

provided $D \geqslant 1$. If no failures due to Cause 1 occurred, then the MLE of θ_1 does not exist. Here, D is a binomial $\mathcal{B}(m,p)$ random variable, where $p = \theta_2/(\theta_1+\theta_2)$, but is restricted to be at least 1. Kundu and Basu (2000) conjectured that, for fixed θ_2, $\hat{\theta}_1$ is stochastically increasing in θ_1, but they could not provide a mathematical proof. However, this result can be easily proved by using TML as follows:

(M1) Under the above assumptions, $\sum_{i=1}^n X_i \sim \mathcal{G}(n,\beta)$, where $\beta = \theta_1\theta_2/(\theta_1+\theta_2)$. Hence, the conditional distribution of $\hat{\theta}_1$, given $D = d$, follows a $\mathcal{G}(n, m\beta/(nd))$ distribution. Since the scale parameter of this gamma distribution is increasing in θ_1, the result follows immediately.

(M2) Similarly, the result follows by observing that the scale parameter of $\mathcal{G}(n, m\beta/(nd))$ is decreasing in d.

(M3) The probability of success p of the binomial distribution of D is strictly decreasing in θ_1 and this implies that D is stochastically decreasing in θ_1.

Hence, the required monotonicity result for $\hat{\theta}_1$ follows immediately.

Appendix

A Distribution of the number of failures

Let $X_1,\ldots,X_r,X_{r+1},\ldots,X_n \overset{\text{iid}}{\sim} \mathcal{E}(\theta)$, $\theta > 0$, where $1 \leqslant r \leqslant n$. Let T_1 and T_2 be some fixed constants with $0 = T_0 < T_1 < T_2$ and $\Delta_j = T_j - T_{j-1}$, $j = 1,2$. Define $N_j = \#\{X\text{'s} \in (T_{j-1},T_j]\}$ and $D_j = \#\{X\text{'s} \leqslant T_j\}$, $j = 1,2$. Clearly, $(D_1,D_2) = (N_1, N_1+N_2)$. Then, the following hold true.

Lemma A.1. *(a) The probability mass function (pmf) of* (N_1,N_2) *is*

$$\mathsf{P}_\theta(N_1 = n_1, N_2 = n_2) = \frac{n!}{n_1!n_2!(n-n_1-n_2)!}\times$$
$$(1-e^{-\Delta_1/\theta})^{n_1}e^{-(n-n_1)\Delta_1/\theta}(1-e^{-\Delta_2/\theta})^{n_2}e^{-(n-n_1-n_2)\Delta_2/\theta}, \quad 0 \leqslant n_1,n_2,n_1+n_2 \leqslant n.$$

(b) The probability mass function (pmf) of (D_1,D_2) *is*

$$\mathsf{P}_\theta(D_1 = d_1, D_2 = d_2) = \frac{n!}{d_1!(d_2-d_1)!(n-d_2)!}\times$$
$$(1-e^{-\Delta_1/\theta})^{d_1}e^{-(n-d_1)\Delta_1/\theta}(1-e^{-\Delta_2/\theta})^{d_2-d_1}e^{-(n-d_2)\Delta_2/\theta}, \quad 0 \leqslant d_1 \leqslant d_2 \leqslant n.$$

(c) The marginal distribution of D_j *is binomial* $\mathcal{B}(n, 1-e^{-T_j/\theta})$, $j = 1,2$.

Proof. The proofs are straightforward. □

Next, we provide some statistical properties of the distribution of (D_1, D_2).

Lemma A.2. *(a) The distribution of D_j has the monotone likelihood ratio property with respect to θ, i.e., the ratio $\frac{\mathsf{P}_\theta(D_j=d)}{\mathsf{P}_{\theta'}(D_j=d)}$ is strictly increasing in d for any $\theta < \theta'$. The result does not change even if D_j is restricted in some subset of $\{0, 1, \ldots, n\}$.*

(b) For any $d_1 \in \{0, 1, \ldots, n\}$, the conditional distribution of D_2, given $D_1 = d_1$, has the monotone likelihood ratio property with respect to θ, i.e., the ratio $\frac{\mathsf{P}_\theta(D_2=d_2|D_1=d_1)}{\mathsf{P}_{\theta'}(D_2=d_2|D_1=d_1)}$ is strictly increasing in d_2 for any $\theta < \theta'$. The result does not change even if D_2 is restricted in some subset of $\{0, 1, \ldots, n\}$.

(c) For any fixed $r \in \{1, \ldots, n\}$ and $\theta > 0$, the ratio $\frac{\mathsf{P}_\theta(D_1=d_1|D_2\geqslant r)}{\mathsf{P}_\theta(D_1=d_1|D_2=r-1)}$ is strictly increasing in $d_1 \in \{1, \ldots, r-1\}$.

(d) For any fixed $r \in \{1, \ldots, n\}$ and $\theta < \theta'$, the ratio $\frac{\mathsf{P}_\theta(D_1=d_1|D_2\geqslant r)}{\mathsf{P}_{\theta'}(D_1=d_1|D_2\geqslant r)}$ is strictly increasing in $d_1 \in \{1, \ldots, r\}$.

(e) For any fixed $r \in \{1, \ldots, n\}$ and $\theta < \theta'$, we have

$$\frac{\mathsf{P}_\theta(D_2 = r-1)}{\mathsf{P}_{\theta'}(D_2 = r-1)} \leqslant \frac{\mathsf{P}_\theta(D_1 \leqslant r-1, D_2 \geqslant r)}{\mathsf{P}_{\theta'}(D_1 \leqslant r-1, D_2 \geqslant r)} \leqslant \frac{\mathsf{P}_\theta(D_1 = r)}{\mathsf{P}_{\theta'}(D_1 = r)}.$$

Proof. (a) Let $\mathcal{D}^*$ be any subset of $\{0, 1, \ldots, n\}$. Then,

$$\frac{\mathsf{P}_\theta(D_j = d|D_j \in \mathcal{D}^*)}{\mathsf{P}_{\theta'}(D_j = d|D_j \in \mathcal{D}^*)} \propto \frac{\binom{n}{d}(1-e^{-T_j/\theta})^d e^{-(n-d)T_j/\theta}}{\binom{n}{d}(1-e^{-T_j/\theta'})^d e^{-(n-d)T_j/\theta'}} \propto \left(\frac{e^{T_j/\theta}-1}{e^{T_j/\theta'}-1}\right)^d,$$

which is strictly increasing in d, since $(e^{T_j/\theta}-1)/(e^{T_j/\theta'}-1) > 1$.
(b) Similar to (a), for $d_2 \in \{d_1, \ldots, n\} \cap \mathcal{D}^*$,

$$\frac{\mathsf{P}_\theta(D_2 = d_2|D_1 = d_1, D_2 \in \mathcal{D}^*)}{\mathsf{P}_\theta(D_2 = d_2|D_1 = d_1, D_2 \in \mathcal{D}^*)} \propto \frac{\mathsf{P}_\theta(D_1 = d_1, D_2 = d_2)}{\mathsf{P}_{\theta'}(D_1 = d_1, D_2 = d_2)} \propto \left(\frac{e^{\Delta_2/\theta}-1}{e^{\Delta_2/\theta'}-1}\right)^{d_2}.$$

(c) For some positive constants C_1 and C_2 that do not depend on d_1, we have

$$\begin{aligned}
\frac{\mathsf{P}_\theta(D_1=d_1+1|D_2\geqslant r)}{\mathsf{P}_\theta(D_1=d_1+1|D_2=r-1)} &= C_1\sum_{d_2=r}^{n}\frac{\mathsf{P}_\theta(D_1=d_1+1,D_2=d_2)}{\mathsf{P}_\theta(D_1=d_1+1,D_2=r-1)}\\
&= C_2\sum_{d_2=r}^{n}\frac{(r-1-d_1-1)!(e^{\Delta_2/\theta}-1)^{d_2}}{(d_2-d_1-1)!(n-d_2)!}\\
&= C_2\sum_{d_2=r}^{n}\frac{d_2-d_1}{r-1-d_1}\times\frac{(r-1-d_1)!(e^{\Delta_2/\theta}-1)^{d_2}}{(d_2-d_1)!(n-d_2)!}\\
&> C_2\sum_{d_2=r}^{n}\frac{(r-1-d_1)!(e^{\Delta_2/\theta}-1)^{d_2}}{(d_2-d_1)!(n-d_2)!}\\
&= \frac{\mathsf{P}_\theta(D_1=d_1|D_2\geqslant r)}{\mathsf{P}_\theta(D_1=d_1|D_2=r-1)}.
\end{aligned}$$

(d) Observe that

$$\begin{aligned}
&\mathsf{P}_\theta(D_1=d_1+1,D_2\geqslant r)\\
&=\sum_{d_2=r}^{n}\frac{n!}{(d_1+1)!(d_2-d_1-1)!(n-d_2)!}\\
&\quad\times(1-e^{-\Delta_1/\theta})^{d_1+1}e^{-(n-d_1-1)\Delta_1/\theta}(1-e^{-\Delta_2/\theta})^{d_2-d_1-1}e^{-(n-d_2)\Delta_2/\theta}\\
&=\frac{e^{\Delta_1/\theta}-1}{(d_1+1)(1-e^{-\Delta_2/\theta})}\sum_{d_2=r}^{n}(d_2-d_1)\frac{n!}{d_1!(d_2-d_1)!(n-d_2)!}\\
&\quad\times(1-e^{-\Delta_1/\theta})^{d_1}e^{-(n-d_1)\Delta_1/\theta}(1-e^{-\Delta_2/\theta})^{d_2-d_1}e^{-(n-d_2)\Delta_2/\theta}\\
&=\frac{e^{\Delta_1/\theta}-1}{(d_1+1)(1-e^{-\Delta_2/\theta})}\\
&\quad\times\left\{\sum_{d_2=r}^{n}d_2\mathsf{P}_\theta(D_1=d_1,D_2=d_2)-d_1\mathsf{P}_\theta(D_1=d_1,D_2\geqslant r)\right\}
\end{aligned}$$

and by Lemma C.1,

$$\frac{\mathsf{P}_\theta(D_1=d_1+1,D_2\geqslant r)}{\mathsf{P}_{\theta'}(D_1=d_1+1,D_2\geqslant r)}>\frac{\sum_{d_2=r}^{n}d_2\mathsf{P}_\theta(D_1=d_1,D_2=d_2)-d_1\mathsf{P}_\theta(D_1=d_1,D_2\geqslant r)}{\sum_{d_2=r}^{n}d_2\mathsf{P}_{\theta'}(D_1=d_1,D_2=d_2)-d_1\mathsf{P}_{\theta'}(D_1=d_1,D_2\geqslant r)}.$$

The right hand side of the above inequality is greater than or equal to $\frac{\mathsf{P}_\theta(D_1=d_1,D_2\geqslant r)}{\mathsf{P}_{\theta'}(D_1=d_1,D_2\geqslant r)}$ if and only if

$$\frac{\sum_{d_2=r}^{n}d_2\mathsf{P}_\theta(D_1=d_1,D_2=d_2)}{\mathsf{P}_\theta(D_1=d_1,D_2\geqslant r)}-\frac{\sum_{d_2=r}^{n}d_2\mathsf{P}_{\theta'}(D_1=d_1,D_2=d_2)}{\mathsf{P}_{\theta'}(D_1=d_1,D_2\geqslant r)}\geqslant 0.$$

However, the last difference equals $\mathsf{E}_\theta(D_2|D_1=d_1,D_2\geqslant r)-\mathsf{E}_{\theta'}(D_2|D_1=d_1,D_2\geqslant r)$ which is strictly positive by Part (b). Thus, the assertion is proved.

(e) The second inequality arises from (d), since $\mathsf{P}_\theta(D_1 = r) = \mathsf{P}_\theta(D_1 = r, D_2 \geqslant r)$. In order to prove the first inequality, we will first show that

$$\frac{\mathsf{P}_\theta(D_2 = r-1)}{\mathsf{P}_{\theta'}(D_2 = r-1)} \leqslant \frac{\mathsf{P}_\theta(D_1 \leqslant r-1, D_2 = d_2)}{\mathsf{P}_{\theta'}(D_1 \leqslant r-1, D_2 = d_2)} \tag{14}$$

for any $d_2 \geqslant r$. This, in conjuction with Lemma C.2, will give the result. Observe that

$$\begin{aligned}
&\frac{\mathsf{P}_\theta(D_1 \leqslant r-1, D_2 = d_2)}{\mathsf{P}_\theta(D_2 = r-1)} \\
&= \frac{\sum_{d_1=0}^{r-1} \frac{n!}{d_1!(d_2-d_1)!(n-d_2)!}(1-e^{-\Delta_1/\theta})^{d_1} e^{-(n-d_1)\Delta_1/\theta}(1-e^{-\Delta_2/\theta})^{d_2-d_1} e^{-(n-d_2)\Delta_2/\theta}}{\frac{n!}{(r-1)!(n-r+1)!}(1-e^{-T_2/\theta})^{r-1} e^{-(n-r+1)T_2/\theta}} \\
&= (e^{\Delta_2/\theta}-1)^{d_2-r+1}\frac{(n-r+1)!}{(n-d_2)!} \\
&\qquad \times \sum_{d_1=0}^{r-1} \frac{(r-1-d_1)!}{(d_2-d_1)!}\binom{r-1}{d_1}\left(\frac{1-e^{-T_1/\theta}}{1-e^{-T_2/\theta}}\right)^{d_1}\left(1-\frac{1-e^{-T_1/\theta}}{1-e^{-T_2/\theta}}\right)^{r-1-d_1} \\
&= (e^{\Delta_2/\theta}-1)^{d_2-r+1}\frac{(n-r+1)!}{(n-d_2)!}\mathsf{E}_\theta\left\{\frac{(r-1-Y)!}{(d_2-Y)!}\right\},
\end{aligned}$$

where $Y \sim \mathcal{B}\big(r-1, \frac{1-e^{-T_1/\theta}}{1-e^{-T_2/\theta}}\big)$. It is easy to show that the probability of success of this distribution is strictly decreasing in θ. Hence, Y is stochastically decreasing in θ. Moreover,

$$\frac{(r-1-(y+1))!}{(d_2-(y+1))!} = \frac{(r-1-y)!}{(d_2-y)!} \times \frac{d_2-y}{r-1-y} > \frac{(r-1-y)!}{(d_2-y)!}$$

for $d_2 \geqslant r$, which means that $\frac{(r-1-y)!}{(d_2-y)!}$ is a strictly increasing function. Thus,

$$\mathsf{E}_\theta\left\{\frac{(r-1-Y)!}{(d_2-Y)!}\right\} > \mathsf{E}_{\theta'}\left\{\frac{(r-1-Y)!}{(d_2-Y)!}\right\}. \tag{15}$$

Now, for any $d_2 \geqslant r$ and $\theta < \theta'$, we have

$$(e^{\Delta_2/\theta}-1)^{d_2-r+1} > (e^{\Delta_2/\theta'}-1)^{d_2-r+1}. \tag{16}$$

Since (15) and (16) imply (14), the inequality is proved. □

B Some results on order statistics

Lemma B.1. *Let X, Y be absolutely continuous random variables with $X \leqslant_{\text{st}} Y$. For any fixed integer n, let $X_1, \dots, X_n$ be independent copies of X and $Y_1, \dots, Y_n$ be independent copies of Y. Then, for any $(w_0, w_1, \dots, w_n) \in \mathbb{R} \times [0,\infty)^n$, we have $w_0 + \sum_{i=1}^n w_i X_{i:n} \leqslant_{\text{st}} w_0 + \sum_{i=1}^n w_i Y_{i:n}$.*

Proof. The assertion is a consequence of the fact that $X \leqslant_{\rm st} Y$ implies $(X_{1:n}, \ldots, X_{n:n}) \leqslant_{\rm st} (Y_{1:n}, \ldots, Y_{n:n})$; see Belzunce *et al.* (2005). □

Lemma B.2. *Let $X_1, X_2, \ldots$ be iid from an absolutely continuous distribution. Then, for any $1 \leqslant r \leqslant n$, and $(w_1, \ldots, w_r) \in [0, \infty)^r$, $S_n = \sum_{i=1}^{r} w_i X_{i:n}$ is stochastically decreasing in n.*

Proof. This follows from the fact that for all $n \geqslant 1$ it holds $(X_{1:n+1}, \ldots, X_{n:n+1}) \leqslant_{\rm st} (X_{1:n}, \ldots, X_{n:n})$; see Zhuang and Hu (2007). □

Lemma B.3. *Let $X_1, X_2, \ldots$ be iid random variables from an absolutely continuous distribution on a subset of non-negative reals. For any $1 \leqslant r \leqslant n$ and $(w_1, \ldots, w_r) \in (0, \infty)^r$, let $S_1 = \sum_{i=1}^{r-1} w_{i+1} X_{i:n-1}$ and $S_2 = \sum_{i=1}^{r} w_i X_{i:n}$. Then, $S_1 \leqslant_{\rm st} S_2$.*

Proof. Let x be any point in the support of the distribution. Conditional on $X_{1:n} = x$, $(X_{2:n}, \ldots,\ X_{r:n}, \ldots, X_{n:n})$ has the same distribution as the order statistics in sample of size $n-1$ from the underlying distribution but left truncated at x. Denote by Y a random variable from this left truncated distribution and recall that $X \leqslant_{\rm st} Y$. By Lemma B.1, we have $\sum_{i=1}^{r-1} w_{i+1} X_{i:n-1} \leqslant_{\rm st} \sum_{i=1}^{r-1} w_{i+1} Y_{i:n-1}$. Clearly, the latter is smaller than $w_1 x + \sum_{i=1}^{r-1} w_{i+1} Y_{i:n-1}$ which has exactly the conditional distribution of S_2, given $X_{1:n} = x$. Thus, for any integrable increasing function h, $\mathsf{E}\{h(S_1)\} \leqslant \mathsf{E}\{h(S_2) | X_{1:n} = x\}$. Since this inequality is true for all x, we have $\mathsf{E}\{h(S_1)\} \leqslant \mathsf{E}\{h(S_2)\}$, and the required result follows. □

Lemma B.4. *Let $X_1, \ldots, X_n$ be iid from some absolutely continuous distribution with pdf f and cdf F. For some fixed real T, let $D = \#\{X\text{'s} \leqslant T\}$. Then, conditional on $D = d$, the random vectors $(X_{1:n}, \ldots, X_{d:n})$ and $(X_{d+1:n}, \ldots, X_{n:n})$ are independent. Moreover, conditional on $D = d$,*

$$
\begin{aligned}
(X_{1:n}, \ldots, X_{d:n}) &\stackrel{\rm d}{=} (U_{1:d}, \ldots, U_{d:d}), \\
(X_{d+1:n}, \ldots, X_{n:n}) &\stackrel{\rm d}{=} (V_{1:n-d}, \ldots, V_{n-d:n-d}),
\end{aligned}
$$

where $U_1, \ldots, U_d \stackrel{\rm iid}{\sim} f(x) I(x \leqslant T)$ and $V_1, \ldots, V_{n-d} \stackrel{\rm iid}{\sim} f(x) I(x > T)$.

Proof. The conditional joint density of the ordered sample is

$$\begin{aligned} f(x_1,\ldots,x_n|D=d) &= \frac{n!\prod_{i=1}^n f(x_i)}{\mathsf{P}(D=d)} I(x_1<\cdots<x_d\leqslant T<x_{d+1}<\cdots<x_n) \\ &= \frac{n!\prod_{i=1}^n f(x_i)}{\frac{n!}{d!(n-d)!}F(T)^d\{1-F(T)\}^{n-d}} I(x_1<\cdots<x_d\leqslant T<x_{d+1}<\cdots<x_n) \\ &= \left\{d!\prod_{i=1}^d \frac{f(x_i)}{F(T)} I(x_1<\cdots<x_d\leqslant T)\right\} \\ &\quad\times\left\{(n-d)!\prod_{i=d+1}^n \frac{f(x_i)}{1-F(T)} I(T<x_{d+1}<\cdots<x_n)\right\} \end{aligned}$$

which proves the required result. □

C Two useful lemmas

Lemma C.1. *For any $a,b>0$, the function $h(x)=\frac{1-e^{-bx}}{e^{ax}-1}$ is strictly decreasing in $(0,\infty)$.*

Proof. After some algebra, we get the derivative of $h(x)$ to be

$$h'(x)=\frac{1-e^{-(a+b)x}}{x(1-e^{-ax})}\left\{\frac{(a+b)x}{e^{(a+b)x}-1}-\frac{ax}{e^{ax}-1}\right\}.$$

Since $\frac{y}{e^y-1}$ is strictly decreasing in $y>0$, the result follows. □

Lemma C.2. *Let $a,a_1,\ldots,a_m$, $b,b_1,\ldots,b_m$ be positive real numbers such that $a_i/b_i\leqslant a/b$, $i=1,\ldots,m$. Then, $\sum_{i=1}^m a_i/\sum_{i=1}^m b_i\leqslant a/b$.*

Proof. The proof is straightforward. □

References

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, John Wiley & Sons, New York.

Balakrishnan, N., Brain, C. and Mi, J. (2002). Stochastic order and mle of the mean of the exponential distribution, *Methodology and Computing in Applied Probability*, **4**, 83–93.

Bartholomew, D. J. (1963). The sampling distribution of an estimate arising in life testing, *Technometrics*, **5**, 361–374.

Belzunce, F., Mercader, J. A. and Ruiz, J. M. (2005). Stochastic comparisons of generalized order statistics, *Probability in the Engineering and Informational Sciences*, **19**, 99–120.

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

Chandrasekar, B., Childs, A. and 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.

Chen, S. and Bhattacharyya, G. K. (1988). Exact confidence bounds for an exponential parameter under hybrid censoring, *Communications in Statistics – Theory and Methods*, **17**, 1857–1870.

Childs, A., Chandrasekar, B., Balakrisnan N. and 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.

Kundu, D. and Basu, S. (2000). Analysis of incomplete data in presence of competing risks, *Journal of Statistical Planning and Inference*, **87**, 221–239.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*, Springer-Verlag, New York.

Spurrier, J. D. and Wei, L. J. (1980). A test of the parameter of the exponential distribution in the Type-I censoring case, *Journal of the American Statistical Association*, **75**, 405–409.

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