Stochastic monotonicity of the MLEs of parameters in exponential simple step-stress models under Type-I and Type-II censoring

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

Abstract

In two recent papers by Balakrishnan, Kundu, Ng and Kannan (*Journal of Quality Technology*, 2007) and Balakrishnan, Xie and Kundu (*Annals of the Institute of Statistical Mathematics*, 2009), the maximum likelihood estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of the parameters θ_1 and θ_2 have been derived in the framework of exponential simple step-stress models under Type-II and Type-I censoring, respectively. Here, we prove that these estimators are stochastically monotone with respect to θ_1 and θ_2, respectively, which has been conjectured in these papers and then utilized to develop exact conditional inference for the parameters θ_1 and θ_2. For proving these results, we have established a multivariate stochastic ordering of a particular family of trinomial distributions under truncation, which is also of independent interest.

Keywords: Exponential distribution; maximum likelihood estimation; step-stress models; Type-II censoring; Type-I censoring; exact confidence intervals; trinomial distribution; multivariate stochastic ordering

1 Introduction

An estimator $\hat{\theta}$ of a scalar parameter θ is said to be *stochastically increasing* in θ if its survival function $\mathsf{P}_\theta(\hat{\theta} > x)$ is an increasing function of θ for any fixed x. This property intuitively means that for larger values of θ we will tend to observe larger values for $\hat{\theta}$. However, besides intuition, the stochastic increasingness of $\hat{\theta}$ with respect to θ also provides a straightforward method of constructing confidence intervals for θ. The method, called *pivoting the cumulative distribution function (cdf)* or, equivalently, the survival function (cf. Casella and Berger, 2002, p. 432) proceeds as follows. Let $\hat{\theta}_{\text{obs}}$ denote the observed

[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`

value of $\hat{\theta}$. Choose α_1, α_2 satisfying $\alpha_1+\alpha_2 = \alpha$ (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. The existence and uniqueness of the solutions of these equations are guaranteed, of course, by the monotonicity of $\mathsf{P}_\theta(\hat{\theta} > \hat{\theta}_{\text{obs}})$ with respect to θ. If we 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 a $100(1-\alpha)\%$ confidence interval for θ.

In the literature, a series of papers appeared discussing the construction of confidence intervals for the parameters of interest relying on the assumed stochastic monotonicity of the corresponding maximum likelihood estimators (MLEs), but did not prove this property. Among these are Chen and Bhattacharyya (1988), Kundu and Basu (2000), Childs et al. (2003), and Chandrasekar et al. (2004) who derived the MLEs of the underlying parameter as well as their exact conditional distributions under different scenarios involving censored samples from an exponential distribution, numerically verified that these MLEs are stochastically increasing with respect to the parameter, and then assumed it to develop exact inference for the parameter. In all these cases, the survival function of the MLE has the mixture form

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

where $\mathcal{D}$ is a finite set. Balakrishnan and Iliopoulos (2009) recently established a lemma concerning the stochastic monotonicity of such mixtures, which proves the required monotonicity of the MLEs in all the above mentioned cases.

Along the lines of developments mentioned above in the case of exponential distribution under different forms of censored data, Balakrishnan et al. (2007, 2009) derived the MLEs of the parameters θ_1 and θ_2 of an exponential simple step-stress model and their exact conditional distributions under Type-II and Type-I censoring, respectively. Once again, being unable to formally prove the stochastic monotonicity of these MLEs and verifying it only through extensive numerical computations, these authors used the monotonicity to develop exact conditional inference for the parameters θ_1 and θ_2. In this paper, we prove formally the required stochastic monotonicity results, thus justifying the exact conditional inference developed in Balakrishnan et al. (2007, 2009).

The rest of the paper is organized as follows. In Section 2, we give a brief description of simple step-stess models and detail the results concerning exact inference for the exponential simple step-stress model under Type-I and Type-II censoring. In Section 3, we provide a slight generalization of the lemma proved by Balakrishnan and Iliopoulos (2009) so that it becomes applicable in the case of mixtures wherein the mixing distribution is multivariate. We then apply this generalized lemma to two exponential simple step-stress models. Specifically, in Section 4, we consider the situation of Type-II censoring discussed by Balakrishnan et al. (2007), while in Section 5 we consider the situation of Type-I censoring discussed by Balakrishnan et al. (2009). We conclude the main part of the paper in Section 6 with some final remarks. Finally, an appendix contains the primary work in developing these results, namely, establishing the stochastic monotonicity of a particular family of trinomial distributions under truncation, which by itself is of independent interest.

2 Step-stress accelerated life tests under censoring

Step-stress testing is a special case of an accelerated life testing experiment. Interested readers may refer to the books by Nelson (1990), Meeker and Escobar (1998), and Bagdonavicius and Nikulin (2002) for an elaborate treatment on accelerated life testing and associated inferential issues. Under such an experiment, n identical units are placed on a life test at an initial stress level l_0. The stress level is successively changed to $l_1, \cdots, l_m$ at some (possibly random) timepoints $0 < T_1 < \cdots < T_m$ and the successive failure times are recorded. The so-called *simple step-stress model*, corresponding to the case $m = 1$ in this set-up, has been studied extensively in the literature.

Sedyakin (1966) and Nelson (1990) have considered the *cumulative exposure model* which can be described as follows. Denoting by F_j the distribution function of the lifetimes at stress level l_{j-1}, $j = 1, 2$, the distribution function of the lifetimes under the simple step-stress model is given by

$$F(x) = \begin{cases} F_1(x), & 0 < x \leqslant T_1, \\ F_2(x - T_1 + T_1^*), & x > T_1, \end{cases}$$

where T_1^* is the solution to the equation $F_1(T_1) = F_2(T_1^*)$. Note that this guarantees the continuity of the distribution function at the point T_1.

Xiong (1998) considered an exponential simple step-stress model with Type-II censoring at the second level of stress. More specifically, let $X_1, \cdots, X_n$ be the lifetimes of the n identical units under test, $T_1 > 0$ a fixed time point, and $r \in \{2, \cdots, n\}$ a pre-fixed integer. The test starts at the stress level l_0 which is changed to l_1 at time T_1 and continues until the r-th failure is observed, at which time the test gets terminated. Xiong used a simple linear regression model for the logarithms of the exponential mean lifetimes and developed inference for its parameters. However, upon noting that the MLEs of the exponential mean lifetimes θ_1 and θ_2 exist only when the number of failures occurring at the first level (with the corresponding lifetimes being at most T_1) is at least 1 and at most $r-1$, Balakrishnan et al. (2007) developed the exact conditional distributions of the MLEs of the mean lifetimes and discussed exact as well as asymptotic inferential procedures and bootstrap methods. Inference for this model has also been studied beyond the context of exponential distribution; for example, Kateri and Balakrishnan (2008) recently discussed the case of Weibull lifetimes.

Balakrishnan et al. (2009) considered a variation of the above model under time constraint, i.e., involving Type-I censoring rather than Type-II censoring at the second level of stress. More specifically, they fixed another time point $T_2 > T_1$ at which the life test gets terminated instead of waiting until the r-th failure to occur. By assuming exponential lifetimes once again, they developed exact inferential procedures for the model parameters.

In both these works, however, the validity of the exact inferential procedures relies on the stochastic monotonicity of the MLEs of the exponential means at the two stress levels which was only verified numerically by these authors and not proved formally. In the following sections, we establish these monotonicity results formally.

3 Preliminaries and the basic lemma

For any $\mathbf{x} = (x_1, \cdots, x_k)$ and $\mathbf{y} = (y_1, \cdots, y_k) \in \mathbb{R}^k$, we write $\mathbf{y} \geqslant \mathbf{x}$ if $y_i \geqslant x_i$ for all $i = 1, \cdots, k$. A set $U \subseteq \mathbb{R}^k$ is called an *upper set* if $\mathbf{x} \in U$ and $\mathbf{y} \geqslant \mathbf{x}$ implies $\mathbf{y} \in U$. Two random vectors $\mathbf{X} = (X_1, \cdots, X_k)$ and $\mathbf{Y} = (Y_1, \cdots, Y_k)$ are ordered in the usual multivariate stochastic order, denoted by $\mathbf{X} \leqslant_{\text{st}} \mathbf{Y}$, if for any measurable upper set U, we have $\mathsf{P}(\mathbf{X} \in U) \leqslant \mathsf{P}(\mathbf{Y} \in U)$. An upper orthant is a special case of upper set, containing all $\mathbf{x} \in \mathbb{R}^k$ such that $\mathbf{x} \geqslant \mathbf{a}$ for some $\mathbf{a} \in \mathbb{R}^k$. We will use the notation $O(\mathbf{a}) \equiv O(a_1, \cdots, a_k)$ to denote the upper orthant with minimum point $\mathbf{a} = (a_1, \cdots, a_k)$. A random vector $\mathbf{X}$ is said to be stochastically smaller in the upper orthant ordering than another random vector $\mathbf{Y}$ if $\mathsf{P}(\mathbf{X} \in O) \leqslant \mathsf{P}(\mathbf{Y} \in O)$ for all upper orthants O. Clearly, the latter is weaker than the usual multivariate stochastic ordering.

The usual multivariate stochastic order is characterized by the following: $\mathbf{X} \leqslant_{\text{st}} \mathbf{Y}$

is equivalent to $\mathsf{E}\{\phi(\mathbf{X})\} \leqslant \mathsf{E}\{\phi(\mathbf{Y})\}$ for any coordinatewise increasing function $\phi = \phi(x_1, \cdots, x_k)$. Obviously, this is equivalent to $\mathsf{E}\{\phi(\mathbf{X})\} \geqslant \mathsf{E}\{\phi(\mathbf{Y})\}$ for any coordinatewise decreasing function ϕ. Note also that $\mathbf{X} \leqslant_{\text{st}} \mathbf{Y}$ if and only if there exist random vectors $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$ defined in the same probability space such that $\hat{\mathbf{X}} \overset{\text{d}}{=} \mathbf{X}$, $\hat{\mathbf{Y}} \overset{\text{d}}{=} \mathbf{Y}$ and $\hat{\mathbf{X}} \leqslant \hat{\mathbf{Y}}$ (a.s.).

Suppose now that the survival function of a particular estimator $\hat{\theta}$ of a scalar parameter θ has the form

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

where $\mathcal{D} \subset \mathbb{R}^k$. Balakrishnan and Iliopoulos (2009) proved a lemma, called *Three Monotonicities Lemma* (TML), which provides sufficient conditions for the stochastic monotonicity of $\hat{\theta}$ with respect to θ in the special case when $k = 1$. Here, we first extend this result to any $k \geqslant 1$.

Lemma 1. [Three Monotonicities Lemma – General Case] *Assume that the following hold true:*

(M1) For all $\mathbf{d} = (d_1, \cdots, d_k) \in \mathcal{D}$, the conditional distribution of $\hat{\theta}$, given $\mathbf{D} = \mathbf{d}$, is stochastically increasing in θ, i.e., the function $\mathsf{P}_\theta(\hat{\theta} > x | \mathbf{D} = \mathbf{d})$ is increasing in θ for all x and $\mathbf{d} \in \mathcal{D}$;

(M2) For all x and $\theta > 0$, the conditional distribution of $\hat{\theta}$, given $\mathbf{D} = \mathbf{d}$, is stochastically decreasing in $\mathbf{d}$, i.e., the function $\phi(\mathbf{d}) = \mathsf{P}_\theta(\hat{\theta} > x | \mathbf{D} = \mathbf{d})$ is decreasing in every d_i, $i = 1, \cdots, k$;

(M3) $\mathbf{D}$ is stochastically decreasing in θ, i.e., $\mathsf{E}_\theta\{\phi(\mathbf{D})\} \leqslant \mathsf{E}_{\theta'}\{\phi(\mathbf{D})\}$ when $\theta < \theta'$ for any coordinatewise decreasing function ϕ.

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

Proof. The proof follows exactly along the lines of TML established by Balakrishnan and Iliopoulos (2009) and is therefore omitted here. □

Due to this result, a proof of the stochastic monotonicity of $\hat{\theta}$ with respect to θ may be completed in three steps, viz., by establishing the three conditions of Lemma 1. Since the above lemma coincides with the original TML of Balakrishnan and Iliopoulos (2009) when $k = 1$, we will refer to this lemma also as TML in the sequel.

4 Simple step-stress model under Type-II censoring

We begin with the case of Type-II censoring as it is more easy to follow and simpler to handle in this situation. By denoting $X_{1:n} < \cdots < X_{n:n}$ for the ordered lifetimes and defining $N_1 = \#\{X's \leqslant T_1\}$, the experimenter will observe one among the following three situations:

$$X_{1:n} < \cdots < X_{r:n} \leqslant T_1,$$
$$X_{1:n} < \cdots < X_{N_1:n} \leqslant T_1 < X_{N_1+1:n} < \cdots < X_{r:n},$$
$$T_1 < X_{1:n} < \cdots < X_{r:n}.$$

By writing down the corresponding likelihood function, it is easy to see that the MLEs of both θ_1 and θ_2 exist only in the second case, i.e., when $1 \leqslant N_1 \leqslant r-1$. Balakrishnan et al. (2007) showed that these MLEs are given by

$$\hat{\theta}_1 = \frac{1}{N_1}\left\{\sum_{i=1}^{N_1} X_{i:n} + (n-N_1)T_1\right\} \tag{2}$$

and

$$\hat{\theta}_2 = \frac{1}{r-N_1}\left\{\sum_{i=N_1+1}^{r} (X_{i:n}-T_1) + (n-N_1-r)(X_{r:n}-T_1)\right\}. \tag{3}$$

They then proceeded to discuss exact inference for the parameters θ_1 and θ_2 just by verifying the stochastic monotonicity of these MLEs through extensive numerical computations. We will now formally establish the stochastic monotonicity of $\hat{\theta}_1$ and $\hat{\theta}_2$ with respect to θ_1 and θ_2, respectively, by using the TML presented in the preceding section.

4.1 Stochastic monotonicity of $\hat{\theta}_1$

The survival function of $\hat{\theta}_1$ can be expressed as

$$\mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | 1 \leqslant N_1 \leqslant r-1) = \sum_{n_1=1}^{r-1} \frac{\mathsf{P}_{\theta_1}(N_1 = n_1)}{\mathsf{P}_{\theta_1}(1 \leqslant N_1 \leqslant r-1)} \mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | N_1 = n_1). \tag{4}$$

Since the survival function is of the form (1) with $\mathcal{D} = \{1, \ldots, r-1\}$, we can apply the TML with $k=1$.

(M1) We have to show that the conditional distribution of $\hat{\theta}_1$, given $N_1 = n_1$, is stochastically increasing in θ_1. To this end, recall that conditional on $N_1 = n_1$, $(X_{1:n}, \cdots, X_{n_1:n})$

have the same distribution as $(Z_{1:n_1}, \cdots, Z_{n_1:n_1})$, where $Z_1, \cdots, Z_{n_1} \overset{\text{iid}}{\sim} \mathcal{E}(\theta_1)I(Z \leqslant T_1)$, that is, exponential with parameter θ_1 but right truncated at T_1; see, for example, Arnold et al. (2008). Hence, conditional on $N_1 = n_1$, $\sum_{i=1}^{N_1} X_{i:n} \overset{\text{d}}{=} \sum_{i=1}^{n_1} Z_{i:n_1} \equiv \sum_{i=1}^{n_1} Z_i$. Since the above right truncated exponential distribution is stochastically increasing in θ_1 and Z's are independent, the required result follows immediately.

(M2) Next, we have to prove that the conditional distribution of $\hat{\theta}_1$, given $N_1 = n_1$, is stochastically decreasing in n_1. In order to prove this assertion, we will use standard coupling. For any $n_1 \in \{1, \cdots, r-2\}$, let $Z_1, \cdots, Z_{n_1}, Z_{n_1+1}$ be iid from $\mathcal{E}(\theta_1)I(Z \leqslant T_1)$. Then,

$$\hat{\theta}_1|(N_1 = n_1) \quad \text{has the same distribution as} \quad \frac{1}{n_1}\left\{\sum_{i=1}^{n_1} Z_i + (n - n_1)T_1\right\}$$

while

$$\hat{\theta}_1|(N_1 = n_1 + 1) \quad \text{has the same distribution as} \quad \frac{1}{n_1+1}\left\{\sum_{i=1}^{n_1+1} Z_i + (n - n_1 - 1)T_1\right\}.$$

But,

$$\frac{1}{n_1}\left\{\sum_{i=1}^{n_1} Z_i + (n - n_1)T_1\right\} - \frac{1}{n_1+1}\left\{\sum_{i=1}^{n_1+1} Z_i + (n - n_1 - 1)T_1\right\} = \frac{\sum_{i=1}^{n_1} Z_i + (n - n_1)T_1 + n_1(T - Z_{n_1+1})}{n_1(n_1+1)} > 0,$$

which implies that $\mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | N_1 = n_1) > \mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | N_1 = n_1 + 1)$ for all $x, \theta_1 > 0$.

(M3) Finally, we need to verify that N_1 is stochastically decreasing in θ_1. Note that (the untruncated) N_1 follows the binomial distribution $\mathcal{B}(n, 1 - e^{-T_1/\theta_1})$. This distribution has the monotone likelihood ratio property with respect to θ_1. It is well known that for univariate random variables, this property is not affected by truncation and consequently, $N_1|(1 \leqslant N_1 \leqslant r-1)$ is stochastically decreasing in θ_1.

Thus, the stochastic monotonicity of $\hat{\theta}_1$ with respect to θ_1 follows.

Remark 1. Note that, in almost every respect, $\hat{\theta}_1$ is very similar to the MLE of θ_1 under standard Type-I censoring. Consider a random sample $Y_1, \cdots, Y_n$ from the exponential distribution $\mathcal{E}(\theta_1)$ which is observed up to the point T_1. Setting $N_1 = \#\{Y's \leqslant T_1\}$, the MLE of θ_1 exists for $N_1 \geqslant 1$ and is equal to $\hat{\theta}_1$ (with Y's in the place of X's). Its exact distribution has the form

$$\mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | N_1 \geqslant 1) = \sum_{n_1=1}^{n} \frac{\mathsf{P}_{\theta_1}(N_1 = n_1)}{\mathsf{P}_{\theta_1}(N_1 \geqslant 1)} \mathsf{P}_{\theta_1}(\hat{\theta}_1 > x | N_1 = n_1).$$

This is the same as the expression in (4), but with a slightly different mixing distribution.

4.2 Stochastic monotonicity of $\hat{\theta}_2$

Balakrishnan et al. (2007) showed that the survival function of $\hat{\theta}_2$ is given by

$$\mathsf{P}_{\theta_2}(\hat{\theta}_2 > x | 1 \leqslant N_1 \leqslant r-1) = \sum_{n_1=1}^{r-1} \frac{\mathsf{P}_{\theta_1}(N_1 = n_1)}{\mathsf{P}_{\theta_1}(1 \leqslant N_1 \leqslant r-1)} \mathsf{P}_{\theta_2}(\hat{\theta}_2 > x | N_1 = n_1),$$

with the conditional distribution of θ_2, given $N_1 = n_1$, being gamma $\mathcal{G}\left(r - n_1, \frac{\theta_2}{r-n_1}\right)$. The gamma distribution is stochastically increasing in its scale parameter, and so $\mathsf{P}_{\theta_2}(\hat{\theta}_2 > x | N_1 = n_1)$ is increasing in θ_2. By noting that the mixing distribution does not depend on θ_2, we conclude that the above survival function is increasing in θ_2 for any fixed x, i.e., $\hat{\theta}_2$ is stochastically increasing in θ_2, as required.

5 Simple step-stress model under Type-I censoring

Balakrishnan et al. (2009) considered the following exponential simple step-stress model under time constraint. Let $0 < T_1 < T_2$ be two pre-specified time points. Then, n identical units are placed on a life test at some stress level. At time T_1, the stess level is changed and the experiment terminates at time T_2. Let $X_1, \cdots, X_n$ be the lifetimes of the units, N_1 be the number of failures at the first stress level, and N_2 be the number of failures at the second stress level. Clearly, there is a chance of observing $N_1 = 0$ or/and $N_2 = 0$. However, we will restrict our discussion to the most interesting case when $N_j \geqslant 1$, $j = 1, 2$, in which the experimenter observes data of the form

$$X_{1:n} < \cdots < X_{N_1:n} \leqslant T_1 < X_{N_1+1:n} < \cdots < X_{N_1+N_2:n} \leqslant T_2.$$

In fact, this is the only case in which the MLEs of both θ_1 and θ_2 exist, and are given by

$$\hat{\theta}_1 = \frac{1}{N_1}\left\{ \sum_{i=1}^{N_1} X_{i:n} + (n - N_1)T_1 \right\} \tag{5}$$

and

$$\hat{\theta}_2 = \frac{1}{N_2}\left\{\sum_{i=N_1+1}^{N_1+N_2}(X_{i:n}-T_1)+(n-N_1-N_2)(T_2-T_1)\right\}. \tag{6}$$

Note that $N_j \geqslant 1$, $j=1,2$, implies that the random vector (N_1, N_2) is truncated in the upper orthant $O(1,1)$.

5.1 Stochastic monotonicity of $\hat{\theta}_1$

Let us first consider the MLE $\hat{\theta}_1$ in (5). It is quite easy to see that its survival function has the form

$$\begin{aligned}
&\mathsf{P}_\theta\{\hat{\theta}_1 > x|(N_1,N_2)\geqslant(1,1)\} = \\
&\quad = \sum_{(n_1,n_2)\geqslant(1,1)} \frac{\mathsf{P}_\theta(N_1=n_1,N_2=n_2)}{\mathsf{P}_\theta(N_1\geqslant 1, N_2\geqslant 1)}\mathsf{P}_\theta(\hat{\theta}_1 > x|N_1=n_1) \\
&\quad = \sum_{n_1=1}^{n-1}\frac{k(n_1)\mathsf{P}_{\theta_1}(N_1=n_1)}{\mathsf{P}_\theta(N_1\geqslant 1,N_2\geqslant 1)}\mathsf{P}_{\theta_1}(\hat{\theta}_1 > x|N_1=n_1),
\end{aligned}$$

where $k(n_1)=\sum_{n_2=1}^{n-n_1}\mathsf{P}_{\theta_2}(N_2=n_2|N_1=n_1) = 1-e^{-(n-n_1)(T_2-T_1)/\theta_2}$ does not depend on θ_1. Observe that the mixing distribution in the above mixture is univariate (depending only on n_1) and so we may use the TML with $k=1$.

It is easy to see that **(M1)** and **(M2)** follow exactly along the lines in Subsection 4.1. This is the case for **(M3)** as well, with the only difference being that the truncation set for N_1 is $\{1,\cdots,n-1\}$ instead of $\{1,\cdots,r-1\}$.

Thus follows the stochastic monotonicity of $\hat{\theta}_1$ with respect to θ_1.

Remark 2. As in the case of the step-stress model with Type-II censoring (see Remark 1), the distribution of $\hat{\theta}_1$ is very similar to the MLE of θ_1 under standard Type-I censoring. Yet again, the only difference is in the mixing distribution with $\mathsf{P}_{\theta_1}(N_1=n_1)/\mathsf{P}_{\theta_1}(N_1\geqslant 1)$, $n_1=1,\cdots,n$, being replaced by $k(n_1)\mathsf{P}_{\theta_1}(N_1=n_1)/\mathsf{P}_{\theta_1}(N_1\geqslant 1, N_2\geqslant 1)$, $n_1=1,\cdots,n-1$.

5.2 Stochastic monotonicity of $\hat{\theta}_2$

Now, let us consider the MLE $\hat{\theta}_2$ in (6). Its survival function takes the form

$$\mathsf{P}_\theta\{\hat\theta_2 > x | (N_1,N_2) \geqslant (1,1)\} = \sum_{(n_1,n_2)\geqslant(1,1)} \frac{\mathsf{P}_\theta(N_1=n_1,N_2=n_2)}{\mathsf{P}_\theta(N_1\geqslant 1, N_2\geqslant 1)} \mathsf{P}_{\theta_2}(\hat\theta_2 > x | N_1=n_1, N_2=n_2).$$

Note that in this case we can not reduce the dimension of the mixing distribution and so we have to apply the TML with $k = 2$.

(M1) By conditioning on $(N_1,N_2) = (n_1,n_2) \geqslant (1,1)$, the random vector $(X_{n_1+1:n} - T_1, \cdots, X_{n_1+n_2:n} - T_1)$ has the same distribution as $(Z_{1:n_2}, \cdots, Z_{n_2:n_2})$, where $Z_1, \cdots, Z_{n_2} \overset{\text{iid}}{\sim} \mathcal{E}(\theta_2) I(Z \leqslant T_2 - T_1)$. Now, by using an argument similar to that of **(M1)** in Subsection 4.1, we arrive at the result.

(M2) For any fixed $N_2 \geqslant 1$, $\hat\theta_2$ is clearly a decreasing function of N_1. On the other hand, for any fixed $N_1 \geqslant 1$, the situation is analogous to that of **(M2)** in Subsection 4.1.

(M3) This is the crucial part of the proof since proving that (N_1,N_2) is stochastically decreasing in θ_2 is not simple at all. As already mentioned, the distribution of (N_1,N_2) is trinomial but truncated in the upper orthant $O(1,1)$. In this case, the following result holds.

Result *The conditional distribution of* (N_1,N_2)*, given* $(N_1 \geqslant 1, N_2 \geqslant 1)$*, is stochastically decreasing in* θ_2*.*

Proof. The result is a special case of Theorem 1 which has been established in the Appendix. To see this, all we have to do is to replace p_1 and p_2 by $1 - e^{-T_1/\theta_1}$ and $1 - e^{-(T_2-T_1)/\theta_2}$, respectively. □

Thus follows the stochastic monotonicity of $\hat\theta_2$ with respect to θ_2.

6 Discussion and some final remarks

It is clear that the models introduced by Xiong (1998) and Balakrishnan et al. (2009) can be naturally extended to the case of $m+1 > 2$ stress levels $l_0, l_1, \ldots, l_m$. Then, in order for the MLEs of all the corresponding parameters $\theta_1, \ldots, \theta_{m+1}$ to exist, at least one observation at each stress level must be observed. If N_j denotes the number of observed failures at stress level l_{j-1}, $j = 1, \ldots, m+1$, then the distribution of $\hat\theta_j$ can be expressed as a mixture with the conditional distribution of $(N_1, \ldots, N_{m+1})$, given $(N_1 \geqslant 1, \ldots, N_{m+1} \geqslant 1)$, as the mixing distribution. Hence, the stochastic monotonicity of $\hat\theta_j$ with respect to θ_j can be established using TML. However, **(M3)** would require one to establish that the above truncated (multinomial) distribution is stochastically decreasing in θ_j. As can be seen from the Appendix, proving this particular property is involved even in the case when $m = 1$. Therefore, the proof of the corresponding result for the general case should be quite complicated, although we feel that the truncated multinomial distribution does satisfy the required property.

In concluding this paper, we would like to mention that Balakrishnan and Xie (2007a,b) considered hybrid censoring schemes (cf. Childs et al., 2003) in the context of exponential simple step-stess models. Specifically, let $0 < T_1 < T_2$ be two pre-fixed time points and $r \in \{1, \cdots, n\}$ be a fixed integer. The life test starts at the stress level l_0 which is changed at time T_1 to the stress level l_1. In the first scheme studied by Balakrishnan and Xie (2007b), the experiment continues until the random time $T_2^* = \min\{X_{r:n}, T_2\}$, whereas in the second scheme studied by Balakrishnan and Xie (2007a), the life test continues until the random time $T_2^{**} = \max\{X_{r:n}, T_2\}$. In both these situations, the authors considered the case of exponential lifetimes and developed exact inference as well as asymptotic inference and also bootstrap methods for the underlying parameters. It is also possible to introduce some other forms of censoring such as generalized hybrid censoring (cf. Chandrasekar et al., 2004) in the framework of step-stress models. In all these situations, it will naturally be of great interest to establish the required monotonicity properties for the MLEs in order to formally provide justification for the exact methods of inference developed in these situations.

Appendix

Some properties of a family of trinomial distributions

Let (N_1, N_2) be a random vector with probability mass function (pmf)

$$\begin{aligned} g(n_1, n_2) &= \mathsf{P}(N_1 = n_1, N_2 = n_2) \\ &= \binom{n}{n_1} p_1^{n_1}(1-p_1)^{n-n_1} \binom{n-n_1}{n_2} p_2^{n_2}(1-p_2)^{n-n_1-n_2}, \quad 0 \leqslant n_1, n_2, n_1+n_2 \leqslant n, \end{aligned}$$

i.e., multinomial (trinomial) distribution with cell probabilities p_1, $p_2(1-p_1)$, $1 - p_1 - p_2(1-p_1) = (1-p_1)(1-p_2)$, respectively. It is more convenient for our purposes to express the distribution in this form with p_2 denoting the probability of success of the conditional binomial distribution of N_2 given N_1, since this way the parameters p_1 and p_2 become free of each other. Let us denote the above distribution by $\mathcal{M}(n; p_1, p_2)$.

Fix p_1, and let $p_2 < p_2'$. If $(N_1, N_2) \sim \mathcal{M}(n; p_1, p_2)$ and $(N_1', N_2') \sim \mathcal{M}(n; p_1, p_2')$, then $(N_1, N_2) \leqslant_{\mathrm{st}} (N_1', N_2')$. To see this, notice that $N_1 \stackrel{\mathrm{d}}{=} N_1'$ and conditional on $N_1 = n_1$ (for any n_1), $N_2 \sim \mathcal{B}(n - n_1, p_2)$ and $N_2' \sim \mathcal{B}(n - n_1, p_2')$ which means that $N_2|(N_1 = n_1) \leqslant_{\mathrm{st}} N_2'|(N_1' = n_1)$. So, there is an easy construction in the same probability space such that $(\hat{N}_1, \hat{N}_2) \leqslant (\hat{N}_1', \hat{N}_2')$.

Suppose now, instead of the original multinomial distribution $\mathcal{M}(n; p_1, p_2)$, we have to deal with its truncated version in an upper orthant, i.e., the distribution that has pmf $g(n_1, n_2)/\sum_{(n_1,n_2)\in O(s,t)} g(n_1, n_2)$ when $(n_1, n_2) \in O(s, t)$, for some $0 \leqslant s, t, s + t \leqslant n$. Since the usual multivariate stochastic order (or even the upper orthant one) does not maintain in general under truncation, we can not say at once that the truncated versions of (N_1, N_2) and (N_1', N_2') are still ordered. However, this fact is proved below in the concluding Theorem 1 which requires the following lemmas and propositions.

Lemma 2. *For any n_1 and $s \leqslant t \leqslant n - n_1 - 1$, we have*

$$\frac{\mathsf{P}(N_1' = n_1, s \leqslant N_2' \leqslant t)}{\mathsf{P}(N_1 = n_1, s \leqslant N_2 \leqslant t)} \leqslant \frac{\mathsf{P}(N_1' = n_1 + 1, s \leqslant N_2' \leqslant t)}{\mathsf{P}(N_1 = n_1 + 1, s \leqslant N_2 \leqslant t)}.$$

Proof. We will show that the ratio

$$\frac{\mathsf{P}(N_1 = n_1 + 1, s \leqslant N_2 \leqslant t)}{\mathsf{P}(N_1 = n_1, s \leqslant N_2 \leqslant t)}$$

is increasing in p_2. We have

$$\begin{aligned}
&\frac{\mathsf{P}(N_1 = n_1 + 1, s \leqslant N_2 \leqslant t)}{\mathsf{P}(N_1 = n_1, s \leqslant N_2 \leqslant t)} \\
&\quad = \frac{\sum_{k=s}^{t} \binom{n}{n_1+1} p_1^{n_1+1}(1-p_1)^{n-n_1-1}\binom{n-n_1-1}{k} p_2^k (1-p_2)^{n-n_1-1-k}}{\sum_{k=s}^{t} \binom{n}{n_1} p_1^{n_1}(1-p_1)^{n-n_1}\binom{n-n_1}{k} p_2^k (1-p_2)^{n-n_1-k}} \\
&\quad \propto \frac{\sum_{k=s}^{t} \binom{n-n_1-1}{k} p_2^k (1-p_2)^{-1-k}}{\sum_{k=s}^{t} \binom{n-n_1}{k} p_2^k (1-p_2)^{-k}} \\
&\quad = \frac{\sum_{k=s}^{t} \binom{m-1}{k} \lambda^k (1+\lambda)}{\sum_{k=s}^{t} \binom{m}{k} \lambda^k} \qquad \left(\begin{array}{l} \text{where } \lambda = p_2/(1-p_2),\ 1+\lambda = 1/(1-p_2) \\ \text{and } m = n - n_1 \geqslant t+1 \end{array} \right) \\
&\quad = \frac{\sum_{k=s}^{t} \binom{m-1}{k} \lambda^k + \sum_{k=s}^{t} \binom{m-1}{k} \lambda^{k+1}}{\sum_{k=s}^{t} \binom{m}{k} \lambda^k} \\
&\quad = \frac{\sum_{k=s}^{t} \binom{m-1}{k} \lambda^k + \sum_{k=s+1}^{t+1} \binom{m-1}{k-1} \lambda^{k}}{\sum_{k=s}^{t} \binom{m}{k} \lambda^k}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\binom{m-1}{s}\lambda^s + \sum_{k=s+1}^{t}\left[\binom{m-1}{k} + \binom{m-1}{k-1}\right]\lambda^k + \binom{m-1}{t}\lambda^{t+1}}{\sum_{k=s}^{t}\binom{m}{k}\lambda^k} \\
&= \frac{\binom{m-1}{s}\lambda^s + \sum_{k=s+1}^{t}\binom{m}{k}\lambda^k + \binom{m-1}{t}\lambda^{t+1}}{\sum_{k=s}^{t}\binom{m}{k}\lambda^k} \\
&= \frac{\sum_{k=s}^{t}\binom{m}{k}\lambda^k + \binom{m-1}{t}\lambda^{t+1} - \binom{m-1}{s-1}\lambda^s}{\sum_{k=s}^{t}\binom{m}{k}\lambda^k} \\
&= 1 + \frac{\binom{m-1}{t}\lambda^{t+1} - \binom{m-1}{s-1}\lambda^s}{\sum_{k=s}^{t}\binom{m}{k}\lambda^k}. \qquad (7)
\end{aligned}
$$

Differentiation of the above quantity with respect to λ results in a fraction with numerator

$$
\begin{aligned}
&\left\{(t+1)\binom{m-1}{t}\lambda^t - s\binom{m-1}{s-1}\lambda^{s-1}\right\}\sum_{k=s}^{t}\binom{m}{k}\lambda^k \\
&\quad -\left\{\binom{m-1}{t}\lambda^{t+1} - \binom{m-1}{s-1}\lambda^s\right\}\sum_{k=s}^{t}k\binom{m}{k}\lambda^{k-1} \\
&\quad = (t+1)\binom{m-1}{t}\sum_{k=s}^{t}\binom{m}{k}\lambda^{t+k} - s\binom{m-1}{s-1}\sum_{k=s}^{t}\binom{m}{k}\lambda^{s+k-1} \\
&\qquad - \binom{m-1}{t}\sum_{k=s}^{t}k\binom{m}{k}\lambda^{t+k} + \binom{m-1}{s-1}\sum_{k=s}^{t}k\binom{m}{k}\lambda^{s+k-1} \\
&\quad = \binom{m-1}{t}\sum_{k=s}^{t}(t+1-k)\binom{m}{k}\lambda^{t+k} + \binom{m-1}{s-1}\sum_{k=s}^{t}(k-s)\binom{m}{k}\lambda^{s+k-1} \;>\; 0.
\end{aligned}
$$

Since λ is strictly increasing in p_2, the lemma follows. □

Lemma 3. *For any n_1 and $s \leqslant n - n_1 - 1$, we have*

$$
\frac{\mathsf{P}(N_1' = n_1, s \leqslant N_2' \leqslant n - n_1)}{\mathsf{P}(N_1 = n_1, s \leqslant N_2 \leqslant n - n_1)} \leqslant \frac{\mathsf{P}(N_1' = n_1 + 1, s \leqslant N_2' \leqslant n - n_1 - 1)}{\mathsf{P}(N_1 = n_1 + 1, s \leqslant N_2 \leqslant n - n_1 - 1)}.
$$

Proof. The proof is similar to that of Lemma 2. We need to show that the ratio

$$
\frac{\mathsf{P}(N_1 = n_1 + 1, s \leqslant N_2 \leqslant n - n_1 - 1)}{\mathsf{P}(N_1 = n_1, s \leqslant N_2 \leqslant n - n_1)}
$$

is increasing in p_2. Following exactly the same steps as in the proof of Lemma 2, we arrive at the ratio, instead of (7),

$$
\frac{\sum_{k=s}^{m}\binom{m}{k}\lambda^k - \binom{m-1}{s-1}\lambda^s}{\sum_{k=s}^{m}\binom{m}{k}\lambda^k} = 1 - \frac{\binom{m-1}{s-1}\lambda^s}{\sum_{k=s}^{m}\binom{m}{k}\lambda^k} = 1 - \frac{\binom{m-1}{s-1}}{\sum_{k=s}^{m}\binom{m}{k}\lambda^{k-s}}
$$

which is obviously increasing in λ. Since λ is strictly increasing in p_2, the lemma follows.

□

Lemma 4. *For any n_2 and $s \leqslant t \leqslant n - n_2 - 1$, we have*

$$
\frac{\mathsf{P}(s \leqslant N_1' \leqslant t, N_2' = n_2)}{\mathsf{P}(s \leqslant N_1 \leqslant t, N_2 = n_2)} \leqslant \frac{\mathsf{P}(s \leqslant N_1' \leqslant t, N_2' = n_2 + 1)}{\mathsf{P}(s \leqslant N_1 \leqslant t, N_2 = n_2 + 1)}.
$$

Proof. We will show that the ratio

$$\frac{\mathsf{P}(s \leqslant N_1 \leqslant t, N_2 = n_2 + 1)}{\mathsf{P}(s \leqslant N_1 \leqslant t, N_2 = n_2)}$$

is increasing in p_2. Recall that $\binom{n}{n_1}\binom{n-n_1}{n_2} = \binom{n}{n_2}\binom{n-n_2}{n_1}$. Then, we have

$$\begin{aligned}
&\frac{P(s \leqslant N_1 \leqslant t, N_2 = n_2 + 1)}{P(s \leqslant N_1 \leqslant t, N_2 = n_2)} \\
&\qquad = \frac{\sum_{k=s}^{t} \binom{n-n_2-1}{k} p_1^k (1-p_1)^{n-k} \binom{n}{n_2+1} p_2^{n_2+1} (1-p_2)^{n-k-n_2-1}}{\sum_{k=s}^{t} \binom{n-n_2}{k} p_1^k (1-p_1)^{n-k} \binom{n}{n_2} p_2^{n_2} (1-p_2)^{n-k-n_2}} \\
&\qquad \propto \frac{\sum_{k=s}^{t} \binom{n-n_2-1}{k} p_1^k (1-p_1)^{-k} p_2 (1-p_2)^{-k-1}}{\sum_{k=s}^{t} \binom{n-n_2}{k} p_1^k (1-p_1)^{-k} (1-p_2)^{-k}} \\
&\qquad = \frac{\sum_{k=s}^{t} \binom{n-n_2-1}{k} \lambda^k (\lambda-\mu)/\mu}{\sum_{k=s}^{t} \binom{n-n_2}{k} \lambda^k} \quad \left(\begin{array}{l}\text{where } \lambda = p_1/[(1-p_1)(1-p_2)] \\ \text{and } \mu = p_1/(1-p_1)\end{array}\right) \\
&\qquad \propto \frac{(\lambda-\mu)\sum_{k=s}^{t} \binom{m-1}{k} \lambda^k}{\sum_{k=s}^{t} \binom{m}{k} \lambda^k} \quad (\text{with } m = n - n_2 \geqslant t+1).
\end{aligned}$$

Differentiation of the above quantity with respect to λ results in a fraction with numerator

$$\begin{aligned}
&\Big\{ \sum_{k=s}^{t} \binom{m-1}{k} \lambda^k + (\lambda-\mu) \sum_{k=s}^{t} \binom{m-1}{k} k \lambda^{k-1} \Big\} \sum_{k=s}^{t} \binom{m}{k} \lambda^k \\
&\quad - (\lambda-\mu) \sum_{k=s}^{t} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \nu \lambda^{\nu-1} \\
&= \Big\{ \sum_{k=s}^{t} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \lambda^\nu + \sum_{k=s}^{t} \binom{m-1}{k} k \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \lambda^\nu \\
&\quad - \sum_{k=s}^{t} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \nu \lambda^\nu \Big\} \\
&\quad + \mu \Big\{ \sum_{k=s}^{t} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \nu \lambda^{\nu-1} - \sum_{k=s}^{t} \binom{m-1}{k} k \lambda^{k-1} \sum_{\nu=s}^{t} \binom{m}{\nu} \lambda^\nu \Big\}. \qquad (8)
\end{aligned}$$

We will now show that the quantities in the brackets are both positive. The quantity in the first bracket equals

$$\begin{aligned}
&\Big\{ \binom{m-1}{t} \lambda^t \sum_{k=s}^{t} \binom{m}{k} \lambda^k + \binom{m-1}{t} t \lambda^t \sum_{k=s}^{t} \binom{m}{k} \lambda^k - \binom{m-1}{t} \lambda^t \sum_{k=s}^{t} \binom{m}{k} k \lambda^k \Big\} \\
&\quad + \Big\{ \sum_{k=s}^{t-1} \binom{m-1}{k} \lambda^k \binom{m}{s} \lambda^s + \sum_{k=s}^{t-1} \binom{m-1}{k} k \lambda^k \binom{m}{s} \lambda^s - \sum_{k=s}^{t-1} \binom{m-1}{k} \lambda^k \binom{m}{s} s \lambda^s \Big\} \\
&\quad + \Big\{ \sum_{k=s}^{t-1} \binom{m-1}{k} \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \lambda^\nu + \sum_{k=s}^{t-1} \binom{m-1}{k} k \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \lambda^\nu \\
&\quad - \sum_{k=s}^{t-1} \binom{m-1}{k} \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \nu \lambda^\nu \Big\} \\
&= \binom{m-1}{t} \lambda^t \sum_{k=s}^{t} \binom{m}{k} (1+t-k) \lambda^k + \binom{m}{s} \lambda^s \sum_{k=s}^{t-1} \binom{m-1}{k} (1+k-s) \lambda^k \\
&\quad + \lambda^{-1} \Big\{ \sum_{k=s+1}^{t} \binom{m-1}{k-1} \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \lambda^\nu \\
&\qquad + \sum_{k=s+1}^{t} \binom{m-1}{k-1} (k-1) \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \lambda^\nu
\end{aligned}$$

$$- \sum_{k=s+1}^{t} \binom{m-1}{k-1} \lambda^k \sum_{\nu=s+1}^{t} \binom{m}{\nu} \nu \lambda^{\nu} \}.$$

The first two terms of the last sum are clearly positive. Moreover, the quantity in the brackets equals

$$\begin{aligned} &\sum_{k=s+1}^{t} \sum_{\nu=s+1}^{t} \binom{m-1}{k-1} \binom{m}{\nu} \{1 + (k-1) - \nu\} \lambda^{k+\nu} \\ &\quad = \sum_{k=s+1}^{t} \sum_{\nu=s+1}^{t} \binom{m-1}{k-1} \binom{m}{\nu} (k-\nu) \lambda^{k+\nu} \\ &\quad = \sum_{k=s+1}^{t-1} \sum_{\nu=k+1}^{t} \left\{ \binom{m-1}{k-1} \binom{m}{\nu} (k-\nu) + \binom{m-1}{\nu-1} \binom{m}{k} (\nu - k) \right\} \lambda^{k+\nu}. \end{aligned} \tag{9}$$

In the above sum, for any pair $k < \nu$, $\lambda^{k+\nu}$ is multiplied by

$$\binom{m-1}{k-1} \binom{m}{\nu} (k-\nu) + \binom{m-1}{\nu-1} \binom{m}{k} (\nu - k) = \frac{1}{m} \binom{m}{k} \binom{m}{\nu} (k-\nu)^2 > 0$$

and thus we conclude that the sum is positive. Hence, the quantity inside the first bracket in (8) is positive. Next, the quantity inside the second bracket equals λ^{-1} times

$$\begin{aligned} &\sum_{k=s}^{t} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \nu \lambda^{\nu} - \sum_{k=s}^{t} \binom{m-1}{k} k \lambda^k \sum_{\nu=s}^{t} \binom{m}{\nu} \lambda^{\nu} \\ = \; &\sum_{k=s}^{t} \sum_{\nu=s}^{t} \binom{m-1}{k} \binom{m}{\nu} (k-\nu) \lambda^{k+\nu} \end{aligned}$$

which is the same quantity as that in (9) with $s+1$ replaced by s. Since $\lambda^{-1} > 0$, the quantity inside the second bracket in (8) is also positive, and this proves the lemma. □

Lemma 5. *For any n_2 and $s \leqslant n - n_2 - 1$, we have*

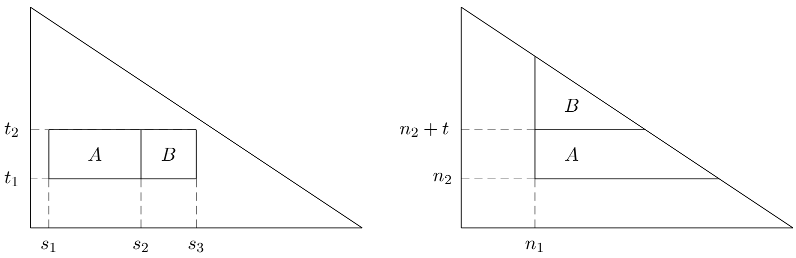
$$\frac{\mathsf{P}(s \leqslant N_1' \leqslant n - n_2, N_2' = n_2)}{\mathsf{P}(s \leqslant N_1 \leqslant n - n_2, N_2 = n_2)} \leqslant \frac{\mathsf{P}(s \leqslant N_1' \leqslant n - n_2 - 1, N_2' = n_2 + 1)}{\mathsf{P}(s \leqslant N_1 \leqslant n - n_2 - 1, N_2 = n_2 + 1)}.$$

Proof. The proof is similar to that of Lemma 4. We need to show that the ratio

$$\frac{\mathsf{P}(s \leqslant N_1 \leqslant n - n_2 - 1, N_2 = n_2 + 1)}{\mathsf{P}(s \leqslant N_1 \leqslant n - n_2, N_2 = n_2)}$$

is increasing in p_2. This ratio is proportional to $(\lambda - \mu) \sum_{k=s}^{m-1} \binom{m-1}{k} \lambda^k / \sum_{k=s}^{m} \binom{m}{k} \lambda^k$, where λ, μ and m are the same as in the proof of Lemma 4. The derivative of this term with respect to λ yields a ratio with numerator

$$\begin{aligned} &\left\{ \sum_{k=s}^{m-1} \binom{m-1}{k} \lambda^k + (\lambda - \mu) \sum_{k=s}^{m-1} \binom{m-1}{k} k \lambda^{k-1} \right\} \sum_{k=s}^{m} \binom{m}{k} \lambda^k \\ &\quad -(\lambda - \mu) \sum_{k=s}^{m-1} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{m} \binom{m}{\nu} \nu \lambda^{\nu-1} \\ &= \left\{ \sum_{k=s}^{m-1} \binom{m-1}{k} \lambda^k \sum_{\nu=s}^{m} \binom{m}{\nu} \lambda^{\nu} + \sum_{k=s}^{m-1} \binom{m-1}{k} k \lambda^k \sum_{\nu=s}^{m} \binom{m}{\nu} \lambda^{\nu} \right. \end{aligned}$$



(a) (b)

Figure 1: (a) Proposition 1(a): The probability ratio on B is larger than or equal to the probability ratio on A. (b) Proposition 2(b): The probability ratio on $B = O(n_1, n_2 + t)$ is larger than or equal to the probability ratio on $A \cup B = O(n_1, n_2)$.

$$
\begin{aligned}
&-\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\sum_{\nu=s}^{m}\binom{m}{\nu}\nu\lambda^\nu\big\} \\
&+\mu\big\{\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\sum_{\nu=s}^{m}\binom{m}{\nu}\nu\lambda^{\nu-1}-\sum_{k=s}^{m-1}\binom{m-1}{k}k\lambda^{k-1}\sum_{\nu=s}^{m}\binom{m}{\nu}\lambda^\nu\big\}.
\end{aligned}
$$

The quantity inside the first bracket equals

$$
\begin{aligned}
&\big\{\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\binom{m}{s}\lambda^s+\sum_{k=s}^{m-1}\binom{m-1}{k}k\lambda^k\binom{m}{s}\lambda^s-\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\binom{m}{s}s\lambda^s\big\} \\
&+\big\{\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\lambda^\nu+\sum_{k=s}^{t-1}\binom{m-1}{k}k\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\lambda^\nu \\
&\quad-\textstyle\sum_{k=s}^{t-1}\binom{m-1}{k}\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\nu\lambda^\nu\big\} \\
=&\textstyle\binom{m}{s}\lambda^s\sum_{k=s}^{m-1}\binom{m-1}{k}(1+k-s)\lambda^k+\lambda^{-1}\big\{\sum_{k=s+1}^{m}\binom{m-1}{k-1}\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\lambda^\nu \\
&+\textstyle\sum_{k=s+1}^{m}\binom{m-1}{k-1}(k-1)\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\lambda^\nu-\sum_{k=s+1}^{m}\binom{m-1}{k-1}\lambda^k\sum_{\nu=s+1}^{m}\binom{m}{\nu}\nu\lambda^\nu\big\} \\
=&\textstyle\binom{m}{s}\lambda^s\sum_{k=s}^{m-1}\binom{m-1}{k}(1+k-s)\lambda^k+\lambda^{-1}\sum_{k=s+1}^{m}\sum_{\nu=s+1}^{m}\binom{m-1}{k-1}\binom{m}{\nu}(k-\nu)\lambda^{k+\nu}
\end{aligned}
$$

and is therefore positive. The quantity inside the second bracket λ^{-1} times

$$
\begin{aligned}
&\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\sum_{k=s}^{m}\binom{m}{k}k\lambda^k-\sum_{k=s}^{m-1}\binom{m-1}{k}k\lambda^k\sum_{k=s}^{m}\binom{m}{k}\lambda^k \\
&=\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k m\lambda^m-\sum_{k=s}^{m-1}\binom{m-1}{k}k\lambda^k\lambda^m \\
&\quad+\textstyle\sum_{k=s}^{m-1}\binom{m-1}{k}\lambda^k\sum_{k=s}^{m-1}\binom{m}{k}k\lambda^k-\sum_{k=s}^{m-1}\binom{m-1}{k}k\lambda^k\sum_{k=s}^{m-1}\binom{m}{k}\lambda^k \\
&=\textstyle\lambda^m\sum_{k=s}^{m-1}\binom{m-1}{k}(m-k)\lambda^k+\sum_{k=s}^{m-1}\sum_{\nu=s}^{m-1}\binom{m-1}{k}\binom{m}{\nu}(\nu-k)\lambda^{k+\nu}
\end{aligned}
$$

which is positive as well, and this completes the proof of the lemma. □

Proposition 1. *Assume that $\mathcal{M}(n; p_1, p_2)$ is truncated in an upper orthant O. We then have the following:*

(a) For any $s_1 \leqslant s_2 < s_3$ *and* $t_1 \leqslant t_2$ *such that the rectangle* $\{(i,j):\ s_1 \leqslant i \leqslant s_3, t_1 \leqslant j \leqslant t_2\}$ *is a subset of* O*, we have*

$$\frac{\mathsf{P}(s_1 \leqslant N_1' \leqslant s_2, t_1 \leqslant N_2' \leqslant t_2)}{\mathsf{P}(s_1 \leqslant N_1 \leqslant s_2, t_1 \leqslant N_2 \leqslant t_2)} \leqslant \frac{\mathsf{P}(s_2+1 \leqslant N_1' \leqslant s_3, t_1 \leqslant N_2' \leqslant t_2)}{\mathsf{P}(s_2+1 \leqslant N_1 \leqslant s_3, t_1 \leqslant N_2 \leqslant t_2)};$$

(b) For any $(n_1, n_2) \in O$ *and* $1 \leqslant s \leqslant n - n_1 - n_2$*, we have*

$$\frac{\mathsf{P}\{(N_1', N_2') \in O(n_1, n_2)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1, n_2)\}} \leqslant \frac{\mathsf{P}\{(N_1', N_2') \in O(n_1+s, n_2)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1+s, n_2)\}}.$$

Proof. (a) By Lemma 2, we have

$$\frac{\mathsf{P}(N_1' = u, t_1 \leqslant N_2' \leqslant t_2)}{\mathsf{P}(N_1 = u, t_1 \leqslant N_2 \leqslant t_2)} \leqslant \frac{\mathsf{P}(N_1' = v, t_1 \leqslant N_2' \leqslant t_2)}{\mathsf{P}(N_1 = v, t_1 \leqslant N_2 \leqslant t_2)}$$

for all pairs $u \in \{s_1, \cdots, s_2\}$ and $v \in \{s_2+1, \cdots, s_3\}$, and the result follows.
(b) By Lemma 3, we have

$$\frac{\mathsf{P}(N_1' = u, n_2 \leqslant N_2' \leqslant n-u)}{\mathsf{P}(N_1 = u, n_2 \leqslant N_2 \leqslant n-u)} \leqslant \frac{\mathsf{P}(N_1' = v, n_2 \leqslant N_2' \leqslant n-v)}{\mathsf{P}(N_1 = v, n_2 \leqslant N_2 \leqslant n-v)}$$

for all pairs $u \in \{n_1, \cdots, n_1+s-1\}$ and $v \in \{n_1+s, \cdots, n-n_2\}$, and the result follows.

□

Proposition 2. *Assume that* $\mathcal{M}(n; p_1, p_2)$ *is truncated in an upper orthant* O*. We then have the following:*
(a) For any $s_1 \leqslant s_2$ *and* $t_1 \leqslant t_2 < t_3$ *such that the rectangle* $\{(i,j):\ s_1 \leqslant i \leqslant s_2, t_1 \leqslant j \leqslant t_3\}$ *is a subset of* O*, we have*

$$\frac{\mathsf{P}(s_1 \leqslant N_1' \leqslant s_2, t_1 \leqslant N_2' \leqslant t_2)}{\mathsf{P}(s_1 \leqslant N_1 \leqslant s_2, t_1 \leqslant N_2 \leqslant t_2)} \leqslant \frac{\mathsf{P}(s_1 \leqslant N_1' \leqslant s_2, t_2+1 \leqslant N_2' \leqslant t_3)}{\mathsf{P}(s_1 \leqslant N_1 \leqslant s_2, t_2+1 \leqslant N_2 \leqslant t_3)};$$

(b) For any $(n_1, n_2) \in O$ *and* $1 \leqslant t \leqslant n - n_1 - n_2$*, we have*

$$\frac{\mathsf{P}\{(N_1', N_2') \in O(n_1, n_2)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1, n_2)\}} \leqslant \frac{\mathsf{P}\{(N_1', N_2') \in O(n_1, n_2+t)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1, n_2+t)\}}.$$

Proof. The proof is similar to that of Proposition 1. □

Propositions 1 and 2 state that the probability ratio increases as we move either to the "right" or "up"; see Figure 1.

The following proposition gives an intermediate result which is essential for proving our final result in Theorem 1.

Proposition 3. *Assume that the multinomial distribution is truncated in the upper orthant* $O(n_1, n_2)$. *Then,* (N_1, N_2) *is stochastically smaller than* (N_1', N_2') *in the upper orthant order.*

Proof. By Propositions 1(b) and 2(b), for any upper orthant $O(n_1+s, n_2+t)$ with $s, t \geqslant 0$, we have

$$\frac{\mathsf{P}\{(N_1', N_2') \in O(n_1+s, n_2+t)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1+s, n_2+t)\}} \geqslant \frac{\mathsf{P}\{(N_1', N_2') \in O(n_1, n_2+t)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1, n_2+t)\}} \geqslant \frac{\mathsf{P}\{(N_1', N_2') \in O(n_1, n_2)\}}{\mathsf{P}\{(N_1, N_2) \in O(n_1, n_2)\}}$$

and thus,

$$\mathsf{P}\{(N_1', N_2') \in O(n_1+s, n_2+t) | (N_1', N_2') \in O(n_1, n_2)\} \geqslant \mathsf{P}\{(N_1, N_2) \in O(n_1+s, n_2+t) | (N_1, N_2) \in O(n_1, n_2)\}.$$

□

Any upper subset U of the lattice $\mathcal{L}(n)$ is a union of a finite number of upper orthants, since $U = \cup_{(n_1,n_2)\in U} O(n_1, n_2)$. However, there exists a minimal representation of U. Let $U^* = \{(n_1, n_2) \in U : \nexists (i,j) \in U \text{ with } (i,j) < (n_1, n_2)\}$ be the set of "minima" of U. Sort the ordinates of the points in U^* in increasing order, and denote this sequence by $n_{21} < \cdots < n_{2m}$. Denote also by n_{1k} the abscissa corresponding to n_{2k}, $k = 1, \cdots, m$. Note that we necessarily have $n_{11} > \cdots > n_{1m}$. Clearly, $U = \cup_{k=1}^m O(n_{1k}, n_{2k})$, since $(n_1, n_2) \notin \cup_{k=1}^m O(n_{1k}, n_{2k})$ if and only if $(n_1, n_2) < (n_{1k}, n_{2k})$ for all $k = 1, \cdots, m$, and then $(n_1, n_2) \notin U$ by the definition of U^*. Furthermore, U can not be represented as a union of less than m upper orthants since then it would exclude some point of U^*.

Theorem 1. *For any upper subset* U *of* $\mathcal{L}(n)$, *we have* $\mathsf{P}\{(N_1', N_2') \in U\} \geqslant \mathsf{P}\{(N_1, N_2) \in U\}$.

Proof. Let U have the minimal representation $U = \cup_{k=1}^m O(n_{1k}, n_{2k})$, $m \geqslant 1$. If $m = 1$, then U is an upper orthant itself for which the assertion has been already proved. So, let $m \geqslant 2$. Moreover, assume that U is connected, i.e., $n_{1k} \leqslant n - n_{2,k+1}$, $k = 1, \cdots, m-1$, since otherwise it can be partitioned into (at least) two connected upper sets and then the proof can be applied to each part separately.

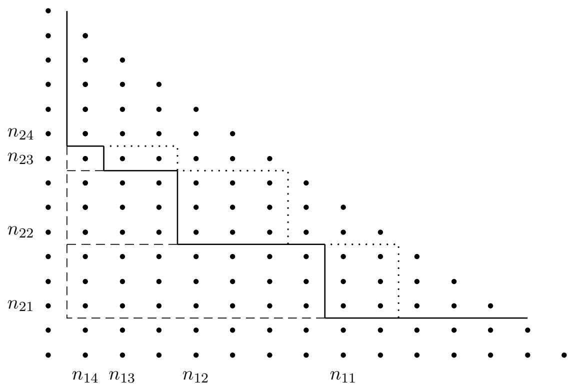


Figure 2: The upper set $U = \cup_{k=1}^{4} O(n_{1k}, n_{2k})$ is above the solid line. The rectangles from the bottom to the top are A_1, A_2, A_3 (dashed lines) and B_2, B_3, B_4 (dotted lines).

Let $R\{a, b, c, d\}$ denote the rectangle with vertices at the points a, b, c, d. For $k = 1, \cdots, m-1$, define

$$A_k = R\{(n_{1m}, n_{2k}), (n_{1m}, n_{2,k+1} - 1), (n_{1k} - 1, n_{2,k+1} - 1), (n_{1k} - 1, n_{2k})\}$$

and notice that $O(n_{1m}, n_{2m}) = \cup_{k=1}^{m-1} A_k \cup U$ is the smallest upper orthant containing U. Set further

$$U_1 = O(n_{1m}, n_{2m}) \quad \text{and} \quad U_k = U_{k-1} \backslash A_{k-1}, \quad j = 2, \cdots, m.$$

Then, $U_1, \cdots, U_m$ is a decreasing sequence with $U_m = U$ and the minimal represenatation of U_k is a union of k upper orthants.

For any measurable set A, let us denote $\mathsf{P}'(A) \equiv \mathsf{P}\{(N_1', N_2') \in A\}$ and $\mathsf{P}(A) \equiv \mathsf{P}\{(N_1, N_2) \in A\}$. Since U_1 is an upper orthant, we have $\mathsf{P}'(U_1) \geqslant \mathsf{P}(U_1)$. We will show that $\mathsf{P}'(U_k) \geqslant \mathsf{P}(U_k)$, $k = 2, \cdots, m$, by dropping the A_k's one after another. Notice here that after completing the k-steps, we would have shown that P' dominates P on any union of k upper orthants.

Since U_2 and A_1 are disjoint with $U_2 \cup A_1 = U_1$, we have $\mathsf{P}'(U_2) + \mathsf{P}'(A_1) \geqslant \mathsf{P}(U_2) + \mathsf{P}(A_1)$. Clearly, if $\mathsf{P}'(A_1) < \mathsf{P}(A_1)$, then $\mathsf{P}'(U_2) \geqslant \mathsf{P}(U_2)$ holds. Assume now that $\mathsf{P}'(A_1) \geqslant \mathsf{P}(A_1)$. Then, U_2 can be partioned as

$$U_2 = O(n_{1m}, n_{22}) \cup O(n - n_{22} + 1, n_{21}) \cup B_2,$$

where $B_2 = R\{(n_{11}, n_{21}), (n_{11}, n_{22} - 1), (n - n_{22}, n_{22} - 1), (n - n_{22}, n_{21})\}$. By Proposition 1(a), $\mathsf{P}'(B_2)/\mathsf{P}(B_2) \geqslant \mathsf{P}'(A_1)/\mathsf{P}(A_1)$ and so $\mathsf{P}'(B_2) \geqslant \mathsf{P}(B_2)$. Since U_2 has the above partition, we conclude that $\mathsf{P}'(U_2) \geqslant \mathsf{P}(U_2)$.

It will now be instructive to consider the case $m = 3$. Since $U_2 = U_3 \cup A_2$, we have $\mathsf{P}'(U_3) + \mathsf{P}'(A_2) \geqslant \mathsf{P}(U_3) + \mathsf{P}(A_2)$. Thus, if $\mathsf{P}'(A_2) < \mathsf{P}(A_2)$, then $\mathsf{P}'(U_3) \geqslant \mathsf{P}(U_3)$ holds. So, assume that $\mathsf{P}'(A_2) \geqslant \mathsf{P}(A_2)$. There are now two possibilities: either $n_{11} \leqslant n - n_{23}$ or $n_{11} > n - n_{23}$. In the first case, partition U_3 as

$$U_3 = \{O(n_{11}, n_{21}) \cup O(n_{13}, n_{23})\} \cup B_3,$$

where $B_3 = R\{(n_{11} - 1, n_{22}), (n_{11} - 1, n_{23} - 1), (n_{12}, n_{23} - 1), (n_{12}, n_{22})\}$. The first part of U_3 is the union of two upper orthants wherein P' dominates P (by the previous step). Moreover, $\mathsf{P}'(B_3)/\mathsf{P}(B_3) \geqslant \mathsf{P}'(A_2)/\mathsf{P}(A_2)$ and so $\mathsf{P}'(B_3) \geqslant \mathsf{P}(B_3)$. Hence, we conclude that $\mathsf{P}'(U_3) \geqslant \mathsf{P}(U_3)$. In the second case, partition U_3 as

$$U_3 = \{O(n_{11}, n_{12}) \cup O(n - n_{23} + 1, n_{22})\} \cup O(n_{13}, n_{23}) \cup B_3,$$

where now $B_3 = R\{(n_{12}, n_{22}), (n_{12}, n_{23} - 1), (n - n_{23}, n_{23} - 1), (n - n_{23}, n_{22})\}$. Once again, use Proposition 1(a) to get $\mathsf{P}'(B_3) \geqslant \mathsf{P}(B_3)$. Since P' dominates P also on the remaining two parts of U_3 (i.e., the union of the two upper orthants and the single upper orthant), we arrive at the result.

Consider now the general case. Given that $\mathsf{P}'(U_k) \geqslant \mathsf{P}(U_k)$, we will show that $\mathsf{P}'(U_{k+1}) \geqslant \mathsf{P}(U_{k+1})$. Since $U_k = U_{k+1} \cup A_k$ and $U_{k+1} \cap A_k = \varnothing$, we have $\mathsf{P}'(U_{k+1}) + \mathsf{P}'(A_k) \geqslant \mathsf{P}(U_{k+1}) + \mathsf{P}(A_k)$. If $\mathsf{P}'(A_k) < \mathsf{P}(A_k)$, then we are done. On the other hand, if $\mathsf{P}'(A_k) \geqslant \mathsf{P}(A_k)$, consider first the case $n_{1,k-1} \leqslant n - n_{2,k+1}$ and partition U_{k+1} as

$$U_{k+1} = \big\{\big[\cup_{j=1}^{k-1} O(n_{1j}, n_{2j})\big] \cup O(n_{1,k+1}, n_{2,k+1})\big\} \cup B_{k+1},$$

where $B_{k+1} = R\{(n_{1,k-1}-1, n_{2k}), (n_{1,k-1}-1, n_{2,k+1}-1), (n_{1k}, n_{2,k+1}-1), (n_{1k}, n_{2k})\}$. Note that on the above union of the k upper orthants, P' dominates P. Moreover, $\mathsf{P}'(B_{k+1}) \geqslant \mathsf{P}(B_k)$ by Proposition 1(a) and the result then follows. When $n_{1,k-1} > n - n_{2,k+1}$, partition U_{k+1} as

$$U_{k+1} = \big\{\big[\cup_{j=1}^{k-1} O(n_{1j}, n_{2j})\big] \cup O(n - n_{2,k+1} + 1, n_{2k})\big\} \cup O(n_{1,k+1}, n_{2,k+1}) \cup B_{k+1},$$

where now $B_{k+1} = R\{(n_{1k}, n_{2k}), (n_{1k}, n_{2,k+1}-1), (n-n_{2,k+1}, n_{2,k+1}-1), (n-n_{2,k+1}, n_{2k})\}$. The first part of U_{k+1} consists of the union of $k-1$ upper orthants wherein P' dominates P.

The second part is an upper orthant and finally by Proposition 1(a), $\mathsf{P}'(B_{k+1})/\mathsf{P}(B_{k+1}) \geqslant \mathsf{P}'(A_k)/\mathsf{P}(A_k)$ which implies that $\mathsf{P}'(B_{k+1}) \geqslant \mathsf{P}(B_{k+1})$ as well. Hence, $\mathsf{P}'(U_{k+1}) \geqslant \mathsf{P}(U_{k+1})$.

By induction, it then follows that $\mathsf{P}'(U) \geqslant \mathsf{P}(U)$, and the theorem is thus established. $\square$

References

Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (2008). *A First Course in Order Statistics*, Classic Edition, SIAM, Philadelphia.

Bagdonavicius, V. and Nikulin, M. (2002). *Accelerated Life Models: Modeling and Statistical Analysis*, Chapman & Hall/CRC Press, Boca Raton, Florida.

Balakrishnan, N. and Iliopoulos, G. (2009). Stochastic monotonicity of the MLE of exponential mean under different censoring schemes, *Annals of the Instute of Statistical Mathematics* (to appear).

Balakrishnan, N., Kundu, D., Ng, H.K.T. and Kannan, N. (2007). Point and interval estimation for a simple step-stress model with Type-II censoring, *Journal of Quality Technology*, **39**, 35–47.

Balakrishnan, N. and Xie, Q. (2007a). Exact inference for a simple step-stress model with Type-II hybrid censored data from the exponential distribution, *Journal of Statistical Planning and Inference*, **137**, 2543–2563.

Balakrishnan, N. and Xie, Q. (2007b). Exact inference for a simple step-stress model with Type-I hybrid censored data from the exponential distribution, *Journal of Statistical Planning and Inference*, **137**, 3268–3290.

Balakrishnan, N., Xie, Q. and Kundu, D. (2009). Exact inference for a simple step-stress model from the exponential distribution under time constraint, *Annals of the Institute of Statistical Mathematics* (to appear).

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

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., Balakrishnan, 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.

Kateri, M. and Balakrishnan, N. (2008). Inference for a simple step-stress model with Type-II censoring and Weibull distributed lifetimes, *IEEE Transactions on Reliability*, **57**, 616–626.

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

Meeker, W.Q. and Escobar, L.A. (1998). *Statistical Methods for Reliability Data*, John Wiley & Sons, New York.

Nelson, W.B. (1990). *Accelerated Life Testing, Statistical Models, Test Plans, and Data Analysis*, John Wiley & Sons, New York.

Sedyakin, N.M. (1966). On the physical principle in reliability theory (in Russian), *Techn. Cybernetics*, **3**, 80–87.

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

Xiong, C. (1998). Inferences on a simple step-stress model with Type-II censored exponential data, *IEEE Transactions on Reliability*, **47**, 142–146.