



# Maximum likelihood estimators based on discrete component lifetimes of a $k$ -out-of- $n$ system

Anna Dembińska<sup>1</sup> · Krzysztof Jasiński<sup>2</sup>

Received: 3 December 2019 / Accepted: 19 June 2020 / Published online: 5 July 2020  
© The Author(s) 2020

## Abstract

This paper deals with parametric inference about the independent and identically distributed discrete lifetimes of components of a  $k$ -out-of- $n$  system. We consider the maximum likelihood estimation assuming that the available data consists of component failure times observed up to and including the moment of the breakdown of the system. First, we provide general conditions for the almost sure existence of a strongly consistent sequence of maximum likelihood estimators (MLE's). Then, we focus on three typical discrete failure distributions—the Poisson, binomial and negative binomial distributions—and prove that in these cases the MLE's are unique, provided they exist, and that they are strongly consistent. Finally, we complete our results by Monte Carlo simulation study. Interestingly, the inference considered in the paper can be viewed as equivalent to one based on Type-II right censored discrete data. Therefore, our results can as well be applied to the case when Type-II right censored sample from a discrete distribution is observed.

**Keywords** MLE · Inference ·  $k$ -Out-of- $n$  system · Poisson distribution · Binomial distribution · Negative binomial distribution

**Mathematics Subject Classification** 62F10 · 62N02 · 62N05

---

✉ Krzysztof Jasiński  
krzys@mat.umk.pl

Anna Dembińska  
dembinsk@mini.pw.edu.pl

<sup>1</sup> Faculty of Mathematics and Information Science, Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland

<sup>2</sup> Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland

## 1 Introduction

An important class of systems studied in reliability theory is one containing so called  $k$ -out-of- $n$  systems. Such systems consist of  $n$  elements and work as long as at least  $k$  of the elements function. As technical structures having some redundancy, they find various applications in engineering when highly reliable products are needed, for example, they are used in design of servers in internet service or in design of automotive and aeronautic engines. Consequently, they have attracted substantial interest—a vast literature on  $k$ -out-of- $n$  systems is available. One stream of this literature concerns inference about component lifetimes based on failure times in a  $k$ -out-of- $n$  system or in a sample of  $k$ -out-of- $n$  systems. Classic works of Halperin (1952) and Bhattacharyya (1985) describe asymptotic properties of MLE's based on failure times of components of a  $k$ -out-of- $n$  system. Generalizations of this results to the case when some of the failure times are censored can be found, among others, in Kong and Fei (1996) and Lin and Balakrishnan (2011). There are also works developing estimation methods for the distribution of components in a system based on a sample of lifetimes of systems; see, for example, Ng et al. (2012), Navarro et al. (2012), Hermanns and Cramer (2018) and the references therein.

All the above-mentioned results, however, concern only the case when the component lifetimes have absolutely continuous distributions. Yet, in some applications the continuity assumption is not adequate. This is the case, for instance, when the system performs a task repetitively and its components have certain probabilities of breakdown upon each cycle or when the component lifetimes represent the numbers of turn-on and switch-off up to failures. While reliability properties of  $k$ -out-of- $n$  systems consisting of components with discrete lifetimes have been studied over the years; see Weiss (1962), Young (1970), Tank and Eryilmaz (2015), Dembińska (2018), Dembińska and Goroncy (2020) and Dembińska et al. (2019), to the best of our knowledge results concerning inference about discrete lifetimes of components based on failure times in  $k$ -out-of- $n$  systems are not known.

The aim of this paper is to fill in this gap in the literature. We focus on maximum likelihood estimation of an unknown parameter of discrete distribution of component lifetimes of a  $k$ -out-of- $n$  system. The estimation is based on failure times of components observed up to and including the system breakdown. In Sect. 2, we set our notation, describe the inference problem under consideration and point out that this problem can be viewed as equivalent to inference from a Type-II right censored sample. Next, in Sect. 3, we present a theorem asserting that under some mild regularity conditions the MLE's of interest exist almost surely for all sufficiently large  $n$  and are strongly consistent. The proof of this theorem is postponed to the "Appendix". In Sect. 4, we choose three typical discrete failure distributions—Poisson, binomial and negative binomial—to be the distributions of lifetimes of the components and show that then the MLE's are unique if they exist, their values can be obtained easily by numerical methods and obtained MLE's are strongly consistent. In Sect. 5, we perform Monte Carlo simulation study to investigate finite-sample properties of MLE's discussed in Sect. 4. Section 6 contains an illustrative example based on real failure data while in Sect. 7 we give concluding remarks and problems for future investigations.

## 2 Maximum likelihood point estimation

Let  $\mathcal{F} = \{F(\theta, \cdot), \theta \in \Theta\}$  be a family of discrete cumulative distribution functions (cdf's), where  $\theta \in \mathbf{R}$  is the parameter of interest. Consider a  $k$ -out-of- $n$  system which consists of  $n$  two-state (i.e., working or failed) components. We assume that the lifetimes of the components,  $T_1, T_2, \dots, T_n$ , are independent and identically distributed (iid) random variables (rv's) with the common cdf  $F(\theta, \cdot) \in \mathcal{F}$ , so that  $F(\theta, t) = P_\theta(T_1 \leq t)$ . Next, we denote  $f(\theta, t) = P_\theta(T_1 = t)$ , i.e.,  $f(\theta, \cdot)$  is the probability mass function (pmf) corresponding to  $F(\theta, \cdot)$ , and  $\bar{F}(\theta, t) = 1 - F(\theta, t)$ . Moreover, for simplicity of notation we require that for any  $\theta \in \Theta$  the support of  $F(\theta, \cdot)$ , denoted by  $\text{supp } F(\theta, \cdot)$ , is of the form  $\{0, 1, \dots, M\}$ , where  $M \leq \infty$ . Yet, it is easily seen that the results of Sect. 3 hold more generally in the case when  $\text{supp } F(\theta, \cdot) = \{x_0, x_1, \dots, x_M\}$ ,  $M \leq \infty$ , where  $x_0 < x_1 < \dots < x_M$  and if  $M = \infty$  then the sequence  $(x_n, n \geq 0)$  has no accumulation points.

Our aim is to use the maximum-likelihood approach to estimate the unknown parameter  $\theta$  from the failure data collected up to and including a breakdown of a  $k$ -out-of- $n$  system. Let  $T_{1:n} \leq T_{2:n} \leq \dots \leq T_{n:n}$  denote the order statistics corresponding to  $T_1, T_2, \dots, T_n$ . A  $k$ -out-of- $n$  system works as long as at least  $k$  of its components work. It fails when the  $(n - k + 1)$ th component failure occurs. Thus, the lifetime of  $k$ -out-of- $n$  system is the  $(n - k + 1)$ th smallest of the component lifetimes, i.e.,  $T_{n-k+1:n}$ . However, in the case of discretely operating elements if  $k \neq 1$  then at the moment of the system failure we do not necessarily have exactly  $n - k + 1$  inoperative elements—due to possible ties between component failures with non-zero probability the number of inoperative elements can be larger than  $n - k + 1$ ; see Davies and Dembińska (2019) for details. Therefore, collecting data up to and including a breakdown of a  $k$ -out-of- $n$  system we can register not only the values of  $T_{1:n}, T_{2:n}, \dots, T_{n-k+1:n}$  but also the value of  $S$ —the number of failed components at the moment of failure of the system. This means that we observe

$$S, T_{1:n}, T_{2:n}, \dots, T_{n-k+1:n},$$

or equivalently,

$$S, T_{1:n}, T_{2:n}, \dots, T_{S:n}.$$

To express in a closed form the joint pmf of  $S, T_{1:n}, T_{2:n}, \dots, T_{S:n}$ ,

$$P_\theta(S = s, T_{1:n} = t_1, \dots, T_{S:n} = t_s),$$

and consequently to find the likelihood function of interest we follow an approach proposed by Gan and Bain (1995) based on the concept of tie-runs. Let  $s \in \{n - k + 1, \dots, n\}$  and  $t_1 \leq t_2 \leq \dots \leq t_s$  have  $m$  tie-runs with lengths  $z_1, z_2, \dots, z_m$  ( $z_1 + \dots + z_m = s$ ), i.e.,

$$t_1 = \dots = t_{z_1} < t_{z_1+1} = \dots = t_{z_1+z_2} < \dots < t_{z_1+\dots+z_{m-1}+1} = \dots = t_{z_1+\dots+z_m} (= t_s).$$

Then the observed likelihood function of  $S, T_{1:n}, T_{2:n}, \dots, T_{s:n}$ , given by

$$\begin{aligned} L(\theta) &= L(\theta; s, t_1, \dots, t_s) \\ &= P_\theta(S = s, T_{1:n} = t_1, \dots, T_{n-k+1:n} = t_{n-k+1}, \dots, T_{s:n} = t_s) \\ &= P_\theta(S = s, T_{1:n} = t_1, \dots, T_{n-k+1:n} = t_{n-k+1}, \dots, T_{s:n} = t_s, T_{s+1:n} > t_s), \end{aligned}$$

where  $T_{n+1:n} = \infty$ , has the form

$$L(\theta) = \frac{n!}{(n-s)! \prod_{i=1}^m z_i!} \left( \prod_{i=1}^m [f(\theta, t_{z_1+\dots+z_i})]^{z_i} \right) [\bar{F}(\theta, t_s)]^{n-s} \tag{1}$$

if  $s \in \{n-k+1, \dots, n\}$  and  $t_{n-k+1} = \dots = t_s$ . Otherwise the right-hand side of (1) reduces to 0.

If the derivatives  $\frac{\partial}{\partial \theta} f(\theta, t), t \in \{0, 1, \dots, M\}$ , exist, then the observed likelihood equation  $\frac{\partial}{\partial \theta} \log L(\theta; s, t_1, \dots, t_s) = 0$  can be written as

$$\sum_{i=1}^m z_i \frac{\partial \log f(\theta, t_{z_1+\dots+z_i})}{\partial \theta} + (n-s) \frac{\partial \log \bar{F}(\theta, t_s)}{\partial \theta} = 0, \tag{2}$$

where  $\frac{\partial}{\partial \theta} \log \bar{F}(\theta, t_s)$  is defined to be equal to 0 if  $t_s = M < \infty$ .

In Sect. 3, we will prove that under some simple regularity conditions concerning the family  $\mathcal{F}$ , the likelihood equation

$$\frac{\partial}{\partial \theta} \log L(\theta; S, T_{1:n}, \dots, T_{s:n}) = 0 \tag{3}$$

with  $P_\theta$ -probability 1, for all sufficiently large  $n$ , has a solution  $\hat{\theta}_n$  such that the sequence  $(\hat{\theta}_n, n \geq 1)$  of estimators of  $\theta$  is strongly consistent. Next, in Sect. 4, we will show that for three families of typical discrete lifetime distributions the MLE of the parameter of interest is unique. Hence, by the result of Sect. 3 in the case of the three families the sequence of MLE's is strongly consistent.

It is worth pointing out that the results presented in this paper, even though formulated in terms of inference from failure times of components of a  $k$ -out-of- $n$  system up to and including its breakdown, can as well be applied to inference based on Type-II right censored discrete data. Indeed, during an experiment in which Type-II right censoring is applied  $n$  items with iid lifetimes  $T_1, T_2, \dots, T_n$  are placed on a test. Due to budget or time limitations or on account of ethical decisions in biomedical problems, the experiment is terminated at the moment of the  $r$ th failure, where  $r < n$  is fixed in advance. If the lifetimes  $T_i, i = 1, \dots, n$ , are discrete rv's, then with non-zero probability it may happen that at the moment of the  $r$ th failure more than  $r$  items are broken. Clearly, in order not to lose any information it is reasonable to make the inference not only from the values of  $T_{1:n}, T_{2:n}, \dots, T_{r:n}$  but to include also the value of  $S$ —the number of failed items at the time of the  $r$ th failure. Therefore, the problem is equivalent to inference based on  $S, T_{1:n}, T_{2:n}, \dots, T_{r:n}$ , and with  $r = n - k + 1$  it is

exactly the same problem as inference from failure times of components of a  $k$ -out-of- $n$  system up to and including its breakdown. To the best of our knowledge maximum likelihood inference for discrete distributions based on censored data has not been studied before in the literature.

### 3 Strong consistency

The standard theorems of asymptotic theory of MLE's constructed from iid observations do not apply to our problem in which we make inference from dependent and non-identically distributed rv's  $S, T_{1:n}, \dots, T_{n-k+1:n}$ . However, as will be shown later on, the basic machinery of proving that under some regularity conditions MLE's from iid observations are strongly consistent can be modified to derive the following analogous result for MLE's obtained from failure data of a  $k$ -out-of- $n$  system.

**Theorem 1** *Assume that the family  $\mathcal{F} = \{F(\lambda), \lambda \in \Theta\}$  satisfies the following three conditions:*

- (A1)  $\Theta \subset \mathbf{R}$  is an open interval (possibly infinite);
- (A2) for all  $\lambda \in \Theta$  and  $j \in \{0, 1, \dots, M\}$ ,  $\frac{\partial^3 f(\lambda, j)}{\partial \lambda^3}$  exists and is a continuous function of  $\lambda \in \Theta$ ;
- (A3)  $\frac{\partial \log f(\lambda, 0)}{\partial \lambda} \neq 0$  for all  $\lambda \in \Theta$ .

Let  $T_1, \dots, T_n$  be iid rv's with cdf  $F(\theta, \cdot)$  for some  $\theta \in \Theta$ . If  $k = k(n) = [(1 - q)n]$ ,  $n \geq 1$ , where  $q \in (0, 1)$  and  $[x]$  stands for the largest integer not exceeding  $x$ , then there exists a sequence  $(\hat{\theta}_n, n \geq 1)$  such that, with  $P_\theta$ -probability 1,

- for all sufficiently large  $n$ ,  $\hat{\theta}_n$  is a solution to the likelihood equation (3);
- $\hat{\theta}_n \rightarrow \theta$  as  $n \rightarrow \infty$ .

**Proof** See the "Appendix". □

Theorem 1 can be used in practice, because for a given family  $\mathcal{F} = \{F(\theta, \cdot), \theta \in \Theta\}$  we can check if its assumptions are satisfied without knowing the value of the true parameter  $\theta$ . In particular, this theorem will allow us to deduce that the MLE's obtained in the next section are strongly consistent.

### 4 MLE's for some specific families of distributions

In this section, we will consider the Poisson  $\text{Poiss}(\theta)$ ,  $\theta > 0$ , binomial  $b(w, \theta)$ ,  $\theta \in (0, 1)$ , and negative binomial  $\text{nb}(w, \theta)$ ,  $\theta \in (0, 1)$ , distributions as possible component lifetime distributions of a  $k$ -out-of- $n$  system. These three distributions, besides the geometric one, are listed by Barlow and Proschan (1996) as typical discrete failure distributions widely used in reliability engineering. We will prove that in the case of all these discrete distributions if the MLE of the parameter  $\theta$  based on observed values of  $S, T_{1:n}, \dots, T_{S:n}$  exists then it is unique. Hence, Theorem 1 will guarantee that in these cases the MLE of  $\theta$  exists almost surely for sufficiently large  $n$  and is strongly consistent.

It is worth pointing out that, since the geometric distribution is a special case of the negative binomial distribution, results proved here for the negative binomial lifetimes of components hold in particular for geometrically distributed lifetimes. Yet the geometric case is easier—then a closed-form formula for the MLE of  $\theta$  can be obtained and hence not only asymptotic but also exact distributional properties of this estimator can be given. For this purpose, the geometric case will be considered in details in a separate paper.

To prove results of this section, we will make use of the following two lemmas. The first one is taken from Pólya and Szegő (1998, p. 41).

**Lemma 1** *Let the radius of convergence of the power series  $\sum_{i=0}^{\infty} \alpha_i x^i$  be  $\rho \in (0, \infty]$ , let the number of its zeros in the interval  $0 < x < \rho$  be  $Z$  and let the number of changes of sign in the sequence of its coefficients be  $C$ . Then  $Z \leq C$ .*

The second lemma concerns linear combinations of Bernstein polynomials and was first proved by Schoenberg (1959). Recall that Bernstein polynomials of degree  $w$  are defined by  $B_{j,w}(x) = \binom{w}{j} x^j (1-x)^{w-j}$ ,  $x \in (0, 1)$ ,  $j = 0, \dots, w$ .

**Lemma 2** *The number of zeros of a given nonzero linear combination of Bernstein polynomials  $B(x) = \sum_{i=0}^n \beta_i B_{i,n}(x)$ ,  $x \in (0, 1)$ , does not exceed the number of sign changes of the sequence  $\beta = (\beta_0, \dots, \beta_n)$ . The first and the last signs of the sum are identical to the signs of the first and the last nonzero element of  $\beta$ , respectively.*

From now on we will assume that the observed values of  $S, T_{1:n}, \dots, T_{s:n}$  are equal to  $s, t_1, \dots, t_s$ , respectively, where  $s \in \{n - k + 1, \dots, n\}$ ,  $t_{n-k+1} = \dots = t_s$  and the chain  $t_1 \leq t_2 \leq \dots \leq t_s$  have  $m$  tie-runs with lengths  $z_1, z_2, \dots, z_m$  ( $z_1 + \dots + z_m = s$ ). Furthermore, for simplicity of notation, we will write

$$\delta = z_1 t_{z_1} + z_2 t_{z_1+z_2} + \dots + z_m t_s. \tag{4}$$

**4.1 Poisson distribution**

Let the component lifetimes  $T_i, i = 1, \dots, n$ , have the Poisson distribution  $\text{Poiss}(\theta)$  with a pmf

$$f(\theta, t) = e^{-\theta} \frac{\theta^t}{t!}, \quad t \in \{0, 1, 2, \dots\}, \tag{5}$$

where  $\theta > 0$  is the parameter to estimate. Then by (1) and (4) the observed likelihood function of  $S, T_{1:n}, \dots, T_{s:n}$  can be written as

$$L(\theta) = C_1 e^{-\theta n} \theta^\delta \left( \sum_{j=t_s+1}^{\infty} \frac{\theta^j}{j!} \right)^{n-s}, \quad \theta > 0,$$

where  $C_1$  does not depend on  $\theta$ . Simple calculations show that

$$\frac{d}{d\theta} \log L(\theta) = -n + \frac{\delta}{\theta} + (n-s) \sum_{j=t_s}^{\infty} \frac{\theta^j}{j!} \left( \sum_{j=t_s+1}^{\infty} \frac{\theta^j}{j!} \right)^{-1}, \quad \theta > 0. \tag{6}$$

From (6) it is easily seen that if  $s = n$ , then the function  $L(\theta), \theta > 0$ , has a global maximum at  $\delta/n$  if  $\delta > 0$ , and does not attain a global maximum if  $\delta = 0$ . Hence, the MLE does not exist when  $T_1 = T_2 = \dots = T_n = 0$ . We see at once that the probability of such an event,  $e^{-n\theta}$ , decreases to 0 as  $n \rightarrow \infty$ . It is worth also noting that if  $s = n$  and  $\delta > 0$ , that is if we observe the following event  $\{T_{n-k+1:n} = T_{n-k+2:n} = \dots = T_{n:n} > 0\}$ , then the MLE is just equal to the sample mean.

It remains to consider the case when  $s \in \{n - k + 1, \dots, n - 1\}$ . For this purpose, note that (6) can be rewritten as

$$\begin{aligned} & \frac{d}{d\theta} \log L(\theta) \\ &= \left( \sum_{j=t_s+1}^{\infty} \frac{\theta^{j+1}}{j!} \right)^{-1} \left( -n \sum_{j=t_s+1}^{\infty} \frac{\theta^{j+1}}{j!} + \delta \sum_{j=t_s+1}^{\infty} \frac{\theta^j}{j!} + (n - s) \sum_{j=t_s}^{\infty} \frac{\theta^{j+1}}{j!} \right) \\ &= \left( \sum_{j=t_s+1}^{\infty} \frac{\theta^{j+1}}{j!} \right)^{-1} \left\{ (\delta + (n - s)(t_s + 1)) \frac{\theta^{t_s+1}}{(t_s + 1)!} + \sum_{j=t_s+2}^{\infty} (\delta - js) \frac{\theta^j}{j!} \right\} \\ &= \left( \sum_{j=t_s+1}^{\infty} \frac{\theta^{j+1}}{j!} \right)^{-1} h(\theta), \text{ say.} \end{aligned} \tag{7}$$

From (7) it is clear that  $\frac{d}{d\theta} \log L(\theta)$  has the same sign as  $h(\theta)$ . But  $h(\theta)$  can be represented as

$$h(\theta) = \sum_{j=0}^{\infty} \alpha_{s,t_1,\dots,t_s}(j) \theta^j, \theta > 0, \tag{8}$$

where

$$\alpha_{s,t_1,\dots,t_s}(j) = \begin{cases} 0, & j = 0, \dots, t_s, \\ ((n - s)(t_s + 1) + \delta)/(j!), & j = t_s + 1, \\ (\delta - sj)/(j!), & j = t_s + 2, t_s + 3, \dots \end{cases}$$

We see at once that  $((n - s)(t_s + 1) + \delta)/(j!) > 0$  since  $s < n$  and  $\delta \geq 0$ . Moreover,  $(\delta - sj)/(j!) < 0$  for  $j \geq t_s + 2$ , because

$$\delta - sj = z_1(t_{z_1} - j) + z_2(t_{z_1+z_2} - j) + \dots + z_m(t_s - j)$$

and  $t_{z_1+\dots+z_i} - j < 0$  for  $i = 1, \dots, m$  and  $j \geq t_s + 2$ , which is due to the fact that  $0 \leq t_{z_1} < t_{z_1+z_2} < \dots < t_{z_1+\dots+z_m} = t_s$ . Consequently, the number of sign changes in the sequence  $(\alpha_{s,t_1,\dots,t_s}(j), j \geq 0)$  equals one. The radius of convergence of the power series in (8) is  $\rho = \infty$ . Therefore, Lemma 1 guarantees that the number of zeros of  $h(\theta)$  in the interval  $(0, \infty)$  is at most one. But from (7) we see that  $h(\theta) = (n - s)\theta^{t_s+1}/(t_s!) + (\delta - s\theta) \sum_{j=t_s+2}^{\infty} \theta^j/(j!)$  and consequently

$$h(\delta/s) = (n - s) \frac{(\delta/s)^{t_s+1}}{t_s!} > 0 \tag{9}$$

and

$$h\left(\frac{\delta + (n - s)(t_s + 1)}{s}\right) = -(n - s)(t_s + 1) \sum_{j=t_s+2}^{\infty} \frac{(\delta + (n - s)(t_s + 1))^j}{s^j j!} < 0. \tag{10}$$

Hence,  $h(\theta)$  has exactly one zero in  $(0, \infty)$ , which by (7) shows that the likelihood equation

$$\frac{d}{d\theta} \log L(\theta) = 0 \tag{11}$$

has exactly one solution in  $(0, \infty)$ . Moreover, the function  $L(\theta)$  is first increasing, then decreasing, which implies it attains its global maximum and the observed MLE of  $\theta$ , being the solution to (11), is unique. From (9) and (10) we know that the observed MLE of  $\theta$  belongs to the finite interval  $\left(\frac{\delta}{s}, \frac{\delta+(n-s)(t_s+1)}{s}\right)$  and therefore can be found easily through numerical methods.

Thus, we have proved the following theorem

**Theorem 2** *From the Poisson distribution with pmf given in (5), suppose we have observed failure times of components of a  $k$ -out-of- $n$  system up to and including the breakdown of the system  $S = s, T_{1:n} = t_1, \dots, T_{s:n} = t_s$ .*

- (1) *Then  $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$ , the observed MLE of  $\theta$ , is unique provided it exists. More precisely, we have*
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$  does not exist if  $s = n$  and  $\delta = 0$  (i.e., if  $t_1 = t_2 = \dots = t_n = 0$ ),
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s) = \delta/n$  if  $s = n$  and  $\delta > 0$ ,
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$  is unique, belongs to the interval  $\left(\frac{\delta}{s}, \frac{\delta+(n-s)(t_s+1)}{s}\right)$  and hence can be obtained easily by numerical methods if  $s \in \{n - k + 1, \dots, n - 1\}$ .
- (2) *Moreover, by Theorem 1, if  $k = [np], 0 < p < 1$ , then almost surely, for all sufficiently large  $n$ ,  $\hat{\theta}_{ML,n} = \hat{\theta}_{ML,n}(S, T_{1:n}, \dots, T_{S:n})$  exists and  $\hat{\theta}_{ML,n}$  is a strongly consistent estimator of  $\theta$ .*

### 4.2 Binomial distribution

Now suppose that the component lifetimes  $T_i, i = 1, \dots, n$ , of a  $k$ -out-of- $n$  system have the binomial distribution  $b(w, \theta)$  with the following pmf

$$f(\theta, t) = \binom{w}{t} \theta^t (1 - \theta)^{w-t}, \quad t \in \{0, 1, \dots, w\}, \tag{12}$$



where  $w \in \{1, 2, \dots\}$  is known and  $\theta \in (0, 1)$  is the parameter to estimate. With the notation (4), the observed likelihood function (1) is given by

$$L(\theta) = C_2 \theta^\delta (1 - \theta)^{ws - \delta} \left[ \sum_{j=t_s+1}^w \binom{w}{j} \theta^j (1 - \theta)^{w-j} \right]^{n-s}, \quad \theta \in (0, 1), \quad (13)$$

where  $C_2$  does not depend on  $\theta$ . An easy computation shows that, for  $\theta \in (0, 1)$ ,

$$\frac{d}{d\theta} \log L(\theta) = \frac{\delta}{\theta} - \frac{ws - \delta}{1 - \theta} + (n-s) \frac{\sum_{j=t_s+1}^w (j - w\theta) \binom{w}{j} \theta^{j-1} (1 - \theta)^{w-j-1}}{\sum_{j=t_s+1}^w \binom{w}{j} \theta^j (1 - \theta)^{w-j}}. \quad (14)$$

If  $s = n$  then we see from (14) that the function  $L(\theta)$ ,  $\theta \in (0, 1)$ , has a global maximum at  $\delta/(wn)$  if  $\delta > 0$ , and does not attain a global maximum if  $\delta = 0$ . Hence, similarly to the Poisson case, the MLE does not exist when  $T_1 = T_2 = \dots = T_n = 0$  and it is easily seen that the probability of non-existence,  $(1 - \theta)^{nw}$ , approaches 0 as  $n \rightarrow \infty$ . Moreover, if  $T_{n-k+1:n} = T_{n-k+2:n} = \dots = T_{n:n} > 0$ , then the MLE is equal to the sample mean divided by  $w$ .

The case of  $s \in \{n - k + 1, \dots, n - 1\}$  requires more effort. Note that (14) can be rewritten as

$$\frac{d}{d\theta} \log L(\theta) = (1 - \theta)^{-1} \left\{ \sum_{j=t_s+1}^w \binom{w}{j} \theta^j (1 - \theta)^{w-j} \right\}^{-1} g(\theta), \quad (15)$$

where

$$g(\theta) = \left\{ \frac{1 - \theta}{\theta} \delta - (ws - \delta) \right\} \sum_{j=t_s+1}^w \binom{w}{j} \theta^j (1 - \theta)^{w-j} + (1 - \theta)(n - s) \sum_{j=t_s+1}^w \binom{w}{j} (j - w\theta) \theta^{j-1} (1 - \theta)^{w-j-1}.$$

Clearly  $g(\theta)$  has the same sign as  $\frac{d}{d\theta} \log L(\theta)$ . Moreover, it is easy to check that  $g(\theta)$  can be represented as the following linear combination of Bernstein polynomials  $B_{j,w}(x) = \binom{w}{j} x^j (1 - x)^{w-j}$ ,  $x \in (0, 1)$ ,  $j = 0, 1, \dots, w$ ,

$$\begin{aligned}
 g(\theta) &= \sum_{j=t_s}^{w-1} \left( \delta \frac{w-j}{j+1} + (n-s)(w-j) \right) B_{j,w}(\theta) \\
 &\quad - \sum_{j=t_s+1}^w ((ws-\delta) + (n-s)(w-j)) B_{j,w}(\theta) \\
 &= \sum_{j=0}^w \beta_{s,t_1,\dots,t_s}(j) B_{j,w}(\theta), \quad \theta \in (0, 1),
 \end{aligned}$$

where

$$\beta_{s,t_1,\dots,t_s}(j) = \begin{cases} 0, & j = 0, \dots, t_s - 1, \\ \frac{w-t_s}{t_s+1} (\delta + (n-s)(t_s+1)), & j = t_s, \\ \frac{\delta(w+1)-ws(j+1)}{j+1}, & j = t_s + 1, \dots, w - 1. \\ -(ws-\delta), & j = w. \end{cases}$$

The coefficient  $\beta_{s,t_1,\dots,t_s}(t_s)$  is positive since  $t_s < w$  when  $s < n$ . Now for  $j = t_s + 1, \dots, w - 1$  we check that  $\beta_{s,t_1,\dots,t_s}(j) < 0$ , which is equivalent to the inequality

$$\begin{aligned}
 &\delta(w+1) - ws(j+1) \\
 &= z_1\{w(t_{z_1} - j) + (t_{z_1} - w)\} + \dots + z_m\{w(t_s - j) + (t_s - w)\} < 0. \tag{16}
 \end{aligned}$$

But

$$t_{z_1} < t_{z_1+z_2} < \dots < t_s < w, \tag{17}$$

which shows that for  $j = t_s + 1, \dots, w - 1$  the expressions in the braces in (16) are negative and hence (16) holds. Finally, we verify that  $\beta_{s,t_1,\dots,t_s}(w) < 0$ . This corresponds to the inequality

$$z_1(t_{z_1} - w) + z_2(t_{z_1+z_2} - w) + \dots + z_m(t_s - w) < 0,$$

which is true because of the relation (17). Summarizing we have proved that

$$\text{sgn}(\beta_{s,t_1,\dots,t_s}(j)) = \begin{cases} 0, & j = 0, \dots, t_s - 1, \\ +1, & j = t_s, \\ -1, & j = t_s + 1, \dots, w. \end{cases}$$

Lemma 2 now ensures that the sign of the derivative (15) is first positive and then negative on  $(0, 1)$ . From this we conclude that the observed likelihood function (13) is first increasing and then decreasing there. Hence, it has a global maximum in  $(0, 1)$  which is attained at the point being the only solution to the observed likelihood equation. Therefore, the observed MLE of  $\theta$  exists and is unique.

Thus, we have the following analogue of Theorem 2.

**Theorem 3** *From the binomial distribution with pmf given in (12), suppose we have observed failure times of components of a k-out-of-n system up to and including the breakdown of the system  $S = s, T_{1:n} = t_1, \dots, T_{s:n} = t_s$ .*

- (1) *Then  $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$ , the observed MLE of  $\theta$ , is unique provided it exists. More precisely, we have*
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$  *does not exist if  $s = n$  and  $\delta = 0$ ,*
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s) = \frac{\delta}{nw}$  *if  $s = n$  and  $\delta > 0$ ,*
  - $\hat{\theta}_{ML,n}(s, t_1, \dots, t_s)$  *is unique and can be obtained easily by numerical methods if  $s \in \{n - k + 1, \dots, n - 1\}$ .*
- (2) *Moreover, conclusion (2) of Theorem 2 holds.*

**4.3 Negative binomial distribution**

Consider a k-out-of-n system composed of n components whose lifetimes  $T_i, i = 1, \dots, n$ , have the negative binomial distribution  $nb(w, \theta)$  with a pmf

$$f(\theta, t) = \binom{t + w - 1}{w - 1} \theta^t (1 - \theta)^w, \quad t \in \{0, 1, 2, \dots\}, \tag{18}$$

where  $w \in \{1, 2, \dots\}$  is known and  $\theta \in (0, 1)$  is the parameter to estimate. Then the observed likelihood function (1) takes on the form

$$\begin{aligned} L(\theta) &= C_3 \theta^\delta (1 - \theta)^{sw} \left\{ 1 - \sum_{j=0}^{t_s} \binom{j + w - 1}{w - 1} \theta^j (1 - \theta)^w \right\}^{n-s} \\ &= C_3 \theta^\delta (1 - \theta)^{nw} \left\{ \sum_{j=t_s+1}^{\infty} \binom{j + w - 1}{w - 1} \theta^j \right\}^{n-s}, \quad \theta \in (0, 1), \end{aligned} \tag{19}$$

where  $C_3$  does not depend on  $\theta$  and  $\delta$  is given in (4). The observed likelihood equation (2) becomes

$$\frac{\delta}{\theta} - \frac{nw}{1 - \theta} + (n - s) \sum_{j=t_s+1}^{\infty} \binom{j + w - 1}{w - 1} j \theta^{j-1} \left\{ \sum_{j=t_s+1}^{\infty} \binom{j + w - 1}{w - 1} \theta^j \right\}^{-1} = 0, \tag{20}$$

or equivalently

$$\{\delta(1 - \theta) - nw\theta\} \sum_{j=t_s+1}^{\infty} \binom{j + w - 1}{w - 1} \theta^{j+(1-\theta)(n-s)} \sum_{j=t_s+1}^{\infty} \binom{j + w - 1}{w - 1} j \theta^j = 0. \tag{21}$$

If  $s = n$  and  $\delta = 0$ , then the function  $L(\theta)$  given in (19) is decreasing and consequently the observed MLE of  $\theta$  does not exist. It is obvious that the probability of non-existence,

$P(T_1 = T_2 = \dots = T_n = 0) = (1 - \theta)^{nw}$ , decreases to 0 as  $n \rightarrow \infty$ . Otherwise, that is when  $s < n$  or  $\delta > 0$ , we have

$$\lim_{\theta \searrow 0} L(\theta) = 0 = \lim_{\theta \nearrow 1} L(\theta)$$

and, since  $L(\theta)$ ,  $\theta \in (0, 1)$ , is continuous and positive, it has a global maximum. In the case when  $s = n$  and  $\delta > 0$  we easily see from (20) that this global maximum is attained at  $\theta = \delta / (nw + \delta)$  and hence that the observed MLE of  $\theta$  is equal to

$$\hat{\theta}_{ML,n}(s, t_1, \dots, t_s) = \frac{\delta}{nw + \delta} = \frac{\bar{t}}{w + \bar{t}} \text{ if } s = n \text{ and } \delta > 0,$$

where  $\bar{t}$  is the observed sample mean.

It remains to consider the case when  $s \in \{n - k + 1, \dots, n - 1\}$ . For this purpose, note that the left-hand side of (21) can be represented as the following power series

$$\begin{aligned} & \sum_{j=t_s+1}^{\infty} (\delta + j(n - s)) \binom{j + w - 1}{w - 1} \theta^j \\ & - \sum_{j=t_s+2}^{\infty} (\delta + nw + (n - s)(j - 1)) \binom{j + w - 2}{w - 1} \theta^j \\ & = \sum_{j=0}^{\infty} \gamma_{s,t_1,\dots,t_s}(j) \theta^j, \end{aligned} \tag{22}$$

where

$$\gamma_{s,t_1,\dots,t_s}(j) = \begin{cases} 0, & j = 0, \dots, t_s, \\ \binom{w+t_s}{w-1} (\delta + (n - s)(t_s + 1)), & j = t_s + 1, \\ \binom{j+w-2}{w-1} \frac{\delta(w-1)-swj}{j}, & j = t_s + 2, t_s + 3, \dots \end{cases}$$

Hence

$$\text{sgn}(\gamma_{s,t_1,\dots,t_s}(j)) = \begin{cases} 0, & j = 0, \dots, t_s, \\ +1, & j = t_s + 1, \\ -1, & j = t_s + 2, t_s + 3, \dots \end{cases} \tag{23}$$

Indeed, if  $s < n$  then  $\delta + (n - s)(t_s + 1) > 0$ , which implies  $\gamma_{s,t_1,\dots,t_s}(t_s + 1) > 0$ . Moreover, if  $j \geq t_s + 2$  then  $j > t_s > \dots > t_{z_1+z_2} > t_{z_1}$  and consequently

$$\begin{aligned} \delta(w - 1) - swj &= z_1((w - 1)t_{z_1} - wj) \\ &+ z_2((w - 1)t_{z_1+z_2} - wj) + \dots + z_m((w - 1)t_s - wj) < 0, \end{aligned}$$

which shows that  $\gamma_{s,t_1,\dots,t_s}(j) < 0$  for  $j \geq t_s + 2$ .

Since the radius of convergence of the power series in (22) is  $\rho = 1$ , from Lemma 1 and (23) we obtain that the left-hand side of (21) [or equivalently of (20)] considered

as a function of  $\theta$  has at most one zero in the interval  $(0, 1)$ . But from the previous discussion, we know that the function  $L(\theta)$ ,  $\theta \in (0, 1)$ , has a global maximum. Therefore, the left-hand side of (20) has exactly one zero in  $(0, 1)$  and this zero is a point at which the likelihood function  $L(\theta)$  attains its global maximum. The observed MLE of  $\theta$  is unique and can be obtained easily by numerical methods.

Thus, we have proved the following result.

**Theorem 4** *From the negative binomial distribution with pmf given in (18), suppose we have observed failure times of components of a  $k$ -out-of- $n$  system up to and including the breakdown of the system  $S = s$ ,  $T_{1:n} = t_1, \dots, T_{s:n} = t_s$ .*

- (1) *Then conclusion (1) of Theorem 3 holds with “ $\frac{\delta}{nw}$ ” replaced by “ $\frac{\delta}{nw+\delta}$ ”.*
- (2) *Moreover, conclusion (2) of Theorem 2 is valid.*

## 5 Monte Carlo simulation study

From Sect. 4 we know that in the case of Poisson  $\text{Pois}(\theta)$ , binomial  $b(w, \theta)$  and negative binomial  $\text{nb}(w, \theta)$  distributions the maximum likelihood estimators of  $\theta$  based on failure times of components of a  $k$ -out-of- $n$  system observed up to and including the breakdown of the system are strongly consistent as  $n \rightarrow \infty$  and  $k = [pn]$ , where  $p \in (0, 1)$  is fixed. The aim of this section is to investigate finite-sample properties of these estimators via Monte Carlo simulation study. For this purpose we assume Poisson  $\text{Pois}(\theta = 1)$ , binomial  $b(w = 4, \theta = 0.5)$  and negative binomial  $\text{nb}(w = 5, \theta = 0.15)$  component lifetimes. The parameters of these distributions were chosen so that the corresponding variances are equal (in the case of the Poisson and binomial distributions) or approximately equal (in the case of the negative binomial distribution) to one. The almost equal variances allow to make comparisons between the three considered cases. Next, for each of the chosen distributions and for some selected values of  $n$  and  $k$  we generate  $N = 1000$  times the failure times of components of a  $k$ -out-of- $n$  system observed up to and including the system breakdown obtaining the data of the form  $s^{(i)}, t_1^{(i)} \leq \dots \leq t_{n-k+1}^{(i)} = \dots = t_{s^{(i)}}^{(i)}$ ,  $i = 1, \dots, N$ . For each  $i = 1, \dots, N$  we then compute  $\hat{\theta}_{\text{ML}}^{(i)}$ , the maximum likelihood estimator of  $\theta$ , using numerical methods if necessary. More precisely, to solve the corresponding likelihood equation we use the method of finding the unique root of a continuous function in a finite interval, such as the bisection method. Finally, we compute the mean and standard deviation of  $\hat{\theta}_{\text{ML}}^{(i)}$ ,  $i = 1, \dots, N$ . These values can be treated as the simulated expectation and standard deviation of  $\hat{\theta}_{\text{ML}}$ . The obtained results are presented in Tables 1, 2 and 3. It is interesting that during the simulations we did not encounter samples with non-existing MLE's. This was so because for the cases considered in the tables the probabilities of non-existence are very small as Table 4 shows.

In the simulation study we observe that even for small  $n$  ( $n = 15$ ) the bias of  $\hat{\theta}_{\text{ML}}$  is small—the simulated expectations of  $\hat{\theta}_{\text{ML}}$  are close to the true values of  $\theta$ . As  $n$  and  $k$  increases in such a way that  $k/n$  is kept fixed both the bias and standard deviation of  $\hat{\theta}_{\text{ML}}$  decreases. Moreover, from Tables 1 and 3 we see that for the same values of  $n$  the

**Table 1** Simulated means and standard deviations of  $\hat{\theta}_{ML}$  for various values of  $n$  and  $k$  when component lifetimes have the Poiss( $\theta = 1$ ) distribution

The case when $k = \frac{1}{3}n$					The case when $k = \frac{2}{3}n$				
$n$	$k$	$n - k + 1$	Mean	SD	$n$	$k$	$n - k + 1$	Mean	SD
15	5	11	1.0030	0.2646	15	10	6	0.9789	0.2821
30	10	21	0.9976	0.1896	30	20	11	0.9806	0.2038
45	15	31	1.0022	0.1514	45	30	16	0.9848	0.1738
60	20	41	0.9988	0.1317	60	40	21	0.9851	0.1421
120	40	81	0.9988	0.0964	120	80	41	0.9884	0.1070

**Table 2** Simulated means and standard deviations of  $\hat{\theta}_{ML}$  for various values of  $n$  and  $k$  when component lifetimes have the binomial  $b(w = 4, \theta = 0.5)$  distribution

The case when $k = \frac{11}{16}n$					The case when $k = \frac{2}{3}n$				
$n$	$k$	$n - k + 1$	Mean	SD	$n$	$k$	$n - k + 1$	Mean	SD
16	11	6	0.4974	0.0687	15	10	6	0.4983	0.0705
32	22	11	0.4975	0.0483	30	20	11	0.4985	0.0488
48	33	16	0.4978	0.0404	45	30	16	0.4989	0.0411
64	44	21	0.4980	0.0341	60	40	21	0.4990	0.0346
128	88	41	0.4983	0.0236	120	80	41	0.4990	0.0252

**Table 3** Simulated means and standard deviations of  $\hat{\theta}_{ML}$  for various values of  $n$  and  $k$  when component lifetimes have the negative binomial distribution  $nb(w = 5, \theta = 0.15)$  distribution

The case when $k = \frac{1}{3}n$					The case when $k = \frac{2}{3}n$				
$n$	$k$	$n - k + 1$	Mean	SD	$n$	$k$	$n - k + 1$	Mean	SD
15	5	11	0.1515	0.0180	15	10	6	0.1538	0.0218
30	10	21	0.1510	0.0126	30	20	11	0.1516	0.0152
45	15	31	0.1504	0.0099	45	30	16	0.1516	0.0120
60	20	41	0.1502	0.0087	60	40	21	0.1509	0.0107

**Table 4** Selected approximate values of probabilities of non-existence of MLE (formulas for these probabilities are taken from Sect. 4)

Distribution Formula	Poiss( $\theta = 1$ ) $e^{-n\theta} = e^{-n}$	$b(w = 4, \theta = 0.5)$ $(1 - \theta)^{nw} = 16^{-n}$	$nb(w = 5, \theta = 0.15)$ $(1 - \theta)^{nw} = (0.85)^{5n}$
$n = 15$	$3 \times 10^{-7}$	$9 \times 10^{-19}$	$5 \times 10^{-6}$
$n = 30$	$9 \times 10^{-14}$	$8 \times 10^{-37}$	$3 \times 10^{-11}$
$n = 60$	$9 \times 10^{-27}$	$6 \times 10^{-73}$	$7 \times 10^{-22}$

bias and standard deviation are smaller when  $n - k + 1$  is larger. This is so because the single experiment terminates at the moment of the  $(n - k + 1)$ th component failure and larger  $n - k + 1$  allows to collect more information and thus to obtain a better precision of estimation.

In Table 2, we see a surprising situation. For the same values of  $n - k + 1$  we obtain larger bias when  $n$  is larger. This may not agree with our intuition—larger  $n$  means that more elements are involved in a single experiment and thus we may expect a better estimation. Yet this is not the case. The reason is that if  $k = k(n) = [(1 - q)n]$ , where  $q \in (0, 1)$  is such that the  $q$ th quantile of  $F(\theta, \cdot)$  is not unique, then due to (38) the behavior of  $T_{n-k+1:n}$  is unstable causing worse behavior of  $\hat{\theta}_{ML}$ . Note that the  $q$ th quantile of the binomial  $b(w = 4, \theta = 0.5)$  distribution is not unique when  $1 - q = 11/16$  and is so when  $1 - q = 2/3$ . Therefore, biases presented in the left-hand side of Table 2 are greater than the corresponding ones given in the right-hand side of this table.

## 6 Illustrative example

The following are times until breakdown in days of air monitors operated at a nuclear power plant:  $T_1(\omega) = 8$ ,  $T_2(\omega) = 26$ ,  $T_3(\omega) = 10$ ,  $T_4(\omega) = 8$ ,  $T_5(\omega) = 29$ ,  $T_6(\omega) = 20$ ,  $T_7(\omega) = 10$ , for fixed  $\omega \in \Omega$ ; see Bickel and Doksum (1977, p. 189). Assuming that the sample is from a Poisson  $\text{Poiss}(\theta)$  population and considering 3 scenarios we will find MLE's of  $\theta$ .

1. For an uncensored sample it is well known that the MLE of  $\theta$  is equal to the mean. Therefore, based on the whole sample we obtain  $\hat{\theta}_{ML}^{(1)} = 15.86$ .
2. Now suppose that we terminate the experiment at the moment of the  $r = 5$ th failure, that is after 20 days. Then we have exactly  $s = 5$  air monitors broken. Using the inference based on  $S(\omega) = 5$ ,  $T_{1:7}(\omega) = 8$ ,  $T_{2:7}(\omega) = 8$ ,  $T_{3:7}(\omega) = 10$ ,  $T_{4:7}(\omega) = 10$ ,  $T_{5:7}(\omega) = 20$ , we get  $\hat{\theta}_{ML}^{(2)} = 12.80$ .
3. Finally, let us consider censoring by terminating the experiment at the moment of the  $r = 3$ th failure, that is after 10 days. Then we observe  $s = 4$  air monitors breakdowns. Thus, we collect the following data:  $S(\omega) = 4$ ,  $T_{1:7}(\omega) = 8$ ,  $T_{2:7}(\omega) = 8$ ,  $T_{3:7}(\omega) = 10$ ,  $T_{4:7}(\omega) = 10$ . The MLE based on this data is equal to  $\hat{\theta}_{ML}^{(3)} = 10.87$ .

We see that the value of the MLE changes significantly when we change the censoring scenario. This unpleasurable feature is due to the fact that  $n = 7$  is very small. Apparently, to obtain more reliable estimates we need to conduct an experiment with a larger number of air monitors.

## 7 Conclusions

In this paper, we have focused on maximum likelihood inference of the discrete lifetime distribution of components of a  $k$ -out-of- $n$  system in the case when failure times of the components observed up to and including the moment of the breakdown of the

system are available. Another problem of interest is the inference in the case when a sample of lifetimes of  $k$ -out-of- $n$  systems and numbers of broken components at the moment of the system failure is given. We are currently working on the latter problem and planning to report our findings in a forthcoming paper.

It is also worth pointing out that the new results we obtained for the discrete case are analogous to that known in the literature for the continuous case in the sense that in both the cases under some regularity conditions the MLE's of interest exist almost surely for sufficiently large  $n$  and are strongly consistent. Yet, the regularity conditions for the two cases are different and in the proofs different techniques are needed.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

The Appendix is devoted to the proof of Theorem 1. From now on we assume that  $\theta$  is the true value of the unknown parameter. For  $q \in (0, 1)$ , let  $\underline{\gamma}_q$  and  $\overline{\gamma}_q$  stand for the lower and upper  $q$ th quantile of  $F(\theta, \cdot)$ , respectively, i.e.,

$$\underline{\gamma}_q = \inf\{x \in \mathbf{R}: F(\theta, x) \geq q\} \quad \text{and} \quad \overline{\gamma}_q = \sup\{x \in \mathbf{R}: F(\theta, x) \leq q\}. \tag{24}$$

If  $\underline{\gamma}_q = \overline{\gamma}_q$  then we say that the  $q$ th quantile of  $F(\theta, \cdot)$  is unique and denote it by  $\gamma_q$ . Otherwise we say that this quantile is not unique. For example, if  $F(\theta, \cdot)$  is the cdf of  $T_1$  such that  $P_\theta(T_1 = 0) = 1 - P_\theta(T_1 = 1) = 0.5$ , then

- for  $q = 0.5$  the  $q$ th quantile is not unique since  $\underline{\gamma}_{0.5} = 0 < \overline{\gamma}_{0.5} = 1$ ,
- for  $q \neq 0.5$  the  $q$ th quantile is unique and  $\gamma_q = 0$  if  $q \in (0, 0.5)$  and  $\gamma_q = 1$  if  $q \in (0.5, 1)$ .

The symbol  $\xrightarrow{P_\theta\text{-a.s.}}$  denotes convergence with  $P_\theta$ -probability 1. If  $M = \infty$  then  $M - 1$  is understood as  $\infty$ .  $I(\cdot)$  stands for the indicator function, that is  $I(A) = 1$  if the event  $A$  occurs and  $I(A) = 0$  otherwise. Moreover, to simplify notation, from now on we adopt the convention that

$$\frac{\partial^j \log \overline{F}(\lambda, \beta)}{\partial \lambda^j} \Big|_{\lambda=\lambda_0} \cdot I(T_i > \beta) = 0, \quad j = 1, 2, 3, \quad \text{whenever } \overline{F}(\lambda_0, \beta) = 0.$$

We start with a technical lemma which will simplify our discussion.

**Lemma 3** *Assume that the family  $\mathcal{F} = \{F(\lambda), \lambda \in \Theta\}$  satisfies conditions (A1)–(A3) of Theorem 1. Let  $T_1, \dots, T_n$  be iid rv's with cdf  $F(\theta, \cdot)$  for some  $\theta \in \Theta$ . Then*



- (1)  $\frac{\partial^3 \log f(\lambda, j)}{\partial \lambda^3}$  exists for any  $\lambda \in \Theta$  and  $j \in \{0, 1, \dots, M\}$ , and  $\frac{\partial^3 \log \bar{F}(\lambda, j)}{\partial \lambda^3}$  exists for all  $\lambda \in \Theta$  and  $j \in \{0, 1, \dots, M - 1\}$ ;
- (2) for  $\beta = 0, 1, \dots, M$ ,

$$A_n^{(1)}(\beta) \xrightarrow{P_\theta\text{-a.s.}} 0 \text{ and } A_n^{(2)}(\beta) \xrightarrow{P_\theta\text{-a.s.}} -v_\theta(\beta) \text{ as } n \rightarrow \infty, \tag{25}$$

where

$$A_n^{(1)}(\beta) = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\partial \log f(\lambda, T_i)}{\partial \lambda} \Big|_{\lambda=\theta} \cdot I(T_i \leq \beta) + \frac{\partial \log \bar{F}(\lambda, \beta)}{\partial \lambda} \Big|_{\lambda=\theta} \cdot I(T_i > \beta) \right\}, \tag{26}$$

$$A_n^{(2)}(\beta) = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\partial^2 \log f(\lambda, T_i)}{\partial \lambda^2} \Big|_{\lambda=\theta} \cdot I(T_i \leq \beta) + \frac{\partial^2 \log \bar{F}(\lambda, \beta)}{\partial \lambda^2} \Big|_{\lambda=\theta} \cdot I(T_i > \beta) \right\}, \tag{27}$$

$$v_\theta(\beta) = \sum_{j=0}^\beta \left[ \frac{\partial \log f(\lambda, j)}{\partial \lambda} \right]^2 \Big|_{\lambda=\theta} \cdot f(\theta, j) + \left[ \frac{\partial \log \bar{F}(\lambda, \beta)}{\partial \lambda} \right]^2 \Big|_{\lambda=\theta} \cdot \bar{F}(\theta, \beta) \in (0, \infty); \tag{28}$$

- (3) for a fixed neighborhood  $N(\theta)$  of  $\theta$  such that the closure of  $N(\theta)$  is contained in  $\Theta$ , and for every  $c \in N(\theta)$ ,  $\beta = 0, 1, \dots, M$ ,  $n = 1, 2, \dots$  there exists a rv  $\tilde{A}_n^{(3)}(\beta)$  satisfying  $|A_n^{(3)}(\beta)| \leq \tilde{A}_n^{(3)}(\beta)$  and

$$\tilde{A}_n^{(3)}(\beta) \xrightarrow{P_\theta\text{-a.s.}} w_\theta(\beta) \text{ as } n \rightarrow \infty, \tag{29}$$

where

$$A_n^{(3)}(\beta) = \frac{1}{n} \sum_{i=1}^n \left\{ \frac{\partial^3 \log f(\lambda, T_i)}{\partial \lambda^3} \Big|_{\lambda=c} \cdot I(T_i \leq \beta) + \frac{\partial^3 \log \bar{F}(\lambda, \beta)}{\partial \lambda^3} \Big|_{\lambda=c} \cdot I(T_i > \beta) \right\} \tag{30}$$

and  $w_\theta(\beta) \in [0, \infty)$  is a non-random constant depending only on  $\theta, \beta$  and  $N(\theta)$ .

**Proof** Part (1) is a consequence of assumption (A2) and the fact that  $\bar{F}(\lambda, j) = 1 - \sum_{i=0}^j f(\lambda, i)$ .

To prove part (2), note that  $A_n^{(1)}(\beta)$  and  $A_n^{(2)}(\beta)$  are means of iid rv's with  $P_\theta$ -expectations equal to

$$\begin{aligned} & \sum_{j=0}^{\beta} \frac{\partial \log f(\lambda, j)}{\partial \lambda} \Big|_{\lambda=\theta} \cdot f(\theta, j) + \frac{\partial \log \bar{F}(\lambda, \beta)}{\partial \lambda} \Big|_{\lambda=\theta} \cdot \bar{F}(\theta, \beta) = \sum_{j=0}^{\beta} \frac{\partial f(\lambda, j)}{\partial \lambda} \Big|_{\lambda=\theta} \\ & + \frac{\partial \bar{F}(\lambda, \beta)}{\partial \lambda} \Big|_{\lambda=\theta} = \frac{\partial}{\partial \lambda} \left\{ \sum_{j=0}^{\beta} f(\lambda, j) + \bar{F}(\lambda, \beta) \right\} \Big|_{\lambda=\theta} = \frac{\partial}{\partial \lambda} (1) = 0 \end{aligned}$$

and

$$\begin{aligned} & \sum_{j=0}^{\beta} \frac{\partial^2 \log f(\lambda, j)}{\partial \lambda^2} \Big|_{\lambda=\theta} \cdot f(\theta, j) + \frac{\partial^2 \log \bar{F}(\lambda, \beta)}{\partial \lambda^2} \Big|_{\lambda=\theta} \cdot \bar{F}(\theta, \beta) \\ & = \frac{\partial^2}{\partial \lambda^2} \left\{ \sum_{j=0}^{\beta} f(\lambda, j) + \bar{F}(\lambda, \beta) \right\} \Big|_{\lambda=\theta} - \left\{ \sum_{j=0}^{\beta} \left( \frac{1}{f(\lambda, j)} \frac{\partial f(\lambda, j)}{\partial \lambda} \right)^2 \right\} \Big|_{\lambda=\theta} \cdot f(\theta, j) \\ & + \left\{ \frac{1}{\bar{F}(\lambda, \beta)} \frac{\partial \bar{F}(\lambda, \beta)}{\partial \lambda} \right\}^2 \Big|_{\lambda=\theta} \cdot \bar{F}(\theta, \beta) \Big\} = -v_\theta(\beta), \end{aligned}$$

respectively. Clearly  $v_\theta(\beta) < \infty$ . Moreover, assumption (A3) guarantees that  $v_\theta(\beta) > 0$ . The strong law of large numbers gives (25).

For the proof of part (3) first observe that the continuity of  $\frac{\partial^3 f(\lambda, j)}{\partial \lambda^3}$  with respect to  $\lambda \in \Theta$  implies that of  $\frac{\partial^3 \log f(\lambda, j)}{\partial \lambda^3}$  and  $\frac{\partial^3 \log \bar{F}(\lambda, \beta)}{\partial \lambda^3}$ . In particular, these functions are continuous on the closure of  $N(\theta)$  and hence bounded there. This means that there exist functions  $g_\theta: \{0, 1, \dots, M\} \rightarrow [0, \infty)$  and  $h_\theta: \{0, 1, \dots, M - 1\} \rightarrow [0, \infty)$  satisfying

$$\left| \frac{\partial^3 \log f(\lambda, j)}{\partial \lambda^3} \right| \leq g_\theta(j) \quad \text{and} \quad \left| \frac{\partial^3 \log \bar{F}(\lambda, \beta)}{\partial \lambda^3} \right| \leq h_\theta(j) \quad \text{for all } \lambda \in N(\theta).$$

Consequently, for  $c \in N(\theta)$ ,

$$|A_n^{(3)}(\beta)| \leq \frac{1}{n} \sum_{i=1}^n \{g_\theta(T_i)I(T_i \leq \beta) + h_\theta(\beta)I(T_i > \beta)\}, \tag{31}$$

where, in the case of  $M < \infty$ , we can define  $h_\theta(M) = 0$ . We will show that  $\tilde{A}_n^{(3)}(\beta)$  can be taken to be equal to the right-hand side of (31). For this purpose, note that the right-hand side of (31) is a mean of iid rv's with  $P_\theta$ -expectation given by

$$\sum_{j=0}^{\beta} g_\theta(j) f(\theta, j) + h_\theta(\beta) \bar{F}(\theta, \beta). \tag{32}$$

Set  $w_\theta(\beta)$  equal to (32). It is easily seen that  $w_\theta(\beta) \in [0, \infty)$ . By the strong law of large numbers, we obtain (29), and the proof of part (3) is complete.  $\square$

Now we are ready to prove Theorem 1. We will divide the proof into two parts. In the first part we will consider the simpler case when the  $q$ th quantile of  $F(\theta, \cdot)$  is unique. In the second part we will deal with the more difficult case when this quantile is not unique.

**Proof of Theorem 1** *Part 1 for the case when the  $q$ th quantile of  $F(\theta, \cdot)$ ,  $\gamma_q$ , is unique*

Analyzing (1) we see that the log-likelihood function of  $S, T_1, \dots, T_S$  can be written in terms of  $T_1, \dots, T_n$  as

$$C + \sum_{i=1}^n \log f(\theta, T_i) \cdot I(T_i \leq T_{n-k+1:n}) + \sum_{i=1}^n \log \bar{F}(\theta, T_{n-k+1:n}) \cdot I(T_i > T_{n-k+1:n}) = l(\theta; T_1, \dots, T_n), \text{ say,}$$

where  $C$  does not depend on  $\theta$ . On account of part (1) of Lemma 3, we can use Taylor expansion of the function  $\frac{\partial}{\partial \lambda} l(\lambda; t_1, \dots, t_n)$  around  $\theta$  to get

$$\frac{1}{n} \frac{\partial}{\partial \lambda} l(\lambda; T_1, \dots, T_n) = A_n^{(1)}(T_{n-k+1:n}) + (\lambda - \theta) A_n^{(2)}(T_{n-k+1:n}) + \frac{1}{2} (\lambda - \theta)^2 A_n^{(3)}(T_{n-k+1:n}), \tag{33}$$

where  $A_n^{(i)}(\cdot), i = 1, 2, 3$ , are defined in (26), (27) and (30), respectively, and  $c$  is a point between  $\lambda$  and  $\theta$ .

Now recall that the uniqueness of the  $q$ th quantile of  $F(\theta, \cdot)$  implies  $T_{n-k+1:n} \xrightarrow{P_\theta\text{-a.s.}} \gamma_q$ , as  $n \rightarrow \infty$ ; see Smirnov (1952). Consequently, since  $\gamma_q$  is not an accumulation point of the support of  $F(\theta, \cdot)$ , we must have, with  $P_\theta$ -probability 1,

$$T_{n-k+1:n} = \gamma_q \text{ for all sufficiently large } n; \tag{34}$$

see Dembińska (2012, proof of Proposition 2.1). Therefore, with  $P_\theta$ -probability 1,  $A_n^{(k)}(T_{n-k+1:n}) = A_n^{(k)}(\gamma_q)$  for all sufficiently large  $n, k = 1, 2, 3$ , which implies

$$A_n^{(k)}(T_{n-k+1:n}) - A_n^{(k)}(\gamma_q) \xrightarrow{P_\theta\text{-a.s.}} 0, \quad k = 1, 2, 3. \tag{35}$$

Using (35) and part (2) of Lemma 3 we get

$$A_n^{(1)}(T_{n-k+1:n}) \xrightarrow{P_\theta\text{-a.s.}} 0 \text{ and } A_n^{(2)}(T_{n-k+1:n}) \xrightarrow{P_\theta\text{-a.s.}} -v_\theta(\gamma_q) \text{ as } n \rightarrow \infty, \tag{36}$$

where  $v_\theta(\cdot)$  is given by (28) and  $v_\theta(\gamma_q) \in (0, \infty)$ . Fix a neighborhood  $N(\theta)$  of  $\theta$  such that the closure of  $N(\theta)$  is contained in  $\Theta$ . Assuming that (33) holds for  $\lambda \in N(\theta)$

and applying (35) and part (3) of Lemma 3 we obtain

$$|A_n^{(3)}(T_{n-k+1:n})| \leq |A_n^{(3)}(T_{n-k+1:n}) - A_n^{(3)}(\underline{\gamma}_q)| + \tilde{A}_n^{(3)}(\underline{\gamma}_q) \xrightarrow{P_\theta\text{-a.s.}} w_\theta(\underline{\gamma}_q), \tag{37}$$

where  $w_\theta(\underline{\gamma}_q)$  is a non-negative real constant.

Having proved (33), (36) and (37) we can now proceed as in the proof of theorem from Section 4.2.2 of Serfling (2001) to show that the conclusion of Theorem 1 is true. *Part 2 for the case when the  $q$ th quantile of  $F(\theta, \cdot)$ ,  $\gamma_q$ , is not unique*

In this case, that is when  $\underline{\gamma}_q < \bar{\gamma}_q$ , where  $\underline{\gamma}_q$  and  $\bar{\gamma}_q$  are defined in (24), we have with  $P_\theta$ -probability 1

$$\liminf_{n \rightarrow \infty} T_{n-k+1:n} = \underline{\gamma}_q \quad \text{and} \quad \limsup_{n \rightarrow \infty} T_{n-k+1:n} = \bar{\gamma}_q; \tag{38}$$

see Dembińska (2014, Section 3). Using the same arguments as in the proof of (34) we see that (38) and the fact that  $\underline{\gamma}_q$  nor  $\bar{\gamma}_q$  are not accumulation points of the support of  $F(\theta, \cdot)$  give, with  $P_\theta$ -probability 1,

$$T_{n-k+1:n} \in \{\underline{\gamma}_q, \bar{\gamma}_q\} \quad \text{for all sufficiently large } n. \tag{39}$$

Now we fix a neighborhood  $N(\theta)$  of  $\theta$  such that the closure of  $N(\theta)$  is contained in  $\Theta$ , and write Taylor expansion (33) for  $\lambda \in N(\theta)$ . Next, we set  $\lambda_1 = \theta - \varepsilon$ ,  $\lambda_2 = \theta + \varepsilon$ ,

where  $0 < \varepsilon < \frac{\min(v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q))}{w_\theta(\underline{\gamma}_q) + w_\theta(\bar{\gamma}_q)}$  and  $\varepsilon$  is small enough that  $\lambda_1, \lambda_2 \in N(\theta)$ . The functions  $v_\theta(\cdot)$  and  $w_\theta(\cdot)$  are specified by Lemma 3. From (33) we get

$$\begin{aligned} & \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} \cdot I(T_{n-k+1:n} = \underline{\gamma}_q) \\ &= \left\{ A_n^{(1)}(\underline{\gamma}_q) - \varepsilon A_n^{(2)}(\underline{\gamma}_q) + \frac{1}{2} \varepsilon^2 A_n^{(3)}(\underline{\gamma}_q) \right\} \cdot I(T_{n-k+1:n} = \underline{\gamma}_q), \end{aligned}$$

which implies

$$\begin{aligned} & \left| \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} - \varepsilon v_\theta(\underline{\gamma}_q) \right| \cdot I(T_{n-k+1:n} = \underline{\gamma}_q) \\ & \leq \left\{ |A_n^{(1)}(\underline{\gamma}_q)| + \varepsilon |A_n^{(2)}(\underline{\gamma}_q)| + v_\theta(\underline{\gamma}_q) \right\} + \frac{1}{2} \varepsilon^2 \tilde{A}_n^{(3)}(\underline{\gamma}_q) \cdot I(T_{n-k+1:n} = \underline{\gamma}_q), \end{aligned} \tag{40}$$

where  $\tilde{A}_n^{(3)}(\cdot)$  is as in part (3) of Lemma 3. Clearly, inequality (40) still holds if we replace “ $\underline{\gamma}_q$ ” by “ $\bar{\gamma}_q$ ”. Hence, setting  $\hat{v}_\theta = v_\theta(\underline{\gamma}_q) \cdot I(T_{n-k+1:n} = \underline{\gamma}_q) + v_\theta(\bar{\gamma}_q) \cdot$

$I(T_{n-k+1:n} = \bar{\gamma}_q)$  and using (39) we obtain, with  $P_\theta$ -probability 1, for all sufficiently large  $n$ ,

$$\begin{aligned} & \left| \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} - \varepsilon \hat{v}_\theta \right| = \left| \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} - \varepsilon v_\theta(\underline{\gamma}_q) \right| \\ & \cdot I(T_{n-k+1:n} = \underline{\gamma}_q) + \left| \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} - \varepsilon v_\theta(\bar{\gamma}_q) \right| \cdot I(T_{n-k+1:n} = \bar{\gamma}_q) \\ & \leq |A_n^{(1)}(\underline{\gamma}_q)| + |A_n^{(1)}(\bar{\gamma}_q)| + \varepsilon \left( |A_n^{(2)}(\underline{\gamma}_q) + v_\theta(\underline{\gamma}_q)| + |A_n^{(2)}(\bar{\gamma}_q) + v_\theta(\bar{\gamma}_q)| \right) \\ & + \frac{1}{2} \varepsilon^2 \left( \tilde{A}_n^{(3)}(\underline{\gamma}_q) + \tilde{A}_n^{(3)}(\bar{\gamma}_q) \right) \xrightarrow{P_\theta\text{-a.s.}} \frac{1}{2} \varepsilon^2 \left( w_\theta(\underline{\gamma}_q) + w_\theta(\bar{\gamma}_q) \right) \leq \frac{1}{2} \varepsilon \min \left( v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q) \right), \end{aligned}$$

where the above convergence is due to parts (2) and (3) of Lemma 3. Consequently, with  $P_\theta$ -probability 1, for all sufficiently large  $n$ ,

$$\left| \frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} - \varepsilon \hat{v}_\theta \right| \leq \frac{3}{4} \varepsilon \min \left( v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q) \right),$$

which yields

$$\frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_1} \geq -\frac{3}{4} \varepsilon \min \left( v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q) \right) + \varepsilon \hat{v}_\theta \geq \frac{1}{4} \varepsilon \min \left( v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q) \right). \tag{41}$$

Similarly, we can show that, with  $P_\theta$ -probability 1, for all sufficiently large  $n$ ,

$$\frac{1}{n} \frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) \Big|_{\lambda=\lambda_2} \leq -\frac{1}{4} \varepsilon \min \left( v_\theta(\underline{\gamma}_q), v_\theta(\bar{\gamma}_q) \right). \tag{42}$$

Since  $I(\lambda; T_1, \dots, T_n)$  is a continuous function with respect to  $\lambda$ , (41) and (42) show that, with  $P_\theta$ -probability 1, for all sufficiently large  $n$ , the likelihood equation  $\frac{\partial}{\partial \lambda} I(\lambda; T_1, \dots, T_n) = 0$  has at least one solution  $\lambda \in (\theta - \varepsilon, \theta + \varepsilon)$ .

Now, repeating the same arguments as that used in Section 4.2.2 of Serfling (2001) we can construct a sequence of estimators  $(\hat{\theta}_n, n \geq 1)$  satisfying conditions from the conclusion of Theorem 1. □

### References

Barlow RE, Proschan F (1996) Mathematical theory of reliability. SIAM, New York  
 Bhattacharyya GK (1985) The asymptotics of maximum likelihood and related estimators based on Type II censored data. J Am Stat Assoc 80:398–404  
 Bickel PJ, Doksum KA (1977) Mathematical statistics: basic ideas and selected topics. Holden-Day, San Francisco  
 Davies K, Dembińska A (2019) On the number of failed components in a  $k$ -out-of- $n$  system upon system failure when the lifetimes are discretely distributed. Reliab Eng Syst Saf 188:47–61  
 Dembińska A (2012) Limit theorems for proportions of observations falling into random regions determined by order statistics. Aust N Z J Stat 54:199–210  
 Dembińska A (2014) Asymptotic behavior of central order statistics from stationary processes. Stoch Process Appl 124:348–372

- Dembińska A (2018) On reliability analysis of  $k$ -out-of- $n$  systems consisting of heterogeneous components with discrete lifetimes. *IEEE Trans Reliab* 67:1071–1083
- Dembińska A, Goroncy A (2020) Moments of order statistics from DNID discrete random variables with application in reliability. *J Comput Appl Math* 371:112703
- Dembińska A, Nikolov NI, Stoimenova E (2019) Reliability properties of  $k$ -out-of- $n$  systems with one cold standby unit. [arXiv:1911.11818](https://arxiv.org/abs/1911.11818)
- Halperin M (1952) Maximum likelihood estimation in truncated samples. *Ann Math Stat* 23:226–238
- Gan G, Bain LJ (1995) Distribution of order statistics for discrete parents with applications to censored sampling. *J Stat Plan Inference* 44:37–46
- Hermanns M, Cramer E (2018) Inference with progressively censored  $k$ -out-of- $n$  system lifetime data. *TEST* 27:787–810
- Kong F, Fei H (1996) Limit theorems for the maximum likelihood estimate under general multiply Type II censoring. *Ann Inst Stat Math* 48:731–755
- Lin CT, Balakrishnan N (2011) Asymptotic properties of maximum likelihood estimators based on progressive Type-II censoring. *Metrika* 74:349–360
- Navarro J, Ng HKT, Balakrishnan N (2012) Parametric inference for component distributions from lifetimes of systems with dependent components. *Naval Res Logist* 59:487–496
- Ng HKT, Navarro J, Balakrishnan N (2012) Parametric inference from system lifetime data with signatures available under a proportional hazard rate model. *Metrika* 75:367–388
- Pólya G, Szegő G (1998) *Problems and theorems in analysis*. Springer, Berlin
- Smirnov NV (1952) Limit distributions for the terms of a variational series. *Am Math Soc Transl Ser* 1(11):82–143
- Serfling RJ (2001) *Approximation theorems of mathematical statistics*. Wiley, New York
- Schoenberg IJ (1959) On variation diminishing approximation methods. In: Langer RE (ed) *On numerical approximation*. University of Wisconsin Press, Madison Press, Madison
- Tank F, Eryilmaz S (2015) The distributions of sum, minima and maxima of generalized geometric random variables. *Stat Pap* 56:1191–1203
- Weiss G (1962) On certain redundant systems which operate at discrete times. *Technometrics* 4:169–174
- Young D (1970) The order statistics of the negative binomial distribution. *Biometrika* 57:181–186

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.