Ann Inst Stat Math (2016) 68:639–657
DOI 10.1007/s10463-015-0508-y

Inference in a model of successive failures with shape-adjusted hazard rates

Stefan Bedbur · Marco Burkschat · Udo Kamps

Received: 11 February 2014 / Revised: 27 June 2014 / Published online: 3 March 2015

Abstract For successive failure times of components in a technical system, a flexible model based on sequential order statistics is proposed. Beyond the common assumption of proportionality, this model allows for structural adjustments of the hazard rates of the underlying lifetime distributions in situations, where failures have an impact on the entire shape of the hazard rate of remaining components. The hazard rates may be chosen, e.g., as strictly ordered bathtub curves. The general structure is analysed, and maximum likelihood estimators are stated for both, unrestricted and order restricted model parameters, as well as for parameters connected by a linear link function. Several properties of the estimators are obtained. Utilizing the maximum likelihood estimators, simultaneous confidence regions based on the Jeffreys–Kullback–Leibler distance and the Hellinger distance are examined.

Keywords Sequential order statistic · Load sharing system · Proportional hazard rate · Bathtub shape · Maximum likelihood estimation · Order restricted inference · Link function · Confidence region · Distance measure

1 Introduction

In technical systems consisting of several units, the failure of a component may have an impact on the lifetimes of the surviving components. For describing the ordered

S. Bedbur · M. Burkschat · U. Kamps (✉)
Institute of Statistics, RWTH Aachen University, 52056 Aachen, Germany
e-mail: udo.kamps@rwth-aachen.de

S. Bedbur
e-mail: stefan.bedbur@rwth-aachen.de

M. Burkschat
e-mail: marco.burkschat@rwth-aachen.de

lifetimes of the components in such a system, sequential order statistics can be applied. This model of ordered data allows for the description of failure situations, where, upon a failure of some component, the remaining components are supposed to have a possibly different underlying lifetime distribution (see Kamps 1995a, b). If the different lifetime distribution functions are denoted by $F_1, F_2, \ldots$, then this leads to the interpretation that, after the jth failure, the hazard rate of (remaining) components will change from λ_j to λ_{j+1}, where $\lambda_j = f_j/(1 - F_j)$ denotes the hazard rate and f_j the density function of F_j, $j = 1, 2, \ldots$.

In the literature on sequential order statistics, it is usually assumed that there exists some common baseline hazard rate h with

$$\lambda_j = \vartheta_j h, \quad j = 1, 2, \ldots,$$

for positive constants $\vartheta_1, \vartheta_2, \ldots$ and, consequently, the hazard rates λ_j are proportional to each other (cf., e.g., Cramer and Kamps 1996, 2001a, b; Belzunce et al. 2005; Hu and Zhuang 2005; Beutner and Kamps 2009; Balakrishnan et al. 2008, 2010, 2011; Bedbur 2010; Burkschat et al. 2010; Deshpande et al. 2010; Bedbur et al. 2012). However, in practical applications, the failure of a unit may lead to a change in the shape of the hazard rate. Consider, for instance, a bathtub shape of the hazard rate, which may be encountered in the case of electronic products (see, e.g., Ch. 3 Lai and Xie 2006). If the load on a component increases, then we may expect not only a larger hazard rate, but also that the flat part of the bathtub curve becomes shorter (cf. Fig. 1 in Gurgenci and Guan 2001 in the context of increasing duty levels of equipment). Corresponding shape adjustments of the hazard rates are illustrated in Fig. 1 for bathtub curves by utilizing a modified Weibull distribution of Lai et al. (2003) and in Fig. 2 for increasing hazard rate curves of Gompertz distributions. Obviously, such differences in the shape cannot be described by proportional hazard rates.

In the present paper, we focus on estimation in the general model of sequential order statistics without imposing proportional hazard rates. We assume that the hazard rates λ_j are given in the form

$$\lambda_j = \vartheta_j h_j, \quad j = 1, 2, \ldots,$$

with unknown positive parameters $\vartheta_1, \vartheta_2, \ldots$ of interest, and pre-fixed, possibly different functions $h_1, h_2, \ldots$. Due to the assumed increase of the load, the hazard rates should be modelled as being ordered, i.e., $\lambda_j(t) \le \lambda_{j+1}(t)$ for $t > 0$. Since the difference between the functions $h_1, h_2, \ldots$ may become arbitrarily small (see, e.g., Figs. 1, 3 as well as Sect. 3.2), the parameters $\vartheta_1, \vartheta_2, \ldots$ should also be assumed to be increasingly ordered. Then, ordered functions $h_1, h_2, \ldots$, i.e., $h_j(t) \le h_{j+1}(t)$ for $t > 0$, will guarantee ordered hazard rates λ_j. The general shape of the hazard rate λ_j, that is, of the function h_j, may be known from previous experiments. In addition to the bathtub hazard rate curves to be ordered (see, e.g., the practical example in Gurgenci and Guan 2001) and the operating periods becoming shorter, the early failure period as well as the wear-out period may be differently modelled such as being characterized by a less steep decrease or a more steep increase, respectively (see, e.g., the hazard rates given in Fig. 1). Alternatively, hazard rates may be desired that possess the same

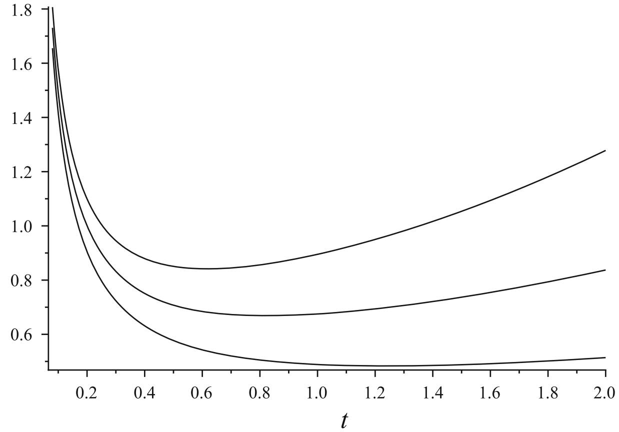

Fig. 1 Hazard rates of the modified Weibull distribution $F(t) = 1 - \exp[-t^{\beta} \exp(\delta t)]$ by Lai et al. 2003 with $\beta = 0.2$ and (in increasing order of the curves) $\delta_1 = 0.2$, $\delta_2 = 0.3$, and $\delta_3 = 0.4$

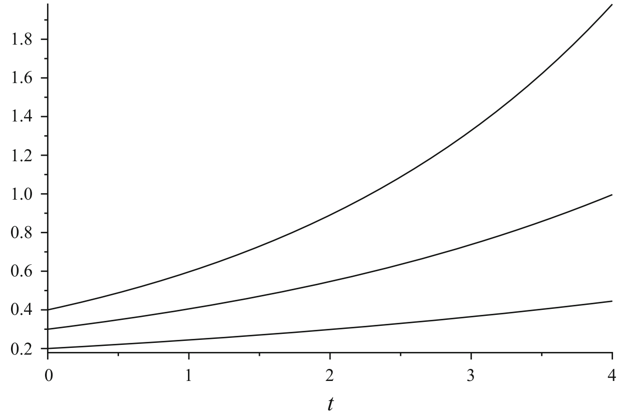

Fig. 2 Hazard rates of the Gompertz distribution $F(t) = 1 - \exp[-(\exp(\delta t) - 1)]$ with (in increasing order of the curves) $\delta_1 = 0.2$, $\delta_2 = 0.3$, and $\delta_3 = 0.4$

limiting behaviour for large t (see, e.g., Fig. 3 for an example based on the Hjorth distribution). For general motivations and reviews on bathtub-shaped hazard rates, we refer to Chapter 33 in Johnson et al. (1995), to Chapter 3 in Lai and Xie (2006) and to Chapter 4 in Marshall and Olkin (2007). For systems with a given number of components there may be additional information available on the time period, when the hazard rate of a contained component is usually nearly flat. However, the actual stress levels $\vartheta_1, \vartheta_2, \ldots$ need to be estimated.

The remaining part of the article is organized as follows. In Sect. 2, we introduce the model along with its exponential family structure and address further properties, which turn out to be helpful in subsequent sections when dealing with statistical inference for unknown parameters. Maximum likelihood estimation of the model parameters $\vartheta_1, \vartheta_2, \ldots$ and related properties are shown in Sect. 3. In particular, estimation under

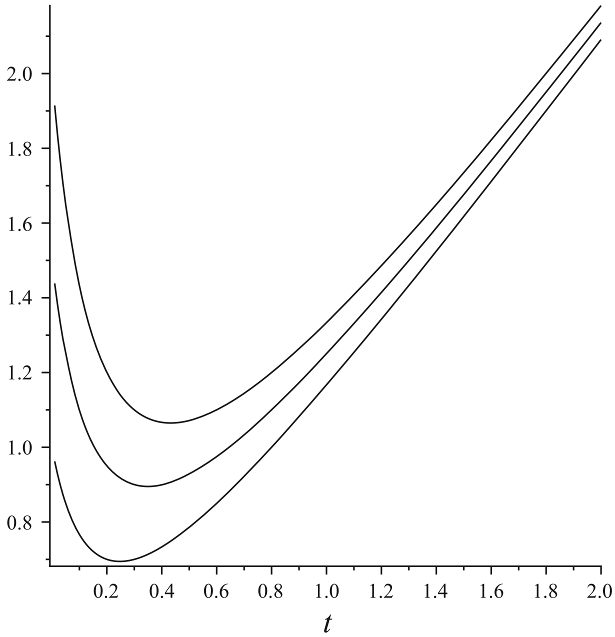

Fig. 3 Hazard rates of the Hjorth distribution $F(t) = 1 - (1+\beta t)^{\delta} \exp(-t^2/2)$ with $\beta = 5$ and (in increasing order of the curves) $\delta_1 = 0.2$, $\delta_2 = 0.3$, and $\delta_3 = 0.4$

simple order restriction is considered in Sect. 3.2. If the parameters are assumed to fulfil a linear relationship, estimation of the remaining two unknown quantities of the link function is subject matter of Sect. 3.3. Finally, in Sect. 4, we focus on simultaneous confidence regions for the model parameters based on two distance measures, the Jeffreys–Kullback–Leibler distance and the Hellinger distance.

2 Model and structural properties

Based on absolutely continuous distribution functions $F_1, \ldots, F_n$ satisfying $F_1^{-1}(1) \leq \ldots \leq F_n^{-1}(1)$ with corresponding density functions $f_1, \ldots, f_n$, the joint density function of the first r $(\leq n)$ sequential order statistics $X_{1,n}^*, \ldots, X_{r,n}^*$ (SOSs) is represented by

$$\frac{n!}{(n-r)!} \prod_{j=1}^{r} \left(\frac{1 - F_j(x_j)}{1 - F_j(x_{j-1})} \right)^{n-j} \frac{f_j(x_j)}{1 - F_j(x_{j-1})} \tag{1}$$

on the cone $-\infty = x_0 < x_1 \leq \cdots \leq x_r$. These SOSs form a Markov chain with transition probabilities

$$P\left(X_{j,n}^* > t \mid X_{j-1,n}^* = s\right) = \left(\frac{1 - F_j(t)}{1 - F_j(s)} \right)^{n-j+1}, \quad t \geq s, \quad 2 \leq j \leq n \tag{2}$$

(cf. Kamps 1995a, b).

In the following, we focus on a particular representation of $F_1, \ldots, F_n$, namely

$$F_j(t) = 1 - \exp\left\{-\vartheta_j H_j(t)\right\}, \quad t \in (c, d), \quad 1 \le j \le n, \tag{3}$$

where $\vartheta_1, \ldots, \vartheta_n$ denote positive parameters, and $H_j : (c, d) \to (0, \infty)$ are cumulative hazard rates, i.e., increasing and absolutely continuous functions on (c, d), $0 \le c < d \le \infty$, with one-sided limits $H_j(c+) = 0$ and $H_j(d-) = \infty$ for $1 \le j \le n$ (typically, $(c, d) = (0, \infty)$). The corresponding hazard rates h_j are given by $h_j(t) = H_j'(t)$, $t \in (c, d)$. Then, the hazard rate of F_j turns out to be

$$\lambda_j(t) = \vartheta_j h_j(t), \quad t \in (c, d), \quad 1 \le j \le n,$$

allowing for modelling increasing stress of remaining components of a system beyond the assumption of proportional hazard rates. The respective conditional hazard rate associated with (2) is thus given by

$$(n - j + 1)\, \lambda_j(t), \quad 2 \le j \le n .$$

This can be interpreted such that the underlying failure rate successively changes from λ_{j-1} to λ_j immediately after failure $j-1$ has occurred (cf. Cramer and Kamps 2001b). Thus, upon each failure, the remaining components are supposed to operate under a possibly different failure rate which, e.g., takes an increased load put on remaining components or their partial damage into account. As a submodel, the proportional hazard rates result by choosing the Hs identically. More precisely, the usual proportional hazard rate model with absolutely continuous baseline distribution function F, respective density function f, and model parameters $\vartheta_1, \ldots, \vartheta_n$ is obtained by setting $H_1 = \ldots = H_n = -\log(1 - F)$ on the domain $(c, d) = (F^{-1}(0+), F^{-1}(1))$. In this particular setting and due to the above interpretation, the SOSs form a model with conditional proportional hazard rates, and

$$\lambda_j(t) = \vartheta_j \frac{f(t)}{1 - F(t)}, \quad 2 \le j \le n .$$

In Table 1, some distribution functions of the type (3), i.e., of the form $F(t) = 1 - \exp\{-\vartheta H(t)\}$, are shown along with some properties. In particular, w.r.t. at least one of their parameters, the hazard rates are ordered as required in our model (cf. Sect. 1); increasing and decreasing shapes are denoted by $\nearrow$ and $\searrow$, respectively. Moreover, the shapes of hazard rates are described by increasing failure rate (IFR), decreasing failure rate (DFR), bathtub shape (BT), and upside down bathtub shape (UBT). For a survey on the Gompertz distribution in row 1 of Table 1, we refer to Marshall and Olkin (2007), Chapter 10; distribution 2 is introduced and studied in Lai et al. (2003); distribution 3 is a particular parametrization of a family of distributions shown in Hjorth (1980) (see also Johnson et al. 1995, pp. 664, 645); for $\delta \ge 2$, distribution 4 is an example of a family of distributions considered in Gaver and Acar 1979; for the Burr XII distribution (row 6) we also refer to Mudholkar et al. (1996), and

Table 1 Distribution functions of the type $F(t) = 1 - \exp\{-\vartheta H(t)\}$ with parameters $\beta > 0,\ \delta > 0$, and associated properties

	$H(t)$	Domain	Distribution	$h(t)$	Ageing properties	Hazard rate ordering
1	$e^{\delta t} - 1$	$(0, \infty)$	Gompertz (double exponential)	$\delta e^{\delta t}$	IFR	$\nearrow$ in δ
2	$t^{\beta} e^{\delta t}$	$(0, \infty)$	Modified Weibull	$t^{\beta-1} e^{\delta t}(\beta + \delta t)$	$\beta \in (0, 1)$: BT $\beta \geq 1$: IFR	$\nearrow$ in δ
3	$\frac{t^2}{2} + \delta \log(1 + \beta t)$	$(0, \infty)$	Hjorth	$t + \frac{\beta\delta}{1+\beta t}$	$\beta^2\delta \in (0, 1]$: IFR $\beta^2\delta > 1$: BT	$\nearrow$ in β, $\nearrow$ in δ
4	$\beta t^{\delta} + \log(t + 1)$	$(0, \infty)$		$\beta\delta t^{\delta-1} + \frac{1}{t+1}$	$\delta \in (0, 1]$: DFR $\delta \geq 2$: BT	$\nearrow$ in β
5	$\log\left(1 + \frac{t-\delta}{\beta}\right)$	(δ, ∞)	Lomax	$(t + \beta - \delta)^{-1}$	DFR	$\searrow$ in β, $\nearrow$ in δ
6	$\log\left(1 + \frac{t^{\delta}}{\beta}\right)$	$(0, \infty)$	Burr XII	$\dfrac{\delta t^{\delta-1}}{\beta + t^{\delta}}$	$\delta \in (0, 1]$: DFR $\delta > 1$: UBT	$\searrow$ in β
7	$-\log\left(1 - t^{\delta}\right)$	$(0, 1)$	Particular generalized Weibull, exponentiated power function	$\dfrac{\delta t^{\delta-1}}{1 - t^{\delta}}$	$\delta \in (0, 1)$: BT $\delta \geq 1$: IFR	$\searrow$ in δ

Table 2 Ordering $F_1 \leq_{af} F_2$ of distribution functions F_1 and F_2 as in Table 1

Distribution condition	1	2	3	4	5	6	7
	$\delta_1 \geq \delta_2$	$\delta_1 \geq \delta_2$	$\delta_1 \leq \delta_2$	$\beta_1 \geq \beta_2$ if $\delta_1 = \delta_2 \geq 1$	$\beta_1 \geq \beta_2$	$\beta_1 \geq \beta_2$	$\delta_1 \geq \delta_2$

Nikulin and Haghighi (2006) in a different parametrization; for a particular case of the generalized Weibull distribution in row 7 see Mudholkar et al. (1996), and for $\delta = 1$ the distribution is also known as J-shaped beta.

Clearly, the shapes of the particular underlying distribution functions $F_1, F_2, \ldots,$ also determine the shapes of the hazard rates of SOSs. For example, we consider the IFR property of SOSs. In Burkschat and Navarro (2011) and Torrado et al. (2012), a corresponding condition is given, which can be expressed in terms of a notion of relative ageing. We write $F_1 \leq_{af} F_2$ for distribution functions F_1 and F_2 of the type (3) if the ratio λ_1/λ_2 of their hazard rates (or equivalently the ratio h_1/h_2) is an increasing function on (c, d). For further information on this notion, see, e.g., Sengupta and Deshpande (1994) and Rowell and Siegrist (1998). In Burkschat and Navarro (2011) and Torrado et al. (2012), it is shown that the rth SOS $X^*_{r,n}$ has the IFR property if $F_1 \leq_{af} F_2 \leq_{af} \ldots \leq_{af} F_r$ and F_r is IFR.

Table 2 illustrates that, upon choosing an appropriate parameter, we may find $F_1 \leq_{af} F_2 \leq_{af} \ldots$ for distribution functions of the type (3). In the table, the relations are summarized for F_1 and F_2 as in Table 1 with parameters $\vartheta_1, \beta_1, \delta_1$ and $\vartheta_2, \beta_2, \delta_2$, respectively. In all cases, the respective other parameters (except for ϑ_1 and ϑ_2) are assumed to be identical. Hence, as an example, the preceding result can be applied to distributions 1, 2 ($\beta \geq 1$), 3 ($\beta^2\delta \leq 1$) and 7 ($\delta \geq 1$) in Table 1 to conclude that some SOS possesses the IFR property. Related conditions such that SOSs have a decreasing reversed hazard rate can be found in Burkschat and Torrado (2014). Moreover, further results on stochastic orderings in the general setting of SOSs are given in Zhuang and Hu (2007).

We now focus on the joint distribution of SOSs in the present situation. By inserting formula (3) in (1), the joint density function of the first r SOSs $X^*_{1,n}, \ldots, X^*_{r,n}$ based on $F_1, \ldots, F_n$ can be rewritten in the form

$$\left(\prod_{j=1}^{r} \vartheta_j\right) \exp\left\{\vartheta_1 \left[-n\, H_1(x_1)\right] + \sum_{j=2}^{r} \vartheta_j \left[-(n-j+1)\left(H_j(x_j) - H_j(x_{j-1})\right)\right]\right\}$$
$$\times \left(\frac{n!}{(n-r)!} \prod_{j=1}^{r} h_j(x_j)\right)$$

on the support $\mathcal{X}_r = \{(x_1, \ldots, x_r) \in (c, d)^r : x_1 \leq \ldots \leq x_r\}$ which leads to the short representation

$$f_{\boldsymbol{\vartheta}}(\boldsymbol{x}) = \exp\left\{\boldsymbol{\vartheta}^t \boldsymbol{T}(\boldsymbol{x}) - \kappa(\boldsymbol{\vartheta})\right\} b(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathcal{X}_r\,, \tag{4}$$

with vector $\boldsymbol{T} = (T_1, \ldots, T_r)^t$ of statistics on $\mathcal{X}_r$ defined as

$$T_1(\boldsymbol{x}) = -n\,H_1(x_1)\,,$$
$$T_j(\boldsymbol{x}) = -(n-j+1)\,\left(H_j(x_j) - H_j(x_{j-1})\right)\,,\quad 2 \le j \le r\,,$$

and mappings

$$b(\boldsymbol{x}) = \frac{n!}{(n-r)!}\prod_{j=1}^{r} h_j(x_j)\,,\quad \boldsymbol{x} = (x_1, \ldots, x_r) \in \mathcal{X}_r\,,$$

and

$$\kappa(\boldsymbol{\vartheta}) = -\sum_{j=1}^{r}\log\left(\vartheta_j\right)\,,\quad \boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_r)^t \in \mathbb{R}_+^r \equiv (0, \infty)^r\,. \tag{5}$$

Here and in the following, superscript t on a vector or a matrix denotes its transpose.

In the context of the common proportional hazard rate model, the exponential family structure of the joint density function of the first r SOSs has been recognized in Bedbur (2010) and Bedbur et al. (2012). It turns out that in the present more flexible model, this structure appears again and useful inferential results are near at hand as we will exemplarily demonstrate in Sect. 3 in the framework of point estimation of the ϑs.

For this purpose, we first state some fundamental results with respect to the recognized exponential family structure.

Theorem 1 *Let $\mathcal{P} = \{f_{\boldsymbol{\vartheta}}\,\mu_r : \boldsymbol{\vartheta} \in \mathbb{R}_+^r\}$ be the exponential family of distributions with densities as in (4), where μ_r denotes the r-dimensional Lebesgue measure on the Borel sets of $\mathcal{X}_r$. Then, we obtain the following properties:*

1, $\mathcal{P}$ is full and regular.
2. The statistics $T_1, \ldots, T_r$ on $\mathcal{X}_r$ are independent, and $-T_j$ is exponentially distributed with mean $1/\vartheta_j$ for $1 \le j \le r$.
3. $\mathcal{P}$ is minimal and of full rank.
4. $\boldsymbol{T}$ is minimal sufficient and complete for $\mathcal{P}$.

Proof 1. Let $\Theta^* = \{\boldsymbol{\vartheta} : 0 < \int_{\mathcal{X}_r}\exp\{\boldsymbol{\vartheta}^t\boldsymbol{T}\}\,b\,d\mu_r < \infty\}$ denote the natural parameter space of $\mathcal{P}$. Evidently, $\mathbb{R}_+^r \subseteq \Theta^*$. Now, let $\boldsymbol{\vartheta} \in [0, \infty)^r$ with $\vartheta_j = 0$ for (at least) one index $j \in \{1, \ldots, r\}$, and let j_0 be maximal with that property. Then, by denoting the one-dimensional Lebesgue measure by ν,

$$\int_{(x_{j-1}, d_j)} \exp\left\{\vartheta_j T_j(\boldsymbol{x})\right\}\,h_j(x_j)\,d\nu(x_j) \;=\; \begin{cases} \frac{1}{(n-j+1)\vartheta_j}\,, & j_0 < j \le r, \\ \infty, & j = j_0\,, \end{cases}$$

and applying Fubini's theorem (for non-negative measurable functions) leads to

$$\int_{\mathcal{X}_r} \exp\left\{\boldsymbol{\vartheta}^t \boldsymbol{T}(\boldsymbol{x})\right\} b(\boldsymbol{x})\, d\mu_r(\boldsymbol{x})$$

$$= \frac{n!}{(n-r)!} \int_{\mathcal{X}_r} \prod_{j=1}^{r} \left[\exp\{\vartheta_j T_j(\boldsymbol{x})\} h_j(x_j)\right] d\mu_r(x_1, \ldots, x_r)$$

$$= \frac{n!}{(n-j_0)!} \left(\prod_{j=j_0+1}^{r} \frac{1}{\vartheta_j}\right) \int_{\mathcal{X}_{j_0-1}} \prod_{j=1}^{j_0-1} \left[\exp\{\vartheta_j T_j(\boldsymbol{x})\} h_j(x_j)\right]$$

$$\times \left(\int_{(x_{j_0-1}, d_{j_0})} h_j(x_{j_0})\, d\nu(x_{j_0})\right) d\mu_{j_0-1}(x_1, \ldots, x_{j_0-1})$$

$$= \infty.$$

Hence, $\boldsymbol{\vartheta} \notin \Theta^*$. As a consequence, every $\boldsymbol{\vartheta}$ having some negative component cannot be an element of Θ^*, since the natural parameter space of an exponential family is always convex (see, e.g., Lemma 2.7.1 in Lehmann and Romano 2005). Thus, $\Theta^* = \mathbb{R}_+^r$ is shown implying that $\mathcal{P}$ is full and, moreover, regular (since Θ^* is open).

2. The statement follows directly from the fact that, by utilizing the exponential family structure, the moment generating function of $-\boldsymbol{T}$ is obtained according to

$$E_{\boldsymbol{\vartheta}} \exp\left\{\boldsymbol{z}^t(-\boldsymbol{T})\right\} = \exp\left\{\kappa(\boldsymbol{\vartheta} - \boldsymbol{z}) - \kappa(\boldsymbol{\vartheta})\right\} = \prod_{j=1}^{r} \left(1 - \frac{z_j}{\vartheta_j}\right)^{-1},$$

for $\boldsymbol{z} = (z_1, \ldots, z_r)^t \in \mathbb{R}^r$ with $z_j < \vartheta_j$, $1 \leq j \leq r$ (see., e.g., Kotz et al. 2000, p. 664; cf. Bedbur et al. 2012).

3. Statement (2) implies that $\mathbf{Cov}_{\boldsymbol{\vartheta}}(\boldsymbol{T}) = \mathbf{diag}(1/\vartheta_1^2, \ldots, 1/\vartheta_r^2)$ is positive definite. Thus, the statistics $T_1, \ldots, T_r$ do not satisfy a linear constraint, and it follows that (the present representation of) $\mathcal{P}$ is minimal. Moreover, since the interior of $\Theta^* = \mathbb{R}_+^r$ is not empty, $\mathcal{P}$ is of full rank.
4. The statement follows directly from (3) in combination with Corollary 1.6.16 and Theorem 1.6.22 in Lehmann and Casella (1998).

□

As a preliminary work for the following sections, we briefly address the corresponding properties related to the product case.

Let $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(s)}$ be a sample of independent random vectors with density function (4), each. Then, the overall joint density function is given by

$$f_{\boldsymbol{\vartheta}}^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right) = \exp\left\{\boldsymbol{\vartheta}^t \boldsymbol{T}^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) - s\kappa(\boldsymbol{\vartheta})\right\} \prod_{i=1}^{s} b(\boldsymbol{x}^{(i)}),\ \tilde{\boldsymbol{x}}^{(s)} = \left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(s)}\right) \in \mathcal{X}_r^s, \tag{6}$$

with vector $\boldsymbol{T}^{(s)} = (T_1^{(s)}, \ldots, T_r^{(s)})^t$ of statistics on $\mathcal{X}_r^s$, which are given by

$$T_j^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right) = \sum_{i=1}^{s} T_j\left(\boldsymbol{x}^{(i)}\right), \quad \tilde{\boldsymbol{x}}^{(s)} = \left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(s)}\right) \in \mathcal{X}_r^s, \quad 1 \le j \le r.$$

Consequently, denoting by $\mu_r^s = \otimes_{i=1}^s \mu_r$ the s-fold product measure of μ_r, the family $\mathcal{P}^{(s)} = \{f_{\boldsymbol{\vartheta}}^{(s)} \mu_r^s : \boldsymbol{\vartheta} \in \mathbb{R}_+^r\}$ of joint distributions of $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(s)}$ forms a multivariate exponential family in $\vartheta_1, \ldots, \vartheta_r$ and statistics $T_1^{(s)}, \ldots, T_r^{(s)}$. Replacing $\boldsymbol{T}$ by $\boldsymbol{T}^{(s)}$ and $\mathcal{P}$ by $\mathcal{P}^{(s)}$, it is seen that Theorem 1 remains valid with the only difference that, for $1 \le j \le r$, the statistic $-T_j^{(s)} \sim \Gamma(s, 1/\vartheta_j)$ now follows a gamma distribution with density function

$$\frac{\vartheta_j^s}{(s-1)!} u^{s-1} \exp\left\{-\vartheta_j u\right\}, \quad u > 0. \tag{7}$$

Remark 1 The above model with non-proportional hazard rates may also be applied as a particular step-stress scheme within accelerated life testing (cf. Bagdonavičius and Nikulin 2002). Such schemes are applied for testing highly reliable products, where too few or even no failures would occur under normal operating conditions (cf., e.g., Balakrishnan 2009 and the references therein). The proportional hazard rates case with $F_1 = F_2 = \ldots$ is studied in Balakrishnan et al. (2012). In a multiple sample setup, a number of s (possibly type-II right censored) lifetime tests are run with n items put on test in experiment i, $1 \le i \le s$. Then, immediately after every failure in experiment i, the underlying level of stress is increased to cause earlier failures; in sample i, $r \le n$ failures are supposed to be observed, $1 \le i \le s$. An increasing stress level may be implemented by choosing the underlying distribution function F_{j+1} after failure j in the ith sample such that its hazard rate h_{j+1} (strictly) exceeds h_j, i.e. $h_{j+1}(t) \ge (>)\, h_j(t)$ for all $t > 0$, and assuming additionally $\vartheta_{j+1} \ge \vartheta_j$. Families of distributions with this property are shown in Table 1. For a model in the multiple samples case by means of common order statistics, we refer to Kateri et al. (2009, 2010).

3 Point estimation

When applying the model of SOSs with distribution functions according to (3), the model parameters $\vartheta_1, \vartheta_2, \ldots$ and/or the functions $H_1, H_2, \ldots$ will usually be unknown and have to be estimated based on data as in (6).

Throughout this section, the functions $H_1, \ldots, H_r$ are supposed to be known (e.g., specified based on some previous experiment or prior information about the shapes of the corresponding hazard rates), and we assume that the uncertainty of the model is totally captured within the vector $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_r)^t$ of unknown model parameters. Starting from sample situation (6), point estimators of $\boldsymbol{\vartheta}$ with desirable properties are presented, where optimality properties are also obtained.

3.1 Unrestricted maximum likelihood estimation

First, point estimation without imposing any further conditions on the relationship of the ϑs or the hs is considered. The results are essentially based on the structural findings of Sect. 2, and are summarized in the following theorem.

Theorem 2 *Let sample situation (6) be given with pre-fixed functions $H_1, \ldots, H_r$. Then, we find:*

1. $\hat{\boldsymbol{\vartheta}}^{(s)} = (\hat{\vartheta}_1^{(s)}, \ldots, \hat{\vartheta}_r^{(s)})^t$ *with* $\hat{\vartheta}_j^{(s)} = -s/T_j^{(s)}$, $1 \leq j \leq r$, *is the unique maximum likelihood estimator (MLE) of* $\boldsymbol{\vartheta}$. *The univariate estimators* $\hat{\vartheta}_1^{(s)}, \ldots, \hat{\vartheta}_r^{(s)}$ *are jointly independent, and* $\hat{\vartheta}_j^{(s)}$ *has an inverse gamma distribution with density function*

$$\frac{(s\vartheta_j)^s}{(s-1)!}\left(\frac{1}{u}\right)^{s+1}\exp\left\{-\frac{s\vartheta_j}{u}\right\}, \quad u > 0, \quad \textit{for } 1 \leq j \leq r.$$

2. $-(s-1)/T_j^{(s)}$ *is the uniformly minimum variance unbiased estimator of* ϑ_j *for* $1 \leq j \leq r$ *and* $s > 2$.
3. *The vector* $(1/\hat{\vartheta}_1^{(s)}, \ldots, 1/\hat{\vartheta}_r^{(s)})^t$ *of reciprocals of the univariate MLEs estimates* $(1/\vartheta_1, \ldots, 1/\vartheta_r)^t$ *efficiently, i.e., its covariance matrix is minimum in the sense of the Löwner ordering among all unbiased estimators of* $(1/\vartheta_1, \ldots, 1/\vartheta_r)^t$ *based on* $\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(s)}$.
4. *The sequence* $\hat{\boldsymbol{\vartheta}}^{(s)}$, $s \in \mathbb{N}$, *is asymptotically unbiased* ($E_{\boldsymbol{\vartheta}}(\hat{\boldsymbol{\vartheta}}^{(s)}) \to \boldsymbol{\vartheta}$, $s \to \infty$), *strongly consistent* ($\hat{\boldsymbol{\vartheta}}^{(s)} \to \boldsymbol{\vartheta}$ *almost sure*, $s \to \infty$), *and asymptotically efficient* ($\sqrt{s}(\hat{\boldsymbol{\vartheta}}^{(s)} - \boldsymbol{\vartheta})$ *converges in distribution to a multivariate normal distribution with mean zero and covariance matrix* $\mathbf{diag}(\vartheta_1^2, \ldots, \vartheta_r^2)$ *which is the inverse of the Fisher information matrix of* $\mathcal{P}$ *at* $\boldsymbol{\vartheta}$).

Proof By means of the form (6) of the overall joint density function and the exponential family properties as shown in Theorem 1 in Sect. 2, it turns out that, although considering a more general model, we arrive at the same mathematical structure as in Bedbur et al. (2012) in the context of SOSs with conditional proportional hazard rates. Hence, substituting α_j by ϑ_j, Ψ by κ, h by $b, \ldots$, we can follow respective arguments therein to obtain the above statements. □

3.2 Maximum likelihood estimation under simple order restrictions

As has been indicated in Sect. 1, one may assume that the hazard rates are ordered, i.e., $\lambda_1(t) \leq \ldots \leq \lambda_r(t)$ for $t \in (c, d)$, and take this information into account when estimating the ϑs. If the functions $h_1, \ldots, h_r$ become arbitrarily close, e.g., if $h_{j+1}(t)/h_j(t) \to 1$ for t tending to a boundary point of the domain of H and $1 \leq j \leq r-1$, then the model parameters $\vartheta_1, \ldots, \vartheta_r$ have to be estimated under simple ascending order. An example thereof is the modified Weibull case with $H_j(t) = t^\beta e^{\delta_j t}$, $t > 0$, and parameters $\beta \in (0, 1)$, $0 < \delta_1 < \ldots < \delta_r$, where $h_{j+1}(t)/h_j(t) \to 1$

for $t \to 0+$ (see row 2 in Table 1; Fig. 1). Another example is provided by the Hjorth case with $H_j(t) = t^2/2 + \delta_j \log(1 + \beta t)$, $t > 0$, and parameters $\beta > 0$, $1/\beta^2 < \delta_1 < \ldots < \delta_r$, where $h_{j+1}(t)/h_j(t) \to 1$ for $t \to \infty$ (see row 3 in Table 1; Fig. 3).

Theorem 3 shows the MLE of $\boldsymbol{\vartheta}$ with respect to the constraint $\vartheta_1 \le \cdots \le \vartheta_r$.

Theorem 3 *Let sample situation (6) be given with pre-fixed functions $H_1, \ldots, H_r$. Then, we find:*

1. *The MLE of $\boldsymbol{\vartheta}$ under the simple order restriction $\vartheta_1 \le \ldots \le \vartheta_r$ uniquely exists and is given by $\tilde{\boldsymbol{\vartheta}}^{(s)} = (\tilde{\vartheta}_1^{(s)}, \ldots, \tilde{\vartheta}_r^{(s)})^t$ with components*

$$\tilde{\vartheta}_j^{(s)} = \min_{j \le w \le r} \max_{1 \le v \le j} \frac{w - v + 1}{\sum_{k=v}^{w} 1/\hat{\vartheta}_k^{(s)}}, \qquad 1 \le j \le r.$$

2. *$\tilde{\boldsymbol{\vartheta}}^{(s)}$, $s \in \mathbb{N}$, is strongly consistent provided that $\vartheta_1 \le \ldots \le \vartheta_r$.*

Proof 1. Using representations (5) and (6), the strict convexity of κ ensures that there exists at most one MLE of $\boldsymbol{\vartheta}$ within the convex set $\Delta = \{\boldsymbol{\vartheta} \in \Theta^* : \vartheta_1 \le \ldots \le \vartheta_r\}$. We introduce the bijective endofunction g on Δ in virtue of $g(\boldsymbol{\vartheta}) = (1/\vartheta_r, \ldots, 1/\vartheta_1)^t$. Then, following the same arguments as in the proof of Theorem 3.1.(1) in Balakrishnan et al. (2008), the MLE $\widetilde{g(\boldsymbol{\vartheta})}^{(s)} = (\widetilde{g(\boldsymbol{\vartheta})}_1^{(s)}, \ldots, \widetilde{g(\boldsymbol{\vartheta})}_r^{(s)})^t$ of $g(\boldsymbol{\vartheta})$ within Δ is given by

$$\widetilde{g(\boldsymbol{\vartheta})}_j^{(s)} = \max_{1 \le l \le j} \min_{j \le t \le r} \frac{\sum_{k=l}^{t} 1/\hat{\vartheta}_{r-k+1}^{(s)}}{t - l + 1}, \qquad 1 \le j \le r.$$

Hence, the MLE $\tilde{\boldsymbol{\vartheta}}^{(s)} = (\tilde{\vartheta}_1^{(s)}, \ldots, \tilde{\vartheta}_r^{(s)})^t$ of $\boldsymbol{\vartheta}$ within Δ exists and turns out to have the components

$$\begin{aligned} \tilde{\vartheta}_j^{(s)} &= \frac{1}{\widetilde{g(\boldsymbol{\vartheta})}_{r-j+1}^{(s)}} = \min_{1 \le l \le r-j+1} \max_{r-j+1 \le t \le r} \frac{t - l + 1}{\sum_{k=l}^{t} 1/\hat{\vartheta}_{r-k+1}^{(s)}} \\ &= \min_{j \le r-l+1 \le r} \max_{1 \le r-t+1 \le j} \frac{t - l + 1}{\sum_{k=r-t+1}^{r-l+1} 1/\hat{\vartheta}_k^{(s)}}, \qquad 1 \le j \le r. \end{aligned}$$

Substituting $r - l + 1$ and $r - t + 1$ by w and v, respectively, the proof of (1) is completed.

2. Strong consistency of the sequence of unrestricted MLEs [see Theorem 2 (4)] implies that $\tilde{\vartheta}_j^{(s)} \to \min_{j \le w \le r} \max_{1 \le v \le j} (w - v + 1)/\sum_{k=v}^{w} 1/\vartheta_k$ almost sure, $1 \le j \le r$, and, by elementary calculation, the limit can be computed to ϑ_j under the assumption that $\vartheta_1 \le \ldots \le \vartheta_r$.

□

3.3 Link functions

In small data situations, where simultaneous estimation of $\vartheta_1, \ldots, \vartheta_r$ is problematic, additional model assumptions may be required. On the other hand, prior information about the model parameters may happen to be available. In both cases, a link function could be part of the statistical model.

We assume that the parameters $\vartheta_1, \ldots, \vartheta_r$ are connected via the linear link function

$$\vartheta_j = \tau_1 + \tau_2 y_j, \quad j = 1, \ldots, r, \tag{8}$$

where $y_1, \ldots, y_r$ are fixed real numbers, and τ_1 and τ_2 are unknown model parameters. Under model assumption (8) it turns out that, again, an exponential family structure in the parameters τ_1 and τ_2 results, i.e., the right-hand side of equation (6) can be rewritten as

$$\tilde{f}^{(s)}_{\boldsymbol{\tau}}\left(\tilde{\boldsymbol{x}}^{(s)}\right) = \exp\left\{\boldsymbol{\tau}^t \tilde{\boldsymbol{T}}^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right) - s\tilde{\kappa}(\boldsymbol{\tau})\right\} \prod_{i=1}^{s} b\left(\boldsymbol{x}^{(i)}\right), \tilde{\boldsymbol{x}}^{(s)} = \left(\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(s)}\right) \in \mathcal{X}_r^s \tag{9}$$

with vector $\boldsymbol{\tau} = (\tau_1, \tau_2)^t$ of model parameters,

$$\tilde{\kappa}(\boldsymbol{\tau}) = -\sum_{j=1}^{r} \log(\tau_1 + \tau_2 y_j),$$

$$\boldsymbol{\tau} = (\tau_1, \tau_2)^t \in \tilde{\Theta} = \Big\{(a_1, a_2)^t \in \mathbb{R}^2 : a_1 + a_2 y_j > 0, \quad j = 1, \ldots, r\Big\},$$

and vector $\tilde{\boldsymbol{T}}^{(s)} = (\tilde{T}_1^{(s)}, \tilde{T}_2^{(s)})^t$ of statistics defined via

$$\tilde{T}_1^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right) = \sum_{j=1}^{r} T_j^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right), \quad \tilde{T}_2^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \sum_{j=1}^{r} y_j\, T_j^{(s)}\left(\tilde{\boldsymbol{x}}^{(s)}\right), \quad \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s.$$

Thus, $\tilde{\mathcal{P}}^{(s)} = \{\tilde{f}_{\boldsymbol{\tau}} \mu_r^s : \boldsymbol{\tau} \in \tilde{\Theta}\}$ forms a full and regular two-parameter exponential family in the natural parameters τ_1 and τ_2 and the statistics $\tilde{T}_1^{(s)}$ and $\tilde{T}_2^{(s)}$. Since $T_1^{(s)}, \ldots, T_r^{(s)}$ are jointly independent with $-T_j^{(s)} \sim \Gamma(s, 1/(\tau_1 + \tau_2 y_j))$ for $1 \leq j \leq r$ [see (7)], we obtain

$$\mathrm{Var}_{\boldsymbol{\tau}}\left(\tilde{T}_1^{(s)}\right) = s\sum_{j=1}^{r}\left(\tau_1 + \tau_2 y_j\right)^{-2}, \quad \mathrm{Var}_{\boldsymbol{\tau}}\left(\tilde{T}_2^{(s)}\right) = s\sum_{j=1}^{r} y_j^2 \left(\tau_1 + \tau_2 y_j\right)^{-2},$$

$$\text{and} \quad \mathrm{Cov}_{\boldsymbol{\tau}}\left(\tilde{T}_1^{(s)}, \tilde{T}_2^{(s)}\right) = s\sum_{j=1}^{r} y_j \left(\tau_1 + \tau_2 y_j\right)^{-2}. \tag{10}$$

The Cauchy–Schwarz inequality then says that $\mathbf{Cov}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{T}}^{(s)})$ is positive definite if and only if $|\{y_1, \ldots, y_r\}| \geq 2$. In that case, $\tilde{\mathcal{P}}^{(s)}$ is minimal and of full rank, and, as a consequence, $\tilde{\boldsymbol{T}}^{(s)}$ turns out to be minimal sufficient and complete for $\tilde{\mathcal{P}}^{(s)}$ by the same arguments as in the proof of Theorem 1 in Sect. 2.

Theorem 4 *Let sample situation* (9) *be given with pre-fixed functions* $H_1, \ldots, H_r$ *and pre-fixed real numbers* $y_1, \ldots, y_r$ *satisfying* $|\{y_1, \ldots, y_r\}| \geq 2$. *Then, we find:*

1. *The MLE* $\hat{\boldsymbol{\tau}}^{(s)} = (\hat{\tau}_1^{(s)}, \hat{\tau}_2^{(s)})^t$ *of* $\boldsymbol{\tau}$ *uniquely exists and is the only solution of the equations*

$$\tau_1 = -\frac{\tau_2 \tilde{T}_2^{(s)} + rs}{\tilde{T}_1^{(s)}} \quad \text{and} \quad \sum_{j=1}^{r} \frac{s}{\left(\tilde{T}_2^{(s)} - y_j \tilde{T}_1^{(s)}\right)\tau_2 + rs} = 1$$

with respect to $\boldsymbol{\tau} = (\tau_1, \tau_2)^t \in \tilde{\Theta}$.

2. *The sequence* $\hat{\boldsymbol{\tau}}^{(s)}$, $s \in \mathbb{N}$, *is strongly consistent and asymptotically efficient, i.e.,* $\sqrt{s}(\hat{\boldsymbol{\tau}}^{(s)} - \boldsymbol{\tau})$ *is asymptotically normal distributed with mean zero and covariance matrix given by the inverse matrix of* $\mathbf{Cov}_{\boldsymbol{\tau}}(\tilde{\boldsymbol{T}}^{(1)})$ *[see* (10)*].*

Proof 1. Let $\mathbf{A} \in \mathbb{R}^{r \times r}$ be a matrix of rank r with first row $(1, \ldots, 1)$ and second row $(y_1, \ldots, y_r)$. Due to the density transformation theorem, $\mathbf{A}\boldsymbol{T}^{(s)}$ and, in particular and more important, $\tilde{\boldsymbol{T}}^{(s)}$ have density functions with respect to r- and 2-dimensional Lebesgue measure, respectively. Consequently, applying Theorem 2.3.2 in Bickel and Doksum (2001) for $k = 2$, the MLE $\hat{\boldsymbol{\tau}}^{(s)}$ of $\boldsymbol{\tau}$ exists and is necessarily a solution of the equation $\nabla(s\tilde{\kappa})(\boldsymbol{\tau}) = \tilde{\boldsymbol{T}}^{(s)}$ with respect to $\boldsymbol{\tau} \in \tilde{\Theta}$, where $\nabla(s\tilde{\kappa})$ denotes the gradient of $s\tilde{\kappa}$. Note that from a general result for minimal regular exponential families, $\nabla(s\tilde{\kappa})$ is a (infinitely differentiable) diffeomorphism from $\tilde{\Theta}$ to the interior of the convex hull of the support of $\tilde{\boldsymbol{T}}^{(s)}$, where the latter is independent of $\boldsymbol{\tau}$ (see Kotz et al. 2000, pp. 667–669). Here, $\nabla(s\tilde{\kappa})(\boldsymbol{\tau}) = \tilde{\boldsymbol{T}}^{(s)}$ is equivalent to

$$\sum_{j=1}^{r} \frac{1}{\tau_1 + \tau_2 y_j} = -\frac{\tilde{T}_1^{(s)}}{s} \quad \text{and} \quad \sum_{j=1}^{r} \frac{y_j}{\tau_1 + \tau_2 y_j} = -\frac{\tilde{T}_2^{(s)}}{s}, \tag{11}$$

which implies that

$$\left(-\frac{\tilde{T}_1^{(s)}}{s}\right)\tau_1 + \left(-\frac{\tilde{T}_2^{(s)}}{s}\right)\tau_2 = r\,.$$

Solving this equation for τ_1 and substituting the term in the first equation of (11) yield statement (1).

2. Strong consistency of the sequence of MLEs is obtained from the representation $\hat{\boldsymbol{\tau}}^{(s)} = (\nabla\tilde{\kappa})^{-1}(\tilde{\boldsymbol{T}}^{(s)}/s)$ and the continuity of $(\nabla\tilde{\kappa})^{-1}$ by application of the strong law of large numbers. Asymptotic efficiency of $\hat{\boldsymbol{\tau}}^{(s)}$, $s \in \mathbb{N}$, then follows from the central limit theorem by the usual arguments (see Theorem 6.5.1 in Lehmann and Casella 1998, cf. Bedbur et al. 2012).

□

4 Confidence regions

In the previous sections, we addressed point estimation of the vector $\boldsymbol{\vartheta}$ of model parameters. To derive confidence regions for $\boldsymbol{\vartheta}$, we may proceed as in Vuong et al. (2013) (see also Bedbur et al. 2013) in the framework of the usual proportional hazard rate model and choose a setup by means of distance measures between density functions. Briefly, we illustrate the procedure by exemplarily utilizing the Jeffreys–Kullback–Leibler distance and the Hellinger distance.

Given densities h_1, h_2 w.r.t. Lebesgue measure ν^k, $k \in \mathbb{N}$, having the same support with interior S, then these distance measures are defined by

$$D_J(h_1, h_2) = \int_S \left(h_1(\boldsymbol{x}) - h_2(\boldsymbol{x})\right)\left(\log h_1(\boldsymbol{x}) - \log h_2(\boldsymbol{x})\right) d\nu^k(\boldsymbol{x}) \quad (12)$$

and

$$D_H(h_1, h_2) = \left(\int_S \left(h_1^{1/2}(\boldsymbol{x}) - h_2^{1/2}(\boldsymbol{x})\right)^2 d\nu^k(\boldsymbol{x})\right)^{1/2}, \quad (13)$$

respectively.

Consider two joint density functions of the form (6) with the same underlying distribution functions $F_1, \ldots, F_r$ and parameter vector $\boldsymbol{\vartheta} = (\vartheta_1, \ldots, \vartheta_r)^t$ and $\boldsymbol{\vartheta}' = (\vartheta_1', \ldots, \vartheta_r')^t$, respectively. Applying (12) and (13) to these density functions, it turns out that both expressions are free of $F_1, \ldots, F_r$ and depend on $\boldsymbol{\vartheta}$ and $\boldsymbol{\vartheta}'$ only by means of componentwise ratios.

Lemma 1 *For density functions as in* (6)*, Jeffreys–Kullback–Leibler and Hellinger distance are given by*

$$D_J\left(f_{\boldsymbol{\vartheta}}^{(s)}, f_{\boldsymbol{\vartheta}'}^{(s)}\right) = s\sum_{j=1}^{r}\left(\frac{\vartheta_j}{\vartheta_j'} + \frac{\vartheta_j'}{\vartheta_j} - 2\right) = D_J^{(s)}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}'), \text{ say,} \quad (14)$$

and

$$D_H\left(f_{\boldsymbol{\vartheta}}^{(s)}, f_{\boldsymbol{\vartheta}'}^{(s)}\right) = \left(2 - 2^{rs+1}\prod_{j=1}^{r}\left(\left(\frac{\vartheta_j}{\vartheta_j'}\right)^{1/2} + \left(\frac{\vartheta_j'}{\vartheta_j}\right)^{1/2}\right)^{-s}\right)^{1/2}$$
$$= D_H^{(s)}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}'), \text{ say.} \quad (15)$$

Proof Since we have exponential family structure, we find with (7)

$$D_J\left(f_{\boldsymbol{\vartheta}}^{(s)}, f_{\boldsymbol{\vartheta}'}^{(s)}\right) = (\boldsymbol{\vartheta} - \boldsymbol{\vartheta}')^t\left(E_{\boldsymbol{\vartheta}}\boldsymbol{T}^{(s)} - E_{\boldsymbol{\vartheta}'}\boldsymbol{T}^{(s)}\right)$$
$$= s(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}')^t\left(\left(1/\vartheta_1', \ldots, 1/\vartheta_r'\right) - \left(1/\vartheta_1, \ldots, 1/\vartheta_r\right)\right)$$

which leads to formula (14) (cf. Kullback 1959, p. 45).

Representation (15) directly follows with (5) by exploiting

$$\int_{\mathcal{X}_r^s} \left(f_{\boldsymbol{\vartheta}}^{(s)}(\tilde{\boldsymbol{x}}^{(s)})\, f_{\boldsymbol{\vartheta}'}^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) \right)^{1/2} d\mu_r^s(\tilde{\boldsymbol{x}}^{(s)})$$
$$= \left(\exp\left\{ \kappa\left(\frac{\boldsymbol{\vartheta} + \boldsymbol{\vartheta}'}{2} \right) - \frac{\kappa(\boldsymbol{\vartheta}) + \kappa(\boldsymbol{\vartheta}')}{2} \right\} \right)^s .$$

□

The results in Lemma 1 can be applied to derive simultaneous confidence regions for $\vartheta_1, \ldots, \vartheta_r$ based on data as in (6). From Theorem 2 (1), the MLEs of the ϑs are independent and inverse gamma distributed. Hence, a direct approach to a joint confidence region for $\vartheta_1, \ldots, \vartheta_r$ is to consider an r-dimensional rectangle. Then, however, confidence levels for all r axis have to be chosen in advance. To avoid this restriction, we propose an alternative procedure.

In the above situation, with $D \in \{D_J, D_H\}$, and the MLE $\hat{\boldsymbol{\vartheta}}^{(s)}$ of $\boldsymbol{\vartheta}$ as in Theorem 2 (1), we consider balls

$$R^{(s)}(c) = \left\{ \boldsymbol{\vartheta} \in \mathbb{R}_+^r \;:\; D^{(s)}\left(\hat{\boldsymbol{\vartheta}}^{(s)}, \boldsymbol{\vartheta} \right) \le c \right\}$$

with radius $c > 0$ about the MLE. Since, from Theorem 2 (1), the distribution of the componentwise ratios $\hat{\vartheta}_j^{(s)}/\vartheta_j$ do neither depend on $F_1, \ldots, F_r$ nor on ϑ_j for $1 \le j \le r$, the same holds for the distribution of $D^{(s)}(\hat{\boldsymbol{\vartheta}}^{(s)}, \boldsymbol{\vartheta})$ and its $(1-p)$-quantile c_{1-p}, $p \in (0, 1)$, which can be determined numerically without much effort by generating random samples from gamma distributions.

Thus, by introducing the notation $P_{\boldsymbol{\vartheta}}^{(s)} = f_{\boldsymbol{\vartheta}}^{(s)} \mu_r^s$ for $\boldsymbol{\vartheta} \in \mathbb{R}_+^r$, we conclude that

$$P_{\boldsymbol{\vartheta}}^{(s)}\left(\boldsymbol{\vartheta} \in R^{(s)}(c_{1-p}) \right) = P_{\boldsymbol{\vartheta}}^{(s)}\left(D^{(s)}(\hat{\boldsymbol{\vartheta}}^{(s)}, \boldsymbol{\vartheta}) \le c_{1-p} \right) = 1 - p\,, \qquad \boldsymbol{\vartheta} \in \mathbb{R}_+^r\,,$$

and $R^{(s)}(c_{1-p})$ is a $(1-p)$-confidence region for the vector $\boldsymbol{\vartheta}$ of model parameters.

When assuming ascendingly ordered parameters, i.e., $\vartheta_1 \le \ldots \le \vartheta_r$, as done in Sect. 3.2, the construction of confidence regions for $\boldsymbol{\vartheta}$ taking into account this information in advance is not straightforward. Here, we content ourselves with the remark that the intersection $R^{(s)}(c_{1-p}) \cap \Delta$, where Δ is defined as in the proof of Theorem 3, represents a $(1-p)$-confidence region for $\boldsymbol{\vartheta} \in \Delta$ since

$$P_{\boldsymbol{\vartheta}}^{(s)}\left(\boldsymbol{\vartheta} \in R^{(s)}(c_{1-p}) \cap \Delta \right) = P_{\boldsymbol{\vartheta}}^{(s)}\left(\boldsymbol{\vartheta} \in R^{(s)}(c_{1-p}) \right) = 1 - p\,, \qquad \boldsymbol{\vartheta} \in \Delta\,,$$

which contains only parameter vectors with ascendingly ordered components.

In the situation of Sect. 3.3, where the ϑs are connected via a linear link function, the Jeffreys–Kullback–Leibler distance $D_J^{(s)}(\boldsymbol{\tau}, \boldsymbol{\tau}') = D_J(\tilde{f}_{\boldsymbol{\tau}}^{(s)}, \tilde{f}_{\boldsymbol{\tau}'}^{(s)})$ and the Hellinger distance $D_H^{(s)}(\boldsymbol{\tau}, \boldsymbol{\tau}') = D_H(\tilde{f}_{\boldsymbol{\tau}}^{(s)}, \tilde{f}_{\boldsymbol{\tau}'}^{(s)})$ are obtained by setting $\vartheta_j = \tau_1 + \tau_2 y_j$ and $\vartheta_j' = \tau_1' + \tau_2' y_j$, $1 \le j \le r$, in formula (14) and (15), respectively. Similarly as above, one may consider the distances $D_J^{(s)}(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})$ and $D_H^{(s)}(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})$ to construct simultaneous confidence regions for τ_1 and τ_2. Here, however, the explicit form of the MLE $\hat{\boldsymbol{\tau}}^{(s)}$ of $\boldsymbol{\tau}$ and its distribution theoretical properties are not available (see Theorem 3.3) which makes it difficult to determine the radius c of a ball about $\hat{\boldsymbol{\tau}}^{(s)}$ for a desired confidence level. However, asymptotical results can be stated. In what follows, let $D_\bullet = D_\bullet^{(s)}$ to simplify notation.

The Jeffreys–Kullback–Leibler distance belongs to the broader class of ϕ-divergence measures and results from the general definition by choosing the function ϕ according to $\phi(x) = (x-1)\log(x)$ (cf. Pardo 2006, pp. 3–6). Hence, by application of Theorem 9.1 in Pardo (2006), p. 409, we obtain that the asymptotic distribution of $2sD_J(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})/\phi''(1) = sD_J(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})$ is an exponential distribution with mean 2. We conclude that

$$\left\{\boldsymbol{\tau} \in \tilde{\Theta} \,:\, sD_J\left(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau}\right) \le -2\log(p)\right\}$$

forms a confidence region for $\boldsymbol{\tau}$ with asymptotic confidence level $1-p$.

The Hellinger distance is closely related to another ϕ-divergence measure, i.e., the Cressie–Read power divergence $D_{CR_{-1/2}}(\boldsymbol{\tau}, \boldsymbol{\tau}') = 2D_H(\boldsymbol{\tau}, \boldsymbol{\tau}')^2$ with parameter $-1/2$ which is obtained by setting $\phi(x) = -4(x^{1/2} - x + (x-1)/2)$ (cf. Pardo 2006, pp. 5-7). Again, Theorem 9.1 in Pardo (2006) then states that $2sD_{CR_{-1/2}}(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})$ is asymptotically exponentially distributed with mean 2, and the confidence region

$$\left\{\boldsymbol{\tau} \in \tilde{\Theta} \,:\, D_H\left(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau}\right) \le \left(-\frac{\log(p)}{2s}\right)^{1/2}\right\} \tag{16}$$

for $\boldsymbol{\tau}$ turns out to have asymptotic confidence level $1-p$.

Alternatively and in a similar manner, one may use the relationship between the Hellinger distance and the Rényi divergence $D_{R_{1/2}}(\boldsymbol{\tau}, \boldsymbol{\tau}') = -4\log(1 - D_H(\boldsymbol{\tau}, \boldsymbol{\tau}')^2/2)$ with parameter 1/2, where the latter is known to be a particular (h, ϕ)-divergence measure with functions h and ϕ specified as $h(x) = -4\log(-x/4+1)$ and $\phi(x) = -4(x^{1/2} - (x-1)/2 - 1)$ [(cf. Pardo 2006, formula (1.3), and pp. 7/8]. From Remark 9.1 in Pardo (2006) then follows that $2sD_{R_{1/2}}(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})/(h'(0)\phi''(1)) = 2sD_{R_{1/2}}(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau})$ is asymptotically exponentially distributed with mean 2. Hence,

$$\begin{aligned}&\left\{\boldsymbol{\tau} \in \tilde{\Theta} \,:\, 2sD_{R_{1/2}}\left(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau}\right) \le -2\log(p)\right\}\\ &\quad = \left\{\boldsymbol{\tau} \in \tilde{\Theta} \,:\, D_H\left(\hat{\boldsymbol{\tau}}^{(s)}, \boldsymbol{\tau}\right) \le \left(2\left(1 - p^{1/(4s)}\right)\right)^{1/2}\right\}\end{aligned}$$

provides another confidence region for $\boldsymbol{\tau}$ with asymptotic confidence level $1-p$ which, as the power series expansion of $\log(p^{1/(4s)})$ shows, forms a subset of the one specified in formula (16) and, thus, has smaller area along with a correspondingly smaller actual confidence level.

References

Bagdonavičius, V., Nikulin, M. (2002). *Accelerated life models—modeling and statistical analysis*. Boca Raton: Chapman & Hall/CRC Press.

Balakrishnan, N. (2009). A synthesis of exact inferential results for exponential step-stress models and associated optimal accelerated life-tests. *Metrika*, *69*, 351–396.

Balakrishnan, N., Beutner, E., Kamps, U. (2008). Order restricted inference for sequential k-out-of-n systems. *Journal of Multivariate Analysis*, *99*, 1489–1502.

Balakrishnan, N., Belzunze, F., Hami, N., Khaledi, B. E. (2010). Univariate and multivariate likelihood ratio ordering of generalized order statistics and associated conditional variables. *Probability in the Engineering and Informational Sciences*, *24*, 441–455.

Balakrishnan, N., Beutner, E., Kamps, U. (2011). Modeling parameters of a load-sharing system through link functions in sequential order statistics models and associated inference. *IEEE Transactions on Reliability TR-60*, 605–611.

Balakrishnan, N., Kamps, U., Kateri, M. (2012). A sequential order statistics approach to step-stress testing. *Annals of the Institute of Statistical Mathematics*, *64*, 303–318.

Bedbur, S. (2010). UMPU tests based on sequential order statistics. *Journal of Statistical Planning and Inference*, *140*, 2520–2530.

Bedbur, S., Beutner, E., Kamps, U. (2012). Generalized order statistics: an exponential family in model parameters. *Statistics*, *46*(2), 159–166.

Bedbur, S., Lennartz, J. M., Kamps, U. (2013). Confidence regions in models of ordered data. *Journal of Statistical Theory and Practice*, *7*(1), 59–72.

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

Beutner, E., Kamps, U. (2009). Order restricted statistical inference for scale parameters based on sequential order statistics. *Journal of Statistical Planning and Inference*, *139*(2), 2963–2969.

Bickel, P. J., Doksum, K. A. (2001). *Mathematical statistics: Basic ideas and selected topics* (2nd ed.). Upper Saddle River: Prentice Hall.

Burkschat, M., Navarro, J. (2011). Aging properties of sequential order statistics. *Probability in the Engineering and Informational Sciences*, *25*, 449–467.

Burkschat, M., Torrado, N. (2014). On the reversed hazard rate of sequential order statistics. *Statistics & Probability Letters*, *85*, 106–113.

Burkschat, M., Kamps, U., Kateri, M. (2010). Sequential order statistics with an order statistics prior. *Journal of Multivariate Analysis*, *101*, 1826–1836.

Cramer, E., Kamps, U. (1996). Sequential order statistics and k-out-of-n systems with sequentially adjusted failure rates. *Annals of the Institute of Statistical Mathematics*, *48*, 535–549.

Cramer, E., Kamps, U. (2001a). Estimation with sequential order statistics from exponential distributions. *Annals of the Institute of Statistical Mathematics*, *53*, 307–324.

Cramer, E., Kamps, U. (2001b). Sequential k-out-of-n systems. In N. Balakrishnan & C. R. Rao (Eds.), *Handbook of Statistics* (vol. 20, pp. 301–372). Amsterdam: Advances in Reliability Elsevier.

Deshpande, J., Dewan, I., Naik-Nimbalka, U. (2010). A family of distributions to model load sharing systems. *Journal of Statistical Planning and Inference*, *140*, 1441–1451.

Gaver, D. P., Acar, M. (1979). Analytical hazard representations for use in reliability, mortality, and simulation studies. *Communications in Statistics Simulation and Computation*, *8*, 91–111.

Gurgenci, H., Guan, Z. (2001). Mobile plant maintenance and the dutymeter concept. *Journal of Quality in Maintenance Engineering*, *7*, 275–285.

Hjorth, U. (1980). A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics*, *22*, 99–107.

Hu, T., Zhuang, W. (2005). A note on stochastic comparisons of generalized order statistics. *Statistics & Probability Letters*, *72*, 163–170.

Johnson, N. L., Kotz, S., Balakrishnan, N. (1995). *Continuous Univariate Distributions* (Vol. 2). New York: Wiley.

Kamps, U. (1995a). *A concept of generalized order statistics*. Stuttgart: Teubner.

Kamps, U. (1995b). A concept of generalized order statistics. *Journal of Statistical Planning and Inference*, *48*, 1–23.

Kateri, M., Kamps, U., Balakrishnan, N. (2009). A meta-analysis approach for step-stress experiments. *Journal of Statistical Planning and Inference*, *139*, 1907–2919.

Kateri, M., Kamps, U., Balakrishnan, N. (2010). Multi-sample simple step-stress experiment under time constraints. *Statistica Neerlandica*, *64*, 77–96.

Kotz, S., Balakrishnan, N., Johnson, N. L. (2000). *Continuous Multivariate Distributions: methods and applications*, 2nd ed (vol. 1). New York: Wiley.

Kullback, S. (1959). *Information theory and statistics*. New York: Wiley.

Lai, C. D., Xie, M. (2006). *Stochastic aging and dependence for reliability*. New York: Springer.

Lai, C. D., Xie, M., Murthy, D. N. P. (2003). A modified Weibull distribution. *IEEE Transactions on Reliability*, *52*, 33–37.

Lehmann, E. L., Casella, G. (1998). *Theory of point estimation* (2nd ed.). New York: Springer.

Lehmann, E. L., Romano, J. P. (2005). *Testing statistical hypotheses* (3rd ed.). New York: Springer.

Marshall, A. W., Olkin, I. (2007). *Life distributions: structure of nonparametric, semiparametric, and parametric families*. New York: Springer.

Mudholkar, G. S., Srivastava, D. K., Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, *91*, 1575–1583.

Navarro, J., Burkschat, M. (2011). Coherent systems based on sequential order statistics. *Naval Research Logistics*, *58*, 123–135.

Nikulin, M., Haghighi, F. (2006). A chi-squared test for the generalized power Weibull family for the head-and-neck cancer censored data. *Journal of Mathematical Sciences*, *133*, 1333–1341.

Pardo, L. (2006). *Statistical Inference Based on Divergence Measures*. Boca Raton: Chapman & Hall/CRC Press.

Rowell, G., Siegrist, K. (1998). Relative aging of distributions. *Probability in the Engineering and Informational Sciences*, *12*, 469–478.

Sengupta, D., Deshpande, J. V. (1994). Some results on the relative ageing of two life distributions. *Journal of Applied Probability*, *31*, 991–1003.

Torrado, N., Lillo, R. E., Wiper, M. P. (2012). Sequential order statistics: ageing and stochastic orderings. *Methodology and Computing in Applied Probability*, *14*, 579–596.

Vuong, Q. N., Bedbur, S., Kamps, U. (2013). Distances between models of generalized order statistics. *Journal of Multivariate Analysis*, *118*, 24–36.

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