LINK FUNCTIONS FOR PARAMETERS OF SEQUENTIAL ORDER STATISTICS AND CURVED EXPONENTIAL FAMILIES

BY

GRIGORIY VOLOVSKIY (Aachen), STEFAN BEDBUR (Aachen), AND
UDO KAMPS (Aachen)

Abstract. Estimation of model parameters of sequential order statistics under linear and nonlinear link function assumptions is considered. Utilizing the arising curved exponential family structure, conditions for existence and uniqueness as well as the validity of asymptotic properties of maximum likelihood estimators are stated. Minimal sufficiency and completeness of the associated canonical statistics are discussed.

2020 Mathematics Subject Classification: Primary 62F10; Secondary 62N05, 62N02.

Key words and phrases: sequential order statistic, maximum likelihood estimation, curved exponential family, link function.

1. INTRODUCTION

Sequential order statistics (SOSs) constitute a general model for ordered statistical data and provide a natural modeling framework for multi-unit technical systems, in which the failure of a component may have an impact on the lifetimes of the surviving components; see, e.g., [18], [19], [14], [25], [2], [13], [16], [7], [5], [6], [11]. The modeling is based on the assumption that, upon a failure of some component, the remaining components have a possibly different underlying lifetime distribution. In the case of an $(n - r + 1)$-out-of-n system with n components, say, let $F_1, \ldots, F_n$ be the respective distribution functions. Commonly, a semi-parametric setting with proportional hazard rates is chosen by supposing that $F_j = 1 - (1 - F)^{\alpha_j}$, $1 \leqslant j \leqslant n$, where F denotes some absolutely continuous distribution function with corresponding density function f and $\alpha_1, \ldots, \alpha_n$ are positive model parameters. The baseline distribution function F is chosen according to a model assumption or to some prior information. If, e.g., F is chosen to be the standard exponential distribution function, then F_j is an exponential distribution function with rate parameter α_j, $1 \leqslant j \leqslant n$. In the semi-parametric setting, the joint density function of the first $r \leqslant n$ SOSs $X_{\star 1}, \ldots, X_{\star r}$ based on $F_1, \ldots, F_n$,

which describe the first r ordered component failure times, is given by

$$(1.1) \qquad f_{\boldsymbol{\alpha}}(\boldsymbol{x}) = \exp\Big\{ \sum_{j=1}^{r} \alpha_j T_j(\boldsymbol{x}) - \kappa(\boldsymbol{\alpha}) \Big\} h(x)$$

for $\boldsymbol{x} \in \mathcal{X}_r = \{(x_1, \ldots, x_r) \in \mathbb{R}^r : F^{-1}(0+) < x_1 < \cdots < x_r < F^{-1}(1)\}$ with $\kappa(\boldsymbol{\alpha}) = -\sum_{j=1}^{r} \log(\alpha_j)$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_r) \in (0, \infty)^r$, and

$$T_j(\boldsymbol{x}) = (n - j + 1) \log\left(\frac{\overline{F}(x_j)}{\overline{F}(x_{j-1})} \right), \quad 1 \leqslant j \leqslant r,$$

$$h(\boldsymbol{x}) = \frac{n!}{(n-r)!} \prod_{j=1}^{r} \frac{f(x_j)}{\overline{F}(x_j)}, \quad \boldsymbol{x} \in \mathcal{X}_r,$$

where F^{-1} denotes the quantile function of F, $\overline{F} = 1 - F$, and $\overline{F}(x_0) \equiv 1$; see [4]. In applications of SOSs, the model parameters $\alpha_1, \alpha_2, \ldots$, which describe successive changes in the system due to component failures, and the baseline distribution function F will usually be unknown and will have to be estimated based on the available data. In a non-parametric approach such as in [9], [10], and [23], the distribution function F can be estimated. However, in lifetime experiments, the number of observations is commonly fairly small. If prior information (e.g. from previous experiments) is available with respect to the shape of F, or equivalently the shape of the hazard rate function $f/\overline{F}$, the only uncertainty regarding the true distribution of the lifetimes of the components is captured by the vector $\boldsymbol{\alpha}$ of unknown model parameters (another option is to consider $\overline{F}^{\alpha_1}$ as given and the parameters $\alpha_2, \ldots, \alpha_r$ to be unknown).

In what follows, we will consider this situation and assume that F is known. Then the family $\mathcal{P} = \{f_{\boldsymbol{\alpha}}\, d\lambda_r : \boldsymbol{\alpha} \in (0, \infty)^r\}$ exhibits an exponential family structure, which allows for convenient derivation of useful inferential results. Here and in the following, λ_r denotes the Lebesgue measure on the Borel sets of the sample space $\mathcal{X}_r$ and the notation $g\, d\nu$ is used for the measure with ν-density function g. In particular, since the exponential family $\mathcal{P}$ is regular and representation (1.1) is minimal, the vector $\boldsymbol{T} = (T_1, \ldots, T_r)$ is minimal sufficient and complete for $\mathcal{P}$; for details, see [4]. Moreover, as far as point estimation of the model parameters is concerned, maximum likelihood estimators (MLEs) and uniformly minimum variance unbiased estimators can be easily obtained, and, in a multi-sample setup, the asymptotic efficiency can be established in a straightforward way; see [4]. Regarding statistical tests utilizing the exponential family structure of $\mathcal{P}$, we refer to, e.g., [3].

In situations with a small amount of data, arising for instance in accelerated life testing experiments (see, e.g., [7]), it is necessary to impose additional model assumptions to make simultaneous estimation of $\alpha_1, \ldots, \alpha_r$ feasible. On the other hand, prior information about the model parameters may be present. In either case, an injective parametric function $\boldsymbol{\psi} : \Xi \to (0, \infty)^r$ may be incorporated into the

model, which then induces a subfamily $\mathcal{P}_\psi = \{f_{\psi(\boldsymbol{\xi})} \, d\lambda_r : \boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \Xi \subseteq \mathbb{R}^m\}$, $m < r$, of the exponential family $\mathcal{P}$. Then, in the $(n - r + 1)$-out-of-n model, the parameters $\alpha_1, \ldots, \alpha_r$ are connected with a smaller number of parameters $\xi_1, \ldots, \xi_m$, say, via ψ, which is therefore referred to as a link function; see, e.g., [2]. Depending on the choice of ψ, the resulting family $\mathcal{P}_\psi$ may again possess a regular exponential family structure (now in the parameters $\xi_1, \ldots, \xi_m$) or be a genuine curved exponential family. In the latter case, some or most of the nice properties of $\mathcal{P}$ mentioned above may get lost, so that the associated statistical inference may become more involved. For instance, in a curved exponential family, in general, completeness of $\boldsymbol{T}$ does no longer hold true, and the existence and uniqueness of the MLE are not evident at all and have to be examined case-by-case. Nonetheless, there are results in the literature, e.g., in [12], providing useful tools for dealing with maximum likelihood estimation in curved exponential families, which however have not yet been applied to parametric inference with SOSs under link function assumptions.

The goal of this paper is to demonstrate the feasibility of maximum likelihood estimation in SOSs for quite a general choice of the link function ψ and with the given underlying distribution function F. Minimal sufficiency and completeness of the canonical statistic $\boldsymbol{T}$ will be discussed as well. Previous results in the literature are included in the proposed general approach as particular examples; see, for instance, [7], [26].

2. MAXIMUM LIKELIHOOD ESTIMATION UNDER LINK FUNCTION ASSUMPTIONS

In what follows, we assume that a sample $\boldsymbol{X}_\star^{(1)}, \ldots, \boldsymbol{X}_\star^{(s)}$ of independent and identically distributed random vectors with density function (1.1) each is given. The density function of the whole sample is then represented by

$$(2.1) \qquad f_{\boldsymbol{\alpha}}^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \exp\Big\{ \sum_{j=1}^{r} \alpha_j T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) - s\kappa(\boldsymbol{\alpha}) \Big\} h^{(s)}(\tilde{\boldsymbol{x}}^{(s)}),$$

$$\tilde{\boldsymbol{x}}^{(s)} = (\boldsymbol{x}^{(1)}, \ldots, \boldsymbol{x}^{(s)}) \in \mathcal{X}_r^s = \bigtimes_{i=1}^{s} \mathcal{X}_r,$$

with

$$T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \sum_{i=1}^{s} T_j(\boldsymbol{x}^{(i)}), \quad 1 \leqslant j \leqslant r,$$

$$h^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \prod_{i=1}^{s} h(\boldsymbol{x}^{(i)}), \quad \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s.$$

It is readily seen that the family $\mathcal{P}^{(s)} = \{f_{\boldsymbol{\alpha}}^{(s)} \, d\lambda_{rs} : \boldsymbol{\alpha} \in (0, \infty)^r\}$ of joint distributions of $\boldsymbol{X}_\star^{(1)}, \ldots, \boldsymbol{X}_\star^{(s)}$ forms an exponential family in $\alpha_1, \ldots, \alpha_r$ and canonical statistics $T_1^{(s)}, \ldots, T_r^{(s)}$. The statistics $T_1^{(s)}, \ldots, T_r^{(s)}$ on $\mathcal{X}_r^s$ are independent,

$-T_j^{(s)}$ has a gamma distribution with scale parameter $1/\alpha_j$ and shape parameter s, $1 \leqslant j \leqslant r$, and the family $\mathcal{P}^{(s)}$ inherits all regularity properties from $\mathcal{P}$; see [4]. In the following, let $\boldsymbol{T}^{(s)} = (T_1^{(s)}, \ldots, T_r^{(s)})$ be the vector of statistics. For a given link function ψ, the induced subfamily of $\mathcal{P}^{(s)}$ will be denoted by $\mathcal{P}_\psi^{(s)}$. Since we will be exclusively concerned with estimation in $\mathcal{P}_\psi^{(s)}$, we will denote the associated likelihood function of $\boldsymbol{\xi} \in \Xi$ by $l(\cdot|\tilde{\boldsymbol{x}}^{(s)})$ without making any reference to the function ψ.

Before discussing results for subfamilies $\mathcal{P}_\psi^{(s)}$, which are curved exponential families, we present an example demonstrating that a link function can induce a regular exponential family structure. In that case, there is no need to invoke results for curved exponential families to find properties of MLEs of $\xi_1, \ldots, \xi_m$; this can be done by standard arguments.

EXAMPLE 2.1. Assume that the parameters $\alpha_1, \ldots, \alpha_r$ are linked via $\boldsymbol{\alpha} = \psi(\boldsymbol{\xi})$ with linear link function $\psi = (\psi_1, \ldots, \psi_r)$ defined by

$$(2.2) \qquad \psi_j(\boldsymbol{\xi}) = \sum_{i=1}^m \xi_i y_{j,i}, \quad j = 1, \ldots, r,$$

for some $m < r$, where $y_{j,i}$, $1 \leqslant j \leqslant r$, $1 \leqslant i \leqslant m$, denote known real numbers, and $\xi_1, \ldots, \xi_m$ are unknown parameters. We further assume that the matrix $\boldsymbol{A} = (y_{j,i})_{1\leqslant j\leqslant r, 1\leqslant i\leqslant m}$ in this multiple linear regression approach has full rank m. Since $\psi(\boldsymbol{\xi}) = \boldsymbol{A\xi}$, this already implies that ψ is one-to-one. Under the link function assumption (2.2) and by setting $\tilde{f}_{\boldsymbol{\xi}}^{(s)} = f_{\psi(\boldsymbol{\xi})}^{(s)}$, the density function (2.1) becomes

$$\tilde{f}_{\boldsymbol{\xi}}^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \exp\Big\{ \sum_{i=1}^m \xi_i \tilde{T}_i^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) - s\tilde{\kappa}(\boldsymbol{\xi}) \Big\} h^{(s)}(\tilde{\boldsymbol{x}}^{(s)}), \quad \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s,$$

with

$$\tilde{\kappa}(\boldsymbol{\xi}) = - \sum_{j=1}^r \log\Big(\sum_{i=1}^m \xi_i y_{j,i} \Big),$$

$$\boldsymbol{\xi} \in \Xi = \Big\{ (\eta_1, \ldots, \eta_m) \in \mathbb{R}^m : \sum_{i=1}^m \eta_i y_{j,i} > 0, \ j = 1, \ldots, r \Big\},$$

and

$$\tilde{T}_i^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = \sum_{j=1}^r y_{j,i} T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}), \quad \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s, \quad i = 1, \ldots, m.$$

Since $\Xi \subseteq \mathbb{R}^m$ is open and the matrix $\boldsymbol{A}$ has full rank, the family $\mathcal{P}_\psi^{(s)}$ forms a regular exponential family. In particular, the vector of statistics $\tilde{\boldsymbol{T}}^{(s)} = (\tilde{T}_1^{(s)}, \ldots, \tilde{T}_m^{(s)})$

is minimal sufficient and complete for $\mathcal{P}_{\psi}^{(s)}$. Moreover, the MLE $\hat{\boldsymbol{\xi}}^{(s)}$ of $\boldsymbol{\xi}$ uniquely exists almost surely and is given as the only solution of the likelihood equations

$$\sum_{j=1}^{r} \frac{y_{j,i}}{\sum_{l=1}^{m} y_{j,l} \xi_l} = -\frac{\tilde{T}_i^{(s)}(\tilde{\boldsymbol{x}}^{(s)})}{s}, \quad i = 1, \ldots, m,$$

with respect to $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_m) \in \Xi$. Furthermore, the sequence $(\hat{\boldsymbol{\xi}}^{(s)})_{s \in \mathbb{N}}$ is strongly consistent and asymptotically efficient with asymptotic covariance matrix given by the inverse matrix of $\mathbf{Cov}_{\boldsymbol{\xi}}(\tilde{\boldsymbol{T}}^{(1)})$; see [4]. Here, using the independence as well as the distributional properties of the statistics $T_1^{(1)}, \ldots, T_r^{(1)}$, we obtain

$$\mathrm{Var}_{\boldsymbol{\xi}}(\tilde{T}_k^{(1)}) = \sum_{j=1}^{r} y_{j,k}^2 \Big(\sum_{i=1}^{m} y_{j,i} \xi_i\Big)^{-2}, \quad k = 1, \ldots, m,$$

$$\mathrm{Cov}_{\boldsymbol{\xi}}(\tilde{T}_k^{(1)}, \tilde{T}_l^{(1)}) = \sum_{j=1}^{r} y_{j,k} y_{j,l} \Big(\sum_{i=1}^{m} y_{j,i} \xi_i\Big)^{-2}, \quad k, l = 1, \ldots, m.$$

REMARK 2.1. (i) As a particular case of Example 2.1 with $m = 2 < r$ and $y_{j,1} = 1$, $y_{j,2} = y_j$ for $j = 1, \ldots, r$, we obtain the linear link function $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_r)$ with

$$\psi_j(\boldsymbol{\xi}) = \boldsymbol{\psi}(\xi_1, \xi_2) = \xi_1 + y_j \xi_2, \quad j = 1, \ldots, r,$$

where $y_1, \ldots, y_r$ are known real numbers satisfying $y_i \neq y_j$ for at least two indices $i, j \in \{1, \ldots, r\}$, and $\Xi = \{(\eta_1, \eta_2) \in \mathbb{R}^2 : \eta_1 + y_j \eta_2 > 0, \ j = 1, \ldots, r\}$. This setup is related to the one studied in [2] and [5], and the results stated in Example 2.1 hold true in the present situation.

(ii) Let us assume that the model parameters $\alpha_1, \ldots, \alpha_r$ are linked via $\boldsymbol{\alpha} = \boldsymbol{\psi}(\boldsymbol{\xi})$ with linear link function $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_r)$ defined by

$$\psi_j(\boldsymbol{\xi}) = \xi_1 + y_j \xi_2 + y_j^2 \xi_3 + \cdots + y_j^{m-1} \xi_m, \quad j = 1, \ldots, r,$$

for some $m < r$, where $y_1, \ldots, y_r$ are known real numbers taking at least m different values, and $\xi_1, \ldots, \xi_m$ are unknown link function parameters. Setting $y_{j,i} = y_j^{i-1}$, $1 \leqslant j \leqslant r$, $1 \leqslant i \leqslant m$, the present setup is seen to be another particular case of Example 2.1, where $\boldsymbol{A}$ is a rank m Vandermonde matrix, and the results stated there may be applied. In the particular example $m = 3$, $r = 4$, $y_j = j$ for $j = 1, 2, 3, 4$, the parameter space $(0, \infty)^4$ is restricted via $\boldsymbol{\alpha} = \boldsymbol{A}\boldsymbol{\xi}$ with

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \end{pmatrix}.$$

From the above examples it is evident that the family $\mathcal{P}_{\boldsymbol{\psi}}^{(s)}$ is again a regular exponential family provided the link function takes the form $\boldsymbol{\psi}(\boldsymbol{\xi}) = \boldsymbol{A}\boldsymbol{\xi}$ for some matrix $\boldsymbol{A}$ of full rank, in which case inference via the maximum likelihood method is straightforward, and existence and uniqueness of the MLE of $\boldsymbol{\xi}$ is guaranteed.

The following result yields sufficient conditions for the MLE to exist uniquely if $\boldsymbol{\psi}$ is a nonlinear function. The result reveals that existence of the MLE in curved exponential submodels of SOSs is essentially linked with topological properties of $\boldsymbol{\psi}(\Xi)$. Moreover, an important condition for the uniqueness of the MLE is that $\boldsymbol{\psi}$ locally takes the from $\boldsymbol{\psi}(\boldsymbol{\xi}) = \boldsymbol{A}\boldsymbol{\xi}$ with a full rank matrix $\boldsymbol{A}$. Below, we use the notation $\mu = h^{(s)} \, d\lambda_{rs}$; note that the measure μ is equivalent to every distribution $P \in \mathcal{P}^{(s)}$. Furthermore, we call a function $\boldsymbol{g} : V \to W$ with $V \subseteq \mathbb{R}^m$ open, $W \subseteq \mathbb{R}^n$, $m \leqslant n$, a C^k-immersion, $k \in \mathbb{N}$, if $\boldsymbol{g}$ is a C^k-function and the Jacobian matrix $D_{\boldsymbol{g}}(\boldsymbol{x})$ of $\boldsymbol{g}$ at $\boldsymbol{x}$ has full rank for all $\boldsymbol{x} \in V$.

THEOREM 2.1. *Let $\tilde{\boldsymbol{X}}_{\star}^{(s)} = (\boldsymbol{X}_{\star}^{(1)}, \ldots, \boldsymbol{X}_{\star}^{(s)})$ have density function (2.1). Further, let $\boldsymbol{\psi} : \Xi \to (0, \infty)^r$ with $\Xi \subseteq \mathbb{R}^m$, $m < r$, be one-to-one, and let $\mathcal{P}_{\boldsymbol{\psi}}^{(s)}$ be the associated subfamily of $\mathcal{P}^{(s)}$.*

(i) *Existence: If $\boldsymbol{\psi}(\Xi)$ is a relatively closed subset of $(0, \infty)^r$, i.e., $\boldsymbol{\psi}(\Xi) = (0, \infty)^r \cap U$ for a closed subset $U \subseteq \mathbb{R}^r$, then an MLE in Ξ exists almost surely.*

(ii) *Uniqueness: If Ξ is an open subset of $\mathbb{R}^m$, $\boldsymbol{\psi}$ a C^2-immersion, and F a C^2-function with positive Lebesgue-density function f, then the MLE is unique with probability one, i.e.,*

$$D = \Big\{ \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s : \exists\, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in \Xi \text{ with } \boldsymbol{\xi}_1 \neq \boldsymbol{\xi}_2 \text{ and}$$

$$l(\boldsymbol{\xi}_1 | \tilde{\boldsymbol{x}}^{(s)}) = \max_{\boldsymbol{\xi} \in \Xi} l(\boldsymbol{\xi} | \tilde{\boldsymbol{x}}^{(s)}) = l(\boldsymbol{\xi}_2 | \tilde{\boldsymbol{x}}^{(s)}) \Big\}$$

is Borel-measurable and $\mu(D) = 0$.

Moreover, under the assumptions in (i) *and* (ii)*, the MLE can be obtained as a solution of the likelihood equations*

$$(2.3) \qquad \sum_{j=1}^{r} \left(\frac{\partial}{\partial \xi_i} \psi_j(\boldsymbol{\xi}) \right) T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = -s \sum_{j=1}^{r} \frac{\partial/\partial \xi_i \, \psi_j(\boldsymbol{\xi})}{\psi_j(\boldsymbol{\xi})}, \quad i = 1, \ldots, m.$$

Proof. (i) Since $\mathcal{P}^{(s)}$ is a regular exponential family and $\boldsymbol{T}^{(s)}$ almost surely lies in $(-\infty, 0)^{rs}$, which is the interior of the convex hull of the support of $\mu^{\boldsymbol{T}^{(s)}}$, existence of an MLE directly follows from [12, Theorem 5.7]; see also Definition 3.2 and Assumptions 5.2 therein.

(ii) For $\tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s$, $i = 1, \ldots, s$, and $j = 2, \ldots, r$, we have

$$\frac{\partial}{\partial x_1^{(i)}} T_1^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = -n\frac{f(x_1^{(i)})}{\overline{F}(x_1^{(i)})} = c_{1i}, \text{ say}, \quad \frac{\partial}{\partial x_j^{(i)}} T_1^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = 0,$$

$$\left(\frac{\partial}{\partial x_{j-1}^{(i)}} T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}), \frac{\partial}{\partial x_j^{(i)}} T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)})\right)$$

$$= (n - j + 1)\left(\frac{f(x_{j-1}^{(i)})}{\overline{F}(x_{j-1}^{(i)})}, -\frac{f(x_j^{(i)})}{\overline{F}(x_j^{(i)})}\right) = (b_{ji}, c_{ji}), \text{ say},$$

$$\frac{\partial}{\partial x_l^{(i)}} T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = 0, \quad l \in \{1, \ldots, r\} \setminus \{j - 1, j\}.$$

Thus, the Jacobian matrix $D_{\boldsymbol{T}^{(s)}}$ of $\boldsymbol{T}^{(s)}$ is of the form

$$D_{\boldsymbol{T}^{(s)}}(\tilde{\boldsymbol{x}}^{(s)}) = (\boldsymbol{L}(\boldsymbol{x}^{(1)})|\cdots|\boldsymbol{L}(\boldsymbol{x}^{(s)})), \quad \tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s,$$

where for $1 \leqslant i \leqslant s$,

$$\boldsymbol{L}(\boldsymbol{x}^{(i)}) = \begin{pmatrix} c_{1i} & 0 & 0 & & 0 \\ b_{2i} & c_{2i} & 0 & & 0 \\ 0 & b_{3i} & c_{3i} & & 0 \\ & & \ddots & \ddots & \\ 0 & & & b_{ri} & c_{ri} \end{pmatrix} \in \mathbb{R}^{r \times r}, \quad \boldsymbol{x}^{(i)} \in \mathcal{X}_r,$$

is a lower bidiagonal matrix with determinant

$$\det(\boldsymbol{L}(\boldsymbol{x}^{(i)})) = \prod_{j=1}^r c_{ji} = \prod_{j=1}^r \left(-(n - j + 1)\frac{f(x_j^{(i)})}{\overline{F}(x_j^{(i)})}\right).$$

In particular, since f is positive by assumption, $\det(\boldsymbol{L}(\boldsymbol{x}^{(i)})) \neq 0$, $\boldsymbol{x}^{(i)} \in \mathcal{X}_r$, $1 \leqslant i \leqslant s$. Hence, $\boldsymbol{T}^{(s)}$ is a submersion, i.e., $D_{\boldsymbol{T}^{(s)}}(\tilde{\boldsymbol{x}}^{(s)})$ induces a surjective linear mapping for all $\tilde{\boldsymbol{x}}^{(s)} \in \mathcal{X}_r^s$. The assertion now follows from [20, Theorems 4.1 and 6.1].

Finally, under the assumptions in (i) and (ii), the MLE is necessarily a solution of the likelihood equations, which are readily obtained from formula (2.1) with $\alpha_j = \psi_j(\boldsymbol{\xi})$, $1 \leqslant j \leqslant r$. ∎

The following theorems concern the asymptotic properties of MLEs in curved exponential submodels of SOSs induced by link functions ψ. In order to ensure strong consistency of the MLEs, the conditions stated in Theorem 2.1 should be supplemented with the requirement that ψ forms a homeomorphic transformation of Ξ into $\psi(\Xi)$.

THEOREM 2.2. *Let the assumptions of Theorem 2.1 including those stated in parts* (i) *and* (ii) *be fulfilled, and suppose $\boldsymbol{\psi}$ has a continuous inverse $\boldsymbol{\psi}^{-1}$: $\boldsymbol{\psi}(\Xi) \to \Xi$. Then, for $s \in \mathbb{N}$, the MLE $\hat{\boldsymbol{\xi}}^{(s)}$ of $\boldsymbol{\xi} \in \Xi$ exists and is unique almost surely, and the sequence $(\hat{\boldsymbol{\xi}}^{(s)})_{s \in \mathbb{N}}$ is strongly consistent.*

Proof. Existence and uniqueness of the MLEs follow from Theorem 2.1. Since $\psi(\Xi)$ is homeomorphic to Ξ, it is locally compact. Hence, strong consistency follows from [8, Theorem 4.1 and Corollary 3.3]. ■

Finally, we present a result establishing the asymptotic normality of the MLE. Below, $D_g^2(\boldsymbol{x})$ denotes the Hessian matrix of a twice differentiable function $\boldsymbol{g}$ at $\boldsymbol{x}$, and $\boldsymbol{M}^T$ is the transpose of a matrix $\boldsymbol{M}$.

THEOREM 2.3. *Let Ξ be an open subset of $\mathbb{R}^m$, $m < r$, and $\boldsymbol{\psi} : \Xi \to (0, \infty)^r$ be a C^1-immersion. Then any sequence $(\hat{\boldsymbol{\xi}}^{(s)})_{s \in \mathbb{N}}$ of strongly consistent MLEs is asymptotically efficient, i.e., $\sqrt{s}\,(\hat{\boldsymbol{\xi}}^{(s)} - \boldsymbol{\xi})$ converges in distribution to an m-dimensional normal distribution with zero mean and covariance matrix given by the inverse of the Fisher information matrix $\boldsymbol{I}(\boldsymbol{\xi}) = D_{\boldsymbol{\psi}}(\boldsymbol{\xi})^T D_\kappa^2(\boldsymbol{\psi}(\boldsymbol{\xi})) D_{\boldsymbol{\psi}}(\boldsymbol{\xi})$ at $\boldsymbol{\xi} \in \Xi$.*

Proof. Under the stated assumptions and by recalling the properties of the enclosing exponential family $\mathcal{P}$, the conditions in [8, Section 6] are all met. The assertion then directly follows from Theorem 6.1 therein. ■

The next example demonstrates the applicability of the above results for link function assumptions, where the log-linear link function is a standard choice in the literature.

EXAMPLE 2.2. In the following examples, the baseline distribution function F is a C^2-function with positive Lebesgue-density function f.

(i) In [7] (see also [26, Section 4.5] for a slightly more general setup), a log-linear link function of the form

$$\boldsymbol{\psi}(\boldsymbol{\xi}) = \boldsymbol{\psi}(\xi_1, \xi_2) = (e^{\xi_1 + \xi_2 y_1}, \dots, e^{\xi_1 + \xi_2 y_r}), \quad \boldsymbol{\xi} \in \Xi = \mathbb{R}^2,$$

with positive and strictly ordered known numbers $y_1, \dots, y_r$ is used and the existence and uniqueness of the MLE of $\boldsymbol{\xi}$ are shown using standard tools from differential calculus. With the use of the preceding results, the proof is even simpler. Consequently, by Theorem 2.2, for any $s \in \mathbb{N}$, there exists an almost surely unique MLE $\hat{\boldsymbol{\xi}}^{(s)}$ of $\boldsymbol{\xi} \in \Xi$ and the sequence $(\hat{\boldsymbol{\xi}}^{(s)})_{s \in \mathbb{N}}$ is strongly consistent. According to formula (2.3), for any $s \in \mathbb{N}$, $\hat{\boldsymbol{\xi}}^{(s)}$ can be obtained as a solution to the equations

$$\sum_{j=1}^r y_j^{i-1}\, e^{\xi_1 + y_j \xi_2}\, T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = -s \sum_{j=1}^r y_j^{i-1}, \quad i = 1, 2.$$

Moreover, by Theorem 2.3, the sequence of MLEs is asymptotically efficient with asymptotic covariance matrix at $\boldsymbol{\xi} \in \Xi$ given by

$$
\boldsymbol{I}^{-1}(\boldsymbol{\xi}) = \begin{pmatrix} r & \sum_{i=1}^r y_i \\ \sum_{i=1}^r y_i & \sum_{i=1}^r y_i^2 \end{pmatrix}^{-1}
$$

$$
= \frac{1}{r \sum_{i=1}^r y_i^2 - (\sum_{i=1}^r y_i)^2} \begin{pmatrix} \sum_{i=1}^r y_i^2 & -\sum_{i=1}^r y_i \\ -\sum_{i=1}^r y_i & r \end{pmatrix},
$$

which is seen to be free of $\boldsymbol{\xi}$.

(ii) In [17], SOSs with power trend conditional hazard rates (PTCHR) are introduced (see also [1], [15]), which amounts to assuming $\alpha_j = \xi^j$, $j = 1, \ldots, r$, for some parameter $\xi > 0$. In this situation, the link function takes the form $\boldsymbol{\psi}(\xi) = (\psi_1(\xi), \ldots, \psi_r(\xi)) = (\xi, \xi^2, \ldots, \xi^r)$, $\xi \in \Xi = (0, \infty)$. Hence, by Theorem 2.2, for any $s \in \mathbb{N}$, there exists an almost surely unique MLE $\hat{\xi}^{(s)}$ of $\xi \in \Xi$ and the sequence $(\hat{\xi}^{(s)})_{s \in \mathbb{N}}$ is strongly consistent. According to formula (2.3), for any $s \in \mathbb{N}$, the MLE can be obtained as a solution to the polynomial equation

$$
\sum_{j=1}^r j \xi^j T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) + \frac{s r (r+1)}{2} = 0
$$

with respect to $\xi \in \Xi$. Moreover, by Theorem 2.3, the sequence of MLEs is asymptotically efficient with asymptotic variance given by

$$
I^{-1}(\xi) = \frac{6 \xi^2}{r(r+1)(2r+1)}, \quad \xi \in \Xi.
$$

(iii) If the distribution function F is the standard exponential one, then F_i corresponds to an exponential distribution with hazard rate α_i and expected value $1/\alpha_i$, $i = 1, \ldots, r$; see Section 1. Hence, instead of connecting $\alpha_1, \ldots, \alpha_r$ as in Example 2.1, a linear link function for their reciprocals may also be reasonable. Consider the link function

$$
\boldsymbol{\psi}(\boldsymbol{\xi}) = \boldsymbol{\psi}(\xi_1, \xi_2) = ((\xi_1 + \xi_2 y_1)^{-1}, \ldots, (\xi_1 + \xi_2 y_r)^{-1}), \quad (\xi_1, \xi_2) \in \Xi,
$$

with $\Xi = \{(\eta_1, \eta_2) \in \mathbb{R}^2 : \eta_1 + y_j \eta_2 > 0, \ j = 1, \ldots, r\}$, where $y_1, \ldots, y_r$ are known real numbers taking at least m different values. Again, the assumptions of Theorem 2.2 are satisfied, and we conclude that, for any $s \in \mathbb{N}$, there exists an almost surely unique MLE $\hat{\boldsymbol{\xi}}^{(s)}$ of $\boldsymbol{\xi} \in \Xi$ and the sequence $(\hat{\boldsymbol{\xi}}^{(s)})_{s \in \mathbb{N}}$ is strongly consistent. According to formula (2.3), for any $s \in \mathbb{N}$, $\hat{\boldsymbol{\xi}}^{(s)}$ can be obtained as a solution to the equations

$$
\sum_{j=1}^r \frac{y_j^{i-1}}{(\xi_1 + y_j \xi_2)^2} T_j^{(s)}(\tilde{\boldsymbol{x}}^{(s)}) = -s \sum_{j=1}^r \frac{y_j^{i-1}}{\xi_1 + y_j \xi_2}, \quad i = 1, 2.
$$

Moreover, by Theorem 2.3, the sequence of MLEs is asymptotically efficient with asymptotic covariance matrix at $\boldsymbol{\xi} \in \Xi$ given by

$$
\boldsymbol{I}^{-1}(\boldsymbol{\xi}) = \begin{pmatrix} \sum_{i=1}^{r} \frac{1}{(\xi_1 + \xi_2 y_i)^2} & \sum_{i=1}^{r} \frac{y_i}{(\xi_1 + \xi_2 y_i)^2} \\ \sum_{i=1}^{r} \frac{y_i}{(\xi_1 + \xi_2 y_i)^2} & \sum_{i=1}^{r} \frac{y_i^2}{(\xi_1 + \xi_2 y_i)^2} \end{pmatrix}^{-1} .
$$

3. MINIMAL SUFFICIENCY AND COMPLETENESS

Verifying minimal sufficiency of the canonical statistic $\boldsymbol{T}^{(s)}$ of a curved exponential family $\mathcal{P}_{\boldsymbol{\psi}}^{(s)}$ relies on standard results for general exponential families; see, e.g., [21, Corollary 6.16]). In particular, for minimal sufficiency to hold true, $\boldsymbol{\psi}(\Xi)$ should not be contained in any $(r-1)$-dimensional affine subspace of $\mathbb{R}^r$. In the following example we demonstrate how one would proceed in the setup of Example 2.2(i).

EXAMPLE 3.1. Assume the link function is as in Example 2.2(i). Moreover, let $c_0, c_1, \ldots, c_r \in \mathbb{R}$ with

(3.1) $$ c_1 e^{\xi_1 + y_1 \xi_2} + c_2 e^{\xi_1 + y_2 \xi_2} + \cdots + c_r e^{\xi_1 + y_r \xi_2} = c_0 $$

for all $\boldsymbol{\xi} \in \Xi$. To verify that $\boldsymbol{\psi}(\Xi)$ is not contained in an $(r-1)$-dimensional affine subspace of $\mathbb{R}^r$, we show that formula (3.1) already implies that $c_0 = c_1 = \cdots = c_r = 0$. First, differentiating equation (3.1) with respect to ξ_1 shows that necessarily $c_0 = 0$. Moreover, differentiating equation (3.1) $r-1$ times with respect to ξ_2 and setting $\xi_1 = \xi_2 = 0$, we find that $(c_1, \ldots, c_r)^T$ is a solution to the equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$, $\boldsymbol{x} \in \mathbb{R}^r$, where

(3.2) $$ \boldsymbol{A} = \begin{pmatrix} 1 & \cdots & 1 \\ y_1 & \cdots & y_r \\ \vdots & & \vdots \\ y_1^{r-1} & \cdots & y_r^{r-1} \end{pmatrix} . $$

Since $\boldsymbol{A}^T$ is a Vandermonde matrix and $y_1, \ldots, y_r$ are pairwise distinct by assumption, we have $\det(\boldsymbol{A}) \neq 0$. Consequently, $c_1 = \cdots = c_r = 0$, and hence $\boldsymbol{\psi}(\Xi)$ is not contained in any $(r-1)$-dimensional affine subspace of $\mathbb{R}^r$. Hence, the canonical statistic $\boldsymbol{T}^{(s)}$ is minimal sufficient for $\mathcal{P}_{\boldsymbol{\psi}}^{(s)}$.

In general, the canonical statistic of a curved exponential family is not complete, though examples to the contrary exist; see [22]. If the statistic is indeed not complete, a formal proof of this fact may be required. A sufficient condition for non-completeness is presented in [24, Theorem 1.6.23], which is used in the following example.

EXAMPLE 3.2. Assume the link function is as in Example 2.2(ii). Then the multivariate polynomial $\Pi(x_1, \ldots, x_r) = x_1 \cdots x_r - x_1^{r(r+1)/2}$ satisfies $\Pi(\psi_1(\xi), \ldots, \psi_r(\xi)) = 0$, $\xi \in \Xi$. Consequently, by [24, Theorem 1.6.23], the statistic $\boldsymbol{T}^{(s)}$ is not complete for $\mathcal{P}_{\psi}^{(s)}$. Further, $\boldsymbol{T}^{(s)}$ is not boundedly complete either, which follows from the fact that, for each $\xi \in \Xi$, the density function of $\boldsymbol{T}^{(s)}$,

$$(t_1, \ldots, t_r) \mapsto \exp\left\{ \sum_{j=1}^{r} \xi^j t_j + \frac{sr(r+1)\log(\xi)}{2} \right\}, \quad (t_1, \ldots, t_r) \in (-\infty, 0)^r,$$

is bounded from below on any r-dimensional rectangular region; see [24, Theorem 1.6.23].

The next example demonstrates that the criterion in [24, Theorem 1.6.23] may be of limited applicability.

EXAMPLE 3.3. Assume the link function is as in Example 2.2(i) with $r = 3$ and choose $(y_1, y_2, y_3) = (\sqrt{2}, \sqrt{3}, \sqrt{5})$. Moreover, let $\Pi(x_1, x_2, x_3) = \sum_{l_1, l_2, l_3} a_{l_1, l_2, l_3} x_1^{l_1} x_2^{l_2} x_3^{l_3}$ be a multivariate polynomial satisfying $\Pi(\psi_1, \psi_2, \psi_3) \equiv 0$, i.e.,

$$(3.3) \qquad \sum_{l_1, l_2, l_3} a_{l_1, l_2, l_3} \exp\{\xi_1(l_1 + l_2 + l_3) + \xi_2(y_1 l_1 + y_2 l_2 + y_3 l_3)\} = 0$$

for all $(\xi_1, \xi_2) \in \Xi$. Since $\sqrt{2}, \sqrt{3}, \sqrt{5}$ considered as elements of the splitting field $\mathbb{Q}[\sqrt{2}, \sqrt{3}, \sqrt{5}]$ are linearly independent over $\mathbb{Q}$, we have $y_1 l_1 + y_2 l_2 + y_3 l_3 \neq y_1 \tilde{l}_1 + y_2 \tilde{l}_2 + y_3 \tilde{l}_3$ for $(l_1, l_2, l_3) \neq (\tilde{l}_1, \tilde{l}_2, \tilde{l}_3)$. Let now q be the number of non-zero coefficients of Π. Differentiating equation (3.3) $q - 1$ times with respect to ξ_2 and subsequently setting $\xi_1 = \xi_2 = 0$, we obtain a system of linear equations whose matrix of coefficients takes the form (3.2) with pairwise distinct coefficients in the second row. Arguing in the same way as in Example 3.1, we conclude that all coefficients of Π are zero. Hence, [24, Theorem 1.6.23] cannot be applied here.

4. CONCLUSION

Upon observing that link function assumptions on model parameters of sequential order statistics induce a curved exponential family structure, inference on the underlying parameters in a multi-sample situation is considered. For a rich class of link functions, the curved exponential family structure is utilized to obtain existence and uniqueness results for the maximum likelihood estimators of the link function parameters along with useful asymptotic properties such as strong consistency and asymptotic efficiency. Former results from the literature concerning particular link functions are included in this setup. Finally, minimal sufficiency and completeness of the corresponding canonical statistics are discussed.

REFERENCES

[1] K. Ahmadi, M. Rezaei and F. Yousefzadeh, *Progressively type-II censored competing risks data for exponential distributions based on sequential order statistics*, Comm. Statist. Theory Methods 47 (2018), 1276–1296.

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

[3] S. Bedbur, *UMPU tests based on sequential order statistics*, J. Statist. Plann. Inference 140 (2010), 2520–2530.

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

[5] S. Bedbur, M. Burkschat and U. Kamps, *Inference in a model of successive failures with shape-adjusted hazard rates*, Ann. Inst. Statist. Math. 68 (2016), 639–657.

[6] S. Bedbur, M. Johnen and U. Kamps, *Inference from multiple samples of Weibull sequential order statistics*, J. Multivariate Anal. 169 (2019), 381–399.

[7] S. Bedbur, U. Kamps and M. Kateri, *Meta-analysis of general step-stress experiments under repeated Type-II censoring*, Appl. Math. Model. 39 (2015), 2261–2275.

[8] R. H. Berk, *Consistency and asymptotic normality of MLE's for exponential models*, Ann. Math. Statist. 43 (1972), 193–204.

[9] E. Beutner, *Nonparametric inference for sequential k-out-of-n systems*, Ann. Inst. Statist. Math. 60 (2008), 605–626.

[10] E. Beutner, *Nonparametric model checking for k-out-of-n systems*, J. Statist. Plann. Inference 140 (2010), 626–639.

[11] M. Bieniek, M. Burkschat and T. Rychlik, *Comparisons of the expectations of system and component lifetimes in the failure dependent proportional hazard model*, Methodol. Comput. Appl. Probab. 22 (2020), 173–189.

[12] L. D. Brown, *Fundamentals of Statistical Exponential Families: With Applications In Statistical Decision Theory*, Institute of Mathematical Statistics, Hayward, CA, 1986.

[13] M. Burkschat and J. Navarro, *Dynamic signatures of coherent systems based on sequential order statistics*, J. Appl. Probab. 50 (2013), 272–287.

[14] E. Cramer and U. Kamps, *Sequential k-out-of-n systems*, in: Handbook of Statistics, Vol. 20, Advances in Reliability, N. Balakrishnan and C. R. Rao (eds.), Elsevier, Amsterdam, 2001, 301–372.

[15] M. Doostparast, M. Hashempour and E. Velayati Moghaddam, *Weibull analysis with sequential order statistics under a power trend model for hazard rates*, arXiv:1912.07967 (2019).

[16] M. Esmailian and M. Doostparast, *Estimation based on sequential order statistics with random removals*, Probab. Math. Statist. 34 (2014), 81–95.

[17] M. Esmailian and M. Doostparast, *Estimates based on sequential order statistics with the two-parameter Pareto distribution*, Bull. Malays. Math. Sci. Soc. 42 (2018), 2897–2914.

[18] U. Kamps, *A concept of generalized order statistics*, J. Statist. Plann. Inference 48 (1995), 1–23.

[19] U. Kamps, *A Concept of Generalized Order Statistics*, Teubner, Stuttgart, 1995.

[20] V. Krätschmer, *The uniqueness of extremum estimation*, Statist. Probab. Lett. 77 (2007), 942–951.

[21] E. L. Lehmann and G. Casella, *Theory of Point Estimation*, 2nd ed., Springer, New York, 1998.

[22] M. A. Messig and W. E. Strawderman, *Minimal sufficiency and completeness for dichotomous quantal response models*, Ann. Statist. 21 (1993), 2149–2157.

[23] F. Mies and S. Bedbur, *Exact semiparametric inference and model selection for load-sharing systems*, IEEE Trans. Reliab. 69 (2020), 863–872.

[24] J. Pfanzagl, *Parametric Statistical Theory*, De Gruyter, Berlin, 1994.

[25] T. Rychlik, *Evaluations of generalized order statistics from bounded populations*, Statist. Papers 51 (2010), 165–177.

[26] A. Schmiedt, *Statistical modeling of non-metallic inclusions in steels and extreme value analysis*, Ph.D. thesis, RWTH Aachen Univ., 2012.

Grigoriy Volovskiy, Stefan Bedbur, Udo Kamps
Institute of Statistics
RWTH Aachen University
52056 Aachen, Germany
E-mail: volovskiy@isw.rwth-aachen.de
          bedbur@isw.rwth-aachen.de
          kamps@isw.rwth-aachen.de