Linear estimation of location and scale parameters using partial maxima

Nickos Papadatos

Abstract Consider an i.i.d. sample $X_1^*, X_2^*, \ldots, X_n^*$ from a location-scale family, and assume that the only available observations consist of the partial maxima (or minima) sequence, $X_{1:1}^*, X_{2:2}^*, \ldots, X_{n:n}^*$, where $X_{j:j}^* = \max\{X_1^*, \ldots, X_j^*\}$. This kind of truncation appears in several circumstances, including best performances in athletics events. In the case of partial maxima, the form of the BLUEs (best linear unbiased estimators) is quite similar to the form of the well-known Lloyd's (in Biometrica 39:88–95, 1952) BLUEs, based on (the sufficient sample of) order statistics, but, in contrast to the classical case, their consistency is no longer obvious. The present paper is mainly concerned with the scale parameter, showing that the variance of the partial maxima BLUE is at most of order $O(1/\log n)$, for a wide class of distributions.

Keywords Partial Maxima BLUEs · Location-scale family · Partial Maxima Spacings · S/BSW-type condition · NCP and NCS class · Log-concave distributions · Consistency for the scale estimator · Records

1 Introduction

There are several situations where the ordered random sample,

$$X_{1:n}^* \leq X_{2:n}^* \leq \cdots \leq X_{n:n}^*, \tag{1.1}$$

Devoted to the memory of my six-years-old, little daughter, Dionyssia, who left us on August 25, 2010, at Cephalonia island.

N. Papadatos (✉)
Department of Mathematics, Section of Statistics and O.R.,
University of Athens, Panepistemiopolis, 157 84 Athens, Greece
e-mail: npapadat@math.uoa.gr

corresponding to the i.i.d. random sample, $X_1^*, X_2^*, \ldots, X_n^*$, is not fully reported, because the values of interest are the higher (or lower), up-to-the-present, record values based on the initial sample, i.e., the partial maxima (or minima) sequence

$$X_{1:1}^* \le X_{2:2}^* \le \cdots \le X_{n:n}^*, \tag{1.2}$$

where $X_{j:j}^* = \max\{X_1^*, \ldots, X_j^*\}$. A situation of this kind commonly appears in athletics, when only the best performances are recorded.

Through this article we assume that the i.i.d. data arise from a location-scale family,

$$\{F((\cdot - \theta_1)/\theta_2); \; \theta_1 \in \mathbb{R}, \theta_2 > 0\},$$

where the d.f. $F(\cdot)$ is free of parameters and has finite, non-zero variance (so that F is non-degenerate), and we consider the partial maxima BLUE (best linear unbiased estimator) for both parameters θ_1 and θ_2. This consideration is along the lines of the classical Lloyd (1952) BLUEs, the only difference being that the linear estimators are now based on the "insufficient sample" (1.2), rather than (1.1), and this fact implies a substantial reduction on the available information. Tryfos and Blackmore (1985) used this kind of data to predict future records in athletic events, Samaniego and Whitaker (1986, 1988) estimated the population characteristics, while Hofmann and Nagaraja (2003) investigated the amount of Fisher Information contained in such data; see also Arnold et al. (1998, Sect. 5.9).

A natural question concerns the consistency of the resulting BLUEs, since too much lack of information presumably would result to inconsistency (see at the end of Sect. 6). Thus, our main focus is on conditions guaranteeing consistency, and the main result shows that this is indeed the case for the scale parameter BLUE from a wide class of distributions. Specifically, it is shown that the variance of the BLUE is at most of order $O(1/\log n)$, when $F(x)$ has a log-concave density $f(x)$ and satisfies the Von Mises-type condition (5.11) or (6.1) (cf. Galambos (1978)) on the right end-point of its support (Theorem 5.2, Corollary 6.1). The result is applicable to several commonly used distributions, like the Power distribution (Uniform), the Weibull (Exponential), the Pareto, the Negative Exponential, the Logistic, the Extreme Value (Gumbel) and the Normal (see Sect. 6). A consistency result for the partial maxima BLUE of the location parameter would be desirable to be included here, but it seems that the proposed technique (based on partial maxima spacings, Sect. 4) does not suffice for deriving it. Therefore, the consistency for the location parameter remains an open problem in general, and it is just highlighted by a particular application to the Uniform location-scale family (Sect. 3).

The proof of the main result depends on the fact that, under mild conditions, the partial maxima spacings have non-positive correlation. The class of distributions having this property is called NCP (negative correlation for partial maxima spacings). It is shown here that any log-concave distribution with finite variance belongs to NCP (Theorem 4.2). In particular, if a distribution function has a density which is either log-concave or non-increasing then it is a member of NCP. For ordinary spacings, similar sufficient conditions were shown by Sarkadi (1985) and Bai et al. (1997) – see

also David and Nagaraja (2003, pp. 187–188), Burkschat (2009), Theorem 3.5. These will be referred as "S/BSW-type conditions".

In every experiment where the i.i.d. observations arise in a sequential manner, the partial maxima data describe the best performances in a natural way, as the experiment goes on, in contrast to the first n record values, $R_1, R_2, \ldots, R_n$, which are obtained from an inverse sampling scheme – see, e.g., Berger and Gulati (2001). Due to the very rare appearance of records, in the latter case it is implicitly assumed that the sample size is, roughly, e^n. This has a similar effect in the partial maxima setup, since the number of different values are about $\log n$, for large sample size n. Clearly, the total amount of information in the partial maxima sample is the same as that given by the (few) record values augmented by record times. The essential difference of these models (records / partial maxima) in statistical applications is highlighted, e.g., in Tryfos and Blackmore (1985), Samaniego and Whitaker (1986), Samaniego and Whitaker (1988), Smith (1988), Berger and Gulati (2001) and Hofmann and Nagaraja (2003)—see also Arnold et al. (1998, Chap. 5).

2 Linear estimators based on partial maxima

Consider the random sample $X_1^*, X_2^*, \ldots, X_n^*$ from $F((x - \theta_1)/\theta_2)$ and the corresponding partial maxima sample $X_{1:1}^* \leq X_{2:2}^* \leq \cdots \leq X_{n:n}^*$ ($\theta_1 \in \mathbb{R}$ is the location parameter and $\theta_2 > 0$ is the scale parameter; both parameters are unknown). Let also $X_1, X_2, \ldots, X_n$ and $X_{1:1} \leq X_{2:2} \leq \cdots \leq X_{n:n}$ be the corresponding samples from the completely specified d.f. $F(x)$, that generates the location-scale family. Since

$$(X_{1:1}^*, X_{2:2}^*, \ldots, X_{n:n}^*)' \overset{\mathrm{d}}{=} (\theta_1 + \theta_2 X_{1:1}, \theta_1 + \theta_2 X_{2:2}, \ldots, \theta_1 + \theta_2 X_{n:n})',$$

a linear estimator based on partial maxima has the form

$$L = \sum_{i=1}^n c_i X_{i:i}^* \overset{\mathrm{d}}{=} \theta_1 \sum_{i=1}^n c_i + \theta_2 \sum_{i=1}^n c_i X_{i:i},$$

for some constants c_i, $i = 1, 2, \ldots, n$.

Let $X = (X_{1:1}, X_{2:2}, \ldots, X_{n:n})'$ be the random vector of partial maxima from the known d.f. $F(x)$, and use the notation

$$\boldsymbol{\mu} = \mathbb{E}[X], \quad \boldsymbol{\Sigma} = \mathbb{D}[X] \text{ and } \boldsymbol{E} = \mathbb{E}[XX'], \tag{2.1}$$

where $\mathbb{D}[\boldsymbol{\xi}]$ denotes the dispersion matrix of any random vector $\boldsymbol{\xi}$. Clearly,

$$\boldsymbol{\Sigma} = \boldsymbol{E} - \boldsymbol{\mu}\boldsymbol{\mu}', \quad \boldsymbol{\Sigma} > 0, \quad \boldsymbol{E} > 0.$$

The linear estimator L is called BLUE for θ_k ($k = 1, 2$) if it is unbiased for θ_k and its variance is minimal, while it is called BLIE (best linear invariant estimator) for θ_k if it is invariant for θ_k and its mean squared error, $\mathrm{MSE}[L] = \mathbb{E}[L - \theta_k]^2$, is minimal. Here

"invariance" is understood in the sense of location-scale invariance as it is defined, e.g., in Shao (2005, p. xix).

Using the above notation it is easy to verify the following formulae for the BLUEs and their variances. They are the partial maxima analogues of Lloyd's (1952) estimators and, in the case of partial minima, have been obtained by Tryfos and Blackmore (1985), using least squares. A proof is attached here for easy reference.

Proposition 2.1 *The partial maxima BLUEs for θ_1 and for θ_2 are, respectively,*

$$L_1 = -\frac{1}{\Delta}\boldsymbol{\mu}'\boldsymbol{\Gamma}\boldsymbol{X}^* \quad and \quad L_2 = \frac{1}{\Delta}\boldsymbol{1}'\boldsymbol{\Gamma}\boldsymbol{X}^*, \tag{2.2}$$

where $\boldsymbol{X}^* = (X_{1:1}^*, X_{2:2}^*, \ldots, X_{n:n}^*)'$, $\Delta = (\boldsymbol{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{1})(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}) - (\boldsymbol{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})^2 > 0$, $\boldsymbol{1} = (1, 1, \ldots, 1)' \in \mathbb{R}^n$ and $\boldsymbol{\Gamma} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{1}\boldsymbol{\mu}' - \boldsymbol{\mu}\boldsymbol{1}')\boldsymbol{\Sigma}^{-1}$. The corresponding variances are

$$\mathrm{Var}[L_1] = \frac{1}{\Delta}(\boldsymbol{\mu}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu})\theta_2^2 \quad and \quad \mathrm{Var}[L_2] = \frac{1}{\Delta}(\boldsymbol{1}'\boldsymbol{\Sigma}^{-1}\boldsymbol{1})\theta_2^2. \tag{2.3}$$

Proof Let $\boldsymbol{c} = (c_1, c_2, \ldots, c_n)' \in \mathbb{R}^n$ and $L = \boldsymbol{c}'\boldsymbol{X}^*$. Since $\mathbb{E}[L] = (\boldsymbol{c}'\boldsymbol{1})\theta_1 + (\boldsymbol{c}'\boldsymbol{\mu})\theta_2$, L is unbiased for θ_1 iff $\boldsymbol{c}'\boldsymbol{1} = 1$ and $\boldsymbol{c}'\boldsymbol{\mu} = 0$, while it is unbiased for θ_2 iff $\boldsymbol{c}'\boldsymbol{1} = 0$ and $\boldsymbol{c}'\boldsymbol{\mu} = 1$. Since $\mathrm{Var}[L] = (\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c})\theta_2^2$, a simple minimization argument for $\boldsymbol{c}'\boldsymbol{\Sigma}\boldsymbol{c}$ with respect to $\boldsymbol{c}$, using Lagrange multipliers, yields the expressions (2.2) and (2.3). □

Similarly, one can derive the partial maxima version of Mann's (1969) best linear invariant estimators (BLIEs), as follows.

Proposition 2.2 *The partial maxima BLIEs for θ_1 and for θ_2 are, respectively,*

$$T_1 = \frac{\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{X}^*}{\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{1}} \quad and \quad T_2 = \frac{\boldsymbol{1}'\boldsymbol{G}\boldsymbol{X}^*}{\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{1}}, \tag{2.4}$$

where $\boldsymbol{X}^*$ and $\boldsymbol{1}$ are as in Proposition 2.1 and $\boldsymbol{G} = \boldsymbol{E}^{-1}(\boldsymbol{1}\boldsymbol{\mu}' - \boldsymbol{\mu}\boldsymbol{1}')\boldsymbol{E}^{-1}$. The corresponding mean squared errors are

$$\mathrm{MSE}[T_1] = \frac{\theta_2^2}{\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{1}} \quad and \quad \mathrm{MSE}[T_2] = \left(1 - \frac{D}{\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{1}}\right)\theta_2^2, \tag{2.5}$$

where $D = (\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{1})(\boldsymbol{\mu}'\boldsymbol{E}^{-1}\boldsymbol{\mu}) - (\boldsymbol{1}'\boldsymbol{E}^{-1}\boldsymbol{\mu})^2 > 0$.

Proof Let $L = L(\boldsymbol{X}^*) = \boldsymbol{c}'\boldsymbol{X}^*$ be an arbitrary linear statistic. Since $L(b\boldsymbol{X}^* + a\boldsymbol{1}) = a(\boldsymbol{c}'\boldsymbol{1}) + bL(\boldsymbol{X}^*)$ for arbitrary $a \in \mathbb{R}$ and $b > 0$, it follows that L is invariant for θ_1 iff $\boldsymbol{c}'\boldsymbol{1} = 1$ while it is invariant for θ_2 iff $\boldsymbol{c}'\boldsymbol{1} = 0$. Both (2.4) and (2.5) now follow by a simple minimization argument, since in the first case we have to minimize the mean squared error $\mathbb{E}[L - \theta_1]^2 = (\boldsymbol{c}'\boldsymbol{E}\boldsymbol{c})\theta_2^2$ under $\boldsymbol{c}'\boldsymbol{1} = 1$, while in the second one, we have to minimize the mean squared error $\mathbb{E}[L - \theta_2]^2 = (\boldsymbol{c}'\boldsymbol{E}\boldsymbol{c} - 2\boldsymbol{\mu}'\boldsymbol{c} + 1)\theta_2^2$ under $\boldsymbol{c}'\boldsymbol{1} = 0$. □

The above formulae (2.2)–(2.5) are well-known for order statistics and records—see David (1981, Chap. 6), Arnold et al. (1992, Chap. 7), Arnold et al. (1998, Chap. 5), David and Nagaraja (2003, Chap. 8). In the present setup, however, the meaning of X^*, X, μ, Σ and E is completely different. In the case of order statistics, for example, the vector μ, which is the mean vector of the order statistics $X = (X_{1:n}, X_{2:n}, \ldots, X_{n:n})'$ from the known distribution $F(x)$, depends on the sample size n, in the sense that the components of the vector μ completely change with n. In the present case of partial maxima, the first n entries of the vector μ, which is the mean vector of the partial maxima $X = (X_{1:1}, X_{2:2}, \ldots, X_{n:n})'$ from the known distribution $F(x)$, remain constant for all sample sizes n' greater than or equal to n. Similar observations apply for the matrices Σ and E. This fact seems to be quite helpful for the construction of tables giving the means, variances and covariances of partial maxima for samples up to a size n. It should be noted, however, that even when $F(x)$ is absolutely continuous with density $f(x)$ (as is usually the case for location-scale families), the joint distribution of $(X_{i:i}, X_{j:j})$ has a singular part, since $\mathbb{P}[X_{i:i} = X_{j:j}] = i/j > 0, i < j$. Nevertheless, there exist simple expectation and covariance formulae (Lemma 2.2).

As in the order statistics setup, the actual application of formulae (2.2) and (2.4) requires closed forms for μ and Σ, and also to invert the $n \times n$ matrix Σ. This can be done only for very particular distributions (see next section, where we apply the results to the Uniform distribution). Therefore, numerical methods should be applied in general. This, however, has a theoretical cost: It is not a trivial fact to verify consistency of the estimators, even in the classical case of order statistics. The main purpose of this article is in verifying consistency for the partial maxima BLUEs. Surprisingly, it seems that a solution of this problem is not well-known, at least to our knowledge, even for the classical BLUEs based on order statistics. However, even if the result of the following lemma is known, its proof has an independent interest, because it proposes alternative (to BLUEs) $n^{-1/2}$-consistent unbiased linear estimators and provides the intuition for the derivation of the main result of the present article.

Lemma 2.1 *The classical BLUEs of θ_1 and θ_2, based on order statistics from a location-scale family, created by a distribution $F(x)$ with finite non-zero variance, are consistent. Moreover, their variance is at most of order $O(1/n)$.*

Proof Let $X^* = (X_{1:n}^*, X_{2:n}^*, \ldots, X_{n:n}^*)'$ and $X = (X_{1:n}, X_{2:n}, \ldots, X_{n:n})'$ be the ordered samples from $F((x-\theta_1)/\theta_2)$ and $F(x)$, respectively, so that $X^* \stackrel{\mathrm{d}}{=} \theta_1 \mathbf{1} + \theta_2 X$. Also write $X_1^*, X_2^*, \ldots, X_n^*$ and $X_1, X_2, \ldots, X_n$ for the corresponding i.i.d. samples. We consider the linear estimators

$$S_1 = \overline{X}^* = \frac{1}{n} \sum_{i=1}^n X_i^* \stackrel{\mathrm{d}}{=} \theta_1 + \theta_2 \overline{X}$$

and

$$S_2 = \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n |X_j^* - X_i^*| \stackrel{\mathrm{d}}{=} \frac{\theta_2}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n |X_j - X_i|,$$

i.e., S_1 is the sample mean and S_2 is a multiple of Gini's statistic. Observe that both S_1 and S_2 are linear estimators in order statistics. [In particular, S_2 can be written as $S_2 = 4(n(n-1))^{-1} \sum_{i=1}^{n} (i - (n+1)/2) X_{i:n}^*$.] Clearly, $\mathbb{E}(S_1) = \theta_1 + \theta_2 \mu_0$, $\mathbb{E}(S_2) = \theta_2 \tau_0$, where μ_0 is the mean, $\mathbb{E}(X_1)$, of the distribution $F(x)$ and τ_0 is the positive finite parameter $\mathbb{E}|X_1 - X_2|$. Since F is known, both $\mu_0 \in \mathbb{R}$ and $\tau_0 > 0$ are known constants, and we can construct the linear estimators $U_1 = S_1 - (\mu_0/\tau_0) S_2$ and $U_2 = S_2/\tau_0$. Obviously, $\mathbb{E}(U_k) = \theta_k$, $k = 1, 2$, and both U_1, U_2 are linear estimators of the form $T_n = (1/n) \sum_{i=1}^{n} \delta(i, n) X_{i:n}^*$, with $|\delta(i, n)|$ uniformly bounded for all i and n. If σ_0^2 is the (assumed finite) variance of $F(x)$, it follows that

$$
\begin{aligned}
\mathrm{Var}[T_n] &\leq \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} |\delta(i, n)| |\delta(j, n)| \mathrm{Cov}(X_{i:n}^*, X_{j:n}^*) \\
&\leq \frac{1}{n^2} \left(\max_{1 \leq i \leq n} |\delta(i, n)| \right)^2 \mathrm{Var}(X_{1:n}^* + X_{2:n}^* + \cdots + X_{n:n}^*) \\
&= \frac{1}{n} \left(\max_{1 \leq i \leq n} |\delta(i, n)| \right)^2 \theta_2^2 \sigma_0^2 = O(n^{-1}) \to 0, \quad \text{as } n \to \infty,
\end{aligned}
$$

showing that $\mathrm{Var}(U_k) \to 0$, and thus U_k is consistent for θ_k, $k = 1, 2$. Since L_k has minimum variance among all linear unbiased estimators, it follows that $\mathrm{Var}(L_k) \leq \mathrm{Var}(U_k) \leq O(1/n)$, and the result follows.   □

The above lemma implies that the mean squared error of the BLIEs, based on order statistics, is at most of order $O(1/n)$, since they have smaller mean squared error than the BLUEs, and thus they are also consistent. More important is the fact that, with the technique used in Lemma 2.1, one can avoid all computations involving means, variances and covariances of order statistics, and it does not need to invert any matrix, in order to prove consistency (and in order to obtain $O(n^{-1})$-consistent estimators). Arguments of similar kind will be applied in Sect. 5, when the problem of consistency for the partial maxima BLUE of θ_2 will be taken under consideration.

We now turn in the partial maxima case. Since actual application of partial maxima BLUEs and BLIEs requires the computation of the first two moments of $X = (X_{1:1}, X_{2:2}, \ldots, X_{n:n})'$ in terms of the completely specified d.f. $F(x)$, the following formulae are to be mentioned here (cf. Jones and Balakrishnan 2002).

Lemma 2.2 *Let $X_{1:1} \leq X_{2:2} \leq \cdots \leq X_{n:n}$ be the partial maxima sequence based on an arbitrary d.f. $F(x)$.*

(i) *For $i \leq j$, the joint d.f. of $(X_{i:i}, X_{j:j})$ is*

$$
F_{X_{i:i}, X_{j:j}}(x, y) = \begin{cases} F^j(y) & \text{if } x \geq y, \\ F^i(x) F^{j-i}(y) & \text{if } x \leq y. \end{cases} \tag{2.6}
$$

(ii) *If F has finite first moment, then*

$$\mu_i = \mathbb{E}[X_{i:i}] = \int_0^\infty (1 - F^i(x)) \, dx - \int_{-\infty}^0 F^i(x) dx \qquad (2.7)$$

is finite for all i.

(iii) *If F has finite second moment, then*

$$\sigma_{ij} = \text{Cov}[X_{i:i}, X_{j:j}] = \iint_{-\infty < x < y < \infty} F^i(x)(F^{j-i}(x)$$

$$+ F^{j-i}(y))(1 - F^i(y)) \, dy \, dx \qquad (2.8)$$

is finite and non-negative for all $i \leq j$. In particular,

$$\sigma_{ii} = \sigma_i^2 = \text{Var}[X_{i:i}] = 2 \iint_{-\infty < x < y < \infty} F^i(x)(1 - F^i(y)) \, dy \, dx. \qquad (2.9)$$

Proof (i) is trivial and (ii) is well-known. (iii) follows from Hoeffding's identity

$$\text{Cov}[X, Y] = \int_{-\infty}^\infty \int_{-\infty}^\infty (F_{X,Y}(x, y) - F_X(x) F_Y(y)) \, dy \, dx$$

(see Hoeffding 1940; Lehmann 1966; Jones and Balakrishnan 2002, among others), applied to $(X, Y) = (X_{i:i}, X_{j:j})$ with joint d.f. given by (2.6) and marginals $F^i(x)$ and $F^j(y)$. □

Formulae (2.7)–(2.9) enable the computation of means, variances and covariances of partial maxima, even in the case where the distribution F does not have a density. Tryfos and Blackmore (1985) obtained an expression for the covariance of partial minima involving means and covariances of order statistics from lower sample sizes.

3 A tractable case: the Uniform location-scale family

Let $X_1^*, X_2^*, \ldots, X_n^* \sim U(\theta_1, \theta_1 + \theta_2)$, so that $(X_{1:1}^*, X_{2:2}^*, \ldots, X_{n:n}^*)' \overset{d}{=} \theta_1 \mathbf{1} + \theta_2 X$, where $X = (X_{1:1}, X_{2:2}, \ldots, X_{n:n})'$ is the partial maxima sample from the standard Uniform distribution. Simple calculations, using (2.7)–(2.9), show that the mean vector $\boldsymbol{\mu} = (\mu_i)$ and the dispersion matrix $\boldsymbol{\Sigma} = (\sigma_{ij})$ of X are given by (see also Tryfos and Blackmore (1985), eq. (3.1))

$$\mu_i = \frac{i}{i+1} \quad \text{and} \quad \sigma_{ij} = \frac{i}{(i+1)(j+1)(j+2)} \quad \text{for } 1 \leq i \leq j \leq n.$$

Therefore, $\boldsymbol{\Sigma}$ is a patterned matrix of the form $\sigma_{ij} = a_i b_j$ for $i \leq j$, and thus, its inverse is tridiagonal; see Graybill (1969, Chap. 8), Arnold et al. (1992, Lemma 7.5.1). Specifically,

$$
\boldsymbol{\Sigma}^{-1} = \begin{pmatrix}
\gamma_1 & -\delta_1 & 0 & \cdots & 0 & 0 \\
-\delta_1 & \gamma_2 & -\delta_2 & \cdots & 0 & 0 \\
0 & -\delta_2 & \gamma_3 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \gamma_{n-1} & -\delta_{n-1} \\
0 & 0 & 0 & \cdots & -\delta_{n-1} & \gamma_n
\end{pmatrix}
$$

where

$$
\gamma_i = \frac{4(i+1)^3(i+2)^2}{(2i+1)(2i+3)}, \quad \delta_i = \frac{(i+1)(i+2)^2(i+3)}{2i+3}, \quad i = 1, 2, \ldots, n-1,
$$

$$
\text{and} \quad \gamma_n = \frac{(n+1)^2(n+2)^2}{2n+1}.
$$

Setting $a(n) = \mathbf{1}'\boldsymbol{\Sigma}^{-1}\mathbf{1}$, $b(n) = (\mathbf{1} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\mathbf{1} - \boldsymbol{\mu})$ and $c(n) = (\mathbf{1} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}\mathbf{1}$, we get

$$
a(n) = \frac{(n+1)^2(n+2)^2}{2n+1} - 2\sum_{i=1}^{n-1} \frac{(i+1)(i+2)^2(3i+1)}{(2i+1)(2i+3)} = n^2 + o(n^2),
$$

$$
b(n) = \frac{(n+2)^2}{2n+1} - 2\sum_{i=1}^{n-1} \frac{(i-1)(i+2)}{(2i+1)(2i+3)} = \frac{1}{2}\log n + o(\log n),
$$

$$
c(n) = \frac{(n+1)(n+2)^2}{2n+1} - \sum_{i=1}^{n-1} \frac{(i+2)(4i^2+7i+1)}{(2i+1)(2i+3)} = n + o(n).
$$

Applying (2.3) we obtain

$$
\text{Var}[L_1] = \frac{a(n) + b(n) - 2c(n)}{a(n)b(n) - c^2(n)}\theta_2^2 = \left(\frac{2}{\log n} + o\left(\frac{1}{\log n}\right)\right)\theta_2^2, \quad \text{and}
$$

$$
\text{Var}[L_2] = \frac{a(n)}{a(n)b(n) - c^2(n)}\theta_2^2 = \left(\frac{2}{\log n} + o\left(\frac{1}{\log n}\right)\right)\theta_2^2.
$$

The preceding computation shows that, for the Uniform location-scale family, the partial maxima BLUEs are consistent for both the location and the scale parameters, since their variance goes to zero with the speed of $2/\log n$. This fact, as expected, contradicts the behavior of the ordinary order statistics BLUEs, where the speed of convergence is of order n^{-2} for the variance of both Lloyd's estimators. However, the comparison is quite unfair here, since Lloyd's estimators are based on the complete sufficient statistic $(X_{1:n}^*, X_{n:n}^*)$, and thus the variance of order statistics BLUE is minimal among all unbiased estimators.

On the other hand we should emphasize that, under the same model, the BLUEs (and the BLIEs) based solely on the first n upper records are not even consistent. In fact, the variance of both BLUEs converges to $\theta_2^2/3$, and the MSE of both BLIEs approaches $\theta_2^2/4$, as $n \to \infty$; see Arnold Arnold et al. (1998, Examples 5.3.7 and 5.4.3).

4 Scale estimation and partial maxima spacings

In the classical order statistics setup, Balakrishnan and Papadatos (2002) observed that the computation of BLUE (and BLIE) of the scale parameter is simplified considerably if one uses spacings instead of order statistics – cf. Sarkadi (1985). Their observation applies here too, and simplifies the form of the partial maxima BLUE (and BLIE).

Specifically, define the partial maxima spacings as $Z_i^* = X_{i+1:i+1}^* - X_{i:i}^* \geq 0$ and $Z_i = X_{i+1:i+1} - X_{i:i} \geq 0$, for $i = 1, 2, \ldots, n-1$, and let $\mathbf{Z}^* = (Z_1^*, Z_2^*, \ldots, Z_{n-1}^*)'$ and $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_{n-1})'$. Clearly, $\mathbf{Z}^* \stackrel{\mathrm{d}}{=} \theta_2 \mathbf{Z}$, and any unbiased (or even invariant) linear estimator of θ_2 based on the partial maxima sample, $L = \mathbf{c}' \mathbf{X}^*$, should necessarily satisfy $\sum_{i=1}^{n} c_i = 0$ (see the proofs of Propositions 2.1 and 2.2). Therefore, L can be expressed as a linear function on Z_i^*'s, $L = \mathbf{b}' \mathbf{Z}^*$, where now $\mathbf{b} = (b_1, b_2, \ldots, b_{n-1})' \in \mathbb{R}^{n-1}$. Consider the mean vector $\mathbf{m} = \mathbb{E}[\mathbf{Z}]$, the dispersion matrix $\mathbf{S} = \mathbb{D}[\mathbf{Z}]$, and the second moment matrix $\mathbf{D} = \mathbb{E}[\mathbf{Z}\mathbf{Z}']$ of $\mathbf{Z}$. Clearly, $\mathbf{S} = \mathbf{D} - \mathbf{m}\mathbf{m}'$, $\mathbf{S} > 0$, $\mathbf{D} > 0$, and the vector $\mathbf{m}$ and the matrices $\mathbf{S}$ and $\mathbf{D}$ are of order $n-1$. Using exactly the same arguments as in Balakrishnan and Papadatos (2002), it is easy to verify the following.

Proposition 4.1 *The partial maxima BLUE of θ_2, given in Proposition 2.1, has the alternative form*

$$L_2 = \frac{\mathbf{m}' \mathbf{S}^{-1} \mathbf{Z}^*}{\mathbf{m}' \mathbf{S}^{-1} \mathbf{m}}, \quad with \quad \mathrm{Var}[L_2] = \frac{\theta_2^2}{\mathbf{m}' \mathbf{S}^{-1} \mathbf{m}}, \tag{4.1}$$

while the corresponding BLIE, given in Proposition 2.2, has the alternative form

$$T_2 = \mathbf{m}' \mathbf{D}^{-1} \mathbf{Z}^*, \quad with \quad \mathrm{MSE}[T_2] = (1 - \mathbf{m}' \mathbf{D}^{-1} \mathbf{m})\theta_2^2. \tag{4.2}$$

It should be noted that, in general, the non-negativity of the BLUE of θ_2 does not follow automatically, even for order statistics. In the order statistics setup, this problem was posed by Arnold et al. (1992), and the best known result, till now, is the one given by Bai et al. (1997) and Sarkadi (1985). Even after the slight improvement, given by Balakrishnan and Papadatos (2002) and by Burkschat (2009), the general case remains unsolved. The same question (of non-negativity of the BLUE) arises in the partial maxima setup, and the following theorem provides a partial positive answer. We omit the proof, since it again follows by a straightforward application of the arguments given in Balakrishnan and Papadatos (2002).

Theorem 4.1   (i)   *There exists a constant $a = a_n(F)$, $0 < a < 1$, depending only on the sample size n and the d.f. $F(x)$ (i.e., a is free of the parameters θ_1 and θ_2),*

such that $T_2 = a L_2$. *This constant is given by* $a = \boldsymbol{m}' \boldsymbol{D}^{-1} \boldsymbol{m} = \boldsymbol{m}' \boldsymbol{S}^{-1} \boldsymbol{m}/(1 + \boldsymbol{m}' \boldsymbol{S}^{-1} \boldsymbol{m})$.

(ii) *If either* $n = 2$ *or the (free of parameters) d.f.* $F(x)$ *is such that*

$$\text{Cov}[Z_i, Z_j] \leq 0 \quad \text{for all } i \neq j, \ i, j = 1, \ldots, n - 1, \qquad (4.3)$$

then the partial maxima BLUE (and BLIE) of θ_2 *is non-negative.*

Note that, as in order statistics, the non-negativity of L_2 is equivalent to the fact that the vector $\boldsymbol{S}^{-1} \boldsymbol{m}$ (or, equivalently, the vector $\boldsymbol{D}^{-1} \boldsymbol{m}$) has non-negative entries; see Balakrishnan and Papadatos (2002) and Sarkadi (1985).

Since it is important to know whether (4.3) holds, in the sequel we shall make use of the following definition.

Definition 4.1 A d.f. $F(x)$ with finite second moment (or the corresponding density $f(x)$, if exists)

(i) belongs to the class NCS (negatively correlated spacings) if its order statistics have negatively correlated spacings for all sample sizes $n \geq 2$.
(ii) belongs to the class NCP if it has negatively correlated partial maxima spacings, i.e., if (4.3) holds for all $n \geq 2$.

An important result by Bai et al. (1997) states that a log-concave density $f(x)$ with finite variance belongs to NCS – cf. Sarkadi (1985). We call this sufficient condition as the S/BSW-condition (for ordinary spacings). Burkschat (2009, Theorem 3.5) showed an extended S/BSW-condition, under which the log-concavity of both F and $1 - F$ suffice for the NCS class. Due to the existence of simple formulae like (4.1) and (4.2), the NCS and NCP classes provide useful tools in verifying consistency for the scale estimator, as well as, non-negativity. Our purpose is to prove an S/BSW-type condition for partial maxima (see Theorem 4.2, Corollary 4.1, below). To this end, we first state Lemma 4.1, that will be used in the sequel.

Only through the rest of the present section, we shall use notation $Y_k = \max\{X_1, \ldots, X_k\}$, for any integer $k \geq 1$.

Lemma 4.1 *Fix two integers* i, j, *with* $1 \leq i < j$, *and suppose that the i.i.d. r.v.'s* $X_1, X_2, \ldots$ *have a common d.f.* $F(x)$. *Let* I(*expression*) *denoting the indicator function taking the value* 1, *if the expression holds true, and* 0 *otherwise.*

(i) *The conditional d.f. of* Y_{j+1} *given* Y_j *is*

$$\mathbb{P}[Y_{j+1} \leq y | Y_j] = \begin{cases} 0, & \text{if } y < Y_j \\ F(y), & \text{if } y \geq Y_j \end{cases} = F(y) I(y \geq Y_j), \quad y \in \mathbb{R}.$$

If, in addition, $i + 1 < j$, *then the following property (which is an immediate consequence of the Markovian character of the extremal process) holds:*

$$\mathbb{P}[Y_{j+1} \leq y | Y_{i+1}, Y_j] = \mathbb{P}[Y_{j+1} \leq y | Y_j], \quad y \in \mathbb{R}.$$

(ii) *The conditional d.f. of Y_i given Y_{i+1} is*

$$\mathbb{P}[Y_i \le x | Y_{i+1}] = \begin{cases} \dfrac{F^i(x)}{\sum_{j=0}^i F^j(Y_{i+1}) F^{i-j}(Y_{i+1}-)}, & \text{if } x < Y_{i+1} \\ 1, & \text{if } x \ge Y_{i+1} \end{cases}$$

$$= I(x \ge Y_{i+1}) + I(x < Y_{i+1})$$

$$\times \frac{F^i(x)}{\sum_{j=0}^i F^j(Y_{i+1}) F^{i-j}(Y_{i+1}-)}, \quad x \in \mathbb{R}.$$

If, in addition, $i + 1 < j$, then the following property (which is again an immediate consequence of the Markovian character of the extremal process) holds:

$$\mathbb{P}[Y_i \le x | Y_{i+1}, Y_j] = \mathbb{P}[Y_i \le x | Y_{i+1}], \quad x \in \mathbb{R}.$$

(iii) *Given (Y_{i+1}, Y_j), the random variables Y_i and Y_{j+1} are independent.*

We omit the proof since the assertions are simple by-products of the Markovian character of the process $\{Y_k, \ k \ge 1\}$, which can be embedded in a continuous time extremal process $\{Y(t), \ t > 0\}$; see Resnick (1987, Chap. 4). We merely note that a version of the Radon–Nikodym derivative of F^{i+1} w.r.t. F is given by

$$h_{i+1}(x) = \frac{dF^{i+1}(x)}{dF(x)} = \sum_{j=0}^i F^j(x) F^{i-j}(x-), \quad x \in \mathbb{R}, \tag{4.4}$$

which is equal to $(i + 1)F^i(x)$ only if x is a continuity point of F. To see this, it suffices to verify the identity

$$\int_B dF^{i+1}(x) = \int_B h_{i+1}(x) \, dF(x) \quad \text{for all Borel sets } B \subseteq \mathbb{R}. \tag{4.5}$$

Now (4.5) is proved as follows:

$$\int_B dF^{i+1} = \mathbb{P}(Y_{i+1} \in B)$$

$$= \sum_{j=1}^{i+1} \mathbb{P}\left[Y_{i+1} \in B, \sum_{k=1}^{i+1} I(X_k = Y_{i+1}) = j \right]$$

$$= \sum_{j=1}^{i+1} \sum_{1 \le k_1 < \cdots < k_j \le i+1} \mathbb{P}[X_{k_1} = \cdots = X_{k_j} \in B,$$

$$X_s < X_{k_1} \text{ for } s \notin \{k_1, \ldots, k_j\}]$$

$$= \sum_{j=1}^{i+1} \binom{i+1}{j} \mathbb{P}[X_1 = \cdots = X_j \in B, X_{j+1} < X_1, \ldots, X_{i+1} < X_1]$$

$$= \sum_{j=1}^{i+1} \binom{i+1}{j} \int_B \mathbb{E}\left[\left(\prod_{k=2}^{j} I(X_k = x)\right)\left(\prod_{k=j+1}^{i+1} I(X_k < x)\right)\right] dF(x)$$

$$= \int_B \left(\sum_{j=1}^{i+1} \binom{i+1}{j}(F(x) - F(x-))^{j-1} F^{i+1-j}(x-)\right) dF(x)$$

$$= \int_B h_{i+1}(x) dF(x),$$

where we used the identity $\sum_{j=1}^{i+1} \binom{i+1}{j}(b-a)^{j-1}a^{i+1-j} = \sum_{j=0}^{i} b^j a^{i-j}, a \le b$.

We can now show the main result of this section, which presents an S/BSW-type condition for the partial maxima spacings.

Theorem 4.2 *Assume that the d.f. $F(x)$, with finite second moment, is a log-concave distribution (in the sense that $\log F(x)$ is a concave function in J, where $J = \{x \in \mathbb{R} : 0 < F(x) < 1\}$), and has not an atom at its right end-point, $\omega(F) = \inf\{x \in \mathbb{R} : F(x) = 1\}$. Then, $F(x)$ belongs to the class NCP, i.e., (4.3) holds for all $n \ge 2$.*

Proof For arbitrary r.v.'s $X \ge x_0 > -\infty$ and $Y \le y_0 < +\infty$, with respective d.f.'s F_X, F_Y, we have

$$\mathbb{E}[X] = x_0 + \int_{x_0}^{\infty} (1 - F_X(t)) \, dt \quad \text{and} \quad \mathbb{E}[Y] = y_0 - \int_{-\infty}^{y_0} F_Y(t) \, dt \qquad (4.6)$$

(cf. Papadatos 2001; Jones and Balakrishnan 2002). Assume that $i < j$. By Lemma 4.1(i) and (4.6) applied to $F_X = F_{Y_{j+1}|Y_j}$, it follows that

$$\mathbb{E}[Y_{j+1}| Y_{i+1}, Y_j] = \mathbb{E}[Y_{j+1}| Y_j] = Y_j + \int_{Y_j}^{\infty} (1 - F(t)) \, dt, \quad \text{w.p. 1.}$$

Similarly, by Lemma 4.1(ii) and (4.6) applied to $F_Y = F_{Y_i|Y_{i+1}}$, we conclude that

$$\mathbb{E}[Y_i| Y_{i+1}, Y_j] = \mathbb{E}[Y_i| Y_{i+1}] = Y_{i+1} - \frac{1}{h_{i+1}(Y_{i+1})} \int_{-\infty}^{Y_{i+1}} F^i(t) \, dt, \quad \text{w.p. 1,}$$

where h_{i+1} is given by (4.4). Note that F is continuous on J, since it is log-concave there, and thus, $h_{i+1}(x) = (i+1)F^i(x)$ for $x \in J$. If $\omega(F)$ is finite, $F(x)$ is also continuous at $x = \omega(F)$, by assumption. On the other hand, if $\alpha(F) = \inf\{x : F(x) > 0\}$

is finite, F can be discontinuous at $x = \alpha(F)$, but in this case, $h_{i+1}(\alpha(F)) = F^i(\alpha(F)) > 0$; see (4.4). Thus, in all cases, $h_{i+1}(Y_{i+1}) > 0$ w.p. 1.

By conditional independence of Y_i and Y_{j+1} [Lemma 4.1(iii)], we have

$$
\begin{aligned}
\mathrm{Cov}\,(Z_i, Z_j \,|\, Y_{i+1}, Y_j) &= \mathrm{Cov}\,(Y_{i+1} - Y_i, Y_{j+1} - Y_j \,|\, Y_{i+1}, Y_j) \\
&= -\mathrm{Cov}\,(Y_i, Y_{j+1} | Y_{i+1}, Y_j) = 0, \quad \text{w.p. 1,}
\end{aligned}
$$

so that $\mathbb{E}[\,\mathrm{Cov}\,(Z_i, Z_j \,|\, Y_{i+1}, Y_j)] = 0$, and thus,

$$
\begin{aligned}
\mathrm{Cov}\,[Z_i, Z_j] &= \mathrm{Cov}\,[\mathbb{E}(Z_i | Y_{i+1}, Y_j), \mathbb{E}(Z_j | Y_{i+1}, Y_j)] \\
&\quad + \mathbb{E}[\,\mathrm{Cov}\,(Z_i, Z_j | Y_{i+1}, Y_j)] \\
&= \mathrm{Cov}\,[\mathbb{E}(Y_{i+1} - Y_i | Y_{i+1}, Y_j), \mathbb{E}(Y_{j+1} - Y_j | Y_{i+1}, Y_j)] \\
&= \mathrm{Cov}\,[Y_{i+1} - \mathbb{E}(Y_i | Y_{i+1}, Y_j), \mathbb{E}(Y_{j+1} | Y_{i+1}, Y_j) - Y_j] \\
&= \mathrm{Cov}\,[g(Y_{i+1}), h(Y_j)], \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (4.7)
\end{aligned}
$$

where

$$
g(x) = \begin{cases} \dfrac{1}{(i+1)F^i(x)} \int_{-\infty}^{x} F^i(t)\,dt, & x > \alpha(F), \\[2mm] 0, & \text{otherwise,} \end{cases}
\qquad
h(x) = \int_{x}^{\infty} (1 - F(t))\,dt.
$$

Obviously, $h(x)$ is non-increasing. On the other hand, $g(x)$ is non-decreasing in $\mathbb{R}$. This can be shown as follows. First observe that $g(\alpha(F)) = 0$ if $\alpha(F)$ is finite, while $g(x) > 0$ for $x > \alpha(F)$. Next observe that g is finite and continuous at $x = \omega(F)$ if $\omega(F)$ is finite, as follows by the assumed continuity of F at $x = \omega(F)$ and the fact that F has finite variance. Finally, observe that $F^i(x)$, a product of log-concave functions, is also log-concave in J. Therefore, for arbitrary $y \in J$, the function $d(x) = F^i(x) / \int_{-\infty}^{y} F^i(t)dt, x \in (-\infty, y) \cap J$, is a probability density, and thus, it is a log-concave density with support $(-\infty, y) \cap J$. By Prèkopa (1973) or Dasgupta and Sarkar (1982) it follows that the corresponding distribution function, $D(x) = \int_{-\infty}^{x} d(t)dt = \int_{-\infty}^{x} F^i(t)dt / \int_{-\infty}^{y} F^i(t)dt, x \in (-\infty, y) \cap J$, is a log-concave distribution, and since y is arbitrary, $H(x) = \int_{-\infty}^{x} F^i(t)dt$ is a log-concave function, for $x \in J$. Since F is continuous in J, this is equivalent to the fact that the function

$$
\frac{H'(x)}{H(x)} = \frac{F^i(x)}{\int_{-\infty}^{x} F^i(t)\,dt}, \quad x \in J,
$$

is non-increasing, so that $g(x) = H(x)/((i+1)H'(x))$ is non-decreasing in J.

The desired result follows from (4.7), because the r.v.'s Y_{i+1} and Y_j are positively quadrant dependent (PQD—Lehmann 1966), since it is readily verified that $F_{Y_{i+1}, Y_j}(x, y) \geq F_{Y_{i+1}}(x) F_{Y_j}(y)$ for all x and y [Lemma 2.2(i)]. This completes the proof. $\qquad\qquad\square$

The restriction $F(x) \to 1$ as $x \to \omega(F)$ cannot be removed from the theorem. Indeed, the function

$$F(x) = \begin{cases} 0 & x \leq 0, \\ x/4, & 0 \leq x < 1, \\ 1, & x \geq 1, \end{cases}$$

is a log-concave distribution in $J = (\alpha(F), \omega(F)) = (0, 1)$, for which $\mathrm{Cov}[Z_1, Z_2] = \frac{59}{184320} > 0$. The function g, used in the proof, is given by

$$g(x) = \begin{cases} \max\left\{0, \frac{x}{(i+1)^2}\right\}, & x < 1, \\[2mm] \frac{x-1}{i+1} + \frac{1}{(i+1)^2 4^i}, & x \geq 1, \end{cases}$$

and it is not monotonic.

Since the family of densities with log-concave distributions contains both families of log-concave and non-increasing densities (see, e.g., Prèkopa 1973; Dasgupta and Sarkar 1982; Sengupta and Nanda 1999; Bagnoli and Bergstrom 2005), the following corollary is an immediate consequence of Theorems 4.1 and 4.2.

Corollary 4.1 *Assume that $F(x)$ has finite second moment.*

(i) *If $F(x)$ is a log-concave d.f. (in particular, if $F(x)$ has either a log-concave or a non-increasing (in its interval support) density $f(x)$), then the partial maxima BLUE and the partial maxima BLIE of θ_2 are non-negative.*

(ii) *If $F(x)$ has either a log-concave or a non-increasing (in its interval support) density $f(x)$ then it belongs to the NCP class.*

Sometimes it is asserted that "the distribution of a log-convex density is log-concave" (see, e.g., Sengupta and Nanda (1999), Proposition 1(e)), but this is not correct in its full generality, even if the corresponding r.v. X is non-negative. For example, let $Y \sim$ Weibull with shape parameter $1/2$, and set $X \overset{d}{=} Y | Y < 1$. Then X has density f and d.f. F given by

$$f(x) = \frac{\exp(-\sqrt{1-x})}{2(1-e^{-1})\sqrt{1-x}}, \quad F(x) = \frac{\exp(-\sqrt{1-x}) - e^{-1}}{1 - e^{-1}}, \quad 0 < x < 1,$$

and it is easily checked that $\log f$ is convex in $J = (0, 1)$, while F is not log-concave in J. However, we point out that if $\sup J = +\infty$ then any log-convex density, supported on J, has to be non-increasing in J and, therefore, its distribution is log-concave in J. Examples of log-convex distributions having a log-convex density are given by Bagnoli and Bergstrom (2005).

5 Consistent estimation of the scale parameter

Through this section we always assume that $F(x)$, the d.f. that generates the location-scale family, is non-degenerate and has finite second moment. The main purpose

is to verify consistency for L_2, applying the results of Sect. 4. To this end, we firstly state and prove a simple lemma that goes through the lines of Lemma 2.1. Due to the obvious fact that $\text{MSE}[T_2] \leq \text{Var}[L_2]$, all the results of the present section apply also to the BLIE of θ_2.

Lemma 5.1 *If $F(x)$ belongs to the NCP class then*

(i)

$$\text{Var}[L_2] \leq \frac{\theta_2^2}{\sum_{k=1}^{n-1} m_k^2 / s_k^2},$$ (5.1)

where $m_k = \mathbb{E}[Z_k]$ is the k-th component of the vector **m** and $s_k^2 = s_{kk} = \text{Var}[Z_k]$ is the k-th diagonal entry of the matrix **S**.

(ii) *The partial maxima BLUE, L_2, is consistent if the series*

$$\sum_{k=1}^{\infty} \frac{m_k^2}{s_k^2} = +\infty.$$ (5.2)

Proof Observe that part (ii) is an immediate consequence of part (i), due to the fact that, in contrast to the order statistics setup, m_k and s_k^2 do not depend on the sample size n. Regarding (i), consider the linear unbiased estimator

$$U_2 = \frac{1}{c_n} \sum_{k=1}^{n-1} \frac{m_k}{s_k^2} Z_k^* \overset{\text{d}}{=} \frac{\theta_2}{c_n} \sum_{k=1}^{n-1} \frac{m_k}{s_k^2} Z_k,$$

where $c_n = \sum_{k=1}^{n-1} m_k^2 / s_k^2$. Since $F(x)$ belongs to NCP and the weights of U_2 are positive, it follows that the variance of U_2, which is greater than or equal to the variance of L_2, is bounded by the RHS of (5.1); this completes the proof. □

The proof of the following theorem is now immediate.

Theorem 5.1 *If $F(x)$ belongs to the NCP class and if there exists a finite constant C and a positive integer k_0 such that*

$$\frac{\mathbb{E}[Z_k^2]}{k\mathbb{E}^2[Z_k]} \leq C, \quad \text{for all } k \geq k_0,$$ (5.3)

then

$$\text{Var}[L_2] \leq O\left(\frac{1}{\log n}\right), \quad \text{as } n \to \infty.$$ (5.4)

Proof Since for $k \geq k_0$,

$$\frac{m_k^2}{s_k^2} = \frac{m_k^2}{\mathbb{E}[Z_k^2] - m_k^2} = \frac{1}{\frac{\mathbb{E}[Z_k^2]}{\mathbb{E}^2[Z_k]} - 1} \geq \frac{1}{Ck - 1},$$

the result follows by (5.1). $\qquad\square$

Thus, for proving consistency of order $1/\log n$ into NCP class it is sufficient to verify (5.3) and, therefore, we shall investigate the quantities $m_k = \mathbb{E}[Z_k]$ and $\mathbb{E}[Z_k^2] = s_k^2 + m_k^2$. A simple application of Lemma 2.2, observing that $m_k = \mu_{k+1} - \mu_k$ and $s_k^2 = \sigma_{k+1}^2 - 2\sigma_{k,k+1} + \sigma_k^2$, shows that

$$\mathbb{E}[Z_k] = \int_{-\infty}^{\infty} F^k(x)(1 - F(x)) \, dx, \tag{5.5}$$

$$\mathbb{E}^2[Z_k] = 2 \iint_{-\infty < x < y < \infty} F^k(x)(1 - F(x))F^k(y)(1 - F(y)) \, dy \, dx, \tag{5.6}$$

$$\mathbb{E}[Z_k^2] = 2 \iint_{-\infty < x < y < \infty} F^k(x)(1 - F(y)) \, dy \, dx. \tag{5.7}$$

Therefore, all the quantities of interest can be expressed as integrals in terms of the (completely arbitrary) d.f. $F(x)$ (cf. Jones and Balakrishnan 2002).

For the proof of the main result we finally need the following lemma and its corollary.

Lemma 5.2 (i) *For any* $t > -1$,

$$\lim_{k \to \infty} k^{1+t} \int_0^1 u^k (1 - u)^t \, du = \Gamma(1 + t) > 0. \tag{5.8}$$

(ii) *For any* t *with* $0 \leq t < 1$ *and any* $a > 0$, *there exist positive constants* C_1, C_2, *and a positive integer* k_0 *such that*

$$0 < C_1 < k^{1+t} (\log k)^a \int_0^1 \frac{u^k (1 - u)^t}{L^a(u)} \, du < C_2 < \infty, \text{ for all } k \geq k_0, \tag{5.9}$$

where $L(u) = -\log(1 - u)$.

Proof Part (i) follows by Stirling's formula. For part (ii), with the substitution $u = 1 - e^{-x}$, we write the integral in (5.9) as

$$\frac{1}{k+1} \int_0^\infty (k+1)(1-e^{-x})^k e^{-x} \frac{\exp(-tx)}{x^a} \, dx = \frac{1}{k+1} \mathbb{E}\left[\frac{\exp(-tT)}{T^a}\right],$$

where T has the same distribution as the maximum of $k+1$ i.i.d. standard exponential r.v.'s. It is well-known that $\mathbb{E}[T] = 1^{-1} + 2^{-1} + \cdots + (k+1)^{-1}$. Since the second derivative of the function $x \to e^{-tx}/x^a$ is $x^{-a-2}e^{-tx}(a+(a+tx)^2)$, which is positive for $x > 0$, this function is convex, so by Jensen's inequality we conclude that

$$k^{1+t}(\log k)^a \int_0^1 \frac{u^k(1-u)^t}{(L(u))^a} \, du \geq \frac{k}{k+1} \left(\frac{\log k}{1+1/2+\cdots+1/(k+1)}\right)^a$$

$$\times \exp\left[-t\left(1 + \frac{1}{2} + \cdots + \frac{1}{k+1} - \log k\right)\right],$$

and the RHS remains positive as $k \to \infty$, since it converges to $e^{-\gamma t}$, where $\gamma = .5772\ldots$ is Euler's constant. This proves the lower bound in (5.9). Regarding the upper bound, observe that the function $g(u) = (1-u)^t/L^a(u)$, $u \in (0, 1)$, has second derivative

$$g''(u) = \frac{-(1-u)^{t-2}}{L^{a+2}(u)} [t(1-t)L^2(u) + a(1-2t)L(u) - a(a+1)], \quad 0 < u < 1,$$

and since $0 \leq t < 1$, $a > 0$ and $L(u) \to +\infty$ as $u \to 1-$, it follows that there exists a constant $b \in (0, 1)$ such that $g(u)$ is concave in $(b, 1)$. Split now the integral in (5.9) in two parts,

$$I_k = I_k^{(1)}(b) + I_k^{(2)}(b) = \int_0^b \frac{u^k(1-u)^t}{L^a(u)} \, du + \int_b^1 \frac{u^k(1-u)^t}{L^a(u)} \, du,$$

and observe that for any fixed $s > 0$ and any fixed $b \in (0, 1)$,

$$k^s I_k^{(1)}(b) \leq k^s b^{k-a} \int_0^b \frac{u^a(1-u)^t}{L^a(u)} \, du \leq k^s b^{k-a} \int_0^1 \frac{u^a(1-u)^t}{L^a(u)} \, du \to 0, \quad \text{as } k \to \infty,$$

because the last integral is finite and independent of k. Therefore, $k^{1+t}(\log k)^a I_k$ is bounded above if $k^{1+t}(\log k)^a I_k^{(2)}(b)$ is bounded above for some $b < 1$. Choose b close enough to 1 so that $g(u)$ is concave in $(b, 1)$. By Jensen's inequality and the fact

that $1 - b^{k+1} < 1$ we conclude that

$$I_k^{(2)}(b) = \frac{1 - b^{k+1}}{k+1} \int_b^1 f_k(u)g(u)\, du = \frac{1 - b^{k+1}}{k+1} \mathbb{E}[g(V)] \le \frac{1}{k+1} g[\mathbb{E}(V)],$$

where V is an r.v. with density $f_k(u) = (k+1)u^k/(1 - b^{k+1})$, for $u \in (b, 1)$. Since $\mathbb{E}(V) = ((k+1)/(k+2))(1 - b^{k+2})/(1 - b^{k+1}) > (k+1)/(k+2)$, and g is positive and decreasing (its first derivative is $g'(u) = -(1 - u)^{t-1}L^{-a-1}(u)(tL(u) + a) < 0, 0 < u < 1$), it follows from the above inequality that

$$k^{1+t}(\log k)^a I_k^{(2)}(b) \le \frac{k^{1+t}(\log k)^a}{k+1} g[\mathbb{E}(V)] \le \frac{k^{1+t}(\log k)^a}{k+1} g\left(\frac{k+1}{k+2}\right)$$

$$= \frac{k}{k+1} \left(\frac{\log k}{\log(k+2)}\right)^a \left(\frac{k}{k+2}\right)^t \le 1.$$

This shows that $k^{1+t}(\log k)^a I_k^{(2)}(b)$ is bounded above, and thus, $k^{1+t}(\log k)^a I_k$ is bounded above, as it was to be shown. The proof is complete. $\qquad\square$

Corollary 5.1 (i) *Under the assumptions of Lemma 5.2(i), for any $b \in [0, 1)$ there exist positive constants A_1, A_2 and a positive integer k_0 such that*

$$0 < A_1 < k^{1+t} \int_b^1 u^k(1 - u)^t\, du < A_2 < +\infty, \quad \text{for all } k \ge k_0.$$

(ii) *Under the assumptions of Lemma 5.2(ii), for any $b \in [0, 1)$ there exist positive constants A_3, A_4 and a positive integer k_0 such that*

$$0 < A_3 < k^{1+t}(\log k)^a \int_b^1 \frac{u^k(1 - u)^t}{L^a(u)}\, du < A_4 < \infty, \quad \text{for all } k \ge k_0.$$

Proof The proof follows from Lemma 5.2 in a trivial way, since the corresponding integrals over $[0, b]$ are bounded above by a multiple of b^{k-a}, of the form Ab^{k-a}, with $A < +\infty$ being independent of k. $\qquad\square$

We can now state and prove our main result:

Theorem 5.2 *Assume that $F(x)$ lies in NCP, and let $\omega = \omega(F)$ be the upper end-point of the support of F, i.e., $\omega = \inf\{x \in \mathbb{R} : F(x) = 1\}$, where $\omega = +\infty$ if $F(x) < 1$ for all x. Suppose that $\lim_{x \to \omega-} F(x) = 1$, and that $F(x)$ is differentiable in a left neighborhood (M, ω) of ω, with derivative $f(x) = F'(x)$ for $x \in (M, \omega)$. For $\delta \in \mathbb{R}$ and $\gamma \in \mathbb{R}$, define the (generalized hazard rate) function*

$$L(x) = L(x; \delta, \gamma; F) = \frac{f(x)}{(1 - F(x))^\gamma(-\log(1 - F(x)))^\delta}, \quad x \in (M, \omega), \quad (5.10)$$

and set

$$L_* = L_*(\delta, \gamma; F) = \liminf_{x \to \omega-} L(x; \delta, \gamma, F), \quad L^* = L^*(\delta, \gamma; F) = \limsup_{x \to \omega-} L(x; \delta, \gamma, F).$$

If either (i) for some $\gamma < 3/2$ and $\delta = 0$, or (ii) for some $\delta > 0$ and some γ with $1/2 < \gamma \le 1$,

$$0 < L_*(\delta, \gamma; F) \le L^*(\delta, \gamma; F) < +\infty, \tag{5.11}$$

then the partial maxima BLUE L_2 (given by (2.2) or (4.1)) of the scale parameter θ_2 is consistent and, moreover, $\mathrm{Var}[L_2] \le O(1/\log n)$.

Proof First observe that for large enough $x < \omega$, (5.11) implies that $f(x) > (L_*/2)(1 - F(x))^\gamma (-\log(1 - F(x)))^\delta > 0$, so that $F(x)$ is eventually strictly increasing and continuous. Moreover, the derivative $f(x)$ is necessarily finite since $f(x) < 2L^*(1 - F(x))^\gamma (-\log(1 - F(x)))^\delta$. The assumption $\lim_{x \to \omega-} F(x) = 1$ now shows that $F^{-1}(u)$ is uniquely defined in a left neighborhood of 1, that $F(F^{-1}(u)) = u$ for u close to 1, and that $\lim_{u \to 1-} F^{-1}(u) = \omega$. This, in turn, implies that $F^{-1}(u)$ is differentiable for u close to 1, with (finite) derivative $(F^{-1}(u))' = 1/f(F^{-1}(u)) > 0$. In view of Theorem 5.1, it suffices to verify (5.3), and thus we seek for an upper bound on $\mathbb{E}[Z_k^2]$ and for a lower bound on $\mathbb{E}[Z_k]$. Clearly, (5.3) will be deduced if we shall verify that, under (i), there exist finite constants $C_3 > 0$, $C_4 > 0$ such that

$$k^{3-2\gamma} \mathbb{E}[Z_k^2] \le C_3 \quad \text{and} \quad k^{2-\gamma} \mathbb{E}[Z_k] \ge C_4, \tag{5.12}$$

for all large enough k. Similarly, (5.3) will be verified if we show that, under (ii), there exist finite constants $C_5 > 0$ and $C_6 > 0$ such that

$$k^{3-2\gamma} (\log k)^{2\delta} \mathbb{E}[Z_k^2] \le C_5 \quad \text{and} \quad k^{2-\gamma} (\log k)^\delta \mathbb{E}[Z_k] \ge C_6, \tag{5.13}$$

for all large enough k. Since the integrands in the integral expressions (5.5)–(5.7) vanish if x or y lies outside the set $\{x \in \mathbb{R} : 0 < F(x) < 1\}$, we have the equivalent expressions

$$\mathbb{E}[Z_k] = \int_\alpha^\omega F^k(x)(1 - F(x)) \, dx, \tag{5.14}$$

$$\mathbb{E}[Z_k^2] = 2 \iint_{\alpha < x < y < \omega} F^k(x)(1 - F(y)) \, dy \, dx, \tag{5.15}$$

where α (resp., ω) is the lower (resp., the upper) end-point of the support of F. Obviously, for any fixed M with $\alpha < M < \omega$ and any fixed $s > 0$, we have, as in the proof

of Lemma 5.2(ii), that

$$\lim_{k \to \infty} k^s \int_{\alpha}^{M} F^k(x)(1 - F(x)) \, dx = 0,$$

$$\lim_{k \to \infty} k^s \int_{\alpha}^{M} \int_{x}^{\omega} F^k(x)(1 - F(y)) \, dy \, dx = 0,$$

because $F(M) < 1$ and both integrals (5.14), (5.15) are finite for $k = 1$, by the assumption that the variance is finite ((5.15) with $k = 1$ just equals to the variance of F; see also (2.9) with $i = 1$). Therefore, in order to verify (5.12) and (5.13) for large enough k, it is sufficient to replace $\mathbb{E}[Z_k]$ and $\mathbb{E}[Z_k^2]$, in both formulae (5.12), (5.13), by the integrals $\int_M^\omega F^k(x)(1 - F(x)) dx$ and $\int_M^\omega \int_x^\omega F^k(x)(1 - F(y)) dy dx$, respectively, for an arbitrary (fixed) $M \in (\alpha, \omega)$. Fix now $M \in (\alpha, \omega)$ so large that $f(x) = F'(x)$ exists and it is finite and strictly positive for all $x \in (M, \omega)$, and make the transformation $F(x) = u$ in the first integral, and the transformation $(F(x), F(y)) = (u, v)$ in the second one. Both transformations are now one-to-one and continuous, because both F and F^{-1} are differentiable in their respective intervals (M, ω) and $(F(M), 1)$, and their derivatives are finite and positive. Since $F^{-1}(u) \to \omega$ as $u \to 1-$, it is easily seen that (5.12) will be concluded if it can be shown that for some fixed $b < 1$ (which can be chosen arbitrarily close to 1),

$$k^{3-2\gamma} \int_{b}^{1} \frac{u^k}{f(F^{-1}(u))} \left(\int_{u}^{1} \frac{1-v}{f(F^{-1}(v))} \, dv \right) du \leq C_3 \quad \text{and} \qquad (5.16)$$

$$k^{2-\gamma} \int_{b}^{1} \frac{u^k(1-u)}{f(F^{-1}(u))} \, du \geq C_4, \qquad (5.17)$$

holds for all large enough k. Similarly, (5.13) will be deduced if it will be proved that for some fixed $b < 1$ (which can be chosen arbitrarily close to 1),

$$k^{3-2\gamma}(\log k)^{2\delta} \int_{b}^{1} \frac{u^k}{f(F^{-1}(u))} \left(\int_{u}^{1} \frac{1-v}{f(F^{-1}(v))} \, dv \right) du \leq C_5 \quad \text{and} \quad (5.18)$$

$$k^{2-\gamma}(\log k)^{\delta} \int_{b}^{1} \frac{u^k(1-u)}{f(F^{-1}(u))} \, du \geq C_6, \qquad (5.19)$$

holds for all large enough k. The rest of the proof is thus concentrated on showing (5.16) and (5.17) (resp., (5.18) and (5.19)), under the assumption (i) (resp., under the assumption (ii)).

Assume first that (5.11) holds under (i). Fix now $b < 1$ so large that

$$\frac{L_*}{2}(1 - F(x))^\gamma < f(x) < 2L^*(1 - F(x))^\gamma, \quad \text{for all } x \in (F^{-1}(b), \omega);$$

equivalently,

$$\frac{1}{2L^*} < \frac{(1-u)^\gamma}{f(F^{-1}(u))} < \frac{2}{L_*}, \quad \text{for all } u \in (b, 1). \tag{5.20}$$

Due to (5.20), the inner integral in (5.16) is

$$\int_u^1 \frac{1-v}{f(F^{-1}(v))}\, dv = \int_u^1 (1-v)^{1-\gamma} \frac{(1-v)^\gamma}{f(F^{-1}(v))}\, dv \leq \frac{2(1-u)^{2-\gamma}}{(2-\gamma)L_*}.$$

By Corollary 5.1(i) applied for $t = 2 - 2\gamma > -1$, the LHS of (5.16) is less than or equal to

$$\frac{2k^{3-2\gamma}}{(2-\gamma)L_*} \int_b^1 u^k (1-u)^{2-2\gamma} \frac{(1-u)^\gamma}{f(F^{-1}(u))}\, du \leq \frac{4k^{3-2\gamma}}{(2-\gamma)L_*^2} \int_b^1 u^k (1-u)^{2-2\gamma}\, du \leq C_3,$$

for all $k \geq k_0$, with $C_3 = 4A_2 L_*^{-2}(2-\gamma)^{-1} < \infty$, showing (5.16). Similarly, using the lower bound in (5.20), the integral in (5.17) is

$$\int_b^1 \frac{u^k(1-u)}{f(F^{-1}(u))}\, du = \int_b^1 u^k(1-u)^{1-\gamma} \frac{(1-u)^\gamma}{f(F^{-1}(u))}\, du \geq \frac{1}{2L^*} \int_b^1 u^k(1-u)^{1-\gamma}\, du,$$

so that, by Corollary 5.1(i) applied for $t = 1 - \gamma > -1$, the LHS of (5.17) is greater than or equal to

$$\frac{k^{2-\gamma}}{2L^*} \int_b^1 u^k(1-u)^{1-\gamma}\, du \geq \frac{A_1}{2L^*} > 0, \quad \text{for all } k \geq k_0,$$

showing (5.17).

Assume now that (5.11) is satisfied under (ii). As in part (i), choose a large enough $b < 1$ so that

$$\frac{1}{2L^*} < \frac{(1-u)^\gamma L^\delta(u)}{f(F^{-1}(u))} < \frac{2}{L_*}, \quad \text{for all } u \in (b, 1), \tag{5.21}$$

where $L(u) = -\log(1-u)$. Due to (5.21), the inner integral in (5.18) is

$$\int\limits_u^1 \frac{(1-v)^{1-\gamma}}{L^\delta(v)} \frac{(1-v)^\gamma L^\delta(v)}{f(F^{-1}(v))}\,dv \le \frac{2}{L_*}\int\limits_u^1 \frac{(1-v)^{1-\gamma}}{L^\delta(v)}\,dv \le \frac{2(1-u)^{2-\gamma}}{L_* L^\delta(u)},$$

because $(1-u)^{1-\gamma}/L^\delta(u)$ is decreasing (see the proof of Lemma 5.2(ii)). By Corollary 5.1(ii) applied for $t = 2 - 2\gamma \in [0, 1)$ and $a = 2\delta > 0$, the double integral in (5.18) is less than or equal to

$$\frac{2}{L_*}\int\limits_b^1 \frac{u^k(1-u)^{2-2\gamma}}{L^{2\delta}(u)}\frac{(1-u)^\gamma L^\delta(u)}{f(F^{-1}(u))}\,du \le \frac{4}{L_*^2}\int\limits_b^1 \frac{u^k(1-u)^{2-2\gamma}}{L^{2\delta}(u)}\,du \le \frac{C_5}{k^{3-2\gamma}(\log k)^{2\delta}},$$

for all $k \ge k_0$, with $C_5 = 4A_4 L_*^{-2} < \infty$, showing (5.18). Similarly, using the lower bound in (5.21), the integral in (5.19) is

$$\int\limits_b^1 \frac{u^k(1-u)^{1-\gamma}}{L^\delta(u)}\frac{(1-u)^\gamma L^\delta(u)}{f(F^{-1}(u))}\,du \ge \frac{1}{2L^*}\int\limits_b^1 \frac{u^k(1-u)^{1-\gamma}}{L^\delta(u)}\,du,$$

and thus, by Corollary 5.1(ii) applied for $t = 1 - \gamma \in [0, 1)$ and $a = \delta > 0$, the LHS of (5.19) is greater than or equal to

$$\frac{k^{2-\gamma}(\log k)^\delta}{2L^*}\int\limits_b^1 \frac{u^k(1-u)^{1-\gamma}}{L^\delta(u)}\,du \ge \frac{A_3}{2L^*} > 0, \quad \text{for all } k \ge k_0,$$

showing (5.19). This completes the proof. $\qquad\square$

Remark 5.1 Taking $L(u) = -\log(1-u)$, the limits L_* and L^* in (5.11) can be rewritten as

$$L_*(\delta, \gamma; F) = \liminf_{u\to 1-} \frac{f(F^{-1}(u))}{(1-u)^\gamma L^\delta(u)} = \left(\limsup_{u\to 1-}(F^{-1}(u))'(1-u)^\gamma L^\delta(u)\right)^{-1},$$

$$L^*(\delta, \gamma; F) = \limsup_{u\to 1-} \frac{f(F^{-1}(u))}{(1-u)^\gamma L^\delta(u)} = \left(\liminf_{u\to 1-}(F^{-1}(u))'(1-u)^\gamma L^\delta(u)\right)^{-1}.$$

In the particular case where F is absolutely continuous with a continuous density f and interval support, the function $f(F^{-1}(u)) = 1/(F^{-1}(u))'$ is known as the density-quantile function (Parzen 1979), and plays a fundamental role in the theory of order statistics. Theorem 5.2 shows, in some sense, that the behavior of the density-quantile function at the upper end-point, $u = 1$, specifies the variance behavior of the partial maxima BLUE for the scale parameter θ_2. In fact, (5.11) (and (6.1), below) is a Von Mises-type condition [cf. Galambos (1978, Sects. 2.7, 2.11)].

Remark 5.2 It is obvious that condition $\lim_{x \to \omega-} F(x) = 1$ is necessary for the consistency of BLUE (and BLIE). Indeed, the event that all partial maxima are equal to $\omega(F)$ has probability $p_0 = F(\omega) - F(\omega-)$ (which is independent of n). Thus, a point mass at $x = \omega(F)$ implies that for all n, $\mathbb{P}(L_2 = 0) \geq p_0 > 0$. This situation is trivial. Non-trivial cases also exist, and we provide one at the end of next section.

6 Examples and conclusions

In most commonly used location-scale families, the following corollary suffices for concluding consistency of the BLUE (and the BLIE) of the scale parameter. Its proof follows by a straightforward combination of Corollary 4.1(ii) and Theorem 5.2.

Corollary 6.1 *Suppose that F is absolutely continuous with finite variance, and that its density f is either log-concave or non-increasing in its interval support $J = (\alpha(F), \omega(F)) = \{x \in \mathbb{R} : 0 < F(x) < 1\}$. If, either for some $\gamma < 3/2$ and $\delta = 0$, or for some $\delta > 0$ and some γ with $1/2 < \gamma \leq 1$,*

$$\lim_{x \to \omega(F)-} \frac{f(x)}{(1 - F(x))^\gamma (-\log(1 - F(x)))^\delta} = L \in (0, +\infty), \tag{6.1}$$

then the partial maxima BLUE of the scale parameter is consistent and, moreover, its variance is at most of order $O(1/\log n)$.

Corollary 6.1 has immediate applications to several location-scale families. The following are some of them, where (6.1) can be verified easily. In all these families generated by the distributions mentioned below, the variance of the partial maxima BLUE L_2 (see (2.3) or (4.1)), and the mean squared error of the partial maxima BLIE T_2 (see (2.5) or (4.2)) of the scale parameter is at most of order $O(1/\log n)$, as the sample size $n \to \infty$.

1. Power distribution (Uniform). $F(x) = x^\lambda$, $f(x) = \lambda x^{\lambda-1}$, $0 < x < 1$ ($\lambda > 0$), and $\omega(F) = 1$. The density is non-increasing for $\lambda \leq 1$ and log-concave for $\lambda \geq 1$. It is easily seen that (6.1) is satisfied for $\delta = \gamma = 0$ (for $\lambda = 1$ (Uniform) see Sect. 3).

2. Logistic distribution. $F(x) = (1+e^{-x})^{-1}$, $f(x) = e^{-x}(1+e^{-x})^{-2}$, $x \in \mathbb{R}$, and $\omega(F) = +\infty$. The density is log-concave, and it is easily seen that (6.1) is satisfied for $\delta = 0, \gamma = 1$.

3. Pareto distribution. $F(x) = 1 - x^{-a}$, $f(x) = ax^{-a-1}$, $x > 1$ ($a > 2$, so that the second moment is finite), and $\omega(F) = +\infty$. The density is decreasing, and it is easily seen that (6.1) is satisfied for $\delta = 0, \gamma = 1 + 1/a$. Pareto case provides an example which lies in NCP and not in NCS class—see Bai et al. (1997).

4. Negative Exponential distribution. $F(x) = f(x) = e^x$, $x < 0$, and $\omega(F) = 0$. The density is log-concave and it is easily seen that (6.1) is satisfied for $\delta = \gamma = 0$. This model is particularly important, because it corresponds to the partial minima model from the standard exponential distribution – see Samaniego and Whitaker (1986).

5. Weibull distribution (Exponential). $F(x) = 1 - e^{-x^c}$, $f(x) = cx^{c-1}\exp(-x^c)$, $x > 0$ $(c > 0)$, and $\omega(F) = +\infty$. The density is non-increasing for $c \leq 1$ and log-concave for $c \geq 1$, and it is easily seen that (6.1) is satisfied for $\delta = 1 - 1/c$, $\gamma = 1$. It should be noted that Theorem 5.2 does not apply for $c < 1$, since $\delta < 0$.

6. Gumbel (Extreme Value) distribution. $F(x) = \exp(-e^{-x}) = e^x f(x)$, $x \in \mathbb{R}$, and $\omega(F) = +\infty$. The distribution is log-concave and (6.1) holds with $\gamma = 1$, $\delta = 0$ $(L = 1)$. This model is particularly important for its applications in forecasting records, especially in athletic events—see Tryfos and Blackmore (1985).

7. Normal Distribution. $f(x) = \varphi(x) = (2\pi e^{x^2})^{-1/2}$, $F = \Phi$, $x \in \mathbb{R}$, and $\omega(F) = +\infty$. The density is log-concave and Corollary 6.1 applies with $\delta = 1/2$ and $\gamma = 1$. Indeed,

$$\lim_{+\infty} \frac{\varphi(x)}{(1 - \Phi(x))(-\log(1 - \Phi(x)))^{1/2}} = \lim_{+\infty} \frac{\varphi(x)}{x(1 - \Phi(x))} \frac{x}{(-\log(1 - \Phi(x)))^{1/2}},$$

and it is easily seen that

$$\lim_{+\infty} \frac{\varphi(x)}{x(1 - \Phi(x))} = 1, \quad \lim_{+\infty} \frac{x^2}{-\log(1 - \Phi(x))} = 2,$$

so that $L = \sqrt{2}$.

In many cases of interest (especially in athletic events), best performances are presented as partial minima rather than maxima; see, e.g., Tryfos and Blackmore (1985). Obviously, the present theory applies also to the partial minima setup. The easiest way to convert the present results for the partial minima case is to consider the i.i.d. sample $-X_1, \ldots, -X_n$, arising from $F_{-X}(x) = 1 - F_X(-x-)$, and to observe that $\min\{X_1, \ldots, X_i\} = -\max\{-X_1, \ldots, -X_i\}$, $i = 1, \ldots, n$. Thus, we just have to replace $F(x)$ by $1 - F(-x-)$ in the corresponding formulae.

There are some related problems and questions that, at least to our point of view, seem to be quite interesting. One problem is to verify consistency for the partial maxima BLUE of the location parameter. Another problem concerns the complete characterizations of the NCP and NCS classes (see Definition 4.1), since we only know S/BSW-type sufficient conditions. Also, to prove or disprove the non-negativity of the partial maxima BLUE for the scale parameter, outside the NCP class (as well as for the order statistics BLUE of the scale parameter outside the NCS class).

Some questions concern lower variance bounds for the partial maxima BLUEs. For example we showed in Sect. 3 that the rate $O(1/\log n)$ (which, by Theorem 5.2, is just an upper bound for the variance of L_2) is the correct order for the variance of both estimators in the Uniform location-scale family. Is this the usual case? If it is so, then we could properly standardize the estimators, centering and multiplying them by $(\log n)^{1/2}$. This would result to limit theorems analogues to the corresponding ones for order statistics—e.g., Chernoff et al. (1967), Stigler (1974)—or analogues to the corresponding ones of Pyke (1965), Pyke (1980), for partial maxima spacings instead of ordinary spacings. However, note that the Fisher-Information approach,

in the particular case of the one-parameter (scale) family generated by the standard Exponential distribution, suggests a variance of about $3\theta_2^2/(\log n)^3$ for the minimum variance unbiased estimator (based on partial maxima)—see Hoffman and Nagaraja (2003, Eq. (15) on p. 186).

A final question concerns the construction of approximate BLUEs (for both location and scale) based on partial maxima, analogues to Gupta (1952) simple linear estimators based on order statistics. Such a kind of approximations and/or limit theorems would be especially useful for practical purposes, since the computation of BLUE via its closed formula requires inverting an $n \times n$ matrix. This problem has been partially solved here: For the NCP class, the estimator U_2, given in the proof of Lemma 5.1, is consistent for θ_2 (under the assumptions of Theorem 5.2) and can be computed by a simple formula if we merely know the means and variances of the partial maxima spacings.

Except of the trivial case given in Remark 5.2, above, there exist non-trivial examples where no consistent sequence of unbiased estimators exist for the scale parameter. To see this, we make use of the following result.

Theorem 6.1 (Hofmann and Nagaraja 2003, p. 183) *Let $X_1^*, X_2^*, \ldots, X_n^*$ be an i.i.d. sample from the scale family with distribution function $F(x; \theta_2) = F(x/\theta_2)$ ($\theta_2 > 0$ is the scale parameter) and density $f(x; \theta_2) = f(x/\theta_2)/\theta_2$, where $f(x)$ is known, it has a continuous derivative f', and its support, $J(F) = \{x : f(x) > 0\}$, is one of the intervals $(-\infty, \infty)$, $(-\infty, 0)$ or $(0, +\infty)$.*

(i) *The Fisher Information contained in the partial maxima data $X_{1:1}^* \leq X_{2:2}^* \leq \cdots \leq X_{n:n}^*$ is given by*

$$I_n^{\max} = \frac{1}{\theta_2^2} \sum_{k=1}^n \int_{J(F)} f(x) F^{k-1}(x) \left(1 + \frac{xf'(x)}{f(x)} + \frac{(k-1)xf(x)}{F(x)}\right)^2 dx.$$

(ii) *The Fisher Information contained in the partial minima data $X_{1:1}^* \geq X_{1:2}^* \geq \cdots \geq X_{1:n}^*$ is given by*

$$I_n^{\min} = \frac{1}{\theta_2^2} \sum_{k=1}^n \int_{J(F)} f(x)(1 - F(x))^{k-1} \left(1 + \frac{xf'(x)}{f(x)} - \frac{(k-1)xf(x)}{1 - F(x)}\right)^2 dx.$$

It is clear that for fixed $\theta_2 > 0$, $I_n^{\max}$ and $I_n^{\min}$ both increase with the sample size n. In particular, if $J(F) = (0, \infty)$ then, by Beppo-Levi's Theorem, $I_n^{\min}$ converges (as $n \to \infty$) to its limit

$$I^{\min} = \frac{1}{\theta_2^2} \int_0^\infty \left\{ \mu(x) \left(1 + \frac{xf'(x)}{f(x)} - x\mu(x)\right)^2 + x^2\mu^2(x)\left(\lambda(x) + \mu(x)\right) \right\} dx,$$

(6.2)

where $\lambda(x) = f(x)/(1 - F(x))$ and $\mu(x) = f(x)/F(x)$ is the failure rate and reverse failure rate of f, respectively. Obviously, if $I^{\min} < +\infty$, then the Cramér-Rao inequality shows that no consistent sequence of unbiased estimators exists. This, of course, implies that in the corresponding scale family, any sequence of linear (in partial minima) unbiased estimators is inconsistent. The same is clearly true for the location-scale family, because any linear unbiased estimator for θ_2 in the location-scale family is also a linear unbiased estimator for θ_2 in the corresponding scale family.

In the following we show that there exist distributions with finite variance such that $I^{\min}$ in (6.2) is finite: Define $s = e^{-2}$ and

$$F(x) = \begin{cases} 0, & x \leq 0, \\ \frac{1}{1 - \log(x)}, & 0 < x \leq s, \\ 1 - (ax^2 + bx + c)e^{-x}, & x \geq s, \end{cases}$$

where

$$a = \frac{1}{54}\exp(e^{-2})(18 - 6e^2 + e^4) \simeq 0.599,$$

$$b = -\frac{2}{27}\exp(-2 + e^{-2})(9 - 12e^2 + 2e^4) \simeq -0.339,$$

$$c = \frac{1}{54}\exp(-4 + e^{-2})(18 - 42e^2 + 43e^4) \simeq 0.798.$$

Noting that $F(s) = 1/3$, $F'(s) = e^2/9$ and $F''(s) = -e^4/27$, it can be easily verified that the corresponding density

$$f(x) = \begin{cases} \frac{1}{x(1 - \log(x))^2}, & 0 < x \leq s, \\ (ax^2 + (b - 2a)x + c - b)e^{-x}, & x \geq s, \end{cases}$$

is strictly positive for $x \in (0, \infty)$, processes finite moments of any order, and has continuous derivative

$$f'(x) = \begin{cases} \frac{1 + \log(x)}{x^2(1 - \log(x))^3}, & 0 < x \leq s, \\ -(ax^2 + (b - 4a)x + 2a - 2b + c)e^{-x}, & x \geq s. \end{cases}$$

Now the integrand in (6.2), say $S(x)$, can be written as

$$S(x) = \begin{cases} \frac{1 - 2\log(x)}{x(-\log(x))(1 - \log(x))^3}, & 0 < x \leq s, \\ A(x) + B(x), & x \geq s, \end{cases}$$

where as $x \to +\infty$, $A(x) \sim Ax^4 e^{-x}$ and $B(x) \sim Bx^6 e^{-2x}$, with A, B being positive constants. Therefore, $\int_0^s S(x)dx = \int_2^\infty \frac{1+2y}{y(1+y)^3}dy = \log(3/2) - 5/18 \simeq 0.128$. Also, since $S(x)$ is continuous in $[s, +\infty)$ and $S(x) \sim Ax^4 e^{-x}$ as $x \to +\infty$, it follows that $\int_s^\infty S(x)dx < +\infty$ and $I^{\min}$ is finite.

Numerical integration shows that $\int_s^\infty S(x)dx \simeq 2.77$ and thus, $I^{\min} \simeq 2.9/\theta_2^2 < 3/\theta_2^2$. In view of the Cramér-Rao bound this means that, even if a huge sample of partial minima has been recorded, it is impossible to construct an unbiased scale estimator with variance less than $\theta_2^2/3$. Also, it should be noted that a similar example can be constructed such that $f''(x)$ exists (and is continuous) for all $x > 0$.

Of course the above example can be adapted to the partial maxima case by considering the location-scale family generated by the distribution function

$$F(x) = \begin{cases} (ax^2 - bx + c)e^x, & x \le -s, \\ \frac{-\log(-x)}{1 - \log(-x)}, & -s \le x < 0, \\ 1, & x \ge 0, \end{cases}$$

with s, a, b and c as before.

References

Arnold BC, Balakrishnan N, Nagaraja HN (1992) A first course in order statistics. Wiley, New York

Arnold BC, Balakrishnan N, Nagaraja HN (1998) Records. Wiley, New York

Bai Z, Sarkar SK, Wang W (1997) Positivity of the best unbiased L-estimator of the scale parameter with complete or selected order statistics from location-scale distribution. Stat Probab Lett 32:181–188

Bagnoli M, Bergstrom T (2005) Log-concave probability and its applications. Econ Theory 26:445–469

Balakrishnan N, Papadatos N (2002) The use of spacings in the estimation of a scale parameter. Stat Probab Lett 57:193–204

Berger M, Gulati S (2001) Record-breaking data: a parametric comparison of the inverse-sampling and the random-sampling schemes. J Stat Comput Simul 69:225–238

Burkschat M (2009) Multivariate dependence of spacings of generalized order statistics. J Multivariate Anal 100:1093–1106

Chernoff H, Gastwirth JL, Johns MVJr (1967) Asymptotic distributions of linear combinations of functions of order statistics with applications to estimation. Ann Math Stat 38:52–72

Dasgupta S, Sarkar SK (1982) On TP$_2$ and log-concavity. In: Inequalities in statistics and probability, IMS Lecture Notes-Monograph Series, vol 5, pp 54–58

David HA (1981) Order statistics, 2nd ed. Wiley, New York

David HA, Nagaraja HN (2003) Order statistics, 3rd ed. Wiley, Hoboken

Galambos J (1978) The asymptotic theory of extreme order statistics. Wiley, Belmont

Graybill FA (1969) Introduction to matrices with applications in statistics. Wadsworth, Belmont

Gupta AK (1952) Estimation of the mean and standard deviation of a normal population from a censored sample. Biometrika 39:260–273

Hoeffding W (1940) Masstabinvariante korrelations-theorie. Scrift Math Inst Univ Berlin 5:181–233

Hofmann G, Nagaraja HN (2003) Fisher information in record data. Metrika 57:177–193

Jones MC, Balakrishnan N (2002) How are moments and moments of spacings related to distribution functions? J Stat Plann Inference (C.R. Rao 80th birthday felicitation volume, Part I) 103:377–390

Lehmann EL (1966) Some concepts of dependence. Ann Math Stat 37:1137–1153

Lloyd EH (1952) Least-squares estimation of location and scale parameters using order statistics. Biometrika 39:88–95

Mann NR (1969) Optimum estimators for linear functions of location and scale parameters. Ann Math Stat 40:2149–2155

Papadatos N (2001) Distribution and expectation bounds on order statistics from possibly dependent vatiates. Stat Probab Lett 54:21–31

Parzen E (1979) Nonparametric statistical data modeling (with comments by Tukey JW, Welsch RE, Eddy WF, Lindley DV, Tarter ME, Crow EL, and a rejoinder by the author). J Am Stat Assoc 74:105–131

Prèkopa A (1973) On logarithmic concave measures and functions. Act Sci Math (Szeged) 34:335–343

Pyke R (1965) Spacings (with Discussion). J R Stat Soc Ser B 27:395–449

Pyke R (1980) The asymptotic behavior of spacings under Kakutani's model for interval subdivision. Ann Probab 8:157–163

Resnick SI (1987) Extreme values, regular variation, and point processes. Springer, New York

Samaniego FJ, Whitaker LR (1986) On estimating population characteristics from record-breaking observations I. Parametric results. Naval Res Log Quart 33:531–543

Samaniego FJ, Whitaker LR (1988) On estimating population characteristics from record-breaking observations II. Nonparametric results. Naval Res Log 35:221–236

Sarkadi K (1985) On an estimator of the scale parameter. Stat Decis Suppl 2:231–236

Sengupta D, Nanda AK (1999) Log-concave and concave distributions in reliability. Naval Res Log 46:419–433

Shao J (2005) Mathematical statistics: exercises and solutions. Springer, New York

Smith RL (1988) Forecasting records by maximum likelihood. J Am Stat Assoc 83:331–338

Stigler SM (1974) Linear functions of order statistics with smooth weight functions. Ann Stat 2:676–693. Correction **7**, 466

Tryfos P, Blackmore R (1985) Forecasting records. J Am Stat Assoc 80:46–50