Probab. Th. Rel. Fields 88, 261–268 (1991)

Global sufficiency of extreme order statistics in location models of Weibull type

Frank Marohn
Mathematisch-Geographische Fakultät, Katholische Universität Eichstätt, Ostenstrasse 26–28, W-8078 Eichstätt, Federal Republic of Germany

Received November 15, 1988; in revised form September 25, 1990

Summary. In Janssen and Reiss (1988) it was shown that in a location model of a Weibull type sample with shape parameter $-1<a<1$ the $k(n)$ lower extremes are asymptotically local sufficient. In the present paper we show that even global sufficiency holds. Moreover, it turns out that convergence of the given statistical experiments in the deficiency metric does not only hold for compact parameter sets but for the whole real line.

1. Introduction

In the asymptotic theory of statistics, local as well as global results where established. Take, for example, Le Cam's local and global asymptotic bounds for risk functions of estimates or the local and global asymptotic normality of statistical experiments. In the present paper we formulate a global version of a local result by Janssen and Reiss (1988) in a location model of Weibull type with shape parameter a strictly between -1 and 1. We adopt their notation.

We consider a location family P_t, $t\in\mathbb{R}$, with Lebesgue density $f_t(x)=f(x-t)$, where f is of Weibull type i.e., f has a representation

$$f(x)=\begin{cases} x^a r(x) & \text{for } x>0 \\ 0 & \text{for } x\leqq 0, \end{cases}$$

where r is slowly varying at zero. Throughout, we assume that the density f is known.

Now let $X_1, \ldots, X_n$ be i.i.d. random variables with common distribution P_t and denote by $X_{1:n}, \ldots, X_{n:n}$ the pertaining order statistics arranged in the increasing order. It is well known that the set of order statistics is sufficient i.e., it contains all the information about the unknown parameter. In the present paper we reduce the number of order statistics to the $k(n)$ lower extremes $X_{1:n}, \ldots, X_{k(n):n}$ and calculate an upper bound for the loss of information. We restrict ourselves to the non-regular case $-1<a<1$. Notice that for $a\geqq 1$ the LAN condition holds.

The loss of information is measured in terms of the deficiency of statistical experiments in the sense of Le Cam (1986), see also Strasser (1985). Let $E=(\Omega_1, \mathscr{A}_1, \{P_t: t\in T\})$ and $F=(\Omega_2, \mathscr{A}_2, \{Q_t: t\in T\})$ be statistical experiments, where $T\neq\emptyset$ is an arbitrary set. Denote by $\|\ \|$ the variational distance between probability measures. Moreover, let $\delta(E, F)$ be the deficiency of E w.r.t. F and let $\Delta(E, F)=\max\{\delta(E, F), \delta(F, E)\}$ be the deficiency between E and F. If E is dominated and $(\Omega_2, \mathscr{A}_2)$ is Polish then the Markov kernel criterion holds i.e.,

$$\delta(E, F)=\inf_K \sup_{t\in T} \|Q_t - KP_t\|$$

where the infimum is taken over all Markov kernels $K: \mathscr{A}_2\times\Omega_1\to[0, 1]$ from $(\Omega_1, \mathscr{A}_1)$ to $(\Omega_2, \mathscr{A}_2)$ and, by definition,

$$KP_t(\cdot)=\int K(\cdot|\omega_1)\, dP_t(\omega_1).$$

It is well known that the deficiency Δ is a pseudodistance on the set $\mathscr{E}(T)$ of all experiments given a parameter set T.

In general it is not possible to calculate the deficiency; some exceptions can be found in the papers by Luschgy (1987) and Torgersen (1972). Due to the Markov kernel criterion an upper bound for $\delta(E, F)$ is given by $\sup_{t\in T}\|Q_t - K^* P_t\|$ where K^* is any Markov kernel. The choice of K^* is crucial.

We consider the statistical experiments

$$E_n=(\mathbb{R}^n, \mathbb{B}^n, \{P^n_{\delta_n t}: t\in\mathbb{R}\})$$

$$E_{n,k}=(\mathbb{R}^k, \mathbb{B}^k, \{\mathscr{L}(\delta_n^{-1}(X_{1:n}, \ldots, X_{k:n})|P^n_{\delta_n t}): t\in\mathbb{R}\})$$

$$G=(\mathbb{R}^{\mathbb{N}}, \mathbb{B}^{\mathbb{N}}, \{\mathscr{L}((S_j^{1/(1+a)}+t)_{j\in\mathbb{N}}): t\in\mathbb{R}\})$$

where $(\delta_n)_{n\in\mathbb{N}}$ is specified in Sect. 3 and S_j is the j-th partial sum of independent, standard exponential random variables.

The present paper is organized as follows. In Sect. 2 we establish an upper bound of the deficiency between E_n and $E_{n,k}$. In Sect. 3 the asymptotic global sufficiency of the $k(n)$ smallest order statistics is shown. Some parts of the present paper are proved in a similar way as results in Janssen and Reiss (1988). The crucial problem is to establish a bound for $\delta(E_{n,k}, E_n)$ (see Theorem 2.1).

2. Upper bound for the deficiency between E_n and $E_{n,k}$

Let $-1<a<1$ and let P_t be defined as in Sect. 1. Denote by

$$K_t^{(n,k)}(\cdot|\underline{x})$$

the conditional distribution of $(X_{1:n}, \ldots, X_{n:n})$ given $(X_{1:n}, \ldots, X_{k:n})=(x_1, \ldots, x_k)=:\underline{x}$ under the parameter t. It is well known (see Reiss 1989, Theorem 1.8.1) that

$$K_t^{(n,k)}(\cdot|\underline{x})=\varepsilon_{x_1}\times\ldots\times\varepsilon_{x_k}\times\mathscr{L}(Y_{1:n-k}, \ldots, Y_{n-k:n-k}),$$

where ε_y denotes the Dirac measure at y and the Y_i, $i \in \{1, \ldots, n-k\}$, are i.i.d. random variables with common distribution P_{t,x_k} (the restriction of P_t to the interval (x_k, ∞)).

Denote by F the distribution function of P_0. If $F(x_k - t) < 1$ then the distribution P_{t,x_k} has the Lebesgue density

$$f_{t,x_k} = \frac{f_t}{1 - F(x_k - t)} 1_{[x_k, \infty)}.$$

To obtain an upper bound for $\Delta(E_n, E_{n,k})$ we choose a kernel of the type

$$K^{(n,k)}_{\hat{\kappa}(X_{1:n}, \ldots, X_{k:n})}(\cdot \mid X_{1:n}, \ldots, X_{k:n}),$$

where $\hat{\kappa}$ is an appropriate estimator of the unknown location parameter t. Janssen and Reiss (1988) considered the kernel $K_0^{n,k}$ for their local treatment at 0. In our situation a plausible choice of $\hat{\kappa}$ is the minimum

$$\hat{\kappa}(X_{1:n}, \ldots, X_{k:n}) = X_{1:n}.$$

Using the kernel

$$K_{x_1}^{(n,k(n))}(\cdot \mid \underline{x})$$

we will be able to verify the global sufficiency of the $k(n)$ smallest order statistics. Kernels from an initial parameter estimate have been successfully applied for the derivation of bounds for the deficiency in various cases, see e.g. Helgeland (1982), Mammen (1986), and Weiss (1979). The upper bound for $\Delta(E_n, E_{n,k})$ will depend on the following three auxiliary functions h, g, and ψ:

$$h(y) = y^{\frac{a}{2}} - (y-1)^{\frac{a}{2}}, \qquad y \geqq 1,$$

$$g(x) = \int_1^\infty ((y^{\frac{a}{2}} r^{\frac{1}{2}}(xy) - (y-1)^{\frac{a}{2}} r^{\frac{1}{2}}(x(y-1)))/r^{\frac{1}{2}}(x) - h(y))^2 \, dy,$$

$$\psi(z) = \int_z^\infty h^2(y) \, dy, \qquad z \geqq 1.$$

We remark that $h \in L_2$ and $\psi \equiv 0$ for $a = 0$.

2.1. Theorem. *For* $k \in \{1, \ldots, n\}$, $t \in \mathbb{R}$, *and* $\varepsilon > 0$ *with* $k/n \leqq F(\varepsilon) < 1$ *the following inequality holds:*

$$\|\mathscr{L}((X_{1:n}, \ldots, X_{n:n}) \mid P_t^n) - K_{x_1}^{(n,k)} \mathscr{L}((X_{1:n}, \ldots, X_{k:n}) \mid P_t^n)\| \qquad (2.1)$$
$$\leqq (1 - F(\varepsilon))^{-\frac{1}{2}} (n-k)^{\frac{1}{2}} I_{1,n}(I_{2,n} + I_{3,n,k}) + R_{n,k}$$

where

$$I_{1,n}^2 = \int_{(0,\varepsilon)} r(x_1) x_1^{a+1} \, d\mathscr{L}(X_{1:n} \mid P_0^n)(x_1),$$

$$I_{2,n}^2 = \int_{(0,\varepsilon)} g(x_1) \, d\mathscr{L}(X_{1:n} \mid P_0^n)(x_1),$$

$$I_{3,n,k}^2 = \int_{(0,\varepsilon)} \int_{(0,\varepsilon)} \psi\left(\frac{x_k}{x_1}\right) d\mathscr{L}((X_{1:n}, X_{k:n}) \mid P_0^n)(x_1, x_k),$$

and

$$R_{n,k}=\exp(-n(F(\varepsilon)-k/n)^2/3).$$

Notice that the right-hand side of (2.1) is independent of the parameter t.

Proof of Theorem 2.1. For convenience, we abbreviate the left-hand side of (2.1) by $\rho(n,k,t)$. First note that for $x_1, x_k \in (t, \varepsilon+t)$ we have $F(x_k-x_1)\leqq F(x_k-t)\leqq F(\varepsilon)<1$. Repeating the arguments in Janssen and Reiss (1988) we obtain

$$\rho(n,k,t)\leqq(1-F(\varepsilon))^{-\frac{1}{2}}(n-k)^{\frac{1}{2}}\int_t^{\varepsilon+t}\int_t^{\varepsilon+t}\tilde{d}(x_1-t,x_k-t)$$

$$d\mathscr{L}((X_{1:n},X_{k:n})\,|\,P_t^n)(x_1,x_k)+\mathscr{L}(X_{k:n}\,|\,P_t^n)((\varepsilon+t,\infty)),$$

where

$$\tilde{d}^2(x_1-t,x_k-t)=\int(f^{\frac{1}{2}}(y)-f^{\frac{1}{2}}(y-(x_1-t)))^2\,1_{(x_k-t,\infty)}(y)\,dy.$$

Since $\{P_t\}$ is a location family we obtain

$$\rho(n,k,t)\leqq(1-F(\varepsilon))^{-\frac{1}{2}}(n-k)^{\frac{1}{2}}\int_{(0,\varepsilon)}\int_{(0,\varepsilon)}\tilde{d}(x_1,x_k)$$

$$d\mathscr{L}((X_{1:n},X_{k:n})\,|\,P_0^n)(x_1,x_k)+P_0^n\{X_{k:n}>\varepsilon\}.$$

From Lemma 2.18 in Janssen and Reiss (1988) we know that

$$P_0^n\{X_{k:n}>\varepsilon\}\leqq\exp(-n(F(\varepsilon)-k/n)^2/3) \tag{2.2}$$

for $k/n\leqq F(\varepsilon)$. By substitution and the Minkowski inequality we obtain

$$\tilde{d}(x_1,x_k)=x_1^{\frac{1}{2}}\left(\int_{x_k/x_1}^{\infty}(f^{\frac{1}{2}}(x_1y)-f^{\frac{1}{2}}(x_1(y-1)))^2\,dy\right)^{\frac{1}{2}}$$

$$=r^{\frac{1}{2}}(x_1)\,x_1^{\frac{1+a}{2}}\left(\int_{x_k/x_1}^{\infty}\left(\frac{y^{\frac{a}{2}}r^{\frac{1}{2}}(x_1y)-(y-1)^{\frac{a}{2}}r^{\frac{1}{2}}(x_1(y-1))}{r^{1/2}(x_1)}\right)^2dy\right)^{\frac{1}{2}}$$

$$\leqq r^{\frac{1}{2}}(x_1)\,x_1^{\frac{1+a}{2}}\left(\int_{x_k/x_1}^{\infty}\left(\frac{y^{\frac{a}{2}}r^{\frac{1}{2}}(x_1y)-(y-1)^{\frac{a}{2}}r^{\frac{1}{2}}(x_1(y-1))}{r^{1/2}(x_1)}-h(y)\right)^2dy\right)^{\frac{1}{2}}$$

$$+r^{\frac{1}{2}}(x_1)\,x_1^{\frac{1+a}{2}}\left(\int_{x_k/x_1}^{\infty}h^2(y)\,dy\right)^{\frac{1}{2}}.$$

By the Cauchy-Schwarz inequality and the definition of g and ψ the proof is complete. □

It is obvious that if we utilize the kernel $K_0^{(n,k)}$ instead, we cannot establish an upper bound for the deficiency which is independent of the parameter, so one has to restrict the parameter space to compact subsets in this case.

Let us denote the right-hand side of (2.1) by $D(n,k)$. Notice that $D(n,k)$ is an upper bound for the deficiency between E_n and $E_{n,k}$.

3. Global sufficiency of the $k(n)$ smallest order statistics

Denote again by F the distribution function of P_0. Recall that for the normalizing sequence $\delta_n = F^{-1}(1/n)$ occurring in E_n and $E_{n,k}$ we have (see Bingham et al. 1987, Theorem 1.5.12) $\delta_n = n^{-1/(1+a)} l(1/n)$ where l is a further slowly varying function (at zero). In the sequel we assume that

(A 1) $r(x) = c\exp(\tilde{h}(x))$ for $0 < x < x_0$

where $c > 0$ and $\tilde{h}$ satisfies the condition $|\tilde{h}(x)| \leqq Lx^\gamma$ for some constant $L > 0$ and $\gamma > 0$. Under condition (A 1) we may choose the normalizing sequence

$$\tilde{\delta}_n = ((1+a)/c)^{1/(1+a)} n^{-1/(1+a)}.$$

Note that $\lim_{x \downarrow 0} r(x) = c$ and $\tilde{\delta}_n \sim \delta_n$. From Corollary 5.5.5 in Reiss (1989) one easily deduces

$$\|\mathscr{L}(\tilde{\delta}_n^{-1}(X_{1:n}, \ldots, X_{k:n}) | P_0^n) - \mathscr{L}(S_1^{1/(1+a)}, \ldots, S_k^{1/(1+a)})\| \leqq C\left(\left(\frac{k}{n}\right)^{\gamma/(1+a)} k^{\frac{1}{2}} + \frac{k}{n}\right). \tag{3.1}$$

In addition to (A 1), we require the following conditions:

(A 2) r is absolutely continuous on $(0, \infty)$ and $\int_0^\infty ((r'(x))^2/r(x)) x^a \, dx < \infty$.

(A 3) $\lim_{n \to \infty} \int x^{1+a} \, d\mathscr{L}(\tilde{\delta}_n^{-1} X_{1:n} | P_0^n) < \infty$.

Assumption (A 3) is valid whenever the second moment of X_1 exists. This follows from limit theorems for moments of extremes (Polfeld 1970, p. 45; Resnick 1987, Chap. 2).

We treat the case $a \neq 0$. For $a = 0$ the convergence rates of Janssen and Reiss (1988) carry over.

3.1. Theorem. *Suppose that* (A 1)–(A 3) *hold. Then for every* $\lambda \in (0, 1)$ *there exists a constant* $C > 0$, *such that for* $n \in \mathbb{N}$ *and* $k \leqq \lambda n$ *the following inequality holds:*

$$\Delta(E_n, E_{n,k}) \leqq C((k^{(a-1)/(1+a)} + (k/n)^{\gamma/(1+a)} k^{\frac{1}{2}} + k/n)^{\frac{1}{2}} + n^{\max(a-1, -2\gamma)/(2(1+a))}).$$

Theorem 3.1 implies that the $k(n)$ smallest order statistics $X_{1:n}, \ldots, X_{k(n):n}$ are asymptotically global sufficient in the sense that

$$\lim_{n \to \infty} \Delta(E_n, E_{n,k(n)}) = 0$$

whenever $n \to \infty$ and $k(n) \to \infty$.

Proof. Throughout, C denotes a generic constant which does not depend on n and $k \leqq \lambda n$. Assumptions (A1) and (A3) imply

$$\limsup_{n \to \infty} (n-k)^{1/2} I_{1,n} < \infty.$$

Under assumptions (A1) and (A2) we have $g(x) \leqq C x^{\min\{1-a, 2\gamma\}}$ which was proved in Janssen and Reiss (1988). By the definition of $\mathfrak{F}_n$ we obtain

$$I_{2,n}^2 \leqq C n^{\max\{a-1, -2\gamma\}/(1+a)}.$$

Moreover,

$$\begin{aligned} I_{3,n,k}^2 \leqq & E\psi((S_k/S_1)^{1/(1+a)}) \\ & + \|h\|_2^2 \, \|\mathscr{L}(\mathfrak{F}_n^{-1}(X_{1:n}, X_{k:n}) \mid P_0^n) - \mathscr{L}(S_1^{1/(1+a)}, S_k^{1/(1+a)})\|. \end{aligned}$$

We show

$$E\psi((S_k/S_1)^{1/(1+a)}) \leqq C k^{(a-1)/(1+a)}. \tag{3.2}$$

An application of the mean value theorem yields

$$\psi(z) = \int_z^\infty h^2(y)\, dy \leqq \frac{a^2}{4(1-a)} (z-1)^{a-1}.$$

Let $\delta \in (\frac{1}{2}, 1)$. Since S_1/S_k is equals in distribution $U_{1:k-1}$, where $U_{1:k-1}$ is the minimum of $k-1$ i.i.d. $(0,1)$-uniform random variables ($U_{1:0} := 1$), we get for $k \geqq 2$

$$\begin{aligned} E\psi(U_{1:k-1}^{-1/(1+a)}) \leqq & \frac{a^2}{4(1-a)} \int_{(0,\delta)} (u^{-1/(1+a)} - 1)^{a-1} \, d\mathscr{L}(U_{1:k-1})(u) \\ & + \|h\|_2^2 \, P\{U_{1:k-1} > \delta\}. \end{aligned} \tag{3.3}$$

Applying the exponential bound for order statistics as given in Lemma 3.1.1 of Reiss (1989) [compare with (2.2)] we see that

$$P\{U_{1:k-1} > \delta\} \leqq \exp(-(k-1)(\delta - 1/(k-1))^2/3).$$

Substituting u by $u/(k-1)$ we obtain for the integral on the right-hand side of (3.3)

$$\begin{aligned} & \int_{(0,\delta)} (u^{-1/(1+a)} - 1)^{a-1} \, d\mathscr{L}(U_{1:k-1})(u) \\ & = (k-1)^{\frac{a-1}{1+a}} \int_0^{(k-1)\delta} (u^{\frac{-1}{1+a}} - (k-1)^{\frac{-1}{1+a}})^{a-1} \, d\mathscr{L}((k-1)U_{1:k-1})(u) \\ & \leqq (k-1)^{\frac{a-1}{1+a}} (1 - \delta^{\frac{1}{1+a}})^{a-1} \int_0^{(k-1)\delta} u^{\frac{1-a}{1+a}} \, d\mathscr{L}((k-1)U_{1:k-1})(u), \end{aligned}$$

where the last inequality follows from the fact that for $u \in (0, (k-1)\delta)$ we have

$$u^{-1/(1+a)} - (k-1)^{-1/(1+a)} \geqq (1 - \delta^{1/(1+a)})\, u^{-1/(1+a)}.$$

Since the uniform distribution has moments of arbitrary order, we deduce from Proposition 2.1 in Resnick (1987)

$$\limsup_{k\to\infty} \int_{(0,(k-1)\delta)} u^{(1-a)/(1+a)}\, d\mathscr{L}((k-1)\, U_{1:k-1})(u) < \infty,$$

which completes the proof of (3.2). Since

$$\exp(-n(F(\varepsilon)-k/n)^2/3)) \leqq \exp(-k(F(\varepsilon)-\lambda)^2/(3\lambda))$$

the assertion of Theorem 3.1 follows from the inequalities (2.1) and (3.1). □

In Janssen (1990) it is shown that $\Delta_s(E_n, G) \to 0$ and $\Delta_s(E_{n,k(n)}, G) \to 0$ if $n\to\infty$ and $k(n)\to\infty$, where Δ_s indicates the restriction to the compact parameter set $[0, s]$. A different proof of this result is given in Janssen and Reiss (1988). In the next theorem, we establish the analogous result for the whole real line. Notice that the assumptions of the convergence result of Lindae (Le Cam 1986, Theorem 2, p. 92, see also Remark 2, p. 93) hold only for compact parameter spaces.

3.2. Theorem (Strong convergence). *Assume that* (A 1)–(A 3) *are valid. Then*

(i) $\Delta(E_n, G) \to 0$ *as* $n \to \infty$.

(ii) $\Delta(E_{n,k(n)}, G) \to 0$ *as* $n\to\infty$

whenever $k(n) \leqq n$ *and* $k(n)\to\infty$ *as* $n\to\infty$.

Proof. Theorem 3.1 states that $\Delta(E_n, E_{n,k(n)}) \to 0$ as $n\to\infty$, $k(n)\to\infty$. Now the proof of (i) is a repetition of the proof of Lemma 5.21 in Janssen and Reiss (1988). Assertion (ii) is immediate from $\Delta(E_{n,k(n)}, G) \leqq \Delta(E_{n,k(n)}, E_n) + \Delta(E_n, G)$. □

The last result is concerned with the rate of convergence of $\Delta(E_n, G)$. It follows from straightforward calculations similar to those in the proof of Theorem 5.14 in Janssen and Reiss (1988).

3.3. Theorem. *Under conditions* (A 1)–(A 3) *we have*

$$\Delta(E_n, G) = O(n^{\beta(\gamma, a)}),$$

where

$$\beta(\gamma, a) = \begin{cases} \dfrac{a-1}{2(3+a)} & \text{for } \gamma \geqq \dfrac{2(1+a)}{1-\alpha} \\ \dfrac{(a-1)\gamma}{4(1+a)(1+\gamma)} & \text{for } 0<\gamma<\dfrac{2(1+a)}{1-a}. \end{cases}$$

Acknowledgement. The author would like to thank R.-D. Reiss for stimulating this research work.

References

Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation. Cambridge: Cambridge University Press 1987

Helgeland, J.: Additional observations and statistical information in the case of 1-parameter exponential distributions. Z. Wahrscheinlichkeitstheor. Verw. Geb. **59**, 77–100 (1982)

Janssen, A.: Statistical experiments with nonregular densities. Appendix to the monograph: Janssen, A., Mason, D.M.: Non-standard rank tests. (Lect. Notes Statist., vol. 65) Berlin Heidelberg New York: Springer 1990

Janssen, A., Reiss, R.-D.: Comparison of location models of Weibull type samples and extreme value processes. Probab. Th. Rel. Fields **78**, 273–292 (1988)

LeCam, L.: Asymptotic methods in statistical decision theory. (Springer Series in Statistics) Berlin Heidelberg New York: Springer 1986

Luschgy, H.: Comparison of shift experiments on a Banach space. In: Puri, M.L. et al. (eds.) Mathematical statistics and probability theory, vol. A, pp. 217–230. Dordrecht: Reidel 1987

Mammen, E.: The statistical information contained in additional observations. Ann. Stat. **14**, 665–678 (1986)

Polfeldt, Th.: Asymptotic results in non-regular estimation. Scand. Acturial J. 1–2 (Suppl.), 1–78 (1970)

Reiss, R.-D.: Approximate distributions of order statistics (with applications to nonparametric statistics). (Springer Series in Statistics) Berlin Heidelberg New York: Springer 1989

Resnick, S.I.: Extreme values, regular variation, and point processes. (Appl. Probab., vol. 4) Berlin Heidelberg New York: Springer 1987

Strasser, H.: Mathematical theory of statistics. (Studies in Math., vol. 7) Berlin New York: W. de Gruyter 1985

Torgersen, E.N.: Comparison of translation experiments. Ann. Math. Stat. **43**, 1383–1399 (1972)

Weiss, L.: The asymptotic sufficiency in a class of nonregular cases. Sci. Stat. Can. **V**, 141–150 (1979)