



**HAL**  
open science

# The usage of the largest observations for parameter and quantile estimation for the Gumbel distribution; an efficiency analysis

J Hüsler, J. Tiago de Oliveira

## ► To cite this version:

J Hüsler, J. Tiago de Oliveira. The usage of the largest observations for parameter and quantile estimation for the Gumbel distribution; an efficiency analysis. *Annales de l'ISUP*, 1988, XXXIII (1), pp.41-56. <hal-03672681>

**HAL Id: hal-03672681**

**<https://hal.science/hal-03672681v1>**

Submitted on 19 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

The usage of the largest observations for parameter and quantile estimation for the Gumbel distribution; an efficiency analysis

J. Hüsler and J. Tiago de Oliveira\*

### Abstract

We analyze the efficiency of the method proposed by I. Weissman for the estimation of the parameters and quantiles for the Gumbel distribution. We compare Weissman's method which uses the  $k$  largest observations, with the known (Gumbel) method and prove that essentially, Gumbel's method is more efficient than Weissman's method. This comparison is based on the Cramér-efficiency and on the efficiency with respect to the quantile estimation.

KEY WORDS: Extreme-value distribution, Gumbel distribution, Estimation, Efficiency, Cramér-efficiency.

\*J. Hüsler is Dozent at the Department of Statistics, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland. J. Tiago de Oliveira is Professor at the Department of Statistics, University of Lisbon, 58 Rua da Escola Politecnica, 1294 Lisboa Codex, Portugal. The work was done while the second author was Visiting Professor at the Department of Statistics, University of Bern.

### 1) Introduction

Suppose that we observe  $N=n \cdot k$  random variables  $X_{ij}$ ,  $i=1, \dots, n$  and  $j=1, \dots, k$ . Usually  $k$  denotes the number of observed blocks (years) and  $n$  the number of observed random variables per block (year). We let  $Y_j = \max(X_{ij}, i \leq n)$  and assume that  $n$  is sufficiently large and the  $X_{ij}$ 's such that approximately (for  $n$  large)

$$P\{Y_j < y\} = \Lambda((y - \lambda_n) / \delta_n) = \exp(-\exp(-(y - \lambda_n) / \delta_n))$$

(e.g. we may assume that the  $X_{ij}$ 's are iid. with marginal distribution belonging to the domain of attraction of the extreme value distribution  $\Lambda$ ).

Then we may use the  $k$  values  $Y_j$  for estimating the values  $\lambda_n$  and  $\delta_n$ . This follows the classical method described e.g. in Gumbel (1958). Recently, Weissman (1978) proposed the estimation based on the  $k$ -largest observations of all  $nk$  values  $X_{ij}$ , denoted by

$$Z_{1,N} > Z_{2,N} > \dots > Z_{k,N}.$$

The purpose of this paper is therefore to analyze the efficiency of Weissman's method with respect to the classical one. There are situations where only the maxima  $Y_j$  per block (year) are recorded; in this case the method of Weissman is not applicable, but the classical one. On the other side there are reasonable applications of the second method (Weissman(1978)). For our comparison we assume that both methods might be used in a given application i.e. that the  $Z_{i,N}$ ,  $i \leq k$ , as well as  $Y_j$ ,  $j \leq k$ , can be observed with  $n$  sufficiently large. The comparison uses asymptotic results of both methods with  $n \rightarrow \infty$  and  $k$  large. Hence we assume for the analytic treatment that  $\Lambda$  is the distribution of the  $X_{ij}$ 's.

In the next section we deal with the particular case with  $\delta_n$  known and show that both methods are equally efficient. For the general case with unknown  $\lambda_n$  and  $\delta_n$  we show in Section 3 that the classical method is in general more efficient than Weissman's proposal. This comparison is based on the Cramér-efficiency and the efficiency for the estimation of the quantiles  $\lambda_n + \chi \cdot \delta_n$ , (see Tiago de Oliveira (1982)). A few computations are given for these quantile-estimates in Gomes (1985), comparing the two methods for the more general von-Mises-Jenkinson distribution. Our analysis is much more detailed, giving a full analytic description of the efficiency in the case of the extreme value - Gumbel distribution. A similar comparison between "annual maximum" and "peaks over threshold" methods (which corresponds in some way with Weissman's method) was discussed for flood frequencies by Cunnane (1973) and Rosbjerg (1985); their analysis is also less detailed than ours.

Note that both methods lose strongly information with respect to the full sample; the efficiencies of both methods with respect to the full sample converge to 0 as  $n \rightarrow \infty$ , since the procedures are based on very small particular subsets of the full sample.

2) Estimation of the location

We consider the particular case to estimate  $\lambda$  with  $\delta$  known, say  $\delta=1$ . For simplicity, we assume that  $X_{ij}$  are iid., with marginal distribution  $\Lambda(\lambda, 1)$ :  $P(X_{ij} \leq x) = \exp(-\exp(-x+\lambda))$ . By the independence  $P(Y_j \leq y) = \exp(-\exp(-x+\lambda+\log n)) = \exp(-\exp(-x+\lambda_n))$  with  $\lambda_n = \lambda + \log n$ .

To estimate  $\lambda$  from the  $k$  values  $Y_j$  (method A) we use the maximum likelihood estimator (MLE)  $\hat{\lambda}_A = \hat{\lambda}_n - \log n = -\log(\sum_{j=1}^k e^{-Y_j} / k) - \log n$ . Since  $Y_j \sim \Lambda(\lambda + \log n, 1)$ , it follows that

$$(2.1) \quad n \exp(-Y_j) \sim \text{Exp}(e^{-\lambda})$$

with  $E(e^{-Y_j}) = e^{-\lambda}/n$  and  $\text{Var}(e^{-Y_j}) = e^{-2\lambda}/n^2$ ; in addition we find that  $\log k + \lambda - \hat{\lambda}_A$  has a log  $\Gamma(k, 1)$  distribution with the following moments

$$E(\hat{\lambda}_A) = \log k - \Psi(k) + \lambda = \lambda + \frac{1}{2k} + \frac{1}{12k^2} + O(k^{-4})$$

and

$$\text{Var}(\hat{\lambda}_A) = \Psi'(k) = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} + O(k^{-4})$$

where  $\Psi(\cdot)$  and  $\Psi'(\cdot)$  are the digamma and trigamma functions (see e.g. Abramowitz and Stegun (1965) or Lawless (1982)),

and

$$\begin{aligned} P(\hat{\lambda}_A \leq x) &= P\left(\sum_{j=1}^k n e^{-Y_j} \geq k e^{-x}\right) = P\left(\sum_{j=1}^k n e^{-Y_j + \lambda} \geq k e^{-x + \lambda}\right) = \\ &= \int_{k e^{\lambda - x}}^{\infty} e^{-t} t^{k-1} dt / \Gamma(k). \end{aligned}$$

Hence the mean-squared error of  $\hat{\lambda}_A$  is equal to

$$\begin{aligned} \text{MSE}(\hat{\lambda}_A) &= \Psi'(k) + (E(\hat{\lambda}_A) - \lambda)^2 = \Psi'(k) + (\log k - \Psi(k))^2 \\ &= \frac{1}{k} + \frac{3}{4k^2} + \frac{1}{4k^3} + O(k^{-4}). \end{aligned}$$

If  $k$  is also large, then by the central limit theorem

$$W_k = \sqrt{k} e^{\lambda + \log n} \left( \frac{1}{k} \sum_{j=1}^k e^{-Y_j} - e^{-\lambda - \log n} \right) \xrightarrow[k \rightarrow \infty]{} N(0, 1);$$

since

$$W_k = \sqrt{k} (e^{-(\hat{\lambda}_A - \lambda)} - 1) = \sqrt{k} (\lambda - \hat{\lambda}_A + o_p(\hat{\lambda}_A - \lambda))$$

it implies also

$$\sqrt{k} (\hat{\lambda}_A - \lambda) \xrightarrow[k \rightarrow \infty]{} N(0, 1).$$

The method of Weissman (method B) uses the largest values of all  $N=nk$  observations  $X_{ij}$  which we denote simply by  $Z_1, Z_2, \dots, Z_k$  with

$$Z_1 > Z_2 > \dots > Z_k, \text{ with probability } 1.$$

By our assumptions on the  $X_{ij}$ , we get the likelihood function based on the  $Z_j$ 's

$$L(\lambda) = \frac{N!}{(N-k)!} \exp\left(-\sum_{i=1}^k (Z_i - \lambda) - \sum_{i=1}^k e^{-(Z_i - \lambda)} - (N-k)e^{-(Z_k - \lambda)}\right).$$

Hence, the MLE  $\hat{\lambda}_B$  of  $\lambda$  becomes in this case

$$\hat{\lambda}_B = -\log\left(\frac{\sum_{i=1}^k e^{-Z_i} + (N-k)e^{-Z_k}}{k}\right)$$

which is the usual, known MLE for the extreme-value distribution in a censored sample.

On the other side we may consider the MLE  $\hat{\lambda}_C$  based on the asymptotic properties (as  $n \rightarrow \infty$ ) given in Weissman (1978, Theorem 5)

$$\hat{\lambda}_C = Z_k - \log n$$

We find easily that  $\hat{\lambda}_C - \hat{\lambda}_B = \log(1 - \frac{1}{n}) + \log(1 + \sum_{i=1}^k e^{-(Z_i - Z_k)}) / ((n-1)k)$   
 $= O_p(1/n)$  since  $e^{-Z_i + Z_k} \leq 1$  for all  $i \leq k$ . Hence, asymptotically  $\hat{\lambda}_B$  and  $\hat{\lambda}_C$  have the same properties, and in addition  $n(\hat{\lambda}_B - \hat{\lambda}_C) = O_p(1)$

Following Weissman, the distribution of  $\hat{\lambda}_C$  is asymptotically for  $n \rightarrow \infty$ .

$$P(\hat{\lambda}_C \leq y + \lambda) = \sum_{i=1}^{k-1} \frac{1}{i!} (k e^{-y})^i e^{-k e^{-y}} = \int_0^{\infty} \frac{e^{-t} t^{k-1}}{k e^{-y}} dt / \Gamma(k)$$

and thus

$$P(\hat{\lambda}_C \leq y) = \int_{k e^{\lambda - y}}^{\infty} e^{-t} t^{k-1} dt / \Gamma(k)$$

This implies that  $\hat{\lambda}_C$  and  $\hat{\lambda}_A$  have asymptotically (as  $n \rightarrow \infty$ ) the same distribution, and thus also  $\hat{\lambda}_B$ . Hence, both methods have asymptotically the same efficiency. This holds certainly also in more general cases where the distribution  $F (\neq \Lambda)$  of  $X_{ij}$  belongs to the domain of attraction of  $\Lambda$  with a sufficiently fast rate of convergence of  $F^n$  to  $\Lambda$ .

3) The two-parametric case

We compare now the two methods for the estimation of both location and scale parameters. Again, we assume that the  $X_{ij}$  are iid., with marginal distribution  $\Lambda(\lambda, \delta)$ . Still, we have  $P(Y_j \leq x) = \exp(-\exp(-(x-\lambda-\delta \cdot \log n)/\delta)) = \exp(-\exp(-(x-\lambda_n)/\delta))$  with  $\lambda_n = \lambda + \delta \log n$ .

Hence, by denoting the usual MLE of  $\lambda_n$  and  $\delta$  by  $\hat{\lambda}_n$  and  $\hat{\delta}$ , we get in this case

$$\hat{\lambda}_A = \hat{\lambda}_n - \hat{\delta} \log n \quad \text{and} \quad \hat{\delta}_A = \hat{\delta}$$

By using the well-known asymptotic ( $k \rightarrow \infty$ ) covariance matrix of  $\hat{\lambda}_n$  and  $\hat{\delta}$  (see e.g. Tiago de Oliveira (1963, 1972), Bain (1978)) we find the approximate covariance matrix  $\Sigma_A$  of  $\hat{\lambda}_A$  and  $\hat{\delta}_A$ :

$$\text{Var}(\hat{\lambda}_A) = \delta^2 [1 + 6(1-\gamma - \log n)^2 / \pi^2] / k$$

$$\text{Var}(\hat{\delta}_A) = 6\delta^2 / \pi^2 k,$$

$$\text{Cov}(\hat{\lambda}_A, \hat{\delta}_A) = 6\delta^2(1-\gamma - \log n) / \pi^2 k \quad (< 0 \text{ for } n \geq 2)$$

$$\rho(\hat{\lambda}_A, \hat{\delta}_A) = -(1 + \pi^2 / 6(1-\gamma - \log n)^2)^{-1/2} \quad \text{for } n \geq 2,$$

with  $\gamma = 0.5772$ , Euler's constant. This implies that the correlation coefficient of  $\hat{\lambda}_A$  and  $\hat{\delta}_A$  tends slowly to  $-1$  as  $n \rightarrow \infty$ .

We consider now the estimation based on the  $k$  largest values  $Z_1 > Z_2 > \dots > Z_k$  of the  $X_{ij}$ ,  $i \leq n$ ,  $j \leq k$ . For fixed  $n$  and  $k$  we get the usual MLE  $\hat{\lambda}_B$  and  $\hat{\delta}_B$  of  $\lambda$  and  $\delta$  as in a left-censored sample of size  $n \cdot k$  by using the values  $Z_1, \dots, Z_k$ .

As  $n$  tends to  $\infty$ , we have asymptotically that

$$\hat{\delta}_B = \bar{Z}_k - Z_k + O_p\left(\frac{\log n}{n}\right) \quad \text{with} \quad \bar{Z}_k = \frac{1}{k} \sum_{i=1}^k Z_i$$

$$\hat{\lambda}_B = Z_k - \hat{\delta}_B \log n + O_p(n^{-1})$$

and

$$(Z_1 - \delta \log nk, \dots, Z_k - \delta \log nk) \xrightarrow{d} (m_1, \dots, m_k)$$

where the distribution of  $(m_1, \dots, m_k)$  is given in Weissman. Based on the asymptotic values  $m_i$ , Weissman derived the MVUE's  $\lambda^*$  and  $\delta^*$  of  $\lambda$  and  $\delta$ , respectively,

$$\lambda^* = \delta^* (S_k^{(1)} - \gamma) + m_k \quad \text{and} \quad \delta^* = \bar{m}_{k-1} - m_k$$

where  $S_k^{(1)} = \sum_{j=1}^{k-1} 1/j$ ,  $(S_1^{(1)} = 0)$ , and  $\bar{m}_{k-1} = \sum_{i=1}^{k-1} m_i / (k-1)$ ,

and also the MLE's  $\hat{\lambda}$ ,  $\hat{\delta}$

$$\hat{\lambda} = \lambda^* + a_k \delta^* \quad \text{and} \quad \hat{\delta} = b_k \delta^* = \bar{m}_k - m_k$$

where  $a_k = (k-1)(\log k)/k - (S_k^{(1)} - \gamma)$  and  $b_k = (k-1)/k$ .

Note that  $a_k \rightarrow 0$  and  $b_k \rightarrow 1$  as  $k \rightarrow \infty$ .

The asymptotic covariance-matrices  $\Sigma^*$  and  $\hat{\Sigma}$  of these estimates are given by

$$\text{Var}(\delta^*) = \delta^2 / (k-1)$$

$$\text{Var}(\lambda^*) = \delta^2 [\sigma_k^2 + (S_k^{(1)} - \gamma)^2 / (k-1)]$$

$$\text{Cov}(\lambda^*, \delta^*) = (S_k^{(1)} - \gamma) \delta^2 / (k-1)$$

$$\rho(\lambda^*, \delta^*) = (1 + (k-1) \sigma_k^2 / (S_k^{(1)} - \gamma)^2)^{-1/2} \quad (> 0 \text{ for } k > 2)$$

and

$$\text{Var}(\hat{\delta}) = (k-1) \delta^2 / k^2$$

$$\text{Var}(\hat{\lambda}) = \delta^2(\sigma_k^2 + (k-1)(\log k)^2 / k^2)$$

$$\text{Cov}(\hat{\lambda}, \hat{\delta}) = \delta^2(k-1)(\log k) / k^2,$$

$$\rho(\hat{\lambda}, \hat{\delta}) = (1+k^2\sigma_k^2 / (k-1)(\log k)^2)^{-1/2} \quad (> 0 \text{ for } k \geq 2),$$

where  $\sigma_k^2 = \frac{\pi^2}{6} - s_k^{(2)}$  with  $s_k^{(2)} = \sum_{j=1}^{k-1} 1/j^2$ , ( $s_1^{(2)} = 0$ ).

In both cases the correlation coefficients of  $\lambda^*$  and  $\delta^*$ , respectively  $\hat{\lambda}$  and  $\hat{\delta}$  converge to 1 as  $k \rightarrow \infty$ . In addition, it can be shown that  $\rho(\lambda^*, \delta^*) < \rho(\hat{\lambda}, \hat{\delta})$  for any  $k$ .

By these formulas we find easily the asymptotic covariance-matrix  $\Sigma_B$ . The distribution of  $\hat{\delta}_B \cong (\bar{Z}_k - \log nk) - (Z_k - \log nk)$  is asymptotically equal to the one of  $\bar{m}_k - m_k = \hat{\delta}$ . Similar the distributions of  $\hat{\lambda}_B$  and of  $\lambda^* - \delta^*(S_k^{(1)} - \gamma + b_k \log n) + \delta \log nk$  are asymptotically equal. Hence we have for large  $n$

$$\text{Var}(\hat{\delta}_B) = (k-1)\delta^2 / k^2$$

$$\text{Var}(\hat{\lambda}_B) = \delta^2(\sigma_k^2 + (k-1)(\log n)^2 / k^2)$$

$$\text{Cov}(\hat{\lambda}_B, \hat{\delta}_B) = -\delta^2(k-1)(\log n) / k^2 \quad (< 0 \text{ for } k \geq 2),$$

and

$$\rho(\hat{\lambda}_B, \hat{\delta}_B) = -(1+k^2\sigma_k^2 / (k-1)(\log n)^2)^{-1/2}.$$

It follows that also  $\rho(\hat{\lambda}_B, \hat{\delta}_B)$  converges slowly to -1 as  $n \rightarrow \infty$ , but  $\rho(\hat{\lambda}_B, \hat{\delta}_B) < \rho(\hat{\lambda}_A, \hat{\delta}_A)$ . To compare the two procedures, we use the following efficiency concepts. Well-known is the Cramér-efficiency which is in our case the relative joint efficiency of  $(\hat{\lambda}_B, \hat{\delta}_B)$  with respect to  $(\hat{\lambda}_A, \hat{\delta}_A)$ , denoted by  $\text{Ceff}(B, A)$ ,

$$\text{Ceff}(B, A) = \det(\Sigma_A) / \det(\Sigma_B).$$

Using our evaluations of  $\Sigma_A$  and  $\Sigma_B$ , we get

$$\text{Ceff}(B, A) = \epsilon / (\pi^2(k-1)\sigma_k^2) \rightarrow \epsilon / \pi^2 = 0.6079 \text{ as } k \rightarrow \infty,$$

which is also true for the Cramér-efficiency of  $(\hat{\lambda}, \hat{\delta})$  and  $(\hat{\lambda}_A, \hat{\delta}_A)$ , as well as for estimation of  $\lambda_n$  and  $\delta$  instead of  $\lambda$  and  $\delta$ .

More precisely,  $C\text{eff}(B,A) \sim 6/\pi^2$ , since  $(k-1)\sigma_k^2 = (k-1)(\pi^2/6 - S_k^2) < 1$  and  $k\sigma_k^2 > 1$  for any  $k$  (see e.g. Fichtenholz (1970)).

This shows that method A is much more efficient than method B, for any  $k$  (see Table 1).

Similarly, we find that for the asymptotic estimates  $(\lambda^*, \delta^*)$  and  $(\hat{\lambda}, \hat{\delta})$  of Weissman

$$C\text{eff}(\Sigma^*, \hat{\Sigma}) = (k-1)^2 / k^2 < 1 \text{ for any } k > 2.$$

We note that  $C\text{eff}(B,A)$  is the same also if we estimate  $\lambda_n = \lambda + \delta \log n$  instead of  $\lambda$ . As mentioned in the introduction, the estimation of  $\lambda_n$  (and  $\delta$ ) is used for inference on the yearly maximum  $Y_j$ 's.

Methods A and B can be compared also based on the efficiency presented in Tiago de Oliveira (1982) which leads to define the efficiency as the worst possible in all comparison and is the smallest root of  $\det(\Sigma_A - \mu \Sigma_B) = 0$ , denoted by  $\text{eff}(B,A)$ . It was shown that the product of the two roots is the Cramér-efficiency. Evidently, we get from the above derivations that  $\mu_{\min} \cdot \mu_{\max} = C\text{eff}(B,A) \cong 6/\pi^2$ , which gives that  $\text{eff}(B,A) =$

$\mu_{\min} < \sqrt{6/\pi^2} = 0.7797$  for large  $k$ . To find the exact values of these efficiencies we have following the paper of Tiago de Oliveira (1982) to compare only the behaviour of the estimators of the quantiles. The quantile  $\lambda + X \cdot \delta$  (with probability  $\Lambda(X)$ ) has the estimators  $\hat{\lambda}_A + X\hat{\delta}_A$  and  $\hat{\lambda}_B + X\hat{\delta}_B$  respectively. It follows that  $\text{eff}(B,A)$  is the smallest value of

$$\text{Var}(\hat{\lambda}_A + X\hat{\delta}_A) / \text{Var}(\hat{\lambda}_B + X\hat{\delta}_B) =: \text{eff}(X)$$

where  $X$  varies on  $\mathbf{R}$ . For applications and inference on  $Y_j$ ,

we have to consider the quantile  $\lambda_n + \chi \delta_n = \lambda_n + \chi \delta = \lambda + (X + \log n) \delta$ , which is estimated by  $\hat{\lambda}_A + (X + \log n) \hat{\delta}_A$  and  $\hat{\lambda}_B + (X + \log n) \hat{\delta}_B$ , respectively. Obviously, the quotient of their variances is equal to  $\text{eff}(X + \log n)$ . It is easily seen that

$$\text{Var}(\hat{\lambda}_A + (X + \log n) \hat{\delta}_A) = \delta^2(1 + 6(X + 1 - \gamma)^2 / \pi^2) / k$$

$$\text{Var}(\hat{\lambda}_B + (X + \log n) \hat{\delta}_B) = \delta^2(\sigma_k^2 + (k-1)X^2/k^2),$$

and thus

$$\text{eff}(X + \log n) = (k / (k-1)) (1 + 6(X + 1 - \gamma)^2 / \pi^2) / (X^2 + \sigma_k^2 k^2 / (k-1)).$$

Obviously,  $\text{eff}(X + \log n) \rightarrow (1 + 6(X + 1 - \gamma)^2 / \pi^2) / (X^2 + 1)$  if  $k \rightarrow \infty$ , and  $\text{eff}(X + \log n) \rightarrow 6k / \pi^2(k-1)$  as  $X \rightarrow \pm\infty$  which is smaller than 1 for all  $k > 2$ . Figure 1 gives the typical behaviour of  $\text{eff}(X)$ , for  $k=15$ .

To get  $\text{eff}(B, A)$  we have to find the smallest value of  $\text{eff}(X + \log n)$ . A little algebra shows that the smallest and largest values are attained at  $X_1$  respectively  $X_2$  where

$$X_{1,2} = \frac{k^2 \sigma_k^2 / (k-1) - (1-\gamma)^2 - \pi^2 / 6}{2(1-\gamma)} \mp \left( \left( \frac{k^2 \sigma_k^2 / (k-1) - (1-\gamma)^2 - \pi^2 / 6}{2(1-\gamma)} \right)^2 + \frac{k^2 \sigma_k^2}{k-1} \right)^{1/2}$$

whose limits are  $X_{1,2} = A \mp (1+A^2)^{1/2}$  with  $A = \frac{1 - (1-\gamma)^2 - \pi^2 / 6}{2(1-\gamma)} = -.9741$ , as  $k \rightarrow \infty$ .

The efficiencies of these values are given in Table 2. We find that method A is not always more efficient than method B. The domain  $(\chi_a, \chi_b)$  of the  $\chi$ -values for which method B has a higher efficiency is tabulated in Table 3. Since for inference on  $Y_j$  we use  $P(Y_j \leq \lambda + (X + \log n) \delta) = \exp(-\exp(-\chi))$ , these values are calculated also for the upper bound  $\chi_b$  of this domain. The  $\exp(-\exp(-\chi_b))$ -values decrease to .8 as  $k \rightarrow \infty$ ; thus

for the estimation of the  $p$ -quantiles of the annual maximum, the method A is better than method B for  $p > .9$  (Table 3 for  $k > 7$ ), which is contrary to what is thought usually. Since these values are based on the above mentioned approximation of  $\Sigma_A$  for large  $k$ , we think that the values are sufficiently adequate for  $k > 15$ , showing the definite superiority of method A, if both methods can be applied.

These results are based on the assumption that  $X_{ij}$  has the distribution  $\Lambda$ . We think that our results hold even with the assumption that the distribution  $F$  of  $X_{ij}$  belongs to the domain of attraction of  $\Lambda$  with a sufficiently fast rate of convergence of  $F^n$  to  $\Lambda$  (as in Section 2). This will be true for  $F$  which is rather close to  $\Lambda$  without defining the appropriate closeness. To discuss the robustness properties analytically is more difficult and is not the aim of this paper.

One might think that method B can always be applied for small  $k$ 's. But the assumption of the independence of the  $X_{ij}$ 's is rarely satisfied in applications. The crucial point for method B is mainly the local dependence of the  $X_{ij}$ 's which influences the asymptotic behaviour of the  $m_i$ 's. Often the largest values occur in clusters which prevents the application of Weissman's method; but method A can still be applied.

Finally we mention that our results are rather useful for the design of experimentation. They are important for cases where only a small amount of information is available, where only  $k$  instead of  $nk$  data are recorded, i.e. where only the extreme values are observed following the design of Gumbel or of Weissman.

Figure 1: Efficiency of quantiles: Eff (X+log n), k=15.

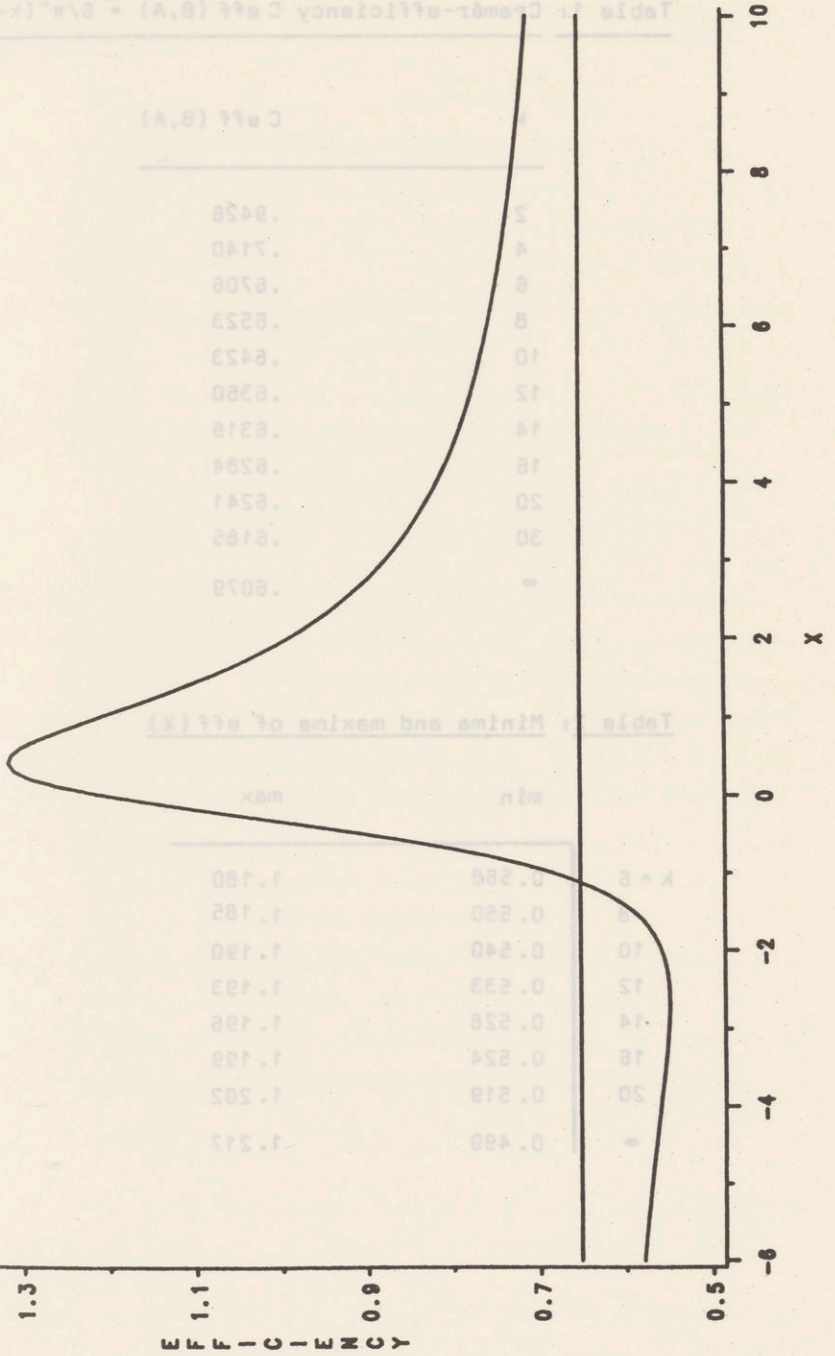


Table 1: Cramér-efficiency  $C_{\text{eff}}(B,A) = \frac{6}{\pi^2(k-1)\sigma_k^2}$

k	$C_{\text{eff}}(B,A)$
2	.9426
4	.7140
6	.6706
8	.6523
10	.6423
12	.6360
14	.6316
16	.6284
20	.6241
30	.6185
$\infty$	.6079

Table 2: Minima and maxima of  $\text{eff}(X)$

	min	max
k = 6	0.568	1.180
8	0.550	1.185
10	0.540	1.190
12	0.533	1.193
14	0.528	1.196
16	0.524	1.199
20	0.519	1.202
$\infty$	0.499	1.217

Table 3: Domain of  $X$  with  $\text{eff}(X) \geq 1 : [X_a + \log n, X_b + \log n]$ 

$k$	$X_a$	$X_b$	$\exp(-\exp(-X_b))$
6	-.0396	2.320	0.906
8	-.0813	2.006	0.874
10	-.1047	1.865	0.857
12	-.1195	1.784	0.846
14	-.1298	1.733	0.838
16	-.1373	1.697	0.833
20	-.1475	1.650	0.825
$\infty$	-.1852	1.496	0.799

Table 4: Eff  $(X_p + \log n)$  for some  $p$ -values

$k$	$X_p = -\log(-\log p)$		
	$p = 0.9$	0.95	0.99
6	1.0068	0.9477	0.8726
8	0.9723	0.9106	0.8344
10	0.9527	0.8896	0.8130
12	0.9400	0.8761	0.7993
14	0.9312	0.8667	0.7898
16	0.9247	0.8598	0.7828
20	0.9157	0.8503	0.7733
$\infty$	0.8813	0.8144	0.7372

References:

- Abramowitz M. and Stegun I A. Eds. (1965) Handbook of mathematical functions , Dover, New York.
- Bain L J (1978) Statistical analysis of reliability and life-testing models. Dekker, New York.
- Cunnane C. (1973) A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. J Hydrology 18: 257-271.
- Fichtenholz G M (1970) Infinite series: rudiments. Gordon and Breach, Pocket Math. Library, New York.
- Gomes M I (1985) Statistical theory of extremes - comparison of two approaches. Statistics and Decisions Suppl Issue 2: 33-38.
- Gumbel E J (1958) Statistics of extremes . Columbia Univ. Press, New York.
- Lawless J F (1982) Statistical models and methods for lifetime data . Wiley, New York.
- Rosbjerg D (1985) Estimation in partial duration series with independent and dependent peak values . J Hydrology 76, 183-195.
- Tiago de Oliveira J (1963) Decision results for the parameters of the extreme value (Gumbel) distribution based on the mean and standard deviation . Trab Estad Inv Oper 14: 61-81.
- Tiago de Oliveira J (1972) Statistics for Gumbel and Fréchet distributions . Structural Safety and Reliability, ed. H. Freudenthal, Pergamon Press Oxford and New York, pp.91-105.
- Tiago de Oliveira J (1982) A definition of estimator efficiency in k-parameter case . Ann Inst Statist Math 34 A, 411-421.
- Weissman I (1978) Estimation of parameters and large quantiles based on the k largest observations. J Amer Statist Assoc 73, 812-815.

Reçu en Octobre 1987