519.271.8

MEDEDELINGEN LANDBOUWHOGESCHOOL
WAGENINGEN • NEDERLAND • 73-18 (1973)

AN ASYMMETRIC TEST ON THE TYPE OF THE DISTRIBUTION OF EXTREMES

M. A. J. VAN MONTFORT*

Department of Mathematics, Agricultural University, Wageningen, The Netherlands

(Received 25-IX-1973)

1. INTRODUCTION

The problem may be introduced by looking at fig. 1, in which (some points of) the empirical distribution of a sample of size 47 has been plotted on Gumbel probability paper. The data are annual maxima of discharges of the North Saskatchewan River at Edmonton, Canada, see table 3. In the same figure the maximum likelihood fitted distribution functions have been drawn assuming that the variate itself (I) or its logarithm (II) has a double exponential distribution. Except in the right tail the fitted distribution functions are nearly the same, but for extrapolation purposes in the right tail (a favourite activity in e.g. operational hydrology) there is a big difference. In this case the .99-points of the fitted distribution functions differ a factor 2.

Roughly speaking this paper deals with the problem of detecting evidence for II with respect to the commonly used I.

In this study only largest extremes are considered. All comments made on largest extremes are applicable to smallest extremes after well-known modifications.

Only two types of limiting distribution functions of extremes for initial distributions with $+\infty$ as upperbounds of their carriers exist (see e.g. KENDALL and STUART, vol. I, chapter 14), namely $\Phi_{\mathrm{I}}\left(\frac{x-\mu}{\sigma}\right)$ and $\Phi_{\mathrm{II}}\left(\frac{x-\mu}{\sigma}\right)$ (with $-\infty < \mu < +\infty$, $0 < \sigma < +\infty$) defined as

*Agricultural University, c.o. De Dreijen 8, Wageningen, the Netherlands. This investigation was performed during a visit at Princeton University on a NATO grant supplied by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

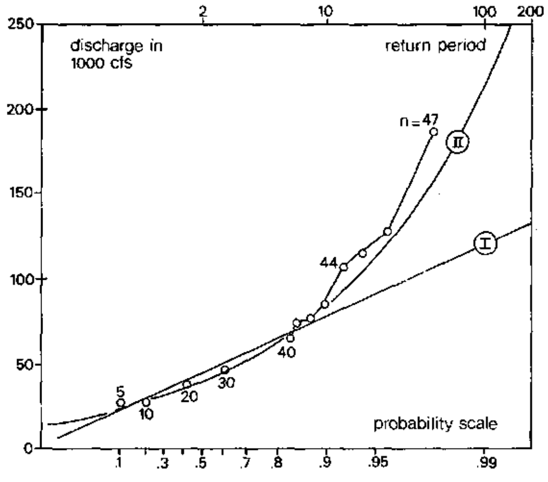

FIG. 1. A probability plot of annual maxima of discharges of a river with fitted distribution functions of two different types (see text).

$$\Phi_I(y) = \exp(-e^{-y}) \quad , -\infty < y < +\infty \tag{1}$$

$$\Phi_{II}(y) = \begin{cases} 0 & , -\infty < y \leq 0 \\ \exp(-y^{-\alpha}) & , 0 < y < \infty, 0 < \alpha < \infty \end{cases} \tag{2}$$

The parameter α is the shape parameter of Φ_{II}; μ is a location parameter and σ a scale parameter.

The limiting value Q of the critical quotient $Q(x)$, defined on the initial distribution (with density f and distribution function F) as

$$Q = \lim_{x \to \infty} Q(x) = \lim_{x \to \infty} \frac{-f^2(x)}{f'(x)\{1-F(x)\}} \tag{3}$$

distinguishes between the two types: if $Q = 1$ then Φ_I; if $Q < 1$ then Φ_{II}. The shape parameter of Φ_{II} can be found from

$$\lim_{x \to \infty} \frac{1-F(x)}{1-F(kx)} = k^{\alpha} \text{ for } k>0 \tag{4}$$

Some relations between Φ_I and Φ_{II} will be established.

(i) If $Y = \log(X-\mu)$ and

$$\Pr(X<x) = \exp\left\{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}\right\}$$

unless $x \leqslant \mu$, where $\Pr(X<x) = 0$,

then

$$\Pr(Y<y) = \exp\{-e^{-\alpha(y-\log\sigma)}\} \tag{5}$$

So, for known μ, statistical problems concerning Φ_{II} can be translated into problems concerning Φ_I using a logarithmic transformation. Whether or not such a transformation including μ is needed is the question.
(ii) Φ_I can be considered as a limiting distribution of Φ_{II}. When X is the prototype of Φ_{II}, then ($\alpha > 0$):

$$\Pr\{\alpha(X-1)<x\} = \begin{cases} 0 & , -\infty < x \leq -\alpha \\ \exp\left\{-\left(1+\frac{x}{\alpha}\right)^{-\alpha}\right\}, & -\alpha < x < +\infty \end{cases} \tag{6}$$

and

$$\lim_{\alpha\to\infty} \Pr\{\alpha(X-1)<x\} = \exp(-e^{-x}),\ -\infty < x < +\infty \tag{7}$$

A variant of (ii) is the following: $\Phi_{II}((x-\mu)/\sigma)$ can be translated and rescaled so that $\Phi_{II}((0-\mu)/\sigma) = e^{-1}$ and $\Phi_{II}((z-\mu)/\sigma) = p > e^{-1}$, $z = -\log(-\log p)$. Then Φ_{II} and the standard Φ_I coincide in two points,

$$\sigma = -\mu = z/\{\exp(z/\alpha)-1\} \tag{8}$$

and

$$\lim_{\alpha\to\infty} \Phi_{II}\left(\frac{x-\mu}{\sigma}\right) = \lim_{\alpha\to\infty} \exp\left[-\left\{1+\frac{x}{\alpha}\left(1+\frac{z}{2!\alpha}+\frac{z^2}{3!\alpha^2}+\ldots\right)\right\}^{-\alpha}\right]$$

$$= \exp(-e^{-x}) = \Phi_I(x),\ -\infty < x < +\infty \tag{9}$$

Note that $\Phi_{II}((x-\mu)/\sigma) < \Phi_I(x)$ for $x > z$; thus especially for extrapolation in the right tail it is highly important to use the correct type.

The difference in tail length in the right tail between Φ_I and Φ_{II} can be expressed as a function of the expectations of the quantities $X_{(n)} - X_{(n-1)}$ and $X_{(n-1)} - X_{(n-2)}$, where $X_{(i)}$ is the i th order statistic, $i = 1,\ldots, n$. An origin and scale invariant measure of tail length may be introduced:

$$\rho = \lim_{n\to\infty} \frac{\mathscr{E}(X_{(n)}-X_{(n-1)})}{\mathscr{E}(X_{(n-1)}-X_{(n-2)})} \tag{10}$$

This results in $\rho(\Phi_I) = 2$ and $\rho(\Phi_{II}) = 2/(1-\alpha^{-1}), \alpha > 1$. In this sense Φ_{II} has a longer tail than Φ_I.

Because of the numerical problems involved in using the likelihood ratio test, derived in appendix 1, and in calculating (expectations and) covariance matrices of ordered samples, a statistic W (see 2.3) will be proposed which does not need the covariance matrix and still has a reasonable power when compared with a statistic W_1 (see 2.2) which does need the expectations and the covariance matrix. The notation W is used while W is analogous to the Shapiro Wilk test statistic for normality, see SHAPIRO and WILK (1965); both can be interpreted as measures for straightness of the probability plot of the empirical distribution function on appriate probability paper.

2. Description of the test statistics used

2.1 *Notation*

Let X_t be the vector of ordered observations $X_{(1)} \leqslant \ldots \leqslant X_{(n)}$ corresponding to an ordered random sample from $X = \mu + \sigma Y$, where

$$Y \sim \begin{cases} \Phi_{\mathrm{I}}(y),\ t = 0 \\ \Phi_{\mathrm{II}}(1+ty),\ t>0 \text{ with } t = \alpha^{-1} \end{cases}$$

The parameters t, μ, σ are unknown. The null hypothesis H_0 is $t = 0$ whilst the alternative is $t>0$

Let

1: vector with all the n elements equal to 1,

I: the identity matrix of order n,

$$\eta_t = \mathscr{E} Y_t = \mathscr{E}\left(\frac{X_t - \mu\mathbf{1}}{\sigma}\right)$$

$$\xi = \left(\frac{\mathrm{d}}{\mathrm{dt}}\eta_t\right)_{t=0}$$

$e_t = Y_t - \eta_t$, the vector of residuals with zero means and covariance matrix V_t,

E: the lower triangular matrix such that $E'E = V_0^{-1}$

The statistics described below can easily be modified to be used in censored samples, i.e. when one or more ordered values are missing. Missing values in the right tail will reduce the power considerably.

2.2 *The W_1 statistic*

For arbitrary μ and σ we have

$$X_t = \mu\mathbf{1} + \sigma(\eta_t + e_t) \tag{11}$$

In order to avoid the necessity of having different statistics for different values of t we will seek a preferential direction within the space of residuals with t small.

Therefore, we approximate η_t, which is continuously differentiable at $t = 0$, by $\eta_t \approx \eta_0 + t\xi$, dropping the terms of higher order in t. After some algebra, based on Lieblein (1953, 1955), the ith element of ξ, called ξ_i, turns out to be

$$\xi_i = \frac{\pi^2/6-\gamma^2}{2} + \gamma\eta_i + i\binom{n}{i}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}\frac{\log^2(i+j)}{2(i+j)} \tag{12}$$

where η_i is the ith element of η_0, and $\sum_{i=1}^{n} \xi_i = n(\pi^2/6+\gamma^2)/2$ and γ is Euler's constant.

The algebra to be done needs the identity

$$S_{n,i} = i\binom{n}{i}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}/(i+j) = 1, \ (i=1, ..., n) \tag{13}$$

Note that $S_{n,n} = 1$ and $S_{n,i} = \frac{n}{n-i} S_{n-1,i} - \frac{i}{n-i} S_{n,i+1}$

and the proof can be completed by means of induction.

Thus

$$EX_t \approx \mu E\mathbf{1} + \sigma E\eta_0 + \sigma t E\xi + \sigma E e_t \tag{14}$$

Note that $\mathrm{cov}(Ee_0) = I$ and that $\sigma t > 0$ for $t > 0$.

As a reference for W, to be introduced in 2.3, a statistic W_1 is chosen. W_1 is analogous to the standard Student t-test of the nulhypothesis that σt (i.e. the regression coefficient of $E\xi$) equals zero, with a rightsided critical region; W_1 behaves on H_0 hopefully like t_{n-3}. It can easily be verified that W_1 is origin and scale invariant by replacing x by $\gamma x+\delta, \gamma>0$ which does not influence W_1. Note that W_1 needs V, η, and ξ of which V (and $E'E = V^{-1}$) needs much calculation time. This is the reason for introducing the statistic W, which needs only computationally simple approximations of η, ξ and V.

2.3. *The W statistic*

Using origin invariant tests, all observations can be reduced by subtracting $x_{(1)}$; then the first element of EX_t is zero and thus e_{11} of E is irrelevant. A striking fact is that e_{ij} turns out to be relatively small for $j \leqslant i$-2 ($i > 2$) and that $e_{ii} \approx -e_{i,\,i-1}$ for $i = 2, ..., n$. Therefore the remaining elements of EX_t can be approximated very well by a multiple of the gaps, where the i-th gap g_i is defined as $X_{(i+1)}-X_{(i)}$. Gaps are rescaled to leaps (a term due to J. W. Tukey) by dividing them by the corresponding expectations of the gaps in an ordered sample from $Y \sim \Phi_1(y)$. On H_0 the leaps turn out (i) to be nearly uncorrelated, (ii) to have expectations equal to σ, (iii) to have nearly the same variance (nearly σ^2), and (iv) their distributions are nearly exponential, see PYKE (1965). (It would be interesting to consider square root or logarithmic transformations of the leaps, both of which have more symmetric distributions than the untransformed leaps; the square root transformation is preferable when data have rounding errors in which case gaps can have the value zero). The vector of leaps ℓ_t can be obtained as

$$\ell_t = AX_t \tag{15}$$

where A is a $(n-1) \times n$ matrix with

$$a_{i,\, i+1} = (\eta_{i+1} - \eta_i)^{-1} = -a_{ii},\ i = 1, ..., (n-1)$$

$$a_{ij} = 0 \text{ elsewhere} \tag{16}$$

and η_i is the i-th element of η_0. Then $A\mathbf{1} = 0$ and $A\eta_0 = 1$
So

$$\ell_t = AX_t = \sigma A\eta_t + \sigma Ae_t \approx \sigma\mathbf{1} + \sigma tA\xi + \sigma Ae_t \tag{17}$$

Note that $\mathrm{cov}(Ae_0) \approx \sigma^2 \mathrm{I}$. Then the regression coefficient of $A\xi$ could be estimated using the *non*-generalized least squares method without a great loss in efficiency for small t. Now a close approximation of ξ, and at the same time for A, is derived using the general approximation of $\mathscr{E}X_{(i)}$ by $\mu + \sigma\Phi^{-1}(i/(n+1))$.

Defining

$$\psi_i = -\log\left(-\log\left(\frac{i}{n+1}\right)\right) \tag{18}$$

the result of this approximation for the prototypes I and II on the basis of formula (6) is

$$\mathscr{E}_0\,(X_{(i)}) \approx \psi_i \tag{19.1}$$

$$\mathscr{E}_t(X_{(i)}) \approx (\exp\,(t\psi_i) - 1)/t \rightarrow \psi_i \text{ when } t \rightarrow 0 \tag{19.2}$$

It is in general true that

$$\lim_{t=0} \mathscr{E}_t X_{(i)} = \mathscr{E}_0 X_{(i)} \tag{20}$$

An approximation of $\mathscr{E}_t X_{(i)}$ by a series expansion of two terms is

$$\mathscr{E}_t X_{(i)} \approx \mathscr{E}_0 X_{(i)} + t\left(\frac{d}{dt}\,\mathscr{E}_t X_{(i)}\right)_{t=0} = \psi_i + t\psi_i^2/2 \tag{21}$$

For the expectation of the i-th element $(i = 1,..., n-1)$ of the vector of leaps, we can write

$$\mathscr{E}_0 \ell_i = \sigma \tag{22.1}$$

$$\mathscr{E}_t \ell_i \approx \frac{\mu + \sigma(\psi_{i+1} + t\psi_{i+1}^2/2) - \mu - \sigma(\psi_i + t\psi_i^2/2)}{\psi_{i+1} - \psi_i} =$$

$$= \sigma\left(1 + \frac{t}{2}\,(\psi_i + \psi_{i+1})\right)$$

$$\approx \sigma\,(1+t\psi_{i+\frac{1}{2}}) = \sigma\,\left(1-t\log\left(-\log\left(\frac{i+\frac{1}{2}}{n+1}\right)\right)\right) \tag{22.2}$$

Thus the expectations of the leaps are constant under H_0 and approximately linearly increasing with $-\log(-\log((i+\frac{1}{2})/(n+1)))$ under the alternative hypothesis. As a result of the fact that the leaps are nearly uncorrelated and have nearly the same variance under H_0, we may propose the new statistic W, which is a function of the correlation coefficient r between ℓ_i and $\psi_{i+\frac{1}{2}}$ ($i = 1, 2,..., n-1$):

$$W = \tfrac{1}{2}\log\frac{1+r}{1-r} \tag{23}$$

which looks more normal than r. As a result of the behavior under H_t the origin and scale invariant statistic W will have a rightsided critical region.

3. Results and an example

Because of the algebraic problems involved in evaluating the distributions of W and W_1 the critical values and the power have been approximated by simulation. The tables 1 and 2 give the results of simulation for samples of size n of the statistics W_1 and W for $t = 0$ (corresponding to the distribution under H_0) and some values of $t > 0$ (corresponding to the distribution of extremes when sampling largest values from e.g. Student's t_ν distribution with $\nu = 1/t$). The critical values (Table 1) and the power (Table 2) have been estimated from the empirical distribution functions based on $N = 1999$ n-tuples. The p-points have been estimated as the $p(N+1)$-th observations of the ordered samples of N realizations of the statistics. A 95% confidence interval for the true size of the tests can be approximated by $p \pm 1.96\sqrt{p(1-p)/N}$, which interval is probably short enough for practical applications. For each n the empirical distribution functions are the results of the same sequence of 'random' numbers.

TABLE 1. Critical Values at Level .05 (.05) .15, .50 for Sample Size n.

statistic	W_1				W			
critical level	0.05	0.10	0.15	0.50	0.05	0.10	0.15	0.50
$n =$ 5	2.93	2.02	1.47	−0.08	1.49	1.16	0.97	+0.03
10	2.08	1.44	1.14	−0.09	0.77	0.58	0.49	+0.03
15	1.78	1.30	1.01	−0.09	0.55	0.42	0.35	+0.02
20	1.81	1.32	1.04	−0.08	0.47	0.37	0.30	+0.03
25	1.72	1.35	1.00	−0.07	0.40	0.33	0.26	+0.03
50					0.27	0.21	0.17	+0.02
100					0.18	0.15	0.12	+0.02

TABLE 2. Power of the Statistics (in 0/00)

size of test		0.05		0.10		0.15	
n	t	W_1	W	W_1	W	W_1	W
5	.1	070	065	135	125	185	190
	.2	085	075	160	160	230	230
	1/3	110	095	210	195	295	295
	.5	135	110	275	250	360	365
10	.1	085	080	175	170	230	230
	.2	150	140	255	255	330	330
	1/3	250	225	380	370	460	455
	.5	375	350	520	570	590	585
15	.1	120	120	200	205	280	270
	.2	220	215	335	335	430	420
	1/3	375	370	520	570	590	580
	.5	555	540	695	675	760	740
20	.1	145	140	230	220	305	300
	.2	260	265	390	380	495	475
	1/3	470	465	590	580	670	655
	.5	670	660	785	760	840	820
25	.1	165	170	235	235	335	320
	.2	325	320	430	415	545	525
	1/3	570	560	675	650	750	740
	.5	785	760	855	840	905	890
50	.1		235		370		470
	.2		530		670		745
100	.1		390		520		600
	.2		785		865		910

W turns out to be nearly as powerfull as W_1. Approximation reduces the power a bit but the amount of calculation is reduced greatly.

The expectations η_0 and the covariance matrix V_0 are available for $n = 2(1)20$, see MANN (1965); and for $n = 2(1)25$, see MANN (1968); the expectations and variances for $n = 1(1)50(5)100$ could be obtained from WHITE, see WHITE (1969). The asymptotic direction ξ for $n = 10(10)100$ can be obtained from the author. The critical values of W_1 show a reasonable accordance with those of t_{n-3}. The W-values are in close relationship with those of a normal distribution with mean and standard deviation to be fitted from the values of Table 1. Critical values at non-tabulated and not too small values of n can be obtained from linear interpolation of the given values with respect to $1/\sqrt{n}$.

The problem dealt with in this paper was suggested to the author by the pages 171–182 of STATISTICAL METHODS IN HYDROLOGY (1967). Its data, see

TABLE 3. Ordered Annual Maxima of the Discharges of the North Saskatchewan River at Edmonton (Canada) in 1000 cu. feet/sec.

19.885	20.940	21.820	23.700	24.888	25.460	25.760
26.720	27.500	28.100	28.600	30.200	30.380	31.500
32.600	32.680	34.400	35.347	35.700	38.100	39.020
39.200	40.000	40.400	40.400	42.250	44.020	44.730
44.900	46.300	50.330	51.442	57.220	58.700	58.800
61.200	61.740	65.440	65.597	66.000	74.100	75.800
84.100	106.600	109.700	121.970	185.560		

Table 3, results in $W = 0.556$, which is highly significant, see table 1. So a logarithmic transformation to type I is desirable.

A referee suggested looking at a simpler test like the Lilliefors modification of the Kolmogorov Smirnov test (LKS-test), see LILLIEFORS (1967), in which parameters are replaced by their estimated values. Three estimators are used: the estimator by the method of moments (ME), the best linear invariant estimator (BLIE, see MANN (1967)) and the maximum likelihood estimator (MLE). At sample size $n = 10$ and at a size of the test of 0.10 the power (simulated by 1000 samples) of the LKS-tests at $t = 0.1$ and $t = 0.5$ turned out to be

power	ME	BLIE	MLE	W_1
$t = 0.1$	.135	.135	.145	.175
0.5	.455	.500	.505	.520

In order to understand the first three columns it is noticed that ME applied to this skew distribution has a low efficiency compared with BLIE and MLE. The power of W_1 at $t = 0.5$ is hardly higher than the LKS-tests using BLIE and MLE; this is not so unexpected while the alternative ($t = 0.5$) deviates largely from the null hypothesis ($t = 0$) and W_1 was constructed in order to be powerfull for *small* values of t. The power at $t = 0.1$ confirms this comment. A disadvantage of the LKS-tests in this study is the lack of an alternative.

4. SOME GENERAL REMARKS ON USING DISTRIBUTIONS OF EXTREMES

(i) When $Q(X) = 1 - \varepsilon(x)$ for large x with $\varepsilon(x) > 0$ and $\lim \varepsilon(x) = 0$ then the distribution of extremes will behave like Φ_{II} when N (the number of samples of which the maximum is taken) is too small for using the limiting distribution Φ_I.

(ii) In practical applications it is not known whether X or $\log X$ is expected to have a type I distribution. This can be considered as a restricted choice from e.g. the following class of order preserving transformations (cfr BOX and COX, 1965):

$$x^{(\lambda)} = \begin{cases} (x^{\lambda}-1)/\lambda, & \lambda \neq 0 \\ \log x, & \lambda = 0 \end{cases}$$

where X corresponds with $\lambda = 1$ and $\log X$ with $\lambda = 0$.

(iii) Another reasonable class of order preserving transformations containing X and $\log(X-\mu)$ except for a linear transformation is

$$x^{(\nu)} = \begin{cases} \log(1+\nu x)/\nu, & 1+\nu x>0,\ \nu \neq 0 \\ -\infty, & 1+\nu x \leq 0,\ \nu \neq 0 \\ x, & \nu = 0 \end{cases}$$

where X corresponds to $\nu = 0$ and $\log(X-\mu)$ to $\nu \neq 0$.

Note that μ is the lower bound of the carrier of the distribution.

The distributions of $X^{(\lambda)}$ and $X^{(\nu)}$ contain three parameters. In (ii) and (iii) we have looked for transformations which can result in a critical quotient equal to 1 or at least give a better fit to data.

Using the calculation technique of Appendix 1 and the data of Table 3 results in Table 4 and gives rise to the following conclusions and comments.

TABLE 4. Numerical Results of Transforming the Extremes of Table 3.

Model	parameters fixed	optimal values			L_{max}	$\hat{\Phi}^{-1}(0.99)$
Φ_I	–	$\mu = 38.15$	$\sigma = 17.74$		−213.55	119.75
Φ_{II}	$\theta_1 = 0$		$\theta_3 = 2.524$	$\log\theta_2 = 3.549$	−208.08	215.23
	−75		7.133	4.717	−209.98	138.03
	+15		1.360	2.870	−210.05	535.00
	–	$\theta_1 \approx .76$	2.471	3.525	−208.07	215.23
$x^{(\lambda)}$	$\lambda = -0.5$	$\mu = 1.654$	$\sigma^{-1} = 15.479$		−209.83	1672.1
	0	3.549	2.524		−208.08	215.23
	+0.6	12.806	0.266		−209.79	135.97
	–	4.272	1.745	$\lambda = +0.1$	−208.07	191.05

(i) The likelihood ratio test, which needs extensive calculations, results in $2(213.55–208.07) = 10.96$, $\Pr(\chi_1^2 > 10.96) \approx 0.001$ which relates reasonably to the result of the previous test. So not X but perhaps $\log(X-\mu)$ has a Φ_I distribution.

(ii) An estimate of μ is then 0.76, which can be rounded off to 0, and a 95% confidence interval is approximately $-15 < \mu < +75$.

(iii) An estimate of the transformation parameter is $\lambda = +0.1$ and a 95% confidence interval is approximately $-.5 < \lambda < +.6$.

(iv) An often used parameter in hydrology is Φ^{-1} (0.99). The estimates under the λ-and ν-transformations are nearly the same, but the 95% confidence intervals are (138, 535) and (135, 1672) respectively.

(v) To estimate the lower bound of the carrier of the distribution (result +0.76) when the real lower bound is known (in this case 0) is realistic when the right tail is of special interest and the validity of the assumptions in the left tail is debatable.

Acknowledgment

The author thanks D. F. Andrews and J. W. Tukey of the Department of Statistics of Princeton University for helpful discussions.

Summary

The type of the distribution of extremes depends on the behaviour of the initial distribution in the tails which is often not known completely; then only a statistical conjecture is available based on fitting the initial distribution mainly outside of the tails. This conjecture often results in using the type I distribution of extremes of largest values whether for untransformed or for log-transformed data. This study deals with the situation of unlimited largest extremes where type I is suspected and type II with an unknown but not 'too small' shape parameter is the competitor. An arithmetically simple but still reasonably powerful origin and scale invariant tests based on complete samples is proposed. The configuration statistic used is a function of differences of adjacent order statistics. Finally the problem is considered how to transform data so that a type I distribution can be used.

References

1. Box, G. E. P. and D. R. Cox, An analysis of transformations, Journal of the Royal Statistical Society, series B, **26** (1964) 211–252.
2. Kendall, M. G. and A. Stuart, The advanced theory of statistics, vol. I, C. Griffin and Company Ltd., London (1969).
3. Lieblein, J., On the exact evaluation of the variances and covariances of order statistics in samples of the extreme value distribution. Ann. Math. Stat. **24** (1953) 282–287.
4. Lieblein, J., On moments of order statistics from the Weibull distribution, Ann. Math. Stat. **26** (1955) 330–333.
5. Lilliefors, H. W., On the Kolmogorov-Smirnov test for normality with mean and variance unknown. J. Amer. Stat. Ass. **62** (1967) 399–402.
6. Mann, N. R., Point and interval estimates for reliability parameters when failure times have the two parameter Weibull distribution. Unpublished thesis, University of California at Los Angeles (1965).
7. Mann, N. R., Results on location and scale parameter estimation with application to the extreme value distribution. ARL 67–0023, Aerospace Research Laboratories, Ohio (1967).

8. MANN, N. R., Results on statistical estimation and hypothesis testing with application to the Weibull and extreme value distribution, ARL 68–0068, Aerospace Research Laboratories, Ohio (1968).
9. PYKE, R., Spacings. J. Roy. Statist. Soc. B, **27** (1965) 421–427.
10. SHAPIRO S. S. and M. B. WILK, Analysis of variance test for normality (complete samples). Biometrika **52** (1965) 591–611.
11. STATISTICAL METHODS IN HYDROLOGY, Queen's Printer, Ottawa, Canada (1967).
12. WHITE, J. S., The moments of log Weibull order statistics. Technometrics **11** (1969) 373–386.

Appendix 1 The likelihood ratio test

(i) The general distribution is

$$F(x) = \begin{cases} \exp\left\{-\left(1+\dfrac{x-\mu}{\alpha\sigma}\right)^{-\alpha}\right\}, & x>\mu-\alpha\sigma \\ 0 & \text{elsewhere} \end{cases}$$

The restricted distribution is

$$\lim_{\alpha\to\infty} F(x) = \exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right), \quad -\infty<x<+\infty$$

Reparameterization of the general distribution

$$\begin{aligned} \theta_1 &= \mu - \alpha\sigma & \mu &= \theta_1 + \theta_2 \\ \theta_2 &= \alpha\sigma & \sigma &= \theta_2/\theta_3 \\ \theta_3 &= \alpha & \alpha &= \theta_3 \end{aligned}$$

leads to the fact, that θ_3 $\{\log (X-\theta_1)-\log(\theta_2)\}$ has a standard Φ_1-distribution, and results in

$$F(x) = \begin{cases} \exp\left\{-\left(\dfrac{x-\theta_1}{\theta_2}\right)^{-\theta_3}\right\}, & x>\theta_1 \\ 0 & \text{elsewhere} \end{cases}$$

$$f(x) = F'(x) = \begin{cases} \theta_3/\theta_2 \cdot \left(\dfrac{x-\theta_1}{\theta_2}\right)^{-\theta_3-1} \exp\left\{-\left(\dfrac{x-\theta_1}{\theta_2}\right)^{-\theta_3}\right\}, & x>\theta_1 \\ 0 \quad \text{elsewhere} \end{cases}$$

$$L = \sum_{i=1}^{n} \log f(x_i) = n(\log\theta_3 - \log\theta_2) - (\theta_3+1)\sum \log t_i - \sum t_i^{-\theta_3}$$

with $t_i = (x_i-\theta_1)/\theta_2$ and L is the log (likelihood).

Let

$$L_i = \frac{dL}{d\theta_i}, \quad L_{ij} = \frac{d^2L}{d\theta_i\, d\theta_j}$$

then (dropping the indices of t)

$$\begin{aligned} L_1 &= \theta_2^{-1}(\theta_3+1)\sum t^{-1} - \theta_2^{-1}\theta_3\sum t^{-\theta_3-1} \\ L_2 &= (n-\sum t^{-\theta_3})\,\theta_2^{-1}\,\theta_3 \\ L_3 &= n/\theta_3 - \sum \log(t) + \sum t^{-\theta_3}\log(t) \end{aligned}$$

$$-L_{11} = \{\theta_3 \sum t^{-\theta_3-2} - \sum t^{-2}\}\,(1+\theta_3)/\theta_2^2$$

$$-\mathscr{E}L_{11} = n\,(1+\theta_3)^2\,\Gamma\,(1+2/\theta_3)/\theta_2^2$$

$$-L_{12} = \theta_3^2 \sum t^{-\theta_3-1}/\theta_2^2$$

$$-\mathscr{E}L_{12} = n\,\theta_3^2\,\Gamma\,(2+1/\theta_3)/\theta_2^2$$

$$-L_{13} = \{-\sum t^{-1}+\sum t^{-\theta_3-1}-\theta_3 \sum t^{-\theta_3-1}\log(t)\}/\theta_2$$

$$-\mathscr{E}L_{13} = n\,\{\Gamma\,(1+1/\theta_3)/\theta_3 + \Gamma'\,(2+1/\theta_3)\}/\theta_2$$

$$-L_{22} = \{n+(\theta_3-1)\sum t^{-\theta_3}\}\,\theta_3/\theta_2^2$$

$$-\mathscr{E}L_{22} = n\,\theta_3^2/\theta_2^2$$

$$-L_{23} = \{-n+\sum t^{-\theta_3}-\theta_3 \sum t^{-\theta_3}\log(t)\}/\theta_2$$

$$-\mathscr{E}L_{23} = n\,(1-\gamma)/\theta_2$$

$$-L_{33} = n/\theta_3^2 + \sum t^{-\theta_3}\log^2(t)$$

$$-\mathscr{E}L_{33} = n\,\{\pi^2/2+(1-\gamma)^2\}/\theta_3^2$$

An outline of the iterative estimation procedure of $\{\theta\} = (\theta_1, \theta_2, \theta_3)$ is

$$\{\theta\}_{j+1} = \{\theta\}_j + \{-L_{ij}\}^{-1}\,\{L_i\}$$

where $\{\theta\}_0$ is a vector of suitably chosen starting values and $\{\theta\}_i$ is the result of the i-th iteration; L_{ij} in the above equation can be replaced by $\mathscr{E}L_{ij}$. Replacing $\{\theta\}$ by $\{\theta\}_\infty$ in the log (likelihood) gives a part of the likelihood ratio test statistic, denoted as L_G.

(ii) Replacing σ by β^{-1} in the restricted distribution results in the log likelihood

$$L = n \log \beta - \beta \sum x_i + n\beta\mu - e^{\beta\mu} \sum e^{-\beta x}$$

Equating first derivatives to zero results in

$$\varphi(\beta) \equiv \beta^{-1} - (\sum x_i)/n + S_1/S_0 = 0$$

$$\mu = \log(n/S_0)/\beta$$

where

$$S_k = \sum x_i^k e^{-\beta x_i},\ k = 0, 1, 2$$

Note

$$\varphi'(\beta) = -\beta^{-2} + (S_1^2 - S_0 S_2)/S_0^2$$

Using the Newton-Raphson technique for solving $\varphi(\beta) = 0$ will give fast convergence and the restricted maximum of the log likelihood L_R can be calculated. Using the fact that $2(L_G - L_R)$ behaves asymptotically as χ_1^2, the test can be completed. Note that a shift does not influence L_G and L_R, and that a change in scale does not influence $L_G - L_R$, so $2(L_G - L_R)$ only depends on α and n.