Goodness-of-fit tests in extreme value applications

András Zempléni

Department of Probability Theory and Statistics
Eötvös Loránd University, Budapest

May 12, 2004

1 Introduction

Extreme value theory has been one of the most quickly developing area of mathematical statistics in the last decades. The methods, based on the Fisher-Tippett theorem about the possible limits of normalized maxima of iid random variables is now routinely applied in very different branches of areas from financial mathematics to environmetrics, from reliability to internet traffic.

This theorem states that the possible limit distributions of

$$M_n = \max\{(X_i - a_n)/b_n : i = 1, \dots n\} \tag{1}$$

and the the so-called generalized extreme value (GEV) distribution functions (cdf) coincide. This class of the possible (nondegenerate) limit distributions of (1) can be given by the Jenkinson-von Mises representation (2)

$$H_\xi(x) = \begin{cases} \exp\{-(1+\xi x)^{-1/\xi}\} & : \text{ if } \xi \neq 0 \\ \exp\{-\exp\{-x\}\} & : \text{ if } \xi = 0 \end{cases} \tag{2}$$

for $1 + \xi x > 0$, where ξ is the so-called shape parameter. In practice location- and scale parameters (denoted by μ and σ, respectively) are also included in the model. For a review of both the theory and methods see Embrecths et al [4].

As the conditions which ensure the convergence to a GEV distribution are rather mild for absolutely continuous, iid observations, it is common practice to assume the asymptotics to hold and estimate the parameters of the limiting distribution H_ξ in (2) if one has practically a data set like block maxima of random variables (similar to (1)). In most cases some graphical methods, like Q-Q plots, are used for model validation. These are rather subjective, formal tests based on the plot are rarely applied. The typical goodness-of-fit procedures, like Kolmogorov-Smirnov statistics do not have distribution-free limits in case of estimated parameters and are not worked out for this case. We give simulated critical values for the case under consideration, but the method itself is not a strong one.

However, if one intends to be more cautious with his approach, may indeed wish to test if there is an asymptotic distribution of the normalized maxima (which is necessarily of the form (2)), i.e. is the MDA condition satistifed. This problem is tackled in a paper by Dietrich et al [3].

In Section 2 we summarize the available tests for checking if the distribution of the sample is indeed that of a GEV. A specific test, developed exactly for this problem was originally presented in Zempléni [10], see also Kotz and Nadarajah [6]. Now we propose also an alternative version to it as well as a modified Anderson-Darling test, adapted to the problems, where the emphasis is on the fit for one of the tails of the distribution.

We also compare the performances of the above tests to the classical Anderson-Darling, Kolmogorov-Smirnov and Cramér-von Mises tests, where the parameter estimation has also to be taken into account.

Another important approach to extreme-value analysis is, where the values higher than a given threshold are applied for inference. In this case the so-called generalized Pareto distribution (GPD) is the possible limit. This case is considered in Section 3.

In several financial applications, the risk is traditionally measured by a high quantile, the so-called Value-at-Risk (VaR). Backtesting the estimators of this value is a common practice. However, the VaR does not give information about the actual loss which can occur if the VaR is exceeded. This amount is called the expected shortfall, which can also be estimated (using the fact that it is related to the GPD). We deal with its tests in Section 4.

2 Generalized Extreme Value distributions

2.1 Specific tests

The Anderson-Darling test for goodness-of-fit is based on the test statistics (not denoting the dependence on the parameters)

$$A^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{F(x)(1 - F(x))} dF(x) \tag{3}$$

where F_n is the cdf of the sample and F is the cdf which is to be fitted. A^2 can be computed - based on the probability integral transform - as

$$A^2 = -n - \sum_{i=1}^{n} (2i-1)\left(\log(z_i) + \log(1 - z_{n+1-i})\right)/n \tag{4}$$

where $z_i = F(x_{(i)})$, the d.f. evaluated at the i^{th} element of the ordered sample.

On the other hand, in most of the cases only one of the tails is important (maximum for enviromental or insurance data, minimum for financial time series or in reliability), so we propose the following test

$$B^2 = n \int_{-\infty}^{\infty} \frac{(F_n(x) - F(x))^2}{(1 - F(x))} dF(x) \tag{5}$$

for the case of maximum and with $F(x)$ in the place of $(1 - F(x))$ for minimum. The advantage of (5) in comparison to (3) is that its sensitivity is concentrated to discrepancies at the relevant tail of the distribution. The statistics B^2 can also be computed easily as the following formula shows.

$$B^2 = n/2 - \sum_{i=1}^{n} (2i-1)\log(1 - z_{n+1-i})/n - \sum_{i=1}^{n} z_i/n. \tag{6}$$

The slightly more complicated form of (6) in comparison to (4) is due to the fact that A^2 is the sum of B^2 and the analogous statistics suggested for minimum, and the cross products vanish for this sum, while they are present in B^2.

Another test, proposed in [10] is based on the stability property of the GEV distributions, (de Haan, [5]): for any $m \in \mathbb{N}$ there exist a_m, b_m such that

$$F(x) = F^m(a_m x + b_m) \tag{7}$$

for all $x \in \mathbb{R}$.

The original test statistics was defined in [10] as the minimal value of

$$h(a, b) = \max_x \mid F_n(x) - F_n^2(ax + b) \mid .$$

Its motivation is (7) applied to the case $m = 2$. In order to find the optimal parameters a, b a computer-intensive method was needed. The aim of this procedure was to find an alternative method to maximum likelihood, which works well also for the cases $\xi < -1$, when in general there is no ML estimator (see Smith, [8]). The power of this test, for which the critical values are based on a distribution-free estimator of the test statistics (see also [6]), has already been presented in [10].

Here we consider an alternative, more in accordance with the other tests in this paper: we plug in the maximum likelihood estimator and calculate $h(\hat{a}, \hat{b})$ as the test statistics, since the cases $\xi < -1$ rarely occur in practice.

For all these statistics we give the simulated critical values in the following subsection. In the sequel we investigate the power of these goodness-of-fit tests for several alternatives.

Table 1: Critical values for B_n^2 in case of known parameters

Probability level/ n	100	200	400	800	1600
0.5	0.373	0.376	0.375	0.374	0.382
0.9	0.998	1.012	0.996	0.989	1.025
0.95	1.312	1.323	1.298	1.288	1.307
0.99	2.083	2.108	2.054	2.066	2.135

2.2 Limit distributions, critical values

In order to obtain critical values, one may simulate the finite sample distribution of the test statistics under the nullhypothesis, or alternatively, for large samples the asymptotic distribution may be used. To accomplish this latter task, one can use the methods, presented in Stephens, [9]. First we note that in the case of known parameters, the limit distribution of B^2 is $\int_0^1 K^2(x)dx$, where K is a mean-zero Gaussian process with covariance structure

$$cov(K(s), K(t)) = \frac{st - s \vee t}{(1-s)(1-t)}. \tag{8}$$

Similarly, for h we have the following result:

$$\lim_{n\to\infty} \sup_x \sqrt{n} \mid F_n(x) - F_n^2(a_2x + b_2) \mid = \sup_y \mid G(y) - 2\sqrt{y}G(\sqrt{y}) \mid \tag{9}$$

where G denotes the Brownian Bridge over $[0, 1]$.

(8) and (9) allow for the simulation of the critical values, but one is interested in the finite sample properties as well, so we have rather simulated the values of the statistics under the nullhypothesis for different sample sizes n. The results confirm earlier observations with respect of the statistics A^2 about its quick convergence to the asymptotic distribution. In Table 1 we present the critical values of B^2 for different sample sizes the number of repetitions in all the simulations was 50000).

For the more realistic case, when the parameters are estimated, one can apply the method in [1], which is based on calculating first the modified covariance structures of the Gaussian process in (8) and then the characteristic function of its integral. As a last step, the critical values can be got by approximating this distribution by a weighted sum of iid χ-squared distributions. As this procedure is tedious and it results in an approximation as well, we preferred simply to simulate the critical values of our statistics. These values are given in Tables 2 and 3.

In Tables 4 and 5 we present the critical values of B^2 for different sample sizes for the case when $\xi = 0.5$ and $\xi = 0.2$, and $\xi = -0.2$ and $\xi = -0.6$, respectively.

The estimation of the parameters has been done by the maximum likelihood method,

Table 2: Critical values for A_n^2 in case of estimated parameters, $\xi = 0.5$ and $\xi = 0.2$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.271	0.277	0.279	0.282	0.284	0.271	0.278	0.280	0.284	0.285
0.9	0.474	0.487	0.494	0.499	0.506	0.474	0.489	0.497	0.504	0.508
0.95	0.555	0.572	0.584	0.590	0.598	0.556	0.576	0.583	0.595	0.595
0.99	0.769	0.773	0.796	0.797	0.839	0.742	0.781	0.796	0.801	0.812

Table 3: Critical values for A_n^2 in case of estimated parameters, $\xi = -0.2$ and $\xi = -0.6$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.285	0.290	0.293	0.297	0.300	0.341	0.325	0.325	0.326	0.330
0.9	0.505	0.516	0.529	0.532	0.538	0.858	0.603	0.594	0.595	0.606
0.95	0.597	0.608	0.624	0.630	0.639	1.148	0.727	0.706	0.712	0.720
0.99	0.839	0.824	0.844	0.857	0.865	1.905	1.095	0.984	0.979	0.987

Table 4: Critical values for B_n^2 in case of estimated parameters, $\xi = 0.5$ and $\xi = 0.2$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.138	0.140	0.142	0.143	0.144	0.137	0.140	0.141	0.142	0.142
0.9	0.256	0.261	0.264	0.270	0.271	0.250	0.258	0.263	0.265	0.267
0.95	0.307	0.311	0.317	0.323	0.324	0.297	0.308	0.312	0.316	0.320
0.99	0.433	0.437	0.446	0.451	0.466	0.407	0.427	0.429	0.432	0.435

Table 5: Critical values for B_n^2 in case of estimated parameters, $\xi = -0.2$ and $\xi = -0.6$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.143	0.144	0.146	0.148	0.149	0.179	0.166	0.165	0.166	0.167
0.9	0.263	0.269	0.276	0.278	0.284	0.613	0.321	0.317	0.321	0.323
0.95	0.313	0.320	0.331	0.333	0.340	0.812	0.390	0.383	0.385	0.389
0.99	0.444	0.442	0.459	0.460	0.472					

Table 6: Critical values for h_n in case of estimated parameters, $\xi = 0.5$ and $\xi = 0.2$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.832	0.866	0.896	0.910	0.920	0.832	0.865	0.896	0.910	0.924
0.9	1.200	1.200	1.224	1.237	1.250	1.152	1.191	1.224	1.237	1.254
0.95	1.272	1.318	1.329	1.344	1.362	1.272	1.304	1.329	1.352	1.362
0.99	1.600	1.573	1.581	1.564	1.611	1.528	1.556	1.576	1.587	1.582

Table 7: Critical values for h_n in case of estimated parameters, $\xi = -0.2$ and $\xi = -0.6$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.808	0.874	0.900	0.920	0.935	0.832	0.880	0.909	0.928	0.946
0.9	1.128	1.200	1.231	1.258	1.268	1.200	1.257	1.241	1.270	1.288
0.95	1.248	1.312	1.351	1.366	1.378	1.312	1.344	1.361	1.385	1.397
0.99	1.528	1.556	1.596	1.592	1.610	1.608	1.683	1.624	1.622	1.638

known to possess the usual regularity properties for $\xi \geq -0.5$ and being superefficient for the case $-1 > \xi \geq -0.5$ (Smith, [8]).

Next we give the critical values for the h-test. In Tables 6 and 7 we present the critical values of h for different sample sizes for the case when $\xi = 0.5$ and $\xi = 0.2$, and $\xi = -0.2$ and $\xi = -0.6$, respectively.

It can be observed that the values in Tables 6 and 7 are just slightly higher than those presented in [10] in spite of the non-tailored estimation procedure applied here, since there a universal, shape-independent estimator was applied.

Finally, we give the critical values for the Kolmogorov-Smirnov test in the case of estimated parameters. In Tables 8 and 9 we present the critical values of the K-S test for different sample sizes for the case when $\xi = 0.5$ and $\xi = 0.2$, and $\xi = -0.2$ and $\xi = -0.6$, respectively.

Table 8: Critical values of the K-S test in case of estimated parameters, $\xi = 0.5$ and $\xi = 0.2$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.535	0.547	0.557	0.564	0.568	0.535	0.548	0.557	0.565	0.567
0.9	0.713	0.732	0.747	0.752	0.759	0.710	0.728	0.742	0.750	0.753
0.95	0.773	0.794	0.809	0.815	0.823	0.768	0.786	0.802	0.811	0.816
0.99	0.904	0.924	0.932	0.943	0.968	0.881	0.905	0.922	0.937	0.941

Table 9: Critical values for the K-S test in case of estimated parameters, $\xi = -0.2$ and $\xi = -0.6$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.543	0.555	0.564	0.573	0.576	0.576	0.577	0.586	0.592	0.596
0.9	0.724	0.741	0.751	0.763	0.768	0.803	0.783	0.788	0.793	0.800
0.95	0.783	0.800	0.813	0.825	0.831	0.891	0.853	0.854	0.861	0.870
0.99	0.906	0.918	0.940	0.951	0.958	1.138	1.012	0.992	0.997	1.005

Table 10: Power of the test for samples at the level of 95%

Test / n	100	200	400	100	200	400	100	200	400	50	100	200
Distr.	NB(2,0.1)			exponential			normal			uniform		
B_n	0.07	0.29	0.17	0.58	0.97	0.01	0.03	0.08	0.12	0.61	0.98	1
A_n	0.31	0.62	0.96	0.72	0.97	1	0.14	0.21	0.34	0.72	0.99	1
h_n	0.67	0.87	0.99	0.75	0.91	0.99	0.08	0.10	0.14	0.75	0.93	0.997

2.3 Power studies

We have compared the performance of the tests for some distributions, including a discrete one: the rejection probabilities for negative binomial, exponential or normal samples are given in Table 10. The critical values have been calculated by linear interpolation from those given in the tables of the previous subsection (and analogously from similar tables for the A test).

As the next step, we applied the new test to real data and their modifications. As a basis, we took Hungarian water level data from the river Tisza, collected during the years 1901-2000. The station we use here is called Vásárosnamény, it is located upstream at this river, which caused several major floodings in the last decade.

Not surprisingly, the original data is very well aproximated by the GEV distribution (see Figure 1), so it is obviously accepted by any tests. However, if we start moving some values to the observed maximum, thus step-by step increasing the deviation from the GEV distribution, then we can measure the sensitivity of the test to such departures. The results are presented in Table 11. In the first three columns we consider no new estimates (i.e. the original distribution, without re-estimating the parameters is considered). This corresponds to a hypothetical question: when can we detect a change in the underlying distribution? The next three columns give the results for the case, when the parameters are always re-estimated for the samples. The results confirm our conjecture: the modified test B_n^2 is quicker in detecting the changes in the upper tail. Similar results can be observed for the case when the largest value is increased by 10cm each step (in this case the re-estimation is obligatory, since the increased values soon move out from the region, where the density is positive). Figure 2 gives the Q-Q plot of GEV-fit for this modified data, where the GEV-hypothesis was just rejected at the 95%-level by the B-test.

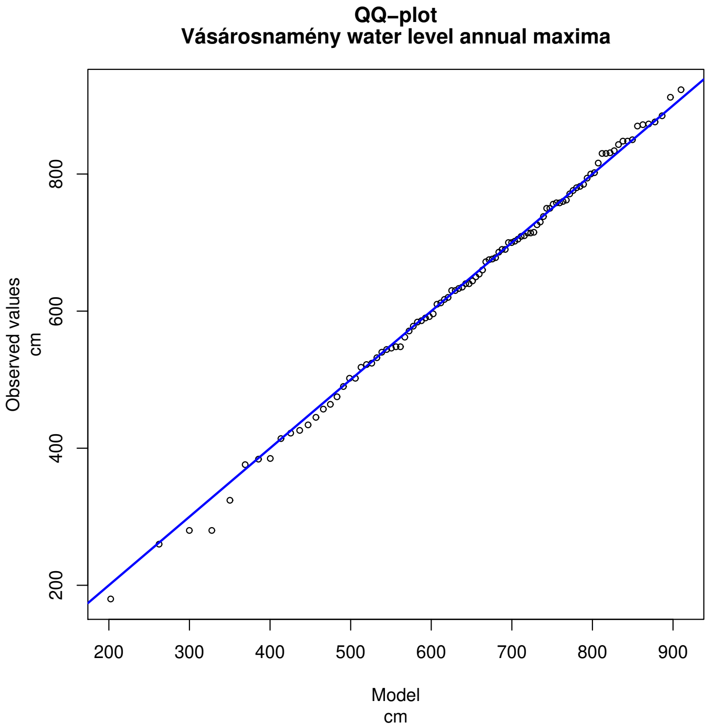

Figure 1: QQ plot of GEV fit to the Vásárosnamény data

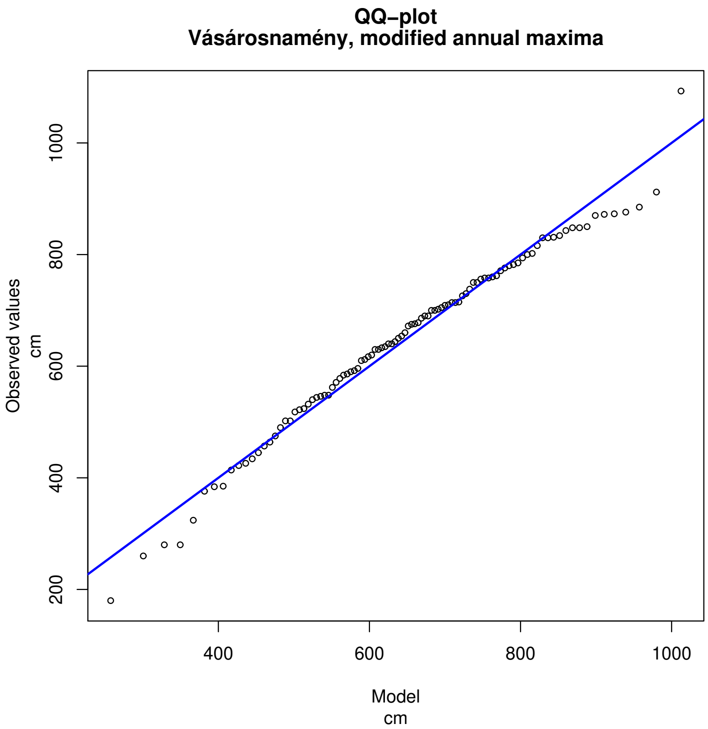

Figure 2: QQ plot of GEV fit to the modified Vásárosnamény data (the maximal value has been increased until it was rejected by the most powerful test)

Table 11: Number of changes needed for the different tests to reject the GEV hypothesis at the level of 95%

Test	B_n^2	A_n^2	C-v M	B_n^2	A_n^2	C-v M	B_n^2	A_n^2	C-v M
	$x_i \to \max$, noest			$x_i \to \max$, reest			$\max \to \max + i * 10$, reest		
k	9	11	25	8	11	17	16	19	36

Table 12: Percentiles of samples rejected by the different tests at the level of 95%

	window size: 25		window size: 50		window size: 75	
Test	streamflow	water level	streamflow	water level	streamflow	water level
A-D	0	0	1	1	0	0
B	0	0	2	3	0	0
h	0	0	1	4	0	1

2.4 Applications to real data

Here we continue the analysis of the flood data, which was already introduced in the previous subsection. Now we are mainly interested if the GEV can be accepted for the annual maxima series for different stations even if we consider shorter sequences of observation (useful for backtesting our procedures).

In Table 12 we give the results for windows 25, 50 and 75 years for the 6 stations over the river Tisza. We see that in most cases the GEV distribution can be accepted. The most sensitive are the two new tests.

3 Generalized Pareto Distributions

If one wants to use not only the block maxima, but all values that exceed a high threshold u, the following characterisation of the possible limit distributions (as $u \to \infty$) give the theoretical background. Under the same mild conditions which ensure the Fisher-Tippett theorem, we have the following as the asymptotic distribution function of the excesses

$$P(X - u > z | X > u) \approx G(z) = 1 - \left(1 + \frac{\xi z}{\sigma}\right)^{-1/\xi}$$

for $\{z : z > 0 \text{ and } (1 + \xi z/\sigma) > 0\}$. G is the so-called generalized Pareto distribution; GPD, see [2] for example.

The same questions as in Section 2 are also appropriate here: if we can reject the fit of the GPD family, then either the threshold u has not been chosen high enough, or possibly the conditions of the Fisher-Tippett theorem are not fulfilled. In a recent paper Choulakian and Stephens [1] investigate the goodness-of-fit tests for the GPD. They

Table 13: Critical values for B_n^2 in case of estimated parameters, $\xi = 0.5$ and $\xi = 0.2$

Prob. / n	25	50	100	200	400	25	50	100	200	400
0.5	0.149	0.149	0.149	0.149	0.150	0.156	0.154	0.154	0.155	0.155
0.9	0.282	0.283	0.282	0.284	0.284	0.298	0.294	0.296	0.297	0.298
0.95	0.339	0.340	0.339	0.340	0.341	0.362	0.355	0.356	0.359	0.359
0.99	0.478	0.474	0.484	0.476	0.470	0.550	0.492	0.498	0.505	0.503

Table 14: Critical values for B_n^2 in case of estimated parameters, $\xi = -0.2$ and $\xi = -0.6$

Prob. / n	25	50	100	200	400	50	100	200	400
0.5	0.176	0.169	0.170	0.170	0.171	0.214	0.199	0.198	0.197
0.9	0.380	0.333	0.331	0.335	0.339	0.585	0.411	0.406	0.405
0.95	0.584	0.406	0.401	0.406	0.411	0.994	0.504	0.496	0.496
0.99	0.563	0.585	0.566	0.583	0.580	1.453	0.753	0.710	0.710

apply the Anderson-Darling test, together with the Cramér-von Mises test and consider the cases of both known and unknown parameters. They show that the Anderson-Darling test outperforms the Cramér-von Mises test in detecting tail discrepancies.

In the sequel we compare the B_n^2-test from the previous section to the A-D test. An analogue of the h-test may be based on the property that if the GPD fits to the excesses of the sample from F, then the excesses of F^2 also belong to the GPD family with the same ξ, but we do not elaborate this idea here.

In Tables 13 and 14 we present the critical values of B_n^2 for different sample sizes for the case when $\xi = 0.5$ and $\xi = 0.2$, and $\xi = -0.2$ and $\xi = -0.6$, respectively. There are some irregular behaviour of the statistics to be observed in the case of $\xi = -0.6$ for small samples; that is why we omitted the values for $n = 25$.

We have applied both the A and the B-test to the flood data of the previous section. Here we have observed a smaller difference in favour of B: in the same experiment as in Subsection 3.3 (with re-estimation), $i = 9$ changes were needed for detecting the departure from the GPD for B^2, compared to $i = 10$ for A.

4 Expected shortfall

McNeil and Frey [7] investigate a model, where the GPD parameters are estimated from the observations of the last n days. These methods lead to VaR ($1 - \alpha$ quantile) and expected shortfall ($E(X_{t+1} - v_t \mid X_{t+1} > v_t)$) estimators.

For the VaR, a simple permutation test is available for checking the hypothesis that $P(X_{t+1} > v_t) = \alpha$ (where v_t is the estimator for VaR at level α based on days $t, t -$

Table 15: Percentiles of samples rejected by the likelihood ratio test at different levels

Level	$\xi = 0.2$	$\xi = -0.2$	$\xi = -0.5$
95%	10	20	40
99%	0	10	25

$1, \ldots t - n + 1$ and X_{t+1} is the observation on day $t+1$). The backtesting means that we compare the observed frequency of the event $\{X_{t+1} > v_t\}$ to α by a simple permutation test (see [7], for example).

The estimation of the expected shortfall can also be based on the GPD, let us denote the estimator by S_t (For details see [7].) The backtesting is more complicated, since we have non-identically distributed observed shortfalls. In McNeil and Frey a bootstrap test is proposed.

We suggest another approach: the profile likelihood is widely applied to confidence interval construction as well as testing hypotheses in the extreme-value setting, see for example Coles [2]. The data used are the increments $(X_{t+1} - S_t)/\sigma_t$ (condition: $\{X_{t+1} > v_t\}$). Let us suppose the independence of these standardized increments and investigate its distribution. The conditional distribution of X_{t+1} on $\{X_{t+1} > v_t\}$ can be considered as GPD (we have a variable, but high threshold v_t) so the standardization results again in a GPD. However, here we have a three-parameter family of distributions instead of the more usual two-parameter family:

$$F(x) = 1 - (1 + \xi(x - \mu)/\sigma)^{-1/\xi} \tag{10}$$

where $x \geq \mu$ if $\xi \geq 0$ and $x \in [\mu, \mu - \sigma/\xi]$ if $\xi < 0$. This gives that there is no local maximum of the likelihood function in μ, but the absolute maximum is achieved for $\hat{\mu} = \min x_i$. Plugging this estimator into the likelihood function, the other two parameters σ and ξ can then be estimated on the usual way by maximum likelihood.

The likelihood ratio test can be constructed as follows: let us consider the maximum of the log-likelihood function for different values of μ (let us denote it by m_μ). Those values of μ can be accepted as plausible values, for which $m_\mu > m_{\hat{\mu}} - c_\alpha/2$, where c_α is the $1 - \alpha$-quantile of the χ-squared distribution with one degrees of freedom.

In a simulation study we investigated the model, where the Var estimates are based on a sequence of iid GPD observations. In case of a shortfall, the difference between the actual estimator and the observed value is recorded. These differences are though not independent, but due to the windows of size n they are in the worst case n-dependent (in practice due to the rare event of shortfall it is very near to independence), so we may hope for acceptable properties of the proposed test.

The results of the simulation study are given in Table 15. The results show that the proportion of rejection is substantially higher for $\xi < 0$ than for $\xi > 0$.

Acknowledgements

This work was done while the author held a research grant of the DAAD at the Technical University of Munich. Support from the SFB 386 is also gratefully acknowledged.

References

[1] Choulakian, V. and Stephens, M.A. (2001): Goodness-of-fit tests for the genaralized Pareto distribution *Technometrics* **43** p. 478-484.

[2] Coles, S. (2001) *An Introduction to Statistical Modeling of Extreme Values* Springer:London.

[3] Dietrich,D. - de Haan, L. and Hüsler, J. (2002): Testing extreme value conditions *Extremes* **5** p. 71-85.

[4] Embrechts, P. - Klüppelberg, K. and Mikosch, T. (1997): *Modelling Extremal Events* Springer: Berlin, Heidelberg.

[5] de Haan, L. (1984): A spectral representation for max-stable processes. *Ann. Prob.* **12**, p. 1194-1204.

[6] Kotz, S. and Nadarajah, S. (2000): *Extreme Value Distributions* Imperial College Press London.

[7] McNeil, A.J. and Frey, R. (2000) Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme approach. *J. of Empirical Finance* **7**, p. 271-300.

[8] Smith, R.L. (1985): Maximum likelihood estimator in a class of nonregular cases *Biometrika* **72** p. 67-90.

[9] Stephens, M.A. (1976): Asymptotic results for goodness-of-fit statistics with unknown parameters *The Annals of Statistics* **4** p. 357-369.

[10] Zempléni, A. (1996): Inference for Generalized Extreme Value Distributions *Journal of Applied Statistical Science* **4** p. 107-122.