

On testing extreme value conditions

Jürg Hüsler · Deyuan Li

Received: 19 October 2005 / Revised: 11 August 2006 /
Accepted: 11 August 2006 / Published online: 28 October 2006
© Springer Science + Business Media, LLC 2006

Abstract Applications of univariate extreme value theory rely on certain assumptions. Recently, two methods for testing these extreme value conditions are derived by [Dietrich, D., de Haan, L., Hüsler, J., *Extremes* 5: 71–85, (2002)] and [Drees, H., de Haan, L., Li, D., *J. Stat. Plan. Inference*, 136: 3498–3538, (2006)]. In this paper we compare the two tests by simulations and investigate the effect of a possible weight function by choosing a parameter, the test error and the power of each test. The conclusions are useful for extreme value applications.

Keywords Extreme value conditions · Test statistic · Weight function · Power

AMS 2000 Subject Classification 62G32

1 Introduction

In order to estimate quantiles and distribution tails outside of the range of the available data, it is usually assumed that the underlying distribution function (d.f.) F belongs to the max-domain of attraction of an extreme value distribution G_γ with $\gamma \in \mathbb{R}$, denoted by $F \in \mathcal{D}(G_\gamma)$, i.e., there exist constants $a_n > 0$, $b_n \in \mathbb{R}$ such that

$$\lim_{n \rightarrow \infty} F^n(a_n x + b_n) = G_\gamma(x) := \exp(-(1 + \gamma x)^{-1/\gamma}), \quad (1.1)$$

for all $1 + \gamma x > 0$. In case of $\gamma = 0$, $G_0(x)$ is interpreted as $\exp(-e^{-x})$.

J. Hüsler (✉) · D. Li
Department of Mathematical Statistics and Actuarial Science, University of Bern,
3012 Bern, Switzerland
e-mail: juerg.huesler@stat.unibe.ch

Condition (1.1) is called the extreme value condition. Under this condition the common approach usually estimates the extreme value index γ and normalizing constants a_n and b_n for a given sample of size n (e.g., Hill (1975), Pickands (1975), Hall (1982), Smith (1985) and Dekkers et al. (1989)). Then based on these estimators, the quantiles or tail probabilities can be derived (e.g., de Haan and Rootzén (1993)). However the extreme value condition is not always fulfilled, for example, the Poisson distribution, the negative binomial distribution and the truncated exponential distribution $F \sim X = \exp[E]$ are not members of $\mathcal{D}(G_\gamma)$ for any $\gamma \in \mathbb{R}$ (e.g., see Anderson (1970) and Leadbetter et al. (1983)), where E is a standard exponential distribution random variable (r.v.) and $[x]$ means the largest integer smaller than x . So if we have no a priori knowledge on whether F belongs to the max-domain, before application we have to check the assumption:

$$H_0 : F \in \mathcal{D}(G_\gamma) \quad \text{for some } \gamma \in \mathbb{R}.$$

Let X_1, X_2, \dots, X_n be independent random variables with common distribution function F . Dietrich et al. (2002) present the following approach to test $H_0: F \in \mathcal{D}(G_\gamma)$ for some $\gamma \in \mathbb{R}$ with some additional (second order) conditions, then for each $\eta > 0$, the test statistic

$$E_n := k \int_0^1 \left(\frac{\log X_{n-[kt],n} - \log X_{n-k,n}}{\hat{\gamma}_+} - \frac{t^{-\hat{\gamma}_-} - 1}{\hat{\gamma}_-} (1 - \hat{\gamma}_-) \right)^2 t^\eta dt, \quad (1.2)$$

where the integer k satisfies that $k \rightarrow \infty$, $k/n \rightarrow 0$ and $k^{1/2}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$ and A is related to the second order condition, converges in distribution to

$$E_\gamma := \int_0^1 \left((1 - \gamma_-)(t^{-\gamma_- - 1}W(t) - W(1)) - (1 - \gamma_-)^2 \frac{t^{-\gamma_-} - 1}{\gamma_-} \right. \\ \left. + \frac{t^{-\gamma_-} - 1}{\gamma_-} R + (1 - \gamma_-)R \int_t^1 s^{-\gamma_- - 1} \log s ds \right)^2 t^\eta dt,$$

where $\gamma_+ = \max\{\gamma, 0\}$, $\gamma_- = \min\{\gamma, 0\}$, W is Brownian motion, and r.v.'s P , R are some integrals involving W (for details see Dietrich et al. (2002)). If $\gamma = 0$, then the term $(t^\gamma - 1)/\gamma$ is interpreted as $\log t$. Moreover, if we assume $\gamma \geq 0$, we may apply the test statistic

$$\text{PE}_n := k \int_0^1 \left(\frac{\log X_{n-[kt],n} - \log X_{n-k,n}}{\hat{\gamma}_+} + \log t \right)^2 t^\eta dt \quad (1.3) \\ \xrightarrow{d} \text{PE} := \int_0^1 \left(t^{-1}W(t) - W(1) + \log t \int_0^1 (s^{-1}W(s) - W(1))ds \right)^2 t^\eta dt.$$

Note that in Dietrich et al. (2002), the estimates for γ_+ and γ_- are fixed to be the moment estimators (Dekkers et al. (1989)). In general, the limiting r.v. E_γ only depends on γ and η . Dietrich et al. (2002) state the result for $\eta = 2$, but it is easy to extend the result for any $\eta > 0$ which parameterize the

weight function t^η . Here we correct one *sign* error (the sign before $(1 - \gamma_-)R$ should be $+$) in the formula of E_γ in Dietrich et al. (2002), and recalculate the quantiles $Q_{p,\gamma}(= Q_{p,\gamma,\eta})$ of the limiting r.v. E_γ (see Table 1).

For testing H_0 , we first choose an η and continue as follows:

- First, estimate $\hat{\gamma}_+$ and $\hat{\gamma}_-$ by the moment estimator and calculate the value of the test statistic E_n .
- Secondly, determine the corresponding quantile $Q_{1-\alpha,\hat{\gamma}}$ of the distribution $E_{\hat{\gamma}}$, by linear interpolation if $\hat{\gamma} < 0$, i.e.,

$$Q_{1-\alpha,\hat{\gamma}} = Q_{1-\alpha,\gamma_1} + \frac{\hat{\gamma} - \gamma_1}{\gamma_2 - \gamma_1}(Q_{1-\alpha,\gamma_2} - Q_{1-\alpha,\gamma_1}), \quad \text{if } \hat{\gamma} \in [\gamma_1, \gamma_2],$$

using the values of Table 1. Here $\hat{\gamma} = \hat{\gamma}_+ + \hat{\gamma}_-$ and α is usually 0.05. If $\hat{\gamma}$ is outside of the range of γ 's in Table 1, we let $Q_{1-\alpha,\hat{\gamma}}$ be the nearest tabulated quantile. For example, if $\hat{\gamma} = -0.75$, let $Q_{1-\alpha,\hat{\gamma}} = Q_{1-\alpha,-0.7}$.

- Finally, compare the value of E_n with the value of $Q_{1-\alpha,\hat{\gamma}}$. If $E_n > Q_{1-\alpha,\hat{\gamma}}$, then reject H_0 with type I error α . Otherwise we have no reason to reject H_0 .

In the same way the test statistic PE_n can be used to test that $F \in \mathcal{D}(G_\gamma)$ assuming $\gamma \geq 0$. In abbreviation, we denote the two tests in Eq. 1.2 and Eq. 1.3 as test E and test PE, respectively.

Drees et al. (2006) propose another method to test $F \in \mathcal{D}(G_\gamma)$ assuming $\gamma > -1/2$: if $F \in \mathcal{D}(G_\gamma)$ for some $\gamma > -1/2$ with some additional (second order) conditions, then for each $\eta > 0$, the test statistic

$$T_n := k \int_0^1 \left(\frac{n}{k} \bar{F}_n(\hat{a}_{n/k} \frac{x^{-\hat{\gamma}} - 1}{\hat{\gamma}} + \hat{b}_{n/k}) - x \right)^2 x^{\eta-2} dx, \tag{1.4}$$

where the integer k satisfies that $k \rightarrow \infty$, $k/n \rightarrow 0$ and $k^{1/2}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$ and A is related to the second order condition, converges in distribution to

$$T_\gamma := \int_0^1 (W(x) + L_\gamma(x))^2 x^{\eta-2} dx,$$

Table 1 Quantiles $Q_{p,\gamma}$ of the limiting r.v. E_γ for test E with $\eta = 2$

γ	p							
	0.10	0.30	0.50	0.70	0.90	0.95	0.975	0.99
≥ 0	0.028	0.042	0.057	0.078	0.122	0.150	0.181	0.222
-0.1	0.027	0.041	0.054	0.074	0.116	0.144	0.174	0.213
-0.2	0.027	0.040	0.053	0.072	0.114	0.141	0.169	0.208
-0.3	0.027	0.040	0.054	0.073	0.113	0.140	0.168	0.206
-0.4	0.027	0.040	0.054	0.073	0.114	0.141	0.169	0.207
-0.5	0.027	0.040	0.054	0.073	0.115	0.141	0.169	0.208
-0.6	0.027	0.040	0.054	0.074	0.116	0.144	0.173	0.212
-0.7	0.028	0.041	0.055	0.074	0.118	0.147	0.176	0.218

where W is Brownian motion, and the process L_γ depends on the asymptotic distribution of $(\hat{\gamma}, \hat{a}, \hat{b})$, which is some \sqrt{k} -consistence estimator of (γ, a, b) (for details see Drees et al. (2006)). In that paper the exact formulas of L_γ and T_γ are given for the maximum likelihood estimates (e.g., see Smith (1985) and Drees et al. (2004)), which depend only on γ and η . The test statistic T_n can be used to test $F \in \mathcal{D}(G_\gamma)$ for some $\gamma > -1/2$. In abbreviation we denote this test (1.4) as test T in the following comparison.

Small simulations are done in both papers: $\eta = 2$ in Dietrich et al. (2002) and $\eta = 1$ in Drees et al. (2006). But both papers do not discuss the choice of η , the parameter of the weight function. Theoretically we can choose any positive η . A small value of η means that we put more weight on the larger order statistics. Here we compare the behaviour of the three tests T, E and PE for $\eta = 0.5, 1$ and 2 and $\alpha = 0.05$.

We discuss extensively the three tests and answer the following questions by simulations:

- (i) What is a good choice of η for each test, and how close is the 0.95 quantile of the test statistic for finite sample sizes to the corresponding limiting quantile?
- (ii) What is the type I error (TIE) of each test for distributions belonging to the max-domain, and what is the power of each test for distributions not belonging to the max-domain?

2 Quantiles of the Limiting r.v.'s

As mentioned above, we need to calculate the quantiles of the limiting r.v.'s. Note that the limiting r.v.'s for the three tests: T, E and PE, only depend on γ and η . To simulate the limiting r.v., the Brownian motion on $[0, 1]$ is simulated on a grid with 100,000 points, and the integral is approximated by a Riemann sum. For example, write $E_\gamma = \int_0^1 f(\gamma, \eta, t, W(t))dt$ and assume $W(i/m), 0 \leq i \leq m = 100,000$ is a realization of a Brownian motion W at the discrete points in $[0, 1]$. Then E_γ is approximated by

$$\frac{1}{m} \sum_{i=1}^m f\left(\gamma, \eta, \frac{i-1/2}{m}, \frac{W((i-1)/m) + W(i/m)}{2}\right).$$

Repeating the simulation 50,000 times, we use the order statistics as the corresponding quantiles. Thus we receive the quantiles for each test and three

Table 2 Quantiles $Q_{p,\gamma}$ of the limiting r.v. PE for test PE ($\gamma \geq 0$)

η	p							
	0.10	0.30	0.50	0.70	0.90	0.95	0.975	0.99
0.5	0.348	0.493	0.639	0.836	1.250	1.515	1.792	2.145
1	0.107	0.159	0.211	0.284	0.438	0.534	0.663	0.793
2	0.034	0.053	0.074	0.105	0.176	0.222	0.272	0.341

Table 3 Quantiles $Q_{p,\gamma}$ of the limiting r.v. T_γ for test T with $\eta = 1.0$

γ	p							
	0.10	0.30	0.50	0.70	0.90	0.95	0.975	0.99
4	0.086	0.123	0.161	0.212	0.322	0.393	0.462	0.558
3	0.085	0.120	0.156	0.205	0.307	0.372	0.440	0.532
2	0.083	0.116	0.150	0.195	0.286	0.344	0.402	0.489
1.5	0.082	0.115	0.148	0.192	0.282	0.340	0.400	0.480
1	0.082	0.114	0.146	0.189	0.276	0.330	0.388	0.466
0.5	0.083	0.116	0.149	0.194	0.285	0.343	0.404	0.481
0.25	0.085	0.119	0.153	0.200	0.295	0.355	0.415	0.499
0	0.089	0.126	0.163	0.213	0.319	0.388	0.455	0.542
-0.1	0.091	0.129	0.168	0.221	0.330	0.400	0.471	0.569
-0.2	0.093	0.133	0.174	0.231	0.350	0.425	0.500	0.604
-0.3	0.096	0.139	0.183	0.242	0.369	0.449	0.531	0.653
-0.4	0.100	0.145	0.192	0.256	0.393	0.484	0.576	0.690
-0.45	0.103	0.150	0.199	0.320	0.416	0.511	0.605	0.735
-0.499	0.107	0.157	0.210	0.338	0.439	0.546	0.652	0.799

η 's (see Tables 1, 2 and 3). Here we only show three important quantile tables (more tables are available on www.imsv.unibe.ch/~deyuan/research.html). From the convergence of the simulated results, the first two digits of these simulated quantiles are accurate. Compared to the quantile tables in Dietrich et al. (2002) and Drees et al. (2006), Tables 1, 2 and 3 give the quantiles for more values of γ and p , being more accurate than the former derived values.

3 Choice of η

Now we give some comments on the choice of η . There are several criteria for the selection of η . A good η should be such that the $1 - \alpha$ quantile of the test statistic for a finite sample size is rather close to the corresponding limiting quantile. But from a pure statistical point of view the η should be such that the nominal α is close to the true α and that the power is large in all or many cases, even for not very large sample sizes.

We begin our discussion by measuring the closeness of the quantiles by the relative errors (T R E, see below). Of course the good choice of η depends on many factors such as the sample size n , proportion of k/n and the underlying d.f. F .

In order to investigate the choice of η for each test, we vary n from 500 to 10,000, which is the usual sample size bound for application. Considering the condition $k \rightarrow \infty, k/n \rightarrow 0$ as $n \rightarrow \infty$ and according to the experience, we choose k based on the following three cases:

Case 1: n (k/n) = 500(5%), 1,000(4%), 2,000(3%), 5,000(2%), 10,000(1.5%);

Case 2: n (k/n) = 500(8%), 1,000(6%), 2,000(4%), 5,000(3%), 10,000(2%);

Case 3: n (k/n) = 500(10%), 1,000(8%), 2,000(5%), 5,000(3%), 10,000(2%).

As underlying distributions we use the Cauchy, Log-gamma, the Burr, the Weibull and EV distributions. All these distributions are in the max-domain $\mathcal{D}(G_\gamma)$ with the extreme value index γ varying from 2 to -0.2 , and satisfy the second order condition required both in Dietrich et al. (2002) and Drees et al. (2006). Definitions of these distributions are given in the last section, in the Appendix.

For each chosen distribution F we simulate the 0.95 quantile of the test statistic, say $\hat{Q}_{0.95,F}$, by generating 500,000 r.v.'s of the test statistic, and calculate the sum of the relative errors on all cases, resulting in the total relative error (T R E), i.e.

$$\text{T R E}(F) := \sum_{\text{Case } j} \sum_{n_i(k_i/n_i)} \left| \hat{Q}_{0.95,F} - Q_{0.95,\gamma(F)} \right| / Q_{0.95,\gamma(F)}. \tag{3.1}$$

Then, a good choice of η for each test should have a small or the smallest T R E for most distributions. We compared also the relative errors in each case with a given sample size, but found that summarizing is a reasonable approach here. The values of T R E for the three tests are listed in Table 4.

From Table 4, we note that:

- For test T, $\eta = 1$ is much better than $\eta = 0.5$ and slightly better than $\eta = 2$. So we might choose $\eta = 1$.
- For test E, $\eta = 2$ is much better than $\eta = 0.5$ and 1 for Log-gamma(2, 5), Burr(1, 2, 2), Weibull(1, 0.5), EV(-0.2) distributions, but is not as good as $\eta = 0.5$ and 1 in case of Cauchy distribution. So in general, we might choose $\eta = 2$.

Table 4 T R E for different distributions and three tests

Distribution	γ	Test	$\eta = 0.5$	$\eta = 1$	$\eta = 2$	Good η
Burr(1, 0.5, 1)	2	T	0.501	0.091	0.092	1 or 2
		E	0.427	0.318	0.455	1
		PE	0.800	0.430	0.763	1
Cauchy	1	T	0.596	0.070	0.094	1
		E	0.148	0.157	0.457	0.5 or 1
		PE	0.562	0.440	0.447	1 or 2
Log-gamma(2, 5)	0.5	T	0.794	0.391	0.330	1 or 2
		E	1.675	0.872	0.486	2
		PE	1.272	0.844	2.217	1
Burr(1, 2, 2)	0.25	T	0.896	0.097	0.451	1
		E	2.047	1.055	0.413	2
		PE	1.067	1.452	3.369	0.5
Weibull(1, 0.5)	0	T	1.924	1.255	1.414	1
		E	6.188	3.079	0.569	2
		PE	7.281	13.08	18.68	0.5
EV(-0.2)	-0.2	T	1.573	0.666	0.664	1 or 2
		E	2.373	1.037	0.431	2

- For test PE, $\eta = 0.5$ is better for Burr (1, 2, 2) and Weibull(1, 0.5) distributions, but for other distributions $\eta = 1$ is better. So we observe that $\eta = 0.5$ is better for small γ , in the sense of T R E. Since often in applications γ is small, $\gamma < 1$, we prefer $\eta = 0.5$ as a reasonable value.

From now on, we fix $\eta = 1$ for test T, $\eta = 2$ for test E, and $\eta = 0.5$ for test PE.

From Table 4, we see that the values of T R E for test E (with $\eta = 2$) vary from 0.42 to 0.57, that those values for test PE (with $\eta = 0.5$) vary from 0.56 to 7.29, and that those values for test T (with $\eta = 1$) vary from 0.07 to 0.40 for $\gamma > 0$ and vary from 0.66 to 1.26 for $\gamma \leq 0$. Hence in view of T R E, we might conclude that (i) test E is always better than test PE; (ii) test T is better than test E for $\gamma > 0$ and test E is better than test T for $\gamma \leq 0$. Possibly, a smaller T R E implies a more accurate test, but this has to be analyzed with respect to type I error α .

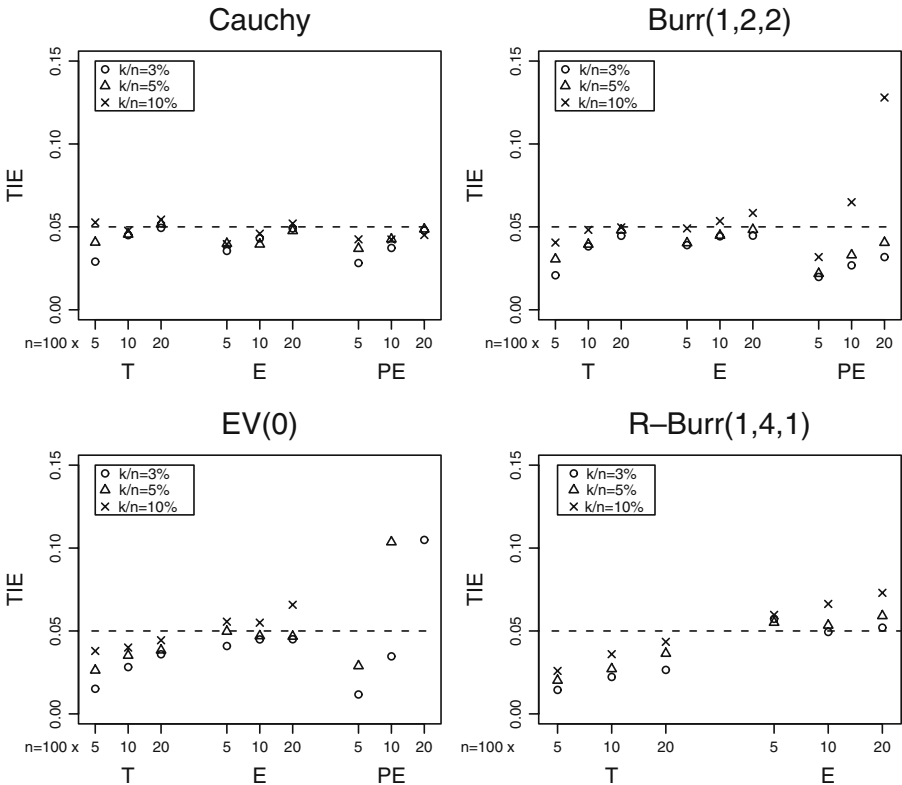


Fig. 1 Type I errors for the distributions satisfying the second order condition

4 Test Error and Power

In both Dietrich et al. (2002) and Drees et al. (2006), a second order condition on F is required. The second order condition introduced and discussed in detail by de Haan and Stadtmüller (1996), is one common assumption and often used in extreme value theory. Most distributions satisfy this second order condition. For such distributions we first analyze the type I error for finite sample sizes. But there still exist some distributions which belong to the max-domain but do not satisfy the second order condition, for example the Pareto distribution. So the question arises: what is the type I error for a distribution which is in the max-domain but does not satisfies the second order condition? We consider secondly the type I error for these particular distributions for finite sample sizes. Finally we investigate for finite sample sizes, the power of each test for distributions which are not in the max-domain.

Again, the choice of η should be discussed also with respect to the type I error or the power. The findings of Section 3 where the choice of η is based on small T R E might be unappropriate considering type I error and power. This is answered in the next section. The following simulations are based on

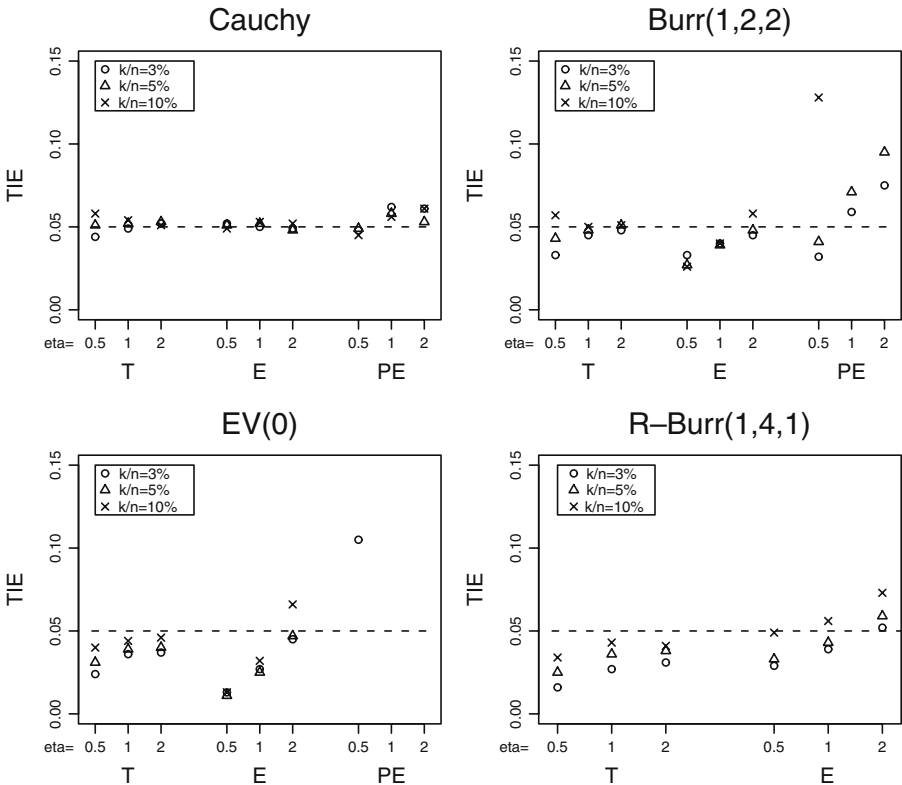


Fig. 2 Type I error for different η with sample size $n = 2,000$

the choices: $\eta = 1$ for test T, $\eta = 2$ for test E and $\eta = 0.5$ for test PE. The test errors and powers are calculated based on the 10,000 replicates of the test statistics.

4.1 F Satisfies the Second Order Condition

We simulate samples of the Burr(1,0.5,1), Cauchy, Log-gamma(2,5), Burr(1,2,2), Weibull(1,0.5), EV(0), EV(-0.1), EV(-0.2) and R-Burr(1,4,1) distributions. All the distributions belong to the max-domain and satisfy the second order condition. By deriving the test statistics and comparing them with the (asymptotic) 0.95 quantiles, we get all the type I errors. The results show that the errors are close to the given $\alpha (= 0.05)$. In Fig. 1 we plot type I errors for several distributions, where the values (crosses and triangles) for EV(0) distribution are missing because they are larger than 0.15. The same happens in other figures again. Note that the type I error depends obviously on the sample size n and the choice of k . Again test E seems slightly better than test T for $\gamma \leq 0$.

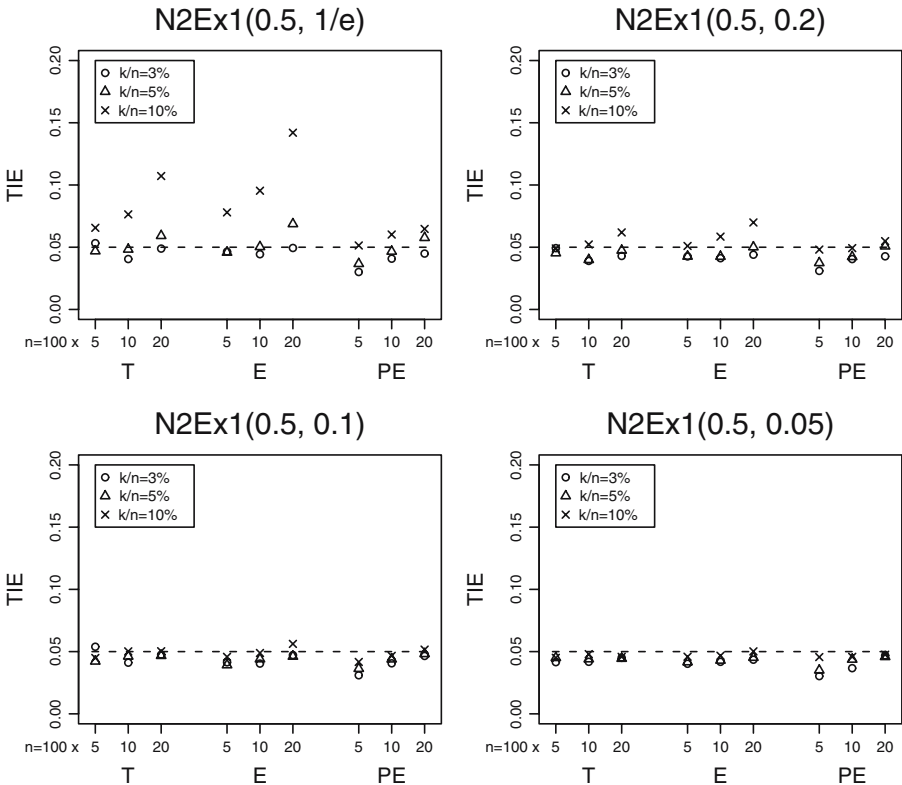


Fig. 3 Type I errors for the N2Ex1(0.5, λ) distributions with $\lambda = 1/e, 0.2, 0.1$ and 0.05

Figure 2 shows the type I errors for different η 's, again $\eta=0.5, 1,$ and 2 . We selected for this figure the sample size $n = 2,000$. We simulated also other sample sizes to observe the dependence of the type I error on the η . Details can be found at <http://www.imsv.unibe.ch/~deyuan/research.html>. From these simulations we can state that test E with $\eta = 2$ has very often the smallest type I errors which are close to the nominal $\alpha = 0.05$ for not large k (with $k/n \leq 0.05$). The type I errors of test T is best with $\eta = 1$ and 2 for positive γ . Also $\eta = 2$ is better in these cases for small sample sizes. But for $\gamma \leq 0$ test T is too conservative, $\eta = 1$ or 2 would be better. Finally for test PE η should be selected as 0.5 or 1 for smaller sample sizes. Also k/n should be ≤ 0.05 .

4.2 F belongs to the Max-domain but does not Satisfy the Second Order Condition

We mention that a second order condition is not the necessary condition for the limit theorems sketched in Section 1. In fact, one can prove that for a slowly increasing k , these limit theorems remain valid if F satisfies Eq. 1.1 without the second order condition. For example, Lemma 2.1 of Dietrich et al. (2002) holds

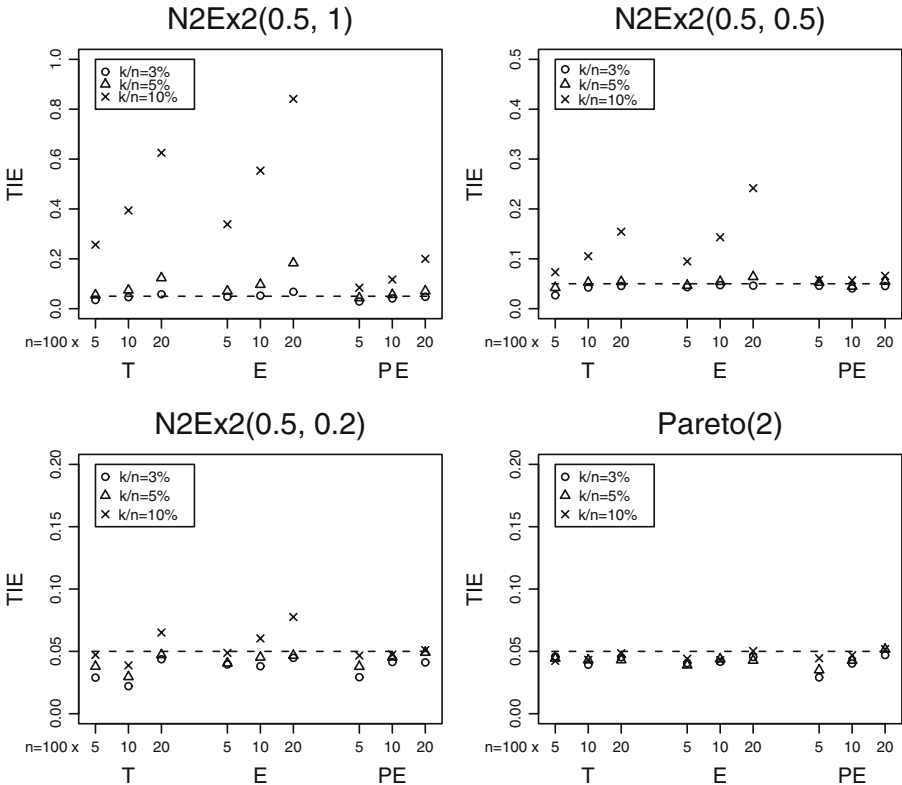


Fig. 4 Type I errors for the $N2Ex2(0.5, \lambda)$ distributions with $\lambda = 1, 0.5, 0.2$ and 0

under this weaker condition. A second order condition provides an easier way to specify the growth rate of k .

To construct a distribution which does not satisfy the second order condition, let $U(t) := F^{\leftarrow}(1 - 1/t)$, $t \geq 1$, where the left arrow \leftarrow denotes the generalized inverse function. It is well known that $F \in \mathcal{D}(G_\gamma)$ is equivalent to $U \in \text{RV}(\gamma)$ in case of $\gamma > 0$ or equivalent to $U(\infty) - U(\cdot) \in \text{RV}(\gamma)$ in case of $\gamma < 0$. In brief, the second order condition implies that the function U (for most cases) can be written as $U(t) = ct^\gamma(1 + A(t))$ in case of $\gamma > 0$ or $U(\infty) - U(t) = ct^\gamma(1 + A(t))$ in case of $\gamma < 0$ for large t , where $c > 0$, $A(t) \rightarrow 0$ as $t \rightarrow \infty$ and $|A| \in \text{RV}(\rho)$ with $\rho \leq 0$. Geluk and de Haan (1987) give several type functions which are not regular varying functions. Thus by taking A a non-regular varying function, we can easily construct the distributions which are in the max-domain but do not satisfy the second order condition. Here are three examples.

- N2Ex1(γ, λ) distribution with $\gamma > 0$ and $0 < \lambda \leq 1/e$:

$$U(t) = t^\gamma(\gamma^{-1} + \lambda t^{-1} \exp(\sin t)), \quad t \geq 1.$$

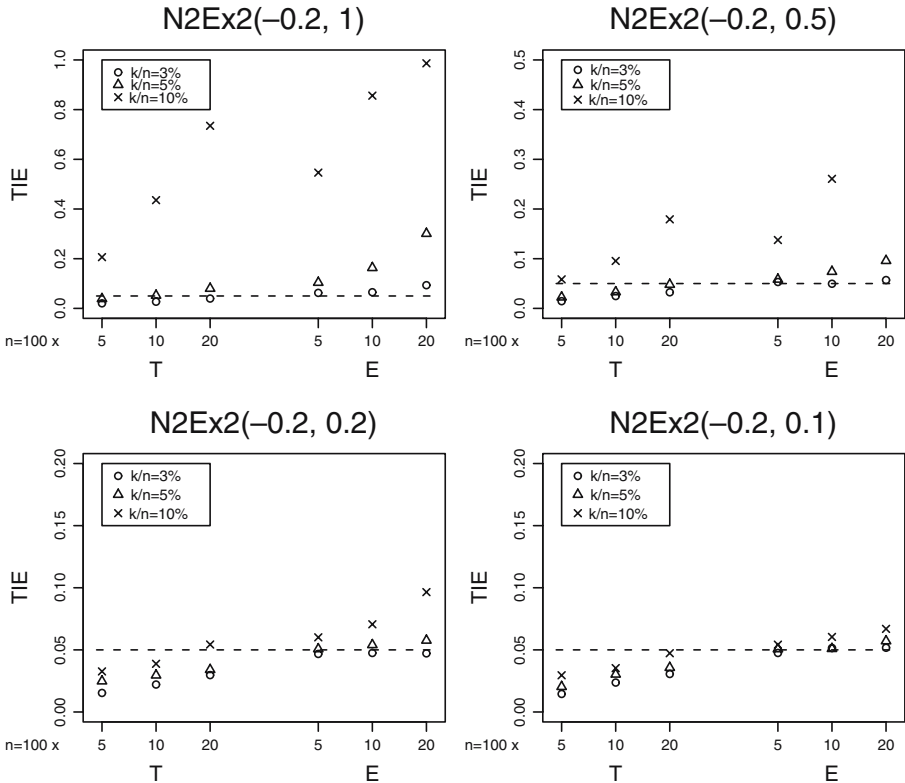


Fig. 5 Type I errors for the N2Ex2(-0.2, λ) distributions with $\lambda = 1, 0.5, 0.2$ and 0.1

- N2Ex2(γ, λ) distribution with $\gamma \in \mathbb{R}$ and $0 < \lambda \leq 1$: for $t \geq 1$,

$$U(t) = \begin{cases} (t + \lambda \sin t)^\gamma, & \gamma > 0, \\ 1 - (t + \lambda \sin t)^\gamma, & \gamma < 0. \end{cases}$$

- N2Ex3(γ, λ) distribution with $\gamma > 0$ and $0 < \lambda \leq 1$:

$$U(t) = t^\gamma + \lambda \sin(t^\gamma), \quad t \geq 1.$$

In the three examples above, it is not difficult to check that each U is non-decreasing. Hence the function F , corresponding to U , is non-decreasing thus a distribution function. Note that in each U the second term is responsible for the reason that F does not satisfy the second order condition, and that for small λ the second term is less important to U and it converges to zero as $\lambda \rightarrow 0$. Also note that the function U of Pareto distribution can be rewritten as $U(t) = ct^\gamma$ for large t with $c > 0$ and $\gamma > 0$ (i.e., no second term or regarding $\rho = -\infty$!), thus the three distributions converge to Pareto distribution as $\lambda \rightarrow 0$. Hence the value of λ reflects how close the distribution is to Pareto distribution (though we do not give a measure for the closeness between two distributions).

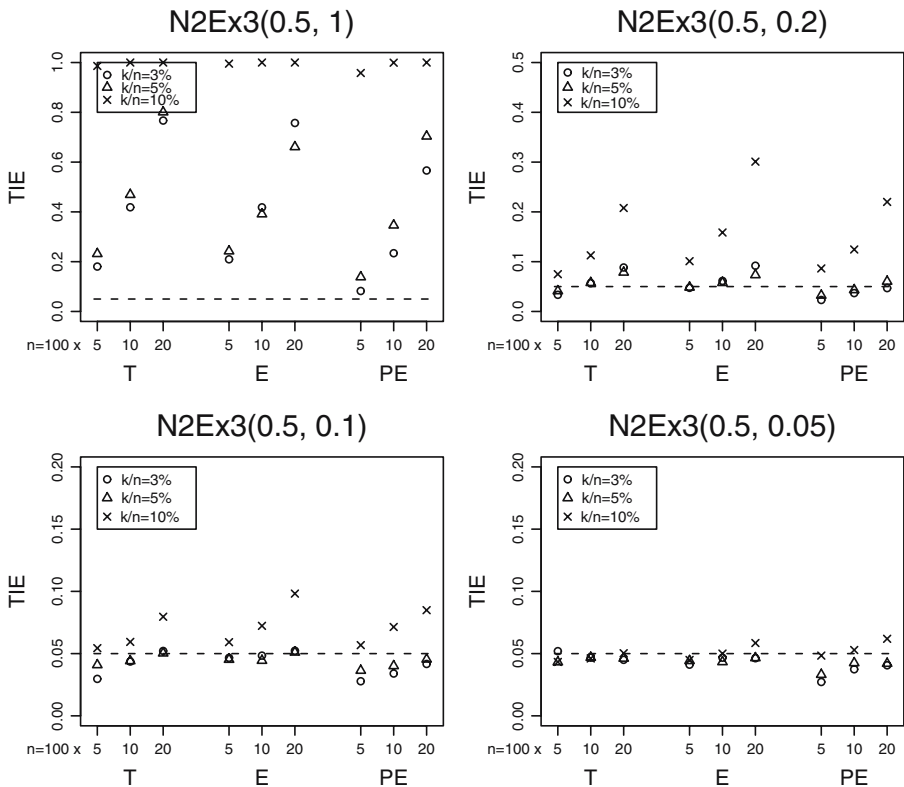


Fig. 6 Type I errors for the N2Ex3(0.5, λ) distributions with $\lambda = 1, 0.2, 0.1$ and 0.05

We simulate samples from the $N2Ex1(0.5, \lambda)$ distributions with $\lambda = 1/e, 0.2, 0.1$ and 0.05 , from the $N2Ex2(0.5, \lambda)$ distributions with $\lambda = 1, 0.5, 0.2$ and 0 , from the $N2Ex2(-0.2, \lambda)$ distributions with $\lambda = 1, 0.5, 0.2$ and 0.1 , and from the $N2Ex3(0.5, \lambda)$ distributions with $\lambda = 1, 0.2, 0.1$ and 0.05 . By Figs. 3, 4, 5 and 6 (note that the scales of y-axis's in each figure are different), it follows that most type I errors decrease and are close to the nominal $\alpha(=0.05)$ as λ decreases. On the other hand, we see that smaller k (e.g., $k/n = 3\%$ or 5%) behaviors better than larger k (e.g., $k/n = 10\%$), which also shows that a second order condition is not the necessary condition for the limit theorems in Section 1. Based on these results we should select a smaller k for large n 's meaning $k/n \leq 5\%$ to get conservative not liberal test.

4.3 F does not belong to the Max-domain

Beside the well known Poisson distribution, the negative binomial distribution and the truncated exponential distribution ($X = \exp[E]$), we consider several distributions which do not belong to the max-domain $\mathcal{D}(G_\gamma)$ for any $\gamma \in \mathbb{R}$:

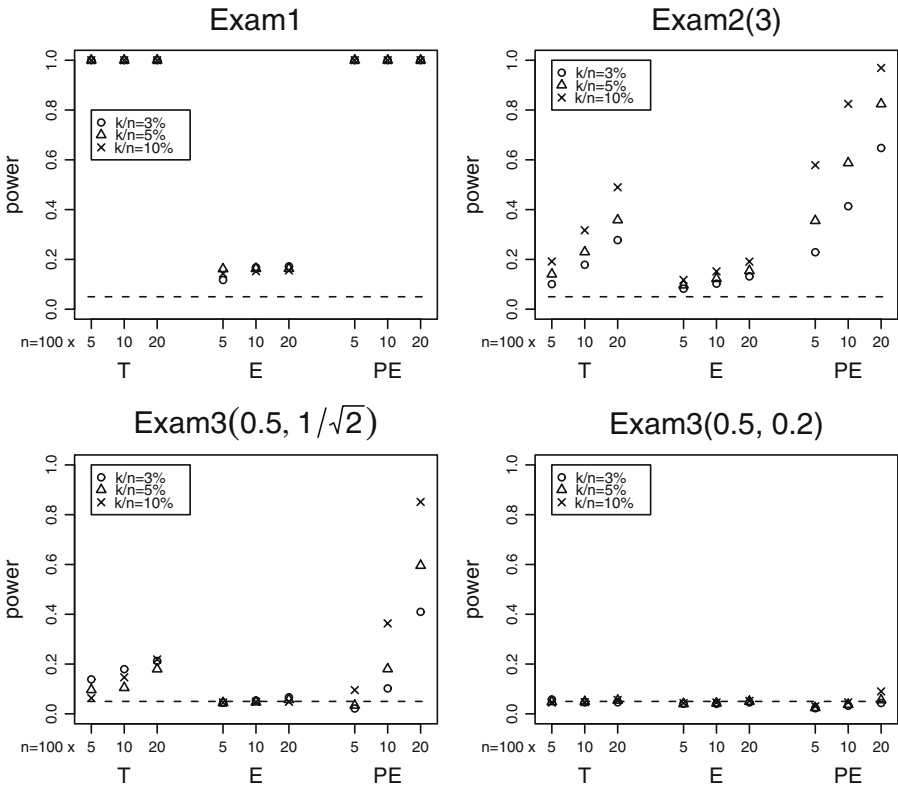


Fig. 7 Powers for the distributions not belonging to the max-domain

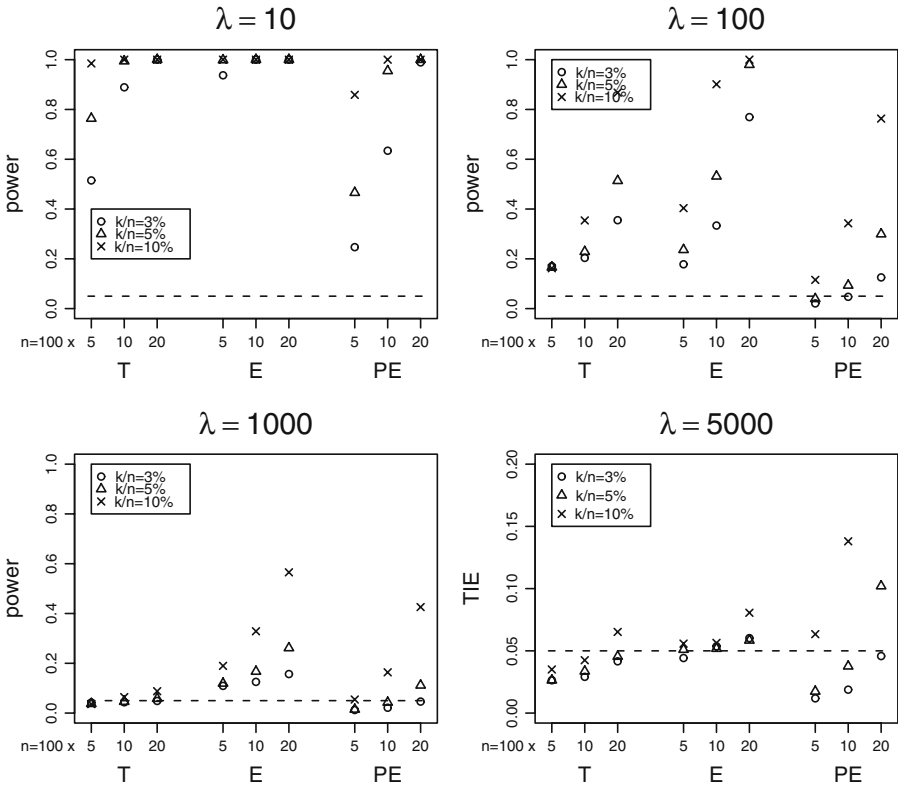


Fig. 8 Powers for the Poisson distributions with $\lambda = 10, 100, 1,000$ and $5,000$

- Exam1 distribution: $F(x) = 1 - \exp(-\sqrt{-\log(1-x)})$, $0 \leq x < 1$.
- Exam2(β) distribution with $\beta > 0$: $F(x) = 1 - (\log x)^{-\beta}$, $x \geq e$.
- Exam3(β, λ) distribution with $\beta > 0$ and $0 < \lambda \leq 1/\sqrt{2}$: $U(t) = t^\beta(1 + \lambda \sin(\beta \log t))$, $t \geq 1$.

It is shown (Falk et al., 2004) that the Exam1 and Exam2(β) distributions with $\beta > 0$, do not belong to the max-domain but belong to the p -max domain (c.f. page 64 in Falk et al., 2004) with power normalization. Note that the Exam3(β, λ) distribution with $\beta > 0$ is not in the max-domain because of the sin function, and that the term $\lambda \sin(\beta \log t)$ is less important to U for small λ . Again, the Exam3(β, λ) distribution converges to Pareto($1/\beta$) distribution as $\lambda \rightarrow 0$. Thus the value of λ reflects how close the distribution is to Pareto distribution, also partially reflects how close the distribution is to the max-domain.

We simulate samples from the Poisson distribution, the negative binomial distribution, the truncated exponential distribution, Exam1, Exam2(3), and Exam3(0.5, λ) distributions with $\lambda = 1/\sqrt{2}$ and 0.2. From the simulated results, it follows that for the truncated exponential distribution the powers of the three tests T, E and PE are very close to 1 (the difference between the

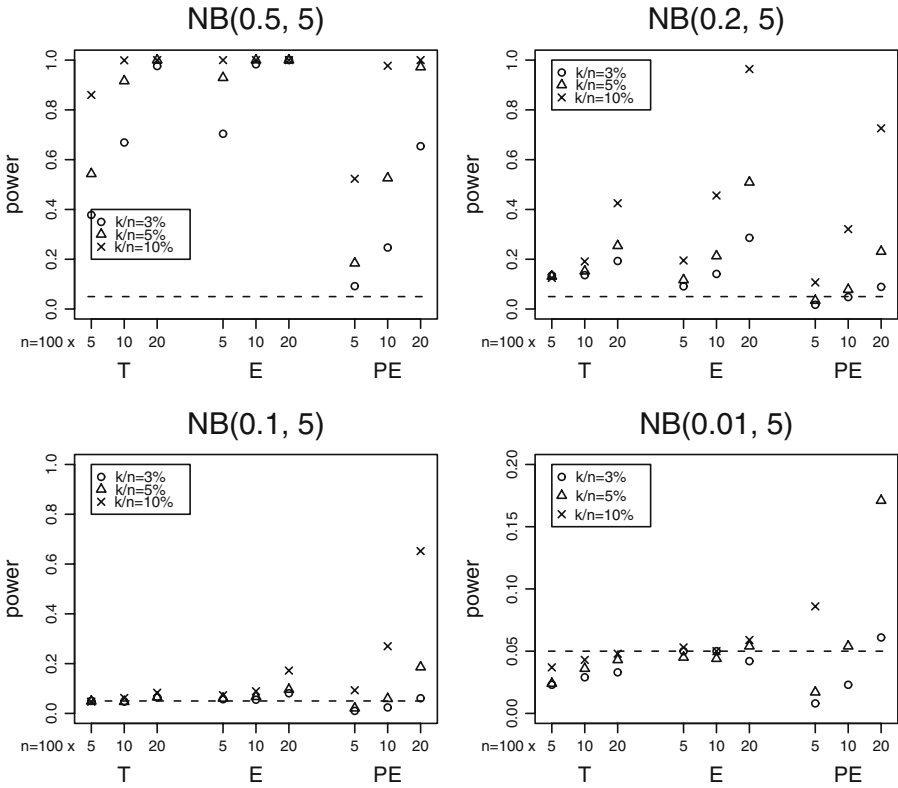


Fig. 9 Powers for the negative binomial distributions $NB(p, r)$ with $r = 5$, $p = 0.5, 0.2, 0.1$ and 0.01

power and 1 is smaller than 0.01). For the other distributions the powers are presented in Figs. 7, 8 and 9. Compared with Exam3(0.5, $1/\sqrt{2}$) distribution, Exam3(0.5, 0.2) distribution is closer to the Pareto(2) distribution (or it is closer to the max-domain), so its powers are smaller and close to the nominal α . By Proposition 1 in Anderson et al. (1997), the Poisson(λ_n) distribution approximately belongs to the max-domain $\mathcal{D}(G_\gamma)$ with $\gamma = 0$ as λ_n goes to infinity at some rate. Also, by Theorem 4 in Nadarajah and Mitov (2003), the negative binomial distribution $NB(p_n, r)$ with fixed integer $r \geq 1$, approximately belongs to the max-domain $\mathcal{D}(G_\gamma)$ with $\gamma = 0$ as p_n converges to zero at certain rate. Figures 8 and 9 show that for larger λ and smaller p , the powers are closer to the nominal α . Note the different scales in these figures.

Again, we simulated the power of the tests also with the three $\eta = 0.5, 1$ and 2 and different sample sizes. We found that the power of test T is larger for $\eta = 2$ in the Poisson(10) and Poisson (100) cases. For these two distributions test E has the largest power, always with $\eta = 2$. For the distributions Exam3(0.5, $1/\sqrt{2}$) and Exam3(0.5,0.2) the power of test E was largest with $\eta = 2$, whereas the power of test T with the choices $\eta = 1$ or 2 were rather

similar. For test PE one should select in the last two cases $\eta = 2$, which is rather different to the choice found in the section on the type I error. Details can be found at <http://www.imsv.unibe.ch/~deyuan/research.html>.

5 Conclusion

From Sections 3 to 4 we conclude that

- Among $\{0.5, 1, 2\}$ the good choice of η for the three tests T, E and PE are $\eta = 1$, $\eta = 2$ and $\eta = 0.5$ respectively. The best choice for test E was always $\eta = 2$. But based on the type I errors and the power derivations, for test T $\eta = 1$ or 2 should be selected. For the test PE a different η should be selected for a good power.
- In sense of T R E, test E with $\eta = 2$ is better than test PE with $\eta = 0.5$ even if we know $\gamma \geq 0$, meaning test PE is converging slower to the asymptotic test than test E. Test T with $\eta = 1$ is comparable to test E with $\eta = 2$. Test T seems better than test E for positive γ , otherwise test E seems better. The power is often larger for test E, but also in some cases test T has a larger power.
- The second order condition is not necessary in testing the null hypothesis H_0 . Type I error for the distribution F belonging to the max-domain but not satisfying the second order condition depends on how important its second term is. If the second term is less important, the type I error is closer to the nominal α . The power for the distribution F not belonging to the max-domain depends on how far F is away from the max-domain. If F is farther away from the max-domain, the power is larger. If F is very close to the max-domain, the power is very close to the nominal α , which means in this case we cannot detect that F is not in the max-domain.

For application we suggest that

- (i) Estimate the extreme value index γ based on both the maximum likelihood estimator and the moment estimator.
- (ii) If the extreme value index can be believed to be positive (for example, both estimators of γ are larger than 0.05), then it might be better to use test T with $\eta = 1$ to test H_0 . Otherwise use test E with $\eta = 2$ to test H_0 .
- (iii) Since our intention is to apply the extreme value theory, we want to confirm H_0 . So, none of the tests should indicate a strong deviation from H_0 by a large test statistic.

The R program code for applying these tests to real data is provided at www.imsv.unibe.ch/~deyuan/research.html.

Acknowledgement We are grateful to the referee and the associate editor for their valuable comments. Partially supported by Swiss National Science foundation.

Appendix

The following distribution functions are used in the simulation.

- Cauchy distribution ($\gamma = 1, \rho = -2$):

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \quad x \in \mathbb{R}.$$

- Log-gamma(λ, m) distribution with $\lambda, m > 0$ ($\gamma = 1/\lambda, \rho = 0$):

$$f(x) = \frac{\lambda^m}{\Gamma(m)} (\log x)^m x^{-\lambda-1}, \quad x \geq 1.$$

- Burr(β, τ, λ) distribution with $\beta, \tau, \lambda > 0$ ($\gamma = 1/(\tau\lambda), \rho = -1/\lambda$):

$$F(x) = 1 - \left(\frac{\beta}{\beta + x^\tau} \right)^\lambda, \quad x > 0.$$

- EV(γ) distribution ($\gamma \in \mathbb{R}, \rho = -1$):

$$F(x) = \exp\left(- (1 + \gamma x)^{-1/\gamma}\right), \quad 1 + \gamma x > 0.$$

- Weibull(λ, τ) distribution with $\lambda, \tau > 0$ ($\gamma = 0, \rho = 0$):

$$F(x) = 1 - \exp(-\lambda x^\tau), \quad x > 0.$$

- Reversed Burr(β, τ, λ) distribution ($\gamma = -1/(\tau\lambda), \rho = -1/\lambda$):

$$F(x) = 1 - \left(\frac{\beta}{\beta + (x_+ - x)^{-\tau}} \right)^\lambda, \quad x < x_+,$$

with $\beta, \tau, \lambda > 0$ and $x_+ = 1$.

- Pareto(β) distribution with $\beta > 0$ ($\gamma = 1/\beta$):

$$F(x) = 1 - x^{-\beta}, \quad x \geq 1.$$

- Negative Binomial distribution, $X \sim \text{NB}(p, r)$:

$$P(X = k) = C_{k-1}^{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \dots$$

References

- Anderson, C.W.: Extreme value theory for a class of discrete distributions with application to some stochastic processes. *J. Appl. Probab.* **7**, 99–113 (1970)
- Anderson, C.W., Coles, S., Hüslér, J.: Maxima of Poisson-like variables and related triangular arrays. *Ann. Appl. Probab.* **7**, 953–971 (1997)
- Dekkers, A.L.M., de Haan, L., Einmahl, J.H.J.: A moment estimator for the index of an extreme-value distribution. *Ann. Stat.* **17**, 1833–1855 (1989)
- Dietrich, D., de Haan, L., Hüslér, J.: Testing extreme value conditions. *Extremes* **5**, 71–85 (2002)
- Drees, H., Ferreira, A., de Haan, L.: On the maximum likelihood estimation of the extreme value index. *Ann. Appl. Probab.* **14**, 1179–1201 (2004)
- Drees, H., de Haan, L., Li, D.: Approximations to the tail empirical distribution function with application to testing extreme value conditions. *J. Stat. Plan. Inference* **136**, 3498–3538 (2006)
- Falk, M., Hüslér, J., Reiss, R.D.: *Laws of Small Numbers: Extremes and Rare Events*. Birkhäuser, Switzerland (2004)

- Geluk, J., de Haan, L.: Regular Variation, Extensions and Tauberian Theorems. CWI Tract 40, Amsterdam (1987)
- de Haan, L., Rootzén, H.: On the estimation of high quantiles. *J. Stat. Plan. Inference* **35**, 1–13 (1993)
- de Haan, L., Stadtmüller, U.: Generalized regular variation of second order. *J. Aust. Math. Soc., Ser. A* **61**, 381–395 (1996)
- Hall, P.: On estimating the endpoint of a distribution. *Ann. Stat.* **10**, 556–568 (1982)
- Hill, B.M.: A simple approach to inference about the tail of a distribution. *Ann. Stat.* **3**, 1163–1174 (1975)
- Leadbetter, M.R., Lindgren, G., Rootzén, H.: *Extremes and Related Properties of Random Sequences and Processes*. Springer, Berlin Heidelberg New York (1983)
- Nadarajah, S., Mitov, K.: Asymptotics of maxima of discrete random variables. *Extremes* **5**, 287–294 (2003)
- Pickands, J.: Statistical inference using extreme order statistics. *Ann. Stat.* **3**, 119–131 (1975)
- Smith, R.L.: Maximum likelihood estimation in a class of nonregular cases. *Biometrika* **72**, 67–90 (1985)