Extremes (2008) 11:99–111
DOI 10.1007/s10687-007-0052-0

Review of testing issues in extremes: in honor of Professor Laurens de Haan

Jürg Hüsler · Liang Peng

Received: 21 October 2007 / Revised: 26 October 2007 /
Accepted: 29 October 2007 / Published online: 16 February 2008

Abstract As a leading statistician in extreme value theory, Professor Laurens de Haan has made significant contribution in both probability and statistics of extremes. In honor of his 70th birthday, we review testing issues in extremes, which include research done by Professor Laurens de Haan and many others. In comparison with statistical estimation in extremes, research on testing has received less attention. So we also point out some practical questions in this direction.

Keywords Extremes · Goodness-of-fit test · Domain of attraction · Extreme value distribution

AMS 2000 Subject Classification 60G70 · 62F03 · 62G63

1 Introduction

As a mathematically sound way of extrapolating data, extreme value theory has been applied e.g. to environmental science, computer science, graph theory, economics, insurance, finance, risk management. This is noticeable from the fair amount of monographs given in the reference list (Arnold and Balakrishnan 1989; Balakrishnan and Cohen 1990; Beirlant et al. 2004; Castillo

J. Hüsler (✉)
Department of Mathematical Statistics, University of Bern,
Sidlerstr. 5, CH-3012 Bern, Switzerland
e-mail: huesler@stat.unibe.ch

L. Peng
School of Mathematics, Georgia Institute of Technology,
Atlanta, GA 30332-0160, USA
e-mail: peng@math.gatech.edu

1988; Castillo et al. 2004; David and Nagaraja 2003; Embrechts et al. 1997; Falk et al. 2005; Finkenstadt and Rootzen 2003; Galambos 1987; Galambos et al. 1994; Kotz and Nadarajah 2001; Leadbetter et al. 1983; McCormick and Sun 2008; Reiss 1989; Reiss and Thomas 2007; Resnick 1987; Salvadori et al. 2007; Sivakumar et al. 2005; Tiago de Oliveira and Gomes 1984). Indeed the basic assumption of extreme value theory is on the stability of normalized maxima, that is, suppose there are $a_{n1}, \cdots, a_{nd} > 0$ and $b_{n1}, \cdots, b_{nd} \in R$ such that the normalized maxima, $(\frac{\max_{1\le i\le n} X_{i1}-b_{n1}}{a_{n1}}, \cdots, \frac{\max_{1\le i\le n} X_{id}-b_{nd}}{a_{nd}})^T$, converge in distribution to a non-degenerate distribution G, where $\{X_i = (X_{i1}, \cdots, X_{id})^T\}_{i=1}^n$ is an independent sequence of random vectors with distribution function F, with $d \ge 1$. In this case, we say G is a multivariate extreme value distribution and F is in the domain of attraction of G [notation: $F \in D(G)$]. Under the above stability assumption, researchers have characterized G, derived rates of convergence, found conditions for F to be in the domain of attraction of G, proposed ways to estimate G so as to extrapolate F. In contrast to statistical estimation, the study of testing conditions in extremes has not received enough attention in the literature. Recently, Fraga et al. (2006) gave a brief review on tests in extremes with focus on their recent research. Here we give a more comprehensive review with focus on the significant work done by Professor Laurens de Haan for honoring his 70th birthday. As a leading statistician, Professor Laurens de Haan has made significant contribution to both probability and statistics of extremes, and he continues to contribute to every statistical aspect of extremes.

In this paper, we focus on independent data since the study of tests for dependent extremes is almost empty. It is known that powerful tests do depend on both the null hypothesis and alternative hypothesis. Before reviewing tests for various different hypotheses, we start with a summary of different models employed in analyzing extremes in Section 2. Section 3 reviews all tests for univariate extremes according to different statistical models. In Section 4, we review some tests for multi/bi-variate extremes. We also pose some practical testing questions for both univariate and multivariate extremes in Sections 3 and 4 so as to stimulate more research in this direction.

2 Models for Analyzing Univariate Extremes

Throughout this section we assume that $X_1, \cdots, X_n$ are independent and identically distributed (iid) random variables with distribution function F. We summarize the following four models for analyzing univariate extremes.

Model A When $X_i's$ are annual maxima, one could assume that the distribution F of the X_i's can be approximated by a generalized extreme value distribution, i.e. by

$$F(x) = \exp\left\{-\left(1+\gamma\frac{x-\mu}{\sigma}\right)^{-1/\gamma}\right\}, \tag{2.1}$$

where γ, μ, σ are called shape, location and scale parameters, respectively. In this model we assume that F is such a generalized extreme value distribution. We refer to Gumbel (1958) for details.

Model B Suppose F is in the domain of attraction of an extreme value distribution

$$G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}.$$

Then there exists a positive nondecreasing function f such that

$$\lim_{t\to x^*} P\left(\frac{X_1 - t}{f(t)} > x | X_1 > t\right) = (1+\gamma x)^{-1/\gamma}$$

for all $x \le x^*$ for which $1+\gamma x > 0$, where $x^* = \sup\{x : F(x) < 1\}$. Therefore the distribution of an exceedance over a high threshold can be approximated by the so-called generalized Pareto distribution

$$H(x) = 1 - (1+\gamma x/\sigma)^{-1/\gamma}, \tag{2.2}$$

see Chapter 3.1 of de Haan and Ferreira (2006). Thus, Model B assumes that exceedances over a fixed high threshold follow exactly the generalized Pareto distribution. We refer to Coles (2001) for details.

Model C This model assumes that F is in the domain of attraction of an extreme value distribution. Thus, unlike the parametric models in Models A and B, Model C may be called a semi-parametric model. We refer to de Haan and Ferreira (2006) for details.

Model D This model allows that parameters in extremes depend on other covariates such as e.g. time and/or space. One reason for the attractiveness of this method is that this model relaxes the assumption of identical distribution. For example, (1) fit a generalized extreme value distribution in Eq. 2.1 to annual maximum temperatures, but allow parameters γ, μ, σ are functions of time; (2) given a covariate Z, assume that the conditional distribution $P(X_1 \le x|Z = z)$ satisfies

$$P(X_1 > x|Z = z) = a(z,x)\{g(z) - x\}^{\alpha(z)}\{1 + o(1)\}$$

as $x \uparrow g(z)$, where $a(z,x) > 0$ and $\alpha(z) > 0$. Under this setup, function g is called a frontier function or boundary. Estimating the unknown function g plays a role in productivity study. Some references on Method D include Beirlant and Goegebeur (2004), Chavez-Demoulin and Embrechts (2004), Chavez-Demoulin and Davison (2005), Cheng and Peng (2007), Coles and Tawn (1990), Davison and Ramesh (2000), Davison and Smith (1990), Dixon and Tawn (1999), Gijbels and Peng (2000), Hall et al. (1997), Hall and Tajvidi (2000), Hall and Van Keilegom (2006), Ramesh and Davison (2002), and Peng (2004).

3 Tests for Univariate Extremes

Suppose we have iid observations $X_1, \cdots, X_n$ with distribution function F. We review tests according to Models A, B, and C of Section 2. There is almost no study on tests for Model D in the literature.

3.1 Tests for Model A

Throughout this subsection, we assume that F is a generalized extreme value distribution, i.e., Eq. 2.1 holds.

A1) Due to the connection between $\gamma = 0$ and the Gumbel distribution, researchers have paid much attention on testing H_0 : Eq. 2.1 holds with $\gamma = 0$ against H_a : any other distributions; see Cabana and Quiroz (2005), Hassanein et al. (1986), Janic-Wroblewska (2004), Lawless (1978), Liao and Shimokawa (1999), Lockhart et al. (1986), Mann et al. (1973), Öztürk (1986), Öztürk and Korukoglu (1988), Shi (1988) and Stephens (1977). Note that the null hypothesis is a parametric model (i.e., the support of F is independent of parameters). So, all standard tests can be employed here.

Open Question on the Case A1) Recently, empirical likelihood method has been applied to conduct a likelihood ratio test under a nonparametric or semiparametric setup. A nice review on empirical likelihood method can be found in Owen (2001). Some recent work on empirical likelihood ratio test includes Cao and Van Keilegom (2006), Chen and Gao (2007), Chen et al. (2003), Chen and Van Keilegom (2006), Einmahl and McKeague (2003), Li and Van Keilegom (2002). As a powerful test, it would be interesting to see how empirical likelihood ratio test can be employed in extremes.

A2) Another question is how to select a particular subclass from a generalized extreme value distribution. That is, how to test (1) $H_0 : \gamma = 0$ against $H_a : \gamma \neq 0$, or against $H_a : \gamma > 0$, or against $H_a : \gamma < 0$; (2) $H_0 : \gamma \geq 0$ against $H_a : \gamma < 0$. The study on case (1) includes Gomes (1989), Gomes and Teresa (1986), Hosking (1984), Tiago de Oliveira and Gomes (1984), Wang et al. (1996). For case (2), we refer to Marohn (2000). Although the model is parametric under both the null and alternative hypotheses, the alternative hypothesis is irregular (i.e., the support of F depends on parameters). Hence a parametric likelihood ratio test has to be restricted to the case $\gamma > -1/2$.

A3) The question here is how to test that two extreme value distributions have some equal parameters; see Hassanein and Saleh (1992), Lawless and Mann (1976). In general, two sample issues receive almost no attention in extremes. This may be due to the lack of practical motivations.

We mention that the application of these methods for block-maxima X_j, e.g. yearly maxima as mentioned, depends on the block length.

3.2 Tests on Model B

This model goes as follows: pick up a high threshold u, consider those $X_i's$ above u, say $Y_1, \cdots, Y_m$, and model the distribution of the exceedances $Y_j - u$ by a generalized Pareto distribution in Eq. 2.2. Under this setup, one wants to test whether the exceedances $Y_i - u$ have a generalized Pareto distribution, see Beisel et al. (2007), Castillo and Hadi (1997), Choulakian and Stephens (2001), Falk (1995), Marohn (2000, 2002). A serious drawback of this setup is to determine how the threshold affects the limit distribution of the test statistics, i.e., the critical values. As before, it would be interesting to develop an empirical likelihood ratio test for this model and compare this new test with others.

3.3 Tests on Model C

Throughout this subsection we assume that F is in the domain of attraction of an extreme value distribution $G_\gamma(x) = \exp\{-(1+\gamma x)^{-1/\gamma}\}$. Let $\Omega_1 \cup \Omega_2 = R$ and $\Omega_1 \cap \Omega_2 = \emptyset$.

C1) The first question is how to test $H_0 : F \in D(G_\gamma), \gamma \in \Omega_1$ against $H_a : F \in D(G_\gamma), \gamma \in \Omega_2$. This was investigated e.g. in see Fraga et al. (1996), Hasofer and Li (1999), Hasofer and Wang (1992), Marohn (1998a, b), Segers and Teugels (2000), Neves et al. (2006), Neves and Fraga Alves (2007).

Open Question on the Case C1) Motivated by the maximum likelihood estimation in Drees et al. (2004), it is natural to seek the parametric likelihood ratio test for testing $H_0 : \gamma \in \Omega_1$ against $H_a : \gamma \in \Omega_2$, where now $\Omega_1 \cup \Omega_2 = \{x : x > -1/2\}$.

C2) The second question is how to test (1) $H_0 : F \in D(G_\gamma)$ for some γ against $H_a : F \notin D(G_\gamma)$ for any γ; or (2) $H_0 : F \in D(G_\gamma), \gamma \geq 0$ against $H_a : F \notin D(G_\gamma), \gamma \geq 0$. There are two ways in conducting such tests: using tail quantile processes (Dietrich et al. 2002) or tail empirical processes (Drees et al. 2006). Main theoretical techniques in deriving asymptotic limits under the null hypothesis are weighted approximations of the tail quantile processes and the tail empirical processes, respectively. Let's summarize the results in Dietrich et al. (2002) and Drees et al. (2006) before posting some practical questions.

Define for $k < n$ and $j = 1, 2$:

$$M_{k,n}^{(j)} = \frac{1}{k}\sum_{i=0}^{k}\{\log X_{n,n-i} - \log X_{n,n-k}\}^j,$$

$$\hat{\gamma}_+ = M_{k,n}^{(1)}, \quad \hat{\gamma}_- = 1 - \frac{1}{2}\left\{1 - \left(M_{k,n}^{(1)}\right)^2 / M_{k,n}^{(2)}\right\}^{-1}.$$

Dietrich et al. (2002) considered the test statistics

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

and

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

with $\eta = 2$. The test statistics were generalized by Hüsler and Li (2006) by replacing the particular weight t^2 by the more general weight with $\eta > 0$, and correcting a sign error.

For the following results a second order condition is assumed for the function $\log U$ where $U := \left(\frac{1}{1-F}\right)^{\leftarrow}$. Suppose that for some function A of constant sign and converging to zero

$$\lim_{t\to\infty} \frac{\frac{\log U(tx) - \log U(t)}{a(t)/U(t)} - \frac{x^{\gamma_-} - 1}{\gamma_-}}{A(t)} = \frac{1}{\rho}\left(\frac{x^{\gamma_- + \rho} - 1}{\gamma_- + \rho} - \frac{x^{\gamma_-} - 1}{\gamma_-} \right) \tag{3.1}$$

for all $x > 0$ where $\rho \le 0$ and $\gamma_- := \min(\gamma, 0)$.

Theorem 1 of Dietrich et al. (2002), Husler and Li (2006) Assume that Eq. 3.1 holds. If $k = k(n) \to \infty$, $\sqrt{k}A(n/k) \to 0$ as $n \to \infty$, then for $n \to \infty$

$$kE_{k,n} \xrightarrow{d} \int_0^1 \left\{ (1-\gamma_-)(t^{-\gamma_- - 1}W(t) - W(1)) - (1-\gamma_-)^2 \frac{t^{-\gamma_-} - 1}{\gamma_-} P \right.$$
$$\left. + \frac{t^{-\gamma_-} - 1}{\gamma_-} R + (1-\gamma_-) R \int_t^1 s^{-\gamma_- - 1} \log s \, ds \right\}^2 t^\eta \, dt,$$

where

$$P = \int_0^1 (t^{-\gamma_- - 1} W(t) - W(1)) \, dt,$$

$$Q = 2 \int_0^1 \frac{t^{-\gamma_-} - 1}{\gamma_-} (t^{-\gamma_- - 1} W(t) - W(1)) \, dt,$$

$$R = (1-\gamma_-)^2 (1 - 2\gamma_-)\{(1/2 - \gamma_-)Q - 2P\},$$

and $W(\cdot)$ is a Brownian motion. Moreover for $\gamma \ge 0$

$$kT_{k,n} \xrightarrow{d} \int_0^1 \left\{ B(t) + t \log t \int_0^1 s^{-1} B(s) \, ds \right\}^2 dt,$$

where $B(\cdot)$ is a Brownian bridge.

Note that the limit of $kT_{k,n}$ is independent of the unknown parameters. Thus the critical values can be simulated for both tests. The correct critical values are given in Hüsler and Li (2006). However, for testing $H_0 : F \in D(G_\gamma)$ for some γ, against $H_a : F \notin D(G_\gamma)$ for any γ, the limit of $kE_{k,n}$ depends on the unknown parameter γ_-. An algorithm for applications is given in Hüsler and Li (2006) with a recommendation on choosing η.

$F \in D(G_\gamma)$ means that $t\bar{F}(a(t)x + b(t)) \to (1 + \gamma x)^{-1/\gamma}$ as $t \to \infty$, for some functions $a(t)(> 0)$ and $b(t)$, for all x with $1 + \gamma x > 0$, where $\bar{F} = 1 - F$. Thus Drees et al. (2006) considered the test statistics

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

where $\hat{\gamma}, \hat{a}, \hat{b}$ are suitable estimators for the shape γ, for the scale function $a(\cdot)$ and the location function $b(\cdot)$ with certain properties. For instance, ML-estimates can be used. They derived the asymptotic distribution of T_n.

Theorem 2.2 of Drees et al. (2006) Assume the second order condition (3.1) with the function A. If $k \to \infty$ and $\sqrt{k}A(n/k) \to 0$, then

$$k_n T_n - \int_0^1 \{W_n(x) + L_n(x)\}^2 x^{\eta-2}\, dx \stackrel{p}{\to} 0$$

for all $\eta > 0$ when $\gamma \neq 0$ or $\rho < 0$, and all $\eta \geq 1$ when $\gamma = \rho = 0$, where $W_n(\cdot)$ is a Brownian motion and L_n depends on the estimators $\hat{\gamma}, \hat{a}, \hat{b}$.

Note that the above limit depends on the unknown parameter γ. Critical values for different γ using the maximum likelihood estimators $\hat{a}(\cdot)$ and $\hat{b}(\cdot)$, are simulated in Table 1 of Drees et al. (2006) with $\eta = 1$. As stated in Remark 2.2 and the paragraph right after Remark 2.5 of Drees et al. (2006), this test requires $\gamma > -1/2$. Different η's were considered in Hüsler and Li (2006), with recommendations on choosing η, and the three statistics were compared. Some further comments and applications are given in Hüsler and Li (2007a).

Open Question on the Case C2) Motivated by the maximum likelihood estimation in Drees et al. (2004) and the advantage of empirical likelihood ratio test, it would be interesting to see how empirical likelihood ratio test can be employed to test $H_0 : F \in D(G_\gamma)$ with $\gamma > -1/2$ against H_a : any other distributions.

C3) Recently, Jureckova (2003), Beirlant et al. (2006) and Koning and Peng (2007) studied how to test $H_0 : F \in D(G_\gamma)$ with $\gamma > 0$ against H_a : any other distributions. Indeed, Koning and Peng (2007) compared several tests via Bahadur efficiency and demonstrated that the score test is most efficient both in terms of the Bahadur efficiency and the empirical power. Another test, related with testing of heavy tails, is given by Jureckova and Picek (2001) for testing $H_0 : x^{\alpha_0}(1 - F(x)) \geq 1$ for all $x \geq x_0$ with some $x_0 \geq 0$ and a given $\alpha_0 > 0$ against

$$H_a : \limsup_{x\to\infty} x^{\alpha_0}(1 - F(x)) < 1.$$

For testing heavy tails under the setup of linear model or AR model, we refer to Jureckova (2000) and Jureckova et al. (2007).

Open Question on the Case C3) It is known that Berk–Jones test (Berk and Jones 1979) is an empirical likelihood ratio test for testing $H_0 : F = F_0$ against $H_a : F \neq F_0$, where F_0 is a given distribution. However, the Berk–Jones test studied in Koning and Peng (2007) is a type of conditional empirical likelihood ratio test, which is less powerful than the score test. It would be interesting to derive the empirical likelihood ratio test, and comparing it with the score test in Koning and Peng (2007).

A Common Question on Tests for Model C The distributions of the test statistics do depend on the threshold even when the asymptotic bias is negligible. Is there any theoretical optimal threshold (or k) for such tests and how could such an optimality be achieved when n is finite?

4 Tests for Multivariate Extremes

Throughout this section we focus on the bivariate case and assume that we have independent observations $(X_1, Y_1), \cdots, (X_n, Y_n)$ with a continuous distribution function F. Let $F_1(x) = F(x, \infty)$ and $F_2(y) = F(\infty, y)$. The tail dependence function and tail copula of F are defined as

$$l(x, y) = \lim_{t\to 0} t^{-1} P(F_1(X_1) > 1 - tx \quad \text{or} \quad F_2(Y_1) > 1 - ty)$$

and

$$r(x, y) = \lim_{t\to 0} t^{-1} P(F_1(X_1) > 1 - tx, F_2(Y_1) > 1 - ty),$$

respectively. Hence $r(x, y) = x + y - l(x, y)$.

Since fitting a parametric class to $l(x, y)$ or $r(x, y)$ is popular in parametric statistics, one question is the goodness-of-fit test. Recently, de Haan et al. (2007) first derived the asymptotic limit of pseudo maximum likelihood estimation with random thresholds and then proposed a goodness-of-fit test. The main technique is the weighted approximation of the tail copula processes derived in Einmahl et al. (2006). Since the limit depends on the unknown function l (or r), a naive parametric bootstrap method was proposed in de Haan et al. (2007) to obtain the critical values. Here, an open question is again how to develop an empirical likelihood ratio test.

The second question is how to test whether F is in the domain of attraction of a bivariate extreme value distribution. This has been investigated by Einmahl et al. (2006). The idea is as follows.

Let R_i^X denote the rank of X_i among $X_1, \cdots, X_n$ and R_i^Y denote the rank of Y_i among $Y_1, \cdots, Y_n$. Define

$$\hat{\Phi}(\theta) = \frac{1}{k}\sum_{i=1}^{n} I\left(R_i^X \vee R_i^Y \geq n+1-k, n+1-R_i^Y \leq \left(n+1-R_i^X\right)\tan\theta\right),$$

$$\hat{l}_1(x, y) = \int_0^{\pi/2} \{x(1 \wedge \tan\theta)\} \vee \{y(1 \wedge \cot\theta)\}\, \hat{\Phi}(d\theta),$$

$$\hat{l}_2(x, y) = \frac{1}{k}\sum_{i=1}^{n} I\left(R_i^X > n+1-kx \quad \text{or} \quad R_i^Y > n+1-ky\right).$$

Einmahl et al. (2006) considered the test statistics

$$L_n = \int\int_{0<x,y<1} \left\{\hat{l}_1(x, y) - \hat{l}_2(x, y)\right\}^2 (x \vee y)^{-\beta}\, dxdy,$$

and showed that

Theorem 2.3 of Einmahl et al. (2006) Under some regularity conditions,

$$kL_n \xrightarrow{d} \int\int_{0<x,y<1} \{A(x, y) + B(x, y)\}^2 (x \vee y)^{-\beta}\, dxdy,$$

where $\beta \in [0, 3)$, and A and B depend on the tail dependence function and its partial derivatives.

In view of the complications of the above limit, Einmahl et al. (2006) also provided a way to approximate the above limit, which can be employed to obtain critical values. One of the main techniques of the proofs is the weighted approximation of the tail copula processes.

The third question is how to test the independence among marginals when observations follow a bivariate extreme value distribution. See Deheuvels and Martynov (1996) for an answer to this question.

The fourth question is how to test asymptotic independence when observations follow a multivariate extreme value distribution, see Falk and Michel (2006).

Recently, Ramos and Ledford (2005) studied tests for the asymptotic independence when a parametric class is fitted to the region where both variables are above fixed thresholds. Therefore, this approach tests the independence for a parametric model with left censoring.

The last tricky, but important issue is how to test asymptotic independence when we assume that F is in the domain of attraction of a multivariate extreme value distribution. In this case, the coefficient of tail dependence was introduced to classify asymptotic independence and estimators were proposed, see e.g. Draisma et al. (2004), Ledford and Tawn (1996, 2003), Peng (1999). Recently, Hüsler and Li (2007b) provided a way for testing the asymptotic independence when F is in the domain of attraction of a bivariate extreme value distribution. Some missing issues in multivariate extremes are how to compare tests theoretically and how to select threshold.

References

Arnold, B.C., Balakrishnan, N.: Relations, Bounds, and Approximations for Order Statistics. Springer (1989)
Balakrishnan, N., Cohen, A.: Order Statistics and Inference: Estimation Methods. Academic Press (1990)
Beirlant, J., de Wet, T., Goegebeur, Y.: A goodness-of-fit statistic for Pareto-type behaviour. J. Comput. Appl. Math. **186**, 99–116 (2006)
Beirlant, J., Goegebeur, Y.: Local polynomial maximum likelihood estimation for Pareto-type distributions. J. Multi. Anal. **89**, 97–118 (2004)
Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.: Statistics of Extremes: Theory and Applications. Wiley (2004)
Beisel, C.J., Rokyta, D.R., Wichman, H.A., Joyce, P.: Testing the extreme value domain of attraction for distributions of beneficial fitness effects. Genetics **176**, 2441–2449 (2007)
Berk, R.H., Jones, D.H.: Goodness-of-fit-statistics that dominates the Kolmogorov statistics. Z. Wahrscheinlichkeitstheorie Verw. Geb. **47**, 47–59 (1979)
Berman, S.: Sojourns and Extremes of Stochastic Processes. Chapman & Hall/CRC (1992)
Cabana, A., Quiroz, A.J.: Using the empirical moment generating function in testing for the Weibull and the type I extreme value distributions. Test **14**, 417–431 (2005)
Cao, R., Van Keilegom, I.: Empirical likelihood tests for two-sample problems via nonparametric density estimation. Canad. J. Statist. **34**, 61–77 (2006)
Castillo, E.: Extreme Value Theory in Engineering. Academic Press. (1988)
Castillo, E., Hadi, A.S.: Fitting the generalized Pareto distribution to data. J. Amer. Statist. Assoc. **92**(440), 1609–1620 (1997)
Castillo, E., Hadi, A.S., Balakrishnan, N., Sarabia, J.M.: Extreme Value and Related Models with Applications in Engineering and Science. Wiley-Interscience (2004)
Chavez-Demoulin, V., Davison, A.C.: Generalized additive modelling of sample extremes. J. Roy. Statist. Soc. Ser. C **54**, 207–222 (2005)
Chavez-Demoulin, V., Embrechts, P.: Smooth extremal models in finance. J. Risk Insurance **71**, 183–199 (2004)
Chen, S.X., Gao, J.: An adaptive empirical likelihood test for time series models. J. Econom. **141**, 950–972 (2007)
Chen, S.X., Härdle, W., Li, M.: An empirical likelihood goodness-of-fit test for time series. J. Roy. Statist. Soc. Ser. B **65**, 663–678 (2003)
Chen, S.X., Van Keilegom, I.: A goodness-of-fit test for parametric and semiparametric models in multiresponse regression. Technical report (2006)
Cheng, M., Peng, L.: Variance reduction in multivariate likelihood models. J. Amer. Statist. Assoc. **102**(477), 293–304 (2007)
Choulakian, V., Stephens, M.A.: Goodness-of-fit tests for the generalized Pareto distribution. Technometrics **43**(4), 478–484 (2001)
Coles, S.: An Introduction to Statistical Modeling of Extreme Values. Springer (2001)
Coles, S.G., Tawn, J.A.: Statistics of coastal flood prevention. Phil. Trans. Royal Soc. London A **332**, 457–476 (1990)
David, H.A., Nagaraja, H.N.: Order Statistics. Wiley-Interscience, 3rd edn. (2003)
Davison, A.C., Ramesh, N.I.: Local likelihood smoothing of sample extremes. J. Roy. Statist. Soc. Ser. B **62**, 191–208 (2000)
Davison, A.C., Smith, R.L.: Models for exceedances over high thresholds (with discussion). J. Roy. Statist. Soc. Ser. B **52**, 393–442 (1990)
de Haan, L., Ferreira, A.: Extreme Value Theory: An Introduction. Springer (2006)
de Haan, L., Peng, L., Neves, C.: Parametric tail copula estimation and model testing. Technical report (2007)
Deheuvels, P., Martynov, G.V.: Cramer-von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions. Comm. Statist. Theory Methods **25**(4), 871–908 (1996)
Dietrich, D., de Haan, L., Hüsler, J.: Testing extreme value conditions. Extremes **5**, 71–85 (2002)
Dixon, M.J., Tawn, J.A.: The effect of non-stationarity on extreme sea-level estimation. Appl. Statist. **48**, 135–151 (1999)

Draisma, G., Drees, H., Ferreira, A., de Haan, L.: Bivariate tail estimation: dependence in asymptotic independence. Bernoulli **10**, 251–280 (2004)
Drees, H., Ferreira, A., de Haan, L.: On maximum likelihood estimation of the extreme value index. Ann. Appl. Prob. **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. Statist. Plann. Inf. **136**, 3498–3538 (2006)
Einmahl, J.H.J., de Haan, L., Li, D.: Weighted approximations of tail copula processes with application to testing the bivariate extreme value condition. Ann. Statist. **34**, 1987–2014 (2006)
Einmahl, J.H.J., McKeague, I.W.: Empirical likelihood based hypothesis testing. Bernoulli **9**, 267–290 (2003)
Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling Extremal Events for Insurance and Finance. Springer (1997)
Falk, M.: On testing the extreme value index via the POT-method. Ann. Statist. **23**, 2013–2035 (1995)
Falk, M., Hüsler, J., Reiss, R.D.: Laws of Small Numbers: Extremes and Rare Events. Birkhauser Basel, 2nd edn. (2005)
Falk, M., Michel, R.: Testing for tail independence in extreme value models. Ann. Inst. Statist. Math. **58**, 261–290 (2006)
Finkenstadt, B., Rootzen, H.: Extreme Values in Finance, Telecommunication and the Environment. Chapman & Hall (2003)
Fraga Alves, M.I., Gomes, M.I.: Statistical choice of extreme value domain of attraction—a comparative analysis. Comm. Statist.- Theory Methods **25**, 789–811 (1996)
Fraga Alves, M.I., Neves, C.: Testing extreme value conditions – an overview and recent approaches. Technical report (2006)
Galambos, J.: The Asymptotic Theory of Extreme Order Statistics. Krieger Pub. Co., 2nd ed. (1987)
Galambos, J., Lechner, J., Simiu, E.: Extreme Value Theory and Applications. Springer (1994)
Gijbels, I., Peng, L.: Estimation of a support curve via order statistics. Extremes **3**, 251–277 (2000)
Gomes, M.I.: Generalized Gumbel and likelihood ratio test statistics in the multivariate GEV model. Comput. Stat. Data Anal. **7**, 259–267 (1989)
Gomes, M.I., Teresa, A.M.: Inference in a multivariate generalized extreme value model—asymptotic properties of two test statistics. Scand. J. Statist. **13**, 291–300 (1986)
Gumbel, E.J.: Statistics of Extremes. Columbia University Press, New York (1958)
Hall, P., Nussbaum, M., Stern, S.E.: On the estimation of a support of indeterminate sharpness. J. Multi. Anal. **62**, 304–232 (1997)
Hall, P., Tajvidi, N.: Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. Statist. Sci. **15**, 153–167 (2000)
Hall, P., Van Keilegom, I.: Nonparametric "regression" when errors are centred at endpoints. Technical report (2006)
Hasofer, A.M., Li, S.: Estimation for type II domain of attraction based on the W statistic. Austral. & New Zealand J. Statist. **41**(2), 223–232 (1999)
Hasofer, A.M., Wang, Z.: A test for extreme value domain of attraction. J. Amer. Statist. Assoc. **87**(417), 171–177 (1992)
Hassanein, K.M., Saleh, A.K.: Testing equality of locations and quantiles of several extreme-value distributions by use of few order statistics of samples from extreme-value and Weibull distributions. In: Order Statistics and Nonparametrics: Theory and Applications, pp. 115–132. North-Holland, Amsterdam (1992)
Hassanein, K.M., Saleh, A.K., Brown, E.F.: Estimation and testing of quantiles of the extreme-value distribution. J. Statist. Plann. Inf. **14**, 389–400 (1986)
Hosking, J.R.M.: Testing whether the shape parameter is zero in the generalized extreme-value distribution. Biometrika **71**, 367–374 (1984)
Hüsler, J., Li, D.: On testing extreme value conditions. Extremes **9**, 69–86 (2006)
Hüsler, J., Li, D.: Testing extreme value conditions with applications. In Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields. R.-D. Reiss, Thomas, M. (2007). Birkhäuser Boston. 3rd ed. pp. 144–151 (2007a)
Hüsler, J., Li, D.: Testing asymptotic independence in bivariate extremes. Technical report (2007b)
Janic-Wroblewska, A.: Data-driven smooth tests for the extreme value distribution. Statistics **38**, 413–426 (2004)

Jureckova, J.: Test of tails based on extreme regression quantiles. Stat. Probab. Lett. **49**, 53–61 (2000)
Jureckova, J.: Statistical tests on tail index of a probability distribution. Metron **61**, 151–175 (2003)
Jureckova, J., Koul, H.L., Picek, J.: Testing the tail index in autoregressive models. Ann. Inst. Statist. Math. (2007) (to appear)
Jureckova, J., Picek, J.: A class of tests on the tail index. Extremes **4**, 165–183 (2001)
Koning, A., Peng, L.: Goodness-of-fit tests for a heavy tailed distribution. Technical report (2007)
Kotz, S., Nadarajah, S.: Extreme Value Distributions: Theory and Applications. World Scientific Publishing Company (2001)
Lawless, J.F.: Confidence interval estimation for the Weibull and extreme value distributions. Technometrics **20**, 355–365 (1978)
Lawless, J.F., Mann, N.R.: Tests for homogeneity of extreme value scale parameters. Comm. Statist.—Theory Methods A **5**, 389–405 (1976)
Leadbetter, M.R., Lindgren, G., Rootzen, H.: Extremes and related properties of stationary sequences and processes. Springer, New York (1983)
Ledford, A.W., Tawn, J.A.: Statistics for near independence in multivariate extreme values. Biometrika **83**, 169–187 (1996)
Ledford, A.W., Tawn, J.A.: Diagnostics for dependence within time series extremes. J. R. Statist. Soc. B **65**, 521–543 (2003)
Li, G., Van Keilegom, I.: Likelihood ratio confidence bands in nonparametric regression with censored data. Scand. J. Statist. **29**, 547–562 (2002)
Liao, M., Shimokawa, T.: A new goodness-of-fit test for type-I extreme-value and 2-parameter Weibull distributions with estimated parameters. J. Statist. Comput. Simul. **64**, 23–48 (1999)
Lockhart, R.A., O'Reilly, F., Stephens, M.A.: Tests for the extreme value and Weibull distributions based on normalized spacings. Naval Res. Logist. Quart. **33**, 413–421 (1986)
Mann, N.R., Scheuer, E.M., Fertig, K.W.: A new goodness-of-fit test for the two-parameter Weibull or extreme-value distribution with unknown parameters. Comm. Statist. **2**, 383–400 (1973)
Marohn, F.: Testing the Gumbel hypothesis via the POT-method. Extremes **1**, 191–213 (1998a)
Marohn, F.: An adaptive efficient test for Gumbel domain of attraction. Scand. J. Statist. **25**, 311–324 (1998b)
Marohn, F.: Testing extreme value models. Extremes **3**, 363–384 (2000)
Marohn, F.: A characterization of generalized Pareto distributions by progressive censoring schemes and goodness-of-fit tests. Commun. Statist.—Theory Meth. **31**(7), 1055–1065 (2002)
McCormick, W.P., Sun, J.: Extreme Value Theory with S Programming. Chapman & Hall/CRC (2008)
Neves, C., Fraga Alves, M.I.: Semi-parametric approach to Hasofer-Wang and Greenwood statistics in extremes. Test **16**, 297–313 (2007)
Neves, C., Picek, J., Fraga Alves, M.I.: The contribution of the maximum to the sum of excesses for testing max-domains of attraction. J. Statist. Plann. Inf. **136**, 1281–1301 (2006)
Owen, A.B.: Empirical Likelihood. Chapman & Hall/CRC (2001)
Öztürk, A.: On the W test for the extreme value distribution. Biometrika **73**, 738–740 (1986)
Öztürk, A., Korukoglu, S.: A new test for the extreme value distribution. Comm. Statist. Simulation Comput. **17**, 1375–1393 (1988)
Peng, L.: Estimation of the coefficient of tail dependence in bivariate extremes. Stat. Probab. Lett. **43**, 399–409 (1999)
Peng, L.: Bias-corrected estimators for monotone and concave frontier functions. J. Statist. Plann. Inf. **119**, 263–275 (2004)
Ramesh, N.I., Davison, A.C.: Local models for exploratory analysis of hydrological extremes. J. Hydrology **256**, 106–119 (2002)
Ramos, A., Ledford, A.: Regular score tests of independence in multivariate extreme values. Extremes **8**, 5–26 (2005)
Reiss, R.D.: Approximate Distributions of Order Statistics. Springer (1989)
Reiss, R.D., Thomas, M.: Statistical Analysis of Extreme Values: with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhauser Boston. 3rd edn. (2007)
Resnick, S.I.: Extreme Values, Regular Variation and Point Processes. Springer-Verlag (1987)
Salvadori, G., de Michele, C., Kottegoda, N.T., Rosso, R.: Extremes in Nature: An Approach Using Copulas. Springer (2007)

Segers, J., Teugels, J.: Testing the Gumbel hypothesis by Galton's ratio. Extremes **3**, 291–303 (2000)
Shi, D.J.: An analysis of variance test for the extreme value distribution. J. Tianjin Univ. **2**, 116–121 (1988)
Sivakumar, M.V.K., Motha, R.P., Das, H.P.: Natural Disasters and Extreme Events in Agriculture: Impacts and Mitigation. Springer (2005)
Stephens, M.A.: Goodness of fit for the extreme value distribution. Biometrika **64**, 583–588 (1977)
Tiago de Oliveira, J.: Statistical Extremes and Applications. Springer (1984)
Tiago de Oliveira, J., Gomes, M.I.: Two test statistics for choice of univariate extreme models. In: Tiago de Oliveira, J. (ed.) Statistical Extremes and Applications, D. Reidel, Dordrecht, pp. 651–668 (1984)
Wang, J.Z., Cooke, P., Li, S.: Determination of domains of attraction based on a sequence of maxima. Austral. J. Statist. **38**, 173–181 (1996)