Testing Exponentiality Versus Pareto Distribution via Likelihood Ratio

Tomasz J. Kozubowski, Anna K. Panorska, Fares Qeadan, Alexander Gershunov, Debra Rominger

Department of Mathematics and Statistics, University of Nevada, Reno, Nevada, USA
Climate Research Division, Scripps Institution of Oceanography, San Diego, California, USA
Game Design, International Game Technology, Reno, Nevada, USA

Online Publication Date: 01 January 2009

To cite this Article

To link to this Article: DOI: 10.1080/03610910802439121
URL: http://dx.doi.org/10.1080/03610910802439121

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Testing Exponentiality Versus Pareto Distribution via Likelihood Ratio

TOMASZ J. KOZUBOWSKI1, ANNA K. PANORSKA1, FARES QEADAN1, ALEXANDER GERSHUNOV2, AND DEBRA ROMINGER3

1Department of Mathematics and Statistics, University of Nevada, Reno, Nevada, USA
2Climate Research Division, Scripps Institution of Oceanography, San Diego, California, USA
3Game Design, International Game Technology, Reno, Nevada, USA

We consider the problem of maximum likelihood estimation of the parameters of the Pareto Type II (Lomax) distribution. We show that in certain parametrization and after modification of the parameter space to include exponential distribution as a special case, the MLEs of parameters always exist. Moreover, the MLEs have a non standard asymptotic distribution in the exponential case due to the lack of regularity. Further, we develop a likelihood ratio test for exponentiality versus Pareto II distribution. We emphasize that this problem is non standard, and the limiting null distribution of the deviance statistic is not chi-square. We derive relevant asymptotic theory as well as a convenient computational formula for the critical values for the test. An empirical power study and power comparisons with other tests are also provided. A problem from climatology involving precipitation data from hundreds of meteorological stations across North America provides a motivation for and an illustration of the new test.

Keywords Generalized Pareto distribution; Heavy-tails; Likelihood ratio test; Lomax distribution; Maximum likelihood estimation; Non regular case; Pareto type II; Peaks-over-threshold; Test for exponentiality.

Mathematics Subject Classification Primary 62F03; Secondary 62F10, 62F12, 62G32, 62P12, 62C05.

Received September 18, 2007; Accepted August 27, 2008
Address correspondence to Tomasz J. Kozubowski, Department of Mathematics and Statistics, University of Nevada, Reno, NV 89557, USA; E-mail: tkozubow@unr.edu

118
1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample from the classical Pareto distribution, given by the survival function

$$S(x) = \left( \frac{\sigma}{x} \right)^\alpha, \quad x \geq \sigma. \quad (1)$$

It is well known that the maximum likelihood estimators (MLEs) of the shape parameter $\alpha > 0$ and the scale parameter $\sigma > 0$ are uniquely available in explicit forms (see, e.g., Arnold, 1983, p. 194). On the other hand, this is no longer so if the sample is from the Pareto II distribution (see, e.g., Arnold, 1983, p. 13) given by the survival function

$$S(x) = \left( \frac{1}{1 + x/\sigma} \right)^\alpha, \quad x \geq 0, \quad (2)$$

also known as the Lomax distribution (see Johnson et al., 1994, p. 575). Here, the likelihood equations must be solved numerically (see, e.g., Silcock, 1954), and, as commented by Arnold (1983, p. 209), “it is, unfortunately, certainly possible to encounter a data set for which no solution exists.” We revisit this problem in Sec. 2. We start with a striking illustration when the sample consists of only one observation $X > 0$ from (2). We show that in this case the MLEs of $\alpha$ and $\sigma$ never exist. Instead, the likelihood converges to its maximum value of $1/(eX)$ as $\alpha$ and $\sigma$ converge to infinity while $\sigma/\alpha$ converges to $X$. Incidentally, this maximal likelihood coincides with that of an exponential model given by the survival function

$$S(x) = e^{-x/s}, \quad x \geq 0, \quad (3)$$

when the scale parameter $s > 0$ is estimated from one observation $X$. Note that the exponential model can be considered as the limiting case of (2) when $\alpha$ and $\sigma$ converge to infinity while $\sigma/\alpha$ converges to $s > 0$. Thus, when we re-parameterize with $\omega = 1/\alpha$, $s = \sigma/\alpha$, and combine (2) and (3) into a single model given by the survival function

$$S(x) = \left( \frac{1}{1 + \omega x} \right)^{1/\omega}, \quad x \geq 0, \quad (4)$$

then the MLEs of $\omega$ and $s$ always exist (with the understanding that $\omega = 0$ corresponds to the limiting exponential (3) as $\omega \to 0$ in (4)). Let us note that in this parameterization, and with an additional case $\omega < 0$ (for which the support of the distribution is the set $0 < x < -s/\omega$), the distribution with the survival function (4) is known as the generalized Pareto distribution (GPD). Since its introduction in Balkema and de Haan (1974) and Pickands (1975) as a model for excesses over a high threshold, GPD played an important role in applications in connection with peaks over threshold (POT) modeling (see, e.g., Davison and Smith, 1990, and references therein).

The remainder of Sec. 2 is devoted to maximum likelihood estimation of the parameters $\omega \in [0, \infty)$ and $s \in (0, \infty)$ of (4). In particular, we show that with the addition of $\omega = 0$ (exponential distribution) to the parameter space of the Pareto
distribution, the MLEs $\hat{\omega}_n$ and $\hat{s}_n$ always exist for any sample size $n$. If the sample from the exponential distribution has size greater than one, then the estimate of $\omega$ may or may not equal zero, depending on the particular sample. A simulation study included in this section shows that as $\omega$ approaches the boundary value of zero, the frequency with which the MLE of $\omega$ is equal to zero increases, and approaches its limiting value of 50% when $\omega = 0$ and the sample size increases. This is in agreement with the limiting distribution of $\hat{\omega}_n$, which will be shown to be a mixture of an atom at zero and a continuous half-normal distribution, both with probabilities $1/2$. The non standard nature of the limiting distribution of $\hat{\omega}_n$ is due to the lack of regularity of this family: the parameter space is not an open set. This result follows the work of Self and Liang (1987) who studied properties of MLEs and likelihood ratio statistics under such non standard conditions (see also Chant, 1974; Chernoff, 1954; Moran, 1971).

In Sec. 3, we consider the problem of testing exponentiality versus Pareto distribution. Our approach, which closely follows Rominger (2005), is the classical likelihood ratio test for nested families. Here again the problem is non standard and the limiting null distribution of the deviance statistic is not chi-square. Instead, in view of the results of Self and Liang (1987), it is a mixture of an atom at zero and a chi-square distribution with equal probabilities. Although this approach has been used before for testing exponentiality versus generalized Pareto distribution (where the survival function is given by (4) with $s > 0$ and either $\omega > 0$ or $\omega < 0$), to the best of our knowledge the distribution of the test statistic in our case (exponential versus Pareto) is derived here for the very first time. A simulation study that concludes this section compares the actual critical values for finite samples from the exponential distribution with those obtained from the limiting distribution of the test statistic. The fact that the limiting distribution of the deviance statistic for a likelihood ratio test is not chi-square is important to note in the age when likelihood ratio tests are often performed numerically, and chi-square percentiles are commonly used to make decisions. It serves as an example of a test important in practice and useful in numerous application areas (e.g., via the POT theory connection), where the standard large sample theory for likelihood ratio tests does not apply.

In Sec. 4, we present the results of a simulation study where power of our test is compared with that of numerous other tests for exponentiality. We considered several tests designed for testing exponentiality versus heavy tail or generalized Pareto distributions (see Brilhante, 2004; Bryson, 1974; Gomes and Van Montfort, 1987; Van Montfort and Witter, 1985) as well as a handful of other general tests for exponentiality. Since the literature on testing exponentiality is quite extensive (see, e.g., Ch. 10 of the monograph of D’Agostino and Stephens, 1986 or more recent surveys Ascher, 1990 and Henze and Meintanis, 2005), we had to be selective in our choices of tests. Our analysis reveals that the likelihood ratio test for exponentiality is overall the most powerful one among all tests considered when the sample is Pareto. This holds across all sample sizes and values of the parameters considered in our study.

The motivation for this work came directly from climatological questions about the nature of daily precipitation distribution tails. Climate is the statistics of weather. The tails of daily precipitation determine the size of hydrological weather extremes and these are necessary for determination of the associated climatic risk. Realistic estimates of high precipitation percentiles depend on the nature of the tails.
A key practical question in climatology and hydrology, specifically in the fields of flood risk estimation, water resources management, safe engineering design, and hazard assessment is whether daily precipitation has exponential or power tails. The reality and prospects of climatic change provide extra urgency to answering this question. Our approach to this decision problem was to establish a likelihood ratio test for exponential versus Pareto distributions. In Sec. 5, we present results of analysis of daily precipitation from hundreds of meteorological stations across North America using our test. We show that power tails appear at a majority of locations. Moreover, we argue that the statistical analysis results agree with climatological reasons for high/low volatility in precipitation. We conclude this work with proofs, collected in Sec. 6.

2. ML Estimation of Pareto II Parameters

Let $X_1, X_2, \ldots, X_n$ be a random sample from the Pareto II distribution given by the survival function (4) with $(\omega, s) \in \Omega$, where

$$\Omega = \{ (\omega, s) : \omega \geq 0, s > 0 \}. \quad (5)$$

The exponential distribution (3) is included here as a special case with the parameters $(\omega, s) \in \Omega_0$ falling on the boundary of the parameter space,

$$\Omega_0 = \{ (\omega, s) : \omega = 0, s > 0 \}. \quad (6)$$

Then the log-likelihood function is

$$L(\omega, s) = -n \left\{ \log s + (1 + 1/\omega) \frac{1}{n} \sum_{j=1}^{n} \log(1 + \omega X_j/s) \right\}. \quad (7)$$

Rather then setting the partial derivatives equal to zero and solving the resulting equations, we follow a two-step procedure (which is essentially the same as the re-parameterization noted by Davison, 1984 in the context of generalized Pareto distribution) that reduces the dimensionality of the maximization problem. In Step 1 we maximize the log-likelihood (7) along the line $\omega = s/\sigma$, $0 < s < \infty$, where $\sigma \in (0, \infty]$ is held fixed. The value $\sigma = \infty$ yields the positive half-line, corresponding to the exponential distribution. It is not hard to see that there is a unique point $(s = s(\sigma), \omega = s(\sigma)/\sigma)$ on the line $\omega = s/\sigma$ that maximizes the log-likelihood (7), where

$$s(\sigma) = \frac{\sigma}{n} \sum_{j=1}^{n} \log(1 + X_j/\sigma) \in (0, \infty) \quad (8)$$

and $s(\infty)$ is understood as the limit of $s(\sigma)$ at infinity, equal to the sample mean $\bar{X}_n$. The latter case is simply the MLE of the scale parameter $s$ based on a random sample from the exponential distribution (3). Thus, for any $\sigma \in (0, \infty]$, we have the relation $\max_{0 < s < \infty} L(s/\sigma, s) = L(s(\sigma)/\sigma, s(\sigma)) = Q(\sigma)$, where

$$Q(\sigma) = -n \left\{ 1 + \log \sigma + \log \left[ \frac{1}{n} \sum_{j=1}^{n} \log(1 + X_j/\sigma) \right] + \frac{1}{n} \sum_{j=1}^{n} \log(1 + X_j/\sigma) \right\}. \quad (9)$$
Step 2 of maximization of the log-likelihood function consists of maximizing \( Q(\sigma) \) with respect to \( \sigma \in (0, \infty) \). The following result provides important properties of the function \( Q \) related to this problem.

**Lemma 2.1.** If \( Q \) is given by (9) with integer \( n \geq 1 \) and non negative \( X_1, X_2, \ldots, X_n \) such that \( \sum_{i=1}^{n} X_i > 0 \), then

(i) \( Q \) is continuous and differentiable on \((0, \infty)\) with

\[
\lim_{\sigma \to 0^+} Q(\sigma) = -\infty \quad \text{and} \quad \lim_{\sigma \to \infty} Q(\sigma) = -n(1 + \log X_n); \tag{10}
\]

(ii) \( Q \) is monotonically increasing when \( n = 1 \).

It follows that when the sample size \( n = 1 \), the function \( Q(\sigma) \) in (9) is maximized by the boundary value of infinity, leading to \( \hat{\sigma}_n = 0 \) and \( \hat{s}_n = X_n \). These parameters correspond to the exponential distribution (3). Unless the latter is included in the Pareto II model (4) as the boundary case \( \omega = 0 \), the MLEs of \( \omega \) and \( s \) do not exist for \( n = 1 \). In general, in view of Part (i) of Lemma 2.1, it is clear that the global maximum of \( Q \) always exists. It occurs at either a finite \( \hat{\sigma}_n \), if the function \( Q \) crosses over the level \( r = -n(1 + \log X_n) \), or at \( \hat{\sigma}_n = \infty \), if \( Q(\sigma) < r \) for all \( \sigma \in (0, \infty) \) (see Fig. 1). The MLEs of \( \omega \) and \( s \) are then given by \( \hat{\omega}_n = s(\hat{\sigma}_n)/\hat{s}_n, \hat{s}_n = s(\hat{\sigma}_n) \), with \( s \) defined by (8). Note that when the function \( Q \) is increasing, then the MLEs are given by \( \hat{\omega}_n = 0 \) and \( \hat{s}_n = X_n \), as they are when \( n = 1 \), pointing towards an exponential sample. This is consistent with the fact that the limit \( r = -n(1 + \log X_n) \) is the maximum value of the log-likelihood if the sample is from the exponential distribution (3), in which case the MLE of \( s \) is the sample mean. Figure 1 illustrates two common behaviors of the pivotal function \( Q \). In the left panel, where the sample is standard exponential, \( Q \) is increasing and \( \hat{\sigma}_n = \infty \). In contrast, in the right panel, where the sample is Pareto II (4) with \( \omega = 1/2 \) and \( s = 1 \), the function is unimodal.

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Two cases of behavior of \( Q(\sigma) \) given by (9), where \( n = 10,000 \). Left panel: samples are from the standard exponential; Right panel: samples are from Pareto II (4) with \( \omega = 1/2 \) and \( s = 1 \). Horizontal line on both graphs indicates the limiting value of \( Q \) when \( \sigma \to \infty \).
Testing Exponentiality Versus Pareto Distribution

with the mode at \( \hat{\sigma}_e \in (0, \infty) \). In general, one can observe both behaviors of \( Q \) for exponential as well as Pareto samples.

We ran a small simulation study to investigate the frequency of the monotonicity of \( Q \) across various values of the tail parameter \( \omega \) and sample size \( n \). We generated \( k = 10,000 \) samples from the Pareto II distribution for each combination of \( \omega \) and \( n \) and recorded the relative frequency \( f(\omega, n) \) with which the function \( Q \) was strictly increasing, which points towards the exponential distribution (see Table 1). It appears that, generally, \( f(\omega, n) \) is rather low when the samples are bona fide Pareto \((0 < \omega < \infty)\) with large \( \omega \) (i.e., “far” from the exponential distribution for which \( \omega \) is close to zero). However, \( f(\omega, n) \) increases when \( \omega \) decreases and approaches the boundary value of zero. In the limit \((\omega = 0)\) the relative frequency of strictly increasing \( Q \)s can be substantial. Moreover, when the samples are exponential \((\omega = 0)\) the frequencies \( f(\omega, n) \) decrease from 1 to 1/2 as the sample size increases.

These results are consistent with the large-sample properties of the MLEs, discussed below. It is well known that if the parameters are restricted to the set

\[
\Omega_1 = \{(\omega, s) : \omega > 0, s > 0\},
\]

so that the distribution is Pareto, then the densities are regular as far as maximum likelihood estimation is concerned (see, e.g., Arnold, 1983, p. 210). In this case, the MLEs of \( \omega \) and \( s \) are consistent and asymptotically normal, with variance–covariance matrix \( \frac{1}{n} I^{-1}_{\omega, s} \), where

\[
I_{\omega, s} = \left[ -\mathbb{E}\left\{ \frac{\partial^2 \log g(X)}{\partial \gamma_i \partial \gamma_j} \right\}_{i, j=1} \right]^2
\]

is the Fisher information matrix. Here, \( X \) has the Pareto II distribution (4) with vector parameter \( \gamma = (\gamma_1, \gamma_2) = (\omega, s) \) and density \( g \). The matrix (12) can be obtained by adapting the Fisher information matrix in the \((\alpha, \sigma)\) parameterization (2), given on p. 210 in Arnold (1983). Routine calculations lead to

\[
I_{\omega, s} = \frac{1}{(1 + \omega)(1 + 2\omega)s^2} \begin{bmatrix} 2s^2 & s \\ s & 1 + 2\omega \end{bmatrix}.
\]

Table 1

The entries are the relative frequencies \( f(\omega, n) \) with which the function \( Q \) given by (9) was strictly increasing based on \( k = 10,000 \) random samples of size \( n \) from the Pareto II distribution (4) with scale \( s = 1 \) and given \( \omega \). In case \( \omega = 0 \), the distribution is exponential (3).

\[
\begin{array}{cccccccccc}
\hline
n \setminus \omega & 10 & 2 & 1 & 0.5 & 0.1 & 0.01 & 0.05 & 0 \\
\hline
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
2 & 0.2867 & 0.6911 & 0.8146 & 0.8844 & 0.9284 & 0.9323 & 0.9344 & 0.9370 \\
10 & 0.0001 & 0.0433 & 0.1732 & 0.3830 & 0.6570 & 0.7338 & 0.7488 & 0.7393 \\
100 & 0 & 0 & 0 & 0.0016 & 0.2634 & 0.5480 & 0.5747 & 0.5884 \\
1,000 & 0 & 0 & 0 & 0 & 0.0034 & 0.4092 & 0.4690 & 0.5304 \\
10,000 & 0 & 0 & 0 & 0 & 0 & 0.1663 & 0.3136 & 0.5141 \\
\hline
\end{array}
\]
In practice, when the sample is truly Pareto II ($\omega \neq 0$) and the sample size $n$ is large, the function $Q$ given by (9) crosses the level $r = -n(1 + \log \bar{X}_n)$ and attains its maximum value at some $\hat{\sigma}_n \in (0, \infty)$, leading to the MLEs $(\hat{\omega}_n, \hat{s}_n) \in \Omega_1$, where $\hat{\omega}_n = s(\hat{\sigma}_n)/\hat{s}_n$, $\hat{s}_n = s(\hat{\sigma}_n)$, and $s$ is given by (8). The value of $\hat{\sigma}_n$ can be obtained by solving the equation $Q'(\sigma) = 0$ by standard numerical methods (such as Newton–Raphson algorithm and its modifications).

In contrast, when the parameter space is the set $\Omega$ given by (5), which is not an open set, and the true distribution is exponential (given by the survival function (4) with parameters $\omega_0 = 0$ and $s_0 > 0$), we have a non regular (boundary) case, where standard large-sample theory does not apply (see, e.g., Chant, 1974; Chernoff, 1954; Moran, 1971; Self and Liang, 1987). Under certain conditions on the derivatives of the relevant densities (see, e.g., Self and Liang, 1987, p. 605), which can be shown to hold in our case, the limiting distribution of $\sqrt{n}[((\hat{\omega}_n, \hat{s}_n) - (\omega_0, s_0)]$ is the same as that of the random vector

$$[Z_1, Z_2] \cdot I[Z_1 > 0] + [0, Z_2 - (I^{21}/I^{11})Z_1] \cdot I[Z_1 < 0],$$

where $[Z_1, Z_2]$ has a multivariate normal distribution with (vector) mean zero and variance–covariance matrix $I^{-1}_{0,s_0} = [I^j_{ij}]_{i,j=1,2}$, and $I(A)$ is the indicator function of the set $A$ (see, e.g., Self and Liang, 1987, Theorem and Case 2, p. 606). One might expect that the Fisher information matrix (13) is continuous at the boundary $\omega_0 = 0, s_0 > 0$,

$$I_{0,s_0} = \lim_{\omega \to 0} I_{\omega,s_0} = \frac{1}{s^2} \begin{bmatrix} 2s^2 & s \\ s & 1 \end{bmatrix},$$

leading to the inverse

$$I^{-1}_{0,s_0} = \begin{bmatrix} 1 & -s \\ -s & 2s^2 \end{bmatrix}.$$

The following result shows that this is indeed the case.

**Lemma 2.2.** The Fisher information matrix (12) corresponding to the Pareto II distribution with $\omega_0 = 0$ and $s_0 > 0$ is given by (15).

In view of the above lemma, the variance–covariance matrix of $(Z_1, Z_2)$ is given by (16) and the quantity $I^{21}/I^{11}$ in (14) is equal to $-s$. In particular, when the true vector-parameter is in the set $\Omega$, the limiting distribution of the MLE of $\omega$ coincides with that of $IZ$, where $I$ has Bernoulli distribution with parameter $1/2$, $Z$ has a standard normal distribution truncated below at the origin, and the two variables are independent.

### 3. Likelihood Ratio Test of Exponentiality Versus Pareto

We now consider the problem of testing

$$H_0 : \omega = 0 \text{ (Exponentiality)} \quad \text{vs.} \quad H_1 : \omega > 0 \text{ (Pareto II)}$$

(17)
based on a random sample \(X_1, \ldots, X_n\) from Pareto II distribution given by the survival function (4) and density
\[
g(x; \omega, s) = \frac{1}{s} \left( \frac{1}{1 + \frac{x}{s}} \right)^{1+1/\omega}, \quad x \geq 0. \tag{18}
\]

The likelihood ratio test rejects the null hypothesis in favor of the alternative whenever the test statistic
\[
\hat{\lambda}_n = \sup_{(\omega, s) \in \Omega_0} g_n(X_1, \ldots, X_n; \omega, s)
\]
\[
\sup_{(\omega, s) \in \Omega} \frac{g_n(X_1, \ldots, X_n; \omega, s)}{g_n(X_1, \ldots, X_n; \hat{\omega}_n, \hat{s}_n)} \tag{19}
\]
is less than a suitable constant \(C\), where
\[
g_n(X_1, \ldots, X_n; \omega, s) = \prod_{j=1}^{n} g(X_j; \omega, s)
\]
is the likelihood function and the sets \(\Omega\) and \(\Omega_0\) are given by (5) and (6), respectively. To evaluate the statistic \(\hat{\lambda}_n\) one needs to find the MLEs of the relevant parameters. In our case, the numerator of \(\hat{\lambda}_n\) is simply the exponential likelihood function,
\[
g_n(X_1, \ldots, X_n; 0, s) = \prod_{j=1}^{n} \frac{1}{s} e^{-X_j/s} = \left( \frac{1}{s} \right)^n e^{-n\bar{X}_n/s},
\]
evaluated at \(\hat{s}_n = \bar{X}_n\) (the MLE of \(s\) under the exponential model), leading to
\[
\sup_{(\omega, s) \in \Omega_0} g_n(X_1, \ldots, X_n; \omega, s) = (e\bar{X}_n)^{-n}. \tag{20}
\]
The denominator is evaluated as
\[
\sup_{(\omega, s) \in \Omega} \frac{g_n(X_1, \ldots, X_n; \omega, s)}{g_n(X_1, \ldots, X_n; \hat{\omega}_n, \hat{s}_n)} \tag{21}
\]
where \(\hat{\omega}_n, \hat{s}_n\) are the MLEs of the parameters of the Pareto II distribution discussed in Sec. 2. In practice, numerical methods are necessary to compute the likelihood ratio statistic \(\hat{\lambda}_n\). The boxplots in Fig. 2 show several properties of our test. First, the critical region for the test on significance level \(\alpha\) is one sided, and consists of values above the \(1 - \alpha\) quantile of \(\hat{\lambda}_n\). Second, the power of the test decreases as \(\omega \to 0\). This is expected since then the Pareto distribution converges weakly to exponential and they become harder to distinguish.

The critical value \(C\) corresponding to a given significance level \(\alpha\) is the \(1 - \alpha\) quantile of the distribution of the test statistic \(\hat{\lambda}_n\) under the null hypothesis. This quantity is often found via Monte Carlo simulations, as the probability distribution of \(\hat{\lambda}_n\) is often untraceable analytically. Alternatively, for large samples one can use the limiting distribution of the related deviance statistic \(-2 \log \hat{\lambda}_n\), which (under certain regularity conditions) is known to be chi-square (see, e.g., Wilks, 1938). We follow these two approaches, noting that in our case the standard regularity conditions are not satisfied (as was already observed in Sec. 2). When we recall the
nature of the asymptotic distribution of the MLEs in the Pareto case, it becomes clear that the deviance statistic $-2 \log \hat{\lambda}_n$ cannot have a chi-square distribution. The reason is that in large samples, about half of the time the MLEs are $\hat{\omega}_n = 0$ and $\hat{s}_n = X_n$, so that the denominator of $\hat{\lambda}_n$ given by (21) coincides with its numerator (20). Instead, the asymptotic distribution of $\hat{\lambda}_n$ (if exists) will have an atom at 1, while that of the deviance statistic will have an atom at zero. The following result formalizes these heuristic observations.

**Proposition 3.1.** Under the null hypothesis in (17), the asymptotic distribution of the deviance $-2 \log \hat{\lambda}_n$ with $\hat{\lambda}_n$ as in (19) is the same as that of $\mathcal{IW}$, where $I$ and $W$ are two independent variables, having a Bernoulli distribution with parameter 1/2 and a chi-square distribution with one degree of freedom, respectively.

We computed the critical values of the likelihood ratio test described above (the deviance statistic) for selected values of the significance level $\alpha$ and sample size $n$ (see Table 2, top of each cell). For every $\alpha$ and finite $n$, the critical values were obtained as medians $C_{\alpha,n}$ of 100 simulated critical values $C_{i,\alpha,n}$ $i = 1, \ldots, 100$, for that combination of $\alpha$ and $n$. Each $C_{i,\alpha,n}$ was obtained as an appropriate sample quantile of the deviance statistic $-2 \log \hat{\lambda}_n$ computed for $k = 10,000$ random samples from the standard exponential distribution (since the test statistic is scale-invariant we do not loose generality by setting $s = 1$). When $n = \infty$, we use the asymptotic result given in Proposition 3.1. Here, the critical value $C$ corresponding to a given significance level $\alpha < 1/2$ is obtained from the relation $\mathbb{P}(W \leq C) = 1 - 2\alpha$, where
Table 2

The entries are (top) the (median) critical values $C_{z,n}$ of the likelihood ratio test for exponentiality versus Pareto II distribution (deviance) for selected values of the significance level $z$ and sample size $n$, and (bottom, in parentheses): standard deviations of the estimates, that is sample standard deviations of $C_{z,n}$.

<table>
<thead>
<tr>
<th>$n\backslash z$</th>
<th>0.10</th>
<th>0.075</th>
<th>0.05</th>
<th>0.025</th>
<th>0.01</th>
<th>0.005</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.48966</td>
<td>0.76077</td>
<td>1.21672</td>
<td>2.09260</td>
<td>3.35497</td>
<td>4.40451</td>
</tr>
<tr>
<td></td>
<td>(0.02632)</td>
<td>(0.03531)</td>
<td>(0.04256)</td>
<td>(0.08688)</td>
<td>(0.13989)</td>
<td>(0.21966)</td>
</tr>
<tr>
<td>20</td>
<td>0.71342</td>
<td>1.02369</td>
<td>1.51814</td>
<td>2.42177</td>
<td>3.75248</td>
<td>4.82679</td>
</tr>
<tr>
<td></td>
<td>(0.03212)</td>
<td>(0.04079)</td>
<td>(0.05394)</td>
<td>(0.09954)</td>
<td>(0.14651)</td>
<td>(0.22657)</td>
</tr>
<tr>
<td>30</td>
<td>0.83550</td>
<td>1.15881</td>
<td>1.65907</td>
<td>2.63144</td>
<td>3.97630</td>
<td>5.11526</td>
</tr>
<tr>
<td></td>
<td>(0.03368)</td>
<td>(0.03391)</td>
<td>(0.05535)</td>
<td>(0.10301)</td>
<td>(0.13937)</td>
<td>(0.24541)</td>
</tr>
<tr>
<td>40</td>
<td>0.90421</td>
<td>1.23871</td>
<td>1.75436</td>
<td>2.71400</td>
<td>4.07980</td>
<td>5.14015</td>
</tr>
<tr>
<td></td>
<td>(0.03675)</td>
<td>(0.03995)</td>
<td>(0.06135)</td>
<td>(0.08630)</td>
<td>(0.14783)</td>
<td>(0.21673)</td>
</tr>
<tr>
<td>50</td>
<td>0.98556</td>
<td>1.30685</td>
<td>1.82447</td>
<td>2.79843</td>
<td>4.21818</td>
<td>5.30660</td>
</tr>
<tr>
<td></td>
<td>(0.04144)</td>
<td>(0.04678)</td>
<td>(0.05315)</td>
<td>(0.09559)</td>
<td>(0.15549)</td>
<td>(0.21899)</td>
</tr>
<tr>
<td>100</td>
<td>1.11028</td>
<td>1.48304</td>
<td>2.03136</td>
<td>3.04784</td>
<td>4.41923</td>
<td>5.55750</td>
</tr>
<tr>
<td></td>
<td>(0.03985)</td>
<td>(0.04337)</td>
<td>(0.04865)</td>
<td>(0.08712)</td>
<td>(0.17191)</td>
<td>(0.24359)</td>
</tr>
<tr>
<td>250</td>
<td>1.27428</td>
<td>1.65673</td>
<td>2.23145</td>
<td>3.27145</td>
<td>4.73036</td>
<td>5.87243</td>
</tr>
<tr>
<td></td>
<td>(0.04286)</td>
<td>(0.04191)</td>
<td>(0.06262)</td>
<td>(0.08931)</td>
<td>(0.16028)</td>
<td>(0.22884)</td>
</tr>
<tr>
<td>500</td>
<td>1.36213</td>
<td>1.76641</td>
<td>2.35324</td>
<td>3.41594</td>
<td>4.91281</td>
<td>6.08266</td>
</tr>
<tr>
<td></td>
<td>(0.04348)</td>
<td>(0.04765)</td>
<td>(0.07481)</td>
<td>(0.10838)</td>
<td>(0.15557)</td>
<td>(0.25426)</td>
</tr>
<tr>
<td>1,000</td>
<td>1.43876</td>
<td>1.84184</td>
<td>2.44156</td>
<td>3.53349</td>
<td>5.04854</td>
<td>6.21052</td>
</tr>
<tr>
<td></td>
<td>(0.04202)</td>
<td>(0.04751)</td>
<td>(0.05882)</td>
<td>(0.08113)</td>
<td>(0.17075)</td>
<td>(0.25488)</td>
</tr>
<tr>
<td>5,000</td>
<td>1.54383</td>
<td>1.97032</td>
<td>2.58501</td>
<td>3.69381</td>
<td>5.22504</td>
<td>6.43516</td>
</tr>
<tr>
<td></td>
<td>(0.04106)</td>
<td>(0.04645)</td>
<td>(0.06926)</td>
<td>(0.09685)</td>
<td>(0.18795)</td>
<td>(0.25958)</td>
</tr>
<tr>
<td>10,000</td>
<td>1.56847</td>
<td>2.00018</td>
<td>2.61434</td>
<td>3.72613</td>
<td>5.29026</td>
<td>6.46242</td>
</tr>
<tr>
<td></td>
<td>(0.04298)</td>
<td>(0.05346)</td>
<td>(0.06192)</td>
<td>(0.10075)</td>
<td>(0.16187)</td>
<td>(0.22427)</td>
</tr>
<tr>
<td>$\infty$</td>
<td>1.64237</td>
<td>2.07225</td>
<td>2.70554</td>
<td>3.84146</td>
<td>5.41189</td>
<td>6.63490</td>
</tr>
</tbody>
</table>

$W$ has a chi-square distribution with one degree of freedom. The critical numbers increase monotonically as $n$ increases and converge to the critical numbers from the limiting distribution of the deviance statistic.

To get an idea about the variability of the (simulated) critical values $C_{z,n}$ presented in Table 2, we computed sample standard deviations of the $C_{z,n}$. These are reported in parentheses below the quantities $C_{z,n}$. Generally, variability of the estimated critical numbers increases with decreasing significance level and with increasing sample size. However, their increase with decreasing significance level is much more pronounced than with the increasing sample size.

The monotone behavior of the critical numbers in Table 2 for a fixed significance level $z$ provided an idea to derive a “formula” for the critical number as a function of the sample size $n$. Such a formula provides a practical and quick way of getting accurate critical numbers for any sample size, without having to approximate them from the values available in Table 2. As the function relating critical values $C_{z,n}$ to $n$ ($n > 3$) and $z$, we suggest the use of a “working” formula of

\[ C_{z,n} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{i} \]
the form

\[ C_{x,n} = \frac{a_{1,x}}{(n + a_{2,x} \log n)^{a_{3,x}}} + C_{x,\infty}, \]  

(22)

where the parameters \( a_{i,x} (i = 1, 2, 3) \) are estimated by fitting a nonlinear regression model via least squares (we used the “nls” procedure in R), and \( C_{x,\infty} \) is the critical number obtained from the asymptotic distribution of the deviance statistic \(-2 \log \lambda_n\).

To fit the nonlinear regression, for each sample size and a given \( x \) we used the same procedure as for the computation of the critical numbers \( C_{x,n} \) for Table 2. That is, we computed medians of 100 realizations of the critical values for that combination of \( n \) and \( x \) using Monte Carlo simulation. Each of the 100 critical numbers was computed based on 10,000 values of the test statistic under the null hypothesis. For each \( x \) we computed critical numbers for the following 53 sample sizes \( n_i \), \( i = 1, \ldots, 53 \), ranging from 5 to 12,000: 5, 10, 20, 30, 40, 50, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 250, 260, 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 480, 500, 600, 700, 800, 900, 1000, 1200, 1400, 1600, 1800, 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 8000, 10,000, 12,000.

Then, the regression was fit to 53 data points \((n_i, C_{x,n_i})\), resulting in the coefficients \( a_{i,x} \) summarized in Table 3.

Figure 3 illustrates the fit of the regression model to the estimated (median) critical numbers (data for regression) for \( x = 0.05, 0.025 \). We note that the fit appears quite tight. In practice, we recommend using the formula (22) with coefficients from Table 3 to find critical numbers for our test for all sample sizes.

4. Power Analysis

Here we present the results of a simulation study where the power of the likelihood ratio test developed in Sec. 3 is compared with that of several other common tests of exponentiality, briefly described below.

4.1. Tests of Exponentiality Used in Our Study

Here we briefly describe the tests for exponentiality used in this study. Our focus is on tests with heavy-tail or Pareto alternative hypotheses. In all cases the tests are based on a complete sample \( X_1, \ldots, X_n \) and the null hypothesis states that the

<table>
<thead>
<tr>
<th>( x )</th>
<th>( a_{1,x} )</th>
<th>( a_{2,x} )</th>
<th>( a_{3,x} )</th>
<th>( C_{x,\infty} )</th>
<th>Res. st. error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>-4.817390</td>
<td>5.988430</td>
<td>0.455834</td>
<td>1.642374</td>
<td>0.00580231</td>
</tr>
<tr>
<td>0.075</td>
<td>-5.698590</td>
<td>6.193300</td>
<td>0.464417</td>
<td>2.072251</td>
<td>0.00597795</td>
</tr>
<tr>
<td>0.05</td>
<td>-6.256970</td>
<td>5.780480</td>
<td>0.458182</td>
<td>2.705543</td>
<td>0.00664779</td>
</tr>
<tr>
<td>0.025</td>
<td>-7.563470</td>
<td>6.315480</td>
<td>0.459347</td>
<td>3.841459</td>
<td>0.00922085</td>
</tr>
<tr>
<td>0.01</td>
<td>-9.999120</td>
<td>8.384410</td>
<td>0.473707</td>
<td>5.411894</td>
<td>0.0150891</td>
</tr>
<tr>
<td>0.005</td>
<td>-8.787270</td>
<td>6.377840</td>
<td>0.435622</td>
<td>6.634897</td>
<td>0.0223808</td>
</tr>
</tbody>
</table>
sample is exponential (with location zero and an unknown scale parameter). The order statistics for sample $X_1, \ldots, X_n$ are denoted by $X_{(1)} < \cdots < X_{(n)}$. Figure 6 shows the distributions of the considered test statistics (best six out of seven considered tests) under the null hypothesis of exponentiality with varying values of $\omega$. The plots help decide if the tests are one- or two-sided, and give an idea about their abilities to distinguish between the Pareto and exponential distributions.

4.1.1. Bryson Tests for Exponentiality Against Pareto Distribution. Bryson (1974) considered a problem of testing exponentiality versus heavy-tailedness, defined through the property of increasing mean residual life. His test for exponentiality against the Lomax distribution is based on the statistic

$$B_n = \frac{\overline{X}_n X_{(n)}}{(n-1)\overline{X}_{GA}^2}, \quad (23)$$

where

$$\overline{X}_{GA} = \left[ \prod_{j=1}^{n} \left( X_j + \frac{X_{(n)}}{n-1} \right) \right]^{1/n} \quad (24)$$

is (an adjusted) geometric mean of the $\{X_j\}$. The test rejects the null hypothesis of exponentiality when $B_n$ is greater than a critical value. Bryson (1974) also suggested two other possible test statistics for exponentiality vs. heavy tails. The first one is the ratio of the largest observation to the sample mean, $R_n = \frac{X_{(n)}}{\overline{X}_n}$. The second one is a statistic proposed by Gnedenko (mentioned among several tests for exponentiality in Fercho and Ringer, 1972),

$$Q_n(r) = \frac{\sum_{j=1}^{r} S_j/r}{\sum_{j=r+1}^{n} S_j/(n-r)}, \quad (25)$$

where $S_j = (n - j + 1)(X_{(j)} - X_{(j-1)})$ and $X_{(0)} = 0$. Under the null hypothesis of exponentiality, the quantities $Q_n$ have an F distribution with $2r$ and $2(n - r)$ degrees
of freedom. Bryson (1974) recommended choosing $r = n - 1$ or $n - 2$ to get a plausible test for exponentiality vs. heavy tailed (increasing mean residual life) distribution.

4.1.2. Tests for Exponentiality Against Generalized Pareto Distribution. Here we consider several known tests for testing exponentiality vs. generalized Pareto distribution (corresponding to (4) with either $\omega < 0$ or $\omega > 0$), adapted for the case where the alternative hypothesis is Pareto ($\omega > 0$). The first one, introduced by Van Montfort and Witter (1985), is based on the statistic $MW_n = r_n \sqrt{n}$, where $r_n$ is the sample correlation coefficient of the pairs $(l_j, -H_{j+1/2})$, $j = 0, 1, \ldots, n - 1$. Here, $l_j = (n - j)(X_{(j+1)} - X_{(j)})$, $X_{(0)} = 0$, and

$$H_j = -\log[1 - j/(n + 1)]. \quad (26)$$

As noted in Van Montfort and Witter (1985), when $\omega > 0$ ($\omega < 0$) the values of $MW_n$ tend to be negative (positive). Thus, when the alternative hypothesis is the Pareto distribution ($\omega > 0$), the critical region consists of the left tail of the null distribution of the test statistic $MW_n$, with the critical values determined by simulations.

The second test, introduced by Gomes and Van Monfort (1987) (see also Van Montfort and Witter, 1985), is the $G$-test based on the ratio of the largest observation to the sample median, $G_n = X_{(n)}/M_n$. As noted in Van Montfort and Witter (1985), when $\omega > 0$ ($\omega < 0$) the values of $G_n$ tend to be large (small). Thus, when the alternative is the Pareto distribution ($\omega > 0$), the null hypothesis is rejected for large values of $G_n$, with the critical values determined by simulations.

The next two tests, both used in a simulation study of Brilhante (2004) and Gomes (1982), have the test statistics

$$V_n = \frac{X_{(n)} - M_n}{M_n - X_{(1)}} \quad \text{and} \quad T_n = \frac{F_U - M_n}{M_n - F_L}. \quad (27)$$

The quantities $F_U$ and $F_L$ separate the upper and lower fourths of the data,

$$F_U = X_{n-[n/4]+1}, \quad F_L = X_{[n/4]}, \quad (28)$$

where $\{x\}$ denotes $x$ rounded to the nearest integer. Both tests are one sided and reject the null hypothesis for large values of the test statistics.

4.1.3. The de Wet–Venter Test. de Wet and Venter (1973) proposed a test statistic based on the ratio of two estimators of the scale parameter,$^1$

$$DW_n = \frac{nX_n}{\left(\sum_{j=1}^{n} X_{(j)}^2/H_j\right)^{1/2}}, \quad (29)$$

where the $\{H_j\}$ are given by (26). This is a one-sided test that rejects the null hypothesis for low values of the $DW_n$.

$^1$There seems to be a misprint in the formula cited in D’Agostino and Stephens (1986, p. 222).
4.1.4. The Jackson Test. Jackson (1967) proposed a test for exponentiality based on a comparison of the ordered observations with their expected values,

\[ J_n = \frac{\sum_{j=1}^{n} m_j X_{(j)}}{n X_n}, \tag{30} \]

where \( m_j = \sum_{i=1}^{j} (n - i + 1)^{-1}, \ j = 1, 2, \ldots, n. \) This is a one-sided test, which rejects the null hypothesis for small values of \( J_n. \)

4.2. Power Analysis

We ran a simulation study to assess the power of the likelihood ratio test and compare its performance with the other tests for exponential versus Pareto distribution. Table 4 summarizes the results of the power study for our test for varying \( \omega \) and sample size \( n. \) As expected, the power increases relatively fast with the sample size and decreases with \( \omega. \) Overall, the power of the likelihood ratio test is quite satisfactory.

To compare the power of our test with that of the tests mentioned earlier in this section, we plotted power functions for varying \( \omega \) and \( n. \) Since the power of \( R_n \) and \( Q_n \) was much lower than that of the other tests, these were omitted from the graph for the clarity of the exposition. Figure 4 shows power plots of the likelihood ratio test and seven other tests discussed above as functions of the sample size

<table>
<thead>
<tr>
<th>( 1/\omega ) ( n )</th>
<th>20</th>
<th>50</th>
<th>100</th>
<th>250</th>
<th>500</th>
<th>1000</th>
<th>5000</th>
<th>10000</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.9918</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.8762</td>
<td>0.9957</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>0.8075</td>
<td>0.9858</td>
<td>0.9999</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>0.7095</td>
<td>0.9491</td>
<td>0.9963</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>1.8</td>
<td>0.6195</td>
<td>0.9062</td>
<td>0.9911</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.554</td>
<td>0.8656</td>
<td>0.9821</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.376</td>
<td>0.6479</td>
<td>0.8814</td>
<td>0.9957</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.2878</td>
<td>0.5007</td>
<td>0.732</td>
<td>0.9519</td>
<td>0.9988</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.2231</td>
<td>0.3931</td>
<td>0.6026</td>
<td>0.8871</td>
<td>0.9895</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.1869</td>
<td>0.3256</td>
<td>0.5013</td>
<td>0.7970</td>
<td>0.9624</td>
<td>0.9992</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>0.1698</td>
<td>0.2751</td>
<td>0.422</td>
<td>0.6960</td>
<td>0.9088</td>
<td>0.9927</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.1492</td>
<td>0.2353</td>
<td>0.3751</td>
<td>0.6067</td>
<td>0.845</td>
<td>0.9791</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.1341</td>
<td>0.2059</td>
<td>0.3206</td>
<td>0.5352</td>
<td>0.7740</td>
<td>0.9503</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>0.1228</td>
<td>0.1915</td>
<td>0.285</td>
<td>0.4842</td>
<td>0.6995</td>
<td>0.9147</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.1151</td>
<td>0.1714</td>
<td>0.2634</td>
<td>0.4299</td>
<td>0.6327</td>
<td>0.8705</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.1105</td>
<td>0.1619</td>
<td>0.2309</td>
<td>0.3860</td>
<td>0.5820</td>
<td>0.8212</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>0.1061</td>
<td>0.1497</td>
<td>0.2162</td>
<td>0.3419</td>
<td>0.5369</td>
<td>0.7708</td>
<td>0.9999</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>0.0994</td>
<td>0.1423</td>
<td>0.203</td>
<td>0.3234</td>
<td>0.4919</td>
<td>0.7221</td>
<td>0.9996</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0.0972</td>
<td>0.1301</td>
<td>0.1846</td>
<td>0.2958</td>
<td>0.4559</td>
<td>0.6747</td>
<td>0.9978</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 4. Power curves for the likelihood ratio test (thick line) as compared to other tests discussed in this section (thin lines). The tests used for comparison include (from best to worst): $J_n$, $MW_n$, $DW_n$, $B_n$, $V_n$, $G_n$, and $R_n$. Left panel: power computed on 10,000 Pareto samples with the parameters $\omega = 1/3$ and $s = 1$ at significance level of 0.05 for variable sample size. Right panel: power computed on 10,000 Pareto samples with $s = 1$ and sample size $n = 100$ at significance level of 0.05 for variable shape parameters $\omega$.

(left panel) and shape parameter (right panel). In both cases, $s = 1$ and the tests are performed on the significance level of $\alpha = 0.05$. The figure shows that the likelihood ratio test’s power is the highest (for this combination of the parameters). The two tests with very comparable power are the Jackson test and the Van Montfort and Witter test.

All computations in the preceeding sections were done using the GNU compiler for the R package on a Sun Fire X4600 ×64 server of a parallel system with 52 processors (2× AMD Opteron Model dual-core 285 (2.6GHz/1MB, 16GB PC3200 DDR1/400 (8× 2GB) memory)). Typical execution times for the computations were: Table 1 (rel. frequencies): 4h; Fig. 2 (distribution of the deviance statistics): 4h; Table 4: 13h; Table 3 (critical numbers): 425h per significance level; Fig. 4 (power comparisons): 19h.

5. An Illustration

The results of this work are directly applicable to several scientific research areas. Below we present an example adapted from Panorska et al. (2007) concerned with climate and hydrology applications and focusing on daily precipitation records over North America. Although the majority of hydrologic literature, following tradition, deals with exponentially tailed models, recent studies include several works pointing to the possibility of heavy tails in the hydro-meteorological data at specific locations (see, e.g., Anderson and Meerschaert, 1998; Katz et al., 2002; Smith, 1989, 2001). Our aim, therefore, was classification of the excesses over high threshold for daily precipitation.

precipitation efficiently at many locations. Days with no recorded precipitation were, of course, excluded from analysis. For the purposes of this work we think of extreme events as those given not only by the maxima, but also by the high percentiles. Our goal was to apply statistical decision theory to daily precipitation at hundreds of meteorological stations covering an entire continent in a relatively simple and computationally efficient way which would enhance our understanding of the connection between nature and mathematics as well as encourage adoption of the method by the wider hydrologic and climate community.

We applied the test developed in Sec. 3 to daily total precipitation from 560 meteorological stations across North America, including Canada (Vincent and Gullett, 1999), U.S. (Groisman et al., 2004), and Mexico (Miranda, 2003), all quality controlled and homogenized at the National Climatic Data Center (NCDC). These stations were selected to give a reasonable coverage over North America (typically one station within a 70 km radius). The selection used the most stringent quality standards in well-sampled regions and relaxed these standards to include all stations in regions with the poorest coverage (i.e., Alaska, Northern Canada, and parts of Mexico). Orography exerts an important control on precipitation. Thus, in mountainous regions, where coverage allowed, we always included the highest elevation station along with the best quality station within the 70 km radius. At least 80% of the data was required to be present at all stations for the common observational time period: January 1, 1950 to December 31, 2001. In what follows, we test the distribution of excesses over local 75th percentile threshold. Computation time was negligible.

Figure 5 shows the spatial distribution of the decision (exponential or Pareto) regarding the tails of the excesses. The majority of the stations show Pareto tail behavior. For example, on 5% significance level, 81% of stations are classified as Pareto tail. Heavy tails occur in regions where strong precipitation volatility reflects a large variety of the synoptic systems producing precipitation. These regions are notably the Gulf Coast where precipitation is produced by frontal systems in winter, thunderstorms and hurricanes in summer; the Midwest and northern plains where both light snowfall and convective downpours are common; the steep Front Range

Figure 5. X’s represent stations at which exponential tail cannot be rejected at the 5% significance level. Circles represent stations at which exponentiality can be rejected at the 5% significance level in favor of the Pareto alternative.
of the Rocky Mountains where typically light precipitation from orographically "squeezed" midlatitude cyclones occurs alongside with orographically enhanced heavy snow from the leading edge of anticyclones descending along the eastern edge of the Rockies which channel them, as well downpours from thunderstorms in summer. Exponential tails exist only in regions where precipitation volatility is quite low because of the similarity of the synoptic systems causing it. These regions include the gentle west-facing slopes of the high western mountains where the bulk of precipitation falls from the orographically enhanced midlatitude cyclones; the high plateau of central Mexico where precipitation is predominantly summertime convective; at the Canadian Atlantic and Pacific coasts, and around drizzly Nova Scotia and Queen Charlotte Islands, where precipitation is predominantly frontal. It appears that diversity in causes and types of precipitation, sometimes due to orography, of a particular region enhance the volatility in precipitation leading to heavy tailed excesses.

The test presented in Fig. 5 shows only limited binary information. Our more detailed work (Panorska et al., 2007) includes more comprehensive information on spatial structure of the relative volatility or heaviness of the tail, as well as a seasonal breakdown of the results. These more detailed results lend more support to our interpretation that diversity in precipitation-producing meteorological systems gives rise to greater volatility. Moreover, seasonal results (Panorska et al., 2007) are absolutely consistent with theory in that distributional mixtures can produce heavy tails only if at least one of the distributions in the mixture is heavy tailed. The spatial structure and coherence of our results and their solid grounding in climate and weather patterns over the North American continent are quite striking, encouraging a clear climatic and geographic interpretation.

Consideration of heavy-tailed distributions in addition to the traditional exponential ones as possible models for precipitation leads to more realistic estimates of extreme event probabilities, return periods (average time until the next extreme event of a given size), size of the $n$-year events such as 100-year floods (Panorska et al., 2007), with vital implications for water resources management, structural design limits and guidelines for safe engineering design, hazard assessment, and other applications including insurance against environmental risk. The correct choice of the stochastic model for the extreme precipitation or stream flow is also essential for scientific investigations of extreme weather events in the context of climate variability and change.

Possible temporal dependence of the data on daily as well as interannual and longer timescales provides impetus for further research. The real prospects of climatic change emphasize the need for mathematical models of extremes consistent with reality. The urgency of this problem is amplified by the prospect that global hydrologic change may disproportionately manifest itself in increased frequency of extreme precipitation. This view is supported by theoretical reasoning and climate projections as well as empirical evidence for increasing trends in the frequency of extreme daily precipitation worldwide (see, e.g., Groisman et al., 2005). A reasonable stochastic model for weather and climate extremes, however, is needed to define and account for high precipitation percentiles. A mathematical model relevant to reality is also required to estimate/project future changes in precipitation extremes. Additionally, our findings can be used as a first step to determine the extent and possible causes of dynamical model (numerical climate and weather models) shortcomings as well as to develop statistical correction schemes for global and
6. Proofs

Proof of Lemma 2.1. The continuity and differentiability of the function \( Q \) on \((0, \infty)\) is easy to see. The limits in (10) follow from the fact that for any \( x > 0 \) the limits of the function \( u(\sigma) = \sigma \log(1 + x/\sigma) \) at zero and infinity are equal to zero and \( x \), respectively. This takes care of Part (i). To establish Part (ii), observe that when we have just one observation \( X > 0 \) then the relation \( Q'(\sigma) > 0 \) is equivalent to

\[
\frac{X/\sigma}{1 + X/\sigma} \left(1 + \frac{1}{\log(1 + X/\sigma)}\right) > 1.
\]

The latter relation follows from the inequality \( \log(1 + X/\sigma) < X/\sigma \), which holds for any positive \( X \) and \( \sigma \). The following result, which can be established by straightforward application of De l’Hôpital’s rule, is useful in proving Lemma 2.2.

Lemma 6.1. For any \( a > 0 \) we have the relations:

\[
\begin{align*}
(\text{i}) \quad & \lim_{\omega \to 0} \frac{\log(1 + a\omega)}{\omega} = a, \\
(\text{ii}) \quad & \lim_{\omega \to 0} \frac{1}{\omega} \left( a - \frac{\log(1 + a\omega)}{\omega} \right) = \frac{a^2}{2}, \\
(\text{iii}) \quad & \lim_{\omega \to 0} \frac{1}{\omega} \left( \frac{\log(1 + a\omega)}{\omega^2} - \frac{a}{\omega + a\omega} - \frac{a}{1 + a\omega} - \frac{a^2}{2} + a \right) = a^2 - \frac{2a^3}{3}.
\end{align*}
\]

Proof of Lemma 2.2. For any \( x > 0 \), let \( h(\omega, s) = \log g(x) \), where \( g \) is the density of the Pareto II distribution (4) with parameters \( \omega \geq 0 \) and \( s > 0 \). Since

\[
h(\omega, s) = \begin{cases} 
- \log s - (1 + 1/\omega) \log(1 + \omega x/s) & \text{for } \omega > 0, \ s > 0 \\
- \log s - x/s & \text{for } \omega = 0, \ s > 0,
\end{cases}
\]

it is clear that the partial derivatives of \( h(\omega, s) \) with respect to \( \omega \) and \( s \) exist for any \( \omega, s > 0 \). Straightforward calculations lead to

\[
\begin{align*}
\frac{\partial h(\omega, s)}{\partial \omega} &= \frac{1}{\omega^2} \log \left(1 + \frac{\omega x}{s}\right) - \frac{x}{s} \left(1 + \frac{\omega x}{s}\right)^{-1} \left(1 + \frac{1}{\omega}\right), \\
\frac{\partial h(\omega, s)}{\partial s} &= -\frac{1}{s} + \frac{x(1 + \omega)}{s^2} \left(1 + \frac{\omega x}{s}\right)^{-1}.
\end{align*}
\]

On the other hand, at the boundary \( \omega = 0, s > 0 \) we have

\[
h'(0, s) = \left. \frac{\partial h(\omega, s)}{\partial \omega} \right|_{(\omega, s) = (0, s)} = \lim_{\omega \to 0^+} \frac{h(\omega, s) - h(0, s)}{\omega} = \lim_{\omega \to 0^+} \frac{-(1 + \frac{1}{\omega}) \log \left(1 + \frac{\omega x}{s}\right) + \frac{x}{s}}{\omega}.
\]

Testing Exponentiality Versus Pareto Distribution
Figure 6. Box plots of the sampling distributions (for varying $\omega$) of the test statistics under the null hypothesis for several of the tests described in Section 4. Left panel: (top) Jackson, (middle) de Vet-Venter, (bottom) Gomes; Right panel: (top) Van Montfort and Witter, (middle) Bryson, (bottom) Gomes and Van Montfort.

which by Parts (i) and (ii) of Lemma 6.1 with $a = x/s$ equals

$$h'(0, s) = \frac{1}{2} \left( \frac{x}{s} \right)^2 - \frac{x}{s}.$$ (35)
The calculation of the partial derivative with respect to $s$ at the boundary is straightforward, leading to

$$h'_s(0, s) = \frac{d}{ds} h(0, s) = -\frac{1}{s} + \frac{x}{s^2}. \quad (36)$$

Next, we have

$$h''_\omega(0, s) = \frac{\hat{c}^2 h(\omega, s)}{\hat{c}^2 \omega^2} \bigg|_{(\omega, x)=(0, x)} = \lim_{\omega \to 0^+} \frac{h'_\omega(\omega, s) - h'_\omega(0, s)}{\omega}, \quad (37)$$

which by (33), (35), and Part (iii) of Lemma 6.1 with $a = x/s$, leads to

$$h''_\omega(0, s) = \left(\frac{x}{s}\right)^2 - \frac{2}{3} \left(\frac{x}{s}\right)^3. \quad (38)$$

Similarly, the 2nd mixed partial derivative at the boundary is

$$h''_{\omega,s}(0, s) = \frac{d}{ds} h'_\omega(0, s) = \frac{x}{s^2} - \frac{x^2}{s^3}. \quad (39)$$

Finally, in view of (36), the second partial derivative (at the boundary) with respect to $s$ is

$$h''_s(0, s) = \frac{d}{ds} h'_s(0, s) = \frac{1}{s^2} - \frac{2x}{s^3}. \quad (40)$$

Now, replacing $x$ in (38)–(40) with $X$ having exponential distribution (3) and taking the expectations produces the Fisher information matrix (15).

Proof of Proposition 3.1. In the notation of Self and Liang (1987), we have $p = 2$, $q = 1$, $s = 0$, and $t = 0$. That is, we have the total of two parameters ($p = 2$), of which one is the parameter of interest ($q = 1$) with true value on the boundary ($\omega = 0$ is on the boundary of the half-closed interval $[0, \infty)$) while the other one is a nuisance parameter ($p - q - s - t = 1$) with true value not on the boundary ($s > 0$ is not on the boundary of the open interval $0, \infty$)). Thus, the result is an immediate consequence of Case 5 of Theorem 3 in Self and Liang (1987).

Acknowledgments

This research was partially supported by NSF grant ATM-0231781. The computational hardware used for this publication was made possible by the NIH Grant No: P20 RR-016464 from the INBRE Program of the National Center for Research Resources. We thank Professor Fred Harris, University of Nevada at Reno, for his kind help with computational resources. We also thank the referee for useful comments.

References


