TESTING THE MULTIVARIATE REGULAR VARIATION MODEL

By

John H.J. Einmahl, Fan Yan,
Chen Zhou

29 October 2018

ISSN 0924-7815
ISSN 2213-9532
Testing the multivariate regular variation model

John H.J. Einmahl*  Fan Yang†
Tilburg University, The Netherlands  University of Waterloo, Canada

Chen Zhou‡
De Nederlandsche Bank and Erasmus University Rotterdam, The Netherlands

October 27, 2018

Abstract

In this paper, we propose a test for the multivariate regular variation model. The test is based on testing whether the extreme value indices of the radial component conditional on the angular component falling in different subsets are the same. Combining the test on the constancy across different conditional extreme value indices with testing the regular variation of the radial component, we obtain the test for testing multivariate regular variation. Simulation studies demonstrate the good performance of the proposed tests. We apply this test to examine two data sets used in previous studies that are assumed to follow the multivariate regular variation model.

Keywords: Extreme value statistics; Hill estimator; local empirical process

JEL Codes: C12, C14.

*E-mail address: j.h.j.einmahl@uvt.nl
†E-mail address: fan.yang@uwaterloo.ca
‡Corresponding author. Postal address: Economics and Research Division, De Nederlandsche Bank, 1000AB Amsterdam, The Netherlands. Phone: +31(20)5243189. E-mail addresses: c.zhou@dnb.nl; zhou@ese.eur.nl
1 Introduction

The multivariate regular variation (MRV) has been widely applied in modeling the dependence across tail events. For instance, various applications of the MRV model in finance and insurance can be found in Mainik and Rüschendorf (2010), Zhou (2010), and Mainik and Embrechts (2013), among others. As a semi-parametric model, the MRV model assumes only a limit relation in the tail region of a multivariate distribution. Consequently it allows for a flexible dependence structure across several heavy-tailed random variables. Theoretical properties of the MRV model and its characterizations can be found in, e.g., Basrak et al. (2002a), Basrak et al. (2002b), Lindskog (2004), Barbe et al. (2006), and Resnick (2007).

In most of the applications, the MRV model is assumed for the underlying random variables without a formal validation. This might be due to the fact that there is no formal goodness-of-fit test of the MRV model in the literature. The only exception is Einmahl and Krajina (2018), which provides a formal test for the MRV model when the underlying random vector is bivariate. In this paper, we construct a formal goodness-of-fit test for the MRV model. Our proposed test can be applied in any dimension.

Without having a formal test, the existing studies employing the MRV model usually apply a simple, informal, check for the validity of the MRV model. The simple check is on the equality of all the extreme value indices of the left and right tails of all marginal distributions because this is implied by the MRV model. Some other application studies conduct a more careful test by comparing extreme value indices beyond the marginals, albeit still informal; see for example Cai et al. (2011). The rationale behind the careful test is as follows. By using polar coordinates, the random variables following a MRV model can be mapped into a univariate radial component and an angular component that can be multivariate. The radius follows a univariate regular variation model with a positive extreme value index, and is asymptotically independent from the angular component. The independence in the limit guarantees that the extreme value index of the radius conditional on the angular component is the same regardless where the conditioning angular component lies. The marginals can be viewed as the directions lining up with the axes in the original coordinate system. Consequently, it would be useful to compare the extreme value indices along other directions, beyond the axes. The main idea in this paper is to formalize such a comparison into a goodness-of-fit test for the MRV model. More specifically, our proposed test combines testing the constancy of the extreme value indices of the radii conditional on
various directions of the angular component with testing the regular variation of the radius.

Testing the constancy in the conditional extreme value indices in all “directions” of the angular component is similar to the constant extreme value index test in Einmahl et al. (2016); see $T_3$ and $T_4$ therein. In their null hypothesis, the observations are generated from different univariate distributions with the same extreme value index but different “scale”. In other words, the extreme value indices are the same at all locations, while the scale varies according to a fixed covariate indicating the location. Our test can be viewed as testing the constancy in the conditional extreme value index across random covariates, i.e., the angular component induces the scale. More specifically, we employ a test that is parallel to the $T_4$ test in Einmahl et al. (2016), but with random covariates.

The estimation of the extreme value index with a random covariate has been studied under various parametric setups. Wang and Tsai (2009) investigated the tail index of a Pareto-type distribution when it is linked by the logarithmic function to the linear predictor induced by covariates. Daouia et al. (2011) used nonparametric conditional quantile estimators to estimate the conditional tail index. Follow-up works generalize this approach to the case that the conditional distribution of the response variable belongs to the max-domain of attraction of the extreme value distribution; see Daouia et al. (2013), Stupfler (2013), Goegebeur et al. (2014), and Goegebeur et al. (2014). In contrast to these studies, we do not impose a parametric model between the extreme value index and the covariates. Neither do we emphasize on the estimation of the conditional extreme value index. Instead, we focus on testing the constancy of the “directional” extreme value indices. We demonstrate the finite sample performance of the proposed tests through various models that either satisfy the null hypothesis or fall in the alternative. Especially, simulations based on 3-dimensional MRV models are also performed to illustrate how our testing procedure works in higher dimension. Finally, we apply the test to two real data sets.

The rest of the paper is organized as follows. Section 2 provides the main theoretical results: the constancy test of the conditional extreme value indices and how to combine it with testing the regularly variation of the radius. The simulation study and application can be found in Sections 3 and 4 respectively. Section 5 concludes the paper. The proofs are deferred to Appendix A.
2 Methodology

We define MRV via a transformation to polar coordinates. For an arbitrary norm \( \| \cdot \| \), the polar coordinate transform of a vector \( x \) is defined as

\[
P(x) = \left( \| x \|, \| x \|^{-1} x \right),
\]

(1)

where \( \| x \| \) is called the radial component and \( \| x \|^{-1} x \) is called the angular component of \( x \). A random vector \( X \) with polar transformation \( P(X) \) is said to be multivariate regularly varying, if there exists a probability measure \( \Psi \) on the Borel \( \sigma \)-algebra \( B(S^{d-1}) \), where \( S^{d-1} = \{ s \in \mathbb{R}^d : \| s \| = 1 \} \), and \( \gamma > 0 \), such that, for all \( x > 0 \), as \( t \to \infty \),

\[
\frac{\Pr \left( \| X \| > tx, \| X \|^{-1} X \in \cdot \right)}{\Pr \left( \| X \| > t \right)} \xrightarrow{\nu} x^{-1/\gamma} \Psi(\cdot), \quad \text{on } B \left( S^{d-1} \right),
\]

(2)

where \( \xrightarrow{\nu} \) denotes vague convergence; \( \Psi \) is called the spectral measure.

With a random sample of observations drawn from the distribution of \( X \), we intend to test whether the underlying distribution satisfies the MRV model defined by (2). It is straightforward to derive from (2) that for any Borel set \( B \in B \left( S^{d-1} \right) \), if \( \Psi(B) > 0 \), then

\[
\lim_{t \to \infty} \frac{\Pr \left( \| X \| > tx, \| X \|^{-1} X \in B \right)}{\Pr \left( \| X \| > t \right)} \| X \|^{-1} X \in B \right) = x^{-1/\gamma},
\]

which implies that \( \| X \| \) is regularly varying in any “direction” defined by \( B \). Therefore, we shall estimate the extreme value index \( \gamma = \gamma(B) \) using the observations of \( \| X \| \) conditioning on \( \| X \|^{-1} X \in B \) and further test whether \( \gamma(B) \) is constant across various (disjoint) sets \( B \) with \( \Psi(B) > 0 \). Besides, we need to test whether the radius \( \| X \| \) possesses a regularly varying tail.

The rest of Section 2 is organized in the following way. Firstly, in Section 2.1, we establish a test in the 2-dimensional setup for the null hypothesis of having a constant \( \gamma(B) \). Secondly, testing the univariate regular variation of \( \| X \| \) is well established in the literature. The difficulty here is to avoid a multiple testing problem, that is, we need to be able to combine the two tests into one. We shall establish this in Section 2.2. Although these two subsections focus on the bivariate case, our testing procedure can be extended to the higher dimensional case. Section 2.3 explains the test for higher dimensional MRV.
2.1 Testing the bivariate MRV model

Consider the following polar transformation based on the $L_2$-norm. For a bivariate random vector $(X, Y)^T$, define $R = \sqrt{X^2 + Y^2}$ and

$$\Theta = \begin{cases} 
\arccot \frac{X}{Y}, & Y > 0, \\
\arccot \frac{X}{Y} + \pi, & Y < 0, \\
\pi, & X < 0, Y = 0, \\
2\pi, & X \geq 0, Y = 0.
\end{cases}$$

(3)

Then $R \geq 0$ and $\Theta \in [0, 2\pi]$. For convenience we assume that $F_R$, the distribution function of $R$, is continuous. Write $U_R = 1/(1 - F_R)^-$, where ‘←’ denotes the left-continuous inverse function.

Let $(X_1, Y_1)^T, \ldots, (X_n, Y_n)^T$ be i.i.d. observations. By the polar transformation (3), we obtain the transformed pairs $(R_1, \Theta_1)^T, \ldots, (R_n, \Theta_n)^T$, which is the starting point for constructing the test. We first define the estimator of the extreme value index $\gamma$ in a subregion. Order $R_1, \ldots, R_n$ as $R_{1,n} \leq \ldots \leq R_{n,n}$ and take $R_{n-k,n}$ ($k \in \{1, \ldots, n-1\}$) as the common threshold.

For any $\delta > 0$ and $0 \leq \theta_1 < \theta_2 \leq 2\pi$ satisfying $\Psi(\theta_2) - \Psi(\theta_1) > \delta$, we define a Hill estimator $\hat{\gamma}(\theta_1, \theta_2)$ as the estimator using the observations corresponding to $\theta_1 < \Theta_i \leq \theta_2$ as follows

$$\hat{\gamma}(\theta_1, \theta_2) = \frac{\sum_{i=1}^n (\log R_i - \log R_{n-k,n}) 1\{R_i > R_{n-k,n}, \theta_1 < \Theta_i \leq \theta_2\}}{\sum_{i=1}^n 1\{R_i > R_{n-k,n}, \theta_1 < \Theta_i \leq \theta_2\}}.$$ 

Observe that $\Psi(\theta_2) - \Psi(\theta_1) > \delta$ guarantees $\sum_{i=1}^n 1\{\theta_1 < \Theta_i \leq \theta_2\} \overset{P}{\to} \infty$, as $n \to \infty$. Denote the distribution function of the spectral measure also with $\Psi$. A natural estimator for $\Psi$ (see Einmahl et al. (1993)) is given by

$$\hat{\Psi}(\theta) = \frac{1}{k} \sum_{i=1}^n 1\{R_i > R_{n-k,n}, \Theta_i \leq \theta\}.$$ 

In order to test the constancy of $\gamma(B)$, we estimate $\gamma(B)$ from various subsamples and compare these estimators. More specifically, first for a fixed integer $m$, we split the data with largest $k$ radii into $m$ disjoint parts with about equal number of observations. The cutoff points are defined as follows. Denote $\theta_j = \Psi^- (j/m)$ and $\hat{\theta}_j = \hat{\Psi}^- (j/m)$ for $j = 0, 1, \ldots, m$. Clearly $\theta_0 = \hat{\theta}_0 = 0$ and $\theta_m = \hat{\theta}_m = 2\pi$. Define $\hat{\gamma}_j := \hat{\gamma}(\hat{\theta}_{j-1}, \hat{\theta}_j)$ and $\hat{\gamma}_{alt} := \hat{\gamma}(0, 2\pi)$. 

5
Next, we define the test statistic as

\[ T_n := \frac{k}{m} \sum_{j=1}^{m} \left( \frac{\hat{\gamma}_j \hat{\gamma}_{all} - 1}{\hat{\gamma}_{all}} \right)^2. \]

Clearly, it compares all the \( \hat{\gamma}_j \) obtained in the \( m \) subregions to \( \hat{\gamma}_{all} \) which uses all peaks over threshold.

To establish the asymptotic theory of the test statistic \( T_n \), we assume a second-order condition as follows.

**Assumption 2.1** There exists a function \( \beta \) such that \( \beta(t) \to 0 \) as \( t \to \infty \) and for any \( x_0 > 0 \), as \( t \to \infty \),

\[
\sup_{x > x_0, 0 \leq \theta \leq 2\pi} \left| x^{1/\gamma} \frac{\Pr(R > tx, \Theta \leq \theta)}{\Pr(R > t)} - \Psi(\theta) \right| = O(\beta(t)).
\]

Further assume that \( \Psi \) is continuous on \([0, 2\pi]\).

Now we are ready to present the asymptotic behavior of \( T_n \) under the null hypothesis; the proof of this theorem is deferred to Appendix A.1.

**Theorem 1** If Assumption 2.1 holds and the sequence \( k \) satisfies \( k \to \infty, k/n \to 0 \) and \( \sqrt{k} \beta(U_R(n/k)) \to 0 \) as \( n \to \infty \), then for a fixed integer \( m \geq 2 \), we have that as \( n \to \infty \),

\[ T_n \overset{d}{\to} \chi^2_{m-1}. \]

We remark that the construction method and the asymptotic result for \( T_n \) resemble those for \( T_4 \) in Einmahl et al. (2016), but the proof of Theorem 1 is substantially different from the proof there.

### 2.2 Dealing with the radial component

Besides testing for the same extreme value index in every direction, we also need to test whether the radial component \( R \) possesses a regularly varying tail. We use the PE test in (1.3) in Hüsler and Li (2006). The test statistic is defined as

\[ Q_n = k \int_0^1 \left( \frac{\log R_n(kt) - \log R_{n-k,n}}{\hat{\gamma}_{all}} + \log t \right)^2 t^n dt. \] (4)
Under the null hypothesis that $R$ possesses a regularly varying tail and a restriction on $k$, $Q_n \overset{d}{\to} Q$ as $n \to \infty$, with

$$Q = \int_0^1 \left( t^{-1}B(t) + \log t \int_0^1 s^{-1}B(s)ds \right)^2 t^\eta dt,$$

where $B$ is a standard Brownian bridge. According to Hüsler and Li (2006), $\eta = 0.5$ is a good choice.

To avoid a multiple testing problem, we need to investigate the joint asymptotic behavior of our test statistic $T_n$ in Theorem 1 and $Q_n$. The following theorem shows that the two are asymptotically independent. The proof is again deferred to Appendix A.2.

**Theorem 2** Under the conditions of Theorem 1, we have that

$$(T_n, Q_n) \overset{d}{\to} (T, Q), \quad n \to \infty,$$

where $T \sim \chi_{m-1}^2$ and $Q$ is as in (5), and $T$ and $Q$ are independent.

Following Theorem 2, we can construct a combined test based on $T_n$ and $Q_n$. For a significance level $\alpha \in (0, 1)$, this combined test rejects if the test based on $T_n$ or that on $Q_n$ rejects for significance level $1 - \sqrt{1 - \alpha}$. The combined test has a $p$-value

$$1 - (1 - \min(p_1, p_2))^2,$$

where $p_1$ and $p_2$ are the $p$-values of the $T_n$ and $Q_n$ tests, respectively.

### 2.3 Dealing with higher dimensions

In Sections 2.1 and 2.2, we constructed tests for the bivariate MRV model. The same method can be applied in higher dimensions. In this section, we discuss the general idea and some practical suggestions for higher dimensional cases.

Suppose $X = (X_1, X_2, ..., X_d)^T$ is a $d$-dimensional random vector. With the polar transformation (1), we can decompose $X$ into a radial component $\|X\|$ and an angular component $\|X\|^{-1}X \in S^{d-1}$. Testing whether $X$ follows a MRV model boils down to testing whether $\|X\|$ possesses a regularly varying tail and whether the extreme value indices are the same in any “direction” specified by a Borel set $B \in \mathcal{B}(S^{d-1})$. For the former testing problem, we refer to the test in Section 2.2. Here we only focus on the latter.
To construct a test for the constancy of the extreme value index, we need to divide the
unit sphere \( S^{d-1} \) into \( m \) subregions containing about equal number of exceedances. One can
achieve this by processing the division dimension by dimension. We illustrate the idea for
dimension 3.

Let \((X, Y, Z)^T\) be a 3-dimensional random vector. Consider the following polar transfor-
mation
\[
\begin{align*}
X &= R \cos \Omega \cos \Theta, \\
Y &= R \cos \Omega \sin \Theta, \\
Z &= R \sin \Omega.
\end{align*}
\]
Clearly, its inverse transformation projects any \((X, Y, Z)^T\) into \((R, \Theta, \Omega)^T\) with \( R \geq 0, \Theta \in [0, 2\pi] \) and \( \Omega \in [-\pi/2, \pi/2] \) as
\[
\begin{align*}
R &= \sqrt{X^2 + Y^2 + Z^2}, \\
\Theta &\text{ given by (3)}, \\
\Omega &= \arcsin \frac{Z}{R}.
\end{align*}
\]
Suppose we observe an i.i.d. sample drawn from the distribution of \((X, Y, Z)^T\). We transform
each observation \((X_i, Y_i, Z_i)^T\) into the polar coordinates \((R_i, \Theta_i, \Omega_i)^T\) for \( i = 1, 2, \ldots, n \).
Again order \( R_1, \ldots, R_n \) as \( R_1 \leq \cdots \leq R_n \).

Let \( m = m_1m_2 \) with \( m_1, m_2 \) positive integers. We intend to find cutoff points
\( \hat{\theta}_j \) and \( \hat{\omega}_{j,l} \), \( j = 0, 1, \ldots, m_1 \) and \( l = 0, 1, \ldots, m_2 \), to split the observations into \( m \)
blocks such that there are about \( k/m \) exceedances falling into each block of the form
\[
\{ \hat{\theta}_{j-1} < \Theta_i \leq \hat{\theta}_j, \hat{\omega}_{j-1,l} < \Omega_i \leq \hat{\omega}_{j,l} \},
\]
for any \( j = 1, 2, \ldots, m_1 \) and \( l = 1, 2, \ldots, m_2 \).

Consider the distribution function \( \Psi \) of the spectral measure for \( \theta \in [0, 2\pi] \) and \( \omega \in
[-\pi/2, \pi/2] \). A natural estimator for \( \Psi \) is
\[
\hat{\Psi}(\theta, \omega) = \frac{1}{k} \sum_{i=1}^n 1\{R_i > R_{n-k,n}, \Theta_i \leq \theta, \Omega_i \leq \omega\}.
\]
Write \( \hat{\Psi}_{\theta}(\theta) = \hat{\Psi}(\theta, \pi/2) \). In the first step, we define the cutoff points \( \hat{\theta}_j = \hat{\Psi}_{\theta}(j/m_1) \), for
\( j = 1, 2, \ldots, m_1 \). In the second step, for each given \( j = 1, 2, \ldots, m_1 \), denote
\[
\hat{\Psi}_{\theta,j}(\omega) = \hat{\Psi}(\hat{\theta}_j, \omega) - \hat{\Psi}(\hat{\theta}_{j-1}, \omega).
\]
Then, the cutoff points are \( \hat{\omega}_{j,l} = \hat{\Psi}_{\theta,j}(l/m_2) \), for \( l = 1, 2, \ldots, m_2 \). Lastly, we can construct
the extreme value index estimator in each subregion as
\[
\hat{\gamma}_{j,l} = \frac{\sum_{i=1}^n (\log R_i - \log R_{n-k,n}) 1\{R_i > R_{n-k,n}, \hat{\theta}_{j-1} < \Theta_i \leq \hat{\theta}_j, \hat{\omega}_{j-1,l} < \Omega_i \leq \hat{\omega}_{j,l}\}}{\sum_{i=1}^n 1\{R_i > R_{n-k,n}, \hat{\theta}_{j-1} < \Theta_i \leq \hat{\theta}_j, \hat{\omega}_{j-1,l} < \Omega_i \leq \hat{\omega}_{j,l}\}}.
\]
for all \( j = 1, 2, \ldots, m_1 \) and \( l = 1, 2, \ldots, m_2 \). Similarly, we denote the Hill estimator of the radii with \( \hat{\gamma}_{\text{all}} = \frac{1}{k} \sum_{i=1}^{n} (\log R_i - \log \hat{R}_{n-k,n})1\{R_i > \hat{R}_{n-k,n}\} \). The test statistic \( T_n \) in the 3-dimensional case is given by

\[
T_n := \frac{k}{m} \sum_{j=1}^{m_1} \sum_{l=1}^{m_2} \left( \frac{\hat{\gamma}_{j,l}}{\hat{\gamma}_{\text{all}}} - 1 \right)^2.
\]

To establish the asymptotic behavior of \( T_n \), we need a corresponding second-order condition in the 3-dimensional case as follows.

**Assumption 2.2** There exists a function \( \beta(t) \) such that \( \beta(t) \to 0 \) as \( t \to \infty \) and for any \( x_0 > 0 \), as \( t \to \infty \),

\[
\sup_{x > x_0, 0 \leq \theta \leq 2\pi, -\pi/2 \leq \omega \leq \pi/2} \left| \frac{x^{1/\gamma} \Pr(R > tx, \Theta \leq \theta, \Omega \leq \omega)}{\Pr(R > t)} - \Psi(\theta, \omega) \right| = O(\beta(t)).
\]

Further assume that \( \Psi \) is continuous on \([0, 2\pi] \times [-\pi/2, \pi/2]\).

**Theorem 3** If Assumption 2.2 holds and the sequence \( k \) satisfies \( k \to \infty \), \( k/n \to 0 \) and \( \sqrt{k}\beta(U_R(n/k)) \to 0 \) as \( n \to \infty \), then for a fixed positive integer \( m \geq 2 \), we have that as \( n \to \infty \),

\[
T_n \overset{d}{\to} \chi^2_{m-1}.
\]

Moreover the statement of Theorem 2 remains true in dimension 3.

The proof of this theorem is very much the same as that of Theorems 1 and 2 and will hence be omitted.

Higher dimensions can be treated in a similar way. We shall consider the 3-dimensional case in the simulation study in detail; see Section 3.

### 3 Simulation

In this section, we demonstrate the finite sample performance of our proposed tests for MRV. We simulate \( l = 1000 \) samples with sample size \( n = 5000 \). For each sample, we perform the tests for each (asymptotic) significance level \( \alpha = 1\%, 5\%, \text{ and } 10\% \). We report the number of samples for which we reject the null.
3.1 Simulations under the null hypothesis, dimension 2

We first consider two bivariate distributions under the null hypothesis.

**Distribution 1.** Let \((X,Y)^T\) follow a centered Student-\(t\) distribution with \(\nu\) degrees of freedom and \(2 \times 2\) scale matrix with 1 as diagonal elements and \(s \in (-1,1)\) as off-diagonal elements. It is known that \((X,Y)^T\) follows a MRV distribution with extreme value index \(1/\nu\) and the corresponding spectral measure has a positive density. We vary the degrees of freedom \((\nu = 0.5, 2)\) and take \(s = 0.3, 0.7\) to examine the impact of these parameters.

**Distribution 2.** Consider the polar coordinates \((R, \Theta)^T\) of \((X,Y)^T\) following the transformation in (3). Assume \(U\) and \(V\) are two independent uniform-(0,1) random variables. Let \(\Theta = 2\pi V\), and

\[
R = \begin{cases} 
F_1^+(1-U), & V \leq 1/2, \\
F_2^+(1-U), & 1/2 < V \leq 1,
\end{cases}
\]

with \(F_i(x) = 1 - \left(\frac{1}{x+1}\right)^{\beta_i}\) for \(x > 0\) and \(i = 1,2\). If \(\beta_1 \neq \beta_2\), then \((X,Y)^T\) follows a MRV distribution that has a spectral measure with zero density on half of the unit circle. We consider different combinations of the extreme value indices (\(\beta_1 = 0.5, \beta_2 = 2\) and \(\beta_1 = 1, \beta_2 = 3\)).

Since the \(Q_n\) test has been well studied in the literature, for the null distributions we only study the \(T_n\) test. We choose \(m = 4\) and \(m = 6\) in Tables 1 and 2, respectively.

The \(T_n\) test performs well for all 6 distributions under the null hypothesis. In particular, it performs better when \(m = 4\) under the current sample size of 5000. Also, it performs better for heavier tailed marginals.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>(k = 250)</th>
<th>(k = 500)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(10%)</td>
<td>(5%)</td>
</tr>
<tr>
<td>D1</td>
<td>(s = 0.7, \nu = 0.5)</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>(s = 0.7, \nu = 2)</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>(s = 0.3, \nu = 0.5)</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>(s = 0.3, \nu = 2)</td>
<td>93</td>
</tr>
<tr>
<td>D2</td>
<td>(\beta_1 = 0.5, \beta_2 = 2)</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(\beta_1 = 1, \beta_2 = 3)</td>
<td>104</td>
</tr>
</tbody>
</table>

Table 1: The total number of rejections under the null \((m = 4)\)
3.2 Simulations under the alternative; dimension 2

We consider two bivariate distributions under the alternative. We choose \( m = 4 \) below because of the better behavior than \( m = 6 \) under the null. Besides the \( T_n \) test, we also check the performance of the combined test for the alternative distributions. Recall that to achieve a significance level of \( \alpha = 1\% \), 5\%, or 10\%, we should reject the combined null if either of the \( T_n \) or \( Q_n \) test rejects at the level \( 1 - \sqrt{1 - \alpha} \approx 0.5\%, 2.5\%, \) or 5.1\%, respectively.

**Distribution 3.** Consider the polar coordinates \((R, \Theta)\) of \((X,Y)\) following the transformation in (3). Let \( U \) and \( V \) be i.i.d. uniform-(0,1) and set \( R = U^{-1/\beta} \), which implies that \( R \) is regularly varying with extreme value index \( 1/\beta \). Define

\[
\Theta = \begin{cases} 
\pi V, & \frac{1}{2^n} < U \leq \frac{1}{2^{n-1}} \text{ with an odd integer } n, \\
\pi + \pi V, & \frac{1}{2^n} < U \leq \frac{1}{2^{n-1}} \text{ with an even integer } n,
\end{cases}
\]

Then the distribution of \((X,Y)^T\) is not MRV. We choose \( \beta = 0.5, 1 \).

**Distribution 4.** Let \( Z_1 \) and \( Z_2 \) be i.i.d. Pareto with extreme value index \( 1/\beta \). We consider two cases.

**Distribution 4.1.** Let \( (X,Y)^T = (Z_1, 2Z_2)^T \). Then \( (X,Y)^T \) possesses a spectral measure with unequal masses \( 1/(1 + 2^\beta) \) and \( 2^\beta/(1 + 2^\beta) \) at 0 and \( \pi/2 \) respectively.

**Distribution 4.2.** Let \( (X,Y)^T = (Z_1, Z_2)^T \). Then \( (X,Y)^T \) possesses a spectral measure with mass \( 1/2 \) at 0 and at \( \pi/2 \).

For both Distributions 4.1 and 4.2, the spectral measure is not continuous, which falls in the alternative. Again we take \( \beta = 0.5, 2 \).

The simulation results for Distributions 3 and 4 are shown in Table 3. For data simulated from these alternative distributions, the powers of both \( T_n \) and the combined test are high,
The simulation results for Distributions 5 and 6 are shown in Table 4. Again, the numbers

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
Distribution & $\alpha$ & $k = 250$ & $k = 500$ \\
\hline
 & $10\%$ & $5\%$ & $1\%$ & $10\%$ & $5\%$ & $1\%$ \\
\hline
D3 & & & & & & & & & & \\
$\beta = 0.5$ & $T_n$ & 728 & 631 & 403 & 950 & 900 & 793 & 657 & 535 & 346 & 911 & 861 & 739 \\
 & Combined & & & & & & & & & & & \\
$\beta = 1$ & $T_n$ & 741 & 631 & 425 & 955 & 923 & 785 & 656 & 555 & 344 & 929 & 881 & 722 \\
 & Combined & & & & & & & & & & & \\
D4.1 & $\beta = 0.5$ & $T_n$ & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 \\
 & Combined & & & & & & & & & & & \\
 & $\beta = 2$ & $T_n$ & 977 & 955 & 860 & 1000 & 1000 & 1000 & 960 & 929 & 781 & 1000 & 1000 & 999 \\
 & Combined & & & & & & & & & & & \\
D4.2 & $\beta = 0.5$ & $T_n$ & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 & 1000 \\
 & Combined & & & & & & & & & & & \\
 & $\beta = 2$ & $T_n$ & 988 & 973 & 886 & 1000 & 1000 & 999 & 977 & 943 & 841 & 1000 & 1000 & 999 \\
 & Combined & & & & & & & & & & & \\
\hline
\end{tabular}
\caption{The total number of rejections under the alternative}
\end{table}

except when using a lower $k$ and $\alpha$ for Distribution 3.

### 3.3 Dimension 3

In dimension 3 we consider the following two distributions, one falls in the null hypothesis, whereas the other one falls in the alternative. Again we take $m = 4$ ($m_1 = m_2 = 2$).

**Distribution 5.** Let $(X, Y, Z)^T$ follow a centered Student-$t$ distribution with $\nu$ degrees of freedom and scale matrix

$$
\Sigma = \begin{pmatrix}
1 & s & 0 \\
s & 1 & s \\
0 & s & 1
\end{pmatrix},
$$

with $s \in (-1, 1)$. Similar to Distribution 1, this distribution is MRV with extreme value index $1/\nu$ and the corresponding spectral measure has a positive density. We choose $\nu = 0.5, 1$ and $s = 0.3, 0.7$.

**Distribution 6.** Let $X, Y$ and $Z$ be three independent random variables following Pareto distributions with extreme value indices $1/\beta_1$, $1/\beta_2$, and $1/\beta_3$, respectively. In this case the distribution function of the spectral measure is not continuous, which falls in the alternative.

The simulation results for Distributions 5 and 6 are shown in Table 4. Again, the numbers

12
of rejections match the significance levels under the null and the power is high under the alternative.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$\alpha$</th>
<th>$k = 250$</th>
<th>$k = 500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s = 0.7, \nu = 0.5$</td>
<td>$T_n$</td>
<td>99</td>
<td>59</td>
</tr>
<tr>
<td>$s = 0.7, \nu = 1$</td>
<td>$T_n$</td>
<td>102</td>
<td>48</td>
</tr>
<tr>
<td>$s = 0.3, \nu = 0.5$</td>
<td>$T_n$</td>
<td>99</td>
<td>50</td>
</tr>
<tr>
<td>$s = 0.3, \nu = 1$</td>
<td>$T_n$</td>
<td>102</td>
<td>54</td>
</tr>
<tr>
<td>$\beta_1 = \beta_2 = \beta_3 = 0.5$</td>
<td>$T_n$</td>
<td>673</td>
<td>564</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>586</td>
<td>467</td>
</tr>
<tr>
<td>$\beta_1 = \beta_2 = \beta_3 = 1$</td>
<td>$T_n$</td>
<td>594</td>
<td>481</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>526</td>
<td>417</td>
</tr>
<tr>
<td>$\beta_1 = \beta_2 = \beta_3 = 2$</td>
<td>$T_n$</td>
<td>448</td>
<td>334</td>
</tr>
<tr>
<td>Combined</td>
<td></td>
<td>392</td>
<td>300</td>
</tr>
</tbody>
</table>

Table 4: The total number of rejections in dimension 3

4 Application

In this section, we test two data sets that are claimed to be MRV in Cai et al. (2011) and He and Einmahl (2017), respectively.

The first data set we consider is the one used in Cai et al. (2011): daily exchange rates of Yen-Dollar and Pound-Dollar from January 4, 1999 to July 31, 2009. Cai et al. (2011) considered daily log returns, i.e.,

$$X_i = \log \left( \frac{P_{i+1}}{P_i} \right)$$

where $P_i$ is the exchange rate on day $i$. We obtain the data, which consist of 2758 observations, from Thomson Reuters. The left panel of Figure 1, presents the scatter plot of the pair (Yen-Dollar, Pound-Dollar).

We show the Hill estimates of the extreme value index of the radius $R$ by varying $k$, the $p$-values of our $T_n$ test by varying $k$, and the $p$-values of the combined test (combining $T_n$ and $Q_n$ tests) by varying $k$. We take 4 blocks ($m = 4$) in conducting the $T_n$ test and the combined test.
Figure 1: Scatter plots for (Yen-Dollar, Pound-Dollar) and (S&P, FTSE, Nikkei).

According to Cai et al. (2011), the estimated extreme value index for $R$ is $\hat{\gamma}_R = 0.256$, which corresponds to a threshold $k$ around 70-80 in the upper graph of Figure 2. At this level of $k$, from both $T_n$ and combined tests we do not reject the null at a significance level of 5%, see the middle and lower graphs in Figure 2. In general, we do not reject that (Yen-Dollar, Pound-Dollar) follows an MRV distribution for $k$ less than 200.

The second data set is from He and Einmahl (2017) and consists of daily international market price indices of the Standard and Poors (S&P) 500 index from the USA, the Financial Times Stock Exchange FTSE 100 index from the UK and the Nikkei 225 index from Japan. The sample period is from July 2nd, 2001, to June 29th, 2007. Again, daily log returns are constructed. We obtain the data set, which has in total 1564 observations, from the accompanying file of that paper.

We consider the triplet (S&P, FTSE, Nikkei) and test whether it follows an MRV distribution using our tests. The right panel of Figure 1 presents the scatter plot of the triplet. Again, our tests are carried out by plotting the $p$-values against various levels of $k$. Our analysis for the triplet is shown in Figure 3. In He and Einmahl (2017), when estimating the left and right extreme value indices of the three series, the threshold $k$ is chosen at 80. At $k = 80$, we do not reject the null that the triplet follows an MRV distribution at the 5% level by both tests. In general, we do not reject for $k$ less than 150.

One potential drawback of our analysis is that we regard the observations as independent.
Figure 2: The pair (Yen-Dollar, Pound-Dollar). The upper graph shows the Hill estimates for the radius $R$. The middle graph shows the $p$-values of the $T_n$ test. The lower graph shows the $p$-values of the combined test.

without accounting for the potential serial dependence. When the data possess weak serial dependence, e.g. satisfying $\beta$-mixing conditions, the test might be still valid subject to some adjustment. More specifically, we conjecture that the statistic $T_n/\sigma^2$ converges to the same $\chi_{m-1}^2$-distribution limit, where $\sigma^2$ is an adjusting factor determined by the serial dependence. Here for “positive” serial dependence, i.e. when extremes are likely to occur on consecutive days, we have $\sigma^2 > 1$. In that case, the current test can be regarded as a conservative test: if we do not reject the null for the data using the current test, we will not reject the
null after adjusting for serial dependence. Given that for both data sets we consider, we do not reject the null by regarding the data as independent, we conjecture that a proper test accounting for serial dependence will not reject the null either. Had we observed a result rejecting the null, we would have to account for the impact of serial dependence. For that purpose, one may consider the observations on even (or odd) days only and carry out the tests by regarding those observations as independent. We have performed such an analysis and obtained the same conclusion.
5 Conclusion

In this paper, we construct a goodness-of-fit test for the multivariate regular variation model. The test is based on comparing the extreme value indices of the radial component conditional on the angular component falling in different, disjoint subsets. This results in the $T_n$ test. In addition, we test whether the radius follows a univariate regular variation model by the $Q_n$ test. The two tests can be easily combined thanks to their asymptotic independence. The proposed tests can be extended to higher dimensional cases. Simulation studies for both 2-dimensional and 3-dimensional cases show that the $T_n$ test performs well and has a high power. The combined test is applied to a few data sets in the literature that are assumed to be MRV. Our test supports making the MRV assumptions for these data sets.

As in any test in extreme value analysis, one needs to choose the tuning parameters. Besides the usual parameter $k$, here one also needs to choose the number of blocks $m$. The higher $m$, the more directions are being compared. In practice, one has to choose a low $m$ to ensure sufficient observations in each block. A good choice of $m$ depends on both the number of observations $n$ and the underlying probability distribution. In applications, it is recommended to choose a few values for both tuning parameters $k$ and $m$.

A Proofs

A.1 Proof of Theorem 1

We start by indicating the steps towards proving Theorem 1. Firstly, we establish the main tool used in the proof, the asymptotic behavior of the following local empirical process:

$$S_n(x, \theta) := \sqrt{k} \left( \frac{1}{k} \sum_{i=1}^{n} 1\{R_i > U_R(\frac{n}{k}x), \Theta_i \leq \theta \} - \frac{n}{k} \Pr \left( R > U_R \left( \frac{n}{k}x \right), \Theta \leq \theta \right) \right),$$

for $x \geq x_1(> 0), \theta \in [0, 2\pi]$. The following proposition gives its asymptotic behavior.

**Proposition 1** Under the conditions of Theorem 1, there exists a sequence of bivariate Wiener processes $W_n$, defined on the probability space accommodating $(R_1, \Theta_1), \ldots, (R_n, \Theta_n)$, with $\text{Cov}(W_n(x_1, \theta_1), W_n(x_2, \theta_2)) = (x_1 \land x_2) \Psi(\theta_1 \land \theta_2)$, such that for any given $x_1 > 0$ and
0 < \zeta \leq 1/2, as n \to \infty,
\sup_{x \geq x_1, 0 < \theta \leq 2\pi} x^{1/2-\zeta} |S_n(x, \theta) - W_n(1/x, \theta)| \xrightarrow{P} 0. \quad (6)

Secondly, by applying Proposition 1, we prove the joint asymptotic normality for the estimators of the “directional” extreme value indices with fixed cutoff points, \( \hat{\gamma}(\theta_1, \theta_2) \).

**Theorem 4** Under the conditions of Theorem 1, with the same sequence of bivariate Wiener processes \( W_n \) as in Proposition 1, for any \( \delta > 0 \) and uniformly for all \( \theta_1, \theta_2 \) satisfying \( 0 \leq \theta_1 < \theta_2 \leq 2\pi \) and \( \Psi(\theta_2) - \Psi(\theta_1) > \delta \), as \( n \to \infty \),
\[ \sqrt{k}(\hat{\gamma}(\theta_1, \theta_2) - \gamma) = \gamma \left( \int_0^1 \frac{W_n(x, \theta_2) - W_n(x, \theta_1)}{\Psi(\theta_2) - \Psi(\theta_1)} \, dx - \frac{W_n(1, \theta_2) - W_n(1, \theta_1)}{\Psi(\theta_2) - \Psi(\theta_1)} \right) + o_P(1). \]

Lastly, we apply Theorem 4 to handle the estimators of the “directional” extreme value indices when using estimated cutoff points, i.e., the \( \hat{\gamma}_j \) in Section 2.1. Consequently, we obtain the asymptotic behavior of the test statistic \( T_n \) which will complete the proof of Theorem 1.

**Proof of Proposition 1.** We start by proving (6) without the weight function \( x^{1/2-\zeta} \). This is achieved by applying Lemma 3.1 in Einmahl et al. (1997). Write \( U_i = 1 - F_R(R_i) \), then \( U_1, \ldots, U_n \) are i.i.d. uniform-(0,1). Further write \( Y^{(n)}_i = (\frac{n}{k} U_i, \Theta_i) \) and consider the sets \( A(y, \theta) = [0, y] \times [0, \theta], y \leq 1/x_1, 0 < \theta \leq 2\pi \). Then we can rewrite the local empirical process as
\[
S_n(x, \theta) = \frac{n}{\sqrt{k}} \left( \frac{1}{n} \sum_{i=1}^n 1_{\{Y_i^{(n)} \in A(1/x, \theta)\}} - \Pr \left( Y^{(n)}_i \in A(1/x, \theta) \right) \right).
\]

In order to apply Lemma 3.1 in Einmahl et al. (1997), we only need to check that as \( n \to \infty \),
\[ \sup_{y \leq 1/x_1, 0 < \theta \leq 2\pi} \left| \frac{n}{k} \Pr(Y_i^{(n)} \in A(y, \theta)) - \mu(A(y, \theta)) \right| \to 0, \quad (7) \]
for some finite measure \( \mu \).

By taking \( \theta = 2\pi \) in Assumption 2.1, we obtain that as \( t \to \infty \),
\[ \sup_{x > x_0} \left| \frac{\Pr(R > tx)}{\Pr(R > t)} - x^{-1/\gamma} \right| = O(\beta(t)), \]
which implies a second order result for the \( U_R \) function:
\[ \sup_{x \geq x_1} \left| \frac{U_R(tx)}{U_R(t)} - x^{\gamma} \right| = O(\beta(U_R(t))), \quad (8) \]

18
where $x_1$ is any positive constant such that $x_1 > x_0^{1/\rho}$. Replacing $t$ and $tx$ by $U_R(n/k)$ and $U_R(n/(ky))$ respectively, in Assumption 2.1, and by (8), we obtain that as $n \to \infty$,

$$\sup_{y \leq 1/x_1, 0 < \theta \leq 2\pi} \left| \frac{n}{k} \text{Pr}(U_i < ky/n, \Theta_i \leq \theta) - y\Psi(\theta) \right| = O(\beta(U_R(n/k))) \to 0,$$

which verifies (7) with $\mu(A(y, \theta)) = y\Psi(\theta)$. Consequently, we obtain that as $n \to \infty$, for any $x_1 > 0$,

$$\sup_{x \geq x_1, 0 < \theta \leq 2\pi} |S_n(x, \theta) - W_n(1/x, \theta)| \to 0,$$

(9)

where $W_n$ is a series of bivariate Wiener processes as in Proposition 1. (In order to return to the original probability space of the $\left( R_i, \Theta_i \right)$, see p. 52 of Einmahl (1997).)

Next, we introduce the weight function and write $y = 1/x$. Given (9), for a proof of (6) it suffices to prove that for any given $\varepsilon > 0$ and $0 < \zeta < 1/2$, there exists $\eta = \eta(\varepsilon, \zeta) > 0$ such that for sufficiently large $n$,

$$\text{Pr}\left( \sup_{y \leq 1/2 + \varepsilon} |S_{n}(1/y, \theta)| > \varepsilon \right) < 3\varepsilon,$$

(10)

$$\text{Pr}\left( \sup_{y \leq 1/2 + \varepsilon} |W_{n}(y, \theta)| > \varepsilon \right) < \varepsilon.$$ 

(11)

The inequality in (11) is well-known, see, e.g., Theorem 2.2 in Orey and Pruitt (1973). To prove (10), we split the interval $(0, \eta]$ into three parts $I_1 := (0, \tau/k], I_2 := (\tau/k, 1/k\alpha]$ and $I_3 := (1/k\alpha, \eta]$, with $\alpha = (1 + 2\zeta)^{-1}$ and $\tau > 0$. We prove that for all $i = 1, 2, 3$, for large $n$,

$$\text{Pr}\left( \sup_{y \in I_i, 0 < \theta \leq 2\pi} y^{-1/2 + \zeta} |S_{n}(1/y, \theta)| > \varepsilon \right) < \varepsilon.$$

Firstly, we deal with $y \in I_1$. Observe that if $\min_{1 \leq i \leq n} U_i > \tau/n$, then for $y \leq \tau/k$ we have $|S_{n}(1/y, \theta)| \leq \sqrt{ky}$. Therefore, by choosing $\tau$ small enough

$$\text{Pr}\left( \sup_{y \in I_1, 0 < \theta \leq 2\pi} y^{-1/2 + \zeta} |S_{n}(1/y, \theta)| > \varepsilon \right) \leq \text{Pr}\left( \sup_{y \in I_1} \sqrt{ky}^{1/2 + \zeta} > \varepsilon, \min_{1 \leq i \leq n} U_i > \tau/n \right) + \text{Pr}\left( \min_{1 \leq i \leq n} U_i \leq \tau/n \right)$$

$$= \text{Pr}\left( \min_{1 \leq i \leq n} U_i \leq \tau/n \right) < \varepsilon.$$

To deal with $I_2$ and $I_3$, we need the following lemma. Consider the empirical process

$$\alpha_n(x, \theta) = \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^{n} 1_{\{U_i \leq x, \Theta_i \leq \theta\}} - \text{Pr}(U_1 \leq x, \Theta_1 \leq \theta) \right)$$
Lemma 5  For $0 < b_1 < b_2 \leq 1/4$, $0 \leq \xi \leq 1/2$ and $\lambda \geq 0$,

$$\Pr \left( \sup_{b_1 \leq x \leq b_2, 0 < \theta \leq 2\pi} x^{-1/2 + \xi} |\alpha_n(x, \theta)| \geq \lambda \right) \leq C \int_{b_1/2}^{b_2} \frac{1}{s} \exp \left( -\frac{\lambda^2}{4} \frac{1}{s^{2\xi}} \psi \left( \frac{\lambda}{n^{1/2}b_1^{1/2 + \xi}} \right) \right) ds, \tag{12}$$

where $C > 0$ is a constant, and $\psi(\lambda) = 2\lambda^{-2}[(1 + \lambda) \log(1 + \lambda) - \lambda]$ is a continuous, decreasing function defined on $[-1, \infty)$.

We will omit the proof of this lemma, but just mention that it follows that of Inequality 2.6 in Einmahl (1987) for dimension 1, but then uses Inequality 2.5 in there for dimension 2.

Next, we deal with $y \in I_2$. Since $S_n(1/y, \theta) = \sqrt{n^2} \alpha_n \left( \frac{ky}{n}, \theta \right)$, by applying Lemma 5 with $\xi = 0$, we have that for large $n$

$$\Pr \left( \sup_{y \in I_2, 0 < \theta \leq 2\pi} y^{-1/2 + \zeta} |S_n(1/y, \theta)| > \varepsilon \right)$$

$$= \Pr \left( \sup_{\frac{k - 1 - a}{n} \leq x \leq \frac{kn}{n}, 0 < \theta \leq 2\pi} x^{-1/2 + \zeta} |\alpha_n(x, \theta)| > \varepsilon \left( \frac{k}{n} \right)^{\zeta} \right)$$

$$\leq \Pr \left( \sup_{\frac{k - 1 - a}{n} \leq x \leq \frac{kn}{n}, 0 < \theta \leq 2\pi} x^{-1/2} |\alpha_n(x, \theta)| > \varepsilon k^{\zeta a} \right)$$

$$\leq C \int_{\frac{kn}{n}}^{\frac{2k^{1-a}}{n}} \frac{1}{s} ds \cdot \exp \left( -\frac{1}{4} \varepsilon^2 k^{2\zeta a} \psi \left( \frac{\varepsilon k^{\zeta a}}{\tau^{1/2}} \right) \right) \leq C_1 (\log k) \exp \left( -k^{\zeta a} \right) < \varepsilon$$

where $C_1$ is some constant.

Lastly, we deal with $y \in I_3$ by directly applying Lemma 5 with $\xi = \zeta$. We have that

$$\Pr \left( \sup_{y \in I_3, 0 < \theta \leq 2\pi} y^{-1/2 + \zeta} |S_n(1/y, \theta)| > \varepsilon \right)$$

$$= \Pr \left( \sup_{\frac{k - 1 - a}{n} \leq x \leq \frac{kn}{n}, 0 < \theta \leq 2\pi} x^{-1/2 + \zeta} |\alpha_n(x, \theta)| > \varepsilon \left( \frac{k}{n} \right)^{\zeta} \right)$$

$$\leq C \int_{0}^{\frac{2k}{n}} \frac{1}{s} \exp \left( -\frac{1}{4} \varepsilon^2 k^{2\zeta a} \psi \left( \frac{\varepsilon k^{\zeta a}}{k^{(1/2)\zeta} (1-a)} \right) \right) ds$$

$$\tau = \frac{n}{k} \int_{0}^{\frac{2\eta}{k}} \frac{1}{t} \exp \left( -\frac{\varepsilon^2}{4t^{2\zeta}} \psi(\varepsilon) \right) dt.$$

By choosing a sufficiently small $\eta = \eta(\varepsilon, \zeta)$, this bound is less than $\varepsilon$. ■

Next, we prove Theorem 4. The proof is similar to that for the asymptotic normality of the Hill estimator using the tail empirical process; see Example 5.1.5 de Haan and Ferreira (2006).
Proof of Theorem 4. We obtain from (6) that

\[ \sup_{x \geq x_1} \pi^{1/2 - \zeta} \left| \left( S_n(x, \theta_2) - S_n(x, \theta_1) \right) - (W_n(1/x, \theta_2) - W_n(1/x, \theta_1)) \right| = o_P(1), \]  

where the \( o_P(1) \)-term should be read as uniformly for all \( 0 \leq \theta_1 < \theta_2 \leq 2\pi \) such that \( \Psi(\theta_2) - \Psi(\theta_1) > \delta \). In the sequel of the proof all \( o_P(1) \)-terms should be read as uniformly for such \( \theta_1 \) and \( \theta_2 \).

Now consider (13) with \( x \) replaced by \( \frac{k}{n} \frac{1}{1 - F_R(U_R(n/k)u)} \), with \( u \geq u_0 \) for any \( u_0 > 0 \). Assumption 2.1 implies that as \( n \to \infty \),

\[ u^{1/\gamma} \frac{1 - F_R(U_R(n/k)u)}{k/n} = 1 + O(\beta(U_R(n/k))) \to 1. \]

Hence, we can replace the weight function in this new version of (13) by \( u^{(1/2 - \zeta)/\gamma} \).

For the two \( S_n \) terms we have uniformly for all \( u \geq u_0 \), as \( n \to \infty \),

\[ u^{(1/2 - \zeta)/\gamma} \left( S_n \left( \frac{k}{n} \frac{1}{1 - F_R(U_R(n/k)u)}, \theta_2 \right) - S_n \left( \frac{k}{n} \frac{1}{1 - F_R(U_R(n/k)u)}, \theta_1 \right) \right) \]

\[ = u^{(1/2 - \zeta)/\gamma} \sqrt{k} \left( \frac{1}{k} \sum_{i=1}^{n} 1\{R_i > U_R(\frac{u}{k})u, \theta_1 < \theta_i \leq \theta_2 \} - \frac{n}{k} \Pr \left( R > U_R(\frac{n}{k})u, \theta_1 < \Theta \leq \theta_2 \right) \right) \]

\[ = u^{(1/2 - \zeta)/\gamma} \sqrt{k} \left( \frac{1}{k} \sum_{i=1}^{n} 1\{R_i > U_R(\frac{u}{k})u, \theta_1 < \theta_i \leq \theta_2 \} - u^{-1/\gamma}(\Psi(\theta_2) - \Psi(\theta_1)) \right) + o(1). \]

Finally, by the modulus of continuity results for Wiener processes (see Theorem 2.1 in Orey and Pruitt (1973)) we have that as \( n \to \infty \),

\[ \sup_{u \geq u_0, 0 < \theta \leq 2\pi} u^{(1/2 - \zeta)/\gamma} \left| W_n \left( \frac{R}{k} \frac{1 - F_R(U_R(n/k)u)}{1 - F_R(U_R(n/k)u)}, \theta \right) - W_n \left( u^{-1/\gamma}, \theta \right) \right| \]

\[ \leq u^{(1/2 - \zeta)/\gamma} (u^{-1/\gamma} o_P(1))^{1/2 - \zeta} = o_P(1). \]

Together with (14), we obtain that the new version of (13) now reads as

\[ \sup_{u \geq u_0} u^{(1/2 - \zeta)/\gamma} \left| \sqrt{k} \left( \frac{1}{k} \sum_{i=1}^{n} 1\{R_i > U_R(\frac{u}{k})u, \theta_1 < \theta_i \leq \theta_2 \} - u^{-1/\gamma}(\Psi(\theta_2) - \Psi(\theta_1)) \right) \right. \]

\[ \left. - \left( W_n \left( u^{-1/\gamma}, \theta_2 \right) - W_n \left( u^{-1/\gamma}, \theta_1 \right) \right) \right| = o_P(1). \]

By taking \( \theta_1 = 0 \) and \( \theta_2 = 2\pi \) in (15) and using the Vervaat (1972) lemma, we obtain that

\[ \sqrt{k} \left( \frac{R_{n-k,n}}{U_R(\frac{n}{k})} - 1 \right) = \gamma W_n(1, 2\pi) + o_P(1). \]
Notice that the result in (15) is parallel to Theorem 5.1.4 in de Haan and Ferreira (2006). Therefore, using (15) and (16) the proof of the theorem can be completed along similar lines as in the proof of Example 5.1.5 there.

**Proof of Theorem 1.** By taking $u = \frac{R_{n-k,n}}{U_n(\frac{2}{\pi})}$ and $\theta_1 = 0$ in (15), and further applying (16), we obtain that as $n \to \infty$,

$$\sup_{\theta \in [0,2\pi]} \left| \sqrt{k} \left( \frac{\Psi(\theta)}{\Psi(\theta)} - (W_n(1,\theta) - \Psi(\theta)W_n(1,2\pi)) \right) \right| = o_P(1).$$

Using this in conjunction with Theorem 4 we have as $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}_j - \gamma) = m\gamma \left( \int_0^1 (W_n(x,\hat{\theta}_j) - W_n(x,\hat{\theta}_{j-1})) \frac{dx}{x} - (W_n(1,\hat{\theta}_j) - W_n(1,\hat{\theta}_{j-1})) \right) + o_P(1),$$

$$= m\gamma \left( \int_0^1 (W_n(x,\theta_j) - W_n(x,\theta_{j-1})) \frac{dx}{x} - (W_n(1,\theta_j) - W_n(1,\theta_{j-1})) \right) + o_P(1),$$

$$=: \sqrt{m\gamma}N_j + o_P(1),$$

where $N_j \sim N(0,1)$ and the second step is due to the uniform continuity of $W_n$. Next, by applying Theorem 4 for $\theta_1 = 0$ and $\theta_2 = 2\pi$, we obtain that as $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}_{all} - \gamma) = \sqrt{m\gamma}N + o_P(1),$$

with $N = \frac{1}{m} \sum_{j=1}^m N_j$. Hence we have as $n \to \infty$,

$$T_n = \sum_{j=1}^m (N_j - \bar{N})^2 + o_P(1).$$

Using the independent increments property of the Wiener processes $W_n$ we have that $N_1, \ldots, N_m$ are independent, which yields the stated $\chi^2_{m-1}$-limit.

**A.2 Proof of Theorem 2**

**Proof of Theorem 2.** From the proof of Theorem 1, we obtain that the limit of $T_n$ depends on the process $W_n$ defined in Proposition 1. Notice that the limit of the $Q_n$ statistics is related to the asymptotic expansion of the tail quantile process based on $R_1, \ldots, R_n$; see e.g. Theorem 5.2.12 in de Haan and Ferreira (2006). In our setup, this refers the approximating univariate Wiener process is $W_n(\cdot, 2\pi)$. Therefore, with the same steps as in the proof of Theorem 5.2.12 in de Haan and Ferreira (2006), we obtain that as $n \to \infty$,

$$\left| Q_n - \int_0^1 \left( B_n(t, 2\pi) + \log t \int_0^1 B_n(s, 2\pi) ds \right)^2 t^\beta dt \right| \overset{P}{\to} 0,$$
where

$$B_n(s, \theta) = s^{-1}W_n(s, \theta) - W_n(1, \theta).$$

To prove the theorem, it suffices to show that the constructing component for the limit of $Q_n$

$$L_n(t) = B_n(t, 2\pi) + \log t \int_0^1 B_n(s, 2\pi)ds$$

is independent of the constructing component of $T_n$ in Theorem 1

$$M_n(\theta) = \int_0^1 B_n(u, \theta)du.$$

Since $(L_n, M_n)$ is a Gaussian process, it suffices to show that $E[L_n(t)M_n(\theta)] = 0$, for $t \in [0, 1], \theta \in [0, 2\pi]$. This easily follows from

$$E[B_n(s, \theta)B_n(u, \zeta)] = \left(\frac{s \land u}{su} - 1\right)\Psi(\theta \land \zeta).$$

References


